

## CORRESPONDENCE OPEN



## Validation and functional characterization of GWAS-identified variants for chronic lymphocytic leukemia: a CRuCIAL study

© The Author(s) 2022

*Blood Cancer Journal* (2022)12:79; <https://doi.org/10.1038/s41408-022-00676-8>**Dear Editor,**

During the past years, considerable efforts have been made to uncover the genetic component of chronic lymphocytic leukemia (CLL) susceptibility. To date, several genome-wide association studies (GWAS) and their meta-analysis have identified not only single-nucleotide polymorphisms (SNPs) associated with CLL risk [1] but also patient survival [2]. However, despite these noticeable results, it becomes evident that both validation and functional characterization of the genetic variations identified are still required before they can be used in a clinical setting. Hence, we decided to validate the association of 41 GWAS-identified hits for CLL in 1158 CLL cases and 1947 controls ascertained through the Consortium for Research in Chronic lymphocytic Leukemia (CRuCIAL) and to investigate their impact on modulating host immune responses and their utility to predict disease onset. Study participants were of European ancestry and gave their written informed consent to participate in the study, which was approved by the ethical review committee of participant institutions. CLL patients had often Binet stage A and Rai stage I (67.00% and 79.83%) and, compared to controls, had a higher mean age ( $66.19 \pm 12.66$  vs.  $55.60 \pm 11.50$ ) and an increased male/female ratio (1.54 vs. 0.91). SNPs selection was based on published GWAS, functionality according to *HaploReg* data, and linkage disequilibrium between the SNPs. Genotyping of genetic variants was performed using KASP<sup>TM</sup> and Taqman<sup>®</sup> assays. Hardy–Weinberg equilibrium was assessed in the controls ( $P > 0.001$ ) and the association between CLL and SNPs was tested using a multivariate unconditional logistic regression analysis adjusted for age, sex, and country of origin. A meta-analysis of the CRuCIAL results with those from previous GWAS was conducted to validate genetic associations and the  $I^2$  statistic was used to assess statistical heterogeneity between the studies ( $P_{\text{Het}} > 0.01$ ). The pooled OR was computed using the fixed-effect model and the significance threshold for the meta-analysis was set to  $5.0 \times 10^{-8}$ . Mechanistically, we evaluated the correlation of the GWAS-identified SNPs with a production of nine cytokines after in vitro stimulation of whole blood, peripheral mononuclear cells, and monocyte-derived macrophages from 408 healthy subjects of the Human Functional Genomic Project (HFGP) with LPS, PHA, Pam3Cys, CpG and *Borrelia burgdorferi* and *Escherichia coli*. In parallel, we also tested the correlation between selected SNPs and circulating concentrations of 103 serum and plasmatic inflammatory proteins, 7 plasma steroid hormones, and absolute numbers of 91 blood-derived immune cell populations. The HFGP study was approved by the Arnhem-Nijmegen Ethical Committee (42561.091.12) and biological specimens were collected after informed consent was obtained. A detailed description of the study population and

participating centers, selected SNPs and protocols and reagents used in the functional experiments are included in the Supplementary Material available on the *Blood Cancer Journal* website. In order to account for multiple comparisons, we used a significance threshold of  $2.3 \times 10^{-5}$ ,  $1.2 \times 10^{-5}$ ,  $1.34 \times 10^{-5}$ , and  $1.74 \times 10^{-4}$  for the cytokine quantitative trait loci, proteomic, blood cell counts, and steroid hormone analyses, respectively.

Logistic regression analyses confirmed the association of 21 SNPs with CLL risk at  $P < 0.05$  level in the CRuCIAL cohort. The strongest association was found for SNPs located in the *GRAMD1B* locus ( $P = 6.2 \times 10^{-16}$  and  $6.0 \times 10^{-4}$ ) that was further validated through meta-analysis (Table 1). The *GRAMD1B* locus (11q24.1) encodes for a transporter mediating the non-vesicular transport of cholesterol from the plasma membrane to the endoplasmic reticulum. Our experiments revealed that carriers of the *GRAMD1B*<sub>rs35923643G</sub> allele had increased numbers of transitional CD24<sup>+</sup>CD38<sup>+</sup> B cells ( $P = 4.25 \times 10^{-5}$ ; Fig. 1A), which have an IL10-dependent immunosuppressive effect on pro-inflammatory responses against cancer cells. We also found that carriers of the *GRAMD1B*<sub>rs35923643G</sub> allele had increased serum concentrations of IL18R1 ( $P = 0.00085$ ; Fig. 1B), a receptor found to be dysregulated in CLL and that contributes to tumor escape from the immune system [3]. In support of the association of the *GRAMD1B*<sub>rs35923643G</sub> SNP with CLL risk, we found that this genetic variant is located among histone marks for primary B cells and it determines altered motifs for PU1, MEF2A, POU2F2, NKFB, OCT2 and IRF4, which is linked to CLL onset [1]. Moreover, we observed that carriers of the *GRAMD1B*<sub>rs2953196G</sub> allele had decreased circulating concentrations of SIRT2 and ADA ( $P = 0.00037$  and  $0.00079$ ; Fig. 1C, D). SIRT2 is overexpressed in primary CLL cells and plays a key role in determining cell survival [4]. Recent studies have shown that increased serum levels of SIRT2 were associated with longer overall survival [5] whereas SIRT2 inhibitors induced cell death in leukemic cell lines [6]. Similarly, ADA, an enzyme of the purine metabolism related to lymphoid T cell differentiation and tumor cellular responses, has been found to be overexpressed in CLL patients and correlates with longer survival [7]. Another study showed that blockade of A2A adenosine receptors made CLL cells more susceptible to pharmacological treatments while restoring immune competence and T cell proliferation [8]. Serra and coworkers also showed that activation of the ADO receptors inhibited chemotaxis and limited drug-induced apoptosis of CLL cells [9]. Finally, we found that carriers of the *GRAMD1B*<sub>rs2953196G</sub> allele had decreased serum concentrations of STAMBP protein ( $P = 0.00033$ ; Fig. 1E), a key protein involved in the control of autophagy flux and the NLRP3 inflammasome. These results suggest that the *GRAMD1B* locus might exert its biological function on CLL by modulating SIRT2, STAMBP, and ADA, which is a diagnostic biomarker for CLL that has been included in a new prognosis score designed to optimize the patient risk stratification [7].

Besides these findings, the meta-analysis confirmed the association of 29 additional SNPs with the risk of developing the

Received: 20 February 2022 Revised: 26 April 2022 Accepted: 4 May 2022  
Published online: 17 May 2022

Table 1. Validation of GWAS-identified variants for CLL.

SNP	Chr.	Nearby gene	Risk allele	CRUCLAL consortium (n = 3105; 1158 CLL cases and 1947 controls)		Previously published GWAS <sup>b</sup>		Meta-analysis		P <sub>Het</sub>
				OR (95% CI) <sup>a</sup>	P	OR (95% CI)	P	OR (95% CI)	P	
rs4368253	18	AC107990.1  NFE2L3P1	C	<b>1.30</b> (1.13–1.49)	<b>2.0E-04</b>	1.18 (1.11–1.26)	<b>8.00E-07</b>	<b>1.20</b> (1.13–1.27)	<b>5.55E-10</b>	0.212
rs58055674	2	ACOXL	C	<b>1.17</b> (1.01–1.35)	<b>0.032</b>	1.44 (1.33–1.56)	<b>5.00E-20</b>	<b>1.37</b> (1.28–1.47)	<b>7.01E-19</b>	0.014
rs1439287	2	ACOXL	T	<b>1.26</b> (1.11–1.41)	<b>0.0002</b>	1.37 (1.26–1.47)	<b>5.00E-15</b>	<b>1.34</b> (1.25–1.43)	<b>1.62E-18</b>	0.249
rs7944004	11	ASCL2  C11orf21	T	1.10 (0.97–1.24)	0.12	1.20 (1.13–1.27)	<b>2.00E-10</b>	<b>1.18</b> (1.12–1.25)	<b>6.44E-10</b>	0.209
rs4987855	18	BCL2	G	<b>1.42</b> (1.15–1.76)	<b>0.0012</b>	1.47 (1.32–1.61)	<b>3.00E-12</b>	<b>1.46</b> (1.34–1.60)	<b>1.50E-16</b>	0.773
rs2651823	11	C11orf2  TSPAN32	A	1.09 (0.97–1.23)	0.15	1.18 (1.13–1.25)	<b>5.00E-11</b>	<b>1.17</b> (1.11–1.22)	<b>9.62E-11</b>	0.228
rs1476569	4	CAMK2D	G	0.97 (0.86–1.10)	0.66	1.18 (1.12–1.25)	<b>6.00E-10</b>	1.14 (1.09–1.20)	2.01E-07	<b>0.004</b>
rs3769825	2	CASP8	T	1.05 (0.93–1.19)	0.40	1.19 (1.12–1.25)	<b>3.00E-09</b>	<b>1.17</b> (1.11–1.23)	<b>2.12E-09</b>	0.069
rs7558911	2	CFLAR	A	1.02 (0.91–1.16)	0.71	1.18 (1.12–1.24)	<b>5.00E-11</b>	<b>1.16</b> (1.10–1.21)	<b>1.94E-09</b>	0.030
rs1036935	18	CXXC1	A	<b>1.20</b> (1.05–1.38)	<b>0.0091</b>	1.15 (1.10–1.21)	<b>3.00E-08</b>	<b>1.16</b> (1.10–1.21)	<b>3.02E-10</b>	0.564
rs1359742	9	DMRTA1	G	<b>1.18</b> (1.04–1.32)	<b>0.0078</b>	1.20 (1.12–1.28)	<b>7.00E-09</b>	<b>1.20</b> (1.13–1.27)	<b>1.98E-09</b>	0.809
rs6546149	2	DTNB	G	1.05 (0.92–1.20)	0.46	1.09 (1.01–1.17)	2.14E-02	1.08 (1.01–1.15)	0.018	0.629
rs9880772	3	EOMES  LINC01980	T	<b>1.27</b> (1.13–1.43)	<b>5.97E-05</b>	1.19 (1.13–1.25)	<b>2.55E-11</b>	<b>1.20</b> (1.15–1.26)	<b>7.39E-15</b>	0.319
rs13015798	2	FAM126B	A	0.98 (0.86–1.13)	0.82	1.20 (1.14–1.30)	<b>3.00E-08</b>	1.15 (1.09–1.23)	1.76E-06	<b>0.009</b>
rs6586163	10	FAS	A	<b>1.29</b> (1.14–1.46)	<b>4.50E-05</b>	1.23 (1.17–1.29)	<b>1.00E-15</b>	<b>1.24</b> (1.18–1.30)	<b>3.20E-20</b>	0.483
rs2267708	7	GPR37	T	<b>1.22</b> (1.08–1.37)	<b>0.0012</b>	1.16 (1.10–1.22)	<b>9.00E-09</b>	<b>1.17</b> (1.12–1.23)	<b>1.04E-10</b>	0.446
rs35923643	11	GRAMD1B	G	<b>1.93</b> (1.66–2.24)	<b>6.20E-16</b>	1.66 (1.54–1.79)	<b>2.00E-40</b>	<b>1.71</b> (1.60–1.83)	<b>2.76E-55</b>	0.078
rs2953196	11	GRAMD1B	G	<b>1.29</b> (1.12–1.49)	<b>6.00E-04</b>	1.30 (1.22–1.38)	<b>5.00E-16</b>	<b>1.30</b> (1.23–1.37)	<b>1.45E-19</b>	0.922
rs3800461	6	ILRUN	C	1.12 (0.90–1.38)	0.32	1.20 (1.13–1.28)	<b>1.97E-08</b>	<b>1.19</b> (1.12–1.27)	<b>6.80E-09</b>	0.544
rs9392504	6	IRF4	A	<b>1.46</b> (1.29–1.64)	<b>1.75E-09</b>	1.33 (1.26–1.40)	<b>1.00E-28</b>	<b>1.35</b> (1.29–1.42)	<b>3.14E-34</b>	0.163
rs391855	16	IRF8	A	<b>1.20</b> (1.06–1.36)	<b>0.0031</b>	1.37 (1.28–1.45)	<b>1.00E-22</b>	<b>1.33</b> (1.26–1.41)	<b>3.94E-24</b>	0.062
rs898518	4	LEF1	A	0.95 (0.84–1.08)	0.41	1.20 (1.14–1.27)	<b>4.00E-10</b>	<b>1.15</b> (1.10–1.21)	<b>1.16E-08</b>	<b>0.001</b>
rs34676223	1	MDS2	C	<b>1.17</b> (1.03–1.33)	<b>0.015</b>	1.19 (1.14–1.25)	<b>5.04E-13</b>	<b>1.19</b> (1.14–1.24)	<b>7.23E-15</b>	0.807
rs57214277	4	MYL12BP2  LINC02363	T	1.06 (0.92–1.23)	0.43	1.13 (1.08–1.18)	<b>3.69E-08</b>	1.12 (1.08–1.17)	6.50E-08	0.409
rs10936599	3	MYNN	C	1.12 (0.96–1.29)	0.14	1.26 (1.17–1.35)	<b>1.74E-09</b>	<b>1.22</b> (1.16–1.31)	<b>2.11E-10</b>	0.160
rs11715604	3	NCK1	T	0.98 (0.84–1.14)	0.80	NA (NA-NA)	<b>1.97E-08</b>	NA (NA-NA)	NA	NA
rs6489882	12	OAS3	G	1.09 (0.96–1.23)	0.19	1.16 (1.10–1.22)	<b>5.00E-08</b>	<b>1.15</b> (1.10–1.21)	<b>1.13E-08</b>	0.364
rs140522	22	ODF3B	T	<b>1.17</b> (1.03–1.32)	<b>0.016</b>	1.15 (1.10–1.20)	<b>2.70E-09</b>	<b>1.15</b> (1.11–1.20)	<b>1.35E-11</b>	0.797
rs2236256	6	OPRM1  PCEF1	C	<b>1.20</b> (1.06–1.35)	<b>0.0037</b>	1.23 (1.15–1.30)	<b>1.50E-10</b>	<b>1.22</b> (1.16–1.29)	<b>4.49E-13</b>	0.721
rs11637565	15	PCAT29  LOC107984788	G	<b>1.20</b> (1.06–1.36)	<b>0.0040</b>	1.35 (1.28–1.42)	<b>2.00E-31</b>	<b>1.33</b> (1.27–1.39)	<b>6.19E-31</b>	0.087
rs17246404	7	POT1	C	<b>1.28</b> (1.12–1.46)	<b>3.68E-04</b>	1.22 (1.14–1.31)	<b>3.40E-08</b>	<b>1.23</b> (1.16–1.31)	<b>2.71E-11</b>	0.530
rs2511714	8	POU5F1P2  ODF1	G	1.08 (0.95–1.22)	0.23	1.19 (1.11–1.27)	<b>2.00E-07</b>	1.16 (1.10–1.24)	4.89E-07	0.181
rs11083846	19	PRKD2	A	<b>1.17</b> (1.01–1.33)	<b>0.030</b>	1.35 (1.22–1.49)	<b>3.96E-09</b>	<b>1.29</b> (1.19–1.39)	<b>1.24E-09</b>	0.099
rs888096	2	QPCT  IRNU6-1116P	A	1.09 (0.97–1.24)	0.16	1.15 (1.09–1.21)	<b>5.00E-08</b>	<b>1.14</b> (1.09–1.20)	<b>5.37E-08</b>	0.431
rs41271473	1	RHOJ	G	0.95 (0.80–1.12)	0.52	1.19 (1.13–1.26)	<b>1.06E-10</b>	<b>1.17</b> (1.11–1.23)	<b>7.76E-09</b>	0.013
rs73718779	6	SERPINB6	A	0.93 (0.76–1.14)	0.47	<b>1.26</b> (1.16–1.36)	<b>1.97E-08</b>	1.21 (1.12–1.30)	4.51E-07	<b>0.006</b>

Table 1. continued

SNP	Chr.	Nearby gene	Risk allele	CRuCIAL consortium (n = 3105; 1158 CLL cases and 1947 controls)		Previously published GWAS <sup>b</sup>		Meta-analysis		
				OR (95% CI) <sup>a</sup>	P	OR (95% CI)	P	OR (95% CI)	P	P <sub>Het</sub>
rs12638862	3	TERC	A	1.09 (0.94–1.25)	0.25	1.15 (1.09–1.19)	2.00E–11	1.15 (1.10–1.19)	2.72E–10	0.481
rs7705526	5	TERT	A	1.27 (1.11–1.45)	5.0E–04	1.18 (1.12–1.25)	6.00E–10	1.19 (1.13–1.26)	1.06E–11	0.319
rs61904987	11	TMPRSS5  DRD2	T	1.24 (1.02–1.52)	0.032	1.24 (1.16–1.32)	2.46E–11	1.24 (1.17–1.32)	6.90E–12	1.000
rs926070	6	TSBP1-AS1	A	1.08 (0.95–1.24)	0.24	1.27 (NA–NA)	4.00E–08	NA (NA–NA)	NA	NA
rs7254272	19	ZBTB7A MAP2K2	A	1.06 (0.91–1.22)	0.46	1.17 (1.10–1.23)	4.67E–08	1.15 (1.10–1.21)	5.93E–08	0.202

OR and 95% CI was not reported in the original study and, therefore, association estimates were not available in the GWAS catalog. In bold in the CRuCIAL cohort, SNPs with  $P$  values  $<0.05$ . In bold in the meta-analyses, SNPs with a  $P$  value  $<5.5 \times 10^{-8}$  (gold standard significance threshold for GWAS). Considering the relatively low power of the CRuCIAL cohort,  $P_{Het}$  threshold was set to 0.01.

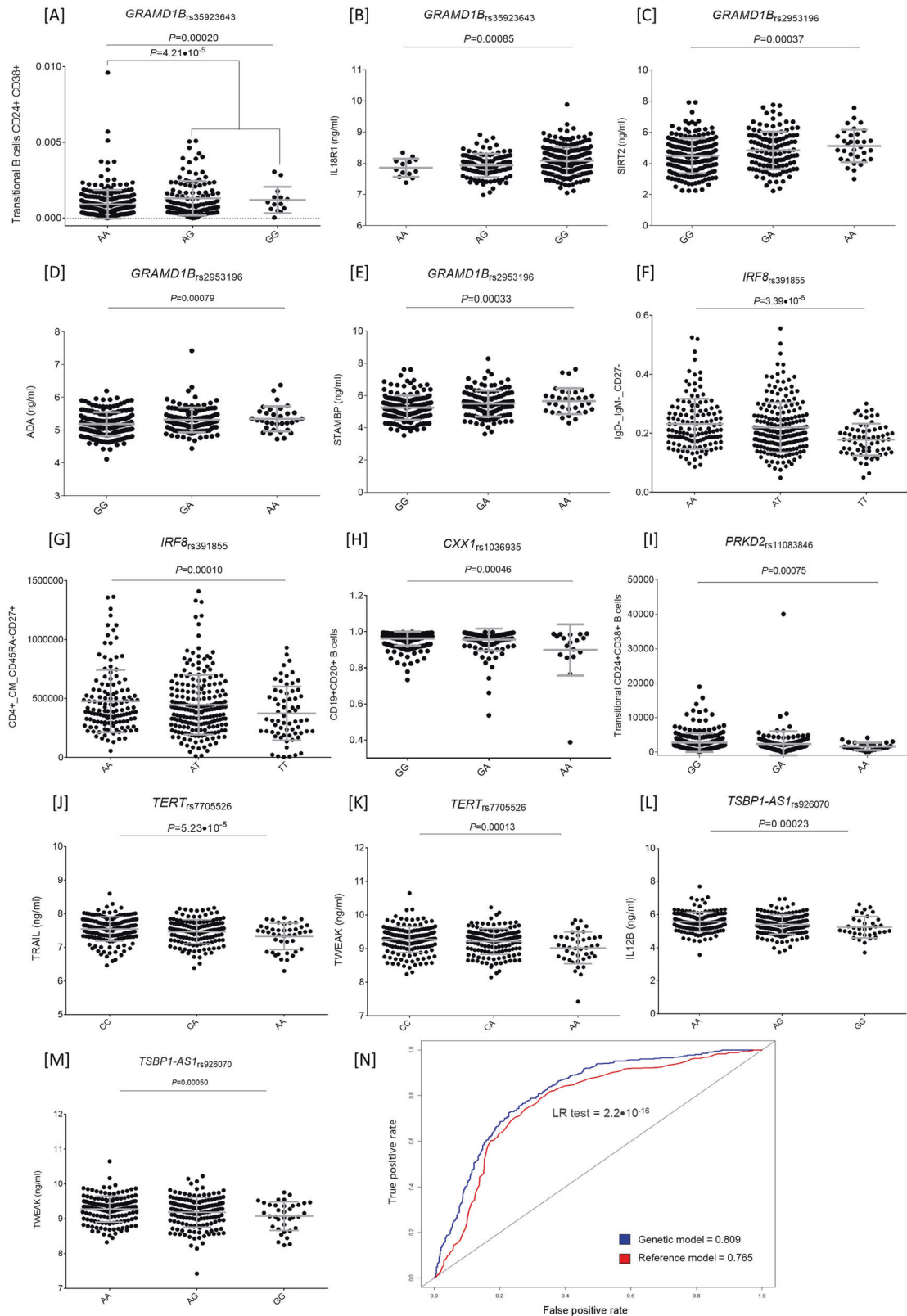
NA not available.

<sup>a</sup>Association estimates calculated according to a log-additive model of inheritance and adjusted for age, sex and country of origin.

<sup>b</sup>References (Pubmed ID) to previously published GWAS are included in the Supplementary Material available on the Blood Cancer Journal website.

disease ( $OR_{Meta} = 1.15–1.71$ ; Table 1), which suggested a functional role of these markers in modulating CLL risk. In this regard, our experiments revealed that carriers of the *IRF8*<sub>rs391855A</sub> allele showed increased numbers of class-switched CD27<sup>+</sup>IgM<sup>+</sup>IgD<sup>+</sup> memory B cells ( $P = 3.39 \times 10^{-5}$ ; Fig. 1F) and central memory CD4<sup>+</sup>CD45RA<sup>−</sup>CD27<sup>+</sup> T cells ( $P = 0.0001$ ; Fig. 1G), whereas carriers of the *CXXC1*<sub>rs1036935A</sub> allele had decreased numbers of CD19<sup>+</sup>CD20<sup>+</sup> B cells ( $P = 0.00075$ ; Fig. 1H), a subset of cells poorly expressed in CLL patients [10]. The *IRF8* locus encodes for a transcription factor exclusively expressed in immune cells that regulate B cell-activating factor (BAFF)-mediated B cell activation, cell survival, adaptive NK cell responses, and CD8/CD4 T cell differentiation. In line with these findings, we also found that carriers of the *PRKD2*<sub>rs11083846A</sub> allele showed decreased numbers of transitional CD24<sup>+</sup>CD38<sup>+</sup>B cells ( $P = 0.00046$ ; Fig. 1I), whereas carriers of the *ILRUN*<sub>rs3800461C</sub> allele had decreased levels of HLA-DR<sup>+</sup> T regulatory and conventional CD4<sup>+</sup> T cells. Finally, we also observed that carriers of the *POU5F1P2*<sub>rs2511714G</sub> allele showed increased numbers of CD8<sup>+</sup> effector memory (CD45RA<sup>−</sup>CD27<sup>−</sup>) T cells. The *POU5F1P2*<sub>rs2511714G</sub> SNP is located among histone marks in primary B cells whereas the *PRKD2*<sub>rs11083846</sub> SNP is an eQTL for the *PRKD2* gene in whole blood but also *SLC1A5*, *CALM3*, and *FKRP* genes that have been associated with CLL onset [11]. We hypothesize that the *IRF8*, *CXX1*, *ILRUN*, and *POU5F1P|ODF1* loci might influence CLL risk by modulating specific subsets of B and T cells and regulatory T cells that play critical roles in the pathogenesis of the disease [12] and influence prognosis. In fact, it is known that peripheral regulatory T cell populations expressing CD4<sup>+</sup> in CLL are associated with disease progression and exhibit a prognostic value [13]. In addition, we found a correlation between the *TERT*<sub>rs7705526A</sub> allele and decreased serum concentrations of TRAIL and TWEAK ( $P = 5.23 \times 10^{-5}$  and 0.0001; Fig. 1J, K), which are involved in the regulation of key cell functions including immune responses, inflammation, proliferation, differentiation, and apoptosis. These results are in agreement with those showing that CLL patients exhibit reduced serum TRAIL both before and after treatment [14] and that its aberrant expression in CLL promotes cell survival [15]. Similarly, we found a correlation of the *TSBP1-AS1*<sub>rs926070G</sub> allele with decreased concentrations of IL12 and TWEAK proteins ( $P = 0.00023$  and 0.00050; Fig. 1L, M), which reinforced the idea of an implication of TWEAK and TWEAK-mediated immune responses in CLL. In support of this finding, it has been reported that TWEAK attenuates the transition from innate to adaptive immunity, which might affect blood cell populations, immune responses, and, consequently, influence the susceptibility to CLL. On the other hand, we found that carriers of the *ILRUN*<sub>rs3800461C</sub> allele had decreased numbers of conventional CD4<sup>+</sup> T cells and HLA-DR<sup>+</sup> T regulatory cells ( $P = 0.00041$  and 0.00058), whereas carriers of the *POU5F1P2*<sub>rs2511714G</sub> allele showed increased numbers of CD8<sup>+</sup> effector memory CD45RA<sup>−</sup>CD27<sup>−</sup> cells ( $P = 0.00053$ ; Supplementary Material). No functional effect for the rest of SNPs was observed.

Considering the number of variants that showed significant associations with CLL risk, we attempted to establish the clinical usefulness of genetic biomarkers in predicting disease onset by using a double approach that consisted of building a predictive model using demographic variables and SNPs significantly associated with CLL risk and weighted and unweighted polygenic risk scores (PRSs; Supplementary Material). The area under the curve (AUC) of a receiver operating characteristic curve analysis and  $-2 \log$ -likelihood ratio (LR) tests showed that a model including age, sex, and 16 SNPs significantly improved the ability to predict the onset of the disease when compared with the reference model including only demographic variables (AUC = 0.809 vs AUC = 0.765;  $P_{LRtest} = 2.2 \times 10^{-16}$ ; Fig. 1N). We also computed weighted and unweighted PRSs in a subset of 806 CLL cases and 1417 controls from the CRuCIAL cohort and we found an OR = 6.81, 95% CI 4.65–9.96,  $P = 2.0 \times 10^{-21}$  for the



**Fig. 1 Functional characterization of GWAS-identified variants for CLL (A–M) and receiver operating characteristic (ROC) curve analysis (N).** Correlation between functional data and GWAS-identified SNPs was evaluated by linear regression analysis adjusted for age and sex. ROC curve summarizes the accuracy of prediction for each particular model. The model including SNPs significantly associated with the risk of developing CLL and demographic variables (marked in blue) showed a significantly improved predictive capacity compared with a reference model including only age and gender as covariates (marked in red). AUC = 0.809 vs. AUC = 0.765;  $N = 2123$  subjects; LR test =  $2.2 \times 10^{-16}$ .



highest vs. lowest quintile of the unweighted score and OR = 10.45, 95% CI 6.96–15.70,  $P = 2.0 \times 10^{-27}$  for the highest vs. lowest quintile of the weighted score. Strong associations were also observed when weighted scores were built using ORs from the original GWASs. The best AUC value was observed for the weighted score computed in the CRuCIAL cohort (AUC = 0.68, 95% CI 0.65–0.70).

In conclusion, this study confirmed the association of 31 GWAS-identified SNPs with CLL risk and shed some light on the function of some of these biomarkers in the modulation of  $T_{Reg}$ , B, and T cell differentiation and proliferation, blood concentrations of B cell-related proteins, cell survival, and the expression of immune- and non-immune-related loci. Though outside the scope of the current study, it is important to mention that additional functional studies using blood samples from CLL patients are still required to validate our findings and to decipher the exact biological mechanisms behind the observed associations. A potential limitation of this work was the relatively small population size of the CRuCIAL cohort that hampered the validation of the SNPs showing modest associations.

Paloma García-Martín<sup>1,36</sup>, Ana Moñiz Díez<sup>2,3,4,36</sup>, José Manuel Sánchez Maldonado<sup>2,3,4,36</sup>, Antonio José Cabrera Serrano<sup>2,3,4</sup>, Rob ter Horst<sup>5,6</sup>, Yolanda Benavente<sup>7,8</sup>, Stefano Landi<sup>9</sup>, Angelica Macaúda<sup>10</sup>, Alyssa Clay-Gilmour<sup>11</sup>, Francisca Hernández-Mohedo<sup>3,4</sup>, Yasmeen Niazi<sup>12,13</sup>, Pedro González-Sierra<sup>3,4</sup>, Blanca Espinet<sup>14,15</sup>, Juan José Rodríguez-Sevilla<sup>16</sup>, Rossana Maffei<sup>17</sup>, Gonzalo Blanco<sup>14,15</sup>, Matteo Giaccherini<sup>9</sup>, Anna Puiggros<sup>14,15</sup>, James Cerhan<sup>18</sup>, Roberto Marasca<sup>17</sup>, Marisa Cañadas-Garre<sup>2,3,4</sup>, Miguel Ángel López-Nevot<sup>19</sup>, Tzu Chen-Liang<sup>20</sup>, Hauke Thomsen<sup>21</sup>, Irene Gámez<sup>20</sup>, Víctor Moreno<sup>8,22</sup>, Rafael Marcos-Gragera<sup>8,23,24</sup>, María García-Álvarez<sup>25</sup>, Javier Llorca<sup>8,26</sup>, Andrés Jerez<sup>20</sup>, Sonja Berndt<sup>27</sup>, Aleksandra Butrym<sup>28</sup>, Aaron D. Norman<sup>18</sup>, Delphine Casabonne<sup>7,8</sup>, Mario Luppi<sup>17</sup>, Susan L. Slager<sup>29,30</sup>, Kari Hemminki<sup>31,32</sup>, Yang Li<sup>5,33</sup>, Miguel Alcoceba<sup>24</sup>, Daniele Campa<sup>9</sup>, Federico Canzian<sup>10</sup>, Silvia de Sanjosé<sup>7,8</sup>, Asta Försti<sup>12,13</sup>, Mihai G. Netea<sup>5,34</sup>, Manuel Jurado<sup>2,3,4</sup> and Juan Sainz<sup>2,3,4,35</sup>✉

<sup>1</sup>Hospital Campus de la Salud, PTS Granada, Granada, Spain.

<sup>2</sup>Genomic Oncology Area, GENYO, Centre for Genomics and Oncological Research: Pfizer/University of Granada/Andalusian Regional Government, PTS, Granada, Spain. <sup>3</sup>Hematology department, Virgen de las Nieves University Hospital, Granada, Spain.

<sup>4</sup>Instituto de Investigación Biosanitaria IBs.Granada, Granada, Spain.

<sup>5</sup>Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, The Netherlands.

<sup>6</sup>CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. <sup>7</sup>Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP) and University of Barcelona, Barcelona, Spain.

<sup>8</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.

<sup>9</sup>Department of Biology, University of Pisa, Pisa, Italy. <sup>10</sup>Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany.

<sup>11</sup>Department of Epidemiology & Biostatistics, Arnold School of Public Health, University of South Carolina, Greenville, SC, USA. <sup>12</sup>Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Heidelberg, Germany.

<sup>13</sup>Hopp Children's Cancer Center (KITZ), Heidelberg, Germany. <sup>14</sup>Molecular Cytogenetics Laboratory, Pathology Department, Hospital del Mar, Barcelona, Spain.

<sup>15</sup>Translational Research on Hematological Neoplasms Group, Cancer Research Program, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain. <sup>16</sup>Hematology Department, Hospital del Mar, Barcelona, Spain. <sup>17</sup>Department of Medical and

Surgical Sciences, University of Modena and Reggio Emilia, AOU Policlinico, Modena, Italy. <sup>18</sup>Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, USA. <sup>19</sup>Immunology Department, Virgen de las Nieves University Hospital, Granada, Spain. <sup>20</sup>Hematology Department, Morales Meseguer University Hospital, Murcia, Spain. <sup>21</sup>ProtaGene CGT GmbH, Heidelberg, Germany. <sup>22</sup>Unit of Biomarkers and Susceptibility, Cancer Prevention and Control Program, IDIBELL, Catalan Institute of Oncology; Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain. <sup>23</sup>Epidemiology Unit and Girona Cancer Registry, Oncology Coordination Plan, Department of Health, Autonomous Government of Catalonia, Catalan Institute of Oncology, Girona Biomedical Research Institute (IdiBGI), and Universitat de Girona, Girona, Spain. <sup>24</sup>Josep Carreras Leukemia Research Institute, Girona, Spain. <sup>25</sup>Department of Hematology, University Hospital of Salamanca (HUS/IBSAL), CIBERONC and Cancer Research Institute of Salamanca-IBMCC (USAL-CSIC), Salamanca, Spain. <sup>26</sup>University of Cantabria, Santander, Spain. <sup>27</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. <sup>28</sup>Medical University of Wroclaw, Wroclaw, Poland. <sup>29</sup>Division of Computational Genomics, Mayo Clinic, Rochester, MN, USA. <sup>30</sup>Division of Hematology, Mayo Clinic, Rochester, MN, USA. <sup>31</sup>Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. <sup>32</sup>Faculty of Medicine and Biomedical Center in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic. <sup>33</sup>Centre for Individualised Infection Medicine (CiiM) & TWINCORE, joint ventures between the Helmholtz-Centre for Infection Research (HZI) and the Hannover Medical School (MHH), Hannover, Germany. <sup>34</sup>Department for Immunology & Metabolism, Life and Medical Sciences Institute (LIMES), University of Bonn, 53115 Bonn, Germany. <sup>35</sup>Department of Biochemistry and Molecular Biology I, Faculty of Sciences, University of Granada (UGR), Granada, Spain. <sup>36</sup>These authors contributed equally: Paloma García-Martín, Ana Moñiz Díez, José Manuel Sánchez Maldonado. ✉email: juan.sainz@genyo.es

## DATA AVAILABILITY

The genotype data used in the present study are available by the corresponding authors upon reasonable request. Functional data used in this project have been meticulously cataloged and archived in the BBMRI-NL data infrastructure (<https://hfgp.bbMRI.nl/>) using the MOLGENIS open-source platform for scientific data. This allows flexible data querying and download, including sufficiently rich metadata and interfaces for machine processing (R statistics, REST API) and using FAIR principles to optimize Findability, Accessibility, Interoperability, and Reusability.

## REFERENCES

- Berndt SI, Camp NJ, Skibola CF, Vijai J, Wang Z, Gu J, et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun.* 2016;7:10933. <https://doi.org/10.1038/ncomms10933>.
- Lin WY, Fordham SE, Sunter N, Elstob C, Rahman T, Willmore E, et al. Genome-wide association study identifies risk loci for progressive chronic lymphocytic leukemia. *Nat Commun.* 2021;12:665. <https://doi.org/10.1038/s41467-020-20822-9>.
- Airoldi I, Raffaghello L, Cocco C, Guglielmino R, Roncella S, Fedeli F, et al. Heterogeneous expression of interleukin-18 and its receptor in B-cell lymphoproliferative disorders deriving from naive, germinal center, and memory B lymphocytes. *Clin Cancer Res.* 2004;10:144–54. <https://doi.org/10.1158/1078-0432.ccr-1026-3>.
- Bhalla S, Gordon LI. Functional characterization of NAD dependent de-acetylases SIRT1 and SIRT2 in B-cell chronic lymphocytic leukemia (CLL). *Cancer Biol Ther.* 2016;17:300–9. <https://doi.org/10.1080/15384047.2016.1139246>.
- Van Damme M, Crompot E, Meuleman N, Mineur P, Bron D, Lagneaux L, et al. HDAC isoenzyme expression is deregulated in chronic lymphocytic leukemia B-cells and has a complex prognostic significance. *Epigenetics.* 2012;7:1403–12. <https://doi.org/10.4161/epi.22674>.
- Dan L, Klimiankova O, Klimiankou M, Klusman JH, van den Heuvel-Eibrink MM, Reinhardt D, et al. The role of sirtuin 2 activation by nicotinamide phosphoribosyltransferase in the aberrant proliferation and survival of myeloid leukemia

- cells. *Haematologica*. 2012;97:551–9. <https://doi.org/10.3324/haematol.2011.055236>.
7. Yun X, Sun X, Hu X, Zhang H, Yin Z, Zhang X, et al. Prognostic and therapeutic value of apolipoprotein A and a new risk scoring system based on apolipoprotein A and adenosine deaminase in chronic lymphocytic leukemia. *Front Oncol*. 2021;11:698572. <https://doi.org/10.3389/fonc.2021.698572>.
  8. Serra S, Vaisitti T, Audrito V, Bologna C, Buonincontri R, Chen SS, et al. Adenosine signaling mediates hypoxic responses in the chronic lymphocytic leukemia microenvironment. *Blood Adv*. 2016;1:47–61. <https://doi.org/10.1182/bloodadvances.2016000984>.
  9. Serra S, Horenstein AL, Vaisitti T, Brusa D, Rossi D, Laurenti L, et al. CD73-generated extracellular adenosine in chronic lymphocytic leukemia creates local conditions counteracting drug-induced cell death. *Blood*. 2011;118:6141–52. <https://doi.org/10.1182/blood-2011-08-374728>.
  10. Ginaldi L, De Martinis M, Matutes E, Farahat N, Morilla R, Catovsky D. Levels of expression of CD19 and CD20 in chronic B cell leukaemias. *J Clin Pathol*. 1998;51:364–9. <https://doi.org/10.1136/jcp.51.5.364>.
  11. Sille FC, Thomas R, Smith MT, Conde L, Skibola CF. Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. *PLoS ONE*. 2012;7:e29632. <https://doi.org/10.1371/journal.pone.0029632>.
  12. De Matteis S, Molinari C, Abbati G, Rossi T, Napolitano R, Ghetti M, et al. Immunosuppressive Treg cells acquire the phenotype of effector-T cells in chronic lymphocytic leukemia patients. *J Transl Med*. 2018;16:172. <https://doi.org/10.1186/s12967-018-1545-0>.
  13. Mpakou VE, Ioannidou HD, Konsta E, Vikentiou M, Spathis A, Kontsioti F, et al. Quantitative and qualitative analysis of regulatory T cells in B cell chronic lymphocytic leukemia. *Leuk Res*. 2017;60:74–81. <https://doi.org/10.1016/j.leukres.2017.07.004>.
  14. Jablonska E, Kiersnowska-Rogowska B, Aleksandrowicz-Bukin M, Rogowski F, Sawicka-Powierza J. TRAIL receptors in the serum of patients with B-cell chronic lymphocytic leukemia. *Neoplasma*. 2008;55:51–4.
  15. Secchiero P, Tiribelli M, Barbarotto E, Celeghini C, Michelutti A, Masolini P, et al. Aberrant expression of TRAIL in B chronic lymphocytic leukemia (B-CLL) cells. *J Cell Physiol*. 2005;205:246–52. <https://doi.org/10.1002/jcp.20392>.

## ACKNOWLEDGEMENTS

We thank all study participants who agreed to participate in the study.

## AUTHOR CONTRIBUTIONS

JS designed and coordinated the study. PG-M, AMD, MC-G, JMSM, and JS were involved in the generation of genetic data from the CRuCIAL cohort and drafted the manuscript. JS obtained funding and performed data quality control and genetic association analyses. MGN, Rth, and YL provided the functional raw data from the HFGP cohort and JS performed the statistical analysis of the functional data. YB, SL, AM, FH-M, YN, PG-S, BE, RM, MG, AP, GB, RM, Mál-N, TC-L, HT, IG, JJR-S, VM, RM-G, MG-Á, JL, AJ, AB, DC, ML, MA, KH, YL, SdS, DC, FC, AF, and MJ provided CLL patients and/or healthy controls. FC, AC-G, JC, ADN, SB, and SS provided data and participated in the computation of the PRSs. All authors contributed to and edited and approved the final version of the manuscript.

## FUNDING

This work was partially supported by the European Union's Horizon 2020 research and innovation program (grant No 856620); grants from the Instituto de Salud Carlos III (Madrid, Spain; PI17/02256 and PI20/01845); Consejería de Economía, Conocimiento, Empresas y Universidad (Granada, Spain; A-CTS-448-UGR18); Consejería de Transformación Económica, Industria, Conocimiento y Universidades (Sevilla, Spain; PY20/01282); Generalitat de Catalunya (17SGR437); Gilead Sciences Fellowship (GLD17/00282); the “Xarxa de Bancs de tumors” sponsored by Pla Director d'Oncologia de Catalunya (XBTC); the Associazione Italiana per la Ricerca sul Cancro and Fondazione Cariplo (TRIDEO 16923 and AIRC IG21436); the Spanish Association Against Cancer (AECC) Scientific Foundation grant GCTRA18022MORE; and the Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), action Genrisk. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41408-022-00676-8>.

**Correspondence** and requests for materials should be addressed to Juan Sainz.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022