Chronic lymphocytic leukemia patients with *IGH* translocations are characterized by a distinct genetic landscape with prognostic implications

Claudia Pérez-Carretero[1,2], María Hernández-Sánchez[1,2,3], Teresa González[1,2], Miguel Quijada-Álamo[1,2], Marta Martín-Izquierdo[1,2], Jesús-María Hernández-Sánchez[1,2], María-Jesús Vidal[4], Alfonso García de Coca[5], Carlos Aguilar[6], Manuel Vargas-Pabón[7], Sara Alonso[8], Magdalena Sierra[9], Araceli Rubio-Martínez[10], Julio Dávila[11], José R. Díaz-Valdés[12], José-Antonio Queizán[12], José-Ángel Hernández-Rivas[13], Rocío Benito[1,2], Ana E. Rodríguez-Vicente[1,2*] and Jesús-María Hernández-Rivas[1,2*]

1. Universidad de Salamanca, IBSAL, Centro de Investigación del Cáncer, IBMCC-CSIC, Salamanca, Spain.
2. Servicio de Hematología, Hospital Universitario de Salamanca, Salamanca, Spain.
3. Department of Medical Oncology, Dana Farber Cancer Institute, Boston, Massachusetts, USA.
4. Servicio de Hematología, Hospital Universitario, León, Spain.
5. Servicio de Hematología, Hospital Clínico, Valladolid, Spain.
6. Servicio de Hematología, Complejo Hospitalario de Soria, Soria, Spain.
7. Servicio de Hematología, Hospital Jarrio, Asturias, Spain.
8. Servicio de Hematología, Hospital Universitario Central de Asturias, Oviedo, Spain.
9. Servicio de Hematología, Hospital Virgen de la Concha, Zamora, Spain.
10. Servicio de Hematología, Hospital Miguel Servet, Zaragoza, Spain.
11. Servicio de Hematología, Hospital Nuestra Señora de Sonsoles, Ávila, Spain.
12. Servicio de Hematología, Hospital General de Segovia, Segovia, Spain.
13. Servicio de Hematología. Hospital Universitario Infanta Leonor. Universidad Complutense. Madrid, Spain.

* AERV and JMHR contributed equally to this work

Corresponding author

Ana E. Rodríguez Vicente, PhD

IBSAL, IBMCC-Centro de Investigación del Cáncer (USAL-CSIC)

Campus Miguel de Unamuno

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ijc.33235

37007 Salamanca (Spain)

Phone: +34 923294812

E-mail: anaerv@hotmail.com

Running title

Genetic landscape of CLL with *IGH* rearrangement

Keywords

Chronic lymphocytic leukemia, chromosomal translocations, high-throughput sequencing, prognostic biomarkers, cytogenetics, clinical molecular genetics.

Scientific category: Research Article

Text word count (without references): 4377

Abstract word count: 239

Figures: 3; 7 supplementary

Tables: 3; 8 supplementary

Reference count: 50

Abbreviations

CLL: chronic lymphocytic leukemia; COSMIC: Catalogue of Somatic Mutations in Cancer; DLBCL: diffuse large B-cell lymphoma; ExAC: Exome Aggregation Consortium; FISH: Fluorescence in situ hybridization; FL: follicular lymphoma; GATK: Genome Analysis Toolkit; ICGC: International Cancer Genome Consortium; IGH: immunoglobulin heavy chain; IGHR-CLLs: CLL patients with IGH rearrangements;

IGHV: immunoglobulin heavy-chain variable; IGV: Integrative Genomics Viewer; IWCLL: International Workshop on CLL; MACS: magnetically activated cell sorting; NGS: next-generation sequencing; NHL: non-Hodgkin lymphoma; OS: Overall survival; PCAWG : Pan-Cancer Analysis of Whole Genomes.; TFT: Time to first treatment; VAF: variant allele frequency; VCF: variant calling files.

Novelty and Impact

In this work, we assess for the first time the genetic landscape of CLLs with *IGH* rearrangements by targeted NGS, characterizing recurrently mutated genes with prognostic implications and demonstrating that these entities exhibit an intermediate mutational profile between CLL and non-Hodgkin lymphoma. Moreover, our findings showed that the incorporation of NGS and the IGH-probe in the CLL-FISH panel used in clinical routine could be extremely useful, especially for elucidating prognosis in 'normal FISH' cases.

ABSTRACT

Chromosome 14q32 rearrangements/translocations involving the immunoglobulin heavy chain (*IGH*) are rarely detected in chronic lymphocytic leukemia (CLL). The prognostic significance of the *IGH* translocation is controversial and its mutational profile remains unknown. Here, we present for the first time a comprehensive next-generation sequencing (NGS) analysis of 46 CLL patients with *IGH* rearrangement (IGH*R*-CLLs) and we demonstrate that IGH*R*-CLLs have a distinct mutational profile with recurrent mutations in *NOTCH1, IGLL5, POT1, BCL2, FBXW7, ZMYM3, MGA, BRAF* and *HIST1H1E* genes. Interestingly, *BCL2* and *FBXW7* mutations were significantly associated with this subgroup and almost half of *BCL2, IGLL5* and *HISTH1E* mutations reported were previously identified in non-Hodgkin lymphomas (NHL). Notably, *IGH/BCL2* rearrangements were associated with a lower mutation frequency and carried *BCL2* and *IGLL5* mutations, while the other IGH*R*-CLLs had mutations in genes related to poor prognosis (*NOTCH1, SF3B1* and *TP53*) and shorter time to first treatment (TFT). Moreover, IGH*R*-CLLs patients showed a shorter TFT than CLL patients carrying 13q-, normal FISH and +12 CLL, being this prognosis particularly poor when *NOTCH1, SF3B1, TP53, BIRC3* and *BRAF* were also mutated. The presence of these mutations not only was an independent risk factor within IGH*R*-CLLs, but also refined the prognosis of low-risk cytogenetic patients (13q-/normal FISH). Hence, our study demonstrates that IGH*R*-CLLs have a distinct mutational profile from the majority of CLLs and highlights the relevance of incorporating NGS and the status of *IGH* by FISH analysis to refine the risk-stratification CLL model.

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is a disease that displays extreme clinical heterogeneity, clearly reflecting the marked biological diversity, which has led to the identification of a plethora of prognostic markers[1-4]. Chromosomal abnormalities are the hallmark of the disease and their correlation to the clinical course has contributed to patients risk stratification since the 2000s[5]. In the last years, CLL molecular and cellular biology has been enriched by seminal insights that have led to a better understanding of CLL pathogenesis[2] and, consequently, to the identification of molecular markers whose evaluation is well-established in clinical routine, such as the *IGHV* mutational status or *TP53* gene abnormalities. The integration of these markers together with the new relevant genetic alterations reported in next-generation sequencing (NGS) studies, specifically those of *NOTCH1*, *SF3B1* and *BIRC3* genes, could be used to refine Döhner hierarchical cytogenetic model[2,6-12].

Although more than 80% of CLL patients carry cytogenetic alterations, chromosome 14q32 rearrangements/translocations involving the immunoglobulin heavy chain gene (*IGH*) was considered a rare aberration affecting fewer than 4% of CLL patients[5,13]. Nevertheless, with the emergence of new molecular approaches and large-scale genomic studies in CLL, a higher incidence of *IGH* rearrangements has been reported in the recent years (5-15%) [14-16]. This cytogenetic abnormality contributes to CLL pathogenesis by deregulating the *IGH*-partner genes[17,18] and their prognostic significance is still controversial. Previous studies have shown that patients carrying 14q32 rearrangements (IGH*R*-CLLs) have an intermediate-adverse outcome[19-21], particularly when compared with favorable and intermediate-risk cytogenetics[22,23]. However, some studies have reported that patients carrying 14q32 rearrangements with *BCL2* have a better clinical course[24,25].

CLL patients with *IGH* rearrangements are still poorly characterized at the molecular level, partly due to the low incidence of cases, the IGH*R* co-occurrence with other cytogenetic alterations, and the difficulty of distinguishing between IGH*R*-CLLs and

forms of non-Hodgkin lymphoma (NHL)[26]. Furthermore, the *IGH* probe is not included in the classic four-probe CLL FISH panel for the 13q14, 12p11.1-q11, 11q22 and 17p13 regions used in routine clinical practice[26], which is partially responsible for this subgroup passing unnoticed. In this study, we characterize the genetic landscape of CLL patients with 14q32/*IGH* translocations for the first time, demonstrating that IGH*R*-CLLs have a distinct mutational profile from other classic cytogenetic groups of CLLs, dependent on whether *BCL2* is involved or not in the *IGH* rearrangement, and as well as the presence of certain mutations. Taken together, our results improve our understanding of the molecular underpinnings of this cytogenetic CLL subgroup, allowing us to refine the prognosis of IGH*R*-CLL patients.

METHODS

Patients

The study was based on 862 CLL patients, diagnosed according to the International Workshop on CLL (IWCLL) criteria[27,28]. All of them were screened for *IGH* translocation and positive cases for *IGH* rearrangement were individually reviewed to rule out the possibility that they represented a different lymphoproliferative disorder (see **Supplementary Methods**). Samples and clinical data were collected from 16 Spanish institutions.

Mutational analysis was performed in 233 untreated CLL patients: 46 with 14q32/*IGH* rearrangements and 187 as the control group. Patients in the control group were selected according to sample and clinical data availability and absence of treatment and were representative of the disease in terms of demographic and clinical characteristics **(Supplementary Table S1)**. Patients risk classification criteria is described in **Supplementary Methods** and a diagram of the patients included in the different outcome analyses is shown in **Supplementary Fig. S1**.

In the IGH*R*-CLL group, the median time between diagnosis and *IGH* rearrangement detection was 1 month (range, 0-117 months), and the median follow-up was 57 months (range, 1-157 months). Within IGH*R*-CLLs, 31/46 (67.4%) received treatment after FISH

test, with a median TFT of 19 months (range: 7-30). Most of them (93.5%) received conventional chemoimmunotherapy and 2 patients were treated with ibrutinib.

The study was approved by the local ethical committee (*Comité Ético de Investigación Clínica, Hospital Universitario de Salamanca*). Written informed consent was obtained from all participants before they entered the study.

Fluorescence *in situ* hybridization (FISH)

Interphase FISH was performed on peripheral blood or bone marrow samples using the following commercially probes: ATM, CEP12, D13S319 and TP53 (Vysis, Abbott Laboratories, IL, USA). Dual color break-apart FISH probes were performed for *IGH*/*BCL2* and *IGH*/*BCL6* translocations. The methods used for FISH analysis have been described elsewhere[29]. Signal screening was carried out in at least 200 cells with well-delineated fluorescent spots. In all cases, a score of ≥10% was considered positive, based on the cut-off value used by our laboratory.

Next-generation sequencing

NGS studies were performed in 233 cases and in the same sample as the FISH test. Genomic DNA was isolated from peripheral blood or bone marrow by magnetically activated cell sorting (MACS) CD19+ B-lymphocytes. B cell purity was greater than 98% by flow cytometry, as previously described in our group[30]. The Agilent SureSelect^{QXT} Target Enrichment system for Illumina Multiplexed Sequencing (Agilent Technologies, Santa Clara, CA, USA) was used to produce libraries of exonic regions from 54 genes CLL-related as well as from *BCL2*, *IGLL5* and *NOTCH1* UTR regions (**Supplementary Methods**). Genes included in the custom-designed panel[31,32] are involved in CLL pathogenesis and the UTR regions were considered due to the previous identification of *IGLL5*, *BCL2* and *NOTCH1* UTRs somatic mutations in CLL[8,33,34]. (**Supplementary Table S2**). Paired-end sequencing (151-bp reads) was run on the Illumina NextSeq instrument (Illumina, San Diego, CA, USA).

Data analysis

Raw data quality control was performed with FastQC (v0.11.8) and Picard tools (v2.2.4) to collect sequencing metrics. Demultiplexed files (FASTQ) were aligned to the reference genome (GRCh37/hg19 genome), read duplicates were marked with SAMTools (v1.3.1) and post-alignment was performed with GATK (v3.5). Coverage for each region was assessed using BEDTools (v.2.26.0). A minimum quality score of Q30 was required for ensuring high-quality sequencing results. Finally, somatic variant calling, and annotation were performed using an in-house pipeline, based on VarScan (v2.4) and ANNOVAR (v.2017Jul16), respectively.

Median coverage of target regions was 600 reads/base, with at least 100X in 97% of them. To validate variants detected with VAF <5% using the custom panel, samples were conducted to resequencing using different amplicon-based approaches (Illumina Nextera XT/454 Roche[30]) with read depth above 1000X, allowing to report variants down to 2% (**Supplementary Methods**).

Data was then filtered according to the severity of the consequence, considering variants that lead to an amino acid change in the protein sequence (missense, nonsense, frameshift) and those in the splice site and UTRs. To discard single nucleotide polymorphisms (SNPs), minor allelic frequencies (MAFs) were consulted in several databases (dbSNP, 1000 genomes, ExAC and our in-house database) and only variants with a MAF of <0.01 were selected for further analysis. In addition, variants with a VAF between 40-60% or greater than 90% were manually reviewed prioritizing variants described in *in silico* tools (Polymorphism Phenotyping v2 (PolyPhen-2), Sorting Intolerant From Tolerant (SIFT) and ClinVar) as deleterious, damaging, pathogenic or likely pathogenic.

Aligned reads were manually reviewed with the Integrative Genomics Viewer (IGV) to confirm and interpret variant calls and reduce the risk of false positives. Variants described in the Catalogue of Somatic Mutations in Cancer database (COSMIC82 database) or mutations in driver genes previously described in seminal papers were rescued for the analysis (CLL and non-Hodgkin lymphoma)[7,8,33,35-37]. Manually screening

in VarSome and International Cancer Genome Consortium (ICGC) Databases was performed for assessing the functional impact of mutations.

Statistical analysis

Statistical analyses were performed using IBM SPSS v23.0 for Windows (IBM Corp., Armonk, NY, USA) and SDM-PSI v6.21 software for the false discovery rate (FDR) correction in multiple comparisons. Continuous variables were analyzed with the Mann-Whitney U test, while the chi-square and Fisher's exact tests were used to assess associations between categorical variables. Overall survival (OS) and time to first therapy (TFT) were calculated from the date FISH test was performed to the date of death, first treatment or last follow-up (considering disease-unrelated deaths as competing events). Statistically significant variables related to OS and TFT were estimated by the Kaplan-Meier method, using the log-rank test to compare the curves of each group. Univariate and multivariate analyses of the OS and TFT employed the Cox regression method. Results were considered statistically significant for values of $P<0.05$. FDR was used to correct *P*-values for multiple hypothesis testing when appropriate, by applying the Benjamini and Hochberg method[38]. Adjusted *P*-values (*Q*-values) were considered significant when $Q < 0.1$.

RESULTS

CLL patients with *IGH* translocations have a distinct mutational profile with high mutation frequencies in *NOTCH1, BCL2, FBXW7, ZMYM3* and *MGA*

NGS analysis of the 233 CLL patients revealed that 75% of cases had at least one mutation in any of the 54 genes included in the targeted-NGS approach, and the median frequency of mutations per patient was 2 (range: 0-7). The most frequently mutated genes were *NOTCH1* (19.3%), *IGLL5* (15%), *SF3B1* (10.7%), *TP53* (10%), *ATM* (9%), *POT1* (8.5%), *RPS15* (6.9%), *CHD2* (6%), *NFKBIE* (5.1%), *BIRC3* (5.1%) and *XPO1* (4.3%).

Regarding the 46 IGH*R*-CLLs, we identified a total of 109 mutations located in 35 genes. The median frequency of mutations per patient was 2 (range: 0-6), and 82% of patients (38/46) harbored at least one mutation. Moreover, 61% of patients (28/46) presented more than one mutated gene. The most frequently mutated genes in this cohort were *NOTCH1* (30.4%), *IGLL5* (17.4%), *SF3B1* (13%), *POT1* (13%), *TP53*, *BCL2*, *FBXW7*, *ZMYM3* and *MGA* (8.7% each) followed by *BRAF*, *EGR2* and *RPS15* (6.5% each) (**Fig. 1A; Supplementary Table S3**). Other genes such as *ATM* (4.3%) or *CHD2* and *MYD88* (2.2% each) were mutated at low frequencies.

The comparison between the mutational profiles of IGH*R*s-CLLs and the control group showed higher mutation frequencies in *NOTCH1*, *BCL2*, *FBXW7*, *ZMYM3* and *MGA* within IGH*R*-CLLs, especially those of *BCL2* and *FBXW7* (Q=0.048, Q=0.06, respectively) **(Fig. 1A; Supplementary Table S3).**

Furthermore, 61% of IGH*R*-CLLs (28/46) carried additional FISH alterations (**Fig 1B**). Their mutational profile was analyzed with respect to the presence of IGH*R* together with 13q, 11q, 17p deletion or trisomy 12, and only *TP53* mutations were significantly associated with 17p or 11q deletion in IGH*R*-CLLs (Q=0.048). We observed that the mutational profile of patients with IGH*R* as a sole aberration (18/46) was similar to that of the entire IGH*R*-CLL cohort: *NOTCH1* (33.3%), *IGLL5* (27.8%), *SF3B1* (16.7%), *BCL2*, *ZMYM3*, *MGA* and *FUBP1* (11.1% each) followed by *FBXW7* and *BRAF* (5.6% each). All mutation frequencies are shown in **Supplementary Table S4**.

Interestingly, we reported a higher incidence of *IGLL5*, *BCL2* and *HIST1H1E* mutations in this subgroup compared to the described in previous large-scale CLL studies[7,8] (**Fig. 1B**). IGH*R*-CLL patients showed *IGLL5* mutations targeting the signal peptide domain (4/10) and the 5'UTR region (3/10), *BCL2* mutations affecting the 5'UTR region (2/6) and the exon 2 (4/6), and *HIST1H1E* mutations located in the exon 1 (**Fig 2**). According to the ICGC Database, most of the coding mutations in *IGLL5* (6/7), *BCL2* (3/4) and *HIST1H1E* (1/2) identified in our study, had functional impact in the gene function (**Table 1**). In addition, six out of 17 mutations detected in the aforementioned three genes were

previously described in non-Hodgkin lymphomas (as reported in the COSMIC and ICGC database and whole-exome and whole-genome data from NHL patients[35-37,39,40]).

Moreover, five of the mutations reported in *IGLL5* and *BCL2* were located in the 5'UTR of the gene. Specifically, the novel *BCL2* recurrent mutation identified in the 5'UTR region (genomic position chr18:60985900) was exclusively found in IGH*R*-CLLs when compared with the control group (P=0.048) **(Supplementary Fig. S2**).

Detailed lists of the mutations detected in the IGH*R*-CLLs and the control group are shown in **Supplementary Tables S5, S6 and S7**.

CLL patients with *IGH/BCL2* exhibit a lower mutation frequency and a different mutational profile than patients with other *IGH* translocations

We next sought to assess whether the mutational landscape changes depend on the *IGH* translocation partner, e.g. *BCL2* and *BCL6*. In our study, 13/46 patients (28%) carried *IGH/BCL2* translocation (**Figure 1B**) and 2/46 (4.3%) harbored an *IGH/BCL6* rearrangement (ID 8 and 20). Due to the small number of *BCL6* rearrangements, we performed further analysis comparing *IGH/BCL2* vs. the rest of IGH*R* cases.

In the IGH*R* patients, fewer CLLs with *IGH/BCL2* translocation had mutations in at least one gene compared with the subgroup with other *IGH* translocations (7/13, 54% *vs.* 31/33, 94%; P=0.001). The median mutation frequency per patient was significantly lower in the group with *IGH/BCL2* compared with that without it (1 *vs.* 2, P=0.030).

The most frequently mutated genes in the *IGH/BCL2* group were *BCL2* (23%), *IGLL5* (23%), *HIST1H1E* (15%) and *NOTCH1* (15%), whereas for all other IGH*R*-CLLs, the most frequently mutated genes were *NOTCH1* (36%), *SF3B1* (18%), *POT1* (18%), *TP53* (12%), and *FBXW7* (12%) (**Supplementary Fig. S3A**). It is worth mentioning that neither *TP53* nor *SF3B1* mutations, widely associated with poor prognosis, were detected in CLL patients with an *IGH/BCL2* translocation, reflecting a different mutational profile from all other IGH*R*-CLLs. The mutational analysis of 9 *IGH/BCL2* cases previously reported in a WES/WGS study of CLL[8] also showed the presence of mutations in *BCL2*,

and *NOTCH1*, and the absence of poor-prognosis genes such as *TP53* or *SF3B1* (**Supplementary Fig. S3B).** They have been previously reported in WES of However, no statistically significant associations were detected in our analysis, probably due to the small number of cases (**Supplementary Table S8).**

In *IGH/BCL2* cases that also harbored *BCL2* mutations, we observed that 60-87% of the cells carried the rearrangement, while *BCL2* mutations VAFs range from 11% to 40%, suggesting that somatic mutations occurred later in time than the rearrangement (**Table 1**).

Patients carrying *IGH* translocations exhibit an intermediate-adverse outcome

We also analyzed the clinical and biological characteristics of IGH*R*-CLLs within the entire cohort (n=862) (**Table 2)**. Patients carrying this cytogenetic alteration showed a higher incidence of poor prognosis markers such as Binet stage B or C (Q=0.039), high β2-microglobulin (Q=0.0007) and lactate dehydrogenase levels (Q=0.054), unmutated *IGHV*) (Q=0.054) and need for treatment (Q=0.007). In addition, two IGH*R*-CLLs developed Richter syndrome during follow-up (patient IDs: 18 and 35). Regarding the presence of additional cytogenetic alterations, 34.8% of IGH*R*-CLL patients (16/46) carried trisomy 12, showing significant co-occurrence of the two events (trisomy 12 and IGH*R*) (Q=0.0007). By contrast, the presence of the 13q deletion in IGH*R*-CLLs was significantly less frequent than in CLLs without *IGH* rearrangements (Q=0.0014) (**Table 2)**.

Within IGH*R*-CLLs, 31/46 (67.4%) received treatment after FISH test, with a median TFT of 19 months (95% CI, 7-30 months). Patients with an *IGH* translocation showed shorter TFT than the 13q- and normal FISH subgroups (median: 19 *vs.* 120 and 184 months; p<0.0001, p<0.0001), and longer TFT than the 11q- and 17p- subgroups (19 *vs.* 5, 6 months; p=0.042, p=0.31). The median TFT of the +12 subgroup was slightly higher than that of IGH*R*-CLLs (28 *vs.* 19 months; p=0.37). In terms of OS, we observed similar trends (**Supplementary Fig. S4A**). Differences in outcome among the cytogenetic subgroups were consistent with the prevalence of unfavorable clinical and biological

features in IGH*R*-CLLs, suggesting that this subgroup exhibits an intermediate-adverse prognosis. In addition, the clinical comparison between IGHR-CLLs and control CLLs selected for the mutational analysis (n=233) showed quite similar results to the presented in this section, also demonstrating that control group was representative of the entire cohort **(Supplementary Table S1)** (**Supplementary Fig. S4B**).

In our entire cohort (n=862), 31% of patients showed no alterations using 13q14/ D13S319, 12p11.1-q11/CEP12, 11q22/ATM, 17p13/P53 probes. However, it is worth mentioning that 6.7% of patients who would be classified as normal FISH in our cohort using the four-probe CLL FISH panel customarily used in routine clinical practice, actually carried an *IGH* rearrangement. The presence of this cytogenetic alteration had a negative effect on the time to first treatment within this group of patients: CLL patients with IGH*R* as the sole FISH abnormality had a significantly shorter TFT than those without any FISH aberration (23 *vs.* 120 month, P =0.01) (**Fig. 3A**).

The presence of the *IGH/BCL2* translocation was associated with mutated *IGHV* (P=0.001), and patients with this alteration showed a longer TFT than those with another IGH*R* (56 *vs.* 4 months, P=0.05). By contrast, the presence of *IGH/BCL2* rearrangement was not associated to any additional cytogenetic alteration (13q-, Q=0.822; +12, Q=0.822; 11q-/17p-, Q=0.822) and there was no significant difference in terms of OS between patients with *IGH/BCL2* and patients with other *IGH* rearrangements (P=0.433) (**Supplementary Fig. S5**).

Genetic mutations refine the prognosis of IGH*R* and low-risk cytogenetic CLL patients

IGH*R*-CLL untreated patients with at least one mutated gene showed a shorter TFT than IGH*R*-CLLs without gene mutations (10 months *vs.* median TFT not reached, P=0.026) (**Fig. 3B**). These differences were more significant among recurrent gene mutations previously associated with worse prognosis (*NOTCH1, SF3B1, TP53, BIRC3* and *BRAF*) (2 *vs.* 88 months, P <0.0001) **(Fig. 3C**). Specifically, TFT was shorter in IGH*R* patients with *TP53* mutations (0 *vs.* 23 months, P <0.0001) as well as with *BRAF* mutations (2 *vs.*

23 months, *P* =0.042) (**Supplementary Fig. S6**). In contrast, the presence of *IGLL5* or *BCL2* mutations showed a better impact in terms of TFT, as IGH*R*-CLL patients with mutated *IGLL5* or *BCL2* showed a longer TFT than those without mutations in any of these genes (median TFT not reached *vs.* 9 months, *P*=0.001) **(Fig. 3D**).

In the univariate analysis, other variables associated with a shorter TFT were Binet's stage B/C (*P*=0.001), splenomegaly (*P*=0.025), unmutated *IGHV* status (*P*=0.013), *TP53* disruption/mutation (*P*=0.003) and the absence of *IGLL5/BCL2* mutations (*P*=0.008). Only the presence of mutations in *NOTCH1, SF3B1, TP53, BIRC3* and *BRAF* was significantly related to a shorter TFT within IGH*R*-CLL patients in the multivariate analysis (HR=0.255, 95% CI=0.07-0.9, *P*=0.030) (**Table 3**).

Since the presence of mutations in these five genes has a prognostic impact within IGH*R*-CLL patients as well as in the control group (median TFT not reached *vs.* 12 months, *P*<0.0001) (**Supplementary Fig. S7**), we propose an integrated mutational and cytogenetic model to account for our observations in the studied cohort (187 control and 46 IGHR-CLLs). Low-risk patients in our control series (13q-/normal FISH, n=134) segregated into two groups according to the presence of mutations in *NOTCH1, SF3B1, TP53, BIRC3* and *BRAF* (median TFT not reached *vs.* 24 months, *P*<0.0001) **(Fig. 3E)**. These mutations also contributed to a worse outcome in intermediate-risk patients (IGH*R* /+12, n=72) (56 *vs.* 2 months, *P*<0.0001). However, the small number of cases with these mutations was insufficient to demonstrate a statistically significant difference in clinical impact in the high-risk cytogenetics subgroup (11q-/17p-, n=27) (4 *vs.* 0 months, *P*=0.580). The median TFT of those intermediate-risk patients with mutations was similar to that of patients with high-risk cytogenetic alterations (2 *vs.* 5 months; *P*=0.548), and the TFT of low-risk patients with mutations was not significantly different from that of intermediate-risk patients without mutations in any of the five genes (24 *vs.* 56, *P*=0.210). Therefore, by including *NOTCH1, SF3B1, TP53, BIRC3* and *BRAF* mutations in the cytogenetic model, approximately 27.6% (37/134) of low-risk patients were reclassified into an intermediate-risk subgroup, and 51% (35/72) of intermediate-risk patients were reclassified into a high-risk subgroup **(Fig. 3E)**.

DISCUSSION

The identification of novel recurrent mutations in CLL has provided a more comprehensive perspective on the genomic landscape and the biological mechanisms underlying the clinical heterogeneity of the disease[2-4,7,8]. Previous studies have shown that CLLs carrying *IGH* rearrangement could have a worse outcome than low-risk cytogenetic CLL patients[22,23]. However, their clinical course and molecular characteristics are not well defined[19-24]. Here, we adopted a targeted NGS approach to assess for the first time the mutational profile of 46 IGH*R*-CLL patients.

Overall, the mutational analysis revealed that IGH*R*-CLL patients had a high incidence of mutations, not only in well-known CLL drivers such as *NOTCH1, SF3B1, POT1, TP53* and *FBXW7* —previously described in unselected large CLL cohorts—, but also in less commonly mutated genes such as *BCL2*, *FBXW7*, *ZMYM3* and *MGA*[7,8,33], being *BCL2* and *FBXW7* significantly associated with the *IGH* translocation (**Fig 1A**). Although we observed the co-occurrence between IGH*R* and trisomy 12 previously described, we demonstrated that IGH*R*-CLLs mutational profile did not depend on the presence of additional cytogenetic aberrations: CLLs with only IGH*R* also exhibited a high mutation frequency in genes well-known associated with trisomy 12 such as *NOTCH1*[8], as well as in the majority of recurrently mutated genes in the entire IGH*R*-CLL cohort (**Fig. 1B**).

Strikingly, several IGH*R*-CLLs showed mutations in *IGLL5, BCL2* and *HIST1H1E*, mainly those with *IGH/BCL2* translocations **(Supplementary Fig. S3A)**. Although these gene mutations have been detected at low frequencies in other CLL cohorts[7,8], they have been extensively reported in other hematological malignances such as diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL)[35-37,39-44] **(Table 1)**. Considering that our series of IGH*R*-CLL patients is well characterized at the immunophenotypic and clinical levels, the presence of mutations previously described in lymphomas suggests that patients with *IGH* translocation are a cytogenetic subgroup with a mutational profile distinct from the other CLLs and, probably, with different genetic mechanisms underlying their disease pathogenesis. Here, we demonstrated that IGH*R*-CLLs had an intermediate

genetic landscape between those of CLL and non-Hodgkin lymphoma, and we suggest that the mutational analysis of patients with *IGH* translocations such as *IGH/BCL2* could help distinguish between CLL and NHL cases.

In this work, we have reported mutations in coding and non-coding regions of *BCL2* and *IGLL5* and, specifically, we identified a novel recurrent 5'UTR mutation in *BCL2*. Although this mutation has not been previously described, proximal mutations have been detected in other CLL and NHL studies[8,34,44]. Moreover, the vast majority of *BCL2* mutations reported in Puente *et al*. were detected in cases harboring *IGH/BCL2* translocations[8], which is consistent with our results (**Supplementary Fig. S3**). Regarding *IGLL5*, a previous study identified 5'UTR and coding mutations in low-risk *IGHV*-mutated CLLs as well as in the presence of rearrangements[33]. These statements are consistent with our results, as most of patients harboring mutations in both genes exhibited mutated *IGHV* and were associated with longer time to first treatment **(Fig. 3D**). In addition, data compiled in the ICGC repository suggested a functional impact of the *BCL2* and *IGLL5* coding mutations in the gene function. However, the functional impact of UTR somatic mutations has not been well-established yet. Puente *et al.* demonstrated the negative impact of the 3'UTR mutation in *NOTCH1*[8] and recent findings from the Pan-Cancer Analysis of WholeGenomes (PCAWG) Consortium identified novel driver candidates, including mutations in UTR regions, with a potential role in CLL pathogenesis[45]. Nevertheless, further investigation is needed in order to determine the importance of *BCL2* and *IGLL5* non-coding mutations in CLL.

The clinical impact of *IGH* translocations is currently under discussion[19-25]. The median TFT was shorter in IGH*R*-CLLs than that of patients with low-risk cytogenetic alterations, but similar to that of patients with trisomy 12 (**Supplementary Fig. S4A)**, indicating that *IGH* translocations could be associated with an intermediate-adverse outcome. Indeed, 6,7% of CLL patients who would be considered 'normal FISH' using the customary four-probe CLL FISH panel in our study, carried the *IGH* translocation and also had a worse prognosis than CLLs lacking IGH*R* (**Fig. 3A)**, thus highlighting the

value of including the *IGH* probe in the CLL FISH panel to improve patient outcome prediction[25].

Previous studies have shown that patients with a *IGH/BCL2* translocation had a favorable clinical course, similar to that of patients with low-risk chromosomal alterations, whereas patients with other *IGH* rearrangements had a similar prognosis to the high-risk subgroups[24,25]. In our study, patients with *IGH/BCL2* not only were associated with *IGHV*-M and longer TFT (**Supplementary Fig. S5A)**, but also exhibited lower mutation rate compared with other *IGH* translocations that may contribute to understand why these entities have a better prognosis than the rest of IGH*R*. Regarding the mutational profile of *IGH/BCL2* translocations, a previous large-scale CLL study showed mutations in *BCL2, IGLL5* and *NOTCH1* within 9 *IGH/BCL2* cases, which strongly supports our findings (**Supplementary Fig. S3B)**. On the other hand, IGH*R*-CLLs without *IGH/BCL2* rearrangement presented higher mutation frequencies in genes related to bad prognosis, such as *NOTCH1, SF3B1, TP53, BRAF* and *RPS15* (**Fig. 1B; Supplementary Fig. S3A**). The high frequency of these mutations may reflect a genomic instability in IGH*R*-CLLs without *IGH/BCL2*, which could be also influenced by the role of the translocated partner in the rearrangement. Furthermore, two IGH*R*-CLLs developed Richter transformation to DLBCL. One of them harbored *IGH/BCL2* rearrangement together with trisomy 12 and *NOTCH1*, and the other patient had an IGH*R* with unknown partner, *NOTCH1* and *TP53* mutations. These observations are in line with previous findings regarding the molecular pathways frequently altered at transformation[46,47]. Altogether, these molecular characteristics could be the underlying mechanisms of the IGH*R*-CLLs poorer outcome[25].

In our cohort, patients harboring *NOTCH1, SF3B1, TP53, BIRC3* or *BRAF* mutations experienced an adverse clinical course (**Fig. 3C**), which is consistent with previous studies[6-8,11,48]. Within IGH*R*-CLLs, the presence of these mutations contributed to shorter TFT being identified as an independent adverse prognostic factor (**Table 3**). Specifically, IGH*R*-CLL patients harboring *BRAF* mutations exhibited an adverse outcome **(Supplementary Fig. S6A**), which corroborates previous results showing that patients carrying these mutations display an aggressive disease [49,50].

Therefore, the present study proposes an integrated mutational and cytogenetic model for CLL prediction that includes IGH*R* and *BRAF* mutational status as novel components with respect to previous prognostic models[9,10]. The presence of mutations in any of the aforementioned five genes caused a significant shift to a more aggressive outcome in low (13q-/normal FISH) and intermediate-risk (+12/IGH*R*) CLLs, refining their prognosis and providing information that could help in therapeutic decisions. Interestingly, low-risk patients with mutations in *NOTCH1*, *SF3B1*, *TP53*, *BIRC3* or *BRAF* still had a significantly better outcome than did intermediate-risk patients with any of those mutations (**Fig. 3E**). These results may indicate that the co-occurrence of cytogenetic abnormalities and gene mutations could have different clinical impacts, depending on the type of the genetic alterations involved.

In conclusion, our study revealed significant differences in the mutational profile and the frequencies of CLL-mutated genes in patients with *IGH* rearrangements. The distribution of genetic mutations differed within the IGH*R*-CLL subgroup: patients with *IGH/BCL2* translocation had higher frequencies of *BCL2* and *IGLL5* mutations than those without the translocation. Conversely, patients with other IGH*R* showed higher mutation frequencies of genes related to bad prognosis (*NOTCH1*, *SF3B1*, *TP53*, *BIRC3* and *BRAF*) than did those with the *IGH/BCL2*. Notably, the presence of those somatic mutations enables us to refine not only the prognosis of IGH*R*-CLLs but also the outcome of low-risk cytogenetic patients. Thus, this mutational analysis improves our understanding of the molecular heterogeneity of CLL patients and could help improve prognostic stratification of CLLs.

Conflict of interest

The authors declare no competing interests

Data Accessibility

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics statement

The present study was approved by the local ethics committee (Comité Ético de Investigación Clínica, Hospital Universitario de Salamanca). Written informed consent was obtained from all participants before they entered the study.

Acknowledgments

We are grateful to S. Santos, C. Miguel, I. Rodríguez, S. González, T. Prieto, M.Á. Ramos, A. Martín, A. Díaz, A. Simón, M. del Pozo, V. Gutiérrez and S. Pujante from the Centro de Investigación del Cáncer, Salamanca, for their technical assistance.

Funding

This work was supported by grants from the Spanish Fondo de Investigaciones Sanitarias PI15/01471, PI18/01500, Instituto de Salud Carlos III (ISCIII), European Regional Development Fund (ERDF) "Una manera de hacer Europa", "Consejería de Educación, Junta de Castilla y León" (SA271P18), "Proyectos de Investigación del SACYL", Spain: GRS 1847/A/18, GRS1653/A17, "Fundación Memoria Don Samuel Solórzano Barruso", by grants (RD12/0036/0069) from Red Temática de Investigación Cooperativa en Cáncer (RTICC) and Centro de Investigación Biomédica en Red de Cáncer (CIBERONC CB16/12/00233). C Pérez Carretero was supported by an "Ayuda predoctoral en Oncología" (AECC) and is a recipient of a PFIS grant (FI19/00191) from Instituto de Salud Carlos III; M Hernández Sánchez was supported by a grant from FEHH/Janssen ("Sociedad Española de Hematología y Hemoterapia") and now holds a Sara Borrell post-doctoral contract (CD19/00222) from the Instituto de Salud Carlos III (ISCIII). PFIS grant and Sara Borrell post-doctoral contrat is co-founded by Fondo Social Europeo (FSE) "El Fondo Social Europeo invierte en tu futuro"; M Quijada-Álamo is supported by an "Ayuda Predoctoral de la Junta de Castilla y León" (JCYL-EDU/529/2017), JM

Hernández Sánchez and AE Rodríguez Vicente are supported by research grants of FEHH ("Fundación Española de Hematología y Hemoterapia").

Author Contribution

CPC, AERV and JMHR designed research. CPC performed the research and statistical analyses, analyzed the data, and drafted the manuscript. MHS analyzed the data and critically reviewed the manuscript. TG selected the samples and individually reviewed all IGH*R*-CLL cases by FISH and immunophenotypic studies. MQA contributed to the analysis and interpretation of the results and critically reviewed the manuscript. MMI and JMHS proceeded the samples and performed the next-generation sequencing studies. MJV, AGC, CA, MVP, SA, MS, ARM, JD, JRDV, JAQ, and JAHR provided the patients' data. RB and MHS designed the custom-CLL panel and the sequencing studies. AERV and JMHR performed research and critically reviewed and approved the final version of the manuscript.

References

1. Shanafelt TD, Geyer SM, Kay NE. Prognosis at diagnosis: integrating molecular biologic insights into clinical practice for patients with CLL. *Blood.* 2004;103(4):1202-1210.
2. Fabbri G, Dalla-Favera R. The molecular pathogenesis of chronic lymphocytic leukaemia. *Nat Rev Cancer.* 2016;16(3):145-162.
3. Rodríguez-Vicente AE, Bikos V, Hernández-Sánchez M, Malcikova J, Hernández-Rivas JM, Pospisilova S. Next-generation sequencing in chronic lymphocytic leukemia: recent findings and new horizons. *Oncotarget.* 2017;8(41):71234-71248.
4. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell.* 2013;152(4):714-726.
5. Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* 2000;343(26):1910-1916.
6. Guièze R, Wu CJ. Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood.* 2015;126(4):445-453.
7. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature.* 2015;526(7574):525-530.
8. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature.* 2015;526(7574):519-524.
9. Rossi D, Rasi S, Spina V, et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood.* 2013;121(8):1403-1412.
10. Jeromin S, Weissmann S, Haferlach C, et al. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia.* 2014;28(1):108-117.
11. Baliakas P, Hadzidimitriou A, Sutton LA, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia.* 2015;29(2):329-336.
12. Weissmann S, Roller A, Jeromin S, et al. Prognostic impact and landscape of NOTCH1 mutations in chronic lymphocytic leukemia (CLL): a study on 852 patients. *Leukemia.* 2013;27(12):2393-2396.
13. Nowakowski GS, Dewald GW, Hoyer JD, et al. Interphase fluorescence in situ hybridization with an IGH probe is important in the evaluation of patients with a clinical diagnosis of chronic lymphocytic leukaemia. *Br J Haematol.* 2005;130(1):36-42.
14. Nelson BP, Gupta R, Dewald GW, Paternoster SF, Rosen ST, Peterson LC. Chronic lymphocytic leukemia FISH panel: impact on diagnosis. *Am J Clin Pathol.* 2007;128(2):323-332.
15. Berkova A, Pavlistova L, Babicka L, et al. Combined molecular biological and molecular cytogenetic analysis of genomic changes in 146 patients with B-cell chronic lymphocytic leukemia. *Neoplasma.* 2008;55(5):400-408.
16. Lu G, Kong Y, Yue C. Genetic and immunophenotypic profile of IGH@ rearrangement detected by fluorescence in situ hybridization in 149 cases of B-cell chronic lymphocytic leukemia. *Cancer Genet Cytogenet.* 2010;196(1):56-63.
17. Willis TG, Dyer MJ. The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. *Blood.* 2000;96(3):808-822.

18. DE Braekeleer M, Tous C, Guéganic N, et al. Immunoglobulin gene translocations in chronic lymphocytic leukemia: A report of 35 patients and review of the literature. *Mol Clin Oncol.* 2016;4(5):682-694.
19. Mayr C, Speicher MR, Kofler DM, et al. Chromosomal translocations are associated with poor prognosis in chronic lymphocytic leukemia. *Blood.* 2006;107(2):742-751.
20. Van Den Neste E, Robin V, Francart J, et al. Chromosomal translocations independently predict treatment failure, treatment-free survival and overall survival in B-cell chronic lymphocytic leukemia patients treated with cladribine. *Leukemia.* 2007;21(8):1715-1722.
21. Jimenez-Zepeda VH, Chng WJ, Schop RF, et al. Recurrent chromosome abnormalities define nonoverlapping unique subgroups of tumors in patients with chronic lymphocytic leukemia and known karyotypic abnormalities. *Clin Lymphoma Myeloma Leuk*. 2013;13(4):467-476.
22. Gerrie AS, Bruyere H, Chan MJ, et al. Immunoglobulin heavy chain (IGH@) translocations negatively impact treatment-free survival for chronic lymphocytic leukemia patients who have an isolated deletion 13q abnormality. *Cancer Genet.* 2012;205(10):523-527.
23. Cavazzini F, Hernandez JA, Gozzetti A, et al. Chromosome 14q32 translocations involving the immunoglobulin heavy chain locus in chronic lymphocytic leukaemia identify a disease subset with poor prognosis. *Br J Haematol.* 2008;142(4):529-537.
24. Davids MS, Vartanov A, Werner L, Neuberg D, Dal Cin P, Brown JR. Controversial fluorescence in situ hybridization cytogenetic abnormalities in chronic lymphocytic leukaemia: new insights from a large cohort. *Br J Haematol.* 2015;170(5):694-703.
25. Fang H, Reichard KK, Rabe KG, et al. IGH translocations in chronic lymphocytic leukemia (CLL): clinicopathologic features and clinical outcomes. *Am J Hematol.* 2018.
26. Reddy KS. Chronic lymphocytic leukaemia profiled for prognosis using a fluorescence in situ hybridisation panel. *Br J Haematol.* 2006;132(6):705-722.
27. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood.* 2016;127(20):2375-2390.
28. Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* 2008;111(12):5446-5456.
29. González MB, Hernández JM, García JL, et al. The value of fluorescence in situ hybridization for the detection of 11q in multiple myeloma. *Haematologica.* 2004;89(10):1213-1218.
30. Quijada-Álamo M, Hernández-Sánchez M, Robledo C, et al. Next-generation sequencing and FISH studies reveal the appearance of gene mutations and chromosomal abnormalities in hematopoietic progenitors in chronic lymphocytic leukemia. *J Hematol Oncol.* 2017;10(1):83.
31. Hernández-Sánchez M, Rodríguez-Vicente AE, González-Gascón Y Marín I, et al. DNA damage response-related alterations define the genetic background of patients with chronic lymphocytic leukemia and chromosomal gains. *Exp Hematol.* 2019.
32. Quijada-Álamo M, Hernández-Sánchez M, Alonso-Pérez V, et al. CRISPR/Cas9-generated models uncover therapeutic vulnerabilities of del(11q) CLL cells to dual BCR and PARP inhibition. *Leukemia.* 2020.
33. Kasar S, Kim J, Improgo R, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun.* 2015;6:8866.

34. Mosquera Orgueira A, Rodríguez Antelo B, Díaz Arias J, et al. Novel Mutation Hotspots within Non-Coding Regulatory Regions of the Chronic Lymphocytic Leukemia Genome. *Sci Rep.* 2020;10(1):2407.
35. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* 2011;476(7360):298-303.
36. Pasqualucci L, Khiabanian H, Fangazio M, et al. Genetics of follicular lymphoma transformation. *Cell Rep.* 2014;6(1):130-140.
37. de Miranda NF, Georgiou K, Chen L, et al. Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. *Blood.* 2014;124(16):2544-2553.
38. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. In. Vol 1. Journal of the Royal Statistical Society1995:289-300.
39. Li H, Kaminski MS, Li Y, et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood.* 2014;123(10):1487-1498.
40. Pillonel V, Juskevicius D, Ng CKY, et al. High-throughput sequencing of nodal marginal zone lymphomas identifies recurrent BRAF mutations. *Leukemia.* 2018;32(11):2412-2426.
41. Correia C, Schneider PA, Dai H, et al. BCL2 mutations are associated with increased risk of transformation and shortened survival in follicular lymphoma. *Blood.* 2015;125(4):658-667.
42. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A.* 2013;110(4):1398-1403.
43. Okosun J, Bödör C, Wang J, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet.* 2014;46(2):176-181.
44. Batmanov K, Wang W, Bjørås M, Delabie J, Wang J. Integrative whole-genome sequence analysis reveals roles of regulatory mutations in BCL6 and BCL2 in follicular lymphoma. *Sci Rep.* 2017;7(1):7040.
45. Rheinbay E, Nielsen MM, Abascal F, et al. Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature.* 2020;578(7793):102-111.
46. Fabbri G, Khiabanian H, Holmes AB, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J Exp Med.* 2013;210(11):2273-2288.
47. Rossi D, Gaidano G. Richter syndrome. *Adv Exp Med Biol.* 2013;792:173-191.
48. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood.* 2016;127(17):2122-2130.
49. Herling CD, Abedpour N, Weiss J, et al. Clonal dynamics towards the development of venetoclax resistance in chronic lymphocytic leukemia. *Nat Commun.* 2018;9(1):727.
50. Giménez N, Martínez-Trillos A, Montraveta A, et al. Mutations in the RAS-BRAF-MAPK-ERK pathway define a specific subgroup of patients with adverse clinical features and provide new therapeutic options in chronic lymphocytic leukemia. *Haematologica.* 2019;104(3):576-586.

Figure legends

Figure 1. Mutational profile of CLL patients with *IGH* rearrangements. A) Mutational frequencies and associations in the CLL cohort according to the presence of *IGH* rearrangements. Significant p/q-values are annotated with asterisks (N=233). B) Each column represents a patient; each row corresponds to a genomic alteration. Patients are clustered according to the IGH*R* (*IGH/BCL2* translocation is indicated in light blue; other *IGH* translocations are shown in dark blue). Missense, frameshift, nonsense, splicing and UTR mutations are reported in red, green, yellow, pink and brown, respectively. The presence of a cytogenetic alteration is shown in gray and the *IGHV* unmutated status is represented in purple (N=46).

Figure 2. Schematic representation of *BCL2, IGLL5* and *HIST1H1E* mutations. Positions of coding mutations are indicated according to the aminoacid change at the protein level; positions of UTR mutations are indicated according to the nucleotide change in the DNA sequence (GRCh37/hg19 genome); with respect to the UTR regions, only *BCL2* and *IGLL5* 5'UTR regions were covered in the sequencing analysis (see Supplementary Table S2). Number of cases are denoted by circles in each mutation line and the color of the circles indicates the mutation subtype (missense, frameshift and nonsense). Mutations identified in the COSMIC database in non-Hodgkin lymphomas (NHL) are represented with red lines; mutations reported in the COSMIC database in CLL are indicated with blue lines; all other mutations are shown in black. Abbreviation: Aa, aminoacid.

Figure 3. Clinical impact of IGH*R* and genetic mutations in CLL patients. A) Kaplan-Meier analysis of TFT according to the presence of *IGH* translocation in CLLs with normal FISH (N=268). Kaplan-Meier analysis of TFT in IGHR-CLL patients with B) any mutation, C) *NOTCH1, SF3B1, TP53, BIRC3* or *BRAF* mutations and D) *IGLL5/BCL2* mutations (N=46). E) Kaplan-meier analysis of TFT in the three risk stratifications subgroups according to the presence of mutations in *NOTCH1, SF3B1, BIRC3, TP53* and *BRAF* genes. In low-risk patients, the presence of mutations in some of

these five genes is significantly associated with shorter TFT (median not reached *vs.* 24 months, $P<0.0001$) as well as in the intermediate-risk subgroup (56 *vs.* 2 months, $P<0.0001$).

Table 1. *IGLL5*, *BCL2* and *HISTH1E* mutations identified in IGHR-CLLs (N=46).

ID Patient	*IGH/BCL2* translocation (% of cells)	Gene	DNA[a]/cDNA change: AA change	VAF, %	Function	ID COSMIC[b]	Previously described	SIFT/ PolyPhen-2 pathogenicity prediction	Reported as somatic in VarSome	Functional impact[b] ICGC
10	NO	*BCL2*	C60985900G	12.74	5' UTR	-	-	-	YES	-
11	YES (85)	*BCL2*	C60985900A	33.87	5' UTR	-	-	-	YES	-
41	YES (85)	*BCL2*	c.G405T:p.E135D	30.15	exonic	-	-	T/P	YES	UNKNOWN
42	YES (60)	*BCL2*	c.G140A:p.G47D	11.39	exonic	COSM220809	DLBCL[35]	T/P	YES	YES
35	YES (87)	*BCL2*	c.G589A:p.G197S	40.1*	exonic	COSM5947452	DLBCL/FL[36]	-/B	YES	YES
45	YES (87)	*BCL2*	c.C175T:p.P59S	37.82*	exonic	COSM4170930	DLBCL/FL[36]	T/B	YES	YES
28	NO	*IGLL5*	c.G26C:p.G9A	51.09	exonic	COSM5713869	DLBCL	T/B	YES	-
15	NO	*IGLL5*	c.G312T:p.K104N	9.42	exonic	-	-	D/P	YES	YES
32	NO	*IGLL5*	c.G72A:p.W24X	42.86	exonic	-	-	-	-	YES
16	NO	*IGLL5*	C23230223-	26.12	5' UTR	-	CLL[33]	-	-	-
16	NO	*IGLL5*	G23230229C	26.51	5' UTR	-	-	-	YES	-
41	YES (85)	*IGLL5*	c.G88A:p.G30S	33.93	exonic	-	-	D/P	YES	YES
35	YES (87)	*IGLL5*	c.T167G:p.V56G	19.88	exonic	-	CLL[8]	T/B	YES	YES
35	YES (87)	*IGLL5*	c.C182T:p.S61F	19.88	exonic	COSM3357314	CLL[8]/DLBCL	T/B	YES	YES
45	YES (87)	*IGLL5*	A23230172C	17.58	5' UTR	-	CLL[8]	-	-	-
43	YES (41)	*IGLL5*	c.G94A:p.A32T	43.48	exonic	COSM5949859	CLL	D/B	YES	YES
34	YES (77)	*HIST1H1E*	c.G515C:p.S172T	42.33	exonic	-	-	D/D	-	-
37	YES (98)	*HIST1H1E*	c.C500T:p.A167V	41.74	exonic	COSM1292261	FL[39]/CLL[7]	T/B	YES	YES

*Confirmed as somatic in the matched CD19- cell fraction.

[a]Positions of UTR mutations are indicated according to the nucleotide change in the DNA sequence (GRCh37/hg19 genome)(reference transcripts: see Suppl. Table S5).

[b]Haematopoietic and lymphoid tissue

Abbreviations: AA=aminoacid; VAF=variant allele frequency; DLBCL=diffuse large B cell lymphoma; FL=follicular lymphoma; CLL=chronic lymphocytic leukemia; T=Tolerable; B=Benign; D=Damaging; P=Pathogenic.

UNKNOWN: reported in ICGC database with unknown functional impact in the gene. "-" indicates the variant has not been previously reported in the databases or seminal papers.

Table 2. Clinical and biological characteristics of CLL patients depending on the presence of *IGH* rearrangements (N=862).

Characteristic	*IGH*-translocation (N=46)	no *IGH*-translocation (N=816)	*P*	*Q*
Median age at diagnosis, years (range)	69 (43-89)	66 (25-97)	0.112[b]	0.542
Gender Male, %	63	63.8	0.698[c]	0.814
Median time from diagnosis to FISH, months (range)	1 (0-117)	1 (0-253)	0.568[b]	0.808
Binet B or C, %	38.6	22.2	**0.014[c]**	**0.039**
Median WBC[a] count, $\cdot 10^9$/L (range)	17.6 (2.3-196)	17.8 (2.4-964)	0.721[b]	0.841
Median lymphocytes count, $\cdot 10^9$/L (range)	11.6 (0.6-186)	12.2 (0.8-960)	0.874[b]	0.874
Median platelet count, $\cdot 10^9$/L (range)	172 (55-295)	187 (2-587)	0.456[b]	0.808
Median hemoglobin level, g/dL (range)	14.1 (6.6-16.5)	14.2 (4.4-18.9)	0.577[b]	0.808
High β2-microglobulin level, %	67.4	36.3	**<0.0001[c]**	**0.0007**
High lactate dehydrogenase level, %	27.3	15.7	**0.027[c]**	**0.054**
Hepatomegaly, %	7.1	6.9	0.824[c]	0.852
Splenomegaly, %	15.9	16.8	0.852[c]	0.852
B symptoms, %	11.1	7.9	0.595[c]	0.757
Richter transformation	4.3	1.7	0.148[c]	0.259
IGHV-Unmutated, *%*	60.6	44.9	**0.025[c]**	**0.054**
13q deletion, %	26.1	43.1	**0.0003[c]**	**0.0014**
trisomy 12, %	34.8	14.5	**<0.0001[c]**	**0.0007**
11q deletion, %	4.3	10.9	0.426[c]	0.596
17p deletion, %	6.5	4.3	0.334[c]	0.520
Need for treatment, %	67.4	44.0	**0.002[c]**	**0.007**
Median follow-up, months (range)	57 (1-157)	133 (106-159)	0.155[b]	0.543

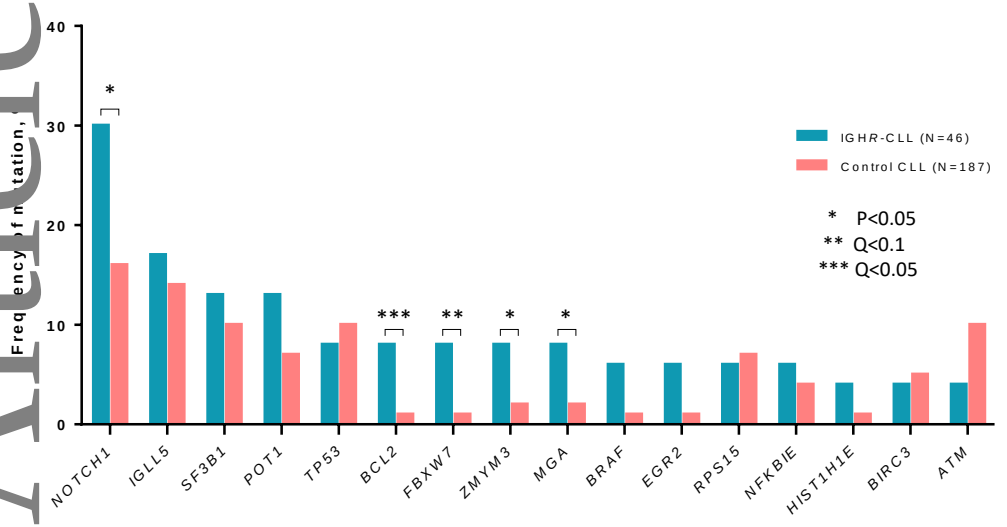
[a] WBC: white blood cells
[b] Mann Whitney U test
[c] Fisher test

Table 3. Univariate and multivariate analysis for Time to First Treatment (TFT) in IGH*R*-CLL patients (N=46).

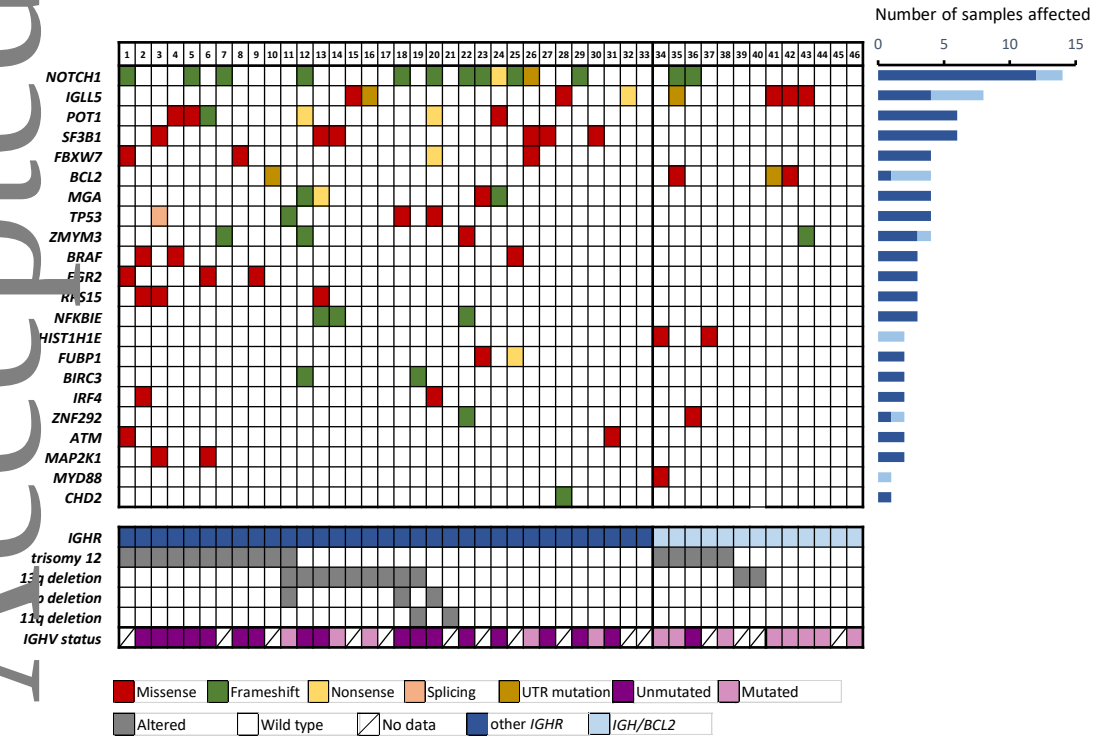
	Univariate				Multivariate			
		95% CI				95% CI		
	HR	**Lower**	**Upper**	***P***	**HR**	**Lower**	**Upper**	***P***
Male Gender	0.675	0.33	1.37	0.276				
Binet B/C	0.257	0.12	0.56	**0.001**	0.558	0.22	1.43	0.221
CD38 positivity	0.810	0.36	1.82	0.614				
IGHV-unmutated	0.325	0.13	0.79	**0.013**	0.566	0.20	1.64	0.29
LDH high	0.726	0.32	1.66	0.448				
β2M high	0.558	0.25	1.27	0.163				
Hepatomegaly	0.340	0.10	1.16	0.085				
Splenomegaly	0.368	0.15	0.88	**0.025**	0.403	0.14	1.19	0.099
B symptoms	0.483	0.18	1.27	0.141				
11q deletion	0.421	0.09	1.82	0.246				
IGH/BCL2 translocation absence	0.443	0.18	1.09	0.076				
IGLL5/BCL2 mutations absence	0.139	0.03	0.60	**0.008**	0.821	0.14	4.76	0.828
TP53 disruption/mutation	0.143	0.04	0.51	**0.003**	0.304	0.07	1.28	0.105
BRAF mutations	0.325	0.09	1.13	0.076				
NOTCH1/SF3B1/TP53/BIRC3/BRAF mutations	0.204	0.08	0.47	**0.0002**	0.255	0.07	0.88	**0.030**
Presence of mutation	0.238	0.05	10.0	0.051				

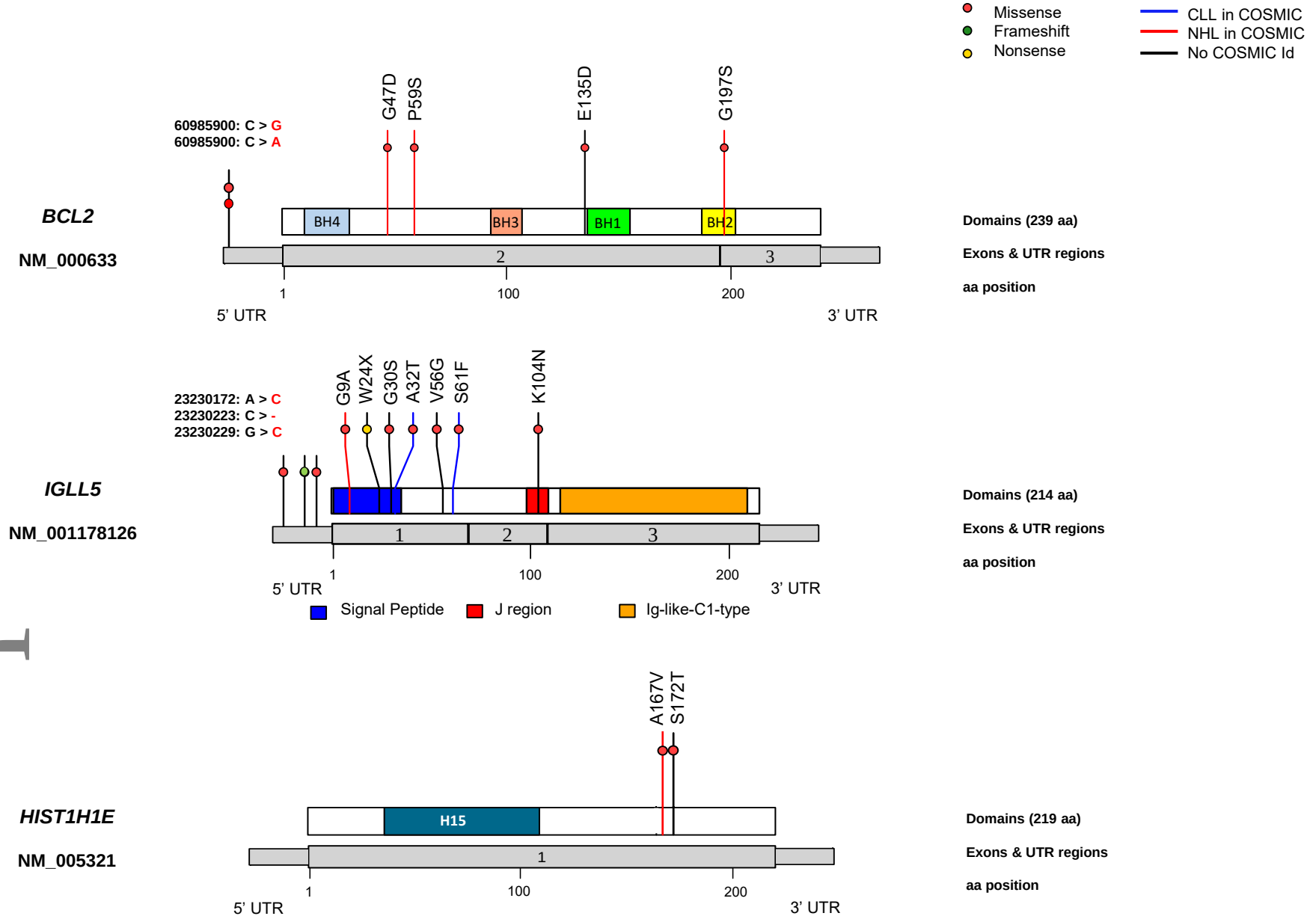
β2M, β2-microglobulin level; LDH, lactate dehydrogenase level.

A)

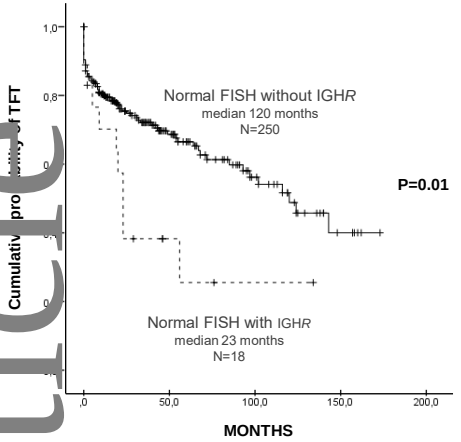


B)

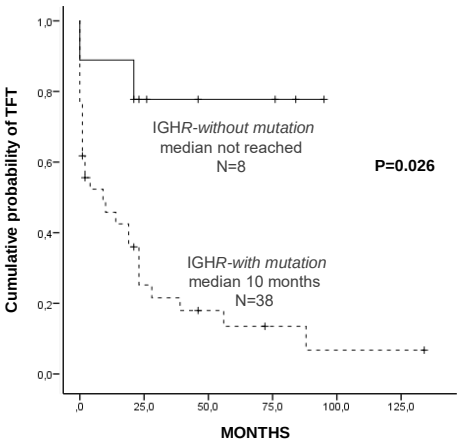




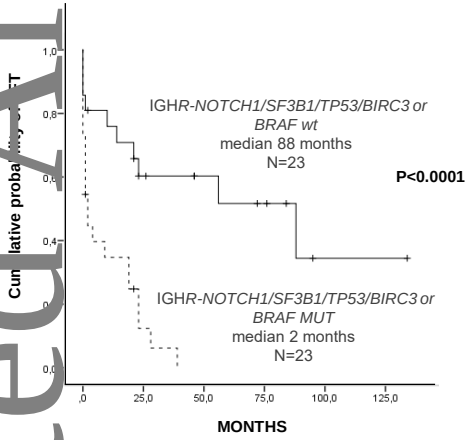
A)



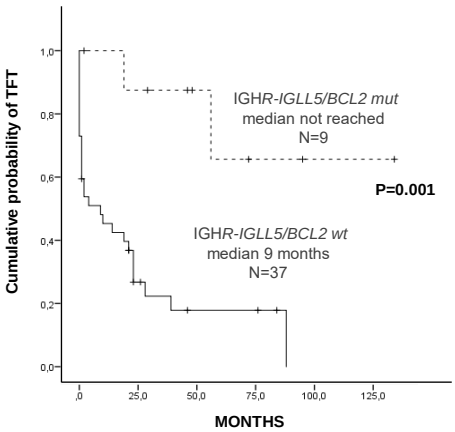
B)



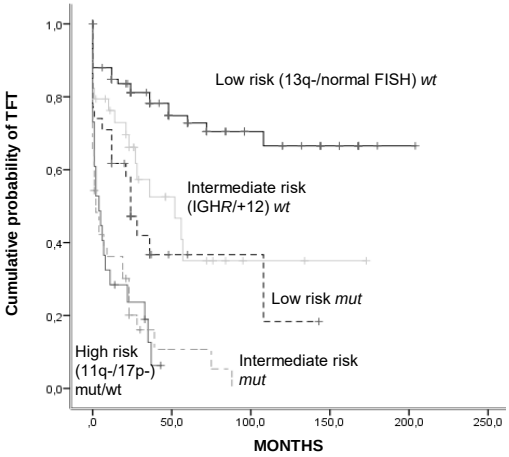
C)



D)



E)



	Total	Median	95 % CI
	97	Not reached	-
	35	56	22.7-89.3
	26	24	17.4-30.6
	37	24	17.4-30.6
	37	2	0-4.9

Low risk wt vs low risk mut	**p<0.0001**
Interm. risk wt vs interm. risk mut	**p<0.0001**
Low risk mut vs Interm. risk wt	**p=0.210**
Interm. Risk mut vs high risk mut/wt	**p=0.548**