의학박사 학위논문

Genomic Profile of Chronic Lymphocytic Leukemia in Korea Identified by Targeted Exome Sequencing

한국인 만성 림프구성 백혈병 환자의 표적 엑솜 시퀀싱을 이용한 유전체 분석

2016 년 2 월

서울대학교 대학원
임상의과학과
김 정 아

A thesis of the Degree of Doctor of Philosophy

한국인 만성 림프구성 백혈병 환자의 표적 엑솜 시퀀싱을 이용한 유전체 분석

Genomic Profile of Chronic Lymphocytic Leukemia in Korea Identified by Targeted Exome Sequencing

February 2016

Department of Clinical Medical Sciences

Seoul National University

College of Medicine

Jung-Ah Kim

ABSTRACT

Background: Chronic lymphocytic leukemia (CLL) is the most common form of leukemia in Western countries but is extremely rare in Asian countries. Therefore, genetic profiles of CLL have not been reported in Asians. Our study aimed to characterize the genomic profiles of Korean patients with CLL and to identify ethnic differences in somatic mutations with prognostic implications.

Methods: Seventy-one patients with CLL were enrolled (median age at diagnosis 61 years, range 23-81 years). We performed targeted exome sequencing for an 87-gene panel using next-generation sequencing (n=48) on an Illumina HiSeq 2500, and the sequencing reads were analyzed using a bioinformatics pipeline. Candidate mutations were confirmed by Sanger sequencing. We performed G-banding (n=60) and fluorescence *in situ* hybridization (FISH) for enumeration of chromosome 12 and for detection of 13q14.3, 17p13, and 11q22 deletions (n=51) in the 71 patients.

Results: G-banding and FISH identified aberrant karyotypes in 28.3% and 66.7% of patients, respectively. Targeted exome sequencing analysis revealed 71 substitutions and insertion/deletions. The average target region coverage was 231-fold. On average, 1.6 mutations per patient were observed among the 48 patients (range, 0-6), and 36 of the 48 patients (75%) carried at least one mutation; the mean number of mutations per patient with at least one mutation was 2.1 (range, 1-6). The most common recurrent mutation (>5% frequency) was *ATM* (20.8%), followed by TP53 (14.6%), *SF3B1* (10.4%), *KLHL6* (8.3%), *BCOR* (6.25%), *LAMB4* (6.25%), and *NOTCH1* (6.25%). The *TP53*

and *MYD88* mutations were associated with a poor prognosis (*P*=0.023, *P*=0.005; conventional Cox regression model, *P*<0.05). The *TP53* and *SF3B1* mutations showed an incidence similar to that observed in Caucasians, while the *ATM*, *NOTCH1*, *LRO1B*, *CHD2*, *POT1* and *TGM7* mutations were present at a rate higher than that observed in Caucasians. Notably, the *ATM* mutation exhibited a 2-fold higher incidence (20.8% vs. 9%) in Koreans compared with Caucasians. The *LAMB4*, *SH2B3*, *RUNX1*, *SCRIB*, *KIT*, *GATA2*, *CEBPA*, *TCF12*, *STAG2*, *ZRSR2*, *SF1*, *CSF1R*, *SETBP1*, *CSF1R*, and *SETBP1* mutations are novel and have not been reported in Caucasians.

Conclusions: The mutation profile of Korean patients with CLL differed from that of Caucasian patients with CLL, but their cytogenetic aberration profiles were similar. Novel mutations discovered in this study must be validated through large cohort studies and may offer clues to the mechanisms underlying the ethnic difference in CLL incidence. In the future, therapeutic strategies targeting these genes should be evaluated and considered.

Keywords: Chronic lymphocytic leukemia; Chromosomal abnormality; Mutation analysis; Ethnicity

Student number: 2012-30782

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CLL	Chronic lymphocytic leukemia
SLL	Small lymphocytic leukemia
MDS	Myelodysplastic syndrome
NGS	Next-generation sequencing
SNV	Single-nucleotide variant
BM	Bone marrow
LN	Lymph node
WHO	World Health Organization
FISH	Fluorescence *in situ* hybridization
Hb	Hemoglobin
PLT	Platelet
OS	Overall survival
AML	Acute myeloid leukemia

1. INTRODUCTION

Chronic lymphocytic leukemia/small lymphocytic leukemia (CLL/SLL) is a clonal B cell proliferative disorder. CLL is a typical malignancy with ethnic differences: it is one of the most common leukemias in Western countries but is extremely rare in Asian countries [1-9]. The incidence rate (per 100,000 persons per year) of CLL is 0.04 in Korea [10], 0.08~0.48 in Japan [7, 11], 0.6 in Koreans residing in America, 0.8 in Chinese residing in America [12], and 5.6 in Caucasians [9]. The incidence rate of CLL is 10- to 20-fold higher in Western than in Eastern populations. Additionally, ethnic differences in the mean age of disease development are observed [2]. This type of ethnic difference is also observed in myelodysplastic syndrome (MDS), which has a mean age of development of 71 years in the West and 57 years in Asia [13]. The causes of these ethnic differences may be related to the etiology of CLL. Chronic persistent viral infection, familial inheritance, and ethnicity (in the context of disease susceptibility) have been suggested as underlying causes, but the exact cause of most cases of CLL is not known. CLL exhibits marked genetic heterogeneity, with a relatively large number of genes showing recurrent mutations at a low frequency. Recently, genetic alterations were elucidated using next-generation sequencing (NGS), and mutations in the *SF3B1*, *TP53*, *NOTCH1*, *MYD88*, *ATM*, *XPO1*, and *KLHL6* genes were the most frequently involved recurrent mutations [14, 15]. The *NOTCH1* mutation is an independent poor prognostic factor [16], while the *ATM* mutation is associated with rapid disease progression [17]. The *TP53* mutation

is notoriously related to adverse survival outcomes and drug resistance [18, 19]. The SF3B1 mutation, which is well known in association with MDS, is also associated with rapid progression and adverse survival outcomes in CLL [20]. Differences in single-nucleotide variants (SNVs) have been reported between Caucasian and Asian populations [21, 22]. Specifically, 6p25.3 in the *IRF4* gene region, 2q37.1 in the *SP140* gene region and 2q13 in the *ACOXL* gene region have been associated with ethnic differences in incidence rates. Recently, several categories and classes of therapeutic drugs have been used to treat CLL; therefore, prediction of drug resistance prior to treatment is essential [23]. Because some mutations are associated with drug resistance, evaluating disease-related mutations is important.

Although there have been few reports on genetic changes in Asian patients with CLL, in one report characterizing Chinese patients with CLL [24], *TP53* and *NOTCH1* were demonstrated to be poor prognostic factors. However, several poor prognostic factors reported in Caucasian patients (*SF3B1*, *ATM* and *BIRC3*) did not show the same tendency in this population.

The modern Eastern diet and environment have become increasingly westernized. This change in life style and absolute increase in the elderly population will affect the incidence rate of CLL in Asian countries, especially in Korea, where the incidence of CLL has been recently increasing [25]. To investigate whether genetic mutations and the prognostic impact of known adverse mutations differ between Korean and Caucasian patients with CLL, we performed target-capture sequencing of 87 hematologic malignancy-related genes using a custom-designed capture panel for 48 CLL cases. To the

best of our knowledge, this study provides the first comprehensive mutation analysis of Asian patients with CLL using NGS.

2. MATERIALS AND METHODS

2.1. Study populations

A series of 58 patients who had undergone bone marrow (BM) examination and been diagnosed with CLL/SLL between September 2001 and October 2013 at Seoul National University Hospital (SNUH) were selected. Five patients who had been diagnosed with CLL/SLL prior to visiting SNUH were also included (n=5). Another group of 13 patients had received lymph node (LN) biopsies and had been diagnosed with CLL/SLL between April 1999 and October 2013 at Asan Medical Center (AMC) (n=13); three of these patients had been diagnosed with CLL/SLL prior to visiting AMC. All of the patients were Korean.

The diagnosis of CLL/SLL was based on the World Health Organization (WHO, 2008) classification criteria and the 2008 International Workshop on Chronic Lymphocytic Leukemia-National Cancer Institute criteria (IWCLL-NCI) [26]. Fluorescence *in situ* hybridization (FISH) for IgH/CCND1 translocations was performed to confirm that the disease was not a leukemic phase of mantle cell lymphoma. Clinical staging was performed using the Binet staging system (classes A, B and C). This staging system is based on the hemoglobin (Hb) count and platelet (PLT) count as well as the number of involved areas (i.e., head and neck, axillae, and groin) and organomegaly [26]. Laboratory data including age, sex, diagnosis and therapy start date, complete blood count, BM morphology, BM CLL cell count percentage sorted by flow

cytometry, cytogenetic analysis, and FISH were obtained for each patient. Mononuclear cells from the initial BM aspirates of patients who underwent BM aspiration and biopsy were fixed in Carnoy's solution and stored at −70℃ for cytogenetic analysis. All BM and LN samples were collected with informed consent, and the study was reviewed and approved by the Institutional Review Board of Seoul National University Hospital.

2.2. Bone marrow examination

Hematopathologists reviewed the Wright-stained BM smears and hematoxylin and eosin (H&E)-stained sections of BM trephine biopsies to determine the percentages and patterns of BM infiltration by lymphocytes. Immunohistochemical (IHC) staining was performed using antibodies targeting CD3, CD5, CD20 (all from Dako, Glostrup, Denmark) and ZAP-70 (Cell Marque, Rocklin, CA, USA). We calculated the tumor burden as the percentage of CD20+ cells. CD20-stained samples were available for 49 of the 58 patients who had undergone BM examination at SNUH. CD20 staining for the remaining 9 patients was not possible due to plastic embedding of the BM biopsy (n=8) or due to inadequate BM sample (n=1); the tumor burden in these patients was calculated as the lymphoid cell percentage in the aspiration samples (n=8). The median lymphoid cell percentage was 70% (range, 5–95%). ZAP-70 staining was performed for 62 patients. BM cellularity was reviewed in 58 samples, three of which were inadequate for cellularity counts. The median cellularity was 65% (range, 15–95%).

2.3. Leukemia-lymphoma marker study

The leukemia-lymphoma marker study using BM aspirates was performed using the Navios Cytometer (Beckman Coulter, Villepinte, France). The antibodies used were specific for TdT, CD2, CD3, CD5, CD7, CD10, CD19, CD20, CD22, CD23, FMC7, CD45, CytoCD3 (all from Beckman Coulter), CD56, Kappa, and Lambda (Becton, Dickinson and Company, BD Biosciences, San Jose, CA, USA). Navios software (Beckman Coulter) was used for data analysis. Among the 58 BM aspiration samples, 39 samples were available for the leukemia-lymphoma marker study. Thirty-eight samples were positive for CD19, and 33 samples were positive for CD5. Six samples were negative for CD5 and were categorized as indicative of atypical CLL. One sample did not express CD19 or CD20 but did express CD2, CD5 and CD7. We assumed that this phenomenon was a consequence of peripheral blood dilution. Because this patient's BM biopsy sample showed 50% CD20+ cells, we diagnosed this patient as having CLL/SLL.

2.4. Lymph node biopsy examination

A pathologist reviewed the formalin-fixed and paraffin-embedded LN sections. IHC staining was performed using antibodies specific for CD3, CD5, CD10, CD20, CD23, cyclin D1, BCL-6, and Ki-67 (all from DAKO, Glostrup, Denmark).

2.5. Conventional karyotyping by G-banding

Conventional cytogenetic data were available for 60 of the 71 patients. Cytogenetic studies using the standard G-banding technique on heparinized BM samples were performed as part of the diagnostic work-up. To stimulate B cells, tetradecanoylphorbol acetate (TPA; phorbol-12-myristate-13-acetate) was added, and the cells were cultured for four days. At least 20 cells in metaphases were analyzed whenever possible. Clonal abnormalities were defined as two or more cells with the same chromosomal gain or structural rearrangement or at least three cells with the same chromosome deletion. Karyotypes were recorded according to the International System for Human Cytogenetic Nomenclature (ISCN) 2013.

2.6. Fluorescence *in situ* hybridization

Common chromosomal abnormalities were investigated using commercial FISH probes. Data were available for 51 of the 71 patients. We used the following probes for enumeration of chromosome 12 and for detection of 13q14.3 deletion, 17p13 deletion, 11q22 deletion and IgH/CCND1 translocations (to exclude mantle cell lymphoma): the LSI D13S319/LSI13q34/CEP12 Multi-color Probe, LSI TP53 (17p13.1) SpectrumOrange Probe, Vysis IGH/CCND1 XT DF FISH Probe (all from Abbott Molecular/Vysis, Des Plaines, IL, USA), and XL ATM/TP53 Probe (Metasystems, GmbH, Altlussheim, Germany). Interphase FISH was performed on stored patient BM aspirate specimens. Slides were stained with

FISH probes and counter-stained with 4',6-diamidino-2-phenylindole (DAPI), and fluorescence signals were analyzed by fluorescent microscopy (Zeiss, Göttingen, Germany). The FISH results were recorded according to the ISCN 2013, and 200 cells in interphase were analyzed. The cut-off values for the deletion, amplification, or translocation of chromosomal regions were based on the mean of the normal controls (20 cytogenetically normal individuals) plus three standard deviations. FISH was successfully performed twice; the cut-off values for patients 1 to 11 were 1.06% for trisomy 12, 4.58% for 13q14.3 deletion, 7.39% for 17p13 deletion, and 5.59% for 11q22 deletion. The cut-off values for patients 12 to 71 were 1.5% for trisomy 12, 4.01% for 13q14.3 deletion, 1.7% for 17p13 deletion, and 5% for 11q22 deletion.

2.7. DNA extraction

DNA was extracted from various samples. Sample quality was evaluated using the Agilent 2200 TapeStation System (Santa Clara, CA, USA). Seventy-one samples were subjected to DNA extraction. DNA extraction was successful in 70 of the samples (28 samples of frozen BM mononuclear cells, 19 BM aspirate smears, and 23 formalin-fixed paraffin-embedded BM and LN biopsies), but only 48 of the 70 samples passed the DNA quality control criteria for multi-gene targeted sequencing. Twenty-two samples showed too low a quality for targeted sequencing (slides (n=4), FFPE (n=18)). Among the frozen BM mononuclear cells, DNA was extracted from 100% (28/28) of the samples, while the extraction percentages for unstained BM aspirate smear slides and FFPE samples were 79% (15/19) and 22% (5/23), respectively.

DNA was extracted from frozen BM mononuclear cells using a MagNA Pure LC DNA Isolation Kit (Roche Diagnostics GmbH, Mannheim, Germany) with the MagNA Pure LC 2.0 System (Roche) according to the manufacturer's instructions. DNA was extracted from the BM aspirate smear unstained slides using a QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) and Tissue Lysis Buffer (Qiagen). DNA was extracted from the FFPE BM biopsy samples using a WaxFree™ Kit (TrimGen Genetic Diagnostics; Sparks, MD, USA) according to the manufacturer's instructions. All samples were stored at −20°C.

In total, 37 patients were subjected to all of the conventional karyotyping, FISH and target-capture sequencing analyses. Ten patients were subjected only to conventional karyotyping, 1 patient only to FISH, and 5 patients only to target-capture sequencing (Figure 1).

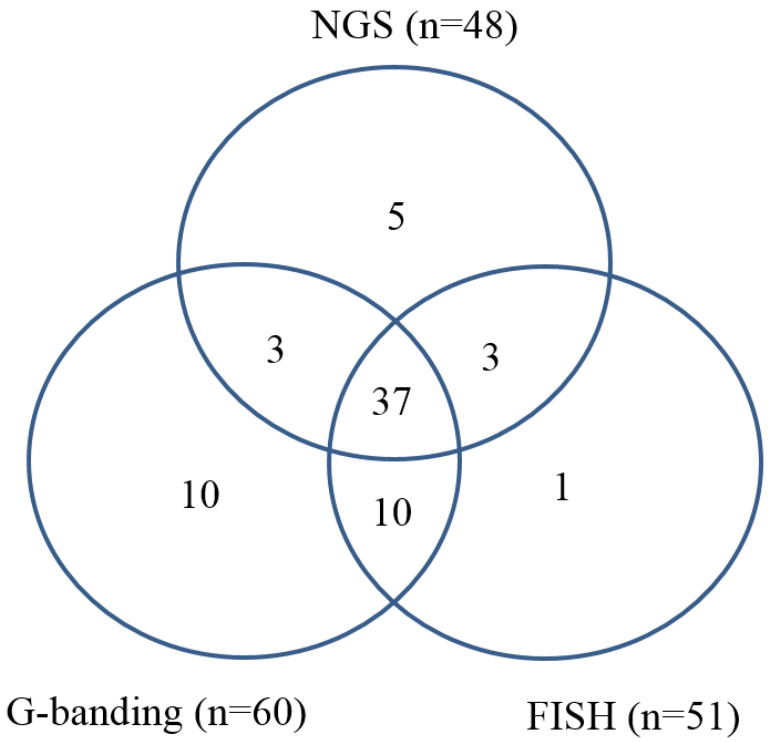


Figure 1. Venn diagram of examination distribution. Due to the available samples and sample quality, not all patients were subjected to all analyses. Among the total study population, only 37 patients were subjected to all three examinations.

NGS: next-generation sequencing; FISH: fluorescence *in situ* hybridization.

2.8. Targeted sequencing (next-generation sequencing, NGS)

To gain insight into the genetic lesions that drive CLL, we performed targeted sequencing. Our work represents the first such investigation in a Korean cohort. We manually prioritized 87 hematology malignancy-related genes that were recently implicated in lymphocytic leukemia (Table 1). Of the 48 samples, two samples were subjected to whole-genome amplification. gDNA shearing to generate the standard library and the hybridization step targeting only exonic regions were performed by Celemics Inc. (Seoul, Korea). The final quality was assessed using the Agilent 2200 TapeStation System (Santa Clara, CA, USA). We sequenced a total target length of 259-kb regions using the paired-end 150-bp rapid-run sequencing mode on an Illumina HiSeq 2500 platform. We achieved over 10x coverage for greater than 97% of targeted regions for each sample. The mean sequencing depth for the targeted regions (259 kb) was 231-fold (n=48). Because a matched control sample was not included in this study, we applied a stringent variant selection pipeline to prioritize the high-confidence set of somatic mutations (Figure 2).

Table 1. Gene panel for targeted sequencing.

Genes	NCBI Id.	Position	Pathway/Ontology
ASXL1	171023	20q11.1	Chromatin modification
ATM	472	11q22.3	DNA repair
ATRX	546	Xq21.1	Chromatin modification
BARD1	580	2q35	DNA repair
BCOR	54880	Xp11.14	Transcription
BIRC3	330	11q22.2	Receptor/Kinases
BRAF	673	7q34	RAS pathway
BRCC3	79184	Xq28	DNA repair
BRD2	6046	6p21.3	Transcription
BRD4	23476	19p13.1	Other
CARD6	84674	5p13.1	Other
CBL	867	11q23.3	RAS pathway
CCND1	595	11q13.3	Cell cycle
CDKN2A	1029	9p21	Cell cycle
CEBPA	1050	19q13.1	Transcription
CHD2	1106	15q26.1	Other
CSF1R	1436	5q32	Receptor/Kinases
CSF3R	1441	1p34.3	Receptor/Kinases
DAP3	7818	1q22	Other
DDX3X	1654	Xp11.4	Other
DIS3	22894	13q22.1	Other
DNMT3A	1788	2p23	DNA methylation
EEF1E1	9521	6p24.3	Other
EGR2	1959	10q21.3	Transcription
ETV6	2120	12p13.2	Transcription
EZH2	2146	7q35-36	Chromatin modification
FAM46C	54855	1p12	Other
FAT4	79633	4q28.1	Other
FBXW7	55294	4q31.3	Receptor/Kinases
FLT3	2322	13q12	Receptor/Kinases
GATA1	2623	Xp11.23	Transcription
GATA2	2624	3q21.3	Transcription
HIST1H1E	3008	6p22.2	Other
IDH1	3417	2q33.3	DNA methylation
IDH2	3418	15q26.1	DNA methylation
IKZF1	10320	7p13	Transcription
ITPKB	3707	1q42.12	Signaling
JAK2	3717	9p24	Receptor/Kinases
KIAA0355	9710	19q13.11	Other
KIT	3815	4q12	Receptor/Kinases
KLHL6	89857	3q27.1	Other
KRAS	3845	12p12.1	RAS pathway
LAMB4	22798	7q31.1	Other
LRP1B	53353	2q21.2	Other
MAPK1	5594	22q11.22	Signal/Kinase

MED12	9968	Xq13.1	Other
MPL	4352	1p34.2	Receptor/Kinases
MYD88	4615	3p22.2	Signaling
NF1	4763	17q11.2	RAS pathway
NFKBIE	4794	6p21.1	Other
NOTCH1	4851	9q34.3	Receptor/Kinases
NPM1	4869	5q35	Transcription
NRAS	4893	1p13.2	RAS pathway
PHF6	84295	Xq26.2	Transcription
PLEKHG5	57449	1p36.31	Other
POLG	5428	15q25	Other
POT1	25913	7q31.33	Other
PRKD3	23683	2p22.2	Signaling
PRPF40B	25766	12q13.12	Splicing
PTEN	5728	10q23.3	Other
PTPN11	5781	12q24.1	RAS pathway
RAD21	5885	8q24.11	Cohesin
RB1	5925	13q14	Cell cycle
RIPK1	8737	6p25.2	Other
RUNX1	861	21q22.3	Transcription
SAMHD1	25939	20q11.23	Other
SCRIB	23513	8q24.3	Other
SETBP1	26040	18q12.3	Other
SF1	7536	11q13.1	Splicing
SF3A1	10291	22q12.2	Splicing
SF3B1	23451	2q33.1	Splicing
SH2B3	10019	12q24.12	Signaling
SMARCA2	6595	9p24.3	Other
SMC1A	8243	Xp11.22	Cohesin
SMC3	9126	10q25.2	Cohesin
SRSF2	6427	17q25.1	Splicing
STAG2	10735	Xq25	Cohesin
TCF12	6938	15q21.3	Transcription
TET2	54790	4q24	DNA methylation
TGM7	116179	15q15.2	Other
TP53	7157	17p13.1	Transcription
U2AF1	7307	21q22.3	Splicing
U2AF2	11338	19q13.42	Splicing
WT1	7490	11p13	Transcription
XPO1	7514	2p15	Other
ZMYM3	9203	Xq13.1	Other
ZRSR2	8233	Xp22.1	Splicing

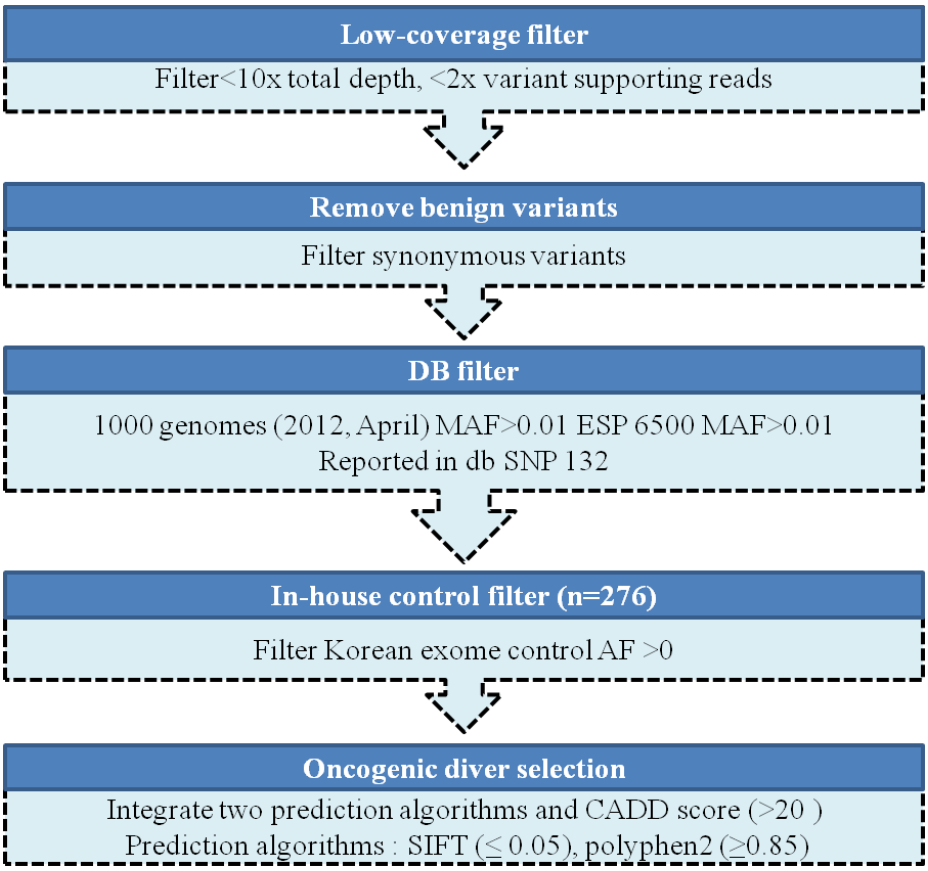


Figure 2. Summary of the variant-filtering pipeline.

MAF: Minor allele frequency; dbSNP: Single Nucleotide Polymorphism Database; AF: Allele frequency; CADD: Combined Annotation Dependent Depletion.

2.9. Statistical analysis

Fisher's exact test was used to test for significant associations among mutations and between mutations and cytogenetic abnormalities. Correlation coefficients of mutations among genes were calculated by Pearson's product-moment correlation. Survival was estimated using the Kaplan-Meier method, and differences between the survival curves and hazard ratios were analyzed using the Cox proportional hazards model. Overall survival (OS) was calculated from the date of diagnosis until the date of death. Statistical analyses were performed using R software (http://www.r-project.org). Because we had a limited number of samples, we did not apply multi-test corrections in our analyses. *P*-values<0.05 were considered statistically significant.

3. RESULTS

3.1. Patient characteristics

Among the 71 patients, 63.9% were men and 36.1% were women. Their median age was 61 years (range, 23-81). More than half of the patients were under 65 years old (63.9%), while 37.5% were older than 65 years. Most of the patients were male (63.9%), and the male:female ratio was 1.8:1. The median clinical characteristics were $15{,}040\times10^9$/L white blood cells (WBCs), $11{,}248\times10^9$/L lymphocytes, 12.8 g/L Hb, and 163×10^9/L PLTs. In total, 64.8% of the patients were Binet stage A, while 14.1% were stage B and 21.1% were stage C. The median OS for all of the patients was 112 months (Table 2). Eight percent (6/71) of the patients progressed to Richter's syndrome. Four patients exhibited transformation to diffuse large B cell lymphoma; 1 patient, to prolymphocytic leukemia; and 1 patient, to composite lymphoma consisting of peripheral T cell lymphoma (unspecified) and large B cell lymphoma.

Table 2. Baseline clinical features of patients with CLL.

Parameter	Patient numbers (ratios, ranges or percentages)
Demographic characteristics	71
Median age (years) at diagnosis	61 (23-81)
<65	44 (62.0%)
≥65	27 (38.0%)
Sex	
Male	45 (63.4%)
Female	26 (36.6%)
Clinical characteristics*	71
WBCs ($\times 10^9$/L)	15,040 (1,340-353,050)
Lymphocytes ($\times 10^9$/L)	11,248 (670-247,135)
Hb (g/L)	12.8 (7-17)
PLTs ($\times 10^9$/L)	163 (41-389)
Binet stage	
A	46 (64.8%)
B	10 (14.1%)
C	15 (21.1%)
Cytogenetics	60
Normal karyotype	39 (65.0%)
Aberrant karyotype	17 (28.3%)
Failed*	4 (6.6%)
Cases with follow-up	71
Median follow-up, in months	37
Median OS, in months	112
5-year OS	63.2%
10-year OS	49.4%

Abbreviations: WBCs, white blood cells; Hb, hemoglobin; PLTs, platelets; OS, overall survival.

* Four samples showed no mitosis.

3.2. Conventional karyotyping and fluorescence *in situ* hybridization

Conventional karyotyping was performed for 60 of the 71 patients; 65.0% (39/60) showed a normal karyotype, while 28.3% (17/60) show ed an aberrant karyotype, and 6.6% (4/60) showed no mitosis (Table 3). Ten patients showed three or more chromosomal aberrations representing complex karyotypes. FISH was performed for 51 of the 71 patients, 34 (66.7%, 34/51) of whom showed chromosomal aberrations. Of the 51 patients, trisomy 12 was investigated in 50, the 13q14 deletion was investigated in 48, 17p13 deletion was investigated in 51, and 11q22 deletion was investigated in 44. Trisomy 12 was detected in 15 cases (30.0%, 15/50), 13q14 was positive in 22 cases (45.8%, 22/48), 17p13 deletion was positive in 12 cases (23.5%, 12/51), and 11q22 deletion was positive in 8 cases (18.2%, 8/44).

Twenty-two cases showed one abnormality (43.1%, 22/51), while two abnormalities were observed in 13 cases (25.5%, 13/51), and three abnormalities were observed in 3 patients (5.9%, 5/51). In our study, the most frequent aberration was 13q14 deletion, while the second most frequent was trisomy 12 followed by 17p13 deletion and 11q22 deletion in order.

Samples from 47 patients were used for both conventional karyotyping and FISH, of which 8 (17.0%, 8/47) were negative by both methods and 23 showed aberrations that were detected only by FISH (48.9%, 23/47).

Table 3. Cytogenetic aberrations detected by G-banding and FISH studies in 60 Korean Patients with CLL

Patient No.	Age/ Sex	Binet Stage	Karyotype	FISH			
				+12 (%)	D13S 319 (%)	17p13 del (%)	11q22 del (%)
1	54/F	A	46,XX[20]	0	2.0	2.0	1.5
2	71/F	A	46,XX[20]	0	0.5	3	0.5
3	60/M	A	46,XY[19]	0	1.0	0.5	0.5
4	64/M	C	insufficient mitosis	0	51.0	0.5	ND
5	67/M	A	46,XY[18]	69.5	0.0	2	0
6	62/M	B	44,XY,,der(13;14)(q10;q10), -14[12]	68.5	24.5	1.5	0
7	60/M	A	46,XY[15]	0	27.0	2	0
8	79/M	A	46,XY[19]	0	78.5	1.5	0.5
9	67/F	A	no mitosis	0	4.5	4.5	1.0
10	38/M	A	ND	0.5	60.0	6	55.0
11	66/F	A	46,XX[18]	0	2.5	4.5	0
12	59/F	C	46,XX[20]	8	0	0	3.5
13	66/F	B	46,XX,add(2)(q37)[4],add(3)(q29)[5],del(4)(q10)[4][cp5]/46,XX[8]	0	0	0	1.0
14	43/M	A	46,XY[20]	0	14	0	1.0
15	70/F	A	46,XX[14]	3	6	0	0.5
16	62/M	A	46,XY[20]	0	5	40	0.5
17	63/M	A	No mitosis	62	0	0	1.0
19	76/M	C	ND	0	ND	0	ND
20	51/M	A	46,XY[10]	0	0	0	ND
21	46/M	B	46~47,XY,add(7)(p22),del(10)(p13),+mar[cp3]/46,XY[8]	0	0	0	91.0
22	61/M	A	46,XY[21]	0	0	0	0
23	61/M	A	46,XY[13]	0	36	0	0
24	53/F	A	46,XX[21]	ND	ND	ND	ND
26	53/M	A	46,XY[20]	4	50	4	1.0

27	57/F	A	46,XX[20]	0	0	0	2.5
28	63/F	A	46,XX[11]	ND	ND	ND	ND
29	67/F	C	ND	3.7	ND	0	ND
30	78/F	A	47,XX,+12[2]/46,XX[18]	40.5	0	0	1.0
31	66/M	A	46,XY[6]	0	87.0	0	89.0
32	36/M	C	46,XY,add(14)(q32),inc[1]/46,XY[19]	0	0	0	1.0
33	79/F	A	46,XX[10]	2	13	0	4.0
34	65/M	C	46,XY[20]	0	55	0	1.0
35	66/M	A	ND	ND	ND	91.5	ND
36	58/M	A	46,XY[21]	0	0	5	ND
37	56/M	A	47,XY,-2,+3,+add(3)(p?13),der(?3;15)(q10;q10),-5,-6,del(?10)(p10),del(13)(q12q14),add(?14)(p10),-17,+3mar[3]/46,XY[25]	0	33	7	0
38	55/F	A	46,XX[17]	0	57	77	0
39	55/M	A	46,XY,der(15)t(8;15)(q21.2;q26.1)[6]/46,XY[14]	0	6	0	11.5
41	62/F	C	48,XX,t(9;14)(p13;q32),dup(12)(q13q24.3),+15,+16[5]/47,sl,dup(1)(q?25q21),t(2;3)(p25;q21),-16[8]/47,sdl1,add(1)(p35)[3]/47,sdl1,+6,der(6;21)(p10;q10)[3]/48,sl,dup(1)(q21q32)[2]/46,XX[2]	0	0	0	1.5
42	51/M	B	46,XY[20]	2	0	4	1.5
44	71/F	C	46,XX,21pstk+[20]	0	0	0	ND
45	81/F	C	46,XX[20]	0	49	0	ND
46	55/F	A	46,XX[20]	0	35	0	3.5
47	70/M	A	46,XY[5]	7	9	5	0.5
49	46/F	A	47,XX,+12[1]/46,XX[28]	37.5	0	0	0
50	55/M	A	46,XY[20]	0	56	0	1.0
51	68/F	A	46~47,XX,add(1)(p32),add(2)(p10),add(7)(q?32),add(17)(p10),add(19)(q13.3),+mar,inc[2]/46,XX[18]	ND	ND	ND	ND
52	72/F	A	52~56,X,+add(X)(q22),del(X)(q26),+der(1;16)(q10;p10),+3,+3,del(4)(q31),+add(6)(q23),+del(6)(q15),add(6)(q2	ND	ND	ND	ND

			3)x2,+7, +del(7)(q22),+del(11)(q23),+13,del(15)(q15q25) ,+21,-22,+mar,inc[2]/46,XX[4]				
54	47/M	B	46,XY[20]	ND	ND	ND	ND
55	49/F	A	46,XX[20]	ND	ND	ND	ND
56	43/M	A	46,XY[20]	ND	ND	ND	ND
57	54/M	B	46,XY[30]	0	0	1.5	36.0
58	74/M	B	46,XY[20]	ND	ND	ND	ND
59	49/M	A	46,XY[20]	0	3	6	1.0
60	49/M	A	47,XY+12[2]/46,XY[18]	ND	ND	ND	ND
61	66/F	A	No mitosis	5	0	3	2.0
62	66/M	C	46,X,-Y,der(1)add(1)(p22)add(1)(q?42),add(2)(q?23),add(3)(p25),+4,-13,add(14)(p10),add(17)(p11.2),+mar[12]/46~48 ,idem, add(6)(q21),del(6)(q?23q25),i(8)(q10),+14,-add(14),+1~2mar,inc[cp8]	6	59	61	2.0
63	55/M	C	38~46,XY,del(1)(p34),del(2)(p21),add(3)(q21),add(6)(p23),del(11)((q?24),add(16)(q22),del(17)(p12),+mar[cp5]/46,XY[9]	0	0	29	66.0
64	46/F	C	46,XX[20]	0	26	0	12.0
65	63/F	A	46,XX[20]	1	3.5	0	5.0
66	46/M	A	46,XY[30]	ND	ND	ND	ND
67	23/F	B	46,XX,i(8)(q10),del(11)(q13q23),add(18)(p11.3)[cp8]/46,XX[12]	ND	ND	ND	ND
69	67/M	B	46,XY[20]	ND	ND	ND	ND
70	71/M	C	48,XY,+3,+12,t(14;19)(q32;q13.3)[9]/46,XY[6]	70	0	0	1.0
71	46/M	B	46,XY[20]	ND	ND	ND	ND

3.3 Gene mutations

A total of 6.6 million reads were obtained for each patient, and 98.9% of the reads mapped to the target region (Table 4). The 71 detected mutations represented 35 mutated genes, most of which were novel (84.5%, 60/71) (Table 5). An average of 1.6 mutations per patient was detected among the 48 patients (range 0-6). A total of 36 of 48 patients (75%) carried at least one mutation, and an average of 2.1 mutations were detected (range 1–6). For the final 71 variant sets, we validated somatic mutations through Sanger sequencing. Sixteen randomly selected mutations were successfully confirmed (Table 6).

Among the 76 mutations, 49 were missense mutations, while 19 were frameshift mutations, and 4 were non-frameshift mutations; the remaining 5 were stop, gain or loss mutations. In more than 5% of the patients, mutations were detected in *ATM*, *TP53*, *SF3B1*, *KLHL6*, *BCOR*, *LAMB4*, and *NOTCH1* (in 20.8%, 14.6%, 10.4%, 8.3%, 6.25%, 6.25%, and 6.25% of the patients, respectively) (Figure 3).

ATM (n=4) and *SF3B1* (n=3) were the most frequent mutations in patients in Binet stage C, and *TP53* (n=4) was the most frequent mutation in patients in Binet stage A. The mean number of mutations according to Binet staging was 1.4 (43 mutations in 31 patients) for stage A, 2.2 (11 mutations in 5 patients) for stage B, and 1.8 (22 mutations in 12 patients) for stage C. There was no correlation between Binet stage and mutation numbers (Figure 4).

Among the patients who exhibited transformation to Richter's syndrome, an average of 1.8 mutations was detected per sample (range, 1–3). Among these

patients, 10 different gene mutations were identified (*ZMYM3*, *CDKN2A*, *ATM*, *TP53*, *NOTCH1*, *SF3B1*, *SAMD1*, *MYD88*, *DDX3X* and *RUNX1*). The *TP53* mutation was detected in two patients, corresponding to a frequency of 33.3% (2/6), which was higher than in patients who did not exhibit transformation to Richter's syndrome. Patients who underwent transformation to Richter's syndrome showed shorter survival time compared with patients with CLL (Figure 5). Allele burden did not differ between patients with Richter's syndrome and patients without Richter's syndrome (P=0.574).

Fifteen genes had at least one insertion or deletion. There were 17 indels, none of which had been previously reported in CLL studies, with the exception of one mutation in *NOTCH1* (p.R2515Rfs*4). The other mutations known from the literature (excluding indels) were *TP53* (H82R, Y88C, G113D, R141C and D149N), EGR2 (E356K), *MYD88* (X160R), *ATM* (K2749I) and *SF3B1* (K666E and K700E). Among substitutions, transitions (68.5%) were more prevalent than transversions (31.5%). Next, we classified genes into nine biological functional categories. The most frequently mutated targets were involved in transcription (33.3% of cases) and signaling (33.3%), followed by DNA repair (20.8%), splicing (14.6%), cell cycle (14.6%), receptor/kinase (10.4%), other functions (6.25%), chromatin modification (6.25%) and cohesin (2.08%).

Table 4. Summary of target-based sequencing results

Variable	Finding
Total no. of reads	6,644,998
Reads mapped to the target region, %	98.9
Target coverage per base (mean)	231 X
Target base pairs covered ≥10 X, %	97.5
Target base pairs covered ≥100 X, %	85.4
Total % of variants discovered	
Missense	69.0 (49/71)
Frameshift	18.3 (13/71)
Stop, gain or loss	7.0 (5/71)
Non-frameshift	5.6 (4/71)

Table 5. Somatic mutations found in 48 patients with CLL

Chr	Gene	Pos	Ref	Alt	Mutation Type	AA Change	Patient No.
1	ITPKB	226923392	G	A	nonsynonymous SNV	p.R590W	19
2	PRKD3	37543553	G	A	stopgain SNV	p.R39X	29
2	LRP1B	141459813	TCTC	-	frameshift deletion	p.2066_2067del	40
2	SF3B1	198266834	T	C	nonsynonymous SNV	p.K700E	48, 64, 65
2	SF3B1	198267361	T	C	nonsynonymous SNV	p.K666E	16
2	SF3B1	198267484	G	C	nonsynonymous SNV	p.R625G	25
3	MYD88	38182641	T	C	stoploss SNV	p.X160R	34, 62
3	GATA2	128205685	C	T	nonsynonymous SNV	p.A64T	16
3	KLHL6	183212036	T	C	nonsynonymous SNV	p.Y394C	8
3	KLHL6	183273170	G	A	nonsynonymous SNV	p.A91V	65
3	KLHL6	183273185	T	A	nonsynonymous SNV	p.H86L	6

3	KLHL6	183273188	C	T	nonsynonymous SNV	p.C85Y	23
3	KLHL6	183273189	A	G	nonsynonymous SNV	p.C85R	65
4	KIT	55604628	C	T	stopgain SNV	p.R946X	15
4	FAT4	126239804	ACAAGAATGG	-	frameshift deletion	p.746_749del	64
4	FAT4	126241813	A	T	nonsynonymous SNV	p.N1416I	64
4	FAT4	126373317	C	A	nonsynonymous SNV	p.R3716S	15
5	CSF1R	149457767	C	T	nonsynonymous SNV	p.V213M	35
6	BRD2	32945698	GAG	-	nonframeshift deletion	p.498_499del	8, 35
7	LAMB4	107703233	T	TA	frameshift substitution	NA	57
7	LAMB4	107706935	C	T	nonsynonymous SNV	p.G853S	8
7	LAMB4	107732794	G	A	nonsynonymous SNV	p.A513V	35
7	POT1	124499011	C	A	nonsynonymous SNV	p.K103N	28
7	EZH2	148506437	G	A	nonsynonymous SNV	p.A636V	41

7	EZH2	148514402	CAGCACCACTCCAC TCCACATTCTCAG	-	nonframeshift deletion	p.388_397del	38
8	SCRIB	144889100	G	T	nonsynonymous SNV	p.D754E	42
9	CDKN2A	21974808	T	A	nonsynonymous SNV	p.S7C	60
9	NOTCH1	139390648	AG	-	frameshift deletion	p.P2515*fs	40, 63
9	NOTCH1	139390815	G	-	frameshift deletion	p.Q2459fs	43
9	NOTCH1	139402795	C	T	nonsynonymous SNV	p.G1072S	43
10	EGR2	64573332	C	T	nonsynonymous SNV	p.E356K	21
11	SF1	64534502	A	AGGC	nonframeshift substitution	NA	65
11	ATM	108115594	C	T	stopgain SNV	p.R248X	57
11	ATM	108163400	A	T	nonsynonymous SNV	p.L1497F	64
11	ATM	108183152	A	T	nonsynonymous SNV	p.E1978V	5
11	ATM	108186757	G	A	nonsynonymous SNV	p.E2039K	48

11	ATM	108196912	T	C	nonsynonymous SNV	p.L2312P	21
11	ATM	108201089	C	G	nonsynonymous SNV	p.R2486G	63
11	ATM	108205790	AG	-	frameshift deletion	p.2702_2702del	21
11	ATM	108206579	A	G	nonsynonymous SNV	p.D2720G	42
11	ATM	108206581	G	A	nonsynonymous SNV	p.D2721N	64
11	ATM	108206666	A	T	nonsynonymous SNV	p.K2749I	19
11	ATM	108216576	C	T	nonsynonymous SNV	p.P2842L	31
11	ATM	108216601	G	T	nonsynonymous SNV	p.L2850F	26
11	ATM	108224508	A	C	nonsynonymous SNV	p.Q2896P	48
11	ATM	108235838	G	T	nonsynonymous SNV	p.W2960C	64
12	SH2B3	111856250	G	A	nonsynonymous SNV	p.E101K	31
12	SH2B3	111885306	T	TG	frameshift substitution	NA	38
13	RB1	49039195	C	T	nonsynonymous SNV	p.S758L	34

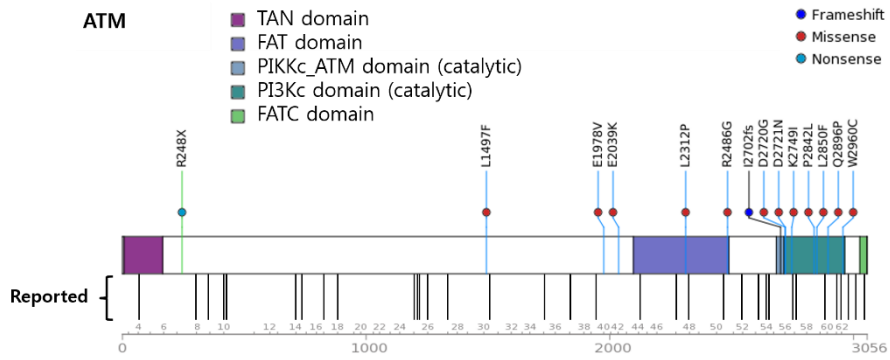
15	TCF12	57489989	A	G	nonsynonymous SNV	p.K182R	57
15	CHD2	93563337	C	T	nonsynonymous SNV	p.H1668Y	30
17	TP53	7576874	TCCAGTGGTTTCT	-	frameshift deletion	p.188_192del	20
17	TP53	7577097	C	T	nonsynonymous SNV	p.D149N	62
17	TP53	7577121	G	A	nonsynonymous SNV	p.R141C	16
17	TP53	7577511	A	G	nonsynonymous SNV	p.L125P	35
17	TP53	7577547	C	T	nonsynonymous SNV	p.G113D	33
17	TP53	7578190	T	C	nonsynonymous SNV	p.Y88C	33
17	TP53	7578208	T	C	nonsynonymous SNV	p.H82R	63
17	TP53	7578212	G	A	stopgain SNV	p.R81X	13
18	SETBP1	42643500	C	G	nonsynonymous SNV	p.P1543R	21
19	CEBPA	33792731	G	GGCGGGT	nonframeshift substitution	CEBPA	61
20	SAMHD1	35579956	T	-	frameshift deletion	p.D31fs	25

21	RUNX1	36259207	G	A	nonsynonymous SNV	p.P68L	51
X	ZRSR2	15818044	CT	-	frameshift deletion	p.57_58del	5
X	BCOR	39914637	A	-	frameshift deletion	p.M1523fs	43
X	BCOR	39923684	C	T	nonsynonymous SNV	p.R1118H	11
X	BCOR	39933386	CTGGGCACCTTCGC	-	frameshift deletion	p.400_405del	65
X	DDX3X	41203568	A	AAC	frameshift substitution	NA	36
X	MED12	70356335	C	G	nonsynonymous SNV	p.L1744V	39
X	ZMYM3	70468137	C	T	nonsynonymous SNV	p.R617H	60
X	STAG2	123200027	C	A	nonsynonymous SNV	p.A700D	5

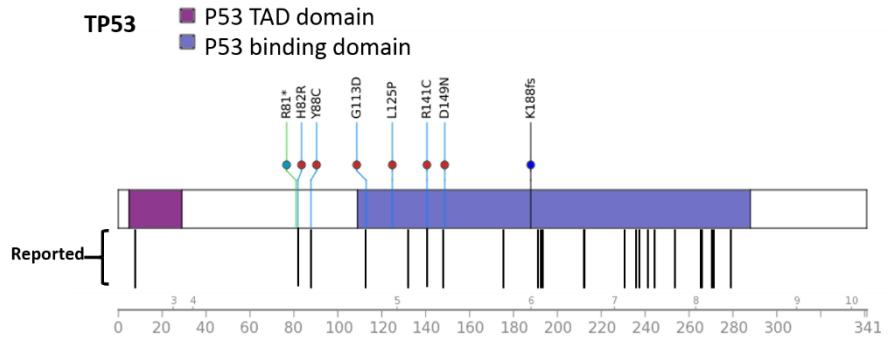
Table 6. Primer sequences used for Sanger sequencing of 16 selected loci

Chr	Gene	Position	Ref	Alt	Forward (5'→3')	Reverse (5'→3')
2	SF3B1	198267484	G	C	AAGAATAGCTATCTGTTGTACA	TGTTTATGGAATTGATTATGGAAA
7	LAMB4	107706935	C	T	AAAATTAAAGTGCATATGAATTCC	GCTCCAGCATACTTCTCTTT
7	EZH2	148514402	CAGCACCACTCCACTCCACATTCTCAG	-	CCTGTCTACATGTTTTGGTCCC	ACGATGGGTTAGTGTTTTGCC
11	ATM	108183152	A	T	TAAGTGATTTATTCTGTTTTGTTTG	CTGTACAGTGTCTATAACAAAATAA
11	ATM	108186757	G	A	GTGGAGGGAAGATGTTACAA	CCAACATACTGAAATAACCTCA
11	ATM	108201089	C	G	GATACACAGTAAAGGTTCAGC	TACAAAGAGGTATACACGATTC
11	ATM	108206579	A	G	AGGTATTTAATTATTTGGGAGACT	TTATATGTTTTTGGTGAACTAACA
11	ATM	108206581	G	A	GTATTTAATTATTTGGGAGACTGT	TTTTATATGTTTTTGGTGAACTAAC
11	ATM	108206666	A	T	ATAAACTGTACTTGTTTATTCATGC	GCCTCCCAAAGCATTATGAA
11	ATM	108216576	C	T	TATATTCTCTATTTAAAGGAGGTGC	ACTCAGAATGTAGAAAAAGTGC
11	ATM	108224508	A	C	AACTACTGTACATACTAGTGTTC	ATTTTGACATCAAAAATTATTTCCC
11	ATM	108235838	G	T	CCCCATCAACTACCATGTGA	ATCTGAAAAACTGACAACAGG
17	TP53	7577121	G	A	CTTCTTGTCCTGCTTGCTTA	TAGGCTCCAGAAAGGACAAG
17	TP53	7578190	T	C	ATTTACTTTGCACATCTCATG	GCCTCTGATTCCTCACTGAT
17	TP53	7578208	T	C	ATGGGGTTATAGGGAGGTCA	GCCTCTGATTCCTCACTGAT
17	TP53	7578212	G	A	GGTTATAGGGAGGTCAAATAAG	GCCTCTGATTCCTCACTGAT

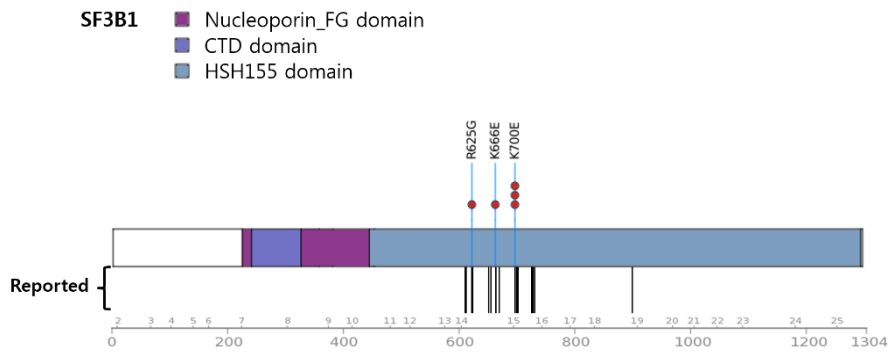
A


ATM
TAN domain
FAT domain
PIKKc_ATM domain (catalytic)
PI3Kc domain (catalytic)
FATC domain
Frameshift
Missense
Nonsense
R248X
L1497F
E1978V
E2039K
L2312P
R2486G
I2702fs
D2720G
D2721N
K2749I
P2842L
L2850F
Q2896P
W2960C
Reported
0
1000
2000
3056

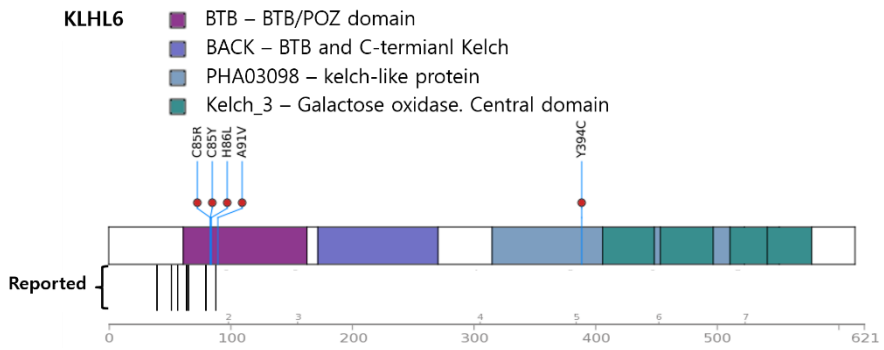
B


TP53
P53 TAD domain
P53 binding domain
R81*
H82R
Y88C
G113D
L125P
R141C
D149N
K188fs
Reported
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
341

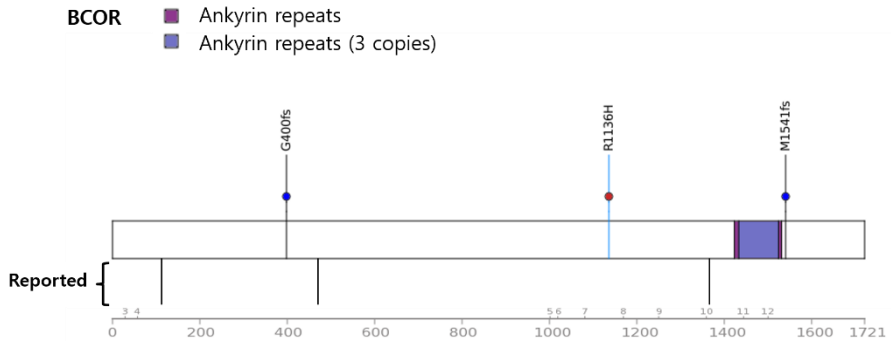
C


SF3B1
Nucleoporin_FG domain
CTD domain
HSH155 domain
R625G
K666E
K700E
Reported
0
200
400
600
800
1000
1200
1304

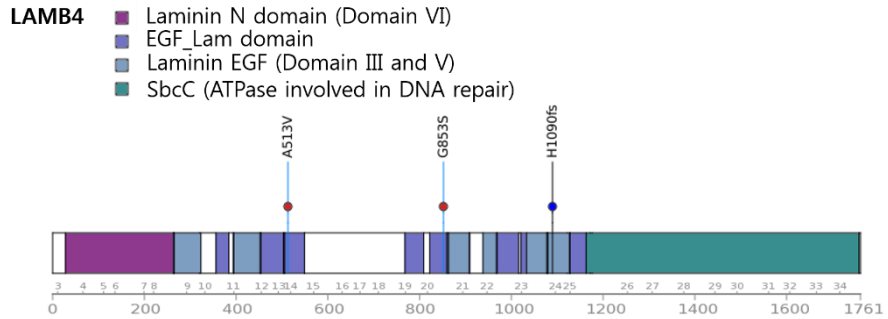
D


KLHL6
BTB – BTB/POZ domain
BACK – BTB and C-termianl Kelch
PHA03098 – kelch-like protein
Kelch_3 – Galactose oxidase. Central domain
C85R
C85Y
H86L
A91V
Y394C
Reported
0
100
200
300
400
500
621

E


BCOR
Ankyrin repeats
Ankyrin repeats (3 copies)
G400fs
R1136H
M1541fs
Reported
0
200
400
600
800
1000
1200
1400
1600
1721

F


LAMB4
Laminin N domain (Domain VI)
EGF_Lam domain
Laminin EGF (Domain III and V)
SbcC (ATPase involved in DNA repair)
A513V
G853S
H1090fs
0
200
400
600
800
1000
1200
1400
1600
1761

G

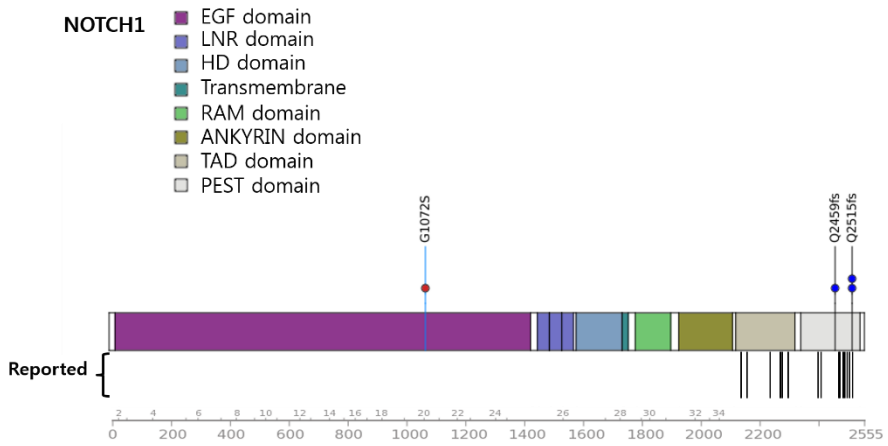
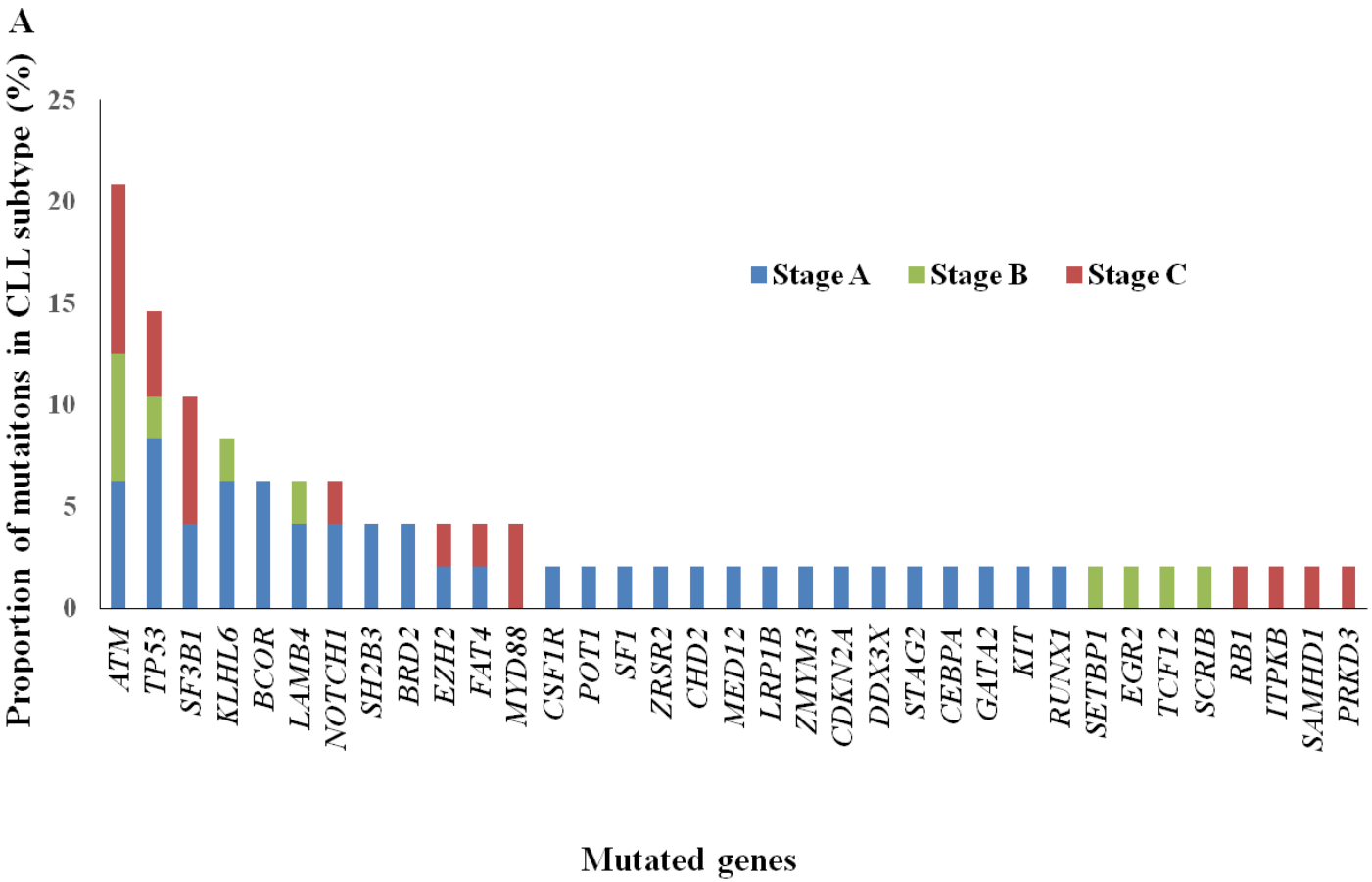


Figure 3. Diagrams for mutations observed in more than 5% of cases. Genes include (A) *ATM*, (B) *TP53*, (C) *SF3B1*, (D) *KLHL6*, (E) *BCOR*, (F) *LAMB4* and (G) *NOTCH1* (Transcript IDs: ATM, NM_000051; TP53, NM_001126114; SF3B1, NM_024582; KLHL6, NM_130446; BCOR, NM_001123383; LAMB4, NM_007356; NOTCH1, NM_017617).


A
Proportion of mutaitons in CLL subtype (%)
25
20
15
10
5
0
Stage A
Stage B
Stage C
ATM
TP53
SF3B1
KLHL6
BCOR
LAMB4
NOTCH1
SH2B3
BRD2
EZH2
FAT4
MYD88
CSF1R
POT1
SF1
ZRSR2
CHD2
MED12
LRP1B
ZMYM3
CDKN2A
DDX3X
STAG2
CEBPA
GATA2
KIT
RUNX1
SETBP1
EGR2
TCF12
SCRIB
RB1
ITPKB
SAMHD1
PRKD3
Mutated genes

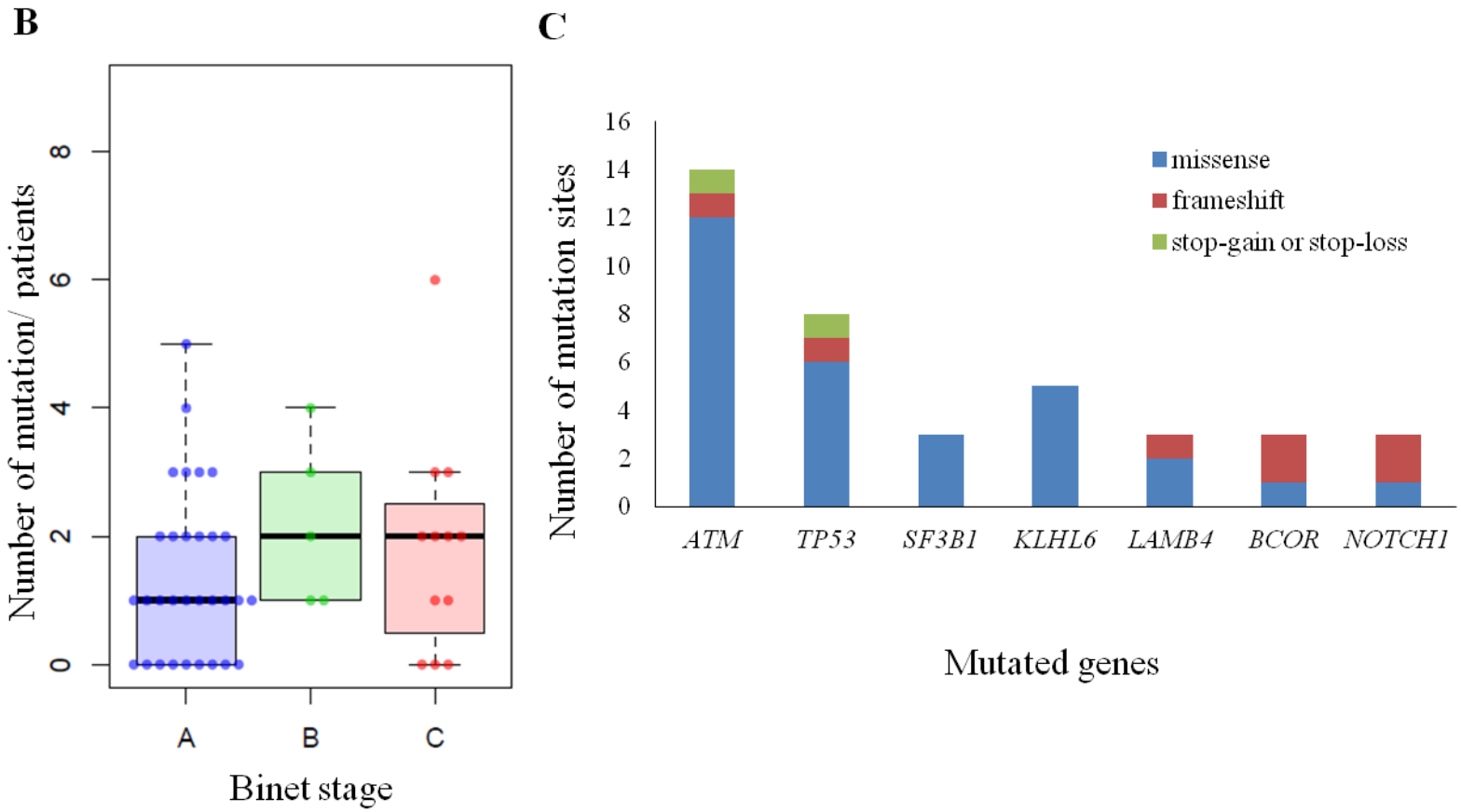
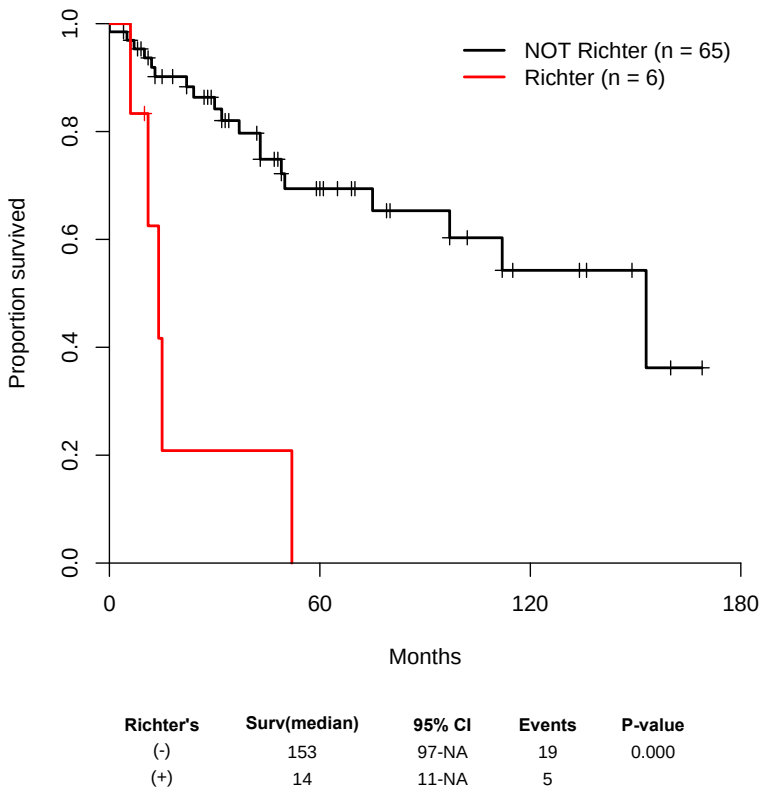

B
Number of mutation/ patients
0
2
4
6
8
A
B
C
Binet stage
C
Number of mutation sites
0
2
4
6
8
10
12
14
16
missense
frameshift
stop-gain or stop-loss
ATM
TP53
SF3B1
KLHL6
LAMB4
BCOR
NOTCH1
Mutated genes

Figure 4. Genomic landscape of the Korean cohort. (A) Frequency of gene mutations broken down by CLL stage. (B) Number of mutations according to the Binet stage. (C) Number of mutation sites occurring in more than 5% of the population.



Richter's	Surv(median)	95% CI	Events	P-value
(-)	153	97-NA	19	0.000
(+)	14	11-NA	5	

Figure 5. Kaplan-Meier survival curves of patients transformed to Richter's syndrome and non-transformed patients

3.4. Correlations between various gene mutations and between gene mutations and cytogenetics

We investigated the correlations between combinations of mutated genes to detect functional interactions among the mutations associated with CLL pathogenesis (Figure 6). Strong positive correlations were observed for the following genes: *EGR2* and *SETBP1* (r=1), *KIT* and *FAT4* (r=0.43), *RB1* and *MYD88* (r=0.70), *BRD2* and *LAMB4* (r=0.81), *BRD2* and *CSF1R* (r=0.70), *ZMYM3* and *CDKN2A* (r=1), and *STAG2* and *ZRS*R2 (r=1) (P<0.05). Additionally, 11q22 deletion and *ATM* (r=0.62) showed strong positive correlations.

We calculated hazard ratios for the genes using the conventional Cox regression model and found several that were significant: *ITPKB* (HR=22.82, P=0.011), *SF3B1* (HR=10.98, P=0.032), *EGR2* (HR=10.98, P=0.032), *MYD88* (HR=9.85, P=0.005), *SAMHD1* (HR=8.64, P=0.049) and *TP53* (HR=3.34, P=0.023) (Figure 7).

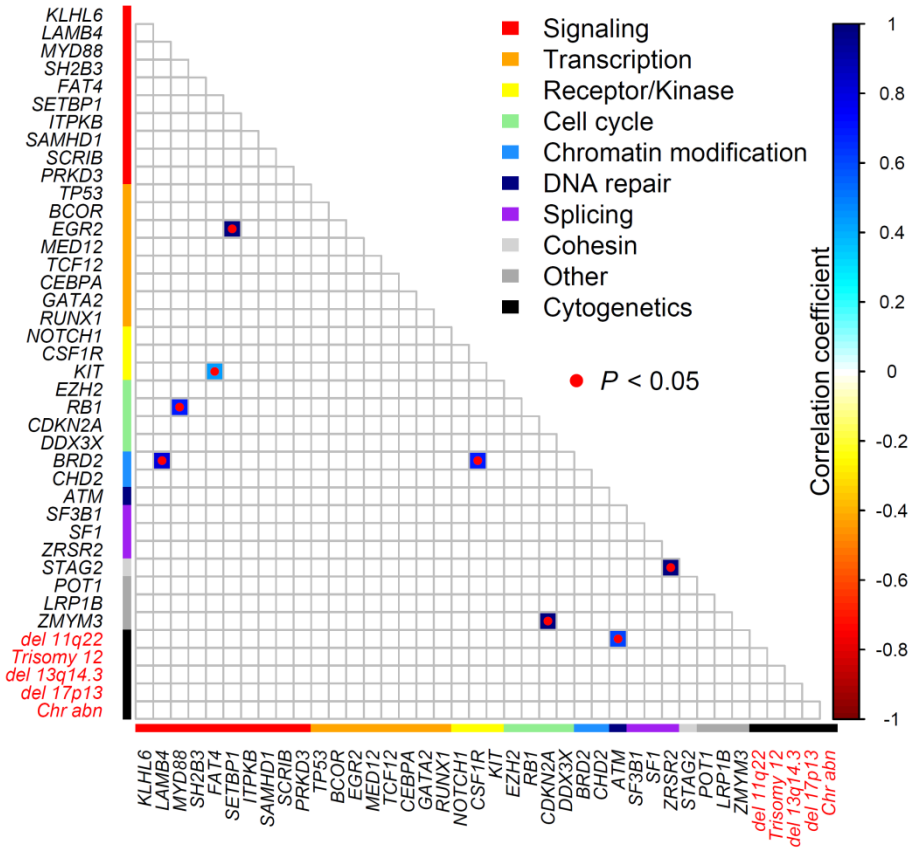


Figure 6. Correlations between various gene mutations and between gene mutations and cytogenetics. Correlation coefficients with *P*-values<0.05 are presented as colored boxes.

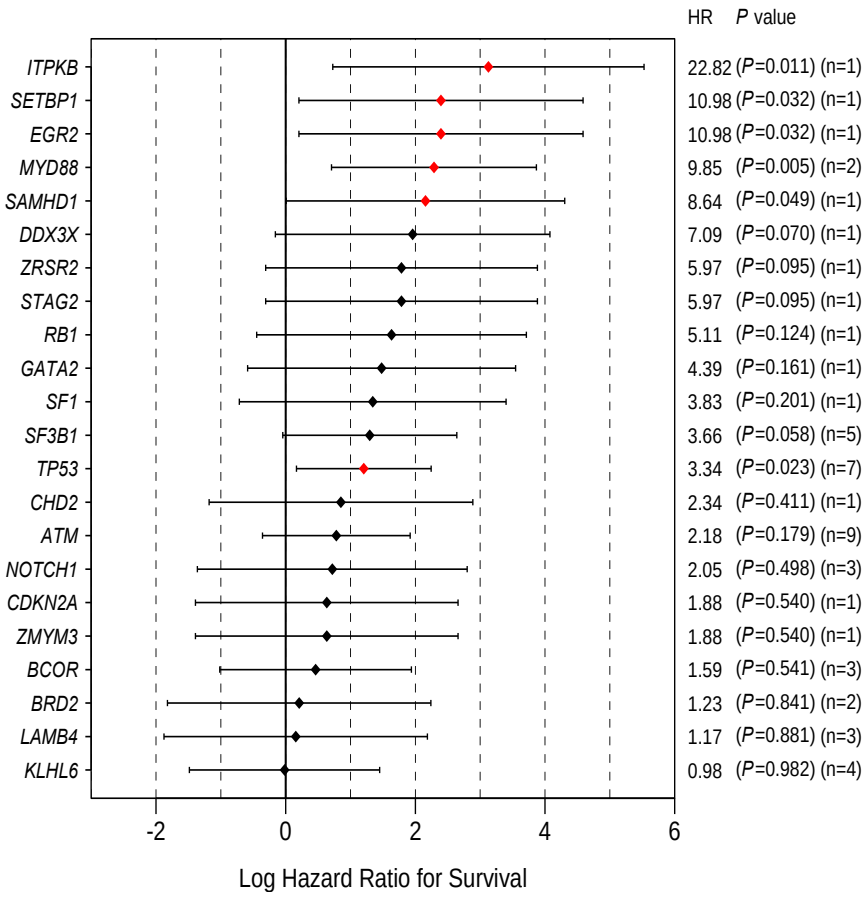


Figure 7. Hazard ratios with 95% confidence intervals for overall survival for each gene variant.

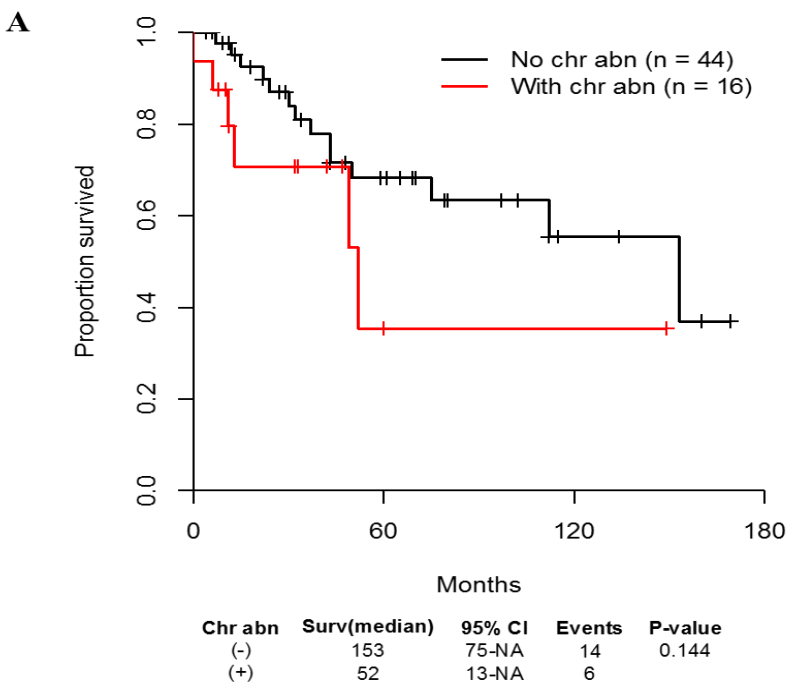
3.5. Prognostic relevance of cytogenetic abnormalities and gene mutations

There was no correlation between aberrant karyotype and survival (*P*=0.144), although complex karyotypes were associated with poor prognosis (*P*=0.017) (Figure 8). Abnormal and normal FISH results did not correlate with survival (*p*=0.491). Additionally, trisomy 12, 13q14 deletion 17p13 deletion and 11q23 deletion did not correlate with survival (*P*=0.392, *P*=0.556, *P*=0.708, and *P*=0.176, respectively). There was no correlation between prognosis and the detection of one, two or three abnormalities by FISH (*P*=0.812) (Figure 9). Finally, the presence of ZAP-70 mutation-positive (>20% tumor cells) and ZAP-70 mutation-negative (<20% tumor cells) cells showed no correlation with survival (*P*=0.744)

A greater number of mutations (n≥2) per patient was linked to poor survival (*P*=0.008) (Figure 10). Patients carrying $TP53_{mut}$ and $MYD88_{mut}$ exhibited shorter survival compared with patients without these mutations (*P*=0.023 and *P*=0.005, respectively) (Figure 11). $SETBP1_{mut}$, $ITPKB_{mut}$, $SAMHD1_{mut}$ and $EGR2_{mut}$ were associated with shorter survival (*P*=0.032, *P*=0.011, *P*=0.049 and *P*=0.032, respectively); however, these results require careful interpretation because only one patient displayed a mutation in each gene. $TP53_{mut}$ and $MYD88_{mut}$ were also associated with low rates of disease-free survival (*P*=0.011, *P*=0.006) (Figure 12). We excluded $ITPKB_{mut}$, $TCF12_{mut}$, and $RUNX1_{mut}$ because they were singletons, although they were related to

low rates of disease-free survival (*P*=0.030, *P*=0.010 and *P*=0.030, respectively).

In Caucasians, mutations in the *ATM*, *SF3B1*, and *NOTCH1* genes have been reported as poor prognostic factors. *ATM* mutations are related to rapid disease progression and a shorter treatment–free interval, and *SF3B1* mutations are related to rapid disease progression and poor OS. *NOTCH1* mutations have been reported as independent predictors of poor survival. However, mutations in these 3 genes were not associated with poor survival in the present study (Figure 13).



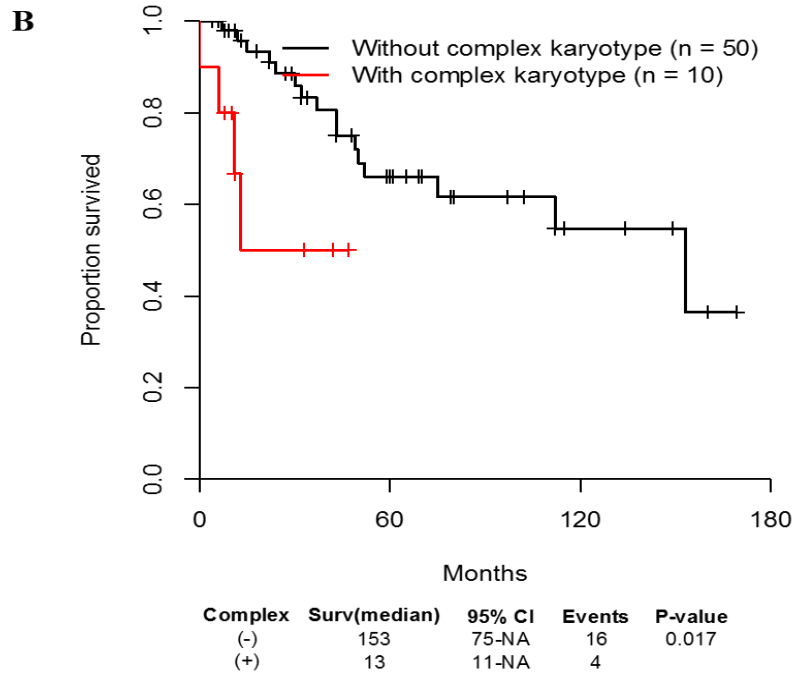


Figure 8. Kaplan-Meier survival curves according to cytogenetic abnormality.

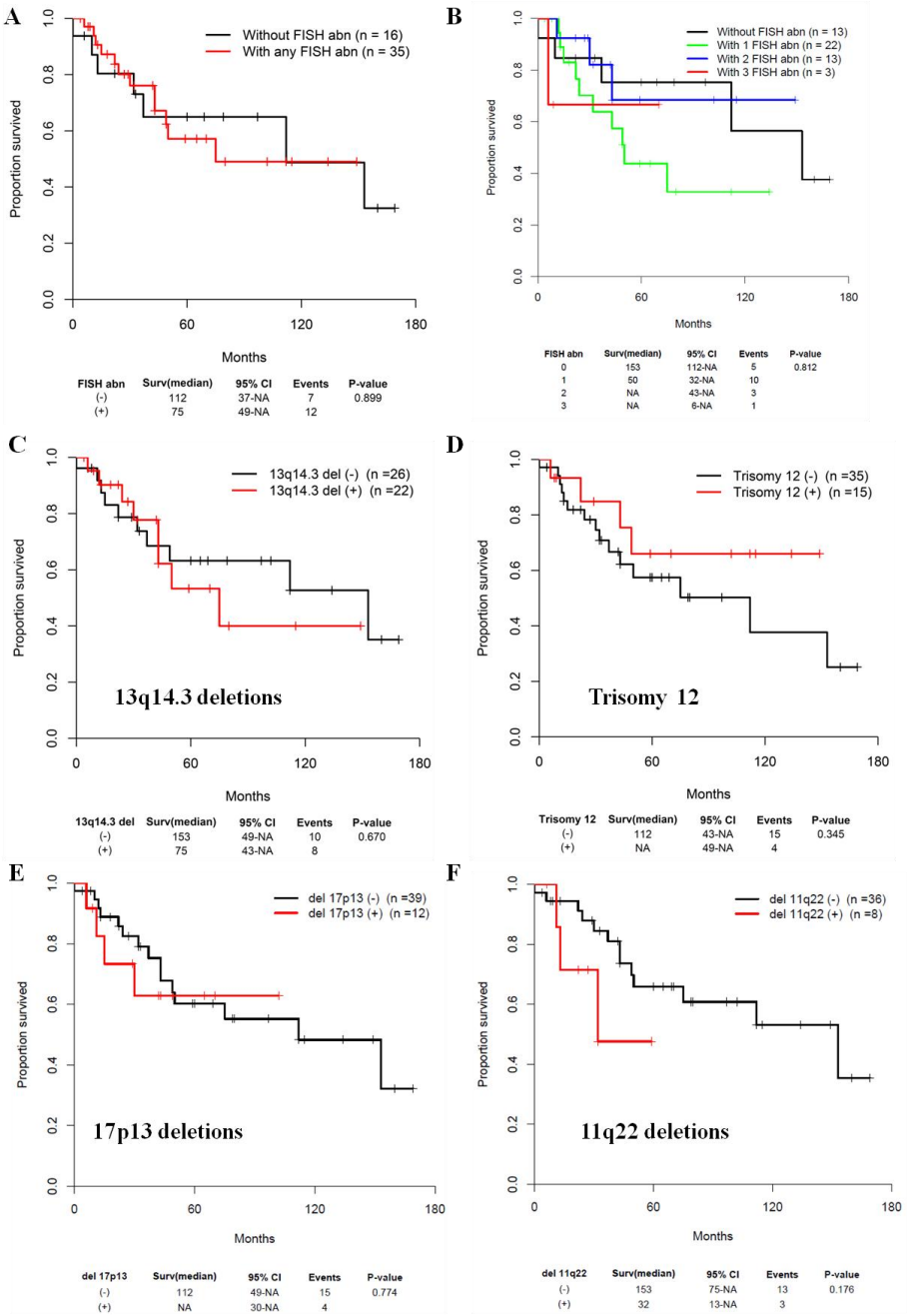
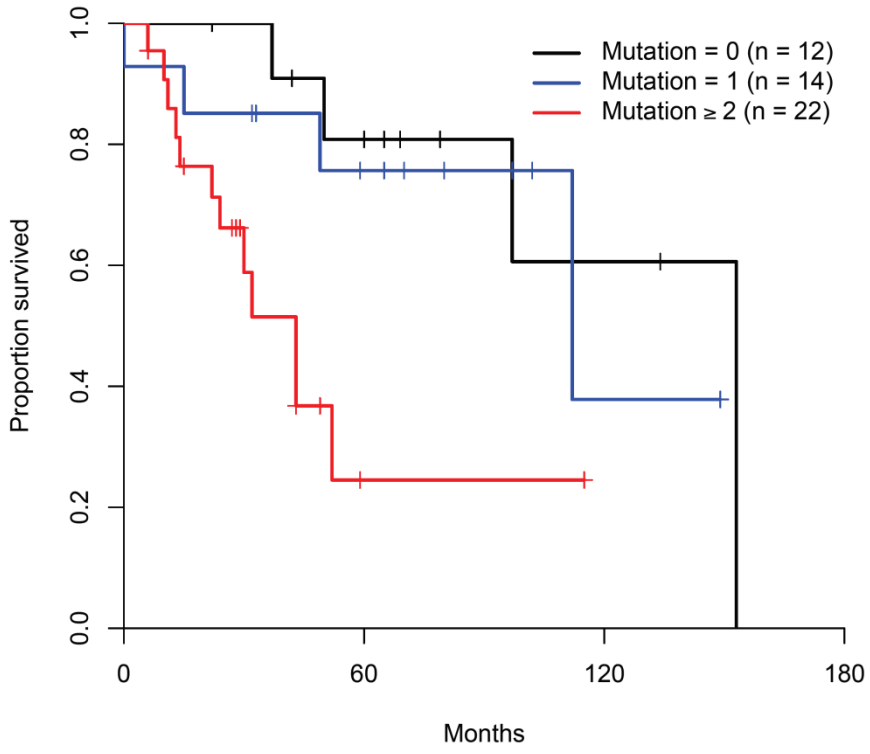


Figure 9. Kaplan-Meier survival curves according to abnormalities detected by FISH.



Mutation	Surv(median)	95% CI	Events	P-value
0	153	97-NA	4	0.008
1	153	112-NA	4	
≥ 2	43	24-NA	12	

Figure 10. Kaplan-Meier survival curves according to the number of mutations.

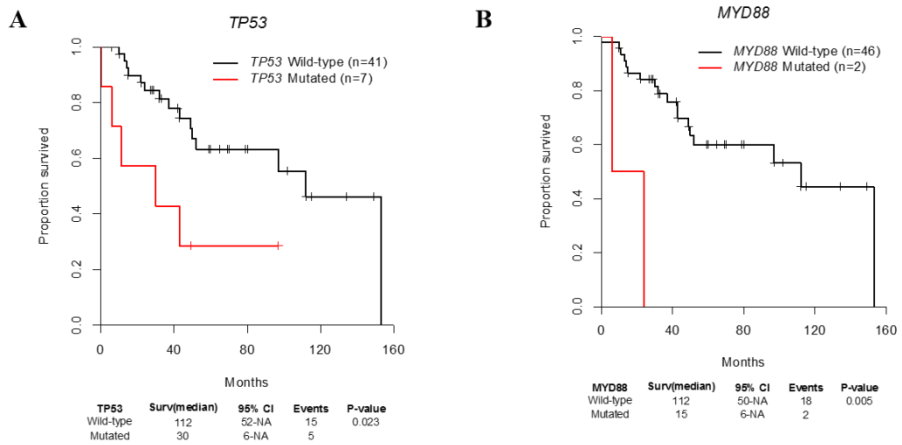


Figure 11. Kaplan-Meier survival curves for overall survival for patients with mutations in the (A) *TP53* and (B) *MYD8*8 genes.

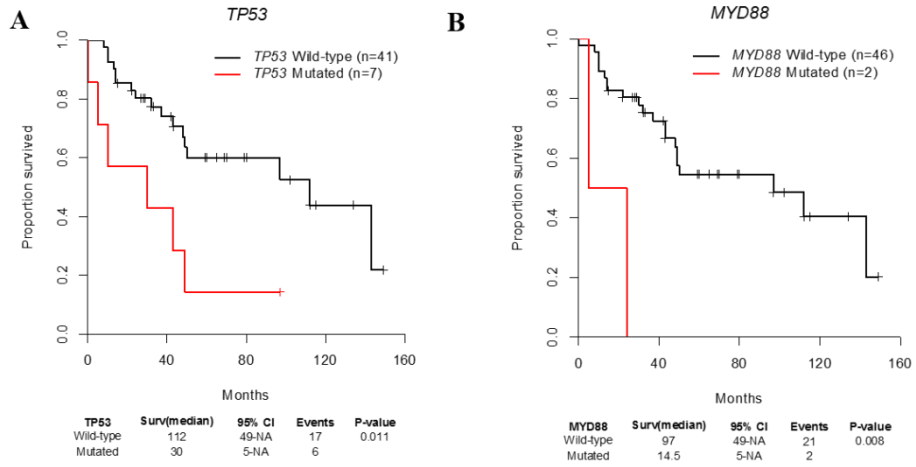


Figure 12. Disease-free survival curves for patients with mutations in the (A) *TP53* and (B) *MYD88* genes.

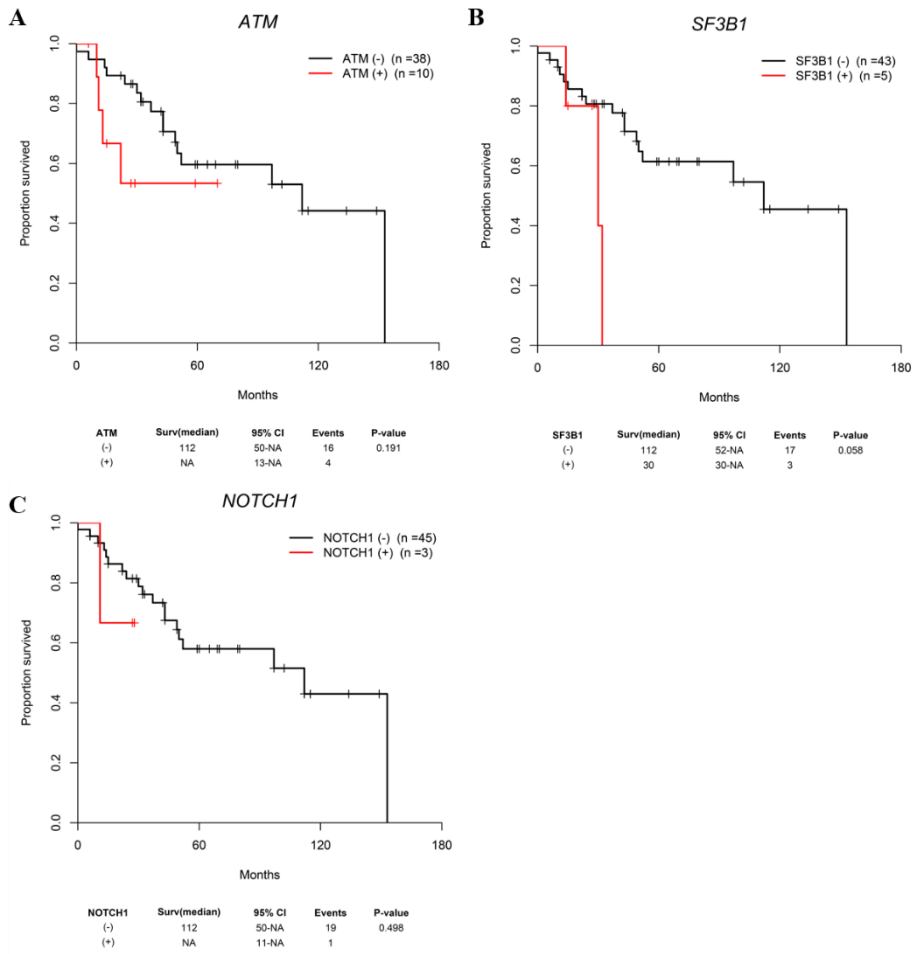


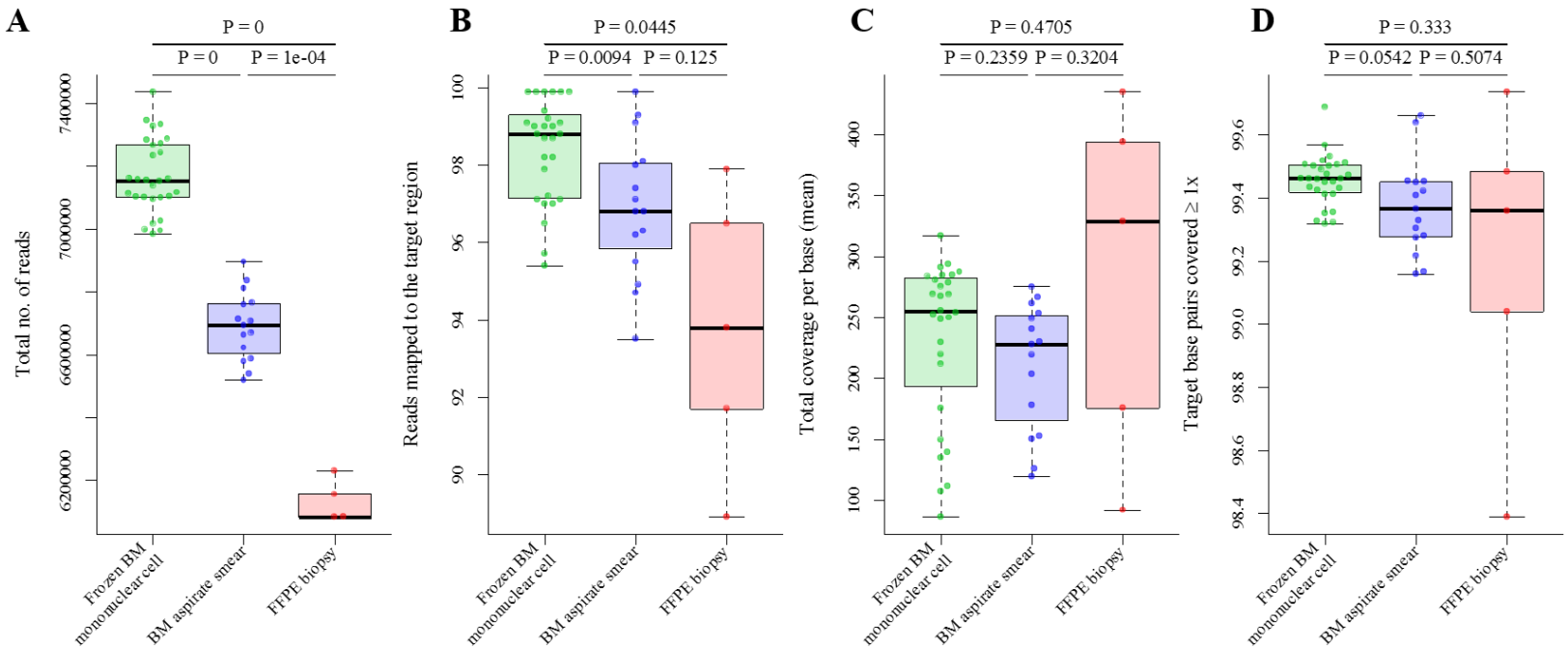
Figure 13. Kaplan-Meier survival curves for genes previously associated with a poor prognosis in Caucasians. (A) *ATM*, (B) *SF3B1*, and (C) *NOTCH1*.

3.6. Comparison of the target-based sequencing quality of DNA from various origins

We subjected three different type of samples to DNA extraction: frozen BM mononuclear cells, unstained BM aspirate smear slides, and FFPE BM and LN biopsies. The different sample types showed different results regarding the quality of target-based sequencing (Table 7). The total number of reads (frozen BM mononuclear cells and unstained BM aspirate smear slides, $P<0.001$; unstained BM aspirate smear slides and FFPE BM and LN biopsies, $P<0.001$; frozen BM mononuclear cells and FFPE BM and LN biopsies, $P<0.001$) and the reads that mapped to the target region (frozen BM mononuclear cells and unstained BM aspirate smear slides, $P<0.009$; frozen BM mononuclear cells and FFPE BM and LN biopsies, $P<0.05$) significantly differed between the sample types, but the target coverage per base and target base pairs covered >1x, >5x, >10x, and >100x showed no significant differences (Figure 14).

Table 7. Summary of target-capture sequencing results for various sample types.

	Frozen BM mononuclear cells (n=28)	Unstained BM aspirate smear slides (n=15)	Formalin-fixed paraffin-embedded (FFPE) biopsies (n=5)
Total no. of reads	7,168,599	6,691,043	6,085,536
Reads mapped to the target region, %	98.4	96.9	93.8
Target coverage per base (mean)	232.5	210.2	285.1
Target base pairs covered ≥10 X, %	97.9	97.7	95
Target base pairs covered ≥100 X, %	86.2	84.7	86.3


A
P = 0
P = 0
P = 1e-04
Total no. of reads
6200000
6600000
7000000
7400000
Frozen BM mononuclear cell
BM aspirate smear
FFPE biopsy
B
P = 0.0445
P = 0.0094
P = 0.125
Reads mapped to the target region
90
92
94
96
98
100
Frozen BM mononuclear cell
BM aspirate smear
FFPE biopsy
C
P = 0.4705
P = 0.2359
P = 0.3204
Total coverage per base (mean)
100
150
200
250
300
350
400
Frozen BM mononuclear cell
BM aspirate smear
FFPE biopsy
D
P = 0.333
P = 0.0542
P = 0.5074
Target base pairs covered ≥ 1x
98.4
98.6
98.8
99.0
99.2
99.4
99.6
Frozen BM mononuclear cell
BM aspirate smear
FFPE biopsy

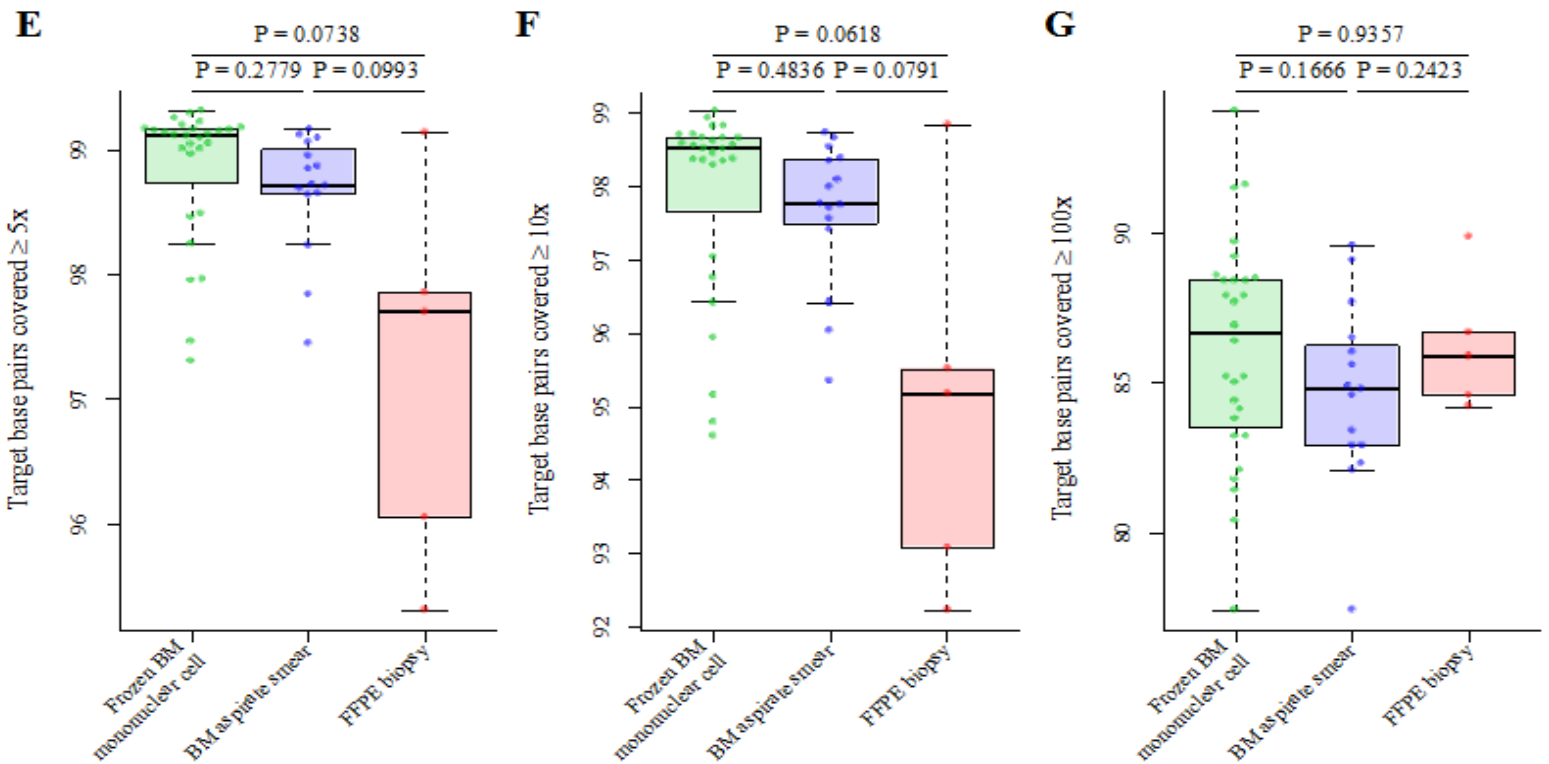


Figure 14. Comparison of target-based sequencing quality between various sample types.

4. DISCUSSION

This study represents the first use of targeted sequencing analysis to examine mutations in Korean patients with CLL. The distribution of cytogenetic aberrations was similar between Caucasian and Korean patients, although the distribution of genetic aberrations was different. Among Koreans, *ATM*, *TP53* and *SF3B1* represented more than 10% of the mutated genes (20.8%, 14.6%, and 10.4%, respectively) (Table 8). The frequencies of $TP53_{mut}$ and $SF3B1_{mut}$ were similar in Koreans and Caucasians, although ATM_{mut} was more common in Koreans (Figure 13). $KLHL6_{mut}$ and $BCOR_{mut}$ were more common in Koreans, while $NOTCH1_{mut}$, $LRP1B_{mut}$, $CHD2_{mut,}$ $POT1_{mut}$ and $TGM7_{mut}$ were more common in Caucasians. The frequencies of $MYD88_{mut}$, $FAT4_{mut}$, $SAMHD1_{mut}$, $DDX3X_{mut}$, $ITPKB_{mut}$, and $MED12_{mut}$ were similar in Koreans and Caucasians. Mutations in the *LAMB4*, *SH2B3*, *RUNX1*, *SCRIB*, *KIT*, *GATA2*, *CEBPA*, *TCF12*, *STAG2*, *ZRSR2*, *SF1*, *CSF1R*, and *SETBP1* genes were identified in this study but have not been previously reported in Caucasian patients with CLL. One Chinese study reporting mutation patterns in patients with CLL [24] found that *TP53* and *NOTCH1* showed mutation rates similar to those that we detected in Korean patients (15% vs. 15% and 8% vs. 6%), although mutations in *SF3B1* were twice as common in Koreans (5% vs. 10%) and mutations in *MYD88* were half as common in Koreans (8% vs. 4%). Thus, differences in the distribution of mutations were observed between Asian populations.

Studies of Caucasian populations have reported that mutations in *TP53*, *BIRC3*, *ATM*, *NOTCH1*, and *SF3B1* are associated with poor prognosis [27]. In this study, mutations in *TP53* and *MYD88* were associated with poor prognosis; only mutations in *TP53* had similar prognostic associations in our study population and in Caucasians. While mutations in both *TP53* and *NOTCH1* were identified as poor prognostic factors in a Chinese study [24], *TP53* was the only gene found to show a poor prognostic association in the Korean patients. $MYD88_{mut}$ has been previously reported in Caucasian patients with CLL [15], and the same mutation (p.L265P) was shown in this study. Puente et al. revealed that $MYD88_{mut}$ is related to mutated immunoglobulin genes [15]; mutated immunoglobulin genes are related to better prognosis [28]. As in this study, MYD88 is associated with poor prognosis. The discrepancy between prognoses may be the difference between ethnicities; thus, further studies should be conducted.

Several genes were reported only in Korean patients with CLL and have not been reported in Caucasians. A *LAMB4* mutation has been reported in MDS patients [29]; this protein-coding gene mediates the attachment and migration of cells into tissues during embryonic development. Laminins are important for tumor invasion as well as the maintenance of normal epithelial-cell structures because they are associated with cell adhesion. The putative effect of this mutation on protein function is unknown. *RUNX*, *ZRSR2*, and *SF1* have also been reported in MDS [30, 31]; *RUNX1* plays an important role in

regulating the transcription of many tumor-suppressor genes [32], while *ZRSR2* is an essential component of the splicing machinery [31], and SF1 is a component of the RNA-splicing machinery [31]. *GATA2*, *CEBPA* and *CSF1R* have been reported in MDS and acute myeloid leukemia (AML) [33-35]. *GATA2* is required for the proliferation and survival of early hematopoietic cells [36], *CEBPA* is a transcription factor that plays important roles in myeloid differentiation [37], and *CSF1R* instructs myeloid lineage fate decisions in hematopoietic stem cells [38]. *SH2B3* and *TCF12* have been linked to acute lymphocytic leukemia [39, 40]; *SH2B3* regulates integrin signaling in endothelial cells [41], and *TCF12* is associated with lymphoid differentiation [40]. *KIT* encodes a transmembrane glycoprotein [42] that has been associated with AML [43]. *SCRIB* has been reported in myeloproliferative neoplasms and regulates the differentiation of planar cell polarity [44]. *STAG2* has been reported in myeloid diseases, such as MDS, AML and chronic myelomonocytic leukemia, and encodes components of the cohesion complex [45]. SETBP1 mutations are observed in atypical chronic myeloid leukemia [46].

Binet staging is one of the clinical staging systems that is commonly used to predict prognosis. In the present study, no correlation between Binet stage and mutation numbers was observed. Additionally, the number of patients who had mutations in genes related to poor prognosis, such as *TP53* and *MYD88*, between stage C and other stages did not correlate (P=0.1434). Furthermore,

allele burden also showed no correlation between stages. These results can be used in consideration of gene mutation status when predicting prognosis and determining treatment strategies.

In the case of Richter's syndrome, a previous study in a Caucasian population used whole exome sequencing and reported ~20 genetic lesions per case [47]. While there was an average of 1.8 mutations per case in the previous study, the number of mutations markedly differed between studies. *TP53* disruption, c-MYC abnormalities [48], *NOTCH1* mutation [49], *BRAF* [50], and *CDKN2A/B* mutation [47] were reported to correlate with Richter's syndrome. In this study, *TP53*, *NOTCH1*, and *CDKN2A* mutations were identified in Richter's syndrome cases, but these genes did not differ between patients with Richter's syndrome and patients without Richter's syndrome (P=0.2061, P=0.3363, and P=0.1250, respectively).

For the cytogenetic aberrations shown in Table 9, most studies detected more abnormalities by FISH than by conventional cytogenetic methods. The proportions of aberrations detected by FISH were somewhat similar to those reported in previous studies in Western countries: 13q14 deletion was the most frequent chromosomal aberration, followed by trisomy 12, 17p deletion and 11q deletion. In a study in Korean patients (n=48), trisomy 12 was the most frequently observed abnormality [51], while another Korean study (n=16) [8] reported rankings similar to those reported here. This discrepancy may merit further in-depth studies in additional Korean patients with CLL.

In Western CLL studies, trisomy 12 is associated with advanced disease and a less favorable prognosis [52, 53]. patients with CLL with chromosomal aberrations such as the 11q deletion or 17p deletion had poorer outcomes compared with patients with normal karyotypes [54]. However, in this study, abnormalities detected by FISH were not associated with significant differences.

Massively parallel sequencing provides new insights that enable the systematic discovery of the genetic aberrations underlying diseases and lead to the identification of new clinically relevant targets. Even with the major breakthroughs achieved through recent whole-genome and exome studies, there had not been sufficient evidence of recurrent mutations associated with CLL. Because CLL is clinically and biological heterogeneous, no comprehensive molecular characterization of CLL has yet been fully undertaken, especially in Asian populations due to its low incidence in Asia.

In conclusion, this study is the first comprehensive NGS-based study of Korean cohorts. Although a limited number of cases were analyzed, the results of this study contribute to the characterization of the mutational landscape of the Asian population and identify new gene variants involved in the pathogenesis of CLL. Further NGS-based sequencing studies in other Asian countries will provide valuable insight into the genetic architecture of this disease. The true biological nature of CLL is still unknown; thus, studies including clinical variables based on well-defined CLL cohorts represent an

area of interest for future work.

Table 8. Comparison of frequencies of mutated genes (%) between Caucasian and Korean populations

Gene	Landau et al. [14], USA (n=160)	Puente et al. [15], Spain*	Fabbri et al. [16], USA (n=53)	Quesada et al. [20], Spain†	Wang et al. [55], USA (n=91)	Domenech et al. [56], Spain (n=10)	Xia et al. [24], China ‡ (n=307)	present study, Korea (n=48)
NOTCH1	10%	12.2%	15%	12.1%	4%		8%	6%
ATM	8%				9%			21%
TP53	13%		7%		15%		15%	15%
SF3B1	14%			9.7%	15%		5%	10%
MYD88	8%	2.9%			10%		8%	4%
BIRC3			4%				2%	0%
FAT4					5%			4%
LRP1B				4.8%				2%
CHD2	4%			4.8%				2%
POT1	3%			4.8%				2%
XPO1	4%	2.4%						0%
ZMYM3	3%				4%			2%
FBXW7					4%			0%
DDX3X	2%				3%			2%
PLEKHG5			4%					0%
TGM7			4%					0%
MAPK1					3%			0%
HIST1H1E	3%							0%
NRAS	3%							0%
BCOR	3%							6%
RIPK1	3%							0%
SAMHD1	3%							2%
KLHL6		1.8%						8%
MED12	2%							2%
ITPKB	2%							2%
KRAS	2%							0%
EGR2	1%							2%
NFKBIE						Novel gene		0%
PRKD3						Novel gene		2%

*n=255 for *NOTCH1*, 310 for *MYD88*, 165 for *XPO1*, and 160 for *KLHL6*

† n=255 for *NOTCH1*, 279 for *SF3B1*, 105 for *POT1*, *CHD2*, and *LRP1B*

‡ Sanger sequencing was performed in this study.

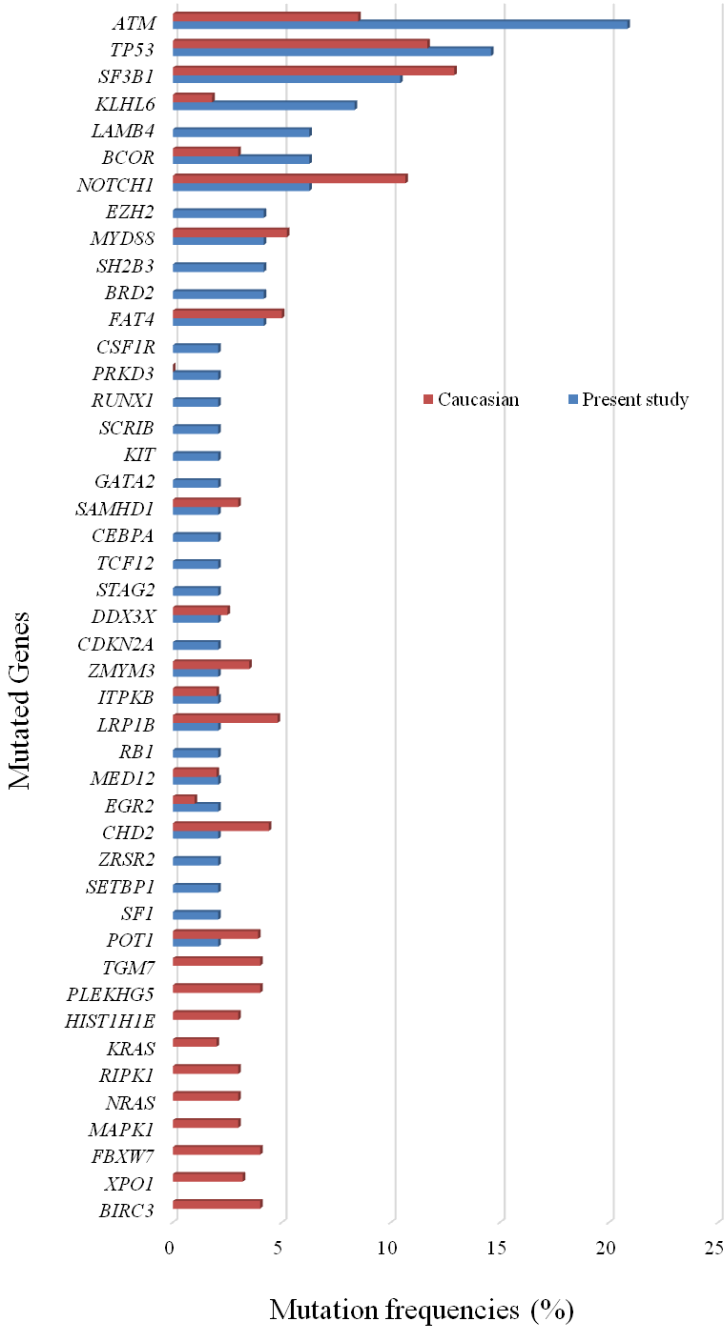


Figure 15. Comparison of mutation frequencies (%) between Caucasian and Korean populations.

Table 9. Incidence (%) of chromosomal abnormalities in patients with CLL (previous and present studies).

Aberration	Wu et al. [57], Taiwan (n=83)	Dong et al. [58], China (n=173)	Döhner et al. [59] Germany (n=325)	Chena et al. [60] Argentina (n=57)	Glassman and Hayes [61] USA (n=100)	Grever et al. [62], USA (n=235)	Chang et al. [8] Korea (n=16)	Yoon et al. [51] Korea (n=48)	present study Korea (n=72)
13q14 deletion	45.8	27.7	55	63.2	40	34.0	69	12.5	45.8
Trisomy 12	20.5	21.9	16	35	11	20.4	19	29.1	30
17p deletion	10.8	7.5	7	11	12	8.1	0	10.4	23.5
11q deletion	13.3	10.9	18	ND	23	17.0	14	12.5	18.2
By conventional chromosome analysis									
Clonal abnormalities	42.2	ND	NA	48.8	28	NA	15	25.5	28
Normal karyotype	57.8	ND	NA	51.2	72	NA	85	74.5	65
By fluorescence *in situ* hybridization									
Clonal abnormalities	69.9	ND	82	80.7	64		84.3 63	52	66.7

Normal karyotype	30.1	ND	18	19.3	36	15.7	37	48	33.3
Cytogenetic methods	conventional cytogenetic studies	Conventional cytogenetic studies	NA	Conventional cytogenetic studies	Conventional cytogenetic studies	NA	Conventional cytogenetic studies	Conventional cytogenetic studies	Conventional cytogenetic studies
B cell stimulator	12-O-tetradecanoyl phorbol-13-acetate	interleukin 2 and CpG-oligonucleotide DSP30	NA	pokeweed mitogen, and lipopolysaccharide	lipopolysaccharide (LPS) mitogen	NA	NA	Lipopolysaccharide	TPA
FISH	done	done	done	done	done	done	done	done	done

REFERENCES

1. Weiss, N.S., *Geographical variation in the incidence of the leukemias and lymphomas.* National Cancer Institute Monograph, 1979(53): p. 139-142.
2. Boggs, D.R., et al., *Chronic lymphocytic leukemia in China.* American journal of hematology, 1987. 25(3): p. 349-354.
3. Kobayashi, T., et al., *[Chronic lymphocytic leukemia in Japan].* [Rinsho ketsueki] The Japanese journal of clinical hematology, 1990. 31(5): p. 554-563.
4. Yavorkovsky, L., et al., *B-chronic lymphocytic leukaemia in Latvia: Epidemiological aspects.* European journal of haematology, 1993. 51(4): p. 214-217.
5. Groves, F., M. Linet, and S. Devesa, *Patterns of occurrence of the leukaemias.* European Journal of Cancer, 1995. 31(6): p. 941-949.
6. Chen, P.-m., et al., *Genotypic characterization and multivariate survival analysis of chronic lymphocytic leukemia in Taiwan.* Acta haematologica, 1997. 97(4): p. 196-204.
7. Tamura, K., et al., *Chronic lymphocytic leukemia (CLL) is rare, but the proportion of T-CLL is high in Japan.* European journal of haematology, 2001. 67(3): p. 152-157.
8. Chang, Y.H., et al., *Korean patients with chronic lymphocytic leukemia show the similar types of chromosomal aberrations as those in Europe and North America.* Leuk Res, 2006. 30(6): p. 695-9.
9. Dores, G.M., et al., *Chronic lymphocytic leukaemia and small lymphocytic lymphoma: overview of the descriptive epidemiology.* British journal of haematology, 2007. 139(5): p. 809-819.
10. Ahn, Y.O., et al., *Incidence estimation of leukemia among Koreans.* Journal of Korean medical science, 1991. 6(4): p. 299.
11. FINCH, S.C., et al., *Chronic lymphocytic leukemia in Hiroshima and Nagasaki, Japan.* Blood, 1969. 33(1): p. 79-86.
12. Gale, R.P., et al., *Decreased chronic lymphocytic leukemia incidence in Asians in Los Angeles County.* Leukemia research, 2000. 24(8): p.

665-669.

13. Matsuda, A., et al., *Difference in clinical features between Japanese and German patients with refractory anemia in myelodysplastic syndromes.* Blood, 2005. 106(8): p. 2633-2640.
14. Landau, D.A., et al., *Evolution and impact of subclonal mutations in chronic lymphocytic leukemia.* Cell, 2013. 152(4): p. 714-26.
15. Puente, X.S., et al., *Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia.* Nature, 2011. 475(7354): p. 101-5.
16. Fabbri, G., et al., *Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation.* J Exp Med, 2011. 208(7): p. 1389-401.
17. Skowronska, A., et al., *ATM germline heterozygosity does not play a role in chronic lymphocytic leukemia initiation but influences rapid disease progression through loss of the remaining ATM allele.* Haematologica, 2012. 97(1): p. 142-6.
18. Zenz, T., et al., *Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up.* Blood., 2008. 112(8): p. 3322-9. doi: 10.1182/blood-2008-04-154070. Epub 2008 Aug 8.
19. El Rouby, S., et al., *p53 gene mutation in B-cell chronic lymphocytic leukemia is associated with drug resistance and is independent of MDR1/MDR3 gene expression.* Blood, 1993. 82(11): p. 3452-3459.
20. Quesada, V., et al., *Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia.* Nat Genet, 2012. 44(1): p. 47-52.
21. Lan, Q., et al., *Genetic susceptibility for chronic lymphocytic leukemia among Chinese in Hong Kong.* European journal of haematology, 2010. 85(6): p. 492-495.
22. Di Bernardo, M.C., et al., *A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia.* Nature genetics, 2008. 40(10): p. 1204-1210.

23. Vaqué, J.P., et al., *B-cell lymphoma mutations: improving diagnostics and enabling targeted therapies.* haematologica, 2014. 99(2): p. 222-231.
24. Xia, Y., et al., *Frequencies of SF3B1, NOTCH1, MYD88, BIRC3 and IGHV mutations and TP53 disruptions in Chinese with chronic lymphocytic leukemia: disparities with Europeans.* Oncotarget, 2014. 31: p. 31.
25. Lee, S.J., et al., *Gradual Increase of Chronic Lymphocytic Leukemia Incidence in Korea, 1999-2010: Comparison to Plasma cell myeloma.* Leuk Lymphoma, 2015. 1: p. 1-19.
26. Hallek, M., 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): p. 5446-56.
27. Puiggros, A., G. Blanco, and B. Espinet, *Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go.* Biomed Res Int, 2014. 2014:435983.(doi): p. 10.1155/2014/435983. Epub 2014 May 22.
28. Hamblin, T.J., et al., *Unmutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia.* Blood, 1999. 94(6): p. 1848-1854.
29. Grossmann, V., et al., *Landscape of genetic lesions in 944 patients with myelodysplastic syndromes.* Blood, 2013. 122(21): p. 521-521.
30. Bejar, R., et al., *Clinical effect of point mutations in myelodysplastic syndromes.* N Engl J Med., 2011. 364(26): p. 2496-506. doi: 10.1056/NEJMoa1013343.
31. Yoshida, K., et al., *Frequent pathway mutations of splicing machinery in myelodysplasia.* Nature, 2011. 478(7367): p. 64-69.
32. Silva, F.P., et al., *Identification of RUNX1/AML1 as a classical tumor suppressor gene.* Oncogene, 2003. 22(4): p. 538-547.
33. Hyde, R.K. and P.P. Liu, *GATA2 mutations lead to MDS and AML.* Nature genetics, 2011. 43(10): p. 926-927.

34. Fuchs, O., et al., *CEBPA polymorphisms and mutations in patients with acute myeloid leukemia, myelodysplastic syndrome, multiple myeloma and non-Hodgkin's lymphoma.* Blood Cells, Molecules, and Diseases, 2008. 40(3): p. 401-405.
35. Bernasconi, P., *Molecular pathways in myelodysplastic syndromes and acute myeloid leukemia: relationships and distinctions–a review.* British journal of haematology, 2008. 142(5): p. 695-708.
36. Tsai, F.-Y. and S.H. Orkin, *Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation.* Blood, 1997. 89(10): p. 3636-3643.
37. Tenen, D.G., et al., *Transcription factors, normal myeloid development, and leukemia.* Blood, 1997. 90(2): p. 489-519.
38. Mossadegh-Keller, N., et al., *M-CSF instructs myeloid lineage fate in single haematopoietic stem cells.* Nature, 2013. 497(7448): p. 239-243.
39. Perez-Garcia, A., et al., *Genetic loss of SH2B3 in acute lymphoblastic leukemia.* Blood, 2013. 122(14): p. 2425-2432.
40. Tijchon, E., et al., *B-lineage transcription factors and cooperating gene lesions required for leukemia development.* Leukemia, 2013. 27(3): p. 541-552.
41. Devallière, J., et al., *LNK (SH2B3) is a key regulator of integrin signaling in endothelial cells and targets α-parvin to control cell adhesion and migration.* The FASEB Journal, 2012. 26(6): p. 2592-2606.
42. Yarden, Y., et al., *Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand.* EMBO J., 1987. 6(11): p. 3341-51.
43. Goemans, B., et al., *Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia.* Leukemia, 2005. 19(9): p. 1536-1542.
44. Tenedini, E., et al., *Targeted cancer exome sequencing reveals recurrent mutations in myeloproliferative neoplasms.* Leukemia.,

2014. 28(5): p. 1052-9. doi: 10.1038/leu.2013.302. Epub 2013 Oct 22.

45. Rocquain, J., et al., *Alteration of cohesin genes in myeloid diseases.* American journal of hematology, 2010. 85(9): p. 717-719.

46. Piazza, R., et al., *Recurrent SETBP1 mutations in atypical chronic myeloid leukemia.* Nature genetics, 2013. 45(1): p. 18-24.

47. Fabbri, G., et al., *Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome.* J Exp Med, 2013. 210(11): p. 2273-88.

48. Rossi, D., et al., *The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation.* Blood, 2011. 117(12): p. 3391-401.

49. Rossi, D., et al., *Different impact of NOTCH1 and SF3B1 mutations on the risk of chronic lymphocytic leukemia transformation to Richter syndrome.* Br J Haematol, 2012. 158(3): p. 426-9.

50. Sellar, R.S., et al., *BRAF(V600E) mutations are found in Richter syndrome and may allow targeted therapy in a subset of patients.* Br J Haematol, 2015. 170(2): p. 282-5.

51. Yoon, J.H., et al., *Validation of Western common recurrent chromosomal aberrations in Korean chronic lymphocytic leukaemia patients with very low incidence.* Hematol Oncol, 2013. 7(10).

52. Gonzalez-Gascon, Y.M.I., et al., *A high proportion of cells carrying trisomy 12 is associated with a worse outcome in patients with chronic lymphocytic leukemia.* Hematol Oncol, 2015. 17(10).

53. Escudier, S.M., et al., *Fluorescent in situ hybridization and cytogenetic studies of trisomy 12 in chronic lymphocytic leukemia.* Blood., 1993. 81(10): p. 2702-7.

54. Nelson, B.P., et al., *Chronic lymphocytic leukemia FISH panel impact on diagnosis.* American journal of clinical pathology, 2007. 128(2): p. 323-332.

55. Wang, L., et al., *SF3B1 and other novel cancer genes in chronic lymphocytic leukemia.* New England Journal of Medicine, 2011. 365(26): p. 2497-2506.

56. Doménech, E., et al., *New mutations in chronic lymphocytic leukemia*

identified by target enrichment and deep sequencing. PloS one, 2012. 7(6): p. e38158.

57. Wu, S.J., et al., *Chromosomal abnormalities by conventional cytogenetics and interphase fluorescence in situ hybridization in chronic lymphocytic leukemia in Taiwan, an area with low incidence--clinical implication and comparison between the West and the East.* Ann Hematol., 2013. 92(6): p. 799-806. doi: 10.1007/s00277-013-1700-x. Epub 2013 Feb 17.

58. Dong, H.J., et al., *The prognostic significance of TP53 mutations in Chinese patients with chronic lymphocytic leukemia is independent of del(17p13).* Ann Hematol., 2011. 90(6): p. 709-17. doi: 10.1007/s00277-010-1125-8. Epub 2010 Nov 27.

59. Döhner, H., et al., *Genomic aberrations and survival in chronic lymphocytic leukemia.* New England Journal of Medicine, 2000. 343(26): p. 1910-1916.

60. Chena, C., et al., *Interphase cytogenetic analysis in Argentinean B-cell chronic lymphocytic leukemia patients: association of trisomy 12 and del (13q14).* Cancer genetics and cytogenetics, 2003. 146(2): p. 154-160.

61. Glassman, A.B. and K.J. Hayes, *The value of fluorescence in situ hybridization in the diagnosis and prognosis of chronic lymphocytic leukemia.* Cancer genetics and cytogenetics, 2005. 158(1): p. 88-91.

62. Grever, M.R., et al., *Comprehensive assessment of genetic and molecular features predicting outcome in patients with chronic lymphocytic leukemia: results from the US Intergroup Phase III Trial E2997.* J Clin Oncol., 2007. 25(7): p. 799-804. Epub 2007 Feb 5.

국문 초록

서론: 만성 림프구성 백혈병은 서양에서는 가장 흔한 백혈병이나, 아시아권의 국가에서는 매우 드문 질환이다. 이에 따라, 아직까지 아시아권의 국가에서는 만성 림프구성 백혈병에 관한 유전적 프로파일이 보고되지 않고 있다. 이번 연구의 목적은, 한국인 만성 림프구성 백혈병 환자의 유전체 프로파일을 특징 짓고, 예후에 미치는 의미와 체세포 돌연변이의 인종차이를 밝혀내 보고자 하였다.

방법: 총 71 명의 만성 림프구성 백혈병 환자가 이 연구에 포함되었다. 환자들의 진단 당시의 평균연령은 61 세였으며, 나이 범위는 23 세에서 81 세였다. 표적 유전자 염기서열분석은 총 87 개의 유전자에 대해서, 48 명의 환자를 대상으로 시행하였다. 표적 유전자 염기서열 분석은 IlluminaHiSeq 2500 으로 검사를 진행하였으며, 변이로 나온 후보 유전자들은 생거 염기서열분석법으로 확인하였다. 71 명의 환자에서, 60 명은 G-분염법으로 염색체 검사를 시행하였고, 51 명의 환자에게는 12 번 염색체 숫자확인, 13q14.3 부위 결실, 17p13 부위 결실, 그리고 11q22 부위 결실 확인을 위한 형광 제자리 부합법을 시행하였다.

결과: G-분염법을 이용한 염색체 검사에서 관찰된 비정상소견은 28.3% 였으며 형광 제자리 부합법으로 확인한 비정상소견은 66.7% 였다. 표적 유전자 염기서열 분석법에서는 71 개의 염기치환 및

삽입/결손이 관찰되었으며 표적 부위의 평균 커버리지는 231X 였다. 48 명의 환자에서는 평균 1.6 개의 변이(범주: 0-6 개)가 관찰되었으며, 48 명의 환자 중 36 명의 환자 (75%)가 적어도 1 개 이상의 변이를 가지고 있었고 변이를 보인 환자당 평균 변이 수는 2.1 개 (범주: 1-6 개) 였다. 상위 변이는 (>5%) *ATM* (20.8%), *TP53* (14.6%), *SF3B1* (10.4%), *KLHL6* (8.3%), *BCOR* (6.25%), *LAMB4* (6.25%), *NOTCH1* (6.25%) 순서였다. *TP53* 과 *MYD88* 유전자 변이는 좋지 않은 예후와 관련이 있었다. *TP53* 과 *SF3B1* 유전자 변이 발생률은 서양인의 발생률과 비슷하였으나, *ATM, NOTCH1, LRO1B, CHD2, POT1, TGM7* 은 한국인에서 그 변이 발생률이 더 높았다. 특히 *ATM* 변이는 한국인이 서양인보다 발생률이 2 배 높았다 (20.8% vs. 9%). 서양인 만성 림프구성 백혈병 환자에서 한번도 보고되지 않았으나, 이번 연구에서 새로이 관찰된 변이는 *LAMB4, SH2B3, RUNX1, SCRIB, KIT, GATA2, CEBPA, TCF12, STAG2, ZRSR2, SF1* 였다.

결론: 한국인 만성 림프구성 백혈병 환자의 세포유전학적 이상소견은 서양인의 이상소견과 비슷하였으나, 변이 프로파일은 서양인과 달랐다. 이번 연구에서 관찰된 새로운 변이들은 더 큰 코호트를 통하여 확인되어야 하겠으며, 이러한 차이는 질병 발생률에서 인종차이가 나는 점을 설명할 수 있는 사항으로 제안할 수 있겠다. 또한 추후 치료 전략을 세움에 있어서 이러한 유전자 변이에 대한 사항도 고려해 보아야 하겠다.

주요어 : 만성 림프구성 백혈병, 염색체 이상, 유전자 분석, 민족성

학 번 : 2012-73082