Tailored approaches grounded on immunogenetic features for refined prognostication in chronic lymphocytic leukemia

Panagiotis Baliakas,[1*] Theodoros Moysiadis,[2*] Anastasia Hadzidimitriou,[1,2] Aliki Xochelli,[1,2] Sabine Jeromin,[3] Andreas Agathangelidis,[2] Mattias Mattsson,[1] Lesley-Ann Sutton,[1] Eva Minga,[2] Lydia Scarfò,[4] Davide Rossi,[5] Zadie Davis,[6] Neus Villamor,[7] Helen Parker,[8] Jana Kotaskova,[9] Evangelia Stalika,[2,10] Karla Plevova,[9] Larry Mansouri,[1] Diego Cortese,[1] Alba Navarro,[7] Julio Delgado,[11] Marta Larrayoz,[8] Emma Young,[1] Achilles Anagnostopoulos,[10] Karin E. Smedby,[12] Gunnar Juliusson,[13] Oonagh Sheehy,[14] Mark Catherwood,[14] Jonathan C. Strefford,[8] Niki Stavroyianni,[10] Chrysoula Belessi,[15] Sarka Pospisilova,[9] David Oscier,[6] Gianluca Gaidano,[16] Elias Campo,[7,17] Claudia Haferlach,[3] Paolo Ghia,[4] Richard Rosenquist,[18] Kostas Stamatopoulos;[1,2] on behalf of The European Research Initiative on CLL (ERIC)

[1]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Sweden; [2]Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece; [3]MLL Munich Leukemia Laboratory, Munich, Germany; [4]Division of Experimental Oncology, IRCCS Istituto Scientifico San Raffaele and Università Vita-Salute San Raffaele, Milan, Italy; [5]Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; [6]Department of Haematology, Royal Bournemouth Hospital, UK; [7]Hemopathology Unit, Hospital Clinic, Barcelona, Spain; [8]Cancer Genomics, Academic Unit of Cancer Sciences, Cancer Research UK Centre and Experimental Cancer Medicine Centre, Faculty of Medicine, University of Southampton, UK; [9]Central European Institute of Technology, Masaryk University and University Hospital Brno, Czech Republic; [10]Hematology Department and HCT Unit, G. Papanicolaou Hospital, Thessaloniki, Greece; [11]Hematology Department, Hospital Clinic, Barcelona, Spain; [12]Department of Medicine, Solna, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden; [13]Lund University and Hospital Department of Hematology, Lund Stem Cell Center, Sweden; [14]Department of Hemato-Oncology, Belfast City Hospital, UK; [15]Hematology Department, Nikea General Hospital, Pireaus, Greece; [16]Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; [17]Department of Pathology, University of Barcelona, Spain and [18]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

**PB and TM contributed equally to this work as first authors.*

 doi:10.3324/haematol.2018.195032

Received: April 12, 2018.

Accepted: September 25, 2018.

Pre-published: September 27, 2018.

Correspondence: kostas.stamatopoulos@gmail.com

Tailored approaches grounded on immunogenetic features for refined prognostication in chronic lymphocytic leukemia

Supplementary Appendix

Supplementary Material includes detailed information regarding methodology as well as Supplementary Figures 1-6 and Supplementary Tables 1-5**.**

Patients-Methods

Overall, 2366 general practice patients from 10 academic institutions in Europe who were diagnosed with CLL according to the 2008 International Workshop on CLL (iwCLL) diagnostic criteria and for whom immunogenetic data was available were included in this multicenter retrospective study. Information about the evaluated cohort and biomarkers is provided in Supplementary Table 1. Following the 98% germline identity (GI) cut off, 1364 (58%) patients were classified as mutated (M-CLL) and 1002 (42%) as unmutated (U-CLL). Information regarding gender as well as age and clinical stage at diagnosis were available for the entire cohort. Data on FISH detected abnormalities were available for 1825/2366 (77%) cases with 1162/1260 (92%) of the treated cases being tested before the administration of any treatment. Mutations within the *TP53, NOTCH1, SF3B1, MYD88* and *BIRC3* genes were evaluated in 1544 (65%), 2097 (89%), 1449 (61%), 929 (39%) and 830 (35%) respectively. Regarding time of testing for each mutation, the proportion of treated cases tested before the administration of any treatment ranged from 77-98%. CD38 expression at the time of diagnosis was available in 1649 (70%) patients. A cut-off of 30% was used to indicate CD38 positivity.

Evaluation of biological markers

PCR amplification of IGHV-IGHD-IGHJ rearrangements - Sequence analysis

PCR amplification and sequence analysis of IGHV-IGHD-IGHJ rearrangements were performed on either genomic DNA (gDNA) or complementary DNA (cDNA) as previously reported[1-4]. PCR amplicons were subjected to direct sequencing on both strands. Sequence data were analyzed using the IMGT® databases and the IMGT/V-QUEST tool (http://www.imgt.org). Only productive rearrangements were evaluated. Output data from IMGT/V-QUEST for all productive IGHV-IGHD-IGHJ rearrangements were parsed, reorganized, and exported to a spreadsheet through the use of computer programming. Information was extracted regarding IG gene repertoires, VH CDR3 length and amino acid sequence and SHM; to identify and cluster stereotyped rearrangements, we used an in-house purpose-built bioinformatics method. In brief VH CDR3 sequences were assigned to stereotyped subsets according to the following criteria: (i) Cases were initially clustered together only if they share at least 50% amino acid identity and 70% similarity within their

respective VH CDR3s; (ii) clustered sequences must have identical VH CDR3 lengths and identical locations of shared patterns; (iii) only sequences carrying IGHV genes of the same phylogenetic clan be placed in the same cluster. Iterative clustering ultimately leads to higher levels of hierarchy describing more distant, and thus relaxed, sequence relationships with more widely shared sequence patterns (affecting only the number - and rarely the location - of these patterns, but neither the VH CDR3 length nor the phylogenetic makeup of the cluster) in progressively larger clusters, which eventually form the collection of subsets.

CD38 expression

CD38 expression was assessed with flow-cytometry. The cut-off for positivity was 30%.

FISH analysis

Preparations for FISH analysis were counterstained with 4,6-diamidino-phenyl-indole (DAPI) and a minimum of 200 interphase nuclei were examined using commercially available probes for chromosomal bands 13q14-34, 11q22, 17p13 and chromosome 12.

Analysis of gene mutations

Mutational screening was performed for the following genes: *NOTCH1* (n=1229, 90%): entire exon 34 or targeted analysis for del7544-45/p.P2514Rfs*4; *TP53* (n=743, 54%): exons 4-8 but also exons 9-10 for some centers; *SF3B1* (n=840, 62%): exons 14-16 and, *MYD88* (n=506, 37%): exons 3 and 5 or targeted analysis for p.L265P.

Statistical analysis

Proportional hazard regression

The proportional hazard (PH) assumption was assessed for both the univariable and multivariable case, using the function *cox.zph*(), which correlates for each covariate the corresponding set of scaled Schoenfeld residuals with time, to test for independence between residuals and time. In the multivariable case it also performs a global test for the model as a whole. Within early stage M-CLL patients, from the six variables that were included as predictors in the multivariable model, only Male did not satisfy the PH assumption in both cases. The global test indicated that the PH assumption was satisfied. Within early stage U-CLL patients, all five variables that were included as predictors in the multivariable model satisfied the PH assumption in both cases. The global test also indicated that the PH assumption was satisfied. Harrell's C-index and its standard error were calculated to assess the discriminatory ability of the Cox model within both early stage M-CLL and U-CLL patients for two scenarios. The first is when a multivariable model included as predictors the important factors according to the univariable model and was based on Binet A cases. The second is when a univariable model included as sole predictor the final prognostic index (four and five categories respectively in M-CLL and U-CLL) and was based on all cases.

Internal validation

A bootstrapping procedure was applied separately for M-CLL and U-CLL cases to validate internally the stability of the multivariable Cox model[5]. Initially, 1000 bootstrap samples equal in size to the original CLL population were randomly generated with replacement from the original CLL population. Subsequently, for each bootstrap sample, the multivariable Cox regression model was applied using the same predictors as in the original model. For each predictor, the percentage of cases it was considered statistically significant and included in the model was recorded, as well as the average number of significant predictors per bootstrap sample. A prognostically important predictor would be expected to be included in the multivariable model in the majority of bootstrap samples. In a subsequent step, 1000 additional bootstrap samples were randomly generated following the same procedure. The multivariable Cox model was applied to each bootstrap sample with the same predictors as in the original modeling. The mean of the hazard ratio and the respective 95% confidence interval were recorded for each predictor based on the 1000 bootstrap samples.

The first step of internal validation within early stage M-CLL patients showed that within the 1000 randomly generated bootstrap samples, the average number of predictors included in the multivariable Cox model was 3.2; three variables exhibited selection percentages greater than 60%, i.e. *TP53*abn, +12 and subset #2. Therefore, we argued that these three predictors were the most important according to Cox regression analysis. The first step of internal validation within early stage U-CLL patients, showed that the average number of predictors considered significant in the multivariable Cox model was 3.5. Four variables exhibited selection percentages greater than 60%, i.e. *TP53*abn, *SF3B1*mut, del(11q) and male sex. Therefore, we argued that these four covariates were the most important within this mutational group.

Binary Recursive partitioning

Recursive partitioning was performed using tree-structured regression models that describe the conditional distribution of TTFT given the same predictors that were included in the multivariable Cox model.

We followed Hothorn et al. who proposed a recursive binary partitioning approach within a theoretically structured conditional inference framework[6]. The algorithm used for the partitioning is briefly described below:

1. Test the global null hypothesis of independence between TTFT and any of the covariates. Stop if this hypothesis cannot be rejected. Otherwise, select the covariate which exhibits the strongest association to TTFT.
2. Once the best covariate is selected, the optimal binary split is determined.
3. Recursively repeat steps 1 and 2.

This procedure enables the hierarchical classification of the significant covariates, from the most important, which splits the primary node (entire population), to those which extend to the terminal nodes. By separating the covariate selection and following the splitting

procedure algorithm, the two main problems when fitting such models are addressed, i.e. overfitting and selection bias towards covariates with many possible splits or missing values. The *ctree* function from the package *party* in *R* was applied. Notable parameters that control aspects of the tree construction are: (i) the minimum sum of patients in a node in order to be considered for splitting (set to 20); (ii) the minimum sum of patients in a terminal node (set to 9); and (iii) the split criterion according to a log-rank scores-based statistic, which was set as $p<0.05$.

Amalgamation

Any two nodes within the tree that arise from the same parent node exhibit significantly different survival behavior. This is not the case for each pair of the terminal nodes that do not share the same parent. Therefore, an amalgamation algorithm is applied to merge terminal nodes that exhibit similar survival behavior[7]. At first, the log-rank test is applied, for each different pair of terminal nodes, to test the null hypothesis that their survival distributions are the same, against the alternative that they differ. The p-value is recorded for each comparison and the maximum p-value of all possible comparisons is considered. When the latter is greater than 0.05, the corresponding nodes are merged to a new terminal node and the procedure is repeated until the maximum p-value is less than 0.05.

Supplementary Table 1. Main clinicobiological features of the entire cohort (n=2366).

	Entire cohort n, %	M-CLL n, %	U-CLL n, %	X^2 test p-value
Clinical stage				
Binet A	1900/2366, 80%	1224/1364, 90%	676/1002, 68%	<0.001
Binet B	287/2366, 12%	83/1364, 6%	204/1002, 20%	<0.001
Binet C	179/2366, 8%	57/1364, 4%	122/1002, 12%	<0.001
Gender				
Male	1449/2366, 61%	806/1364, 59%	643/1002, 64%	0.233
Age at Diagnosis				
Median age (years)	64.3 (22-92)	63.7 (22-91)	63.5 (26-92)	0.575
CD38 expression				
High	293/1649, 18%	167/1319, 13%	126/330, 38%	<0.001
FISH detected abnormalities				
idel(13q)	671/1373, 49%	503/1013, 50%	168/360, 47%	0.603
Trisomy 12	263/1798, 15%	114/1043, 11%	149/755, 20%	<0.001
del(11q)	220/1813, 12%	35/1047, 3.3%	185/766, 24%	<0.001
del(17p)	114/1825, 6%	34/1057, 3.2%	80/768, 10.5%	<0.001
Recurrent gene mutations				
MYD88	21/929, 2.2%	21/506, 4.1%	0/423, 0%	<0.001
NOTCH1	166/2097, 8%	22/1229, 1.8%	144/868, 16.5%	<0.001
SF3B1	115/1449, 8%	31/840, 3.7%	84/609, 14%	<0.001
TP53	137/1535, 9%	42/743, 5.6%	95/801, 12%	<0.001
BIRC3	24/830, 3%	7/458, 1%	17/372, 5%	0.020
***TP53*abn**	183/2095, 9%	55/1154, 4.8%	128/941, 14%	<0.001
Immunogenetic features				
GI: 97-97.99%	104/2366, 4.4%	104/1364, 7.6%	-	-
GI: 100%	750/2366, 32%	-	750/1002, 75%	-
Stereotyped #1	55/2366, 2.3%	-	55/1002, 5.5%	-
Stereotyped #2	33/2366, 1.5%	27/1364, 2%	6/1002, 0.6%	0.009
Stereotyped #4	35/2366, 1.5%	35/1364, 2.6%	-	

High CD38 expression: positivity >30%, idel(13q): isolated deletion of chromosome 13q, del(11q): deletion of chromosome 11q, del(17p): deletion of chromosome 17p, *TP53*abn: deletion of chromosome 17p (del(17p)) and/or *TP53* mutation, GI: germline identity, Stereotyped #2: assignment to stereotyped subset #2, stereotyped #4: assignment to stereotyped subset #4, M-CLL: patients carrying mutated IGHV genes, U-CLL: cases carrying unmutated IGHV genes.

Supplementary Table 2. Main clinicobiological features of the validation cohort. The p-value stems from the comparison of each biomarker between the main and the validation cohort.

	n, %	p-value (X^2 test)
Gender		
Male	397/649, 62%	1
Age at diagnosis		
Median age	63.6 (29-89) years	0.630
FISH detected abnormalities		
idel(13q)	301/522, 58%	0.061
Trisomy 12	65/598, 11%	0.049
del(11q)	59/598, 10%	0.203
del(17p)	13/598, 2%	<0.001
Recurrent gene mutations		
SF3B1	28/553, 5%	0.046
TP53	27/623, 4%	<0.001
***TP53*abn**	29/632, 5%	0.002
Immunogenetic features		
M-CLL	442/649, 68%	0.020
Stereotyped #2	20/649, 3%	0.008

idel(13q): isolated deletion of chromosome 13q, del(11q): deletion of chromosome 11q, del(17p): deletion of chromosome 17p, *TP53*abn: deletion of chromosome 17p (del(17p)) and/or *TP53* mutation, Stereotyped #2: assignment to stereotyped subset #2, M-CLL: patients carrying mutated IGHV genes.

Supplementary Table 3. Overview of the entire cohort and the subgroup of cases included in the multivariable analysis. No significant difference is detected regarding the evaluated biomarkers.

Feature	Entire cohort, M-CLL n, %	Cases included in the multivariable -analysis, M-CLL; n, %	p-value
Gender			
Male	711/1224, 58%	530/919, 58	0.84
Age at diagnosis			
Median (years)	64.6	64.5	0.99
CD38 expression			
High	127/1186, 11%	97/864, 11%	0.71
FISH detected abnormalities			
Idel(13q)	432/731, 59%	432/731, 59%	-
Trisomy 12	100/972, 10.3%	119/919, 12.5%	0.07
del(11q)	23/939, 2.4%	34/904, 3.5%	0.1
Recurrent gene mutations			
MYD88	16/436, 3.7%	14/435, 3.2%	0.71
NOTCH1	15/1104, 1.3%	12/902, 1.3%	0.95
SF3B1	20/750, 2.7%	22/686, 3.2%	0.54
***TP53*abn**	46/1035, 4.4%	36/919, 4%	0.56
Immunogenetic features			
GI: 97-97.99%	80/1224, 6.5%	57/919, 6.2%	0.75
Stereotyped #2	13/1224, 1%	12/919, 1.3	0.60
	Entire cohort, M-CLL n, %	**Cases included in the multivariable analysis, U-CLL; n, %**	**p-value**
Gender			
Male	406/676, 61%	233/384, 61%	0.84
Age at diagnosis			0.79
Median (years)	64.4	64.5	
FISH detected abnormalities			
idel(13q)	125/266, 47%	97/194, 50%	0.52
Trisomy 12	108/523, 21%	80/384, 21%	0.94
del(11q)	118/529, 22%	89/384, 23%	0.75
Recurrent gene mutations			
NOTCH1	88/587, 15%	69/380, 18%	0.19
SF3B1	45/416, 11%	43/384, 12%	0.86
BIRC3	9/241, 4%	9/236, 3.8%	0.96
***TP53*abn**	73/636, 12%	36/384, 9.5%	0.29
Immunogenetic features			
GI: 100%	502/676, 74%	284/384, 74%	0.91
Stereotyped #1	34/555, 6%	17/303, 5.5%	0.26

idel(13q): isolated deletion of chromosome 13q, del(11q): deletion of chromosome 11q, del(17p): deletion of chromosome 17p, *TP53*abn: deletion of chromosome 17p (del(17p)) and/or *TP53* mutation, Stereotyped #2: assignment to stereotyped subset #2, M-CLL:

patients carrying mutated IGHV genes, U-CLL: patients carrying unmutated IGHV genes, MV-analysis: multivariable analysis.

Supplementary Table 4. Main clinicobiological features of M-CLL cases assigned to different Binet clinical stages. Bonferroni correction was applied and the significance level was set at p < 0.017.

	Binet A n=1224	**Binet B n=83**	**Binet C n=57**	**p-value A vs B**	**p-value A vs C**	**p-value B vs C**
Male	711/1224, 58%	58/83, 70%	37/57, 65%	0.034	0.31	0.53
CD38 expression	127/1186, 11%	22/78, 28%	18/55, 32.8%	<0.0001	<0.0001	0.57
NOTCH1	15/1104, 1.3%	4/75, 5.3%	3/50, 6%	0.008	0.009	0.87
MYD88	16/436, 3.7%	3/41, 7.3%	2/29, 6.8%	0.25	0.38	0.94
SF3B1	20/750, 2.7%	6/53, 11.3%	5/37, 13.5%	0.0005	0.0002	0.75
idel(13q)	459/906, 51%	27/64, 42%	17/43, 40%	0.2	0.15	0.78
+12	100/972, 10.3%	8/66, 12%	10/44, 22.7%	0.63	0.009	0.14
del(11q)	23/939, 2.4%	7/65, 10.7%	5/43, 11.6%	0.0001	0.0004	0.88
***TP53*abn**	46/1035, 4.4%	5/72, 7%	4/47, 8.5%	0.32	0.19	0.75
GI: 97-97.99%	80/1224, 6.5%	13/83, 15.7%	11/57, 19.2%	0.001	0.0002	0.57
Subset #2	13/1224, 1%	8/83, 9.6%	6/57, 10.5%	<0.0001	<0.0001	0.86

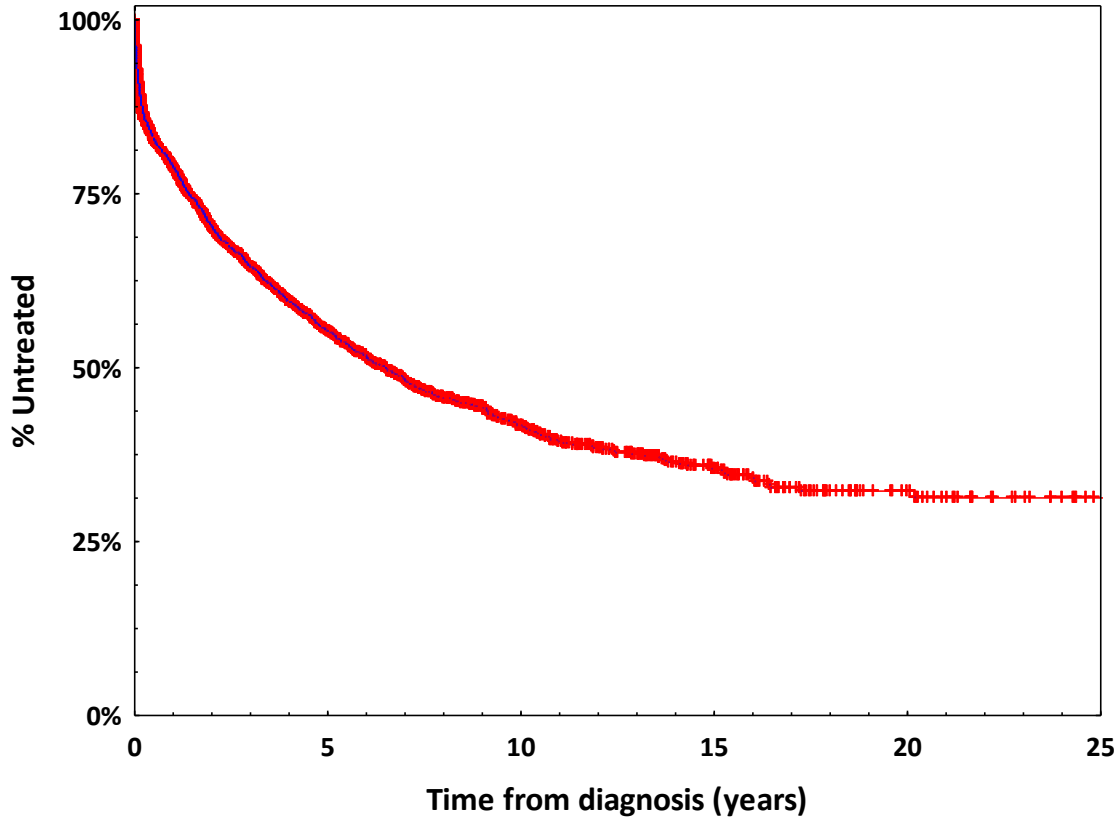
CD38 expression: cut-off for positivity >30%, idel(13q): isolated deletion of chromosome 13q, +12: trisomy 12, del(11q): deletion of chromosome 11q, *TP53*abn: deletion of chromosome 17p [del(17p)] and/or *TP53* mutation, GI: germline identity, subset #2: assignment to stereotyped subset #2.

Supplementary Table 5. Main clinicobiological features of U-CLL cases assigned to different Binet clinical stages. Bonferroni correction was applied and the significance level was set at $p < 0.017$.

	Binet A n=676	**Binet B n=204**	**Binet C n=122**	**p-value A vs B**	**p-value A vs C**	**p-value B vs C**
Male	406/676, 61%	148/204, 73%	89/122, 73%	0.001	0.006	0.93
CD38 expression	78/213, 37%	36/82, 44%	12/35, 34%	0.24	0.79	0.33
NOTCH1	88/587, 15%	33/174, 19%	23/107, 21%	0.23	0.09	0.61
BIRC3	9/241, 4%	4/71, 6%	4/60, 7%	0.48	0.31	0.81
SF3B1	45/416, 11%	25/118, 21%	14/75, 19%	0.003	0.05	0.67
idel(13q)	125/266, 47%	24/54, 44%	19/40, 48%	0.73	0.95	0.91
+12	108/523, 21%	29/138, 21%	12/94, 13%	0.92	0.075	0.11
del(11q)	118/529, 22%	42/143, 29%	25/94, 27%	0.078	0.36	0.64
***TP53*abn**	73/636, 12%	28/188, 15%	27/117, 23%	0.21	0.0007	0.07
GI: 100%	502/676, 74%	152/204, 75%	96/122, 79%	0.94	0.29	0.39
IGHV1-69	160/676, 24%	54/204, 26%	28/122, 23%	0.41	0.86	0.48
Subset #1	34/555, 6%	12/177, 7%	9/115, 8%	0.75	0.49	0.73

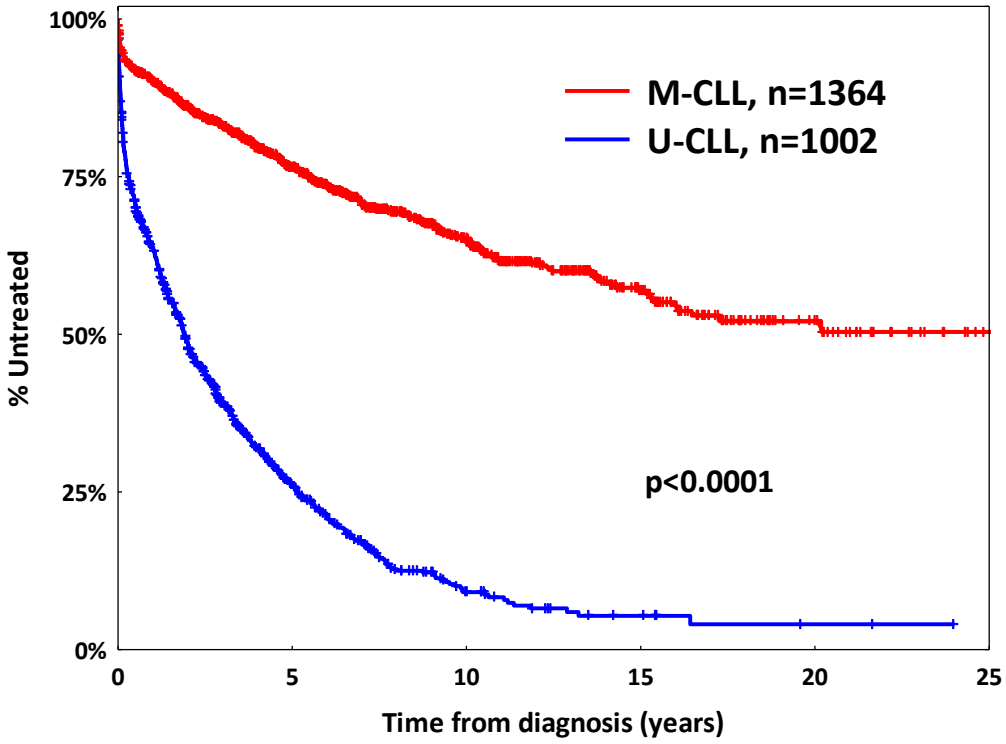
CD38 expression: positivity >30%, idel(13q): isolated deletion of chromosome 13q, +12: trisomy 12, del(11q): deletion of chromosome 11q, *TP53*abn: deletion of chromosome 17p [del(17p)] and/or *TP53* mutation, GI: germline identity, subset #1: assignment to stereotyped subset #1.

Supplementary Figure 1. Kaplan Meier curve for time-to-first-treatment (TTFT) for the entire cohort (n=2366).

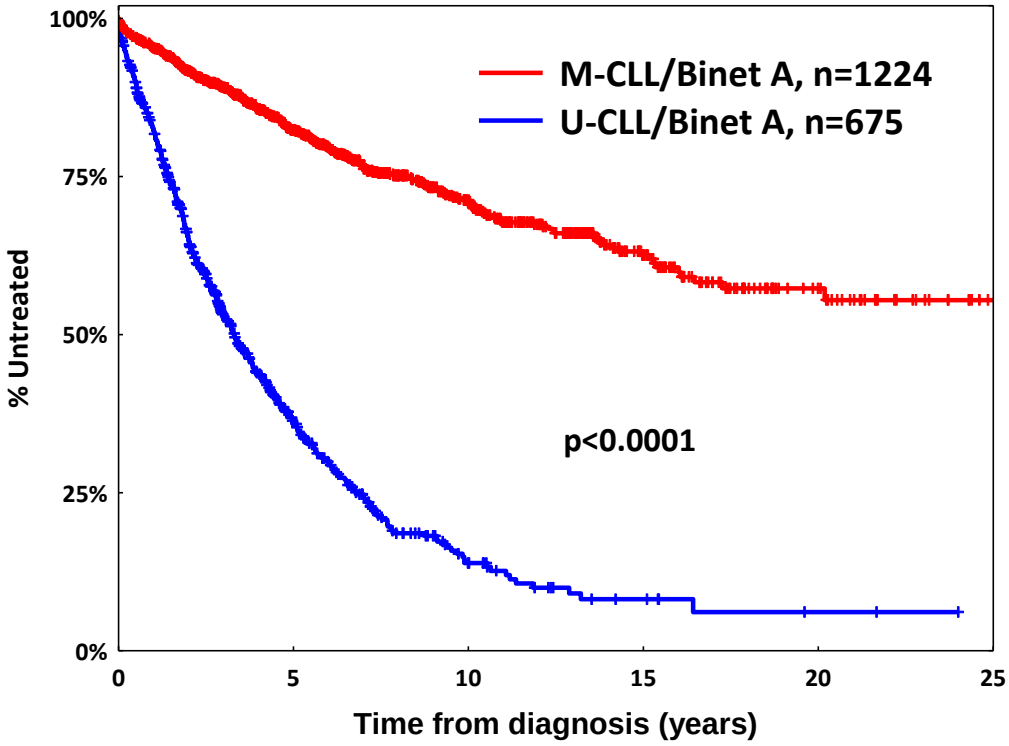


Supplementary Figure 2. (A) Kaplan Meier curves for time-to-first-treatment (TTFT) for M-CLL and U-CLL. (B) Kaplan Meier curves for TTFT for Binet A M-CLL and U-CLL.

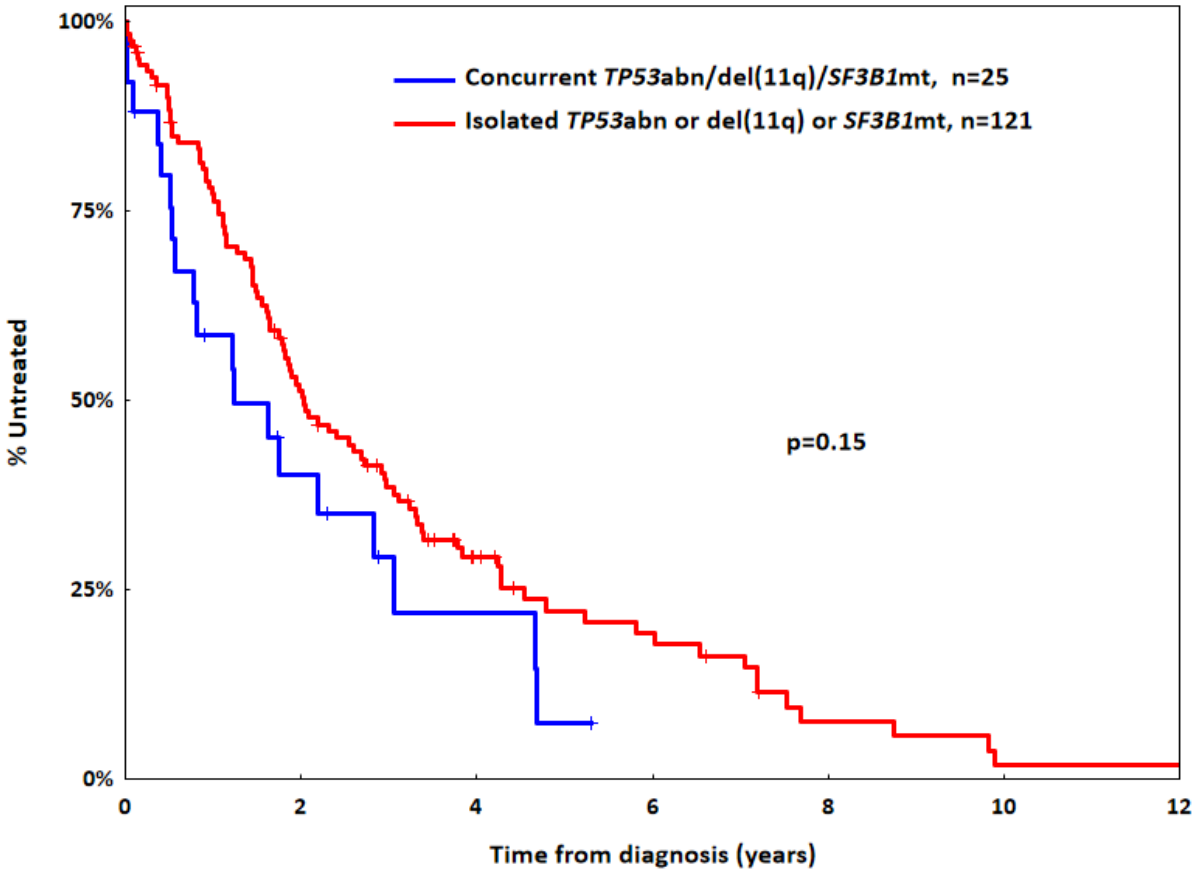
2A



2B

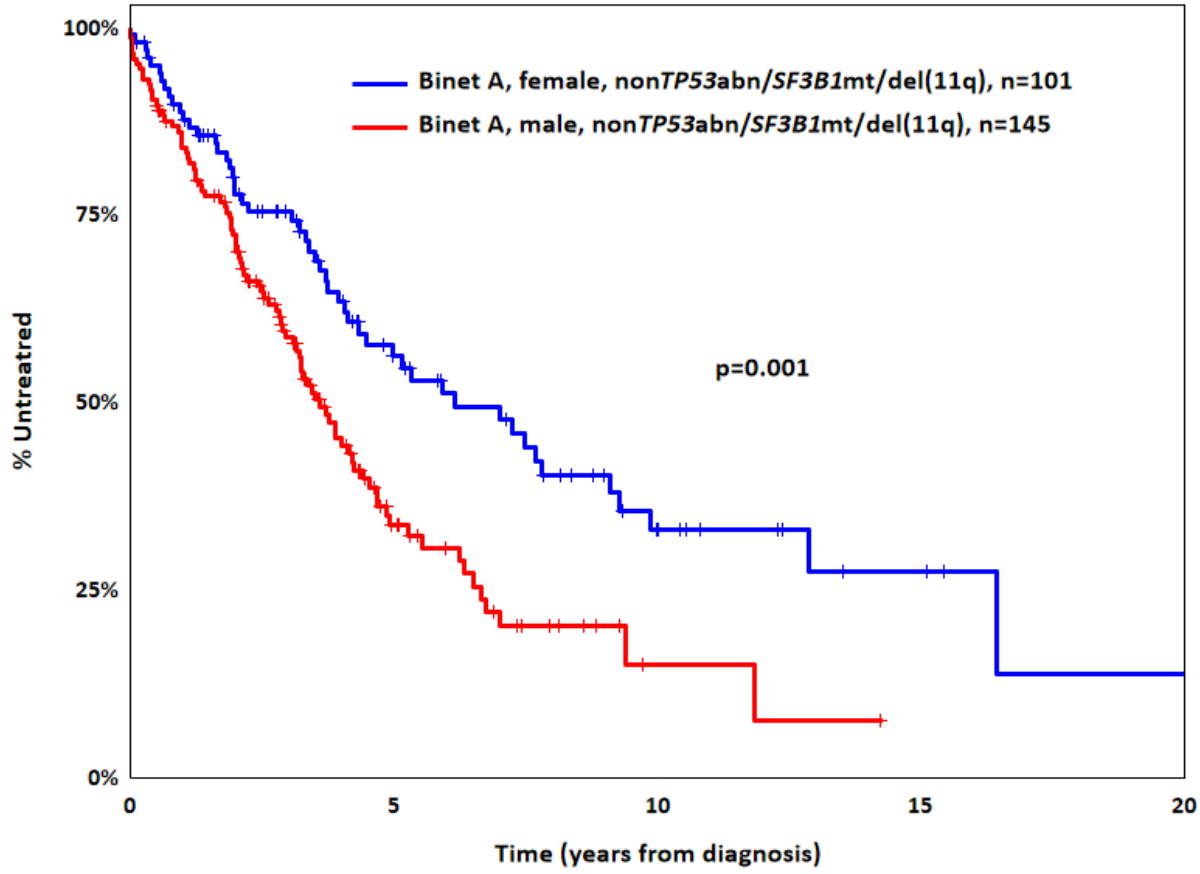

100%
75%
50%
25%
0%
% Untreated
0
5
10
15
20
25
Time from diagnosis (years)
M-CLL/Binet A, n=1224
U-CLL/Binet A, n=675
p<0.0001

Supplementary Figure 3. Kaplan Meier curves for time-to-first-treatment (TTFT) for Binet A U-CLL cases with either isolated *TP53*abn, *SF3B1* mutations and del(11q) vs cases with 2 concurrent of the mentioned aberrations

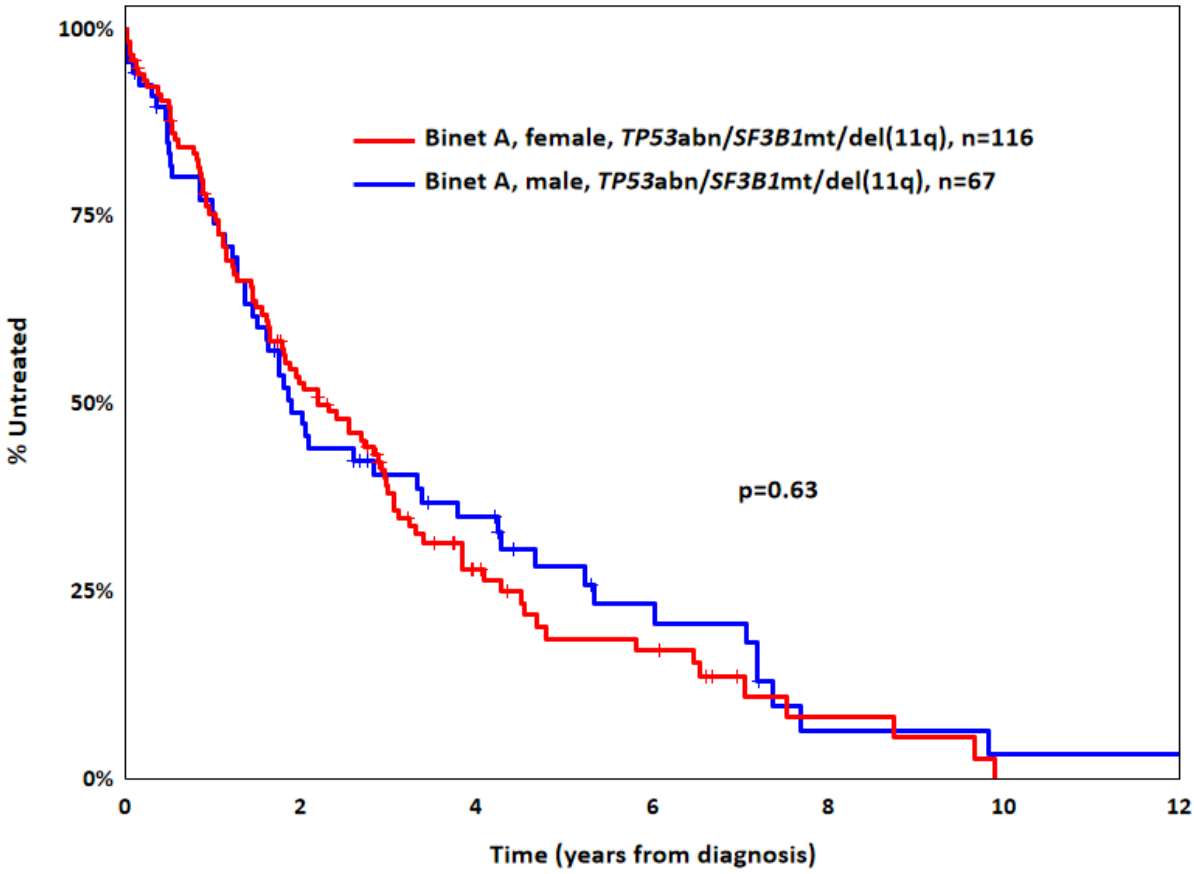


Supplementary Figure 4. Kaplan Meier curves for time-to-first-treatment (TTFT) in Binet A U-CLL. (A): Male sex is correlated with shorter TTFT within the non *TP53*abn/*SF3B1*mut/del(11q) Binet A. (B): No impact of male sex within the *TP53*abn/*SF3B1*mut/del(11q) Binet A cases.

4A

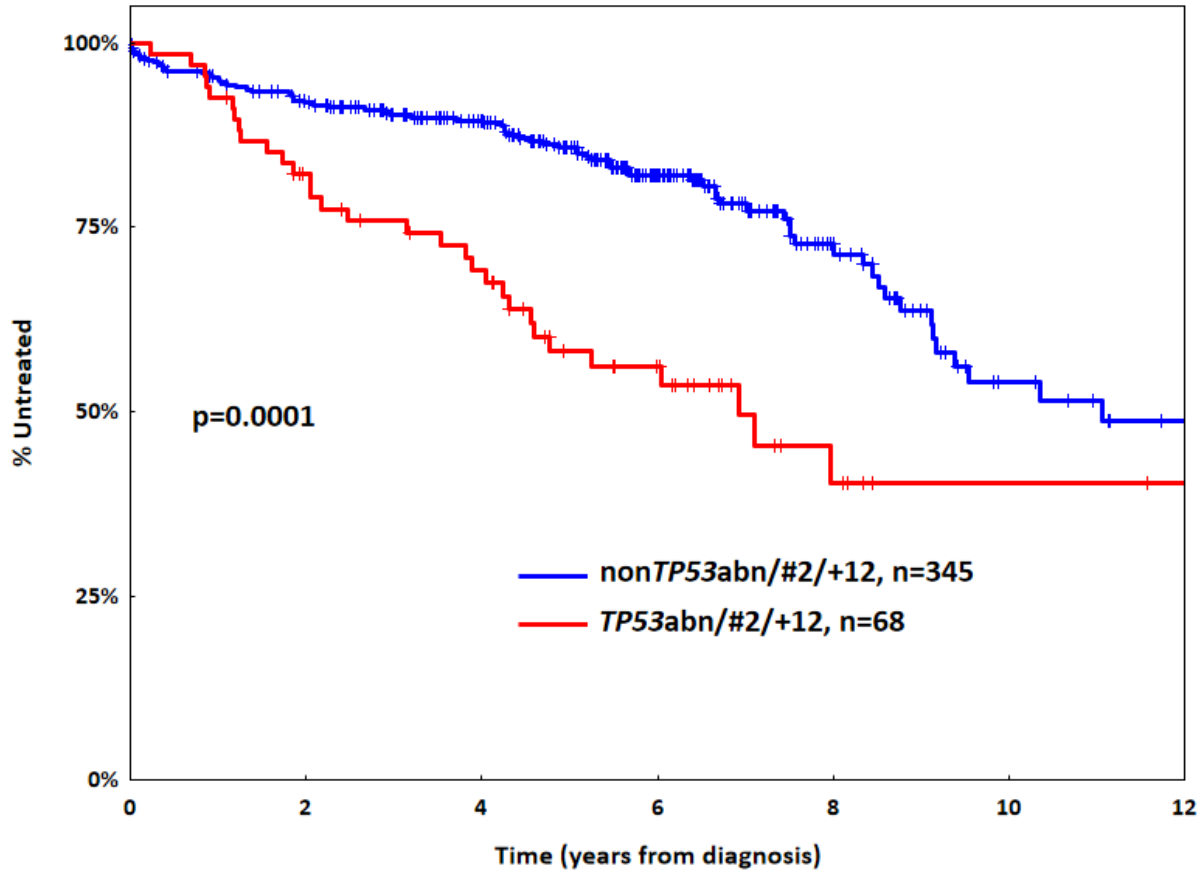


4B

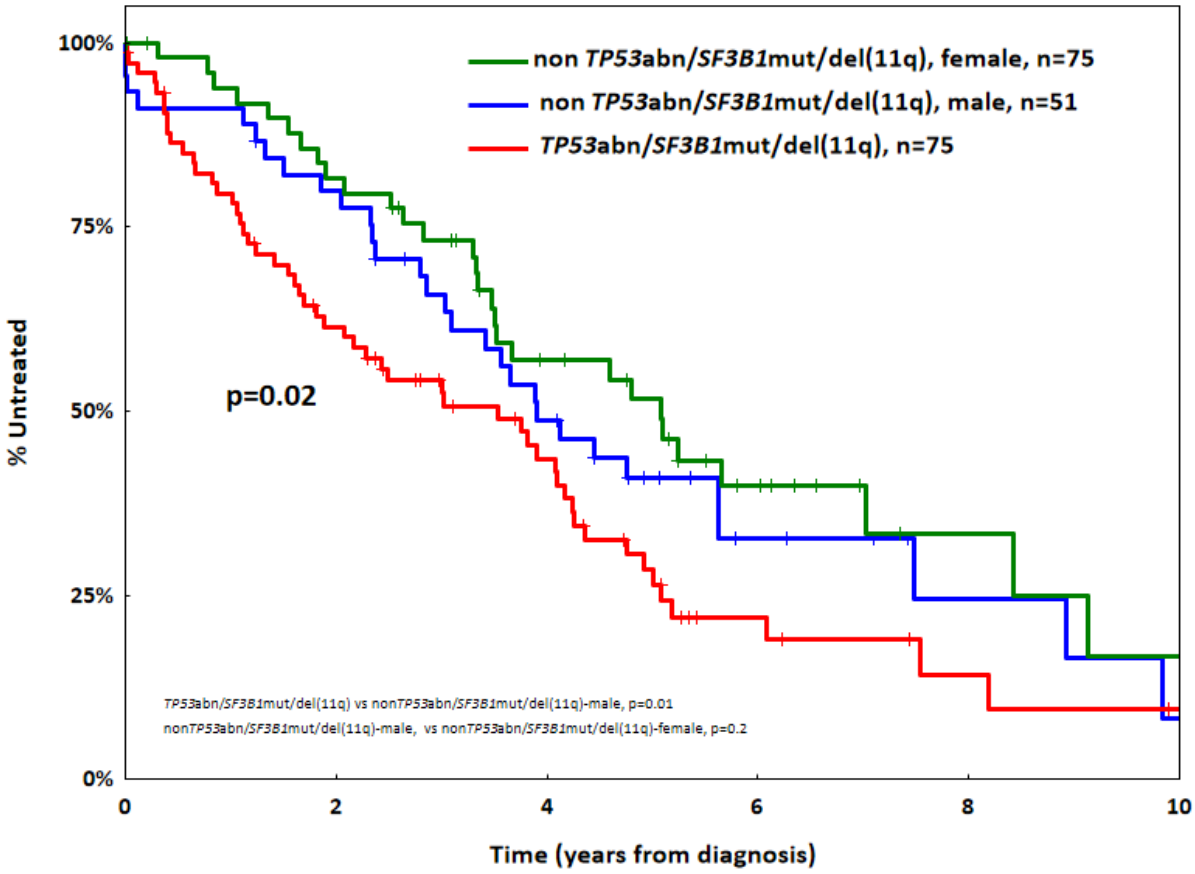

100%
75%
50%
25%
0%
% Untreated
Binet A, female, TP53abn/SF3B1mt/del(11q), n=116
Binet A, male, TP53abn/SF3B1mt/del(11q), n=67
p=0.63
0
2
4
6
8
10
12
Time (years from diagnosis)

Supplementary Figure 5. Kaplan Meier curves for time-to-first-treatment (TTFT) in the validation cohort (n=649). (A) M-CLL cases carrying *TP53*abn, trisomy 12 (+12) or assigned to stereotyped subset #2, display shorter TTFT compared to non *TP53*abn/#2/+12 cases; (B) Within Binet A U-CLL, *TP53*abn/*SF3B1*mut/del(11q) cases exhibit the shortest TTFT. Within the remaining cases the difference between male and female patients does not exhibit statistical significance.

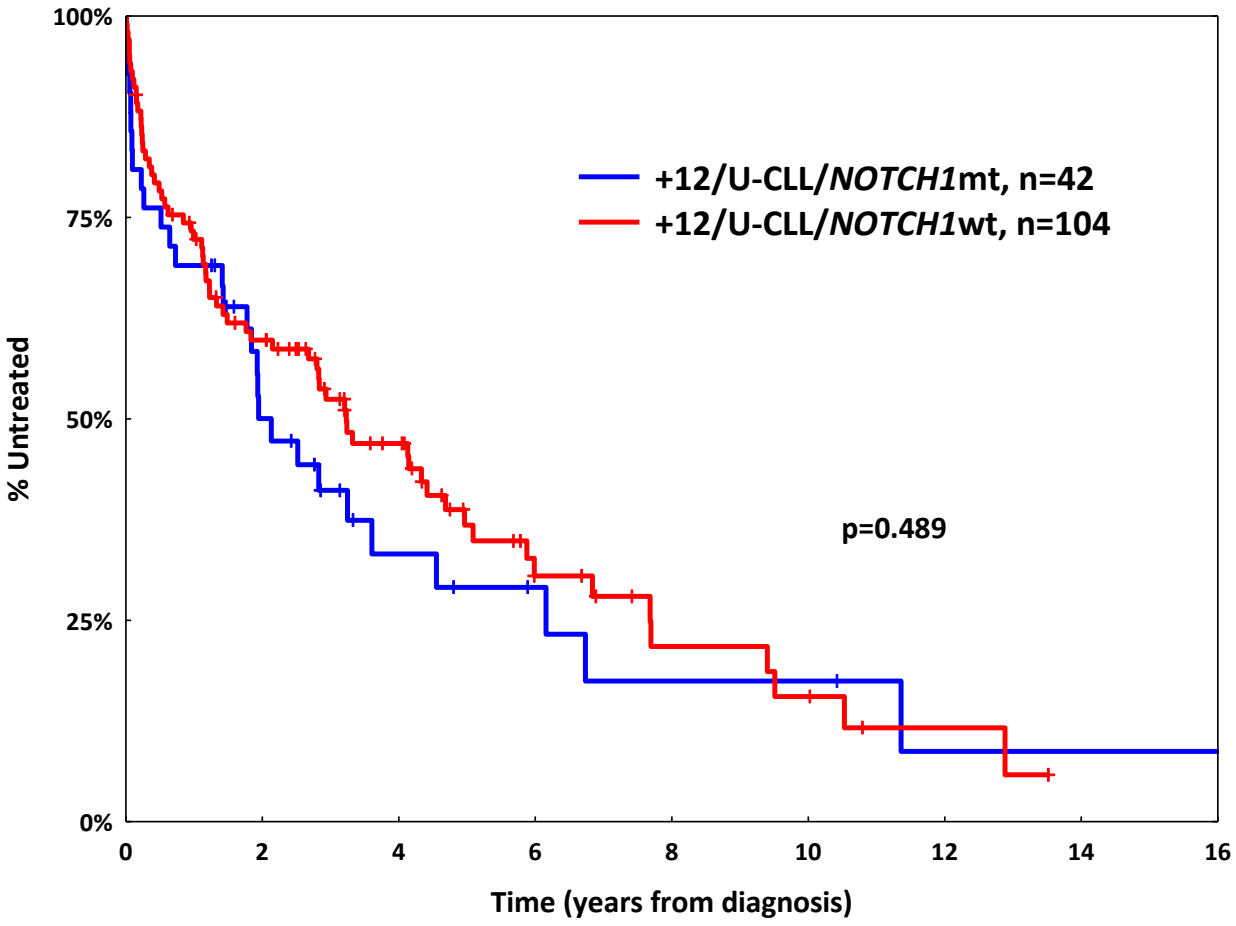
5A



5B


non TP53abn/SF3B1mut/del(11q), female, n=75
non TP53abn/SF3B1mut/del(11q), male, n=51
TP53abn/SF3B1mut/del(11q), n=75
p=0.02
TP53abn/SF3B1mut/del(11q) vs nonTP53abn/SF3B1mut/del(11q)-male, p=0.01
nonTP53abn/SF3B1mut/del(11q)-male, vs nonTP53abn/SF3B1mut/del(11q)-female, p=0.2
% Untreated
100%
75%
50%
25%
0%
0
2
4
6
8
10
Time (years from diagnosis)

Supplementary Figure 6. Kaplan Meier curves for time-to-first-treatment (TTFT) within cases carrying trisomy 12 (+12). No impact of *NOTCH1* mutation in cases with unmutated IGHV genes (U-CLL).



References

1.Agathangelidis A, Darzentas N, Hadzidimitriou A, et al: Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. Blood 119:4467-75, 2012

2.Ghia P, Stamatopoulos K, Belessi C, et al: ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia 21:1-3, 2007

3.Langerak AW, Davi F, Ghia P, et al: Immunoglobulin sequence analysis and prognostication in CLL: guidelines from the ERIC review board for reliable interpretation of problematic cases. Leukemia 25:979-984, 2011

4.Murray F, Darzentas N, Hadzidimitriou A, et al: Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. Blood 111:1524-33, 2008

5.Chen CH, George SL: The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. Stat Med 4:39-46, 1985

6.Hothorn T, Hornik K, Zeileis A: Unbaised Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics 15:651-674, 2006

7.Rossi D, Rasi S, Spina V, et al: Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. Blood 121:1403-12, 2013