NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL

Ilaria Del Giudice,[1]* Davide Rossi,[2]* Sabina Chiaretti,[1]* Marilisa Marinelli,[1] Simona Tavolaro,[1] Sara Gabrielli,[1] Luca Laurenti,[3] Roberto Marasca,[4] Silvia Rasi,[2] Marco Fangazio,[2] Anna Guarini,[1] Gianluca Gaidano,[2]§ and Robin Foà[1]§

[1]Division of Hematology, Department of Cellular Biotechnologies and Haematology, University "Sapienza", Rome; [2]Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara; [3]Institute of Hematology, Catholic University of the Sacred Heart, Rome; and [4]Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy

Citation: Del Giudice I, Rossi D, Chiaretti S, Marinelli M, Tavolaro S, Gabrielli S, Laurenti L, Marasca R, Rasi S, Fangazio M, Guarini A, Gaidano G, and Foà R. NOTCH1 *mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. Haematologica 2012;97(3):437-441. doi:10.3324/haematol.2011.060129*

Online Supplementary Design and Methods

Mutation analysis of NOTCH1

The *NOTCH1* (exon 34; RefSeq NM_017617.2) mutation hotspot previously identified[1] in CLL was analyzed in blind with respect to clinical data by PCR amplification and direct sequencing of genomic DNA extracted from fresh or frozen peripheral blood mononuclear cells (PBMC) obtained at presentation. Purified amplicons were subjected to conventional DNA Sanger sequencing using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, CA, USA). Sequences were compared to the corresponding germline RefSeq sequences using the Mutation Surveyor Version 2.41 software package (SoftGenetics, State College, PA, USA) after both automated and manual curation. All variants were sequenced from both strands on independent PCR products. Synonymous mutations, reported polymorphisms and changes present in the matched normal DNA were removed from the analysis. All PCR primers and conditions are available upon request.

The presence of *NOTCH1* c.7544_7545delCT alleles was also investigated by ARMS PCR approach. ARMS PCR was performed utilizing two forward (For) and one common reverse primers (Rev). The first forward primer (ForMUT) is specific for the mutant allele and yields a 183 bp amplimer. The second forward primer (ForC) amplifies a 283 bp product from both mutant and wild-type alleles and serves as an internal PCR control. Sequences of PCR primers are: ForC, 5'-GTGACCGCAGCCCAGTT-3'; ForMUT, 5'-TCCTCACCCCGTCCCGA-3'; Rev, 5'-AAGGCTTGGGAAAGGAAGC-3'. ARMS was performed with the Go Taq Hot Start polymerase (Promega, Madison, WI, USA) on 25 ng of genomic DNA under the following conditions: denaturing step at 95°C for 3 min followed by 30 cycles at 95°C (30 sec per cycle), annealing step at 57°C (40 sec per cycle), and extension at 72°C (40 sec per cycle), in a BioRad C1000 thermal cycler. The final concentrations of the ForC, ForMUT and the Rev primers were 0.1 μM, 0.4 μM and 0.5 μM, respectively, in a final volume of 30 microliters. Products were resolved on 2% agarose gels by electrophoresis and visualized after staining with ethidium bromide. Based on a dilution curve, ARMS is capable of detecting a mutation present in at least 10% of the alleles.

Statistical analysis

Overall survival (OS) was measured from the date of initial presentation to the date of death (event) or last follow up (censoring). Survival analysis was performed by the Kaplan-Meier method. The crude association between exposure variables and outcome was estimated by univariate Cox's regression analysis. The independence of *NOTCH1* mutations as a predictor of OS in CLL harboring +12 was estimated after controlling for confounding variables by multivariate Cox's regression analysis. None of the covariates violated the proportional hazard assumption as documented by plotting the smoothed Schoenfeld residuals, and by performing a correlation test between time and residuals. The assumption of effect additivity of predictors was not violated, as documented by a global test of additivity including interactions between *NOTCH1* mutations and other covariates. None of the covariates showed colinearity. Model discrimination and calibration was assessed to provide a bias-corrected estimate of prediction accuracy and to protect against overfitting. Recursive-partitioning analysis for censored survival data was performed to hierarchically classify CLL patients into risk categories based on *NOTCH1* mutations, 13q14 deletion, 11q22-q23 deletion and 17p13 deletion. Categorical variables were compared by χ^2 test and Fisher's exact test when appropriate. Mutual information between *NOTCH1* mutations and presence or absence of additional cytogenetic abnormalities among +12 CLL was estimated according to the entropy estimator method. All statistical tests were two-sided. Statistical significance was defined as $P<0.05$. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.18.0 (Chicago, IL, USA) and with R statistical package 2.13.0 (*http://www.r-project.org*).

RNA extraction, microarray preparation and data analysis

Total RNA was extracted from fresh leukemia cells using the RNeasy mini procedure (Qiagen, Hilden, Germany) according to the manufacturer's instructions with minor modifications. No further purification was performed since the percentage of leukemic cells was greater than 90%.

For oligonucleotide array analysis, the HGU133 Plus 2.0 gene chips (Affymetrix, Santa Clara, CA, USA) were used. Sample preparation and microarray processing were performed as pre-

viously described.[2] Raw data were previously reported in the National Center for Biotechnology Information's Gene Expression Omnibus database (*http://www.ncbi.nlm.nih.gov/geo/*) under series accession number GSE13204.[3]

Gene expression data were analyzed with the dChip software (*http://www.dchip.org*). To identify differentially expressed genes between patients harboring the *NOTCH1* ΔCT (c.7544_7545delCT) mutation *versus NOTCH1* wild-type patients, based on the ARMS approach, a t-test was applied with the following criteria: average expression ≥100 in at least one group, $P<0.05$ and fold change ≥1.5.

Functional annotation analysis was performed using the DAVID database (*http://david.abcc.ncifcrf.gov*).

References

1. Rossi D, Rasi S, Fabbri G, Spina V, Fangazio M, Forconi F, et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. Blood. 2012;119(2):521-9.
2. Chiaretti S, Tavolaro S, Marinelli M, Messina M, Del Giudice I, Mauro FR, et al. Evaluation of TP53 mutations with the AmpliChip p53 research test in chronic lymphocytic leukemia: correlation with clinical outcome and gene expression profiling. Genes Chromosomes Cancer. 2011;50(4):263-74.
3. Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Béné MC, et al. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. J Clin Oncol. 2010;28(15):2529-37.

Online Supplementary Table S1. Clinical and biological characteristics of the CLL patients analyzed by gene expression profile.

Sample ID	Gender	Age at diagnosis	IGHV status	*TP53* status	ZAP-70 expression	CD38%	FISH analysis
***NOTCH1* ΔCT mutated CLL**							
CLL_00026	F	49	UM	wt	N.E.	69	+12 (62%)
CLL_00255	M	57	UM	wt	0	65	+12 (51%)
CLL_00269	F	43	UM	wt	0	10	+12 (61%)
CLL_00276	M	49	UM	wt	32	8	+12 (74%)
CLL_00355	M	54	UM	wt	51	10	+12 (57%)
CLL_00370	M	49	UM	wt	0	21	+12 (58%)
CLL_00435	F	48	M	wt	6	62	+12 (80%)
***NOTCH1* ΔCT wt CLL**							
CLL_00036	F	39	M	wt	0	0	+12 (10%); del(11q) (5%)
CLL_00245	F	55	M	wt	3	1	+12 (74%); del(13q) (89%)
CLL_00248*	M	45	M	wt	88	86	+12 (60%); del(13q) (83%)
CLL_00358	M	55	UM	wt	46	88	+12 (58%)
CLL_00361	F	47	M	wt	0	1	+12 (46%)
CLL_00394	F	N.A.	UM	wt	26	82	+12 (60%)
CLL_00441	M	30	UM	wt	32	78	+12 (47%); del(13q) (15%); del(6q) (20%)

*M: mutated; UM: unmutated; N.E.: not evaluated; N.A.: not available. *This patient later acquired a TP53 mutation and developed a myelodysplastic syndrome.*

Online Supplementary Table S2. Differentially expressed genes between *NOTCH1* ΔCT mutated and *NOTCH1* ΔCT wt CLL. Genes are rank-ordered according to their chromosomal location.

Probeset	Gene symbol	Chromosomal location	*P* value	Fold change	Gene function	Expression in *NOTCH1* mutated CLL
222482_at	*SSBP3*	1p32.3	0.034075	2.75	Regulation of transcription	High
221078_s_at	*CCDC88A*	2p16.1	0.005693	1.79	Regulation of DNA replication	High
219387_at	*CCDC88A*	2p16.1	0.01884	1.67	Regulation of DNA replication	High
212350_at	*TBC1D1*	4p14	0.026777	1.52	Regulation of Rab GTPase activity	High
226735_at	*TAPT1*	4p15.32	0.017456	1.79	Multicellular organismal development	High
209110_s_at	*RGL2*	6p21.3	0.01825	1.62	Ras protein signal transduction	High
205859_at	*LY86*	6p25.1	0.022251	1.91	Apoptosis	High
1569974_x_at	*SEPT7P2*	7p12.3	0.021272	1.56	Unknown	High

continued on next page

continued from previous page

225081_s_at	*CDCA7L*	7p15.3	0.039691	1.68	Microtubule-based movement	High
216242_x_at	*POLR2J2*	7q22.1	0.038912	1.55	Transcription	High
204082_at	*PBX3*	9q33.3	0.033762	1.72	Regulation of transcription, DNA-dependent	High
220091_at	*SLC2A6*	9q34	0.007643	1.7	Transmembrane transport	High
635_s_at	*PPP2R5B*	11q12	0.00112	2.37	Signal transduction	High
204611_s_at	*PPP2R5B*	11q12	0.001456	2.24	Signal transduction	High
202645_s_at	*MEN1*	11q13	0.016521	1.83	Chromatin remodeling	High
208149_x_at	*DDX11*	12p11	0.014739	1.95	Cell cycle	High
208159_x_at	*DDX11*	12p11	0.015661	1.94	Cell cycle	High
226517_at	*BCAT1*	12p12.1	0.035518	2.51	Cell proliferation	High
225285_at	*BCAT1*	12p12.1	0.040839	1.77	Cell proliferation	High
218102_at	*DERA*	12p12.3	0.020246	1.76	Metabolic process	High
225958_at	*PHC1*	12p13	0.013816	1.86	Multicellular organismal development	High
200822_x_at	*TPI1*	12p13	0.017911	1.71	Carbohydrate metabolic process	High
201161_s_at	*CSDA*	12p13.1	0.009578	1.67	Regulation of transcription, DNA-dependent	High
226295_at	*ITFG2*	12p13.33	0.026071	1.81	Integrin-mediated signaling pathway	High
224906_at	*ANO6*	12q12	0.024548	2.02	Ion transport	High
203046_s_at	*TIMELESS*	12q12-q13	0.008725	2.06	Cell cycle	High
204173_at	*MYL6B*	12q13.13	0.015571	2.18	Muscle contraction	High
232543_x_at	*ARHGAP9*	12q14	0.015602	1.56	Signal transduction	High
224451_x_at	*ARHGAP9*	12q14	0.011538	1.54	Signal transduction	High
214085_x_at	*GLIPR1*	12q21.2	0.03418	2.06	Unknown	High
203795_s_at	*BCL7A*	12q24.13	0.032166	4.69	Negative regulation of transcription	High
221909_at	*RNFT2*	12q24.22	0.01115	2.89	Unknown	High
220137_at	*VSIG10*	12q24.23	0.002776	5	Unknown	High
208722_s_at	*ANAPC5*	12q24.31	0.032658	1.65	G2/M transition of mitotic cell cycle	High
200098_s_at	*ANAPC5*	12q24.31	0.038259	1.56	G2/M transition of mitotic cell cycle	High
211036_x_at	*ANAPC5*	12q24.31	0.034302	1.5	G2/M transition of mitotic cell cycle	High
204072_s_at	*FRY*	13q13.1	0.002288	2.59	Regulation of transcription	High
218352_at	*RCBTB1*	13q14	0.036476	1.57	Cell cycle	High
203940_s_at	*VASH1*	14q24.3	0.046779	2.67	Angiogenesis	High
215621_s_at	*IGHD*	14q32.33	0.032654	4.64	Immune response	High
213674_x_at	*IGHD*	14q32.33	0.036887	4.37	Immune response	High
209374_s_at	*IGHM*	14q32.33	0.034933	2.52	Immune response	High
212827_at	*IGHM*	14q32.33	0.046768	2.18	Immune response	High
202771_at	*FAM38A*	16q24.3	0.028396	1.79	Unknown	High
226143_at	*RAI1*	17p11.2	0.028611	1.6	Regulation of transcription from RNA polymerase II promoter	High
205414_s_at	*ARHGAP44*	17p12	0.011301	3.03	Signal transduction	High
224686_x_at	*LRRC37A2*	17q21.31	0.038497	1.65	Unknown	High
205594_at	*ZNF652*	17q21.32	0.037232	1.77	Regulation of transcription	High
213730_x_at	*TCF3*	19p13.3	0.017214	1.66	B-cell lineage commitment	High
213811_x_at	*TCF3*	19p13.3	0.021386	1.61	B-cell lineage commitment	High
210776_x_at	*TCF3*	19p13.3	0.010805	1.59	B-cell lineage commitment	High
224882_at	*ACSS1*	20p11.23-p11.21	0.033755	1.7	Acetyl-CoA biosynthetic process	High
208689_s_at	*RPN2*	20q12-q13.1	0.046071	1.57	Protein modification process	High
225744_at	*ZDHHC8*	22q11.21	0.018907	2.19	Behavior	High
201540_at	*FHL1*	Xq26	0.020768	3.4	Cell differentiation	High
221740_x_at	*LOC100506162*	Unknown	0.021473	1.86	Unknown	High
209728_at	Unknown	Unknown	0.04703	41.53	Unknown	High
1568983_a_at	Unknown	Unknown	0.004061	2.74	Unknown	High
235274_at	Unknown	Unknown	0.033678	1.86	Unknown	High
244429_at	Unknown	Unknown	0.033447	1.8	Unknown	High
230180_at	Unknown	Unknown	0.029016	1.74	Unknown	High

continued on next page

continued from previous page

239893_at	Unknown	Unknown	0.04423	1.62	Unknown	High
1557987_at	Unknown	Unknown	0.037781	1.58	Unknown	High
229028_s_at	Unknown	Unknown	0.028794	2.47	Unknown	High
205263_at	*BCL10*	1p22	0.038434	1.62	Toll-like receptor signaling pathway	Low
227150_at	*MTF1*	1p33	0.002254	1.61	Regulation of transcription, DNA-dependent	Low
228528_at	*LOC100286909*	1q32.2	0.037179	2.17	Unknown	Low
38149_at	*ARHGAP25*	2p13.3	0.035704	1.75	Signal transduction	Low
204882_at	*ARHGAP25*	2p13.3	0.016663	1.81	Signal transduction	Low
212675_s_at	*CEP68*	2p14	0.042511	1.92	Centrosome organization	Low
228286_at	*GEN1*	2p24.2	0.008045	2.5	DNA repair	Low
240432_x_at	*KLF7*	2q32	0.030033	2.9	Regulation of transcription from RNA polymerase II promoter	Low
233632_s_at	*XRN1*	3q23	0.019577	1.51	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	Low
227119_at	*CNOT6L*	4q13.3	0.017056	1.57	mRNA processing	Low
227601_at	*METTL14*	4q26	0.038673	1.84	RNA methylation	Low
203403_s_at	*RNF6*	13q12.2	0.019639	1.51	Ubiquitin-dependent protein catabolic process	Low
220330_s_at	*SAMSN1*	21q11	0.042021	2.21	Unknown	Low