Leukemia (2007) 21, 1984–1991

www.nature.com/leu

ORIGINAL ARTICLE

Molecular heterogeneity in chronic lymphocytic leukemia is dependent on BCR signaling: clinical correlation

A Rodríguez[1], R Villuendas[1], L Yáñez[2], ME Gómez[3], R Díaz[4], M Pollán[5], N Hernández[3], P de la Cueva[1], MC Marín[1], A Swat[1], E Ruiz[6], MA Cuadrado[2], E Conde[2], L Lombardía[7], F Cifuentes[8], M Gonzalez[9], JA García-Marco[10] and MA Piris[1] for Spanish National Cancer Centre (CNIO)

[1]Molecular Pathology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain; [2]Hematology Department, Hospital Marqués de Valdecilla, Santander, Spain; [3]Hematology Department, Hospital Gregorio Marañón, Madrid, Spain; [4]Bioinformatic Program, Spanish National Cancer Centre (CNIO), Madrid, Spain; [5]National Centre for Epidemiology, Carlos III Institute of Health, Madrid, Spain; [6]Genetics and Pathology Department, Hospital Virgen de la Salud, Toledo, Spain; [7]Biotechnology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain; [8]Agilent Technologies, Palo Alto, CA, USA; [9]Hematology Department, Hospital Clínico Universitario, Salamanca, Spain and [10]Hematology Department, Hospital Universitario Puerta de Hierro, Madrid, Spain

Chronic lymphocytic leukemia (CLL), the most frequent form of adult leukemia in Western countries, is characterized by a highly variable clinical course. Expression profiling of a series of 160 CLL patients allowed interrogating the genes presumably playing a role in pathogenesis, relating the expression of functionally relevant signatures with the time to treatment. First, we identified genes relevant to the biology and prognosis of CLL to build a CLL disease-specific oligonucleotide microarray. Second, we hybridized a training series on the CLL-specific chip, generating a biology-based predictive model. Finally, this model was validated in a new CLL series. Clinical variability in CLL is related with the expression of two gene clusters, associated with B-cell receptor (BCR) signaling and mitogen-activated protein kinase (MAPK) activation, including nuclear factor-κB1 (NF-κB1). The expression of these clusters identifies three risk-score groups with treatment-free survival probabilities at 5 years of 83, 50 and 17%. This molecular predictor can be applied to early clinical stages of CLL. This signature is related to immunoglobulin variable region somatic hypermutation and surrogate markers. There is a molecular heterogeneity in CLL, dependent on the expression of genes defining BCR and MAPK/NF-κB clusters, which can be used to predict time to treatment in early clinical stages.
Leukemia (2007) **21,** 1984–1991; doi:10.1038/sj.leu.2404831; published online 5 July 2007
Keywords: CLL; pathogenesis; prognosis; BCR; microarray

Introduction

Chronic lymphocytic leukemia (CLL), the most frequent form of adult leukemia in Western countries, is characterized by a highly variable clinical course. Almost one-third of patients present relatively stable forms of the disease, with long survival and no requirement for treatment, another one-third of patients progress after an indolent period and the rest of the patients have aggressive forms of the disease and progress quickly from the initial diagnosis to death from disease-related causes within a few years.[1,2] Several clinical and biological prognostic factors have been identified, such as the Rai and Binet clinical staging systems,[3–5] specific cytogenetic alterations,[6] mutational status of immunoglobulin (Ig) genes[7,8] and the expression level of CD38, ZAP70 and LPL.[9–16] The implication of these biological predictor factors with the molecular pathogenesis of the CLL is still under investigation, and an eventual therapeutic application is not yet developed. The recognition of novel molecular variables identified through the use of high-throughput molecular analytical techniques could contribute to a better knowledge of the disease pathogenesis, the development of more accurate biological predictive factors, the adjustment of therapies to the specific risk, and the identification of new therapeutic targets in CLL. Initial expression-profiling analysis in CLL suggests that CLL cases could be considered as a single entity with a homogeneous signature, in opposition to the heterogeneity suggested by immunoglobulin variable region (IgVH) and ZAP70 findings.[17,18]

We have performed genome-wide expression profiling in a large series of CLL patients, under the hypothesis that a comparative analysis of cases with different outcome would depict functional clusters of genes reporting on the most relevant cell functions regulating cell survival.

To demonstrate the reproducibility of the data generated, the analysis was performed in three steps: first, we identified genes relevant to the biology and prognosis of CLL to build a disease-specific oligonucleotide microarray. We then hybridized a new series of cases on the CLL-specific chip, generating a biology-based predictive model capable of defining three different risk groups. Gene clusters included in the final model were involved in B-cell receptor (BCR) signaling and mitogen-activated protein kinase (MAPK)/nuclear factor-κB (NF-κB) activation, highlighting the role of these pathways in pathogenesis and prognosis of CLL. Finally, this predictive statistical model was validated in a new and independent series of CLL patients.

Materials and methods

Patients and samples

A series of 160 consecutive untreated patients with diagnosis of CLL according to National Cancer Institute-sponsored Working Group criteria[19] was obtained from two institutions: Hospital

Correspondence: Dr MA Piris, Molecular Pathology Program, Spanish National Cancer Centre (CNIO), Melchor Fernandez Almagro 3, Madrid 28029, Spain.
E-mail: mapiris@cnio.es
Financial support: This study was supported by grants from the Ministerio de Sanidad y Consumo (G03/179, PI051623) and the Ministerio de Ciencia y Tecnología (SAF2005-00221, SAF2004-04286), Spain. Antonia Rodríguez was supported by a grant from the Asociación Española Contra el Cáncer (AECC).
Received 12 December 2006; revised 28 May 2007; accepted 30 May 2007; published online 5 July 2007

Universitario Puerta de Hierro, Madrid, Spain and Hospital Marqués de Valdecilla, Santander, Spain, under the supervision of the local Ethical Committees. The series was divided into a training set (98 patients from Hospital Universitario Puerta de Hierro) and a validation set (62 patients from Hospital Marqués de Valdecilla). Mononuclear cells were isolated by density-gradient centrifugation from peripheral blood samples at diagnosis. Additional B-cell selection was performed in 20 cases from the training set using anti-CD19 microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany).

The median age at diagnosis was 63 years (range, 33–88 years) for the training set and 73 years (range, 37–87 years) for the validation set. The median follow-up time was 61 months (range, 2–112 months) and 79 months (range, 4–209 months) in the training and validation sets, respectively. First-line treatment[19] needed to be initiated in 47/98 patients of the training set and in 40/62 patients of the validation set (range times to treatment from diagnosis of 1–85 and 1–135 months, respectively). These series were considered to reflect a reasonable variability of CLL patients, diagnosed in different institutions along specific time intervals, thus assuring the applicability of these results to other CLL patients. Patients were treated when active disease developed, according to National Cancer Institute-sponsored Working Group criteria using standard therapies, including alkylating agents and purine analogs, alone or in combination. Additional clinical and biological characteristics are listed in Supplementary Table 1.

Gene-expression profiling

RNA was extracted as described previously[20,21] and hybridized for gene-expression profiling on the Agilent Human 1A 22K Oligonucleotide Microarray (23 cases) and on the CLL-specific 1.9K. Oligonucleotide Microarray (160 cases) following manufacturer's instructions (Agilent, Santa Clara, CA, USA). Briefly, 1 μg of total RNA was amplified and fluorescence-labeled with Agilent's Low RNA Input Fluorescence Linear Amplification kit (Agilent Technology). cRNA product was hybridized following the manufacturer's instructions. Universal Human Reference RNA (Stratagene, La Jolla, CA, USA) was used as a reference. Scanning was carried out using the Agilent G2565AA Microarray Scanner System (Agilent Technologies) and data were collected with Feature Extraction v7.1 software (Agilent Technologies). Inconsistent duplicates were discarded, all consistent duplicates were averaged and genes for which complete data were not available were excluded from further analysis (http://asterias.bioinfo.cnio.es).

Raw data from CLL-specific microarray hybridizations were normalized using a set of 526 genes that exhibited low variability of expression (range, -0.15 to $+0.15$) in more than 30% of cases of the first-step process. Gene expression of the training and validation sets was compared with the median expression level of the training set in order to provide a measure of the variability of expression throughout the series.

Genes associated with CLL and progression: gene selection

We used a multistep approach to develop a predictive model for progression in CLL. First, a high-throughout analysis was performed in a previous series of 23 cases of CLL[21] using the Agilent Human 1A 22K Oligonucleotide Microarray, identifying 88 genes statistically associated with survival (adjusted *P*-value <0.2). After a comprehensive literature search, we also selected all the genes previously described as being associated with pathogenesis and prognosis of CLL, 409 genes derived from 23 publications.

In total, 497 genes (Supplementary Table 5) were obtained from the first step and the literature search and were included in the CLL-specific oligonucleotide microarray. For control purposes, 58 of these selected genes were printed in duplicate or triplicate on the microarray. The 526 genes with a low variability of expression in the 23 cases hybridized in the first step were included as normalization genes. Oligonucleotide sequences as internal controls of hybridization (323 sequences) were also incorporated. Finally, 1900 sequences were printed in known positions on the microarray using an eight-pack microarray format. The array description has been submitted to ArrayExpress (accession number: A-MEXP-328).

This CLL-specific microarray was then used to generate a predictive model in a training set of 98 patients. To eliminate the inherent variability of the different proportion of T cells in the peripheral blood of CLL patients (Supplementary Table 2), only genes found to be expressed by sorted B cells in a previous analysis performed on CD19-positive B-CLL cells were considered (413 genes), including ZAP70 and other genes co-expressed by B cell and other cell subsets. Cluster analysis was used separately with genes that were positively and negatively associated with progression in order to identify groups of coregulated genes. These groups were then used to build a predictive model, as explained in the next section. Finally, the predicted model from the training set was confirmed in the validation series, a new and independent series of 62 patients. Details of these processes are described below.

Statistical analysis

A Cox univariate analysis using the Pomelo and SAM statistical tools was performed to identify genes associated with survival (false discovery rate, FDR <0.2 and *q*-value <0.2, respectively) in the first-step process (http://pomelo2.bioinfo.cnio.es and http://www-stat.stanford.edu/~tibs/SAM/index.html).

To identify genes expressed exclusively by B cells and to exclude information derived from non-B cells, we performed a comparative analysis of sorted B cells from a subset of 20 consecutive cases with the remaining unsorted samples from the training set. A permutation-based *t*-test (http://pomelo2.bioinfo.cnio.es) was used to compare the average expression of B-cell-sorted and -unsorted cases, using a value of FDR <0.2. We selected only genes expressed by B cells for further analysis.

To generate the predictive model, we used a new web tool, http://signs.bioinfo.cnio.es, which uses a combination of statistical gene selection, clustering and survival model building, adapted from the described method by Dave *et al*.[22] First, we identified 50 genes associated with disease progression ($P<0.01$) in the training set using a univariate Cox analysis. These genes were classified as good- or poor-prognosis genes if expression levels were associated with long or short treatment-free survival (TFS), respectively. Prognosis genes were then separately clustered by a complete-linkage hierarchical method to identify associations with the expression patterns that could represent gene-expression signatures. A correlation coefficient of $r>0.6$ and a cluster size of 2–10 genes were chosen as the criteria for this analysis.

For each group of genes or signature, the average expression value was calculated and used as a new explanatory variable to construct a multivariate Cox model for progression, assigning a progression-predictive score to each patient. The linear predictor of this multivariate model served to determine a predictive score, a higher score corresponding to a poorer

prognosis. The predictive model was applied in the validation set to confirm the reproducibility of our results. Finally, we recalculated the risks of the model in the complete series including training and validation sets.

We used the SPSS program to obtain the relative risk of progression, to derive Kaplan–Meier curves and to estimate TFS with regard to three different risk-score groups of patients. All microarray data were submitted to the ArrayExpress (accession number: E-TABM-80; username: e-tabm-80, password: UHrt7Edi).

Results

Progression-predictive model in CLL based on expression profiling

A multistep gene expression analysis was performed, identifying 11 clusters of genes associated with the clinical course, and yielding a model composed of two clusters whose expression was independently associated with variations in the time to treatment.

We performed genomic-scale gene-expression profiling of a series of 160 untreated CLL patients using the original CLL-specific microarray. This was divided between a training set (98 patients) to create a predictive model and a validation set (62 patients) from a different Institution to corroborate the results. The Cox model identified 50 genes associated with TFS ($P<0.01$) in the training set (Supplementary Table 3). These progression-predictive genes were classified as unfavorable or favorable genes if expression levels were associated with short or long TFS, respectively. Both sets of genes were separately hierarchically clustered according to the correlation between expression patterns, as described in the experimental procedures. Unfavorable or favorable genes were grouped in seven and four clusters, respectively (Figures 1 and 2; Table 1).

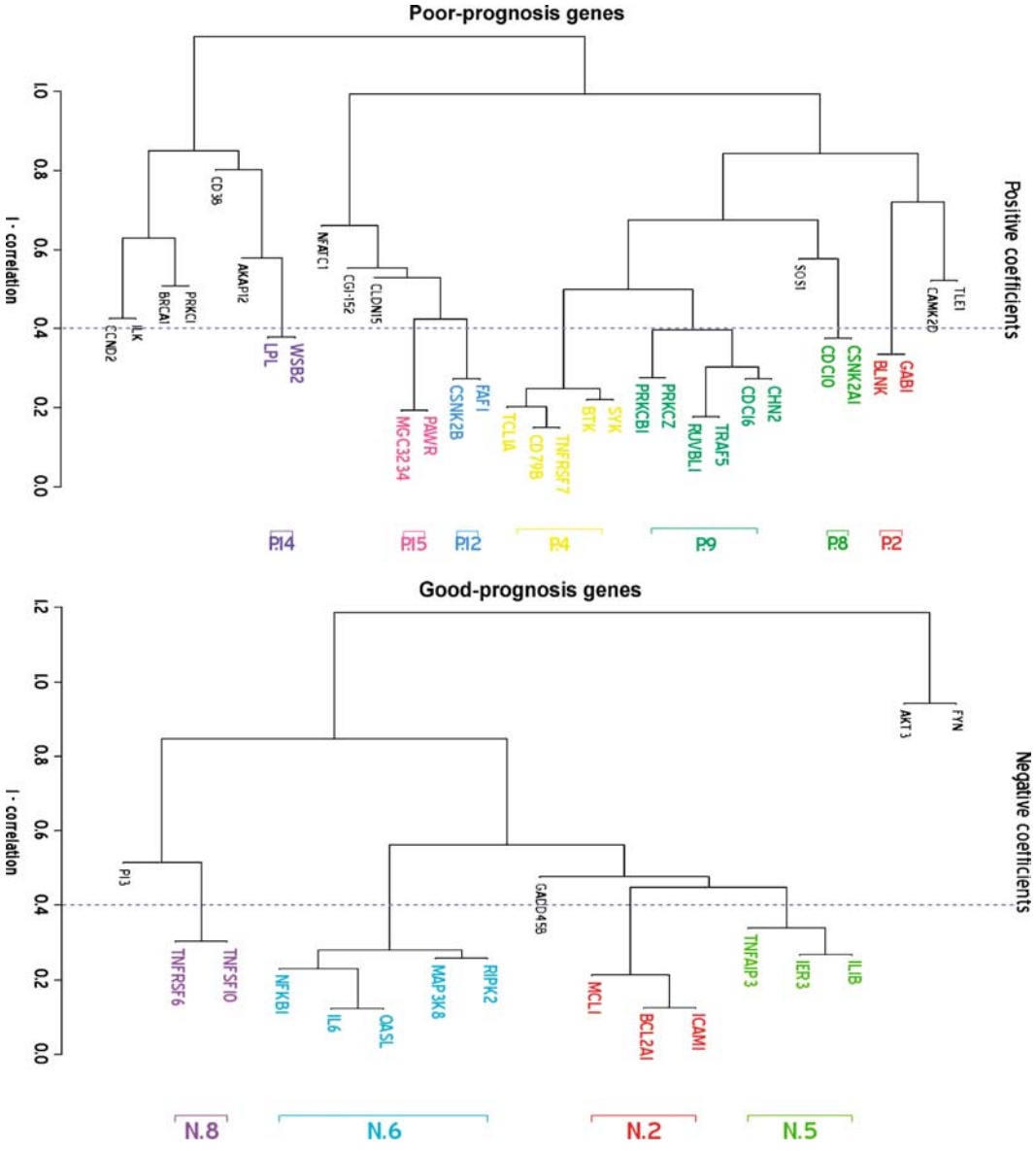

Figure 1 Prognosis gene clusters. Clustering of the 50 genes found to be significantly predictive of clinical course. Genes were clustered based on the following criteria: $P<0.01$, correlation coefficient across gene expression of training samples >0.6 and maximum and minimum number of 10 and 2 genes, respectively. Distance across gene expression is measured as 1-correlation, dashed line represent 1–0.6 (significant clusters have r values >0.6).

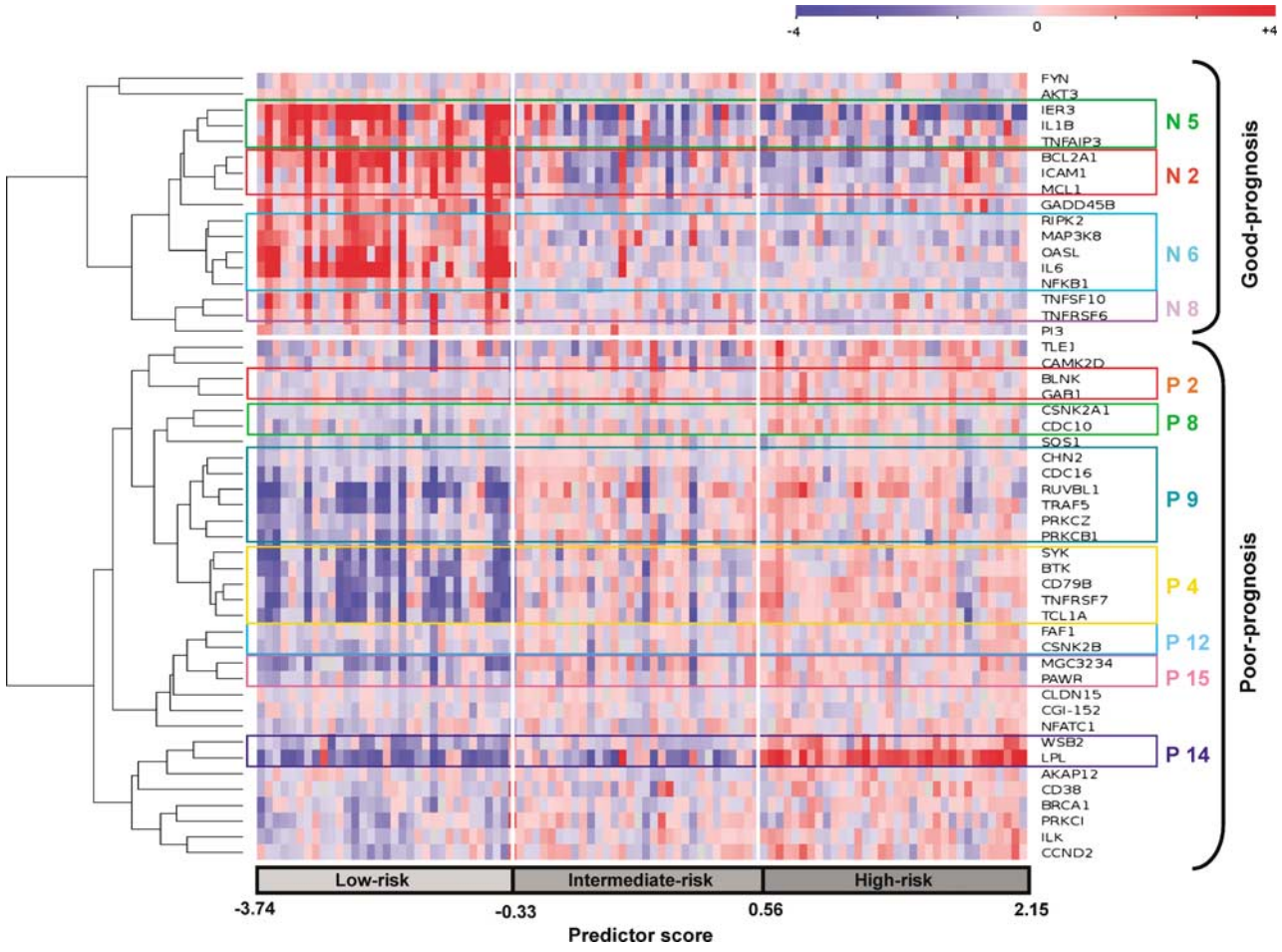

Figure 2 Heat map of prognosis genes. Identification of the signatures associated with variation in time to progression in training series. The series has been divided into terciles. Cases with increased expression of the genes associated with BCR signaling have a shorter time to progression, while cases with greater expression of MAPK/NF-κB genes have a more favorable outcome. Gene expressions are represented as median center values.

Table 1 Identification and predictive power of the clusters of genes associated with variations in the TFS in the training series

Cluster of genes	*Risk of progression (95% CI)*	P-*value*
Clusters of genes associated with longer TFS		
N2: ICAM1, MCL1, BCL2A1	0.65 (0.48–0.86)	0.003
N5: TNFAIP3, IER3, IL1B	0.62 (0.48–0.79)	<0.001
N6: NF-κB1, IL6, OASL, MAP3K8, RIPK2	0.36 (0.19–0.68)	0.002
N8: TNFSF10, TNFRSF6	0.37 (0.19–0.71)	0.003
Clusters of genes associated with shorter TFS		
P2: BLNK1, GAB1	3.7 (1.93–7.1)	<0.001
P4: CD79B, BTK, TNFRSF7 (CD27), SYK, TCL1A	2.11 (1.37–3.25)	0.001
P8: CDC10, CSNK2A1	4.88 (2.2–10.9)	<0.001
P9: PRKCB1, PRKCZ, RUVBL1, TRAF5, CDC16, CHN2	2.90 (1.68–4.99)	<0.001
P12: CSNK2B, FAF1	2.81 (1.35–5.86)	0.006
P14: LPL, WSB2	2.02 (1.57–2.61)	<0.001
P15: PAWR, MGC3234	2.26 (1.41–3.62)	0.001

Abbreviations: CI, confidence interval; TFS, treatment-free survival.

Expression levels of genes within a cluster were averaged for each sample. Different combinations of the 11 gene-expression averages were tested to create a multivariate model for TFS. The final statistical model ($P<0.001$) included two clusters, namely the P14 (LPL and WSB2) and N6 (NF-κB1, IL6, OASL, MAP3K8 and RIPK2) signatures, which were associated with shorter or longer time to treatment, respectively. The relative risks associated with the expression of the clusters N6 and P14 are shown in Table 2. This model was used to assign a progression-predictive score to each patient in the training set: $(2.41 \times \text{P14 cluster average}) + (-3.29 \times \text{N6 cluster average})$. Patients with higher scores also had higher rates of progression. Dividing the series into three terciles, we identified three different risk-score groups of patients with TFS probabilities of 87, 57 and 20% at 5

Table 2 Final predictive model. Predictive power of the two clusters after multivariate analysis

Gene cluster	*Training series*		*Validation series*		*Total (training+validation) series*	
	HR (95% CI)	P-*value*	*HR (95% CI)*	P-*value*	*HR (95% CI)*	P-*value*
N6	0.04 (0.01–0.45)	0.01	0.04 (0.00–0.56)	0.017	0.03 (0.01–0.2)	<0.001
P14	11.15 (3.95–31.41)	<0.001	4.58 (1.81–11.57)	0.001	6.82 (3.52–13.19)	<0.001

Abbreviations: CI, confidence interval; HR, hazard ratio.

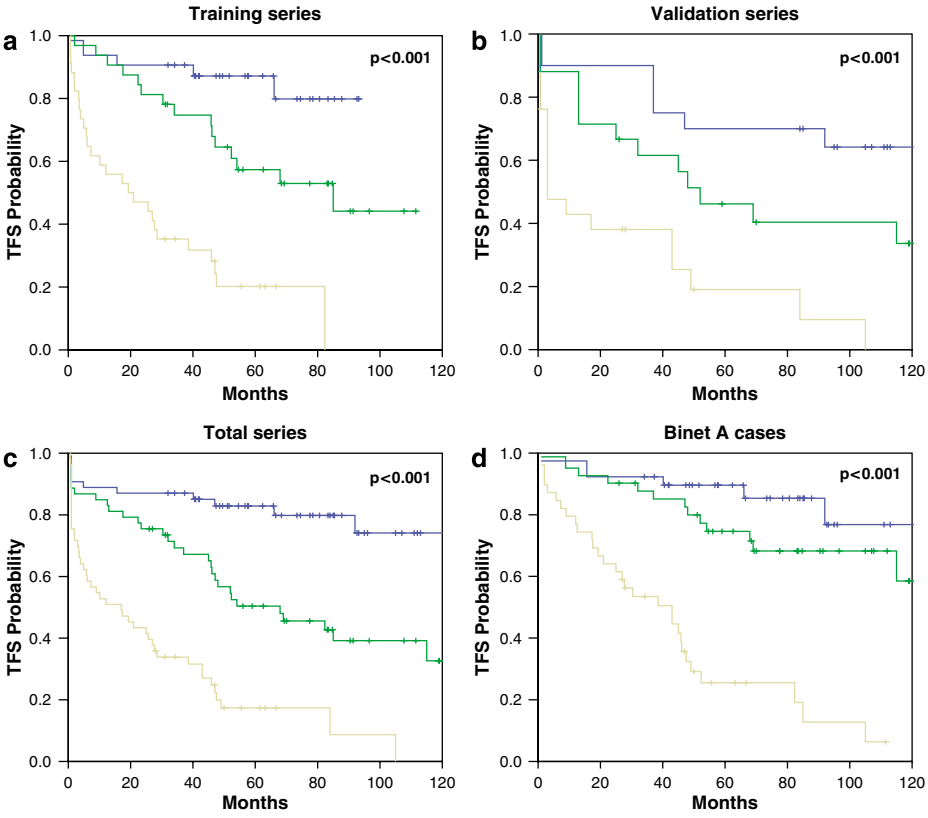

Figure 3 CLL survival analysis. Kaplan–Meier analysis estimates TFS probabilities with regard to the three different risk-score groups in the training set (**a**), validation set (**b**), total series (**c**) and Binet stage A patients (**d**). Cases were divided into terciles using the progression-predictive score (2.41 × P14 cluster average) + (−3.29 × N6 cluster average).

years from diagnosis (Figure 3a). This model was validated in a new independent series of 62 patients, where the model also identified three risk-score groups of patients with TFS probabilities of 70, 46 and 19% at 5 years (Figure 3b). Finally, risk of progression obtained by the generated model in the training set was recalculated in the complete series of the training and validation sets of patients. Kaplan–Meier analysis of the three risk-score groups gave TFS probabilities at 5 years of 83, 50 and 17% (Figure 3c). This biological predictive score was independent of all the clinical and cytogenetic variables found to predict clinical course. An analysis of the model for the patients diagnosed at Binet stage A gave TFS probabilities of 90, 75 and 26% at 5 years from diagnosis ($P<0.001$), thus confirming the applicability of the model to patients diagnosed in early stages (Figure 3d).

The genes here found were analyzed using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways classification (Supplementary Table 4). In the group of genes associated with progression, the main pathways are B-cell receptor signaling and Wnt signaling pathways. In the case of genes associated with no progression, the main pathways are MAPK signaling pathway and apoptosis. The MAPK pathway includes NF-κB1 (Table 3).

To compare this multigene prognostic score with the already existing biological markers, we analyzed the predictive power of the mutational status of immunoglobulin genes (IgVH), ZAP-70 and CD38 expression in the validation series using a univariate Cox regression analysis. IgVH mutational status (mutated if germ-line sequence homology was <98%, unmutated if homology was ⩾98%) and expression of CD38 and ZAP-70 (measured by flow cytometry, with a threshold of 20% for both variables) were found to predict TFS in this series ($P<0.01$). There was a strong correlation between the described multigene prognostic score and IgVH mutational status, with most of cases (44/55, $P<0.001$) exhibiting a positive correlation between the expression of the BCR signaling genes and unmutated IgVH genes. Thus, IgVH was found mutated in 25/29 good prognosis patients whereas IgVH was unmutated in 19/26 bad prognosis

Table 3 KEGG pathway annotation for 50 genes associated with TFS

KEGG pathways	*Genes*	*No. of genes*
Genes associated with shorter time to progression		
B-cell receptor signaling pathway	BTK, CD79B, BLNK, NFATC1 and SYK	5
Wnt signaling pathway	CCND2, CSNK2B, RUVBL1, CAMK2D and NFATC1	5
Tight junction	PRKCZ, CSNK2B, PRKCI and CLDN15	4
Focal adhesion	CCND2, SOS1 and ILK	3
Insulin signaling pathway	SOS1, PRKCZ and PRKCI	3
Calcium signaling pathway	CD38, CAMK2D and NFATC1	3
T-cell receptor signaling pathway	SOS1 and NFATC1	2
Jak–STAT signaling pathway	CCND2 and SOS1	2
Genes associated with longer time to progression		
MAPK signaling pathway	GADD45B, MAP3K8, IL-1B, TNFRSF6, NF-κB1 and AKT3	6
Apoptosis	IL-1B, TNFRSF6, TNFSF10, NF-κB1 and AKT3	5
Toll-like receptor signaling pathway	IL-6, IL-1B, NF-κB1 and AKT3	4
Cytokine–cytokine receptor interaction	IL-6, IL-1B, TNFRSF6 and TNFSF10	4
T-cell receptor signaling pathway	FYN, MAP3K8, NF-κB1 and AKT3	4
Focal adhesion	FYN and AKT3	2
Hematopoietic cell lineage	IL-6 and IL-1B	2
Adipocytokine signaling pathway	NF-κB1 and AKT3	2
B-cell receptor signaling pathway	NF-κB1 and AKT3	2
Jak-STAT signaling pathway	IL-6 and AKT3	2

Abbreviations: KEGG, Kyoto Encyclopedia of Genes and Genomes; TFS, treatment-free survival.

Table 4 Multivariate analysis of the predictor score together with ZAP70, CD38 expression and IgVH mutational status performed in the validation series

Variable	*Validation series*	
	HR (95% CI)	P-*value*
ZAP70	3.44 (1.37–8.63)	<0.05
CD38	3.26 (1.22–8.72)	<0.05
PP-score	1.84 (1.05–3.26)	<0.05
IgVH status	—	NS

Abbreviations: CI, confidence interval; HR, hazard ratio; IgVH, immunoglobulin variable region; NS, not significant; PP-SCORE, progression-predictor score.

patients (Supplementary Figure 1). Multivariate analysis identified the prognosis score, CD38 and ZAP-70 as independent variables predicting time to progression (Table 4).

Discussion

The study here performed identifies multiple biological prognostic markers, illustrating the complexity of this condition and recognizes BCR signaling and aspects of NF-κB regulation as major determinants of CLL cell survival. A predictor model based on these multiple biological markers allowed the division of a training series of CLL patients into three groups with TFS probabilities of 87, 57 and 20% at 5 years from diagnosis. This finding is independent of any clinical and molecular variable and was reproduced in a blind series of CLL patients. It is particularly applicable to patients diagnosed in early clinical stages of CLL (Binet stage A), where the analysis identified three terciles with TFS at 5 years of 90, 75 and 26%. This result differs only slightly from that obtained for the whole series and suggests that this subset of early-stage CLL patients with progressive disease could be more accurately identified.

This analysis in CLL takes advantage of the fact that these patients have not been treated until required, and this permits to compare strictly groups of homogeneously staged patients, without taking into consideration individual variations in the response to treatment.

Tumoral cells in high-risk patients exhibit a molecular signature composed of genes mainly implicated in BCR signaling, such as those defined in the clusters P4, P2, P8 and P9. At the same time, this subset of patients with a shorter time to treatment showed increased expression of negative regulators of the NF-κB pathway, such as those included in clusters P15 and P12. These observations are consistent with the interpretation that stimulatory and growth signals, delivered by the BCR, allow CLL cells to avoid apoptosis and proliferate, thus constituting the driving force for CLL survival.[2,23]The genes whose expression appears associated with more aggressive behavior are all functionally related. Thus, genes expressed within the P4 component, such as CD79B, BTK, TNFRSF7 (CD27) and SYK, are all members of the BCR complex or molecules activated by BCR signaling. This also applies to BLNK and GAB1, which are genes identified in the P2 cluster, a correlated cluster. The Gab1 docking protein links the B-cell antigen receptor to the phosphatidylinositol 3-kinase/Akt signaling pathway and to the SHP2 tyrosine phosphatase.[24] Additionally, the regulation of CSNK2A1 (CK2), the P8 cluster, plays a role in CD5-dependent antigen-receptor activation in B cells.[25] Aggressive CLL cases also show increased expression of a cluster of genes that are comprised of several PKCs, such as PRKCB1 and PRKCZ**,** which are included in the P9 cluster. PRKCB (PKC-beta) is required for BCR-mediated NF-κB activation.[26] Targeted disruption of PRKCZ results in phenotypic alterations in secondary lymphoid organs reminiscent of those observed in TNFR1 or LTB receptor gene-deficient mice. Although with a slightly lower level of significance, these aggressive CLL cases also showed increased expression of PKCI, the second atypical PKC. The lack of zetaPKC in embryonic fibroblasts impairs NF-κB-dependent transcriptional activity as well as cytokine-induced phosphorylation of p65.[27] TRAF5, another component of the same cluster (P9), also plays a role as an adapter protein in

the signal transduction pathway of CD27 that leads to the activation of NF-κB and MAPK8/JNK.[28] Interestingly, three of the genes identified here (RUVBL1, CHN2 and CDC10) have recently been found to be overexpressed in advanced-stage DLBCL.[29] Negative regulators of the NF-κB pathway, such as PAWR (PAR4), have been shown to interact with and inhibit atypical protein kinase C isoforms, functioning as a negative regulator of the NF-κB pathway.[30] Additionally, Fas-associated factor-1, FAF1, inhibits NF-κB activity by interfering with nuclear translocation of the RelA (p65) subunit of NF-κB.[31] This group of patients is also recognized by the expression of the genes defined in the P14 cluster: LPL, WSB2. LPL could play a role in lipid-raft formation or stabilization, a biological process known to be important in B-cell activation.[15] This gene has been reported to be associated with absence of IgVH somatic mutations and shorter TFS in CLL patients.[14–16]Although with a lower level of significance, these patients have increased expression of CD38 and ZAP70, two genes whose expression is associated with adverse outcome in B-CLL.[8,9]

This analysis also identified a group of patients with a more indolent variant of the disease and subsequent longer time to treatment, whose CLL cells express the genes in the N2, N8, N5 and N6 clusters. Genes defining this group of patients are essentially NF-κB genes (NF-κB1), NF-κB upstream regulators (IL6, FAS, RIPK2), NF-κB partners (MAP3K8) and NF-κB targets, such as MCL1, ICAM1, MCL1, BCL2A1 and TNFAIP3, IER3, IL1B or TRAIL. The existence of a subset of CLL cases with an increased level of NF-κB activation and a better outcome has already been described in a different series of patients.[21] Nevertheless, data here obtained need to be interpreted with caution when considering that other NF-κB subunits are not differentially expressed between the CLL patients, and there is a striking cross talking between NF-κB and other regulatory transcription factors. Nevertheless, it appears that CLL cases with more aggressive behavior carry a balanced molecular signature made up of BCR and BCR signaling genes and components of the MAPK/NF-κB pathway. This equilibrium determines the final status of the activation of NF-κB and other transcriptional regulators. This finding is consistent with the slight clinical benefit observed in CLL patients treated with NF-κB inhibitors, and implies that BCR signaling could lead to the activation of other transcription factors responsible for B-cell survival. The increased expression of TCL1, an AKT activator, and GAB1 strongly suggests that these cases could be characterized by PKB/AKT activation. This increased expression of TCL1 also establishes a link with the role of microRNAs in the pathogenesis of CLL, as demonstrated by Croce *et al*.[32]

The interpretation of the data here provided needs to take into consideration a potential caveat, since the study could not potentially completely eliminate the impact of the variable presence of T cells on the expression of genes co-expressed simultaneously by both B and T cells. Nevertheless, the main data and conclusions here provided are supported by the multivariate analysis performed, and by the previous analysis done on peripheral blood B cells, showing that increased BCR signaling genes and the genes involved in the MAPK/NF-κB pathway are expressed by the tumoral B cells.[21,33]

These findings confirm that the already identified heterogeneity of CLL cases, when considering IgVH somatic mutation or the expression of surrogate markers, can be extended to the molecular signatures of CLL, this allowing to assign CLL patients to specific risk groups depending on the degree of activation of BCR signaling genes, and some genes involved in the MAPK/NF-κB pathway.

When these data are considered in the context of the previously available information about other prognostic indicators in CLL,[33] it becomes clear that tumors characterized by the overexpression of genes associated with BCR signaling have a lower rate of somatic hypermutation of the IgVH genes and an increased expression of CD38 and ZAP70, and may allow to propose that both IgVH-unmutated sequence and ZA70 expression could merely be surrogate markers for increased BCR signaling dependent on the recognition of unknown antigens.[33,34]

Conclusion

This study recasts CLL as a disease with a wide pathogenic spectrum, whose recognition may have a twofold clinical benefit, through a more precise stratification of patients, and the ability to identify multiple potential therapeutic targets.

Acknowledgements

We thank the Tumour Bank Network of the CNIO for the coordination of shipment of samples. Francisco Cifuentes is a full-time employee of the company Agilent Technologies, whose custom product '8-pack microarray' platform was used in the present study.

References

1 Dighiero G. CLL biology and prognosis. *Hematology (Am Soc Hematol Educ Program)* 2005; **1**: 278–284.
2 Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med* 2005; **352**: 804–815.
3 Binet JL, Lepoprier M, Dighiero G, Charron D, D'Athis P, Vaugier G *et al*. A clinical staging system for chronic lymphocytic leukemia: prognostic significance. *Cancer* 1977; **40**: 855–864.
4 Rai KR, Montserrat E. Prognostic factors in chronic lymphocytic leukemia. *Semin Hematol* 1987; **24**: 252–256.
5 Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood* 1975; **46**: 219–234.
6 Dohner H, Stilgenbauer S, Benner A, Leupolt E, Krober A, Bullinger L *et al*. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* 2000; **343**: 1910–1916.
7 Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 1999; **94**: 1848–1854.
8 Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL *et al*. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 1999; **94**: 1840–1847.
9 Crespo M, Bosch F, Villamor N, Bellosillo B, Colomer D, Rozman M *et al*. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N Engl J Med* 2003; **348**: 1764–1775.
10 Durig J, Naschar M, Schmucker U, Renzing-Kohler K, Holter T, Huttmann A *et al*. CD38 expression is an important prognostic marker in chronic lymphocytic leukaemia. *Leukemia* 2002; **16**: 30–35.
11 Durig J, Nuckel H, Cremer M, Fuhrer A, Halfmeyer K, Fandrey J *et al*. ZAP-70 expression is a prognostic factor in chronic lymphocytic leukemia. *Leukemia* 2003; **17**: 2426–2434.
12 Ibrahim S, Keating M, Do KA, O'Brien S, Huh YO, Jilani I *et al*. CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. *Blood* 2001; **98**: 181–186.
13 Wiestner A, Rosenwald A, Barry TS, Wright G, Davis RE, Henrickson SE *et al*. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin

genes, inferior clinical outcome, and distinct gene expression profile. *Blood* 2003; **101**: 4944–4951.
14 Van't Veer MB, Brooijmans AM, Langerak AW, Verhaaf B, Goudswaard CS, Graveland WJ *et al.* The predictive value of lipoprotein lipase for survival in chronic lymphocytic leukemia. *Haematologica* 2006; **91**: 56–63.
15 Oppezzo P, Vasconcelos Y, Settegrana C, Jeannel D, Vuillier F, Legarff-Tavernier M *et al.* The LPL/ADAM29 expression ratio is a novel prognosis indicator in chronic lymphocytic leukemia. *Blood* 2005; **106**: 650–657.
16 Heintel D, Kienle D, Shehata M, Krober A, Kroemer E, Schwarzinger I *et al.* High expression of lipoprotein lipase in poor risk B-cell chronic lymphocytic leukemia. *Leukemia* 2005; **19**: 1216–1223.
17 Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, Husson H *et al.* Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med* 2001; **194**: 1625–1638.
18 Rosenwald A, Alizadeh AA, Widhopf G, Simon R, Davis RE, Yu X *et al.* Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med* 2001; **194**: 1639–1647.
19 Cheson BD, Bennett JM, Grever M, Kay N, Keating MJ, O'Brien S *et al.* National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. *Blood* 1996; **87**: 4990–4997.
20 Tracey L, Villuendas R, Ortiz P, Dopazo A, Spiteri I, Lombardia L *et al.* Identification of genes involved in resistance to interferon-alpha in cutaneous T-cell lymphoma. *Am J Pathol* 2002; **161**: 1825–1837.
21 Rodriguez A, Martinez N, Camacho FI, Ruiz-Ballesteros E, Algara P, Garcia JF *et al.* Variability in the degree of expression of phosphorylated IkappaBalpha in chronic lymphocytic leukemia cases with nodal involvement. *Clin Cancer Res* 2004; **10**: 6796–6806.
22 Dave SS, Wright G, Tan B, Rosenwald A, Gascoyne RD, Chan WC *et al.* Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. *N Engl J Med* 2004; **351**: 2159–2169.
23 Stevenson FK, Caligaris-Cappio F. Chronic lymphocytic leukemia: revelations from the B-cell receptor. *Blood* 2004; **103**: 4389–4395.
24 Ingham RJ, Santos L, Dang-Lawson M, Holgado-Madruga M, Dudek P, Maroun CR *et al.* The Gab1 docking protein links the b cell antigen receptor to the phosphatidylinositol 3-kinase/Akt signaling pathway and to the SHP2 tyrosine phosphatase. *J Biol Chem* 2001; **276**: 12257–12265.
25 Raman C, Kuo A, Deshane J, Litchfield DW, Kimberly RP. Regulation of casein kinase 2 by direct interaction with cell surface receptor CD5. *J Biol Chem* 1998; **273**: 19183–19189.
26 Su TT, Guo B, Kawakami Y, Sommer K, Chae K, Humphries LA *et al.* PKC-beta controls I kappa B kinase lipid raft recruitment and activation in response to BCR signaling. *Nat Immunol* 2002; **3**: 780–786.
27 Leitges M, Sanz L, Martin P, Duran A, Braun U, Garcia JF *et al.* Targeted disruption of the zetaPKC gene results in the impairment of the NF-kappaB pathway. *Mol Cell* 2001; **8**: 771–780.
28 Nakano H, Oshima H, Chung W, Williams-Abbott L, Ware CF, Yagita H *et al.* TRAF5, an activator of NF-kappaB and putative signal transducer for the lymphotoxin-beta receptor. *J Biol Chem* 1996; **271**: 14661–14664.
29 Nishiu M, Yanagawa R, Nakatsuka S, Yao M, Tsunoda T, Nakamura Y *et al.* Microarray analysis of gene-expression profiles in diffuse large B-cell lymphoma: identification of genes related to disease progression. *Jpn J Cancer Res* 2002; **93**: 894–901.
30 Garcia-Cao I, Lafuente MJ, Criado LM, Diaz-Meco MT, Serrano M, Moscat J. Genetic inactivation of Par4 results in hyperactivation of NF-kappaB and impairment of JNK and p38. *EMBO Rep* 2003; **4**: 307–312.
31 Park MY, Jang HD, Lee SY, Lee KJ, Kim E. Fas-associated factor-1 inhibits nuclear factor-kappaB (NF-kappaB) activity by interfering with nuclear translocation of the RelA (p65) subunit of NF-kappaB. *J Biol Chem* 2004; **279**: 2544–2549.
32 Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE *et al.* A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005; **353**: 1793–1801.
33 Huttmann A, Klein-Hitpass L, Thomale J, Deenen R, Carpinteiro A, Nuckel H *et al.* Gene expression signatures separate B-cell chronic lymphocytic leukaemia prognostic subgroups defined by ZAP-70 and CD38 expression status. *Leukemia* 2006; **20**: 1774–1782.
34 Chen L, Widhopf G, Huynh L, Rassenti L, Rai KR, Weiss A *et al.* Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* 2002; **100**: 4609–4614.

Supplementary Information accompanies the paper on the Leukemia web site (http://www.nature.com/leu)