Association between B-cell receptor responsiveness and disease progression in B-cell chronic lymphocytic leukemia: results from single cell network profiling studies

Alessandra Cesano,[1] Omar Perbellini,[2] Erik Evensen,[1] Charles C. Chu,[3] Federica Cioffi,[2] Jason Ptacek,[1] Rajendra N Damle,[3] Roberto Chignola,[4] James Cordeiro,[1] Xiao-jie Yan,[3] Rachael E. Hawtin,[1] Ilaria Nichele,[2] Jodi R. Ware,[1] Chiara Cavallini,[2,5] Ornella Lovato,[5] Roberta Zanotti,[2] Kanti R. Rai,[3] Nicholas Chiorazzi,[3] Giovanni Pizzolo,[2] and Maria T. Scupoli[2,5,6]

[1]Nodality Inc., South San Francisco, CA, USA; [2]Department of Medicine, Section of Hematology, University of Verona, Verona, Italy; [3]Feinstein Institute for Medical Research, Long Island Jewish Health System, Manhasset, NY, USA; [4]Department of Biotechnology, University of Verona, Verona, Italy; [5]Interdepartmental Laboratory for Medical Research (LURM), University of Verona, Verona, Italy; and [6]Applied Research on Cancer–Network (ARC-NET), University of Verona, Verona, Italy

 doi:10.3324/haematol.2012.071910

Online Supplementary Design and Methods

Experimental and gating analysis in the exploratory study

Cryopreserved peripheral blood mononuclear cells were thawed at 37°C, washed in phosphate-buffered saline/10% fetal bovine serum (FBS), resuspended in RPMI 1640/10% FBS, and aliquoted to $4x10^6$ cells per condition. Cells were rested for 1 h at 37°C followed by B-cell receptor (BCR) modulation, incubating cells with 20 μg/mL polyclonal goat $F(ab')_2$ anti-human IgM (Southern Biotech, Birmingham, AL, USA) for 10 min in standard tubes. This time point was chosen on the basis of preliminary time-course results (at 10, 30, 60 min) showing that anti-IgM responses at 10 min were representative, although not necessarily maximal, for each phosphoprotein (*Online Supplementary Figure S2*). The cells were fixed with pre-warmed Fix Buffer I (Becton Dickinson, San Jose, CA, USA) at 37°C for 10 min and then pelleted, resuspended and permeabilized with ice-cold 50% methanol for 30 min on ice. The permeabilized cells were washed with fluorescence activated cell sorting (FACS) buffer (phosphate-buffered saline/0.5% bovine serum albumin/0.05% NaN_3), pelleted, and stained with a cocktail of fluorochrome-conjugated antibodies (*Online Supplementary Table S4A*). Approximately 10,000 gated events were acquired for each sample on a FACSCalibur dual-laser cytometer (Becton Dickinson).

Flow cytometry data were gated using FlowJo software (TreeStar, Ashland, OR, USA). Dead cells and debris were excluded based upon forward scatter and side scatter measurements. All analyses were gated on B-cell chronic lymphocytic leukemia (B-CLL) cells, which were identified on the basis of CD5 and CD19 co-expression. Fluorescence signals were normalized with respect to the controls by calculating the ratio between the median fluorescence intensity of each phosphoprotein and the respective isotype-matched, monoclonal antibody control (relative median fluorescence intensity, RMFI).

Changes in phosphorylation of phosphoproteins were expressed by calculating the $\log_2$Fold difference in RMFI of modulated cells divided by unmodulated cells ("$\log_2$Fold change"). A number of samples (both in unmodulated and modulated conditions) showed a bimodal distribution of fluorescence, with a proportion of cells displaying higher fluorescence intensity of p-SYK (5/27 samples in the unmodulated condition and 4/27 samples in the modulated condition) and p-NF-κB (1/27 sample in the modulated condition) (see Figure 1A for representative histograms). The proportion of positive cells was >50% in each case and the bimodal samples included both *IGHV* unmutated and mutated cases. In bimodal histograms, the median of the positive population was considered as an indication of the value of protein phosphorylation status for analysis.

Experimental and gating analyses in the test studies

Cryopreserved peripheral blood mononuclear cells were thawed at 37°C, stained for viability with Amine Aqua (Invitrogen, Carlsbad, CA, USA), resuspended in RPMI 1640/10% FBS and aliquoted at 100,000 cells per well into two replicate (test 1) or single (test 2) 96-deepwell plates. Cells were rested for 2 h at 37°C followed by BCR modulation for 10 min, incubating cells with 20 μg/mL polyclonal goat $F(ab')_2$ anti human IgM (Southern Biotech). Cells were fixed with paraformaldehyde at a final concentration of 1.6% for 10 min at 37°C, pelleted, resuspended and permeabilized with 100% methanol, then stored at -80°C overnight. The permeabilized cells were washed with FACS buffer, pelleted, and stained with a cocktail of fluorochrome-conjugated antibodies (*Online Supplementary Table S4B,C*). Approximately 30,000 gated events were acquired for each sample on an LSR II three-laser cytometer (Becton Dickinson).

Flow cytometry data were gated using WinList (Verity House Software, Topsham, ME, USA). Dead cells and debris were excluded by forward scatter, side scatter, and Amine Aqua viability dye. All analyses were gated on B-CLL cells, which were identified as CD3 negative cells exhibiting CD5 and co-expressing either the B cell marker CD20 (test 1) or CD19 (test 2). The raw instrument fluorescence intensities were converted to calibrated intensity metrics (equivalent number of reference fluo-

rophores, ERF).[1,2] The calibration was applied on a plate-by-plate basis using the rainbow calibration particles included on each plate. This correction ensures that data across the plate and between plates are calibrated to the same values, regardless of the instrument used for acquisition.

Determination of the cut-point for predicting time to first treatment in test studies

To implement the anti-IgM→p-ERK | Uu finding as a prognostic classifier, an optimal cut-point for separating patients into groups with a favorable or unfavorable prognosis was identified in test study 1 by a linear search of all possible cut-offs with the *P*-value of the log rank test for difference in Kaplan-Meier time to first treatment (TTFT) estimates as the objective. A cut-off of 0.66 (Uu metric) was determined to yield the best separation in Kaplan-Meier TTFT estimates for favorable (anti-IgM→p-ERK | Uu < 0.66) and unfavorable (anti-IgM→p-ERK | Uu ≥ 0.66) prognostic groups.[3] As multiple cut-points yielded the minimum log rank *P*-value, the cut-point was chosen by examining the plot of *P*-values *versus* cut-points (*data not shown*).

The performance of this cut-point in an independent sample set was assessed by applying the cut-point to the test 2 cohort and computing the log rank test for the difference in Kaplan-Meier TTFT estimates for patients with anti-IgM→p-ERK | Uu less than the cut-off (favorable prognostic group) compared to those with signaling greater than or equal to the cut-off (unfavorable prognostic group).

Repeatability assay analysis in the test study 1

Assay repeats were run on different plates on the same day with fixed operators. Following laboratory processing, linear regression and Bland-Altman analyses[4,5] were performed to assess assay repeatability. In replicate single cell network profiling (SCNP) experiments, anti-IgM→p-ERK | log_2Fold signaling was highly repeatable. Linear regression of the repeats displayed high concordance with an R^2=0.96, slope=0.98, and an intercept of 0.02 (*Online Supplementary Figure S3A*). The bias estimate of the log_2Fold metric from the Bland-Altman analysis was -0.01 (*data not shown*), indicating excellent repeatability of the assay. Similar results were obtained for the same analysis using the Uu metric; R2 = 0.96, slope = 0.99, and intercept = 0.0 (*Online Supplementary Figure S3B*); for the Bland-Altman analysis of the Uu metric data, the bias estimate was 0.0003 (*data not shown*).

Patients' samples in the bridging study

The bridging study was performed using ten patients' samples that had been previously analyzed in the exploratory study. The Verona and Nodality laboratories independently applied the SCNP assay using their own procedures, in order to determine differences due to the methodologies. The goal of the bridging study was to confirm assay reproducibility between laboratories and between associated measurements using two different methodologies.

Experimental methodology in the bridging study

The bridging study performed at Nodality evaluated two methods for phospho-specific flow cytometry using duplicate aliquots of cryopreserved samples from ten B-CLL donors processed in the exploratory study at Verona University. The notable differences between the two methods are that Verona University used BD Fix Buffer 1 compared to 1.6% paraformaldehyde by Nodality for fixation, cell permeabilization was performed by incubating on ice with 50% methanol for 30 min (Verona University) *versus* 100% methanol at -80°C for ~16 h (Nodality), and CD19 (Verona University) and CD20 (Nodality) was used for B-cell delineation (*Online Supplementary Table S3*).

Briefly, the cryopreserved samples were thawed at 37°C, washed with RPMI with 1% FBS, stained for viability with Amine Aqua (Invitrogen), filtered to remove debris, and counted. Cells were resuspended in RPMI with 10% FBS, and aliquoted at 100,000 cells per well into two replicate 96-deep-well plates with one replicate processed with much of the same methodology used in the exploratory study (Verona methodology) and the second replicate plate processed using techniques optimized at Nodality. The Verona plate was rested for 1 h at 37°C; the Nodality plate was rested for 2 h at 37°C. Modulation of BCR signaling pathways was performed by the addition of polyclonal goat $F(ab')_2$ anti-human IgM (Southern Biotech) at 20 μg/mL for 10 min to crosslink the BCR and activate downstream signaling pathways. Cells and their signaling pathways were fixed by the addition of either BD Fix Buffer 1 (Verona plate) or paraformaldehyde at a final concentration of 1.6% (Nodality plate) for 10 min at 37°C. The cells were pelleted and resuspended in 50% ice-cold methanol for 30 min (Verona plate) or 100% methanol and stored at -80°C for ~16 h (Nodality plate). Following permeabilization, the cells were washed with FACS buffer (phosphate-buffered saline, 0.5% bovine serum albumin, 0.05% NaN_3), and stained with a set of antibody cocktails. The fluorochrome-conjugated antibodies included were specific for: CD3 (clone UCHT1), CD19 (clone SJ25C1), CD20 (clone H1), p-ERK (T202/Y204) (clone 20A), p-ZAP70/p-SYK (pY319/pY352) (clone 17a/P-ZAP70) (Becton Dickson); CD5 (clone UCHT2), (Biolegend); p-ERK (T202/Y204) (clone D13.14.4E), ZAP-70 (clone 136F12) (Cell Signaling Technologies). Antibody concentrations and specificity had been evaluated previously as detailed by Palazzo *et al.*[6]

Flow cytometry data were acquired on LSRII Flow Cytometers equipped with a high throughput sampler using FACS DIVA software (Becton Dickson). All flow cytometry data were analyzed with WinList (Verity House Software). Dead cells and debris were excluded by forward scatter, side scatter, and the amine aqua viability dye. Leukemic cells were identified as $CD3^-$, $CD19^+$ or $CD20^+$ and $CD5^+$.

Data analysis in bridging study

The assay results were bridged by comparison of the SCNP assay readouts obtained using the experimental process conducted in the exploratory study and that used in the test studies. Following laboratory processing, linear regression was performed to assess assay concordance between the University of Verona and Nodality SCNP procedures.

Results and conclusions from the bridging study

Nodality performed SCNP experiments using ten B-CLL exploratory study samples for which additional aliquots were

available. Assay concordance for anti-IgM-induced fold changes for p-ERK and p-SYK were R=0.89 and R=0.66 for p-SYK, respectively (*Online Supplementary Figure S5* for p-ERK results).

The bridging study showed good correspondence and high reproducibility between the measurements conducted by the two laboratories on the same samples despite some differences in processes, reagents and instruments, thus indicating the robustness of the biology measured for the signaling nodes that were tested.

References

1. Purvis N, Stelzer G. Multi-platform, multi-site instrumentation and reagent standardization. Cytometry. 1998;33(2):156-65.
2. Wang L, Gaigalas AK, Marti G, Abbasi F, Hoffman RA. Toward quantitative fluorescence measurements with multicolor flow cytometry. Cytometry A. 2008;73(4):279-88.
3. Koller C, Bekele BN, Zhou X, Park C, Estrov Z, O'Brien S, et al. Plasma thrombopoietin compared with immunoglobulin heavy-chain mutation status as a predictor of survival in chronic lymphocytic leukemia. Blood. 2006;108(3):1001-6.
4. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician. 1983;32:307-17.
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.
6. Palazzo AI, Evensen E, Huang YG, Cesano A, Nolan GP, Fantl WJ. Association of reactive oxygen species-mediated signal transduction with in vitro apoptosis sensitivity in chronic lymphocytic leukemia B cells. PLoS One. 2011;6(10):e24592.
7. Zanotti R, Ambrosetti A, Lestani M, Ghia P, Pattaro C, Remo A, et al. ZAP-70 expression, as detected by immunohistochemistry on bone marrow biopsies from early-phase CLL patients, is a strong adverse prognostic factor. Leukemia. 2007;21(1):102-9.

Online Supplementary Table S1. **B-CLL patients and sample eligibility criteria for the exploratory and test studies.**

	Exploratory (n=27)	Test 1 (n=30)	Test 2 (n=37)
PBMC collection and cryopreservation			
Dates	August 2003 to December 2008	January 2004 to September 2009	2006 and 2007
Process	Isolated from patients' samples by Ficoll-hypaque centrifugation (Lymphoprep, Nicomed, Oslo, Norway), suspended in freezing medium (FBS, 10% DMSO), and stored in liquid nitrogen		
Inclusion criteria for B-CLL patients' samples			
Cell viability		> 85%	
Data analysis inclusion	27 B-CLL patients included in exploratory analysis	30 Binet stage A patients included in test 1 analysis	37 Rai stage 0 & 1 patients included in test 2 analysis
Prior treatment	Lack of prior treatment for at least 6 months prior to sample collection	No prior treatment	No prior treatment
Clinical annotation	Availability, including patient and disease characteristics, date of diagnosis, clinical staging, serum biomarker concentrations (e.g., β_2 microglobulin, LDH), results from FISH analysis for cytogenetics, biological prognostic parameters (e.g., *IGHV* status, ZAP-70 and CD38 expression*), date of sample collection, TTFT and list of treatments with respective dates of administration		

*IGHV *mutational status, ZAP-70 and CD38 expression were determined as previously described*[7]. *PBMC: peripheral blood mononuclear cells; FBS: fetal bovine serum; DMSO: dimethylsulfoxide*

Online Supplementary Table S2. Characteristics of patients included in the exploratory and test studies.

	Exploratory study evaluable All Stage B-CLL (n=27)	Exploratory study evaluable Binet Stage A B-CLL (n=21)	Test study 1 evaluable Binet Stage A B-CLL (n=30)	*P* value Binet Stage A B-CLL	Test study 2 evaluable Rai Stage 0,1 B-CLL (n=37)	*P* value Rai Stage 0,1 B-CLL *vs.* exploratory study Binet Stage A
Age at diagnosis Median (range), years	55 (36 to 73)	54 (36 to 73)	67 (51 to 83)	0.00022	56 (34-81)	0.33
Age at sample collection Median (range), years	59 (45 to 86)	59 (47 to 86)	70 (53 to 83)	0.0064	58 (34-88)	0.94
Gender Male Female	15 (56%) 12 (44%)	10 (48%) 11 (52%)	21 (70%) 9 (30%)	0.19	28 (76%) 9 (24%)	0.045
IGHV Mutated Unmutated Unavailable	13 (48%) 14 (52%) 0	12 (57%) 9 (43%) 0	22 (74%) 7 (23%) 1 (3%)	0.27*	20 (54%) 17 (46%) 0	1.0
Binet/Rai stage	Binet Stage A: 21 (78%) B: 3 (11%) C: 3 (11%)	Binet Stage A: 21 (100%) B: NA C: NA	Binet Stage A: 30 (100%) B: NA C: NA	NA	Rai Stage 0: 21 (57%) 1: 16 (43%)	NA
ZAP-70 (>20%) Negative Positive Unavailable	14 (52%) 13 (48%) 0	13 (62%) 8 (38%) 0	20 (67%) 9 (30%) 1 (3%)	0.83*	16 (43%) 19 (52%) 2 (5%)	0.28
CD38 (>30%) Negative Positive	18 (67%) 9 (33%)	16 (76%) 5 (24%)	23 (77%) 7 (23%)	0.78	23 (62%) 14 (38%)	0.38
Cytogenetics Favorable Neutral Unfavorable Unavailable	5 (18%) 17 (63%) 5 (18%) 0	3 (14%) 14 (67%) 4 (19%) 0	7 (23%) 21 (70%) 2 (7%) 0	0.34	8 (22%) 11 (30%) 8 (22%) 10 (26%)	0.24
Treated prior to sample collection	9 (33%)	6 (29%)	0	0.0075	0	0.0013
Follow up Median (range), months	164 (28 to 430)	164 (29 to 430)	57 (4-179)	0.049	89 (4-308)	0.38
Progressed to active disease, requiring treatment	20 (74%)	14 (67%)	8 (27%)	0.011	15(41%)	-0.10
TTFT Median (range), months	42 (1-415)	50 (6-415)	149 (9-164)	0.067	102 (11-162)	0.089

** Leaving out one unavailable case from the test set. NA, not applicable*

Online Supplementary Table S3. Summary of key differences between the experimental methodologies of Verona University and the Nodality Laboratory.

	Verona University Laboratory (Exploratory Study)	Nodality Laboratory (Test Studies)
Reagents	Materials prepared and used by Verona University	Materials prepared and used by Nodality using standard Nodality procedures
Antibody clones for p-ERK	BD Biosciences Clone 20A on Alexa 647	Cell Signaling Technology (CST) Clone D13.14.4E on Alexa 488
Cell concentration for modulation	$4x10^6$ cells/mL	$1x10^6$ cells/mL
Staining	Approximately $3x10^5$ cells/tube utilizing a tube-based format	Approximately $1x10^5$ cells/well utilizing a 96-well plate format
Fixation	BD Fix Buffer 1	1.6% paraformaldehyde
Permeabilization	50% MeOH on ice for 30 minutes	100% MeOH at -80°C overnight
Controls	19 healthy control PBMC processed in individual tubes	1 healthy control PBMC on each plate Ramos control cell lines on each plate for qualitatively evaluating modulation
Staining controls	Isotype	Autofluorescence

Online Supplementary Table S4A. Antibodies used in the exploratory study.

Antibody	Species and Isotype	Manufacturer	Clone
CD5	Mouse IgG2$_a$, κ	Becton Dickson	L17F12
CD19	Mouse IgG1, κ	Becton Dickson	SJ25C1
p-SYK (Y353)	Mouse IgG1	Becton Dickson	17A/P-ZAP70
p-NF-κB p65 (S529)	Mouse IgG2$_b$, κ	Becton Dickson	K10-895.12.50
p-ERK (T202/Y204)	Mouse IgG1	Becton Dickson	20A
p-p38 (T180/Y182)	Mouse IgG1, κ	Becton Dickson	36/p38(pT180/pY182)
p-JNK (T183/Y185)	Mouse IgG1	Cell Signaling Technology	G9

Online Supplementary Table S4B. Antibodies used in test study 1.

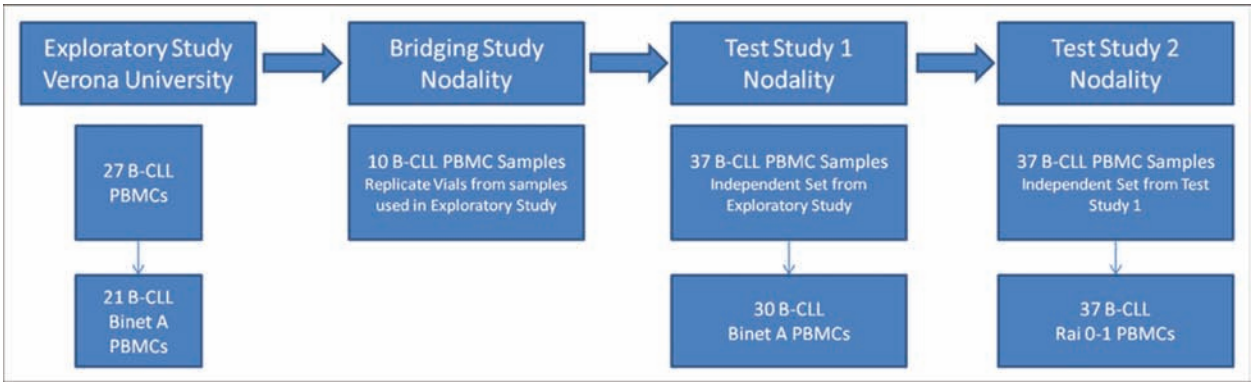
Antibody	Species and Isotype	Manufacturer	Clone
CD3	Mouse IgG1, κ	Becton Dickson	UCTH1
CD5	Mouse IgG1, κ	Biolegend	UCHT2
CD20	Mouse IgG2$_a$, κ	Becton Dickson	H1
p-SYK (Y352)	Mouse IgG1	Becton Dickson	17A/P-ZAP70
p-ERK (T202/Y204)	Rabbit IgG	Cell Signaling Technology	D13.14.4E

Online Supplementary Table S4C. Antibodies used in test study 2.

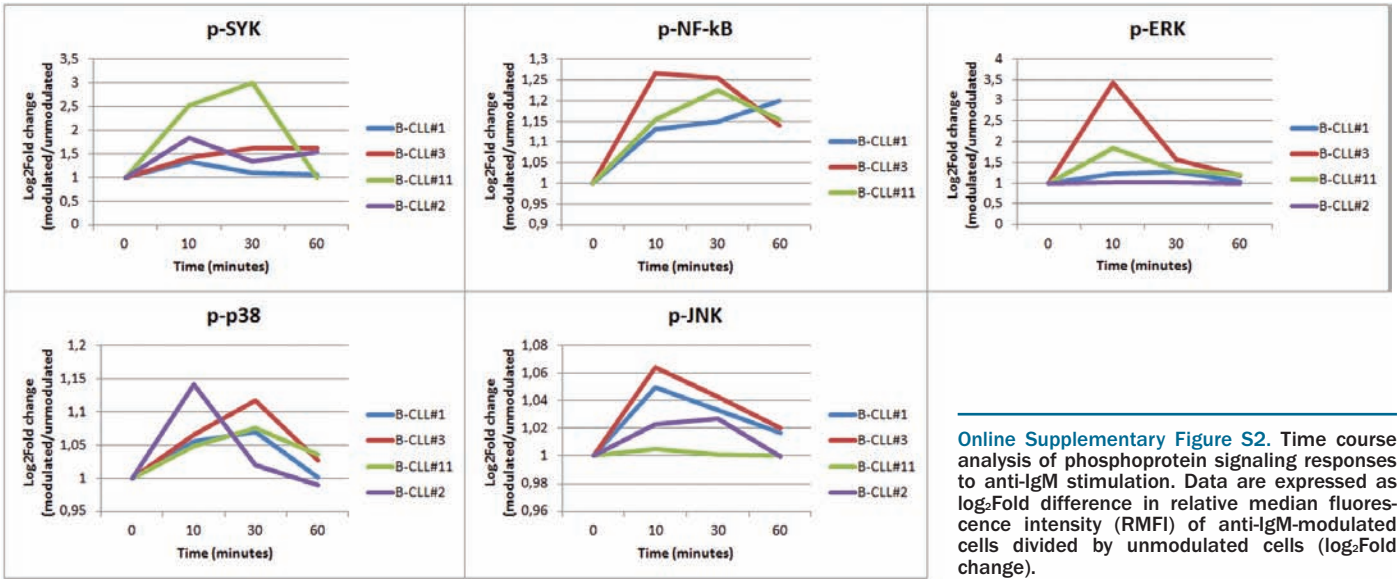
Antibody	Species and Isotype	Manufacturer	Clone
CD3	Mouse IgG1, κ	Becton Dickson	UCTH1
CD5	Mouse IgG1, κ	Biolegend	UCHT2
CD19	Mouse IgG1, κ	Becton Dickson	SJ25C1
p-SYK (Y352)	Mouse IgG1	Becton Dickson	17A/P-ZAP70
p-ERK (T202/Y204)	Rabbit IgG	Cell Signaling Technology	D13.14.4E

Online Supplementary Table S5. Activation of downstream BCR signaling pathways in mutated and unmutated CLL samples.

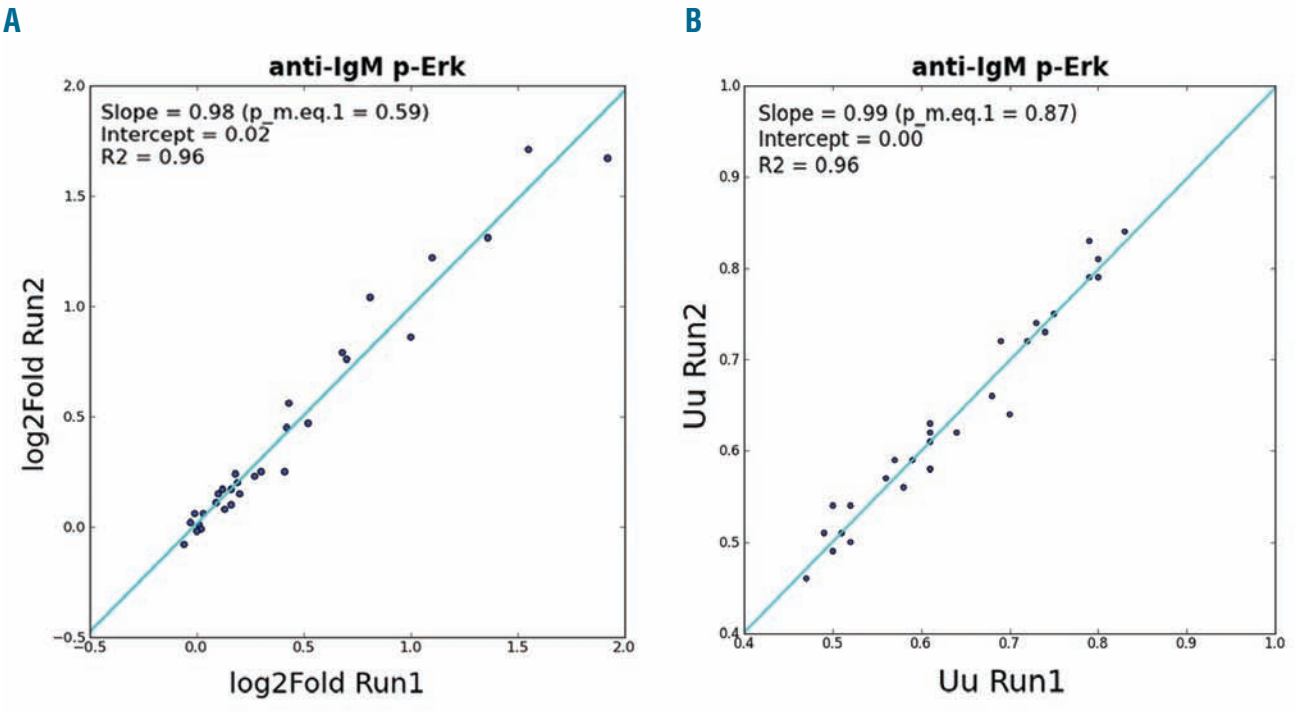
	p-SYK	p-NF-κB	p-ERK	p-p38	p-JNK
Mutated positive/total cases (%)	4/13 (31%)	1/13 (8%)	3/13 (23%)	2/13 (15%)	0/13 (0%)
Unmutated positive/total cases (%)	14/14 (100%)	6/14 (43%)	8/14 (57%)	1/14 (7%)	0/14 (0%)



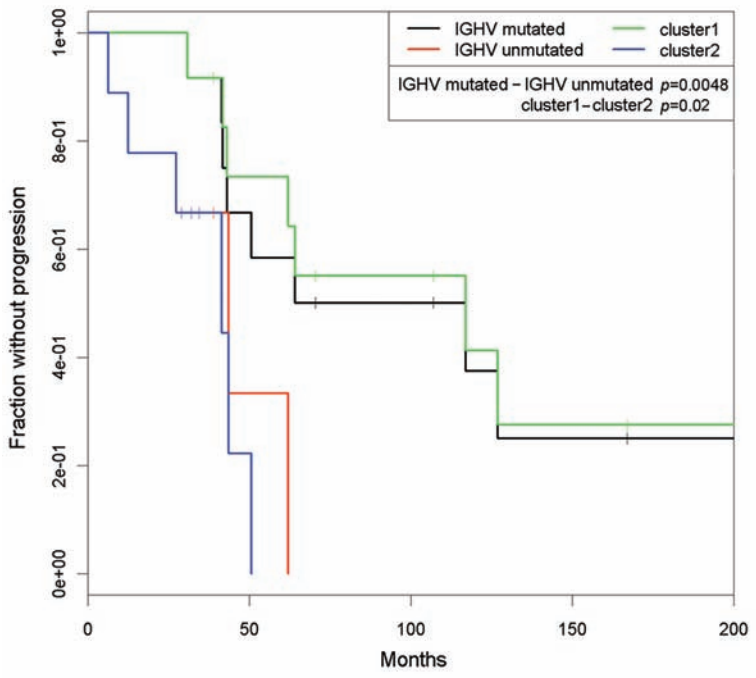
Online Supplementary Figure S1. Study schema.



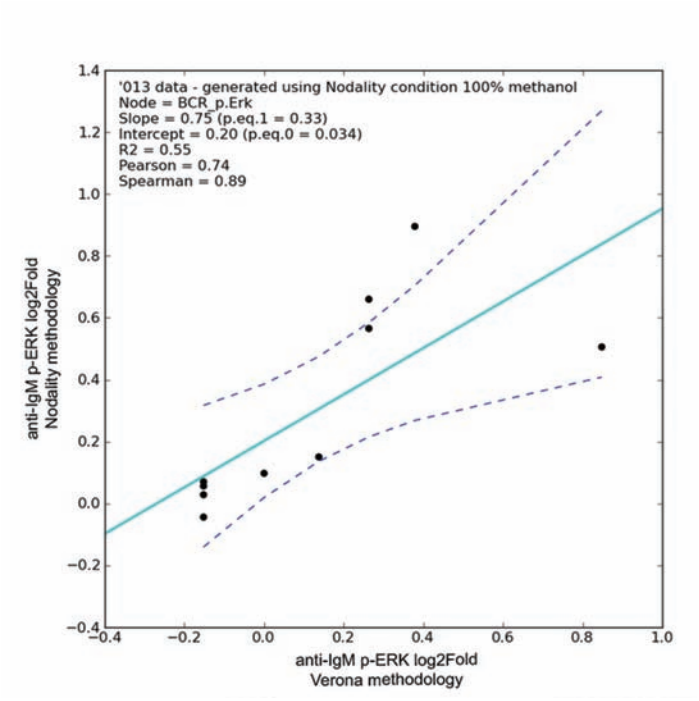
Online Supplementary Figure S2. Time course analysis of phosphoprotein signaling responses to anti-IgM stimulation. Data are expressed as $\log_2$Fold difference in relative median fluorescence intensity (RMFI) of anti-IgM-modulated cells divided by unmodulated cells ($\log_2$Fold change).



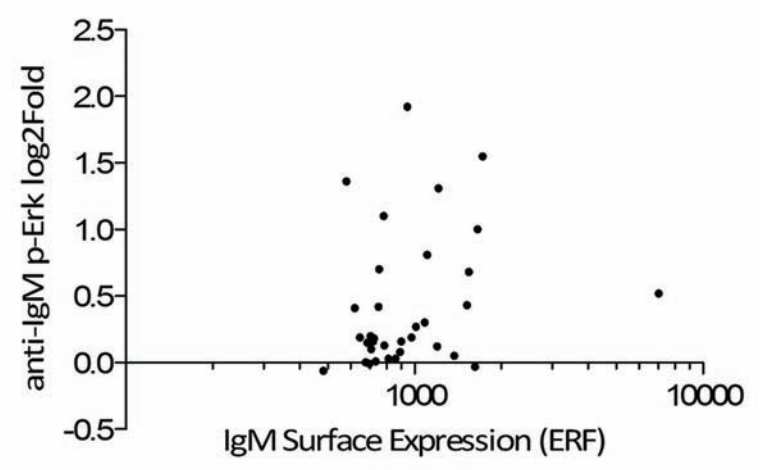
Online Supplementary Figure S3. Linear regression of the repeatability of the anti-IgM→p-ERK | log_2Fold (A) and anti-IgM→p-ERK | Uu (B) node-metric. Data show the correlation, slope, intercept and goodness of fit.



Online Supplementary Figure S4. Comparison of Kaplan-Meier curves of TTFT for subgroups of Binet stage A patients defined by BCR response to anti-IgM (clustering) and *IGHV* mutational status in the exploratory study. The curves are truncated (the last event is at 421 months). *P* values are from the log-rank test.



Online Supplementary Figure S5. Assay concordance in the bridging study comparing assay readouts of the anti-IgM→p-ERK | log_2Fold node-metric. Data show the correlation, slope, intercept, and goodness of fit.



Online Supplementary Figure S6. Correlation between the p-ERK signal induced by cross-linking IgM and the surface expression of IgM on B-CLL cells. Samples show a range of both surface expression of IgM immunoglobulin and anti-IgM→p-ERK response. The correlation did not reach statistical significance (Spearman r = 0.063). Each point represents one sample.