

Supplemental Materials

Seda et al. FoxO1-GAB1 Axis Regulates Homing Capacity and Tonic AKT Activity in Chronic Lymphocytic Leukemia

Correspondence:

Marek Mraz, M.D., Ph.D.
Associate Professor of Oncology
Central European Institute of Technology, Masaryk University
Kamenice 5, 625 00 Brno, Czech Republic
E-mail: marek.mraz@email.cz
Tel.: +420 549498143

1) Supplemental Methods

2) Supplemental Figures: 29

3) Supplemental Tables: 8

1) Supplemental Methods

CLL samples

Primary CLL samples were collected with written informed consent, and the Institutional Review Board approved the study. Peripheral blood samples were obtained from patients who did not receive any therapy for at least six months (except the ibrutinib/idelalisib treated patients). The samples from ibrutinib/idelalisib-treated patients were collected at a day before the drug administration and during therapy at the time points indicated in the figure legend. All primary CLL samples were purified by negative selection with RosetteSep Human B Cell Enrichment Cocktail (Stemcell Technologies) and RosetteSep Human CD3 Depletion Cocktail (Stemcell Technologies) and Ficoll-Paque PLUS (Sigma Aldrich) according to manufacturer's protocol and the final purity was >95% of CD5+19+ cells. In some cases PBMCs (Fig. 5E, samples CLL_113-118) were isolated and subsequently purified by magnetic separation using anti-CD3 MicroBeads (Miltenyi Biotec). Final purity was ≥95% of CD5+19+ cells in all cases (evaluated by flow cytometry). Purified primary CLL cells were cultivated in RPMI-1640

(Sigma-Aldrich) supplemented with 10% fetal bovine serum (Biosera) and 100 U·ml⁻¹ / 100 µg·ml⁻¹ of penicillin/streptomycin (Sigma Aldrich) and in 5% CO₂ at 37°C.

Cell culture

MEC1 cell line was obtained from the German Collection of Microorganisms and was cultivated in Iscove's Modified Dulbecco's Medium (IMDM, Biosera) supplemented with 10% fetal bovine serum (Biosera) and 100 U·ml⁻¹ / 100 µg·ml⁻¹ penicillin/streptomycin (Sigma Aldrich) and further cultivated in 5% CO₂ at 37°C. OSU-CLL cell line was a kind gift from Dr. Byrd (The Ohio State University) and was cultivated in RPMI-1640 media (Biosera) supplemented with 10% fetal bovine serum (Biosera) and 100 U·ml⁻¹ / 100 µg·ml⁻¹ penicillin/streptomycin (Sigma Aldrich) and further cultivated in 5% CO₂ at 37°C. HS5 cell line as well as HEK293-FT cell line were obtained from the German Collection of Microorganisms and both were cultivated in Dulbecco's Modified Eagle's Medium (DMEM, Biosera) supplemented with 10% fetal bovine serum (Biosera) and 100 U·ml⁻¹ / 100 µg·ml⁻¹ penicillin/streptomycin (Sigma Aldrich) and further cultivated in 5% CO₂ at 37°C.

Gene expression analysis

Total RNA was isolated by TRIzol (Molecular Research Center) as described previously.¹ Protein-coding gene expression analyses was performed using TaqMan Gene Expression Assays (Thermo Fisher Scientific) and normalized to an endogenous control (*GAPDH*, *HPRT*), as described previously.² Probes used in this study: *GAB1* (Hs00157646_m1), *FoxO1* (Hs00231106_m1), *NCK2* (Hs02561903_s1), *MYLK* (Hs00364926_m1), *GPR183* (Hs00270639_s1), *FLNA* (Hs00924645_m1), *SERPINF1* (Hs01106937_m1), *SEMA4F* (Hs01586592_m1), *HPRT1* (Hs02800695_m1), *GAPDH* (4333764F), mouse *CXCL12* (Mm00445553_m1), mouse *CXCL13* (Mm00444534_m1), mouse *HPRT* (Mm03024075_m1).

mRNA-seq and data analysis

The sequencing of poly(A) transcriptome of CXCR4^{dim}CD5^{bright} and CXCR4^{bright}CD5^{dim} CLL intracлонаl subpopulations was performed as described previously.³ Briefly, libraries were prepared using TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) according to the manufacturer instructions. The final libraries were diluted to the desired concentration of 2 pM and were used to prepare a paired-end flow cell using a cBot System (Illumina) and then sequenced with Illumina NextSeq 500/550 High Output v2.5 kit (Illumina) generating 75-bp paired-end sequences.

The quality of the sequencing data was checked using the FastQC⁴ (v0.11.3) and Kraken package⁵ (v13-274). The presence of adapters was scanned using Minion and Swan (Kraken package,⁵ v13-274). Pre-processing of the raw reads was done using Trimmomatic⁶

(v0.36) in the following steps (i) *N* and very low-quality bases (Phred < 3) from both 5' and 3' ends were removed, (ii) Poly(A) and poly(T) tails were removed, (iii) Sequencing adapters were removed, (iv) Reads were trimmed to a maximal length of 75 bp, (v) Low quality ends with average Phred score < 5 of 4 consecutive bases were trimmed using the sliding window approach, (vi) Reads shorter than 15 bp and without a proper pairing after the pre-processing were removed. The pre-processed reads were then mapped to the *Homo sapiens* reference genome (Ensembl⁷ release 84, primary assembly) together with corresponding Ensembl genome annotation (Ensembl release 84) using STAR⁸ (v2.5.2a) in 2-pass mapping approach. Exon-mapped reads were summarized to genes using STAR (a similar approach to the HTSeq-count⁹ 'union') . Only uniquely mapped reads were used. Strandedness of the sequencing protocol was taken into account and 'reverse' counts were used. Differential gene expression was analyzed using DESeq2¹⁰ (v1.13.8) Bioconductor¹¹ package for genes with a base mean count >200 per million. P-values were adjusted for multiple testing using the Benjamini-Hochberg method. The fold change was calculated as CXCR4^{dim}CD5^{bright}/CXCR4^{bright}CD5^{dim} cells. The heatmap visualization was created using the Heatmapper tool.¹¹

siRNA transfection

MEC1 cells or freshly isolated primary CLL cells (5x10⁶) were electroporated (Neon Transfection System, Thermo Fisher Scientific (program for MEC1 cells: 1500 V / 20 ms / 2 pulses; program for primary CLL cells: 1900V / 30ms / 1 puls) with a short interfering RNA (siRNA) against *GAB1*, or against *FoxO1* (both On-Target Plus Smartpool siRNAs, 1000 nM, Dharmacon), or by On-Target Plus Smartpool Negative Control (1000 nM, Dharmacon). The cells were harvested 48 hrs after transfection for further analyses.

Flow cytometry and cell sorting

The primary CLL cells or cell lines were labeled with one or combination of these antibodies: CD5 (FITC, UCHT2; Sony), CD184 [CXCR4] (PE, 12G5; Sony), CD19 (PECy7, JC-119, Beckman Coulter), CD45 (BV510, HI30; SONY), CD105 (A647, 43A3; SONY) and labeled with SYTOX-BLUE (Invitrogen) for viability. For intracellular staining, the cells were further fixed using 4% paraformaldehyde (PFA) and 50% methanol and stained with primary rabbit antibodies against *GAB1* (#3232, 1:100; Cell Signaling) or *FoxO1* (#2880, 1:100; Cell Signaling) and secondary anti-rabbit antibody conjugated with Alexa-647 (#4414, 1:500; Cell Signaling). Data were collected from at least 20,000 events using BD FACSVerser and analyzed using FlowJo software (version 10). Normalised expression of intracellular proteins was calculated as MFI of the protein of interest relative to MFI for secondary antibody only.

Cell viability was investigated by DiOC6 (3,3'-dihexyloxacarbocyanine iodide) together with PI (propidium iodide) staining (Thermo Fisher Scientific). All measurements were performed on FACS Verse (BD Biosciences-US) and cell sorting on BD FACSAria Fusion (BD Biosciences-US).

Calcium flux assay

The primary CLL cells (1×10^6 /ml) were incubated in Hanks Balanced Salt Solution (HBSS) containing 2 mM Fluo-4AM (Molecular Probes) for 30 min at 37°C. We established a baseline fluorescence threshold for each sample (by acquiring samples for 60 seconds without any stimulus). We then calculated the peak Fluo-4AM signal and fluorescence intensity above baseline Fluo-4AM signal following stimulation.

HS5 coculture and conditioned media

HS5 cells were seeded at 2×10^6 in 75 cm² culture flasks in DMEM with 10% FBS. The next day, DMEM medium was replaced by IMDM or RPMI media containing 10% FBS and 2.5×10^6 /ml of MEC1 cells (in IMDM) or 5×10^6 /ml primary CLL cells (in RPMI). The HS5 coculture was allowed for the specified time and then harvested and stained with Sytox blue and CD105 antibody. The purity of B cells after sorting was validated by anti-CD19 antibody. The conditioned media was harvested after cultivating confluent HS5 cells in DMEM with 10% FBS for 48 hrs.

To prepare conditioned media, 3.5×10^6 of HS5 cells was seeded in T75 cultivation flask (TPP) in 12 ml of DMEM media. The next day, original media was discarded, and the cells attached to the bottom of the flask were washed once with 10 ml of PBS. Subsequently, 25 ml of RPMI media was added to the attached HS5 cells. The conditioned media was collected after 48 hrs, separated from cell-debris by 10 minutes centrifugation at 2000 RCF and diluted with fresh RPMI media containing 10% FBS and $100 \text{ U} \cdot \text{ml}^{-1}$ / $100 \text{ } \mu\text{g} \cdot \text{ml}^{-1}$ penicillin/streptomycin to the final concentration of 50% conditioned media. This media was immediately used for experiments.

Primary human mesenchymal stromal cells and conditioned media

Human bone marrow was obtained from a healthy male donor with a written consent. Cells were resuspended in osmolysis buffer and spun (300g for 10 min). This step was repeated till all erythrocytes were removed. Cells were subsequently washed with PBS and plated in alphaMEM media (Sigma-Aldrich) containing 10% FBS and $100 \text{ U} \cdot \text{ml}^{-1}$ / $100 \text{ } \mu\text{g} \cdot \text{ml}^{-1}$ of penicillin/streptomycin (Sigma-Aldrich). Cells were cultured at 37°C for 48 hrs and subsequently washed (3x) with PBS in order to remove non-adherent cells. Adherent cells were resupplemented in fresh media, and incubated for another 48 hrs. For all experiments

only cells within 2nd-6th passage were used. To prepare conditioned media, 4×10^6 of the primary human bone marrow mesenchymal stromal cells were seeded in T75 cultivation flask (TPP) in 12 ml of full alpha-MEM media. The next day, original media was discarded, and the cells attached to the bottom of the flask were washed once with 10 ml of PBS and supplemented with fresh 12 ml of alpha-MEM media. The conditioned media was collected after 72 hrs, separated from cell-debris by 10 min centrifugation at 2000 RCF and immediately used for experiment or directly frozen at -80°C .

Chemokine concentration measurements by ELISA

The levels of chemokines SDF1 (CXCL12) and CXCL13 (BLC/BCA-1) were measured in conditioned media produced by HS5 or by primary human bone marrow mesenchymal stromal cells (for conditioned media preparation see above), or in serum from CLL peripheral blood or paired serum sample from sternal puncture in CLL patients. All measurements were performed according to manufacturer's protocol (Quantikine, R&D Systems, Minneapolis, MN). The absorbance at 450 and 570 nm was measured by TECAN SPARK.

***In vitro* migration assay and F-actin polymerization**

The competitive migration assay was performed by staining of MEC1 cells or primary CLL cells with CFSE or FarRed CellTrace dye according to manufacturer protocol and incubated overnight. The next day, the stained control cells and GAB1-KO cells were analyzed for viability (stained with Sytox Blue; Thermo Fisher Scientific). The paired samples with low viability (<80%) were excluded from further experiments as well as samples where the difference in viability between paired samples was higher than 5%. To compare the migration, the stained cells were mixed in ratio 1:1 and loaded to transwell inserts with 5 μm pores (Falcon) according to manufacturer protocol. The migration towards SDF1 (100 ng/ml for CLL cells or 250 ng/ml for MEC1 cells), CXCL13 (250 ng/ml for CLL cells or 500 ng/ml for MEC1 cells), 100% conditioned media produced by primary human bone marrow mesenchymal stromal cells, or 50% conditioned media produced by HS5 cells was allowed for the indicated time and subsequently analyzed by flow cytometry. The F-actin polymerization was performed by staining with Phalloidin-iFluor488 (Abcam) and visualized by EVOS Cell Imaging Systems (Thermo Fisher Scientific).

Plasmid transfections, CRISPR/Cas9, transduction of shRNAs

The GAB1^{MYC} plasmid was kindly provided by Armelle Yart¹² and empty control plasmid was a gift from Adam Antebi (Addgene plasmid #52535). pSpCas9(BB)-2A-GFP (PX458) was a gift from Feng Zhang (Addgene plasmid #48138;).

The gRNA sequences targeting the second exon of the *GAB1* gene or first exon of *FoxO1* gene were designed in Benchling (Biology Software, 2019) online tool. The sequence 5'-CTACTTGGTAGCAGACAGCG-3' for GAB1-gRNA or 5'-ATCCACATCGAGGCTCCTCG-3' for FoxO1-gRNA was cloned into pSpCas9(BB)-2A-GFP vector (#48138, Addgene). To obtain the mutant variant, the MEC1 wild-type cells (MEC1^{WT}) cells were electroporated as described above for siRNAs. After 48 hrs viable cells were sorted according to GFP-positivity, expanded, and validated for absence of GAB1 or FoxO1 .

To perform inducible shRNA expression, the sequence 5'-CCGGAGTTAACACACTCGTAGTATTCTCGAGAATACTACGAGTGTGTTAACTTTTTTG-3' was purchased from Sigma Aldrich and subsequently cloned into the empty plasmid Tet-pLKO-puro for shRNA expression.¹³ This plasmid was purchased from Addgene (Addgene plasmid #21915) as well as the envelope plasmid pCMV-VSV-G (Addgene plasmid #8454). The packaging plasmid dR8.91 was kindly provided by M. Smida (Masaryk University). Virus particles were produced in HEK-293-FT (obtained from ATCC) after plasmids transfection by Dharmacon (Horizon) and added to desired B cell lines. The transduced cells were subsequently selected by puromycin (3 µg/ml, Sigma Aldrich) for 5 days. The expression of shRNA in B cell lines was induced by adding tetracycline (1 µg/ml; Sigma Aldrich) into media (tet+) for 48 hrs, and control cells were treated with vehicle (tet-).

For *in vivo* growth assay MEC1^{WT} and MEC1^{GAB1-KO} cell lines were transduced by GFP or Azurit encoding vectors (both obtained from Addgene; #36083 and #36086, respectively) as described above. Cells were subsequently sorted for GFP or Azurit positive cells.

Immunoblotting

Cells were lysed in lysis buffer (1% SDS, 50 mM TRIS-HC pH 6.8, 10% glycerol) with phosphatase and protease inhibitors (Sigma Aldrich) and protein concentration was determined using DC Protein Assay (BioRad). Equal amounts of protein were separated by SDS-PAGE and transferred to the PVDF membrane (0.45 µm pore size, Millipore). The membranes were incubated with the Cell Signaling antibodies specific for the following immobilized proteins: GAB1 (#3232, 1:2000), pGAB1^{Y307} (#3234, 1:2000), FoxO1 (#2880, 1:1000), pFoxO1 (#9461, 1:2000), AKT (#2920, 1:3000), pAKT (#4060, 1:3000), BCL_{XL} (#2764, 1:2000), SYK (#12358, 1:2000), PLC γ (#3872, 1:2000), CD79 α (#3351, 1:2000), ERK (#4696, 1:3000) pERK (#4377, 1:3000), B-actin (#4970, 1:5000), GAPDH (#2118, 1:4000), or with antibodies obtained from Abcam - pGAB1^{Y627} (ab131458, 1:2000), anti-CD20 antibody (Ab78237, 1:2000). Secondary horse-radish peroxidase (HRP)-conjugated anti-mouse or anti-rabbit antibodies (both Cell Signaling) were used to detect primary antibodies. Immunocomplexes were detected using ECL (BioRad), and the chemiluminescent signal was digitally detected with UVItec Alliance 4.7 (UVItec).

ChIP

Chromatin immunoprecipitation was performed as previously described by Carey *et al.*¹⁴ Briefly, 20x10⁶ MEC1 cells were fixed in 1% paraformaldehyde. The fixation was quenched by glycine after 10 min. After washing with cold PBS, the fixed cells were lysed in 10ml of ChIP-lysis buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40) and incubated on ice for another 10 min. The cell nuclei were pelleted by centrifugation and subsequently lysed in 1ml of nuclei-lysis buffer (50 mM TRIS-Cl pH 8.0, 10 mM EDTA, 1% SDS) containing protease inhibitors. Sonication was performed by the Covaris S220 sonicator in original glass tubes. The instrument was set according to manufacturer protocol (PIP:140, DF: 5, CBP: 200, treatment time: 20 min, temperature min/max: 3/9°C). The cell debris was subsequently discarded by centrifugation and DNA concentration was measured by Nanodrop 2000 (Thermo Fisher Scientific). One-hundred µg of chromatin was aliquoted per antibody and diluted with dilution buffer (16.7 mM TRIS-Cl pH 8.0, 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% TRITON X-100, containing protease inhibitors) to a final volume of 300 µl. Chromatin was precleared with 50 µl of protein A/G ChIP-grade magnetic beads (Thermo Fisher Scientific). The magnetic beads were discarded after 2 hrs of incubation, and 10 µl of anti-FoxO1A antibody (ab39670; ChIP Grade, Abcam) or 10 µl of control rabbit IgG (#2729; Cell Signaling) was added to chromatin. The samples were rotated and incubated at 4°C overnight. The next day, the immunocomplexes were pulled down by 50 µl of A/G ChIP-grade magnetic beads (Thermo Fisher Scientific) blocked by salmon sperm and BSA. The beads were washed three times in 1ml of High-salt wash buffer (50 mM HEPES pH 7.9, 500 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% TRITON X-100, 0.1% deoxycholate). The elution was performed in 300 µl of Elution buffer (50 mM TRIS-Cl, 10 mM EDTA, 1% SDS) supplemented with 20 µg of proteinase K. Samples were incubated for 2 hrs at 55°C and then overnight at 65°C to reverse crosslinking. The DNA was purified by QIAquick PCR Purification Kit (Qiagen). The qRT-PCR was performed by using Luna Universal qPCR Master Mix (New England Biolabs), and three different sets of primers (see supplemental Table 5), which were designed using Primer3 online tool. The prediction of possible FoxO1-binding sites in the 3.5 kb promoter region of the *GAB1* gene was performed with ContraV3 online tool. The non-specific primer-set was designed into the 2nd intron of *GAB1* where there is no predicted binding site for FoxO1 (forward: 5'-CAGTCAGTCTAGTAGGAGGGACAA-3', reverse: 5'-CTGAAAAGATCACTGCTGCTGCTA-3'). Two sets of primer were designed for the promoter region with predicted binding sites for FoxO1 (TBS no. 4: forward 5'-TTTGGGGCTTTTTACAAAT-3', reverse 5'-TGCCTACTTCCTTGTGTGTTTC-3'; TBS no.10: forward 5'-ATGCTGCCTATGGGAGAAGA-3', reverse 5'-AAACCAAAGCACTTCTTTTTAAGT-3'). The results were calculated as % of input.

Confocal microscopy

The MEC1 cells were plated on poly-L-lysine coated (0.01%) glass for 30 min. Glass-attached MEC1 cells or primary CLL cells in suspension were subsequently fixed with 1% paraformaldehyde for 15 min/room temperature and washed (2x) with TBS-T. The 1 hr long blocking step in 5% BSA was followed by overnight incubation in primary antibody against FoxO1 (#2880, 1:100; Cell Signaling). The next day cells were washed (3x) in TBS-T and incubated in a mixture of secondary antibody (anti-Rabbit, Alexa-647 #4414, 1:250; Cell Signaling) and Phalloidin-iFLuor488 (1:1000; Abcam) for 90 min/room temperature (in dark). Subsequently, three wash steps with TBS-T were followed by staining with DAPI (1:1000; 15 min; Abcam), and cells were mounted to glass slides using Vectashield (Vector laboratories), and visualized on confocal microscope Leica DM5500 Q with TCS SPE confocal laser scanning system.

Cell motility by holographic microscopy

Cell motility speed and other measurements were performed as described previously.¹⁵⁻¹⁷ Briefly, tissue culture treated μ -Slide (Ibidi) with 0.8 mm channel height was first filled with the culture medium to coat the substrate with serum proteins. After several minutes, the medium was removed and the μ -Slide channel was filled with 200 μ l of cell suspension containing 10 000 exponentially growing MEC1 cells in a conditioned medium to achieve a subconfluent culture. Cells were incubated for 1 h at 37°C in a humidified atmosphere enriched with 5% CO₂. The μ -Slide with cells was then positioned on the stage of our Multimodal Holographic Microscope Q-PHASE (Tescan/Telight) and subjected to time-lapse recording for 1 h with 10 s lapse-interval using 10 \times objective lenses 0.3 NA using the holographic mode measuring the optical thickness of cells with an accuracy of 2 nm. The recording was repeated 3 times for each cell type on two different days producing 12 time-lapse sequences which were processed using in-house developed image processing procedure using Mathematica (Wolfram Research), which removes background fluctuations, reconstructs and subtracts the background, converts the optical thickness of cells into their dry mass distributions, segments the cells, calculates centroids of dry cell mass, and morphometric parameters of *Circularity*, *Elongation*, *Mass* and *Spreading* defined as:

$$Circularity = \sqrt{\frac{(4 \pi A)}{P^2}}$$

where A is cell area in μm^2 and P is the perimeter length in μm ;

$$Elongation = 1 - \left(\frac{Width}{Length}\right)$$

where *Width* and *Length* are the smallest and largest axes of the best fit ellipse respectively;

$$Spreading = A/Mass$$

where *Mass* is the total dry mass of a cell expressed in pg (picogram = 10^{-12} g).

The software also tracks the cells in time on the basis of the nearest cell with a similar dry mass in the next frame. Tracked centroids were processed in order to calculate *Displacements* and *Persistences*. *Persistence* was defined as the ratio of direct translocation in 60 s and the wiggly track distance was determined as the sum of six 10 s displacements. *The Persistence* of a cell moving in a straight line is 1, and bending of the trajectory reduces the *Persistence* towards 0.

Immunohistochemistry of paraffin-embedded tissues

Immediately after the tissue of experimental animals was dissected out, it was fixed in 4% paraformaldehyde pH 7.4 at room temperature overnight. The FFPE tissue sections (1 μ m) were automatically processed and stained in VENTANA BenchMark Ultra system. Universal DAB Detection Kit (Ventana) was used for staining in combination with UltraView Amplification Kit (Ventana). Counterstaining was achieved by using Hematoxylin II (Ventana) and post-counterstaining by using Bluing reagent (Ventana). Immunostaining was performed using the anti-CD20 (for MEC1 cells; M0701, DAKO) or anti-CD45 (for primary CLL cells; M0755, DAKO) antibody.

Immunohistochemistry of cryopreserved tissues

Immediately after the tissue of experimental animals was dissected out, it was fixed in 4% paraformaldehyde pH 7.4 at 4 °C for 5 hrs. Subsequently, liver and spleen were cryopreserved in 30% sucrose at 4 °C overnight and embedded in OCT medium (Tissue-Tek, 4583) on dry ice. Femurs were decalcified in 10% EDTA pH 7.4 for 7 days at 4 °C prior to cryopreservation and OCT embedding. Samples were sectioned on a cryostat (Leica CM1850UV) as 14- μ m thick slices and collected on SuperFrost Plus Adhesion slides (Thermo Fisher Scientific, J1800BMNZ). Staining with primary antibody anti-mouse collagen IV (COL4, AbD Serotec, 2150-1470; 1:500) was performed at 4 °C overnight, followed by incubation with Alexa-conjugated secondary antibody (Invitrogen; 1:1000) at room temperature for 2 hrs. Cell nuclei counterstaining was performed with DRAQ5 Fluorescent Probe Solution (Thermo Fisher Scientific, 62251), and slides were mounted with Fluoromount Aqueous Mounting Medium (Sigma–Aldrich; F4680). Imaging was performed using Zeiss LSM880 laser scanning confocal microscope. ZEN2.1 (ZEISS) and Imaris (Bitplane) software was used for image processing.

Statistical analysis



Differences in expression levels between paired samples were compared using paired t-test or Wilcoxon matched matched-pairs test and differences in non-matched groups were compared using unpaired t-test or Mann-Whitney U test. The appropriate tests were chosen based on the results of Shapiro-Wilk test of normality of input data. Statistical analysis of GAB1 inhibitors' effect on cell viability in comparison to their combination with ibrutinib was analysed by 2-way ANOVA including Tukey test. All analysis were performed in GraphPad v8.0.1.

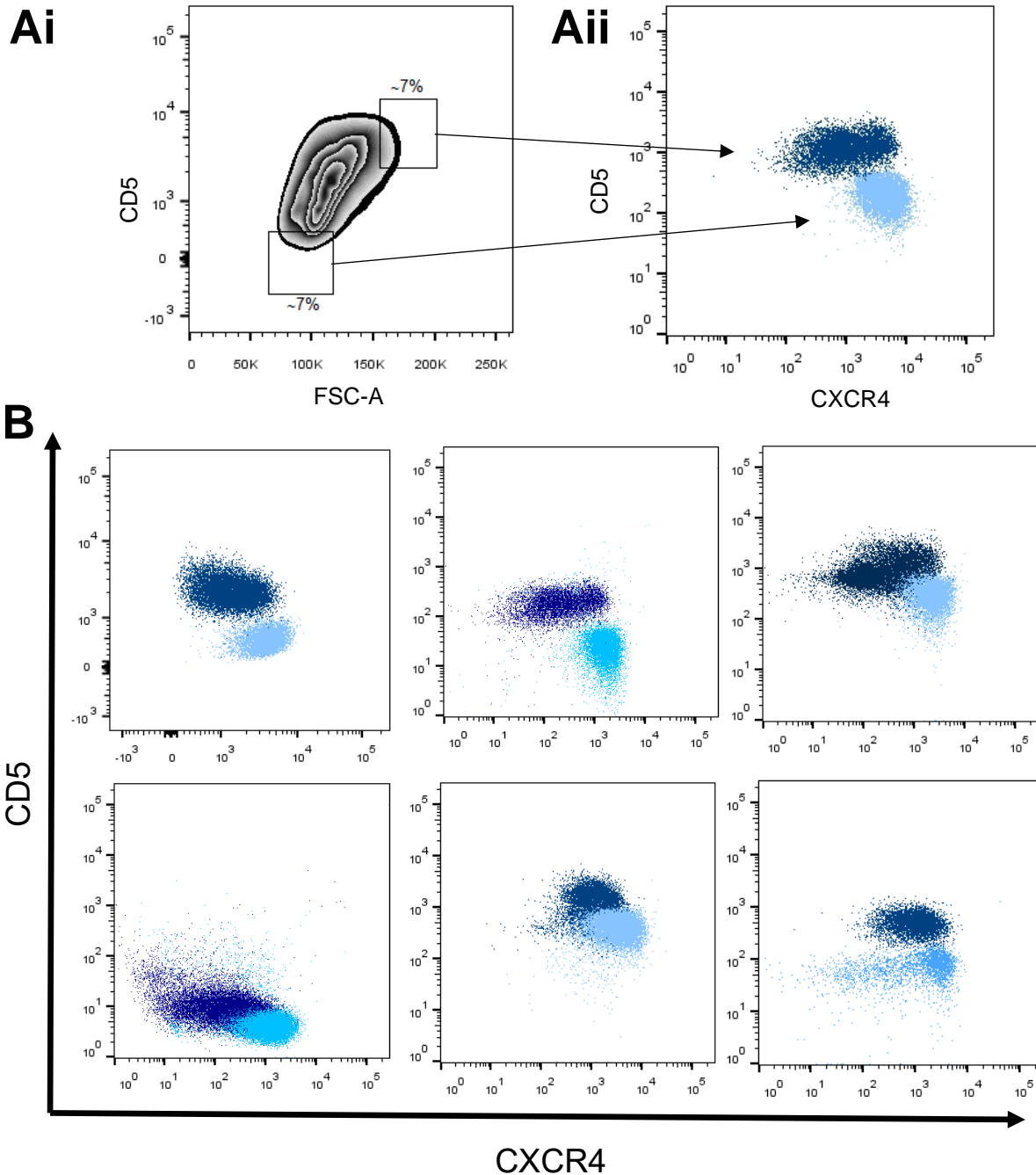
Data Sharing Statement

The results of the RNAseq analysis and data deposition may be found above. For other original data or detailed protocols, please contact the corresponding author.

2) Supplemental Figures

Supplemental Figure 1

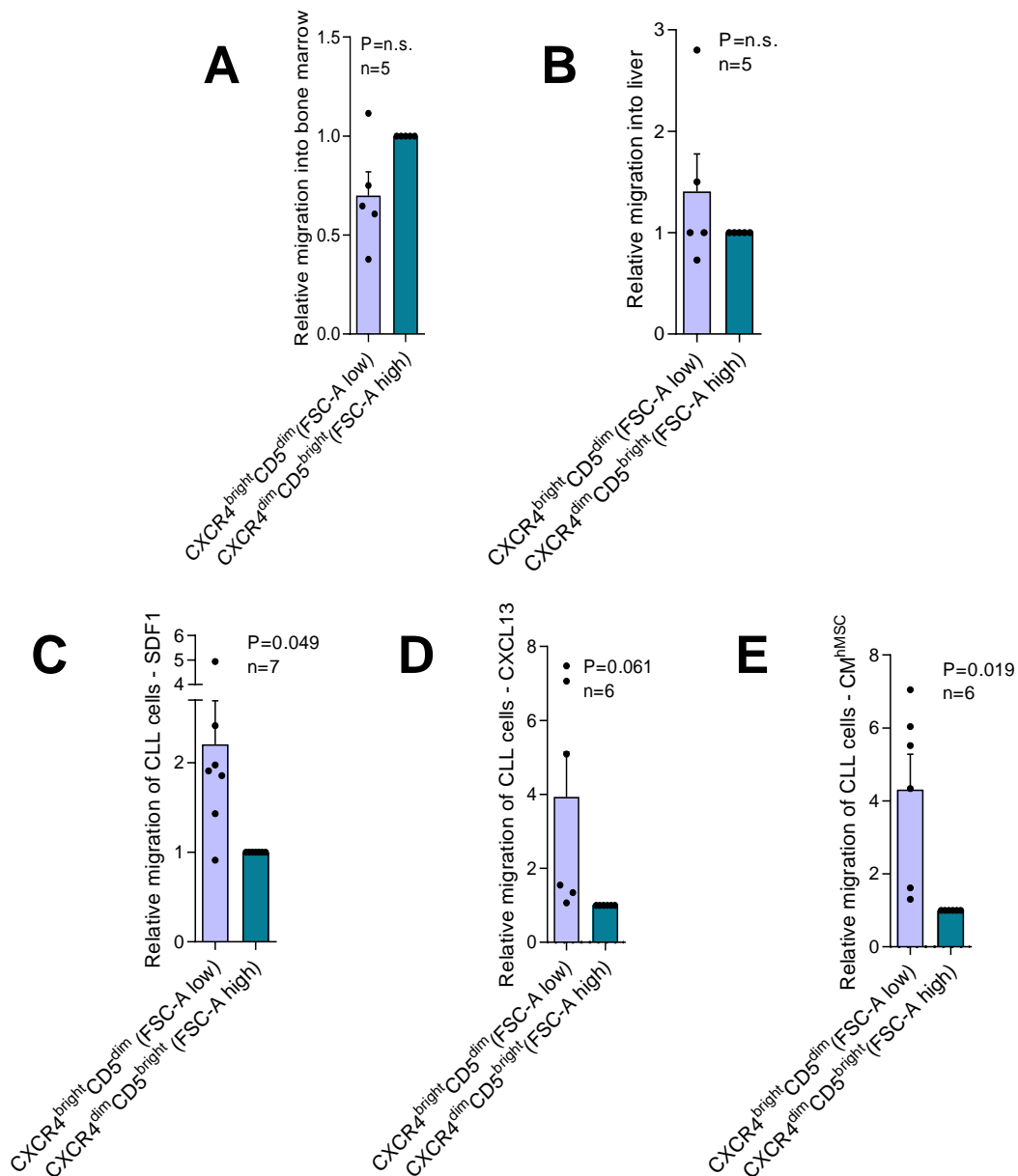
Color	Sample
	CD5 ^{dim} FSC-A ^{low}
	CD5 ^{bright} FSC-A ^{high}



Supplemental Figure 1. (Ai) Representative example of a gating strategy for sorting of CLL cells according to CD5 and FCS-A (forward scatter area). CXCR4/CD5 subpopulations were sorted according to CD5 and forward scatter area (CXCR4^{dim} cells have a higher FSC-A). The CD5/FSC-A was used to prevent labeling by an antibody binding to the CXCR4 receptor, which would interfere with the migration to SDF1 and potentially other chemokines. **(Aii)** Evaluation of CXCR4 and CD5 expression in cells sorted according to the gating strategy shown in [Ai]. Small aliquotes of sorted CD5^{dim}FSC-A^{low} and CD5^{bright}FSC-A^{high} cells

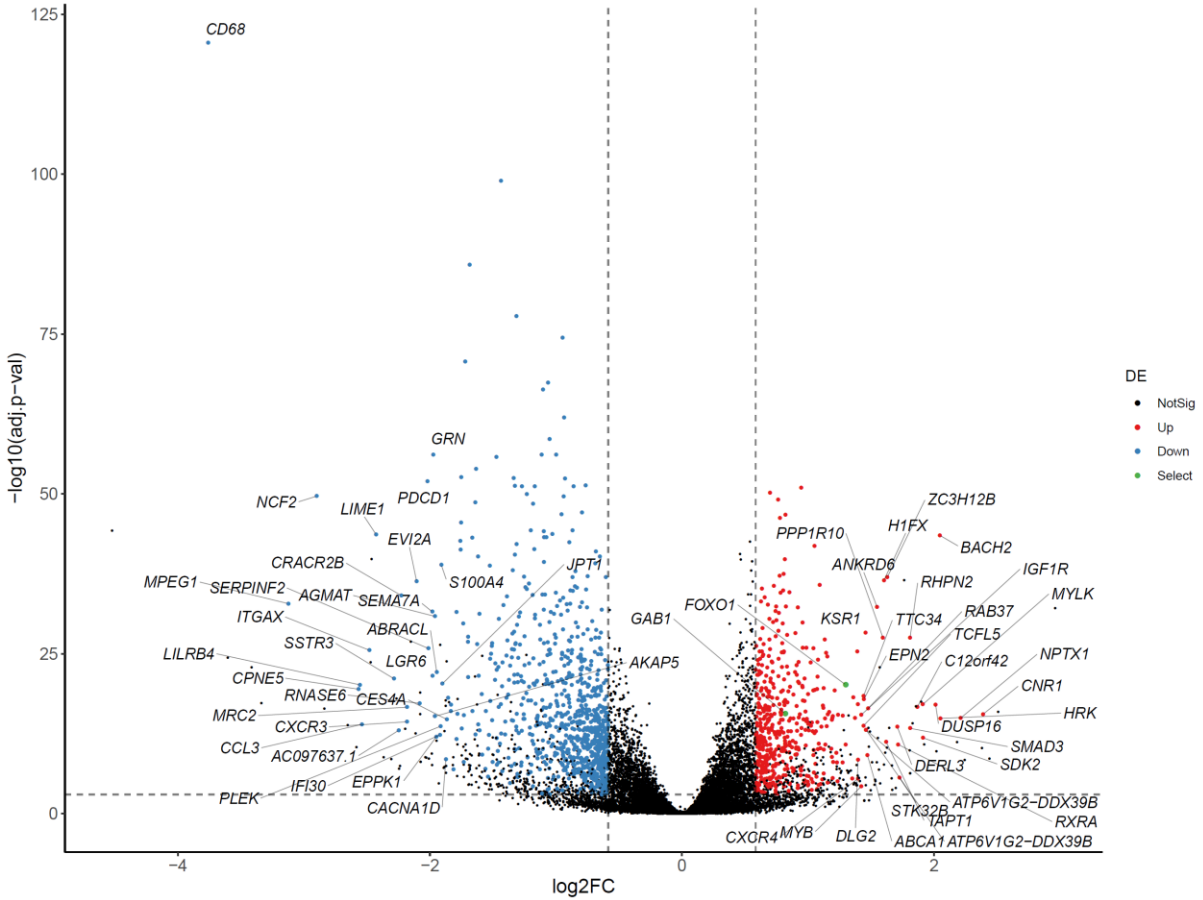
(correspond to CD5^{dim}CXCR4^{bright} and CD5^{bright}CXCR4^{dim} cells, respectively) were stained for CXCR4 and measured by flow cytometry to validate CXCR4 cell-surface levels and separation of CXCR4/CD5 subpopulations (this validation was performed in all samples used for experiments below). The sorting approach using CD5/FSC-A was successful in separating CXCR4/CD5 subpopulations in approx. 75% of CLL cases and only the samples with an acceptable profile of CXCR4/CD5 after validation by flow cytometry were used in further experiments. **(B)** Validation of CXCR4/CD5 expression in other CLL samples sorted according to FCS-A/CD5.

Supplemental Figure 2



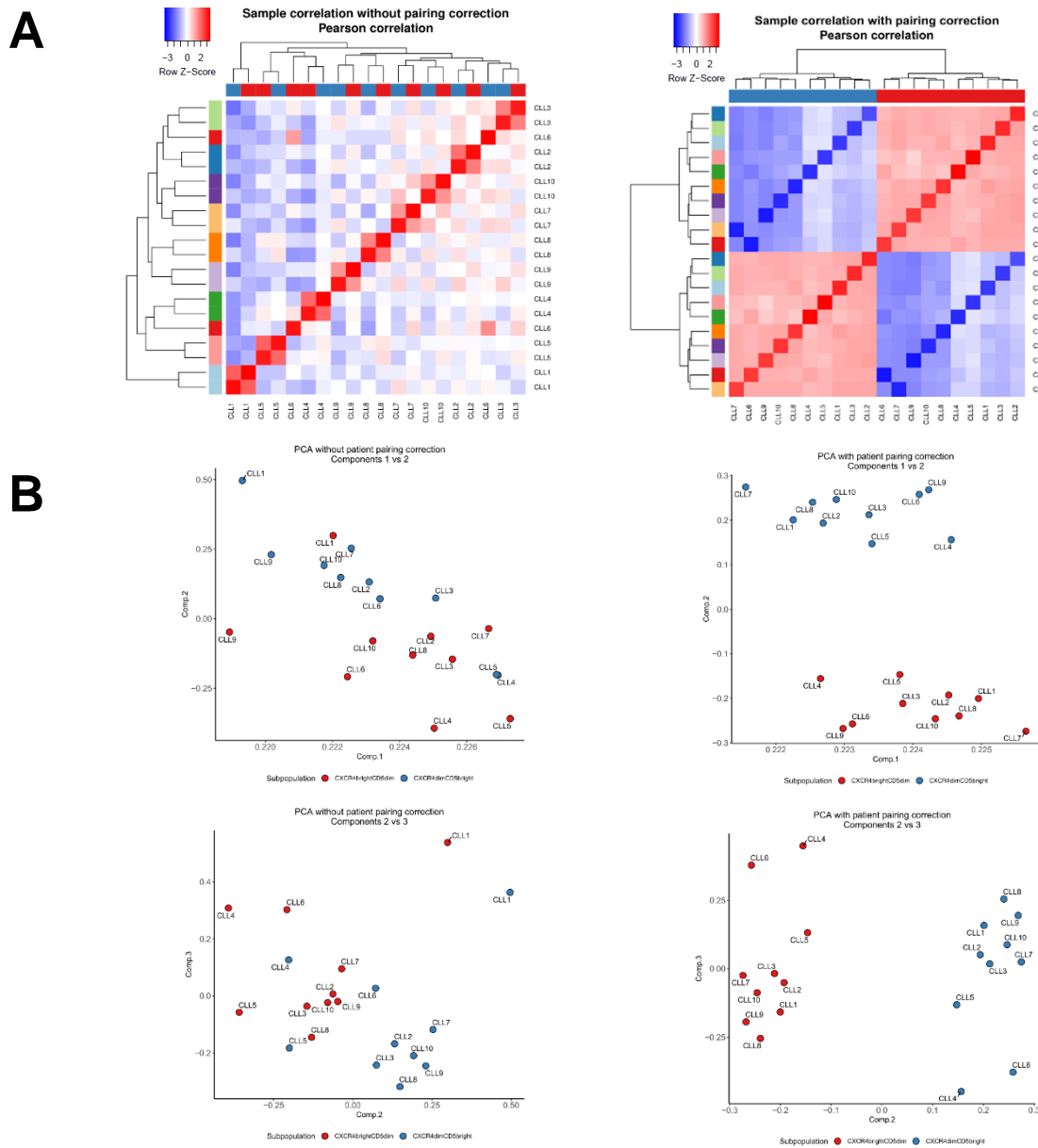
Supplemental Figure 2. Competitive migration assay of primary CLL cells sorted according to CD5/FSC-A (as surrogate markers for CD5/CXCR4 subpopulations, the cell-surface levels of CXCR4/CD5 were validated after sorting, see Supplemental Fig. 1). **(A-B)** Sorted cells were stained with two different CellTrace dyes, mixed in a 1:1 ratio, and injected into the tail vein of NGS mice (see Fig. 2A). Migration was allowed for 4 hrs, and then bone marrow **(A)** and liver **(B)** were analyzed for the presence of CLL cells (n=5). The numbers of CLL cells infiltrating bone marrow or liver were markedly lower in comparison to the spleen (Fig. 1A). **(C-E)** Competitive migration assay of CXCR4/CD5 CLL cell subpopulations towards SDF1 **(C)**, CXCL13 **(D)**, or conditioned media **(E)** produced by bone marrow mesenchymal stromal cells from a healthy donor (CM^{hMSC}). Sorted and validated CXCR4/CD5 subpopulations were stained with two different CellTrace dyes, mixed in a 1:1 ration, and loaded into transwells.

Supplemental Figure 3



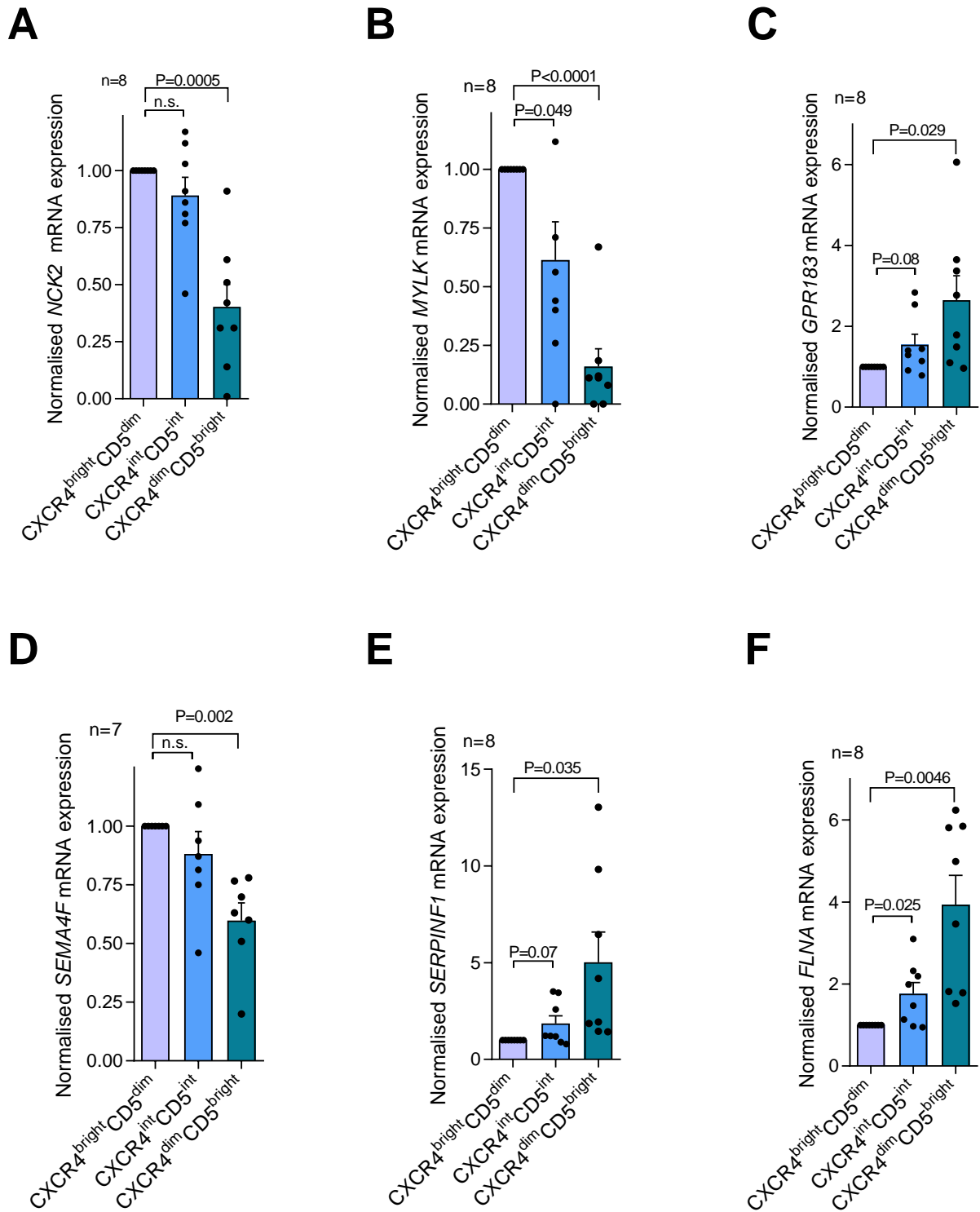
Supplemental Figure 3. Volcano plot depicting differentially expressed genes in CXCR4/CD5 subpopulations. Vertical dashed lines represent a fold-change of >1.5-fold. The horizontal dashed line represents adjusted P-value <0.001. Blue color represents genes relatively down-regulated in CXCR4^{bright}CD5^{dim} cells, and the red color represents genes relatively up-regulated in CXCR4^{bright}CD5^{dim} cells (cell sorting strategy shown in Fig. 1B). FoxO1 and GAB1 are depicted in green color. Log2FC stands for a fold-change in log2 scale, -log10(adj.p-val) stands for adjusted P-value in a log10 scale.

Supplemental Figure 4



Supplemental Figure 4. (A) Pearson correlation for gene expression between all samples used for RNAseq analysis. The correlation was calculated from log₂ normalized gene counts (DESeq) + 1 to avoid logarithms of 0 counts. Correlation heatmaps for both sample pairing uncorrected (top) and corrected for paired CXCR4/CD5 subpopulations (bottom) are shown. The heatmap colors were normalized to row Z-score. Red color depicts CXCR4^{bright}CD5^{dim} cells, and blue color depicts CXCR4^{dim}CD5^{bright} cells, row colors represent individual patients. The data were corrected for pairing by limma (v3.40.6) „removeBatchEffect“ function.¹⁸ **(B)** Principal Component Analysis (PCA) of all samples used for RNAseq analysis was calculated from log₂ normalized counts (DESeq1) + 1 to avoid logarithms of 0 counts. PCA components for both sample pairing uncorrected (left) and corrected for paired CXCR4/CD5 subpopulations (right) are shown. Red color depicts CXCR4^{bright}CD5^{dim} cells, and blue color depicts CXCR4^{dim}CD5^{bright} cells. The data were corrected for pairing by limma (v3.40.6) „removeBatchEffect“ function.¹⁸

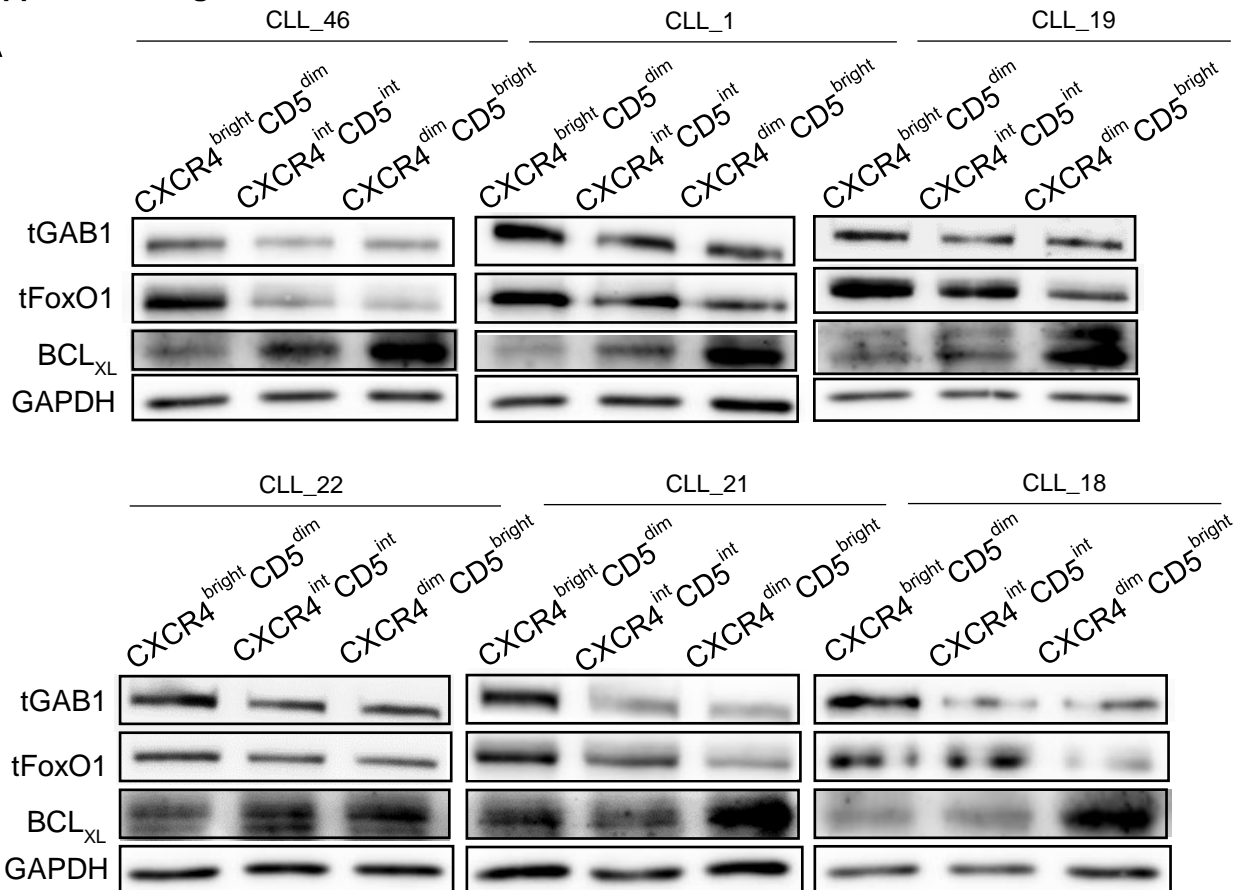
Supplemental Figure 5



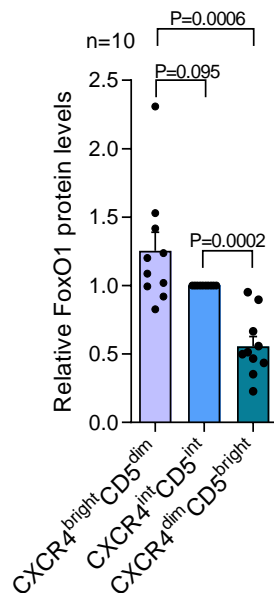
Supplemental Figure 5. Analysis of six selected migration-related genes by qRT-PCR in sorted CXCR4/CD5 subpopulations (for cell sorting strategy, see Fig. 1B). The mRNAs were randomly selected from differentially expressed migration-related genes identified by RNAseq (Fig. 1C), and their expression was validated in an independent set of sorted samples. The CXCR4^{int}CD5^{int} represents a transitional intermediate (int) subpopulation between CXCR4^{dim}CD5^{bright} and CXCR4^{bright}CD5^{dim} cells.

Supplemental Figure 6

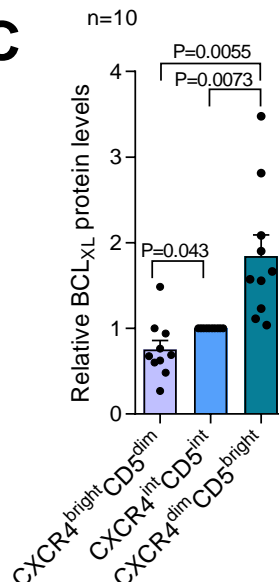
A



B

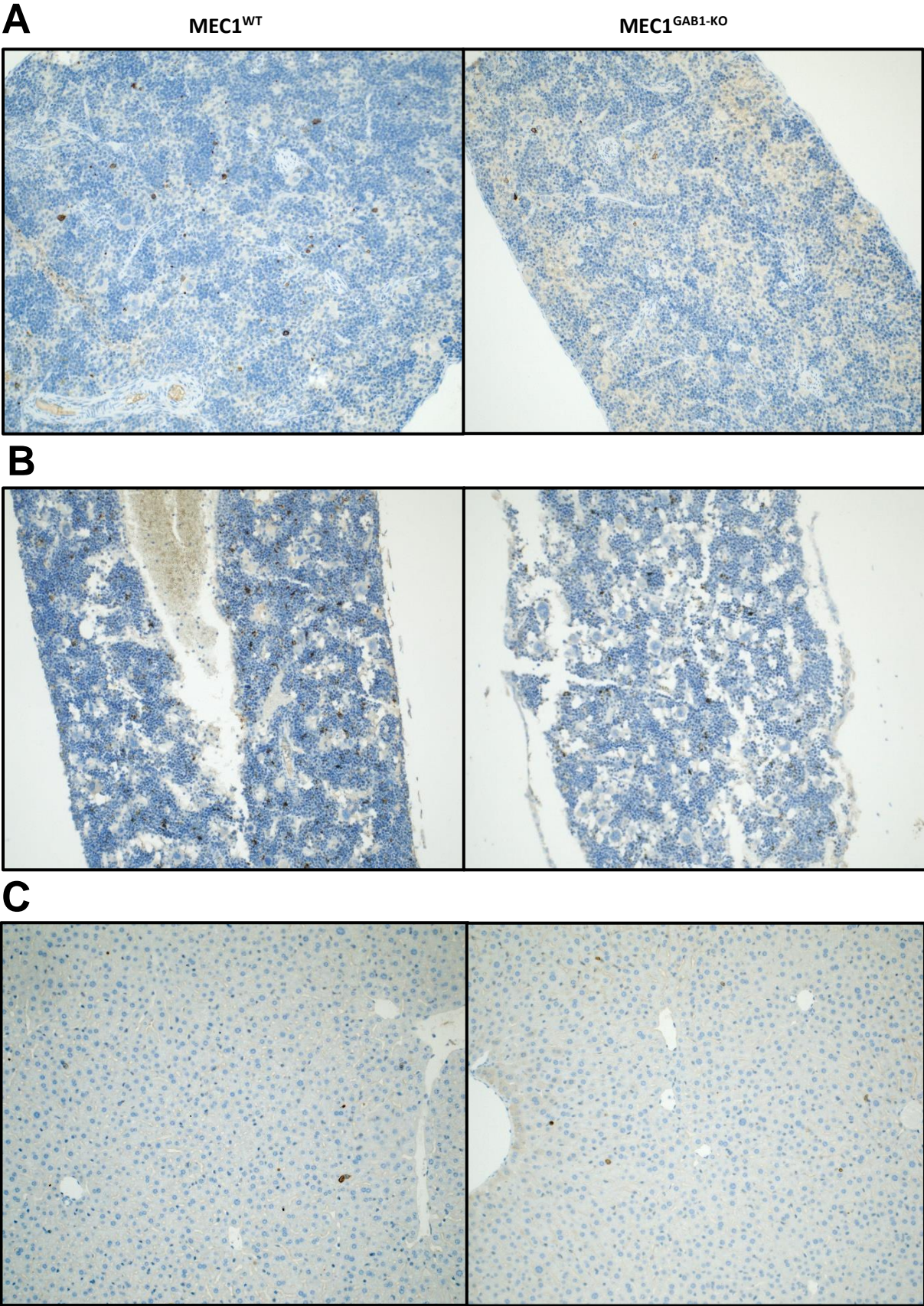


C

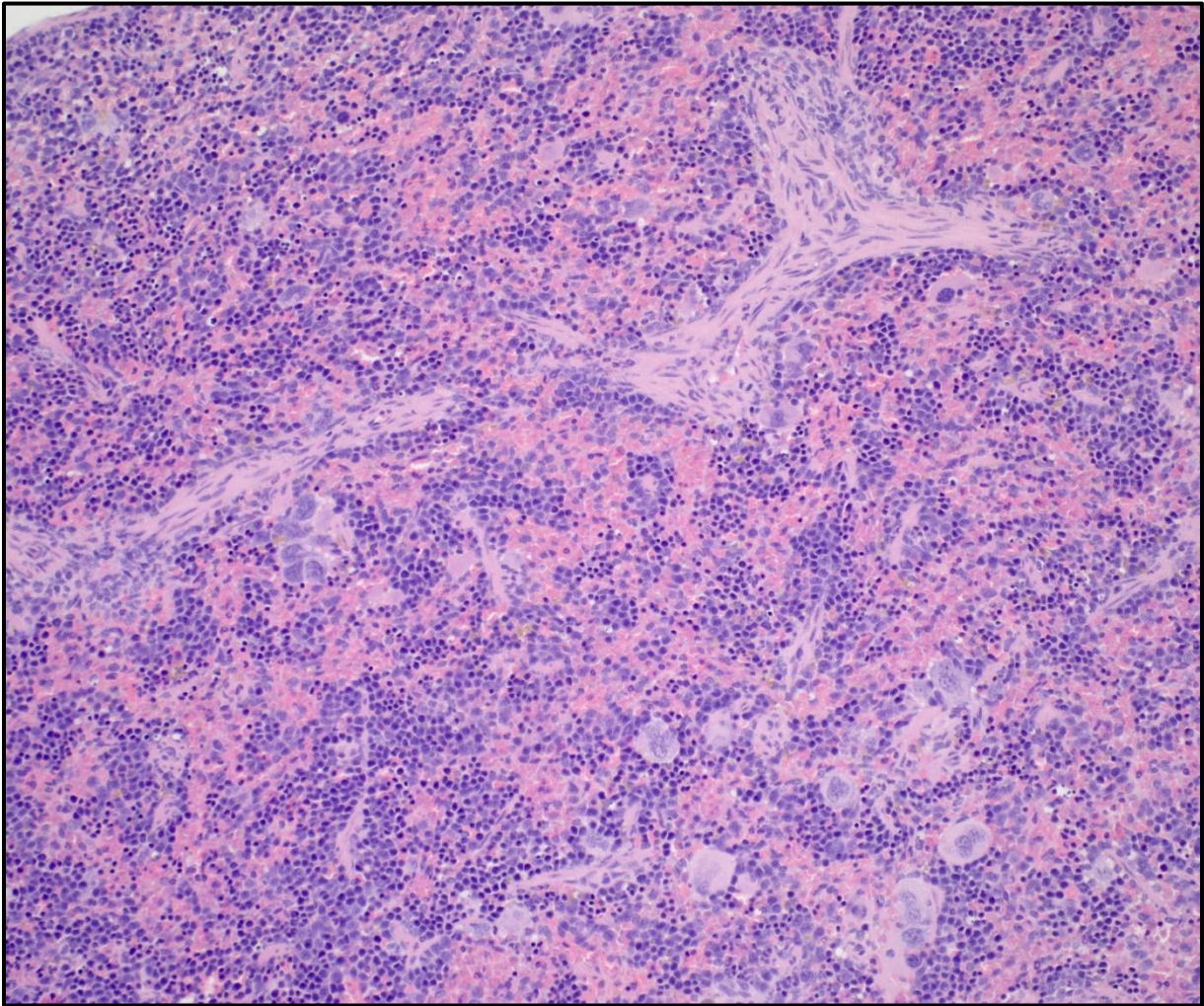


Supplemental Figure 6. (A) Additional immunoblots for proteins of interest from sorted CXCR4/CD5 subpopulations. The CXCR4^{int}CD5^{int} represents a transitional intermediate (int) subpopulation between CXCR4^{dim}CD5^{bright} and CXCR4^{bright}CD5^{dim} cells. **(B-C)** and Densitometric quantification of FoxO1 (B) and BCL_{XL} (C) protein levels in sorted CXCR4/CD5 subpopulations as analyzed by immunoblot. BCL_{XL}, an antiapoptotic protein, serves as a positive control for microenvironmental interactions. For cell sorting see Fig. 1B and for densitometric analysis of protein levels, see Fig. 1G

Supplemental Figure 7

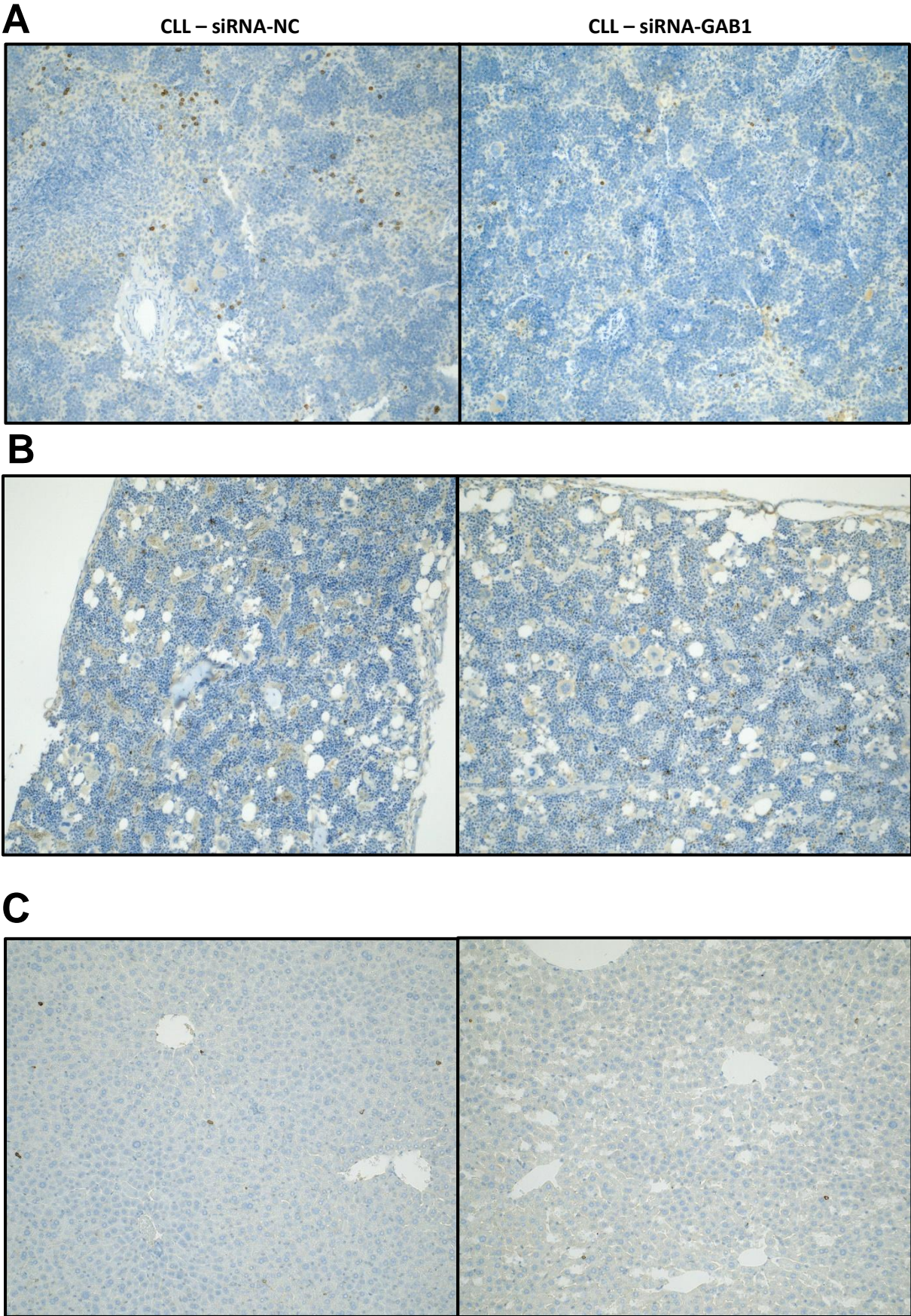


D



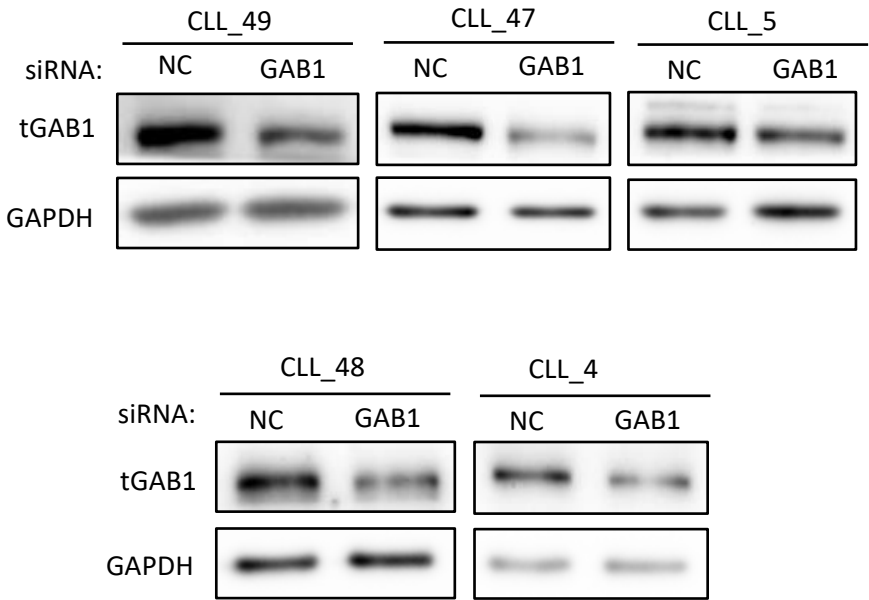
Supplemental Figure 7. Histological evaluation of spleen **(A)**, bone marrow **(B)**, and liver **(C)** of NSG mice transplanted with MEC1^{wt} (*left*) or MEC1^{GAB1-KO} (*right*) cells. In this experiment, MEC1^{wt} or MEC1^{GAB1-KO} cells were transplanted separately to individual NSG mice via a tail vein injection and sacrificed after 6 hrs. Tissue sections were stained with anti-human-CD20 antibody (brown signal) and the post-counterstaining by Bluing reagent (Ventana). The composition of the spleen structure in NSG mice is different from normal mouse spleen and reflects the absence of mature B/T/NK cells in these animals (see **D** for hematoxylin-eosin staining of the spleen, 20x). Therefore, the white pulp of the spleen is absent, and the organ is mainly formed by red pulp with a high content of overactive erythron and excessive erythropoiesis (including the presence of megakaryocyte). The migrated MEC1 cells were localized diffusely in the spleen, and there was no difference in tissue localization for GAB1-knockout versus GAB1-wt cells.

Supplemental Figure 8



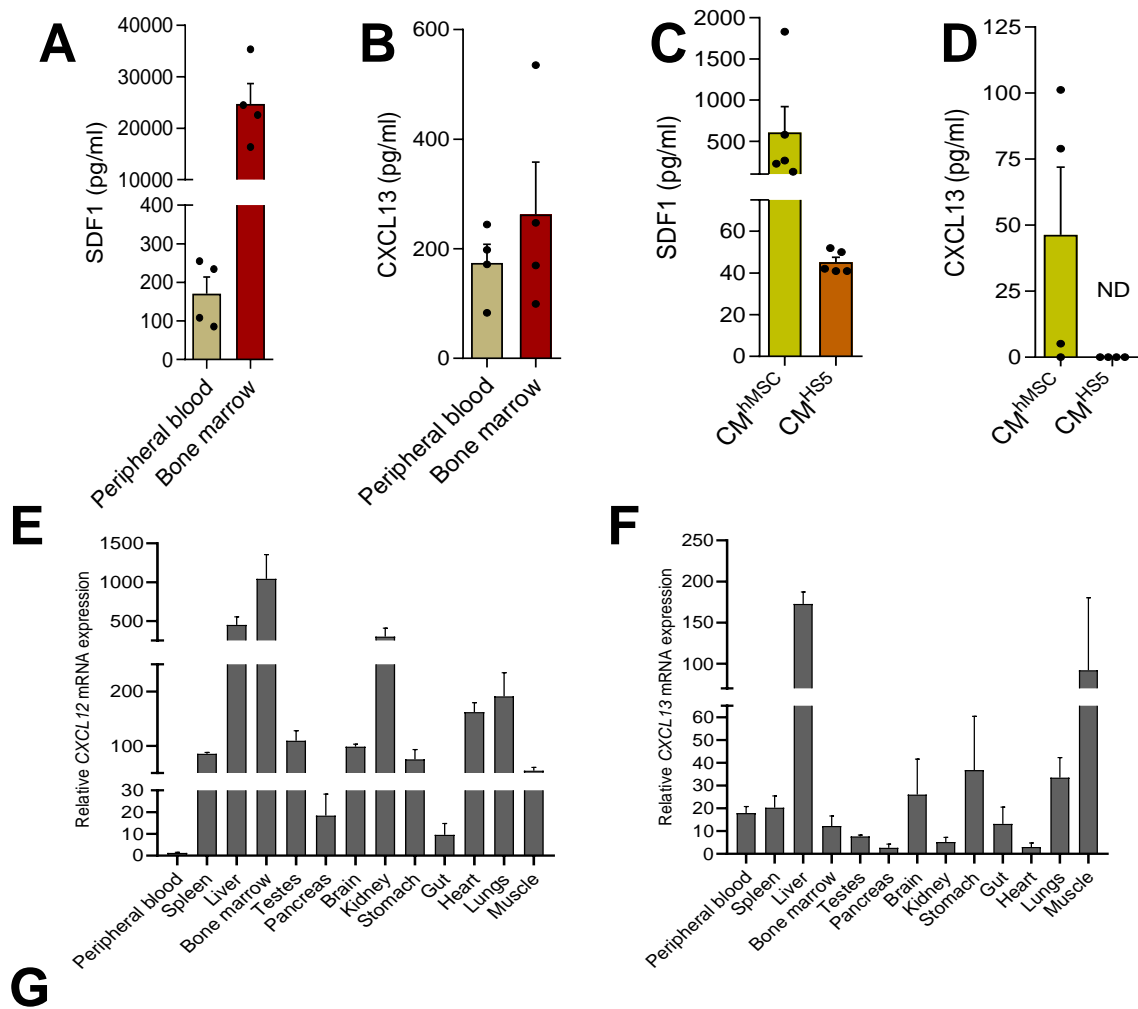
Supplemental Figure 8. Histological evaluation of spleen **(A)**, bone marrow **(B)**, and liver **(C)** of NSG mice transplanted with CLL cells transfected with siRNA against *GAB1* (siRNA-GAB1, *right*) or control siRNA (siRNA-NC, *left*). Cells were incubated for 48 hrs after transfection and then stained with Sytox Blue viability dye and sorted according to viability. Viable CLL cells (10×10^6 of siRNA-GAB1 or siRNA-NC transfected cells) were transplanted to individual NSG mice via a tail vein injection and sacrificed after 6 hrs. Tissue sections were stained with anti-human-CD45 antibody (brown signal) and the post-counterstaining by using Bluing reagent (Ventana). The composition of the spleen structure in NSG mice is different from normal mouse spleen and reflects the absence of mature B/T/NK cells in these animals (see Supplemental Fig. 10D for hematoxylin-eosin staining of the spleen, 20x). Therefore, the white pulp of the spleen is absent, and the organ is mainly formed by red pulp with a high content of overactive erythron and excessive erythropoiesis (including the presence of megakaryocyte). The migrated CLL cells were localized diffusely in the spleen, with a tendency for higher CLL cell infiltration in the trabecular system near the trabecular veins; however, there was no difference in the localization inside the studied organs for siRNA-GAB1 or siRNA-NC CLL cells. For patient characteristics, see Supplemental Tab1, CLL_3.

Supplemental Figure 9



Supplemental Figure 9. Control immunoblots from primary CLL cells transfected with control siRNA (NC) or siRNA against *GAB1* (GAB1). Cells were incubated for 48 hrs and then stained with Sytox Blue viability dye and sorted according to viability. A small aliquot of sorted cells, was used for these control immunoblots, and the remaining cells were used for competitive migration assay in NSG mice.

Supplemental Figure 10

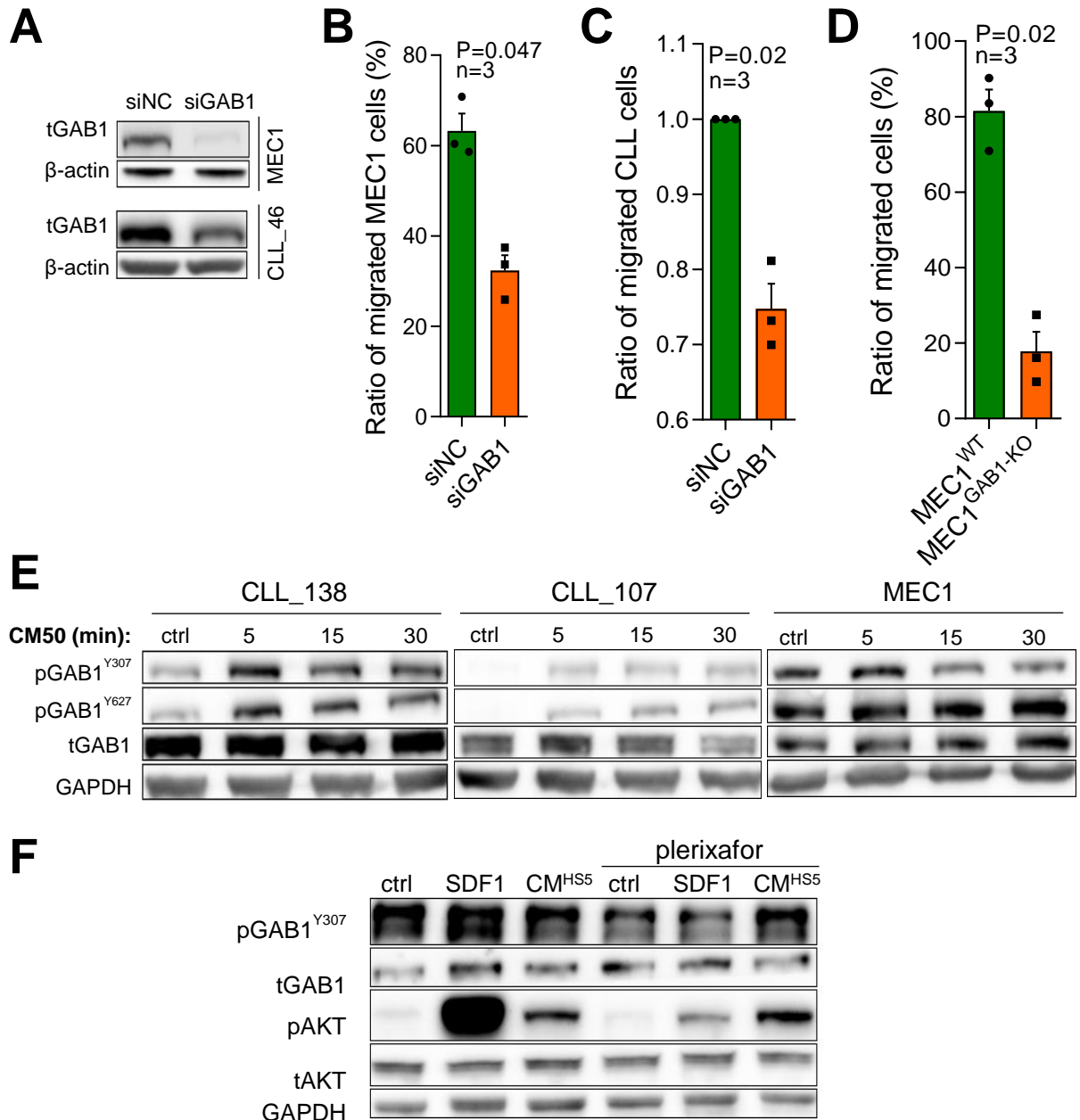


G

Gene	Peripheral blood cells	Spleen	Liver (neonatal)	Bone marrow	Testes	Pancreas	Brain (cortex)	Kidney (neonatal)	Stomach	Gut	Heart (neonatal)	Lungs	Muscle (juvenile)	Axillary Lymph Node	Thymus
CXCL12	NA	404	726	705	14	49	67	289	141	50	134	163	141	733	304
CXCL13	NA	190	0.5	4	NA	82	NA	4	29	15	2	29	3	2039	18

Supplemental Figure 10. (A-B) Serum isolated from peripheral blood or bone marrow of CLL patients (n=4) was tested for SDF1 **(A)** or CXCL13 **(B)** levels by ELISA (R&D Systems). **(C-D)** Conditioned media produced by primary human mesenchymal stromal cells (CM^{hMSC}) or by human stromal cell line HS5 (CM^{HS5}) was analyzed for levels of SDF1 **(C)** or CXCL13 **(D)**. The replicates in C and D represent analyses of conditioned media from various passages of the cells. ND stands for “not-detectable”. **(E-F)** Relative expression of CXCL12 **(E)** and CXCL13 **(F)** mRNAs in different organs of the NSG mice. For this experiment, 3 NSG mice (without transplantation of malignant cells) were sacrificed, organs homogenized, RNA isolated, and subjected to qRT-PCR. **(G)** Publicly available data on CXCL12 and CXCL13 mRNA expression in different compartments of C57BL/6J mice (mice strain not used in this study). NA stands for “data not available”. Data are presented as TPM (Transcripts per Kilobase Million) and were obtained from *Expression Atlas database* (ArrayExpress: experiment E-MTAB-3579).¹⁹

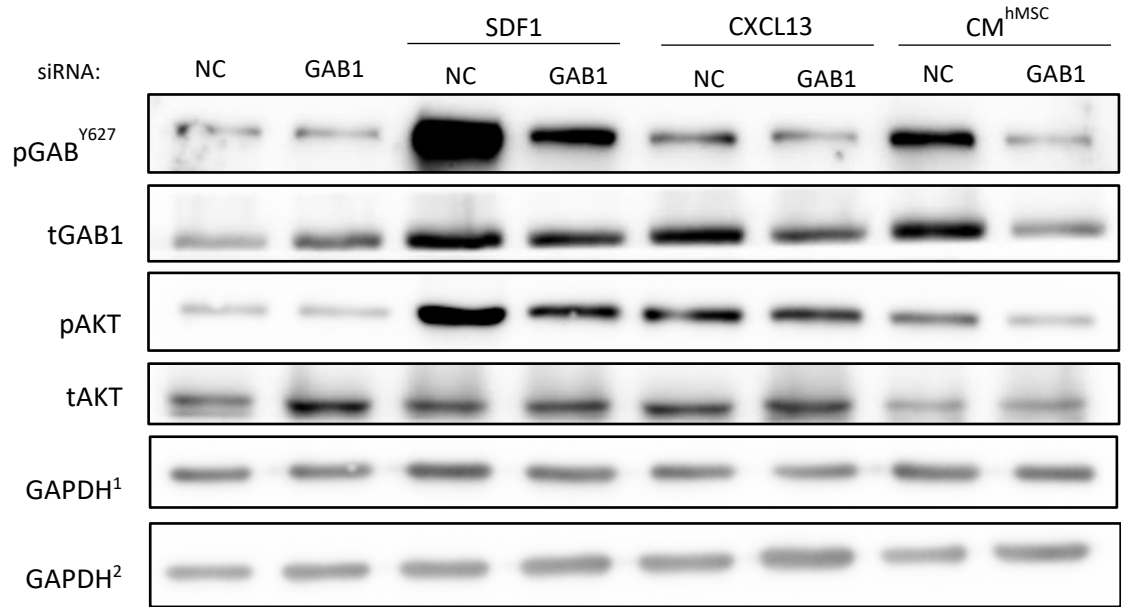
Supplemental Figure 11



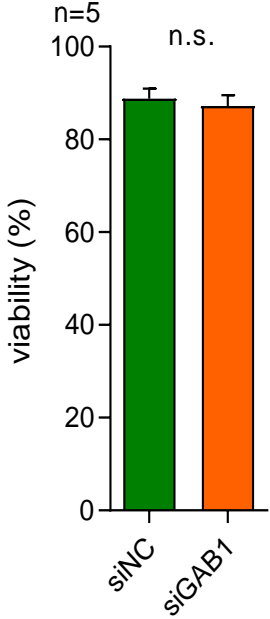
Supplemental Fig. 11 (A) Representative example of GAB1 silencing by siRNA in MEC1 (*top*) and primary CLL cells (*bottom*). **(B-C)** Migration of MEC1 cells (**B**) and CLL cells (**C**) towards 50% conditioned media produced by HS5 cells. CLL or MEC1 cells were transfected with siRNA against *GAB1* (siGAB1) or control siRNA (siNC). Migration across the transwell towards the 50% conditioned media produced by HS5 cells was allowed for 6 hrs (n=3), and the ratio of passed cells was measured by flow cytometry. **(D)** In vitro competitive migration assay of MEC1^{WT} vs. MEC1^{GAB1-KO} cells (cells stained as shown in Fig. 2B). Migration was allowed for 6 hrs towards conditioned media produced by HS5 stromal cells. **(E)** Representative immunoblot for primary CLL cells stimulated for various times (3-60 min) with recombinant SDF1 (100 ng/ml). **(F)** Representative immunoblot from primary CLL cells treated with 50% conditioned media produced by human stromal cell line HS5 (CM^{HS5}), SDF1 (100 ng/ml), or their combination with CXCR4-inhibitor plerixafor. Plerixafor (5 μg/ml) was added 2 hrs prior to the cell treatment with SDF1 or conditioned media to ensure full CXCR4 inhibition. The control stands for cells treated with vehicle and cultured on plastic.

Supplemental Figure 12

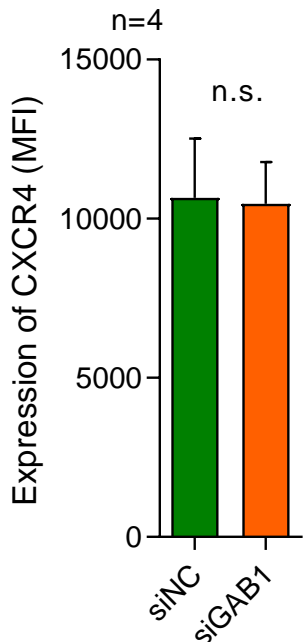
A



B

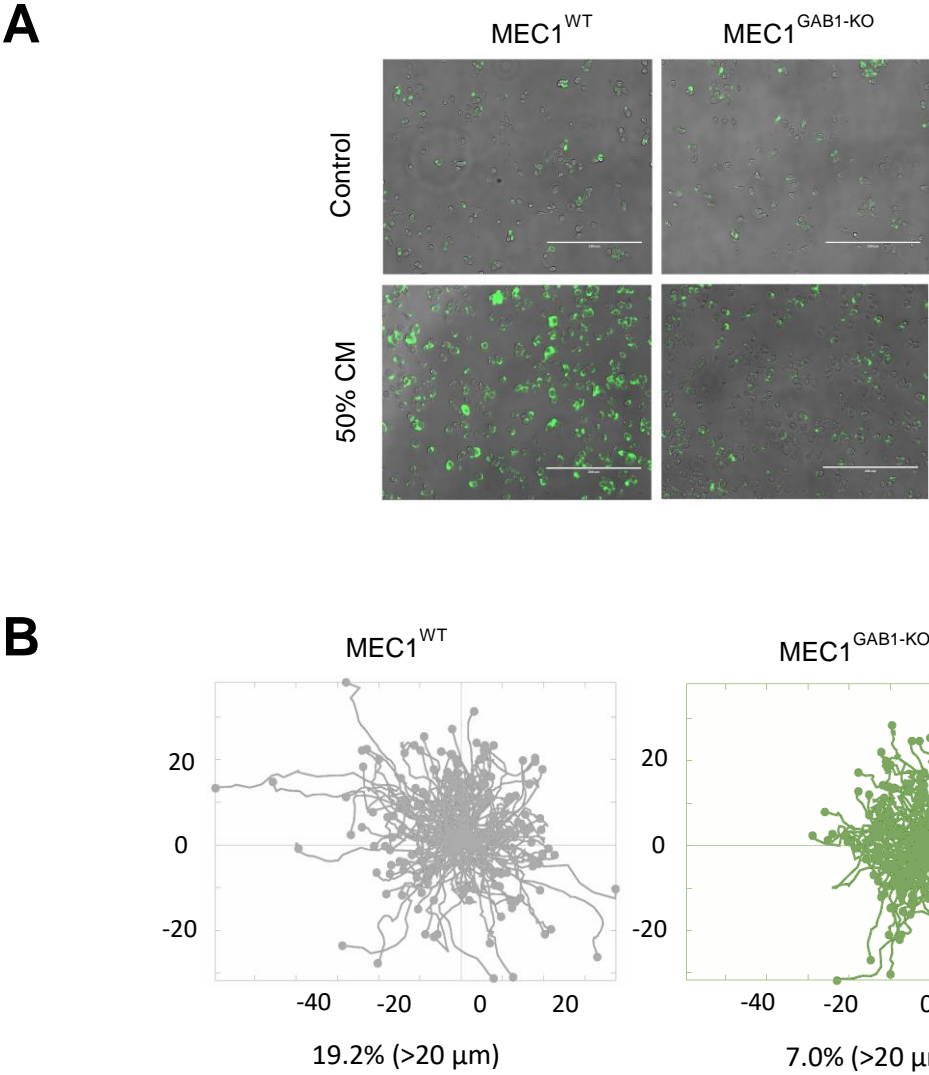


C



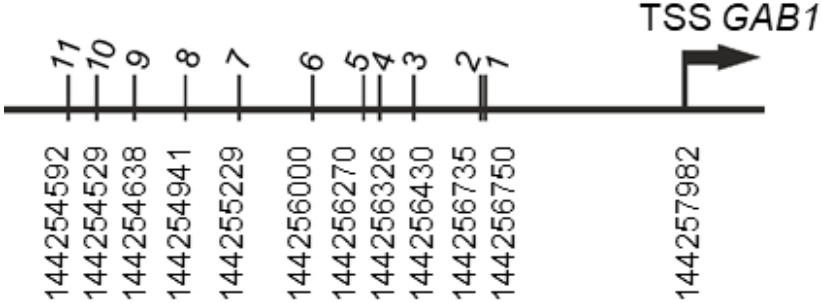
Supplemental Figure 12. (A) Additional immunoblot from primary CLL cells (CLL_133) transfected with siRNA against GAB1 (GAB1) or a negative control (NC). Cells were (48 hrs post-transfection) stimulated for 10 min by SDF1 (100 ng/ml), or CXCL13 (250 ng/ml), or conditioned media produced by primary human mesenchymal stromal cells (CM^{hMSC}). The immunoblot contains 2 endogenous controls (GAPDH) marked by upper index, because for technical reasons pGAB1 and pAKT (loading control GAPDH¹) were analyzed on first gel and remaining proteins (loading control GAPDH²) on the second gel (identical protein loading and conditions). **(B-C)** Analysis of cell viability (A) and CXCR4 cell-surface expression (B) in primary CLL cells transfected with siRNA against GAB1 (siGAB1) or with control scrambled siRNA (siNC). Viability was measured by DiOC6/PI staining.

Supplemental Figure 13



Supplemental Figure 13. (A) Representative image of F-actin polymerization in MEC1^{wt} and MEC1^{GAB1-KO} cells after treatment with 50% conditioned media produced by HS5 cells (15 min). MEC1 cells were subsequently fixed, and polymerized F-actin was visualized by fluorescent microscopy after Phalloidin-488 staining. **(B)** Representative rose plots with 5 min migration trajectories of MEC1^{wt} vs. MEC1^{GAB1-KO} cells in the presence of conditioned media produced by HS5 cells (shifted to a common origin, axis in μm). The percentage below the figure indicates the fraction of cells with trajectory >20 μm at 5 min in MEC1^{wt} vs. MEC1^{GAB1-KO} cells (19.2 vs. 7%, respectively; P<0.0001).

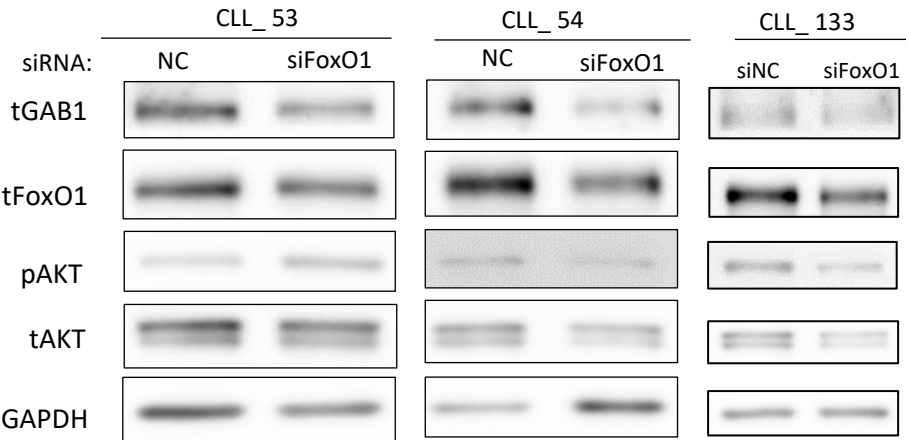
Supplemental Figure 14



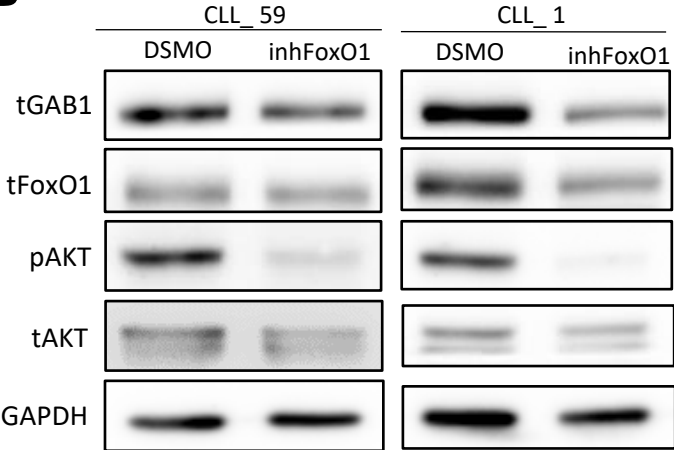
Supplemental Figure 14. Map of predicted transcription binding sites of FoxO1 protein in 3.5 kb promoter region of the *GAB1* gene (ContraV3). Predicted transcription binding sites are number 1-11 (at the top), and their genomic position is indicated.

Supplemental Figure 16

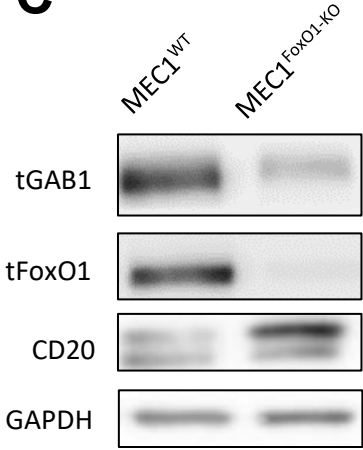
A



B

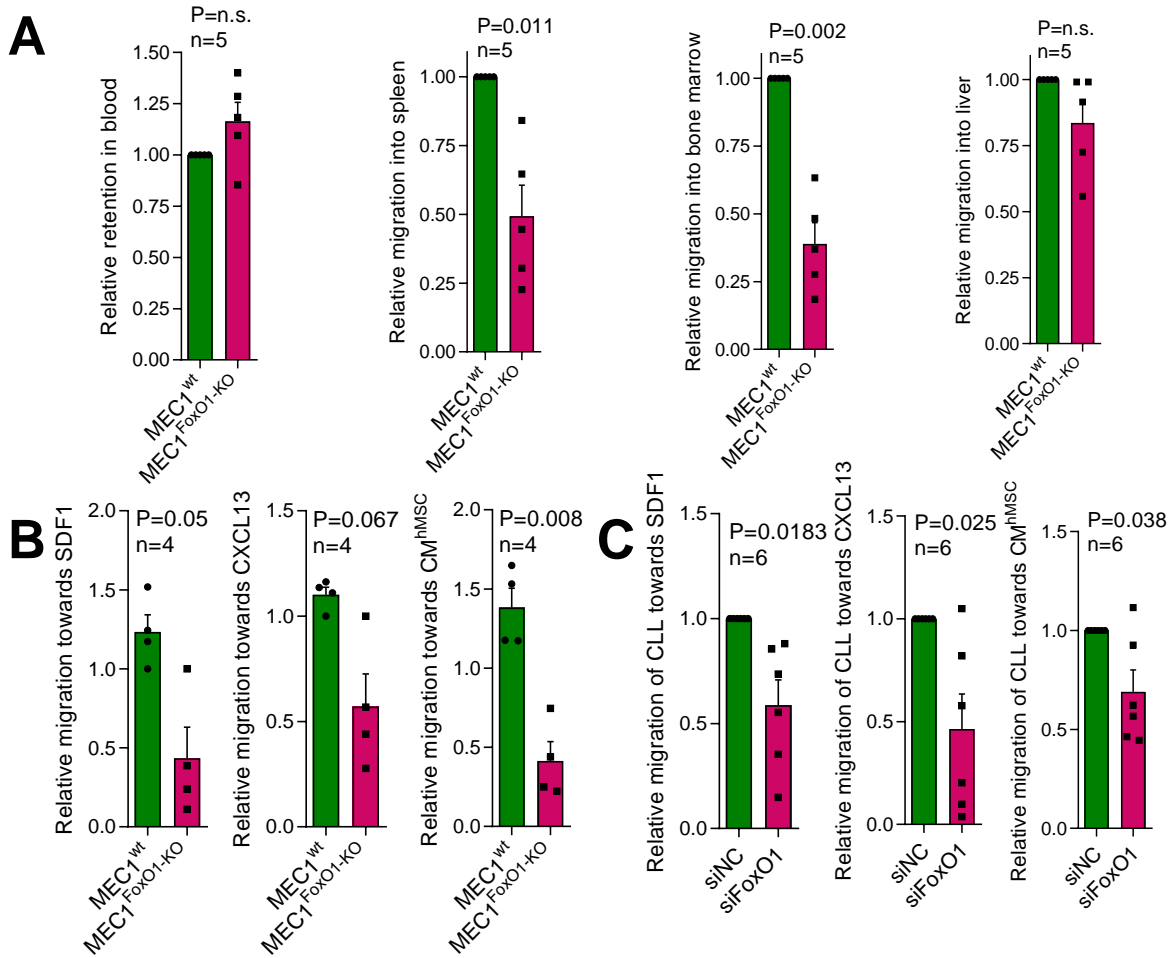


C



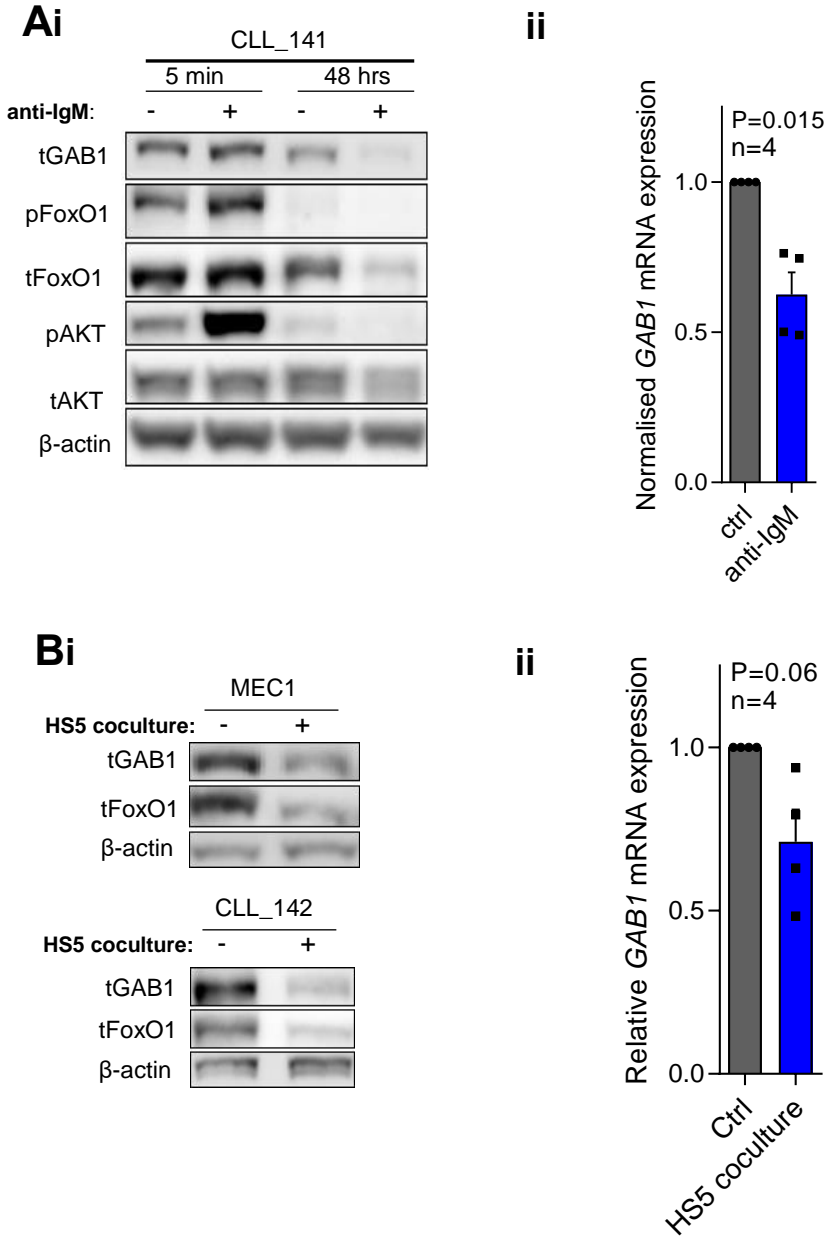
Supplemental Figure 16. (A) Immunoblots from primary CLL cells transfected with non-specific control siRNA (NC) or siRNA against *FoxO1* (siFoxO1). Cells were harvested 48 hrs post transfection. **(B)** Immunoblots from primary CLL cells treated with FoxO1 inhibitor (0,5 μM, 48 hrs). **(C)** Representative immunoblot for MEC1^{wt} and MEC1^{FoxO1-KO} cells. CD20 was used as a positive control of FoxO1 repression since FoxO1 is a negative regulator of CD20 transcription.²¹ The same samples were used in Fig. 4F.

Supplemental Figure 17



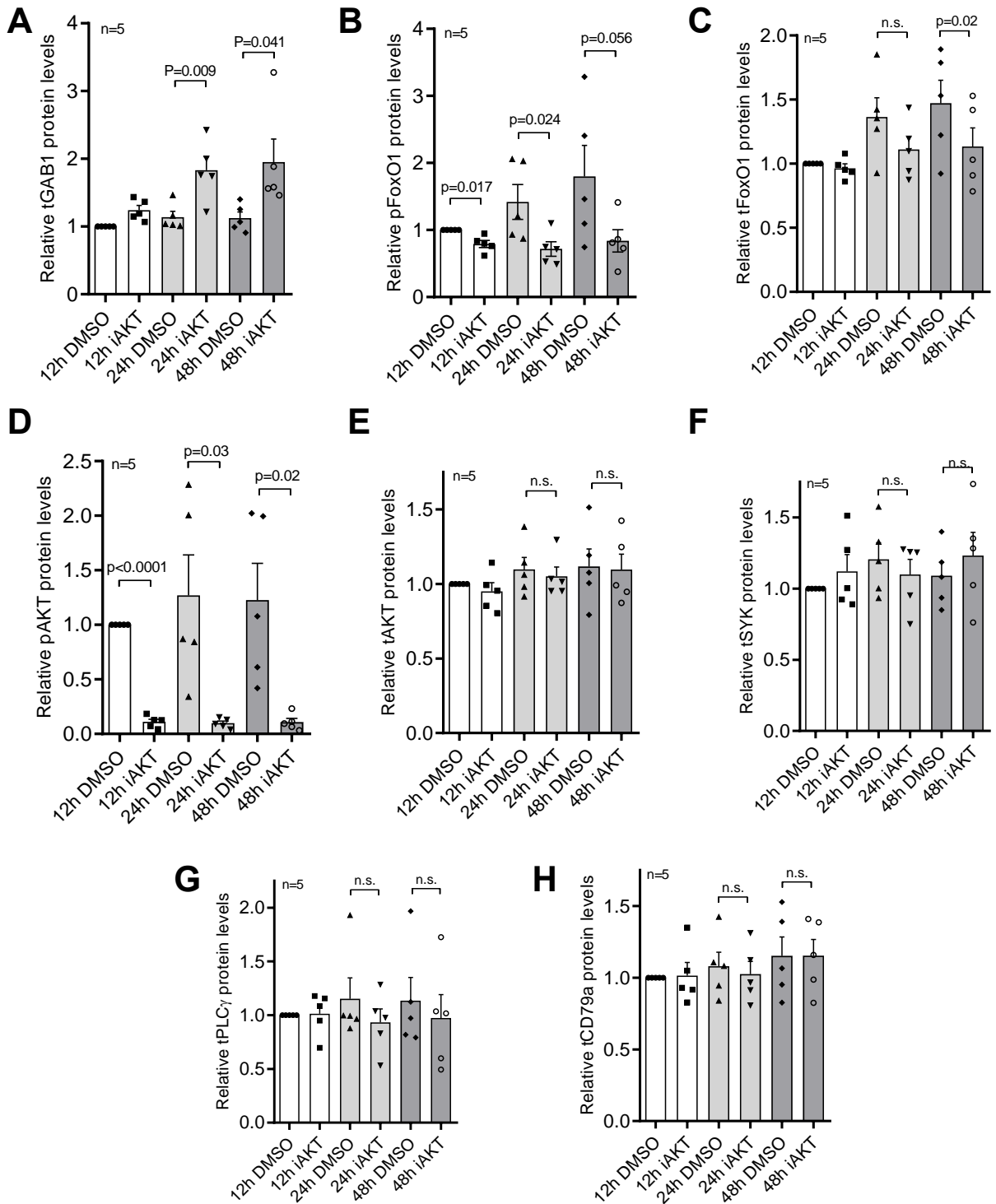
Supplemental Figure 17. (A) *In vivo* competitive migration assay of MEC1^{wt} vs. MEC1^{FoxO1-KO} cells. MEC1^{wt} and MEC1^{FoxO1-KO} were stained with CFSE or FarRed CellTrace dye, respectively, and incubated overnight. The next day equal numbers of CFSE stained and FarRed stained cells (each 50×10^6) were mixed to obtain a population with their 1:1 ratio (validated by flow cytometry). A mixture of stained cells was injected into NSG mice via the tail vein (total 100×10^6 cells, $n=5$). Mice were sacrificed after 6 hrs, and blood, spleen, bone marrow, and liver were analyzed by flow cytometry for the presence of viable CFSE or FarRed positive cells. The amount of MEC1 cells in each examined site is presented as relative migration between CFSE vs. FarRed positive cells. **(B)** *In vitro* competitive migration assay of MEC1^{wt} vs. MEC1^{FoxO1-KO}. MEC1^{wt} and MEC1^{FoxO1-KO} were stained with CFSE or FarRed CellTrace dye, respectively, and incubated overnight. The next day equal numbers of CFSE stained and FarRed stained cells were loaded to transwell. Migration ($n=4$) towards SDF1 (250 ng/ml), or CXCL13 (500 ng/ml), or conditioned media produced by human bone mesenchymal stromal cells (CM^{hMSC}) was allowed for 6 hrs, and the ratio of passed cells was assessed by flow cytometry. The number of migrated cells was normalized to the migration of cells applied to transwell without chemokines in the bottom chamber. **(C)** Competitive migration assay of CLL cells after *FoxO1* silencing by siRNA (siFoxO1) compared to cells transfected with control siRNA (siNC). Cells were incubated for 48 hrs and then stained with two different CellTrace dyes and loaded in 1:1 ratio on transwell. Migration ($n=6$) towards SDF1 (100 ng/ml), or CXCL13 (250 ng/ml), or conditioned media produced by human bone mesenchymal stromal cells (CM^{hMSC}) was allowed for 6 hrs, and the ratio of migrated cells was assessed by flow cytometry.

Supplemental Figure 18



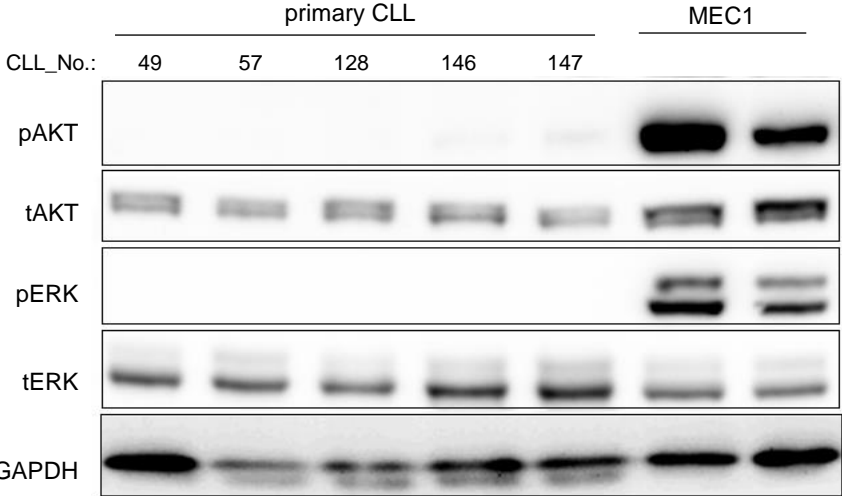
Supplemental Figure 18. (A) (i) Representative immunoblot for FoxO1 and GAB1 levels in primary CLL cells stimulated with anti-IgM (10 µg/ml) for 5 min or 48 hrs. **(ii)** GAB1 mRNA expression from CLL cells (n=4) treated with anti-IgM for 48 hrs. Ctrl stands for cells incubated in culture media only. **(B) (i)** Representative immunoblot of GAB1 and FoxO1 protein levels after coculture of MEC1 cells or primary CLL cells with HS5 stromal cells for 48 hrs. Primary CLL cells were seeded on a confluent layer of HS5 cells in an approx. 5:1 ratio (CLL:HS5). After 48 hrs cell suspension was stained with antibody against CD105 (a specific marker of stromal cells) and sorted by FACS (purity > 99,9% of CD19+CD5+ cells). **(ii)** GAB1 mRNA expression after coculture of primary CLL cells (n=4) with HS5 stromal cells for 48 hrs.

Supplemental Figure 19



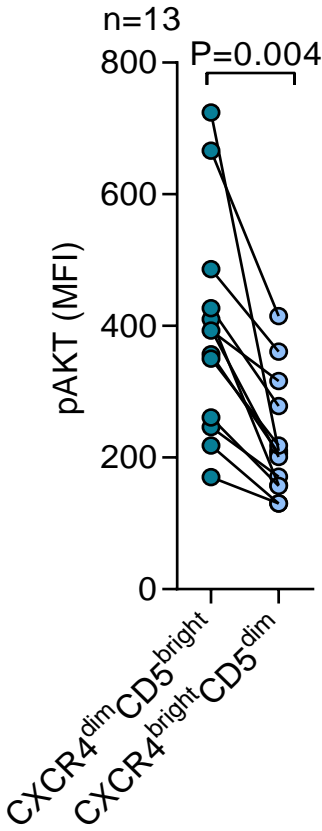
Supplemental Figure 19. Densitometric quantification of tGAB1 (**A**), pFoxO1 (**B**), tFoxO1 (**C**), pAKT (**D**), tAKT (**E**), tSYK (**F**), tPLC γ (**G**), tCD79a (**H**) and protein levels in MEC1 cells treated with 1 μ M inhibitor of AKT (iAKT) or vehicle (DMSO) for 12-48 hrs. MEC1 cells were used in this experiment, since primary CLL cells have relatively very low basal AKT phosphorylation levels to be inhibited (supplemental Figure 20).

Supplemental Figure 20



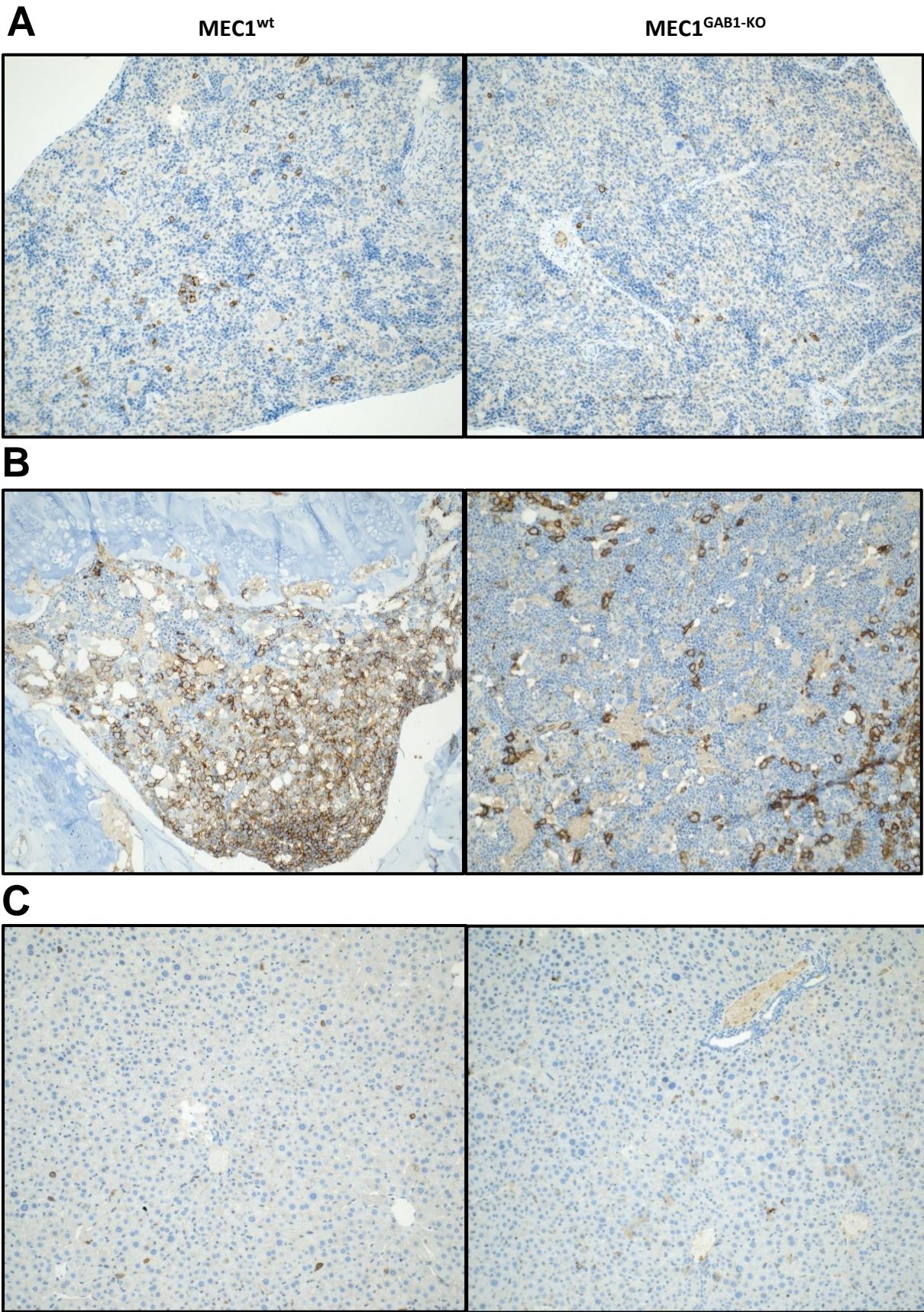
Supplemental Figure 20. Representative immunoblot for freshly isolated primary CLL cells and MEC1 cell line (samples from 2 different cultures of MEC1 cells).

Supplemental Figure 21



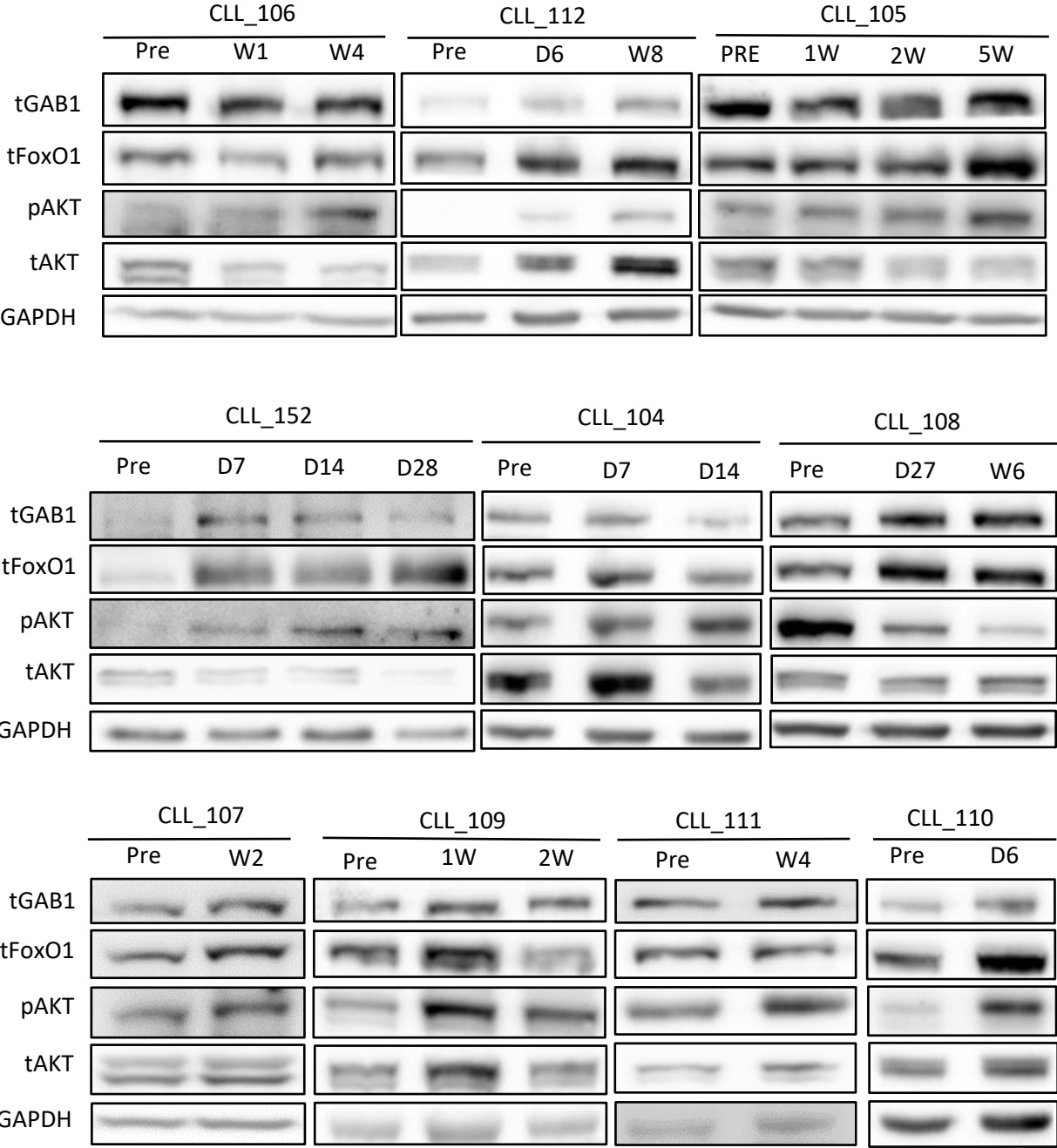
Supplemental figure 21. Intracellular staining of pAKT in CXCR4/CD5 CLL subpopulations.

Supplemental Figure 22



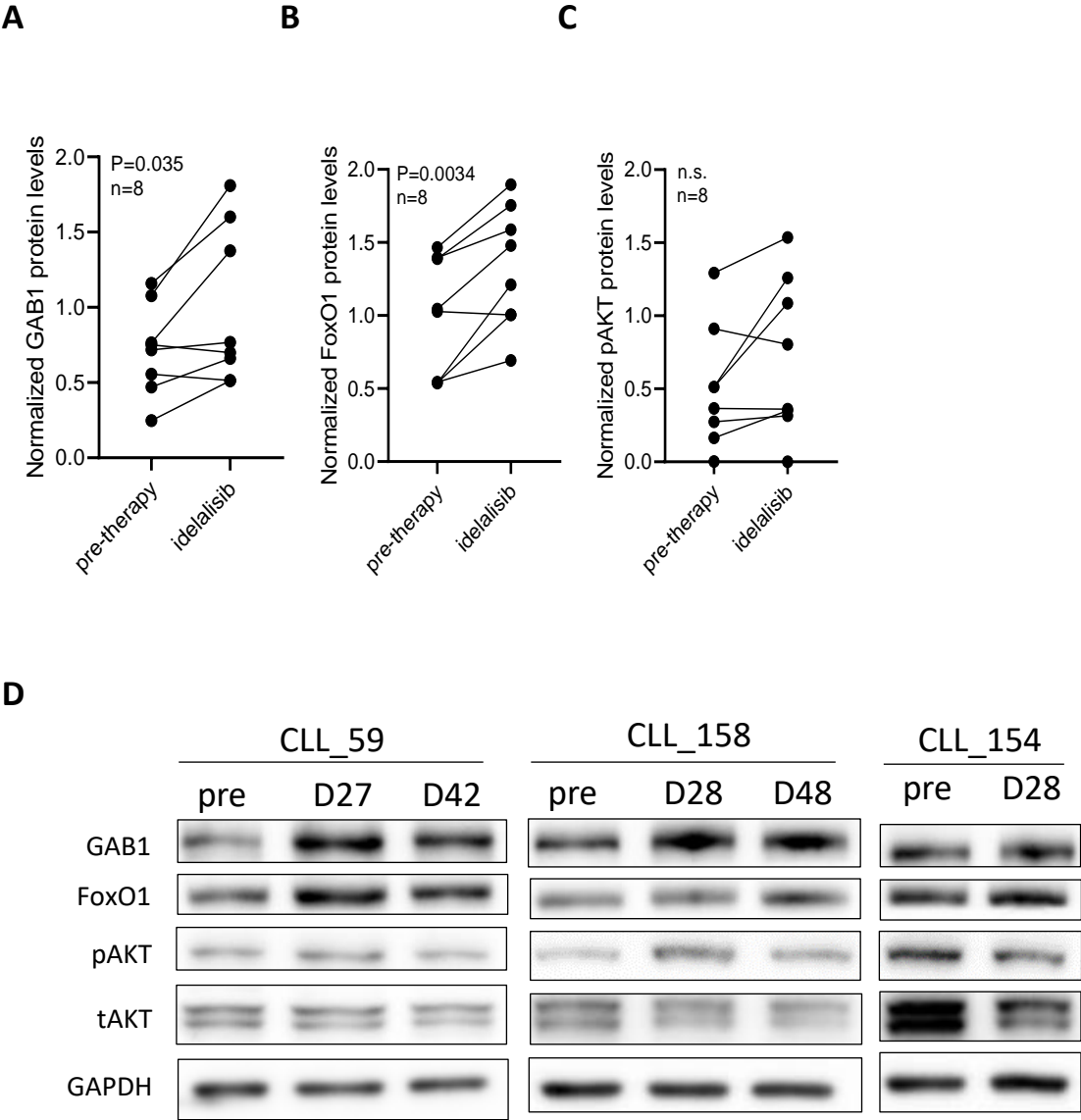
Supplemental Figure 22. Histological evaluation of spleen **(A)**, bone marrow **(B)**, and liver **(C)** of NSG mice transplanted with MEC1^{wt} (left) or MEC1^{GAB1-KO} (right) cells. In this experiment, MEC1^{wt} or MEC1^{GAB1-KO} cells were transplanted separately to individual NSG mice via a tail vein injection and sacrificed after 3 weeks. Tissue sections were stained with anti-human-CD20 antibody (brown signal) and the post-counterstaining by Bluing reagent (Ventana). The composition of the spleen structure in NSG mice is different from normal mouse spleen and reflects the absence of mature B/T/NK cells in these animals (see Supplemental Fig. 10D for hematoxylin-eosin staining of the spleen). Therefore, the white pulp of the spleen is absent, and the organ is mainly formed by red pulp with a high content of overactive erythron and excessive erythropoiesis (including the presence of megakaryocyte). The migrated MEC1 cells were localized diffusely in the spleen, and there was no difference in tissue localization for GAB1-knockout versus GAB1-wt cells (data not shown).

Supplemental Figure 23



Supplemental Fig. 23 Additional immunoblots from primary CLL cells isolated from peripheral blood of patients before and during ibrutinib therapy. “Pre” stands for samples isolated immediately before the first administration of ibrutinib, and samples obtained during therapy are indicated by the number of days (D) or weeks (W) on therapy.

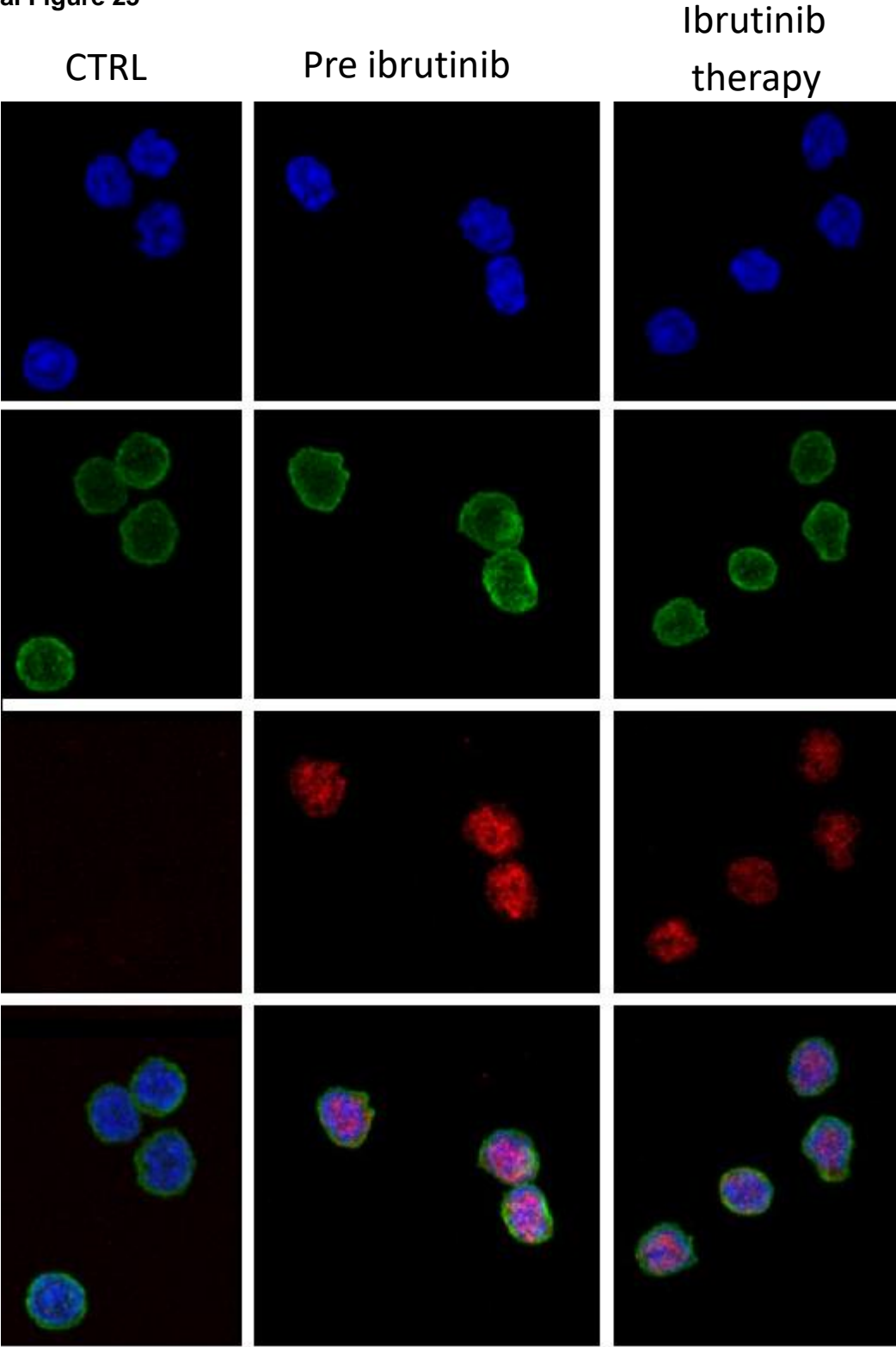
Supplemental Figure 24



Supplemental Fig. 24 (A-E) Densitometric quantification of tGAB1 (A), tFoxO1 (B), pAKT (C) protein levels in primary CLL cells before (pre-therapy) and during idelalisib therapy as a single agent (week 4). The density of each protein band was normalized to endogenous control. (D) Representative immunoblots from primary CLL cells isolated from peripheral blood of patients before (pre) and during idelalisib therapy (D stands for the number of days on therapy).

Supplemental Figure 25

A



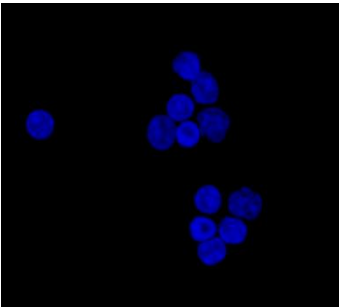
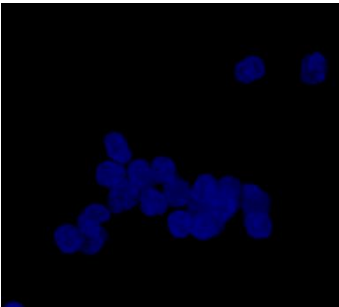
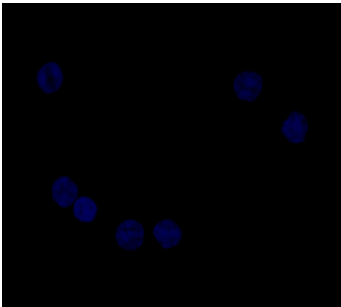
B

CTRL

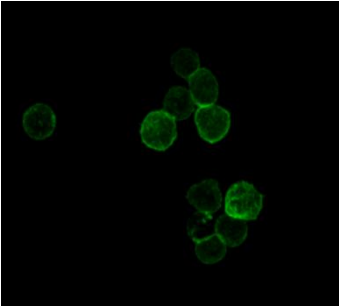
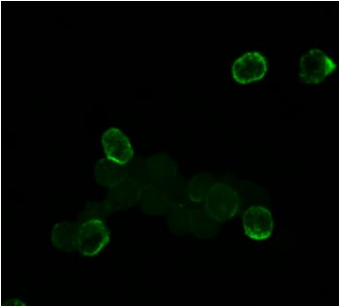
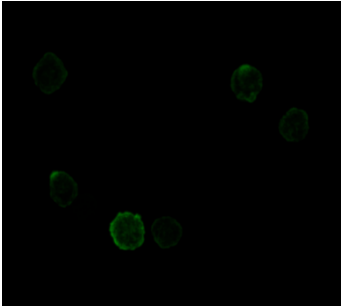
Pre ibrutinib

Ibrutinib
therapy

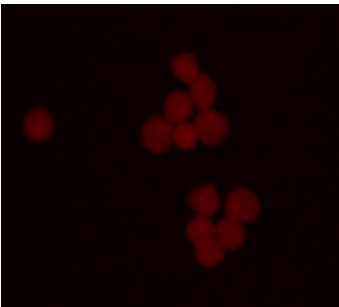
DAPI



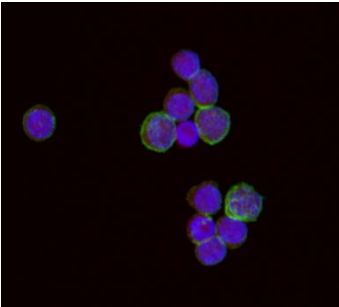
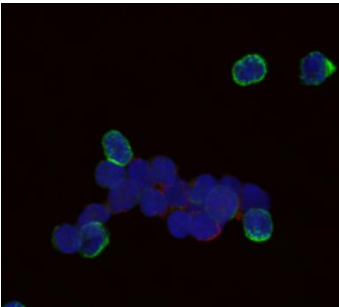
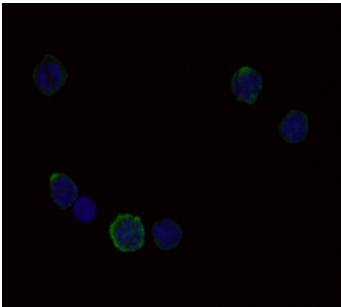
Actin

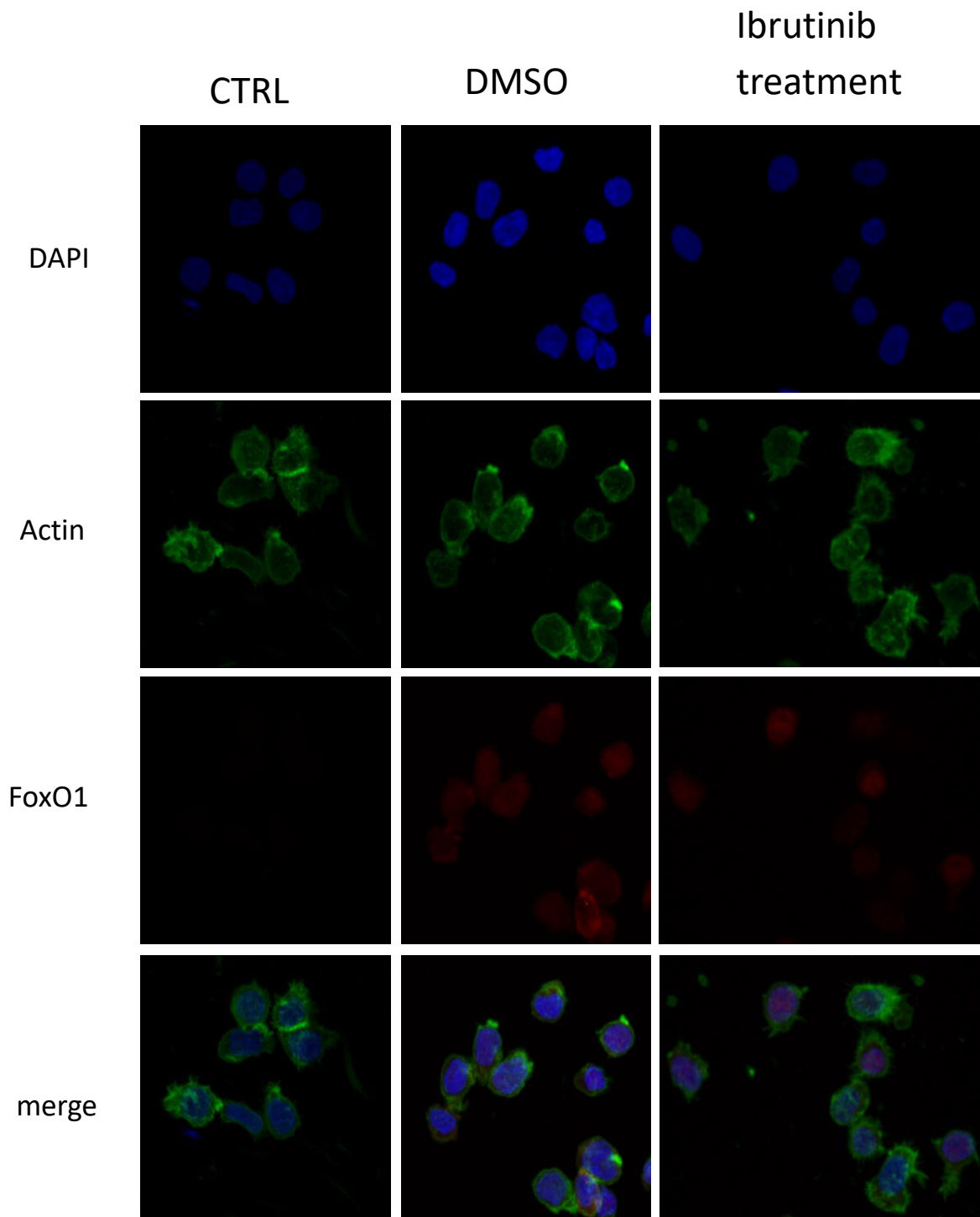


FoxO1



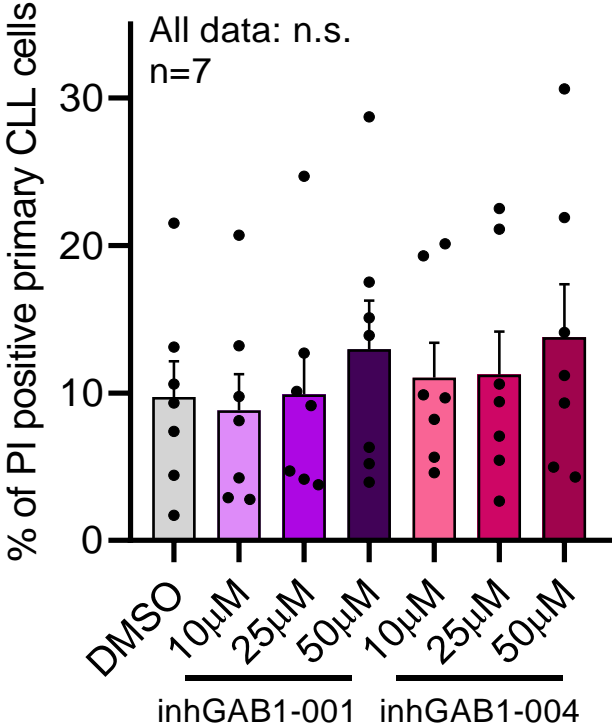
merge



C

Supplemental Figure 25. (A-B) Representative confocal microscopy picture of 2 primary CLL samples isolated from the patient before (pre) and during ibrutinib therapy (**[A]** week 2, **[B]** week 8; for patient characteristic, see sample CLL_160 and CLL_112, respectively, Supplemental Tab.1). **(C)** Representative confocal microscopy of MEC1 cells treated with 1 μ M ibrutinib or with an equal volume of control (DMSO) for 5 days. In control sample (CTRL) no anti-FoxO1 primary antibody was used. Blue staining (DAPI) represents the nucleus, green represents the cytoskeleton (actin), and red staining represents FoxO1.

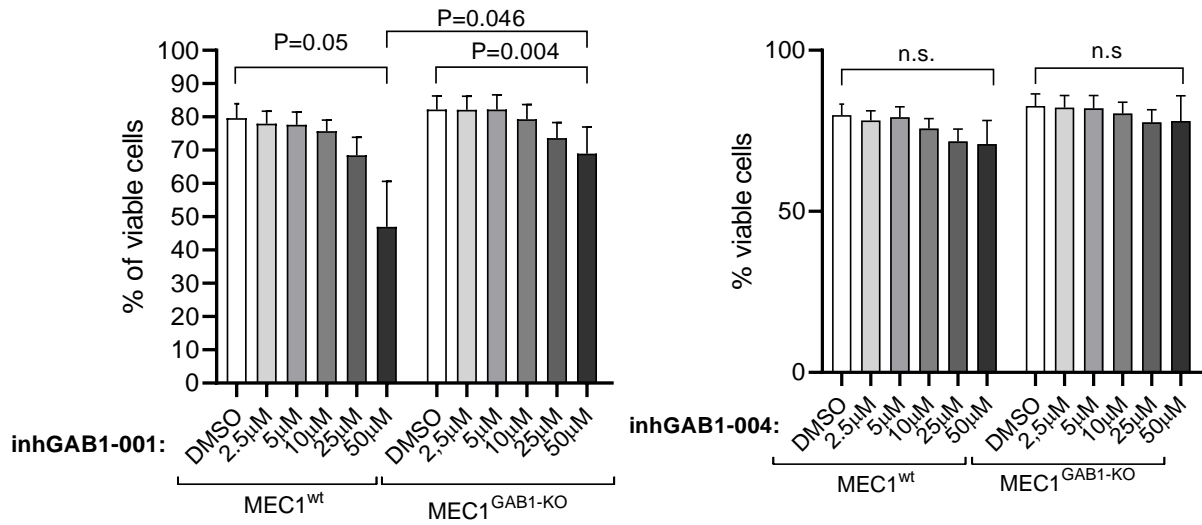
Supplemental Figure 26



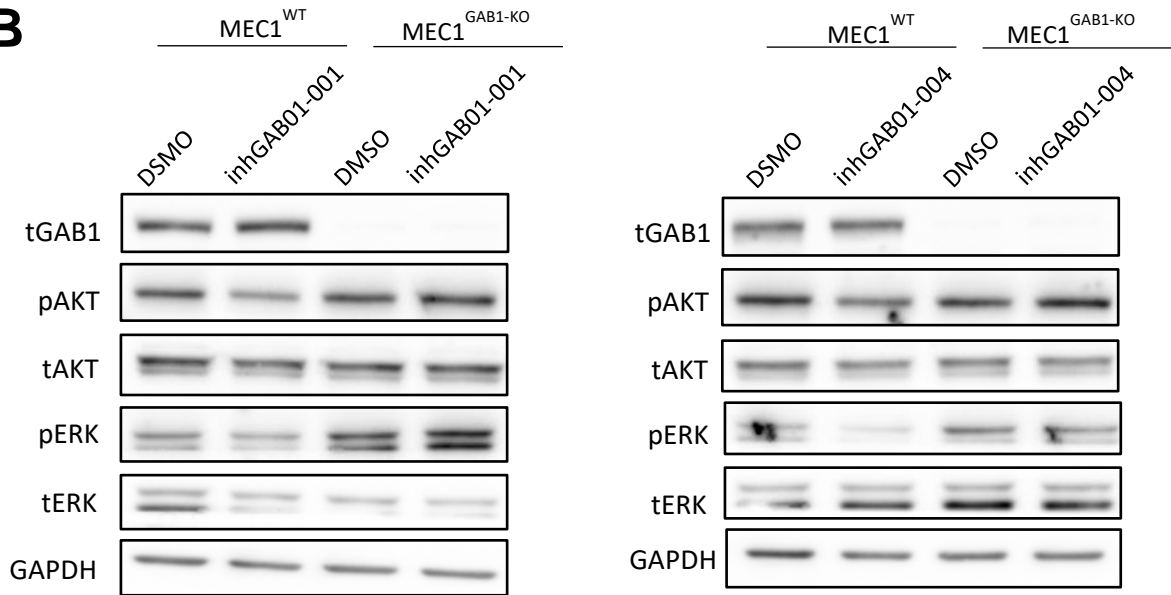
Supplemental Figure 26. Analysis of cell viability in primary CLL cells treated for 6 hrs with different doses of GAB1 inhibitor GAB1-001 or GAB1-004.

Supplemental Figure 27

A



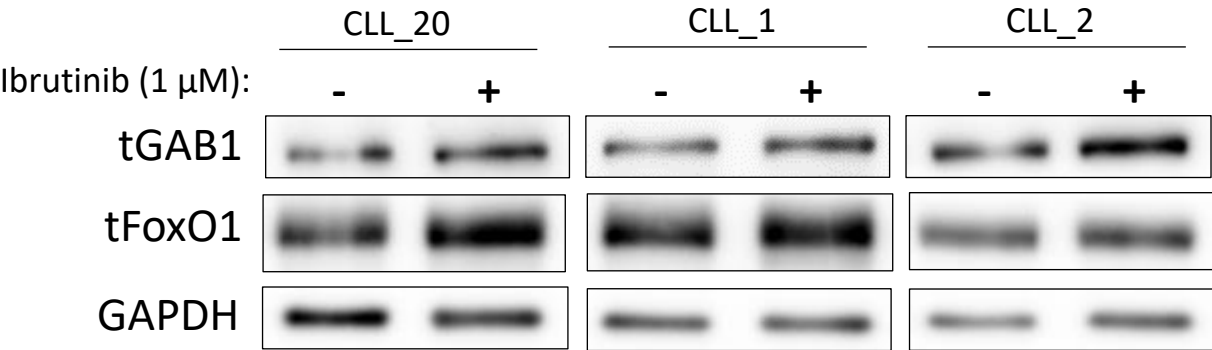
B



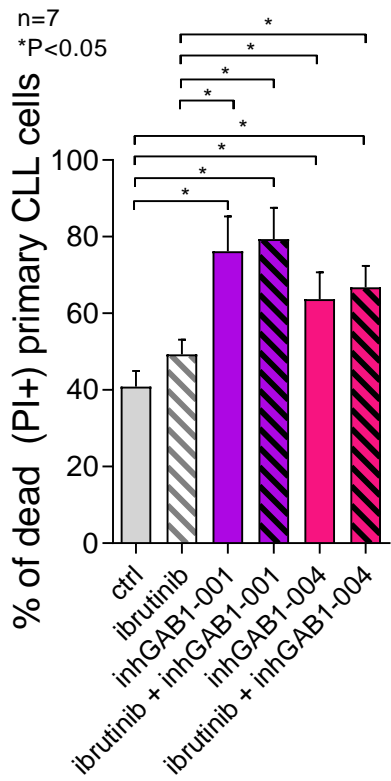
Supplemental Figure 27. (A) The effect of various doses (2.5-50 μM) of inhibitor GAB1-001 (*left*) or GAB1-004 (*right*) on the viability of MEC1^{wt} and MEC1^{GAB1-KO} cells (72 hrs, n=7 replicates). **(B)** Immunoblot for MEC1^{WT} and MEC1^{GAB1-KO} cells treated for 6 hrs with 50 μM inhibitor GAB1-001 (*left*) or GAB1-004 (*right*).

Supplemental Figure 28

A

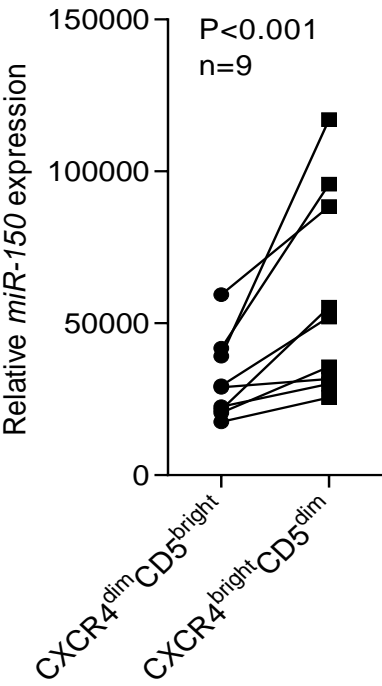


B



Supplemental Figure 28. (A) Representative immunoblot of primary CLL cells treated with 1 μM ibrutinib (+) or equal volume of DMSO (-) for 6 days. Cells were incubated in media containing 5 ng/ml IL4. **(B)** Statistical analysis of primary CLL cells' viability (n=7) after 48 hrs treatment with ibrutinib (1 μM) or two different GAB1 inhibitors (inhGAB1-001 or inhGAB1-004; 25 μM) or their combinations.

Supplemental Figure 29



Supplemental Figure 29. The relative expression of *miR-150-5p* in CXCR4/CD5 CLL subpopulations (n=9).

Supplemental Table 1: Characteristics of CLL patients

Patient ID	Sample used in figure	Gender	Age at time of diagnosis	Rai stage (at sample collection)	FISH results				TP53 mutation status	IGHV status
					del17p13 (%)	del11q23 (%)	trisomy 12 (%)	del13q14 (%)		
CLL 1	1A; 1D; 1G; S1; S2; S5; S6; S6; S16; S17; S28A	F	63	II B	neg	neg	neg	neg	wt	unmut
CLL 2	1A; 1D; 6I; 6J; S1; S2; S5; S26; S28A	M	66	IA	neg	pos (95)	neg	pos (96)	wt	unmut
CLL 3	1A; 2D; S1; S2; S8	M	56	II B	neg	neg	neg	neg	wt	unmut
CLL 4	1A; 2D; S1; S2; S9	F	54	0A	neg	neg	neg	pos (62)	wt	unmut
CLL 5	S9	M	56	II B	neg	neg	neg	neg	wt	unmut
CLL 6	1A; S1; S2	M	64	I A	neg	pos (94)	neg	pos (96)	wt	unmut
CLL 7	1A; S1; S2	F	68	IV	neg	pos (87)	neg	pos (10)	wt	---
CLL 8	1C; 1F; 1G; 3A; 3I; S3; S4; S6; S28	F	59	I A	neg	pos (64)	neg	pos (97)	wt	unmut
CLL 9	1C; 3A; 3I; S3; S4	M	74	IV A	neg	neg	neg	pos (94)	mut100	unmut
CLL 10	1C; 3A; 3I; S3; S4	F	64	III B	neg	pos (9)	neg	pos (47)	wt	---
CLL 11	1C; 3A; 3I; S3; S4; S28	M	71	II A	neg	neg	neg	neg	wt	---
CLL 12	1C; 3A; 3I; S3; S4; S28	F	59	III A-B	neg	neg	neg	pos (72)	wt	unmut
CLL 13	1C; 3A; 3I; S3; S4; S28	F	71	I A	neg	neg	neg	pos (59)	wt	unmut
CLL 14	1C; 3A; 3I; S3; S4; S28	M	79	IIIA	neg	pos (7)	neg	neg	wt	unmut
CLL 15	1C; 3A; 3I; S3; S4; S28	F	62	I A	neg	neg	neg	pos (7)	wt	unmut
CLL 16	1C; 3A; 3I; S3; S4; S28	M	70	I A	neg	neg	neg	pos (99)	wt	unmut
CLL 17	1C; 3A; 3I; S3; S4	F	65	II A	neg	neg	neg	pos (95)	wt	unmut
CLL 18	1D; 1G; 1E; 1F; 3F; 3H; 6C; S2; S5; S6; S28	M	58	I A	neg	neg	neg	43% cells with monoalelic deletion	wt	unmut
CLL 19	1D; 1G; 2H; 6D; 6E; 6I; 6J; S5; S6; S6; S19; S28A	M	62	II A	neg	neg	neg	pos (95)	wt	mut
CLL 20	1D; 1G; 3G; 6I; 6J; S5; S6; S11; S12; S28A	M	64	I A	neg	neg	neg	pos (9)	wt	---

CLL 21	1D; 1G; S2; S5; S6; S6	M	73	III A	neg	pos (74)	neg	neg	wt	unmut
CLL 22	1D; 1G; S5; S6; S6	M	68	IV A	neg	pos (98)	neg	neg	wt	
CLL 23	1D; 6C; S1; S2; S5; S18; S28	F	70	IA	neg	pos (95)	neg	pos (36)	wt	mut
CLL 24	1E; 2H; 3F; 3H; S19; S21	F	77	IVA	neg	neg	neg	81% cells with monoalelic deletion	wt	unmut
CLL 25	1E; 3F; 3H; S21	M	64	ND	tetrasomy (14%)	---	---	---	---	---
CLL 26	1E; 3F; 3H	M	90	IVA/B	neg	pos (34)	neg	pos (42)	wt	---
CLL 27	1E; 3F; 3H	M	58	IIIA	neg	neg	neg	96% cells with monoalelic deletion	mut (ND)	unmut
CLL 28	1E; 3F; 3H	M	68	IIA	neg	pos (97)	neg	neg	wt	unmut
CLL 29	1E; 3F; 3H	M	52		neg	neg	pos (65)	neg	wt	unmut
CLL 30	1E; 3F; 3H; S21	F	71	---	---	---	---	---	---	---
CLL 31	1E; 3F; 3H	M	46	IA	neg	neg	neg	pos (74)	wt	mut
CLL 32	1E; 3F; 3H	F	56	IIIB	neg	neg	neg	27% cells with monoalelic deletion	wt	unmut
CLL 33	1E; 3F; 3H; S21	M	69	OA	neg	neg	neg	neg	wt	unmut
CLL 34	1E; 3F; 3H; S21	F	75	III B	pos (del, 94)	neg	neg	91 % cells with bialelic deletion	mut (77)	mut
CLL 35	1E; 3F; 3H	M	64	OA	neg	neg	neg	pos (10)	wt	mut
CLL 36	1E; 3F; 3H; S21	F	70	A	neg	neg	pos (50)	neg	ND	ND
CLL 37	1E; 3F; 3H; S21	F	75	IIIB	pos (13)	4 copies of gene (63)	tetrasomy (62)	deletion (11); 4 copies (30); 3 copies (32)	wt	---
CLL 38	1E; 3F; 3H; S21	M	55	ND	neg	pos (98)	neg	neg	wt	unmut
CLL 39	1E; 3F; 3H	M	67	OA	neg	neg	neg	pos (41)	wt	mut
CLL 40	1E; 3F; 3H; 6C, S28	M	78	IVA	neg	pos (99)	neg	99% cells with monoalelic deletion	wt	unmut
CLL 41	1E; 3F; 3H; 6C, S28	F	54	IA	neg	neg	neg	25 % cells with monoalelic deletion	wt	mut
CLL 42	1E; 3F; 3H; 6D; 6E	F	66	IA	neg	neg	pos (89)	neg	wt	unmut

CLL 43	1E; 3F; 3H; 6I; 6J; S2	M	30	IIB	neg	neg	neg	neg	wt	unmut
CLL 44	1G; 3D; 6C; 6D; 6E; S6; S28	F	52	IA	neg	neg	neg	97% cells with bialelic deletion	wt	mut
CLL 45	1G; 3G; S6	M	76	IVB	neg	pos (66)	neg	pos (75)	mut	ND
CLL 46	1G; S8; S6; S11; S12	M	73	IIA	neg	neg?	neg	---	wt	unmut
CLL 47	2D; 2G; 4F; 6D; 6E; 6I; 6J; S9; S18; S26	F	54	IIIA	neg	neg	pos (85)	neg	wt	---
CLL 48	2D; S9	F	68	IV B	neg	pos (92)	neg	pos (88)	wt	mut
CLL 49	2D; S9; S20	M	70	IIB	neg	neg	neg	pos (8)	wt	mut
CLL 50	2H	F	76	ND	neg	neg	neg	pos (92)	mut (85)	unmut
CLL 51	2H	M	66	III A	neg	pos (28)	neg	pos (96)	wt	unmut
CLL 52	2H	F	73	II A	neg	neg	neg	neg	wt	unmut
CLL 53	2H; 6I; 6J; S16; S17; S19	M	73	---	neg	neg	neg	pos (95)	wt	---
CLL 54	2H; 6I; 6J; S16; S17; S19; S26	F	73	I A	neg	neg	pos (20)	pos (69)	wt	mut
CLL 55	2H; 6I; 6J; S17; S19; S26	F	75	0A	neg	neg	neg	neg	wt	mut
CLL 56	2H; S19	F	61	I A	neg	neg	neg	pos (46)	wt	mut
CLL 57	2H; S20	M	69	IV	neg	neg	neg	pos (6)	wt	unmut
CLL 58	3D	F	62	1A	neg	neg	neg	pos (81)	wt	mut
CLL 59	3D; 6A; 6B; S16; S17; S24; S26	F	59	---	neg	pos (93)	neg	neg	wt	unmut
CLL 60	3E	F	60	0	neg	neg	neg	pos (85)	wt (10)	mut
CLL 61	3E	F	67	0A	pos (62)	neg	neg	pos (71)	mut (63)	unmut
CLL 62	3E	M	59	IIA	neg	neg	neg	neg	wt	unmut
CLL 63	3E	M	50	IIB	neg	pos (95)	neg	pos (80)	wt	unmut
CLL 64	3E	F	61	IIIA	pos (98)	neg	neg	neg	mut (49)	unmut
CLL 65	3E	M	56	IIIA	pos (84)	neg	neg	pos (86)	mut (75)	unmut
CLL 66	3E	M	67	III	pos (93)	neg	neg	neg	wt	unmut
CLL 67	3E	F	58	IB	pos (47)	neg	neg	pos (95)	mut (95)	mut
CLL 68	3E	M	67	IVA	neg	neg	neg	pos (99)	mut (48)	unmut
CLL 69	3E	M	76	IVB	pos (63)	neg	neg	pos (62)	mut (95)	unmut

CLL 70	3E	F	57	II	neg	neg	neg	neg	mut (44)	unmut
CLL 71	3E	F	63	III	neg	neg	neg	neg	wt	unmut
CLL 72	3E	M	58	IIA	pos (97)	neg	neg	pos (18)	mut (91)	unmut
CLL 73	3E	M	62	IIB	neg	neg	neg	neg	wt	unmut
CLL 74	3E	F	67	I	neg	neg	neg	pos (98)	wt	mut
CLL 75	3E	M	66	0	neg	neg	neg	neg	wt	unmut
CLL 76	3E	M	63	IA-B	neg	neg	pos (81)	neg	wt	unmut
CLL 77	3E	F	56	IVA	neg	neg	neg	pos (15)	wt	unmut
CLL 78	3E	F	56	N/A	pos (66)	neg	neg	pos (94)	mut (47)	mut
CLL 79	3E	M	62	IV	neg	pos (100)	pos (ND)	neg	mut (89)	unmut
CLL 80	3E	F	68	III	neg	pos (96)	neg	pos (49)	wt	unmut
CLL 81	3E	F	57	II	neg	pos (93)	neg	pos (63)	wt	unmut
CLL 82	3E	M	60	II B	neg	pos (98)	neg	pos (99)	wt	unmut
CLL 83	3E	M	54	IVA	neg	neg	neg	neg	wt	unmut
CLL 84	3E	M	67	IIA	pos (93)	neg	pos (76)	pos (97)	mut (46)	mut
CLL 85	3E	M	72	IVA-B	neg	neg	neg	neg	del (82)	---
CLL 86	3E	F	56	III	neg	neg	pos (88)	pos (90)	mut (67)	mut
CLL 87	3E	F	54	IV	neg	neg	neg	pos (62)	wt	unmut
CLL 88	3E	M	49	0	neg	neg	neg	pos (96)	wt	mut
CLL 89	3E	M	65	IV	neg	pos (88)	neg	neg	wt	mut
CLL 90	3E	M	47	IVB	pos (72)	pos (38)	pos (35)	pos (35)	mut (90)	unmut
CLL 91	3E	M	77	III	pos (9)	neg	neg	pos (92)	mut (54)	unmut
CLL 92	3E	M	38	IIA	pos (88)	pos (11)	neg	pos (35)	mut (15)	unmut
CLL 93	3E	M	45	IV	neg	neg	neg	neg	mut (22)	mut
CLL 94	3E	M	55	III	neg	neg	pos (28)	neg	mut (46)	unmut
CLL 95	3E	M	75	III-IV	neg	pos (90)	neg	neg	wt	unmut
CLL 96	3E	M	38	I	neg	pos (99)	neg	pos (50)	wt	unmut
CLL 97	3E	F	59	IIB	neg	neg	neg	pos (21)	mut (99)	unmut
CLL 98	3E	M	43	I	neg	pos (15)	neg	neg	wt	unmut

CLL 99	3E	F	58	IIIA	neg	neg	neg	neg	wt	unmut
CLL 100	3E	M	64	IVA-B	neg	neg	neg	neg	wt	unmut
CLL 101	4A; S12	F	75	---	neg	neg	neg	mut (93)	wt	mut
CLL 102	5Di; 5Dii	M	47	IB	pos	neg	neg	neg	wt	unmut
CLL 103	1E; 3E; 3F; 3H; 5Dii	M	39	IVB	pos (12)	pos (85)	neg	pos (15)	mut (15)	unmut
CLL 104	5Dii; 5E; S23	M	59	IVB	neg	pos (80)	neg	pos (83)	wt	unmut
CLL 105	5Dii; 5E; S23	M	43	ND	neg	pos (96)	pos (80)	neg	wt	unmut
CLL 106	5Dii; 5E; S23	M	48	IVB	neg	neg	neg	pos (85)	wt	unmut
CLL 107	5Dii; 6B; S11; S23	M	65	IA	neg	neg	pos (86)	neg	wt	unmut
CLL 108	5Dii; S23	M	51	IVA	neg	neg	neg	neg	wt	unmut
CLL 109	5Dii; S23	F	64	---	neg	neg	neg	pos (8)	wt	mut
CLL 110	5Dii; S23	M	71	IVB	pos (85)	pos (15)	neg	neg	mut (95)	unmut
CLL 111	5Dii; S23	M	67	IA	neg	pos (86)	neg	pos (11)	wt	unmut
CLL 112	5Dii; S23; S25	M	58	IIB	neg	neg	neg	neg	wt	unmut
CLL 113	5E	M	34	I	pos (16)	neg	neg	neg	ND	mut
CLL 114	5E	F	51	IV	neg	neg	pos (52)	neg	ND	unmut
CLL 115	5E	M	57	IV	pos (60.5)	neg	neg	pos (96)	mut	mut
CLL 116	5E	M	75	I	neg	neg	neg	pos (78.5)	ND	mut
CLL 117	5E	M	56	I	pos (35.5)	neg	neg	pos (4.5)	mut	unmut
CLL 118	5E	F	67	IV	pos (93.5)	pos (93.5)	neg	neg	ND	unmut
CLL 119	5E	M	54	IA	neg	pos (60)	neg	pos (54)	wt	unmut
CLL 120	5E	F	53	IIIA	pos (8)	neg	neg	neg	wt	unmut
CLL 121	6B	M	66	IA	neg	neg	pos (77)	neg	wt	unmut
CLL 122	6B	M	64	IA	neg	neg	pos (81)	neg	wt	unmut
CLL 123	6B	M	64	IV-A	neg	neg	neg	neg	wt	ND
CLL 124	6C; S28	M	72	III A	neg	pos (0.7)	neg	73 % monoalelic deletion	wt	unmut

CLL 125	6C; S28	F	54	IIIB	neg	neg	pos (0.74)	neg	wt	biclonal /unmut
CLL 126	6D; 6E	M	51	0A	neg	neg	pos (82)	neg	wt	unmut
CLL 127	6D; 6E; 6I; 6J; S18	F	46	IVA-B	neg	neg	neg	pos (89)	wt	mut
CLL 128	6G; 6F; S11; S20	F	79	IA-B	neg	neg	neg	pos (8)	wt	mut
CLL 129	6I; 6J; S26	F	60	II B	neg	pos (16)	neg	neg	wt	unmut
CLL 130	6I; 6J	F	72	III A	neg	neg	neg	pos (52)	wt	unmut
CLL 131	6I; 6J	M	87	IV B	---	---	---	---	---	---
CLL 132	6I; 6J; S2	M	67	III B	neg	neg	neg	neg	---	mut
CLL 133	6I; 6J; S2; S12; S16; S26	M	77	I A	neg	neg	neg	pos (9)	wt	---
CLL 134	S10	M	76	0A	neg	neg	neg	pos (20)	wt	---
CLL 135	S10	M	61	IIIB	neg	pos (44)	neg	pos (9)	wt	unmut
CLL 136	S10	M	71	IV B	neg	pos (73)	neg	neg	mut	unmut
CLL 137	S10	F	66	IIIB	neg	pos (70)	neg	neg	wt	unmut
CLL 138	S11	F	80	IIIA	neg	neg	pos (63)	neg	wt	---
CLL 139	S11; S12	F	60	IIB	neg	pos (95)	neg	pos (8)	wt	unmut
CLL 140	S12	F	55	0A	neg	neg	neg	mut (60)	wt	mut
CLL 141	S18	F	73	IVA	neg	neg	pos (70)	neg	wt	---
CLL 142	S18	F	57	0A	neg	neg	neg	pos (45)	mut (40)	mut
CLL 143	S18	M	83	---	neg	neg	pos (78)	neg	wt	---
CLL 144	S18	F	74	IB	neg	neg	pos (80)	neg	wt	---
CLL 145	S18	M	70	IVA	neg	pos (52)	neg	neg	mut (35)	unmut
CLL 146	S20	M	69	IA	neg	neg	pos (81)	neg	wt	unmut
CLL 147	S20	M	68	IIB	neg	neg	pos (67)	pos (66)	wt	mut
CLL 148	S21	M	75	IIB	neg	neg	neg	neg	wt	unmut
CLL 149	S21	F	64	IA	neg	neg	neg	neg	wt	unmut
CLL 150	S21	M	71	IIA	neg	neg	neg	neg	wt	mut
CLL 151	S21	F	69	IIA	neg	neg	pos (72)	pos (61)	wt	mut
CLL 152	S23	M	64	III B	neg	neg	neg	neg	wt	unmut

CLL 153	S24	F	50	IV A	neg	pos (49)	neg	neg	mut	---
CLL 154	S24	F	74	---	neg	neg	pos (78)	neg	wt	---
CLL 155	S24	M	61	IV A	neg	pos (86)	neg	neg	wt	unmut
CLL 156	S24	M	---	---	neg	neg	neg	pos (94)	wt	unmut
CLL 157	S24	M	---	---	neg	pos (96)	neg	pos (94)	wt	mut
CLL 158	S24	F	63	IV A	neg	neg	neg	pos (55)	mut	---
CLL 159	S24	F	82	II A	neg	neg	neg	pos (13)	mut (22)	---
CLL 160	S25	M	34	IA	neg	pos (74)	neg	pos (15)	minor mut (2.5)	unmut
CLL 161	S2	M	80	---	neg	neg	neg	neg	wt	unmut
CLL 162	S29	M	69	IV B	neg	pos (91)	neg	pos (96)	wt	unmut
CLL 163	S29	M	68	III B	neg	neg	neg	pos (98)	wt	unmut

Supplemental Table 2: List of top differentially expressed genes in CXCR4/CD5 subpopulations (top 30 upregulated and top 30 downregulated mRNAs)

Gene ID	Gene name	Base Mean (reads per million)	Fold Change (log2) *	adjusted P-value
ENSG00000118432	CNR1	1722.72	2.39	2.80E-16
ENSG00000171246	NPTX1	1555.61	2.21	1.15E-15
ENSG00000135116	HRK	1111.617	2.05	1.28E-15
ENSG00000112182	BACH2	10947.54	2.05	2.99E-44
ENSG00000111266	DUSP16	219.72	2.01	8.96E-18
ENSG00000069188	SDK2	1503.15	1.91	1.39E-12
ENSG00000065534	MYLK	2140.31	1.91	8.31E-18
ENSG00000179088	C12orf42	256.13	1.87	2.05E-17
ENSG00000166949	SMAD3	289.63	1.81	4.59E-14
ENSG00000131941	RHPN2	276.96	1.81	3.23E-28
ENSG00000152953	STK32B	462.16	1.73	2.35E-06
ENSG00000186350	RXRA	2740.28	1.72	1.55E-11
ENSG00000099958	DERL3	1749.33	1.71	2.61E-14
ENSG00000102053	ZC3H12B	895.52	1.63	1.08E-37
ENSG00000169762	TAPT1	2568.98	1.62	6.45E-12
ENSG00000184897	H1FX	7058.84	1.60	3.21E-37
ENSG00000204569	PPP1R10	2626.45	1.59	2.89E-28
ENSG00000135299	ANKRD6	273.52	1.55	5.14E-33
ENSG00000172794	RAB37	1088.77	1.48	3.46E-17
ENSG00000165029	ABCA1	3209.94	1.47	6.96E-10
ENSG00000254870	ATP6V1G2- DDX39B	227.76	1.46	8.12E-14
ENSG00000141068	KSR1	690.79	1.46	5.01E-29
ENSG00000072134	EPN2	271.56	1.44	1.34E-18
ENSG00000101190	TCFL5	715.48	1.44	1.75E-14
ENSG00000215912	TTC34	440.05	1.44	3.82E-19
ENSG00000118513	MYB	290.40	1.42	5.37E-05
ENSG00000121966	CXCR4	82583.46	1.40	3.26E-06
ENSG00000150672	DLG2	570.30	1.40	3.76E-09
ENSG00000099953	MMP11	421.29	1.40	7.99E-18
ENSG00000166348	USP54	347.07	1.39	4.46E-26
ENSG00000129226	CD68	953.26	-3.76	2.05E-125
ENSG00000197629	MPEG1	2503.31	-3.12	1.13E-35
ENSG00000116701	NCF2	1227.26	-2.90	4.51E-53
ENSG00000124772	CPNE5	2444.89	-2.57	1.17E-21
ENSG00000186818	LILRB4	949.50	-2.56	2.30E-22
ENSG00000277632	CCL3	316.78	-2.54	7.13E-16

ENSG00000140678	ITGAX	7274.70	-2.48	4.60E-28
ENSG00000203896	LIME1	225.21	-2.43	6.65E-47
ENSG00000278195	SSTR3	255.24	-2.29	2.25E-23
ENSG00000173366	AC097637.1	240.20	-2.25	7.01E-15
ENSG00000177685	CRACR2B	1095.87	-2.23	5.75E-37
ENSG00000011028	MRC2	200.59	-2.18	1.07E-18
ENSG00000186810	CXCR3	307.94	-2.18	2.69E-16
ENSG00000126860	EVI2A	579.76	-2.11	2.71E-39
ENSG00000169413	RNASE6	813.61	-2.08	2.40E-19
ENSG00000188389	PDCD1	361.91	-2.02	1.42E-55
ENSG00000167711	SERPINF2	275.58	-2.01	2.32E-28
ENSG00000138623	SEMA7A	724.22	-1.98	2.46E-34
ENSG00000133067	LGR6	228.25	-1.98	7.69E-24
ENSG00000030582	GRN	6864.43	-1.97	6.96E-60
ENSG00000179841	AKAP5	255.83	-1.96	3.41E-17
ENSG00000116771	AGMAT	230.14	-1.96	1.40E-33
ENSG00000261150	EPPK1	281.68	-1.95	8.57E-14
ENSG00000146386	ABRA1	290.48	-1.95	1.78E-24
ENSG00000216490	IFI30	1022.84	-1.92	3.86E-14
ENSG00000115956	PLEK	2534.33	-1.91	1.46E-15
ENSG00000196154	S100A4	3060.76	-1.91	6.17E-42
ENSG00000189159	JPT1	358.79	-1.90	1.57E-22
ENSG00000157388	CACNA1D	574.72	-1.87	4.32E-10
ENSG00000172824	CES4A	426.87	-1.87	1.12E-16

*30 most up-regulated genes (positive fold change) and down-regulated genes (negative fold change) in CXCR4^{bright}CD5^{dim} subpopulation.

Supplemental Table 3: List of migration-related genes which were identified as differentially expressed in CXCR4/CD5 subpopulations

Gene ID	Gene name	Base Mean (reads per million)	Fold Change (log ₂) *	adjusted P-value
ENSG00000114353	GNAI2	8685	-0.95	4.08E-75
ENSG00000100823	APEX1	3708	-1.05	2.8E-59
ENSG00000030582	GRN	6864	-1.97	7.4E-57
ENSG00000128340	RAC2	9335	-1.75	2.46E-53
ENSG00000182054	IDH2	1225	-1.23	1.14E-50
ENSG00000108518	PFN1	8083	-0.96	1.63E-47
ENSG00000100292	HMOX1	1000	-1.76	2.45E-43
ENSG00000159842	ABR	11899	1.05	1.36E-42
ENSG00000205542	TMSB4X	27296	-1.33	3.42E-41
ENSG00000103490	PYCARD	302	-1.53	1.59E-39
ENSG00000197043	ANXA6	6852	-1.34	8.87E-39
ENSG00000172725	CORO1B	2034	-0.75	7.85E-38
ENSG00000149781	FERMT3	2224	-1.24	3.7E-36
ENSG00000176973	FAM89B	519	-1.23	7.76E-36
ENSG00000108561	C1QBP	984	-1.00	2.72E-35
ENSG00000002586	CD99	2088	-1.11	5.98E-35
ENSG00000131375	CAPN7	2115	0.92	5.99E-33
ENSG00000149925	ALDOA	7757	-0.64	1.32E-32
ENSG00000138623	SEMA7A	724	-1.98	2.66E-32
ENSG00000100092	SH3BP1	1891	-0.96	7.89E-32
ENSG00000121879	PIK3CA	3805	0.75	5.77E-31
ENSG00000134242	PTPN22	5538	-0.84	1.36E-30
ENSG00000139626	ITGB7	6295	-1.74	1.79E-30
ENSG00000105429	MEGF8	1215	0.80	1.85E-30
ENSG00000149177	PTPRJ	8064	-0.83	6.66E-30
ENSG00000196712	NF1	2136	.1.13	5.83E-28
ENSG00000019995	ZRANB1	2053	0.69	6.03E-28
ENSG00000064666	CNN2	5355	-1.32	2.19E-27
ENSG00000142798	HSPG2	725	-1.51	6.47E-27
ENSG00000140678	ITGAX	7275	-2.48	2.76E-26
ENSG00000023902	PLEKHO1	3666	-0.87	2.88E-26
ENSG00000177169	ULK1	2386	0.62	4.53E-26
ENSG00000136167	LCP1	12293	-1.04	1.06E-25
ENSG00000108854	SMURF2	1459	0.69	2.23E-25
ENSG00000144711	IQSEC1	2265	-0.91	4.25E-25
ENSG00000172757	CFL1	9706	-0.61	9.42E-25
ENSG00000143322	ABL2	2082	0.93	2.11E-24
ENSG00000164190	NIPBL	8141	0.61	4.29E-24
ENSG00000111348	ARHGDIB	9596	-0.75	1.85E-23
ENSG00000163466	ARPC2	10620	-0.81	2.38E-23
ENSG00000132386	SERPINF1	2035	-1.47	3.17E-23
ENSG00000157764	BRAF	3089	0.63	8.28E-23
ENSG00000134318	ROCK2	1826	0.60	1.91E-22
ENSG00000155366	RHOC	1754	-1.33	1.98E-22
ENSG00000133067	LGR6	228	-1.98	2.87E-22
ENSG00000197471	SPN	1125	-1.27	1.59E-21
ENSG00000103769	RAB11A	1740	-0.95	1.78E-21
ENSG00000184481	FOXO4	969	0.70	2.34E-21
ENSG00000139146	SINHCAF	4318	0.92	2.49E-21
ENSG00000139725	RHOF	893	-0.92	4.00E-21

ENSG00000162704	ARPC5	2307	-1.17	7.62E-21
ENSG00000123338	NCKAP1L	12107	-0.74	9.13E-21
ENSG00000160007	ARHGAP35	1749	0.96	1.04E-20
ENSG00000122694	GLIPR2	363	-1.59	1.93E-19
ENSG00000134954	ETS1	21200	1.00	1.97E-19
ENSG00000131504	DIAPH1	5515	-0.60	4.96E-19
ENSG00000100300	TSPO	986	-0.89	4.48E-18
ENSG00000136286	MYO1G	3398	-1.44	7.97E-18
ENSG00000065534	MYLK	2140	.1.91	8.31E-18
ENSG00000110934	BIN2	3042	-1.04	8.99E-18
ENSG00000185950	IRS2	2382	1.12	1.19E-17
ENSG00000102879	CORO1A	13492	-0.65	1.38E-17
ENSG00000083642	PDS5B	2239	0.70	2.36E-17
ENSG00000213654	GPSM3	2977	-0.63	2.51E-17
ENSG00000188906	LRRK2	5584	-0.64	1.35E-16
ENSG00000109919	MTCH2	672	-0.96	1.97E-16
ENSG00000109458	GAB1	6770	0.82	2.44E-16
ENSG00000177697	CD151	339	-1.78	2.55E-16
ENSG00000140443	IGF1R	344	1.42	3.55E-16
ENSG00000079385	CEACAM1	1015	-1.15	5.01E-16
ENSG00000173531	MST1	346	-1.15	7.41E-16
ENSG00000168961	LGALS9	4699	-0.75	1.28E-15
ENSG00000196924	FLNA	29974	-1.15	1.33E-15
ENSG00000109756	RAPGEF2	6076	0.65	1.46E-15
ENSG00000151651	ADAM8	1768	-1.25	1.46E-15
ENSG00000108819	PPP1R9B	5964	-0.89	1.64E-15
ENSG00000160255	ITGB2	480	-1.34	3.23E-15
ENSG00000166974	MAPRE2	1919	-0.66	3.96E-15
ENSG00000186810	CXCR3	308	-2.18	4.32E-15
ENSG00000026025	VIM	17798	-0.88	5.35E-15
ENSG00000277632	CCL3	317	-2.54	1.09E-14
ENSG00000115310	RTN4	1720	-0.63	1.5E-14
ENSG00000087589	CASS4	391	-1.16	1.87E-14
ENSG00000143669	LYST	9027	1.15	2.23E-14
ENSG00000213066	FGFR1OP	873	0.63	4.07E-14
ENSG00000166949	SMAD3	290	1.81	4.59E-14
ENSG00000163832	ELP6	294	-0.65	5.63E-14
ENSG00000146535	GNA12	4735	0.98	8.25E-14
ENSG00000150093	ITGB1	8907	-0.69	8.38E-14
ENSG00000136950	ARPC5L	1205	-0.78	8.98E-14
ENSG00000035664	DAPK2	292	-1.31	9.39E-14
ENSG00000118689	FOXO3	1987	1.06	1.26E-13
ENSG00000039523	RIPOR1	9249	0.84	1.46E-13
ENSG00000137509	PRCP	1533	-0.61	1.56E-13
ENSG00000131236	CAP1	5959	-0.74	5.82E-13
ENSG00000174483	BBS1	532	0.84	9.33E-13
ENSG00000261150	EPPK1	282	-1.95	1.00E-12
ENSG00000108840	HDAC5	2714	0.65	1.29E-12
ENSG00000169508	GPR183	1004	-1.79	2.1E-12
ENSG00000177105	RHOG	886	-0.81	2.31E-12
ENSG00000142347	MYO1F	1331	-0.61	7.07E-12
ENSG00000071051	NCK2	2153	1.09	1.1E-11
ENSG00000079337	RAPGEF3	7587	0.82	2.54E-11
ENSG00000150347	ARID5B	6961	0.60	3.47E-11
ENSG00000100219	XBP1	2708	0.91	4.1E-11
ENSG00000104998	IL27RA	774	-0.99	5.8E-11

ENSG00000174175	SELP	566	1.10	5.86E-11
ENSG00000134352	IL6ST	2266	0.80	1.77E-10
ENSG00000123358	NR4A1	578	1.23	3.3E-10
ENSG00000115109	EPB41L5	254	1.04	6.93E-10
ENSG00000149418	ST14	3406	-0.72	1.02E-09
ENSG00000196220	SRGAP3	520	-0.83	1.66E-09
ENSG00000140564	FURIN	1265	0.66	1.93E-09
ENSG00000127603	MACF1	7205	-0.65	5.98E-09
ENSG00000135622	SEMA4F	544	0.62	6.28E-09
ENSG00000275342	PRAG1	1966	-0.65	9.72E-09
ENSG00000107104	KANK1	730	0.61	9.96E-09
ENSG00000112139	MDGA1	2082	1.06	1.71E-08
ENSG00000152127	MGAT5	6438	0.63	3.27E-08
ENSG00000004399	PLXND1	2216	0.95	3.51E-08
ENSG00000144791	LIMD1	6186	-0.78	4.12E-08
ENSG00000166313	APBB1	1070	0.63	4.25E-08
ENSG00000102760	RGCC	1179	-0.93	7.29E-08
ENSG00000089159	PXN	1592	-0.64	9.01E-08
ENSG00000157193	LRP8	412	-0.67	1.01E-07
ENSG00000163697	APBB2	2695	0.89	1.09E-07
ENSG00000135503	ACVR1B	392	0.89	1.66E-07
ENSG00000163932	PRKCD	1513	-0.87	1.97E-07
ENSG00000079691	CARMIL1	1689	0.60	2.09E-07
ENSG00000068024	HDAC4	237	0.88	2.54E-07
ENSG00000130669	PAK4	208	0.82	1.01E-06
ENSG00000261371	PECAM1	4511	-0.78	1.9E-06
ENSG00000169896	ITGAM	1899	-1.30	1.91E-06
ENSG00000121966	CXCR4	82583	1.40	3.26E-06
ENSG00000107738	VSIR	416	-0.77	6.09E-06
ENSG00000117020	AKT3	1058	0.83	7.23E-06
ENSG00000160683	CXCR5	5662	-0.61	8.47E-06
ENSG00000143537	ADAM15	362	-0.67	1.45E-05
ENSG00000197381	ADARB1	2338	1.10	2.34E-05
ENSG00000158321	AUTS2	492	0.74	3.89E-05
ENSG00000160789	LMNA	334	-1.18	5.73E-05
ENSG00000198648	STK39	201	-0.68	6.67E-05
ENSG00000069702	TGFBR3	1979	-0.61	7.03E-05
ENSG00000182957	SPATA13	874	-0.62	0.000104
ENSG00000171522	PTGER4	847	-1.06	0.000138
ENSG00000067141	NEO1	305	0.62	0.000356
ENSG00000085733	CTTN	295	0.76	0.000739

* positive numbers mean a relative upregulation in CXCR4^{bright}CD5^{dim} subpopulation

Supplemental Table 4: GO term analysis of differentially expressed genes in CXCR4/CD5 subpopulations

GO Term*	Description	P-value	FDR q-value	Enrichment	B	b
GO:0110053	regulation of actin filament organization	6.39E-12	7.89E-9	2.72	228	54
GO:0032956	regulation of actin cytoskeleton organization	8.98E-12	1.02E-8	2.48	292	63
GO:0032970	regulation of actin filament-based process	6.71E-11	4.52E-8	2.33	320	65
GO:0030832	regulation of actin filament length	9.03E-11	5.57E-8	3.00	157	41
GO:0008064	regulation of actin polymerization or depolymerization	2.79E-10	1.59E-7	2.94	156	40
GO:0007015	actin filament organization	3.46E-9	1.55E-6	2.46	224	48
GO:0030833	regulation of actin filament polymerization	1.62E-8	5.84E-6	2.79	144	35
GO:0030029	actin filament-based process	1.52E-7	4.68E-5	2.00	345	60
GO:0030036	actin cytoskeleton organization	1.95E-7	5.9E-5	2.10	284	52
GO:0030838	positive regulation of actin filament polymerization	1.63E-6	3.71E-4	2.19	71	20
GO:0032233	positive regulation of actin filament bundle assembly	3.86E-5	3.86E-3	2.91	54	15
GO:0034314	Arp2/3 complex-mediated actin nucleation	7.41E-5	6.42E-3	1.73	14	7
GO:0008154	actin polymerization or depolymerization	1.01E-4	8.13E-3	2.83	52	14
GO:0045010	actin nucleation	3.41E-4	2.02E-2	1.97	22	8
GO:0030834	regulation of actin filament depolymerization	4.3E-4	2.44E-2	2.93	46	12
GO:0032231	regulation of actin filament bundle assembly	7.97E-4	4.13E-2	1.29	84	17
GO:0040012	regulation of locomotion	1.35E-5	1.8E-3	3.44	813	106
GO:0040017	positive regulation of locomotion	1.21E-4	9.35E-3	1.92	461	64
GO:0040011	locomotion	1.65E-4	1.14E-2	1.73	835	103
GO:2000145	regulation of cell motility	3.78E-6	6.75E-4	1.62	753	102
GO:2000147	positive regulation of cell motility	4.5E-5	4.38E-3	2.11	437	63
GO:0048870	cell motility	3.31E-4	1.98E-2	1.25	763	94
GO:0030334	regulation of cell migration	3.58E-6	6.63E-4	2.76	706	97
GO:0030335	positive regulation of cell migration	2.91E-5	3.19E-3	2.11	422	62
GO:1903039	positive regulation of leukocyte cell-cell adhesion	1.77E-6	3.98E-4	1.77	210	40
GO:0030155	regulation of cell adhesion	4.11E-6	7.25E-4	3.42	608	86
GO:0045785	positive regulation of cell adhesion	1.1E-5	1.53E-3	4.51	366	57
GO:0022409	positive regulation of cell-cell adhesion	1.27E-5	1.72E-3	3.04	243	42
GO:1903037	regulation of leukocyte cell-cell adhesion	7.02E-5	6.19E-3	5.74	286	45
GO:0022407	regulation of cell-cell adhesion	2.55E-4	1.63E-2	2.11	373	53

GO:0032956	regulation of actin cytoskeleton organization	8.98E-12	1.02E-8	2.48	292	63
GO:0030036	actin cytoskeleton organization	1.95E-7	5.9E-5	2.10	284	52
GO:0051493	regulation of cytoskeleton organization	5.68E-7	1.5E-4	1.80	467	73
GO:0051495	positive regulation of cytoskeleton organization	1.55E-6	3.63E-4	3.13	179	36
GO:0007010	cytoskeleton organization	1.7E-4	1.17E-2	1.31	739	93
GO:0110053	regulation of actin filament organization	6.39E-12	7.89E-9	2.72	228	54
<p>*GO Term analysis was performed by GOrilla online tool and showed GO Terms are referring to group „Processes“.</p> <p>Enrichment = $(b/n) / (B/N)$ N - is the total number of genes B - is the total number of genes associated with a specific GO term n - is the number of genes in the top of the user's input genes b - is the number of genes in the intersection</p>						

Supplemental Table 5: Statistical analysis of all measurements for cell motility assay

<i>Parameter</i>	MEC1^{wt} cells (mean ± s.e.m.)	MEC1^{GAB1-KO} cells (mean ± s.e.m.)	ANOVA P-value
<i>Circularity*</i>	0.652 ± 0.004	0.653 ± 0.004	n.s.
<i>Elongation*</i>	0.323 ± 0.008	0.332 ± 0.008	n.s.
<i>Mass [pg]*</i>	200.6 ± 5.3	198.7 ± 4.8	n.s.
<i>Spreading [µm²/pg]*</i>	1.75 ± 0.02	1.57 ± 0.02	n.s.
<i>Displacement [µm/60 s]*</i>	3.75 ± 0.14	2.71 ± 0.10	0.035
<i>Persistence*</i>	0.69 ± 0.01	0.60 ± 0.01	0.036
Number of movies	6	6	
Number of tracked cells	195	186	
Number of time point measurements	11 939	15 184	
s.e.m. stands for standard error of the mean			
* For the description of the categories see <i>Supplemental Methods</i> above			

Supplemental Table 6: Transcription factors differentially expressed in CXCR4/CD5 subpopulations

Gene ID	Gene name	Base Mean (reads per million)	Fold Change (log2) *	adjusted P-value
ENSG00000112182	BACH2	10948	2.05	2.99E-44
ENSG00000166949	SMAD3	290	1.81	4.59E-14
ENSG00000186350	RXRA	2740	1.72	1.55E-11
ENSG00000101190	TCFL5	715	1.44	1.75E-14
ENSG00000118513	MYB	290	1.42	5.37E-05
ENSG00000150907	FOXO1	3615	1.30	7.29E-21
ENSG00000123358	NR4A1	578	1.23	3.30E-10
ENSG00000177683	THAP5	919	1.08	3.70E-22
ENSG00000118689	FOXO3	1987	1.06	1.26E-13
ENSG00000182568	SATB1	1161	1.05	4.81E-20
ENSG00000160961	ZNF333	1259	1.04	5.76E-20
ENSG00000106624	AEBP1	4585	1.03	4.06E-06
ENSG00000197647	ZNF433	497	1.02	1.89E-14
ENSG00000074657	ZNF532	1534	1.00	3.05E-05
ENSG00000102870	ZNF629	365	1.00	9.21E-15
ENSG00000134954	ETS1	21200	1.00	1.97E-19
ENSG00000137309	HMG1A1	1173	-1.05	2.13E-26
ENSG00000198715	GLMP	432	-1.20	2.04E-18
ENSG00000198176	TFDP1	1457	-1.26	2.90E-21
ENSG00000011590	ZBTB32	2944	-1.48	3.1715E-24
* positive numbers mean a relative upregulation in CXCR4 ^{bright} CD5 ^{dim} subpopulation				

Supplemental Table 7: Numbers of predicted transcriptional binding sites in promotor region of *GAB1* gene for transcription factors differentially expressed in CXCR4/CD5 subpopulations

Gene ID	Gene name	Number of predicted TBS*
ENSG00000112182	BACH2	1
ENSG00000166949	SMAD3	6
ENSG00000186350	RXRA	15
ENSG00000101190	TCFL5	25
ENSG00000118513	MYB	15
ENSG00000150907	FOXO1	11
ENSG00000123358	NR4A1	5
ENSG00000177683	THAP5	Not in database
ENSG00000118689	FOXO3	7
ENSG00000182568	SATB1	25
ENSG00000160961	ZNF333	4
ENSG00000106624	AEBP1	Not in database
ENSG00000197647	ZNF433	Not in database
ENSG00000074657	ZNF532	Not in database
ENSG00000102870	ZNF629	Not in database
ENSG00000134954	ETS1	58
ENSG00000137309	HMGA1	5
ENSG00000198715	GLMP	Not in database
ENSG00000198176	TFDP1	0
ENSG00000011590	ZBTB32	Not in database
*performed with ContraV3 online tool		

Supplemental Table 8: Table of predicted FoxO1 binding sites in the upstream region of *GAB1* gene

Transcription binding site (TBS) number	Approximate position (hg19 version)	ChIP results for FoxO1-enrichment in the region upstream of <i>GAB1</i>
1	chr4:144256738	No enrichment
2	chr4:144256758	No enrichment
3	chr4:144256434	Not investigated
4	chr4:144256269	Enrichment (Fig. 3B)
5	chr4:144256326	Enrichment (Fig. 3B)
6	chr4:144255992	No enrichment
7	chr4:144255229	Not investigated
8	chr4:144254941	Not investigated
9	chr4:144254638	Not investigated
10	chr4:144254529	Enrichment (Fig. 3B)
11	chr4:144254457	Not investigated
<p>Note: TBS 1 + 2 and 4 + 5 were analyzed with the same primer sets due to close vicinity from each other.</p>		

Supplemental References

1. Mraz M, Malinova K, Kotaskova J, et al. miR-34a, miR-29c and miR-17-5p are downregulated in CLL patients with TP53 abnormalities. *Leukemia*. 2009;23(6):1159-1163.
2. Mraz M, Malinova K, Mayer J, Pospisilova S. MicroRNA isolation and stability in stored RNA samples. *Biochem Biophys Res Commun*. 2009;390(1):1-4.
3. Pavlasova G, Borsky M, Svobodova V, et al. Rituximab primarily targets an intra-clonal BCR signaling proficient CLL subpopulation characterized by high CD20 levels. *Leukemia*. 2018;32(9):2028-2031.
4. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*. 2011;17(1):3 %J EMBnet.journal.
5. Hackenberg M, Sturm M, Langenberger D, Falcon-Perez JM, Aransay AM. miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. *Nucleic Acids Res*. 2009;37(Web Server issue):W68-76.
6. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120.
7. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42(Database issue):D68-73.
8. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
9. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
10. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-121.
11. Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*. 2016;44(W1):W147-153.
12. Yart A, Laffargue M, Mayeux P, et al. A critical role for phosphoinositide 3-kinase upstream of Gab1 and SHP2 in the activation of ras and mitogen-activated protein kinases by epidermal growth factor. *J Biol Chem*. 2001;276(12):8856-8864.
13. Wiederschain D, Wee S, Chen L, et al. Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle*. 2009;8(3):498-504.
14. Carey MF, Peterson CL, Smale ST. Chromatin immunoprecipitation (ChIP). *Cold Spring Harb Protoc*. 2009;2009(9):pdb prot5279.
15. Strbkova L, Zicha D, Vesely P, Chmelik R. Automated classification of cell morphology by coherence-controlled holographic microscopy. *J Biomed Opt*. 2017;22(8):1-9.
16. Zicha D, Genot E, Dunn GA, Kramer IM. TGFbeta1 induces a cell-cycle-dependent increase in motility of epithelial cells. *J Cell Sci*. 1999;112 (Pt 4):447-454.
17. Dunn GA, Zicha D. Dynamics of fibroblast spreading. *J Cell Sci*. 1995;108 (Pt 3):1239-1249.
18. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
19. Papatheodorou I, Moreno P, Manning J, et al. Expression Atlas update: from tissues to single cells. *Nucleic Acids Res*. 2020;48(D1):D77-D83.
20. Kreft L, Soete A, Hulpiau P, Botzki A, Saeys Y, De Bleser P. ConTra v3: a tool to identify transcription factor binding sites across species, update 2017. *Nucleic Acids Res*. 2017;45(W1):W490-W494.
21. Pyrzynska B, Dwojak M, Zerrouqi A, et al. FOXO1 promotes resistance of non-Hodgkin lymphomas to anti-CD20-based therapy. *Oncoimmunology*. 2018;7(5):e1423183.