Supplemental Data

Supplemental figure legends and Tables

Supplemental Figure 1 GM-CSF does not activate the GM-CSFR signaling pathway in chronic lymphocytic leukemia (CLL) cells. (**A**) Western blot analysis of peripheral blood (PB) CLL cells. As shown, a 24-hour incubation of CLL low-density cells with increasing concentrations of GM-CSF did not induce tyrosine phosphorylation of STAT3 or STAT5 or affect the phosphorylation status of AKT or ERK. As positive controls (c), protein extracts of HeLa (for STAT3 and STAT5) and Jurkat (for AKT and ERK) cells were used. (**B**) JAK2 immunoprecipitation and detection of phosphotyrosine JAK2 following proteins separation by SDS-PAGE. CLL cells were incubated with increasing concentrations of GM-CSF as described above, and following immunoprecipitation, pJAK2 was detected using mouse anti-human antibodies. As shown, pJAK was not detected following incubation with GM-CSF (I.P., immunoprecipitate; B, beads control). (**C**) Flow cytometry analysis of CLL cells incubated with GM-CSF or GM-CSFR neutralizing antibodies. CLL cells were incubated with 100 ng GM-CSF for 24 or 48 h. As shown in the upper panel, incubation with GM-CSF did not increase the levels of the cell surface protein CD20, and as shown in the lower panel, incubation of CLL cells with GM-CSFRα antibodies for 24 hours did not affect cell surface CD20 levels.

Supplemental Figure 2 Sequence of the 5'-flanking region of the human GM-CSFRα promoter. Analysis of the promoter region was performed using the TFSEARCH (Transcription factor Search) version 1.3 and TESS (Transcription Element Search System) version 6.0 software. Based on the findings from the above search, we designed the primers and probes to identify which of the GAS-like elements binds STAT3. The putative STAT3 binding sites in the proximal region of the GM-CSFRα promoter are shown in red, and those in the distal region of the GM-CSFRα promoter are shown in underlined black. Green denotes the first exon region, and blue denotes a TG-rich sequence.

Supplemental Figure 3 Sequence of the 5'-flanking region of the human GM-CSFRα promoter. The putative STAT3-binding sites that we studied are highlighted in red.

Supplemental Figure 4 The proteins pulled-down together with GM-CSFRα and detected by mass spectrometry analysis.

Supplemental Table 1. Patient Characteristics

CLL pt. #	Sex	Age (yr)	WBC (10^9/l)	BM Lymph. %	Hb. (g/dl)	Plts. (10^9/l)	Rai stage	β_2M (mg/dl)	V_H mutation	Cytogenetics /FISH
1	F	70	23	78	14	354	0	2.3	N	13q
2	M	61	17	68.4	16.7	164	1	1.6	Y	T12
3	M	69	57.1	73	13.3	166	1	3.8	Y	ND
4	M	56	48.3	79	13.1	236	1	2	Y	46, XY
5	F	65	33	80	12.8	310	0	2.5	Y	ND
6	M	70	33	65	15.6	192	1	4.7	N	46,XY
7	M	64	21.4	ND	14.6	243	0	2.5	ND	46,XY
8	M	85	24	57	12.4	163	1	3.9	Y	13q
9	F	46	55.8	58	13.4	248	1	2.2	Y	13q
10	F	81	47.3	69	11.9	132	1	4.1	N	11q
11	F	69	116.8	84	13	169	1	2.9	Y	13q
12	M	68	26.8	52	14.7	251	1	2.2	N	13q
13	M	78	40.9	58	14.6	169	1	2	Y	13q
14	M	70	17.4	54	14.8	147	1	1.9	Y	13q
15	F	67	14.8	77	11.4	535	1	1.8	Y	46,XX
16	F	59	50.9	63	14.2	306	1	2	Y	46,XX
17	M	57	85.9	82	15.1	241	1	2.5	N	46,XY
18	M	82	139.1	79	12.7	122	1	3.9	Y	T12
19	F	69	28.7	80	12.4	119	2	4.4	Y	T12

Abbreviations; Yr., years; WBC, white blood cells; lymph., Lymphocytes; Hb, hemoglobin; Pts., platelets; β_2M, β_2 microglobulin; M, male F, female; ND, not done or not available; V_H mutation, hypermutation of the immunoglobulin heavy chain gene presented as N (negative; if % derivation from the germline sequence is $\leq$ 2%); or Y (positive, if % derivation from germline sequence is > 2%); FISH, fluorescence in situ hybridization.

Supplemental Table 2 Primers and Probes

Primers used in the luciferase assay

Hu-GM-CSF -4012-5 GCG GGT ACC GGT GAA AGT CAG AAA CAG GGT

Hu-GM-CSF -2965-5 GCG GGT ACC CGT TGT TGG ACA TTT AGG TTG

Hu-GM-CSF -2517-5 GCG GGT ACC GAA TGG ACT AAG ACA GCT CCT

Hu-GM-CSF -496-5 GCG GGT ACC TGT ATG CCC ATG TGC ACT GTG

Hu-GM-CSF -23-3 GCG AGA TCT CTT CTG AGT AGC TCC CTT CAG

Primers used for the ChIP assay

GM-CSF-1-F -3605 ATT CTC AGT TTC CAG GAG AGC

GM-CSF-1-R AGA CAT TGT GTG TGG AGA GAC -3502

GM-CSF-2-F -3418 GGA GAA CCA CTA TGA GCC GAA

GM-CSF-2-R GTG CAC ACG TAC CCT AGA ACT -3259

GM-CSF-3-F -3018 GCT GCA TAG TAT TCC ATG GTG

GM-CSF-3-R CAA CCT AAA TGT CCA ACA ACG -2945

GM-CSF-4-F -2551 TAC CCA ACC TCA GGT ATT CCT

GM-CSF-4-R GCA AAG GTG AAA GAT GGC TGA -2436

GM-CSF-5-F -94 CAA TGA ACT CAC GGA GCA AT

GM-CSF-5-R CCC TTC AGA GTT CCT CTG TGT C +9

Probes used for EMSA

Probe-1 GAGCTGG -3581 TTGTTGAAAA -3572 GACCC

Probe-1-mutant GAGCTGG -3581 T*AGA*AG AAAA -3572 GACCC

Probe-3 TGCCACA -2984 TTTTCTTAA -2976 TCCAGTC

Probe-3-mutant TGCCACA -2984 T*AAA*CTT AA -2976 TCCAGTC

Probe-5 TTACAGG -68 TTTCCCAA -61 TCCTATGAAA

Probe-5-mutant TTACAGG -68 T*GG*CCCAA -61 TCCTATGAAA

Probe-2 TTACC -3372 TTTTTTTAAAAA -3383 TTTTA

Probe-4 ATCAA -2551 TTACCCAA TTTCCAGCAA -2521 CATGA

Supplemental Methods

Expression of Human GM-CSFRα in 293FT Cells

To characterize the epitope specificity of the GM-CSFRα antibody and to explore the intracellular localization of GM-CSFRα, we used CLL-cell DNA to generate full-length (amino acids 1–400) GM-CSFRα and its extracellular (amino acids 1–323) and intracellular (amino acids 347–400) regions in 293FT cells. To generate these constructs, we designed primers containing overhangs that included restriction sites for *Hin*dIII or BamHI on the 5' and 3' end, a Kozac sequence upstream of the ATG start codon to enhance the efficiency of the *expression in mammalian cells, and a start codon*. The primers used were full-length (from the start codon ATG to 1492bp): forward, 5´-GCGGCGAAGCTTCGCCACCATGCTTCTCCTGGTGACAAGC-3' and reverse, 5´-GCGGCGGGATCCGGTAATTTCCTTCACGGTCAAG-3´; the primers for the extracellular domain (from ATG to 1261bp) were forward, 5'-GCGGCGAAGCTTCGCCACCATGCTTCTCCTGGTGACAAGC-3' and reverse, 5'-GCGGCGGGATCCGCCGAGGTTCCCGTCGTCAG-3'; and the primers for the intracellular domain (from 1331bp to 1492bp) were forward, 5'-GCGGCGAAGCTTCGCCACCATGAAAAGGTTCCTTAGGATACAG-3' and reverse, 5'-GCGGCGGGATCCGGTAATTTCCTTCACGGTCAAG-3'. For the fusion of GFP, we mutated the stop codon for the full-length and intracellular domain of human GM-CSFRα. GFP was pre-cloned into pcDNA3 vector at the BamHI/XhoI site. Human GM-CSFRα and its truncated fragments were cloned into the *Hin*dIII*/BamH*I site. After verifying each cloned cDNA by direct sequencing, the full-length and the GM-CSFRα fragments were transfected into 293FT cells using Lipofectamine 2000 (Invitrogen) and incubated for 72 hours, in accordance with the manufacturer's instructions. Following incubation, the cells were harvested, and transfection efficiency was assessed by flow cytometry.

Epitope Mapping of the GM-CSFRα Antibodies

The coding sequences of full-length extracellular and intracellular GM-CSFRα domains were generated by PCR and cloned into a mammalian expression system using the pcDNA 3.1 vector (Invitrogen) as described above. The pcDNA3.1 vector was modified by adding the sequence coding of GFP. Human 293FT cells were transfected using recombinant plasmids, incubated for 72 hours, and their cell pellets stored at -80°C and analyzed by Western immunoblotting using anti-GFP antibodies.

Photoactivatable Ribonucleoside-Enhanced Crosslink and Immunoprecipitation (PAR-CLIP)

We used the PAR-CLIP method to incorporate the photoreactive ribonucleoside analog 4-thiouridine into nascent RNA transcripts of 293FT cells. GFP-tagged human GM-CSFRα was expressed in 293FT cells as described above. The cells were incubated overnight in DMEM supplemented with 10% FBS in the presence of 100 µM of 4-thiouridine (4SU, Sigma-Aldrich). Viable cells were washed with cold PBS, exposed to 302-nm UV light for 1 minute, lysed in NP-40 lysis buffer on ice for 10 minutes, and digested with 1 U/µl of RNase T1 (Invitrogen/Ambion) at room temperature for 15 minutes. Sepharose beads coated with anti-GFP antibodies (Abcam) were washed, suspended in RNase T1-treated cell lysates, and incubated overnight on a rotating shaker. Then the beads with the captured proteins were washed in washing buffer and digested with 100 U/µl of RNase T1 for 15 minutes at room temperature. Treated beads were washed with high-salt wash buffer 3 times and resuspended with Proteinase K buffer followed by Proteinase K (Roche) at 55°C for 30 minutes. Total RNA was extracted by using TRIzol LS reagent (Invitrogen). The cDNA library was prepared with a SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen) for Illumina Next Generation Sequencing.

Bioinfomatics and Data Analysis

To align RNA sequence reads to the reference human genome 19 (Hg19; February 2009, GRCh37 (http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/index.shtml)), we used the Partek Flow software (version 2.1, 2012; Partek) bowtie algorithm. Annotated mRNA reads were exported to the Partek Genomic Suite Package (version 6.6, 2012; Partek) for downstream analysis. To identify differentially expressed genes, we used the 1-way ANOVA and the Benjamini-Hochberg correction for multiple comparisons with a false discovery rate of 0.05, and we created a heat map with an average-linkage clustering of selected genes. To identify pathways enriched for differentially expressed genes in the 2 experimental conditions, we used the KEGG pathway database (http://www.genome.jp/kegg/pathway.html).

Mass Spectrometry Analysis

The silver-stained bands from the pull-down assay were destained and subjected to in-gel digestion with 200 ng of modified trypsin (sequencing grade, Promega) at 37°C for 24 hours. The resulting peptides were analyzed by nano-liquid chromatography-coupled ion trap mass spectrophotometry with online desalting on a system consisting of a FAMOS autosampler, an UltiMate Nano liquid

chromatography module, and a Switchos precolumn switching device on a 75-µm by 150-mm C_{18} column (all from Dionex). Electrospray ion trap mass spectrometry was performed on a LTQ linear ion trap mass spectrometer (Thermo Scientific). Resulting proteins were then identified by a database search for the fragment spectra using the National Center for Biotechnology Information nonredundant protein database, using Mascot (Matrix Science) and SEQUEST (Thermo Scientific) results. Resulting peptide matches were manually curated.

Apoptosis Assay

The rate of cellular apoptosis was analyzed using double staining with a Cy5-conjugated Annexin V kit and PI (BD Biosciences), according to the manufacturer's instructions. Briefly, 1×10^6 cells were washed once with PBS and suspended in 200 µl of binding buffer with 0.5 µg/ml of Annexin V-Cy5 and 2 µg/ml of PI. After incubation for 15 minutes in the dark at room temperature, the samples were analyzed on a FACSCalibur flow cytometer (Becton Dickinson). Cell viability was calculated as the percentage of Annexin V-positive cells, and statistical analysis was performed using the CellQuest Pro program.

Fig. S1

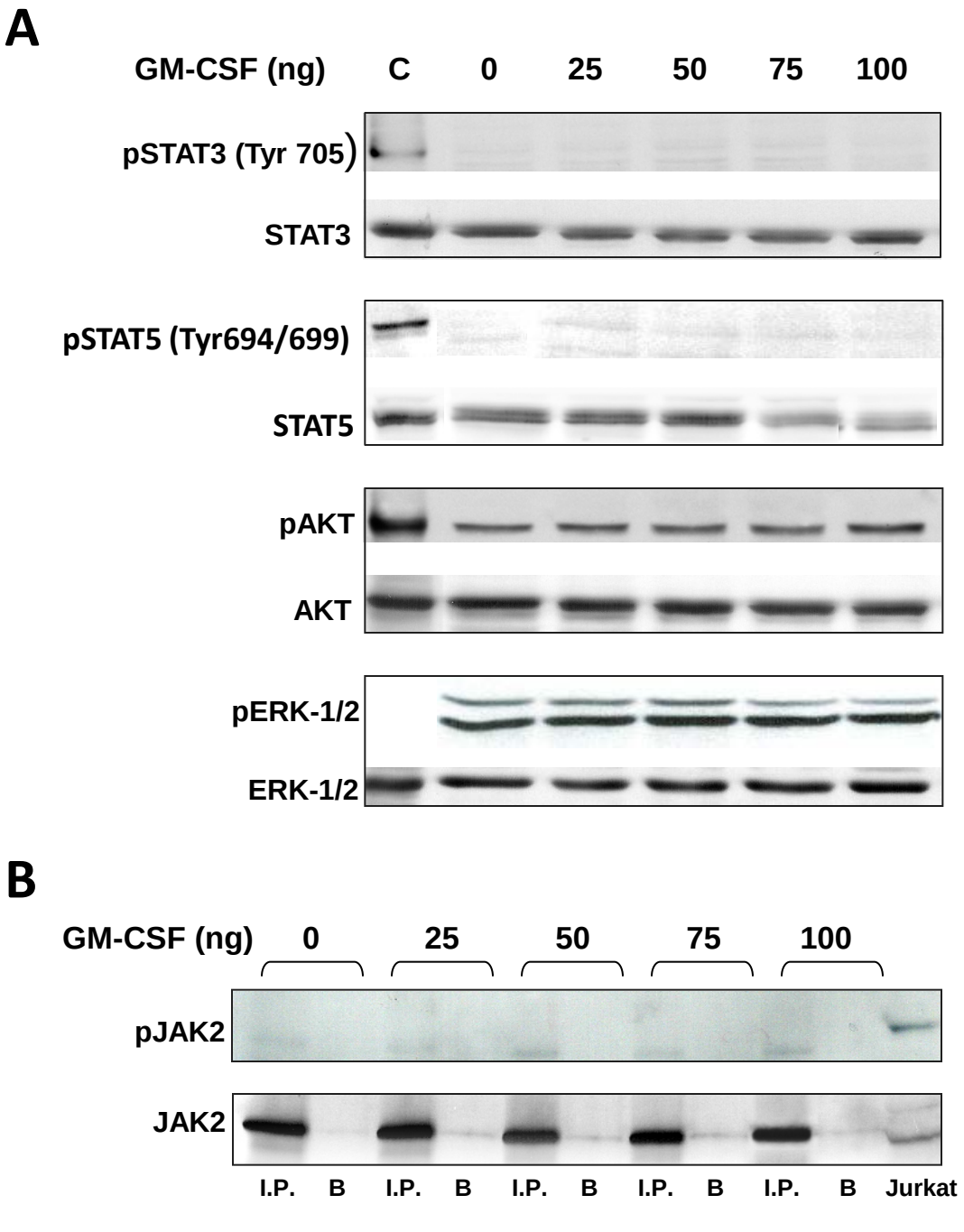
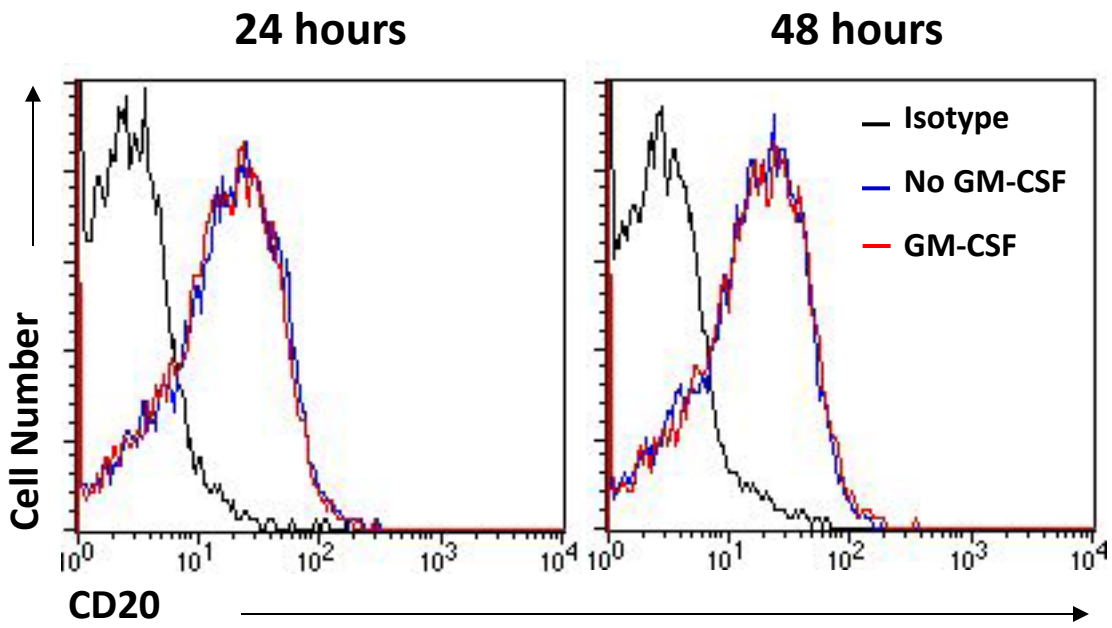

A
GM-CSF (ng)
C
0
25
50
75
100
pSTAT3 (Tyr 705)
STAT3
pSTAT5 (Tyr694/699)
STAT5
pAKT
AKT
pERK-1/2
ERK-1/2
B
GM-CSF (ng)
0
25
50
75
100
pJAK2
JAK2
I.P.
B
I.P.
B
I.P.
B
I.P.
B
I.P.
B
Jurkat

Fig. S1

C


24 hours
48 hours
Isotype
No GM-CSF
GM-CSF
Cell Number
CD20
10^0
10^1
10^2
10^3
10^4

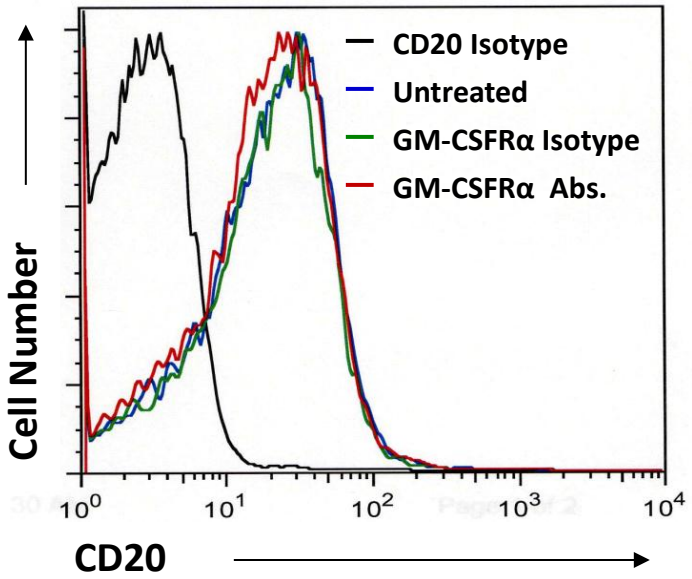

CD20 Isotype
Untreated
GM-CSFRα Isotype
GM-CSFRα Abs.
Cell Number
CD20
10^0
10^1
10^2
10^3
10^4

Fig. S2


-5000 ACAGTGCCAGTTCCAGCAGCCTTGGTGAACCGTCTGGTTGAGCTCTTGGGAAGATAAACT -4941
-4940 TCTCTACATCAATGTCAAAAGGAAGATATGAGCCTTTGACAAGAATATCCACCATGATGC -4881
-4880 CCAGTTCCAGCAGCCTGGGTGAGCCATCTGGTTGAGCTCTTGGGAAGATAAACTTCTCTA -4821
-4820 CATCAATGTCAACAAGAAGATATGAGCCTTTGACAAAAAGATCCACCATGATGCCCAGTT -4761
-4760 CCAGCAGCCTGGGTGAGCCATCTGGTTGAGCTCTTGGGAAGATAAACTTCTCTACATCAA -4701
-4700 TGTCAACAAGAAGATATGAGCCTTTGACAAAAAGATCCACCGTGATGCCCAGTTCCAGCA -4641
-4640 GCCTGGCTGAGCCATCTGGTTGAGCTCTTGGGAAGATAAGCTTCTCTACATCAATGTCAA -4581
-4580 CAAGAAGATATGAGCCTTTGACAAAAAGATCCACCGTGATGCCCAGTTCCAGCAGCCTGG -4521
-4520 CTGAGCCATCTGGTTGAGCTCTTGGGAAGATAAACTTCTCTACATCAATGTCAACAAGAA -4461
-4460 GATATGAGCCTTTGACAAAAAGATCCACCATGATGCCCAGTTCCAGCAGCCTGGCTGAGC -4401
-4400 CATCTGGTTGAGCTCTTGGGAAGATAAACTTCTCTACATCAATGTCAACAAGAAGATATG -4341
-4340 AGCCTTTGACAAAAAGATCCACCATGATGCCCAGTTCCAGCAGCCTGGCTGAGCCATCTG -4281
-4280 GTTGAGCTCTTGGGAAGATAAACTTCTCTACATCAATGTCAACAAGAAGATATGAGCCTT -4221
-4220 TGACAAAAAGATCCACCATGATGCCCAGTTCCAGCAGCCTGGCTGAGCCATCTGGTTGAA -4161
-4160 TTCTTGGGAAGATAAACTTCTTTACATTGATGTCAACAACAGGGCATGATTCATTAAAGA -4101
-4100 CAAAAAGGCCCATTATTGTGCCCACTTCCAGCAGCCTTGGTGAGTCATCTGATTTGGCTC -4041
-4040 TTGGAAAGACAAACTTCTCTACATGAATGGTGAAAGTCAGAAACAGGGTATGAGCCATTG -3981
-3980 ACAAGCAGACCCACTATAGTTTGCAGTTTGAGCCTTGGTAAGCCATCTGGTTGAACTCTT -3921
-3920 GGAAAGTTAAATTTCTCTACATCGATGTCAGCAACAAAGTATGAGCCATTCACAAGAAGA -3861
-3860 TCCACCATGGTGGCCATTTCCAGCAGCCTTCGTGAGCCACCTGCTTGAGCTCTTGGGAGA -3801
-3800 TAACCTTCTCTACATCAATTCTGGAGAGAGGCAGCATGCTACAGTTTGGATACGGTTTGT -3741
-3740 TTGGTCCTGCCAGGTTTCATGTTGAAATTTGACCCCCTGTGTGGTAGGTGGGACCTGGTG -3681
-3680 GGAGATGTTTAAACCATGTGGGTGAATCCCGTATGAATGGCTTGGTCCTGCCCCCATAGT -3621
-3620 CATGAGTGAGTTCTCATTCTCAGTTTCCAGGAGAGCTGGTTGTTGAAAAGACCCTGGCAC -3561
-3560 CTCCCCCCGCTCTCTCTTGCTTCCTTTCTGCCCATGTGGTCTCTCCACACACAATGTCTT -3501
-3500 CCCTTCACTTTCTGTCATGACTGGAAGCAGCGTGACGTCCTCACCACAAGCAGGTCCAAA -3441
-3440 CACCACGCTTCCTGGACAACCTGGAGAACCACTATGAGCCGAATAAATAATCTCTTTTTT -3381
-3380 TTTTTACCTTTTTTTAAAAATTTTATTATTATTATACTTTAAGTTCTAGGGTACGTGTGC -3321
-3320 ACAACGTGCAGCTTTGTTACATATGTATCCATGTGCCATGTTGGTGTGCTGCACCCATTA -3261
-3260 ACTCGTCATTTAGCATTAGGTATGTCTCCTAATGCTGTCCCTCCCCCCTCCCCCACCCC -3201
-3200 ACAACAGTCCCCGGTGTGTGATGTTCCCCTTCCTGTGTCCATGTGTTCTCATGGTTCAAT -3141
-3140 TCCCACCTATGAGTGAGAACATGCGGTGTTTGGTTTTTTGTCCTTGCGATAGTTTGCTCA -3081
-3080 GAATGATGGTTTCCAGCTTCAACCCTGTCCCTACAAAGGACATGAACTCATCATGTTTTA -3021
-3020 TGGCTGCATAGTATTCCATGGTGTATATGTGCCACATTTTTCTTAATCCAGTCTATCGTTG -2961
-2960 TTGGACATTTAGGTTGGTTCCAAGTCTTTGCTATTGTGAATAGTGCCATTATAAACATAT -2901
-2900 GTGTGCATGTGTCTTTATAGCAGCATGATTTATAATCTTCTTTTTTTTTTGAGACACAGT -2841
-2840 CTCGCTCTGTCGCCGGGCTGGAGTGCAGTGGCAGCATCTCGGCTCACTGCAAGCTCCGCC -2781
-2780 TCCCGGGTTCAATTGATTCTCCTGCCTCAGCTTCCCAAGTAGCTGGGACTGCAGGTGCAC -2721
-2720 GCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCC -2661
-2660 AGGCTGGTCTTGATCTCTTGACCTCATGATCCGCCTGCCTCGGTCTCCCAAAGTGCTGGG -2601
-2600 ATTACAGGCGTGAGCCACCGCGCCCAGCCAATGATCTCTTCTTTATCAATTACCCAACCT -2541
-2540 CAGGTATTCCTTTCCAGCAACATGAATGGACTAAGACAGCTCCTGCAGGAATGGGGAGGC -2481
-2480 TAAGACGGTAGAGGTGCAGCCTGGTCAGCCATCTTTCACCTTTGCTGATGTTGCTATCCA -2421
-2420 GGTGTTTTCCATTGCATGTGCATGAAGCCCTATGCTTTGTTTAGGCGCCAACTGTTAGTA -2361
-2360 TCGTAACACATGTCCAGTCTGTTTGGTTGTTTTTTTTGAGACGGAGTTTTGCTCTTGTTG -2301
-2300 CCCAGGCTGGAGGGCAATGGCACCATCTCAGCTCACTGCAACCTCCATCTCCTGGGTTCA -2241
-2240 AGGGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACTCGCCACCACGCC -2181
-2180 TGGCTAATTTTTGTATTTTTTTTTTTTTTAGTGAGCCACTGCCCCTGGCCTGCGTGATGTT -2121
-2120 TCTTATGATTTTGTAGACAGCACAAAAGGCATCTGTATTTGTGTGTGTGCAAGTGTCTGT -2061
-2060 GCAGGTGTGTGTACACCAATATGCTTGTGCAGAGGCGTGTGCAGATGTGTGTGTAGTTGT -2001
-2000 GTGTATGGATAGGTGCACTGTGTGCAGGTGTGTGGGGATGTGTGTGTGCAGGTCTGTGTG -1941
-1940 TGCAGACATGCACGTGTATGCAGGTGTGTGTGTGTGTACAGATGTGTGTGCAGGTGTGTG -1881
-1880 TGCAGATGTGTGTGTAGTTGTGTGTATGCACATGTGCATTGTGTGCAGGTGTGTGGGGAT -1821
-1820 GTGTGTGTGCAGGTCTGTGTGTGCAGACATGCATGTGTATGGAGGTGTGTGTACAGACGT -1761
-1760 GCGTGTAGGTGTGTGTGCATGTGTGTGTGTGCAGATGTGTGTATGCACATGTGCACTGTG -1701
-1700 TGCAGGTGTGTGAGGATGTGTGTGTGCAGGTCTGTGTGTGCAGACATGCGTGTGTGTGCA -1641
-1640 GGTCTGTGTGTGCAGACATGCATGTGTATGCAGGTGTGTGTGTGTACAGATGTGTGTGCA -1581
-1580 GGTGTGTGTTTGCAGATGTGTGTGCAGATGTGCATGTGTCCTTGTGTATATGCACATGTG -1521
-1520 CACTGTGTGCAGGTGTGTGTGTGTACAGATATGTGTGCAGGTGTGTGTGCAGATGTGCTT -1461
-1460 GTGTACAGATGTGTGTGTGTAGTTCTATGTACGCCCATGTGCACTGTGTGCAGGTGTGTG -1401
-1400 GGGATGTGTGTGTGCAGGTCTGTGTGTGCAGTTGTGTGTGCAGACATGCATGTGTGTGCA -1341
TG Rich Seqence
-1340 GGTGTGTGTGCAGGTGTGCGTGTGTAGACAAGTGTGTGTGTGCCATTGTGCTTGTTTGTG -1281
-1280 CAGACGTACGTGCAGGTGTGTGTGCAGATGTGTGTGTGTAGCTGTGTGTACACACATGTG -1221
-1220 CCCTGTGTGCAGGTGTGTGGGGATGTGTGCGCACGTCTGTGTGTGCAGACATGCATGTGC -1161
-1160 CTGCAGGAAAGTGGGGATGTGTGAGTGTGCAGGTCTGTGTGTGCAGATGTGTTTGCAGAC -1101
-1100 GTATGTGTGTGCAGGTGTGTGTGCAGGACTGTGCCAGTGTGCTTGTTTGTGCAGATGTGT -1041
-1040 GTGCAGATGTGTGTGTAGCTGTGTGTACGCCCGTGTGCACTGTGTGCAGGTGTGTGGGGA -0981
-0980 TGTGTGTGCAGTTGTGTGTGCAGACATGCATGTGTGTGCAGGTGTGTGCGCAGATGTGTG -0921
-0920 TGTAGTTGTGTGTATGCACATGTGCATTGTGTGCAGGTGTGTGGGGATGTGTGTGCAGGT -0861
-0860 CTGTGTGTGCAGATGTGTGTGCAAACATGCATGTGTGTGCAGGTGTGTGTGCAGATGTGC -0801
-0800 GTGTGTAGACAAGTGTGTGTGCCATTGTGCTTGTTTGTGCAGACGTACGTGCAGGTGTGT -0741
-0740 GTGCAGATGTGTGTGTGTAGCTGTGTGTACACACATGTGCCCTGTGTGCAGGTGTGTGGG -0681
-0680 GATGTGTGCGCACGTCTGTGTGTGCAGACATGCATGTGCCTGCAGGAAAGTGGGGATGTG -0621
-0620 TGAGTGTGCAGGTCTGTGTGTGCAGATGTGTGTGCAGACATGTATGTGTGTGCAGGTGTG -0561
-0560 TGTGCAGGACTGTGCCAGTGTGCTTGTTTGTGCAGATGTGTGTGCAGATATGTGTGTAGC -0501
-0500 TGTGTGTATGCCCATGTGCACTGTGTGCAGGTGTGTGGGGATGTGTGTGTGCAGGTCTGT -0441
-0440 GTGTGCAGTTGTGTGTGCAGACATGCATATGTGTGCAGGTGTGTGTGCAGATGTGTGTGT -0381
-0380 AGCTGTGTGTACGCCCGTGCGCACTGTGTGCAGGTGTGTGGAGATGTGTGTGTGTGCAGG -0321
-0320 TCTGTGTGTGCAGGTGTGTGTGCAAACATGCACACGCGTGCAAGCATGTGTGCAGGAGTG -0261
-0260 TGTGTGCACGTGTGTGCGGGCGTGTGGGAATGTGTGCGTGTGCAGGTCTGTGTGTGCAGA -0201
-0200 TGTGTGTGCAGACATACATGTGTGTGCAGGTGTGTGTGTGTAGATGTGTGTGGGACTCAC -0141
-0140 ACTTGAGCTGTTCCACGGGCCAGTCTTTGGAGACGAGGTCAGGAGCCAATGAACTCACGG -0081
-0080 AGCAATTACAGGTTTCCCAATCCTATGAAATGAGGAAGCAGGGGAGGGAGGGGCAGGGAG -0021
-0020 TGTATGATGACACAGAGGAACTCTGAAGGGAGCTACTCAGAAGCGGGAGTCTCCGAGAGA +0040

Fig. S3

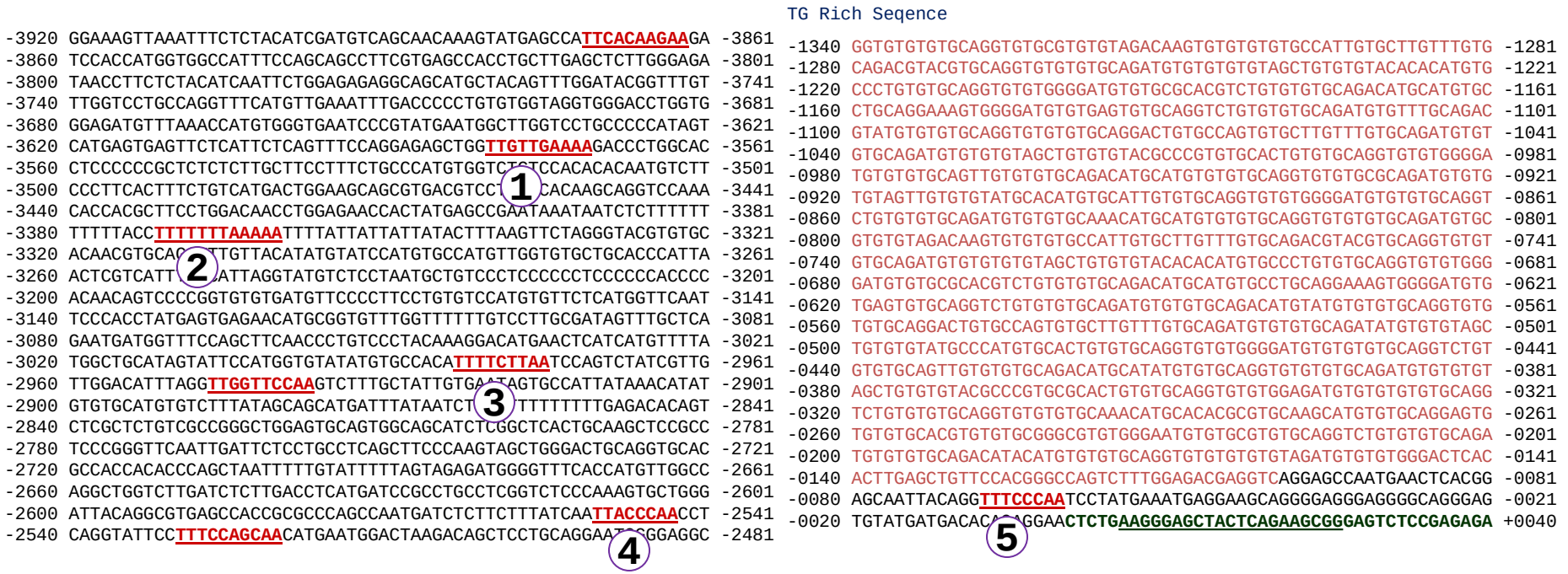

TG Rich Seqence
-3920 GGAAAGTTAAATTTCTCTACATCGATGTCAGCAACAAAGTATGAGCCATTCACAAGAAGA -3861
-3860 TCCACCATGGTGGCCATTTCCAGCAGCCTTCGTGAGCCACCTGCTTGAGCTCTTGGGAGA -3801
-3800 TAACCTTCTCTACATCAATTCTGGAGAGAGGCAGCATGCTACAGTTTGGATACGGTTTGT -3741
-3740 TTGGTCCTGCCAGGTTTCATGTTGAAATTTGACCCCCTGTGTGGTAGGTGGGACCTGGTG -3681
-3680 GGAGATGTTTAAACCATGTGGGTGAATCCCGTATGAATGGCTTGGTCCTGCCCCCATAGT -3621
-3620 CATGAGTGAGTTCTCATTCTCAGTTTCCAGGAGAGCTGGTTGTTGAAAAGACCCTGGCAC -3561
-3560 CTCCCCCCGCTCTCTCTTGCTTCCTTTCTGCCCATGTGGTCCTTCCACACACAATGTCTT -3501
-3500 CCCTTCACTTTCTGTCATGACTGGAAGCAGCGTGACGTCC[1]CACAAGCAGGTCCAAA -3441
-3440 CACCACGCTTCCTGGACAACCTGGAGAACCACTATGAGCCGAATAAATAATCTCTTTTTT -3381
-3380 TTTTTACCTTTTTTTAAAAATTTTATTATTATTATACTTTAAGTTCTAGGGTACGTGTGC -3321
-3320 ACAACGTGCAC[2]TTGTTACATATGTATCCATGTGCCATGTTGGTGTGCTGCACCCATTA -3261
-3260 ACTCGTCATT[2]ATTAGGTATGTCTCCTAATGCTGTCCCTCCCCCCTCCCCCCACCCC -3201
-3200 ACAACAGTCCCCGGTGTGTGATGTTCCCCTTCCTGTGTCCATGTGTTCTCATGGTTCAAT -3141
-3140 TCCCACCTATGAGTGAGAACATGCGGTGTTTGGTTTTTTGTCCTTGCGATAGTTTGCTCA -3081
-3080 GAATGATGGTTTCCAGCTTCAACCCTGTCCCTACAAAGGACATGAACTCATCATGTTTTA -3021
-3020 TGGCTGCATAGTATTCCATGGTGTATATGTGCCACATTTTTCTTAATCCAGTCTATCGTTG -2961
-2960 TTGGACATTTAGGTTGGTTCCAAGTCTTTGCTATTGTGAATAGTGCCATTATAAACATAT -2901
-2900 GTGTGCATGTGTCTTTATAGCAGCATGATTTATAATCT[3]TTTTTTTTGAGACACAGT -2841
-2840 CTCGCTCTGTCGCCGGGCTGGAGTGCAGTGGCAGCATCTCGGCTCACTGCAAGCTCCGCC -2781
-2780 TCCCGGGTTCAATTGATTCTCCTGCCTCAGCTTCCCAAGTAGCTGGGACTGCAGGTGCAC -2721
-2720 GCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCC -2661
-2660 AGGCTGGTCTTGATCTCTTGACCTCATGATCCGCCTGCCTCGGTCTCCCAAAGTGCTGGG -2601
-2600 ATTACAGGCGTGAGCCACCGCGCCCAGCCAATGATCTCTTCTTTATCAATTACCCAACCT -2541
-2540 CAGGTATTCCTTTCCAGCAACATGAATGGACTAAGACAGCTCCTGCAGGAA[4]GGAGGC -2481
-1340 GGTGTGTGTGCAGGTGTGCGTGTGTAGACAAGTGTGTGTGTGCCATTGTGCTTGTTTGTG -1281
-1280 CAGACGTACGTGCAGGTGTGTGTGCAGATGTGTGTGTGTAGCTGTGTGTACACACATGTG -1221
-1220 CCCTGTGTGCAGGTGTGTGGGGATGTGTGCGCACGTCTGTGTGTGCAGACATGCATGTGC -1161
-1160 CTGCAGGAAAGTGGGGATGTGTGAGTGTGCAGGTCTGTGTGTGCAGATGTGTTTGCAGAC -1101
-1100 GTATGTGTGTGCAGGTGTGTGTGCAGGACTGTGCCAGTGTGCTTGTTTGTGCAGATGTGT -1041
-1040 GTGCAGATGTGTGTGTAGCTGTGTGTACGCCCGTGTGCACTGTGTGCAGGTGTGTGGGGA -0981
-0980 TGTGTGTGCAGTTGTGTGTGCAGACATGCATGTGTGTGCAGGTGTGTGCGCAGATGTGTG -0921
-0920 TGTAGTTGTGTGTATGCACATGTGCATTGTGTGCAGGTGTGTGGGGATGTGTGTGCAGGT -0861
-0860 CTGTGTGTGCAGATGTGTGTGCAAACATGCATGTGTGTGCAGGTGTGTGTGCAGATGTGC -0801
-0800 GTGTGTAGACAAGTGTGTGTGCCATTGTGCTTGTTTGTGCAGACGTACGTGCAGGTGTGT -0741
-0740 GTGCAGATGTGTGTGTGTAGCTGTGTGTACACACATGTGCCCTGTGTGCAGGTGTGTGGG -0681
-0680 GATGTGTGCGCACGTCTGTGTGTGCAGACATGCATGTGCCTGCAGGAAAGTGGGGATGTG -0621
-0620 TGAGTGTGCAGGTCTGTGTGTGCAGATGTGTGTGCAGACATGTATGTGTGTGCAGGTGTG -0561
-0560 TGTGCAGGACTGTGCCAGTGTGCTTGTTTGTGCAGATGTGTGTGCAGATATGTGTGTAGC -0501
-0500 TGTGTGTATGCCCATGTGCACTGTGTGCAGGTGTGTGGGGATGTGTGTGTGCAGGTCTGT -0441
-0440 GTGTGCAGTTGTGTGTGCAGACATGCATATGTGTGCAGGTGTGTGTGCAGATGTGTGTGT -0381
-0380 AGCTGTGTGTACGCCCGTGCGCACTGTGTGCAGGTGTGTGGAGATGTGTGTGTGTGCAGG -0321
-0320 TCTGTGTGTGCAGGTGTGTGTGCAAACATGCACACGCGTGCAAGCATGTGTGCAGGAGTG -0261
-0260 TGTGTGCACGTGTGTGCGGGCGTGTGGGAATGTGTGCGTGTGCAGGTCTGTGTGTGCAGA -0201
-0200 TGTGTGTGCAGACATACATGTGTGTGCAGGTGTGTGTGTGTAGATGTGTGTGGGACTCAC -0141
-0140 ACTTGAGCTGTTCCACGGGCCAGTCTTTGGAGACGAGGTCAGGAGCCAATGAACTCACGG -0081
-0080 AGCAATTACAGGTTTCCCAATCCTATGAAATGAGGAAGCAGGGGAGGGAGGGGCAGGGAG -0021
-0020 TGTATGATGACACAAGGAACTCTGAAGGGAGCTACTCAGAAGCGGGAGTCTCCGAGAGA +0040
[5]

Fig. S4

HSP71_HUMAN
Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5
K2C1_HUMAN
Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
K1C9_HUMAN
Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3
K1C10_HUMAN
Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6
K2C6A_HUMAN
Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3
K2C6B_HUMAN
Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5
HSP7C_HUMAN
Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1
GRP78_HUMAN
78 kDa glucose-regulated protein OS=Homo sapiens GN=HSPA5 PE=1 SV=2
K2C5_HUMAN
Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3
K22E_HUMAN
Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2
FUBP2_HUMAN
Far upstream element-binding protein 2 OS=Homo sapiens GN=KHSRP PE=1 SV=4
TRAP1_HUMAN
Heat shock protein 75 kDa, mitochondrial OS=Homo sapiens GN=TRAP1 PE=1 SV=3
K1C14_HUMAN
Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4
KRT84_HUMAN
Keratin, type II cuticular Hb4 OS=Homo sapiens GN=KRT84 PE=1 SV=2
TRI25_HUMAN
E3 ubiquitin/ISG15 ligase TRIM25 OS=Homo sapiens GN=TRIM25 PE=1 SV=2
CSF2R_HUMAN
Granulocyte-macrophage colony-stimulating factor receptor subunit alpha OS=Homo sapiens GN=CSF2RA PE=1 SV=1
PABP1_HUMAN
Polyadenylate-binding protein 1 OS=Homo sapiens GN=PABPC1 PE=1 SV=2
ALBU_HUMAN
Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2
TIF1B_HUMAN
Transcription intermediary factor 1-beta OS=Homo sapiens GN=TRIM28 PE=1 SV=5
CH079_HUMAN
Uncharacterized protein KIAA1456 OS=Homo sapiens GN=KIAA1456 PE=2 SV=1
SPIR1_HUMAN
Protein spire homolog 1 OS=Homo sapiens GN=SPIRE1 PE=1 SV=2