Cell Stem Cell, Volume 7

Supplemental Information

Combinatorial Transcriptional Control

In Blood Stem/Progenitor Cells: Genome-wide

Analysis of Ten Major Transcriptional Regulators

Nicola K. Wilson, Samuel D. Foster, Xiaonan Wang, Kathy Knezevic, Judith Schütte, Polynikis Kaimakis, Paulina Chilarska, Sarah Kinston, Willem H. Ouwehand, Elaine Dzierzak, John E. Pimanda, Marella F.T.R. de Bruijn, and Berthold Göttgens

4 Supplemental Figures:

Supplemental Figure 1; relates to Figure 2 in paper

Supplemental Figure 2; relates to Figure 3 in paper

Supplemental Figure 3; relates to Figure 6 in paper

Supplemental Figure 4; relates to Figure 7 in paper

4 Supplemental Tables:

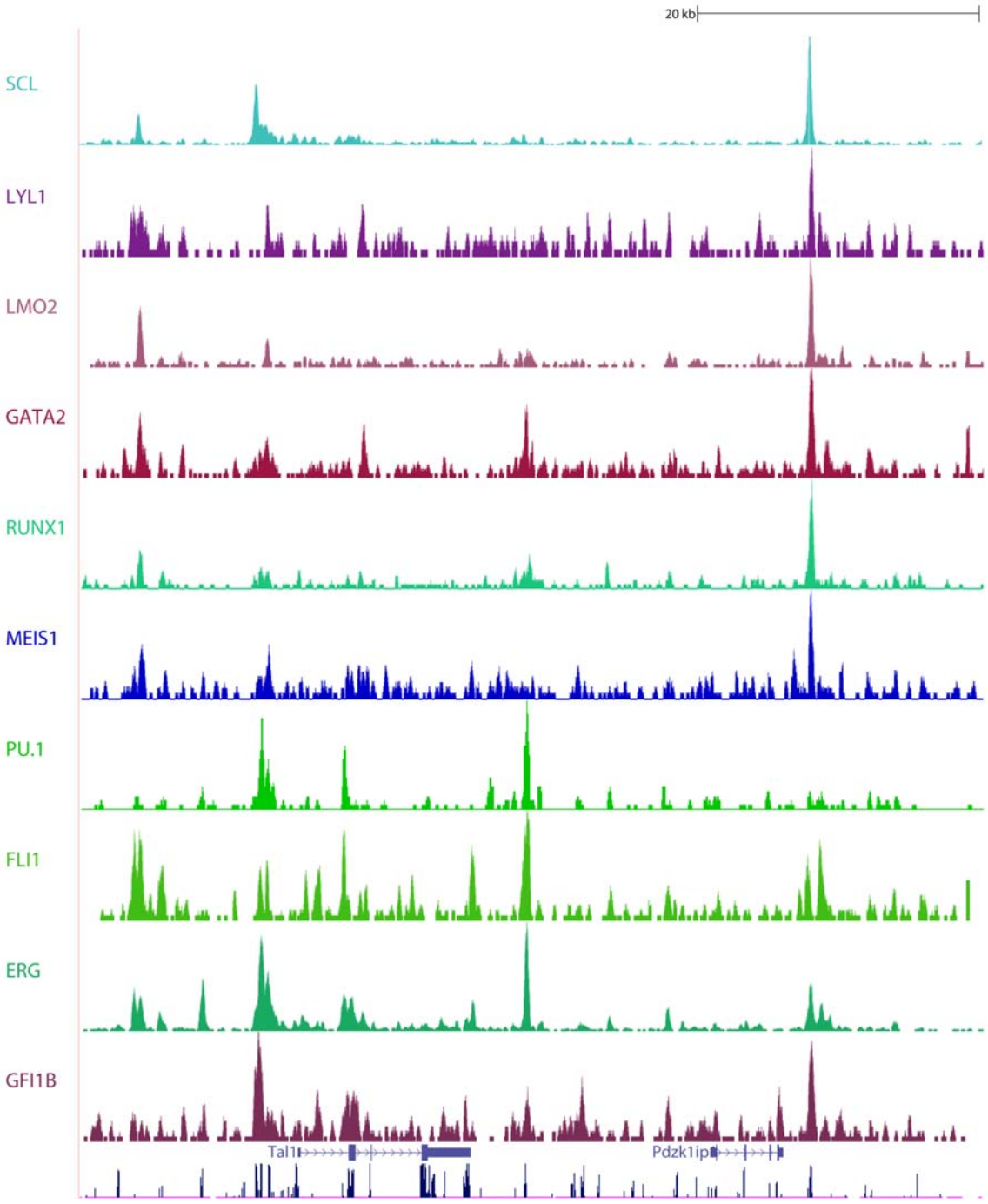
Supplemental Table 1: Excel spreadsheet with ChIP-Seq peak positions for all 10 transcription factors; relates to Figure 1 in paper

Supplemental Table 2: Excel spreadsheet with information for each mouse gene on ChIP-Seq peaks for all transcription factors; relates to Figure 1 in paper

Supplemental Table 3: Excel spreadsheet containing heptad peak positions; relates to Figure 4 in paper

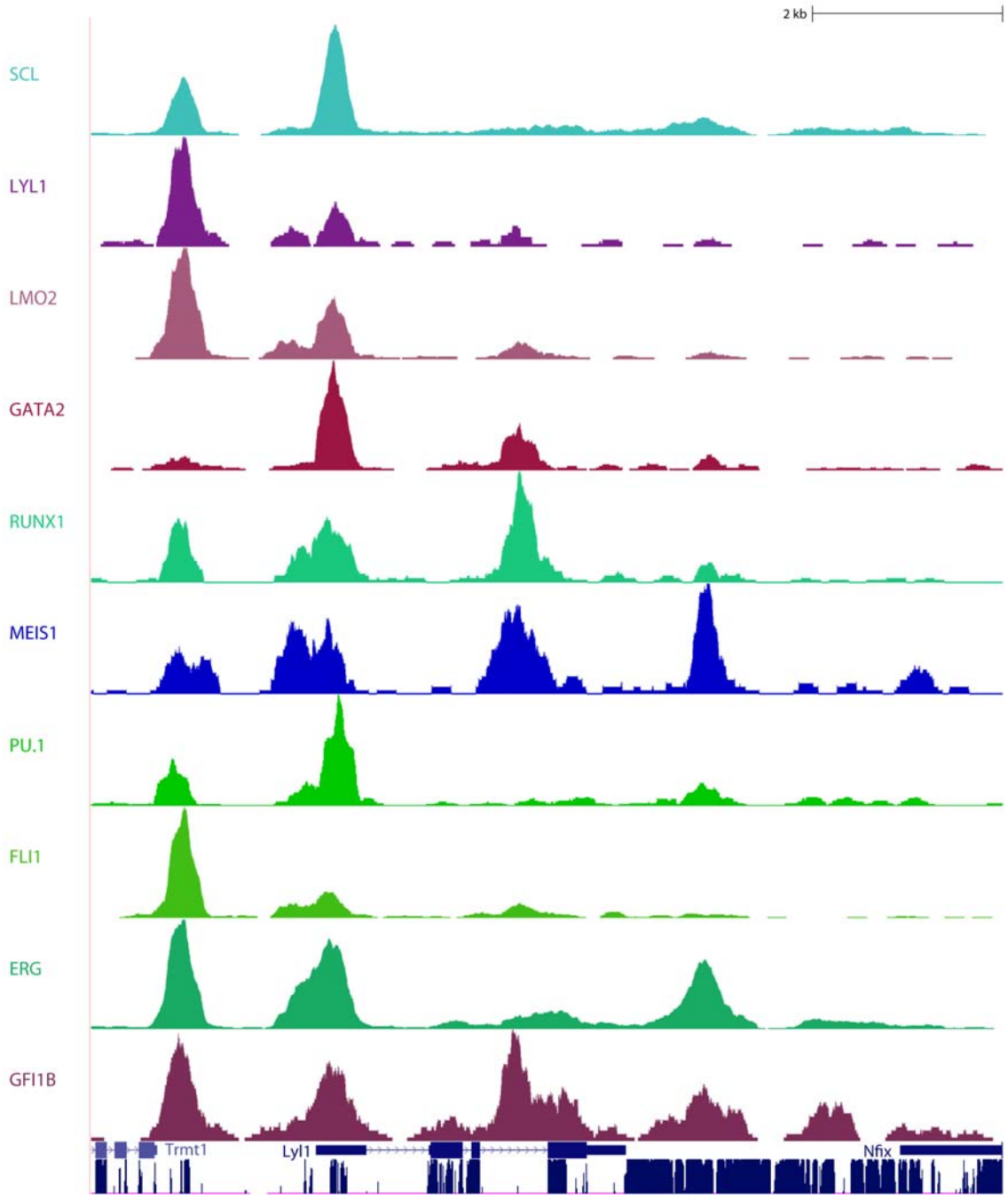
Supplemental Table 4: Excel spreadsheet containing list of heptad peaks and associated genes which do not contain either an E-box or a Runx motif; relates to Figure 5 in paper

A) i)



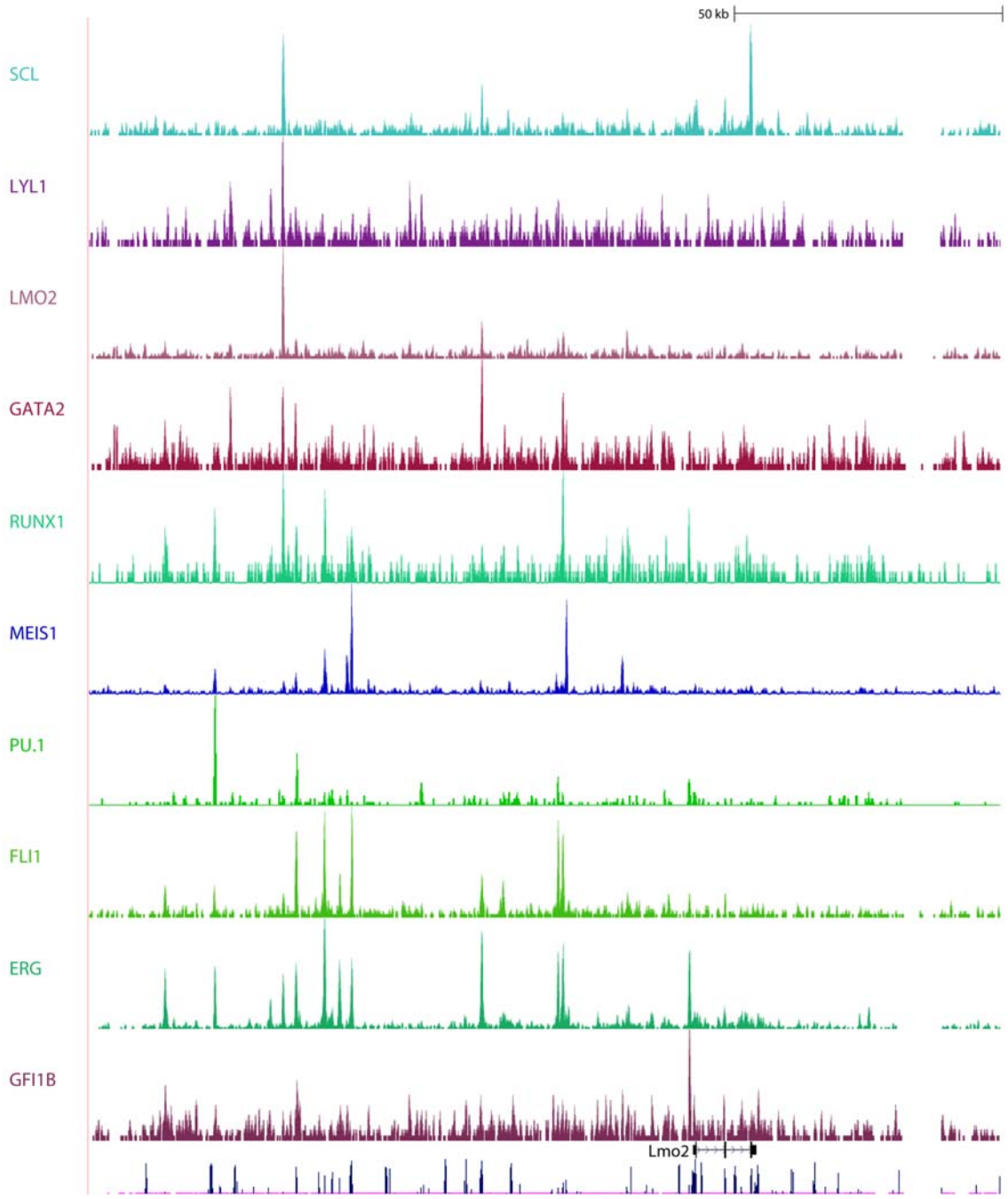
Supplementary Figure 1.

A) ii)



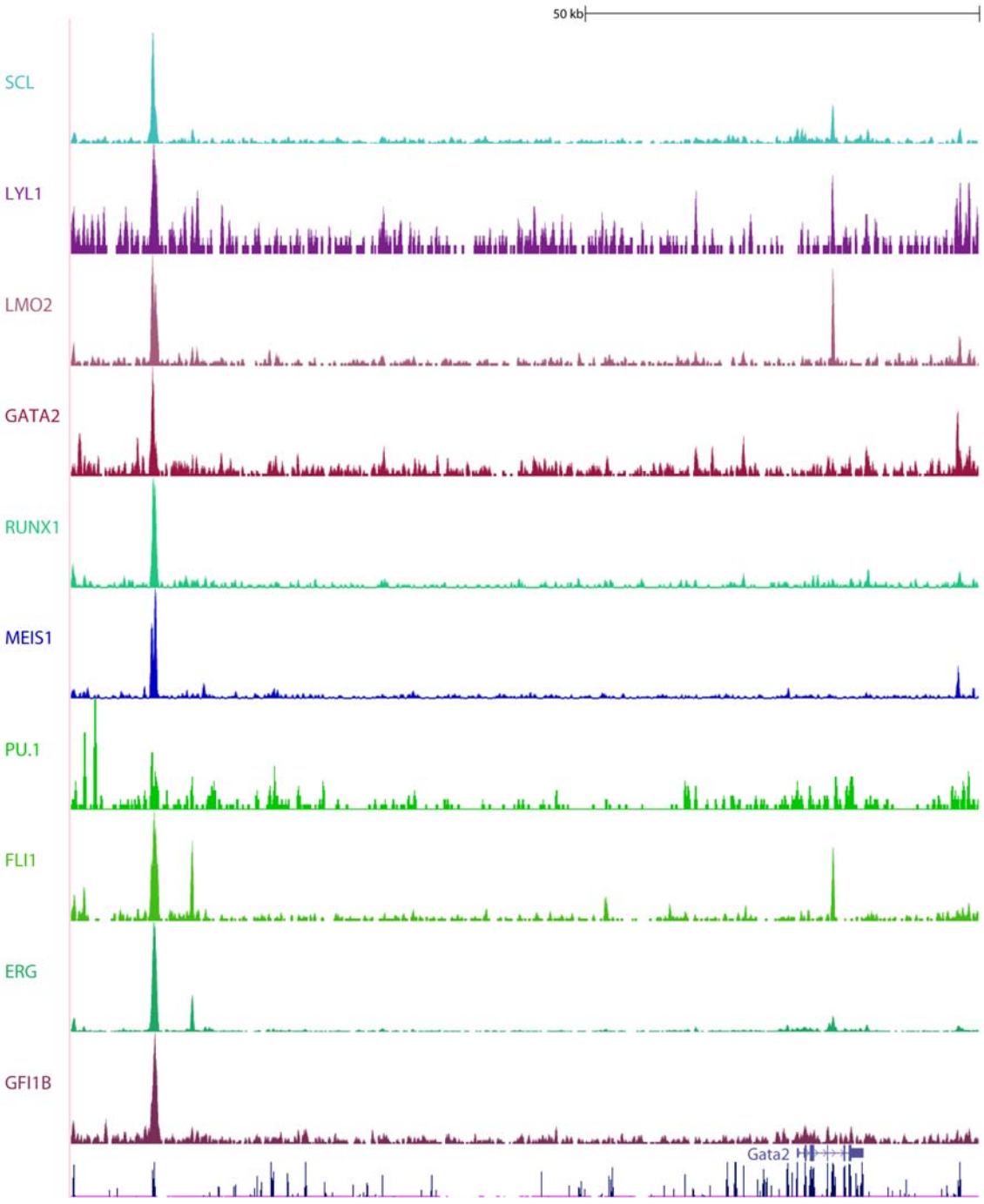
Supplementary Figure 1 continued.

A) iii)



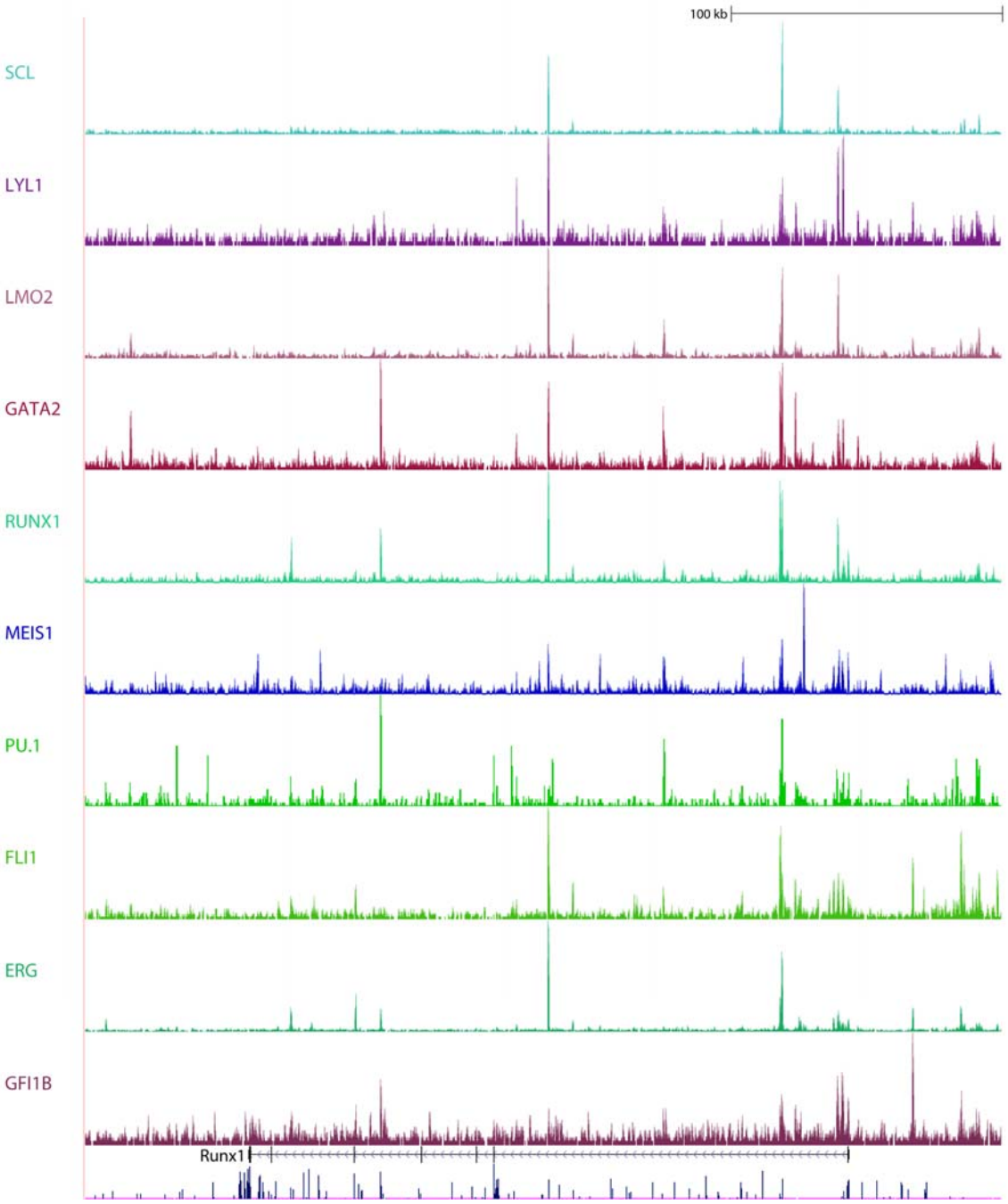
Supplementary Figure 1 continued.

A) iv)



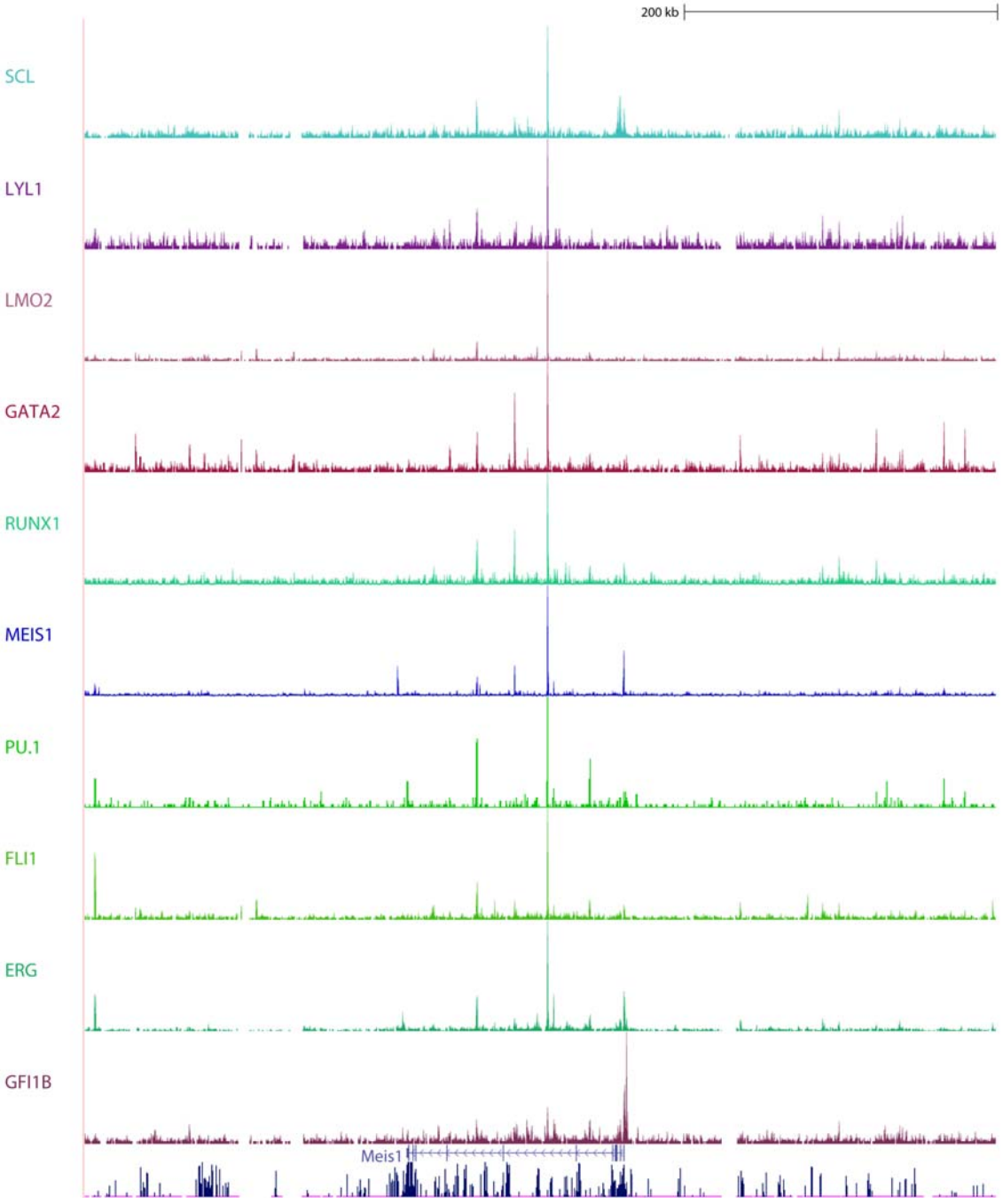
Supplementary Figure 1 continued.

A) v)



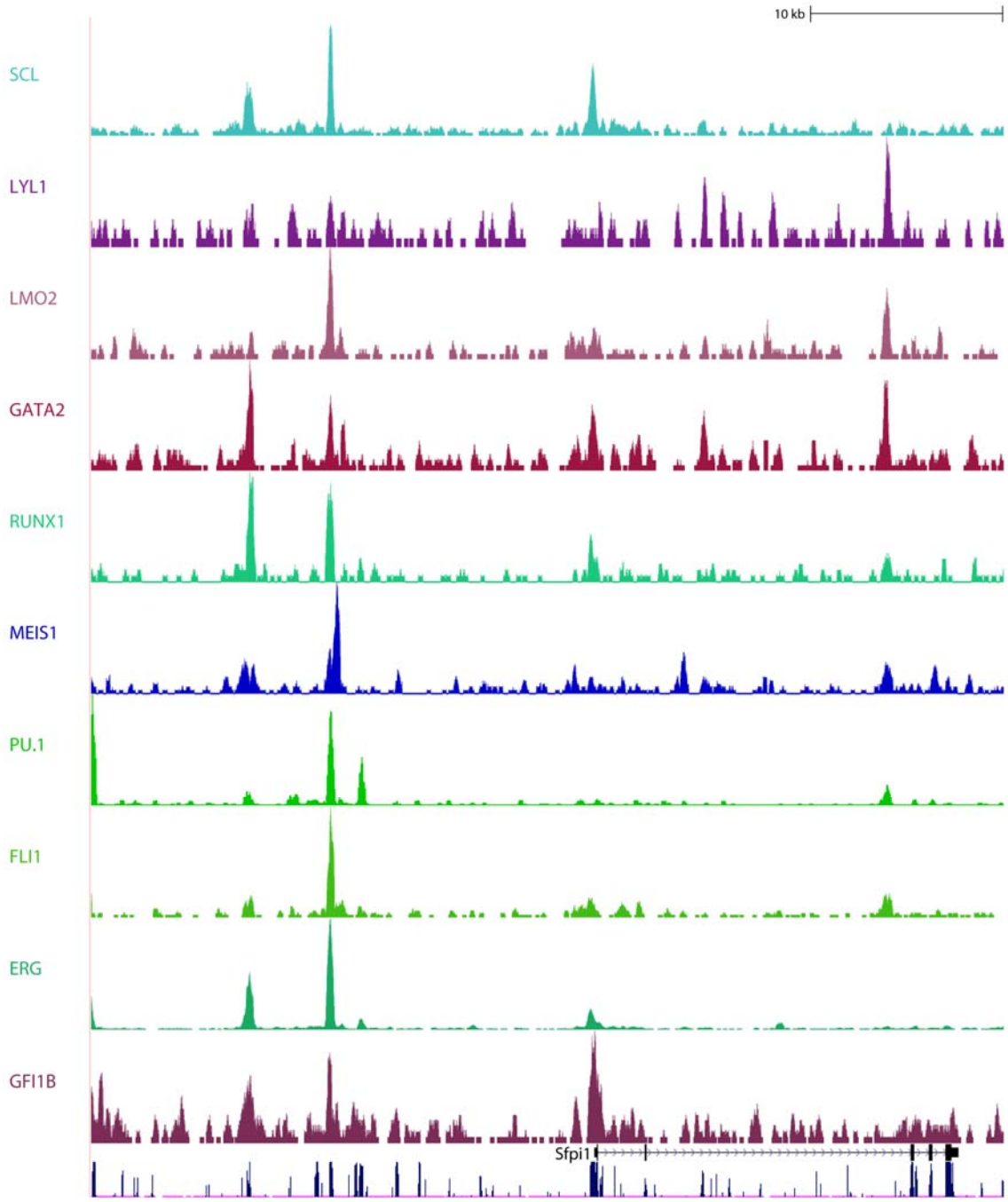
Supplementary Figure 1 continued.

A) vi)



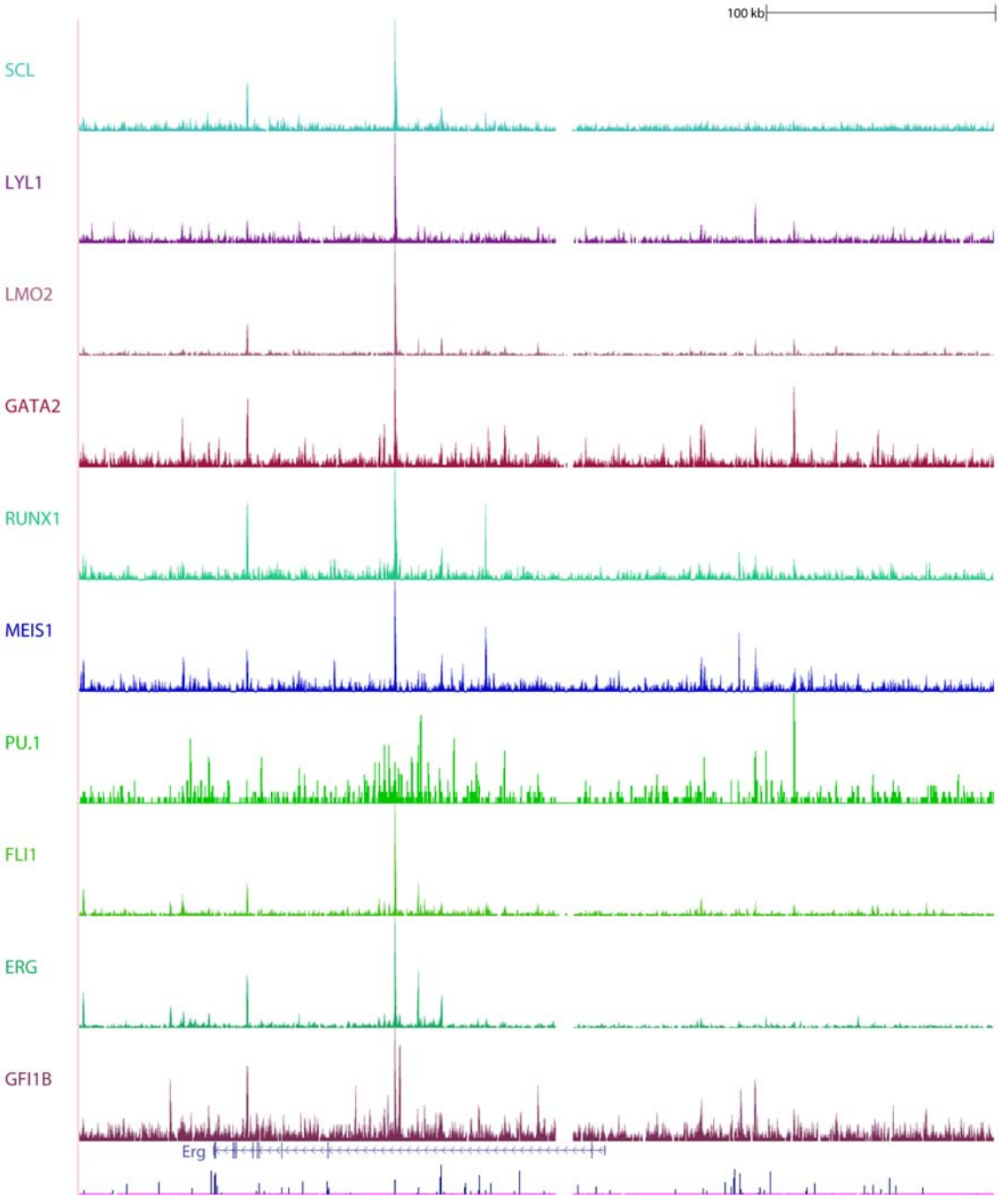
Supplementary Figure 1 continued.

A) vii)



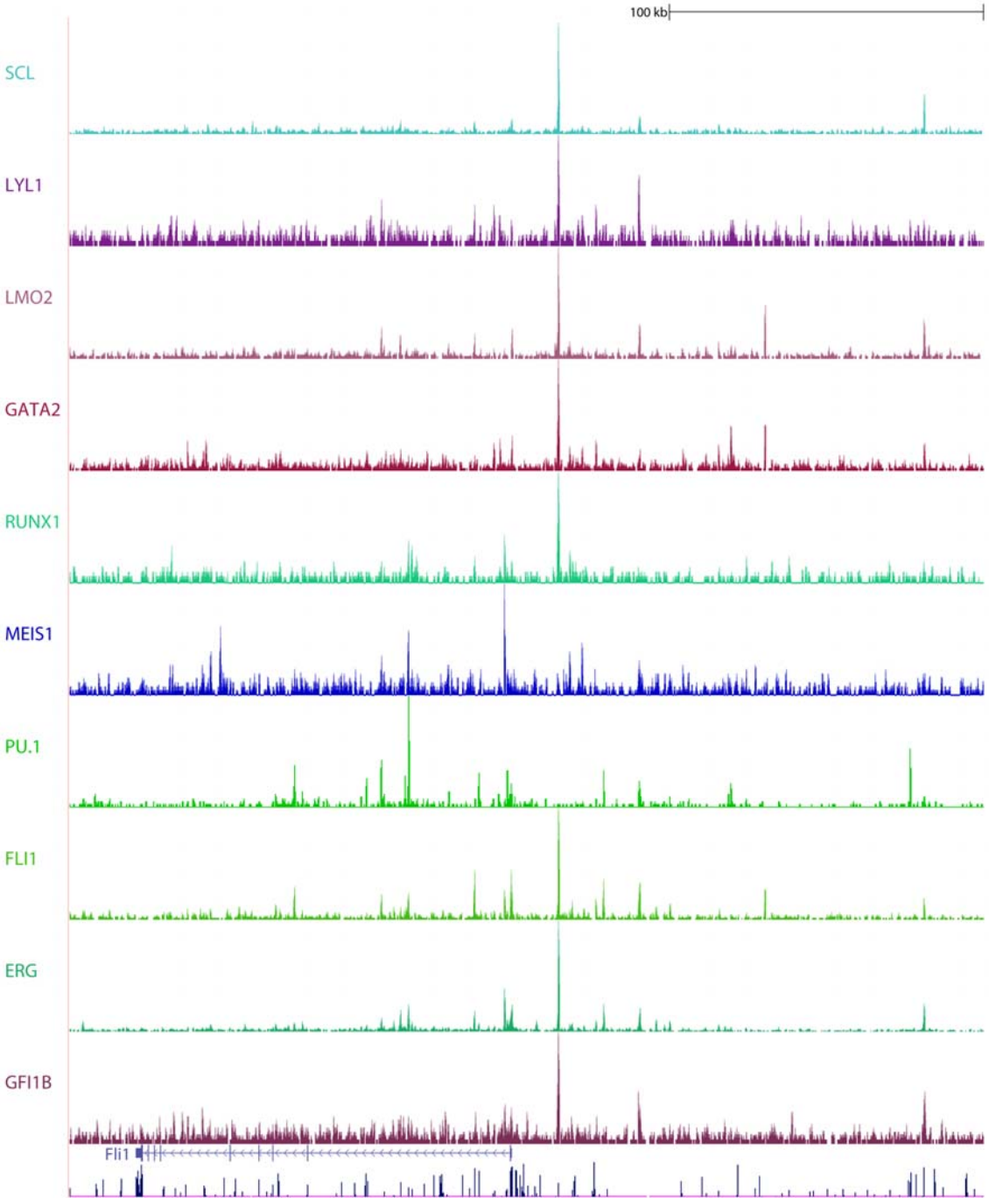
Supplementary Figure 1 continued.

A) viii)



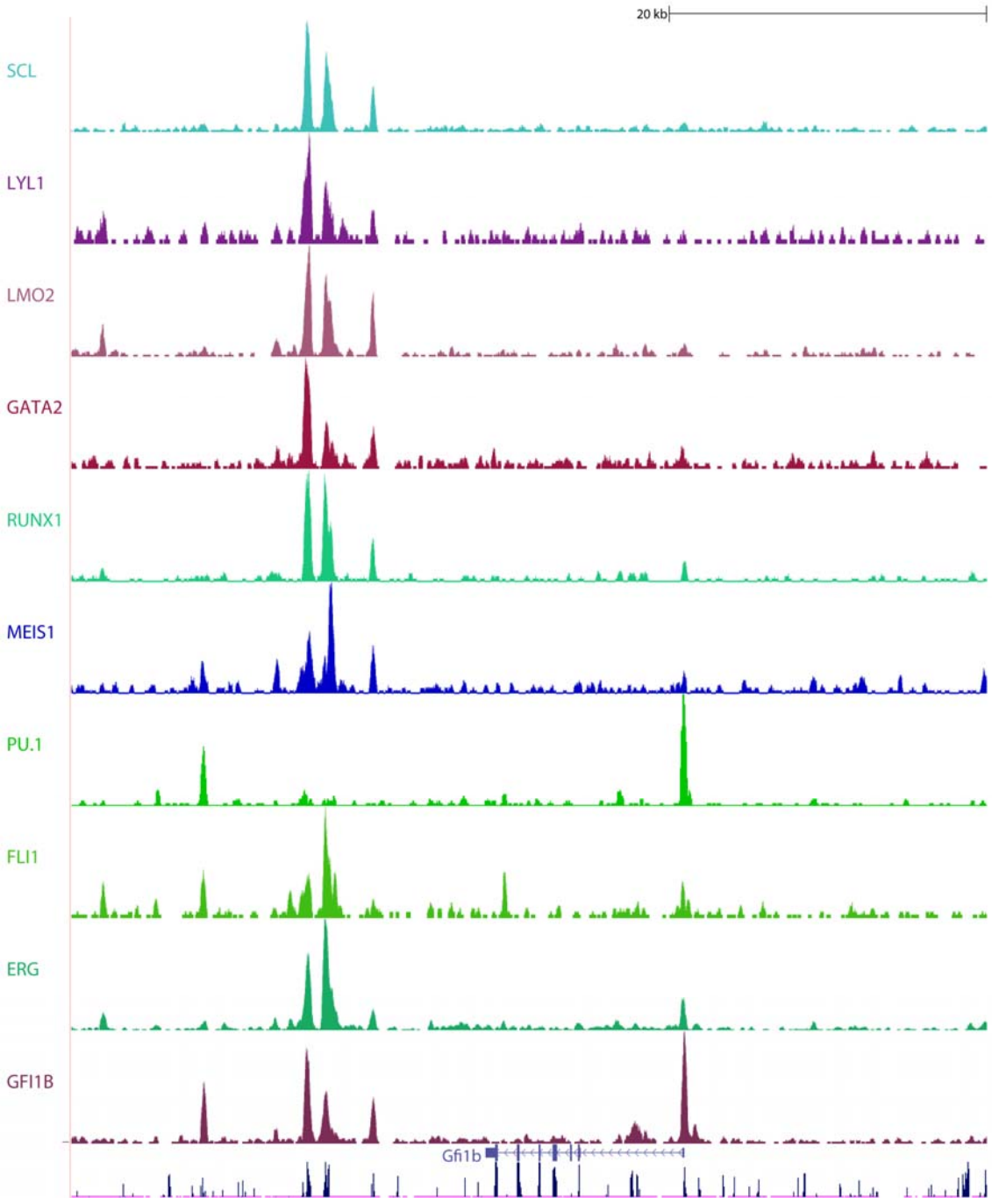
Supplementary Figure 1 continued.

A) ix)



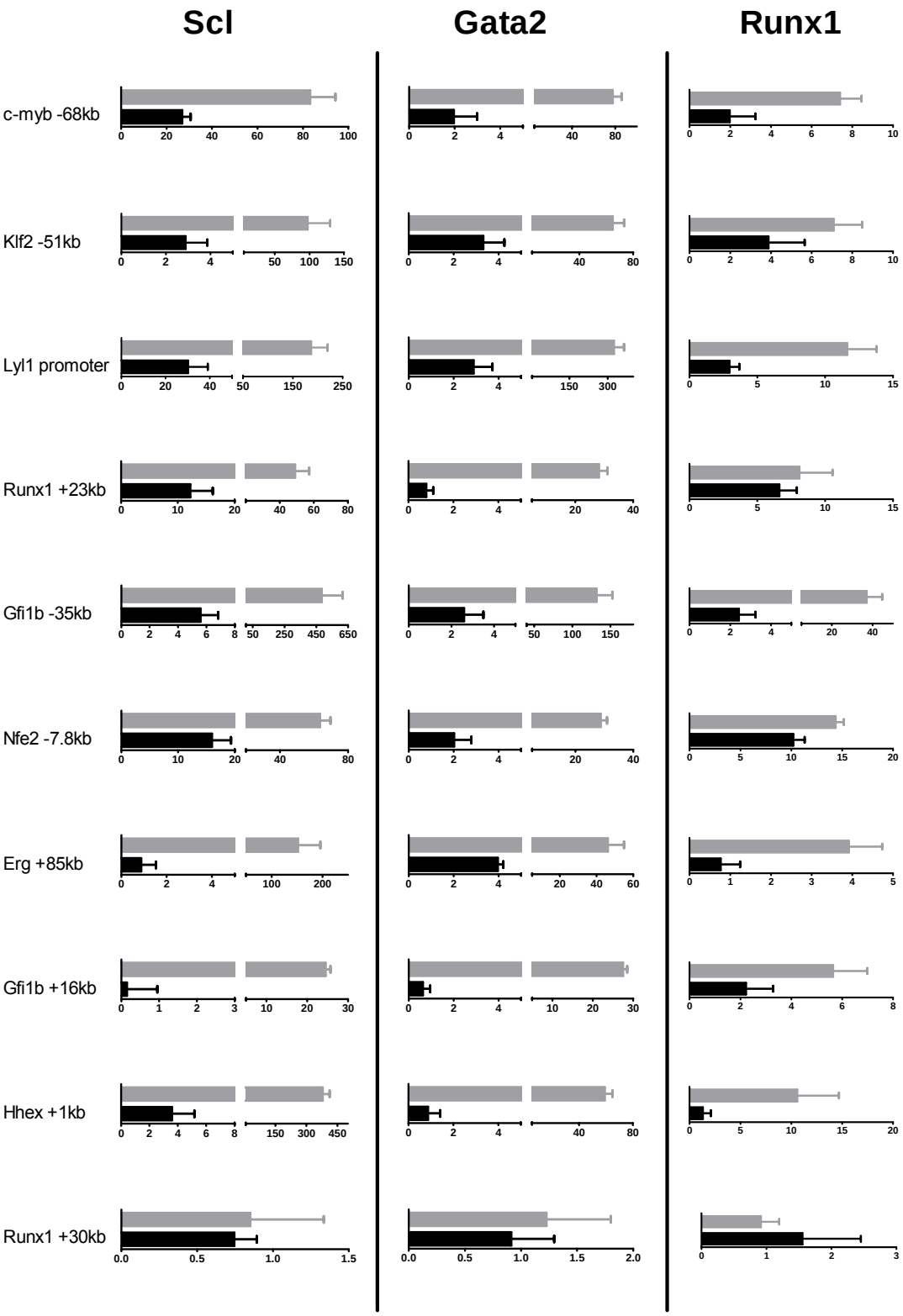
Supplementary Figure 1 continued.

A) x)



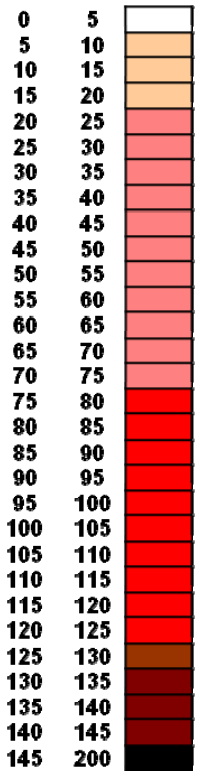
Supplementary Figure 1 continued.

B)



Supplementary Figure 1 continued.

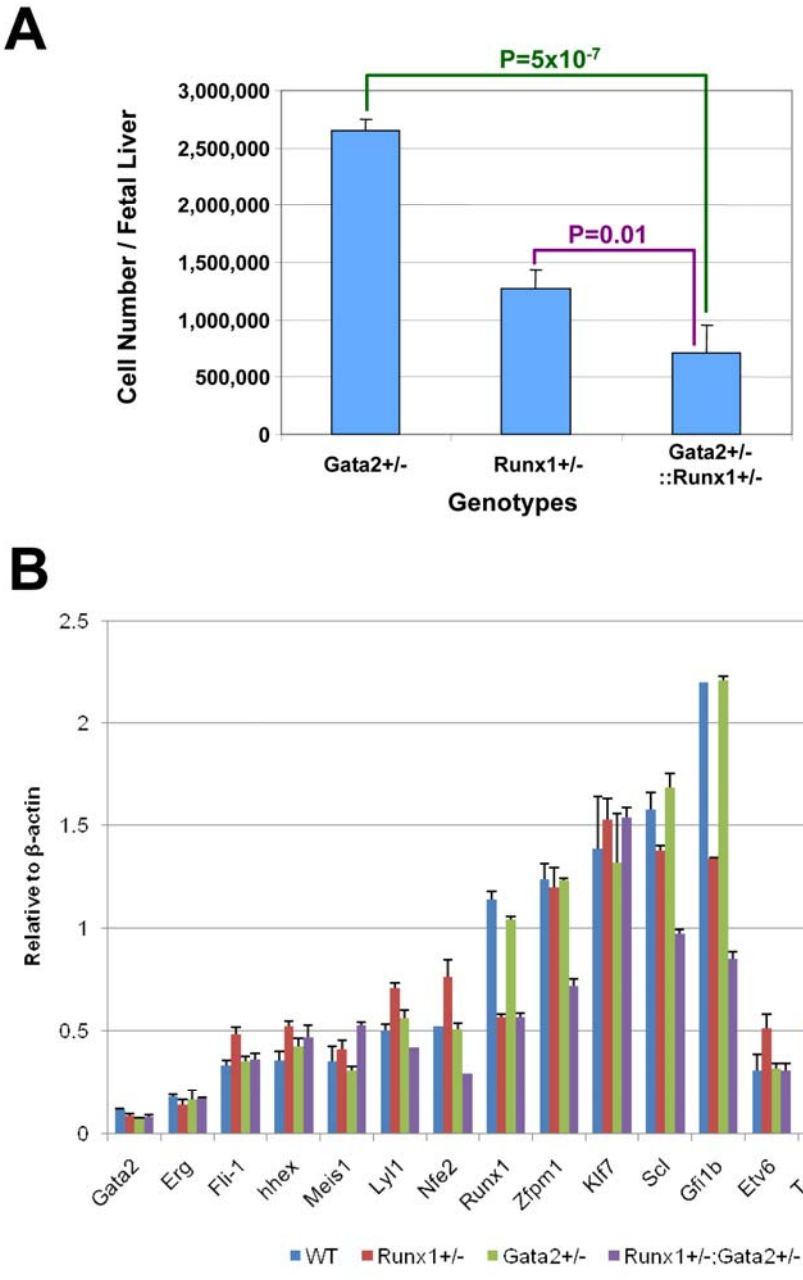
Factor	Scl/Tal 1	Lyl1	Lmo2	Gata2	Runx1	Meis1	Pu.1	Erg	Fli-1	Gfi1b
Scl overlap		113.1	153.9	73.1	121.5	53.1	30.8	91.2	86.3	80.2
Lyl1 overlap			137.7	123.6	119.4	45.7	8.7	26.8	77.0	43.8
Lmo2 overlap				111.6	145.8	56.2	6.7	53.2	117.6	48.3
Gata2 overlap					89.3	31.9	4.1	21.1	56.1	29.4
Runx1 overlap						90.7	19.9	75.1	107.6	68.8
Meis1 overlap							14.3	50.3	67.8	40.7
Pu.1 overlap								66.6	48.1	32.7
Erg overlap									135.7	74.3
Fli-1 overlap										50.1



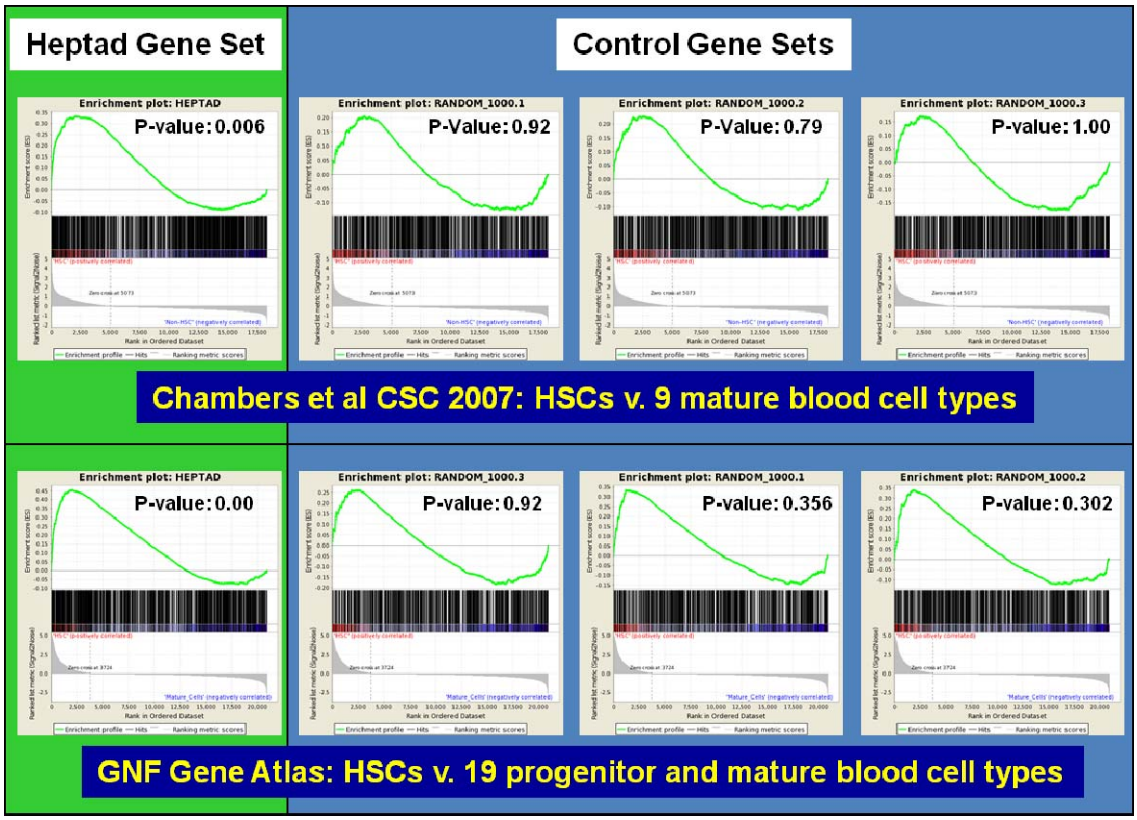
Shown at the top are the same pair-wise comparisons as in Figure 3A, but this time the Z-scores (from boot-strapping analysis) are shown rather than the number of overlapping peaks.

Shown to the left is a table with the heatmap intensities used to rank Z-scores and peak overlaps.

Supplementary Figure 2.



Supplementary Figure 3.



Supplementary Figure 4.

Supplementary Figure Legends, Figures and Supplementary Experimental Procedures

Figure S1: 10 haematopoietic gene loci exhibiting a wide variety of combinatorial binding patterns (relates to Figure 2).

A) Raw ChIP-Seq read data was transformed into a density plot for each TF and loaded into the UCSC genome browser as custom tracks above the UCSC tracks for gene structure and mammalian homology. Visual inspection of gene loci showed many different combinations of binding for the 10 different transcription factors. Shown here are the loci of the 10 TF, i) Scl, ii) Lyl1, iii) Lmo2, iv) Gata2, v) Runx1, vi) Meis1, vii) Pu,1/Sfpi1, viii) Erg, ix) Fli-1 and x) Gfi1b. B) ChIP assays were performed in HPC-7 (grey bars) and E11.5 FL (black bars). Material was validated by real-time PCR using primer pairs for regions bound by the heptad in the ChIP-Seq analysis which either have the presence or absence of a Runx motif (Runx1, Gfi1, Nfe2, Gfi1b and c-myb, Klf2, Lyl1, Erg, Hhex respectively). For lists of primers used see Supplementary Experimental Procedures.

Figure S2: Preferential transcription factor binding pairs (relates to Figure 3).

The table shows the actual z-scores for the pairwise overlaps obtained by bootstrapping. Colour coding matches Figure 3. Full colour coded legend is shown below the table.

Figure S3: The Scl/Lyl1/Gata2/Runx1/Lmo2/Fli-1/Erg heptad displays previously unrecognised protein-protein as well as genetic interactions (relates to Figure 6).

A) Cell numbers from the E12.5 fetal livers were counted before the colony assays were performed. The data shown correspond to three, four and six livers respectively, each analysed in duplicate. Fetal livers from compound heterozygous embryos are statistically significantly reduced in cell number when compared to fetal livers from either Gata2+/- or Runx1+/- embryos. B) qRT-PCR was performed on Wild type (WT), $Runx1^{+/-}$, $Gata2^{+/-}$ and $Runx1^{+/-}$::$Gata2^{+/-}$ compound heterozygotes. No dramatic reductions in expression were observed. However, the expected reductions in the expression of Runx1 and Gata2 can be seen in the $Runx1^{+/-}$ and $Gata2^{+/-}$ genotypes and some genes show expression changes specifically in compound heterozygotes.

Figure S4: GSEA analysis using control gene sets to validate HSC-specific expression of genes associated with heptad-bound peaks (relates to Figure 7).

Shown to the left are the results of GSEA analysis with the gene set defined through peaks bound by the heptad of TFs (927 genes) also shown in Figure 7. Shown to the right are the GSEA results when using three randomly generated gene sets of equivalent size. These control gene sets were chosen from a list of all genes bound by at least 3 of the 10 TFs (so as to exclude those genes that are in heterochromatin and therefore not accessible for any TF binding). The graphical outputs show enrichment (green curve) of the heptad gene set along ranked lists of all genes ordered based on their HSC-specificity of expression in the two datasets. None of the three control gene sets show any significant enrichment.

Supplemental Experimental Procedures

Chromatin imunoprecipitation assay

10^8 HPC7 cells were incubated in (0.4% - Ack9, Scl, Gata2, Meis1 and Pu.1 or 1% - Runx1, Erg, Lmo2, Fli-1, Lyl1 and Gfi1b) formaldehyde for 10 minutes at room temperature. To quench the reaction 0.125M glycine was added and the cells were incubated at room temperature for 5 minutes. Cells were immediately harvested and washed in cold 1x PBS. Cells were harvested and resuspended in 1.5 x pellet volume of cell lysis buffer (10mM Tris pH 8.0, 10mM NaCl and 0.2% NP40) containing protease inhibitors (leupeptin, NaBu and PMSF), the cells were homogenised and incubated on ice for 10 minutes. The nuclei were harvested at 600 x g for 5 minutes at 4^{o}C and resuspended in 1 ml of nuclei lysis buffer (50mM Tris pH 8.0, 10mM EDTA, 1% SDS) containing protease inhibitors (leupeptin, NaBu and PMSF), the cells were homogensised and incubated on ice for 10 minutes. An equal volume of IP dilution buffer (20mM Tris pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.01% SDS) containing protease inhibitors (leupeptin, NaBu and PMSF) was added. Cells were sonicated on ice-water (Biorupter, Diagenode) for 5 cycles (30s on, 30s off). The chromatin solution was centrifuged for 10 minutes at 3220 x g. After transferring the chromatin solution to a clean falcon tube, a further 3 mls of IP buffer was added. The chromatin solution was pre-cleared by the addition of 50µl of rabbit IgG (2 µg/µl) and incubated at 4^{o}C for 1 hour, then 200µl of Protein G sepharose beads (1:1 slurry in IP dilution buffer) were added to the chromatin solution and further incubated at 4^{o}C for 2 hours. The beads/IgG were collected by centrifugation at 1791 x g for 2 minutes. The chromatin was transferred to 1.5 ml tubes, an input sample was removed and antibodies/IgG were added then incubated overnight at 4^{o}C with rotation. 60µl of protein G agarose beads (1:1 slurry in IP dilution buffer) were added and incubated with the samples for 2 hours. The beads were harvested at 5400 x g for 2 minutes and washed twice with low salt buffer (20mM Tris pH 8.0, 2mM EDTA, 50mM NaCl, 1% Triton X-100, 0.1% SDS), then once with LiCl buffer (10mM Tris pH 8.0, 1mM EDTA, 0.25M LiCl, 1% NP40, 1% Sodium deoxycholate monohydrate) and twice with 1x TE pH 8.0. The complexes were eluted twice from the beads by adding 150µl elution buffer (100mM $NaHCO_3$, 1% SDS). To reverse the cross-linking 0.3M NaCl was added to all the IP samples and input, RNase was added and the samples were incubated at 65^{o}C overnight. The samples were then treated with Proteinase K for 2 hours at 45^{o}C. DNA was purified using Qiagen PCR clean up columns.

Primer name	Sequence Forward Primer 5' – 3'	Sequence Reverse Primer 5' – 3'
ChIP		
c-myb -68kb	GGGCATCCTGATTGTGCTAA	CCAAGTGACGGTGACAGAGA
Klf2 -51kb	TCCTTGGCCTGATATTTTGG	GCTTCTGAGGGATGTGGAAA
Lyl1 promoter	AGCCAGACCCTTATCTGCAC	CTGGCTTCCTCCCTCTTACC
Runx1 +23kb	AAGCTGCCCACGTTATCAGT	CAGATGGAGGCATCCTGTTT
Gfi1 -35kb	CCACAAACAGAACAGCTGGA	CCACATGACCTCATGAATGC
Nfe2 -7.8kb	TCACACCAGTAGGCAATCCA	GTGGCTAGAGGTGGAACCAG
Erg +85kb	GCTGAACACTCGTTACAAGAT	GTCCCAGAGTGACCCACC
Gfi1b +16kb	GCTATTTCTGCCAAGGGTGA	GGGGTCTGAGGACCATGATA
Hhex +1	CGACGTCTGATAGCCAGGAT	AGGCAGAGAGGAAGCAACTG
Runx1 +30kb	GCCACAGAGAGGATGTGGAC	AGCACCTGCCAAGAGACATC
Expression analysis		
β-actin	TCCTGGCCTCACTGTCCAC	GTCCGCCTAGAAGCACTTGC
Gata2	GCACCTGTTGTGCAAATTGT	GCCCCTTTCTTGCTCTTCTT
Erg	GGCAGCTACATGGAGGAGAA	TATTCTTTCACCGCCCACTC
Fli-1	CCAACGAACGGAGAGTCATT	ATTCCTTGCCATCCATGTTC
Hhex	TCAGAATCGCCGAGCTAAAT	CTGTCCAACGCATCCTTTTT
Meis1	TAACTGACCAGCCCTCTTGG	GTTGTCCAAGCCATCACCTT
Lyl1	CAGGACCCTTCAGCATCTTC	ACGGCTGTTGGTGAACACTC
Nfe2	TGAGCTGCAGGGTCTAAATG	ACAGGGGCAATATGTTGGAG
Runx1	CTCCGTGCTACCCACTCACT	ATGACGGTGACCAGAGTGC
Zfpm1	CCCTGTGCAGGAACCAGTAG	TACCAGATCCCGCAGTCTTT
Klf7	CACTTAAAGGCCCACCAGAG	CTTAAAGGGCTTTGCACCTG
Scl	CATGTTCACCAACAACAACCG	GGTGTGAGGACCATCAGAAATCTC
Gfi1b	AGCACAGAGTCTCCCTTGGA	CATGCACTTCTAGCCCATGA
Etv6	TCCCTTTCGCTGTGAGACAT	GGGCGTGTATGAAATTCGTT
Tet2	TGTTTGGGTCTGAAGGAAGG	TGAGGGTGACCACCACTGTA
Itga2b	ACCTCAACCGAGACGGCTAT	TGGACTCAGCCCTTCACTCT
Itgb3	ACAGAGCGTGTCCCGTAATC	CAATATGGGTCTTGGCATCC
Stat1	TGCCGAGAACATACCAGAGA	AGTTCGCTTAGGGTCGTCAA
Stat3	TGACCCTTAGGGAGCAGAGA	ATCTTGAGGCCTTGGTGGTA