www.elsevier.com/locate/developmentalbiology

A somatic mutation of GFI1B identified in leukemia alters cell fate via A SPI1 (PU.1) centered genetic regulatory network

Eduardo Anguita, Rajeev Gupta, Victor Olariu, Peter J. Valk, Carsten Peterson, Ruud Delwel, Tariq Enver

PII: S0012-1606(16)30072-0
DOI: http://dx.doi.org/10.1016/j.ydbio.2016.02.002
Reference: YDBIO7013

To appear in: *Developmental Biology*

Received date: 20 April 2015
Revised date: 1 February 2016
Accepted date: 2 February 2016

Cite this article as: Eduardo Anguita, Rajeev Gupta, Victor Olariu, Peter J. Valk, Carsten Peterson, Ruud Delwel and Tariq Enver, A somatic mutation of GFI1E identified in leukemia alters cell fate via A SPI1 (PU.1) centered genetic regulatory network, *Developmental Biology* http://dx.doi.org/10.1016/j.ydbio.2016.02.002

A somatic mutation of GFI1B identified in leukemia alters cell fate via a SPI1 (PU.1) centered genetic regulatory network

Running title: Dysregulation of a GFI1B-SPI1 subnetwork in AML

Eduardo Anguita[a*], Rajeev Gupta[b], Victor Olariu[c], Peter J. Valk[d], Carsten Peterson[c], Ruud Delwel[d], Tariq Enver[b*]

[a]Hematology Department, IdISSC, Hospital Clinico San Carlos, 28040, Madrid, Spain.

[b]UCL Cancer Institute, Paul O'Gorman Building 72 Huntley St., London WC1E6BT.

[c]Computational Biology and Biological Physics Division, Lund University, Lund, Sweden.

[d]Department of Hematology Erasmus University Medical Center, Rotterdam, Netherlands.

e-mail address of authors:

eduardo.anguita@salud.madrid.org

rajeev.gupta@ucl.ac.uk

victor.olariu@thep.lu.se

p.valk@erasmusmc.nl

carsten@thep.lu.se

h.delwel@erasmusmc.nl

t.enver@ucl.ac.uk

Author to whom correspondence and proofs are to be sent to communicate with the Editorial and Production offices:

Tariq Enver. e-mail: t.enver@cancer.ucl.ac.uk. Mailing address: UCL Cancer Institute, Paul O'Gorman Building 72 Huntley St., London WC1E6BT. Phone: 00 44 207 679 6545. Fax: 00 44 207 679 6817

Also Corresponding author:

Eduardo Anguita. e-mail: eanguita.hcsc@salud.madrid.org. Mailing address: Hematology Department, Hospital Clinico San Carlos, 28040, Madrid, Spain. Phone: 00 34 91330 3321. Fax: 00 34 91330 3322

ABSTRACT

We identify a mutation (D262N) in the erythroid-affliated transcriptional repressor GFI1B, in an acute myeloid leukemia (AML) patient with antecedent myelodysplastic syndrome (MDS). The GFI1B-D262N mutant functionally antagonizes the transcriptional activity of wild-type *GFI1B*. GFI1B-D262N promoted myelomonocytic versus erythroid output from primary human hematopoietic precursors and enhanced cell survival of both normal and MDS derived precursors. Re-analysis of AML transcriptome data identifies a distinct group of patients in whom expression of wild-type *GFI1B* and *SPI1* (*PU.1*) have an inverse pattern. In delineating this GFI1B-SPI1 relationship we show that (i) *SPI1* is a direct target of GFI1B, (ii) expression of GFI1B-D262N produces elevated expression of *SPI1*, and (iii) SPI1-knockdown restores balanced lineage output from GFI1B-D262N-expressing precursors. These results table the SPI1-GFI1B transcriptional network as an important regulatory axis in AML as well as in the development of erythroid versus myelomonocytic cell fate.

Keywords: GFI1B; SPI1; PU.1; transcriptional networks; acute myeloid leukemia; myelodysplastic syndrome.

INTRODUCTION

Transcription factor mutation is a major cause of cell fate dysregulation in leukemia. Indeed many key regulators of hematopoiesis were identified through analysis of leukemia patients.[1] Myelodysplastic syndromes (MDS) by contrast, are acquired clonal diseases regarded broadly as 'pre-leukemic' conditions that can be conservatively managed. However, approximately 10% of all MDS cases evolve into acute myeloid leukemia (AML) within the course of two years.[2] Key challenges in the field are to understand how individual regulators control cell fate at the level of lineage choice, self-renewal, cell cycle or survival, how different regulators interact in transcriptional networks and how circuit behavior is altered in leukemia.[3,4]

Transcription factor GFI1B was identified by sequence homology with oncogene *GFI1* (Growth Factor Independence 1).[5] Both GFI1 and GFI1B have six C-terminal C2H2 zinc-fingers and an N-terminal SNAG transcriptional repression domain.[5–7]

In mice, *Gfi1* is essential for neutrophil differentiation[8,9] and in humans severe congenital neutropenia is associated with mutations in *GFI1*.[10] Gfi1 is also required for B and T lymphopoiesis.[11,12] In contrast knockout mice have demonstrated that *Gfi1b* is required for development of both erythroid and megakaryocytic lineages,[13] but the mechanism is not understood.[14] Loss of *Gfi1b* in adult mice increases absolute numbers of hematopoietic stem cells (HSC), these are less quiescent than wild type HSCs.[15] In partial agreement with the mouse models, human congenital mutations of GFI1B produce hemorrhagic conditions with abnormal platelet function.[16,17]

Against this backdrop, we have studied a previously unreported acquired point

mutation (D262N) in the transcription factor GFI1B in a patient with AML arising from an antecedent MDS and explored its biological properties.

This mutation leads to an amino acid change in the fourth zinc finger domain and results in a mutant protein that when expressed in normal human Common Myeloid Progenitors (CMPs) skews lineage output towards myelopoiesis. Expression of GFI1B (D262N) in human bone marrow (BM) cells from healthy individuals and other MDS patients promotes their survival. We show that these effects are mediated by upregulation of *SPI1* (*PU.1*).

SPI1 is a critical factor in blood cell differentiation.[18,19] In order to form macrophages, granulocytes and B cells SPI1 levels must be exquisitely regulated.[20,21] Our data, point to a SPI1/GFI1B axis regulating erythroid/myeloid commitment, which when dysregulated may play a role in leukemogenesis.

MATERIAL AND METHODS

Human Material

Human material and clinical information was obtained with informed consent and approval of institutional ethics committee.

Point mutation analysis

PCR fragments were analyzed by DHPLC as before (Supplementary Table 1).[22]

Target Region Capture Sequencing

Genomic DNA was randomly fragmented by Covaris (Woburn). DNA was amplified by ligation-mediated PCR (LM-PCR), purified, and hybridized to the NimbleGen (Roche) human custom array. High-throughput sequencing was performed with Hiseq2000 (Illumina).

Cloning

GFI1B cDNA was amplified with primers indicated in Supplementary Table 2, cloned in pGEM-T easy vector and sequenced. Point mutation was introduced by PCR. Wild-type or mutated coding cDNA was amplified from this plasmid with oligonucleotides shown in Supplementary Table 2 and cloned into pHRSINCSGW lentivirus vector, expressing green fluorescent protein (GFP).[23]

The *SPI1* hairpin has been published,[24] it was cloned into the SLX vector,[23] with or without wild-type or mutant *GFI1B*.

Lentivirus

Were prepared as described.[25]

Enrichment of hematopoietic stem and progenitors

Enrichment of CD34+ cells obtained from G-CSF mobilized adult human peripheral blood from patients undergoing autograft procedures for non-hematological disorders was by magnetic selection (Miltenyi Biotec).

Cord blood progenitors were isolated as described.[26]

Mobilised adult CD34+ cells and cord blood derived hematopoietic progenitors were maintained in MyeloCult® H5100 (StemCell Technologies) supplemented with 100ng/mL each human SCF, TPO, Flt3L (Peprotech).

Isolation of lentivirally transduced cells

Three days after lentiviral infection transduced cells were labeled with APC anti human CD34 (Cat 345802, BD Pharmingen) and CD34+/GFP+ cells were isolated using a Mo-Flo cell sorter (Dako).

MDS cell culture

Bone marrow mononuclear cells were enriched for CD34+ cells by magnetic separation and cultured in MyeloCult® H5100 with 10% (v/v) HS-5 conditioned medium, human SCF and IL-3 (both 50 ng/ml, Peprotech).

Methylcellulose Colony-Forming Cell (CFC) assays

CFC assays were in MethoCult H4436 (StemCell Technologies), 15 days. We performed 2 biological replicates for each experiment and plated 1000 cells in each experiment.

Liquid differentiation assays

These were performed in Myelocult® H5100 supplemented with human cytokines: SCF, FLT3L (both 50 ng/ml), IL-3, IL-6, GM-CSF, M-CSF (all 10 ng/ml): (Peprotech); G-CSF (0.1 MUI) (Ratiopharm); EPO (Amgen) (100 ng/ml)

Chromatin immunoprecipitation (ChIP) assay

ChIPs and immunoprecipitated DNA analysis were performed as previously described.[27–29] *SPI1* promoter primers and probes are shown in Supplementary Table 3.

AML patients gene expression profiling and data analysis.

This has been published (www.ncbi.nlm.nih.gov/geo, GSE1159).[30]

Computational model

We applied deterministic rate equations.[3,31] Additional information can be found in the Supplementary Methods and Supplementary Figure 1.

RESULTS

A GFI1B zinc finger mutation arising on transformation of MDS into AML

Using DHPLC , we analyzed the *GFI1B* coding regions in DNA from the diagnostic bone marrow of 69 previously described patients with AML.[22] In this analysis, we identified a case with an abnormal pattern (Figure 1A). Immunophenotyping data from the diagnostic bone marrow aspirate (which showed 48% blasts) are provided in Supplementary Table 4. Bacterial cloning allowed the identification of a novel G>A mutation in exon six (position +3198 from the start codon, chromosome nine coordinate 135865264 GRCh37/hg19 assembly, NM_004188.5:c.784G>A), Figure 1B, in 12% of clones. This mutation is predicted to lead to an Asp (D) to Asn (N) transversion at aa262 (p.Asp262Asn), in the phylogenetically conserved fourth zinc finger motif of both GFI1B and GFI1 (Figure 1C). To accurately quantify the proportion of mutated DNA in the marrow, we performed a target region capture and next generation sequencing (NGS) of coordinates 135851000-135883000 chromosome nine GRCh37/hg19 assembly. This identifies the mutation in 13% of the total analysed DNA, corresponding to 26% of marrow cellularity. It is therefore, is a major clone in the leukemia (Supplementary Table 5). In contrast, we were unable to detect the mutation in samples taken some 18 months previously, when the patient initially presented with MDS, refractory anemia with excess

blasts (RAEB), (Figure 1D-F). This suggests that in this case, the acquisition of D262N is associated with disease progression.

GFI1B D262N impairs the function of the wild-type GFI1B.

To evaluate the impact of the GFI1B D262N (GFI1Bm) mutation we generated lentiviral vectors containing *GFI1B* and *GFI1Bm* under the control of either the spleen focus forming virus (CSI) or the Wiskott Aldrich Syndrome Protein promoters (CWI) allowing us to study the effects of relatively high (CSI) or low (CWI) levels of transgene over-expression (Figure 2A). Transduced cells were marked by an emerald-GFP reporter. GFP positive primary cells and cell lines were sorted and evaluated for expression of GFI1B targets.

GFI1B down-regulates its own expression.[28,32] Consistent with this, we observed reduced levels of endogenous *GFI1B* mRNA after enforced expression of a *GFI1B* transgene in the leukemic cell line K562. This effect was proportional to transgene levels (Figure 2B) and similar results were obtained in primary CD34+ BM cells (Figure 2C). Enforced expression of *GFI1Bm*, in contrast, failed to repress endogenous *GFI1B*; indeed transcript levels increased (Figure 2B, C). This effect was magnified by the increase in *GFI1B* expression (Figure 2B). These data are consistent with the notion that expression of *GFI1B*m interferes with the repressor functions of wild-type GFI1B.

Functional antagonism of GFI1Bm on GFI1B was also observed in the derepression of other known GFI1B targets (Figure 2D).[33–35]

To explore mechanism of the differential effect of GFI1B and GFI1Bm on gene expression, we studied epigenetic marks within the endogenous *GFI1B* (Figure 2E) locus by ChIP in the hematopoietic cell line K562. Histone H3K4me3, associated with active transcription,[36] was significantly reduced at, and immediately downstream of, the endogenous promoter in cells transduced with *GFI1B*, but not *GFI1Bm* (Figure 2F). Consistent with previous observations,[28,37] H3K4me2 did not change (Figure 2G).

Taken together these results indicate that GFI1B D262N lacks the transcriptional repressor functions associated with the wild-type protein and that this alters histone methylation.

GFI1B D262N impacts lineage output from human primary hematopoietic cells

We tested the functional impact of *GFI1Bm* in hematopoiesis, by transducing primitive CD34+ cells derived from adult bone marrow with CWI or CSI based lentiviral vectors (Figure 3A) and assessing cell fate *in vitro*.

When CD34+ cells expressing *GFI1Bm* were maintained in liquid culture conditions that support multi-lineage differentiation, we observed a decrease in erythroid and an increase in myeloid output compared to cells transduced with wild-type *GFI1B*. This was assessed by expression of erythroid markers CD36 and CD235a and the myeloid marker CD13, Figure 3B.

Human CD34+ cells comprise several primitive hematopoietic progenitor types. In order to further define the cells in which *GFI1Bm* has functional impact, we transduced cord-blood derived CMP (defined as CD34+38+10-7-135+45RA-,

Supplementary Figure 2). Consistent with our initial observations, *GFI1Bm* expression skewed lineage output towards myeloid in this compartment (Figure 3C). This shows that GFI1B D262N has a biological impact in highly purified bipotential progenitors. This skewing of lineage output was also seen in colony assays in methylcellulose with the same cytokines under conditions where total colony numbers were unaltered (Figure 3D, Supplementary Figure 3).

We next explored the impact of *GFI1Bm* on progenitors under conditions that promote erythroid output. These results are presented in Figure 3E and show that expression of *GFI1Bm* also produces a reduction in erythroid output. Transduction with the CWI construct driving lower expression levels revealed that this phenomenon was partly *GFI1Bm* dose dependent. The persistence of the erythroid differentiation bias under this experimental setting, which is closer to the physiological conditions, underlines the *in vivo* relevance of this finding.

GFI1B exhibits pro-apoptotic activity in human CD34+ cells.[38] Consistent with GFI1Bm's functional antagonism of wild-type GFI1B activity, *GFI1Bm* enhanced the resistance of normal bone marrow CD34+ cells to apoptosis when incubated in the absence of cytokines (Figure 3F, G). Similarly, GFI1Bm also enhanced the resistance of CD34+ cells from MDS patients to apoptosis (Supplementary Figure 4 and Supplementary Table 6). Interestingly, colonies derived from cells transduced with *GFI1B* appeared smaller than those from *GFI1Bm* transduced cells (data not shown).

These data show that GFI1B D262N promotes the survival of normal human BM CD34+ cells and skews lineage output towards myeloid under conditions that support multilineage differentiation, implying a link between this mutation

and the malignant phenotype.

Cross regulatory relationships of GFI1B and SPI1 in normal and AML cells

Although the *GFI1B* mutation we report is thus far restricted to a single case, we reasoned that the pathways that it dysregulates could have a wider role in AML. In support of this, analysis of transcriptional profiles from 285 previously characterized AMLs[30] revealed a striking inverse relationship in the patterns of *GFI1B* and *SPI1* expression (Figure 4A). This suggests that *GFI1B* antagonizes *SPI1* expression and that, simplistically at least, *GFI1Bm* could increase *SPI1* levels. However, the presence of wild type and mutant *GFI1B* alleles in the same cell complicates the situation (Figure 4B, C) and, therefore, we initially analyzed the behavior of this circuit *in silico* using a deterministic rate equation approach,[3,31] (Supplementary Methods, Supplementary Figure 1 and Supplementary Table 7). Computational analysis indicates that even heterozygous expression of *GFI1Bm* enhances *SPI1* expression (Figure 4D). Manipulation of *GFI1B* expression in K562 cells provides concrete experimental evidence for this prediction (Figure 4E). Studies in primary human cells showed a similar trend (Figure 4F), small divergences between the two most likely reflect the inherent heterogeneity in primary cells.

Gfi1b binding to the *Spi1* promoter in mouse,[39] suggests that the effects on *SPI1* expression that we observed may in part stem from a direct regulation of *SPI1* locus by GFI1B. We found a conserved GFI1B binding sequence at the same location in the human promoter (Figure 4G), and confirmed by ChIP that

GFI1B binds to this site (Figure 4H).

We next explored the AML data set to ask what other molecular features correlated with high levels of *SPI1* expression and observed a strong correlation with CD4 (Figure 5A). Interestingly, the leukemic blasts in the index case, were CD4+ (Supplementary Table 4). We have previously shown *CD4* to be direct target of Spi1 in multipotent mouse progenitor cells as well as in their myelomonocytic-differentiated progeny,[40] and consistent with our hypothesis that *GFI1Bm* in part contributes to leukemogenesis though manipulation of the SPI1 programme we demonstrated that forced expression of *GFI1Bm* increases cell-surface CD4 expression in normal human CD34+ cells (Figure 5B, C).

The effects of *GFI1Bm* may be in part due to the actions of transcription factors other than SPI1, and this could explain small deviations from the predictions of our model. To explore this possibility, we re-analysed the AML data, and noted both a positive relationship between the expression of *GFI1B and* the megakaryocytic/erythroid affiliated transcription factor *MLLT3*[41] (Figure 5D), and an inverse relationship between *SPI1* and *MLLT3*. Consistent with this, in both a hematopoietic cell line and in primary cells, *GFI1B* increased and *GFI1Bm* reduced *MLLT3* expression (Figure 5E and F).

Taken together, these results suggest that high level *SPI1* expression is associated with myelomonocytic affiliated AML samples. Expression of *GFI1Bm* dramatically enhances *SPI1* expression and also results in up-regulation of the SPI1 target gene *CD4*. Interference with endogenous *GFI1B* activity is a

plausible explanation for these findings.

The effects of GFI1B D262N are routed through *SPI1*

We tested whether the effects of *GFI1Bm* were mediated by increased *SPI1* expression using a previously validated shRNA sequence.[24] With this reagent, we achieved knockdown of *SPI1* RNA to approximately 15% of its starting level in adult human primary cells (CD13+/235-) (Figure 6A). We constructed a lentiviral vector containing this short-hairpin as well as either *GFI1B* or *GFI1Bm* under the control of an *SFFV* promoter. This strategy ensured co-expression within the same cells of the *SPI1* shRNA and either the wild type or mutant *GFI1B*. These vectors were validated in K562 cells where *SPI1* and *GFI1B*/*GFI1Bm* expression were simultaneously assessed (Figure 6B). The relative levels of *SPI1* and *GFI1B*/*GFI1Bm* expression observed are consistent with the cross regulatory relationships we have proposed. Thus, total *GFI1B* levels are lower than those of *GFI1Bm*, consistent with its negative-autoregulatory behaviour. In contrast the extent of *SPI1* knockdown observed is more marked in the *GFI1B* than the *GFI1Bm* setting, which is consistent with the proposed repressive role of GFI1B on *SPI1*.

We next transduced both adult mobilized CD34+ cells and cord blood derived CMPs with these 'dual' vectors and cultured them post transduction under cytokine conditions supportive of multineage differentiation, which was assessed with lineage-affiliated cell surface markers. Strikingly, and in both primary cell settings, knockdown of *SPI1* activity, almost completely abrogated

the lineage impact of *GFI1Bm* expression (Figure 6, C, D and E). Similarly, the effect of *GFI1Bm* on CD4 expression could be attenuated by *SPI1* knockdown (Figure 6F). SPI1 has a half-life of approximately 70 hours.[42] We sorted CD34+/GFP+ cells 72 hours after lentivirus transduction and then cultured them for 5-6 days under differentaiting conditions before analyzing only the GFP+ cells. Thus, our experiments were performed more than 3 SPI1 protein half lives after viral transduction, thereby excluding the possibility residual endogenous protein could be interfering in our experiments. These results suggest that most - if not all - of *GFI1Bm*'s activity in this experimental setting is routed through the master regulator *SPI1*.

DISCUSSION

Human cancer cells carry an extensive number of “driver” and neutral or “passenger” mutations.[43] Therefore, the frequent descriptions of mutations in malignant samples, do not necessarily imply a causal relationship between a genetic defect and a clinical phenotype. Establishing a link between such mutations and the malignant process requires functional assessment of their biological impact in the context of appropriate normal and malignant primary cells. Against this backdrop, we have identified a human *GFI1B* point mutation (D262N) in AML and explored its biological properties in both normal and pre-leukemic (MDS) cells. We show that (i) the mutant GFI1B molecule impairs the function of wild-type GFI1B, (ii) its expression in primary hematopoietic progenitor cells of bone marrow or cord blood derivation, leads to a block in erythroid differentiation, an increase in myeloid output and a reduction in cell

death and (iii) it enhances cell survival of primary MDS samples under conditions of limiting growth factor availability. Taken together these findings establish a causal link between this mutation and the malignant process. Through survey of gene expression data from primary AMLs we noted that leukemias with low *GFI1B* levels displayed high expression of *SPI1*, prompting us to test, and indeed prove, the hypothesis that the effects of this new mutation in *GFI1B* are routed through the nodal master regulator *SPI1*. From a mechanistic and transcriptional network perspective, these results suggest that pathways leading to enhanced expression of *SPI1* may have a broad role in the pathogenesis of AML.

High levels of expression of *GFI1B* have been observed in some patients with leukemia and also in leukemic cell lines.[44,45] However, the recent findings showing that Gfi1b is a direct repressor of oncogene *Meis1* supports the notion that GFI1B can be involved in leukemia when its repressor function is abolished.[46] Other genes including *SPI1*.[47-50] are associated with hematological malignancy both when up regulated and functionally inactivated This can be the case in proteins in which dose can have a deep impact on cell function and could explain the tight regulation of GFI1B expression level.[22]

Spi1 over expression produces erythroleukemia in mice.[47] Furthermore, *SPI1* is over expressed in human *KMT2A* (*MLL)*-rearranged AMLs when compared to normal bone marrow CD34 cells,[49] and in mice, it has been shown to play an essential role in the initiation and maintenance of these leukemias.[49,50] A subset of SPI1 target genes relevant to KMT2A leukemias is also regulated by

MEIS1.[49] That GFI1B regulates both *SPI1* and *MEIS1* places it upstream in the same leukemogenic pathway. Therefore, our data uncover a previously unrecognized route to the dysregulation of SPI1 in AML. Recently, mutations in SPI1 have been found in 3 of 31 (10%) *KMT2A*-rearranged and 1 in 384 non-*KMT2A*-rearranged AMLs.[51] Two of the mutations in KMT2A leukemias were associated with reduced expression of SPI1 target genes. However, these authors did not find *SPI1* among the genes whose expression level is altered in KMT2A AML compared to other AMLs.[51] These data underline the heterogeneity of AML, in which seemingly disparate subsets of patients may share specific transcriptional networks that could be therapeutically targeted.

The mutation described here appeared at the evolution of AML from a case of MDS, and it was not detectable during this earlier phase of the disease. NGS indicates that it is present in 54% of *leukemic* cells. It is known that clonal heterogeneity is common in AML, importantly, different subclones are relevant at different stages of the disease,[52] and similarly new clonal events almost certainly drive the evolution of MDS to AML.

Although this is an infrequent mutation, we have demonstrated clear biological activity and a link between GFIB mutations and SPI1 deregulation. Our findings reinforce the importance of the SPI1 pathway in leukemogenesis.

CONCLUSIONS

Our data link GFI1B to both the myeloid and erythroid transcriptional networks

by its repression of *SPI1* and promotion of *MLLT3* expression. This increases the understanding of its role in lineage specification and its potential in promoting blood malignancy. Our clinical findings and experimental data show that the GFI1B D262N mutant plays a role in AML in humans and does so primarily through the agency of the master transcriptional regulator SPI1, reflecting the physiological function of GFI1B in *SPI1* regulation.

Funding

This work was supported in part by the Spanish Ministry of Science (Ref. SAF2005-04709/), Madrid Regional Government/UCM (Ref. CCG08-UCM/BIO-3997), LLR, MRC Molecular Hematology Unit Intramural Programme, UCL CBRC and BLUEPRINT (UK). EA was supported by the Ramón y Cajal and José Castillejo Programs and SEHH and ICRETT (UICC) traveling fellowships. VO and CP were supported by Swedish Foundation for Strategic Research (SSF).

ACKNOWLEDGEMENTS

The authors thank BGI (Tai Po, Hong Kong) for NGS; Douglas R Higgs and Veronica Buckle for support and discussions and Kevin Clarke for cell sorting assistance (MHU, WIMM); Cristina Pina (UCL) for assistance with colony counting and discussions; Kath Stacey (National Blood Service, John Radcliffe Hospital, Oxford) for human cells and Jose Teles (Lund University, SE) for discussion.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Sanders MA, Valk PJ. The evolving molecular genetic landscape in acute myeloid leukaemia. *Curr Opin Hematol* 2013; **20**: 79–85. DOI: 10.1097/MOH.0b013e32835d821c.
2. Germing U, Strupp C, Kuendgen A, Isa S, Knipp S, Hildebrandt B Giagounidis A, Aul C, Gattermann N, Haas R. Prospective validation of the WHO proposals for the classification of myelodysplastic syndromes. *Haematologica* 2006; **91**: 1596–604.
3. Chickarmane V, Enver T, Peterson C. Computational Modeling of the Hematopoietic Erythroid-Myeloid Switch Reveals Insights into Cooperativity, Priming, and Irreversibility. *PLoS Comput Bio* 2009; **5**: e1000268. DOI: 10.1371/journal.pcbi.1000268.
4. Graf T, Enver T. Forcing cells to change lineages. *Nature* 2009; **462**: 587–94. DOI: 10.1038/nature08533.
5. Tong B, Grimes HL, Yang TY, Bear SE, Qin Z, Du K, El-Deiry WS, Tsichlis PN. The Gfi-1B proto-oncoprotein represses p21WAF1 and inhibits myeloid cell differentiation. *Mol Cell Biol* 1998; **18**: 2462–2473.
6. Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Mol Cell Biol* 1996; **16**: 6263–6272.
7. Zweidler-Mckay PA, Grimes HL, Flubacher MM, Tsichlis PN. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Mol Cell Biol* 1996; **16**: 4024–34.

8. Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, Dührsen U, Möröy T. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. *Nat Genet* 2002; **30**: 295-300. DOI:10.1038/ng831.
9. Hock H, Hamblen MJ, Rooke HM, Traver D, Bronson RT, Cameron S, Orkin SH. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity* 2003; **18**: 109–20. DOI:10.1016/S1074-7613(02)00501-0.
10. Person RE, Li FQ, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL, Horwitz M. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. *Nat Genet* 2003; **34**: 308–12. DOI:10.1038/ng1170.
11. Yücel R, Karsunky H, Klein-Hitpass L, Möröy T. The transcriptional repressor Gfi1 affects development of early, uncommitted c-Kit+ T cell progenitors and CD4/CD8 lineage decision in the thymus. *J Exp Med* 2003; **197**: 831–44. DOI: 10.1084/jem.20021417.
12. Spooner CJ, Cheng JX, Pujadas E, Laslo P, Singh H. A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. Immunity 2009; **31**:576–586. DOI: 10.1016/j.immuni.2009.07.011.
13. Saleque S, Cameron S, Orkin SH. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* 2002; **16**: 301–6. DOI: 10.1101/gad.959102.

14. Vassen L1, Beauchemin H, Lemsaddek W, Krongold J, Trudel M, Möröy T. Growth factor independence 1b (gfi1b) is important for the maturation of erythroid cells and the regulation of embryonic globin expression. *PLoS One* 2014 May 6; **9**: e96636. DOI: 10.1371/journal.pone.0096636. eCollection 2014.

15. Khandanpour C, Sharif-Askari E, Vassen L, Gaudreau MC, Zhu J, Paul WE, Okayama T, Kosan C, Möröy T. Evidence that growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. *Blood* 2010; **116**: 5149–61. DOI: 10.1182/blood-2010-04-280305.

16. Stevenson WS1, Morel-Kopp MC, Chen Q, Liang HP, Bromhead CJ, Wright S, Turakulov R, Ng AP, Roberts AW, Bahlo M, Ward CM. GFI1B mutation causes a bleeding disorder with abnormal platelet function. J *Thromb Haemost* 2013;**11**: 2039–47. DOI: 10.1111/jth.12368.

17. Monteferrario D1, Bolar NA, Marneth AE, Hebeda KM, Bergevoet SM, Veenstra H, Laros-van Gorkom BA, MacKenzie MA, Khandanpour C, Botezatu L, Fransen E, Van Camp G, Duijnhouwer AL, Salemink S, Willemsen B, Huls G, Preijers F, Van Heerde W, Jansen JH, Kempers MJ, Loeys BL, Van Laer L, Van der Reijden BA. A dominant-negative GFI1B mutation in the gray platelet syndrome. *N Engl J Med* 2014; 370: 245–53. DOI: 10.1056/NEJMoa1308130.

18. Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 1994; **265**: 1573-7. DOI: 10.1126/science.8079170.

19. McKercher SR, Torbett BE, Anderson KL, Henkel GW, Vestal DJ, Baribault H, Klemsz M, Feeney AJ, Wu GE, Paige CJ, Maki RA. Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J* 1996; **15**: 5647–58.

20. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science* 2000; **288**:1439–41. DOI:10.1126/science.288.5470.1439

21. Dahl R, Walsh J C, Lancki D, Laslo P, Iyer SR, Singh H, Simon MC. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBP ratio and granulocyte colony-stimulating factor. *Nature Immunology* 2003; **4**: 1029–36. DOI: 10.1038/ni973

22. Hernández A, Villegas A, Anguita E. Human promoter mutations unveil Oct-1 and GATA-1 opposite action on Gfi1b regulation. *Ann Hematol* 2010; **89**: 759–65. DOI: 10.1007/s00277-009-0900-x.

23. Gupta R, Hong D, Iborra F, Sarno S, Enver T. NOV (CCN3) functions as a regulator of human hematopoietic stem or progenitor cells. *Science* 2007; **316**: 590–3. DOI:10.1126/science.1136031

24. Rosa A, Ballarino M, Sorrentino A, Sthandier O, De Angelis FG, Marchioni M, Masella B, Guarini A, Fatica A, Peschle C, Bozzoni I. The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. *Proc Natl Acad Sci U S A* 2007; **104**: 19849–54. DOI: 10.1073/pnas.0706963104.

25. Guenechea G, Gan OI, Inamitsu T, Dorrell C, Pereira DS, Kelly M, Naldini L, Dick JE. Transduction of human CD34+ CD38- bone marrow and cord blood-derived SCID-repopulating cells with third-generation

lentiviral vectors. *Mol Ther* 2000; **1**: 566–73. DOI: 10.1006/mthe.2000.0077.

26. Doulatov S, Notta F, Eppert K, Nguyen LT, Ohashi PS, Dick JE. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nat Immunol* 2010; **11**: 585–93. DOI: 10.1038/ni.1889.
27. Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, Higgs DR. Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J* 2004; **23**: 2841–52. DOI: 10.1038/sj.emboj.7600274.
28. Anguita E, Villegas A, Iborra F, Hernández A. Gfi1b controls its own expression binding to multiple sites. *Haematologica* 2010; **95**: 36–46. DOI: 10.3324/haematol.2009.012351.
29. De Gobbi M, Anguita E, Hughes J, Sloane-Stanley JA, Sharpe JA, Koch CM, Dunham I, Gibbons RJ, Wood WG, Higgs DR. Tissue-specific histone modification and transcription factor binding in alpha globin gene expression. *Blood* 2007; **110**: 4503–10. DOI: 10.1182/blood-2007-06-097964.
30. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Löwenberg B, Delwel R. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 2004; **350**: 1617–28. DOI: 10.1056/NEJMoa040465.
31. Shea MA, Ackers GK. The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J Mol Biol* 1985; **181**:

211–30. DOI:10.1016/0022-2836(85)90086-5

32. Huang DY, Kuo YY, Chang ZF. GATA-1 mediates auto-regulation of Gfi-1B transcription in K562 cells. *Nucleic Acids Res* 2005; **33**: 5331–42. DOI: 10.1093/nar/gki838.

33. Doan LL, Porter SD, Duan Z, Flubacher M.M, Montoya D, Tsichlis PN, Horwitz M, Gilks CB, Grimes HL. Targeted transcriptional repression of Gfi1 by GFI1 and GFI1B in lymphoid cells. *Nucleic Acids Res* 2004; **32**: 2508–19. DOI: 10.1093/nar/gkh570.

34. Vassen L, Fiolka K, Mahlmann S, Möröy T. Direct transcriptional repression of the genes encoding the zinc-finger proteins Gfi1b and Gfi1 by Gfi1b. *Nucleic Acids Res* 2005; **33**: 987–98. DOI: 10.1093/nar/gki243.

35. Randrianarison-Huetz V, Laurent B, Bardet V, Blobe GC, Huetz F. Duménil D. Gfi-1B controls human erythroid and megakaryocytic differentiation by regulating TGF-beta signaling at the bipotent erythro-megakaryocytic progenitor stage. *Blood* 2010; **115**: 2784–95. DOI: 10.1182/blood-2009-09-241752.

36. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 2007; **130**: 77—88. DOI:10.1016/j.cell.2007.05.042.

37. Laurent B, Randrianarison-Huetz V, Kadri Z, Roméo PH, Porteu F, Duménil D. Gfi-1B promoter remains associated with active chromatin marks throughout erythroid differentiation of human primary progenitor cells. *Stem Cells* 2009; **27**: 2153—2162. DOI: 10.1002/stem.151.

38. Osawa M, Yamaguchi T, Nakamura Y, Kaneko S, Onodera M, Sawada K, Jegalian A, Wu H, Nakauchi H, Iwama A. Erythroid expansion

mediated by the Gfi-1B zinc finger protein: role in normal hematopoiesis. *Blood* 2002; **100**: 2769–77. DOI: 10.1182/blood-2002-01-0182.

39. Wilson NK, Foster SD, Wang X, Knezevic K, Schütte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E, Pimanda JE, de Bruijn MF, Göttgens B. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* 2010; **7**: 532–44. DOI: 10.1016/j.stem.2010.07.016.
40. May G, Soneji S, Tipping AJ, Teles J, McGowan SJ, Wu M, Guo Y, Fugazza C, Brown J, Karlsson G, Pina C, Olariu V, Taylor S, Tenen DG, Peterson C, Enver T. Dynamic analysis of gene expression and genome-wide transcription factor binding during lineage specification of multipotent progenitors. *Cell Stem Cell* 2013; **13**: 754–68. DOI: 10.1016/j.stem.2013.09.003.
41. Pina C, May G, Soneji S, Hong D, Enver T. MLLT3 regulates early human erythroid and megakaryocytic cell fate. Cell *Stem Cell* 2008; **2**: 264–73. DOI: 10.1016/j.stem.2008.01.013.
42. Kueh HY, Champhekar A, Nutt SL, Elowitz MB, Rothenberg EV. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* 2013; **341**: 670–3. DOI: 10.1126/science.1240831.
43. Greenman C1, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R,

Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR. Patterns of somatic mutation in human cancer genomes. *Nature* 2007; **446**: 153–8. DOI: 10.1038/nature05610

44. Elmaagacli AH, Koldehoff M, Zakrzewski JL, Steckel NK, Ottinger H, Beelen DW. Growth factor-independent 1B gene (GFI1B) is overexpressed in erythropoietic and megakaryocytic malignancies and increases their proliferation rate. *Br J Haematol* 2007; **136**: 212–19. DOI: 10.1111/j.1365-2141.2006.06407.x.
45. Vassen L, Khandanpour C, Ebeling P, van der Reijden BA, Jansen JH, Mahlmann S, Dührsen U, Möröy T. Growth factor independent 1b (Gfi1b) and a new splice variant of Gfi1b are highly expressed in patients with acute and chronic leukemia. *Int J Hematol* 2009; **89**: 422–30. DOI: 10.1007/s12185-009-0286-5.
46. Chowdhury AH, Ramroop JR, Upadhyay G, Sengupta A, Andrzejczyk A, Saleque S. Differential Transcriptional Regulation of meis1 by Gfi1b and Its Co-Factors LSD1 and CoREST. *PLoS One* 2013; **8**: e53666. DOI: 10.1371/journal.pone.0053666.
47. Moreau-Gachelin F, Wendling F, Molina T, Denis N, Titeux M, Grimber G Briand P, Vainchenker W, Tavitian A. Spi-1/PU.1 transgenic mice develop multistep erythroleukemias. *Mol Cell Biol* 1996; **16**: 2453–63.

48. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, Akashi K, Fiering S, Tenen DG. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* 2004; **36**: 624–30. DOI: 10.1038/ng1361.

49. Zhou J, Wu J, Li B, Liu D, Yu J, Yan X, Zheng S, Wang J, Zhang L, Zhang L, He F, Li Q, Chen A, Zhang Y, Zhao X, Guan Y, Zhao X, Yan J, Ni J, Nobrega MA, Löwenberg B, Delwel R, Valk PJ, Kumar A, Xie L, Tenen DG, Huang G, Wang QF. PU.1 is essential for MLL leukemia partially via crosstalk with the MEIS/HOX pathway. *Leukemia* 2014; **28**: 1436–48. DOI: 10.1038/leu.2013.384.

50. Aikawa Y, Yamagata K, Katsumoto T, Shima Y, Shino M, Stanley ER, Cleary ML, Akashi K, Tenen DG, Kitabayashi I. Essential role of PU.1 in maintenance of mixed lineage leukemia-associated leukemic stem cells. *Cancer Sci* 2015; **106**: 227–36. DOI: 10.1111/cas.12593.

51. Lavallée VP, Baccelli I, Krosl J, Wilhelm B, Barabé F, Gendron P, Boucher G, Lemieux S, Marinier A, Meloche S, Hébert J, Sauvageau G. The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet* 2015; **47**: 1030–7. DOI: 10.1038/ng.3371.

52. Graubert TA, Mardis ER. Genomics of acute myeloid leukemia. *Cancer J* 2011; **17**: 487–91. DOI: 10.1097/PPO.0b013e31823c5652.

FIGURE LEGENDS

Figure 1. A point mutation in *GFI1B* alters a conserved amino acid in a patient with AML arising from MDS. Analysis of the presentation MDS suggests the *GFI1B* mutation appeared on leukemic transformation.

(**A**) DHPLC traces from the index patient at the diagnosis of AML (red) and a normal control (blue). The diagram on the top represents the result after running the samples at 63ºC and at the bottom are the tracings performing the experiment at 65ºC. This is a representative result from four independent experiments.

(**B**) Sequence analysis of the patient at the diagnosis of AML. Top, direct sequence of PCR products; middle, mutant allele after subcloning the PCR products (2 mutant clones in 17 total analyzed were found two independent experiments); bottom, wild-type allele. Arrow indicates the mutation.

(**C**) Protein sequence alignment of *GFI1B* orthologs including the mutated amino acid (arrow). Human (Homo sapiens); Mouse (Mus musculus); Rat (Rattus norvegicus); Dog (Canis familiaris); Opossum (Monodelphis domestica); Platypus (Ornithorhynchus anatinus); Chicken (Gallus gallus); Anole lizard (Anolis carolinensis); Xenopus (Xenopus tropicalis); Fugu (Takifugu rubripes); C elegans (Caenorhabditis elegans); D melanogaster (Drosophila melanogaster); sens-2 (senseless-2); sens (senseless).

(**D**) Photomicrographs of peripheral blood (PB) slides (Wright's stain) from the MDS diagnostic sample of the index patient showing anisopoikilocytosis, dysplastic features and circulating myeloblasts consistent with the diagnosis of RAEB. Images have been taken with an Olympus BX 60 microscope and Q

imaging, QICAM, camera and Openlab 5.5.1 software. Top, 40x and bottom 100x objective, respectively.

(**E**) DHPLC tracing of archival DNA from the index patient at the time of diagnosis of MDS (red) and a normal control (blue). The plot on the top shows the result after running the samples at 63ºC and at the bottom are the records performing the experiment at 65ºC. The abnormal peaks subsequently observed when the patient developed AML were not present at this point.

(**F**) Direct sequencing at MDS diagnosis. The arrow indicates the base where the mutation was found, undetected in this experiment.

Figure 2 GFI1B D262N impairs the function of the normal protein.

(**A**) Western blot of total protein extracts from K562 cells transduced with either *GFI1B* or *GFI1Bm* in either CWI-emerald (CWI-em, Wiskott Aldrich promoter) or CSI-emerald (CSI-em, spleen focus forming virus promoter) vectors, for low and high level of enforced gene expression respectively. Ev, empty vector; SIN3A, loading control; UT, untransduced.

(**B**) Quantitative real time RT-PCR analysis of endogenous *GFI1B* (e*GFI1B*) expression normalized to the empty vector mRNA level (arbitarily set to 1) and *GAPDH* in K562 cells transduced with either CWI-em or CSI-em. Error bars indicate mean +SEM of 3 and 2 independent experiments, respectively. P value for *GFI1Bm vs GFI1B*.

(**C**) Quantitative RT-PCR assay of e*GFI1B* expression in CD34+GFP+ primary cells infected with the CSI-em lentiviruses. Analysis as in Figure 2B. Error bars indicate mean +SEM of 3 independent infections. P value for *GFI1Bm vs GFI1B*.

(**D**) Quantitative real time RT-PCR analysis of GFI1B targets *GFI1* and *TGFBR3* analyzed as in Figure 2B of K562 cells transduced with CWI-em or CSI-em lentiviruses, as indicated. Error bars indicate mean +SEM 3 independent experiments.

(**E**) Diagram of human *GFI1B* locus; coordinates relative to ATG start codon; promoter (P) is pointed with a black arrow; the downstream regulatory elements (Reg. Elements) are shown with black boxes over DNA sequence (black horizontal line); grey boxes on the sequence are *GFI1B* exons; black boxes numbered 1-6 indicate the position of amplicons used in real time ChIP PCR.

(**F-G**) Chromatin immunoprecipitation (ChIP) assay of lysine 4 of histone 3 trimethylation (H3K4me3) (F) and histone 3, lysine 4 dimethylation (H3K4me2) (G). Numbers correspond to the amplicons indicated in Figure 2E. Results in *HBZ* (*ζ globin*, *ζ Gl.*) and *HBG* (*γ globin*, *γ Gl.*) promoter controls are shown. Error bars indicate mean +SEM of 3 to 4 experiments.

Figure 3. *GFI1B* mutant blocks erythroid differentiation and promotes myeloid output and cell survival in normal human primary cells.

(**A**) Total *GFI1B* expression in adult CD34+ cells, obtained from G-CSF mobilized adult human peripheral blood from patients undergoing autograft procedures for non-hematological disorders, transduced with either *GFI1B* or *GFI1Bm* in CSI-emerald (CSI-em) or CWI-emerald (CWI-em) vectors. Ev, empty vectors. Analysis as in Figure 2B, error bars indicate mean +SEM of 2 independent experiments.

(**B**) Representative histograms showing lineage output of virally transduced (GFP+) adult bone marrow derived CD34+ cells (CSI-em vector) after 6 days

liquid culture with SCF, FLT3L, IL-3, GM-CSF, G-CSF and EPO (multipotential cytokines); (total number of experiments (n)= 5). Erythroid differentiation was assesed by surface expression of CD36 and CD235a and myeloid differentiation by CD13. CD36 identifies early erythroid cells as well as monocytes, CD235a is a specific marker of terminal erythroid differentiation.

(**C**) Representative flow cytometric profiles of GFP+ cord blood CMPs following transduction with the indicated CSI-em viruses and liquid culture with multipotential cytokines for 5 days.

(**D**) Myeloid and erythroid colony output from bone marrow-derived CD34+ cells transduced with CSI vectors carrying either *GFI1B* or *GFI1Bm*. Transduced cells were maintained in liquid culture with multipotential cytokines for 3 days before GFP+ cells were isolated, and 1000 plated in methylcellulose containing multipotential cytokines. Colonies were scored after a further for 15 days (in each experiment, identical numbers of GFP+ cells were plated, error bars indicate mean +SEM of 5 experiments).

(**E**) CD36+/CD235a+ output from GFP+ CD34+ cells transduced with either *GFI1B* or *GFI1Bm* in CSI-em or CWI-em vectors following 7 days liquid culture with IL-3, IL-6 and EPO (mean +SEM of 3 and 5 independent experiments, respectively).

(**F**) Annexin V staining of transduced (CSI-em) CD34+ cells after 5 days in the absence of cytokines (mean +SEM of 3 experiments).

(**G**) Representative histograms showing annexin V and Hoechst 33258 staining. *p<0.05; **p<0.005; ***p<0.0005.

Figure 4. GFI1B-GFI1Bm versus SPI1 cross regulation in normal and AML cells suggest that GFI1Bm functions through *SPI1* deregulation.

(**A**) Expression of *SPI1*, *GFI1B* and *GATA1* in 285 unselected AML patients. Pairwise correlations between samples calculated on the basis of 2856 probe sets are displayed as described.[30] Colors of boxes visualize Pearson's correlation coefficient: red indicates higher positive correlation, blue indicates higher negative correlation. 16 clusters of AMLs samples with similar gene expression profiles were uncovered.[30] The histograms next to each sample represent levels of *SPI1, GFI1B* and *GATA1* mRNA expression for each patient sample.

(**B**) The transcription factor GFI1B-SPI1 interaction circuit for a heteroygous *GFI1B* mutation. The dashed lines indicate the effect of mutation on the *GFI1B* repression of itself, the *GFI1B* mutant and *SPI1*.

(**C**) Time series concentrations of *GFI1B* (blue) and S*PI1* (gray) indicating the final steady state values. In the normal case (no mutations) the steady state is low for *SPI1* and high for *GFI1B*.

(**D**) Time series concentrations of *GFI1B* wild-type (blue), *GFI1B* mutant (red) and *SPI1* (gray) mRNA showing the effect of the *GFI1B* mutation. The presence of the *GFI1B* mutant allele results in elevated expression of *SPI1* and both mutant and wild-type *GFI1B*. The expression values of the transcription factors in this circuit are increased as compared to the initial conditions. Note that *GFI1B* and *SPI1* curves are superimposed. The dynamical equations underlying panels (C) and (D) are described in the Supplementary Methods.

(**E**) Quantitative RT-PCR assay of *SPI1* relative (R) expression in K562 cells infected with the CWI-em viruses and plotted as in Figure 2B. Error bars indicate mean +STD of 2 independent experiments *p<0.05 (0.0079).

(**F**) Quantitative RT-PCR of *SPI1* expression in GFP+/CD13-/CD235a- cells derived from adult CD34+ cells infected with CSI-em viruses and plotted as in Figure 2B.

(**G**) Cross-species sequence alignment at the *SPI1* promoter. Solid bar indicates the core motif of GFI1B binding site (AATC). The coordinates of the start of this sequence in human, GRCh/hg19, (H); and mouse, NCBI37/mm9, (M) are shown on top of the alignment. H, human (Homo sapiens); M, mouse (Mus musculus); D, dog (Canis familiaris); C, cow (Bos taurus); Op, opossum (Monodelphis domestica).

(**H**) ChIP assay with antibody anti-GFI1B (α-GFI1B) and no antibody control (K562 cells, CSI vectors). Results at *SPI1* promoter, *GFI1B* promoter (positive control) and *HBZ* (ζ *globin*, ζ *Gl.*) and *HBG* (γ *globin*, γ *Gl.*) promoter (negative controls) are shown. Error bars indicate mean +SEM of 2 independent experiments, these data was confirmed in K562 cells transduced with CWI Ev and overexpresing *GFI1B* (data not shown). *p<0.05, **p<0.005.

Figure 5. **SPI1 regulated gene *CD4* expression correlates with SPI1 levels in AML and increases in the presence *of GFI1B*m. In contrast, the expression of positive regulator of erythroid and megakaryocytic differentiation *MLLT3* is increased by high *GFI1B* expression and reduced by *GFI1Bm*.**

(**A**) Correlation view of 285 AML patients. See legend Figure 4A; mRNA expression of *SPI1* and *CD4* is shown.

(**B-C**) Representative histograms (B) and bar charts (C) showing CD4 expression by the progeny of adult CD34+ cells transduced with either *GFI1B* or *GFI1Bm* in CSI-em and maintained in liquid culture with multipotential cytokines for 5 days; Ev, empty vector; error bars indicate mean +SEM of 3 experiments.

(**D**) Correlation view of 285 AML patients, see legend Figure 4A; mRNA expression of *GFI1B*, *MLLT3* and *SPI1* is shown. Correlation between *MLLT3* and *GFI1B*: Both are low in clusters 5, 9, 11, 12 and 13. Both are high in clusters 7, 8 and 10. Anti-correlation with *SPI1*: *SPI1* is higher in the clusters with low *GFI1B*/*MLLT3*, i.e. clusters 5, 9 and 11, 12 and 13. *SPI1* is low in the clusters with high *GFI1B*/*MLLT3*, i.e. clusters 7, 8 and 10. In other clusters expression of all three genes is intermediate.

(**E-F**) Quantitative real time RT-PCR analysis of *MLLT3* in K562 (**E**) and GFP+ adult CD34+ (**F**) cells transduced with CSI-em lentiviruses. Expression normalized relative to the CSI-em mRNA level (set to 1). Number of experiments (n) is indicated, **p<0.005.

Figure 6. GFI1B D262N differentiation disruption is greatly restored by knockdown of *SPI1*.

(**A**) Quantitative RT-PCR assay of *SPI1* expression in adult GFP+/CD13+/CD235a- cells transduced with either control (SLX) or *SPI1* hairpin (KD) vectors. Analysis as before.

(**B**) Quantitative RT-PCR analysis of *SPI1* and *GFI1B* relative expression (R. Expression) in K562 transduced with control vector (SLX), and *SPI1* hairpin[24]

(KD) vectors carrying either *GFI1B* (KD/*GFI1B*) or *GFI1B*m (KD/*GFI1Bm*) transgenes.

(**C**) Analysis of erythroid output (assessed with erythroid markers CD71 and CD235a) from adult CD34+ cells from G-CSF mobilized peripheral blood with CSI-em and SLX-*SPI1* knock down (KD) lentiviruses, over-expressing either wild-type (*GFI1B*) or mutant (*GFI1Bm*), or CSI-em empty vector (Ev) and SLX controls, and maintained in liquid culture for 5 days with multipotential cytokines. Histograms representative of 5 independent experiments.

(**D**) Representative histograms (left) showing CD13 and CD235a expression in the progeny of normal adult CD34+ cells transduced with CSI-em and SLX-*SPI1* knock down (KD) lentiviruses, over-expressing either wild-type (*GFI1B*) or mutant (*GFI1Bm*) *GFI1B* and maintained in liquid culture with multipotential cytokines for 5 days. Quantitation of CD235a+ expression from 5 such experiments (right); error bars indicate mean +SEM, *p<0.05.

(**E**) Representative histograms (left) showing CD13 and CD235a expression in the progeny of CMPs transduced as in (D) and maintained in liquid culture with multipotential cytokines for 5 days. Quantitation of CD235a+ expression from 2 such experiments (right); error bars indicate mean +SEM, *p<0.05.

(**F**) Left, representative flow cytometric profiles showing CD4 expression in the progeny of normal adult CD34+ cells transduced as indicated. Right, quantitation of CD4+ expression from 3 such experiments; error bars indicate mean +SEM.

Highlights:

1. Novel mutation in GFI1B (D262N) associated with transition to AML from MDS
2. GFI1B-D262N promotes myeloid versus erythroid output and enhances cell-survival
3. *SPI1* (*PU.1*) is a direct target of GFI1B & GFI1B-D262N dysregulates *SPI1* expression
4. Biological impact of the GFI1B-D262N mutation is routed through SPI1

FIGURE 1

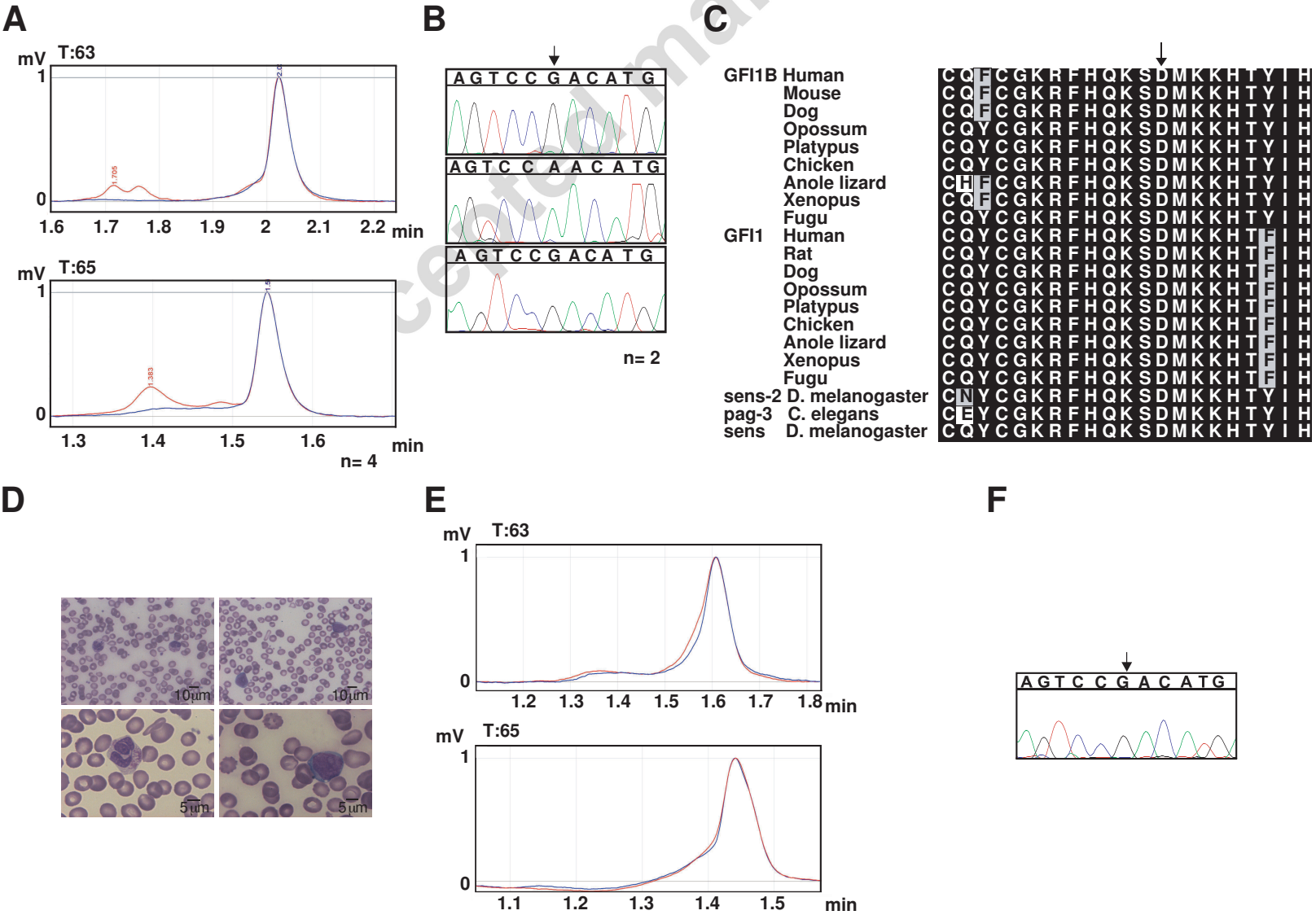


Figure 2

FIGURE 2

Click here to download Figure: Figure 2.ppt

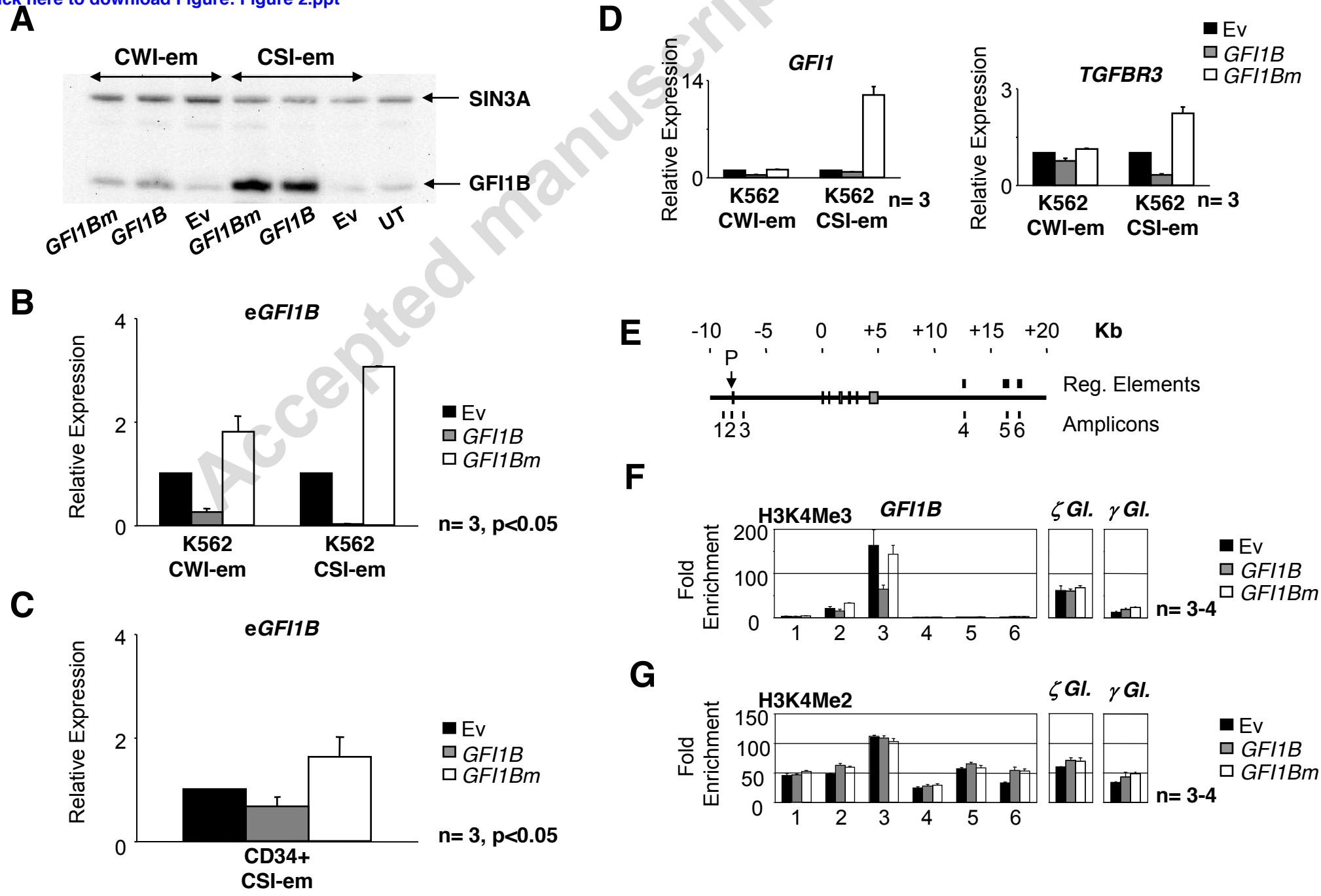


Figure 3

FIGURE 3

Click here to download Figure: Figure 3 .ppt

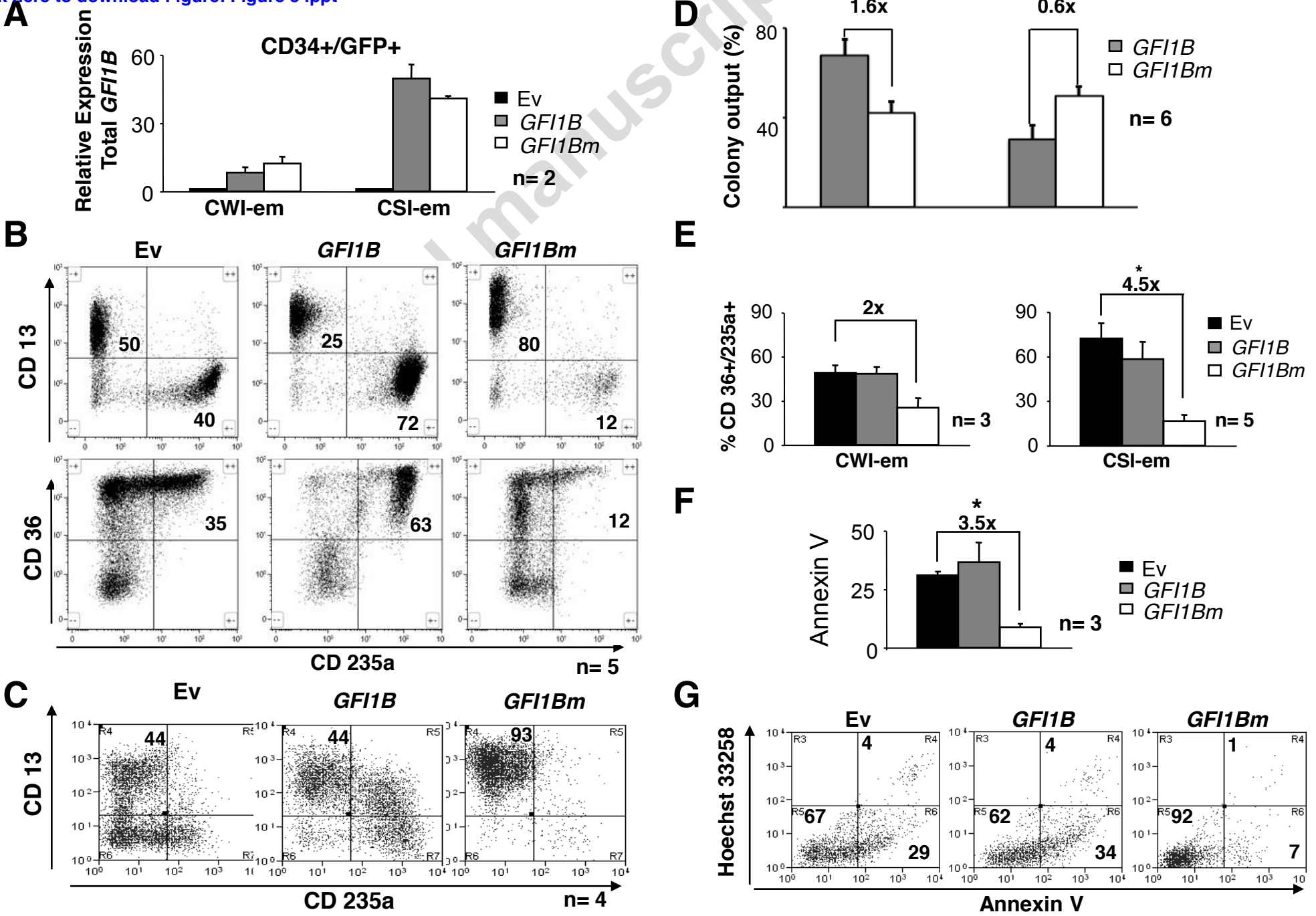


Figure 4

FIGURE 4

Click here to download Figure: Figure 4.ppt

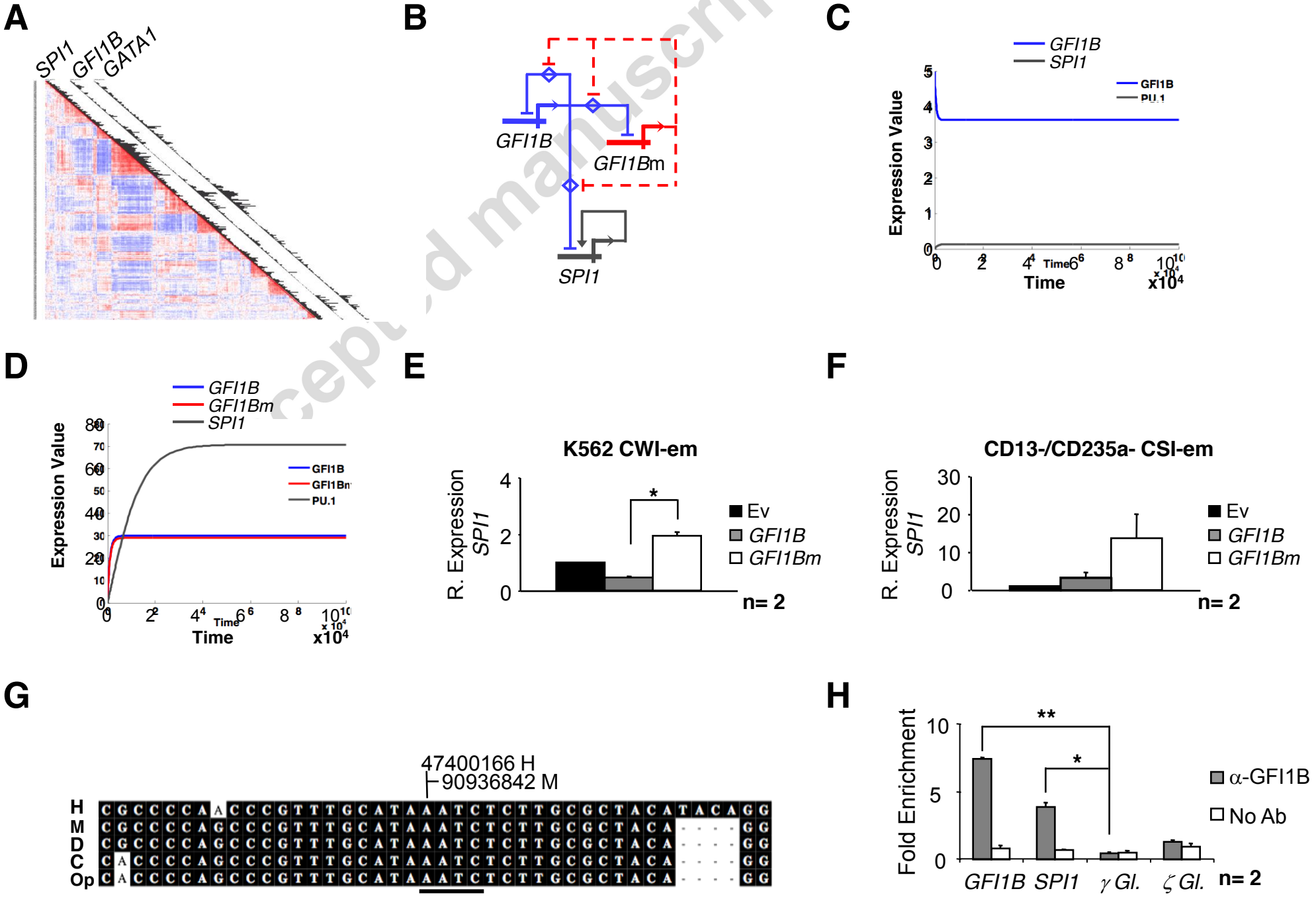
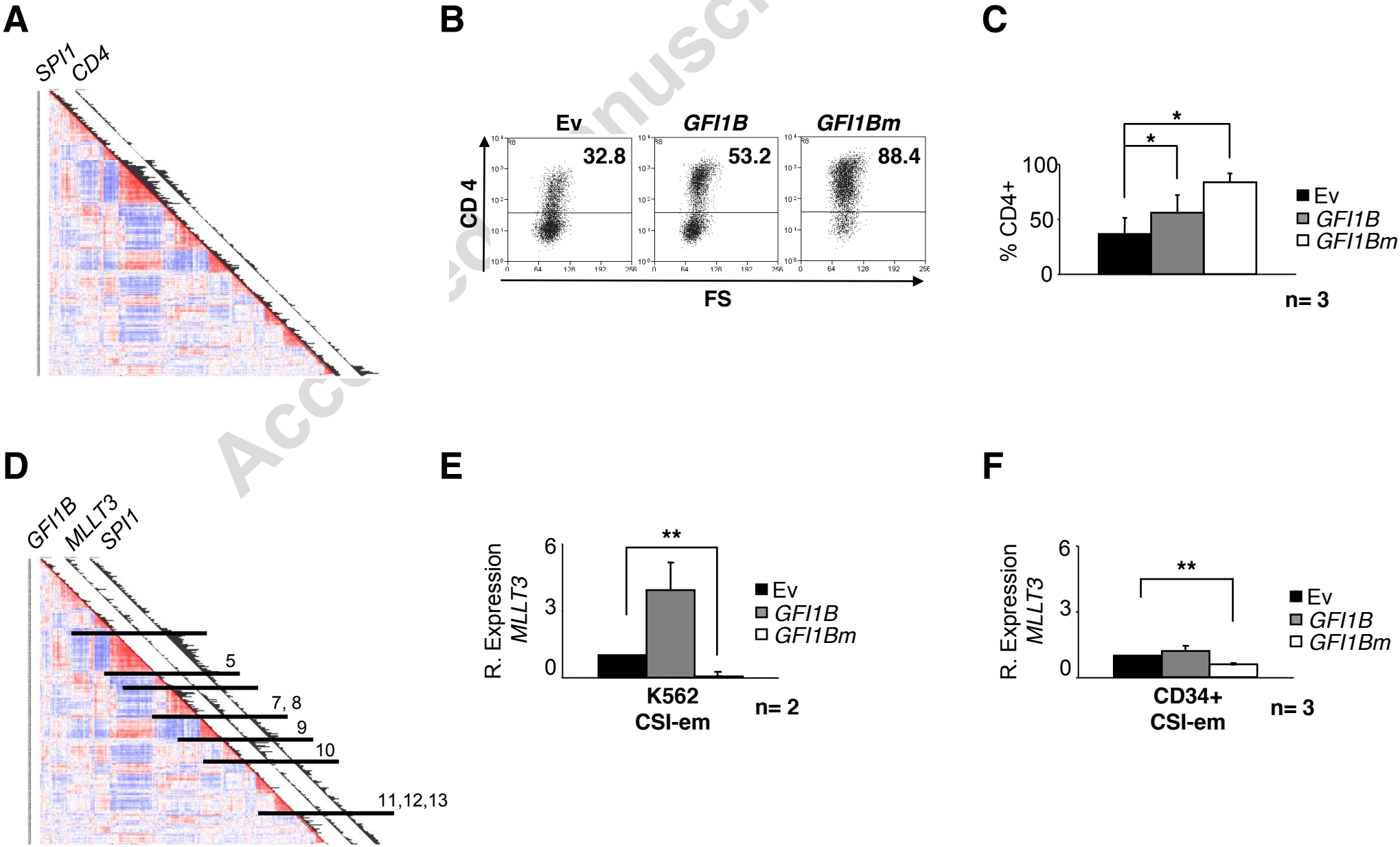
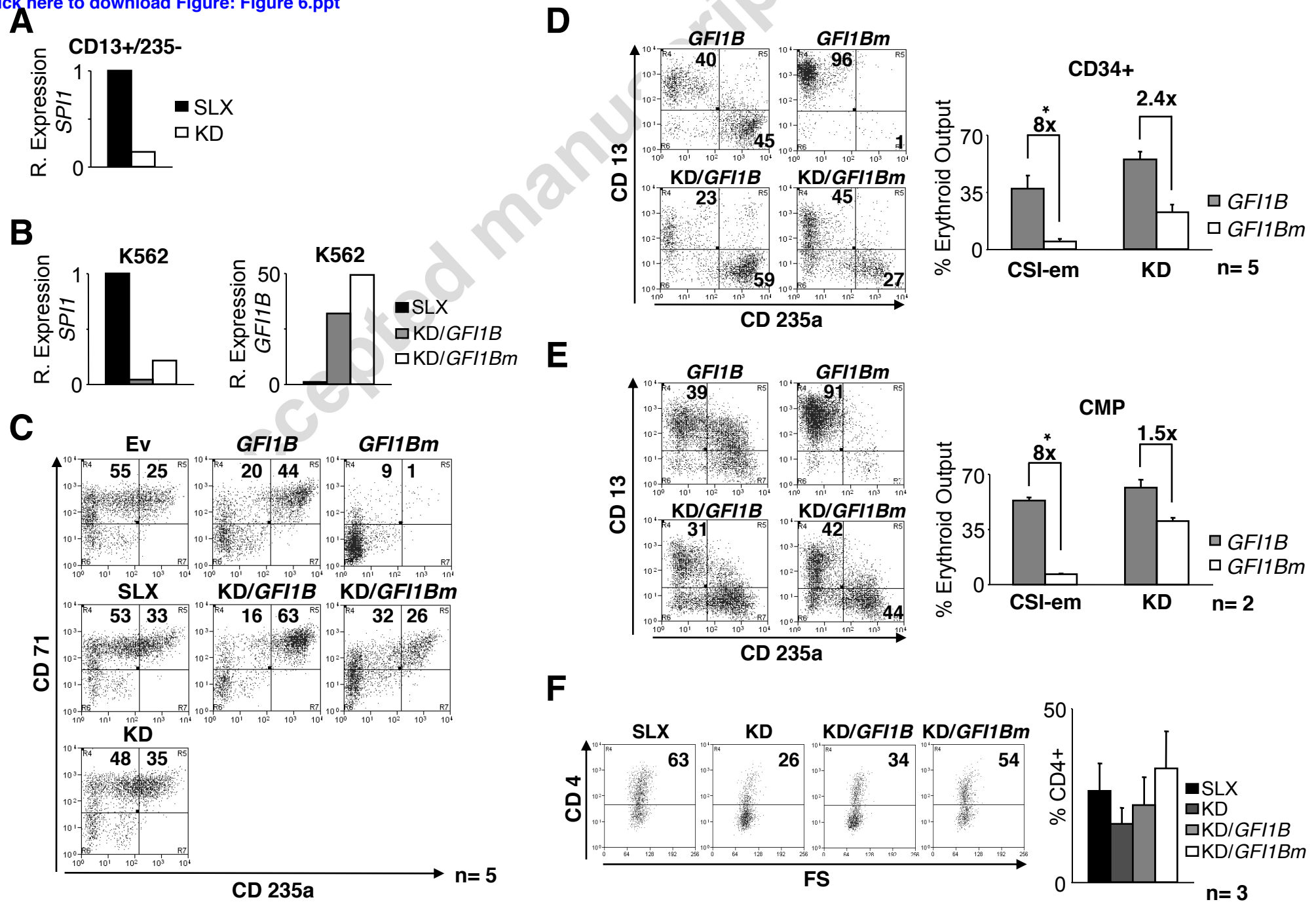


Figure 5

FIGURE 5

Click here to download Figure: Figure 5.ppt





Graphical Abstract