© Ferrata Storti Foundation

Retroviral insertional mutagenesis identifies the del(5q) genes, *CXXC5*, *TIFAB* and *ETF1*, as well as the Wnt pathway, as potential targets in del(5q) myeloid neoplasms

An interstitial deletion of the long arm of chromosome 5, del(5q), is a recurring abnormality in myeloid disorders, including myelodysplastic syndromes (MDS), *de novo* acute myeloid leukemia (AML), and therapy-related myeloid neoplasms (t-MN) comprising therapy-related MDS and AML (t-MDS/t-AML).[1] In t-MN, a del(5q) occurs in approximately 40% of patients and is associated with prior therapy with alkylating agents, a complex karyotype, *TP53* mutations, a strong propensity to progress to t-AML, and a poor outcome.[2,3] We previously identified two haploinsufficient tumor suppressor genes on 5q, the early growth response gene, *EGR1* [(5q31.2, deleted in all t-MN with a del(5q)] and the adenomatous polyposis coli gene, *APC* [5q22.2, deleted in >95% of t-MNs with a del(5q)], and showed that heterozygous loss of *Egr1* and *Apc* in mice promote the pathogenesis of MDS/AML, in co-operation with knockdown of *Trp53* and/or alkylating agent therapy.[4] EGR1 is a member of the WT-1 family of transcription factors and is a transcriptional regulator of many tumor suppressor genes, including *TP53, CDKN1A/p21* and *TGFB*.[5] The APC protein acts as a tumor suppressor and is a negative regulator of the WNT signaling pathway.[6] Adding to the difficulty

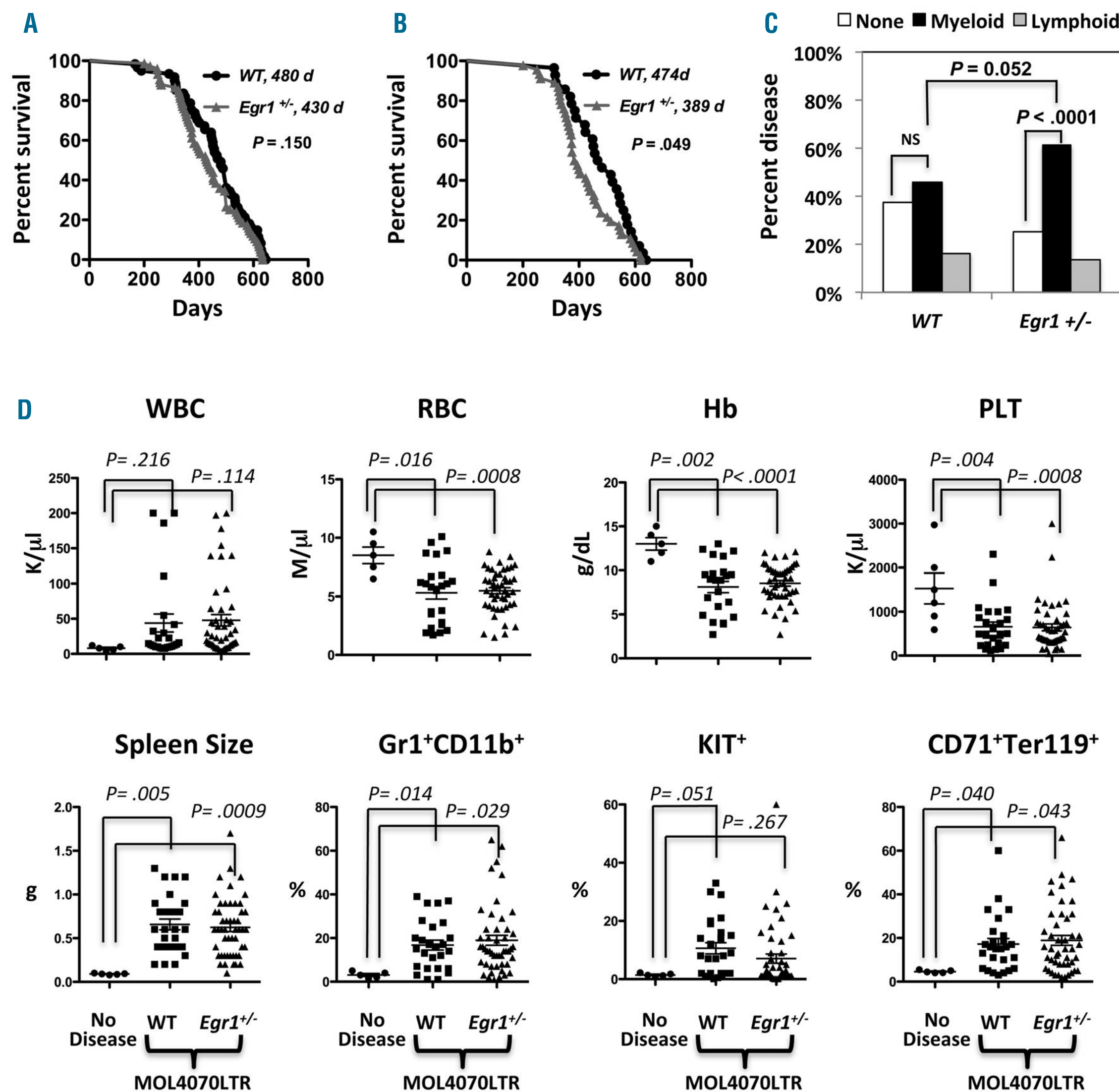
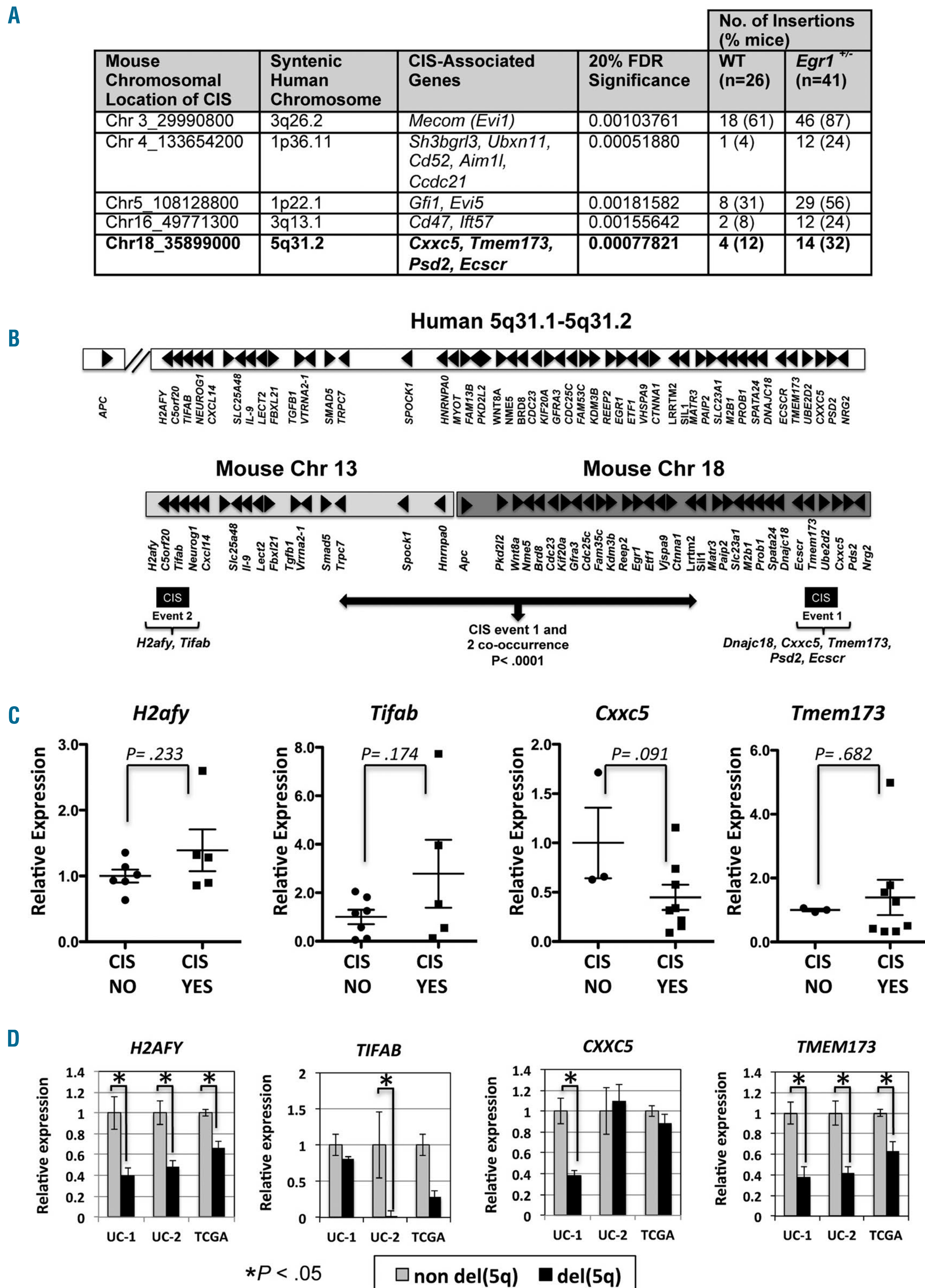


Figure 1. *Egr1* haploinsufficiency accelerates MOL4070LTR-induced myeloid disease. (A and B) Kaplan-Meier survival curves of WT (n=61) and *Egr1*$^{+/-}$ (n=77) neonatal mice injected with MOL4070LTR retrovirus. (A) Overall survival and (B) survival of mice with myeloid neoplasms. Median survival of all mice or mice that specifically developed myeloid disease was 430 and 389 days for *Egr1*$^{+/-}$ and 480 and 474 days for WT mice. Myeloid neoplasms were identified by splenomegaly, leukocytosis, anemia, thrombocytopenia, and enrichment of myeloid cells in hematopoietic and non-hematopoietic tissue. (C) Distribution of the disease phenotypes observed in MOL4070LTR-treated mice. In *Egr1*$^{+/-}$ mice, there is an increase in neoplastic diseases concomitant with an increase in myeloid *versus* lymphoid neoplasms. A trend towards a significant increase in myeloid disease in *Egr1*$^{+/-}$ mice compared to WT was observed (P=0.052 by one-tailed Fisher's exact test). (D) Complete blood counts (CBC) and phenotype analysis of *WT* and *Egr1*$^{+/-}$ mice treated with MOL4070LTR retrovirus, when the mice developed a myeloid neoplasm. "No disease" mice were *Egr1*$^{+/-}$ mice that were not treated with retrovirus and remain healthy. WBC: white blood cells; RBC: red blood cells; Hb: hemoglobin; PLT: platelets. Spleen size (g) and flow cytometric analysis showing granulocytes (Gr1$^+$, CD11b/Mac1$^+$), KIT$^+$ myeloid cells, and immature erythroblasts (CD71$^+$, Ter119$^+$) in the spleen. Significant differences using pairwise 2-sample Student's *t*-tests (Control *vs.* WT; Control *vs.* *Egr1*$^{+/-}$) are shown.

© Ferrata Storti Foundation



Mouse Chromosomal Location of CIS	Syntenic Human Chromosome	CIS-Associated Genes	20% FDR Significance	No. of Insertions (% mice) WT (n=26)	*Egr1*$^{+/-}$ (n=41)
Chr 3_29990800	3q26.2	*Mecom (Evi1)*	0.00103761	18 (61)	46 (87)
Chr 4_133654200	1p36.11	*Sh3bgrl3, Ubxn11, Cd52, Aim1l, Ccdc21*	0.00051880	1 (4)	12 (24)
Chr5_108128800	1p22.1	*Gfi1, Evi5*	0.00181582	8 (31)	29 (56)
Chr16_49771300	3q13.1	*Cd47, Ift57*	0.00155642	2 (8)	12 (24)
Chr18_35899000	**5q31.2**	***Cxxc5, Tmem173, Psd2, Ecscr***	**0.00077821**	**4 (12)**	**14 (32)**

Figure 2. Co-occurring CISs target human del(5q) genes. (A) Common insertion site (CIS) analysis was performed on mice that developed myeloid neoplasms. List of CISs found statistically to occur more frequently in *Egr1*$^{+/-}$ than WT mice. Some mice had multiple insertions for the same CIS. The CIS on chromosome 18 is syntenic to human 5q31.2, and includes *Dnajc18, Cxxc5, Tmem173, Psd2*, and *Ecscr*. (B) Human 5q31.1-q31.2 is syntenic to mouse chromosomes 13 and 18. Fisher's exact test with multiple testing correction revealed that CISs on chromosome 18 (event 1) and on chromosome 13 (event 2) co-occur more frequently than predicted by chance (P<0.0001). The second CIS (event 2) was found with equal frequency in *Egr1*$^{+/-}$ and WT mice. (C) Expression of candidate del(5q) genes was examined by Q-PCR in *Egr1*$^{+/-}$ mice that had proviral integrations proximal to candidate gene (CIS YES) *versus* those that did not (CIS NO). (D) Fold expression of genes was measured by comparing gene expression in MNs with del(5q) compared to non-del(5q). Expression in 2 independent cohorts for the University of Chicago (UC-1; GSE39991 and UC-2; SRA061655) as well as the TCGA database were examined. UC-1: 10 del(5q) and 28 non-del(5q); UC-2: 12 del(5q) and 16 non-del(5q); TCGA: 18 del(5q) and 161 non-del(5q). Evidence for haploinsufficient expression in del(5q) therapy-related myeloid neoplasm patients patients was observed for *H2AFY, TIFAB, TMEM173*, and *CXXC5*.

© Ferrata Storti Foundation

of identifying additional genes on 5q that may be contributing to the pathogenesis of MDS/AML, most patients have large deletions encompassing over 70 Mb (5q14-q33). The complex cytogenetic abnormalities and recurrent somatic mutations, such as *ASXL1, NF1, TET2* and *TP53* associated with advanced MDS and AML with a del(5q), has provided some insight; however, the complete genetic profile, and consequences thereof, are not yet known.[3,7] Whether prior cytotoxic therapy induces mutations and/or promotes expansion of pre-existing genetic mutations in t-MNs is also poorly understood. We used a mutagenesis approach in mice, heterozygous for Egr1 (*Egr1*$^{+/-}$) to identify genes and signaling pathways that predispose to developing del(5q) myeloid neoplasms (MNs).

Forward genetic screens with the recombinant retrovirus MOL4070LTR create a sensitized background for the development of MNs, and have been shown to be a powerful approach for the identification of secondary co-operating mutations.[8] Untreated WT and *Egr1*$^{+/-}$ mice do not develop disease up to two years of age;[5] however, MOL4070LTR-treated WT and *Egr1*$^{+/-}$ mice developed T-cell or myeloid neoplasms with a median survival of 480 and 430 days (d), respectively (*P*=0.150) (Figure 1A). *Egr1*$^{+/-}$ mice developed transplantable MNs with a shorter latency (389 d *vs.* 474 d; *P*=0.049) and at a higher overall frequency than WT littermate controls (Figure 1B and *Online Supplementary Figure S1*). There was also a significant increase in myeloid *versus* no disease in *Egr1*$^{+/-}$ mice (*P*<0.001), but not WT mice (*P*=0.36) (Figure 1C), indicating that loss of one allele of *Egr1* shifts the disease spectrum to favor the development of MNs. A Fisher's exact test indicated that the higher frequency of MNs observed in *Egr1*$^{+/-}$ mice compared to WT showed a trend towards significance (*P*=0.052). The morphological and phenotypic features of the diseases were similar in the *Egr1*$^{+/-}$ and the WT mice (Figure 1D).

We mapped the common retroviral integration sites (CISs) in MNs from WT and *Egr1*$^{+/-}$ mice using bar-coded splinkerette PCR and Illumina high-throughput sequencing. TAPDANCE software, a fully automated process to identify and annotate CISs,[9] identified 159 and 365 CISs (*P*<0.05) in myeloid neoplasms from WT (n=29) mice or *Egr1*$^{+/-}$ mice (n=46), respectively. Whereas none of the CISs were significantly associated with the WT genotype, 5 CISs showed significant associations with the *Egr1*$^{+/-}$genotype (Figure 2A). None of the genes implicated in cell growth and cancer within the common insertion sites (*Gfi1, Ccdc21, Cd47, Cd52, Cxxc5, Evi5, Psd2, Tmem173* and *Ubxn11*) displayed a significant change in expression in myeloid neoplasms due to retroviral integration (*Online Supplementary Figure S2A*) except for *Evi1* (*Mecom*: MDS1 and EVI1 complex locus), which showed a significant increase (*Online Supplementary Figure S2B*). We showed that self-renewal of *Evi1*-expressing progenitors was enhanced by loss of *Egr1 in vitro*; however, there was no apparent co-operation between *Evi1* overexpression and *Egr1* haploinsufficiency leading to hematopoietic neoplasms using an *in vivo* mouse model (*Online Supplementary Figure S2C and D*).

Identifying which genes on 5q play a critical role in the development of MDS and AML continues to be a major challenge. Interestingly, one CIS significantly associated with MNs in *Egr1*+/- mice mapped to a region of mouse chromosome 18 that is syntenic to human 5q31.2 and proximal to the *DNAJC18, ECSCR, TMEM173, CXXC5* and *PSD2* genes (Figure 2B). Of relevance to MNs with a del(5q), this CIS co-occurred with a CIS on chromosome 13 (syntenic to 5q31.1 and proximal to the *TIFAB*, and *H2AFY* genes) (*P*=0.001). In Gr1^{+}CD11b^{+} myeloid bone marrow cells isolated from *Egr1*$^{+/-}$ mice that had proviral integrations proximal to the 5q genes (CIS YES), compared to mice that did not (CIS NO), *Tifab* expression was slightly higher and *Cxxc5* lower (though this was not significant), raising the possibility that aberrant expression of these genes may be important (Figure 2C). In leukemia samples from t-MN patients with a del(5q), *CXXC5, TMEM173, TIFAB*, and *H2AFY* showed approximately 50% lower expression relative to non-del(5q) samples, consistent with haploinsufficiency (Figure 2D). Recent studies suggest that *TIFAB* (TRAF-interacting protein with forkhead-associated domain, family member B) and *CXXC5* (CXXC finger protein 5) may be important genes for the pathogenesis of myeloid disease. Deletion of *Tifab* contributes to an MDS-like phenotype in mice by changing the dynamic range of innate immune pathway activation.[10] *CXXC5*, also known as *RINF*, is a candidate tumor suppressor in AML, since it inhibits leukemia cell proliferation and Wnt signaling.[11,12] Taken together, *TIFAB* and *CXXC5* warrant strong consideration for future studies of myeloid neoplasms associated with a del(5q).

Traditional CIS analysis reveals a potentially important gene (e.g. *Evi1*) only if insertions proximal to a genomic region occur in multiple mice more frequently than could be expected by chance. It does not, however, take into account proviral insertions, proximal to multiple different genes within the same signaling pathway or ontology term, in multiple mice. We applied the Genomic Regions Enrichment of Annotations Tool (GREAT) to evaluate whether *Egr1*$^{+/-}$ mice with myeloid disease had proviral integrations proximal to multiple genes within the same signaling pathway or gene ontology.[13] GREAT analysis revealed the twenty most significant Molecular Signature Database (MSigDB) pathway and gene ontology molecular terms that were enriched in WT and *Egr1*$^{+/-}$ mice that developed MNs (*Online Supplementary Tables S1 and S2*). As expected, several pathway terms were shared in both WT and *Egr1*$^{+/-}$ mice; however, others were unique (Figure 3A). Shared pathways included an enrichment of cancer signatures, such as AML and CML, as well as MAPK, JAK-STAT and cytokine signaling.

Biological processes and pathways enriched only in *Egr1*$^{+/-}$ mice included hematopoiesis, PIP_3 signaling, and Wnt/β-catenin signaling, as well as protein kinase and cytokine binding. We previously observed that concurrent haploinsufficiency of *Egr1* and *Apc*, a key regulator of the Wnt/β catenin pathway, co-operate in the pathogenesis of MDS and AML. In this regard, we observed an increase in expression of two major transcriptional mediators of Wnt/β catenin signaling: *Lef1* (*P*<0.0001) and *Tcf7* (*P*=0.01014). This is consistent with Wnt activation in MOL4070LTR *Egr1*$^{+/-}$ mice (Figure 3B). *LEF1* (Lymphoid Enhancer-Binding Factor 1) was also very highly expressed (2-15-fold) in del(5q) t-MN patients (University of Chicago series 1 and 2) and del(5q) AML patients in the TCGA database, and we recently identified significant changes in the expression of Wnt pathway genes, consistent with activation, in del(5q) patients.[6]

EGR1, a transcriptional regulator expressed at haploinsufficient levels in del(5q) t-MN patients, is involved in the homeostasis of hematopoietic stem cells, in myeloid differentiation, and in neoplastic transformation.[4,14] We hypothesize that dosage-dependent deregulation of hematopoietic transcriptional programs contributes to the development of myeloid disease; however, the affected *Egr1* targets genes are still not known. Interestingly, there was a significant enrichment of provirally-targeted genes that contain binding sites for at least one or more

© Ferrata Storti Foundation

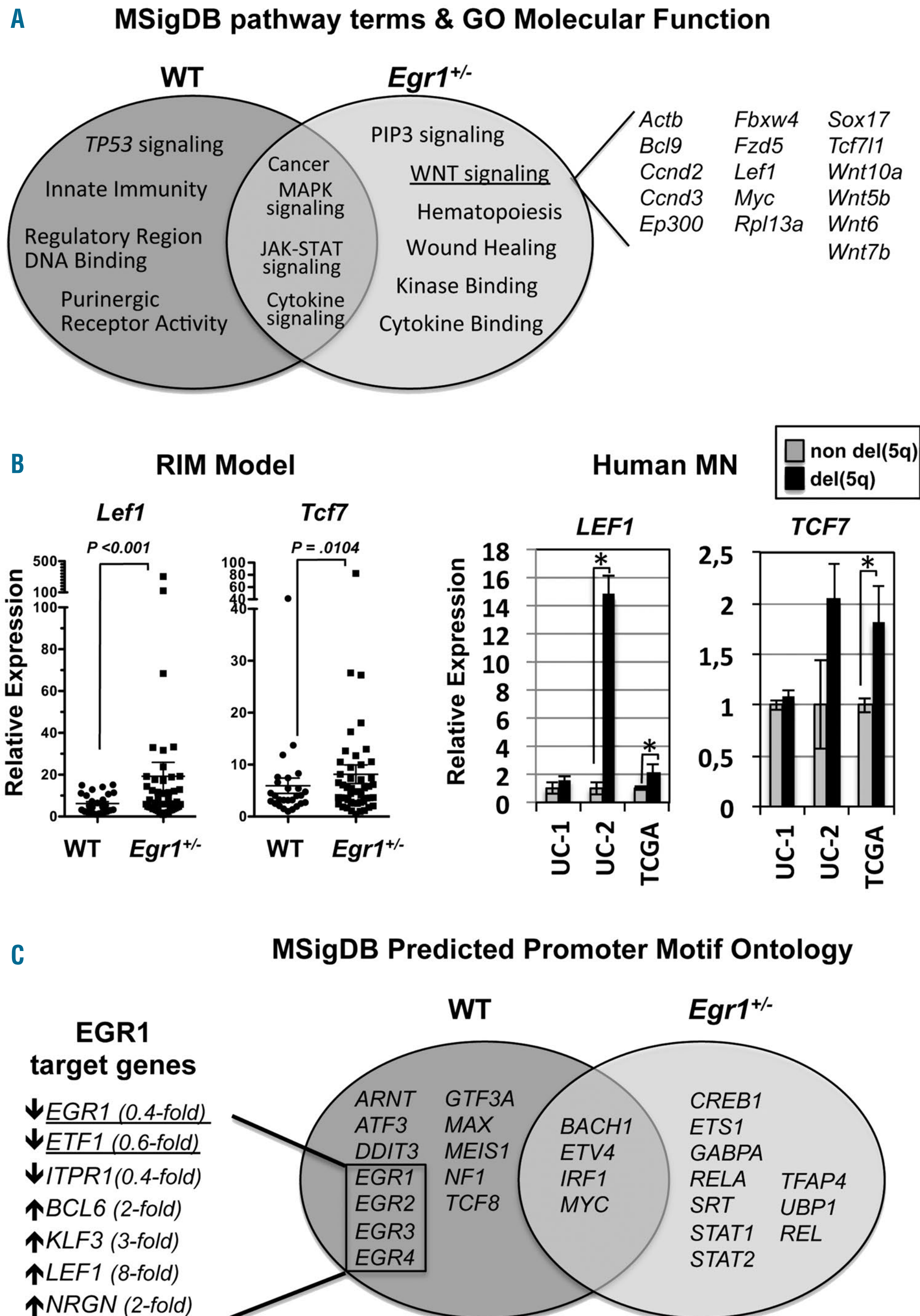


Figure 3. GREAT analysis of proviral insertions in *Egr1+/-* mice reveals deregulation of the Wnt signaling pathway. (A) GREAT was used to search for enrichment of proviral insertions near genes that target common signaling pathways, biological terms and shared transcription factor binding sites. Venn diagram illustrates the terms, assigned by GREAT analysis, which were enriched in WT, *Egr1+/-* or both mouse genotypes. Interestingly, an enrichment of proviral insertions near Wnt signaling pathway genes (right) was observed in *Egr1+/-* mice. (B) Expression of *Lef1* and *Tcf7* (major mediators of Wnt signaling) was assessed in spleen cells isolated from mice that developed myeloid disease after exposure to MOL4070LTR virus. *Lef1* and *Tcf7* are expressed at significantly higher levels in *Egr1+/-* mice compared to WT mice consistent with activation of the Wnt/β-catenin pathway in *Egr1+/-* mice. (C) The MSigDB predicted promoter motifs ontology set contains genes that share a transcription factor binding site in their promoters. Venn diagram illustrates that genes with promoter motifs that bind EGR transcription factors were specifically targeted in WT mice suggesting that a disruption of EGR-regulated genes, through proviral insertion, promotes myeloid disease. EGR1 target genes (confirmed by EGR1 ChIP analysis in K562 cells) with significant changes in gene expression of del(5q) *versus* non-del(5q) MNs is shown on the left. Deregulation of these genes, possibly through *EGR1* haploinsufficiency, may play an essential role in the malignant transformation process observed in patients with a del(5q). Of note, *ETF1* maps to 5q31.2 immediately distal to *EGR1* within the commonly deleted segment in MNs.

EGR family members (*EGR1, EGR2, EGR3, EGR4)* in WT, but not *Egr1*$^{+/-}$ mice (Figure 3C and *Online Supplementary Tables S3 and S4*). Of these 40 potential EGR target genes, we identified 7 genes that were confirmed *EGR1* targets by chromatin immunoprecipitation in K562 cells, and were differentially expressed in del(5q) *versus* non-del(5q) MNs. These include *EGR1, ETF1* and *ITPR1* [down-regulated in del(5q) neoplasms] and *BCL6, KLF3, LEF1* and *NRGN* [up-regulated in del(5q) neoplasms]. These genes warrant consideration as potential *EGR1* target genes that are deregulated during leukemogenesis in t-MN del(5q) patients. Notably, *ETF1* (Eukaryotic Translation Termination Factor 1) has previously been suggested to act as a haploinsufficient tumor suppressor protein in MDS/AML patients with a del(5q), possibly through production of potentially oncogenic, aberrant proteins.[15] In addition, *ETF1* (5q31) and *LEF1* (WNT target gene) expression may be deregulated in del(5q) patients due to *EGR1* haploinsufficiency and/or loss of one copy of *ETF1* and *APC* (5q22.2), respectively.

Patients with high-risk MDS, AML or t-MN with a del(5q) present with a very complex genetic profile and have a poor response to current therapies and a poor prognosis. In this study, we used our previously described mouse model of t-MN, with loss of the critical del(5q) haploinsufficient myeloid tumor suppressor gene, *EGR1*. Haploinsufficiency of *Egr1*, does not cause phenotypic abnormalities in mice under basal conditions; however, a reduction in Egr1 levels creates a sensitized background that co-operates with secondary mutations to give rise to myeloid malignancies. Moreover, our findings provide evidence that deregulation of genes at 5q31.1-5q31.2 (with *TIFAB, CXXC5* and *ETF1* being the most likely candidates) and activation of specific signaling pathways (with Wnt/β-catenin pathway the most likely aberrant pathway) are critical alterations that promote the development of MNs in co-operation with *EGR1* haploinsufficiency. Notably, our studies also identify potential *EGR1* target genes that may be deregulated during the progression towards t-MN, with *ETF1* and *LEF1* warranting further investigation.

Understanding the pathobiology of MN with a del(5q) has been challenging due to the large size of deletion and the ensuing haploinsufficient loss of multiple genes, and the additional genetic alterations that co-operate with 5q loss. This study is a step towards defining the critical pathways at play in MNs with a del(5q), which may inform the development of effective therapeutic strategies for this genetic subset of patients.

Angela Stoddart,[1] Zhijian Qian,[2] Anthony A. Fernald,[1] Rachel J. Bergerson,[3] Jianghong Wang,[1] Theodore Karrison,[4,5] John Anastasi,[5,6] Elizabeth T. Bartom,[7] Aaron L. Sarver,[3] Megan E. McNerney,[5,8] David A. Largaespada,[3] and Michelle M. Le Beau[1,5]

[1]Department of Medicine, University of Chicago, Chicago, IL; [2]Department of Medicine, and University of Illinois Cancer Center, University of Illinois at Chicago, Chicago, IL; [3]Department of Pediatrics, Masonic Cancer Center, University of Minnesota, Minneapolis, MN; [4]Department of Public Health Sciences, University of Chicago, Chicago, IL; [5]University of Chicago Medicine Comprehensive Cancer Center, Chicago, IL: [6]Department of Pathology, University of Chicago, Chicago, IL; [7]Department of Biochemistry and Molecular Genetics, Northwestern University Feinburg School of Medicine, Chicago, IL; and [8]Departments of Pathology and Pediatrics, and Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL, USA

Acknowledgments: we thank Dr. Linda Wolff for the generous gift of the MOL4070LTR virus. Her contribution merits co-authorship; however, her retirement from NIH precluded her review of this manuscript.

Funding: this work was supported by National Institutes of Health, National Cancer Institute (P01 CA40046 and RO1 CA190372) (M.M.L.B.), and the Integrated Microscopy and Cytometry and Antibody Technologies Core Facilities of the Comprehensive Cancer Center (P30 CA14599).

The online version of this letter has a Supplementary Appendix.

Correspondence: astoddar@bsd.uchicago.edu

doi:10.3324/haematol.2015.139527

Key words: myeloid neoplasms, del(5q), Wnt pathway, retroviral, insertional mutagenesis.

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. Blood. 2012;120(12):2454-2465.
2. Christiansen DH, Andersen MK, Pedersen-Bjergaard J. Mutations with loss of heterozygosity of p53 are common in therapy-related myelodysplasia and acute myeloid leukemia after exposure to alkylating agents and significantly associated with deletion or loss of 5q, a complex karyotype, and a poor prognosis. J Clin Oncol. 2001;19(5):1405-1413.
3. Walter MJ, Shen D, Shao J, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. Leukemia. 2013;27(6):1275-1282.
4. Stoddart A, Fernald AA, Wang J, et al. Haploinsufficiency of del(5q) genes, Egr1 and Apc, cooperate with Tp53 loss to induce acute myeloid leukemia in mice. Blood. 2014;123(7):1069-1078.
5. Joslin JM, Fernald AA, Tennant TR, et al. Haploinsufficiency of EGR1, a candidate gene in the del(5q), leads to the development of myeloid disorders. Blood. 2007;110(2):719-726.
6. Stoddart A, Nakitandwe J, Chen SC, Downing JR, Le Beau MM. Haploinsufficient loss of multiple 5q genes may fine-tune Wnt signaling in del(5q) therapy-related myeloid neoplasms. Blood. 2015;126(26):2899-2901.
7. Fernandez-Mercado M, Burns A, Pellagatti A, et al. Targeted resequencing analysis of 25 genes commonly mutated in myeloid disorders in del(5q) myelodysplastic syndromes. Haematologica. 2013;98(12):1856-1864.
8. Wolff L, Koller R, Hu X, Anver MR. A Moloney murine leukemia virus-based retrovirus with 4070A long terminal repeat sequences induces a high incidence of myeloid as well as lymphoid neoplasms. J Virol. 2003;77(8):4965-4971.
9. Sarver AL, Erdman J, Starr T, Largaespada DA, Silverstein KA. TAPDANCE: an automated tool to identify and annotate transposon insertion CISs and associations between CISs from next generation sequence data. BMC Bioinformatics. 2012;13:154.
10. Varney ME, Niederkorn M, Konno H, et al. Loss of Tifab, a del(5q) MDS gene, alters hematopoiesis through derepression of Toll-like receptor-TRAF6 signaling. J Exp Med. 2015;212(11):1967-1985.
11. Kuhnl A, Valk PJ, Sanders MA, et al. Downregulation of the Wnt inhibitor CXXC5 predicts a better prognosis in acute myeloid leukemia. Blood. 2015;125(19):2985-2994.
12. Treppendahl MB, Mollgard L, Hellstrom-Lindberg E, Cloos P, Gronbaek K. Downregulation but lack of promoter hypermethylation or somatic mutations of the potential tumor suppressor CXXC5 in MDS and AML with deletion 5q. Eur J Haematol. 2013;90(3):259-260.
13. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28(5):495-501.
14. Min IM, Pietramaggiori G, Kim FS, Passegue E, Stevenson KE, Wagers AJ. The transcription factor EGR1 controls both the proliferation and localization of hematopoietic stem cells. Cell Stem Cell. 2008;2(4):380-391.
15. Dubourg C, Toutain B, Helias C, et al. Evaluation of ETF1/eRF1, mapping to 5q31, as a candidate myeloid tumor suppressor gene. Cancer Genet Cytogenet. 2002;134(1):33-37.

© Ferrata Storti Foundation