Aberrant ERG expression associates with downregulation of miR-4638-5p and selected genomic alterations in a subset of diffuse large B-cell lymphoma

Shanxiang Zhang MD, PhD, Lin Wang MD, PhD, Liang Cheng, MD. Department of Pathology & Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN.

Correspondence:

Shanxiang Zhang, MD, PhD

Department of Pathology and Laboratory Medicine

Indiana University

350 West 11th Street, Room 5042

Indianapolis, IN 46202

USA

Email: sz5@iupui.edu

Tel: 001-317-491-6175

FAX: 001-317-491-6114

This is the author's manuscript of the article published in final edited form as:

Zhang, S., Wang, L., & Cheng, L. (2019). Aberrant ERG expression associates with downregulation of miR-4638-5p and selected genomic alterations in a subset of diffuse large B-cell lymphoma. Molecular Carcinogenesis, 58(10), 1846–1854. https://doi.org/10.1002/mc.23074

Shortened title: ERG involvement in diffuse large B-cell lymphoma.

Key words: ERG, diffuse large B-cell lymphoma, microRNA, genomic pathway, whole exome sequencing

List of abbreviation:

ERG+, aberrant ERG expression; AML, acute myeloid leukemia; T-ALL, T lymphoblastic leukemia; B-ALL, B lymphoblastic leukemia; DLBCL, diffuse large B-cell lymphoma; TMA, tissue microarrays; GCB, germinal center B-cell-like; non-GCB, non-germinal center B-cell-like; FFPE, formalin fixed paraffin embedded; FISH, fluorescence in situ hybridization; RT-PCR, reverse transcriptase polymerase chain reaction; WES, whole exome sequencing.

Funding: The study was supported by Indiana University Health Value Grant (VFR-415-Zhang) awarded to Shanxiang Zhang

Abstract

ERG, an oncoprotein in prostate carcinoma and Ewing's sarcoma is associated with poor prognosis in patients with acute myeloid leukemia and T lymphoblastic leukemia. However little is known about ERG in lymphoma. Here we studied ERG in diffuse large B-cell lymphoma (DLBCL) by immunohistochemistry, fluorescence in situ hybridization (FISH), genome-wide miRNA expression profiling, real-time reverse transcriptase polymerase chain reaction (RT-PCR) and whole exome sequencing (WES). Approximately 30% of de novo DLBCLs (37/118) expressed ERG (ERG+). ERG expression showed no significant correlation with DLBCL cell-of-origin classification, patient's age, gender, nodal or extranodal disease status, tumor expression of p53 or p63. There was no *ERG* rearrangement in ten randomly selected ERG+ DLBCLs by FISH. Forty-three miRNAs showed significant differential expression between ERG+ and ERG-negative DLBCLs. Downregulation of miR-4638-5p was confirmed by real-time RT-PCR. WES not only confirmed known gene mutations in DLBCLs but also revealed multiple novel gene mutations in *POLA1*, *E2F1*, *PSMD8*, *AXIN1*, *GAB2* and *GNB2L1*, which occur more frequently in ERG+ DLBCLs. In conclusion, our studies demonstrated aberrant ERG expression in a subset of DLBCL, which is associated with downregulation of miR-4638-5p. In comparison with ERG-negative DLBCL, ERG+ DLBCL more likely harbors mutations in genes important in cell cycle control, B-cell receptor-mediated signaling and degradation of beta-catenin. Further clinicopathological correlation and functional studies of ERG-related miRNAs and pathways may provide new insight into the pathogenesis of DLBCL and reveal novel targets for better management of patients with DLBCL.

Background

ERG (avian v-ets erythroblastosis virus E26 oncogene homolog), a member of ETS family transcription factor is physiologically expressed in endothelial cells and hematopoietic precursor cells. [1-3] In hematopoietic system, ERG expression is essential for normal hematopoiesis including megakaryopoiesis. [4]Aberrant ERG expression via ERG-involved translocation has been well documented in prostate carcinoma, Ewing's sarcoma and rare cases of acute myeloid leukemia (AML). [5-8]Aberrant ERG expression (ERG+) is associated with poor prognosis in a subgroup of cytogenetically normal AML and T lymphoblastic leukemia (T-ALL) while deletion of *ERG* locus is associated with a subset of B lymphoblastic leukemia (B-ALL) with good prognosis. [9-13] Targeted therapy against ERG either directly or indirectly to ERG-involved pathway has shown promising results in prostate carcinoma. [14-16]However little is known about ERG expression in mature hematopoietic cell neoplasm.

Diffuse large B-cell lymphoma (DLBCL) is one of the most common lymphomas with close to 50% of treatment failure in patients receiving current standard chemotherapy. [17]Novel treatment modality particularly targeted therapy based on specific pathways involved in DLBCL are currently being actively pursued. Following our initial observation of ERG expression in an index DLBCL case, we decided to study systematically ERG expression in DLBCL.

miRNAs are small non-coding RNAs that regulate gene expression at the post-transcription level. They are involved in many important biological processes such as development, differentiation, apoptosis and proliferation. Systematic studies such as miRNA library profiling and miRNA sequencing have identified many miRNAs, which are involved in tumor initiation, progression and are used as markers for cancer diagnosis, prognosis stratification and targeted

therapy. [18]Though there are several miRNA studies in DLBCLs, there is no reported miRNA study about ERG+ DLBCLs.[19-23]

In the current study, we first studied ERG expression in normal hematopoietic tissue and then in DLBCL. miRNA profiling was then performed in 16 ERG+ and 23 ERG- DLBCLs. Furthermore, whole exome sequencing data from 5 ERG+ and 13 ERG- DLBCLs were compared. We established ERG aberrant expression in a subset of DLBCLs, which is associated with differential expression of a group of miRNAs. The genetic mutation and its associated pathways that may are more commonly seen in ERG+ DLBCLs were identified.

Materials and methods

Clinical cases and immunohistochemistry

Tissue microarrays (TMAs) were constructed from our archival cases of *de novo* DLBCL diagnosed during 2000-2014 following approval by the Institutional Review Board of our Institute. Clinical information was collected from the medical records.

TMAs for DLBCLs were stained with antibodies against CD20, BCL-2, CD10, BCL-6, and MUM-1. Prediluted, ready-to-use antibodies (CD20, GA604; CD10, GA648; BCL2, IR614; BCL6, GA625; MUM-1, GA644; ERG, IR659; CD31, GA610 from Dako, Santa Clara, CA; LMO2, 370R and HGAL, 375M from Cell Marque, St. Louis, MO) were used with a Dako Autostainer Plus instrument following the manufacturer's protocols (www.dako.com). DLBCL cases were classified as either germinal center B-cell-like (GCB) or non-germinal center B-cell-like (non-GCB) subgroup based on Han's algorithm. [24]To confirm the ERG expression identified in TMAs, tissue sections from 10 randomly selected ERG+ DLBCLs based on the TMA study were directly stained for ERG.

A DLBCL is considered ERG+ when at least one of the two duplicate cores (1.0 mm in diameter) in the TMAs were positive for CD20 and showed moderate to strong nuclear expression of ERG in over 50% of tumor cells. The background weak scattered staining of ERG in reactive germinal centers from tonsils present in each TMA was defined as negative. Accordingly, a total of 118 DLBCL individual cases were successfully determined for its ERG expression status. In addition, tissue sections from another eight DLBCL cases were directly stained for ERG. A total of126 DLBCL cases were included in Kaplain-Meier survival analysis. Among the 118 DLBCL in TMAs, 94 of them were successfully stained for CD10, BCL-6 and MUM-1 and subsequently classified as either GCB or non-GCB DLBCLs. These 94 cases were included in the clinicopathologic correlation study.

***ERG*-rearrangement study by fluorescence in situ hybridization (FISH)**

FISH study was performed as described before. [25]Briefly, 4-micron tumor tissue sections from 10 randomly selected ERG+ DLBCLs were mounted on positively- charged glass slides. An accompanying H & E slide was reviewed to determine tumor location. Slides were baked at 65ºC +/- 5ºC for 1 hour, de-paraffinized in xylene at room temperature and washed in 100% ethanol and air-dried. Pretreatment included exposure to 0.2 N HCl, followed by 1M sodium thiocyanate and pepsin at 37ºC to digest proteins in the tissue sample. Following pepsin treatment, slides were viewed with phase microscopy to ensure adequate digestion of the tumor tissue. Denaturation, ethanol dehydration, and hybridization with a probe cocktail including BAC clones RP11-476D17-gold (3' ERG) and RP11-95121-green (5' ERG) (Empire Genomics, Buffalo, NY, USA) diluted 1:25 in tDenHyb2 (Insitus, Albuquerque, NM, USA). Cells were counterstained with DAPI and observed under a Leica DMRXA2 fluorescence microscope. Two hundred cells were analyzed by two readers (100 cells apiece) for each of the three probes.

Scoring criteria included the selection of single nuclei with representation of at least one signal for each color/probe/ nucleus. Tissue sections from prostate carcinoma with known ERG rearrangement were used as the positive control.

miRNA microarray analysis

Total RNA was extracted from formalin fixed paraffin embedded (FFPE) tissue using RecoverALL™ Total Nucleic Acid Isolation kit (Invitrogen, Carlsbad, California) following the manufacturer's instructions. RNA quality and quantity were measured using NanoDrop spectrophotometer. 1.0 μg total RNA samples were labeled using T4 Rnl2tr-K227Q kit (New England BioLabs, Ipswich, MA) according to the manufacturer's instructions. Human miRNA Array (miRHuman_21, LC Sciences, Houston, Texas), which contained the entire 1881 miRNAs annotated in Sanger miRBase Release 21 were used. Hybridization was performed overnight on a μParaflo microfluidic chip using a micro-circulation pump (Atactic Technologies, Houston, TX). [26]Fluorescence images were collected using a laser scanner (GenePix 4000B, Molecular Device, San Jose, CA) and digitized using Array-Pro image analysis software (Media Cybernetics, Rockville, MD). Data were analyzed by first subtracting the background and then normalizing the signals using a LOWESS filter (Locally-weighted Regression).[27]

Quantitative real-time reverse transcriptase polymerase chain reaction (RT-PCR) assay

Total RNA including miRNAs were extracted from FFPE tissue samples using miRNeasy FFPE kit (QIAGEN, Hilden, Germany). The quantity of total RNA was determined by Qubit RNA HS Assay kit (Thermo Fisher Scientific, Waltham, MA). Reverse transcription was performed using miScript II RT kit (QIAGEN) with HiFlex buffer. Real-time qPCR was performed using miScript SYBR Green kit (QIAGEN) according to the manufacturer's protocol on a QuantStudio 5 Real-time PCR system (Thermo Fisher Scientific). Universal forward primer included in the kit

and primers included in the miScript primer assay (QIAGEN) were used to quantify the expression of miRNAs. Two housekeeping small RNAs, RNU6B and SNORD68, were also included as internal controls. Relative expression (RE) of miRNAs was measured as $RE = 2^{-\Delta\Delta Ct}$.

Whole exome sequencing, variant calling and pathway prediction

The whole exome sequencing (WES) of 5 ERG+ and 13 ERG- DLBCLs was performed at the Technology Center for Genomics & Bioinformatics at University of California at Los Angeles. Whole-exome DNA was captured from total genomic DNA using the SeqCap EZ System from NimbleGen according to the manufacturer's instructions. Briefly, genomic DNA extracted from FFPE tumor tissue was sheared, size selected to roughly 200-250 base pairs, and the ends were repaired and ligated to specific adapters and multiplexing indexes. Fragments were then incubated with SeqCap biotinylated DNA baits followed by the LM-PCR, and the RNA-DNA hybrids were purified using streptavidin-coated magnetic beads. The RNA baits were then digested to release the targeted DNA fragments, followed by a brief amplification of 15 or less PCR cycles. The libraries were then sequenced on the HiSeq 3000 platform from Illumina, using 150-bp pair-ended reads. The sequence data were aligned to the GRCh37 human reference genome using BWA v0.7.7-r411. PCR duplicates were marked using MarkDuplicates program in Picard-tools-1.115 tool set. Samtools was used to call the SNVs (single nucleotide variants) and small INDELs (insertions and deletions). Varscan2 was used to call the somatic SNVs. All variants were annotated using the Annovar program. The top 203 genes showing greater than 0.5 difference in the frequency of the nonsynonymous SNV in genes between ERG+ and ERG- DLBCLs were used for pathway prediction. StrandNGS software and the Wikipathways database were used for pathway analysis.

Statistical analysis

SPSS Statistics24 software (IBM Corporations, New York, NY) was used to generate Kaplan-Meier survival curve. Pearson's chi-squared test was used for comparison with a p value of less than 0.05 being considered as significant.

Results

ERG expression in normal hematopoietic tissue

To characterize ERG expression in normal hematopoietic tissue, two reactive lymph nodes, two tonsils and two spleens were stained for ERG. Reactive follicles with well-defined germinal centers from a reactive lymph node were shown in Fig.1A. As shown in Fig.1B and 1C, ERG was expressed as a nuclear protein, predominantly in cells outside of the germinal centers. The weak nuclear and cytoplasmic stain noted in germinal center cells may represent non-specific background staining. The ERG expression within reactive lymphoid tissues is similar to that of BCL-2 (Fig.1D). There is similar ERG staining pattern in reactive tonsils and spleens (data not shown).

ERG expression in a subset of diffuse large B-cell lymphoma

As we initially made an incidental observation of high ERG expression in a bladder DLBCL (Fig. 2A-D), we decided to study ERG expression by staining TMAs containing a series of de novo DLBCLs (Fig. 2E-F). ERG is expressed in approximately one-third of DLBCLs. Tissue sections from ten randomly selected ERG+ DLBCLs were also directly stained for ERG, which showed complete correlation with ERG expression observed in TMAs (result not shown).

To investigate if ERG expression may be due to *ERG* rearrangement as seen in approximately 50% of prostate carcinomas, FISH using a break-apart *ERG* probe was performed on these 10 ERG+ DLBCL cases. Prostate carcinoma case with confirmed ERG translocation was used as the positive control in the FISH study. None of DLBCLs tested showed *ERG* rearrangement

(data not shown), indicating that other mechanism may contribute to ERG overexpression in ERG+ DLBCLs.

To correlate ERG expression with the clinicopathological features of DLBCL, TMAs were stained for CD20, CD10, BCL-6, MUM-1, p53, p63 as well as ERG. A total of 94 de novo DLBCLs were successfully stained and classified as GCB or non-GCB subgroups based on Hans' algorithm. [24]As shown in Table 1, ERG moderate to strong expression was seen in both GCB and non-GCB DLBCL cases with no significant difference (p=0.603) and showed no significant correction with patient's age (p=0.688), gender (p=0.826), nodal or extranodal disease status (p=0.903), tumor expression of p53 (p=0.672) or p63 (p=0.640). A total of 126 patients with ERG determined status were included in the overall survival analysis. ERG+ DLBCL showed no difference in overall survival compared with ERG- DLBCL (p=0.940). However, in patients with treatment of rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone (R-CHOP, n=54), ERG+ DLBCL patients started to show a trend toward inferior overall survival (p=0.192) though the patient cohort became much smaller (Fig. 3).

miRNA signatures of ERG+ vs ERG- DLBCLs and downregulation of miR-4638-5p in ERG+ DLBCL

To study the miRNAs expression in ERG+ DLBCLs, we performed whole-genome miRNA library profiling on 16 ERG+ and 23 ERG- DLBCLs using Human miRNA Array (miRHuman_21, LC Sciences, Houston, Texas), which contained the entire 1881 miRNAs annotated in Sanger miRBase Release 21. To investigate if the miRNA array screening was working adequately, we compared first the miRNA expression between GCB and non-GCB DLBCLs. In our current study, the following 10 miRNAs (miR-221-3p, miR-155-5p, miR-4267, miR-148a-3p, miR-151a-5p, miR-222-3p, miR-151-b, miR-615-5p, miR-138-5p, and miR-3197)

showed most significant difference in the expression level between GCB and non-GCB DLBCLs. Six of them (miR-221-3p, miR-155-5p, miR-151a-5p, miR-222-3p, miR-151-b, and miR-138-5p) have been reported in the previous studies, which confirmed the validity of our current assay. [28] There is no available information about the remaining four miRNAs in DLBCLs. Next we compared the ERG+ versus ERG- DLBCLs. We identified forty-three miRNAs, which showed significantly different expression levels ($p<0.05$) (Fig. 4A). Twenty-six miRNAs were downregulated and 17 upregulated in ERG+ DLBCLs compared with ERG- DLBCLs. Four miRNAs miR-4638-5p, miR-1273c, miR-6780a-5p and miR-203a-3p showed more than 2-fold reduction in ERG+ DLBCL and had relatively high signals in the array study. miR-4638-5p was the only one with mean signal > 500 in both ERG+ and ERG- DLBCLs, which was the recommended cutoff by the manufacturer. To confirm the above findings, we performed quantitative real-time RT-PCR analysis for the above four miRNAs in the same 39 samples used for miRNA array analysis. Expression of miR-4638-5p was significantly lower and there was a trend of downregulation of miR-6780a-5p in ERG+ DLBCLs. There was essentially no difference for the expression of miR-1273c and miR-203a-3p between ERG+ vs ERG- DLBCLs (Fig. 4B).

Identification of genes with genomic alterations in ERG+ DLBCL

To gain an insight on the potential pathways associated with ERG+ DLBCLs, we analyzed WES data obtained from 5 ERG+ and 13 ERG- DLBCLs. To validate and compare our WES sequencing and analysis with previously published ones in DLBCLs, we analyzed the genomic alterations of a group of 15 genes known to be important in DLBCLs. As shown in Table 2, our study showed significant overlap with the previous studies. [29, 30]As our study did not include any normal matched DNAs for normal individual variant control, the much higher alteration

frequency seen in genes *TNFRSF14*, *TP53*, *PCLO* may reflect the normal individual variation that was not completely filtered during our analysis. It may also reflect random sample bias introduced by our relative small sample size. To mitigate this effect and identify genes that show difference in mutation frequencies between ERG+ and ERG- DLBCLs, we focused on 203 genes, which showed more than 0.5 in genomic alteration frequency between ERG+ and ERG- DLBCLs (Table S1). To identify the genomic pathways, which may be significantly impacted by alterations in these 203 genes, we calculated the p-value, which reflects how significant the pathway change is based on a gene's variants by combining what mutation a gene has and how significant the mutated gene is in the pathway. The predicted pathways affected by these 203 genes in ERG+ and ERG- DLBCLs were summarized in Table S2. The detailed information about the underlying genes was summarized in Table S3 for ERG+ DLBCLs and Table S4 for ERG- DLBCLs. In addition, we reported for the first time mutations in six genes *POLA1*, *E2F1*, *PSMD8*, *AXIN1*, *GAB2* and *GNB2L1*, which occur more frequently in ERG+ DLBCLs. Their mutation frequencies vary from 0.6 to 0.8 in ERG+ DLBCLs and 0 to 0.23 in ERG- DLBCLs. The types of mutations include silent, missense, nonsense and frameshift (Table S5). Five of 13 pathways were significantly affected by ERG expression in DLBCLs and their involved genes were summarized in Table 3. The Table S6 summarized the detailed findings for the involved genes.

Discussion

ERG is a well-established oncogene in prostate carcinoma. ERG overexpression in a subset of AML and T-ALL impacts negatively towards patient's prognosis, possibly by inducing chemoresistance in leukemia cells.[31]Recently ERG dysregulation at both transcriptional and translational levels is documented in a subset of B lymphoblastic leukemia as well. [12]ERG is

emerging as an important prognostic marker as well as a novel therapeutic target. However, there has been no systematic study about ERG in lymphomas. Here we reported ERG expression in a subset of DLBCL.

Our study showed moderate to strong ERG expression in approximately 30% of de novo DLBCLs including both GCB and non-GCB groups without significant difference. The ERG overexpression in ERG+ DLBCLs is unlikely due to *ERG* gene rearrangement as 10 randomly selected ERG+ DLBCL cases were all negative for *ERG* rearrangement by FISH study. Though our FISH study for *ERG* rearrangement is small, it nevertheless suggests that other mechanisms may be involved as seen in most ERG+ AML and T-ALL. [12, 32-34]Though ERG expression appears not to correlate with any clinicopathological features we have studied, it does point to a trend that patients with ERG+ DLBCL might behave more poorly compared to those with ERG- DLBCLs when receiving current standard chemotherapy (rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone; R-CHOP), especially in the first three years after the diagnosis. However, the cohort is too small and we did not have the follow-up on many of our patients as they were treated in the outside facilities after the diagnosis. Additional studies with much larger cohort of patients are needed to investigate if ERG expression may confer a worse prognosis.

As miRNA's role in the regulation of gene expression at the post-transcription level is well established, we decided to study miRNAs expression through genome-wide miRNA profiling. Of the 43 miRNAs showing differential expression in ERG+ vs ERG-DLBCLs, only miR339-3p and miR-320b have been previously reported to be involved in DLBCL. [19, 35] miR-106b-5p, mir-5787, miR-192-5p, miR106-5p, miR-378-3p have been however implicated in cellular growth, tumorigenesis, tumor progression, metastasis and sensitivity to chemotherapy in other

types of tumor. [36-43]Interestingly none of the 43 miRNAs were predicted to directly regulate ERG expression based on miRDB curation (www.mirdb.org), suggesting either their indirect role in the regulation of ERG overexpression or their expression is impacted by aberrant ERG expression in DLBCLs.

While miRNA array expression profiling is a very sensitive and powerful technology in identifying the differentially expressed miRNAs among different study groups, it is important to validate these findings by another methodology, such as quantitative real-time RT-PCR. Among the four miRNAs selected for RT-PCR study, the significant downregulation of miR-4638-5p in ERG+ DLBCLs was confirmed, while a trend of downregulation of miR-6780a-5p in ERG+ DLBCLs was observed by real-time RT-PCR. There were essentially no differences for the expressions of miR-1273c and miR-203a-3p between ERG+ vs ERG- DLBCLs. While the result was not entirely surprising as miR-4638-5p was the only miRNA with mean signal > 500 in both ERG+ and ERG- DLBCLs in miRNA array study, which was the recommended cutoff by the manufacturer, it emphasizes the importance of cross validation by different methodologies in miRNA study.

In prostate carcinoma, downregulation of miR-4638-5p is associated with castration resistance though its direct repression of protein kinase D-interacting substrate of 220 KDa (kidins200) and PI3K/AKT pathway activity. [44]In resting and stimulated B cells, kidins200 interacts with B-cell antigen receptor (BCR) and positively regulates pre –BCR and BCR functioning. [45]In T cells kidins200 associates with B-Raf and the T-cell antigen receptor to promote sustained Erk signaling. [46]Whether similar mechanism may be involved in ERG+ DLBCL warrants further studies.

In summary, we report for the first time that ERG is expressed in up to one-third of DLBCL, which raises the caution in its use as a diagnostic marker for cutaneous leukemia involvement. [47] Through genome-wide miRNA study, we identified and confirmed that miR-4638-5p was significantly associated with aberrant ERG expression in DLBCLs. WES and subsequent pathway analysis reveal several cellular pathways significantly associated with genomic alterations preferentially seen in ERG+ DLBCLs. Further studies will not only decide whether ERG may be used as a novel prognostic and/or therapeutic marker in DLBCL, but also help us understand the general oncogenic role of ERG in tumorigenesis.

Authors' contributions: SZ designed, collected and analyzed data and wrote the manuscript; LC performed and analyzed FISH study; LW performed RNA extraction and helped with data collection. We all contributed to writing and approved the manuscript.

Acknowledgements: None

Declarations

Ethics approval and consent to participate: The study was approved by the Institutional Review Board of Indiana University.

Consent for publication: Not applicable

Availability of data and material: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests: No conflict of interest from all participating authors.

References

1. Baltzinger, M., A.M. Mager-Heckel, and P. Remy, *Xl erg: expression pattern and overexpression during development plead for a role in endothelial cell differentiation.* Developmental dynamics : an official publication of the American Association of Anatomists, 1999. **216**(4-5): p. 420-33.
2. Loughran, S.J., et al., *The transcription factor Erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells.* Nature immunology, 2008. **9**(7): p. 810-9.
3. McLaughlin, F., et al., *Combined genomic and antisense analysis reveals that the transcription factor Erg is implicated in endothelial cell differentiation.* Blood, 2001. **98**(12): p. 3332-9.
4. Salek-Ardakani, S., et al., *ERG is a megakaryocytic oncogene.* Cancer research, 2009. **69**(11): p. 4665-73.
5. Tomlins, S.A., et al., *Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer.* Science, 2005. **310**(5748): p. 644-8.
6. Ida, K., et al., *EWS-FLI-1 and EWS-ERG chimeric mRNAs in Ewing's sarcoma and primitive neuroectodermal tumor.* International journal of cancer, 1995. **63**(4): p. 500-4.
7. Sorensen, P.H., et al., *A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG.* Nature genetics, 1994. **6**(2): p. 146-51.
8. Ichikawa, H., et al., *An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation.* Cancer research, 1994. **54**(11): p. 2865-8.
9. Baldus, C.D., et al., *High expression of the ETS transcription factor ERG predicts adverse outcome in acute T-lymphoblastic leukemia in adults.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2006. **24**(29): p. 4714-20.
10. Baldus, C.D., et al., *Low ERG and BAALC expression identifies a new subgroup of adult acute T-lymphoblastic leukemia with a highly favorable outcome.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2007. **25**(24): p. 3739-45.
11. Marcucci, G., et al., *High expression levels of the ETS-related gene, ERG, predict adverse outcome and improve molecular risk-based classification of cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B Study.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2007. **25**(22): p. 3337-43.
12. Zhang, J., et al., *Deregulation of DUX4 and ERG in acute lymphoblastic leukemia.* Nature genetics, 2016. **48**(12): p. 1481-1489.
13. Clappier, E., et al., *An intragenic ERG deletion is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions.* Leukemia, 2014. **28**(1): p. 70-7.
14. Chatterjee, P., et al., *PARP inhibition sensitizes to low dose-rate radiation TMPRSS2-ERG fusion gene-expressing and PTEN-deficient prostate cancer cells.* PloS one, 2013. **8**(4): p. e60408.
15. Rahim, S., et al., *YK-4-279 inhibits ERG and ETV1 mediated prostate cancer cell invasion.* PloS one, 2011. **6**(4): p. e19343.
16. Nhili, R., et al., *Targeting the DNA-binding activity of the human ERG transcription factor using new heterocyclic dithiophene diamidines.* Nucleic acids research, 2013. **41**(1): p. 125-38.
17. Roschewski, M., L.M. Staudt, and W.H. Wilson, *Diffuse large B-cell lymphoma-treatment approaches in the molecular era.* Nature reviews. Clinical oncology, 2014. **11**(1): p. 12-23.
18. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers.* Nature reviews. Cancer, 2006. **6**(11): p. 857-66.
19. Lim, E.L., et al., *Comprehensive miRNA sequence analysis reveals survival differences in diffuse large B-cell lymphoma patients.* Genome biology, 2015. **16**: p. 18.
20. Shepshelovich, D., et al., *MicroRNA signature is indicative of long term prognosis in diffuse large B-cell lymphoma.* Leukemia research, 2015. **39**(6): p. 632-7.
21. Montes-Moreno, S., et al., *miRNA expression in diffuse large B-cell lymphoma treated with chemoimmunotherapy.* Blood, 2011. **118**(4): p. 1034-40.

22. Lawrie, C.H., et al., *MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma.* International journal of cancer, 2007. **121**(5): p. 1156-61.
23. Roehle, A., et al., *MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas.* British journal of haematology, 2008. **142**(5): p. 732-44.
24. Hans, C.P., et al., *Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray.* Blood, 2004. **103**(1): p. 275-82.
25. Williamson, S.R., et al., *ERG-TMPRSS2 rearrangement is shared by concurrent prostatic adenocarcinoma and prostatic small cell carcinoma and absent in small cell carcinoma of the urinary bladder: evidence supporting monoclonal origin.* Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc, 2011. **24**(8): p. 1120-7.
26. Gao, X., E. Gulari, and X. Zhou, *In situ synthesis of oligonucleotide microarrays.* Biopolymers, 2004. **73**(5): p. 579-96.
27. Bolstad, B.M., et al., *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.* Bioinformatics, 2003. **19**(2): p. 185-93.
28. Larrabeiti-Etxebarria, A., et al., *Systematic Review of the Potential of MicroRNAs in Diffuse Large B Cell Lymphoma.* Cancers, 2019. **11**(2).
29. Lohr, J.G., et al., *Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(10): p. 3879-84.
30. Zhang, J., et al., *Genetic heterogeneity of diffuse large B-cell lymphoma.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(4): p. 1398-403.
31. Mochmann, L.H., et al., *ERG induces a mesenchymal-like state associated with chemoresistance in leukemia cells.* Oncotarget, 2014. **5**(2): p. 351-62.
32. Bock, J., et al., *ERG transcriptional networks in primary acute leukemia cells implicate a role for ERG in deregulated kinase signaling.* PloS one, 2013. **8**(1): p. e52872.
33. Bohne, A., et al., *Epigenetic control of differential expression of specific ERG isoforms in acute T-lymphoblastic leukemia.* Leukemia research, 2009. **33**(6): p. 817-22.
34. Haferlach, C., et al., *Gene expression of BAALC, CDKN1B, ERG, and MN1 adds independent prognostic information to cytogenetics and molecular mutations in adult acute myeloid leukemia.* Genes, chromosomes & cancer, 2012. **51**(3): p. 257-65.
35. Wang, J., et al., *AKT Hyperactivation and the Potential of AKT-Targeted Therapy in Diffuse Large B-Cell Lymphoma.* The American journal of pathology, 2017. **187**(8): p. 1700-1716.
36. Li, H., et al., *Clinical and biological significance of miR-378a-3p and miR-378a-5p in colorectal cancer.* European journal of cancer, 2014. **50**(6): p. 1207-21.
37. Liu, F., et al., *MicroRNA-106b-5p boosts glioma tumorigensis by targeting multiple tumor suppressor genes.* Oncogene, 2014. **33**(40): p. 4813-22.
38. Lu, J., et al., *miR-106b-5p promotes renal cell carcinoma aggressiveness and stem-cell-like phenotype by activating Wnt/beta-catenin signalling.* Oncotarget, 2017. **8**(13): p. 21461-21471.
39. Xiang, W., et al., *miR-106b-5p targets tumor suppressor gene SETD2 to inactive its function in clear cell renal cell carcinoma.* Oncotarget, 2015. **6**(6): p. 4066-79.
40. Yan-Chun, L., et al., *MicroRNA-192-5p Promote the Proliferation and Metastasis of Hepatocellular Carcinoma Cell by Targeting SEMA3A.* Applied immunohistochemistry & molecular morphology : AIMM, 2015.
41. Yoo, H., et al., *The hsa-miR-5787 represses cellular growth by targeting eukaryotic translation initiation factor 5 (eIF5) in fibroblasts.* Biochemical and biophysical research communications, 2011. **415**(4): p. 567-72.

42. Yu, S., et al., *MicroRNA-106b-5p regulates cisplatin chemosensitivity by targeting polycystic kidney disease-2 in non-small-cell lung cancer.* Anti-cancer drugs, 2017. **28**(8): p. 852-860.
43. Garbicz, F., et al., *Increased expression of the microRNA 106b~25 cluster and its host gene MCM7 in corticotroph pituitary adenomas is associated with tumor invasion and Crooke's cell morphology.* Pituitary, 2017. **20**(4): p. 450-463.
44. Wang, Y., et al., *MiR-4638-5p inhibits castration resistance of prostate cancer through repressing Kidins220 expression and PI3K/AKT pathway activity.* Oncotarget, 2016. **7**(30): p. 47444-47464.
45. Fiala, G.J., et al., *Kidins220/ARMS binds to the B cell antigen receptor and regulates B cell development and activation.* The Journal of experimental medicine, 2015. **212**(10): p. 1693-708.
46. Deswal, S., et al., *Kidins220/ARMS associates with B-Raf and the TCR, promoting sustained Erk signaling in T cells.* Journal of immunology, 2013. **190**(5): p. 1927-35.
47. Xu, B., et al., *ERG Is a Useful Immunohistochemical Marker to Distinguish Leukemia Cutis From Nonneoplastic Leukocytic Infiltrates in the Skin.* The American Journal of dermatopathology, 2016. **38**(9): p. 672-7.

Legends of figures:

Figure 1. ERG expression in reactive lymph node. A, hemotoxylin-eosin stained lymph node section showing reactive follicles with well-defined germinal centers, 100X. B to D-immunohistochemistry. B, C-ERG expression, B-100X, C-400X. D, BCL-2 expression, 200X.

Figure 2. ERG expression in a subset of diffuse large B-cell lymphoma. A-D tissue section (x 100): A- H&E, B-CD20, C-BCL-6, D-ERG; E-F TMA section (x 40): E-CD20, F-ERG.

Figure 3. Patients with ERG+ DLBCL associated with a tendency for inferior survival when treated with R-CHOP. Kaplain-Meier survival analysis.

Figure 4. A Differential expression of 43 miRNAs in ERG+ vs ERG- DLBCLs identified by miRNA library profiling; B Downregulation of miR-4638-5p in ERG+ vs ERG- DLBCLs measured by real-time RT-PCR.