博士論文

Chromatin Modifications in B-Cell Acute Lymphoblastic Leukemia Carrying DUX4 Fusions

（ＤＵＸ４融合遺伝子陽性Ｂ細胞性急性リンパ性白血病におけるクロマチン修飾の解析）

田中　庸介

Chromatin Modifications in B-Cell Acute Lymphoblastic Leukemia Carrying DUX4 Fusions

（ＤＵＸ４融合遺伝子陽性Ｂ細胞性急性リンパ性白血病における クロマチン修飾の解析）

分子細胞生物学専攻　がん細胞情報学分野

指導教員　　間野 博行　教授

田中 庸介

Contents

Abstract

DUX4 fusions are novel oncogenes encoding truncated forms of the transcription factor

DUX4. They are preferentially found in B-cell acute lymphoblastic leukemia (B-ALL)

in children and young adults, shaping a new clinical entity. Although insertion of the

DUX4 gene into the *IGH* locus is implicated as the leading event, the mechanism of

leukemogenesis is not completely understood. I thus explored the transcriptional

regulation and chromatin modification by *DUX4* fusions in B-ALL. I revealed that the

genomic binding sites and the binding motif of DUX4-IGH were similar to those of the

wild-type (WT) DUX4, and DUX4-IGH holds histone modification ability. *DUX4-IGH*

knockdown in NALM6 cells indicated that the transcription of genes characteristic to B-

ALL carrying *DUX4* fusions were positively regulated by DUX4-IGH. The

transcriptional activity of DUX4-IGH regarding WT DUX4 target genes was severely

attenuated compared with that of WT DUX4. Through ATAC-seq combined with RNA

seq, I identified *BCL2* as a candidate gene for tumorigenesis driven by DUX4-IGH.

DUX4-IGH knockdown resulted in the increased expression of genes associated with

hematopoietic cell differentiation, suggesting that these genes were negatively regulated

by DUX4-IGH, that was further supported by ATAC-seq. These results provide new

insight into the mechanisms of transcriptional perturbation in B-ALL carrying *DUX4*

fusions.

Introduction

Acute lymphoblastic leukemia (ALL) is the most common childhood cancer, and B-cell precursor ALL (B-ALL) comprises more than 80 % of ALL[1]. Through the cytogenetic investigation for over half a century, most of the ALL cases have been identified to harbor specific oncogenic genomic alterations, mainly aneuploidy and chromosomal rearrangements creating chimeric fusion genes, and each genetic alterations were proven to define the characteristics of each ALL subtype[2]. Thus, identification of genetic alteration in ALL potentiates the better classification, risk stratification, and targeted therapy. The advent of next generation sequencing technology enabled us to identify novel oncogenes which had not been captured by the conventional method[3] such as karyotype analysis, and further led to the more subdivided classification of ALL (Figure 1). In the context of increasing novel oncogenes, functional analysis of these newly identified oncogenes is imperative, given that understanding of the oncogenic pathway can lead to the development of novel therapeutic intervention[4].
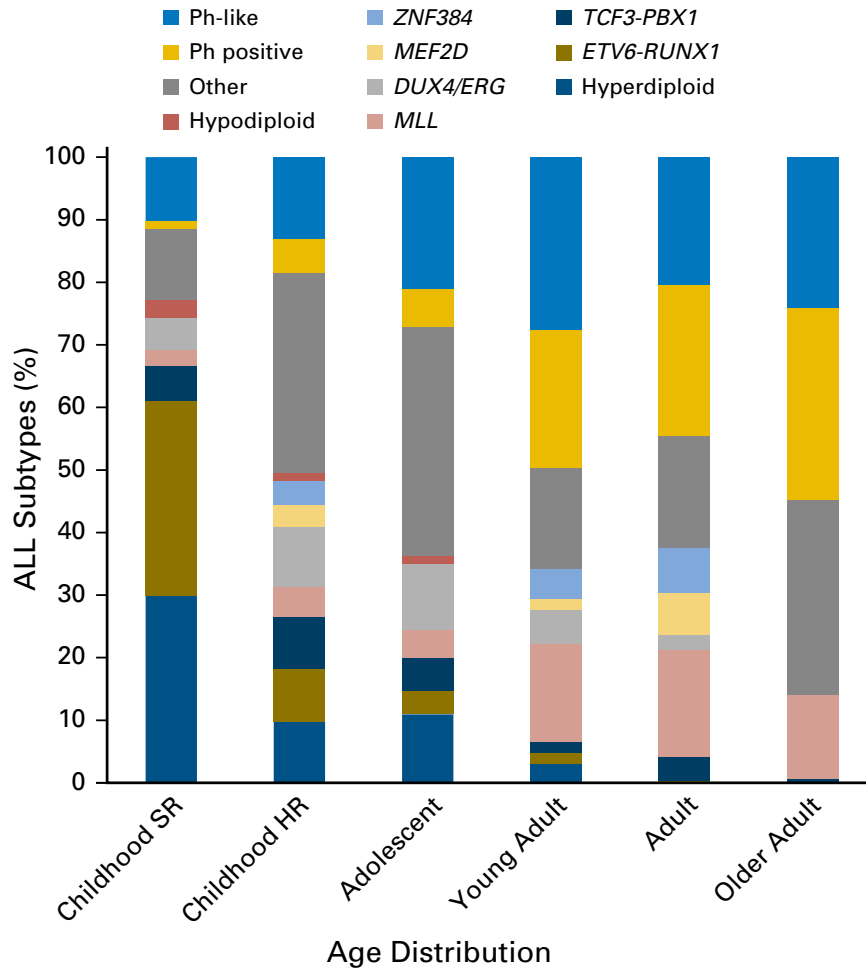
Figure 1. The distribution of genetic alterations in B-ALL. The distribution of the subtype-defining genetic alteration of B-ALL is age group specific[2].

Our group previously identified novel oncogenes, *DUX4* fusions, *ZNF384* fusions, and *MEF2D* fusions, through RNA-sequencing using the samples of adolescent and young adult patients who suffered from B-ALL without known genetic alterations[5]. Previous expression data analysis of B-ALL in 2010 had already identified a cluster of unique

gene expression pattern[6] that was characterized by the outlier expression genes

including *AGAP1, CCNJ, CHST2/7, CLEC12A/B, PTPRM,* and genetic deletion of

ERG, however, the causative genetic alteration had not been elucidated. Our group and

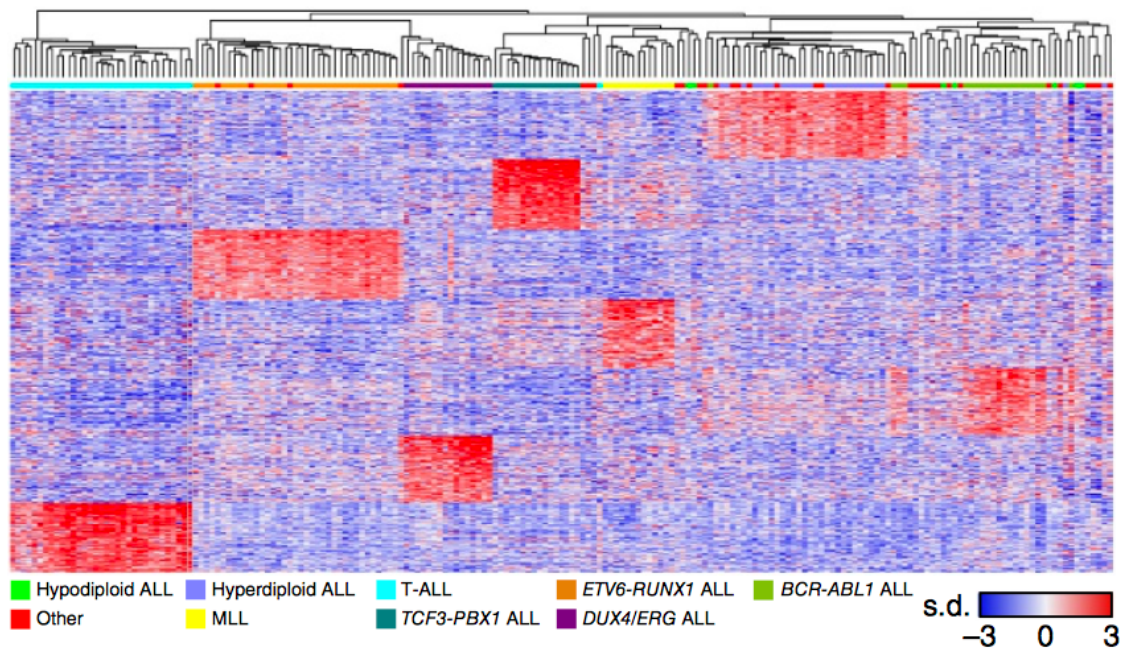others identified that this cluster corresponds to the one carrying *DUX4* fusions (Figure

2)[7, 8].

Hypodiploid ALL   Hyperdiploid ALL   T-ALL   *ETV6-RUNX1* ALL   *BCR-ABL1* ALL
Other   MLL   *TCF3-PBX1* ALL   *DUX4/ERG* ALL   s.d. −3 0 3

Figure 2. Hierarchical clustering of the transcriptome of ALL. Hierarchical
clustering of ALL identified a distinct expression cluster of patient samples with *DUX4*
fusions[7].

Among the oncogenic fusion genes in B-ALL, *DUX4* fusion is conspicuous

for some reasons. First, *DUX4* fusions are identified in the specific age group,

especially in the older pediatrics patients and in adolescent and young adult patients

(Figure 1). Second, the genomic function of *DUX4* gene is not fully understood. Third,

multiple copies of *DUX4* gene are coded repeatedly on the subtelomeric region of

chromosome 4 and chromosome 10, and it is challenging to capture the fusion gene

even with RNA-seq[9].

WT DUX4 is a transcription factor with a double-homeobox domain[9], coded

in a macrosatellite repeat region called D4Z4 in subtelomeric regions of chromosomes 4

and 10. WT DUX4 had long been thought to be a gene with unknown function, and its

accurate role in human normal physiology is still under investigation. As a disease

causing gene, WT *DUX4* first attracted attention as a causative gene of

facioscapulohumeral muscular dystrophy (FSHD)[10]. In normal physiology, WT

DUX4 is reported to be expressed only in limited areas such as testis, while its

expression is usually epigenetically repressed in somatic tissues[11]. De-repression of

WT DUX4 protein in myocytes causes cell death, leading to irreversible muscle loss in

FSHD[10]. The cytotoxicity of WT DUX4 when expressed in various types of somatic

tissues including hematopoietic cells are experimentally indicated[8].

Recent investigations have identified a critical role of WT *DUX4* in embryonic cleavage-specific transcriptional programs. A transient expression wave of WT *DUX4*, or *dux* in mouse, leads to the activation of genes necessary for phase-specific function in human and mouse embryos[12-14], called embryonic genome activation (EGA) (Figure 3). As a strong transcription factor, WT DUX4 activates genes required for the chromatin modification in cell division stages. These target genes include *ZSCAN4*, that is known to play an important role in telomere maintenance in embryonic cells[15], *LEUTX*, a homeobox gene[16], and *KDM4E*, a histone demethylase gene[17]. Furthermore, DUX4 activates Human Endogenous Retrovirus (HERV) retrotransposons[12]. These activated genes cooperate to cause chromatin modification required for further embryonic cell division.
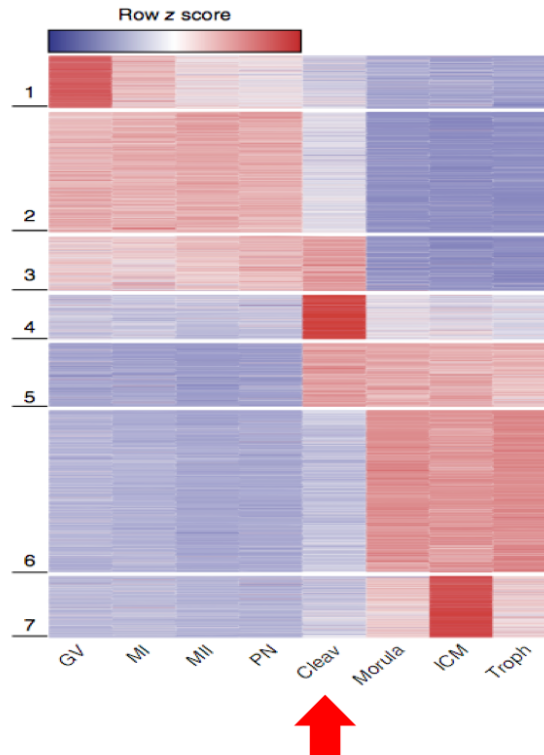
Figure 3. Hierarchical clustering of the transcriptome in embryonic cell division stages. Embryonic gene activation is driven by the activation of DUX4 specifically in cleavage stage[12]. GV, germinal vesicle; MI, metaphase I; MII, metaphase II; PN, pronuclear stage; Cleav, cleavage; ICM, inner cell mass; Troph, trophectoderm.

As an example of the oncogenic aspect of *DUX4* other than B-ALL, *CIC-DUX4* fusions were reported as a disease defining oncogene in Ewing-like sarcoma[18]. The resultant fusion protein is CIC protein fused to C terminus of DUX4, that is an opposite pattern of *DUX4* fusions identified in B-ALL cases. CIC-DUX4 fusion protein is speculated to show its transcriptional role mainly through the cooperation of C terminal transactivating domain of DUX4, and DNA binding domain of CIC[19].

9

From the perspective of the structure of *DUX4* fusions in B-ALL, *DUX4*

fusion transcripts are generated as a result of insertions of wild-type (WT) *DUX4*,

mainly into the *IGH* locus (Figure 4). Notably, the translocation replaces the 3′ end of

the WT *DUX4* coding region with a fragment of *IGH* or another gene, producing DUX4

out-of-frame fusion proteins devoid of the C terminus of WT DUX4 protein. DNA

binding domains of WT DUX4 were preserved in DUX4 fusion protein in every case.
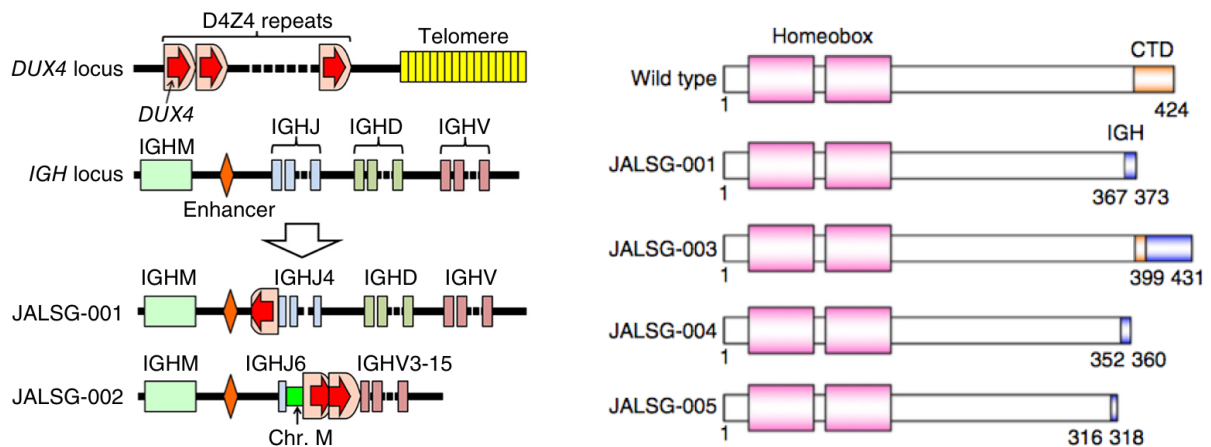

Figure 4. Gene and protein structures of *DUX4-IGH* in B-ALL. *DUX4* gene is
translocated to immunoglobulin loci (left), and produces DUX4-IGH protein devoid of
C terminus. Wild-type DUX4 is shown in the top (right)[5].

We previously identified that knockdown of *DUX4-IGH* in NALM-6 cells, a

B-ALL cell line carrying *DUX4-IGH* fusion gene, decreased the cell number of

NALM6 cells[5]. Furthermore, the introduction of *DUX4-IGH* into a mouse pro-B cell

results in defective differentiation of B cells, and ultimately developed leukemic

transformation.

These results indicate the oncogenic role of *DUX4-IGH* in B-ALL carrying

DUX4 fusions. To date, insertion of the *DUX4* gene has been implicated as the leading

event in the pathogenesis of B-ALL carrying *DUX4* fusions. The resultant DUX4 fusion

protein possesses transforming abilities, even though WT DUX4 is presumed to be pro-

apoptotic. Recently, the other group reported that DUX4 fusions transactivate *ERG*

gene, resulting in its alternative form that harbor oncogenic potential[7]. However, the

detailed mechanism of *DUX4* fusions has not been fully elucidated, given that the

functional analysis of WT *DUX4* is still underway.

In this regard, I performed functional analysis of *DUX4* fusions through

comprehensive analysis in order to identify the target gene of *DUX4-IGH*. Because the

induction of WT *DUX4* in somatic cells or knockdown of *DUX4-IGH* in the *DUX4-IGH*

dependent cell line lead to cell death, the materials and experimental methods available

were limited. To overcome this limitation, I utilized the assay that can be applied to a

small number of cells such as Assay for Transposase-Accessible Chromatin with

sequencing (ATAC-seq). Finally, I could identify the target gene of *DUX4-IGH* through

the integrated analyses of RNA-seq, chromatin immunoprecipitation assay coupled with

sequencing (ChIP-seq), and ATAC-seq.

Materials and Methods

Cell culture

Human leukemia cell lines, NALM6 and Reh, were obtained from the American Type

Culture Collection (ATCC) and cultured in RPMI 1640 supplemented with 10% fetal

bovine serum (FBS) and penicillin/streptomycin (all from Life Technologies, Carlsbad,

CA, USA). Human embryonic kidney (HEK) 293T cells were obtained from ATCC,

and maintained in Dulbecco's modified Eagle's medium (DMEM)-F12 (Life

Technologies) supplemented with 10% FBS, L-glutamine (Life Technologies), and

penicillin/streptomycin.

Retrovirus and lentivirus production and infection

Retroviral shRNA hairpins targeting *DUX4-IGH* in NALM6 was described previously

[5], and was constitutively expressed from pMKO.1 GFP (Addgene plasmid # 10676).

Complementary DNA (cDNA) for *DUX4-IgH* used in this study is derived from

NALM6 cells and patient sample in JAPAN ADULT LEUKEMIA STUDY GROUP

(JALSG) study[5]. These cDNA fragments were inserted into the retroviral plasmid

pMXs-ires-EGFP (Cell Biolabs, San Diego, CA, USA) or the lentiviral plasmid

pINDUCER20 (Addgene, #44012)[20]. To produce infectious viral particles, HEK293T

cells were transfected with such plasmids together with the retroviral packaging

plasmids (Takara Bio, Otsu, Japan), or lentiviral packaging plasmids (psPAX2,

Addgene, #12260, and pMD2.G, Addgene, #12259).

Establishment of antibody against DUX4

Polyclonal antibodies to an N-terminal portion of DUX4 (DUX4-N) were raised in

rabbits in Medical and Biological Laboratories, Japan. Two clones (clone A and clone

B) were used for the experiments.

ChIP-seq

ChIP was performed using a SimpleChIP Plus Enzymatic Chromatin IP Kit (Cell

Signaling Technology, Danvers, MA, USA), in accordance with the manufacturer's

instructions. Briefly, crosslinking of chromatin was achieved by incubation in 1%

formaldehyde at room temperature for 10 min. Crosslinked chromatin was fragmented

by enzymatic treatment and sonication. Sheared chromatin was immunoprecipitated

overnight using the following antibodies: anti-DUX4 antibody (in house, clone A), anti-

acetyl-histone H3 (Lys27, H3K27ac) rabbit mAb (Cell Signaling Technology, #8173),

and anti-trimethylated-histone H3 (Lys4, H3K4me3) rabbit mAb (Cell Signaling

Technology, #9751). Two percent of the sheared chromatin was used as a reference.

After reversal of the crosslinking of immunoprecipitated chromatin, genomic DNA was

extracted and then prepared for high-throughput sequencing. Two independent

experiments were performed. Paired-end reads were aligned to the human reference

genome (hg19) using Burrows-Wheeler Aligner (BWA v0.7.12)[21]. Downstream

analysis was performed on bam files, with duplicates removed using the samtools

rmdup command. Peaks were determined using MACS (version 1.4.2)[22], and filtered

to remove peaks that overlap with ENCODE blacklisted regions. The HOMER pipeline

(v4.9, 2-20-2017)[23] was used for annotation, *de novo* motif discovery, and calculation

of read density. Integrative Genomics Viewer (IGV version 2.3.34)[24] was used for

visualization.

ATAC-seq

To profile for accessible chromatin regions, we used ATAC-seq as reported[25]. Briefly, cells (50,000) were washed with $1 \times$ phosphate-buffered salts (PBS) by centrifugation using 5 min at 500 g. Cell pellets were re-suspended in 50 μl of cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl$_2$, 0.1% IGEPAL CA-630) and nuclei were pelleted by centrifugation for 10 min at 500 g, 4°C. Supernatant was discarded and nuclei were re-suspended in 25 μl reaction buffer containing 2.5 μl of Tn5 transposase and 12.5 μl of TD buffer (Nextera Sample preparation kit, Illumina). The reaction was incubated at 37°C for 30 min. Tagmented DNA was purified using a QIAGEN MiniElute PCR Purification Kit. For library amplification, PCRs were conducted with indexing primers included in the Nextera Index kit (Illumina, CA, USA) and NEBNext High-Fidelity $2 \times$ PCR Master Mix. The libraries were sequenced on an Illumina HiSeq2500.

ATAC-seq analyses

All the experiments were performed in duplicate. Briefly, using Bowtie2, reads were mapped onto the human genome hg19 assembly[26]. The duplicated reads and the reads

from mitochondria were removed. Reads in ENCODE blacklisted regions were also

removed. Replicate bam files were merged for downstream analysis. Reads enrichment

were called by MACS2 (v2.1.0)[22] with a q value cut off of 1×10^{-3} and the

parameters --nomodel --shift -150 –extsize 200.

Immunoprecipitation

Immunoprecipitations were performed by washing cells with cold PBS and resuspended

in buffer containing 0.1% NP-40 with protease inhibitors (cOmplete, EDTA-free, Roche,

Basel, Switzerland). Nuclear extract was acquired by sonication. Supernatants were

quantified, and 200 mg of protein was used for immunoprecipitation with 1−2 mg of

antibodies against DUX4 (in-house, clone A), rabbit IgG overnight at 4 ˚C and then for 2

hours with 20 mL Protein A Dynabeads (Thermo Fisher, MA, USA). Beads were eluted

with loading LDS and loaded onto SDS-PAGE.

Western blot analyses

Standard protocols were used for the detection of proteins by immunoblot (IB) analysis.

The following antibodies were used. Anti-DUX4 (in-house, clone B), anti-CBP (Cell

Signaling Technology, 7389), and anti-BCL2 (Cell Signaling Technology, 4223).

Membranes were developed using secondary antibodies for visualization by the

ImageQuant LAS4000 (GE Healthcare, Japan)

RNA extraction and quantification of transcripts

RNA was isolated using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) and

quantified by NanoDrop (Thermo Fisher, MA, USA). Reverse transcription was

performed using Superscript VILO (Thermo Fisher, MA, USA). Quantitative

polymerase chain reaction was performed using Power SYBR Green PCR Master Mix

(Thermo Fisher, MA, USA) on the 7500HT Fast Real-Time PCR System (Thermo

Fisher, MA, USA).

RNA-seq analyses

The cDNA library was made using the NEBNext Ultra Directional RNA Library Prep

Kit (New England BioLabs, MA, USA) and subjected to massively parallel sequencing

with the HiSeq2500 platform (Illumina, CA, USA). For expression profiling, paired-end

reads were aligned to the hg19 human genome assembly using TopHat2 (v2.0.9)[27].

The expression level of each gene was calculated using the DESeq2 package with

variance stabilizing transformation[28]. Expression of repetitive elements was

calculated using featureCounts[29].

Clinical sequence data

Sequencing data of JALSG B-ALL clinical samples[5] were obtained from the Japanese

Genotype–Phenotype Archive (accession JGAS00000000047), which is hosted by the

DNA Databank of Japan (DDBJ).

Gene pathway analysis and gene set enrichment analysis

Gene pathway analysis was performed using Metascape[30]. Gene set enrichment

analysis (GSEA) was performed as described previously[31]. Genes were ranked based

on the log2 fold change of the gene expression intensity determined with the DESeq2

package, and gene enrichment was calculated based on the rank of genes and gene sets.

We used gene sets of the Reactome Pathway Database (http://www.reactome.org),

combined with the dataset from previous reports[7, 32, 33].

Human cord blood cell culture

Human umbilical cord blood (CB) cells were obtained from Riken BRC or the Japanese

Red Cross Kanto Koshinetsu Cord Blood Bank (Tokyo, Japan). CD34+ cells were

separated and transduced with vector. All experiments using human cord blood cells were

approved by the Ethics Committee at the Institute of Medical Science, the University of

Tokyo (approval number: 27-34-1225).

Luciferase assay

A Genic region of *ZSCAN4* containing DUX4 binding sites was integrated upstream of

the luciferase-coding sequence in the pGL4.10 firefly luciferase vector (Promega,

Madison, WI, USA). The reporter construct was then cotransfected with the expression

vectors and control pGL4.74 (Promega) vector expressing *Renilla* luciferase. The signal

intensity of the firefly luciferase was normalized to the signal intensity of *Renilla*

luciferase and then normalized to that in a mock experiment.

Short hairpin RNA (shRNA) knockdown vectors

Target sequences of shRNA were as follows: *DUX4*, 5′-

CCCUGUGUGUCUCAGUUCA-3′; *STAP1*-1, 5′-GAACCCUUCUUUGGGAAAU-3′;

STAP1-2, 5′-CCAAGAAUCAAGCACUACA-3′; *STAP1*-3, 5′-

CCUGUAACACUCCCAAACC-3′; *STAP1*-4, 5′-GCUCUACCUUUGUACUUUG-3′;

CCNJ, 5′-CCCUUGUAUUACUCCAUGU-3′; *CLEC12A*, 5′-

GCAGCCUUGUUUCUGACUC-3′; *BCL2*-1, 5'-GUGAUGAAGUACAUCCAUU-3';

BCL2-2, 5'-CCGGGAGAUAGUGAUGAAG-3'

Individual sequences were cloned into the pMKO.1-GFP retroviral vector (Addgene

#10676, a gift from Prof. William Hahn, Dana-Farber Cancer Institute). NALM6 cells

were infected with each virus vector, and then the GFP-positive fraction of the cells was

measured and sorted for RNA extraction using flow cytometry.

Results

DNA binding capacity of DUX4-IGH

To identify the DUX4-IGH, I first investigated the DNA binding capacity and

histone modification ability of DUX4-IGH. I performed ChIP-seq with anti-DUX4

antibody, anti-trimethylation of histone H3 lysine 4 (H3K4me3) and anti-acetylation of

histone H3 lysine 27 (H3K27ac) antibody, using NALM6, a B-cell acute lymphoblastic

leukemia cell line carrying *DUX4-IGH* fusion. I detected 1,215 peaks (local rate cut-off

$<1.0 \times 10^{-5}$) with ChIP-seq against DUX4-IGH. Ninety three percent of the binding

sites of DUX4-IGH were located in intergenic and intronic regions, implying that

DUX4-IGH preferentially binds enhancer regions.

Motif analysis using DUX4-IGH binding regions identified a motif that was

almost identical to that recognized within WT DUX4 binding regions in human

myoblast and embryonic cells which was also recapitulated by reanalysis of the

published data on ChIP-seq of WT DUX4 in induced pluripotent stem (iPS) [11, 12].

Furthermore, ChIP-seq using anti-DUX4 antibody in Reh cell, a B-cell acute

lymphoblastic leukemia cell line without *DUX4-IGH* fusion, that is ectopically

expressing another type of DUX4-IGH identified in a clinical sample (JALSG 005)

indicated the similar motif. In every case, only the DUX4 motif was statistically

significant, suggesting the absence of a transcription factor that colocalizes with DUX4-
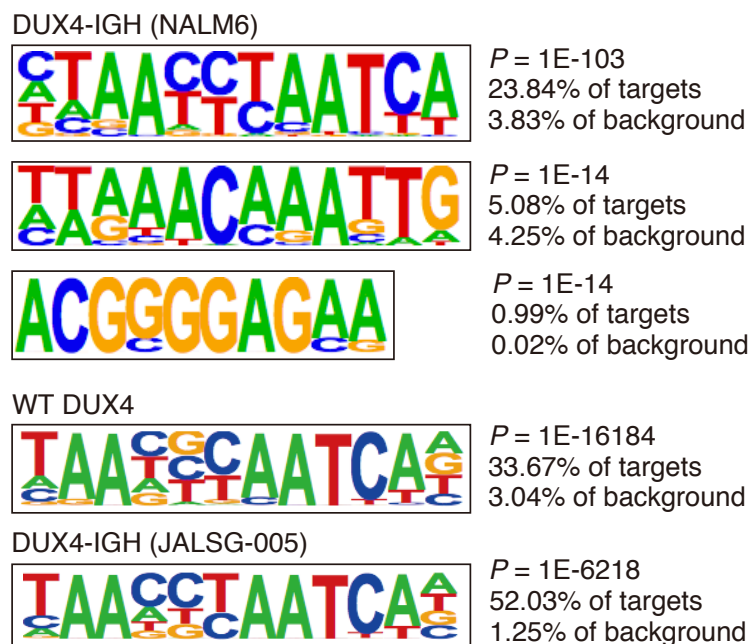
IGH (Figure 5).

DUX4-IGH (NALM6)

P = 1E-103
23.84% of targets
3.83% of background

P = 1E-14
5.08% of targets
4.25% of background

P = 1E-14
0.99% of targets
0.02% of background

WT DUX4

P = 1E-16184
33.67% of targets
3.04% of background

DUX4-IGH (JALSG-005)

P = 1E-6218
52.03% of targets
1.25% of background

5. *De novo* motif-discovery analysis using HOMER software. Motif analysis
4-IGH binding sites in NALM6 cells identified DUX4 motif as only significant
pper panel). The motifs of WT DUX4 (ectopic, Hendrickson et al. 2017) in
pluripotent stem cells (middle panel), and DUX4-IGH (JALSG 005-type,
ectopic) in Reh cells (lower panel) are almost identical to DUX4 motif in NALM6 cells.

Ninety-seven percent (1,057/1,082) of genes detected in our ChIP-seq against DUX4-

IGH were included among the detected genes in the ChIP-seq analysis against WT

DUX4 in iPS cells.

Histone modifications by DUX4-IGH

When the reads marked by H3K4me3 and H3K27ac were aligned against read

peaks of DUX4-IGH, they were preferentially clustered toward the center of the DUX4-

IGH peaks, indicating that H3K4me3 and H3K27ac are colocalized with DUX4-IGH

(Figure 6). Even though the 76% of the peaks of DUX4-IGH are accompanied by the

peaks of H3K4me3 and H3K27ac, only 4.6% of these overlapped regions are located in

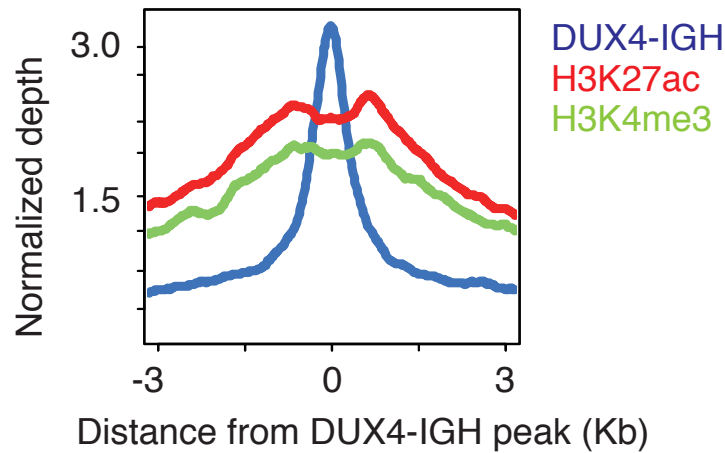promoter regions, and the most are located in intergenic and intron regions.

Figure 6. ChIP-seq in NALM6 cells. The average binding densities around DUX4-IGH peaks are shown: blue, DUX4-IGH; red, H3K27ac; green, H3K4me3. Y axis: normalized read number, X axis: distance from the peak of DUX4-IGH.

Next, to assess the histone activating capacity of DUX4-IGH at the binding loci, I ectopically expressed DUX4-IGH in Reh cells using inducible vector expressing DUX4-IGH and performed ChIP-seq against H3K27ac, comparing between with and without DUX4-IGH expression status. I observed partial activation of H3K27ac in the DUX4 bound regions corresponding to those in NALM6 region. (Figure 7).

Taken together, these data suggest that DNA binding capacity of DUX4-IGH is not altered compared with that of WT DUX4, and DUX4-IGH retains the capacity of augmenting active histone marks.
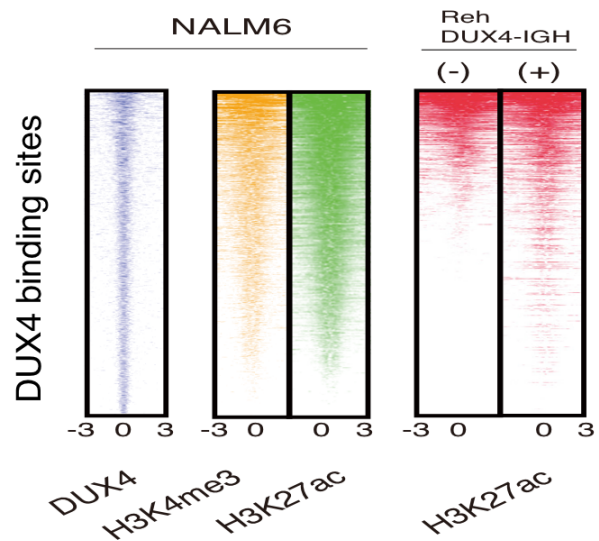
Figure 7. Heat maps of ChIP-seq in NALM6 and Reh. Heat maps of ChIP-seq against DUX4-IGH, H3K4me3, H3K27ac in NALM6 cells (left), and H3K27ac in Reh cells in which expression of DUX4-IGH was induced by doxycycline (Dox) (right).

Downstream targets of *DUX4-IGH*

To investigate the genes regulated by DUX4-IGH, NALM6 cells were infected with a retroviral vector expressing shRNA against *DUX4-IGH* or control shRNA. The expression level of *DUX4-IGH* was found to be decreased to 52% with the shRNA. I identified 279 significantly overrepresented genes (*DUX4-IGH* shRNA / control shRNA >1.5, an adjusted P-value of <0.05) and 190 significantly underrepresented genes (control shRNA / *DUX4-IGH* shRNA >1.5, an adjusted P-value of <0.05). (Figure 8)
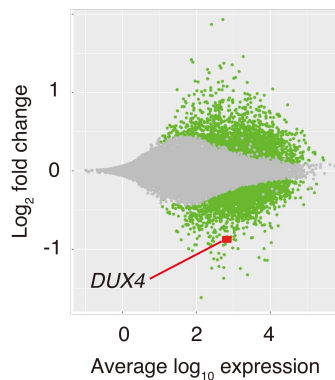
Figure 8. MA plot of genes with expression change after *DUX4-IGH* knockdown in NALM6 cells. MA plot showing the change in the transcripts with *DUX4-IGH* knockdown. Genes with a significant change (p<0.05) in expression are plotted in green, and *DUX4* is plotted in red. Y axis: fold change of the genes with expression change, X axis: average expression of each gene.

The list of the representative genes underrepresented in NALM6 cells with *DUX-IGH* knockdown is shown in supplementary table 1. The genes underrepresented in NALM6 cells with *DUX4-IGH* knockdown were supposed to comprise genes positively regulated by DUX4-IGH. I compared these genes positively regulated by DUX4-IGH with the DUX4-IGH bound genes proximally annotated by HOMER software. Only 14% (26/190) of them were overlapped, indicating that the majority of the genes dysregulated by *DUX4-IGH* knockdown were not direct transcriptional targets of DUX4-IGH. Among the direct target genes of DUX4-IGH, 93% (24/26)

were marked with H3K4me3, indicating that DUX4-IGH substantially contributed to
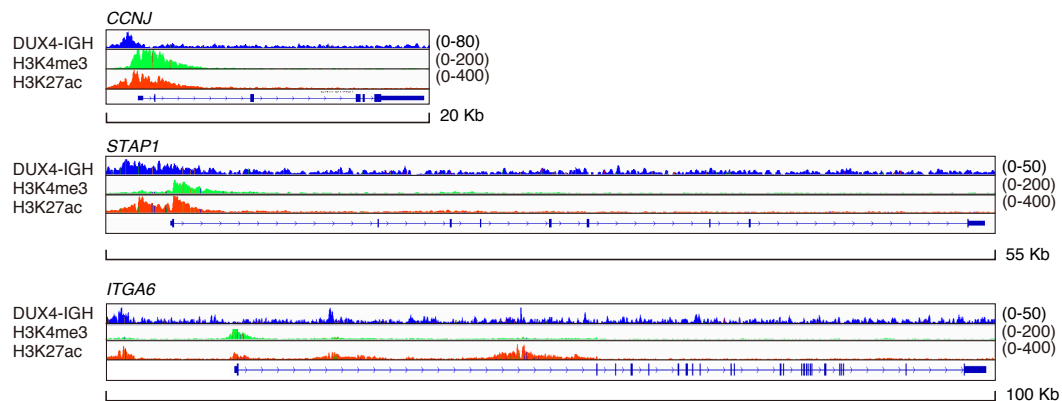
transcriptional activity in NALM6 cells (Figure 9).

Figure 9. ChIP-seq in the target gene loci of DUX4-IGH in NALM6 cells. Mapped
reads at t⸺us obtained with ChIP-seq against DUX4-IGH
(blue), H⸺(red) in NALM6 cells.

Direct ta⸺hange in expression with *DUX4-IGH*

knockdo⸺*TGA6,* whereas indirect target genes with the

largest fold change in expression include *CLEC12A*, *PTPRM*, and *DDIT4L*. *STAP1*,

CCNJ, and *ITGA6* were also bound by DUX4-IGH (JALSG 005-type) (Figure 10).
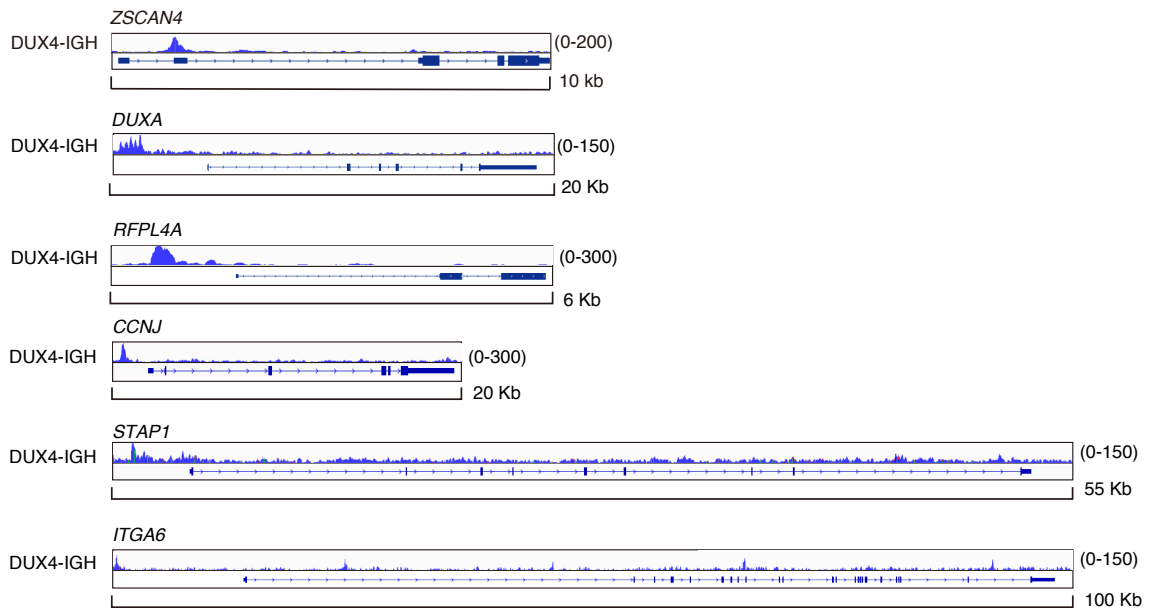
Figure 10. ChIP-seq against ectopically expressed DUX4-IGH in Reh cells. Mapped reads at the *ZSCAN4, DUXA, RFPL4A, CCNJ, STAP1*, or *ITGA6* locus obtained with chromatin immunoprecipitation coupled with sequencing (ChIP-seq) against DUX4-IGH (JALSG 005-type) in Reh cells.

In Reh cells, ectopic expression of DUX4-IGH induced expression of these direct and indirect target genes. In the RNA-seq from JALSG B-ALL clinical samples[5], I observed significantly higher expression of these genes in B-ALL cases with DUX4 fusion, compared with other types of B-ALL cases, further validating the specific transcriptional activity of DUX4-IGH (Figure 11).
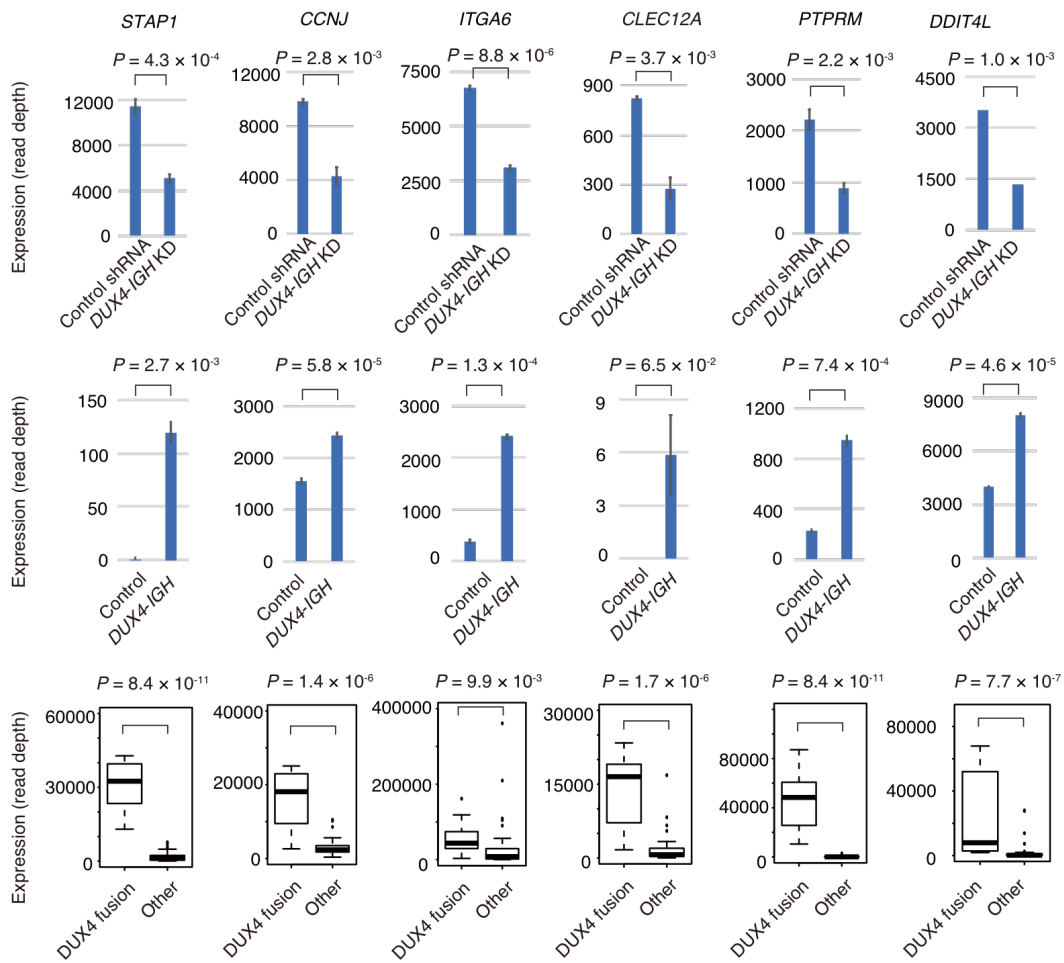
Figure 11. Effect on expression of target genes by DUX4-IGH in NALM6 cells, Reh cells and patient samples. Expression of representative genes in NALM6 infected with control shRNA or shRNA against *DUX4-IGH* (upper panel). Expression of the genes in Reh infected with control vector or *DUX4-IGH* expression vector (middle panel). Expression of the genes in clinical samples of B-ALL with DUX4 (n=13) or without DUX4 fusion (n=69) (lower panel).

Furthermore, the same pattern of expression was recapitulated regarding the other genes

such as *AGAP1*, *CHST2, GPR155,* and *PAM,* reported to be distinctly expressed in a

cluster of leukemia with *ERG* deletion[6], the cluster recently proved to overlap the

cluster carrying *DUX4* fusions. (Figure 12)
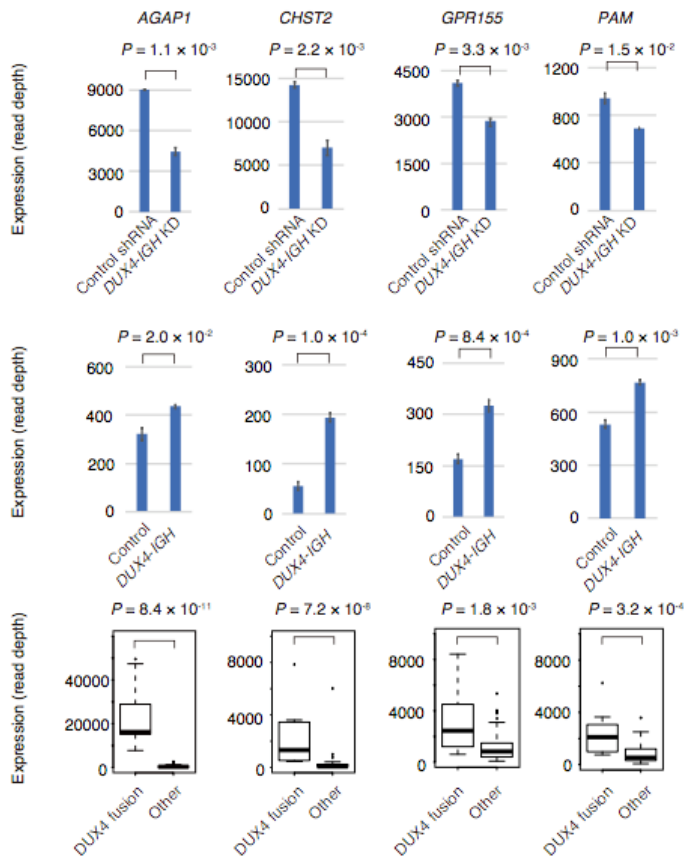
Figure 12. Effect on expression of target genes by DUX4-IGH in NALM6 cells, Reh cells and patient samples. Expression of representative genes in NALM6 infected with control shRNA or shRNA against *DUX4-IGH* (upper panel). Expression of the genes in Reh infected with control vector or *DUX4-IGH* expression vector (middle panel). Expression of the genes in clinical samples of B-ALL with DUX4 (n=13) or without DUX4 fusion (n=69) (lower panel).

GSEA revealed that genes upregulated in *DUX4* fusion carrying B-ALL[7]

were enriched in the genes overrepresented in NALM6 cells with *DUX4-IGH*
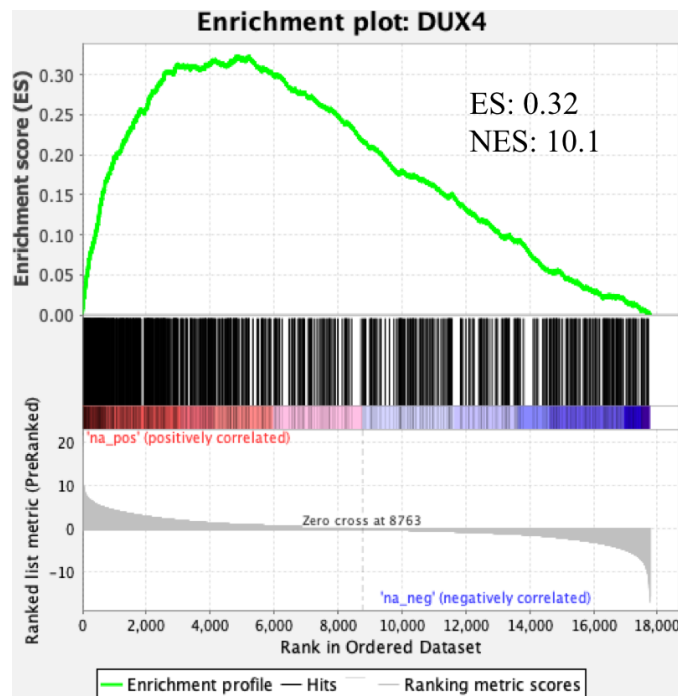
knockdown (Figure 13).

Figure 13. Gene set enrichment analysis using the gene set unique in *DUX4* fusion carrying B-ALL. Gene set enrichment analysis of overrepresented genes in NALM6 cells with *DUX4-IGH* knockdown against gene sets upregulated in *DUX4* fusion carrying B-ALL[7] showed enrichment of the *DUX4-IGH* target genes.

Pathway analysis of the genes underrepresented in NALM6 cells with *DUX4-IGH* knockdown revealed that these genes are associated with neural differentiation and cell adhesion (Figure 14).
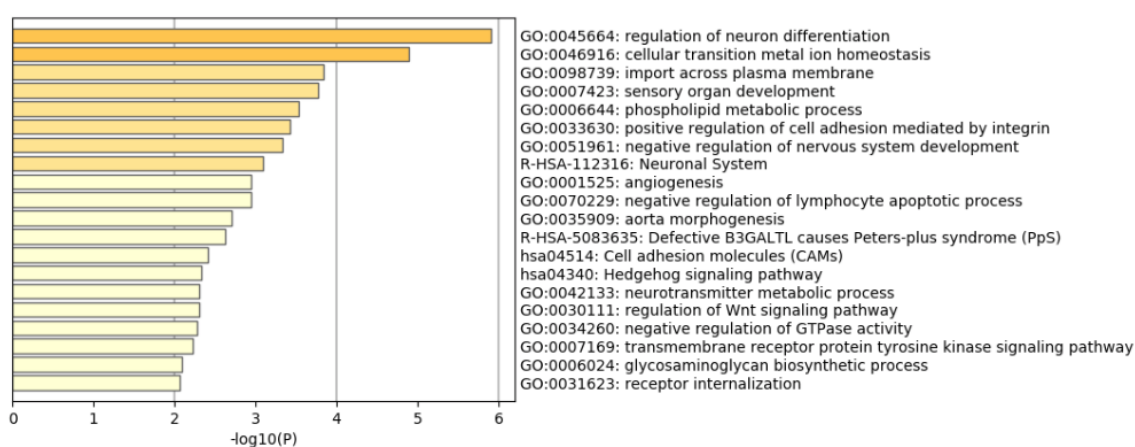

GO:0045664: regulation of neuron differentiation
GO:0046916: cellular transition metal ion homeostasis
GO:0098739: import across plasma membrane
GO:0007423: sensory organ development
GO:0006644: phospholipid metabolic process
GO:0033630: positive regulation of cell adhesion mediated by integrin
GO:0051961: negative regulation of nervous system development
R-HSA-112316: Neuronal System
GO:0001525: angiogenesis
GO:0070229: negative regulation of lymphocyte apoptotic process
GO:0035909: aorta morphogenesis
R-HSA-5083635: Defective B3GALTL causes Peters-plus syndrome (PpS)
hsa04514: Cell adhesion molecules (CAMs)
hsa04340: Hedgehog signaling pathway
GO:0042133: neurotransmitter metabolic process
GO:0030111: regulation of Wnt signaling pathway
GO:0034260: negative regulation of GTPase activity
GO:0007169: transmembrane receptor protein tyrosine kinase signaling pathway
GO:0006024: glycosaminoglycan biosynthetic process
GO:0031623: receptor internalization

-log10(P)

Figure 14. Pathway analysis of overrepresented genes in NALM6 cells with *DUX4-IGH* knockdown. Pathway analysis using overrepresented genes in NALM6 cells with *DUX4-IGH* knockdown identified pathway related to neural differentiation and cell adhesion.

Taken together, DUX4-IGH exerts the transcriptional activity in forming the characteristic expression signature of B-ALL carrying DUX4-fusions, directly and indirectly. To investigate the role of these genes in the survival of NALM6 cells, I performed knockdown of some of these genes using shRNA, especially focusing on *STAP1*, which is the downstream of Tec kinase in B-cell receptor signaling[34], *CCNJ,*

a member of the cyclin family which is involved in epithelial cancers[35], and

CLEC12A, which is associated with leukemia stem cell in acute myeloid leukemia[36].

However, single knockdown of these genes did not inhibit the growth of NALM6 cells
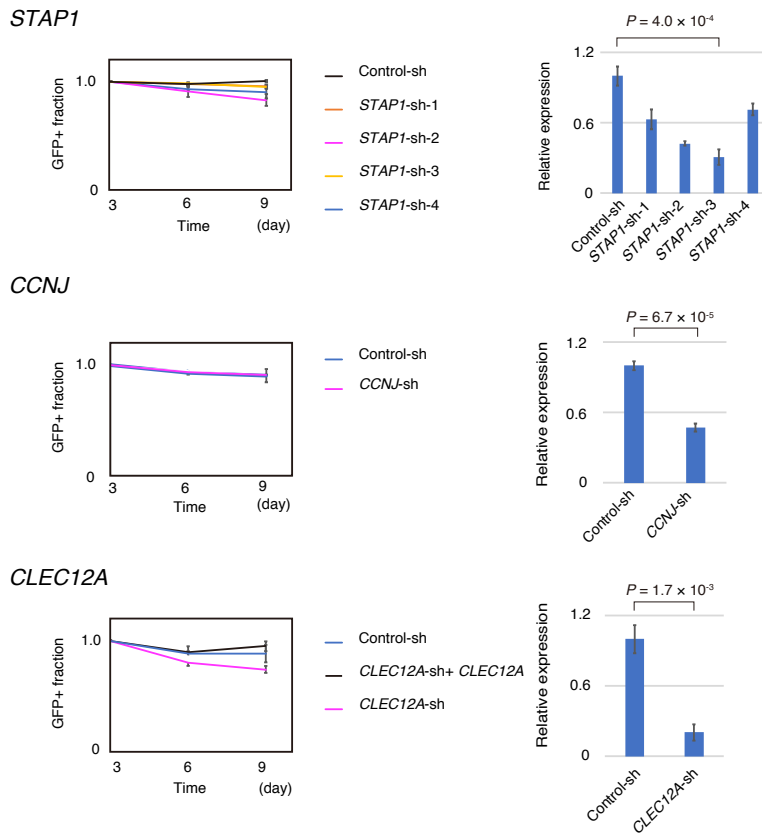
(Figure 15).

Figure 15. Knockdown of target genes in NALM6 cells. NALM6 cells were infected with a recombinant retrovirus expressing GFP and either a control shRNA (Control-sh) or an shRNA targeting *STAP1* (*STAP1*-sh-1,2,3,4), *CCNJ* (*CCNJ*-sh), or *CLEC12A* (*CLEC12A*-sh), with or without a virus coding shRNA-resistant versions of the target genes. The GFP+ cell fraction was evaluated by flow cytometry after culture of the cells for the indicated times (left). Data were normalized to the value for day 3. Knockdown efficiency was evaluated using quantitative polymerase chain reaction (right).

Target genes of DUX4-IGH and WT DUX4

While DUX4-IGH has the capacity to transform pro-B cells, induction of WT DUX4 causes cell death in somatic tissues. To identify the difference of transactivation ability between DUX4 fusions and WT DUX4, I compared the genes regulated by DUX4-IGH and WT DUX4. As ectopic expression of WT DUX4 causes rapid cell death in cell lines other than NALM6 cells, I introduced inducible vectors encoding DUX4-IGH or WT DUX4 into NALM6 cell, where DUX4-IGH is expressed endogenously in NALM6. Induction of DUX4-IGH and WT DUX4 in NALM6 increased the expression of DUX4 to 10-fold and 2-fold, respectively. In NALM6 cells, the expression of 149 genes was upregulated (>1.5-fold, an adjusted P-value of <0.05) by DUX4-IGH, while 1519 genes were upregulated by WT DUX4 (>1.5-fold, an adjusted P-value of <0.05). Only a few genes were upregulated by both DUX4-IGH and WT DUX4 (Figure 16). This result suggested that transcriptional targets of DUX4-IGH and WT DUX4 are different.
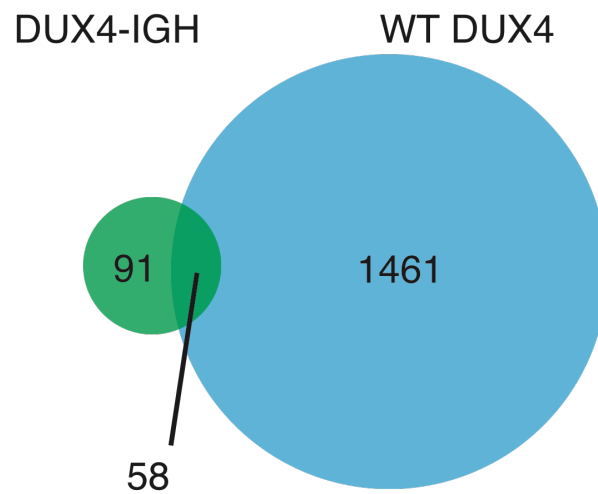
Figure 16. Comparison of target genes between WT DUX4 and DUX4-IGH. Venn diagram showing the number of genes upregulated by ectopic expression of DUX4-IGH (green) or wild-type (WT) DUX4 (light blue) in NALM6 cells.

Regarding the genes positively regulated by WT DUX4 that are identified in Figure 16, pathway analysis revealed that these genes are mainly associated with proteolysis or autophagy, possibly reflecting the ability to cause cell death by WT DUX4 (Figure 17).
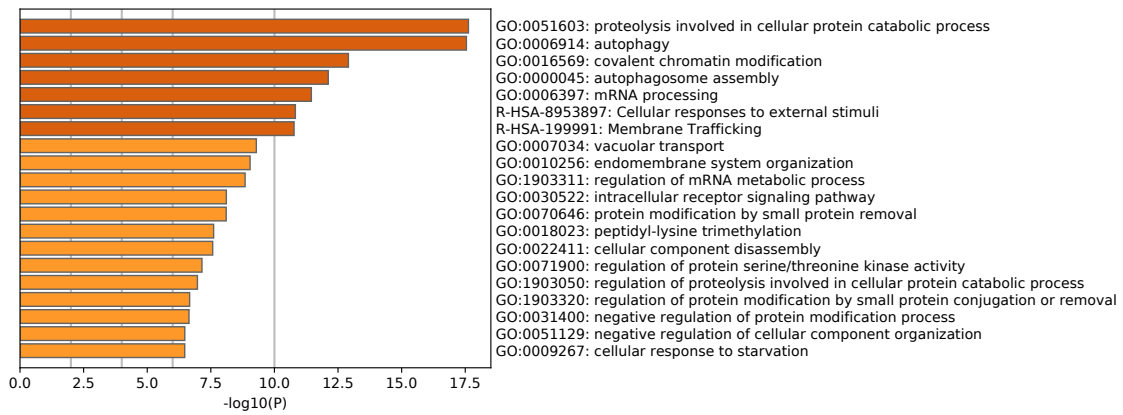
Figure 17. Pathway analysis of the gene sets positively regulated by WT DUX4 using NALM6 cells. Pathway analysis using WT DUX4 target genes identified the pathway related to proteolysis or autophagy.

To identify the transcriptional ability of DUX4 fusions on the target genes of WT DUX4, I analyzed the known target genes of WT DUX4 that are bound by WT DUX4, such as *ZSCAN4*, *DUXA,* and *REFPL4A.* The expression of these genes were prominently upregulated by the ectopic expression of WT DUX4, but not by DUX4-IGH (NALM6-type) (Figure 18). Similarly, these genes were not upregulated by ectopic expression of DUX4-IGH in Reh cells.
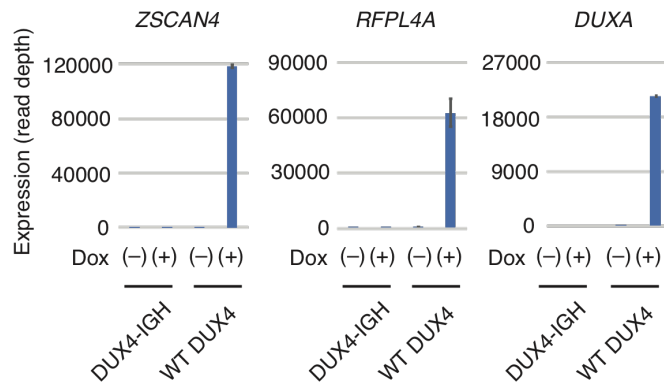
Figure 18. Comparison of expression levels of WT DUX4 target genes by DUX4-IGH or WT DUX4 in NALM6 cells. Expression of the representative target genes of WT DUX4 in NALM6 cells, in which expression of DUX4-IGH or WT DUX4 was induced by doxycycline (Dox).

Moreover, luciferase assay of DUX4-IGH or WT DUX4 using *ZSCAN4* promoter

indicated decreased transcriptional activity of various types of DUX4-IGH fusions

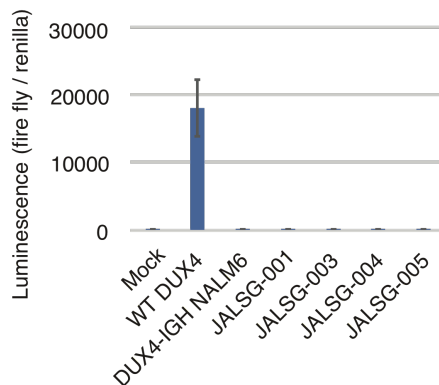compared with that of WT DUX4 (Figure 19).

Figure 19. Luciferase assay using the genic region of DUX4 binding sites in *ZSCAN4*. Luciferase assay was performed using the genic region of DUX4 binding sites in *ZSCAN4* by ectopically expressing WT DUX4 and DUX4 fusions from various samples.

H3K4me3 and H3K2

NALM6 cells, confir

NALM6 cells (Figure

ZSCAN4
DUX4-IGH   (0-100)
H3K4me3   (0-200)
H3K27ac   (0-400)
12 kb

DUXA
DUX4-IGH   (0-50)
H3K4me3   (0-200)
H3K27ac   (0-400)
20 Kb

RFPL4A
DUX4-IGH   (0-150)
H3K4me3   (0-200)
H3K27ac   (0-400)
6 Kb
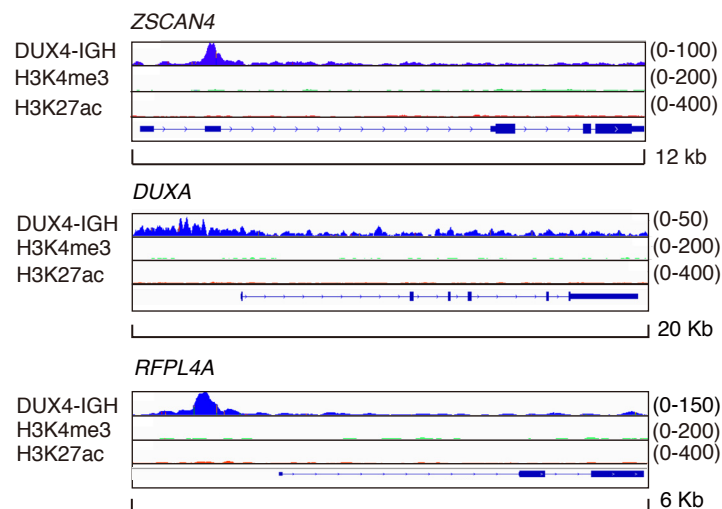
Figure 20. ChIP-seq in the WT DUX4 target gene loci of DUX4-IGH in NALM6 cells. Mapped reads at the *ZSCAN4, DUXA, RFPL4A* locus obtained with ChIP-seq against DUX4-IGH, H3K4me3 and H3K27ac in NALM6 cells.

Effect of DUX4-IGH on Human Endogenous Retroviral (HERV) elements

Gene activation in transposable elements is another hallmark of WT DUX4 in

EGA, namely in Human Endogenous Retrovirus (HERV) retrotransposons that are

essential for the subsequent gene activation[12]. Therefore, I sought the relevance of

ERV element activation in NALM6 cells. As I mentioned before, DUX4-IGH binding is

enriched in LTR regions compared with the general distribution in the human genome,

implying the preference for binding of DUX4-IGH to these regions. When I calculated

the RNA expression in NALM6 cells ectopically expressing DUX4-IGH, WT DUX4,

and Reh cells ectopically expressing DUX4-IGH, the expression of various HERV

elements were activated only by WT DUX4, and not by DUX4-IGH in NALM6 cells
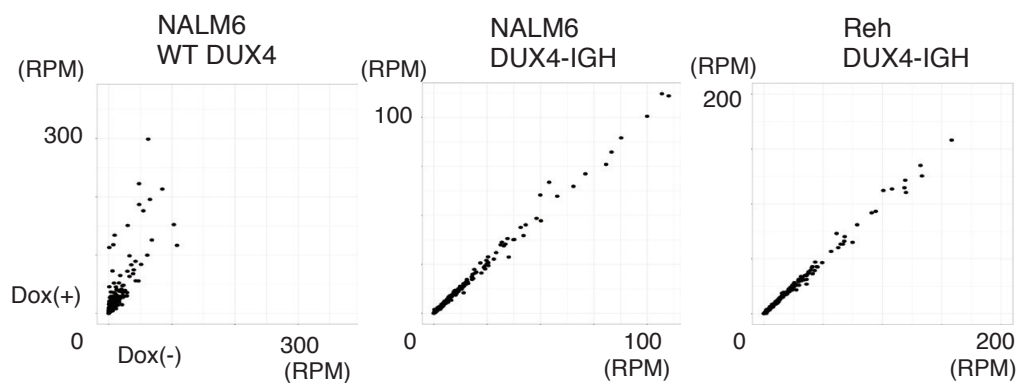
and Reh cells (Figure 21).

**Figure 21. Comparison of expression levels of HERV elements by DUX4-IGH or
WT DUX4 in NALM6 cells and by DUX4-IGH in Reh cells.** Expression of HERV
elements in NALM6 cells in which expression of WT DUX4 was induced by Dox,
NALM6 cells in which expression of DUX4-IGH was induced by Dox, and Reh cells in
which expression of DUX4-IGH was induced by Dox (right)

Furthermore, there was no difference in the expression of the representative WT DUX4

target HERV elements such as *HERVL-int*, *MLT2A1*, between the B-ALL patient

samples with and without *DUX4*-fusions, indicating that HERV elements were not

activated in B-ALL carrying *DUX4* fusions (Figure 22). These results suggest that

DUX4-IGH almost lost the transactivation activity at the HERV element in B-ALL.

Transcription of the other repetitive elements was not activated even with WT DUX4

(data not shown).
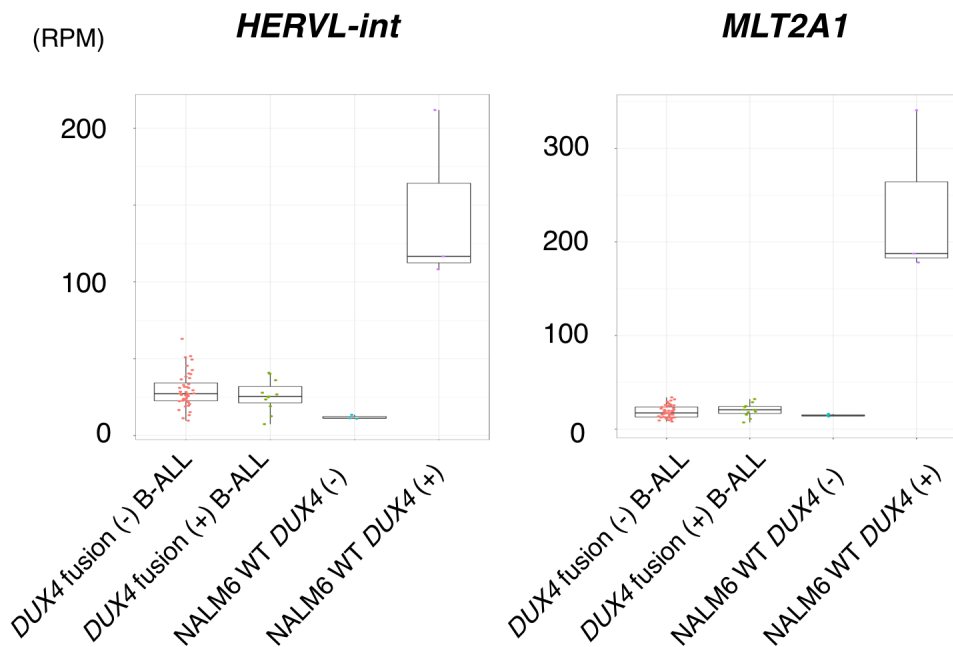

Figure 22. Comparison of expression levels of the representative HERV elements.
Expression of *HERVL-int* and *MLT2A1* in B-ALL with *DUX4* fusion, B-ALL without
DUX4 fusion, NALM6 cells with or without ectopic expression of WT *DUX4*.

Taken together, these results suggested that transactivation ability of DUX4-IGH against WT DUX4 targets is attenuated compared with that of WT DUX4. This is in good agreement with a previous report describing that DUX4-s[37], a short variant of DUX4 that maintains the N-terminal double-homeobox domain but lacks the C-terminal transactivation domain, loses transactivation ability through an inability to bind histone acetyltransferases such as EP300 and CBP. Actually, I confirmed that CBP is not the binding partner of DUX4-IGH through immunoprecipitation against DUX4 in NALM6 cells (Figure 23). I was unable to perform immunoprecipitation against WT DUX4 with ectopically WT DUX4 expressing NALM6 cells, because WT DUX4 expression led to early cell death of NALM6 cells.
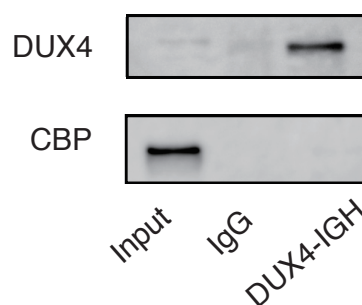

Figure 23. Western blots of anti-DUX4 and anti-CBP. Immunoprecipitation against DUX4 in NALM6 cells. Western blots of anti-DUX4 and anti-CBP are shown. Different antibody clones were used for immunoprecipitation and western blotting. DUX4 (clone A) was used for immunoprecipitation, and DUX4 (clone B) was used for western blotting.

Chromatin structure change driven by DUX4-IGH

As DUX4-IGH binding is not always associated with the transcriptome

change, the identification of the direct targets of DUX4-IGH cannot be achieved only

with ChIP-seq against DUX4. I speculated that the locus with transcriptional activity

should be accompanied by the open chromatin status, and to this end, I compared the

chromatin accessibility of the genome of NALM6 cells with that of *DUX4-IGH*

depleted NALM6 cells with ATAC-seq. NALM6 cells were collected 3 days after the

induction of the control shRNA vector or shRNA vector targeting *DUX4-IGH.* ATAC-

seq revealed 69,206 peaks and 76,696 peaks in with and without DUX4-IGH

knockdown. When I mapped the ATAC reads to the peak of DUX4-IGH in NALM6

cells in ChIP-seq, I identified that DUX4-IGH bound region showed a decrease of

ATAC signaling, which support ability of DUX4-IGH to open the chromatin, that is

consistent with the previous report of the chromatin opening ability of WT DUX4 in

early embryo[38]. In addition, Reh cells collected 2 days after induction of *DUX4-IGH*

expression vector, were subjected to ATAC-seq (Figure 24). When I focused on the

DUX4-IGH bound regions overlapping repetitive elements such as LTR, LINE, and

SINE, ATAC signal was not affected by the existence of DUX4-IGH. On the other

hand, ATAC-seq identified that chromatin opening ability in the intron and intergenic

regions are retained, suggesting that transactivating ability of DUX4-IGH is exerted

only in intron and intergenic regions.
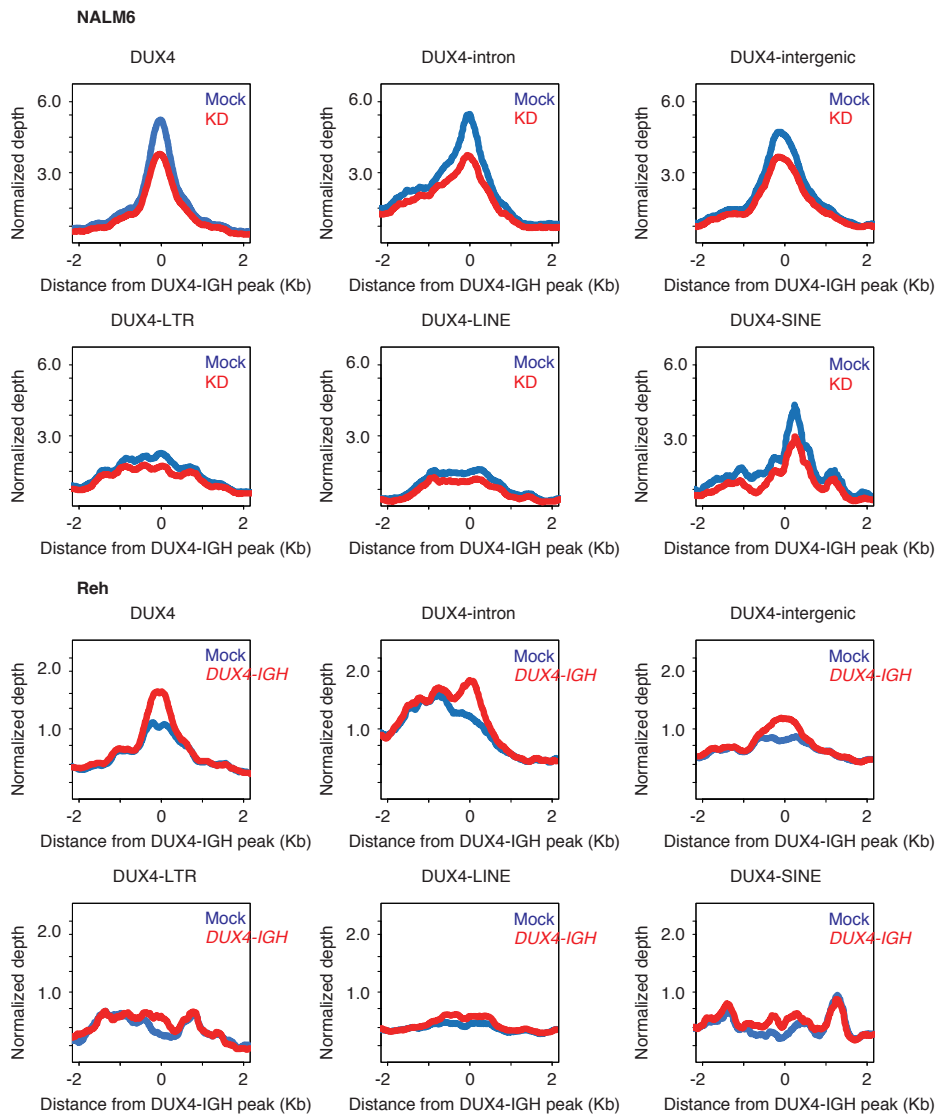
Figure 24. ATAC-seq signals in DUX4-IGH binding regions in NALM6 cells and Reh cells. Changes of ATAC signal with knockdown (KD) of DUX4-IGH in NALM6 cells (upper panel) and with ectopic expression of DUX4-IGH in Reh cells (lower panel) at the overlapping regions between the binding sites of DUX4-IGH and intron regions, intergenic regions, LTR regions, LINE regions, and SINE regions.

I next performed differential peak calling between the two status, and found that

substantial change in open chromatin status between with and without DUX4-IGH

knockdown, using the region with a substantial change (more than 2-fold intensity). I

identified 243 peaks lost and 3285 peaks gained by DUX4-IGH knockdown (Figure

25).
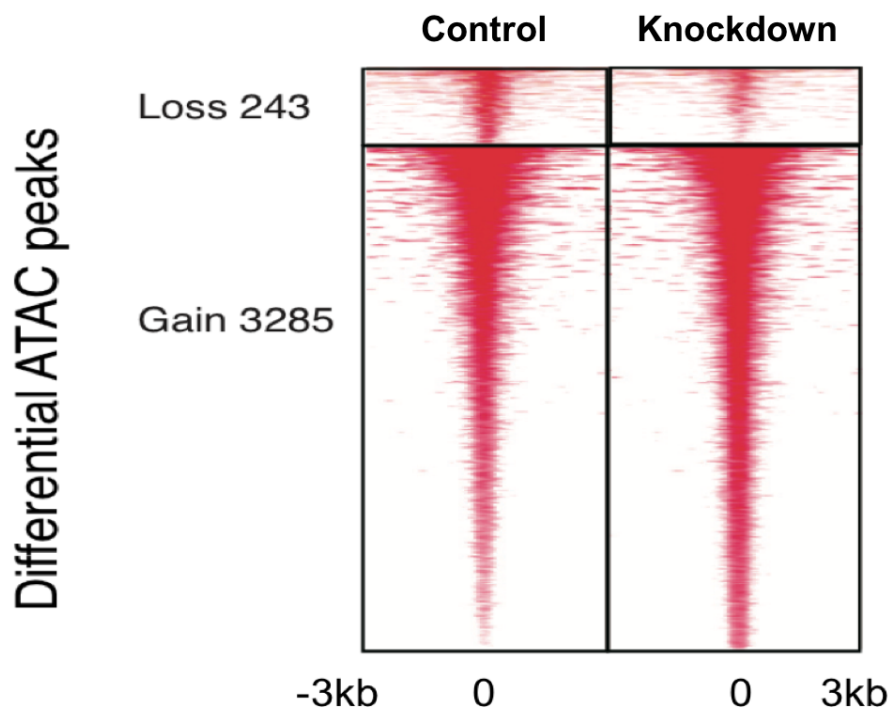
Figure 25. Heat maps of differential peaks of ATAC-seq. Heat maps of gained and
lost ATAC signal peak regions in NALM6 cells with (right) and without (left)
knockdown of *DUX4-IGH*.

Motif analysis of lost peaks revealed a DUX4 motif in the region lost by knockdown of

DUX4-IGH in NALM6 cells. The identification of DUX4 motifs is compatible with the

result of ChIP-seq motif analysis of DUX4-IGH peak that identified only DUX4 motif

(Figure 5), and support the idea that DUX4-IGH opens the DUX4 bound area and

activates transcription, possibly without coexistence of other transcription factors. On

the other hand, PU.1, encoded by *SPI1*, CTCF, TCF3, EBF1 motifs were identified in

the region gained by knockdown of DUX4-IGH (Figure 26). The motifs gained by

knockdown of DUX4-IGH may imply that existence of DUX4-IGH repress the

expression or binding of transcription factors relevant to B cell differentiation such as

SPI1, TCF3, EBF1, and loss of DUX4-IGH restores their activity.
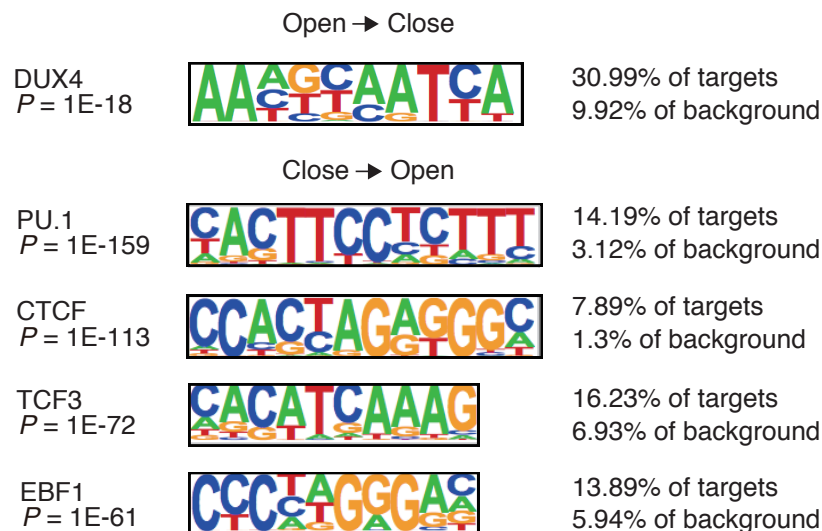
Figure 26. *De novo* motif-discovery analysis using HOMER software. Motif analyses of ATAC peaks lost by DUX4-IGH knockdown in NALM6 (upper panel), and gained (lower panel).

Next, I searched for the genes with expression change whose open chromatin

status is influenced by *DUX4-IGH*. Comparing genes where ATAC signal decreases

with knockdown of *DUX4-IGH* in NALM6 cells and increases with induction of *DUX4-*

IGH in Reh cells, I identified *BCL2, ITGA6, and TSACC* as DUX4-IGH target

candidates. Among them, *BCL2* is of interest from the perspective of hematologic

malignancies (Figure 27).
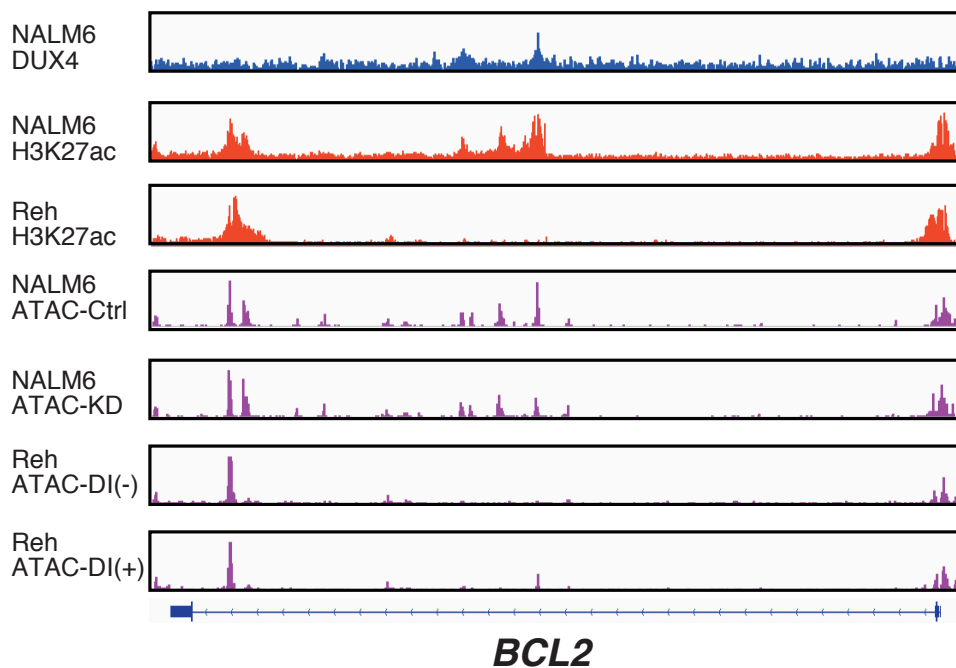
Figure 27. ChIP-seq and ATAC-seq in *BCL2* locus. Mapped reads at the *BCL2* locus
obtained with ChIP-seq against DUX4-IGH, H3K27ac in NALM6 cells, ChIP-seq
against H3K27ac in Reh cells, ATAC-seq with or without knockdown of DUX4-IGH in
NALM6 cells, ATAC-seq with or without ectopic expression of DUX4-IGH in Reh
cells. Ctrl, control; KD, knockdown; DI, DUX4-IGH

The transcript expression level and protein level of BCL2 was decreased by *DUX4-IGH*

knockdown in NALM6 cells (Figure 28).
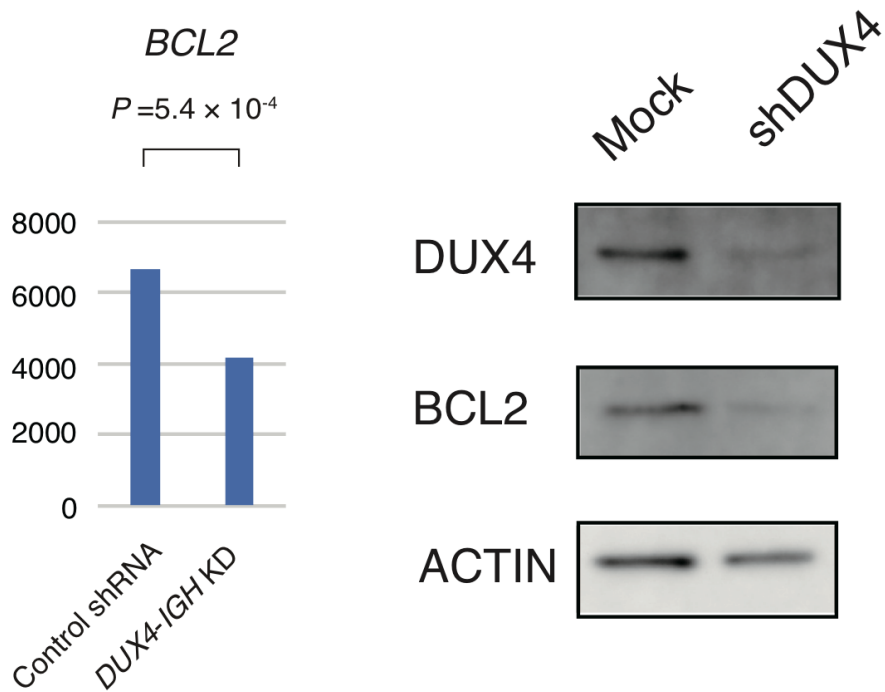
Figure 28. Expression and protein levels of *BCL2* after knockdown of *DUX4-IGH*.
Transcriptome level and protein level of *BCL2* in NALM6 infected with control short
hairpin RNA (shRNA) or shRNA against *DUX4-IGH*.

Number of NALM6 cells was decreased with knockdown of *BCL2* (Figure 29).
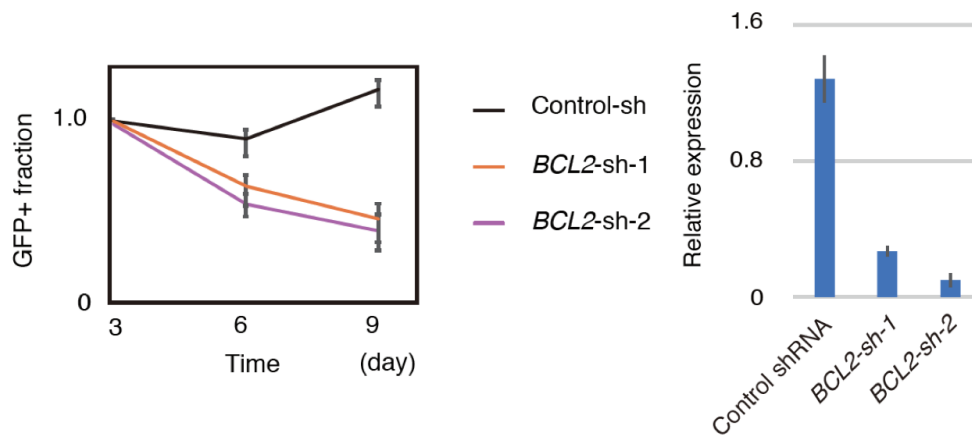
Figure 29. Knockdown of *BCL2* in NALM6 cells. NALM6 cells were infected with a recombinant retrovirus expressing GFP and either a control shRNA (Control-sh) or an shRNA targeting *BCL2* (*BCL2*-sh-1,2). The GFP+ cell fraction was evaluated by flow cytometry after culture of the cells for the indicated times (left). Data were normalized to the value for day 3. Knockdown efficiency was evaluated using quantitative polymerase chain reaction (right).

Finally, FDA approved BCL2 inhibitor, ABT-199, inhibited the cell growth of NALM-6 cells in vitro (Figure 30), suggesting that NALM6 cells are dependent on the increased expression of *BCL2* by transactivation by DUX4-IGH.
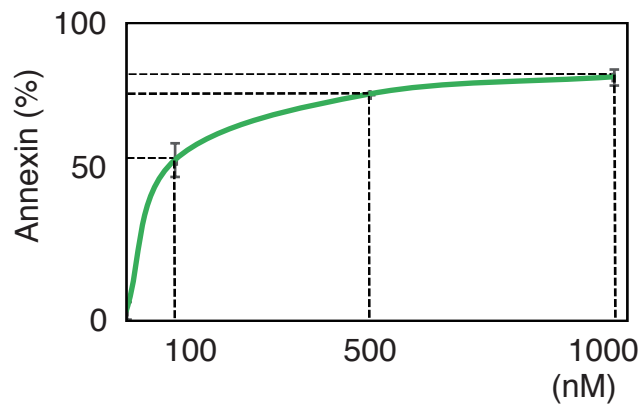
Figure 30. Administration of ABT-199 to NALM6 cells. Administration of ABT-199 to NALM6 cells showed dose dependent increase of cell death after 24 hours.

Genes associated with hematopoietic cell function regulated by DUX4-IGH

To explore the effect of *DUX4-IGH* on human hematopoietic cell, I ectopically expressed *DUX4-IGH* in human CB CD34+ cells. While human CB cells introduced with *DUX4-IGH* maintained the expression of CD34 at least for 14 days, those with the mock vector rapidly lost CD34 expression. (Figure 31).
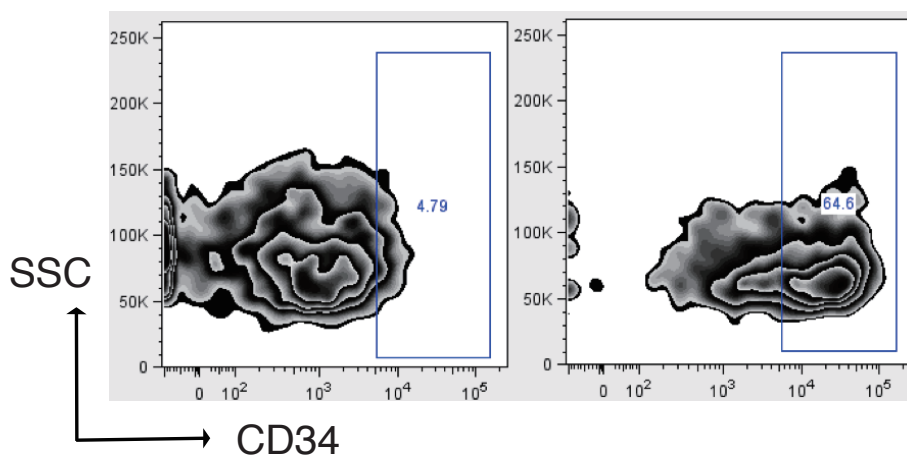
Figure 31. Induction of *DUX4-IGH* in CD34+ cells. Surface marker CD34 expression of human cord blood CD34+ cells cells infected with a mock vector (left) and a vector expressing DUX4-IGH (right), followed by 14 days *in vitro* culture.

CD34 was underrepresented in NALM6 cells with *DUX4-IGH* knockdown, and was overrepresented in Reh cells with DUX4-IGH ectopic expression. *CD34* is also highly expressed in clinical samples with DUX4 fusion (Figure 32).
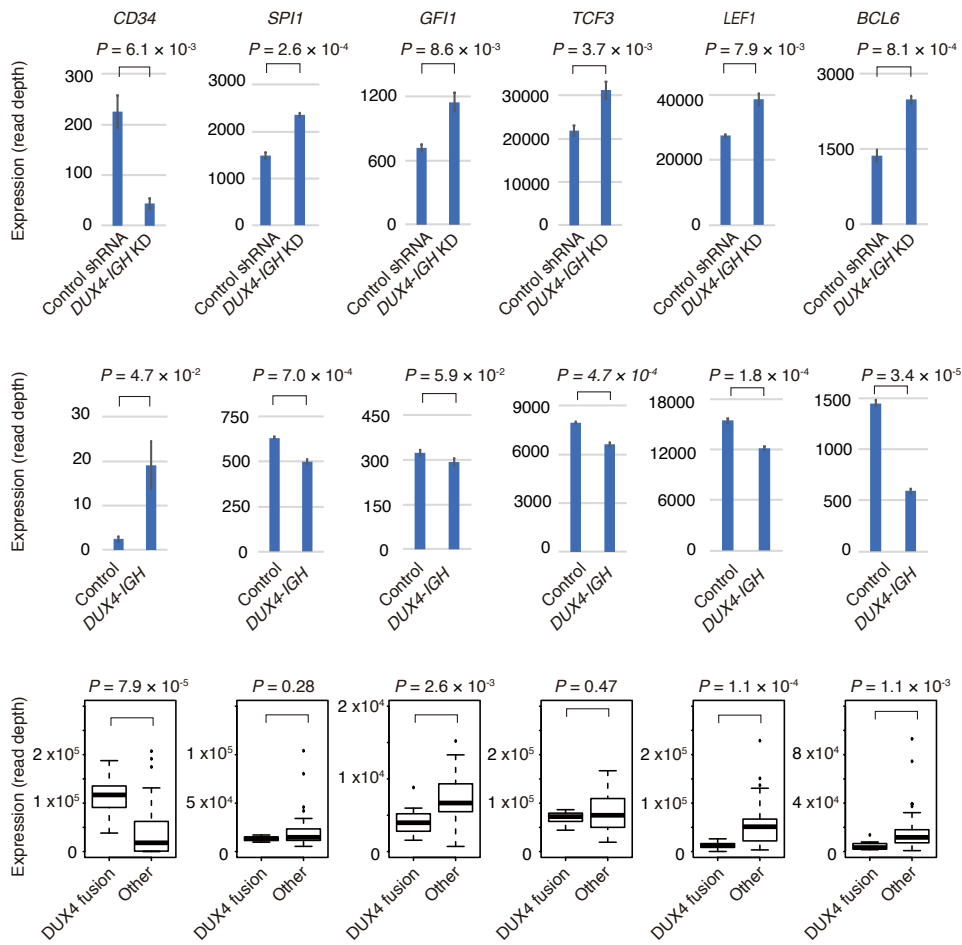
Figure 32. Effect on expression of genes related to hematopoiesis by DUX4-IGH in NALM6 cells, Reh cells and patient samples. Expression of representative genes related to hematopoietic differentiation in NALM6 infected with control short hairpin RNA (shRNA) or shRNA against *DUX4-IGH* (upper panel). Expression of the genes in Reh infected with control vector or *DUX4-IGH* expression vector (middle panel). Expression of the genes in clinical samples of B-ALL with DUX4 (n=13) or without DUX4 fusion (n=69) (lower panel).

Interestingly, genes overrepresented in NALM6 cells with *DUX4-IGH* knockdown

included those associated with hematopoietic cell differentiation, and exhibit a low

expression level specifically in clinical samples with DUX4 fusion. Expression levels of

these genes were moderately repressed in Reh cells and NALM6 cells by DUX4-IGH

ectopic expression. Furthermore, these genes include those involved in early B-

lymphoid development, such as *TCF3* (E2A) and *SPI1* (PU.1). In addition, transcription

factors regulated through various development stages of B lymphocytes, such as *GFI1*,

LEF1, and *BCL6* were significantly decreased in B-ALL cells with DUX4 fusions.

These genes also play important roles in phases before lymphoid lineage priming. To

clarify the effect of *DUX4-IGH* knockdown on B-cell differentiation, we performed

GSEA using the datasets from Hystad et al.[32] and Laurenti et al.[33]. GSEA revealed

that genes associated with pro-B cells, pre-B cells, and multilineage progenitor (MLP)

cells were enriched in the genes overrepresented in NALM6 cells with DUX4-IGH

knockdown, while genes associated with megakaryocyte-erythroid progenitor (MEP)

cells and common myeloid progenitor (CMP) cells were enriched in the genes

underrepresented in NALM6 cells with DUX4-IGH knockdown (Figure 33). This result

indicated that DUX4-IGH enables progenitor cells to retain an immature status, and

knockdown of DUX4-IGH drives NALM6 cells toward B-cell differentiation, which is

compatible with our previous results *in vivo* demonstrating that murine B-lymphocyte

differentiation is abrogated by the expression of DUX4-IGH [5].
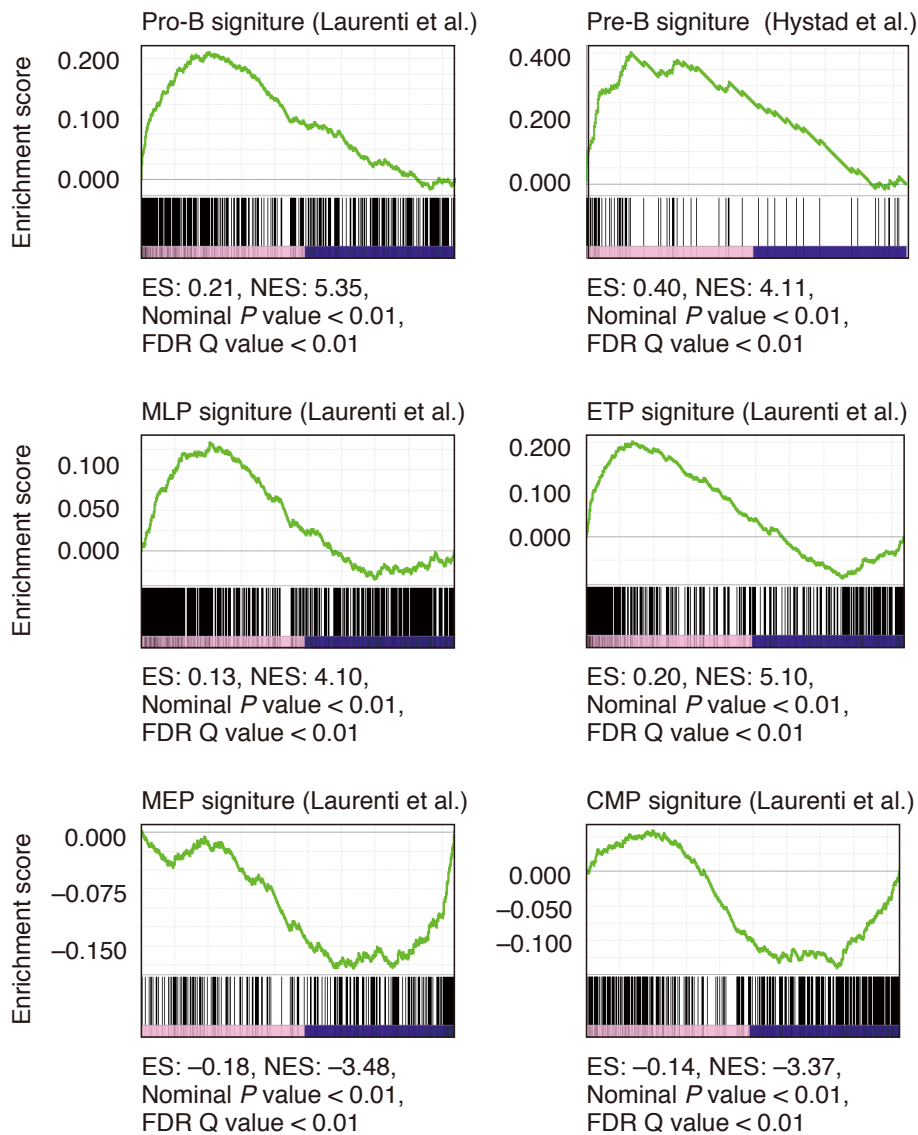
Figure 33. Gene set enrichment analysis using the gene set related to hemaptopoiesis. Gene set enrichment analysis of the underrepresented or overrepresented genes in NALM6 cells with *DUX4-IGH* knockdown using gene set related to hematopoiesis. ES, enrichment score; NES, nominal enrichment score; FDR, false discovery rate.

Discussion

I analyzed the transcriptional activity of DUX4-IGH and the gene expression

profile regulated by DUX4-IGH, providing several important findings that shed light on

the mechanism by which DUX4-IGH is associated with leukemogenesis.

WT DUX4 is known as a transcription factor that induce transcription of genes

necessary for embryonic gene activation[12]. Therefore, I first analyzed the

transcriptional ability and histone modification ability of DUX4-IGH. I revealed that

DNA binding motif of DUX4-IGH is similar to WT DUX4, and DUX4-IGH holds a

substantial histone modification capacity.

Several observations indicated that DUX4-IGH exhibited substantial

transcriptional activities. First, a large proportion of genomic regions bound by DUX4-

IGH were marked with H3K4me3 and H3K27ac. Second, *DUX4-IGH* knockdown

resulted in decreased expression of a subset of genes, some of which were bound by

DUX4-IGH. Third, the previous report identified that *ERG* alternative isoforms, that are

characteristic to the leukemia cases with *DUX4* rearrangement, are expressed from the

binding site of DUX4-IGH, supporting the transcriptional activity of DUX4-IGH[7].

These data suggest that DUX4-IGH holds transcriptional ability.

The ability of DUX4-IGH to induce the transcription of known WT DUX4

target genes was severely attenuated compared with that of WT DUX4. Target genes of

WT DUX4 including the HERV elements are not activated by DUX4-IGH, even though

these genes are bound by DUX4-IGH. Previous report identified that DUX4-s, a short

variant of DUX4 that maintains the N-terminal double-homeobox domain but lacks the

C-terminal transactivation domain, had lost transactivation ability[37]. The structure of

DUX4-s is similar to DUX4-IGH in the point that they have lost the C terminal

transactivation domain. The report showed that histone acetyltransferases can bind to

WT DUX4 but not to DUX4-s, and I confirmed that CBP did not bind to DUX4-IGH in

NALM6 cells. These results partially explain the attenuated transcriptional activity of

DUX4-IGH against WT DUX4 targets.

The question is whether the transcriptional activity of DUX4-IGH is the

weaker version of WT DUX4, or transcriptional activity of DUX4-IGH is dependent on

the different mechanisms. Recent study identified that translocation of *DUX4-IGH* in

DUX4 rearranged leukemia cases predominantly occurs in the silenced allele of

immunoglobulin heavy chain locus, not in the active immunoglobulin heavy chain

locus, where immunoglobulin heavy chain RNA is actively transcribed[39]. This data

suggests that expression of *DUX4-IGH* is regulated to the "just-right" levels, because

the high expression levels of *DUX4-IGH* will lead to the cytotoxicity, while *DUX4-IGH*

is essential for the proliferation of leukemia cells. In this regard, the hypothesis that

DUX4-IGH lost its C-terminal transactivation domain to attenuate transcriptional ability

and avoid the unnecessary cytotoxicity but utilize the DNA binding domain, is

plausible. However, there are some issues that need to be addressed. First, the genes

underrepresented in NALM6 cells with *DUX4-IGH* knockdown differed from the genes

whose expression was induced by the ectopic expression of WT *DUX4* in NALM6

cells, implying that the transcriptional targets of DUX4-IGH differ from those of WT

DUX4. Second, there is an imbalance of the attenuation of transcription in DUX4-IGH

bound genes; transcription of the known target genes of WT DUX4, such as *ZSCAN4*, is

totally lost, while there is a substantial transcriptional ability of DUX4-IGH for the

target genes of DUX4-IGH, such as *STAP1, CCNJ*. The mechanism WT DUX4 and

DUX4-IGH transactivate different targets is not clear, thus requires further investigation

including the chromatin structure analyses.

ATAC-seq analysis identified that the existence of DUX4-IGH caused drastic

change in the chromatin status. WT DUX4 is reported to open chromatin in early

embryo[38], and I confirmed the chromatin opening ability of DUX4-IGH. On the other

hand, I observed more regions with increased ATAC signals after knockdown of

DUX4-IGH in NALM6 cells. Motif analysis of these regions identified PU.1, TCF3,

EBF1. Expression of *SPI1* and *TCF3* are increased after knockdown of *DUX4-IGH* in

NALM6 cells, and expression of *SPI1* is repressed in B-ALL samples carrying *DUX4*

fusions. Previous report showed that *SPI1*-deficient mice also failed to generate the

earliest lymphoid progenitors[40]. My results suggest that *SPI1* is repressed when

DUX4-IGH is expressed and knockdown of *DUX4-IGH* de-represses the expression of

SPI1. Consequently, PU.1 binds to the genomes and leads to the differentiation of B

cells. Higher expression levels of *CD34* in B-ALL cells with *DUX4* fusion, together

with GSEA results showing the loss of myeloid trait and the gain of lymphoid trait in

response to *DUX4-IGH* knockdown, suggested that the transcriptional network

perturbation by DUX4-IGH may result in a more primitive hematopoietic stage

associated with the potential for multi-lineage differentiation.

Although DUX4-IGH induced the expression of the target genes, the

contribution of those genes to leukemogenesis is not clear. Gene groups induced by

DUX4-IGH did not mainly include the genes related to B cell function, but include

some genes whose function is reported to be related to cancer. I knocked down some of

these genes, but knockdown of the single gene did not affect the proliferation of

NALM6 cells. ATAC-seq analysis identified *BCL2* as a target candidate of DUX4-IGH.

BCL2 is a regulator of mitochondrial apoptotic pathway[41], and many hematopoietic

malignancies are dependent of dysregulation of *BCL2*. Recently, FDA approved ABT-

199, BCL2 inhibitor[42], for some of the hematopoietic malignancies[43], making the

high expression status of *BCL2* as a therapeutic target. Actually, I confirmed in vitro

efficacy of ABT-199 against NALM 6 cells, suggesting the utility of BCL2 inhibitor for

B-ALL carrying *DUX4* fusions.

In conclusion, I revealed that DUX4-IGH holds a transactivating activity

related to leukemogenesis. DUX4-IGH also drives drastic chromatin changes both at

DUX4-IGH bound genomic regions or other genomic regions, that are associated with

the defective differentiation of B-ALL carrying *DUX4* fusions. Further understanding of

the mechanism of transcriptional network perturbation by DUX4-IGH should provide

clues for a treatment strategy for B-ALL carrying *DUX4* fusions.

Supplementary table 1

Gene	Average expression	Log2FoldChange	DUX4-IGH binding
CD34	225.6334617	-1.617812116	
CLEC12A	822.04585	-1.370271783	
DDIT4L	3509.575647	-1.326547522	
PTPRM	2219.460342	-1.246392686	
SPINK4	394.3958937	-1.241855631	
TMPRSS3	812.9792336	-1.235237097	○
RP11-100K18.1	593.1789554	-1.219809986	
STAP1	11441.2949	-1.141183318	○
CCNJ	9834.356308	-1.140027973	○
RP11-253E3.3	529.5596913	-1.135540043	
TMIE	435.8760683	-1.122892836	
CNTN2	659.9037998	-1.110666711	
ITGA6	6770.22983	-1.105954375	○
ENPP2	792.429675	-1.069495838	○
B4GALT6	568.7718143	-1.051171147	
TMEM132A	350.6217895	-1.041906104	○
LRRC7	120.7734429	-1.01927024	○
NRCAM	97.72115254	-0.993693485	○
IPCEF1	1810.23307	-0.992869845	○
AGAP1	9047.681329	-0.992816514	
FJX1	1311.496222	-0.984604183	
TTYH2	735.3067165	-0.948544733	
NAALADL2	146.197193	-0.944111783	
KIAA1671	132.3182423	-0.943672829	
GJA3	337.7887934	-0.941218236	
ARRB1	816.1881348	-0.934138116	
ROR2	297.7530584	-0.933612735	
GPR124	284.9829263	-0.926007933	
CHST2	14211.06891	-0.915945241	
TBX2	399.9735095	-0.901518219	

RP11-366M4.3	367.7231247	-0.892098747	
ANKRD46	1528.45267	-0.883522167	
RP11-17A4.2	63.31002515	-0.880023557	
DUX4L13	914.9370747	-0.874615128	
PPARGC1B	1367.565422	-0.874288229	
XKR4	146.76515	-0.863125466	
DGAT2	388.7694135	-0.857972983	
CLEC14A	3956.282891	-0.854899303	
SOGA2	4271.083401	-0.85165436	
TRIM71	300.8004136	-0.843353128	
NT5DC3	1135.841047	-0.84177213	
ADAMTS15	474.5772341	-0.839401335	
THNSL1	898.9672454	-0.838532321	
PLD6	1385.218978	-0.836572468	
GNPDA1	319.5365243	-0.831986662	
OSCAR	165.6557575	-0.831248282	
SPR	1444.429312	-0.825107371	
ARRDC4	552.1419483	-0.815729072	
SIRPB2	89.78398789	-0.809491906	
KCNJ12	2498.069952	-0.803959326	
MAP6D1	278.7394993	-0.79295811	
ALS2CL	431.0395314	-0.786085483	
RASL10B	1358.62467	-0.785195856	
ZFYVE28	379.3885004	-0.784787879	○
PCTP	432.3875951	-0.781080458	
TPTE2P2	51.62469381	-0.779751201	
RP11-61L19.2	127.0178231	-0.779293507	
PTGFRN	422.8560065	-0.773111275	
SLC16A1	9046.30012	-0.770445264	
HPDL	879.7809664	-0.769114939	
KCNK3	47.8980093	-0.767619432	
TMEM170B	769.7106828	-0.766400435	○

RASA4	1357.737596	-0.755826922	
SLC31A2	163.2466353	-0.754887561	
CTC-523E23.1	406.4494921	-0.754304544	
TSACC	191.8134501	-0.746770768	○
GLDC	570.6321455	-0.745783495	
KAZN	1517.206699	-0.745475955	○
TMEM132B	98.23755946	-0.741800398	○
RAB31	707.7449705	-0.741166405	
CSRNP3	1745.617279	-0.740187374	○
CCDC3	158.7011345	-0.737320706	
EGFL7	120.1303436	-0.731358182	
CD3EAP	2366.80456	-0.730313114	
PIEZO1	1284.388706	-0.728758655	
PXDN	13653.11501	-0.728611829	
RASEF	624.3930331	-0.722106329	
PCOLCE2	827.7374891	-0.721009563	
MEGF11	155.2116064	-0.720401033	
PLXNA1	156.4536703	-0.720320563	
ARHGEF4	335.8913631	-0.719724356	
MINA	2915.507832	-0.719053088	
KLRF2	119.5173682	-0.718403207	
SOCS2	1547.564519	-0.713272019	
FAM109B	236.8894852	-0.711858836	
PANK1	726.311169	-0.710069969	
HIF3A	66.15050915	-0.708839606	
SLC16A10	1139.4461	-0.707856319	
SKI	705.6083746	-0.705642538	
DAB2IP	180.6145519	-0.704356334	
MLEC	12396.69619	-0.703915584	
TTN	20265.35581	-0.701524179	
OAS1	250.9517048	-0.69931771	
WDR72	191.8085391	-0.698178973	○

B3GALT6	1608.23215	-0.69604274	
NUS1	7508.143855	-0.695063368	
VMO1	148.073131	-0.694629482	○
LA16c-366D3.1	81.70450279	-0.693285239	
AC093668.2	84.80607025	-0.69034824	
WDR43	8605.143186	-0.687671153	
ETNK1	5108.892946	-0.6858543	
LINC00977	73.34838338	-0.685155949	○
IFFO2	139.525642	-0.682175325	
NT5DC2	5042.140286	-0.680862857	
GPRC5B	81.92382698	-0.680722895	
POLR1B	4511.984482	-0.677592626	
MPP3	254.2119119	-0.67580513	
IL17D	364.5321845	-0.675299752	
ABCG2	801.8631686	-0.671184561	○
FRMPD1	110.93264	-0.669726842	
SACS	12115.86136	-0.669188534	
ANKS6	1182.330741	-0.668004202	
C12orf79	245.828585	-0.667114298	
CTC-523E23.4	202.8897987	-0.663928182	
BCL2	6686.970916	-0.660377545	○
UMPS	3115.177872	-0.66024092	
SLC1A3	296.0523098	-0.658927328	
LPCAT1	1656.800357	-0.652966118	
MPV17L	267.8607519	-0.650988801	
FKBP4	4748.154983	-0.649745181	
PMP22	565.7160581	-0.649419523	
DFFB	851.9291722	-0.648530253	
FAM86A	836.8805683	-0.648329448	
NLRP11	95.15550079	-0.645349985	○
TMEM177	437.7831717	-0.64143888	
NDST1	1720.705708	-0.641293444	

RP11-16P6.1	175.0746433	-0.640938824	
ADCY5	154.2810057	-0.640446625	
GNG11	103.4221669	-0.638434312	
LRP3	593.4190595	-0.637698066	
SLC39A14	5997.781002	-0.63752127	
ADAMTS17	549.5703405	-0.636183338	
HEATR3	2227.175789	-0.635663774	
SLC39A10	1509.70147	-0.635540593	
ZSCAN5C	76.20882304	-0.635316517	
C17orf51	337.386972	-0.635313761	
ZNF485	347.752296	-0.634102519	
PODXL	2294.629701	-0.629730675	
RET	107.3821305	-0.62955348	
RP11-452H21.4	147.6497333	-0.629404742	
MAML2	1512.072997	-0.629048154	◯
SAMD10	99.31531448	-0.626961371	
ARL10	284.9875966	-0.625396007	
TNFRSF8	74.18764809	-0.624038997	
RP5-1024G6.8	1590.869745	-0.622328293	
CHRNA5	1024.311265	-0.621505758	
SPIRE1	366.6709386	-0.621301393	
EFNB2	3071.905606	-0.619772989	
TMEM150A	95.68437078	-0.619269415	
MBLAC2	1209.949053	-0.616876067	
RP11-418J17.1	267.1638423	-0.616670306	
HMOX2	1500.465632	-0.615527545	◯
GRB10	4891.843613	-0.614449315	
ARG2	90.24306346	-0.612826568	
CABLES1	564.5159495	-0.612354956	
SLC19A1	2669.308915	-0.612084677	
SOCS2-AS1	181.7857876	-0.610628144	
ZNF285	134.6908377	-0.609610686	

AIMP2	2698.90423	-0.608773608	
ZDHHC9	865.8480342	-0.608248993	
AP001372.2	88.47909466	-0.607839316	
CCDC138	1781.21639	-0.607200505	
MT1H	1719.9637	-0.606311283	
AGPAT5	3630.687157	-0.606055287	
SCO2	533.4180599	-0.603935048	
URB2	4184.809426	-0.602693083	
PTGER2	1151.488259	-0.602454231	○
TMEM120B	2110.98435	-0.601167088	
SLC29A3	260.8686244	-0.600887756	
EEPD1	429.9586292	-0.599819583	
MYCL	165.6233735	-0.599464112	
MIR17HG	490.9441787	-0.598495954	
PTPRS	6622.755634	-0.598082358	
GPR161	101.2321281	-0.596596011	
TRAM2	1613.202163	-0.595488147	
PUS7	4063.305703	-0.59447347	
GHRLOS	102.2242724	-0.593796718	
ADAMTS9	557.7468309	-0.593132862	○
SHANK3	173.4419812	-0.592894208	
DOC2A	128.3005565	-0.591966967	
CHRNA1	233.0186535	-0.591891478	
RPP40	1510.588267	-0.591790075	
PDCD2L	720.0535709	-0.590848635	
CYTIP	657.4639915	-0.589410365	
TRIM37	5009.067581	-0.589227675	○
DIXDC1	607.0190004	-0.588983454	
CPNE2	2217.109215	-0.588964888	
PIGW	1791.030785	-0.585980128	
RPF2	3340.502698	-0.585792211	

Acknowledgements

References

1.   Terwilliger, T. and M. Abdul-Hay, *Acute lymphoblastic leukemia: a comprehensive review and 2017 update.* Blood Cancer J, 2017. **7**(6): p. e577.

2.   Iacobucci, I. and C.G. Mullighan, *Genetic Basis of Acute Lymphoblastic Leukemia.* J Clin Oncol, 2017. **35**(9): p. 975-983.

3.   Hunger, S.P. and C.G. Mullighan, *Acute Lymphoblastic Leukemia in Children.* N Engl J Med, 2015. **373**(16): p. 1541-52.

4.   *The future of cancer genomics.* Nat Med, 2015. **21**(2): p. 99.

5.   Yasuda, T., S. Tsuzuki, M. Kawazu, F. Hayakawa, S. Kojima, T. Ueno, N. Imoto, S. Kohsaka, A. Kunita, K. Doi, T. Sakura, T. Yujiri, E. Kondo, K. Fujimaki, Y. Ueda, Y. Aoyama, S. Ohtake, J. Takita, E. Sai, M. Taniwaki, M. Kurokawa, S. Morishita, M. Fukayama, H. Kiyoi, Y. Miyazaki, T. Naoe, and H. Mano, *Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults.* Nature Genetics, 2016. **48**(5): p. 569-574.

6.   Harvey, R.C., C.G. Mullighan, X. Wang, K.K. Dobbin, G.S. Davidson, E.J. Bedrick, I.M. Chen, S.R. Atlas, H. Kang, K. Ar, C.S. Wilson, W. Wharton, M. Murphy, M. Devidas, A.J. Carroll, M.J. Borowitz, W.P. Bowman, J.R. Downing, M. Relling, J. Yang, D. Bhojwani, W.L. Carroll, B. Camitta, G.H. Reaman, M. Smith, S.P. Hunger, and C.L. Willman, *Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome.* Blood, 2010. **116**(23): p. 4874-84.

7.   Zhang, J., K. McCastlain, H. Yoshihara, B. Xu, Y. Chang, M.L. Churchman, G. Wu, Y. Li, L. Wei, I. Iacobucci, Y. Liu, C. Qu, J. Wen, M. Edmonson, D. Payne-Turner, K.B. Kaufmann, S.I. Takayanagi, E. Wienholds, E. Waanders, P. Ntziachristos, S. Bakogianni, J. Wang, I. Aifantis, K.G. Roberts, J. Ma, G. Song, J. Easton, H.L. Mulder, X. Chen, S. Newman, X. Ma, M. Rusch, P. Gupta, K. Boggs, B. Vadodaria, J. Dalton, Y. Liu, M.L. Valentine, L. Ding, C. Lu, R.S. Fulton, L. Fulton, Y. Tabib, K. Ochoa, M. Devidas, D. Pei, C. Cheng, J. Yang,

W.E. Evans, M.V. Relling, C.H. Pui, S. Jeha, R.C. Harvey, I.L. Chen, C.L. Willman, G. Marcucci, C.D. Bloomfield, J. Kohlschmidt, K. Mrozek, E. Paietta, M.S. Tallman, W. Stock, M.C. Foster, J. Racevskis, J.M. Rowe, S. Luger, S.M. Kornblau, S.A. Shurtleff, S.C. Raimondi, E.R. Mardis, R.K. Wilson, J.E. Dick, S.P. Hunger, M.L. Loh, J.R. Downing, C.G. Mullighan, and P. St. Jude Children's Research Hospital-Washington University Pediatric Cancer Genome, *Deregulation of DUX4 and ERG in acute lymphoblastic leukemia.* Nat Genet, 2016. **48**(12): p. 1481-1489.

8. Tanaka, Y., M. Kawazu, T. Yasuda, M. Tamura, F. Hayakawa, S. Kojima, T. Ueno, H. Kiyoi, T. Naoe, and H. Mano, *Transcriptional activities of DUX4 fusions in B-cell acute lymphoblastic leukemia.* Haematologica, 2018. **103**(11): p. e522-e526.

9. Lemmers, R.J., P.J. van der Vliet, R. Klooster, S. Sacconi, P. Camano, J.G. Dauwerse, L. Snider, K.R. Straasheijm, G.J. van Ommen, G.W. Padberg, D.G. Miller, S.J. Tapscott, R. Tawil, R.R. Frants, and S.M. van der Maarel, *A unifying*

genetic model for facioscapulohumeral muscular dystrophy. Science, 2010. **329**(5999): p. 1650-3.

10.   Snider, L., L.N. Geng, R.J. Lemmers, M. Kyba, C.B. Ware, A.M. Nelson, R. Tawil, G.N. Filippova, S.M. van der Maarel, S.J. Tapscott, and D.G. Miller, *Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene.* PLoS Genet, 2010. **6**(10): p. e1001181.

11.   Geng, L.N., Z. Yao, L. Snider, A.P. Fong, J.N. Cech, J.M. Young, S.M. van der Maarel, W.L. Ruzzo, R.C. Gentleman, R. Tawil, and S.J. Tapscott, *DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy.* Dev Cell, 2012. **22**(1): p. 38-51.

12.   Hendrickson, P.G., J.A. Dorais, E.J. Grow, J.L. Whiddon, J.W. Lim, C.L. Wike, B.D. Weaver, C. Pflueger, B.R. Emery, A.L. Wilcox, D.A. Nix, C.M. Peterson, S.J. Tapscott, D.T. Carrell, and B.R. Cairns, *Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons.* Nat Genet, 2017. **49**(6): p. 925-934.

13.	Whiddon, J.L., A.T. Langford, C.J. Wong, J.W. Zhong, and S.J. Tapscott, *Conservation and innovation in the DUX4-family gene network.* Nat Genet, 2017. **49**(6): p. 935-940.

14.	De Iaco, A., E. Planet, A. Coluccio, S. Verp, J. Duc, and D. Trono, *DUX-family transcription factors regulate zygotic genome activation in placental mammals.* Nat Genet, 2017. **49**(6): p. 941-945.

15.	Amano, T., T. Hirata, G. Falco, M. Monti, L.V. Sharova, M. Amano, S. Sheer, H.G. Hoang, Y. Piao, C.A. Stagg, K. Yamamizu, T. Akiyama, and M.S. Ko, *Zscan4 restores the developmental potency of embryonic stem cells.* Nat Commun, 2013. **4**: p. 1966.

16.	Jouhilahti, E.M., E. Madissoon, L. Vesterlund, V. Tohonen, K. Krjutskov, A. Plaza Reyes, S. Petropoulos, R. Mansson, S. Linnarsson, T. Burglin, F. Lanner, O. Hovatta, S. Katayama, and J. Kere, *The human PRD-like homeobox gene LEUTX has a central role in embryo genome activation.* Development, 2016. **143**(19): p. 3459-3469.

17.   Liu, X., Y. Wang, Y. Gao, J. Su, J. Zhang, X. Xing, C. Zhou, K. Yao, Q. An, and Y. Zhang, *H3K9 demethylase KDM4E is an epigenetic regulator for bovine embryonic development and a defective factor for nuclear reprogramming.* Development, 2018. **145**(4).

18.   Kawamura-Saito, M., Y. Yamazaki, K. Kaneko, N. Kawaguchi, H. Kanda, H. Mukai, T. Gotoh, T. Motoi, M. Fukayama, H. Aburatani, T. Takizawa, and T. Nakamura, *Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation.* Hum Mol Genet, 2006. **15**(13): p. 2125-37.

19.   Okimoto, R.A., W. Wu, S. Nanjo, V. Olivas, Y.K. Lin, R.K. Ponce, R. Oyama, T. Kondo, and T.G. Bivona, *CIC-DUX4 oncoprotein drives sarcoma metastasis and tumorigenesis via distinct regulatory programs.* J Clin Invest, 2019. **129**(8): p. 3401-3406.

20.   Meerbrey, K.L., G. Hu, J.D. Kessler, K. Roarty, M.Z. Li, J.E. Fang, J.I. Herschkowitz, A.E. Burrows, A. Ciccia, T. Sun, E.M. Schmitt, R.J. Bernardi, X. Fu, C.S. Bland, T.A. Cooper, R. Schiff, J.M. Rosen, T.F. Westbrook, and S.J.

Elledge, *The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3665-70.

21.   Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.

22.   Zhang, Y., T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, and X.S. Liu, *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.

23.   Heinz, S., C. Benner, N. Spann, E. Bertolino, Y.C. Lin, P. Laslo, J.X. Cheng, C. Murre, H. Singh, and C.K. Glass, *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.

24.   Robinson, J.T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, and J.P. Mesirov, *Integrative genomics viewer.* Nat Biotechnol, 2011. **29**(1): p. 24-6.

25.   Buenrostro, J.D., P.G. Giresi, L.C. Zaba, H.Y. Chang, and W.J. Greenleaf, *Transposition of native chromatin for fast and sensitive epigenomic profiling of*

open chromatin, DNA-binding proteins and nucleosome position. Nat Methods, 2013. **10**(12): p. 1213-8.

26. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.

27. Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S.L. Salzberg, *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biol, 2013. **14**(4): p. R36.

28. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

29. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.

30. Zhou, Y., B. Zhou, L. Pache, M. Chang, A.H. Khodabakhshi, O. Tanaseichuk, C. Benner, and S.K. Chanda, *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat Commun, 2019. **10**(1): p. 1523.

31.   Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov, *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

32.   Hystad, M.E., J.H. Myklebust, T.H. Bo, E.A. Sivertsen, E. Rian, L. Forfang, E. Munthe, A. Rosenwald, M. Chiorazzi, I. Jonassen, L.M. Staudt, and E.B. Smeland, *Characterization of Early Stages of Human B Cell Development by Gene Expression Profiling.* The Journal of Immunology, 2007. **179**(6): p. 3662-3671.

33.   Laurenti, E., S. Doulatov, S. Zandi, I. Plumb, J. Chen, C. April, J.B. Fan, and J.E. Dick, *The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment.* Nat Immunol, 2013. **14**(7): p. 756-63.

34.   Ohya, K., S. Kajigaya, A. Kitanaka, K. Yoshida, A. Miyazato, Y. Yamashita, T. Yamanaka, U. Ikeda, K. Shimada, K. Ozawa, and H. Mano, *Molecular cloning of a docking protein, BRDG1, that acts downstream of the Tec tyrosine kinase.* Proc Natl Acad Sci U S A, 1999. **96**(21): p. 11976-81.

35.   Venturutti, L., R.I. Cordo Russo, M.A. Rivas, M.F. Mercogliano, F. Izzo, R.H.

Oakley, M.G. Pereyra, M. De Martino, C.J. Proietti, P. Yankilevich, J.C. Roa, P.

Guzman, E. Cortese, D.H. Allemand, T.H. Huang, E.H. Charreau, J.A. Cidlowski,

R. Schillaci, and P.V. Elizalde, *MiR-16 mediates trastuzumab and lapatinib*

response in ErbB-2-positive breast and gastric cancer via its novel targets CCNJ

and FUBP1. Oncogene, 2016. **35**(48): p. 6189-6202.

36.   van Rhenen, A., G.A. van Dongen, A. Kelder, E.J. Rombouts, N. Feller, B.

Moshaver, M. Stigter-van Walsum, S. Zweegman, G.J. Ossenkoppele, and G. Jan

Schuurhuis, *The novel AML stem cell associated antigen CLL-1 aids in*

discrimination between normal and leukemic stem cells. Blood, 2007. **110**(7): p.

2659-66.

37.   Choi, S.H., M.D. Gearhart, Z. Cui, D. Bosnakovski, M. Kim, N. Schennum, and

M. Kyba, *DUX4 recruits p300/CBP through its C-terminus and induces global*

H3K27 acetylation changes. Nucleic Acids Res, 2016. **44**(11): p. 5161-73.

38.   Liu, L., L. Leng, C. Liu, C. Lu, Y. Yuan, L. Wu, F. Gong, S. Zhang, X. Wei, M.

Wang, L. Zhao, L. Hu, J. Wang, H. Yang, S. Zhu, F. Chen, G. Lu, Z. Shang, and

G. Lin, *An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos.* Nat Commun, 2019. **10**(1): p. 364.

39. Tian, L., Y. Shao, S. Nance, J. Dang, B. Xu, X. Ma, Y. Li, B. Ju, L. Dong, S. Newman, X. Zhou, P. Schreiner, E. Tseng, T. Hon, M. Ashby, C. Li, J. Easton, T.A. Gruber, and J. Zhang, *Long-read sequencing unveils IGH-DUX4 translocation into the silenced IGH allele in B-cell acute lymphoblastic leukemia.* Nat Commun, 2019. **10**(1): p. 2789.

40. Iwasaki, H., C. Somoza, H. Shigematsu, E.A. Duprez, J. Iwasaki-Arai, S. Mizuno, Y. Arinobu, K. Geary, P. Zhang, T. Dayaram, M.L. Fenyus, S. Elf, S. Chan, P. Kastner, C.S. Huettner, R. Murray, D.G. Tenen, and K. Akashi, *Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation.* Blood, 2005. **106**(5): p. 1590-600.

41. Certo, M., V. Del Gaizo Moore, M. Nishino, G. Wei, S. Korsmeyer, S.A. Armstrong, and A. Letai, *Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members.* Cancer Cell, 2006. **9**(5): p. 351-65.

42.   Souers, A.J., J.D. Leverson, E.R. Boghaert, S.L. Ackler, N.D. Catron, J. Chen, B.D. Dayton, H. Ding, S.H. Enschede, W.J. Fairbrother, D.C. Huang, S.G. Hymowitz, S. Jin, S.L. Khaw, P.J. Kovar, L.T. Lam, J. Lee, H.L. Maecker, K.C. Marsh, K.D. Mason, M.J. Mitten, P.M. Nimmer, A. Oleksijew, C.H. Park, C.M. Park, D.C. Phillips, A.W. Roberts, D. Sampath, J.F. Seymour, M.L. Smith, G.M. Sullivan, S.K. Tahir, C. Tse, M.D. Wendt, Y. Xiao, J.C. Xue, H. Zhang, R.A. Humerickhouse, S.H. Rosenberg, and S.W. Elmore, *ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets.* Nat Med, 2013. **19**(2): p. 202-8.

43.   Kater, A.P., M.D. Levin, and C.U. Niemann, *Ibrutinib and Venetoclax for First-Line Treatment of CLL.* N Engl J Med, 2019. **381**(8): p. 788-789.