Characterization of epigenetic and transcriptional landscape in infantile hemangiomas with ATAC-seq and RNA-seq

Xueqing Li¹ , Yuanzheng Chen², Cong Fu¹, Hongwen Li¹, Kun Yang¹, Jianhai Bi^{*,1} & Ran Huo^{**,1}

¹Department of Plastic Surgery, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, PR China

²Department of Burns & Plastic Surgery, Linyi People's Hospital, Linyi, PR China

*Author for correspondence: Tel.: +86 151 1008 0702; bijianhai@126.com

**Author for correspondence: Tel.: +86 151 6888 9001; huoran@medmail.com.cn

Aim: This study was conducted to reveal epigenetic landscape in infantile hemangiomas (IHs) and identify transcription factors (TFs) and their downstream genes active in IHs. **Materials & methods:** We performed Assay for Transposase Accessible Chromatin (ATAC-seq) with RNA-seq in three pairs of IHs and their adjacent normal tissues. Functions of candidate TFs were investigated in human umbilical vein endothelial cells (HUVECs). **Results:** Chromatin of IH tissues is less compact. Some candidate genes and TFs were identified. In HUVECs, SPDEF inhibited cell viability and tube formation, and promoted apoptosis; SOX4 exerted the opposite effect. SPDEF may act through *EPHA5*, *ZBTB46* and *SASH1*; SOX4 may act through *MMP12* and *HIVEP3*. **Conclusion:** Epigenetics plays a role in IHs. SPDEF and SOX4 may act in IHs.

First draft submitted: 13 February 2020; Accepted for publication: 13 March 2020; Published online: 30 March 2020

Keywords: ATAC-seq • infantile hemangioma • open chromatin • RNA-seq • SOX4 • SPDEF

Infantile hemangiomas (IHs) are the most common benign vascular tumors in infants. IHs are typically noticed within a few days of birth and undergo a unique course of disease, wherein a rapid proliferating phase is followed by spontaneous regression. In approximately 10–15% of cases, IHs cause disfigurement and dysfunction due to rapid proliferation of the tumor [1]. The pathogenesis of IHs, and factors that cause spontaneous regression after the proliferation remain unclear. The protocell hypothesis has been widely accepted [2,3]. Although both sporadic and familial IHs are known, no causative mutations or chromosome changes have been identified to date [4,5]. Epigenetic regulation is likely to play an important role in the development of IHs. Moreover, it has been confirmed that DNA methylation levels, along with those of miRNAs, circRNAs and lncRNAs are altered in IHs [6–9].

Nuclear DNA in eukaryotic cells undergo complex and orderly folding into chromosomes with the aid of proteins. In order for genes to be expressed, chromatin must be in open conformation. Open chromatin allows regulatory proteins to bind to DNA and modulate its functions [10]. The Assay for Transposase Accessible Chromatin with sequencing (ATAC-seq) identifies the sequences of the accessible open chromatin regions using Tn5 transposase [11,12]. This simple method requires only small amounts of the sample to yield clear and reproducible sequencing results, and closely resembles ChIP-seq [12]. Cost-effective ATAC-seq is a powerful method for obtaining information on the open chromatin regions, as well as genes that are transcribed under disease conditions and their regulatory elements [13]. Importantly, detecting changes in open chromatin can reveal the role of epigenetic regulation in IHs.

To determine the different chromatin accessibility of IHs and normal skin, and understand how gene expression is regulated in IHs, we used ATAC-seq in combination with RNA-seq, which are mutually authenticating [14,15]. By combining analysis of the regulation of the transcription process from DNA to RNA, we identified genes and transcription factors (TFs) which possibly act in the pathogenesis of IHs.

Figure 1. Pathological sections of tumors of infantile hemangiomas (n×zoom level of CaseViewer 2.0). (A) HE stain of sample 1, 10.0×. (B) Immunohistochemistry of GLUT-1 of sample 1, 20.0×. (C) HE stain of sample 2, 10.0×. (D) Immunohistochemistry of GLUT-1 of sample 2, 20.0×. (E) HE stain of sample 3, 10.0×. (F) Immunohistochemistry of GLUT-1 of sample 3, 20.0×. HE: Hematoxylin and eosin.

Table 1. Characteristics of patients.					
Sample	Sex	Age (months)	Location	Classification	Previous therapy
T1/S1	Female	9	Forehead	Mix	None
T2/S2	Male	4	Postauricular region	Mix	None
T3/S3	Female	8	Back	Mix	None

Material & methods

Samples

Proliferative hemangioma was diagnosed by two independent doctors according to clinical features and Doppler sonography. Samples were collected from March 2018 to March 2019 at the Department of Plastic Surgery, Shandong Provincial Hospital. Three paired samples were used, including tumor of proliferative IHs (abbreviated as T) and the normal tissues adjacent to but farther than 1 cm from the outermost edge of the tumors (abbreviated as S). The tumors were divided into two parts after excision: one part was flash-frozen in liquid nitrogen for subsequent experiments, and the other part was paraffin-embedded and sectioned for hematoxylin and eosin and immunohistochemical staining for glucose transporter 1 [16]. The tumors were reconfirmed by pathological examination (Figure 1), and the normal tissues were confirmed to be free of tumor invasion. General information on patients and tumors is shown in Table 1.

ATAC-seq & statistical analysis

The nuclei of cryopreserved tissue was processed according to Omni-ATAC [17]. Approximately 50,000 nuclei obtained from Optiprep (Sigma-Aldrich, MO, USA) differential centrifugation were suspended in transposition reaction mix. After purification, transposed DNA fragments were amplified for 14 cycles with primers designed according to Buenrostro and colleagues [12]. The barcode sequences of PCR Primer 2 are shown in Supplementary File 1.1. The PCR products were eluted and quantified using Agilent Bioanalyzer 2100. The standard libraries were established and sequenced on a HiSeq X Ten using 150-bp paired-end reads (PE150).

Clean reads were obtained using Trimmomatic (v. 0.26) and aligned to hg19 reference genome (RefSeq) by Bowtie 2 [18]. After filtering out non-nucleus and duplication reads, properly paired reads with mapping qualities ≥30 were used for subsequent analyses. The bam files of the T and S samples were merged by Samtools and subsequently styleDnase.pl of Homer2 was used for peaks calling [19]. The procedure to obtain differentially enriched peaks and motif enrichment was conducted according to the protocol provided by Scharer and colleagues [20].

Figure 2. Peaks and genes of ATAC-seq. (A) Distribution of insert size in reads. The short fragments represent sequence reads in open chromatin of sub-nucleosome (<150 bp), while the larger ones represent reads that span one nucleosome. (B) Reads density distribution in chromosomes. In order to visualize the distribution of reads in the gene body and its flanking 2-kb range. (C) Proportion of peaks regions matched to elements in the human genome. (D) Scatter plot of the accessibility at each peak. The red points represent T-S peaks. The green points represent S-T peaks.

Down: Downregulated; TES: Transcription end site; TSS: Transcription start site; Up: Upregulated.

RNA-seq & statistical analysis

Total RNA was extracted with TRIzol reagent (Invitrogen Life Technologies, CA, USA). RNA purity was analyzed using a NanoPhotometer (IMPLEN, CA, USA), concentration was determined using a Qubit 3.0 Fluorometer (Life Technologies, CA, USA) and quality tested with the Agilent 2100 RNA Nano 6000 Assay Kit (Agilent Technologies, CA, USA). Poly(A)⁺ RNA was separated using magnetic beads oligo (dT) from Invitrogen. A cDNA library was prepared using fragmented RNA according to protocols [21], and the RNA was used to generate double-stranded cDNA by reverse transcription with random hexamer primers.

The cDNA libraries were subjected to Novaseq 6000seq with PE150. Raw reads from Illumina were filtered to obtain clean reads, which the subsequent analysis was based on. Clean reads were mapped to RefSeq by HISATS [22]. Fragments per kilobase per million reads (FPKM) was used to quantitatively estimate gene expression values [23]. DESeq2 was used to analyze differential gene expression [24]. Genes with $|\log_2 \text{Fold change}| \geq 1$ and $\text{padj} < 0.05$ were chosen as the differentially expressed genes [23].

Preliminary function of candidate TFs

SPDEF and SOX4 were chosen as candidate TFs. Human umbilical vein endothelial cells (HUVECs) (Procell, Wuhan, China) were cultured under DMEM-high glucose (Hyclone, UT, USA) with 20% fetal bovine serum

Figure 3. Differential expressed transcripts. (A) Volcano plot. The red points in the plot represent the differentially expressed transcripts between IHs and adjacent normal tissue with statistical significance. (B) Proportion of reads matched to elements in the human genome. (C) Hierarchical cluster analysis of all the deregulated genes annotated by reads.

ATAC-seq: Assay for Transposase Accessible Chromatin using sequencing; IH: Infantile hemangioma.

(Hyclone) at 37°C and 5% CO₂. After cell attachment for 24 h, transfection was performed. There were four groups: blank group without treatment (blank group), negative control group with empty plasmid (NC group), SPDEF overexpression plasmid group (SPDEF group) and SOX4 overexpression plasmid group (SOX4 group). The pENTER transfer vector plasmid constructed with SPDEF or SOX4 genes through AsiSI/MluI cutting sites (Vigene Biosciences, Jinan, China) were used. During cell transfection, Opti-MEM (Gibco, NY, USA) and Lipofectamine 2000 Reagent (Thermo Fisher Scientific, MA, USA) were used. The experiment was performed in triplicate. See Supplementary File 2 for western blot of transfection validation. Tukey's multiple comparisons test of GraphPad Prism 8.0.2 was used to analyze the data.

After 24, 48, 72 and 96 h of transfection, MTT assay was used to detect cell viability, and absorbance was measured at 490 nm. After 48 h of transfection, flow cytometry was performed to detect apoptosis (Annexin V assay) and cell cycle (PI method). Matrigel assay was used to detect *in vitro* angiogenesis. Matrigel (10 µl) and cell suspension (50 µl; 300,000 cells/ml) were added to each well of a µ-Slide. Figures were obtained after 8 h of culture.

GRB14, *EPHA5*, *ZBTB46* and *SASH1* were primarily chosen as candidate genes of SPDEF; *MMP12* and *HIVEP3* were chosen as candidate genes of SOX4. After 48 h of transfection, real-time PCR was performed to detect changes in gene expression after HUVECs were transfected by the overexpression plasmid. Takara MiniBEST universal RNA extraction kit (Takara, Tokyo, Japan) was used for RNA extraction. Primescript RT reagent kit with gDNA Eraser (perfect real time) (Takara) was used for genomic DNA removal and reverse transcription reaction. Light cycler 480 SYBR Green I Master (Roche, Basel, Switzerland) was used for real-time PCR, and primer sequences are included in Supplementary File 1.2.

Figure 4. Enrichment analysis of genes integrated by ATAC-seq and RNA-seq. (A) Venn diagram. Overlap of genes that have the same trend in RNA-seq and ATAC-seq. **(B)** Hierarchical cluster analysis of all the deregulated genes annotated by peaks. **(C & D)** Open regions at motif of SPDEF. **(E & F)** Open regions at motif of SOX4.
ATAC-seq: Assay for Transposase Accessible Chromatin using sequencing.

Results

Mapping open chromatin of IHs with ATAC-seq

Processed data files of ATAC-seq are shown in Supplementary File 3.1–3.6. The distinct period distribution of all ATAC-seq libraries showed that most of the insert fragments were small, representing inter-nucleosomal open chromatin (Figure 2A). The accessibility of transcriptional start sites was significantly enriched (Figure 2B & C).

After merging data, we found 4970 differential peaks that represent T to be more open than S (abbreviated as T-S), and 3106 differential peaks that represent S to be more open than T (abbreviated as S-T). In general, differentially open chromatin regions showed more accessibility in IH tissue (Figure 2D). In detail, T-S peaks mapped to 3888 genes, including 708 genes that contained more than one differentially accessible region and 11 peaks mapped to no gene; S-T differential peaks mapped to 2507 genes, including 409 genes that contained more than one differentially accessible region, and 9 peaks mapped to no gene. Moreover, 839 genes had mapped peaks from both T-S and S-T. For T-S and S-T peaks which were annotated as promoters, 75.8 and 78.6% were promoters of protein-coding genes, and 19.8 and 21.4% were promoters of ncRNA genes, respectively. Additionally, the one-to-one pairing comparison of the three specimen pairs showed that the chromatin of T was more accessible than that of S.

Detecting gene expression with RNA-seq

Processed data file of RNA-seq is shown in Supplementary File 4. There were 3286 differentially expressed genes, including 1453 upregulated genes (Figure 3A). In general, most of the sequences mapped to the exon regions. Due to alternative splicing or new transcripts, as well as noisy signals, some sequences may have mapped to intron or intergenic regions, respectively (see Figure 3B). Hierarchical clustering was performed to visualize the differentially expressed genes (see Figure 3C).

Table 2. TFs binding homer known motifs enriched in chromatin regions of infantile hemangiomas.

Factor	Binding motif/name	% of target	p-value
T-S			
SPDEF	SPDEF(ETS)/VCaP-SPDEF-ChIP-Seq(SRA014231)/Homer	9.13	1e-2
S-T			
REVERB	Reverb(NR), DR2/RAW-Reverba.biotin-ChIP-Seq(GSE45914)/Homer	1.41	1e-3
TBX20	TBX20(T-box)/Heart-Tbx20-ChIP-Seq(GSE29636)/Homer	2.23	1e-2
EBF1	EBF1(EBF)/Near-E2A-ChIP-Seq(GSE21512)/HomeT2r	7.03	1e-2
SOX4	SOX4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	8.73	1e-2
GATA:SCL	GATA:SCL(Zf, bHLH)/Ter119-SCL-ChIP-Seq(GSE18720)/Homer	1.41	1e-2
SOX6	SOX6(HMG)/Myotubes-Sox6-ChIP-Seq(GSE32627)/Homer	19.11	1e-2

IH: Infantile hemangioma; TF: Transcription factor.

Integration of ATAC-seq with RNA-seq

Among the genes mapped by T-S peaks, 155 genes had the same trend in RNA-seq, indicating an intersection of mapped genes between more accessible chromatin and upregulated RNA when comparing T with S (Figure 4A). Among them, 45 genes corresponded to more than one peak (Supplementary File 5). As for S-T, 97 genes were found, including 63 genes corresponding to more than one peak (Figure 4A, Supplementary File 6). Hierarchical clustering was performed to visualize the overall integrated differential peaks (Figure 4B). To gain further insight into the function of these genes, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were analyzed (Supplementary File 7 & 8). Rap 1 signaling pathway, Ras signaling pathway, Apelin signaling pathway and MAPK signaling pathway are among the top 20 KEGG enrichment pathways.

Only 5.2 and 7.6% of the ATAC-seq peaks mapped to differentially expressed genes, for T-S and S-T, respectively. Because many of the genes mapped by ATAC-seq peaks were not differentially expressed, we sought to determine whether these differentially accessible peaks were enriched for particular TF-binding sites. Finally, known Homer TFs SPDEF, REVERB, TBX20, EBF1, SOX4, GATA:SCL and SOX6 were identified (Table 2). In the integration data, SPDEF only regulated T-S genes, and the others only regulated S-T genes. Except REVERB, the other TFs exhibited relevant expression in RNA-seq. Possible downstream genes with a motif for these TFs, following integration of results from the two analyses, are marked in Table 3. Based on literature data and integration analysis, we preliminarily considered that the possible TFs were SPDEF and SOX4. Open region patterns were observed at the sequence motifs for SPDEF and SOX4 (Figure 4C–F).

Table 3. TFs and their corresponding genes after integration.

TF	Gene name	ATAC-seq			RNA-seq	
		Peak ID	log ₂ FC	p-value	log ₂ FC	padj
SPDEF	<i>PTCHD4</i>	chr6-15305	2.05	4.89E-06	6.96	6.58E-13
	<i>LINC02109</i>	chr5-60951	1.97	7.60E-05	6.63	2.51E-04
	<i>BBOX1-AS1</i>	chr11-6662	1.55	3.87E-05	4.25	3.83E-07
	<i>AGBL4</i> [†]	chr1-1299	1.30	9.93E-05	3.71	8.10E-06
	<i>EXOC6</i>	chr10-3924	1.79	1.42E-06	3.43	2.28E-20
	<i>LINGO1</i> [†]	chr15-424	1.72	3.55E-07	3.10	3.26E-11
	<i>CLIC5</i>	chr6-9139	1.97	1.40E-06	2.93	1.54E-03
	<i>GRB14</i> [†]	chr2-22015	1.59	5.69E-05	2.93	9.55E-11
	<i>MYO1B</i>	chr2-34907	1.93	5.27E-06	2.91	8.28E-11
	<i>DGKB</i>	chr7-18113	1.81	1.99E-05	2.90	3.30E-04
	<i>NRXN3</i>	chr14-4572	1.69	1.61E-05	2.31	3.22E-05
	<i>ERG</i>	chr21-1828	1.87	1.46E-06	2.22	5.45E-05
	<i>EPHA5</i>	chr4-18836	1.97	1.99E-05	2.16	4.26E-02
	<i>GRIK3</i>	chr1-6048	1.46	2.87E-05	2.16	2.85E-09
	<i>ABCC9</i>	chr12-2075	1.58	1.29E-05	2.16	1.25E-03
	<i>PLA2G4C</i>	chr19-340	1.39	1.26E-05	1.96	3.99E-07
	<i>ZBTB46</i>	chr20-12	1.10	3.27E-05	1.57	9.68E-03
	<i>WDFY3-AS2</i>	chr4-13964	1.81	1.99E-05	1.36	3.15E-02
	<i>SASH1</i>	chr6-533	1.86	1.07E-07	1.31	6.32E-04
	<i>CD200</i>	chr3-667	1.32	3.27E-05	1.25	3.17E-02
REVERB	<i>DBI</i> [†]	chr2-10799	1.86	3.85E-06	-1.41	1.53E-02
TBX20	<i>MOXD1</i>	chr6-7214	1.51	9.61E-05	-4.78	1.14E-04
	<i>DBI</i> [†]	chr2-10799	1.86	3.85E-06	-1.41	1.53E-02
EBF1	<i>LY6D</i> [†]	chr8-41	1.45	7.58E-07	-9.02	1.67E-02
	<i>SDC1</i> [†]	chr2-291	1.36	2.19E-05	-5.77	1.20E-05
	<i>ACSBG1</i> [†]	chr15-3376	1.93	1.02E-06	-5.32	3.67E-04
	<i>ARHGEF4</i>	chr2-48012	1.71	5.25E-05	-4.54	7.44E-08
	<i>PSAT1</i>	chr9-2694	1.93	1.02E-06	-3.42	7.34E-06
	<i>LRRC4C</i> [†]	chr11-16447	1.84	1.48E-05	-3.26	2.52E-03
	<i>TSTD1</i> [†]	chr1-809	1.28	5.63E-05	-3.25	8.42E-07
	<i>DTX2</i> [†]	chr7-1092	1.39	4.71E-05	-1.60	8.48E-03
SOX4	<i>MMP12</i>	chr11-8729	1.77	3.67E-06	-8.85	3.54E-05
	<i>TYRP1</i>	chr9-18264	1.80	1.48E-05	-8.48	3.56E-07
	<i>GDA</i> [†]	chr9-3356	1.43	6.78E-05	-5.72	4.84E-03
	<i>ACSBG1</i>	chr15-3376	1.93	1.02E-06	-5.32	3.67E-04
	<i>PTPRZ1</i> [†]	chr7-18343	1.99	3.87E-06	-4.88	1.81E-04
	<i>DLG2</i>	chr11-10421	1.70	1.27E-05	-3.33	1.35E-03
	<i>LRRC4C</i> [†]	chr11-4226	1.46	6.78E-05	-3.26	2.52E-03
	<i>IMPA2</i>	chr18-125	1.32	6.90E-05	-2.88	5.92E-05
	<i>PACRG</i> [†]	chr6-2985	1.84	1.01E-06	-2.84	3.80E-03
	<i>TMEM132B</i> [†]	chr12-10867	1.62	4.22E-05	-2.78	2.59E-04
	<i>SLC35F3</i>	chr1-5282	1.62	1.02E-05	-2.59	4.04E-02
	<i>SPAG17</i> [†]	chr1-14109	2.32	3.58E-08	-2.42	4.56E-02
	<i>UST</i> [†]	chr6-7317	1.58	3.70E-05	-2.21	1.02E-02
	<i>HIVEP3</i> [†]	chr1-5889	1.43	6.78E-05	-1.73	5.80E-03
	<i>ME1</i>	chr6-16147	2.12	8.24E-07	-1.57	2.29E-02

Different peak sequences can map to the same peak ID, which means the same peak ID can have different Motifs.

[†] genes mapped by more than one peak.

ATAC-seq: Assay for Transposase Accessible Chromatin using sequencing; TF: Transcription factor.

Table 3. TFs and their corresponding genes after integration (cont.).

TF	Gene name	ATAC-seq			RNA-seq	
		Peak ID	log ₂ FC	p-value	log ₂ FC	padj
GATA:SCL	<i>TMEM132B</i> [†]	chr12-4608	2.19	4.60E-08	-2.78	2.59E-04
	<i>HS6ST1</i>	chr2-8835	1.48	9.61E-05	-1.28	4.68E-02
SOX6	<i>TECRL</i>	chr4-30398	1.82	5.67E-05	-7.00	6.05E-03
	<i>GPRIN2</i>	chr10-90	1.25	3.04E-05	-6.60	2.85E-05
	<i>GDA</i> [†]	chr9-3356	1.43	6.78E-05	-5.72	4.84E-03
	<i>LINC00958</i>	chr11-1567	2.39	3.56E-10	-5.17	7.94E-03
	<i>PTPRZ1</i> [†]	chr7-18343	1.99	3.87E-06	-4.88	1.81E-04
	<i>ASPG</i>	chr14-245	1.29	6.90E-05	-4.18	2.60E-02
	<i>KLF5</i>	chr13-7848	1.66	4.75E-05	-3.88	1.14E-03
	<i>CHL1</i>	chr3-5934	1.58	3.70E-05	-3.38	2.17E-03
	<i>LRRC4C</i> [†]	chr11-4226	1.46	6.78E-05	-3.26	2.52E-03
	<i>BMP3</i>	chr4-2969	1.66	1.15E-05	-3.18	2.29E-02
	<i>DNAH9</i>	chr17-328	1.38	3.23E-05	-3.15	4.39E-02
	<i>OCA2</i>	chr15-1473	1.58	9.01E-06	-3.15	1.03E-02
	<i>PNMA6A</i>	chrX-3333	2.20	1.76E-07	3.08	7.14E-03
	<i>PACRG</i> [†]	chr6-11545	1.92	3.93E-06	-2.84	3.80E-03
	<i>TMEM132B</i> [†]	chr12-10867	1.62	4.22E-05	-2.78	2.59E-04
	<i>PCSK6</i>	chr15-183	1.37	2.66E-05	-2.70	8.53E-05
	<i>SPAG17</i> [†]	chr1-14109	2.32	3.58E-08	-2.42	4.56E-02
	<i>UST</i> [†]	chr6-7317	1.58	3.70E-05	-2.21	1.02E-02
	<i>HIVEP3</i> [†]	chr1-5889	1.43	6.78E-05	-1.73	5.80E-03
	<i>MFAP2</i>	chr1-5601	2.00	2.56E-07	-1.71	2.64E-02
	<i>PAK3</i>	chrX-6701	1.76	5.25E-05	-1.45	4.68E-02
	<i>NAMPT</i> [†]	chr7-35170	3.40	5.95E-11	-1.37	1.37E-02
	<i>FHIT</i>	chr3-8179	1.55	3.21E-05	-1.22	3.72E-02

Different peak sequences can map to the same peak ID, which means the same peak ID can have different Motifs.

[†] genes mapped by more than one peak.

ATAC-seq: Assay for Transposase Accessible Chromatin using sequencing; TF: Transcription factor.

Preliminary function of SPDEF & SOX4

In the MTT assay, the cell viability began to show a difference from 48 h of culture. At 48 and 72 h, the cell viability of the SOX4 group was higher than that of the other groups. At 96 h, the cell viability of the SOX4 group was higher than that of the NC and SPDEF groups. The cell viability of the SPDEF group was lower than that of the SOX4 group at 48 and 96 h, and that of other groups at 72 h. The overall trend was that the cell viability of the SPDEF group decreased while that of the SOX4 group increased (Figure 5A–E). Flow cytometry showed that the SPDEF group had an increased apoptotic rate while SOX4 group was just the opposite (Figure 5F). SPDEF group had increased cells in the G0/G1 phase, and decreased cells in the S phase, while the SOX4 group was just the opposite. There was no statistical difference in the number of cells in the G2/M phase between the two groups (Figure 5G). As for *in vitro* angiogenesis, tube formation decreased in the SPDEF group and increased in the SOX4 group (Figure 5H). PCR showed that after the transfection of the SPDEF overexpressed plasmid, *EPHA5*, *ZBTB46* and *SASH1* showed increased expression ($p < 0.05$), while *GRB13* showed decreased expression ($p = 0.004$; Figure 5I–N). With the increased expression of SOX4, *MMP12* and *HIVEP3* showed increased expression ($p < 0.0005$).

Discussion

Our findings are consistent with the effect of many candidate molecules on IHs [25], and provide new insights into the open region of chromatin. The protocell hypothesis, is a generally accepted explanation for the origin of IHs, suggests that IHs are clonal tumors [1–3]. ATAC-seq in this study showed that chromatin of IH tissue to be more open than that of adjacent normal tissue, indicating that the genetic information of hemangiomas is actively expressed, or that a high proportion of the cells in IHs are transcriptionally active. Only few studies have focused

Figure 5. Function of SPDEF & SOX4 (A) Cell viability after 24, 48, 72 and 96 h of transfection. (B) 24 h. (C) 48 h. (D) 72 h. (E) 96 h. (F) Apoptosis after 48 h of transfection. (G) Cell cycle after 48 h of transfection. (H) Tube formation by Matrigel assay. (I–L) The relative mRNA levels under SPDEF overexpression. (M & N) The relative mRNA levels under SOX4 overexpression.

*p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

ATAC-seq: Assay for Transposase Accessible Chromatin using sequencing.

on the TFs of IHs. TFs, including STAT1, STAT3, STAT5, FOXO1, HOXD3, HOXD10, TBX2, have been reported sporadically and not quite systematically in IHs [26–30].

Notably, in our study, only a small number of peaks are mapped to differentially expressed genes. This is consistent with other studies and supports the view that gene activation depends on multiple regulatory regions, many of which are located far from the gene locus itself [14]. According to RNA-seq, both SPDEF and SOX4 were expressed in IHs, but did not show a statistical difference with the control group. TFs do not need to be in a state of differential expression to function. To investigate the role of SPDEF and SOX4, and their downstream genes, we performed experiments in HUVECs. HUVECs have been commonly used *in vitro* in experiments to investigate IHs and other vascular tumors or angiogenesis in tumors [31,32]. The current studies report that the androgen-responsive gene *SPDEF* is mainly related to tumors – *SPDEF* was reported as a suppressor in tumor metastasis and progression, while some studies report otherwise [33]. Reports strongly support that *SOX4* acts as an oncogene that promotes stemness, cancer cell survival, angiogenesis, migration, mesenchymal transition and metastasis. Under normal condition, *SOX4* regulates stemness, differentiation, progenitor development, working as an essential developmental TF [34]. Furthermore, *SOX4* becomes restricted as cells reach their final differentiation state, which is very alike with the life span of IHs – protocell origin, angiogenesis and tumor properties, and self-restricted growth.

According to the performance of HUVECs overexpressing the above two candidate TFs and change in the mRNA levels of candidate genes, we assumed that SPDEF could act through *EPHA5*, *ZBTB46*, or *SASH1* to act as a suppressor in HUVECs, while SOX4 has a promoting effect through regulating *MMP12* and *HIVEP3*. Decreased

EPHA5 level is associated with poorer tumor prognosis [35,36], and *SASH1* acts as a tumor suppressor [37]. *ZBTB46* likely has an inhibitory role in angiogenesis [38]. *MMPs* are involved in angiogenesis during tumor metastasis and invasion [39]. *HIVEP3* plays a critical role in female-biased autoimmune disorders and may make the tumor more progressive by changing the immune environment [40,41].

Chen and colleagues proved that SPDEF is physically associated with the promoter of *ZBTB46*; however, SPDEF transcriptionally represses *ZBTB46* in prostate cancer cells [42]. Further, *ZBTB46* acts as a tumor promoter in prostate cancer [43]. However, *ZBTB4* increases the number of cells in G0/G1 phase of the cell cycle in human aortic endothelial cells [38]. Therefore, the effect of SPDEF on *ZBTB46* in endothelial cells needs further investigation.

In most cases of IHs, by 4 months of age, approximately 80% of the tumor growth is complete, and thereafter the slope of the growth curve becomes smaller. When the child reaches 9 months of age, the tumor reaches its maximum volume, at which point apoptosis begins to prevail. Till the end of involution, protocells of IHs reach their final differentiation state. Even in one tumor, the phases of growth and regression can overlap [1]. According to our research, in IHs under middle and late proliferating phases, *EPHA5*, *ZBTB46* and *SASH1* were upregulated, while *MMP12* and *HIVEP3* were downregulated. The expression pattern of these genes is consistent with the fact that the IH specimens included had begun transitioning into regression.

Conclusion

There are some limitations of our study. The protocells of IHs are clonal cells. Although ATAC-seq uses fewer starting cells, it still cannot solve this problem. Therefore, considering individual differences, a one-to-one specimen Motif analysis without bam file merging is provided in Supplementary File 9 for your reference. Second, early stage (within 3 months) tumor specimens could complement our hypothesis.

In summary, we showed that nuclear chromatin of IHs was more open than that of normal tissues, which confirmed the neovascularized and neoplastic characteristics of IHs and that epigenetics does play a role in IHs. A total of 252 genes were identified along with SPDEF, REVERB, TBX20, EBF1, SOX4, GATA:SCL and SOX6. We suggest that the TFs SPDEF and SOX4 may be involved in the pathogenesis of IHs. Instead of causing differential expression, SPDEF may act through *EPHA5*, *ZBTB46* and *SASH1*; SOX4 may act through *MMP12* and *HIVEP3*. During the middle and late proliferating phases, unregulated *EPHA5*, *ZBTB46* or *SASH1* exerts an inhibitory effect, and the promoting effect of *MMP12* or *HIVEP3* is downregulated. Our results improve our understanding of IHs, a vascular lesion that remains a significant cause of morbidity in newborn children.

Future perspective

Study of the pathophysiology of IHs still has a long way to go, and epigenomic research seems to contribute to some extent. The method used in this study has enriched the database, proposed possible molecules and provided bam files for analysis. We speculate that both the genes and motifs identified in the integration analysis of this study can be further analyzed. Furthermore, the data presented herein can be correlated with those on histone modifications, methylation, etc. by other researchers in order to detect key molecules and their role in IHs.

Summary points

- Detecting changes in open chromatin can reveal the role of epigenetic regulation in infantile hemangioma (IHs).
- Assay for Transposase Accessible Chromatin using sequencing combined with RNA-seq, which are mutually authenticating, can reveal the regulation of the transcription process from DNA to RNA.
- Nuclear chromatin of IHs was more open than that of normal tissues, which confirmed the neovascularized and neoplastic characteristics of IHs and that epigenetics plays a role in IHs.
- A total of 252 genes were identified.
- Homer known motifs of transcription factors were SPDEF, REVERB, TBX20, EBF1, SOX4, GATA:SCL and SOX6.
- We hypothesize that SPDEF and SOX4 could be the transcription factors acting in IHs.
- SPDEF acts as a suppressor, while SOX4 has a promoting effect in human umbilical vein endothelial cells.
- The downstream genes of SPDEF may include *EPHA5*, *ZBTB46* and *SASH1*; possible target genes of SOX4 were *MMP12* and *HIVEP3*.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/epi-2020-0060

Acknowledgments

The authors would like to thank S Liu, X.biol (Jinan, China), for his technical assistance and contribution to the data processing.

Financial & competing interest disclosure

This work was supported by the National Natural Science Foundation of China (grant number 81671927). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

This study was approved by the Institutional Ethics Review Board of Shandong Provincial Hospital Affiliated to Shandong University (No. 2017-051). Written informed consents were signed by the guardians of the children.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

1. Léauté-Labrèze C, Harper JJ, Hoeger PH. Infantile haemangioma. *Lancet* 390(10089), 85–94 (2017).
2. Wong A, Hardy KL, Kitajewski AM, Shawber CJ, Kitajewski JK, Wu JK. Propranolol accelerates adipogenesis in hemangioma stem cells and causes apoptosis of hemangioma endothelial cells. *Plast. Reconstr. Surg.* 130(5), 1012–1021 (2012).
3. Munabi NC, England RW, Edwards AK *et al.* Propranolol targets hemangioma stem cells via cAMP and mitogen-activated protein kinase regulation. *Stem Cells Transl. Med.* 5(1), 45–55 (2016).
4. Castren E, Salminen P, Vikkula M, Pitkaranta A, Klockars T. Inheritance patterns of infantile hemangioma. *Pediatrics* 138(5), pii: e20161623 (2016).
5. Jinnin M, Medici D, Park L *et al.* Suppressed NFAT-dependent VEGFR1 expression and constitutive VEGFR2 signaling in infantile hemangioma. *Nat. Med.* 14(11), 1236–1246 (2008).
6. Zhang L, Zhang J, Lv R *et al.* Role of N6-methyladenine DNA modification in the pathogenesis of infantile hemangioma. *Chin. J. Plast. Surg.* 34(11), 959–964 (2018).
7. Biswas A, Pan X, Meyer M *et al.* Urinary excretion of MicroRNA-126 Is a biomarker for hemangioma proliferation. *Plas. Reconstr. Surg.* 139(6), 1277e–1284e (2017).
8. Liu X, Lv R, Zhang L *et al.* Long noncoding RNA expression profile of infantile hemangioma identified by microarray analysis. *Tumour Biol.* doi:10.1007/s13277-016-5434-y (2016) (Epub ahead of print).
9. Fu C, Lv R, Xu G *et al.* Circular RNA profile of infantile hemangioma by microarray analysis. *PLoS ONE* 12(11), e0187581 (2017).
10. Thurman RE, Rynes E, Humbert R *et al.* The accessible chromatin landscape of the human genome. *Nature* 489(7414), 75–82 (2012).
11. Adey A, Morrison HG, Asan *et al.* Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density *in vitro* transposition. *Genome Biol.* 11(12), R119 (2010).
12. Buenostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10(12), 1213–1218 (2013).
- **Compares ATAC-seq with other methods of studying open chromatin.**
13. Buenostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 109, 21.29.1–21.29.9 (2015).
- **Introduces classic protocol of Assay for Transposase Accessible Chromatin with sequencing (ATAC-seq).**
14. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol. Metab.* 5(3), 233–244 (2016).
- **Introduces the application of integration of ATAC-seq and RNA-seq.**
15. Lake BB, Chen S, Sos BC *et al.* Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* 36(1), 70–80 (2018).

16. Van Vugt LJ, Van Der Vleuten CJM, Flucke U, Blokx WAM. The utility of GLUT1 as a diagnostic marker in cutaneous vascular anomalies: a review of literature and recommendations for daily practice. *Pathol. Res. Pract.* 213(6), 591–597 (2017).
17. Corces MR, Trevino AE, Hamilton EG *et al.* An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14(10), 959–962 (2017).
- **Introduces an improved ATAC-seq protocol.**
18. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9(4), 357–359 (2012).
19. Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38(4), 576–589 (2010).
20. Scharer CD, Blalock EL, Barwick BG *et al.* ATAC-seq on biobanked specimens defines a unique chromatin accessibility structure in naive SLE B cells. *Sci. Rep.* 6, 27030 (2016).
- **We referred to the method in which ATAC-seq data are processed.**
21. Nagalakshmi U, Waern K, Snyder M. RNA-Seq: a method for comprehensive transcriptome analysis. *Curr. Protoc. Mol. Biol.* 89, 4.11.1–4.11.13 (2010).
22. Siren J, Valimäki N, Mäkinen V. Indexing graphs for path queries with applications in genome research. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 11(2), 375–388 (2014).
23. Trapnell C, Williams BA, Pertea G *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28(5), 511–515 (2010).
24. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26(1), 136–138 (2010).
25. Gomez-Acevedo H, Dai Y, Strub G, Shawber C, Wu JK, Richter GT. Identification of putative biomarkers for Infantile Hemangiomas and Propranolol treatment via data integration. *Sci. Rep.* 10(1), 3261 (2020).
26. Sulzberger L, Tan EMS, Davis PF, Brasch HD, Tan ST, Itinteang T. Phosphorylated forms of STAT1, STAT3 and STAT5 are Expressed in Proliferating but Not Involuting Infantile Hemangioma. *Front. Surg.* 5, 31 (2018).
27. Byun JW, An HY, Yeom SD, Lee SJ, Chung HY. NDRG1 and FOXO1 regulate endothelial cell proliferation in infantile haemangioma. *Exp. Dermatol.* 27(6), 690–693 (2018).
28. Zeng Z, Liu S, Cai J *et al.* miR-501 promotes hemangioma progression by targeting HOXD10. *Am. J. Transl. Res.* 11(4), 2439–2446 (2019).
29. Todorovich SM, Khan ZA. Elevated T-box 2 in infantile hemangioma stem cells maintains an adipogenic differentiation-competent state. *Dermatoendocrinol.* 5(3), 352–357 (2013).
30. Hansen SL, Dosanjh A, Young DM, Boudreau N, Hoffman WY. Hemangiomas and homeobox gene expression. *J. Craniofacial Surg.* 17(4), 767–771 (2006).
31. Bota M, Fischer-Fodor E, Bochiş OV *et al.* Combined effect of propranolol, vincristine and bevacizumab on HUVECs and BJ cells. *Exp. Ther. Med.* 17(1), 307–315 (2019).
32. Arnold A, Yuan M, Price A, Harris L, Eberhart CG, Raabe EH. Synergistic activity of mTORC1/2 kinase and MEK inhibitors suppresses pediatric low-grade glioma tumorigenicity and vascularity. *Neuro-oncology* doi:10.1093/neuonc/noz230 (2019) (Epub ahead of print).
33. Mahajan N. Signatures of prostate-derived Ets factor (PDEF) in cancer. *Tumour Biol.* 37(11), 14335–14340 (2016).
- **Reviews SPDEF.**
34. Moreno CS. *SOX4*: the unappreciated oncogene. *Semin. Cancer Biol.* doi:10.1016/j.semcancer.2019.08.027 (2019) (Epub ahead of print).
- **Reviews SOX4.**
35. Li Y, Chu J, Feng W *et al.* EPHA5 mediates trastuzumab resistance in HER2-positive breast cancers through regulating cancer stem cell-like properties. *FASEB J.* 33(4), 4851–4865 (2019).
36. Chen X, Wang X, Wei X, Wang J. EphA5 protein, a potential marker for distinguishing histological grade and prognosis in ovarian serous carcinoma. *J. Ovarian Res.* 9(1), 83 (2016).
37. Wang Q, Dai H. SASH1, a potential therapeutic target for cancer. *Hum. Pathol.* 80, 246 (2018).
38. Wang Y, Sun HY, Kumar S, Puerta MDM, Jo H, Rezvan A. ZBTB46 is a shear-sensitive transcription factor inhibiting endothelial cell proliferation via gene expression regulation of cell cycle proteins. *Lab. Invest.* 99(3), 305–318 (2019).
39. Butler GS, Overall CM. Updated biological roles for matrix metalloproteinases and new “intracellular” substrates revealed by degradomics. *Biochemistry* 48(46), 10830–10845 (2009).
40. Robbins CJ, Bou-Dargham MJ, Sanchez K, Rosen MC, Sang QA. Decoding somatic driver gene mutations and affected signaling pathways in human medulloblastoma subgroups. *J. Cancer* 9(24), 4596–4610 (2018).
41. Young NA, Friedman AK, Kaffenberger B *et al.* Novel estrogen target gene ZAS3 is overexpressed in systemic lupus erythematosus. *Mol. Immunol.* 54(1), 23–31 (2013).

42. Chen WY, Zeng T, Wen YC *et al.* Androgen deprivation-induced ZBTB46-PTGS1 signaling promotes neuroendocrine differentiation of prostate cancer. *Cancer Lett.* 440–441, 35–46 (2019).
43. Fararjeh AS, Liu YN. ZBTB46, SPDEF, and ETV6: novel potential biomarkers and therapeutic targets in castration-resistant prostate cancer. *Int. J. Mol. Sci.* 20(11), 2802 (2019).