Stem Cell Reports, Volume 16

Supplemental Information

Expression dynamics of HAND1/2 in in vitro human cardiomyocyte

differentiation

Chikako Okubo, Megumi Narita, Azusa Inagaki, Misato Nishikawa, Akitsu Hotta, Shinya Yamanaka, and Yoshinori Yoshida

Figure S1. Construction of the *HAND1***mCherry,** *HAND2***EGFP and** *MYH6***-iRFP670 triple reporter line.** Related to Figure 1.

(A and B) Southern blotting analyses of the parental hiPSC line (409B2), *HAND1*mCherry *HAND2*EGFP knocked-in double reporter line (H1mH2E) and Cre-treated double reporter line (H1mH2E+Cre) with external *HAND1* and *HAND2* probes and internal mCherry and EGFP probes, respectively. 409B2 showed two bands for the *HAND1* probe, suggesting a heterozygous mutation in the HindIII region.

(C and D) Karyotypes of parental 409B2 hiPSC and *HAND1*mCherry *HAND2*EGFP reporter lines after removal of the selection cassettes.

(E) A 10-base deletion was generated downstream of the *HAND2* coding region in the double reporter line. HA, HA-tag; 2A, 2A peptide; TSS, transcription start site; TES, transcription end site.

(F) Karyotype of the *HAND1*mCherry *HAND2*EGFP *MYH6*-iRFP670 triple reporter line.

(G) FACS plots show the expression of iRFP670 on day 20 (left) and expression of Troponin T in sorted iRFP670+ (middle) and iRFP670- populations (right).

Figure S2. Flowcytometry analysis and confocal images of the *HAND1***mCherry** *HAND2***EGFP** *MYH6* **iRFP670 triple reporter line in the EB-based differentiation protocol.** Related to Figure 1.

(A) Representative FACS plots of the expression dynamics of mCherry and EGFP from day 1 to day 20.

(B) Representative FACS plots of the expression dynamics of mCherry and iRFP670 from day 1 to day 20.

(C) Representative FACS plots of the expression dynamics of EGFP and iRFP670 from day 1 to day 20. (D) Images of EBs on days 3, 5 and 7. Scale bars, 100 µm.

(E) Images of an EB on day 20 (left). Scale bar, 100 µm. Two sections of the EB (middle and right). Scale bars, 50 µm.

Figure S3. Construction of the AAVS1-CAG-tagBFP triple reporter line from the *HAND1***mCherry** *HAND2***EGFP double reporter line.** Related to Figure 2.

(A) Scheme of the tagBFP knock-in AAVS1 locus. SA, splicing acceptor; 2A, 2A peptide; PuroR, puromycin resistance genes; pA, poly adenosine sequence; CAG, CAG promoter sequence; 5arm, 5' homology arm; 3arm, 3' homology arm. HindⅢ (H) was used to digest genomic DNA for Southern blotting. Orange lines indicate external and internal probes for Southern blotting with expected band sizes.

(B) Southern blots of the external AAVS1 probe (left) and internal puromycin resistance gene probe (right).

(C) The *HAND1*mCherry, *HAND2*EGFP and AAVS1-CAG-tagBFP triple reporter line shows normal karyotype.

Figure S4. Effects of cytokines on *HAND1* **and** *HAND2* **expression.** Related to Figure 3. (A) Representative FACS plots of mCherry expression in the parental hiPSC line (409B2) and triple reporter line on day 5.

(B) Representative FACS plots of iRFP670, mCherry (mC) and EGFP (E) expression in the triple reporter line and parental hiPSC line on day 20.

(C-E) Percentages of iRFP670+ CM subpopulations on day 20 with various concentrations of Activin A (C), BMP4 (D) and bFGF (E) in day 0-1 medium ($n = 4$ independent experiments). The ventricular protocol was used for the other stages.

(F) Percentages of mCherry+ cells on day 5, iRFP670+ cells on day 20 (left), and iRFP670+ CM subpopulations on day 20 (right) at different VEGF and IWP-3 quantities in day 3 medium ($n = 3$) independent experiments). The ventricular protocol was used for the other stages.

(G) Percentages of iRFP670+ CMs on day 20 (left) and iRFP670+ CM subpopulations on day 20 (right) with different quantities of VEGF and bFGF in day7-20 medium ($n = 3$ independent experiments). The ventricular protocol was used for the other stages.

(H-K) Percentages of iRFP670+ CMs on day 20 (left) and iRFP670+ CM subpopulations on day 20 (right) with the chemical inhibitors SB431542 (H), IWP-3 (I), Dorsomorphin (J) , and BMS 493 (K) (n = 4-6 independent experiments). Data represent means \pm SD.

Figure S5. Distribution of CM subpopulations in the atrial differentiation and monolayer differentiation protocol and profiling of the subpopulations in ventricular differentiation protocol. Related to Figures 4 and 5.

(A) Frequency of iRFP670+ CMs on day 20 with the atrial protocol (n = 4 independent experiments). *p < 0.05 , **p < 0.01 by one-way ANOVA with Dunnett's test comparing to 0 μ M.

(B) Percentages of the iRFP670+ CM subpopulations shown in A.

(C) Frequency of iRFP670+ CMs on day 20 with the monolayer protocol ($n = 4$ independent experiments).

(D) Percentages of iRFP670+ CM subpopulations in C.

(E) Representative overlayed FACS plots of the expressions of iRFP670, mCherry and EGFP from the EB protocol and monolayer protocol.

(F) Immunohistochemistry with anti α-actinin antibody of each CM subpopulation. Scale bars, 200 µm.

(G) Expression levels of immature CM marker genes (MYH6 and TNNI1) and mature CM marker genes (MYH7 and TNNI3) from the RNA-seq data of ventricular CM subpopulations on day 20. Data represent means \pm SD.

Figure S6. WNT signaling and knockdown of *HAND1***,** *HAND2* **and** *LEF1***.** Related to Figure 6.

(A) Normalized counts of *LEF1* from the RNA-seq data of ventricular iRFP670+ CM subpopulations. (n $= 3$ independent experiments). *p < 0.05, **p < 0.01 by one-way ANOVA with Tukey's multiple comparisons test. Data represent means \pm SD.

(B) Heatmap and clustering of scaled expression levels of WNT signaling molecules. Days 0, 3, 5 (isolated mCherry- and mCherry+ populations), 9 (isolated iRFP670- and iRFP670+ populations) and 20 (isolated subpopulations in i RFP670+ CMs) were collected from the ventricular CM differentiation (n = 3 independent experiments).

(C) Knockdown efficiency of HAND1 by the Wes automated capillary electrophoresis system. The expression level of HAND1 protein was normalized with total protein.

(D) Knockdown efficiency of LEF1 by western blotting. The expression level of LEF1 protein was normalized with GAPDH.

(E) Heatmap and clustering of scaled expression levels of WNT signaling molecules in the knockdown of *HAND1*, *HAND2* and *LEF1* samples. The sorted CMs were transfected at day 15 with siRNAs for negative control (siNC), *HAND1* (siHAND1), *HAND2* (siHAND2) and *LEF1* (siLEF1), and the RNAs were collected on day 20 ($n = 4$ independent experiments).

(F) ChIP-seq tracks for H3K27Ac, H3K4me3, H3K27me3, HAND1, HAND2 and/or LEF1 at 5 loci (*HAND1*, *HAND2*, *LEF1*, *CCND1* and *CCND2*) in human ESC-derived mesodermal cells, human cancer cell lines (GIST-T1, BE2C, CLB-Ga, K562) and the HEK293 cell line. The dashed red rectangle highlights the binding sites of HAND1 and HAND2 upstream of *LEF1*.

Figure S7. The utility of CD105 as a proliferative cardiomyocyte marker and schematic diagram of the expression patterns of *HAND1* **and** *HAND2* **for cardiac** *in vitro* **differentiation.** Related to Figure 7.

(A and B) Percentages of EdU+ cells on day 23 in lineage (CD90, CD31, CD49a, CD140b)-negative and SIRPA-positive CMs and CD105-APC-high or -low CMs isolated on day 20 from 692D2 (A) and 1390D4 (B) (n = 5 independent experiments). $\rm *p < 0.05$, $\rm *p < 0.01$ by unpaired t-test.

(C and D) Expression level of *HAND1* in CD105-high and -low CMs isolated on day 20 of 692D2 (C) and 1390D4 (D) ($n = 3$ independent experiments). **p < 0.01 by Welch's t-test.

(E) EdU+ ratio with TGF β signal inhibitor (SB431542) treatment (n = 4 independent experiments). All comparisons were not significant by one-way ANOVA with Dunnett's multiple comparisons test (compared to DMSO).

(F) EdU+ ratio with the administration of SB and 5 μ M IWP-3 (n = 5 independent experiments). All comparisons were not significant by one-way ANOVA with Dunnett's multiple comparisons test (compared to DMSO). Data represent means \pm SD.

(G) Schematic representation of the differential stages *in vitro* based on the expressions of *HAND1* and *HAND2* and surface markers.

Table S1. RNA-seq data of genes for each of the 4 subpopulations in day 20 iRFP670+ CMs from the ventricular protocol. Related to Figure 5. Please see associated excel file (TableS1.xlsx).

Table S2. Results of the gene set enrichment analysis of the Reactome pathway using genes in cluster 3. Significantly enriched pathways are listed (adjustment p-value < 0.05). Related to Figure 5. Please see the associated excel file (TableS2.xlsx).

Target	Cell type	Accessions number	PMID (Reference)
H3K27Ac	hESC derived mesoderm (Day 5)	GSM1505669	25693565 (Tsankov et al., 2015)
HAND1	hESC derived mesoderm (Day 5)	GSM1505812	25693565 (Tsankov et al., 2015)
HAND ₂	hESC derived mesoderm (Day 5)	GSM1505811	25693565 (Tsankov et al., 2015)
LEF1	hESC derived mesoderm (Day 5)	GSM1505691	25693565 (Tsankov et al., 2015)
H3K27Ac	GIST-T1	GSM2527250	29866822 (Hemming et al., 2018)
HAND ₁	GIST-T1	GSM2527318	29866822 (Hemming et al., 2018)
H3K27Ac	BE ₂ C	GSM3128275	30127528 (Durbin et al., 2018)
HAND ₂	BE ₂ C	GSM2486155	30127528 (Durbin et al., 2018)
H3K27Ac	CLB-GA	GSM2664317	28740262 (Boeva et al., 2017)
HAND ₂	CLB-GA	GSM2664371	28740262 (Boeva et al., 2017)
H3K27me3	K562	GSM788088	22955616 (Consortium, 2012)
H3K4me3	K562	GSE96303	22955616 (Consortium, 2012)
H3K27Ac	K562	GSM733656	22955616 (Consortium, 2012)
LEF1	HEK293T	GSE105382	22955616 (Consortium, 2012)
LEF1	K562	GSE105908	22955616 (Consortium, 2012)
LEF1	K562	GSE91682	22955616 (Consortium, 2012)

Table S3. List of ChIP-seq data. Related to Figure S6.

Table S4. List of DNA oligos/primers and vectors. Related to Figures 1, 2, S1 and S3. Please see the associated excel file (TableS4.xlsx).

Table S6. List of probes used in TaqMan Gene Expression Assays. Related to Figures 2, 4 and 6.

Target gene name	ID	
ACTB	Hs00357333 g1	
GAPDH	Hs99999905 m1	
$NKX2-5$	HS00231763 m1	
TBX5	Hs00361155 m1	
ISL1	Hs01099686 m1	
TBX20	Hs00396596 m1	
FGF ₈	Hs00171832 m1	
FGF ₁₀	Hs00610298 m1	
ISL1	Hs01099686 m1	
HCN4	Hs00975492 m1	
NR _{2F2}	Hs00819630 m1	
HAND ₁	Hs02330376 s1	
HAND ₂	Hs00232769 m1	
LEF1	Hs01547250 m1	

Table S7. List of siRNAs. Related to Figure 6.

Supplemental Experimental Procedures

Establishment of *HAND1***mCherry and** *HAND2***EGFP double reporter line**

Using the CRISPR-Cas9 system, two guide RNAs (gRNAs) were designed close to the stop codons of *HAND1* and *HAND2* genes, respectively, and cloned into pHL-H1-ccdB-mEF1a-RiH plasmid (Li et al., 2015). We constructed two targeting vectors, a FLAG-2A-mCherry (floxed PGK-Puromycin resistance) donor plasmid containing homology to the 3' of the *HAND1* gene locus and a HA-2A-EGFP (floxed-PGK-Neomycin resistance) donor plasmid containing homology to the 3' end of the *HAND2* gene locus. For homologous recombination, the homology arms were cloned 1000 bases up- and downstream of the *HAND1* and *HAND2* stop codons. For the knock-in experiments, we used 409B2 hiPSCs, which were established by the episomal method and cultured on neomycin and puromycin resistance SNL feeder cells (Okita et al., 2011). Electroporation was done with a NEPA 21 (NEPAGENE) following a previously described method with modifications (Li et al., 2015). Briefly, 5×10^5 hiPSCs were transfected with 2.5 µg Cas9 vector (pHL-EF1a-SphcCas9-iP-A), 3 µg of each targeting vector and 2.5 µg of each gRNA vector simultaneously. After 48 hours, 0.5 μ g/ml puromycin (Sigma-Aldrich) was administered for 5 days, and subsequently 50 μ g/ml G418 (Gibco) was administered until subcloning. The cloned hiPSCs were genotyped by PCR and sequenced at the junction of the homologous arms. Apart from a 10-base deletion downstream of the *HAND2* homology arm, both reporter cassettes were successfully inserted at the target sites. In addition, we performed Southern blotting to choose the hiPSC line with heterozygous knock-in for both *HAND1* and *HAND2* reporter cassettes. All DNA oligos/primers and vectors are listed in Table S4.

Removing the selection cassettes with Cre

To remove the selection cassettes, the hiPSCs were transiently transfected with a Cre-expressing vector (1 µg/ml pCAG-Cre-Blast, kindly provided by Dr. Keisuke Okita) using FuGENE HD (Promega) following the manufacturer's instruction (Table S4). After plating the hiPSCs on dishes coated with 10 µg/ml Matrigel (Corning), the cells were cultured in mouse embryonic fibroblast (MEF)-conditioned medium with 10 µg/ml Blasticidin S (Funakoshi) for 2 days. Blasticidin-resistant subclones were established, and removal of the puromycin and neomycin resistance cassettes was confirmed by PCR, sequencing and Southern blotting.

Establishment of AAVS1-CAG-tagBFP, *HAND1***mCherry and** *HAND2***EGFP triple reporter line**

For the constitutive expression of tagBFP, its coding sequence was knocked into the AAVS1 locus, which allowed for the stable expression of the CAG promoter to drive the transgene during differentiation, of double reporter hiPSCs using TALEN (Oceguera-Yanez et al., 2016). The targeting vector has a splicing acceptor with the puromycin resistance gene connected to CAG-tagBFP-pA, which was originated from AAVS1-CAG-hrGFP (a gift from Dr. Su-Chun Zhang (Addgene plasmid #52344)) (Qian et al., 2014). 2.5 µg of each targeting vector and the TALEN left- and right-arm vectors hAAVS1 1L TALEN and hAAVS1 1R TALEN (gifts from Feng Zhang (Addgene plasmid #35431 and #35432)) were transfected by electroporation as explained above for the CRISPR-Cas9 knock-in (Sanjana et al., 2012). Two days after the electroporation, tagBFP+ cells were sorted and subcloned. The subclones were confirmed by sequencing, karyotyping and Southern blotting. All DNA oligos/primers and vectors used are listed in Table S4.

Establishment of *MYH6* **reporter line**

The *MYH6*-iRFP670 *piggyBac* vector was transfected into the double reporter iPSCs described above. We confirmed normal karyotype by G-banding analysis. Cells undergoing CM differentiation, i.e., iRFP670+ cells, were 96.8% positive for Troponin T (Thermo Scientific, #MS-295-P), whereas iRFP670- cells were only 6.9% positive for Troponin T, confirming the functionality of the *MYH6* CM reporter. All antibodies used are listed in Table S5.

Fluorescence activated cell sorting (FACS) and immunostaining

For FACS, EBs were dissociated with Accumax (Innovative Cell Technologies) for 15 minutes at 37℃. EBs past day 7 were treated with collagenase type I (Sigma-Aldrich) for 4-12 hours before the dissociation. All samples for the FACS were suspended with FACS buffer (5% FBS-PBS) including 0.5 μ g/ml DAPI (Thermo Scientific) and 10 µg/ml DNase (CALBIOCHEM). DAPI+ cells were eliminated from the analysis. For reporter cells, 409B2, the parent hiPSC line of the reporter cells, was used as a negative control. For the flowcytometric analysis of day 5 EBs, the dissociated cells were stained with antibodies against human CD13-BV421, PDGFRA-BV421 and C-KIT-BV421 on ice for 30 minutes. For day 20, the cells were stained with CD105-APC antibody on ice for 30 minutes. For CMs derived from non-reporter hiPSCs, the dissociated cells were stained with CD105-APC, lineage markers-PE (CD140b, CD31, CD49a, CD90) and SIRPA-PE/Cy7 (CD170a/b) antibodies on ice for 30 minutes. All antibodies used are listed in Table S5.

Monolayer differentiation of cardiomyocytes

For the first 5 days, hiPSCs were differentiated to CMs following the EB method. On day 5, the cells were dissociated and seeded into a 24-well plate coated with fibronectin (Sigma-Aldrich) (50×10^4 cells/well) in maintenance medium including 5 ng/ml VEGF. The maintenance medium was refreshed every 2-3 days.

Mixed co-culture of tagBFP triple reporter line

To trace mCherry- cells, day 5 EBs were dissociated into single cells using Accumax for 5 minutes at 37℃, and tagBFP+ cells were sorted as mCherry- or mCherry+ (Figure 2C). Then, the tagBFP-labeled cells (0.8 \times 10⁴ cells/well) and non-labeled cells (7.2 \times 10⁴ cells/well), which were derived from the parental double reporter line, were re-aggregated to a total of 8×10^3 cells/well with 0.5% Matrigel.

Microscopy and immunostaining

The sorted cells were seeded on 24-well plates coated with fibronectin. After 4 days, the cells were fixed with 4% paraformaldehyde (Nacalai) for 30 minutes at room temperature and permeabilized with 0.5% saponin (Sigma-Aldrich). Cells stained with mouse anti-actin (1:800, Sigma-Aldrich, A7811), goat antimouse Alexa Fluor 546 (1:400, Invitrogen, A11030) and 1 µg/ml Hoechst 33342 (Invitrogen) were detected using a fluorescence microscope (KEYENCE BZ-X700) (Table S5).

Confocal imaging

For confocal imaging, the EBs were put on a glass-bottom 96-well plate (Corning) in FACS buffer with 10 µM Y-27632 to stop the beating. The expressions of mCherry, EGFP and iRFP670 in live cells were detected using a confocal microscope (A1R MP+, Nikon).

Southern blotting

Using cell lysis solution (QIAGEN), the genome DNA of the iPSCs was purified using a DNeasy Blood and Tissue Kit (QIAGEN). Genome DNA (3 µg) was digested using HindⅢ (New England Biolabs) overnight, separated on a 0.8% agarose gel, and transferred to a nylon membrane (GE health care). The membrane was incubated at 42℃ overnight with a digozigenin (DIG)-labeled DNA probe in DIG Easy Hyb buffer (Roche Life Science). After washing, the membrane was incubated in skim milk with alkaline phosphatase-conjugated anti-DIG antibody (1:10000, Anti-Digoxigenin-AP, Fab fragments, Roche Life Science). The signals were detected using CDP-*Star*® reagent (Roche Life Science) and a LAS3000 imaging system (FUJI FILM). The two probes were designed in the internal and external regions of the knock-in sequences generated by the PCR using the primer sets described in Table S4.

RNA sequencing and data analysis

Total RNAs were extracted using the miRNeasy Micro Kit and purified by RNase-Free DNase Set (QIAGEN) according to the manufacturer's protocols. Libraries were generated using 100 µg RNA and TruSeq Stranded total RNA with the Ribo-Zero Gold LT Sample Prep Kit, Set A and B (Illumina) according to the manufacturer's instruction. The NextSeq 500/550 High Output Kit v2 (75 Cycles) (Illumina) was used for sequencing. After trimming adapter sequences using cutadapt-1.15 (Martin, 2011), we removed the reads mapped to ribosomal RNA using samtools and bowtie2 (Langmead and Salzberg, 2012; Li et al., 2009). The reads mapped to the human genome (GRCh38 from the UCSC Genome Browser) using STAR (version 2.5.4a), underwent a quality check using RSeQC (version 2.6.4) (Dobin et al., 2013; Wang et al., 2012). The reads were counted using HTSeq (version 0.9.1) with the GENCODE annotation file (version 27) (Anders et al., 2014; Frankish et al., 2019). The counts were normalized using DESeq2 (version 1.24.0) in R (version 3.6.1) (Love et al., 2014). Using the DESeq2 package, PCA and likelihood ratio tests were performed. Gene clustering was performed using DEGreport (version 1.20.0) packages (Pantano, 2017). ClusterProfiler (version 3.12.0) and ReactomePA (version 1.28.0) were used for GO and pathway enrichment analysis, respectively (Yu and He, 2016; Yu et al., 2012). All heatmaps were produced using the pheatmap package. For upstream analysis, we used the geneXplain platform and the genes of cluster 2. We performed "Upstream analysis Transfac and Geneways" with default settings (Kel et al., 2006; Koschmann et al., 2015).

EdU assay

Sorted CMs were seeded on fibronectin-coated plates (5-6 \times 10⁵ cells/well in a 6-well plate or 1.0-1.5 \times 105 cells/well in a 12-well plate). After 2 days for recovery, the cells were treated with 1 µM EdU for 18 hours using the Click-iT® EdU Flow Cytometry Assay Kit (Invitrogen) following the manufacturer's instructions.

Evaluation of knockdown levels

Five days after the transfection of siRNA into purified CMs, the CMs were washed with PBS two times and then resuspended in M-PER (Thermo) containing Protease inhibitor cocktail (1:100, Nacalai). Protein concentrations were determined using the BAC assay (Bio Lad). To detect the knockdown efficiency of HAND1 by siRNA transfection, the 12-230 kDa Separation Module (Protein Simple) and Wes automated capillary electrophoresis system (Protein Simple) were used.

Anti-HAND1 (1:20, AF3168, R&D Systems) was used as the primary antibody, and HRPconjugated secondary antibodies were also used. Protein loading was normalized using the Total Protein Detection Module (Protein Simple). The data were analyzed and visualized using Compass (Protein Simple). For LEF1, conventional western blotting was performed. In brief, the proteins were separated in 10-20% gel (FUJIFILM) with 10x Tris/Glycine/SDS Buffer (BIOLAD) and transferred to a nitrocellulose membrane (Merck Millipore) with 10x Tris/Glycine Buffer (BIOLAD). The membrane was blocked with 5% skim milk (FUJIFILM) for 60 min at room temperature with constant agitation and then incubated with primary anti-LEF1 (1:1000, A303-486A, BETHYL) at 4 ℃ overnight. The membrane was washed three times with Tris-buffered saline. Then, the membrane was incubated with secondary antibody (1:5000, goat anti-rabbit IgG-HRP, sc-2004, Santa Cruz Biotechnology) diluted in Bullet Blocking One for Western Blotting (Nacalai tesque) for 1 hour at room temperature. After the secondary antibody reaction, the membrane was incubated in chemiluminescent HRP substrate for 5 min. Images were obtained using an ImageQuant LAS 4000 (Cytiva).

Chromatin Immunoprecipitation (ChIP) data analysis

All ChIP-seq data (wig and bed files) of HAND1, HAND2, LEF1, H3K27Ac, H3K4me3 and H3K27me3 immunoprecipitated samples mapped to hg19 were obtained from Gene Expression Omnibus (GEO) by NCBI and visualized by Integrative Genomics Viewer (Robinson et al., 2011) . The files are listed in Table S3.

Chemical inhibition and activation of signaling

To inhibit and activate WNT signaling, IWP-3 and CHIR99021 (FUJIFILM), respectively, were added to the maintenance medium from day 16 to day 18. For the knockdown experiments, these compounds were administered from day 17 to day 19. To investigate the effect of the CM sub-population distribution, IWP-3, SB, Dorsomorphin, and BSM 492 were administered from day 7 to day 20.

Supplemental References

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq—a Python framework to work with highthroughput sequencing data. Bioinformatics 31, 166-169.

Boeva, V., Louis-Brennetot, C., Peltier, A., Durand, S., Pierre-Eugene, C., Raynal, V., Etchevers, H.C., Thomas, S., Lermine, A., Daudigeos-Dubus, E., et al. (2017). Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. Nat Genet 49, 1408- 1413.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21.

Durbin, A.D., Zimmerman, M.W., Dharia, N.V., Abraham, B.J., Iniguez, A.B., Weichert-Leahey, N., He, S., Krill-Burger, J.M., Root, D.E., Vazquez, F., et al. (2018). Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. Nat Genet 50, 1240-1246.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 47, D766-D773.

Hemming, M.L., Lawlor, M.A., Zeid, R., Lesluyes, T., Fletcher, J.A., Raut, C.P., Sicinska, E.T., Chibon, F., Armstrong, S.A., Demetri, G.D., et al. (2018). Gastrointestinal stromal tumor enhancers support a transcription factor network predictive of clinical outcome. Proc Natl Acad Sci U S A 115, E5746-E5755.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

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

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17, 10.

Pantano, L. (2017). DEGreport: Report of DEG analysis.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24-26.

Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A., and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. Nature 518, 344-349.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics 28, 2184-2185.

Yu, G., and He, Q.Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst ¹², 477-479.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing

biological themes among gene clusters. OMICS 16, 284-287.