nature research

Corresponding author(s): Nate Hathaway

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	The software for flow cytometry acquisition is Attune NxT 2.6. For sequencing analysis: Illumina barcoded libraries were made and sequenced on an Illumina NextSeq 500, to a depth of 30 million base pairs (150 bp paired end).
Data analysis	The software for flow cytometry analysis is FlowJo 10.4.2. For sequencing analysis reads were mapped to UCSC reference for hg19 using STAR34, then read abundance for each gene counted (TPM method) using RSEM. Data visualization was performed using custom R scripts (available here: https://github.com/GryderArt/VisualizeRNAseq/tree/master/RNAseq_Pipeline).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

No large data sets were generated or analyzed during this current study, other than ChIP-seq and RNA-seq, availability is described in that section below.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. Each condition was analyzed with 3 biological replicates. This is typically the standard for the experiments performs so that a reasonable range of variability between samples can be accounted for.
Data exclusions	All biological replicates are presented as indicated, with the number defined for each experiment.
Replication	All attempts of replication showed a consistent pattern and were deemed successful. Multiple independent experiments were conducted as reported to ensure reproducibility of the experimental findings were validated.
Randomization	Samples of cells were sorted into pre-labeled tubes. Following sorting into tubes, the order of sample preparation was randomized and not chosen in a specific order.
Blinding	Allocation was not blinded to ensure that full sets of biological replicates were analyzed together, rather than multiple biological replicates of a few sample conditions.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	APC anti-human CD184 (CXCR4) was used for flow cytometry (from Biolegend catalog: 306510). BRD4 CHIP-grade antibody was used for ChIP qPCR and ChIP sequencing (Abcam catalogue #ab128874).
Validation	For the CXCR4 antibody, we optimized the incubation conditions in the 293t cells through serial dilutions and determining the concentration at which signal plateaus. For the BRD4 antibody, we used the amount of antibody recommended by the ChIP-IT High Sensitivity Kit from Active Motif (catalogue #53040) and used negative control qPCR primer sets.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	293 [HEK-293] (ATCC® CRL-1573™) Obtained from the UNC Tissue Culture Facility ; Lenti-X™ 293T cell line obtained from Clontech ; HCT116 (ATCC® CCL-247™) Obtained from the UNC Tissue Culture Facility. RH4 human cancer cell line obtained from Dr. Peter Houghton.
Authentication	The cell lines used were not authenticated in our hands.
Mycoplasma contamination	Cell lines in Hathaway lab are tested for mycoplasma contamination by PCR every 6 months or at least annually. All lines to date tested negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://github.com/GryderArt/ChIPseqPipe
Files in database submission	Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7 Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7 Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7
Genome browser session (e.g. UCSC)	https://genome.ucsc.edu

Methodology

Replicates	One replicate was used for each condition.
Sequencing depth	Instrument mode: Illumina NextSeq 475 Read length: 75 Single end
Antibodies	BRD4 CHIP-grade antibody was used for ChIP qPCR and ChIP sequencing (Abcam catalogue #ab128874).
Peak calling parameters	GEO.name bed.path fastq.path bed fastq HEK sgMYOD cntrl BRD4 /khanlab/projects/ChIP_seq/DATA/Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7/ MACS_Out_p_1e-07/Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed /khanlab/ projects/ChIP_seq/DATA/Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7/ Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz HEK sgMYOD CEM87 BRD4 /khanlab/projects/ChIP_seq/DATA/Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7/ MACS_Out_p_1e-07/Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed / khanlab/projects/ChIP_seq/DATA/Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7/ Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz HEK sgNT CEM87 BRD4 /khanlab/projects/ChIP_seq/DATA/Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7/ MACS_Out_p_1e-07/Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed /khanlab/ projects/ChIP_seq/DATA/Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7/ Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7_peaks.narrowPeak.nobl.GREAT.bed Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7_R1.fastq.gz
Data quality	SampleFile PEAKS_PVALUE PEAKS_FOLD_CHANGE5 Sample_HEK_sgMYOD_cntrl_BRD4_037_C_HMJ2HBGX7 21901 21612 Sample_HEK_sgMYOD_CEM87_BRD4_037_C_HMJ2HBGX7 1879 1851 Sample_HEK_sgNT_CEM87_BRD4_037_C_HMJ2HBGX7 523 479
Software	ChIP enriched DNA reads were mapped to reference genome using BWA version 0.7.10 The ChIP-seq and DNase-seq peaks were called by MACS2 version 2.1.0. Enrichment of known and de novo motifs were found using HOMER version 4.8.2

https://github.com/GryderArt/ChIPseqPipe

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were washed with 1X PBS, dissociated, quenched with media. Cells were spun down, washed with 1X PBS, and spun down, and lastly resuspended the cell pellet in 400 uL of FACS buffer. For the experiments with APC-CXCR4, cells were incubated in 8 uL per mL of blocking buffer for 1 hour in the dark at room temperature.
Instrument	The Thermo Fisher Attune NxT is equipped with a 96/384 well plate Autosampler and 4 lasers (405, 488, 561, and 637nm) and can detect up to 14 fluorescent parameters. Its acquisition software is Attune NxT 2.6.
Software	The analysis software that was used is FlowJo.
Cell population abundance	At least 100,000 cells were obtained for each sample and technical replicate.
Gating strategy	Gating was performed using FSC/SSC area, followed by SSC-height/SSC-area. The same gate was applied to all samples of cells within the same experiment.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.