nature portfolio

Corresponding author(s): Ioannis Sanidas

Last updated by author(s): Mar 18, 2026

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☐	☒ A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We utilized the following software and algorithms for data collection: trim-galore (v0.6.4_dev), cutadapt (v2.8), FastQC (v0.11.9), bwa (v0.7.17-r1198-dirty or v0.7.17-r1188), bowtie (v2.3.5.1), sambamba (v0.7.1), samtools (v1.10), htslib (v1.14), bedtools (v2.27.1), deeptools (v3.5.1), pyGenomeTracks (v3.8), STAR (v2.7.1a), STAR-Fusion (v2.7.1a), RSEM (v1.3.1), MACS (v2.2.7.1), bedGraph2Cluster (v5.3), PanChIP (v3.0.10), HOMER (v4.11.1), MEME Suite (v5.3.3), Juicer Tools (v1.22.0), JuiceBox (v1.11.08), HiCCUPS (v1.22.0), HiCCUPS Diff (v1.22.0), APA (v1.22.0), cooler (v0.9.2), coolpup.py (v1.0.0), cooltools (v0.5.2), hic2cool (v0.5.1), HiCExplorer (v2.2.1.1), Stripenn (v1.1.65.15), FitHiChIP (v11.0), ROSE (v0.1), RepEnrich (v2)
Data analysis	Source codes utilized for the bioinformatics analyses are publicly available at our GitHub repository [https://github.com/hanjunlee21/Lee-et-al-2023], which is archived at Zenodo [https://doi.org/10.5281/zenodo.10062483]. We utilized the following software and algorithms for data anaylses: OriginPro (v2023b), FCS Express (v7.18.0025), MATLAB (vR2018b or vR2023a), MATLAB Runtime (v9.5), Perl (v5.30.0), Python (v2.7.18 or v3.8.10), get_qc.py, NumPy (v1.22.3), Pandas (v1.5.2), Seaborn (v0.11.1), Matplotlib (v3.4.2), Bioframe (v0.3.3), R (v4.3.0), tximportData (v1.28.0), tximport (v1.28.0), readr (v2.1.4), DESeq2 (v1.40.2), Integrative Genomics Viewer (v2.16.2), LDC (v1.20.1), DMD (v2.090.1), LLVM (v9.0.1), PGS (v0.0.1), PyMOL (v2.5.2), POV-Ray (v3.7.0.msvc10), ImageJ (v1.54d)

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Source data are provided as a Source Data file. The RB, E2F1, CTCF, c-Jun, H3K4me1, H3K4me3, and H3K27ac ChIP-seq from asynchronous hTERT-RPE1 cells are available at GEO under the Series GSE17603523. The SMC3, acSMC3, H3K27ac, and H3K27me3 ChIP-seq, ATAC-seq, MNase-seq, and RNA-seq datasets from hTERT-RPE1 cells are available at GEO under the SuperSeries GSE246946. Individual accession codes are as follows: ATAC-seq, GSE246942; ChIP-seq, GSE246943; RNA-seq, GSE246944; and MNase-seq, GSE246945. Additionally, STAG1, STAG2, and MED15 ChIP-seq from G1-arrested hTERT-RPE1 cells and SMC3 ChIP-seq from Eg5 inhibitor-treated hTERT-RPE1 cells are available at GEO under the Series GSE315529. Single cell RNA-seq from RBWT and RBKO hTERT-RPE1 cells are available at GEO under the Series GSE315528. The Micro-C dataset is available at the EMBL-EBI ArrayExpress database under the accession code E-MTAB-13566, and the processed .mcool files are available at EMBL-EBI BioStudies database under the accession code S-BSST1221. The H3K27ac HiChIP in G1-arrested hTERT-RPE1 cells is available at GEO under the Series GSE266412. All mass spectrometry data can be accessed through the MassIVE data repository under the accession number MSV000098256.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Not applicable
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable
Population characteristics	Not applicable
Recruitment	Not applicable
Ethics oversight	Not applicable

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	We utilized two different RB knockout clones and at least two different methods of cell cycle arrest for our main observations. For phenotypic assays, we performed n>=8 samples per experiment. These sample sizes were sufficient to assess whether RBKO cells compared to RBWT cells exhibit differential signals in RNA-seq, ChIP-seq, and other phenotypic assays. No statistical method was used to predetermine sample size.
Data exclusions	We included all data collected for analyses, and no data were excluded from the analyses.
Replication	For ChIP-seq, HiChIP, RNA-seq, and proteomics experiments, we performed biological replicates and assessed statistical significance based on replication.
Randomization	Not applicable. Present study is based on computational analyses of ChIP-seq, ATAC-seq, RNA-seq, Micro-C, and HiChIP experiments.
Blinding	Not applicable. The differential analysis of ChIP-seq, ATAC-seq, RNA-seq, Micro-C, and HiChIP experiments was done in silico via computational algorithms, which are blinded to the conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Antibodies

Antibodies used

anti-DYKDDDDK tag (Cell Signaling Technology, Cat. #14793, RRID: AB_2572291), anti-FLAG (M5) (Sigma-Aldrich Cat. #F4042, RRID: AB_439686), anti-RB (Cell Signaling Technology, Cat. #9309S, RRID: AB_10696874), anti-H3K4me3 (Cell Signaling Technology, Cat. #9751, RRID: AB_2616028), anti-H3K4me1 (Abcam, Cat. #ab8895, RRID: AB_306847), anti-H3K27ac (Active Motif, Cat. #39034, RRID: AB_2722569), anti-H3K27ac (Cell Signaling Technology, Cat. #8173, RRID: AB_10949503), anti-H3K27me3 (Cell Signaling Technology, Cat. #12158, RRID: AB_2797834), anti-E2F1 (Cell Signaling Technology, Cat. #3742, RRID: AB_2096936), anti-CTCF (Cell Signaling Technology, Cat. #3418, RRID: AB_2086791), anti-c-Jun (Cell Signaling Technology, Cat. # 9165, RRID: AB_2130165), anti-SMC3 (Abcam, Cat. #ab9263, RRID: AB_307122), anti-acSMC3 K105/K106 (Sigma-Aldrich, Cat. #MABE1073, RRID: N/A), anti-α-tubulin (Sigma-Aldrich Cat. #T9026, RRID: AB_477593

Validation

We utilized INPUT controls for all ChIP-seq experiments.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

Non-transformed human retinal pigment epithelial cells immortalized by human telomerase reverse transcriptase (hTERT1-RPE1) were a gift from Dr. David Pellman's laboratory (Dana Farber Cancer Institute, RRID: CVCL_4388). Human embryonic kidney cells (HEK293T) were obtained from ATCC (Cat. # CRL-3216, RRID: CVCL_0063).

Authentication

hTERT1-RPE1 cells were authenticated by DNA sequencing. For the validation of RB1 mutations, RBWT and RBKO hTERT-RPE1 cells were harvested from culture, and approximately 5 million cells per sample were utilized. After the chromatin was isolated from cells, the chromatin was digested with micrococcal nuclease. 1–2 ng MNase-digested DNA samples were used to prepare sequencing libraries, using the Ovation Ultralow System V2 kit (TECAN, Cat. # 0344NB-A01) and were subsequently sequenced using the NovaSeq 6000 sequencer (Illumina) using the S4 flow cell. Adapter trimming was performed using trim-galore in a paired-end mode to remove Illumina adapter sequences. Sequencing reads were then aligned using bwa mem (v0.7.17-r1188) to the hg38 human reference genome build. After the aligned .bam file was sorted and indexed using samtools (v1.10), sambamba (v0.7.1) was utilized to sort, mark duplicate reads, and index .bam files. Then, we filtered reads in the .bam files to produce a file that only contains reads that are not PCR duplicates, with less than or equal to two mismatches, and with mapping quality that is greater than or equal to 30.

Mycoplasma contamination

These cells were negative for Mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

Not applicable

Plants

Seed stocks

Not applicable

Novel plant genotypes

Not applicable

Authentication

Not applicable

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.*	The RB, E2F1, CTCF, c-Jun, H3K4me1, H3K4me3, and H3K27ac ChIP-seq from asynchronous hTERT-RPE1 cells are available at GEO under the Series GSE17603523. The SMC3, acSMC3, H3K27ac, and H3K27me3 ChIP-seq from hTERT-RPE1 cells are available at GEO under the SuperSeries GSE246946. Additionally, STAG1, STAG2, and MED15 ChIP-seq from G1-arrested hTERT-RPE1 cells and SMC3 ChIP-seq from Eg5 inhibitor-treated hTERT-RPE1 cells are available at GEO under the Series GSE315529.
Files in database submission	Not applicable
Genome browser session (e.g. UCSC)	Not applicable

Methodology

Replicates	For ChIP-seq experiments, biological replicate experiments were performed and their number of replicates are indicated in each figure.
Sequencing depth	Sequencing depth for ChIP-seq experiments was ~30 million reads per condition in a paired-end manner. The read length was 150 bp.
Antibodies	Anti-DYKDDDDK Tag (Cell Signaling, Cat. # 14793) was used for FLAG-RB ChIP, and for other proteins, their primary antibodies listed in the Antibodies subheadings were utilized for ChIP: anti-DYKDDDDK tag (Cell Signaling Technology, Cat. #14793, RRID: AB_2572291), anti-FLAG (M5) (Sigma-Aldrich Cat. #F4042, RRID: AB_439686), anti-RB (Cell Signaling Technology, Cat. #9309S, RRID: AB_10696874), anti-H3K4me3 (Cell Signaling Technology, Cat. #9751, RRID: AB_2616028), anti-H3K4me1 (Abcam, Cat. #ab8895, RRID: AB_306847), anti-H3K27ac (Active Motif, Cat. #39034, RRID: AB_2722569), anti-H3K27ac (Cell Signaling Technology, Cat. #8173, RRID: AB_10949503), anti-H3K27me3 (Cell Signaling Technology, Cat. #12158, RRID: AB_2797834), anti-E2F1 (Cell Signaling Technology, Cat. #3742, RRID: AB_2096936), anti-CTCF (Cell Signaling Technology, Cat. #3418, RRID: AB_2086791), anti-c-Jun (Cell Signaling Technology, Cat. # 9165, RRID: AB_2130165), anti-SMC3 (Abcam, Cat. #ab9263, RRID: AB_307122), anti-acSMC3 K105/K106 (Sigma-Aldrich, Cat. #MABE1073, RRID: N/A), anti-α-tubulin (Sigma-Aldrich Cat. #T9026, RRID: AB_477593
Peak calling parameters	Peak calling was performed on MACS2 with default parameters. The following parameters for FitHiChIP were utilized: BINSIZE, 100 kb; LowDistThr, 200 kb; UppDistThr, 10 Mb; UseP2PBackgrnd, 0; BiasType, 1; MergeInt, 1; and QVALUE, 0.01.
Data quality	Default parameters were used for MACS2. The ChIP-seq peaks between biological replicate experiments were highly correlated (Fig. 3a-c).
Software	Adapter trimming was performed using trim-galore in a paired-end mode to remove Illumina adapter sequences. Sequencing reads were then aligned using bwa mem (v0.7.17-r1188) either to the hg38 human reference genome build or the chimera build of hg38 human reference genome build and NZ_CP080620.1 Escherichia coli TOP10 strain reference genome build. After the aligned .bam file was sorted and indexed using samtools (v1.10), sambamba (v0.7.1) was utilized to sort, mark duplicate reads, and index .bam files. Then, we filtered reads in the .bam files to produce a file that only contains reads that are not PCR duplicates, with less than or equal to two mismatches, and with mapping quality that is greater than or equal to 30. After indexing the .bam file, deeptools bamCoverage (v3.5.1) was utilized to normalize the number of reads based on counts per million (CPM), and .bigWig files were generated at a 1 bp resolution. Peak calling was performed as previously described (15), using bedGraph2Cluster with the following parameters: fold change, 2; normalization, QNorm; distance method, sqeuclidean; and clustering method, symmetry-collapsed profile and scalar. Alternatively, MACS (v2.2.7.1) was utilized to call peaks. For samples with spike-in controls, spike-in normalization was performed by multiplying the data column in each bedGraph file with a scaling factor. The scaling factor was determined by dividing the number of reads that are aligned to the NZ_CP080620.1 Escherichia coli TOP10 strain reference genome build by the total number of aligned reads and further multiplying the value by 1,000,000 and rounding it. For the assessment of repetitive sequences, sequencing reads were realigned to the hg38 human reference genome build via bowtie (v2.3.5.1) and were analyzed using RepEnrich (v2) using the Python programming language (v2.7.18). The fold change upon RB loss was assessed using the read counts calculated from RBWT and RBKO samples in hTERT1-RPTE1 cells, and their z-scores were measured. Statistical significance was measured using the z-score, and the p-values were utilized to generate the Manhattan plot. For the identification of cohesin gain and cohesin loss gene set, a custom MATLAB code was utilized. The code is available at our GitHub repository (https://github.com/hanjunlee21/Lee-et-al-2023). In brief, the list of promoters from Ensembl (v105) and the list of enhancers from histone modification-based clustering analysis of H3K27ac ChIP-seq peaks were assessed. Promoters that were non-RB-bound and non-E2F1-bound and those that were within 60 kb from the nearest enhancer were utilized.

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	The cell cycle was analyzed for hTERT1-RPE1 cells in various conditions. Cells were incubated with 20 µM 5-ethynyl-2'-deoxyuridine (EdU; Life technologies, Cat. # A10044) for 2 h, fixed with 3.7% formaldehyde in PBS for 15 min, blocked with 3% BSA in PBS for 1 min, permeabilized with 0.5% Triton X-100 in PBS for 30 min, and stained with ClickIT reaction (100mM Tris-HCl pH 7.5, 3mM CuSO4, 50 mM Ascorbic Acid, and 2.5 µM Alexa Fluor-647 azide; Life Technologies, Cat. # A10277) for 30 minutes and 3 µM 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI; Life Technologies, Cat. # D1306) in staining buffer (100mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM CaCl2, 0.5 mM MgCl2, and 0.1% NP-40) for 15 minutes.
Instrument	Flow cytometry experiments were performed with LSR II flow cytometer (BD Biosciences).
Software	Flow cytometry experiments were analyzed using the FCS Express software (v7.18.0025).
Cell population abundance	Information regarding each cell population for each gate is available in Supplementary Fig. 2.
Gating strategy	All gating strategy is described in Extended Data Fig. 2. In short, we performed FSC-A/FSC-H and FSC-A/SSC-A gating to identify live singlet cells, and then we performed DAPI/EdU staining for cell cycle analysis.

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