PNAS

Supporting Information for

The human adenovirus E1B-55K oncoprotein coordinates cell transformation through regulation of DNA-bound host transcription factors.

Konstantin von Stromberg[1]*, Laura Seddar[1], Wing-Hang Ip[1], Thomas Günther[2], Britta Gornott[1], Sophie-Celine Weinert[1], Max Hüppner[1], Luca D. Bertzbach[1] and Thomas Dobner[1]*

[1] Leibniz Institute of Virology (LIV), Department of Viral Transformation,
Martinistraße 52, 20251 Hamburg, Germany.
[2] Leibniz Institute of Virology (LIV), Virus Genomics,
Martinistraße 52, 20251 Hamburg, Germany.

* Corresponding authors: konstantin.stromberg@leibniz-liv.de, thomas.dobner@leibniz-liv.de

This PDF file includes:

Generation, production and titration of lentiviral vectors
Lentiviral gene ontology (LeGO) vectors (1) were used to deliver adenoviral oncogene sequences. PCR-amplified genomic HAdV-C5 E1A (GenBank: AY339865; nucleotides (nt) 560 to 1,545) and E1B comprising E1B-19K (GenBank: AAQ19286; nt 1,714 to 2,244) and E1B-55K (GenBank: AAQ19287 nt 2,019 to 3,509) or HAdV-A12 E1B-55K (GenBank: X73487, nt 1,847 to 3,295) were ligated into the respective lentiviral vectors. E1A was cloned into the LeGO-iVLN2 backbone, which allows continuous mRNA expression of the gene of interest together with an internal ribosomal entry site (IRES)-linked venus-neomycin resistance fusion protein under the control of the spleen focus-forming virus (SFFV) promoter. An HA-tag was fused to the N-terminus of any E1B-55K protein via site-directed mutagenesis, thereby deleting its original start codon. HA-tagged E1B-55K proteins and the respective mutant sequences were cloned into the LeGO-iBLB2 vector backbone, analogous to the aforementioned vector, but contains a blue fluorescent protein (BFP)-blasticidin-S resistance gene. Lentiviral particles were produced according to the method described by Weber *et al*. (1). Mutations in the LeGO-iBLB2-E1B plasmid were introduced via site-directed mutagenesis using the oligonucleotides shown in Table S1. The HAdV-C5 E1B-55K mutant K104R contains a single amino acid exchange (K to R) at position 104 (2, 3), whereas HAdV-C5 E1B-55K K101R contains a K to R exchange at position 101 (3, 4). The HAdV-C5 E1B-55K NES carries three amino acid exchanges at positions 83, 87, and 91 (leucines to alanines) (5, 6). The mutations were confirmed by Sanger sequencing.

Transduction of BRK cells
BRK cells were seeded in 6-well plates (Sarstedt) 16-24 h at 80% confluence before lentiviral transduction. Adherent cells were transduced twice, at a multiplicity of infection of 1 and with a 24 h interval between each transduction. The transductions were performed in DMEM supplemented with 10% FCS, 20 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, Sigma-Aldrich) and 8 µg/ml polybrene (Millipore). After 48 h cells were harvested and sorted by fluorescence-activated cell sorting. Double-positive cells (BFP and venus-positive cells) were seeded onto an appropriate plate depending on the cell number. The medium was changed every 48 h for 14 days and the cells were subsequently expanded into polyclonal cell lines.
In the case of the defective HA-tagged E1B-55K mutants K104R the standard transduction protocol was unsuccessful and reverse transduction was applied followed by antibiotic selection. Briefly, 1 ml cell suspension (approximately 350,000 cells) was seeded onto 6-well plates containing 100 µl of lentiviral concentrate. Transduction was performed in DMEM supplemented with 10% heat-inactivated FCS as well as the abovementioned supplements. Cells were incubated with the lentiviruses for 48 h. For selection, 50 µg/ml blasticidin-S (Invitrogen) and 200 µg/ml G418 (Calbiochem) were applied for 4 weeks. The medium was changed every 3 to 4 days. Occurring foci were pooled and expanded into polyclonal cell lines.

Western blotting
Cells were lysed in radioimmunoprecipitation assay buffer (RIPA) buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA,1 mM dithiothreitol (DTT), 0.1% sodium dodecyl sulphate, 1% Nonidet P-40, 0.1% Triton X-100, 0.5% sodium deoxycholate) containing freshly added 1% phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich), 0.1% aprotinin (Sigma-Aldrich), 1 µg/ml leupeptin (Roche), 1 µg/ml pepstatin (Biomol) at 4°C. Protein concentrations were measured with the Bradford Reagent-based Bio-Rad protein assay (Bio-Rad). Equal amounts of total protein were separated by SDS-polyacrylamide gel electrophoresis and transferred to 0.45 µm nitrocellulose membranes (Amersham) by western blotting. These membranes were blocked overnight in phosphate-buffered saline (PBS) containing 5% non-fat dry milk powder. Afterwards, membranes were incubated for 2 h in PBS containing the appropriate primary antibody. Proteins of interest were visualized after 2 h incubation in PBS containing 3% non-fat dry milk powder and 0.1% Tween 20 (PBS-T) with the matching secondary HRP-conjugated antibodies followed by enhanced chemiluminescence (ECL) on medical X-ray films (CEA RP). Autoradiograms were scanned and cropped using Adobe Photoshop CS6 and figures were prepared using Adobe Illustrator CS6.

Antibodies
Primary antibodies specific for HAdV-C5 proteins included E1A mouse monoclonal antibody (mAb) M73 (1:10) (7), E1B-55K mouse mAb 2A6 (1:10) (8), and E1B-19K (1:4,000) (9). Primary antibodies against specific cellular proteins comprise rabbit polyclonal antibody (pAb) p53 (1:1,000, sc-6243, Santa Cruz) or p53 mouse mAb (1:10, PAb421; a generous gift from A. Levine, Princeton University, Princeton), mouse mAb p21 (1:1,000, sc-6246, Santa Cruz) and mouse mAb β-actin (1:5,000, A-5441, Sigma-Aldrich). The p53 mouse mAb was only used for immunofluorescence analysis due to discontinuation and replacement of the antibody during the course of the experiments. Secondary antibodies conjugated with horseradish peroxidase (HRP) against mouse or rabbit IgG were obtained from Jackson ImmunoResearch. Polyclonal Alexa Fluor 488- and 555-labeled secondary antibodies

were obtained from Invitrogen. For ChIP-seq, 4 µg control polyclonal control rabbit IgG antibody (#2729, Cell Signaling) or 4 µg anti-HA rabbit pAb (ab9110, Abcam) per individual ChIP were used.

Indirect immunofluorescence

Cells were seeded onto 18×18 mm glass coverslips placed in 6-well plates and fixed 24 h post plating with 4% paraformaldehyde (PFA) at room temperature (RT) for 15 min. PFA was aspirated and coverslips were blocked for 45 min with Tris-buffered saline-BG (TBS-BG; 20 mM Tris/HCl (pH 7.6), 137 mM NaCl, 3 mM KCl, 1.5 mM $MgCl_2$, 0.05% [v/v] Tween 20, 0.05% [w/v] NaN_3, 5% [w/v] glycine, 5% [w/v] BSA). The blocking solution was discarded and cells were incubated for 1 h with the respective primary antibodies diluted in PBS. The samples were washed three times with PBS-T and subsequently incubated with the corresponding Alexa-Fluor 488- or 555-conjugated secondary antibodies for 1 h at 4°C in a dark chamber. After three final washing steps, coverslips were mounted in Rotimount (Dako) and digital images were acquired on a Nikon A1 confocal microscope (with a Nikon Ti2 frame). In order to visualize A12 E1B-55K, a Leica DMI 6000B was used. Images were cropped and processed with ImageJ 1.51j and assembled in Inkscape 1.1.

Dual chromatin immunoprecipitation (Dual-X-ChIP)

Both formaldehyde (Sigma-Aldrich) and disuccinimidyl glutarate (DSG, ChemCruz) were used in this work to dual cross-link protein-protein complexes to the DNA. One 80% confluent 150 mm_2 dish was harvested and the cell pellet was washed three times with PBS containing 1% phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich). The pellets were resuspended in PBS containing 1% formaldehyde and fixed for 10 min at RT with gentle agitation. For dual cross-linking with DSG and formaldehyde, cells were first resuspended in PBS containing 2 mM DSG and incubated for 30 min at RT while gently swirled, washed twice with PBS-PMSF, and subjected to formaldehyde fixation. Then, the pellets were resuspended in PBS containing 1% formaldehyde and fixed for 10 min at RT with gentle agitation. The reaction was quenched by immediately adding $1:10^{th}$ the volume of a 1.25 M glycine stock solution (PanReac AppliChem) for 5 min at RT while inverting.

All buffers were supplemented with protease inhibitors (final concentrations: 1 µg/ml leupeptin, 1 µg/ml pepstatin, 10 U/ml aprotinin, and 1 mM PMSF). Cells were resuspended with 1 ml buffer I (50 mM HEPES-KOH (PanReac AppliChem, Merck), 140 mM NaCl (Fluka), 1 mM EDTA (Merck), 10% glycerol (PanReac AppliChem), 0.5% Nonidet NP40 (PanReac AppliChem), 0.25% Triton X-100 (Sigma-Aldrich)) and rotated at 4°C for 10 min. After centrifugation (at 2,000 × g and 4°C for 5 min), pelleted nuclei were washed with 1 ml buffer II (10 mM Tris-HCl (PanReac AppliChem), 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA (PanReac AppliChem)). For lysis, pelleted nuclei were vigorously resuspended in 1 ml buffer III (1% SDS (PanReac AppliChem), 10 mM EDTA, 50 mM Tris-HCl) to disrupt the nuclear membrane. Isolated chromatin was sheared by using a Bioruptor Standard System (Diagenode) (15 cycles, 30 seconds intervals on/off; using highest available power settings). 100 µl 10% Triton X-100 was added, cell debris was pelleted (20,000 × g, 4°C), and the chromatin-containing supernatant was collected. For immunoprecipitation, chromatin of 1×10^6 cells was diluted 1:10 in dilution buffer (0.1% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, 167 mM NaCl) and 4 µg of the antibody specific for HA (ab9110, Abcam) and the corresponding control rabbit IgG antibody (#2729, Cell Signaling) were added and incubated for 16 h at 4°C rotating. Afterwards, 50 µl BSA-blocked magnetic Protein-A/G beads (Thermo Fisher) were added for 1 h at 4°C to precipitate antibody-chromatin complexes. Washing steps were performed on a magnetic stand. Beads were then washed once with 1 ml of the following salt buffers: low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, 150 mM NaCl); high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, 500 mM NaCl); LiCl-wash buffer (0.25 M LiCl (Sigma-Aldrich), 1% Nonidet P-40, 1% sodium-deoxycholate (PanReac AppliChem), 1 mM EDTA, 10 mM Tris-HCl). The magnetic beads were then washed twice with TE buffer. To elute chromatin, beads were incubated in 210 µl elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) on a thermal block at 65°C and 1,000 rpm for 30 min. The beads were then separated from the chromatin-containing solution via application of a magnetic stand. The chromatin-containing supernatant was transferred to a reaction tube with 8 µl of 5 M NaCl. De-crosslinking was performed overnight on a thermal block at 65°C and 1,000 rpm. To degrade RNA, 200 µl TE with 8 µl RNAse A (10 mg/ml, Roche) was added before incubation at 37°C for 2 h. Subsequently, 7 µl of 300 mM $CaCl_2$ (Sigma-Aldrich) and 8 µl proteinase K (10 mg/ml, Sigma-Aldrich) were added and tubes were incubated for 1 h at 55°C to degrade proteins in the solution. DNA was purified by phenol-chloroform-isoamyl alcohol (25:24:1; Roth) extraction in phase-lock gel tubes (5Prime, Quantabio) and precipitated with ethanol. Finally, DNA pellets were eluted in 55 µl elution buffer. Input DNA samples were diluted in 200 µl elution buffer and processed just as the ChIP samples starting with the de-crosslinking overnight. The DNA was subsequently subjected to ChIP-seq library preparation.

Library preparation and sequencing

RNA-seq

The polyadenylated (poly(A) mRNA fractions were purified from the total RNA with the NEBnext poly(A) mRNA Magnetic Isolation Module (New England Biolabs, USA), and RNA-seq libraries were generated using the NEXTflex Rapid Directional qRNA-Seq Kit (Perkin Elmer, Bioo Scientific) according to the manufacturer's recommendations. The concentrations and sizes of the final cDNA libraries were measured via RNA High Sensitivity Chip on an Agilent 2100 Bioanalyzer (Agilent). All samples were normalized to 2 nM and pooled at equimolar concentrations. The library pool was sequenced on a NextSeq500 (Illumina, USA) using single read (1 × 75 bp) flow cells via the NextSeq 500/550 High Output Kit v2.5 (Illumina, USA). The samples were sequenced at the high-throughput sequencing technology platform of the LIV.

ChIP-seq

All ChIP and input libraries were generated using the NEXTFLEX ChIP-Seq Library Prep Kit (Perkin Elmer, Bioo Scientific) according to the manufacturer's instructions. The library pool was sequenced on a NextSeq500 (Illumina, USA) using single read (1 × 75bp) flow cells via the NextSeq 500/550 High Output Kit v2.5 (Illumina, USA). The samples were sequenced at the high-throughput sequencing technology platform of the LIV.

Sequencing data analyses

RNA-seq

RNA-seq expression data was generated by quantifying single-end mRNA reads using a decoy-aware mRatBN7.2 (Ensembl 105) transcriptome with the salmon quantifier software (10) via the selective alignment algorithm. This transcriptome was generated by concatenating the respective non-coding- and coding RNA in front of the *Rattus norvegicus* genome, which generates a combined reference file that is subsequently used by salmon to generate an index. This workflow can be obtained from the Zenodo digital library (11) with the DOI 10.5281/zenodo.8047240. The differential gene expression of the quantification data was analyzed with the DESeq2 package according to the developer's vignette (12, 13). The script utilized here can be obtained from the Zenodo digital library (11) with the DOI 10.5281/zenodo.8048294. The output of this DEG analysis, including the normalized count table and all investigated conditions can be found in the Supplementary Data File 3A. These annotated gene lists were frequently sub-selected into upregulated, downregulated or non-significant gene sets, based on their respective $\log_2$ fold expression values, and used within Figures 2, 3, 4 and 5. The Metascape (14) algorithm was used extensively to provide biological context to the generated gene lists using the Reactome- (15), KEGG- (16) and WikiPathway- (17) databases (Supplementary Data File 3B). Heatmaps and other figures associated to selected p53 targets (Figs. 2B and 3A) are based upon a curated (18) list of 116 total genes that were strongly associated with several different p53 ChIP-seq data sets. A slightly reduced list corresponding to their genomic localization in the *Rattus norvegicus* system can be found in Supplementary Data File 4. An overview over these workflows can be found in Fig. S3.

ChIP-seq

Peak identification, visualization and annotation

The single-end reads obtained by sequencing of DNA enriched via HA-E1B-55K-targeted dual-X-ChIP-seq or their inputs and serotype controls in BRK cells were sorted and quality-filtered with samtools (19) using standard settings and subsequently aligned to the *Rattus norvegicus* reference genome mRatBn7.2 (Ensembl 105) using bowtie2 (20) with standard settings. Sequencing-depth normalized genomic coverage files were generated with the deeptools software suite (21) by invoking the bamCoverage tool with –RPGC, -e 150 and –bs 5 settings, assuming an effective genome size of 2.2e9 bp. All heatmaps and profile plots were generated using the deeptools computeMatrix and plotHeatmap tool with varying settings. Significant peaks of individual replicates were identified using the MACS2-algorithm (22) with q-value cut-offs of 0.05, assuming the aforementioned genome size. The peak data from replicates was combined via the MSPC software (23) with standard settings to generate a singular validated peak data set for each E1B-55K protein. In order to reduce background noise and the number of putative false-positive peaks, we discarded any peak that was also identified in the respective serotype control set using the bedtools suite (24), as well as removing any peak that had more than 50% overlap with repetitive genomic regions when compared to an repeat-masked filter file obtained from the UCSC repository for the mRatBn7.2 genome. Any putative peaks that were located on non-canonical chromosome tracks were similarly removed. The remaining peaks were subsequently annotated based on their nearest gene using the ChIPseeker package (25), followed by biological contextualization via Metascape (Supplementary Data File 2A and B), using the Reactome-, KEGG- and WikiPathways-databases. These workflows can be obtained from the Zenodo digital library (11) with the DOI 10.5281/zenodo.8048302 and 10.5281/zenodo.8048509, respectively.

Genome track visualization and region quantification

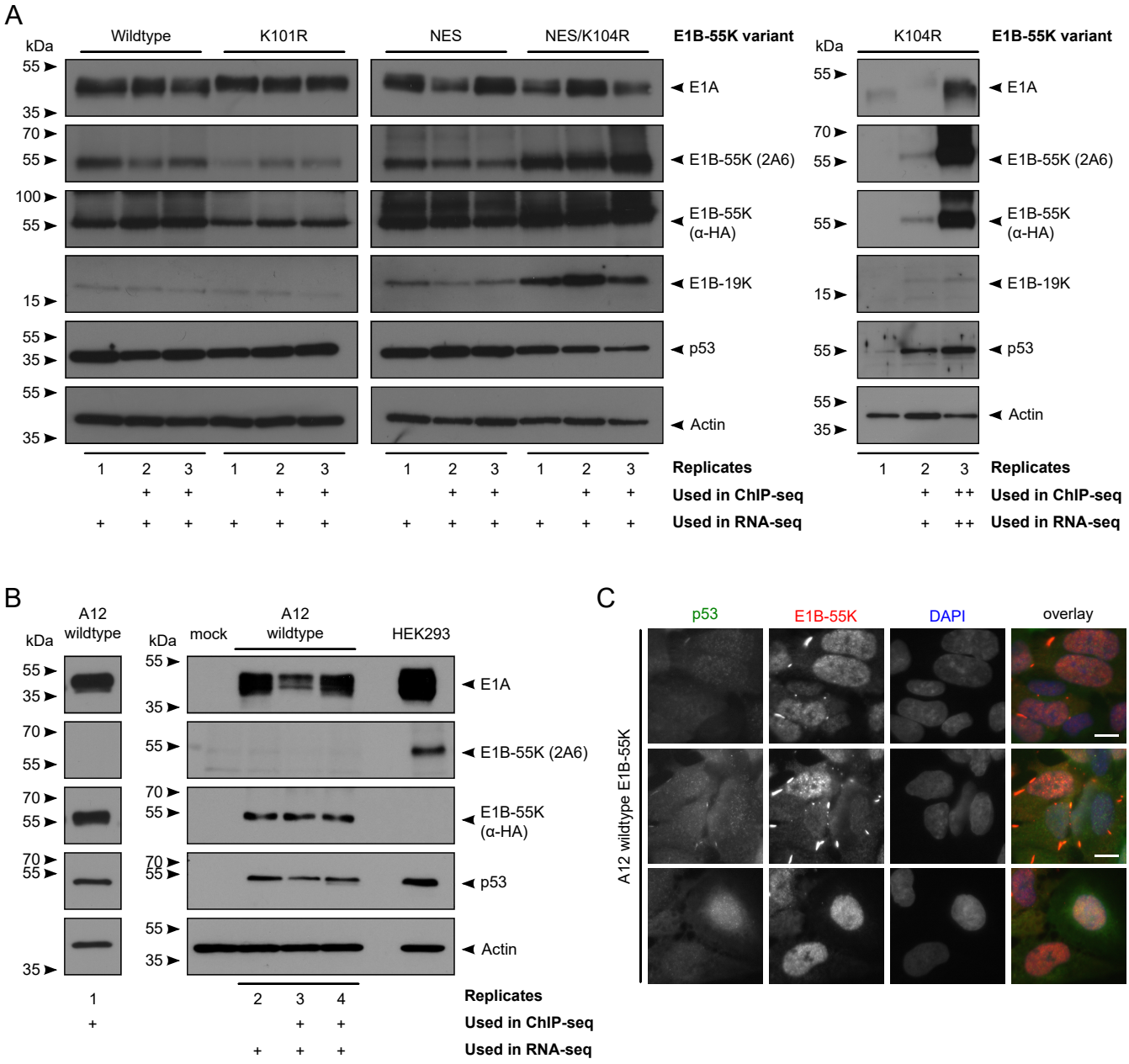
The EaSeq (18) software suite was used to generate the coverage tracks depicted in Fig. 4B. The region quantification program was used to calculate the ChIP-seq signal of all A12 E1B-55K replicates around specific gene sets ranging from the transcriptional start to end sites. The normalization program was subsequently used to perform quantile normalization on these region quantifications prior to statistical analysis and visualization via GraphPad Prism v9.4.0(673).

Motif identification

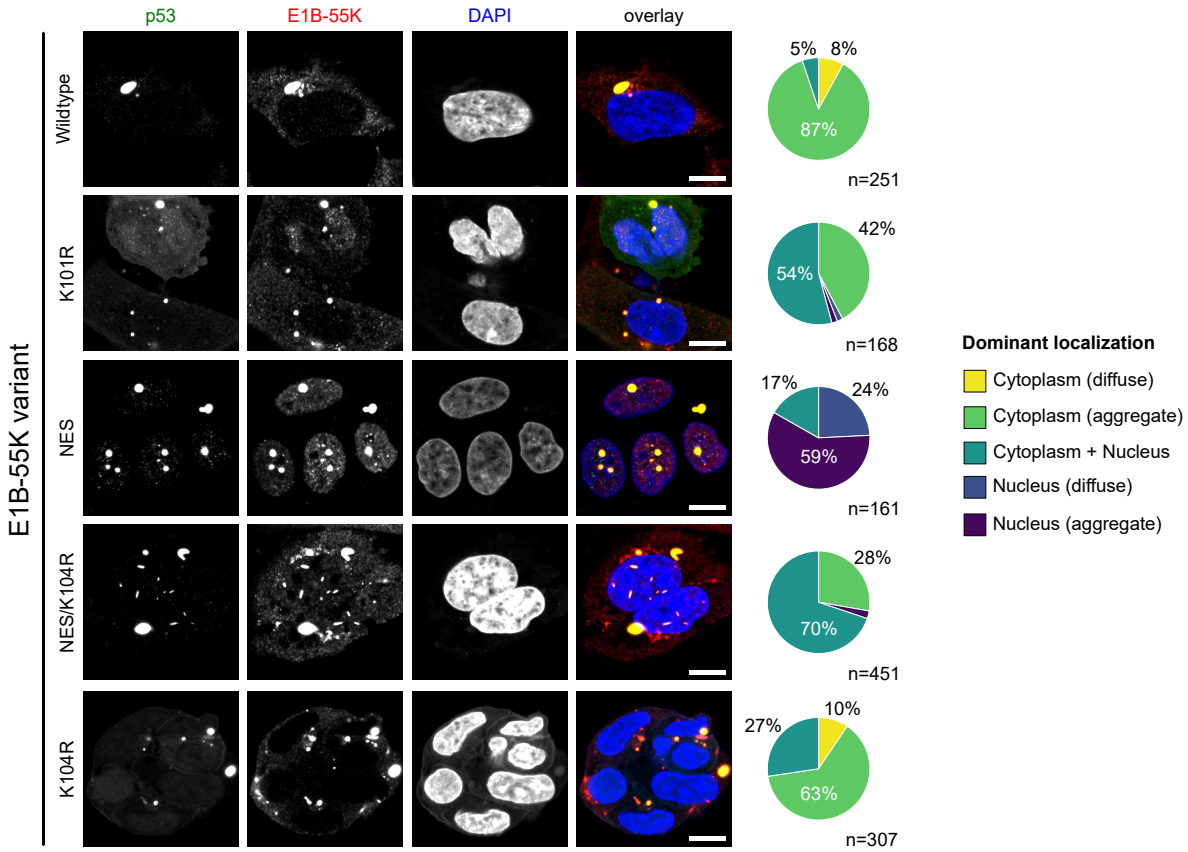
We used the HOMER (26) suite to identify any significantly enriched nucleotide motifs within the final ChIP-seq peak data sets by invoking the findMotifsGenome.pl program on a 200 bp region around the peak summits, allowing for two mismatches. Thereby identified motifs were obtained from the de novo motif discovery, visualized in Figs. 1E, 4D and 5D-F, as well as summarized in Supplementary Data File 1B. Specific peaks that contained either only one or a specific combination of the most prevalent motifs were isolated with the annotatePeaks.pl program contained in the HOMER suite and annotated with their respective nearest genes and their exact genomic location into either promoter or enhancer-located sets. To increase the precision of peak association to functional enhancers, a specific set of putative enhancers identified in eleven different *Rattus norvegicus* cell lines (27) was manually combined to create a consensus-enhancer localization file (see DOI 10.5281/zenodo.8048320). Here, the localization of any peak described as “distal intergenic” was compared to the consensus file and only intersecting peaks were assigned the “enhancer”-tag. This approach removed around 15% of distal-intergenic peaks, which were not clearly associated with a previously identified enhancer region. The enhancer-consensus list can be found in Supplementary Data File 8. Through this localization and expression status of their nearest associated genes, we calculated overlaps based on the amount of genes found in the respective groups. These data were visualized using GraphPad Prism in Fig. 5D-F and Figures S5A and B as well as S6A-C. The individual motif-sets, their annotated genes and expression levels can be found in Supplementary Data File 6B. The data associated with the top three motifs identified by A12 E1B-55K ChIP-seq can be found in Supplementary Data File 2B.

Statistical analyses

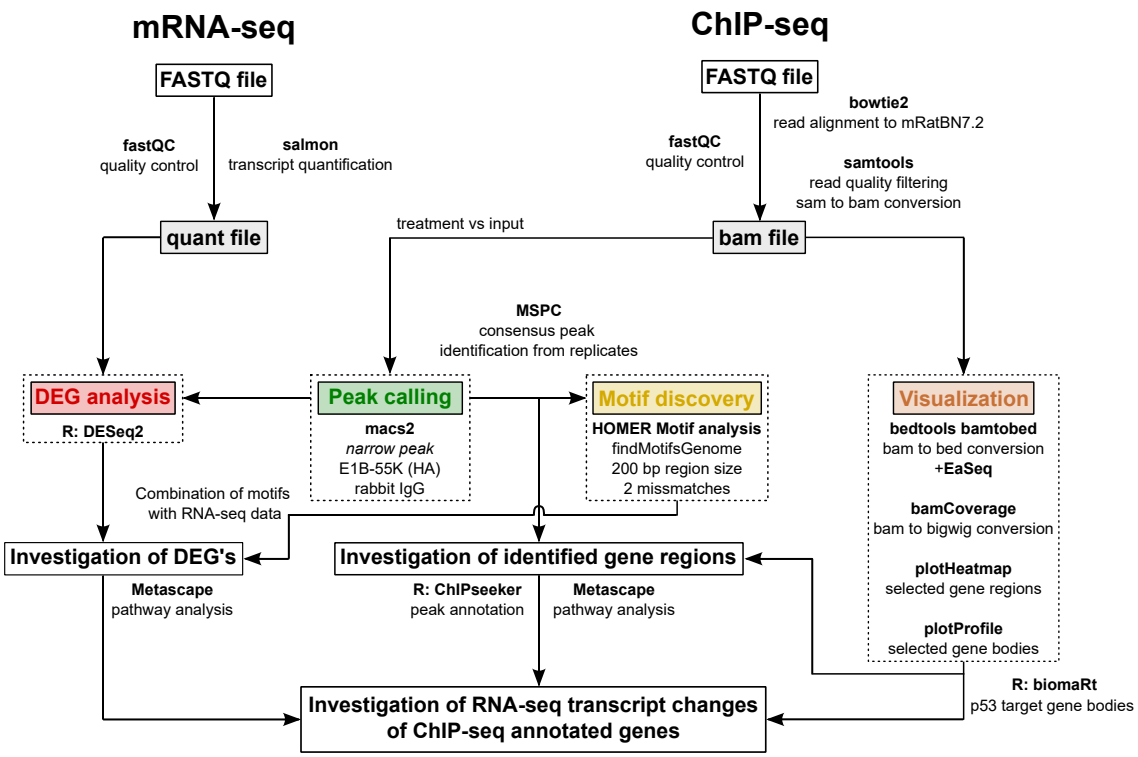
The Pearson correlation coefficient calculations depicted via GraphPad Prism in Fig. 5A-C are presented in Supplementary Data File 6A, where the strength of interaction, based on the peak score, was correlated with the $\log_2$ fold change of corresponding significantly changed genes (adjusted p-value < 0.1). The statistical relationship of the EaSeq-calculated and normalized ChIP-seq signal of three A12 E1B-55K ChIP-seq replicates from three different cell lines between the significantly downregulated, upregulated or non-significant clusters was calculated using a two-way ANOVA test in GraphPad Prism and is presented in Supplementary Data File 5.



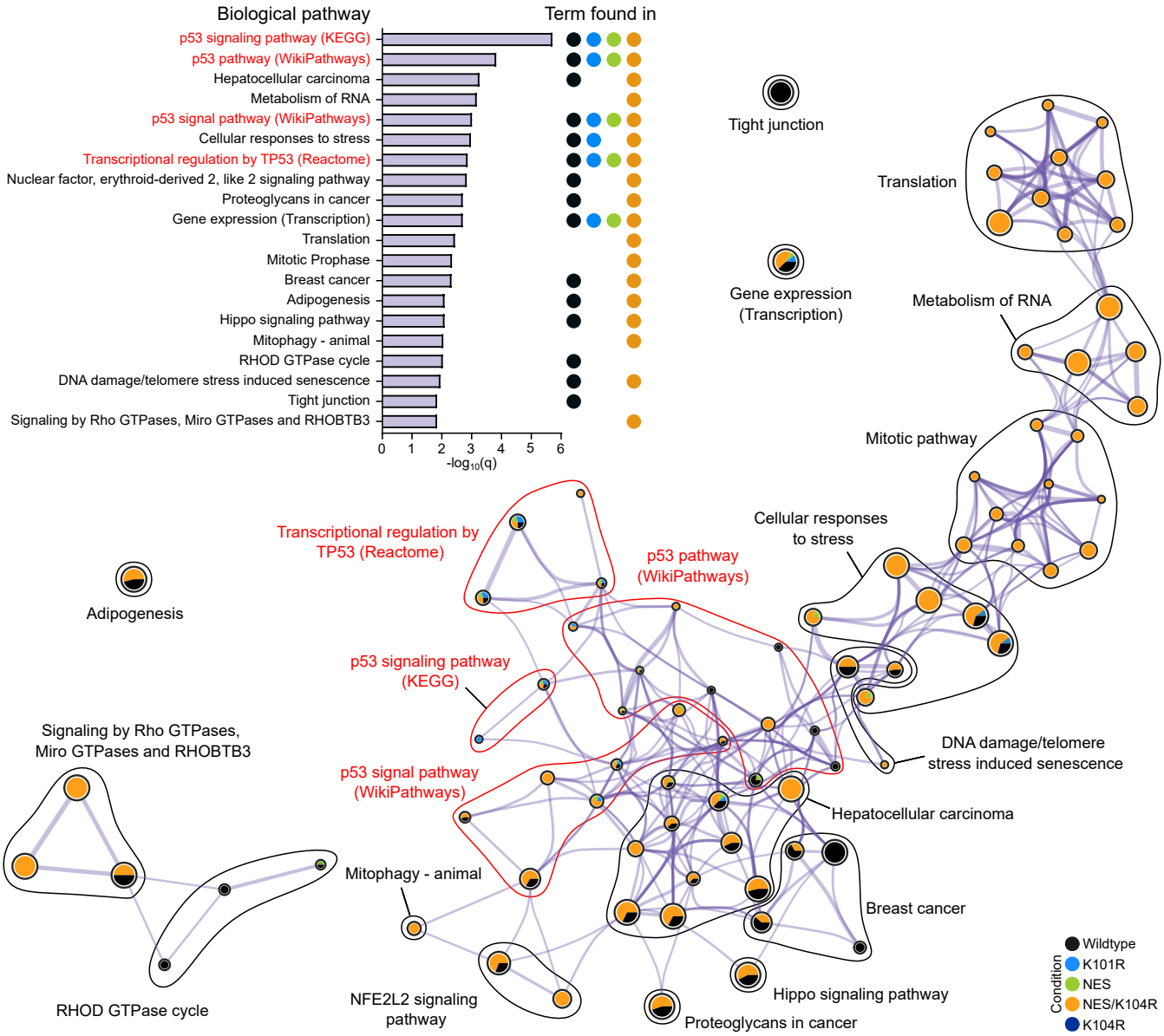
Supplementary Figure 1: Characterization of stable BRK cell lines expressing the annotated HAdV-C5 or -A12 E1B-55K variant. After total protein extraction from BRK cells transformed via transduction of HAdV-C5 E1A and (A) C5 E1B-55K or (B) A12 E1B-55K, expression levels of key proteins were detected using M73 (α-E1A), 2A6 (α-Ad5 E1B55K), ChIP-HA (α-HA-E1B-55K), #490 (α-Ad5 E1B-19K), pAB421 (α-p53) and A5441 (α-β-actin) antibodies. Annotated E1B-55K proteins represent three independently propagated cell pools of one transduction experiment, while the K104R replicates originate from three different transduction experiments. Regarding A12, an initial pioneering experiment was completed with three additional replicates originating from one transduction. The annotation below the individual blots indicates which sample has been used in ChIP-seq and RNA-seq, respectively. (C) E1B-55K (α-HA-E1B-55K) and p53 (pAB421) localization in A12 HA-E1B-55K expressing BRK cells (replicate 3). The dsDNA is visualized by DAPI staining. Imaging was performed with a Leica DMI 6000B microscope at 40x magnification with oil-immersion. The scale bar corresponds to 10 µm.



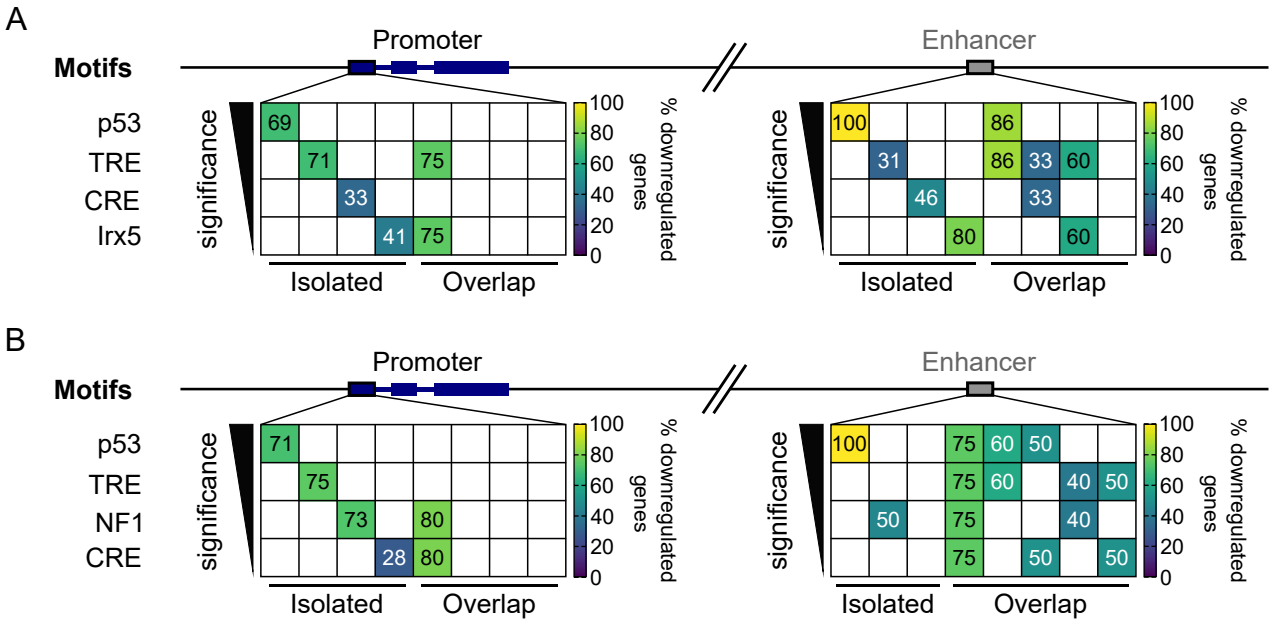
Supplementary Figure 2: Quantification and intracellular localization of p53 and HAdV-C5 E1B-55K in stable BRK cell lines. BRK expressing different variants of C5 E1B-55K were analyzed regarding their E1B-55K (α-HA-E1B-55K) and p53 (pAB421) localization. The dsDNA is visualized by DAPI staining. BRK cells were grouped according to their E1B-55K localization patterns, ranging from diffusely in the cytoplasm (in yellow) to aggregates in the nucleus (in dark purple). Images shown here are examples of the observed major distribution patterns. Imaging was performed with a Nikon A1R HD25 microscope at 60x magnification with oil-immersion. Multiple images were fused by using the Nikon image stitch tool, extracted via the Fiji software and counted by virtue of the cell counter plugin. The scale bar corresponds to 10 µm.



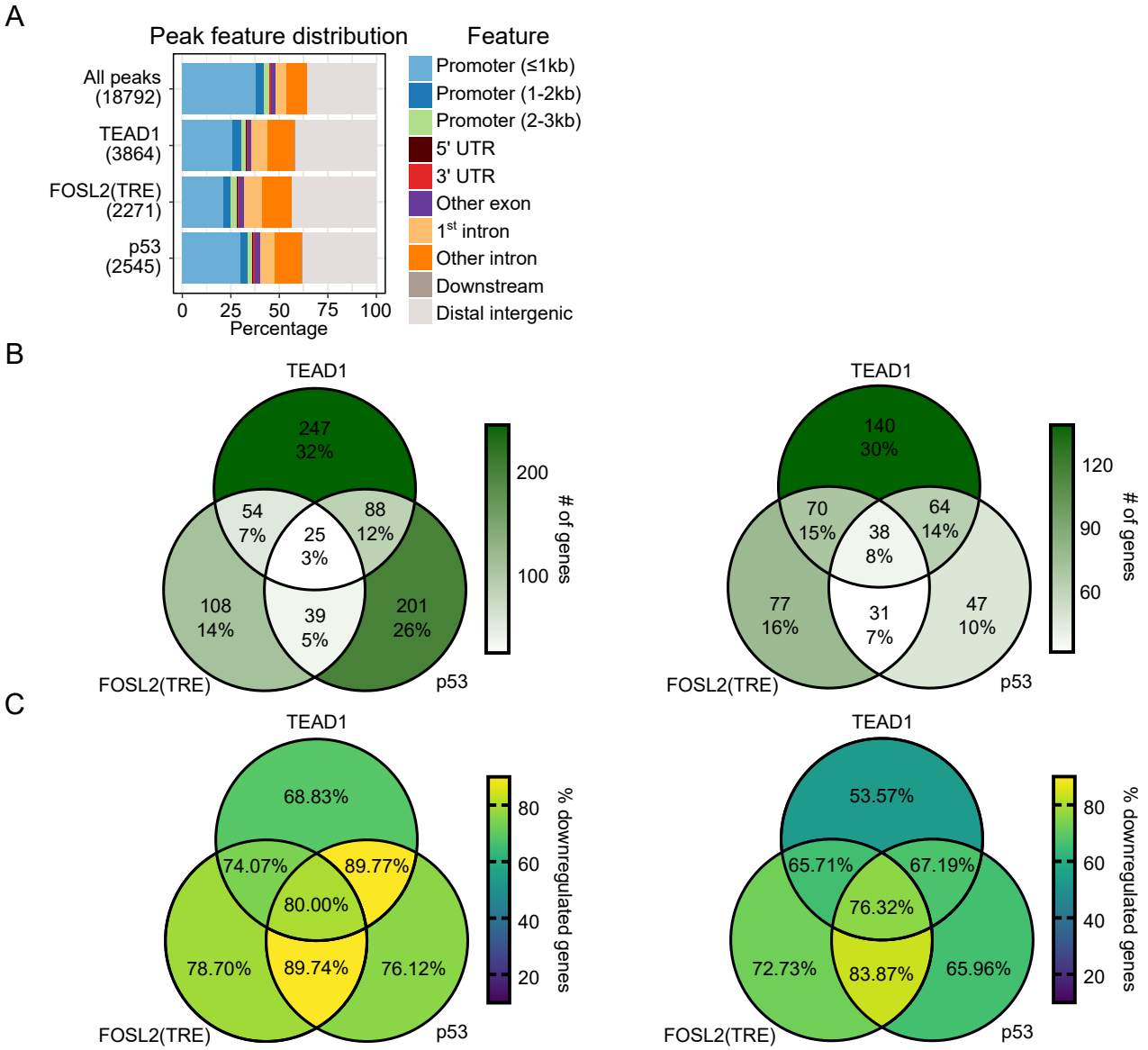
Supplementary Figure 3: Exemplary bioinformatic workflow utilized for ChIP-seq, RNA-seq and combinatorial analysis. Flowchart diagram displaying the utilized bioinformatic software for DNA sequence alignment, MACS2 narrow peak calling, MSPC replicate peak verification, HOMER de novo motif discovery, visualization and R packages used for functional assessment, as well as software used for mRNA transcript quantification and subsequent differential gene expression analyses.



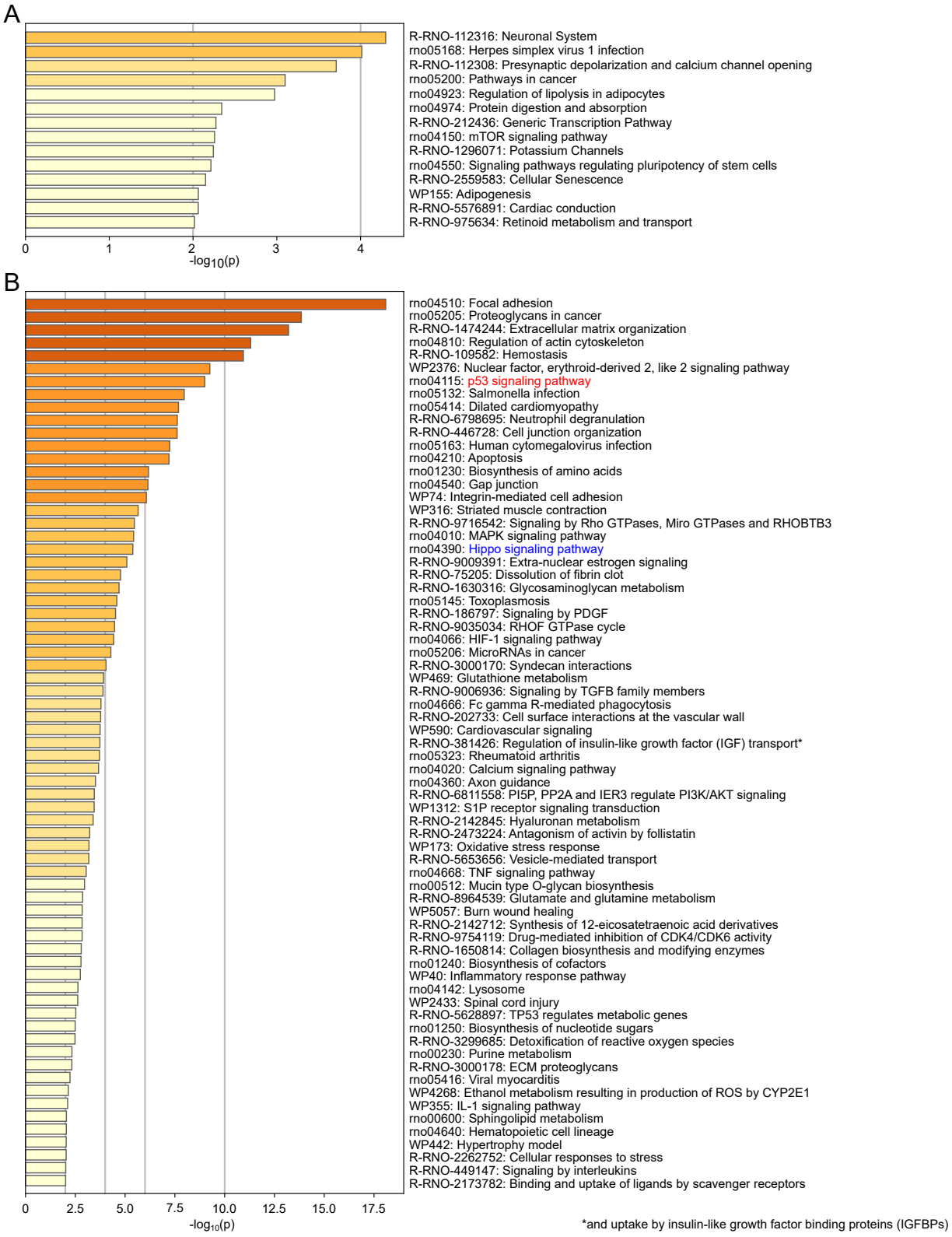
Supplementary Figure 4: Peak-to-gene based pathway analysis reveals potential influence of E1B-55K on vital cellular pathways. Visualization of pathways obtained from genes that are adjacent to the identified peaks. For each given gene list, pathway and process enrichment analyses were performed with the following gene ontology sources: KEGG Pathway, Reactome Gene Sets, and WikiPathways. Shown is a network plot where similar or redundant terms are connected into clusters. Here, pathways with a p-value of < 0.01, a minimum count of three and an enrichment factor of > 1.5 are collected and grouped into clusters based on their membership similarities. The pathway with the best p-value was selected as an annotation for the entire cluster. Each node is represented by a pie chart that is color-coded based on the representation of the different E1B-55K peak datasets. The size of each pie is proportional to the total number of hits that fall into that specific pathway while the size of each slice represents the percentage of genes that belong to the term that originated from the corresponding gene list. Bar charts (upper left) of ranked pathways, sorted by their respective $-\log_{10}$ q-values from most to least significant. A significant presence of E1B-55K proteins in identified pathways is annotated with circles. Pathways directly connected to p53 are highlighted in red. The original data are provided in Supplementary Data File 2a.



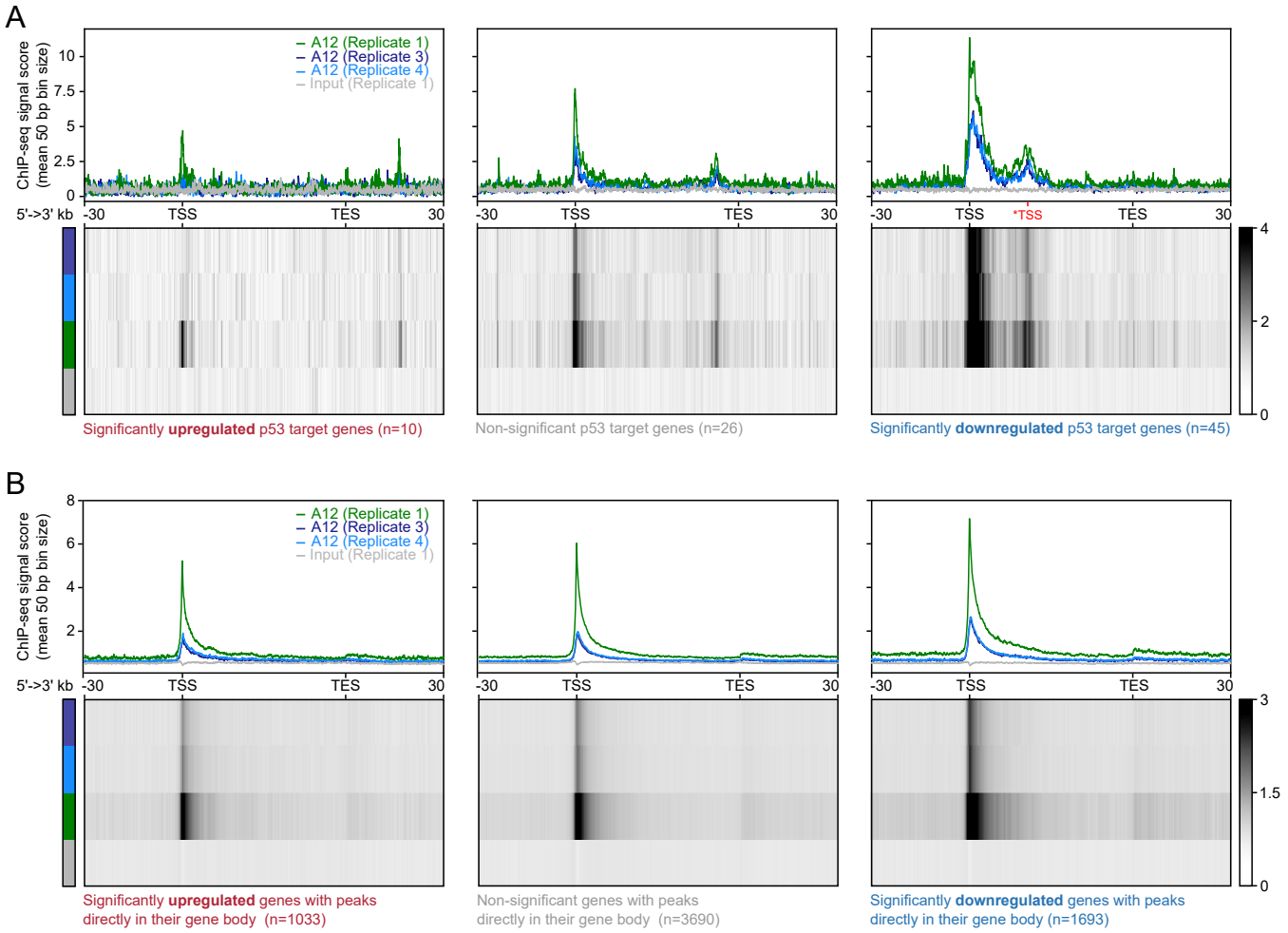
Supplementary Figure 5: Investigation into motif overlap associated with RNA-seq transcriptome changes in peaks through HAdV-C5 E1B-55K ChIP-seq on promoter or enhancer regions. (A) Wildtype- and (B) NES/K104R E1B-55K associated motifs co-occurrence, additionally separated into promoter- and enhancer sets. Peaks with isolated motifs are not co-occupied by any other significant motif, while peaks with overlapping motifs share at least one of the indicated occurance. A coloured panel represents at least four unique motif occurrences in isolated or combined total overlapping motifs. The color code ranges from dark blue (upregulated) to bright yellow (downregulated).



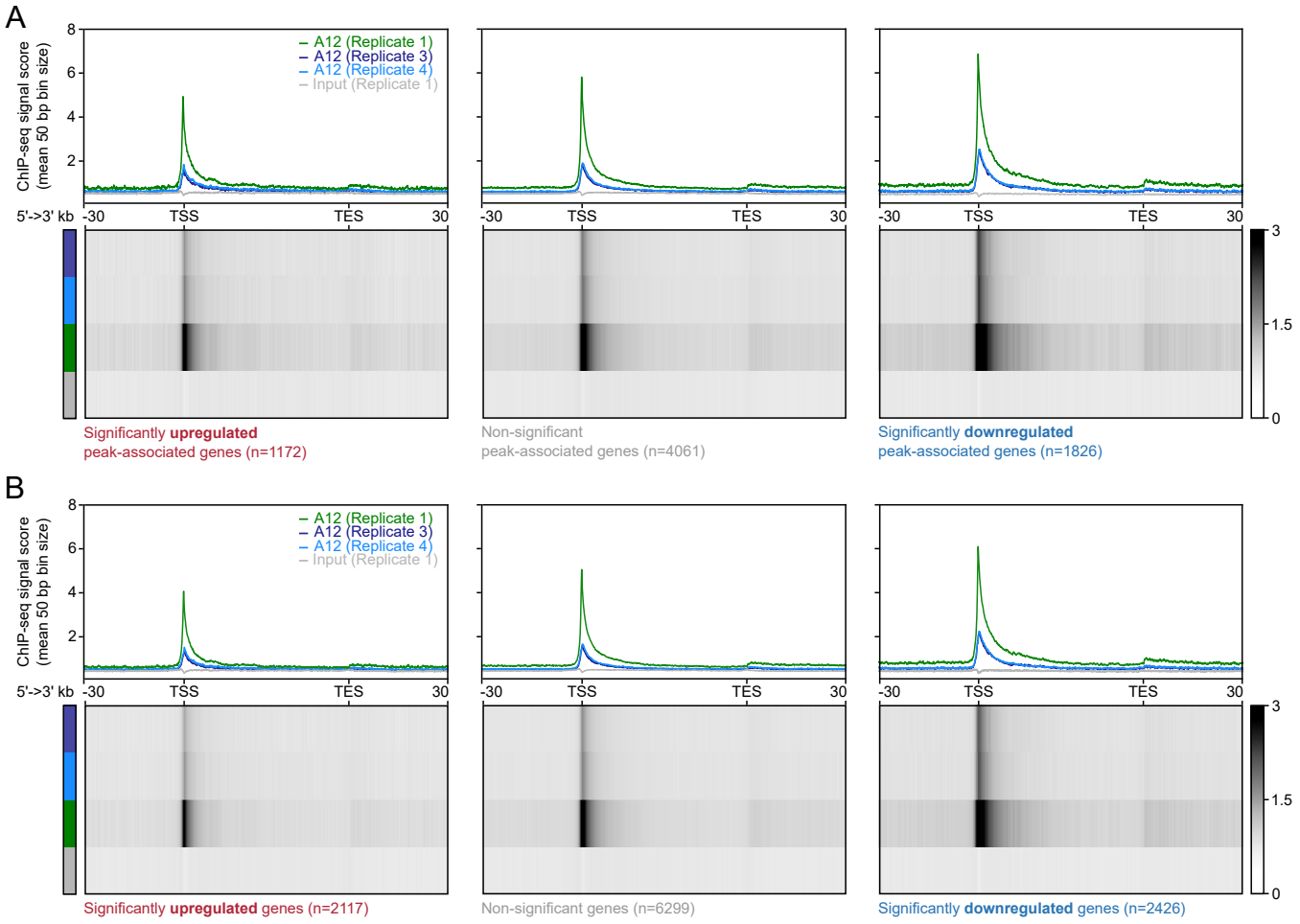
Supplementary Figure 6: Peak annotation and motif overlap associated with RNA-seq transcriptome changes through HAdV-A12 E1B-55K ChIP-seq on promoter or enhancer regions. (A) Feature classification of significantly enriched motifs in HAdV-A12 E1B-55K peaks (q-value ≤ 0.05) with the total numbers of peaks in brackets. (B) Absolute numbers and percentage of associated overlap co-occurrence from the three most significant motifs associated to significantly changed genes (RNA-seq FDR ≤ 0.1), additionally separated into promoter- (*Left Side*) and enhancer (*Right side*) sets. The color code ranges from white to dark green, according to the absolute number of genes in the respective category. (C) Transcriptional status of genes from (B) associated with either of the shown motifs. The color code ranges from dark blue (upregulated) to bright yellow (downregulated).



Supplementary Figure 7: All significantly deregulated pathways by HAdV-A12 E1B-55K (RNA-seq). Pathway analysis of significantly changed genes compared to E1B-negative BRK with an FDR ≤ 0.01 and a $\log_2$ fold change (A) ≥ 1 or (B) ≤ -1. The heatmap plots is coloured and sorted by p-values, with only the top 100 significant pathways presented. The cluster term "*p53 signaling pathway*", is hightlighted in red while "*Hippo signaling pathway*" is highlighted in blue. The complete data is available in Supplementary Data File 7.



Supplementary Figure 8: A12 E1B-55K ChIP-seq signal around gene bodies of deregulated p53 targets and gene sets with binding events within gene bodies. Visualization of A12 E1B-55K ChIP-seq signal around specific gene bodies of genes that are either significantly up- (*Left Window*; $\log_2$ fold change ≥ 0 and FDR ≤ 0.1), non-significant (*Middle Window*; FDR ≥ 0.1) or significantly downregulated (*Right Window*; $\log_2$ fold change ≤ 0 and FDR ≤ 0.1) according to their relative mRNA abundance compared to an E1B-negative cell line. Here, (A) contains a set of specific p53 target genes, while (B) represents genes that contain peaks directly in their gene body. The upper blot shows the ChIP-seq signal profile track of scores calculated over 50 bp bin size windows with the respective track annotated in the top right corner of the first profile plot. All gene bodies are visualized from 5' to 3' direction and normalised to a size of 50 kb. TSS depicts the transcriptional starting site, while TES depicts the transcriptional end site. The signal around the *TSS (in red) is attributed to interaction with an alternative transcriptional starting site of the *CDKN1A* and *RPS27L* genes. The heatmaps visualize the profile below their respective profile plot, the intensity of signal represented by a greyscale, ranging from white (no signal) to black (strong signal). The absolute number of genes within each set is depicted below the respective heatmap plots.



Supplementary Figure 9: A12 E1B-55K ChIP-seq signal around gene bodies of deregulated gene sets with associated binding events or non-pre-selected gene sets. Visualization A12 E1B-55K ChIP-seq signal around specific gene bodies of genes that are either significantly up- (*Left Window*; $\log_2$ fold change ≥ 0 and FDR ≤ 0.1), non-significant (*Middle Window*; FDR ≥ 0.1) or significantly downregulated (*Right Window*; $\log_2$ fold change ≤ 0 and FDR ≤ 0.1) according to their relative mRNA abundance compared to an E1B-negative cell line. Here, (A) contains a set of genes that are generally associated with peaks, while (B) represents all genes that have a mRNA transcript base mean above 50. The upper blot shows the ChIP-seq signal profile track of scores calculated over 50 bp bin size windows with the respective track annotated in the top right corner of the first profile plot. All gene bodies are visualized from 5' to 3' direction and normalised to a size of 50 kb. TSS depicts the transcriptional starting site, while TES depicts the transcriptional end site. The heatmaps visualize the profile below their respective profile plot, the intensity of signal represented by a greyscale, ranging from white (no signal) to black (strong signal). The absolute number of genes within each set is depicted below the respective heatmap plots.

Table S1. Mutagenesis primers.

Primer description[a]	Primer sequence (5'-3')	Reference
E1B-55K K104R fwd	GGGCTAAAGGGGGTAAGGAGGGAGCGGGGG	Kolbe *et al*. (2022), JVI (4)
E1B-55K K104R rev	CCCCCGCTCCCTCCTTACCCCCTTTAGCCC	
E1B-55K NES fwd	GGTGGCTGAAGCGTATCCAGAAGCGAGACGCA TTGCGACAATTACAGAGG	This work
E1B-55K NES rev	CCTCTGTAATTGTCGCAATGCGTCTCGCTTCTG GATACGCTTCAGCCACC	
E1B-55K K101R fwd	GGGCAGGGGCTAAGGGGGGTAAAGAGGG	Kolbe *et al*. (2022), JVI (4)
E1B-55K K101R rev	CCCTCTTTACCCCCCTTAGCCCCTGCCC	
HA insertion fwd	ATGTACCCATACGATGTTCCAGATTACGCTGAG CGAAGAAAC	This work
HA insertion rev	AGCGTAATCTGGAACATCGTATGGGTACATTTA TCCTTTATAAAAC	

[a] fwd, forward primer; rev, reverse primer

SI References

1. K. Weber, U. Bartsch, C. Stocking, B. Fehse, A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. *Mol Ther* **16**, 698-706 (2008).
2. C. Endter, J. Kzhyshkowska, R. Stauber, T. Dobner, SUMO-1 modification required for transformation by adenovirus type 5 early region 1B 55-kDa oncoprotein. *Proc Natl Acad Sci U S A* **98**, 11312-11317 (2001).
3. M. Fiedler *et al.*, Protein–Protein Interactions Facilitate E4orf6-Dependent Regulation of E1B-55K SUMOylation in HAdV-C5 Infection. *Viruses* **14**, 463 (2022).
4. V. Kolbe *et al.*, Conserved E1B-55K SUMOylation in Different Human Adenovirus Species Is a Potent Regulator of Intracellular Localization. *J Virol* **96**, e0083821 (2022).
5. C. Endter, B. Härtl, T. Spruss, J. Hauber, T. Dobner, Blockage of CRM1-dependent nuclear export of the adenovirus type 5 early region 1B 55-kDa protein augments oncogenic transformation of primary rat cells. *Oncogene* **24**, 55-64 (2005).
6. F. Krätzer *et al.*, The adenovirus type 5 E1B-55K oncoprotein is a highly active shuttle protein and shuttling is independent of E4orf6, p53 and Mdm2. *Oncogene* **19**, 850-857 (2000).
7. E. Harlow, B. R. Franza, Jr., C. Schley, Monoclonal antibodies specific for adenovirus early region 1A proteins: extensive heterogeneity in early region 1A products. *J Virol* **55**, 533-546 (1985).
8. P. Sarnow, C. A. Sullivan, A. J. Levine, A monoclonal antibody detecting the adenovirus type 5 E 1 b-58Kd tumor antigen: Characterization of the E 1 b-58Kd tumor antigen in adenovirus-infected and -transformed cells. *Virology* **120**, 510-517 (1982).
9. E. Lomonosova, T. Subramanian, G. Chinnadurai, Mitochondrial localization of p53 during adenovirus infection and regulation of its activity by E1B-19K. *Oncogene* **24**, 6796-6808 (2005).
10. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).
11. European Organization For Nuclear Research, OpenAIRE (2013) Zenodo.
12. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).
13. M. I. Love, S. Anders, W. Huber (2023) Analyzing RNA-seq data with DESeq2.
14. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
15. M. Gillespie *et al.*, The reactome pathway knowledgebase 2022. *Nucleic Acids Res* **50**, D687-D692 (2022).
16. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
17. M. Martens *et al.*, WikiPathways: connecting communities. *Nucleic Acids Res* **49**, D613-D621 (2021).
18. M. Lerdrup, J. V. Johansen, S. Agrawal-Singh, K. Hansen, An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat Struct Mol Biol* **23**, 349-357 (2016).
19. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10** (2021).
20. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
21. F. Ramirez *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
22. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137 (2008).
23. V. Jalili, M. Matteucci, M. Masseroli, M. J. Morelli, A. Valencia, Using combined evidence from replicates to evaluate ChIP-seq peaks. *Bioinformatics* **31**, 2761-2769 (2015).
24. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
25. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).
26. S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).
27. T. Gao, J. Qian, EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Res* **48**, D58-D64 (2020).

Description of Supplementary Data Files

Supplementary Data File 1A. Compiled lists of dual-fixation ChIP-seq peaks targeting HA-E1B-55K wt or any of the other investigated proteins. We obtained the peaks (MACS2, q-value < 0.01) by MSPC-verification (standard settings) of two individual ChIP-seq experiments in order to reduce the number of potential off-target peaks. Annotation of the peaks is based on their respective nearest genes using the R package ChIPseeker. Overlap gene sets between either C5-associated E1B-55K proteins or A12 with C5 wt and NES/K104R are annotated here.

Supplementary Data File 1B. Data frame that contains the functional output of all significant motifs obtained by invoking the HOMER de novo motif analysis software (normal settings) on HA-E1B-55K-associated peaks. Depicted in columns is the consensus motif, the respective p-values, the proportion of motif occurrence in individual target or background sequences (in %), the standard deviation of motif occurrence off the peak centre in individual target (STD) or background sequence (STD; Bg) and the best match compared to the HOMER vertebrate motif library. These motifs are presented in Figures 1E, 4D, and Supplementary Figures 5 and 6.

Supplementary Data File 2A. Data obtained via Metascape-assisted (https://metascape.org/) biological pathway analysis with combined C5 E1B-55K peak-associated gene sets annotated in Supplementary Data File 1. Also contains the top 20 clusters that are illustrated in individual pathway membership data.

Supplementary Data File 2B. Data obtained via Metascape-assisted biological pathway analysis with combined A12 E1B-55K peak-associated gene sets annotated in Supplementary Data File 1. Only contains genes that are associated with TEAD-, TRE- or p53-motifs. Also contains the top 20 clusters that are illustrated in individual pathway membership data.

Supplementary Data File 3A. Data frame that contains the normalized mRNA-seq transcript counts for all mRatBN7.2 (Ensembl 105) genes obtained from DEseq2-analysis. Other tables consist of compiled DEseq2-based differential gene expression analysis from all comparisons (E1A + HA-E1B-55K proteins versus E1B-negative). These genes are used for the volcano blots in Fig. 2B, for the Metascape-assisted gene up- or down-regulated analysis in Fig. 2C and Supplementary Fig. 8.

Supplementary Data File 3B. Data obtained via Metascape-assisted biological pathway analysis with combined C5 E1B-55K peak-associated gene sets obtained from DEseq2-analysis and annotated in Supplementary Data File 3A (E1A + HA-E1B-55K proteins versus E1B-negative). Also contains the top 20 clusters that are illustrated in individual pathway membership data. These pathways are depicted in Fig. 2C and D.

Supplementary Data File 4. Compiled list of p53 target genes in .bed format (mRatBN7.2 genome assembly). This list was originally compiled by Fischer et al. (2017) and is based on genes that were identified as activated p53 target genes in at least 6 out of 16 genome-wide data sets. These genes are used in Figures 2B, 3, and 4F, as well as Supplementary Fig. 8A.

Supplementary Data File 5. EaSeq ChIP-seq gene body quantification data of three A12 wt replicates with their corresponding significance tests. These lists contain either genes annotated as p53-targeted genes (see Supplementary data file 4) or genes that harbour verified peaks within their gene body. These genes were sub-selected into significantly upregulated (FDR < 0.1 and $\log_2$ fold change > 0), non-significant (FDR > 0.1) or significantly downregulated (FDR < 0.1 and $\log_2$ fold change < 0), based on their expression compared to the E1B-negative control cell line. The coordinates of these gene bodies from the TSS (transcription start site) to the TES (transcription end site) were extracted using the Ensembl BioMart data-mining tool. The cumulative ChIP-seq signal

from TSS to TES was calculated via the EaSeq Quantify regions tool with normal settings, normalizing to reads per million and to a gene size of 1000 bp, assuming a fragment size of 150 bp. Areas above 1 Gbp or below 100 bp were set to those sizes. All quantified signals obtained from these genes across the aforementioned sub-selected groups were subsequently quantile normalised using the EaSeq Normalize tool. Statistical significance between the cumulative signals of these groups was tested via a non-parametric Kruskal–Wallis one-way ANOVA, followed by Dunn's multiple comparisons test (left side) or a two-way ANOVA test, followed by Tukey's multiple comparisons test (right side). This data are depicted in Fig. 4F.

Supplementary Data File 6A. Data frames containing the Pearson correlation calculations for C5 wt and NES/K104R E1B-55K-, as well as A12 wt E1B-55K-, associated peaks. These peaks were ranked from highest to lowest based on their respective MACS2 peak score, matched to their nearest gene and combined with the mRNA expression $\log_2$-fold change of that gene, compared to the E1B-negative control. Contains both peaks lists with- (*Right Side*) and without (*Left Side*) significance-cut-off pertaining to the differential gene expression data. The left side is illustrated in Fig. 5A-C. Pearson: Calculation of correlation between peak score and $\log_2$-fold change; N: number of peaks; T stat: t-statistic; DF: degrees of freedom; pvalue: p-value; log2fold: average $\log_2$-fold change of all peak-associated genes.

Supplementary Data File 6B. Data frames containing the complete lists of peak-associated genes for C5 wt and NES/K104R E1B-55K-, as well as A12 wt E1B-55K. These peaks were ranked from highest to lowest based on their respective macs2 peak score, matched to their nearest gene and combined with the mRNA expression $\log_2$-fold change of that gene, compared to the E1B-negative control (FDR < 0.1). Contains all individual motif occurrences of HOMER de novo motif analyses, as well as the occurrence of each significant motif in the individual peaks. The combined tab contains all peak-associated motifs, their individual peak-based functional annotation, as well as the proportion of mRNA-seq-based up- or downregulation of its nearest gene. These data are depicted in Fig. 5D-F.

Supplementary Data File 7. Data obtained via Metascape-assisted biological pathway analysis with A12 E1B-55K peak-associated gene sets obtained from DEseq2-analysis and annotated in Supplementary Data File 3A (E1A + HA-E1B-55K proteins versus E1B-negative). Also contains the top 20 clusters that are illustrated in individual pathway membership data. These pathways are depicted in Supplementary Fig. 7.

Supplementary Data File 8. Consensus enhancer file in bed format containing a compiled high-confidence list of enhancers obtained from different *Rattus norvegicus* cell lines (http://www.enhanceratlas.org/). The predicted enhancers from the following cell lines were combined: chondrosarcoma cells, GC pituitary cells, INS-1E cells, liver cells, neuron cells, ventricular myocyte (NRVM) cells, oligodendrocytes and their precursors, Schwann cells, sciatic nerve cells, and rat aortic smooth muscle cells (RASMC). The exact locations of these regions were initially corresponding to the rn5 genome, which were subsequently lifted over to the rn6 and mRatBN7.2 genomes using the hgLiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) tool.