



Published in final edited form as:

Nat Genet. 2018 December ; 50(12): 1658–1665. doi:10.1038/s41588-018-0263-0.

Retinoic Acid and BMP4 Cooperate with TP63 to alter Chromatin Dynamics during Surface Epithelial Commitment

Jillian M. Pattison^{#1}, Sandra P. Melo^{#1,4}, Samantha N. Piekos^{#1}, Jessica L. Torkelson¹, Elizaveta Bashkirova^{1,5}, Maxwell R. Mumbach^{1,2}, Charlotte Rajasingh¹, Hanson Hui Zhen¹, Lingjie Li¹, Eric Liaw^{1,6}, Daniel Alber¹, Adam J. Rubin¹, Gautam Shankar¹, Xiaomin Bao^{1,7}, Howard Y. Chang^{1,2,3}, Paul A. Khavari¹, and Anthony E. Oro

¹Program in Epithelial Biology, Stanford University School of Medicine, Stanford, California 94305, USA.

²Center for Personal Dynamic Regulomes, Stanford, California 94305, USA.

³Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA.

⁴Current Addresses: Agilent Technologies, Santa Clara, California 95051, USA.

⁵Current Addresses: Columbia Stem Cell Initiative, Columbia University Medical Center, New York, New York 10032, USA.

⁶Current Addresses: David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA.

⁷Current Addresses: Department of Molecular Sciences & Department of Dermatology, Northwestern University, Evanston, Illinois 60208, USA.

These authors contributed equally to this work.

Abstract

Human embryonic stem cell (hESC) differentiation promises advances in regenerative medicine^{1–3}, yet conversion into transplantable tissues remains poorly understood. Using our keratinocyte differentiation system, we employ a multi-dimensional genomics approach to interrogate the contributions of inductive morphogens retinoic acid (RA) and bone morphogenetic protein 4 (BMP4) and the epidermal master regulator p63^{4,5} during surface ectoderm commitment. In contrast to other master regulators^{6–9}, p63 effects major transcriptional changes only after morphogens alter chromatin accessibility, establishing an epigenetic landscape for p63 to modify.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Address correspondence to: Anthony E. Oro oro@stanford.edu.

Author Contributions

S.P.M. and J.M.P. designed and executed experiments, analyzed data, and wrote the manuscript. S.N.P. analyzed data and wrote the manuscript. J.L.T., E.B., M.R.M., C.R., H.H.Z., L.L., and X.B. executed experiments and contributed to experimental design. E.L., D.A., A.J.R., and G.S. contributed to data analysis. H.Y.C. and P.A.K. contributed to experimental design. A.E.O. designed experiments, analyzed data, wrote the manuscript, and conceived the project with S.P.M.

Competing Interests

The authors declare no competing financial interests.

Accession Codes

All sequencing data is available through Gene Expression Omnibus (GEO) accession number: GSE114846.

p63 distally closes chromatin accessibility and promotes accumulation of H3K27me3 modifications. Cohesin HiChIP¹⁰ visualizations of chromosome conformation reveal that p63 and the morphogens contribute to dynamic long-range chromatin interactions, as illustrated with *TFAP2C* regulation¹¹. Our study demonstrates the unexpected dependency of p63 on morphogenetic signaling and provides novel insights into how a master regulator can specify diverse transcriptional programs based on the chromatin landscape induced by specific morphogen exposure.

As published protocols of hESC-derived keratinocytes suffer from excessive heterogeneity^{5,12–15}, we developed a xeno-free, chemically-defined differentiation system using E6 media¹⁶ supplemented with two morphogens, RA and BMP4 (Fig. 1a). RA/BMP4 treatment produced functional keratinocytes that behaved similarly to those described previously⁵ (Supplementary Fig. 1). Differentiating cells progressed through a simple epithelial state as indicated by immunofluorescence (IF) analysis of epithelial markers keratin 18 (K18)¹⁷ and p63^{18,19} at day 7, followed by high levels of p63 and keratinocyte maturation marker keratin 14 (K14)²⁰ at day 45 (Fig. 1a). Robust p63 expression occurred only when both morphogens were present, indicating a synergistic role for p63 accumulation (Fig. 1b,c, Supplementary Fig. 1). As morphogenetic exposure for 7 days induced both uniform p63 expression and subsequent keratinocyte development^{4,5}, we interrogated this key 7-day period with a multi-dimensional genomics approach to understand the functional interaction between p63 and the morphogens.

To assess the individual contributions to chromatin dynamics, we created p63 gain and loss of function hESCs using CRISPR/Cas9 technology (Fig. 1d,f, Supplementary Fig. 2a,b) to yield four cell types: d0 (wild-type hESCs), d0 p63GOF (hESCs ectopically expressing p63), d7 p63WT (wild-type hESCs morphogen-treated, with endogenous p63), and d7 p63KO (hESCs morphogen-treated with no p63 expression). We used the deltaNp63-alpha isoform for our p63GOF cell line because it is predominantly expressed in our system, consistent with published reports of developing keratinocytes^{21–25} (Supplementary Fig. 2c,d). Furthermore, we designed the p63KO to mimic the alleles of the p63-null mouse⁸ (Supplementary Fig. 2d), whereby the DNA-binding domain is deleted. We verified loss of p63 protein in these cell lines through IF, western blot, and sequencing (Fig. 1e,g, Supplementary Fig. 2).

Previous studies indicate that p63 overexpression can drive surface ectoderm commitment²⁶, yet remarkably, expression of p63 in hESCs was insufficient to induce an exit from pluripotency and a switch towards epidermal differentiation (Fig. 1e, Supplementary Fig. 2e). Consistent with this observation, transcriptome analysis using RNA-seq revealed minimal expression changes between d0 and d0 p63GOF (only 320 genes changing), compared to more than 2,400 genes in d7 p63WT vs. d7 p63KO (Fig. 1h). Further, morphogen exposure resulted in a p63-independent exit from pluripotency and was required for p63 regulation of key transcription factors (TFs) associated with epithelial development (Fig. 1h, Supplementary Fig. 2f). These TFs, including *TFAP2C*, *KLF4*, and others, were all repressed by p63 upon morphogen treatment. Notably, other key epidermal and surface ectoderm-promoting developmental TFs, such as *JUN*^{27,28} and *MSX1*²⁹, were p63-

independent at this stage (Supplementary Fig. 2f). The morphogen-dependence yet p63-independence of these factors suggests that these regulators work in parallel to p63 to specify the surface ectoderm fate, further highlighting the importance of the morphogen contributions. d7 p63 overexpression (d7 p63GOF) resulted in no significant changes in expression of p63-dependent genes, indicating the existence of maximal endogenous p63 repression at this timepoint (Supplementary Fig. 2g). Thus we conclude that morphogenetic signaling promotes a simple epithelial state, while enabling p63 to modify the morphogen-induced transcriptome to drive epidermal fates.

The striking influence of morphogens on p63 activity led us to investigate whether differences in p63 genomic occupancy accounted for the altered transcriptional activity. p63 chromatin immunoprecipitation (ChIP-seq) in d0 p63GOF and d7 p63WT revealed 7,960 and 6,097 p63 binding sites, respectively, with the p63 motif significantly enriched under peaks in both datasets (Fig. 1i); remarkably, over 70% of the sites were identical between datasets (Fig. 1j,k), while 17% of peaks were gained in the d0 p63GOF (Supplementary Fig. 3a). Thus, differences in p63 occupancy cannot explain the dramatic morphogen-regulated p63 activity.

We next characterized how the morphogens and p63 affected chromatin accessibility and deposition of four histone modifications (H3K27ac, H3K27me3, H3K4me3, H3K4me1) using the Assay for Transposase Accessible Chromatin followed by sequencing (ATAC-seq) and histone ChIP-seq, respectively. Approximately 20,000 transposase accessible sites changed during the induction phase, with 14,000 opening and 6,000 closing between d0 and d7 p63WT (Fig. 2a). Over one third of the morphogen-dependent accessible sites became more accessible upon p63 loss (Fig. 2a,d). Comparison of established histone modifications in d7 p63WT vs. p63KO revealed significant differences in H3K27me3, yet no observable differences on activating promoter or enhancer marks (Supplementary Fig. 3a-c). p63 absence resulted in a significant decrease in signal of the H3K27me3 mark, whereas H3K27me3 increased in d0 p63GOF (Fig. 2b,d). Interestingly, unlike what might be expected from previous keratinocyte studies^{30,31}, global loss of H3K27me3 in the d7 p63KO did not coincide with decreased expression of *Ezh2* (Supplementary Fig. 3d). Furthermore, H3K27me3 ChIP-seq signal decreased during differentiation: 1396 of 1433 differential regions had higher signal in d0 vs d7 p63WT, and 1283 of these differential regions were p63-dependent. GO terms for the genes associated with morphogen-dependent H3K27me3 regions consisted of cell and neuron fate commitment (Supplementary Fig. 3e). ChromHMM analysis indicated most of the accessibility changes and p63 binding sites occur in enhancers (Supplementary Fig. 3f). We conclude that p63 edits a subset of the morphogen-induced accessibility changes and regulates the accumulation of H3K27me3 histone modifications.

Lineage selectors can act either directly or at a distance on the epigenetic landscape to alter accessibility or histone modification deposition³². To determine how p63 acts, we intersected the p63-dependent H3K27me3 regions and morphogen-dependent accessible sites with p63 binding sites, revealing that few of the p63 binding sites overlapped with either of these changing elements (Fig. 2e). These data indicate that most of the p63 epigenetic regulatory action occurs distal to p63 binding. Interestingly, when we assigned

p63 binding sites, morphogen-dependent accessible sites, and differential H3K27me3 regions to the nearest genes through GREAT, we found that these elements converge on a common gene set, despite each being in distinct genomic regions (Fig. 2f, Supplementary Fig. 3g,h).

To assess the connectivity and dynamics of the three-dimensional architecture between these distinct genomic regions, we employed cohesin HiChIP, a method analogous to Hi-C¹⁰, in all four cell types. We identified high-confidence chromatin contacts with 10 kb resolution using FitHiC³³ (Supplementary Fig. 4) and demonstrated that 54% of p63 ChIP-seq peaks in d7 p63WT participate in these chromatin connections (Fig. 3a). Additionally, most morphogen and p63-dependent dynamic elements participate in looping connections. Notably, only 34% of genes GREAT identified as having transcriptional start sites (TSSs) connected to p63 binding sites were verified by cohesin HiChIP, reinforcing the non-uniformity of the existing chromatin landscape (Supplementary Fig. 3,5).

For the 4,412 protein-coding, p63-dependent genes (> 1.5 FC in gene expression between d7 p63WT and d7 p63KO), we determined the connectivity of their TSSs to a p63 binding site, revealing that 13% of these genes were in direct contact with p63 via chromatin looping (1°) and 11% were in indirect contact via a morphogen-dependent accessible site or p63-dependent H3K27me3 region (2°) (Fig. 3b, Supplementary Table 1). Although more complex conformations through multiple elements (3°) were detected, random simulation demonstrated that p63 was not connected to p63-dependent genes by 3° connections at a frequency above random chance (FDR = 0.173); thus we focused on the 0°, 1°, and 2° p63 connections (Fig. 3b, Supplementary Fig. 6). Analysis of p63 connectivity to p63-independent genes (<1.5 FC in gene expression between d7 p63WT and d7 p63KO) was also conducted (Fig. 3b, Supplementary Table 1).

Interrogating the correlation between p63 connection to any protein-coding TSS and transcriptional regulation, we found that p63 connectivity was insufficient to regulate gene expression (Fig. 3b). Both p63-dependent and -independent genes connected to a p63 site were involved in organ development and cell differentiation, consistent with known p63 function (Fig. 3c)^{7,8}. Although connection of the TSS to p63 was not a guarantee of p63 regulation, the probability of transcriptional repression was significantly higher at genes connected to p63 (Fig. 3d). Furthermore, d7 p63-independent genes connected to p63 include keratinocyte differentiation genes whose expression becomes p63-dependent later during keratinocyte maturation, including *TP63* itself (Supplementary Table 2)^{22,34,35}. These data suggest that p63 and morphogen-regulated chromatin connections foreshadow future gene action. Notably, specific chromatin conformation types were not indicative of positive or negative p63 gene regulation (Supplementary Table 3). In all, a large subset of the morphogen and p63-dependent elements are physically connected at d7 (Fig. 3e), accounting for the ability of p63 to regulate the epigenetic landscape at a distance.

Next, we determined the extent to which p63 and the morphogens influenced connectivity (Fig. 3f). In 1° (middle panel) and 2° (right panel) connections, contacts between morphogen-dependent accessible sites and p63 binding sites were regulated by both the morphogens and p63, with loss of p63 abolishing the connections, and overexpression of

p63 failing to enhance them. Conversely, p63-H3K27me3 and p63-TSS interactions were enhanced by the morphogens and p63 overexpression, and weakened by p63 loss (Supplementary Fig. 7). Finally, we determined greater repression at TSSs connected to both p63 and morphogen-dependent accessible sites (Fig. 3g) than at TSSs connected to both p63 and an H3K27me3 peak (Supplementary Fig. 7b). These findings indicate that morphogen and p63-dependent conformational changes drive optimal p63-regulated transcription.

From our global analyses, we identified *TFAP2C*, a critical epithelial regulator¹¹, as a gene induced by morphogens and repressed by p63 that exhibits a complex chromatin architecture driving its regulation. To illustrate p63-morphogen interactions we dissected the p63 negative feedback regulation of this key developmental regulator (Fig. 4). Cohesin HiChIP analysis (Fig. 4a, Supplementary Fig. 8a-c) revealed a distal p63 binding site with three d7 p63WT connections to the TSS: through a direct contact, the adjacent morphogen-dependent accessible site, and the distal H3K27me3 peak, all within 400 kb. We confirmed our cohesin HiChIP with UMI-4C³⁶ using primer viewpoints around the three connections (Supplementary Fig. 9).

Comparison of the chromosome conformation among the different cell lines indicated that p63 presence enhances connectivity to all three of the main loops at d7, and in the absence of p63, the connections and transcriptional output collapse. Morphogen exposure connects p63 to the induced neighboring morphogen-dependent accessible site, but the connection relies on ongoing p63 expression to maintain it, as loss of p63 fails to uphold it despite morphogen presence.

To validate the importance of the morphogen-dependent accessible site, we removed the region using CRISPR/Cas9 and demonstrated a loss of morphogen-induced *TFAP2C* expression (Supplementary Fig. 8d). Furthermore, we hypothesized that removal of the p63 binding site should drive both *TFAP2C* and p63 expression, given our observation that *TFAP2C* induces p63 expression in hESCs (Li and Oro, unpublished results) and that p63 provides important early negative regulation of *TFAP2C*. To test this, we deleted the p63 binding site (p63BSKO) and found dramatically elevated levels of *TFAP2C* at d7, consistent with the predicted negative feedback modulation of *TFAP2C* by p63 (Fig. 4b,c, Supplementary Fig. 10). Moreover, d7 p63BSKO showed increased expression of p63, demonstrating the need for tight p63-morphogen regulation to control the levels of key developmental factors. Given that high levels of p63 in d7 p63GOF do not further alter gene expression (Supplementary Fig. 2g), these data confirm that p63 tightly modulates *TFAP2C* expression through its binding and connections at the locus. Histone ChIP-qPCR revealed a loss of H3K27me3 accumulation at both the TSS and the distal H3K27me3 site in d7 p63BSKO, while other non p63-connected sites remained unaffected (Fig. 4d). Similarly, the morphogen-dependent accessible site became more accessible in d7 p63BSKO, to levels found in d7 p63KO (Fig. 4d), confirming the connectivity of these distal elements.

Here we deepen our understanding of the interplay between morphogens and lineage selectors and find that morphogens provide the powerful driving force for cell state change by inducing expression of the lineage factor while also altering chromatin accessibility, histone modifications, and chromosome conformation. We speculate that morphogen-

induced TFs and epigenetic regulators control these landscape alterations to poise the chromatin for p63 action. p63 then modifies the morphogen-dependent landscape to drive surface ectoderm differentiation. We demonstrate that p63 cannot function without morphogens, implicating the importance of these downstream factors in helping p63-mediated transcriptional regulation. Sequence analysis of morphogen-dependent accessibility sites implicates combinatorial regulation by multiple TFs with p63 to drive this process.

Our results illustrate how chromatin connections to the lineage selector p63 are necessary and more likely to induce gene expression changes, but are not sufficient. Our finding that p63 at d7 is poised to act on later keratinocyte differentiation genes (Supplementary Table 2)^{22,34,35} suggests the existence of additional inductive influences after addition of RA/BMP that will enable broader p63-dependent transcription. This is functionally similar to “poised” histone modifications and provides a structural explanation of how the order of morphogen exposure can determine downstream transcriptional programs.

Previous studies have highlighted the importance of p63 interactions with chromatin remodelers such as DNMT3A and PRC2, suggesting a mechanism of action for p63 at these distal enhancer regions^{31,37,38}. Interestingly, as protein levels of Ezh2 are unaltered in our p63KO cells (Supplementary Fig. 3d), the reduction in H3K27me3 that we observe is likely dependent upon p63 interactions with a variety of other TFs and chromatin remodelers^{37,39} depending on the stage of differentiation.

This study has important implications for the apparent autonomy of lineage selectors and for the basis of morphogenesis. Our work suggests that small changes in morphogen activity can dramatically alter the induced chromosome landscape and connectivity, explaining how a single lineage selector like p63 can direct a panoply of transcriptional programs depending on specific morphogen exposure.

Methods

CRISPR/Cas9 guided genome editing

gRNAs were designed using the online tool available at <http://crispr.mit.edu/>⁴⁰, selected based on the highest scores and the least off-targets, and incorporated into a DNA fragment bearing all the components necessary for gRNA expression⁴¹. Donor sequences were designed by selecting 700 bp arms flanking left and right of the region to be modified. Both gRNAs and donor sequences were synthesized as 5-phosphorylated gene blocks (IDT) and cloned into a blunted plasmid with puromycin selection, except for gRNAs targeting the AAVS1 locus, which were acquired through Addgene (Plasmid # 72833)⁴². The d0 p63GOF line was generated by integrating the humanized Np63 mouse cDNA under the control of a Tetracycline Responsive Element (TRE) to the AAVS1 locus. Doxycycline (Sigma) was added to the media for 2 days at a concentration of 2 ug/ml to induce expression of p63 in hESCs. Scans of all Western blots used to validate cell lines are in Supplementary Fig. 11.

hESC culture and transfection

H9 human embryonic stem cells were cultured on Vitronectin Recombinant Human Protein (Life Technologies) in Essential 8 medium (E8, Life Technologies) as described previously¹⁶. Cells were passaged every three days as clumps with 0.5 mM EDTA (Lonza). For transfection, 2×10^6 cells were nucleofected using AmaxaTM P3 Primary Cell 4D-Nucleofector (Lonza) as recommended by the manufacturer, with no more than a 10 uL mix of 2 ug of plasmid carrying each gRNA, 2 ug of plasmid carrying hCas9 and 2–4 ug of plasmid carrying the donor DNA to repair the Cas9/gRNA induced break by homologous recombination. Cells were plated and allowed to recover for a minimum of 6h in E8 media supplemented with 2 uM thiazovivin (Stemgent). Drug selection with 1 ug/mL puromycin (InvivoGen) started 48h after transfection and lasted 2 days for loss of function cell lines, or continued for several days for gain of function cell line. Colonies were picked 10 days after selection and genotyped by PCR to confirm homozygosity. To generate the point mutations in the p63BSPM line, gRNAs were synthesized via in vitro transcription and transfected into cells with Lipofectamine CRISPRMAX (Invitrogen) along with Cas9 protein. The donor sequence supplied alongside was designed as a 200 bp single-stranded DNA oligo (an ssODN ultramer from IDT)^{43,44}.

In vitro epithelial hESC differentiation

For differentiation, 6.2×10^3 cells/cm² were plated as colonies on Vitronectin coated plates. Next day, media was changed to Essential 6 (E6, Life Technologies) supplemented with 1 uM RA (Sigma) and 5 ng/mL Recombinant Human BMP-4 (R&D Systems), and replaced every two days for seven days, at which point cells were dissociated with Accutase (StemCell Technologies) and collected for downstream analysis, or media was replaced to Defined Keratinocyte-SFM media (DKSFM, Life Technologies) for terminal differentiation into keratinocytes.

Immunofluorescence staining

Cells were cultured on glass cover slips in 12 wells, subjected to the appropriate treatment and fixed for 10 min at room temperature in 4% paraformaldehyde in PBS. Cells were permeabilized for 10 min with permeabilization buffer (0.1% Triton-X + 0.05% Tween-20 in PBS) and blocked for 30 min with 10% Horse Serum (Vector Laboratories) in permeabilization buffer. Antibodies at appropriate dilutions were incubated overnight at 4°C. Secondary antibodies were added at 1:500 dilution and incubated at room temperature protected from light for 1h. Cells were washed three times in Hoechst 1:10,000 in PBS, and glass cover slips were mounted onto glass slides with mounting medium before imaging. Antibodies were diluted in permeabilization buffer at the indicated dilutions: AP-2 γ (1:100, Cell Signaling 2320S), p63 (1:200, Genetex GTX102425), KRT18 (1:800, R&D AF7619), KRT14 (1:1000, BioLegend 906001), OCT4 (1:100, BioLegend 631902).

RNA extraction and library preparation

For RNA extraction, cells were lysed directly in Trizol (Invitrogen), purified as indicated by the manufacturer, and then run through RNeasy columns (Qiagen). Libraries for RNA-seq were prepared using TrueSeq RNA Library Prep Kit v2 (Illumina) according to the

manufacturer's protocol. Real time PCR was performed with SYBR Green PCR master mix (Life Technologies) and in a Stratagene real time PCR machine. Primer sequences for detecting different Np63 isoform variants were described by Nylander et al (2002)²⁵.

Chromatin immunoprecipitation (ChIP) and library preparation

Cells were cross-linked in suspension for 10 min using freshly prepared 1% formaldehyde (Thermo Scientific) in PBS. Subsequently, glycine was added to a final concentration of 0.125 M to quench formaldehyde, and cells were washed twice with cold PBS. 60×10^6 or 10×10^6 cross-linked cells were used per ChIP for p63 or histone marks, respectively. Cells were lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS, 1X protease inhibitors) for 30 minutes on ice and sonicated for 2h using a Bioruptor (Diagenode) to achieve a chromatin size between 200 and 300 bp. Chromatin was centrifuged to remove debris, quantified and diluted in dilution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1X protease inhibitors) to achieve a 0.1% SDS final concentration. Sheared chromatin was incubated overnight at 4° with appropriate antibodies, followed by incubation with 30 uL of agarose G beads (Invitrogen) for 4h at 4°C. Antibodies were used at the indicated concentrations per ChIP (per 10×10^6 cells): p63 (12 uL, Active Motif 39739), H3K4me3 (5 ug, Abcam ab8580), H3K4me1 (5 ug, Abcam ab8895), H3K27Ac (5 ug, Abcam ab4729), H3K27me3 (5 ug, Millipore 07-449). Beads were washed twice each with low salt buffer (50 mM Tris-HCl pH 8.0, 0.15 M NaCl, 1 mM EDTA pH 8.0, 0.1% SDS, 1% triton X-100, 0.1% sodium deoxycholate), high salt buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 1 mM EDTA pH 8.0, 0.1% SDS, 1% triton X-100, 0.1% sodium deoxycholate), and LiCl buffer (50 mM Tris-HCl pH 8.0, 0.15 M LiCl, 1 mM EDTA pH 8.0, 1% Nonidet P-40, 0.1% sodium deoxycholate). DNA was eluted in 100 uL of elution buffer (50 mM NaHCO₃, 1% SDS) and crosslinks were reversed with 4 uL of 5 M NaCl incubated overnight at 67°C. RNA was removed by adding 1 uL of 10 mg/mL RNase A and incubating for 30 min at 37°C. DNA was cleaned using the Qiagen Qiaquick PCR purification kit and quantified using Qubit (Invitrogen). Between 5 and 10 ng of pooled DNA were used for library preparation using NEBNext kit (New England Biolabs) and Agencourt AMPure XP beads (Beckman) according to the manufacturer's protocol. Single-read libraries were sequenced on Illumina NextSeq 500 sequencer and two independent, biological replicates were sequenced per ChIP.

Assay for Transposase Accessible Chromatin (ATAC-seq)

ATAC-seq was performed as described⁴⁵. Briefly, after treatment with Accutase, 7×10^4 cells were washed with cold PBS and lysed using 0.1% NP40 in RSB buffer. Nuclei pellets were Tn5 transposed using the DNA Sample Preparation Kit from Nextera®. Libraries were amplified for 9–15 total cycles using the Ad1 and Ad 2.1–2.16 barcodes. Libraries were purified using the Min-Elute columns (Qiagen) and eluted with 10 uL of buffer EB. Library DNA concentrations were determined with Bioanalyzer High-Sensitivity DNA analysis kit (Agilent). Paired-end libraries were sequenced initially on a MiSeq sequencer and analyzed using a custom script to determine the signal enrichment over background at TSSs over a 2 kb window. Only libraries that had enrichment scores above 6 were sequenced deeper in an Illumina NextSeq 500 sequencer and two independent, biological replicates were sequenced per sample.

Cohesin HiChIP

In situ chromosome conformation capture (3C) was performed as described earlier¹⁰. Briefly, 25×10^6 cells were crosslinked and digested with MboI (NEB). After digest, biotin was incorporated into the sticky ends of fragments before ligation. Cohesin ChIP was performed to enrich for proximity ligations bound to cohesin, using an SMC1 antibody (Bethyl, A300–055A). The library quality was assessed on a MiSeq sequencer before sequencing on an Illumina HiSeq 4000. Three replicates were pooled and sequenced across two HiSeq lanes for a total of 1200 million reads per sample.

UMI-4C

UMI-4C was performed as described previously³⁶. Briefly, 1×10^7 cells were crosslinked in suspension with 1% formaldehyde then quenched with glycine, and pelleted cells were lysed in 1 mL fresh cold lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1% TX-100, 1x protease inhibitors) on ice. The nuclei were extracted and resuspended in water, DpnII buffer, and 10% SDS for DpnII digestion. Three rounds of DpnII digestion were performed, adding 200 U of HC DpnII (NEB) for 2 hours, incubating overnight, and then 2 more hours all at 37°C with rotation. After inactivation of DpnII, the 3C reactions were diluted to 7 mL and 13.6 uL of HC T4 ligase (NEB) were added for overnight ligation at 16°C. Crosslinks were reversed overnight at 65°C with Proteinase K (Qiagen) and DNA was treated with RNase A (Qiagen) for 45 minutes at 37°C. The DNA was then purified with one phenol-chloroform extraction (ThermoScientific) and ethanol precipitation, and resuspended in 150 uL of 10 mM Tris-HCl, pH 8.0. DNA was quantified using Qubit before proceeding with library preparations. Aliquots of chromatin were taken before and after DpnII digestion and after overnight ligation to determine efficiency of enzymatic reactions. UMI-4C library preparation was performed as described previously³⁶. Briefly, 5–10 ug of 3C library was sonicated in a Diagenode Bioruptor to achieve a chromatin size between 400 and 600 bp. End-repair (NEB kit), A-tailing (NEB kit) and 5'-diphosphorylation (NEB) of ends were performed as recommended by the manufacturer. TruSeq Illumina indexed adapters were ligated to the 3'-end of the DNA using Quick Ligase (NEB). Libraries were generated by nested PCR at particular genomic loci using GoTaq Hot Start Polymerase (Promega) and 200 ng of DNA template (Extended Data Fig. 9). The primer for the second PCR included the Illumina dangling adaptor for enrichment of the product from the first PCR, as described³⁶. Two independent, biological replicates of paired-end libraries were sequenced in the Illumina NextSeq 500 sequencer.

Next Generation Sequencing processing of ChIP-seq and ATAC-seq data

Quality control of fastq files was done with FASTQC⁴⁶. Sequence alignment to hg19 was performed using tophat for RNA-seq (parameters: p 10 --library-type fr-firststrand -r 100 --mate-std-dev 100), or bowtie²⁴⁷ for ChIP-seq (parameters: -p 24 -S -a -m 1 --best --strata) and ATAC-seq (parameters: -p 24 -S -m 1 -X 2000). Aligned reads were processed to remove PCR duplicates using samtools⁴⁸ and mitochondrial DNA (for ATAC-seq datasets). Peak calling was carried out with MACS2⁴⁹ using default settings with a p-value of 0.05. To filter out non-reproducible peaks, called peaks from biological replicates were processed through the Irreproducible Discovery Rate (IDR) framework implemented in R⁵⁰.

Differential counting, heatmaps, and average profiles

For ChIP-seq and ATAC-seq, a union list of the MACS2 called peaks per sample not filtered by IDR was generated using bedtools merge⁵¹, and raw reads covering each region were recovered from bam files using bedtools multicoverage. For RNA-seq, raw reads on reference genes were recovered using HOMER (version 4.8⁵² analyzeRepeats.pl command). To test for differential counting, raw reads were compared using DESeq2 package implemented in R⁵³, and filtered based on an adjusted p-value of < 0.05 and 2 fold change. For heatmaps and average profiles, tag counts were recovered +/- 2 kb from the peak summit using HOMER annotatePeaks.pl command with -hist 25 -ghist or -hist 25 parameters. Heatmap images were generated using Java TreeView⁵⁴. Average profiles and scatter plots were plotted using Python matplotlib.

Motif discovery and Gene Ontology

De novo motif discovery was performed using Homer findMotifsGenome function with -size 200 as a parameter. Gene ontology analysis was performed using DAVID⁵⁵ for RNA-seq data and GREAT⁵⁶ for ChIP-seq and ATAC-seq data.

Chromatin state determination

ChromHMM software⁵⁷ was used to learn and identify chromatin states as instructed in the manual. Encode chromatin segmentation using ChromHMM was used as a reference to label each state using a custom script using bedtools intersect. The enrichment of each state for a set of peaks was calculated using the NeighborhoodEnrichment command and compared among samples using a custom script. Enrichments were plotted using Python matplotlib library.

Analysis of UMI-4C data

UMI-4C data was aligned and analyzed using HiC-Pro⁵⁸ and the DpnII segmented genome annotation file. Interaction matrices of 5 kb resolution were generated and used to create Virtual 4C profiles through a custom python script and the matplotlib library.

Analysis of cohesin HiChIP data

HiChIP paired end reads were aligned to hg19 using HiC-Pro⁵⁸. Duplicate reads were removed, assigned to MboI restriction fragments, filtered for valid interactions, and then used to generate binned interaction matrices of both 5 kb and 10 kb resolution. The 5 kb interaction matrices were used to visualize contacts by Virtual 4C, similar to the UMI-4C analysis. The 10 kb interaction matrix was used to call high confidence contacts (defined as counts > 10, FDR < 0.001) using the contact caller, FitHiC³³. Default FitHiC settings were used to generate an FDR for each bin pair. These high confidence cohesin contacts were used in the subsequent analyses.

Contact connection analyses

An element was considered participating, or anchored, in cohesin connections, if it possessed at least one high confidence contact bin in a given cell type. When considering ways in which p63 was connected to a TSS (defined as TSS +/- 5 kb), four chromatin

conformations were considered. 0° connections were defined as two elements overlapping in physical space (e.g. p63 BS contained within the TSS). 1° connections were defined as one element anchored in one bin of a cohesin contact and the second element anchored in the other bin. More complicated connections between elements were also considered: 2° connections were defined as two elements in distinct physical space both forming 1° connections to the same third element. Finally, 3° connections were when one element formed a 1° connection to a second element, which also formed a 1° connection to a third element, which also had a 1° connection to the fourth (target) element. All elements in both 2° and 3° configurations were in distinct physical space (i.e. non-overlapping).

Empirical cumulative distribution function analyses

ecdf was performed to determine whether the cumulative levels of log₂ FC in a subset of genes was differential compared to all protein-coding genes using a combination of custom unix and python script. Only protein-coding genes for which there was a FC value calculated using DESeq2 on the RNA-seq data were used in constructing the ecdfs.

Differential contact analysis

The Bioconductor package edgeR⁵⁹ was used to perform multiple comparison differential analysis of high confidence FitHiC contacts in d0, d0 p63GOF, d7 p63WT, and d7 p63KO cells. The Anderson-Darling 2-sample test, a modification of the K-S test, which gives greater weight to the tails, was used to calculate statistical significance between populations of the fold change in contact connectivity⁶⁰.

Statistics and Reproducibility

All data represent similar results from 3 independent, biological experiments and cell cultures (n=3) unless otherwise stated. Number of independent, biological experiments for deep sequencing replicates are indicated in the appropriate Methods subsections (n=2 for all, except for HiChIP where n=3). The center values of all graphs depict the mean, unless otherwise stated. Significance values were calculated using a Student's two-sided t-test, unless otherwise indicated. The following annotation applies for all figures: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.005, **** p-value < 1×10⁻¹⁰. Statistical methods for specific analyses are detailed above in the corresponding Methods subsections.

The minimum, maximum, and percentile distribution of the violin plots in Fig. 3f and Supplementary Fig. 7a and 7c can be found in Supplementary Table 4. In all the panels the distributions are the log₂ FC or the FDR for d0 vs d7 p63WT (red), d7 p63WT vs d7 p63KO (green), and d0 vs d0 p63GOF (blue). No statistics were calculated comparing the population distributions in Supplementary Fig. 7c. The following are the exact p-values for comparisons between the population distributions for Fig. 3f and Supplementary Fig 7a by a 2-sample Anderson-Darling test, which calculates the probability that two datasets come from the same population.

For all contacts (Fig 3f, left panel), the p-values are 2.23e-308 (red vs green) 2.23e-308 (red vs blue) and 2.23e-308 (green vs blue). Please note that this is the smallest value that a computer is able to represent. Comparing the populations of p63 – morph-dep ATAC 1°

contacts (Fig 3f, middle panel), the p-values are 1.74e-22 (red vs green), 2.07e-21 (red vs blue), and 2.44e-74 (green vs blue). Whereas the p-values for the subset of p63 - moph-dep ATAC 1° contacts in which both elements are connected to the same TSS (Fig 3f, right panel) are 1.74e-22 (red vs green), 2.07e-21 (red vs blue), 2.44e-74 (green vs blue).

The p-values for comparing the populations of p63 – TSS contacts are 2.23e-308 (red vs green) 6.68e-15 (red vs blue) and 2.23e-308 (green vs blue). For p63 – p63-dep H3K27me3 1° contacts (Supplementary Fig. 7a, middle panel), the p-values are 9.29e-209 (red vs green), 2.53e-26 (red vs blue), and 6.89e-184 (green vs blue). Whereas the p-values for the subset of p63 - p63-dep H3K27me3 1° contacts in which both elements are connected to the same TSS (Fig 7a, right panel) are 2.34e-65 (red vs green), 5.12e-11 (red vs blue), 1.25e-54 (green vs blue).

Code Availability

Custom scripts described in the Methods are available on GitHub – username OroLabStanford.

Data Availability

All sequencing data is available through Gene Expression Omnibus (GEO) accession number: GSE114846.

A **Life Sciences Reporting Summary** for this publication is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of the Oro Laboratory, P. Greenside, J. Wysocka, A. Kundaje, O. Wapinski, and D. Webster for helpful discussions and comments. This work was supported by CIRM Tools grant RT3-07796 (A.E.O.), NIAMS/NIH grant F32AR070565 (J.M.P.), NIAMS/NIH grant AR45192 (P.A.K.), NIH P50 HG007735 (H.Y.C.), and NIH K99/R00 5R00AR065490 (X.B.). H.Y.C. is an Investigator of the Howard Hughes Medical Institute.

References

1. Walmsley GG et al. Induced pluripotent stem cells in regenerative medicine and disease modeling. *Curr Stem Cell Res Ther* 9, 73–81 (2014). [PubMed: 24359141]
2. Inoue H, Nagata N, Kurokawa H & Yamanaka S iPS cells: a game changer for future medicine. *The EMBO journal* 33, 409–417, doi:10.1002/embj.201387098 (2014). [PubMed: 24500035]
3. Hanna J et al. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923 (2007). [PubMed: 18063756]
4. Umegaki-Arao N et al. Induced pluripotent stem cells from human revertant keratinocytes for the treatment of epidermolysis bullosa. *Sci Transl Med* 6, 264ra164, doi:10.1126/scitranslmed.3009342 (2014).
5. Sebastiano V et al. Human COL7A1-corrected induced pluripotent stem cells for the treatment of recessive dystrophic epidermolysis bullosa. *Sci Transl Med* 6, 264ra163, doi:10.1126/scitranslmed.3009540 (2014).
6. Zaret KS & Carroll JS Pioneer transcription factors: establishing competence for gene expression. *Genes & development* 25, 2227–2241, doi:10.1101/gad.176826.111 (2011). [PubMed: 22056668]

7. Yang A et al. p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature* 398, 714–718, doi:10.1038/19539 (1999). [PubMed: 10227294]
8. Mills AA et al. p63 is a p53 homologue required for limb and epidermal morphogenesis. *Nature* 398, 708–713, doi:10.1038/19531 (1999). [PubMed: 10227293]
9. Lupien M et al. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* 132, 958–970, doi:10.1016/j.cell.2008.01.018 (2008). [PubMed: 18358809]
10. Mumbach MR et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nature methods* 13, 919–922, doi:10.1038/nmeth.3999 (2016). [PubMed: 27643841]
11. Qiao Y et al. AP2gamma regulates neural and epidermal development downstream of the BMP pathway at early stages of ectodermal patterning. *Cell research* 22, 1546–1561, doi:10.1038/cr.2012.122 (2012). [PubMed: 22945355]
12. Metallo CM, Ji L, de Pablo JJ & Palecek SP Retinoic acid and bone morphogenetic protein signaling synergize to efficiently direct epithelial differentiation of human embryonic stem cells. *Stem cells* 26, 372–380 (2008). [PubMed: 17962700]
13. Itoh M et al. Generation of 3D skin equivalents fully reconstituted from human induced pluripotent stem cells (iPSCs). *PloS one* 8, e77673, doi:10.1371/journal.pone.0077673 (2013). [PubMed: 24147053]
14. Guenou H et al. Human embryonic stem-cell derivatives for full reconstruction of the pluristratified epidermis: a preclinical study. *Lancet* 374, 1745–1753, doi:10.1016/S0140-6736(09)61496-3 (2009). [PubMed: 19932355]
15. Coraux C et al. Reconstituted skin from murine embryonic stem cells. *Curr Biol* 13, 849–853 (2003). [PubMed: 12747834]
16. Chen G et al. Chemically defined conditions for human iPSC derivation and culture. *Nature methods* 8, 424–429, doi:10.1038/nmeth.1593 (2011). [PubMed: 21478862]
17. Owens DW & Lane EB The quest for the function of simple epithelial keratins. *BioEssays : news and reviews in molecular, cellular and developmental biology* 25, 748–758, doi:10.1002/bies.10316 (2003).
18. Senoo M, Pinto F, Crum CP & McKeon F p63 Is essential for the proliferative potential of stem cells in stratified epithelia. *Cell* 129, 523–536, doi:10.1016/j.cell.2007.02.045 (2007). [PubMed: 17482546]
19. Koster MI & Roop DR Mechanisms regulating epithelial stratification. *Annu Rev Cell Dev Biol* 23, 93–113, doi:10.1146/annurev.cellbio.23.090506.123357 (2007). [PubMed: 17489688]
20. Green H, Easley K & Iuchi S Marker succession during the development of keratinocytes from cultured human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 100, 15625–15630, doi:10.1073/pnas.0307226100 (2003). [PubMed: 14663151]
21. Yang A et al. p63, a p53 homolog at 3q27–29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Molecular cell* 2, 305–316 (1998). [PubMed: 9774969]
22. Truong AB, Kretz M, Ridky TW, Kimmel R & Khavari PA p63 regulates proliferation and differentiation of developmentally mature keratinocytes. *Genes & development* 20, 3185–3197, doi:10.1101/gad.1463206 (2006). [PubMed: 17114587]
23. Romano RA, Ortt K, Birkaya B, Smalley K & Sinha S An active role of the DeltaN isoform of p63 in regulating basal keratin genes K5 and K14 and directing epidermal cell fate. *PloS one* 4, e5623, doi:10.1371/journal.pone.0005623 (2009). [PubMed: 19461998]
24. LeBoeuf M et al. Hdac1 and Hdac2 act redundantly to control p63 and p53 functions in epidermal progenitor cells. *Dev Cell* 19, 807–818, doi:10.1016/j.devcel.2010.10.015 (2010). [PubMed: 21093383]
25. Nylander K et al. Differential expression of p63 isoforms in normal tissues and neoplastic cells. *J Pathol* 198, 417–427, doi:10.1002/path.1231 (2002). [PubMed: 12434410]
26. Aberdam E et al. A pure population of ectodermal cells derived from human embryonic stem cells. *Stem cells* 26, 440–444 (2008). [PubMed: 18032703]
27. Li G et al. c-Jun is essential for organization of the epidermal leading edge. *Dev Cell* 4, 865–877 (2003). [PubMed: 12791271]

28. Mehic D, Bakiri L, Ghannadan M, Wagner EF & Tschachler E Fos and jun proteins are specifically expressed during differentiation of human keratinocytes. *J Invest Dermatol* 124, 212–220, doi: 10.1111/j.0022-202X.2004.23558.x (2005). [PubMed: 15654976]
29. McLarren KW, Litsiou A & Streit A DLX5 positions the neural crest and preplacode region at the border of the neural plate. *Developmental biology* 259, 34–47 (2003). [PubMed: 12812786]
30. Ezhkova E et al. Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell* 136, 1122–1135, doi:10.1016/j.cell.2008.12.043 (2009). [PubMed: 19303854]
31. Rapisarda V et al. p63 transcription factor regulates nuclear shape and expression of nuclear envelope-associated genes in epidermal keratinocytes. *J Invest Dermatol*, doi:10.1016/j.jid.2017.05.013 (2017).
32. Wamstad JA et al. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* 151, 206–220, doi:10.1016/j.cell.2012.07.035 (2012). [PubMed: 22981692]
33. Ay F, Bailey TL & Noble WS Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome research* 24, 999–1011, doi:10.1101/gr.160374.113 (2014). [PubMed: 24501021]
34. Barton CE et al. Novel p63 target genes involved in paracrine signaling and keratinocyte differentiation. *Cell death & disease* 1, e74, doi:10.1038/cddis.2010.49 (2010). [PubMed: 21151771]
35. Koh LF, Ng BK, Bertrand J & Thierry F Transcriptional control of late differentiation in human keratinocytes by TAP63 and Notch. *Exp Dermatol* 24, 754–760, doi:10.1111/exd.12764 (2015). [PubMed: 26013684]
36. Schwartzman O et al. UMI-4C for quantitative and targeted chromosomal contact profiling. *Nature methods* 13, 685–691, doi:10.1038/nmeth.3922 (2016). [PubMed: 27376768]
37. Rinaldi L et al. Dnmt3a and Dnmt3b Associate with Enhancers to Regulate Human Epidermal Stem Cell Homeostasis. *Cell Stem Cell* 19, 491–501, doi:10.1016/j.stem.2016.06.020 (2016). [PubMed: 27476967]
38. Wu H et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 329, 444–448, doi:10.1126/science.1190485 (2010). [PubMed: 20651149]
39. Lin-Shiao E et al. KMT2D regulates p63 target enhancers to coordinate epithelial homeostasis. *Genes & development* 32, 181–193, doi:10.1101/gad.306241.117 (2018). [PubMed: 29440247]

Methods-only References

40. Hsu PD et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* 31, 827–832, doi:10.1038/nbt.2647 (2013).
41. Mali P, Esvelt KM & Church GM Cas9 as a versatile tool for engineering biology. *Nature methods* 10, 957–963, doi:10.1038/nmeth.2649 (2013). [PubMed: 24076990]
42. Natsume T, Kiyomitsu T, Saga Y & Kanemaki MT Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell reports* 15, 210–218, doi: 10.1016/j.celrep.2016.03.001 (2016). [PubMed: 27052166]
43. Jacobi AM et al. Simplified CRISPR tools for efficient genome editing and streamlined protocols for their delivery into mammalian cells and mouse zygotes. *Methods* 121–122, 16–28, doi: 10.1016/j.ymeth.2017.03.021 (2017).
44. Liang X, Potter J, Kumar S, Ravinder N & Chesnut JD Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *Journal of biotechnology* 241, 136–146, doi:10.1016/j.jbiotec.2016.11.011 (2017). [PubMed: 27845164]
45. Buenrostro JD, Giresi PG, Zaba LC, Chang HY & Greenleaf WJ Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods* 10, 1213–1218, doi:10.1038/nmeth.2688 (2013). [PubMed: 24097267]

46. Andrews S FastQC: a quality control tool for high throughput sequence data., <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>> (2010).
47. Hwang S, Kim E, Lee I & Marcotte EM Systematic comparison of variant calling pipelines using gold standard personal exome variants. *Sci Rep* 5, 17875, doi:10.1038/srep17875 (2015). [PubMed: 26639839]
48. Li H et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079, doi:10.1093/bioinformatics/btp352 (2009). [PubMed: 19505943]
49. Zhang Y et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology* 9, R137, doi:10.1186/gb-2008-9-9-r137 (2008). [PubMed: 18798982]
50. Li H A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993, doi:10.1093/bioinformatics/btr509 (2011). [PubMed: 21903627]
51. Quinlan AR & Hall IM BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842, doi:10.1093/bioinformatics/btq033 (2010). [PubMed: 20110278]
52. Heinz S et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* 38, 576–589, doi:10.1016/j.molcel.2010.05.004 (2010). [PubMed: 20513432]
53. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550, doi:10.1186/s13059-014-0550-8 (2014). [PubMed: 25516281]
54. Saldanha AJ Java Treeview--extensible visualization of microarray data. *Bioinformatics* 20, 3246–3248, doi:10.1093/bioinformatics/bth349 (2004). [PubMed: 15180930]
55. Huang da W, Sherman BT & Lempicki RA Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4, 44–57, doi:10.1038/nprot.2008.211 (2009). [PubMed: 19131956]
56. McLean CY et al. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495–501, doi:10.1038/nbt.1630 (2010).
57. Ernst J & Kellis M ChromHMM: automating chromatin-state discovery and characterization. *Nature methods* 9, 215–216, doi:10.1038/nmeth.1906 (2012). [PubMed: 22373907]
58. Servant N et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome biology* 16, 259, doi:10.1186/s13059-015-0831-x (2015). [PubMed: 26619908]
59. Anders S et al. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature protocols* 8, 1765–1786, doi:10.1038/nprot.2013.099 (2013). [PubMed: 23975260]
60. Scholz F & Stephens M K-Sample Anderson-Darling Tests. *J Am Stats Assoc* 82, 918–924 (1987).

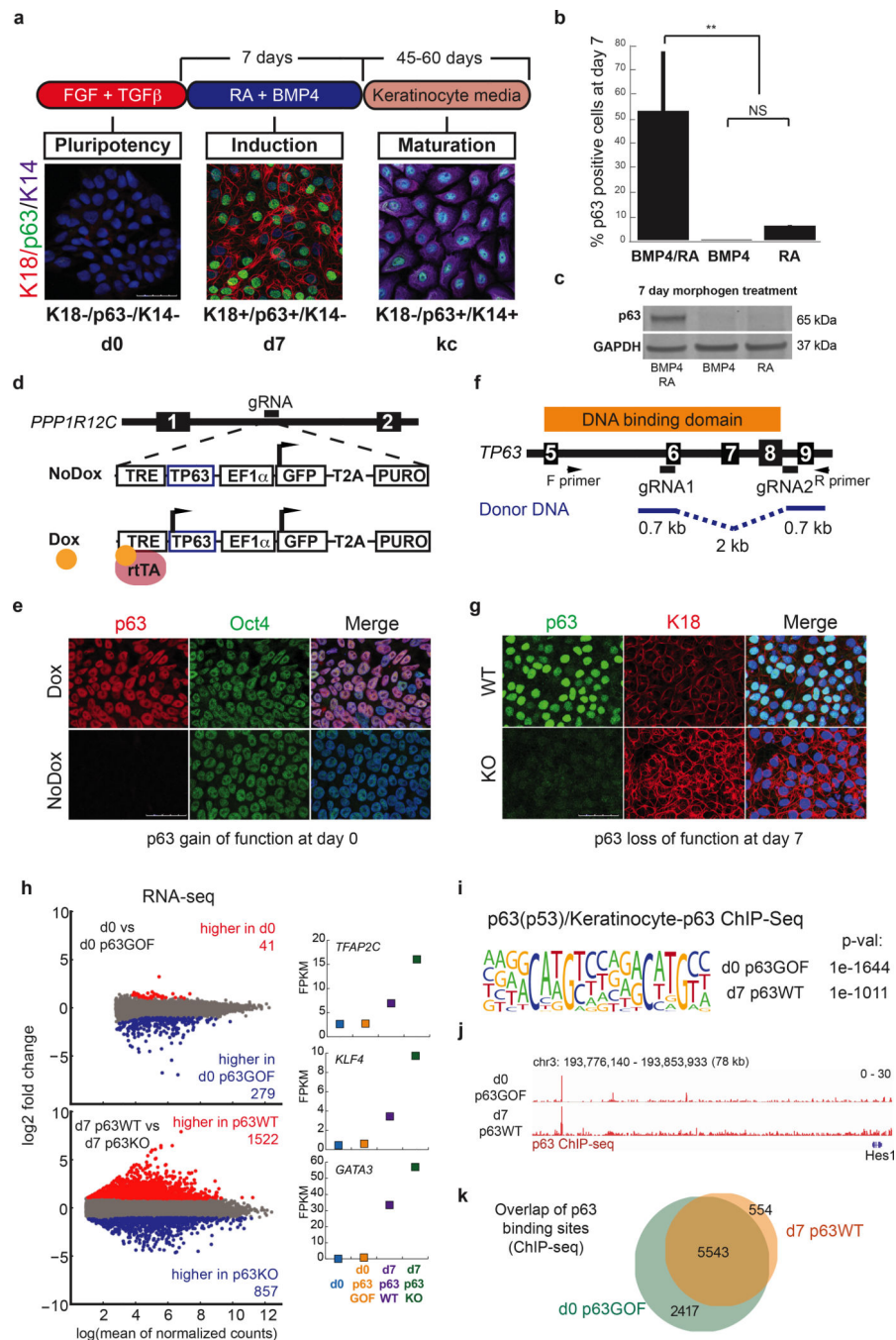


Figure 1. Morphogens and p63 cooperate to drive early epithelial differentiation. (a) RA/BMP4 treatment of hESCs for 7 days induces K18 and p63 expression. Functional keratinocytes (kc) expressing K14 and p63 grow out in kc selection media (n=3). Scale bar: 50 μ m. (b,c) hESCs need exposure to both RA and BMP4 to achieve high p63 expression. Error bars represent s.d., n=3; p-values (Tukey HSD Post-hoc Test): BMP4 vs RA/BMP4 (***,p=0.0011), BMP vs RA (NS-not significant, p=0.7591), RA/BMP4 vs RA (***,p=0.0022). (d) Strategy for generating the d0 p63GOF cell line. Numbered black boxes

signify exons. (e) Expression of p63 in the d0 p63GOF line (n=3). Scale bar: 50 μ m. (f) Strategy for generating the p63KO line. (g) IF validation of the p63KO line (n=3). Scale bar: 50 μ m. (h) Differential expression analysis from RNA-seq (measured by DESeq2) between d0 and d0 p63GOF (upper panel), and d7 p63WT and p63KO (lower panel). Genes either have no change in expression (gray), increased expression (> 2 fold change) in the d0 or d7 p63WT (red), or decreased expression (< -2 fold change) in the d0 or d7 p63WT. Key TFs associated with epithelial development are induced by the morphogens and repressed by p63. (i) Using HOMER analysis, the p63 motif was the most significantly recovered motif under p63 ChIP-seq peaks. (j) p63 binds distal to TSSs, as depicted at the *HES1* locus, and to the same sites in d0 p63GOF and d7 p63WT. Tracks represent n=2. (k) p63 binds to similar sites genome-wide with and without morphogen presence.

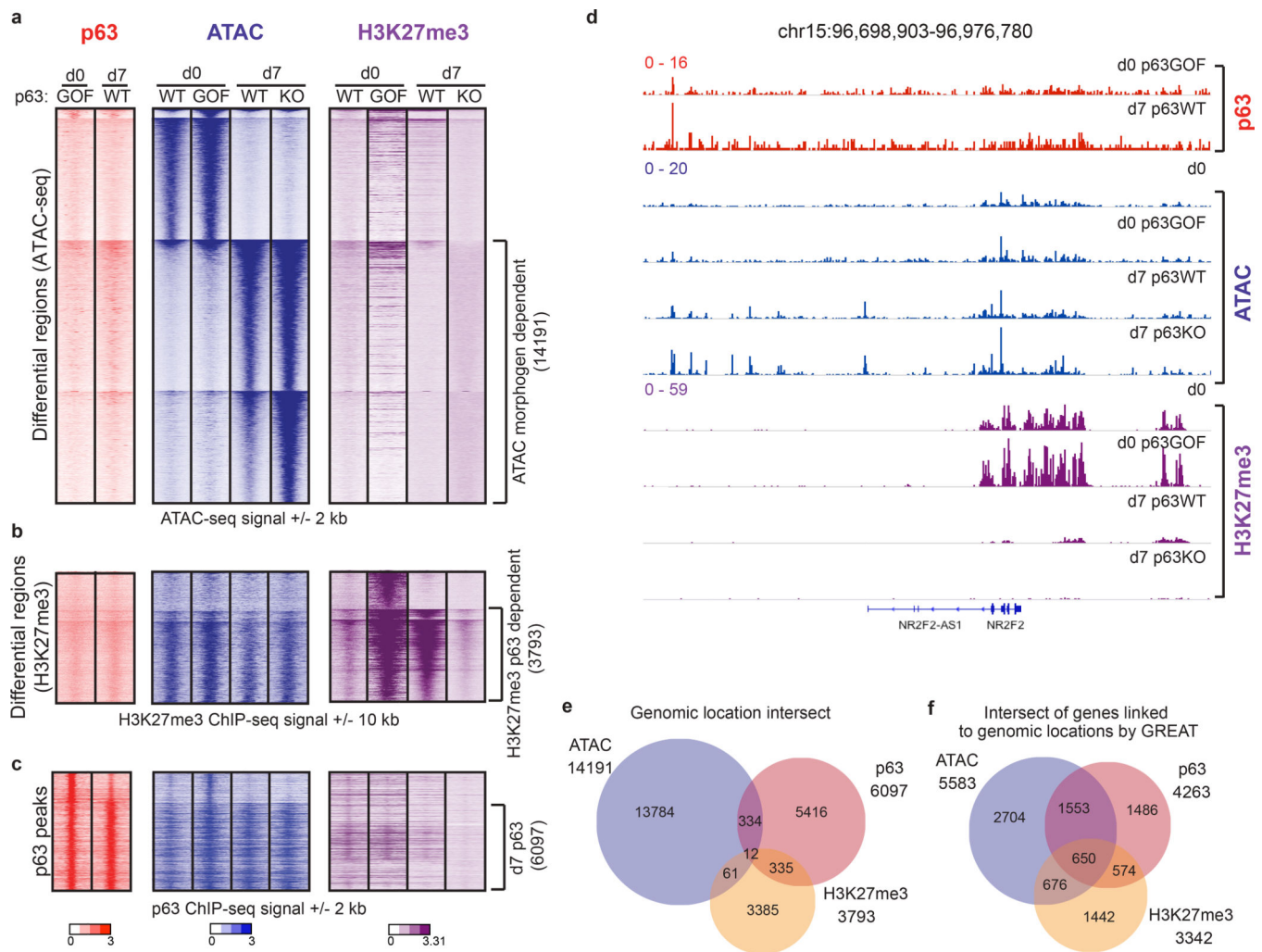


Figure 2.

The morphogens establish an epigenetic landscape that p63 modifies at a distance. (a) Differential accessible regions between d0 and d7 p63WT as analyzed using DESeq2 on ATAC-seq signal. Heatmaps represent the signal at these ATAC regions within the various cell types and assays: p63 ChIP-seq signal (red, left panel), ATAC-seq signal (blue, middle panel), and H3K27me3 ChIP-seq signal (purple, right panel). 14,191 differential regions become more accessible upon morphogen treatment (morphogen-dependent). (b) Differential H3K27me3 regions between d7 p63WT and p63KO as analyzed by DESeq2. Heatmaps represent the same datasets as (a) only signal is shown at the 3,793 differential H3K27me3 sites. (c) ATAC-seq (blue) and H3K27me3 (purple) signal at p63 binding sites (red). (d) Signal intensities of p63 ChIP-seq, ATAC-seq, and H3K27me3 ChIP-seq shown at the *NR2F2* locus. Tracks represent n=2. (e) The overlap of genomic regions that are differential as measured in (a), (b), and (c). The genomic location intersect is very low. (f) GREAT analysis linking the above differential regions to the closest gene shows that these elements converge on a similar gene set. While their physical genomic locations do not overlap, they are linked to a common gene via GREAT.

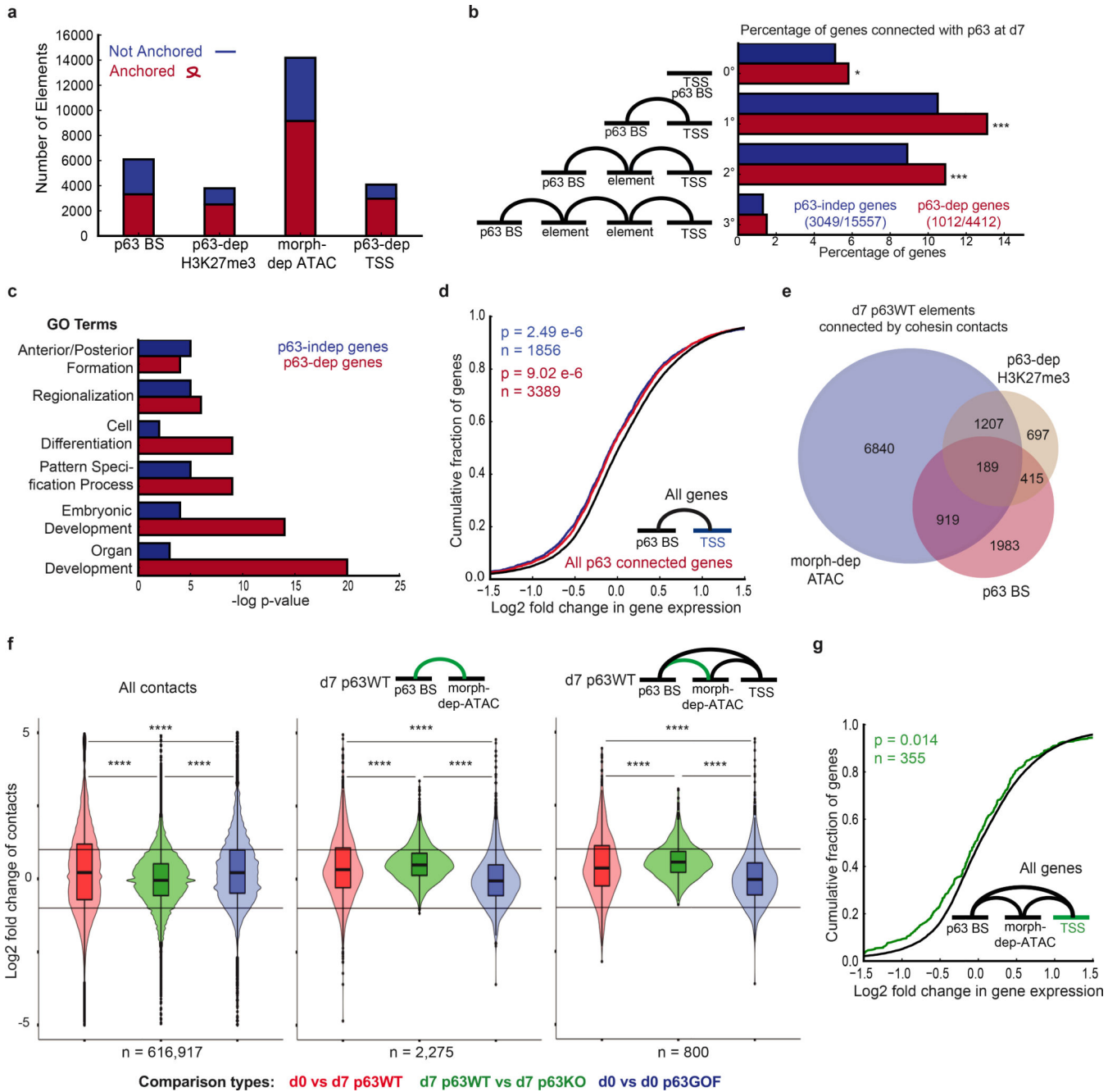


Figure 3. p63 - TSS connections are associated with negative regulation genome-wide. (a) Number of p63 binding sites (BS), p63-dependent (p63-dep) H3K27me3 sites, morphogen-dependent (morph-dep) ATAC sites, and p63-dep TSSs participating in chromatin looping (Anchored, red) vs those that do not (Not Anchored, blue). (b) Percentage of p63-independent (p63-indep) genes (blue) and p63-dep genes (red), whose TSS is connected to p63 by direct binding (0°), direct contact (1°), or via one (2°) or two (3°) morph-dep ATAC and/or p63-dep H3K27me3 elements. 1012/4412 p63-dep genes and 3049/15557 p63-indep genes are connected to p63 via one or more of these chromatin conformations. FDR by monte carlo

simulation *FDR<0.05,***FDR<0.001 (Supplementary Fig. 6a). (c) Gene Ontological Terms associated with p63-connected, p63-indep genes (blue) and p63-dep genes (red). (d) Empirical cumulative distribution function (ecdf) of the log₂ FC in gene expression between d7 p63WT vs d7 p63KO cells for all p63-connected genes (red) and 1° p63-connected genes (blue) compared to all genes (black). n = number of genes. 2-sided t-test. (e) 1° contact connections between p63 BS (red), p63-dep H3K27me3 (gold), and morph-dep ATAC (blue). (f) Change in connectivity strength between various cell types of all contacts (left panel), p63 - morph-dep ATAC contacts (middle panel), and p63 - morph-dep ATAC contacts in which both elements are connected to the TSS (right panel). n = number of contacts. Anderson-Darling 2-samples test ****p<1×10⁻¹⁰. (g) ecdf of the expression level changes (d7 p63WT vs d7 p63KO) of genes whose TSS is connected to a p63 BS and morph-dep ATAC site, which are connected to each other (green) compared to all genes (black). n = number of genes. 2-sided t-test.

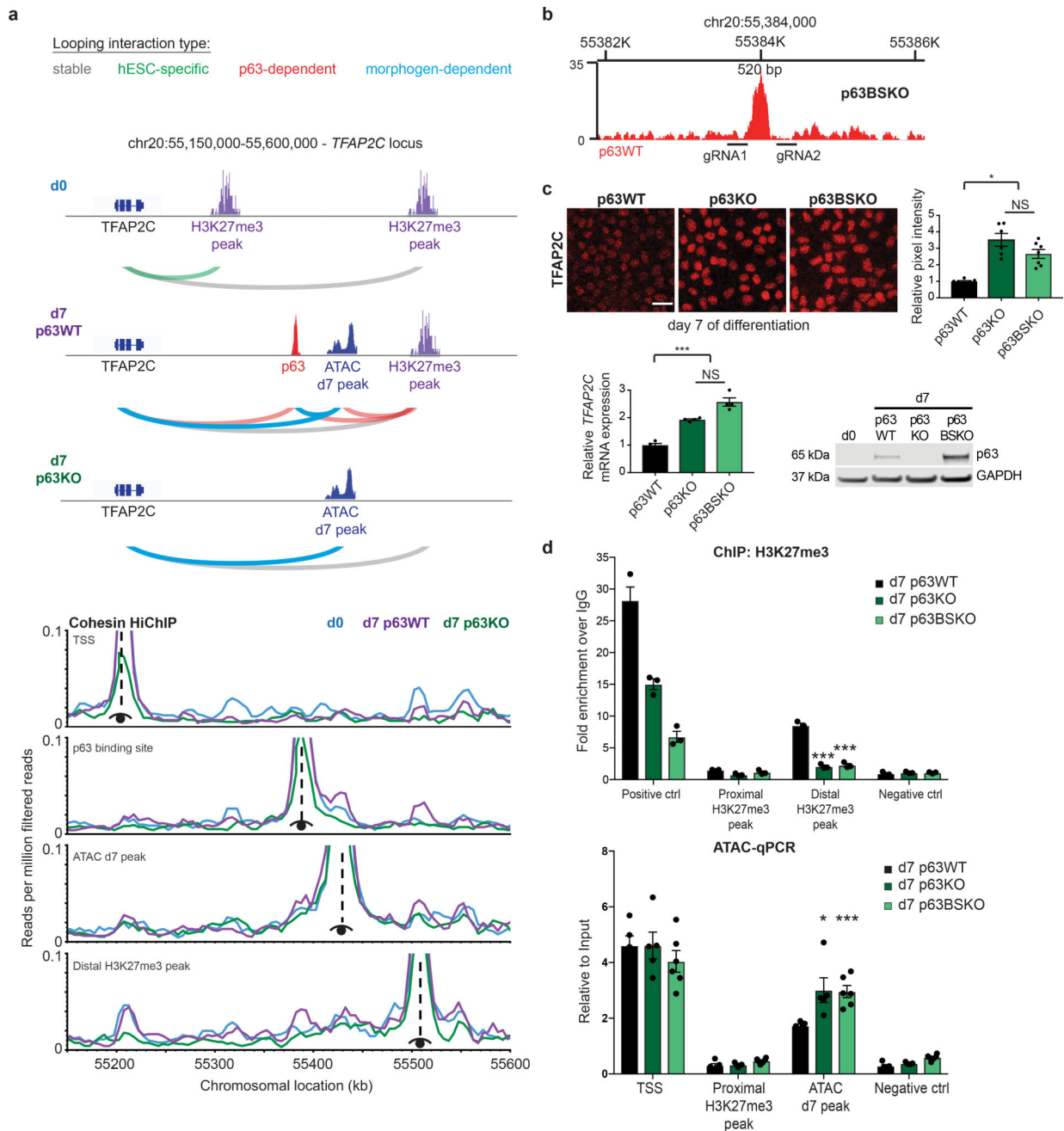


Figure 4. p63 negatively regulates *TFAP2C* expression through morphogen-induced and p63-dependent distal elements and connectivity. (a) Cohesin HiChIP reveals complex looping interactions at the *TFAP2C* locus. Schematic of morphogen and p63-dependent interactions (top panels) with virtual 4C plots of the normalized HiChIP data (bottom panel). In the virtual 4C plots, the dotted line and eye denote the viewpoint from which the interactions are graphed: top plot depicts all interactions from the *TFAP2C* TSS; second plot depicts all interactions from the p63 binding site; third plot depicts all interactions from the ATAC d7 peak; fourth plot depicts all interactions from the distal H3K27me3 peak. Together these

plots represent the full connectivity at the *TFAP2C* locus (n=3). (b) A 520 bp deletion surrounding the p63 binding site (p63BSKO) was generated using CRISPR/Cas9. (c) Deletion of the p63 binding site leads to an increase in *TFAP2C* expression similar to the levels seen in the d7 p63KO cells. Loss of *TFAP2C* expression leads to a dramatic increase in p63 expression. Relative pixel intensity was calculated from 3 independent images (*p-value < 0.05, ***p-value < 0.005, NS-not significant). Error bars represent s.e.m. Scale bar: 20 μ m. (d) ChIP-qPCR for H3K27me3 at the *TFAP2C* locus shows a decrease in the histone mark in the d7 p63BSKO, similar to the d7 p63KO (***p-value < 0.005), n=2. ATAC-qPCR shows an increase in accessibility at the ATAC d7 peak in d7 p63BSKO, again similar to the d7 p63KO cells (*p-value = 0.04, ***p-value = 0.002), n=4. Graphs depict signal relative to input and error bars represent s.e.m.