Pausing of RNA polymerase II (Pol II) close to promoters is a common regulatory step in RNA synthesis, and is coordinated by a ribonucleoprotein complex scaffolded by the noncoding RNA RN7SK. The function of RN7SK-regulated gene transcription in adult tissue homoeostasis is currently unknown. Here, we deplete RN7SK during mouse and human epidermal stem cell differentiation. Unexpectedly, loss of this small nuclear RNA specifically reduces transcription of numerous cell cycle regulators leading to cell cycle exit and differentiation. Mechanistically, we show that RN7SK is required for efficient transcription of highly expressed gene pairs with bidirectional promoters, which in the epidermis co-regulated cell cycle and chromosome organization. The reduction in transcription involves impaired splicing and RNA decay, but occurs in the absence of chromatin remodelling at promoters and putative enhancers. Thus, RN7SK is directly required for efficient Pol II transcription of highly transcribed bidirectional gene pairs, and thereby exerts tissue-specific functions, such as maintaining a cycling cell population in the epidermis.
Regulation of transcription is one of the most important steps in gene expression to ensure coordinated cellular behaviours and fate decisions. Transcription of all protein-coding genes and many noncoding RNAs is carried out by the RNA polymerase II (Pol II) complex. To initiate transcription and maintain elongation, the Pol II complex interacts with a multitude of proteins and protein complexes. Pol II often pauses shortly downstream of transcription initiation sites before beginning productive elongation to ensure proper 5' capping of nascent RNAs, to prevent transcriptional re-initiation by another Pol II enzyme, and to maintain a nucleosome-free promoter. The paused elongation complex is stabilized by the negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF), Release of paused Pol II into productive elongation requires the P-TEFb kinase complex and phosphorylation of serine 2 of the Pol II C-terminal domain (CTD).

The activity of P-TEFb itself is regulated by a ribonucleoprotein complex containing RN7SK, a highly structured, abundant non-coding RNA of 331 nucleotide length. RN7SK is stabilized by methyl phosphate capping enzyme (MePCE) at the 5' end and LARP7 at the 3' end. In the nucleus, RN7SK regulates transcription by sequestering P-TEFb, which is expected to prolong Pol II pausing. Pol II pausing is a common feature of gene regulation during development and embryonic stem cell differentiation, and orchestrates rapid and dynamic changes in transcription, in particular of regulators involved in signal transduction.

Despite the essential roles of Pol II pausing during development and diseases, the underlying molecular roles of RN7SK in regulating the transcriptional processes in adult tissue homeostasis remains largely unexplored. Depletion experiments in mouse embryonic stem cells revealed impaired neuronal differentiation, and identified RN7SK as regulator of bidirectionally transcribed enhancers and transcription termination. However, the direct transcriptional functions of RN7SK during cell differentiation have yet to be identified.

Here, we characterized the functional role of RN7SK in adult tissues using the mammalian epidermis, one of the best-characterized epithelial tissues. Depletion of RN7SK triggered terminal differentiation through transcriptional repression of cell cycle regulators causing cell cycle arrest. The downregulation of genes was independent of chromatin changes in promoters and enhancers. Instead, gene repression occurred at specific highly expressed genes bidirectionally transcribed from exceedingly open, accessible promoters. Thus, our work identifies a functional role of the 7SK snRNP complex in regulating RNA synthesis during cellular differentiation in adult skin.

Results

RN7SK regulates epidermal cellularity. To investigate the function of 7SK in gene transcription, we generated two transgenic mouse lines carrying floxed Rn7sk alleles (Fig. 1a). We either targeted the Rn7sk gene including the TATA-box (Fig. 1a, Line 1) or including the TATA-box and the proximal sequence element (PSE) in the RNA Pol III promoter (Fig. 1a, Line 2). To remove Rn7sk in the interfollicular epidermis (IFE) (Rn7sk cKO), the mice were crossed to a transgenic line carrying an inducible oestrogen receptor domain under the control of the keratin 14 promoter (Krt14:Cre-ERT2), targeting all undifferentiated basal cells in the IFE. To visualize recombined cells, we included a reporter transgene (Rosa26:TdTomato) (Fig. 1b).

Administration of 4-hydroxytamoxifen (4-OHT) for 2 weeks efficiently removed Rn7sk in the epidermis (Fig. 1c, d; Fig. S1a-c, upper panels). After two weeks, we measured a significant loss of cellularity in the Rn7sk cKO epidermis (Fig. 1d, middle and lower panels and Fig. 1e; Fig. S1d). To replace lost epidermal cells, skin often induces wound healing processes. Indeed, we measured increased proliferation and upregulation of the injury marker keratin 6 (K6) after four weeks of 4-OHT treatment (Fig. 1f; Fig. S1d-g). We observed the same phenotype when we deleted Rn7sk during skin morphogenesis starting from postnatal day 4 (P4) (Fig. 1g-i). Unexpectedly, one month after the last 4-OHT application, the skin returned to normal despite the complete absence of Rn7sk in the epidermis (Fig. S1h, i).

To test whether Rn7sk was essential for skin development, we generated Rn7sk knockout mice by crossing animals to a Sox2:Cre transgenic line, in which Cre recombination is expressed in the inner cell mass, leading to total Rn7sk-deletion. Complete absence of Rn7sk was sublethal (Fig. S1j), but the surviving offspring was phenotypically normal, including the skin morphology (Fig. S1k).

We concluded, that acute deletion of Rn7sk in mice reduced epidermal cellularity in the short term, yet this effect was compensated in vivo via a mechanism that resembled a wound healing process (Fig. 1j).

Ablation of RN7SK enhances epidermal cell differentiation. To confirm that the reduction in epidermal cellularity was a direct consequence of RN7SK-deletion, we repressed RN7SK in primary human keratinocytes, a well-characterized in vitro model for IFE cells. To reduce RN7SK expression, we used three different siRNAs targeting RN7SK and one against LARP7 (Fig. S2a-c). LARP7 is essential for stabilizing RN7SK. All siRNAs tested, siRNA5 reduced RN7SK most efficiently and was used throughout this study (Fig. S2b).

To evaluate the cellular effects of reduced RN7SK levels in human epidermal cells. We performed colony-forming assays and skin reconstitution tests on de-epidermized dermis (Fig. 2a-c). The colony-forming efficiency of RN7SK-depleted cells was significantly reduced (Fig. 2a, b), and their ability to reconstitute a multi-stratified epithelium ex vivo was abolished (Fig. 2c). These results are in line with the loss of cellularity observed in our RN7SK cKO mouse epidermis after 2 weeks of 4-OHT treatments (Fig. 1d, e).

Next, we analyzed the consequences of RN7SK-depletion on keratinocyte differentiation. We induced the terminal differentiation programme by calcium induction, the best-studied pro-differentiating stimulus for keratinocytes (Fig. 2d). While expression of the undifferentiation marker ITGA6 decreased two-fold, terminal differentiation markers (INV, TGM1) increased more than five-fold in absence of RN7SK (Fig. 2e, f). RN7SK-depletion also phenotypically enhanced terminal differentiation by inducing the formation of a stratified epithelium in culture (Fig. S2d, e). The increase in differentiation was confirmed using three different RN7SK siRNAs and a siRNA targeting LARP7 (Fig. 2g). We concluded that deletion of RN7SK caused epidermal cell differentiation.

RN7SK maintains robust transcription of highly expressed genes. To understand the underlying molecular mechanisms leading to epithelial differentiation, we first tested how the ablation of RN7SK affected global transcription. ChIP-sequencing experiments confirmed a slight but significantly lower RNA Pol II occupancy at transcription start sites (TSS) after RN7SK knockdown (Fig. 3a). Calculation of the pausing index (number of reads +250 base pairs around the TSS divided by the number of reads across the rest of the gene body) indicated that ablation of RN7SK moved Pol II from initiation into elongation (Fig. 3b). A reduction of Pol II pausing was further confirmed by a two-fold increase in serine 2 phosphorylation levels at its C-terminal domain (CTD) in the absence of RN7SK (Fig. 3c, Ser2), which is
required for productive elongation\textsuperscript{37}. In contrast, transcription initiation requires phosphorylation at serine 5, and we measured no differences in RN7SK-depleted cells (Fig. 3c, Ser5). We concluded that depletion of RN7SK removed Pol II from the transcriptional start sites.

If loss of RN7SK decreased Pol II pausing, RNA synthesis should be enhanced. However, when we profiled newly transcribed metabolically labelled RNA (4SU-seq) (Fig. S3a\textsuperscript{38}), we unexpectedly measured a global decrease of nascent RNAs at TSSs in the absence of RN7SK (Fig. 3d). This decrease of 4SU-labelled RNAs was driven by robust gene-specific repression of highly transcribed genes (Fig. 3e, f). In contrast, the overall nascent transcript levels remained unchanged (Fig. S3b). Counterintuitively, both up- and downregulated genes exhibited a reduction in the pausing index when RN7SK was depleted (Fig. S3c, d).

To confirm that loss of RN7SK inhibited rather than induced transcription, we first identified all common differentially transcribed genes in two independent 4SU RNA-sequencing datasets (Fig. 3g; Fig. S3e–g). Then, we asked whether the differences in new transcription were also found at mature mRNA levels (Fig. 3h). Downregulation of nascent transcripts
correlated with reduced levels of total RNA after 24 and 48 h of RN7SK knockdown (Fig. 3h, lower panel). In contrast, upregulated nascent transcripts only modestly increased on total RNA level (Fig. 3h, upper panel). We concluded that RN7SK was required for efficient transcription of a specific set of highly expressed genes.

7SK-sensitive genes are characterized by bidirectional transcription and open chromatin. In search of a mechanism causing the transcriptional changes in RN7SK-depleted cells, we first inspected the most upregulated newly transcribed RNAs (Fig. S3h, left panel). Upregulated transcripts often contained several alternative start sites (Fig. 3i, AKAP12, upper panel; Fig. S3i), possibly leading to the accumulation of new transcript reads over the gene body.

A noticeable feature of the most repressed transcripts was the enrichment of sequence reads upstream of the TSS, indicating antisense transcription (Fig. 3i, lower panel, CDT1; Fig. S3j). To test whether the downregulated transcripts were commonly bidirectionally transcribed, we quantified all antisense sequence reads. Only downregulated nascent RNAs contained a higher number of antisense transcripts when compared to all genes (Fig. 3j). Using CAGE sequencing data from the FANTOM project, we confirmed that the downregulated genes were twice as correlated with reduced levels of total RNA after 24 and 48 h of RN7SK knockdown (Fig. 3h, lower panel). In contrast, upregulated nascent transcripts only modestly increased on total RNA level (Fig. 3h, upper panel). We concluded that RN7SK was required for efficient transcription of a specific set of highly expressed genes.

7SK-sensitive genes are characterized by bidirectional transcription and open chromatin. In search of a mechanism causing the transcriptional changes in RN7SK-depleted cells, we first inspected the most upregulated newly transcribed RNAs (Fig. S3h, left panel). Upregulated transcripts often contained several alternative start sites (Fig. 3i, AKAP12, upper panel; Fig. S3i), possibly leading to the accumulation of new transcript reads over the gene body.

A noticeable feature of the most repressed transcripts was the enrichment of sequence reads upstream of the TSS, indicating antisense transcription (Fig. 3i, lower panel, CDT1; Fig. S3j). To test whether the downregulated transcripts were commonly bidirectionally transcribed, we quantified all antisense sequence reads. Only downregulated nascent RNAs contained a higher number of antisense transcripts when compared to all genes (Fig. 3j). Using CAGE sequencing data from the FANTOM project, we confirmed that the downregulated genes were twice as correlated with reduced levels of total RNA after 24 and 48 h of RN7SK knockdown (Fig. 3h, lower panel). In contrast, upregulated nascent transcripts only modestly increased on total RNA level (Fig. 3h, upper panel). We concluded that RN7SK was required for efficient transcription of a specific set of highly expressed genes.

7SK-sensitive genes are characterized by bidirectional transcription and open chromatin. In search of a mechanism causing the transcriptional changes in RN7SK-depleted cells, we first inspected the most upregulated newly transcribed RNAs (Fig. S3h, left panel). Upregulated transcripts often contained several alternative start sites (Fig. 3i, AKAP12, upper panel; Fig. S3i), possibly leading to the accumulation of new transcript reads over the gene body.

A noticeable feature of the most repressed transcripts was the enrichment of sequence reads upstream of the TSS, indicating antisense transcription (Fig. 3i, lower panel, CDT1; Fig. S3j). To test whether the downregulated transcripts were commonly bidirectionally transcribed, we quantified all antisense sequence reads. Only downregulated nascent RNAs contained a higher number of antisense transcripts when compared to all genes (Fig. 3j). Using CAGE sequencing data from the FANTOM project, we confirmed that the downregulated genes were twice as correlated with reduced levels of total RNA after 24 and 48 h of RN7SK knockdown (Fig. 3h, lower panel). In contrast, upregulated nascent transcripts only modestly increased on total RNA level (Fig. 3h, upper panel). We concluded that RN7SK was required for efficient transcription of a specific set of highly expressed genes.
well-known gene cluster regulated by bidirectional transcription contains histone genes, where bidirectional promoters are used to maintain stoichiometry. Accordingly, a quarter of all repressed nascent transcripts in the absence of RN7SK were histone genes (Fig. S4a, b). We concluded that RN7SK was required for efficient bidirectional co-expression of highly expressed gene pairs.

RN7SK-mediated gene repression is not caused by changes in chromatin. Since RN7SK regulated epidermal cellularity in mouse, we more closely investigated how transcription of cell cycle genes was affected. We selected consistently downregulated genes (CDK1, CDC25c, CDC45, and MCM10) and confirmed their significant reduction as early as 12 h after RN7SK-depletion (Fig. S4c). Upregulated genes remained largely unaffected within
24 h of RN7SK knock-down (Fig. S4d). The repression of cell cycle regulators in the absence of RN7SK was confirmed in three other independent human keratinocyte lines (Fig. S4e), and we obtained similar results when using two different RN7SK siRNAs (205 and 207) or a siRNA targeting LARP7 (Fig. S4f). Because we found a slight reduction of global Pol II occupancy at transcriptional start sites, we confirmed that GAPDH, used for normalization, was not significantly repressed in any condition (Fig. S4g). Furthermore, co-expression of a wild-type or mutated version of human RN7SK (not targeted by siRNA 5) in RN7SK knock-down cells, prohibited efficient downregulation of cell cycle regulators (Fig. S4h, i). Since not all genes rely on P-TEFb activity for transcription52, we also confirmed that inhibition of P-TEFb with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) caused downregulation of the four cell cycle regulators (Fig. S4j). Thus, our data revealed that removal of RN7SK in primary human epidermal cells directly repressed cell cycle genes.

We next asked whether loss of RN7SK correlated with the formation of repressive chromatin at transcriptional start sites of these cell cycle regulator genes.63. However, when we depleted RN7SK, nucleosome positioning around the TSS remained unaltered (Fig. 5a–d). Our results excluded chromatin remodelling at the TSS, yet high levels of RN7SK occupancy have also been reported at active enhancers, where it limits enhancer-RNA transcription31. However, ChIP-sequencing experiments for two histone modifications commonly found at putative enhancers and promoters (H3K4Me1, H3K27Ac)54, revealed no differences in occupancies (Fig. 5e–j; Fig. S5a, b). Thus, 7SK-driven gene expression appeared to be independent of chromatin remodelling at promoters and putative enhancers.

**RN7SK orchestrates mRNA synthesis and splicing.** Next, we investigated whether mis-regulation of bidirectional transcription was sufficient to explain the epidermal phenotype caused by RN7SK-depletion. We asked how the loss of RN7SK affected mRNA processing and degradation for two reasons. First, downregulated genes were more likely to have another down-stream sense-strand gene in close proximity (<1 kb) (Fig. 4c). Second, 7SK has been described to prevent transcription down-stream of polyadenylation sites30. To determine whether RNA synthesis, processing, or degradation was the most prevalent RNA metabolic pathways affected by depletion of RN7SK, we used INSPEct, a tool that integrates intronic and exonic signals from nascent and total RNA-seq data to derive the rates of pre-mRNA synthesis, processing, and mature mRNA degradation35,36. In both independent RNA sequencing datasets, RNA synthesis was most affected by depletion of RN7SK (Fig. 6a–d; Fig. S6a–d). Since mRNA syntheses take place in the nucleus, we confirmed the downregulation of genes in the nucleus as early as 18 h after RN7SK depletion (Fig. 6e).

Since some gene expression changes were predicted to be caused by RNA processing and degradation, we tested how RN7SK-depletion affected splicing. Indeed, we found increased intron retention levels in downregulated genes in both RNA and 4SU sequencing datasets (Fig. 6f, g). Using rMATS to identify alternative splicing events genome-wide37, we revealed exon skipping as the most prevalent splicing differences (Fig. 6h; Fig. S6e). The splicing differences were consistent in both independent RNA sequencing datasets, but the overlapping genes were not enriched in cell cycle regulation, DNA repair, or chromosome organization (Fig. 6i–k; Fig. S6f–h). Thus, RNA synthesis and splicing were likely to be regulated by independent mechanisms involving 7SK.

**Terminal differentiation is the consequence of cell cycle arrest in the absence of RN7SK.** To understand the molecular mechanisms causing terminal differentiation in the absence of RN7SK, we transcriptionally profiled undifferentiated and calcium-differentiated epidermal cells (Fig. S7a). Several lines of evidence indicated that loss of RN7SK- and calcium-induced differentiation were two distinct regulatory processes. First, the number of differentially expressed genes in response to deletion of RN7SK was three-fold larger than transcriptional changes induced by the calcium-switch alone (Fig. S7b). Second, the differential gene expression profile of RN7SK knock-down cells in calcium-low and -high conditions overlapped by more than 70% (Fig. S7b). Third, while both calcium-induced differentiation and RN7SK-depletion increased expression of genes involved in epidermis differentiation (Fig. 7a; Fig. S7c), only RN7SK depletion repressed cell cycle genes (Fig. 7b; Fig. S7d).

Since cell-cycle withdrawal is an early hallmark of skin differentiation that occurs already in the undifferentiated basal epidermal compartment58, we asked whether cell cycle arrest explained the induction of terminal differentiation in the absence of RN7SK. Indeed, RN7SK-depleted cells accumulated in G2/M phase of the cell cycle 48 h after siRNA transfection (Fig. 7c, d; Fig. S7e). Cell cycle arrest was confirmed in four independent epidermal lines, yet one line arrested in G1 (Fig. 7d; Fig. S7f). As expected, the cell cycle was also affected by calcium-induced differentiation (Fig. S7g). We further confirmed a reduction of
Fig. 4 7SK orchestrates bi-directional transcription of highly expressed gene pairs. a Illustration of the analyses shown in (b, c, b, c Distance (upper panels) and frequency (lower panels) of genes with upstream (b) or downstream (c) antisense (as) (left panels) or sense (ss) (right panels) genes within 1000 bases (1 kb) distance (grey dotted box). Up- (orange) and down- (blue) regulated genes were defined based on two independent 4SU and total RNA-seq datasets. All (grey) = all protein-coding genes; Up = genes with log2FC > 0.3 and padj < 0.05; Down = genes with log2FC < −0.3 and padj < 0.05 (Wald test with FDR correction). The Up and Down groups were compared to all protein-coding genes. Violin plots show the median and interquartile range and width indicating frequency (b, c upper panels). P-value calculation (see "Methods"). b Correlation of RNA levels of downregulated genes with their upstream as genes shown as average log2 fold-changes (FC) in two independent 4SU (left panel) and total RNA (right panel) sequencing datasets. P-value tests slope deviation from 0. e-g UCSC genome browser shots of bi-directional repressed DTL and its upstream antisense gene partner INTS7 (e). RNA levels (reads) of new transcripts of downregulated genes (padj < 0.05; FC < −0.5; n = 485) with bi-directional as gene less than 1 kb away and more than an average of 5 reads (n = 36 genes) (f) or less than 5 reads (n = 33 genes) (g) in upstream antisense (dark blue) or downstream sense (light blue) directions in control (ctr) siRNA transduced cells. h Reactome of all bi-directionally orientated gene pairs less than 1 kb apart. I Gene ontology analyses using all common significantly (padj < 0.5; Wald test with FDR correction) downregulated genes in the 4SU RNA-seq datasets. Background: all expressed genes. Source data are provided as a Source Data file.
cell cycle genes in RN7SK cKO mouse epidermis (Fig. 7e). We concluded that downregulation of cell cycle genes leading to cell cycle arrest induced terminal differentiation of RN7SK-lacking epidermal cells.

To prove that the cell cycle changes were directly caused by loss of RN7SK, we re-expressed a wild-type or mutated RN7SK construct in a RN7SK-depleted human epithelial cancer line (FaDu) (Fig. 7f). As expected, cell cycle regulators were downregulated in the absence of RN7SK (Fig. 7h). Yet, re-expression of wild-type and mutated RN7SK in these depleted cells prohibited the accumulation of cells in the G2/M phase of the cell cycle (Fig. 7g). We concluded that epidermal cells required RN7SK to coordinate the expression of cell cycle genes, and disruption of this co-regulation stimulated terminal differentiation processes through cell cycle exit.

In summary, our data demonstrated that the 7SK snRNP complex orchestrates efficient transcription of highly expressed bidirectionally transcribed gene pairs potentially by tethering Pol II to the transcriptional start sites. Loss of RN7SK in the epidermis specifically represses cell cycle genes causing cell cycle arrest and thereby stimulates differentiation, a process that was reversible in vivo.

**Discussion**

Here, we investigated the transcriptional roles of Rn7sk in adult tissues, using the epidermis as a model system. We show that human RN7SK sustained Pol II activity at highly expressed bidirectionally transcribed gene pairs. Although the 7SK ribonucleoprotein complex regulates P-TEFb activity, we find no evidence that RN7SK-depletion enhances RNA synthesis due to an increased release of Pol II. However, previous studies identified an inhibitory effect of 7SK on P-TEFb in response to stress such as ultraviolet radiation, while uninduced cells showed little evidence that RN7SK-RN7SK had no inhibitory effect of 7SK on P-TEFb in response to stress such as ultraviolet radiation, while uninduced cells showed little changes in global transcription upon RN7SK-depletion. Thus, the release of P-TEFb from 7SK upon stress might rather reflect the transcriptional reprogramming upon the stress signal than a general inhibitory role of the complex on transcriptional elongation.

Our finding that RN7SK was required to maintain robust bidirectional transcription of highly expressed gene pairs implies a structural role for 7SK at promoters with complex and high turnover of Pol II. An unexpectedly high RNA Pol II turnover has also been reported at paused promoters. We propose that the 7SK ribonucleoprotein complex tethers P-TEFb and other transcriptional regulators to highly transcribed bidirectional promoters to regulate Pol II activity. Our data confirm that bidirectionally transcribed genes often regulate DNA repair, cell cycle, and RNA metabolism, a highly efficient way of coordinating the expression of genes acting in the same cellular response pathway. For instance, out of 120 examined human DNA repair genes, 42% are amongst the most sensitive genes to RN7SK-depletion. Histone genes expressing bidirectionally transcribed gene pairs are arranged in a bidirectionally divergent configuration with transcription start sites less than 1 kb apart.
are bidirectionally transcribed to maintain stoichiometry. As a consequence, depletion of RN7SK induced epidermal cells to exit the cell cycle and undergo differentiation. However, in vivo, this loss of cellularity was later compensated by a wound-like response. The different phenotype in response to RN7SK-depletion in vitro versus in vivo can be explained by our finding that RN7SK affects transcription gene-specifically, thereby regulating cell context-specific functions. For instance, in response to acute deletion of Rn7sk in mouse skin, cycling epidermal populations will trigger the terminal differentiation programme due to a synchronized cell cycle exit. However, the cell cycle regulators are still expressed, albeit with lower levels. The overall reduction of cell cycling might lead to a slightly lower, yet sustainable, epidermal turn-over. In contrast, deletion of Rn7sk in mouse embryonic stem cells specifically repressed a different...
cohort of transcriptionally poised genes with bivalent or activating chromatin marks\(^5\). Embryonic stem cells display unique cell-cycle features with a prolonged S-phase but truncated G1 and G2 phases\(^6\).

Although chromatin changes are known to influence Pol II pausing\(^6,2,3\), we detected no epigenetic changes at promoters or enhancers in the absence of RN7SK. However, we cannot exclude a faster transcription elongation rate due to the absence of tightly controlled Pol II activity, which might cause impaired splicing leading to RNA decay. Increased RNA Pol II elongation rates can affect co-transcriptional splicing and splicing efficiency, which then compromises splicing fidelity\(^6\).

In summary, our work demonstrates that the precise co-ordination of highly expressed bidirectional gene pairs required the 7SK ribonucleoprotein complex for epidermal homeostasis.

**Methods**

**Mice.** All mice were housed in the Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute Animal Unit. All mouse husbandry and experiments were carried out in compliance with the Animals (Scientific Procedures) Act 1986 following ethical review and approval by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB) under the terms and conditions of the UK Home Office licences PPL80/2619 and PPL36833804.

**Generation of the Rn7sk cKO transgenic lines.** All transgenic lines were bred on a mixed background (F1 of B6SJL x CBA). To generate Rn7sk conditional knockout mice (cKO), we produced targeting vectors by BAC recombineering\(^6\). To generate the homology arms BACS bMQ337m02 and bMQ215m24 (Source Bioscience) were used. We generated two targeting vectors, which differ for the position of the loxp site located at the 5′ end of the Rn7sk gene locus, whereas the 3′ end loxp site was in the same position in both constructs. In line 1, the 5′ loxp site was located between the TATA Box and the proximal sequence element (TATA and PSE, respectively) of the Rn7sk promoter (RNA Pol II promoter). In line 2, the 5′ loxp site was instead positioned between the PSE and the distal sequence element (DSE). Both lines showed the same phenotype. If not otherwise stated, all displayed data were obtained from line 1.

E14 mouse embryonic stem cells were targeted by homologous recombination with the Rn7sk cKO targeting vector. After 10 days in the selection medium containing G418, around 100 clones per construct were picked and screened by PCR for integration of both 5′ and 3′ arms. Three double-positive clones per construct were injected into blastocysts derived from C57BL6 females. Chimeric offspring was mated with C57BL6 mice and agouti F1s were genotyped for the presence of the targeted Rn7sk allele. One line per construct was further mated with
FLP mice to induce excision of the selection cassette to generate the Rn7sk
+/- mouse. Mouse genotyping was performed by standard PCR (Supplementary
Table 1).

Generation of Rn7sk-knockout mice was achieved by mating Rn7sk+/lox with
Sox2-Cre mice33. To achieve interfollicular epidermis (IFE) specific deletion of
Rn7sk, Rn7sk+/lox mice were further crossed to Tg(KRT14-cre/ERT2)20Efu/J mice
(Krt14Cre/ERT2)30, using a retroviral supernatant of Cre expressing Rn7sk
lox/lox;ERT2Tat (Tato)31. To conditionally delete the Rn7sk gene in the IFE, the back
skin of mice was treated topically every second day for the indicated time with
200 μl of a 4-hydroxy-tamoxifen (Sigma Aldrich) solution of 14.28 mg/ml in
acetone.

Tissue processing and staining. The mouse epidermis was fixed overnight in 4%
paraformaldehyde, dehydrated in ethanol gradient followed by xylene, paraffin-
embedded, and then cut using a microtome into 5 μm thick sections. For
staining, sections were deparaffinised in xylene and rehydrated in an ethanol
gradient. For immunostaining, antigen retrieval was performed by boiling the
sections for 20 min in sodium citrate solution, followed by blocking in 10% FBS in
PBS. Primary antibodies (Supplementary Table 2) were diluted in blocking
solution and incubated overnight at 4 °C. After washing three times in PBS, sec-
tions were incubated with the appropriate secondary antibody (Alexa fluoro-con-
jugated; Thermofisher Scientific) diluted 1:500 in PBS tween and DAPI when
indicated. After washing slides were then mounted in a 1:1 solution of PBS and
Glycerol. For RNA in situ hybridization of Rn7sk, an RNAse probe (ACD bio)
was used following manufacturer instructions. Images were acquired with an
Axioplan2 microscope (Zeiss).

Quantification of epidermal cellularity. Epidermal cellularity was quantified on
images of large areas of HE stained sections using ImageJ (FIJI). Epidermal cell-
ularity was considered normal when a continuous single layer of nuclei was fol-
lowing each other; reduced if gaps were present between nuclei, increased if more
than one layer of nuclei was present. Data are represented as the percentage of the
skin surface that fell into each category.

Quantification of epidermal proliferation. To quantify IFE proliferation, sections of
mouse skin were stained with Ki67 antibody as described above. Images were
processed with Velocity (Perkin Elmer). The entire IFE or only the basal layer was
manually selected for analysis. The software was asked to identify nuclei on the
DAPI channel and then to identify how many of those had also a positive signal in
the Ki67 channel to identify the percentage of Ki67+ cells.

Cell culture, transfaction, and infection. Unless otherwise stated, neonatal pri-
mary human keratinocytes (Sciencell; #2100) were cultured in KGM gold (Lonza)
or EpiLife added with human keratinocytes growth supplements (Thermo Fisher)
in presence of 0.06 mM CaCl2 and without the addition of antibiotics, in a
humidiﬁed incubator at 37 °C with 7% CO2. To induce differentiation, the calcium
concentration was raised to 1.2 mM. siRNAs were transfected using Lipofectamine
RNAi Max (Invitrogen), following manufacturer instructions using 1 μl of Lipof-
fectamine for each 0.6 μl of 10 μM siRNA. siRNAs were added to cells at a final
concentration of 10 nM (Supplementary Table 3). For P-TEFb inhibition, cells were
transfected with either control or Rn7sk siRNAs and treated with 50 μM 5,6-
Dichloro-2'-Deoxyuridine (DDU) (Sigma) for the last 6 h.

FaDu (HTB-43; ATCC) cells were grown in Eagles minimum essential medium
(ATCC) in the presence of Penicillin Streptomycin and 10% FBS, in a human
embryonic fibroblast (HFF) fed, 5% CO2 incubator at 37 °C with 5% CO2. siRNAs
were added to cells at a nal concentration of 20 nM. After washings, cells were
seeded into 10 cm dishes or 15 cm dishes per 30 μg of RNA using Superscript III
(Thermo Fisher Scientific) and following manufacturer instructions. After
incubation overnight at 65 °C, DNA was then puriﬁed and used as a control DNA.

RNA extraction and RT-qPCR. RNA was extracted using TRIZOL reagent (Thermo
Fisher Scientiﬁc) following manufacturer instructions and quantiﬁed using a
nanodrop. cDNA was synthesized from 1 μg of RNA using Superscript II
(Thermo Fisher Scientiﬁc) and following manufacturer instructions. RT-qPCR
were performed using either Fast SYBR green master mix or Taqman fast universal
master mix (Thermo Fisher Scientiﬁc). Taqman probes were purchased from
Thermo Fisher Scientiﬁc (Supplementary Table 4). RT-qPCR results of target
genes were normalized to GAPDH or 18S rRNA (Mm03928990_g1).

For nuclear and cytoplasmic RNA extraction, human primary keratinocytes
were collected by trypsin, washed in PBS, centrifuged and the pellet was
processed with the NE-PER kit (Thermo Fisher Scientiﬁc) following manufacturer
instructions. Four replicates of each sample were processed, and the experiment
was performed twice.

For quantification of Rn7sk levels in mouse IFE, back skin was isolated from
Rn7sk cKO mice and was quickly snap frozen after fat was removed. Frozen skin
was then homogenized in TRIzol reagent (Thermo Fisher Scientiﬁc) and RNA was
then extracted following manufacturer instructions.

Chromatin immunoprecipitation. Epidermal cells were grown to 60–70% con-
fluence and transfected with control or Rn7sk siRNA 5 × 10^6 for each 10 cm dish
per siRNA). 18 h after transfection cells were ﬁxed with 1% formaldehyde for 10 min
at room temperature. Cross-linking was terminated with 2.5 M glycine for 5 min. Cell
pellets were recovered by centrifugation at 1350 × g for 5 min at 4 °C. Cell pellets
were recovered and chromatin was isolated and sonicated for 17 cycles of 30 s
with an output of 30 W, using an automated sonicator (300W; Misonix)18. Immuno-
precipitation was carried out overnight at 4 °C using 5 μg of H3K4me1 (Abcam) or
10 μg of H3K27Ac (Abcam) or RNA Pol II N20 (Santa Cruz) antibodies previously
bound to Dynabeads protein G (Thermo Fisher Scientiﬁc). DNA-protein complex-
es were eluted in 200 μl of elution buffer (50 mM Tris pH 8, 1 mM EDTA, 1%
SDS) by heating at 65 °C with brief agitation every 2–3 min. Cross-links were
reverted both in the immunoprecipitated samples and the whole-cell extract,
separated from the sonicated material before immunoprecipitation, by incubating
overnight at 65 °C. DNA was then puriﬁed by phenol:chloroform extraction.

Libraries were prepared using NEBNext Rapid DNA-Seq Kit (Biso Scientiﬁc). Four
replicates per sample were sequenced on an Illumina HiSeq 2500 platform. Each
ChIP-seq experiment was performed once with four technical replicates per sample.

Cell cycle proﬁling. Human primary keratinocytes and FaDu cells were collected
with trypsin, ﬁxed in ice-cold 70% ethanol overnight. Next day they were cen-
trifuged, resuspended in PBS with DAPI, and analyzed with a LSRFortessa cell
analyzer (BD biosciences). Analysis of FACS data was performed using FlowJo
software (FlowJo LLC). G1, G0, S, and G2/M peaks were manually delimited on the
405 nm histogram. Statistical analysis was performed in Prism (GraphPad
Softwares inc.).
**ARTICLE**

[Link to the article or full text content]

---

**Metabolic RNA labelling and isolation.** 4SU labelling and RNA isolation of newly transcribed was performed as described in ref. 6. Briefly, for new RNA labelling with 4SU, HeLa cells from human keratinocytes were cultured and transfected as described above. At the indicated time point 4SU was added to the culture media at a concentration of 583 μM for 10 min. After the indicated incubation time, cells were quickly harvested by adding TRIzol directly to the plates, and RNA was extracted following the manufacturer protocol, followed by DNase treatment and phenol-chloroform extraction of 4SU. RNA (80 μg of RNA per sample) was isolated per replicate. The incorporated 4SU was biohybridlyated by incubating 1.5 h at room temperature with EZ-link biothin-HDP (Pierce) in biohylination buffer (10 mM Tris Ph 7.4, 1 mM EDTA). Biohylinated RNA was then purified by phenol-chloroform extraction and loaded onto micro-MACS streptavidin columns (Milenyi biotech). After three washes with warm (65°C) and three with room temperature washing buffer (100 mM Tris pH 7.4, 10 mM EDTA, 1 M NaCl. 0.1% Tween 20), 4SU RNA was then eluted directly into RLT buffer (Qiagen) by adding twice 100 μl of 100 mM DTT (Sigma Aldrich) and processed using the kit RNeasy MiniElute (Qiagen), following manufacturer instructions.

**mRNA protection assay.** Cells were collected with trypsin, quantified, and centrifuged. The pellets were then lysed in 1 ml of NP-40 lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.5% NP-40, 0.15 mM spermine, 0.5 mM spermidine) per 10⁶ cells on ice for 5 min. After centrifugation, the pellets were then resuspended in mRNA digestion buffer (10 mM Tris-HCl pH 7.4, 15 mM NaCl, 2.5 mM spermine, 0.5 mM spermidine, 0.01 M EDTA, 0.2 μg/ml RNase K digestion). Monomers were then extracted using the Gel extraction kit (Qiagen) and DNA was diluted to 5 μg/ml prior to quantitative PCR. The signal of the PCR was normalized to total sonicated DNA (Input) prepared as described in the ChIP section. Four replicates of each sample were processed, and the experiment was performed once.

**Statistical analysis.** Quantified data are expressed as mean ± SD, unless otherwise stated in the figure legends. Statistical significance between samples was assessed using unpaired two-tailed Student's t-tests with Welch's correction unless otherwise stated in the figure legends. For cell cycle profiles two-way ANOVA with multiple comparisons was used. Quantitative data were analyzed using Excel and/or Prism software, with the exception for sequencing data.

**RNA library generation and sequencing.** For RNA sequencing, RNA was extracted from cells using TRIZOL (Thermo Fisher Scientific), following manufacturer instructions. After isolation, RNA was treated with turbo DNase (Thermo Fisher Scientific) for 30 min at 37°C. DNase was subsequently removed by phenol-chloroform extraction. For total RNA sequencing at 48 h ribosomal RNA was depleted using the Ribo-zero kit (Cambio), following manufacturer instructions. Ribodepleted RNA was then processed with NEXTFLEX directional RNA-seq kit (dUTP-Based R2) (Bio Scientific), following manufacturer instructions. Four replicates per sample were sequenced on an Illumina HiSeq 2500 platform. 4SU RNA sequencing and total RNA sequencing at 24 h libraries were prepared using the SMARTer Stranded Total RNA-Seq Kit, Pico Input Mammalian kit (Takara Bio Inc). Libraries were prepared using NEXTFLEX TriSeq 1000 (Takara Bio Inc). Sequence data were then sequenced on an Illumina HiSeq 4000 platform. Each experiment was performed once with four technical replicates per sample.

**RNA-seq analysis.** Paired-end RNA-seq reads (for 4SU and total RNA) were quality-trimmed using Trim Galore, and mapped to the human reference genome (GRCh37/hg19) using TopHat2 with the parameters "-m 1 -2" to generate unique sequence alignments. Potential PCR duplicates were removed with MACS2 "filterdup". Narrow peaks for H3K27ac versus WCE control were called by using MACS2 'callpeak' with parameters "-B--call-summits -q 0.05"; broad peaks were called for H3K4me1 versus WCE control with parameters "-b--broad-cutoff 0.05". The overlapping peaks for the replicates were merged using bedops (option -m).

**Bioinformatics analysis.** Genome browser shots and metagene profiles were prepared using DeepTools. For this, the aligned reads were converted to bigWig using bamCoverage. The Pol II ChIP-seq data was processed with the following parameters: "-b-size 1 --ignoreForNormalization chrM --extendReads 150 --centerReads --normalizeUsingPQTP. Read coverage per scaled gene was calculated using computeMatrix with the parameters "scale-region --binSize 10 -b 1000 -a 500 --regionBodyLength 500 --unscaledSPM 1000 --unscaledSPM 500". The regions were defined by the start and end coordinates of all protein-coding genes (Genencode v19). To remove differences caused by variation in ChIP efficiency, we normalized each library using a set of 199 high-expressed genes that were unaunced in RN75K knock-down in total RNA-seq (details in the RNA-seq analysis section). Specifically, we first subtracted the mean baseline signal across the gene body of those genes (excluding the regions closest to the TSS and TTS) and then divided the resulting signal by the mean across the TSS ± 1 kb. The 4SU-seq data was processed using "-b-size 1 --ignoreForNormalization chrM --normal-izeUsingRPKM" for all reads or for the forward and reverse strand separately. Read coverage across all TSSs was calculated using computeMatrix with the parameters "reference-point --binSize 1 b 500 -a 500" and a list of TSS coordinates. For this, all exons in protein-coding genes (n = 1,071,216) were annotated as first, internal and/or last. A set of unique exons that were exclusively first (n = 126,731) were identified and used.

To calculate Pol II travelling ratios (TR), aligned and duplicate-filtered PolII ChIP-seq reads were extended by 150 nt. The PolII travelling ratio (TR) was calculated as described22, as the fraction of the PolII ChIP-seq read counts in the +1 promoter-proximal region to the −1 promoter-proximal region. To calculate the transcriptional start site versus the PolII ChIP-seq read counts in the (2) transcribed region from +250 bp to the annotated transcriptional end site. To generate enhancer heatmaps, cell-type-specific enhancers were defined as H3K4me1 peaks, which did not overlap a promoter region (from −900 to +100 bp of the human transcript start site of ENCODE). Enhancer transcriptional profiles and heatmaps were calculated around the centre of the H3K4me1 peaks ±/+3 kb using deepTools.

To model RNA synthesis, processing and degradation we used the INSPEcT tool. Exon and intron counts were derived for each nascent and total RNA-seq reads using the quantifyExpressionsFromBAM module in strand-specific mode. Exon and intron coverage divided by mean exon coverage, where mean coverage is the intron/exon ratios were calculated based on Gencode v19 annotations. The intron/exon counts were replaced by 1 or 0.001, respectively.

**ChIP-seq analysis.** Single-end ChIP-seq reads for PolII, H3K27ac, H3K4me1, ChIP, and WCE (whole-cell extract) were quality-trimmed using Trim Galore, and reads were mapped to the human reference genome (GRCh37/hg19) using bowtie with the parameters "-p -s 1 -0 -minOverlap 10 -B -C". Read counts were normalized, and the statistical significance of differential expression between RN75K and WT was assessed using the R Bioconductor DESeq2 package. To also identify transcriptional initiation across the genome: http://fantom.gsc.riken.jp/5/sstar/FF:11349-117G8, http://fantom.gsc.riken.jp/5/sstar/FF:11421-118F8, http://fantom.gsc.riken.jp/5/sstar/FF:11272-116H3. 4SU RNA seq up- and downregulated genes (FC 0.3) in RN75K-depleted cells were screened for CAGE peaks on the opposite strand within 100 nucleotides from the annotated TSS.

The intron/exon ratios were calculated based on Gencode v19 annotations. Introns were defined for each gene by subtracting exonic regions from the full transcript and adding the resulting regions as introns to the gff file. Raw intron and exon counts were normal to log2-expressed (log10) with mean ± deviation above 200 were included. Genes with extreme intron/exon ratios resulting from either no exons counts or (nearly) no intron counts were replaced by 1 or 0.001, respectively.

---

**NATURE COMMUNICATIONS** | https://doi.org/10.1038/s41467-021-26083-4 | www.nature.com/naturecommunications
Differential splicing usage was computed with rMATS (v4.1.0) using “--variable-read-length --novelSS --allow-clipping”. Splicing events with FDR < 0.05 and absolute difference inclusion level > 0.1 were reported.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data supporting the findings of this study are available from the corresponding authors upon reasonable request. The next-generation sequencing (NGS) data generated in this study have been deposited in the GEO database under accession code GSE101217.

ATAC-seq data from human keratinocytes are available through GEO (GSE67382), sample GSM1645708 and GSM1645709. CAGE-seq data from human keratinocytes are available through the FANTOM5 Table Extraction Toolkit (sample FF11349-117Gr, FF:11421-118F8, and FF:11272-1161H3; http://fantom.gsc.riken.jp/5/tet). The underlying raw data generated in this study and the subsequent statistical analyses for each experiment are provided in the Supplementary Information/Source Data file. Source data are provided with this paper.

Code availability
Custom codes used in the paper are available on GitHub (https://github.com/susbo/Bandiera-et-al-2021-scripts). Trim Galore! https://www.bioinformatics.babraham.ac.uk/projects/trim_galore
TopHat2 https://ccb.jhu.edu/software/tophat
DESeq2 https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Bowtie http://bowtie.bio.sourceforge.net
MACS2 https://github.com/tauloi/MACS
rMATS https://github.com/Xinglab/rmats-turbo
bedops https://bedops.readthedocs.io
rMATs https://github.com/Xinglab/rmats-turbo
MACS2 https://github.com/tauloi/MACS
TopHat2 https://ccb.jhu.edu/software/tophat

Received: 27 August 2019; Accepted: 10 September 2021; Published online: 07 October 2021

References
1. Schier, A. C. & Taatjes, D. J. Structure and mechanism of the RNA polymerase II transcription machinery. Genes Dev. 34, 465–488 (2020).
2. Rasmussen, E. B. & Lis, J. T. In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proc. Natl. Acad. Sci. USA. 90, 7923–7927 (1993).
3. Tome, J. M., Tippens, N. D. & Lis, J. T. Single-molecule nascent RNA transcription initiation. J. Mol. Biol. 607, e29736 (2017).
4. Vos, S. M. et al. Structure of activated transcription complex Pol II-DSIF-NELF. Nature 560, 601–606 (2018).
5. Shao, W. Q. & Zeitlinger, J. Paused RNA polymerase II inhibits new transcriptional initiation. Nat. Genet. 49, 1045 (2017).
6. Gilchrist, D. A. et al. Pausing of RNA polymerase II disrupts DNA-nicked nucleosome organization to enable precise gene regulation. Cell 143, 540–551 (2010).
7. Core, L. et al. Paused RNA polymerase II as a developmental checkpoint. Nat. Biotechnol. 322, 2229 (2008).
8. Gilchrist, D. A. et al. Differential splicing usage was computed with rMATS (v4.1.0) using “--variable-read-length --novelSS --allow-clipping”. Splicing events with FDR < 0.05 and absolute difference inclusion level > 0.1 were reported.

ARTICLE
45. Sugimoto, M., Oohashi, T. & Ninomiya, Y. The genes Col4a5 and Col4a6, coding for basement-membrane collagen chains alpha-5(IV) and alpha-6(IV), are located head-to-head in close proximity on human-chromosome Xq22 and Col4a6 is transcribed from 2 alternative promoters. Proc. Natl. Acad. Sci. USA. 91, 11679–11683 (1994).

46. Momota, R. et al. Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and alpha4(IV) collagen chains are arranged head-to-head on chromosomes: 2q36, FEBS Lett. 424:11–16 (2018).

47. Guarguaglii, G. et al. Expression of the murine RanBP1 and Htf9-c genes is regulated from a shared bidirectional promoter during cell cycle progression. Biochem. J. 325, 277–286 (1997).

48. Adachi, N. & Lieber, M. R. Bidirectional gene organization: a common architectural feature of the human genome. Cell 109, 807–809 (2002).

49. Albig, W., Kioschis, P., Poustka, A., Meergans, K. & Doenecke, D. Human histone gene organization: nonregular arrangement within a large cluster. Genomics 40, 314–322 (1997).

50. Ahn, J. & Gruen, J. R. The genomic organization of the histone clusters on human 6p21.3. Mamm. Genome 10, 768–770 (1999).

51. Maxson, R., Cohn, R., Kedes, L. & Mohun, T. Expression and organization of histone genes. Annu. Rev. Genet. 17, 239–257 (1983).

52. Gomes, N. P. et al. Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. Genes Dev. 20, 601–612 (2006).

53. Gilchrist, D. A. et al. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximo nuclear assembly. Genes Dev. 22, 1921–1933 (2008).

54. Heintzman, N. D. et al. Histone modifications at open cell-type-specific gene expression. Nature 459, 108–112 (2009).

55. de Pretis, S. et al. INSPEcT: a computational tool to infer mRNA synthesis, processing and degradation dynamics from RNA- and 4sU-seq time course experiments. Bioinformatics 31, 2829–2835 (2015).

56. de Pretis, S. et al. Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. Genome Res. 27, 1658–1664 (2017).

57. Shen, S. et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. USA. 111, E5593–E5601 (2014).

58. Albers, K. M., Greif, F., Setzer, R. W. & Taichman, L. B. Cell-cycle withdrawal in cultured keratinocytes. Differentiation 34, 236–240 (1987).

59. St红豆arek, C. et al. The 7SK/P-TEFb snRNP controls ultraviolet radiation-induced transcriptional reprogramming. Cell Rep. 35, 108965 (2021).

60. Krebs, A. R. et al. Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. Mol. Cell 67, 411 (2017).

61. Ballabeni, A. et al. Cell cycle adaptations of embryonic stem cells. Genes Dev. 24, 133–140 (2010).

62. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).

63. Nascimento, E. M. et al. The opposing transcriptional functions of Sin3a and c-Myc are required to maintain tissue homeostasis. Nat. Cell Biol. 13, 1395–1405 (2011).

64. Radle, B. et al. Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. J. Vis. Exp. 78, 50195 (2013).

Acknowledgements
We thank everybody who provided us with reagents. In particular, we thank Arndt G. Benecke and Sebastian Elbrecht for providing the RN7SK plasmid. We gratefully acknowledge the support of all the Wellcome MRC Cambridge Stem Cell Institute core facility managers, in particular Maike Paramor, Bill Mansfield, Gemini Chu, Peter Humphreys, and Andy Riddell. We also thank John Brown for his help obtaining tissue samples. This work was funded by Cancer Research UK (CR-UK), Worldwide Cancer Research, the Medical Research Council (MRC), the European Research Council (ERC), and the Helmholtz Association. For this research, Michaela Frye’s laboratory was also supported by a core support grant from the Wellcome Trust and MRC to the Wellcome Trust-Medical Research Cambridge Stem Cell Institute. Some parts of figure panels are created with BioRender.com

Author contributions
R.B., M.F., and S.B. designed experiments, analyzed data, and wrote the paper. R.W. performed experiments and analyzed data. T.B.-B. and S.D. performed computational analyses. C.D. performed and supervised analyses.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26083-4.

Correspondence and requests for materials should be addressed to Susanne Bornelov or Michaela Frye.

Peer review information Nature Communications thanks Nicholas Proudfoot and other, anonymous, reviewers for their contributions to the peer review of this work. Peer review reports are available.

Reprints and permission information are available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.