nature genetics

ARTICLES

https://doi.org/10.1038/s41588-020-0635-0

Selective Mediator dependence of cell-type-specifying transcription

Martin G. Jaeger[1], Björn Schwalb[2], Sebastian D. Mackowiak[3], Taras Velychko[2], Alexander Hanzl[1], Hana Imrichova[1], Matthias Brand[1], Benedikt Agerer[1], Someth Chorn[1], Behnam Nabet[4,5], Fleur M. Ferguson[4,5], André C. Müller[1], Andreas Bergthaler[1], Nathanael S. Gray[4,5], James E. Bradner[6], Christoph Bock[1,7], Denes Hnisz[3], Patrick Cramer[2]✉ and Georg E. Winter[1]✉

The Mediator complex directs signals from DNA-binding transcription factors to RNA polymerase II (Pol II). Despite this pivotal position, mechanistic understanding of Mediator in human cells remains incomplete. Here we quantified Mediator-controlled Pol II kinetics by coupling rapid subunit degradation with orthogonal experimental readouts. In agreement with a model of condensate-driven transcription initiation, large clusters of hypophosphorylated Pol II rapidly disassembled upon Mediator degradation. This was accompanied by a selective and pronounced disruption of cell-type-specifying transcriptional circuits, whose constituent genes featured exceptionally high rates of Pol II turnover. Notably, the transcriptional output of most other genes was largely unaffected by acute Mediator ablation. Maintenance of transcriptional activity at these genes was linked to an unexpected CDK9-dependent compensatory feedback loop that elevated Pol II pause release rates across the genome. Collectively, our work positions human Mediator as a globally acting coactivator that selectively safeguards the functionality of cell-type-specifying transcriptional networks.

DNA-binding transcription factors (TFs) guide the activity of RNA Pol II. Seminal in vitro studies established that eukaryotic TFs require intermediary factors to relay their regulatory information to Pol II core components[1,2]. Biochemical purification of this coactivator activity led to the identification of the large, multisubunit Mediator complex in numerous eukaryotic model organisms[2–9]. Mediator is organized into biochemically separable functional modules termed head, middle, tail and kinase modules[10–13]. Initially described as an absolute requirement for activator-dependent transcription, Mediator also stimulates basal transcription in vitro[4]. This prompted the question of whether Mediator is a general or gene-specific factor[14]. Yeast in vivo studies yielded variable results, partly finding that global mRNA levels were decreased as much in Mediator mutants as in Pol II mutants[15] and partly describing mild effects upon nuclear depletion of individual Mediator subunits[16]. Strong transcriptional defects were only observed upon simultaneous depletion of head and tail Mediator subunits[16]. In comparison to the understanding obtained from these and other studies in yeast[17,18], mechanistic understanding of human Mediator is emerging slowly, owing to experimental challenges.

Human transcription regulation differs from that of yeast in several key aspects. Most importantly, human genomes allow the specification of different cell types from the same genomic potential. To this end, human genomes evolved distal *cis*-regulatory enhancer elements, which concentrate a plethora of TFs and coactivators, including Mediator[19]. High levels of coactivator binding were used to classify large stretches of clustered enhancers, termed super-enhancers (SEs), which predominantly drive the expression of cell-type-specific genes[20,21]. A subset of SE-driven TFs interact in interconnected gene-regulatory networks that establish transcriptional circuits to control cell identity[22–24]. Recent models implicate the Mediator subunit MED1 with its intrinsically disordered region (IDR) in the formation of nuclear condensates at SE regions[25,26]. The molecular mechanisms of how Mediator subunits contribute to the organization of these subnuclear structures remain to be elucidated. Also, whether and how nuclear condensates influence Pol II transcriptional dynamics have not been addressed experimentally.

Enhancers are often megabases away from their target genes. Intricate mechanisms have evolved to ensure faithful communication with their target promoters, including tight regulation of chromatin architecture[27,28]. Mediator is located at the nexus of these processes, but whether it directly controls enhancer–promoter looping remains to be unambiguously resolved. Initial studies in mouse cells showed Mediator colocalization with cohesin and implicated both factors in the formation and maintenance of enhancer–promoter contacts at pluripotency genes[29]. A more recent study yielded opposite results, as overall chromatin architecture was largely unchanged after Mediator depletion[30]. Interestingly, the authors of the latter study implicated Mediator as a globally required transcriptional regulator, as they observed sevenfold-lower mRNA levels after long-term MED14 depletion for 60 h. However, extended perturbation periods limit causal conclusions.

Almost three decades after the discovery of Mediator, these unresolved questions call for mechanistic studies of Mediator in

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [2]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany. [3]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany. [4]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA. [5]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. [6]Novartis Institutes for BioMedical Research, Cambridge, MA, USA. [7]Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria. ✉e-mail: patrick.cramer@mpibpc.mpg.de; gwinter@cemm.oeaw.ac.at

human cells. Such studies have been hampered by the difficulty in rapidly perturbing essential Mediator components. Here we unravel direct Mediator functions by using human cells with alleles encoding pharmacologically degradable Mediator subunits. Acute Mediator loss strongly abrogated the transcription of SE-driven, cell-type-specifying genes in gene regulatory networks but only mildly impacted overall transcriptional output and genome architecture. Integrating nascent transcriptional readouts with mathematical modeling and chromatin proteomics led us to propose a two-pronged mechanism for the observed hypersensitivity of gene regulatory networks. First, Mediator degradation disrupted large nuclear Pol II clusters. This was accompanied by a transcriptional collapse of associated constituent genes of gene regulatory networks, which were characterized by highly optimized Pol II dynamics. Second, we observed an unanticipated feedback loop that globally elevated Pol II pause release rates via increased activity of CDK9, a subunit of the P-TEFb complex that was previously shown to physically interact with Mediator[31–34]. This compensatory mechanism was sufficient to maintain transcriptional output at the vast majority of genes. However, it failed to rescue transcription of the constituents of gene regulatory networks, where Mediator optimized transcriptional output by enabling maximal rates of Pol II turnover. Our data position human Mediator as a general coactivator that facilitates transcription globally but is strictly required for the functionality of cell-type-specifying gene regulatory circuits.

Results

Acute Mediator subunit degradation in human cells. To uncover the direct function of human Mediator, we generated eight alleles encoding endogenously tagged, degradation-sensitive Mediator subunits in near-haploid human KBM7 cells by using degradation tag (dTAG) technology (Fig. 1a and Supplementary Table 1)[35,36]. This approach enables ligand-dependent recruitment of tagged subunits to the $CRL4^{CRBN}$ E3 ligase complex, leading to their fast and selective proteasomal degradation. Particular emphasis was put on core Mediator subunits, previously defined to support a minimal functional Mediator[16,37]. All tagged subunits were amenable to target-selective, ligand-dependent proteolysis with near-complete degradation evident after 6 h of ligand exposure (Fig. 1b and Extended Data Fig. 1a,b). This allowed us to record the transcriptional fingerprints of individual subunit degradation through spike-in-normalized 3′ mRNA sequencing (Fig. 1c). We contextualized the results with BRD4/BET protein and CDK9 degradation, which are known to prompt global transcriptional arrest[38,39]. Mediator subunit depletion clearly differed from such global arrest and only mildly affected overall mRNA levels, despite similar degradation kinetics and effectiveness (Fig. 1c,d and Extended Data Fig. 1c,d). Transcriptional profiles after subunit destabilization primarily differed in perturbation strength, that is, quantitatively rather than qualitatively (Fig. 1e). Degradation of MED14 caused the most severe consequences on mRNA levels (Fig. 1c,e). Mediator subunit degradation segregated along principal component 2 (PC2), which distinguished impaired MYC-driven transcription (Fig. 1c and Extended Data Fig. 1e,f)[40]. Time-resolved immunoblots showed MED14 degradation already after 1 h, followed by notable MYC destabilization as early as 2 h (Extended Data Fig. 1g). Thus, MYC protein loss likely accounted for this secondary transcriptional phenotype, highlighting the need for readouts with higher kinetic resolution.

Structurally, MED14 acts as a central scaffold that anchors Mediator's head and middle modules[30,41–43]. In line with this conserved, essential role of MED14, continuous degradation of MED14 abrogated cell growth (Extended Data Fig. 2a). Coupling acute (2-h) MED14 degradation with size-exclusion chromatography showed major impacts on overall complex integrity (Extended Data Fig. 2b). In line with recently described coactivator condensates, we observed a highly focal distribution of both MED14 and MED1, with the majority of signal originating from large clusters of around 500 nm in diameter (Fig. 1f)[25,26]. These data are in good agreement with previously described stable Mediator clusters in mouse embryonic stem cell (mESC)-derived epiblast-like cells[26]. Notably, degradation of MED14 dispersed MED1 signal throughout the nucleus without affecting overall signal intensity (Fig. 1f and Extended Data Fig. 2c). This finding reflects MED14's role as a central Mediator scaffold and suggests an inability of MED1 to maintain clusters, even though it contains an extended intrinsically disordered region (Extended Data Fig. 2d). In accordance, MED14 degradation, but not degradation of MED1, prevented co-precipitation of other Mediator members with biotinylated isoxazole hydrogels, either through direct effects on Mediator integrity or indirect effects on other members of the transcription apparatus (Extended Data Fig. 2e)[44]. Together, these data indicate that acute MED14 degradation serves as a tool to rapidly disrupt overall Mediator integrity in human cells.

Nascent transcriptional profiling after MED14 degradation. The high kinetic resolution afforded by MED14 degradation motivated us to record direct effects on transcriptional activity, rather than mRNA levels, through spike-in-normalized transient transcriptome sequencing (TT-seq)[45]. Mediator disruption resulted in a notably asymmetric response. Transcription of *MYC* and *MYB* had completely collapsed after 1 h of MED14 degradation, whereas most genes, including the expression-matched *RAB3GAP1* control gene, were largely unaffected (Fig. 1g,h). Profound downregulation of additional cell identity genes supported a model in which human Mediator safeguards the functionality of cell-type-specifying transcriptional networks (Extended Data Fig. 2f). Cell-type-specific genes are typically controlled by large, clustered enhancers (SEs)[20,21]. A subset of SE-controlled TFs are required to establish, maintain and alter gene regulatory network topology[24]. These TFs can be further refined to a core set of autoregulatory TFs, which lock in cell identity by binding SE regions that drive their own expression[46]. Such autoregulatory TFs were among the most Mediator-dependent genes independently of their steady-state expression levels and formed a densely connected transcription regulatory circuit in KBM7 cells (Fig. 1i,j, Supplementary Table 2 and Supplementary Note).

To further evaluate whether Mediator selectively drives the expression of cell-type-specifying genes, we also engineered HCT-116 colorectal carcinoma cells to endogenously express MED14–dTAG (Supplementary Table 1). These cells were sensitive to $dTAG^{V}$-1 ligand hijacking the von Hippel–Lindau $CRL2^{VHL}$ E3 ligase, but not to dTAG7 (Extended Data Fig. 3a), likely owing to a previously observed lack of activity of CRBN-based degraders in HCT-116 cells[38,47]. MED14 degradation in HCT-116 cells was rapid, with near-complete degradation evident after 1 h (Extended Data Fig. 3b). Also here, Mediator disruption strongly abrogated nascent transcription of SE-proximal genes and cell-type-specifying autoregulatory TFs. In line with data derived after MED14 degradation in KBM7 cells, global transcription was again much less affected (Extended Data Fig. 3c–f). As expected, autoregulatory TFs in HCT-116 cells were distinct from those in KBM7 cells, with *MYC* being the only TF in common (Extended Data Fig. 3g and Supplementary Table 2). Notably, the transcriptional sensitivity of a given TF gene to Mediator ablation was dependent on participation in a cell-type-specific autoregulatory network (Fig. 1k and Extended Data Fig. 3h,i). These findings argue that lineage-specifying transcription is hypersensitive to Mediator disruption.

Impact of Mediator loss on chromatin conformation. We next addressed whether the observed selective shutdown in cell-type-specifying transcription and Mediator cluster disassembly could be explained by a disruption of SE–promoter contacts.

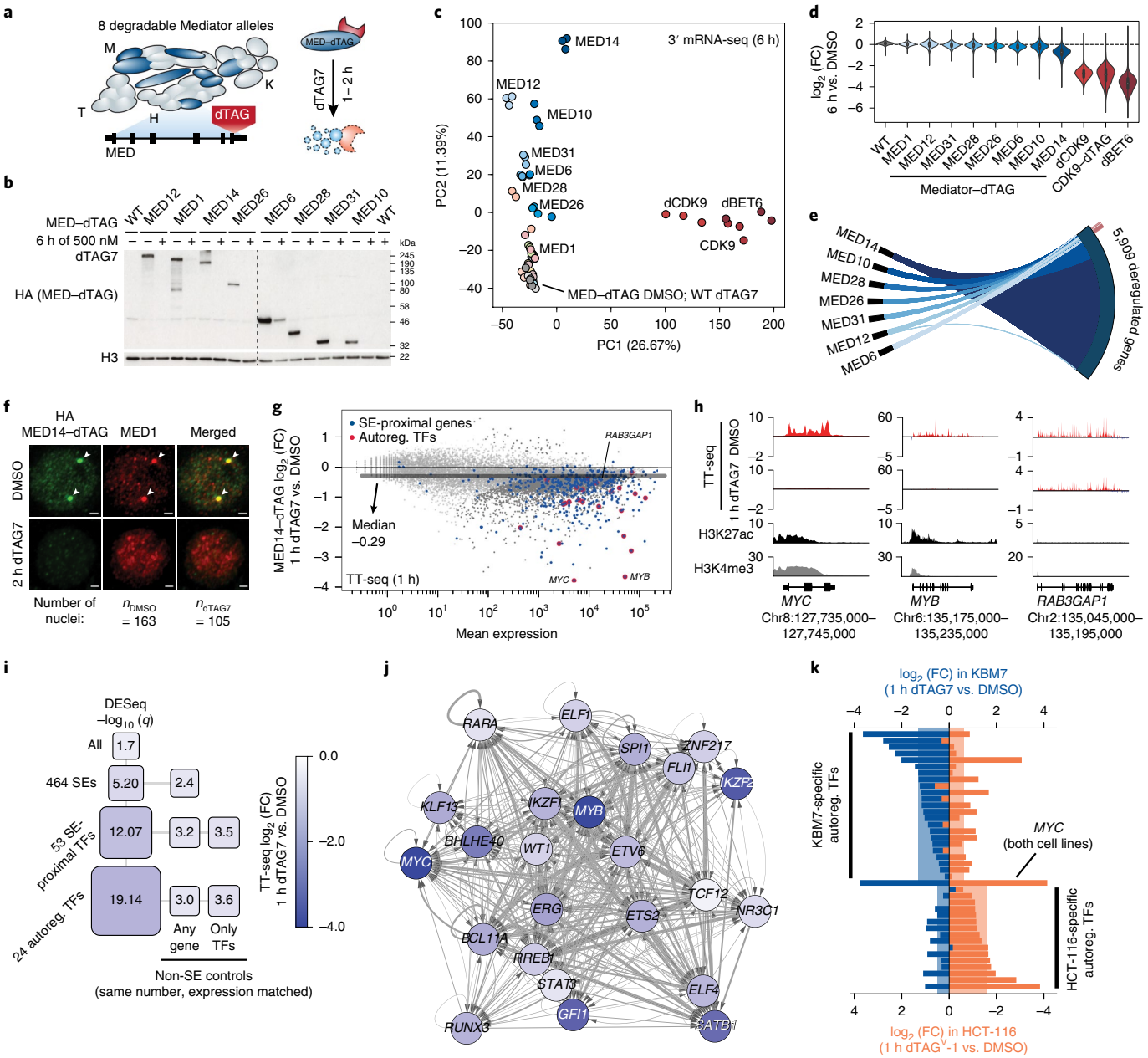


Fig. 1 | Acute Mediator loss selectively abrogates the functionality of cell-type-specifying transcriptional circuits. a, MED–dTAG alleles and the location of the corresponding subunits in the complex. H, head; M, middle; T, tail; K, kinase module. **b**, Destabilization of MED–dTAG fusion proteins. WT, wild type. The unprocessed western blot is available as source data. **c**, Principal-component analysis of transcriptional fingerprints after 6 h ($n = 3$ independent drug treatments). Gene-symbol labels indicate dTAG-carrying cells treated with 500 nM dTAG7. dCDK9 and dBET6 labels indicate direct degradation of CDK9 or BET proteins in wild-type cells with 250 nM THAL-SNS-032 (dCDK9) or 100 nM dBET6. **d**, Spike-in-normalized $\log_2$-transformed fold change (FC) in mRNA of degrader-treated cell lines ($n = 8{,}798$ genes; median of three independent treatments). Violin plots show the approximate density distribution with internal box plots showing medians with interquartile range and whiskers indicating 1.5 times the interquartile range. **e**, Genes deregulated by MED–dTAG degradation (left) converge on consensus targets (right). Red bars, genes deregulated by ≥3 perturbations. **f**, Immunofluorescence of MED1 foci (arrowheads) in MED14–dTAG cells. Maximum-intensity projections of three-dimensional (3D) images are shown. Scale bars, 1 µm. **g**, Differences in TT-seq nascent transcript levels ($n = 2$ independent treatments). Significantly deregulated (DESeq2, $q < 0.01$; dark gray), SE-proximal (blue) and autoregulatory TF (red) genes are highlighted. The dark gray line indicates the median $\log_2$-transformed fold change for all $n = 19{,}559$ transcriptional units. **h**, TT-seq signal of two autoregulatory TFs (*MYC* and *MYB*) and an expression-matched control gene. H3K27ac and H3K4me3 chromatin immunoprecipitation and sequencing (ChIP–seq) signals are from wild-type KBM7 cells. **i**, Fold change (color) and significance (size) of SE-driven cell identity gene sets and expression-matched control gene sets (data as in **g**). **j**, Twenty-four autoregulatory TFs constitute a KBM7 cell-type-specifying gene regulatory network. Arrows indicate that the given TF has binding motifs in the target TF's SE region(s). Edge weight is proportional to the number of motifs. **k**, Cell-type-specific impact of MED14 degradation on KBM7 and HCT-116 autoregulatory TFs. Shaded areas correspond to mean $\log_2$-transformed fold change in expression of the underlying TF genes.

To this end, we coupled acute MED14 degradation in KBM7 cells with acetylated histone H3 at lysine 27 (H3K27ac) HiChIP analysis[48]. In agreement with a recent report[30], Mediator degradation was largely inconsequential for overall chromatin architecture, leaving the majority of enhancer–promoter contacts intact (Fig. 2a–c and Extended Data Fig. 4a–c). Contacts involving SEs did display a

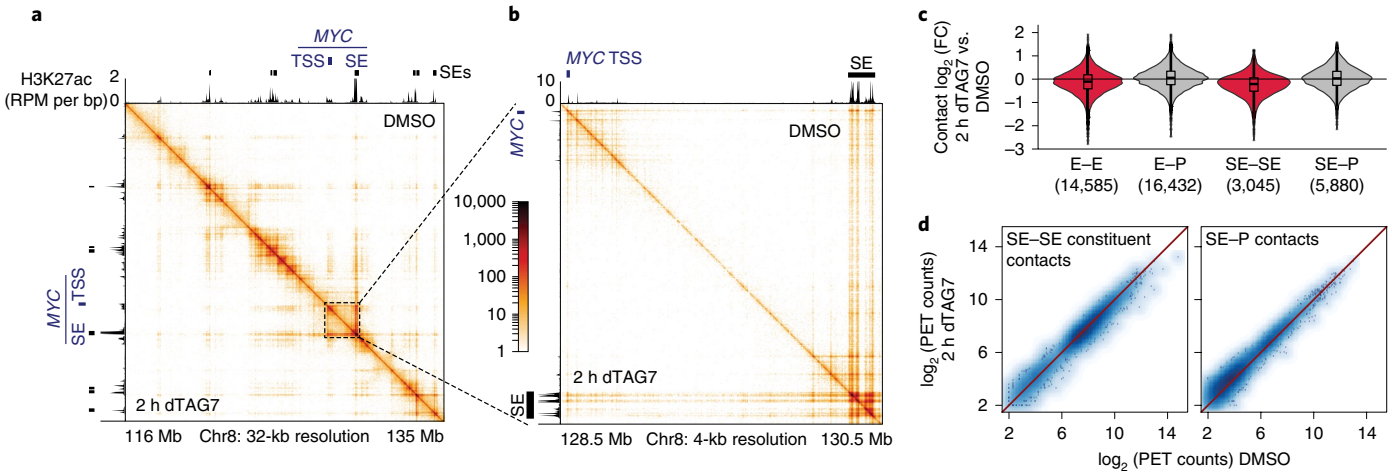


Fig. 2 | Mediator is dispensable for maintenance of enhancer–promoter contacts. a, H3K27ac HiChIP contact matrices for an approximately 20-Mb region encompassing the *MYC* locus. Color corresponds to the normalized number of contacts between a pair of loci. The upper triangle shows contacts after 2 h of DMSO treatment, and the lower triangle shows contacts after 2 h of MED14 degradation. Axis tracks show KBM7 H3K27ac ChIP-seq signal (capped at 2 reads per million (RPM) per bp) and SEs. TSS, transcriptional start site. **b**, Zoomed-in region of **a** highlighting *MYC* SE–promoter and SE constituent–constituent contacts. **c**, Impact of MED14 degradation on H3K27ac HiChIP contact frequency by contact class (E, enhancer; P, promoter; SE, SE constituent). The number of quantified contacts of each type is given in parentheses. Violin plots show the approximate density distribution with internal box plots showing medians with interquartile range and whiskers indicating 1.5 times the interquartile range. **d**, Paired-end tag (PET) counts of contacts involving SE constituents.

subtle downward trend after degradation, albeit at a minor effect size (Fig. 2c,d). Global H3K27ac was only mildly affected at early time points, although differential H3K27ac levels at SE loci might still contribute to the observed changes in contact frequency (Extended Data Fig. 4d). To rule out such immunoprecipitation-based biases of H3K27ac HiChIP, we confirmed these modest changes by using in situ circularized chromosome conformation capture and sequencing (4C–seq) on selected SE-proximal genes (Extended Data Fig. 4e,f)[49]. Nevertheless, on the basis of their minor effect size, changes in chromatin architecture did not seem to account for the drastic defects in cell-type-specifying transcription.

Pol II clusters and enhancer transcription depend on MED14. We next aimed to understand whether the selectively abrogated transcription of cell-type-specifying genes coincided with a loss of enhancer activity. Transcription of enhancer regions has emerged as a key proxy for their activity[50]. To record the genome-wide distribution of transcriptionally engaged polymerases, we used spike-in-normalized precision nuclear run-on sequencing (PRO-seq)[51]. Importantly, this assay does not provide information on Pol II recruitment in non-transcribing states, such as preinitiation complexes[52]. Already at 1 h, MED14 degradation eradicated transcribing polymerase at enhancers and SE constituents in particular, indicating a lack of Pol II activity at these loci (Fig. 3a,b). To test emerging models of condensate-driven transcription initiation, we next wanted to address whether Mediator directs the formation of Pol II clusters[53,54]. We hence engineered an N-terminal monomeric enhanced green fluorescent protein (mEGFP) fusion construct into the endogenous *POLR2A* locus of MED14–dTAG KBM7 cells (Supplementary Table 1). Indeed, large Pol II foci of around 500 nm in diameter[26,55], which overlapped MED14 clusters at steady state, disassembled upon acute MED14 degradation (Fig. 3c and Extended Data Fig. 5a,b). These clusters likely represented non-transcribing Pol II, as evidenced by immunofluorescence experiments using an antibody directed against hypophosphorylated Pol II (Extended Data Fig. 5c). Concomitantly, Mediator loss eradicated Pol II at SE target genes and especially the subset of SE-controlled autoregulatory TFs, in accordance with TT-seq results (Fig. 3d and Extended Data Fig. 5d,e). Pol II density in non-SE gene bodies was less affected, which again shows that the transcriptional output of the majority of genes is only subtly decreased after 1 h of Mediator ablation (Fig. 3d and Extended Data Fig. 5d). At human genes, Pol II pauses close to the promoter for several minutes before productively elongating a transcript. This pause might serve as a quality-control checkpoint and has been shown to sterically limit upstream Pol II initiation[56–58]. We observed a profound and global decrease in promoter-proximal Pol II after MED14 degradation, indicative of a change in pause dynamics (Fig. 3e and Extended Data Fig. 5f). This unexpected finding prompted us to investigate how Mediator loss affects Pol II transcription kinetics.

Measuring Mediator-controlled Pol II turnover dynamics. To derive a quantitative kinetic model of Mediator-dependent human gene control, we performed integrative analysis of polymerase location (PRO-seq) and RNA synthesis activity (TT-seq). Combining information on the number of RNA molecules being synthesized (TT-seq) with genome-wide polymerase density maps (PRO-seq) informs on the number of transcribing polymerases and their turnover dynamics per unit time at near-nucleotide resolution[56,59]. This multi-omics modeling allowed us to calculate genome-wide rates of productive transcription initiation and pause duration[56]. Productive initiation events are defined as events that lead to successful Pol II escape, not only from the promoter but also from the pause site into the gene body (Supplementary Note). Transcriptional output increases with higher productive initiation rates and shorter pause duration and is confined by steric considerations, manifesting as the pause initiation limit[56,59]. Locating genes in the pause initiation landscape revealed that SE-proximal autoregulatory TFs operated more closely to the theoretical pause initiation limit than expected by chance (Fig. 3f and Extended Data Fig. 5g)[56]. This optimized location was predominantly governed by very high productive initiation rates (Fig. 3f). Acute loss of Mediator selectively diminished the productive initiation rates of autoregulatory TFs, whereas most genes maintained normal productive initiation rates (Fig. 3g,h and Extended Data Fig. 5h). Concomitantly, we observed an unexpected global decrease in pause duration, indicating more efficient Pol II release into active elongation (Fig. 3g,h and Extended Data Fig. 5h). Together, these findings indicate that Mediator organizes Pol II clusters and exerts highly selective control over productive initiation at

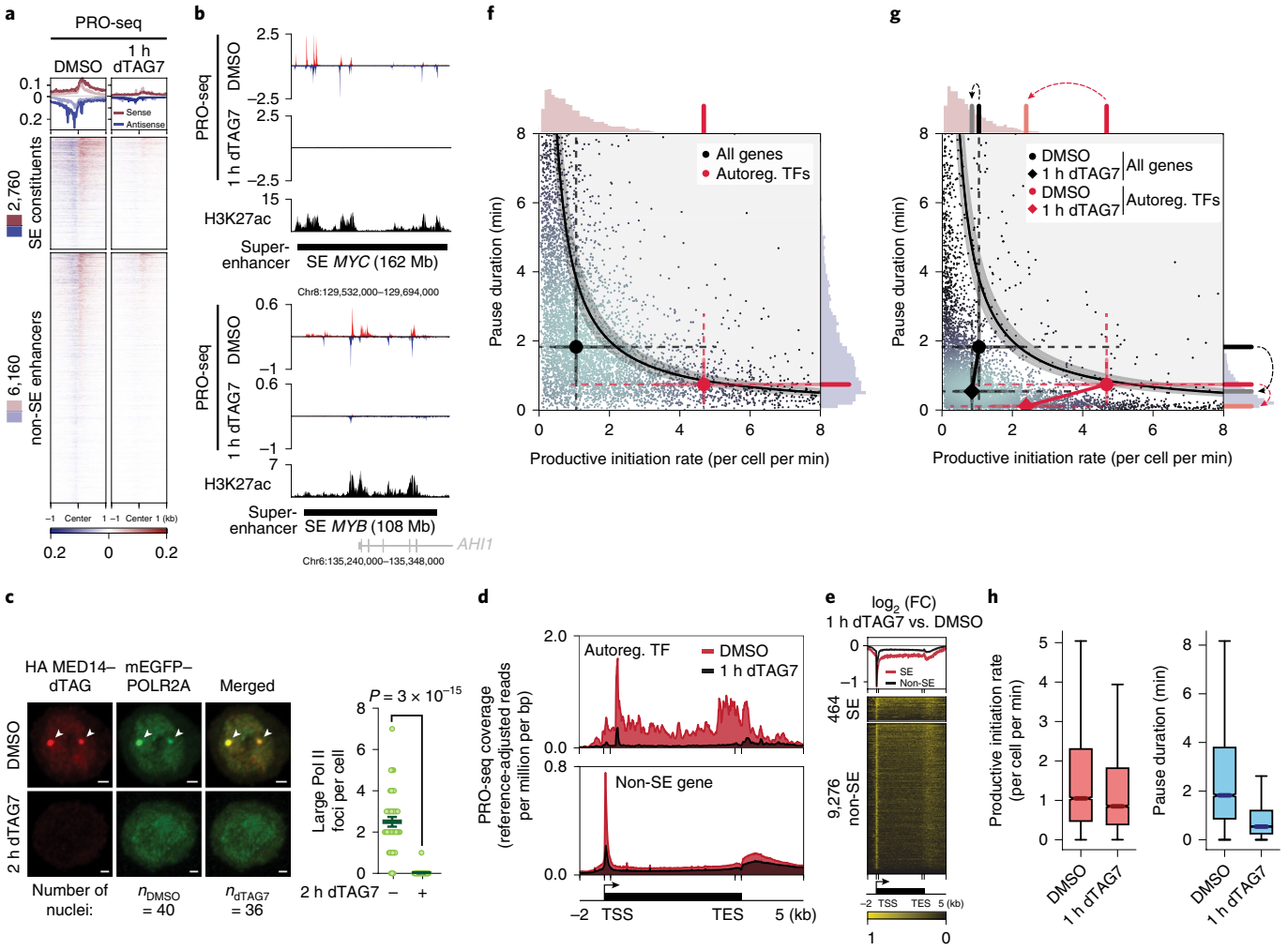


Fig. 3 | Mediator organizes Pol II clusters to optimize the transcriptional dynamics of cell-type-specifying gene regulatory networks. a, Impact of 1 h of MED14 degradation on polymerase engagement at enhancers. **b**, PRO-seq signal for two SE regions. H3K27ac ChIP–seq signal is from unperturbed wild-type KBM7 cells. **c**, Fixed-cell imaging of large Pol II foci (arrowheads) in MED14–dTAG mEGFP–POLR2A KBM7 cells. Maximum-intensity projections of 3D images are shown. Scale bars, 1 µm. Data in the graph are shown as the mean ± s.e.m., with the *P* value calculated by two-sided, unpaired *t* test. Further quantification and controls are shown in Extended Data Fig. 5a,b. **d**, Aggregate PRO-seq coverage of autoregulatory TF and non-SE metagenes. TES, transcriptional end site. **e**, Enrichment heat map of PRO-seq $\log_2$-transformed fold change over SE and non-SE metagenes. **f**, Steady-state productive initiation rates and pause durations for 6,791 transcriptional units in MED14–dTAG cells. The black center line denotes the theoretical pause initiation limit with 15% uncertainty[56]. Markers indicate median gene set values with protruding whisker (dashed) and interquartile range (solid) proxies. Axis histograms highlight the median values for autoregulatory TFs. **g**, Pause initiation values after 1 h of MED14 degradation. Dashed arrows in the histogram indicate the median pause duration and productive initiation rate trajectories from DMSO treatment to 1 h of MED14 degradation. **h**, Changes in productive initiation rates and pause durations for all *n* = 6,791 transcriptional units. Box plots show medians with interquartile range, with whiskers indicating 1.5 times the interquartile range and confidence region notches.

SE-proximal cell-type-specifying genes, while apparently having less influence on the bulk of genes. The global decrease in pause duration led us to hypothesize that defects in preinitiation complex formation without Mediator are partially compensated by more efficient pause release of residually initiating Pol II.

MED14 loss triggers P-TEFb release from 7SK particles. We next aimed to identify factors that might induce the observed changes in pause dynamics. To this end, we subjected purified chromatin fractions to unbiased, label-free mass spectrometry analysis. We intersected the results of three independent data analysis strategies to focus on high-confidence hits (Extended Data Fig. 6a, Supplementary Table 3 and Supplementary Note). Finally, we identified biochemically interacting hits by mapping differentially chromatin-bound proteins to publicly available protein–protein interaction data[60]. MED14 depletion was nearly saturating at 95% and evicted other Mediator subunits from chromatin (Fig. 4a and Extended Data Fig. 6b,c). Binding of general transcription factors remained largely unchanged, and Pol II components were only mildly decreased (Fig. 4a and Extended Data Fig. 6b,c). In contrast, three factors were more stably bound to chromatin with high confidence: the main pause release factor CDK9, its cognate cyclin CCNT1 and MEPCE, all members of the positive transcription elongation factor b (P-TEFb) and 7SK regulatory complexes (Fig. 4a and Extended Data Fig. 6c,d)[61]. As we elaborate in the Discussion, this finding is in apparent disagreement with previous reports, which positioned Mediator as a positive regulator of pause release by directly recruiting P-TEFb via its MED23 or MED26 subunit[31–33]. Unexpectedly, our data imply the presence of a mechanism that leads to a net increase in P-TEFb binding to chromatin upon acute

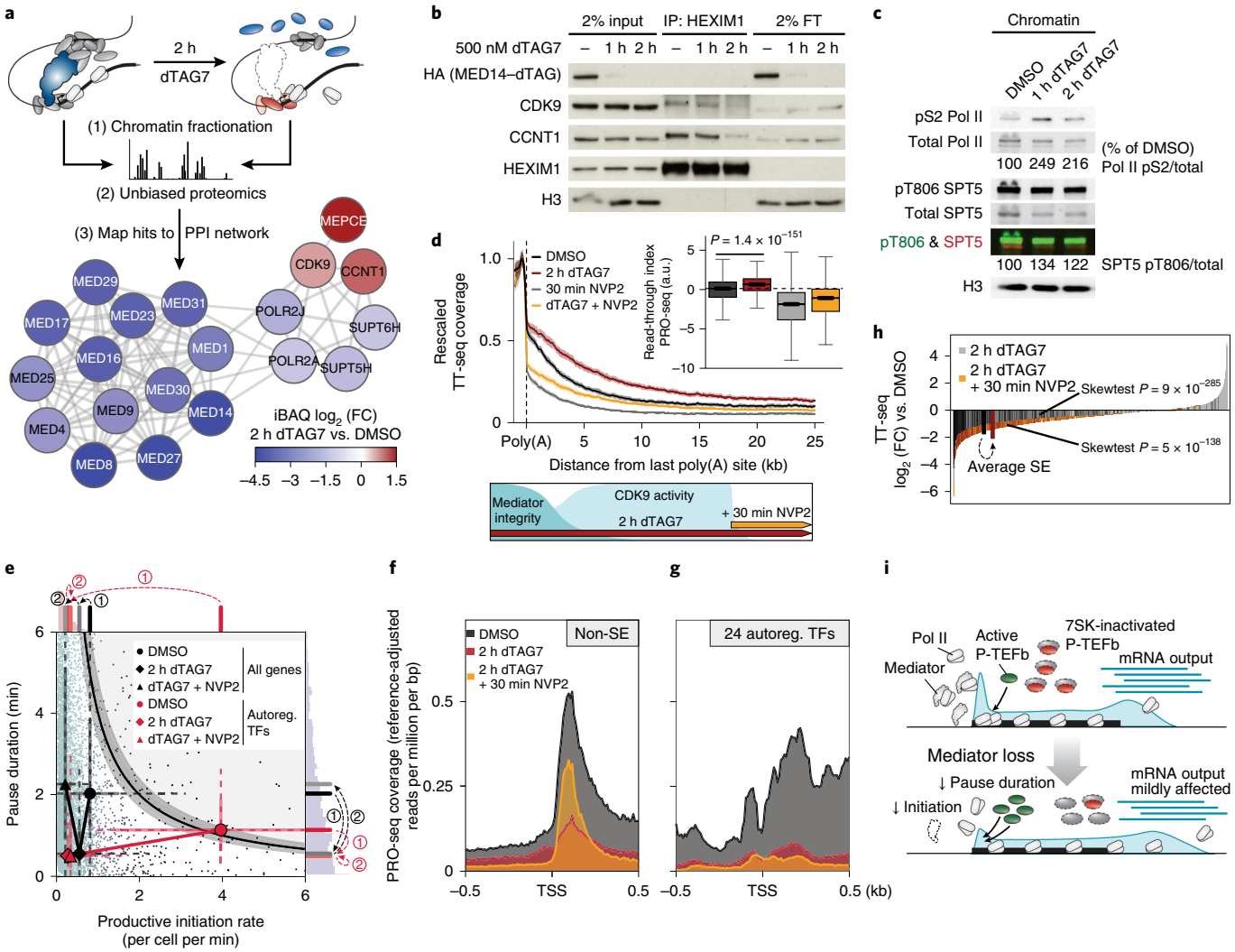


Fig. 4 | Compensatory P-TEFb activation boosts non-super-enhancer output to shape the Mediator hyperdependence of cell-type-specifying transcription. a, Chromatin proteomics strategy and results. Chromatin fractions were subjected to unbiased, label-free mass spectrometry. High-confidence hits were mapped to public protein–protein interaction (PPI) data[60]. The largest connected component network shows differential chromatin binding by isobaric absolute quantification (iBAQ). **b**, CDK9 and CCNT1 co-immunoprecipitation with HEXIM1-containing inhibitory 7SK particles. IP, immunoprecipitation; FT, flow-through. **c**, CDK9 target phosphorylation on purified chromatin fractions. **d**, Aggregate TT-seq coverage past polyadenylation sites ($n = 10{,}949$ transcriptional units; mean ± bootstrapped confidence region). Treatment scheme: 2 h of MED14 degradation (500 nM dTAG7) with or without CDK9 inhibition (500 nM NVP2) in the last 30 min. Box plots show PRO-seq read-through indices ($n = 5{,}558$ genes; Mann–Whitney *U* test) as the median with interquartile range, with whiskers indicating 1.5 times the interquartile range and confidence region notches. **e**, Productive initiation rates and pause durations for 6,954 transcriptional units after combined MED14 and CDK9 perturbation. Markers show median trajectories from DMSO control (circles) to 2 h of MED14 degradation (diamonds) and combined MED14 and CDK9 perturbation (triangles). Axis histograms further illustrate these trajectories. **f**,**g**, Aggregate PRO-seq coverage around the TSSs of non-SE genes (**f**) and autoregulatory TFs (**g**). **h**, Rank-ordered TT-seq differential expression for 2 h of MED14 degradation (SE genes in black) or combined MED14 and CDK9 perturbation (SE genes in red) as compared to DMSO control ($n = 20{,}882$ transcriptional units). Bold bars denote mean SE ranks and $\log_2$-transformed fold changes. *P* values from all-gene distribution, two-sided skewness tests. **i**, Model of how P-TEFb activation might compensate for less efficient initiation in response to Mediator loss. This mechanism falls short of rescuing initiation defects at autoregulatory TFs. Unprocessed western blots for **b**,**c** are available as source data.

Mediator loss. At steady state, the majority of P-TEFb is inactivated by incorporation into 7SK ribonucleoproteins (RNPs), which consist of 7SK RNA, HEXIM1 or HEXIM2, LARP7 and MEPCE[61]. Although several transcriptional stresses, including CDK9 inhibition, are known to release P-TEFb from 7SK particles, functional impacts on Pol II dynamics remain elusive[61]. MED14 degradation also triggered 7SK disassembly, as indicated by decreased CDK9 and CCNT1 binding to HEXIM1 (Fig. 4b). P-TEFb release led to increased phosphorylation of the chromatin-bound CDK9 targets Pol II at C-terminal repeat domain (CTD) residue Ser2 and SPT5 at Thr806 (Fig. 4c)[61,62]. SPT5 phosphorylation is known to regulate Pol II velocity in the termination window at the 3′ ends of genes[63]. In agreement with this, we observed vast CDK9-dependent read-through transcription, even at long genes where newly elongating Pol II had not yet reached the gene end (Fig. 4d and Extended Data Fig. 7a–c). The latter finding rules out the possibility that these phenotypes are due to differences in elongation complex formation close to the promoter. In summary, unbiased chromatin proteomics led us to identify P-TEFb hyperactivation by 7SK disassembly in response to Mediator loss.

Hypersensitive genes are refractory to P-TEFb compensation. We next addressed whether P-TEFb hyperactivation shapes the transcription regulatory consequences of acute Mediator depletion. To this end, we pharmacologically inhibited CDK9 in the last 30 min of 2 h of MED14 degradation. This allowed us to assess the role of CDK9 activity at a time of near-complete Mediator depletion (scheme in Fig. 4d). Pause initiation modeling confirmed that CDK9 activity was indeed responsible for decreased pause duration after Mediator loss (Fig. 4e and Extended Data Fig. 7d,e). Upon CDK9 inhibition, productive initiation rates decreased even further, implying the presence of residual Pol II initiation despite Mediator depletion (Fig. 4e and Extended Data Fig. 7f). We hence approximated promoter initiation activity by measuring whether polymerases continued to enter the pause site. Indeed, blocking pause release after near-complete Mediator degradation substantially increased promoter-proximal Pol II occupancy (Fig. 4f and Extended Data Fig. 7g). We conclude that P-TEFb activation partially compensates for Mediator loss by globally increasing pause release. The efficacy of this compensation is contingent on residual Pol II initiation.

Could this compensatory mechanism contribute to the apparent Mediator hyperdependence of cell-type-specifying transcription? In fact, additional CDK9 inhibition failed to re-accumulate promoter-proximal Pol II at autoregulatory TFs and did not revert their initial drop in pause duration (Fig. 4e,g and Extended Data Fig. 7g). Hence, Mediator appeared to be a strict requirement for successful Pol II initiation at these cell-type-specifying genes. On a transcript level, preventing P-TEFb compensation rendered the Mediator phenotype less SE selective (Fig. 4h and Extended Data Fig. 7h). Together, these experiments uncover a new phenotypic buffering capacity of promoter-proximal pausing, wherein transient P-TEFb activation can partially compensate for defects in transcription initiation (Fig. 4i). Although this mechanism seemingly operates across the genome, it falls short of compensating for initiation defects at cell-type-specifying genes, which are transcriptionally optimized for efficient Mediator-cluster-driven Pol II recruitment. In summary, human Mediator contributes to global transcription but is predominantly deployed to ensure the functionality of cell-type-specifying transcriptional circuits.

Discussion

Here we combined rapid pharmacological degradation with multi-omics kinetic analyses to derive direct Mediator functions in human transcription regulation. We focused on MED14 degradation to disrupt overall Mediator integrity. MED14 was asymmetrically required to maintain Mediator and Pol II clusters and transcription of SE-driven cell-type-specifying genes, but not overall genome architecture. Mediator did stimulate Pol II initiation globally but appeared to be strictly required for the functionality of cell-type-specifying transcriptional circuits. Two mechanisms contributed to the observed hypersensitivity of cell-type-specifying gene regulatory networks. First, Mediator sustained unusually high rates of Pol II turnover at the constituent genes of these networks. Second, a genome-wide increase in Pol II pause release unveiled a new P-TEFb-activating mechanism, which compensated for defective initiation at non-SE genes.

Mechanistic conclusions were enabled by our experimental setup, which coupled precise and fast perturbations with orthogonal nascent transcriptional readouts. Notably, key TF proteins were undetectable within hours of MED14 degradation, likely owing to their transcriptional shutdown, which highlights the need for time-resolved analyses. Previously reported global transcriptional collapse after long-term Mediator depletion[30] might thus result from secondary effects that are due to dysfunctional TF machinery at the time of measurement. Our findings also reinforce the rationale behind and feasibility of pharmacological inhibition of general transcriptional coactivators as a gene-selective therapeutic strategy[64].

We surmised that globally decreased pause duration compensated for compromised Pol II initiation at most genes. We also showed that shorter pausing is most likely due to a CDK9-activity-dependent increase in pause release. However, less efficient establishment of paused Pol II complexes might also be a contributing factor. Of note, previous work has demonstrated a physical interaction between Mediator and P-TEFb and functionally implicated Mediator in contributing to P-TEFb recruitment[31–34]. Specifically, knockdown for 48 h of *MED26* led to decreased P-TEFb binding at selected genes as assayed by ChIP–qPCR[31]. Another study described partially overlapping morphological defects in *Drosophila* embryos depleted of maternally derived *MED26* or *Cdk9* mRNA[33]. In a third study, ChIP–qPCR in *Med23*$^{-/-}$ mESCs showed decreased P-TEFb binding to selected genes[32]. Differences in kinetic resolution and cellular models render it difficult to unambiguously reconcile these findings with the apparently contradictory functional impact of acute MED14 degradation on Pol II pause dynamics. Among other possibilities, complex-independent roles of MED26 or MED23 might be contributing and could be addressed in future studies. Overall, it remains to be addressed how acute Mediator disruption can prompt an increase in P-TEFb occupancy on chromatin and an ensuing increase in global pause release rates.

Our data are consistent with a model where increased CDK9 activity tips the balance to favor elongation licensing over putative early termination in the pause site[56,57,59,65]. These findings imply that pausing can serve as a phenotypic buffering mechanism to bridge periods of defective initiation. Moreover, these findings might directly reflect the evolutionary advantage of why paused Pol II limits upstream transcription initiation, particularly in human cells[56,57]. Systematic biochemical analyses uncovered functional compensation that was driven by P-TEFb shedding from inhibitory 7SK particles. Although many stresses are known to trigger 7SK disassembly[66], we here show a functional impact on pause duration. Notably, our data do not provide direct evidence of the genomic regions that P-TEFb binds in response to Mediator disruption. Spatial information in this study is limited to the functional impacts on transcription dynamics, where we observe decreased pause duration at the vast majority of genes. It thus remains to be fully resolved how initiation defects are sensed in cells and how P-TEFb is recruited on a genome-wide scale[66].

Mediator was predominantly required for productive Pol II initiation at cell-type-specifying genes. We observed that Mediator organized clusters to recruit and supply Pol II, which is consistent with emerging models of condensate-driven transcriptional initiation[53,54]. Notably, our fixed-cell imaging experiments do not allow any conclusions on whether these clusters are truly phase-separated compartments[67,68]. Nonetheless, the fact that Mediator orchestrates nuclear Pol II distribution and optimizes the transcriptional dynamics of SE-driven autoregulatory TF genes argues for a strong mechanistic link between cluster formation and cell-type-specifying transcription. This interpretation extends well-established biochemical principles of preinitiation complex formation to human cells and positions human Mediator as a globally acting cofactor that selectively safeguards the functionality of cell-type-specifying gene regulatory networks[69].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-020-0635-0.

Received: 2 December 2019; Accepted: 24 April 2020;
Published online: 1 June 2020

References

1. Kelleher, R. J. 3rd, Flanagan, P. M. & Kornberg, R. D. A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **61**, 1209–1215 (1990).
2. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**, 235–239 (2005).
3. Thompson, C. M., Koleske, A. J., Chao, D. M. & Young, R. A. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361–1375 (1993).
4. Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599–608 (1994).
5. Fondell, J. D., Ge, H. & Roeder, R. G. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl Acad. Sci. USA* **93**, 8329–8333 (1996).
6. Jiang, Y. W. et al. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc. Natl Acad. Sci. USA* **95**, 8538–8543 (1998).
7. Malik, S. & Roeder, R. G. Dynamic regulation of Pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30**, 256–263 (2005).
8. Conaway, R. C., Sato, S., Tomomori-Sato, C., Yao, T. & Conaway, J. W. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* **30**, 250–255 (2005).
9. Kim, Y. J. & Lis, J. T. Interactions between subunits of *Drosophila* Mediator and activator proteins. *Trends Biochem. Sci.* **30**, 245–249 (2005).
10. Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
11. Soutourina, J. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **19**, 262–274 (2018).
12. Jeronimo, C. & Robert, F. The Mediator complex: at the nexus of RNA polymerase II transcription. *Trends Cell Biol.* **27**, 765–783 (2017).
13. Eychenne, T., Werner, M. & Soutourina, J. Toward understanding of the mechanisms of Mediator function in vivo: focus on the preinitiation complex assembly. *Transcription* **8**, 328–342 (2017).
14. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* **11**, 761–772 (2010).
15. Holstege, F. C. et al. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
16. Petrenko, N., Jin, Y., Wong, K. H. & Struhl, K. Evidence that Mediator is essential for Pol II transcription, but is not a required component of the preinitiation complex in vivo. *eLife* **6**, e28447 (2017).
17. Jeronimo, C. et al. Tail and kinase modules differently regulate core Mediator recruitment and function in vivo. *Mol. Cell* **64**, 455–466 (2016).
18. Petrenko, N., Jin, Y., Wong, K. H. & Struhl, K. Mediator undergoes a compositional change during transcriptional activation. *Mol. Cell* **64**, 443–454 (2016).
19. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286 (2014).
20. Whyte, W. A. et al. Master transcription factors and Mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
21. Loven, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
22. Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
23. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
24. Davidson, E. H. Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920 (2010).
25. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
26. Cho, W.-K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412 (2018).
27. Levine, M., Cattoglio, C. & Tjian, R. Looping back to leap forward: transcription enters a new era. *Cell* **157**, 13–25 (2014).
28. Kim, S. & Shendure, J. Mechanisms of interplay between transcription factors and the 3D genome. *Mol. Cell* **76**, 306–319 (2019).
29. Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
30. El Khattabi, L. et al. A pliable Mediator acts as a functional rather than an architectural bridge between promoters and enhancers. *Cell* **178**, 1145–1158 (2019).
31. Takahashi, H. et al. Human Mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**, 92–104 (2011).
32. Wang, W. et al. Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. *Transcription* **4**, 39–51 (2013).
33. Dahlberg, O., Shilkova, O., Tang, M., Holmqvist, P. H. & Mannervik, M. P-TEFb, the super elongation complex and Mediator regulate a subset of non-paused genes during early *Drosophila* embryo development. *PLoS Genet.* **11**, e1004971 (2015).
34. Conaway, R. C. & Conaway, J. W. The Mediator complex and transcription elongation. *Biochim. Biophys. Acta* **1829**, 69–75 (2013).
35. Erb, M. A. et al. Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* **543**, 270–274 (2017).
36. Nabet, B. et al. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018).
37. Cevher, M. A. et al. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nat. Struct. Mol. Biol.* **21**, 1028–1034 (2014).
38. Winter, G. E. et al. BET bromodomain proteins function as master transcription elongation factors independent of CDK9 recruitment. *Mol. Cell* **67**, 5–18 (2017).
39. Olson, C. M. et al. Pharmacological perturbation of CDK9 using selective CDK9 inhibition or degradation. *Nat. Chem. Biol.* **14**, 163–170 (2018).
40. Muhar, M. et al. SLAM-seq defines direct gene-regulatory functions of the BRD4–MYC axis. *Science* **360**, 800–805 (2018).
41. Plaschka, C. et al. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518**, 376–380 (2015).
42. Nozawa, K., Schneider, T. R. & Cramer, P. Core Mediator structure at 3.4 Å extends model of transcription initiation complex. *Nature* **545**, 248–251 (2017).
43. Tsai, K. L. et al. Mediator structure and rearrangements required for holoenzyme formation. *Nature* **544**, 196–201 (2017).
44. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
45. Schwalb, B. et al. TT-seq maps the human transient transcriptome. *Science* **352**, 1225–1228 (2016).
46. Saint-Andre, V. et al. Models of human core transcriptional regulatory circuitries. *Genome Res.* **26**, 385–396 (2016).
47. Nabet, B. et al. Rapid and direct control of target protein levels with VHL-recruiting dTAG molecules. Preprint at *bioRxiv* https://doi.org/10.1101/2020.03.13.980946 (2020).
48. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
49. Weintraub, A. S. et al. YY1 is a structural regulator of enhancer–promoter loops. *Cell* **171**, 1573–1588 (2017).
50. Li, W., Notani, D. & Rosenfeld, M. G. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* **17**, 207–223 (2016).
51. Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).
52. Wissink, E. M., Vihervaara, A., Tippens, N. D. & Lis, J. T. Nascent RNA analyses: tracking transcription and its regulation. *Nat. Rev.Genet.* **20**, 705–723 (2019).
53. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).
54. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
55. Guo, Y. E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
56. Gressel, S. et al. CDK9-dependent RNA polymerase II pausing controls transcription initiation. *eLife* **6**, e29736 (2017).
57. Shao, W. & Zeitlinger, J. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.* **49**, 1045–1051 (2017).
58. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
59. Ehrensberger, A. H., Kelly, G. P. & Svejstrup, J. Q. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell* **154**, 713–715 (2013).
60. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
61. Zhou, Q., Li, T. & Price, D. H. RNA polymerase II elongation control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
62. Sanso, M. et al. P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev.* **30**, 117–131 (2016).
63. Parua, P. K. et al. A Cdk9–PP1 switch regulates the elongation–termination transition of RNA polymerase II. *Nature* **558**, 460–464 (2018).
64. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional addiction in cancer. *Cell* **168**, 629–643 (2017).
65. Krebs, A. R. et al. Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. *Mol. Cell* **67**, 411–422 (2017).

66. Li, Y., Liu, M., Chen, L. F. & Chen, R. P-TEFb: finding its ways to release promoter-proximally paused RNA polymerase II. *Transcription* **9**, 88–94 (2018).
67. Mir, M., Bickmore, W., Furlong, E. E. M. & Narlikar, G. Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? *Development* **146**, dev182766 (2019).
68. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
69. Roeder, R. G. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791 (2019).

Methods

Cell culture, cellular growth and lentiviral transduction. Human near-haploid chronic myeloid leukemia KBM7 cells were grown in IMDM supplemented with 10% FBS and 1× penicillin-streptomycin (Gibco). Human diploid HCT-116 colorectal carcinoma cells were grown in 15-cm cell culture dishes in RPMI supplemented with 10% FBS and 1× penicillin-streptomycin (Gibco) and passaged every 3–4 d upon reaching ~70% confluence. *Drosophila* S2 cells were grown in Schneider's *Drosophila* medium supplemented with 10% FBS (Gibco) and used as exogenous spike-in for all PRO-seq experiments.

For growth-over-time analysis, 10^5 cells were seeded in triplicate in 1 ml of medium with or without 100 nM dTAG47 in 24-well plates, counted and split back to 10^5 cells every 3 d for a total of 9 d. Total accumulated cells were calculated by assuming exponential growth kinetics.

For CellTiter-Glo (CTG, Promega) viability-based drug dose–response assays, we seeded 6×10^3 MED14–dTAG HCT-116 cells in 50 µl of RPMI into 96-well plates for ten data points in triplicate. Serial drug dilutions ($1 : \sqrt{10}$) were prepared at a 2× concentration in RPMI, starting from a 10 µM final concentration, and 50 µl was added to the 50-µl cell suspension. Cells were incubated at 37 °C for 3 d before the CTG assay was performed according to the manufacturer's instructions. Briefly, CTG substrate was diluted 1:4 with water and added 1:1 to cells before measuring chemical luminescence on an M5 plate reader. Relative viability was calculated by normalizing each data point to the mean luminescence of the lowest drug concentration.

KBM7 cells stably expressing Cas9 were generated with lentivirus produced from pLenti_Cas9_blast (Addgene, 52962). Lentivirus was produced by seeding 4×10^6 HEK293T cells into 10-cm dishes 24 h before transfection with 5 µg transfer plasmid, 2.5 µg pMD2.G (Addgene, 12259) and 3.75 µg psPAX2 (Addgene, 12260) using PolyFect (Qiagen) according to the manufacturer's instructions. The medium was changed to 8 ml fresh DMEM 8 h after transfection. Viral supernatant was collected after 72 h and filtered through a 0.22-µm syringe filter to remove cell debris. Viral supernatant was aliquotted and stored at −80 °C until further use. KBM7 cells (3×10^6 in 3 ml) were transduced with different amounts of viral supernatant in 12-well plates using 8 µg ml^{-1} polybrene and spin-infected for 45 min at 37 °C at 2,000 r.p.m. Antibiotic selection with 10 µg ml^{-1} blasticidin was started 24 h after transduction.

Plasmids, cloning and chemicals used in this study. All plasmids and chemicals used in this study are summarized in Supplementary Table 4. To facilitate cloning of sgRNA cutting plasmids for endogenous dTAG knock-in, we generated a universal pX330A_sgX_sgPITCh cutting plasmid by slightly adapting published protocols[70]. Briefly, the U6-sgPITCh cassette from pX330S-2-PITCh (Addgene, 63670) was cloned into pX330A_1-2x (Addgene, 58766) with BsaI. Targeting sgRNA sequences for endogenous knock-in (Supplementary Table 1) were then introduced under the control of the first U6 cassette of pX330A_sgX_sgPITCh with standard BbsI-mediated oligonucleotide-annealing cloning. Targeting sgRNA sequences were selected on the basis of location (as close as possible to the upstream start codon) and predicted sgRNA cleavage and off-target performance.

pCRIS-PITChv2 repair template plasmids were cloned according to published procedures[71]. Briefly, primers were selected to contain 20- to 22-bp microhomology sequences corresponding to genomic sequences immediately 5′ and 3′ of the sgRNA cleavage sites. Bases were added or removed to conserve the target reading frame, coding back the last cleaved amino acid whenever possible. The resulting Rep_X_F and Rep_X_R primers (Supplementary Table 1) were used to PCR amplify the dTAG-blast/puro cassettes from template plasmids (Addgene, 91792 and 91793), and the cassettes were reintroduced into the MluI-linearized pCRIS-PITChv2 backbone with NEBuilder 2× HiFi assembly (New England Biolabs).

For knock-in of mEGFP to be fused to the N terminus of POLR2A, we modified pCRIS-PITChv2 repair plasmids analogously to the dTAG-cassette-carrying templates (as in Addgene constructs 91792 and 91793). The coding sequence of mEGFP (A206K) was PCR amplified with primers adding sequences encoding N-terminal FLAG and C-terminal $(GGGGS)_2$ linker sequences. The resulting fragment encoding FLAG–mEGFP–SG was introduced into the MluI-linearized pCRIS-PITChv2 backbone with NEBuilder 2× HiFi assembly (New England Biolabs). Twenty-two-base-pair microhomology sequences immediately 5′ and 3′ to the cleavage site for an sgRNA targeting the N terminus of POLR2A (Supplementary Table 1) were then introduced with the same strategy as described for the dTAG cassettes.

Genome editing for endogenous knock-in. For endogenous knock-in of dTAG cassettes, KBM7 cells were seeded to 2×10^5 cells per ml the day before transfection to ensure exponential cell growth. The next day, cells were transfected with PITCh sgRNA/Cas9 and repair template plasmids by using the Amaxa Nucleofector II with Reagent V (Lonza)[70]. Briefly, 2×10^6 cells per reaction were collected by slow centrifugation at 90*g* and resuspended in 1 ml of warm antibiotic-free IMDM per reaction. Cells were aliquotted into 2-ml microcentrifuge tubes and spun down at 100*g* in a tabletop centrifuge. After aspirating the supernatant, cells were resuspended in 100 µl of supplemented Amaxa Reagent V mixed with 6 µg of DNA per plasmid (see Supplementary Table 1 for sgRNA and microhomology sequences). The cell and plasmid suspension was quickly transferred to the reaction cuvette, and electroporation was performed by using the X-001 program. Immediately after the reaction, 700 µl of warm antibiotic-free IMDM was added to the cuvette, and the cell suspension was carefully transferred to 1.5 ml of pre-equilibrated, warm antibiotic-free IMDM in six-well plates using the provided single-use Pasteur pipettes.

Cells were allowed to recover for 5 d at 37 °C before starting antibiotic selection of the pools in 10 ml of IMDM in T25 flasks (1 µg ml^{-1} puromycin or 10 µg ml^{-1} blasticidin, depending on the dTAG cassette used; Supplementary Table 1). After 7 d of selection, surviving cells were seeded into 384-well plates at concentrations of 0.2–1 cells per well (50 µl per well). Single-clone colonies were expanded to 24-well plates after 14 d and initially characterized through HA immunoblotting. HA-positive clones were genotyped by genomic DNA PCR of the integration site followed by Sanger sequencing (Microsynth; see Supplementary Table 1 for genotyping primer sequences and editing outcomes). One working clone and up to two backup clones were selected and kept for further experiments.

Owing to difficulties in obtaining MED12–dTAG clones, we followed a slightly modified targeting strategy for this construct: KBM7_Cas9 cells stably expressing Cas9 (pLenti_Cas9_blast; Addgene, 52962) were transfected with a truncated version of the pX330A_sgMED12_sgPITCh cutting plasmid lacking Cas9. In this case, transfected DNA amounted to 2 µg of the cutting plasmid lacking Cas9 and 10 µg of pCRIS-PITCh_v2_dTAG_puro repair plasmid. Clone selection and screening were performed as described above.

Endogenous knock-in of MED14–dTAG in HCT-116 colorectal carcinoma cells was performed with the same strategy as described above for KBM7 cells with minor adaptions. Briefly, 2×10^6 HCT-116 cells (~70% confluent) were trypsinized and transfected with 6 µg per PITCh sgRNA/Cas9 and repair template plasmid by using the Amaxa Nucleofector II with Reagent V and electroporation program D-032 (Lonza)[70]. Cells were then selected with 10 µg ml^{-1} blasticidin before single-clone expansion and characterization as described above for KBM7 cells.

Secondary knock-in of mEGFP–POLR2A into MED14–dTAG KBM7 cells was achieved by using the same strategy as for the parental MED14–dTAG knock-in cells with the following modifications: PITCh sgRNA/Cas9 and repair template plasmids were transiently transfected into MED14–dTAG KBM7 cells as described above but using a FLAG-mEGFP-SG repair template with microhomologies mapping to the N terminus of POLR2A (Supplementary Table 1). The pool of transfected cells was allowed to recover for 5 d before fluorescence-activated cell sorting of the top 0.15% of GFP-positive cells (Sony SH800 FACS cell sorter). Cells were expanded for 13 d before a second sort, followed by another 6 d of expansion and a third sort. Clones were isolated from GFP-positive pools by plating in limiting dilution to 96-well plates and subsequent clonal expansion. Successful mEGFP knock-in to the endogenous *POLR2A* locus was validated by PCR amplification and Sanger sequencing of the full-length genomic integration fragment (Microsynth; Supplementary Table 1).

Whole-cell lysis and immunoblotting. PBS-washed cell pellets were lysed in RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 1× Halt protease inhibitor cocktail, 25 U ml^{-1} Benzonase) for 15 min on ice. Lysates were cleared by centrifugation for 15 min at 4 °C and 20,000*g*. Protein concentration was measured by BCA before adding 4× LDS sample buffer. Protein (10–30 µg) was loaded on Bolt 4–12% Bis-Tris gels and blotted to nitrocellulose membranes. Membranes were blocked with 5% milk in TBST for 30 min at room temperature. Primary antibodies were added in milk or TBST alone for 1 h at room temperature or 4 °C overnight. Secondary antibodies were added for 1 h at room temperature. Blots were developed with chemiluminescence films for HRP-labeled antibodies or a ChemiDoc fluorescence imaging system for fluorescently labeled antibodies. Where applicable, western blots were quantified with ImageJ software (US National Institutes of Health).

The antibodies used (Supplementary Table 5) were to the following proteins: HA (1:1,000; Cell Signaling Technology, 3724S and 2367S), GAPDH (1:1,000; Santa Cruz Biotechnology, sc-365062), histone H3 (1:10,000; Abcam, ab1791), β-actin (1:10,000; Sigma-Aldrich, A5441), MYC (1:1,000; Santa Cruz Biotechnology, sc-764), MED10 (1:500; Abcam, ab110786), CDK9 (1:1,000; Cell Signaling Technology, 2316S), BRD4 (1:5,000; Bethyl Laboratories, A301-985A100), MED12 (1:2,000; Bethyl Laboratories, A300-774A), MED1 (1:2,000; Bethyl Laboratories, A300-793A), MED14 (1:1,000; Bethyl Laboratories, A301-044A), MED26 (1:1,000; Cell Signaling Technology, 14950S), MED6 (1:1,000; Santa Cruz Biotechnology, sc-390474), MED23 (1:1,000; Bethyl Laboratories, A300-425A), BRD9 (1:1,000; Active Motif, 61537), CDK8 (1:1,000; Cell Signaling Technology, 4101S), H3K27ac (1:1,000; Abcam, ab4729), G3BP1 (1:500; Santa Cruz Biotechnology, sc-365338), α-tubulin (1:1,000; Sigma-Aldrich, T9026), CCNT1 (1:1,000; Cell Signaling Technology, 81464S), HEXIM1 (1:1,000; Abcam, ab25388), MED31 (1:500; Santa Cruz Biotechnology, sc-101189), Pol II phosphorylated at Ser2 (1:1,000; Active Motif, 61083), Pol II (1:250; Santa Cruz Biotechnology, sc-899), SPT5 phosphorylated at Thr806 (1:1,000; laboratory of Robert Fisher, Icahn School of Medicine at Mount Sinai), SPT5 (1:1,000; Santa Cruz Biotechnology, sc-133217) and tubulin (rhodamine conjugated, 1:5,000; Bio-Rad, 12004166).

Secondary antibodies were as follows: anti-mouse-HRP (1:5,000; Jackson ImmunoResearch, 115-035-003), anti-rabbit-HRP (1:5,000;

Jackson ImmunoResearch, 111-035-003), anti-rat-HRP (1:5,000;, Dako, P0450), anti-rabbit-StarBright700 (1:5,000; Bio-Rad, 12004162) and anti-mouse-Dylight800 (1:5,000; Bio-Rad, STAR117D800GA).

Quant-seq 3′ mRNA transcriptional fingerprinting. Spike-in normalized Quant-seq (Lexogen) was performed in biological triplicate to record a transcriptional fingerprint of acute Mediator perturbation in the context of chemical (genetic) perturbations with known transcriptional consequences. Briefly, eight MED–dTAG (MED12, MED1, MED14, MED26, MED6, MED28, MED31 or MED10), CDK9–dTAG and wild-type KBM7 cells were seeded at 5×10^5 cells per ml in 10-ml T25 flasks 2 d before the experiment to avoid major differences in cell cycle and growth kinetics between different clones. On the day of the experiment, cells were counted and 8×10^5 cells per sample were seeded into 24-well plates in 1 ml for triplicate drug treatments. Drug treatments from 1,000× DMSO stocks were performed for 6 h with the following conditions: 500 nM dTAG7 or DMSO for dTAG-carrying clones and 500 nM dTAG7, 100 nM dBET6, 250 nM THAL-SNS-032 (dCDK9) or DMSO for wild-type KBM7 cells. Cell pellets were collected by centrifugation, washed with cold PBS, snap-frozen and stored at −80 °C.

The next day, cell pellets were thawed on ice and total RNA was extracted with RNeasy kits (Qiagen) after addition of SIRV-Set 3 exogenous spike-in mix (Lexogen). Briefly, 350 µl of buffer RLT was added to the pellets and the mix was centrifuged through QIAshredder spin columns according to the manufacturer's instructions. Then, 10 µl of a 1:100 dilution (0.303 ng per sample; according to the manufacturer's protocol) SIRV-Set 3 spike-in mix was added to the homogenized lysate before addition of 350 µl ethanol and column purification according to the manufacturer's instructions. Total RNA was eluted in 30 µl of elution buffer and stored at −80 °C overnight.

The next morning, RNA was thawed on ice and the concentration was measured to 71–102 ng µl^{-1} with a NanoDrop. All samples were then diluted to 70 ng µl^{-1} in 20 µl, and Quant-seq was performed in a 96-well PCR plate using 5 µl (350 ng RNA) as input. Samples were randomly distributed over the plate to avoid edge effects. Quant-seq libraries were performed by carefully following the manufacturer's instructions. Briefly, first-strand (oligo(dT)) cDNA synthesis was followed by RNA removal and second-strand synthesis via random priming. The double-stranded library was bead purified to remove reaction components before PCR amplification with i7 single-index primers for 13 cycles. Amplified libraries were again bead purified according to the manufacturer's protocol, yielding 0.67–3.16 ng µl^{-1} DNA as measured by Qubit assay. A total of 11 representative low-, intermediate- and high-yield samples were checked for fragment size distribution on a Bioanalyzer before pooling for 50-bp single-read sequencing on a HiSeq 3000 or 4000 platform to obtain 2.4–9.6 million raw reads per sample (Supplementary Table 6).

Quant-seq raw data were processed essentially as recommended by the manufacturer of the library preparation kit (Lexogen). Raw reads were trimmed with BBtools version 38.00 'bbduk.sh -ref 'polyA.fa.gz', 'truseq.fa.gz' k=13 ktrim=r useshortkmers=t mink=5 qtrim=r trimq=10 minlength=20' and aligned to a concatenated hg38_SIRV-Set-3 (sequences provided by Lexogen) index using STAR version 2.5.2b with '-outFilterType BySJout -outFilterMultimapNmax 20 -alignSJoverhangMin 8 -alignSJDBoverhangMin 1 -outFilterMismatchNmax 999 -outFilterMismatchNoverLmax 0.6 -alignIntronMin 20 -alignIntronMax 1000000 -alignMatesGapMax 1000000 -outSAMattributes NH HI NM MD -outSAMtype BAM SortedByCoordinate'[72]. Aligned reads were then counted using htslib version 0.6.0 'htseq-count -m intersection-nonempty -s yes -f bam -r pos' on a fused hg38_SIRV-Set-3.gtf ENSEMBL e87 geneID annotation[73]. Raw reads were normalized to counts per million (cpm).

Quant-seq data analysis (PCA, spike-in normalization, differential expression). PCA was performed on cpm values with the scikit-learn Python package (v0.19.0)[74]. First, read counts were normalized by using the StandardScaler.fit_transform function with standard settings. The first ten principal components were calculated by using the decomposition.PCA function. Per-gene principal-component loadings were computed by multiplication of the eigenvectors with the square root of the principal-component eigenvalues. Gene set enrichment analysis (GSEA) of PC2 loadings was performed with GenePattern's GSEAPreranked with collections KEGG_pathway, GO_process and TF_Targets[75]. The top two gene sets containing at least 75 genes were shown. We additionally tested a published set of 100 high-confidence direct MYC target genes[40].

Locally estimated scatterplot smoothing (LOESS) normalization was performed for spike-in normalization by using the normalize.loess function of the affy v1.50.0 R package[76]. The LOESS distribution was fit on the SIRV-Set-3 spike-ins and then applied to the whole dataset. Average $\log_2$-transformed fold change values were calculated from the means of triplicate treatments, omitting genes with normalized $\log_2$(cpm) < 5.0 on average in wild-type cells treated with DMSO.

To identify the target gene spectrum of individual Mediator subunits, we performed DESeq2 v1.23.0 differential expression analysis on normalized counts relative to wild-type DMSO control[77]. The union of significantly deregulated genes ($\log_2$(FC) >1.5 or <−1.5; $P<0.05$) was hierarchically clustered with correlation as the distance metric and the average as linkage method. The clustered list of deregulated genes was used as the coordinates for a long 'target genes scaffold' in the R BioCircos library[78]. Links connect dummy 'MED–dTAG perturbation scaffolds' with the respective coordinate in the target gene scaffold if the gene was deregulated by the given perturbation. A secondary heat map outside the target gene scaffold indicates genes that were deregulated by at least three MED–dTAG perturbations.

Size-exclusion chromatography to assess Mediator complex integrity. Size-exclusion chromatography of Mediator-containing nuclear extracts was performed similarly to published procedures[79]. MED14–dTAG KBM7 cells (100×10^6) were washed with PBS and transferred to 2-ml tubes. Cytoplasmic lysis was performed by resuspending with 500 µl cold buffer A (10 mM Tris-HCl pH 7.4, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 0.1% Triton X-100, 1× Halt protease inhibitor cocktail) and incubating for 10 min on ice. Nuclei were collected by centrifugation at 4 °C and 1,000*g* for 5 min, and the pellet was washed once with 1 ml buffer A. High-salt nuclear extractions were performed by adding 45 µl of buffer C (20 mM Tris-HCl pH 7.4, 400 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 5 U µl^{-1} Benzonase, 1× Halt protease inhibitor cocktail) and incubating for 1 h on ice with occasional vortexing. The lysate was clarified by centrifugation for 15 min at 4 °C and 20,000*g*, yielding about 75 µl of nuclear extract with 5–10 µg ml^{-1} total protein. Potential particles were removed by centrifugation through 0.22-µm spin filter columns before injection into the HPLC system.

Nuclear extracts were loaded onto an Äkta FPLC system via a 50-µl injection loop and separated with SEC running buffer (20 mM Tris-HCl pH 7.4, 200 mM KCl) over a Superose 6 Increase 3.2/300 mm gel filtration column (GE Lifesciences). Run parameters were set to 0.1 column volumes (CVs) of initial equilibration, followed by sample injection and 1.5 CV elution at a constant flow rate of 0.04 ml min^{-1} (resulting in 0.7–0.8 MPa run pressure). Elution fractions of 150 µl were collected in 96-deep-well plates over the entire elution range. Eluates were directly diluted with 4× LDS sample buffer, and 20 µl per fraction (or 1–2 µl of input nuclear extract) was loaded onto 4–12% Bis-Tris gels for western blot analysis.

Immunofluorescence microscopy of Mediator and RNA Pol II. MED14–dTAG KBM7 cells (1×10^6) were spun onto poly(lysine)-precoated (30 min at 37 °C; washed three times with water) coverslips in 24-well plates for 15 min at 1,800 r.p.m. Cells were fixed by directly adding an equal volume of 4% methanol-free paraformaldehyde (final concentration of 2%) and incubating for 10 min at room temperature. Coverslips were washed three times with PBS and then incubated for 5 min in 200 µl per well of permeabilization buffer (0.1% non-neutralized dibasic sodium citrate, 0.2% Triton X-100). After again washing three times with PBS, coverslips were incubated for 30 min in 200 µl per well of blocking buffer (2.5% BSA in PBS). Primary antibodies were added directly to the blocking solution at a dilution of 1:200, and plates were incubated at 4 °C overnight. The next morning, coverslips were washed three times with PBS and incubated for 1–2 h with a 1:250 dilution of secondary antibody at room temperature. Coverslips were again washed three times with PBS and labeled with DAPI for 5 min before mounting onto glass slides with Dako mounting medium.

The same procedure was followed for combined fixed-cell imaging of mEGFP–POLR2A (direct fluorescence) and MED14–dTAG (immunofluorescence; detected with anti-rabbit-AF594) KBM7 cells, including a negative-control condition where anti-HA primary antibody was omitted. For visualization of representative cells, AF594 signal was scaled to allow quantitative assessment of MED14–dTAG degradation, while mEGFP signal was scaled to resolve focal structures or the lack thereof.

Images were acquired as 3D *z* stacks on an LSM780 confocal fluorescence microscope in AiryScan mode (Zeiss) by using an ×63 oil objective. Maximum-intensity projections were derived with Zeiss ZEN Blue software and segmented for image quantification with a custom CellProfiler pipeline[80]. Large Mediator and Pol II foci of around 500 nm in diameter were segregated and classified using area- and intensity-based thresholds to quantify mutual overlap and number of large foci per cell.

Primary antibodies were used to the following proteins: HA (Cell Signaling Technology, 3724 and 2367), MED1 (Abcam, ab64965) and Pol II 8WG16 (Abcam, ab817). Secondary antibodies were as follows: anti-rabbit-AF594 (Invitrogen, A-21442) and anti-mouse-AF488 (Invitrogen, A-11001).

Mediator subunit disorder prediction. Sequence-intrinsic disorder propensity was predicted for all human Mediator subunits. Canonical isoform FASTA sequences were obtained from UniProt and subjected to PONDR-VSL2 disorder prediction software (http://www.pondr.com/)[81]. Results were compiled into a matrix by residue number and smoothed with a running mean window of 20 residues. Only high-confidence (VSL2 > 0.9) disordered regions were highlighted in a heat map.

Biotinylated isoxazole precipitation of low-complexity-region-containing proteins. Washed MED14–dTAG KBM7 cells (75×10^6) were resuspended in 1.5 ml of EE buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 10% glycerol, 0.1% Triton X-100, 1× Halt protease inhibitor

cocktail) and incubated for 30 min at 4 °C on a rotating wheel. Lysates were clarified by centrifugation for 15 min at 4 °C and 20,000*g*. Lysate (75 µl) was removed as input, and the rest was divided into three 400-µl aliquots in fresh 1.5-ml tubes. Biotinylated isoxazole from 100× DMSO stocks (Dalton Pharma Services) was added to the lysates to final concentrations of 0–100 µM, and the tubes were incubated for 2 h at 4 °C on a rotating wheel. Pellets were collected by centrifugation for 15 min at 4 °C and 20,000*g*, supernatant was removed and kept, and the pellet was washed once with 500 µl of EE buffer. The resulting pellets were taken up in 75 µl of 1× LDS in EE buffer and heated for 10 min at 95 °C before loading 15 µl per sample (input, supernatant, b-isox pellet) for western blot analysis.

Statistics and reproducibility. Unpaired two-tailed Student's *t* tests were performed for Fig. 3c and Extended Data Figs. 2c and 5a,c. The Mann–Whitney *U* test in Fig. 4c was performed by using the SciPy v1.2.1 mannwhitneyu function with default settings (*P* = half of the two-sided *P* value)[82]. Skewness in Fig. 4h was tested by using the SciPy v1.2.1 skewtest function with default settings (two-sided *P* value), which tests the null hypothesis that the skewness of the given population is the same as that of a corresponding normal distribution[82].

Box plots were generated by using the Seaborn v.0.9.0 or Matplotlib v.2.2.3 boxplot function with standard settings: center lines display median values and boxes display the quartile 3–quartile 1 interquartile range (IQR), with whiskers protruding 1.5 times the IQR. When included, notches display the confidence region around the median, derived with standard settings[83,84]. Violin plots were generated by using the Seaborn v0.9.0 violinplot function with standard settings, which displays an approximate density distribution of the underlying data points, as well as an internal box plot with the above-mentioned elements[83].

The results in Fig. 1b and Extended Data Figs. 1b,g and 4d are representative of two independently conducted cellular treatments and immunoblots. The results in Fig. 4c are representative of two independently conducted cellular treatments, immunoprecipitations and immunoblots. The results in Fig. 4d are representative of three independently loaded immunoblots from a single chromatin fractionation experiment. The results in Extended Data Figs. 1c and 3b are from a single cellular treatment and immunoblot experiment. The results in Extended Data Fig. 2b are from a single size-exclusion chromatography experiment. The results in Extended Data Fig. 2e are representative of three independent biotinylated isoxazole precipitation experiments. The maximum-intensity projection in Extended Data Fig. 5b is representative of nine nuclei acquired in one field of a single staining replicate and was performed in parallel to the experiment in Fig. 3c. Extended Data Fig. 6d is representative of two independently loaded immunoblots from a single cellular treatment and nuclear fractionation experiment.

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

Data availability

Next-generation sequencing data are available through the NCBI Gene Expression Omnibus under accession code GSE139468. Chromatin proteomics data have been deposited at PRIDE under dataset identifier PXD017611. Source data for Figs. 1 and 4 and Extended Data Figs. 1–4 and 6 are presented with the paper.

Code availability

Custom code used to analyze the data in this study is available at https://github.com/GWinterLab/Jaeger_Mediator_NatureGenetics_2020.

References

70. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I. T. & Yamamoto, T. MMEJ-assisted gene knock-in using TALENs and CRISPR–Cas9 with the PITCh systems. *Nat. Protoc.* **11**, 118–133 (2016).
71. Brand, M. & Winter, G. E. Locus-specific knock-in of a degradable tag for target validation studies. *Methods Mol. Biol.* **1953**, 105–119 (2019).
72. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
73. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
74. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
75. Reich, M. et al. GenePattern 2.0. *Nat. Genet.* **38**, 500–501 (2006).
76. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
77. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
78. Cui, Y. et al. BioCircos.js: an interactive Circos JavaScript library for biological data visualization on web applications. *Bioinformatics* **32**, 1740–1742 (2016).
79. Aranda-Orgilles, B. et al. MED12 regulates HSC-specific enhancers independently of Mediator kinase activity to control hematopoiesis. *Cell Stem Cell* **19**, 784–799 (2016).
80. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
81. Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K. & Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* **7**, 208 (2006).
82. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
83. Waskom, M. et al. mwaskom/seaborn: v0.9.0. *Zenodo* https://doi.org/10.5281/ZENODO.1313201 (2018).
84. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

Acknowledgements

We thank R. Fisher (Icahn School of Medicine at Mount Sinai) for sharing antibody to SPT5 phosphorylated at Thr806. We thank the Biomedical Sequencing Facility at CeMM and the MPIMG sequencing core for assistance with next-generation sequencing. We thank the imaging core facility of the Medical University of Vienna for assistance with microscopy. We thank P. Lenart for critical review of the image quantification procedures. We thank A. Mayer and M. Erb for feedback on this manuscript. M.G.J. was supported by a Boehringer Ingelheim Fonds PhD fellowship. T.V. was supported by the International Max Planck Research School for Genome Science, part of the Göttingen Graduate Center for Neurosciences, Biophysics and Molecular Biosciences. B.A. is supported by the Austrian Science Fund (FWF) and the Medical University of Vienna's joint PhD program in Inflammation and Immunity (FWF1212). C.B. is supported by a New Frontiers Group award of the Austrian Academy of Sciences and by an ERC Starting Grant (European Union's Horizon 2020 research and innovation programme, grant agreement 679146). B.N. was supported by an American Cancer Society Postdoctoral Fellowship (PF-17-010-01-CDD). B.N. and N.S.G. were supported by the Katherine L. and Steven C. Pinard Research Fund. D.H. is supported by the SPP2202 Priority Program Grant (HN 4/1-1) from the Deutsche Forschungsgemeinschaft (DFG). This project was further supported by an FWF Stand-Alone grant (P31690-B) awarded to the Winter laboratory.

Author contributions

M.G.J. and G.E.W. conceptualized this project. M.G.J., T.V., A.H., M.B., B.A. and D.H. designed and conducted experiments. M.G.J., B.S., S.D.M., A.H., H.I. and M.B. analyzed and interpreted original and published omics data. M.G.J., M.B., B.A. and S.C. generated cell lines. M.G.J., B.S., S.D.M. and M.B. visualized data. B.N. and F.M.F. synthesized the dTAG^V-1 reagent. A.M., A.B., J.E.B., N.S.G., C.B., D.H., P.C. and G.E.W. supervised the work. M.G.J. and G.E.W. wrote the manuscript with input from all authors.

Competing interests

G.E.W., J.E.B. and B.N. are inventors on patent applications related to the dTAG system (WO/2017/024318, WO/2017/024319, WO/2018/148443, WO/2018/148440). The dTAG^V-1 molecule is the subject of a patent application filed by Dana-Farber Cancer Institute. N.S.G. is a scientific founder, member of the scientific advisory board (SAB) and equity holder for C4 Therapeutics, Syros, Soltego, B2S, Gatekeeper and Petra Pharmaceuticals. The Gray laboratory receives or has received research funding from Novartis, Takeda, Astellas, Taiho, Janssen, Kinogen, Voroni, Her2llc, Deerfield and Sanofi. J.E.B. is now an executive and shareholder of Novartis and has been a founder and shareholder of SHAPE (acquired by Medivir), Acetylon (acquired by Celgene), Tensha (acquired by Roche), Syros, Regenacy and C4 Therapeutics.

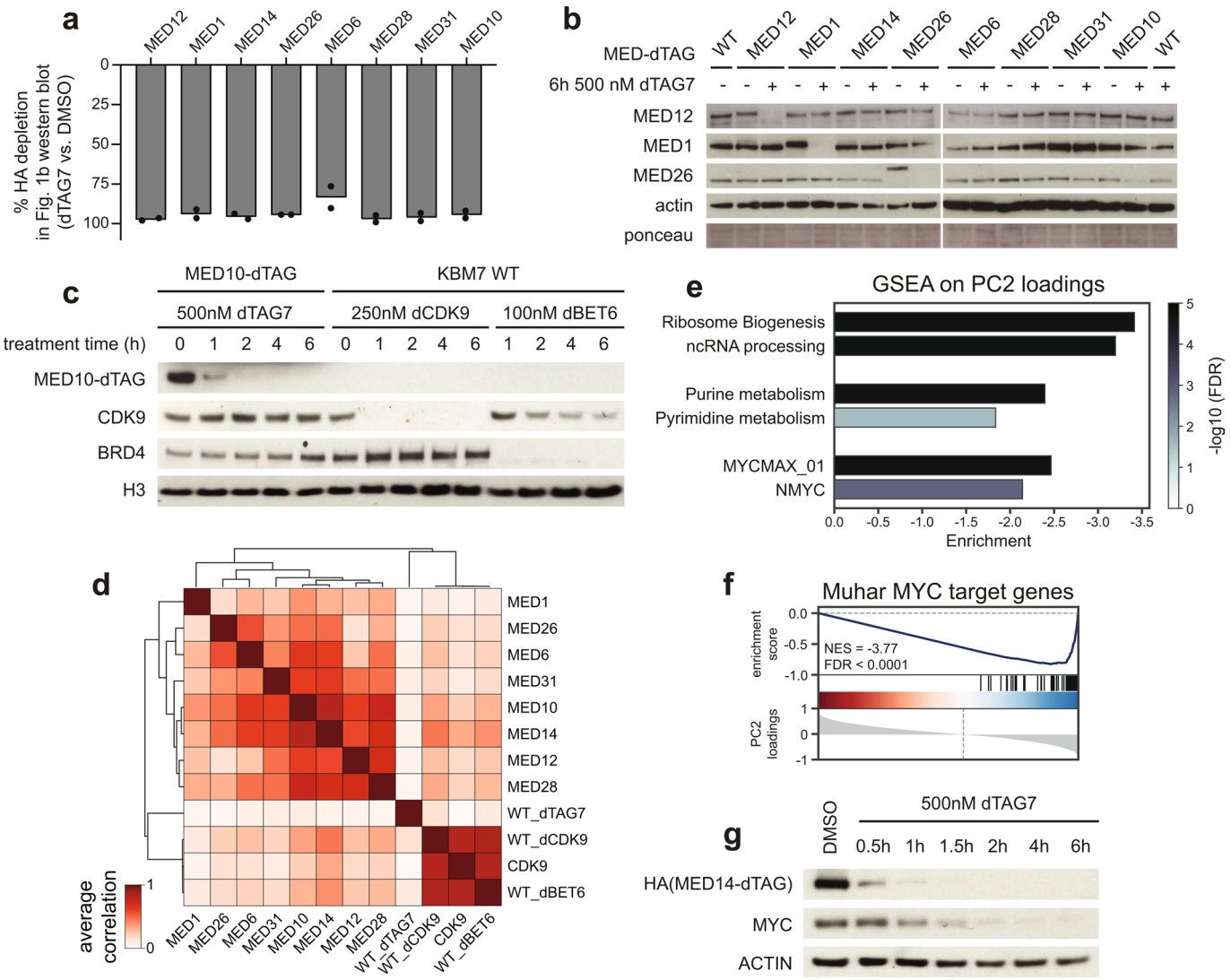
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-020-0635-0.

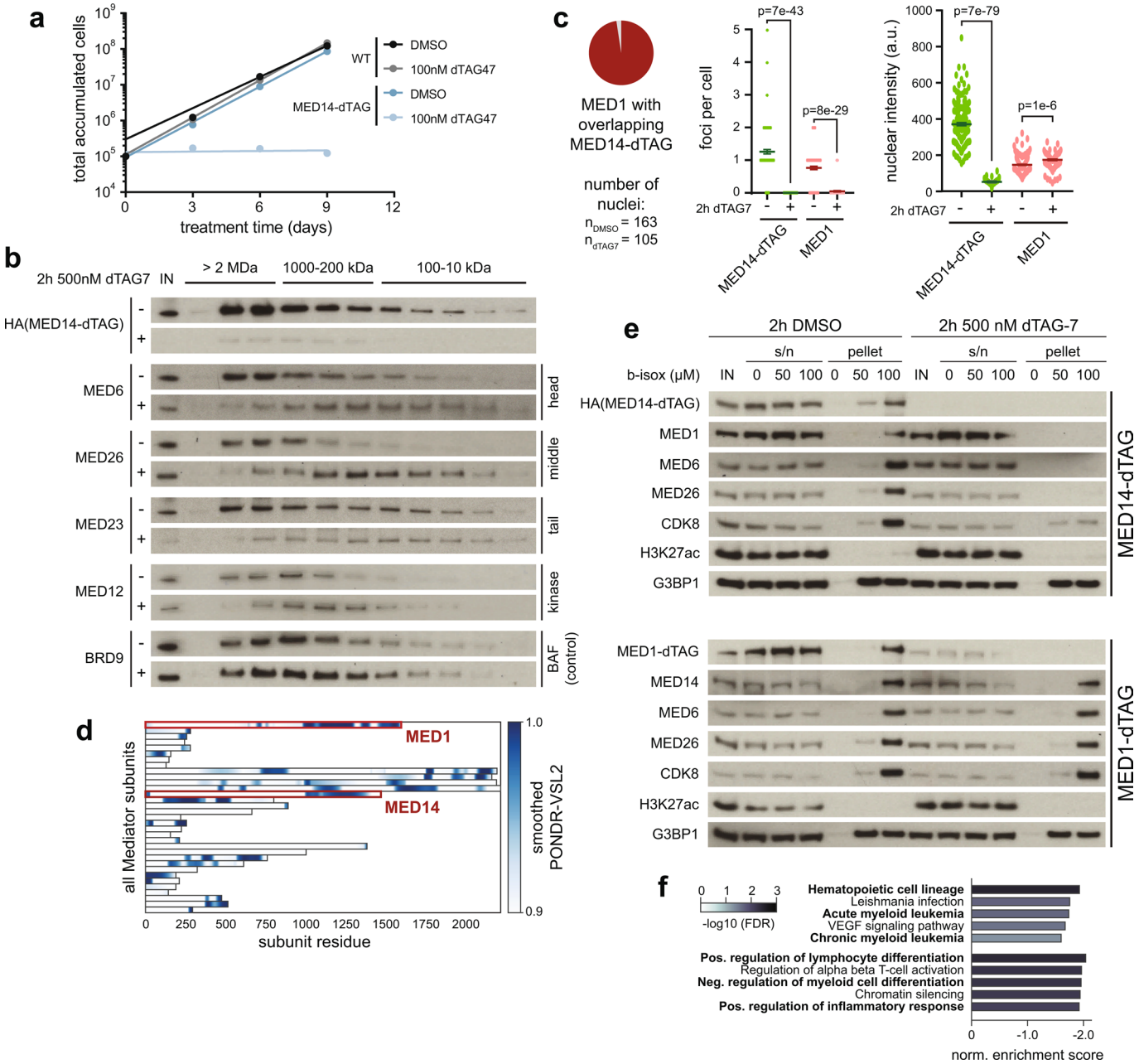
Supplementary information is available for this paper at https://doi.org/10.1038/s41588-020-0635-0.

Correspondence and requests for materials should be addressed to P.C. or G.E.W.

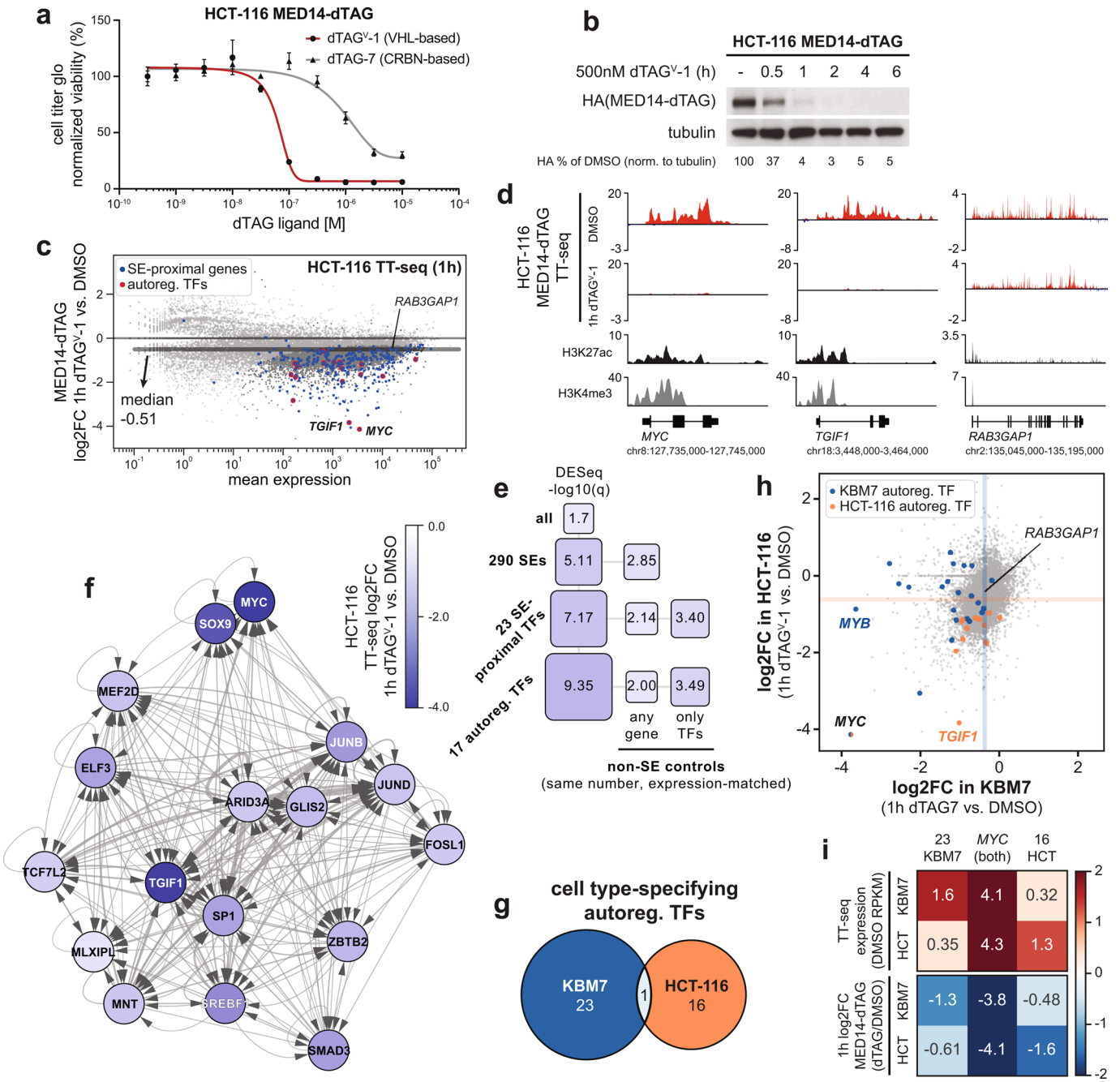
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | Extended characterization of chemically degradable MED-dTAG alleles. a, MED-dTAG depletion mean of two independent image quantifications of the Fig. 1b immunoblot. b, Degrader treatment selectively destabilizes the tagged Mediator subunit without affecting other complex members. c, Time-resolved immunoblot of MED10-dTAG and direct pharmacologic degradation of CDK9 (dCDK9; THAL-SNS-032) or BRD4 (dBET6). d, Pearson correlation of average 3′ mRNA-seq log2 fold changes after 6 h (n = 3 independent drug treatments). For dTAG-carrying cell lines (only gene names shown), we compare dTAG7 vs. vehicle control in the same cell line. Other conditions represent drug vs. vehicle control in wild-type cells. e, Gene ontology (GO) terms enriched among negative PC2 loadings in Fig. 1c. Enrichment was calculated using the GSEAPreranked tool[75]. Negative enrichment indicates a strong influence of these terms on PC2 diversity and that the underlying genes are downregulated. f, Gene set enrichment analysis of top 100 core MYC target genes from (ref. [40]) among PC2 loadings. g, Time-resolved immunoblot of MED14-dTAG degradation kinetics and its influence on MYC protein levels. Unprocessed western blots shown in Source Data.

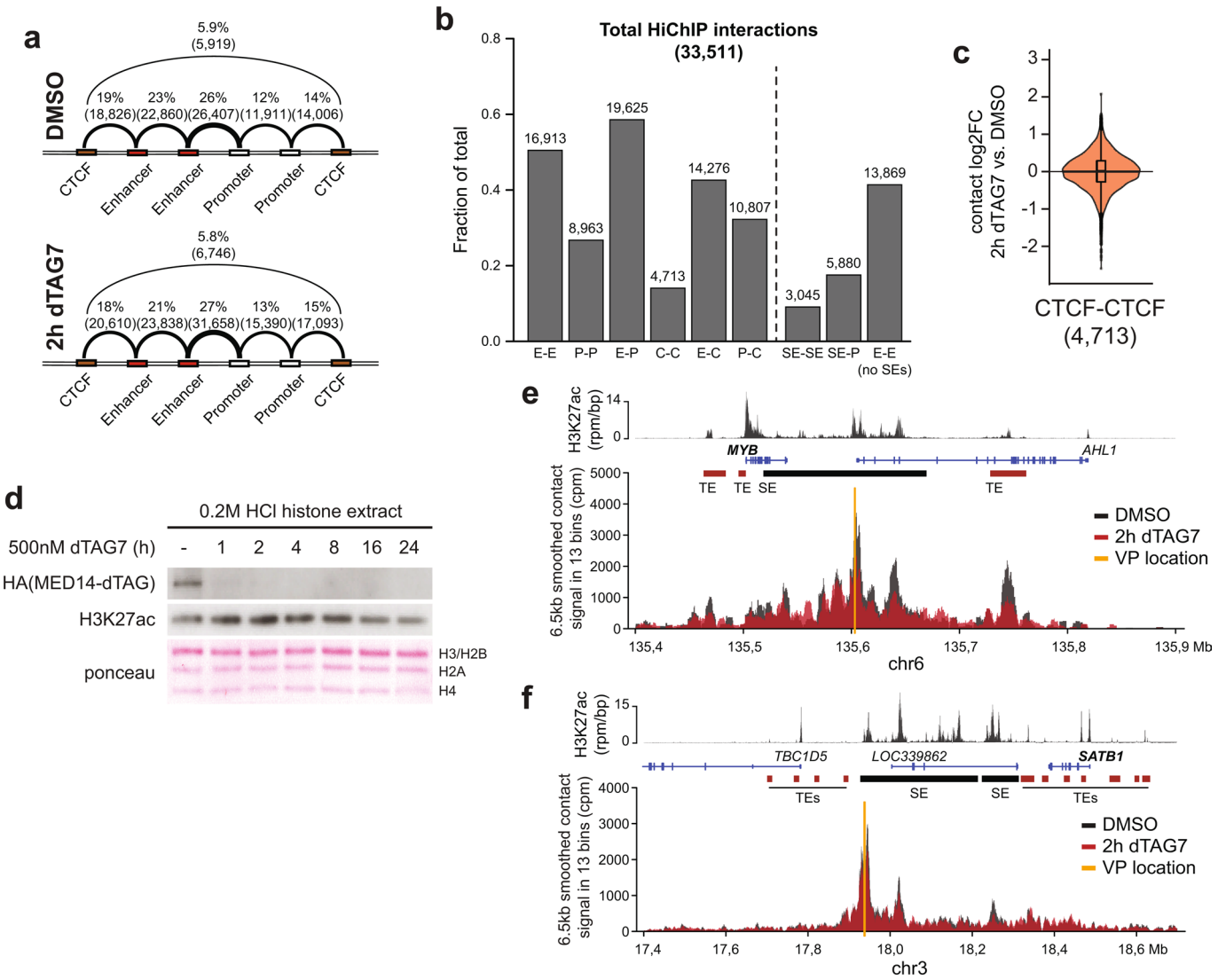


Extended Data Fig. 2 | MED14 degradation disrupts overall Mediator complex integrity. a, Influence of long-term MED14 degradation on cell growth. **b**, Size-exclusion chromatography and western blotting of nuclear extracts after MED14 degradation. Mediator subunits of each submodule shifted to lower apparent molecular weight, indicating complex disassembly. BAF complex member BRD9 serves as negative control. **c**, Image quantification related to Fig. 1f. Pie chart: percent of n = 125 MED1 foci with overlapping MED14-dTAG foci. Middle dot plot: mean±s.e.m number of foci per cell. Swarm plot: mean±s.e.m integrated nuclear fluorescence intensity. Unpaired, two-sided t-tests. 163 nuclei were quantified for DMSO and 105 for dTAG7. **d**, 20-residue running average-smoothed PONDR-VSL2 disorder prediction for human Mediator. Subunits are ordered by ascending index numbers from MED1 to MED31, followed by CDK8, CDK19, and CCNC. **e**, Influence of MED14 or MED1 degradation on co-precipitation of other Mediator subunits with biotinylated isoxazole pellets. MED14, but not MED1 degradation, prevents Mediator co-precipitation with IDR-enriching hydrogels. **f**, Cell identity gene sets enriched among downregulated transcription units in TT-seq after 1h MED14 degradation. Enrichment was calculated using the GSEAPreranked tool[75]. Unprocessed western blots shown in Source Data.

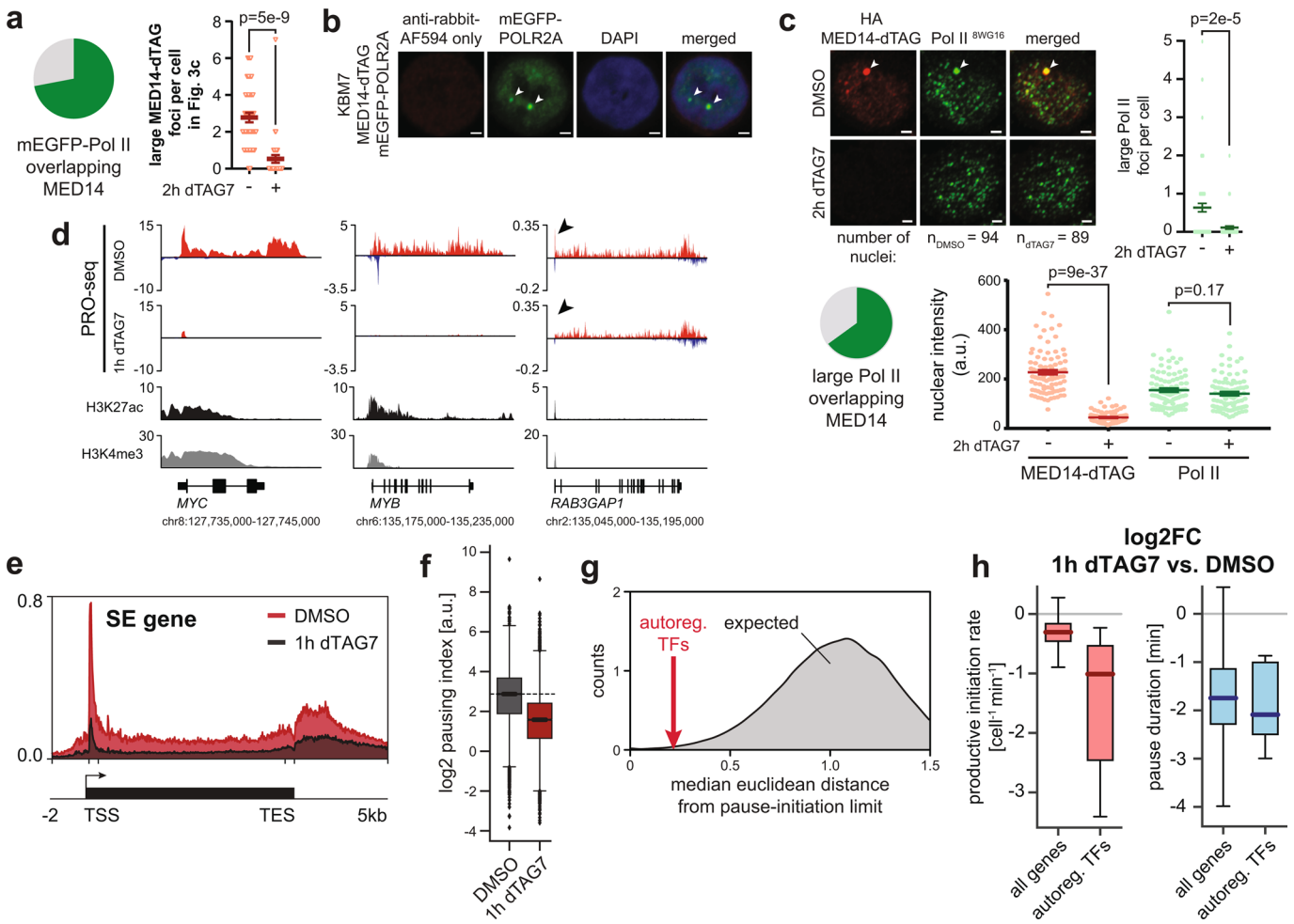


Extended Data Fig. 3 | Acute transcriptional consequences of MED14 degradation in HCT-116 cells. a, CellTiter-Glo viability-based 72 h dose-response of dTAG7 and dTAG^V-1 in HCT-116 MED14-dTAG cells. Mean±s.d. of n = 3 drug treatments. **b**, Time-resolved immunoblot of MED14-dTAG degradation. **c**, Differences in TT-seq nascent transcript levels (n = 2 independent treatments). Significantly deregulated (DESeq2 q < 0.01; dark grey), SE-proximal (blue), and auto-regulatory TF genes (red) are highlighted. Dark grey line: median log2 fold change of all n = 21,629 transcription units. **d**, TT-seq signal of two auto-regulatory TFs, and an expression-matched control gene. H3K27ac and H3K4me3 ChIP-seq signals are from publically available data (GSE72622; see Supplementary Table 7)[85]. **e**, Fold-change (color) and significance (size) of SE-driven HCT-116 cell identity and expression-matched control gene sets (data as in **c**). **f**, Regulatory wiring of 17 auto-regulatory TFs in the HCT-116 cell type-specifying gene regulatory network. Arrows: the given TF has binding motifs in the target TF's SE region(s). Edge weight mirrors number of motifs. **g**, Overlap of KBM7 and HCT-116 auto-regulatory TFs. **h**, Cell type-specific impact of 1 h MED14 degradation. Auto-regulatory TFs in KBM7 (blue, for example *MYB*), HCT-116 (orange, for example *TGIF1*), or *MYC* (black) as the only shared TF are highlighted. Colored lines: median log2FC in the respective cell line. **i**, Mean steady state expression of auto-regulatory TFs in merged 1 h and 2 h DMSO TT-seq conditions and transcriptional defects after 1 h MED14 degradation. Unprocessed western blot shown in Source Data.

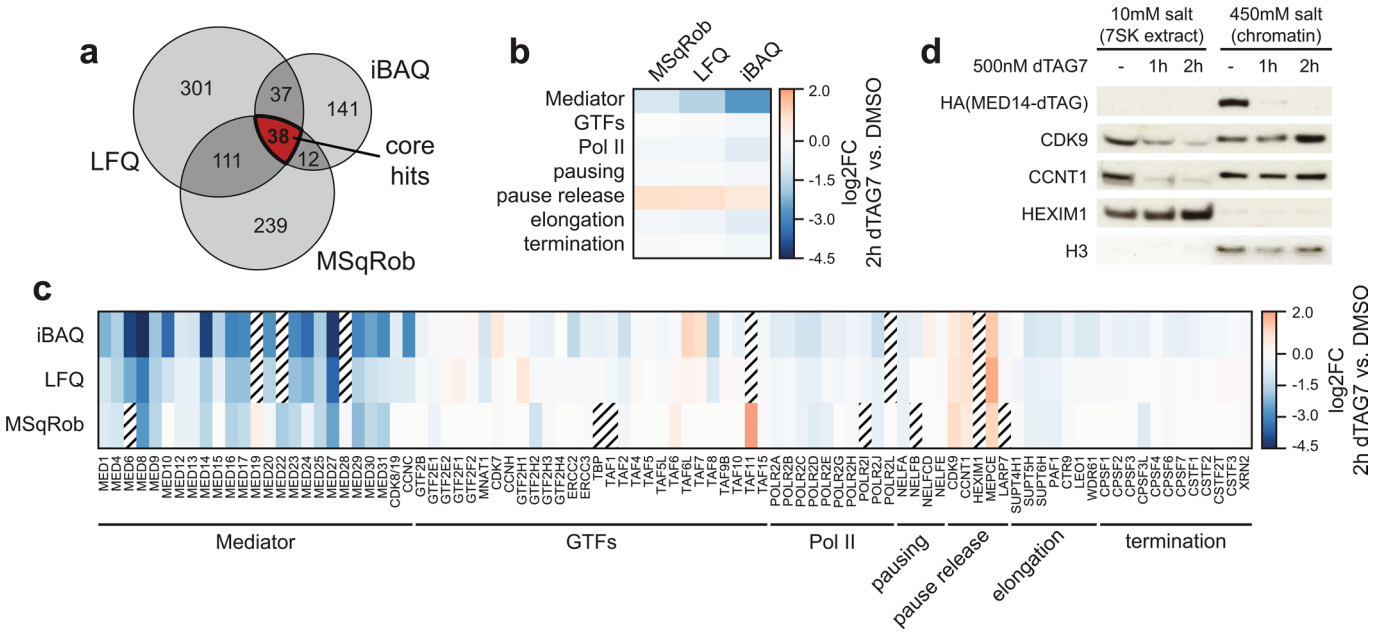
85. McNamara, R. P. et al. KAP1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. *Mol. Cell* **61**, 39–53 (2016).



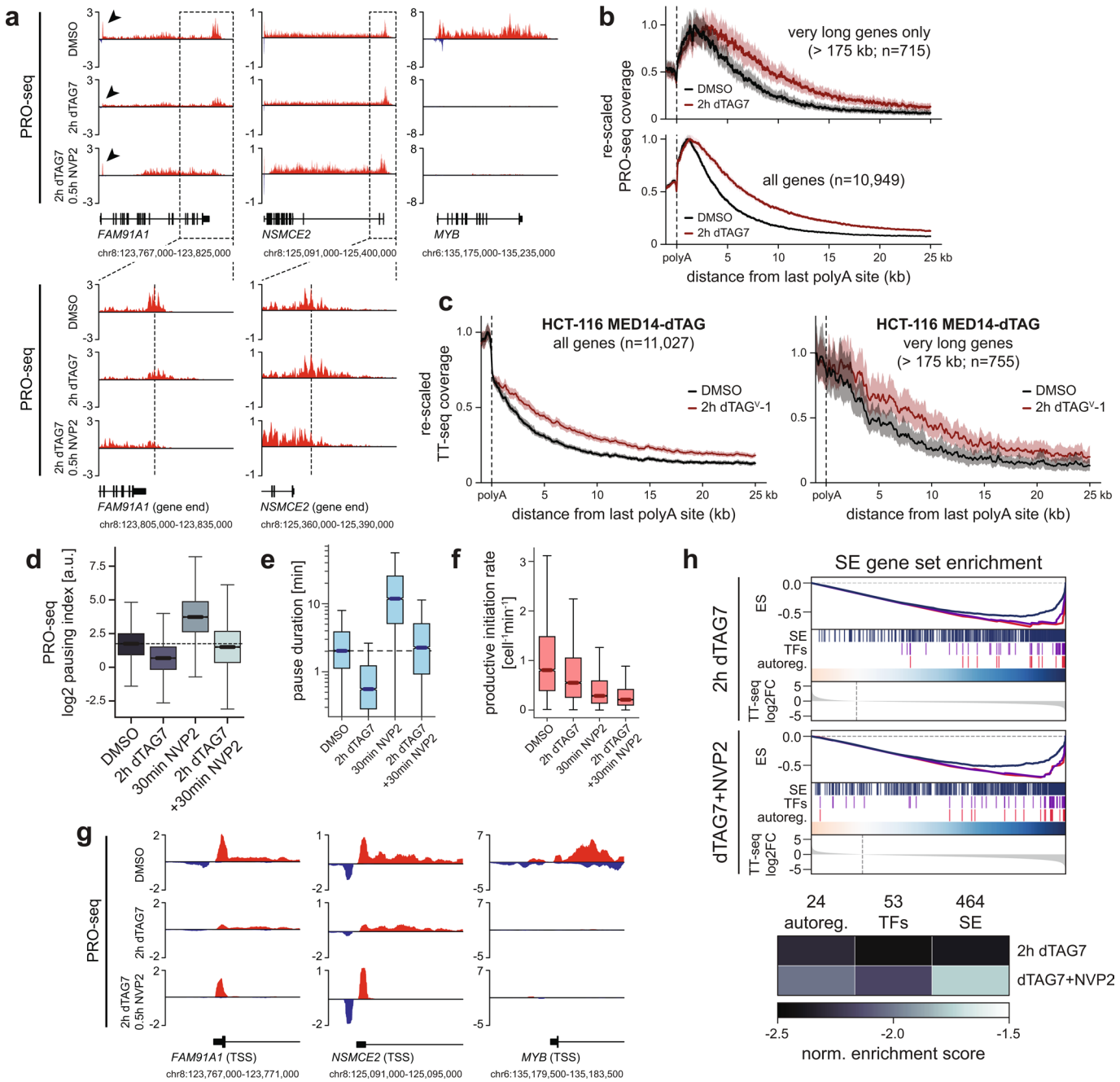
Extended Data Fig. 4 | Impact of MED14 degradation on overall chromatin architecture. a, Genomic feature classes at H3K27ac HiChIP contact anchors. Only significant interactions called by hichipper/mango were used for anchor identification. Arcs indicate the percentage of anchor-anchor pairs annotated with the indicated feature in each of the samples. **b**, Total number of interactions common to DMSO and dTAG7 samples, which were used for quantification (E: enhancer, P: promoter, SE: constituent). **c**, Impact of Mediator loss on CTCF-CTCF contact strength as negative control. Bracket: number of quantified contacts. Violin plot elements: approximated density distribution with internal box plots showing medians with interquartile range and 1.5x whiskers. **d**, Impact of MED14 degradation on H3K27 acetylation. **e**, Pulldown-independent 4C-seq analysis of MYB SE constituent viewpoint (VP) after 2 h MED14 degradation in triplicates. Top track shows KBM7 wild-type H3K27ac ChIP-seq. TE: typical enhancer, SE: super-enhancer **f**, Analogous to (**e**) with a SATB1 SE viewpoint. Unprocessed western blot shown in Source Data.



Extended Data Fig. 5 | Impact of MED14 degradation on Pol II clusters and nascent transcription dynamics. a, Image quantification related to Fig. 3c. Pie chart: percent of n = 100 large Pol II foci, which overlap MED14-dTAG foci. Mean±s.e.m. with two-sided, unpaired t-test (n = 40 nuclei in DMSO; n = 36 nuclei in dTAG7 condition). **b**, Control imaging experiment related to Fig. 3c, omitting anti-HA primary antibody to rule out that Pol II foci are an HA channel bleed through artifact. **c**, Immunofluorescence of large hypo-phosphorylated Pol II foci (8WG16; arrows) in MED14-dTAG KBM7 cells. Maximum intensity projections of 3D images. Scale bars 1 µm. Pie chart: percent of n = 60 large Pol II foci, which overlap MED14-dTAG foci. Dot plots: changes in number of large Pol II foci per cell and integrated nuclear fluorescence intensity. Mean±s.e.m. with unpaired, two-sided t-tests (n = 94 nuclei for DMSO; n = 89 for dTAG7). **d**, PRO-seq signal of auto-regulatory TFs *MYC* and *MYB*, and the expression-matched control gene *RAB3GAP1* after 1h MED14 degradation. Arrows highlight loss of promoter-proximal signal. H3K4me3 and H3K27ac ChIP-seq signal from KBM7 wild-type cells. **e**, Aggregated PRO-seq coverage over an SE-proximal metagene. TSS, transcription start site; TES, transcription end site. **f**, Changes in PRO-seq pausing index at n = 7,643 genes after 1h MED14 degradation. **g**, Observed vs. expected median Euclidean distance of auto-regulatory TFs from the pause-initiation limit in Fig. 3f. The expected distribution was generated by randomly selecting the same number of genes from bulk. **h**, Changes in productive initiation rate and pause duration of all 6,791 genes vs. the 24 auto-regulatory TFs. Productive initiation rates selectively decrease for auto-regulatory TFs, while pause duration decreases globally. Box plot elements: medians with interquartile range and 1.5x whiskers.



Extended Data Fig. 6 | Unbiased proteomics reveal increased P-TEFb levels on chromatin. a, Overlap of three independent data analyses to detect high-confidence differentially chromatin-bound proteins ($p < 0.1$; see Supplementary Note). **b**,**c**, Differential chromatin binding of transcription regulators. Class averages are shown in **b**. Scratched boxes indicate missing values. GTFs: general transcription factors. **d**, Salt-based fractionation of 7SK- and chromatin-bound P-TEFb complexes. Unprocessed western blot shown in Source Data.



Extended Data Fig. 7 | P-TEFb activation shapes the transcriptional response to Mediator loss. a, PRO-seq read-through upon 2 h MED14 degradation. Additionally inhibiting CDK9 with 500 nM NVP2 in the last 30 min reverses the read-through. Zoom-ins show 30 kb windows around the polyadenylation site. Arrows highlight transcription start site (TSS) regions shown in **g**. **b**, Aggregated PRO-seq coverages show read-through even for long genes, where newly initiated Pol II has not yet reached the termination site. Mean±bootstrapped confidence region. **c**, Aggregated TT-seq coverages show read-through transcription after MED14 degradation also in HCT-116 cells. **d**,**e**, Changes in PRO-seq pausing index of all n = 5,558 genes (**d**) and calculated pause duration at all n = 6,954 transcription units (**e**) after MED14/CDK9 perturbation. **f**, Changes in productive initiation rates for all n = 6,954 transcription units. Box plot elements: medians with interquartile range, 1.5x whiskers and confidence region notches. **g**, PRO-seq signal around transcription start sites (TSS) of two non-SE and one auto-regulatory TF gene. Paused polymerase does not re-accumulate at the *MYB* TSS upon combined MED14/CDK9 perturbation. **h**, TT-seq SE-gene set enrichment upon combined MED14/CDK9 perturbation. Less significant enrichment confirms that CDK9 activity aggravated the SE-selectivity of Mediator disruption.

nature research

Corresponding author(s): Dr. Georg E. Winter, Dr. Patrick Cramer

Last updated by author(s): Apr 21, 2020

Reporting Summary

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

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

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

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

Software and code

Policy information about availability of computer code

Data collection	Use of publically available software is fully referenced in the methods section.
Data analysis	Use of publically available software, including versions and key parameters used, is fully referenced in the methods section.

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

Data

Policy information about availability of data

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

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

Next generation sequencing data are available through NCBI Gene Expression Omnibus under accession code GSE139468. Chromatin proteomics data have been deposited at PRIDE under dataset identifier PXD017611.

Field-specific reporting

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

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

Life sciences study design

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

Sample size	Sample sizes were chosen to balance statistical power and technical feasibility, especially with regard to nascent RNA sequencing experiments, which relied on accurate timing across replicates. Quant-seq: biological triplicates TT-seq: biological duplicates HiChIP: biological duplicates 4C-seq: biological triplicates PRO-seq: biological duplicates ChIP-seq: Histone ChIP-seq in KBM7 wild-type cells was performed without biological replicates. Chromatin proteomics: biological triplicates
Data exclusions	No data were excluded in this study.
Replication	All attempts to replicate experiments in this study were successful (especially biochemical experiments, where representative western-blot results are shown).
Randomization	No samples were randomized in this study due to exclusive use of cell lines and comparisons between vehicle control and (short-term) drug treatments.
Blinding	No samples were blinded in this study due to exclusive use of cell lines and comparisons between vehicle control and (short-term) drug treatments.

Behavioural & social sciences study design

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

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

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

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and*

	any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access and import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	HA (1:1000, Cell Signaling Technology, #3724S and #2367S), GAPDH (1:1000, Santa Cruz Biotechnology, sc-365062), Histone H3 (1:10000 Abcam, ab1791), β-ACTIN (1:10000, Sigma Aldrich, A5441), MYC (1:1000, Santa Cruz Biotechnology, sc-764), MED10 (1:500, Abcam, ab110786), CDK9 (1:1000, Cell Signaling Technology, #2316S), BRD4 (1:5000, Bethyl Laboratories, A301-985A100), MED12 (1:2000, Bethyl Laboratories, A300-774A), MED1 (1:2000, Bethyl Laboratories, A300-793A), MED14 (1:1000, Bethyl Laboratories, A301-044A), MED26 (1:1000, Cell Signaling Technology, #14950S), MED6 (1:1000, Santa Cruz Biotechnology, sc-390474), MED23 (1:1000, Bethyl Laboratories, A300-425A), BRD9 (1:1000, Active Motif, #61537), CDK8 (1:1000, Cell Signaling Technology, #4101S), H3K27ac (1:1000, Abcam, ab4729), G3BP1 (1:500, Santa Cruz Biotechnology, sc-365338), α-TUBULIN (1:1000, Sigma Aldrich, T9026), CCNT1 (1:1000, Cell Signaling Technology, #81464S), HEXIM1 (1:1000, Abcam, ab25388), MED31 (1:500, Santa Cruz Biotechnology, sc-101189), pS2 Pol II (1:1000, Active Motif, #61083), Pol II (1:250, Santa Cruz Biotechnology, sc-899), pT806 SPT5 (1:1000, Robert Fisher lab, Icahn School of Medicine at Mount Sinai), SPT5 (1:1000, Santa Cruz Biotechnology, sc-133217), TUBULIN-Rhodamine (1:5000, Bio-Rad, #12004166).

anti-mouse-HRP (1:5000, Jackson ImmunoResearch, 115-035-003), anti-rabbit-HRP (1:5000, Jackson ImmunoResearch, 111-035-003), anti-rat-HRP (1:5000, Dako, P0450), anti-rabbit-StarBright700 (1:5000, Bio-Rad, #12004162), anti-mouse-Dylight800 (1:5000, Bio-Rad, STAR117D800GA).

MED1 (Abcam, ab64965), Pol II 8WG16 (Abcam, ab817). Secondary antibodies: anti-rabbit-AF594 (Invitrogen, A-21442), anti-mouse-AF488 (Invitrogen, A-11001).

Validation

Antibodies (especially against Mediator subunits) were validated by dTAG-mediated target degradation, colocalization with tagged Mediator subunits in immunofluorescence experiments, or phenotypic correlation with antibodies targeting other Mediator subunits (e.g. common molecular-weight shifts in size-exclusion chromatography, co-precipitation with b-isox hydrogels, co-immunopurification with known complex members, etc.).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Human near-haploid chronic myeloid leukemia KBM7 cells were obtained from Carette et al., Nat Biotechnol. 2011 (PMID: 21623355). HCT-116 cells were obtained from the Giulio Superti-Furga lab.

Authentication

KBM7 cell lines were authenticated by flow cytometric verification of (near-)haploidy. HCT-116 cells were authenticated by SNV genotyping (Microsynth, Switzerland) and additionally karyotyped to the following:
MP1: 45,X,-Y,add(10q),add(16p),add(18p)
MP2: 46,XY,add(10q),add(16p),add(18p)
MP3: 46,X,-Y,add(10q),add(16p),add(18p),+mar
MP4: 47,X,-Y,add(10q),add(16p),add(18p),+i(21)?x2
MP5: 46,XY,add(10q),add(16p),add(18p)

Mycoplasma contamination

Cell lines were regularly tested negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

None.

Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

Human research participants

Policy information about studies involving human research participants

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

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

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

Data access links *May remain private before publication.*	To review GEO accession GSE139468: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE139468
Files in database submission	See Supplementary Table 8 for a complete list.
Genome browser session (e.g. UCSC)	http://genome-euro.ucsc.edu/s/WinterLab/JaegerMG_Mediator_2020 This session summarizes all ChIP-seq, PRO-seq and TT-seq data generated (or used) in this study. PRO- and TT-seq tracks show replicate-merged signal for clearer visualization. BigWig signal files for all NGS experiments are available under GEO accession SuperSeries GSE139468.

Methodology

Replicates	Biological singlicates for Histone ChIP-seq.
Sequencing depth	raw_reads mapped_reads sample 58990025 40643907 H3K27Ac_KBM7.fastq 52329027 27913676 H3K4Me3_KBM7.fastq 60407066 44830033 input_KBM7.fastq all 40bp single-read.
Antibodies	H3K27ac: Abcam ab4729 H3K4me3: Millipore 07-473
Peak calling parameters	Raw reads were trimmed with skewer v0.1.126 with parameters "-f sanger -t 8 -m any" and aligned to the hg38 genome assembly using bowtie v2.3.4 with "--very-sensitive". Aligned reads were filtered for duplicates and MAPQ > 30 using sambamba v0.5.5 before calling peaks using macs2 v2.1.0 callpeak with "--fix-bimodal --extsize 180 --bw 200 -c $KBM7_input".
Data quality	Data quality was verified by cross-verification with nascent RNA sequencing data (especially enrichment of bi-directional enhancer transcription with H3K27ac peaks). peaks with FDR<5%: H3K4me3 = 36996, H3K27ac = 96697 peaks with >5-fold: H3K4me3 = 27784, H3K27ac = 41161
Software	see above under "peak calling parameters"

Flow Cytometry

Plots

Confirm that:

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

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

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

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

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

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

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.