Human promoter directionality is determined by transcriptional initiation and the opposing activities of INTS11 and CDK9

Joshua D Eaton[1,2,3], Jessica Board[1,3], Lee Davidson[1], Chris Estell[1] and Steven West[1*]

[1]The Living Systems Institute, University of Exeter, Stocker Rd, Exeter, EX4 4QD, UK.

[2]Current address: MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Trumpington, Cambridge CB2 0QH, UK

[3]Equal contribution

*Corresponding author: s.west@exeter.ac.uk; +44(0)1392727458

ABSTRACT

RNA polymerase II (RNAPII) transcription initiates bidirectionally at many human protein-coding genes. Sense transcription usually dominates and leads to messenger RNA production, whereas antisense transcription rapidly terminates. The basis for this directionality is not fully understood. Here, we show that sense transcriptional initiation is more efficient than in the antisense direction, which establishes initial directionality. After transcription begins, the opposing functions of Integrator (INTS11) and cyclin-dependent kinase 9 (CDK9) maintain directionality. Specifically, INTS11 terminates antisense transcription, whereas sense transcription is protected from INTS11-dependent attenuation by CDK9 activity. Strikingly, INTS11 attenuates transcription in both directions upon CDK9 inhibition, and the engineered recruitment of CDK9 desensitises transcription to INTS11. Therefore, the preferential initiation of sense transcription and the opposing activities of CDK9 and INTS11 explain mammalian promoter directionality.

61 **INTRODUCTION**

62 In humans, most protein-coding genes initiate transcription bidirectionally. Usually,
63 only sense transcription efficiently elongates and leads to mRNA synthesis. Antisense
64 transcription terminates within a few kilobases (kb), and the short non-coding (nc)RNA is
65 degraded[1]. Similar asymmetry frequently occurs in unicellular eukaryotes (e.g., budding
66 yeast) and at some plant promoters[2,3]. Thus, promoter directionality is a broadly observed
67 phenomenon. Interestingly, bidirectionality is the ground state of promoters and directionality
68 is acquired over evolutionary time[4]. This is proposed to be through a combination of DNA
69 sequences and proteins that favour the directional initiation and elongation of transcription.

70 The present explanation for the directionality of mammalian RNAPII promoters
71 involves the arrangement of U1 snRNA binding sites and polyadenylation signals (PASs). U1
72 snRNA promotes RNAPII elongation through protein-coding genes by binding to RNA and
73 preventing early termination[5-7]. RNA-bound U1 prevents early termination by inhibiting PASs
74 and antagonising other attenuation mechanisms, which include the PP1 regulator, PNUTS,
75 and the Restrictor complex[8,9]. In contrast, U1 binding sites are rarer in short antisense
76 transcripts, which are often rich in PAS sequences that are proposed to promote
77 transcriptional termination[10]. This model predicts that polyadenylation factors control a large
78 fraction of antisense transcriptional termination.

79 The multi-subunit Integrator complex also regulates promoter-proximal transcription[11-
80 15]. The Integrator complex comprises the backbone, arm/tail, phosphatase, and endonuclease
81 modules[11]. Its endoribonuclease is INTS11, and its phosphatase activity is mediated by INTS6
82 and protein phosphatase 2A (PP2A). INTS11 endonuclease broadly affects promoter-
83 proximal transcriptional attenuation whereas INTS phosphatase is proposed to regulate the
84 escape of RNAPII into elongation[16,17]. INTS6/PP2A phosphatase functionally antagonises
85 CDK9, which is vital for RNAPII promoter escape and productive elongation[18]. By this model,
86 CDK9 activity promotes elongation across protein-coding genes and INTS6/PP2A opposes it.

87 We tested the prediction that PAS factors control transcription directionality by
88 terminating antisense transcription, but this is not usually the case. Instead, we find that
89 promoter directionality is often conferred by preferential initiation in the sense direction and is
90 thereafter maintained by INTS11 and CDK9. The termination of antisense transcription is
91 constitutively INTS11-dependent, whereas sense transcription is hypersensitive to INTS11
92 only when CDK9 activity is simultaneously inhibited. We hypothesise that CDK9 activity
93 protects sense transcription from attenuation by INTS11 and that reduced CDK9 activity in the
94 antisense direction exposes RNAPII to INTS11-dependent termination.

95

RESULTS

Polyadenylation factor depletion does not increase or extend antisense transcription

The current explanation of mammalian promoter directionality invokes early PAS-dependent termination of antisense transcription[10]. This is partly based on the direct detection of polyadenylated antisense transcripts, which provides evidence that some of their transcriptional termination is PAS-dependent. However, there are non-polyadenylated antisense RNAs that might be attenuated in other ways[19]. To test the contribution of PAS-dependent termination toward antisense transcriptional termination and promoter directionality we tagged *RBBP6* with the dTAG degron. Because RBBP6 is required to activate the PAS cleavage machinery[20,21], its depletion should inhibit any PAS-dependent transcriptional termination. Three homozygous *dTAG-RBBP6* clones were isolated (Supplemental Figure 1A) and tagged RBBP6 was efficiently depleted after exposure to the dTAGv-1 degrader (Figure 1A). To test the contribution of RBBP6 to nascent transcription, we used POINT (Polymerase Intact Nascent Transcript)-seq[22], which maps full-length RNA extracted from immunoprecipitated RNAPII.

Figure 1B shows POINT-seq coverage over *NEDD1* including the upstream antisense transcript. RBBP6 loss causes a termination defect beyond the PAS shown by the extended POINT-seq signal beyond the *NEDD1* gene. However, RBBP6 does not affect the termination of upstream antisense RNA which is not extended when RBBP6 is depleted. Because RBBP6 loss induces strong termination defects beyond the PAS, we generalised its effects using genes separated from their neighbours by ≥10kb. Metaplot and heatmap analyses across these 1316 genes confirm that RBBP6 loss causes a general termination defect downstream of protein-coding genes but has little impact on antisense transcription (Figures 1C and D). Although many antisense transcripts contain multiple AAUAAA sequences, most still terminate RBBP6-independently (Supplementary Figures 1B and C). Similarly, we recently showed that the PAS-dependent 5'→3' exonucleolytic torpedo terminator, XRN2, does not affect antisense transcriptional termination[9]. Therefore, although some antisense transcripts are polyadenylated, most antisense transcription can terminate using PAS-independent mechanisms. Consequently, PAS-dependent termination cannot fully explain the promoter-proximal attenuation of antisense transcription.

Integrator depletion increases and extends antisense transcription

Our data argue that PAS-independent termination mechanisms control a large fraction of antisense transcription. A major PAS-independent termination pathway is driven by the Integrator complex, which was first identified as the 3' end processing complex for snRNAs[23].

132 Several reports show that Integrator terminates transcription from most promoters, including
133 those that initiate antisense transcription, so it might affect promoter directionality[11,16,17]. To
134 analyse this, we tagged its endonucleolytic subunit (INTS11) with a dTAG degron[24], which
135 enables rapid depletion of its catalytic activity (Figure 2A). We then performed POINT-seq on
136 *INTS11-dTAG* cells depleted or not of INTS11 to assay the global impact of INTS11 on RNAPII
137 transcription. Integrator is established to control the transcription of snRNAs, and this was
138 detected in our POINT-seq, which demonstrates a clear extension of read density at *RNU5A-*
139 *1* and *RNU5B-1* loci (Figure 2B).
140
141 As exemplified by *PGK1* (Figure 2C), INTS11 loss does not affect the termination of
142 protein-coding transcription but causes a strong upregulation of antisense transcription. Meta-
143 analysis of the same gene set used for RBBP6 POINT-seq shows the generality of antisense
144 transcriptional attenuation via INTS11 (Figure 2D). A heatmap analysis of transcription 3kb
145 upstream and downstream of the same transcription start sites (TSSs) demonstrates the
146 dominant impact of INTS11 on antisense vs. sense transcription at most of these promoters
147 (Figure 2E). Although the transcription over protein-coding genes is less affected by INTS11
148 loss, an increase in promoter-proximal transcription is apparent in some cases (Supplemental
149 Figure 2A). The most strongly affected genes are lowly expressed, which is consistent with
150 recent findings (Supplemental Figure 2B)[12-17]. Overall, INTS11 frequently attenuates
151 antisense transcription whereas a smaller fraction of sense transcription is affected. RNAPII.
152
153 The hypersensitivity of antisense transcripts to INTS11 might be due to the makeup of
154 antisense RNAPII or some other promoter feature. To interrogate this further, we analysed
155 two other promoter classes: those where protein-coding transcripts are initiated in both
156 directions and those that initiate the bidirectional transcription of unstable enhancer (e)RNAs.
157 In the former case, both transcripts are extended and stable like most sense transcripts at
158 directional promoters. In the latter case, both are short and unstable like most antisense
159 transcripts at directional promoters. When both directions are protein-coding, INTS11
160 depletion causes modest reductions in transcription (Supplemental Figure 2C). Conversely,
161 eRNA transcription is bidirectionally upregulated upon INTS11 elimination like the antisense
162 transcripts in (Supplemental Figure 2D). We conclude that short ncRNAs are more strongly
163 affected by INTS11 than protein-coding transcripts. At directional promoters this results in the
164 attenuation of antisense transcription.
165
166

167 **Transcription initiates more efficiently in the sense direction**

168 The increased antisense POINT-seq signal following INTS11 loss is consistent with
169 defects in transcriptional termination. However, it could result from increased transcriptional
170 initiation. We were also interested in whether preferential sense initiation could contribute to
171 mammalian promoter directionality as was proposed in budding yeast[4]. To precisely resolve
172 directional aspects of initiation and assay any impact of INTS11, we devised a variant of
173 POINT-seq called short (s)POINT. Briefly, sPOINT follows the POINT-seq protocol, but library
174 preparation employs the selective amplification of 5' capped RNAs <150nts (Figure 3A,
175 Supplemental Figure 3A, Experimental Procedures). The sPOINT signal
176 over *PGK1* exemplifies this and demonstrates a very restricted signal close to the TSS (Figure
177 3B). Figure 3C compares the meta profile of POINT- and sPOINT-seq on 684 well-expressed
178 and well-spaced (≥10kb from neighbours) genes to highlight the full read coverage obtained
179 by POINT-seq and the tightly restricted, TSS-proximal, sPOINT signal. Because most capped
180 RNAPII-associated RNA <150nts are promoter-proximal, sPOINT effectively assays the
181 promoter-proximal RNAPII pause.
182
183 The sPOINT-seq metaplot in Figure 3B shows a higher signal in the sense direction,
184 suggesting that more efficient transcriptional initiation establishes promoter directionality. To
185 assay this further and test the impact of INTS11, we performed sPOINT-seq in *INTS11-*
186 *dTAG* cells treated or not with dTAGv-1 and plotted the coverage over the promoters of well-
187 expressed protein-coding genes (3060 promoters, Figure 3D). As sPOINT-seq maps the 5'
188 and 3' ends of these reads, precise TSSs are mapped at single-nucleotide resolution and
189 plotted in the lower meta profile in Figure 3D. This analysis once again shows a higher sPOINT
190 signal in the sense direction. In addition, the lower TSS mapped plot also shows that sense
191 transcription initiated in a more focused manner compared to in the antisense direction. Figure
192 3E (*PGK1*) and supplemental figure 3B (*ACTB*) exemplify these features on individual genes,
193 and the metaplot in Figure 3F highlights the dispersed nature of antisense TSSs. Quantitation
194 of the TSS-derived sense vs. antisense signal confirms a higher read count in the sense
195 direction in untreated cells (Supplemental Figure 3C). The most noticeable (although still
196 modest) effect of INTS11 loss on sPOINT-seq profiles is a mild signal reduction. This shows
197 that INTS11 depletion does not generally enhance transcriptional initiation in either direction.
198 This slight reduction in sPOINT-seq signal could result from less transcriptional initiation or
199 less efficient pausing if, for example, INTS11 loss allows more RNAPII to escape the promoter.
200 Although a lower resolution technique, RNAPII ChIP-seq confirmed these promoter
201 characteristics and the mild impact of INTS11 (Supplemental Figures 3D and E). Overall,

202 these data show that directionality might be established by more efficient initiation in the sense
203 direction.
204
205 **CDK9 inhibition sensitises sense transcription to INTS11**

206 After transcription initiates, further control of directionality is evident because sense
207 transcription goes further than antisense transcription. If INTS11 regulates this, an opposing
208 force is required to promote elongation. INTS11 affects transcription very early, occupies
209 promoters, and becomes less active as RNAPII moves into elongation[16-18,25,26]. Therefore, if
210 INTS11 is counteracted to allow sense transcription, any responsible mechanism needs to act
211 early. One of the first transcriptional checkpoints involves the phosphorylation of RNAPII and
212 other factors by CDK9, which releases promoter-proximally paused RNAPII into elongation[27].
213 During this process, the Integrator-associated phosphatase, PP2A, antagonizes CDK9 and
214 presumably regulates the sensitivity of RNAPII to INTS11[18]. Because Integrator phosphatase
215 has little effect on antisense transcription[16], we hypothesised that INTS11 sensitivity and
216 CDK9 activity are inversely correlated to maintain directionality after initiation.

217 Our hypothesis predicts that sense transcription will be attenuated via INTS11 when
218 CDK9 is inactive. To test this genome-wide, we depleted INTS11 from *INTS11-dTAG* cells in
219 the presence or absence of a specific CDK9 inhibitor (NVP-2[28]) and performed POINT-seq.
220 As exemplified by *TARDBP*, the depletion of INTS11 alone caused an antisense
221 transcriptional termination defect with a milder impact on the protein-coding sense direction
222 (Figure 4A). As expected, NVP-2 treatment reduced transcription over the protein-coding
223 gene body. Antisense transcription also displays some CDK9 sensitivity. Importantly, in NVP-
224 2-treated cells, INTS11 loss increased transcription in both directions. This contrasts with the
225 dominant antisense effect deriving from just depleting INTS11 (refer to Figure 2). As such,
226 CDK9 activity prevents sense transcription from being attenuated by INTS11. This is a
227 genome-wide trend as shown in the metaplots in Figures 4B and C. A heatmap of the effect
228 of INTS11 loss after CDK9 inhibition shows the bidirectional upregulation of transcription at
229 the same promoters assayed in Figure 2 (Figure 4D). Although this experiment employed
230 longer INTS11 depletion than that in Figure 2 (2.5hr vs. 1.5hr to allow concurrent CDK9
231 inhibition), its antisense effects remain general, and the affected protein-coding genes strongly
232 overlapped (Supplemental Figure 4A-C). Lastly, when CDK9 and INTS11 are both
233 compromised, the POINT-seq signal remains higher in the sense vs. antisense direction (black
234 line, Figure 4C). This is consistent with our sPOINT-based finding that initiation of sense
235 transcription is generally more efficient.
236

237 Following CDK9 inhibition and INTS11 loss, the largest recovery of protein-coding
238 transcription is often over the 5' end of genes. This presumably reflects poor elongation
239 without CDK9 activity even when INTS11 is absent. RNAPII elongation is associated with
240 phosphorylation of the C-terminal domain (CTD) of its largest subunit, RPB1, and most
241 frequently occurs on Serine 5 or 2 (Ser5/2p) of its heptad repeat[29,30]. We assayed the effects
242 of CDK9 inhibition and INTS11 depletion on these two modifications by western blotting
243 (Figure 4E). As previously shown[28], CDK9 inhibition by NVP-2 suppresses Ser2p. However,
244 both Ser5p and Ser2p are enhanced by INTS11 loss whether CDK9 is active or not. Thus,
245 unattenuated transcription after INTS11 depletion is associated with some CTD
246 phosphorylation. Where NVP-2 is employed, kinases redundant with CDK9 presumably
247 phosphorylate RNAPII.
248
249 We confirmed the CDK9-mediated suppression of INTS11 on four selected protein-
250 coding genes using an alternative inhibitor, 5, 6-dichloro-1-β-D-ribofuranosylbenzimidazole
251 (DRB) (Supplemental Figure 4D). Interestingly, inhibition of CDK7 with THZ2 sensitised the
252 same selected protein-coding transcripts to INTS11 attenuation (Supplemental Figure 4E).
253 This may be because CDK7 is required to activate CDK9[31], or because Integrator additionally
254 targets RNAPII when CDK7 is inactive. This result reinforces the idea that the successful
255 execution of early transcriptional checkpoints may overcome Integrator-mediated attenuation.
256
257 If CDK9 counteracts INTS11, its recruitment should prevent transcriptional attenuation
258 by Integrator. To assay this, we employed a plasmid where transcription is driven by the
259 human immunodeficiency virus (HIV) promoter. Transcription from the HIV promoter results
260 in synthesis of the trans-activating response (TAR) element and promoter-proximal RNAPII
261 pausing. Pause release requires the trans-activator of transcription (TAT), which promotes
262 RNAPII elongation by recruiting CDK9[32]. INTS11 suppresses transcription from the HIV
263 promoter when TAT is absent[33]. To assay whether CDK9 affects this process, *INTS11-dTAG*
264 cells were transfected with the HIV reporter with or without TAT before treatment or not with
265 dTAGv-1. Transcription from the reporter was then analysed by qRT-PCR (Figure 4F).
266 INTS11 loss induces HIV transcription in the absence of TAT. However, TAT strongly
267 stimulates transcription (~200 fold) and desensitises it to INTS11 loss (slightly less reporter
268 RNA was recovered). Therefore, TAT-mediated CDK9 recruitment alleviates INTS11-
269 dependent attenuation of transcription. In line with endogenous protein-coding genes, CDK9
270 inhibition sensitises TAT-activated transcription to INTS11 (Supplemental Figure 4F). Overall,
271 our data strongly suggest that CDK9 activity counteracts transcriptional attenuation by
272 INTS11.

DISCUSSION

Most studies on mammalian promoter directionality have focused on understanding how antisense transcription terminates and how RNAPII elongation is enabled in the sense direction. However, we show that preferential sense transcriptional initiation significantly contributes toward the directionality of many mammalian promoters. More analyses are required to elucidate the mechanism, but we hypothesise that a unidirectional arrangement of at least some core promoter elements favours transcription in the protein-coding direction. Consistently, the evolution of promoter elements explains how a bidirectional promoter "ground state" acquired directionality in yeast[4]. Consistent with our finding, focused transcription initiation is associated with highly expressed mammalian genes and promoters with clearly defined elements (e.g. TATA)[34]. Thus, the lower level and dispersed nature of antisense initiation may reflect the suboptimal orientation of promoter elements or the opportunistic transcription of open chromatin in promoter regions.

We hypothesise that INTS11 controls transcriptional attenuation via its endonuclease activity. An inactive mutant of INTS11 (E203Q) is widely used to interrogate its catalytic function[23]; however, we discovered that it poorly associates with other Integrator components compared to wild-type INTS11 (Supplemental Figure 4G). Therefore, when employed in cells, it may not isolate the effects of INTS11 activity from those requiring an intact Integrator complex. Similarly, catalytic mutations in the highly related PAS endonuclease, CPSF73, disrupt its association with other cleavage and polyadenylation components[35]. In light of this, it is formally possible that non-catalytic consequences of INTS11 loss explain the upregulation of antisense transcription.

How CDK9 activity opposes INTS11 is unresolved, but INTS11 and SPT5 are adjacent in the RNAPII: Integrator structure[36]. As SPT5 is a prominent substrate of CDK9[37-39], its phosphorylation might evict INTS11 or prevent its association with the complex. Consistently, Integrator is enriched on promoters whereas phosphorylated SPT5 (SPT5p) is most prevalent during elongation over protein-coding gene bodies[25,38]. Furthermore, although antisense RNAPII occupancy is increased by INTS11 loss, published data suggest that this is not paralleled by increases in its Ser5p and Ser2p forms[16]. As these modifications require CDK7 and CDK9, this provides further support that these activities are less prevalent for antisense vs. sense transcription.

While INTS11 and CDK9 activities are the clearest on antisense and sense transcription, respectively, this is not binary. Some antisense transcription is NVP-2 sensitive, which indicates CDK9 activity on some RNAPII (Figure 4 and[40]). CDK9-sensitive antisense transcription could undergo INTS11-independent attenuation, which might account for the

308 polyadenylated antisense RNA observed previously[10,41]. Consistently, some antisense
309 transcription is targeted by the poly(A) exosome targeting connection, which operates on
310 cleaved and polyadenylated transcripts[42,43]. Reciprocally, promoter-proximal transcription
311 from lowly expressed protein-coding transcription is INTS11-sensitive, which could reflect low
312 CDK9 activity at those loci. We and others recently described the Restrictor complex as
313 restraining antisense transcription[9,44,45]. Unlike Integrator, which associates with NELF-bound
314 RNAPII[36], Restrictor is proximal to the distinct PAF-bound RNAPII[9]. Thus, various complexes
315 may target different forms of RNAPII before and after CDK9 activity. Even so, our data
316 suggests that Integrator influences a greater volume of antisense RNAPII than the cleavage
317 and polyadenylation pathway.

318 In addition to CDK9, multiple studies demonstrate a role for U1 snRNA in promoting
319 elongation through protein-coding genes[5,6,8,46]. U1 promotes RNAPII elongation and shields
320 transcription from early PAS-dependent termination and the Restrictor complex[5,9]. The U1-
321 mediated suppression of premature PAS-dependent termination inspired the original model of
322 promoter directionality. As U1 sites are rarer in antisense RNAs, their termination was
323 hypothesised to be driven by early PASs that would consequently be active. Our RBBP6 data
324 (and other published data on PAS termination factors[9]) argues that a large fraction of
325 antisense transcriptional termination is PAS-independent. Nevertheless, the lack of U1 sites
326 may limit RNAPII elongation and render it sensitive to Integrator (or Restrictor). Accordingly,
327 it will be interesting to test whether U1 regulates the sensitivity of transcription to Integrator.

328 In conclusion, we provide a new model for mammalian promoter directionality and the
329 subsequent control of bidirectional transcription (Figure 4G). sPOINT demonstrates more
330 efficient sense transcription initiation, which establishes directionality. Early events dictate the
331 decision to elongate or attenuate transcription. In the sense direction, CDK9 activity prevents
332 attenuation by Integrator to favour productive elongation, whereas antisense transcription is
333 hypersensitive to INTS11 and terminates early. As INTS11 occupies most promoters and
334 CDK9 activity is near-universally required to achieve protein-coding transcription, this model
335 can generally explain how directionality is initiated and maintained. Our new sPOINT-seq
336 approach will be valuable in elucidating additional aspects of promoter-proximal transcriptional
337 regulation.

338 **ACKNOWLEDGEMENTS**

339 We thank Hiroshi Kimura for the total RNAPII antibody used for POINT-seq. Our
340 research is funded by a Wellcome Trust Investigator Award to SW (223106/Z/21/Z). This
341 project used the University of Exeter Sequencing Service, and their equipment was funded by
342 the Wellcome Trust (Multi-User Equipment Grant award number 218247/Z/19/Z).

343 **AUTHOR CONTRIBUTIONS**

344 Conceptualization J.D.E, S.W; data curation J.D.E, J.B, L.D, C.E; formal analysis
345 J.D.E., J.B., L.D; C.E and S.W; methodology, J.D.E. and S.W.; investigation, J.D.E., J.B., L.D;
346 C.E and S.W.; supervision and funding acquisition, S.W.; writing and editing, J.D.E., J.B., L.D;
347 C.E and S.W.

348

349 **EXPERIMENTAL PROCEDURES**

350 **Sequencing data**

351 Deposited at Gene Expression Omnibus under accession: GSE243266.

352 **Cloning**

353 HIV reporter constructs were made by removing the CMV promoter, the entire β-globin
354 sequence, and its PAS from a pcDNA5 FRT/TO plasmid containing the WT β-globin (βWT)
355 gene[47] and inserting an HIV promoter and downstream TAR element derived from βΔ5-7[48].
356 *INTS11* targeting constructs were modified from those we previously described to generate
357 *INTS11-SMASh* cells[49]. The SMASh tag was removed and replaced with 2xHA dTAG derived
358 from Addgene plasmid 91792[24]. Guide RNA expressing Cas9 plasmids to modify *INTS11* or
359 *RBBP6* were made by inserting annealed oligonucleotides, containing the targeting sequence,
360 into px330[50] digested with BbsI.

361 **Cell culture and cell lines**

362 HCT116 cells were maintained in DMEM supplemented with penicillin/streptomycin at
363 37°C, 5% CO_2. *dTAG-RBBP6* cells were generated using the "CHoP in" protocol[51]. Briefly, a
364 24-well dish was transfected with 250ng of px330[50] containing the *RBBP6*-targeting guide and
365 250ng of PCR product containing the dTAG degron preceded by a blasticidin or puromycin
366 selection marker (derived from Addgene plasmid 91792 and 91793[24]). Jetprime (Polyplus)
367 was used for transfection. Three days later, cells were passaged into media containing
368 10μg/ml Blasticidin/1μg/ml Puromycin and colonies were PCR screened ~10 days later.
369 *INTS11-dTAG* cells were generated by homology-directed repair. A 6-well dish of cells was
370 transfected with 1μg px330 containing the *INTS11*-targeting guide (described in[49]) and 1μg
371 each of the repair templates. Three days later, cells were passaged into media containing
372 30μg/ml Hygromycin and 800μg/ml G418. ~10 days later, colonies were picked and screened
373 by PCR. 1μM dTAGv-1 (Tocris) was used for 1-14hrs (see figure legends for timings used in
374 each experiment); NVP-2 was used at 250nM for 2.5hr; DRB was used at 100μM for 2.5hr;
375 THZ2 was used at 5μM for 2.5hr.

376 **POINT-seq and sPOINT-seq**

377 For POINT-seq, we followed the protocol provided in[22]. The only modification was that
378 we started with a confluent 10cm dish of cells and performed the immunoprecipitation with
379 6µg of anti-RNAPII. ~2% cell volume of Drosophila S2 cells was included as a spike in control.
380 Libraries were prepared using the NEBNext Ultra™ II Directional RNA Library Prep Kit for
381 Illumina (New England Biolabs). sPOINT was performed in the same manner with the
382 following difference: a confluent 15cm dish of cells was used and 10µg of anti-RNAPII. In
383 sPOINT, the immunoprecipitated and DNAse treated RNA was treated with Terminator™ 5-
384 Phosphate-Dependent Exonuclease (lucigen) to remove any uncapped transcripts, and
385 libraries were prepared with the SMARTer® smRNA-Seq Kit for Illumina (Takara Bio) to
386 selectively capture transcripts <150nts.

387 **ChIP-seq**

388 For each experiment, 1x10cm dish of cells was used. Protein: DNA crosslinks were
389 formed by adding Formaldehyde (1% v/v) to culture media for 10 mins then quenching with
390 125mM glycine. Cells were rinsed 2x with PBS, scraped off the dish, and pelleted in 10ml
391 PBS at 500xg for 5 mins. We then employed the simple ChIP enzymatic kit (Cell Signalling
392 Technologies) to fragment chromatin and purify RNAPII-bound DNA. We followed the kit
393 protocol except for conjugating 5µg of anti-total RNAPII to sheep anti-mouse dynabeads (Life
394 Technologies). Sequencing libraries were generated using the NEBNext® Ultra™ II DNA
395 Library Prep Kit for Illumina.

396 **Total/chromatin-associated RNA isolation and qRT-PCR**

397 For total RNA, a 24-well dish of cells was transfected with 100ng HIV reporter plasmid
398 and, where co-transfected, 50ng TAT plasmid[52] using Jetprime (Polyplus) following the
399 manufacturer's protocol. Media was refreshed 5 hours post-transfection and dTAGv-1 was
400 added where appropriate. The next day, total RNA was isolated using Trizol (Thermo Fisher)
401 following the manufacturer's protocol. For chromatin-associated RNA (for Supplemental
402 Figure 4F), a 6 well dish of cells was transfected with 300ng HIV reporter plasmid and 150ng
403 TAT. Pelleted cells were resuspended in 800µl hypotonic lysis buffer (HLB: 10 mM Tris-HCl
404 (pH 7.5), 10 mM NaCl and 2.5 mM $MgCl_2$, 0.5% NP40) and underlayered with 200µl HLB
405 +10% sucrose. Nuclei were isolated by centrifugation for 5 mins at 500xg. These were
406 resuspended in 100µl NUN1 (20 mM Tris-HCl (pH 7.9), 75 mM NaCl, 0.5 mM EDTA, and 50%
407 Glycerol). After addition of 1ml of NUN2 (20 mM HEPES-KOH (pH 7.6), 300 mM NaCl, 0.2
408 mM EDTA, 7.5 mM $MgCl_2$, 1% NP-40, 1 M Urea) chromatin was isolated by 10 min incubation
409 on ice followed by 10 min centrifugation at 13000rpm. RNA was isolated from the chromatin
410 pellet using Trizol following the manufacturers' instructions. For qRT-PCR of total and

411 chromatin-associated RNA, RNA was DNase treated then 1µg was reverse transcribed using
412 Protoscript II reverse transcriptase (New England Biolabs). qPCR was performed using LUNA
413 SYBR green reagent (New England Biolabs) on a Qiagen Rotorgene instrument. Quantitative
414 analysis used the ΔCT method.

415 **Co-immunoprecipitation**

416 For each transfection, a semi-confluent 10cm dish of cells was transfected with 5µg of
417 plasmid expressing flag tagged wild type or E203Q INTS11. Immunoprecipitation was based
418 on the ELCAP protocol[53]. The following day, cells were washed 2x in PBS, scrapped into 10ml
419 PBS and spun down for 5 mins at 500xg. Pelleted cells were resuspended in 800µl hypotonic
420 lysis buffer low NP40 (HLB: 10 mM Tris-HCl (pH 7.5), 10 mM NaCl and 2.5 mM $MgCl_2$, 0.1%
421 NP40) and underlayered with 200µl HLB low NP40 +10% sucrose. Nuclei were resuspended
422 in 1ml Chromatin digestion buffer (20 mM HEPES pH 7.9, 1.5 mM MgCl2, 10% (v/v) glycerol,
423 150 mM NaCl, 0.1% (v/v) NP-40 and 250 U/mL Benzonase) and incubated for 1hr at 4°C.
424 Debris was pelleted at 13000rpm for 10 mins and supernatant was incubated with 20µl of anti-
425 flag magnetic beads (Sigma) for 2hrs at 4ºC. Beads were washed 6x with ice-cold chromatin
426 digestion buffer and samples were eluted by boiling in protein loading buffer (100 mM Tris-Cl
427 (pH 6.8) 4% (w/v) sodium dodecyl sulfate 0.2% (w/v) bromophenol blue 20% (v/v) glycerol)
428 before polyacrylamide separation and western blotting.

429 **qRT-PCR primers and gRNA target sites**

430 See supplemental table 1.

431 **BIOINFORMATICS**

432 **POINT-Seq alignment and visualisation**

433 A metagene list of genes with no overlapping regions within 10kb of any other
434 expressed transcription unit (1316). Adapters were removed from raw reads using Trim
435 Galore! and mapped to GRCh38 using HISAT2 using default parameters. Biological replicates
436 were normalised and merged using SAMtools merge. Split strand metagene plots were
437 produced using Drosophila spike-in normalised sense and antisense (scaled to -1) bigwig
438 coverage files separately with further graphical processing performed in R. For heat maps,
439 computematrix (DeepTools) was used to generate score files from the normalised bigwig files
440 using the 10kb non-overlapping gene list. A log2 ratio (depletion/control) was applied to
441 identify changes in reads. Plots were redrawn in R; parameters used for each heat map are
442 detailed in figure legends.

443 **sPOINT**

444 For sPOINT TSS metaplots showing full-length and 5' derived coverage, gene lists
445 were determined by selecting principal protein-coding transcript isoforms from gencode v42
446 human annotation – specifically, those containing both "appris_principal_1" and
447 "Ensembl_canonical" labels (15301 in total). The top 20% expressed (based on -dTAG) were
448 used to generate meta profiles (3060 TSSs in total). The 684 genes used to exemplify the
449 difference between POINT and sPOINT coverage are the top 50% of expressed genes (based
450 on POINT-seq in *INTS11-dTAG)* cells, which are separated from neighbouring transcription
451 usints by at least 10kb.

452 **Antisense PAS analysis**

453 Bed files from Figure 1C were edited to map the antisense transcript from each gene
454 by fitting each gene's start and end point to 3kb upstream of their TSS. Sequences underlying
455 these regions were obtained via BEDTools getFASTA (hg38) and consensus PAS (AATAAA)
456 enumerated in R. PAS frequency per transcript was plotted using ggplot2. Heatmaps for
457 transcripts containing > 1 or no PAS motifs were plotted with DeepTools.

458 **ChIP-seq alignment and plotting**

459 Adapters were removed from raw reads using Trim Galore! and mapped to GRCh38
460 using HISAT2 with default parameters. Reads were also mapped to Dm6 to identify spike-in
461 signal. Reads with MAPQ score of ≤30 were removed with SAMtools[54]. Peaks were called
462 using MACS2 in paired-end mode[55]. Metaplots were created from S2 scaled, merged replicate
463 -log10 q-value bigwig coverage using DeepTools with further processing performed within R.
464 For Supplementary Figure 3E, plots used the same gene list described above for sPOINT
465 promoter analyses (see sPOINT section above).

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (*D. melanogaster*)	S2	This paper	FLYB:FBtc0000181; RRID:CVCL_Z992	Cell line maintained in N. Perrimon lab; FlyBase symbol:

				S2-DRSC.
Cell line (H. Sapiens)	HCT116	In house	This paper	
Cell line (H. Sapiens)	*INTS11-dTAG*	In house	This paper	
Cell line (H. Sapiens)	*dTAG-RBBP6*	In house	This paper	
Antibody	INTS11	Abbexa	abx234340	WB: 1:500
Antibody	INTS10	Proteintech	Cat# 15271-1-AP RRID:AB_2127260	WB: 1:3000
Antibody	INTS8	Proteintech	Cat# 18802-1-AP RRID:AB_10597250	WB: 1:1000
Antibody	INTS9	Proteintech	Cat# 11657-1-AP RRID:AB_2127514	WB: 1:1000
Antibody	INTS1	Bethyl	Cat# A300-361A-T RRID:A	WB: 1:500

			B_2632121 (Now discontinued)	
Antibody	HA	Cell Signalling	Cat# 3724 RRID:AB_1549585	WB: 1:2000
Antibody	RNAPII Ser2p	Active Motif	Cat# 61084 RRID:AB_2687450	WB: 1:2000
Antibody	RNAPII Ser5p	Millipore	Cat# 04-1572-I RRID:AB_2801296	WB: 1:2000
Antibody	Total RNAPII	Hiroshi Kimura Lab	N/A	WB: 1:10,000 IP: 5-10ug (Ab:Beads)
Antibody	SPT5	Santa Cruz Biotech	Cat# sc-136075 RRID:AB_2286723	WB: 1:1000
Antibody	Flag	Proteintech	Cat# 20543-1-AP RRID:AB_11232216	WB: 1:5000

Antibody	anti-Rabbit	Proteintech	Cat# SA00001-2 RRID:AB_2722564	WB: 1:2000
Antibody	anti-Mouse	Proteintech	Cat# SA00001-1-A RRID:AB_2890995	WB: 1:2000
Antibody	anti-Rat	Thermofisher	Cat #A18865 RRID:AB_2535642	WB: 1:2000
Recombinant DNA reagent	px300	Addgene	RRID #: Addgene_42230	
Recombinant DNA reagent	pCRIS-PITChv2-BSD-dTAG (BRD4)	Addgene	RRID:Addgene_91792	
Recombinant DNA reagent	pCRIS-PITChv2-PURO-dTAG (BRD4)	Addgene	RRID:Addgene_91793	
Transfected construct	TAT	[52]		
Transfected construct	INTS11 WT and E203Q	Twist	This paper	Supplemental Table 1
Transfected construct	HIV reporter	In house	This paper	

Commercial assay, kit	Terminator 5-phosphate dependant exonuclease	Lucigen	Cat #TER51020	
Commercial assay, kit	SMARTer smRNA-seq kit for illumina	Takara	Cat #635031	
Commercial assay, kit	Dynabeads M280 sheep anti-mouse	Thermofisher	Cat #11202D	
Commercial assay, kit	anti-FLAG/DYKDDDDK magnetic beads	MilliporeSigma	Cat #M8823	
Commercial assay, kit	SimpleChIP® Plus Enzymatic Chromatin kit	Cell Signalling	Cat #9005	
Commercial assay, kit	NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®	NEB	Cat #E7645S	
Commercial assay, kit	JetPRIME	PolyPlus	Cat #114-01	
Commercial assay, kit	NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina	NEB	Cat #E7765	
Commercial assay, kit	Protoscript II reverse transcriptase	NEB	Cat #M0368	
Commercial assay, kit	LUNA qPCR reagent	NEB	Cat #M3003	

Commercial assay, kit	Ampure XP beads	Beckman Coulter	Cat #A63880	
Other	Benzonase	Sigma	Cat #E1014-5KU	
Chemical compound, drug	dTAGv-1	Tocris	Cat #2624313-15-9	
Chemical compound, drug	NVP2	Merck	Cat #SML3069-5MG	
Chemical compound, drug	THZ2	Medchem Express	Cat #HY-12280	
Chemical compound, drug	DRB	Merck	Cat #D1916	
Chemical compound, drug	Trizol Reagent	Thermofisher	Cat #15596026	
Commercial assay, kit	Turbo DNAse kit	Thermofisher	Cat #AM2238	
software, algorithm	Bamtools	[56]	NA	NA
software, algorithm	BEDtools	[57]	NA	NA
software, algorithm	Cutadapt	Other	NA	https://doi.org/10.14806/ej.17.1.200

software, algorithm	Deeptools	[58]	NA	NA
software, algorithm	FastQC	Other	NA	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
software, algorithm	Hisat2	[59]	NA	NA
software, algorithm	IGV	[60]	NA	NA
software, algorithm	MACS2	[55]	NA	NA
software, algorithm	MultiQC	[61]	NA	NA
software, algorithm	R	Other	NA	https://cran.r-project.org/
software, algorithm	Rstudio	Other	NA	https://www.rstudio.com/
software, algorithm	SAMTools	[54]	NA	NA
software, algorithm	Trim_galore!	Other	NA	https://github.com/FelixKrueger/TrimGalore/

466

467

468 **REFERENCES**

469 1. Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S.,
470 Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion
471 reveals transcription upstream of active human promoters. Science *322*, 1851-1854.
472 1164096 [pii]

473 10.1126/science.1164096.

474 2. Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J.,
475 Regnault, B., Devaux, F., Namane, A., Seraphin, B., et al. (2005). Cryptic pol II
476 transcripts are degraded by a nuclear quality control pathway involving a new poly(A)
477 polymerase. Cell *121*, 725-737. 10.1016/j.cell.2005.04.030.

478 3. Thieffry, A., Vigh, M.L., Bornholdt, J., Ivanov, M., Brodersen, P., and Sandelin, A.
479 (2020). Characterization of Arabidopsis thaliana Promoter Bidirectionality and
480 Antisense RNAs by Inactivation of Nuclear RNA Decay Pathways. Plant Cell *32*,
481 1845-1867. 10.1105/tpc.19.00815.

482 4. Jin, Y., Eser, U., Struhl, K., and Churchman, L.S. (2017). The Ground State and
483 Evolution of Promoter Region Directionality. Cell *170*, 889-898 e810.
484 10.1016/j.cell.2017.07.006.

485 5. Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L., and Dreyfuss, G.
486 (2010). U1 snRNP protects pre-mRNAs from premature cleavage and
487 polyadenylation. Nature *468*, 664-668. nature09479 [pii]

488 10.1038/nature09479.

489 6. Mimoso, C.A., and Adelman, K. (2023). U1 snRNP increases RNA Pol II elongation
490 rate to enable synthesis of long genes. Molecular cell. 10.1016/j.molcel.2023.03.002.

491 7. Chiu, A.C., Suzuki, H.I., Wu, X., Mahat, D.B., Kriz, A.J., and Sharp, P.A. (2018).
492 Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature
493 Polyadenylation Suppressed by U1 snRNP. Molecular cell *69*, 648-663 e647.
494 10.1016/j.molcel.2018.01.006.

495 8. So, B.R., Di, C., Cai, Z., Venters, C.C., Guo, J., Oh, J.M., Arai, C., and Dreyfuss, G.
496 (2019). A Complex of U1 snRNP with Cleavage and Polyadenylation Factors
497 Controls Telescripting, Regulating mRNA Transcription in Human Cells. Molecular
498 cell *76*, 590-599 e594. 10.1016/j.molcel.2019.08.007.

499 9. Estell, C., Davidson, L., Eaton, J.D., Kimura, H., Gold, V.A.M., and West, S. (2023).
500 A restrictor complex of ZC3H4, WDR82, and ARS2 integrates with PNUTS to control
501 unproductive transcription. Molecular cell *83*, 2222-2239 e2225.
502 10.1016/j.molcel.2023.05.029.

503 10. Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B., and Sharp, P.A. (2013). Promoter
504 directionality is controlled by U1 snRNP and polyadenylation signals. Nature *499*,
505 360-363. 10.1038/nature12349.

506 11. Wagner, E.J., Tong, L., and Adelman, K. (2023). Integrator is a global promoter-
507 proximal termination complex. Molecular cell *83*, 416-427.
508 10.1016/j.molcel.2022.11.012.

509 12. Lykke-Andersen, S., Zumer, K., Molska, E.S., Rouviere, J.O., Wu, G., Demel, C.,
510 Schwalb, B., Schmid, M., Cramer, P., and Jensen, T.H. (2021). Integrator is a
511 genome-wide attenuator of non-productive transcription. Molecular cell *81*, 514-529
512 e516. 10.1016/j.molcel.2020.12.014.

513 13. Beckedorff, F., Blumenthal, E., daSilva, L.F., Aoi, Y., Cingaram, P.R., Yue, J., Zhang,
514 A., Dokaneheifard, S., Valencia, M.G., Gaidosh, G., et al. (2020). The Human
515 Integrator Complex Facilitates Transcriptional Elongation by Endonucleolytic
516 Cleavage of Nascent Transcripts. Cell reports *32*, 107917.
517 10.1016/j.celrep.2020.107917.

518 14. Tatomer, D.C., Elrod, N.D., Liang, D., Xiao, M.S., Jiang, J.Z., Jonathan, M., Huang,
519 K.L., Wagner, E.J., Cherry, S., and Wilusz, J.E. (2019). The Integrator complex
520 cleaves nascent mRNAs to attenuate transcription. Genes & development *33*, 1525-
521 1538. 10.1101/gad.330167.119.

522 15. Elrod, N.D., Henriques, T., Huang, K.L., Tatomer, D.C., Wilusz, J.E., Wagner, E.J.,
523 and Adelman, K. (2019). The Integrator Complex Attenuates Promoter-Proximal
524 Transcription at Protein-Coding Genes. Molecular cell *76*, 738-752 e737.
525 10.1016/j.molcel.2019.10.034.

526 16. Hu, S., Peng, L., Song, A., Ji, Y.X., Cheng, J., Wang, M., and Chen, F.X. (2023).
527 INTAC endonuclease and phosphatase modules differentially regulate transcription
528 by RNA polymerase II. Molecular cell *83*, 1588-1604 e1585.
529 10.1016/j.molcel.2023.03.022.

530 17. Stein, C.B., Field, A.R., Mimoso, C.A., Zhao, C., Huang, K.L., Wagner, E.J., and
531 Adelman, K. (2022). Integrator endonuclease drives promoter-proximal termination at
532 all RNA polymerase II-transcribed loci. Molecular cell *82*, 4232-4245 e4211.
533 10.1016/j.molcel.2022.10.004.

534 18. Vervoort, S.J., Welsh, S.A., Devlin, J.R., Barbieri, E., Knight, D.A., Offley, S.,
535 Bjelosevic, S., Costacurta, M., Todorovski, I., Kearney, C.J., et al. (2021). The PP2A-
536 Integrator-CDK9 axis fine-tunes transcription and can be targeted therapeutically in
537 cancer. Cell *184*, 3143-3162 e3132. 10.1016/j.cell.2021.04.022.

538 19. Gockert, M., Schmid, M., Jakobsen, L., Jens, M., Andersen, J.S., and Jensen, T.H.
539 (2022). Rapid factor depletion highlights intricacies of nucleoplasmic RNA
540 degradation. Nucleic acids research *50*, 1583-1600. 10.1093/nar/gkac001.

541 20. Schmidt, M., Kluge, F., Sandmeir, F., Kuhn, U., Schafer, P., Tuting, C., Ihling, C.,
542 Conti, E., and Wahle, E. (2022). Reconstitution of 3' end processing of mammalian
543 pre-mRNA reveals a central role of RBBP6. Genes & development *36*, 195-209.
544 10.1101/gad.349217.121.

545 21. Boreikaite, V., Elliott, T.S., Chin, J.W., and Passmore, L.A. (2022). RBBP6 activates
546 the pre-mRNA 3' end processing machinery in humans. Genes & development *36*,
547 210-224. 10.1101/gad.349223.121.

548 22. Sousa-Luis, R., Dujardin, G., Zukher, I., Kimura, H., Weldon, C., Carmo-Fonseca, M.,
549 Proudfoot, N.J., and Nojima, T. (2021). POINT technology illuminates the processing
550 of polymerase-associated intact nascent transcripts. Molecular cell *81*, 1935-1950
551 e1936. 10.1016/j.molcel.2021.02.034.

552 23. Baillat, D., Hakimi, M.A., Naar, A.M., Shilatifard, A., Cooch, N., and Shiekhattar, R.
553 (2005). Integrator, a multiprotein mediator of small nuclear RNA processing,
554 associates with the C-terminal repeat of RNA polymerase II. Cell *123*, 265-276.
555 10.1016/j.cell.2005.08.019.

556 24. Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett,
557 A.L., Erb, M.A., Lawlor, M.A., Souza, A., et al. (2018). The dTAG system for
558 immediate and target-specific protein degradation. Nature chemical biology *14*, 431-
559 441. 10.1038/s41589-018-0021-8.

560 25. Zheng, H., Qi, Y., Hu, S., Cao, X., Xu, C., Yin, Z., Chen, X., Li, Y., Liu, W., Li, J., et
561 al. (2020). Identification of Integrator-PP2A complex (INTAC), an RNA polymerase II
562 phosphatase. Science *370*. 10.1126/science.abb5872.

563 26. Ramamurthy, L., Ingledue, T.C., Pilch, D.R., Kay, B.K., and Marzluff, W.F. (1996).
564 Increasing the distance between the snRNA promoter and the 3' box decreases the
565 efficiency of snRNA 3'-end formation. Nucleic acids research *24*, 4525-4534.
566 10.1093/nar/24.22.4525.

567 27. Fujinaga, K., Huang, F., and Peterlin, B.M. (2023). P-TEFb: The master regulator of
568 transcription elongation. Molecular cell *83*, 393-403. 10.1016/j.molcel.2022.12.006.

569 28. Olson, C.M., Jiang, B., Erb, M.A., Liang, Y., Doctor, Z.M., Zhang, Z., Zhang, T.,
570 Kwiatkowski, N., Boukhali, M., Green, J.L., et al. (2018). Pharmacological
571 perturbation of CDK9 using selective CDK9 inhibition or degradation. Nature
572 chemical biology *14*, 163-170. 10.1038/nchembio.2538.

573 29. Schuller, R., Forne, I., Straub, T., Schreieck, A., Texier, Y., Shah, N., Decker, T.M.,
574 Cramer, P., Imhof, A., and Eick, D. (2016). Heptad-Specific Phosphorylation of RNA
575 Polymerase II CTD. Molecular cell *61*, 305-314. 10.1016/j.molcel.2015.12.003.

576 30. Suh, H., Ficarro, S.B., Kang, U.B., Chun, Y., Marto, J.A., and Buratowski, S. (2016).
577 Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA
578 Polymerase II. Molecular cell *61*, 297-304. 10.1016/j.molcel.2015.12.021.

579 31. Larochelle, S., Amat, R., Glover-Cutter, K., Sanso, M., Zhang, C., Allen, J.J., Shokat,
580 K.M., Bentley, D.L., and Fisher, R.P. (2012). Cyclin-dependent kinase control of the
581 initiation-to-elongation switch of RNA polymerase II. Nat Struct Mol Biol *19*, 1108-
582 1115. 10.1038/nsmb.2399.

583 32. Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). A novel
584 CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its
585 high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462. 10.1016/s0092-
586 8674(00)80939-3.

587 33. Stadelmayer, B., Micas, G., Gamot, A., Martin, P., Malirat, N., Koval, S., Raffel, R.,
588 Sobhian, B., Severac, D., Rialle, S., et al. (2014). Integrator complex regulates
589 NELF-mediated RNA polymerase II pause/release and processivity at coding genes.
590 Nat Commun *5*, 5531. 10.1038/ncomms6531.

591 34. Rengachari, S., Schilbach, S., Kaliyappan, T., Gouge, J., Zumer, K., Schwarz, J.,
592 Urlaub, H., Dienemann, C., Vannini, A., and Cramer, P. (2022). Structural basis of
593 SNAPc-dependent snRNA transcription initiation by RNA polymerase II. Nat Struct
594 Mol Biol *29*, 1159-1169. 10.1038/s41594-022-00857-w.

595 35. Kolev, N.G., Yario, T.A., Benson, E., and Steitz, J.A. (2008). Conserved motifs in
596 both CPSF73 and CPSF100 are required to assemble the active endonuclease for
597 histone mRNA 3'-end maturation. EMBO Rep *9*, 1013-1018.
598 10.1038/embor.2008.146.

599 36. Fianu, I., Chen, Y., Dienemann, C., Dybkov, O., Linden, A., Urlaub, H., and Cramer,
600 P. (2021). Structural basis of Integrator-mediated transcription regulation. Science
601 *374*, 883-887. 10.1126/science.abk0154.

602 37. Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H.
603 (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for
604 processive transcription elongation. Molecular cell *21*, 227-237. S1097-
605 2765(05)01812-5 [pii]

606 10.1016/j.molcel.2005.11.024.

607 38. Cortazar, M.A., Sheridan, R.M., Erickson, B., Fong, N., Glover-Cutter, K., Brannan,
608 K., and Bentley, D.L. (2019). Control of RNA Pol II Speed by PNUTS-PP1 and Spt5
609 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism.
610 Molecular cell *76*, 896-908 e894. 10.1016/j.molcel.2019.09.031.

611 39. Parua, P.K., Booth, G.T., Sanso, M., Benjamin, B., Tanny, J.C., Lis, J.T., and Fisher,
612 R.P. (2018). A Cdk9-PP1 switch regulates the elongation-termination transition of
613 RNA polymerase II. Nature *558*, 460-464. 10.1038/s41586-018-0214-z.

614 40. Fong, N., Sheridan, R.M., Ramachandran, S., and Bentley, D.L. (2022). The pausing
615 zone and control of RNA polymerase II elongation by Spt5: Implications for the
616 pause-release model. Molecular cell *82*, 3632-3645 e3634.
617 10.1016/j.molcel.2022.09.001.

618 41. Flynn, R.A., Almada, A.E., Zamudio, J.R., and Sharp, P.A. (2011). Antisense RNA
619 polymerase II divergent transcripts are P-TEFb dependent and substrates for the
620 RNA exosome. Proceedings of the National Academy of Sciences of the United
621 States of America *108*, 10460-10465. 10.1073/pnas.1106630108.

622 42. Wu, G., Schmid, M., Rib, L., Polak, P., Meola, N., Sandelin, A., and Jensen, T.H.
623 (2020). A Two-Layered Targeting Mechanism Underlies Nuclear RNA Sorting by the
624 Human Exosome. Cell reports *30*, 2387-2401 e2385. 10.1016/j.celrep.2020.01.068.

625 43. Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-
626 Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., and Jensen, T.H.
627 (2016). Identification of a Nuclear Exosome Decay Pathway for Processed
628 Transcripts. Molecular cell *64*, 520-533. 10.1016/j.molcel.2016.09.025.

629 44. Estell, C., Davidson, L., Steketee, P.C., Monier, A., and West, S. (2021). ZC3H4
630 restricts non-coding transcription in human cells. Elife *10*. 10.7554/eLife.67305.

631 45. Austenaa, L.M.I., Piccolo, V., Russo, M., Prosperini, E., Polletti, S., Polizzese, D.,
632 Ghisletti, S., Barozzi, I., Diaferia, G.R., and Natoli, G. (2021). A first exon termination
633 checkpoint preferentially suppresses extragenic transcription. Nat Struct Mol Biol *28*,
634 337-346. 10.1038/s41594-021-00572-y.

635 46. Vlaming, H., Mimoso, C.A., Field, A.R., Martin, B.J.E., and Adelman, K. (2022).
636 Screening thousands of transcribed coding and non-coding regions reveals
637 sequence determinants of RNA polymerase II elongation potential. Nat Struct Mol
638 Biol. 10.1038/s41594-022-00785-9.

639 47. Muniz, L., Davidson, L., and West, S. (2015). Poly(A) Polymerase and the Nuclear
640 Poly(A) Binding Protein, PABPN1, Coordinate the Splicing and Degradation of a
641 Subset of Human Pre-mRNAs. Molecular and cellular biology *35*, 2218-2230.
642 10.1128/MCB.00123-15.

643 48. Dye, M.J., and Proudfoot, N.J. (2001). Multiple transcript cleavage precedes
644 polymerase release in termination by RNA polymerase II. Cell *105*, 669-681. S0092-
645 8674(01)00372-5 [pii].

646 49. Davidson, L., Francis, L., Eaton, J.D., and West, S. (2020). Integrator-Dependent and
647 Allosteric/Intrinsic Mechanisms Ensure Efficient Termination of snRNA Transcription.
648 Cell reports *33*, 108319. 10.1016/j.celrep.2020.108319.

649 50. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X.,
650 Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using
651 CRISPR/Cas systems. Science *339*, 819-823. science.1231143 [pii]

652 10.1126/science.1231143.

653 51. Manna, P.T., Davis, L.J., and Robinson, M.S. (2019). Fast and cloning-free
654 CRISPR/Cas9-mediated genomic editing in mammalian cells. Traffic *20*, 974-982.
655 10.1111/tra.12696.

656 52. Adams, S.E., Johnson, I.D., Braddock, M., Kingsman, A.J., Kingsman, S.M., and
657 Edwards, R.M. (1988). Synthesis of a gene for the HIV transactivator protein TAT by
658 a novel single stranded approach involving in vivo gap repair. Nucleic acids research
659 *16*, 4287-4298.

53. Gregersen, L.H., Mitter, R., and Svejstrup, J.Q. (2022). Elongation factor-specific capture of RNA polymerase II complexes. Cell Rep Methods *2*, 100368. 10.1016/j.crmeth.2022.100368.

54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.

55. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.

56. Barnett, D.W., Garrison, E.K., Quinlan, A.R., Stromberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691-1692. 10.1093/bioinformatics/btr174.

57. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842. 10.1093/bioinformatics/btq033.

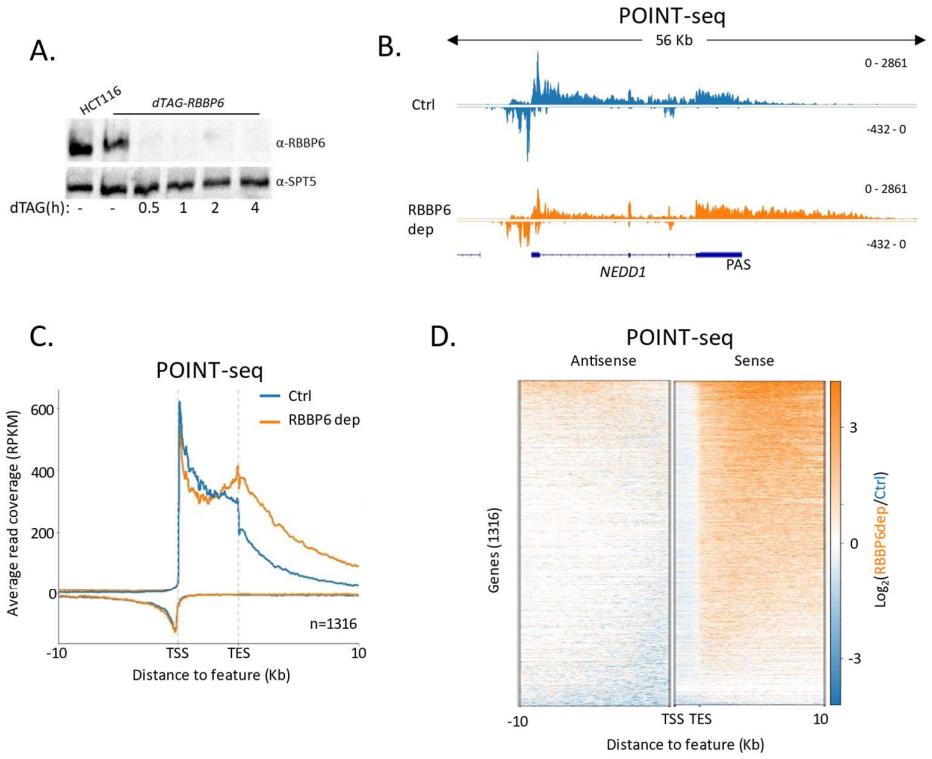
58. Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research *44*, W160-165. 10.1093/nar/gkw257.

59. Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360. 10.1038/nmeth.3317.

60. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26. 10.1038/nbt.1754.

61. Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047-3048. 10.1093/bioinformatics/btw354.

FIGURE 1


A.
HCT116
dTAG-RBBP6
α-RBBP6
α-SPT5
dTAG(h): - - 0.5 1 2 4
B.
POINT-seq
56 Kb
Ctrl
0 - 2861
-432 - 0
RBBP6 dep
0 - 2861
-432 - 0
NEDD1
PAS
C.
POINT-seq
Ctrl
RBBP6 dep
Average read coverage (RPKM)
600
400
200
0
n=1316
-10
TSS
TES
10
Distance to feature (Kb)
D.
POINT-seq
Antisense
Sense
Genes (1316)
3
0
-3
Log2(RBBP6dep/Ctrl)
-10
TSS TES
10
Distance to feature (Kb)

697

698 **FIGURE 1: RBBP6 loss disrupts PAS-dependent termination of sense transcription.**

699 **A.** Western blot demonstrating the depletion of dTAG-RBBP6 over a time course of
700 dTAGv-1 addition. SPT5 serves as a loading control.

701 **B.** Genome browser track of *NEDD1* in POINT-seq data from *dTAG-RBBP6* cells treated
702 or not (2hr) with dTAGv-1. RBBP6 depletion induces a termination defect in the
703 protein-coding direction (downstream of the indicated PAS) but not the upstream
704 antisense direction. The y-axis shows Reads Per Kilobase per Million mapped reads
705 (RPKM).

706 **C.** Metaplot of POINT-seq data from *dTAG-RBBP6* cells treated (RBBP6 dep) or not (Ctrl)
707 (2hr) with dTAGv-1. This shows 1316 protein-coding genes selected as separated
708 from any expressed transcription unit by ≥10kb. Signals above and below the x-axis
709 are sense and antisense reads, respectively. The y-axis scale is RPKM.
710 TSS=transcription start site; TES=transcription end site (this marks the PAS position).
711 This is an average of two biological replicates.

712 **D.** Heatmap representation of the data in C, which displays signal as a log2 fold change
713 (log2FC) in RBBP6 depleted versus un-depleted conditions. This is an average of two
714 biological replicates.

715

716

717

718

719

720

721

722

723

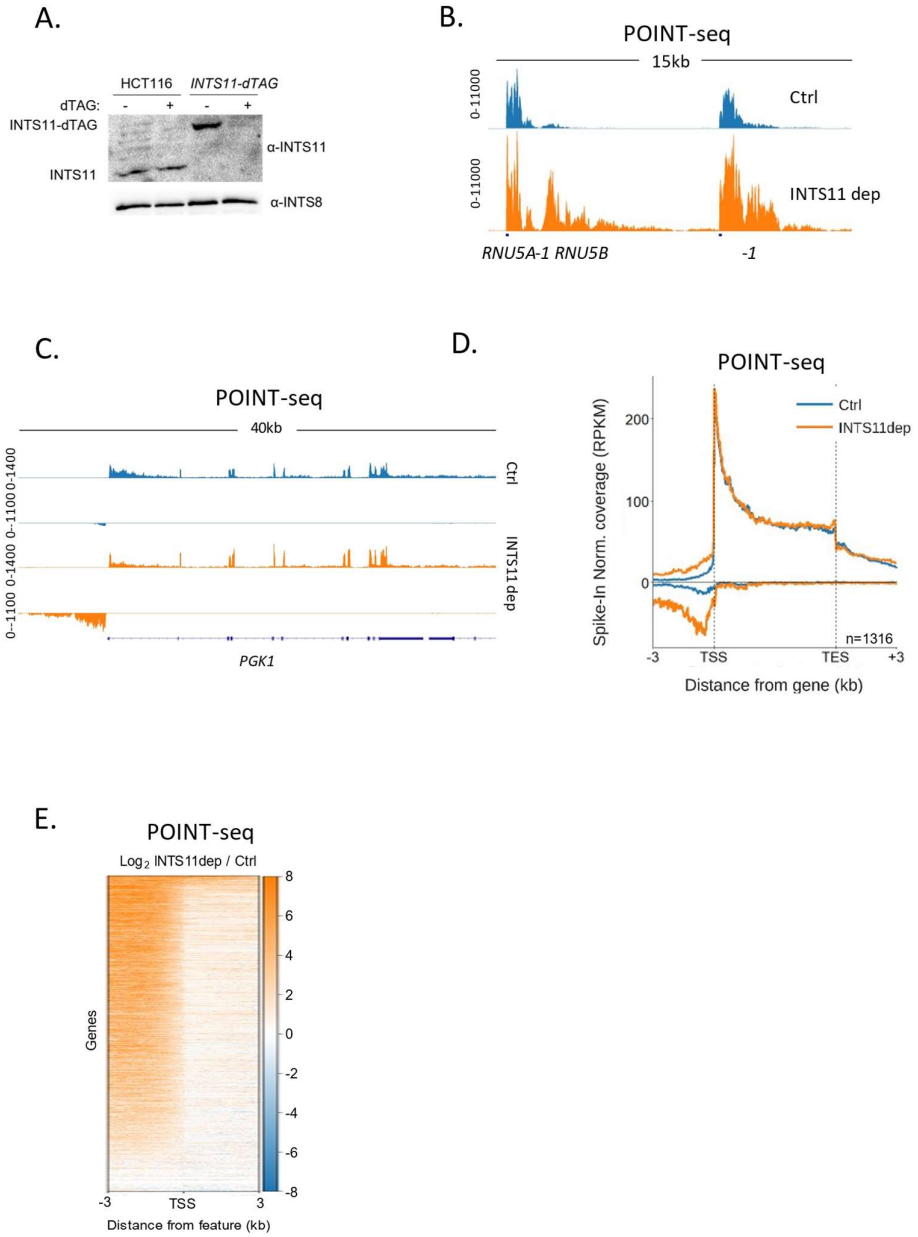
724

725

726

727

FIGURE 2


A.
HCT116
INTS11-dTAG
dTAG:
- + - +
INTS11-dTAG
INTS11
α-INTS11
α-INTS8
B.
POINT-seq
15kb
0-11000
Ctrl
0-11000
INTS11 dep
RNU5A-1 RNU5B
-1
C.
POINT-seq
40kb
0-1400
0-1100
Ctrl
0-1400
0-1100
INTS11 dep
PGK1
D.
POINT-seq
Ctrl
INTS11dep
Spike-In Norm. coverage (RPKM)
200
100
0
n=1316
-3
TSS
TES
+3
Distance from gene (kb)
E.
POINT-seq
Log2 INTS11dep / Ctrl
8
6
4
2
0
-2
-4
-6
-8
Genes
-3
TSS
3
Distance from feature (kb)

728

729 **FIGURE 2: INTS11 loss disrupts the termination of antisense transcription.**

730 **A.** Western blot demonstrating homozygous tagging of *INTS11* with dTAG and the
731 depletion of INTS11-dTAG after 1.5hr treatment with dTAGv-1. INTS8 serves as a
732 loading control.

733 **B.** Genome browser track showing POINT-seq signal over *RNU5A-1* and *RNU5B-1* in
734 *INTS11-dTAG* cells treated or not (1.5hr) with dTAGv-1. Note, that INTS11 depletion
735 induces a termination defect in each case. Y-axis shows RPKM following spike-in
736 normalisation.

737 **C.** Genome browser track showing POINT-seq signal over *PGK1* and its upstream
738 antisense region in *INTS11-dTAG* cells treated or not (1.5hr) with dTAGv-1. Y-axis
739 shows RPKM following spike-in normalisation.

740 **D.** Metaplot of POINT-seq data from *INTS11-dTAG* cells treated or not (1.5hr) with
741 dTAGv-1. This shows the same 1316 genes used in Figure 1C. Signals above and
742 below the x-axis are sense and antisense reads, respectively. Y-axis shows RPKM
743 following spike-in normalisation. This is an average of three biological replicates.

744 **E.** Heatmap representation of the data in D, which displays signal as a log2 fold change
745 (log2FC) in INTS11 depleted versus undepleted conditions over a region 3kb upstream
746 and downstream of annotated TSSs. This is an average of three biological replicates.

747

748

749

750

751

752

753

754

755

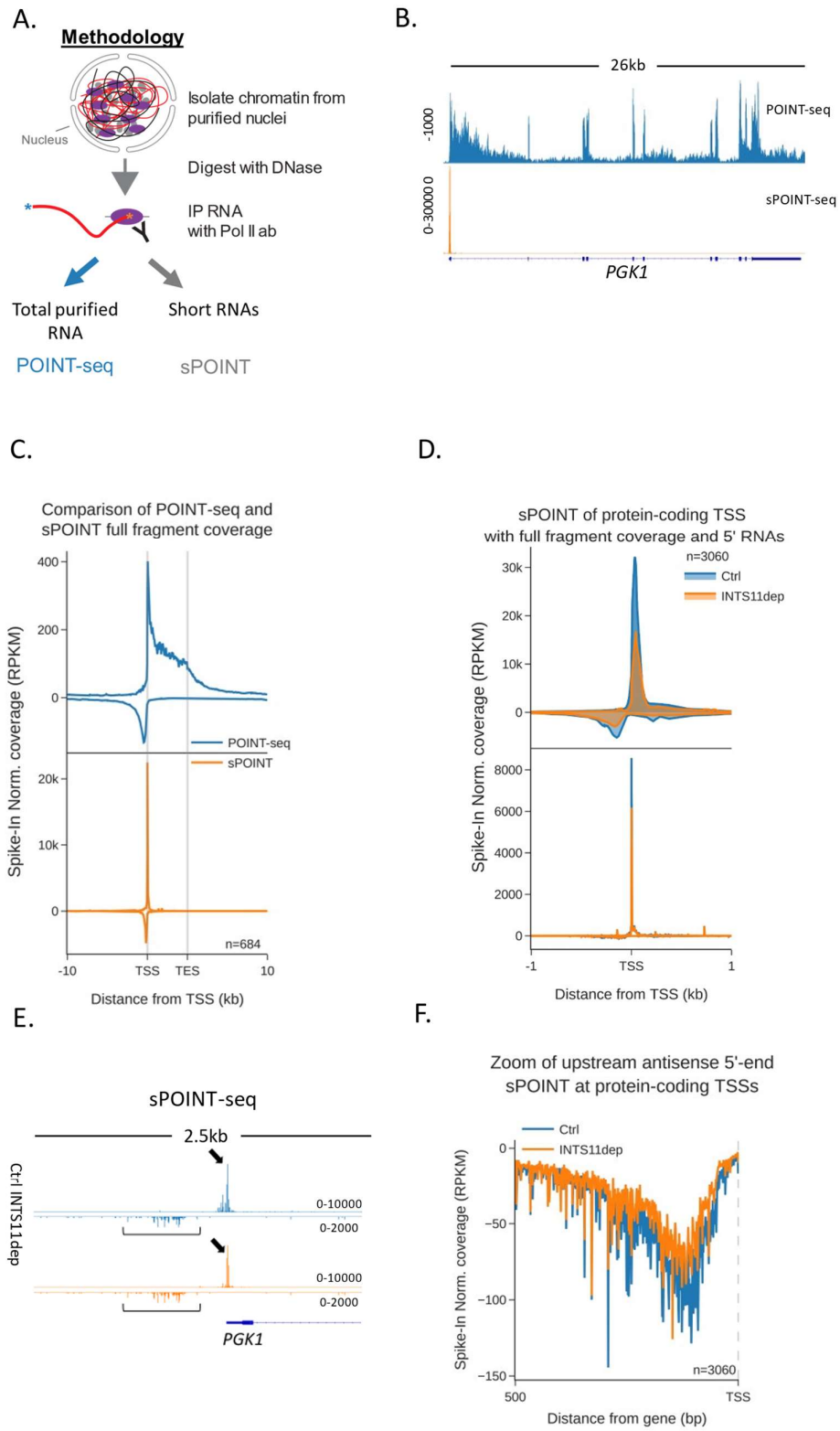
756

757

758

759

FIGURE 3

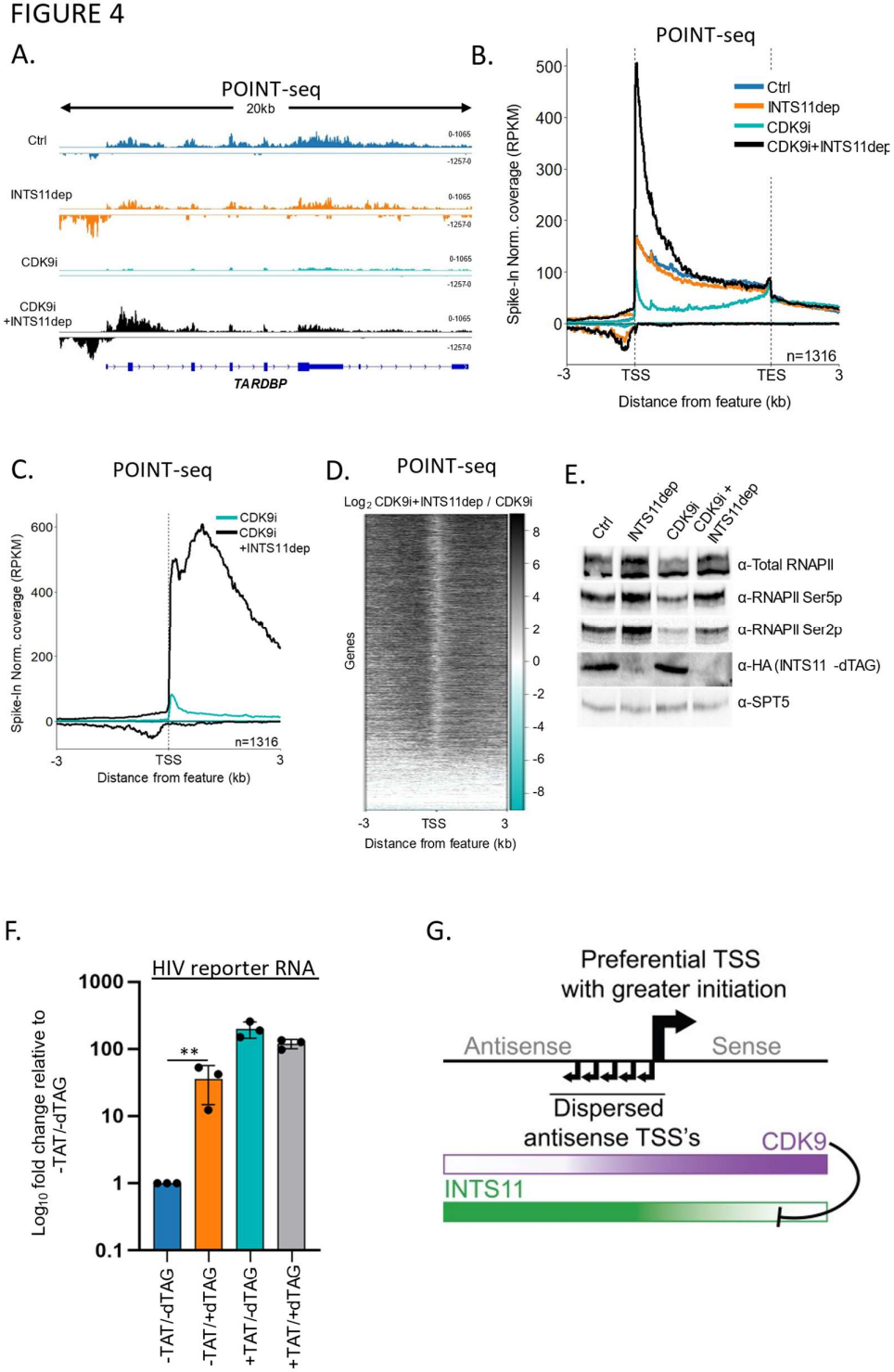

A.
Methodology
Isolate chromatin from purified nuclei
Nucleus
Digest with DNase
IP RNA with Pol II ab
Total purified RNA
Short RNAs
POINT-seq
sPOINT
B.
26kb
POINT-seq
sPOINT-seq
-1000
0-30000 0
PGK1
C.
Comparison of POINT-seq and sPOINT full fragment coverage
Spike-In Norm. coverage (RPKM)
POINT-seq
sPOINT
n=684
-10
TSS
TES
10
Distance from TSS (kb)
D.
sPOINT of protein-coding TSS with full fragment coverage and 5' RNAs
n=3060
Ctrl
INTS11dep
Spike-In Norm. coverage (RPKM)
-1
TSS
1
Distance from TSS (kb)
E.
sPOINT-seq
2.5kb
Ctrl INTS11dep
0-10000
0-2000
PGK1
F.
Zoom of upstream antisense 5'-end sPOINT at protein-coding TSSs
Ctrl
INTS11dep
Spike-In Norm. coverage (RPKM)
n=3060
500
TSS
Distance from gene (bp)

760

761 **FIGURE 3: Sense direction transcription initiation is more efficient and focused.**

762 **A.** Schematic of sPOINT-seq protocol. The POINT-seq protocol is followed, in which
763 chromatin is isolated and engaged RNAPII is immunoprecipitated. Short transcripts
764 are preferentially amplified during library preparation (see Experimental Procedures
765 for full details).
766 **B.** Comparison of POINT-(top trace) and sPOINT-seq (lower trace) on *PGK1*. Y-axis
767 units are RPKM.
768 **C.** Metaplot comparison of POINT- (top plot) and sPOINT-seq (lower plot) profiles across
769 the 684 highest expressed protein-coding that are separated from expressed
770 transcription units by ≥10kb (top ~50%). Signals above and below the x-axis are sense
771 and antisense reads, respectively. Y-axis shows RPKM following spike-in
772 normalisation.
773 **D.** Top metaplot shows full read coverage for sPOINT-seq performed in *INTS11-dTAG*
774 cells treated or not (1.5hr) with dTAGv-1 at the promoters of the top expressed 20% of
775 protein-coding genes. The lower metaplot is the same data but only the 5' end of each
776 read is plotted. The y-axis signals are RPKM following spike-in normalisation. Two
777 biological replicates of sPOINT were performed.
778 **E.** Genome browser track of *PGK1* promoter region in sPOINT-seq. This showcases the
779 focused sense TSS (black arrows) and the dispersed antisense reads (brackets). Note
780 the higher y-axis scale (RPKM) for sense vs. antisense.
781 **F.** Metaplot zoom of the antisense TSS signals deriving from the lower plot in D. This
782 makes clear the dispersed sites of initiation. The Y-axis scale is RPKM following spike-
783 in normalisation. Two biological replicates of sPOINT were performed.
784
785
786
787
788
789
790
791
792

FIGURE 4


A.
POINT-seq
20kb
Ctrl
INTS11dep
CDK9i
CDK9i +INTS11dep
TARDBP
B.
POINT-seq
Ctrl
INTS11dep
CDK9i
CDK9i+INTS11dep
Spike-In Norm. coverage (RPKM)
n=1316
TSS
TES
Distance from feature (kb)
C.
POINT-seq
CDK9i
CDK9i +INTS11dep
Spike-In Norm. coverage (RPKM)
n=1316
TSS
Distance from feature (kb)
D.
POINT-seq
Log2 CDK9i+INTS11dep / CDK9i
Genes
TSS
Distance from feature (kb)
E.
Ctrl
INTS11dep
CDK9i
CDK9i + INTS11dep
α-Total RNAPII
α-RNAPII Ser5p
α-RNAPII Ser2p
α-HA (INTS11 -dTAG)
α-SPT5
F.
HIV reporter RNA
Log10 fold change relative to -TAT/-dTAG
**
-TAT/-dTAG
-TAT/+dTAG
+TAT/-dTAG
+TAT/+dTAG
G.
Preferential TSS
with greater initiation
Antisense
Sense
Dispersed
antisense TSS's
CDK9
INTS11

793

794 **FIGURE 4: Some promoter directionality is lost when CDK9 is inhibited.**

795 **A.** Genome browser track of *TARDBP* in POINT-seq data derived from *INTS11-dTAG*
796 cells either untreated, dTAG treated, NVP-2 treated or dTAG and NVP-2 treated (2.5
797 hr). Signals above and below the Y-axis are sense and antisense reads, respectively.
798 The Y-axis scale shows RPKM following spike-in normalisation.

799 **B.** Metaplot of POINT-seq in *INTS11-dTAG* cells depleted or not of INTS11 and treated
800 or not with NVP-2 (2.5 hr). This uses the same gene set as Figure 1C. The regions
801 3kb upstream and downstream of genes are included. Y-axis units are RPKM following
802 spike-in normalisation. This is an average of two biological replicates.

803 **C.** Metaplot of the same CDK9i + and -dTAG data shown in B but zoomed into the region
804 3kb upstream and downstream of the TSS. This is an average of three biological
805 replicates.

806 **D.** Heatmap representation of the data in C, which displays signal as a log2 fold change
807 (log2FC) in INTS11 depleted versus un-depleted conditions covering a region 3kb
808 upstream and downstream of the TSS. This is an average of three biological replicates.

809 **E.** Western blot for total RNAPII and RNAPII phosphorylated on Ser2/5 (Ser2/5p) in
810 *INTS11-dTAG* cells treated or not with dTAGv-1 and/or NVP-2 (as indicated). dTAG-
811 mediated INTS11 depletion is shown in the anti-HA blot and SPT5 us used as a loading
812 control.

813 **F.** qRT-PCR analysis of *INTS11-dTAG* cells transfected with the HIV reporter construct
814 with or without TAT then depleted or not of INTS11. Quantitation shows signals
815 relative to those obtained in the presence of INTS11 and the absence of TAT after
816 normalising to MALAT1 RNA. n=3. Error bars show standard deviation. ** denoted
817 p=0.01. Note that INTS11 depletion was performed concurrently with transection (14hr
818 in total).

819 **G.** Model for promoter directionality depicting higher levels of focused transcriptional
820 initiation in the sense direction together with opposing gradients of CDK9 and INTS11
821 activity that peak in sense and antisense directions, respectively.

822

823

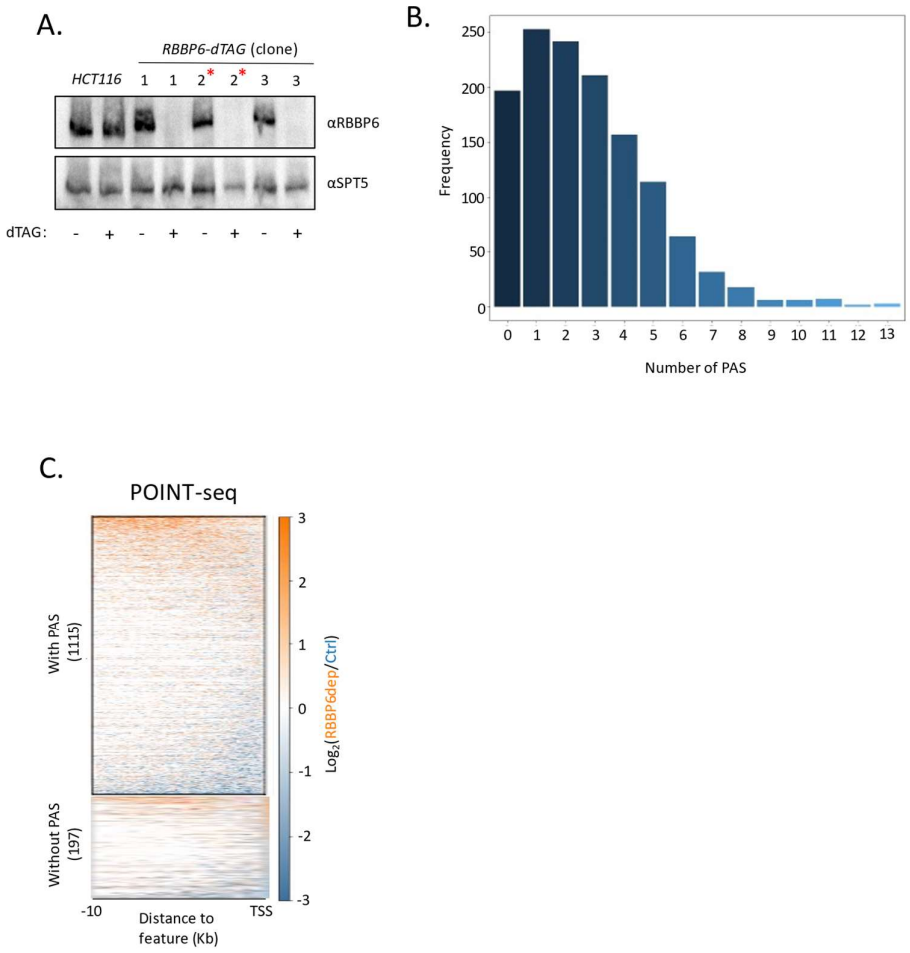
824

825

826

827

SUPPLEMENTAL FIGURE 1, related to figure 1



828

829 **SUPPLEMENTAL FIGURE 1, related to figure 1.**

830 **A.** Western blot showing three separate *dTAG-RBBP6* cell clones. In each case,
831 homozygous tagging is demonstrated by the size shift versus endogenous RBBP6
832 (HCT116 lanes). Tagged RBBP6 is completely depleted by 2hr treatment with dTAGv-
833 1 whereas endogenous RBBP6 is unaffected. SPT5 serves as a loading control.
834 Clone number 2 (red asterisk) was selected for the experiments in Figure 1.

835 **B.** Graph plotting the number of AAUAAA sequences in antisense transcripts derived
836 from a 3kb window upstream of TSSs. Y-axis is the number of transcripts, and the x-
837 axis shows the AAUAAA count per transcript.

838 **C.** Heatmap of control or RBBP6-depleted POINT-seq showing the RBBP6 effect on
839 antisense transcripts without an AAUAAA or those that contain at least one AAUAAA.
840 Most are unaffected by RBBP6 loss. The y-axis is RPKM.

841

842

843

844

845

846

847

848

849

850

851

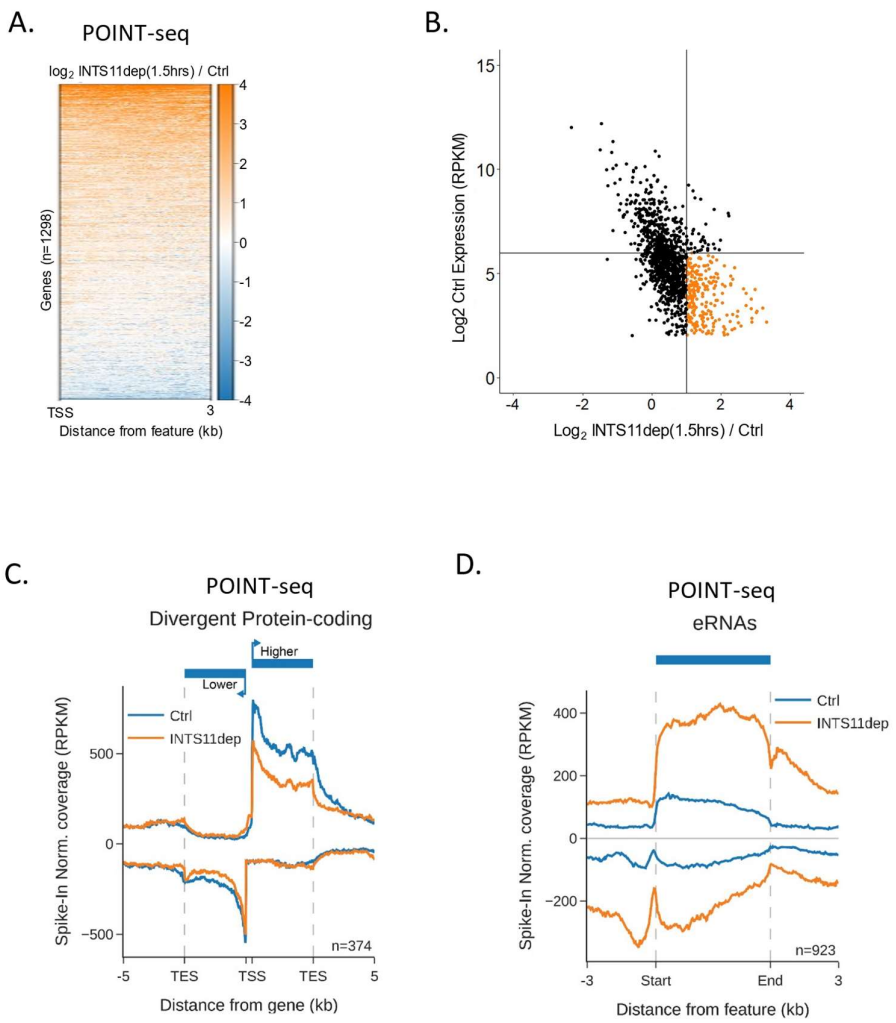
852

853

854

855

SUPPLEMENTAL FIGURE 2, related to figure 2



856

857 **SUPPLEMENTAL FIGURE 2, related to figure 2.**

858 A. Heatmap analysis of the 1316 gene set used in Figure 1C but focused on promoter-
859 proximal transcription of the sense direction. Note the change in scale (Log2FC -
860 INTS11/+INTS11) to reflect the smaller effect sizes vs. those for antisense
861 transcription from the same gene set.
862 B. Comparison of INTS11-dependent changes in promoter-proximal protein-coding
863 transcription with the expression level of each gene using the 1316 gene set as Figure
864 1C. X-axis shows Log2FC in levels (-INTS11/+INTS11) and y axis shows expression
865 level. Note that genes with the largest increase in promoter-proximal signal following
866 INTS11 loss are low expressed (coloured orange). Data derives from POINT-seq
867 following treatment or not with dTAGv-1 (1.5hr).
868 **C.** Metaplot of POINT-seq data from *INTS11-dTAG* cells treated or not (1.5hr) with
869 dTAGv-1. These are protein-coding genes arranged head-to-head. Signals above
870 and below the x-axis are sense and antisense reads, respectively. Y-axis shows
871 RPKM following spike-in normalisation.
872 **D.** Metaplot of POINT-seq data from *INTS11-dTAG* cells treated or not (1.5hr) with
873 dTAGv-1. This shows RNAs derived from enhancer clusters, which generally initiate
874 transcription bidirectionally. Signals above and below the x-axis are sense and
875 antisense reads, respectively. Y-axis shows RPKM following spike-in normalisation.

876

877

878

879

880

881

882

883

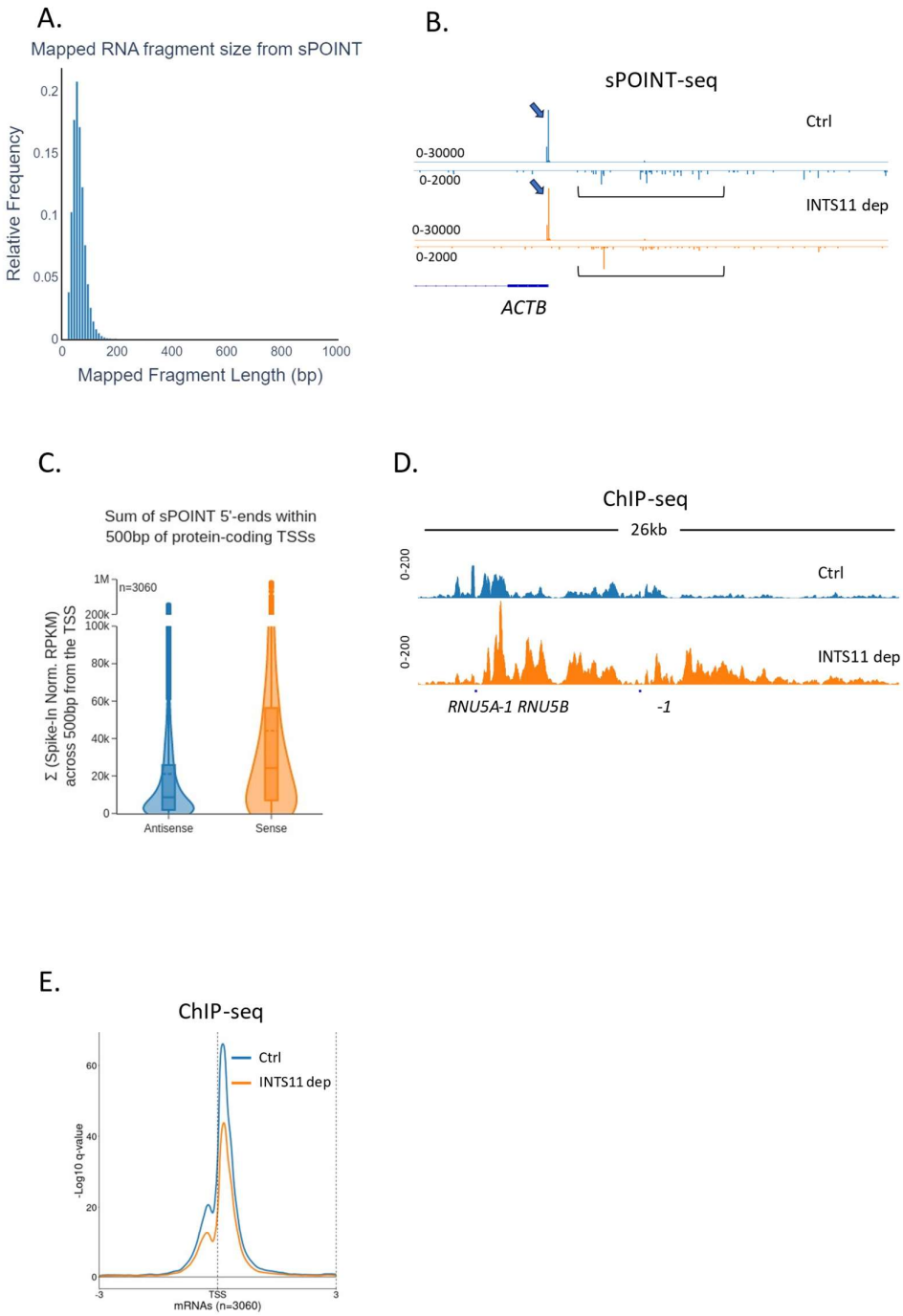
884

885

886

887

SUPPLEMENTAL FIGURE 3, related to figure 3



888

889 **SUPPLEMENTAL FIGURE 3, related to figure 3.**

890 **A.** Plot showing the size distribution of fragments mapped by sPOINT-seq. This
891 demonstrates its selectivity toward transcripts <150nts.
892 **B.** Genome browser track of *ACTB* promoter region in sPOINT-seq from *INTS11-dTAG*
893 cells treated or not with dTAGv-1 (1.5 hr). This showcases the focused sense TSS
894 (black arrows) and the dispersed antisense reads (brackets). Note the different y-axis
895 scales (RPKM) for sense vs. antisense.
896 **C.** Violin plot of the sum sPOINT 5'end spike-in normalised RPKM signal across a 500bp
897 window from the protein-coding TSSs shown in Figure 3D (n=3060). These samples
898 are untreated with dTAG to quantitate the normal levels of sense vs. antisense
899 initiation/pausing in a window -/+ 500bp from the TSS in respective directions for
900 antisense and sense. This shows more initiation/pausing in sense vs. antisense
901 directions.
902 **D.** Genome track of *RNU5A-1* and *RNU5B-1* in RNAPII ChIP-seq performed on *INTS11-*
903 *dTAG* cells treated or not with dTAGv-1 (2hr). INTS11 depletion causes RNAPII build-
904 up beyond both genes indicating a transcription termination defect. The Y-axis scale
905 is RPKM.
906 **E.** Metaplot of RNAPII occupancy of protein-coding TSS regions (the same gene set
907 employed for the sPOINT analysis in Figure 3) derived from RNAPII ChIP-seq
908 performed on *INTS11-dTAG* cells treated or not with dTAGv-1 (2hr). The Y-axis scale
909 is log10 qvalue of peak pileup values normalised to spike in control.

910

911

912

913

914

915

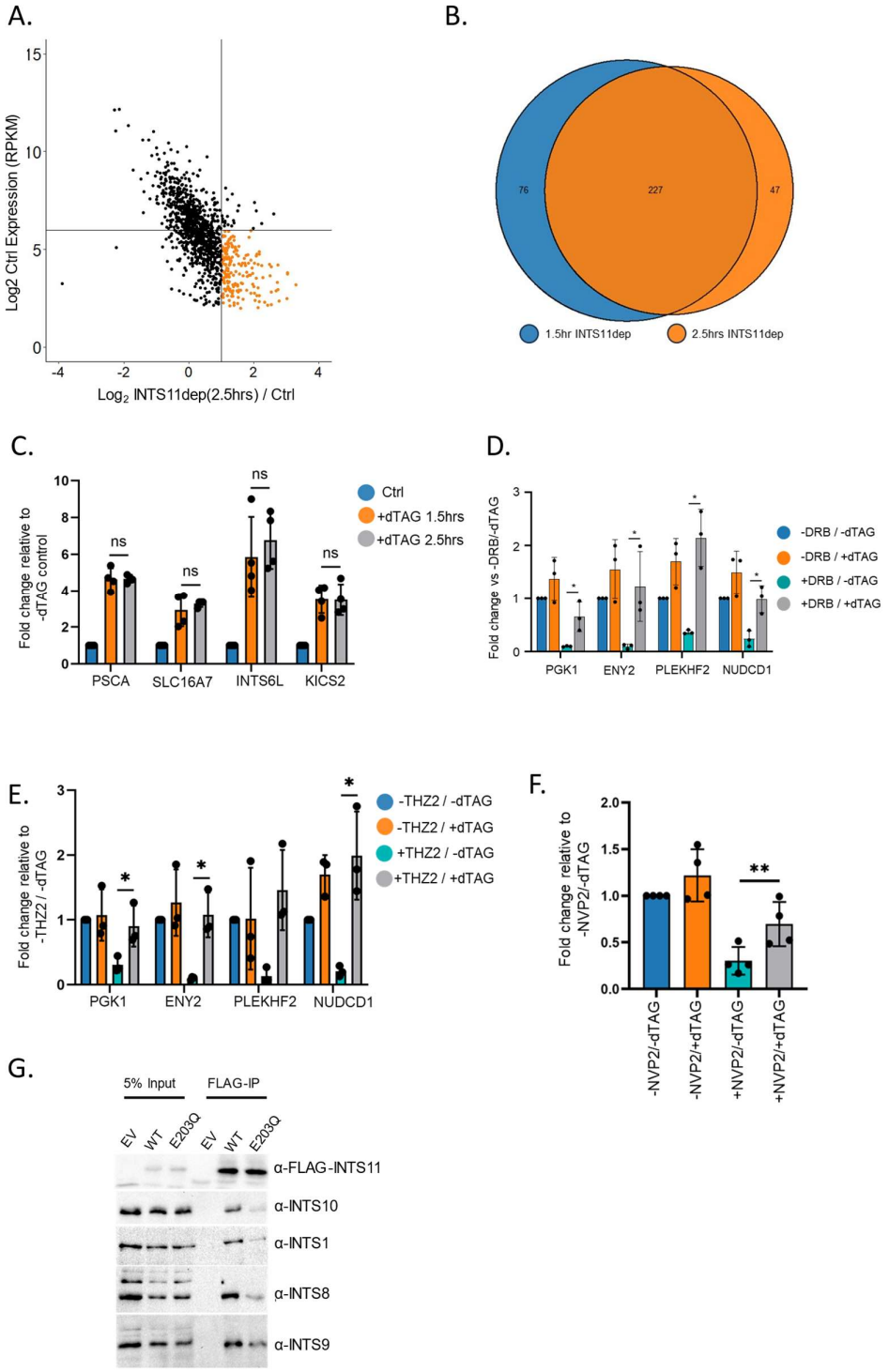
916

917

918

919

SUPPLEMENTAL FIGURE 4, related to figure 4



920

921 **SUPPLEMENTAL FIGURE 4, related to figure 4.**

922 A. Comparison of INTS11-dependent changes in promoter-proximal protein-coding
923 transcription with the expression level of each gene. The gene set is equivalent to the
924 analogous panel in Supplemental Figure 2B. The x-axis shows Log2FC in levels (-
925 INTS11/+INTS11) and the y-axis shows expression level. Note that genes with the
926 largest increase in promoter-proximal signal following INTS11 loss are low expressed
927 (coloured orange). Data derives from POINT-seq following treatment or not with
928 dTAGv-1 (2.5hr).

929 **B.** Venn diagram showing the number of protein-coding genes where promoter-proximal
930 POINT-seq signal increases by ≥Log2FC of 1 after INTS11 loss after 1.5h or 2.5h
931 dTAGv-1 treatment. A strong overlap is seen indicating reproducible effects at the two
932 time points.

933 **C.** qRT-PCR analysis of PSCA, SLC16A7, INTS6L, and KICS2 pre-mRNAs in *INTS11-*
934 *dTAG* cells treated or not with dTAGv-1 for 1.5hr or 2.5hr. To enrich nascent
935 transcripts, primers detect intronic RNA. Quantitation shows fold change versus
936 spliced actin relative to untreated samples. Error bars show standard deviation. n=4.
937 n.s denotes not significant.

938 **D.** qRT-PCR analysis of PGK1, ENY2, PLEKHF2, and NUDCD1 pre-mRNAs in *INTS11-*
939 *dTAG* cells treated or not with dTAGv-1 and at the same time exposed or not to DRB
940 (all 2.5hr). To enrich nascent transcripts, primers detect intronic RNA. Quantitation
941 shows fold change versus spliced actin relative to samples untreated with dTAGv-1 or
942 DRB. DRB treatment substantially reduces signal, which is restored when INTS11 is
943 co-depleted. Error bars show standard deviation. n=3, * denotes p≤0.05.

944 **E.** As for C but following treatment with the CDK7 inhibitor, THZ1, and/or depletion of
945 INTS11 with dTAGv-1 (all treatments 2.5hr). Error bars show standard deviation. n=3,
946 * denotes p≤0.05.

947 **F.** qRT-PCR analysis of chromatin-associated (nascent) RNA isolated from *INTS11-*
948 *dTAG* cells transfected with the HIV reporter construct together with TAT then treated
949 or not with INTS11 and/or NVP-2. Quantitation shows signals relative to those
950 obtained in untreated cells after normalising to MALAT1 RNA. n=4. Error bars show
951 standard deviation. **denotes p≤0.01.

952 **G.** Flag immunoprecipitation performed in *INTS11-dTAG* cells transfected with flag-
953 tagged wild-type or E203Q INTS11. The gel shows input and co-precipitated samples
954 probed for components of different Integrator modules (tail module – INTS10;
955 backbone module – INTS1; phosphatase module – INTS8 and cleavage module –
956 INTS9). These components are poorly associated with E203Q INTS11 vs. wild type.