Integrator mediates the biogenesis of enhancer RNAs

Fan Lai1*, Alessandro Gardini1, Anda Zhang3 & Ramin Shiekhattar1

Integrator is a multi-subunit complex stably associated with the carboxy-terminal domain (CTD) of RNA polymerase II (RNAPII)1. Integrator is endowed with a core catalytic RNA endonuclease activity, which is required for the 3′-end processing of non-polyadenylated, RNAPII-dependent, uridylate-rich, small nuclear RNA genes1. Here we examine the requirement of Integrator in the biogenesis of transcripts derived from distal regulatory elements (enhancers) involved in tissue- and temporal-specific regulation of gene expression in metazoans2–8. Integrator is recruited to enhancers and super-enhancers in a stimulus-dependent manner. Functional depletion of Integrator subunits diminishes the signal-dependent induction of enhancer RNAs (eRNAs) and abrogates stimulus-induced enhancer–promoter chromatin looping. Global nuclear run-on and RNAPII profiling reveals a role for Integrator in 3′-end cleavage of eRNA primary transcripts leading to transcriptional termination. In the absence of Integrator, eRNAs remain bound to RNAPII and their primary transcripts accumulate. Notably, the induction of eRNAs and gene expression responsiveness requires the catalytic activity of Integrator complex. We propose a role for Integrator in biogenesis of eRNAs and enhancer function in metazoans.

To assess the role for Integrator in the biogenesis of eRNAs, we examined the signal-dependent recruitment of Integrator complex to enhancer sites. HeLa cells were starved of serum for 48 h, after which they were stimulated with epidermal growth factor (EGF) to induce immediate early genes (IEGs). We identified 2,029 enhancers based on their occupancy by RNAPII, CBP/p300 and containing acetylated histone H3 lysine 27 (H3K27ac) chromatin modification (see Methods). We found that while assessing steady-state levels of eRNAs provided a measure of EGF-induced eRNAs, we obtained a better read-out of eRNAs after sequencing of the chromatin-enriched RNA fractions (ChromRNA-seq)9. We focused on 91 enhancers that displayed EGF-eRNAs after sequencing of the chromatin-enriched RNA fractions—eRNAs following a measure of EGF-induced eRNAs, we obtained a better read-out of eRNAs after sequencing of the chromatin-enriched RNA fractions (ChromRNA-seq)9. We focused on 91 enhancers that displayed EGF-induced eRNAs in the proximity of EGF-responsive genes following stimulus-dependent recruitment of the Integrator complex at EGF-responsive enhancers.

To examine the functional importance of Integrator at enhancers and its role in the biogenesis of eRNAs, we developed HeLa clones expressing doxycycline-inducible short hairpin RNAs (shRNAs) against INTS11 and INTS1 subunits of the Integrator complex (Extended Data Fig. 3b). Within the time course of these experiments the mature levels of small nuclear RNAs (snRNAs) were not perturbed (data not shown). Twenty minutes of EGF stimulation resulted in the induction of bi-directional eRNAs similar to previous reports (Fig. 1a, f and Extended Data Fig. 1c–h)5,8–11. Depletion of INTS11 diminished the eRNA induction after EGF stimulation (Fig. 1f; as shown at two enhancer loci; enhancers were named after their proximity to an EGF-responsive gene). The fold induction of eRNAs at all EGF-induced enhancers decreased significantly (Fig. 1g, h). We also observed a significant decrease in the transcriptional induction of EGF-responsive protein-coding genes in the proximity of these EGF-induced enhancers (Fig. 1g, h). Interestingly, there was a subtle increase (statistically not significant) in H3K27 acetylation at enhancers following EGF stimulation, which was reduced after Integrator depletion (Fig. 1f and Extended Data Fig. 3c).

To gain further insight into quantitative changes in eRNAs following depletion of Integrator, we depleted INTS11 or INTS1 and performed a time-course analysis of eRNA induction using specific primer sets for each strand. Depletion of either Integrator subunit diminished the EGF-induced increase in eRNA levels from both strands of the enhancers (Extended Data Fig. 4a, b). Analysis of regulatory landscape in the proximity of the EGF-responsive gene ATF3 (activating transcription factor 3) revealed the presence of clusters of acetylated H3K27 and p300 binding sites similar to that described for super-enhancers12–14 (Extended Data Fig. 4c). This region also displayed occupancy by RNAPII at multiple sites, and we could detect additional recruitment of RNAPII and Integrator to these sites following EGF stimulation (Extended Data Fig. 4c). Analysis of eRNA synthesis using strand-specific RNA-seq and real-time PCR (during a time-course experiment) demonstrated a requirement for Integrator in the induction of eRNAs at the super-enhancer sites after EGF stimulation (Extended Data Fig. 4d). Collectively, these results highlight a requirement for Integrator in stimulus-dependent induction of eRNAs from individual enhancers and enhancer clusters.

An important component of enhancer function is the formation of stimulus-dependent chromatin looping, allowing enhancer and promoter communication15–17. We measured chromatin looping between NR4A1 and DUSP1 enhancers and their respective promoters using chromosome conformation capture (3C) following stimulation with EGF (Fig. 2a). We observed a robust association between the enhancer and the promoter regions of NR4A1 and DUSP1 after EGF stimulation (Fig. 2b). Remarkably, depletion of Integrator abrogated the EGF-induced chromatin looping without any effect on non-stimulus-induced chromosomal interactions (Fig. 2b, c and Extended Data Fig. 5a, b). These results demonstrate that Integrator regulates...
RNAPII for eRNAs; from genes (mean density was calculated as follows: 6 kb surrounding the peak of Integrator occupancy at enhancers and their corresponding protein-coding genes (doxycycline (dox)) markedly reduces steady-state levels of eRNAs as measured by total RNA-seq). Data were obtained using a tet-inducible shRNA system, stably transduced in HeLa cells. Acetylation of H3K27 is also shown. g, h. Average expression levels of 91 eRNAs and their neighbouring (<500 kb) 57 protein-coding genes indicate a significant impairment of activation. Box plots represent the expression fold change (log2) before and after EGF treatment in normal conditions (ctrl) and upon depletion of Integrator (dox) (t-test, \( P < 0.0005 \) for all panels). Fold change of RPKM (reads per kilobase of exon per million mapped reads) values was calculated from RNA-seq (f) and ChromRNA-seq (g) data.

To gain an insight into the mechanism by which Integrator regulates enhancer function and eRNA biogenesis, we depleted Integrator and performed RNAPII profiling and global nuclear run-on followed by high-throughput sequencing (GRO-seq) after EGF induction. Notably, Integrator depletion resulted in the increase and spreading of GRO-seq reads throughout the body of eRNA transcripts at both enhancers and super-enhancers, which was mirrored by a concomitant increase and spreading of RNAPII localization (Fig. 3a, b). Indeed, the average profile of depth-normalized reads of 91 EGF-induced enhancers showed a significant accumulation of GRO-seq and RNAPII ChIP-seq reads (Extended Data Fig. 6a, b). Analysis of RNAPII travelling ratio, a measure of RNAPII productive elongation, revealed that in contrast to EGF-responsive protein coding genes, which experience a block in productive elongation after Integrator depletion\(^{18}\), there is increased RNAPII occupancy in the body of eRNA transcripts (Extended Data Fig. 6c, d). The accumulation of RNAPII at eRNA loci after Integrator depletion occurred despite the decreased recruitment of super elongation complex (SEC) to enhancers (Extended Data Fig. 7a, b).

Figure 1 | Integrator mediates induction of eRNAs. a, EGF induction of an enhancer in the vicinity of the NR4A1 gene (see Extended Data Fig. 1i). RNAPII and INTS11 are recruited to the enhancer after 20 min of stimulation and eRNAs are transcribed bi-directionally from the locus (as revealed by deep sequencing of chromatin-associated RNA, ChromRNA-seq). The y axis represents the read counts normalized to sequencing depth. b, Average profile of Integrator recruitment to 91 EGF-responsive enhancers. TSS indicates transcription start site. The y axis shows the average of read density. c, Increased Integrator occupancy at enhancers and their corresponding protein-coding genes (mean density was calculated as follows: 6 kb surrounding the peak of RNAPII for eRNAs; from −0.5 kb to +2.5 kb for coding genes; \( P < 0.001 \)). Whiskers on the box plot indicate the variability in the datasets. d, Average profile of RNAPII upon EGF treatment at enhancers. e, Increased RNAPII occupancy following EGF stimulation at enhancers and their corresponding protein-coding genes (\( P < 0.005 \)). f, Inducible knockdown of INTS11 (doxycycline (dox)) markedly reduces steady-state levels of eRNAs as (as measured by total RNA-seq). Data were obtained using a tet-inducible shRNA system, stably transduced in HeLa cells. Acetylation of H3K27 is also shown. g, h. Average expression levels of 91 eRNAs and their neighbouring (<500 kb) 57 protein-coding genes indicate a significant impairment of activation. Box plots represent the expression fold change (log2) before and after EGF treatment in normal conditions (ctrl) and upon depletion of Integrator (dox) (t-test, \( P < 0.0005 \) for all panels). Fold change of RPKM (reads per kilobase of exon per million mapped reads) values was calculated from RNA-seq (f) and ChromRNA-seq (g) data.

The increased RNAPII occupancy at eRNA loci suggests a block in 3’-end cleavage of primary eRNA transcripts, leading to a defect in termination. To quantitate such a 3’-end cleavage defect, we measured the accumulation of primary levels (or unprocessed levels) of eRNA transcripts after Integrator depletion using semi-quantitative PCR and real-time PCR. We observed a 3- to 10-fold accumulation of unprocessed eRNA transcripts concomitant with the reduction of the processed eRNA levels (Fig. 3c–e and Extended Data Fig. 8a). Previous experiments revealed that the loss of 3’-end cleavage by Integrator led to increased levels of polyadenylated U snRNA transcripts, which are normally not polyadenylated\(^{18}\). Indeed, analysis of the polyadenylated transcripts revealed a robust increase in polyadenylation of eRNAs in the absence of Integrator (Fig. 3f, g). These results attest to Integrator cleavage of the 3’ end of eRNAs leading to a termination of transcription.
We surmised that such a termination defect might result in the inability of RNAPII to dissociate from the eRNAs, leading to accumulation of RNAPII–eRNA complexes and a consequent decrease in mature eRNA levels. We performed ultraviolet (UV) cross-linking followed by RNA immunoprecipitation (UV-RIP) using antibodies against RNAPII to examine increased association of eRNAs with RNAPII after depletion of Integrator. Consistent with a role for Integrator in the processing of eRNAs, depletion of Integrator led to a profound increase in eRNA

**Figure 2** Integrator is required for enhancer–promoter interaction. a, Diagrams of NR4A1 (left) and DUSP1 (right) genomic regions with their respective enhancers (shown in red). The arrowheads depict the position of primers for detection of chromatin looping and the stick bars indicate enzyme digestion sites (named N1–6 and D1–5). E refers to the anchor primer at the enhancer sites; control sites are also indicated. b, Looping events between the promoter region of NR4A1 and its enhancer were detected at N3, N4 and N5 sites after EGF induction (left). A similar interaction was also captured between sites D3 and D4 of DUSP1 promoter and its downstream enhancer after EGF induction (right). c, Knockdown of Integrator abolished chromosomal looping events at both NR4A1 and DUSP1 sites. The interaction frequency between the anchoring points and the distal fragments were determined by real-time PCR and normalized to BAC templates. All sites were assayed in three independent experiments (P < 0.01, two-sided t-test). Control anchors are displayed in Extended Data Fig. 5.

**Figure 3** Integrator has a role in termination of eRNAs. a, b, RNAPII dynamics was analysed by ChIP-seq and GRO-seq at the enhancer regions adjacent to NR4A1 and DUSP1 (a) and at the super-enhancer upstream of DUSP5 (b). The y axis represents the read counts normalized to sequencing depth. c, 3'-end cleavage of eRNAs was examined with semi-quantitative PCR. Primer pairs were designed to amplify a portion of the enhancer transcript as detected in the control GRO-seq experiment (t, total) or a longer template further extending into the 3’ of the enhancer region (u, unprocessed). d, PCR analysis was performed in two independent replicates, before (ctrl) and after (dox) depletion of INTS11 at three eRNAs (sense and antisense strand). e, The housekeeping gene GUSB was used as a cDNA loading control. f, Polyadenylation of eRNAs increases after depletion of Integrator at DUSP1 and CCNL1 enhancer loci. The polyadenylated fraction of RNA from whole-cell lysates was sequenced after EGF stimulation, before and after depletion of INTS11 at three eRNAs (sense and antisense strand). g, Box plot shows significant increase in polyadenylated RNA reads (P < 0.001) across the entire set of EGF responsive enhancers. Whiskers on the box plot indicate the variability in the datasets.
engagement with RNAPII following induction with EGF (Extended Data Fig. 8b–d). We found similar results after analysis of RNAPII interaction with the eRNAs at the ATF3 super-enhancer (Extended Data Fig. 8e–g). Taken together, these results implicate the Integrator complex in the termination of eRNAs and highlight Integrator’s role in the release of eRNA transcripts from transcribing RNAPII.

The catalytic subunit of Integrator is composed of the heterodimer of INTS11 and INTS9 enzymes with close homology to CPSF73 and CPSF100, respectively²⁰. We previously showed that a single point mutation (E203Q) in the catalytic domain of INTS11 leads to impaired processing of small nuclear RNAs. To assess the impact of INTS11 enzymatic activity on eRNA biogenesis, we developed wild type and mutant INTS11 (E203Q) that would be refractory to the action of shRNAs against INTS11, and used these constructs to perform rescue experiments. While ectopic expression of wild-type INTS11 could substantially rescue the EGF-induced eRNA levels after depletion of INST11, the single-point catalytic mutant was without any effect (Fig. 4a and Extended Data Fig. 9a). Interestingly, we observed a similar rescue of the transcriptional activation of EGF-induced genes by the wild-type INTS11 and not its catalytic mutant (Fig. 4b). These results not only demonstrate the requirement of INTS11 catalytic activity in regulating the induction of eRNAs but also highlight the defect in eRNA processing as a contributing factor in the loss of transcriptional responsiveness.

To determine the scope of Integrator function on active enhancers we analysed the 2,029 transcriptionally active enhancers in HeLa cells.

Figure 4 | Integrator has a global role in enhancer regulation. a, Ectopic expression of wild-type INTS11, and not its catalytic mutant (E203Q), following Integrator depletion can rescue eRNA induction by EGF. A similar rescue was observed for wild-type INTS11 on the target protein-coding genes. Real-time PCR analysis was performed on CCNL1 and DUSP1 eRNAs and their corresponding mRNAs before and after EGF stimulation. Each eRNA was assayed with two sets of primers. Error bars represent ± s.e.m. (n = 3 biological independent experiments). **P < 0.01 by two-sided t-test. b, The heat map showcases 2,029 enhancer regions identified using RNAPII extragenic loci enriched in H3K27 acetylation (see Methods). Enhancers were centred at the middle of the RNAPII peak and ranked by transcription activity (GRO-seq). The distribution of p300 and H3K27ac are consistent with a group of active enhancers. Upon Integrator depletion, nascent RNA reads and RNAPII profiles spread beyond the normal 3’ end of eRNAs. c, Model for the role of Integrator at eRNAs. Stimulation of serum-starved cells with EGF triggers recruitment of RNAPII and Integrator to enhancer sites and induces bi-directional transcription of non-polyadenylated eRNAs. Upon EGF stimulation Integrator navigates the enhancers along with RNAPII to promote endonucleolytic cleavage of nascent transcripts, leading to release of the mature eRNAs. Depletion of Integrator elicits a cleavage defect leading to faulty termination, which results in extended eRNA transcripts and accumulation of RNAPII.
We ranked the enhancers based on their transcriptional activity, which mirrored that of RNAPII occupancy (Fig. 4c). Notably, depletion of Integrator resulted in processing defects at all active enhancers, as reflected by the broadening of GRO-seq and RNAPII ChIP-seq reads commensurate with the transcriptional activity of each enhancer site (Fig. 4c). This was in contrast to GRO-seq and RNAPII profiles at transcriptionally active protein-coding genes (Extended Data Fig. 9b). These results demonstrate the generality of Integrator in the processing of eRNAs at enhancers (Fig. 4d).

Recent genome-wide studies have revealed the presence of RNAPII at active enhancers coincident with expression of each enhancer, and these sections appear only in the online paper. Indeed, we found that Integrator is the molecular machine that is recruited to enhancers in a signal-dependent manner and is required for the induction of eRNAs. We surmise that the defect in 3′-end processing following Integrator depletion leads to a termination defect reflected in increased levels of primary eRNA transcripts. It is also likely that Integrator affects the stability of the mature transcripts, since its depletion leads to changes in steady-state levels of mature eRNAs.

Similar to other regulatory complexes, Integrator is also recruited to the promoters of protein-coding genes including IEGs. Importantly, such eRNAs have been shown to have critical roles in transcriptional induction by a variety of signal transduction pathways. We show that Integrator is the molecular machine that is recruited to enhancers in a signal-dependent manner and is required for the induction of eRNAs. We surmise that the defect in 3′-end processing following Integrator depletion leads to a termination defect reflected in increased levels of primary eRNA transcripts. It is also likely that Integrator affects the stability of the mature transcripts, since its depletion leads to changes in steady-state levels of mature eRNAs.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 March 2014; accepted 7 July 2015.

Published online 26 August 2015.

1. Bailat, D. et al. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell 123, 265–276 (2005).
2. Wang, K.C. et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472, 120–124 (2011).
3. Ørom, U.A. et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 143, 46–58 (2010).
4. De Santa, F. et al. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 8, e1000384 (2010).
5. Kim, T.K. et al. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187 (2010).
6. Bhatt, D.M. et al. Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell 150, 279–290 (2012).
7. Lam, M.T. et al. Rev-erbα represses macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498, 511–515 (2013).
8. Li, W. et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498, 516–520 (2013).
9. Wang, D. et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474, 390–394 (2011).
10. Sigova, A.A. et al. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc. Natl Acad. Sci. USA 110, 2876–2881 (2013).
11. Hahn, N. et al. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 145, 622–634 (2011).
12. Whyte, W.A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013).
13. Lovén, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 325–334 (2013).
14. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. Cell 155, 934–943 (2014).
15. Sanyal, A. et al. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 145, 622–634 (2011).
16. Melo, C.A. et al. eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol. Cell 49, 524–535 (2013).
17. Moussavi, K. et al. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol. Cell 51, 606–617 (2013).
18. Gardini, A. et al. Integrator regulates transcriptional initiation and pause release following activation. Mol. Cell 56, 128–139 (2014).
19. Yamamoto, J. et al. DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. Nat. Commun. 5, 4263 (2014).
20. Albrecht, T.R. & Wagner, E.J. snRNA 3′ end formation requires heterodimeric association of integrator subunits. Mol. Cell. Biol. 32, 1112–1123 (2012).
21. Koch, F. et al. Transcription initiation platforms and GTI recruitment at tissue-specific enhancers and promoters. Nature Struct. Mol. Biol. 18, 956–963 (2011).
22. Yang, L. et al. IncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature 500, 598–602 (2013).
23. Stadlmayer, B. et al. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. Nat. Commun. 5, 5531 (2014).
24. Skaar, J.R. et al. The Integrator complex controls the termination of transcription at diverse classes of gene targets. Cell Res. 25, 288–305 (2015).
25. Schaukowitch, K. et al. Enhancer RNA facilitates NELF release from immediate early genes. Mol. Cell 56, 29–42 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank J. M. Marinis and M. A. Lazar for technical support. We thank D. Hsu in A. Shilatifard’s laboratory for performing the SEC-ChIP-seq experiments. We thank the Oncogenomics core facility at the Sylvester Comprehensive Cancer Center for performing high-throughput sequencing. We also thank Shiekhattar laboratory members and P.-J. Hamard for support and discussions. This work was supported by funds from University of Miami Miller School of Medicine, Sylvester Comprehensive Cancer Center and grants R01 GM078455 and R01 GM105754 (R.S.) from the National Institute of Health.

Author Contributions F.L. and A.G. are co-first authors. R.S., F.L. and A.G. conceived and designed the overall project. F.L., A.G. and A.Z. performed the experiments. R.S., F.L. and A.G. analysed the data and wrote the paper.

Author Information High-throughput data are deposited at the Gene Expression Omnibus (GEO) under accession number GSE68401. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.S. (rshekhattar@med.miami.edu).
METHODS

No statistical methods were used to predetermine sample size.

Genome-wide data. High-throughput sequencing data analysed in this study are originally described in ref. 18 and are deposited at the Gene Expression Omnibus with accession number GSE40632.

H3K27ac, H3K4me1 and p300 data sets from HeLa-S3 cells are available as part of the ENCODE project and can be retrieved under the following accession numbers: GSM373684, GSM978322, GSM393550. Additional experiments are deposited at GEO (GSE68401) and include RNA-seq data (chromatin-bound RNA, polyadenylated and non-polyadenylated fractions of total RNA) as well as ChIP-seq experiments (acetylation of H3K27 and occupancy of NELFA). Every genome-wide experiment is performed in two independent biological replicates.

ChIP-seq data were obtained using HiSeq 2000 and ChIP-seq analysis. Briefly, we centred a 6-kb window at the midst of the RNAPII peak and we used ChIP-seq obtained in HeLa-S3 (GEO GSE31477). The analysis resulted in 2,029 ngsplot30.

Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth. Heat maps were generated with ChAsE (http://chase.cs.univie.ac.at/), using default parameters, a 10 kb window and 400 bins and with sequencing depth.
protein–RNA complex was extracted using TRizol method for RNA extraction and subjected to RT–qPCR with corresponding primers.

**Inducible cell lines.** INTS11 and INTS1 knockdown inducible clones were generated from HeLa cells using the Tet-pLKO-puro vector. For EGF induction, cells were serum starved in 0.5% FBS for 48 h and treated with 100 ng ml\(^{-1}\) EGF (Invitrogen) for the indicated time course. All cell lines in this study are mycoplasma negative.

**Transfections.** Cells were treated with doxycycline for 48 h. 24 h before EGF induction, INTS11 and INTS11 (E203Q) mutant protein expression plasmids were transfected using Lipofectamine 2000 (Life Technologies, Inc.) according to the manufacturer’s instruction. Cells were harvested 0 and 20 min after EGF induction.

All the PCR primer sequences are listed in the supplementary Table 2.

26. Gerstein, M. B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91–100 (2012).
27. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
28. Wang, L. G., Wang, S. Q. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
29. Ye, T. *et al.* seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* **39**, e35 (2011).
30. Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014).
31. Lai, F. *et al.* Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494**, 497–501 (2013).
Extended Data Figure 1 | Identification eRNAs responsive to EGF. a, We identified 91 EGF-responsive enhancer regions in HeLa cells. We annotated extragenic RNAPII sites (see Methods) and used the middle of the RNAPII peak as an anchor to display average profiles of p300, H3K27ac and H3K4me1 (data from the ENCODE project). The profiles represent the mean read density of ChIP-seq data. The 91 loci display a typical enhancer signature, with enrichment of p300 and H3K27ac around the TSS and a broader decoration by H3K4me1. 

b, Profiles of H3K27ac were obtained from ChIP-seq analysis of HeLa cells before and after 20 min of EGF induction. Mean read density was normalized to sequencing depth. c, EGF stimulates bi-directional transcription from 91 enhancer regions. We displayed the mean read density obtained from strand-specific sequencing of the chromatin-bound RNA fraction (ChromRNA-seq). 

d, e, Normalized read density (RPKM) was calculated from RNA-seq data for 91 eRNAs (d) and 57 neighbouring protein-coding genes (e) that responded to EGF stimulation (FC >1.6) and mapped within 500 kb from an EGF-responsive eRNA. f, Average profiles of ChromRNA-seq data at 91 enhancer loci (mean density of reads, normalized to total read number). 

g, h, Box plot of 91 eRNAs before and after treatment with EGF shows the average increase of transcription 20 min after stimulation (P < 0.001), matched by an increase in the neighbouring protein-coding genes (P < 0.02). i, NR4A1 is activated by EGF in HeLa cells: RNAPII and INTS11 are recruited to the NR4A1 locus after 20 min of stimulation, with concomitant accumulation of reads from RNA-seq and ChromRNA-seq. A neighbouring eRNA locus also exhibits increased transcription along with RNAPII and INTS11 recruitment. Sequencing tracks are visualized in BigWig format and aligned to the hg19 assembly of the UCSC Genome Browser. Whiskers on the box plots indicate the variability in the datasets.
Extended Data Figure 2 | EGF-induced eRNAs are predominantly non-polyadenylated. a, We examined transcription at three enhancers adjacent to EGF-responsive genes CCNL1, NR4A1 and DUSP1. Total RNA samples were collected before and after EGF induction. Reverse transcription was performed with random hexamer primer or oligo d(T) primer. Each eRNA strand was analysed by real-time PCR with specific primers. Error bars represent ± standard error of the mean (s.e.m., n = 3 biological independent experiments). *P < 0.01 by two-sided t-test. b, c, RNA-seq was performed on the polyadenylated and non-polyadenylated fraction of total RNA. RNA-seq tracks were visualized in BigWig format and aligned to the hg19 assembly of the UCSC Genome Browser. CCNL1 and DUSP1 enhancers were displayed (b) along with a polyadenylated control (DUSP1 protein-coding locus) and a non-polyadenylated transcript (snRNA U12) (c). All EGF-induced eRNAs and protein-coding genes (RefSeq hg19) were examined for their average RPKM throughout the entire locus. d, We compared polyadenylation levels of 225 eRNAs and 150 protein-coding genes (2 fold induction upon EGF, RPKM calculated from ChromRNA-seq data previously described). The box plot shows predominance of non-polyadenylated transcripts mapping to eRNA loci, as opposed to transcripts coding for RefSeq genes. Whiskers on the box plot indicate the variability in the datasets.
Extended Data Figure 3 | The Integrator complex is recruited to enhancers upon EGF stimulation. a, qChIP analysis of Integrator occupancy using INTS11, INTS1 and INST9 antibodies at four eRNA loci. Data were collected during a time course of EGF induction in HeLa cells (0, 20, 40 and 60 min). Error bars represent ± standard error of the mean (s.e.m., n = 3 biological independent experiments). P < 0.01 by two-sided t-test. b, Depletion of INST1 and INST11 protein levels in tet-inducible HeLa clones. The arrow indicates the INTS11-specific signal; the asterisk shows a non-specific band. c, Fold change of H3K27 acetylation (0 min/20 min EGF) before (ctrl) and after (dox) depletion of INTS11. Data were calculated from read density of ChIP-seq experiments across EGF-induced enhancers. Depletion of Integrator significantly affects EGF-dependent increase in H3K27ac (P < 0.05). Whiskers on the box plot indicate the variability in the datasets.
Extended Data Figure 4 | Depletion of Integrator impairs activation of eRNAs by EGF. a, b, Activation of eRNAs near DUSP1, CCNL1 and NR4A1 genes were assayed by qRT–PCR in three independent experiments, using INTS11 (a) or INTS1 (b) inducible shRNA clones. Transcription was followed throughout a 20-min time-course experiment. Each eRNA was amplified with two different sets of specific primers to analyse both strands; dashed lines indicate treatment with doxycycline (dox) to induce shRNAs. Data at every time point are reported as fold change (EGF/non-induced). Error bars represent ± s.e.m. (n = 3 biological independent experiments), P < 0.01 by two-sided t-test. c, Schematic representation of ATF3 and its super-enhancer region located 30 kb upstream (top). Snapshots of ChIP-seq and RNA-seq tracks show EGF-dependent recruitment of RNAPII and INTS11 at the ATF3 locus and at several upstream enhancers. Depletion of INTS11 nearly abolished transcription of eRNAs and ATF3 mRNA. d, Real-time RT–PCR analysis of the ATF3 super-enhancer region upon depletion of INTS11. qPCR analysis was performed before and 5, 10, 15, 20 min after EGF treatment with strand-specific primer sets (indicated below the RNA-seq tracks in c). Error bars represent ± s.e.m. (n = 3 biological independent experiments), P < 0.01 by two-sided t-test.
Extended Data Figure 5  | Chromatin conformation capture at control loci.

**a**, 3C analysis of NR4A1 promoter and control sites. The Con1 site lies 74 kb upstream of the NR4A1 protein-coding gene and the Con2 site is located 42 kb downstream of the enhancer site. There are no looping events between either control sites with the NR4A1 promoter region after EGF induction.

**b**, Similarly, no looping events were detected between the promoter of DUSP1 and a downstream control site (Con). All data were averaged from three independent experiments, $P < 0.01$ by two-sided t-test.
**Extended Data Figure 6 | Integrator has a role in eRNA termination.**
a. Mean density profiles of GRO-seq data at 91 EGF-induced enhancers. Data are presented as strand-specific mean read density, centred at the middle of the RNAPII peak and normalized to sequencing depth. The underlying box plots were used to quantify the enrichment of GRO-seq reads at the 3’ end of both eRNA transcripts (2 kb window, centred 1 kb downstream of the RNAPII peak). b, RNAPII profiling at 91 enhancers after INTS11 depletion shows accumulation of ChIP-seq reads towards the 3’ end. Data are presented as mean read density, centred at the middle of the RNAPII peak and normalized to sequencing depth. Box plots represent the enrichment of RNAPII reads of both eRNA transcripts (2 kb window, centred 1 kb downstream of the RNAPII peak). RNAPII significantly accumulated ($P < 0.004$) after depletion of INTS11. Whiskers on the box plots indicate the variability in the datasets. c, d, RNAPII travelling ratio at enhancers was measured as the ratio between RNAPII density close to the transcription start site (the surrounding 300 bp) and 3 kb downstream. Given the bi-directional nature of transcription at enhancers, travelling ratio was calculated for both sense (c) and antisense (d) transcripts.
Extended Data Figure 7 | Analysis of super elongation complex at enhancers. a, b, Metagene analysis on 91 eRNA loci shows the effect of EGF stimulation and INTS11 depletion on the recruitment of the ELL2 (a) and AFF4 (b) subunits of the super elongation complex (SEC). SEC was recruited to enhancers upon EGF stimulation. Depletion of Integrator decreases AFF4 and ELL2 recruitment. Data were visualized as mean read density, normalized to sequencing depth, across 8 kb surrounding the centre of enhancers. c, To investigate the role of the negative elongation factor (NELF) in induction of eRNAs, we infected HeLa cells with lentiviral shRNAs against NELFA, NELFE and a control GFP. Quantitative RT–PCR analysis shows the extent of NELF depletion 72 h after infection. Error bars represent ± s.e.m. (n = 3 biological independent experiments), P < 0.01 by two-sided t-test. d, Depletion of two different NELF subunits does not significantly impact activation of EGF-responsive eRNAs. Data represent fold change of induction (EGF/not induced) after 20 min of stimulation and were normalized against GUSB expression. Error bars represent ± s.e.m. (n = 3 biological independent experiments), P < 0.01 by two-sided t-test. e, ChIP-seq analysis of NELFA before and after depletion of INTS11. Metagene analysis shows mean read density (normalized to sequencing depth) across 91 eRNAs. NELF occupancy at enhancers was not affected by depletion of Integrator.
Extended Data Figure 8 | Integrator depletion causes accumulation of unprocessed eRNAs and prevents release of RNAPII. a, Termination of eRNAs was examined with quantitative RT–PCR. Primer pairs were designed to amplify a portion of the enhancer transcript detected in normal condition (t, total) or a longer template further extending into the 3’ of the enhancer region (u, unprocessed). qPCR analysis was performed before (ctrl) and after (dox) depletion of INTS11 at three eRNAs (sense and antisense strand), after stimulation with EGF. In the absence of INTS11, we observed accumulation of unprocessed eRNA, suggestive of a termination defect. Error bars represent ± s.e.m. (n = 3 biological independent experiments), *P < 0.01 by two-sided t-test. Release of eRNA transcripts from RNA polymerase was investigated by means of RNAPII immunoprecipitation following UV cross-link (UV-RIP). b–d, After RNAPII immunoprecipitation, eRNAs near DUSP1, CCNL1 and NR4A1 genes were assayed by qRT–PCR and showed increased association with RNAPII in the absence of Integrator. Each eRNA was detected by two different sets of specific primers (sense and antisense). Error bars represent ± s.e.m. (n = 3 biological independent experiments). *P < 0.01, **P < 0.01, ***P < 0.001 by two-sided t-test. e–g, RNAPII UV-RIP analysis was also performed on several eRNAs from the ATF3 super-enhancer. qRT–PCR on the RNA recovered after immunoprecipitation shows increased association between RNAPII and eRNAs in the absence of Integrator. Each eRNA was detected by two different sets of specific primers (sense and antisense). Error bars represent ± s.e.m. (n = 3 three independent experiments). **P < 0.01 by two-sided t-test.
Extended Data Figure 9 | Distribution of RNAPII and nascent RNAs across protein-coding genes. a, Expression level of exogenous INTS11 wild type (WT) and its catalytic mutant (E203Q). Nuclear extracts were subjected to Flag immunoprecipitation and probed with a polyclonal antibody raised against the C terminus of INTS11. b, Heat map of nascent RNA (GRO-seq) and RNAPII ChIP-seq across the 2,000 most active genes in HeLa cells. Gene loci were analysed for their entire gene body, with 3 additional kilobases on both ends. H3K27ac data from ENCODE is shown on the left; genes are ranked according to the intensity of RNAPII signal. Depletion of Integrator does not appear to affect termination at protein-coding genes.