

Molecular Cell, Volume 69

Supplemental Information

Tyrosine-1 of RNA Polymerase II CTD

Controls Global Termination

of Gene Transcription in Mammals

Nilay Shah, Muhammad Ahmad Maqbool, Yousra Yahia, Amal Zine El Aabidine, Cyril Esnault, Ignasi Forné, Tim-Michael Decker, David Martin, Roland Schüller, Stefan Krebs, Helmut Blum, Axel Imhof, Dirk Eick, and Jean-Christophe Andrau

Supplemental Material

Tyrosine-1 of RNA Polymerase II CTD controls global termination of gene transcription in mammals

Nilay Shah^{1#}, Muhammad Ahmad Maqbool^{2#}, Yousra Yahia², Amal Zine El Aabidine², Cyril Esnault², Ignasi Forné³, Tim-Michael Decker¹, David Martin², Roland Schüller¹, Stefan Krebs⁴, Helmut Blum⁴, Axel Imhof³, Dirk Eick^{1*} and Jean-Christophe Andrau^{2,5*}

Affiliations

¹Department of Molecular Epigenetics, Helmholtz Center Munich, Center of Integrated Protein Science Munich, Marchioninistrasse 25, 81377 Munich, Germany;

²Institut de Génétique Moléculaire de Montpellier (IGMM), CNRS-UMR5535, Montpellier, France;

³Biomedical Center Munich, ZFP, Großhaderner Strasse 9, 82152 Planegg-Martinsried, Germany;

⁴Laboratory for functional genome analysis, Gene Center, Ludwig-Maximilians-Universität, Munich, Germany.

These authors contributed equally

⁵Lead Contact

* Corresponding authors: jean-christophe.andrau@igmm.cnrs.fr and eick@helmholtz-muenchen.de

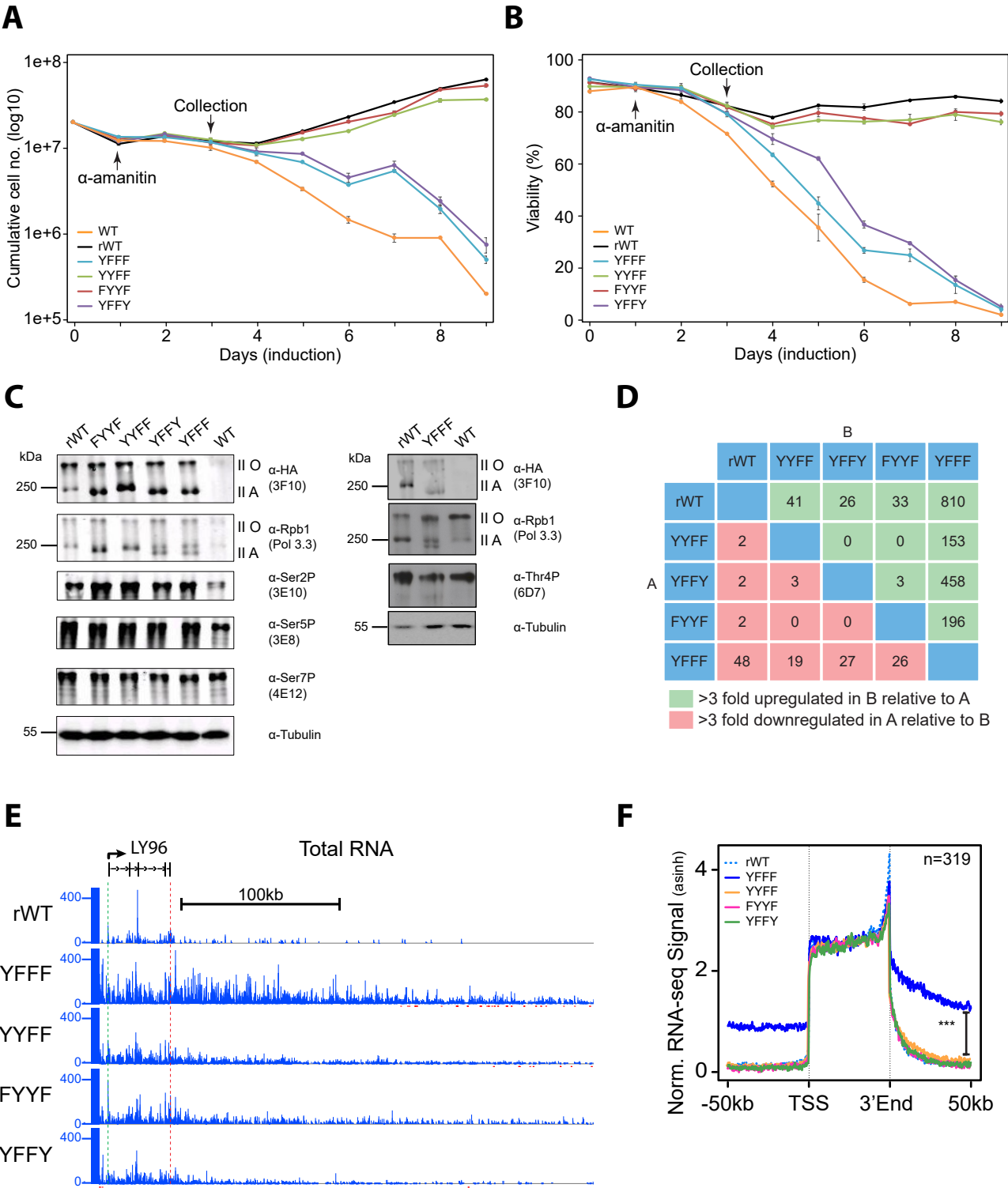


Figure S1: Phenotype and differential transcriptome of CTD Tyrosine mutants, Related to Figure 1

A, B) Proliferation kinetics and viability curve following induction of rWT and tyrosine mutants by removal of tetracycline (tet-off system) and treatment with α -amanitin. The time of sample collection for RNA-seq and ChIP-seq experiments presented further is indicated by an arrow. **C)** Western blots probing for Ser2P, Ser5P and Ser7P Abs in all mutants (left panel) and Thr4P in rWT, YFFF and WT cells (right panel), indicate no major alteration on the phosphorylation pattern of the CTD. **D)** Differential gene expression analysis of genes up and down regulated in the tyrosine mutants relative to rWT (3-fold change, FDR<0.05). **E)** Example of read-through phenotype at 3' end of the LY96 gene. **F)** Average metagene profile of total sense RNA-seq signal (asinh) over the gene bodies and 50kb upstream and downstream regions. The 3 stars indicate a p-value < 2×10^{-16} (2 sided Wilcoxon test) between rWT and YFFF.

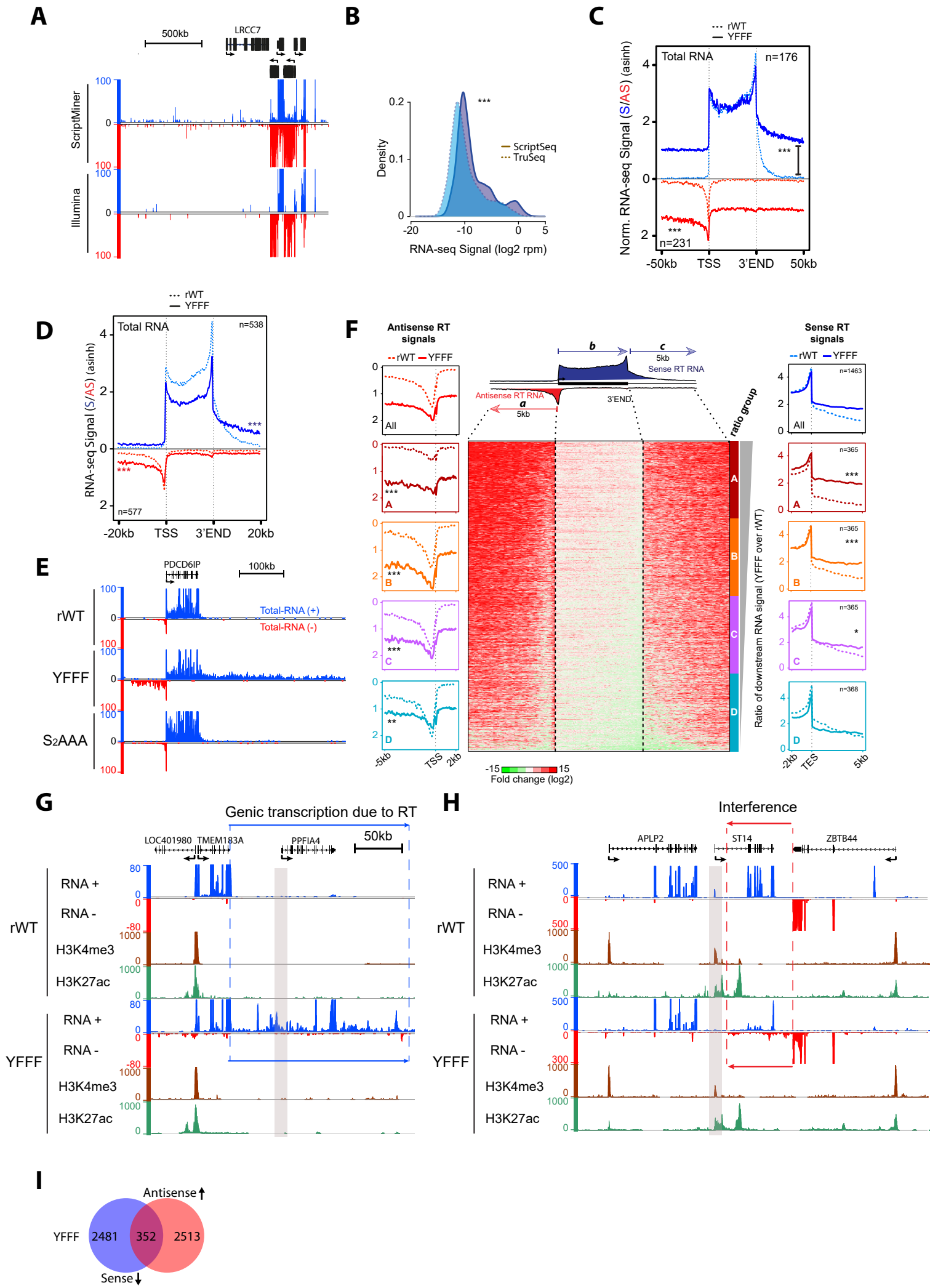


Figure S2: YFFF RT phenotype at 3' sense and 5' antisense transcription is specific and likely causes transcriptional interference, Related to Figure 2

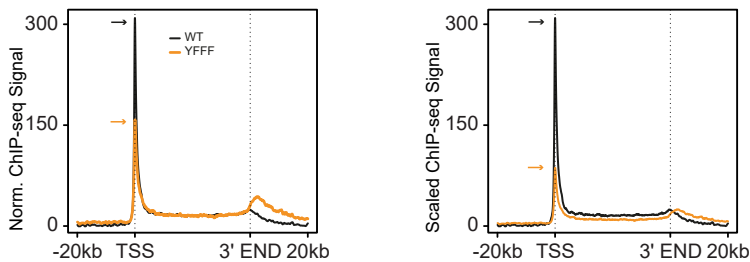
A) Comparison of RNA-seq in WT Raji cells performed by using ScriptMiner RNA Library Prep Kit vs Illumina TruSeq small RNA Library Prep Kit. As compared to ScriptMiner Kit, TruSeq kit reduces the background noise in intergenic regions as well as increases strand-specificity. **B)** Distribution of intergenic RNA-seq signals obtained from TruSeq and ScriptMiner libraries in a selection of 21792 intergenic regions, excluding genes within 100kb for an assessment of the intrinsic experimental noise. The 3 stars indicate a p-value $< 2 \times 10^{-16}$ (2 sided Wilcoxon test). **C)** Average metagene profile of total RNA-seq signal (asinh) in sense (blue) and AS (red) orientation of the gene bodies and 50kb upstream and downstream regions. The 3 stars indicate a p-value $< 2 \times 10^{-16}$ (2 sided Wilcoxon test). **D)** Average metagene profile of total RNA-seq signal (asinh) without normalization on gene bodies, in sense (blue) and AS (red) orientation of the gene bodies and 20kb upstream and downstream regions. **E)** Tyrosine mutations of the YFFF induce a specific 5' AS and 3' sense RT phenotype as exemplified at the PDCD6IP locus and as compared to the S2AAA control mutant (lower panel). The S2AAA mutant has Ser2 positions of the last 3/4 of the CTD heptads replaced by Ala (lower panel). The PDCD6IP gene is representative for the RT phenotype observed genome-wide. **F)** Comparison of the 3' sense and the 5' AS RT phenotypes in the YFFF mutant using total RNA-seq. Genes were ranked according to 3' RT decreasing ratio in the YFFF vs rWT within 5kb after the annotated 3' ends and further divided in 4 equal sized groups A-D (colored profiles on the right). The corresponding ratio profiles in 5' AS RNAs are shown on the left of the density heat maps and in the middle for the gene bodies rescaled (0-100%). The upper plots represent the global average profiles whereas the 4 below, represent the groups A to D (more to less affected in 3' RT from top to bottom). The p-values of the YFFF vs rWT comparison are for (1) 3' S: A $< 2.2 \times 10^{-16}$; B $< 2.2 \times 10^{-16}$; C = 9.715×10^{-11} ; D = 0.037 and (2) for 5' AS: A $< 2.2 \times 10^{-16}$; B $< 2.2 \times 10^{-16}$; C $< 2.2 \times 10^{-16}$; D = 2.35×10^{-15} . **G)** Augmentation of adjacent transcript densities, PPFIA4 as example. The absence of the epigenetic marks H3K4me3 and H3K27ac at the PPFIA4 promoter pleads for a RT effect of TMEM183A rather than a neo-initiation event. **H)** Example of apparent RNA interference resulting in transcription inhibition (ST14 example). **I)** Venn Diagram showing overlap of protein coding genes down regulated with increased AS signal in YFFF mutant. These selections were isolated using DESeq package with \log_2 FC > 1 , FDR 0.05, oval 0.05.

Shah, Maqbool et al, Figure S3

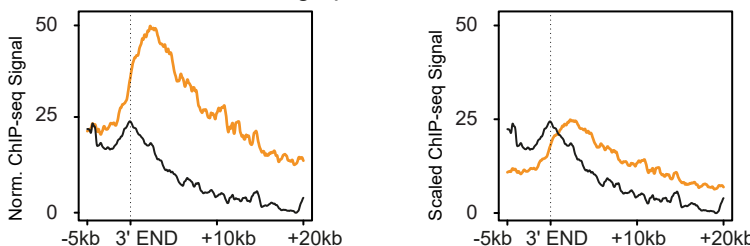
**Scaled on sequenced tag
number and normalised
to gene bodies**

**Scaled on sequenced
tag number**

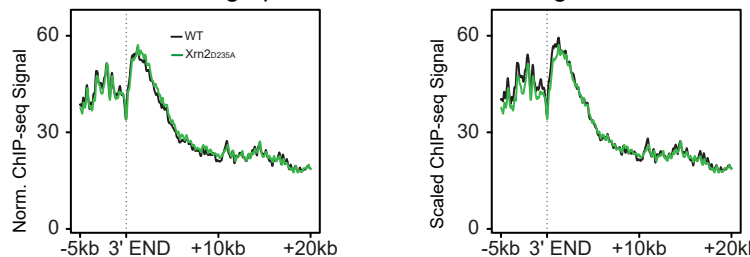
A Composite average profile over gene bodies



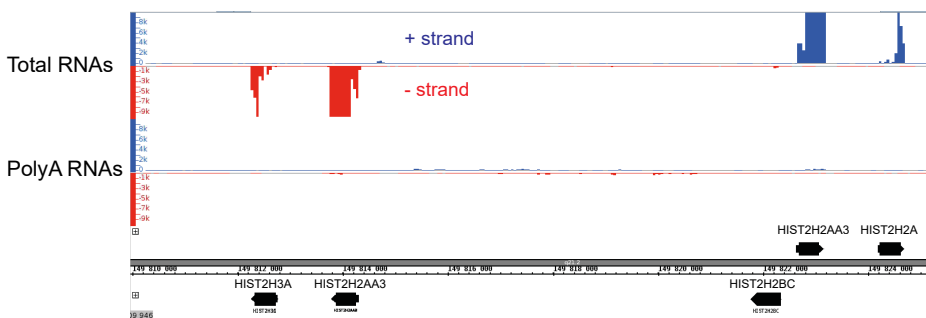
B Average profile over 3'ENDs



C Average profile over 3'ENDs: Fong et al, 2015

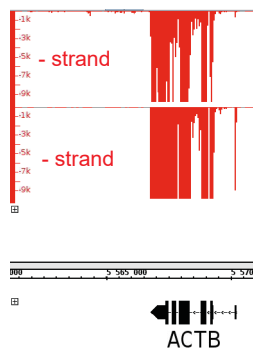


D



Total RNAs

PolyA RNAs



E

chrRNA

polyA-RNA

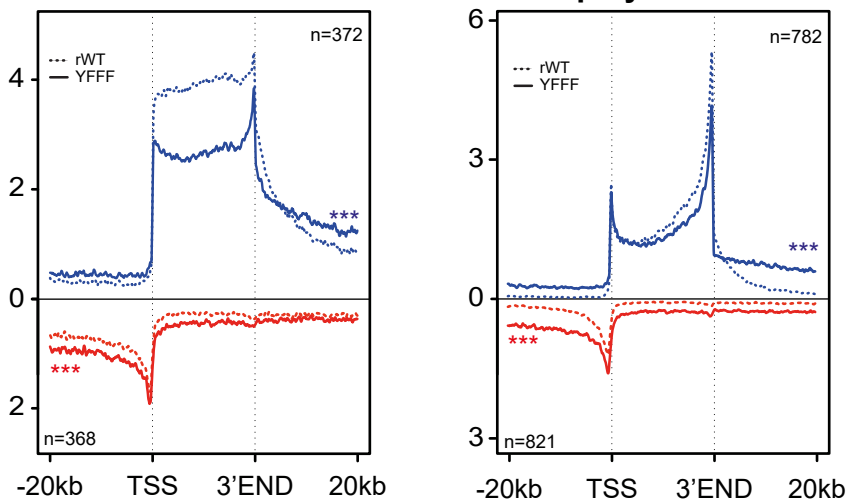


Figure S3: Pol II ChIP-seq and Poly(A) vs total RNA-seq analyses of YFFF mutant phenotype, Related to Figure 3

A) Composite Pol II ChIP-seq average profile (top 30% coding genes) over gene bodies normalized as described in the methods section for rWT and YFFF mutant (left) or just based on the sequence tag counts (right). The orange and black arrows represent the positions of the max peak at TSSs for WT and YFFF Pol II. **B)** Pol II average profiles for rWT and YFFF around 3' end of genes, normalized/scaled as in A). **C)** Pol II average profiles for WT and Xrn2 dominant mutation (Fong et al, 2015) around 3' end of genes, normalized/scaled as in A) in a selection of the top30% of coding genes. **D)** RNA-seq signals in rWT cells over 4 genes (non-polyadenylated histones) of the histone cluster located on chromosome 1. Around 100 times less signal is observed in poly(A) RNA-seq as compared to total RNA-seq suggesting that the protocols used allow discrimination of both populations. The highly transcribed ACTB poly(A) coding gene is shown as control, enriched in both poly(A) and total RNA-seq experiments. **E)** Average metagene profiles of chr- and PolyA- RNA-seq signal (asinh) without normalization on gene bodies, in sense (blue) and AS (red) orientation of the gene bodies and 20kb upstream and downstream regions.

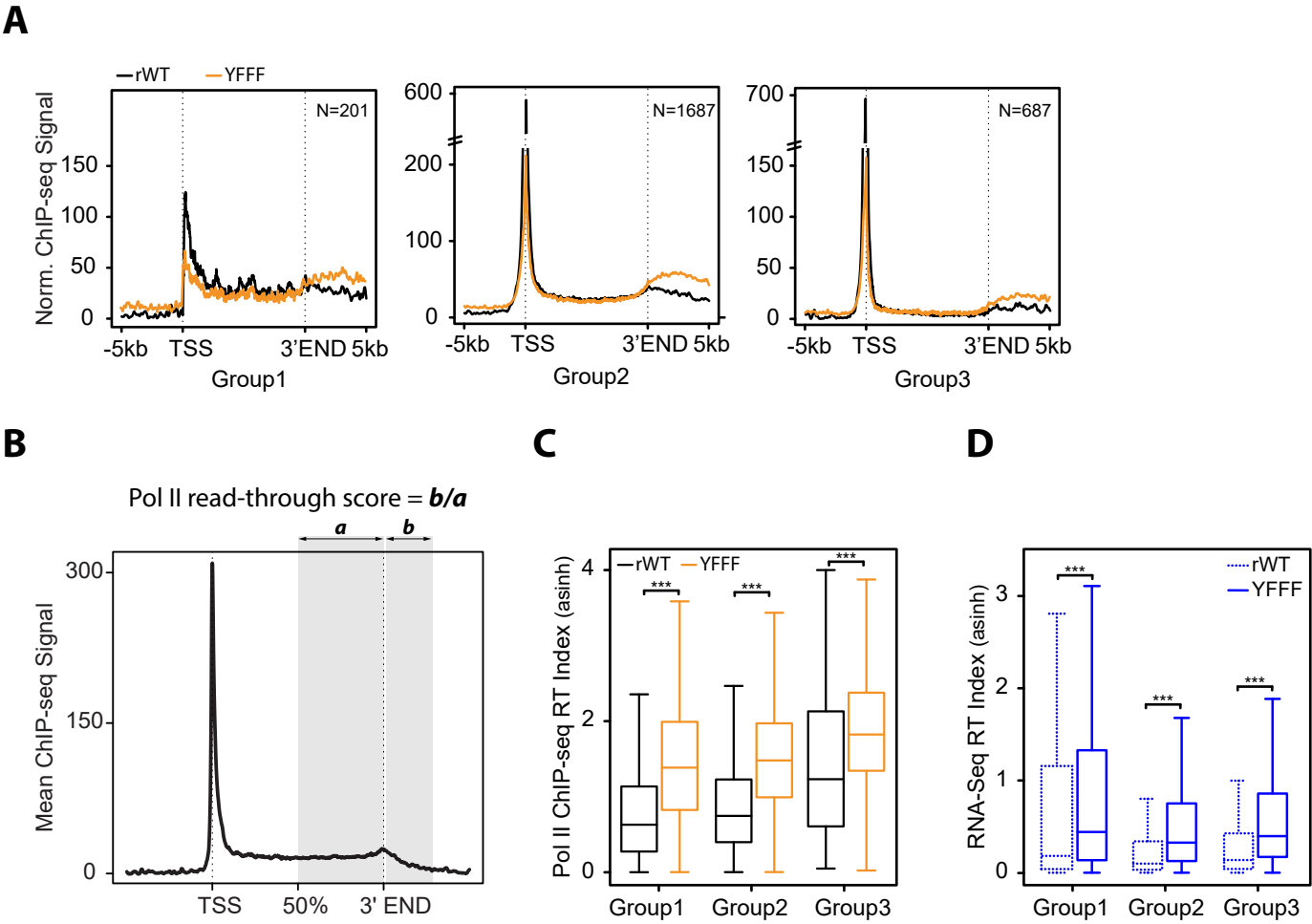


Figure S4: Pol II Pausing and 3' end RT of RNA and Pol II, Related to Figure 4

A) Pol II average profile on the 3 groups of pausing scores defined in Figure 4. **B)** Pol II RT score shown in C) was calculated by dividing the average Pol II signal downstream of 3' end (10kb) with average signal in the second half of gene body (50-100%). **C)** RT indices Box plot in rWT and YFFF calculated with Pol II signal from three groups of genes. **D)** Box plot of read-through score calculated with total RNA signal from three groups of genes.

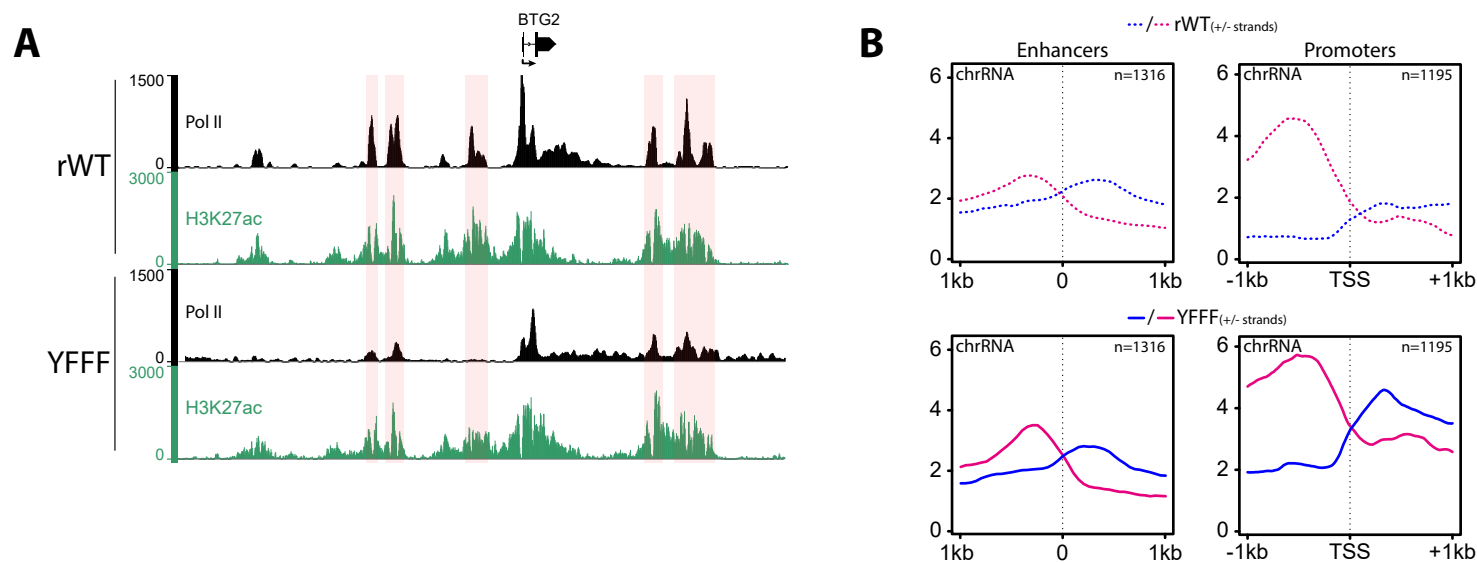
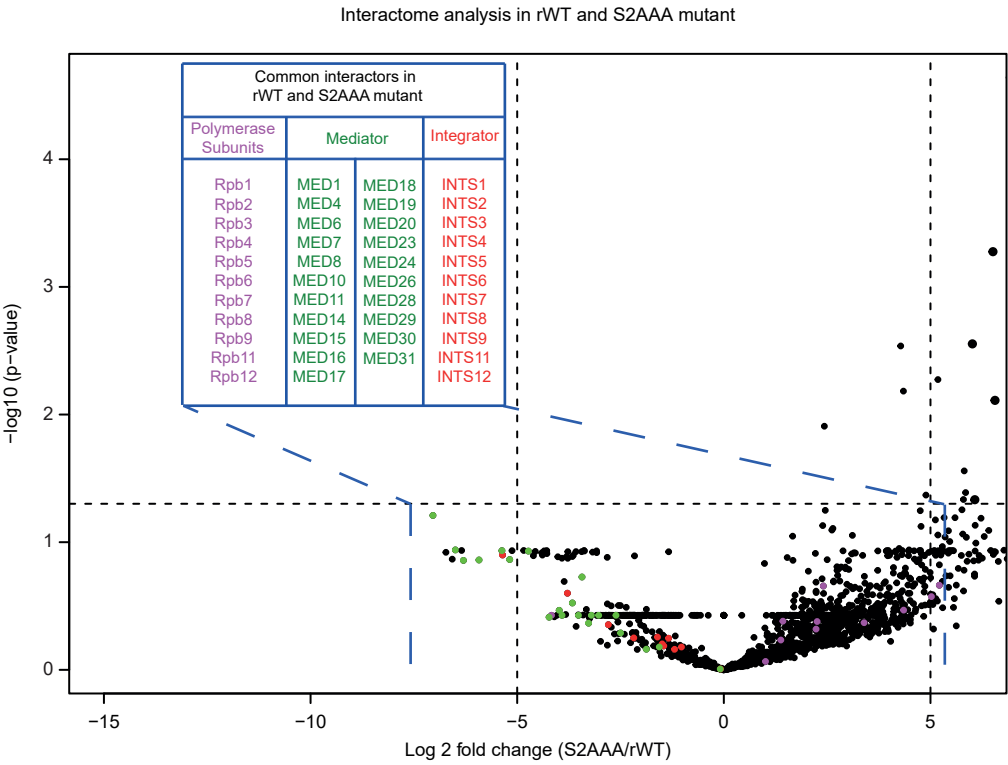


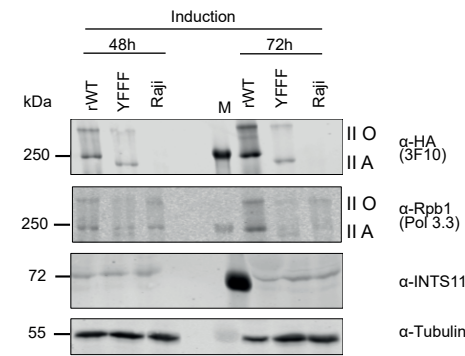
Figure S5: YFFF mutations do not impair transcription at enhancers, Related to Figure 5

A) Additional example of impaired Pol II loading at enhancers around the BTG2 gene indicated by light pink rectangles. **B)** ChrRNA-seq average profiles at enhancers reveal that nascent transcription is not affected when compared to promoters (right).

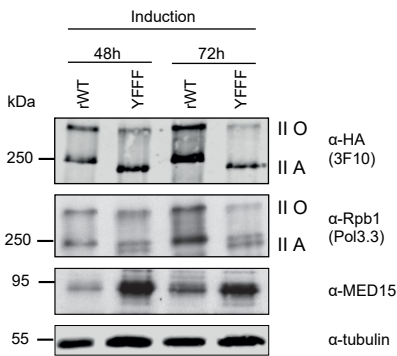
A



B



C



D

Raji

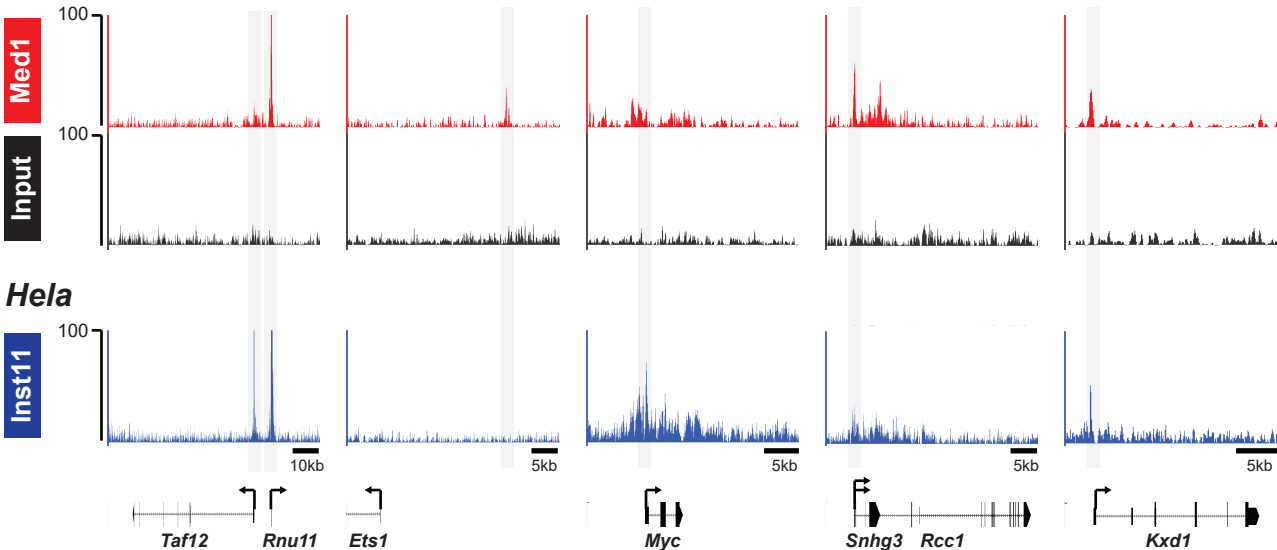


Figure S6: Control experiments for MS specificity and Mediator/Integrator complexes integrity in the YFFF mutant, Related to Figure 6

A) Volcano plot showing MS differential analysis of Mediator and Integrator interactions in a S2AAA mutant. No major interaction loss with Int subunits is observed. **B)** Western blot of HA-Rpb1, Rpb1, Ints11 in rWT or YFFF cells following 48h and 72h of induction (24h and 48h of α -amanitin treatment). **C)** Western blot of HA-Rpb1, Rpb1, Med15 in rWT or YFFF cells following 48h and 72h of induction (24h and 48h of α -amanitin treatment). **C)** ChIP-seq signals at selected loci for ChIP-qPCR analysis. The Med1 (red) and Input (Black) tracks from Raji cells are shown on top, the Ints11 track from HeLa cells is shown at bottom (Stadelmayer et al., 2014). Grey rectangles highlight areas that were used for qPCR analysis (see Figure 6).

Supplementary Tables

Table S1: Peptide counts of proteins and complexes not interacting with the YFFF mutant. Related to Figure 6

Peptide counts of 69 proteins that do not interact with the YFFF mutant for all five biological replicates. Samples in the experiments 1 and 2 were subjected to on-beads trypsin digest, while samples in the experiments 3, 4 and 5 were subjected to in-gel trypsin digest.

			rWT					YFFF				
	Experiment Number		1	2	3	4	5	1	2	3	4	5*
	Uniprot ID	Gene Name	Peptide counts					Peptide counts				
1	Q9NRY2	INIP	2	5	3	5	8	0	0	0	0	0
2	Q9NPJ6	MED4	6	10	13	13	13	0	0	0	0	0
3	Q9UL03	INTS6	10	27	32	41	45	0	0	0	0	0
4	Q6P2C8	MED27	3	9	11	12	10	0	0	0	0	0
5	Q9BUE0	MED18	3	4	3	3	6	0	0	0	0	0
6	O95402	MED26	5	14	14	18	11	0	0	0	0	0
7	Q9BTT4	MED10	1	5	6	6	7	0	0	0	0	0
8	Q5T8T7	MED22	1	2	6	7	6	0	0	0	0	0
9	Q6P9B9	INTS5	3	15	13	17	26	0	0	0	0	0
10	Q68E01	INTS3	12	31	35	34	50	0	1	1	0	0
11	Q9NWA0	MED9	1	1	6	3	6	0	0	0	0	0
12	A0JLT2	MED19	1	3	3	3	5	0	0	0	0	0
13	Q96CB8	INTS12	1	6	9	15	13	0	0	0	0	0
14	Q9H0H0	INTS2	4	11	17	18	30	0	0	0	0	0
15	Q96G25	MED8	4	8	8	9	11	0	0	1	0	0
16	Q8N201	INTS1	16	43	48	60	74	0	0	1	0	0
17	Q9H944	MED20	1	2	4	7	6	0	0	1	0	0
18	Q96HW7	INTS4	2	12	15	31	41	0	0	0	0	0
19	Q9Y3C7	MED31	2	4	5	5	5	0	1	0	0	0
20	Q9H0M0	VWV1	3	12	18	20	27	0	0	0	0	0
21	Q9NVC6	MED17	3	6	19	17	21	0	0	1	0	0
22	Q9NV88	INTS9	2	5	7	13	21	0	0	0	0	0
23	Q96HR3	MED30	1	4	6	6	5	0	0	1	0	0
24	O60244	MED14	6	29	33	30	34	0	0	2	0	0
25	O00308	VWV2	5	12	13	27	24	0	0	0	0	0
26	Q15648	MED1	6	10	28	31	31	0	0	1	0	0
27	O75586	MED6	1	8	9	6	6	0	0	1	0	0
28	Q9BQ15	NABP2	0	5	1	4	7	0	0	0	0	0
29	Q9H204	MED28	0	1	3	3	4	0	0	0	0	0
30	O43513	MED7	0	1	3	7	8	0	0	0	0	0
31	Q9Y2Z0	SUGT1	1	1	2	3	4	1	0	0	0	0
32	Q96J02	ITCH	13	25	26	29	34	0	2	4	1	0
33	Q13503	MED21	1	2	1	4	3	0	0	1	0	0
34	Q75QN2	INTS8	0	6	6	18	23	0	0	0	0	0
35	Q9P086	MED11	0	1	2	4	3	0	0	0	0	0
36	Q96P16	RPRD1A	1	15	23	12	15	0	8	8	0	0
37	Q96RN5	MED15	0	1	5	6	7	0	0	0	0	0
38	Q5TA45	CPSF3L	0	2	8	8	13	0	0	0	0	0
39	Q15369	TCEB1	1	1	2	5	3	0	0	0	3	0
40	O95104	SCAF4	0	2	11	7	5	0	0	0	0	0
41	Q9NX70	MED29	0	3	2	3	3	0	0	0	0	0
42	O75448	MED24	0	4	6	17	20	0	0	0	0	0
43	Q5VT52	RPRD2	0	8	18	15	7	0	0	0	0	0
44	Q6DN90	IQSEC1	0	6	12	6	11	0	0	0	0	0

45	Q9Y2X0	MED16	0	3	2	8	12		0	0	0	0	0
46	Q9NVH2	INTS7	0	6	6	23	30		0	0	0	1	0
47	A8MU58	AIMP2	0	1	2	2	2		0	0	0	0	0
48	Q5JSJ4	INTS6L	2	8	8	10	8		0	0	0	0	0
49	Q5TEJ8	THEMIS2	0	1	3	3	6		0	0	0	0	0
50	P30153	PPP2R1A	0	2	1	4	6		0	0	1	0	0
51	Q99590	SCAF11	0	3	5	3	3		0	0	0	0	0
52	Q13418	ILK	0	3	5	3	3		0	0	2	0	0
53	Q53G59	KLHL12	0	1	2	1	1		0	0	0	0	0
54	O00329	PIK3CD	0	2	3	4	2		0	0	0	0	0
55	Q13049	TRIM32	0	1	1	2	3		0	0	0	0	0
56	Q14145	KEAP1	0	2	8	2	3		0	0	0	0	0
57	Q13501	SQSTM1	0	1	3	0	3		0	0	0	1	0
58	Q14344	GNA13	0	1	2	1	1		0	0	0	0	0
59	H3BQA8	WDR61	1	1	0	1	1		0	0	0	0	0
60	O00505	KPNA3	1	3	4	4	4		0	1	1	0	0
61	Q14157	UBAP2L	0	2	2	1	1		0	0	0	0	0
62	Q15418	RPS6KA1	2	6	6	5	7		0	2	2	0	0
63	Q8ND56	LSM14A	0	2	4	0	1		0	0	0	0	0
64	Q9ULK4	MED23	0	1	1	9	12		0	0	0	0	0
65	Q13451	FKBP5	2	3	2	4	8		1	1	2	0	0
66	P04637	TP53	0	2	1	0	1		0	0	1	0	0
67	Q71RC2	LARP4	0	0	2	3	3		0	0	0	0	0
68	Q16576	RBBP7	0	3	4	4	2		0	1	3	2	0
69	P13807	GYS1	0	3	2	0	1		0	0	0	0	0

Table S2: List of proteins and complexes not interacting with the YFFF mutant. Related to Figure 6

A total of 69 proteins were found that shows loss of interaction with the YFFF Pol II mutant compared to rWT. Listed in the table are 25 subunits of the Mediator complex (green); 11 subunits of the Integrator complex (red); CTD phosphatase (magenta); E3-ubiquitin ligase, components of SOSS complex (blue) and few others. Log2 fold change (YFFF/rWT) and p-values for each protein is shown in the table. Data is based on five independent biological replicates.

	Uniprot ID	Gene Name	Description	Log2Fold Change (YFFF/rWT)	p-value
1	Q9NRY2	INIP	INTS3 and NABP interacting protein	-14.548	5.588E-08
2	Q9NPJ6	MED4	Mediator Complex Subunit 4	-13.671	1.447E-09
3	Q9UL03	INTS6	Integrator Complex Subunit 6	-13.490	3.673E-09
4	Q6P2C8	MED27	Mediator Complex Subunit 27	-13.159	2.356E-09
5	Q9BUE0	MED18	Mediator Complex Subunit 18	-12.981	1.433E-12
6	Q95402	MED26	Mediator Complex Subunit 26	-12.838	1.41E-11
7	Q9BTT4	MED10	Mediator Complex Subunit 10	-12.827	2.053E-08
8	Q5T8T7	MED22	Mediator Complex Subunit 22	-12.637	3.805E-08
9	Q6P9B9	INTS5	Integrator Complex Subunit 5	-12.259	7.930E-08
10	Q68E01	INTS3	Integrator Complex Subunit 3	-12.146	2.420E-05
11	Q9NWA0	MED9	Mediator Complex Subunit 9	-11.932	1.696E-06
12	A0JLT2	MED19	Mediator Complex Subunit 19	-11.858	4.219E-08
13	Q96CB8	INTS12	Integrator Complex Subunit 12	-11.533	2.203E-05
14	Q9H0H0	INTS2	Integrator Complex Subunit 2	-11.520	2.799E-08
15	Q96G25	MED8	Mediator Complex Subunit 8	-11.496	2.292E-04
16	Q8N201	INTS1	Integrator Complex Subunit 1	-11.496	2.449E-06
17	Q9H944	MED20	Mediator Complex Subunit 20	-11.388	2.521E-04
18	Q96HW7	INTS4	Integrator Complex Subunit 4	-11.258	5.507E-06
19	Q9Y3C7	MED31	Mediator Complex Subunit 31	-11.180	5.368E-04
20	Q9H0M0	WWP1	WW Domain containing E3 Ubiquitin Protein Ligase 1	-11.161	3.837E-07
21	Q9NVC6	MED17	Mediator Complex Subunit 17	-11.092	1.153E-04
22	Q9NV88	INTS9	Integrator Complex Subunit 9	-11.079	9.394E-07
23	Q96HR3	MED30	Mediator Complex Subunit 30	-11.069	2.195E-04
24	O60244	MED14	Mediator Complex Subunit 14	-10.979	1.094E-04
25	O00308	WWP2	WW Domain containing E3 Ubiquitin Protein Ligase 2	-10.976	5.113E-07
26	Q15648	MED1	Mediator Complex Subunit 1	-10.736	9.263E-06
27	O75586	MED6	Mediator Complex Subunit 6	-10.708	6.511E-04
28	Q9BQ15	NABP2	Nucleic acid binding protein 2	-10.628	4.230E-03
29	Q9H204	MED28	Mediator Complex Subunit 28	-9.820	4.124E-03
30	O43513	MED7	Mediator Complex Subunit 7	-9.771	3.981E-03
31	Q9Y2Z0	SUGT1	SGT1 Homolog, MIS12 Kinetochore Complex Assembly Cochaperone	-9.705	3.864E-11
32	Q96J02	ITCH	Itchy E3 Ubiquitin Protein Ligase	-9.434	2.206E-03
33	Q13503	MED21	Mediator Complex Subunit 21	-9.335	2.326E-03
34	Q75QN2	INTS8	Integrator Complex Subunit 8	-9.223	4.178E-03
35	Q9P086	MED11	Mediator Complex Subunit 11	-8.966	4.348E-03
36	Q96P16	RPRD1A	Regulation of nuclear pre-mRNA domain containing 1A (CTD phosphatase)	-8.924	1.335E-02
37	Q96RN5	MED15	Mediator Complex Subunit 15	-8.887	4.168E-03
38	Q5TA45	CPSF3L	Cleavage and polyadenylation specificity factor 3-like (Integrator Complex Subunit 11)	-8.729	4.677E-03
39	Q15369	TCEB1	Transcription elongation factor B subunit 1	-8.481	5.145E-03
40	O95104	SCAF4	SR-related CTD associated factor 4	-8.349	4.985E-03
41	Q9NX70	MED29	Mediator Complex Subunit 29	-8.304	4.207E-03
42	O75448	MED24	Mediator Complex Subunit 24	-7.916	4.396E-03
43	Q5VT52	RPRD2	Regulation of nuclear pre-mRNA domain containing 2 (CTD phosphatase)	-7.903	4.556E-03
44	Q6DN90	IQSEC1	IQ motif and Sec7 Domain 1	-7.801	4.395E-03
45	Q9Y2X0	MED16	Mediator Complex Subunit 16	-7.639	4.243E-03
46	Q9NVH2	INTS7	Integrator Complex Subunit 7	-7.255	3.588E-02
47	A8MU58	AIMP2	Aminoacyl tRNA Synthetase Complex-interacting Multifunctional protein 2	-7.053	4.231E-03
48	Q5JSJ4	INTS6L	Integrator Complex Subunit 6 Like	-6.739	4.555E-03

49	Q5TEJ8	THEMIS2	Thymocyte selection associated family member 2	-6.695	4.375E-03
50	P30153	PPP2R1A	Protein phosphatase 2 regulatory subunit A, alpha	-6.577	7.267E-03
51	Q99590	SCAF11	SR-related CTD associated factor 11	-6.371	4.849E-03
52	Q13418	ILK	Integrin linked kinase	-6.116	4.462E-02
53	Q53G59	KLHL12	Kelch like family member 12	-6.051	4.223E-03
54	O00329	PIK3CD	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Delta	-5.961	4.118E-03
55	Q13049	TRIM32	Tripartite Motif Containing 32	-5.938	7.280E-03
56	Q14145	KEAP1	Kelch Like ECH Associated Protein 1	-5.919	2.511E-02
57	Q13501	SQSTM1	Sequestosome 1	-5.918	4.041E-02
58	Q14344	GNA13	G Protein Subunit Alpha 13	-5.908	6.915E-03
59	H3BQA8	WDR61	WD Repeat Domain 61	-5.784	5.656E-03
60	O00505	KPNA3	Karyopherin Subunit Alpha 3	-5.738	3.355E-02
61	Q14157	UBAP2L	Ubiquitin associated protein 2 like	-5.721	6.337E-03
62	Q15418	RPS6KA1	Ribosomal Protein S6 kinase A1	-5.715	3.025E-02
63	Q8ND56	LSM14A	LSM14A mRNA processing body assembly factor	-5.583	4.459E-02
64	Q9ULK4	MED23	Mediator Complex Subunit 23	-5.570	8.126E-03
65	Q13451	FKBP5	FK506 Binding protein 5	-5.353	1.853E-02
66	P04637	TP53	Tumor protein p53	-5.263	4.374E-02
67	Q71RC2	LARP4	La Ribonucleoprotein Domain Family Member 4	-5.112	4.040E-02
68	Q16576	RBBP7	Retinoblastoma Binding Protein 7	-5.100	4.517E-02
69	P13807	GYS1	Glycogen Synthase 1	-5.082	4.381E-02

Table S3: Peptide counts of selected proteins interacting with the Pol II of both, the rWT and the YFFF mutant. Related to Figure 6

Peptide counts of selected proteins and complexes that interact with Pol II of both, the rWT and the YFFF mutant for all five biological replicates. Samples in the experiments 1 and 2 were subjected to on-beads trypsin digest, while samples in the experiments 3, 4 and 5 were subjected to in-gel trypsin digest.

			rWT					YFFF					
Experimental Number			1	2	3	4	5		1	2	3	4	5*
Uniprot ID	Gene Name	Peptide counts						Peptide counts					
		Polymerase Subunits											
1	P24928	RPB1	119	132	156	153	153		98	126	139	102	11
2	P30876	RPB2	33	53	73	68	70		26	49	61	35	1
3	P19387	RPB3	7	15	17	19	16		5	14	13	9	0
4	O15514	RPB4	0	1	3	4	11		0	0	1	2	0
5	P19388	RPB5	4	11	14	11	10		6	11	9	5	0
6	U3KPY1	RPB6	0	0	1	1	2		0	0	1	0	0
7	P62487	RPB7	1	0	3	4	4		0	1	1	0	0
8	P52434	RPB8	8	12	12	11	12		7	11	11	9	2
9	P36954	RPB9	2	5	7	9	8		2	5	5	3	0
10	P62875	RPB10	3	3	1	1	2		3	3	1	1	0
11	P52435	RPB11	5	7	6	4	9		5	6	6	4	0
12	P53803	RPB12	2	1	2	1	3		0	2	2	1	0
			Splicing factors										
13	Q07955	SRSF1	5	14	20	11	15		5	16	20	13	2
14	J3KP15	SRSF2	0	0	7	0	2		0	2	6	1	0
15	P84103	SRSF3	4	11	10	7	6		3	11	10	10	1
16	Q08170	SRSF4	3	8	7	5	4		2	8	8	4	1
17	Q13243	SRSF5	0	4	5	4	3		1	5	7	3	1
18	Q13247	SRSF6	3	9	9	8	8		3	8	9	9	1
19	Q16629	SRSF7	3	10	11	8	11		7	11	10	10	2
20	Q13242	SRSF9	1	9	16	11	12		4	11	20	17	2
21	O75494	SRSF10	1	9	11	8	10		1	9	12	9	0
22	Q5T760	SRSF11	0	1	2	0	0		0	2	4	0	0
23	Q01081	U2AF1	1	7	7	4	5		1	6	7	3	1
24	P26368	U2AF2	2	2	14	2	4		0	4	8	0	0
			3' end processing and termination factors										
25	Q10570	CPSF1	0	3	13	4	3		0	2	8	1	0
26	Q9P2I0	CPSF2	0	1	2	1	1		0	1	3	0	0
27	G5E9W3	CPSF3	0	0	2	1	2		0	0	1	0	0
28	B7Z7B0	CPSF4	0	1	0	0	1		0	0	1	1	0
29	O43809	CPSF5	0	3	3	0	0		0	3	5	0	0
30	F8WJN3	CPSF6	0	1	2	0	1		0	2	1	0	0
31	Q9H0D6	XRN2	0	8	23	12	14		1	8	17	2	0

Table S4: Conditions for chromatin immunoprecipitation experiments. Related to STAR Methods section “ChIP-seq and ChIP-qPCR”

ChIP-ed Protein	Antibody Ref.	Antibody Quantity	# cells/ChIP	Dynabeads Prot. G /ChIP	# RIPA Washes
Pol-II	ab9110	10 µg	25 x 10 ⁶	100 µl	6
H3K4me1	ab8895	2 µg	5 x 10 ⁶	20 µl	6
H3K4me3	ab8580	2 µg	5 x 10 ⁶	20 µl	5
H3K27ac	ab4729	2 µg	5 x 10 ⁶	20 µl	5

Table S5: List of proteins and complexes in MS with the rWT and Ser2AAA mutant Pol II. Related to Figure 6

Log2fold change (S2AAA/rWT) and p-values for subunits of Polymerase, Mediator and Integrator complexes. Data is based on three independent biological replicates.

List of proteins in rWT and the mutant S2AAA					
	Uniprot ID	Gene Name	Description	Log2Fold Change (S2AAA/rWT)	p-value
Polymerase Subunit					
1	P24928	POLR2A	RNA Polymerase II Subunit B1 (RPB1)	1.429	0.415
2	P30876	POLR2B	RNA Polymerase II Subunit B2 (RPB2)	1.380	0.584
3	P19387	POLR2C	RNA Polymerase II Subunit B3 (RPB3)	2.232	0.479
4	O15514	POLR2D	RNA Polymerase II Subunit B4 (RPB4)	1.009	0.859
5	P19388	POLR2E	RNA Polymerases I, II, And III Subunit ABC1 (RPB5)	2.257	0.418
6	U3KPY1	POLR2F	RNA Polymerases I, II, And III Subunit ABC2 (RPB6)	-4.169	0.374
7	P62487	POLR2G	RNA Polymerase II Subunit B7 (RPB7)	3.394	0.427
8	P52434	POLR2H	RNA Polymerases I, II, And III Subunit ABC3 (RPB8)	2.407	0.220
9	P36954	POLR2I	RNA Polymerase II Subunit B9 (RPB9)	4.350	0.340
10	P62875	POLR2L	RNA Polymerases I, II, And III Subunit ABC5 (RPB10)	7.651	0.111
11	P52435	POLR2J	RNA Polymerase II Subunit B11 (RPB11)	5.215	0.217
12	P53803	POLR2K	RNA Polymerases I, II, And III Subunit ABC4 (RPB12)	5.023	0.267
13	Q8N201	INTS1	Integrator Complex Subunit 1	-1.025	0.662
14	Q9H0H0	INTS2	Integrator Complex Subunit 2	-1.192	0.690
15	Q68E01	INTS3	Integrator Complex Subunit 3	-1.337	0.566
16	Q96HW7	INTS4	Integrator Complex Subunit 4	-1.453	0.650
17	Q6P9B9	INTS5	Integrator Complex Subunit 5	-2.179	0.562
18	Q9UL03	INTS6	Integrator Complex Subunit 6	-1.605	0.554
19	Q9NVH2	INTS7	Integrator Complex Subunit 7	-2.794	0.443
20	Q75QN2	INTS8	Integrator Complex Subunit 8	-5.354	0.126
21	Q9NV88	INTS9	Integrator Complex Subunit 9	-3.784	0.251
22	Q5TA45	CPSF3L	Integrator Complex Subunit 11	-1.492	0.626
23	Q96CB8	INTS12	Integrator Complex Subunit 12	-0.082	0.985
24	Q15648	MED1	Mediator Complex Subunit 1	-1.562	0.662
25	Q9NPJ6	MED4	Mediator Complex Subunit 4	-6.496	0.115
26	O75586	MED6	Mediator Complex Subunit 6	-2.498	0.514
27	O43513	MED7	Mediator Complex Subunit 7	-3.205	0.374
28	Q96G25	MED8	Mediator Complex Subunit 8	-4.227	0.387
29	Q9BTT4	MED10	Mediator Complex Subunit 10	-3.271	0.431
30	Q9P086	MED11	Mediator Complex Subunit 11	-5.924	0.138
31	O60244	MED14	Mediator Complex Subunit 14	-3.984	0.342
32	Q96RN5	MED15	Mediator Complex Subunit 15	-5.369	0.117
33	Q9Y2X0	MED16	Mediator Complex Subunit 16	-4.729	0.117
34	Q9NVC6	MED17	Mediator Complex Subunit 17	-5.181	0.137
35	Q9BUE0	MED18	Mediator Complex Subunit 18	-3.517	0.374
36	A0JLT2	MED19	Mediator Complex Subunit 19	-2.609	0.374
37	Q9H944	MED20	Mediator Complex Subunit 20	-1.875	0.689
38	Q9ULK4	MED23	Mediator Complex Subunit 23	-3.432	0.187
39	O75448	MED24	Mediator Complex Subunit 24	-3.665	0.299
40	Q9NX70	MED29	Mediator Complex Subunit 29	-3.028	0.374
41	Q96HR3	MED30	Mediator Complex Subunit 30	-0.086	0.986
42	Q9Y3C7	MED31	Mediator Complex Subunit 31	-3.914	0.374

Table S6: Peptide counts of proteins in MS with the rWT and Ser2AAA mutant Pol II. Related to Figure 6

Peptide counts of subunits for Polymerase, Mediator and the Integrator complexes in rWT and S2AAA mutant.

Gene Name	Description	Peptide counts					
		rWT1	rWT2	rWT3	S2AAA_1	S2AAA_2	S2AAA_3
POLR2A	RNA Polymerase II Subunit B1 (RPB1)	74	115	89	133	107	120
POLR2B	RNA Polymerase II Subunit B2 (RPB2)	14	61	38	53	56	57
POLR2C	RNA Polymerase II Subunit B3 (RPB3)	1	13	11	12	15	15
POLR2D	RNA Polymerase II Subunit B4 (RPB4)	0	4	3	0	6	4
POLR2E	RNA Polymerases I, II, And III Subunit ABC1 (RPB5)	2	5	6	9	8	8
POLR2F	RNA Polymerases I, II, And III Subunit ABC2 (RPB6)	0	0	1	0	0	0
POLR2G	RNA Polymerase II Subunit B7 (RPB7)	0	2	3	4	2	5
POLR2H	RNA Polymerases I, II, And III Subunit ABC3 (RPB8)	4	7	8	10	10	9
POLR2I	RNA Polymerase II Subunit B9 (RPB9)	0	3	5	6	4	5
POLR2L	RNA Polymerases I, II, And III Subunit ABC5 (RPB10)	0	1	0	1	2	2
POLR2J	RNA Polymerase II Subunit B11 (RPB11)	0	3	2	5	6	6
POLR2K	RNA Polymerases I, II, And III Subunit ABC4 (RPB12)	0	1	1	3	1	1
INTS1	Integrator Complex Subunit 1	3	48	23	21	6	22
INTS2	Integrator Complex Subunit 2	0	15	3	2	1	4
INTS3	Integrator Complex Subunit 3	4	29	14	11	10	9
INTS4	Integrator Complex Subunit 4	0	16	6	3	2	1
INTS5	Integrator Complex Subunit 5	0	14	1	1	0	2
INTS6	Integrator Complex Subunit 6	1	26	12	4	5	4
INTS7	Integrator Complex Subunit 7	0	16	5	1	0	1
INTS8	Integrator Complex Subunit 8	0	12	3	1	0	0
INTS9	Integrator Complex Subunit 9	0	3	2	1	0	0
CPSF3L	Integrator Complex Subunit 11	0	1	2	1	1	0
INTS12	Integrator Complex Subunit 12	0	1	7	6	0	3
MED1	Mediator Complex Subunit 1	0	26	12	3	1	4
MED4	Mediator Complex Subunit 4	2	6	6	0	0	1
MED6	Mediator Complex Subunit 6	1	0	1	3	0	0
MED7	Mediator Complex Subunit 7	0	0	2	0	0	0
MED8	Mediator Complex Subunit 8	3	6	0	0	3	0
MED10	Mediator Complex Subunit 10	0	2	2	0	0	1
MED11	Mediator Complex Subunit 11	0	1	2	0	0	0
MED14	Mediator Complex Subunit 14	0	29	10	2	0	2
MED15	Mediator Complex Subunit 15	0	2	2	0	0	0
MED16	Mediator Complex Subunit 16	0	5	4	0	0	0
MED17	Mediator Complex Subunit 17	3	14	14	6	0	1
MED18	Mediator Complex Subunit 18	0	0	2	0	1	0
MED19	Mediator Complex Subunit 19	0	0	2	0	0	0
MED20	Mediator Complex Subunit 20	0	0	4	1	0	0
MED23	Mediator Complex Subunit 23	0	6	3	0	0	1
MED24	Mediator Complex Subunit 24	0	8	4	0	0	1
MED29	Mediator Complex Subunit 29	0	1	2	1	0	1
MED30	Mediator Complex Subunit 30	0	0	3	1	0	1
MED31	Mediator Complex Subunit 31	0	0	3	0	0	0

Table S7: Parameters used for ChIP-seq peak calling with Integrated Genome Browser's *Thresholding* function. Related to STAR Methods section “ChIP-seq data analysis”

Sample	Thresho Id	Max.Ga p	Min.Ru n
rWT_H3K4me1	110	3000	400
rWT_H3K27ac	130	2000	200
rWT_H3K4me3	100	1000	200
rWT_PolII	35	3000	300