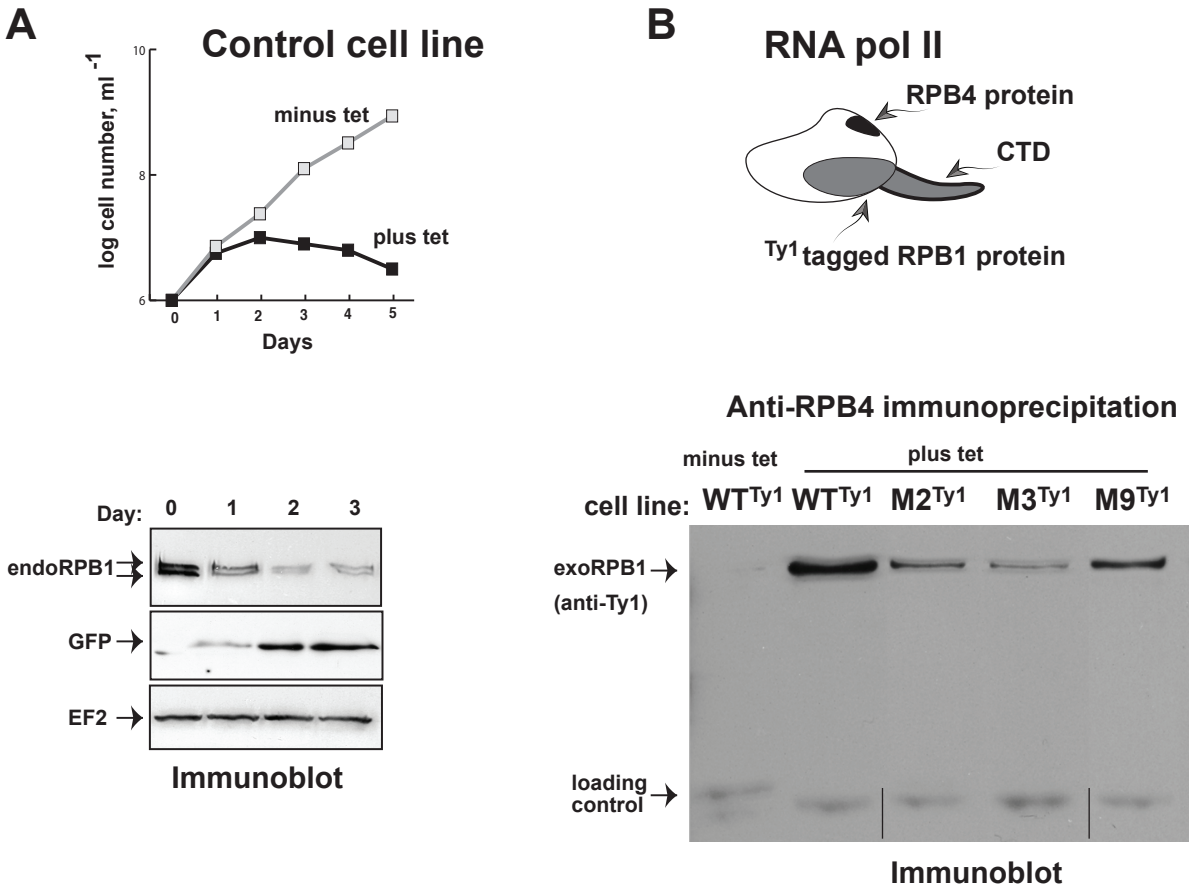

A
Control cell line
log cell number, ml -1
10
8
6
minus tet
plus tet
0 1 2 3 4 5
Days
Day: 0 1 2 3
endoRPB1
GFP
EF2
Immunoblot
B
RNA pol II
RPB4 protein
CTD
Ty1 tagged RPB1 protein
Anti-RPB4 immunoprecipitation
minus tet
plus tet
cell line: WTTy1 WTTy1 M2Ty1 M3Ty1 M9Ty1
exoRPB1
(anti-Ty1)
loading
control
Immunoblot

Figure S1.

RNA polymerase II assays

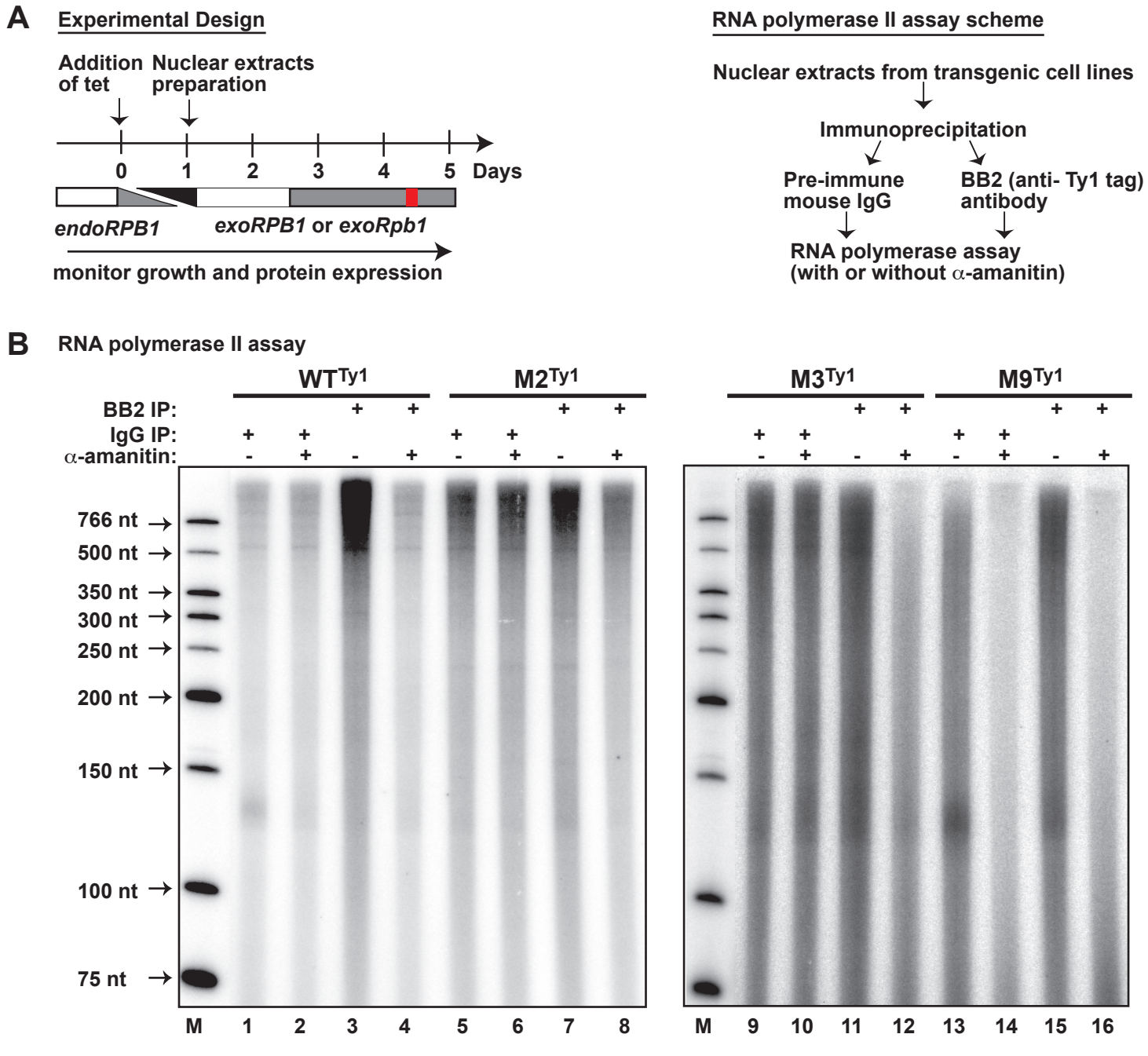
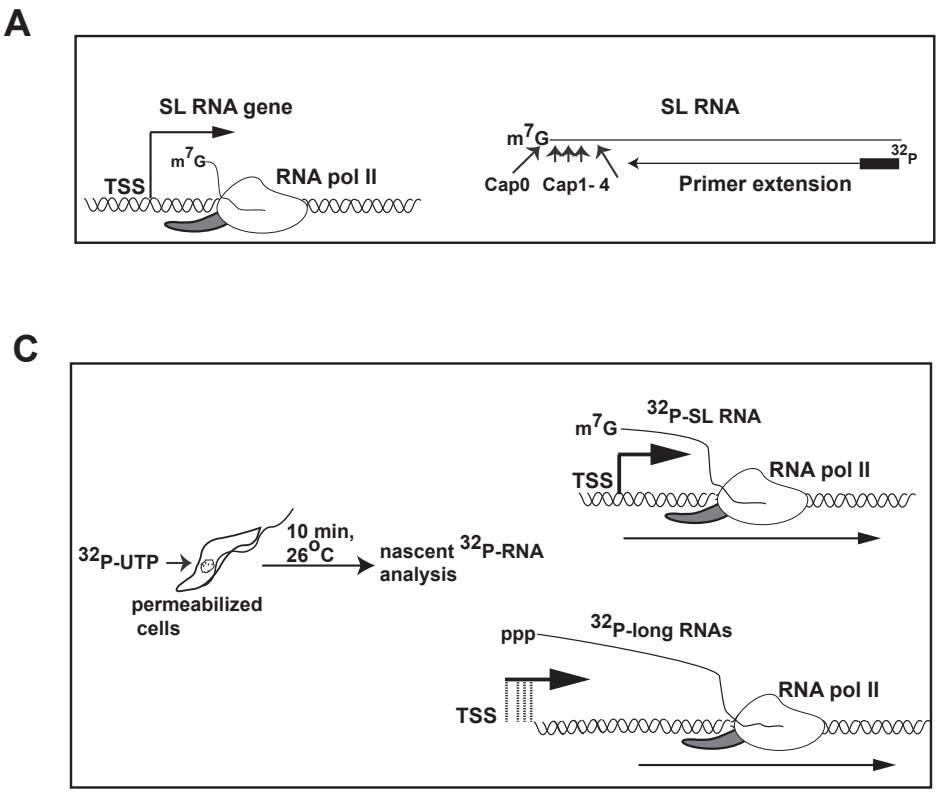
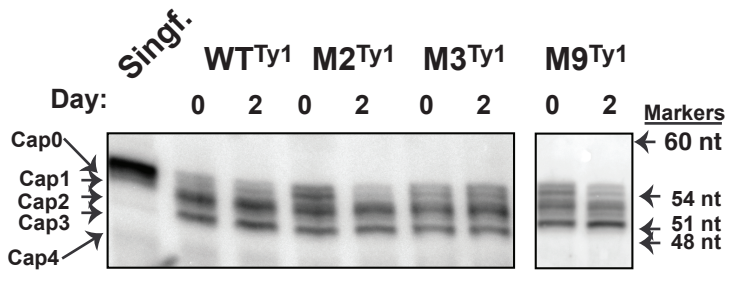


Figure S2.





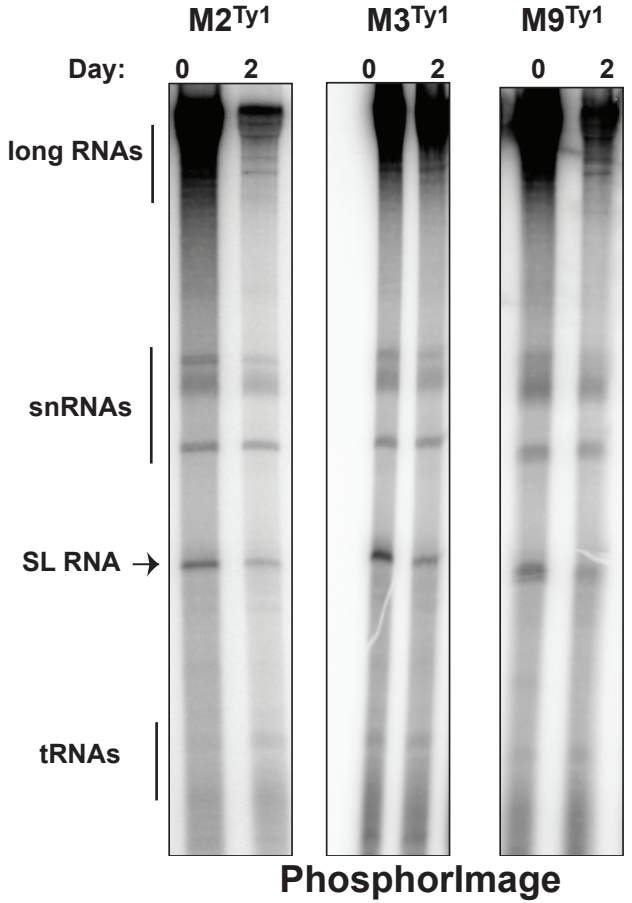


Figure S3.

^{32}P-UTP incorporation on Day 2, compared to Day 0

	M2^{Ty1}	M3^{Ty1}	M9^{Ty1}
% long RNAs	31	91	49
% SL RNA	66	75	81

snRNAs are used as internal control

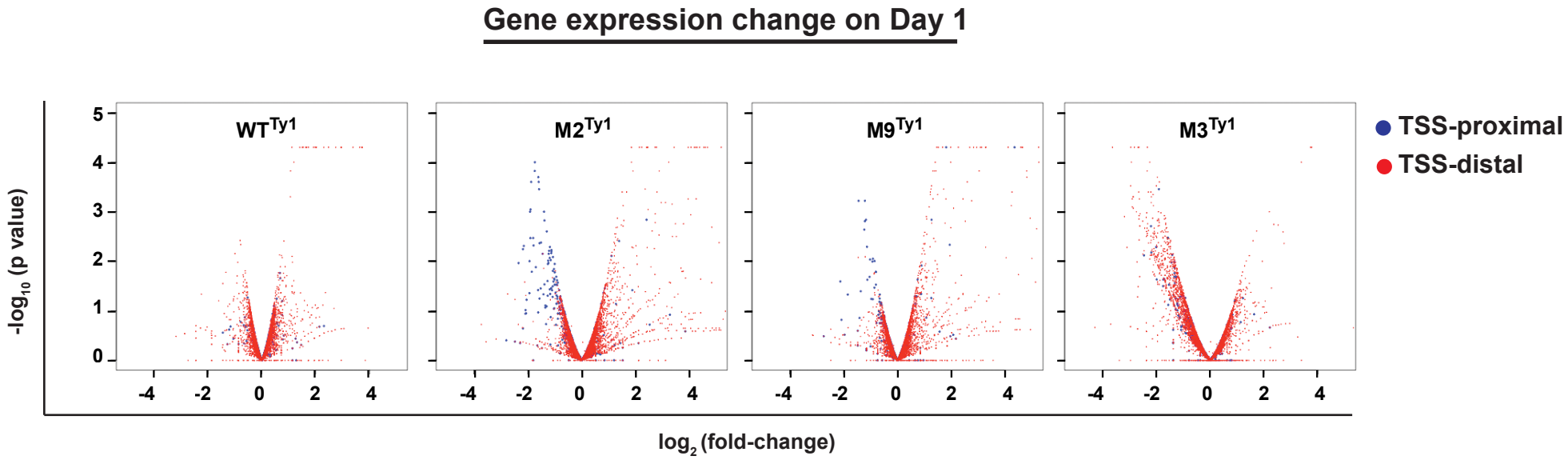

Gene expression change on Day 1
WT Ty1
M2 Ty1
M9 Ty1
M3 Ty1
TSS-proximal
TSS-distal
-log10 (p value)
log2 (fold-change)
0
1
2
3
4
5
-4
-2
0
2
4

Figure S4.

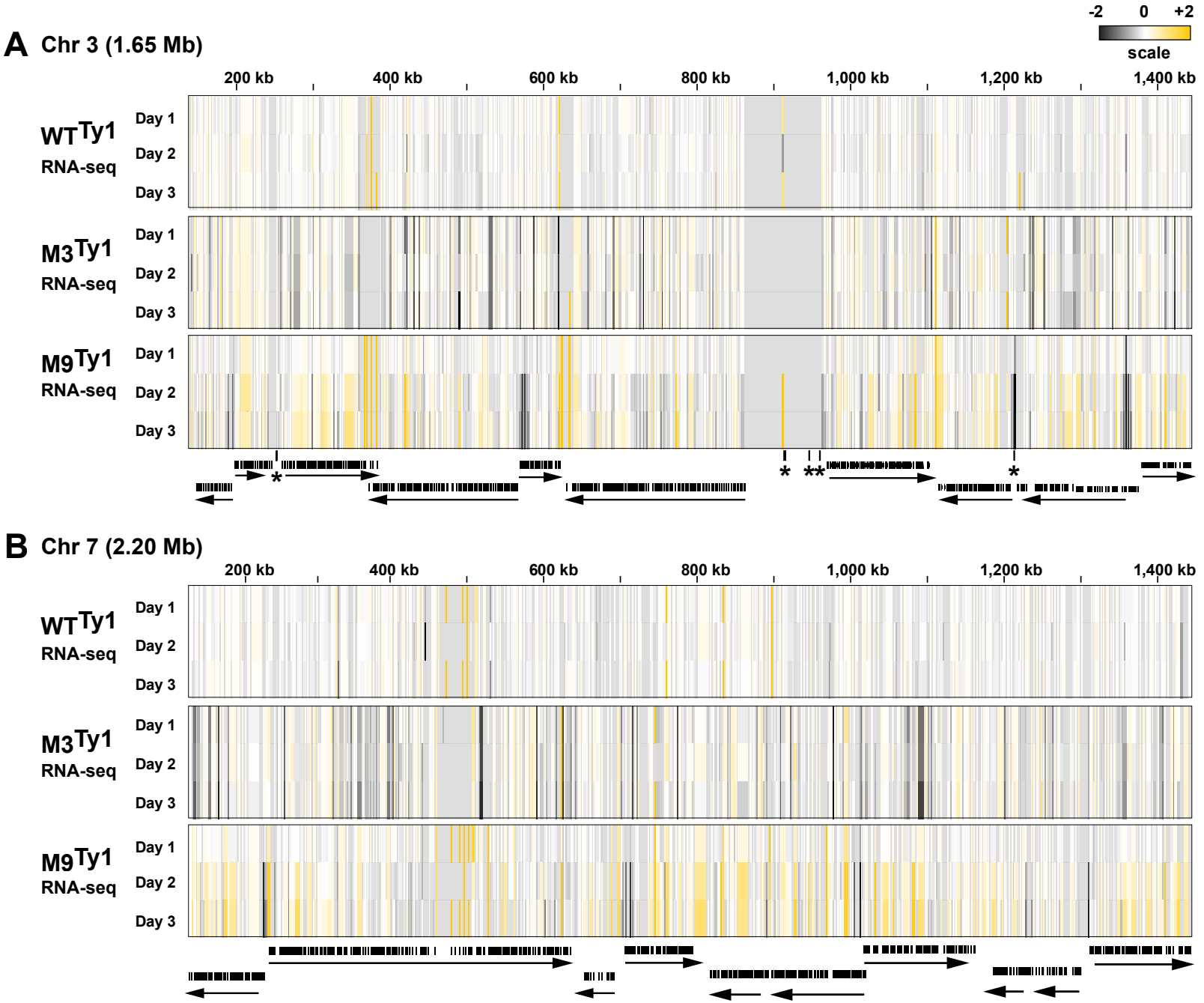

A
Chr 3 (1.65 Mb)
-2
0
+2
scale
200 kb
400 kb
600 kb
800 kb
1,000 kb
1,200 kb
1,400 kb
WT Ty1
RNA-seq
M3 Ty1
RNA-seq
M9 Ty1
RNA-seq
Day 1
Day 2
Day 3
B
Chr 7 (2.20 Mb)
200 kb
400 kb
600 kb
800 kb
1,000 kb
1,200 kb
1,400 kb
WT Ty1
RNA-seq
M3 Ty1
RNA-seq
M9 Ty1
RNA-seq
Day 1
Day 2
Day 3

Figure S5.

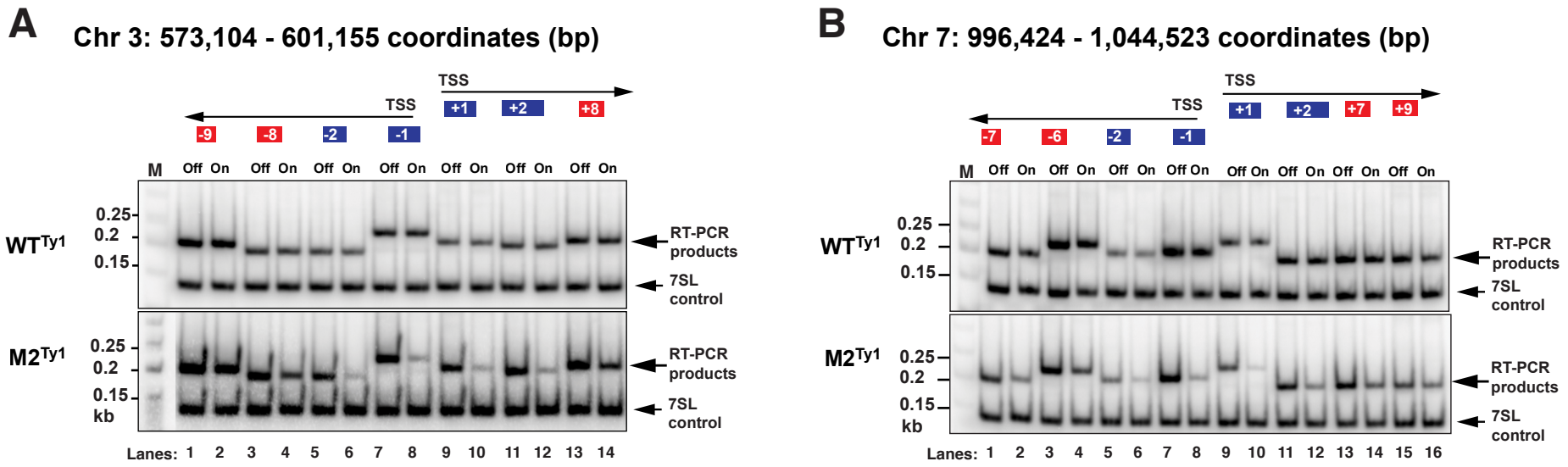
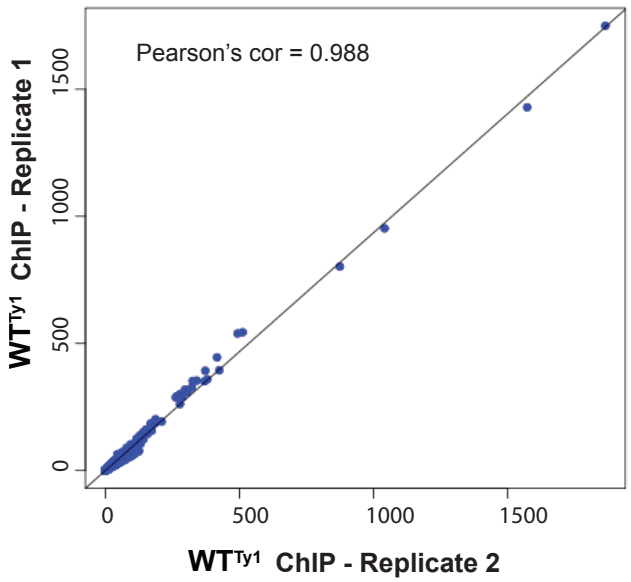
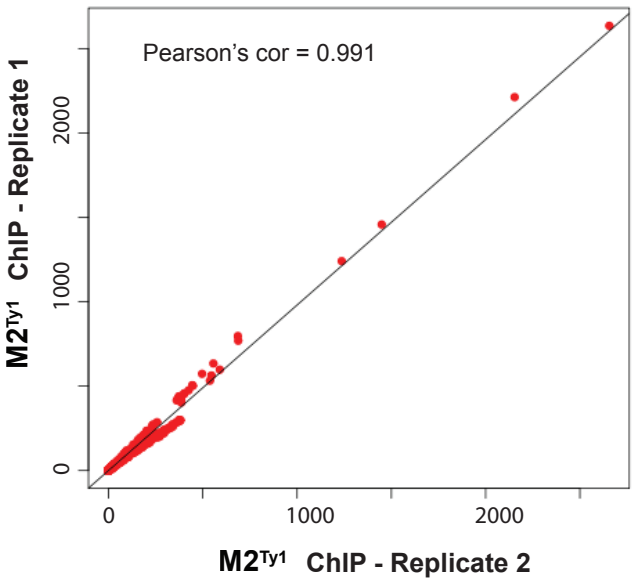
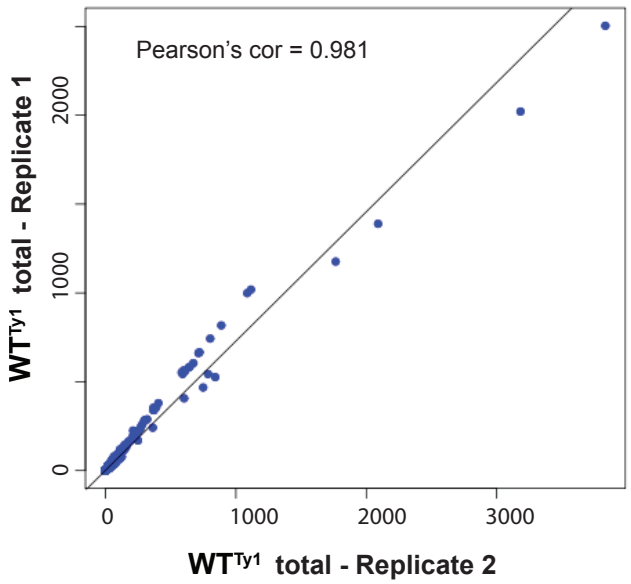

A
Chr 3: 573,104 - 601,155 coordinates (bp)
TSS
TSS
+1
+2
+8
-9
-8
-2
-1
M
Off On Off On Off On Off On Off On Off On Off On
WTTy1
M2Ty1
0.25
0.2
0.15
kb
RT-PCR products
7SL control
Lanes: 1 2 3 4 5 6 7 8 9 10 11 12 13 14
B
Chr 7: 996,424 - 1,044,523 coordinates (bp)
TSS
TSS
+1
+2
+7
+9
-7
-6
-2
-1
M
Off On Off On Off On Off On Off On Off On Off On Off On
WTTy1
M2Ty1
0.25
0.2
0.15
kb
RT-PCR products
7SL control
Lanes: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure S6.


Pearson's cor = 0.988
WTTy1 ChIP - Replicate 1
WTTy1 ChIP - Replicate 2
0
500
1000
1500


Pearson's cor = 0.991
M2Ty1 ChIP - Replicate 1
M2Ty1 ChIP - Replicate 2
0
1000
2000


Pearson's cor = 0.981
WTTy1 total - Replicate 1
WTTy1 total - Replicate 2
0
1000
2000
3000

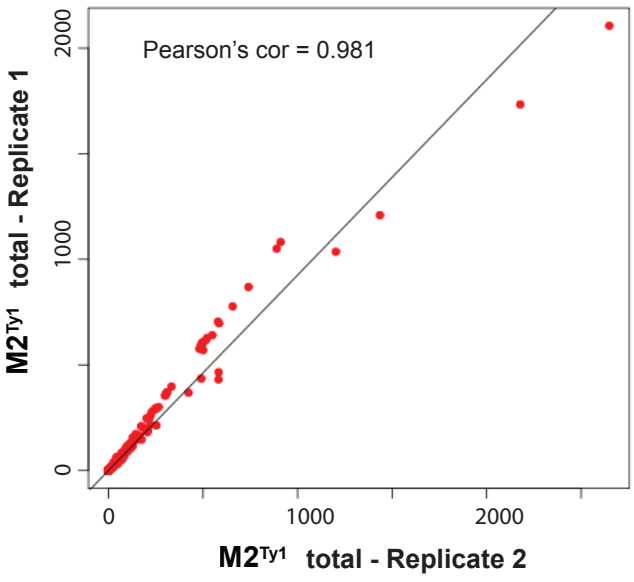
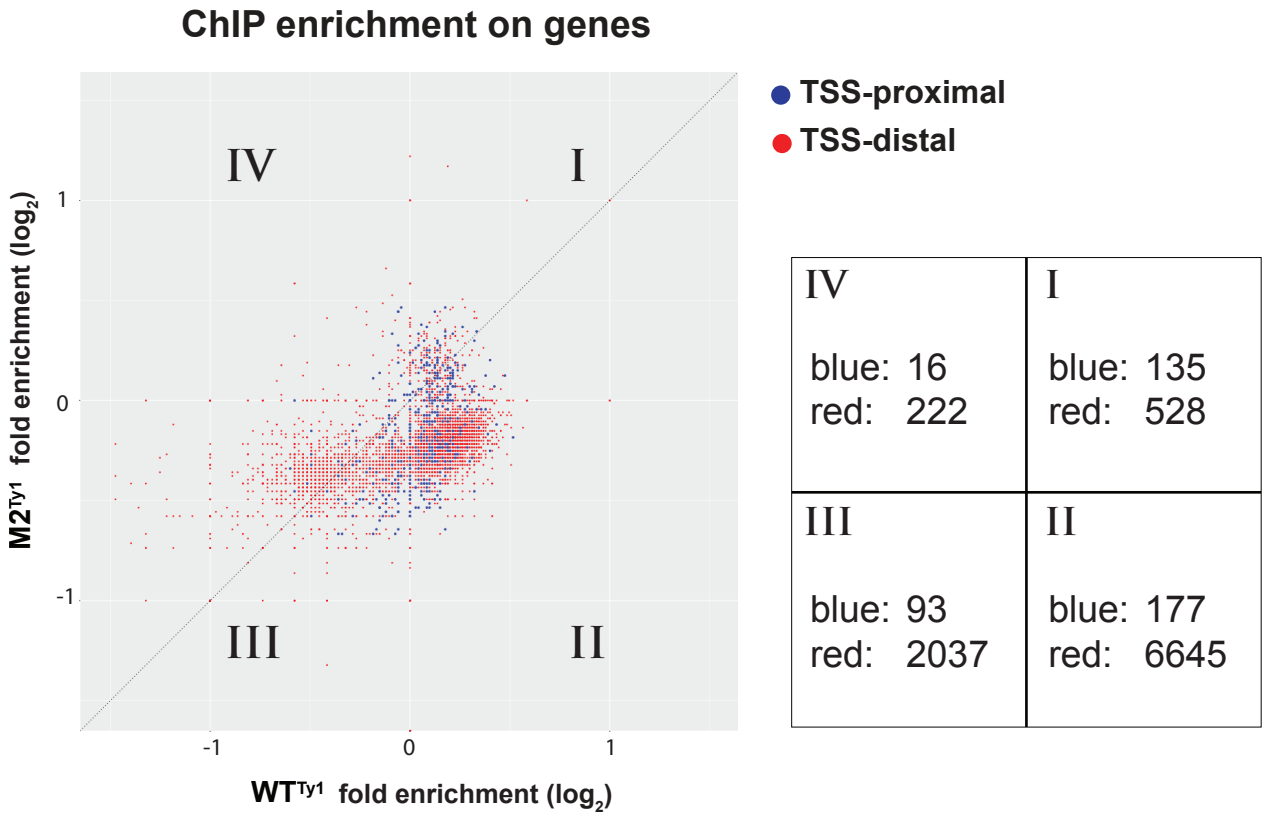
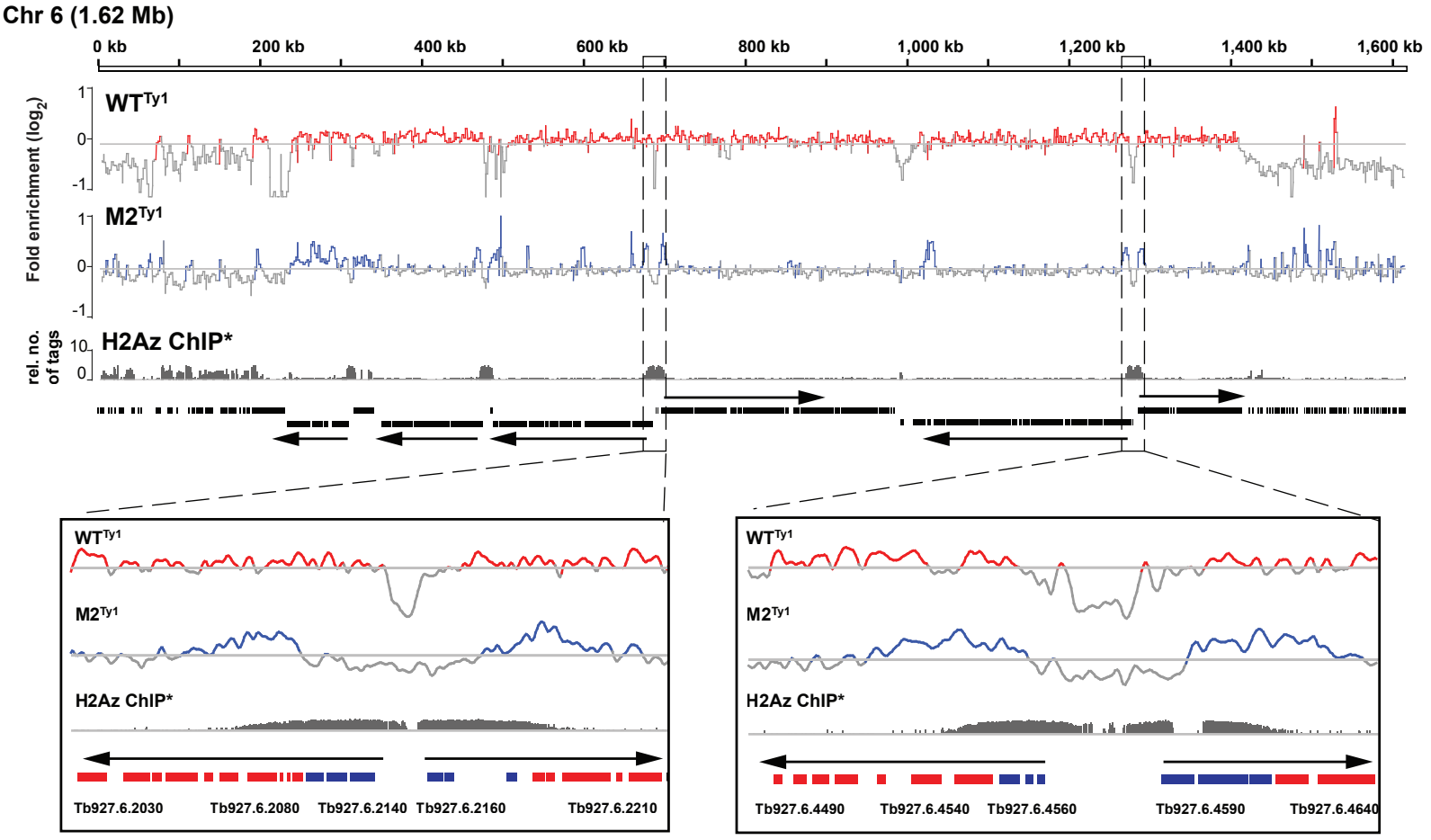

Pearson's cor = 0.981
M2Ty1 total - Replicate 1
M2Ty1 total - Replicate 2
0
1000
2000

Figure S7.



IV	I
blue: 16 red: 222	blue: 135 red: 528
III	II
blue: 93 red: 2037	blue: 177 red: 6645

Figure S8.


Chr 6 (1.62 Mb)
0 kb
200 kb
400 kb
600 kb
800 kb
1,000 kb
1,200 kb
1,400 kb
1,600 kb
Fold enrichment (log2)
WTTy1
M2Ty1
rel. no. of tags
H2Az ChIP*
Tb927.6.2030
Tb927.6.2080
Tb927.6.2140
Tb927.6.2160
Tb927.6.2210
Tb927.6.4490
Tb927.6.4540
Tb927.6.4560
Tb927.6.4590
Tb927.6.4640

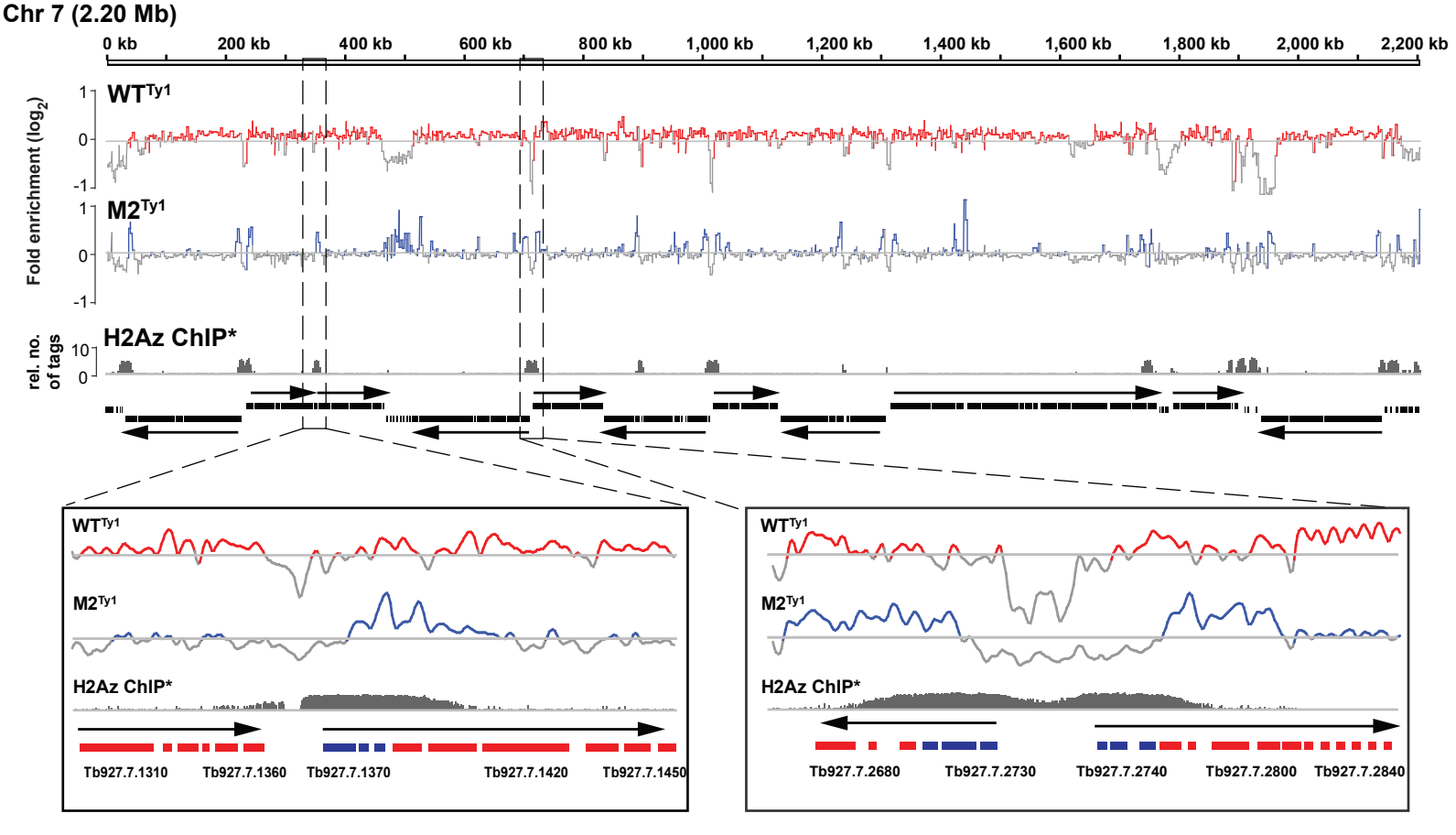

Chr 7 (2.20 Mb)
0 kb
200 kb
400 kb
600 kb
800 kb
1,000 kb
1,200 kb
1,400 kb
1,600 kb
1,800 kb
2,000 kb
2,200 kb
Fold enrichment (log2)
WTTy1
M2Ty1
rel. no. of tags
H2Az ChIP*
Tb927.7.1310
Tb927.7.1360
Tb927.7.1370
Tb927.7.1420
Tb927.7.1450
Tb927.7.2680
Tb927.7.2730
Tb927.7.2740
Tb927.7.2800
Tb927.7.2840

Figure S9.

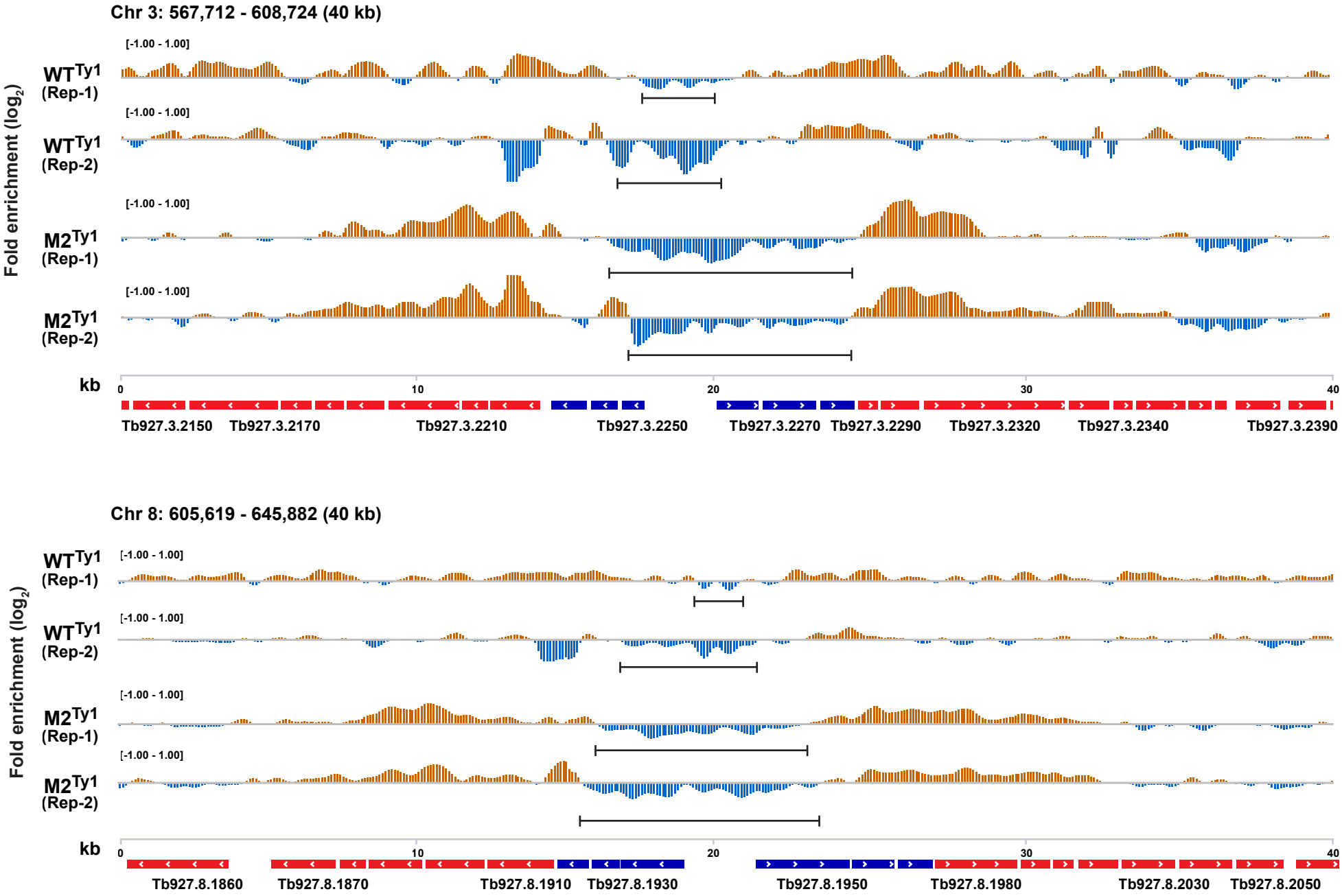

Chr 3: 567,712 - 608,724 (40 kb)
Fold enrichment (log2)
[-1.00 - 1.00]
WTTy1 (Rep-1)
[-1.00 - 1.00]
WTTy1 (Rep-2)
[-1.00 - 1.00]
M2Ty1 (Rep-1)
[-1.00 - 1.00]
M2Ty1 (Rep-2)
kb
0
10
20
30
40
Tb927.3.2150
Tb927.3.2170
Tb927.3.2210
Tb927.3.2250
Tb927.3.2270
Tb927.3.2290
Tb927.3.2320
Tb927.3.2340
Tb927.3.2390
Chr 8: 605,619 - 645,882 (40 kb)
Fold enrichment (log2)
[-1.00 - 1.00]
WTTy1 (Rep-1)
[-1.00 - 1.00]
WTTy1 (Rep-2)
[-1.00 - 1.00]
M2Ty1 (Rep-1)
[-1.00 - 1.00]
M2Ty1 (Rep-2)
kb
0
10
20
30
40
Tb927.8.1860
Tb927.8.1870
Tb927.8.1910
Tb927.8.1930
Tb927.8.1950
Tb927.8.1980
Tb927.8.2030
Tb927.8.2050

Figure S10

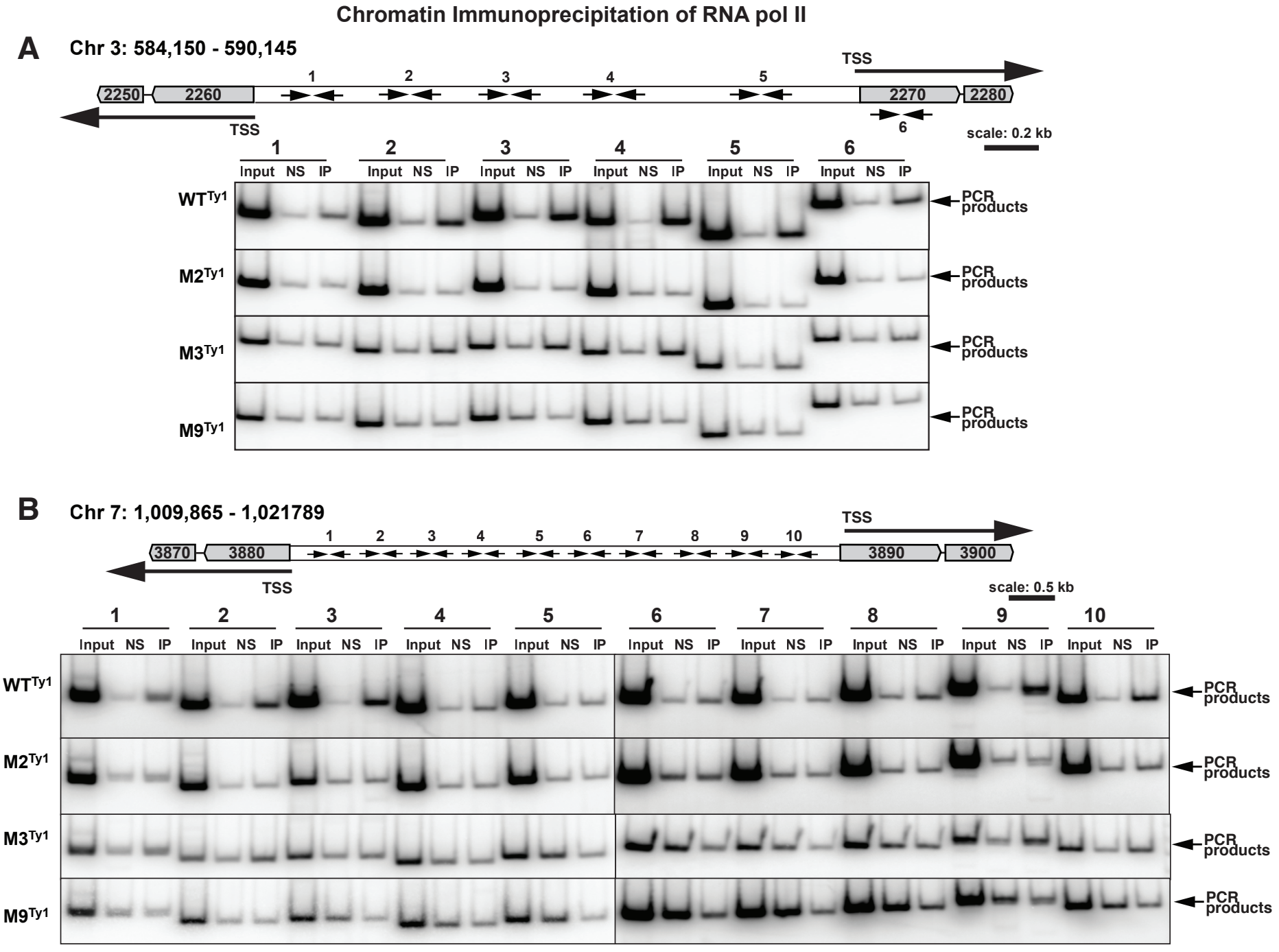

Chromatin Immunoprecipitation of RNA pol II
A
Chr 3: 584,150 - 590,145
TSS
2250
2260
2270
2280
TSS
scale: 0.2 kb
Input NS IP
WT Ty1
M2 Ty1
M3 Ty1
M9 Ty1
PCR products
B
Chr 7: 1,009,865 - 1,021789
3870
3880
3890
3900
TSS
scale: 0.5 kb

Figure S11.

Supplementary Material Figure legends

Figure S1: Description of the molecular genetic system for experimentation.

(A) (Top) Growth curves of Control cell line in the absence (gray squares) and presence (black squares) of tetracycline. (Bottom) Immunoblots show depletion of endoRPB1 (top panel, anti-RPB1) after addition of tetracycline and simultaneous inducible expression of GFP (middle panel, anti-GFP). GFP was the inducible 'control' gene present in the Control cell line. Anti-EF2 immunoblot (bottom panel) is loading control.

(B) Tagged-wild-type CTD or tagged-mutant CTD proteins are incorporated into the RNA pol II complex. (Top) Illustration of RNA pol II showing RPB4 (black; a subunit of RNA pol II), and RPB1 (gray; the largest subunit of RNA pol II) with its CTD (outlined black tail). (Bottom) Immunoblot analysis of RNA pol II, immunoprecipitated with anti-RPB4 antibodies, from whole cell extracts prepared before (minus tet) and after (plus tet) addition with tetracycline. Presence of tagged protein (anti-Ty1) indicates their existence in the RNA pol II complex. An unknown ~25kD protein that cross-reacts with anti-Ty1 is loading control.

Figure S2: Tagged-wild-type and tagged-mutant RNA pol II complexes are transcriptionally competent.

(A) Schematics of RNA pol II activity assay. Tagged-RNA pol II complexes, immunoprecipitated (anti-Ty1) from nuclear extracts, were subjected to transcription (polymerase) assays. Mouse pre-immune IgG served as negative control in immunoprecipitation.

(B) PhosphoImage™ analysis of [32P]-UTP incorporated RNA synthesis indicates transcriptionally competent tagged-RNA pol II complexes in all four cell lines (lanes 3, 7, 11, 15). In each case RNA synthesis activity is inhibited by the RNA pol II inhibitor α-amanitin (100 μg/ml) (lanes 4, 8, 12, 16). RNA synthesis is absent in control reactions (lanes 1, 2, 5, 6, 9, 10, 13, 14). Assay shown is representative of two experiments completed for the cell lines.

Figure S3: SL RNA and long RNA synthesis in transgenic cell lines.

(A) and (B) SL RNA and its cap hypermethylation levels are maintained in all four cell lines, as shown by primer extension analyses of total RNA isolated before (Day 0) and two days after (Day 2) induction of tagged proteins by addition of tetracycline.

(C) and (D) Long RNA synthesis in $M2^{Ty1}$ and $M9^{Ty1}$ cells is significantly reduced after two days of tagged-protein expression, as demonstrated by nascent RNA synthesis ([^{32}P]-UTP incorporation). $M3^{Ty1}$ maintained comparable synthesis of long mRNA. Moderate decreases in SL RNA synthesis were observed in all three cell lines. RNA synthesis activities on Day 2, compared to Day 0, are calculated by

measuring band intensities using ImageJ software. In all cases, data analyses included normalization to comparable cell numbers.

Figure S4: Reduced expression of TSS-proximal genes in $M2^{Ty1}$ and $M9^{Ty1}$ cells is evident after one day of tetracycline induction.

Volcano plots, similar to Figure 2C, for WT^{Ty1}, $M2^{Ty1}$, $M9^{Ty1}$ and $M3^{Ty1}$ cells, showing gene expression (log_2-fold changes) versus adjusted significance p-values (log_{10}) for all protein coding genes (mRNA-Seq). The set of 482 TSS-proximal genes is highlighted in blue; all other genes (~8000) are shown in red.

Figure S5: Heat map views of WT^{Ty1}, $M3^{Ty1}$ and $M9^{Ty1}$ cells.

Heat map views, similar to Figure 2, of gene expression changes (FPKM, log_2 of fold-change) in WT^{Ty1}, $M3^{Ty1}$ and $M9^{Ty1}$ cells from chromosome 3 (A) and chromosome 7 (B) after Day 1, Day 2 and Day 3 of tetracycline induction compared to uninduced cells. Yellow and black show increases and decreases in expression, respectively.

Figure S6: mRNA synthesis alterations from TSS-proximal genes as determined by reverse transcription-PCR analyses.

PhosphorImages™ of reverse transcription-PCR analyses of transcript levels of genes nearby two typical TSSs; one in chromosome 3 (A) and the other in chromosome 7 (B) in WT^{Ty1} (top panel) and $M2^{Ty1}$ (bottom panel) cells before (Off) and after (On) two days of tetracycline addition. Black arrows indicate orientations of polycistronic transcription units, with genes numbered according to their position in each polycistronic unit; positive and negative numbers indicate top and bottom strand coding. Blue identifies TSS-proximal protein coding genes and red indicates TSS-distal protein coding genes. The RNA pol III-dependent 7SL RNA serves as internal control.

Figure S7: Statistical evaluation of ChIP-Seq data.

ChIP-Seq analyses show high correlation between two biological replicate data sets as determined by correlation plot (WT^{Ty1} ChIP, Pearson's cor = 0.988, $M2^{Ty1}$ ChIP, 0.991, WT^{Ty1} total DNA, 0.981, and $M2^{Ty1}$ total DNA, 0.981). Genome-wide reads, covering every 100-base region in the genome, are counted, normalized per million reads and compared.

Figure S8: Scatter plot of ChIP-enrichment for all RNA pol II- dependent coding region genes (open reading frames).

The ChIP data analyzed do not include the RNA pol I-dependent coding regions genes (which exist in a few cases in the Trypanosoma brucei genome), and does not include the RNA pol II-dependent Spliced leader genes (which contain only non-coding RNA). The TSS-proximal genes, defined as the first three open reading frames in a polygenic unit, regardless of their distance from the H2Az, H2BV, H4K10ac and BDF3 marks localized upstream from and within the proximal regions of polygenic transcription units, are in blue (12,13). All other

open reading frames are designated as TSS-distal genes and are shown in red. The purpose of this labeling is to be commensurate with the open reading frame labeling in the RNA-seq analysis of Fig. 2 and 3. Although TSS-distal genes increase in number in the order of quadrant IV, I, III and II, TSS- proximal genes increase in number in the order of quadrant IV, III, I and II. These data suggest differences between the TSS-proximal and distal genes, as defined in the RNA-seq analysis, reflected in the ChIP seq analysis.

Figure S9: Line plot views of chromosome-wide RNA pol II ChIP enrichment data.
Line plot views of chromosome-wide RNA pol II ChIP enrichment relative to total DNA, displayed with IGV browser, on chromosome 6 (top panel) and chromosome 7 (bottom panel) in WT^{Ty1} (positive enrichment, red; negative enrichment, gray) and $M2^{Ty1}$ (positive enrichment, blue; negative enrichment, gray) cells. Histograms (gray) of variant histone H2Az occupancies (relative number of tags) are also shown (14). Black bars and arrows show coding sequences in polycistronic units and their direction of transcription. The four "zoomed-in" views are the regions between the dashed lines. Genes are color coded as in Figure 3.

Figure S10: Reproducibility of ChIP-Seq data.
Histograms, similar to Figure 5, show reproducible RNA pol II ChIP enrichment profile for two biological replicate samples of WT^{Ty1} and $M2^{Ty1}$ cells. Two ~ 40 kb candidate loci, on chromosomes 3 and 8, that are part of Figure 5 are shown. As in Figure 5, the y-axes represent estimated ChIP enrichment (orange, positive; blue, negative). TSS-proximal and TSS-distal genes are shown in blue and red respectively. Arrows indicate orientation of gene transcription. Numbers on top of gene arrays indicate the length (kb) of the chromosomal regions. The horizontal bar under each histogram marks the TSS regions that are deficient in RNA pol II occupancy.

Figure S11: ChIP-PCR analyses of RNA pol II recruitment.
ChIP-PCR analyses indicate RNA pol II recruitment into two TSS regions on chromosome 3 (A), and chromosome 7 (B). PhosphorImages™ of PCR analyses of anti-Ty1 ChIP-enriched DNA from WT^{Ty1}, $M2^{Ty1}$, $M3^{Ty1}$ and $M9^{Ty1}$ cells, following two days of induction are shown. Total DNA (Input), nonspecific (NS) and specific (IP) ChIP reactions are shown. Black arrows indicate PCR primers; numbers indicate corresponding amplicons. Arrows, along with gene IDs, indicate orientations of polycistronic transcription units.

Table S1

Cell line	Background	Genotype	Drug resistance phenotype	Amino acids substitutions in Ty1::rpb1^{Ti} BSD	Growth phenotype
Wild type	Lister 427				normal
29-13	Wild type	TETR HYG T7RNAP NEO	Hygromycin, G418. Parent strain for conditional gene expression.		normal
AD-101	29-13 with a tet-inducible RNAi to target endogenous RPB1 mRNA	TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE	Hygromycin, G418, Phleomycin		lethal
WT^{Ty1}	AD-101	TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::RPB1^{Ti} BSD	Hygromycin, G418, Phleomycin, Blasticidin		normal
M1^{Ty1}	AD-101	TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m1 BSD	Hygromycin, G418, Phleomycin, Blasticidin	$S_{1620}A$ $S_{1621}A$ $S_{1627}A$	lethal
M2^{Ty1}	AD-101	TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m2 BSD	Hygromycin, G418, Phleomycin, Blasticidin	$S_{1591}A$ $S_{1594}A$ $S_{1595}A$ $S_{1597}A$	lethal
M3^{Ty1}	AD-101	TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m3 BSD	Hygromycin, G418, Phleomycin, Blasticidin	$S_{1662}A$ $S_{1663}A$	normal

Table S1

M4^{Ty1}	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m4 BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**	**$S_{1651}A$ $S_{1653}A$**	**normal**
M5^{Ty1}	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m5 BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**	**$S_{1585}A$ $S_{1587}A$ $S_{1591}A$ $S_{1594}A$ $S_{1595}A$ $S_{1597}A$**	**lethal**
M9^{Ty1}	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m9 BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**	**$S_{1651}A$ $S_{1653}A$ $S_{1662}A$ $S_{1663}A$**	**lethal**
M17^{Ty1}	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB^{Ti} BLE *Ty1::rpb1^{Ti}-m17 BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**	**$S_{1612}A$ $S_{1614}A$ $S_{1617}A$ $S_{1620}A$ $S_{1621}A$ $S_{1627}A$**	**lethal**
M20^{Ty1}	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE *Ty1::rpb1^{Ti}-m20 BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**	**$M_{1589}A$ $M_{1593}A$**	**slow**
Control cell line	**AD-101**	**TETR HYG T7RNAP NEO RIT-RPB1^{Ti} BLE GFP^{Ti} BSD**	**Hygromycin, G418, Phleomycin, Blasticidin**		**lethal**

RIT is derived from the term RIT-seq that refers to RNA interference (RNAi) target sequencing. RIT-RPB1^{Ti} describes the use of tetracycline-inducible expression (Ti) of RNAi to ablate endogenous RPB1 mRNA by the targeting of its 3'UTR.

*** Ty1::RPB1^{Ti} BSD and Ty1::rpb1^{Ti} BSD genes are tetracycline-inducible and RNAi-resistant.**

All cell lines are derived from Lister 427 procyclic parasites and are in the 29-13 background, which allows for tetracycline induction of RNAi-target gene knockdown and induction of transgene expression.

Table S2

Cell line	Days after addition of tetracycline	# of biological samples	Total # of reads	# of reads mapped to Ty1 sequence	# of reads mapped to Ty1 sequence, normalized to 10 M	# of reads mapped to endogenous *RPB1* normalized to 10 M	# of reads mapped to mutated RPB1 normalized to 10 M
WT^{Ty1}	0	2	71543618	0	0		
	1	2	136070104	90	7		
	2	2	80145388	56	7		
	3	2	87752232	67	8		
$M2^{Ty1}$	0	2	34920992	1	0	3	1
	1	2	40406616	38	9	0	4
	2	2	32145250	48	15	0	10
	3	2	40412816	27	7	0	5
$M3^{Ty1}$	0	1	9174916	0	0	5	0
	1	1	21832818	9	4	1	3
	2	1	10254464	8	8	2	4
	3	1	22215384	19	9	2	4
$M9^{Ty1}$	0	1	17521490	0	0	2	0
	1	1	24733386	23	9	0	6
	2	1	21927344	9	4	1	7
	3	1	19269712	8	4	1	2

Table S3

Cell line	ChIP-seq library	Biological Replicate #	Total # of reads	% Reads mapped to genome
WT^{Ty1}	Total DNA	1	26346282	84.54
		2	16145206	79.77
	ChIP DNA	1	22906042	79.40
		2	24586106	79.41
$M2^{Ty1}$	Total DNA	1	44508988	78.73
		2	27140334	79.92
	ChIP DNA	1	22840750	66.60
		2	23593758	71.91

Table S4

TSS_ID	1st gene of Leftward TSS	1st gene of Rightward TSS	Start coordinate of first leftward ORF	Start coordinate of first rightward ORF	Comments
Chr 1.1 (L-R)	Tb927.1.890	Tb927.1.990	279928	288567	
Chr 1.2 (R)		Tb927.1.2100		531517	
Chr 1.3 (L-R)	Tb927.1.2880	Tb927.1.2950	651264	655620	
Chr 1.4 (L)	Tb927.1.3560		761284		
Chr 1.5 (R)		Tb927.1.3800		814232	Downstream from rRNA
Chr 1.6 (R)		Tb927.5000		1004241	
Chr 2.1 (L-R)	Tb927.2.1600	Tb927.2.1680	307642	315628	
Chr 2.2 (L-R)	Tb927.2.2590	Tb927.2.2650	511483	519545	
Chr 2.3 (L-R)	Tb927.2.5080	Tb927.2.5120	898935	905712	
Chr 2.4 (L-R)	Tb927.2.5540	Tb927.2.5600	995768	998939	
Chr 2.5 (R)		Tb927.2.5750		1031745	Downstream from U2 snRNAs
Chr 3.1 (L)	Tb927.3.580		128330		
Chr 3.2 (L-R)	Tb927.3.860	Tb927.3.870	201908	206154	
Chr 3.3 (R)		Tb927.3.1070		263813	
Chr 3.4 (L-R)	Tb927.3.2260	Tb927.3.2270	586295	587852	
Chr 3.5 (L)	Tb927.3.3410		874336		
Chr 3.6 (R)		Tb927.3.3440		976028	

Table S4

Chr 3.7 (L)	Tb927.3.4390		1229349		Downstream from tRNA
Chr 3.8 (L-R)	Tb927.3.4890	Tb927.3.4900	1373571	1375295	
Chr 4.1 (L)	Tb927.4.1190		314305		
Chr 4.2 (L-R)	Tb927.4.2080	Tb927.4.2100	535378	541988	
Chr 4.3 (L-R)	Tb927.4.3740	Tb927.3750	947054	971540	
Chr 4.4 (L)	Tb927.4.4660		1282661		
Chr 4.5 (L)	Tb927.4.5070		1390647		
Chr 4.6 (L)	Tb927.4.5390		1467499		
Chr 5.1 (L)	Tb927.5.600		242846		
Chr 5.2 (L-R)	Tb927.5.630	Tb927.5.640	269443	274270	
Chr 5.3 (R)		Tb927.5.670		295064	
Chr 5.4 (R)		Tb927.5.950		377913	
Chr 5.5 (L-R)	Tb927.5.1580	Tb927.5.1590	555436	562254	
Chr 5.6 (L-R)	Tb927.5.2150	Tb927.5.2160	737918	747897	
Chr 5.7 (L-R)	Tb927.5.2900	Tb927.5.2910	973702	980389	
Chr 5.8 (L-R)	Tb927.5.3500	Tb927.5.3510	1157315	1159513	
Chr 5.9 (L-R)	Tb927.5.4530	*Tb927.5.4540	1389431	*1398296	*Rightward 1st ORF is redundant (GRESAG family member)
Chr 6.1 (L-R)	Tb927.6.1290	Tb06.3A7.960	476573	484320	

Table S4

Chr 6.2 (L-R)	Tb927.6.2150	Tb927.6.2160	685951	689752	
Chr 6.3 (L-R)	Tb927.6.4580	Tb927.6.4590	1278304	1285448	
Chr 7.1 (L-R)	Tb927.7.920	Tb927.7.930	229403	235766	
Chr 7.2 (L-R)	Tb927.7.2730	Tb927.7.2740	711160	716949	
Chr 7.3 (L)	Tb927.7.3470		895,705		Downstream from a large non-coding region
Chr 7.4 (L-R)	Tb927.7.3880	Tb927.7.3890	1013751	1019310	
Chr 7.5 (L-R)	Tb927.7.4670	Tb927.4680	1,238,531	1243295	
Chr 7.6 (L-R)	Tb927.7.4960	Tb927.7.4970	1,308,074	1317985	
Chr 7.7 (L-R)	Tb927.7.6430	Tb927.7.6440	1,746,856	1748200	
Chr 7.8 (L)	Tb927.7.6820		1,889,270		
Chr 7.9 (L)	Tb927.7.6840		1,900,547		Downstream from a tRNA
Chr 7.10 (L-R)	Tb927.7.6850#	*Tb927.7.6860	1,920,893	1928354	*7.6860 is followed by rRNAs
Chr 7.11 (L)	Tb927.7.7460		2,142,378		
Chr 7.12 (L-R)	Tb927.7.7480	Tb927.7.7490	2,159,028	2164933	
Chr 8.1 (L-R)	Tb927.8.1370	*Tb927.1410	445557	*464007	*Downstream from rRNA
Chr 8.2 (L-R)	Tb927.8.1940	Tb927.8.1950	624115	626528	
Chr 8.3 (R)		Tb927.8.2670		794576	Downstream from two tRNAs
Chr 8.4 (L-R)	Tb927.8.2880	Tb927.8.2890	876190	877265	
Chr 8.5 (L)	Tb927.8.3930		1172928		

Table S4

Chr 8.6 (L-R)	Tb927.8.4890	Tb927.8.4900	1443256	1450507	
Chr 8.7(L-R)	Tb927.8.5430	Tb927.8.5440	1608351	1611644	
Chr 8.8 (R)		Tb927.8.5940		1742397	
Chr 8.9 (L-R)	Tb927.8.6920	Tb927.8.6930	1994876	1999215	
Chr 8.10 (L-R)	Tb927.8.7350	Tb927.8.7360	2112076	2116042	
Chr 8.11 (R)		Tb927.8.7750		2244583	
Chr 9.1 (L)	Tb927.9.1960		439948		
Chr 9.2 (L)	Tb927.9.3170		686,315		
Chr 9.3 (R)		Tb927.9.3280		693330	
Chr 9.4 (R)		Tb927.9.3740		747221	
Chr 9.5 (R)		Tb927.9.4860		880706	
Chr 9.6 (R)		Tb927.9.6830		1122621	
Chr 9.7 (L)	Tb927.9.8880		1429144		
Chr 9.8 (L)	Tb927.9.9870		1566703		
Chr 9.9 (L)	Tb927.9.11150		1768768		
Chr 9.10 (R)		Tb927.9.11220		1774913	
Chr 9.11 (L-R)	Tb927.9.14520	Tb927.9.14530	2266753	2267602	
Chr 9.12 (R)		Tb927.9.14990		2341418	
Chr 9.13 (L-R)	Tb927.9.15890	Tb927.9.15920	2517584	2524466	

Table S4

Chr 10.1 (L-R)	Tb927.10.120	Tb927.10.130	56675	58919	
Chr 10.2 (L)	Tb927.10.1220		329260		
Chr 10.3 (L-R)	Tb927.10.1670	Tb927.10.1680	448542	453284	
Chr 10.4 (R)		Tb927.10.1800		476008	
Chr 10.5 (L-R)	Tb927.10.2450	Tb927.10.2460	636054	638192	
Chr 10.6 (R)		Tb927.10.3070		802451	
Chr 10.7 (L-R)	Tb927.10.4180	Tb927.10.4190	1100900	1104360	
Chr 10.8 (L-R)	Tb927.10.4950	Tb927.10.4970	1230829	1236490	
Chr 10.9 (L-R)	Tb927.10.6410	Tb927.10.6440	1623215	1635920	
Chr 10.10 (L)	Tb927.10.7500		1885637		
Chr 10.11 (L)	Tb927.10.7630		1923152		
Chr 10.12 (L-R)	Tb927.10.8330	Tb927.10.8350	2058838	2064075	
Chr 10.13 (R)		Tb927.10.9650		2377684	
Chr 10.14 (L)	Tb927.10.10850		2638240		
Chr 10.15 (L-R)	Tb927.10.11270	Tb927.10.11280	2753738	2758131	
Chr 10.16 (L)	Tb927.10.12550		3045879		
Chr 10.17 (L)	Tb927.10.13020		3164862		
Chr 10.18 (L)	Tb927.10.13520		3302553		
Chr 10.19 (L)	Tb927.10.14040		3420396		
Chr 10.20 (L-R)	Tb927.10.14810	Tb927.10.14850	3611738	3620341	

Table S4

Chr 11.1 (L-R)	Tb927.11.720	Tb927.11.730	175002	181241	
Chr 11.2 (L-R)	Tb927.11.1920	Tb927.11.1930	562256	567362	
Chr 11.3 (L-R)	Tb927.11.3200	Tb927.11.3230	909106	922726	
Chr 11.4 (R)		Tb927.11.3570		1057312	
Chr 11.5 (L-R)	Tb927.11.4430	Tb927.11.4450	1287291	1291905	
Chr 11.6 (R)		Tb927.11.4780		1378218	
Chr 11.7 (L)	Tb927.11.6310		1775798		
Chr 11.8 (L)	Tb927.11.6720		1897334		Downstream from tRNA
Chr 11.9 (L-R)	Tb927.11.7970	Tb927.11.8000	2245624	2251912	
Chr 11.10 (R)		Tb927.11.20730		2436211	
Chr 11.11 (L-R)	Tb927.11.9810	Tb927.11.9820	2650229	2657285	
Chr 11.12 (R)		Tb927.11.11080		2966102	
Chr 11.13 (R)		Tb927.11.11430		3058873	
Chr 11.14 (R)		Tb927.11.12020		3222630	
Chr 11.15 (L)	Tb927.11.13120		3510032		Downstream from tRNA
Chr 11.16 (L)	Tb927.11.13150		3526193		
Chr 11.17 (L-R)	Tb927.11.13710	Tb927.11.13730	3651576	3655817	
Chr 11.18 (L)	Tb927.11.15370		4078384		
Chr 11.19 (L)	Tb927.11.16170		4279478		
Chr 11.20 (L-R)	Tb927.11.17020	Tb927.11.17040	4541418	4549949	

Table S5

Primers for RT-PCR analysis		
7SL_SRP RNA	AD703	TTG GTG TTC TGC TTG GTT GC
	AD704	GTG CTT CTG CAA CGG GAA C
ORFs proximal and distal to Chr 3.4 (L-R) TSSs		
Tb927.3.2180	AD714	CGCTTTCCCAATGAATGACC
	AD715	TCCGTTAGTTCCGTTGACGA
Tb927.3.2190	AD691	CTGGGAAGCCCAGTGAATGC
	AD692	GTCACGCTCAAGCGGTCCTT
Tb927.3.2250	AD693	AATGCGTGTCAGGGGCAACT
	AD694	GACATGCGGACGACAAATGC
Tb927.3.2260	AD695	GCCCGAAAACGGTTTGAAGA
	AD696	TACGCTGCTCCCAGCCAGAT
Tb927.3.2270	AD697	ATGCGTGGGTGACGATTTCC
	AD698	ATCGACTGCACCACCACCAA
Tb927.3.2280	AD699	ACACCGTCGCTCACCCAAAT
	AD700	GTGGAGAACGGACCCGACAT
Tb927.3.2340	AD701	CCGCGCTCGTCTTCTTTTGT
	AD702	GCAAGGAAGTGCAGGGTGTG
ORFs proximal and distal to Chr 7.4 (L-R) TSSs		
Tb927.7.3820	AD761	AATGTGGTTCCTCCCAAACG
	AD762	AGACAATGATGCGGACGACA
Tb927.7.3830	AD763	GAACGCACAAACTCGCACTC
	AD764	AAGGCATGCATCACAAGTGG
Tb927.7.3870	AD765	CCCGGTTACGTGGCAATAAG
	AD766	GGATTTCGCTCTGCCGTAAA
Tb927.7.3880	AD767	GCGAAACCCGGTAACTGCTA
	AD768	CCGAAAAGTCTGCTCCCATC
Tb927.7.3890	AD769	CCGCAATTCTGCACTCATGT
	AD770	TCGCAACCGTGGAGAGAATA
Tb927.7.3900	AD771	GGGGCTTGCATCGATATGTT
	AD772	AGGTCCTTTCATTGGCGAAA
Tb927.7.3950	AD773	AGATGCAGGAGGCTCATCGT
	AD774	CACAGCGTTCTTCTCCAACG
Tb927.7.3970	AD775	GCCTTCGGCTATTGAACGTC
	AD776	CTTGCGAAATGCCACAAGAA

Table S5

Primers for ChIP-PCR analysis		
Primers for Chr 3.4 (L-R) region		
Fragment 1	AD746	CAT ATC CAT GCC CGA AAA AG
	AD747	CGG TCC TGT CAC TTT CGA TT
Fragment 2	AD748	ATT GAA GGA ATG GCA GGA TG
	AD749	CTG GCC TTT TCA TTC CCT TT
Fragment 3	AD750	CGA CCC GTT CTA CGT TTG TT
	AD751m	AGT CGC TGT GGG CCG TTA C
Fragment 4	AD752m	ACC GCT ATG TTT GGT TCA CT
	AD753m	G GCA GAT TTT CGT CTT CAA AC
Fragment 5	AD754	GTG CAC ACG CAC GCA TGT AA
	AD755m	CAA GAA GCG CCG TGA AGC AC
Fragment 6	AD756	CAT GGA GCA ACG CTA ACA TT
	AD757	TGC GGC ATA GCT GCT AAG AG
Primers for Chr 7.4 (L-R) region		
Fragment 1	AD777	TAGGAAGCCAAGCGGATGAC
	AD778	CGGTAAGGTTCCGTGCAAGT
Fragment 2	AD779	AAGAATGCGGTGACAGACGA
	AD780	ACATCGGCCCAAGGTAACTC
Fragment 3	AD781	CCACACGCAATTTCAACCAC
	AD782	CTTTGGGTGGGTTTTTGCTC
Fragment 4	AD783	CGAAACTGATCGTGCAAAGC
	AD784	AGTGTCTGAGGTCGCTGCAA
Fragment 5	AD785	CCACGCCTTGCATATTAACG
	AD786	TTACGAGGCCGAAACAAAGG
Fragment 6	AD787	AACGAAGGGTGAGCTCTGGA
	AD788	GGCTCAACTGATGCCTTTCC
Fragment 7	AD789	GACATAACAGTCCGCGGTGA
	AD790	GCGTATGTGCGCTGCTAAAC
Fragment 8	AD791	GGCACCTTTCAAACAACGTG
	AD792	GACCGCACTGTGAAGCAAAA
Fragment 9	AD793	ACTGCGACGTGCTGATGAAG
	AD794	AGCATGGTGACCGTTTGTTG
Fragment 10	AD795	AAGCTGCGGCTACCTAATCG
	AD796	ACCCACCGTTCTCACTCAAG

Supplementary Material Table legends

Table S1: List of cell lines with the corresponding mutations used in this study and their growth phenotypes. All cell lines are derived from *T. brucei* Lister 427 parasites and in the 29-13 background, which enables tetracycline-induction of both RNAi-mediated endoRPB1 knockdown and expression of all exogenous transgenes. RNA interference (RNAi) target (RIT)-RPB1^{Ti} indicates that the cell line contains a tetracycline-inducible RNAi that ablates endogenous RPB1 mRNA by targeting its 3'UTR. Asterisk (*) indicates Ty1::RPB1^{Ti} BSD and Ty1::rpb1^{Ti} BSD genes are tetracycline-inducible and RNAi-resistant.

Table S2: RNA abundance of endogenous RPB1 and tagged proteins in four cell lines. The table shows number of biological samples (column 3) and total number of reads for each RNA-seq data set (column 4). Total read number and normalized read number mapped to the Ty1-tag sequence are shown (columns 5 and 6). Normalized read number mapped to wild-type (column 7) or the mutated RPB1 sequence in the M2^{Ty1}, M3^{Ty1} or M9^{Ty1} cell line (column 8), before and after 1 to 3 days of tetracycline induction, are shown. In the WT^{Ty1} RNA-seq data set, reads mapped to the endogenous and tagged-protein (i.e., exogenous) genes cannot be distinguished.

Table S3: ChIP-seq data from WT^{Ty1} and M2^{Ty1} cell lines. The table shows total number of mapped reads for each ChIP-seq data set (column 4). The percent of mapped reads from total DNA and ChIP DNA from two biological replicate samples of WT^{Ty1} and M2^{Ty1} are shown (column 5).

Table S4: Table contains the 167 TSS used to define genes proximal and distal to putative RNA pol II transcription initiation regions in the genome. This list includes the majority (167 vs. 191) of the putative RNA pol II transcription units identified in 5' end mapping in reference 12 (Table S14. Doi:10.1371). We omitted single-gene units; updated "1st gene of a TSS" based on visual evaluation of current data sets in TriTrypDB. Asterisks in column 2 or 5 refer to information in column 6.

Table S5: The oligonucleotide primers used in the reverse transcription-PCR analysis (Fig. S4) and ChIP-PCR analysis (Fig. S7) are listed.