Supplementary Figure 1: Raw scores used for calculation of TTAS for five transcription factors depending on the distance from transcription start site of the associated gene

Supplementary Figure 2: Comparison of of Pol II occupancy and RNA levels as measured by RNA--seq and microarrays and their reproducibility over 21,854 RefSeq genes

Supplementary figure 3: Examples of Pol II and RNA-seq coverage over differentially expressed genes under two experimental conditions.

Supplementary Figure 4: Correlation of TF occupancy with Pol II ChIP-seq and RNA-seq data from ENCODE datasets.

Supplementary Figure 5: Effect of poised polymerase and polymerase accumulated on 3'-UTRs on the correlation of Pol II ChIP-seq with transcription factor occupancy

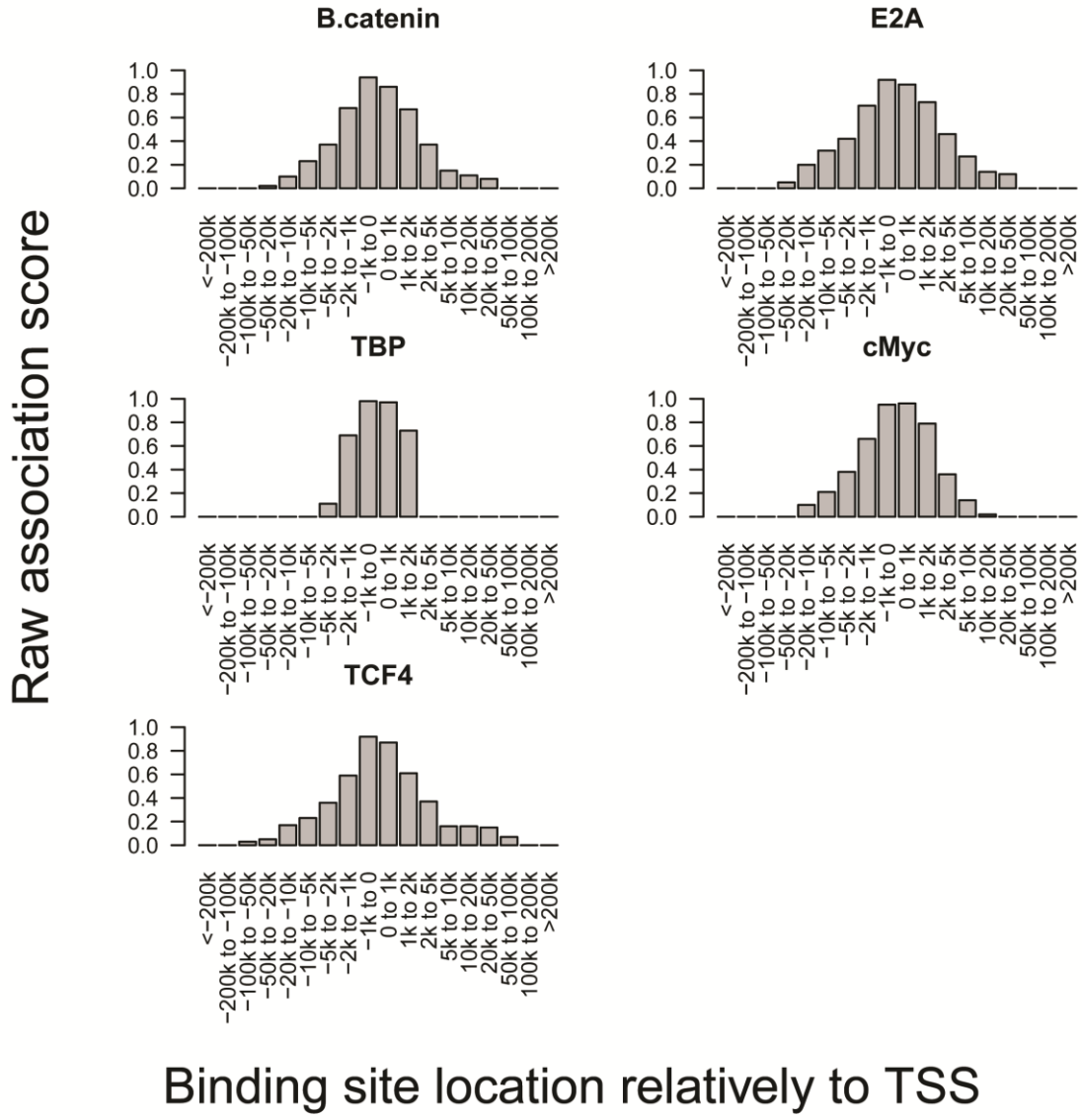
Supplementary Figure 6 Pol II occupancy on chromosome 1

Supplementary Figure 7 Rank analysis of Pol II occupancy with empirically defined gene classes (I, II, III)

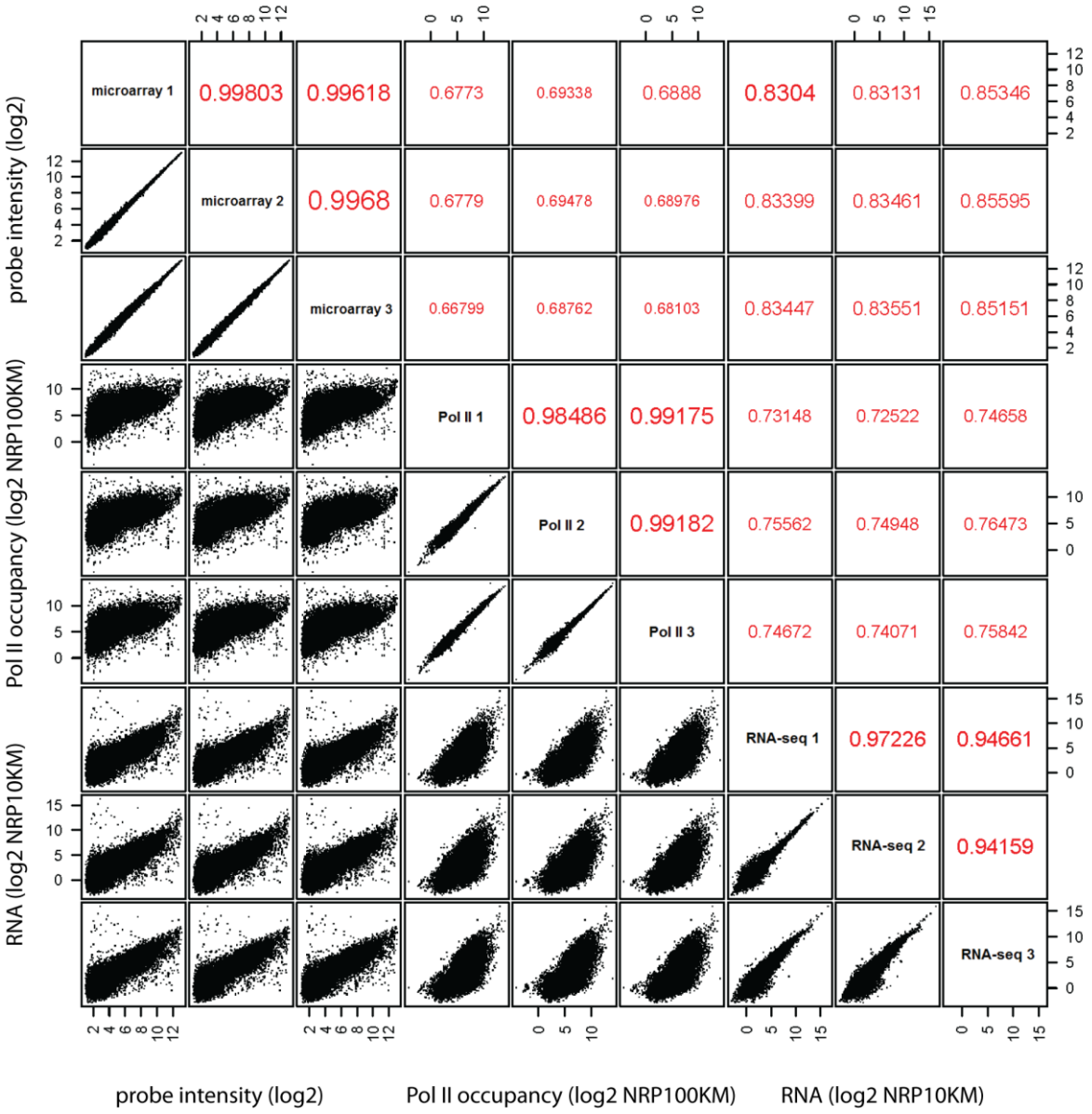
Supplementary Table Sequencing and mapping statistics

Supplementary Material 1: Lists of transcription factor binding sites called from ChIP-seq experiments
Supplementary Material 2: Datasets used from ENCODE project
Supplementary Material 3: Enrichment of GO terms in Wnt target genes identified by 3 different methods
Supplementary Material 4: Enriched GO terms in 3 gene classes

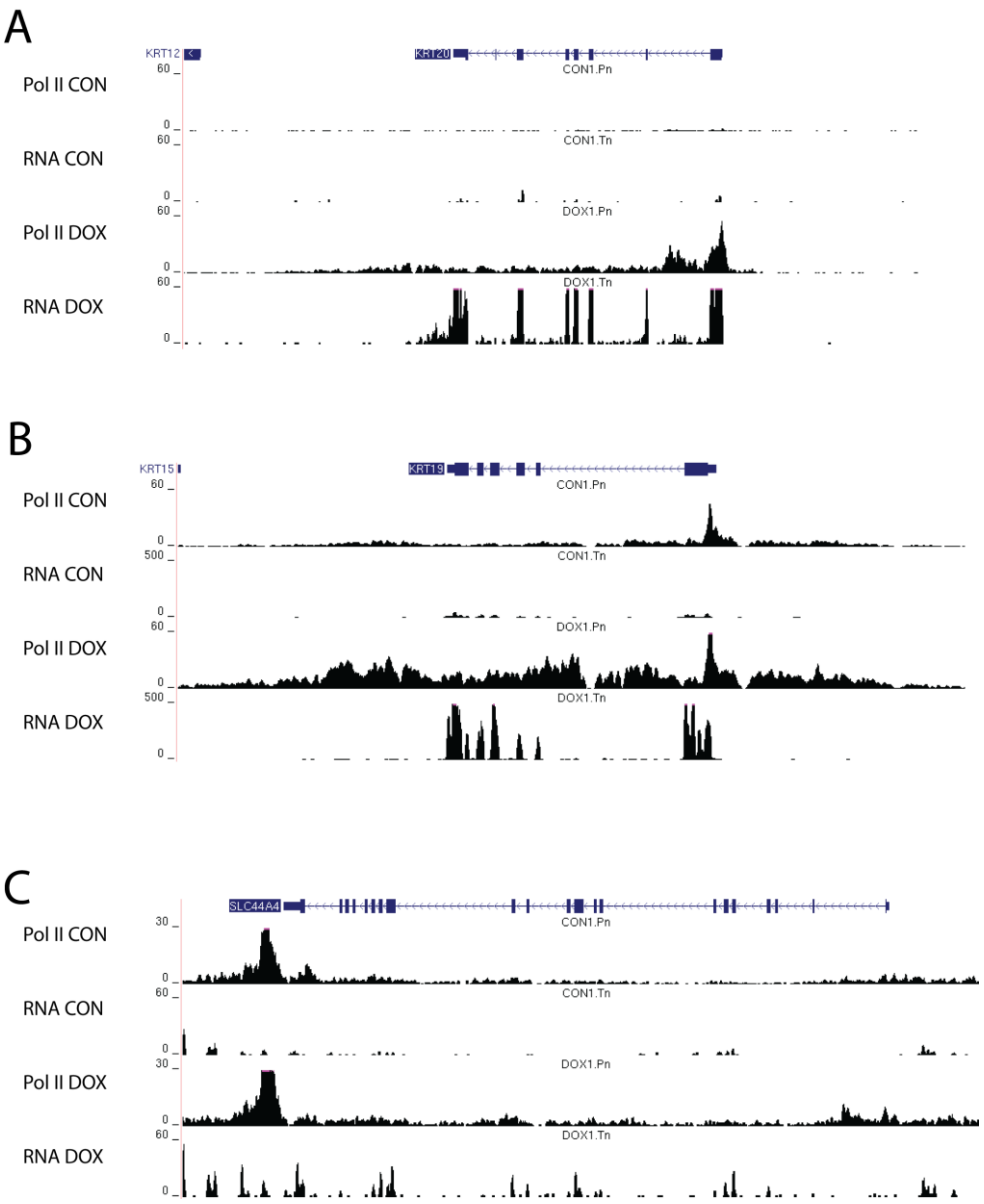
Supplementary Figures



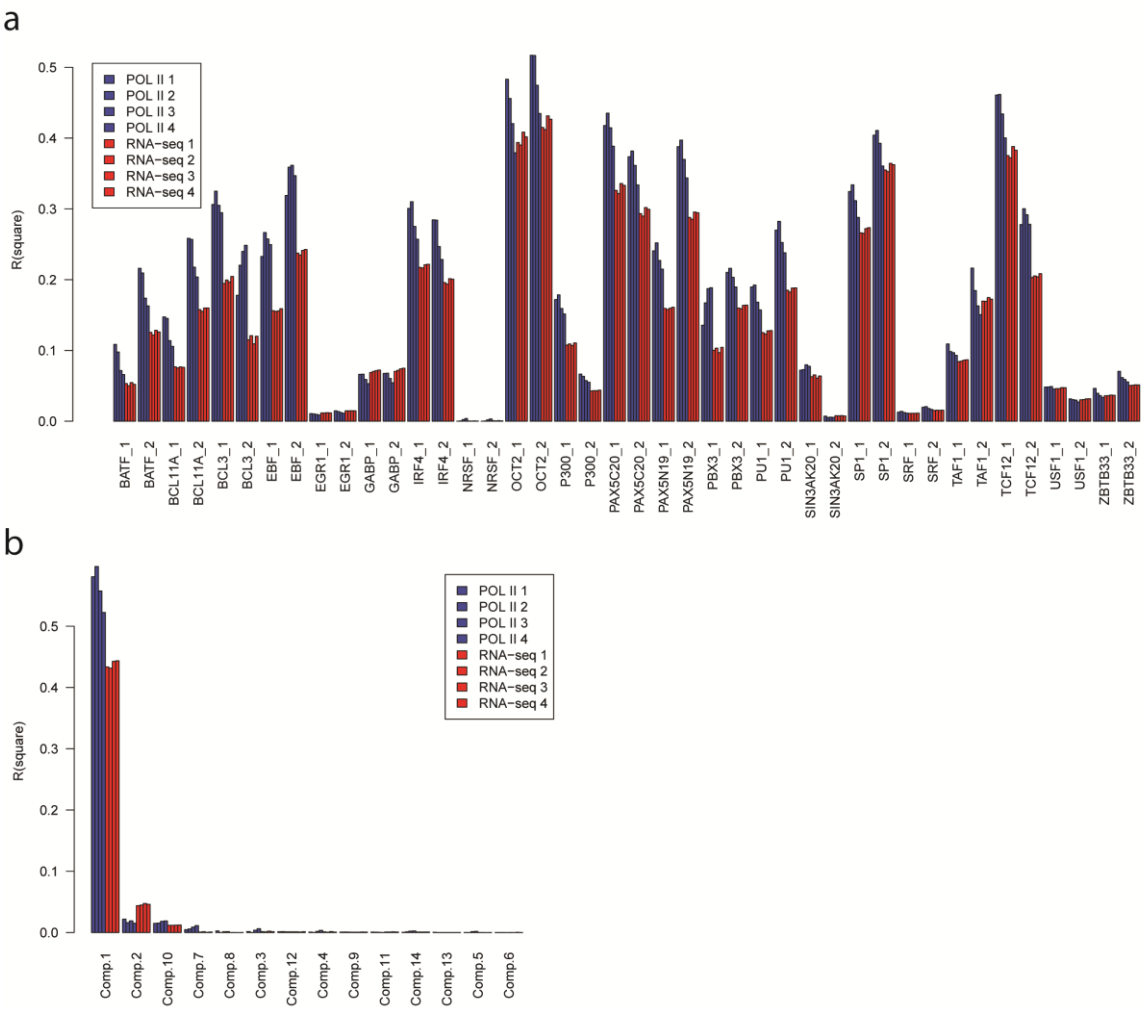
Supplementary Figure 1: Raw scores used for calculation of TTAS for five transcription factors depending on the distance from transcription start site of the associated gene. TTAS scores reflect the likelihood of a gene being regulated by a given transcription factor within set distance from gene's TSS. Some transcription factor binding sites are distributed exclusively close to transcription start sites. Others can regulate more distant genes.



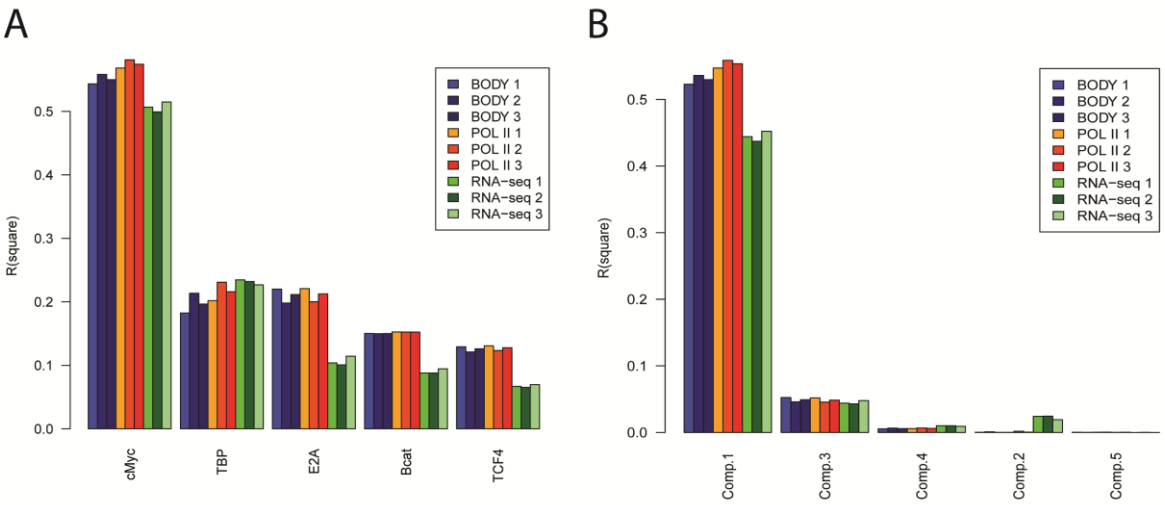
Supplementary Figure 2: Comparison of of Pol II occupancy and RNA levels as measured by RNA--seq and microarrays and their reproducibility over 21,854 RefSeq genes. Each method shows good reproducibility between all 3 replicates. Correlations between different methods are decreased**.** Red numbers represent pair-wise Pearson's correlation coefficients.



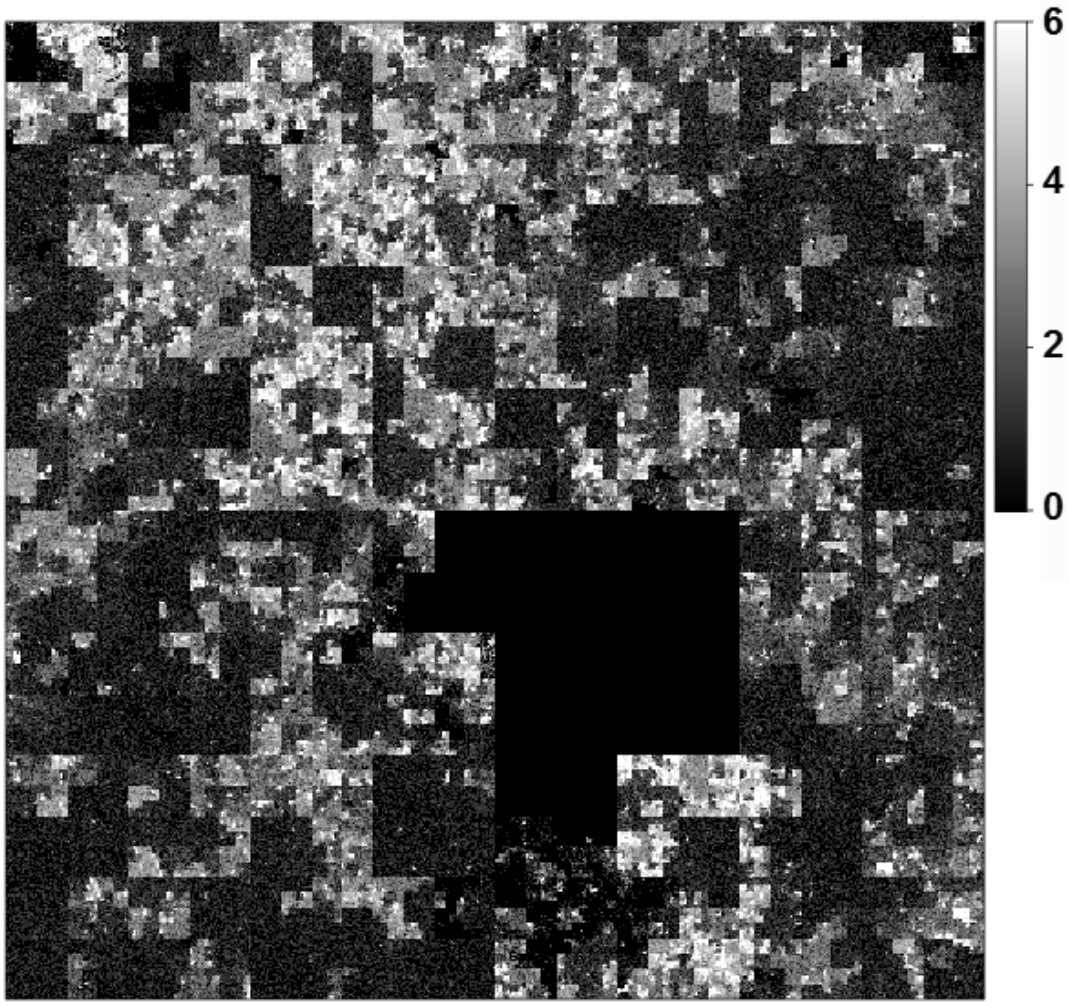
Supplementary figure 3: Examples of Pol II and RNA-seq coverage over differentially expressed genes under two experimental conditions. Vertical axis represents sequencing tags coverage per position A)B) Upregulated genes shows increase in tag density for both Pol II and RNA-seq. C) Pol II tag density over gene body is increased less extensively after doxycycline treatment compared to RNA levels of this gene. Up-regulation of gene comes with increased mRNA stability. Control sample - CON and doxycyclin treated sample - DOX.



Supplementary Figure 4 Correlation of TF occupancy with Pol II ChIP-seq and RNA-seq data from ENCODE datasets. Results of linear regression analysis of Pol II ChIP-seq and RNA-seq with a) transcription factor occupancy and b) individual principal components extracted from TTAS of 21 transcription factors. All transcription factors are presented as two independent biological replicates.



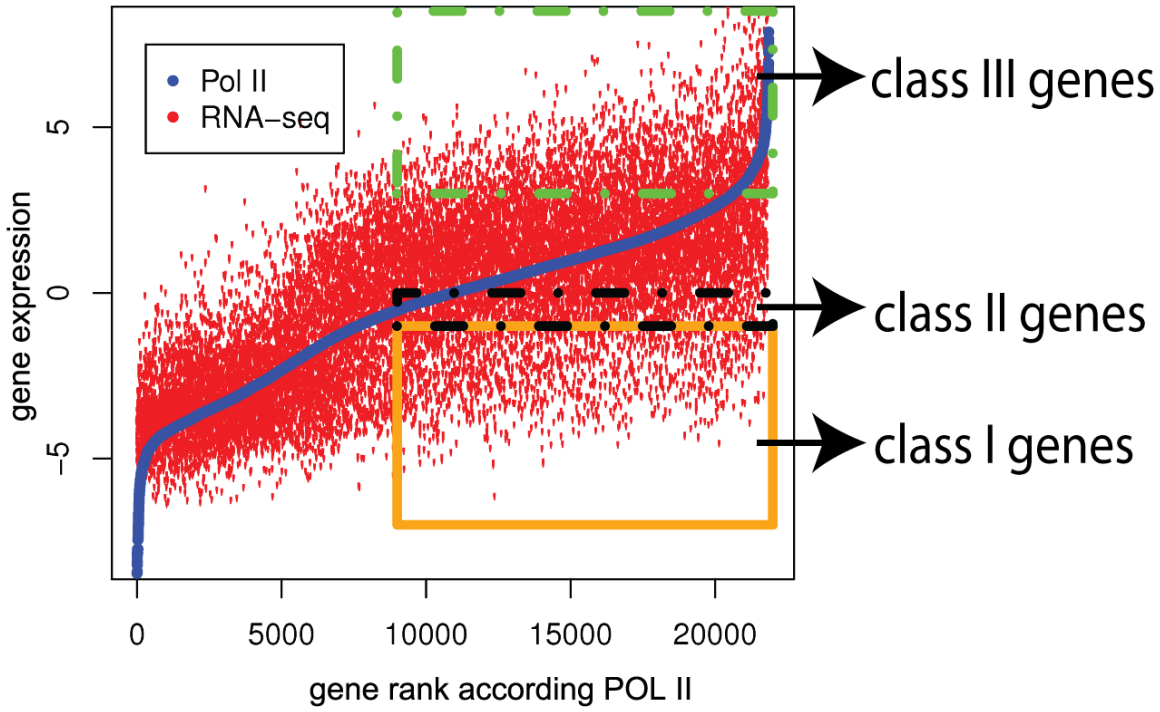
Supplementary Figure 5: Effect of poised polymerase and polymerase accumulated on 3'-UTRs on correlation of Pol II ChIP-seq with transcription factor occupancy A) Correlation of Pol II presence in gene body excluding 3'UTR and first 300 bp ("BODY"), Pol II presence in whole gene transcript ("POL II") and by RNA-seq with transcription factor occupancy B) Correlation of Pol II occupancy with individual principal components extracted from TTAS of 5 transcription factors. Excluding pol II tags bound close to transcription start sites and 3'-UTRs has only a minor influence on the correlation of Pol II ChIP-seq data with TF occupancy.



Supplementary Figure 6 Pol II occupancy on chromosome 1 reveals that major part of chromosome 1 is occupied by Pol II. Pol II occupation is restricted to distinguishable domains. For better visualization of domains, chromosome 1 coordinates are represented as continuous space filling curve (Hilbert curve) (Anders, 2009).Color log2 scale represents quantile normalized number of Pol II tags mapped to 500bp bins. Large area without coverage of POL II sequencing tags is represents centromere.

Reference:

Anders S (2009) Visualization of genomic data with the Hilbert curve. *Bioinformatics* **25:** 1231-1235



Supplementary Figure 7 Rank analysis of Pol II occupancy with empirically defined gene classes (I, II, III) used for gene ontology term enrichment analysis. Genes are ranked according to Pol II ChIP-seq results. Genes in all classes are expressed as defined by Pol II ChIP-seq and seem to be non-expressed (class I), lowly expressed (class II) and very highly expressed (class III) by RNA-seq. All expression values represent median centered and log2 transformed NR100KM (Pol II) and NR10KM (RNA-seq).

Supplementary Table

Supplementary Table, Sequencing and mapping statistics

	Sample	Raw reads	Mapped reads	Mapped reads (q>0)	Percentage mapped	Percentage with q>0
RNA sequencing	CON1	43,519,947	19,409,978	11,469,611	44.6	59.1
	CON2	32,962,479	15,120,183	8,890,432	45.9	58.8
	CON3	46,869,848	20,349,215	11,451,852	43.4	56.3
	DOX1	31,602,468	13,622,531	7,625,974	43.1	56.0
	DOX2	37,447,267	15,428,604	9,565,192	41.2	62.0
	DOX3	42,341,072	15,213,191	8,315,499	35.9	54.7
Pol II ChIP sequencing	CON1	65,455,910	41,065,573	36,186,881	62.7	88.1
	CON2	63,256,626	41,829,937	37,054,927	66.1	88.6
	CON3	52,403,736	33,479,940	29,424,271	63.9	87.9
	DOX1	69,649,772	35,978,193	31,798,338	51.7	88.4
	DOX2	57,203,495	37,304,546	33,257,995	65.2	89.2
	DOX3	60,205,429	34,741,976	30,692,140	57.7	88.3