Ch 3 - L 5.2

Chromatin at Promoters

CpG islands and Cytosine Methylation at Promoters

Broad type promoters are usually CG-rich and lack a clear TATA-box

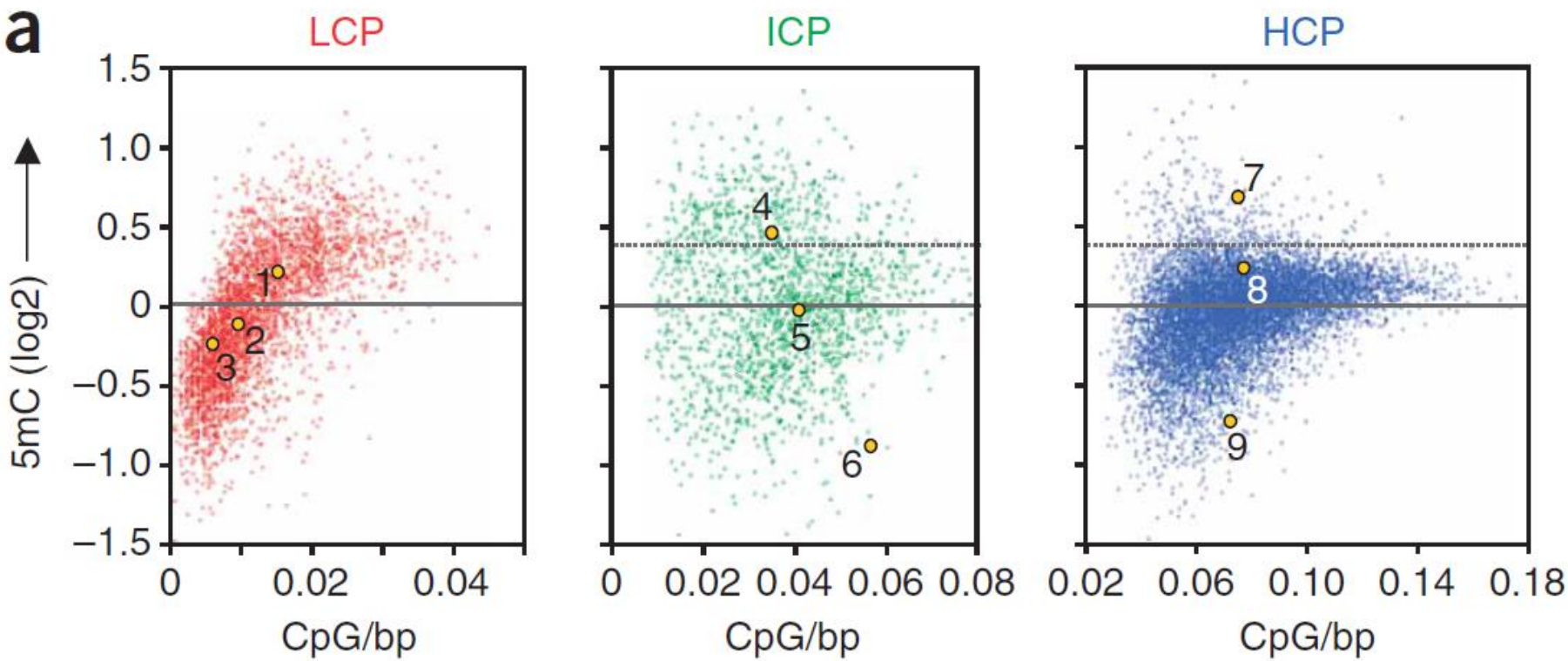
I.e. they contains **CpG-islands**

How is CpG methylation linked to function of these promoters ?

Take again Weber et al. 2007 results.
MeDIP – promoter microarray
Promoters divided in three groups based on % CpG (HCP, ICP and LCP)

Methylation enrichment as a function of CpG % (density):

In LCP, there is a almost linear correlation, i.e. more CpG you have, more meC you get.
In HCP, it seems like a «saturation curve» i.e. increasing CpG density does not change 5mC enrichment (most are unmethylated independently on the CpG density).
In ICP we see rea lack of correlation, indicating that some promoters are hypermethylated, other hypo- , suggesting regulation.



Weber et al, (2007) Nat Genet 39:457.

5mC-data correlated with RNA Pol II occupancy and H3K4me3, by ChIP on the same microarrays.

1) High CpG promoters are mostly hypo-methylated and H3K4me3 +.

2) Even in cases when no PolII ChIP signal, still H3K4me3

3 Low CpG (LCP) promoters unlinked to the degree of methylation

4) Intermediate CpG promoters (ICP): methylation level reflects inversely RNA Pol II occupancy

Nucleosome positioning at promoters

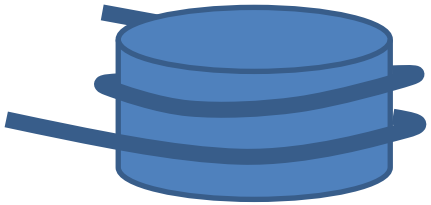
MNase-seq: chromatin is digested using Micrococcal Nuclease (mixed exo- endo-), then the remaining «nucleosomal» DNA is NGS sequenced.

While in most of the genome nucleosomes are random positioned) in functional positions like Enhancers and **Promoters** we see a very strong positioning.

By ChIP-seq these nucleosomes feature histone variants (H3.3 and H2A.z) and the typical PTMs (H3K27ac/H3K42-3me, other PTMs).

MNase-Seq workflow

- Treat Nuclei with MNaseI
- Nucleosome-protected DNA is isolated by deproteinization
- DNA fragements ~150 bp purified
- Ligated to adaptors → Library
- Sequenced by NGS technology



MNase trimmed nucleosome

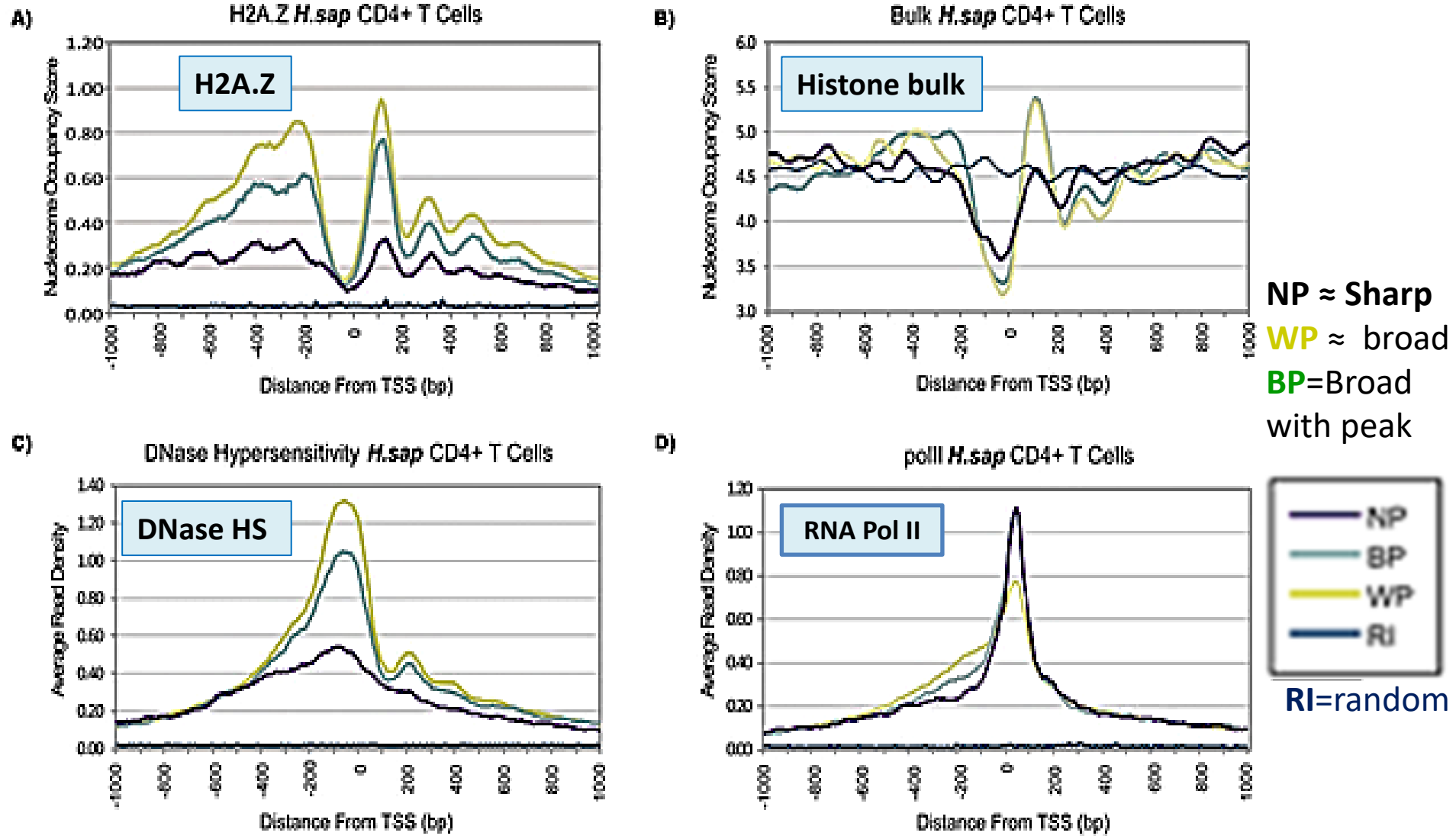
deproteinize

~150 bp

Adapter, library prep. → NGS

After mapping to reference genome, highest peaks signal most «positioned» nucleosomes

Figure 1. Promoter Classes Reflect Distinct Profiles of Nucleosome Organization. Rach 2011 Plos Genet e1001274



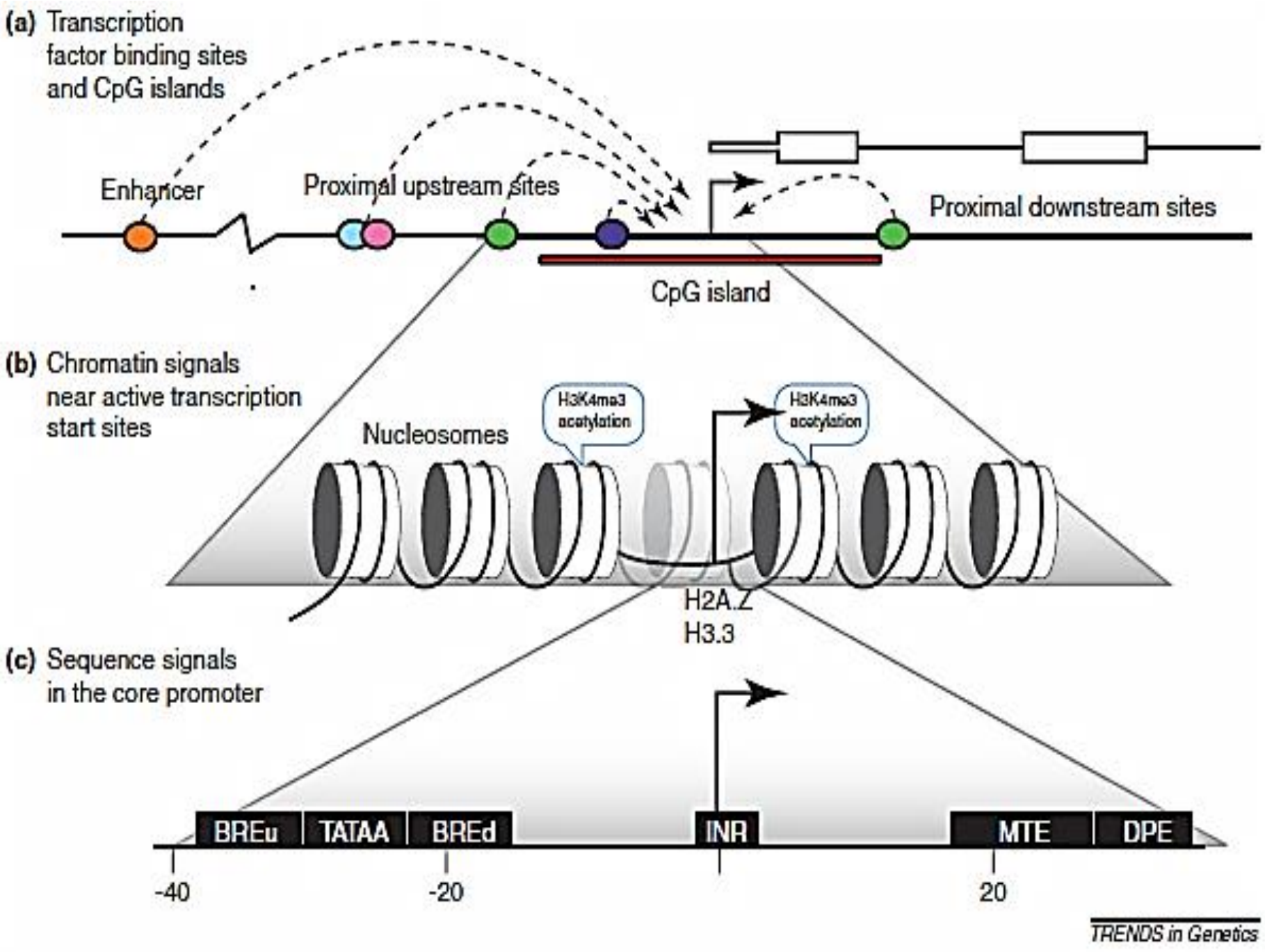
NP ≈ Sharp

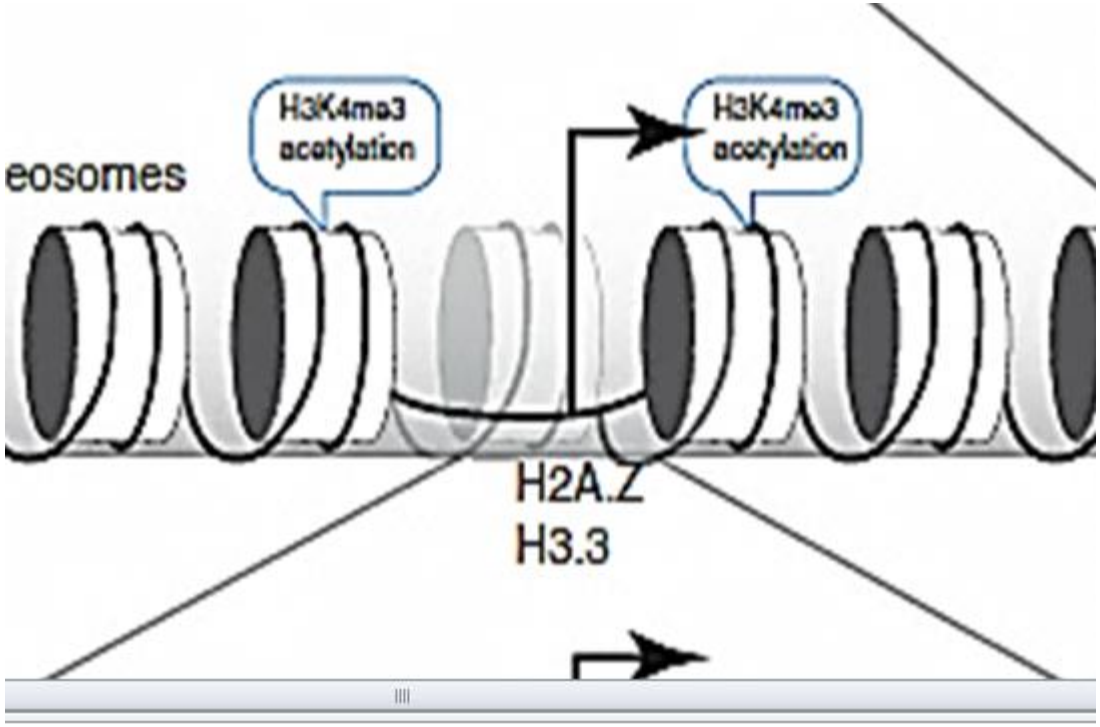
WP ≈ broad

BP=Broad with peak

RI=random

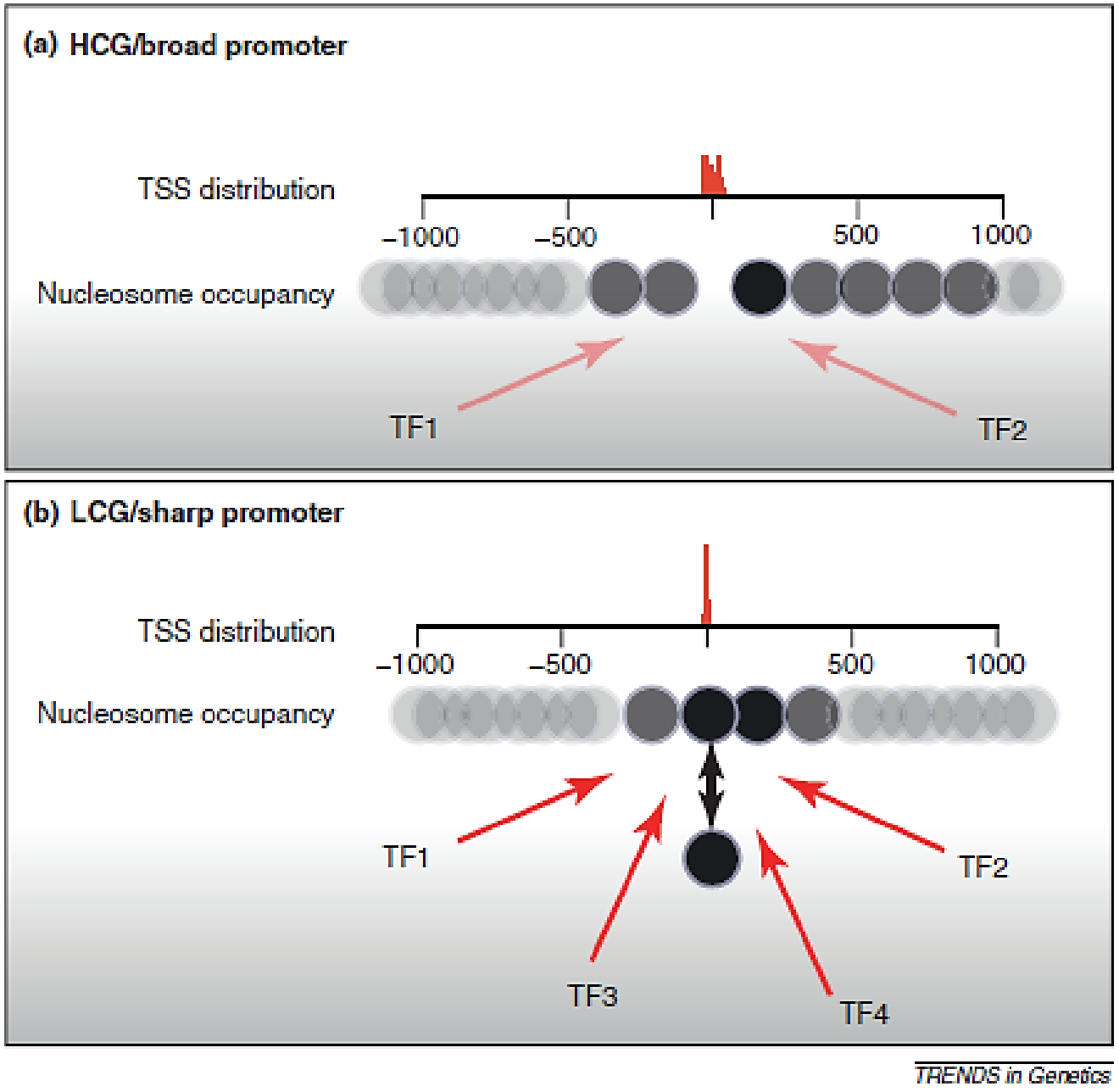
Authors classified Promoters as Narrow Peak (NP), Broad with Peak (BP), and Weak Peak (WP). RI refers to average levels at random intergenic sites, which is used as a baseline. (A) Increased H2A.Z levels ($p<10E-36$), (B) increased bulk levels. DNase hypersensitive sites revealed a more accessible nucleosome-free region at BP and WP but not at NP promoters (C), yet pol II levels were higher at NP promoters (D).


(a) Transcription factor binding sites and CpG islands
Enhancer
Proximal upstream sites
Proximal downstream sites
CpG island
(b) Chromatin signals near active transcription start sites
Nucleosomes
H3K4me3 acetylation
H3K4me3 acetylation
H2A.Z
H3.3
(c) Sequence signals in the core promoter
BREu
TATAA
BREd
INR
MTE
DPE
-40
-20
20
TRENDS in Genetics


eosomes
H3K4me3
acetylation
H3K4me3
acetylation
H2A.Z
H3.3

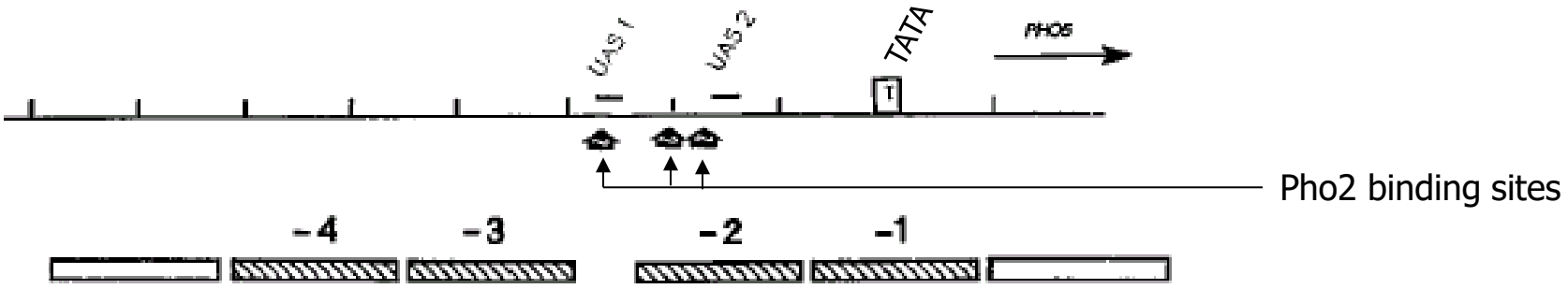
Mostly constitutive

Mostly regulated

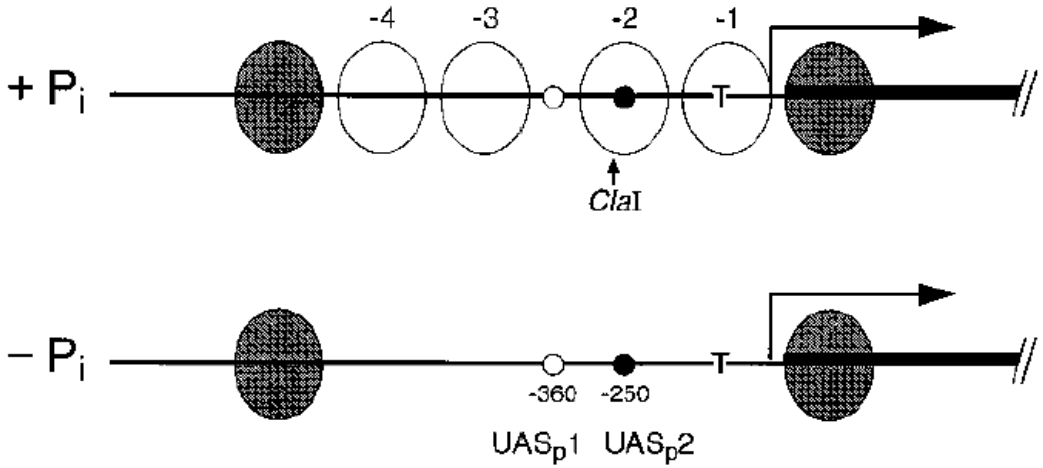


From Valen & Sandelin 20011, Trends in Genetics 27:475

Eviction of promoter nucleosome by activating TFs: the PHO5 gene in the Yeast



Pho4 is the P-sensitive inducer, whereas Pho2 is constitutive



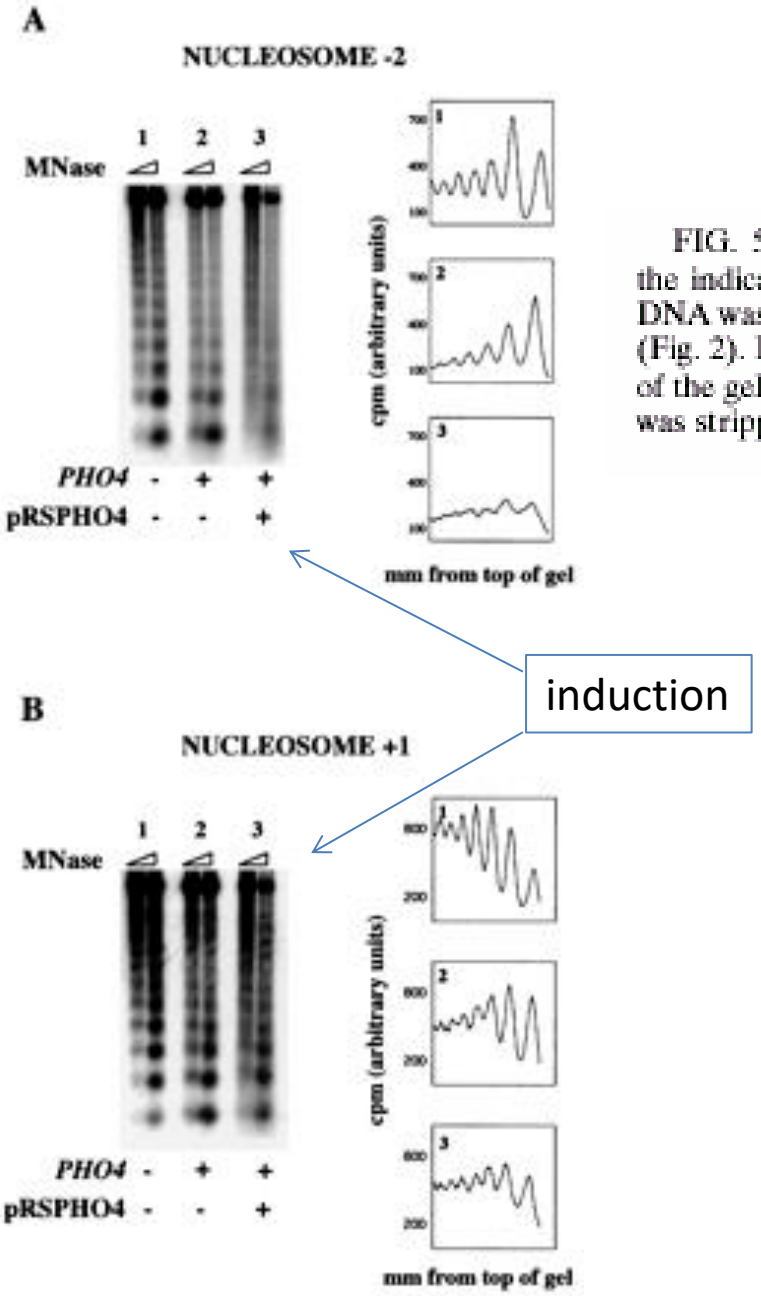


FIG. 5. In vivo chromatin remodeling of episomal *PHO5*. Spheroplasts from the indicated strains were treated with micrococcal nuclease (MNase), and the DNA was purified and Southern blotted. (A) The blot was probed with probe A (Fig. 2). Data from each sample were quantified, and the distance from the top of the gel was graphed against the signal density. (B) The blot shown in panel A was stripped and reprobed with probe B (Fig. 2).

HCG promoters

- «Broad» TSS – several TSS spread over 30-100bp
- CpG-island, under-methylated
- Functions: housekeepers, i.e. constant and ubiquitous expression
- Stable expression levels
- Bound mostly by ubiquitous Transcription Factors (TFBS over-represented close to TSS, e.g. Sp1)
- Nucleosome Free (or Depleted) Region (NDR) evident
- First nucleosome downstream TSS strongly positioned.
- Nucleosomes flanking NFR enriched for H2A.Z + H3.3
- Most expression-predictive PTMs: H3K27ac and H4K20me1
- CGIs are «intrinsically» promoters (CpG recognized by Cfp1/Set1 H3K4 methyltransferase)

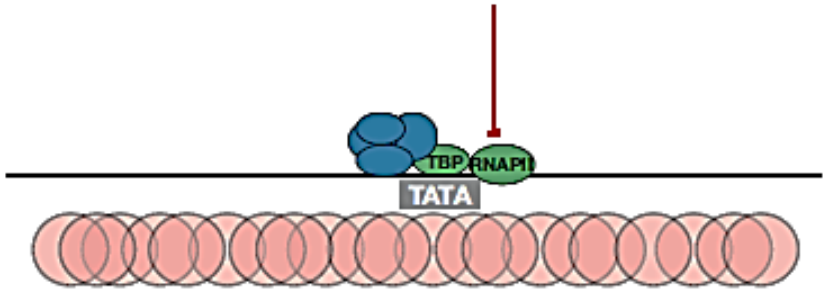
LCG promoters

- «Sharp» TSS – One single TSS at 1-3 adjacent nucleotides
- Clear TATA sequence present (30%)
- Function: Tissue-specific, inducible
- Large variability in expression level
- Bound mainly by tissue-specific and inducible Transcription Factors
- Nucleosome-occupied. NFR less evident
- Nucleosome positioning and PTMs average less evident
- Most expression-predictive marks: H3K4me3 and acetylation
- «Intrinsically» repressed, require TFs and chromatin remodelers to be freed and activated

Promoter types

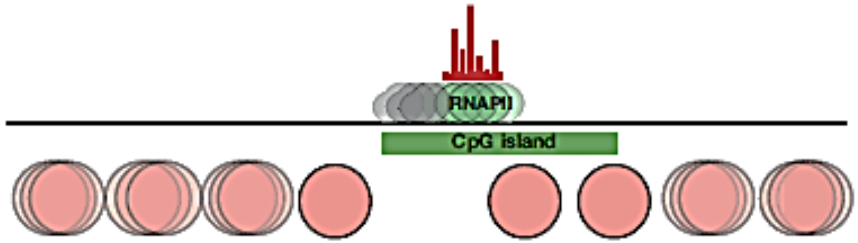
Adult tissue-specific

1) sharp transcription initiation pattern
2) context-specific regulatory input close to TSS
3) TATA-box ~30 bp upstream of TSS
4) disordered nucleosomes



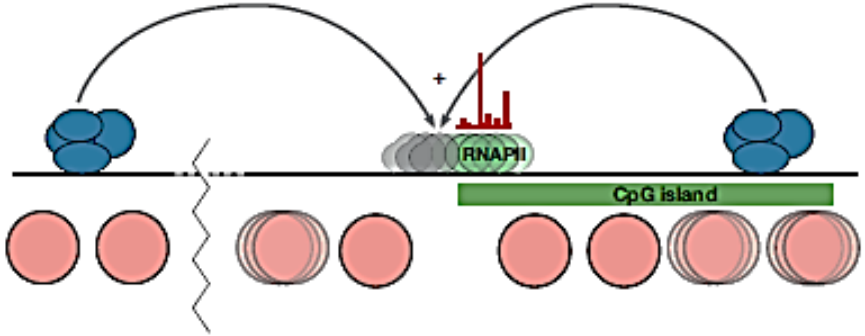
Ubiquitously expressed ("housekeeping")

1) broad transcription initiation pattern
2) no context-specific regulatory input
3) CpG island around TSS
4) nucleosome-free region with precisely positioned -1 and +1 nucleosomes



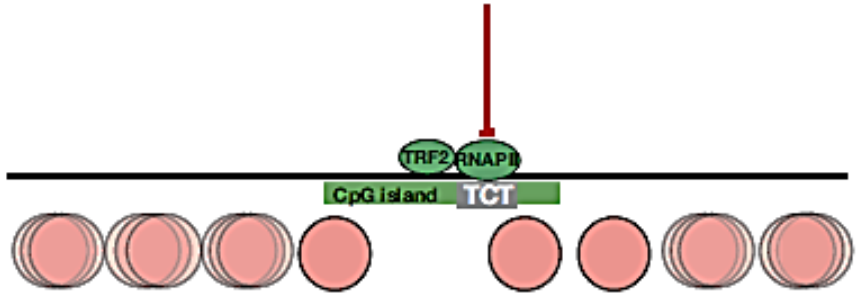
Developmentally regulated

1) broad transcription initiation pattern
2) context-specific regulatory input from distal enhancers
3) long CpG island(s) into gene body
4) nucleosome-free region with precisely positioned -1 and +1 nucleosomes



Translational machinery-specific

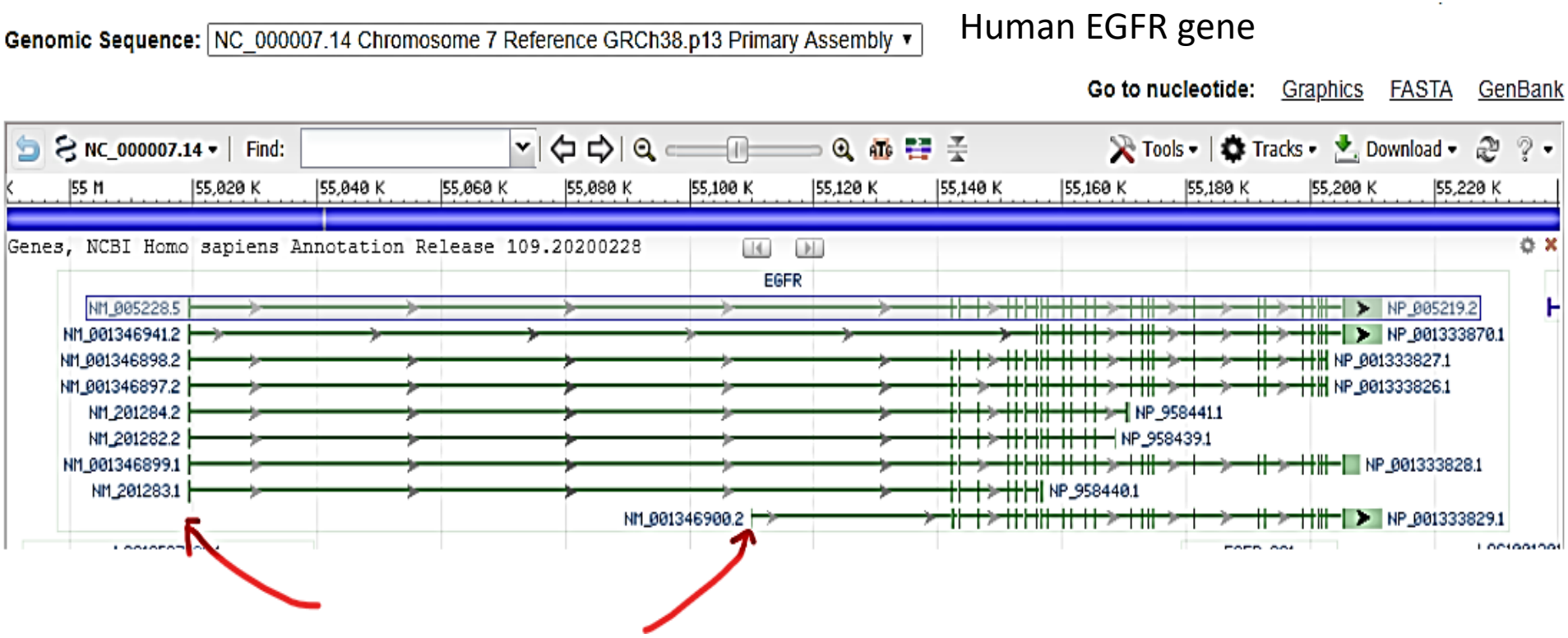
1) sharp transcription initiation pattern
2) no context-specific regulatory input
3) TCT initiator and CpG island
4) nucleosome-free region with precisely positioned -1 and +1 nucleosomes



Alternative promoters and overlapping codes

A number of genes present alternative TSS that depend on alternative promoters, which are used in a differential way.

Human EGFR gene



Broad-type Promoters may have condensed and overlapped promoter codes (a group of promoter elements superimposed to a second group of elements).

Important example comes from oocyte: housekeeper genes here are trancribed using a different machinery and using a promoter «code» that is different form those used by somatic cells at the same genes.
Often the two codes overlap.

Textbook «G» by Levine et al. 2014:

«some unusual features concerning promoters and basal factors».

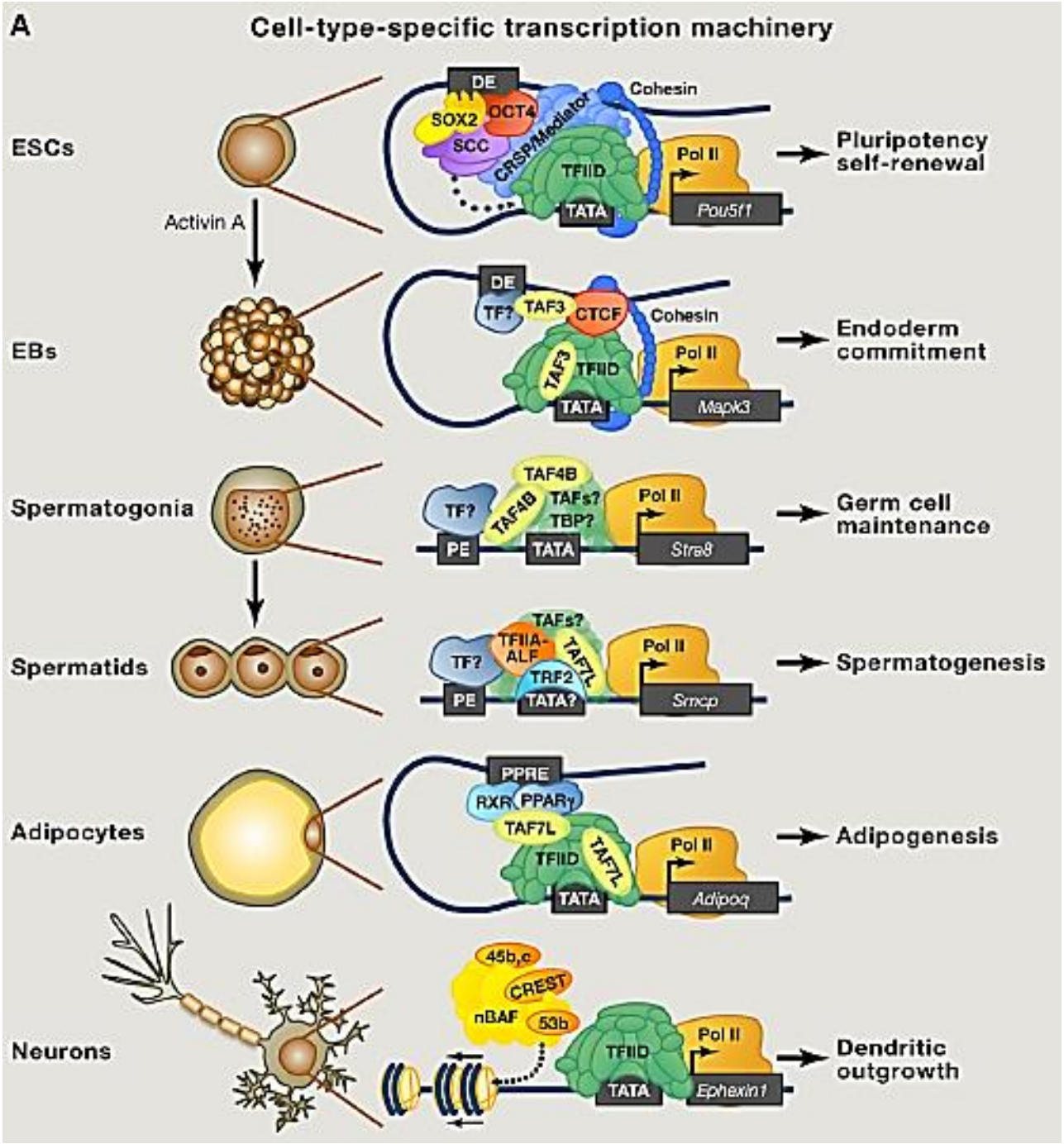
Different basal transcriptional machineries exist in different tissues, which transcribe specific and common genes.

This include a set of Promoters that in spermatids require the TBP-related TRF2 factor instead of the ubiquitous TBP.

We see essentially two (rare) events:

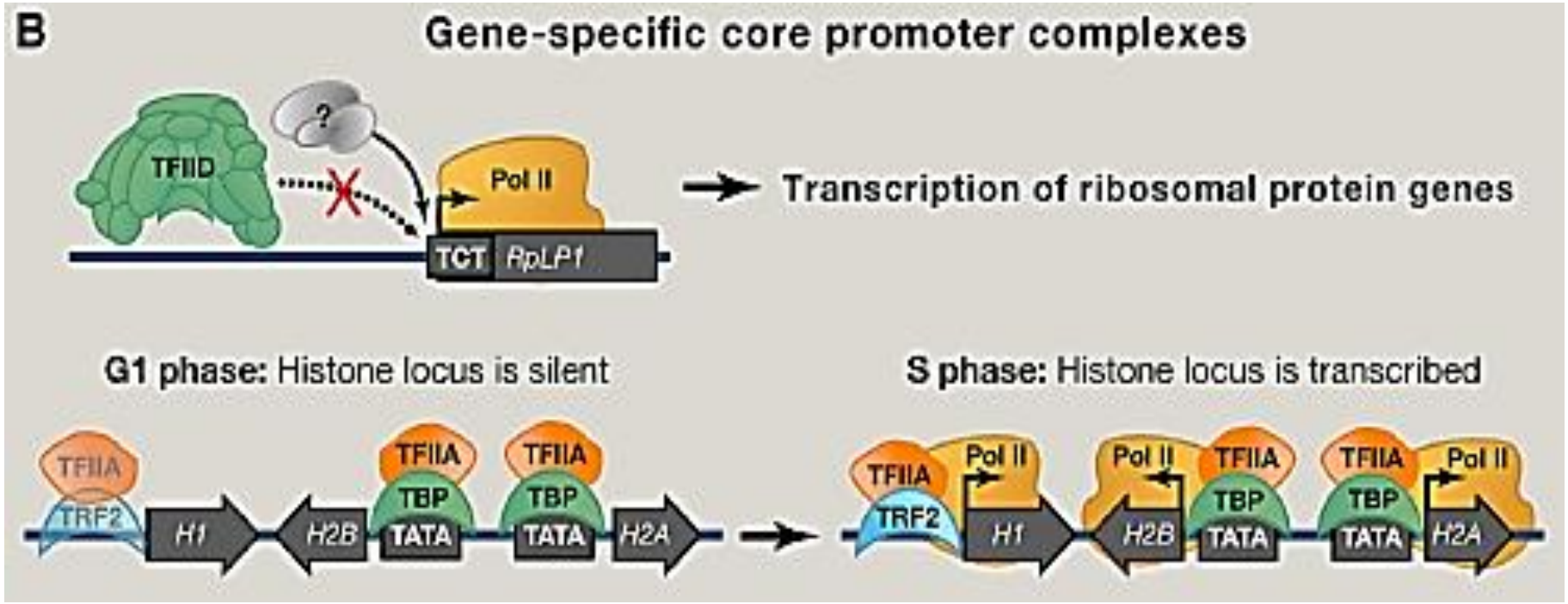
- Tissue-specific transcriptional machineries that transcribe cerain genes

- Gene-specific transcriptional machineries



Levine et al. 2014

Finally, note that a number of special Promoters have aso been described, were the core promoter complex is different from the classical one, as in the case of Histone gene promoters.



Levine et al. 2014