CONTROLE DA EXPRESSÃO GÊNICA EM EUCARIOTOS

Maria Carolina Quecine
Departamento de Genética
mquecine@usp.br

SUMÁRIO

- Genoma x Transcriptoma em eucariotos;
- Controle da transcrição;
- A diversas RNA polimerases;
- Sequências regulatórias da transcrição;
- Fatores de transcrição;
- Enhancers, silenciadores e insulators;
- Próxima aula

Sequencing the Genome of Sitting Bull and Other Famous People

August 22nd, 2010 in Ancient DNA, DNA in the News, DNA Studies, Famous DNA |

ScienceNews reports that researchers led by Eske Willerslev at the University of Copenhagen are attempting to sequence the genome of legendary Native American "Sitting Bull" (see "Genome of a Chief").

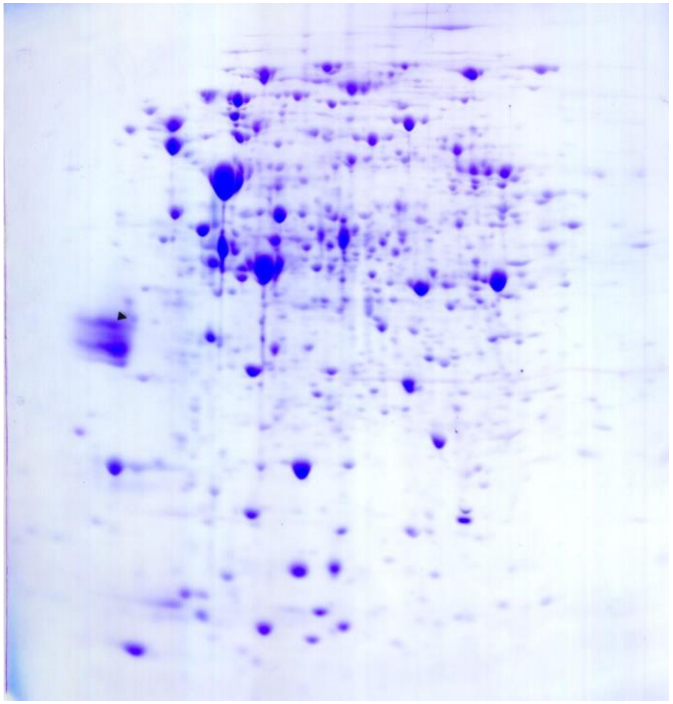
Earlier this year (2010), Eske Willersleve announced the successful sequencing of approximately 80% of the genome of "Inuk," a man from Greenland who left behind a few small fragments of bone and four hairs frozen in permafrost when he died about 4,000 years ago (see "Long-Locked Genome of Ancient Man Sequenced"). Using these ancient DNA sequencing techniques, Willersleve's group is analyzing DNA from other samples.

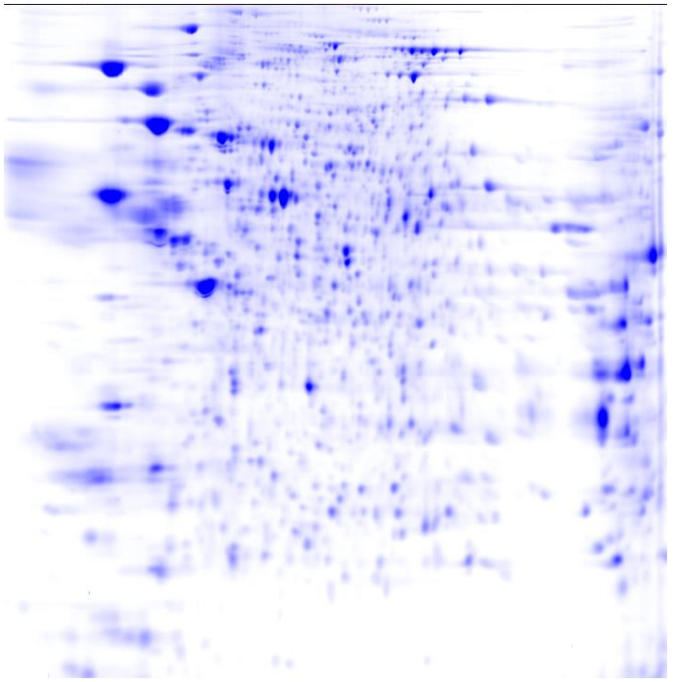
One of these samples is a lock of hair from Sitting Bull.

Sitting Bull (c. 1831 – Dec. 15, 1890) was a Hunkpapa Lokota Sioux born in South Dakota. Sitting Bull played an important role in the June 25, 1876 Battle of the Little Bighorn, and later toured as a performer in Buffalo Bill's Wild West show.

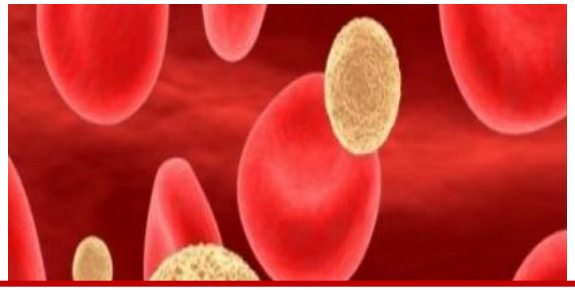
O QUANTO ISSO PODE ME TRAZER DE INFORMAÇÃO?

PERFIS DE PROTEÍNAS DE DOIS ORGANISMOS PRÓXIMOS?



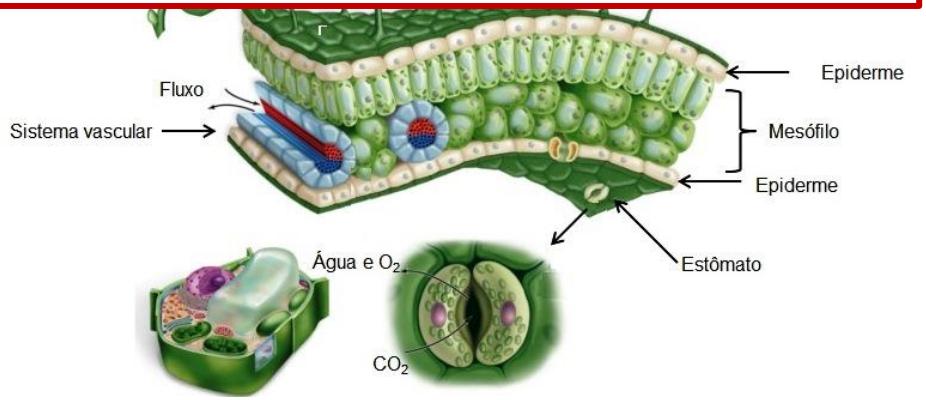


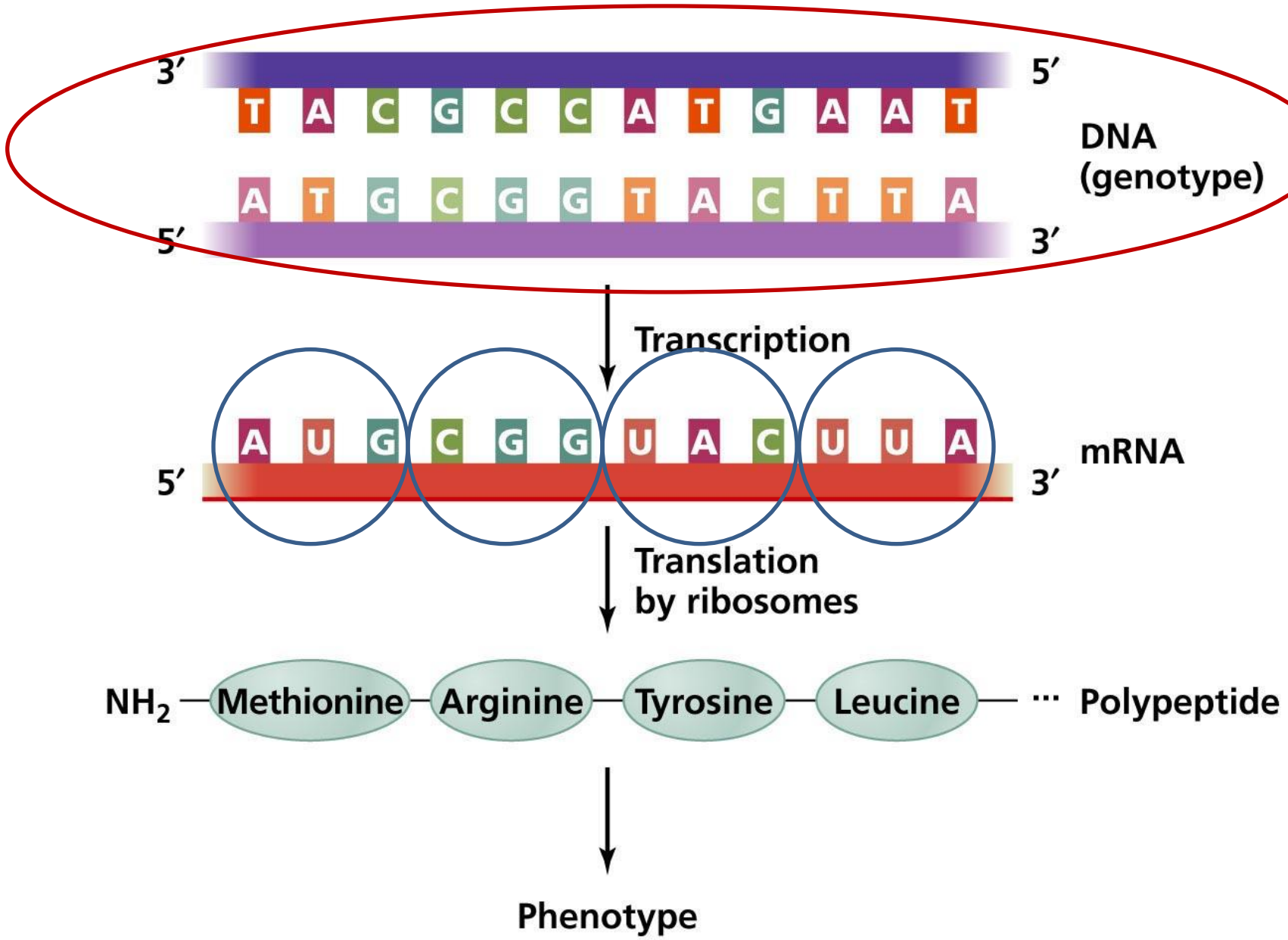
O GENOMA É O MESMO...O QUE MUDA SÃO OS GENES EXPRESSOS EM CADA TECIDO.



Muito mais complexo o controle da expressão gênica em eucariotos:

- compartimentalização do processo;
- diferenciação celular nos organismos pluricelulares.





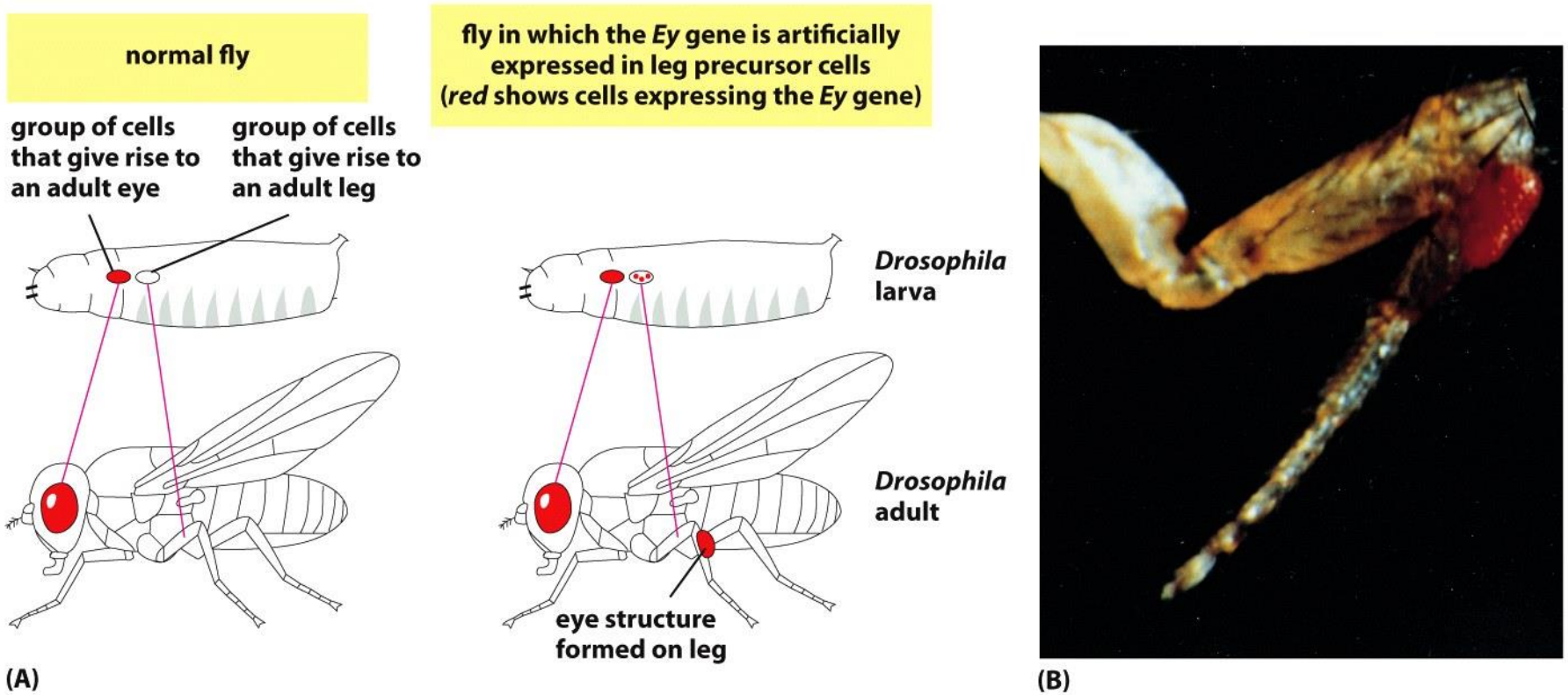
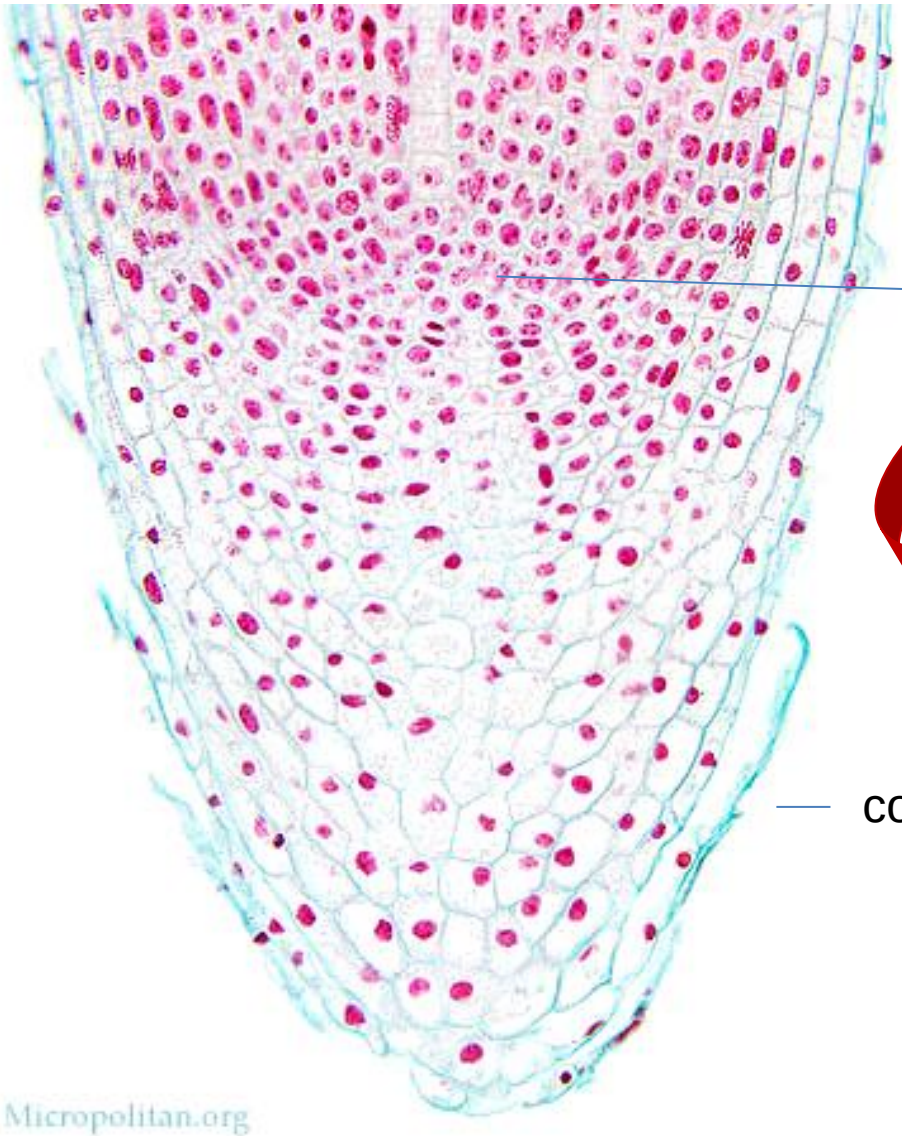
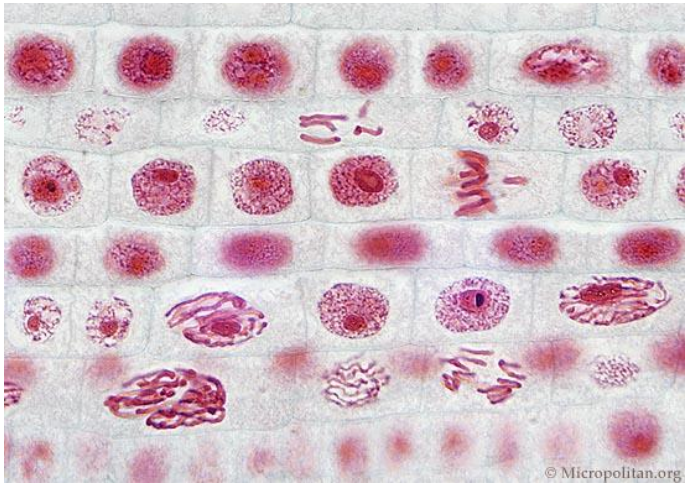


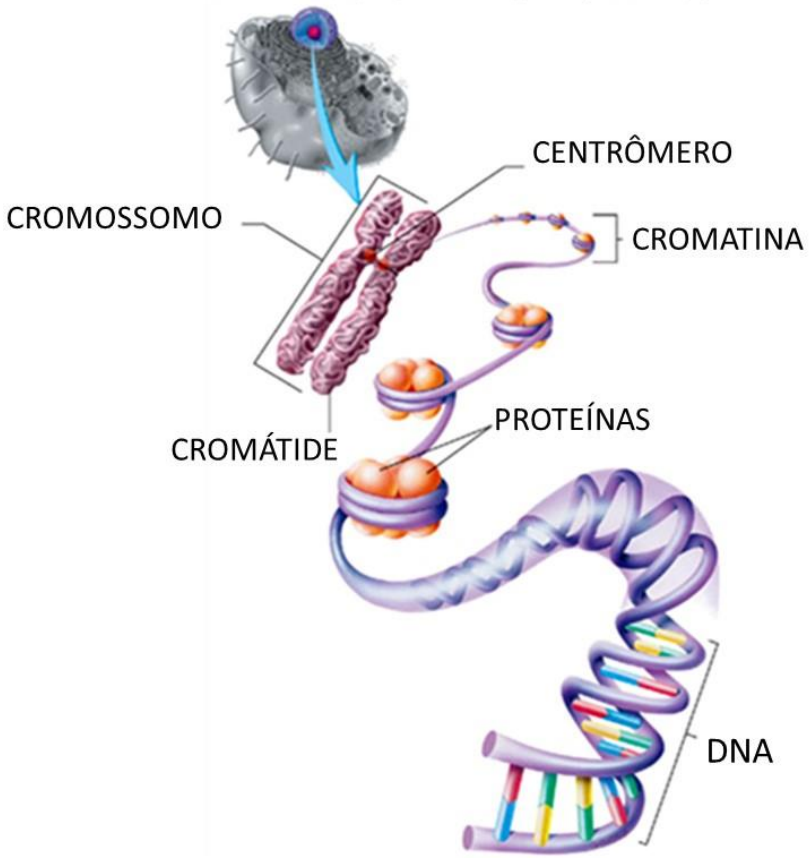
Figure 8-23 Essential Cell Biology 3/e (© Garland Science 2010)

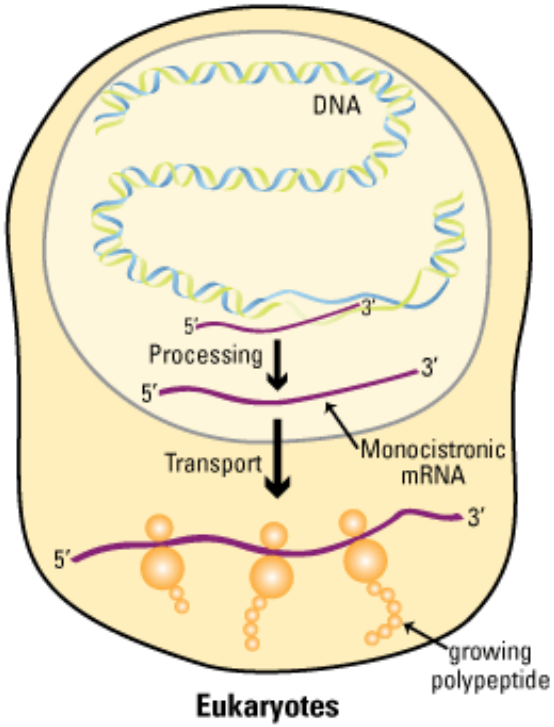


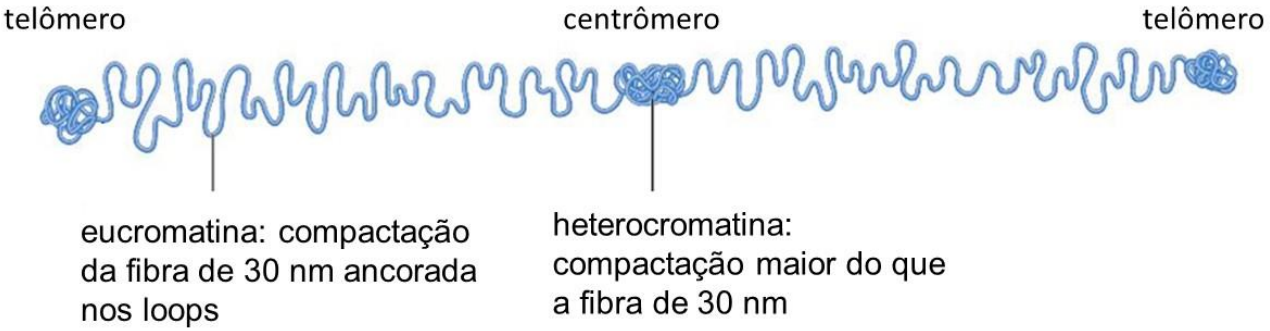


Divisão celular

EM EUCARIOTOS …







ETAPAS NO CONTROLE DA EXPRESSÃO

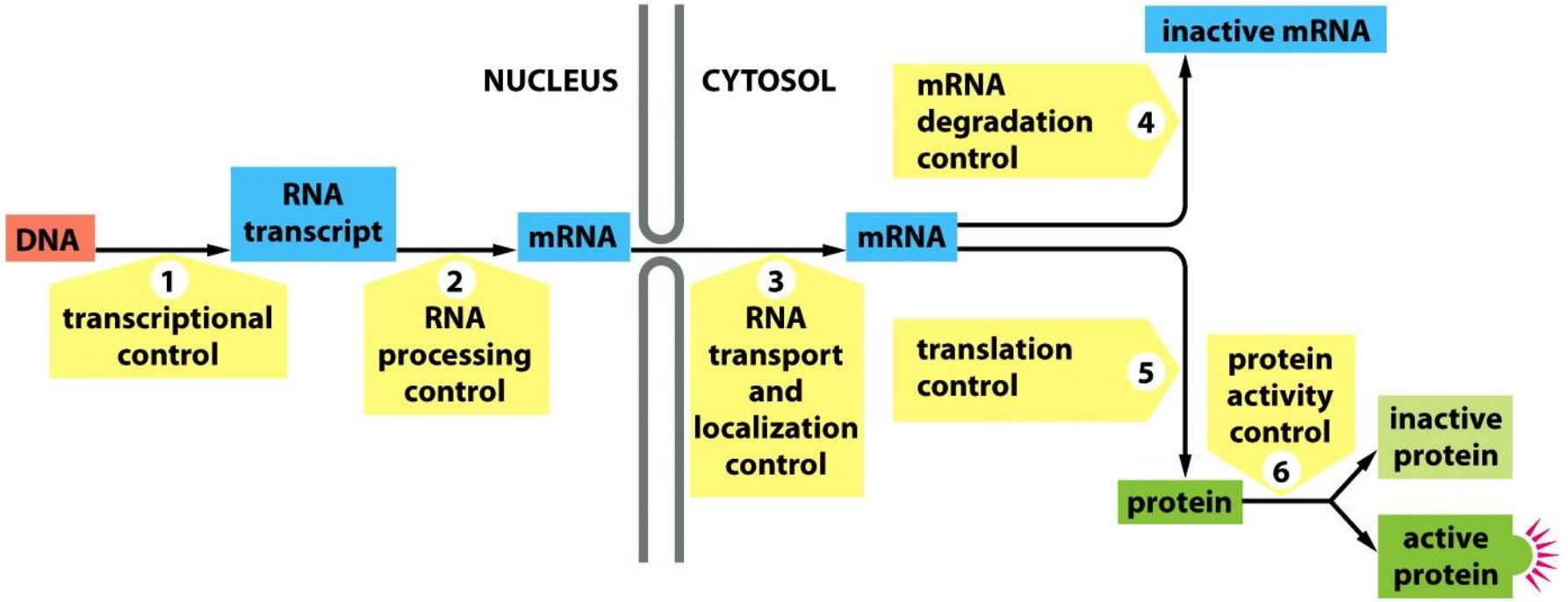


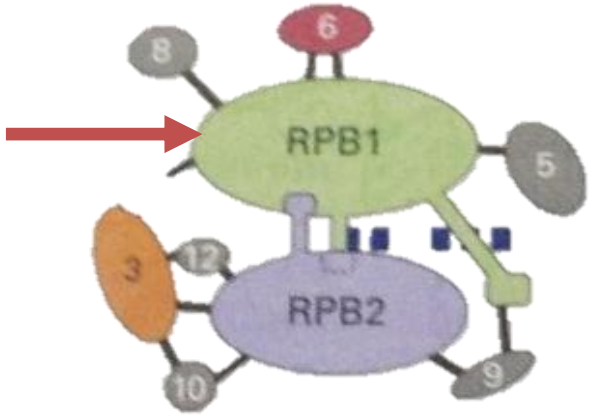
Figure 8-3 Essential Cell Biology 3/e (© Garland Science 2010)

Etapa - controle da transcrição é majoritária!

VÁRIAS RNA POLIMERASES!

Polimerases	RNA transcrito
RNA polimerase I	Pré-rRNA
RNA polimerase II	mRNA, snRNA, siRNA, miRNA
RNA polymerase III	tRNA, 5S-rRNAsn RNA u6, outros pequenos RNA estáveis

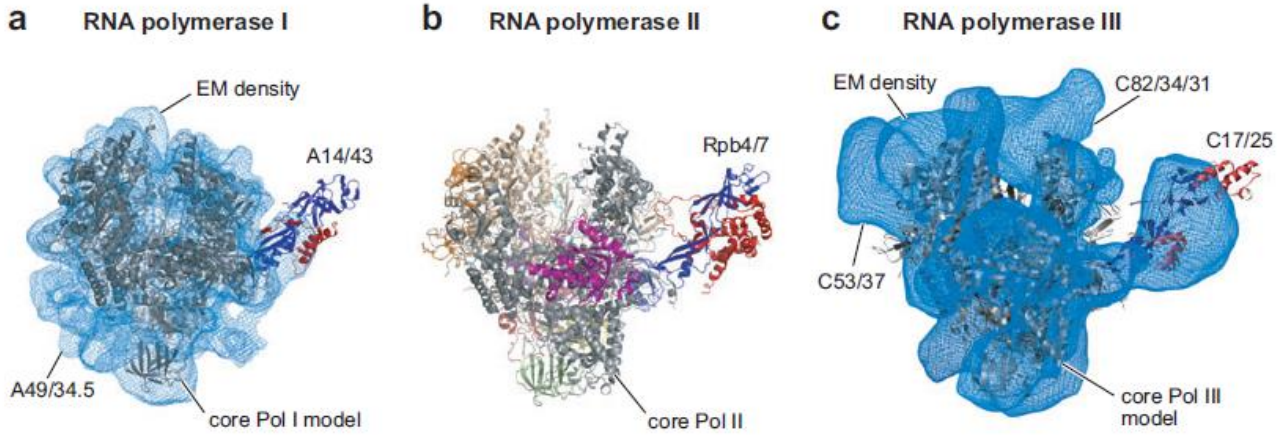
RNAP II – Domínio CTD – Tyr-Ser- Pro-Thr-Ser - Pro -Ser
Vária em número de cópias – em levedura pelo menos 10 para sobreviver.
Ser – fosforilada – se desliga do promotor



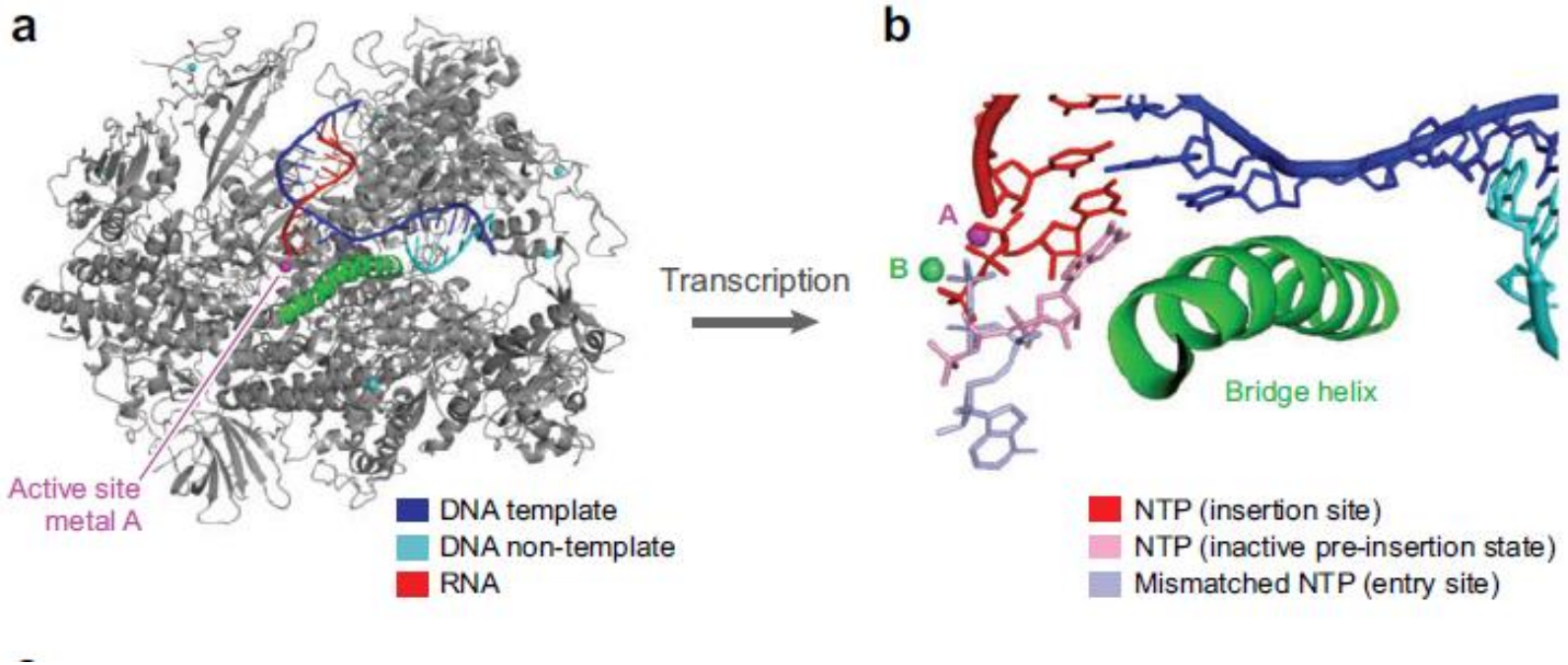
Structure of Eukaryotic RNA Polymerases

P. Cramer, K.-J. Armache, S. Baumli, S. Benkert, F. Brueckner, C. Buchen, G.E. Damsma, S. Dengl, S.R. Geiger, A.J. Jasiak, A. Jawhari, S. Jennebach, T. Kamenski, H. Kettenberger, C.-D. Kuhn, E. Lehmann, K. Leike, J.F. Sydow, and A. Vannini

Gene Center Munich and Center for Integrated Protein Science CIPSM, Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany; email: cramer@lmb.uni-muenchen.de



Essas enzimas são co-reguladas por proteínas acessório especificas relacionadas à transferência de sinais por ativadores e/ou repressores transcricionais.



Mesmo principio dos procariotos no complexo fechado durante a elongação, estabilizando a RNAPII, onde de 8 a 9 pb do híbrido DNA-RNA, bem como o domínio grampo fechando-se sobre o DNA à jusante!

RNA Pol IV and V in Gene Silencing: Rebel Polymerases Evolving Away From Pol II's Rules

Ming Zhou[1] and **Julie A. Law**[1,*]

[1]Plant Biology Laboratory, Salk Institute for Biological Studies, La Jolla, CA, 92037, USA

RNA polymerase IV and transcriptional silencing

Hervé Vaucheret

RNA-based silencing mechanisms suppress gene expression through the sequence-specific activity of small RNAs. New studies in plants now identify atypical RNA polymerases that promote and maintain transcriptional silencing.

nature genetics

RNA polymerase V transcription guides ARGONAUTE4 to chromatin

Andrzej T Wierzbicki, Thomas S Ream, Jeremy R Haag & Craig S Pikaard

COMO ESTUDAR A TRANSCRIÇÃO?

Box 1 | **Techniques used to detect sites of transcription and RNAPII**

BrUTP incorporation

Cells are either injected with, or lightly permeabilized and incubated with, a pulse of 5-bromouridine 5′-triphosphate (BrUTP). The pulse time, typically 2 to 15 minutes, is optimized to be the minimum time required to produce a detectable signal. This ensures that sites of RNA synthesis are labelled rather than distant processing sites. Suboptimal triphosphate concentrations can also be used to limit elongation so that tagged transcripts remain at synthetic sites. Transcription sites are then detected by immunofluorescence with antibodies that recognize bromouridine.

Immunofluorescence

Immunofluorescence with antibodies against elongating forms of RNA polymerase II (RNAPII) can be used to analyse sites of transcription (see BOX 2 for further discussion of antibodies).

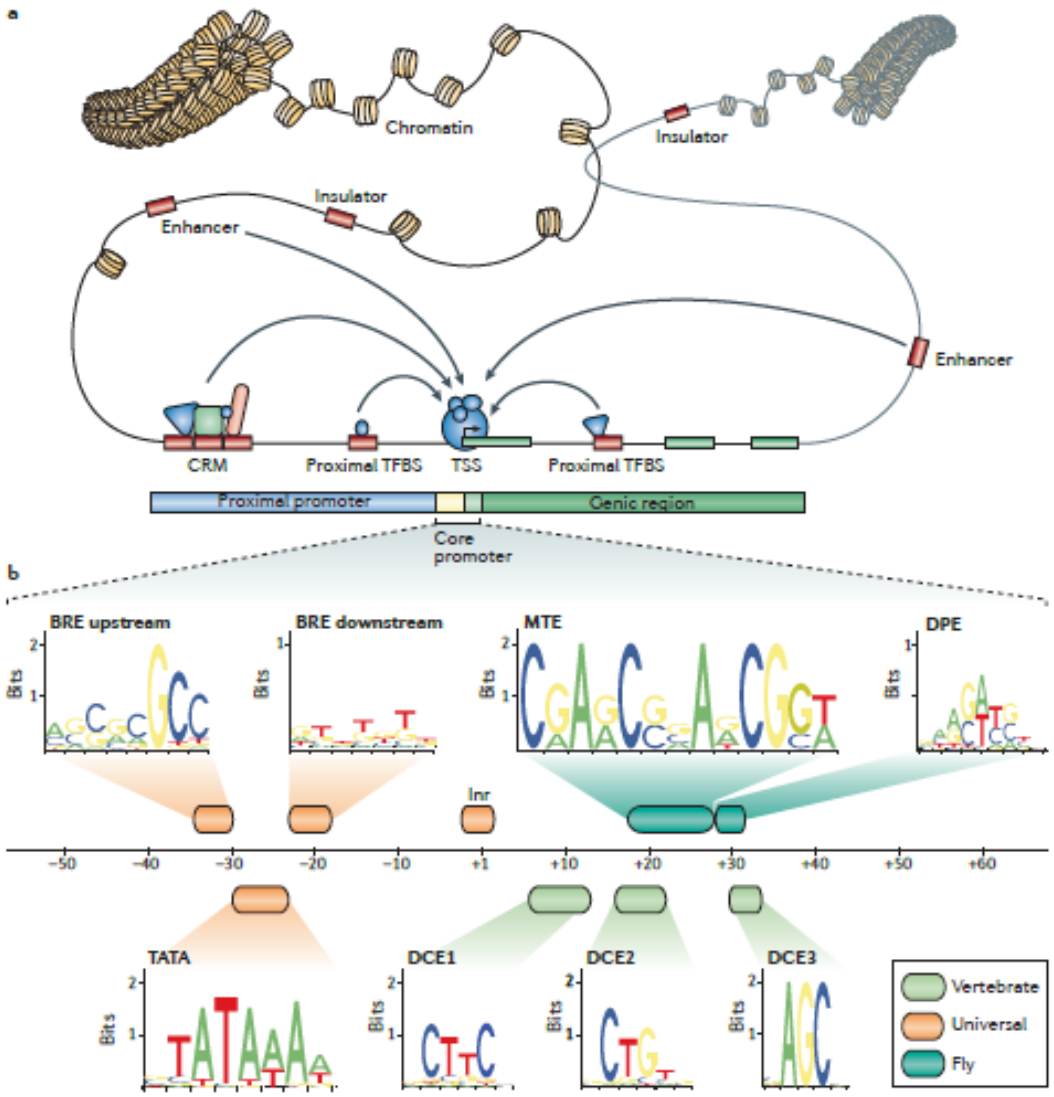
RNA FISH

RNA FISH can pinpoint sites of gene transcription because, generally, the highest concentration of specific RNA species in a cell is at the site of transcription. The absence of a denaturation step ensures that ssRNA is detected and not, for example, dsDNA. If intron probes are used, then production of the primary transcript, prior to splicing and intron degradation, can be visualized. RNA FISH can be combined with immunofluorescence (immuno-FISH) (see FIG. 3a for an example).

GFP-labelled RNAPII

Tagging subunits of RNAPII with GFP (or its variants) allows the dynamics of RNAPII recruitment to be studied when transcription of a gene is induced. Owing to sensitivity issues, most studies have been done with arrays of transgenes or on polytene nuclei. Fluorescence recovery after photobleaching (FRAP) of fluorescently labelled RNAPII can give information about the mobility of RNAPII molecules at various sites in the nucleus, including sites of transcription, and during the transcription cycle.

SEQUÊNCIAS REGULATÓRIAS DA TRANSCRIÇÃO



doi:10.1038/nrg3163

ENCONTRANDO SEQUENCIAS REGULATÓRIAS DA TRANSCRIÇÃO?

Table 1 | **Methods for characterizing promoters and transcription initiation**

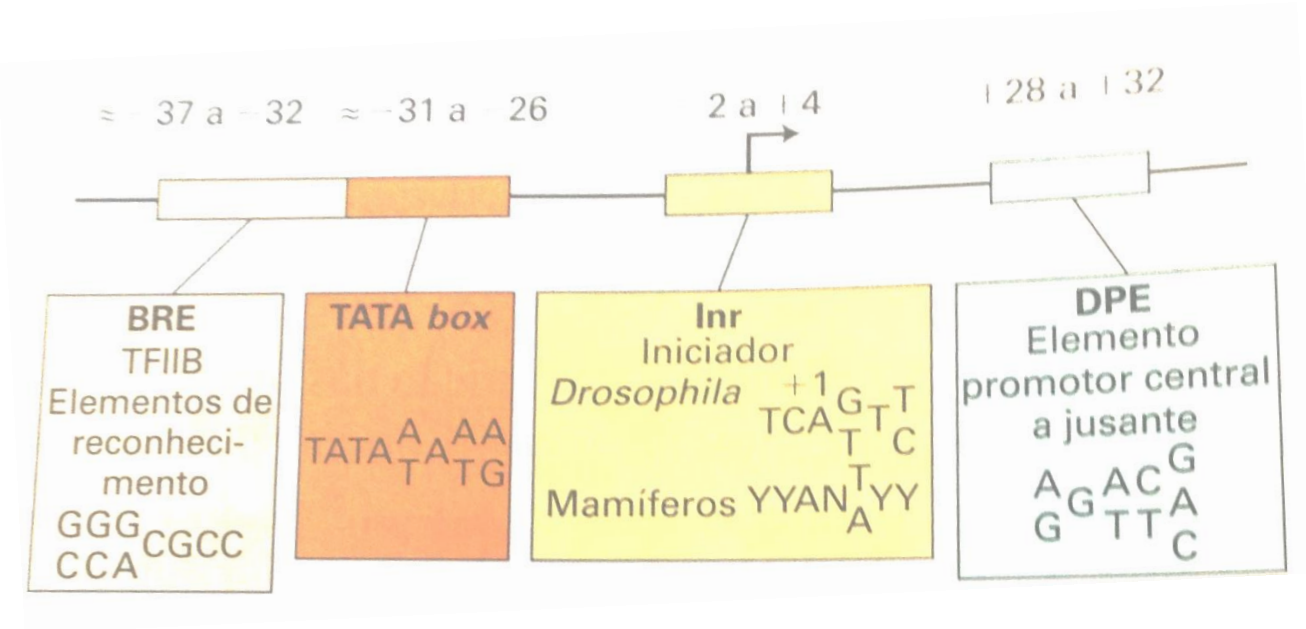
Type	Description	Advantages	Disadvantages	Examples
Methods based on sequencing RNA (or cDNA)				
5′-tag-based methods	A range of methods based on capturing capped mRNAs and isolating the first 20–30 nt of corresponding full-length cDNA, which is sequenced and mapped back to the genome. Some methods can identify 3′ ends (paired-end tags) at the same time	These technologies allow single-nucleotide resolution of TSSs (much higher resolution that ChIP-based methods) and thus enable studying the nucleotides of the initiator sites and thus the detailed architecture of the promoter	The methods do not measure promoter activity directly but rather measure the products from transcription, which may be subject to post-transcriptional regulation. Some protocols may require large amounts of RNAs[82], although methods exist to avoid this[65,83]. The methods cannot distinguish capped RNAs and potentially recapped RNA[84]	CAGE[82,83]; nanoCAGE[70]; oligocapping 5′ end SAGE[85]; template switch[65]; 5′–3′ pair end tags (PET)[86]
Small RNA sequencing	Small RNAs (or cDNAs) of a particular size range, often 18–30 nt, are extracted using gel fractionation and are sequenced with next-generation sequencing[87]. The difference between these and 5′-based methods is the size of the RNA and that there is typically no cap selection	Protocols are flexible, as size range, modifications and cell compartments can be selected for. It is possible to select for small RNAs that are bound to specific molecules. As with 5′ tagging, the resolution is very high. A technical advantage is that most sequencing platforms can sequence these RNAs in their full length	As with 5′ tagging, only the result of transcription or RNA processing is measured. As miRNAs and tRNAs or other structure-derived small RNAs[88] dominate the samples, rare small RNAs are harder to characterize and require very deep sequencing	See REFS 81,86,87,99
RNAPII run-on: GRO–seq	This method sequences RNAs that are within elongating RNAPII based on incorporating a label into nascent RNA. The position of elongating RNAPII can be detected	This method can distinguish RNAs that are within RNAPII from other processed small RNAs. High sensitivity	Technically challenging, including isolation of nuclei	See REF. 56
Methods capturing DNA-bound proteins, including RNAPII, transcription factors and histone modifications				
ChIP–seq or ChIP–chip	Proteins bound to DNA are crosslinked by formaldehyde. DNA is then fragmented, and specific protein–DNA complexes can be extracted and isolated using antibodies. DNA fragments are sequenced from their 5′ ends and mapped to the genome (using ChIP–seq) or identified by microarray (using ChIP–chip)	Captures DNA-bound proteins 'in the act' — includes transcription factors, modified histones or proteins that are part of the transcription machinery, such as RNAPII	Heavily reliant on antibody specificity and quality; this makes comparisons between different ChIP experiments hard, even for the same target. The resolution is much lower than for RNA-based methods, as the ChIP fragments are much larger than the bound site. Most analysis also requires computational peak-calling, which has no general standards	A user manual of several technologies (ChIP–seq, DNase hypersensitive site analysis and DNA methylation analysis) was published by the ENCODE Consortium and provides an overview of these methods and applications[100]
Methods locating transcription 'bubbles'				
Permanganate footprinting	Identifies DNA regions that correspond to the 'melted' transcription bubble and can therefore locate the exact genomic sites at which active RNAPII is located	This method detects active RNAPII and not the products of RNA. This method is best at identifying poised RNAPII	At present, it can only be done by targeting specific genes, so throughput is limited	See REFS 101,102

CAGE, cap analysis of gene expression; ChIP–chip, chromatin immunoprecipitation followed by microarray; ChIP–seq, chromatin immunoprecipitation followed by sequencing; GRO–seq, global run-on followed by sequencing; RNAPII, RNA polymerase II; SAGE, serial analysis of gene expression; TSS, transcription start site.

doi:10.1038/nrg3163

REGIÃO PROMOTORA

Primeiros estudados – virus e genes *housekeeping* – ***TATA box* mais comum??**



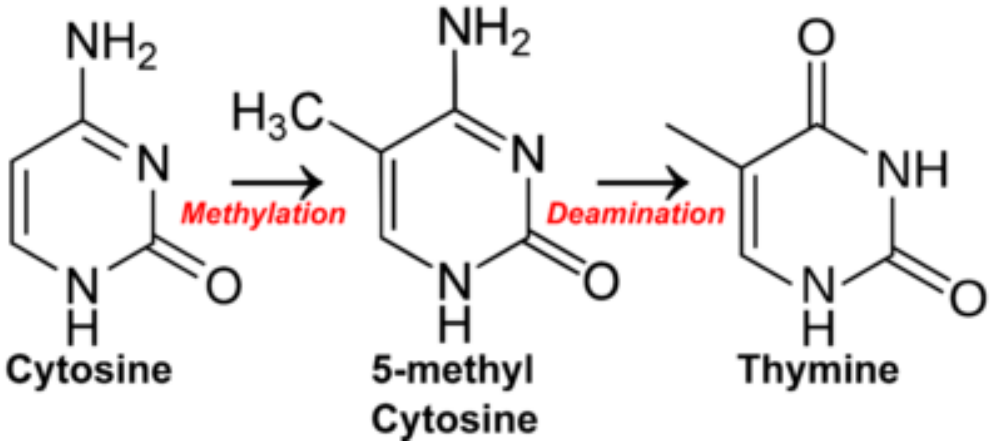
Sequencia iniciadoras – (5’) Y Y A^{+1} N T/A Y Y Y Y (3’)

Y – uma pirimidina

Aumento a força de ligação por fatores de transcrição!

ILHAS CpG

100 a 1000pb

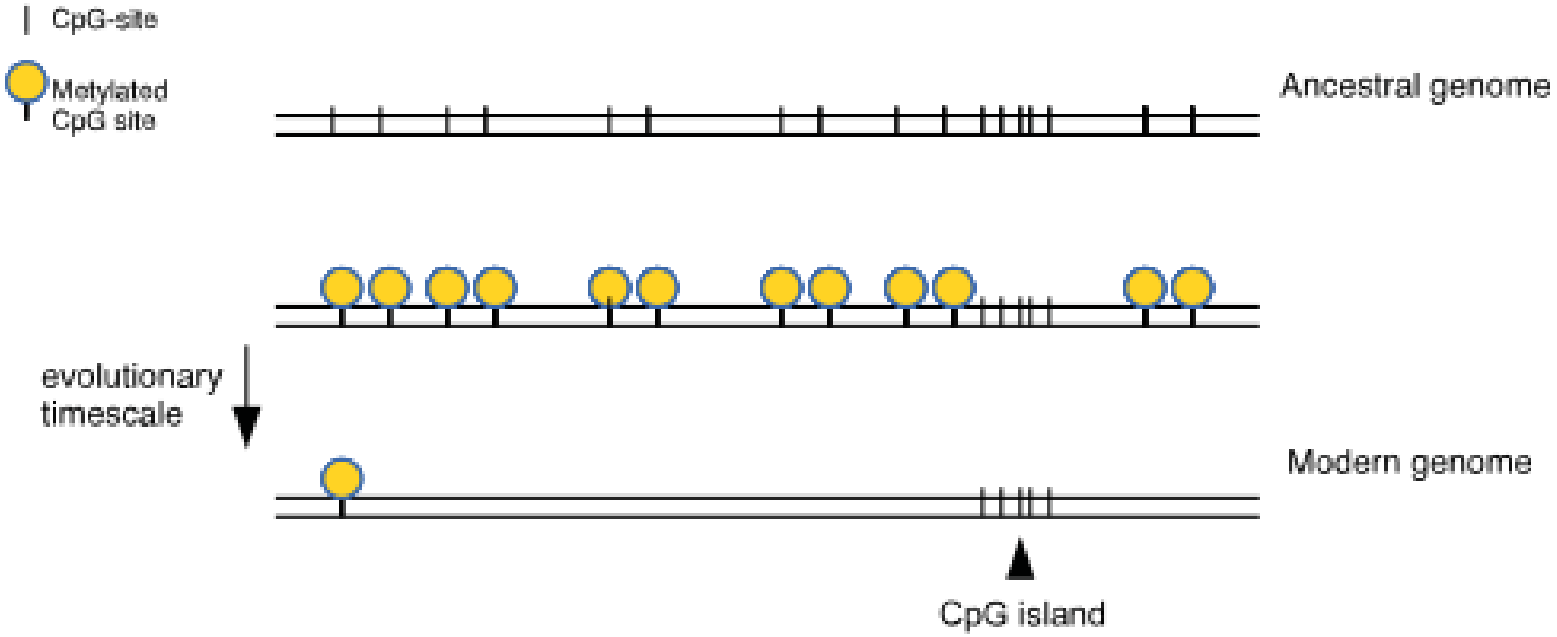


C – metilado – sofre desaminação - TG (21 % a menos do total esperado)

Como mantêm as ilhas CpG???

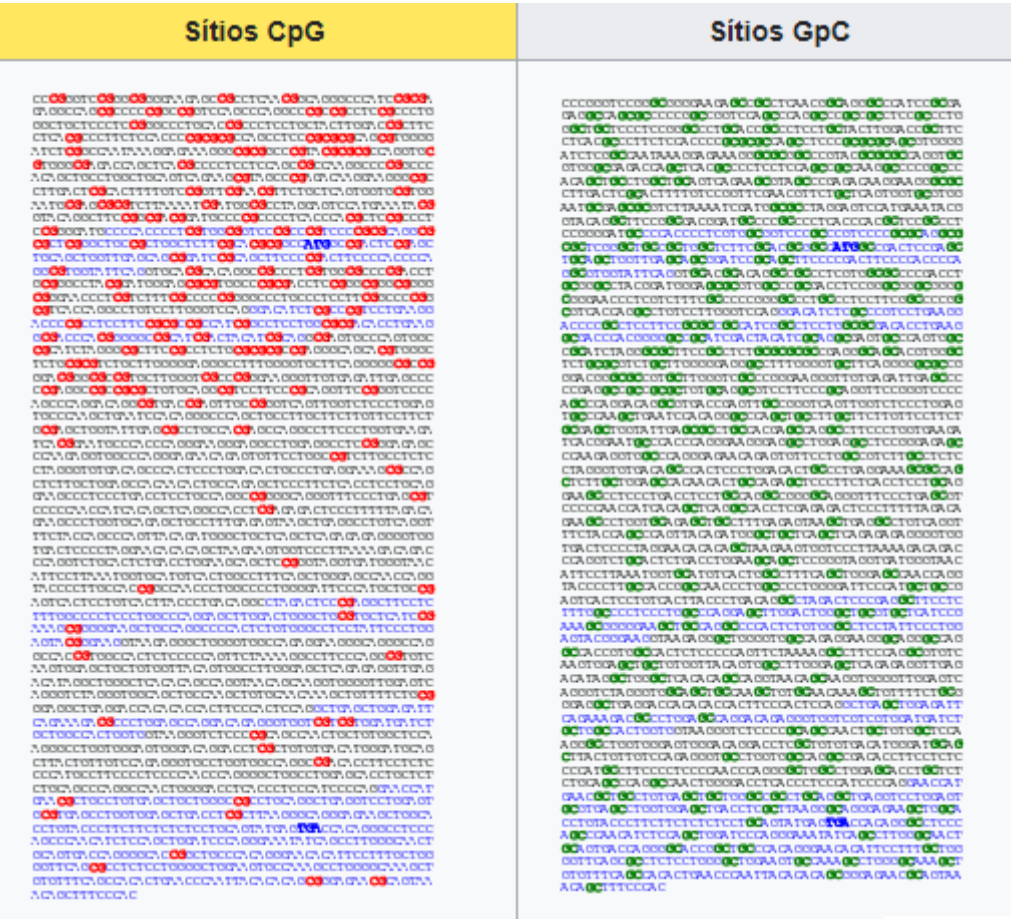
A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters

Serge Saxonov*†, Paul Berg†‡, and Douglas L. Brutlag*†§



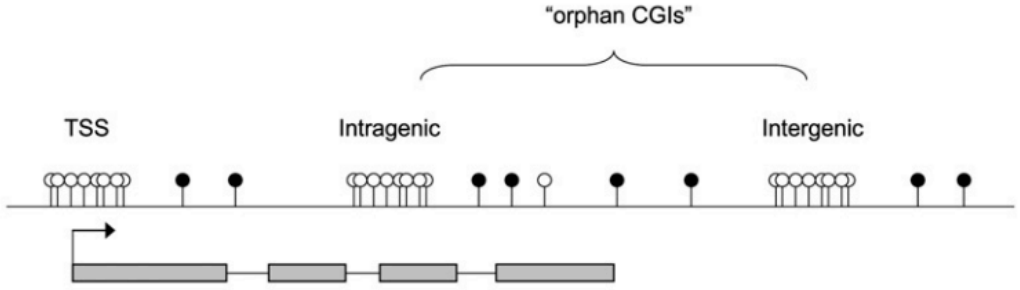
A grande maioria dos promotores em humanos são ilhas CpG – menor afinidade com histonas – fora dos nucleossomos!!!

gene APRT humano



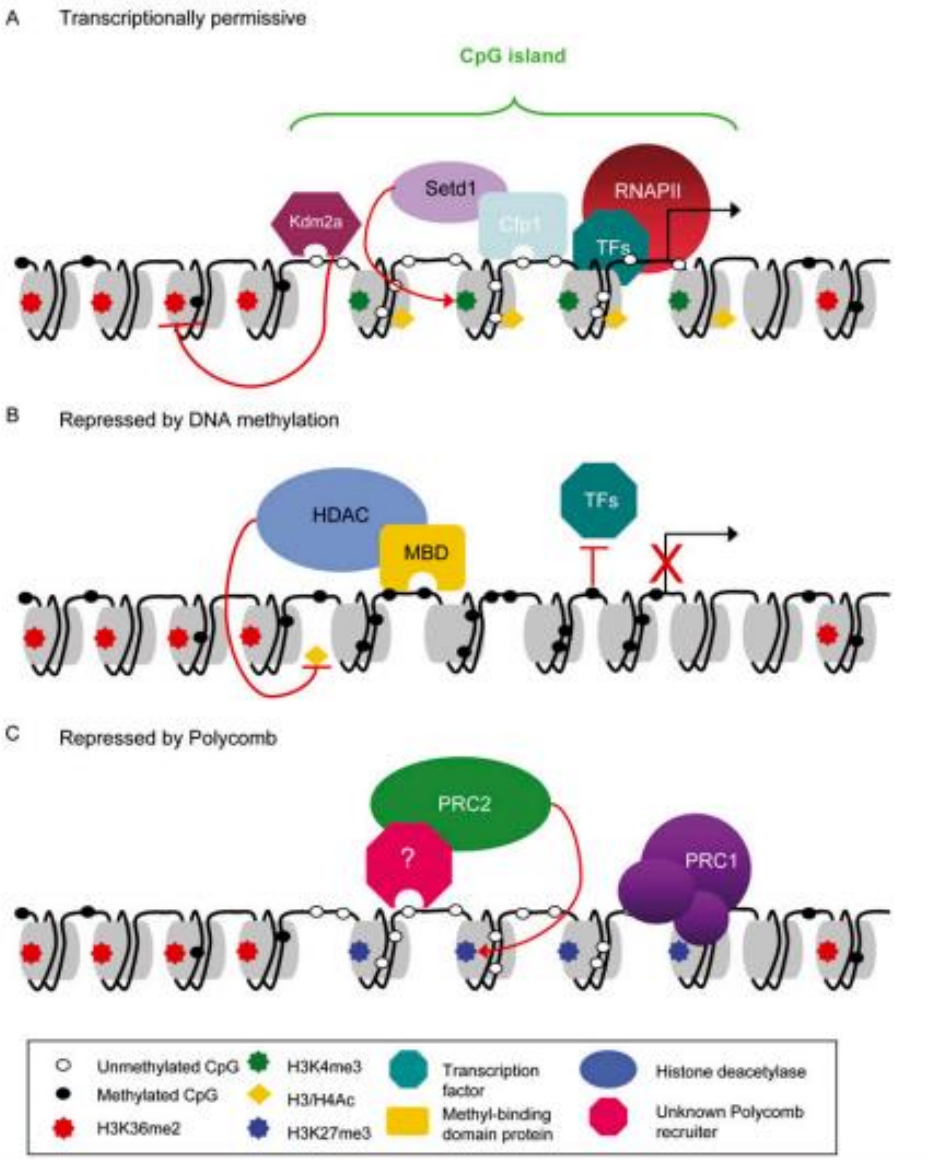
Padrão de hiper/hipometilação em células cancerígenas

Relógios biológicos – CpG metiladas



CpG islands and the regulation of transcription

Aimée M. Deaton and Adrian Bird[1]

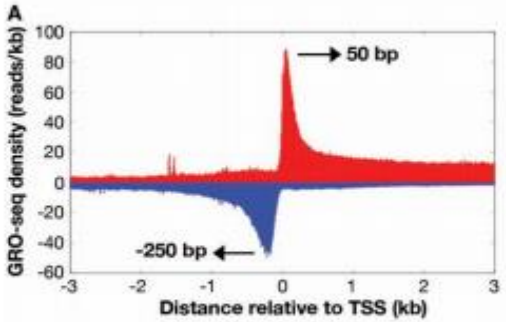


doi:10.1101/gad.2037511

Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters

Leighton J. Core,* Joshua J. Waterfall,* John T. Lis†

RNA polymerases are highly regulated molecular machines. We present a method (global run-on sequencing, GRO-seq) that maps the position, amount, and orientation of transcriptionally engaged RNA polymerases genome-wide. In this method, nuclear run-on RNA molecules are subjected to large-scale parallel sequencing and mapped to the genome. We show that peaks of promoter-proximal polymerase reside on ~30% of human genes, transcription extends beyond pre-messenger RNA 3′ cleavage, and antisense transcription is prevalent. Additionally, most promoters have an engaged polymerase upstream and in an orientation opposite to the annotated gene. This divergent polymerase is associated with active genes but does not elongate effectively beyond the promoter. These results imply that the interplay between polymerases and regulators over broad promoter regions dictates the orientation and efficiency of productive transcription.



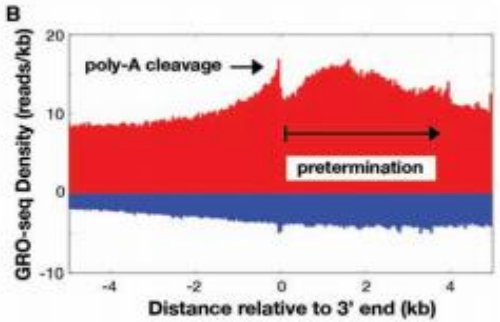


Fig. 2. Alignment of GRO-seq reads to TSSs and 3′ ends. **(A)** GRO-seq reads aligned to Ref-seq TSSs in 10-bp windows in both sense (red) and antisense (blue) directions relative to the direction of gene transcription. **(B)** GRO-seq reads flanking the 3′ ends of genes. The sharp peak coincides with the new 5′ end created after cleavage at the poly-A site. Polymerase density extends considerably downstream before termination.

PRINCIPAIS PROMOTORES EM ANIMAIS

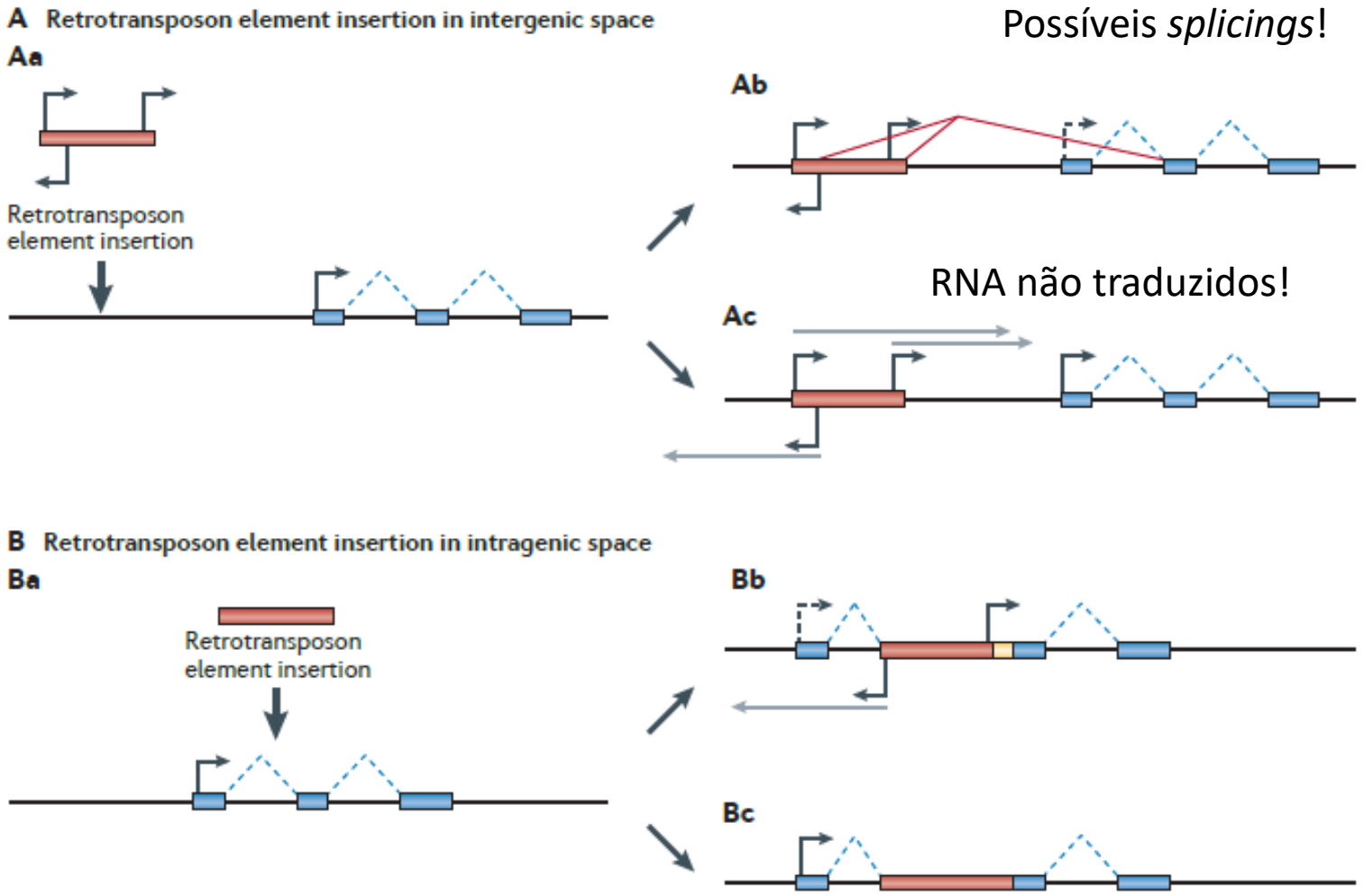
Table 2 | **Promoter types**

Promoter type	Dominant gene function	Common properties	Vertebrate-specific	*Drosophila melanogaster*-specific	Refs
Major promoters					
Type I ('adult')	Tissue-specific expression in adult peripheral tissues	Sharp ('focused') TSS, TATA-box enrichment, disordered nucleosomes	Mostly no CpG islands		8,9,13,17
Type II ('ubiquitous')	Broad expression throughout organismal cycle	Broad ('dispersed') TSS, ordered nucleosome configuration	CpG islands, TATA-depleted	Enrichment of non-positionally fixed motifs (Motif 1 or 6, DRE)	8,9,13,17
Type III ('developmentally regulated')	Differentially regulated genes, often regulators in multicellular development and differentiation	Polycomb repression-regulated genes, broad H3K27me3 marks	Large CpG islands extending into the body of gene	Enriched for DPE	16
Minor promoters					
TCT promoter	Highly expressed genes of translational apparatus	Sharp, pyrimidine-stretch ('TCT') initiator sequence, often full TATA box, ubiquitous-promoter-like nucleosome configuration	CpG island overlapping		23

DPE, downstream promoter element; DRE, DNA recognition element; H3K27me3, histone H3 lysine 27 trimethylation; TSS, transcription start site.

doi:10.1038/nrg3163

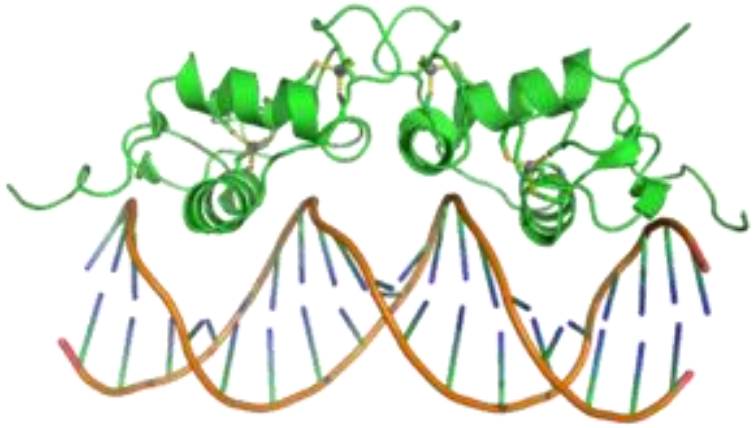
Retrotransposons e seus promotores influênciam a expressão gênica



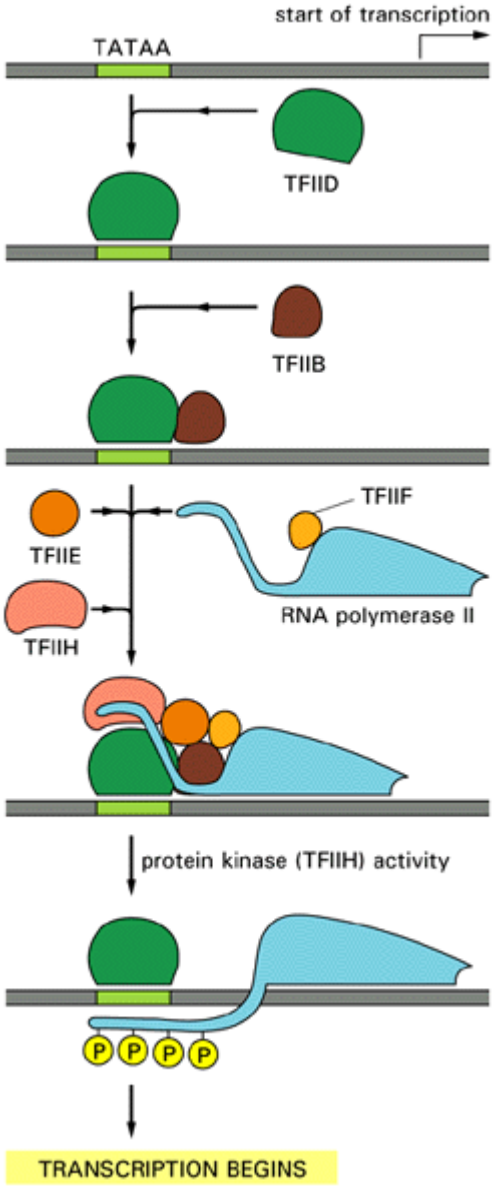
doi:10.1038/nrg3163

FATORES DE TRANSCRIÇÃO

Proteínas que se ligam ao DNA de células eucarióticas para permitir que haja uma ligação entre a enzima RNA-polimerase e o DNA, permitindo assim a transcrição. Qualquer proteína necessária para o início da transcrição, mas que não seja parte integrante do RNA polimerase.



Sequências regulatórias x Reguladores transcricionais



FATORES GERAIS DA TRANSCRIÇÃO

Conservado em todos os eucariotos:

TFIID – composta por uma proteina TBP

TFIIB – N terminal liga a RNPII (canal de saída de RNA)

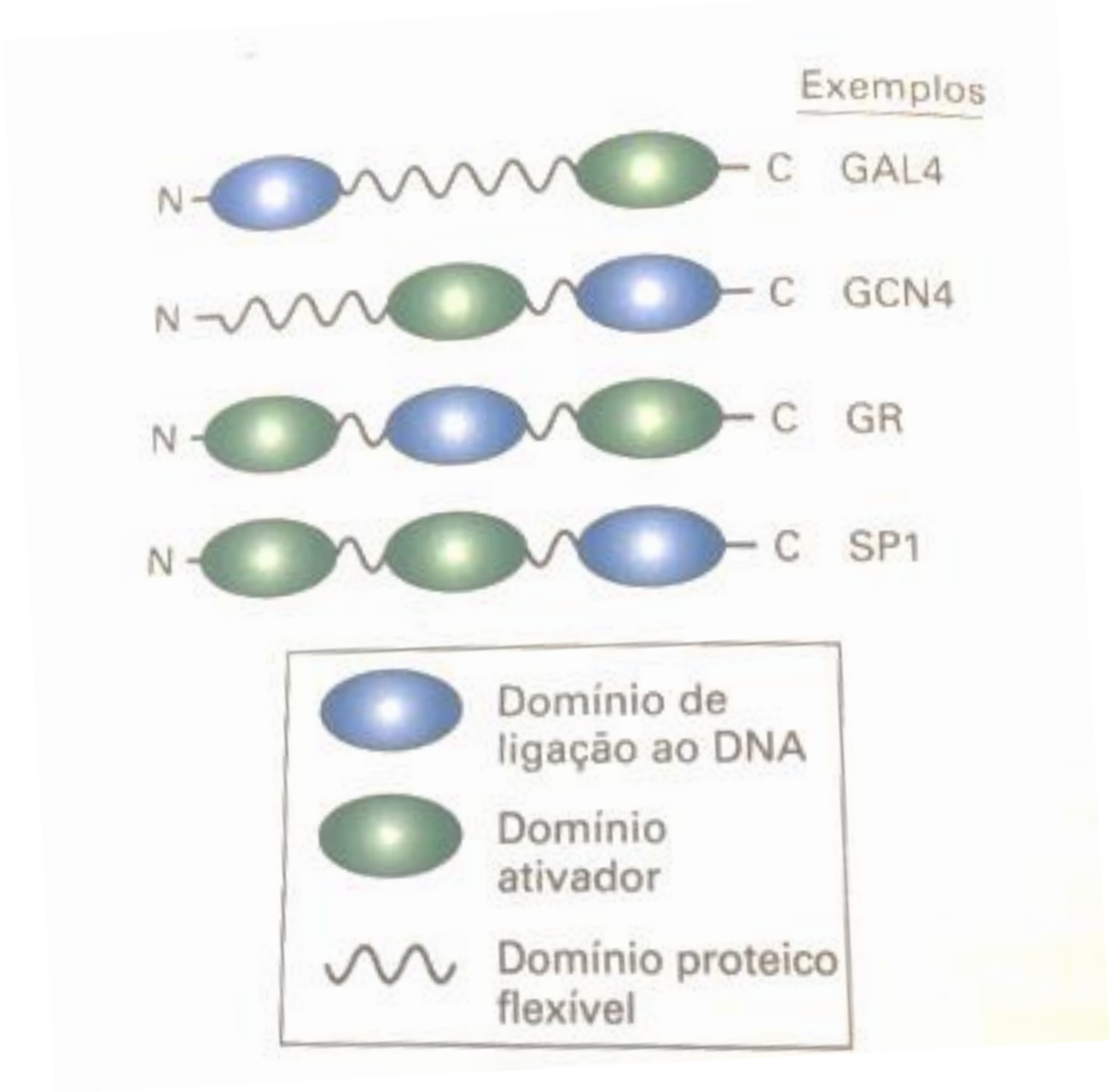
TFIIF – heterodimero

TFIIE – tetramérica

TFIIH – multimérica – fosforilação do CDT da RNPII

Isso *in vitro*...

EXEMPLOS DE TFs E SEUS DOMÍNIOS FUNCIONAIS



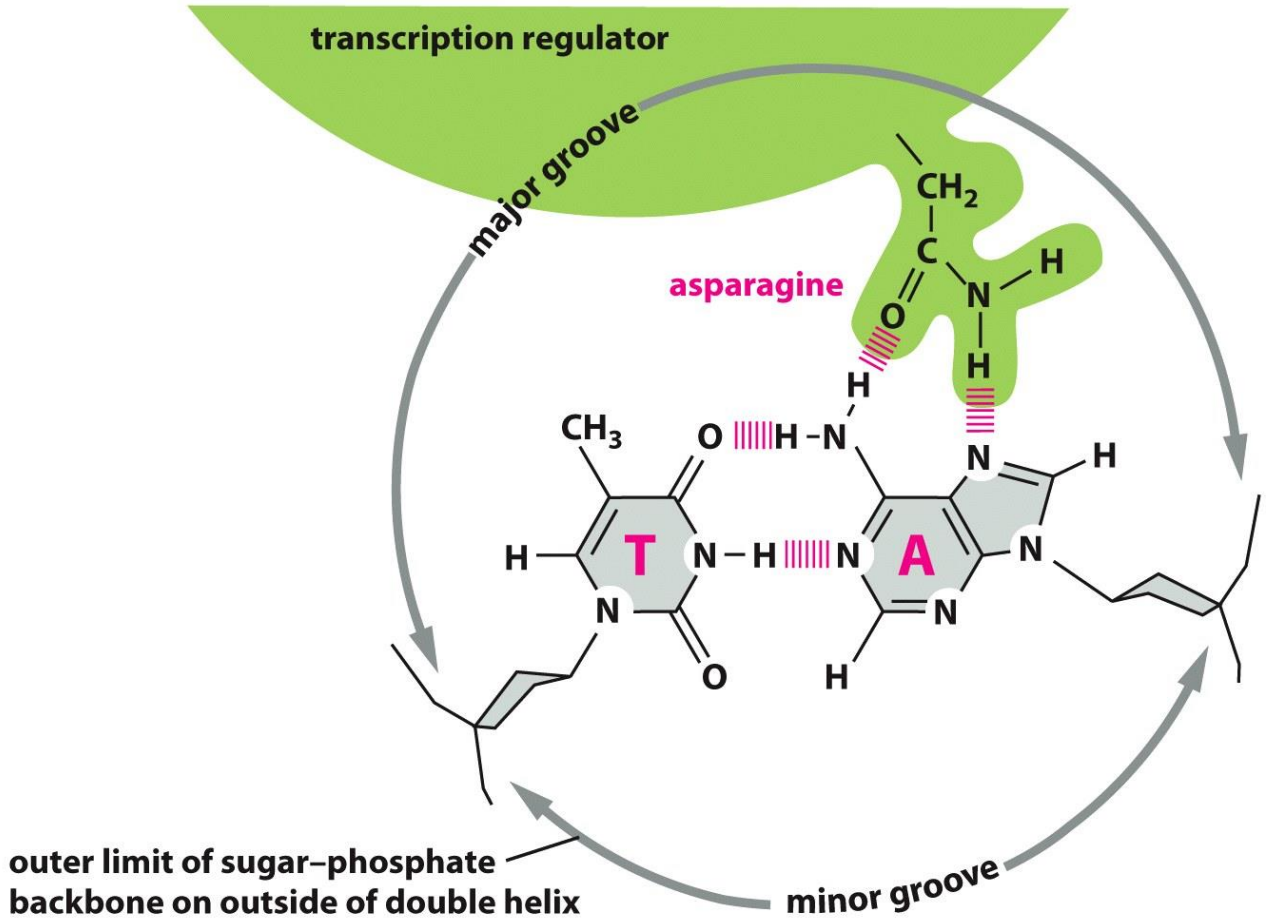
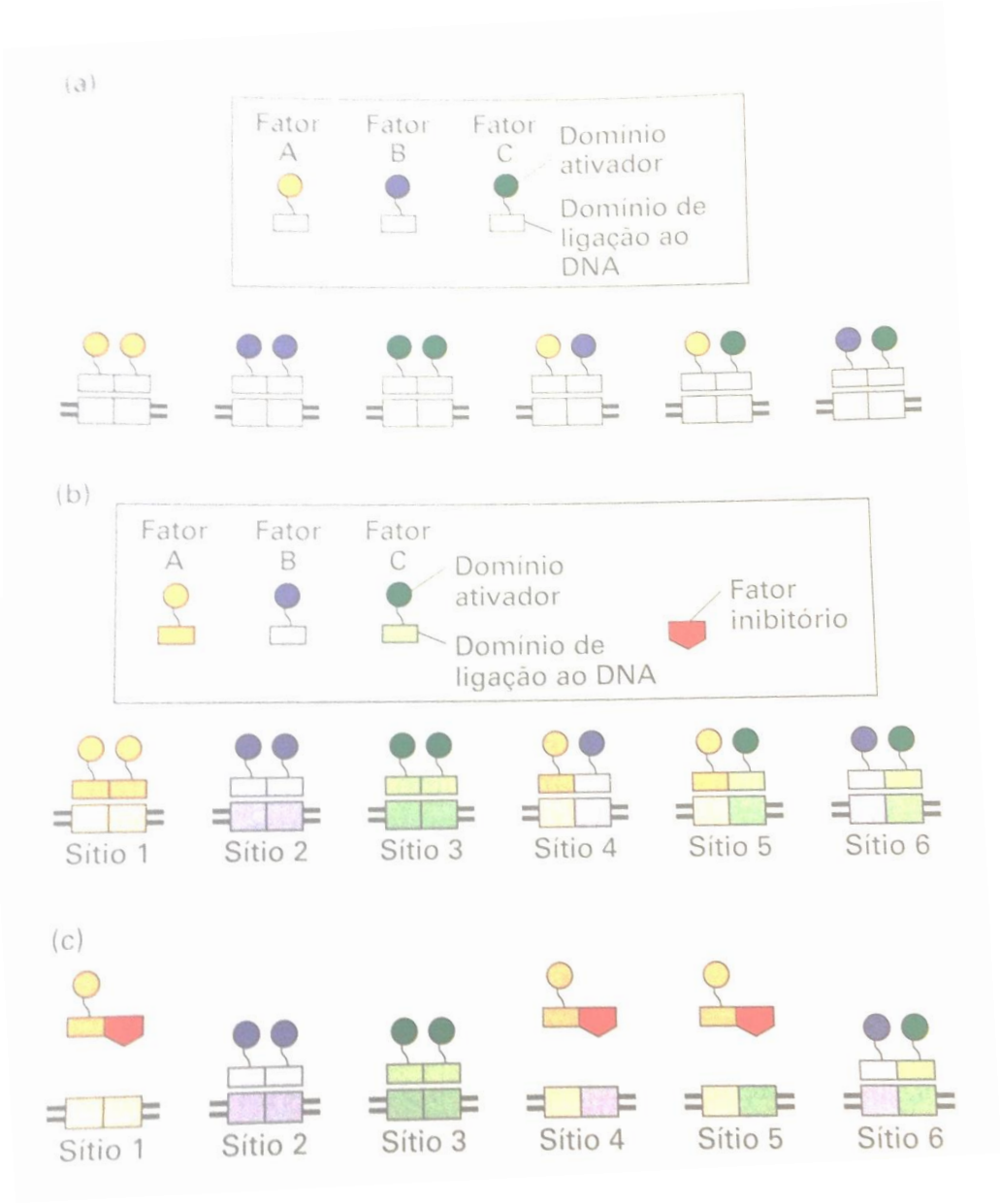


Figure 8-4 Essential Cell Biology 3/e (© Garland Science 2010)

Muito sitios de contato… a união faz a força!!!

Geralmente dímeros – aumento o número de sítio que podem se ligar!!

FATORES DE TRANSCRIÇÃO HETERODIMÉRICOS



FATORES DE TRANSCRIÇÃO COMO ATIVADORES E REPRESSORES

- Geralmente ação a distância;
- Ajuda a associação dos fatores gerais de transcrição e RNApolI;
- Modelagem da cromatina.

TF I - proteínas que são imprescindíveis para a ligação a RNA-polimerase I, responsáveis pela transcrição de genes que resultam em RNAs ribossomais.
TF II - proteínas que são imprescindíveis para a ligação a RNA-polimerase II, responsáveis pela transcrição de genes que resultam em RNAs mensageiros.
TF III - proteínas que são imprescindíveis para a ligação a RNA-polimerase III, responsáveis pela transcrição de genes que resultam principalmente em RNAs transportadores.

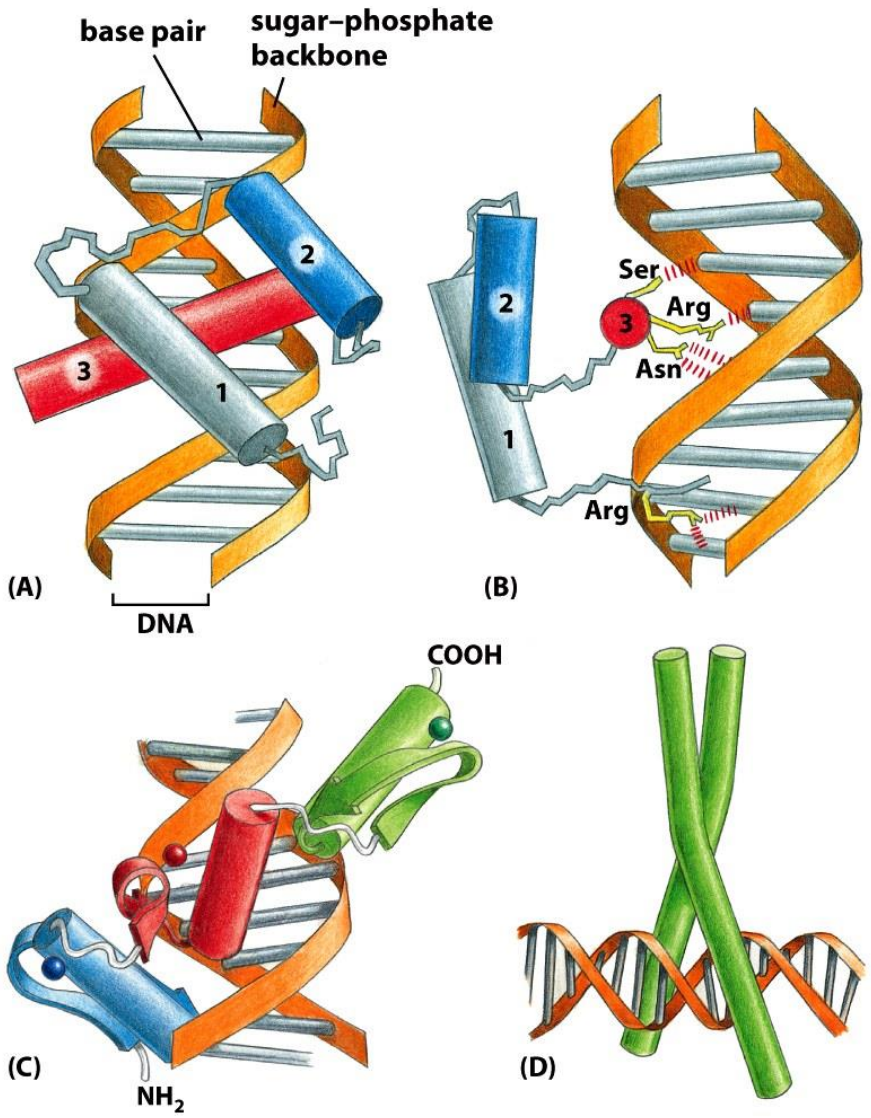


Figure 8-5 Essential Cell Biology 3/e (© Garland Science 2010)

Diferentes motivos de ligação ao DNA:

a e b – comum homeodomíos de α-hélice

c – motivos dedo de zinco

d - Ziper de leucina

MODELO ACEITO PARA O USO DAS SEQUÊNCIAS REGULATÓRIAS!

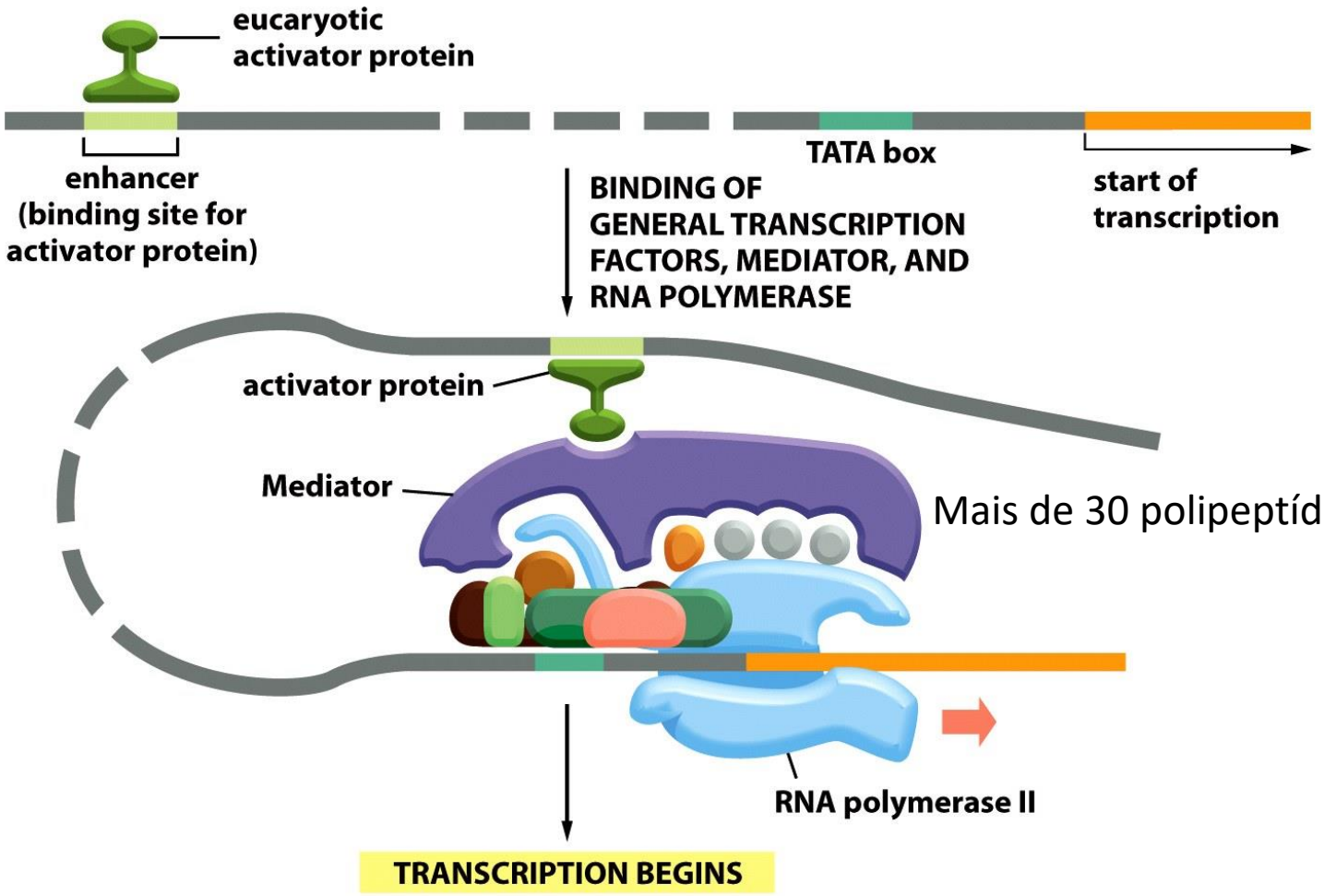
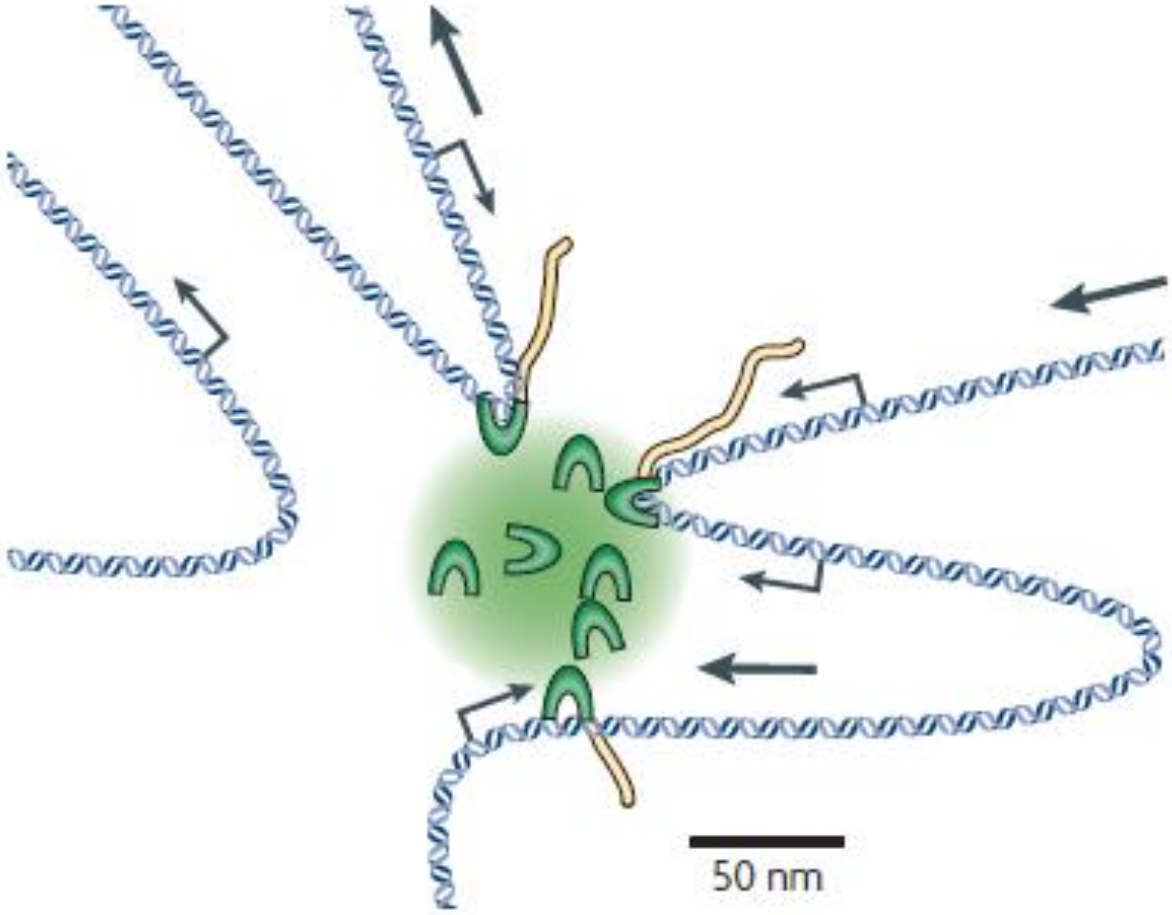


Figure 8-10 Essential Cell Biology 3/e (© Garland Science 2010)

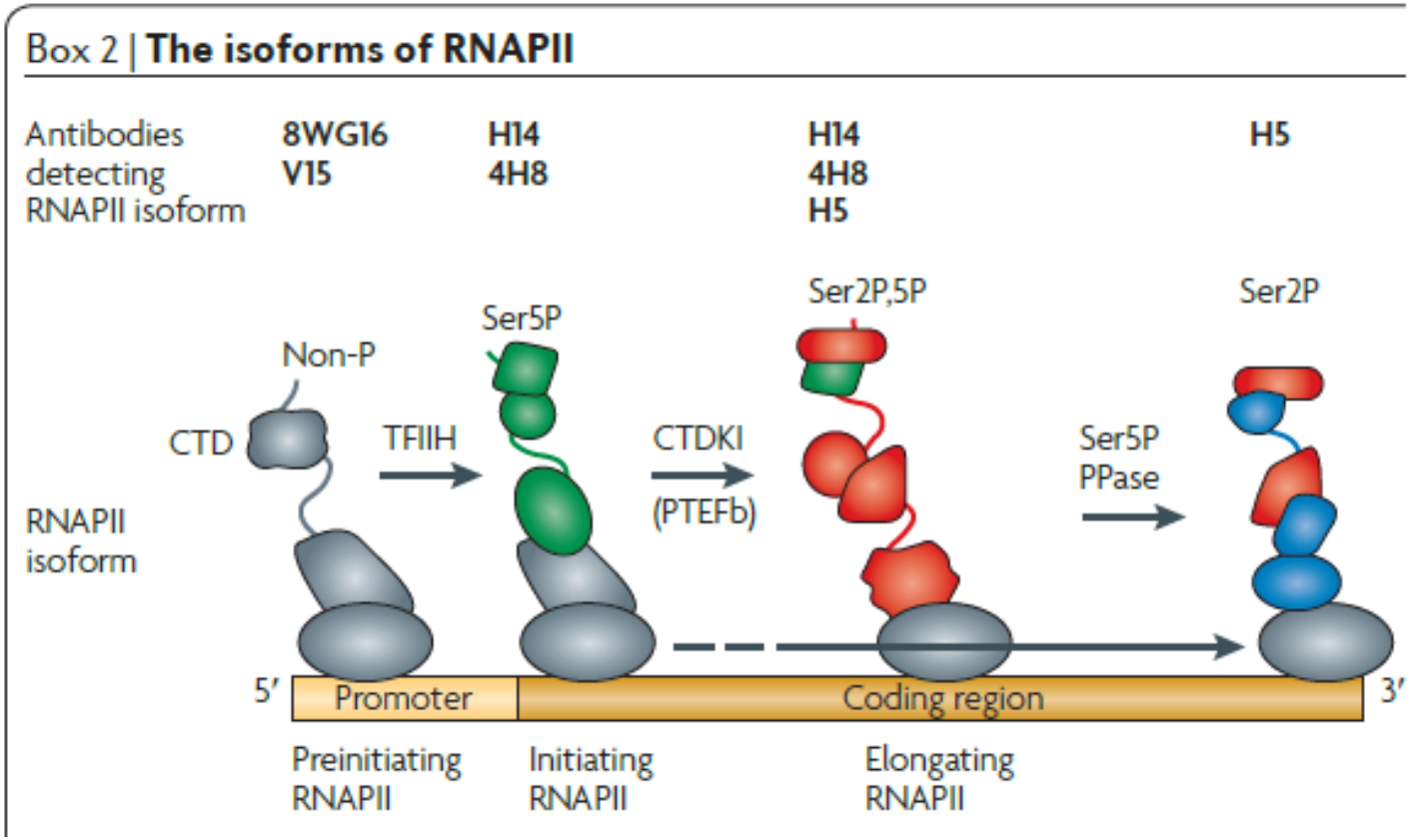
Mais de 30 polipeptídeos!!

MUDANDO O PARADIGMA: A MONTANHA VAI ATÉ MAOMÉ...



Estudos com DNA e RNA-FISH e cromossomos politênicos

doi:10.1038/nrg2592

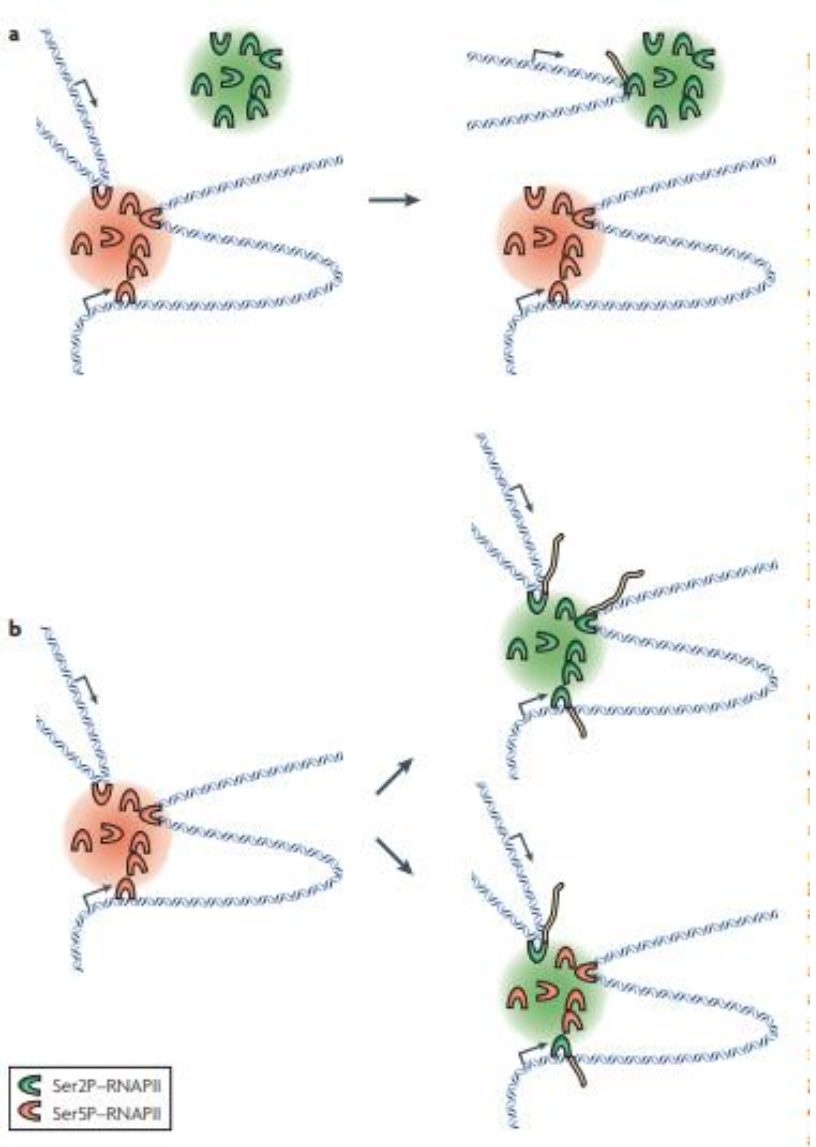


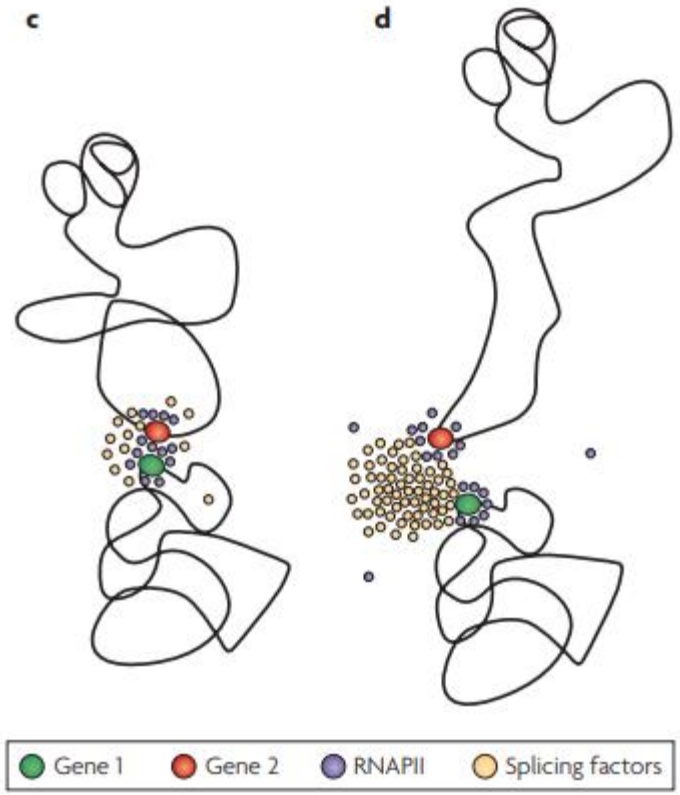
Os fatores de transcrição são responsáveis por obtenção de isoformas da RNAP II – fosforilando resíduos distintos de CDT da Rpb1 nas diferentes etapas da transcrição.

doi:10.1038/nrg2592

Modelo aidna em estudo…

doi:10.1038/nrg2592





Spliceossomos envolvidos?

COORDENAÇÃO DA EXPRESSÃO GÊNICA

Mesma sequência de reconhecimento das proteínas - genes controlados em conjunto!

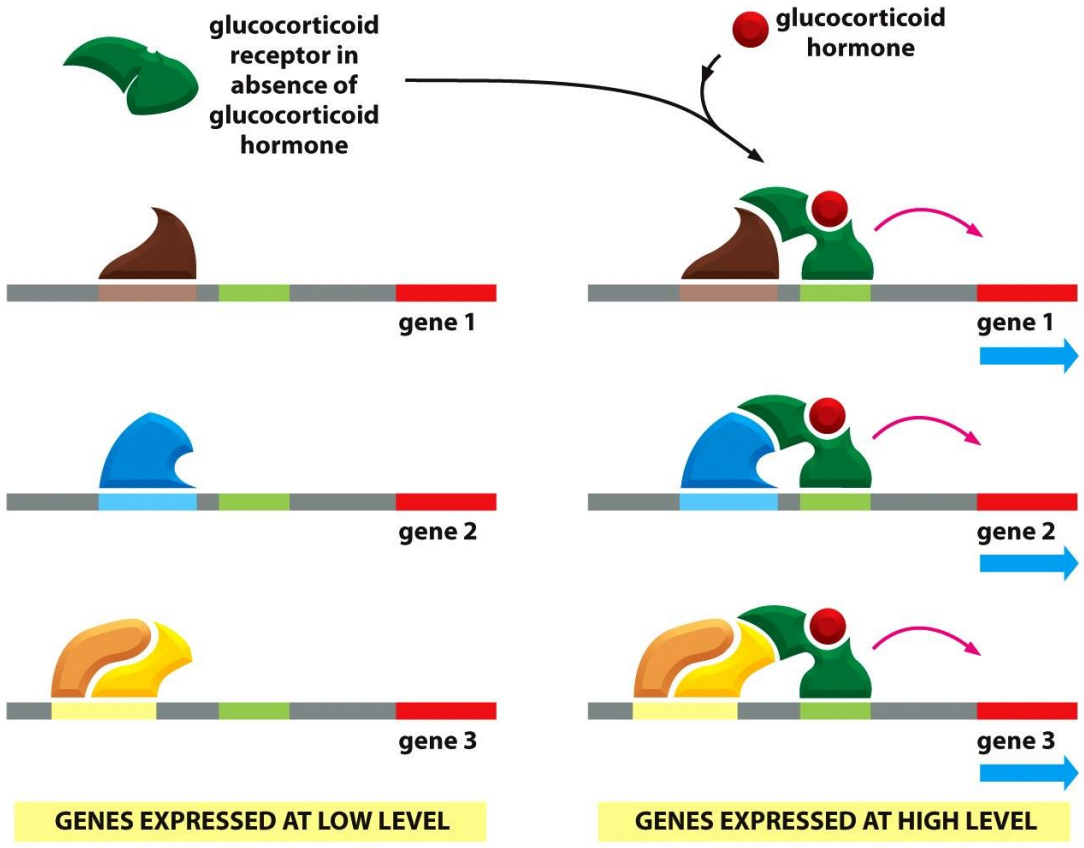


Figure 8-17 Essential Cell Biology 3/e (© Garland Science 2010)

O controle combinatório e as sequências regulatórias compartilhadas permitem que um conjunto limitado de reguladores controlem vários genes…

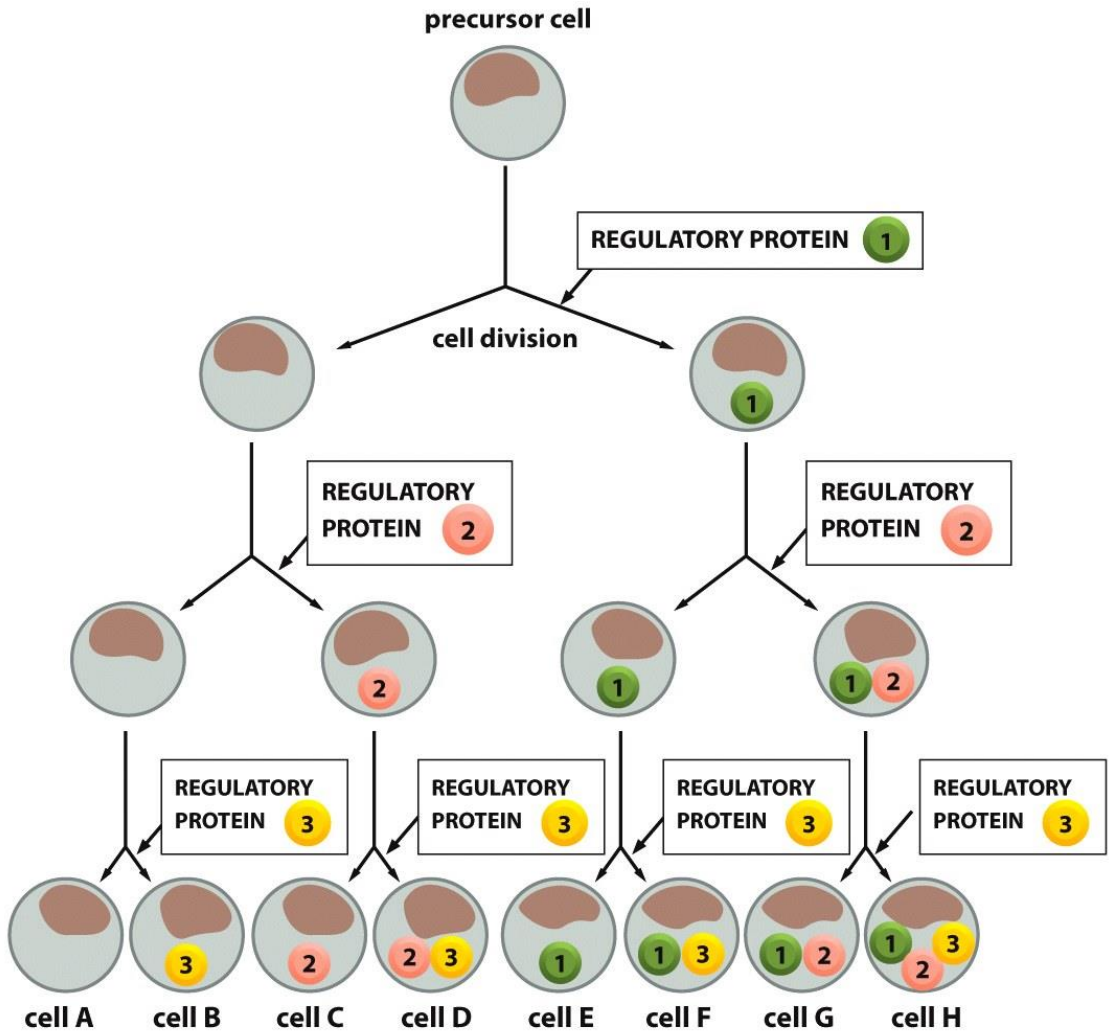


Figure 8-19 Essential Cell Biology 3/e (© Garland Science 2010)

CIRCUITO DE RETROALIMENTAÇÃO POSITIVA

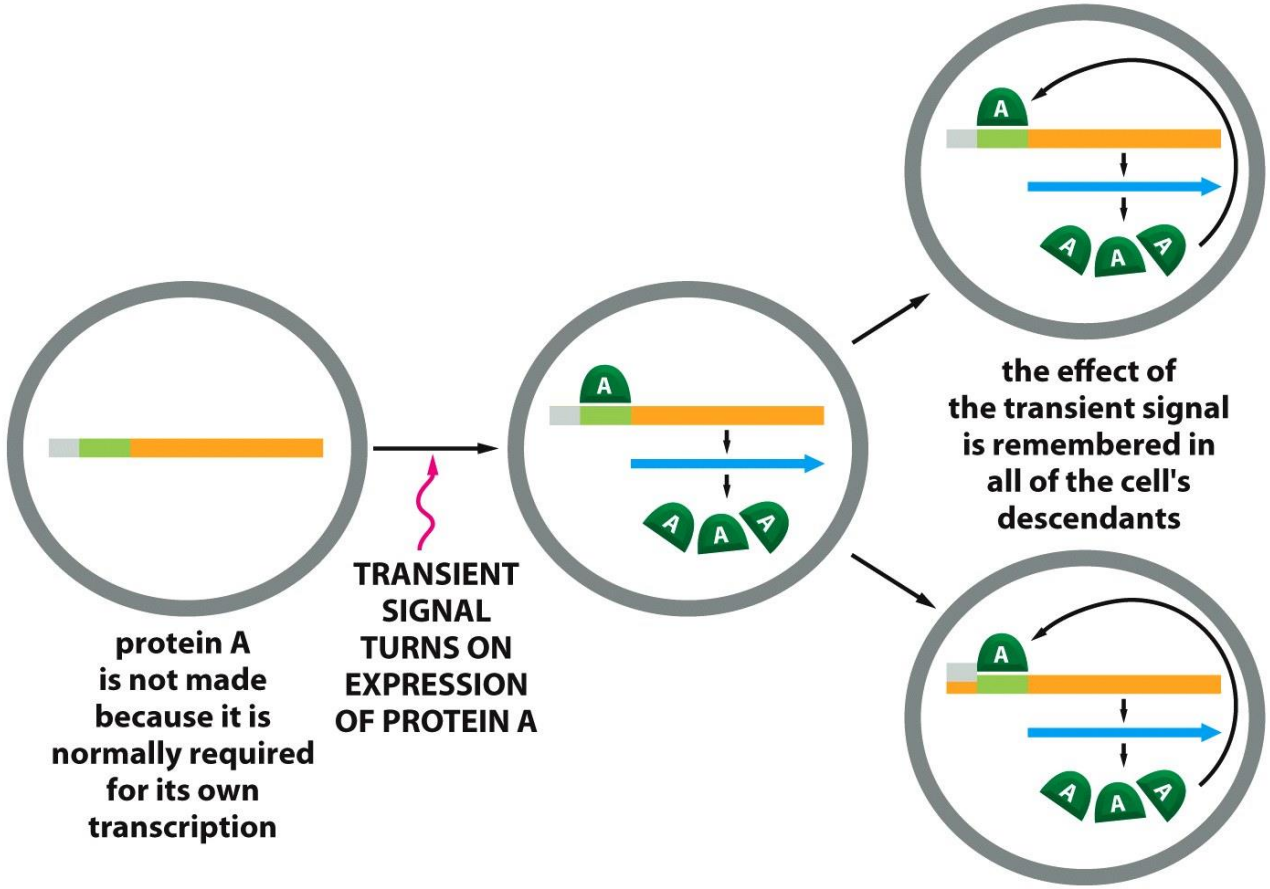


Figure 8-20 Essential Cell Biology 3/e (© Garland Science 2010)

Importante para o entendimento da diferenciação!!!

OPINION

Insulators and promoters: closer than we think

Jesse R. Raab and Rohinton T. Kamakaka

Regiões regulatórias que controlam a expressão de genes restringindo a ação de *enhancers* e silenciadores

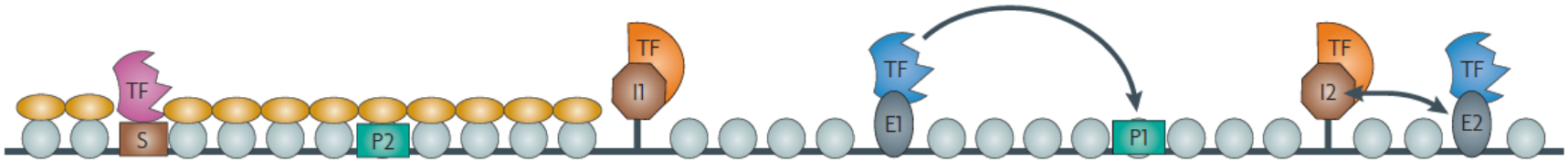
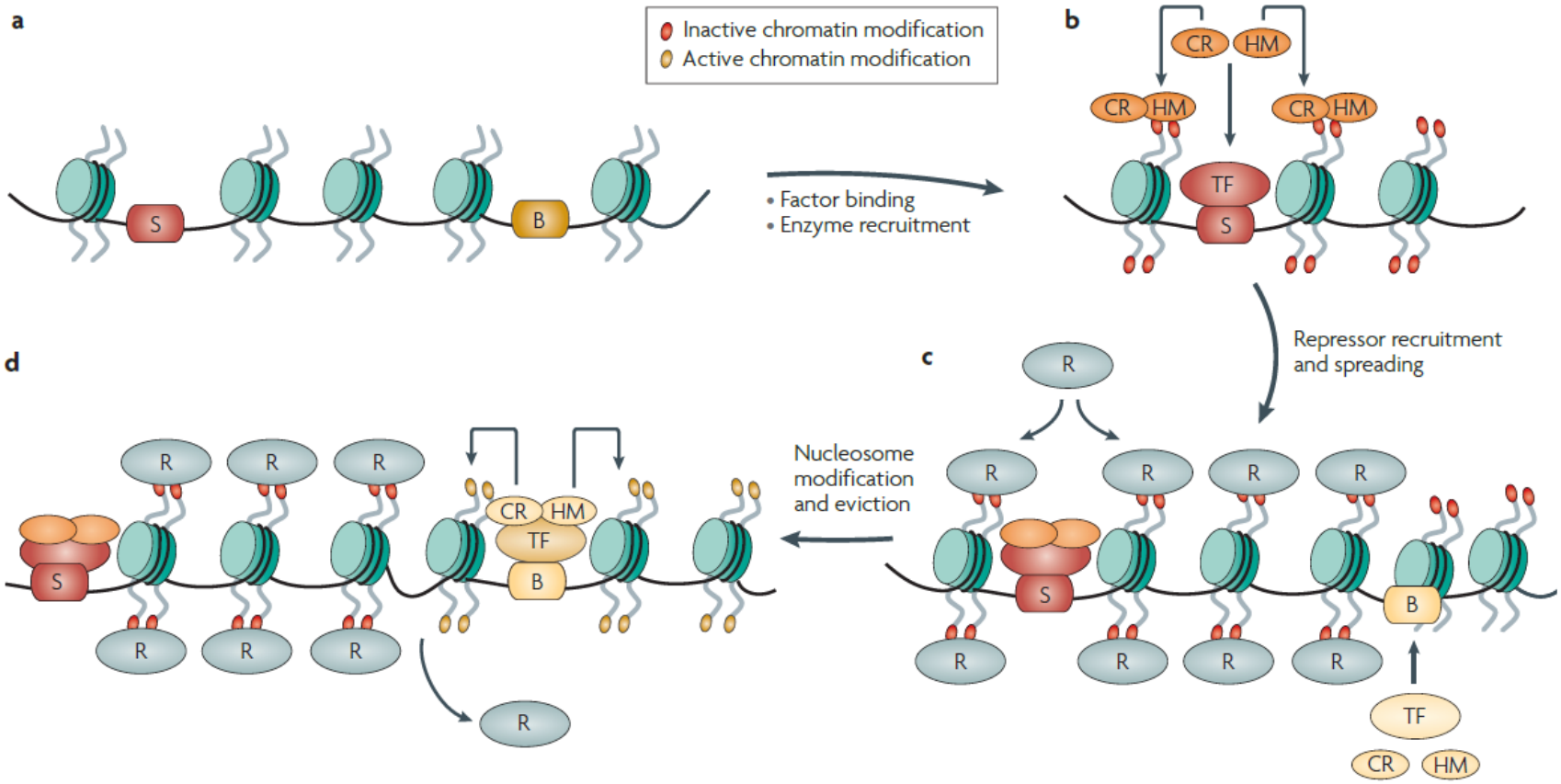
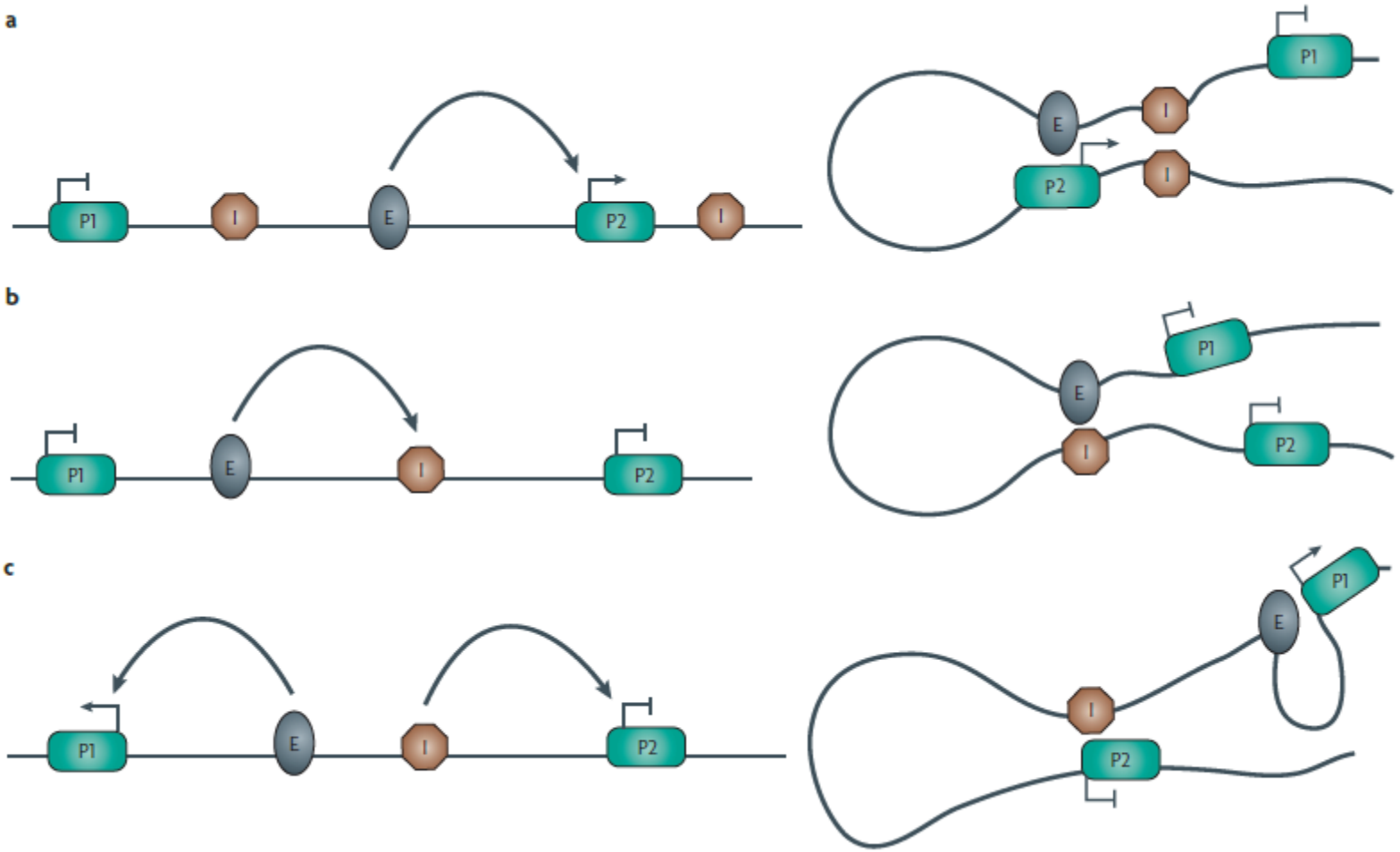


Figure 1 | **Chromatin domains and regulatory elements.** The relationships among silencers, enhancers, promoters and insulators are shown. Light blue circles represent nucleosomes and yellow ovals represent silencing proteins, such as heterochromatin protein 1 (HP1) and Sir3. Silencer elements (S) are sites of initiation of heterochromatin, which spreads and encompasses promoters (P2 in the diagram), silencing transcription. The I1 insulator functions to restrict the spread of heterochromatin. An enhancer (E1) that is present in an active chromatin domain flanked by insulators (I1 and I2) and that is bound by a transcription factor (TF) is able to communicate with a promoter (P1) in the same domain, whereas another enhancer (E2) is unable to communicate with promoter P1 because of an intervening insulator (I2).

AÇÃO DOS *INSULATORS* EM SILENCIADORES



AÇÃO DOS *INSULATORS* EM *ENHANCERS*



Transcriptional enhancers: from properties to genome-wide predictions

Daria Shlyueva, Gerald Stampfel* and Alexander Stark*

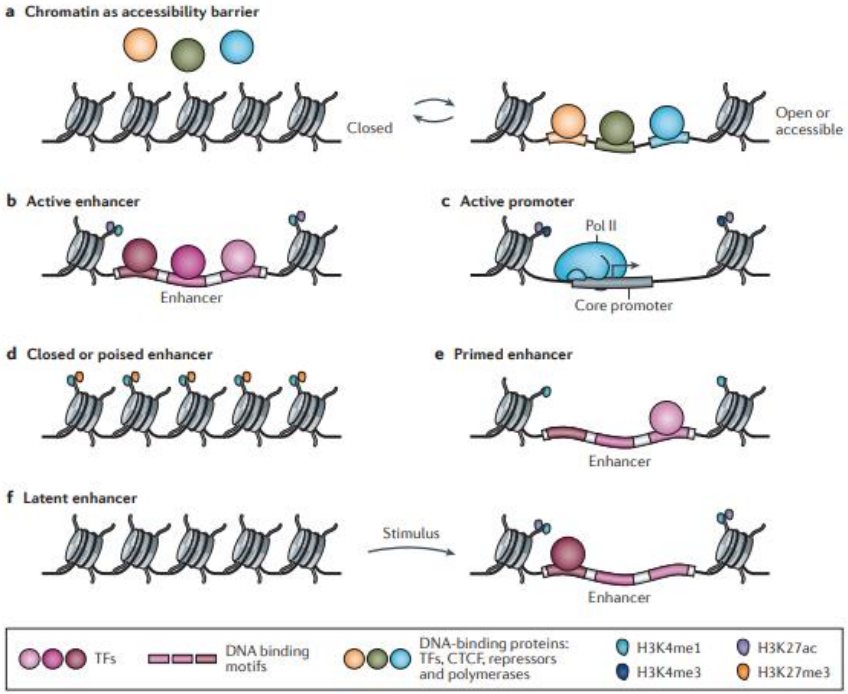


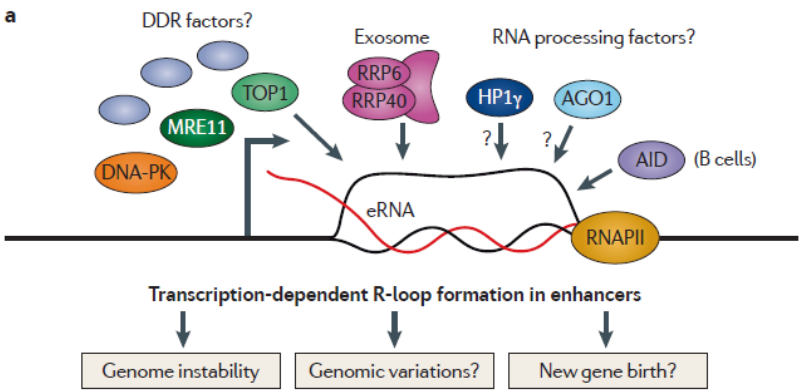
Table 1 | **Overview of enhancer testing and screening methods***

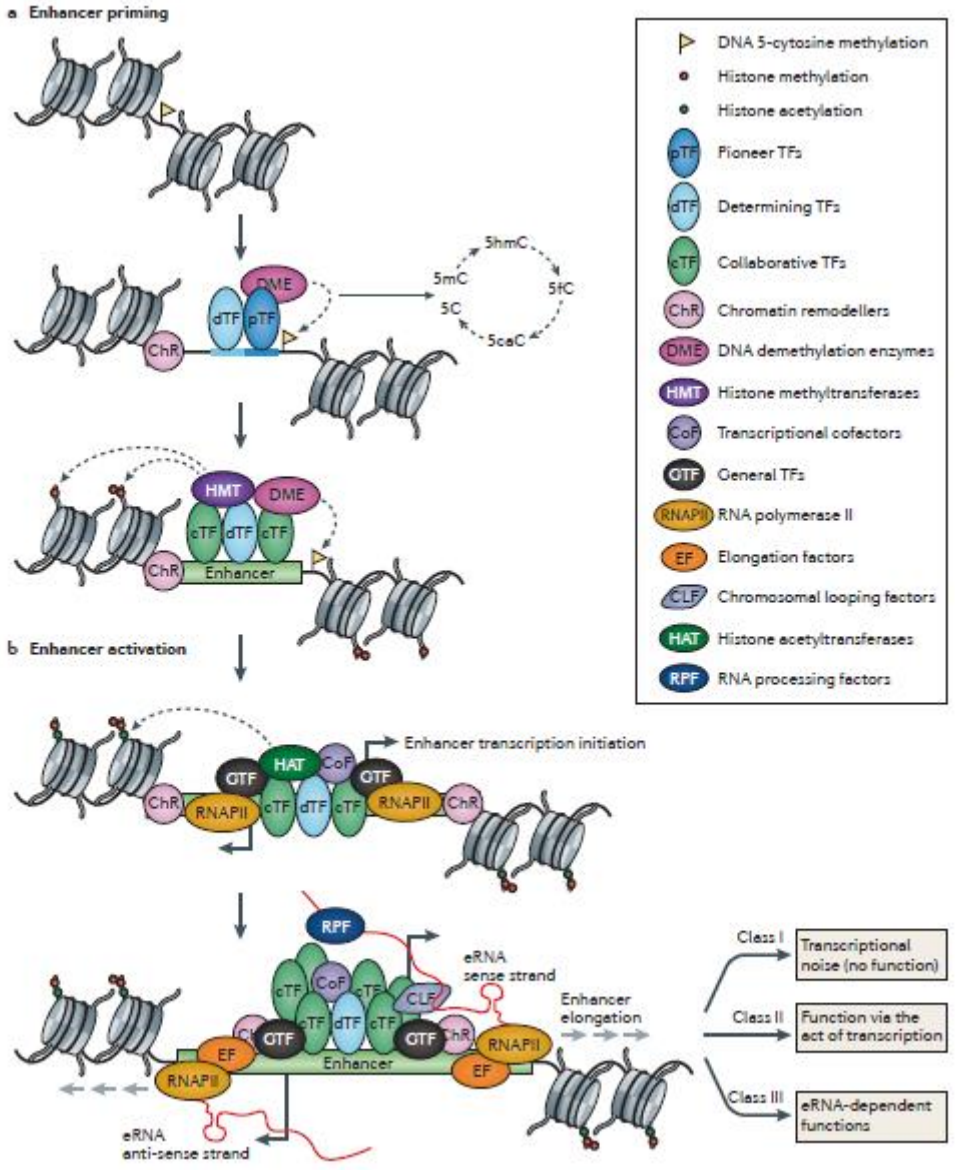
Methods	Throughput	Model system	Plasmid or integrated	Reporter set-up‡	Quantification
Image-based methods[53,121,123] (reviewed in REF. 4)	Up to thousands	Mouse and fly; *in vivo*	Random or site-specific integration		Staining intensity
Enhancer-FACS-seq[124]	Hundreds	Fly; *in vivo*	Site-specific integration		Fluorescence intensity and sequencing
Sharon *et al.*[129]	Thousands	Yeast	Plasmid		Fluorescence intensity
CRE-seq[165]	Thousands	Yeast	Site-specific integration		Sequencing of transcribed barcode
Nam *et al.*[125]	Hundreds	Sea urchin; *in vivo*	Random integration		Sequencing of transcribed barcode
MPRA[126],CRE-seq[128] and MPFD[127]	Thousands	Human and mouse; *in vitro*	Plasmid		Sequencing of transcribed barcode
STARR-seq[72]	Millions	Human and fly; *in vitro*	Plasmid		Sequencing of transcribed enhancer

CRE-seq, *cis*-regulatory element analysis by sequencing; FACS, fluorescence-activated cell sorting; MPFD, massively parallel functional dissection; MPRA, massively parallel reporter assay; STARR-seq, self-transcribing active regulatory region sequencing. *This table provides an overview of image- and deep-sequencing-based methods with their respective throughput and model system: sea urchin[125], yeast[129,165], mouse, human or fly[72,124,126-128]. Important properties of such assays include whether the candidate fragment is tested ectopically on a plasmid or after genomic integration, the reporter set-up with respect to the relative position of the candidate fragments or barcodes, and whether and how enhancer activity is quantified. ‡Tested fragments are shown in blue. Reporter genes are shown in beige; fluorescent or otherwise visually detectable reporter genes are shown in yellow. Barcodes are represented by red bars, and regions used for identifying fragments in a multiplexed sample are denoted by grey lines.

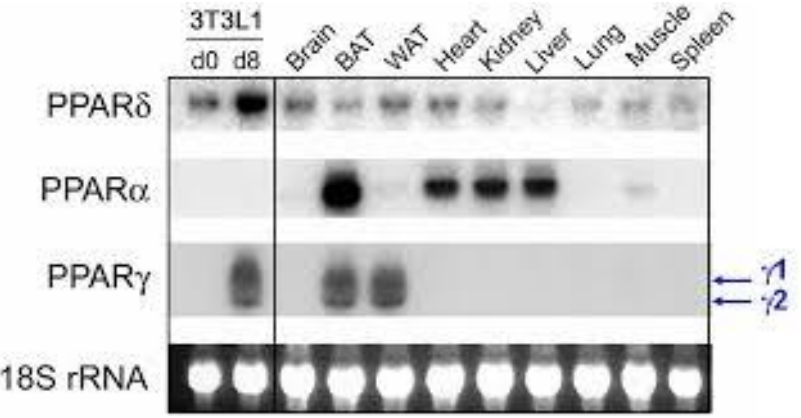
Enhancers as non-coding RNA transcription units: recent insights and future perspectives

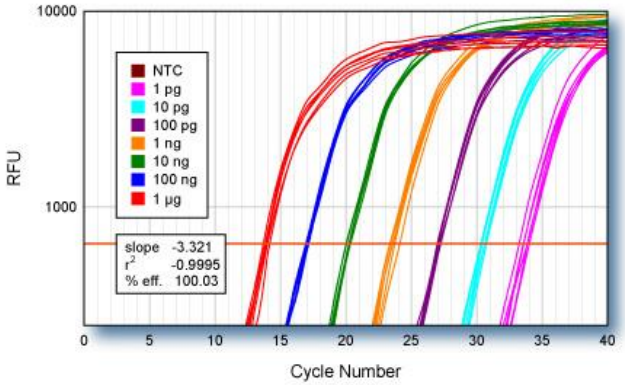
Wenbo Li, Dimple Notani and Michael G. Rosenfeld

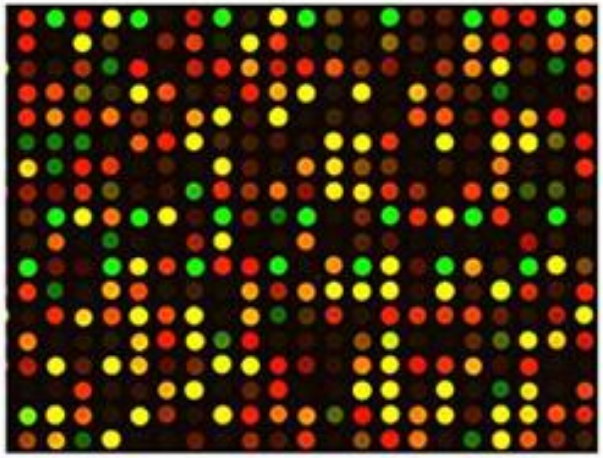




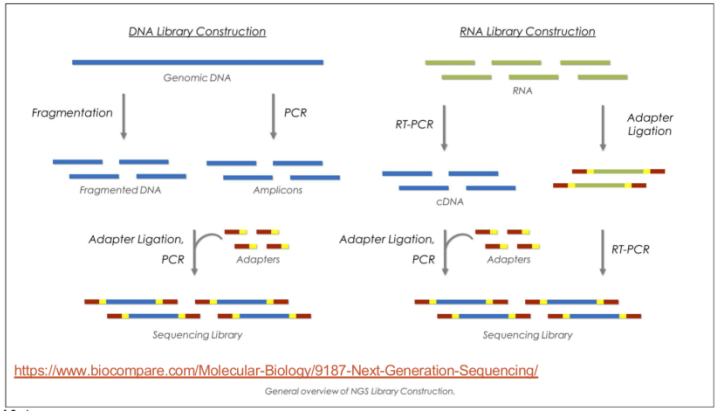
EM EVOLUÇÃO CONSTANTE ….







http://nextgen.mgh.harvard.edu/IlluminaChemistry.html



https://www.biocompare.com/Molecular-Biology/9187-Next-Generation-Sequencing/

General overview of NGS Library Construction.

PRÓXIMA AULA –10 de Maio

Aula Epigenética – 9 horas

Não esqueçam dos relatórios...