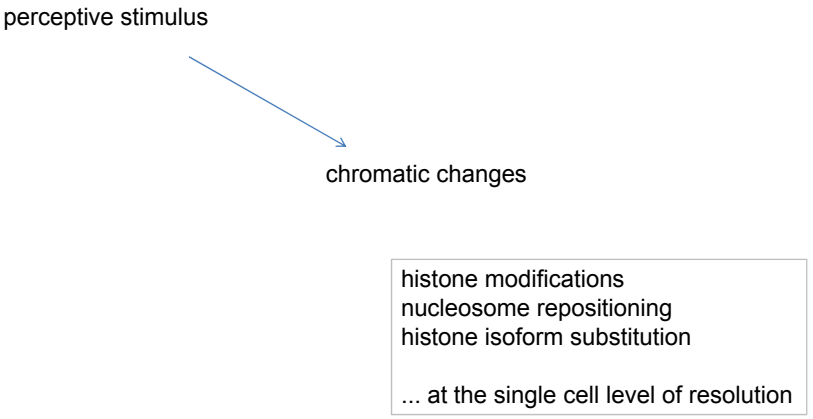
What we have seen yesterday in TV .....

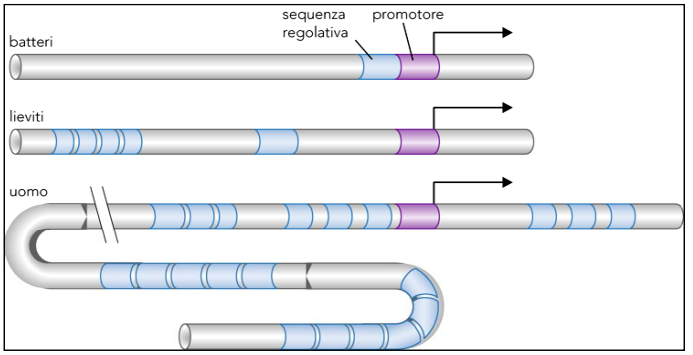



Studies with 3C (chromosome conformation capture) have shown:

- interaction between enhancers and promoters when **proximal** or little distance
- medium long-range interaction between enhancers and promoters to direct developmentally regulated **choice** of genes in clusters (e.g. β-globin)
- intra- and **interchromosomal** long-range interactions between a single enhancer and one out of thousand gene promoters in **mutual exclusive** manner (Odorant receptor choice)
- Inter-chromosomal interaction between different genes **induced** after a stimulatory event (e.g. estrogen-induced genes).

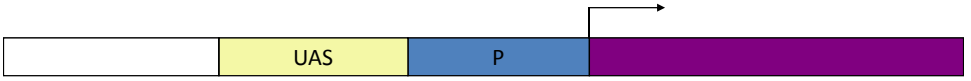
TF binding to specific sequences is the first event in gene activation

Problems:
- Where ?
- How many TFs ?
- Is DNA binding always relevant to gene activation?




L'aumento di complessità dei genomi è parallelo all'aumento delle regioni regolatrici in rapporto alle regioni codificanti

*la situazione più semplice è:*




sebbene vi siano diversi tipi di promotore (es. TATA-dep, Inr-dep, CpG, etc) il promotore minimo è:
1) estremamente inefficiente *in vivo*
2) non regolato

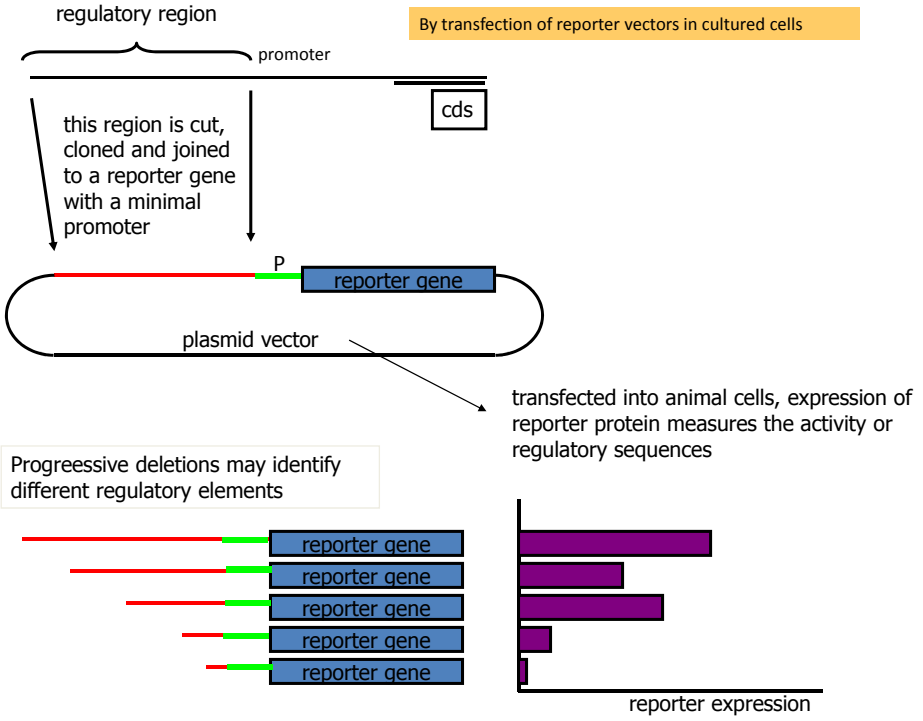
ed è prevalentemente un **elemento di posizione**

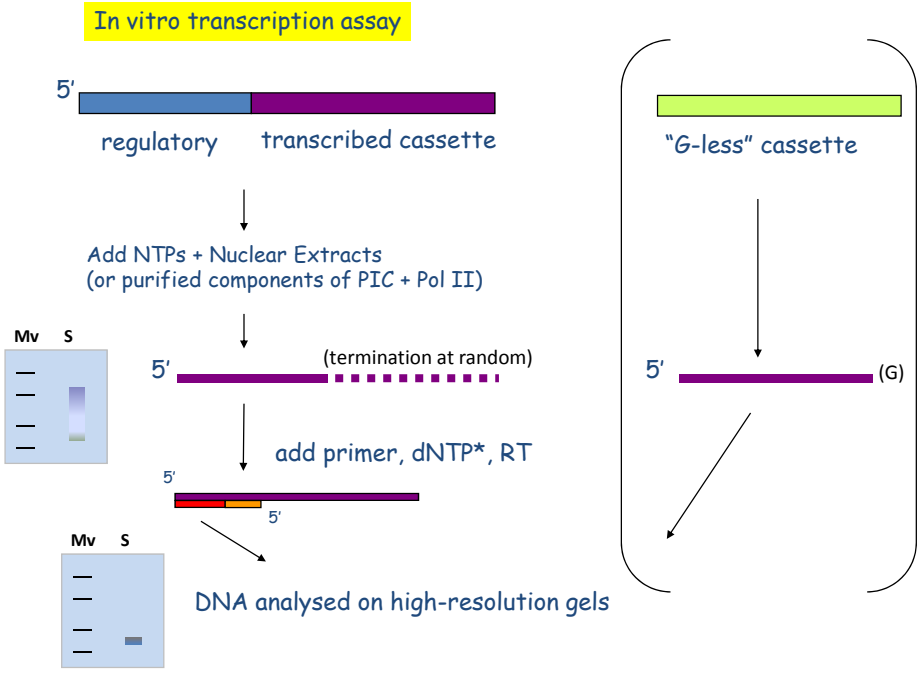
Biochemical definition:

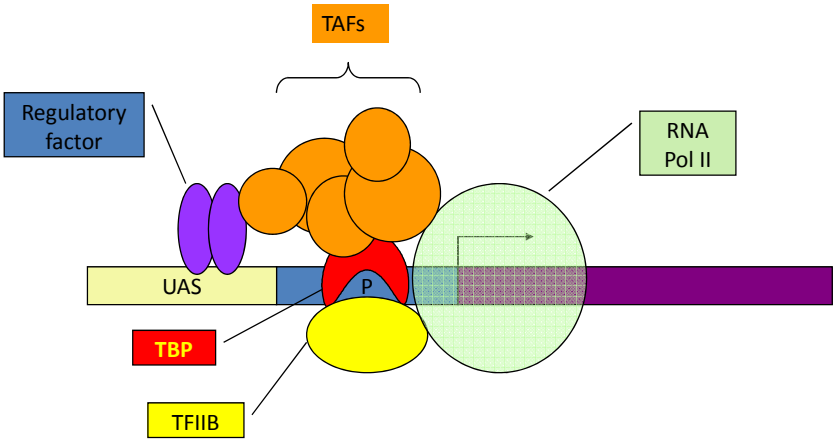
"trans-activation" means activation of transcription by "*trans*" factor (essentially, proteins that recognize DNA or the gene context):

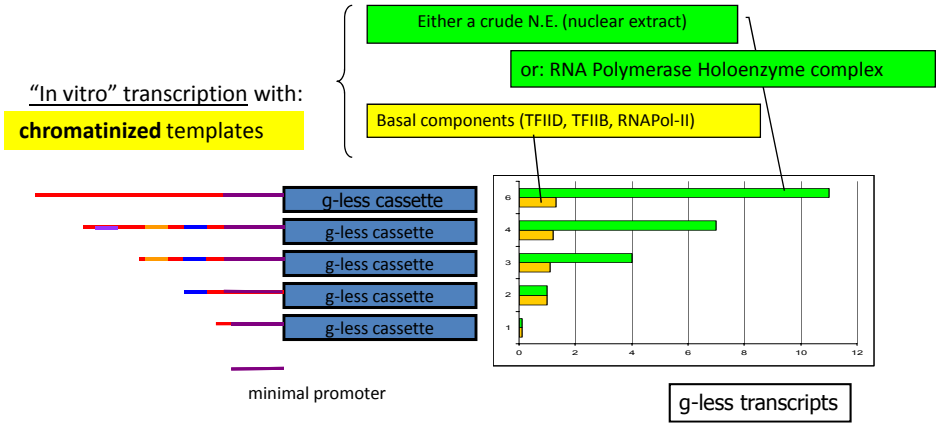
**improvement of the efficiency of a basal promoter**

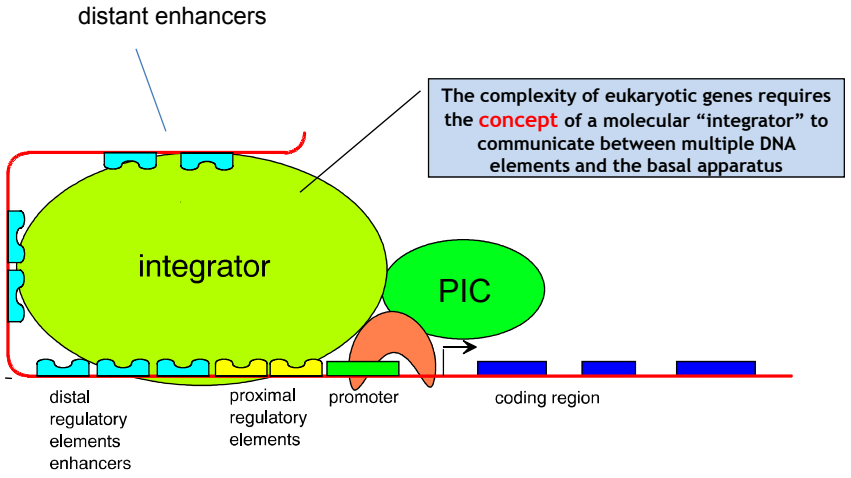
Very important how it is defined experimentally:
1. "in vivo" transcription with reporter vectors
2. "in vitro" transcription with DNA templates


regulatory region
promoter
cds
this region is cut, cloned and joined to a reporter gene with a minimal promoter
P
reporter gene
plasmid vector
By transfection of reporter vectors in cultured cells
transfected into animal cells, expression of reporter protein measures the activity or regulatory sequences
Progreessive deletions may identify different regulatory elements
reporter gene
reporter gene
reporter gene
reporter gene
reporter gene
reporter expression


In vitro transcription assay
5′
regulatory
transcribed cassette
Add NTPs + Nuclear Extracts
(or purified components of PIC + Pol II)
Mv
S
5′
(termination at random)
add primer, dNTP*, RT
5′
5′
Mv
S
DNA analysed on high-resolution gels
"G-less" cassette
5′
(G)


TAFs
Regulatory factor
RNA Pol II
UAS
P
TBP
TFIIB


"In vitro" transcription with:
chromatinized templates
Either a crude N.E. (nuclear extract)
or: RNA Polymerase Holoenzyme complex
Basal components (TFIID, TFIIB, RNAPol-II)
g-less cassette
g-less cassette
g-less cassette
g-less cassette
g-less cassette
minimal promoter
g-less transcripts

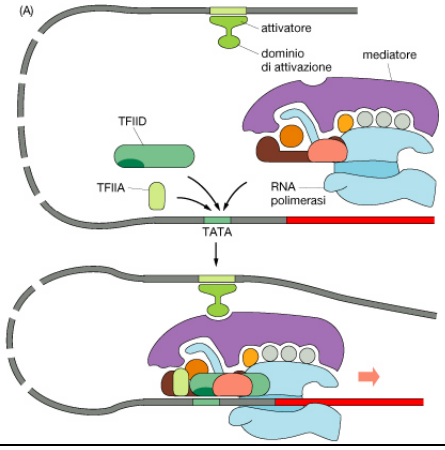



Different approaches used to isolate proteins mediating transactivation yielded a number of multi-component complexes (often **megadalton**).

1) Purification of the RNA Polymerase II holoenzyme
2) Affinity chromatography using antibodies against Transcription Factors (i.e. proteins bound to DNA regulatory elements in the proximal or distant regulatory regions) or "tagged" transcription factors.
3) Affinity chromatography using antibodies against components previously identified (e.g. CBP, GCN5, TAFs, GTFs).

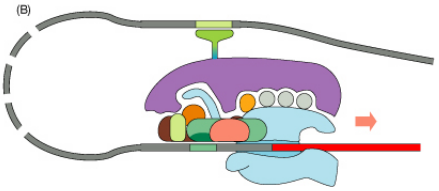
Different complexes were found:

- Complexes with ATP-dep. remodeling activity
- Complexes with Histone Acetyl Transferase activity
- **Complexes with GTF- or RNA PolII-binding activity**
- Mixed activity complexes



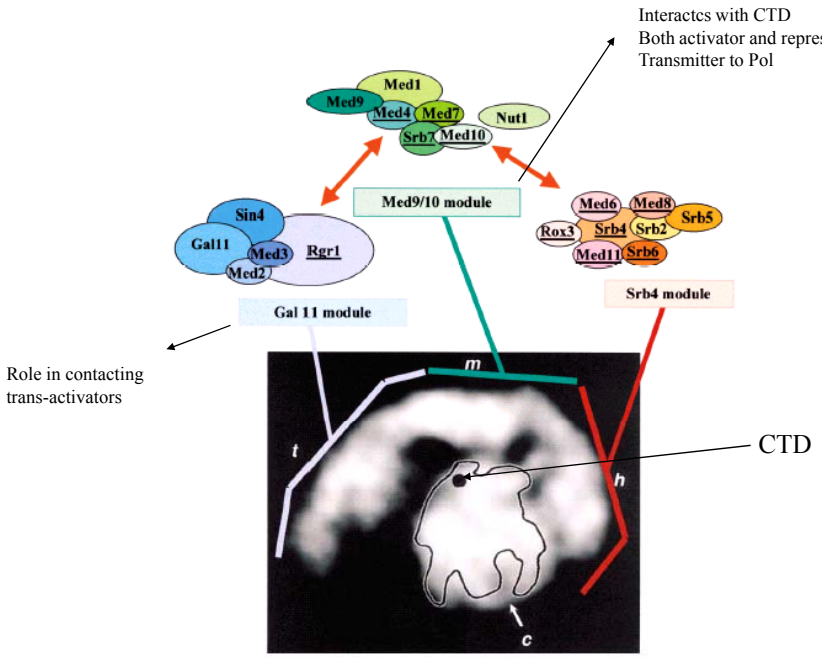

The Mediator, a megadalton complex interacting with TF and PIC components, was isolated ancd characterized in 2004-05.

There are several version of Mediator in the cell nucleus




**Subunit compositions of mediator complexes.** Il complesso del "Mediatore"

| Unified subunit designation | DRIP Mediator-D | ARC Mediator-A | TRAP/ SMCC Mediator-T/S | PC2 Mediator-P | CRSP Mediator-C | NAT Mediator-N | hMediator Mediator-S | Murine Mediator Mediator-M | *S. cerevisiae* | *C. elegans* | *Drosophila* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | CBP/p300 | | | | | | | | | |
| Med240 | DRIP250 | ARC250 | TRAP240 | | | | ND | | | | dTRAP240 |
| Med230 | DRIP240 | ARC240 | TRAP230 | | | p230 | ND | p160a | Nut1 | Sop-1 | dTRAP230 |
| Med220 | DRIP205 | ARC205 | TRAP220 | (TRAP220) | CRSP200 | | ND | p160b | Gal11 | | dTRAP220 |
| Med150 | DRIP150 | ARC150 | TRAP170 | TRAP170 | CRSP150 | p150 | ND | Rgr1/p110 | Rgr1 | | dTRAP170 |
| Med130 | DRIP130 | ARC130 | TRAP150 | TRAP150b | CRSP130 | p140/hSur2 | hSur2 | | | Sur-2 | CG3695 |
| Med105 | | ARC105/ TIG-1 | | | | | | | | | |
| Med100 | DRIP100 | ARC100 | TRAP100 | TRAP100 | | | ND | | Sin4 | | dTRAP100 |
| Med97 | DRIP97 | | TRAP97 | | | p95 | | Ring3/p96a | Srb4 | | |
| Med95 | DRIP92 | ARC92 | TRAP95 | TRAP95 | | p90 | ND | p96b | Med1 | | dTRAP95 |
| | | | TRAP93 | | | | | | | | |
| Med78 | DRIP77 | ARC77 | TRAP80 | TRAP80 | CRSP77 | | ND | p78 | | | dTRAP80 |
| Med70 | DRIP70-2 | ARC70 | | | CRSP70 | p70 | ND | | Med2 | | |
| | | ARC42 | | p37 | | | | p55 | Pgd1/Hrs1 | | |
| Cdk8 | (Cdk8) | (Cdk8) | hSrb10 | | | p56/Cdk8 | Cdk8 | | Srb10 | | dCdk8 |
| Med36 | DRIP36 | ARC36 | | p36 | CRSP34 | p45 | ND | p34 | Med4 | | CG8609 |
| Med34 | DRIP34 | ARC34 | hMed7 | hMed7 | CRSP33 | p37 | Med7 | Med7/p36 | Med7 | ceMed7 | dMed7 |
| | | | | | | p36 | | | Srb5 | | |
| Med33 | DRIP33 | ARC33 | hMed6 | (hMed6) | | p33 | ND | Med6/p32 | Med6 | ceMed6 | CG9473 |
| | | ARC32 | hTRF | hTRF | | | ND | TRF/p28a | Med8 | | |
| Cyclin C | | | hSrb11 | | | p31/ Cyclin C | Cyclin C | | Srb11 | | |
| | | | | | | p30 | | p28b | Rox3 | | |
| | | | | | | p23 | | | Srb2 | | |
| | | | hSoh1 | hSoh1 | | p22 | | | | | |
| | | | | | | p21 | | | Med9/Cse2 | | |
| Med17 | hSrb7 | | hSrb7 | hSrb7 | | p17 | ND | Srb7/p21 | Srb7 | ceSrb7 | CG17397 |
| Med10 | hMed10 | | hNut2 | hNut2 | | p14 | ND | | Med10/Nut2 | ceMed10 | dNut2 |
| | | | | | | | | | Med11 | | |
| | | | | | | | | | Srb6 | | |



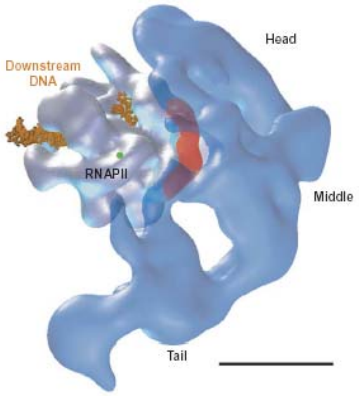




Figure 3. Interaction of Mediator and RNA polymerase II (RNAPII) in the holoenzyme complex. The precise orientation of RNAPII in the holoenzyme complex was established by 2D cross-correlation analysis between holoenzyme and RNAPII projections. The figure shows a cryoelectron microscopy reconstruction of polymerase fitted into the extended Mediator structure in the orientation determined by cross-correlation analysis. Multiple contacts between Mediator and RNAPII are established in the holoenzyme complex, involving mostly the head and middle domains, and distributed around the Rpb3–Rpb11 polymerase subunits (highlighted in red). The small green circle indicates the point where the carboxyterminal domain of Rpb1 (the largest polymerase subunit), crucial for Mediator polymerase interaction, emanates from the surface of the enzyme. The bacterial homolog of the Rpb3–Rpb11 complex, the a2 homodimer, is involved in transcription regulation in bacteria, suggesting a conservation between prokaryotes and eukaryotes of the RNA polymerase surface involved in regulation. The scale bar represents 100 A° .

Different forms of Mediator exist and Mediator conformation is dependent on the kind of **activator** is bound

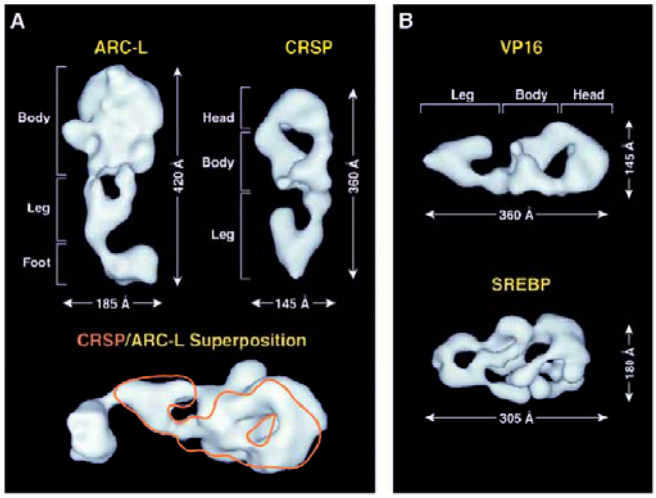



**Fig. 2.** Conformations of the mammalian mediator complexes. (A) EM composites of the ARC-L and CRSP complexes, which illustrate the size and structural differences between the two. (B) EM composites showing the distinctly different structural conformations adopted by CRSP when isolated via affinity interactions with either the VP16 or SREBP activator. EM composites were generously provided by Dylan Taatjes and Bob Tjian (Naar et al., 2002; Taatjes et al., 2002).

Regulatory Transcription Factors
(DNA binding proteins)

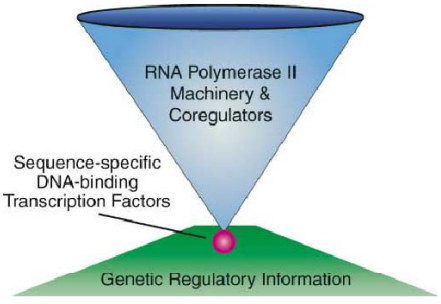
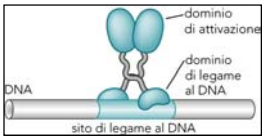



Figure 1. Sequence-Specific DNA Binding Transcription Factors Interpret and Transmit Genetic Regulatory Information

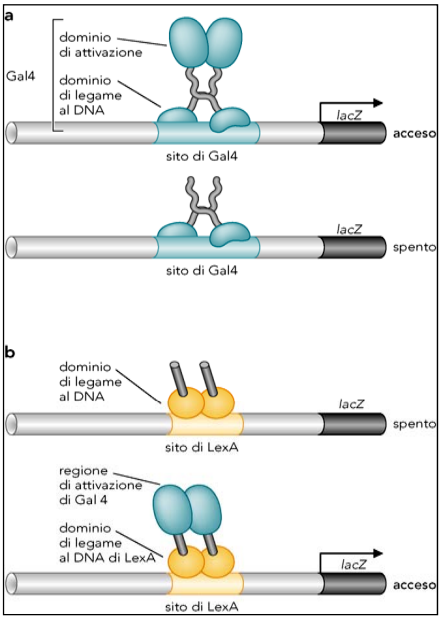
In this diagram, sequence-specific factors are depicted as the apex at the interface of the vast array of genetic regulatory information and the inverted cone of the RNA polymerase II transcriptional machinery and coregulators.

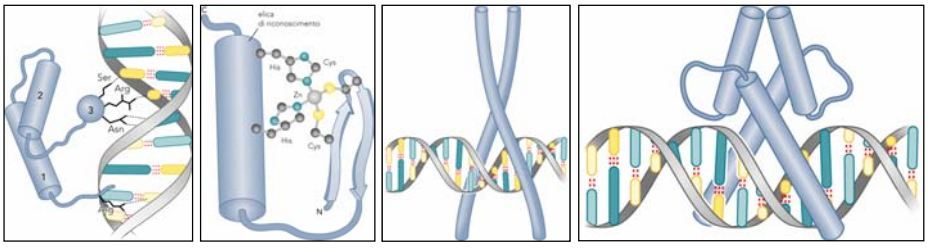
the domain swap experiment




Negli eucarioti, nei fattori trascrizionali di regolazione, il dominio di legame al DNA è necessario per guidare il dominio di attivazione (trans-activating domain) (o di repressione) in posizione produttiva rispetto al promotore del gene.

modularità



Examples:

| | | | |
|---|---|---|---|
| Hox proteins | GAL4 | GCN4 | Myc |
| Antennapedia | Steroid receptors | fos-jun (AP-1) | Myo-D, Neuro-D |
| Matα | Nuclear receptors | CREB | SREBP |

**DNA-binding domains** (as well as dimerization domains, which are very often closely associated in transcription factors, display quite rigid 3D structures.

In sharp contrast, **transactivating domains** have never been resolved by cristallography, i.e. they are flexible and adaptable domains, which most likely assume different conformations, depending on interactions.

Trans-activating domain classification is rather based on aminoacid composition, i.e.:

- acidic
- glutamine-rich
- glutamine/proline rich
- hydrophobic

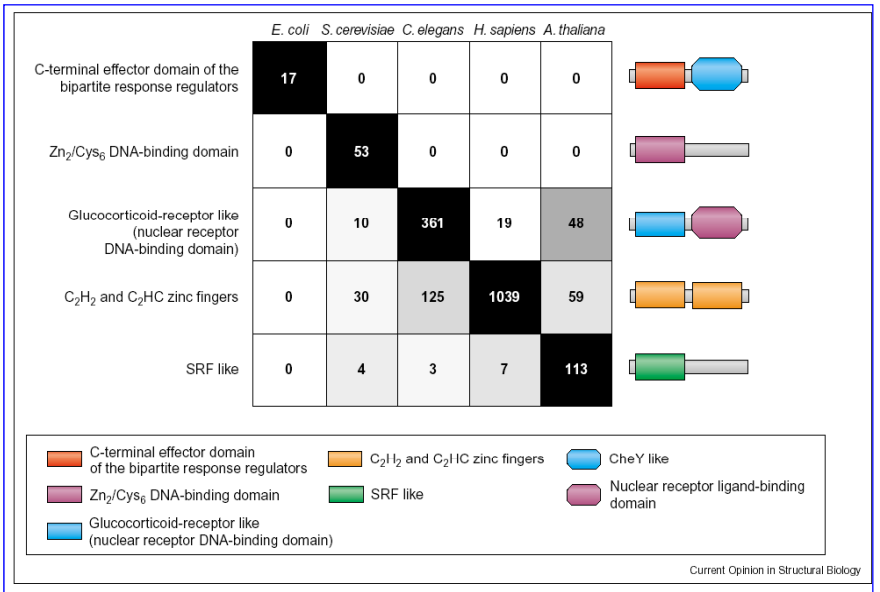
DNA-binding factors

**Table 1**

**Numbers of DNA-binding transcription factors in five organisms[a].**

| Organism | Number of transcripts | Number of proteins with DNA-binding domains | Percentage of transcripts containing DNA-binding domains |
|---|---|---|---|
| *E. coli* | 4280 | 267 | 6.2 |
| *S. cerevisiae* | 6357 | 245 | 3.9 |
| *C. elegans* | 31 677 | 1463 | 4.6 |
| *H. sapiens* | 32 036[b] | 2604 | 8.1 |
| *A. thaliana* | 28 787 | 1667 | 5.7 |

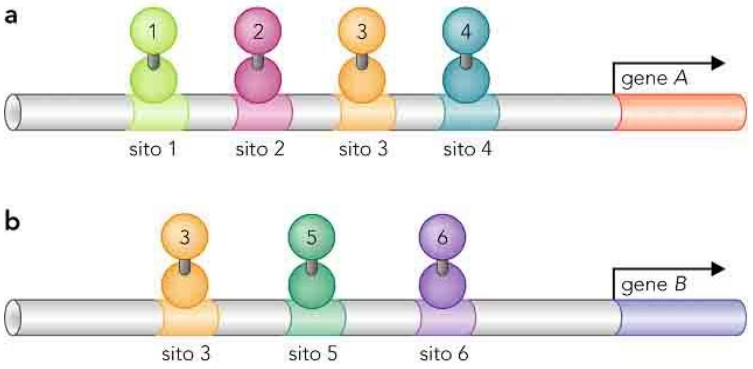
[a]DNA-binding domain assignments from Pfam and SUPERFAMILY are used to establish the repertoire of DNA-binding transcription factors in five model organisms. An expectation value threshold of 0.002 was used in making the assignments. Co-regulators that do not bind DNA directly are excluded. [b]Predicted by Ensembl v19.34a [42].



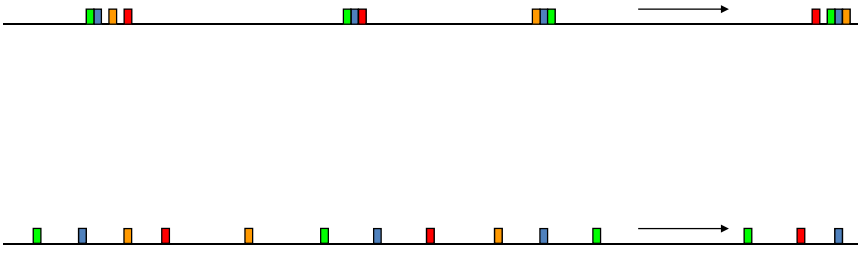

Lineage-specific expansion of DNA-binding domain families. Examples of DNA-binding domain families of transcription factors that are prevalent in one of the five genomes, but are rare in the others. The genomic occurrence of each family is provided in the table and we depict their most common domain architectures alongside. SRF, serum response factor.

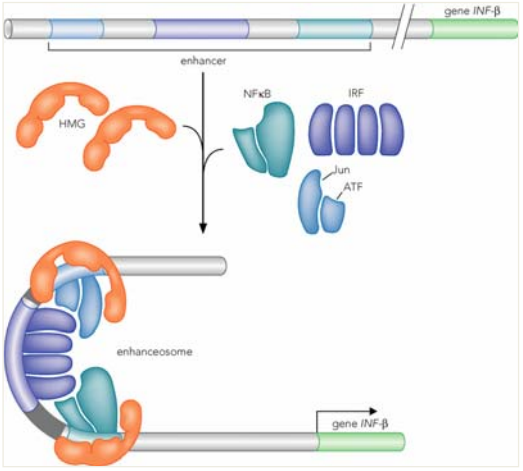
Il controllo trascrizionale viene realizzato con un numero limitato (ancorchè assai grande) di fattori trascrizionali leganti il DNA.

Ogni regione di controllo è formata dalla giustapposizione di diversi elementi in un ordine spaziale specifico.




La cooperazione e composizionalità delle sequenze regolatrici è suggerita anche dal fatto che gli elementi *cis* sono distibuiti in *clusters* e non in modo uniforme.

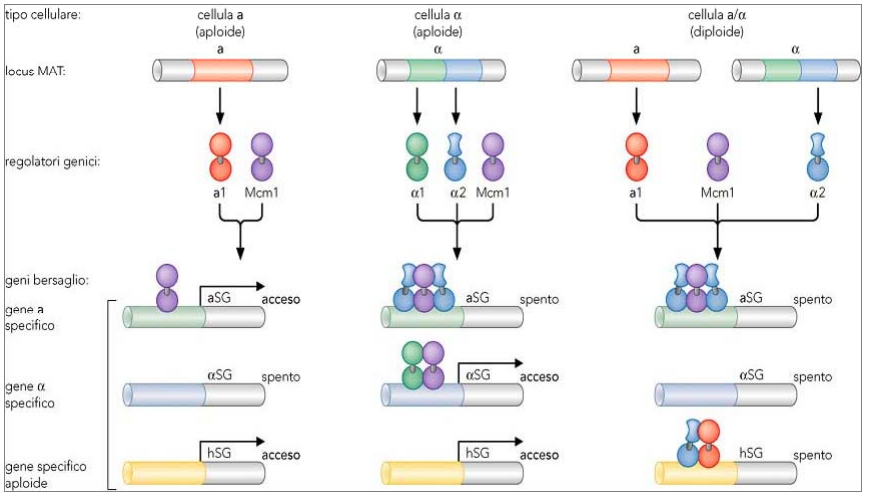





The INF-β “enhanceosome”

In certain cases, the binding of multiple different proteins to adjacent sites in an enhancer is required to make it working. HMG are DNA-binding proteins with no transactivating domain, but display “architectural” functions, e.g. bending the DNA

As clearly esemplified in the case of the MAT locus-encoded transcription factors a1, α1, α2, the activatory or repressory result is given by the **combination** and **geometry** of interaction between different factors, on composed DNA elements.



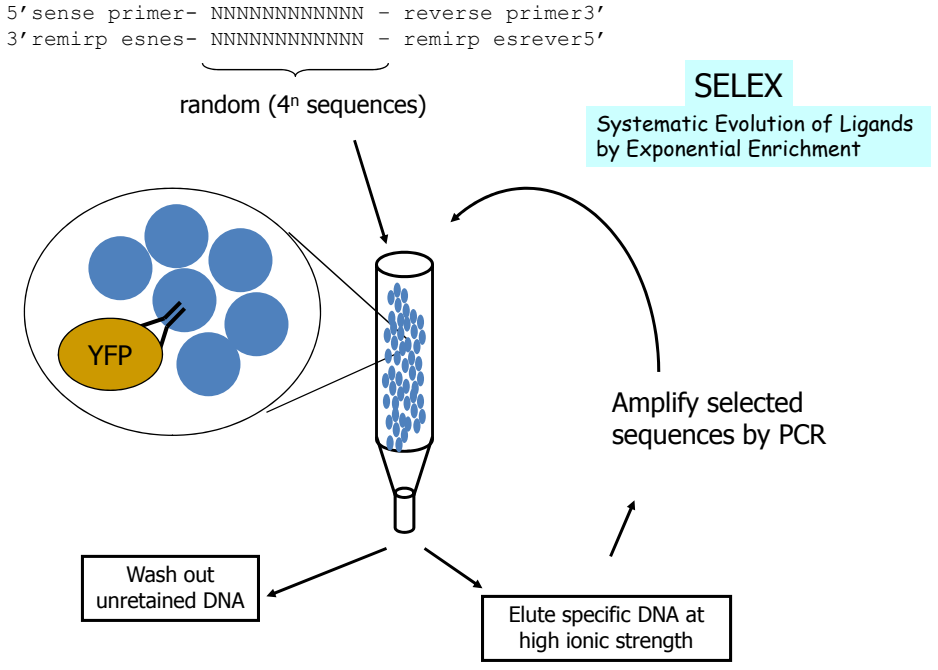
Problems in defining the “response element”

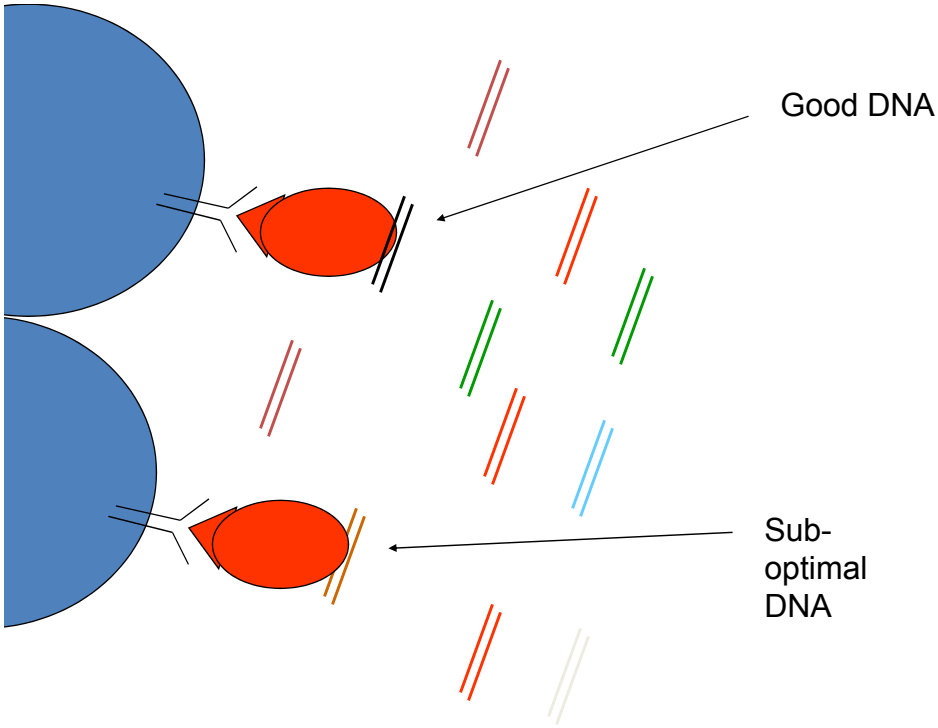
Bioinformatic analysis of binding sites is possible, but does not indicate the “in vivo” situation, or only very partially

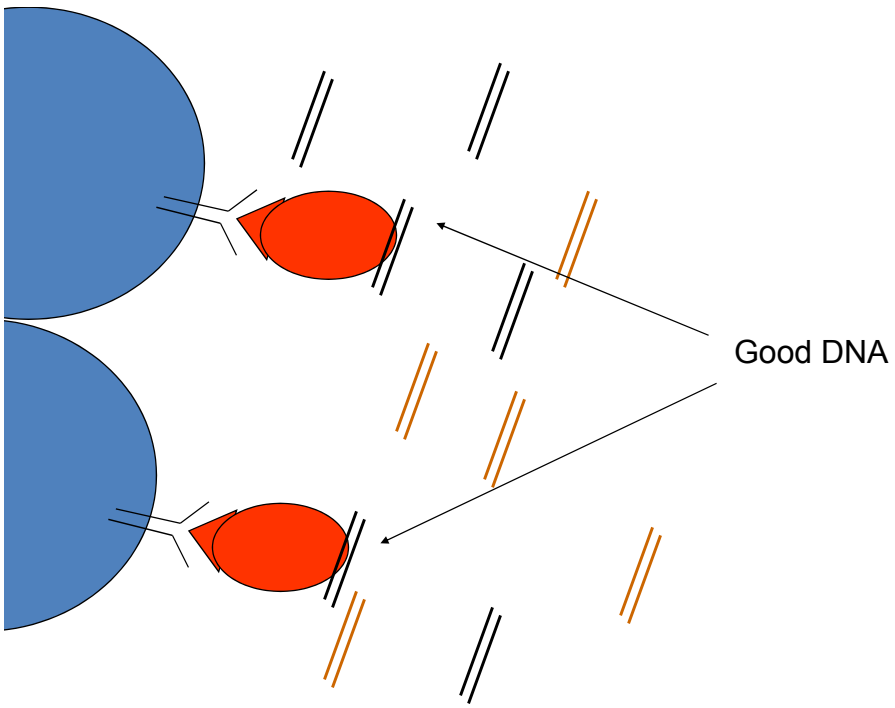
When a DNA binding protein is under study, the sequence of DNA it interacts with can be selected using a process called:

SELEX

Systematic Evolution of Ligands by Exponential Enrichment




Good DNA
Sub-
optimal
DNA


Good DNA

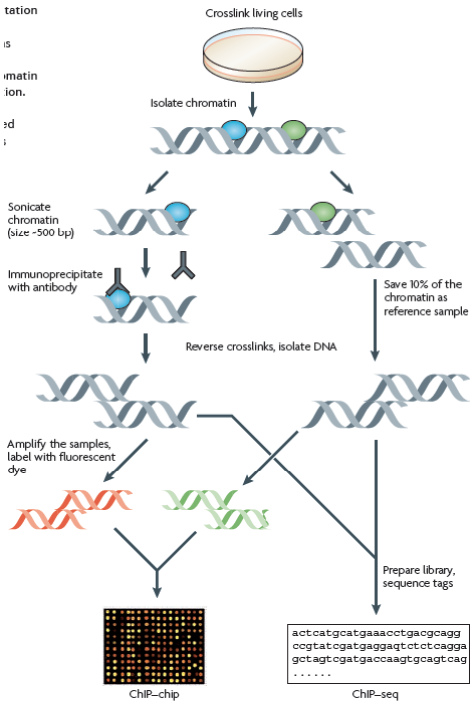
# Insights from genomic profiling of transcription factors

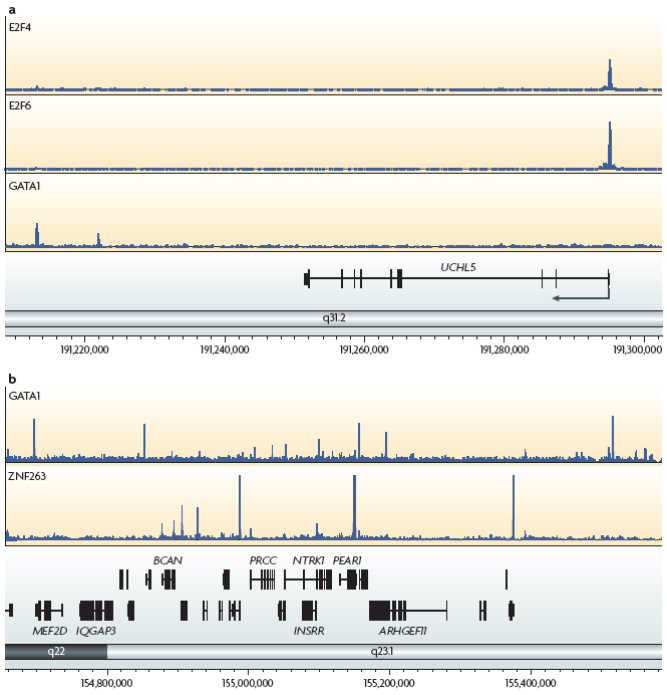
*Peggy J. Farnham*

Abstract | A crucial question in the field of gene regulation is whether the location at which a transcription factor binds influences its effectiveness or the mechanism by which it regulates transcription. Comprehensive transcription factor binding maps are needed to address these issues, and genome-wide mapping is now possible thanks to the technological advances of ChIP–chip and ChIP–seq. This Review discusses how recent genomic profiling of transcription factors gives insight into how binding specificity is achieved and what features of chromatin influence the ability of transcription factors to interact with the genome. It also suggests future experiments that may further our understanding of the causes and consequences of transcription factor–genome interactions.

REVIEW








Article: this article is only suggeted, not mandatory

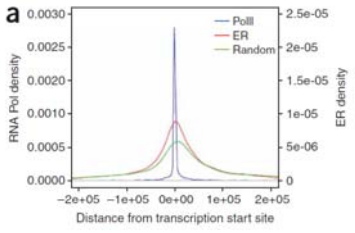
# Genome-wide analysis of estrogen receptor binding sites

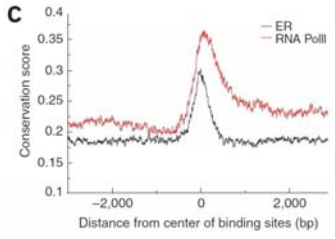
Jason S Carroll[1], Clifford A Meyer[2,3], Jun Song[2,3], Wei Li[2,3], Timothy R Geistlinger[1], Jérôme Eeckhoute[1], Alexander S Brodsky[4], Erika Krasnickas Keeton[1], Kirsten C Fertuck[1], Giles F Hall[5], Qianben Wang[1], Stefan Bekiranov[6,8], Victor Sementchenko[6], Edward A Fox[5], Pamela A Silver[5,7], Thomas R Gingeras[6], X Shirley Liu[2,3] & Myles Brown[1]

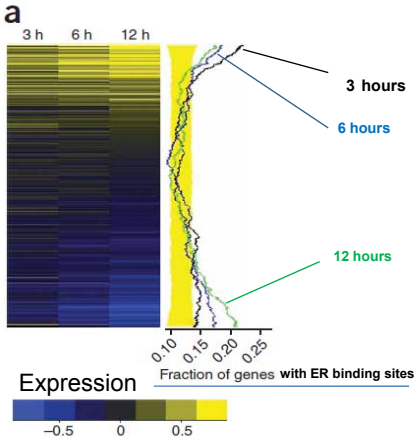
**The estrogen receptor is the master transcriptional regulator of breast cancer phenotype and the archetype of a molecular therapeutic target. We mapped all estrogen receptor and RNA polymerase II binding sites on a genome-wide scale, identifying the authentic *cis* binding sites and target genes, in breast cancer cells. Combining this unique resource with gene expression data demonstrates distinct temporal mechanisms of estrogen-mediated gene regulation, particularly in the case of estrogen-suppressed genes. Furthermore, this resource has allowed the identification of *cis*-regulatory sites in previously unexplored regions of the genome and the cooperating transcription factors underlying estrogen signaling in breast cancer.**

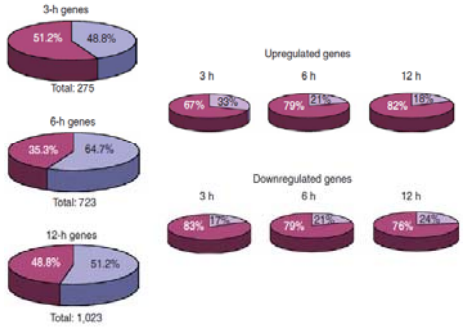
NATURE GENETICS VOLUME 38 | NUMBER 11 | NOVEMBER 2006 1289

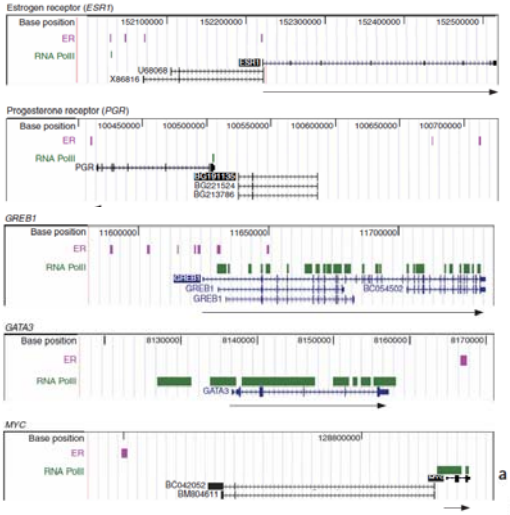
A seminar by Jay Carroll is planned in Torino, February 5.

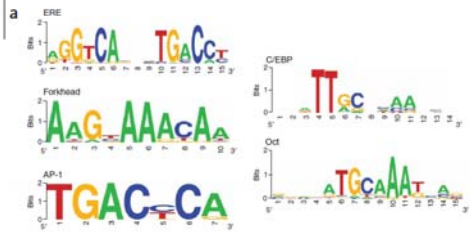

a
PolII
ER
Random
RNA Pol density
ER density
Distance from transcription start site

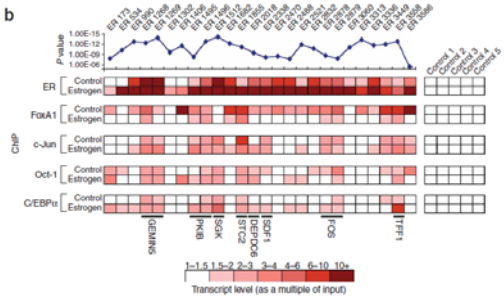

c
ER
RNA PolII
Conservation score
Distance from center of binding sites (bp)


a
3 h
6 h
12 h
3 hours
6 hours
12 hours
Fraction of genes with ER binding sites
Expression

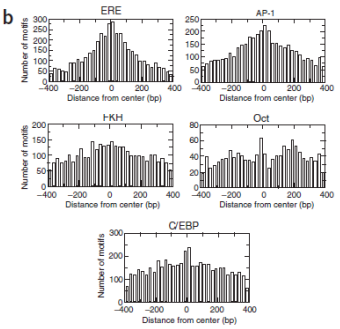

3-h genes
Total: 275
6-h genes
Total: 723
12-h genes
Total: 1,023
Upregulated genes
Downregulated genes
contain ER binding sites
do not contain ER binding sites

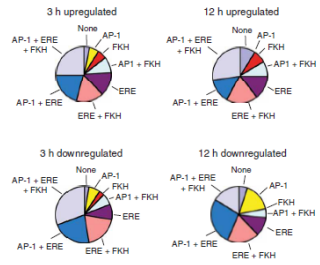

Estrogen receptor (ESR1)
Progesterone receptor (PGR)
GREB1
GATA3
MYC
Base position
ER
RNA PolII


a
ERE
C/EBP
Forkhead
Oct
AP-1




real-time PCR




A clear bias toward AP1 in late downregulated

Corepressors induced by E2, known to bind eityher ER or AP-1: NRIP1

Functional analysis of NRIP1: