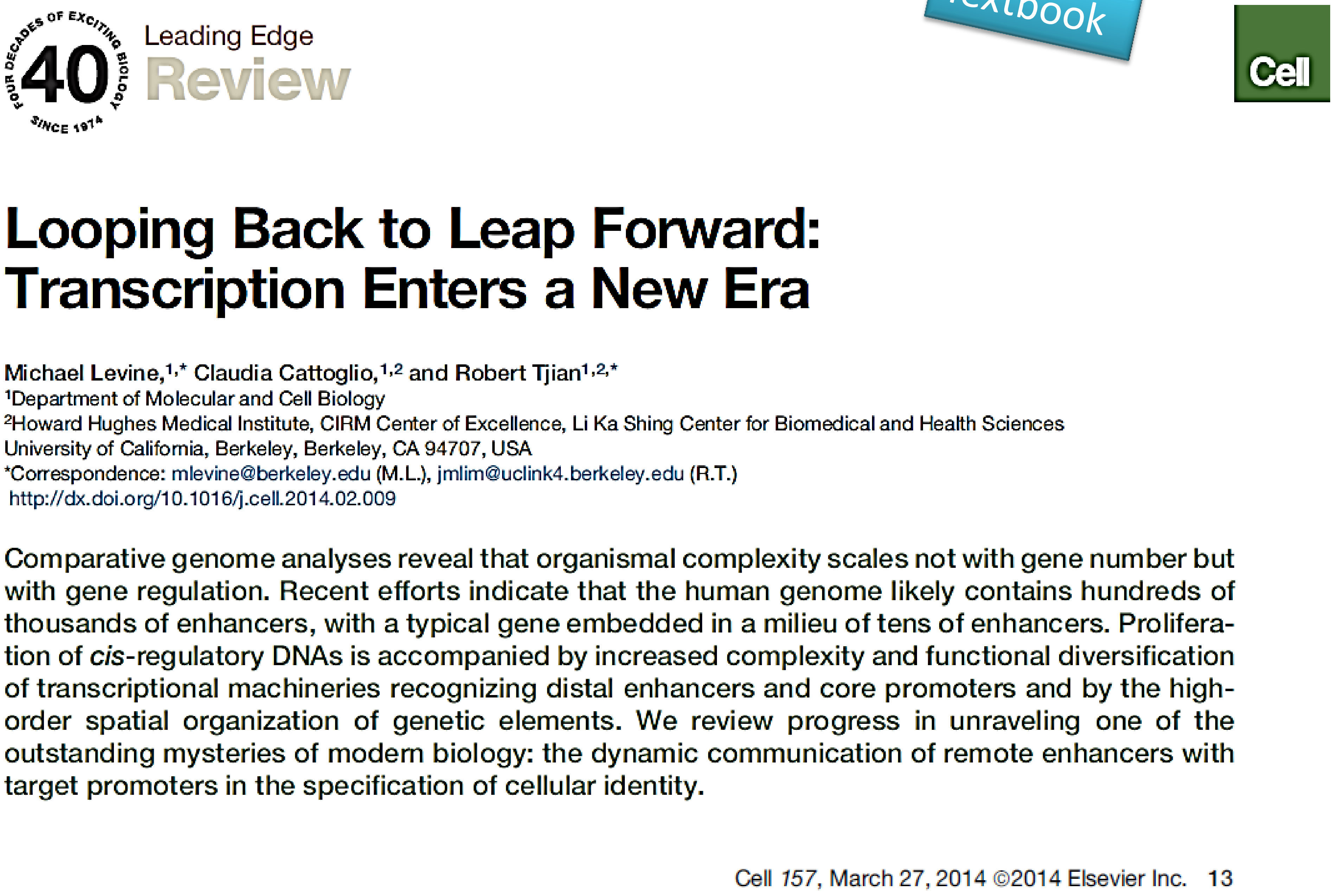
Part 3 – Transcriptional regulation

L3.1

Textbook



FOUR DECADES OF EXCITING BIOLOGY 40 SINCE 1974

Leading Edge
Review

Cell

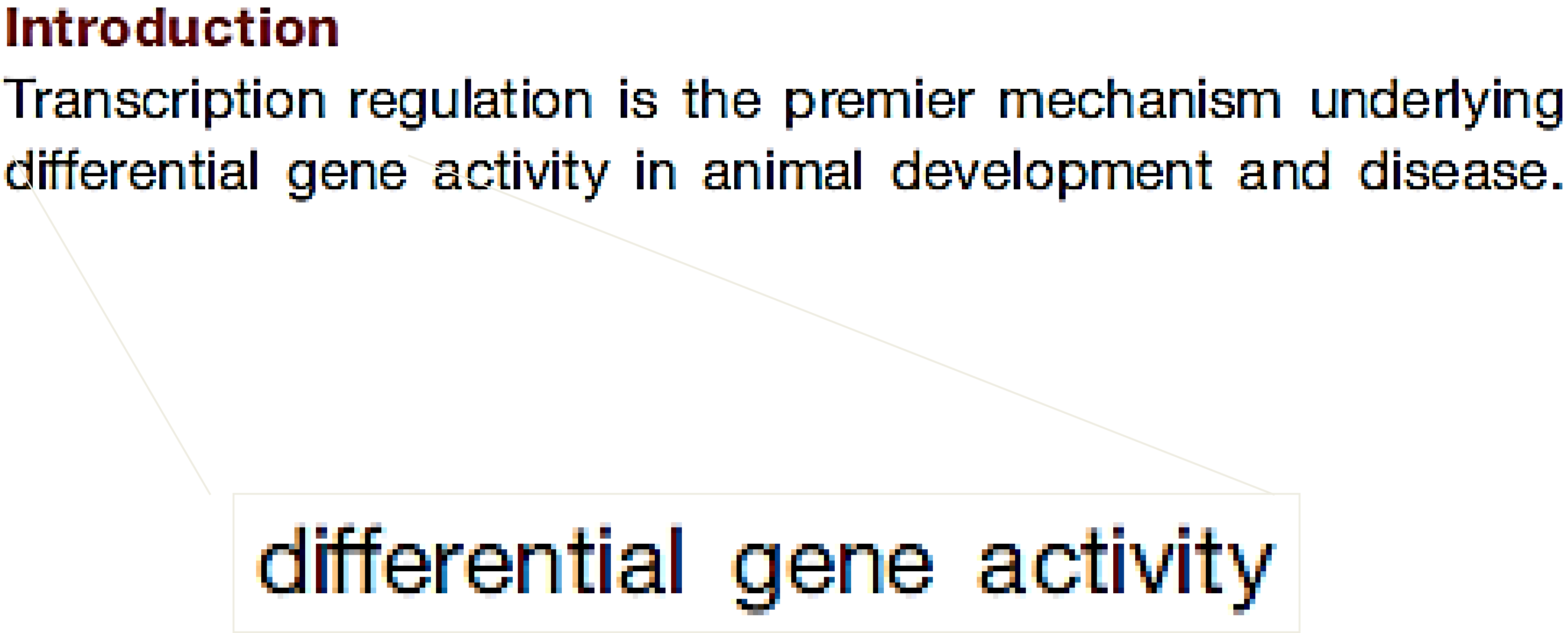
Looping Back to Leap Forward: Transcription Enters a New Era

Michael Levine,[1,*] Claudia Cattoglio,[1,2] and Robert Tjian[1,2,*]
[1]Department of Molecular and Cell Biology
[2]Howard Hughes Medical Institute, CIRM Center of Excellence, Li Ka Shing Center for Biomedical and Health Sciences
University of California, Berkeley, Berkeley, CA 94707, USA
*Correspondence: mlevine@berkeley.edu (M.L.), jmlim@uclink4.berkeley.edu (R.T.)

http://dx.doi.org/10.1016/j.cell.2014.02.009

Comparative genome analyses reveal that organismal complexity scales not with gene number but with gene regulation. Recent efforts indicate that the human genome likely contains hundreds of thousands of enhancers, with a typical gene embedded in a milieu of tens of enhancers. Proliferation of *cis*-regulatory DNAs is accompanied by increased complexity and functional diversification of transcriptional machineries recognizing distal enhancers and core promoters and by the high-order spatial organization of genetic elements. We review progress in unraveling one of the outstanding mysteries of modern biology: the dynamic communication of remote enhancers with target promoters in the specification of cellular identity.

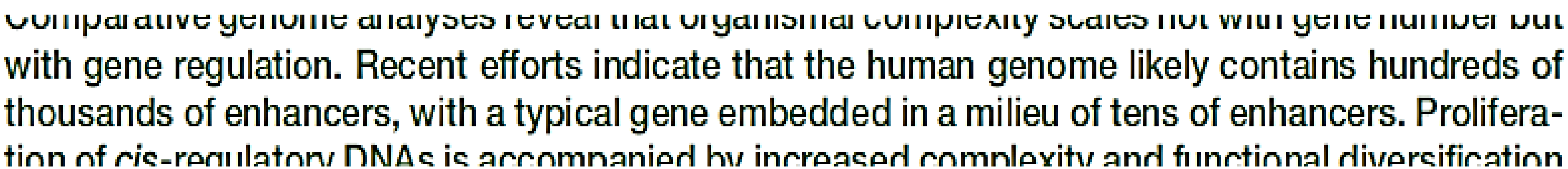
Levine, 2014



Introduction

Transcription regulation is the premier mechanism underlying differential gene activity in animal development and disease.

differential gene activity



Comparative genome analyses reveal that organismal complexity scales not with gene number but with gene regulation. Recent efforts indicate that the human genome likely contains hundreds of thousands of enhancers, with a typical gene embedded in a milieu of tens of enhancers. Proliferation of *cis*-regulatory DNAs is accompanied by increased complexity and functional diversification

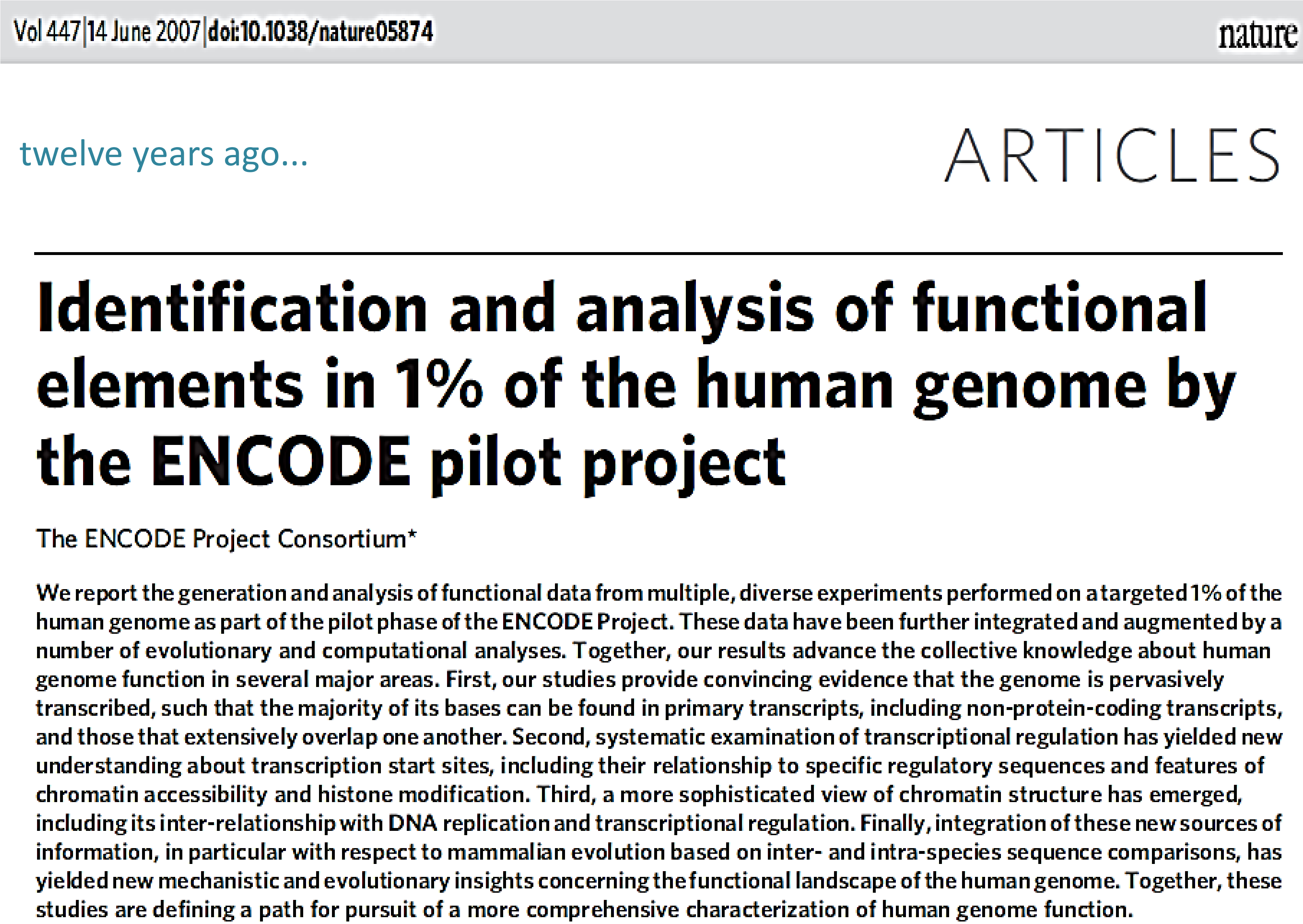
... the human genome likely contains hundreds of thousands of enhancers, with a typical gene embedded in a milieu of tens of enhancers.

This is one of the most impacting results of ENCODE

ENCODE 2007   with traditional sequencing + microarrays
1% of the Human Genome
finding promoters, enhancers, transcripts etc

Scaled rapidly up after NGS

twelve years ago...



Vol 447 | 14 June 2007 | doi:10.1038/nature05874

nature

ARTICLES

Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project

The ENCODE Project Consortium*

We report the generation and analysis of functional data from multiple, diverse experiments performed on a targeted 1% of the human genome as part of the pilot phase of the ENCODE Project. These data have been further integrated and augmented by a number of evolutionary and computational analyses. Together, our results advance the collective knowledge about human genome function in several major areas. First, our studies provide convincing evidence that the genome is pervasively transcribed, such that the majority of its bases can be found in primary transcripts, including non-protein-coding transcripts, and those that extensively overlap one another. Second, systematic examination of transcriptional regulation has yielded new understanding about transcription start sites, including their relationship to specific regulatory sequences and features of chromatin accessibility and histone modification. Third, a more sophisticated view of chromatin structure has emerged, including its inter-relationship with DNA replication and transcriptional regulation. Finally, integration of these new sources of information, in particular with respect to mammalian evolution based on inter- and intra-species sequence comparisons, has yielded new mechanistic and evolutionary insights concerning the functional landscape of the human genome. Together, these studies are defining a path for pursuit of a more comprehensive characterization of human genome function.

ARTICLE

doi:10.1038/nature11247

An integrated encyclopedia of DNA elements in the human genome

The ENCODE Project Consortium*

The human genome encodes the blueprint of life, but the function of the vast majority of its nearly three billion bases is unknown. The Encyclopedia of DNA Elements (ENCODE) project has systematically mapped regions of transcription, transcription factor association, chromatin structure and histone modification. These data enabled us to assign biochemical functions for 80% of the genome, in particular outside of the well-studied protein-coding regions. Many discovered candidate regulatory elements are physically associated with one another and with expressed genes, providing new insights into the mechanisms of gene regulation. The newly identified elements also show a statistical correspondence to sequence variants linked to human disease, and can thereby guide interpretation of this variation. Overall, the project provides new insights into the organization and regulation of our genes and genome, and is an expansive resource of functional annotations for biomedical research.

The human genome sequence provides the underlying code for human biology. Despite intensive study, especially in identifying protein-coding genes, our understanding of the genome is far from complete, particularly with

ENCODE
Encyclopedia of DNA Elements
nature.com/encode

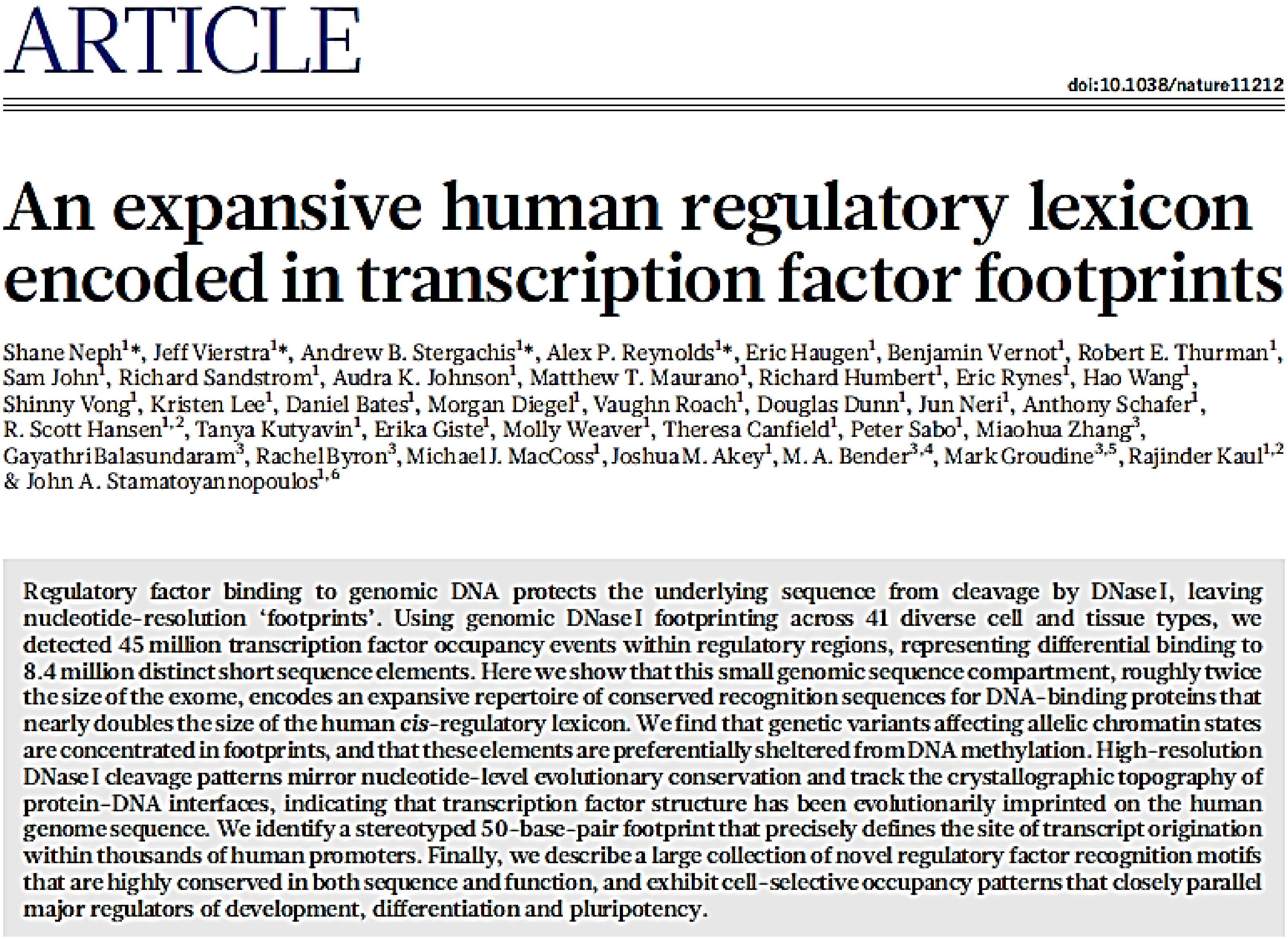
95% of the genome lies within 8 kilobases (kb) of a DNA–protein interaction (as assayed by bound ChIP-seq motifs or DNase I footprints), and 99% is within 1.7 kb of at least one of the biochemical events measured by ENCODE.

in brief...

- ✓ 80.4% of the human genome participate in at least one biochemical event
- ✓ two third of genomic sequences are represented in RNA
- ✓ 400,000 sites have chromatin features of enhancers
- ✓ 70,300 regions have promoter-like features

connect to Encode experiment matrix

https://www.encodeproject.org/matrix/?type=Experiment&status=released



ARTICLE

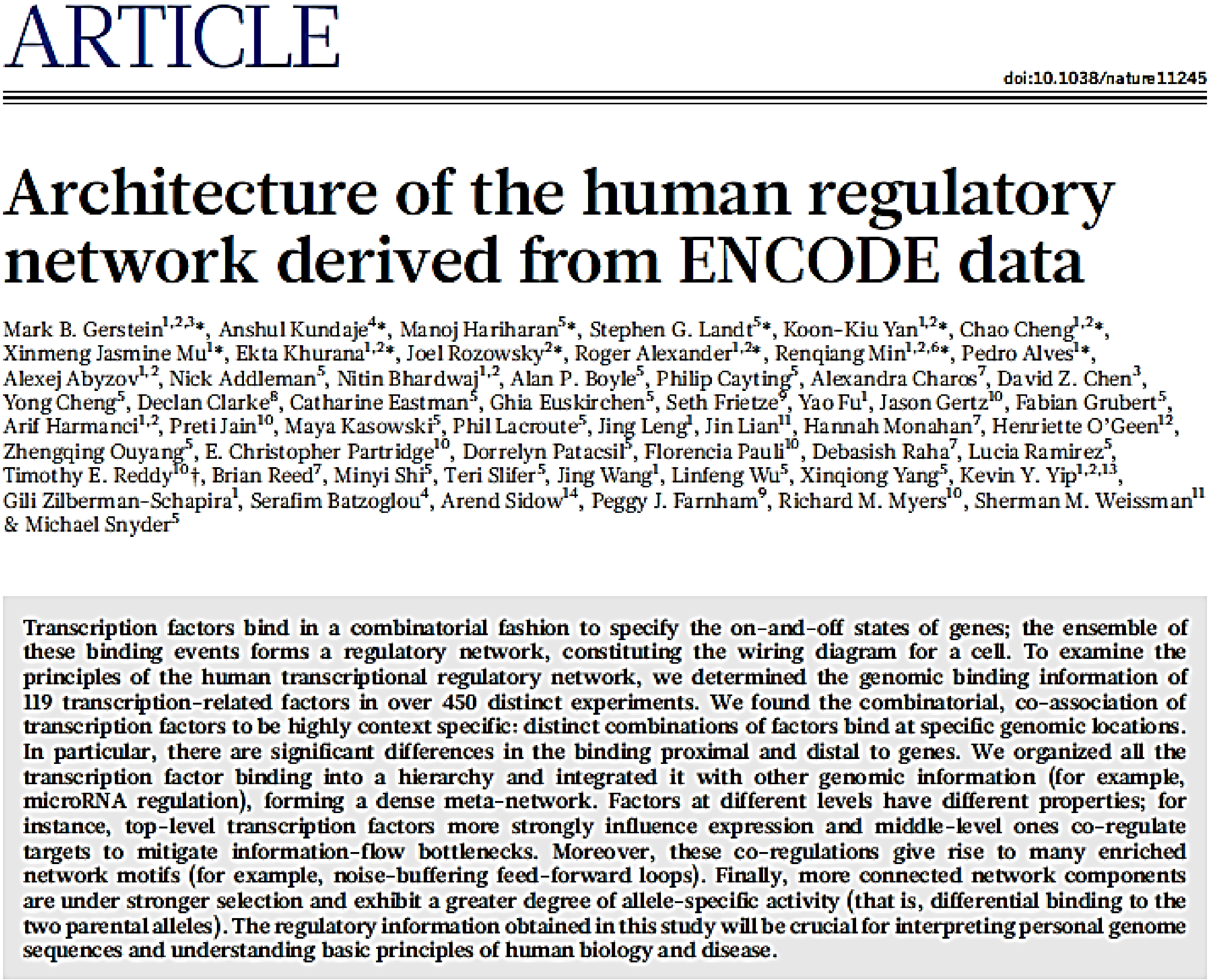
doi:10.1038/nature11212

An expansive human regulatory lexicon encoded in transcription factor footprints

Shane Neph[1]*, Jeff Vierstra[1]*, Andrew B. Stergachis[1]*, Alex P. Reynolds[1]*, Eric Haugen[1], Benjamin Vernot[1], Robert E. Thurman[1], Sam John[1], Richard Sandstrom[1], Audra K. Johnson[1], Matthew T. Maurano[1], Richard Humbert[1], Eric Rynes[1], Hao Wang[1], Shinny Vong[1], Kristen Lee[1], Daniel Bates[1], Morgan Diegel[1], Vaughn Roach[1], Douglas Dunn[1], Jun Neri[1], Anthony Schafer[1], R. Scott Hansen[1,2], Tanya Kutyavin[1], Erika Giste[1], Molly Weaver[1], Theresa Canfield[1], Peter Sabo[1], Miaohua Zhang[3], Gayathri Balasundaram[3], Rachel Byron[3], Michael J. MacCoss[1], Joshua M. Akey[1], M. A. Bender[3,4], Mark Groudine[3,5], Rajinder Kaul[1,2] & John A. Stamatoyannopoulos[1,6]

Regulatory factor binding to genomic DNA protects the underlying sequence from cleavage by DNase I, leaving nucleotide-resolution 'footprints'. Using genomic DNase I footprinting across 41 diverse cell and tissue types, we detected 45 million transcription factor occupancy events within regulatory regions, representing differential binding to 8.4 million distinct short sequence elements. Here we show that this small genomic sequence compartment, roughly twice the size of the exome, encodes an expansive repertoire of conserved recognition sequences for DNA-binding proteins that nearly doubles the size of the human *cis*-regulatory lexicon. We find that genetic variants affecting allelic chromatin states are concentrated in footprints, and that these elements are preferentially sheltered from DNA methylation. High-resolution DNase I cleavage patterns mirror nucleotide-level evolutionary conservation and track the crystallographic topography of protein–DNA interfaces, indicating that transcription factor structure has been evolutionarily imprinted on the human genome sequence. We identify a stereotyped 50-base-pair footprint that precisely defines the site of transcript origination within thousands of human promoters. Finally, we describe a large collection of novel regulatory factor recognition motifs that are highly conserved in both sequence and function, and exhibit cell-selective occupancy patterns that closely parallel major regulators of development, differentiation and pluripotency.

6 SEPTEMBER 2012 | VOL 489 | NATURE | 83



doi:10.1038/nature11245

Architecture of the human regulatory network derived from ENCODE data

Mark B. Gerstein[1,2,3]*, Anshul Kundaje[4]*, Manoj Hariharan[5]*, Stephen G. Landt[5]*, Koon-Kiu Yan[1,2]*, Chao Cheng[1,2]*, Xinmeng Jasmine Mu[1]*, Ekta Khurana[1,2]*, Joel Rozowsky[2]*, Roger Alexander[1,2]*, Renqiang Min[1,2,6]*, Pedro Alves[1]*, Alexej Abyzov[1,2], Nick Addleman[5], Nitin Bhardwaj[1,2], Alan P. Boyle[5], Philip Cayting[5], Alexandra Charos[7], David Z. Chen[3], Yong Cheng[5], Declan Clarke[8], Catharine Eastman[5], Ghia Euskirchen[5], Seth Frietze[9], Yao Fu[1], Jason Gertz[10], Fabian Grubert[5], Arif Harmanci[1,2], Preti Jain[10], Maya Kasowski[5], Phil Lacroute[5], Jing Leng[1], Jin Lian[11], Hannah Monahan[7], Henriette O'Geen[12], Zhengqing Ouyang[5], E. Christopher Partridge[10], Dorrelyn Patacsil[5], Florencia Pauli[10], Debasish Raha[7], Lucia Ramirez[5], Timothy E. Reddy[10]†, Brian Reed[7], Minyi Shi[5], Teri Slifer[5], Jing Wang[1], Linfeng Wu[5], Xinqiong Yang[5], Kevin Y. Yip[1,2,13], Gili Zilberman-Schapira[1], Serafim Batzoglou[4], Arend Sidow[14], Peggy J. Farnham[9], Richard M. Myers[10], Sherman M. Weissman[11] & Michael Snyder[5]

Transcription factors bind in a combinatorial fashion to specify the on-and-off states of genes; the ensemble of these binding events forms a regulatory network, constituting the wiring diagram for a cell. To examine the principles of the human transcriptional regulatory network, we determined the genomic binding information of 119 transcription-related factors in over 450 distinct experiments. We found the combinatorial, co-association of transcription factors to be highly context specific: distinct combinations of factors bind at specific genomic locations. In particular, there are significant differences in the binding proximal and distal to genes. We organized all the transcription factor binding into a hierarchy and integrated it with other genomic information (for example, microRNA regulation), forming a dense meta-network. Factors at different levels have different properties; for instance, top-level transcription factors more strongly influence expression and middle-level ones co-regulate targets to mitigate information-flow bottlenecks. Moreover, these co-regulations give rise to many enriched network motifs (for example, noise-buffering feed-forward loops). Finally, more connected network components are under stronger selection and exhibit a greater degree of allele-specific activity (that is, differential binding to the two parental alleles). The regulatory information obtained in this study will be crucial for interpreting personal genome sequences and understanding basic principles of human biology and disease.

6 SEPTEMBER 2012 | VOL 489 | NATURE | 91

Regulatory regions

Promoters *versus* Enhancers

the main feature is that promoters are always in the region immediately preceding and overlapping the Transcriptional Start Site (TSS)

while

Enhancers are placed in virtualli indifferent regions around the gene, i.e. up to 100,000 bp upstream or downstream, in introns, with aparent no deal of distance with function.

Definitions:

Promoter = the minimal sequence sustaining transcription and correct initiation, usually 50-150 bp 5’-upstream TSS

Upstream regulatory sequence: sequences 5’-adjacent to promoter that regulate promoter utilization (500-2,000 bp, usually a downstream part to +100 is also included). Also sometimes indicated as «UAS», «proximal regulatory element» or «proximal enhancer».

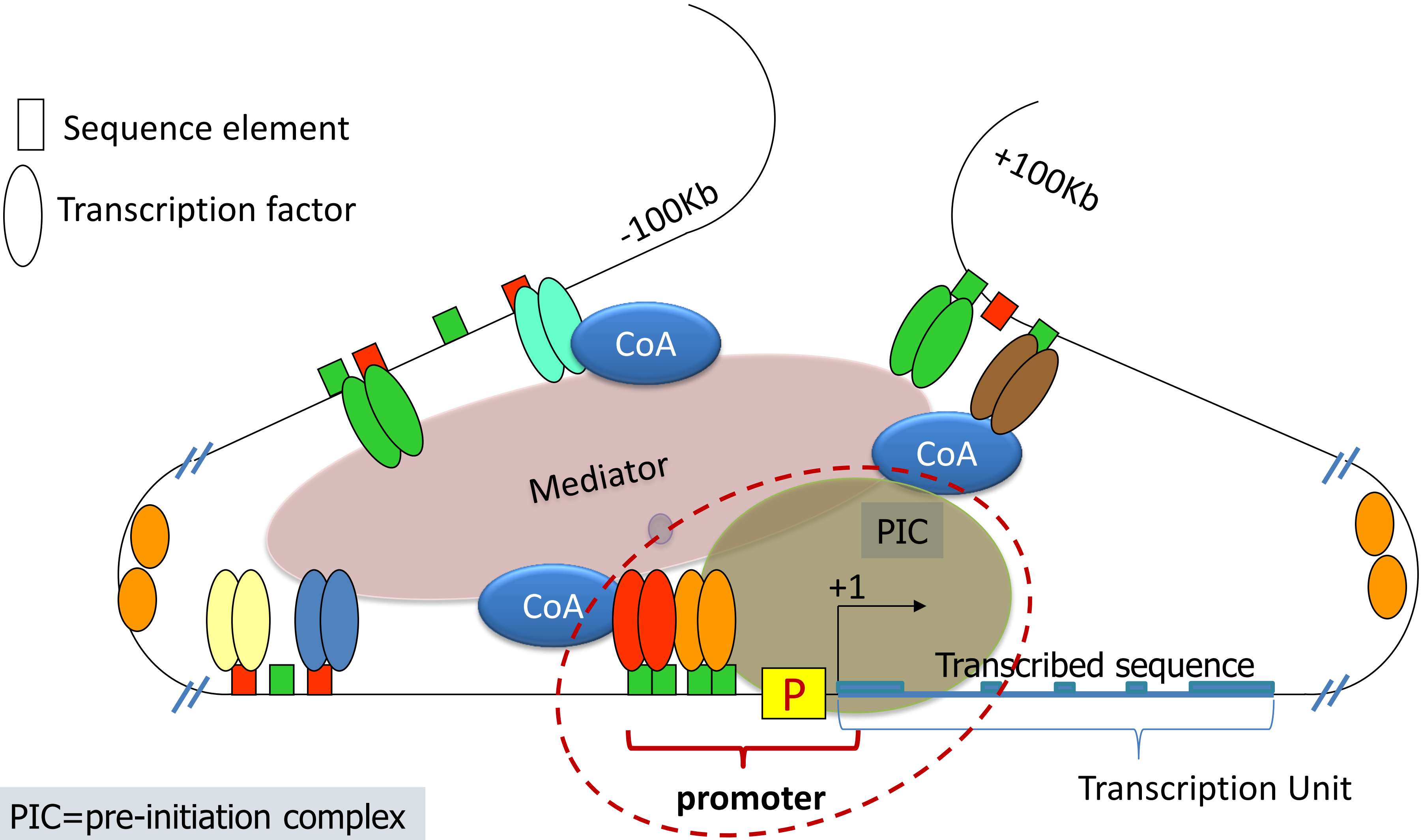
Enhancers: regulatory sequences or «modules» laying virtually at any distance and position from the regulated («cognate») TSS or promoter. Note: even though «enhancer» means «something that increases», enhancers may display repressing activity.

Minimal or «core» promoters are defined as the region bound by General Transcription Factors and RNA Polymerase, that is roughly -40 to +40 bp in respect to TSS.

Essentially, it is the region footprinted by RNA PolII. and GTFs.

Normally however the promoter is accompanied by a proximal regulatory region, that Aa place somewhere at -1,000 to +100 bp respect the TSS.

Schematics of eukaryotic gene regulatory sequences and proteins



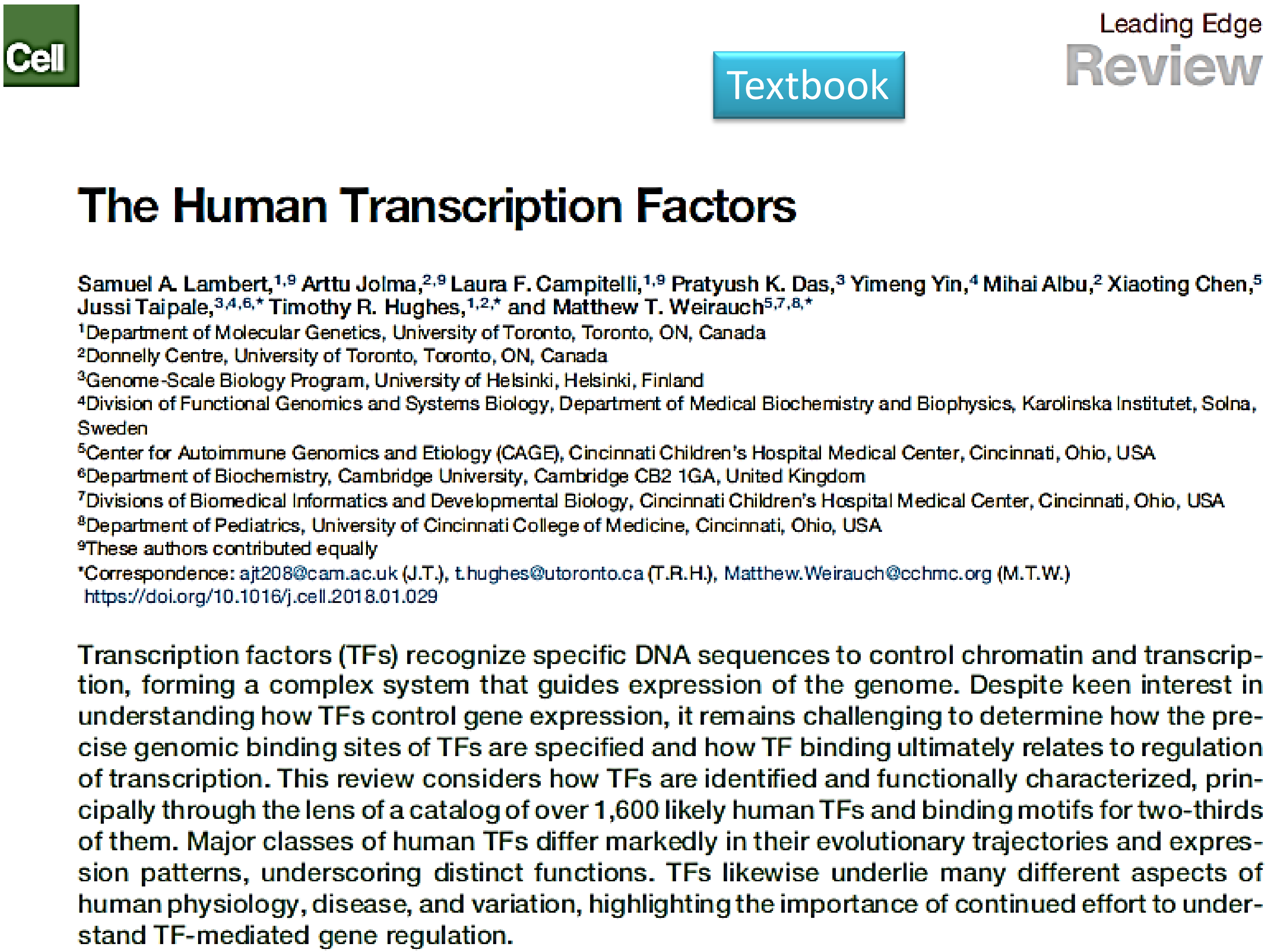
PIC=pre-initiation complex

Regulatory modules (enhancers, proximal regulatory elements, etc.)

DNA segments where short sequence motifs, 4 to 15 base long, called Response Elements and recognized by **Transcription Factors** are juxtapposed.

Response Elements = **TFBS** (Transcription Factor Binding Sites)

Textbook



Cell

Leading Edge
Review

The Human Transcription Factors

Samuel A. Lambert,[1,9] Arttu Jolma,[2,9] Laura F. Campitelli,[1,9] Pratyush K. Das,[3] Yimeng Yin,[4] Mihai Albu,[2] Xiaoting Chen,[5] Jussi Taipale,[3,4,6,*] Timothy R. Hughes,[1,2,*] and Matthew T. Weirauch[5,7,8,*]

[1]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
[2]Donnelly Centre, University of Toronto, Toronto, ON, Canada
[3]Genome-Scale Biology Program, University of Helsinki, Helsinki, Finland
[4]Division of Functional Genomics and Systems Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Solna, Sweden
[5]Center for Autoimmune Genomics and Etiology (CAGE), Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
[6]Department of Biochemistry, Cambridge University, Cambridge CB2 1GA, United Kingdom
[7]Divisions of Biomedical Informatics and Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
[8]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA
[9]These authors contributed equally
*Correspondence: ajt208@cam.ac.uk (J.T.), t.hughes@utoronto.ca (T.R.H.), Matthew.Weirauch@cchmc.org (M.T.W.)

https://doi.org/10.1016/j.cell.2018.01.029

Transcription factors (TFs) recognize specific DNA sequences to control chromatin and transcription, forming a complex system that guides expression of the genome. Despite keen interest in understanding how TFs control gene expression, it remains challenging to determine how the precise genomic binding sites of TFs are specified and how TF binding ultimately relates to regulation of transcription. This review considers how TFs are identified and functionally characterized, principally through the lens of a catalog of over 1,600 likely human TFs and binding motifs for two-thirds of them. Major classes of human TFs differ markedly in their evolutionary trajectories and expression patterns, underscoring distinct functions. TFs likewise underlie many different aspects of human physiology, disease, and variation, highlighting the importance of continued effort to understand TF-mediated gene regulation.

Transcription Factors

>10% of the coding potential (2,000-3,000) of the Human Genome

TFs recognize DNA motifs by multiple chemical interactions between aa residues of their DBD (DNA binding domain) and 4-6 bp in the major groove

TFBS for a specific TF can vary a bit → consensus sequence

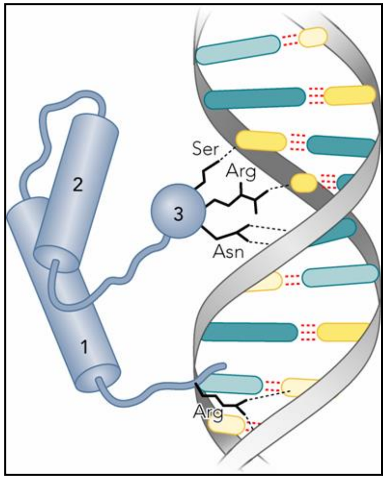
TFs contain one or more transactivating domains (TA) that interact with co-regulatory proteins (Mediator, ATP-dep. remodelers, TFIID, HAT/HDAC, other)

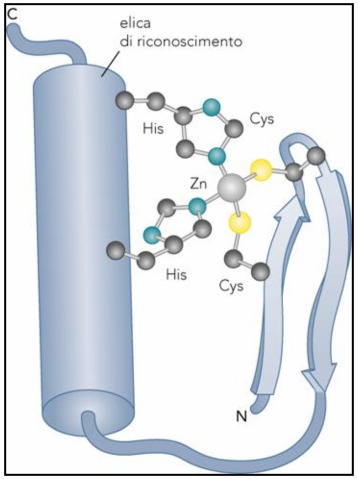
How TFs recognize DNA sequences

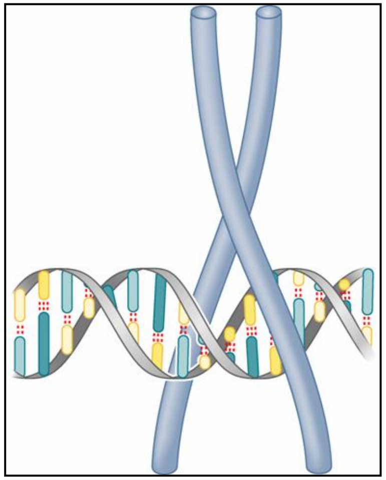
TF database: http://www.transcriptionfactor.org/index.cgi?Home

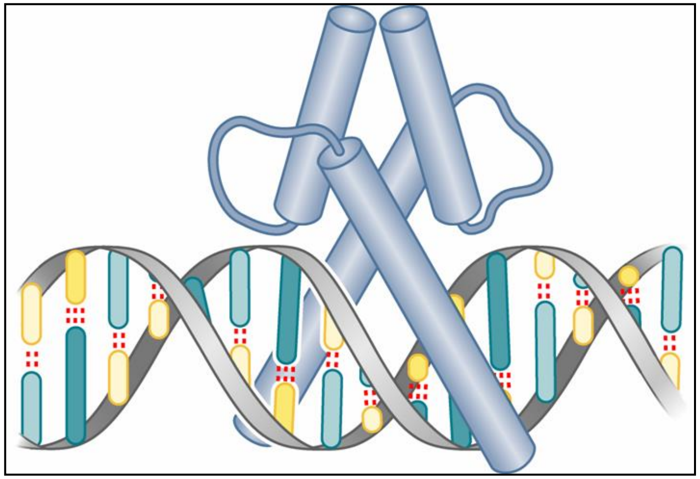
Structures (DBD) http://www.rcsb.org/pdb/home/home.do

Yearly NAR database issues









Examples:

Hox proteins

Antennapedia

Matα

GAL4

Steroid receptors

Nuclear receptors

GCN4

fos-jun (AP-1)

CREB

Myc

Myo-D, Neuro-D

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 Transcription Factors (regulatory factors) (TF)

GO cathegory

Do not include:

- Coactivators or corepressors
- Enzymes
- General Transcription Factors (basal PIC components)

Do include:

- Putative proteins with similitude to known TFs
- Proteins that possess structural domains similar to DNA Binding domain (DBD) of known TFs

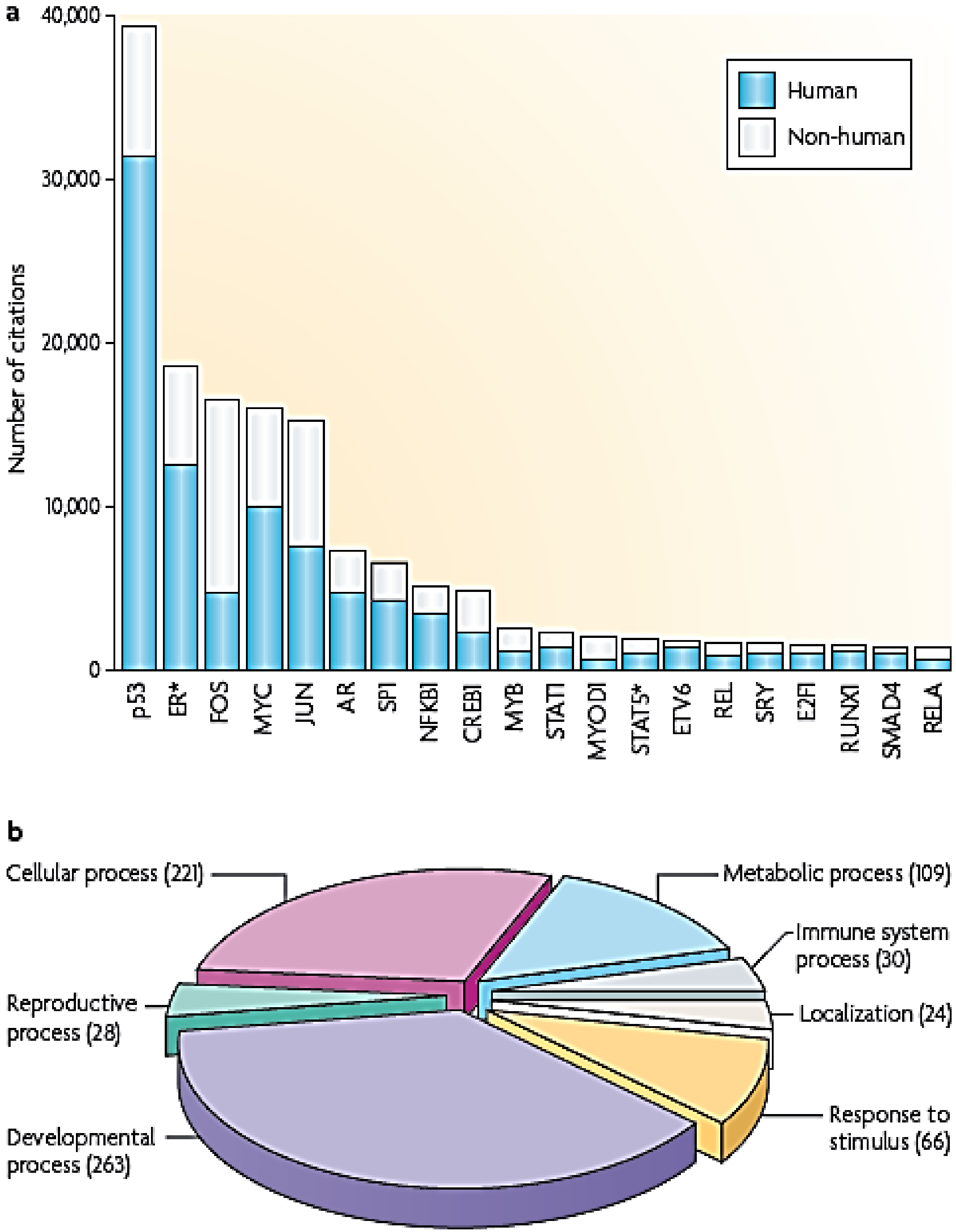
Gene Ontology

Figure 1

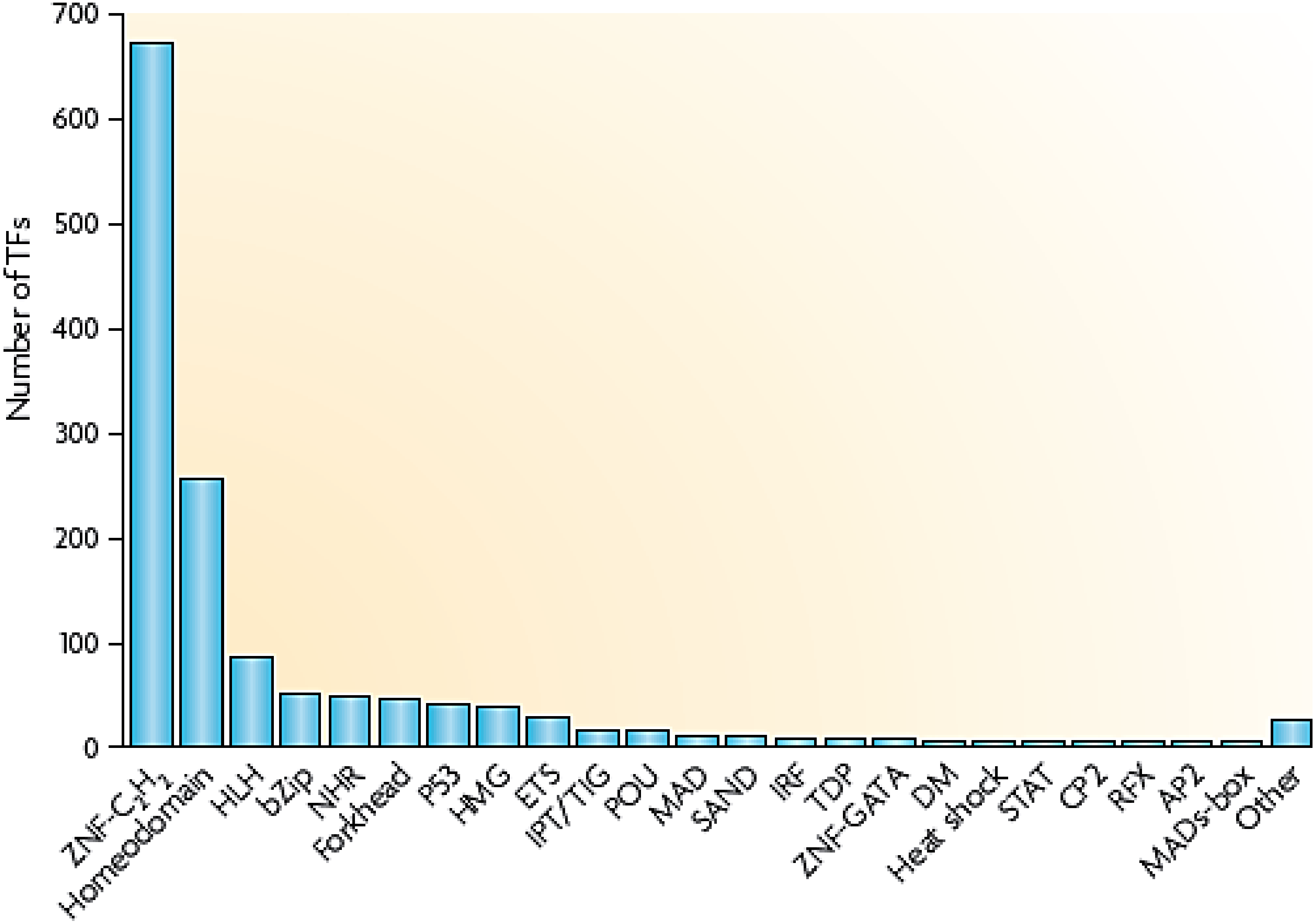
Current state **of knowledge** about transcription factors in the human genome.

a | For the top 20 most cited transcription factors (TFs) in PubMed the number of studies performed in humans (blue bars) and in all other organisms (grey bars) is shown. ER* combines the citations for ERS1 and ERS2, which were indistinguishable in the literature search; similarly, STAT5* includes citations for both STAT5A and STAT5B.

b | summary of biological processes regulated by TFs.
Annotations were obtained from the Gene Ontology database, excluding those based only in electronic annotation. Numbers of annotated TFs are given in parentheses; each gene can be annotated with more than one function.



Vaquerizas, NRG 2009



<5 members

Figure 2 | **Transcription factors classified by DNA-binding domain.** Transcription factors (TFs) were classified into families according to their DNA-binding domain composition. InterPro parent–child relationships between DNA-binding domains were used as the basis for TF family definition (Supplementary information S1 (PDF)). TFs with multiple DNA-binding domains were classified in each of their respective families. Families with less than five members were classified as 'other'.

Vaquerizas, NRG 2009

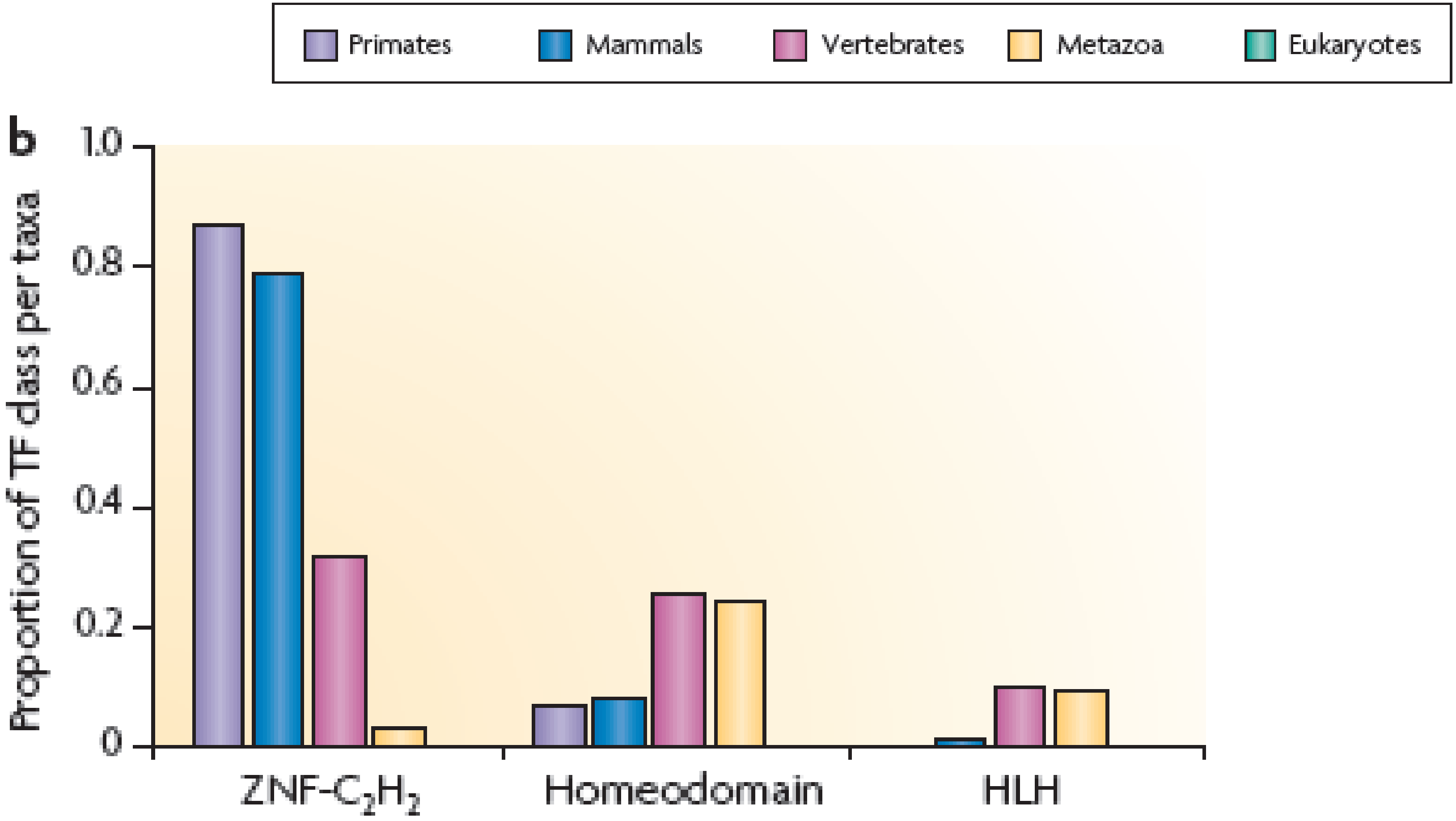
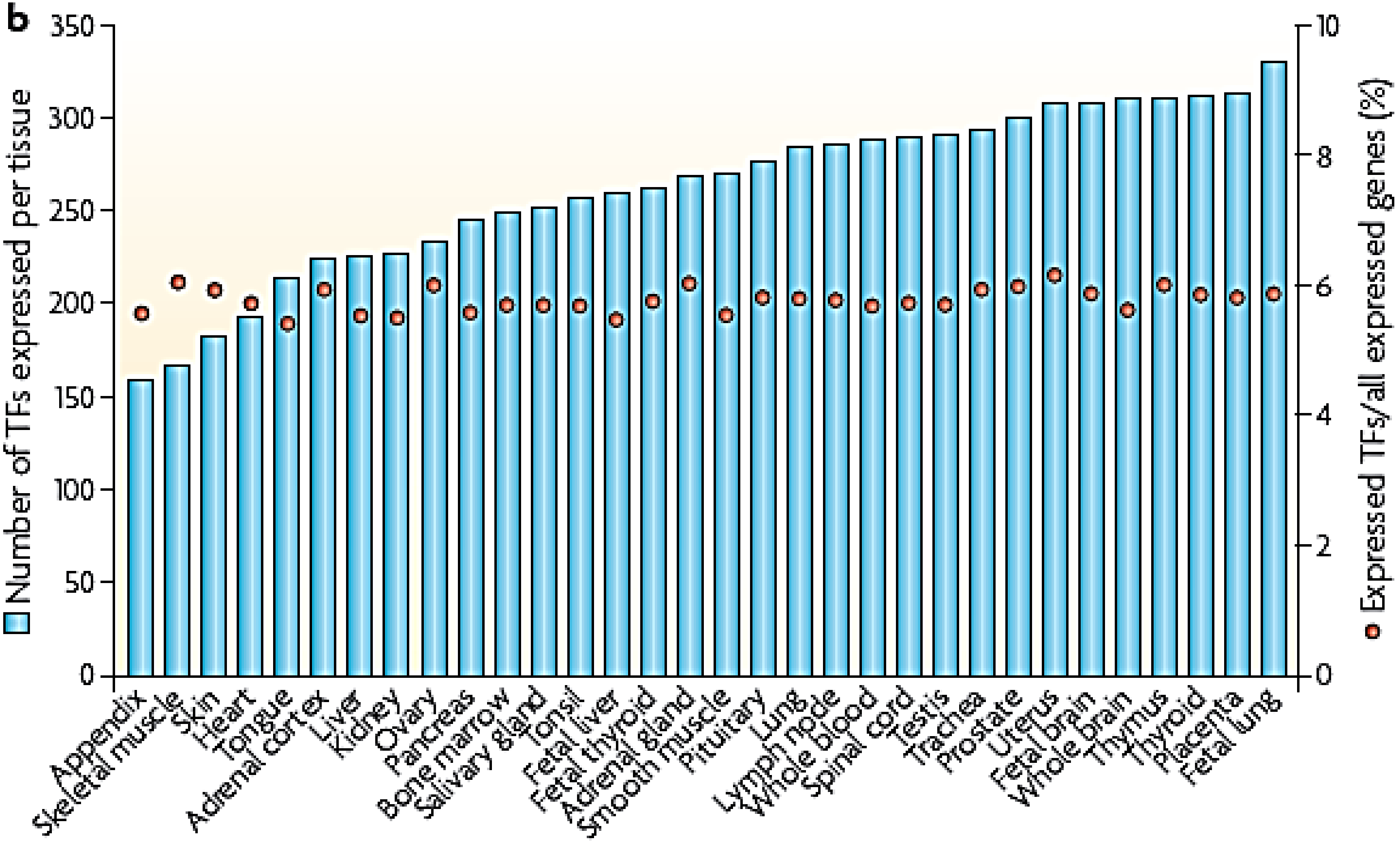


Figure 5 | **conservation of human transcription factors across 24 eukaryotic genomes.** **b** | For human TFs in the three largest families, the proportion that are conserved in each taxonomic group is shown.

Vaquerizas, NRG 2009



b | Numbers of TFs expressed in each sample (blue bars) and the proportion of expressed TFs versus all expressed genes, given as a percentage (red points). The numbers of expressed regulators vary widely, ranging from about 150 in the appendix to over 300 in the fetal lung. However, in all tissues, TFs constitute ~6% of expressed genes.

From Vaquerizas, NRG 2009

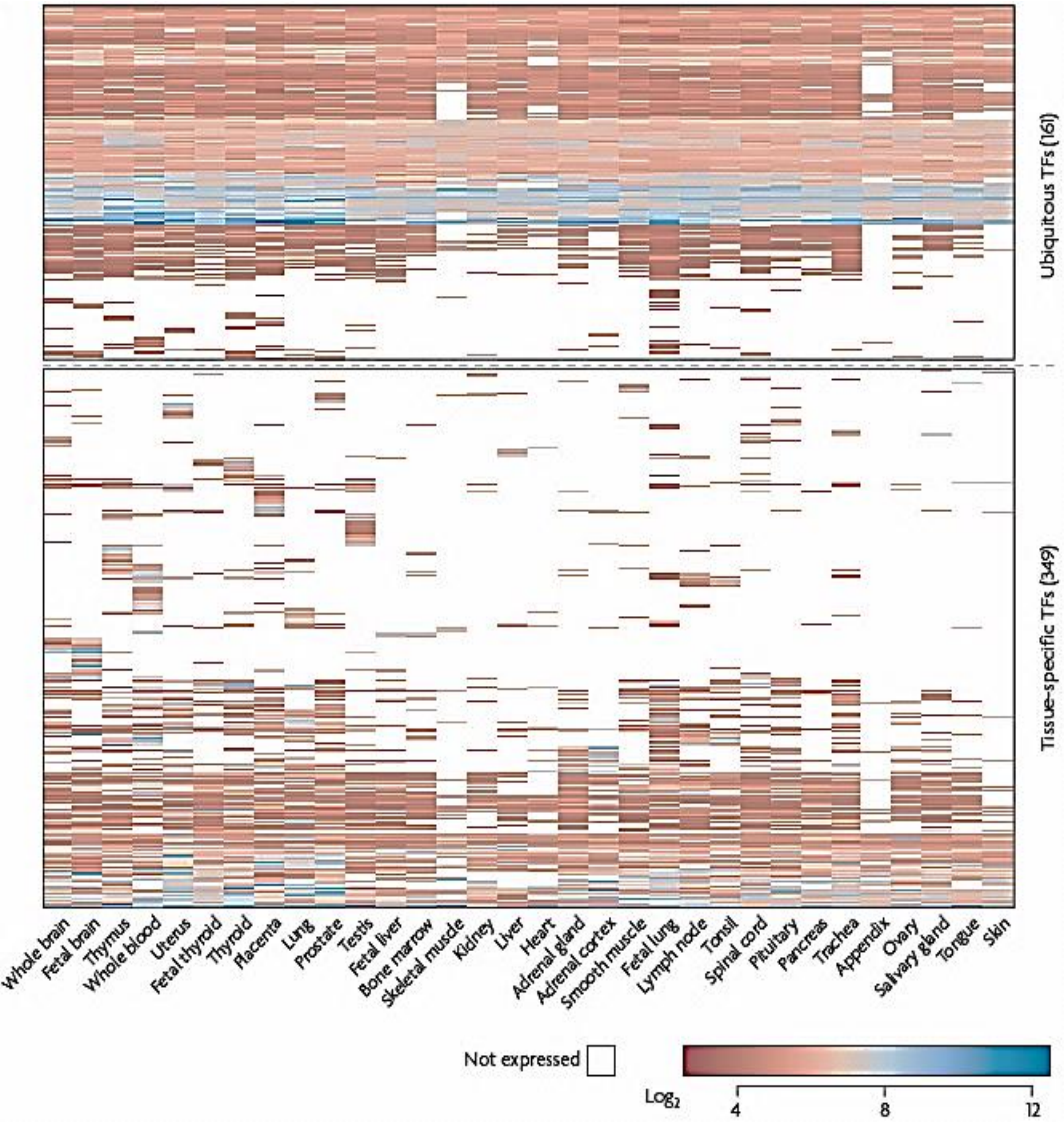


Figure 4 | **Heat map representation of transcription factor expression in 32 human organs and tissues.**

Heat map of transcription factor (TF) expression (rows) in 32 organs and tissues (columns). Intersecting cells are shaded according to expression level (dark red for low expression and blue for high expression). Ubiquitous and specific TFs are grouped according to their expression profiles using hierarchical clustering (before setting an expression level threshold). Ubiquitous regulators are expressed at similar levels across most tissues, whereas specific regulators are expressed at significantly different levels in certain tissues (supplementary information s1 (PDF)). Expression levels below the threshold of detection are depicted as white cells.

Vaquerizas, NRG 2009

Characterizaton of Transcription Factors

- ➢ Analysis of DNA-binding activity
- ➢ Structural analysis – crystallography
- ➢ Analysis of trans-activation properties
- ➢ Identification of DNA response elements = TFBS
- ➢ Identification of genome-wide binding activity
- ➢ Identification of co-operating TFs
- ➢ Protein-protein interactions on DNA

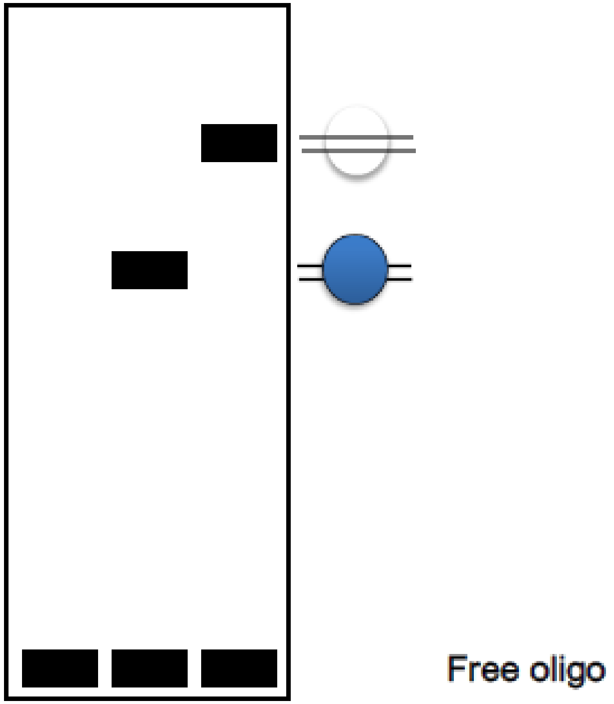
Characterizaton of Transcription Factors

- Analysis of DNA-binding activity

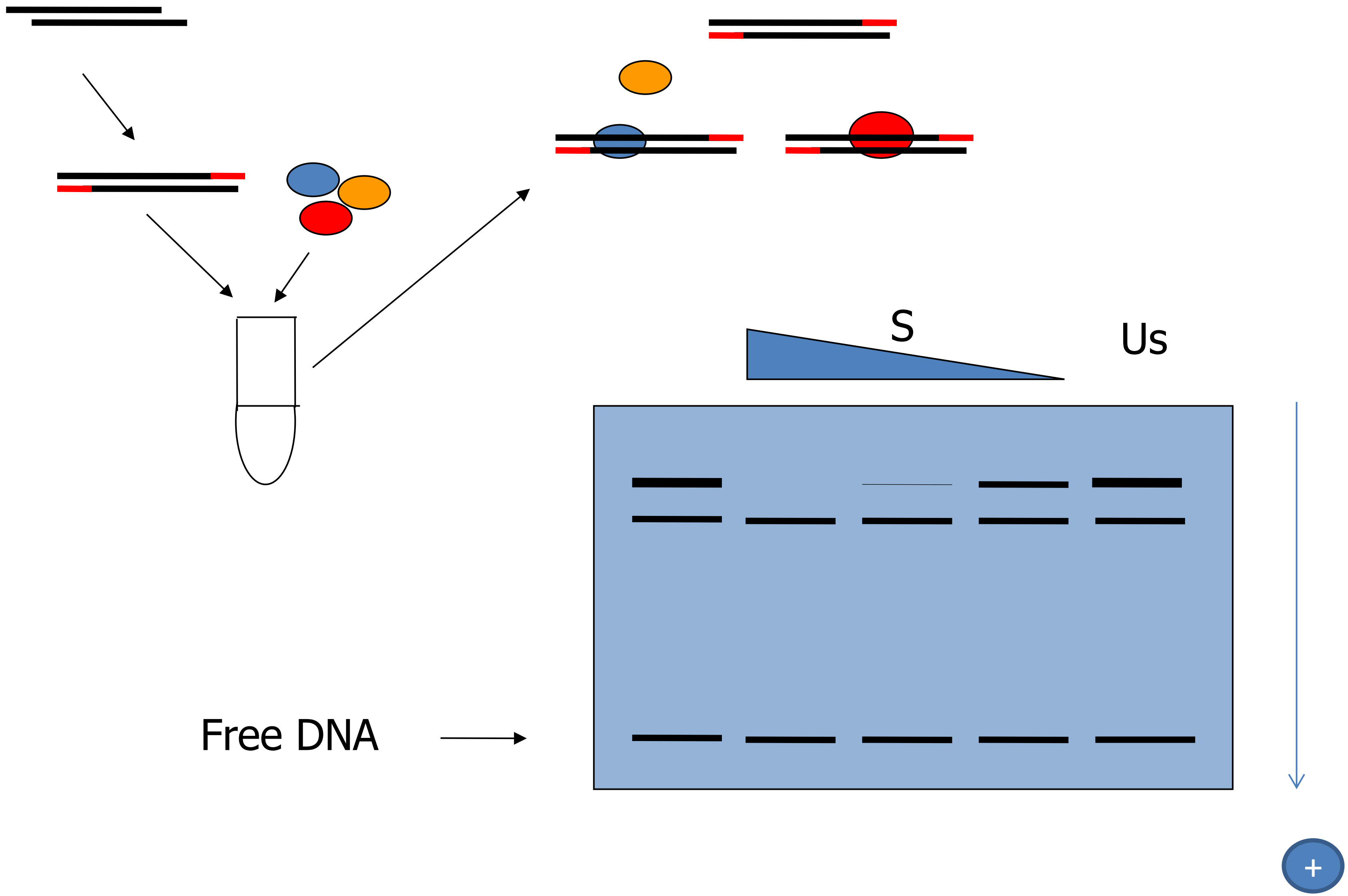
Gel Shifts/Electrophoretic Mobility Shift Assays

- In vitro analysis of the transcriptional factor binding function
- Binding does not always correlate with transcriptional activity

1. Nuclear extracts from cells or tissues
2. Mix with ^{32}P-labeled ds-oligo
3. Run on Native acrylamide gels



Gel-shift assay or Electrophoretic Mobility Shift Assay EMSA



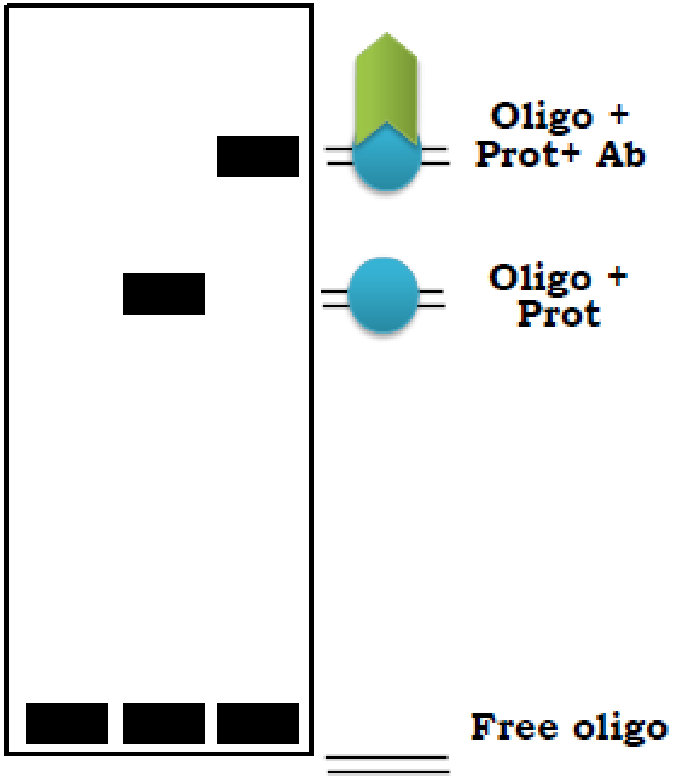
How do we determine the identity of complexes and if they are specific?

- Competition assays

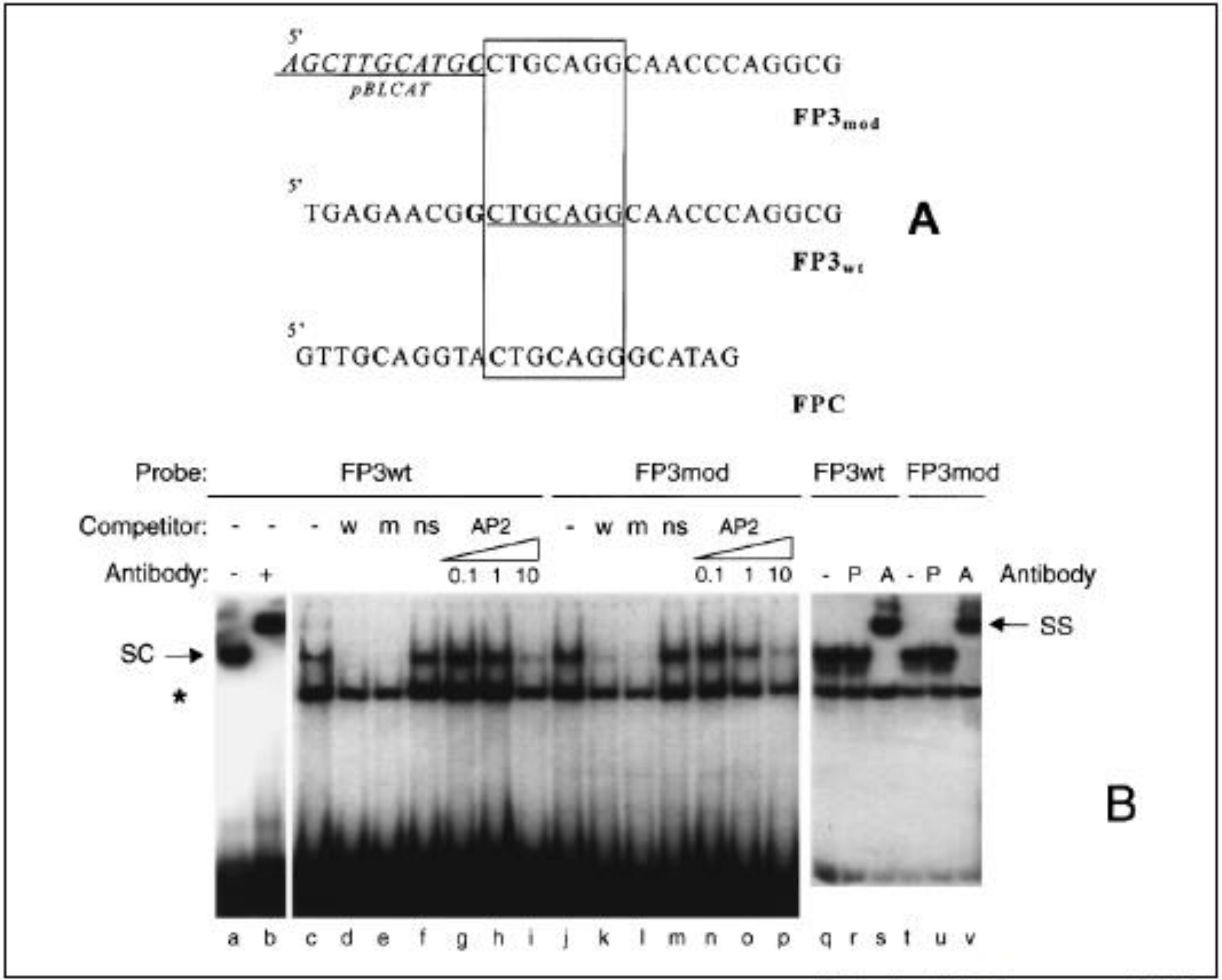
Molar excess of identical, mutant, or consensus site

- Supershift Assays

Add specific antibody to the binding assay



Examples: Gel Shifts/EMSA

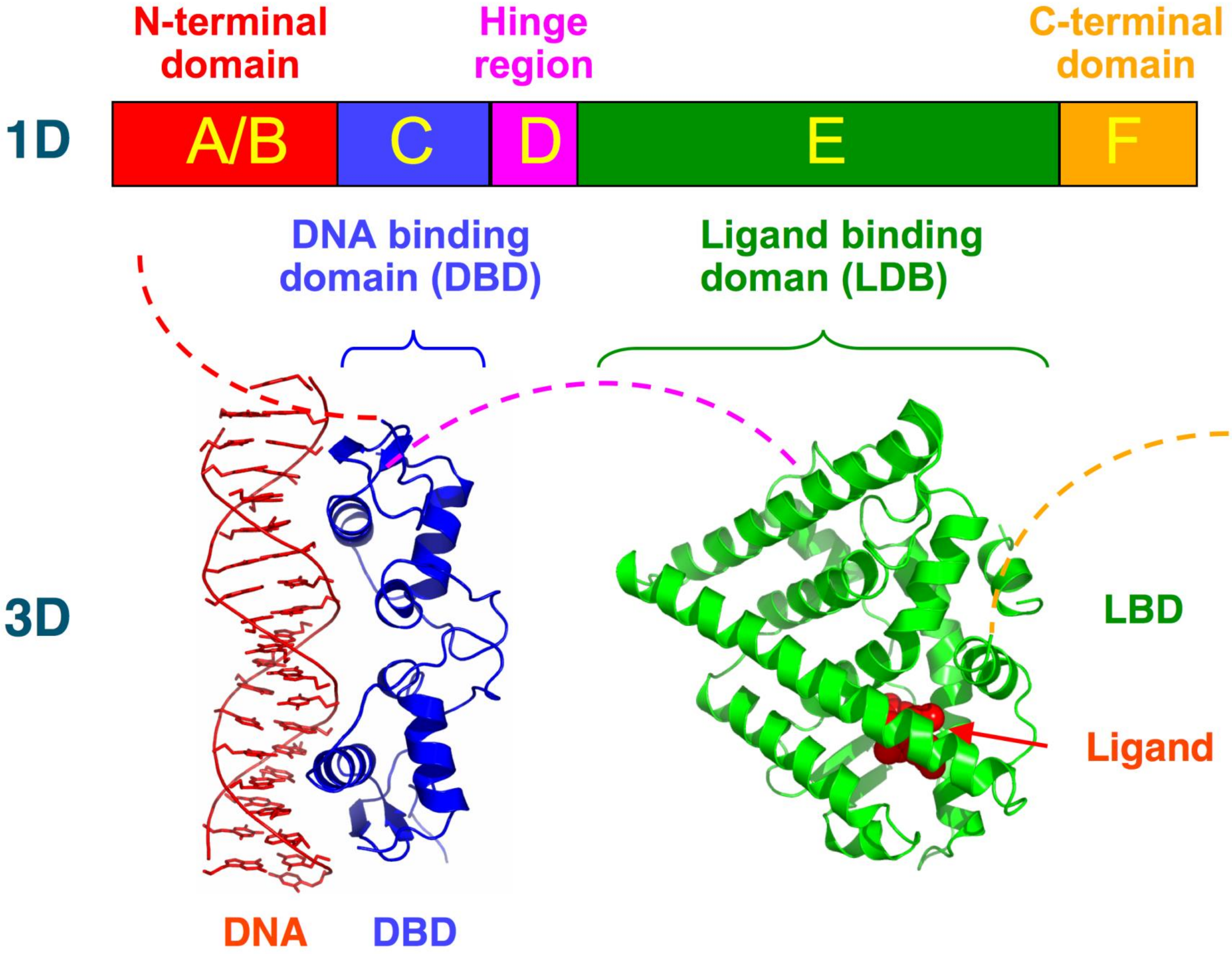

5'
AGCTTGCATGCCTGCAGGCAACCCAGGCG
pBLCAT
FP3mod
5'
TGAGAACGGCTGCAGGCAACCCAGGCG
FP3wt
A
5'
GTTGCAGGTACTGCAGGGCATAG
FPC
Probe:
FP3wt
FP3mod
FP3wt
FP3mod
Competitor: - - - w m ns AP2 - w m ns AP2
Antibody: - +
0.1 1 10
0.1 1 10
- P A - P A Antibody
SC
*
SS
B
a b c d e f g h i j k l m n o p q r s t u v

Perissi et al., Oncogene, 2000

Characterizaton of Transcription Factors

- Structural analysis – crystallography

Structural Organization of Nuclear Receptors



N-terminal domain

Hinge region

C-terminal domain

1D

A/B

C

D

E

F

DNA binding domain (DBD)

Ligand binding doman (LDB)

3D

LBD

Ligand

DNA

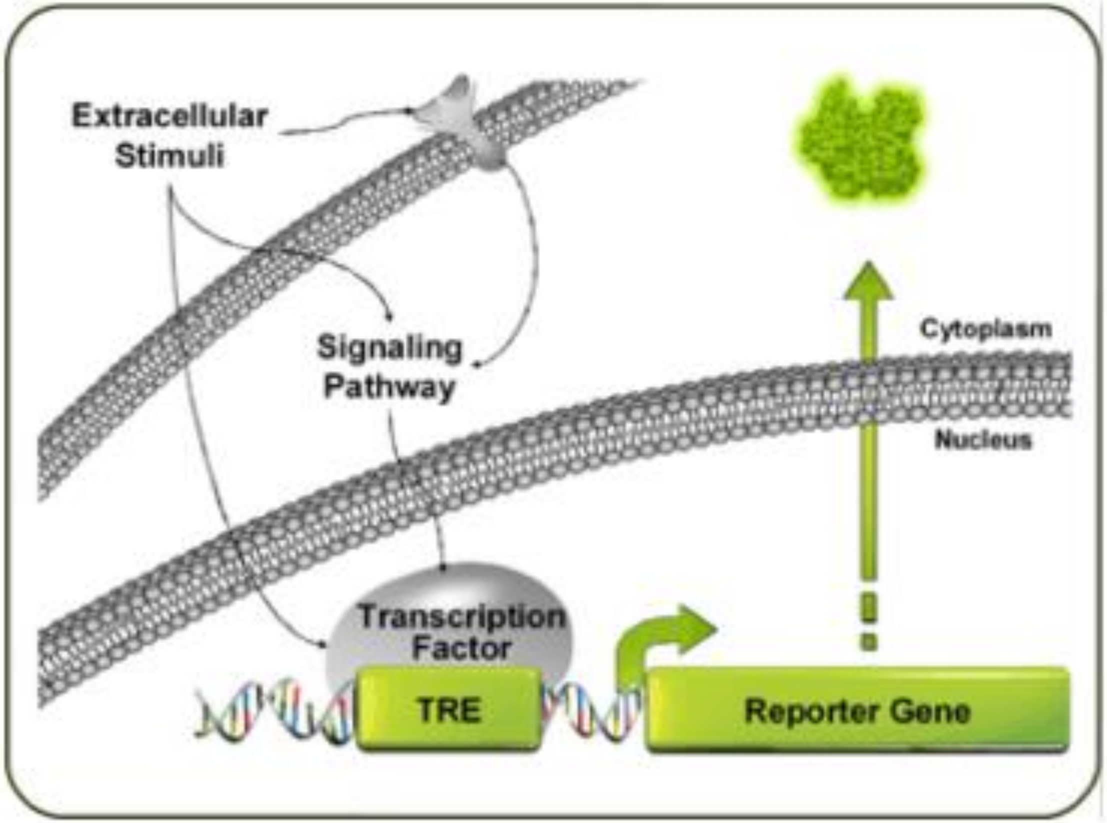
DBD

Structures (DBD)
http://www.rcsb.org/pdb/home/home.do

Characterizaton of Transcription Factors

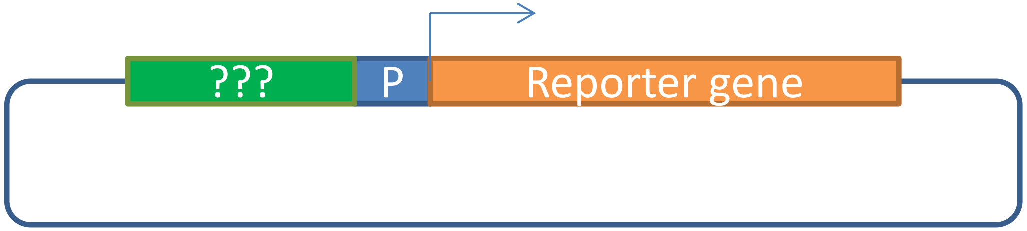
- Analysis of trans-activation properties

What is a Reporter Assay?



Functional validation: **reporter assay**

Plasmid with a **reporter** gene driven by a promoter: clone upstream studied fragment



Transfect into cultured cells, after 24-72 hours measure the product.

Reporter genes are usually nonmammalian genes:

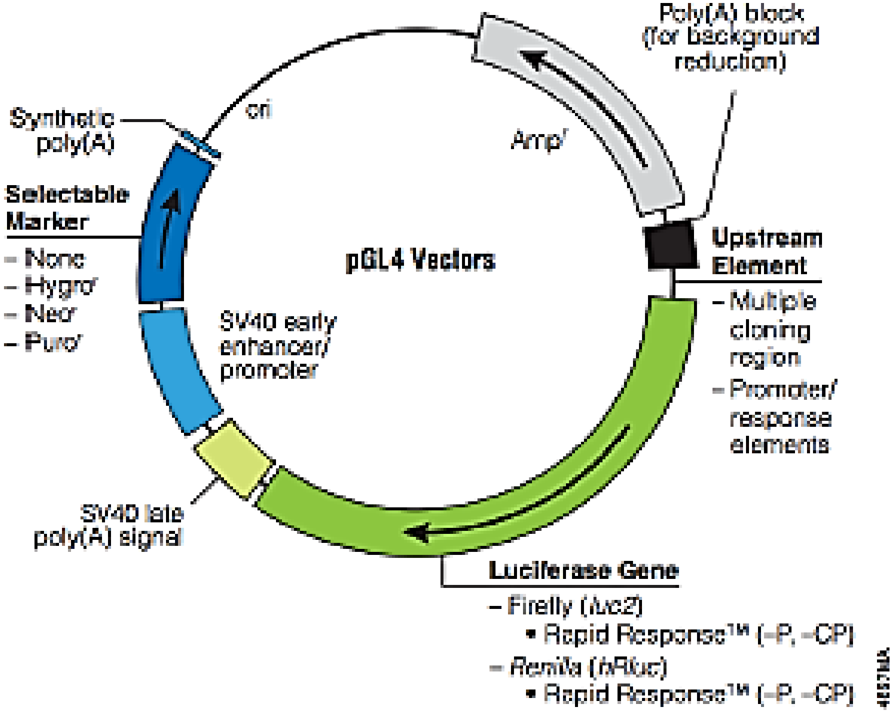
- CAT cloramphenicol acetyltransferase
- Luc firefly luciferase
- GFP Jellyfish green fluorescent protein
- β-gal beta-galactosidase

Deletional + Mutational analysis of fragment may lead to identify important elements

Luciferase Reporters—pGL Family

The pGL4 Vector family includes:

- Basic vectors with no promoter that contain a multiple cloning region for cloning a promoter of choice
- Vectors containing a minimal promoter
- Vectors containing response elements and a minimal promoter
- Promoter-containing vectors that can be used as expression controls or as co-reporter vectors



Reporter Assays

Strengths

High throughput

Can measure function of mutations in promoters

Large dynamic range

Many reporters possible

- GFP
- b-galactosidase
- CAT (chloramphenical acetyl transferase)
- Luciferase (firefly, renilla)

Can be used for in vivo/in cell monitoring.

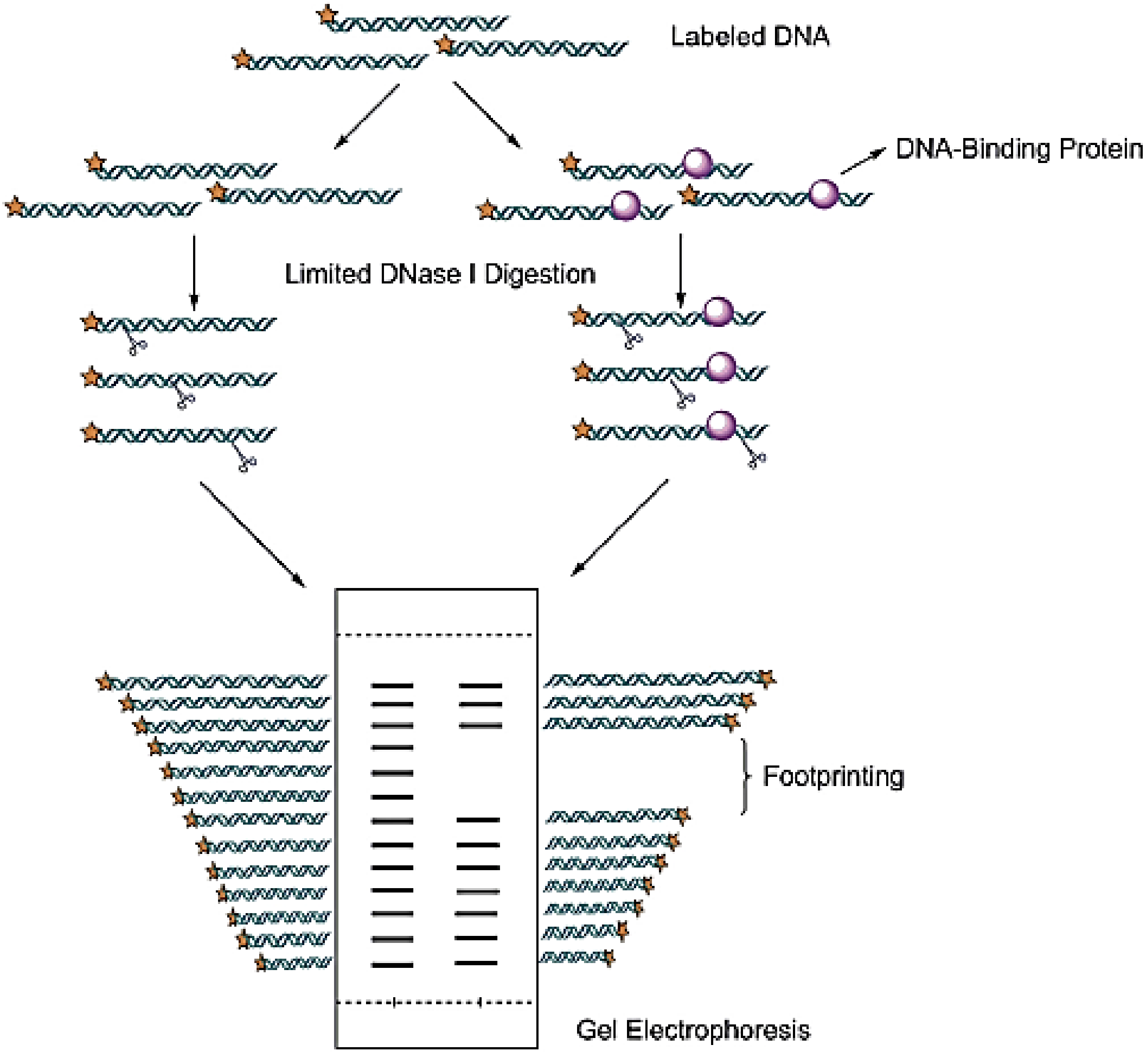
Weaknesses

- uses exogenous DNA, not chromatin
- Gene dosage artifacts are possible
- relies on transfection, not easy for all cells

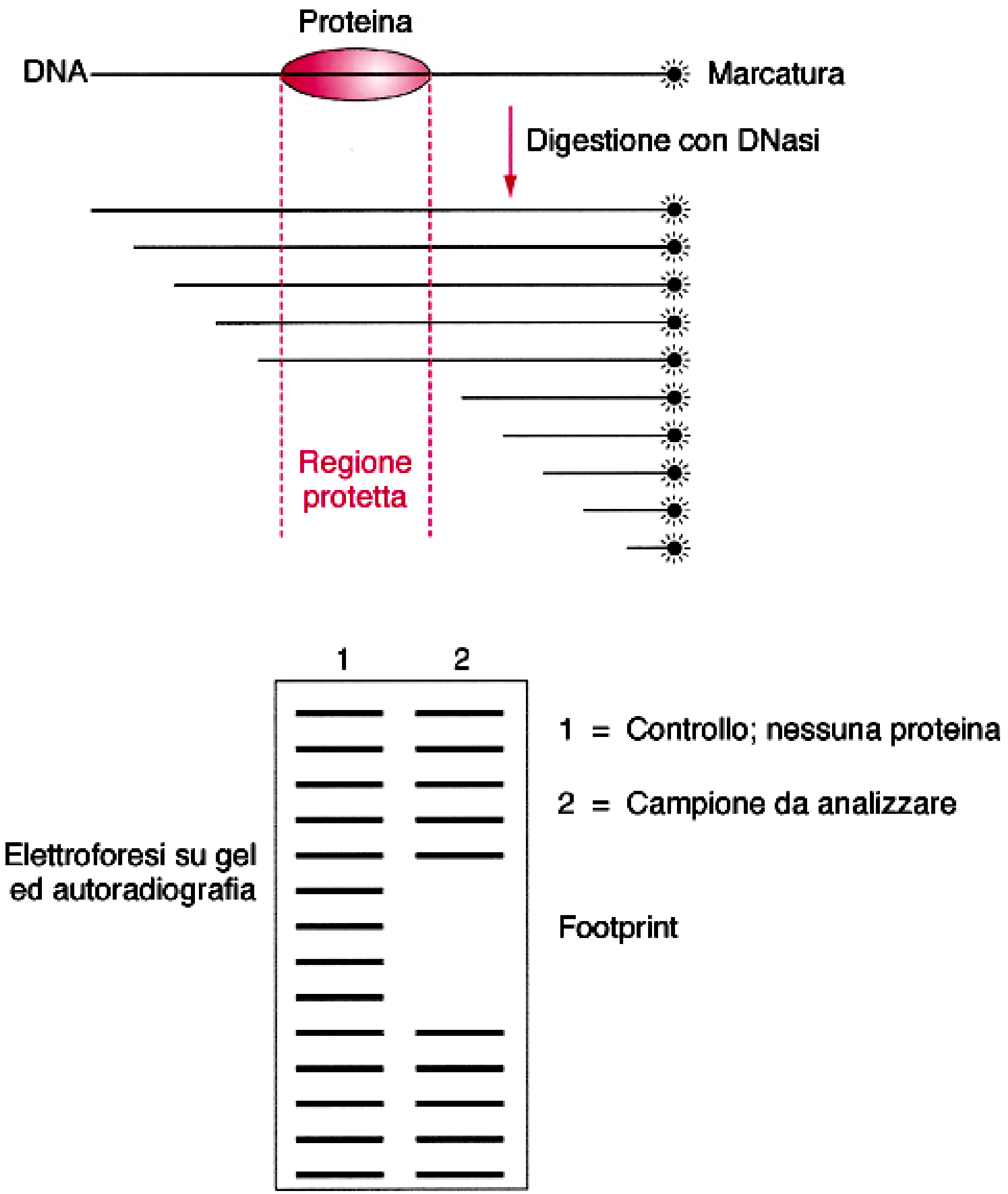
Characterizaton of Transcription Factors

- Identification of DNA response elements = TFBS

DNaseI Footprinting

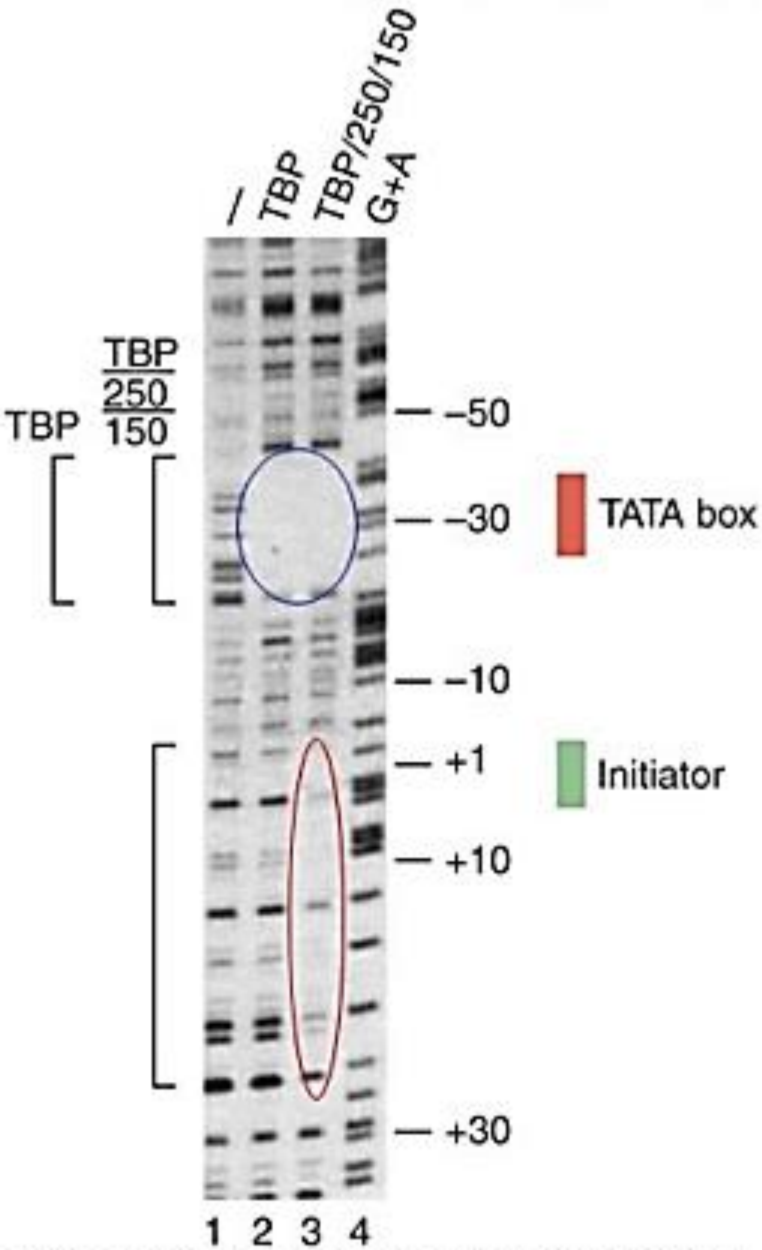


DNaseI Footprinting


Proteina
DNA
Marcatura
Digestione con DNasi
Regione protetta
1
2
1 = Controllo; nessuna proteina
2 = Campione da analizzare
Elettroforesi su gel ed autoradiografia
Footprint

DNaseI Footprinting of hsp70 promoter

TAF$_{II}$250 and TAF$_{II}$150 cooperate in binding to the initiator & DPE



Verrijzer et al., Promoter recognition by TAFs. *Cell* 81 (30 Jun 1995) p. 1117, f. 2c.
Reprinted by permission of Elsevier Science.

Identification of DNA Binding specificity

Historical

1° route:

isolating a promoter sequence, make deletional mutants and identify regulatory elements.
This is paralleled with Dnase I footprinting experiments using whole Nuclear Extract.
Once identified, the response elements are further analyzed by Band-shift (EMSA)
Proteins bound are then isolated by DNA affinity chromatography and identified.

This approach has led to the characterization of several tens of Transcription Factors.

2° route:

Several putative TFs are identified by homology cloning.
The binding site was often identified by **SELEX**
Finally, bioinformatic search for the binding site is performed on known genomic sequences.

3° route:

Conserved, nontranscribed sequences proximal to known genes are explored statistically to describe over-represented sequence "words" as compared to the whole genome.
Experimental proofs that the identified "words" (or motifs) can bind regulatory factors are needed

SELEX

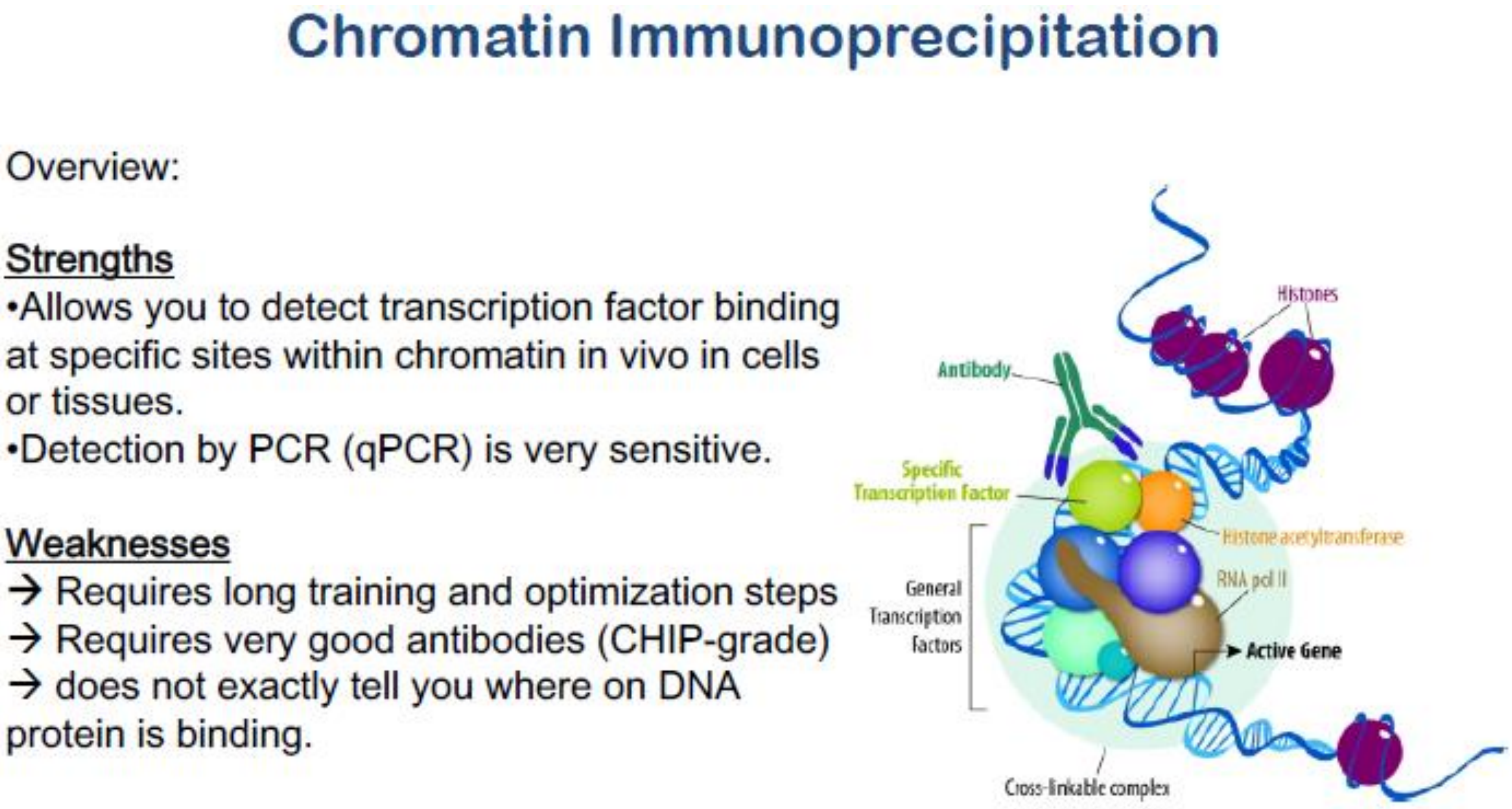
A random sequence oligonucleotide library is explored using a purified or recombinant Transcription Factor

Classical SELEX: many rounds of selection + PCR amplification

SELEX variants: single selection step at high stringency, followed by elution and NGS sequencing

Usually produced a series of short sequences → consensus

Transcription Factors / DNA binding



Chromatin Immunoprecipitation

Overview:

Strengths

•Allows you to detect transcription factor binding at specific sites within chromatin in vivo in cells or tissues.

•Detection by PCR (qPCR) is very sensitive.

Weaknesses

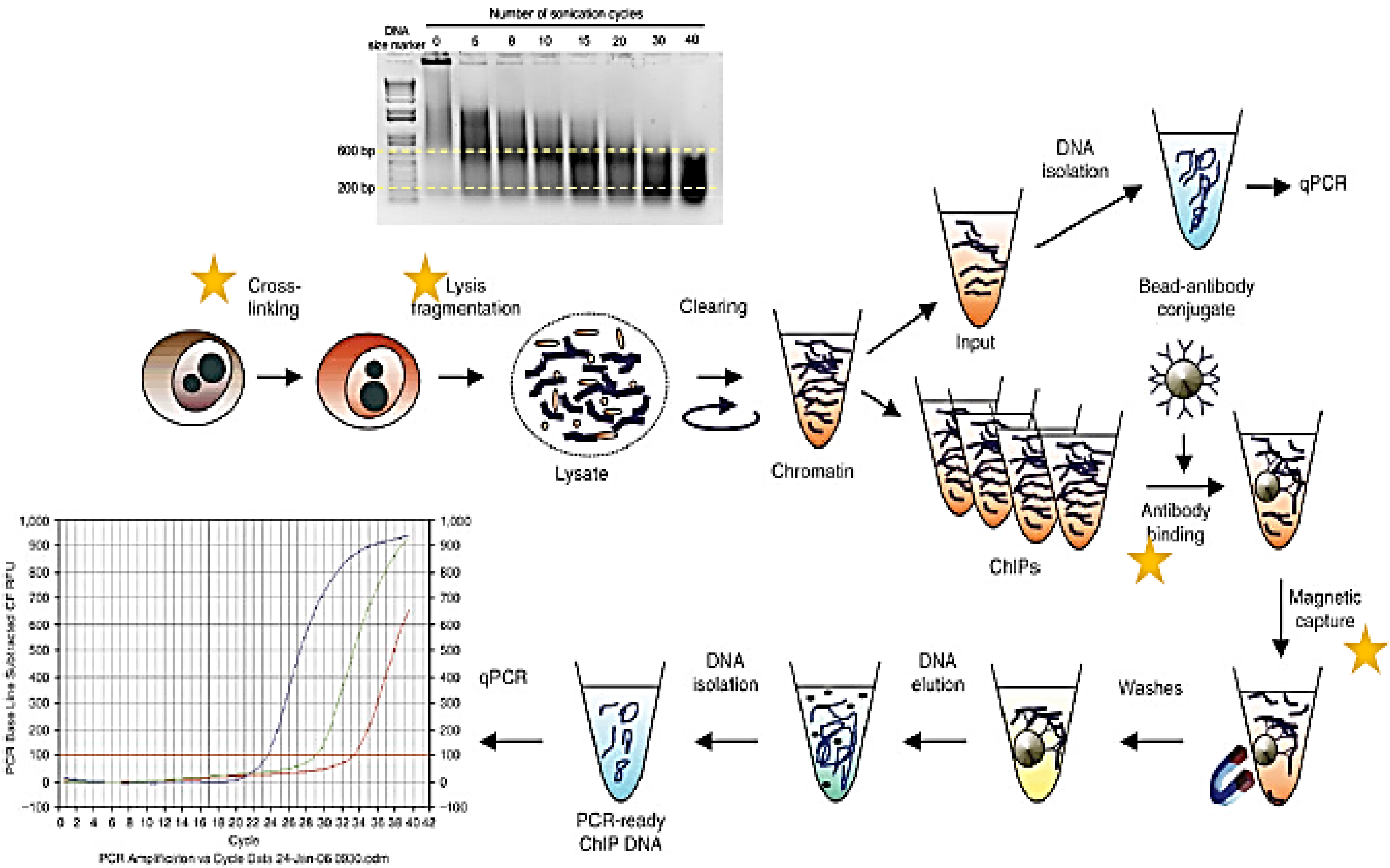
→ Requires long training and optimization steps

→ Requires very good antibodies (CHIP-grade)

→ does not exactly tell you where on DNA protein is binding.

Histones
Antibody
Specific Transcription Factor
Histone acetyltransferase
RNA pol II
General Transcription Factors
Active Gene
Cross-linkable complex

ChIP Steps & Optimization



ChIP Controls

ChIP controls

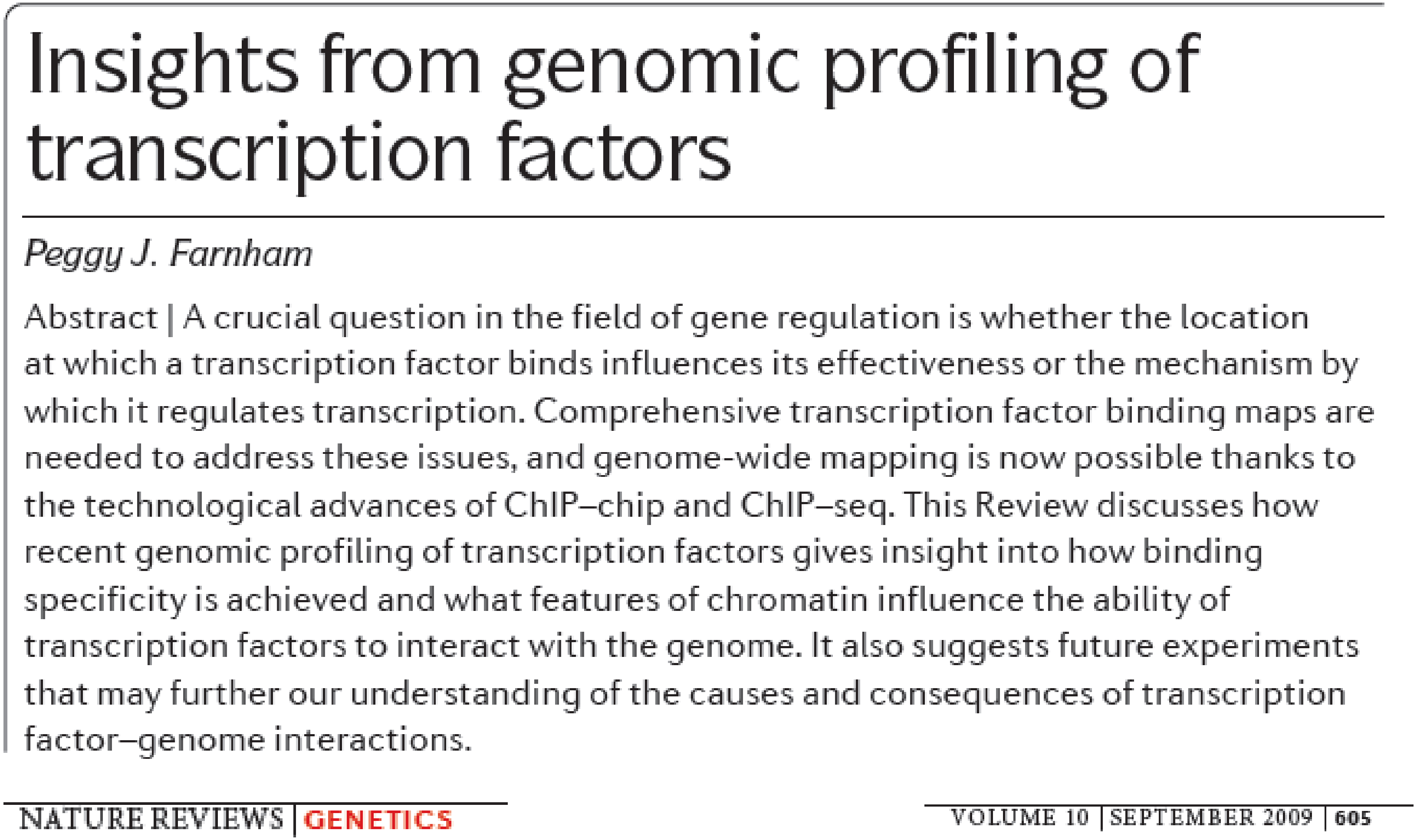
- PolII or histone marks antibodies can be useful if you are unsure about your antibody (positive control for ChIP technique)
- Controls for genes that have previously been shown to be bound by factor if interest (positive control for antibody)
- Controls for unrelated genomic regions that should not bind factor of interest (negative control)
- Normal IgG or pre-immune IgG (negative control for IP)

PCR controls

- Negative PCR controls as usual
- Serial dilutions of input material to calculate reference curve
- Dissociation curve to validate primers

Identification of R.E.

Using ChIP + microarrays or (best) ChIP-Seq it has been quite straightforward to obtain high-resolution maps of TF binding to chromatin using cell lines.



Insights from genomic profiling of transcription factors

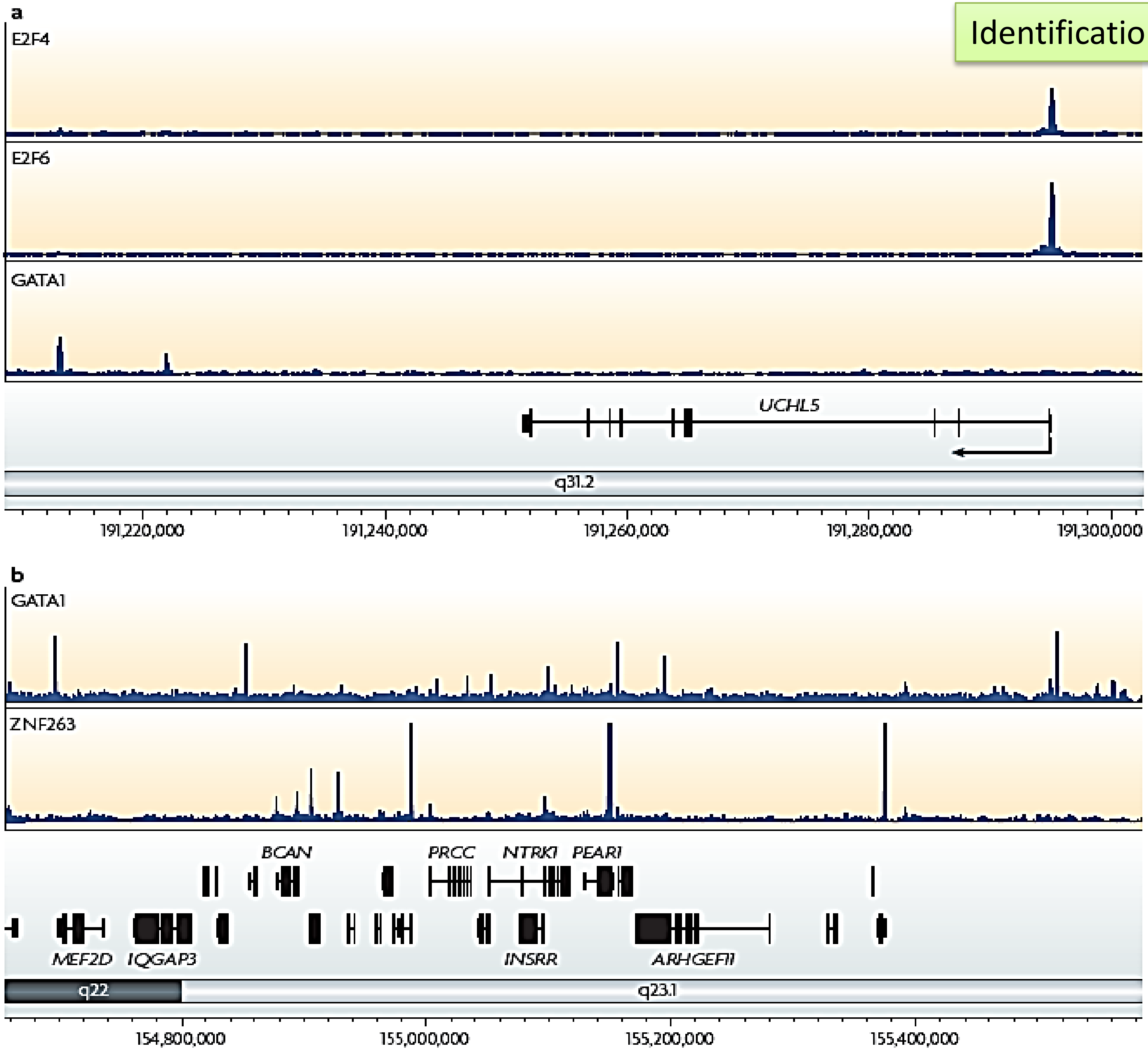
Peggy J. Farnham

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.

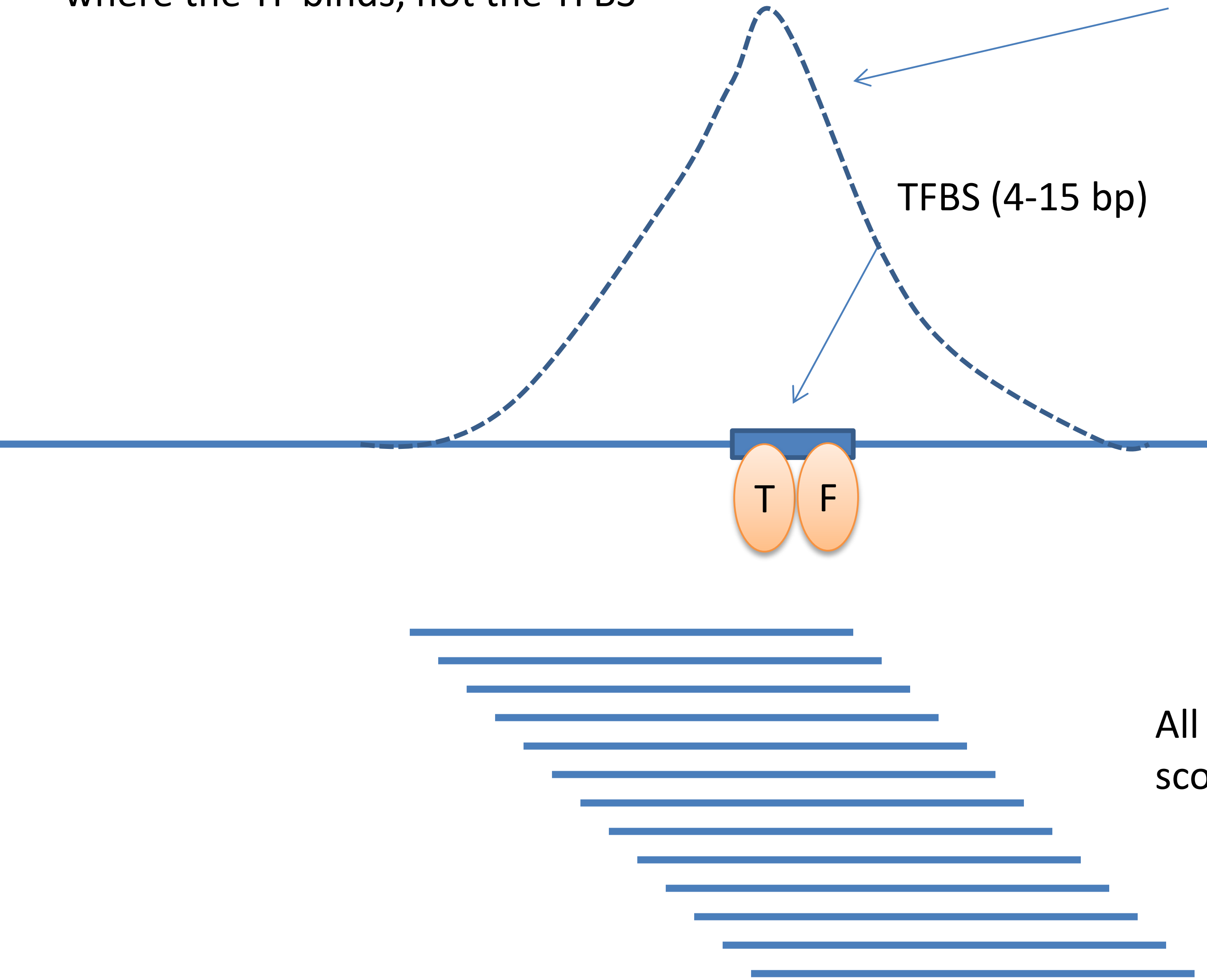
NATURE REVIEWS | GENETICS

VOLUME 10 | SEPTEMBER 2009 | 605

Identification of R.E.



ChIp-Seq analysis identifies a region where the TF binds, not the TFBS



Frequency plot

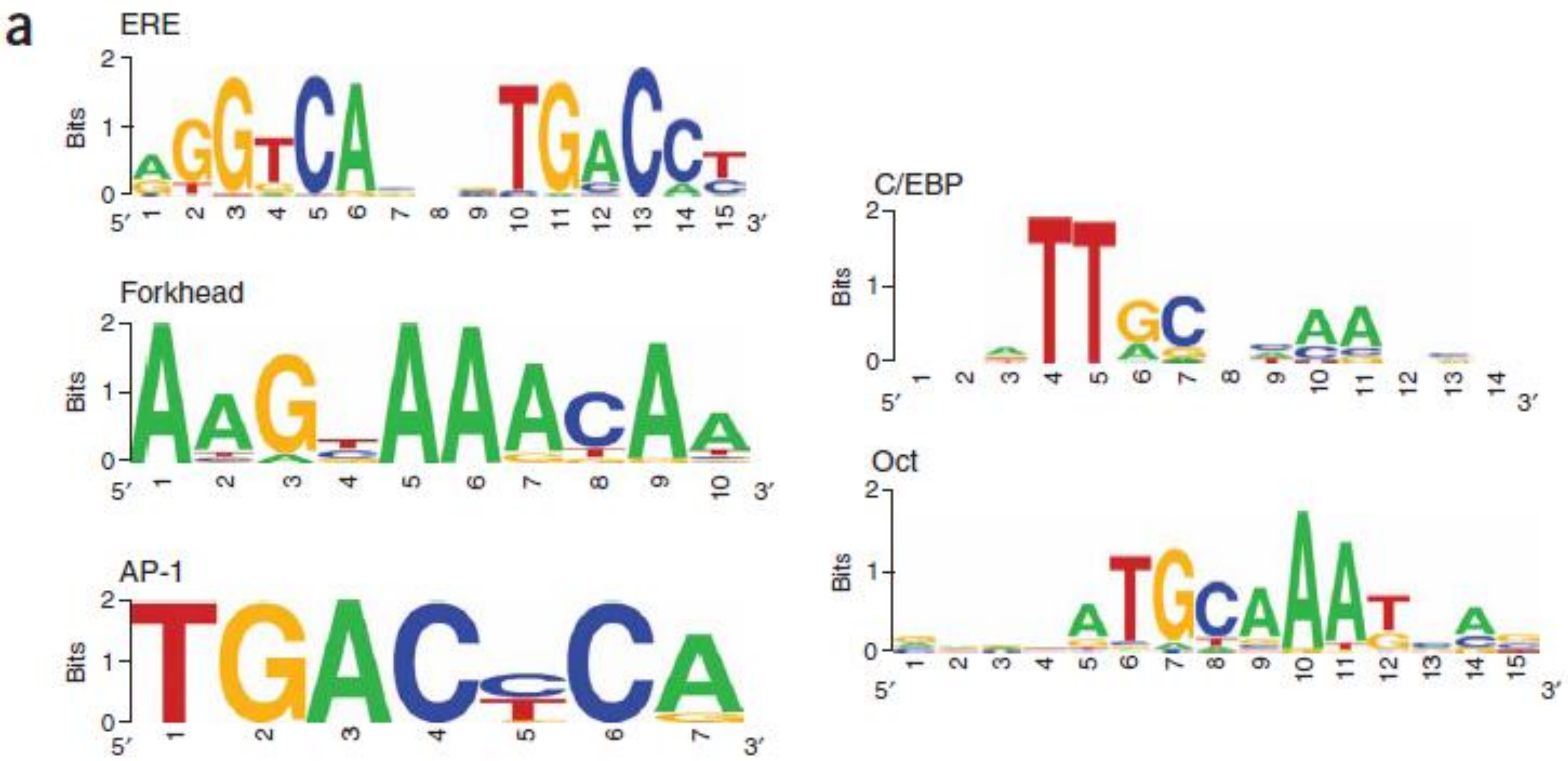
TFBS (4-15 bp)

All these fragment will score positive in ChIP assay

Identification of R.E.

In the **-500, +500** interval around binding peaks, algorithms exist to find unbiased overrepresented motifs, or known motifs based on **positional weight matrices**.

Examples:


a
ERE
Forkhead
AP-1
C/EBP
Oct
Bits
5′
3′