Inteligentny projekt to rzeczywistość – dowody z DNA

Intelligent Design is a Reality: The DNA evidence

How the 'gene' is a multilevel mediator of 'information' that lacks a material description

Richard v. Sternberg

Biologic Institute

Czym dokładnie jest gen?

What exactly is a ‘gene’?

What is a gene, post-ENCODE? History and updated definition

Mark B. Gerstein,[1,2,3,9] Can Bruce,[2,4] Joel S. Rozowsky,[2] Deyou Zheng,[2] Jiang Du,[3] Jan O. Korbel,[2,5] Olof Emanuelsson,[6] Zhengdong D. Zhang,[2] Sherman Weissman,[7] and Michael Snyder[2,8]

17:669–681 ©2007
Genome Research

Theory Biosci.
DOI 10.1007/s12064-008-0025-0

ORIGINAL PAPER

“Genes”

Sonja J. Prohaska · Peter F. Stadler

Theory Biosci.
DOI 10.1007/s12064-009-0067-y

SHORT COMMUNICATION

Defining genes: a computational framework

Peter F. Stadler · Sonja J. Prohaska · Christian V. Forst · David C. Krakauer

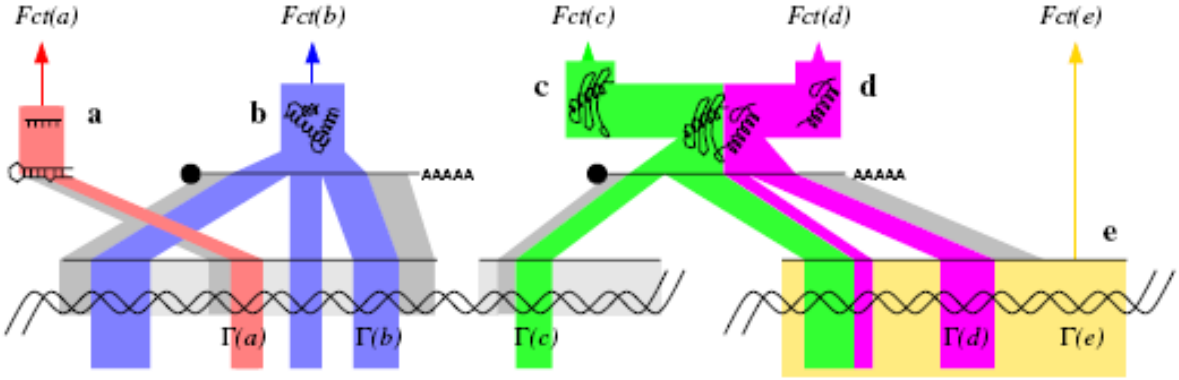


Fig. 1 Functional objects *a* to *e* and relationships with their genomic footprints $\Gamma(a)$ to $\Gamma(e)$. A functional RNA molecule (e.g., a miRNA) with function $\mathrm{Fct}(a)$ is processed in two steps from an intronic sequence. Its image on the DNA is the genomic footprint $\Gamma(a)$. The genomic footprint $\Gamma(b)$ of the functional protein *b* is a discontinuous stretch of DNA corresponding to the coding sequence (CDS) including the start codon but excluding the stop codon. The mRNA includes UTRs that also map back to the DNA as well as parts without footprints on the DNA (the 5′-cap and the poly-A tail). The functional proteins *c* and *d* are obtained by cleavage of the (non-functional) precursor *cd*. The later is encoded by a trans-spliced mRNA. The footprint $\Gamma(c)$ is distributed over two DNA molecules. The primary transcript *e* has an additional function $\mathrm{Fct}(e)$ that is independent of its role as precursor of the mRNA of *cd*. As a consequence, $\Gamma(e)$ overlaps with both, $\Gamma(c)$ and $\Gamma(d)$. In all cases, the gene is the pair $(\Gamma(x),x)$ composed of the genomic footprint $\Gamma(x)$ and the resulting functional molecule *x*

Although the gene has conventionally been viewed as the fundamental unit of genomic organization, on the basis of ENCODE data it is now compellingly argued that this unit is not the gene but rather the transcript (Washietl et al. 2007; Djebali et al. 2012a). On this view, genes represent a higher-order framework around which individual transcripts coalesce, creating a polyfunctional entity that assumes different forms under different cellular states, guided by differential utilization of regulatory DNA.

What does our genome encode?

John A. Stamatoyannopoulos

Genome Res. 2012 22: 1602-1611

Teza

Thesis. In order for a ‘gene’ to be a ‘gene’, to be a “higher-order poly-functional entity” that takes on different forms at different times, it has to be more than an invariant particle.

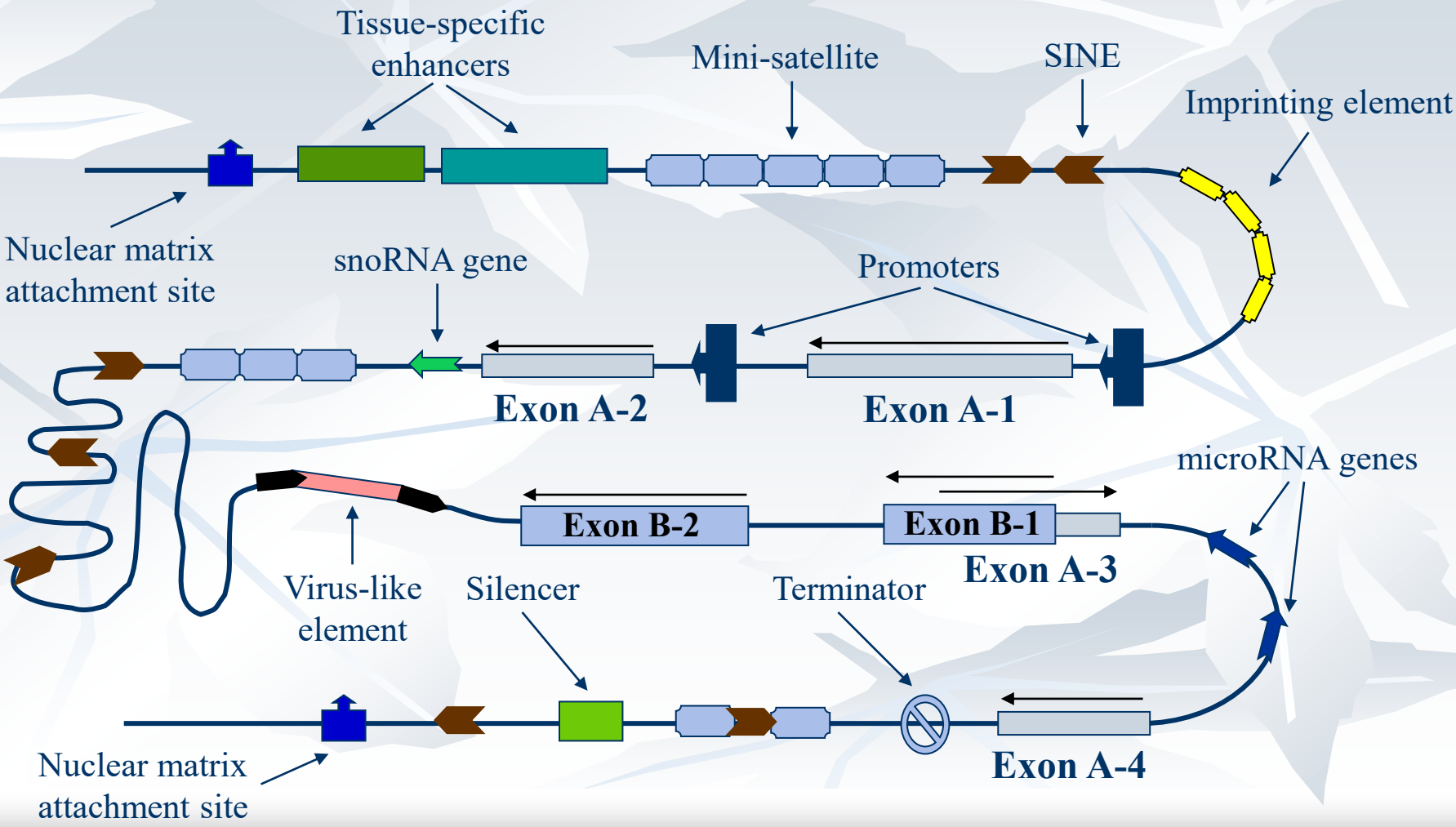
Pytanie

The question is: **What evidence do we have that such a thesis could be (at least partly) correct?**

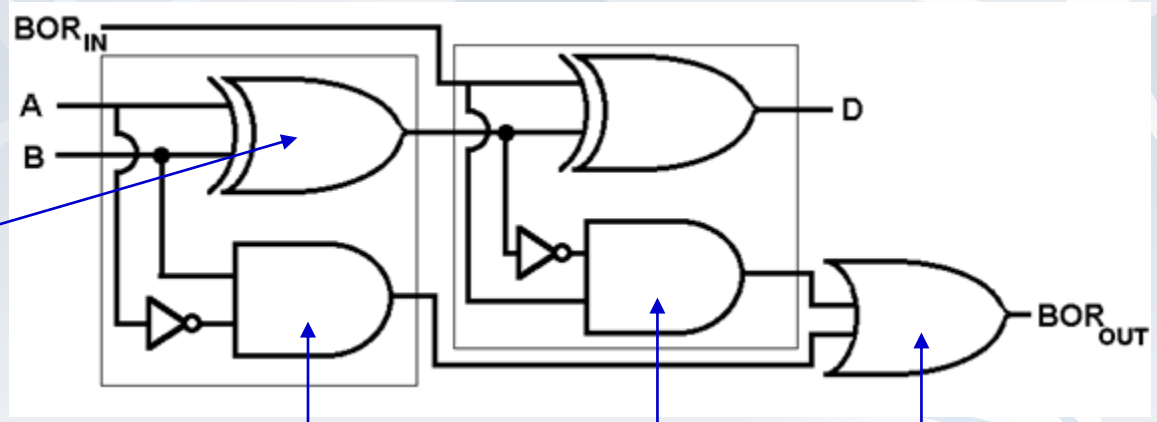
Wskazówka 1

The First Clue

Principle 1: A typical (metazoan) "gene" consists of interleaved, interspersed, multilevel, and overlapping "data files."



Principle 2: This order permits a "gene" to be formed into circuits differentially.



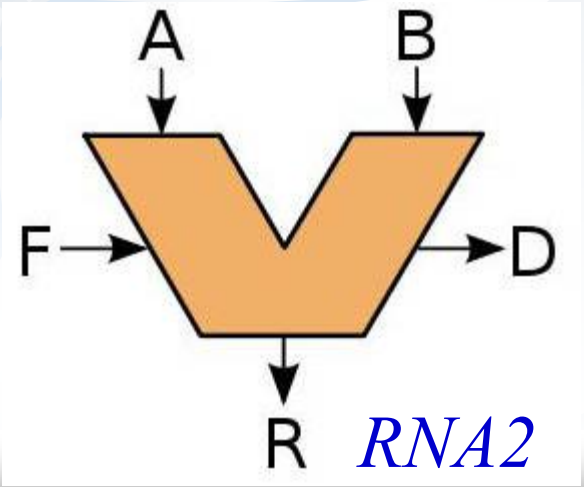
Tissue-specific enhancer

Imprinting element

SINE

Promoter

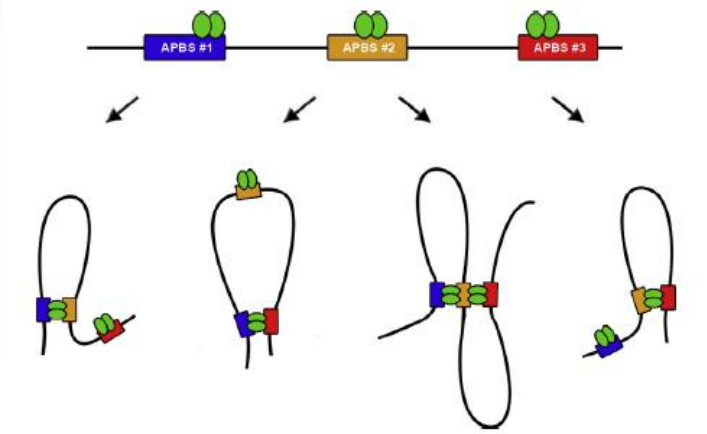
RNAs



RNA1

RNA2

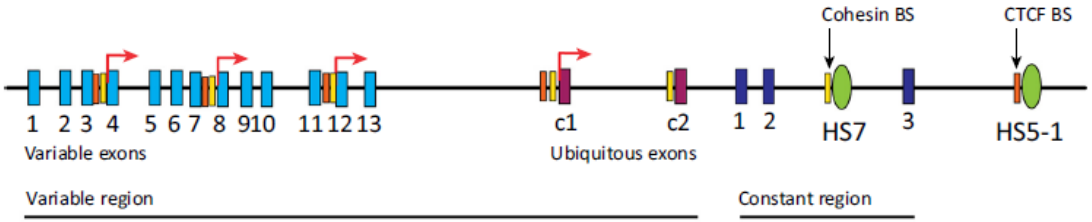
Wikipedia

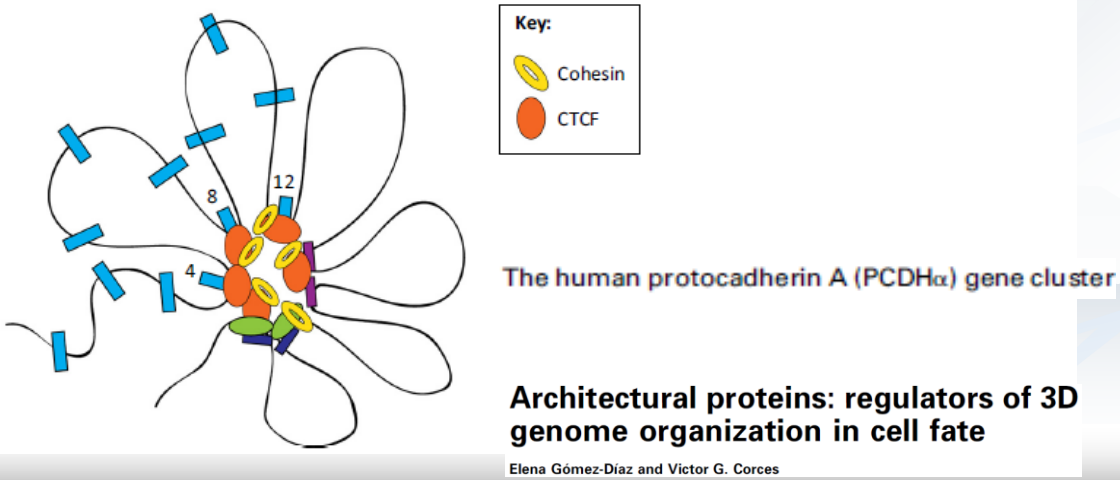


Architectural proteins, transcription, and the three-dimensional organization of the genome

http://dx.doi.org/10.1016/j.febslet.2015.05.025

Caelin Cubeñas-Potts, Victor G. Corces *





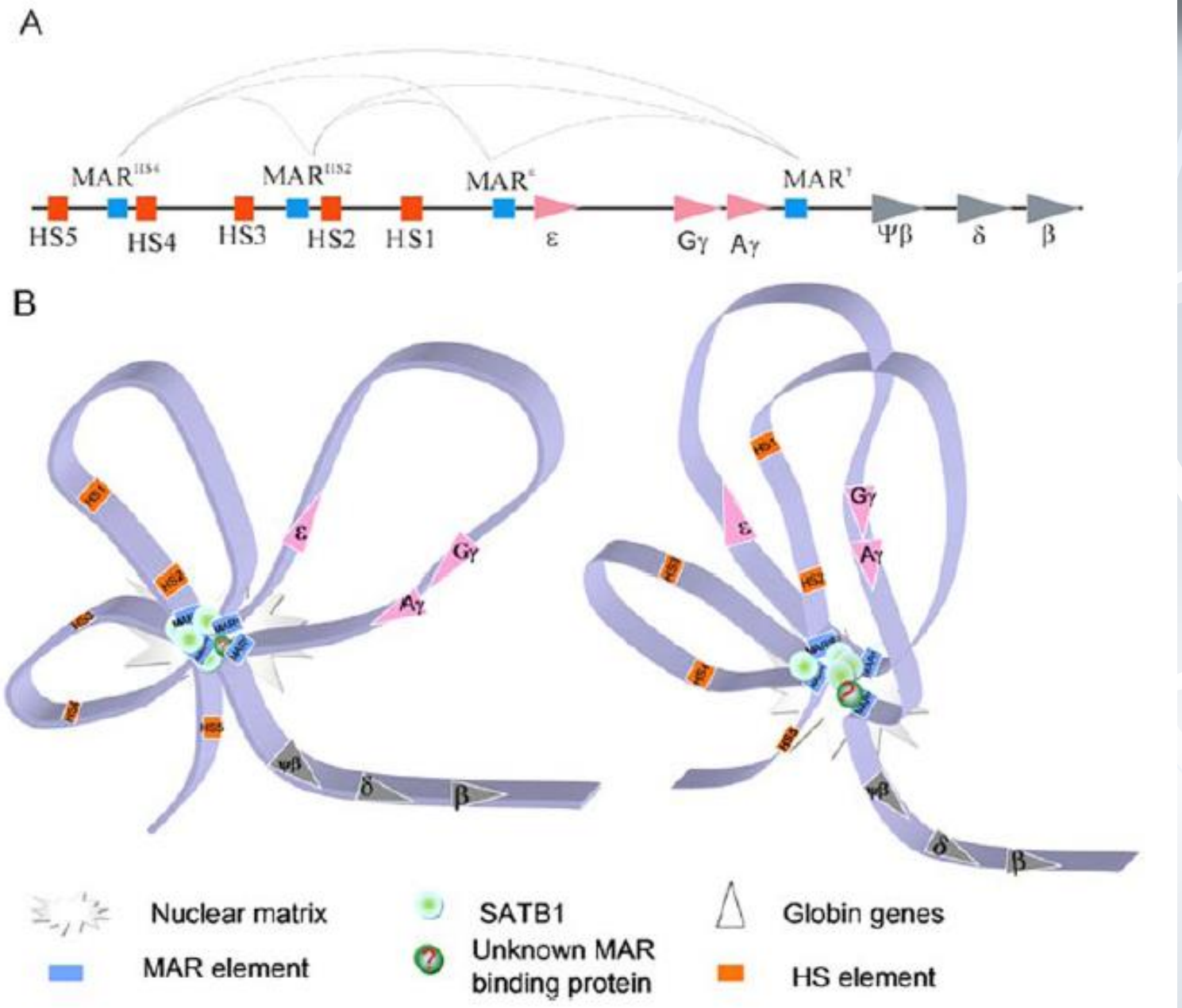
The human protocadherin A (PCDHα) gene cluster

Architectural proteins: regulators of 3D genome organization in cell fate

Elena Gómez-Díaz and Victor G. Corces

Trends in Cell Biology, November 2014, Vol. 24, No. 11

Chromatin folding indeed allows different circuits to be formed.

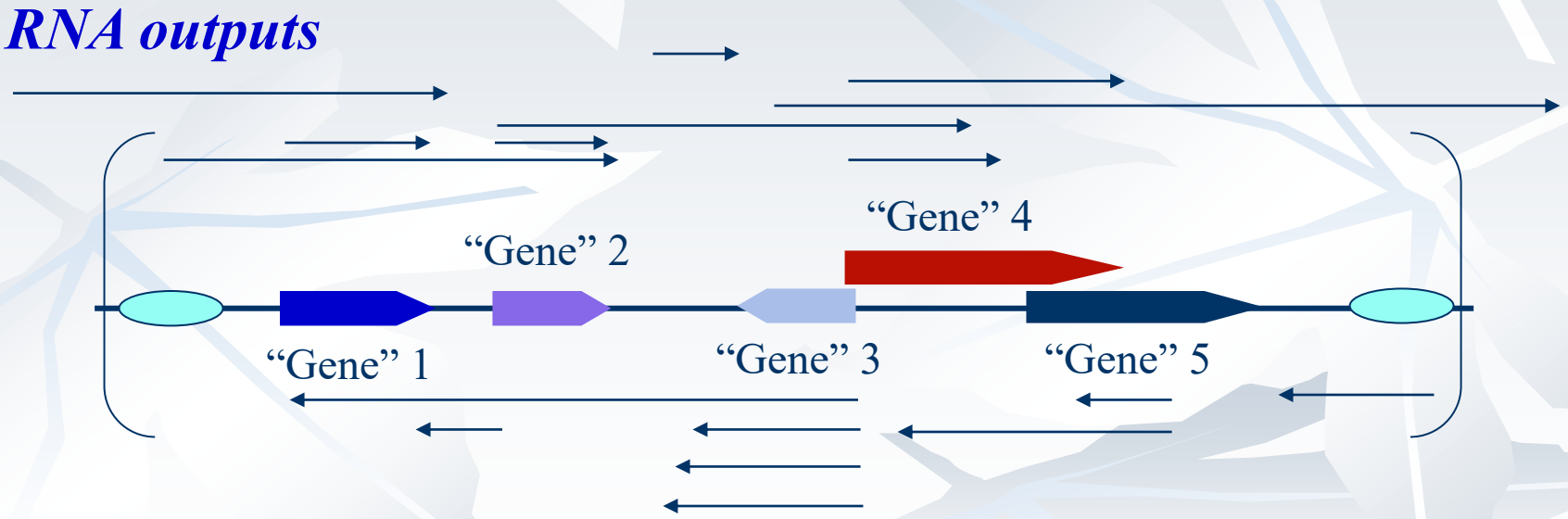


Inter-MAR Association Contributes to Transcriptionally Active Looping Events in Human *β-globin* Gene Cluster

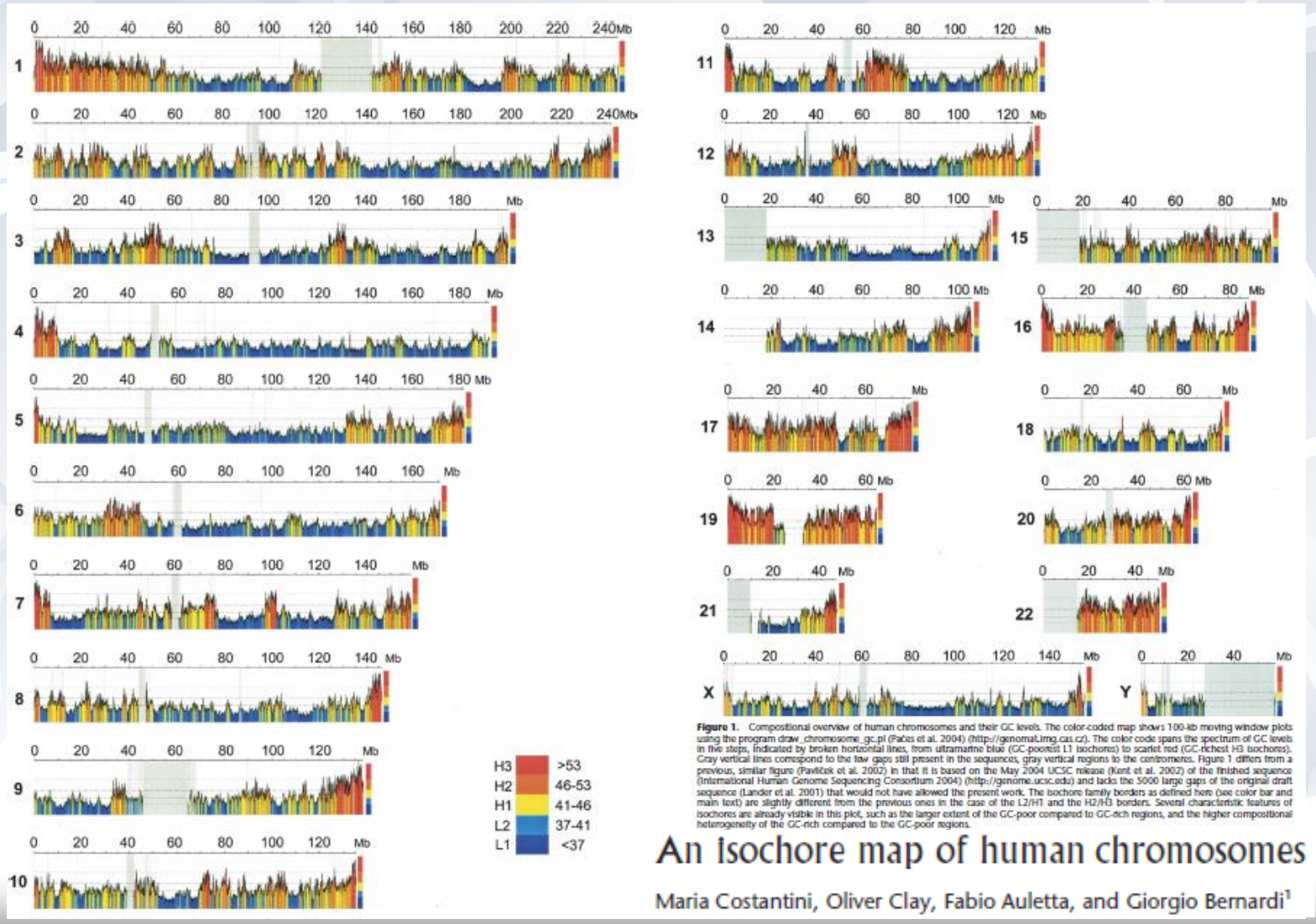
Li Wang, Li-Jun Di, Xiang Lv, Wei Zheng, Zheng Xue, Zhi-Chen Guo, De-Pei Liu*, Chi-Chuan Liang

February 2009 | Volume 4 | Issue 2 | e4629

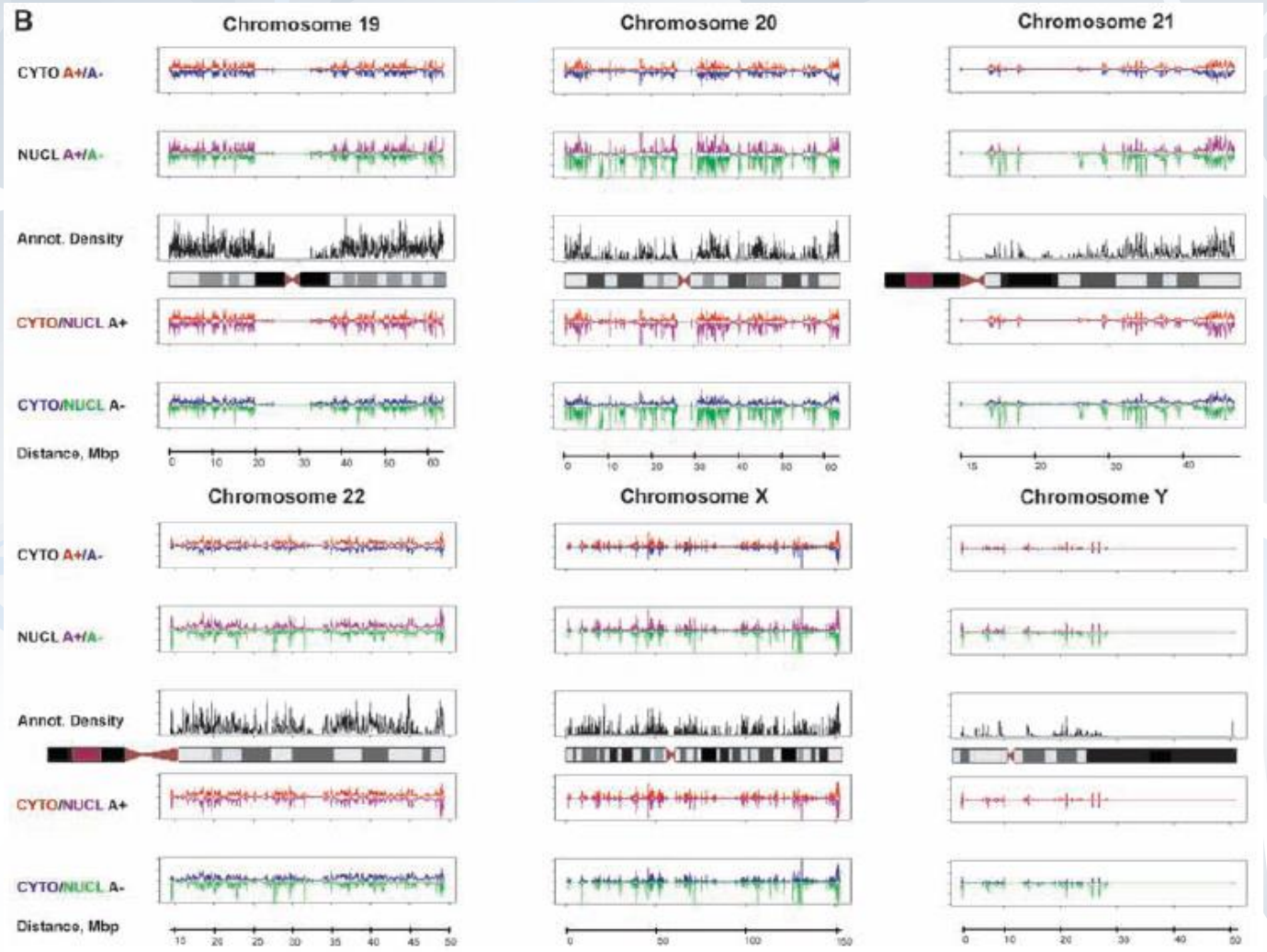
Principle 3: Gene data files are clustered into higher-order "folders" along a chromosome. This arrangement enables different types of RNAs to be encoded on both strands.



Principle 4: Gene folders/ALUs are in turn arranged into “superfolders.”

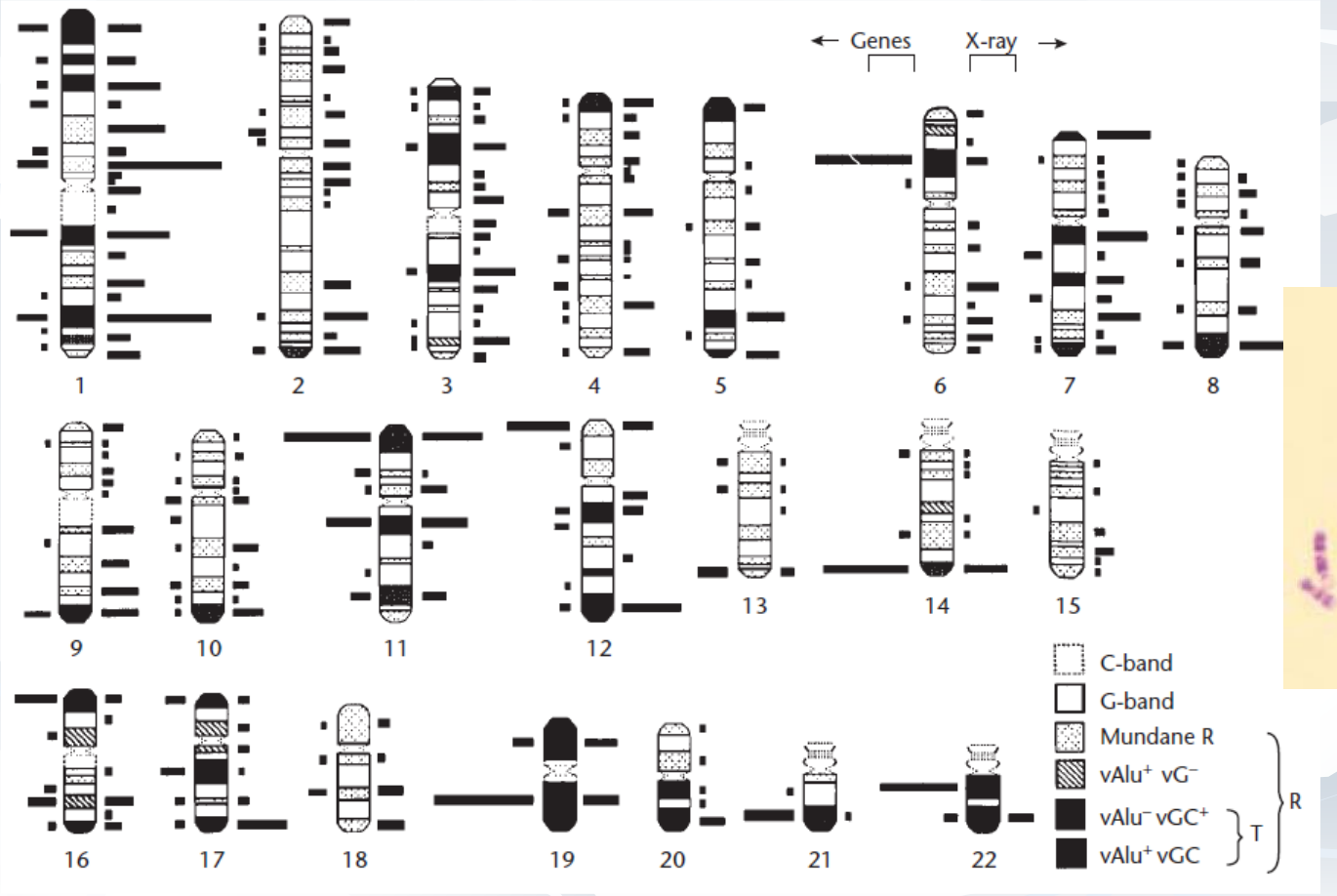


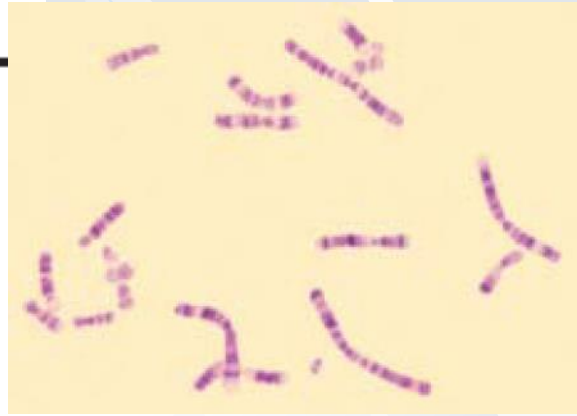
Different “superfolders” encode different classes of RNA outputs.



Transcriptional Maps of 10 Human Chromosomes at 5-Nucleotide Resolution
Jill Cheng, *et al.*
Science **308**, 1149 (2005);

And chromosome "superfolders" are in turn ordered into banding patterns…





Chromosomal Bands and Sequence Features

ENCYCLOPEDIA OF LIFE SCIENCES © 2005,

Gerald P. Holmquist, *City of Hope, Duarte, California, USA*

…such as those of CpG islands.

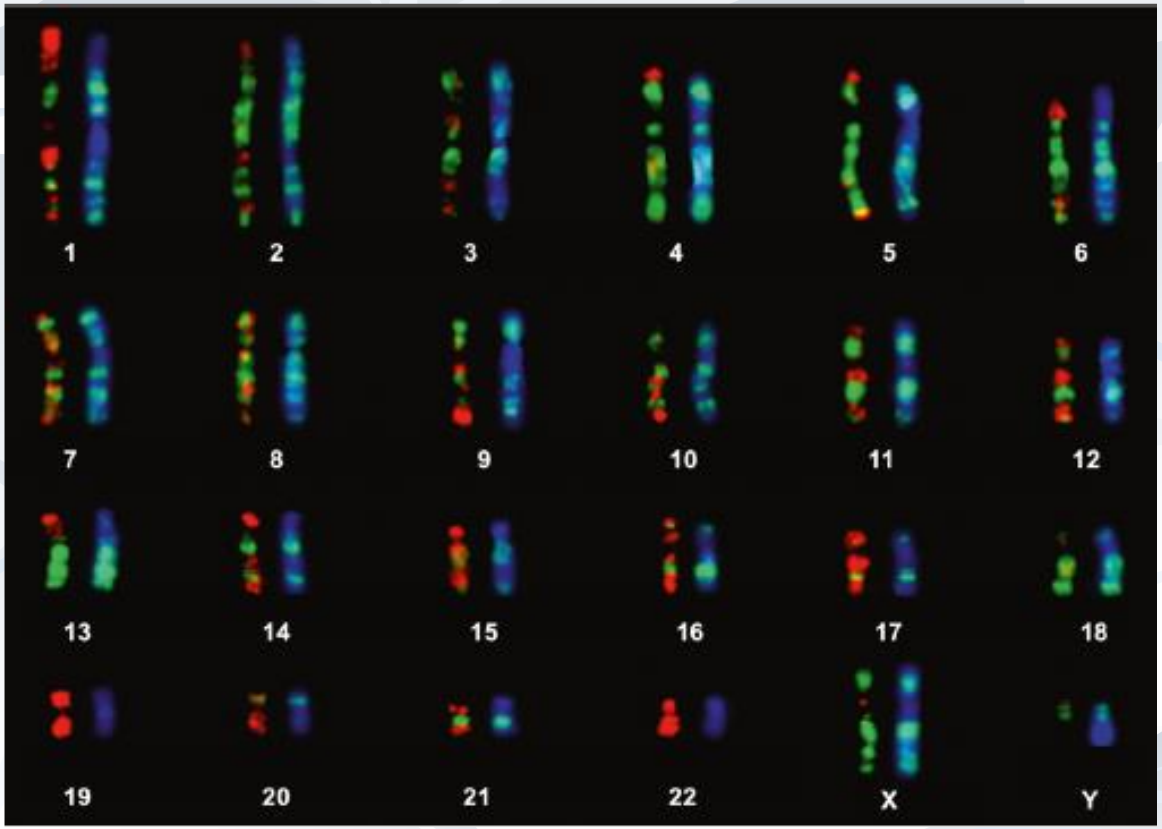


Fig. 2 Fluorescence in situ hybridisation (FISH) reveals the distribution of CpG islands across the human genome. For each metaphase chromosome, the hybridisation signal from CpG islands (red) is shown on the left of each pair. 4,6-Diamidino-2-phenyl indole (DAPI)-stained chromosomes are on the left. Late replicating G-bands are shown in green. Modified from Craig and Bickmore (1994)

Patterns in the genome

Wendy A. Bickmore [1]

Heredity (2019) 123:50–57

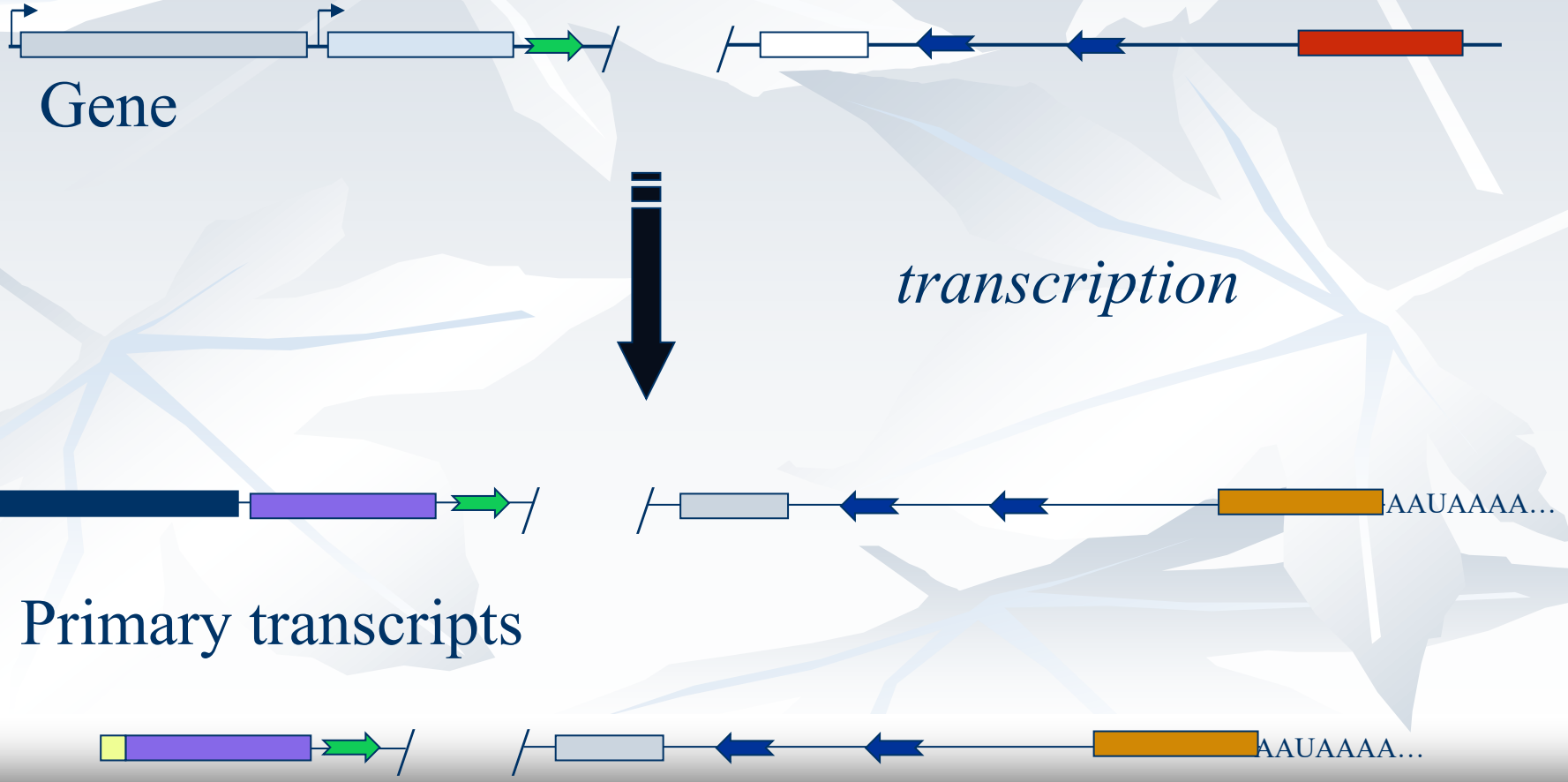
As a consequence of these results, a physical description of the "gene" is currently lacking. What we do know is that each DNA region:

- Is hierarchically ordered;
- Has "multilevel optimization" of many different types of codes; and
- Is connected by "coding chains" with "genes" on the same and other chromosomes.

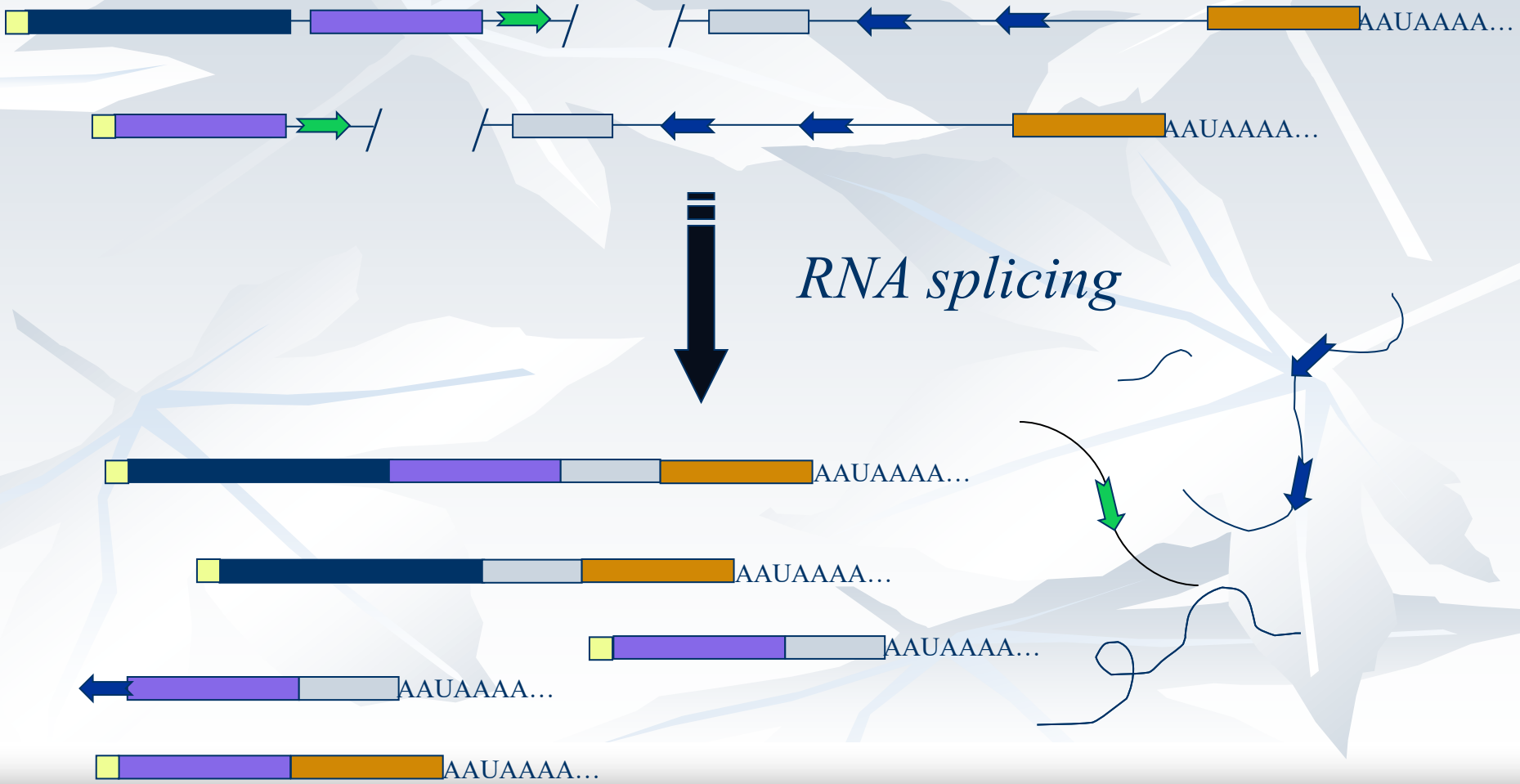
Wskazówka 2

The Second Clue

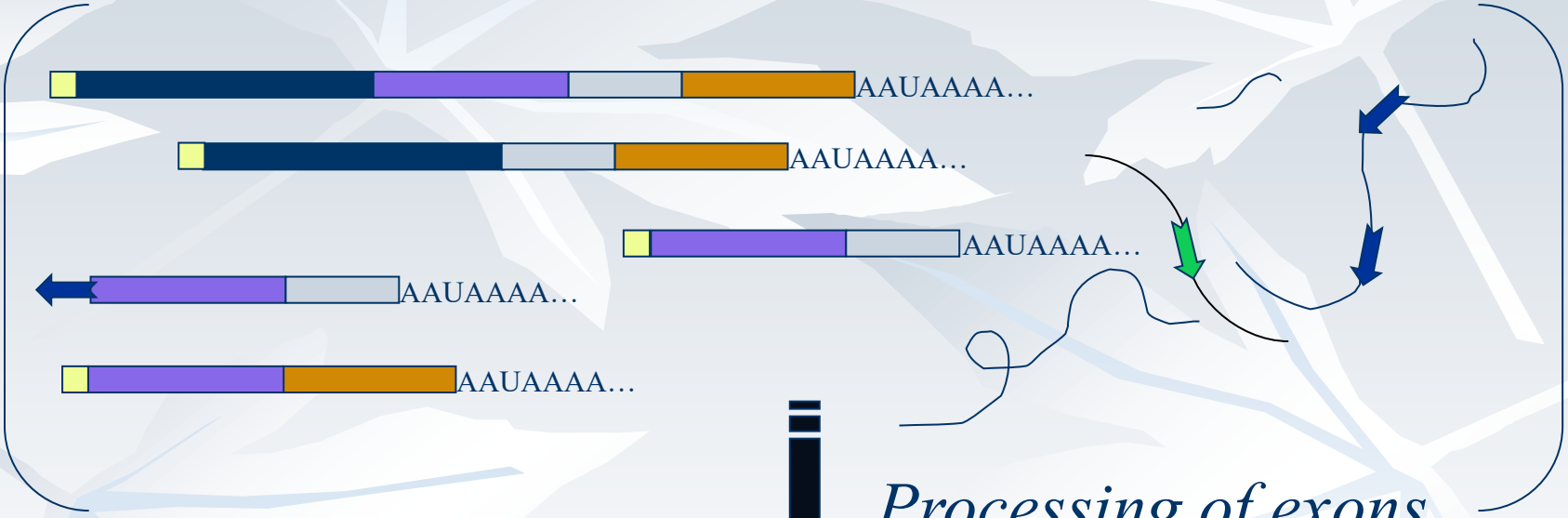
Other key pieces of evidence also began to accumulate. For example, it was found that some "genes" can potentially encode many different transcripts (over 1,000,000 in one case!)



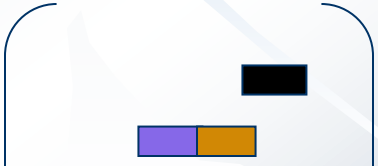
And the splicing of RNAs generates yet more “gene” products



In addition, it was soon realized that the “junk” sections of RNAs are processed into a host of functional sequences



Processing of exons and introns



ncRNAs
(various roles)

snoRNAs
(RNA editing)

microRNAs
(regulatory)

And now it is known that cellular pathways literally rewrite genetic scripts to make new transcripts and proteins, a widespread phenomenon called “RNA editing”

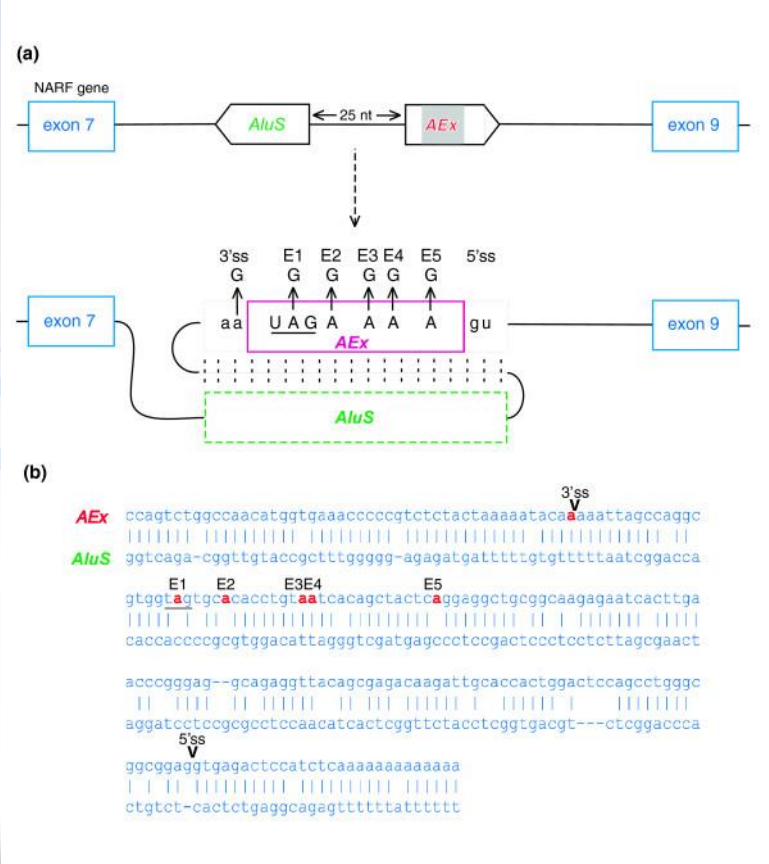


Fig. 1, Lev-Maor, G. et al., 2007. RNA-editing-mediated exon evolution. Genome Biology 8(2): R29.

Indeed, ribosomal and transfer RNAs must be highly edited in order to become functional in all known taxa

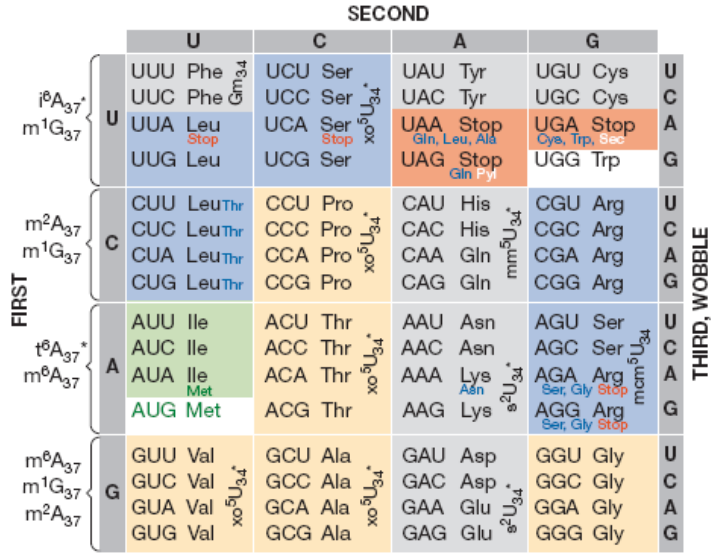


Fig 1 | Universal genetic code. The 64 codes are associated with the transfer RNA (tRNA) modifications that are important for decoding and/or translocation. Twofold degenerate amino-acid codes are highlighted in grey and fourfold degenerate codes are highlighted in tan. Amino acids with six codons are highlighted in blue. The threefold degenerate codons of Ile are highlighted in green, whereas the single codons of Met and Trp are highlighted in white. The three stop codons are highlighted in orange. Non-canonical codon use by some organisms and the mitochondrion is shown by using a small font for the amino acids (blue) or translational stop codons (red). The modified nucleoside abbreviations are defined in the text. Selenocysteine (Sec) and pyrrolysine (Pyl) codons are denoted in white. In the mitochondrion, $tRNA^{Met}$ responds to AUG and AUA, which is not used as an Ile codon (Agris *et al*, 2007; Szymański & Barciszewski, 2007; Björk *et al*, 1987).

Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications

Paul F. Agris

EMBO *reports* VOL 9 | NO 7 | 2008

tRNA's Wobble Decoding of the Genome: 40 Years of Modification

Paul F. Agris*, Franck A. P. Vendeix and William D. Graham

J. Mol. Biol. (2007) 366, 1–13

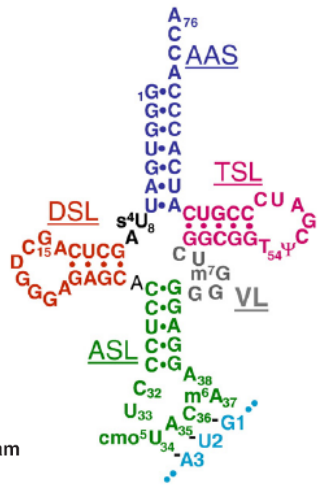


Figure 2. tRNA primary sequence, secondary structure, and codon binding. The sequence and secondary structure of *E. coli* $tRNA^{Val}$. The physical and functional domains of the *E. coli* $tRNA^{Val}_{UAC}$ sequence and secondary structure are the amino acid-accepting stem, AAS (dark blue), the dihydrouridine stem and loop, DSL (red), the anticodon stem and loop, ASL (green), the variable loop, VL (gray), and the thymidine stem and loop, TSL (purple). The modified nucleosides in this tRNA are: s^4U, 4-thiouridine; D, dihydrouridine; cmo^5U, uridine-5-oxyacetic acid; m^6A, N6-methyladenosine; m^7G, 7-methylguanosine; ribothymidine, T; and pseudouridine, Ψ. Because of the wobble nucleoside modification, cmo^5U_{34}, *E. coli* $tRNA^{Val}_{UAC}$ is capable of decoding all of the fourfold degenerate valine codons.[32–34] The tRNA is shown binding the cognate codon for valine, GUA, in light blue.

Clearly, a "gene" provides the substrate for many types of information that are layered on by the cell. In fact…

- Many RNAs, because of being rearranged and edited,* do not mirror *any DNA sequence;

- The RNA-level codes that are formed are often topological in nature; and

- Many RNA-level codes are sequence-independent.

Wskazówka 3

The Third Clue

So-called junk DNA elements are replete with experimentally demonstrated functions:

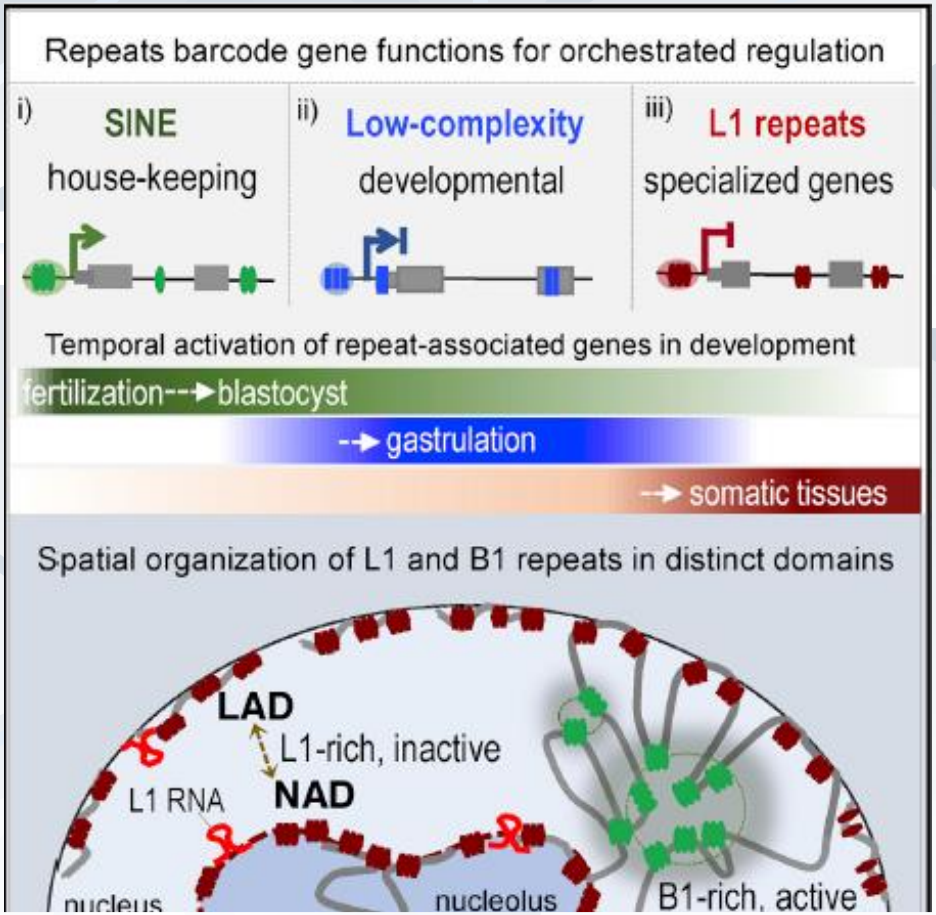
Highlights

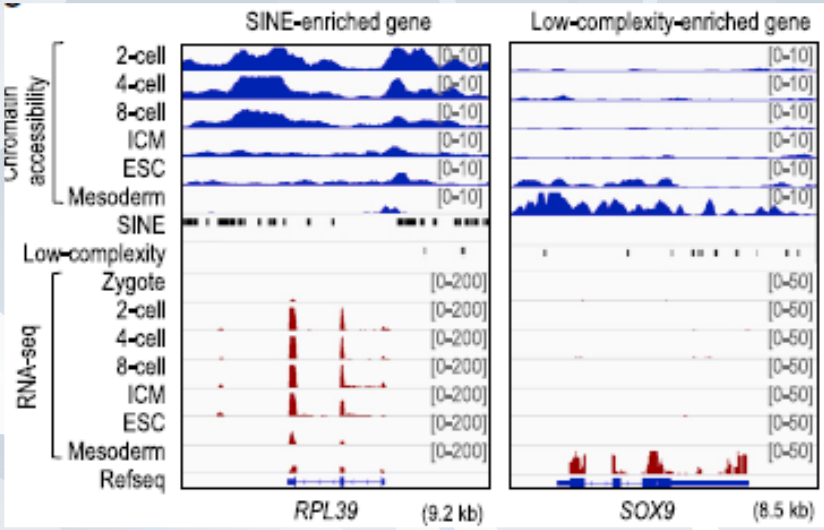
- SINE, L1, and low-complexity repeats barcode genes with distinct functions
- Genomic repeats dictate the time and level of gene expression during development
- L1-enriched genes are sequestered in the inactive NAD/LAD domains for silencing
- L1 RNA promotes the nuclear localization and repression of L1-enriched genes

Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation

J. Yuyang Lu, Wen Shao, Lei Chang, ...,
Miguel Ramalho-Santos, Yujie Sun,
Xiaohua Shen

Lu et al., 2020, Cell Reports *30*, 3296–3311

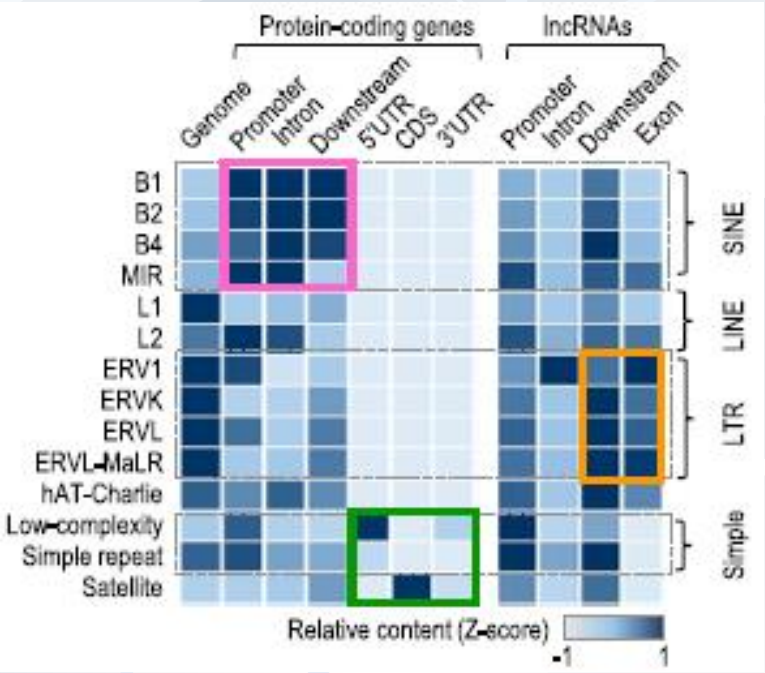


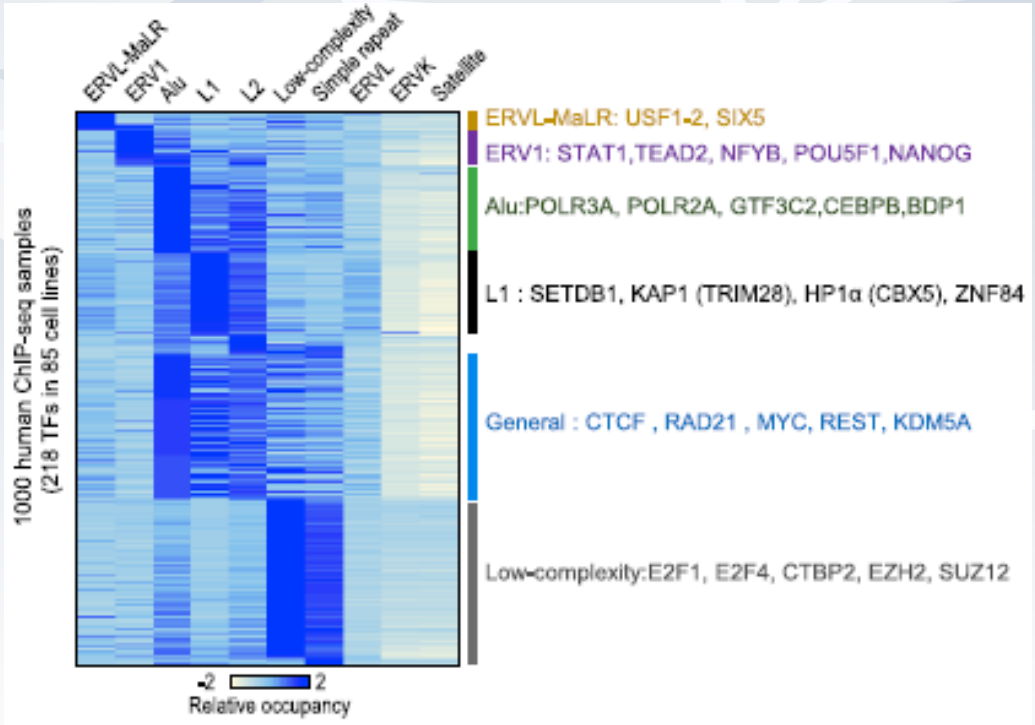


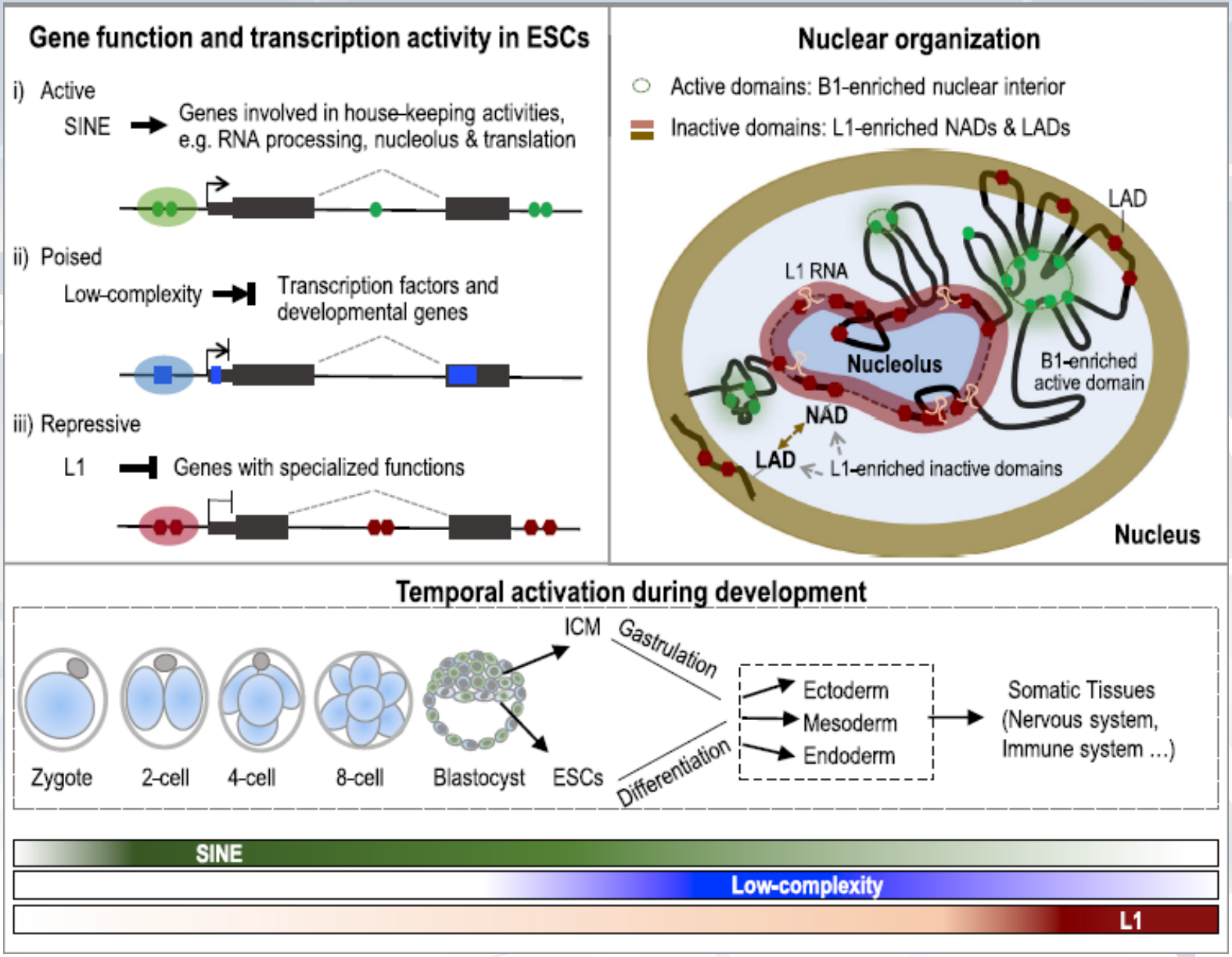
Chromatin accessibility of genic repeats

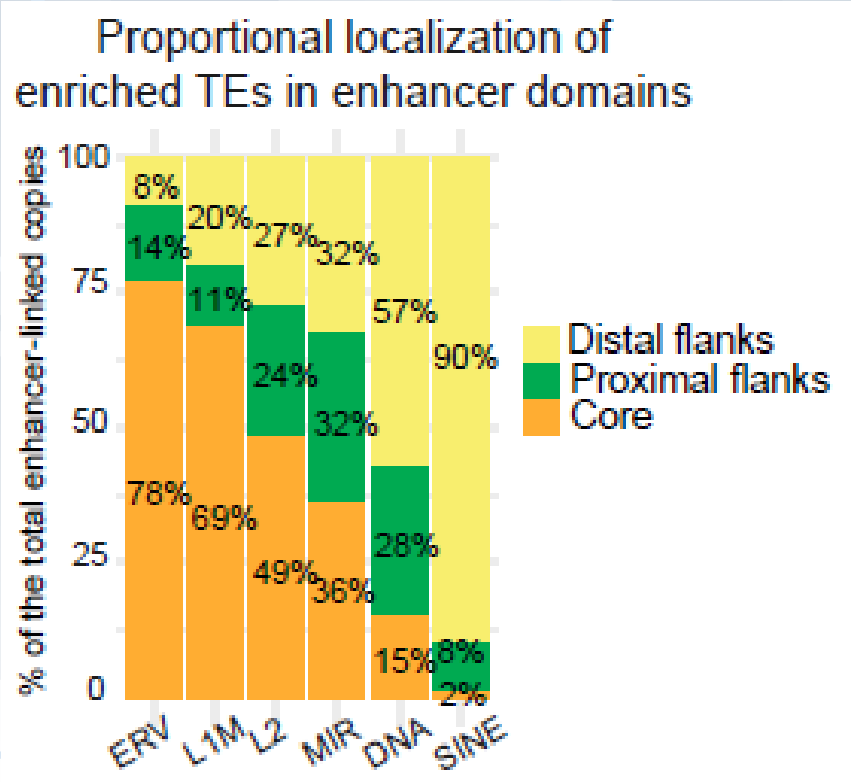
	2-cell					Pluripotent		Lineage-specific		
	Early-I	Early	Late	4-cell	8-cell	ICM	ESC	Meso	Endo	NPC
SINE	0.8	1.6	3.0	2.4	2.4	1.3	1.2	0.8	0.8	0.6
B1	1.1	2.5	4.8	3.7	3.4	1.7	1.6	0.9	1.0	0.8
B2	0.7	1.3	2.5	1.9	2.0	1.1	0.9	0.5	0.5	0.4
B4	0.6	1.0	1.8	1.5	1.7	1.0	0.9	0.6	0.6	0.5
ERVL	0.5	1.0	1.9	1.6	1.5	0.7	0.6	0.3	0.4	0.3
ERV1	0.4	0.7	1.1	0.9	1.0	0.8	0.7	0.4	0.5	0.3
ERVK	0.2	0.4	0.6	0.6	0.7	0.4	0.2	0.1	0.2	0.2
ERVL-MaLR	0.2	0.4	0.7	0.7	0.7	0.3	0.3	0.2	0.2	0.1
L1	0.1	0.1	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1
L2	0.6	0.8	1.1	1.0	1.3	1.0	1.0	0.9	0.8	0.7
Satellite	0.8	0.6	0.8	0.8	1.3	0.8	1.1	0.5	0.6	0.4
Low-complexity	1.2	1.0	1.2	1.2	1.0	1.3	1.7	2.0	2.2	2.4

0 2 (observed / random)


Protein-coding genes
lncRNAs
Genome
Promoter
Intron
Downstream
5'UTR
CDS
3'UTR
Promoter
Intron
Downstream
Exon
B1
B2
B4
MIR
L1
L2
ERV1
ERVK
ERVL
ERVL-MaLR
hAT-Charlie
Low-complexity
Simple repeat
Satellite
SINE
LINE
LTR
Simple
Relative content (Z-score)
-1
1


ERVL-MaLR
ERV1
Alu
L1
L2
Low-complexity
Simple repeat
ERVL
ERVK
Satellite
1000 human ChIP-seq samples
(218 TFs in 85 cell lines)
ERVL-MaLR: USF1-2, SIX5
ERV1: STAT1,TEAD2, NFYB, POU5F1,NANOG
Alu:POLR3A, POLR2A, GTF3C2,CEBPB,BDP1
L1 : SETDB1, KAP1 (TRIM28), HP1α (CBX5), ZNF84
General : CTCF , RAD21 , MYC, REST, KDM5A
Low-complexity:E2F1, E2F4, CTBP2, EZH2, SUZ12
-2
2
Relative occupancy


Gene function and transcription activity in ESCs
i) Active
SINE
Genes involved in house-keeping activities, e.g. RNA processing, nucleolus & translation
ii) Poised
Low-complexity
Transcription factors and developmental genes
iii) Repressive
L1
Genes with specialized functions
Nuclear organization
Active domains: B1-enriched nuclear interior
Inactive domains: L1-enriched NADs & LADs
LAD
L1 RNA
Nucleolus
B1-enriched active domain
NAD
LAD
L1-enriched inactive domains
Nucleus
Temporal activation during development
ICM
Gastrulation
Differentiation
Zygote
2-cell
4-cell
8-cell
Blastocyst
ESCs
Ectoderm
Mesoderm
Endoderm
Somatic Tissues (Nervous system, Immune system …)
SINE
Low-complexity
L1



Specific subfamilies of transposable elements contribute to different domains of T lymphocyte enhancers

Mengliang Ye[a], Christel Goudot[a], Thomas Hoyler[a], Benjamin Lemoine[b], Sebastian Amigorena[a,1], and Elina Zueva[a,1]

www.pnas.org/cgi/doi/10.1073/pnas.1912008117

Predicted TF or TF family	logo	Enhancer e-value	Enhancer %	Gene desert e-value	Gene desert %	similar sequence in the consensus
ORR1						
ETS (Etv-Ets-Gabpa)		2.6e-244	50%	5.1e-30	22%	CAGGAAGT(T/G)
		2.5e-92	49%	5.1e-77	27%	CAGGAAGT(T/G)
		4.5e-28	46%	-	-	TTCCTCT
RUNX (1,2,3)		3.9e-64	16%	-	-	TGTGGTTT (AAACCACA)
Lin54		4.8e-31	27%	4.7e-22	21%	TTTGAATG (CATTCAAA)
Max_Myc		6.3e-18	30%	5.5e-11	14%	ACACTTGGT
MTD						
E2f/Erf_Fli1		2.8e-56	60%	5.1e-30	27%	TTCCTGC (GCAGGAA)
		1.5e-46	53%	-	-	TTCCTGC (GCAGGAA)
		5.6e-41	60%	2.1e-30	57%	TTCCTGC
Runx1		2e-61	43%	1.5e-15	29%	CTGTGGG (CCCACAG)
RMER						
Sp1, Klf, E2f2		1.3e-13	44%	-	-	TCCCTTCCCC
		4.6e-06	44%	-	-	CCCCTCCCC
Rel_Rela_Bcl6		9.8e-22	55%	-	-	TCCCTTCCCC
Tcf7_Lef1		5.8e-07	20%	-	-	AGACCAAC, TTTGGTCT
Tead3		6.8e-09	35%	-	-	ACCATACC
MTE						
ETS (Etv, Gabpa, Elk)		3.7e-44	41%	2e-17	36%	AGGAGAAA, AGGAGACA
Sp1, Klf		2.5e-18	30%	3.3e-13	26%	ACCCACCC
Zbtb26_Smad4		4.7e-07	15%	-	-	ATCTAGAAT
RLTR						
Klf1, RUNX		1.3e-10	53%	-	-	TGTGGTT
Prdm1_RelA		2.8e-06	32%	-	-	GAAAGTC
Zfp523_Zfp143		9.5e-07	34%	-	-	ACTAAAACA
MLT						
Rbpj		0.015	15%	-	-	TCCCCCCA
Sp1/2_, Klf		0.025	12%	-	-	CCCTCCC
Hic1		0.054	11%	-	-	GCCACC
Forkhead, Znf384		0.009	22%	-	-	AAATAAAT

MIR						
Zfp787		9.5e-27	24%	5.8e-13	13%	GGGCCTCAGTTTC (GGAAACTGAG)
Zfp768		4.5e-20	18%	-	-	
Tbp		1.7e-15	16%	-	-	GTAAAATG (CATTTTAC)
Nfat		4.5e-20	16%	6.3e-14	23%	GTAAAATGG
Nr4f2_Essra		4.70e-16	19%	-	-	GTGACCT (AGGTCAC)
Gata		2.6e-11	11%	-	-	AGATGA
L2						
Sry/Zfp422_ 384/Forkhead		3e-19	20%	6.3e-21	13%	AATAAA
		4.7e-48	18%	1.9e-24	10%	AAAAAACAAAAA
Sp1, Klf_Znf263		6.3e-51	15%	-	-	CCCCTCCCC
Fli1		3.1e-11	21%	-	-	AGGAG
		3.5e-46	12%	-	-	CACACA
L1						
Sry/Zfp422_ 384/Forkhead		1e-14	42%	-	-	TATTTTA (ATAAAAT)
		1.1e-34	37%	-	-	AAAAACAAA
Setbp1_Ahctf1		1.5e-81	42%	3e-22	12%	TATTTTAA
Sp1/Klf, Znf148		3.4e-64	40%	ns		CCCCCCT(CT)CCCC
		7.4e-95	26%	9.8e-10	16%	CACACCCA
DNAhat						
Srebf1		7.40e-38	54%	-	-	GGGTCACCACAA (TTGTGGTGACCC)
		5.50e-48	51%	3.5e-31	40%	TTAAAGGGTC
Rxra_Rxrb_Zfp652		1.3e-142	46%	-	-	GACCCCT
B2						
Zfp384/Forkhead		1e-79	30%	1e-56	23%	ATAAAAATAAA
Fos_Jun		7.1e-205	47%	-	-	GATGGCTCA
B1						
		1e-300	10%	1e-300	2%	AAAAAACAAAA
		3.5e-138	18%	2.3e-63	9%	AAAAAACAAAA
B4						
		1e-300	13%	9.8e-300	11%	CACACACACACA

And many taxon-specific repeats have almost "synonymous" chromosomal locations:

Alu and B1 Repeats Have Been Selectively Retained in the Upstream and Intronic Regions of Genes of Specific Functional Classes

Aristotelis Tsirigos*, Isidore Rigoutsos*

PLoS Computational Biology
December 2009 | Volume 5 | Issue 12 | e1000610

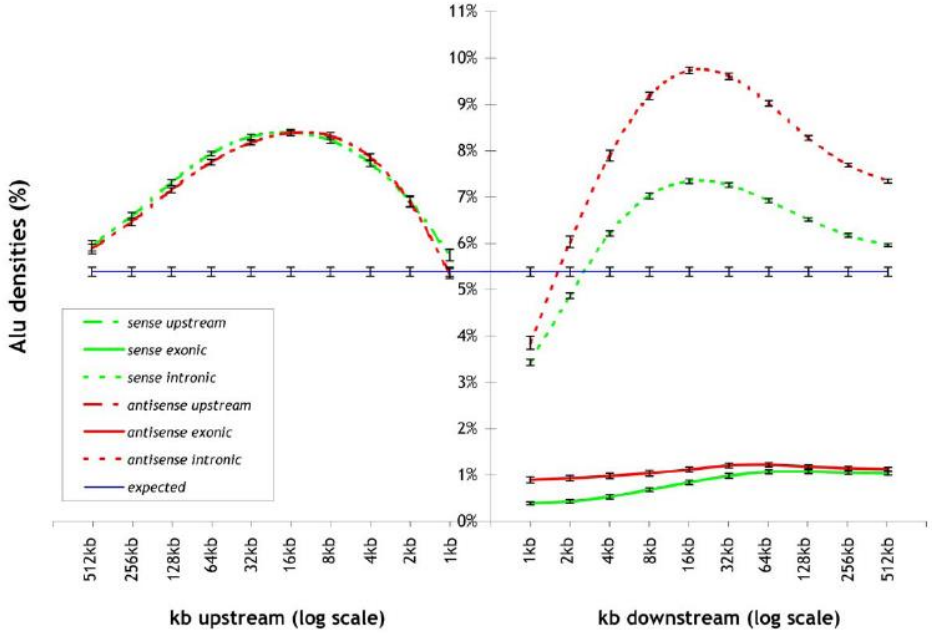


Figure 1. Alu densities upstream and downstream of known genes as a function of distance from the gene transcript start position.

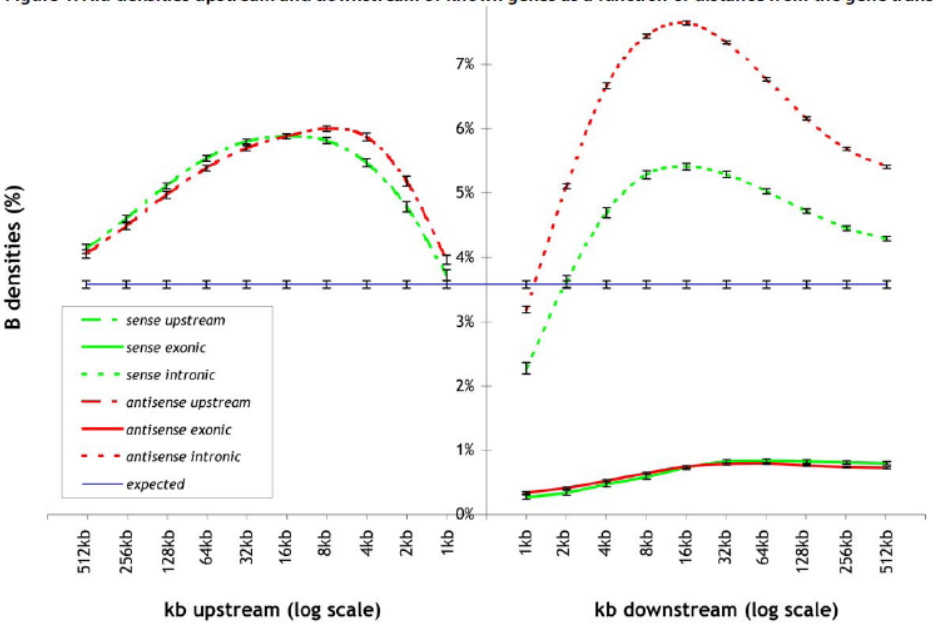
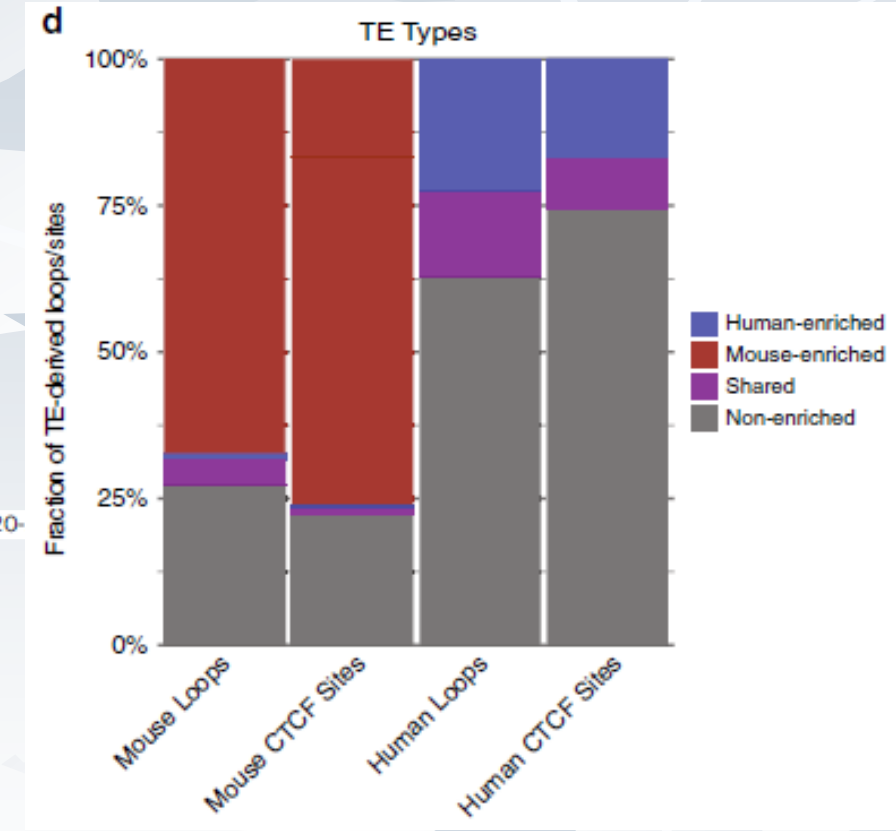


Figure 2. B element (B1, B2, B4) densities upstream and downstream of known genes as a function of distance from the gene transcript start position. Green and red curves correspond to B element instances in the sense and antisense orientation respectively.

Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes

Adam G. Diehl [1], Ningxin Ouyang[1] & Alan P. Boyle [1,2]✉

NATURE COMMUNICATIONS | (2020)11:1796 | https://doi.org/10.1038/s41467-020-15520-



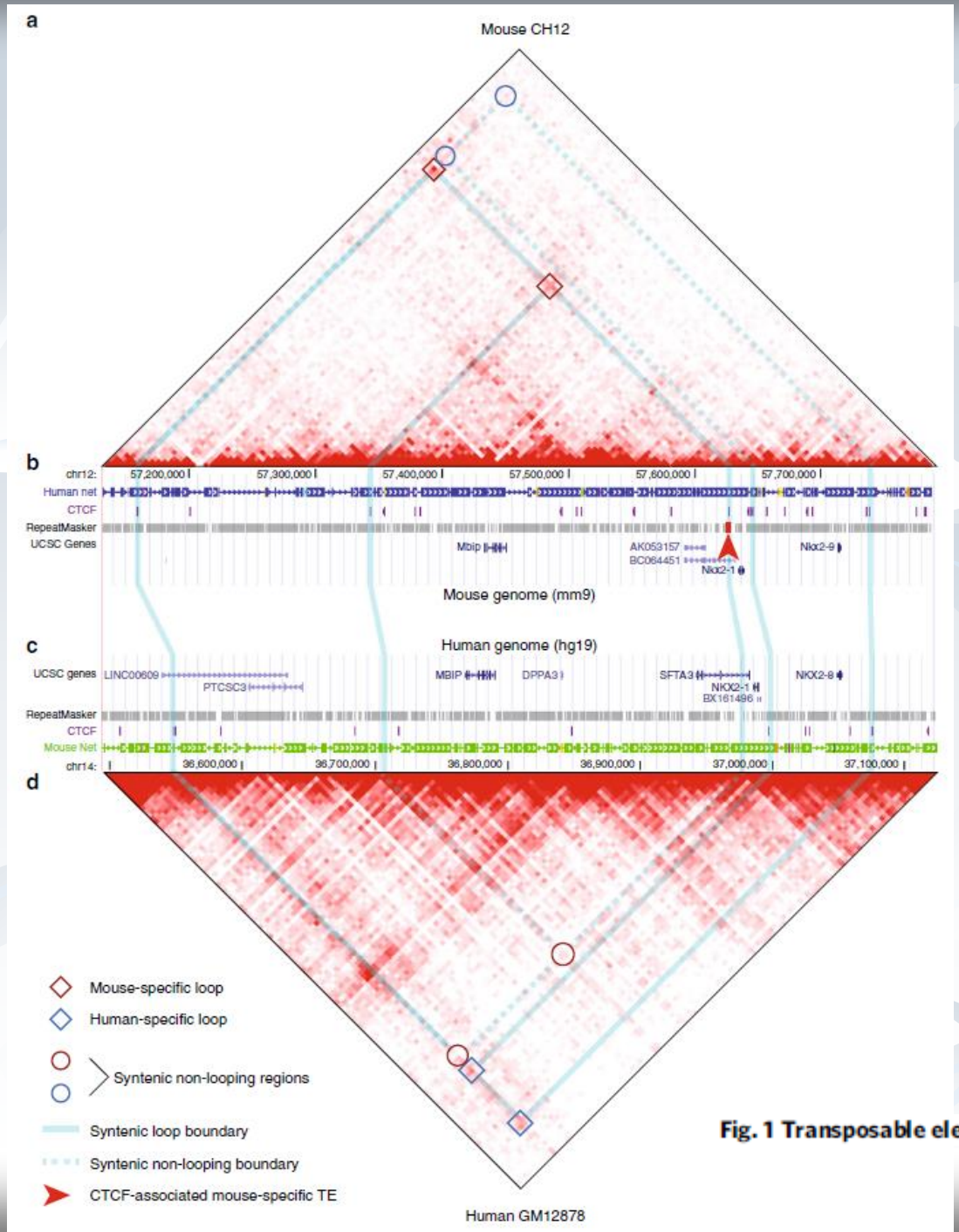
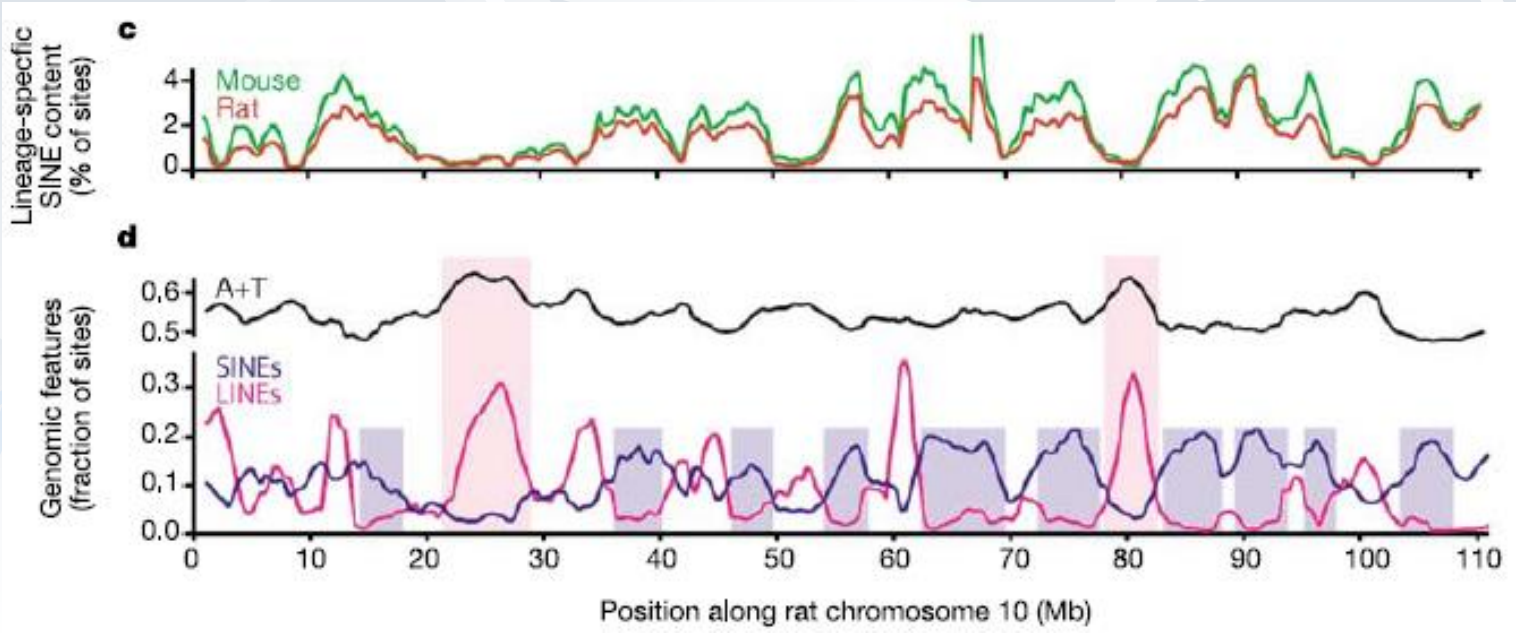


Fig. 1 Transposable element insertions create novel species-specific loop contacts.

The overall "data" pattern along a megafolder is the same *but* the species-specific details of the logic gates are different.



Genome sequence of the Brown Norway rat yields insights into mammalian evolution

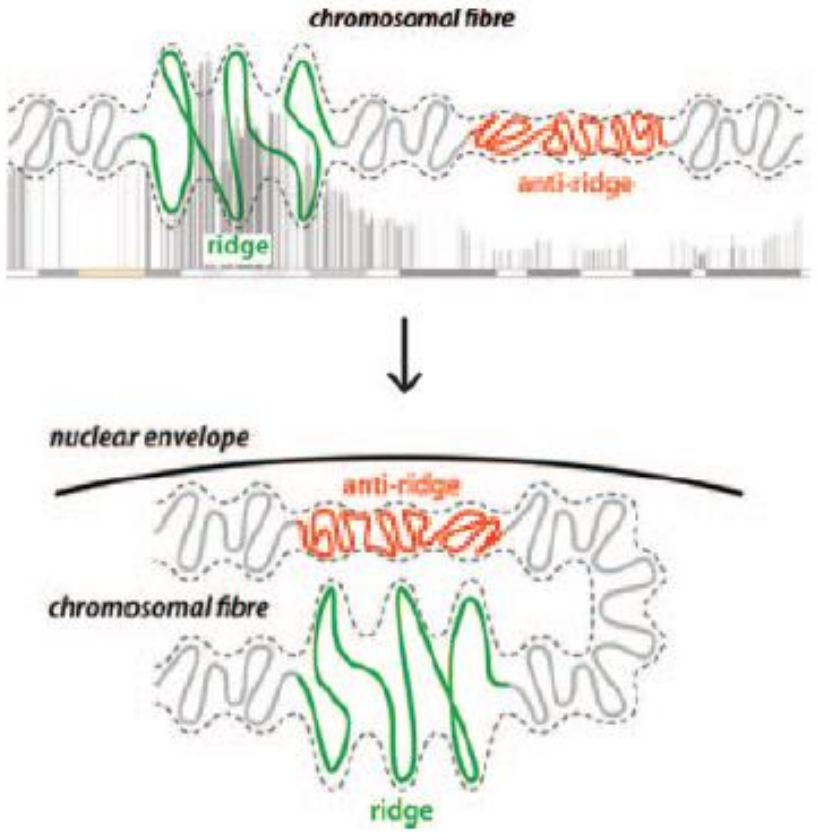
Rat Genome Sequencing Project Consortium*

NATURE | VOL 428 | 1 APRIL 2004

Wskazówka 4

The Fourth Clue

Co-expressed loci are clustered together along in the nucleus, sometimes to “create” genes



The Three-Dimensional Structure of Human Interphase Chromosomes Is Related to the Transcriptome Map[▽]

Sandra Goetze,[1]† Julio Mateos-Langerak,[1]† Hinco J. Gierman,[2] Wim de Leeuw,[3] Osdilly Giromus,[1] Mireille H. G. Indemans,[2] Jan Koster,[2] Vladan Ondrej,[4] Rogier Versteeg,[2] and Roel van Driel[1]*

MOLECULAR AND CELLULAR BIOLOGY, June 2007, p. 4475–4487

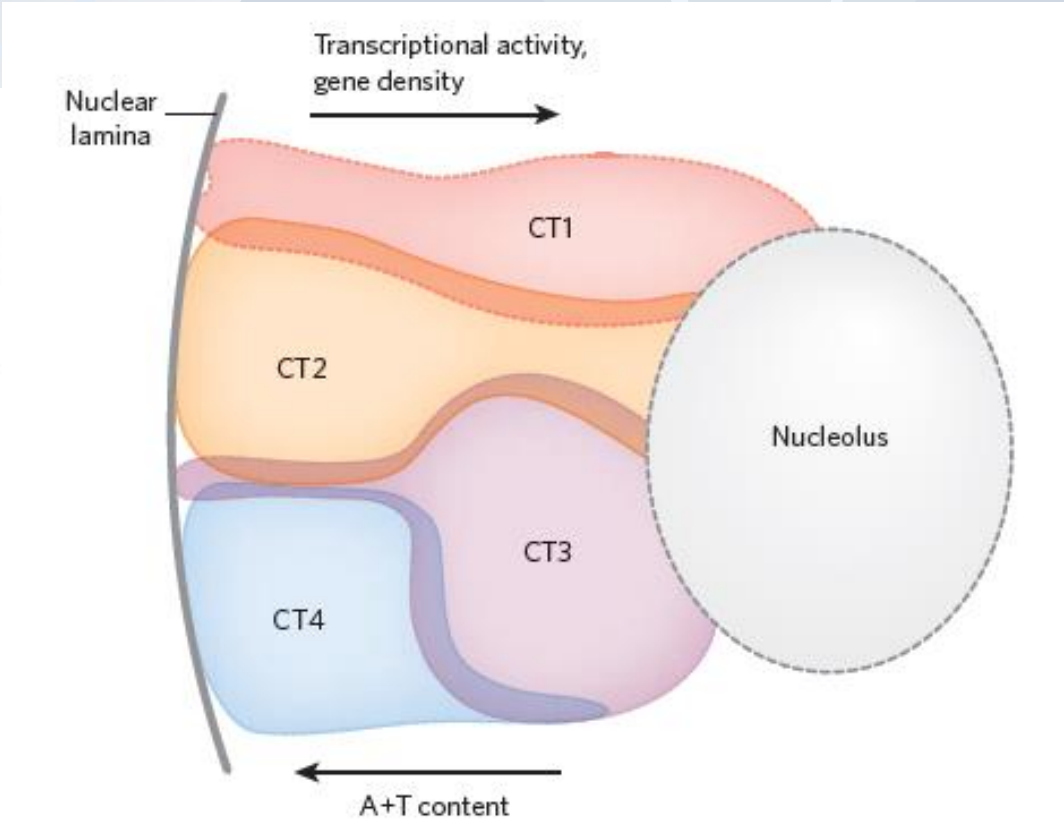


Figure 3 | Radial organization of chromosome territories within the nucleus regulates opportunities for chromatin crosstalk. The relative positions of chromosomes in an interphase nucleus depend on the proportion of genes and the A+T content. The opportunities for chromatin crosstalk between

Chromosome crosstalk in three dimensions

Anita Göndör[1] & Rolf Ohlsson[1]

NATURE|Vol 461|10 September 2009|

And these are in turn organized into "topologically-associating domains" that are cell-specific.

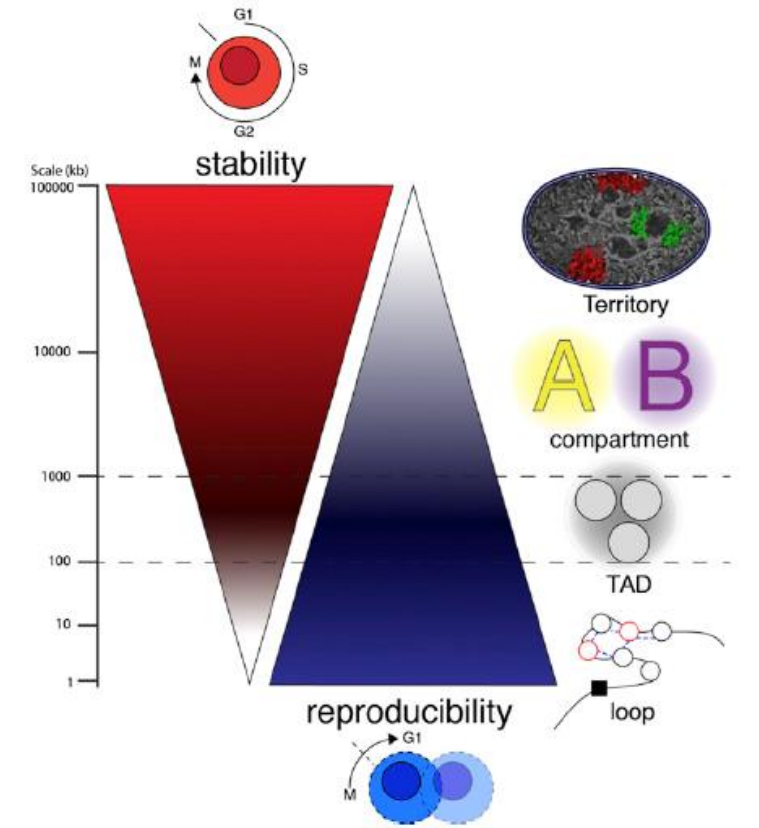


Figure 3. The Stability and Reproducibility of Chromosomal Interactions

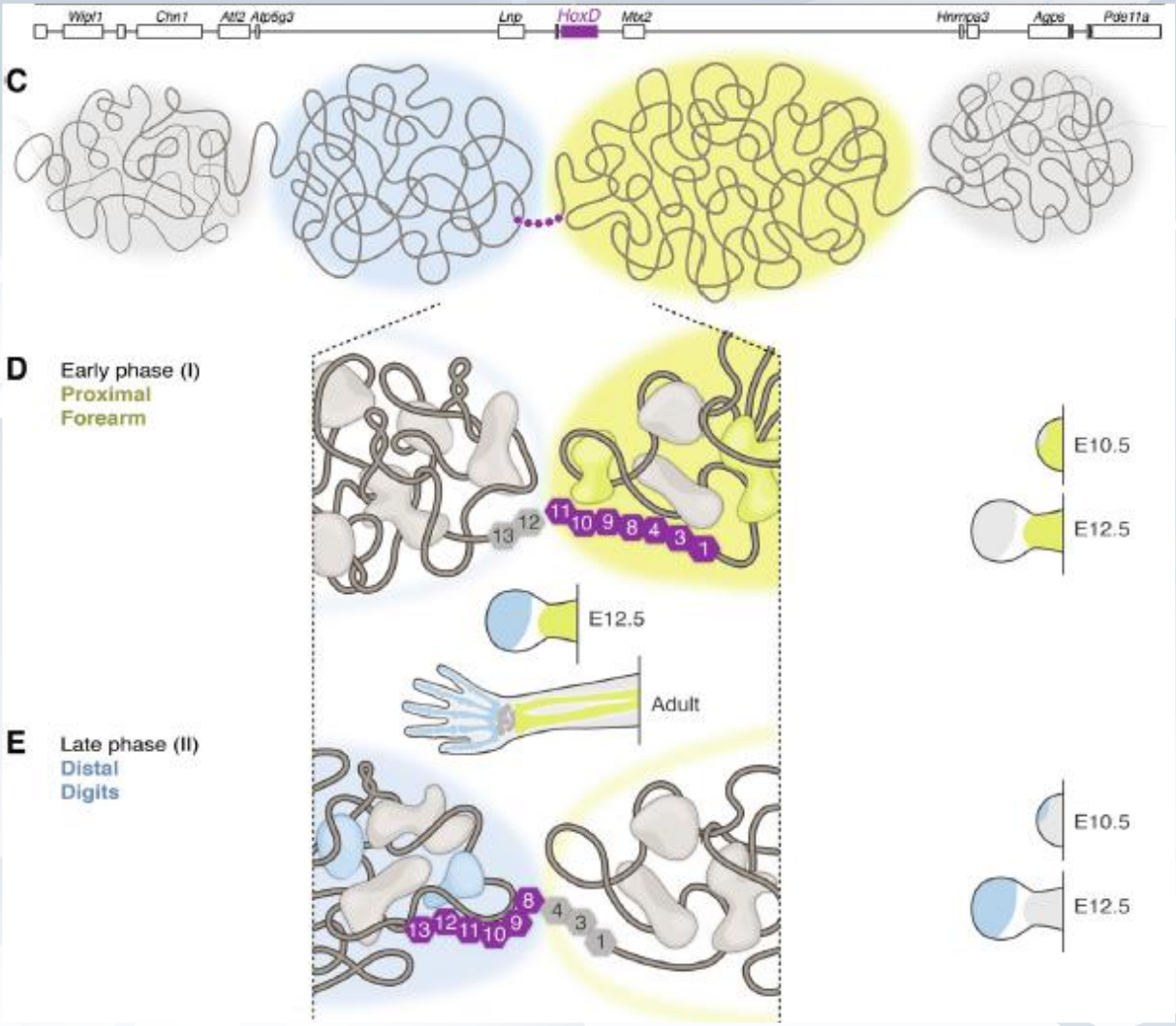
Chromosomal territories and compartments are very stable within one cell cycle of a given cell, but they are unlikely to be reproduced from one cell cycle to the next. Conversely, interactions between loops (within TADs) will be unstable and variable within each cell cycle, but this "instability" is reproducible from one cell cycle to the next. At the junction between stability and reproducibility, TADs confine looping, while maintaining the possibility of compartmentalization.

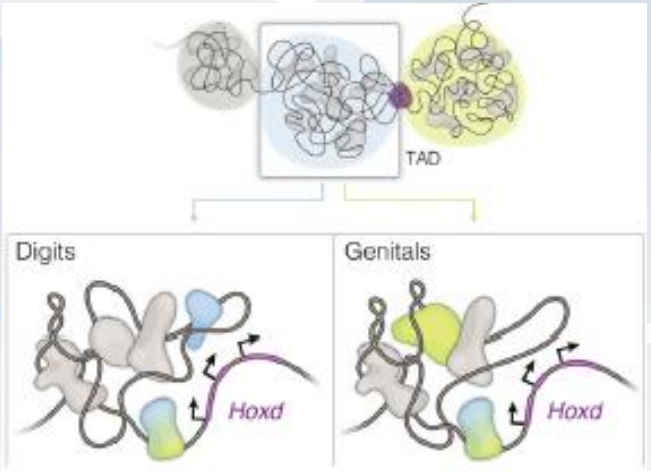
The Hierarchy of the 3D Genome

Johan H. Gibcus[1] and Job Dekker[1,*]

Molecular Cell *49*, March 7, 2013

A regulatory switch between two adjacent TADs underlies the bimodal regulation occurring at the *HoxD* locus during limb development.

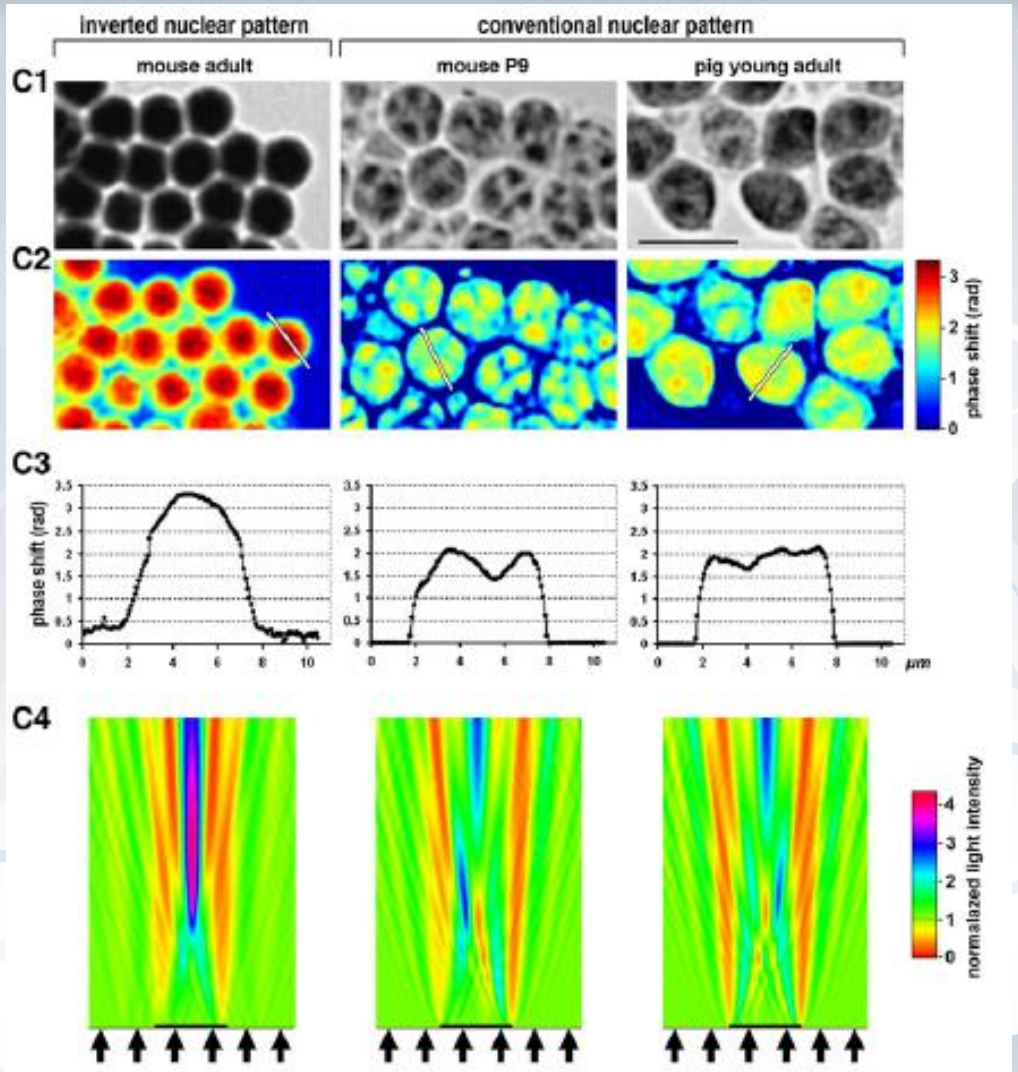




Structure, function and evolution of topologically associating domains (TADs) at *HOX* loci

Nicolas Lonfat [a,1], Denis Duboule [a,b,*]

http://dx.doi.org/10.1016/j.febslet.2015.04.024



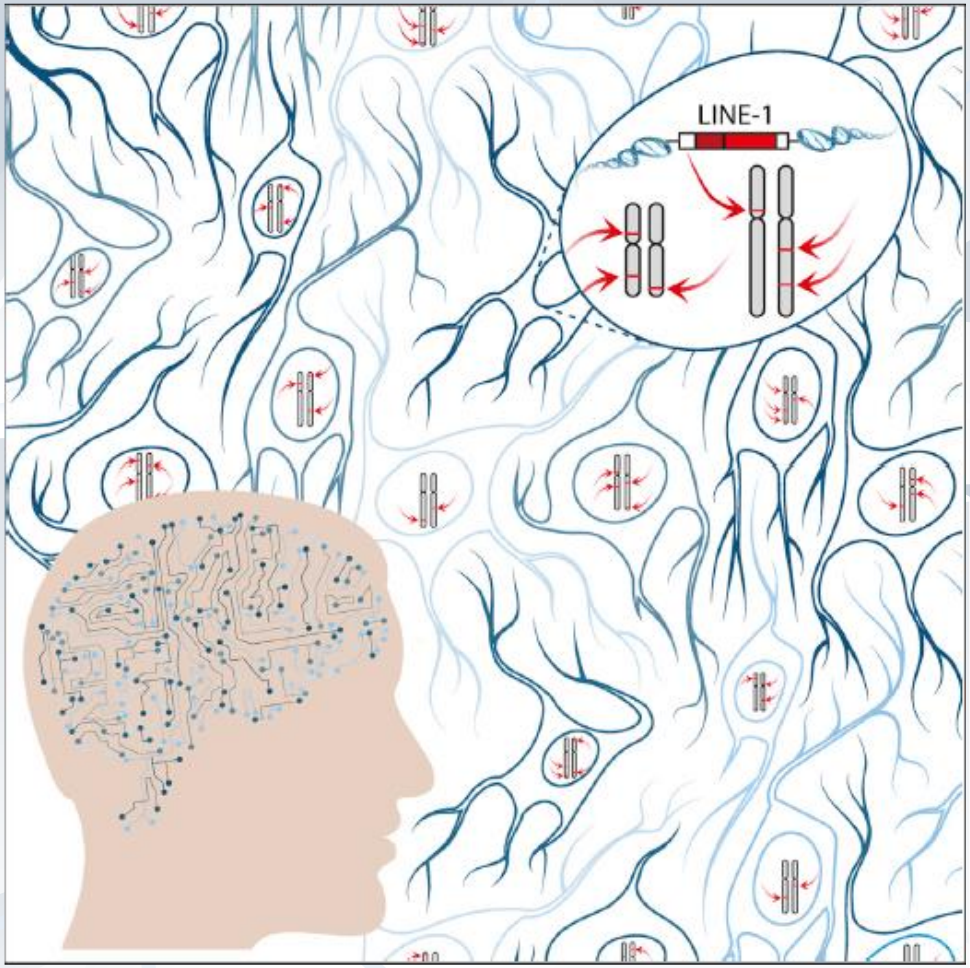
Nuclear Architecture of Rod Photoreceptor Cells Adapts to Vision in Mammalian Evolution

Irina Solovei,[1] Moritz Kreysing,[2] Christian Lanctôt,[1,5] Süleyman Kösem,[1] Leo Peichl,[3] Thomas Cremer,[1,4] Jochen Guck,[2,*] and Boris Joffe[1,*]

Cell *137*, 356–368, April 17, 2009

DNA Sequences as Context-Dependent, Data-Storage Regions

Given all the evidence we now have available, a new model of "genes" is now emerging…



An estimated 13.7 somatic L1 insertions occur per hippocampal neuron, on average

Target-primed reverse transcription drives somatic L1 retrotransposition

Somatic L1 insertions sense oriented to introns are depleted in neurons and glia

Hippocampus genes and enhancers are strikingly enriched for somatic L1 insertions

In Brief

Somatic genome mosaicism among neurons has the potential to impact brain function. L1 retrotransposons mobilize extensively in hippocampal neurons, preferentially in hippocampally expressed loci, and are depleted from mature neurons when oriented in the most deleterious configuration to host genes, suggesting functional significance.

Ubiquitous L1 Mosaicism in Hippocampal Neurons

Kyle R. Upton,[1,6] Daniel J. Gerhardt,[1,6] J. Samuel Jesuadian,[1,6] Sandra R. Richardson,[1] Francisco J. Sánchez-Luque,[1] Gabriela O. Bodea,[1] Adam D. Ewing,[1] Carmen Salvador-Palomeque,[1] Marjo S. van der Knaap,[2] Paul M. Brennan,[3] Adeline Vanderver,[4] and Geoffrey J. Faulkner[1,5,*]

Cell *161*, 228–239, April 9, 2015

…and it is one where we have to attribute the "informing" principle to something other than DNA.

The epigenome and top-down causation

P. C. W. Davies*

Interface Focus (2012) **2**, 42–48
doi:10.1098/rsfs.2011.0070

THE EPIGENOME AS A VIRTUAL OBJECT

. . . we will look in vain for any particular physical object within the cell that we can identify as 'the epigenome.' In the case of epigenetics, *there is no physical headquarters*, no localized commanding officers issuing orders, no geographical nerve centre where the epigenomic 'programme' is stored and from where epigenomic instructions emanate to help run the cell. The epigenome is not to be found at a place and the ultimate information source of epigenetics cannot be located anywhere specifically; rather, it is distributed throughout the cell. To be sure, the epigenome is *manifested* in particular structures (histone tails, nucleosome patterns, methylation patterns, chromatin packing . . .), but it does not *originate* there. The epigenome is everywhere and nowhere; it is a global, systemic entity. Expressed more starkly, *the epigenome is a virtual object*. Given that it calls many, if not most, of the biological shots, its non-existence as a specific physical entity is deeply significant.

. . . Undeniably the genome provides the words, but the epigenome writes the play! For those