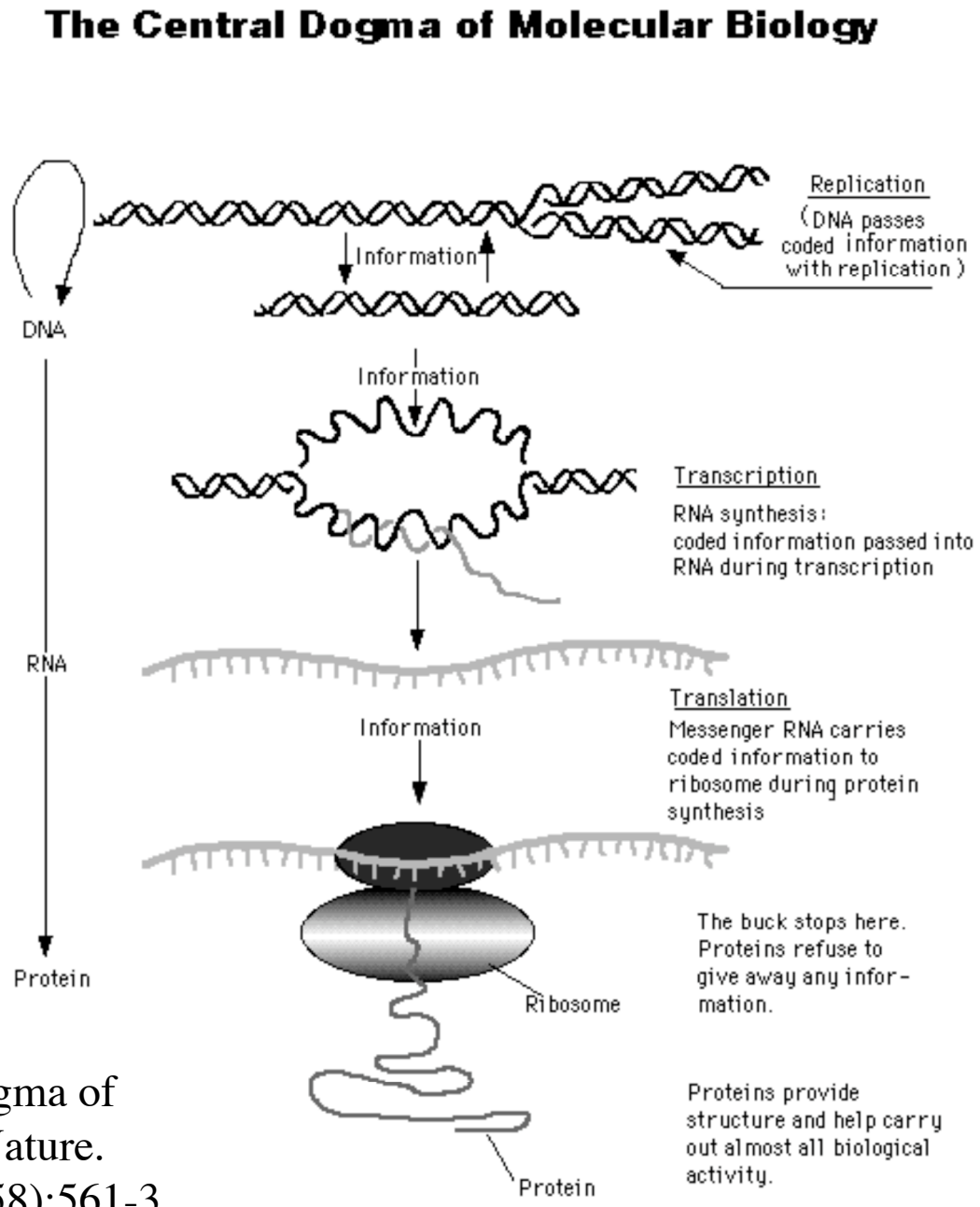
Revisiting the Central Dogma in the 21st Century

James A. Shapiro

University of Chicago

http://shapiro.bsd.uchicago.edu



Crick, F. Central dogma of molecular biology.Nature. 1970 Aug 8;227(5258):561-3.

What has changed since the central dogma?

Basic cell activities

- reverse transcription
- post-transcriptional RNA processing
- catalytic RNA
- genome-wide (pervasive) transcription and trans-splicing
- post-translation protein modification
- DNA proofreading and repair

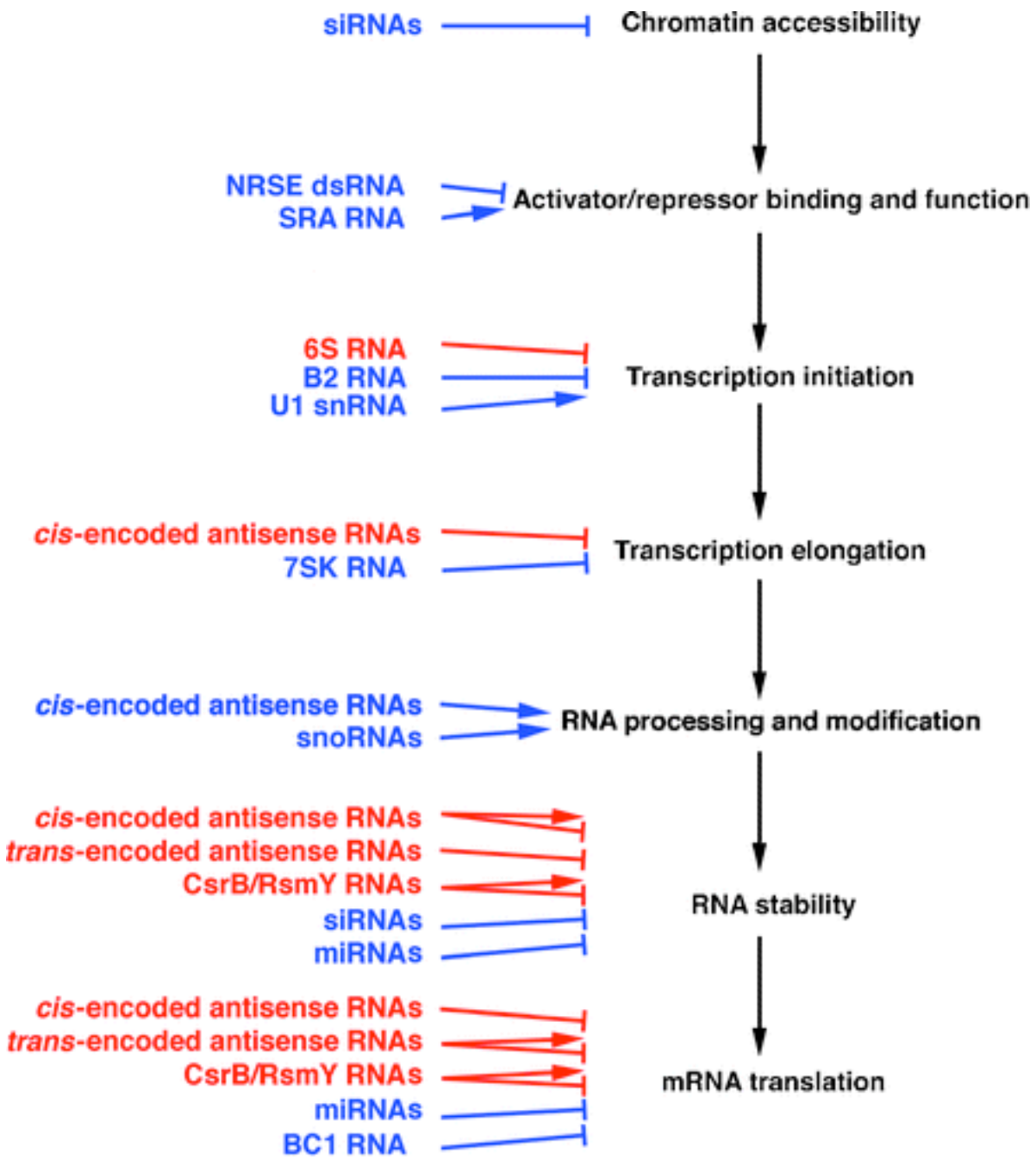
Cellular sensing and intercellular communication

- allosteric binding proteins (LacI)
- riboswitches & ribosensors
- surface and transmembrane receptors
- surface signals (Wnt, Notch)
- intercellular protein transfer
- exported signals (NO, CH_2=CH_2,HSLs, peptide pheromones, cytokines)
- internal monitoring (RecA, MutS, NMD, protein localization, cell cycle progress)

Cellular control regimes

- Feedback regulation circuits
- signal transduction networks
- second messengers (chemical symbols)
- checkpoints
- Chromatin formatting, reformatting and epigenetic regulation
- regulatory RNAs
- subnuclear localization

Small RNA control of protein synthesis



Storz G, Altuvia S, Wassarman KM An abundance of RNA regulators. Annu Rev Biochem. 2005;74:199-217.

Composite organization of macromolecules

- multidomain structure of proteins and RNAs (functionally differentiated domains)
- introns and exons; splicing (alternative and regulated)
- complex nature of genomic elements
- repetitive DNA

Natural genetic engineering

- DNA translocation across membranes (HGT)
- homology-dependent and -independent recombination
- DNA rearrangement modules
- retrotransposition, retrotransduction and reverse splicing
- protein engineering by DNA rearrangements and targeted mutagenesis (bacteria, trypanosomes, lymphocytes)
- genome reorganization in normal life cycles (mating-type switches, chromatin diminution, ciliated protozoa, antigen receptors)
- response to stress and other stimuli
- targeting within the genome

Molecules involved in cellular information transfer

1970:

- (DNA --> 2X DNA) --> RNA --> Protein --> Phenotype (Crick, Central dogma of molecular biology)

2008:

- DNA + 0 --> 0
- DNA + Protein + ncRNA --> chromatin
- Chromatin + Protein + ncRNA --> DNA replication, chromatin maintenance/reconstitution
- Protein + RNA + lipids + small molecules --> signal transduction
- Signals + Chromatin + Protein --> RNA (primary transcript)
- RNA + Protein + ncRNA --> RNA (processed transcript)
- RNA + Protein + ncRNA --> Protein (primary translation product)
- Protein + nucleotides + Ac-CoA + SAM + sugars + lipids --> Processed and decorated protein
- DNA + Protein --> New DNA sequence (mutator polymerases)
- Chromatin + Protein --> New DNA structure (DNA-based rearrangements)
- RNA + Protein + chromatin --> New DNA structure (retrotransposition, retroduction, retrohoming)
- Signals + chromatin + proteins + ncRNA + lipids --> nuclear/nucleoid localization

- Protein + ncRNA + signals + other molecules + structures <--> Phenotype

What lessons have we learned?

1. no central dogma, no unidirectional information flow
2. atomistic view of genome untenable; interactive nature of all genomic functions (inertness of isolated DNA)
3. combinatorial and “fuzzy logic” precision instead of hardwired “lock and key” specificities
4. genome change as a biochemical process subject to regulation
5. informatic rather than mechanical processes control cell functions
6. critical role of “signals” in execution of cell/organismal phenotypes (cell-to-cell indeterminacy)

What new informatic concepts do we need to elaborate?

- cellular cognition & action on genome
    - Sensing, computation and decision-making are central features of cellular functions
    - The cell is an active agent utilizing information stored in genome
- internal symbolic representations of conditions & operations in signal transduction and checkpoints, developmental programs
- genome system architecture for accessing genome space

Genome System Architecture I:
Genome as a RW memory at multiple time scales

- Within cell cycle by adjustment of DNA binding protein complexes (e.g. replication factors, transcription factors, cell cycle and checkpoint monitors, cohesins)
- Over several cell cycles by chromatin reformatting (DNA methylation, histone modification, chromatin binding proteins)
- Over evolutionary time by natural genetic engineering

Genome System Architecture II:

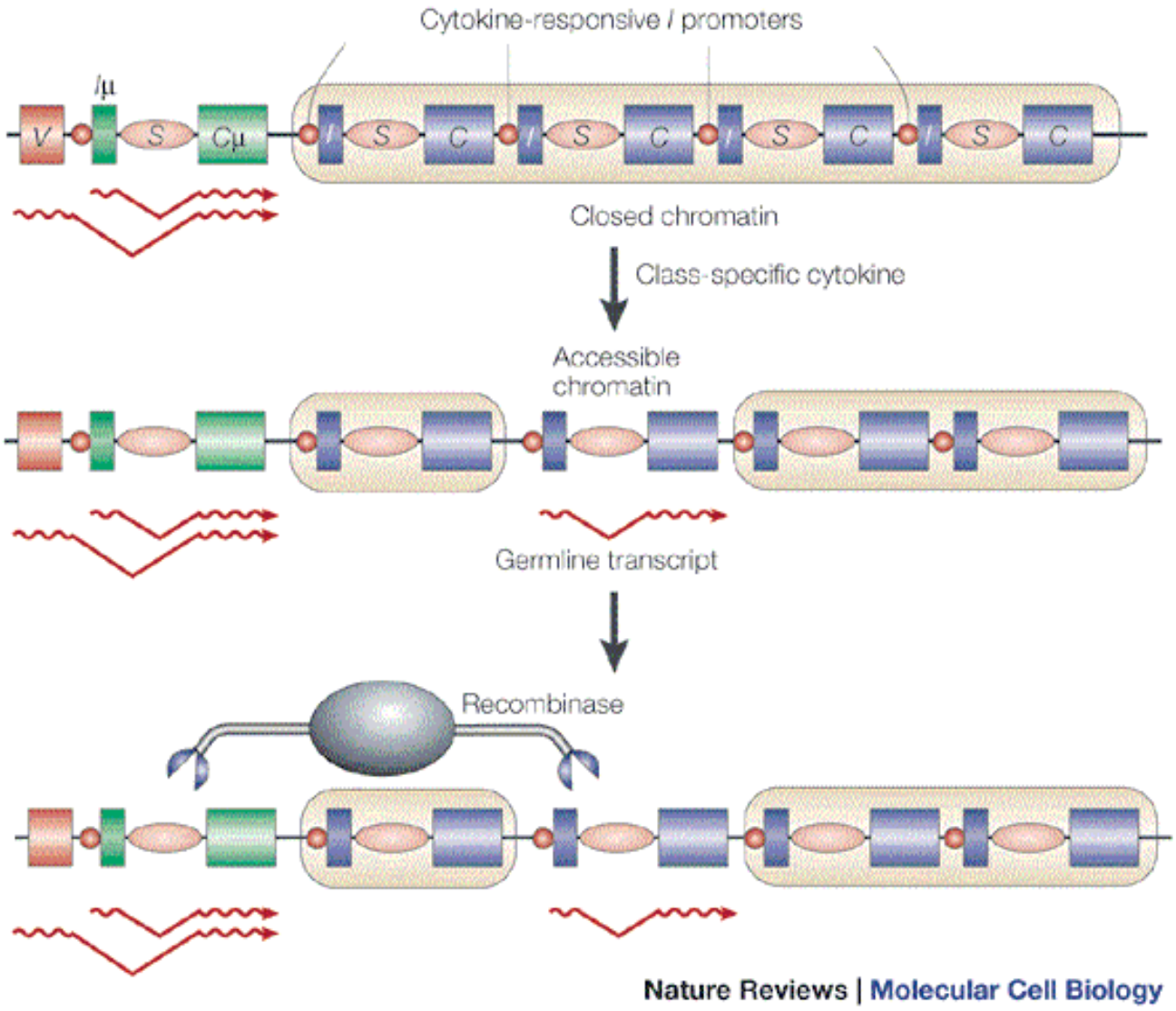
algorithms for searching genome space by control regimes in normal life cycles (e.g. transcription)

- locate locus in nucleus/nucleoid
- adjust chromatin configuration
- adjust transcription factors
- move to proper functional domain [transcription “factory”]
- execute transcription
- process transcription product

Genome System Architecture III:
algorithms for searching genome space by natural genetic engineering functions

- express natural genetic engineering functions
- choose & locate substrate sequences (donor, target)
- move substrates to proper functional domain for rearrangements (e.g. subnuclear DS break repair foci)
- adjust chromatin configuration
- process substrates (e.g. reverse transcription)
- strand joining, replication & sealing

Signals in natural genetic engineering



Kinoshita K, Honjo T. Nature Reviews Molecular Cell Biology 2; 493-503 (2001) LINKING CLASS-SWITCH RECOMBINATION WITH SOMATIC HYPERMUTATION.

Targeting of Natural Genetic Engineering

- protein recognition of DNA sequences and secondary structures (nucleases, recombinases, transposases)
- RNA base-pairing to DNA guide sequences (reverse splicing)
- Coupling to transcription
  - retrotransposon integration (protein-protein interaction)
  - transcription-dependent DS breaks in B cell CSR
- Coupling to chromatin (retrotransposon integration)
- P-element “homing”

Algorithmic control of natural genetic engineering in the normal life cycle

- DNA excision in bacterial differentiation (*B. subtilis* spores, *A. nidulans* heterocysts)
- RNA-guided genome restructuring in ciliated protozoa macronuclear development
- VDJ joining, somatic hypermutation and class switch rearrangements in mammalian lymphocytes