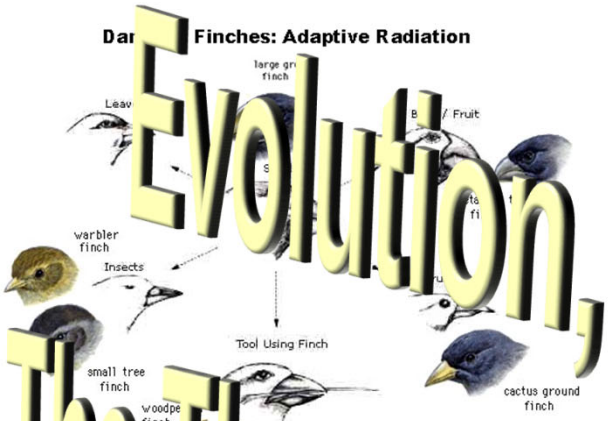
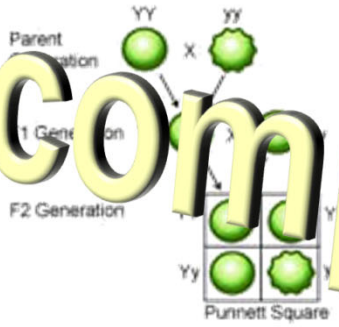
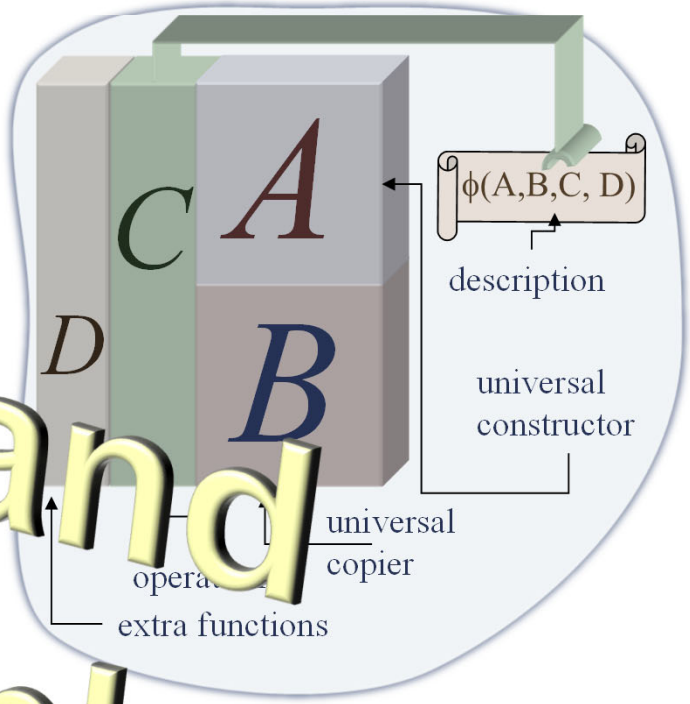
lecture 10: Evolution, Genes, and the Threshold of Complexity

Evolution, Genes, and The Threshold of complexity







BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

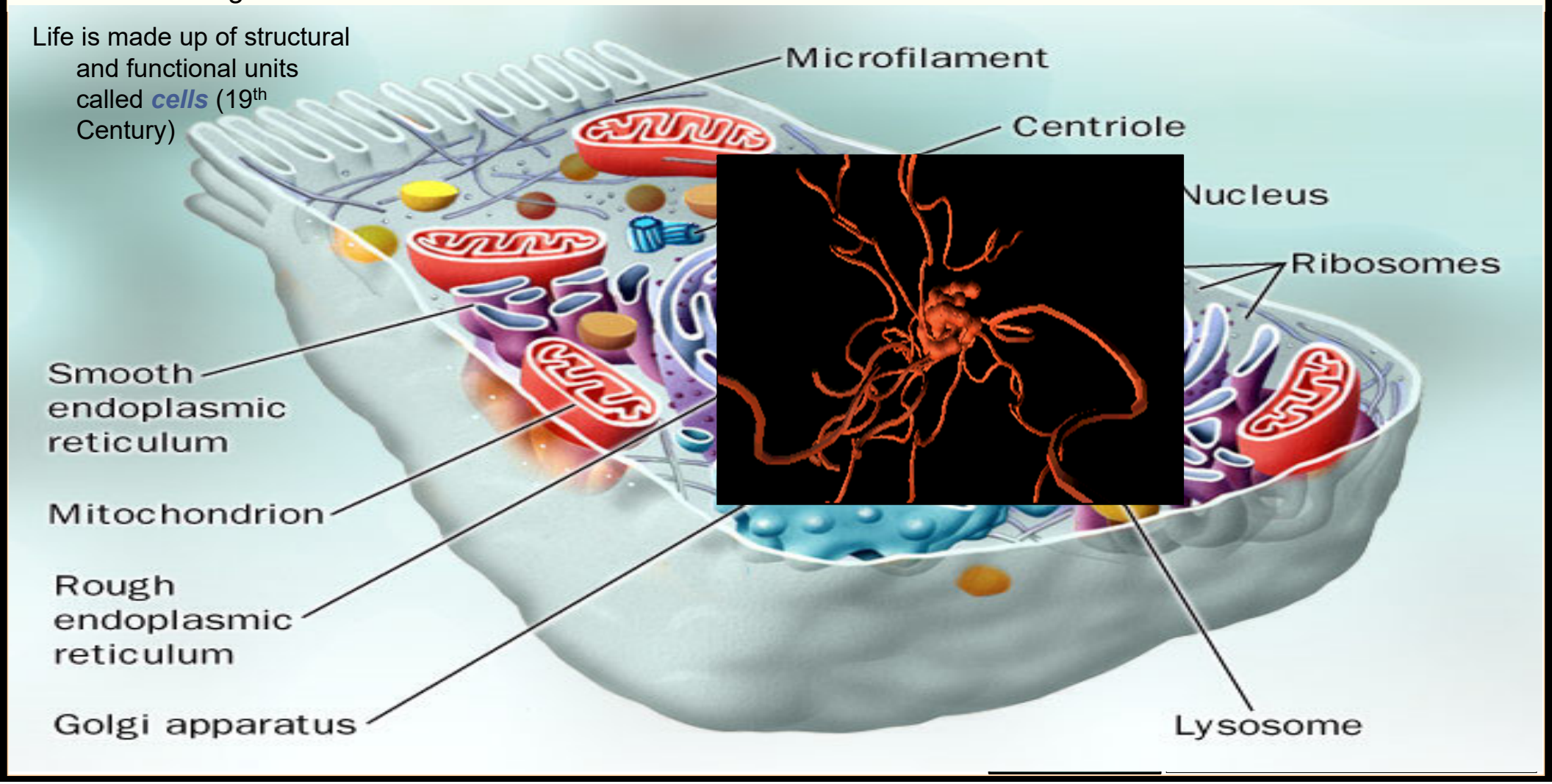
exploring similarities across nature

- self-similar structures
  - Trees, plants, clouds, mountains
    - morphogenesis
  - Mechanism
    - Iteration, recursion, feedback
- dynamical systems and unpredictability
  - From limited knowledge or inherent in nature?
  - Mechanism
    - Chaos, measurement
- self-organization, collective behavior, emergence
  - Complex behavior from collectives of many simple units or agents
    - cellular automata, dynamical networks, morphogenesis, swarms, brains, social systems
  - Mechanism
    - Parallelism, multiplicity, multi-solutions, redundancy
- evolution
  - Adaptation, learning, social evolution
  - Mechanism
    - Reproduction, transmission, variation, selection, Turing's tape
- Network causality (heterogenous complexity)
  - Behavior derived from many inseparable sources
    - Immune system, anticipatory systems, brain-body-environment-culture, embodiment, epigenetics, culture
  - Mechanism
    - Modularity, control, hierarchy, connectivity, stigmergy, redundancy

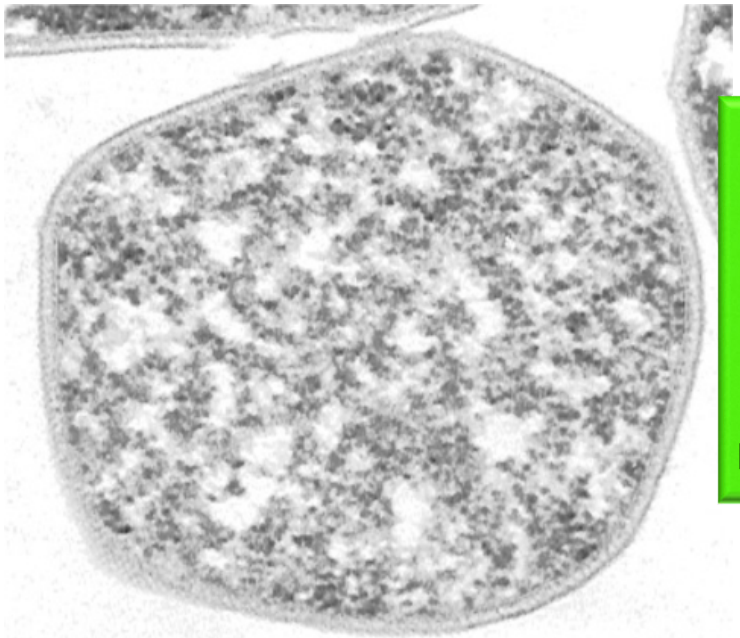
BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

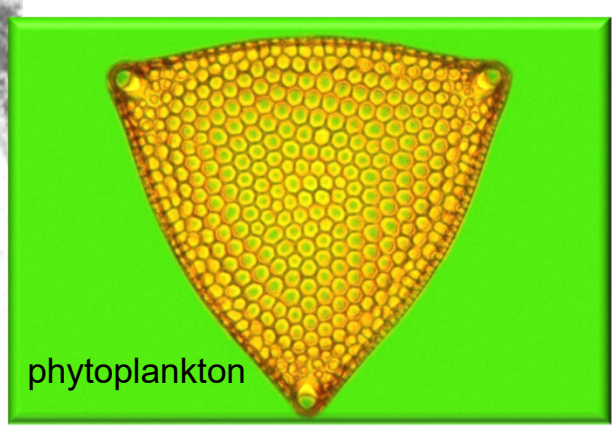
Structure and organelles

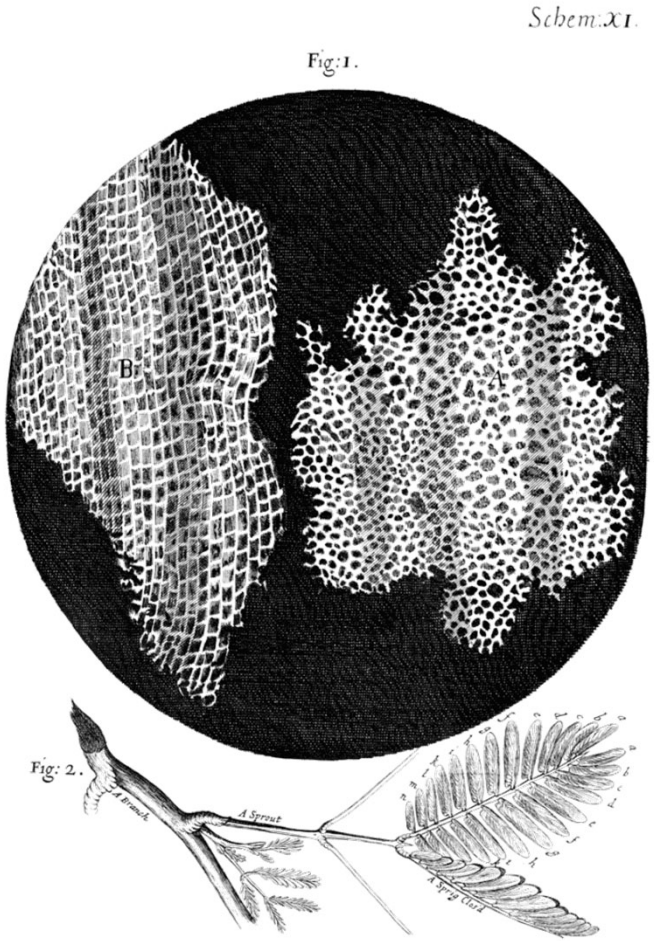
Life is made up of structural and functional units called ***cells*** (19th Century)



- Cell theory
  - Term coined by Robert Hooke (17th century)
  - Matthias Schleiden and Theodor Schwann (19th century)
    - All organisms are composed of one or more cells.
    - All cells come from preexisting cells.
    - All vital functions of an organism occur within cells.
    - Cells contain life's hereditary ***information***

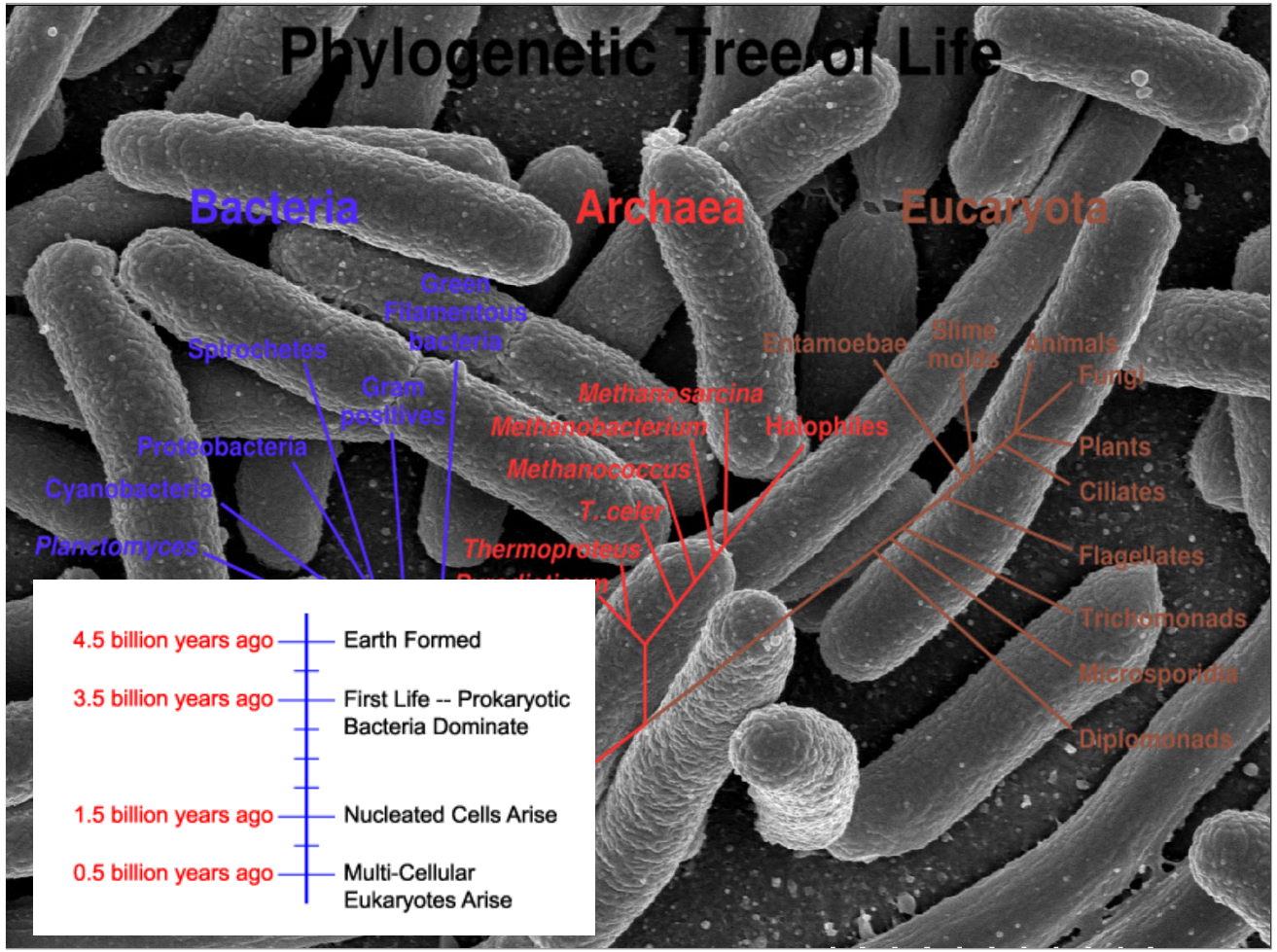






BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

rocha@indiana.edu
casci.binghamton.edu/academics/i-bic



BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

differences and explanations

- Evolution
  - adaptation, learning, innovation, social evolution
- Mechanism
  - Reproduction, transmission, variation, selection
- Design causes
  - Natural selection
- explanation?
  - Contingent, historical, context/specific
  - Does not seem lawful

Natural Selection

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

path to Darwin

- Evolution by natural selection
  - Organisms vary from one another
  - New variation appears from time to time
  - Variation is passed from parent to offspring
  - “struggle for existence” (limited resources)
- Recognized before Darwin
  - Empedocles (490–430 BC)
    - why animals adapt to environment
  - Lucretius (99 - 55 BC) – Epicurus (341-270 BC)
    - Random evolution, free will
  - Al-Jahiz (781 – 869 AD)
    - on the struggle for existence
  - Thomas Hobbes (XVII)
  - Erasmus Darwin (XVIII)
  - Thomas Malthus (XVIII)
    - Populations grow exponentially, re
  - Charles Lyell (XIX)
    - Gradual change in geological lands
  - Jean-Baptiste Lamarck (XIX)
    - Mechanism: mutation and (acquire
  - Alfred Russel Wallace
    - Reached same conclusion as Darv
  - Charles Darwin
    - Evolution, **inevitable**

(Cosma Shalizi citing Aristotle citing) Empedocles:
A difficulty presents itself: why should not nature work, not for the sake of something, nor because it is better so, but just as the sky rains, not in order to make the corn grow, but of necessity? What is drawn up must cool, and what has been cooled must become water and descend, the result of this being that the corn grows. Similarly if a man's crop is spoiled on the threshing-floor, the rain did not fall for the sake of this--in order that the crop might be spoiled--but that result just followed. Why then should it not be the same with the parts in nature, e.g. that teeth should come up of necessity -- the front teeth sharp, fitted for tearing, the molars broad and useful for grinding down the food -- since they did not arise for this end, but it was merely a coincident result; and so with all other parts in which we suppose that there is purpose? **Wherever then all the parts came about just what they would have been if they had come be for an end, such things survived, being organized spontaneously in a fitting way; whereas those which grew otherwise perished and continue to perish**, as Empedocles says his 'man-faced ox-progeny' did.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

path to Darwin

- Evolution by natural selection
  - Organisms vary from one another
  - New variation appears from time to time
  - Variation is passed from parent to offspring
  - “struggle for existence” (limited resources)
- Recognized before Darwin
  - Empedocles (490–430 BC)
    - why animals adapt to environment
  - Lucretius (99 - 55 BC) – Epicurus (341-270 BC)
    - Random *evolution, free will*
  - Al-Jahiz (78
    - on the str
  - Thomas Ho
  - Erasmus Da
  - Thomas Ma
    - Populatio
  - Charles Lye
    - Gradual c
  - Jean-Baptis
    - Mechanis
  - Alfred Russ
    - Reached same conclusion as Darwin (with less evidence)
  - Charles Darwin
    - Evolution, **inevitable**

Lucretius and Epicurism (translated by Stephen Greenblatt):
"... moving randomly through space, like dust motes in a sunbeam, colliding, hooking together, forming complex structures, breaking apart again, in a ceaseless process of creation and destruction. There is no escape from this process. ... There is no master plan, no divine architect, no intelligent design. […] All things, including the species to which you belong, have evolved over vast stretches of time. The **evolution is random**, though in the case of living organisms, it involves **a principle of natural selection**. That is, **species that are suited to survive and to reproduce successfully, endure, at least for a time; those that are not so well suited, die off quickly**. But nothing — from our own species, to the planet on which we live, to the sun that lights our day — lasts forever. Only the atoms are immortal ..."

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

path to Darwin

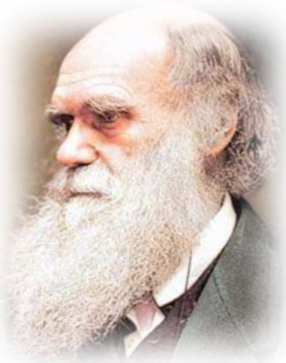
- Evolution by natural selection
  - Organisms vary from one another
  - New variation appears from time to time
  - Variation is passed from parent to offspring
  - "struggle for existence" (limited resources)
- Recognized before Darwin
  - Empedocles (490–430 BC)
    - why animals adapt to environment
  - Lucretius (99 - 55 BC) – Epicurus (341-270 BC)
    - Random evolution, free will
  - Al-Jahiz (781 – 869 AD)
    - on the struggle for existence
  - Thomas Hobbes (XVII)
  - Erasmus Darwin (XVIII)
  - Thomas Malthus (XVIII)
    - Populations grow exponentially, resources lineraly
  - Charles Lyell (XIX)
    - Gradual change in geological landscape
  - Jean-Baptiste Lamarck (XIX)
    - Mechanism: mutation and (acquired) inheritance
  - Alfred Russel Wallace
    - Reached same conclusion as Darwin (with less evidence)
  - Charles Darwin
    - Evolution, **inevitable**

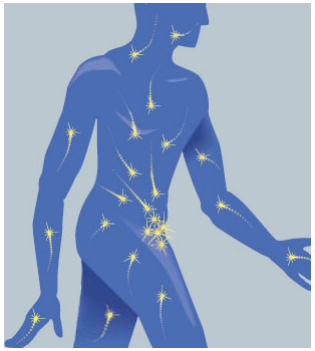
"I happened to read for amusement Malthus on population, and being well prepared to appreciate the struggle for existence…it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The result of this would be the formation of new species." [Charles Darwin]

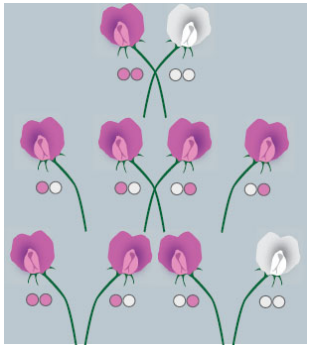
BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

Inheritance mechanism

- XIX Century
  - Evolution of species quickly accepted
  - Natural selection as most important engine of change, was not
    - What was the mechanism?
- Jean-Baptiste Lamarck (XIX)
  - mutation and (acquired) inheritance
- Charles Darwin
  - "gemules" ejected from each tissue and traveling to sex organs
- Gregor Mendel
  - discrete factors corresponding to traits
  - Each individual would carry two copies (one from each parent), but only one would be "expressed"
- "Synthesis" only in the XX century





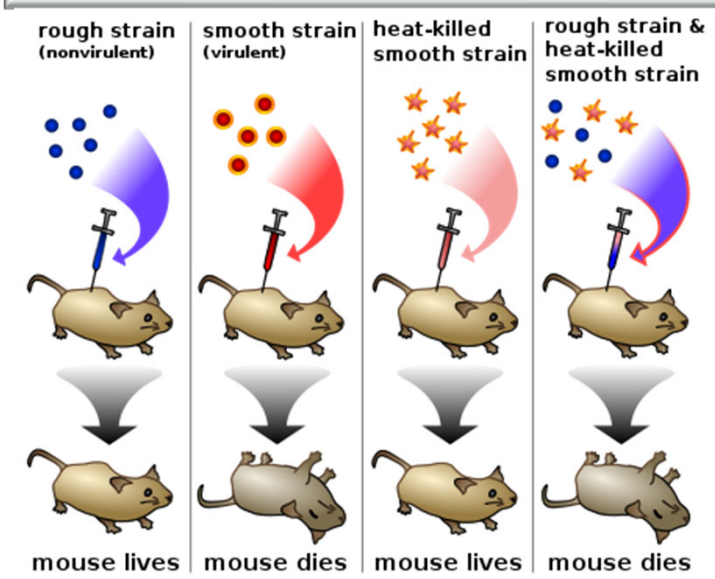


Sci. American, Jan 2009

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

identifying the loci of genetic information

- Frederick Griffith's experiment
  - In 1928: Identified a "transforming principle"
- Avery's experiment
  - Oswald Avery, Colin MacLeod, and Maclyn McCarty
  - 1944: DNA as the loci of "transformation"
    - Chemically knocking off various cellular constituents until trying DNA
    - Considerable resistance in the community accepting this result until the early 1950's (Schrodinger, Delbruck, phage group)



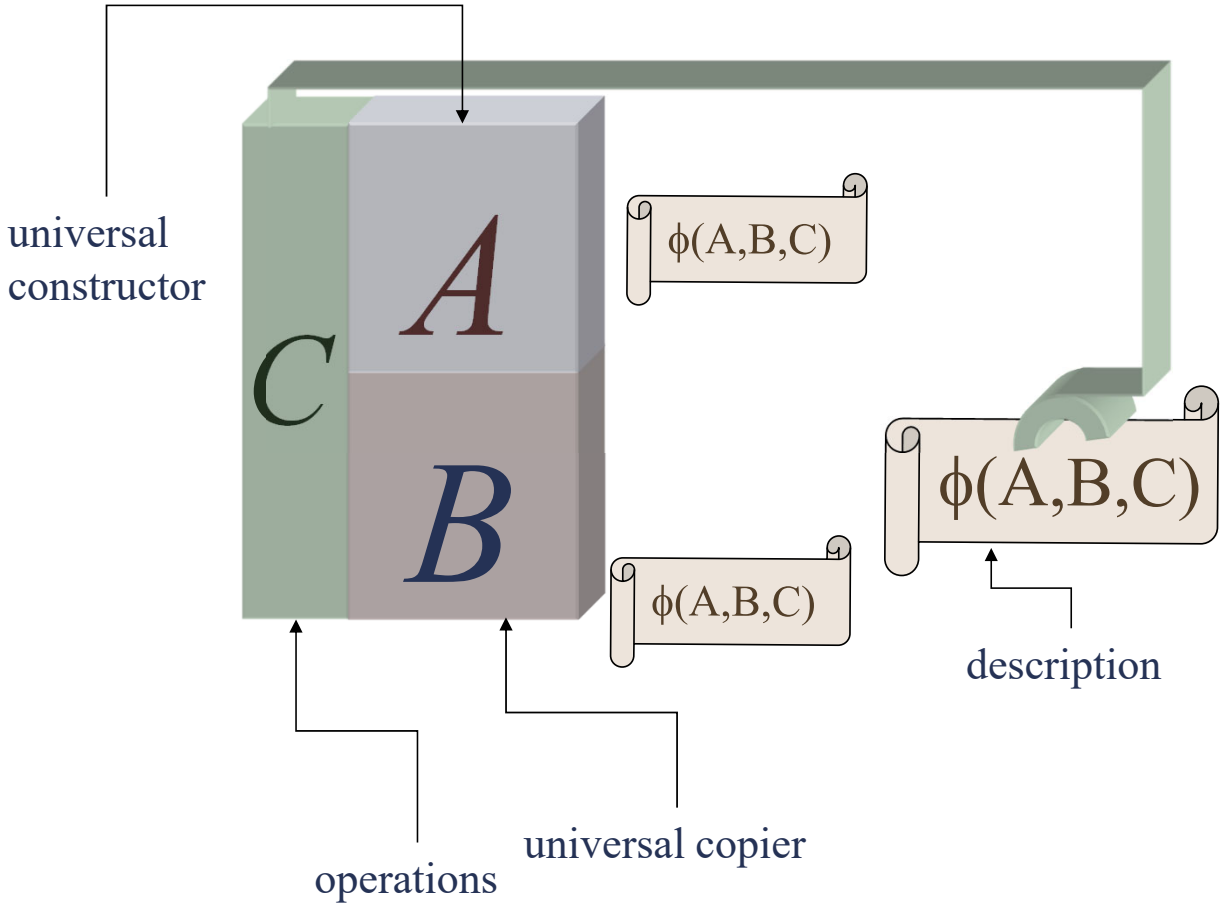
2 different strains of pneumococcus bacteria

initially not well accepted (No auto-catalysis with DNA)

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

Von Neumann's generalization of Turing's tape

as a general principle (system) of **self-replication**

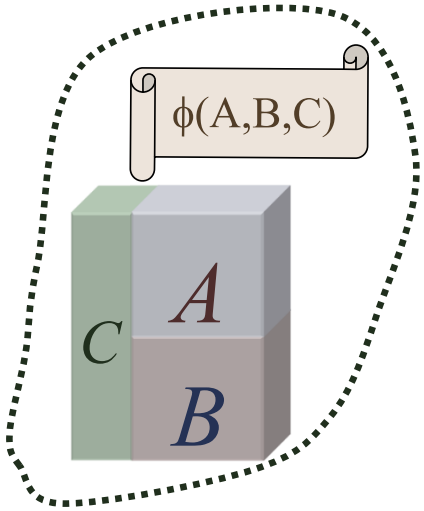


Description is copied **separately**

Construction: interpreted
(horizontal transmission)

Copy: uninterpreted
(vertical Transmission)

code



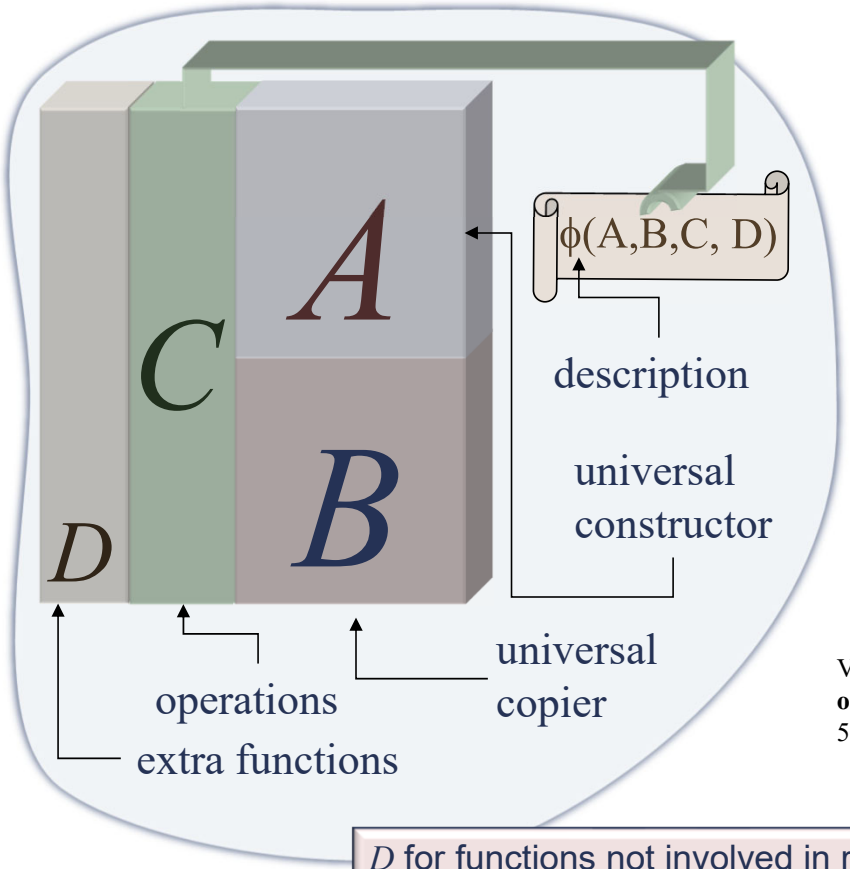
distinction between *numbers that mean things*
and *numbers that do things*.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

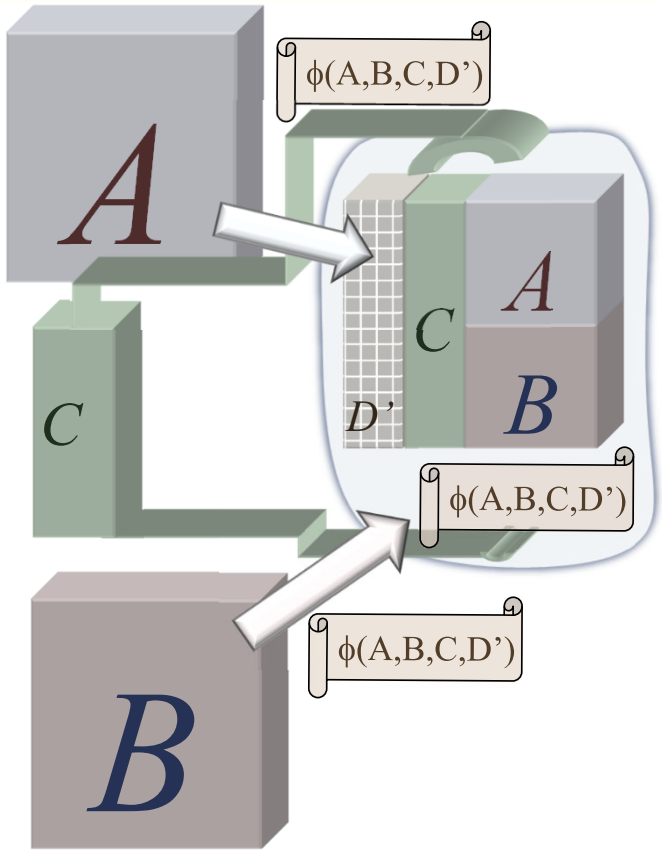
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

Von Neumann's generalization of Turing's tape

as a general principle (system) of evolution or **open-ended complexity**



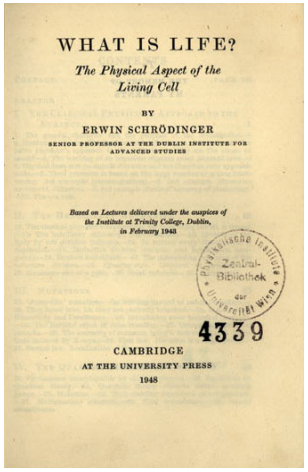
Von Neumann, J. [1949]. "**Theory and organization of complicated automata**." 5 lectures at University of Illinois



D for functions not involved in reproduction
Mutations in *D* can be propagated *vertically*
Leads to **open-ended evolution**

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

Erwin Schrödinger(1943-1944)


WHAT IS LIFE?

The Physical Aspect of the Living Cell

BY

ERWIN SCHRÖDINGER

SENIOR PROFESSOR AT THE DUBLIN INSTITUTE FOR ADVANCED STUDIES

Based on Lectures delivered under the auspices of the Institute at Trinity College, Dublin, in February 1943

CAMBRIDGE
AT THE UNIVERSITY PRESS
1948

- puzzled by the persistence of living structures
  - Call to understand how life stores and perpetuates order
  - "[…] **chromosomes**[…] contain in some kind of **code-script** the entire pattern of the individual's future development."
    - "complete (double) copy of the code-script."
- **aperiodic crystals as structures that can replicate themselves**
  - "We believe a gene—or perhaps the whole chromosome **fiber**—to be an aperiodic solid."
    - "structure without predictable repetition"
  - DNA is entirely regular
    - Instead of "aperiodicity" we have **encoded information**: separated description/construction

"Turing invented the stored-program computer, and von Neumann showed that the description is separate from the universal constructor. This is not trivial. Physicist Erwin Schrödinger confused the program and the constructor in his 1944 book *What is Life?,* in which he saw chromosomes as "*architect's plan and builder's craft in one*". This is wrong. The code script contains only a description of the executive function, not the function itself." (Sydney Brenner)

Brenner, Sydney. [2012]. "Life's code script." *Nature* **482** (7386): 461-461.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

Schrodinger vs. Von Neumann

self-replication vs. decoupled, encoded information

Von Neumann, J. [1949]. "**Theory and organization of complicated automata**." 5 lectures at University of Illinois

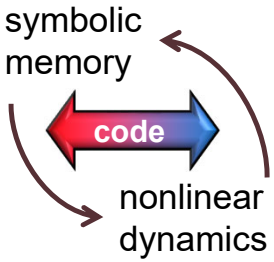
Brenner, Sydney. [2012]. "Life's code script." *Nature* **482** (7386): 461-461.

"Turing invented the stored-program computer, and von Neumann showed that the description is separate from the universal constructor. This is not trivial. Physicist Erwin Schrödinger confused the program and the constructor in his 1944 book *What is Life?,* in which he saw chromosomes as "*architect's plan and builder's craft in one*". This is wrong. The code script contains only a description of the executive function, not the function itself." (Sydney Brenner)

two roles of information
data/program (Turing)
passive/active (Von Neumann)
description/construction-function (Pattee)
genotype/phenotype (Biology)

semiotic closure (semiotic coupling)

fundamental principle of *organized complexity*
Leads to **open-ended evolution**
General principle that includes *Natural Selection*
Von Neumann described this scheme **before** structure of DNA molecule was identified in 1953 by Watson & Crick



Howard Pattee

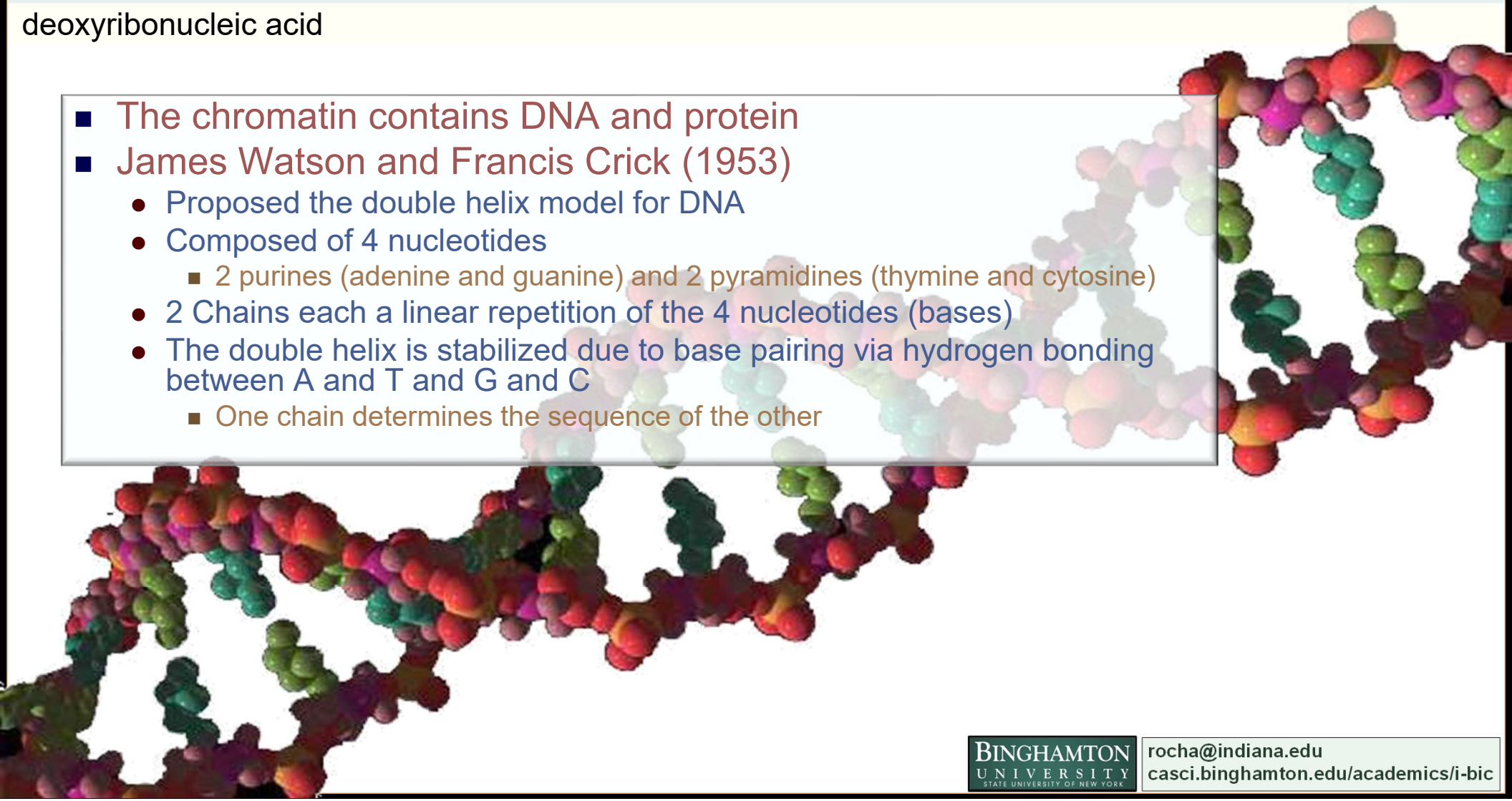
Pattee, HH [2001] *Biosystems* **60** (1):5-21

Rocha, L.M. & W. Hordijk [2005] *Artificial Life* **11**:189 - 214.
Rocha, L.M. [2001] *Biosystems* **60**: 95-121.
Rocha, L.M. [1996] *Systems Research* **13**: 371-384.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

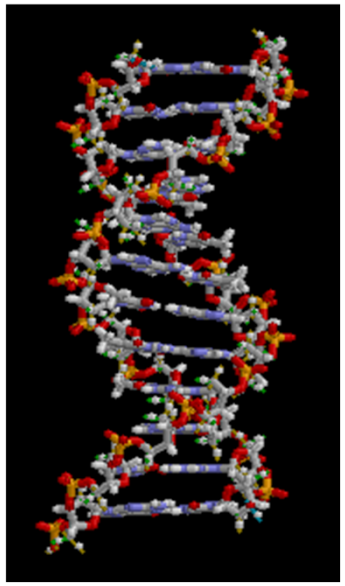
deoxyribonucleic acid

- The chromatin contains DNA and protein
- James Watson and Francis Crick (1953)
  - Proposed the double helix model for DNA
  - Composed of 4 nucleotides
    - 2 purines (adenine and guanine) and 2 pyramidines (thymine and cytosine)
  - 2 Chains each a linear repetition of the 4 nucleotides (bases)
  - The double helix is stabilized due to base pairing via hydrogen bonding between A and T and G and C
    - One chain determines the sequence of the other



BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

a molecular language system: nucleotide “bases” (the genotype “tape”)



Purine (R) → Adenine (A), Guanine (G)

Nucleotides

Pyrimidine (Y) → Cytosine (C), Thymine (T), Uracil (U)

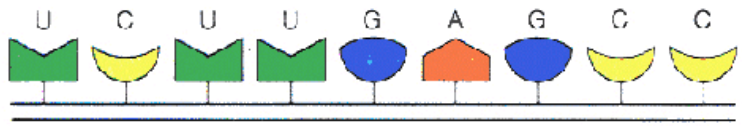
4 Letter Alphabet
DNA: A, G, C, T
RNA: A, G, C, U

Form sequences that can store information

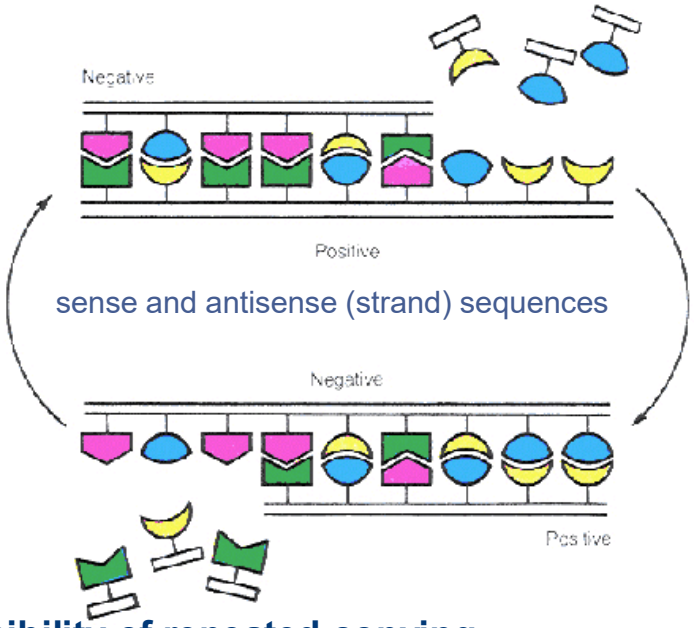
Linear molecules with a phosphate-sugar backbone (deoxyribose and ribose)

Complementary base pairing
(Hydrogen-bonding between purines and pyrimidines)

A-T (U) G-C



Figures from Eigen [1992] . *Steps Towards Life*.

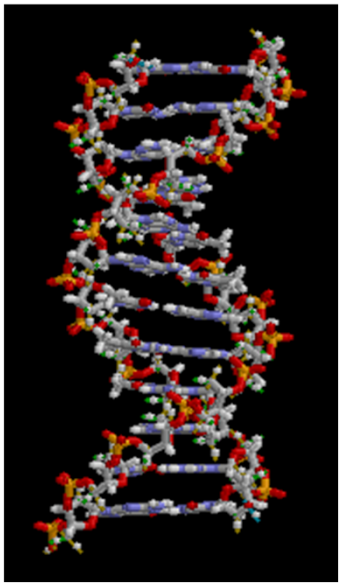


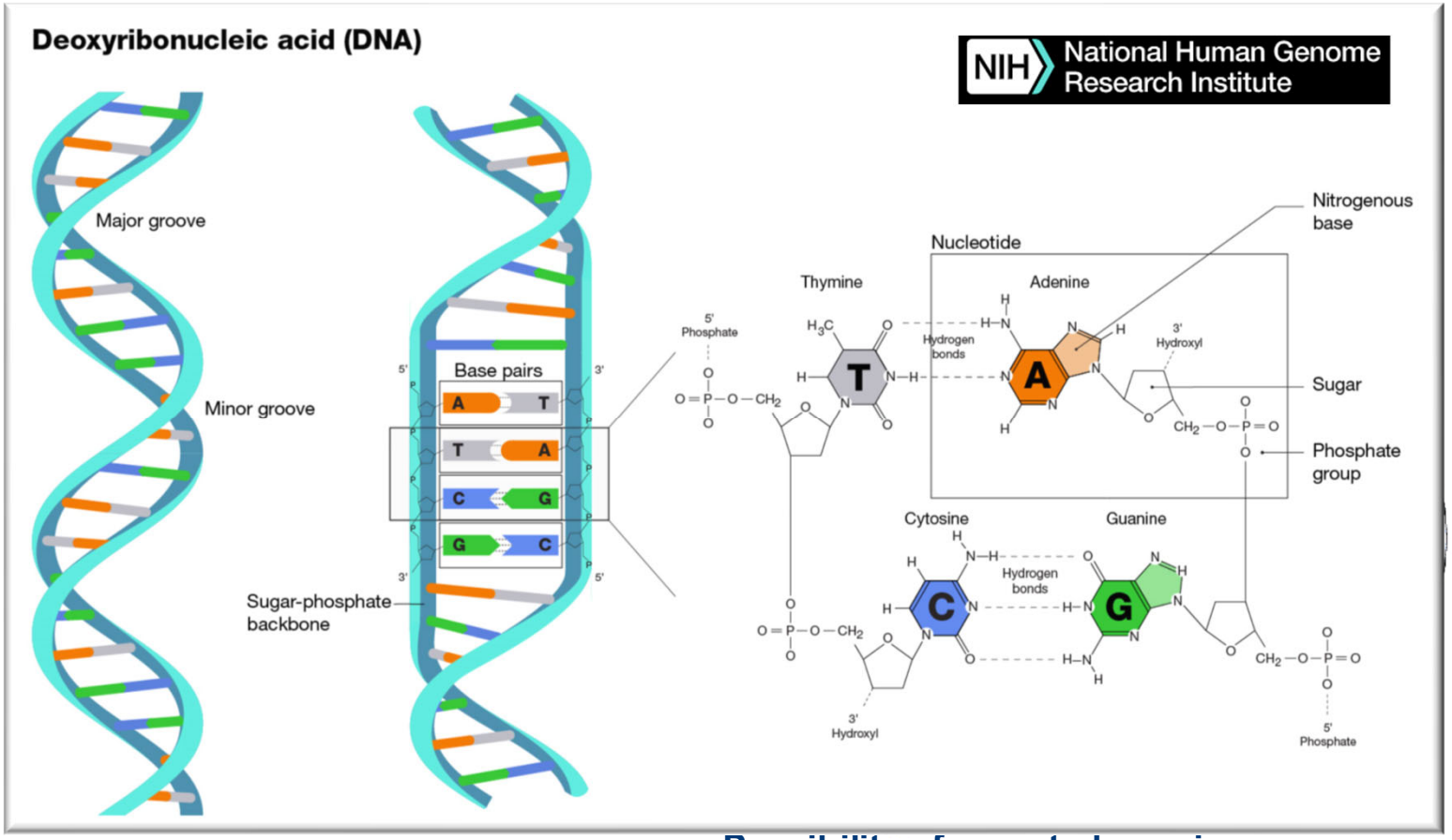
Requirements for structural information

Possibility of repeated copying

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

a molecular language system: nucleotide “bases” (the genotype “tape”)

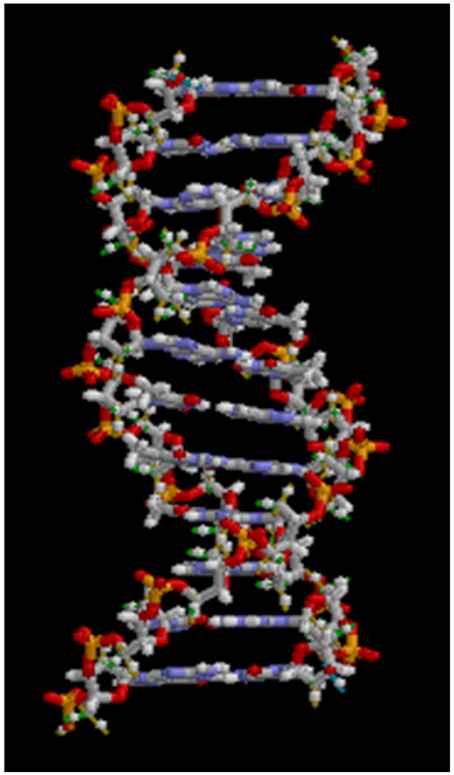


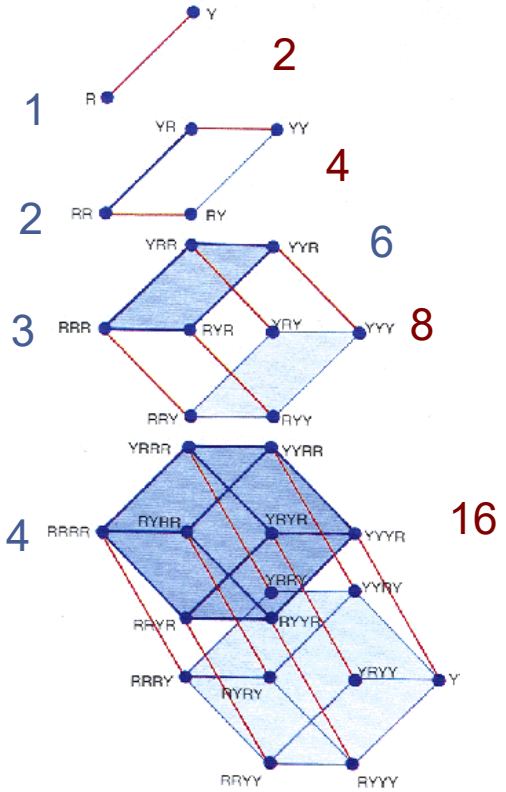


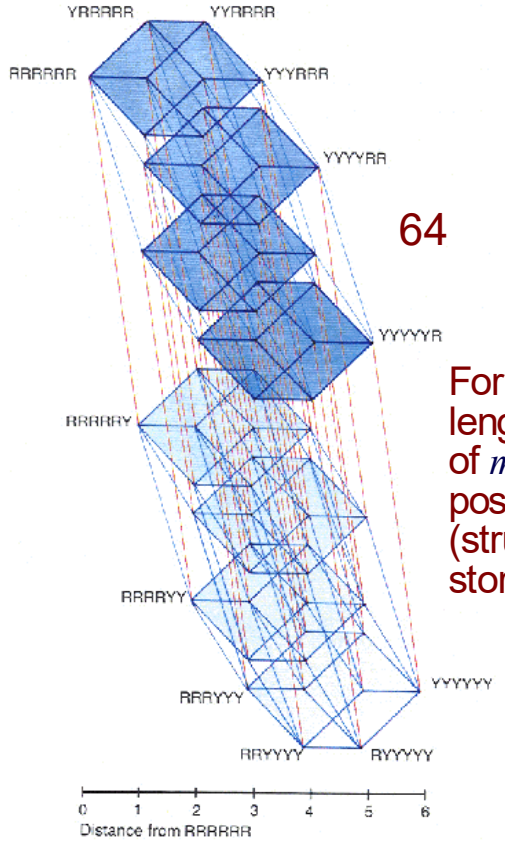
Possibility of repeated copying

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

the genotype “tape” encodes an enormous amount of information





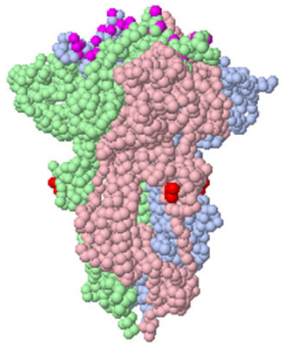


For a sequence of length n, composed of m-ary symbols, m^n possible values (structures) can be stored

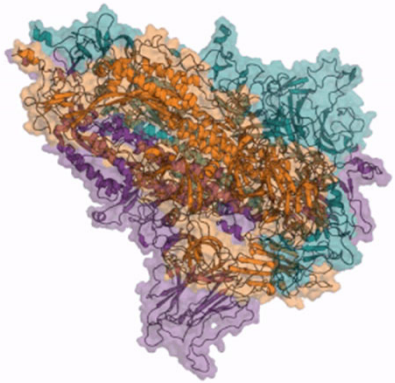
Figures from Eigen [1992] . *Steps Towards Life*.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

functional products that build up (self-organize) the phenotype



Jmol



Polypeptide chains of aminoacids

Primary Structure

Folding

3-dimensional structure

Secondary and tertiary bonds

- In proteins, it is the 3-dimensional structure that dictates function
  - The specificity of enzymes to recognize and react on substrates
- The functioning of the cell is mostly performed by proteins
  - Though there are also ribozymes

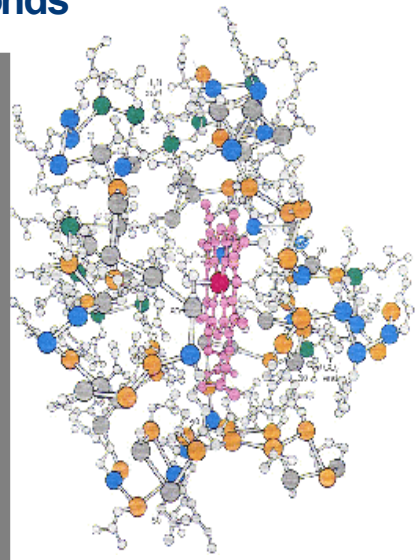


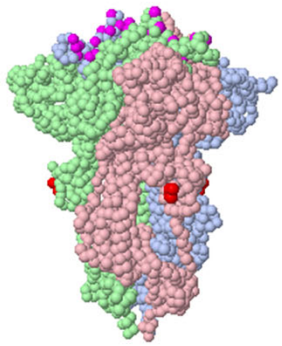
Table 1.4. Amino acid codes

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine
Asx	B	Asn or Asp
Glx	Z	Gln or Glu
Sec	U	Selenocysteine
Unk	X	Unknown

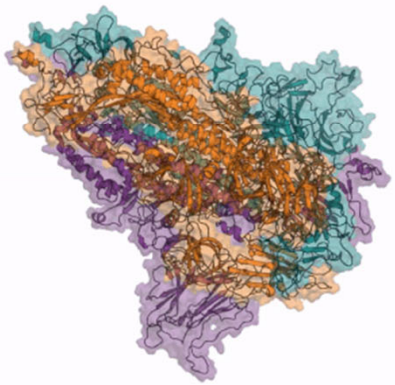
Figures from Eigen [1992] . *Steps Towards Life*.

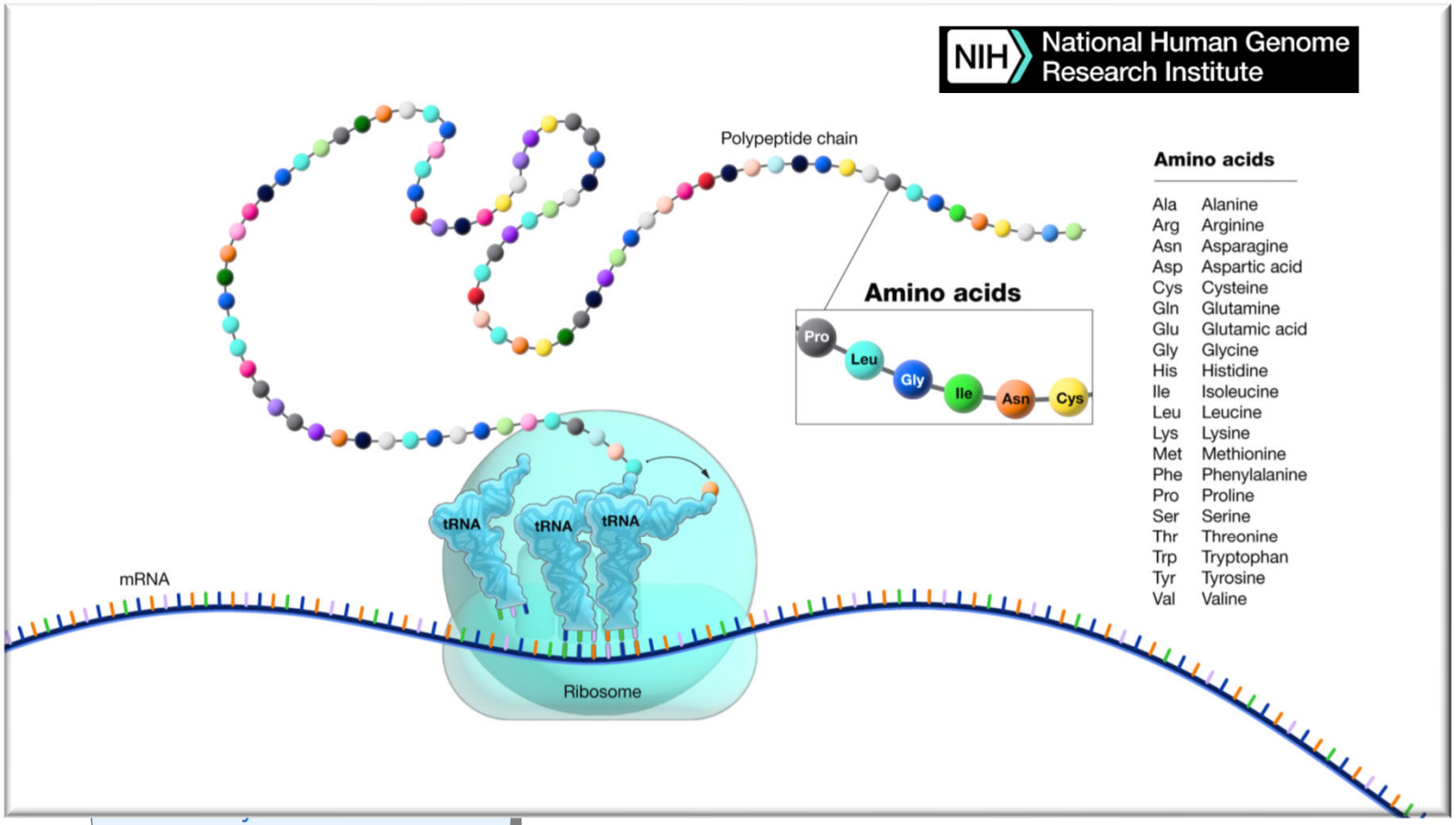
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

functional products that build up (self-organize) the phenotype



Jmol





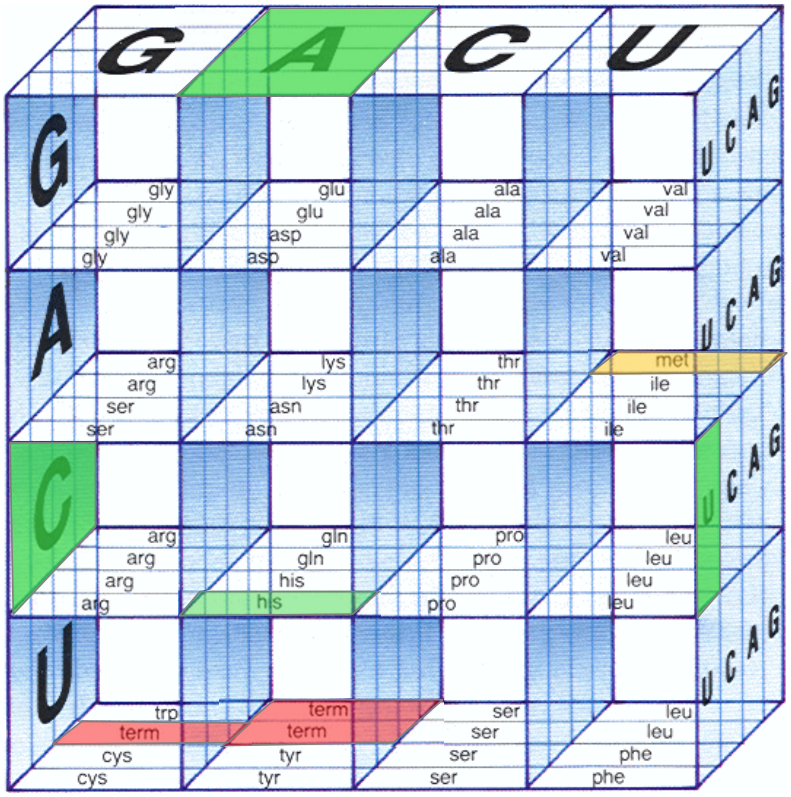
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

between genotype and phenotype

Triplets of 3 Nucleotides can define 64 possible codons, but only 20 amino acids are used (redundancy)

Table 1.4. Amino acid codes

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine
Asx	B	Asn or Asp
Glx	Z	Gln or Glu
Sec	U	Selenocysteine
Unk	X	Unknown



Figures from Eigen [1992] . *Steps Towards Life*.

- The genetic code maps information stored in the genome into functional proteins
  - Triplet combinations of nucleotides into amino acids

Purine (R) → Adenine (A), Guanine (G)

Nucleotides

Pyrimidine (Y) → Cytosine (C), Thymine (T), Uracil (U)

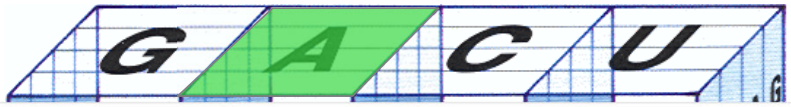
4 Letter Alphabet
DNA: A, G, C, T
RNA: A, G, C, U

Linear molecules w phosphate-sugar

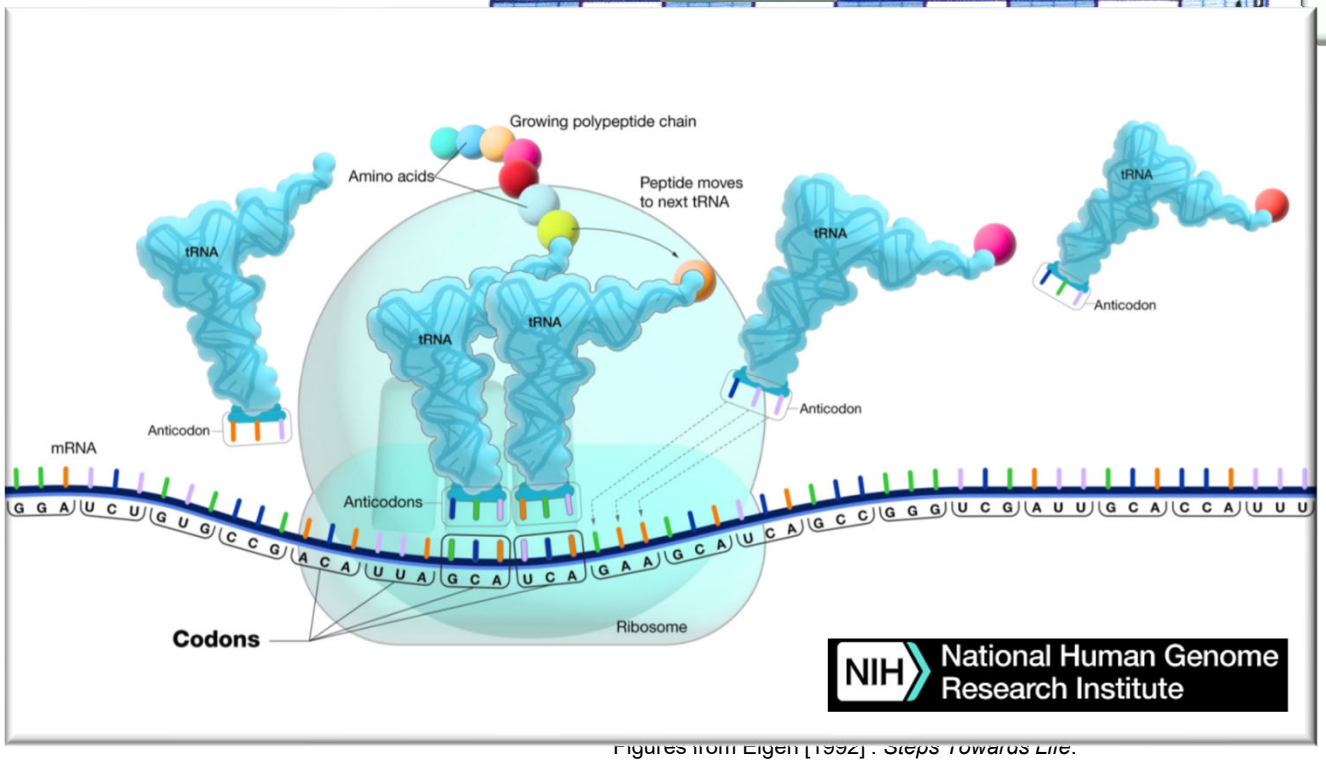
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

between genotype and phenotype

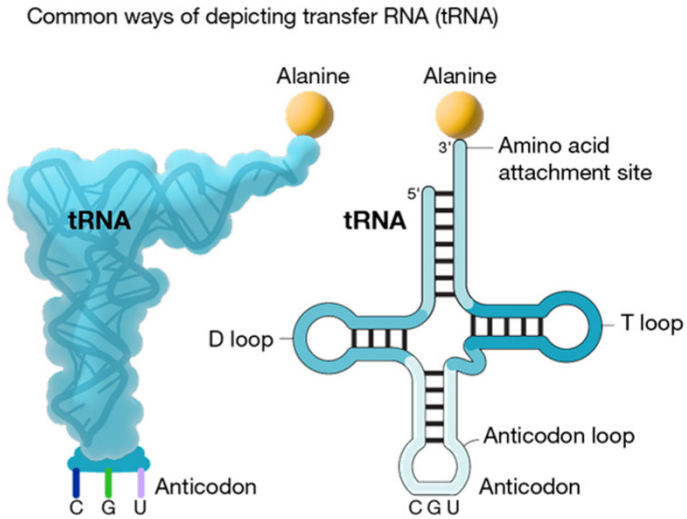
Triplets of 3 Nucleotides can define 64 possible codons, but only 20 amino acids are used (redundancy)



- The genetic code maps information stored in the genome into functional proteins
  - Triplet combinations of nucleotides into amino acids



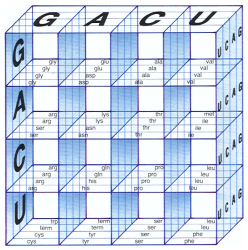
Figures from Eigen [1992]: *Steps Towards Life*.

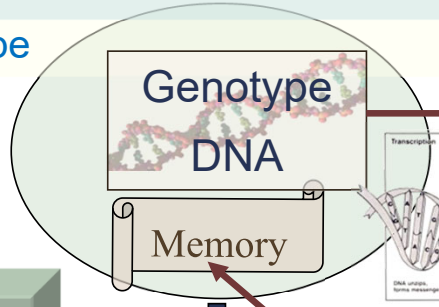


BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

semiotic closure: genetic information at work

genotype/phenotype

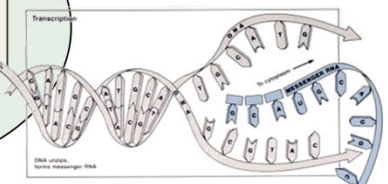




C

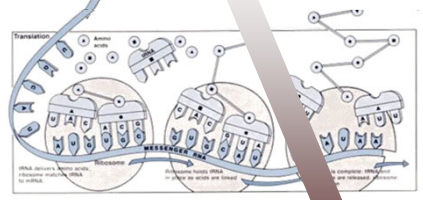
syntax

transcription



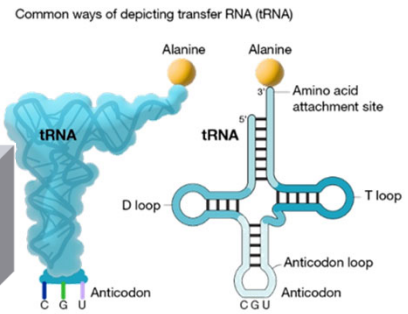
RNA

translation



code

A



semantics

amino acid chains

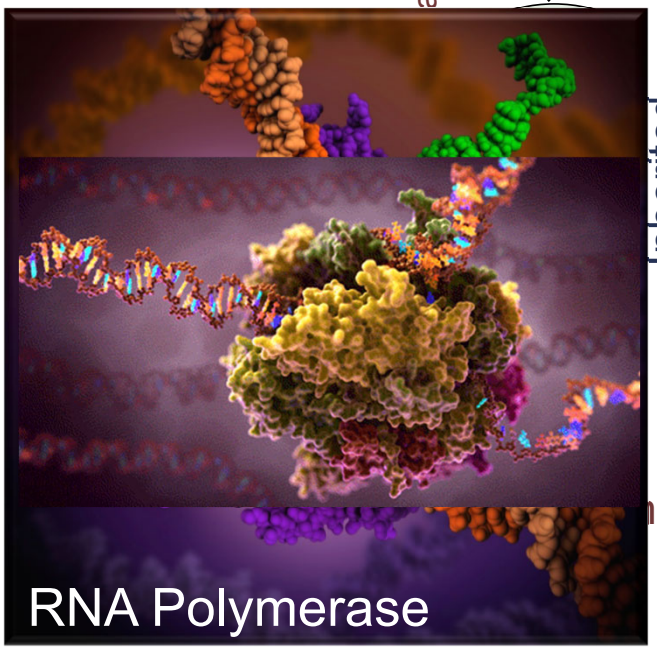
Development, regulation

pragmatics

environmental ramifications

selection

Inherited variation



cell line

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

readings

- Class Book
  - Floreano, D. and C. Mattiussi [2008]. *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*. MIT Press.
    - **Chapters 1 and 4.**
- Lecture notes
  - Chapter 1: What is Life?
  - Chapter 2: The logical Mechanisms of Life
  - Chapter 3: Formalizing and Modeling the World
  - Chapter 4: Self-Organization and Emergent Complex Behavior
  - Chapter 5: Reality is Stranger than Fiction
  - Chapter 6: Von Neumann and Natural Selection
    - posted online @ http://informatics.indiana.edu/rocha/i-bic
- Papers and other materials
  - Optional
    - Nunes de Castro, Leandro [2006]. *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*. Chapman & Hall.
      - Chapter 2, 7, 8
      - **Chapter 3, sections 3.1 to 3.5**
    - Flake's [1998], *The Computational Beauty of Life*. MIT Press.
      - Chapters 10, 11, 14 – Dynamics, Attractors and chaos