information transmission in Biology: the gene

Darwin's Finches: Adaptive Radiation

- Leaves
- Seeds
- Bud/Fruit
- Insect
- Grubs
- Small tree finch
- Woodpecker finch
- Ground finch

Stamen (male)
Carpel (female)

Smooth or Dented Seeds
Green or Yellow Seeds
Axial or Terminal Flowers
Green or Yellow pods
White or Purple Flowers

Parent Generation
F1 Generation Yy
F2 Generation YY Yy
Punnett Square

rocha@indiana.edu
http://informatics.indiana.edu/rocha/i-bic
Assignments: 35%

- Students will complete 4/5 assignments based on algorithms presented in class

Lab meets in I1 (West) 109 on Lab Wednesdays

- Lab 0: January 14th (completed)
  - Introduction to Python (No Assignment)
- Lab 1: January 28th
  - Measuring Information (Assignment 1)
  - Graded
- Lab 2: February 11th
  - L-Systems (Assignment 2)
  - Graded
- Lab 3: March 25th
  - Cellular Automata & Boolean Networks (Assignment 3)
    - Due: April 1st
- Lab 4: April 8th
  - Genetic Algorithms (Assignment 4)
    - Due: April 22nd
Readings until now

- **Class Book**
    - Chapters 1, 2, 3.1-3.4, 7.1-7.5, 8.1-8.2, 8.3.10

- **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”
    - posted online @ [http://informatics.indiana.edu/rocha/i-bic](http://informatics.indiana.edu/rocha/i-bic)

- **Other materials**
    - Chapters 10, 11, 14
Projects

- **Due by May 4th in Oncourse**
  - ALIFE 15 (14)
    - Actual conference due date: 2016
    - [http://blogs.cornell.edu/alife14nyc/](http://blogs.cornell.edu/alife14nyc/)
    - 8 pages (LNCS proceedings format)
    - [http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0](http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0)
  - Preliminary ideas **due by April 1st**!
- **Individual or group**
  - With very definite tasks assigned per member of group
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

Types of Cells
- Prokaryotic (3.5 billion years ago)
  - in single-celled and colonial organisms
    - Bacteria and Archaea, asexual reproduction,
- Eukaryotic cells (aprox. 1.6 - 2.1 billion years ago)
  - Contain organelles with their own membranes
    - Single (amoeba) and multicellular, slime mold, colonial (sponge)

Organisms
- Unicellular, colonial, and multicellular

Chromosome
- Haploid: One copy of each chromosome
  - Fungi, male bees, wasps and ants
- Diploid: Two copies (homologs) of each chromosome
  - One homolog from each parent
colonies, colonial organisms, and clonal colonies from symbiosis to the origin of multicellular organisms?

- Portuguese man-of-war
  - Colony of colonial organism
  - More structurally organized colony of single cell organism

- Scenedesmus
  - Colonial Alga of single cell organisms
  - More structurally organized, specialized and large colony of single cell organisms

- Aspen
  - Clonal colony
  - More structurally organized colony of single cell organism

- Synura

- Volvox

rocha@indiana.edu
http://informatics.indiana.edu/rocha/i-bic
Many systems biology models operate in the ordered regime.

Dynamical systems capable of computation exist well before the edge of chaos.

- A much wider transition? A “band” of chaos.

Most important information transmission and computation in Biology an altogether different process than self-organization.

- Turing/Von Neumann Tape

is self-organization enough?
design principles of computation

Babbage/Lovelace, Turing’s tape, and roles of information

distinction between numbers that mean things and numbers that do things.
biologically inspired computing

biological, social and complexity explanations

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
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

(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 our 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.
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)
  - Random evolution, free will

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 ..."
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)
    - 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 linearly
  - 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]
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*
identifying the loci of genetic information

the discovery of the genetic tape

- 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)
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 pyrimidines (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
nucleic acids as information stores
a molecular language system: nucleotide “bases”

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

Requirements for structural information
Possibility of repeated copying

For a sequence of length \( n \), composed of \( m \)-ary symbols, \( m^n \) possible values (structures) can be stored.
The genetic code maps information stored in the genome into functional proteins

- Triplet combinations of nucleotides into amino acids

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

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

<table>
<thead>
<tr>
<th>AA</th>
<th>Tripeptide Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Gln</td>
<td>Q</td>
</tr>
<tr>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Val</td>
<td>V</td>
</tr>
<tr>
<td>Asx</td>
<td>B</td>
</tr>
<tr>
<td>Glx</td>
<td>Z</td>
</tr>
<tr>
<td>Sec</td>
<td>U</td>
</tr>
<tr>
<td>Unk</td>
<td>X</td>
</tr>
</tbody>
</table>

Functional products

Proteins

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

Figures from Eigen [1992], *Steps Towards Life*.
- Reproduction
  - DNA Polymerase
- Transcription
  - RNA Polymerase
- Translation
  - Ribosome
- Coupling of AA’s to adaptors
  - Aminoacyl Synthetase

generating a message to be expressed
constructing (decoding) the message
Class Book
  - Chapter 2, 7, 8
  - Appendix B.3.2-3 - Turing Machines, Computational complexity
  - Chapter 3, sections 3.1 to 3.5

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”
  - posted online @ http://informatics.indiana.edu/rocha/i-bic

Optional materials
- Flake’s [1998], The Computational Beauty of Life. MIT Press
  - Chapter 20