What is life?

- **Self-sustained heritability**
  - Functionality is *limited* by the genome
  - Life cannot be explained entirely by functionality ("design")
    - Why do snakes have femurs?
    - Why do humans have 5 fingers per hand?
    - Why do mitochondria and chloroplasts have their own rDNA and genetic codes?
  - Because of heredity, **history** is the explanation for **current patterns**
    - "Historical constraint" (phylogenetic or genomic constraint)
  - History of life = evolution
What is evolution?

- **Evolution is a result**: heredity + variation
  - "Descent with modification"
  - **Fact**: evolution has occurred
    - Began as a hypothesis
    - Not falsified, though falsifiable
  - **Theory**: evolution is a body of explanatory principles
    - **Embodies a mechanism**: natural selection, itself a result of:
      - Variation in fitness (ability to survive to reproduce)
      - Heredity (ability to pass alleles to progeny)
      - Reproduction / multiplication (ability for population to grow)
    - **Explains a variety of phenomena**
      - Adaptations
      - Shared nonadaptive features
- **"Origin of species" vs. origin of life**
  - Difficulty of historical reconstruction increases with elapsed time
  - Not possible to reconstruct actual events before heredity originated
Origin of Life

• **Precellular, Cellular**

• **Testing hypotheses**
  – *A priori assumptions* (often not stated) may themselves be tested
  – **Hypotheses** must make falsifiable "predictions"
    • "Null" hypothesis: chance alone is responsible
  – **Predictions** are confirmed (consistent with) or refuted by **data**
  – "**Parsimony**" distinguishes between alternative, unfalsified, working hypotheses
    • "Burden of proof" for less parsimonious (more complex) hypotheses

• **Alternative hypotheses**
  – Intelligent design
    • Least parsimonious (requires a creator)
    • Lack of evidence for "design" does not falsify existence of intelligence
  – Extraterrestrial origin
  – Spontaneous self-organization and natural selection
    • Also embodies many alternative hypotheses
Precellular evolution

• "Prebiotic synthesis"
  – **Hypothesis**: The molecules of life can be formed spontaneously under "prebiotic conditions"
  – **Assumptions**: Prebiotic conditions, "uniformitarianism"
    • (NOT that humans should arise "continuously" from chimpanzees!)
  – **Predictions**
    • Amino acids and purines from prebiotic mixtures (Harold & Urey 1953)
    • Ribose from formaldehyde ("formose" reaction; Butlerow 1861)
    • Polymerization of "activated" nucleoside monomers without proteins

• The "RNA World"
  – **Hypothesis**: RNA preceded proteins and DNA as a primordial, information-bearing, catalytic molecule
  – **Assumptions**: Heredity, uniformitarianism
  – **Predictions**
    • Conserved roles for RNA in fundamental machinery of life
    • At least some RNAs should demonstrate elementary catalytic activity
    • Role for RNA as precursor to DNA may be conserved
Prebiotic conditions?
(a priori assumptions)

- **Molecules** (potentially available from atmosphere, deep ocean rifts, benthic clays)
  - $\text{H}_2$, $\text{CH}_4$, $\text{NH}_3$, $\text{CO}$, $\text{H}_2\text{S}$
  - Mineral catalysts (phosphorus, pyrite, clays)

- **Energy**—potential sources (cal cm$^{-2}$ yr$^{-1}$)
  - Solar radiation $2.6 \times 10^5$
  - UV at wavelengths:
    - 300-400 nm $3.4 \times 10^3$
    - 250-300 nm $5.6 \times 10^2$
    - 200-250 nm $4.1 \times 10^1$
    - $<150$ nm $1.7$
  - Electrical discharges $4.0$
  - Shock waves $1.1$
  - Radiactivity $8 \times 10^1$
  - Volcanoes $1.3 \times 10^{-1}$
  - Cosmic rays $1.5 \times 10^{-3}$

Reasonable energy levels for organic synthesis
Oparin & Haldane (1920) hypothesis

- **Hypothesis:** The origin of life was prebiotic
- **Prediction:** Molecules of life should arise spontaneously in prebiotic conditions
- **Test** (Stanley Miller & Harold Urey 1953)

- **Results:**
  - >10% of C from CH$_4$ was in organic molecules
  - These included amino acids and precursors
    - Amino acids (G, A, D, V, L)
    - HCN and other cyano compounds
    - Aldehydes
Precursors to amino acids

• "Strecker synthesis"
  – Overall reaction:

\[
R-\text{CH} + \text{HC} \equiv \text{N} + \text{H}_2\text{O} \rightarrow \text{H}_2\text{N} \overset{\text{CH}}{\text{C}} \overset{\text{OH}}{\text{O}}
\]

  – In atmosphere:

\[
\begin{align*}
R-\text{CH} & \xrightarrow{+ \text{NH}_3} R-\text{CH} \xrightarrow{- \text{NH}_3} R-\text{CH} \xrightarrow{- \text{H}_2\text{O}} R-\text{CH} \xrightarrow{+ \text{H}_2\text{O}} R-\text{CH} \xrightarrow{+ \text{HCN}} R-\text{CH} \\
& \xrightarrow{- \text{HCN}} 
\end{align*}
\]

  – In ocean:

\[
\begin{align*}
\text{NH}_2 & \xrightarrow{+ \text{H}_2\text{O}} \text{NH}_2 \\
\text{C} \equiv \text{N} & \xrightarrow{- \text{NH}_3} \text{O} = \text{C} \overset{\text{NH}_2}{=} \text{OH}
\end{align*}
\]
Precursors to purines

$4 \text{HCN} \xrightarrow{\text{Sunlight}} \text{HCN}$

$\text{H}_2\text{N}$

$\text{CH}$

$\text{HC}$

$\text{N}$

$\text{N}$

$\text{N}$

$\text{N}$

$\text{N}$

$\text{N}$

$\text{CH}$

$\text{CaPO}_4$

$\text{ATP}$
Precursors to ribose

- **"Formose" synthesis** (Butlerow 1861)
  - Series of condensations beginning with formaldehyde

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{HCOOH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CHOHCHO} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{HCOH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

\[ \text{HCHO} + \text{HCHO} \rightarrow \text{CH}_2\text{OH} \]

Ribose-based nucleoside

Acyclo-nucleoside

Phosphorylation would "activate" monomers
Origin of polymers

- Model for formation of RNA by activated nucleosides

  e.g., methylated nucleoside 5'-phosphorimidazolide:

Without metals  With Zn^{++}

![UV absorption graphs for chain length (poly-G)](image)
Self-assembly of macromolecules

• **Order and complexity result from self-assembly**
  – Proteinoid microspheres with internal structure
  – Multisphere assemblages
  – Membrane-like bilayers with "junctions"

• **Novel microenvironments allow:**
  – Selective permeability via lipid or proteinoid "membranes"
  – Novel (high) concentrations and enhancement of interactions
  – Chained reactions (concentrated products available as substrates)
  – Localized precipitation and organization (compartmentalization)
  – Entropy can *decrease* in subsystems (*not* a violation of the 2nd Law of Thermodynamics)
"RNA World" hypothesis
RNA preceded proteins & DNA (Orgel, Crick, Woese, 1960s)

• RNA has the essential role in peptide assembly
  – mRNA, tRNA, rRNA (which can promote translation even missing some proteins)
  – snRNAs (e.g., U1, U2, U4/6, U5)

• RNA is required for DNA replication and synthesis
  – Primer RNAs required for DNA replication
  – Telomerase RNA required for telomere synthesis
  – Deoxyribonucleotides are *derivatives* of ribonucleotides
  – Reverse transcriptase copies DNA from RNA template

• RNAs are key cofactors
  – 7S RNA (protein secretion), ATP, Coenzyme A

• Some RNAs are catalytic
  – Catalytic unit of RNase P (processes *E. coli* pre-tRNA\textsuperscript{Tyr})
  – Self-splicing of introns

• Some RNAs are regulatory
  – miRNAs
Group 1 self-splicing introns

- Processing occurs as a series of transesterifications
Conservation of Group 1 introns

- Common ancestry allows comparative analysis of function
  - Functionally important sequences/structures are often conserved
  - Group 1 intron structures are conserved:
    - in different genes
    - in different species (slime molds, mitochondria, chloroplasts, some bacteriophage)

Some Gr. 1 introns encode maturases, and some encode endonucleases!
Other introns

- **Group 2 self-splicing a precursor to eukaryotic spliceosome?**
  - Note that Group 2 does not require a cofactor and makes a "lariat"

- **Trans-splicing also occurs** (all mRNAs of trypanosomes, many in *C. elegans*)
  - Important implications for "exon shuffling"
RNA as polymerase

- "L-19" RNA can direct template-dependent extension
  - Depends only on availability of (spontaneous) oligos
RNA as nuclease and ligase

- Nuclease activity is similar to splicing, but site-specific
- Ligase activity is energetically the reverse
RNA as regulator of gene expression

Developmental regulatory gene

stRNA precursor

Transgene, virus, transposon, etc.

dsRNA

Dicer

ALG-1
ALG-2

(stRNA)

(miRNA)

RDE-4

siRNA

RDE-1

Repression of mRNA translation

Degradation of mRNA

Ambros, 2001
"RNA World" hypothesis
RNA preceded proteins & DNA (Orgel, Crick, Woese, 1960s)

- **RNA has the essential role in peptide assembly**
  - mRNA, tRNA, rRNA (which can promote translation even missing some proteins)
  - snRNAs (e.g., U1, U2, U4/6, U5)

- **RNA is required for DNA replication and synthesis**
  - Primer RNAs required for DNA replication
  - Telomerase RNA required for telomere synthesis
  - Deoxyribonucleotides are derivatives of ribonucleotides
  - Reverse transcriptase copies DNA from RNA template

- **RNAs are key cofactors**
  - 7S RNA (protein secretion), ATP, Coenzyme A

- **Some RNAs are catalytic**
  - Catalytic unit of RNase P (processes *E. coli* pre-tRNA\textsuperscript{Tyr})
  - Self-splicing of introns

- **Some RNAs are regulatory**
  - miRNAs
RNA genome

- Early RNA genome could probably self-replicate
  - Template-dependent synthesis, ligation
- Self-splicing would have allowed rapid evolution
  - Different combinations of sequences and thus functions
- Early protein synthesis was directed by RNAs
  - tRNAs, rRNA
- Early gene expression could be regulated by RNAs
  - miRNAs
- Evolution of the Genetic Code
  - Once the codons began to be set up, and complexity of the code increased, it would be difficult to change (historical constraint)
  - Order to codon groupings suggests a possible stepwise adoption of codon assignments...
## Genetic code

<table>
<thead>
<tr>
<th>1st position</th>
<th>2nd position</th>
<th>3rd position</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>Phe</td>
<td>Ser</td>
<td>Tyr</td>
</tr>
<tr>
<td>Phe</td>
<td>Ser</td>
<td>Tyr</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>STOP</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>STOP</td>
</tr>
<tr>
<td>C</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>His</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
</tr>
<tr>
<td>A</td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Asn</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>Met</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>G</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
</tr>
</tbody>
</table>

**Hydrophobic**

**Hydrophilic**
Genetic code

- The code is "degenerate"
  - Third codon position is often completely synonymous
    - Perhaps the original machinery used only 1st & 2nd positions?
- Second position determines hydrophobicity/hydrophilicity
  - If pyrimidine, codon is hydrophobic; if purine, codon is hydrophilic
- Easiest non-enzymatic RNA synthesis encodes protein order (β-sheets)
  - RYR-YRY-RYR... : hydrophobic-hydrophilic-hydrophobic...
- Heritable order is subject to Darwinian selection
  - The stepwise process of selection will always lead to adaptations
    - Novel features arise by (duplication and) modification
Precellular evolution

**"Prebiotic synthesis"**
- **Hypothesis:** The molecules of life can be formed spontaneously under "prebiotic conditions"
- **Assumptions:** Prebiotic conditions, "uniformitarianism"
  - (NOT that humans should arise continuously from chimpanzees!)
- **Predictions**
  - Amino acids and purines from prebiotic mixtures (Harold & Urey 1953)
  - Ribose from formaldehyde ("formose" reaction; Butlerow 1861)
  - Polymerization of "activated" nucleoside monomers without proteins

**The "RNA World"**
- **Hypothesis:** RNA preceded proteins and DNA as a primordial, information-bearing, catalytic molecule
- **Assumptions:** Heredity, uniformitarianism
- **Predictions:**
  - Conserved roles for RNA in fundamental machinery of life
  - At least some RNAs should demonstrate elementary catalytic activity
  - Role for RNA as precursor to DNA may be conserved
Next time

• Cellular evolution
  – Molecular systematics
  – The evolution of plastids
  – Rooting the tree of life with gene duplications
  – The evolution of introns: recent or ancient?
  – Exon shuffling in the evolution of novel functions