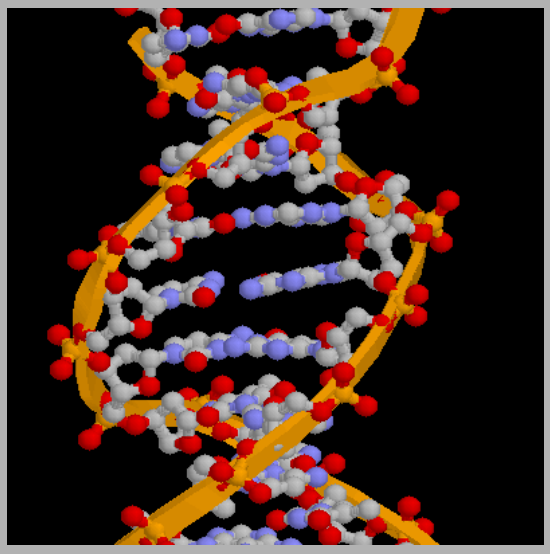
Genes to Life



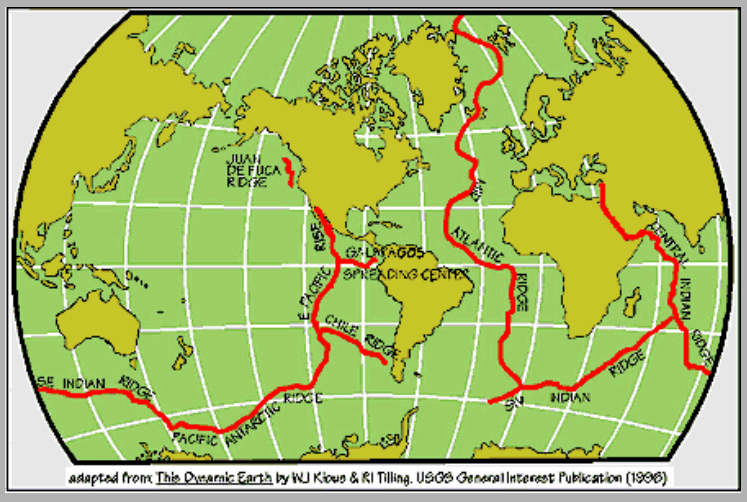
Dr. Adam Marsh, College of Marine Studies, amarsh@udel.edu

Antarctica

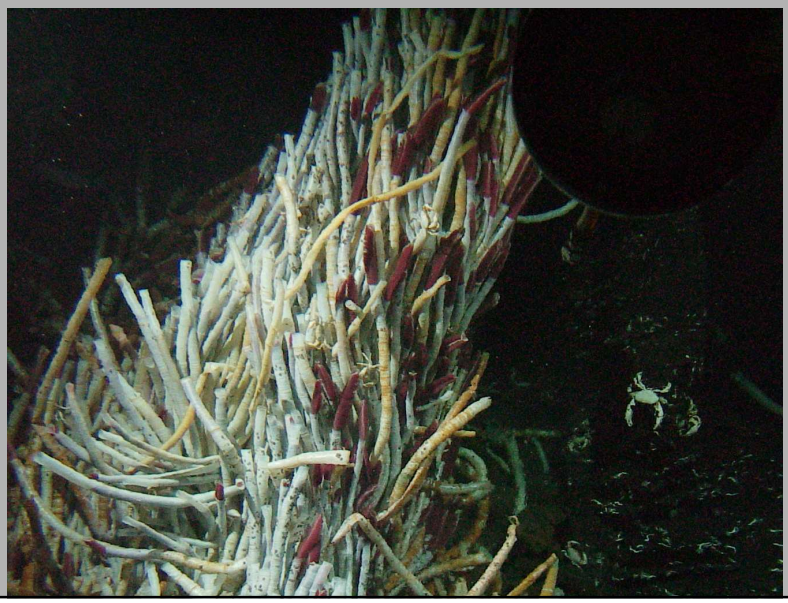
Surface sea ice

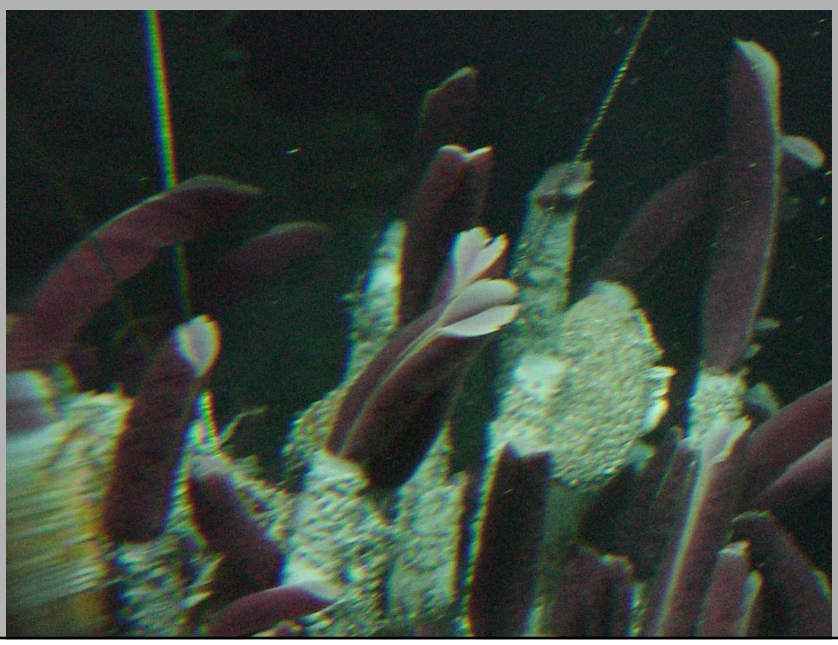
Frozen Ocean

Ocean spreading centers



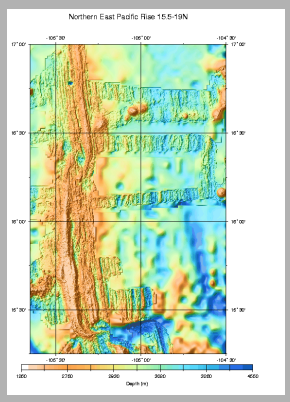
adapted from This Dynamic Earth by WJ Kious & RI Tilling, USGS General Interest Publication (1996)





Genetic Adaptation

How do organisms persist in such diverse environments?



"gene"

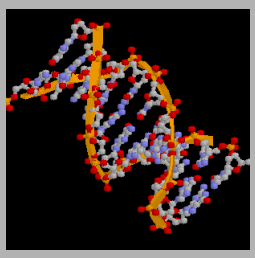
Jacob & Monod (1961):

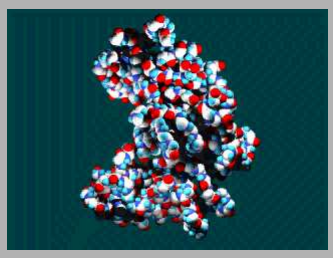
One gene equals one protein

DNA: 1 coding sequence

Protein: 1 biochemical function

Information Expression





1-D information array

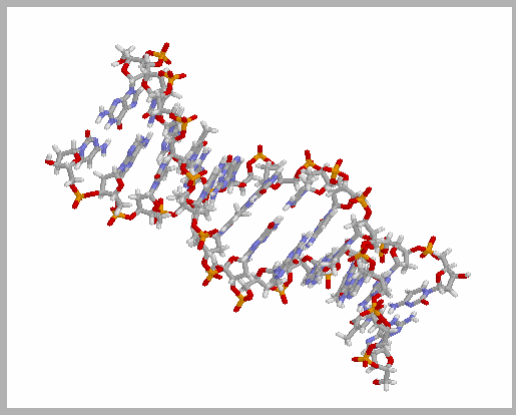
3-D biochemical structure

1 gene = 1 protein

- Genome was a collection of genes
  - coding sequence
  - direct phenotypic function
- Single biochemical function
  - phenotypic trait
  - direct evolutionary selection

DNA → Protein → Phenotype

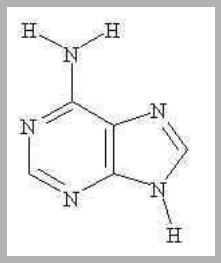
DNA Structure



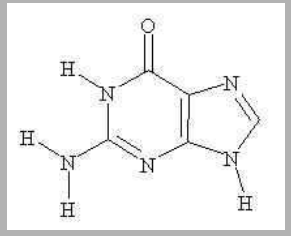
Deoxyribonucleic acids (DNA)

purines

adenosine

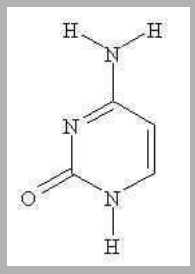


guanine

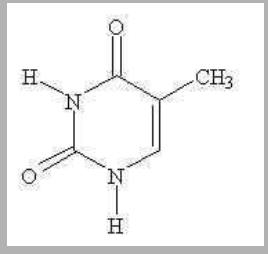


pyrimidines

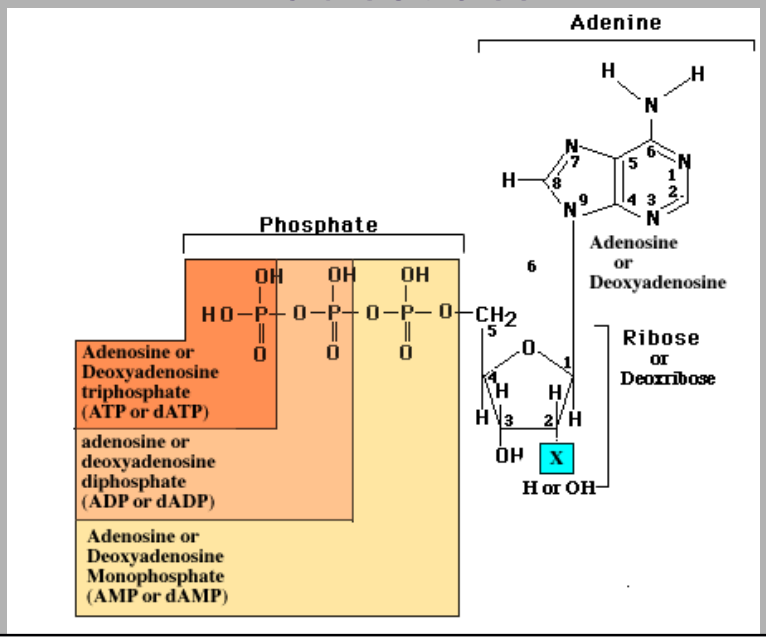
cytosine



thymidine



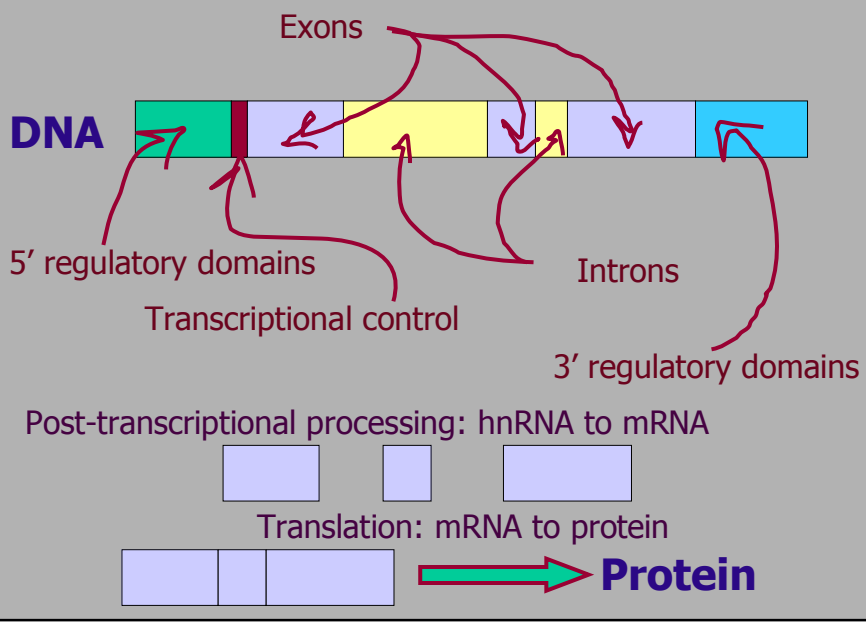
nucleotides



How complex can a 4 letter code really be?

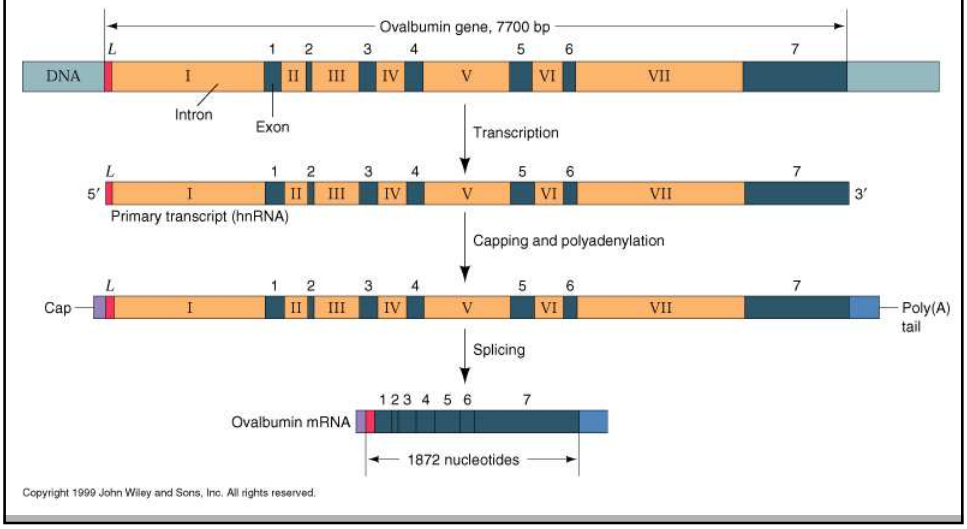
atcgggctatcgatagctatagcgcgatatatcgcgcgtatatgcgcgcatattag
tagctagtgctgattcatctggactgtcgtaatatatacgcgcccggctatcgcgct
atgcgcgatatcgcgcggcgctatataaatattaaaaaataaaatatatatatatgc
tgcgcgatagcgctataggcgcgctatccatatataggcgctcgcccgggcgcga
tgcatcggctacggctagctgtagctagtcggcgattagcggcttatatgcggcga
gcgatgagagtcgcggctataggcttaggctatagcgctagtatatagcggctagc
cgcgtagacgcgatagcgtagctagcggcgcgcgtatatagcgcttaagagcca
aaatgcgtctagcgctataatatgcgctatagctatatgcggctattatatagcgca
gcgctagctagcgtatcaggcgaggagatcgatgctactgatcgatgctagagca
gcgtcatgctagtagtgccatatatatgctgagcgcgcgtagctcgatattacgcta
cctagatgctagcgagctatgatcgtagca.....

Gene Structure

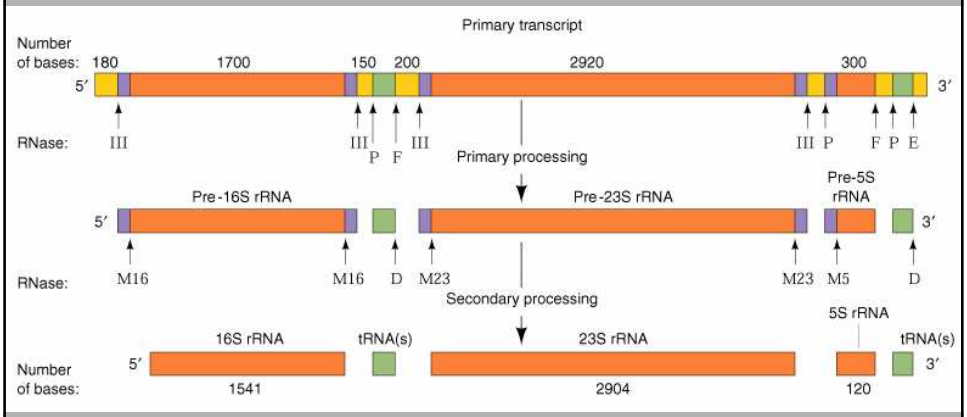


Chicken Ovalbumin Gene

Postranscriptional Processing:



E. coli rRNA gene



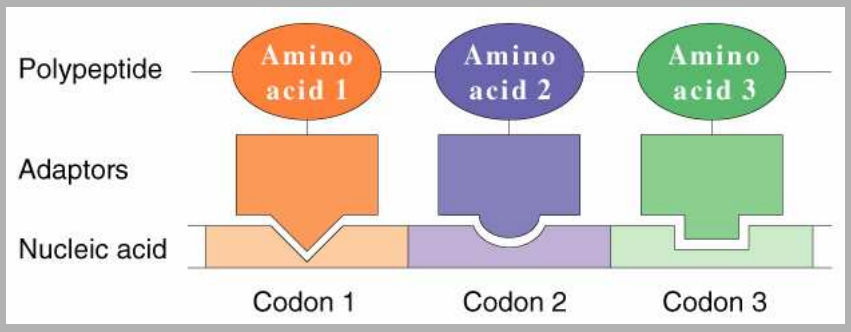
Genetic Code: codons

First Position	Second Position of Codon: T	Second Position of Codon: C	Second Position of Codon: A	Second Position of Codon: G	Third Position
T	TTT Phe [F]	TCT Ser [S]	TAT Tyr [Y]	TGT Cys [C]	T
	TTC Phe [F]	TCC Ser [S]	TAC Tyr [Y]	TGC Cys [C]	C
	TTA Leu [L]	TCA Ser [S]	TAA *Ter* [end]	TGA *Ter* [end]	A
	TTG Leu [L]	TCG Ser [S]	TAG *Ter* [end]	TGG Trp [W]	G
C	CTT Leu [L]	CCT Pro [P]	CAT His [H]	CGT Arg [R]	T
	CTC Leu [L]	CCC Pro [P]	CAC His [H]	CGC Arg [R]	C
	CTA Leu [L]	CCA Pro [P]	CAA Gln [Q]	CGA Arg [R]	A
	CTG Leu [L]	CCG Pro [P]	CAG Gln [Q]	CGG Arg [R]	G
A	ATT Ile [I]	ACT Thr [T]	AAT Asn [N]	AGT Ser [S]	T
	ATC Ile [I]	ACC Thr [T]	AAC Asn [N]	AGC Ser [S]	C
	ATA Ile [I]	ACA Thr [T]	AAA Lys [K]	AGA Arg [R]	A
	ATG Met [M]	ACG Thr [T]	AAG Lys [K]	AGG Arg [R]	G
G	GTT Val [V]	GCT Ala [A]	GAT Asp [D]	GGT Gly [G]	T
	GTC Val [V]	GCC Ala [A]	GAC Asp [D]	GGC Gly [G]	C
	GTA Val [V]	GCA Ala [A]	GAA Glu [E]	GGA Gly [G]	A
	GTG Val [V]	GCG Ala [A]	GAG Glu [E]	GGG Gly [G]	G

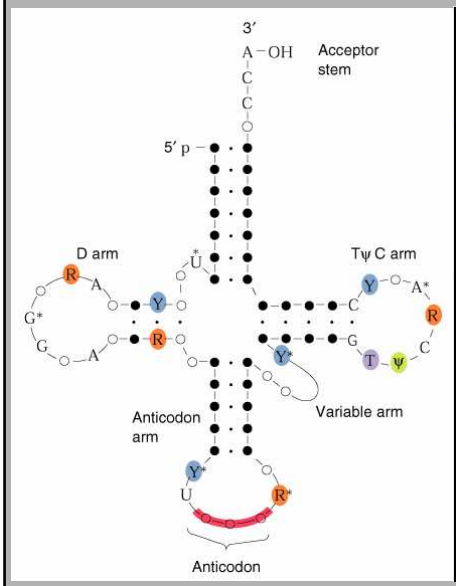
Amino Acids and Proteins

- There are 20 "standard" amino acids that make up all the proteins in all organisms
- There are more potential combinations of these 20 amino acids than visible stars in the sky (100 AA peptide)

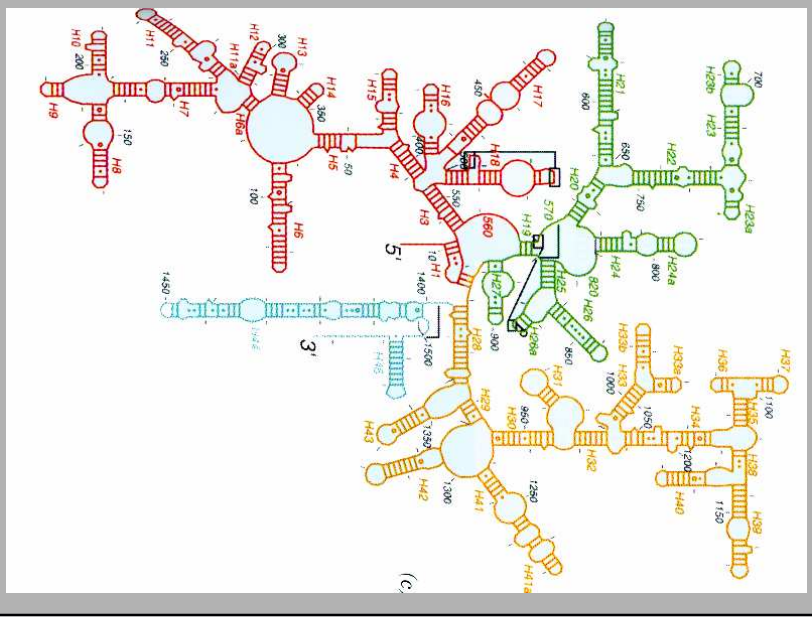
mRNA Translation



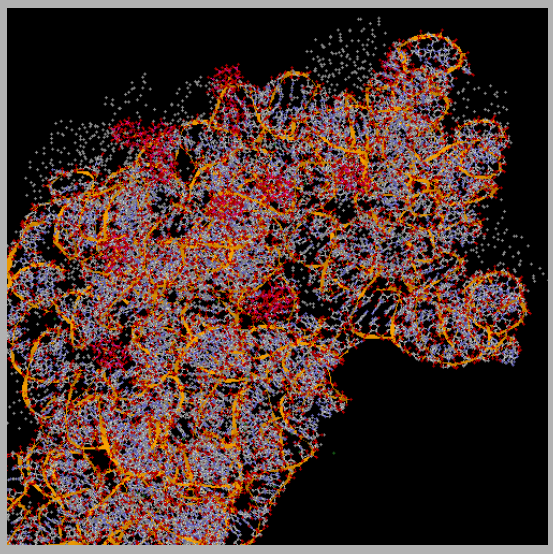
tRNA: codon translators



Ribosomes: structural RNA



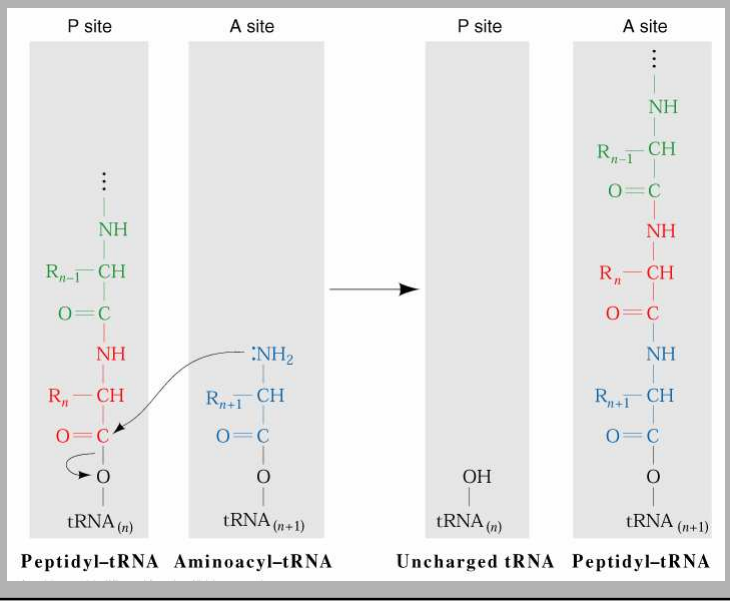
Ribosomes: structural RNA



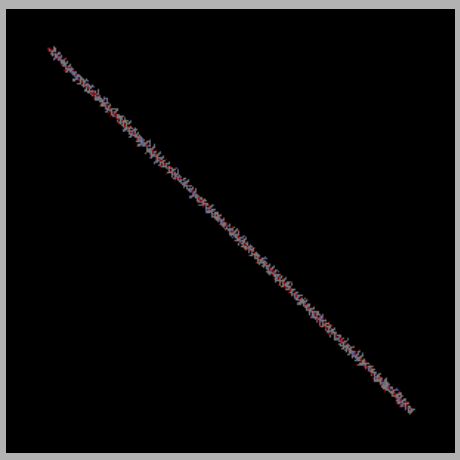
PS: Initiation

	Initiation codon	
araB	– U U U G G A U G G A G U G A A A C G A U G G C G A U U –	
galE	– A G C C U A A U G G A G C G A A U U A U G A G A G U U –	
lacI	– C A A U U C A G G G U G G U G A U U G U G A A A C C A –	
lacZ	– U U C A C A C A G G A A A C A G C U A U G A C C A U G –	
Q β phage replicase	– U A A C U A A G G A U G A A A U G C A U G U C U A A G –	
ϕX174 phage A protein	– A A U C U U G G A G G C U U U U U U A U G G U U C G U –	
R17 phage coat protein	– U C A A C C G G G G U U U G A A G C A U G G C U U C U –	
Ribosomal S12	– A A A A C C A G G A G C U A U U U A A U G G C A A C A –	
Ribosomal L10	– C U A C C A G G A G C A A A G C U A A U G G C U U U A –	
trpE	– C A A A A U U A G A G A A U A A C A A U G C A A A C A –	
trp leader	– G U A A A A A G G G G U A U C G A C A A U G A A A G C A –	
3′ end of 16S rRNA	3′ HO A U U C C U C C A C U A G – 5′	

PS: amino acid oligomerization

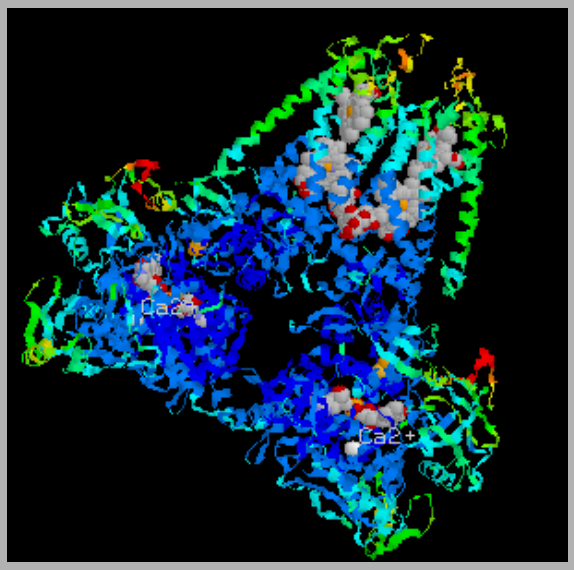


Polypeptide Chain



Linear protein molecule?

Quaternary Protein Structure

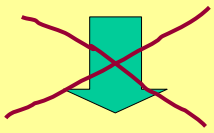


"gene"

Jacob & Monod (1961):

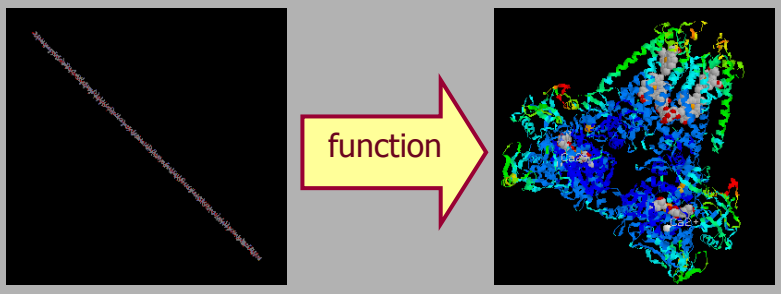
One gene equals one protein

DNA: 1 coding sequence



Protein: 1 biochemical function

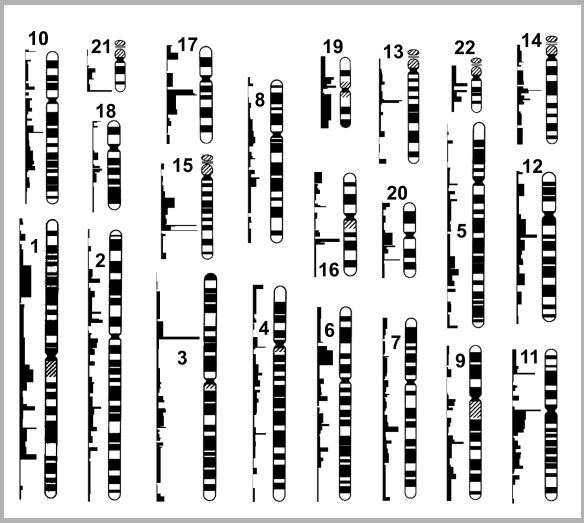
STRUCTURE IS EVERYTHING



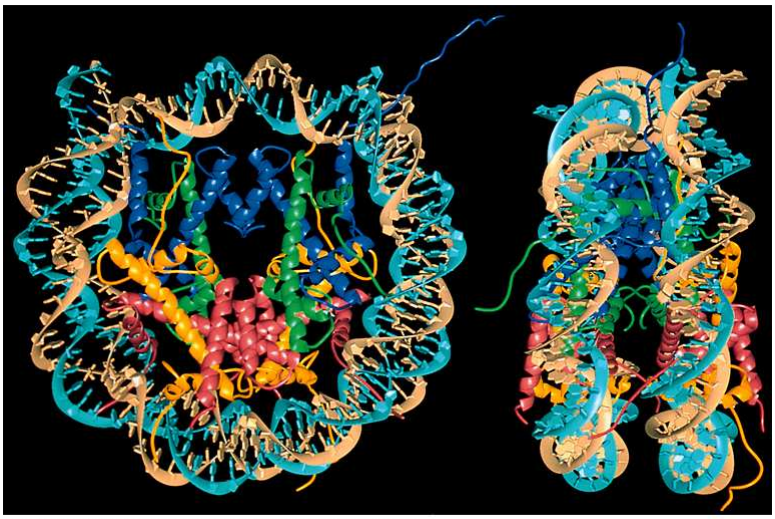
1 dimensional data sets obscure biological function
DNA has structure that is just as critical to its function

Gene Distributions

- Unequal density across chromosomes



DNA Structure: Histones

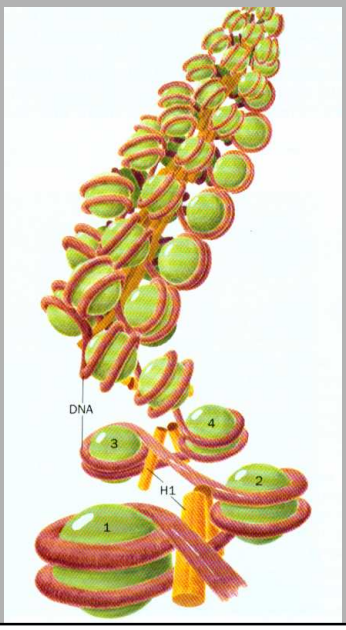


Courtesy of Timothy Richmond, Eidgenössische Technische Hochschule, Switzerland.

Histones and chromatin

- At the simplest level, the suite of DNA in eukaryotic cells would not fit within the nucleus without some further packaging
- DNA is wrapped around histones to form nucleosomes, and then further condensed into chromatin

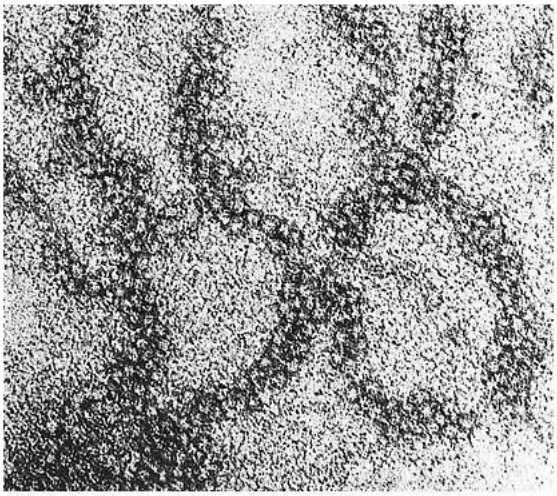
Chromatin



Histone/DNA clusters form aggregation structures called CHROMATIN. Gene expression events are highly impacted by chromatin formation.

All genes are not equally available to be expressed

Histone:DNA 2nd structure



Courtesy of Jerome B. Rattner, University of Calgary, Canada.

"gene"

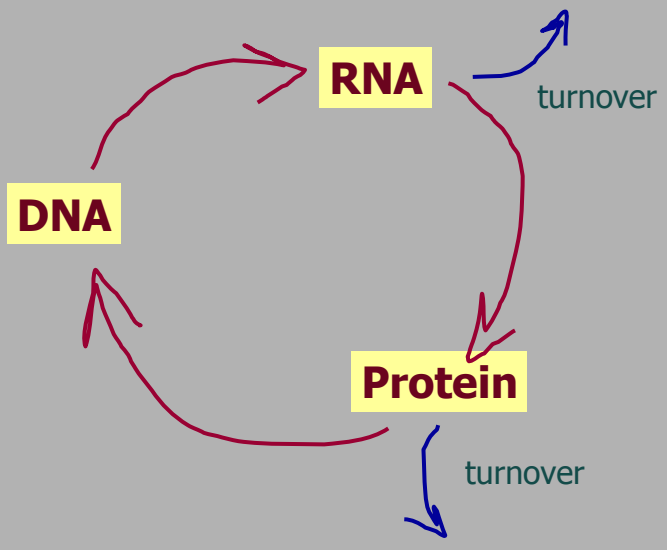
Jacob & Monod (1961):

One gene equals one protein

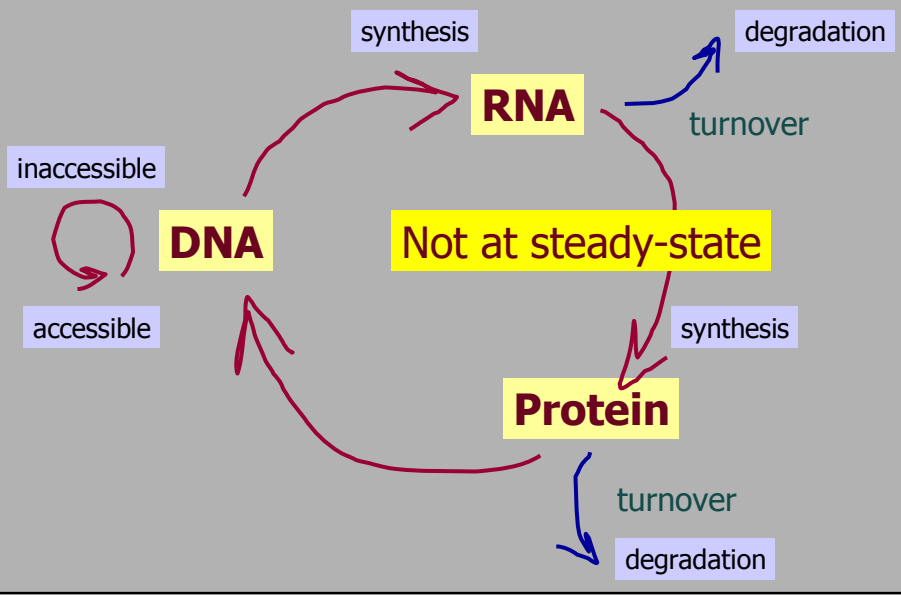
DNA: 1 coding sequence

Protein: 1 biochemical function

Information Flow



Dynamic Information Flow



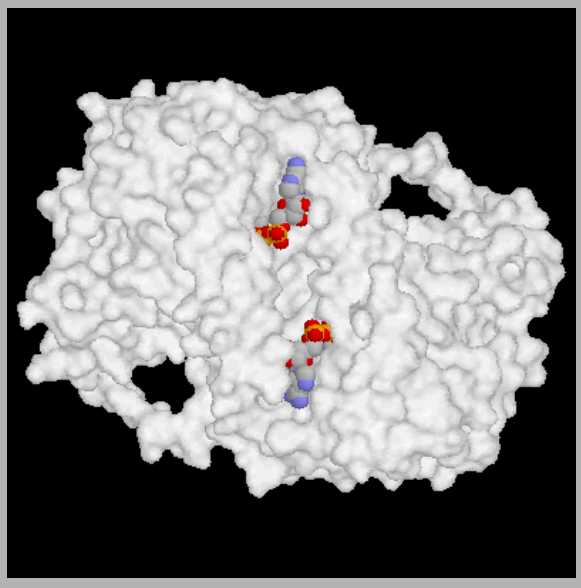
Environmental Genomics

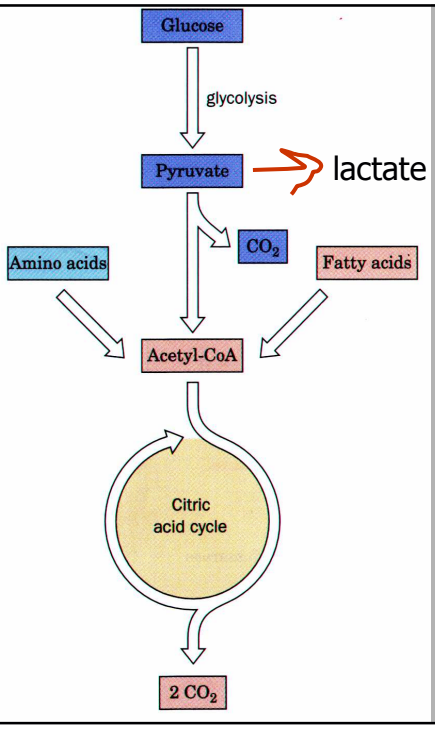
- If an organism is adapted to a specific environment . . .
- Then you should be able to look at any organism's genome and identify some characteristics of the environment that it inhabits.
- How does natural selection impact nucleotide sequence at the level of the whole genome?

Bioinformatics

- Is biological function inherent in primary sequence information?
- No, that's why biologists need sophisticated analytical and computational tools
- How do you account for 'structure' in sequence information content?
- Don't know, that's why biologists and computer scientists need to collaborate.

Lactate Dehydrogenase





Anaerobic Glycolysis

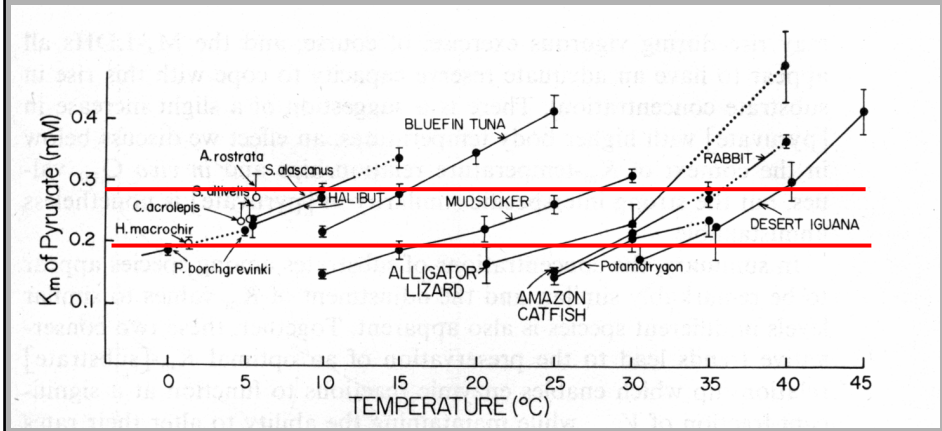
Pyruvate + NADH

← →

Lactate + NAD+

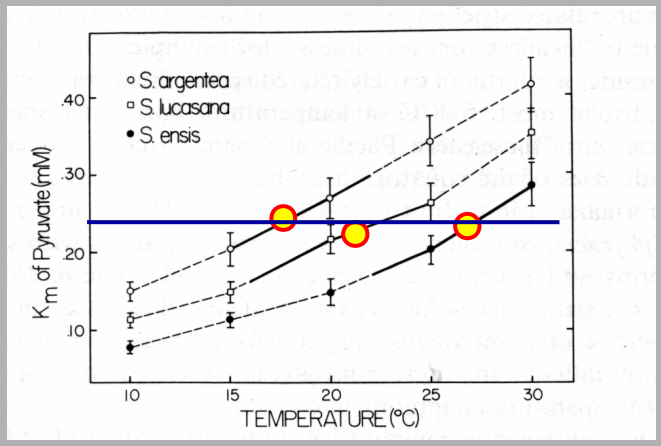
LDH: Structure

Metabolic Temperature Adaptation



- The affinity of LDH for pyruvate only exhibits a 2 fold change across a 45C temperature range in a phylogenetically constrained sample of vertebrates

Pacific Barracuda



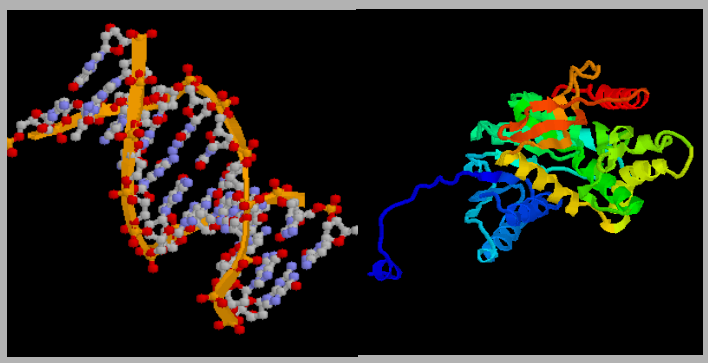
- Despite differences in thermal regimes, there appears to be an overall selective pressure among animals to maintain an LDH with a Km for pyruvate of ~0.25 mM

Barracuda LDH

S. argentea	**G H**	**H G S**
	GGCCAT	CATGGCAGC
S. lucasana	**D H**	**H G G**
	GACCAT	CATGGCGGC

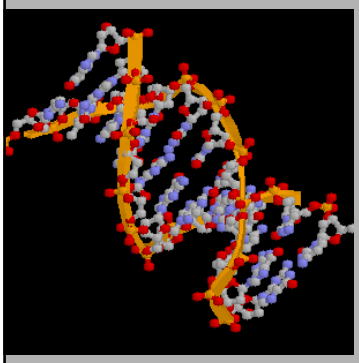
How could you ever decide that
GGCCAT vs. GACCAT
would be indicative of natural selection for
temperature adaptations in LDH?

Bridging the Gap



Genomics **Proteomics**

Bridging the Gap



function

Genomics **Proteomics**

Functional-Informatics

- Biologists need to know function
- Significance of any 'information' feature can only be assessed by describing function
- "Can the informatic feature contribute towards an understanding of how the organism works?"
- Describing a functional role leads to deriving testable hypotheses