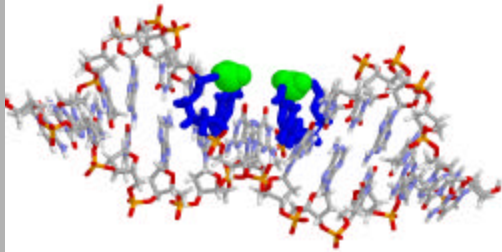
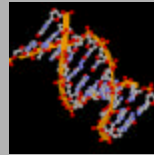


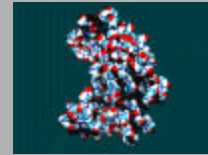
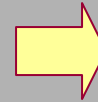
19 Gene Expression



Information Expression



1-D information array



3-D biochemical structure

How complex can a 4 letter code really be?

```

atcgggctatcgatagctatagcgcgatataatcgcgcgtatatgcgcgcattag
tagctagtctgattcatctggactgtcgaataatatacgcgccggctatcgcgct
atgcgcgataatcgcgcggcgcctataaaatataaaaaataaaatataatataatgc
tgcgcgatagcgcctataggcgcgctatccatataataggcgcctcgcggggcgcga
tgcacggctacggctagctgtagctagtcggcgattagcggcttatatcggcga
gcgatgagagtcgcggctataggcttaggctatagcgcctagtatatagcggctagc
cgcgtagacgcgatagcgtagctagcggcgcgcgtatatagcgcttaagagcca
aaatcgcgttagcgcctataatagcgcctatagctatagcggctattatagcga
gcgctagctagcgtatcaggcaggagatcgatgctactgatcgcgatgctagagca
gcgctatgctagtagtccatataatagctgagcgcgcgtagctcgatattacgcta
cctagatgctagcagctatgatcgtagca.....
    
```

Helicobacter pylori
 1.7 MB
 1,590 predicted ORFs
 91% of genome

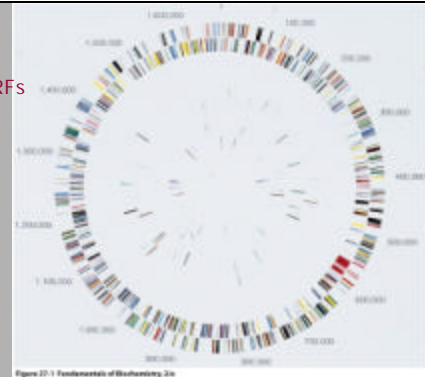


Figure 27-11 *Essentials of Biochemistry*, 2/e

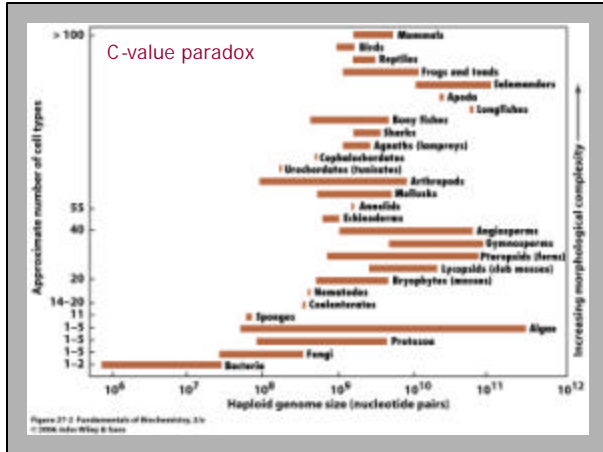
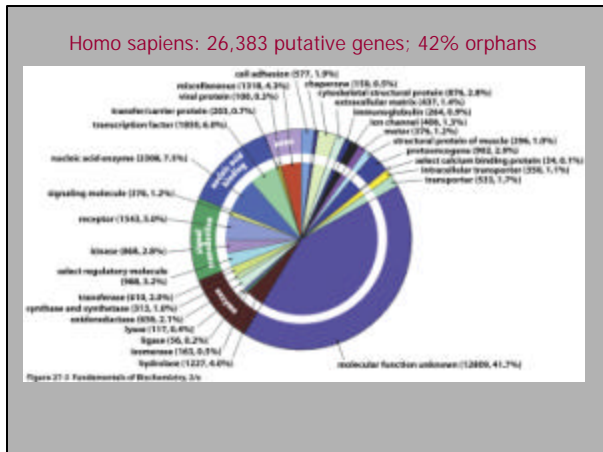


Table 27-1 Genome Size and Gene Number in Some Organisms

Organism	Genome Size (kb)	Number of Genes
<i>Haemophilus influenzae</i> (bacterium)	1,830	1,740
<i>Escherichia coli</i> (bacterium)	4,639	4,289
<i>Saccharomyces cerevisiae</i> (yeast)	11,700	6,034
<i>Caenorhabditis elegans</i> (nematode)	97,000	19,099
<i>Oryza sativa</i> (rice)	430,000	~35,000
<i>Arabidopsis thaliana</i> (mustard weed)	117,000	~26,000
<i>Drosophila melanogaster</i> (fruit fly)	137,000	13,061
<i>Mus musculus</i> (mouse)	2,500,000	~30,000
<i>Homo sapiens</i> (human)	3,200,000	~30,000

Table 27-1 Fundamentals of Biochemistry, 3/e
© 2004 John Wiley & Sons

Sequence divergence between human and chimp: <1%
Sequence divergence between strains of maize: >20%

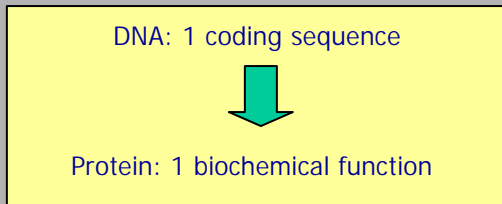


Genomic Isochores

- Most genomic sequence (vertebrates) have a high A-T composition
- Only 3% of the human genome has a high G-C content
- Yet 75% of the identified genes in humans reside within these G-C domains (only 3% of our genome).
- All DNA is not created equal

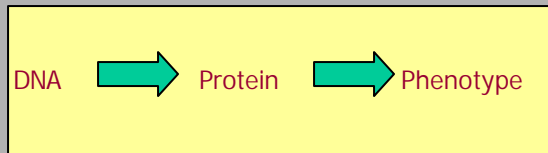
"gene"

Jacob & Monod (1961):
One gene equals one protein



1 gene = 1 protein

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



Absence of inducer

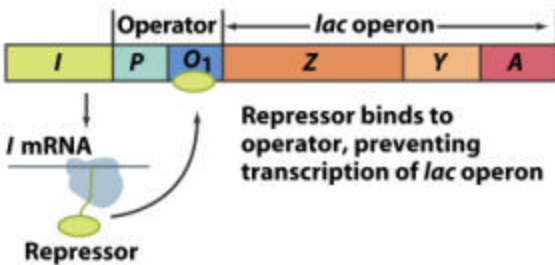


Figure 27-7b. Fundamentals of Biochemistry, 2/e
© 2008 John Wiley & Sons

Presence of inducer

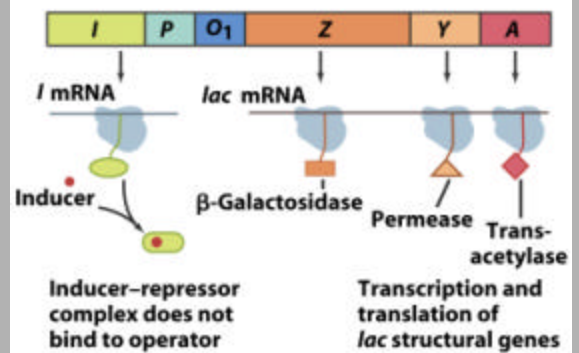


Figure 27-7b. Fundamentals of Biochemistry, 2/e
© 2008 John Wiley & Sons

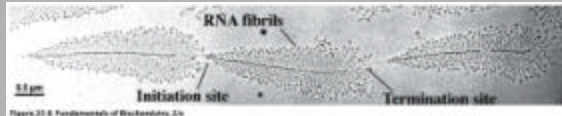
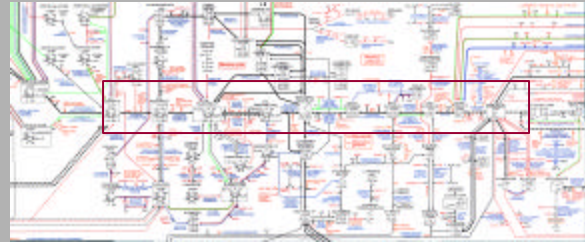


Figure 23-8 Fundamentals of Biochemistry, 5/e

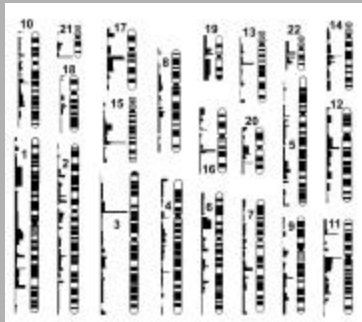
Glycolysis



glucose → pyruvate

Gene Distributions

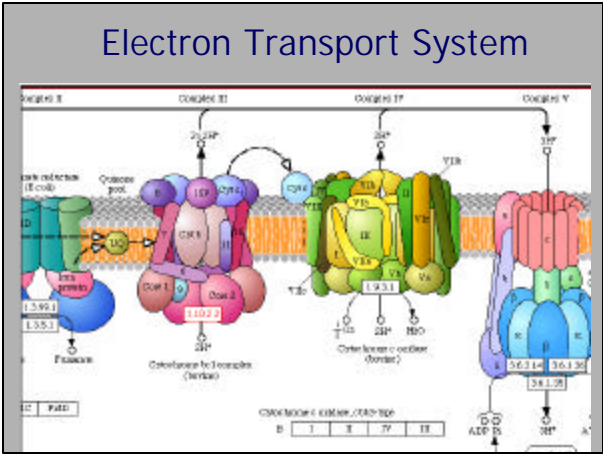
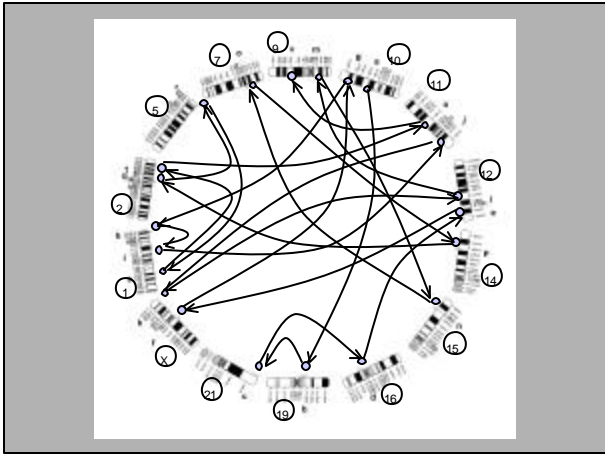
- Unequal density across chromosomes



Glycolytic Enzymes: gene loci

Table 1. A serial list of the enzymes involved in glycolysis and the TCA Cycle. Chromosomal locations were obtained from the LocusLink database and the loci codes used in Fig. 2 are indicated in the "map" column.

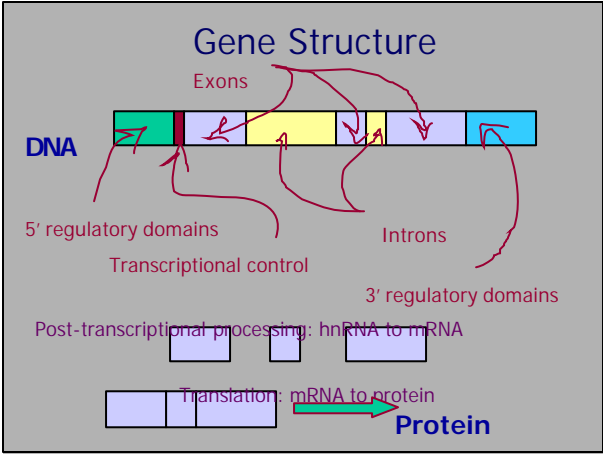
EC#	Enzyme Name	Locationmap
A 2.7.1.1	hexokinase	10q22a
B 5.3.1.9	glucose-P- isomerase	19q13.1b
C 2.7.1.11	6-phosphofruktokinase	21q22.3c
D 4.1.2.13	fructose-bis-P-aldolase	16q22-q24d
E 1.2.1.12	GAP- dehydrogenase	12p13e
F 2.7.2.3	phosphoglycerate kinase	xq13f
G 5.4.2.1	phosphoglycerate mutase	7q31-q34g
H 4.2.1.11	enolase	1p36.3-p36.2h
I 2.7.1.40	pyruvate kinase	1q21f
J 1.1.1.27	lactate dehydrogenase	11p15.4j
K 1.2.1.51	pyruvate dehydrogenase	xp22.2-p22.1k
L 4.1.3.7	citrate synthase	12q11-qterf
M 4.2.1.3	aconitase	9p22-p13m
N 1.1.1.41	isocitrate dehydrogenase	15q25.1n
O 1.2.4.2	α-keto dehydrogenase	7p14-p13o
P 2.3.1.61	S-succinyltransferase	14q24.3p
Q 6.2.1.4	succinylCoA synthetase	2p11.2q
R 1.3.5.1	succinic dehydrogenase	5p15r
S 4.2.1.2	fumarase	1q42.1s
T 1.1.1.37	malate dehydrogenase	2p13.3t
U 6.4.1.1	pyruvate carboxylase	11q13.4u
V 3.1.3.11	fructose-bisphosphatase	9q22.3v



Operon	-35 region	-10 region (Pribnow box)	Initiation site (+1)
<i>lac</i>	ACCCGAGCCTTAACTTTATGCTTCCGGCTCGATGTTGTGTGGATTGTGAGCGG	TATGTTGTGTGG	ATGTAAGCGG
<i>lacI</i>	CCATCGAATGGCCAAACCTTTCCGGTATGGCATGATAGCCCCGAAAGAAATC	TATGTTGTGTGG	ATGTAAGCGG
<i>galP2</i>	ATTTATTCATGTCACACTTTTCGCATCTTTGTATGCTATGGTTATTTTCATACCAT	TATGTTGTGTGG	ATGTAAGCGG
<i>araC</i>	GGATCCTACCTGACGCTTTTATCGCAACTCTCTACTGTTTCTCCATACCGTTTTT	TATGTTGTGTGG	ATGTAAGCGG
<i>araC</i>	BCCGTATTATGACACTTTTGTACGCGTTTTTTCATGCTTTGTTCCGCTTTG	TATGTTGTGTGG	ATGTAAGCGG
<i>trp</i>	AAATGAGCTGTGACATTAAATCATCGAATTAATTAATGACAGATTCACGTA	TATGTTGTGTGG	ATGTAAGCGG
<i>bioA</i>	TTCGAAACGTTTTTGTGTTAATTCGGTGAAGCTGTAACTTAACTT	TATGTTGTGTGG	ATGTAAGCGG
<i>bioB</i>	CATAATCGACTTGTAAACCAATTAAGAAATTAAGTTTACAACTTACACCGAAT	TATGTTGTGTGG	ATGTAAGCGG
<i>rRNA^{Tr}</i>	CAACGTAACACTTTACGCGGCGCTCATTTGATATGATGCGCCCGCTTCCGATA	TATGTTGTGTGG	ATGTAAGCGG
<i>rmsD</i>	CAAAAATACTTTGCGAAAAATGGGATCCCTAATGCGCTCCCTTGAACGA	TATGTTGTGTGG	ATGTAAGCGG
<i>rmsE</i>	CAATTTTTCTATTCGCGCTGCGGAACTCCCTATAATGCGCTCCATCGACAG	TATGTTGTGTGG	ATGTAAGCGG
<i>rmsI</i>	AAAATAAATGCTTAACTCTGATGCGGAAAGGCTATTATGCACACCCGCGCGCTG	TATGTTGTGTGG	ATGTAAGCGG

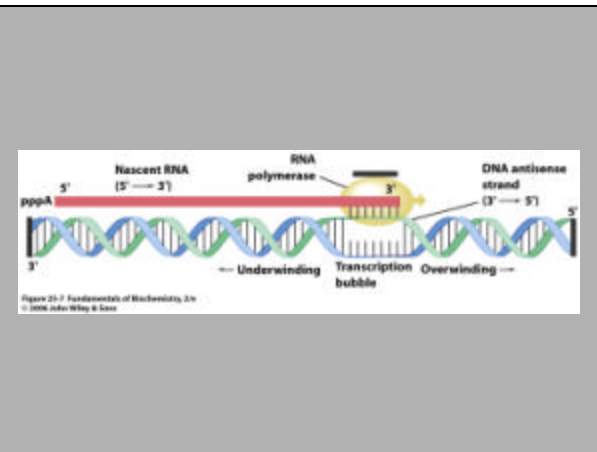
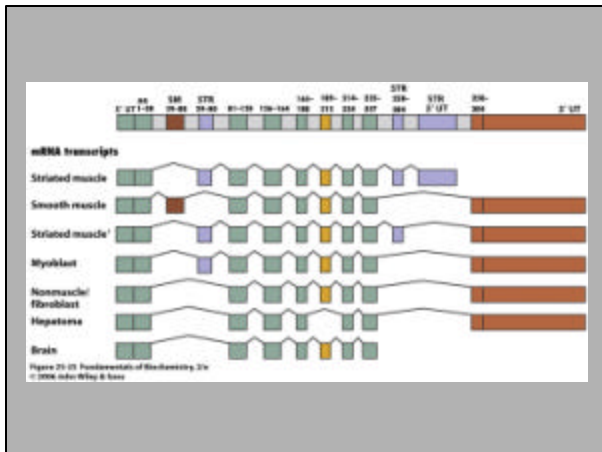
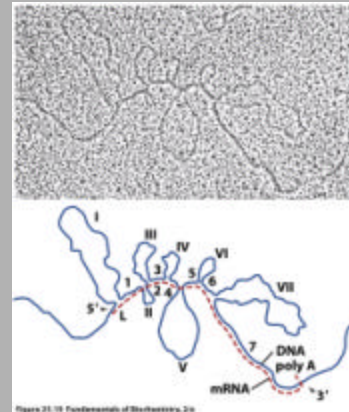
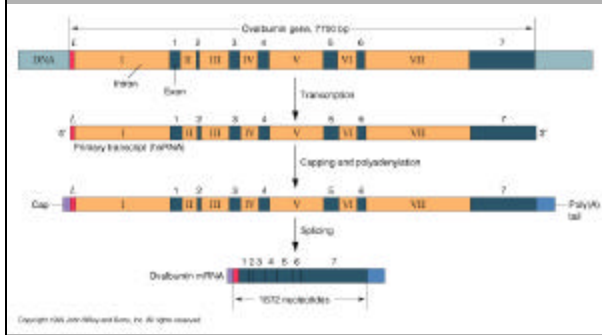
Consensus sequence:	-35 region	-10 region	Initiation site
	T T G A C A ... 16-19 bp ...	T A T A A T ... 5-8 bp ...	A
	69 79 61 56 54 54	77 76 60 61 56 82	31 1
			55 48
			42

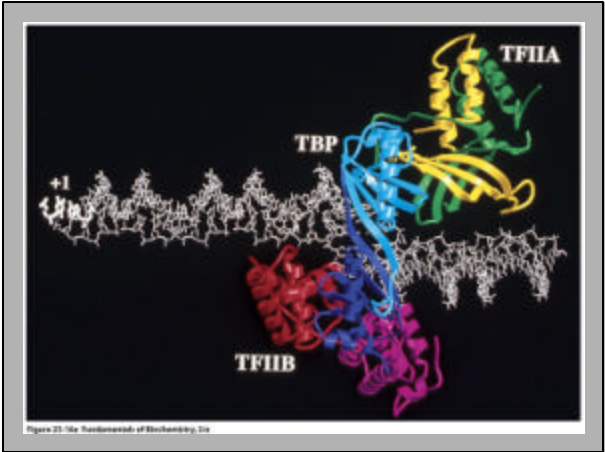
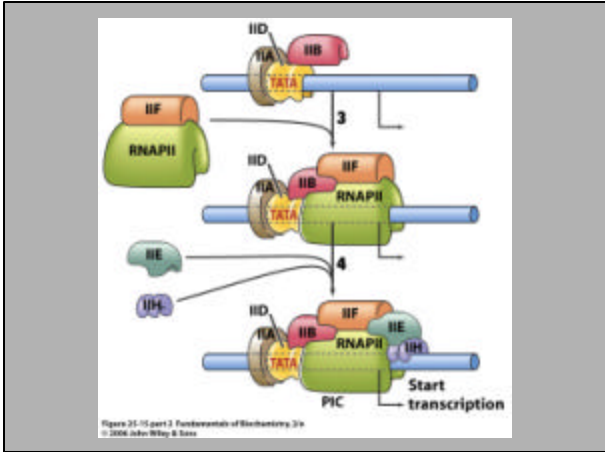
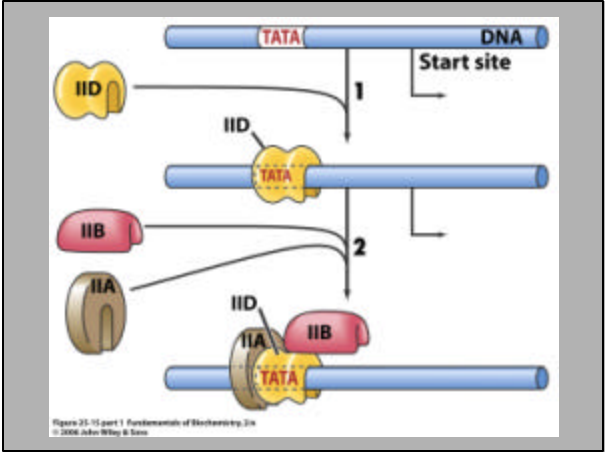
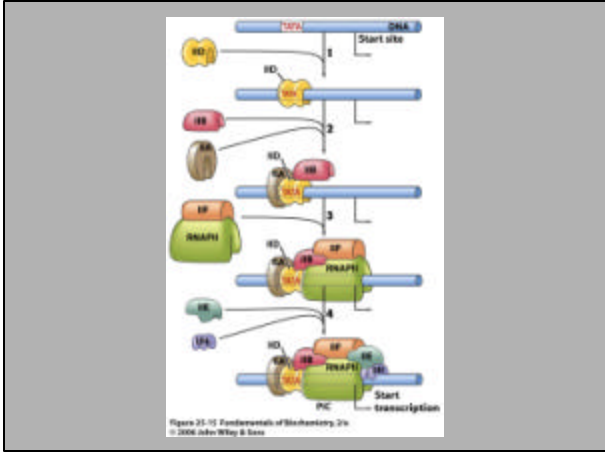
Figure 21-9. Fundamentals of Biochemistry, 5/e
© 2008 John Wiley & Sons



Chicken Ovalbumin Gene

Postranscriptional Processing:



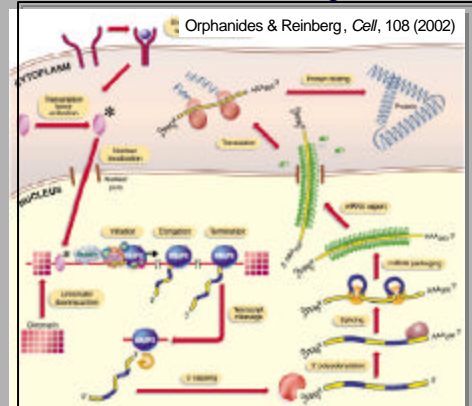


Traditional View of Gene Expression

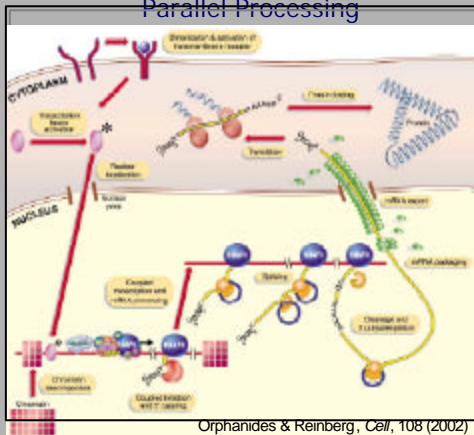
Gene sequence is translated to active protein in a series of discrete processes:

- Nucleus
 - Transcription (3 steps)
 - Capping
 - Polyadenylation
 - Splicing
 - Transport to cytoplasm
- Cytoplasm
 - Translation
 - Protein folding

Serial Processing



Parallel Processing



Epigenetic Regulation: Covalent modifications of DNA to alter expressed phenotypes

Most plants, animals and fungi utilize:

- Acetylation
- Methylation

- Phosphorylation
- ADP-ribosylation

for control of gene expression

Covalent modifications of DNA:

1. Acetylation of histones

Acetylation/De-acetylation

- Hypoacetylation of *Lys* residues
 - Heterochromatin & Silenced euchromatic genes
- Hyperacetylation of *Lys* residues
 - Euchromatin and Inducible genes
 - HS regions

Covalent modifications of DNA:

2. Methylation of histones

H3 methylation at Lysine 9 (H3-mLys9)

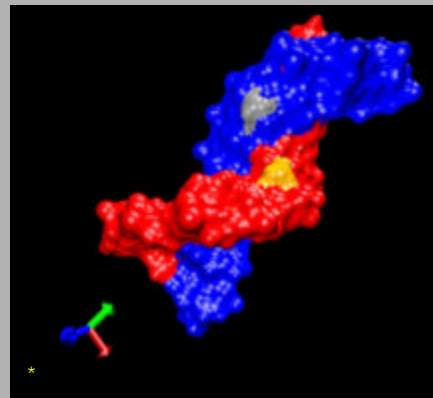
- Characteristic of the heterochromatic state
- Predominantly in pericentric DNA and repetitive DNA sites
- Absent in areas flanking silent genes
- Also associated with silencing of euchromatic genes

Covalent modifications of DNA:

3. Methylation of DNA

5'-Cytosine methylation (5mC)

- Most common form of DNA modification in eukaryotes
- Contributes to the stability of pericentromeric heterochromatin
- Maintains epigenetic expression states
- Silencing of imprinted genes

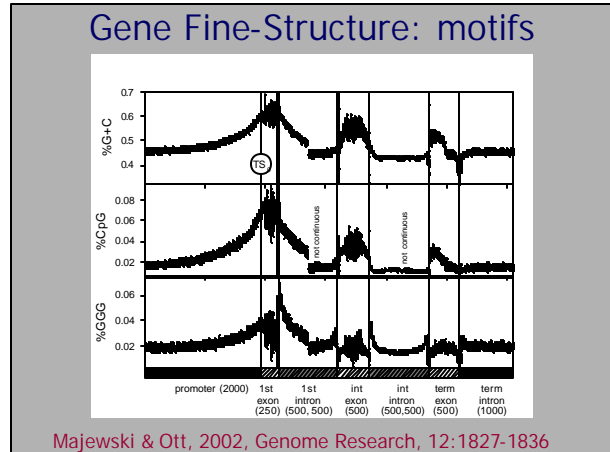
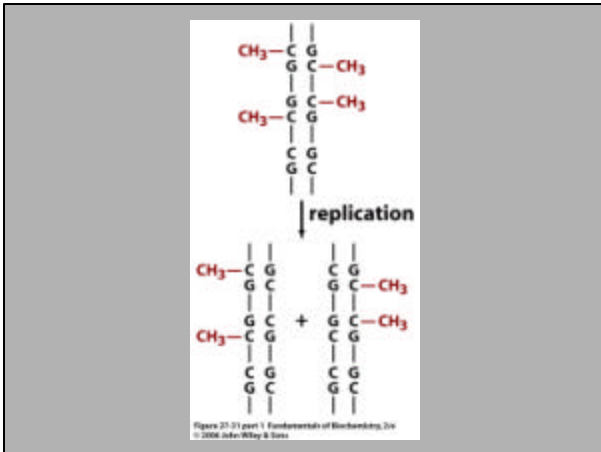
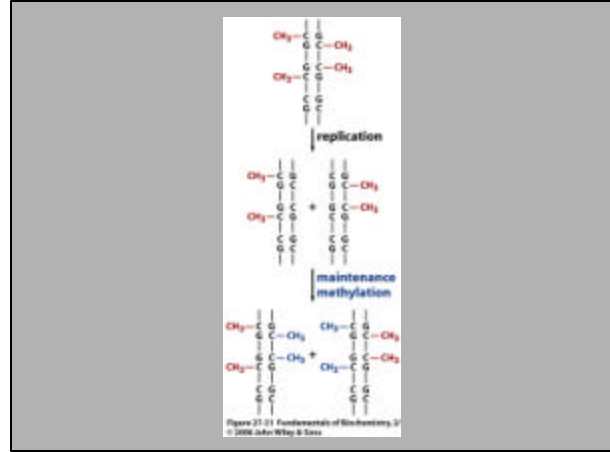
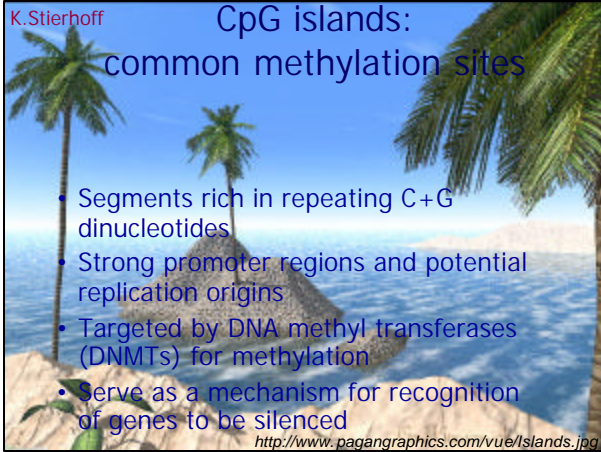


K. Stierhoff

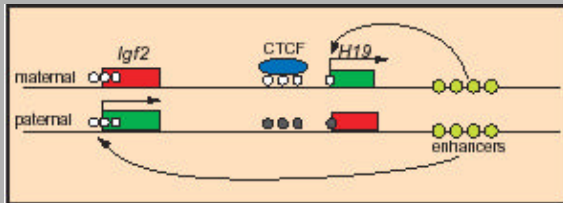
CpG islands: common methylation sites

- Segments rich in repeating C+G dinucleotides
- Strong promoter regions and potential replication origins
- Targeted by DNA methyl transferases (DNMTs) for methylation
- Serve as a mechanism for recognition of genes to be silenced

<http://www.pagangraphics.com/vue/islands.jpg>



Gene Imprinting



Regulation is not encoded in DNA:
Epigenetic determination

Imprinting: "non-genetic" inheritance



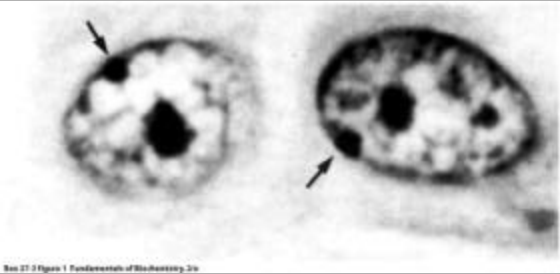
Parents matter. Hybrid of a horse and donkey, the hinny (foreground) differs from the mule because of parent-of-origin effects.

male horse x female donkey = hinny
female horse x male donkey = mule

Cloning alters imprinting



See 27.2 Figure 2. Fundamentals of Biochemistry, 5e



Box 27.3 Figure 1 Fundamentals of Biochemistry, 2/e