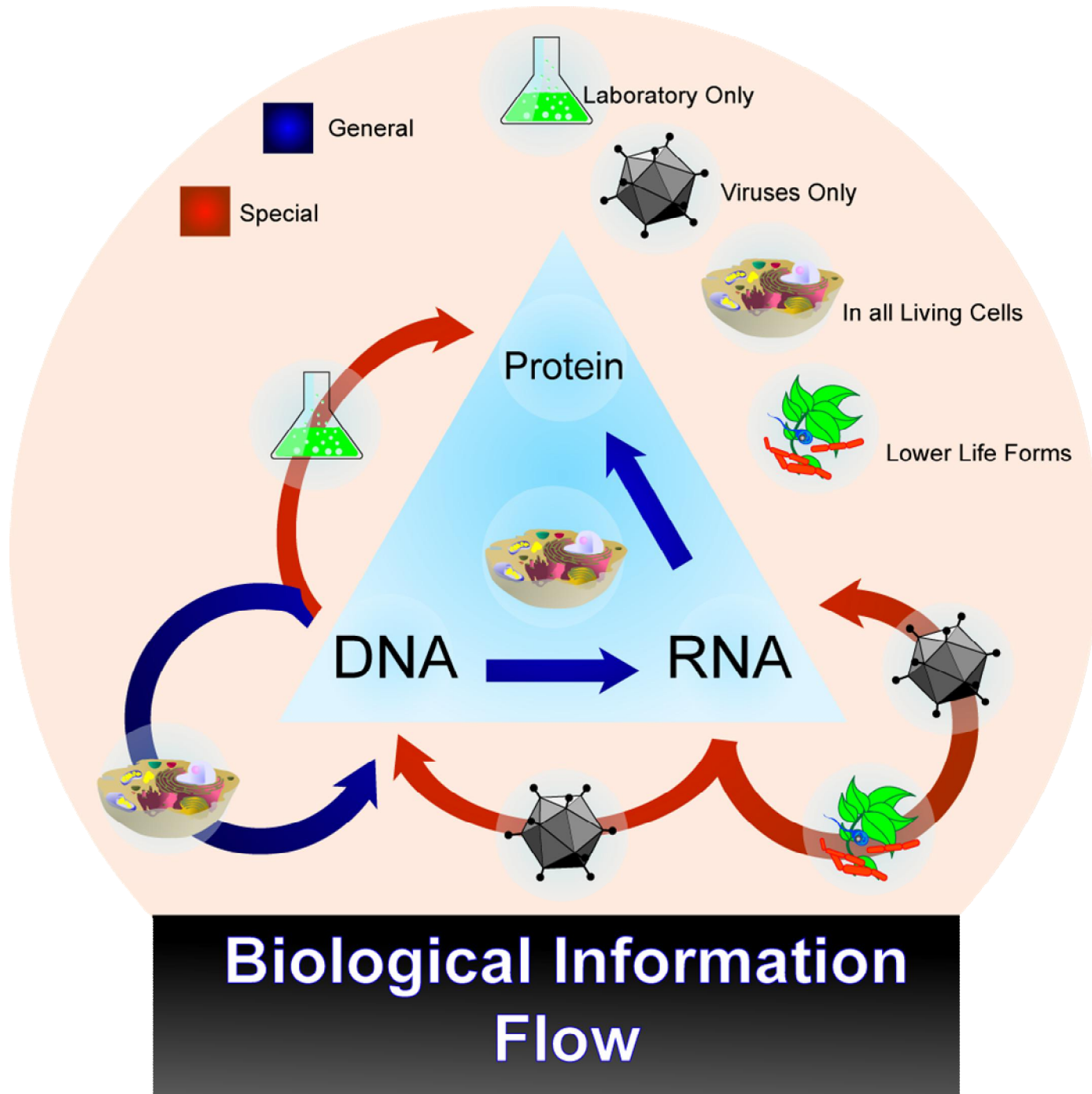


Molecular biology recap

Autumn 2007

Esa Pitkänen

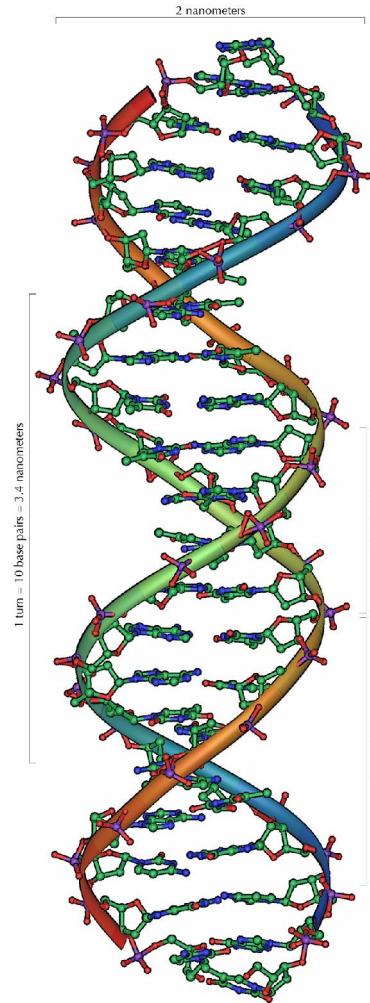
Master's Degree Programme in Bioinformatics (MBI)
Department of Computer Science, University of Helsinki
<http://www.cs.helsinki.fi/mbi/courses/07-08/itb/>



Biological Information Flow

Deoxyribonucleic acid (DNA)

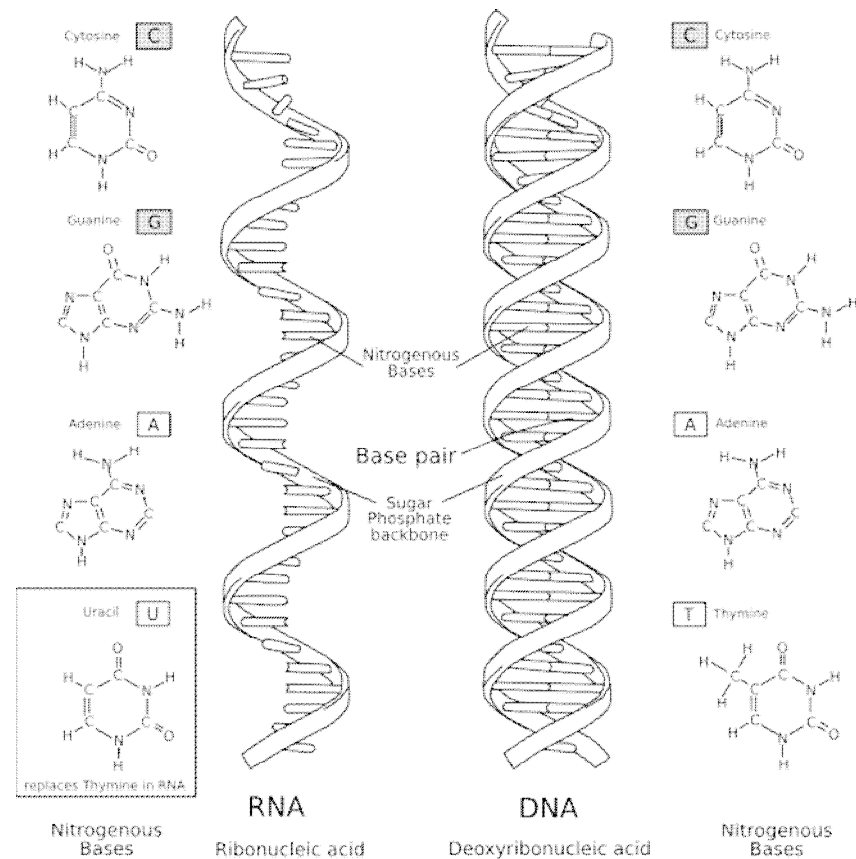
- Long-term storage of genetic information
- Information is encoded using the four bases **a**denine, **c**ytosine, **g**uanine and **t**hymine
- DNA sequence can thus be written as a string:
gcagcgcgcccctgcccagg...
- Utilised by transcribing DNA into RNA



Michael Ströck, <http://en.wikipedia.org/wiki/Dna>

Ribonucleic acid (RNA)

- "Working copy" of genetic instructions
- Four-letter alphabet
- Thymine replaced with **uracil**
 - gcagcgcgccugcccag
 - ...
- Translated into proteins



<http://en.wikipedia.org/wiki/Rna>

Proteins

- Complex macromolecules
- Many functions in the cell
- Composed of a linear sequence of amino acids
- Twenty different amino acids → twenty-letter alphabet
 - AEGLV...WKKLAG



Triplet code for proteins

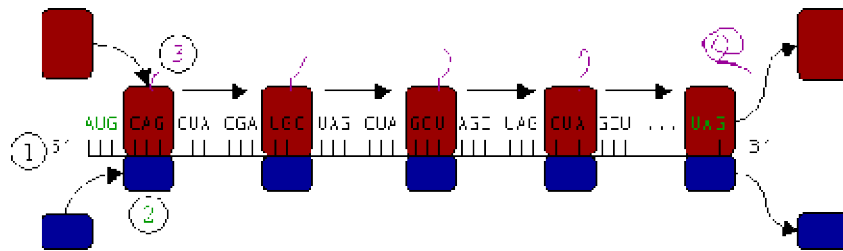
- Three consecutive bases specify one amino acid according to the coding table on the right
- Our example:

gcagcgcgcccugcccag

A A P P A Q

		Second letter				
		U	C	A	G	
First letter	U	UUU Phenyl-alanine UUC UUA Leucine UUG	UCU Serine UCC UCA UCG	UAU Tyrosine UAC UAA Stop codon UAG Stop codon	UGU Cysteine UGC UGA Stop codon UGG Tryptophan	U C A G
	C	CUU Leucine CUC CUA CUG	CCU Proline CCC CCA CCG	CAU Histidine CAC CAA Glutamine CAG	CGU Arginine CGC CGA CGG	U C A G
	A	AUU Isoleucine AUC AUA AUG Methionine; start codon	ACU Threonine ACC ACA ACG	AAU Asparagine AAC AAA Lysine AAG	AGU Serine AGC AGA Arginine AGG	U C A G
G	GUU Valine GUC GUA GUG	GCU Alanine GCC GCA GCG	GAU Aspartic acid GAC GAA Glutamic acid GAG	GGU Glycine GGC GGA GGG	U C A G	

www.bioalgorithms.info



Mark Dominus, <http://en.wikipedia.org/wiki/Ribosome>

From protein structure to function

- Protein 3D structure determined by the amino acid sequence
- Protein function determined by the 3D structure
- By looking at the genome, we can thus learn about the functional level (*phenotype*) of cells

References

- Richard C. Deonier, Simon Tavaré and Michael S. Waterman. Computational Genome Analysis, An Introduction. Springer, 2005.
- Neil C. Jones, Pavel A. Pevzner, “An Introduction to Bioinformatics Algorithms”.
- Alberts, Bruce, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter. Molecular Biology of the Cell. New York: Garland Science. 2002.
- Campbell, Neil. Biology, Third Edition. The Benjamin/Cummings Publishing Company, Inc., 1993.