Gene Expressions

Computational Biology camp for high school students

August 22, 2012

MIRcore

Inhan Lee, Ph.D.
inhan@mircore.org
http://mircore.org

Values

Science

Compassion

Efficiency

In Computational Genomics Camp

- Focus: Humans
- Hypothesis: diseases are conditions that occur when our body fail to maintain homeostasis.
- Ideal: preventive personalized medicine
- Current goal: find potential cure for cerebral palsy in the general population by focusing on one gene per camper (or group)

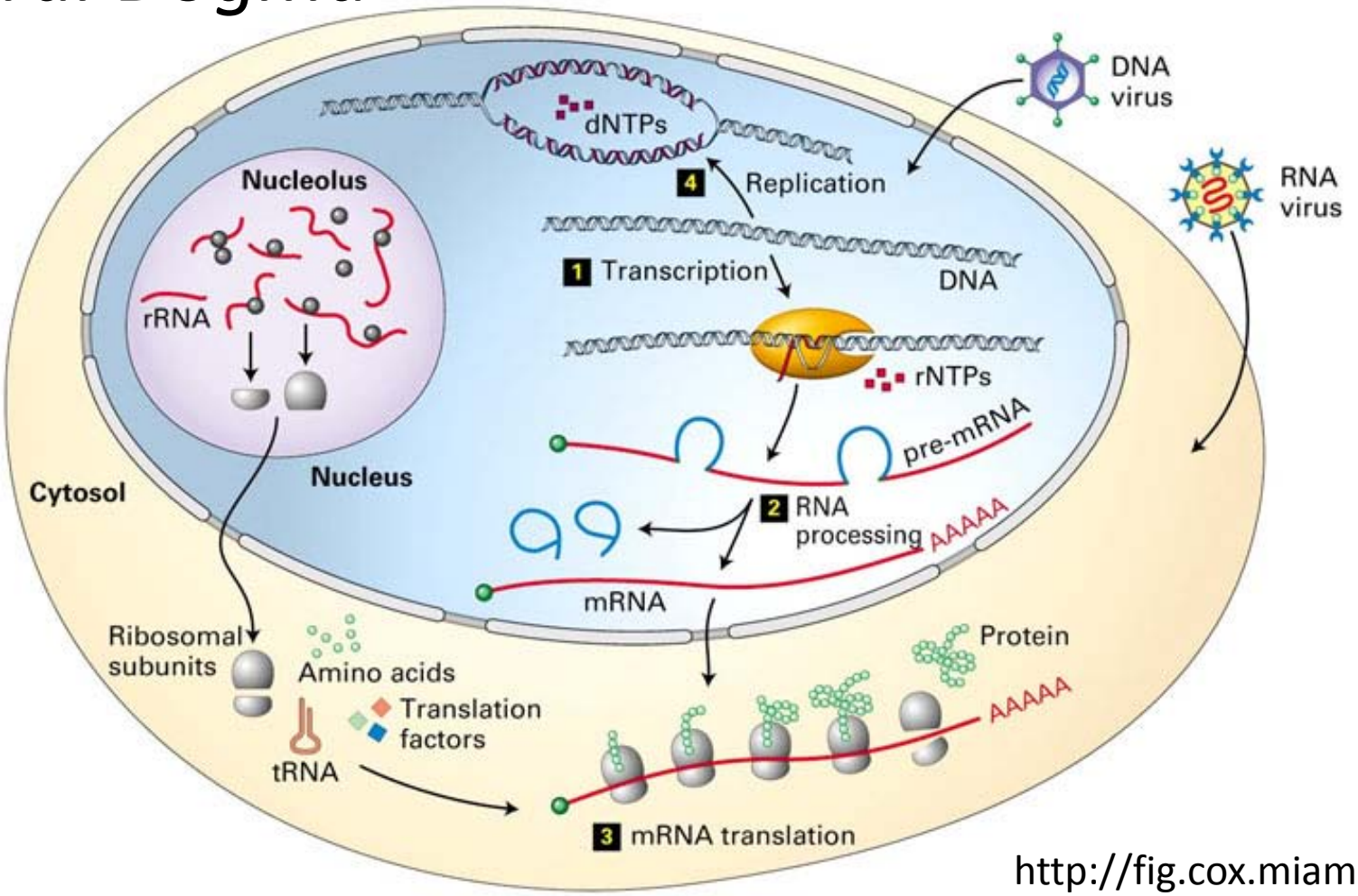
Recap – day 1

currently

- Genes: inherited molecules (DNA sequences) that generate functional entities (proteins or RNAs).
- Genome: The sum of genes
- Genomics: The study of all genes at DNA, RNA, or protein levels.
- Diseases: consider multiple genetic networks

Recap – day 2

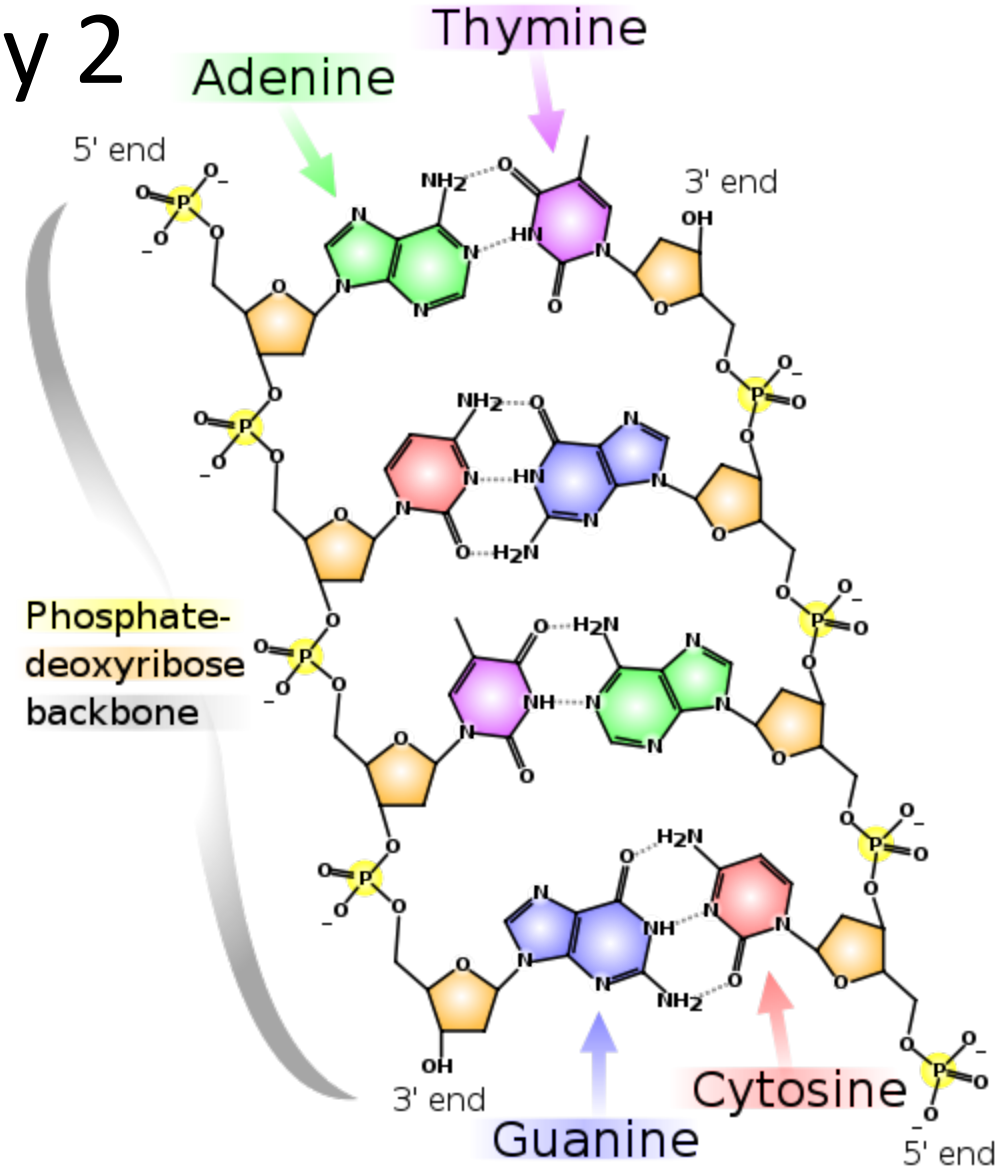
Central Dogma



http://fig.cox.miami.edu

Recap – day 2

DNA



Wikipedia: Madeleine Price Ball

Structures!

- RCSB Protein Data Bank

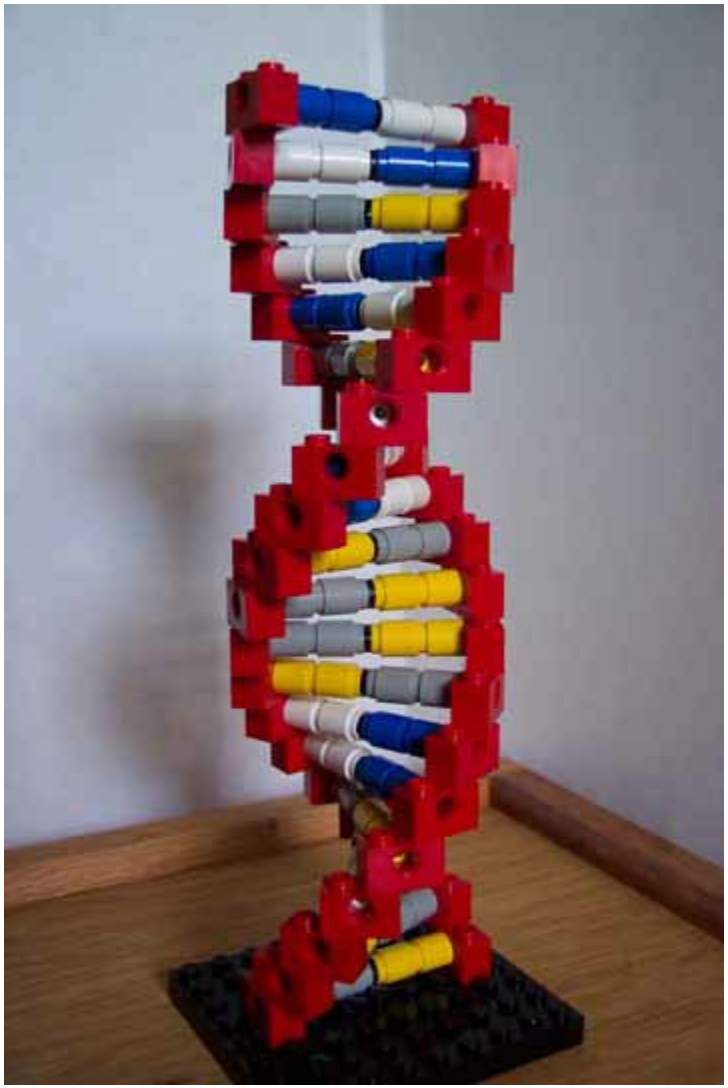
http://www.rcsb.org/pdb/home/home.do

Search double-helix DNA

Or 1CGC

Recap – day 2

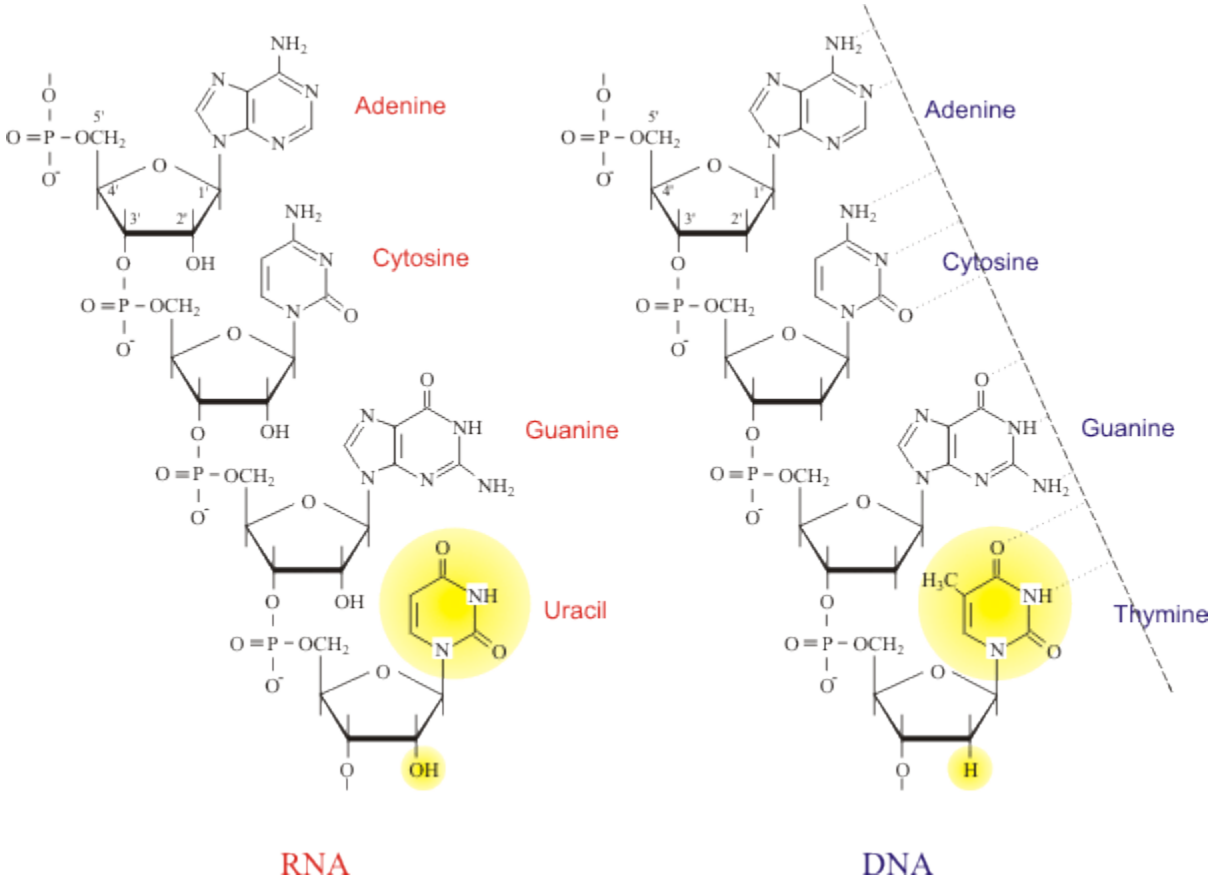
DNA



http://www.ericharshbarger.org/lego/mini_dna.html

Recap – day 2

RNA



Source: periodni.com

Structures!

- RCSB Protein Data Bank

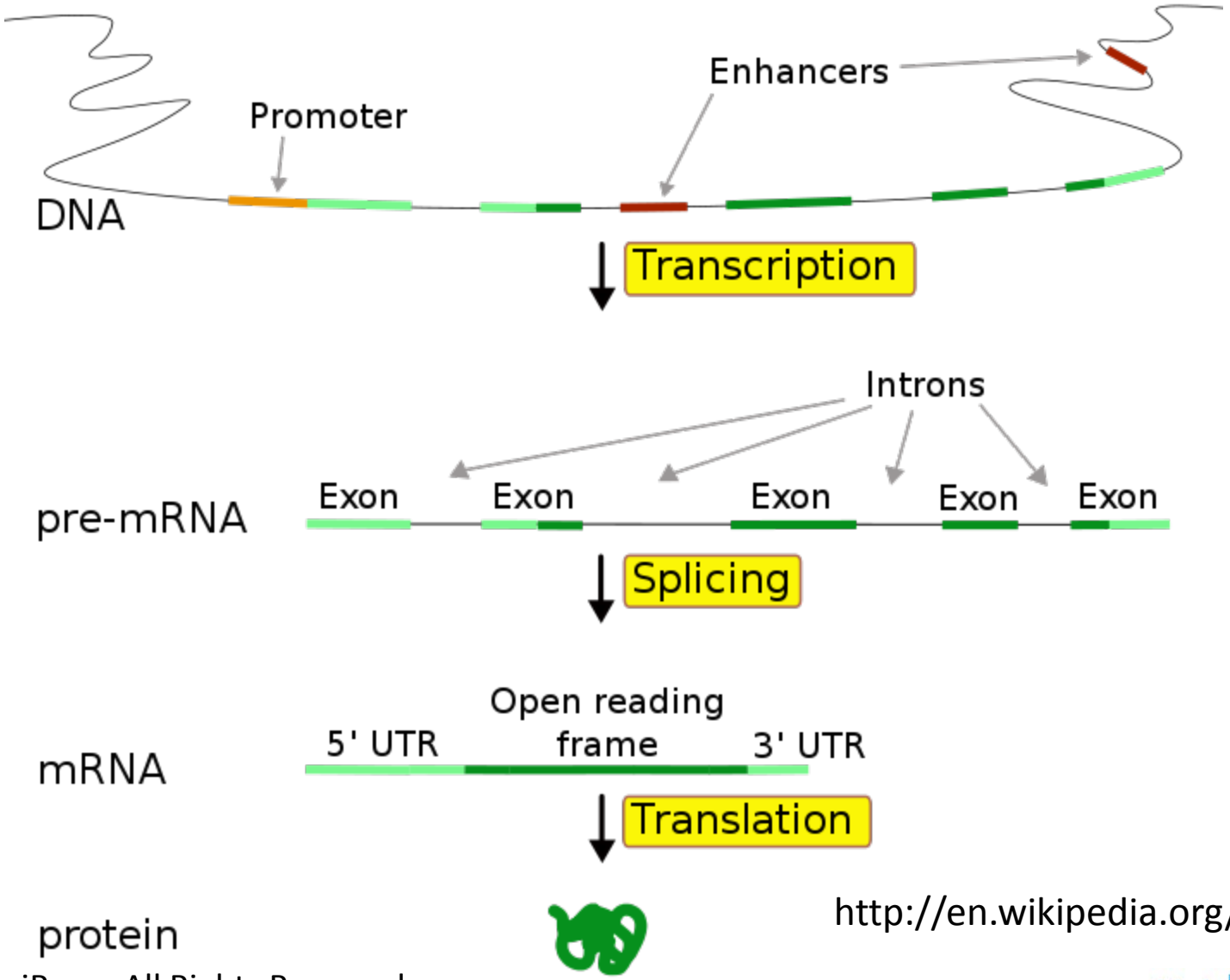
http://www.rcsb.org/pdb/home/home.do

Search RNA

or 4TNA

or 1C2X

Recap – day 2



http://en.wikipedia.org/wiki/Gene

How do we measure them?

DNA

- What do we want to measure?

Sequence Code to Amino Acids

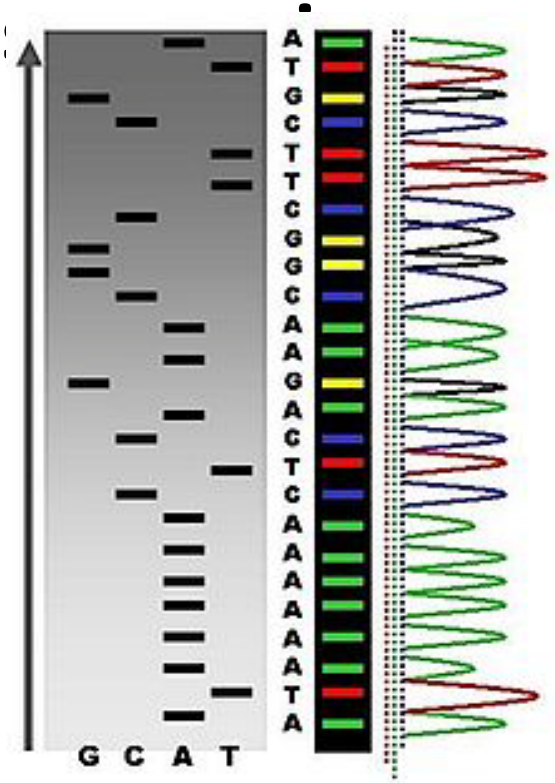
1st base	2nd base								3rd base
	U		C		A		G		
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	(Leu/L) Leucine	UCA		UAA	Stop (*Ochre*)	UGA	Stop (*Opal*)	A
	UUG		UCG		UAG	Stop (*Amber*)	UGG	(Trp/W) Tryptophan	G
C	CUU		CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	(Gln/Q) Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine	A
	AUG[A]	(Met/M) Methionine	ACG		AAG		AGG		G
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	(Glu/E) Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G

Source: http://en.wikipedia.org/wiki/Genetic_code

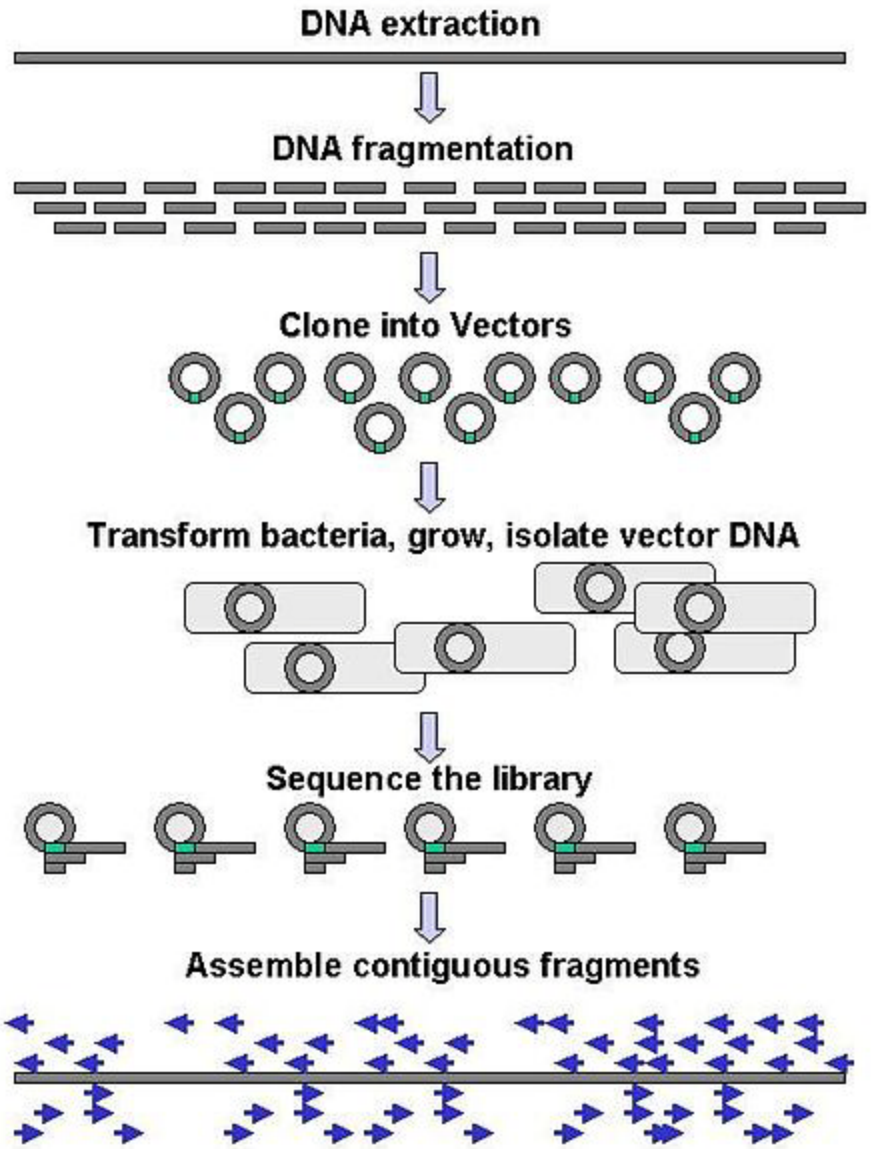
BY PROFESSOR CHARLES PUSEY

DNA sequencing

- **Dye terminator**



Attribution: Abizar at en.wikipedia



Disruptive Technology: High Throughput Sequencing (next generation sequencing)

Prepare gDNA Library
3 hours hands-on, 6 hours total time

Generate Clusters
<1 hour hands-on time, 5 hours total time

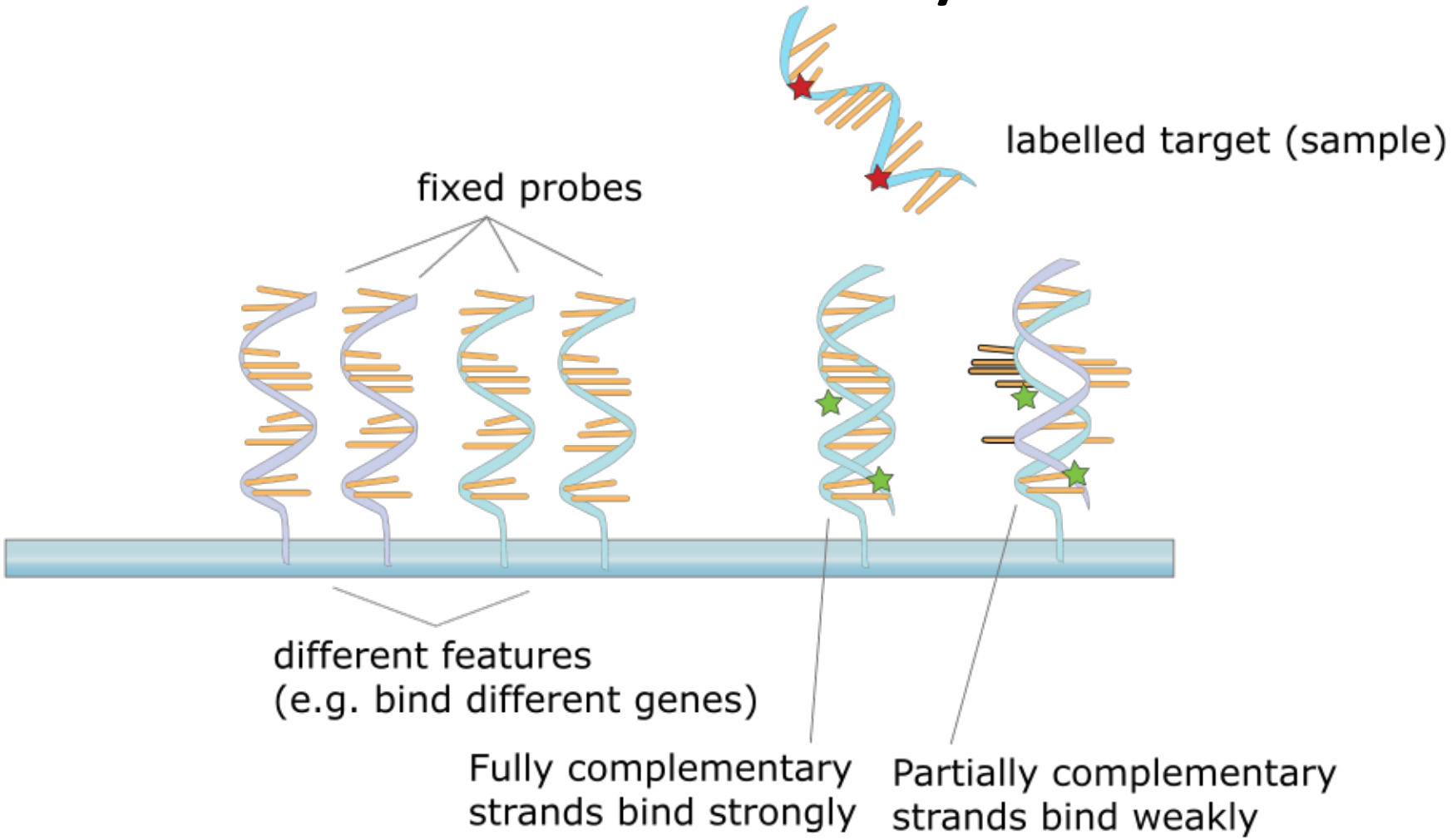
Sequence Clusters
2.5 days single read (36 bases)

Illumina

RNA

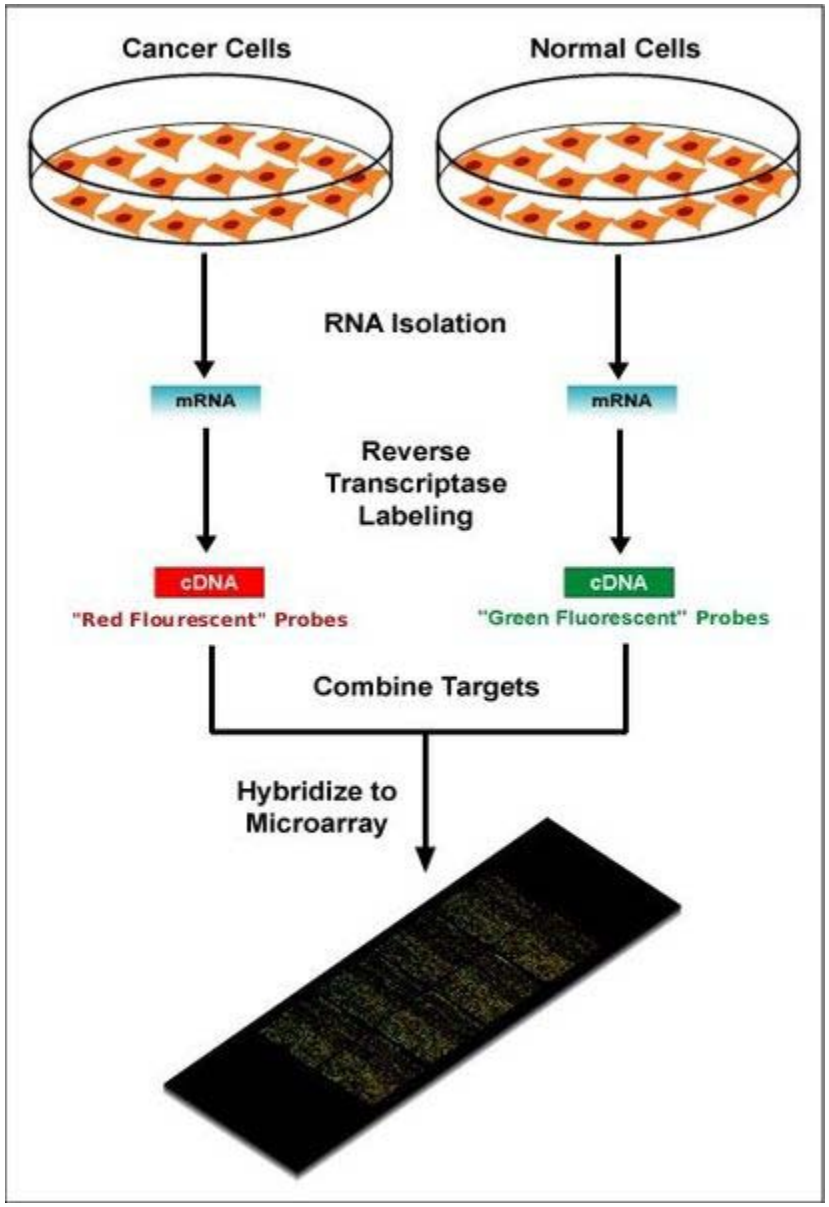
- What do we want to measure?

Microarray



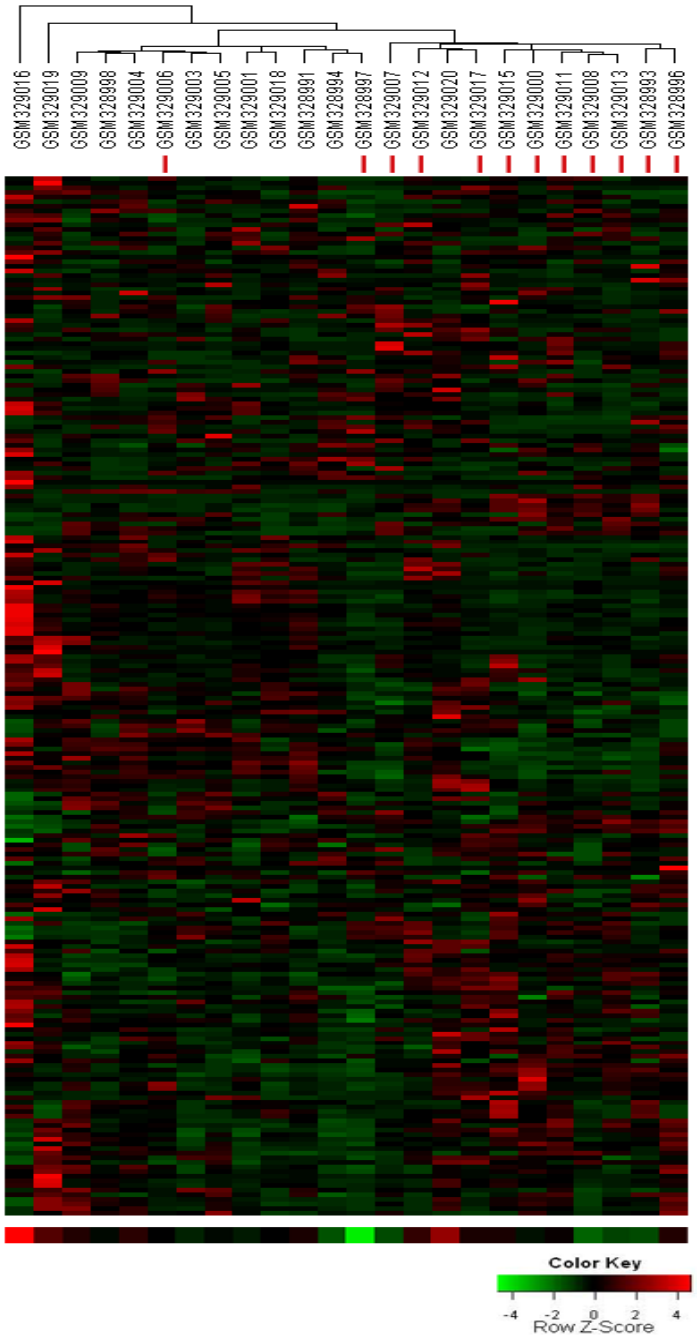
http://en.wikipedia.org/wiki/DNA_microarray

Compare expressions



http://en.wikipedia.org/wiki/DNA_microarray

heatmap

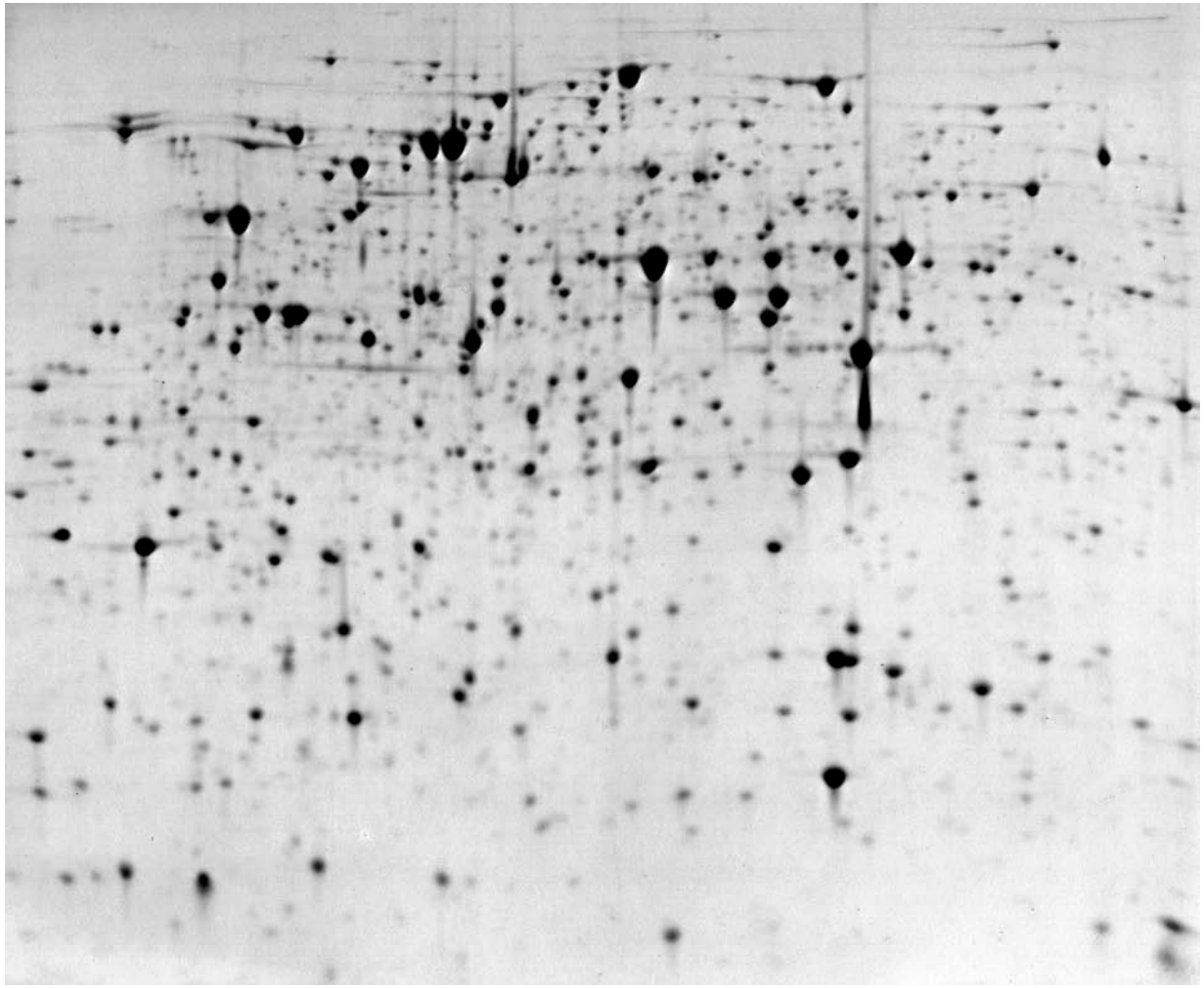


miRcore

Protein

- What do we want to measure?

2D gel electrophoresis



abdn.ac.uk

Tasks for your gene

- Find chromosome positions (start and end)
- Find entire DNA sequence with exons in capitals and introns in lower case.
- Find entire mature mRNA sequence with coding regions in capitals and UTRs in lower case.
- Find only coding regions.
- Find amino acid
- Connect 11^{th} – 20^{th} amino acids to the corresponding RNA coding sequences

Database to use:

http://genome.ucsc.edu/cgi-bin/hgGateway

http://www.ncbi.nlm.nih.gov/gene

Make dsDNA of your gene

- Choose 50 sequences including first exon and first intron
- Make single strand of DNA X of 50 sequences
- Mark exon and intron boundary
- Make complementary DNA strand Y of X
- Mark 5’-end of each strand
- Make dsDNA helical structure