

***SYLICA 2013***  
**Bowater lectures**

**Using 'Omic Technologies to  
Investigate Gene Function**

# Bowater Lectures in Brno, Feb. 2013

4 lectures on linked topics will be delivered during the coming week:

- *Contemporary DNA Sequencing Technologies – 26/2/2013 @ 10:00*
- *Using 'Omic Technologies to Investigate Gene Function – 26/2/2013 @ 14:00*
- *Biophysical Methods to Study Molecular Interactions – 27/2/2013 @ 10:00*
- *Synthetic Biology & Nanotechnology: Tomorrow's Molecular Biology? – 28/2/2013 @ 10:00*

# Genomics, 'Omics & Technology

- Molecular biology: major scientific discipline for past ~50 years
- *Genomics* = “analysis of genomes”: became important science during 1990’s
- Analyses of various other biological molecules have developed into their own scientific disciplines; e.g. *Metabolomics* = “analysis of metabolites”, etc.
- *Transcriptomics/Proteomics*: developed during past 10-15 years
- Bioinformatics: has developed as major branch of science - enables efficient analysis of data from “omics” experiments

# Genomics & Technology

- Significance of “omics” coincides with dramatic improvements in different technologies:
  - molecular biology: increased range of approaches for purification and manipulation of proteins and nucleic acids
  - computers: required for gathering and analysis of data
  - internet: allows data to be shared, quickly and easily
- All developments have increased speed and cost-effectiveness - available to much wider audience

# Transcriptomes

- *Genome*: all of hereditary information encoded in the DNA (or RNA)
- *Transcriptome*: set of all mRNAs ("transcripts") produced from a genome
- Term can be applied to:
  - complete set of transcripts for a given organism
  - specific subset of transcripts present in a particular cell type or under specific growth conditions
- Transcriptome varies because it reflects genes that are actively expressed at any given time

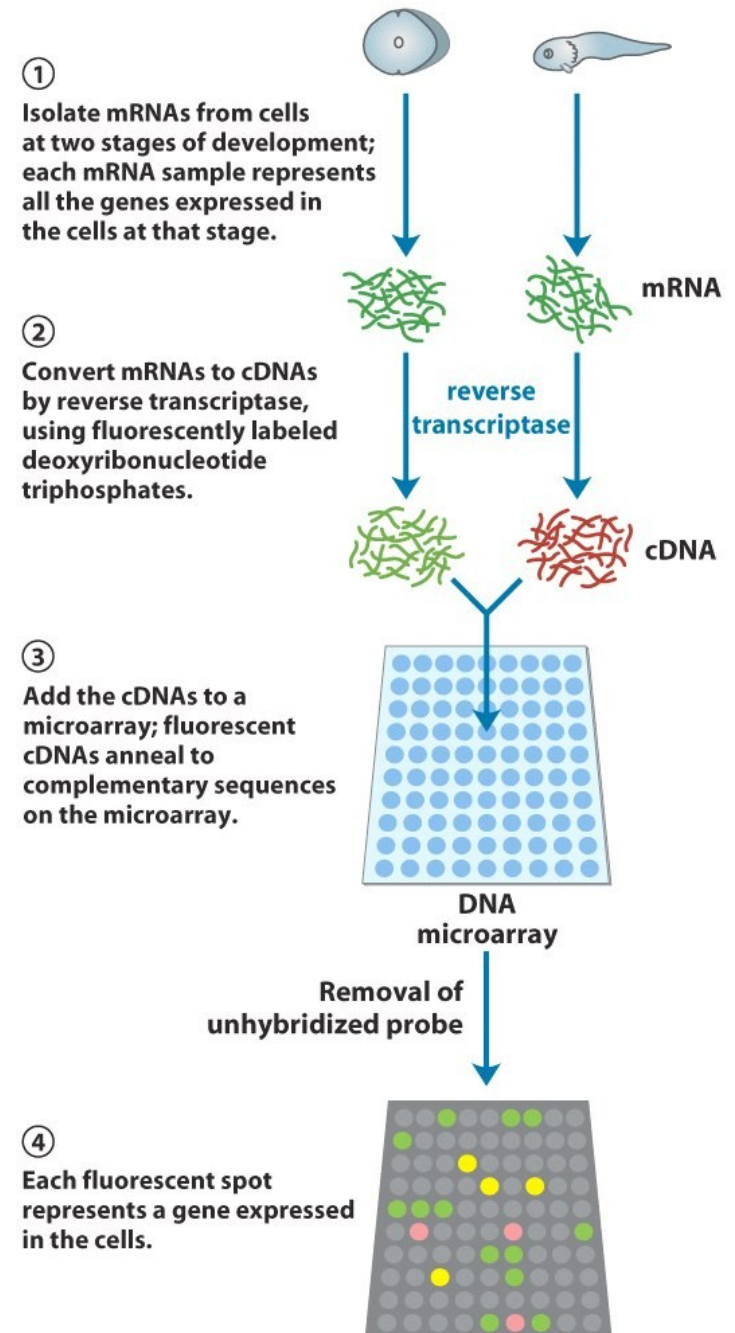
# DNA Microarrays Show Differences in Gene Expression

- Microarray chips contain fragments from genes in the group to be analyzed
  - Full genome of bacteria or yeast, or protein families from larger genomes
- mRNA or cDNA from different samples are differentially tagged
- Analysis on the same chip shows differences

# Transcriptomics

- *Transcriptomics* uses high-throughput techniques based on DNA microarrays
- For further details about microarrays see Lucchini *et al.*, *Microbiology*, **147**, 1403-1414 (2001)

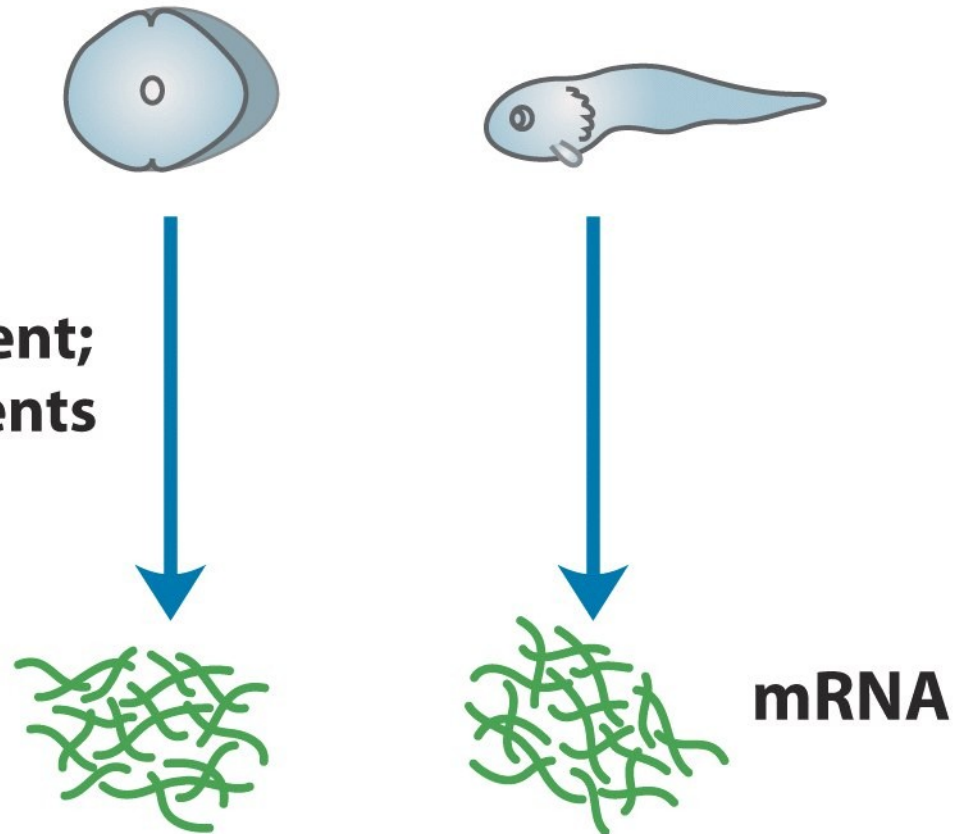
Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328



# Transcriptomics

①

**Isolate mRNAs from cells at two stages of development; each mRNA sample represents all the genes expressed in the cells at that stage.**



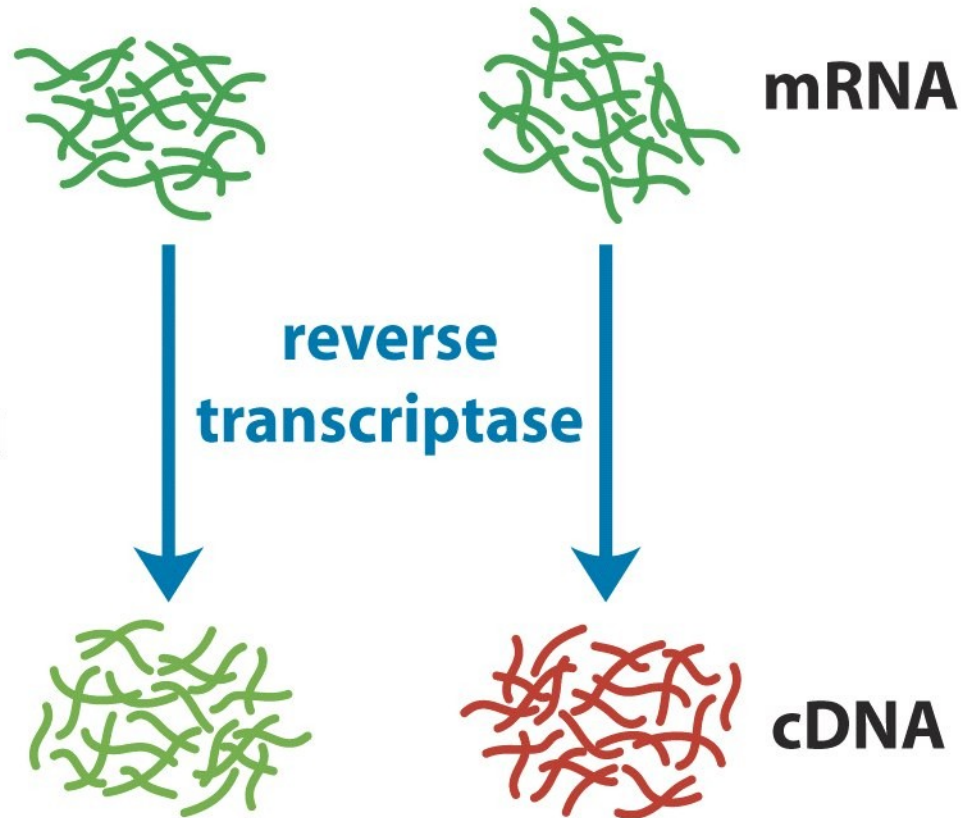
Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328



# Transcriptomics

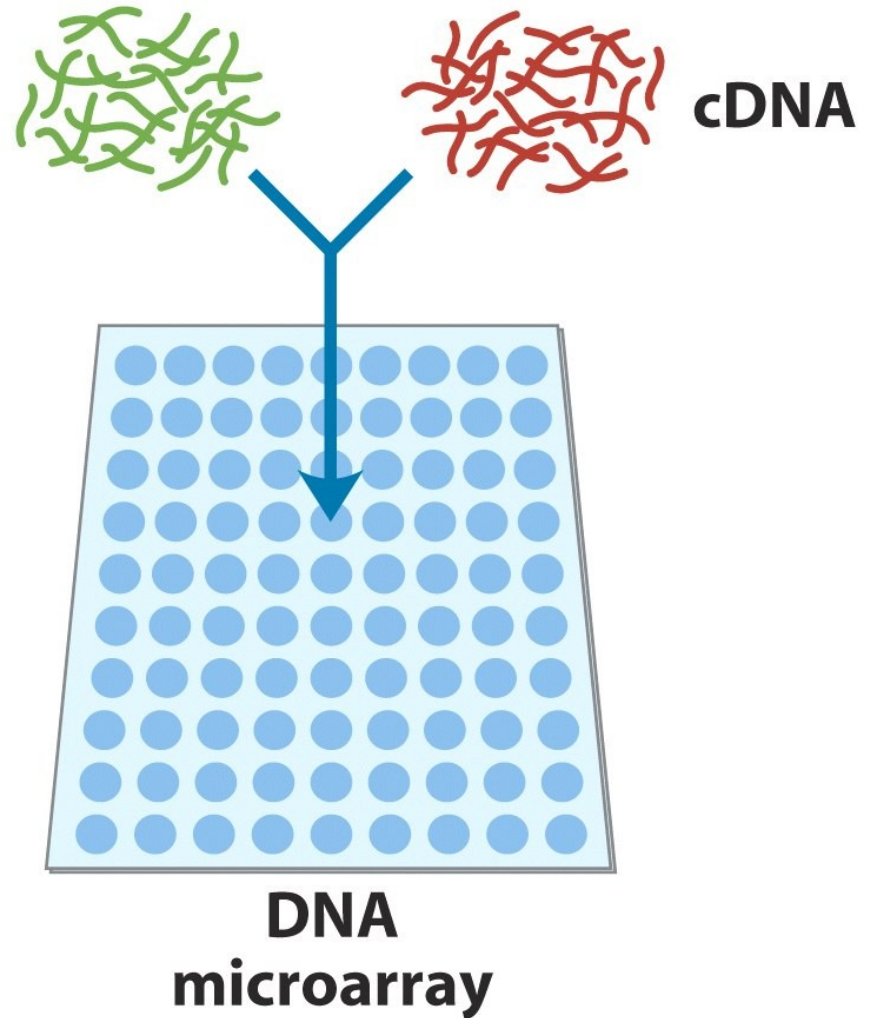
②

**Convert mRNAs to cDNAs by reverse transcriptase, using fluorescently labeled deoxyribonucleotide triphosphates.**



Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328

# Transcriptomics



③

**Add the cDNAs to a microarray; fluorescent cDNAs anneal to complementary sequences on the microarray.**

Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328

# Transcriptomics

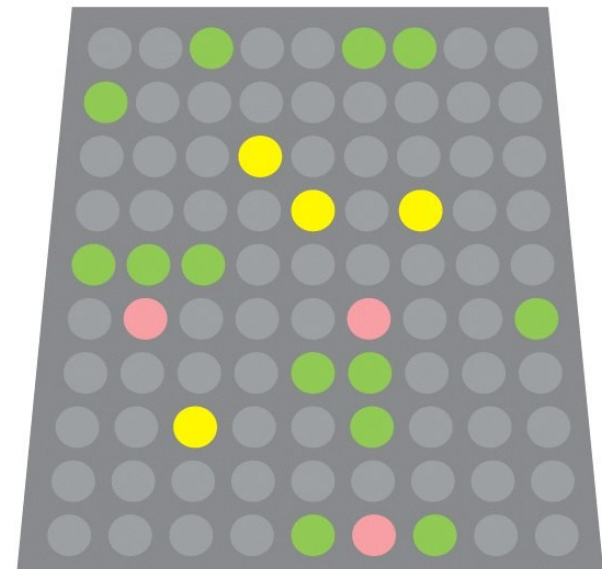
DNA  
microarray

Removal of  
unhybridized probe



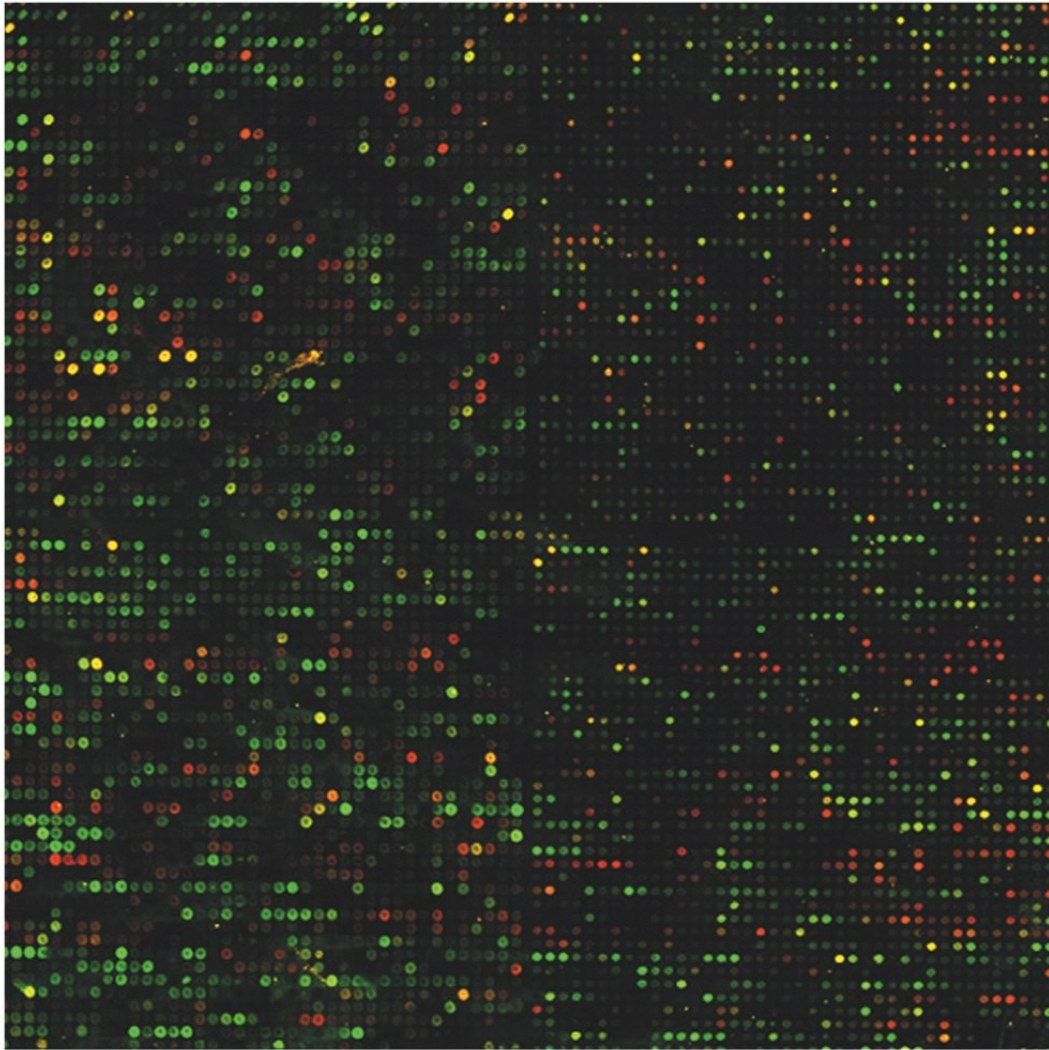
④

Each fluorescent spot  
represents a gene expressed  
in the cells.



Nelson & Cox, "Lehninger, Principles of  
Biochemistry", 4<sup>th</sup> edn, 2004, p. 328

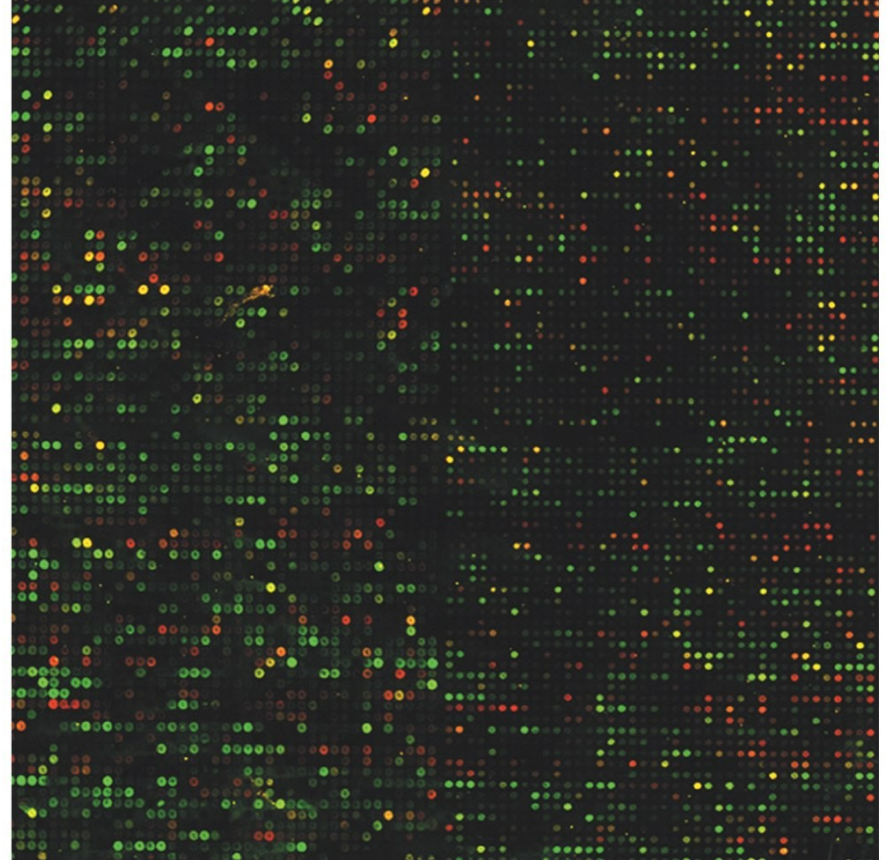
# Transcriptomics



Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328

# Transcriptomics

- Experiments performed under different conditions
- Determines effect of conditions on expression
- Produces huge amount of data
- Lots of repeats required - expensive



Nelson & Cox, "Lehninger, Principles of Biochemistry", 4<sup>th</sup> edn, 2004, p. 328

# Polymerase Chain Reaction (PCR)

- Used to amplify DNA in the test tube
  - Can amplify regions of interest (genes) within DNA
  - Can amplify complete circular plasmids
- Mix together
  - Target DNA
  - Primers (oligonucleotides complementary to target)
  - Nucleotides: dATP, dCTP, dGTP, dTTP
  - Thermostable DNA polymerase
- Place the mixture into thermocycler
  - Melt DNA at  $\sim 95^{\circ}\text{C}$
  - Cool to  $\sim 50\text{--}60^{\circ}\text{C}$ , primers anneal to target
  - Polymerase extends primers in  $5' \rightarrow 3'$  direction
  - After a round of elongation is done, repeat steps

# General Steps of PCR

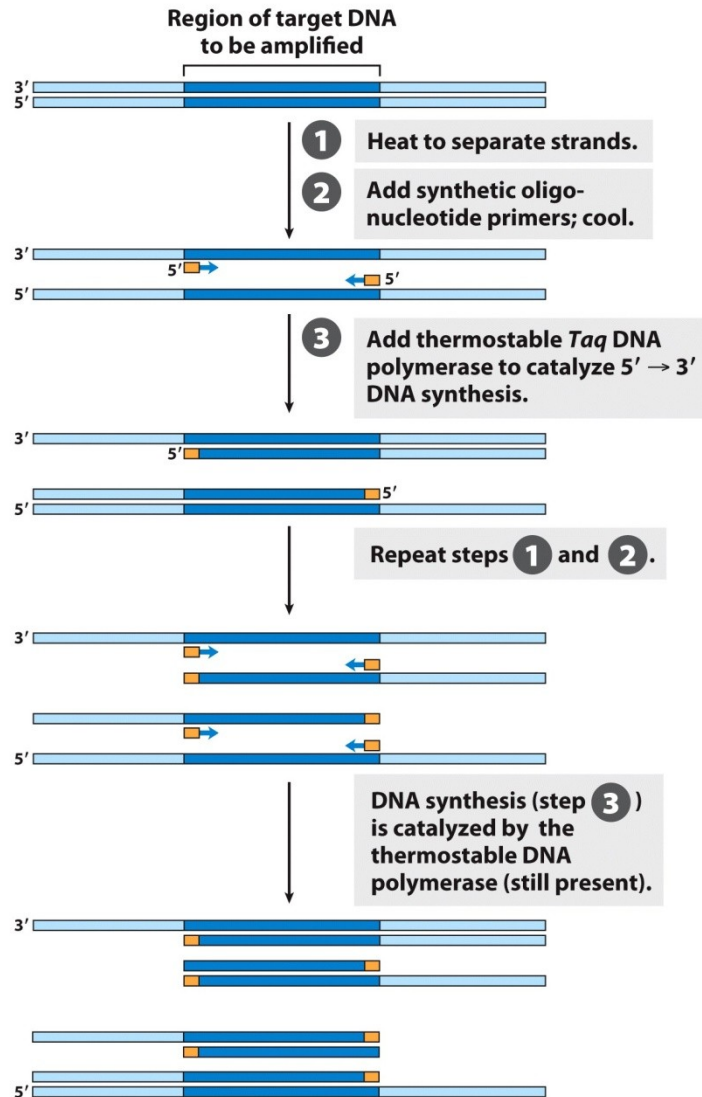
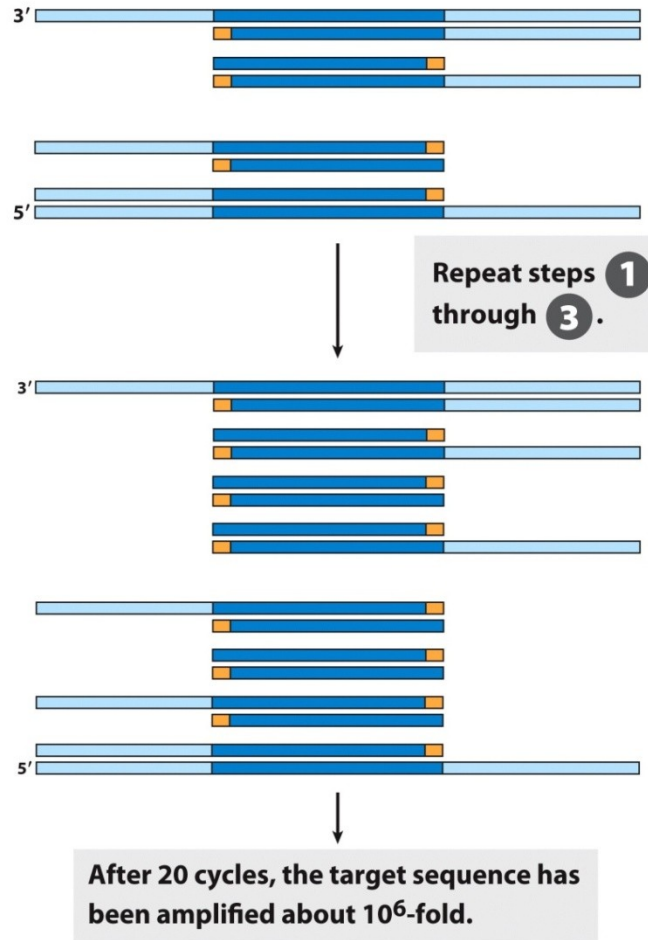


Figure 9-12a part 1

Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

# General Steps of PCR

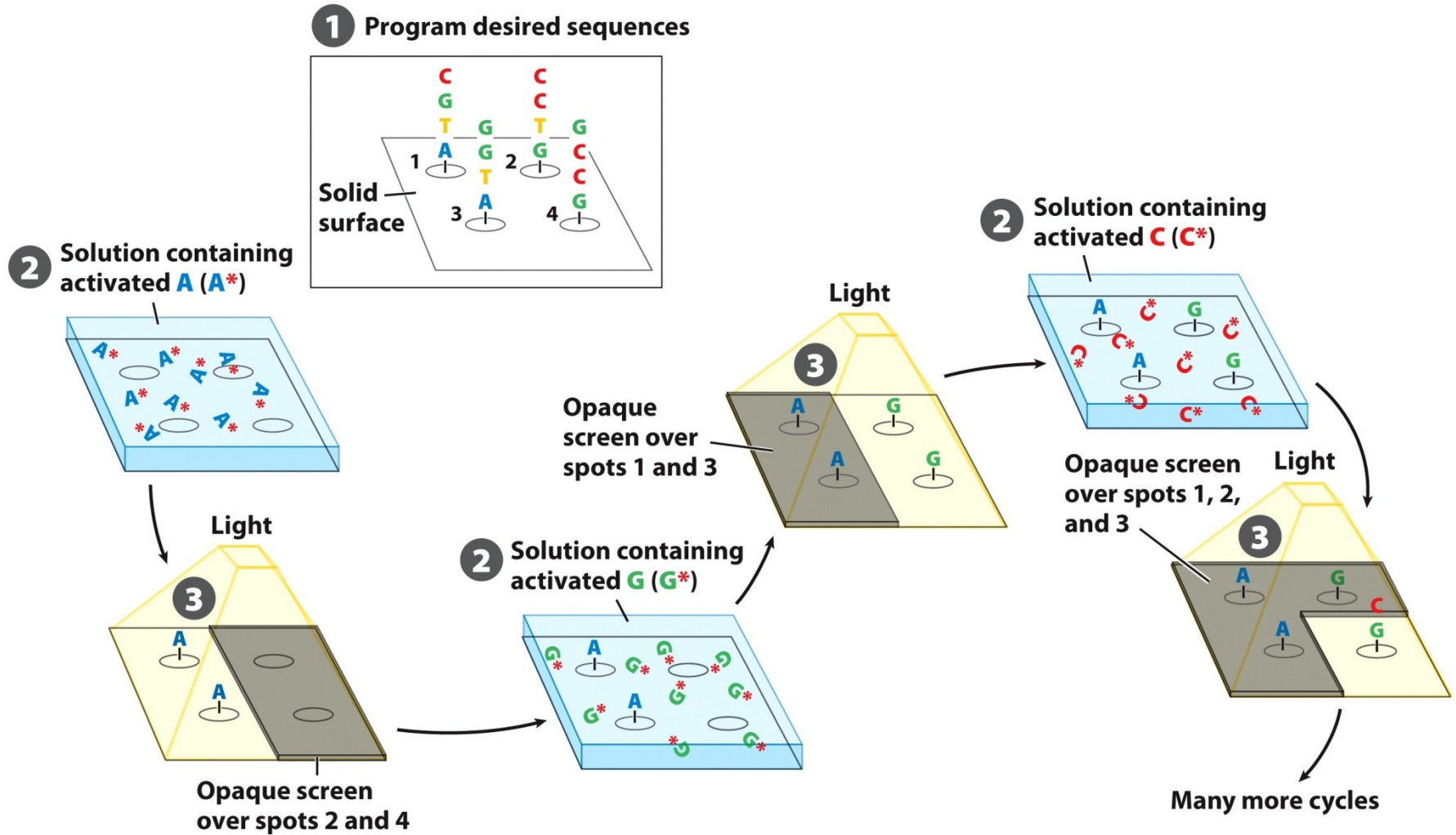
- Repeat steps 1–3 many times:



**Figure 9-12a part 2**  
*Lehninger Principles of Biochemistry, Sixth Edition*  
© 2013 W. H. Freeman and Company



# Photolithographic Synthesis of DNA



**Figure 9-22**  
*Lehninger Principles of Biochemistry, Sixth Edition*  
© 2013 W. H. Freeman and Company

# DNA Microarrays: Applications

DNA microarrays allow simultaneous screening of many thousands of genes: high-throughput screening

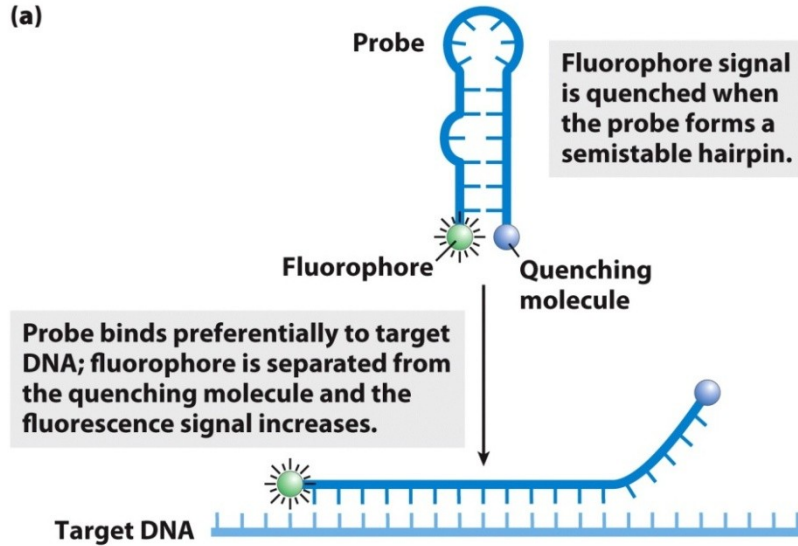
- Genome-wide genotyping
  - Which genes are present in this individual?
- Tissue-specific gene expression
  - Which genes are used to make proteins?
- Mutational analysis
  - Which genes have been mutated?

# Adaptations to PCR

- Reverse Transcriptase PCR (RT-PCR)
  - Used to amplify RNA sequences
  - First step uses reverse transcriptase to convert RNA to DNA
- Quantitative PCR (Q-PCR)
  - Used to show quantitative differences in gene levels

# qPCR

(a)

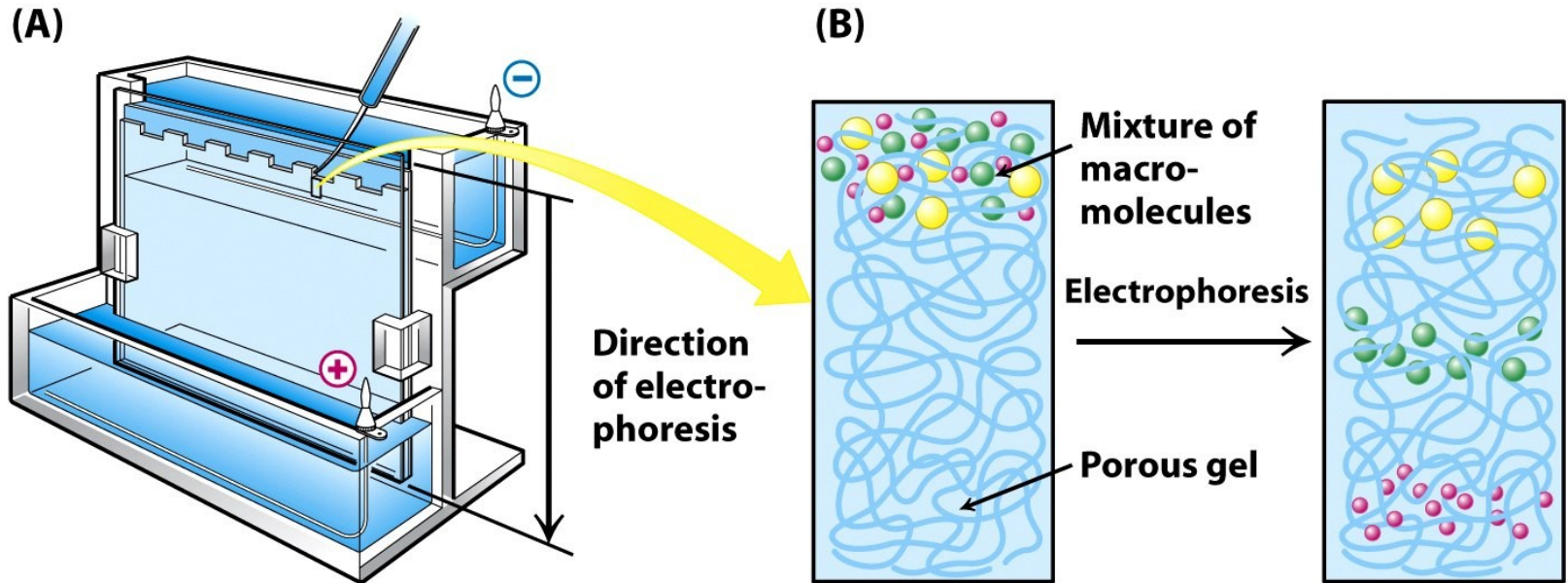


# Proteomes

- *Proteome*: set of all proteins produced under a given set of conditions
- Term can be applied to:
  - complete set of proteins for a given organism
  - specific subset of proteins present in a particular cell type or under specific growth conditions
- Proteome varies because it reflects genes that are actively expressed at any given time
- *Proteomics* analyses many samples using 2D-electrophoresis and mass spectrometry
- High-throughput, but less than transcriptomics

# Gel Electrophoresis

- Electrophoresis separates molecules by size
- Resolution is limited



**Figure 3-7**  
*Biochemistry, Sixth Edition*  
© 2007 W.H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 71

# Isoelectric Focusing

- Electrophoresis across a pH gradient
- Proteins migrate to their isoelectric pH

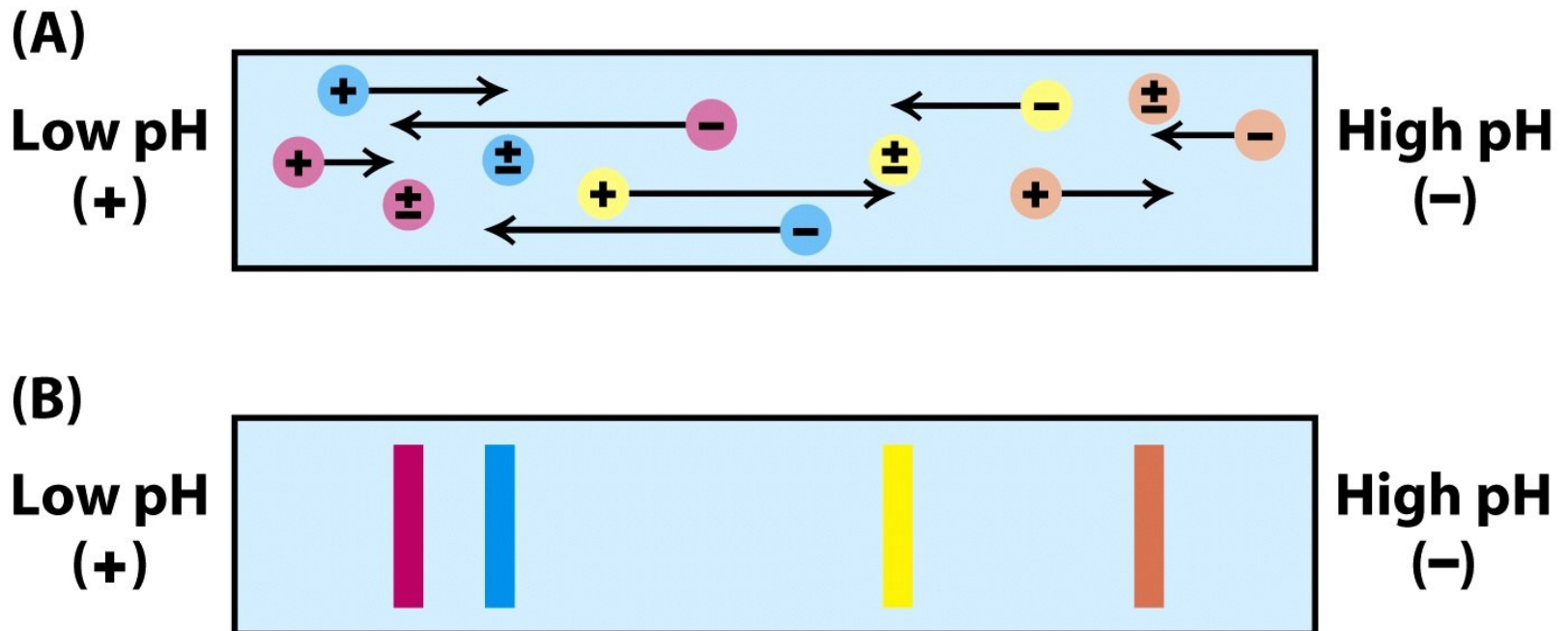


Figure 3-11  
*Biochemistry, Sixth Edition*  
© 2007 W.H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 73

# Two-dimensional Gel Electrophoresis

- Protein sample initially fractionated in one dimension by isoelectric focusing
- SDS-PAGE performed perpendicular to original direction
- Separates proteins according to pI and mass

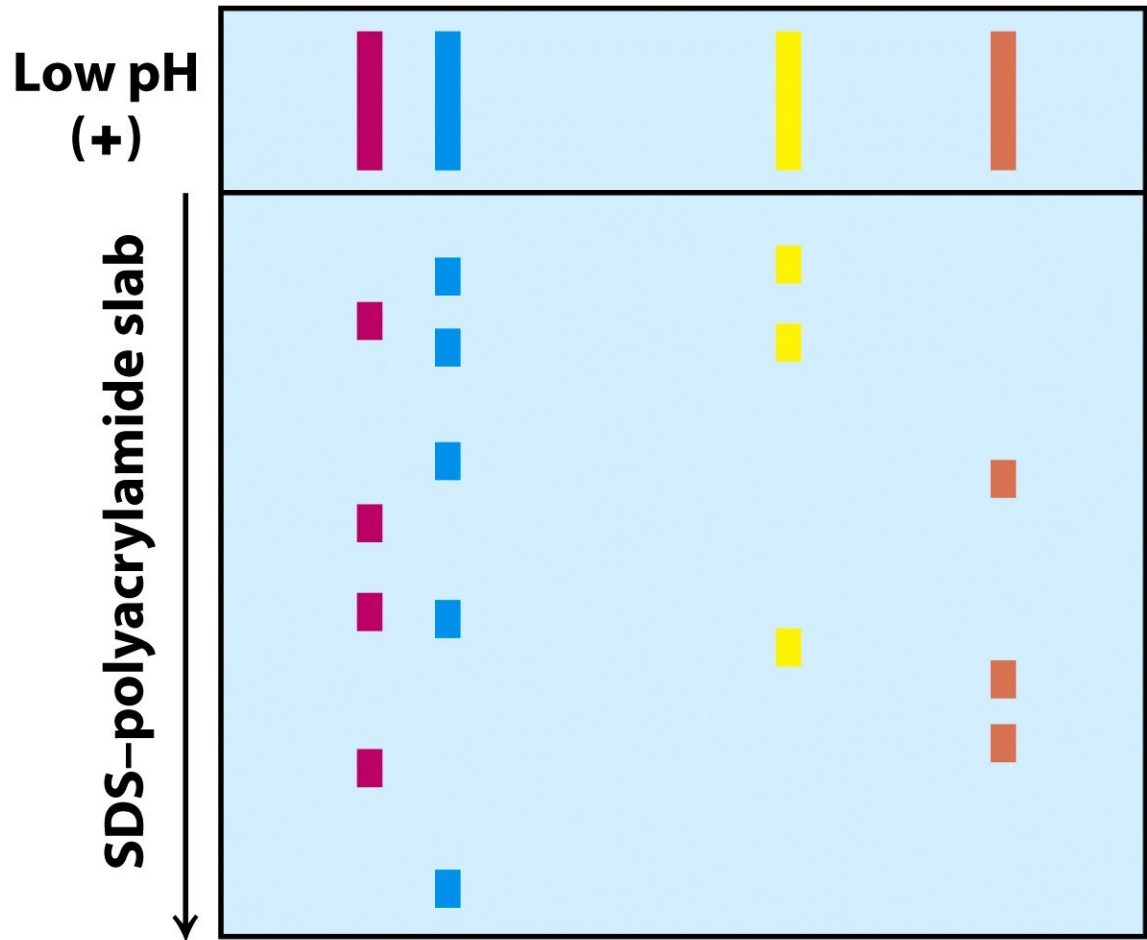


Figure 3-12a  
*Biochemistry, Sixth Edition*  
© 2007 W.H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 74



# Two-dimensional Gel Electrophoresis

## Isoelectric focusing

SDS-PAGE



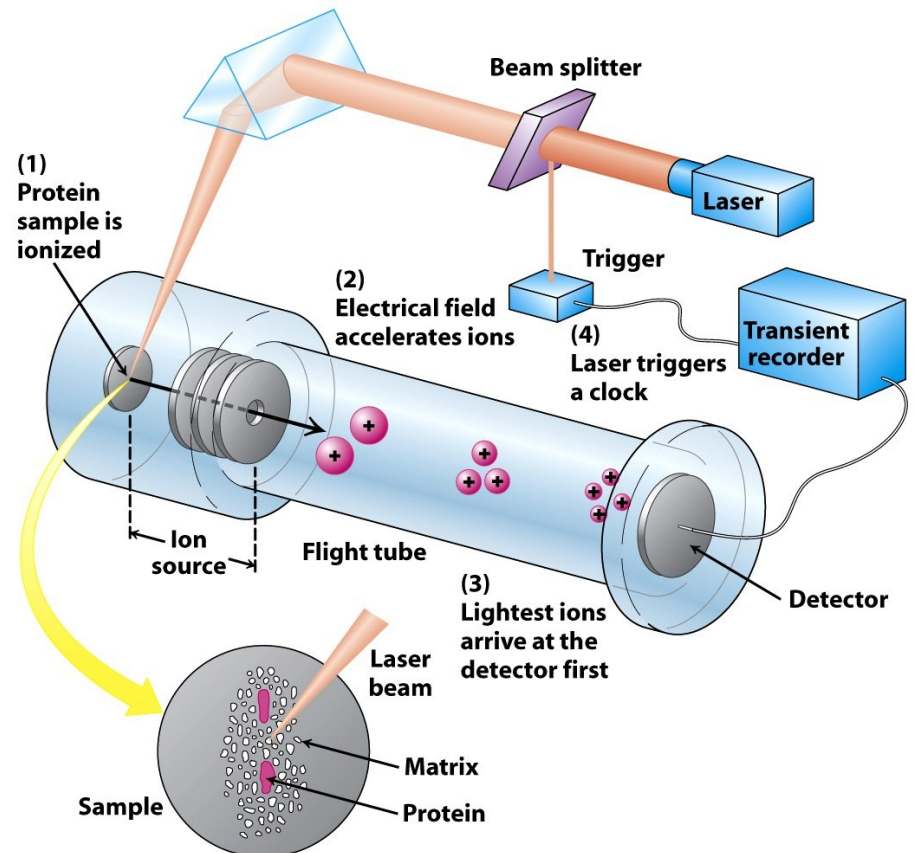
- Proteins from *E. coli* separated by 2D-electrophoresis
- >1,000 proteins can be resolved

Figure 3-12b  
*Biochemistry, Sixth Edition*  
© 2007 W.H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 74

# Mass Spectrometry

- MALDI-TOF mass spectrometry
- Protein sample is ionized and exposed to electrical field
- Ions travel according to size



Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 94

Figure 3-40  
*Biochemistry, Sixth Edition*  
© 2007 W. H. Freeman and Company

# MALDI-TOF Mass Spectrum

- MALDI-TOF gives good estimates of molecular weights
- Can be used to identify a few proteins within a mixture

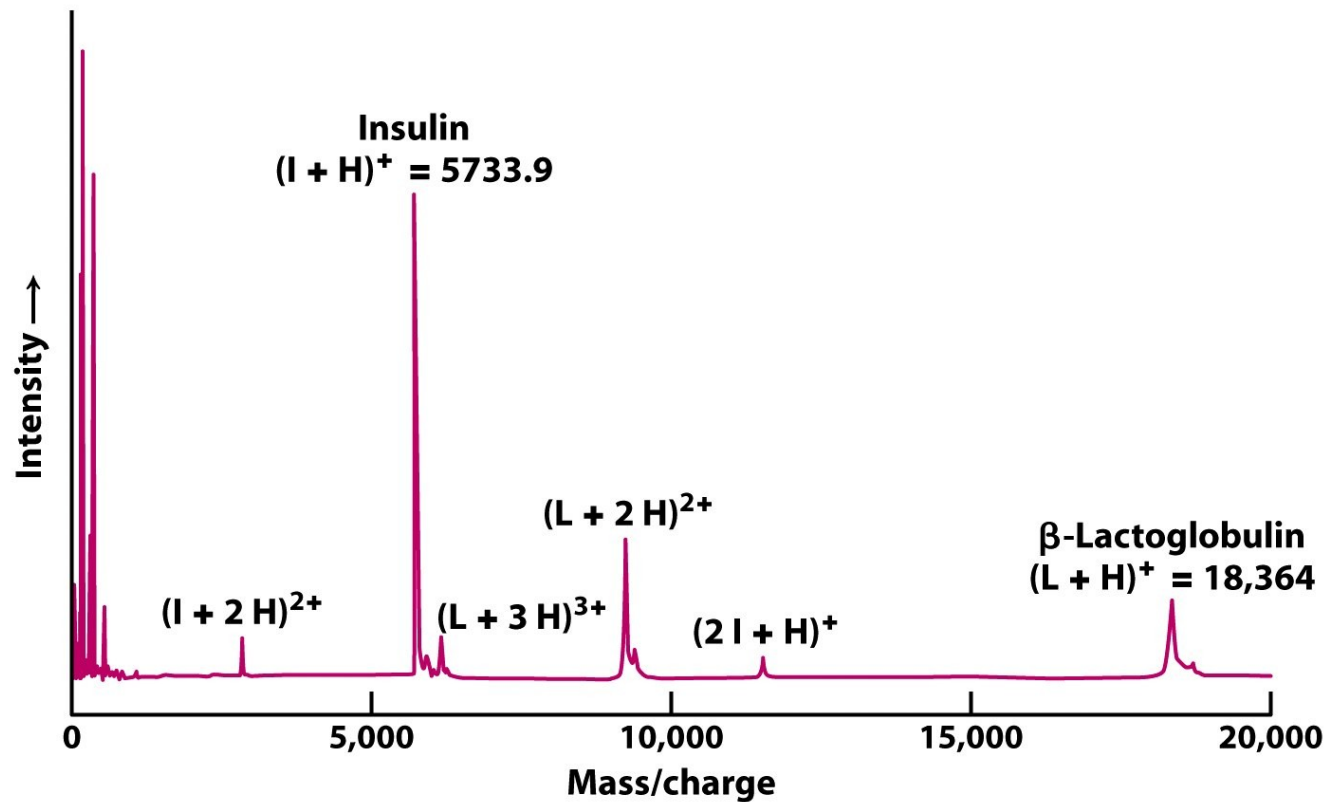


Figure 3-41  
*Biochemistry, Sixth Edition*  
© 2007 W.H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 94

# Proteomic Analysis by Mass Spectrometry

- Proteins separated by 2D electrophoresis
- Single proteins eluted
- Digestion with trypsin will give fragments with unique set of sizes
- Sizes identified by mass spectrometry and matched to database
- Allows identification of unknown proteins

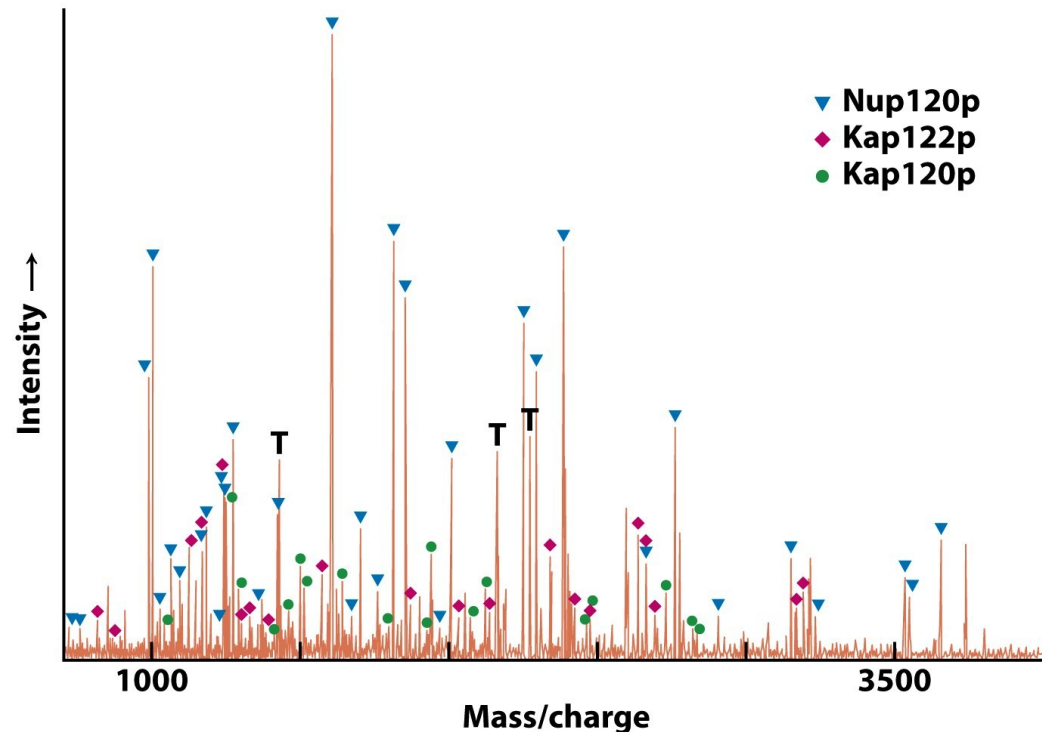


Figure 3-42  
*Biochemistry, Sixth Edition*  
© 2007 W. H. Freeman and Company

Berg, Tymoczko & Stryer, "Biochemistry",  
6<sup>th</sup> edn, 2006, p. 95

# Transcriptomics v Proteomics

- Transcriptomics and proteomics are both very powerful
- Differences in their practical application:
  - Transcriptomics is robust, relatively cost-effective and user-friendly
  - Proteomics still relatively limited – problems can remain with purification and stability of proteins
- Increasing potential to combine and compare data sets - for discussion see Hegde *et al.*, *Curr. Opin. Biotech.*, **14**, 647-651 (2003)

# **Bioinformatics: Mining the Data**

# Bioinformatics & Databases

- Latest biological data is gathered, organised and disseminated through large databases
- Databases include:
  - EBI, NCBI, Pfam, SMART, SWISS-PROT, TAIR
- Information in bioinformatic databases:
  - sequences, structures, homology searches
- Fast search engines allow searches by all with internet access – databases are as useful as the results they help generate!
- Improved tools for analysis of sequences

# Databases – Some URLs

Resource	URL
European Bioinformatics Institute	<a href="http://www.ebi.ac.uk/">www.ebi.ac.uk/</a>
<a href="http://www.ncbi.nlm.nih.gov/Genbank/">GenBank</a>	<a href="http://www.ncbi.nlm.nih.gov/Genbank/">www.ncbi.nlm.nih.gov/Genbank/</a>
NCBI	<a href="http://www.ncbi.nlm.nih.gov/">www.ncbi.nlm.nih.gov/</a>
<a href="http://www.rcsb.org/pdb/home/home.do">Protein DataBank</a>	<a href="http://www.rcsb.org/pdb/home/home.do">http://www.rcsb.org/pdb/home/home.do</a>
Sanger Centre	<a href="http://www.sanger.ac.uk/">www.sanger.ac.uk/</a>
<a href="http://smart.embl-heidelberg.de">SMART</a>	<a href="http://smart.embl-heidelberg.de">smart.embl-heidelberg.de</a>
The Arabidopsis Information Resource (TAIR)	<a href="http://www.arabidopsis.org/">www.arabidopsis.org/</a>



# NCBI: Complete Genomes

The screenshot shows the NCBI Genomes & Maps website. The browser window title is "Genomes & Maps - Site Guide - NCBI - Microsoft Internet Explorer provided by University of East Anglia". The address bar shows "http://www.ncbi.nlm.nih.gov/guide/genomes-maps/". The NCBI logo and "National Center for Biotechnology Information" are visible. A search bar is present with a dropdown menu set to "All Databases". The main content area is titled "Genomes & Maps" and has a navigation bar with tabs for "All", "Databases", "Downloads", "Submissions", "Tools", and "How To". The "Databases" tab is selected, showing a list of databases with descriptions:

- BioProject (formerly Genome Project)**: A collection of genomics, functional genomics, and genetics studies and links to their resulting datasets. This resource describes project scope, material, and objectives and provides a mechanism to retrieve datasets that are often difficult to find due to inconsistent annotation, multiple independent submissions, and the varied nature of diverse data types which are often stored in different databases.
- CloneDB (formerly Clone Registry)**: A database that integrates information about clones and libraries, including sequence data, map positions and distributor information.
- Database of Genome Survey Sequences (dbGSS)**: A division of GenBank that contains short single-pass reads of genomic DNA. dbGSS can be searched directly through the Nucleotide GSS Database.
- Database of Genomic Structural Variation (dbVar)**: The dbVar database has been developed to archive information associated with large scale genomic variation, including large insertions, deletions, translocations and inversions. In addition to archiving variation discovery, dbVar also stores associations of defined variants with phenotype information.
- Epigenomics**: This resource enables users to explore and visualize richly-annotated epigenomics datasets. It provides a unique interface to search and navigate epigenomic data in the context of biological sample information, as well as tools to select, download and view multiple sets of epigenomic data as tracks on genome browsers.

On the left side, there is a "Resource List (A-Z)" menu with "Genomes & Maps" highlighted. On the right side, there is a "Quick Links" section with links to BioProject, dbVar, Genome, Nucleotide Database, PopSet, SRA, Trace Archive, UniSTS, GenBank, ProtMap, Workbench, Map Viewer, ProSplign, and Splign. The browser status bar at the bottom shows "Internet | Protected Mode: On" and "125%".

# NCBI: Eukaryotic Genomes

Genome List - Microsoft Internet Explorer provided by University of East Anglia  
 http://www.ncbi.nlm.nih.gov/genome/browse/

coerulea																	
Aquilegia formosa	PRJNA18745	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Aquilegia formosa	PRJNA168267	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis arenosa	PRJNA37965	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis halleri	PRJEB79	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis lyrata subsp. lyrata	PRJNA49545 PRJNA41137	Plants	Land Plants	183.71	36.10	v.1.0	-	-	-	ADBK01	695	32534	32549	2009/1			
Arabidopsis lyrata	PRJEB83	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana	PRJNA116 PRJNA10719	Plants	Land Plants	119.67	36.05	TAIR10	5	2	-	-	-	33583	35378	1997/0			
Arabidopsis thaliana	PRJNA13190	Plants	Land Plants	93.65	36.04	ASM21127v1	6	-	-	-	-	16842	20111	2000/1			
Arabidopsis thaliana	PRJNA117657	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana	PRJNA169548	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana	PRJNA13192	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana	PRJNA67133	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arachis hypogaea	PRJNA48331	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arbutus unedo	PRJNA81259	Plants	Land Plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arctica islandica	PRJEA89475	Animals	Other Animals	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arothron mappa	PRJNA74295	Animals	Fishes	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Internet | Protected Mode: On 125%

# NCBI: Eukaryotic Genomes

Genome List - Microsoft Internet Explorer provided by University of East Anglia

http://www.ncbi.nlm.nih.gov/genome/browse/

Genome List

<i>Mossiiensis</i> OWVT-1													
<i>Hirsutella</i> <i>vermicola</i> AS3.7877	PRJNA157045	Fungi	Ascomycetes	-	-	-	-	-	-	-	-	-	-
<i>Homalodisca</i> <i>vitripennis</i>	PRJNA162481	Animals	Insects	-	-	-	-	-	-	-	-	-	-
Homo sapiens	PRJNA168 PRJNA31257	Animals	Mammals	3095.69	41.58	GRCh37.p10	24	1	-	-	441	37150	3
Homo sapiens	PRJNA1431	Animals	Mammals	2852.22	41.17	Hs_Celera_WGSA	24	-	-	AADD01	8832	0	
Homo sapiens	PRJNA16133 PRJNA10793	Animals	Mammals	158.33	40.90	CRA_TCAGchr7v2	1	-	-	-	6	1956	
Homo sapiens	PRJNA20837 PRJNA19621	Animals	Mammals	2844	40.93	HuRef HuRefPrime	24	-	-	ABBA01	187643	35738	3
Homo sapiens	PRJNA176729	Animals	Mammals	2831.81	40.90	CHM1_1.0	-	-	-	AMYH01	2178	-	
Homo sapiens	PRJNA28335	Animals	Mammals	41.67	40.80	Watson-partial	-	-	-	ABKV01	-	-	
Homo sapiens	PRJNA42199	Animals	Mammals	2113.77	39.60	YH1	-	-	-	ADDF01	191424	-	
Homo sapiens	PRJNA42201	Animals	Mammals	2083.78	39.30	BGI AF	-	-	-	DAAB01	314786	-	
Homo sapiens	PRJNA59877	Animals	Mammals	2615.02	40.70	HsapALLPATHS1	-	-	-	AEKP01	11393	-	
Homo sapiens	PRJNA28911	Animals	Mammals	-	-	-	-	-	-	-	-	-	
Homo sapiens	PRJNA28919	Animals	Mammals	-	-	-	-	-	-	-	-	-	
Homo sapiens	PRJNA28957	Animals	Mammals	-	-	-	-	-	-	-	-	-	
Homo sapiens	PRJNA29429	Animals	Mammals	-	-	-	-	-	-	-	-	-	

Internet | Protected Mode: On

125%

# NCBI: Microbial Genomes

Genome List - Microsoft Internet Explorer provided by University of East Anglia  
 http://www.ncbi.nlm.nih.gov/genome/browse/

Genome

**Genome Information by organism**

[Download Reports from FTP site](#)

Overview [8493] Eukaryotes [2190] Prokaryotes [14592] Viruses [3267]

First Previous **Shown: 1 - 100 out of 14592 items** Next Last [Download selected records](#)

Organism/Name	BioProject	Group	SubGroup	Size (Mb)	GC%	Chromosomes		Plasmids		WC
						RefSeq	INSDC	RefSeq	INSDC	
Abiotrophia defectiva ATCC 49176	PRJNA55729 PRJNA33011	Firmicutes	Bacilli	3.48	37.10	-	-	-	-	ACII
Acaricomes phytoseiuli DSM 14247	PRJNA174970	Actinobacteria	Actinobacteria	-	-	-	-	-	-	-
Acaryochloris sp. CCME 5410	PRJNA78283 PRJNA16707	Cyanobacteria	Chroococcales	7.88	-	-	-	-	-	AFE
Acaryochloris marina MBIC11017	PRJNA58167 PRJNA12997	Cyanobacteria	Chroococcales	8.36	46.99	NC_009925.1	CP000828.1	NC_009926.1 NC_009929.1 NC_009930.1 NC_009928.1 NC_009927.1 NC_009932.1 NC_009934.1 NC_009931.1 NC_009933.1	CP000838.1 CP000841.1 CP000842.1 CP000840.1 CP000839.1 CP000844.1 CP000846.1 CP000843.1 CP000845.1	-
Acetivibrio cellulolyticus CD2	PRJNA51533 PRJNA45843	Firmicutes	Clostridia	6.15	-	-	-	-	-	AED

Done Internet | Protected Mode: On 125%

# NCBI: Microbial Genomes

Escherichia coli (ID 167) - Genome - NCBI - Microsoft Internet Explorer provided by University of East Anglia

http://www.ncbi.nlm.nih.gov/genome/167

NCBI Resources How To Sign in to NCBI

Genome Genome Search

Limits Advanced Help

Display Settings: Overview Send to:

Organism Overview; Genome Project Report; Genome Annotation Report; Plasmid Annotation Report

**Escherichia coli**  
A well-studied enteric bacterium

Lineage: Bacteria[3413]; Proteobacteria[1423]; Gammaproteobacteria[621]; Enterobacteriales[146]; Enterobacteriaceae[146]; Escherichia[7]; Escherichia coli[1]

**Escherichia coli.** This organism was named for its discoverer, Theodore Escherich, and is one of the premier model organisms used in the study of bacterial genetics, physiology, and biochemistry. This enteric organism is typically present in the lower intestine of humans, where it is the dominant facultative anaerobe present, but it is [More...](#)

Representatives

- Reference genome: Escherichia coli str. K-12 substr. MG1655
- Reference genome: Escherichia coli O157:H7 str. Sakai
- Calculated: Escherichia coli SE11

Dendrogram (based on genomic BLAST)

Internet | Protected Mode: On 125%

# Databases - The Caveats

- Databases contain mistakes (low as a proportion of total data)
  - primary data errors
  - data analysis errors
  - annotation errors
- Errors are difficult to correct
- Make the interpretation of data your own responsibility!!

# NCBI – Useful Links

National Center for Biotechnology Information - Microsoft Internet Explorer provided by University of East Anglia

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

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases Search

Home  
Resource List (A-Z)  
Popular Resources

**Welcome to NCBI**

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [Research](#) | [RSS Feeds](#)

**Popular Resources**

- PubMed
- Bookshelf
- PubMed Central
- PubMed Health
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

**NCBI Announcements**

New version of Genome Workbench available  
06 Sep 2012  
An integrated, downloadable application for viewing and analyzing sequence data. the

NCBI's July Newsletter is on the Bookshelf  
13 Aug 2012  
Introduction to the 1000 Genomes Browser, PubMed's Citation Manager and Favorites Button. new and modified BLAST services.

New Microbial BLAST Page  
28 Jun 2012  
Now easier to use and with the familiar

**Genetic Testing Registry**

A portal to clinical genetics resources with detailed information about genetic tests and laboratories. **GO**

1 2 3 4 5 6 7 8

Genomes & maps  
Homology  
Literature  
Proteins  
Sequence Analysis  
Taxonomy  
Training & Tutorials  
Variation

Links to brief description of all resources at NCBI

Internet | Protected Mode: On 125%

# NCBI – Useful Links

The image shows a screenshot of the NCBI website in a Microsoft Internet Explorer browser window. The browser's address bar shows the URL <http://www.ncbi.nlm.nih.gov/>. The page features a navigation menu with 'All Resources' circled in blue. A 'Welcome to NCBI' message is displayed, along with a 'Popular Resources' section where 'PubMed Central' and 'BLAST' are also circled in blue. Three callout boxes are overlaid on the page: one pointing to 'All Resources' with text about taxonomy, one pointing to 'PubMed Central' with text about domain structure, and one pointing to 'BLAST' with text describing it as a similarity search tool.

**Taxonomy: general information about taxonomy**

**Domain Structure: NCBI Structure Group, including tools to search and display structures genes and genetic disorders**

**BLAST: BLAST® (Basic Local Alignment Search Tool) is a set of similarity search programs**



# Databases Summary

- Many databases are available
  - some have lot of general information (NCBI, EBI)
  - some have specific data (Pfam, SWISS-PROT)
  - some relate to specific research interests (TAIR)
- Become well acquainted with specific databases
- Wide range of databases, web sites and other resources are available for *in silico* analysis of biological data
- Great advantages, but beware caveats and potential pitfalls – understand capabilities and limitations!
- Use information intelligently:
  - always ask if the conclusions make biological sense
  - may require further analyses or experimentation

# “Omics” Overview

- Analyses of various biological molecules have developed into their own scientific disciplines; e.g. *Metabolomics* = “analysis of metabolites”, etc.
- *Transcriptome*: set of all mRNAs (“transcripts”) produced from a genome
- *Proteome*: set of all proteins produced under a given set of conditions
- Both can vary because they reflect genes that are actively expressed at any given time
- Transcriptomics and proteomics are both powerful, but are used differently: transcriptomics is cheaper and more user friendly than proteomics