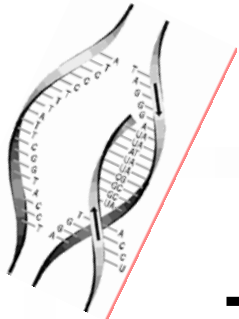


Transcriptomics

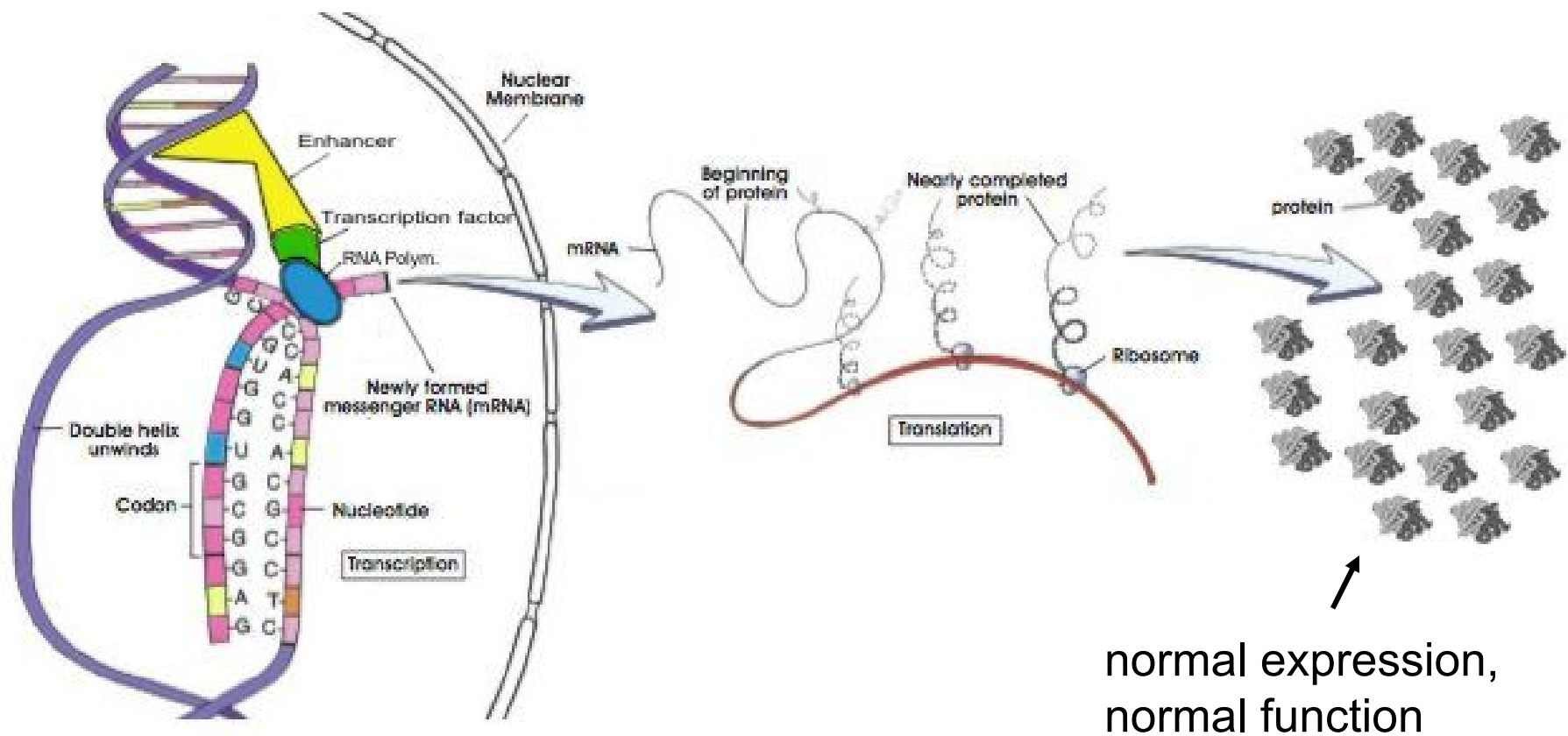
Transcripts – general considerations

- Differential regulation of the same genes has dramatic consequences (e.g. they account for most differences between chimp and man).
- The time axis is critical: how much is the gene expressed, but also **when** and **how long**?
- The analysis at transcript level (mRNA) is linked with the analysis at genetic level (DNA)
- The analysis at transcript level (mRNA) is linked with the analysis at protein level.

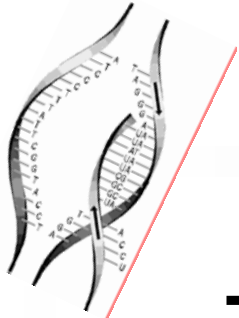
Transcriptomics



Transcripts – versus genes

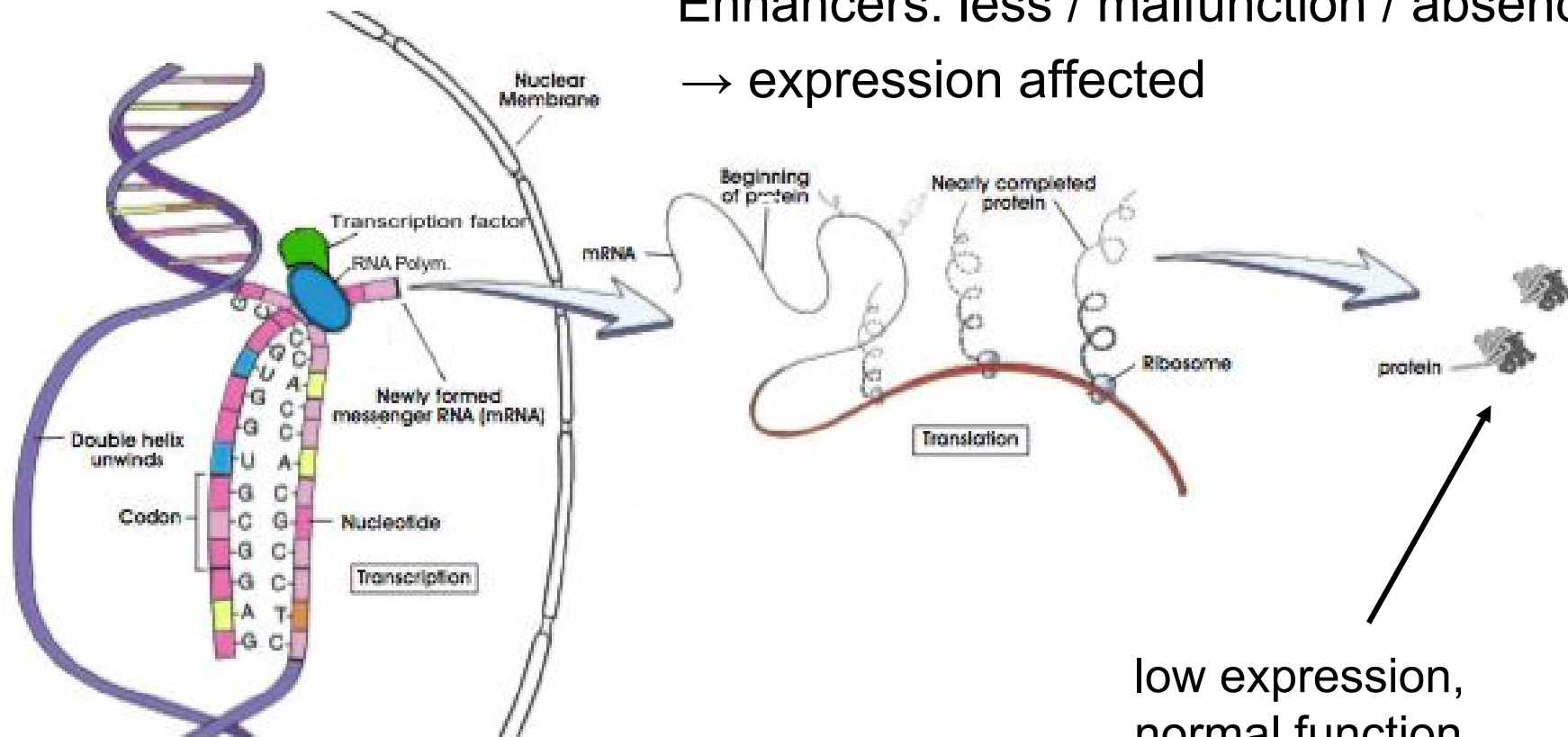


Transcriptomics

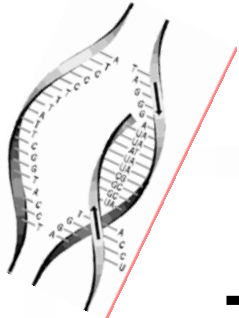


Transcripts – versus genes

Enhancers: less / malfunction / absence
→ expression affected



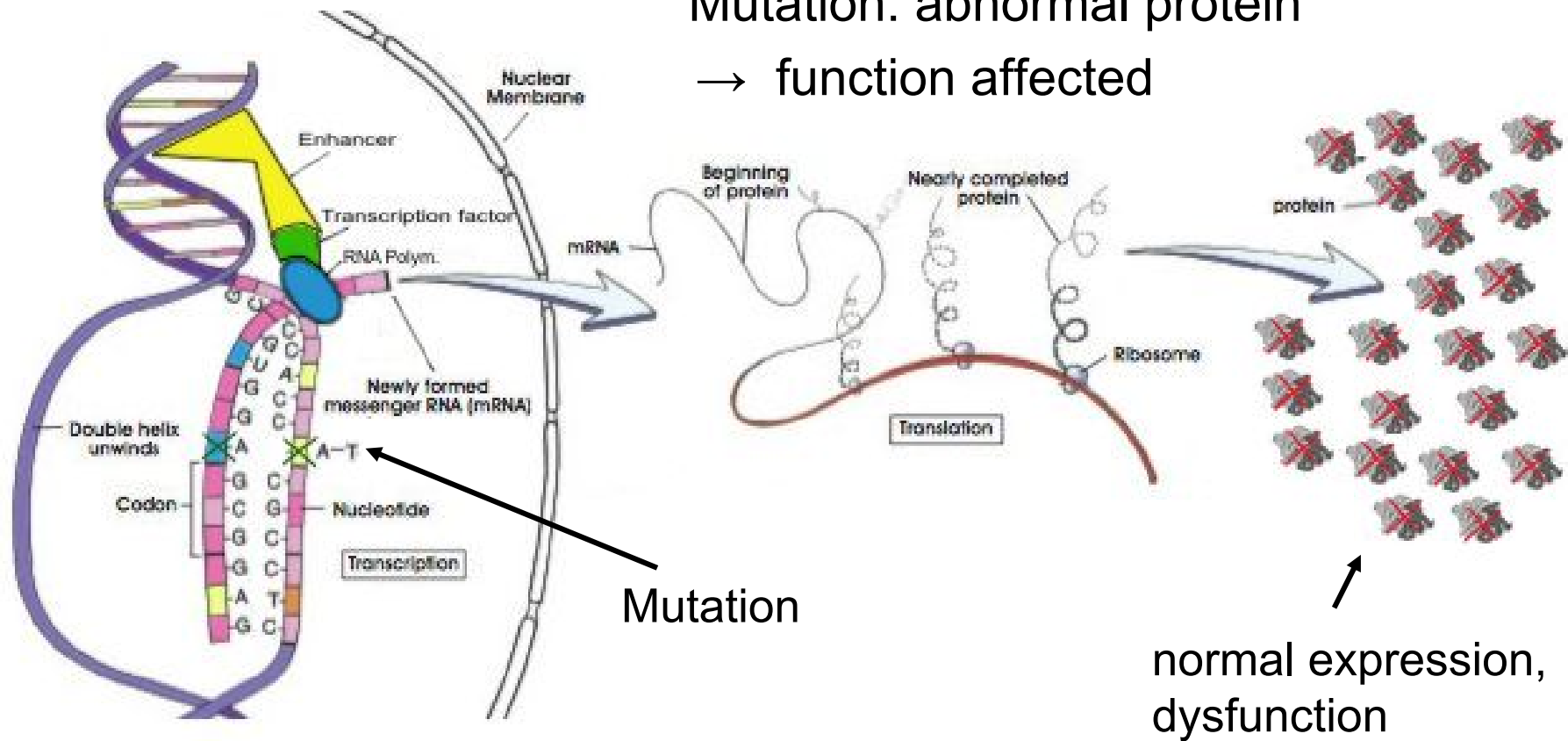
low expression,
normal function

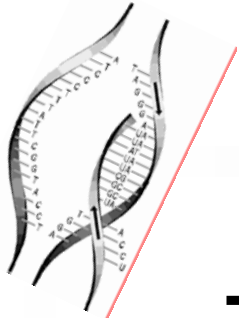


Transcriptomics

Transcripts – versus genes

Mutation: abnormal protein
→ function affected

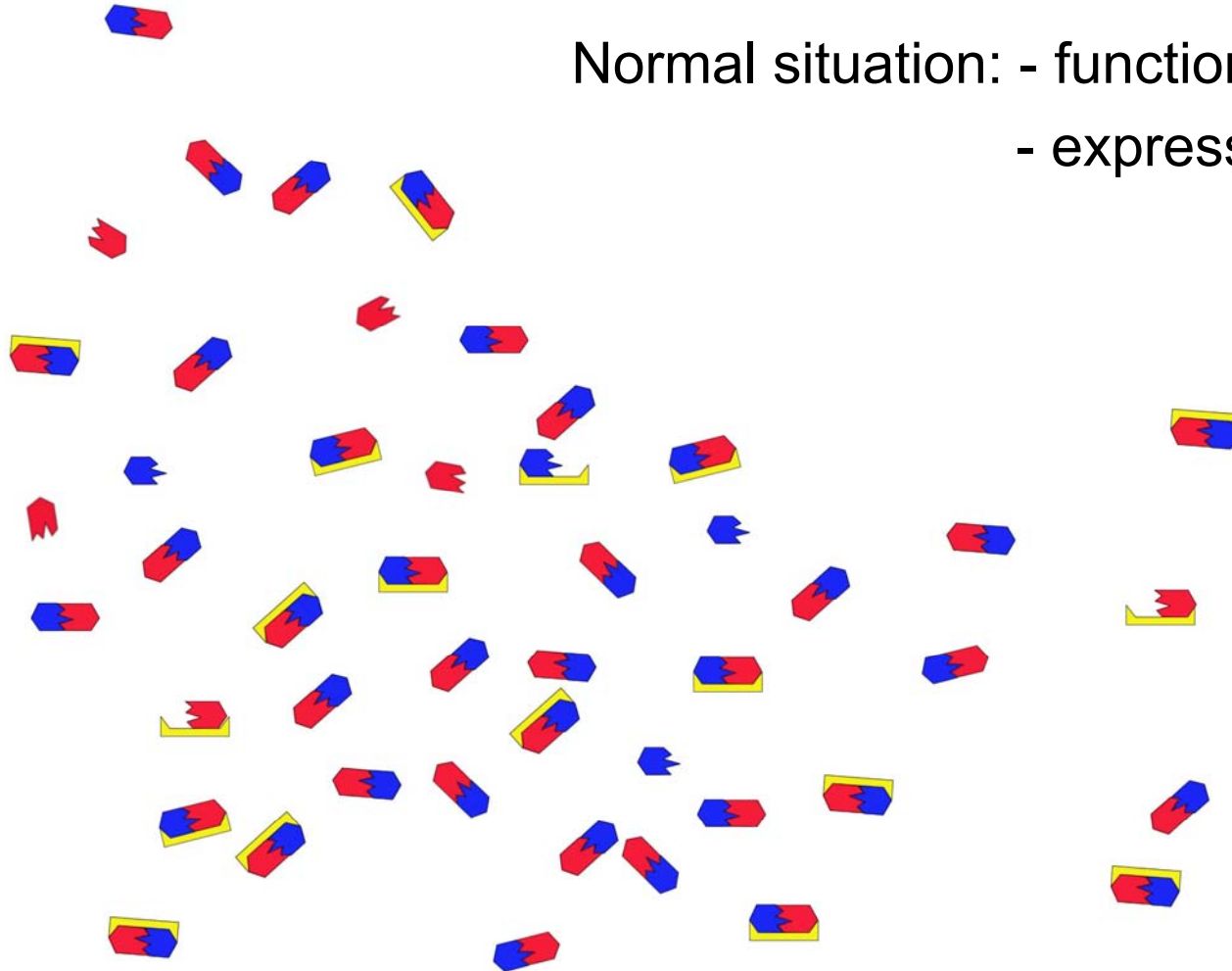


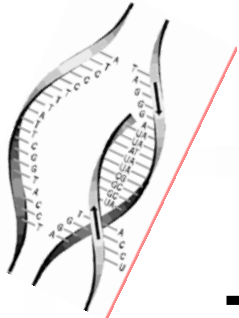


Transcriptomics

Transcripts – versus genes

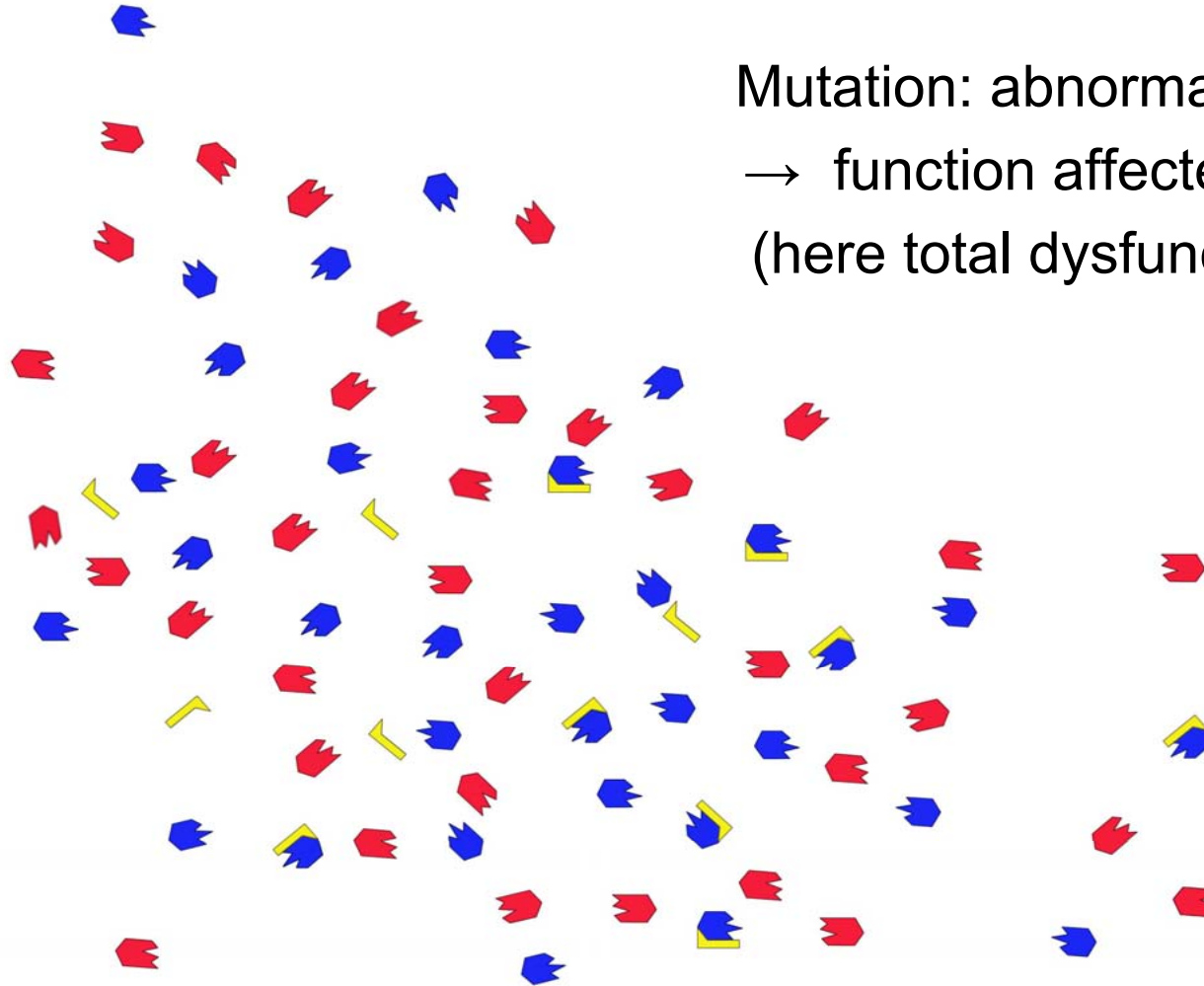
Normal situation: - function ok
- expression ok



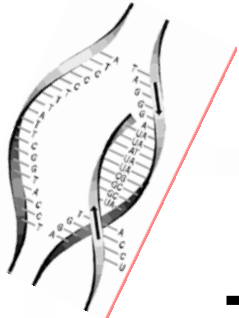


Transcriptomics

Transcripts – versus genes

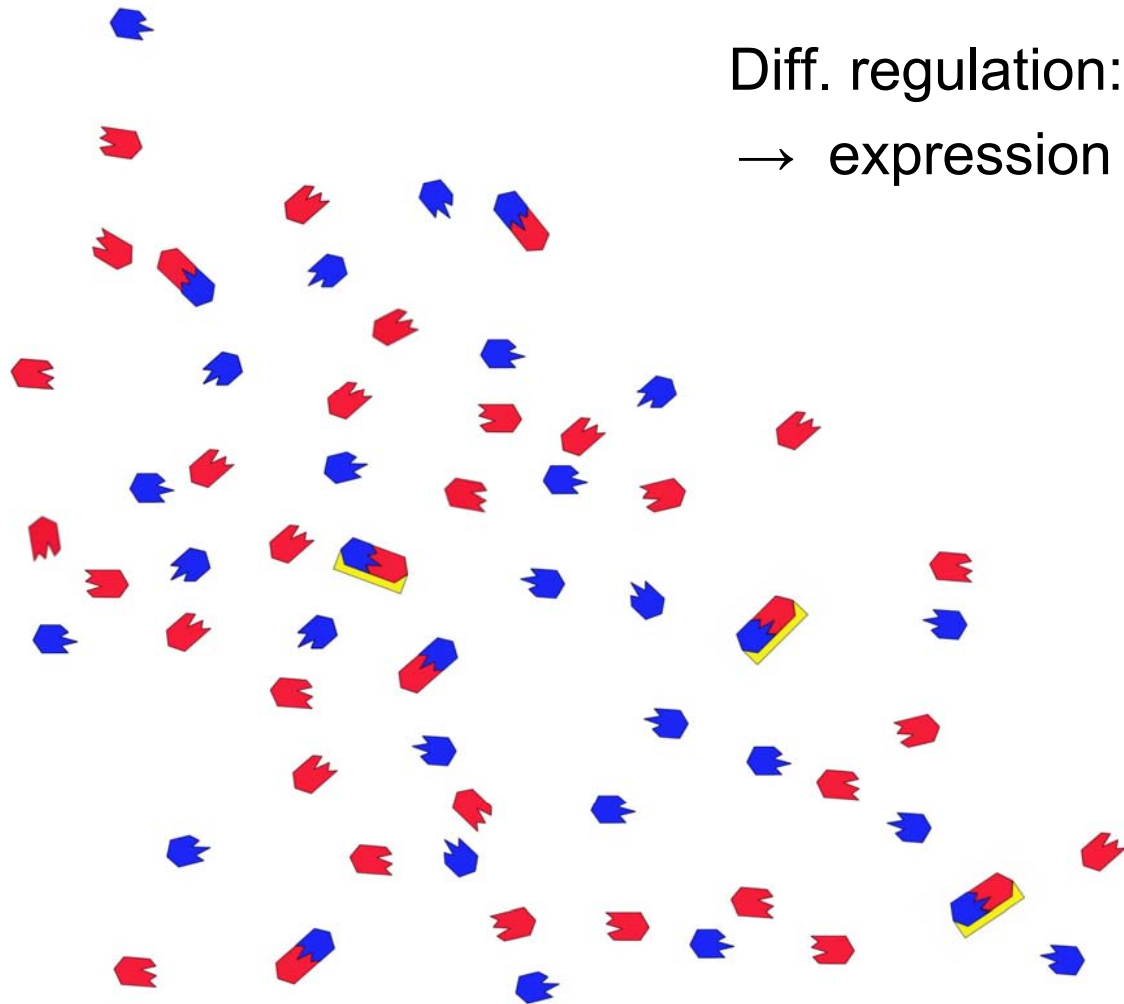


Mutation: abnormal protein
→ function affected
(here total dysfunction)



Transcriptomics

Transcripts – versus genes



Diff. regulation: normal protein
→ expression affected

if the expression is low enough, it can have the same effects as a dysfunctional mutation



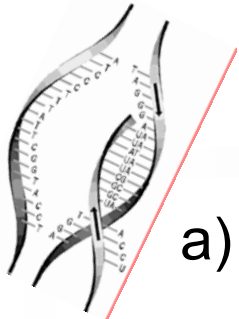
Transcriptomics

Transcripts – identification techniques

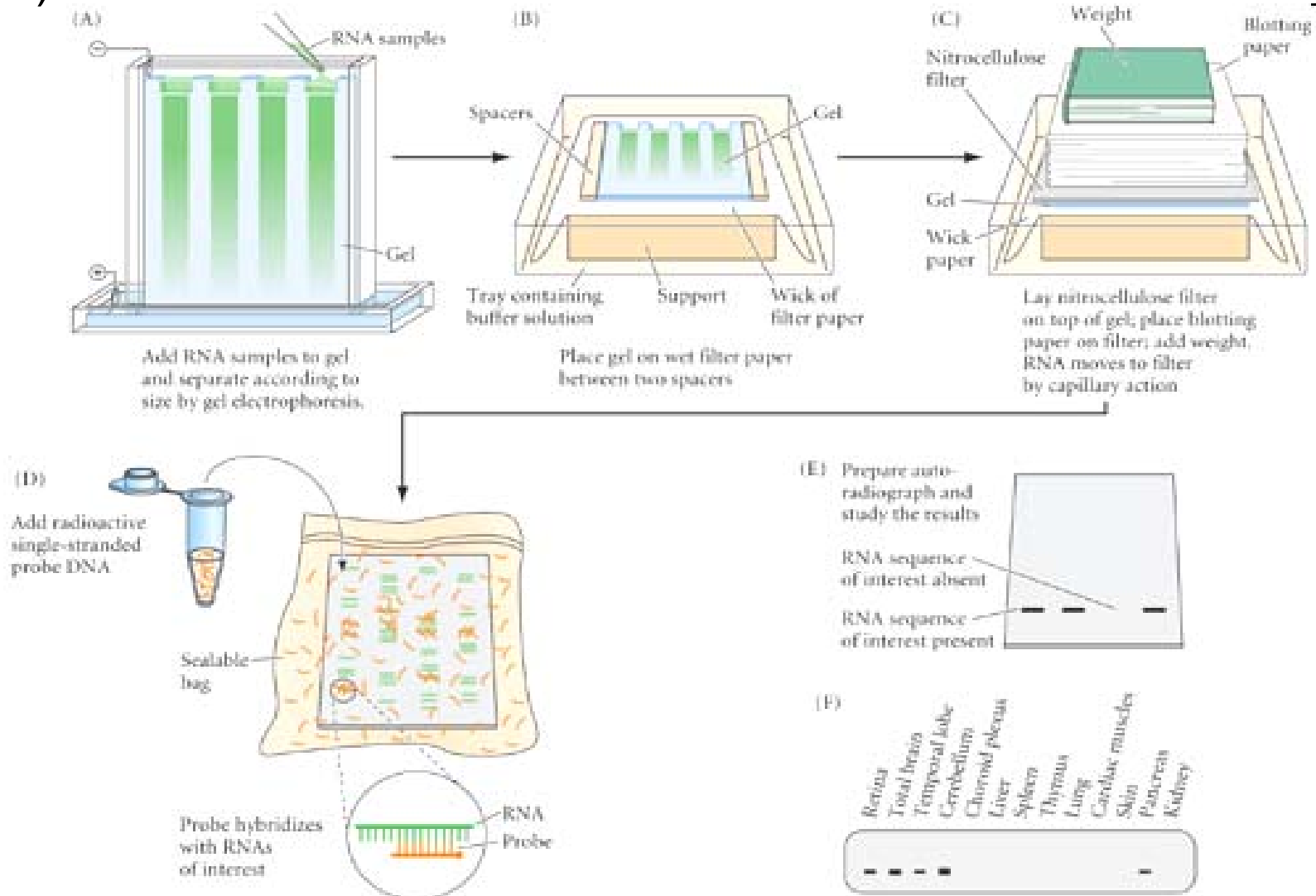
RNA

- a) Northern Blot (single RNAs)
- b) High-throughput (massive)

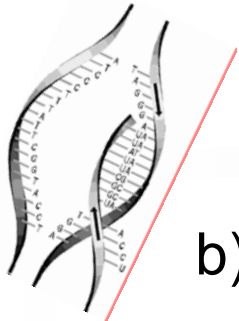
Transcriptomics



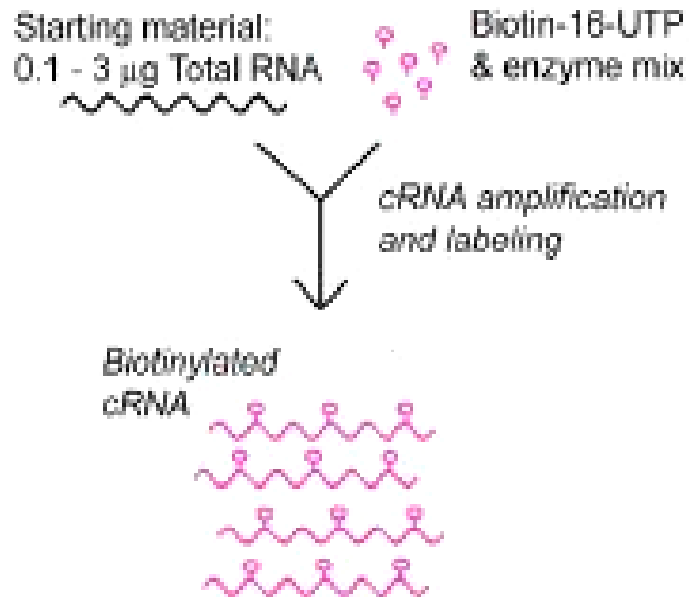
a) Northern Blot



Transcriptomics

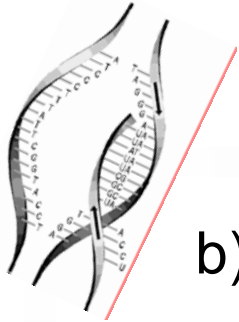


b) High-throughput

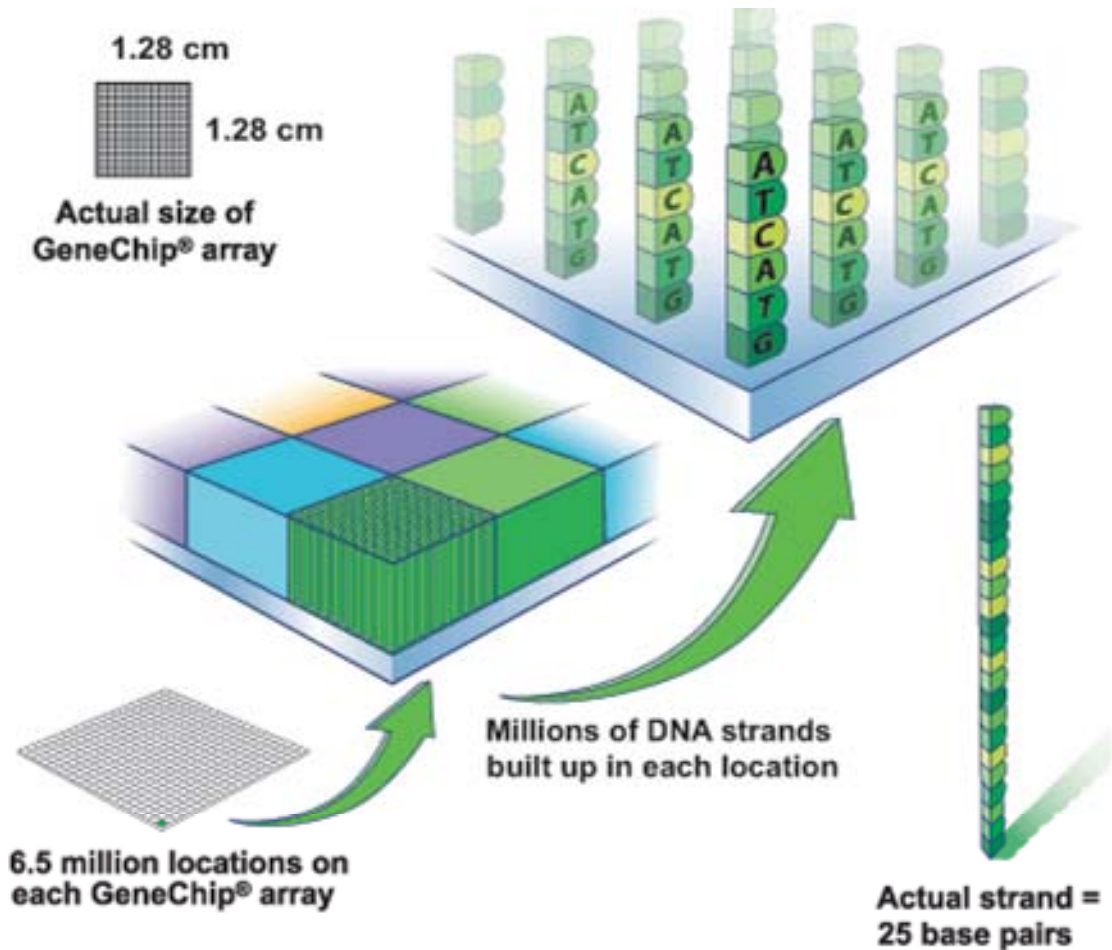


RNA gets to cRNA or cDNA,
is amplified, labelled (biotin) and fragmented

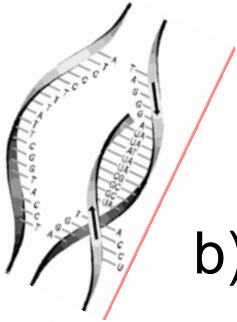
Transcriptomics



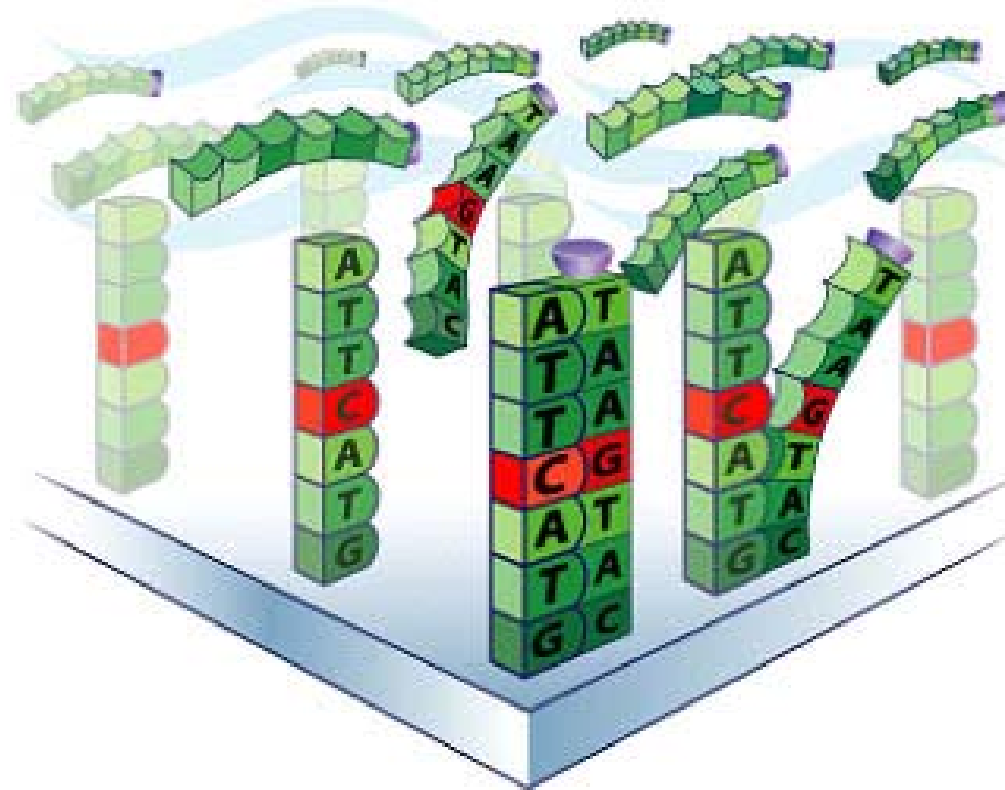
b) High-throughput



Transcriptomics

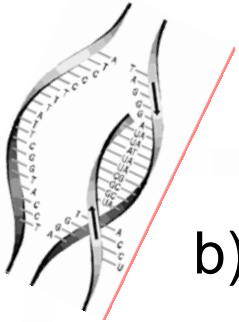


b) High-throughput

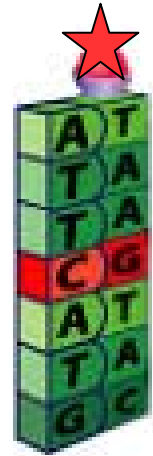


Hybridization

Transcriptomics

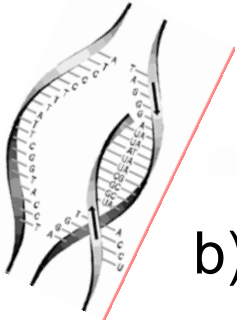


b) High-throughput



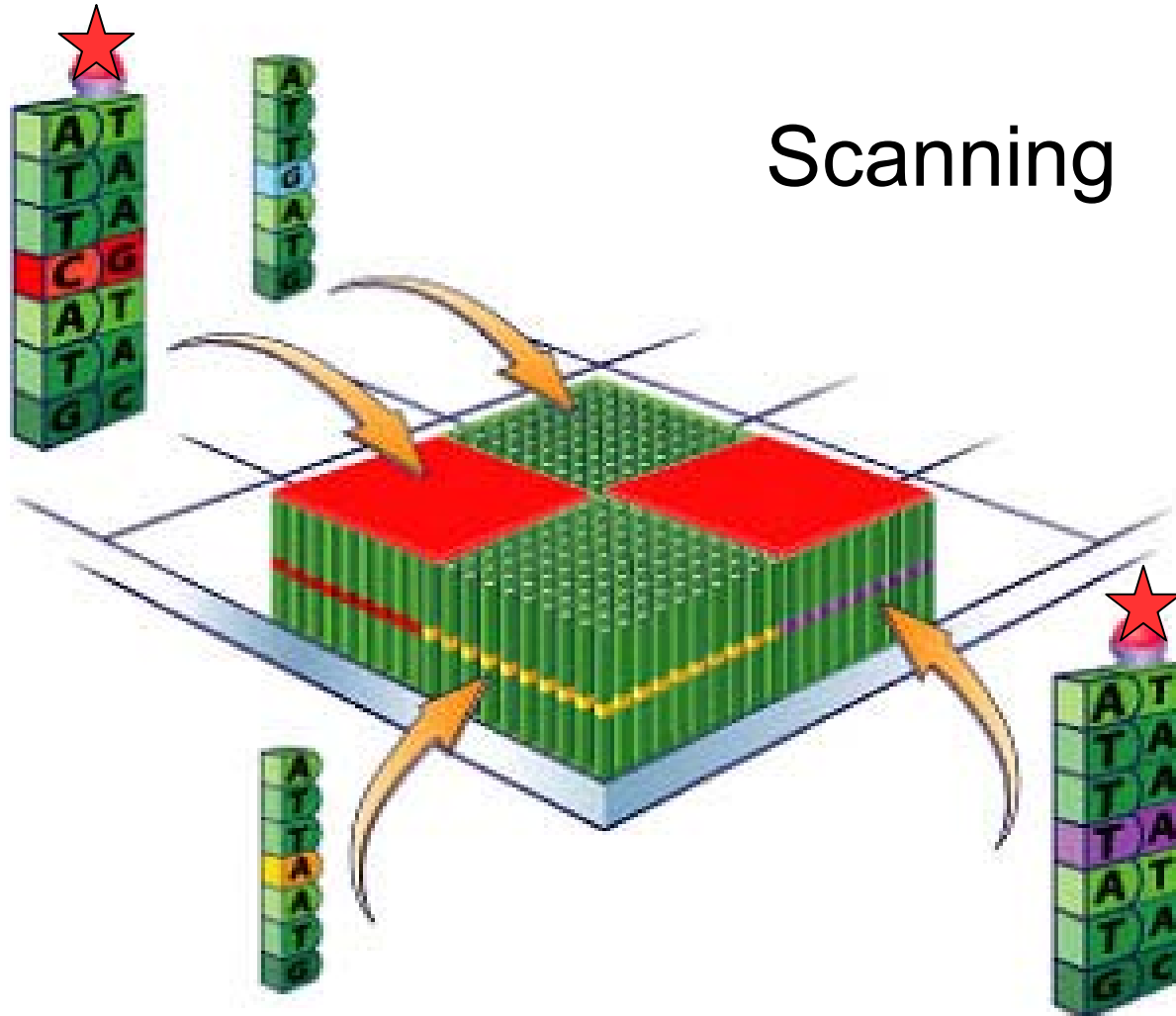
Staining (fluorescent)

Transcriptomics

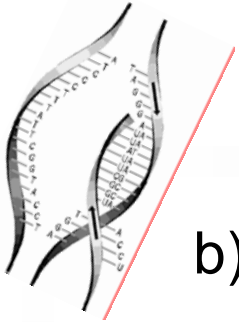


b)

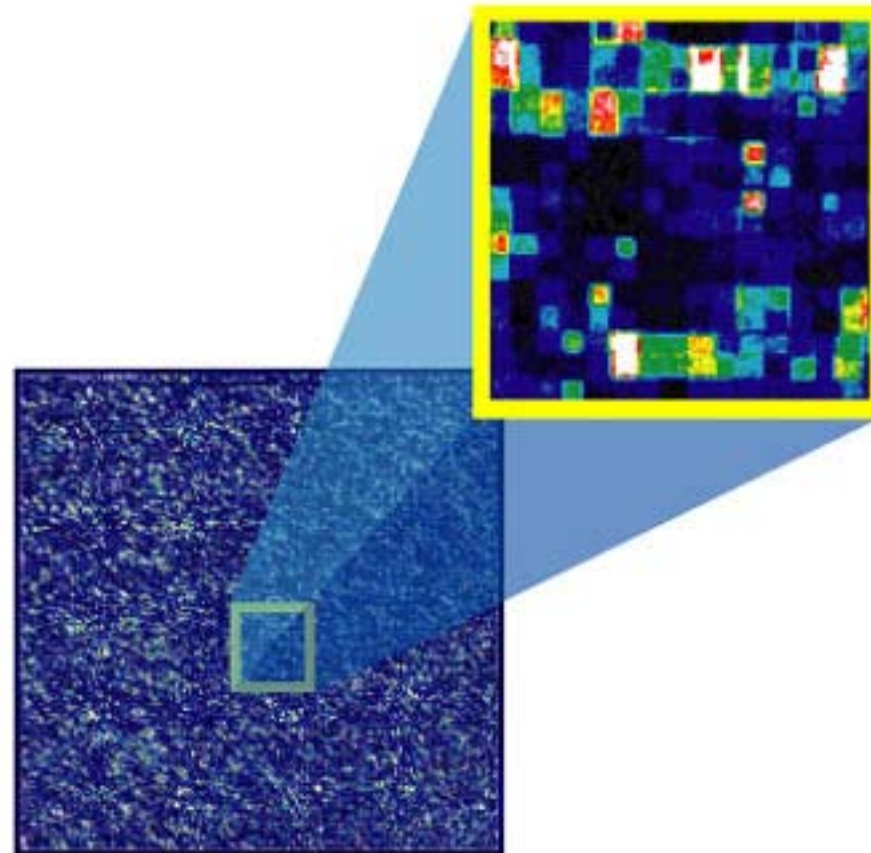
High-throughput



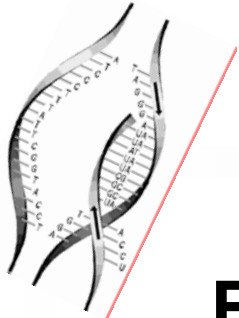
Transcriptomics



b) High-throughput



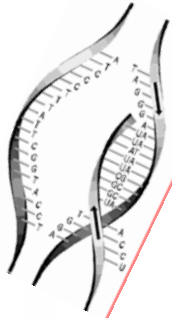
Scanned image



Transcriptomics

Bioinformatic's role

- Interpretation of the scanned image: raw-data are transformed to real values. Multiple measures from the same oligo at different sites of the microarray are summarised.
- Estimation of the intra-individual change (differential expression before-after)
- Estimation of the change in expression levels for the whole sample with several subjects
- Selection of significant results (corrections!)
- Clustering of transcripts and subjects
- In a time-series: time-series analysis (tendencies).



Transcriptomics

R-based software – bioconductor

BIOCONDUCTOR
open source software for **bioinformatics**

Bioconductor is an open source and open development software project for the analysis and comprehension of genomic data.

home | getting started | overview | downloads | documentation | publications | workshops | cabig

project news

- ▶ [2008-05-01](#)
BioC 2.2, consisting of 260 packages and designed to work with R 2.7.0, was released today.
- ▶ [2008-03-04](#)
BioConductor release scheduled for 30 April 2008.

[More...](#)

QUICK LINKS

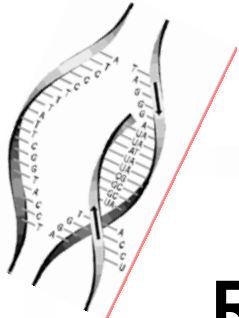
- ▶ Getting Started
- ▶ Installation
- ▶ Downloads
- ▶ Software
- ▶ FAQ

Using Bioconductor for CHIP-seq experiments

Three day course on how CHIP-seq experiments can be analyzed within Bioconductor, November 12-14, 2008 Seattle, WA. [Details and registration](#).

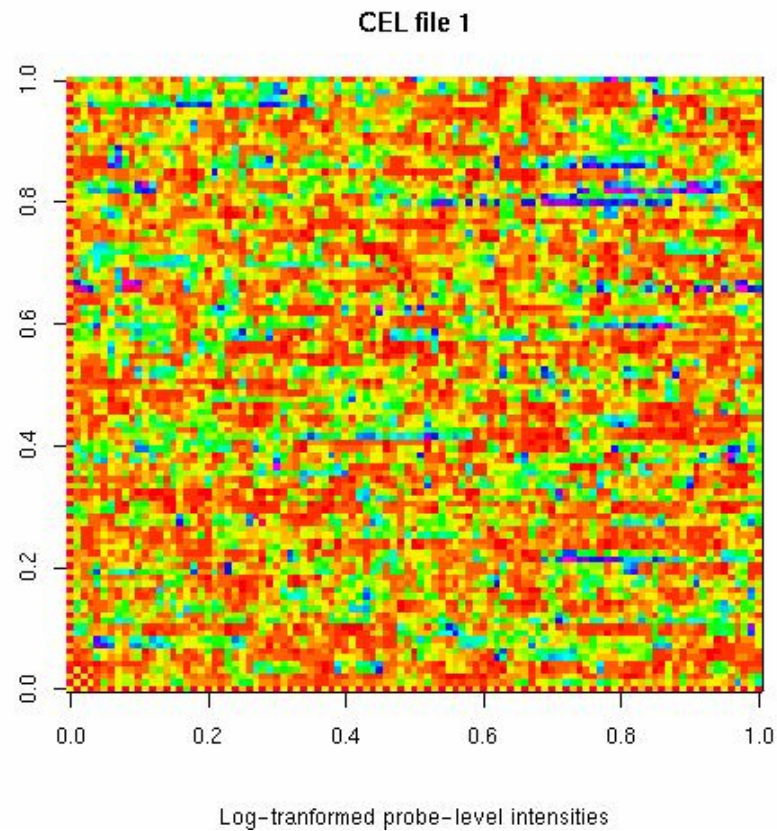
BioC Release 2.2

BioConductor 2.2 was released 1 May, 2008. This release is designed for R 2.7.0. View the packages [here](#)

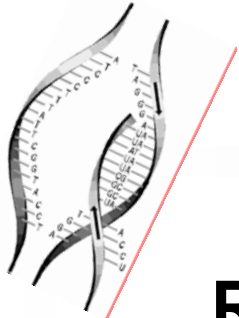


Transcriptomics

R-based software – bioconductor

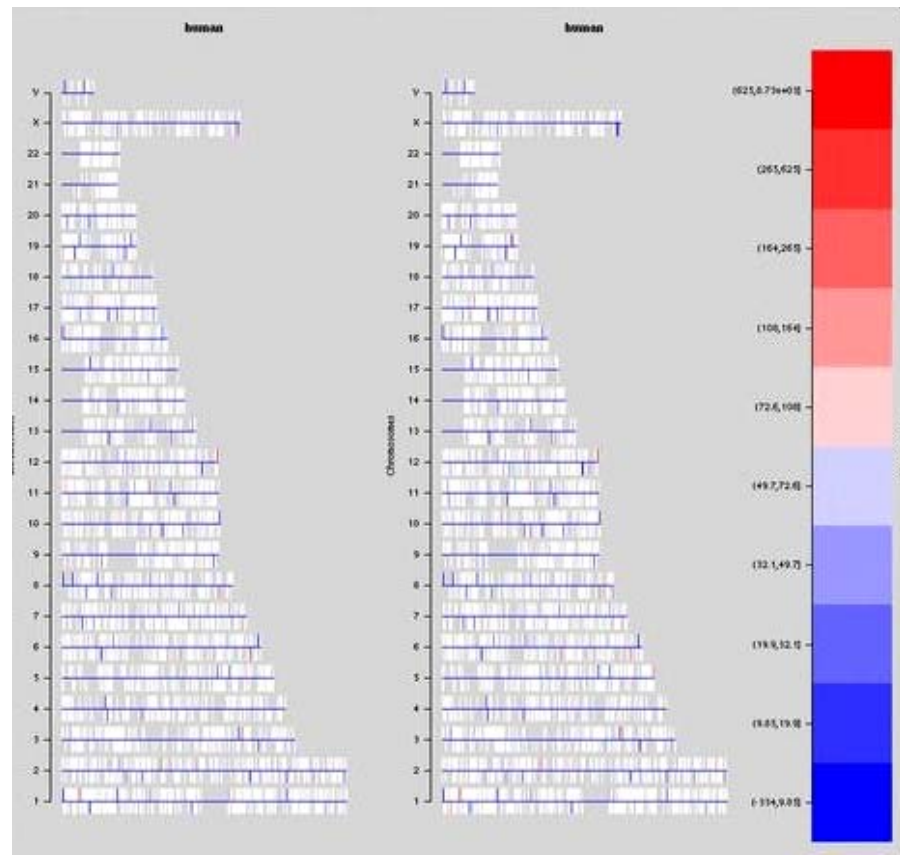


From raw-data transformation (CEL-files) ...

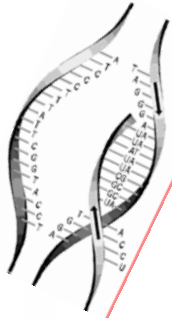


Transcriptomics

R-based software – bioconductor

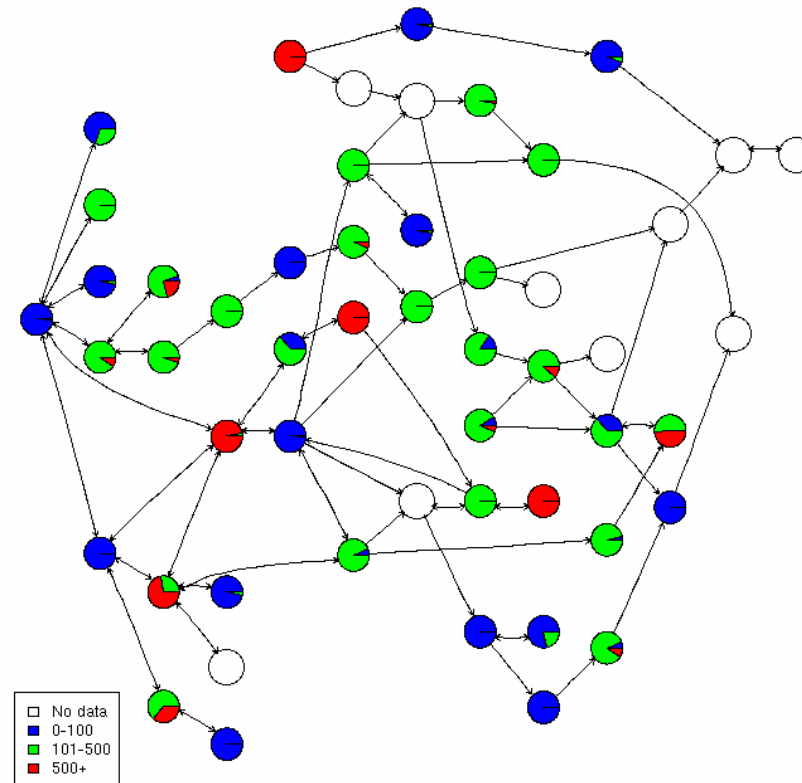


... to final analysis

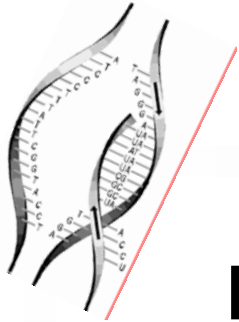


Transcriptomics

R-based software – bioconductor



... and even meta-analysis



Transcriptomics

Databases – GEO

The screenshot shows the NCBI Gene Expression Omnibus (GEO) website. At the top left is the NCBI logo. To the right is the GEO logo with the text "Gene Expression Omnibus". Below the logos is a navigation bar with links: HOME, SEARCH, SITE MAP, Handout, NAR 2006 Paper, NAR 2002 Paper, FAQ, MIAME, and Email GEO. Below the navigation bar is a breadcrumb trail: NCBI > GEO. On the right side of the breadcrumb trail, it says "Not logged in | Login".

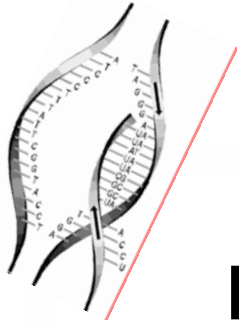
The main content area features a description of GEO: **Gene Expression Omnibus:** a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Below the description is a "GEO navigation" section with two main categories: **QUERY** and **BROWSE**. Under **QUERY**, there are three options: "DataSets" (with a search box and a "GO" button), "Gene profiles" (with a search box and a "GO" button), and "GEO accession" (with a search box and a "GO" button). Below "GEO accession" is a link for "GEO BLAST". Under **BROWSE**, there are two main options: "DataSets" and "GEO accessions". "GEO accessions" is further divided into three sub-options: "Platforms", "Samples", and "Series".

On the right side of the page, there are two summary boxes. The first is titled "Public data" and contains the following information:

GPL Platforms	3129
GSM Samples	128454
GSE Series	5203
<i>Total</i>	136786

The second box is titled "Site contents" and contains a "Documentation" section with links to: Overview | FAQ, Submission guide, Linking & citing, Journal citations, DataSet clusters, GEO announce list, Data disclaimer, and GEO staff. Below this is a "Query & Browse" section with links to: Repository browser, Submitter contacts, SAGEmap, and FTP site.



Transcriptomics

Databases – GEO

The screenshot shows the NCBI GEO Datasets search interface. The search query is 'asthma', and the results are displayed in a table format. The first result is highlighted, showing details for GDS2852.

NCBI GEO DATASETS
Gene Expression Omnibus

All Databases PubMed Nucleotide Proteins Genome Structure PMC Journals

Search GEO DataSets for asthma [Go] [Clear] [Save Search]

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All: 73 DataSets: 20 Platforms: 17 Series: 36

Items 1 - 20 of 73 Page 1 of 4 Next

1: GDS2852 record: Interleukin 13 effect on bronchial cell line: time course [Homo sapiens] [GEO Profiles, Links](#)

Summary: Analysis of bronchial A549 cells at various time points up to 24 hours following treatment with the cytokine interleukin 13 (IL-13). IL-13 contributes to the pathogenesis of asthma.
Parent Platform: [GPL96](#)
Reference Series: [GSE3183](#)

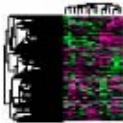
Type: gene expression array-based, count

Subsets: 2 agent, 4 time sets.

Supplementary Files: [CEL download...](#)

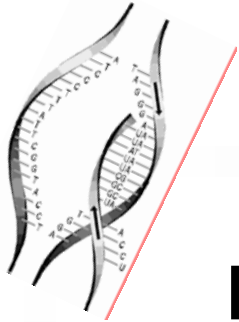
Samples: 15

GSM47462 : PGA-HAH-ctr_0h-1bUA-s2
GSM47468 : PGA-HAH-ctr_0h-1aUA-s2
GSM47472 : PGA-HAH-ctr_0h-1cUA-s2
GSM47465 : PGA-HAH-ctr_24h-1bUA-s2
GSM47469 : PGA-HAH-ctr_24h-1aUA-s2
GSM47470 : PGA-HAH-ctr_24h-1cUA-s2



2: GDS2647 record: Asthma model: lungs [Mus musculus] [GEO Profiles, Links](#)

Summary: Comparison of whole lungs of wild-type and recombinase activating gene (RAG) deficient animals sensitized and then challenged with the allergen ovalbumin. RAG deficient animals do not develop allergic asthma. Results provide insight into the molecular basis of the allergic asthmatic response.
Parent Platform: [GPL1261](#)
Reference Series: [GSE6858](#)

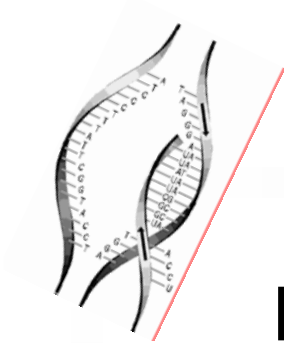


Transcriptomics

Databases – GEO

The screenshot shows the NCBI GEO database interface. At the top left is the NCBI logo. To the right is the GEO logo (Gene Expression Omnibus). Below these are navigation links: HOME, SEARCH, SITE MAP, Handout, NAR 2006 Paper, NAR 2002 Paper, FAQ, MIAME, and Email GEO. The main header shows the path: NCBI > GEO > Accession Display. On the right of the header, it says "Not logged in | Login". Below the header is a search bar with the following fields: Scope: Self, Format: HTML, Amount: Quick, GEO accession: GSM47462, and a GO button. The main content area displays the details for Sample GSM47462, including a link to "Query DataSets for GSM47462".

Sample GSM47462	Query DataSets for GSM47462
Status	Public on May 26, 2005
Title	PGA-HAH-ctr_0h-1bUA-s2
Sample type	RNA
Source Name	Human Cell Line
Organism(s)	Homo sapiens
Extracted molecule	total RNA
Description	Human bronchial cell line A549 harvested at time point 0 without any treatment
Submission date	Apr 05, 2005
Contact name	Eric Hoffman
E-mail(s)	ehoffman@cnmcresearch.org
Phone	202-884-6011
Fax	202-884-6014
URL	http://pepr.cnmcresearch.org
Organization name	Children's National Medical Center
Department	Children's Research Institute
Lab	Research Center for Genetic Medicine
Street address	111 Michigan Ave. 5th Floor
City	Washington
State/province	DC
ZIP/Postal code	20010
Country	USA



Transcriptomics

Databases – GEO

ID_REF
VALUE SIGNAL (See Affymetrix Literature)
ABS_CALL Presence or Absence of gene transcript in sample

Data table

ID_REF	VALUE	ABS_CALL
1007_s_at	366.15045166015625	P
1053_at	106.40161895751953	P
117_at	39.683868408203125	A
121_at	416.8681945800781	P
1255_g_at	33.98335647583008	A
1294_at	111.41187286376953	A
1316_at	40.22675704956055	P
1320_at	42.27968978881836	A
1405_i_at	2.7143423557281494	A
1431_at	14.615274429321289	A
1438_at	40.67414093017578	A
1487_at	197.6019744873047	P
1494_f_at	154.8931427001953	P
1598_g_at	257.0430603027344	A
160020_at	203.87881469726562	A
1729_at	115.28802490234375	P
1773_at	104.36334228515625	A
177_at	53.95371627807617	A
179_at	381.58062744140625	P

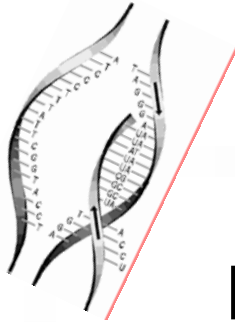
Total number of rows: **22283**

Table truncated, full table size **678 Kbytes**.

[View full table...](#)

Supplementary file	Size	Download	File type/resource
GSM47462.CEL.gz	3.6 Mb	(ftp) (http)	CEL

Raw data provided as supplementary file



Transcriptomics

Databases – GEO

Sample organism:	Homo sapiens	Platform organism:	Homo sapiens
Feature count:	4132	Value type:	transformed count
Series:	GSE8759	Series published:	09/14/2007
Last GDS update:	09/17/2007		

Subset and Sample Info

Sample selection

check all

uncheck all

toggle

Data

download

analysis

Find gene in this DataSet..

2 assigned subsets			Find genes differentially expressed between groups..		
Samples	Type	Description	A	0.050 significance level	B
<input checked="" type="checkbox"/> (41)	disease state	control	<input checked="" type="checkbox"/>	↔	<input type="checkbox"/>
<input checked="" type="checkbox"/> (60)	disease state	Marfan syndrome	<input type="checkbox"/>	↔	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> GDS2960 only <input checked="" type="checkbox"/> ranks <input checked="" type="checkbox"/> values <input type="checkbox"/> subset effects			<input checked="" type="checkbox"/>	Query A vs. B	<input checked="" type="checkbox"/>

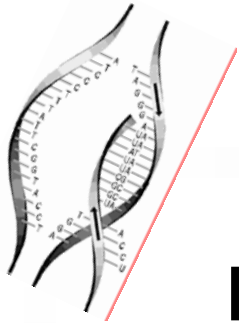
101 samples, order: none

[GSM217644](#) : fibroblast, control subject NF10 rep 1
sro1: cultured skin fibroblasts from control subjects

[GSM217645](#) : fibroblast, control subject NF10 rep 2
sro1: cultured skin fibroblasts from control subjects

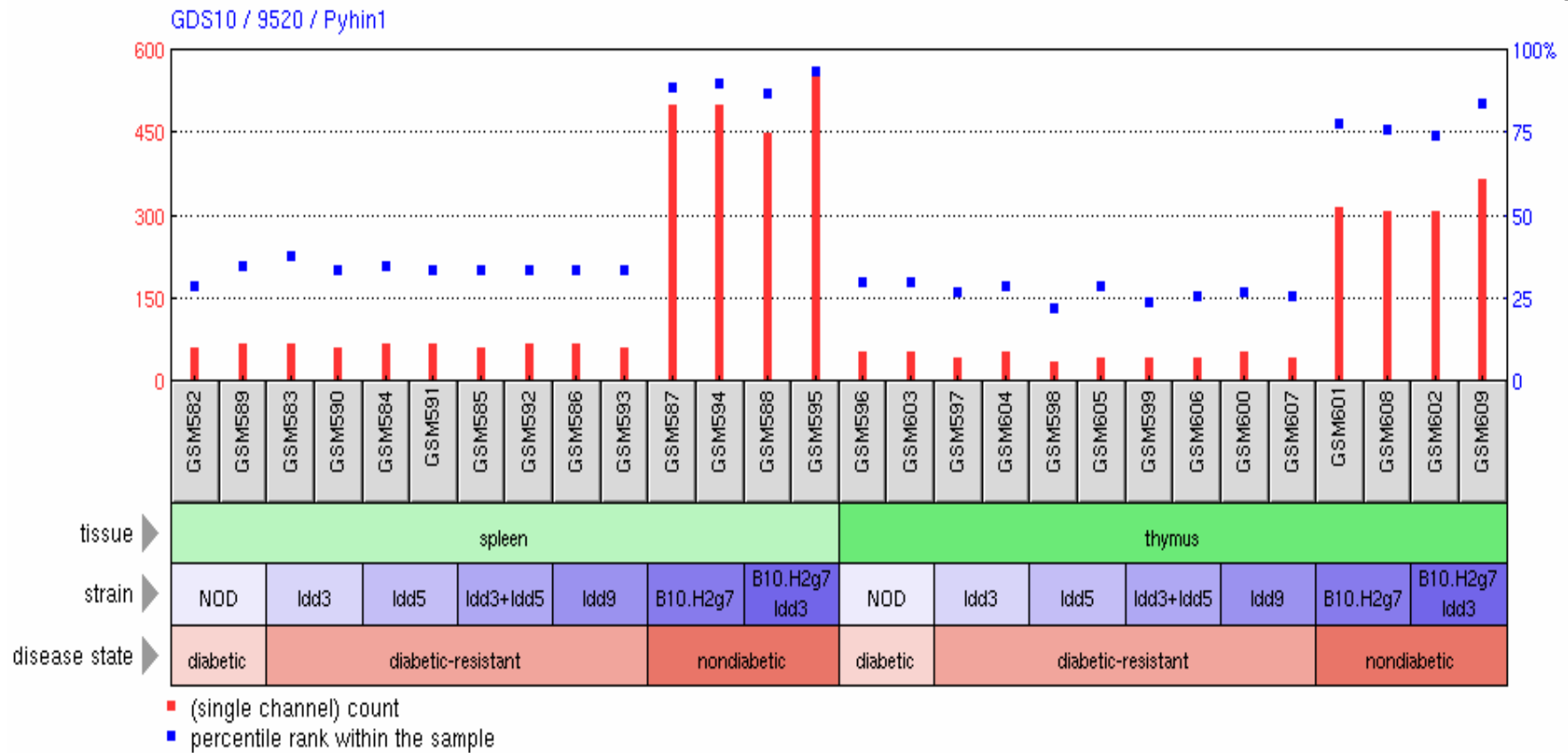
[GSM217646](#) : fibroblast, control subject NF11 rep 1
sro1: cultured skin fibroblasts from control subjects

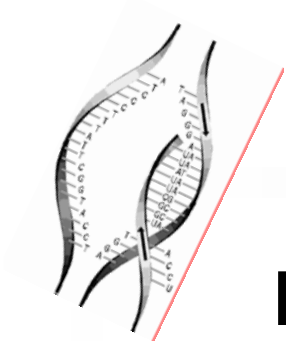
[GSM217647](#) : fibroblast, control subject NF11 rep 2
sro1: cultured skin fibroblasts from control subjects



Transcriptomics

Databases – GEO





Transcriptomics

Databases – AE

The screenshot shows the ArrayExpress database homepage. At the top is a navigation bar with the EMBL-EBI logo, an 'EB-eye Search' button, a dropdown menu set to 'All Databases', a search input field containing 'Enter Text Here', and buttons for 'Go', 'Reset', and 'Advanced Search'. Below this is a secondary navigation bar with links for 'Databases', 'Tools', 'Groups', 'Training', 'Industry', 'About Us', 'Help', and 'Site Index'. The main content area features a large 'ArrayExpress' header with a description: 'ArrayExpress is a public repository for **microarray data**, which is aimed at storing MIAME-compliant data in accordance with MGED recommendations. The ArrayExpress Data Warehouse stores gene-indexed **expression profiles** from a curated subset of experiments in the repository.' To the right of the text is the ArrayExpress logo, a stylized 'AE' in red and green. Below the description is a 'More Info' link. Two search panels are visible: 'Experiments' with an 'RSS' button, a search term 'RNAi', a 'query' button, and links to 'Browse experiments', 'Advanced query interface', and 'Submitter/reviewer login'; and 'Expression Profiles' with 'Gene(s)' set to 'Ctnd2', 'Species' set to 'Homo sapiens', a 'query' button, and a link to 'ArrayExpress Warehouse Homepage'.

EMBL-EBI **EB-eye Search** All Databases **Go** Reset ? Advanced Search

Databases Tools Groups Training Industry About Us Help Site Index

ArrayExpress

ArrayExpress is a public repository for **microarray data**, which is aimed at storing MIAME-compliant data in accordance with MGED recommendations. The ArrayExpress Data Warehouse stores gene-indexed **expression profiles** from a curated subset of experiments in the repository.

[More Info](#)

Experiments RSS

Search term(s)

» Browse experiments

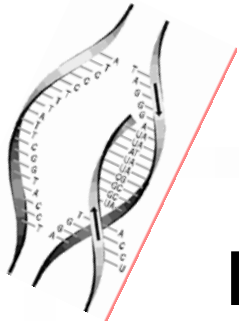
» Advanced query interface
» Submitter/reviewer login

Expression Profiles

Gene(s)

Species

» ArrayExpress Warehouse Homepage



Transcriptomics

Databases – AE

MBL-EBI EB-eye Search All Databases

Databases Tools EBI Groups Training Industry About Us Help Site Index

Experiment, citation, sample and factor annotations [clear] Filter on [reset] Display options [reset]

Match whole words Loaded in ArrayExpress Atlas

Any species Any array Any experiment type

25 experiments per page Detailed view

» ArrayExpress Browse Help

ID	Title	Assays	Species	Date	Processed	Raw	Atlas
E-GEOD-4122	Gene Expression Analysis of Ovarian Samples	67	Homo sapiens	2008-10-16		-	-
E-GEOD-4938	murine CD4+ and CD8+ T-cells from wildtype and Gfi knockout cells	4	Mus musculus	2008-10-16		-	-
E-GEOD-3808	Transcription profiling of embryonic kidneys from wild type and Hoxa	28	Mus musculus	2008-10-14			
E-TOXM-39	Transcription profiling of rat liver and kidney samples performed in :	517		2008-10-09	-	-	-
E-TABM-428	Transcription profiling analysis of the PecS regulon of Erwinia chrysa	4	Erwinia chrysanthemi	2008-10-08			-
E-MTAB-57	Transcription profiling of colon cancer tumor and normal biopsies fro	47	Homo sapiens	2008-10-07			-
E-TABM-451	Transcription profiling of potato germplasm from different cultivars a	16	Solanum tuberosum	2008-10-07			-
E-MEXP-1820	microRNA profiling of human gastric cancer cells after treatment with	2	Homo sapiens	2008-10-06		-	-
E-MEXP-1821	Transcription profiling of human gastric cancer cells after treatment	1	Homo sapiens	2008-10-06	-		-
E-MEXP-1833	Comparison of WT Erk1 and Erk2 deficient CD8 T cells	4	Mus musculus	2008-10-06			-
E-MEXP-1711	Transcription profiling of liver from wild type and RXR-alpha deficien	24	Mus musculus	2008-10-03	-		-
E-MEXP-1827	Transcription profiling of BMM and GM-BMM from wild type and IFNA1	12	Mus musculus	2008-10-03			-
E-MEXP-1690	Transcription profiling of human gangliogliomas and adjacent tissue	12	Homo sapiens	2008-10-02			-
E-MEXP-1826	Transcription profiling time series of human CD34+ stem cells differ	14	Homo sapiens	2008-10-02	-		-
E-TABM-567	Transcription profiling of endometrium biopsies in patients and cont	96	Homo sapiens	2008-10-02			-
E-MEXP-1765	Transcription profiling of budding yeast during meiosis to examine th	16	Saccharomyces cerevisia	2008-09-30			-

6404 experiments, 195218 assays. Displaying experiments 1 to 25. Pages: 1 2 3 4 5 6 7 8 9 10 .. 257