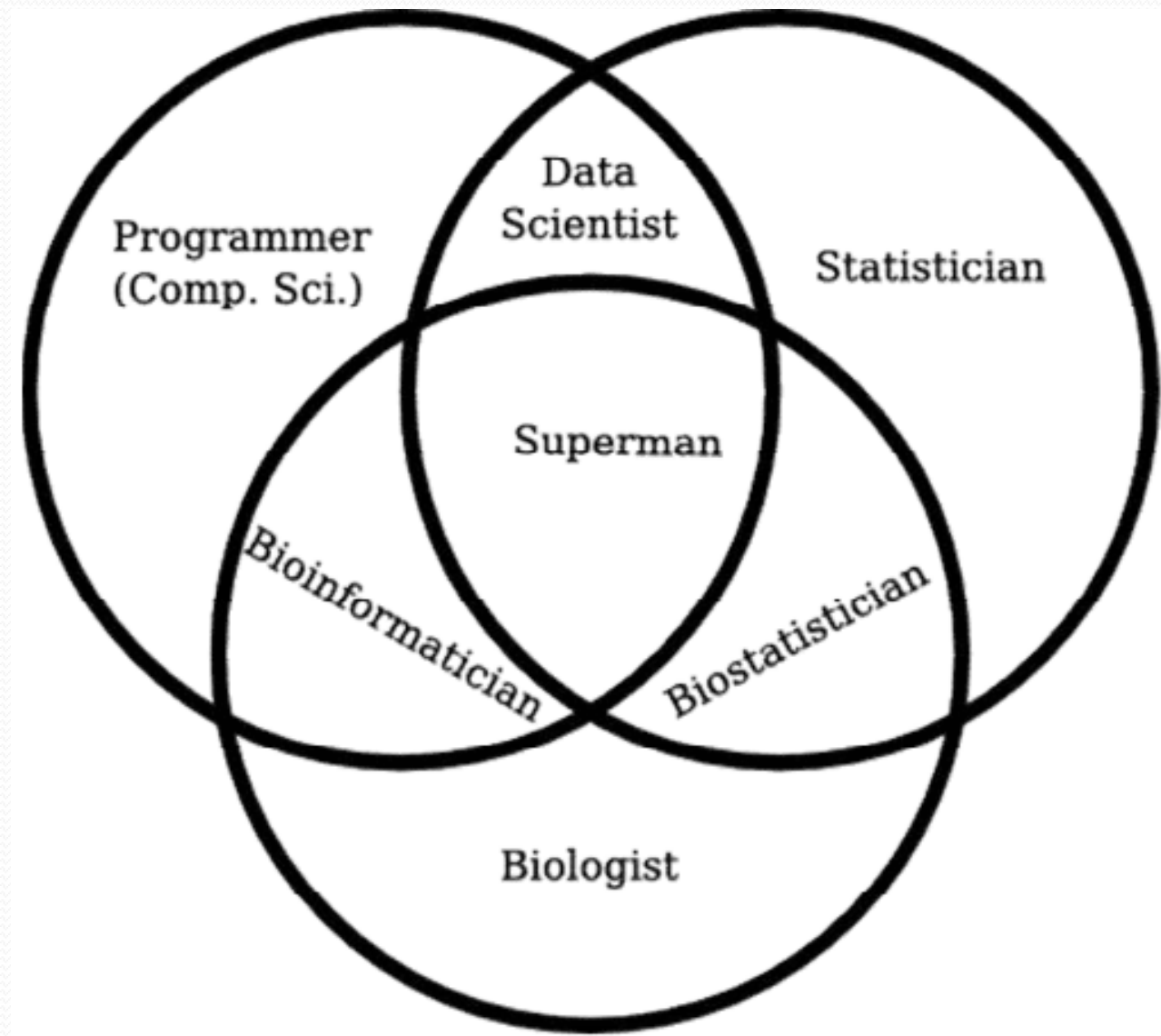


Biomedical Informatics

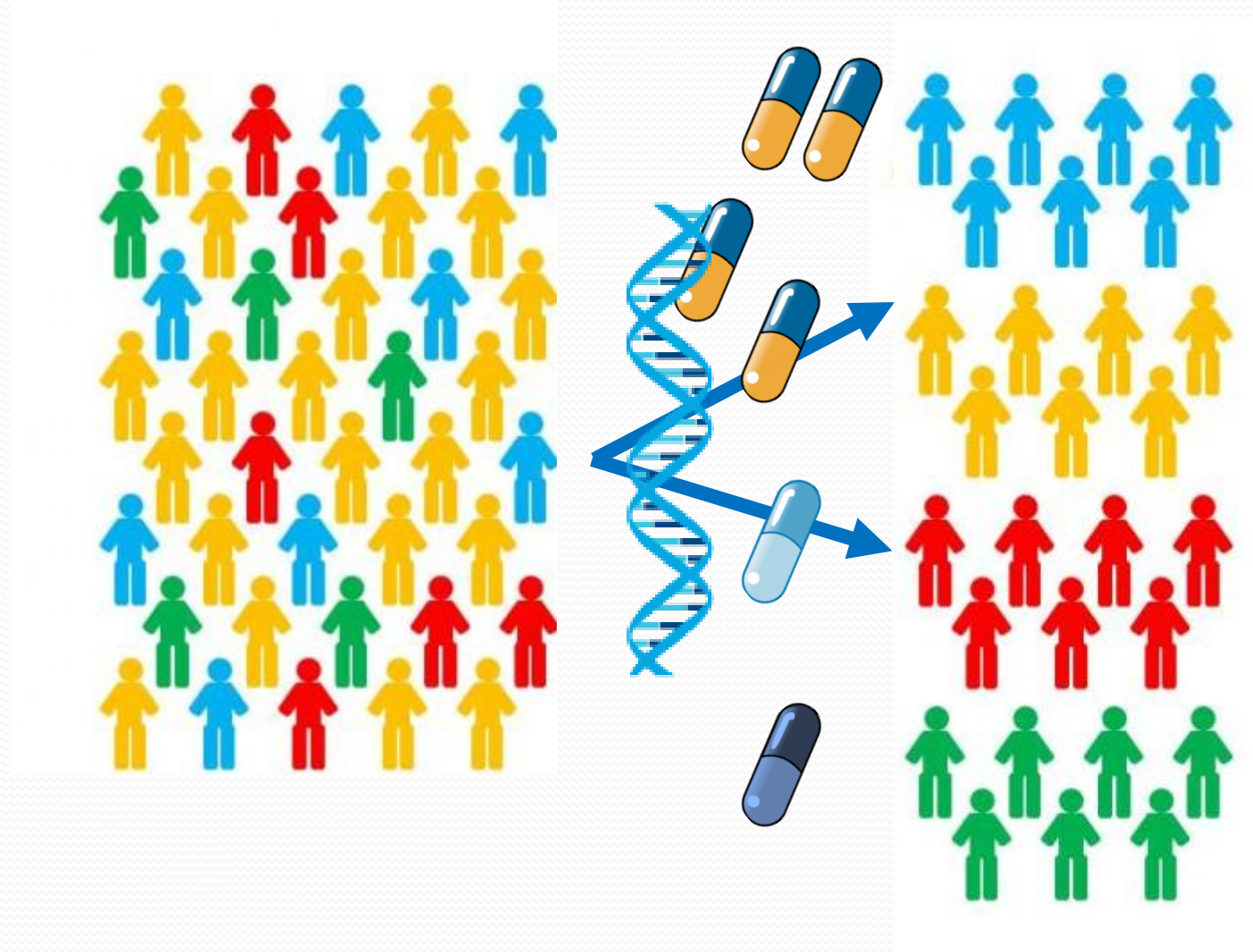
Lefteris Koumakis

Bioinformatics



Biomedical Informatics

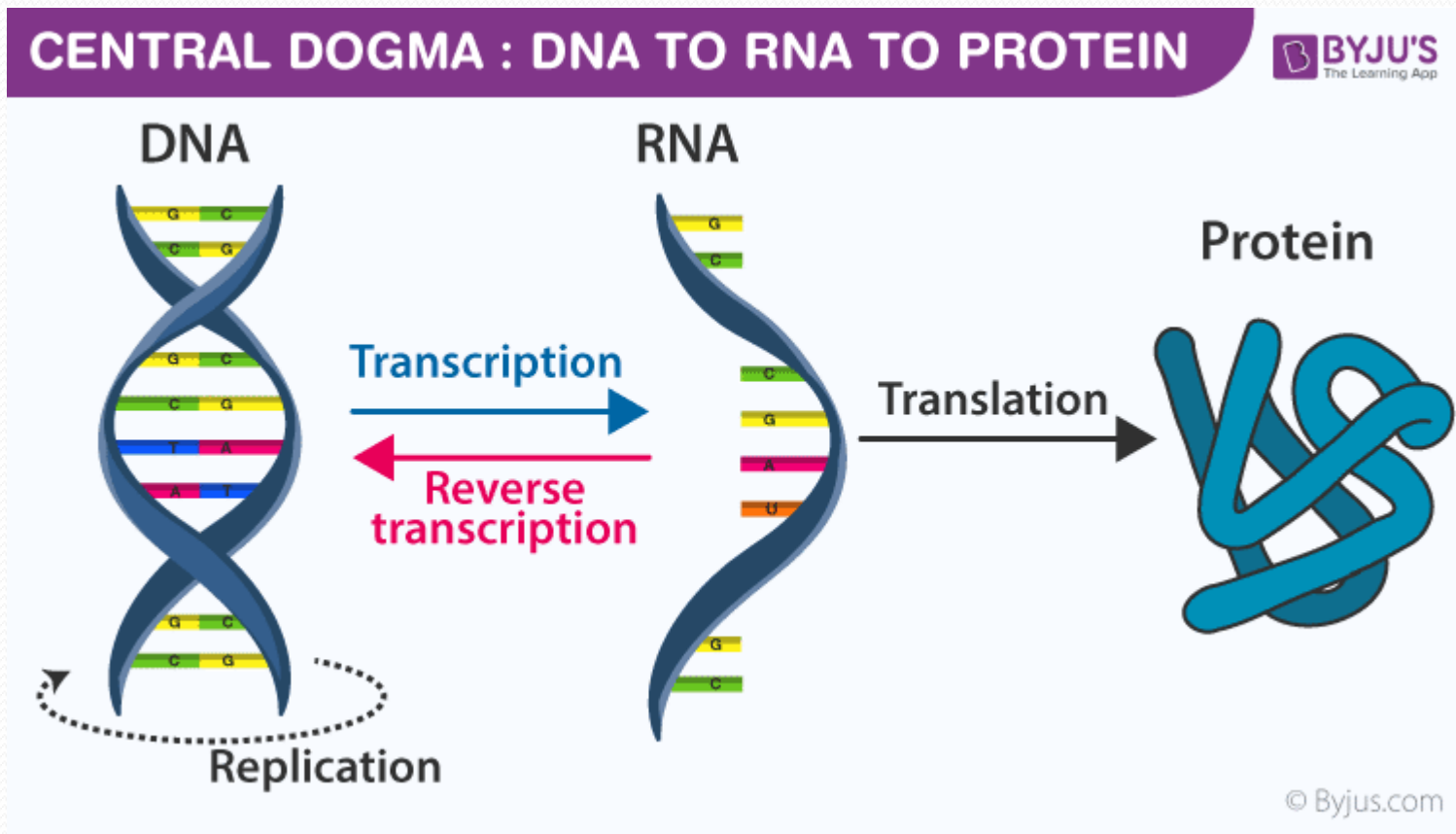
Personalized Medicine



New drugs vs repurposing

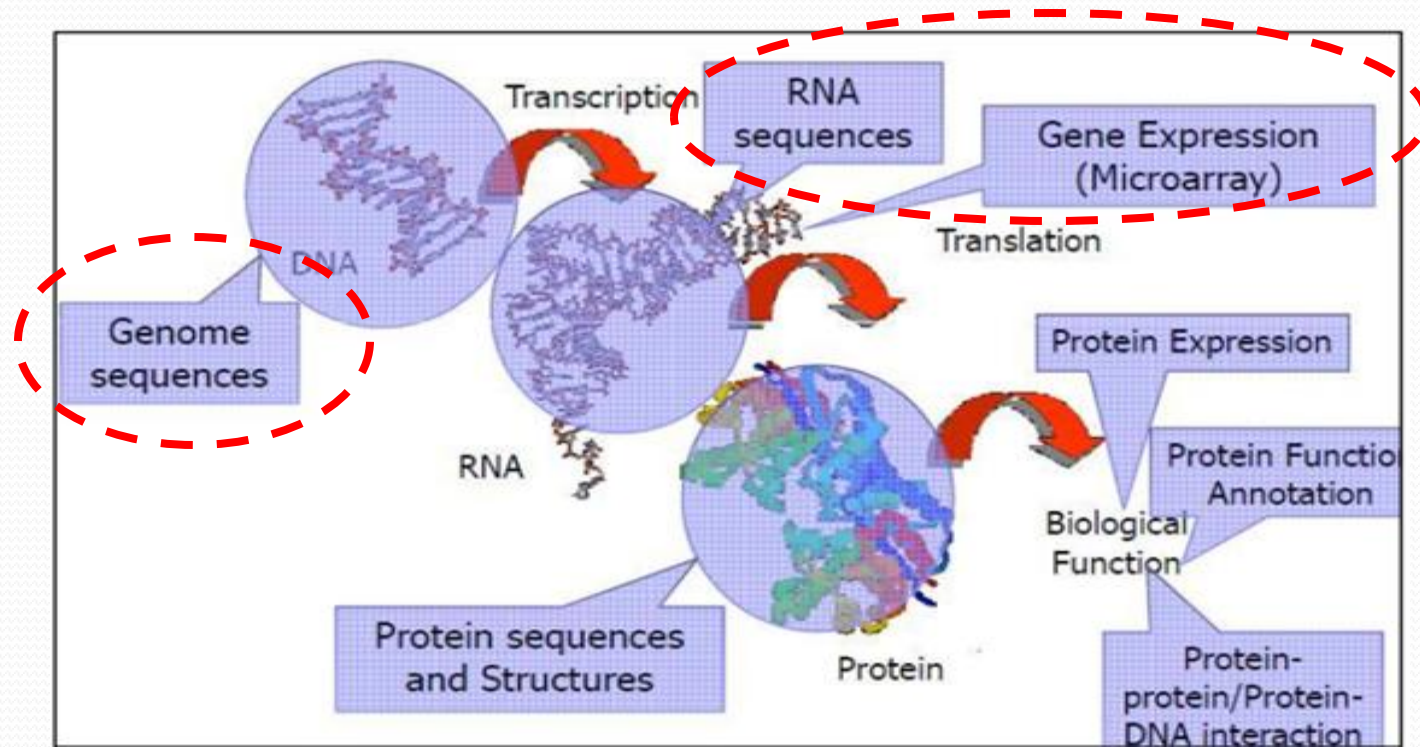


Central dogma of molecular biology



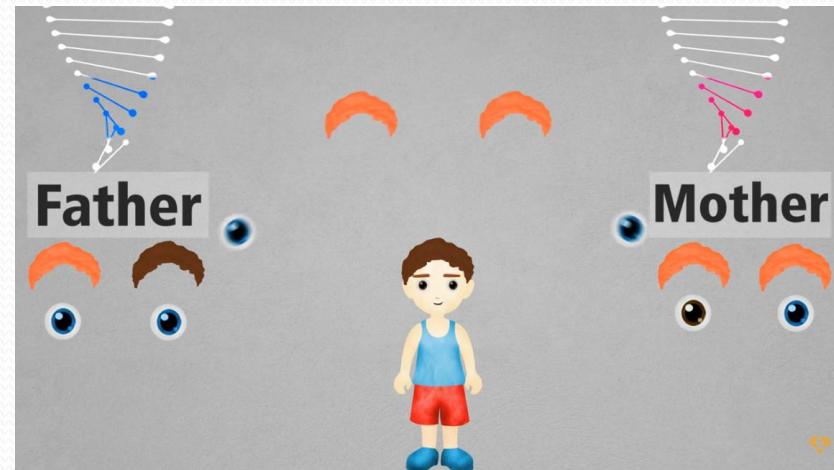
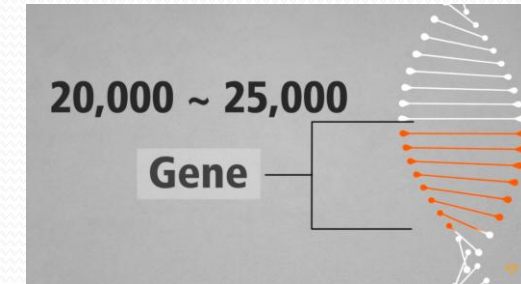
Francis Crick in 1957

Data



Genes

DNA contains all the information necessary for the functioning of an organism, in segmented units of information called genes. Genes are the basic units of heredity. A gene contains the information for the synthesis of a protein.



<https://www.youtube.com/watch?v=5MQdXjRPHmQ>

Genes and evolution

- Genes and evolution

- https://www.youtube.com/watch?v=G4VINRUe_o4

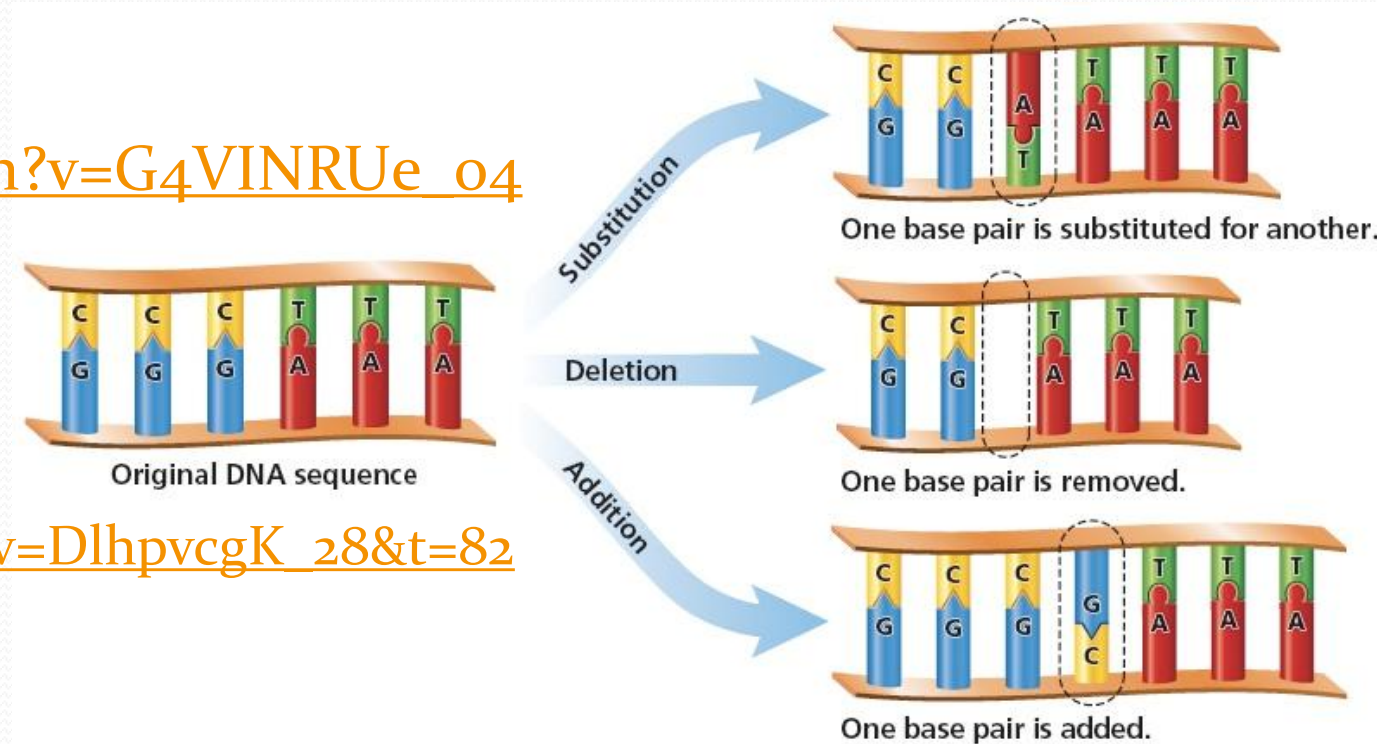
- Types of evolution

- E.g. Point Mutations

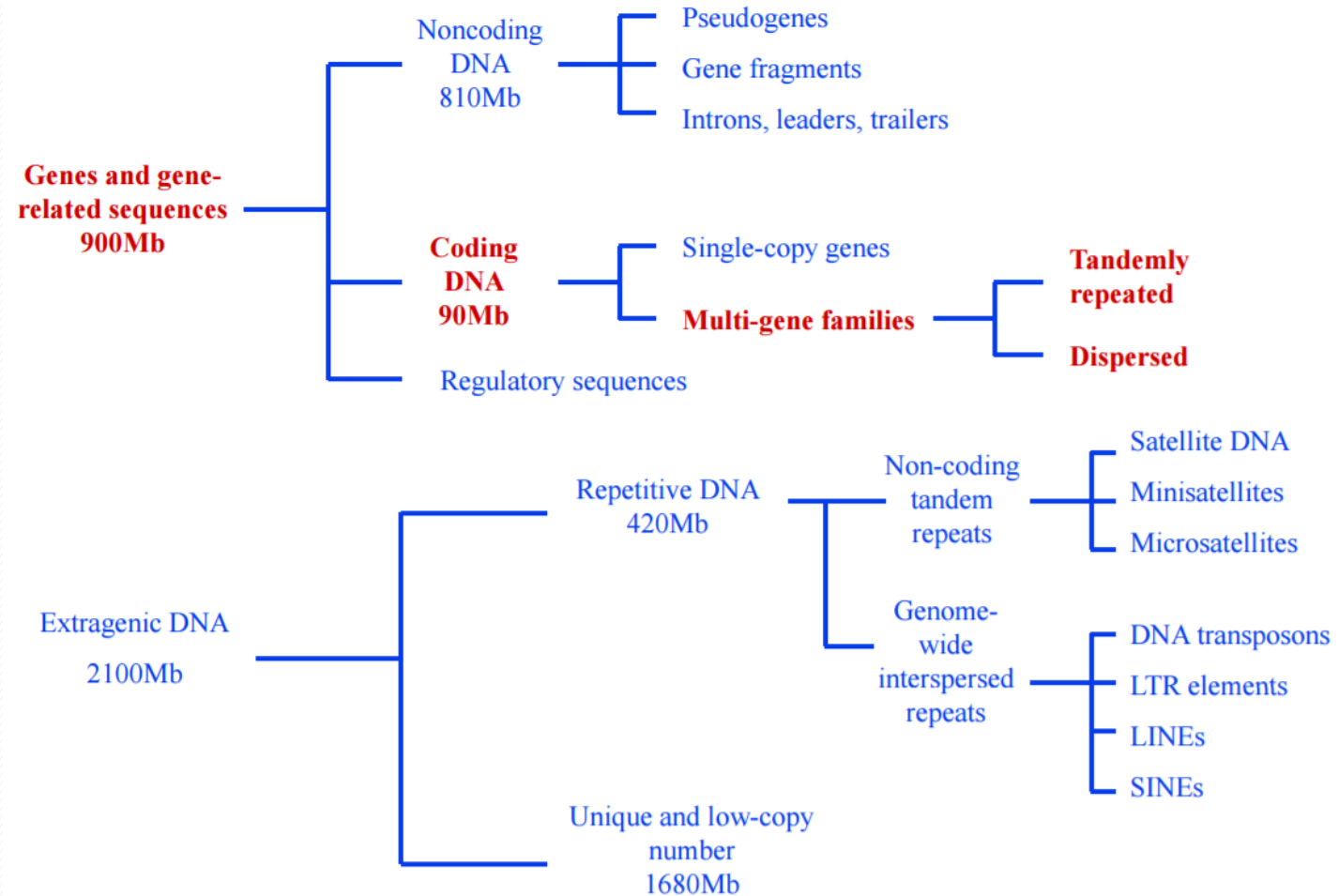
- https://www.youtube.com/watch?v=DlhpvcgK_28&t=82

- Duplication

- https://www.youtube.com/watch?v=G4VINRUe_o4



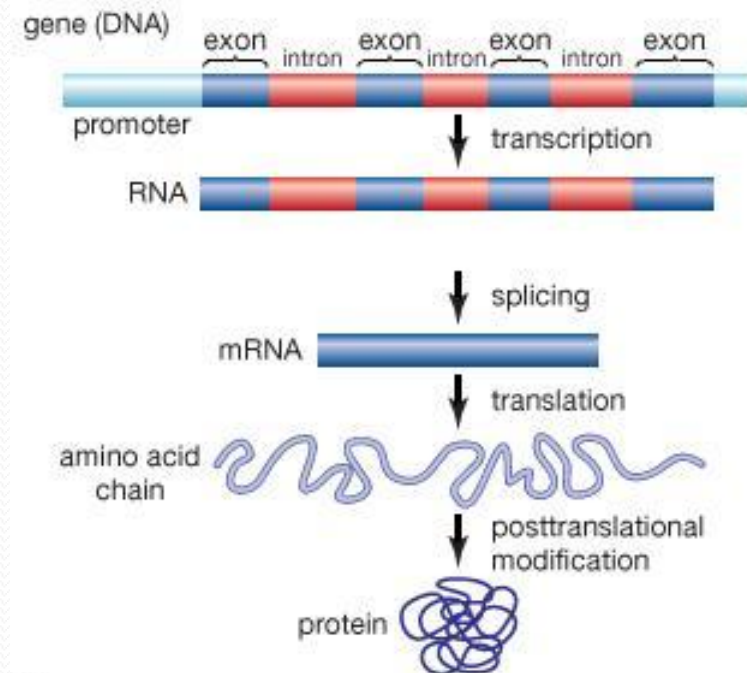
Human Genome



Genes

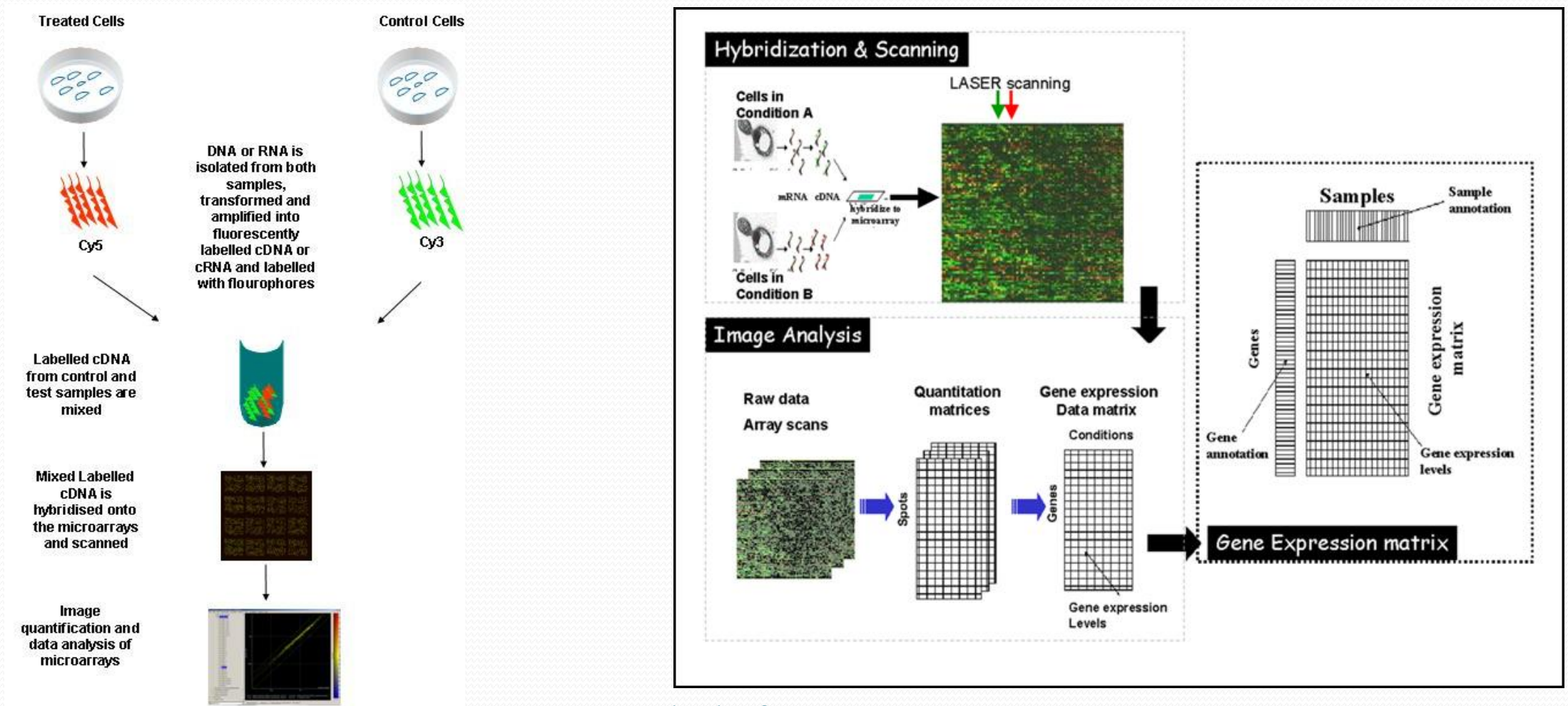
Human Genes

- About 3% of our DNA
- Average length of gene: ~ 8,000 bp
- About 5-6 exons/gene
- Average length of exon : ~200 bp
- Average length of intron : ~2,000 bp
- ~8% contain only one exon
- Some exons can be small 1-3 bp.

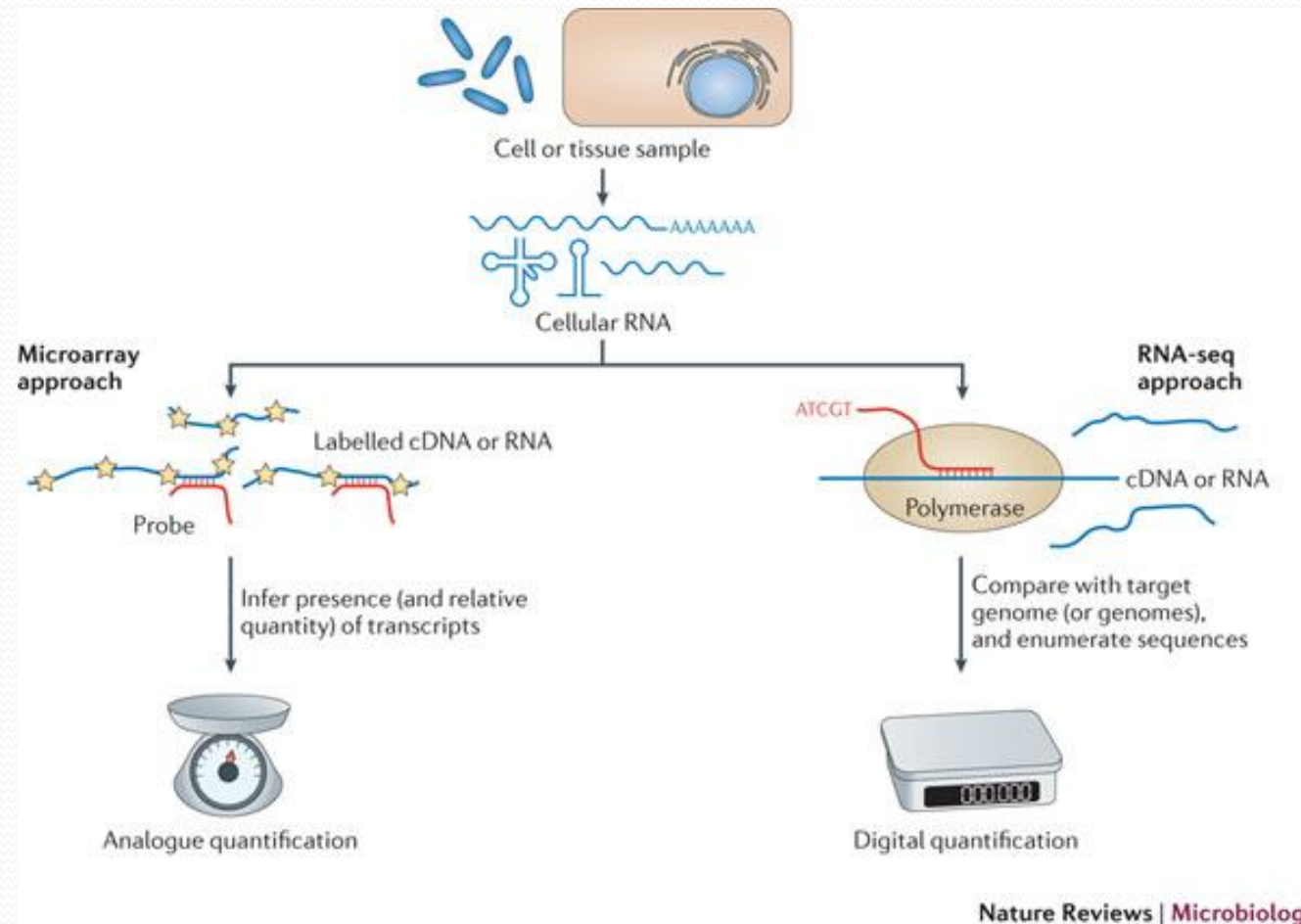


© 2008 Encyclopædia Britannica, Inc.

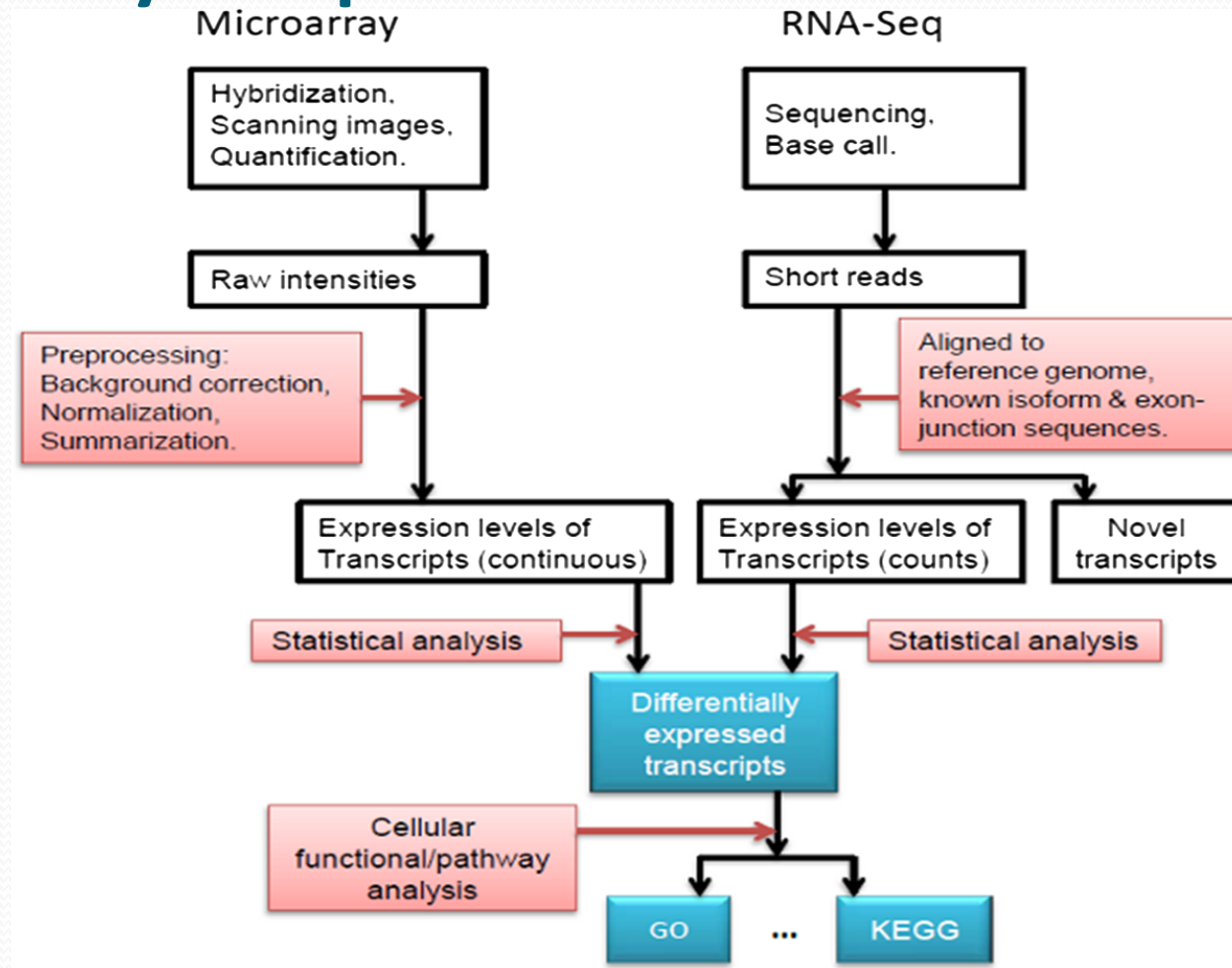
Measuring Gene expressions - Microarrays



Measuring Gene expressions – RNA-seq



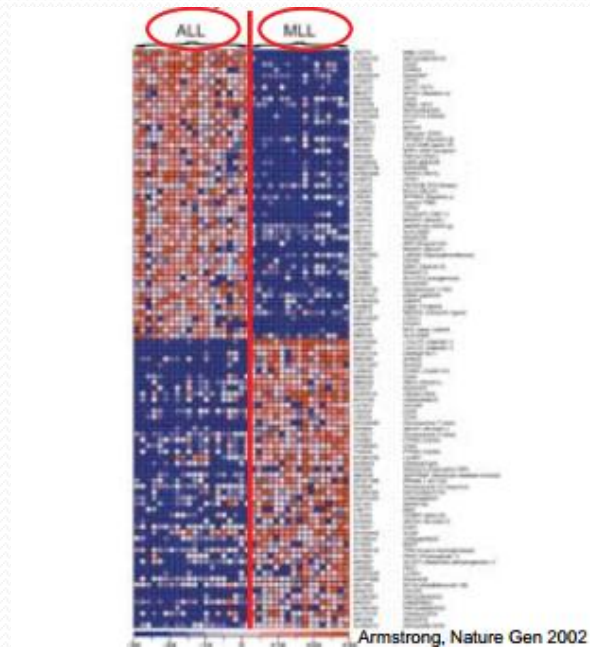
Differentially expressed Genes



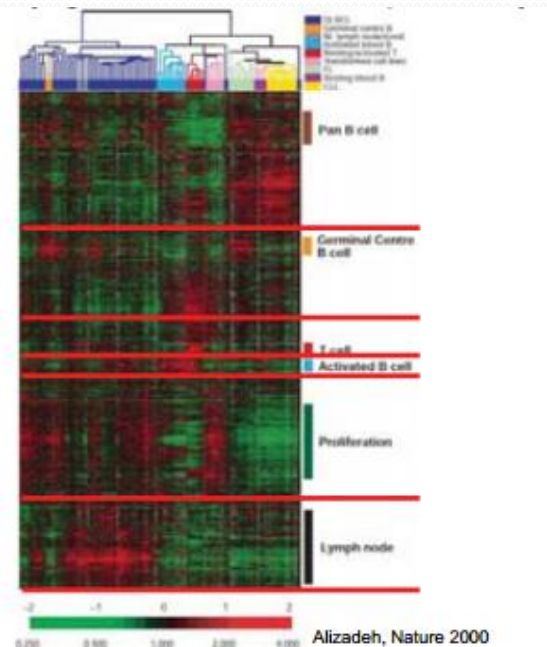
Analysis of gene expression

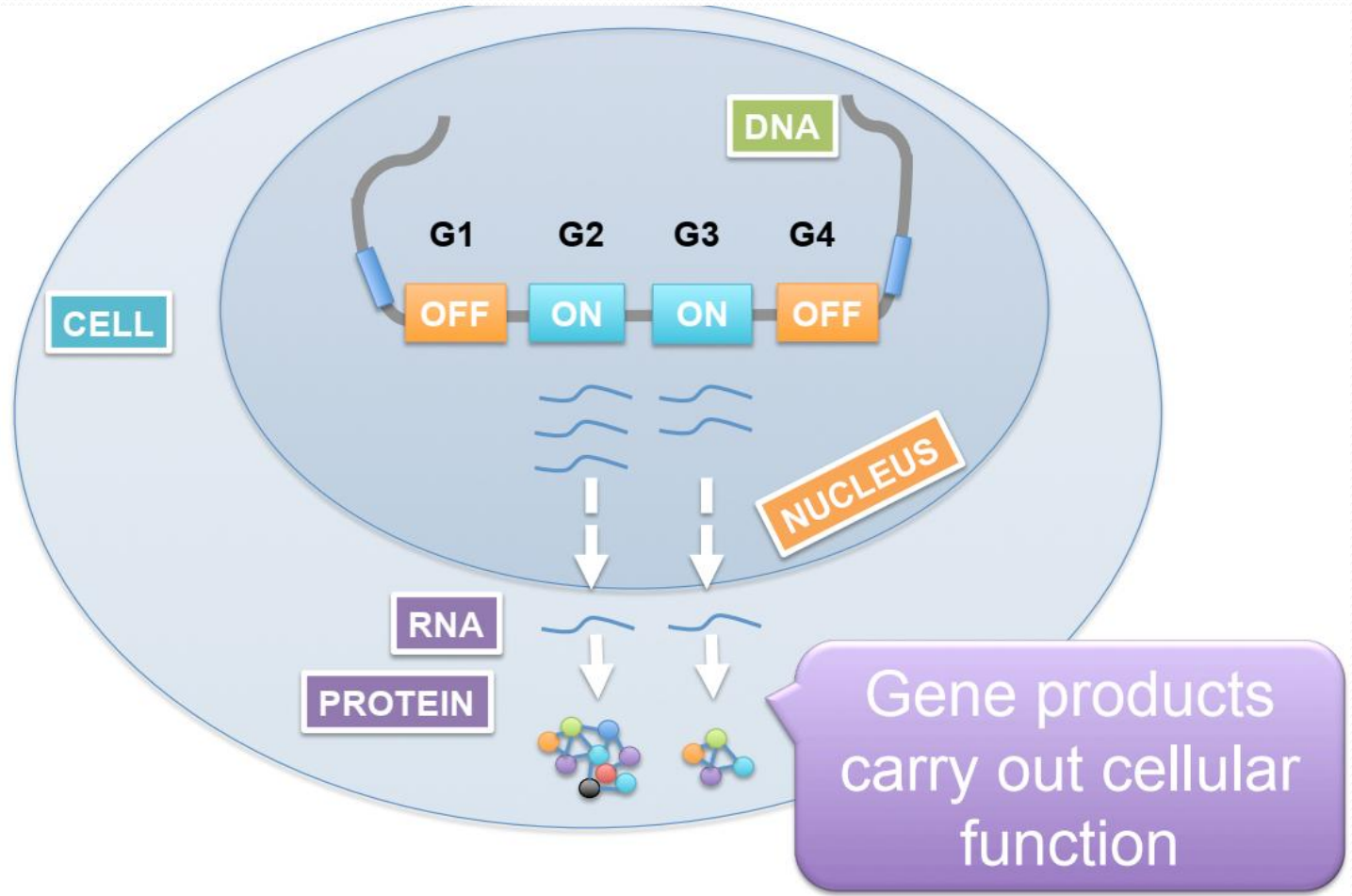
- Classification

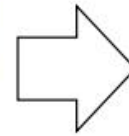
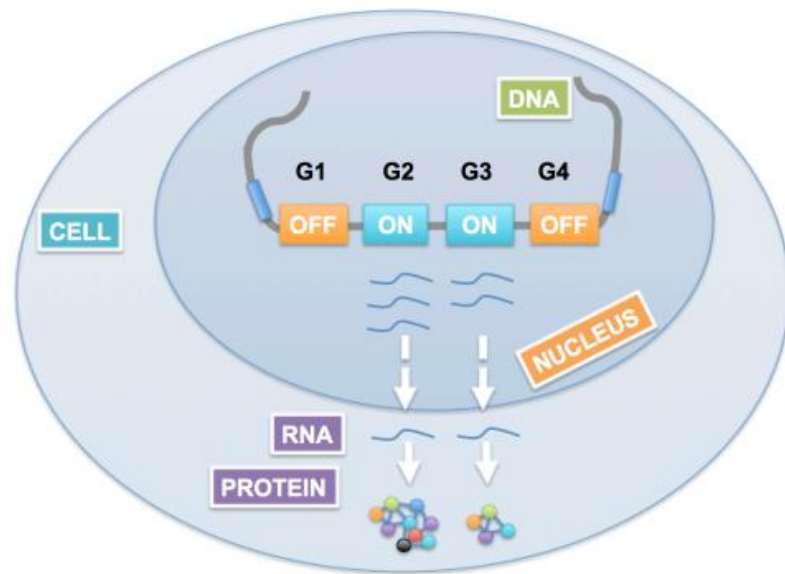
Genes that can discriminate phenotypes



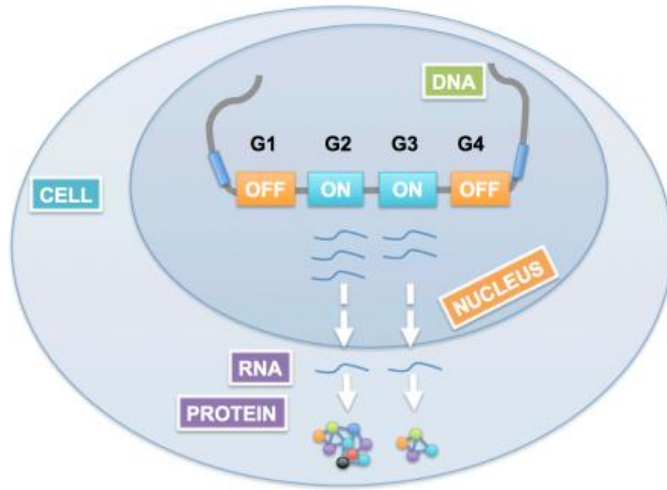
Identification of new genes with the same patterns in specific phenotypes





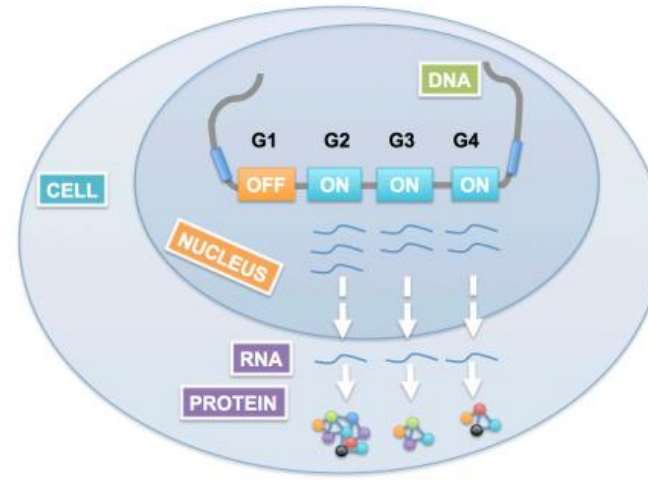


Gene	Expression profile	Expression level
G1	OFF	0
G2	ON	30
G3	ON	20
G4	OFF	0



HEALTHY CELL

Gene	Expression profile	Expression level
G1	OFF	0
G2	ON	30
G3	ON	20
G4	OFF	0



CANCER CELL

Gene	Expression profile	Expression level
G1	OFF	0
G2	ON	30
G3	ON	20
G4	ON	20

Motivation

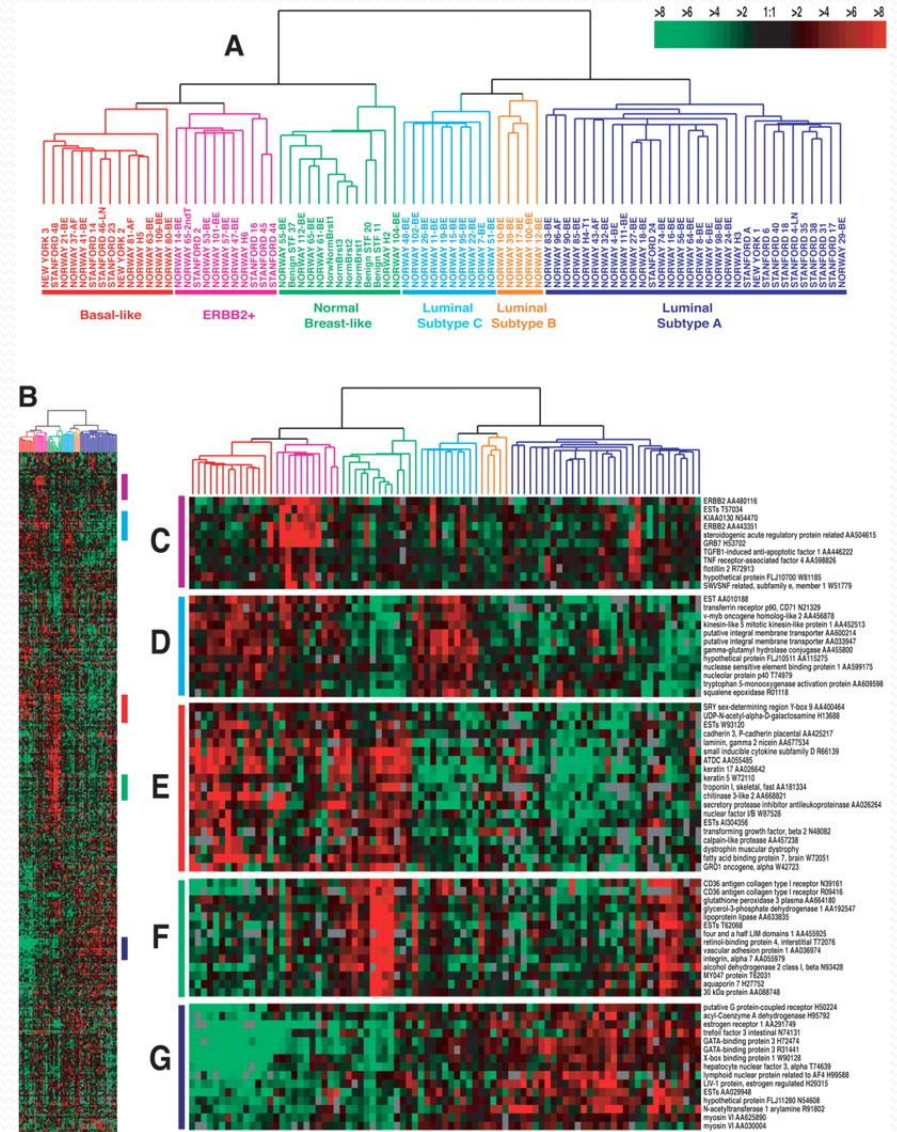
- Genes can be key to understanding what goes wrong / or get fixed under certain condition (cancer, stress etc.)
- In other cases, these genes can be used as ‘features’ for a classifier.
- These genes can also serve as a starting point for a model for the system being studied (e.g. cell cycle, pheromone response etc.)

Our goal is to identify the ‘real’ differences, that is, differences that can be explained by the various errors introduced during the experimental phase.

Sørli et al doi: 10.1073/pnas.191367098

Gene expression patterns of 85 experimental samples representing 78 carcinomas, three benign tumors, and four normal tissues, analyzed by hierarchical clustering using the 476 cDNA intrinsic clone set.

- (A) The tumor specimens were divided into five (or six) subtypes based on differences in gene expression.
- (B) The full cluster diagram scaled down. The colored bars on the right represent the inserts presented in C–G.
- (C) ERBB2 amplicon cluster.
- (D) Novel unknown cluster.
- (E) Basal epithelial cell-enriched cluster.
- (F) Normal breast-like cluster.
- (G) Luminal epithelial gene cluster containing ER.

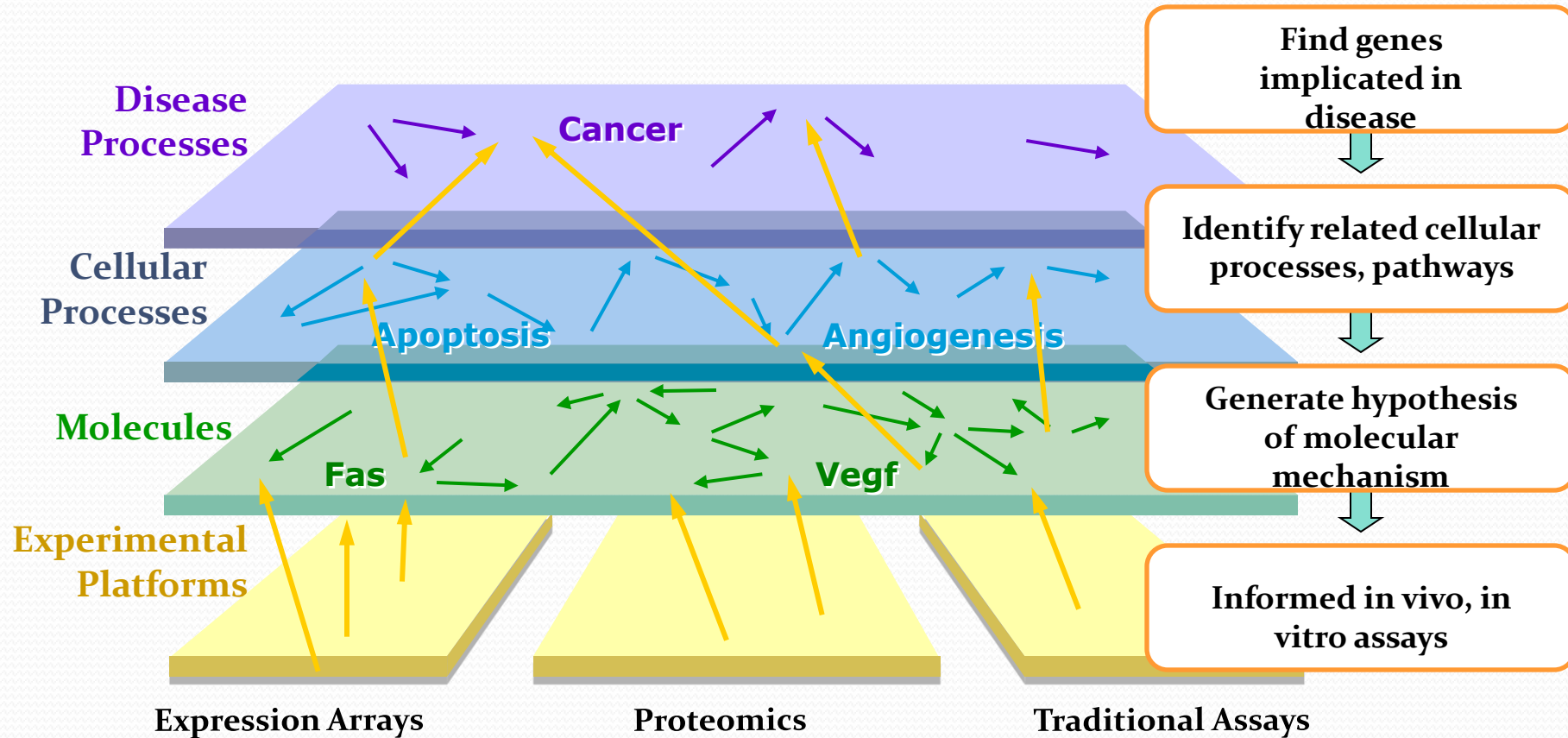


Systems Biology Pathway Analysis

The Challenge

Data Analysis Across Multiple Dimensions of Biology

Systematic Generation of Novel
Biological and Therapeutic Insights



Biomedical Informatics

Pathway analysis - Problem definition

Genes

```

    graph TD
      IL-IR --> TRADO
      TRADO --| FLIP
  
```

	cases				
	POS			NEG	
	case1	case2	case3	case4	case5
IL-IR ON, 1	ON	ON	ON	ON	ON
TRADO ON, 1	ON	ON	ON	OFF	ON
FLIP OFF, 0	OFF	OFF	OFF	OFF	ON
MyD88	ON	ON	ON	ON	ON
NIK	ON	OFF	OFF	ON	OFF

Sub-Paths

	cases				
	POS			NEG	
	case1	case2	case3	case4	case5
IL-IR→TRADO	ON	ON	ON	OFF	ON
✓ IL-IR→TRADO→ FLIP	ON	ON	ON	OFF	OFF
IL-IR→MyD88	ON	ON	ON	ON	ON
IL-IR→MyD88→NIK	ON	OFF	OFF	ON	OFF

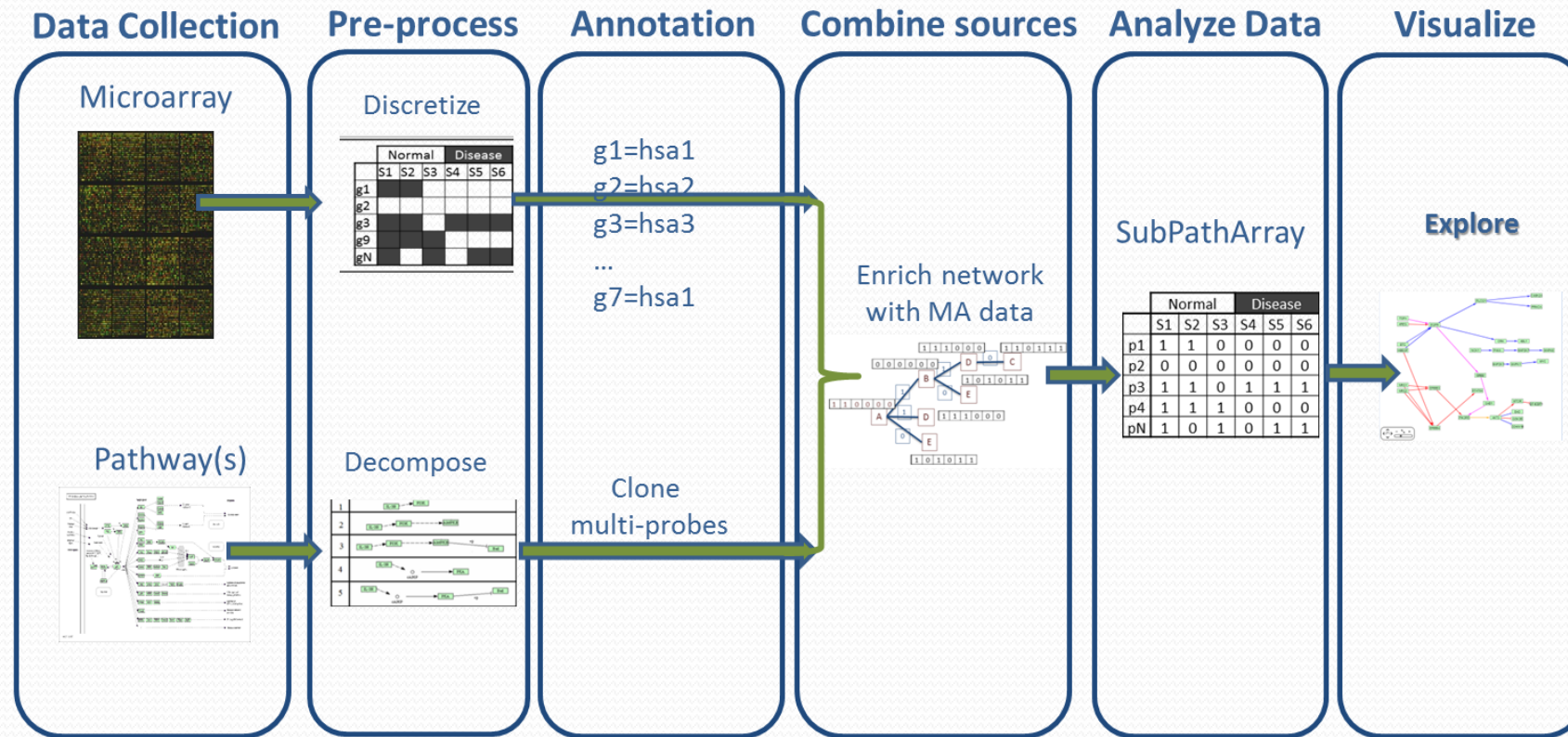
Initial expectation: microarrays would reveal specific gene signatures for various phenotypes

But it seems to be bounded to a number of limitations mainly because of the complexity and the individual variations and heterogeneities associated with the induced gene-signatures

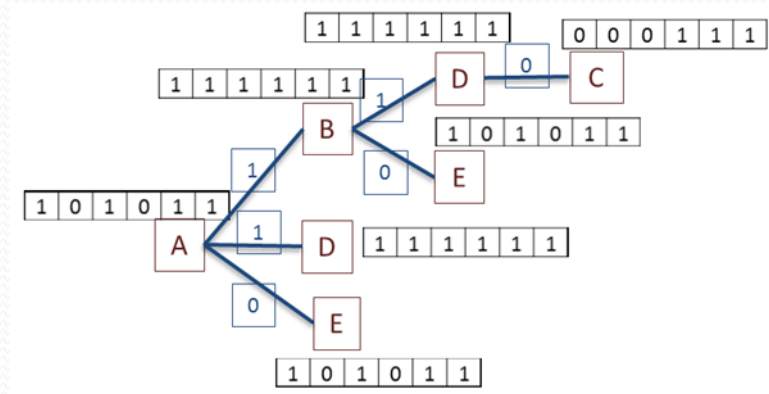
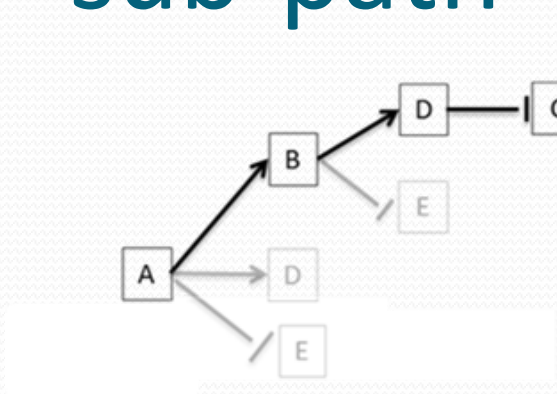
Barabási et al*: "Given the functional interdependencies between the molecular components in a human cell, a **disease is rarely a consequence of an abnormality in a single gene**, but reflects the perturbations of the complex intracellular and intercellular network that links tissue and organ systems."

* Barabási, Albert-László, Natali Gulbahce, and Joseph Loscalzo. "Network medicine: a network-based approach to human disease." *Nature Reviews Genetics* 12, no. 1 (2011): 56-68.

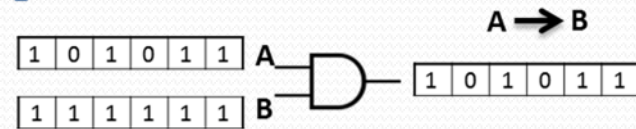
MinePath flow of operations



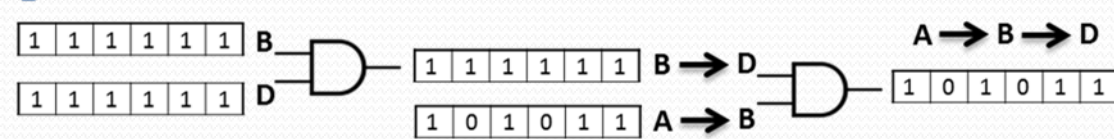
Calculating functional status of a sub-path



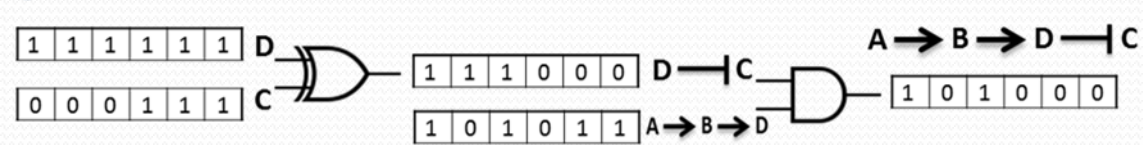
1



2



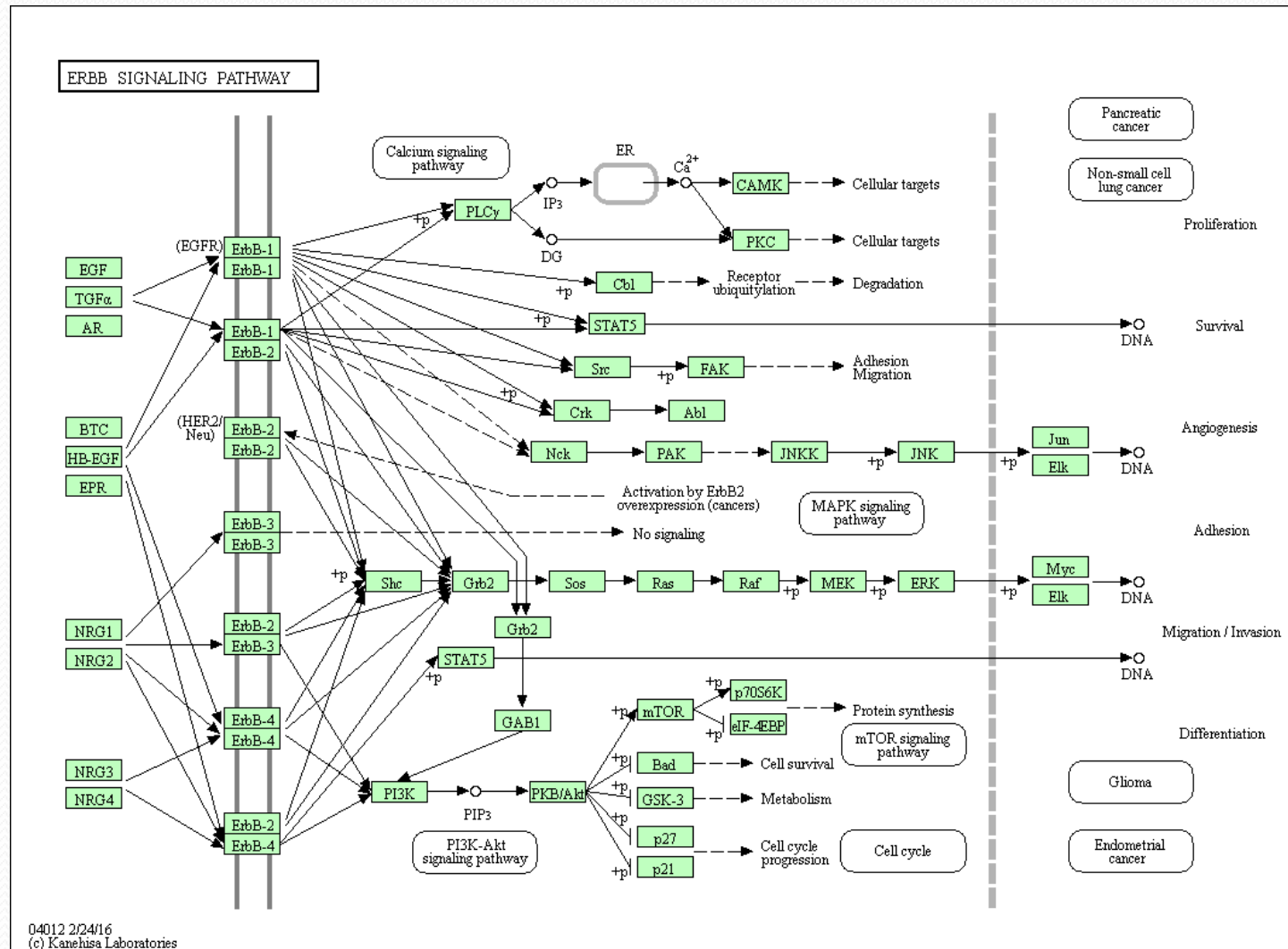
3



	Pheno-1		Pheno-2			
	S1	S2	S3	S4	S5	S6
A → B	1	0	1	0	1	1
A → B → D	1	0	1	0	1	1
A → B → D → C	1	0	1	0	0	0

The result is an array of sub-paths with binary values for every sample in the form of a discretized microarray

ErbB signaling pathway (KEGG)

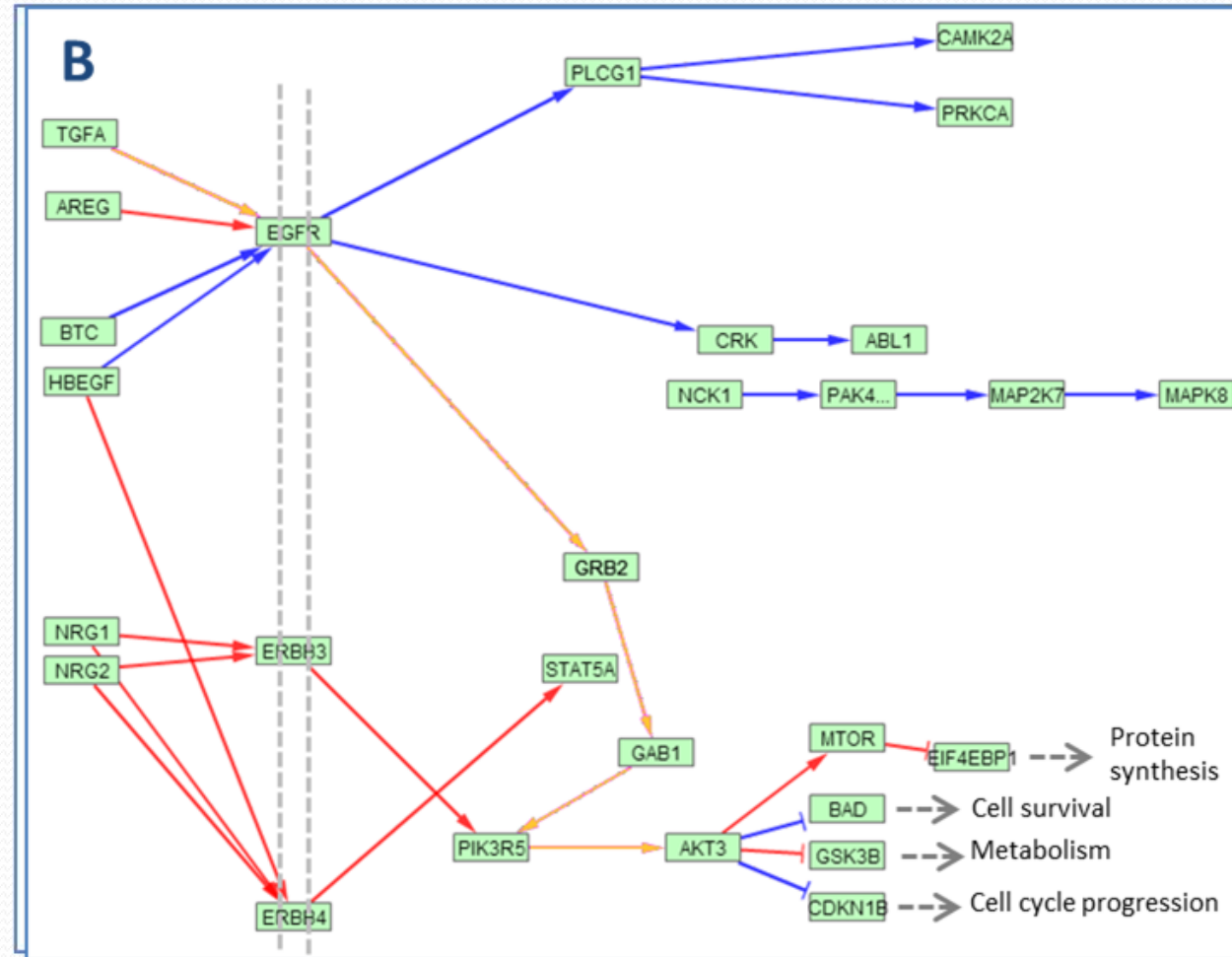


Common sub-paths

Sub-paths which are **always activated** may fill-in the gap (functional interaction) between two sub-paths and reveal a complete functional and biologically valuable route.

Colour coding:

- **Red:** sub-paths active at class 1
- **Blue:** sub-paths active at class 2
- **Orange:** sub-paths that are always active.



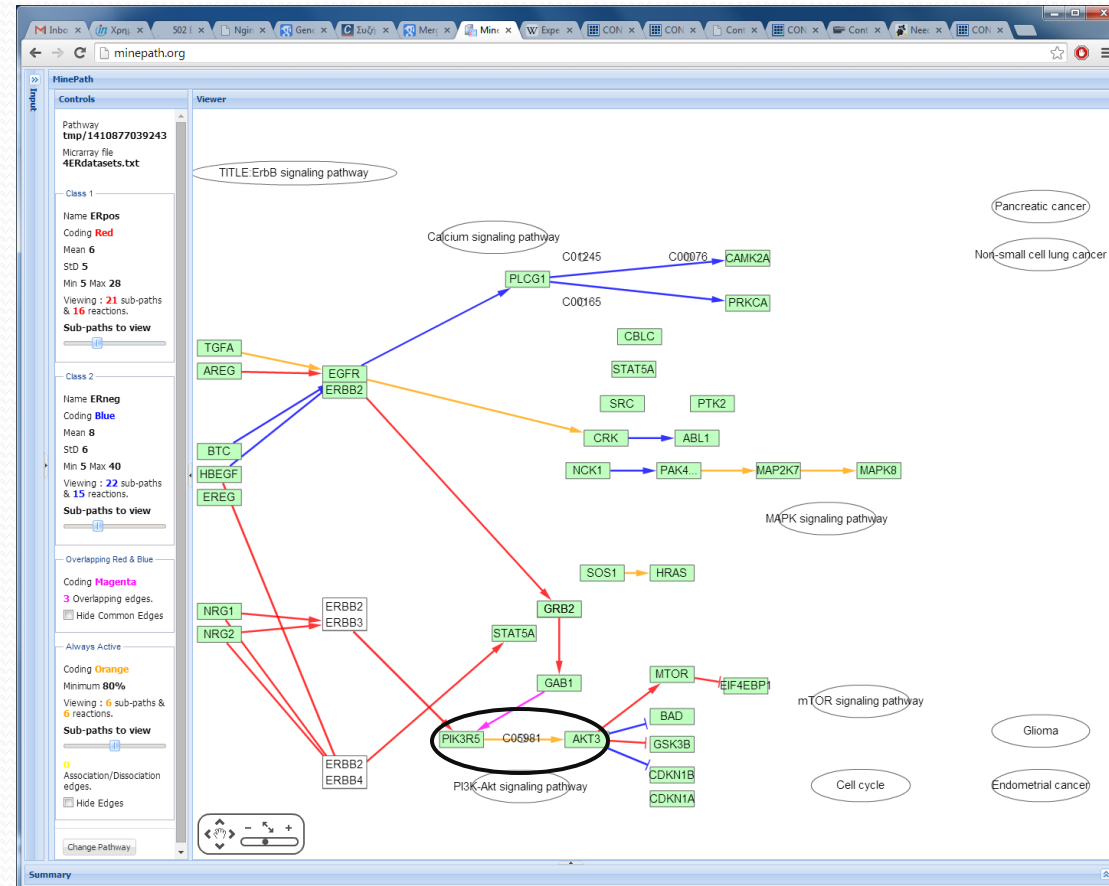
Discovery of New Biological Knowledge

Merged ER datasets & 14 cancer related pathways

1. Load/visualize ErbB signaling pathway
2. Double the thresholds (ER+ from 6 to 12, ER- from 8 to 16 & common to 90%)
3. Delete inactive genes and relations

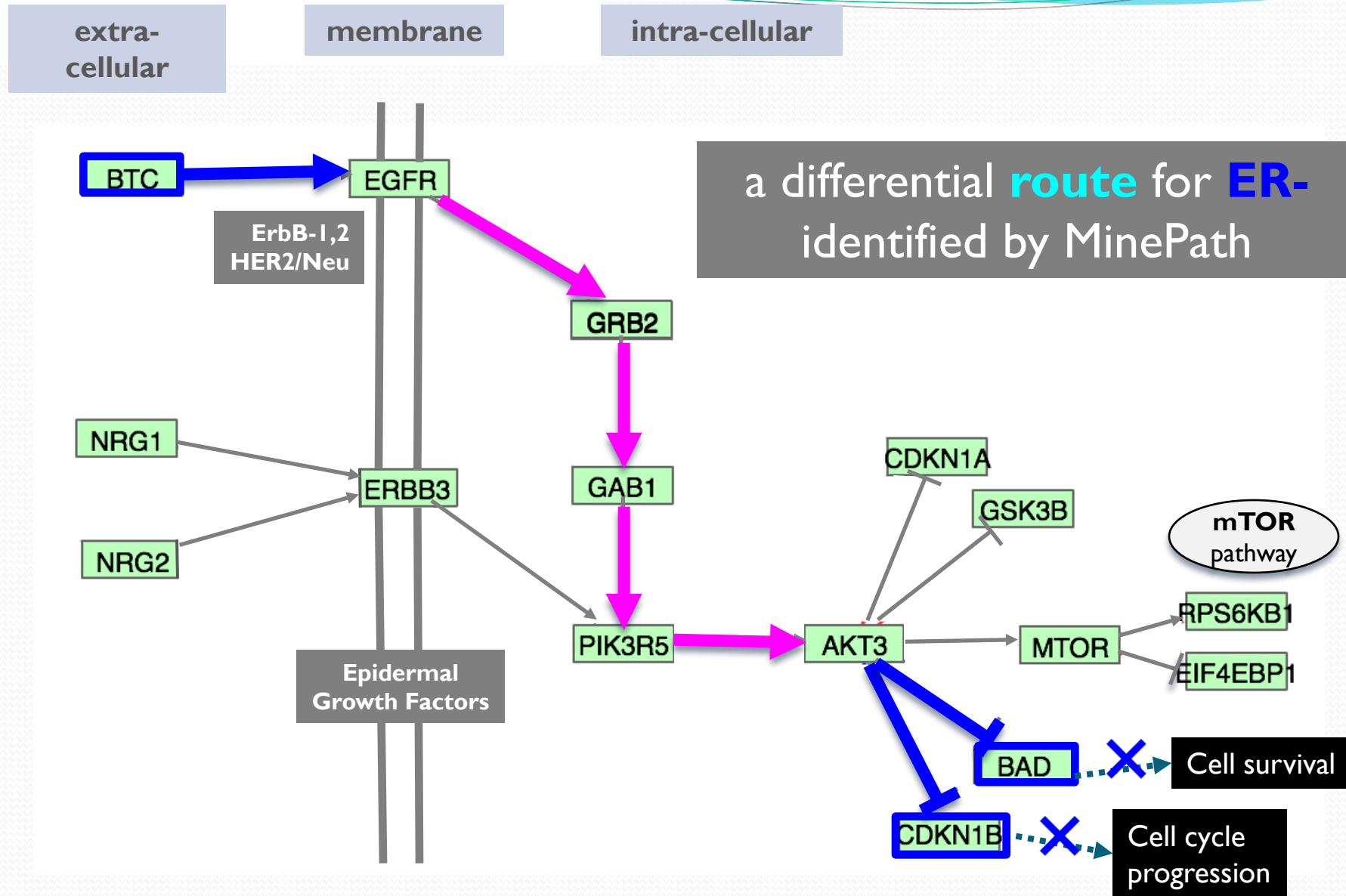
According to the literature, the results are quite relevant to the estrogen-receptor status.

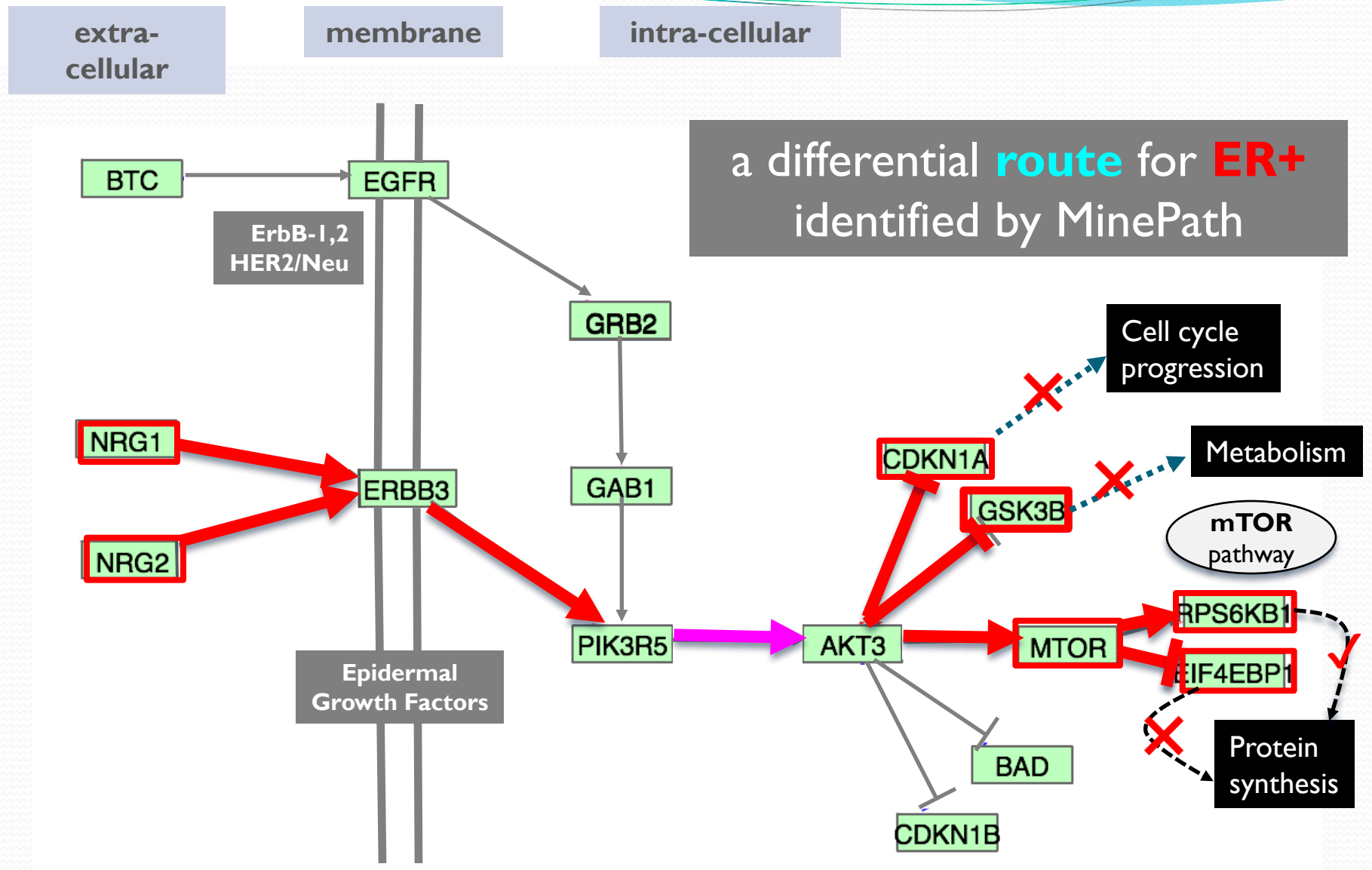
- Hutcheson et al*: “...fulvestrant treatment is sensitive to the actions of the ErbB3/4 ligand HRGβ1 (NRG1) with enhanced ErbB3/4-driven signaling activity, and significant increases in cell proliferation ...”



Exploring ErbB for the 4ERdatasets using MinePath

* Hutcheson, I.R., et al.: Heregulin beta1 drives gefitinib-resistant growth and invasion in tamoxifen-resistant MCF-7 breast cancer cells. *Breast Cancer Research* 9(4):R50, (2007)





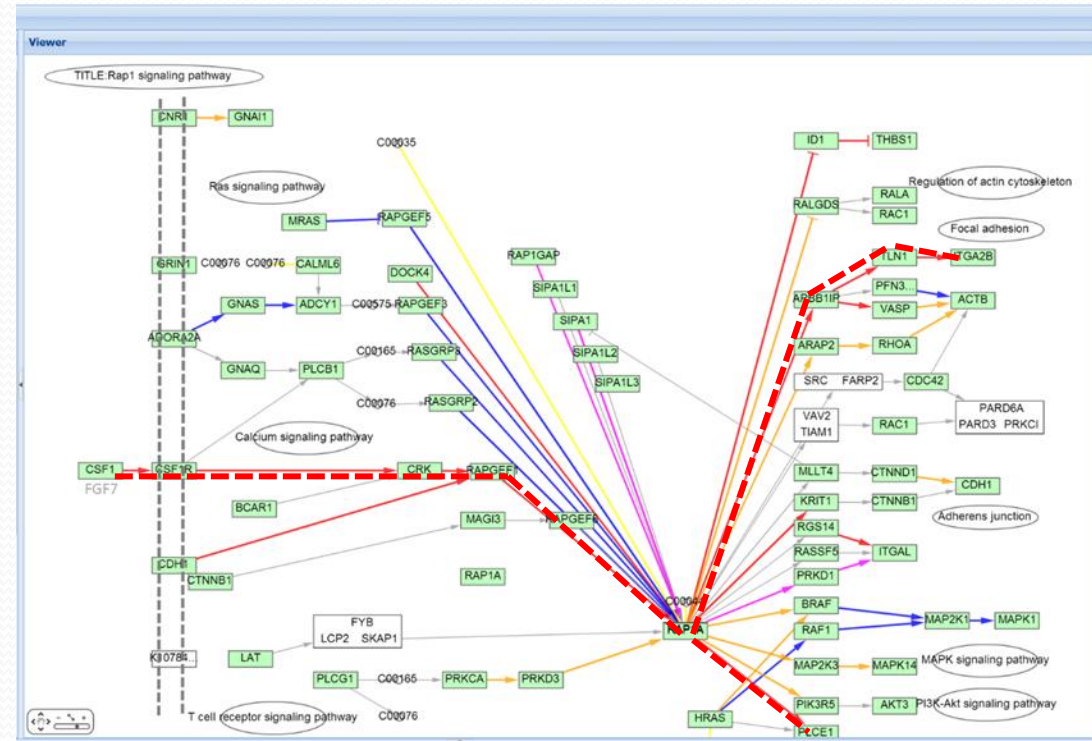
MinePath Biological Validation

Craniosynostosis (GSE27976*)

- 199 patients compared against a control population (n = 50)

MinePath identified **Rap1** signaling pathway as one of the most discriminant pathways and the most informative for Synostosis:

- **CSF1**→**CSF1R**→**CRK**→**RAPGEF1**→**RAP1A**→**APBB1P**→**TLN1**→**ITGA2B** leading to Focal Adhesion
 - Stamper et al* :
 - **FGF7/CSF1** (the most discriminant gene)
 - Focal adhesion pathway (the most discriminant pathway)
- **CSF1**→**CSF1R**→**CRK**→**RAPGEF1**→**RAP1A**→**PLCE1** leading to the PI3K-Akt signaling pathway.
 - Dufour et al** identified that PI3K/Akt attenuation plays important role in the control of osteoblast survival by FGFR2 signaling (member of the fibroblast growth factor FGFR family).



* Stamper, Brendan David, et al. "Transcriptome correlation analysis identifies two unique craniosynostosis subtypes associated with IRS1 activation." *Physiological genomics* 44.23 (2012): 1154-1163

** Dufour, Cécilie, et al. "FGFR2-Cbl interaction in lipid rafts triggers attenuation of PI3K/Akt signaling and osteoblast survival." *Bone* 42.6 (2008): 1032-1039.