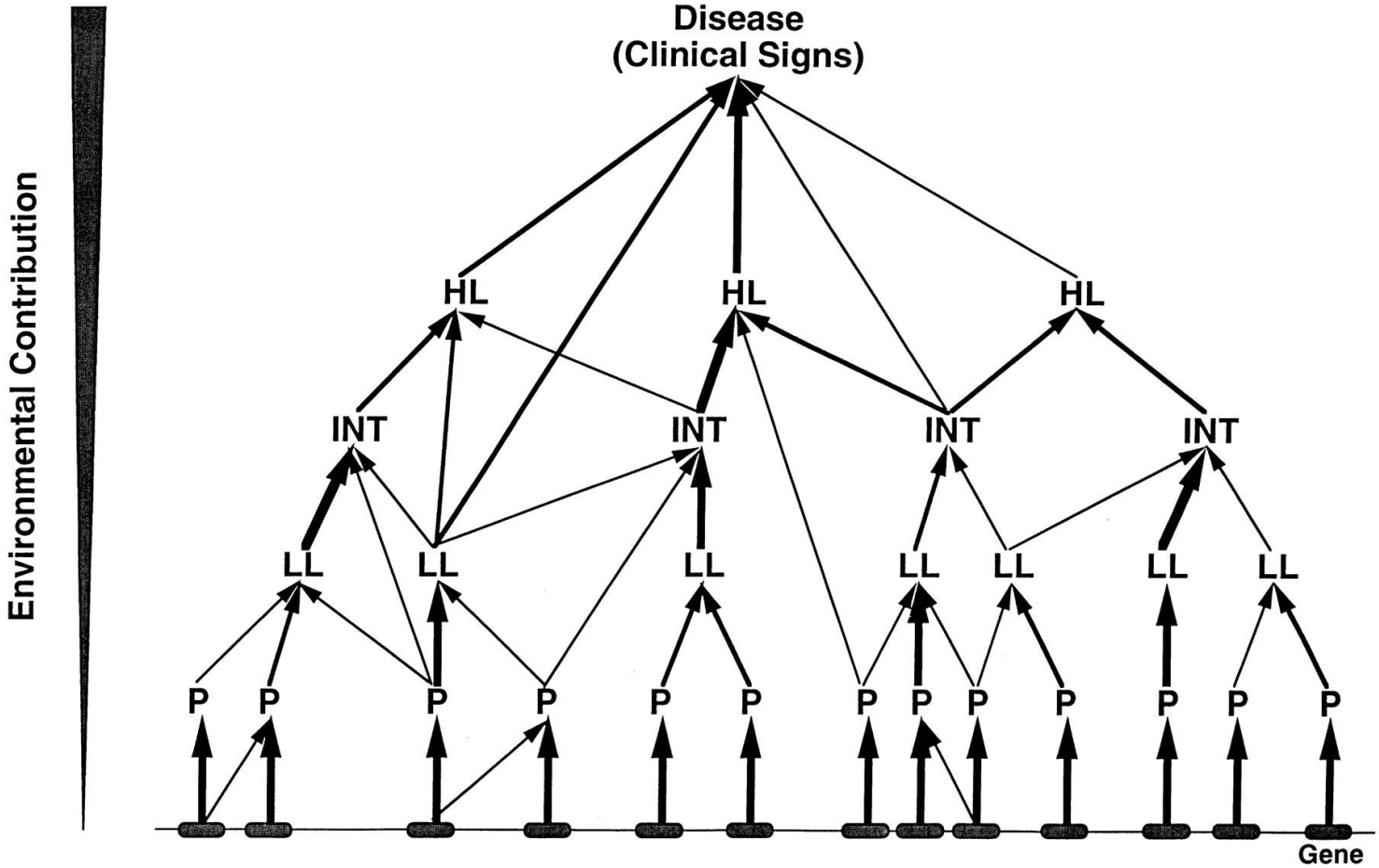
Fenomika

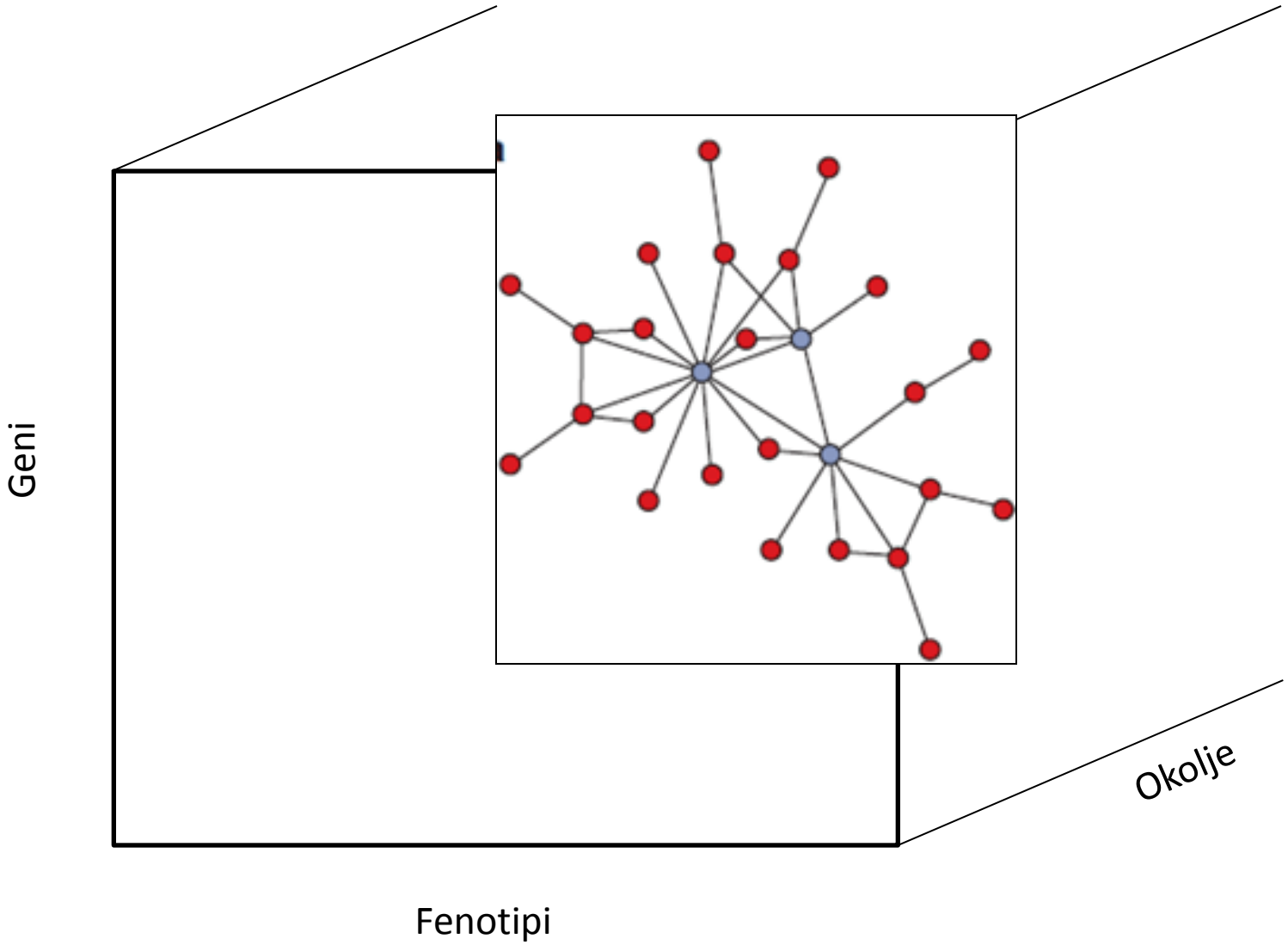
All about Eve (well, Adam too)

Nicholas J. Shork (1997) Genetics of complex disease: approaches, problems, and solutions. *Am J Respir Crit Care* 156:S10

'In addition, a new discipline, "phenomics" or "phenometrics," could be initiated that would complement genomic research as presently performed.'



Pleiotropnost



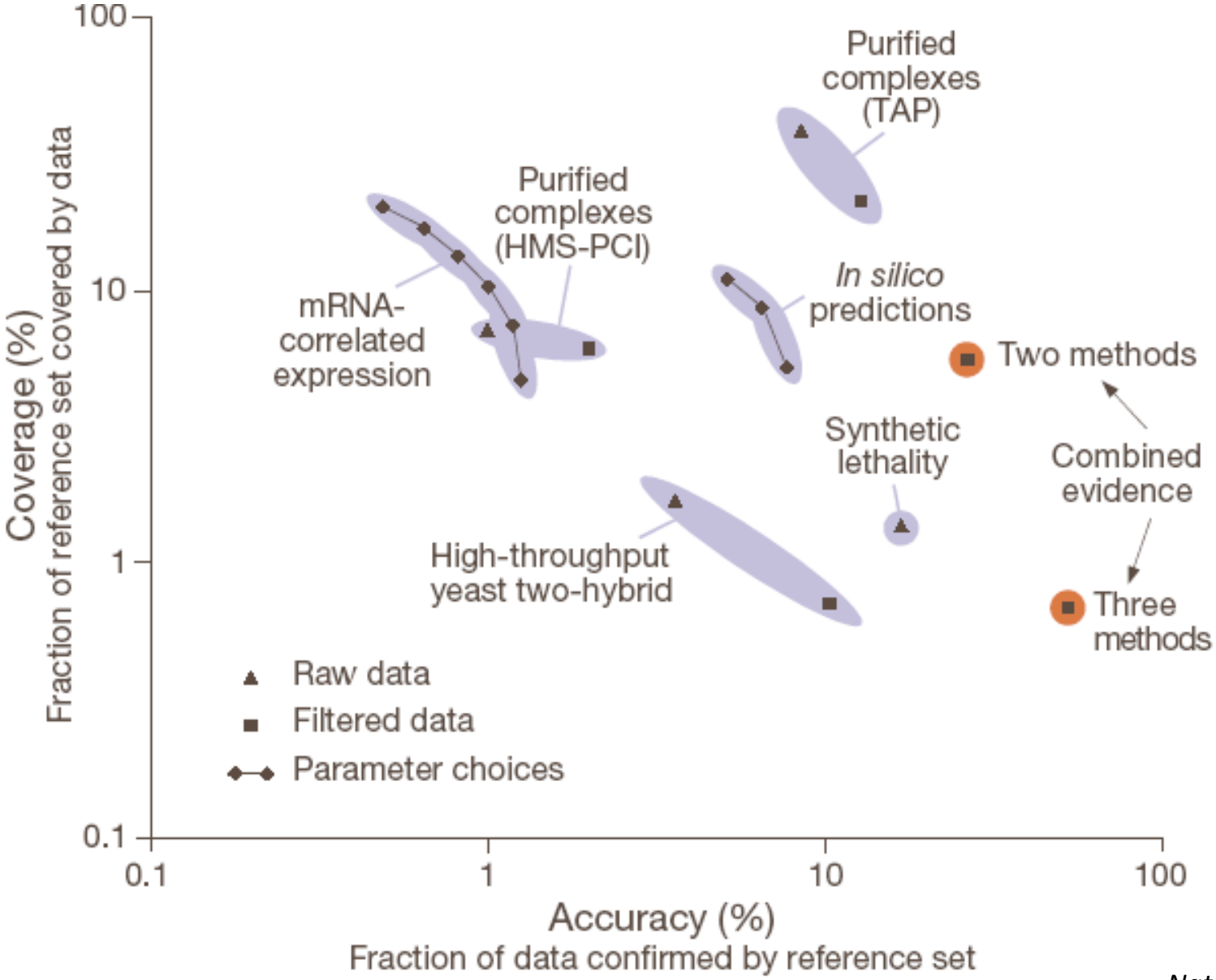
Fenetika & **Fenomika**

Ena definicij fenomike: "**the acquisition of high-dimensional phenotypic data on an organism-wide scale**", ali
pridobivanje visokodimenzionalnih podatkov na ravni celotnega organizma;
Houle D, Govindaraju DR, Omholt S. (2010) Nat Rev Genet. 11:855

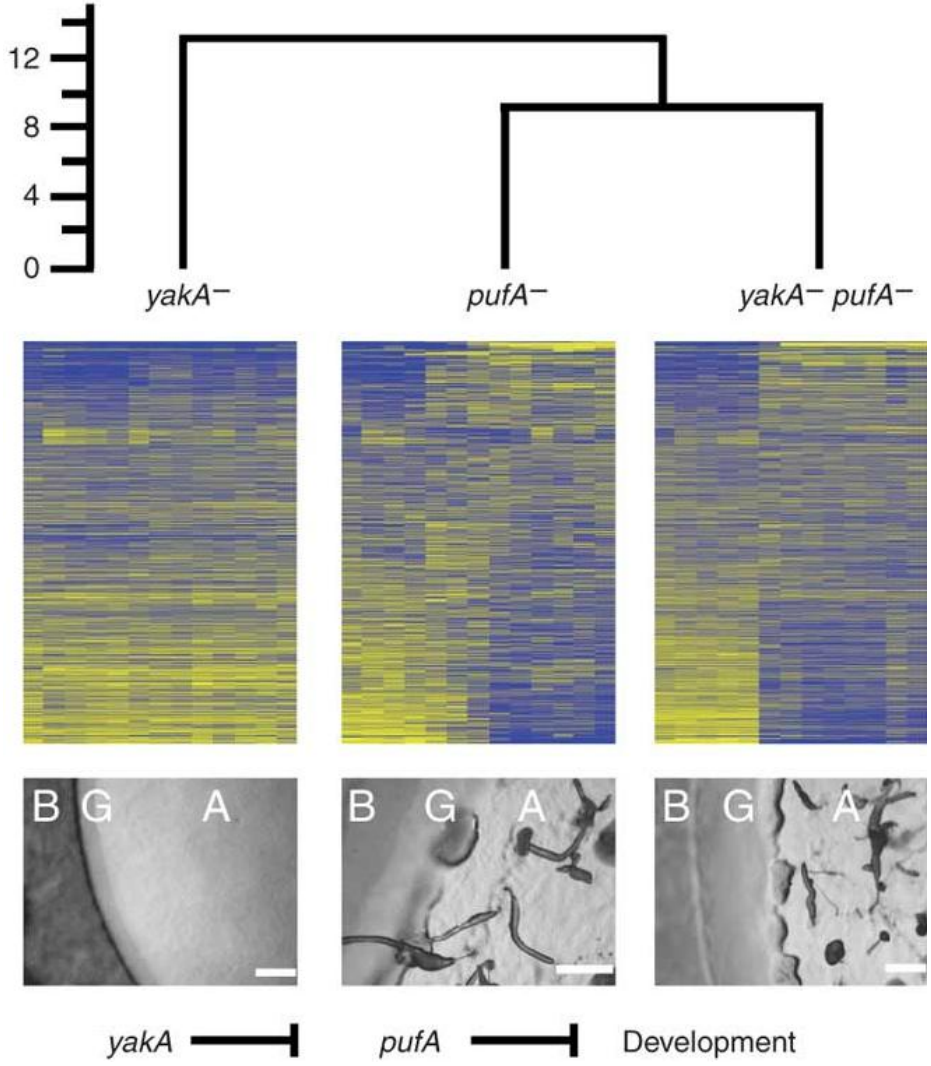
Opcija: merjenje specifične lastnosti pri organizmu, spremenjenem s perturbacijo na ravni celotnega genoma

ekstenzivna vs. **intenzivna** fenomika
eks.: velik nabor fenotipov
int.: velika globina posameznega fenotipa

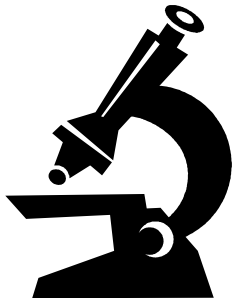
Različni tipi podatkov



Nature (2002) 417:399



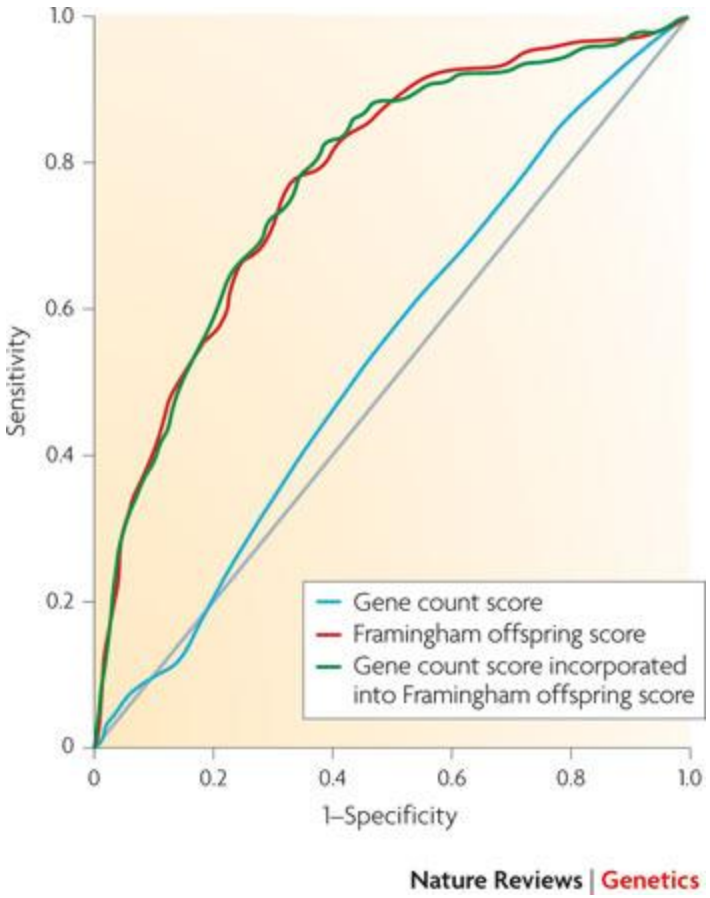
“Mikroskop molekularne biologije”

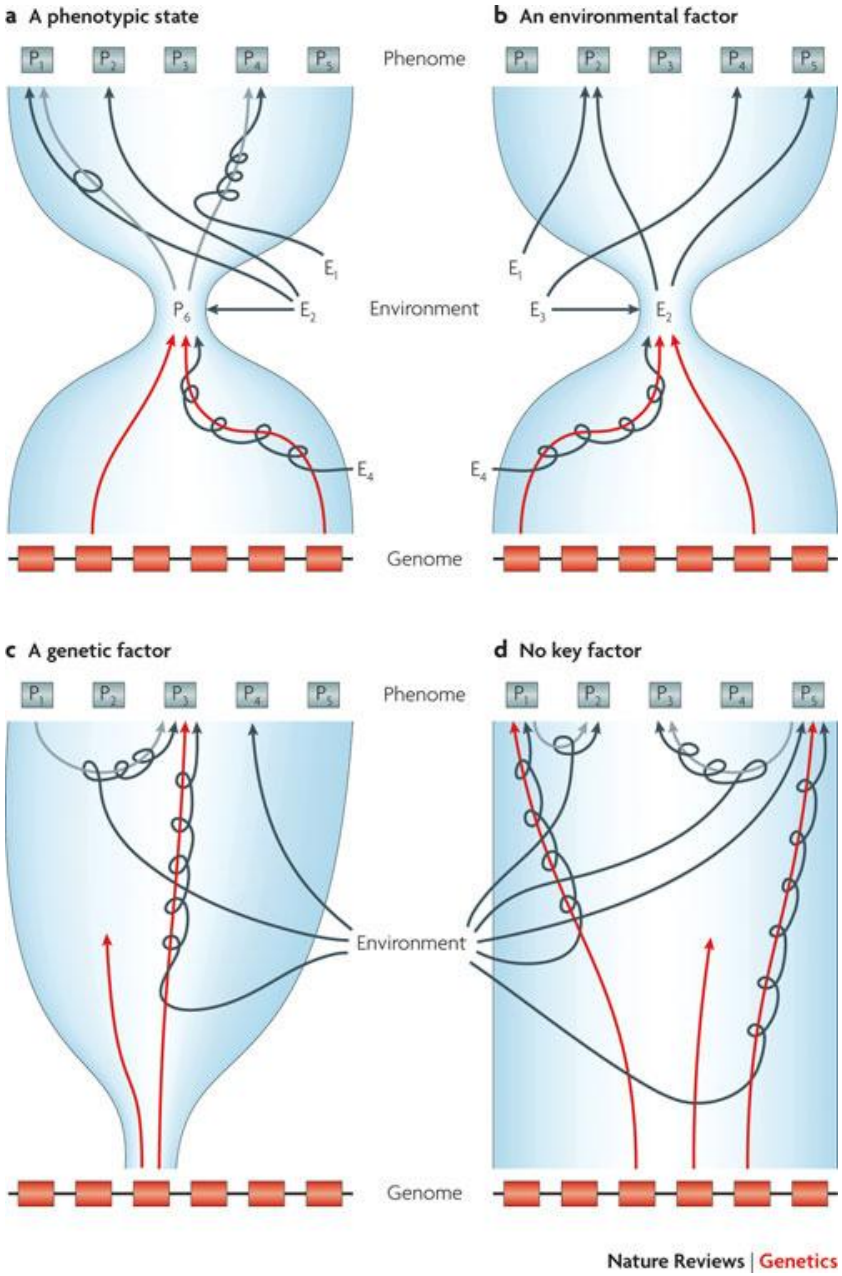


Van Driessche N, Demšar J, Booth EO, Hill P, Juvan P, Zupan B, Kuspa A, Shaulsky G. (2005) *Nature Genetics* 37:471

Ali lahko fenotipe napovedujemo?

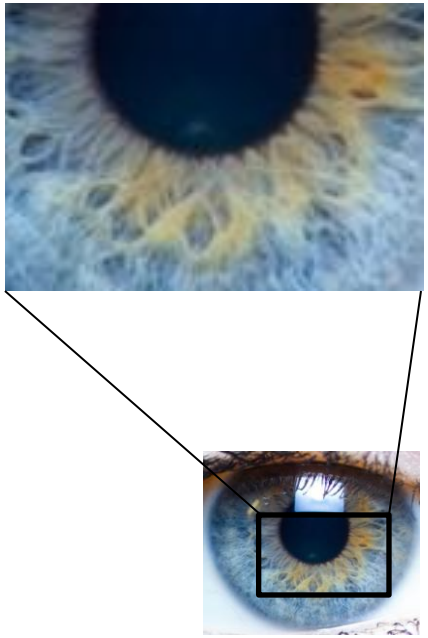
Če da, na osnovi česa?

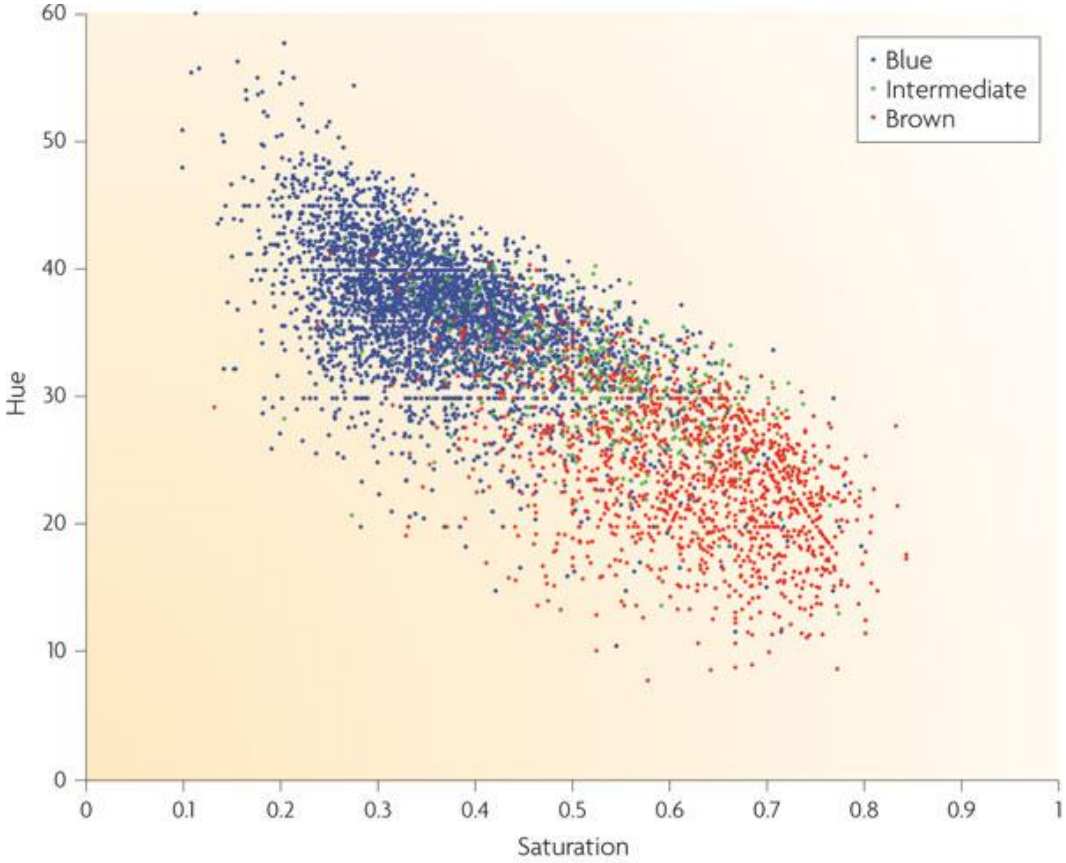




Nat Rev Genet (2010) 11:855

Diskretni in zvezni fenotipi





Nature Reviews | Genetics

Nat Rev Genet (2010) 11:855

Fenotipizacija preko bioloških hierarhij

Level	Extensive	Intensive
DNA, RNA	Solved and costs declining rapidly. Limited by bioinformatics capabilities	Possible but costly; detailed sampling is challenging
Chromatin	Made possible by chromatin immunoprecipitation but with low precision; costly	Possible but costly; detailed sampling is challenging owing to instability
Proteins, metabolites	Mostly solved and capabilities are still improving for rare constituents	Possible but costly; detailed sampling is challenging
Cells	Huge diversity of assays is possible but most are low throughput. Image-based techniques for high throughput are promising	Many temporally continuous assays. Image-based approaches allow some combination of extensive and intensive measurement
Development	Low-throughput measures are abundant. High-throughput image-based methods are possible for small living individuals (embryos and *Caenorhabditis elegans*) and sectioned tissues	Temporal depth by repeated sampling or image-based approaches
Physiology	Huge diversity of assays is possible, but most are low throughput, unless based on proteomic or metabolomic data	Many temporally continuous assays. Spatial sampling is often possible
Morphology	Solved in principle with the use of imaging, but assays often require extensive sample preparation. Post-processing to extract and measure features requires specialized informatics capability	Intensive sampling of morphological form is possible with specialized processing. Temporal depth is limited by destructive sampling
Behaviour	Possible with continuous observation of video of confined or local populations, or for humans, self-reporting. Data extraction from video using specialized software or human labour	Extensive sampling gives intensive coverage. Specific aspects of behaviour can be intensively measured with data loggers

Nat Rev Genet (2010) 11:855

Časovni razpon: ~15 velikostnih razredov
Prostorski razpon: ~9 velikostnih razredov

Podatkovne baze za fenotipe

Genotypic and phenotypic databases	Concepts				Relationships					
					Binary				Ternary	
	Gene	Molecular Class	Phenotype	Disease	Gene-Gene	Gene-Molecular Class	Gene-Phenotype	Gene-Disease	Gene-Molecular Class –Phenotype	Gene-Molecular Class-Disease
Curated Phenomic Databases										
MGI (12)	●	●	●			■	■		■	
OMIM (3, 69)	●		⊙	⊙			◊	◊		
OMIA (55, 71)	●		●	●			□	□		
Other model organism resources (12, 13, 55, 71–76)	●	●	●			□	□		□	□
Computationally-derived phenomic databases										
www.PhenoGO.org (35)	●	●	●	●	■	■	■	■	■	■
PhenomicDB (2)	●		⊙	⊙			◊	◊		
Genetic Association Database (77)	●		○	○						◊
GeneCards (59–61)	●					□		◊		
Phenomic analyses										
QMR-OMIM (63)	●		●	●			■	■		
Mining OMIM (78)										
GenesTrace (64)	●	●	●	●		■	■	■	■	■
UMLS-GEO Network (65)	●		●	●			■	■		
Literature-based techniques (19, 57, 62, 79, 80)	●			⊙	■			◊		□
Rank-based integration / fusion (70, 81)	●	●				■		◊		◊

Proc Am Thorac Soc. (2007) 4:18

Minimum Information Standards

From Wikipedia, the free encyclopedia

Contents [hide]

Enter term(s) or part of a term:

How to search

Example: apoptosis

in these sections:

All

Use wildcard (If unselected wildcard can be used with *)

Search Clear

Select Organisms:

All
Caenorhabditis elegans
Dictyostelium discoideum
Fruit fly
Human
Mouse
Yeast
Zebrafish

Restrict query to:

- genotypes
- phenotypes
- only genotypes with associated phenotypes
- only phenotypes with associated genotypes
- no restriction

Select data fields to show:

Default
Official gene symbol
Gene description
External genotype ID
NCBI Gene ID
mRNA ID
GO term
GO ID

PhenomicDB is a multi-organism phenotype-genotype database including human, mouse, fruit fly, C.elegans, and other model organisms.
The inclusion of gene indeces (NCBI Gene) and orthologues (same gene in different organisms) from HomoloGene allows to compare phenotypes of a given gene over many organisms simultaneously.
PhenomicDB contains data from publicly available primary databases: FlyBase, Flyrnai.org, WormBase, Phenobank, CYGD, MatDB, OMIM, MGI, ZFIN, SGD, DictyBase, NCBI Gene, and HomoloGene.

View cluster statistics

Statistics

Organism	Phenotypes	Genotypes	
		all	with orthologues
Caenorhabditis elegans	130777	49436	8678
Dictyostelium discoideum	11230	13893	0
Fruit fly	154223	34679	9263
Human	8944	42562	18832
Mouse	29444	86976	21470
Yeast	80846	7218	4849
Zebrafish	5736	47742	18668
Other	21985	8068897	154383

How to link to PhenomicDB using e.g. RefSeq identifiers
PhenomicDB in Science: From Gene To Flesh,
Citing PhenomicDB

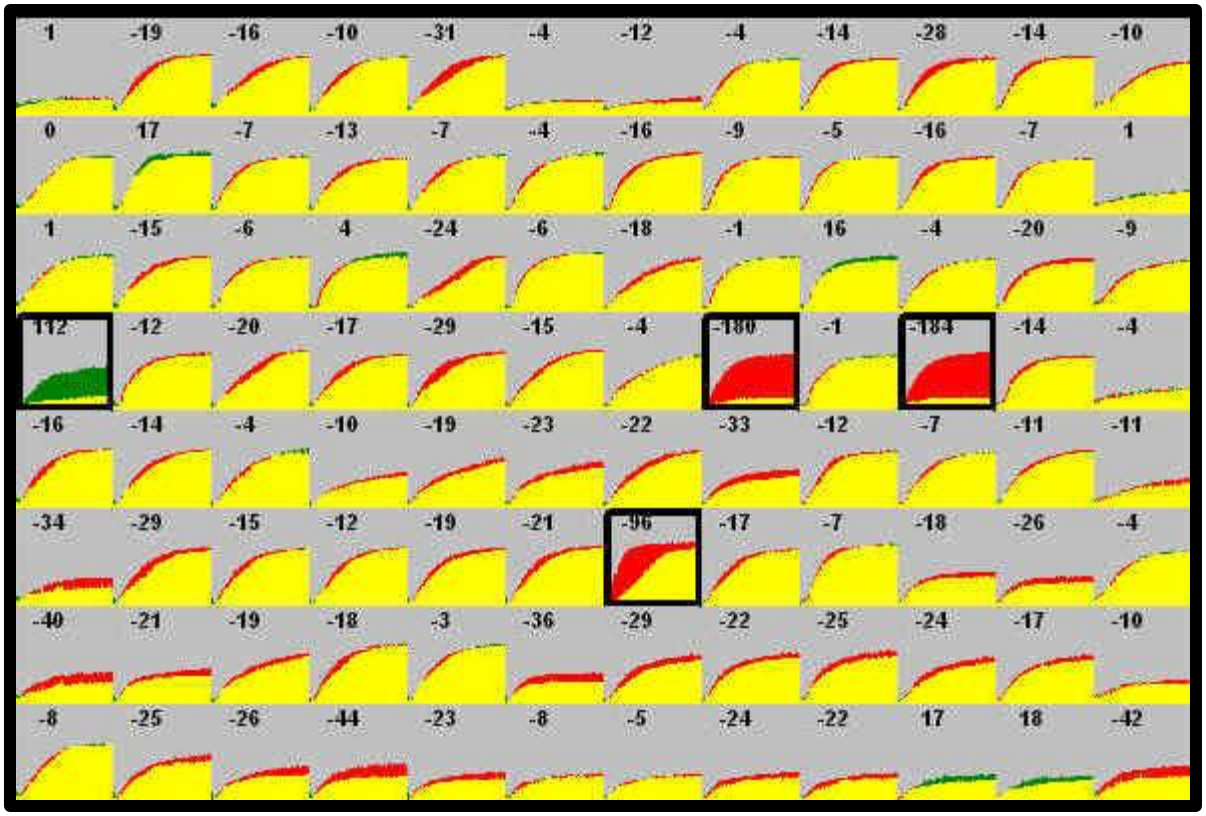
Biobar features PhenomicDB
GET BIOBAR

Last Updated: 6 February 2012; freeze date: 14 November 2011
PhenomicDB Version 3.5 Database release notes

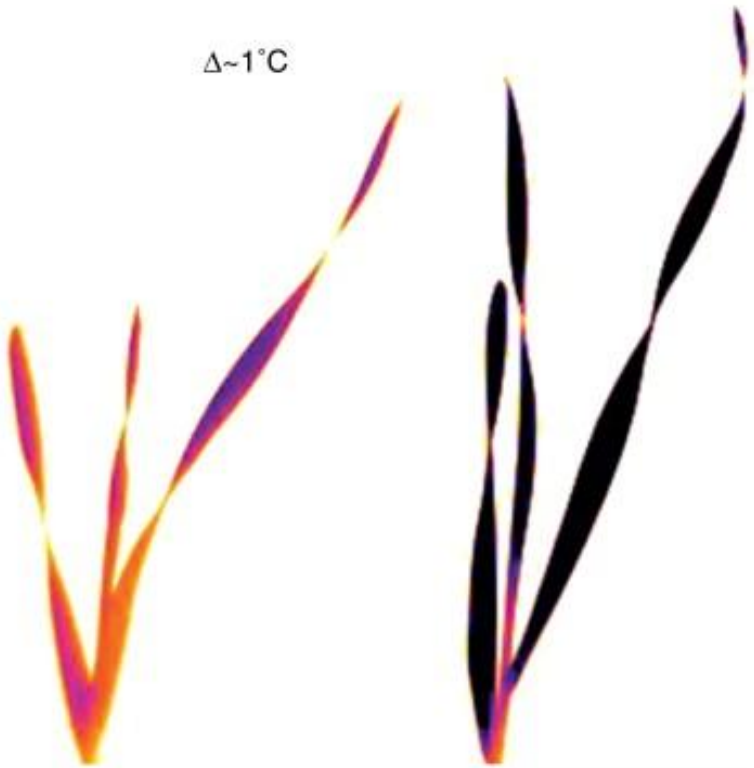
site map | legal

Analiza posameznih celic vs. analiza populacije

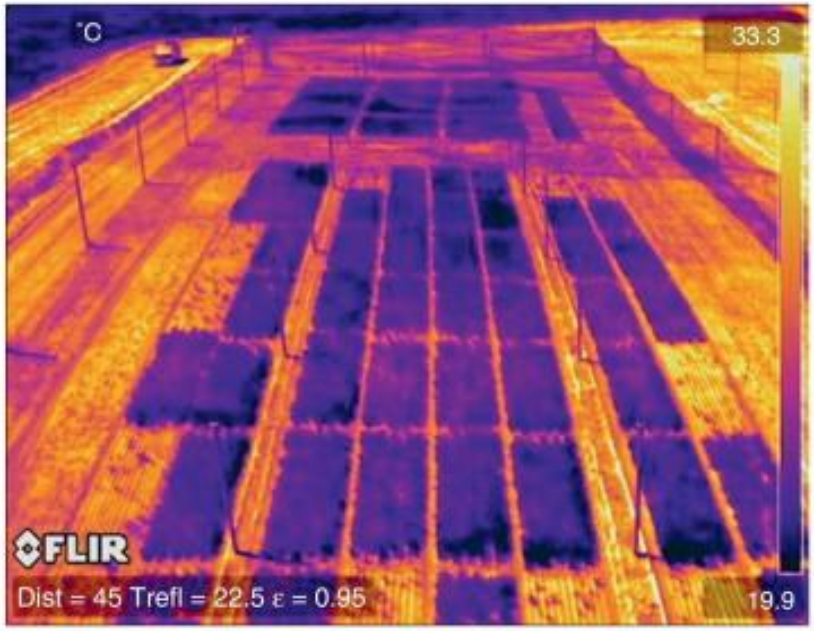
Phenotype microarray
(Fenotipska ‘mikromreža’)



Abiotski stres – termalna fotografija

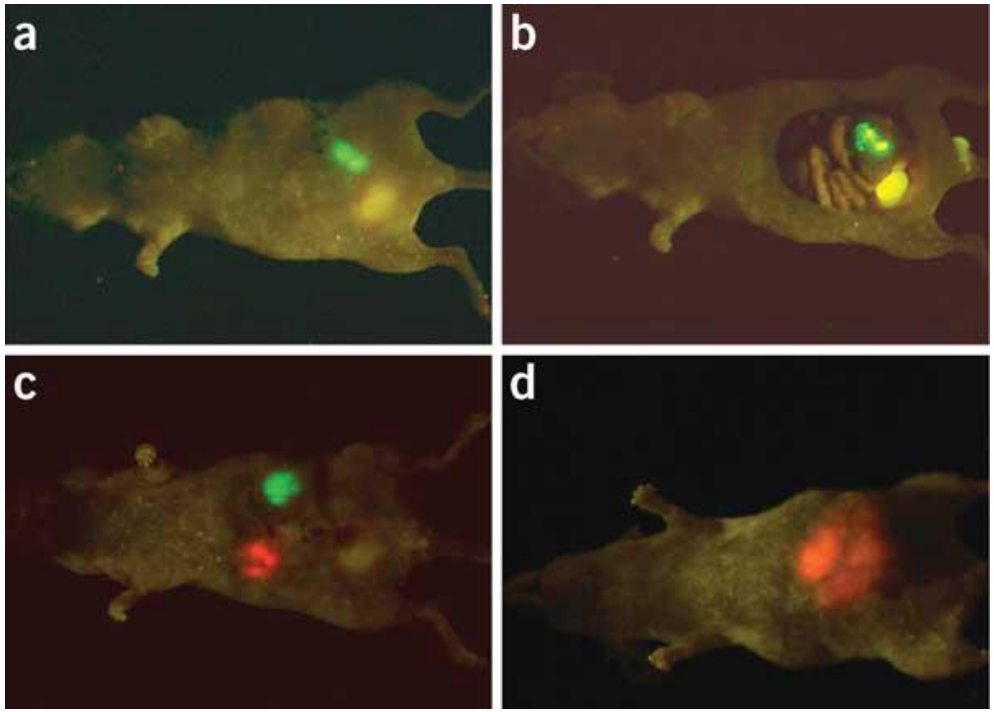


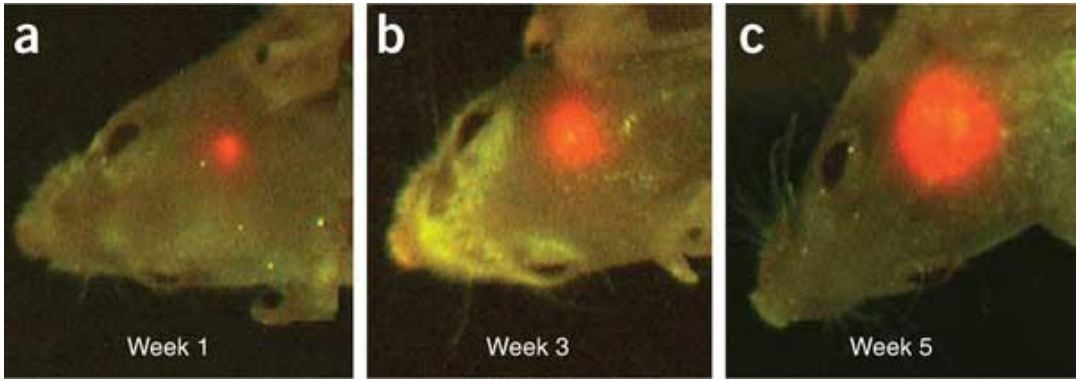
TRENDS in Plant Science

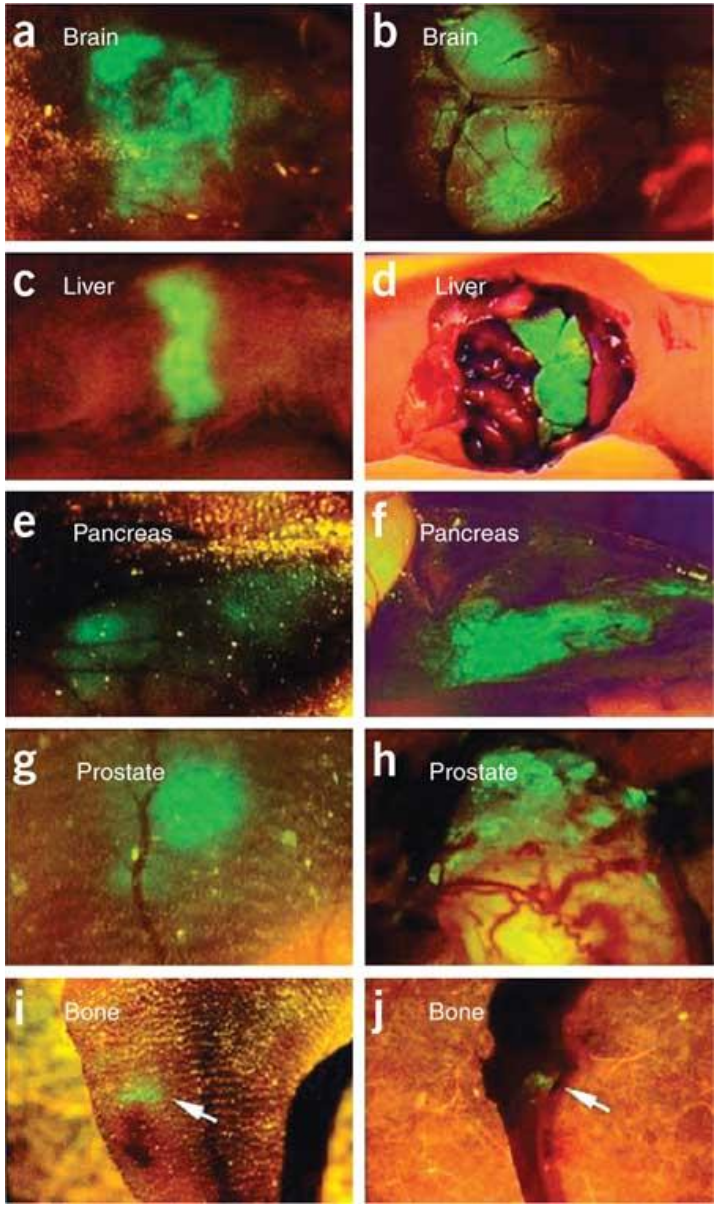


TRENDS in Plant Science

Trends in Plant Science (2011) 16:635

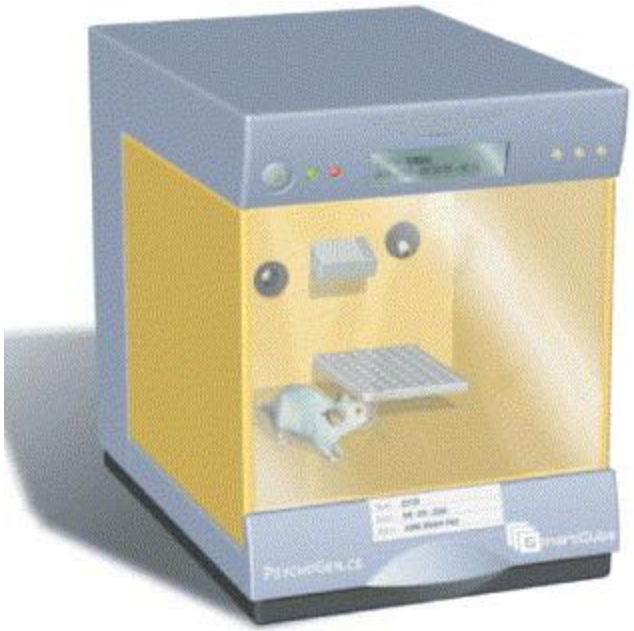






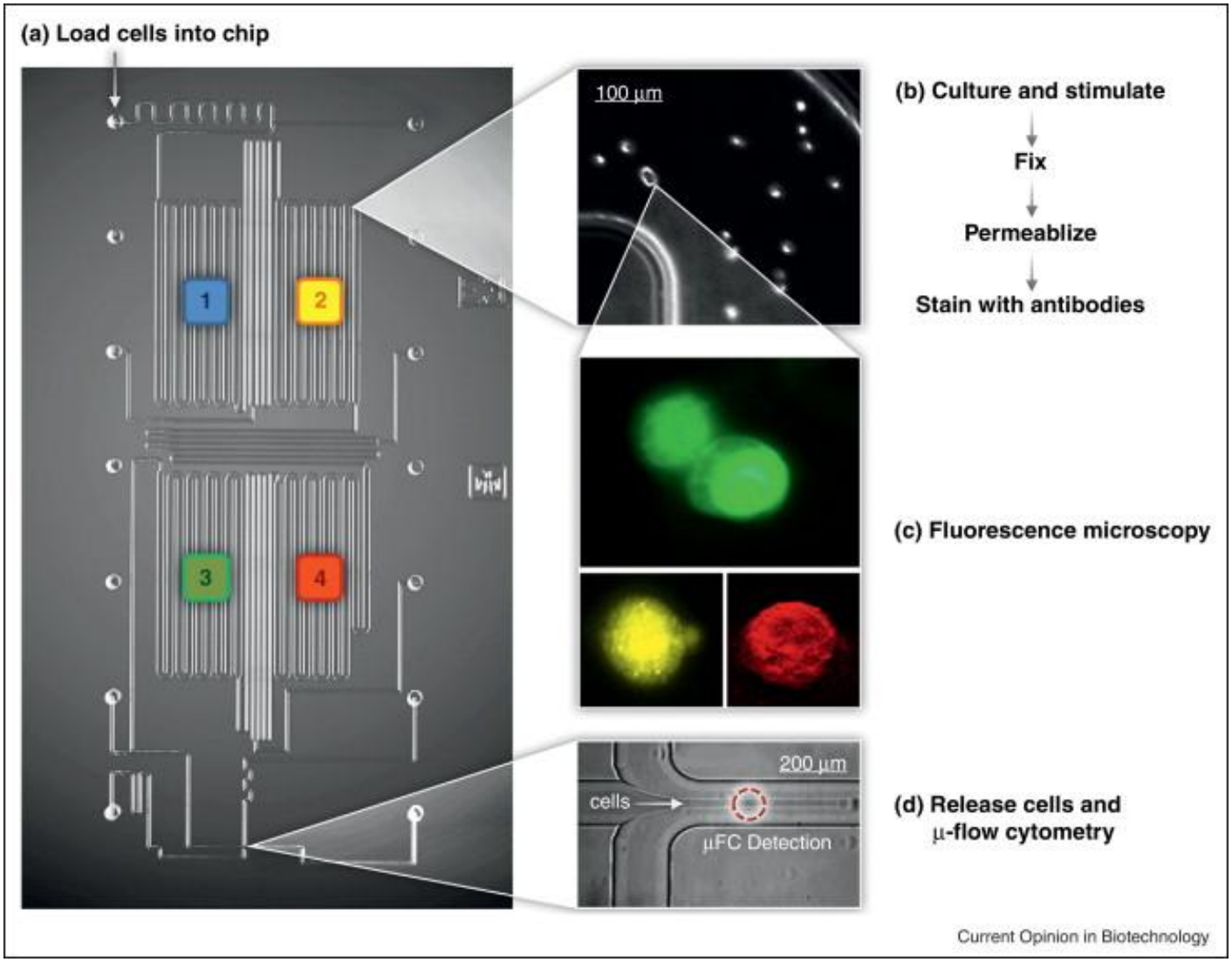
Whole-body imaging with fluorescent proteins. *Nature Protocols* (2006) 1:1429

Sistematično testiranje obnašanja



SmartCube™ (PsychoGenics, Inc.)

Mikrotekočinski sistemi



Curr Opin Biotech (2012) 23:83

Analiza podatkov

- Analiza slik (*Image analysis*)

- Problem LPSN (*large p, small N*)

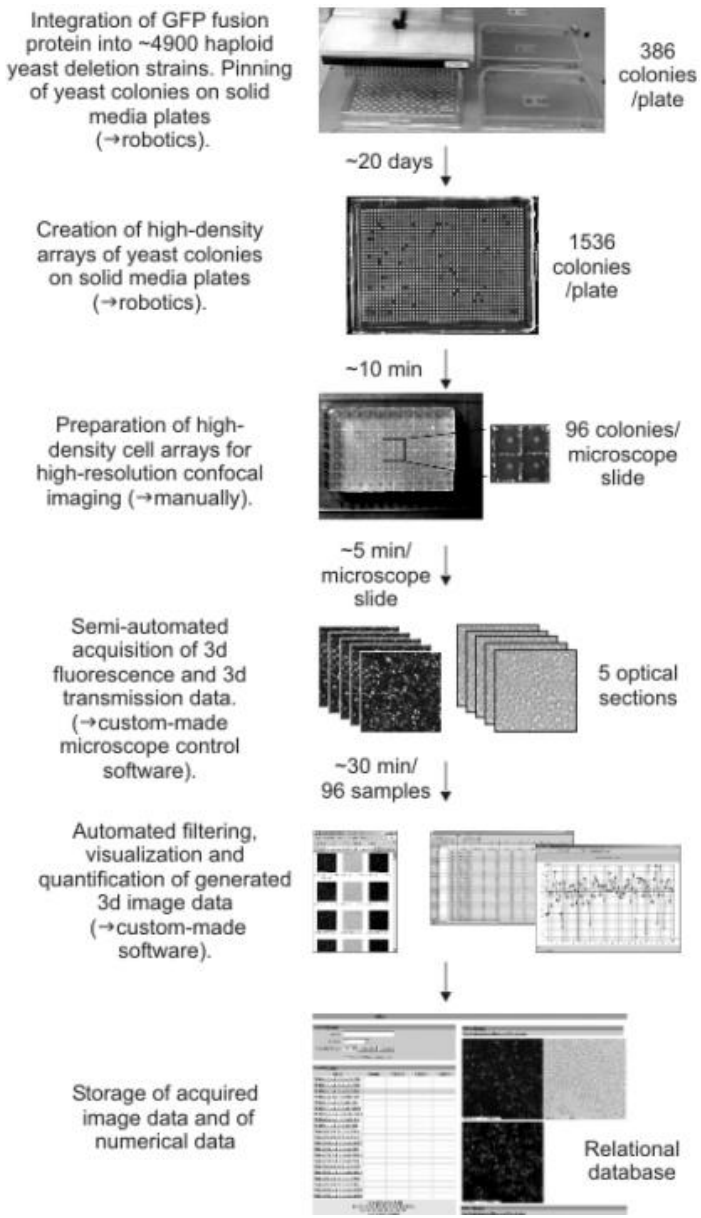
↓

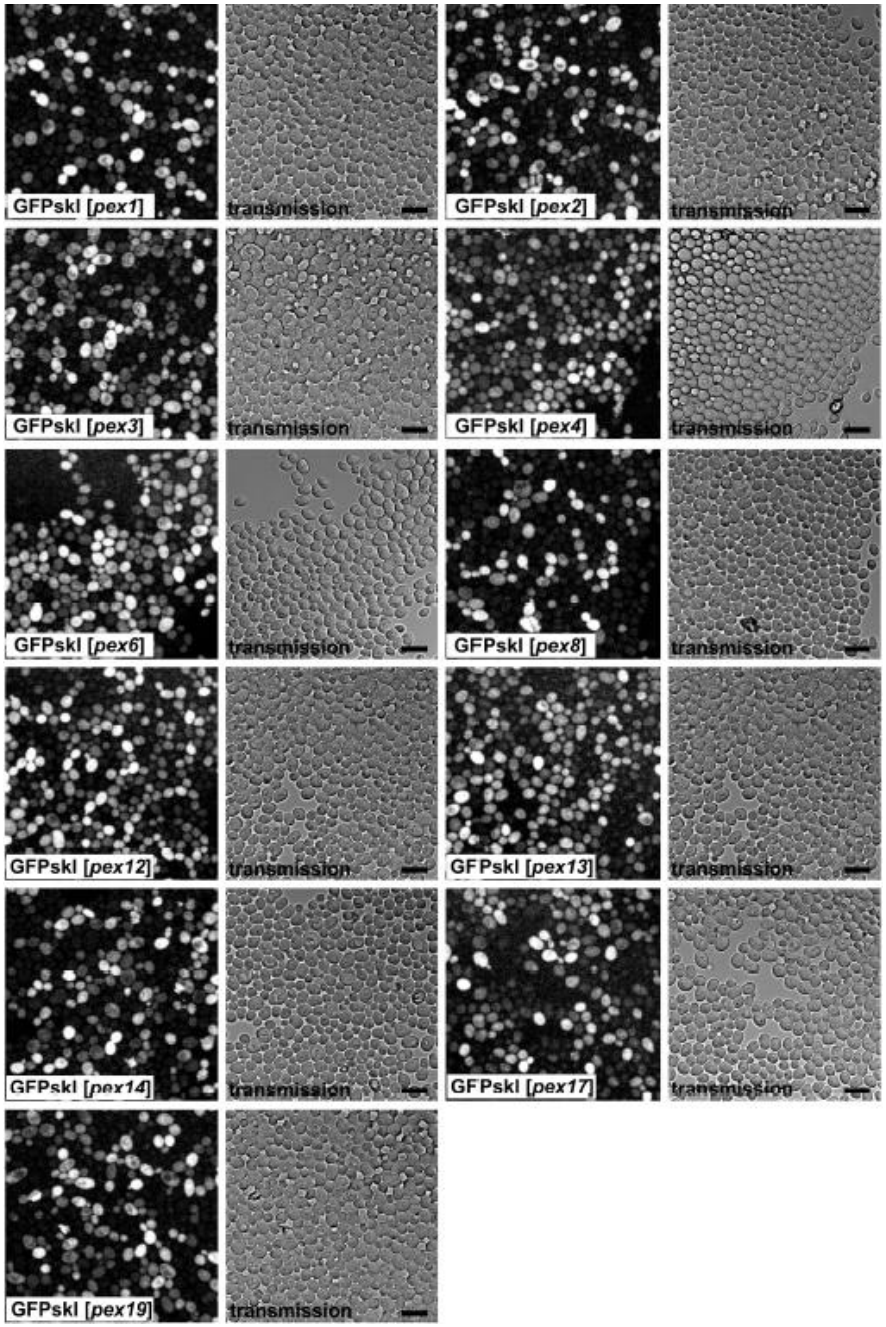
Statistične metode
(*SVM, Random forest, Partial least-square regression*)

→ **“A imate še kaj drugega na zalogi?”**

cG-P modeli

HCM - natančnost

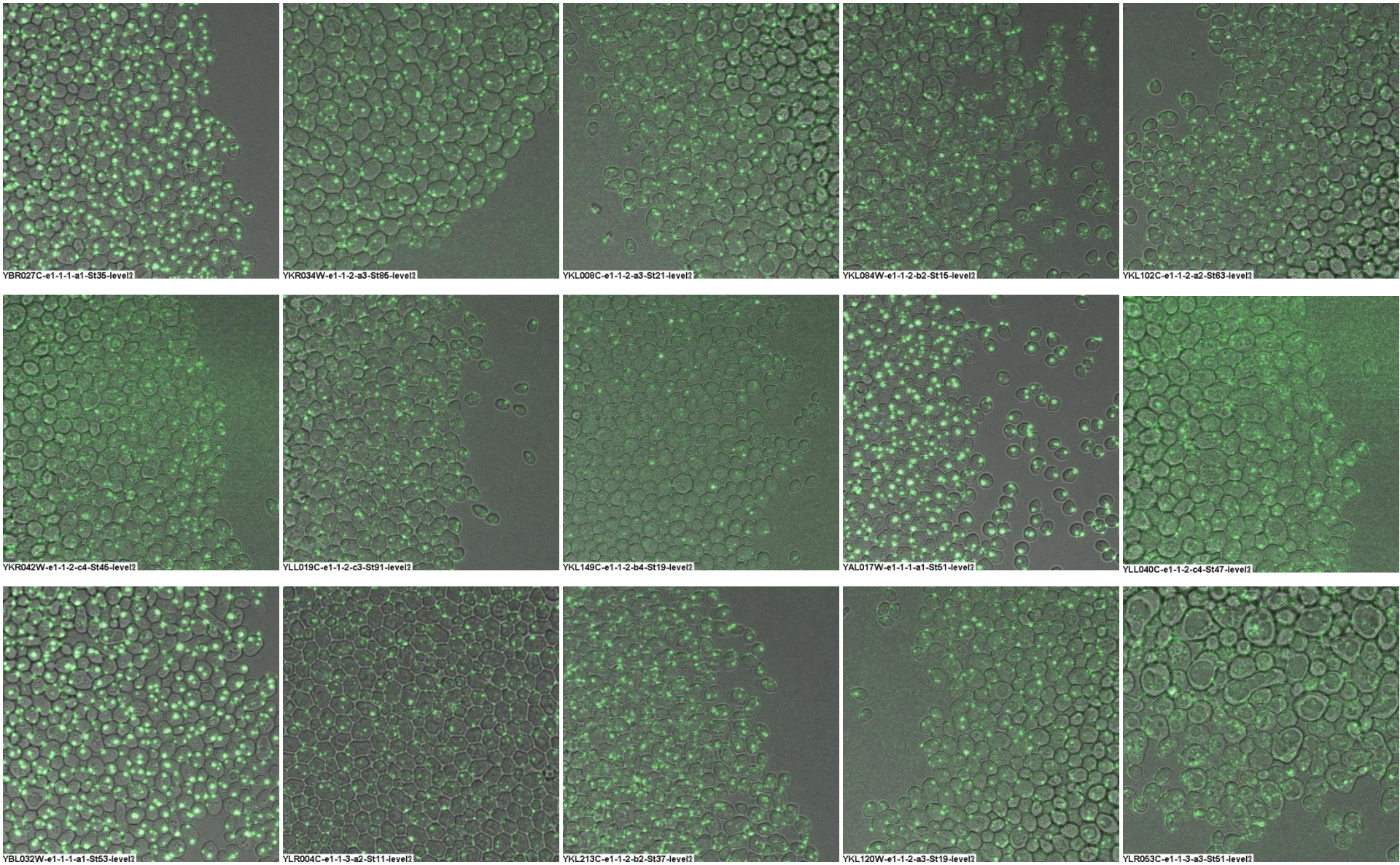


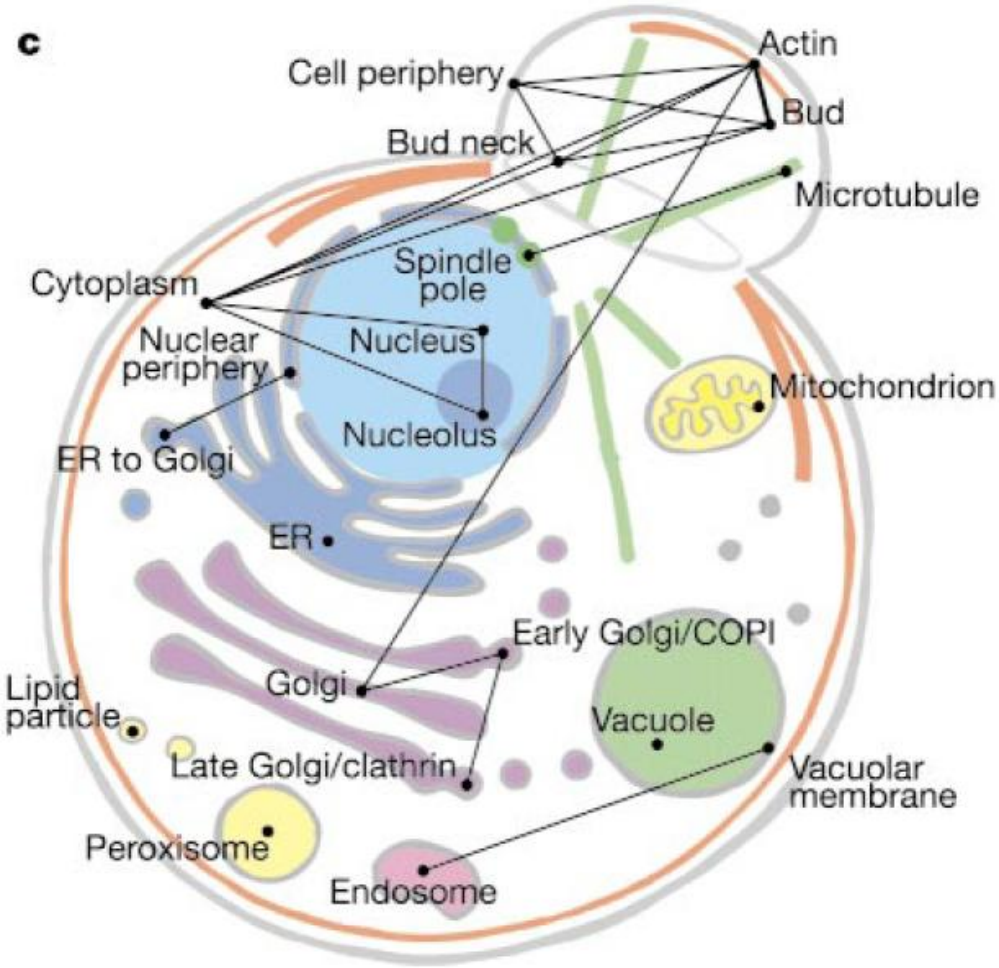


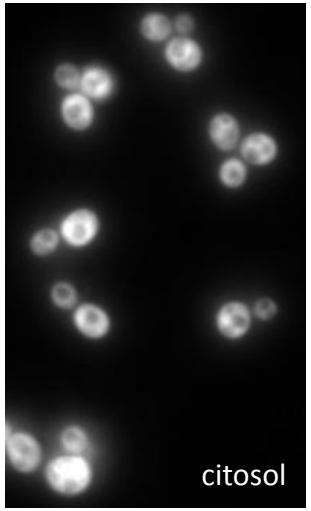
J Proteom Res (2009) 8:20

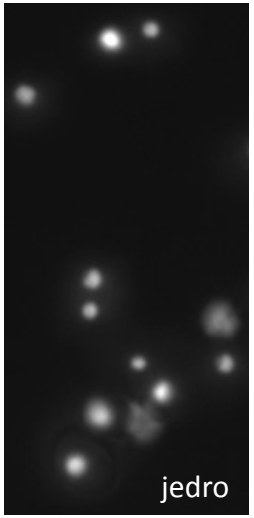
Napake pri metodah funkcijske genomike:
kako izvreči lažno pozitivne oziroma kako ujeti lažno negativne

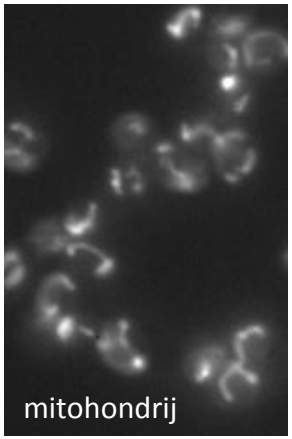
Analiza slike: človek vs. računalnik

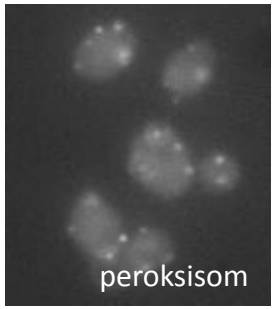


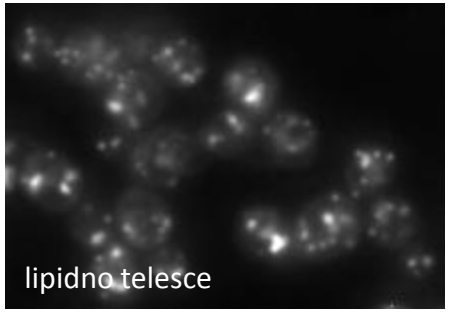


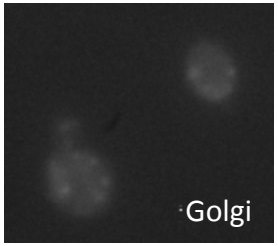


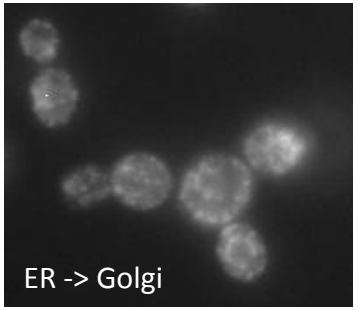






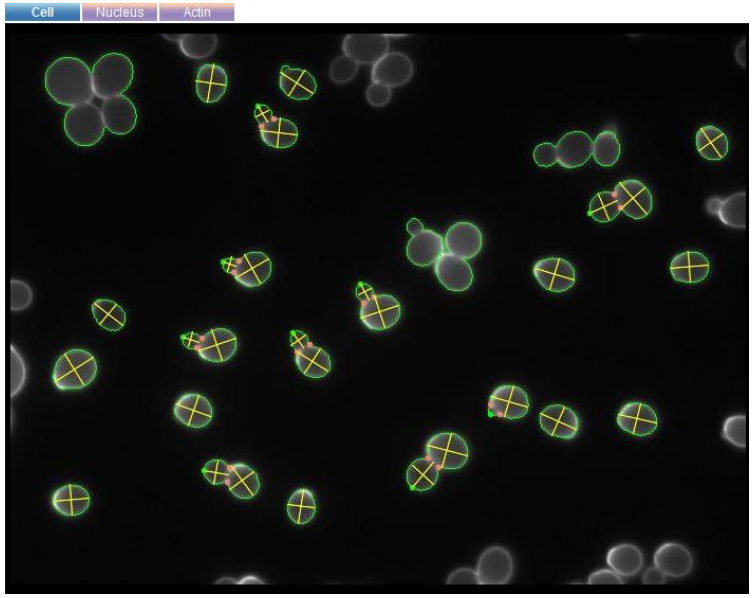


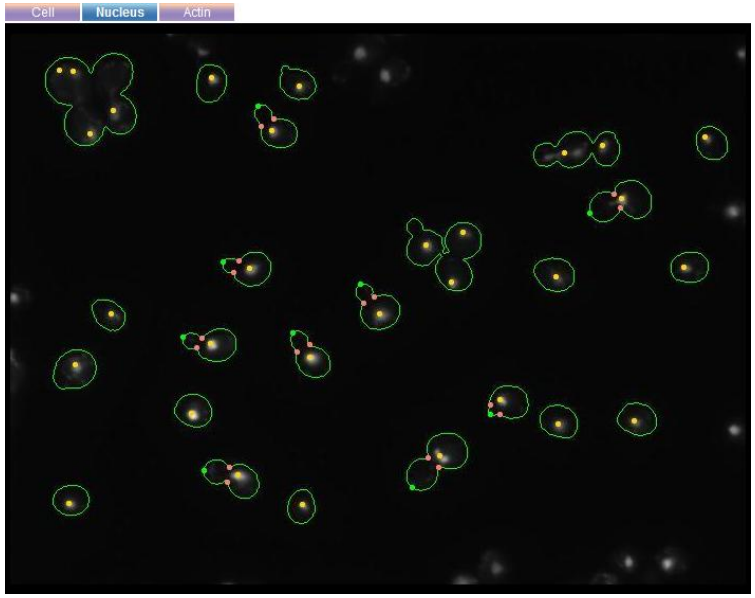


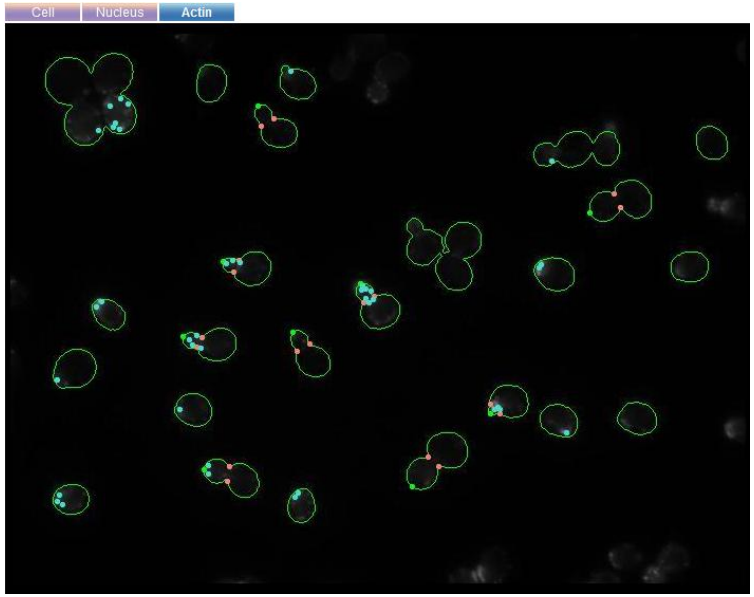


Nature (2003) 425:686

SCMD Saccharomyces Cerevisiae Morphological Database

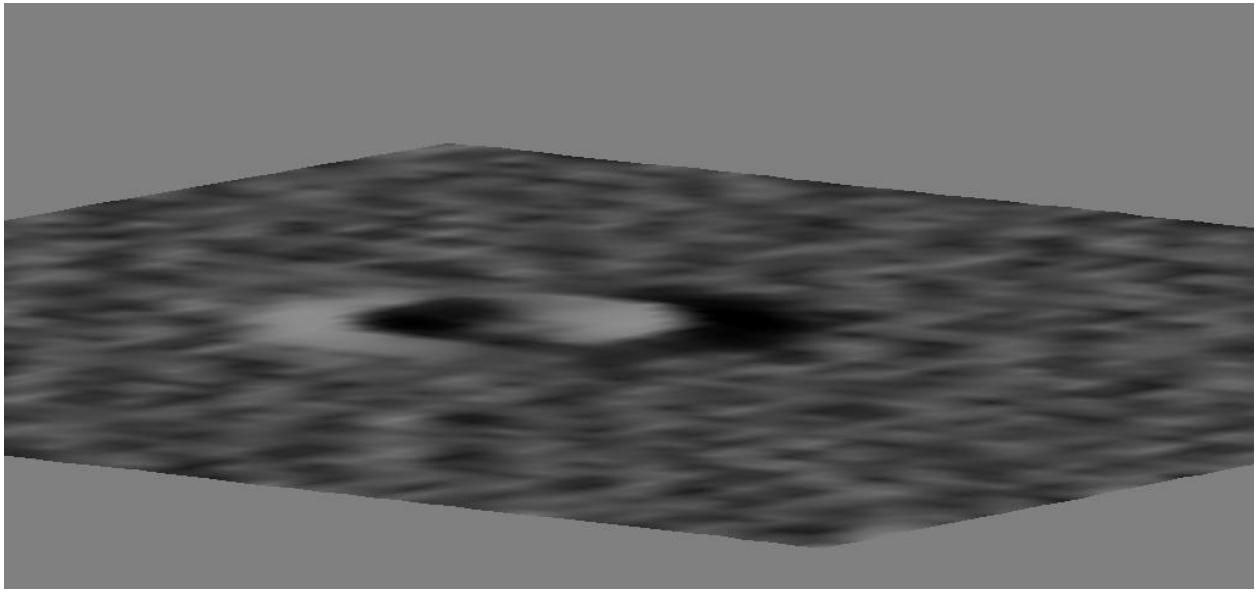


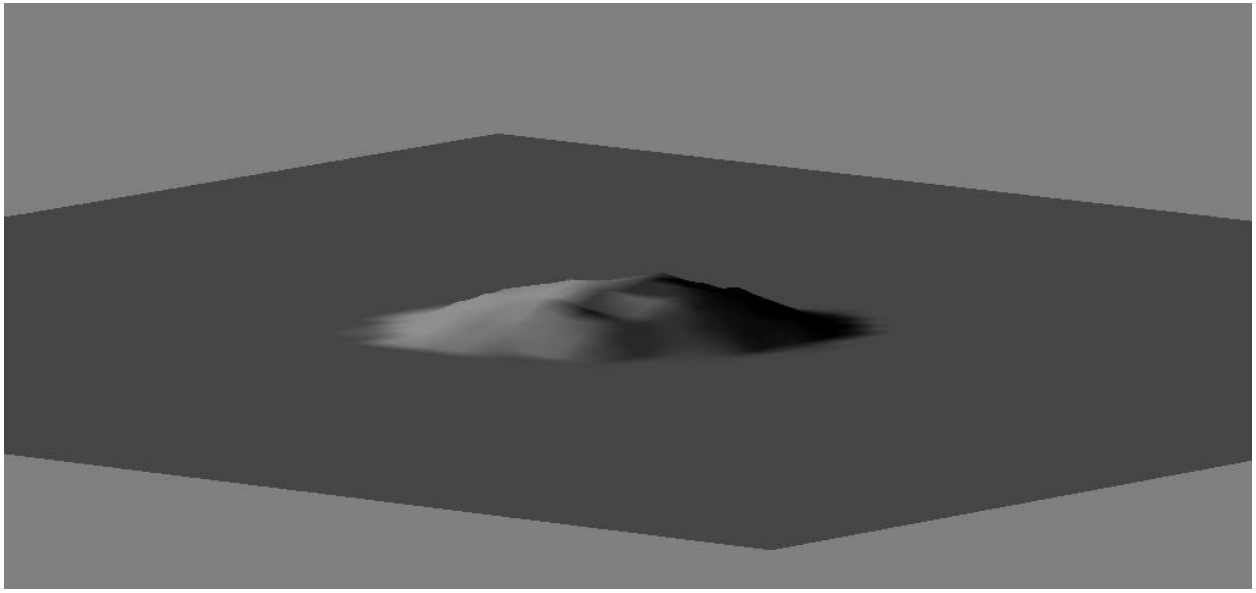


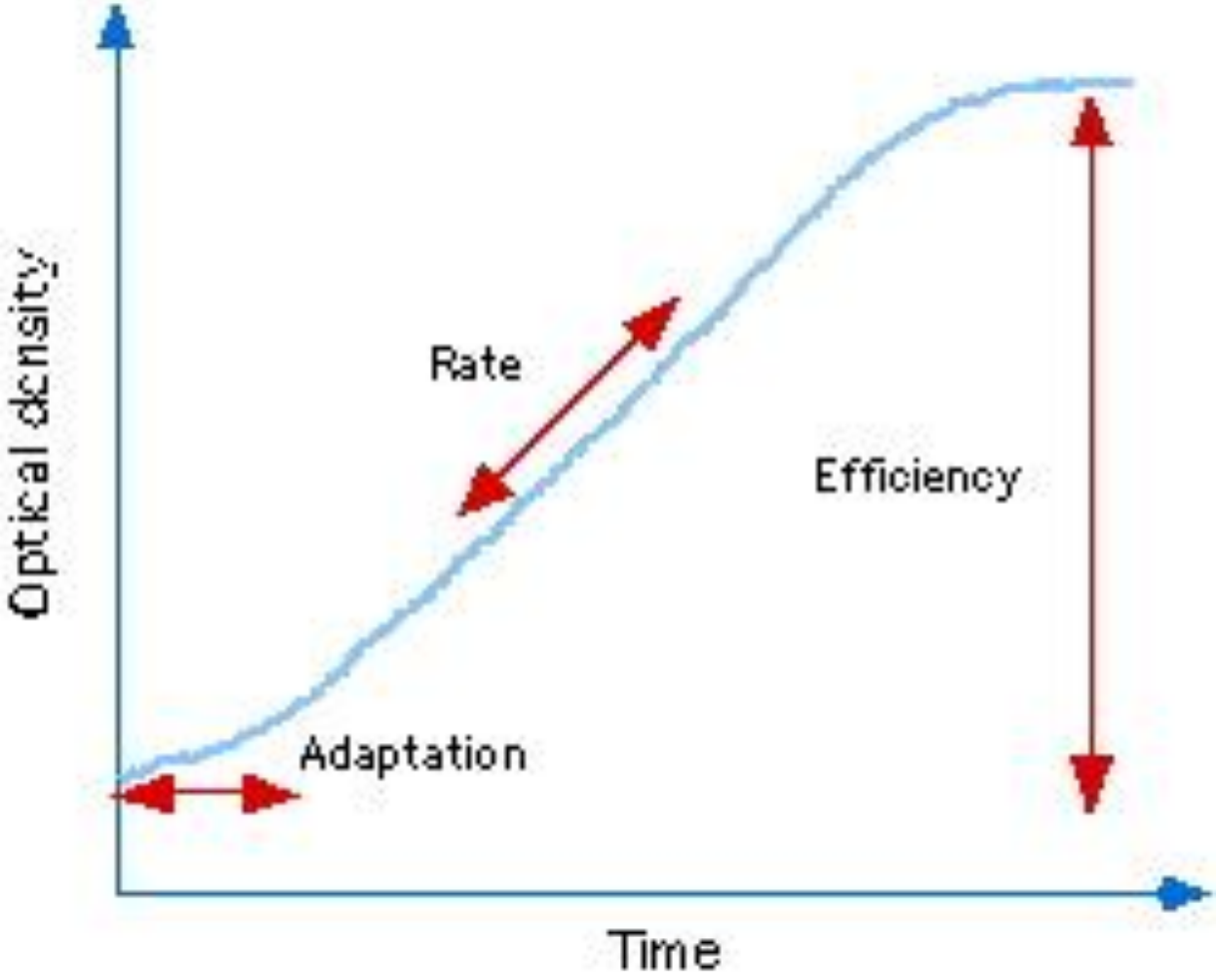


~500 morfoloških parametrov

Kateri so **paralelni** in kateri **ortogonalni**?






Optical density
Rate
Efficiency
Adaptation
Time