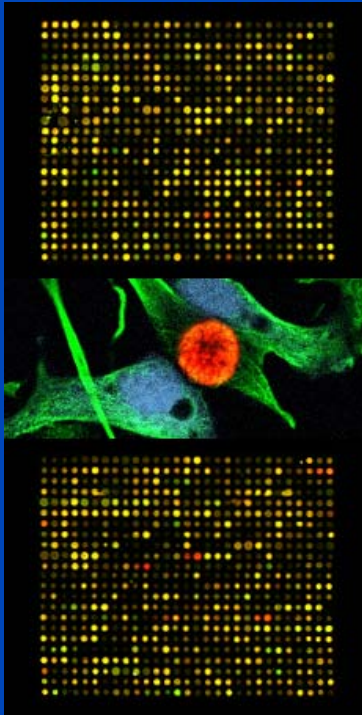




Data analysis in cell-based functional assay

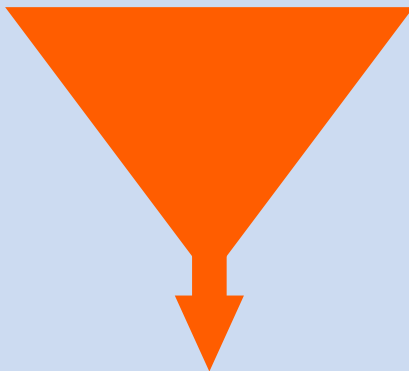
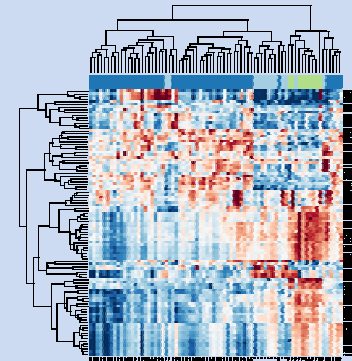
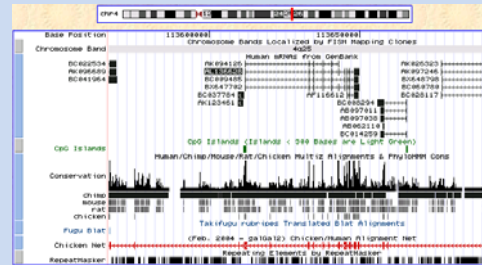
Tools for automated pre-processing, analysis and visualization of high throughput FACS data

- **Challenge and Concept**
- **Assay Design**
- **Data Analysis**

- **Challenge and Concept**
- Assay Design
- Data Analysis

The Challenge: Identification of Disease Genes

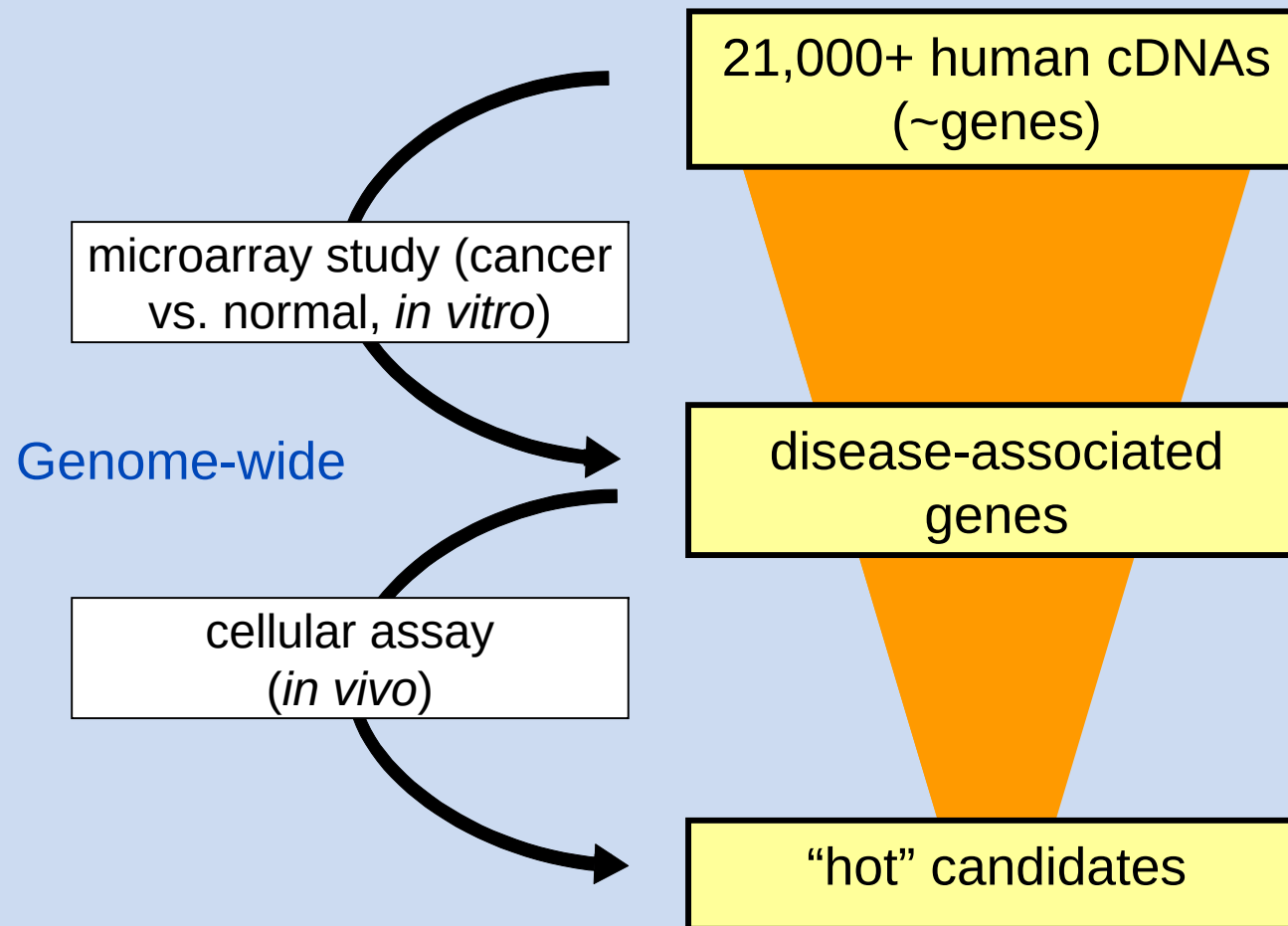
Candidate gene sets from
systematic gene identification
and microarray studies:
dozens...hundreds

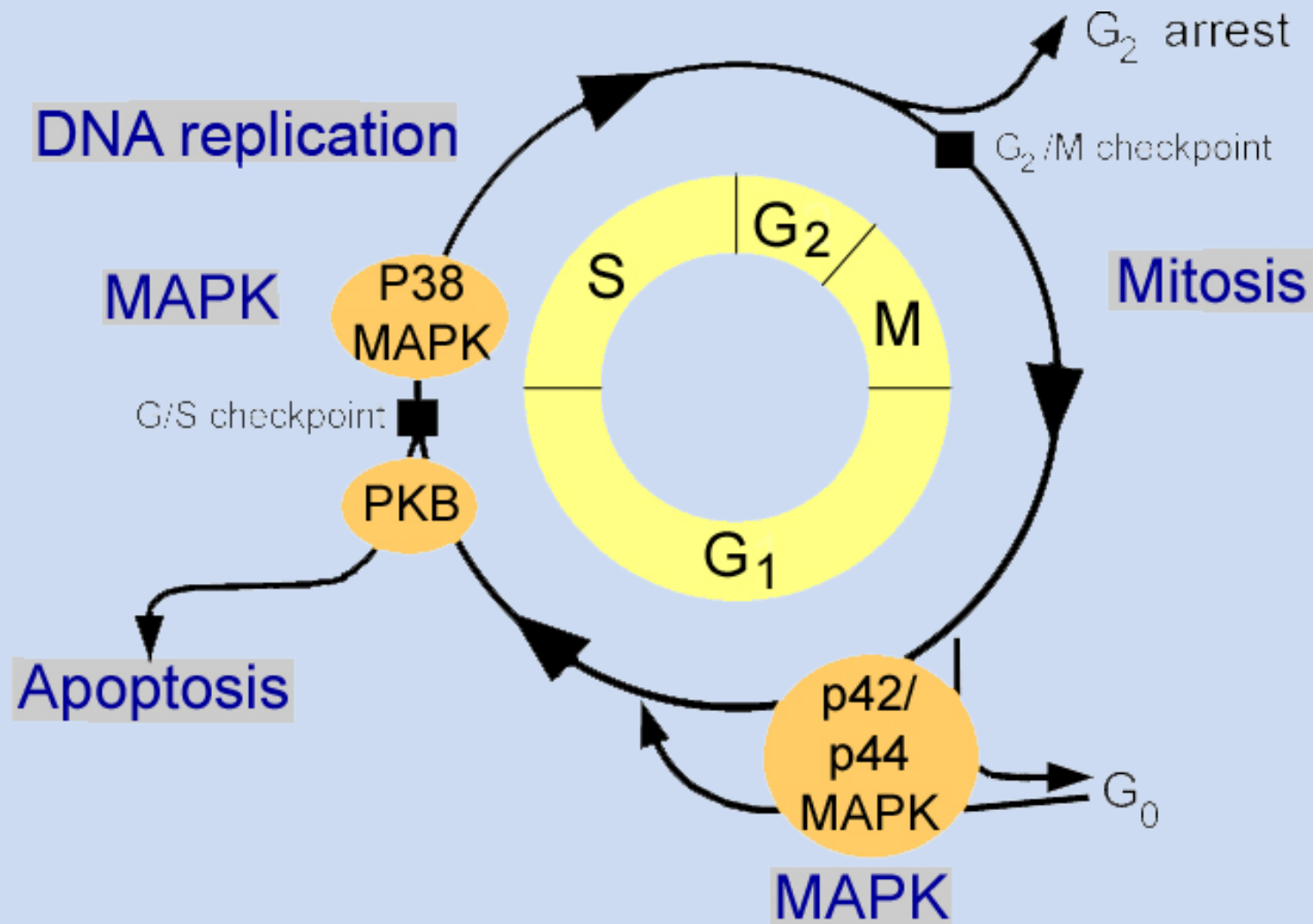


Capacity of *in-vivo*
functional studies: ...few

How close the gap?



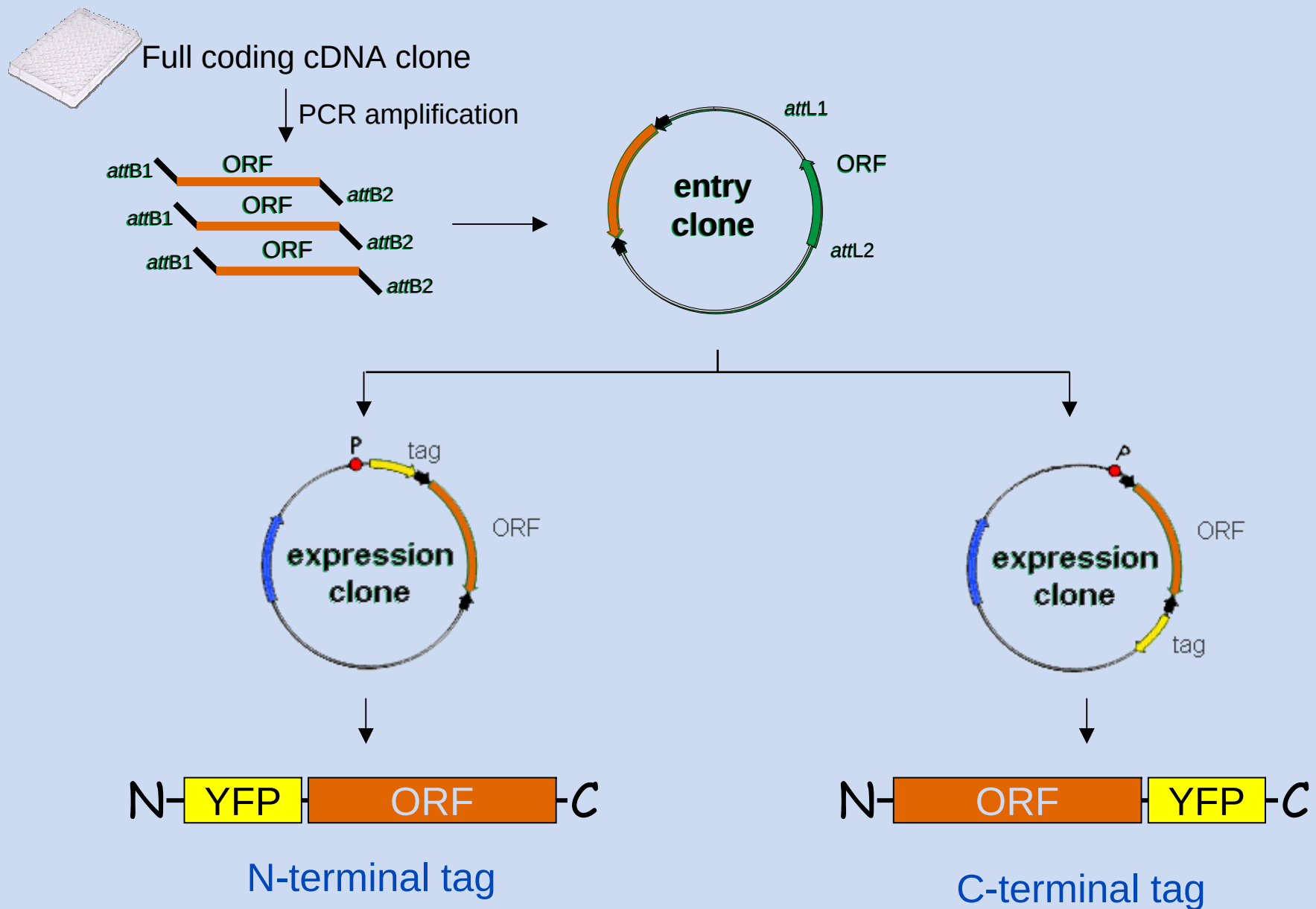




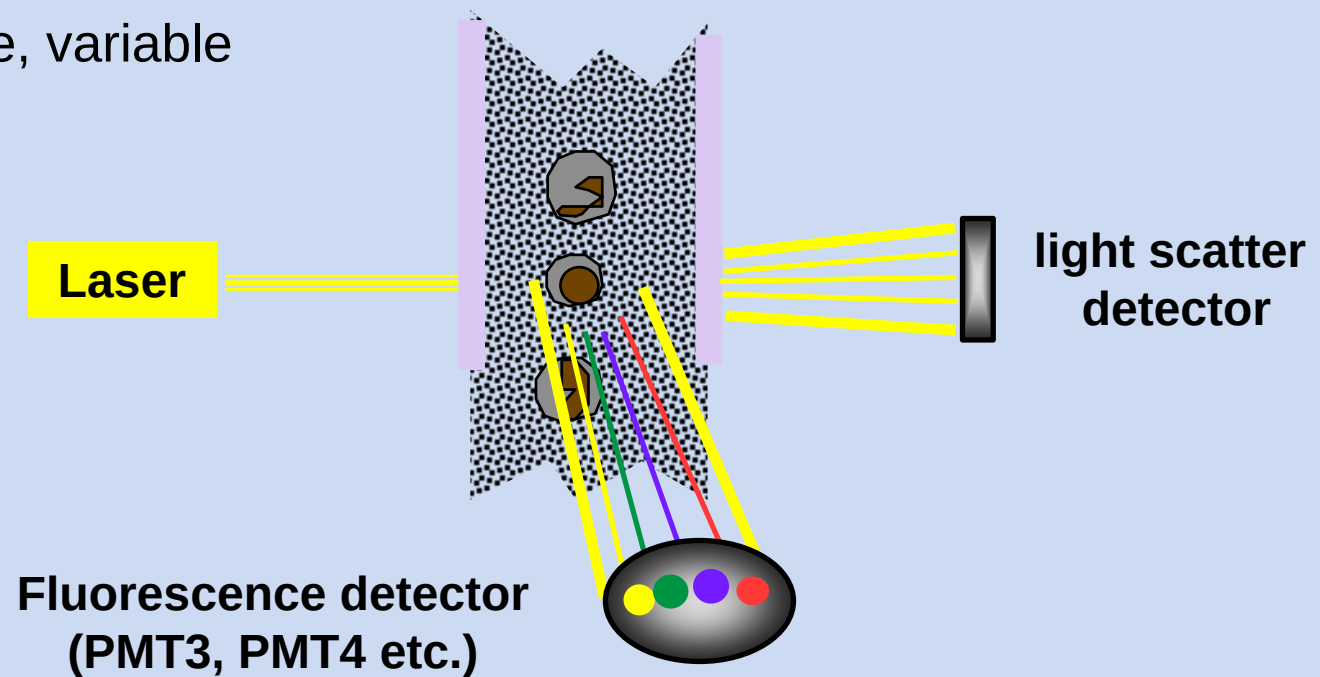
- Challenge and Concept
- **Assay Design**
- Data Analysis

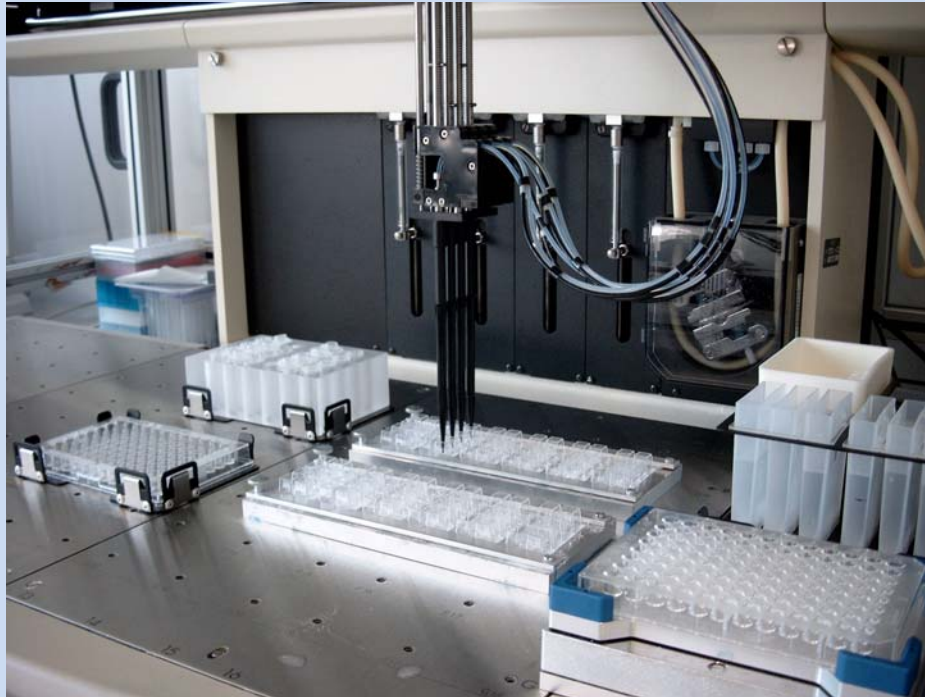
- system to willfully manipulate expression level of certain genes in cells
 - ↑ up regulation (transfection of expression vectors)
 - ↓ down regulation (RNA interference)
- means to monitor perturbation (beneficial but not mandatory)
 - expression of fluorescence protein tag
- means to monitor effect of perturbation
 - expression or activation state of key regulatory proteins (FACS, automated microscope)

ORF cloning: The Gateway™ System



- measures fluorescence intensities as well as morphological parameters on the basis of light emission
- offers single cell resolution
- robust, reliable, variable

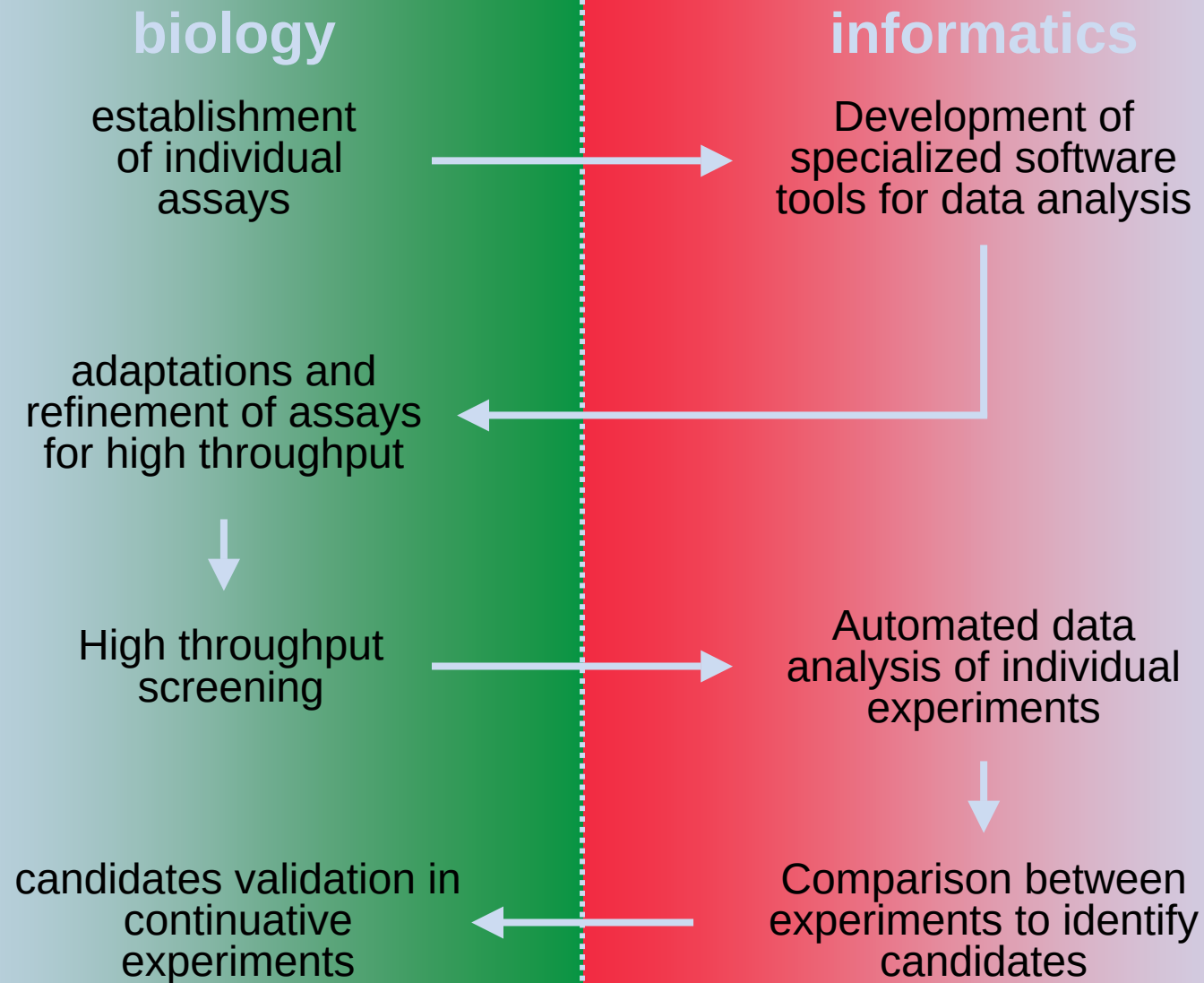




pipetting robot
(liquid handling)

HTS Sampler for
automated flow cytometry



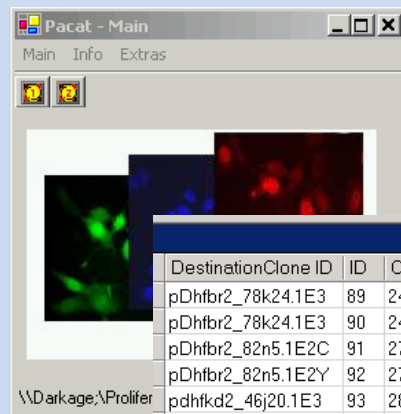


- Challenge and Concept
- Assay Design
- **Data Analysis**

Keeping track of experiments: PACAT

PACAT

(proliferation assay clone administration tool)



Destination	Clone ID	ID	Ort
pDhfr2_78k24.1E3	89	24	
pDhfr2_78k24.1E3	90	24	
pDhfr2_82n5.1E2C	91	27	
pDhfr2_82n5.1E2Y	92	27	
pdhfr2_46j20.1E3	93	28	
pdhfr2_46j20.1E3Y	94	28	
pdYFPPhfr2_4k14	95	29	
pdYFPPhfr2_4k14	96	29	
pdCFPhfr2_4k14	97	29	
pdCFPhfr2_4k14	98	29	
pdhfr2_4k14.1E2Y	99	29	
pdhfr2_4k14.1E2Y	100	29	
pdhfr2_4k14.1E2	101	29	
pdhfr2_4k14.1E2	102	29	
pdhfr2_4k14.1E2	103	29	
pdhfr2_4k14.1E2	104	29	
pDhtes3_35p17.1E3	105	31	
pDhtes3_35p17.1E3	106	31	
pDhtes3_35p17.1E3	107	31	
pDhtes3_35p17.1E4	108	31	
pDhtes3_35p17.1E4	109	31	
pDhtes3_35p17.1E4	110	31	
pDhtes3_35p17.1E4	111	31	
pDhtes3_50j4.1E1C	112	33	
pDhtes3_50j4.1E1C	113	33	
pDhtes3_50j4.1E1C	114	33	
pDhtes3_50j4.1E1Y	115	33	
pDhtes3_50j4.1E1Y	116	33	
pDhtes3_50j4.1E1Y	117	33	
pDhtes3_50j4.1E2C1	118	33	
pDhtes3_50j4.1E2C	119	33	
pDhtes3_50j4.1E2Y	120	33	
pDhtes3_50j4.1E2Y	121	33	

Project: proliferation plate1 Create: 6/25/2003 Deliver: 0 Institute: dkfz Comment:

Name: dorit8 Result: 0 Name: Art

Open Save Export

	0	1	2	3	4	5	6	7	8	9	10	11	12
A	YFP	YCA	YPP2	pdYFPPhamy 2_1p4_1E1 Y1	pdYFPPhamy 2_1p4_1E1 Y2	pDhfr2_2d 15.1E6Y1	pDhfr2_2d1 5.1E6Y2	pDhfr2_78k 24.1E3Y1	pdhfr2_4k1 4.1E2Y1	pdhfr2_4k1 4.1E2Y2	pDhtes3_50j 4.1E1Y2	pDhtes3_50j 4.1E2Y1	
B	YFP	YCA	YPP2	pdYFPPhamy 2_1p4_1E1 Y1	pdYFPPhamy 2_1p4_1E1 Y2	pDhfr2_2d 15.1E6Y1	pDhfr2_2d1 5.1E6Y2	pDhfr2_78k 24.1E3Y1	pdhfr2_4k1 4.1E2Y1	pdhfr2_4k1 4.1E2Y2	pDhtes3_50j 4.1E1Y2	pDhtes3_50j 4.1E2Y1	
C	pDhtes3_7p 9.1E1Y1	pdhfr2_2c1 8.2E2Y1	pdYFPPhfr2 _2c18Gy	YFP	YCA	YPP2	pdYFPPhfr2 _22f21_B	pdYFPPhfr2 _22a1_By	pdYFPPhfr2 _23b10_A	YFP	YCA	YPP2	
D	pDhtes3_7p 9.1E1Y1	pdhfr2_2c1 8.2E2Y1	pdYFPPhfr2 _2c18Gy	YFP	YCA	YPP2	pdYFPPhfr2 _22f21_B	pdYFPPhfr2 _22a1_By	pdYFPPhfr2 _23b10_A	YFP	YCA	YPP2	
E	CFP	CCA	CPP2	pdCFPhamy 2_1p4_Bx	pdCFPhamy 2_1p4_By	pDhfr2_2d 15.1E5C1	pDhfr2_2d1 5.1E5C2	pDhfr2_78k 24.1E3C2	pdhfr2_4k1 4.1E2C1	pdhfr2_4k1 4.1E2C2	pDhtes3_50j 4.1E1C3	pDhtes3_50j 4.1E1C2	
F	CFP	CCA	CPP2	pdCFPhamy 2_1p4_Bx	pdCFPhamy 2_1p4_By	pDhfr2_2d 15.1E5C1	pDhfr2_2d1 5.1E5C2	pDhfr2_78k 24.1E3C2	pdhfr2_4k1 4.1E2C1	pdhfr2_4k1 4.1E2C2	pDhtes3_50j 4.1E1C3	pDhtes3_50j 4.1E1C2	
G	pDhtes3_7p 9.1E1C2	pdhfr2_2c1 8.3E1C1	pdhfr2_2c1 8.2E2C1	CFP	CCA	CPP2	pdCFPhfr2 _22f21_B1	pdCFPhfr2 _22a1_Bx	pdhfr2_23b10.1E1C 1	CFP	CCA	CPP2	
H	pDhtes3_7p 9.1E1C2	pdhfr2_2c1 8.3E1C1	pdhfr2_2c1 8.2E2C1	CFP	CCA	CPP2	pdCFPhfr2 _22f21_B1	pdCFPhfr2 _22a1_Bx	pdhfr2_23b10.1E1C 1	CFP	CCA	CPP2	



package prada contains functionalities for analysis of data derived from cell based assays

→ **modular framework**

- data preprocessing
- data visualization
- data integration

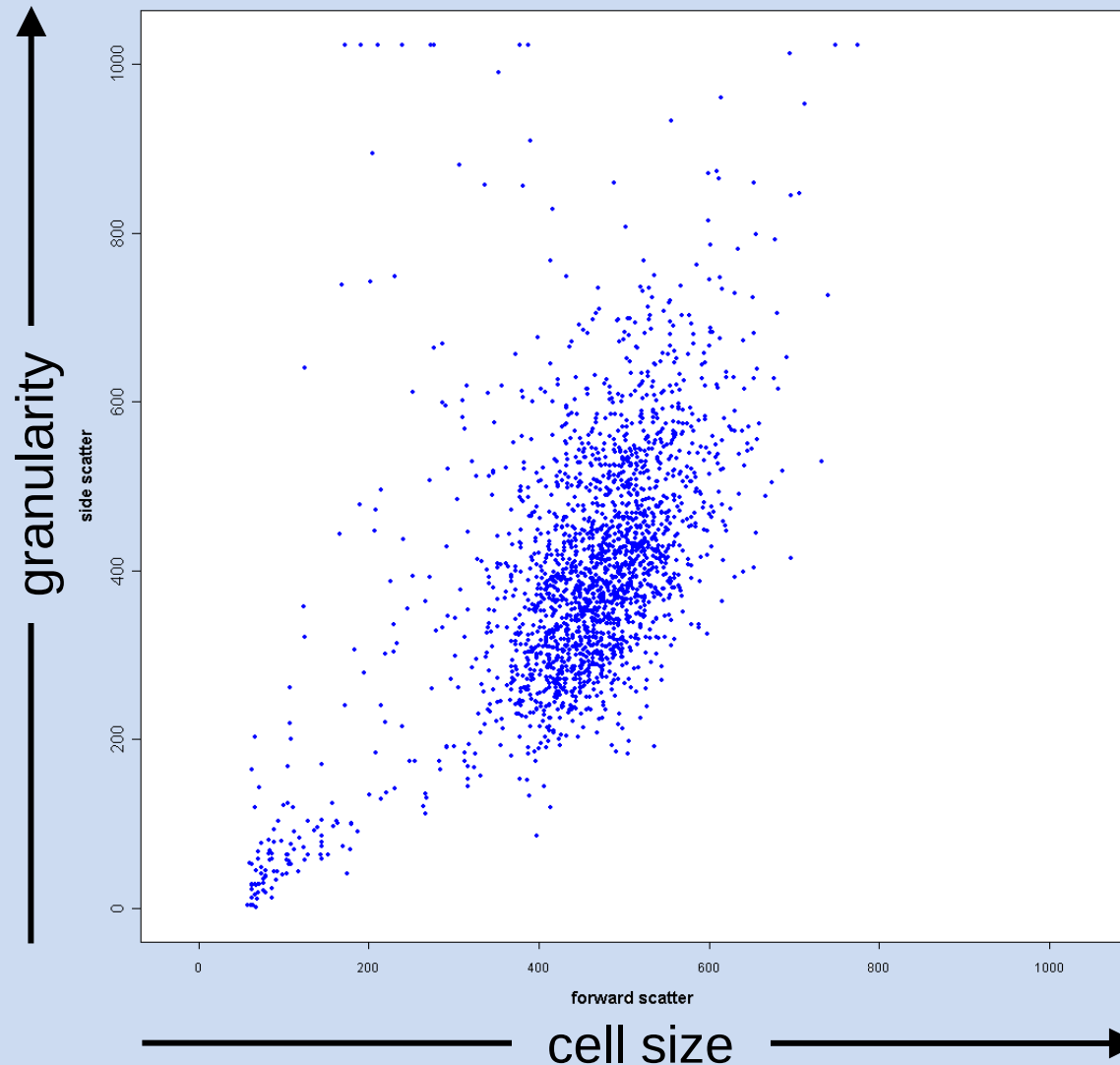
for statistical inference and modeling general purpose tools can be used

- linear and local regression
- hypothesis testing

- **FCS 3.0 files**
 - standardized storage format for FACS data
 - contains fluorescence values in data segment, wealth of meta data in text segment
 - can be imported into R (*function readFCS*)
- **cytoFrame**

R internal representation of data from one FCS file
→ *generic functions*
- **cytoSet**

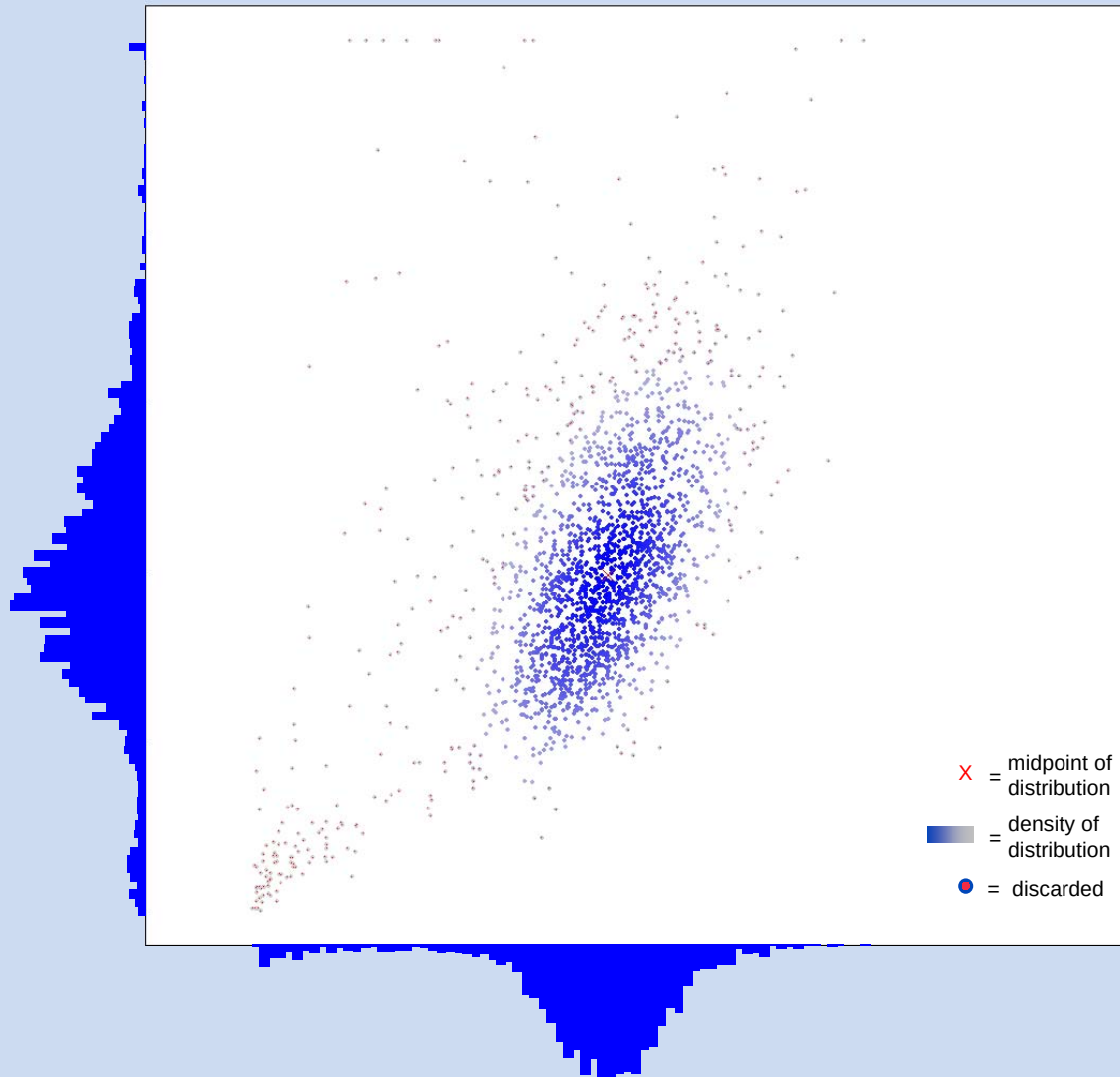
R internal representation of data from several FCS files
(e.g. one 96 well plate)



distinction on basis of morphological properties

strong variation between experiments

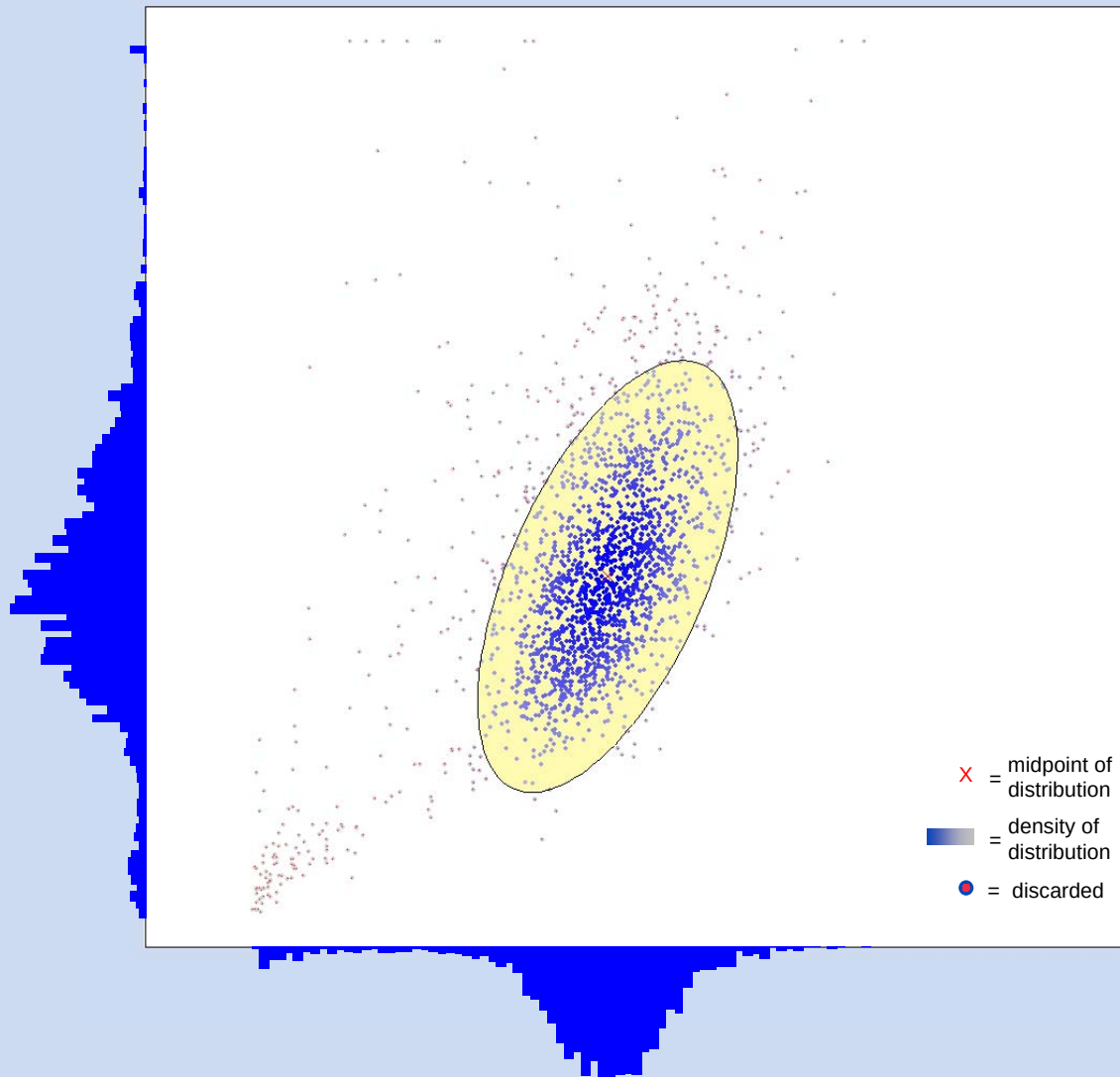
→ dynamic determination



assumption:

bivariate normal distribution
→ robust fitting

discarding cells that do not
lie within some given
boundary of this distribution



assumption:

bivariate normal distribution
→ robust fitting

discarding cells that do not lie within some given boundary of this distribution

shape and localization of main distribution can be used for **quality control**

plate plots as graphical representation of experimental entities

- false color coding for concise display of numeric outcomes from statistical analyses
- HTML image map allows for hyper linking to include further information for each well

quantitative

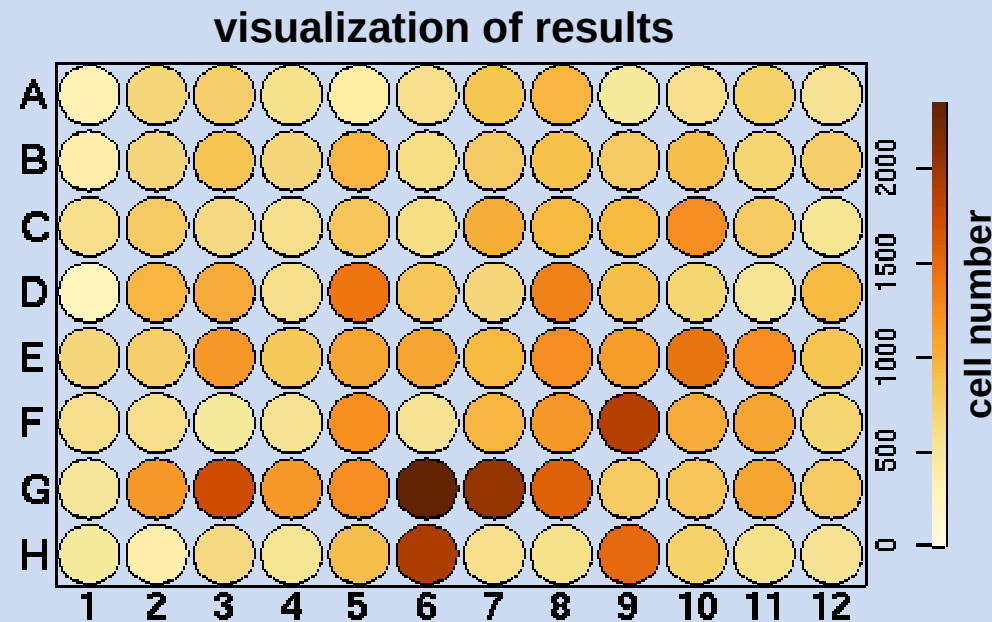
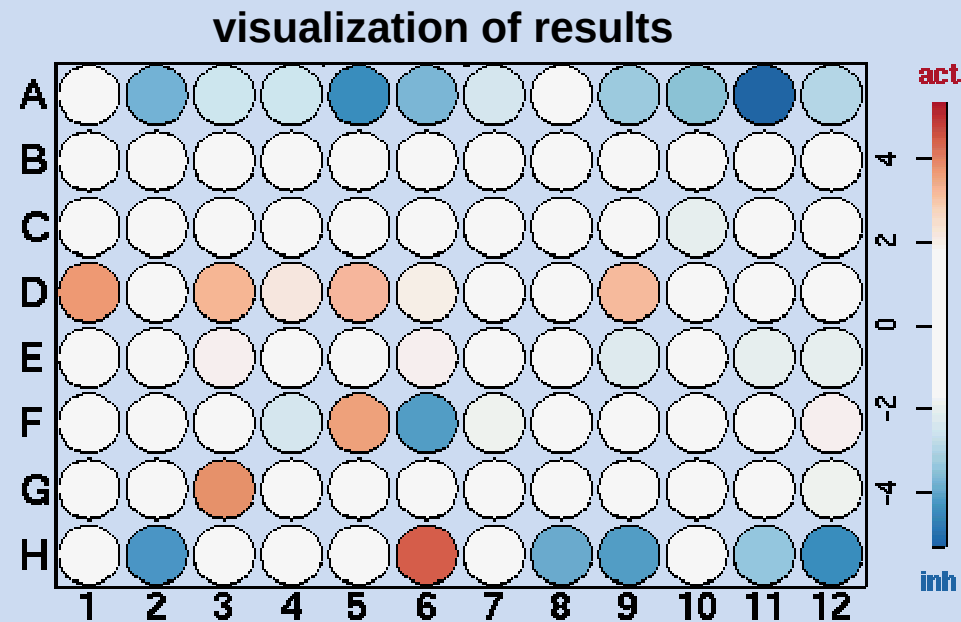


plate plots as graphical representation of experimental entities

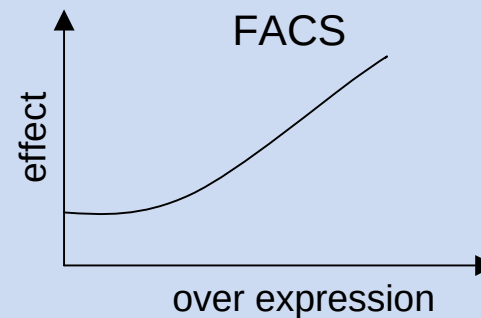
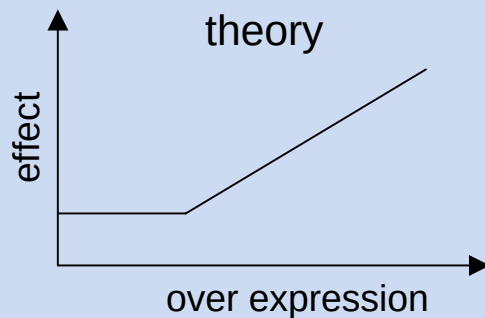
- false color coding for concise display of numeric outcomes from statistical analyses
- HTML image map allows for hyper linking to include further information for each well

qualitative

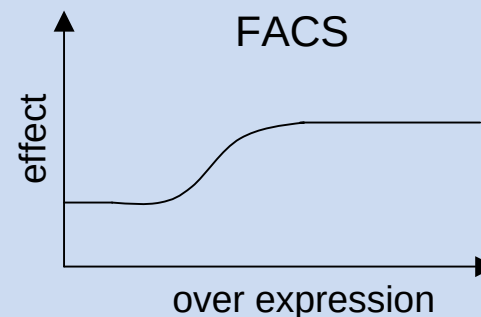
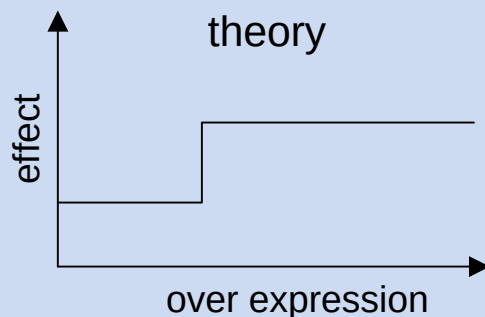


different responses for different assays

- continuous response: concentration dependent (e.g. MAP kinase)



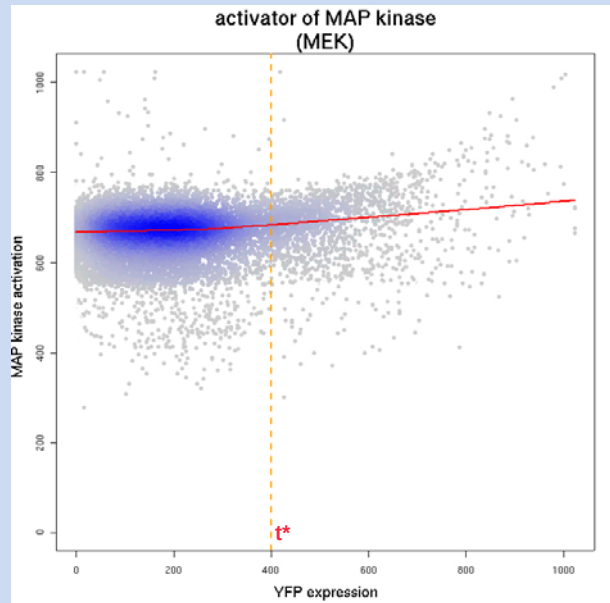
- discrete response: on/off mechanism (e.g. apoptosis, proliferation)



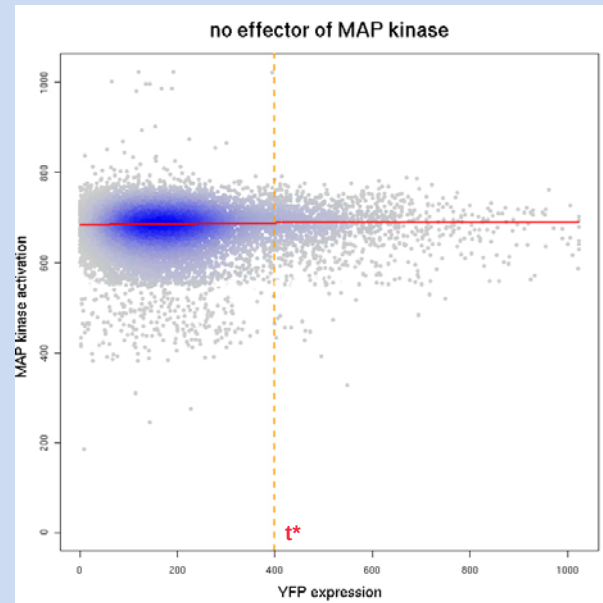
- robust fitting of smoothed local regression function
- z-score as measure of effect:
ratio of estimated slope and its standard error
at YFP intensity t^*

$$z = \frac{\hat{m}'(t^*)}{\hat{\sigma}_{m'}(t^*)}$$

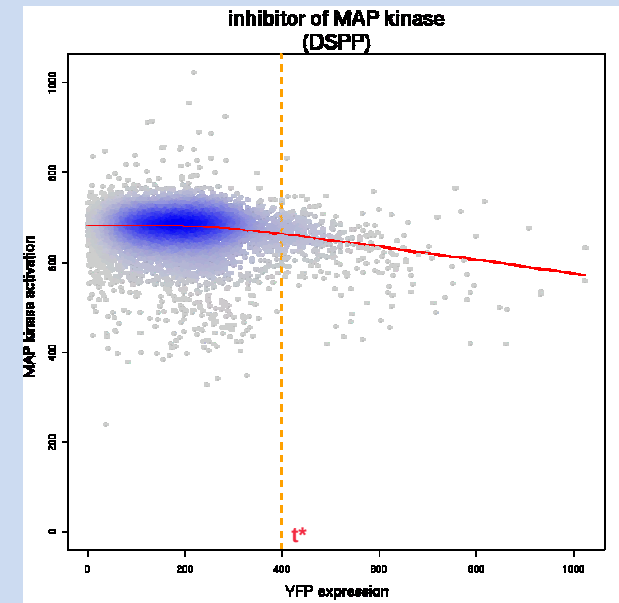
$z = 8.59$



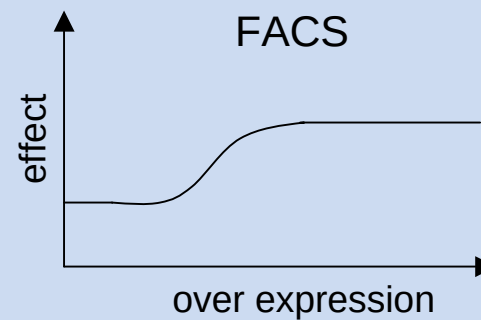
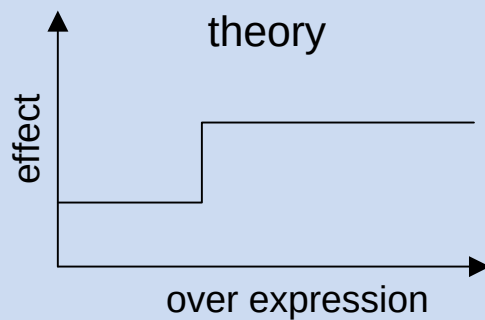
$z = 0.88$



$z = -11.42$



- discrete response: on/off mechanism
(e.g. apoptosis, proliferation)



Fisher's exact test

$$r_1 = \frac{a}{b}$$

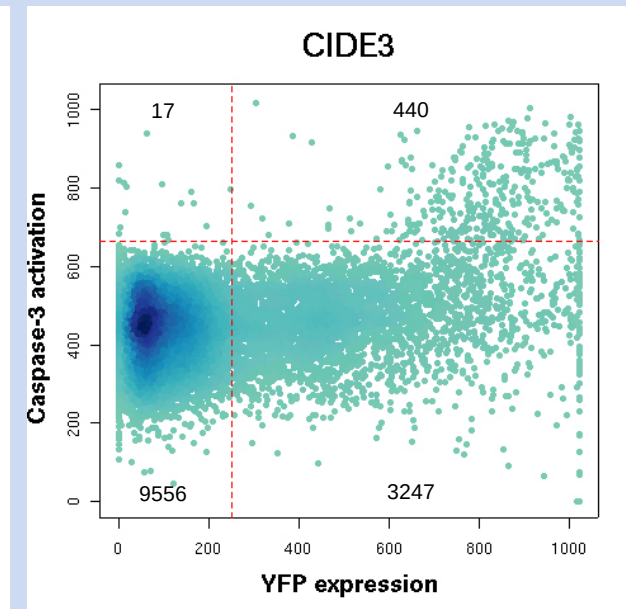
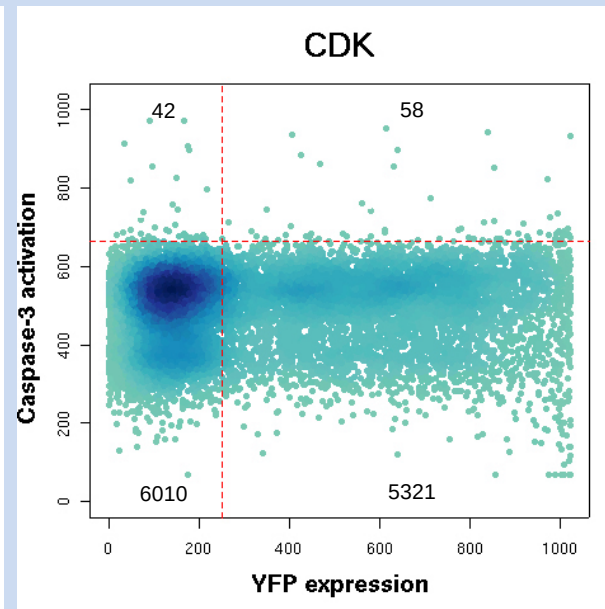
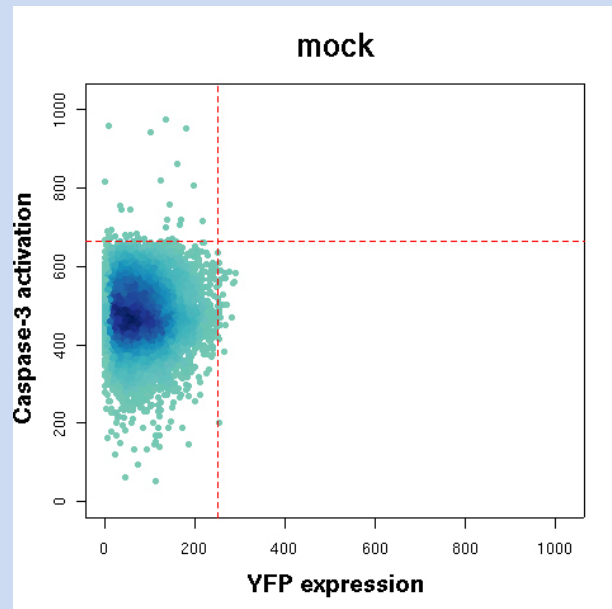
effect	untransfected positive (a)	transfected positive (c)
	untransfected negative (b)	transfected negative (d)
	transfection	

$$r_2 = \frac{c}{d}$$

effect size $\leftarrow$ odds ratio $= \frac{r_1}{r_2}$, p value $\rightarrow$ significance

no effect

activator

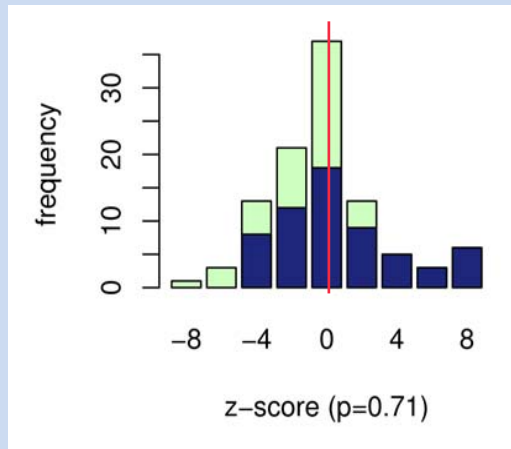


$$-\log(\text{odds ratio}) = 0.44$$
$$(p = 4.4\text{e-}03)$$

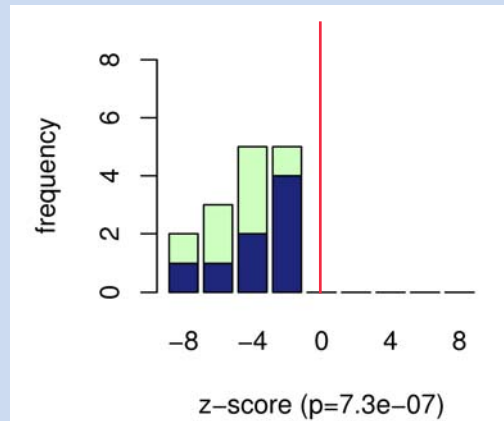
$$-\log(\text{odds ratio}) = 4.33$$
$$(p = 2.2\text{e-}16)$$

MAP kinase

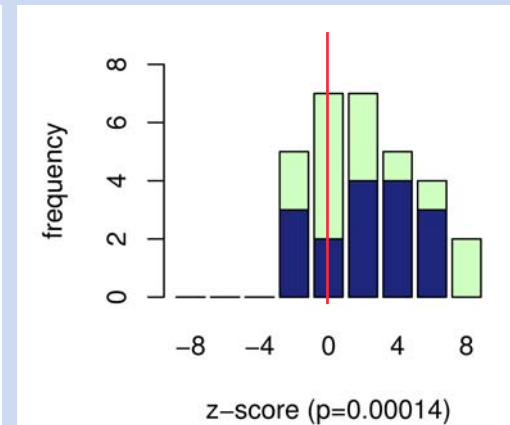
control



inhibitor

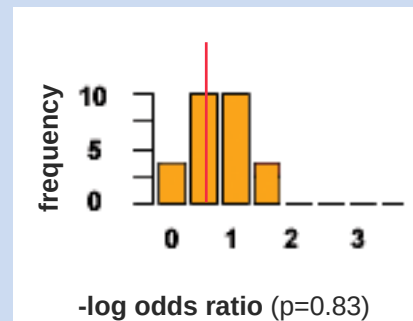


activator

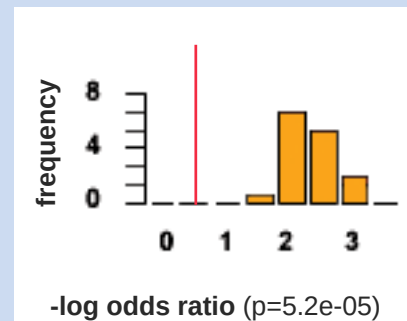


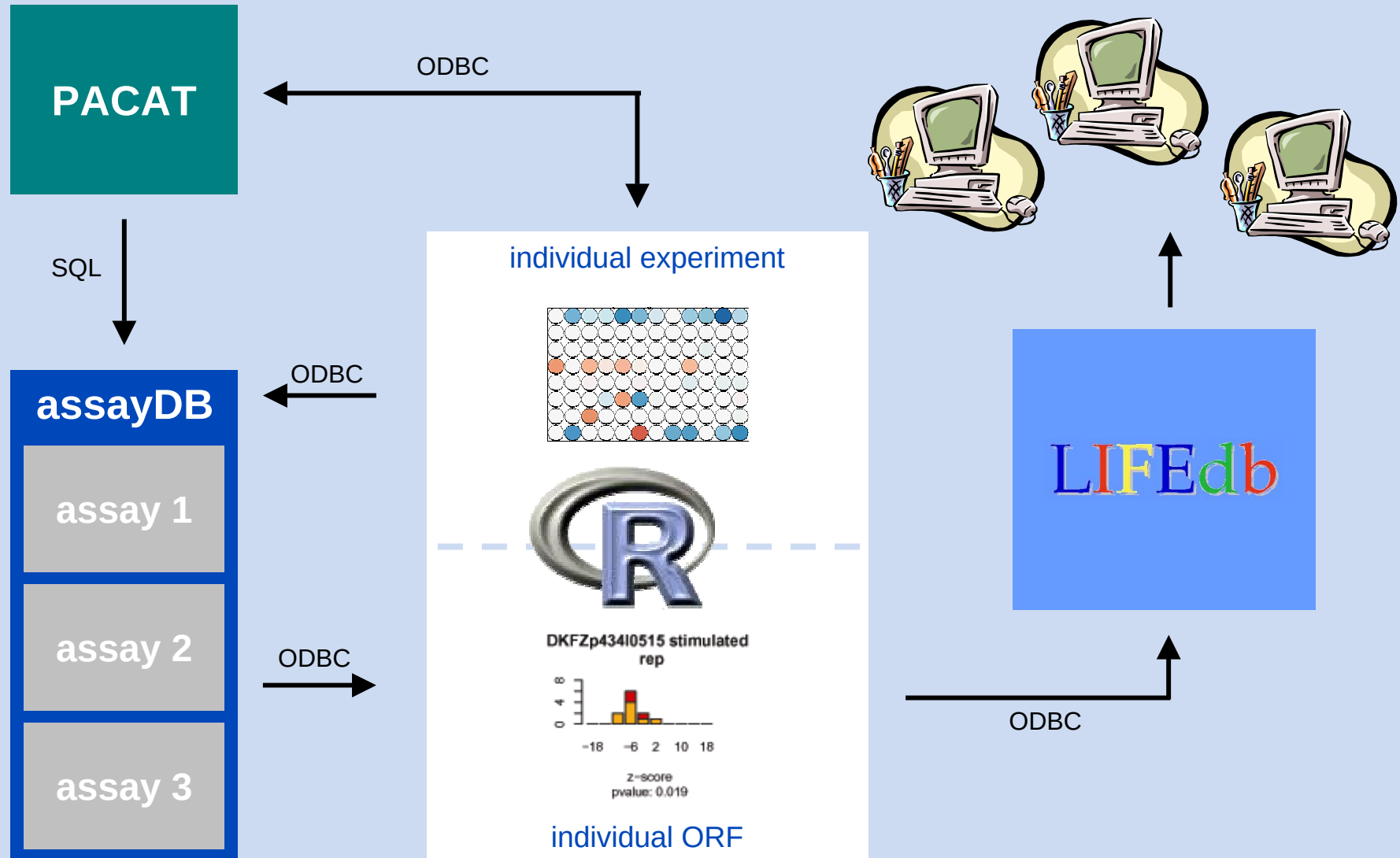
apoptosis

control



activator





- cellular assays help to close the gap between genome-wide large scale studies and analyses on the single molecule level

association/correlation → causal relationships

- FACS has proven to be a capable tool for high throughput analyses with single cell resolution
- package prada provides a framework for integrating various analysis approaches of multiple assays

modular structure

Annemarie Poustka

Stefan Wiemann

Wolfgang Huber

Andreas Buneß

Marcus Ruschhaupt

Heiko Rosenfelder

Alex Mehrle

Dorit Arlt

Meher Majety

Mamatha Sauermann

YOU for the invitation!