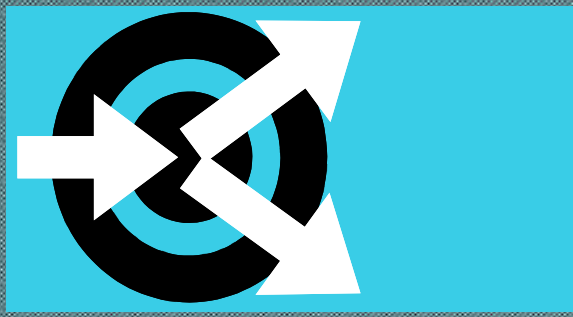
Computing offtarget-effect by cross-analysis of transcriptome-level data from ~~1,118~~ 1,380 studies

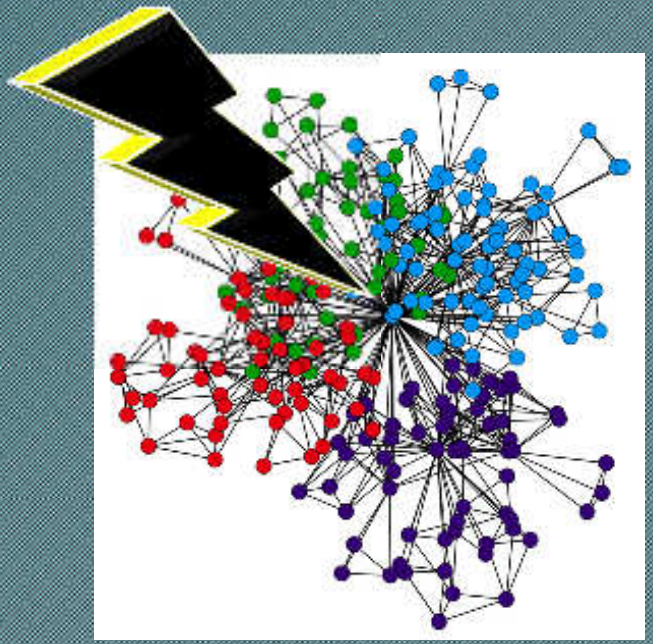


Balázs Győrffy

MTA TTK Lendület Cancer Biomarker Research Group

Semmelweis University 2nd Dept. of Pediatrics

Side effects of molecular targeted therapy



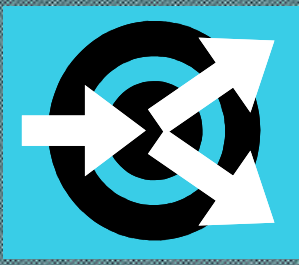
On-target effect:

- EGFR inhibitors -> skin rash
- VEGFR2 blockers -> vasoconstriction -> hypertonia

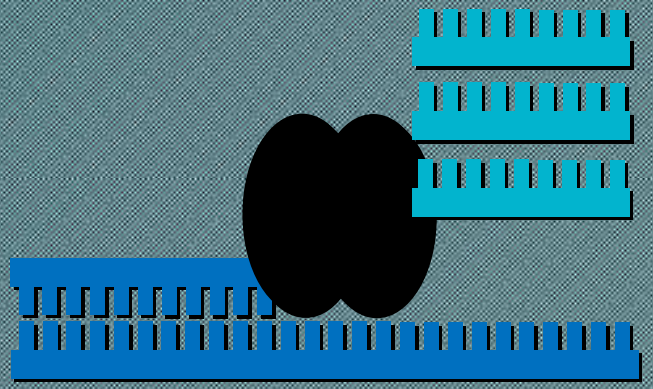
Off-target effect:

- HER2 inhibitors -> heart failure
- mTOR inhibitors -> high cholesterol
- BRAF inhibitors -> epithelial tumors

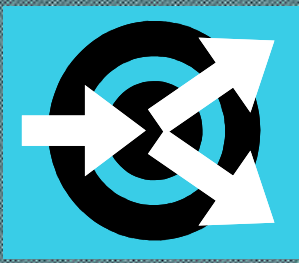
Therapy



1. Drug treatment
   - Approved medicines
   - Agents of preclinical studies
2. RNA interference
   - double-stranded RNA
   - "interfere" with the translation of protein
   - degradation of mRNA at specific sequences
3. Overexpression
   - Experimentally induced excessive expression of a gene



Therapy -> Off-target effects



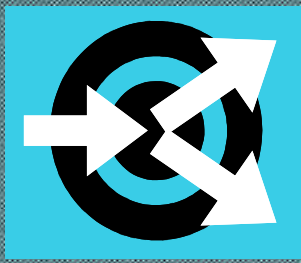
General approach:

treatment  off-target effect

Reversed question:

off-target effect  which treatment?

Study aim



Establish a framework to identify off-target effects for a large number of agents on a genomic scale.

Methods 1/3: database setup

GEO search: all cell culture experiments, 2005-2014, n=145,693

↓

Raw data (CEL file) available n=134,289

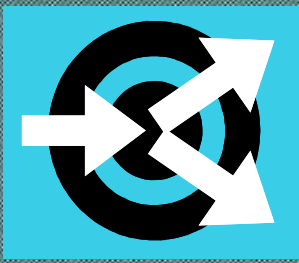
↓

Treatment-control pairs n=68,108

→

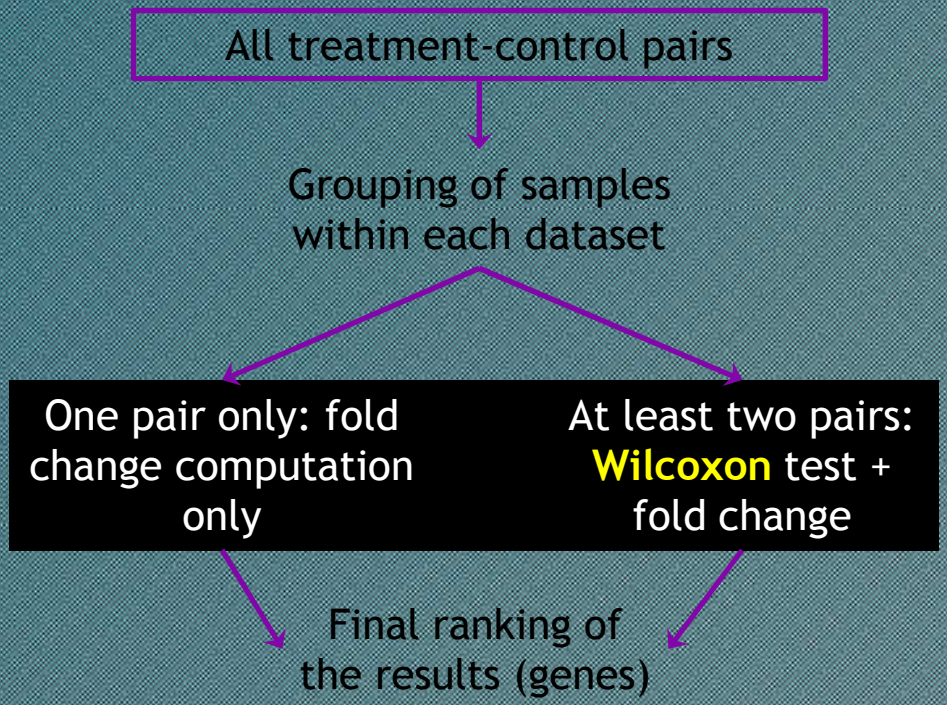
The database in numbers	n
Datasets	1,380
Cell lines	1,823
Drugs	1,920
Silenced / overexpressed genes	759

Methods 2/3: data processing

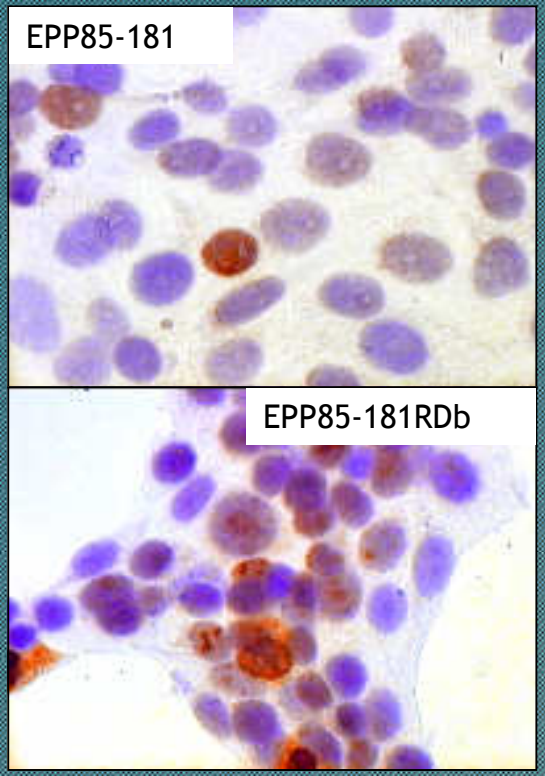


- **Platforms:** Affymetrix HGU133A, HGU133plus2, HGU133Av2
- **R environment**
- **Normalization:** MAS5
- **Quality control:** Affy whitepaper
- **Probe sets:** JetSet (n=12,209)

Methods 3/3: statistical analysis

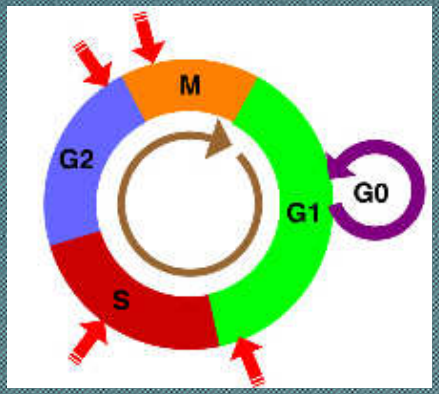


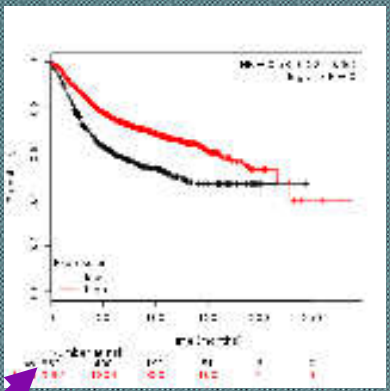
Results 1/3: proliferation (breast cancer)



MKI67

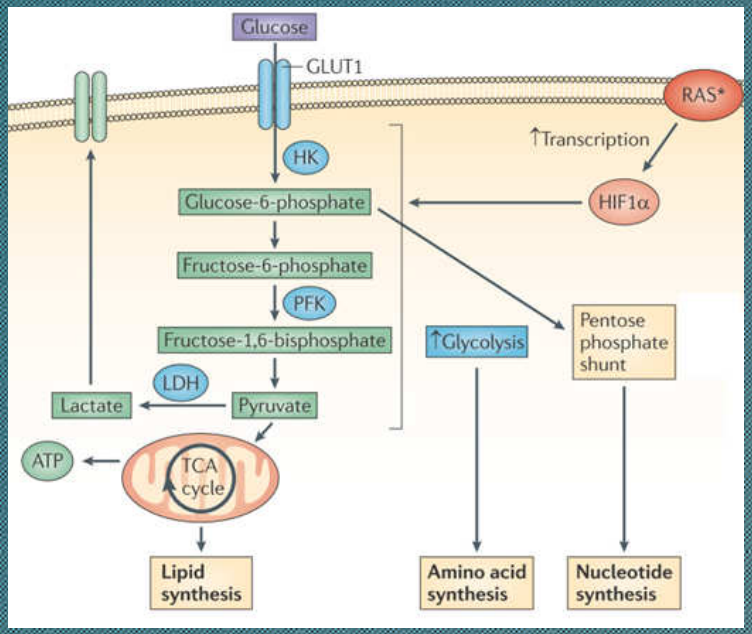
(212023_s_at)





Treatment	Type	P value	Fold change	Datasets
ESR2	Overexpression	1E-16	0.62	GSE9936
Resveratrol	Drug	3E-08	0.43	CMAP
MTX	Drug	1.5E-05	0.45	CMAP
Estrogen	Drug	1.1E-04	1.68	GSE11324, GSE8640

Results 2/3: energy metabolism (drug)



HIF1A

(200989_at)

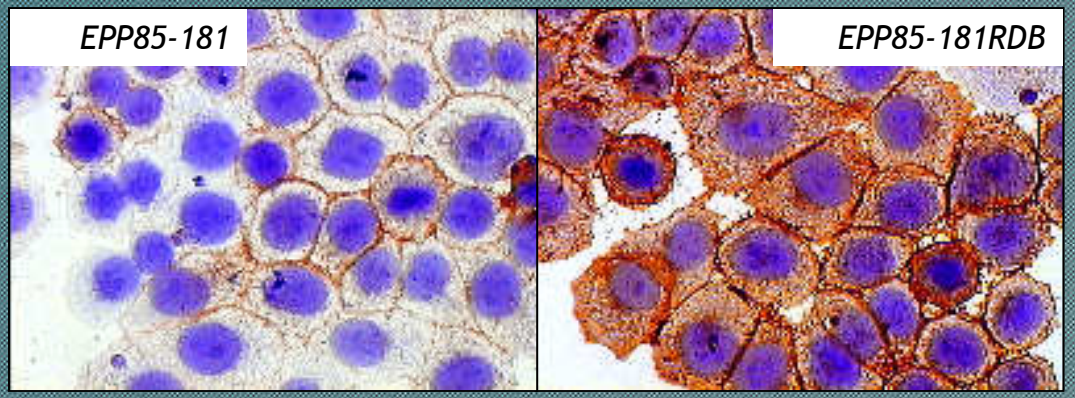
Treatment	Type	P value	Fold change	Datasets
Camptothecin	Drug	0.00049	0.52	GSE1417
PLX4032	Drug	0.016	0.7	GSE20051

Vemurafenib

Topoisomerase inhibitor

Results 3/3: multidrug resistance (silencing)

ABCB1
(209993_at)

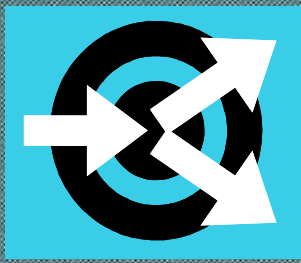


repair of DNA double strand breaks

PI3K-Akt signaling pathway

Treatment	Type	P value	Fold change	Datasets
RAD51	Silencing	0.022	5.6	GSE56940
Cyclin T	Silencing	9.6E-05	6.31	GSE11462
PSIP1	Silencing	0.0039	0.15	GSE7508
NFKB1	Silencing	0.024	0.14	GSE31912

Summary



Established a database and a bioinformatic framework to process a large volume of curated experimental data.

is useful to:

1. **We can identify the treatment inhibiting/activating any selected gene as its off-target effect.**
2. **We can select the most reliable preclinical model for a validation study.**

Acknowledgements

Database construction

Gyöngyi Munkácsy PhD

Péter Herman

Bence Bán

Nóra Szarvas

Zsófia Pénzváltó PhD

Bioinformatics

Zsófia Sztupinszki MD

Boglárka Weltz

Tamás Szabó Miklós