ProLiFiler and Cancer Data Miner, combined platforms for a preclinical investigation to scrutinize the impact of inhibitors on the KRAS, PI3K, and MDM2 signaling pathways

Vincent Vuaroqueaux[1], Daniel Feger[2], Anne-Lise Peille[1], Oliver Siedentopf[2], Hoor Al-Hasani[1], Sarah Ulrich[2], Sebastian Dempe[2], Heinz-Herbert Fiebig[1], Jan Erik Ehlert[2]

[1]4HF Biotec GmbH, Freiburg, Germany; [2]Reaction Biology, Freiburg, Germany

4HF Biotec Cancer Data Mining
www.4hfbiotec.com

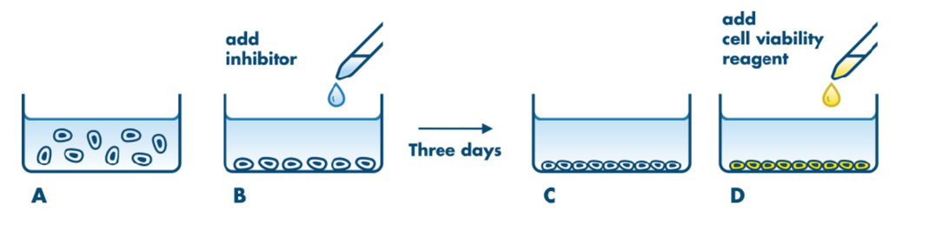
Introduction

Testing novel anti-cancer agents across a large panel of tumor models covering the genetic diversity of cancers is increasingly considered a cornerstone of preclinical development. Reaction Biology developed the ProLiFiler – performing the Cell Proliferation Assay on a tumor cell line panel covering most common cancer types to evaluate the anti-proliferative activity of novel drugs. To understand the molecular basis of drug sensitivity, 4HF Biotec uses their in-silico platform, "4HF-Cancer Data Miner", for bioinformatics analysis. Here we present integrative pharmacogenomic studies on three small molecules targeting major signaling pathways in cancer. It includes SOS1::KRAS interaction inhibitor BI-3406, MDM2 inhibitor Nutlin-3a, and PI3K inhibitor Taselisib. The study's primary goal is to provide meaningful information for these three drugs regarding their efficacy and potency, validate their mechanism of action (MoA), learn about suitable clinical indications, possible drug combinations, and predictive biomarkers of sensitivity or resistance.

ProLiFiler – Testing anti-proliferative efficacy

Assay procedure of the Cell Proliferation Assay

The ProLiFiler assay is performed with a contact-free nano-dispensing system (Tecan D300E) requiring small amounts of your compound.



Assay procedure: Cells are seeded into 384 multi-well plates (A). The next day, compounds are added (B), and cells are incubated for 72 h (C). Subsequently, the cell viability dye CellTiter-Glo is added, and the luminescence is measured as a parameter for cell viability (D). IC_{50} values are determined from 8 concentrations in duplicates.

Origin of cell lines

The panel currently consists of 140 human tumor cell lines derived from 21 tumor types.

Tumor type	#	(sub)type
Breast	13	4 ER^+; 5 $ERBB2^+$; 4 TNBC
Uterus	6	2 cervix; 4 endometrium
Ovary	9	
Prostate	3	1 AR^+; 2 AR^-
Kidney	4	
Colorectal	11	
Stomach	9	
Pancreas	4	
Liver	2	2 hepatocellular carcinoma
Non-Small Cell Lung	27	17 adenocarcinoma; 5 squamous; 3 large cells; 2 unclassified
Small Cell Lung	4	
Melanoma	4	
Central Nervous System	10	5 glioblastoma; 5 neuroblastoma
Sarcoma	6	3 osteosarcoma; 2 Ewing; 1 soft-tissue
miscellaneous	4	1 bladder; 1 duodenum, 1 head & neck; 1 vulva (skin)
Leukemia	16	*3 ALL; 10 AML; 2 CLL; 1 CML
Lymphoma	3	*2 DBLC; 1 ACLC
Myeloma	5	
Grand Total	140	

Nutlin-3a showed moderate potency with a strong selectivity by inhibiting only some CLs.

Efficacy of Taselisib was at a submicromolar range and showed efficacy in a wide variety of CLs.

In our 2D system, anti-proliferative activity of BI-3406 was modest, only few CLs were sensitive.

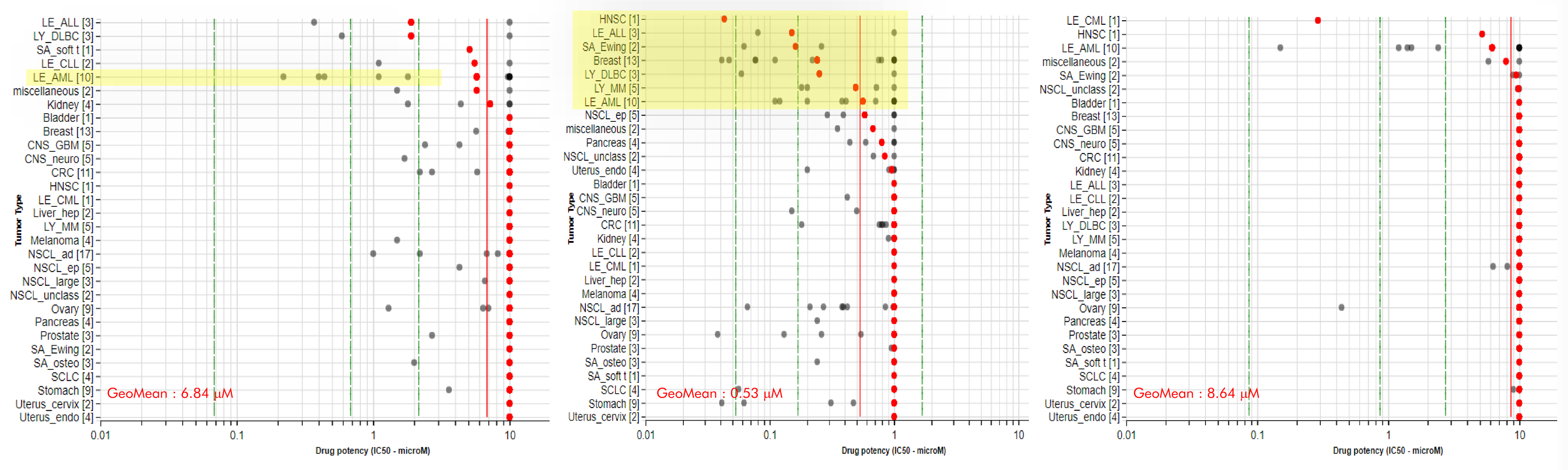


Figure 1: Scatter plots showing Nutlin-3a, Taselisib, and BI-3406 IC_{50} value for each cell line across cancer (sub)types of the 140 CL panel. X-axis: IC_{50} value per CL, y-axis: the histological (sub)types sorted from top to bottom by increasing median IC_{50} values. The red dots are the median Abs IC_{50} value for each tumor (sub)type, and the red line is the overall Geomean IC_{50} value. Between brackets the total number of CLs within a tumor (sub)type.

Abbreviations: CNS_Glioblastoma (CNS_GBM), CNS_Neuroblastoma (CNS_neuro); Colorectal cancer (CRC), Head-Neck Squamous Cell (HNSC), Acute Lymphoblastic Leukemia (LE_ALL), Acute Myeloid Leukemia (LE_AML), Chronic Lymphocytic Leukemia (LE_CLL), Chronic Myelocytic Leukemia (LE_CML) Liver_hepatocellular (Liver_hep), Lymphoma Diffuse Large B Cells (LY_DLBC), Lymphoma_Multiple Myeloma (LY_MM), Lymphoma_unclass (LY_unclass), NSCL_adenocarcinoma (NSCL_ad), NSCL_epidermoid (NSCL_ep), NSCL_Large cells (NSCL_large), NSCL_unclassified (NSCL_unclass), Sarcoma Ewing (SA_Ewing), Osteosarcoma (SA_osteo), Sarcoma soft tissue (SA_soft t), Small Cell Lung Cancer (SCLC), Uterus_endometrium (Uterus_endo).

MoA Finder – Identifying drugs with same inhibitory profiles and MoA

Drug Mechanism of Action

The MoA Finder tool aims to identify the mechanism of action (MoA) of your test compound based on the results obtained in the ProLiFiler cell panel screening. To this end, we compare the sensitivity profile of your drug candidate to more than 700 drugs for which the MoA is known.

- The MoA Finder tool allows the confirmation of the mechanism of action of drugs.
- The data help to identify potential off-targets and side effects.
- The MoA of compounds with unknown mechanisms can be identified and gives suggestions for target deconvolution in a phenotypic approach.

***Database:** The drug sensitivity profile, which defines the mechanism of action of more than 700 anticancer compounds, was collected from both literature and internal results.*

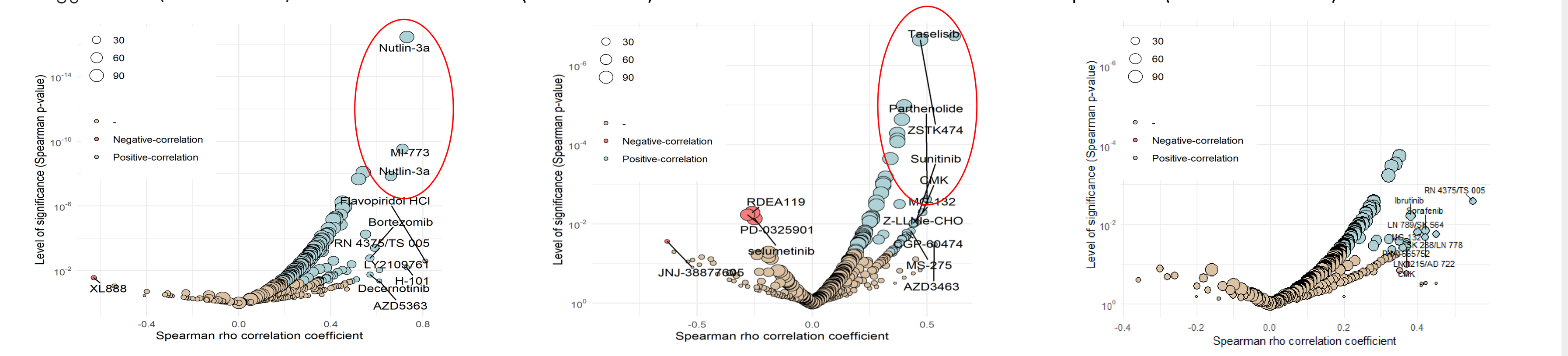
***Approach:** The drug sensitivity profile (IC_{50} values of ProLiFiler screen) of your test compound will be compared to the reference compounds' drug sensitivity profiles. The correlation of the two datasets is evaluated by two statistical values, the Spearman Rho correlation value, and its p-value.*

***Deliverables:** Graphical presentation of the correlation analysis. Excel table with raw data.*

Compounds targeting p53 pathway were those best correlating with our Nutlin-3a IC_{50} data (Rho>0.5).

Top 7 hits best correlating with 140 CL-derived Taselisib IC_{50} data were compounds targeting PI3K signaling (Rho>0.5).

No compound of the drug databases strongly correlated with BI-3406 anti-proliferative profile (all Rho<0.6).


Negative-correlation
Positive-correlation
Level of significance (Spearman p-value)
Spearman rho correlation coefficient
Nutlin-3a
MI-773
Bortezomib
RN 4375/TS 005
LY2109761
H-101
Decernotinib
AZD5363
XL888
Taselisib
ZSTK474
PD-0325901
selumetinib
JNJ-38877605
Z-LLNle-CHO
CGP-60474
MS-275
AZD3463
LN 789/SK 564
SK 288/LN 778
LN 215/AD 722
CMK

Rank	Compound	Mode Of Action	Rho	P-value	N	GeoMean (µM)	Below HalfLog
1	Nutlin-3a (-)	p53 pathway	1	-	140	6.84	30/140
2	Nutlin-3a (-)	p53 pathway	0.74	<0.001	97	33.99	363/867
3	MI-773	p53 pathway	0.71	<0.001	59	7.4	72/274
4	Nutlin-3a	p53 signaling	0.66	<0.001	59	15.71	73/274
5	Bortezomib	Protein degrad.	0.59	<0.001	32	0	0/407
6	RG-7112	p53 pathway	0.55	<0.001	47	6.59	63/243
7	CEP-701	RTK signaling	0.53	<0.001	101	0.54	251/869
8	Camptothecin	DNA replication	0.53	<0.001	102	0.02	0/869
9	Cytarabine	DNA replication, Other	0.48	<0.001	102	1.21	381/866
10	Veliparib	Genome integrity	0.46	<0.001	102	49.78	382/870
11	(5Z)-7-Oxozeaenol	other	0.45	<0.001	109	2.28	378/936

Rank	Compound	Mode Of Action	Rho	P-value	N	GeoMean (µM)	Below HalfLog
1	Taselisib	Inhibitor of PI3K α, β, δ, γ	1	-	140	0.53	26/140
2	Taselisib	Inhibitor of PI3K α, β, δ, γ	0.62	<0.001	59	2.14	116/274
3	ZSTK474	Inhibitor of PI3K β, δ, γ	0.47	<0.001	111	2.73	429/939
4	PIK-93	Inhibitor of PI3K α	0.41	<0.001	111	14.71	428/935
5	GDC0941	PI3K signaling	0.4	<0.001	112	2.66	843/1764
6	KIN001-102	PI3K signaling	0.37	<0.001	111	13.49	442/939
7	GSK690693	PI3K signaling	0.37	<0.001	110	35.24	371/935
8	GSK2126458	PI3K signaling	0.36	<0.001	110	0.05	0/937
9	CP466722	Genome integrity	0.34	<0.001	111	13.71	502/940
10	EKB-569	EGFR signaling	0.31	<0.001	110	4.58	426/936
11	QL-XII-47	other	0.31	<0.001	111	4.4	496/937

Rank	Compound	Mode Of Action	Rho	P-value	N	GeoMean (µM)	Below HalfLog
1	BI-3406	inhibitor SOS1:KRAS	1	-	140	8.64	11/140
2	Foretinib	RTK signaling	0.35	<0.001	111	1.3	425/932
3	XL-184	RTK signaling	0.34	<0.001	111	14.83	461/935
4	KIN001-055	other	0.33	<0.001	111	54.52	419/935
5	FR-180204	ERK MAPK signaling	0.32	<0.001	111	135.01	416/935
6	RN 4375/TS 005	Metabolism; casein Kinase	0.55	0.0027	28	2.48	32/80
7	TPCA-1	inhibitor of IKK-2, other	0.28	0.0027	111	20.35	457/935
8	Y-39983	cytoskeleton	0.28	0.0028	111	74.39	422/935
9	THZ-2-49	cell cycle	0.27	0.0036	111	14.64	476/932
10	PF-4708671	PI3K/MTOR signaling	0.28	0.0037	107	46.6	460/927
11	FMK	ERK MAPK signaling	0.28	0.004	104	151.16	346/806

Figure 2. Volcano plots showing the anti-cancer drugs correlated with Nutlin-3a, Taselisib and BI-3406 IC_{50} (COMPARE analysis*). x-axis: Rho values obtained (Spearman), y-axis: p-values. Blue: anti-cancer drugs with a positive correlation (p-values<0.05), red: negative correlation (p-values<0.05) (light brown: not significant). Drug sensitivity databases included in the study: 4HF Biotec, Sanger GDSC1, and CTRP-CTD^2 databases. The size of the circles is proportional to the numbers of data points compared. Tables show top anti-cancer agents best correlating with each of the three drugs studied.

*Huang, R., Wallqvist, A. & Covell, D.G. Anticancer metal compounds in NCI's tumor-screening database: putative mode of action. Biochem Pharmacol 69, 1009-1039 (2005).
*Paull, K.D. et al. Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. J Natl Cancer Inst 81, 1088-1092 (1989).

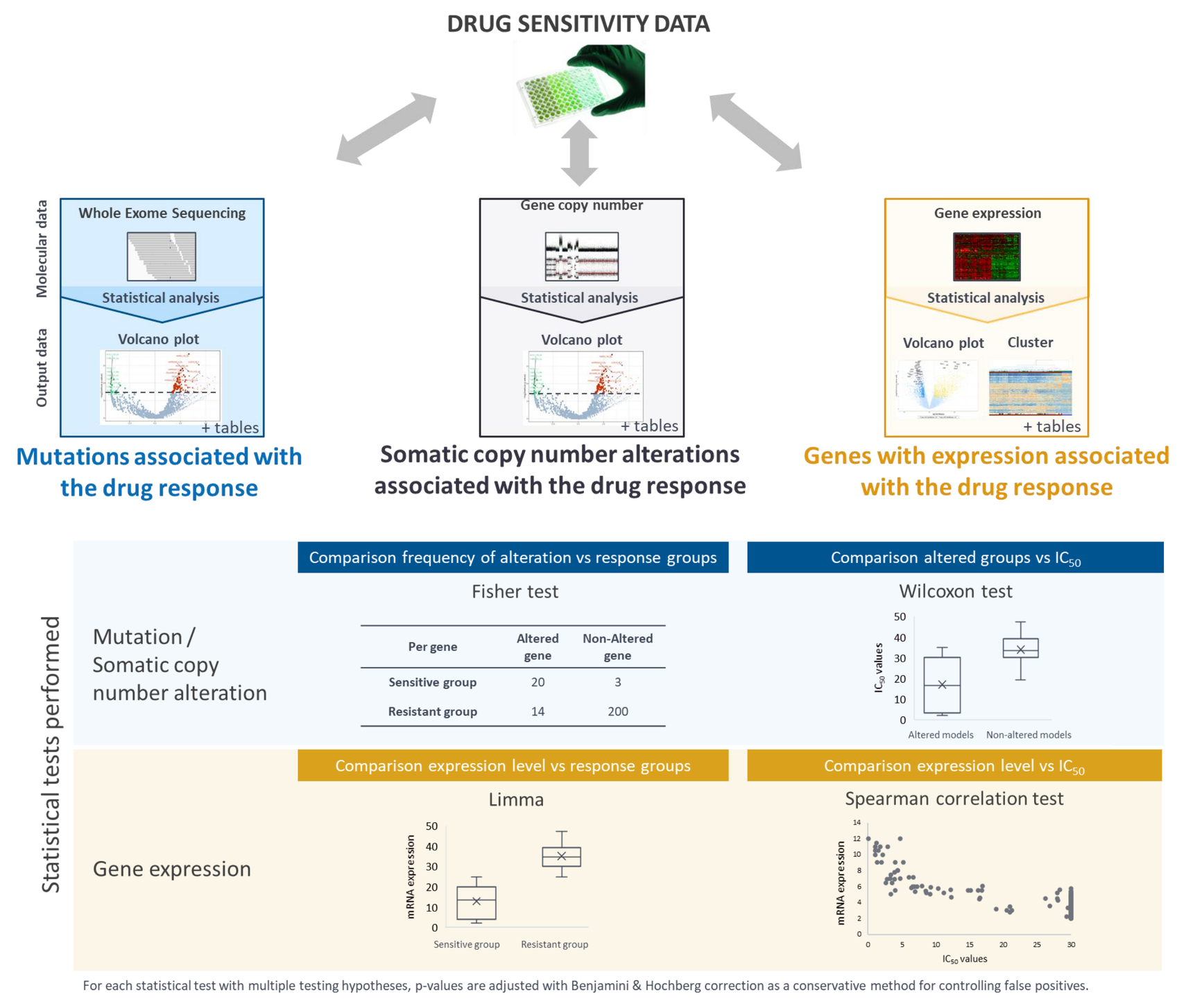
Biomarker Analysis – Discovery of predictors of drug sensitivity

Biomarker Analysis

The Biomarker Analysis tool reveals the specific genomic signature of tumor cell lines that are sensitive to your test drug. To this end, we correlate the drug sensitivity profile (IC_{50} data of ProLiFiler screen) of your test drug with data sets of gene expression and genetic alterations (mutations, deletions, amplifications) known for the tumor cell lines.

Use the Biomarker Analysis tool to identify genes or genetic alterations whose presence is indicative of the efficacy of the test compound and may serve for predictive biomarkers in disease models and patient cohorts.

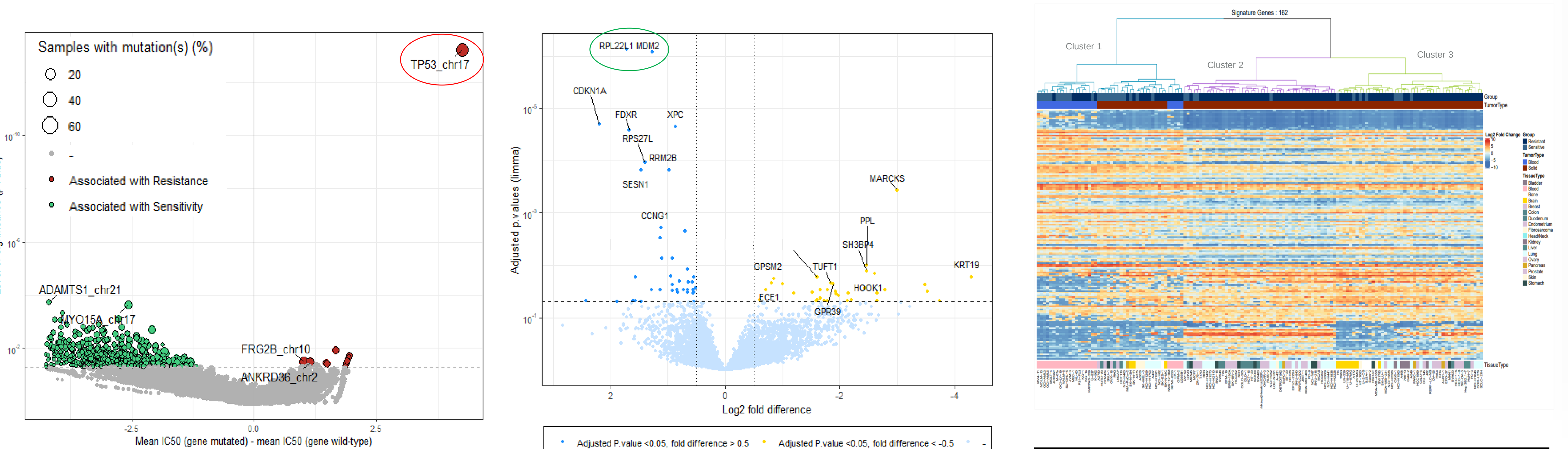
***Approach:** The ProLiFiler results are correlated with the gene expression, whole-exome sequencing data (mutations), and somatic copy number alteration data of the tumor cell lines. An analysis is performed for every cell line to determine the correlation of its drug sensitivity to each gene with regards to the expression level, mutation status, and somatic copy number.*


DRUG SENSITIVITY DATA
Whole Exome Sequencing
Gene copy number
Gene expression
Statistical analysis
Volcano plot
Mutations associated with the drug response
Somatic copy number alterations associated with the drug response
Genes with expression associated with the drug response
Fisher test
Wilcoxon test
Limma
Spearman correlation test
Gene expression
Statistical tests performed

The screening of whole exome mutations of the 140 CLs identified *TP53* mutations as top hit biomarker predicting resistance to Nutlin-3a.

The analysis of the transcriptome highlighted that *MDM2* expression level positively correlates with Nutlin-3a sensitivity.

Unsupervised hierarchical clustering using transcripts positively and negatively associated with Nutlin-3a IC_{50} identified 3 clusters with distinct response rates.


Samples with mutation(s) (%)
Associated with Resistance
Associated with Sensitivity
TP53_chr17
ADAMTS1_chr21
FRG2B_chr10
ANKRD36_chr2
Level of significance (p-value)
Mean IC50 (gene mutated) - mean IC50 (gene wild-type)
RPL22L1
MDM2
CDKN1A
FDXR
XPC
RPS27L
RRM2B
SESN1
CCNG1
MARCKS
PPL
SH3BP4
KRT19
Adjusted P values (limma)
Log2 fold difference
Adjusted P.value <0.05, fold difference > 0.5
Adjusted P.value <0.05, fold difference < -0.5
Total: 15125
Cluster 1
Cluster 2
Cluster 3

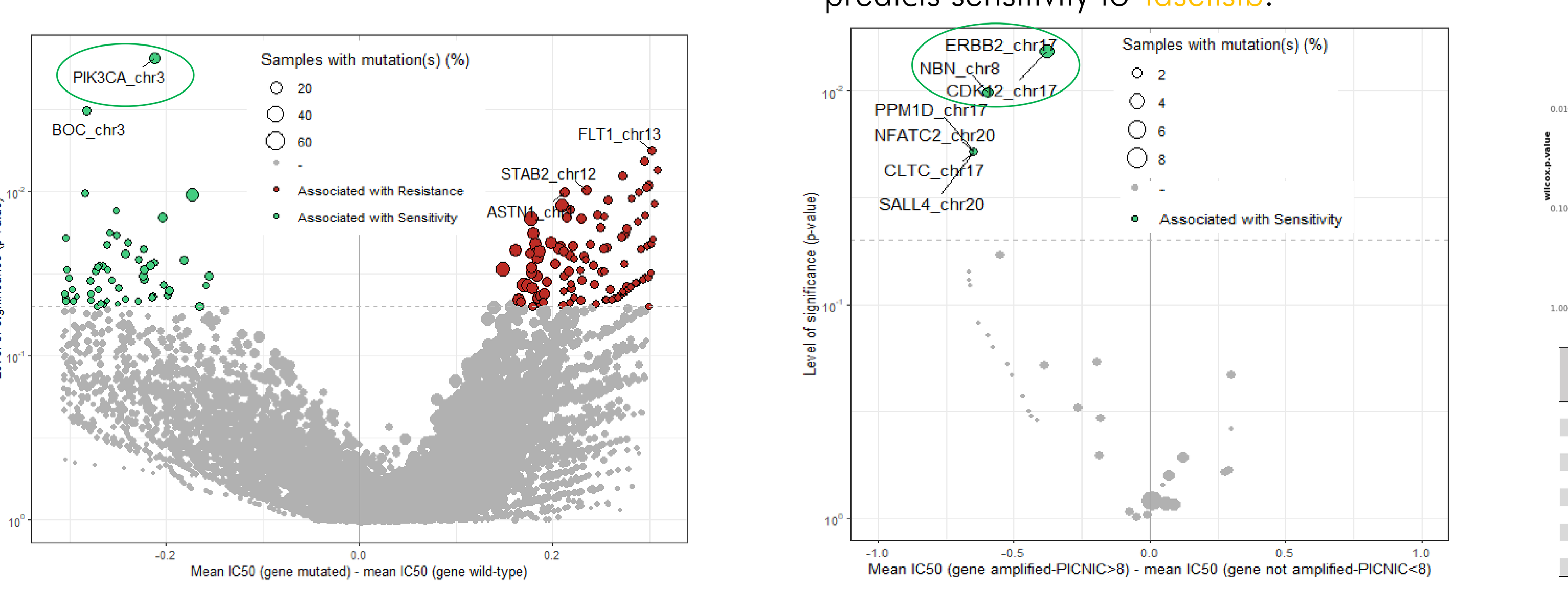
Clusters	1	2	3
Sensitive (IC_{50}<10µM)	23	5	8
Resistant (IC_{50}>10mM)	23	43	38
Sum	46	48	46
Distribution sensitive	**64%**	**14%**	**22%**
***P* value (propotion test)**	**1.89E-05**		

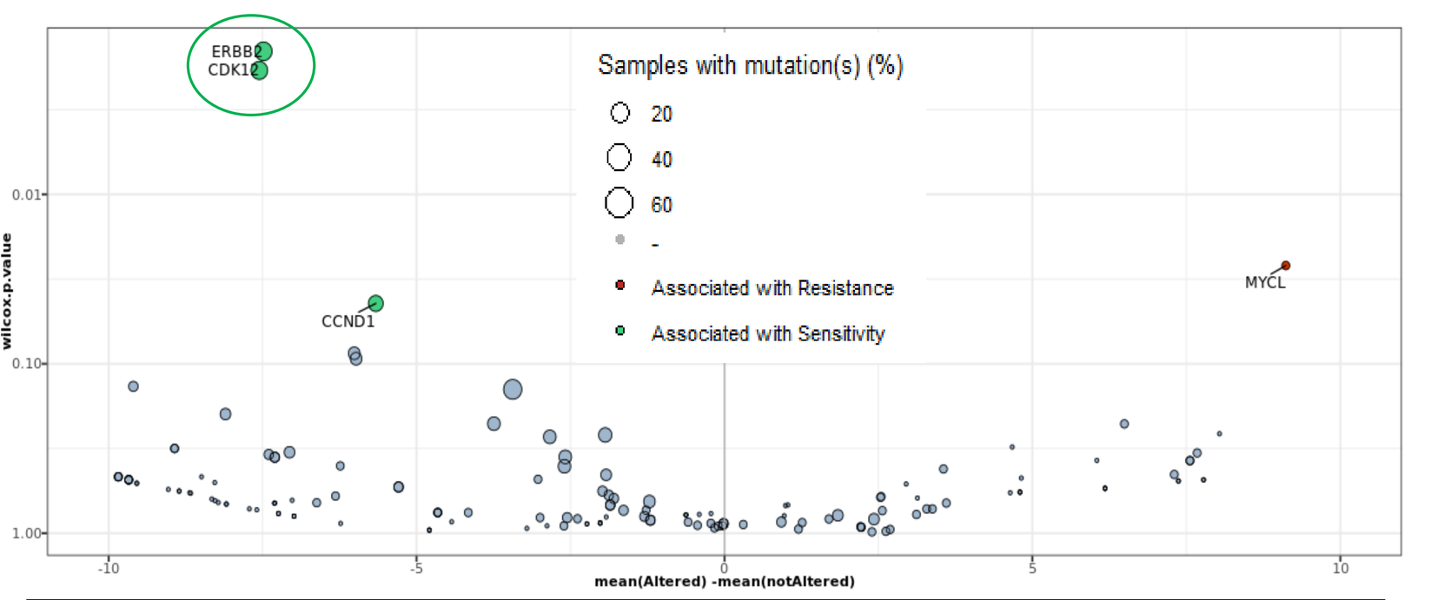
Figure 3a: **Left:** Volcano plot showing the top exome mutations positively and negatively associated with the Nutlin-3a IC_{50} values (Wilcoxon test). **Center:** Volcano plot showing levels of significance of individual genes for the association of their expression levels with Nutlin-3a IC_{50} values. X-axis, log_2 fold difference of gene expression level between Nutlin-3a response groups; y-axis, Limma adjusted p-values on log_{10} scale. The yellow and blue dots show the genes having an expression significantly associated with the response to Nutlin-3a (significant adjusted p-value in Limma test) and a log_2 fold difference >0.5 (blue dots) or <-0.5 (yellow dots) as calculated by Limma. **Right:** Heatmap with unsupervised hierarchical clustering of differentially expressed genes (Limma p-adjusted <0.05 , log_2 fold difference > ±0.5) and table with response rates to Nutlin-3a across the clusters.

Exome mutations in *PIK3CA* gene were the strongest predictors of CL sensitivity to Taselisib.

Whole Somatic Copy Number Alterations (SCNA) analysis showed that amplification of Chr17 part containing *ERBB2* gene predicts sensitivity to Taselisib.

Testing SCNA with Taselisib data from the GDSC2 dataset (Sanger institute) confirms *ERBB2+* as positive biomarker of response.


Samples with mutation(s) (%)
Associated with Resistance
Associated with Sensitivity
PIK3CA_chr3
BOC_chr3
FLT1_chr13
STAB2_chr12
Level of significance (p-value)
Mean IC50 (gene mutated) - mean IC50 (gene wild-type)
ERBB2_chr17
NBN_chr8
CDK12_chr17
PPM1D_chr17
NFATC2_chr20
CLTC_chr17
SALL4_chr20
Mean IC50 (gene amplified-PICNIC>8) - mean IC50 (gene not amplified-PICNIC<8)



Rank	Gene	Low IC50 Ampl.counts	High IC50 Ampl.counts	Low IC50 NoAmpl.counts	High IC50 NoAmpl.counts	Fisher p.value	Fisher p.value.adj	wilcox p.value	wilcox p.value.adj	Median taselisib IC50 (Ampl.)	Median taselisib IC50 (noAmpl.)
1	ERBB2_chr17	12	8	157	388	0.005	0.620	0.001	0.23	2.5	9.0
2	CDK12_chr17	11	4	158	392	0.001	0.136	0.002	0.23	2.4	9.0
3	MYCL_chr1	1	8	168	388	0.291	1	0.03	0.59	30.0	8.6
4	GRM3_chr7	3	0	166	396	0.026	1	0.04	0.59	1.9	8.7
5	CCND1_chr11	8	11	161	385	0.307	1	0.04	0.59	5.0	8.9
6	LZTR1_chr22	0	4	169	392	0.323	1	0.04	0.59	30.0	8.6
7	RARA_chr17	2	3	167	393	0.638	1	0.05	0.59	3.6	8.7
8	CCND2_chr12	0	2	169	394	1	1	0.06	0.59	30.0	8.6
9	CDH17_chr8	0	2	169	394	1	1	0.06	0.59	30.0	8.6
10	COX6C_chr8	0	2	169	394	1	1	0.06	0.59	30.0	8.6

Figure 3b: **Left:** Volcano plot showing the mutated genes positively and negatively associated with the Taselisib IC_{50} values (Wilcoxon test). **Center:** Volcano plot showing the gene amplifications positively and negatively associated with the Taselisib IC_{50} values (Wilcoxon test). **Right:** Volcano plot validating amplification of *ERBB2* gene as positive predictor of CL sensitivity to Taselisib in the in vitro dataset GDSC2-Sanger institute (Wilcoxon test).

Conclusions

The present pharmacogenomic study demonstrated the relevance of combined ProLiFiler and Cancer Data Miner platforms for preclinical profiling of novel anti-cancer agents. It permits early gain of information about tumor (sub)types to be targeted, to reveal or confirm the drug MoA and the identification of a relevant panel of biomarkers predicting response or resistance. We believe that such an approach should be more systematically made in programs of drug development to prepare the next steps of in vivo testing and clinical phases.