Master-regulators driving resistance of non-small cell lung cancer cells to p53 reactivator Nutlin-3

Boyarskikh UA[1], Pintus SS[2], Mandrik NV[2], Stelmashenko DE[2], Kiselev IN[2,3], Sharipov RN[2], Kolpakova AF[2,3], Stegmaier P[4], Filipenko ML[1], Kolpakov FA[2,3], Kel AE[1,2,4,*]

[1] Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia;

[2] BIOSOFT.RU, LLC, Novosibirsk, Russia;

[3] Institute of Computational Technologies, SB RAS, Novosibirsk, Russia;

[4] geneXplain GmbH, D-38302 Wolfenbüttel, Germany.

***** Author to whom correspondence should be addressed; e-mail:
alexander.kel@genexplain.com;
Tel.: +49-5331-992200-13; Fax: +49-5331-992200-22.

Summary

The tumor suppressor protein p53 is one of the most important targets in various cancers including non-small cell lung cancer (NSCLC). The small molecule Nutlin-3 reactivates p53 in cancer cells by interacting with the complex between p53 and its repressor Mdm-2, which leads to increased apoptosis of cancer cells. Nevertheless, clinical and experimental studies of Nutlin-3 have shown that in some cases cancer cells are not sensitive to this compound. In this paper we studied possible mechanisms of such a resistance of cancer cells to Nutlin-3 through increased activity of specific pro-survival pathways that appeared to be more active in the resistant cells, as compared to the cells sensitive to the treatment by this compound. Using genome-wide gene expression profiling we compared several NSCLC cell lines. Nutlin-3 resistant cell lines showed a significantly different gene expression in comparison with sensitive cell lines before and after treatment by Nutlin-3. Using original bioinformatics approaches we analyzed the revealed gene expression patterns and identified complexes of transcription factors specifically regulating different expression of genes in the resistant cell lines. Analysis of the signal transduction network upstream from transcription factors allowed us to identify potential master-regulators responsible for

maintaining this resistant state. Among them the most promising was mTOR acting in the context of the activated PI3K pathway. With the help of chemical inhibitors, we were able to validate these findings experimentally. The Nutlin-3 resistant cell lines showed the highest sensitivity to the inhibitor of mTOR with PI3K, which lead to the rapid death of these cells.

1. Introduction

Target therapy is one of the most promising approaches for the treatment of oncological diseases. It performs a targeted pharmacological interaction with specific proteins in the cancer or associated cells (so-called pharmacological molecular target). One of the most important targets in cancer cells is the well known tumor suppressor protein p53. Among known small molecules interacting with p53 the compound called Nutlin-3 gives rise to promising drug candidates for treatment of various cancers [1-2]. Through interaction with the complex between p53 and its repressor Mdm-2 the compound Nutlin-3 increases the activity of p53 which in turn leads to an increased apoptosis of cancer cells. Nevertheless, clinical and experimental studies of Nutlin-3 showed that in some cases cancer cells are not sensitive to this compound. One of the possible mechanisms of such resistance of cancer cells to Nutlin-3 is an increased activity of several pro-survival pathways that are regulated by specific signaling mechanisms in the cells. These pathways appeared to be more active in the resistant cells compared to the cells sensitive to the treatment by this compound. As a result, these tumor cells can avoid apoptosis due to a compensatory effect of these pathways despite the treatment with cytotoxic compounds. Such a highly active pro-survival pathway can either pre-exist in some types of tumor cells or become active upon treatment with cytotoxic compounds.

The group of drugs that reactivate the transcription factor p53 includes such interesting compounds as: nutlins (Nutlin-3); BDAs (benzodiazepinediones); A series of Mdm2 inhibitors (MI-63, MI-219 and MI-43); Tenovin (Tenovin-6); RITA, PRIMA-1, CP-31398, MIRA-1 and others. These compounds are small molecules with a common antitumor mechanism of action. Namely, the interaction of compounds with the p53 protein leads to the restoration of its functional activity and a subsequent induction of apoptosis in the tumor cell.

In this work, we study the effects of reactivation of p53 in the non-small cell lung cancer (NSCLC) by one of the most promising compounds - Nutlin-3.

The effect of Nutlin-3 leads to the restoration of p53 activity in the cell by inhibiting

the interaction of p53 with the protein Mdm2. The Mdm2 protein is an E3 ubiquitin ligase which plays a key role in the regulation of p53 [1]. In the cell, Mdm2 binds to the N-terminal domain of the p53 protein leading to the inactivation and the subsequent proteolytic ubiquitin-dependent degradation of p53. Very often, in tumor cells there is a substantial increase in the expression of Mdm2 [3]. As a result, in such tumor cells a rapid degradation of p53 occurs, which allows them to escape p53-dependent apoptosis. The destruction of the complex Mdm2-p53 promotes the stabilization of p53 and the restoration of its activity, which in turn leads to an inhibition of proliferation and / or death of tumor cells. To date, three classes of small organic molecules have been identified that, due to their structural features, are capable of inhibiting the interaction of Mdm2 and p53. These include heterocyclic compounds: Nutlins (nutlins) [4], BDAs (benzodiazepindiones) [5] and a series of Mdm2 inhibitors (MI) of spiro-oxindole derivatives including MI-63, MI-219 and MI-43 [6, 7]. All three series of compounds mimic the amino acid residues F19, W23 and L26 of the p53 protein and bind with high affinity to Mdm2 in the p53-specific pocket region and thus displace p53 from the Mdm2-p53 complex.

Of all the compounds inhibiting the interaction of the proteins Mdm2 and p53, nutlins are the best studied. Pre-clinical trial data of the treatment of acute myeloid leukemia with Nutlin-3a for [8, 9] has confirmed the ability of Nutlin-3a to induce apoptosis of tumor cells (but not the normal hematopoietic blood cells). Recently the small molecule drug RG7112 (Roche, Penzberg, Germany) (a derivative compound of Nutlin-3a, which was used as an early lead) was studied in a couple of Phase I clinical trials for advanced solid and hematological cancers, and liposarcoma [1] (clinical trial: NCT00559533), in acute myelogenous leukemia (clinical trial: NCT01635296) and soft tissue sarcoma (clinical trial NCT01605526 done by Roche) delivering reasonably promising results [2]. Still during such studies researchers often observed a decrease of efficiency in cancer inhibition due to the resistance of cancer cells to the treatment by Nutlin-3.

In order to analyze the effect of Nutlin-3 on non-small cell lung cancer and understand the mechanisms of resistance we performed an extensive study of the biological activity of the compounds Nutlin-3 on a number of NSCLC cell lines carrying wild-type TP53 gene. A total of 8 cell lines were analyzed: A549, NCI-H292, A427, COR-L23, DV-90, NCI-H1395, NCI-H1944, NCI-H2228. The cytotoxic activity of Nutlin-3 on NSCLC cells was determined based on the IC50 (inhibitory concentration) parameter (concentration of compound leading to 50% of cell death in a culture). The IC50 parameter was determined using a resazurin viability test. We found that NSCLC cell lines differ significantly from

each other in the sensitivity to Nutlin-3.

We chose one of the most Nutlin-3 sensitive cell lines - H1944, and two moderately resistant cell lines - A427 and NCI-H292 for further studies. We performed microarray experiments in these cell lines before and after treatment by Nutlin-3 in order to reveal different gene expression profiles in resistant cell lines. Comparative analysis of microarray data of these three cell lines was performed using a computational pipeline "From genome to target" (http://my-genome-enhancer.com) implemented using BioUML driven systems biology platform (www.biouml.org, www.genexplain.com). We revealed a number of differentially expressed genes (DEGs) between sensitive and resistant cell lines. Promoter and pathway analysis of these DEGs using the "upstream analysis" approach [10] helped us to identify potential master regulators in these cancer cell lines responsible for the elevated resistance to Nutlin-3. Among them, the most promising was mTOR as one of the most important regulators of pro-survival mechanisms in the cell. We applied specific chemical inhibitors in order to test their effect on these cell lines. We found that Nutlin-3 resistant cell lines exhibit the highest sensitivity to the dual chemical inhibitor of mTOR-PI3K whereas the Nutlin-3 sensitive cell line appeared to be relatively insensitive to this inhibitor. These results confirmed our prediction that the master regulators in the mTOR-PI3K signaling pathway are responsible for the elevated resistance of particular NSCLC cell lines to treatment by the p53-reactivating compound Nutlin-3. As we predicted, the Nutlin-3 resistant cell lines appeared to be highly sensitive to the inhibitors of mTOR-PI3K pathway. These findings open a promising possibility for a combinatory therapy combining Nutlin-3 with mTOR-PI3K inhibitors. Such drug combinations will have a potential to tackle the observed heterogeneity between different cancer cell linages towards p53-reactivators such as Nutlin-3.

2. Data, Materials and Methods

Cell lines. In our study we used the following eight cell lines of non-small cell lung cancer (A549, NCI-H292, A427, COR-L23, DV-90, NCI-H1395, NCI-H1944, and NCI-H2228). The characteristics of these eight cell lines are given in Table 1.

Table 1. NSCLC cell lines used in this work

Cell line	Type of cancer	P53 status
A427	NSCLC adenocarcinoma	wt

A549	NSCLC adenocarcinoma	wt
COR-L23	NSCLC large cell adenocarcinoma	wt
DV-90	NSCLC adenocarcinoma	wt
NCI-H1395	NSCLC adenocarcinoma	wt
NCI-H1944	NSCLC adenocarcinoma	wt
NCI-H2228	NSCLC adenocarcinoma	wt
NCI-H292	NSCLC mucoepidermal adenocarcinoma	wt

Conditions of cultivation. The cells were cultured in DMEM / F12 medium containing 10% FBS, 0.03% glutamine, and kanamycin at a concentration of 100 μg / ml at 37 ° C and 5% CO2 in the atmosphere. When the cell culture reached a density of ~ 70%, the monolayer cells were seeded with 1: 4 culture dilution. The described cultivation conditions allowed maintaining the cell culture in exponential growth.

Test of sensitivity of NSCLC cells to Nutlin-3. Cells from frozen cultures were seeded in 25cm^3-culture flasks. Before testing the compound, the cells must be reseeded at least once. To test the sensitivity to the compound treatment, the NSCLC cells in exponential growth phase were seeded in wells of a 24-well plate at 60,000 cells per well. After 24 hours, the culture medium was replaced with a fresh one containing the compound in the following concentrations: Nutlin-3 (Nutlin3 (±), Cayman Chemical) at concentrations of 34 μM, 17 μM, 8.5 μM, 4.25 μM, 2.2 μM and 0 μM (control).

Each concentration point was represented in triplicates. 48 hours after the addition of the compound the proportion of viable cells was determined by a viability test with resazurin.

Test for vitality with resazurin. Resazurin is an oxidation-reduction dye and is used to analyze the toxicity of compounds in cell culture [11]. The analysis is based on the ability of metabolically active cells to restore resazurin (a blue product) to resorufin (a pink color product, a fluorophore with absorption and emission maxima at 571 and 586 nm, respectively). The conversion of resazurin to resorufin is carried out intracellularly and is provided by mitochondrial, microsomal and cytosolic oxidoreductases. Thus, the amount of resorufin in the culture medium is related to the total number and (or) survival of the cells in the culture.

The viability test was performed 48 hours after the cell culture was incubated with the test compound. To do this, the culture medium with the test compound was removed from the wells of the plate and 500 μl of Dulbecco's solution (1 ×) containing resazurin (resazurin

sodium salt, Diaam) at a concentration of 50 μg / ml was added. A 500 μl Dulbecco solution with resazurin (50 μg / ml) added to an empty well of a 24-well culture plate was used as control (background). The incubation time of cells with resazurin was 1 hour. The incubation was carried out under sterile conditions, at 99% humidity, at 37° C and at a concentration of 5%. Accumulation of resorufin (reduced form of resazurin) was assessed by fluorescence detection in the range of 587/607 nm. To assess the results of the test, a calibration curve was constructed for the change in fluorescence as a function of the number of cells. Using this calibration curve, we determined the relationship between the intensity of fluorescence and the number of viable cells in culture. The proportion of surviving cells was used to construct the dose-effect curve for each cell line. The IC50 value was calculated from these curves obtained using the Probit Analysis 1.0 software [12].

Isolation of cellular RNA and synthesis of cDNA. The cells of the NSCLC lines were lysed with Trisol reagent (Invitrogen). The total RNA was isolated according to the recommendations of Invitrogen. The quality of RNA preparations was assessed by electrophoretic separation in a 1.3% agarose gel.

The reverse transcription reaction was performed at 42° C for 45 minutes in 20 μl of a reaction mixture containing 10 mM Tris-HCl (pH 8.3), 5 mM MgCl2, 10 mM DTT, 50 mM KCl, 0.2 mM dNTP, Stat-9 primer (10ng / Ml), 100 unit act., DNA-dependent RNA polymerase MoMLV (Biosan, ICBPM SORR) and 500ng of the total RNA.

Amplification reactions were performed using thermal cyclers with an optical unit for detecting the fluorescence of iQ5 iCycler or CFX96 (Bio-Rad). The possibility of recording fluorescence in real time was achieved by adding to the reaction mixture the intercalating dye SYBRGreen I. Amplifications of the cDNA of the RPL32 and b-ACT genes were carried out in a single reaction mixture (duplex-PCR). In this case Taq-man probes labeled fluorophores FAM and R6G, respectively, were used to register the fluorescence in real time. To obtain optimal conditions for PCR, the structures and combinations of primers, the composition of the amplification buffer, and the parameters of the amplification cycle were varied.

Microarray analysis

Microarray analysis was performed using the following microarray platform: Human HT-12 v3 Expression BeadChips (Illumina). Three cell lines: A427, H292 and H1944 were treated by Nutlin-3 in a cytotoxic concentration (30μM) and at a concentration that

maximally discriminates the sensitive and moderately resistant cell lines (5μM). The cells were incubated with the compound for 24 hours. The experiment was done in two biological replicates for each condition. In order to detect differentially expressed genes the raw microarray data were normalized and further analyzed using the Limma tools from the R/Bioconductor package integrated into the BioUML driven pipeline "From genome to target". Limma calculated the LogFC (the logarithm on the basis of 2 of the fold change between different conditions) and the p-value and adjusted p-value (corrected to the multiple testing). All these parameters were used to detect differentially expressed genes.

Analysis of Enriched Transcription Factor Binding Sites

Transcription factor binding sites in promoters of differentially expressed genes were analyzed using known DNA-binding motifs described in the TRANSFAC® library [13], release 2017.2 (geneXplain, Wolfenbüttel, Germany) (http://genexplain.com/transfac). The motifs are specified using position weight matrices (PWMs) that assign weights to each nucleotide in each position of the DNA binding motif for a transcription factor or a group of them.

The BioUML environment provides tools to identify transcription factor binding sites (TFBS) that are enriched in the promoter regions under study as compared to a background sequence set such as promoters of genes that were not differentially regulated under the condition of the experiment. We denote study and background sets briefly as Yes and No sets. The algorithm for TFBS enrichment analysis, called F-Match, has been initially described in [14]. Briefly, as it has been described in detail previously [10], the procedure finds a critical value (a threshold) for the score of each PWM in the library that maximizes the Yes/No ratio R_{YN} as defined in Equation (1) under the constraint of statistical significance.

$$R_{YN} = \frac{\#Sites_{Yes} / \#Sites_{No}}{\#Seq_{Yes} / \#Seq_{No}} \quad (1)$$

In Equation (1), *#Sites* and *#Seq* are the sites and sequences counted in Yes and No sets. A high Yes/No ratio indicates strong enrichment of binding sites for a given PWM in the Yes sequences. The statistical significance is computed as follows:

$$P(X \geq x) = \sum_{n=x}^{N} \binom{N}{n} \cdot p^n \cdot (1-p)^{N-n}$$
$$p = \#Seq_{Yes} / (\#Seq_{Yes} + \#Seq_{no})$$
$$N = \#Sites_{Yes} + \#Sites_{No} \quad (2)$$
$$n = \#Sites_{Yes}$$

This binding site enrichment analysis is carried out in the Site Analysis module of the pipeline "From genome to target". For further analysis we consider only those TFBSs that achieved a Yes/No ratio >1 and a P-value < 0.01. The workflow further maps the matrices to respective transcription factors, and generates visualizations of all results.

Finding Master Regulators in Networks

We searched for master regulator molecules in signal transduction pathways upstream of the identified transcription factors using the BioUML platform tools. The master-regulator search uses the TRANSPATH® database (http://genexplain.com/transpath) [15]. A comprehensive signal transduction network of human cells is built by the network analysis module of the BioUML platform using reactions annotated in TRANSPATH®. The main algorithm of the master regulator search has been described earlier [10]. The goal of the algorithm is to find nodes in the global signal transduction network that may potentially regulate the activity of the set of transcription factors found at the previous step of analysis. Such nodes are considered most potent drug targets, since any influence on such a node may switch the transcriptional programs of hundreds of genes that are regulated by the respective TFs. In our analysis, we have run the algorithm with a maximum radius of 12 steps upstream of the TFs.

Computational pipeline «From genome to target»

Comparative analysis of microarray data of these three cell lines was performed using a computational pipeline "From genome to target" (http://my-genome-enhancer.com) implemented on a BioUML-driven systems biology platform (www.biouml.org, www.genexplain.com). The pipeline can take various multi-omics data (such as Genomics, Transcriptomics, Epigenomics, Proteomics and Metabolomics) and automatically perform the "Upstream analysis" [10] to detect master-regulators as potential drug targets. It provides

a flexible graphical tool for describing meta-data of the experiment that helps researchers to define data for the automatic analysis. The pipeline consists of four modules: Statistics, Genome Enhancer, Drugs, Biomarkers, and Targets. Depending on the input data, meta-data and the chosen parameters, the system generates a tailor-made data analysis workflow going through all these four modules. In the first step, the data is statistically analyzed and the lists of differentially expressed genes (DEGs) (in case of Transcriptomics data), proteins (Proteomics) and metabolites (Metabolomics) are prepared for the next steps of analysis. In the cases of Epigenetic data (ChIP-seq, DNA methylation) a list of statistically significant peaks and CpG methylation sites is prepared for the next step. The Genomics data are also analyzed in this step and lists of revealed mutations are computed in the form of VCF files. At the Genome Enhancer module, the analysis of gene regulatory regions of differentially expressed genes is performed using the Match™ [10] and CMA [16] tools (which in turn use the TRANSFAC® [13], HOCOMOCO [17] and GTRD [18] databases of position weight matrices) in order to detect transcription factor binding sites (TFBSs) in the promoters and enhancers of DEGs. ChIP-seq peaks and CpG methylation sites are used in this module to help to define the enhancer and silencer regions of the differentially expressed genes. In addition, those mutations that were found inside enhancers and silencers are used to reveal transcription factor binding sites significantly affected by these mutations. As a result, a list of transcription factors regulating genes through the identified TFBSs is forwarded to the network analysis in the signal transduction network (using the TRANSPATH®, REACTOME, and HumanCyc databases). If genomic data is available, then the observed mutations in the coding regions of proteins, which are involved in signal transduction and which damage the function of these proteins, are taken into account. The network is then modified by exclusion of corresponding nodes and reactions from it. The network analysis algorithm reveals master regulators as described above. In pipeline modules “Drugs” and “Biomarkers and Targets” the revealed master-regulators are interrogated and prioritized in order to select the most promising therapeutic targets. Various properties of these master-regulator proteins, such as potential “drugability” (possibility to find known drugs or novel chemical compounds potentially interacting with these proteins) as well as additional annotation from the HumanPSD® (www.genexplain.com/humanpsd) database about known disease relevance of these proteins, are taken into account for such a prioritization. Finally, the pipeline “From genome to target” reports a short list of most promising targets.

3. Results

Sensitivity of different NSCLC cell lines to Nutlin-3

To test the sensitivity to Nutlin-3 we treated the selected NSCLC cell lines during 24 hours with the compound in the following concentrations: 34 µM, 17 µM, 8.5 µM, 4.25 µM, 2.2 µM and 0 µM (control).

With the help of the resazurin viability test we constructed the curves of percentages of survived cells under increasing concentrations of Nutlin-3 and identified IC50 values of Nutlin-3 in eight cell lines of non-small cell lung cancer (A549, NCI-H292, A427, COR-L23, DV-90, NCI-H1395, NCI-H1944, NCI-H2228). We obtained the following results.

The H1944 line was the most sensitive to the action of Nutlin-3 (IC50 = 4.9 ± 7.6). The other cell lines showed very different cytotoxic effects of Nutlin-3: H1395 (IC50 = 11.4 ± 3.8), DV-90 (IC50 = 12.9 ± 7.6), COR-L23 (IC50 = 15.1 ± 4.6), H292 (IC50 = 15.4 ± 2.6), A427 (IC50 = 18.8 ± 3.4). The A549 and H2228 cell lines were quite resistant to Nutlin-3, with IC50 values of 31.2 ± 5.2 and 33.1 ± 6.7, respectively. For further analysis we selected three cell lines: one, which was the most sensitive to Nutlin-3 – the H1944 cell line, and two, which were still reacting to Nutlin-3, but only under relatively high concentrations – H292 and A427 (moderately resistant). On Fig. 1 one can see the clear differences in the cell survival curve between the sensitive cell line and these two moderately resistant.

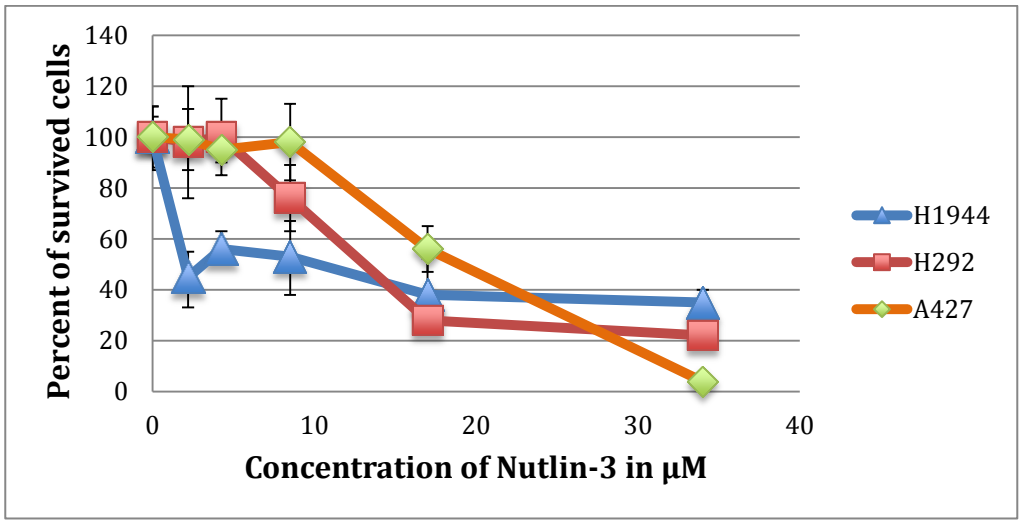

Figure 1. Viability test of three NSCLC cell lines with increasing concentrations of Nutlin-3. Cell line H1944 shows a higher sensitivity to the compound as compared to the cell lines H292 and A427.

Gene expression profiles of Nutlin-3 sensitive and moderately resistant NSCLC cell lines

In order to get a comprehensive gene expression profile of the cell lines under study before and after treatment by Nutlin-3 we applied Illumina microarrays (Human HT-12 v3 Expression BeadChips). We treated three cell lines: A427, H292 and H1944 by Nutlin-3 in a cytotoxic concentration (30μM) and in a concentration, which maximally discriminates the sensitive and moderately resistant cell lines (5μM). After data normalization, we applied Limma tools (R/Bioconductor package integrated into BioUML as a modules of the pipeline "From genome to target") and compared gene expression in the moderately resistant cell lines (A427 and H292) with gene expression in the sensitive cell line (H1944) before treatment and after treatment by two concentration of Nutlin-3. Limma has calculated the LogFC (the logarithm on the basis of 2 of the fold change between different conditions), the p-value and the adjusted p-value (corrected to the multiple testing) of the observed fold change. In Table 2 one can see the results of detection of Up- and Down- regulated genes (adj.p-value<0.05) in all three conditions.

Table 2. The number of differentially expressed genes (UP and DOWN) in Nutlin-3 moderately resistant versus sensitive cell lines in the control and under increasing dosage of Nutlin-3.

Nutlin-3 dosage (μM)	**UP**	**DOWN**	**UP+DOWN**
0 (Control)	872	760	1632
5	454	397	851
30	252	292	544

One can see that the untreated cells are characterized by rather high numbers of DEGs. With increasing concentration of Nutin-3 the number of differently responding genes between these cell lines is becoming smaller.

Comparison between revealed differentially expressed genes in three conditions is shown by Venn diagrams (Figure 2). One can see a significant overlap between all three conditions showing that the cell line type determines quite a big portion of the gene expression profile. Also, it is interesting to see that a lot of genes with significant initial differences in their expression levels between sensitive and resistant cell lines become similarly expressed after the treatment with Nutlin-3 (546 up-regulated and 391 down-

regulated genes). We think that these genes reflect the inducible part of the molecular mechanism of the cell response to treatment, which is rather common between different cell lines.

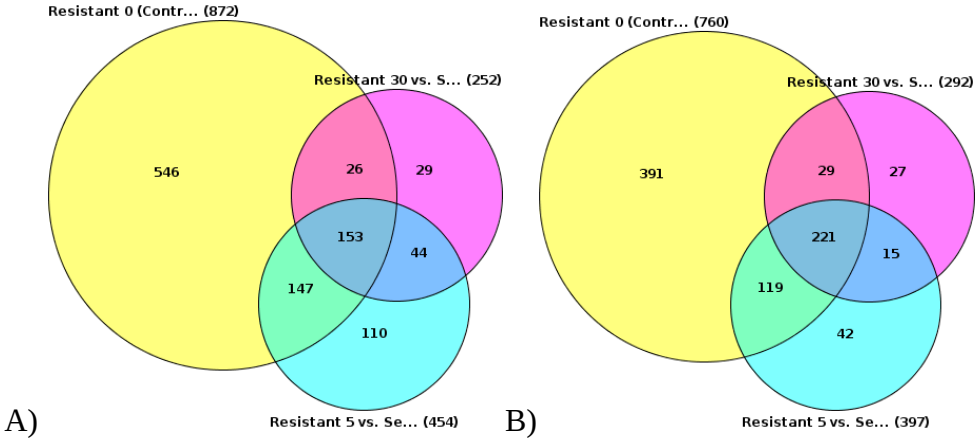

Figure 2. Venn-diagrams of the differentially expressed genes of Nutlin-3 moderately resistant cell lines versus sensitive cell lines after treatment by Nutlin-3 in two different dosages (5μM and 30μM) and in the control. A) Up-regulated genes; B) Down-regulated genes.

The revealed Up- and Down- regulated genes were mapped to different ontologies such as GO, pathways (TRANSPATH®), and diseases (HumanPSD®). The most statistically significant categories are presented in the Table 3 and in Figure 2. We also see a rapid change of the activated pathways with an increased dosage of treatment by Nutlin-3.

Table 3. Mapping of Up- and Down regulated genes (before treatment by Nutlin-3) on various Ontologies (GO, pathways (TRANSPATH®), diseases (HumanPSD®).

Ontology	Terms
	Up-regulated
GO (biological function)	cell cycle process (p<5.93E-50), cell cycle (p<5.52E-48), mitotic cell cycle process (p<3.82E-46), mitotic cell cycle (p<2.34E-42), mitotic cell cycle phase (p<2.70E-35), cell cycle phase (p<1.76E-34), DNA metabolic process (p<5.12E-31), organelle organization (p<1.94E-29), single-organism organelle organization (p<2.36E-27), cellular macromolecule metabolic process (p<3.38E-27), cellular component organization or biogenesis (p<1.10E-25), cell division (p<2.18E-25), cellular component organization (p<1.12E-24), nucleobase-containing compound metabolic process (p<1.54E-24), chromosome organization (p<1.55E-24), mitotic nuclear division (p<2.77E-24), nucleic acid metabolic process (p<3.25E-24), macromolecule metabolic process (p<3.69E-23), heterocycle metabolic process (p<3.69E-23), DNA replication (p<5.32E-23), nuclear division (p<1.26E-22), cellular aromatic compound metabolic process (p<1.85E-22), gene expression (p<4.75E-22), macromolecular complex subunit organization (p<5.32E-22), cellular metabolic process (p<9.19E-22), cellular nitrogen compound metabolic process (p<1.32E-21), mitotic M phase (p<8.69E-21), M phase (p<8.69E-21), organelle fission (p<9.45E-21), organic cyclic compound metabolic process (p<2.10E-19), cell cycle phase transition (p<2.13E-19), mitotic cell cycle phase transition (p<5.69E-19), cell cycle checkpoint (p<1.21E-18),

TRANSPATH pathways	GDP ---> dGTP (p<5.75E-07), ADP ---> dATP (p<5.75E-07), 2'-deoxycytidine-5'-phosphate ---> DNA-P-UC (p<2.44E-06), 2'-deoxyuridine 5'-phosphate ---> deoxythymidine 5'-triphosphate (p<3.66E-06), interconversion and degradation of purine deoxyribonucleotides (p<1.01E-05), Aurora-B cell cycle regulation (p<3.41E-05), Metaphase to Anaphase transition (p<8.94E-05), Aurora-A activation, substrates and degradation (p<2.97E-04), interconversion and degradation of pyrimidine nucleotides and bases (p<3.05E-04), Aurora-A cell cycle regulation (p<3.47E-04), E2F network (p<4.04E-04), HIF-1alpha pathway (p<5.13E-04), Plk1 cell cycle regulation (p<0.0016), AR pathway (p<0.0025), S phase (Cdk2) (p<0.0025), parkin associated pathways (p<0.0036), VHL ---> HIF-1alpha degradation (p<0.0043), G1 phase (Cdk2) (p<0.0047), cyclosome regulation (p<0.0087), G2/M phase (cyclin B:Cdk1) (p<0.0088), cyclosome regulatory network (p<0.0106), ER-alpha pathway (p<0.0106), p53 pathway (p<0.0211)
HumanPSD diseases	Adenocarcinoma, Follicular:Correlative (p<2.29E-09), Colonic Neoplasms:Correlative (p<4.97E-05), Adenocarcinoma:Correlative (p<5.69E-04), Triple Negative Breast Neoplasms:Causal (p<8.52E-04), Neoplasms by Site:Preventative (p<9.29E-04), Carcinoma, Papillary, Follicular:Correlative (p<0.0015), Ovarian Diseases:Correlative (p<0.0015), Neoplasms, Glandular and Epithelial:Correlative (p<0.0016), Adnexal Diseases:Correlative (p<0.0017), Glioblastoma:Preventative (p<0.0017), Genital Neoplasms, Female:Preventative (p<0.0018), Ovarian Neoplasms:Correlative (p<0.0020), DNA Repair-Deficiency Disorders:Correlative (p<0.0020), Wolf-Hirschhorn Syndrome:Causal (p<0.0020), Astrocytoma:Preventative (p<0.0025), Otorhinolaryngologic Neoplasms:Preventative (p<0.0025), Carcinoma:Correlative (p<0.0029), Head and Neck Neoplasms:Correlative (p<0.0030), Breast Neoplasms:Correlative (p<0.0032), Breast Diseases:Correlative (p<0.0032)
Down-regulated	
GO (biological function)	single-organism metabolic process (p<1.50E-07), oxoacid metabolic process (p<1.55E-07), small molecule metabolic process (p<2.16E-07), organic acid metabolic process (p<2.16E-07), metabolic process (p<2.11E-06), carboxylic acid metabolic process (p<4.68E-06), primary metabolic process (p<1.94E-05), response to endoplasmic reticulum stress (p<4.46E-05), organic substance metabolic process (p<4.69E-05), cellular metabolic process (p<7.97E-05), lipid metabolic process (p<7.97E-05), cellular lipid metabolic process (p<2.20E-04), Golgi vesicle transport (p<7.49E-04), autophagy (p<0.00125), cellular response to external stimulus (p<0.00125)
TRANSPATH pathways	PDGF B ---> STATs (p<0.00314), H2SO4 ---> AMP (p<0.00314), IL-10 ---> STAT1alpha, STAT3 (p<0.00527), IL-22 ---> STAT1alpha, STAT3 (p<0.005274), IL-10 pathway (p<0.00527), IL-22 pathway (p<0.00527), sulfur metabolism (p<0.00809), ERK1 ---/ Tau (p<0.00922), hipk2 ---> C/EBPbeta (p<0.00922), metabolism of amino sugars (p<0.01320), vitronectin ---> PAK4 (p<0.01776), autophagy (p<0.0187), D-glucosamine-6-phosphate ---> CMP-N-acetylneuraminate (p<0.0210), acetyl-CoA, acetoacetyl-CoA ---> cholesterol, fatty acid (p<0.0235), cholesterol metabolism (p<0.0235)
HumanPSD diseases	Ovarian Neoplasms:Correlative (p<4.61E-05), Vascular Diseases:Causal (p<5.47E-05), Ovarian Diseases:Correlative (p<7.71E-05), Adnexal Diseases:Correlative (p<8.47E-05), Psoriasis:Correlative (p<1.08E-04), Skin Diseases, Papulosquamous:Correlative (p<1.22E-04), Endocrine Gland Neoplasms:Correlative (p<1.37E-04), Prostatic Neoplasms:Correlative (p<1.96E-04), Leukemia, Myeloid:Correlative (p<2.15E-04), Prostatic Diseases:Correlative (p<2.63E-04), Genital Neoplasms, Male:Correlative (p<2.68E-04), Cerebral Hemorrhage:Causal (p<2.74E-04), Gonadal Disorders:Correlative (p<2.78E-04), Neoplasms:Causal (p<2.98E-04), Glioma:Causal (p<5.81E-04), Proteostasis Deficiencies:Correlative (p<6.85E-04), Pancreatic Diseases:Correlative (p<8.70E-04), Motor Neuron Disease:Correlative (p<8.78E-04), TDP-43 Proteinopathies:Correlative (p<8.78E-04), Amyotrophic Lateral Sclerosis:Correlative (p<9.25E-04), Cardiovascular Diseases:Causal (p<9.30E-04),

On Figure 3 we selected several pathways that demonstrated an interesting dynamic in the change of significance (measured as -Log2(p-value)) of the moderately resistant cell lines in comparison with sensitive cell lines in different conditions with increased dosage of treatment by Nutlin-3. One can see that such pathways as “E2F network” are relatively stable in their significance, which demonstrates that genes belonging to this pathway have shown different expression in the resistant and sensitive cell lines independent of the stimuli. Such pathways as HIF-1alpha, Aurora-A activation and Metaphase to Anaphase transition showed a sharp decrease of significance in the cells under treatment by Nutlin-3. Expression of the genes that belong to these pathways appeared to be different between resistant and sensitive cell lines, but it became similar after treatments by Nutlin-3.

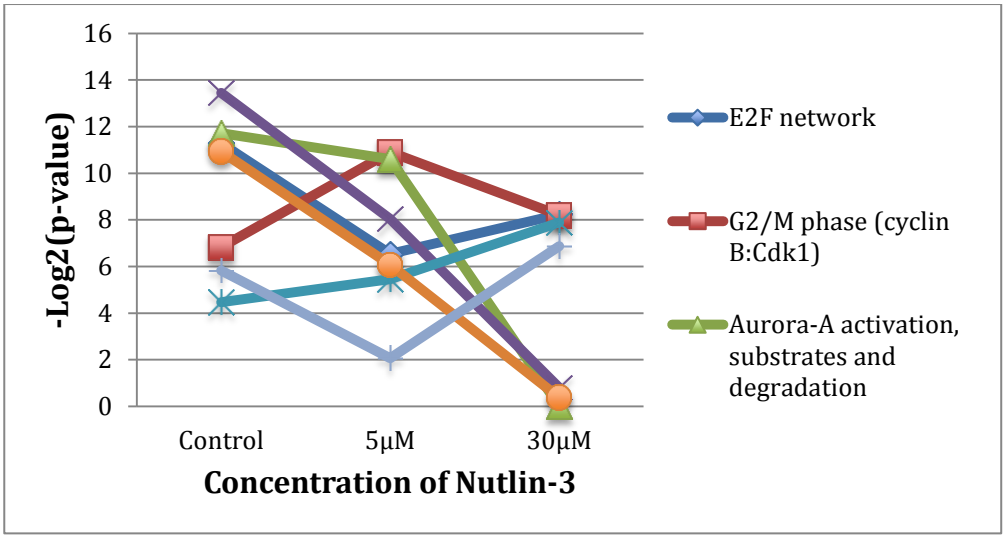

Figure 3. Dynamics of the change of significance of some important pathways in the moderately resistant cell lines in comparison with sensitive cell lines in different conditions with increased dosage of treatment by Nutlin-3.

Enriched Transcription Factor Binding Sites in promoters of Nutlin-3 moderately resistant cell lines

Transcription factor binding sites in promoters of differentially expressed genes were analyzed using known DNA-binding motifs described by PWMs in the TRANSFAC® library, release 2017.2 (geneXplain, Wolfenbüttel, Germany) (http://genexplain.com/transfac). In this study, we used the «Genome Enhancer» module of the pipeline «From genome to target» to identify transcription factor binding sites (TFBS) that are enriched in the promoter regions under study as compared to a background sequence set such as promoters of genes that were not differentially regulated under the conditions of the experiment. We denote the study and background sets briefly as Yes and No sets, respectively. The «Genome Enhancer» module uses two algorithms for TFBS enrichment analysis: F-Match [19] and CMA (Composite Module Analyst) [16]. The F-Match algorithm finds a critical value (a threshold) for the score of each PWM in the library that maximizes the Yes/No ratio under the constraint of statistical significance (see Materials and Methods). The CMA algorithm applies a genetic algorithm approach to find combinations of PWMs for the co-localized sites in promoters of DEGs. Such combinations of PWMs define groups of transcription factors that bind to the promoters of these genes in a synergistic manner (forming enhanceosomes [20]) and regulate the expression of their target genes in very specific conditions.

In this work, we applied these algorithms to analyze promoters of Up-regulated genes in Nutlin-3 moderately resistant NSCLC cell lines (A427 and H292) in comparison to the Nutlin-3 sensitive cell line H1944. We focused our attention on the up-regulated genes in order to reveal the potential resistance-specific positive feedback loops (as it is described in our previous paper [10]) that are maintained mainly through gene up-regulation. For the promoter analysis, we restricted the list of up-regulated genes by two criteria: The FC>1.5 (which corresponds to LogFC>0.58) and CI0.95>2.0 (confidence interval 0.95). The second criterion provides higher evidence that these genes are indeed up-regulated and it is applied here due to the very low number of replicas used (only 2 replicas were used in each condition, which is quite the lowest limit for Limma). As a result, we came up with a list of 179 up-regulated genes. As a control for the promoter analysis we selected 1000 genes that did not significantly change their expression between compared conditions (-0.1<LogFC <0.1).

In order to characterize this more restricted set of upregulated genes a bit further we mapped them to the disease ontology from the HumanPSD® databases and found a very high enrichment of genes related to different types of cancer (the most statistically significant match is the set of 71 genes to the category "Causative biomarkers of Neoplasms" (p-value<10^{-10}). 65 of these genes belong to the biomarkers of Lung neoplasms (p-value < 10^{-4}). We constructed the heatmap for these genes summarizing their expression profiles throughout all obtained microarray data (Fig. 4). One can clearly see the difference between gene expression profiles of the Nutlin-3 sensitive cell line (H1944 in the center of the heatmap) compared to the two moderately resistant cell lines (A427 – left side and H292 – right side of the heatmap).

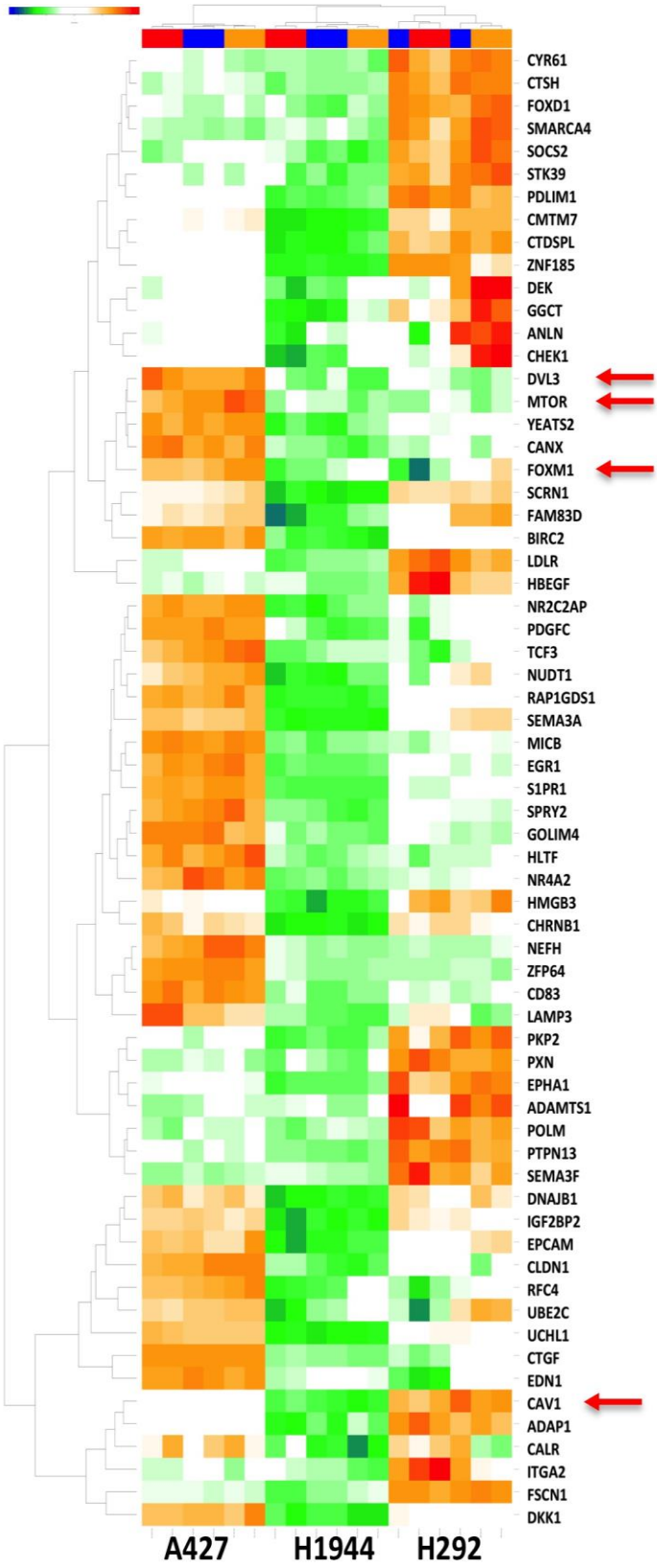

Figure 4. Heatmap of 65 genes belonging to the category of Lung Neoplasm biomarkers (HumanPSD®) that were up-regulated in moderately Nutlin-3 resistant NSCLC

cell lines (A427 and H292) in comparison to the Nutlin-3 sensitive cell line H1944. The top colored bar represents the samples with different treatment conditions: red – 30μM of Nutlin-3; blue - 5μM of Nutlin-3; orange – Control. Red arrows indicate four genes that were selected as potential therapeutic targets as the result of our "Upstream analysis" approach (see in the next sections).

Promoters were extracted from the human genome (build hg38) -1000 nucleotide upstream of TSS (start of transcription) and +100 downstream. Results of F-Match analysis are presented in Table 4 below.

Table 4. Results of F-Match analysis of 179 up-regulated genes in moderately Nutlin-3 resistant NSCLC cell lines.

ID	Yes density per 1000bp	No density per 1000bp	Yes-No ratio	PWM cutoff	p-value
V$CHCH_01	12.864	10.473	1.228	0.9874	2.70E-20
V$MAZ_Q6_01	3.154	2.197	1.436	0.8832	5.28E-15
V$SP1_Q6_01	4.134	3.031	1.364	0.9072	8.72E-15
V$GKLF_Q4	13.215	11.204	1.179	0.9535	3.09E-14
V$ZFP161_04	15.322	13.207	1.160	0.705	1.49E-13
V$CPBP_Q6	10.518	8.946	1.176	0.9972	2.59E-11
V$AP2ALPHA_03	18.593	16.500	1.127	0.7154	4.12E-11
V$EGR1_Q6	1.478	1.051	1.406	0.9123	2.95E-07
V$GLI_Q3	4.093	3.400	1.204	0.8974	1.50E-06
V$HBP1_03	0.071	0.010	7.092	0.9981	2.97E-06
V$RNF96_01	1.635	1.258	1.300	0.8951	2.05E-05
V$P53_04	9.634	8.742	1.102	0.7651	6.80E-05
V$RREB1_01	0.193	0.094	2.056	0.8679	2.38E-04
V$SRF_Q5_02	0.041	0.006	6.368	0.952	6.73E-04
V$FPM315_01	1.356	1.085	1.249	0.8954	7.30E-04
V$MAZR_01	0.020	0.001	22.289	0.9886	0.00235
V$REST_01	0.081	0.033	2.477	0.8308	0.00353
V$DR4_Q2	0.127	0.065	1.962	0.8775	0.00416
V$SP100_04	0.041	0.010	4.053	0.9607	0.00447
V$MEIS1_01	0.574	0.431	1.331	0.9861	0.00451
V$NKX25_Q6	0.549	0.409	1.340	0.9707	0.00454
V$NANOG_01	1.056	0.862	1.225	0.868	0.00516
V$NF1_Q6	0.030	0.006	4.776	0.9905	0.00808
V$RELA_Q6	0.152	0.088	1.723	0.9447	0.00836

ID column gives the PWM ID from the TRANSFAC® database; Yes and No site densities per 1000bp show the frequency of sites found in Yes promoters and No promoters respectively; Yes-No ratio is the ratio between frequencies of the sites in Yes promoters versus No promoters. PWM cutoff – the cutoff of the PWM optimized

by the F-Match algorithm; p-value – statistical significance of the ratio of site frequencies between Yes and No promoters.

We also have applied the CMA (Composite Module Analyst) algorithm, which allows to predict the formation of complexes of transcription factors binding sites that could jointly regulate groups of genes, such as up-regulated in our study. As a result, we identified the potential complexes of TFs that may synergistically bind to promoters of these genes and maintain their elevated expression, which in turn leads to the observed relatively high resistance of these cell lines to Nutlun-3.

Results of the CMA analysis are shown on the Figure 5 below.

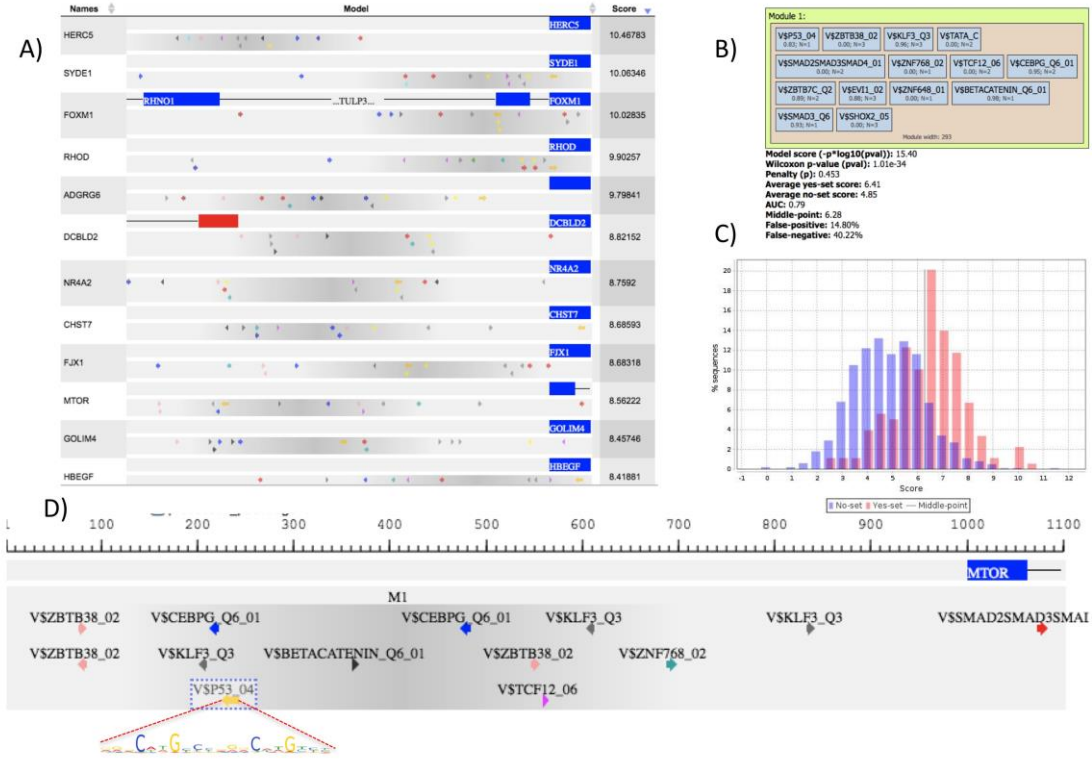

Figure 5. Results of CMA analysis of up-regulated genes in moderately Nutlin-3 resistant cell lines. A) Representation of promoters of top genes with maximal composite score (last column “Score”) containing the maximal number of TFBSs of TFs selected by the genetic algorithm. Sites are shown as colored arrows along the promoters. The blue bars represent the first 100bp of the first exons of the respective genes. B) Combination of PWMs with their optimized cut-offs identified by genetic algorithm. C) The discriminative parameters of the positional weight matrices (PWM) composition (p-value of the Wilcoxon test, AUC, rates of false positives and false negatives) and two distribution histograms of the

composite score values in Yes and No promoters. D) An example of the site location in the promoter of MTOR gene. The p53 site is marked and the logo of the p53 motif is shown below.

We identified a combination of PWMs that discriminates the promoters of Yes and No sets rather well. This combination contains PWMs for such important cancer-related transcription factors as p53, beta-catenin, Smad, Tcf12, Klf3, C/EBP and others. The full list of 25 transcription factor genes encoding TFs of selected combinations of PWMs is given in Table 5. Many of them are associated with various Neoplasms including Lung Neoplasms.

Table 5. The list of 25 transcription factor genes that are linked to the PWM combination found by CMA in the promoters of up-regulated Nutlin-3 moderately resistant genes.

Gene symbol	**Gene description**	**PWMs**	**RNA metabolic process**	**Lung Neoplasms**	**Neoplasms**	**TGFbeta pathway**	**p53 pathway**	**wnt pathway**
CTNNB1	catenin beta 1	V$BETACATENIN_Q6_01	+	+	+			+
CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma	V$CEBPG_Q6_01	+	+	+			
MECOM	MDS1 and EVI1 complex locus	V$EVI1_02	+	+	+	+		
KLF3	Kruppel-like factor 3 (basic)	V$KLF3_Q3	+	+	+			
TP53	tumor protein p53	V$P53_04	+	+	+		+	
SMAD2	SMAD family member 2	V$SMAD2SMAD3SMAD4_01	+	+	+	+		
SMAD4	SMAD family member 4	V$SMAD2SMAD3SMAD4_01	+	+	+	+		
SMAD3	SMAD family member 3	V$SMAD2SMAD3SMAD4_01,V$SMAD3_Q6	+	+	+	+		
TBP	TATA-box binding protein	V$TATA_C	+	+	+			+
TAF9	TATA-box binding protein	V$TATA_C	+				+	

	associated factor 9							
BTAF1	B-TFIID TATA-box binding protein associated factor 1	V$TATA_C	+		+			
TAF1	TATA-box binding protein associated factor 1	V$TATA_C	+	+	+			
TAF11	TATA-box binding protein associated factor 11	V$TATA_C	+		+			
TAF12	TATA-box binding protein associated factor 12	V$TATA_C	+		+			
TAF13	TATA-box binding protein associated factor 13	V$TATA_C	+		+			
TAF6	TATA-box binding protein associated factor 6	V$TATA_C	+	+	+			
TAF10	TATA-box binding protein associated factor 10	V$TATA_C	+					
TAF4	TATA-box binding protein associated factor 4	V$TATA_C	+					
TAF5	TATA-box binding protein associated factor 5	V$TATA_C	+					
TAF7	TATA-box binding protein associated factor 7	V$TATA_C	+					
TCF12	transcription factor 12	V$TCF12_06	+		+			
ZBTB38	zinc finger and BTB domain containing 38	V$ZBTB38_02	+		+			

ZBTB7C	zinc finger and BTB domain containing 7C	V$ZBTB7C_Q2			+			
ZNF648	zinc finger protein 648	V$ZNF648_01	+					
ZNF768	zinc finger protein 768	V$ZNF768_02	+					

The last six columns show with which pathways or GO terms and diseases the selected genes are associated.

Finding Master Regulators in Networks upstream of TFs.

We searched for master regulator molecules in signal transduction pathways upstream of the identified transcription factors. The master-regulator search uses the TRANSPATH® database (http://genexplain.com/transpath) [15] and is implemented in the workflow "From genome to target" as a last step. The main algorithm of the master regulator search has been described earlier [10]. The goal of the algorithm is to find nodes in the global signal transduction network that may potentially regulate the activity of the set of transcription factors found at the previous step of analysis. Such nodes are considered the most potent drug targets, since any influence on such a node may switch the transcriptional programs of hundreds of genes that are regulated by the respective TFs. In our analysis, we have run the algorithm with a maximum radius of 12 steps upstream of the TFs. In order to identify the potential positive feedback loops in the system we applied additional filtering to the obtained master-regulators. We required that the Composite Score of the genes encoding master-regulators in the system were above the critical value (6.28) computed by the CMA algorithm. This requirement enables us to find those master-regulators that regulate elevated expression of their own genes (through multiple TFBSs found in the promoters of these genes), thus leading to the feedback mechanism of maintaining its own elevated expression. In Table 6 below we give the final list of obtained master-regulators potentially involved in maintaining elevated resistance to Nutlin-3. Figure 6 shows the diagram of the network constructed by the algorithm of master-regulator search. The network connects the identified top master-regulators (red) with the transcription factors found in the promoter analysis (blue).

Table 6. List of master-regulator molecules (including complexes and modified forms) that have passed all criteria. The top molecule with the highest Master-regulator Score (0.656) is the complex of mTOR with raptor protein.

Master molecule name	Gene Symbol	logFC	Composite Score	Reached from set	Master-regulator Score	FDR	Z-Score
mTOR:raptor	MTOR	0.773	8.562	11	**0.656**	0.022	1.499
Dvl-3	DVL3	1.115	7.027	11	0.440	0.037	1.617
foxm1{sumo}	FOXM1	1.281	10.028	10	0.378	0.002	3.809
mTOR{pS}:raptor:p70S6K 1	MTOR	0.773	8.562	10	0.292	0.012	2.250
mTOR{pS}:4E-BP1:EIF-4E:raptor	MTOR	0.773	8.562	10	0.292	0.015	2.195
caveolin-1:(eNOS{myrG2}{palC15}{palC26})2:Hsp90	CAV1	1.532	6.704	8	0.239	0.040	1.866
mTOR:raptor:mLST8:SIN 1:Protor-1	MTOR	0.773	8.562	9	0.207	0.034	1.684

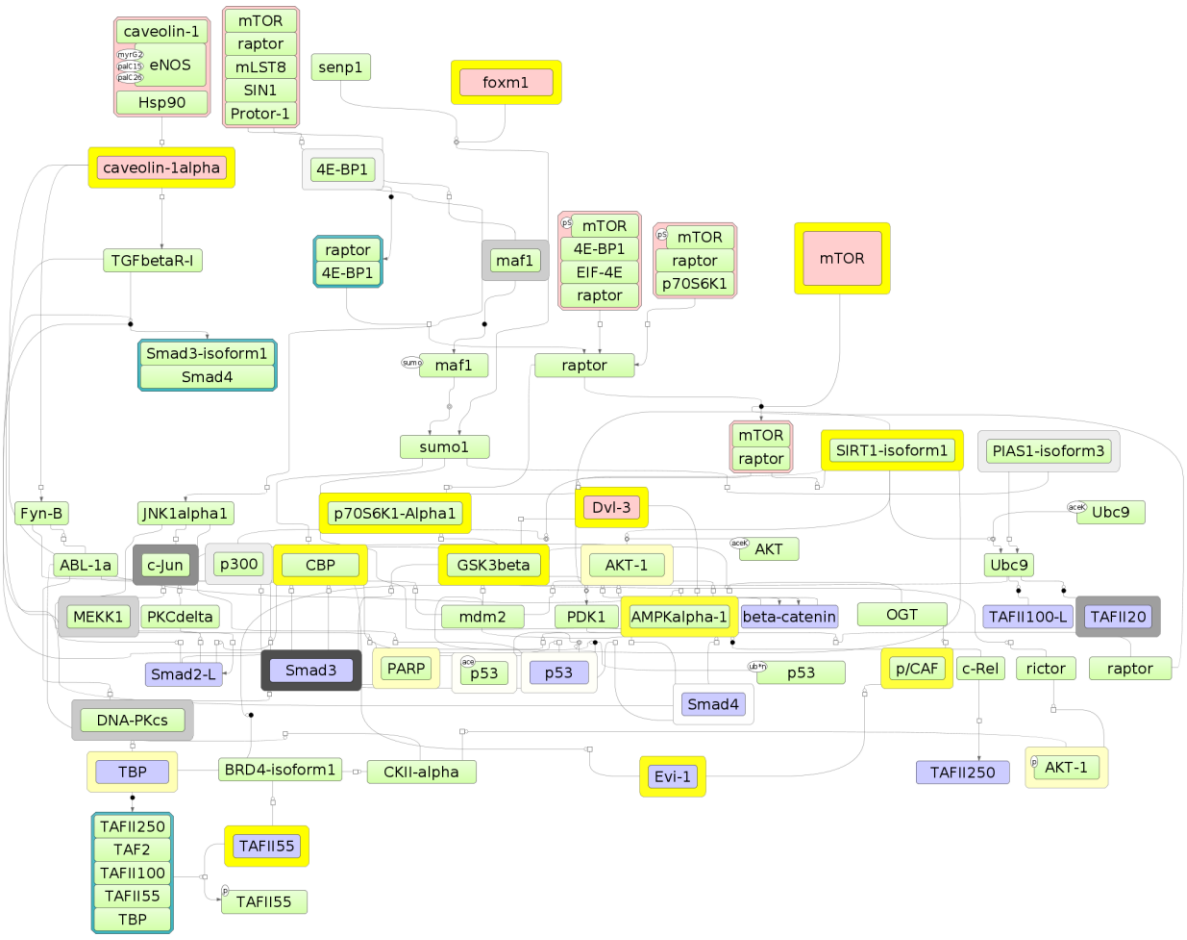

Figure 6. Visualization of that part of signal transduction network which connects identified master-regulators (red) with the transcription factors (blue) found in the promoter analysis of genes upregulated in Nutlin-3 moderately resistant cells. Expression of genes encoding the corresponding proteins on the diagram is shown by the color of the layer around the molecule. Yellow color corresponds to up-regulation, the gray color corresponds to down-regulation. The intensity of the color reflects the fold change.

Comparison of the master-molecule network with canonical pathways.

In order to understand which components of canonical pathways are involved in maintaining the resistance of studied cells to Nutlin-3 we mapped the whole identified network connecting potential master-regulators with transcription factors onto the comprehensive collection of pathways in the TRANSPATH® database.

Table 7. Mapping of the master-regulator network onto the TRANSPATH® collection of canonical pathways. Hit names are the gene names of the master-regulator network mapped to the respective pathway.

TRANSPATH ID	Title	Number of hits	Group size	P-value	Adjusted P-value	Hit names
CH000000658	PDK1 ---> 4E-BP1, p70S6K1	5	9	2.74E-09	7.19E-07	AKT1,MTOR,PDPK1,RPS6KB1,RPTOR
CH000001016	AR pathway	7	46	3.45E-08	4.53E-06	AKT1,EP300,MDM2,PDPK1,SIRT1,SMAD3,UBE2I
CH000004640	transcriptional regulation of ECM components	4	8	2.32E-07	1.78E-05	EP300,SMAD2,SMAD3,SMAD4
CH000000743	tuberin pathway	6	37	2.71E-07	1.78E-05	AKT1,GSK3B,MTOR,PDPK1,RPS6KB1,RPTOR
CH000000710	p53 pathway	8	130	3.86E-06	1.13E-04	AKT1,EP300,KAT2B,MDM2,MTOR,PDPK1,TP53,UBE2I
CH000000968	PI3K ---Mdm2--/ p53	5	33	4.56E-06	1.14E-04	AKT1,MDM2,MTOR,PDPK1,TP53
CH000000651	insulin ---> AKT-1 pathway	5	38	9.39E-06	2.06E-04	AKT1,GSK3B,MTOR,PDPK1,RPS6KB1
CH000003574	SIRT1 ---/ AR	3	7	1.66E-05	3.35E-04	EP300,SIRT1,SMAD3
CH000004522	TLR2 ---Rac1-->AKT	4	21	1.85E-05	3.47E-04	AKT1,GSK3B,MTOR,PDPK1
CH000004188	NRG ---> Akt-1	4	22	2.25E-05	3.94E-04	AKT1,GSK3B,MTOR,PDPK1
CH000004600	mammalian Hippo	5	51	4.11E-05	6.75E-04	AKT1,PDPK1,SMAD2,SMAD3,SMAD4

	network					
CH000000750	insulin pathway	6	89	5.23E-05	8.09E-04	AKT1,GSK3B,MTOR,PDPK1 ,RPS6KB1,RPTOR
CH000004646	Smad2/3 --- TAZ---> cytoplasmic retention	3	11	7.65E-05	0.00100	SMAD2,SMAD3,SMAD4
CH000000574	mTOR ----> S6, eIF-4E	3	12	1.01E-04	0.00121	MTOR,RPS6KB1,RPTOR
CH000004539	AKT-1 --- PRAS40---> mTOR	3	12	1.01E-04	0.00121	AKT1,MTOR,RPTOR

As one can see from the Table 7 the pathways that involve the master regulator mTOR are the well-known pro-survival pathways of PDK1, ACT-1 and PI3K. We found the PI3K sub-pathway especially interesting, which involves mTOR for inhibiting p53 – the main target of Nutlin-3 (see Fig. 7).

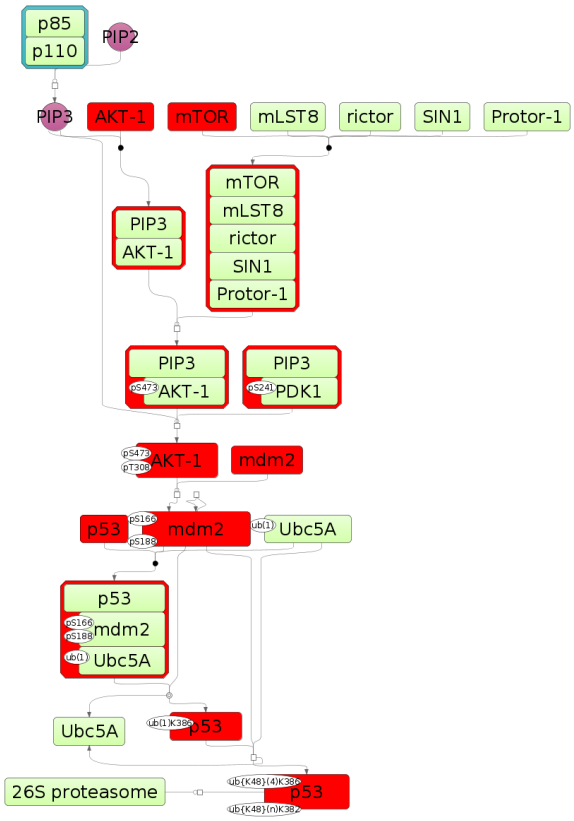

Figure 6. Visualization of the canonical PI3K sub-pathway that leads to the inhibition of p53 (PI3K ---Mdm2---/ p53). PI3K is represented on the diagram as the complex "p85:p110". Red molecules represent the components of the master-regulator network (see Fig.5) mapped to the PI3K sub-pathway.

Experimental validation of mTOR as master-regulator by chemical inhibitors.

After choosing the master-regulator mTOR as the target molecule we applied specific chemical inhibitors in order to test their effect on the survival of these cell lines. We choose the following chemical inhibitors (Table 8).

Table 8. Three chemical inhibitors used in this work for the validation of molecular targets.

#	**cat №**	**Chemical name** (Cayman)	**Target(s)**
C1	Cay70920-5	LY294002	PI3K
C2	Cay10005229-5	2,3-DCPE (hydrochloride)	Bcl-XL
C3	Cay10565-25	NVP-BEZ235	PI3K and mTOR

To validate the effect of the inhibition of mTOR in the context of PI3K pathways we chose the dual inhibitor C3 that is known to effectively inhibit both PI3K and mTOR action. To test the mTOR-specificity of the effect we used the general inhibitor of PI3K (C1). And as a negative test we used the inhibitor of Bcl-XL (C2) that is known to be involved in cancer but was not predicted here as a master-regulator.

We used the following treatment conditions for these three compounds.

- C1 in concentrations: 50μM, 25μM, 12,5μM, 6,25μM, 3,1μM and 0μM (Control);
- C2 in concentrations: 80μM, 40μM, 20μM, 10μM, 5μM and 0μM (Control);
- C3 in concentrations: 50μM, 25μM, 12,5μM, 6,25μM, 3,1μM and 0μM (Control).

The concentration of the compounds was chosen according to their maximal solubility in the medium. DMSO was used to prepare the needed concentration of the compounds and the concentration of DMSO was the same in all tested solutions (including Control) and it was below 0.5%. Each measurement was done in three replicas. After 48 hours from adding of the compound we measured the percentage of the survived cells using the test with rezuverin.

We found that compounds C1 and C2 do not demonstrate any cytotoxic activity on all tested cell lines (data not shown). Under any compound concentrations the number of the survived cells were the same as in the Control and did not differ significantly from 100%.

In contrast, compound C3 demonstrated a moderate cytotoxic activity with the following IC50 values in tested cell lines: A427 (IC50 = 11.8±3.4), H292 (IC50 = 6.05±2.1), DV-90 (IC50 = 19.9±4.6), H1944 (IC50 = 34.9±3.6) и H2228 (IC50 = 27.1±6.6). For the three cell lines of interest the dose-effect curves are shown in Figure 8 below.

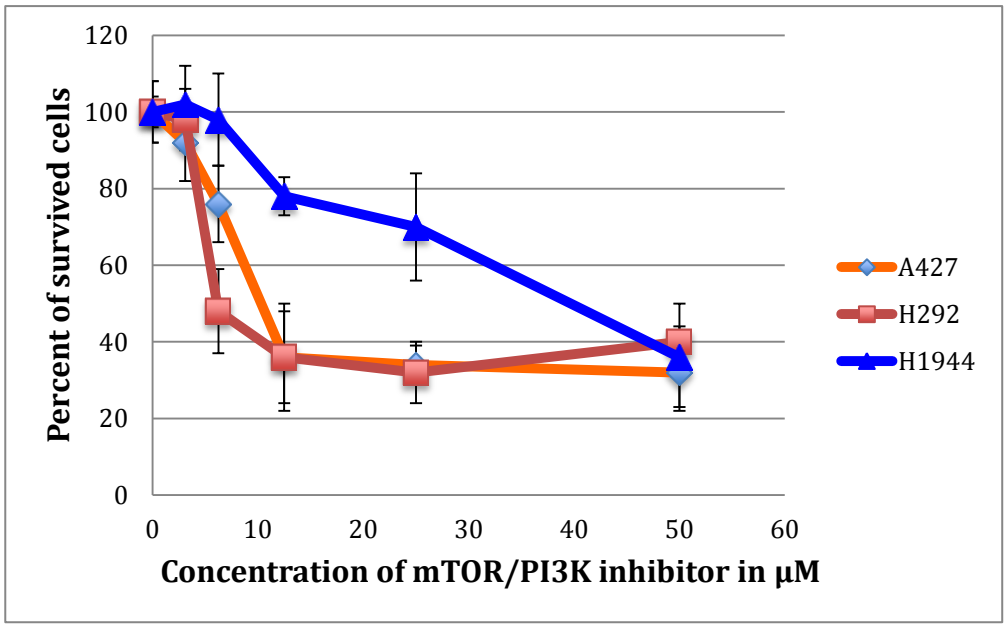

Figure 8. Graphics of the viability test of three NSCLC cell lines to compound NVP-BEZ235 – the dual inhibitor of PI3K and mTOR. The colored lines represent effects of increasing inhibitor concentrations on the survival of three cell lines: sensitive to Nutlin-3 (H1944) and moderately resistant to Nutlin-3 cell lines (A427 and H292).

We found that Nutlin-3 moderately resistant cell lines (A427 and H292) exhibit the highest sensitivity to the dual chemical inhibitor of mTOR-PI3K whereas the Nutlin-3 sensitive cell line (H1944) appeared to be relatively insensitive to this inhibitor. These results confirmed our prediction of the master regulator mTOR in the PI3K signaling pathway, which, most probably, is responsible for the elevated resistance of particular NSCLC cell lines to treatment by the p53-reactivating compound Nutlin-3.

3. Discussion

Resistance to chemotherapy and targeted therapy of cancer cells is one of the biggest problems of cancer treatment. In this work, we studied the molecular mechanisms of resistance of cancer cells to the p53-reactivating compound Nutlin-3 using genome-wide transcriptomics profiling followed by causative computational analysis. In order to analyze

the effect of Nutlin-3 on non-small cell lung cancer (NSCLC) and understand the mechanisms of resistance, we performed an extensive study of the biological activity of the compound Nutlin-3 on a number of NSCLC cell lines carrying a wild-type TP53 gene. We identified that the cell line H1944 is the most sensitive to the treatment by Nutlin-3. Even under the lowest concentration of the compound the cells of this cell line were rapidly dying. Other cell lines, such as A427 and H292 showed relatively high resistance to the treatment by Nutlin-3. The death of cells was triggered only under relatively high concentration of the compound. We compared the gene expression profiles of these cell lines before and after treatment with Nutlin-3 and identified several hundreds of genes whose expression was significantly different between sensitive and the moderately resistant cell lines. The GO and pathway mapping of these genes gave us the first clue about the main processes involved in regulation of resistance to Nutlin-3. Among the most significant pathways were the p53 pathway, the E2F network, the Aurora-B and A cell cycle regulation pathway and other pathways involved in the regulation of different processes related to the regulation of the cell cycle. From the perspective of searching for drug targets we focus our attention first of all on the genes that have higher expression values in the resistant cell lines compared to the sensitive (up-regulated genes). Analysis of promoters of these genes helped us to identify several transcription factors with enriched binding sites found in co-localized clusters. We believe that such clusters reflect position and composition of a very specific enhanceosome that is formed at the promoters of up-regulated genes and controls their elevated expression in the resistant cell line. In the next step of our study we analyzed signal transduction networks upstream of the revealed transcription factors in order to understand the potential molecular mechanism of activation of the revealed set of transcription factors. The goal of such an analysis was to reveal few master-regulators in this network that might exert their control on the transcription factors found in the first step.

Finally, after performing the search for potential master regulators, we checked which of them were actually up-regulated by themselves. So we require that genes, which are expressing proteins that were found by the algorithm as potential master regulators, should have a significantly higher expression in the resistant cells compared to the sensitive cells. This reflects the presence of positive feedback loops in the system. We hypothesized that the observed increase of resistance might be supported by the presence of positive feedback loops. We can observe such loops in the network when the genes, expressing master-regulator proteins, are working under control of transcription factors, which receive activating signals through the signaling cascade, starting from the proteins, which are

expressed by these genes (master regulators). Therefore, the up-regulation of those genes, encoding master regulators in this analysis, indicates the presence of such feedback loops. We think that such positive feedback loops can contribute to the stabilization of the resistance to Nutlin-3 (and potentially to other anticancer compounds with similar mechanisms of action, such as p53 reactivator molecule RITA, which was studied in our previous work [21]), since they maintain the constant activation of a certain set of critically important genes through the auto-activation loop.

As a result, we revealed the following master regulator genes: MTOR, DVL3, FOXM1, CAV1. We noticed that many of the suggested master regulators are very important proteins that are known to be involved in regulating such processes as cell cycle and apoptosis.

Finally, we found that Nutlin-3 moderately resistant cell lines (A427 and H292) exhibit the highest sensitivity to the dual chemical inhibitor of mTOR-PI3K, whereas the Nutlin-3 sensitive cell line (H1944) appeared to be relatively insensitive to this inhibitor. These results confirmed our prediction of the master regulators in the mTOR-PI3K signaling pathway responsible for the elevated resistance of particular NSCLC cell lines to treatment by the p53-reactivating compound Nutlin-3. As we predicted, the Nutlin-3 resistant cell lines appeared to be highly sensitive to the inhibitors of mTOR-PI3K pathway. These findings open a promising possibility for a combinatory therapy combining Nutlin-3 with mTOR-PI3K inhibitors. Such drug combinations will have a potential for tackling the observed heterogeneity between different cancer cell linages towards p53-reactivators such as Nutlin-3.

Funding

This work was supported by the grant of the Skolkovo Foundation to BIOSOFT.RU, LLC, the project title is «Information platform for biomedical research», grant number - Г73/15 (29th of October, 2015).

Acknowledgments

We would like to thank the Skolkovo foundation for financial support of this work (grant No Г73/15 from October 29th 2015, given to BIOSOFT.RU, LLC for the project "Information platform for biomedical research"). We are also very grateful to Niko Voss,

Holger Michael, Jeannette Koschmann, Dr. Olga Kel-Margoulis and Prof. Edgar Wingender for their irreplaceable contributions to the performed research.

Conflicts of Interest

AK is an employee of geneXplain GmbH, which maintains and distributes the BioUML/geneXplain platform used in this study. Alexander Kel is also an employee in BIOSOFT.RU, LLC, which developed the BioUML platform and the pipeline system “From genome to target”, which was used to perform this study.

References

1. Ray-Coquard I, Blay JY, Italiano A, Le Cesne A, Penel N, Zhi J, Heil F, Rueger R, Graves B, Ding M, Geho D, Middleton SA, Vassilev LT, Nichols GL, Bui BN Effect of the MDM2 antagonist RG7112 on the P53 pathway in patients with MDM2-amplified, well-differentiated or dedifferentiated liposarcoma: an exploratory proof-of-mechanism study. Lancet Oncol. 2012 Nov; 13(11):1133-40.
2. Burgess A, Chia KM, Haupt S, Thomas D, Haupt Y, Lim E. Clinical Overview of MDM2/X-Targeted Therapies. Front Oncol. 2016 Jan 27;6:7. doi: 10.3389/fonc.2016.00007. eCollection 2016. Review.
3. Haupt, Y., Maya, R., Kazaz, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. // Nature – 1997. – Vol. 387. P. 296–29
4. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. // Science. – 2004. – Vol. 303(5659). P.844–848.
5. Grasberger BL, Lu T, Schubert C, Parks DJ, Carver TE, Koblish HK, Cummings MD, LaFrance LV, Milkiewicz KL, Calvo RR, et al. Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. // J Med Chem. – 2005. – Vol. 48. P. 909–912.
6. Shangary S, Qin D, McEachern D, Liu M, Miller RS, Qiu S, Nikolovska-Coleska Z, Ding K, Wang G, Chen J, et al. Temporal activation of p53 by a specific MDM2 inhibitor is selectively toxic to tumors and leads to complete tumor growth inhibition. // Proc Natl Acad Sci USA. – 2008. – Vol. 105(10). P. 3933–3938
7. Ding K, Lu Y, Nikolovska-Coleska Z, Wang G, Qiu S, Shangary S, Gao W, Qin D, Stuckey J, Krajewski K, et al. Structure-based design of spiro-oxindoles as potent, specific small-molecule inhibitors of the MDM2-p53 interaction. // J Med Chem/ – 2006. – Vol. 49. P. 3432–3435

8. Kojima, K. et al. MDM2 antagonists induce p53-dependent apoptosis in AML: implications for leukemia therapy. // Blood – 2005. Vol. 106 P. 3150–3159

9. Secchiero, P., di Iasio, M. G., Gonelli, A. & Zauli, G. The MDM2 inhibitor Nutlins as an innovative therapeutic tool for the treatment of haematological malignancies. // Curr. Pharm. Des. – 2008. – Vol. 14. P. 2100–2110

10. Kel, A. E., Stegmaier, P., Valeev, T., Koschmann, J., Poroikov, V., Kel-Margoulis, O. V. and Wingender, E. (2016) Multi-omics "upstream analysis" of regulatory genomic regions helps identifying targets against methotrexate resistance of colon cancer. EuPA Open Proteomics 13, 1-13. doi: 10.1016/j.euprot.2016.09.002.

11. Anoopkumar-Dukie, J B Carey, T Conere et al. Resazurin assay of radiation response in cultured cells. // The British Journal of Radiology – (2005), – Vol. 78. P. 945-947

12. Finney D.J. Probit Analysis. // Cambridge Un. Press (1978) Cambridge.

13. Wingender E (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Brief. Bioinform. 9:326-332.

14. Kel AE, Gössling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res. 31:3576–3579.

15. Krull M, Pistor S, Voss N, Kel A, Reuter I, Kronenberg D, Michael H, Schwarzer K, Potapov A, Choi C, Kel-Margoulis O, Wingender E. TRANSPATH: an information resource for storing and visualizing signaling pathways and their pathological aberrations. // Nucleic Acids Res. - (2006) – Vol. 34(Database issue):D546-51.

16. Waleev T, Shtokalo D, Konovalova T, Voss N, Cheremushkin E, Stegmaier P, Kel-Margoulis O, Wingender E, Kel A. (2006) Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. Nucleic Acids Res.,34c(Web Server issue):W541-5.

17. Kulakovskiy IV, Vorontsov IE, Yevshin IS, Soboleva AV, Kasianov AS, Ashoor H, Ba-Alawi W, Bajic VB, Medvedeva YA, Kolpakov FA, Makeev VJ. HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models. Nucleic Acids Res. 2016 Jan 4;44(D1):D116-25. doi: 10.1093/nar/gkv1249.

18. Yevshin I, Sharipov R, Valeev T, Kel A, Kolpakov F. GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments. Nucleic Acids Res. 2017 Jan 4;45(D1):D61-D67. doi: 10.1093/nar/gkw951.

19. Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E. Beyond microarrays: find key transcription factors controlling signal transduction pathways. // BMC Bioinformatics. – (2006) – Vol. 7 Suppl 2:S13.

20. Chung IM, Ketharnathan S, Kim SH, Thiruvengadam M, Rani MK, Rajakumar G. Making Sense of the Tangle: Insights into Chromatin Folding and Gene Regulation. Genes (Basel). 2016 Sep 23;7(10). pii: E71. doi: 10.3390/genes7100071. Review.

21. Nikulenkov F, Spinnler C, Li H, Tonelli C, Shi Y, Turunen M, Kivioja T, Ignatiev I, Kel A, Taipale J, Selivanova G. Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. Cell Death Differ. 2012 Dec;19(12):1992-2002.