

REPORT - PROJECT 01 - GROUP 5

Carolin Bayan, Savannah Cattarius, Laura Diekmann, Rositsa Todorovska

07/19/2021

INTRODUCTION

Drug repurposing describes the strategy of reusing drugs that have already been approved for other medical applications. These drugs can obtain a new purpose by using them in a different therapeutic indication. This approach offers several advantages compared to the development and research of novel drugs. Drug repurposing allows researchers to use already approved pharmaceuticals, which helps them face less obstacles regarding potential side effects and the general approval of these drugs. This is due to the fact that the safety of the drugs has already been tested and validated in prior researches. Therefore, costs and legal requirements can be reduced which results in a quicker development of new treatment options [Pushpakom2019].

According to the WHO ¹, cancer is the second leading cause of death in the United States and a major global health problem. In addition, the progress of diagnosing and treating this disease was set back due to the coronavirus disease 2019 (COVID-19) pandemic which resulted in reduced health care capacities [https://doi.org/10.3322/caac.21654]. Hence, finding new treatment options by repurposing drugs in order to save resources and time seems to be essential for our society.

For our project, we used data obtained from a large screening of 1396 oncological and non-oncological drugs that were tested on 481 different cancer cell lines [Corsello2020]. While also working with the whole data consisting of several different cancer cell lines, in our analysis we particularly focused on the included 30 ovarian cancer cell lines. The American Cancer Society states that ovarian cancer causes more deaths than any other form of cancer being related to the female reproductive system ².

In our project, we aimed to answer two main questions:

1. How do different drugs influence ovary cancer cell lines and which are particularly noticeable?
How do these drugs affect the other cell lines?
2. Is the sensitivity of different drugs on ovary cancer cell lines connected to specific cancer-related genes or gene knock-outs?

We subdivided our project into smaller steps in order to find answers to these questions.

First, we analyzed the drugs and their influence on the proliferation of cancer cell lines. We were able to identify the most promising drugs and categorized them based on their mechanism of action (MOA). In addition, we analyzed to what extent the used drug dosage had an effect on the proliferation values of the cancer cell lines and found out that in most cases the higher the dosage the greater the effect on proliferation. We also investigated the drug response of certain ovarian cancer cell line clusters based on certain gene expression patterns and knockdown scores which showed less correlations than we initially anticipated. Finally, we used the data to perform a linear regression model in order to predict drug efficiency from certain gene mutations or expressions.

¹https://www.who.int/health-topics/cancer#tab=tab_1

²<https://www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html>

MOST EFFECTIVE DRUGS AMONG OVARIAN CELL LINES

GENERAL OVERVIEW

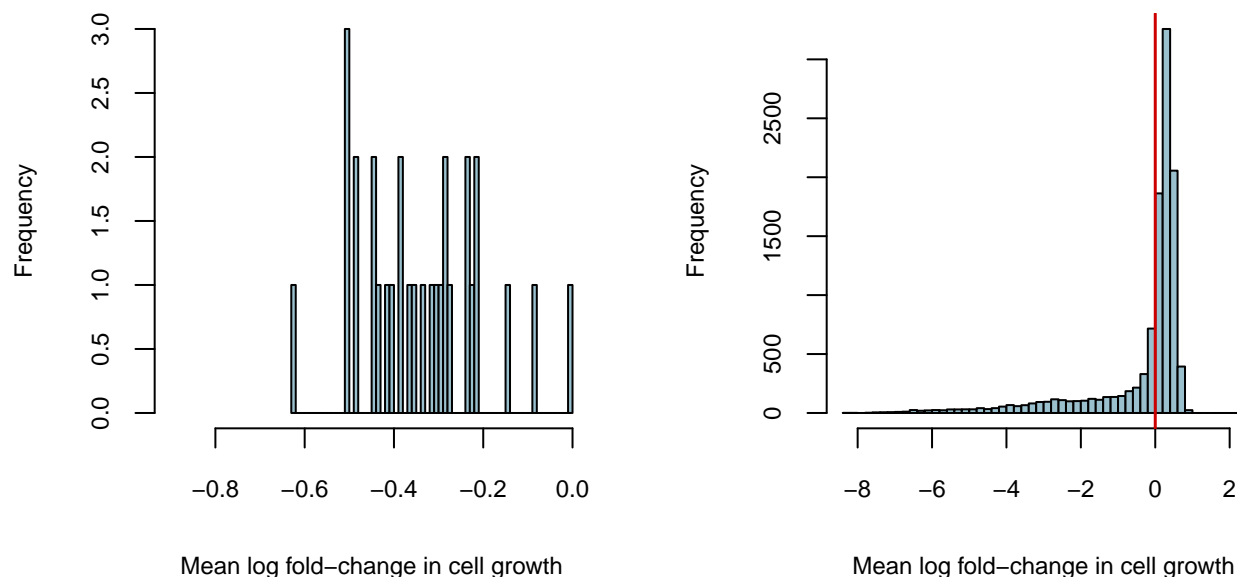


Figure 1: **1.a.** Ovarian Cancer Cell Lines Mean Sensitivity towards Drugs **1.b.** Mean Drug Efficiency regarding Proliferation in Ovarian Cell Lines

Using the provided prism data frame, we extracted the ovarian DepMap_IDs with their proliferation values after drug treatment. In order to get a general overview of our data, we computed two histograms showing the distribution of the mean sensitivity across all ovarian cell lines and all drugs (see Fig. 1). Fig. 1.a. shows that all ovarian cancer cell lines had a mean sensitivity towards the drugs of less than or equal to zero, which showed us that in general, the applied drugs seemed to have a promising effect on their cell growth. The right plot (Fig. 1.b.) shows all values for the mean drug efficiency in ovarian cell lines. Values below zero indicate negative proliferation. In total, 8652 drugs lie below zero. This means that these drugs generally - with regard to their mean values - caused a reduction in cell growth when they were administered to the ovarian cell lines.

SETTING A THRESHOLD

In our next step, we further analyzed the prism data set. Our aim was to determine the strongest effect each drug had on the ovarian cancer cell lines in terms of proliferation values - independent of the dosage that was used. For each drug we chose its minimum value which reflects the strongest negative effect on proliferation of the ovarian cell lines. In order to be able to work with a manageable amount of drugs, we decided to set a threshold at the 10 % quantile (see Fig. 2.a.). By doing so, we could focus on the 140 most effective drugs which caused the lowest negative proliferation values in the ovarian cancer cell lines.

MECHANISM OF ACTION

While further analyzing the drugs that lie below the 10 % quantile, we focused on categorizing them after their mechanism of action which we found in the data set prism.treat. According to the original study, clustering by the MOAs seems to be a good choice in order to determine whether the drugs are normally used for treating cancer or not [Corsello2020].

Our results are shown in Fig. 2.b. These are the 45 most common MOAs regarding the drugs that had the most promising effects on the ovarian cancer cell lines.

We further investigated these MOAs and came to the following results.

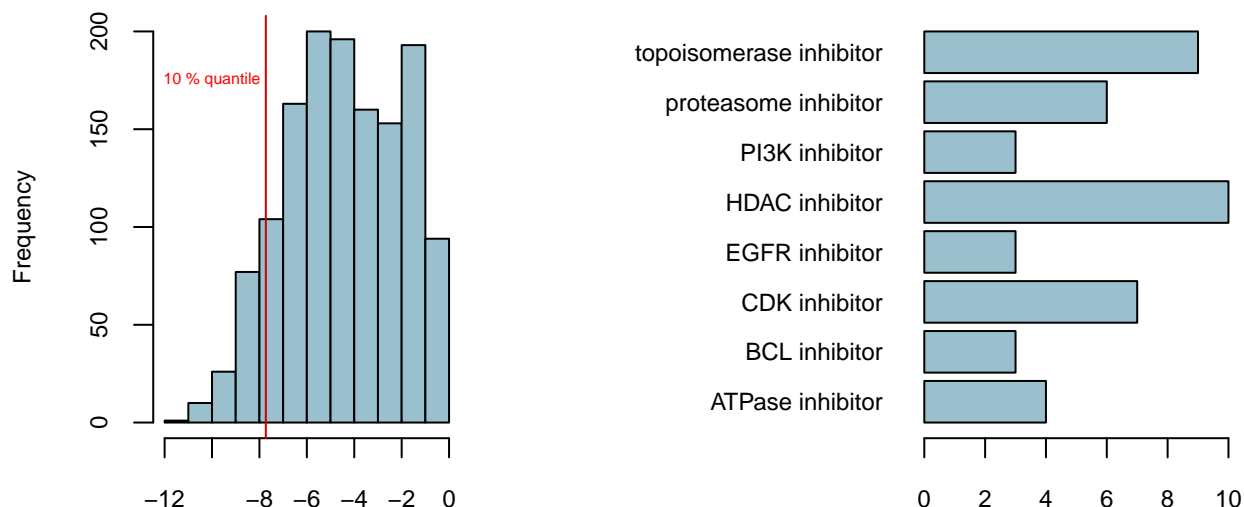


Figure 2: **2.a.** Most Effective Drug-Proliferation Values in Ovarian Cell Lines **2.b.** Most Frequent MOAs Among the Most Effective Drugs

10 out of the 140 drugs below our threshold are Histone deacetylase (HDAC) inhibitors. Their targets include among other things HDAC3, HDAC4, HDAC8. These are epigenetic targets for noncancer drug repurposing candidates [Moreira-Silva2020] and were also found to be potential anti-cancer agents for inhibiting cancer cell migration and invasion in ovarian cancer [AHN2012189].

With 9 occurrences, the topoisomerase inhibitors are the second most frequent MOAs we found. However, this is not surprising since these drugs are commonly used in cancer therapy, thus they showed strong effects on the cancer cell lines [Parchment1998]. Interestingly, one of the drugs, mitoxantrone, is also used in the disease area neurology/psychiatry and multiple sclerosis³.

In addition, cyclin-dependent kinase (CDK) inhibitors appeared seven times in our findings regarding the most effective and common MOAs for treating ovarian cancer. Especially the targets CDK4 and CDK6 showed promising results in a different study [Iyengar2018] as a new therapeutic approach to treat ovarian cancer.

We found 6 proteasome inhibitors to be most effective against ovarian cancer cell lines. They are an important class of drugs against multiple myeloma and cell lymphoma. However, they are currently in clinical trials for other types of cancer [doi:10.1146/annurev-pharmtox-010919-023603] which is why it makes sense that we also found them to have promising effects in our analysis.

In a different study, researchers found that epidermal growth factor receptor (EGFR) is not a suitable target for ovarian cancer therapy [Mehner2017] and seems to be more effective for lung cancer and pancreas cancer therapy. However, we found three EGFR inhibitors with effective impacts on ovarian cancer cell lines.

The remaining three MOAs: Phosphoinositide 3-kinase (PI3K)- [Vanhaesebroeck2021], B-cell lymphoma (BCL)- [Montero2018] and ATPase-inhibitors [doi:10.4155/ppa.12.3], are all cancer related MOAs. Therefore, all of them are commonly used in cancer therapy which is why it is not surprising that they appear in our findings.

DIMENSION REDUCTION

Our next step included performing k-means and finding the optimal number of clusters which we visualized by computing a principal component analysis (PCA).

³National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 4212, Mitoxantrone. Retrieved July 11, 2021 from <https://pubchem.ncbi.nlm.nih.gov/compound/Mitoxantrone>.

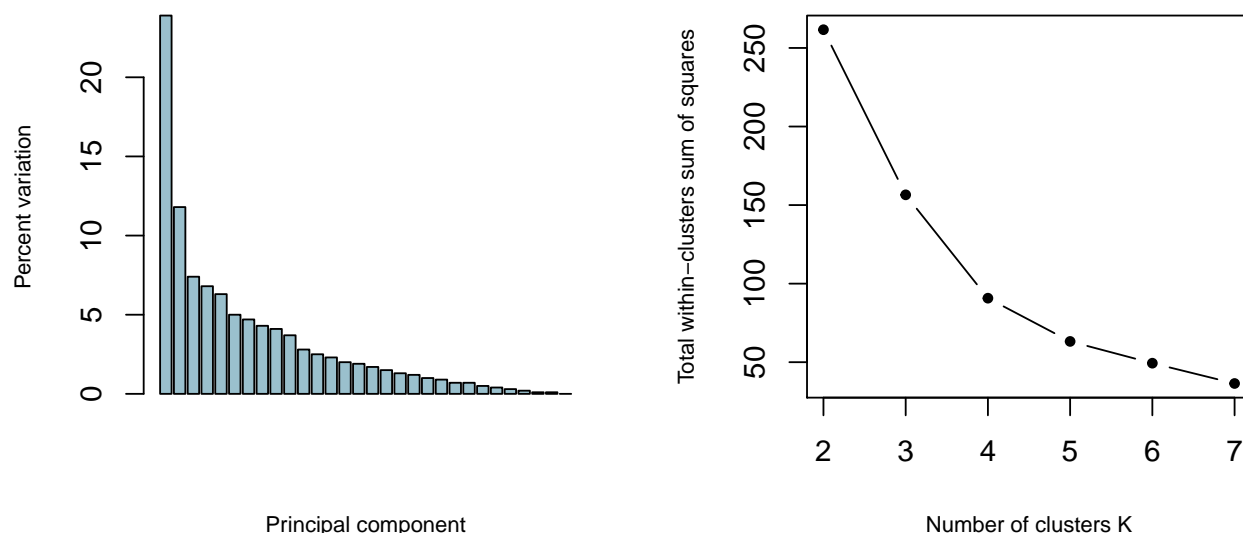


Figure 3: **3.a.** Variance Explained by Mechanism of Action PCs **3.b.** Elbow Method for Finding Optimal Number of Clusters

Fig. 3.a. shows that the majority of the variance is explained by using the first two principal components. In order to find the most optimal number of clusters for our PCA of PC1 and PC2, we used the elbow method. As a result, the optimal number of clusters seems to be 4 (see Fig. 3.b.).

OUTLIER

Fig. 4.a. shows the clustering for the ovarian cell lines according to their reactions towards the 45 selected drugs (see Fig. 2.b.). They do not cluster in completely distinct clusters but one cell line (ACH-000524) forms its own cluster. We found that interesting, which is why we looked further into that one. According to the data set prism.cl, the outlier cell line belongs to the most prominent lineage subtype: ovary adenocarcinoma. This subtype makes up 28 of the 30 lineage subtypes in our data with the remaining two being brenner tumor and ovary carcinoma. This is why we figured that the lineage subtype could not have been the reason for the cell line forming its own cluster.

We performed several PCAs with the other provided data sets (prism.achilles, prism.ecp and prism.cnv) and inspected the behavior of the outlier cell line. When we analyzed the gene transcripts per million (TPM) values and the gene knockdown scores, the cell line ACH-000524 clustered in the biggest clusters together with the other cell lines (not shown). However, the cell line did show a different behavior and clustered separately from the majority according to its gene copy number (CN) values (see Fig. 4.b.). This led us to the hypothesis that its differing behavior (see Fig. 4.a.) could be a result of distinct CN values.

GENERALLY EFFECTIVE DRUGS

In Fig. 5, the contribution to PC1 for ovary cell line clusters (see Fig. 4.a.) is shown. Drugs with values close to zero show a similar effect on the proliferation among all cell lines - thus, these are generally effective when administered to ovary cancer cell lines. This is the case for the following drugs (see highlighted red in Fig. 5).

The ATPase inhibitor ouabain is normally used to treat hypertension and cardiac arrhythmia - however, a study found that it can have promising effects as a cancer treatment for ER+ breast cancer [Chen2006]. In another study, researchers found that one in three ovarian cancers is ER+ [Harding1990], which leads to the conclusion that this is why we found that ouabain generally showed positive effects among the ovarian cancer cell lines.

The drugs JNJ-26481585 [Arts6841] and romidepsin [doi:10.1517/13543784.2011.594437] are HDAC

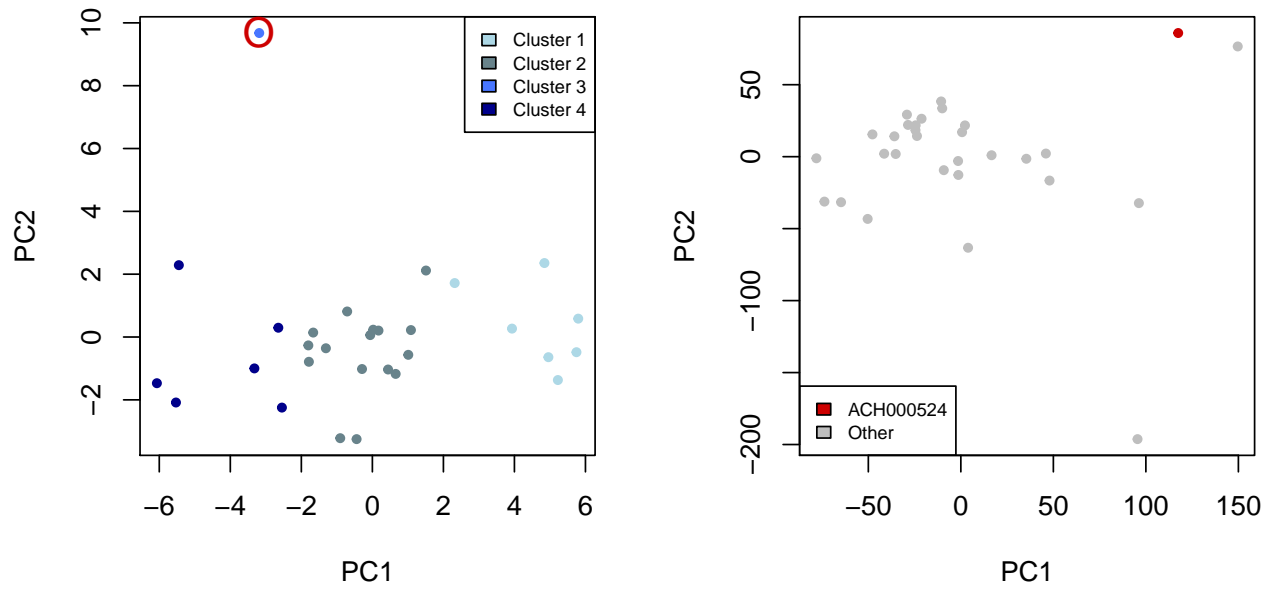


Figure 4: **4.a.** PC1 vs. PC2 for Ovary Cell Line Clusters **4.b.** PC1 vs. PC2 for CN Values in Ovary Cell Lines

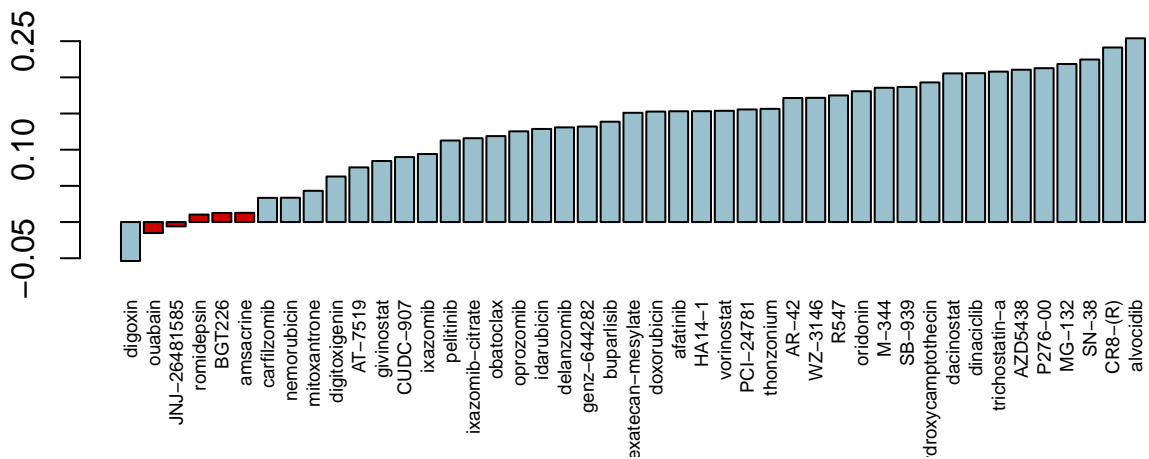


Figure 5: Contribution to PC1

inhibitors that, as mentioned before, in general show promising effects on ovarian cancer cell lines. Therefore, it makes sense that we found them among the five generally most effective drugs.

BGT226 is a PI3K inhibitor that is actually being considered as a potential therapeutic treatment option in different solid tumors, including ovarian cancer [Simioni2015].

Amsacrine, a topoisomerase inhibitor, is a chemotherapy drug that is commonly used for acute leukemia as well as Hodgkin's and non-Hodgkin's lymphomas [Ketron2012]. Hence, a general effect on the proliferation of other cancer cell lines such as ovarian cancer seems to be logical.

DRUG EFFECT ON THE CELL LINE TYPES

MECHANISM OF ACTION

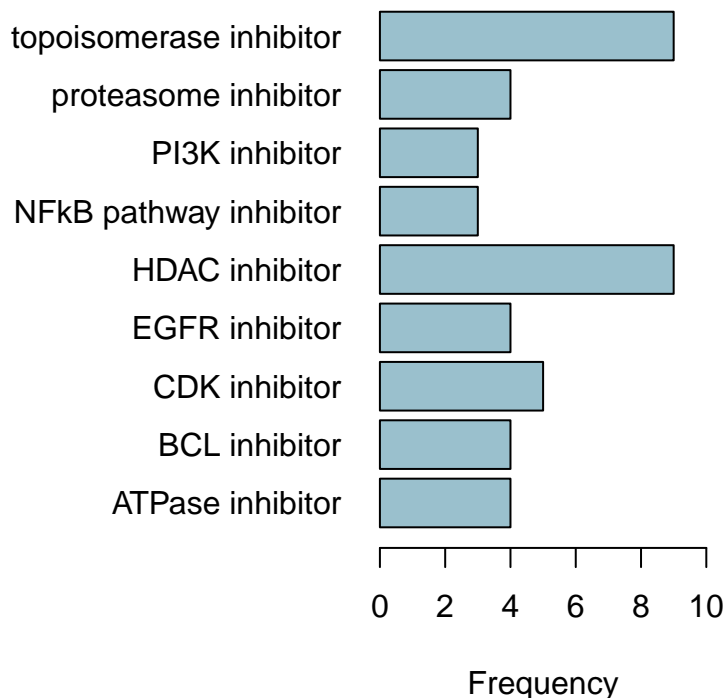


Figure 6: Most frequent MOA's amongst the most effective drugs

After finding the drugs with the most effective response in regards to the ovary cancer cell lines, we now continued our analysis by testing if the efficiency of these drugs is dependent on the cancer cell line type or not. In order to test this we repeated many of the same steps we completed for the ovary cell lines but on the non-ovary cell lines. We extracted the most negative proliferation value for each drug in each non-ovary cell line, set the threshold to the 10% quantile, and then extracted the most common MOAs.

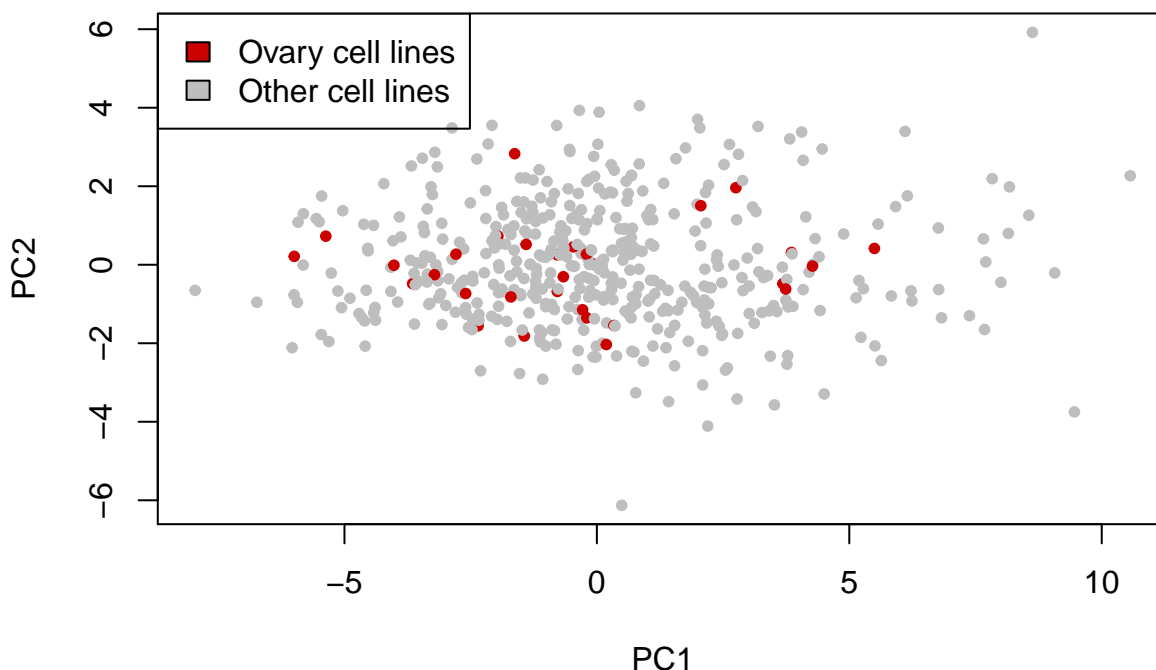
As can be seen in Fig. X, all but one of the most frequently occurring MOAs are the same as the most frequently occurring MOAs in ovary cancer cell lines, however; the frequency with which they occur in the other cell lines differs slightly from the frequency with which they occur in the ovary cell lines. The NFkB pathway inhibitor is the one MOA that appears among the 45 most frequent MOAs in the non-ovary cell lines three times but not at all in the 45 most frequent MOAs in the ovary cell lines. Although NFkB pathway inhibitors do occur in the ovary cell lines twice, this is not often enough to have made the top 45 which explains why they are not shown in Fig 2.b. NFkB transcription factors are often used to regulate inflammation and immune responses, cell growth, apoptosis, and the expression of certain viral genes [Gilmore2006]. Due to the importance of the NFkB transcription factors in regulating these processes, NFkB pathway inhibitors are often used in treating chronic inflammation and cancer [Gilmore2006].

FISCHER'S EXACT TESTS

After finding the most common MOAs among the ovary and all other cell lines respectively, the next step in our analysis was to find out if any of these MOAs showed a statistically significant occurrence in ovary cancer specifically. In order to test this, we decided to perform a fisher-exact-test on all of our selected most common MOAs for ovary cancer and all other cell lines (see Fig. X).

In order for the null hypothesis - that there is no association between the occurrence of specific MOAs and the cell line type (either ovary or non-ovary) - to be rejected, the p-value must be smaller than the confidence level of 0.05 [Biau2010]. This was not the case with any of the fisher tests that we conducted. For the fisher tests conducted on the topoisomerase inhibitor, the PI3K inhibitor, the NFkB pathway inhibitor, the HDAC inhibitor, the EGFR inhibitor, the BCL inhibitor, and the ATPase inhibitor the p-values were equal to one, as were the odds-ratios. This shows without a doubt that the H0-hypothesis cannot be rejected and that there therefore is no association between the occurrence of specific MOAs and the cell line type (either ovary or the others) for these MOAs. The fisher tests conducted on the remaining MOAs all calculate a p-value that is lower than 1, none of which are lower than 0.7 however. The odds ratios for these remaining MOAs are all either larger than one or lower than one, which would suggest a positive/negative association, respectively. Since the corresponding p-values are much larger than the confidence level of 0.5 however, these associations are most likely due to chance and aren't true associations. Conclusively these fisher tests have shown that not a single one of the most common MOAs among our cell lines occur so commonly due to an association to a specific cell line (either ovary or the others).

DIMENSION REDUCTION



As a next step we decided to perform a PCA on all cell lines. We clustered by cell lines in order to test if the ovary cell lines would form their own distinctive cluster regarding the drug response. As can be seen in Fig. X, this was not the case. The red dots representing the ovary cancer cell lines are distributed across the entire plot, rather than concentrated into one distinctive cluster. The fact that the ovary cell lines did not form their own cluster shows that there seems to be no combining trait among them that would trigger a similar drug response.

WELCH'S T-TEST

Additionally to the PCA clustering in Fig. 4.a. and the PCA clustering above, we also conducted a PCA using only non-ovary cell lines (not shown). As a final step to find out if the drug response is related to the cell line type, we decided to perform a Welch's unpaired two sided t-test, which was subjected a Bonferroni correction. In order to do this we reduced the `prism.treat` data frame further. It now included only those cell lines located in clusters two and four from the PCA clusterings for non ovary and ovary cell lines, respectively. With these two data frames we performed a Welch's unpaired two sided t-test, which was subjected a Bonferroni correction. The resulting p-value was higher than the confidence level of 0.05 which means that the null hypothesis - that there is no association between the mean value of the drug response in ovary and non ovary cell lines - could not be rejected. This shows once again that the drug response seems to be independent of the cell line type in regards to the ovary or non-ovary cell lines.

Table 1: Table of the 3 drugs with a significant different response for clusters 1 and 3

	p_value	Bonferroni
trichostatin-a	1.27e-05	0.0176843
semaxanib	1.43e-05	0.0198872
alvocidib	2.45e-05	0.0342022

Table 2: Table of the 2 drugs with a significant different response for clusters 2 and 3

	p_value	Bonferroni
torin-1	1.63e-05	0.0227086
AVN-944	1.74e-05	0.0243205

REFERENCES

PROCESS AND RESULTS

In this step, we used an unpaired two-sided Welch’s t-test, which was subjected to a Bonferroni correction, to show that the different drug response could not be explained by significant differences in specific gene expression rates, gene knockdowns or copy number of variations. We reduced prism.treat to only the cells from the clusters 1 and 3, 1 and 2, as well as 2 and 3, to perform multiple t-tests. We could not perform any tests on cluster 4 due to the only one cell line it consisted of which would provide an unrepresentative result. We also performed t-tests on the pairs of clusters from prism.exp, prism.cnv and prism.achilles. Only the results from the t-tests on prism.treat for the clusters 1-2 and 2-3 gave important information. They demonstrate that there are significant differences in the drug response between the clusters. However, there were not any major differences in knockdown scores and gene expression rates. Thus, this was not the reason for the different drug response. There are significant differences in the drug response for the drugs - “torin-1”, “AVN-944” for the clusters 2-3 and “trichostatin-a”, “semaxanib”, “alvocidib” for the clusters 1-3.