Method Title

Authors

Affiliations

# Summary Sentence:

include 1 short sentence description of your method that can be used in a main text table.

# Background/Introduction

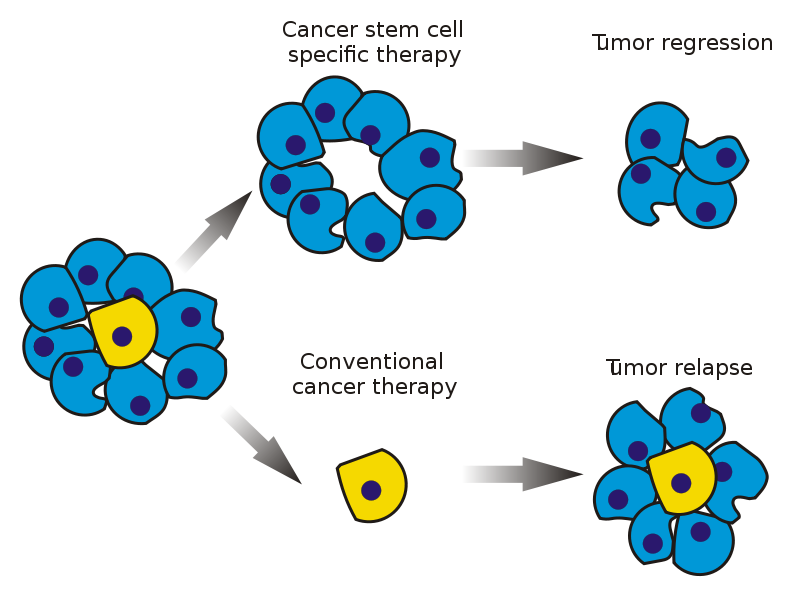
In order to determine which drugs combination are potent at preventing relapse over time, we developed an algorithm based on the following assumptions:

1. Cancer cells always harbor heterogeneity in their transcription profiles. Upon selective pressure (eg Drug exposure), a drug-resistant and a drug-sensitive populations are identified.
2. Prior drug exposure in vitro, rare cancer cells are already intrinsically resistant.
3. The dynamic of transcription profiles are predominantly the result of changes in populations distribution, rather than molecular response of each cancer cell.

Tumors contain heterogeneouscellpopulations that are hierarchically organized. Although individual cancers cells share the same mutations, a rare subpopulation of cancer cells, called cancer stem cells (CSCs) are responsible for the tumor initiation, the generation of non-CSCs that make up the bulk of the tumor and the metastatic colonization. Failure of chemotherapy in cancer patients is frequent and CSCs are suggested to be responsible for tumor recurrence because of their intrinsically resistance. A single chemotherapeutic drug may shrink tumor bulk by eliminating non-CSC, but sparring CSCs. On the other hand, CSC-targeted therapies may be not enough for complete tumor regression because there is a possibility that non-CSCs can become CSCs under critical circumstances owing to their plasticity.

Our approach consists of combining drugs that are the most complementary in terms of population targeting.

NCBI-DREAM data provides information about the nature of both negatively and positively selected subpopulations following drug exposures. We reasoned that the combination of one drug targeting CSCs and a second targeting non-CSCs would be the most effective pair. Our method does not rely on SNP profiles, IC20 values, nor incorporation of external database.



Stem cell specific and conventional cancer therapies

*Source: Wikipedia*

# C:\Users\JEAN-PAUL\Dropbox\Script\dream 2nd challenge\manuscript\Untitled-1.pngMethods

## Identification of differentially expressed genes over times following drug exposure

Static methods, such as Limma and SAM, are used to identify differentially expressed genes in samples with multiple biological conditions. When time-series data are available, the time-dependent structure offer the advantage to improve the sensitivity of detection. Noisy signal is sustained across time point and can be eliminated. This improves detection of small signals that cannot be identified when using only static methods.

Bayesian Estimation of Temporal Regulation (BETR) is a new algorithm that takes into account the correlation between successive time points and estimate the probability that a gene is differentially expressed between two conditions over time, without the need to have balanced sample sizes. We used BETR R package to select probes with differential expression for each drug against DMSO. The cut-off used was >95%.

## Filtering of probes with a dose-response relationship for each time point

In order to select the most relevant probes for each drug, we assume that a linear correlation between drug dosage (0, 1/10 of IC20 and IC20) and subset of probes exists. We perform a linear regression for each probe at a given time and determine quality of the fitted model by one-way ANOVA.

## Clustering of drug-specific probes into drug-response signatures

For a given drug and time point, probes were clustered into two signatures depending on the sign of the slope. In total , 84 signatures were obtained (14 drugs x 3 time points x 2 up/down probes). We considered signatures with down-regulated genes as representative of the drug effect (drug-effect signature, ***DES***) and signatures with up-regulated genes as representative of the compensatory mechanism the tumor cell uses to survive (drug-surviving signature, ***DSS***).

## Derivation of the drug-drug interaction matrix

We cross-compared all drug-specific signatures against all samples by Nearest Template Prediction and used cosine distances (FDR<0.05) to build an interaction matrix

* Cosine distances values are summed up when a treated sample is associated to a given ***DSS***
* Cosine distances values are subtracted when a treated sample is associated to a given ***DES***

Scores are calculated for each timepoint and averaged together to produce a final interaction matrix, describing which combinations are the most potent.

The synergistic effect is predicted when the combination of two different drugs gives a higher score than the drug-specific signature with the same drug-treated sample.

# Discussion

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BETR | Linear drug-gene at 6h | | Linear drug-gene at 12h | | Linear drug-gene at 24h | |
| **# Probe** | **DES** | **DSS** | **DES** | **DSS** | **DES** | **DSS** |
| Aclacinomycin | 633 | 0 | 52 | 0 | 346 | 0 | 6 |
| Blebbistatin | 6519 | 0 | 413 | 5 | 1910 | 0 | 294 |
| Camptothecin | 14701 | 51 | 6240 | 1 | 5231 | 71 | 4177 |
| Cycloheximide | 18177 | 16 | 4671 | 2 | 4329 | 421 | 5602 |
| Doxorubicin | 10211 | 3 | 1439 | 27 | 2408 | 2 | 2779 |
| Etoposide | 2761 | 0 | 701 | 2 | 621 | 3 | 636 |
| Geldanamycin | 2460 | 1 | 958 | 0 | 597 | 7 | 520 |
| H7 | 13368 | 9 | 3772 | 5 | 3017 | 50 | 759 |
| Methotrexate | 2912 | 17 | 166 | 2 | 270 | 12 | 1511 |
| Mitomycin | 2656 | 0 | 218 | 0 | 268 | 27 | 554 |
| Monastrol | 9334 | 1 | 1190 | 6 | 2160 | 178 | 2035 |
| Rapamycin | 8850 | 9 | 1515 | 2 | 1975 | 8 | 1217 |
| Trichostatin | 17440 | 7 | 5557 | 2 | 4982 | 1012 | 2404 |
| Vincristine | 573 | 0 | 59 | 0 | 189 | 0 | 105 |

We observe that BETR algorithm was more efficient to detect more subtle differently expressed genes, compared to commonly used methods. Since drug responses are delayed in time, we suggest that using 6 hour as the baseline could have improved the identification of more relevant genes.

The filtering of genes with linear relationship between probe level and drug concentration has severely affected the predictive power of our algorithm. Indeed, few genes were found in DESs (a.i. positive correlation to drug concentration. We suggest that a logistic model would be more suitable for the dose-response relationship. Additionally, the range of concentration used in NCBI-DREAM data could have been too narrowed to capture those important probes (a.i. IC20 and IC50 instead of IC20 and 1/10th of IC20). The unbalanced signatures sizes must have affected the predictive power of NTPez algorithm.

Then, DSSs could have been reduced by keeping probes with sustained pattern over time, since the cells were keep in the drug over the full timecourse. This could have increased the specificity by eliminating false positive probes and probes with circadian regulation.

Finally, our usage of NTPez algorithm could be improved by pre-processing the drug-treated sample by differential expression against all others-drug treated samples. We suggest this approach improves the ranking prediction of drug pairs without synergistic nor antagonistic effects. We could have also used NTPez without time-matching restriction between DES/DSS and treated samples. Indeed, the drug responses could be delayed in time and pre-conditioning tumor with one drug could potentiate the toxicity of a second drug later on.

# Conclusion

Our ability to understand tumor response offers the possibility to anticipate the escape mechanism that tumors use. We are grateful to NCBI-DREAM organizers for the stimulating challenges they provided and the opportunity to gain insights about cancer resistance.

# References

Betr, NTPez

# Authors Statement