Result Colosys HT29 screen for drug enhancement MK/CHEK combintaion.

Nov 25: Seeding HT29-dox-dCAS9-KRAB16 in 16 10cm dishes

Nov 26: Count HT29 cells in one dish: 5x10^6 cells. Total 75x10^6 cells

Infect with 5 ml lentiviral sgRNA library 1:10

Nov 27: Add 2 ug/ml puromycin to each plate (+control non-infected well)

Nov 28: Trypsinized 15 10 cm dishes (190x10^6 cells) split into 15 15 cm dishes

Dec 1: Trypsinized 15 15cm dishes 3x10^8 cells total, split in 75 15cm dishes with 1.5x10^6 cells per 15cm dish

Dec 2: Addition of drugs to the dishes in 3 replicates (5 plates each) in 2,5ml medium each plate. Doxycycline to 1ug/ml to induce library expression.

5 Arm Screen:

1. DMSO
2. Chk1i 0,25uM
3. Chk1i 0,75uM
4. Mk2i 1uM
5. Chk1i 0,25uM and Mk2i 1uM

Dec 5: Again 1,5 million cells were seeded in 15cm dishes, 5 plates per replicate. The remaining cells were washed with PBS, pelleted and frozen at -80C (T1, 5.12).

Dec 11: Again 1,5 million cells were seeded in 15cm dishes, 5 plates per replicate. The remaining cells were washed with PBS, pelleted and frozen at -80C (T2, 11.12).

Dec 17: Cells were harvested, washed with PBS and frozen at -80C (T3,17.12)

Jan 6-23: Genomic DNA isolation T2, 16 96 well plate PCRs 1, pooling, 32 PCRs 2, normalization, pooling, gel purification.

Jan 23: Hand in samples for sequencing

Feb 10: Sequencing run finished

Feb 11: Mapping completed

Feb 18: Analysis completed

**Quality control**

99% of the library sgRNAs is present in the dataset. In the samples, 90% or more of the sgRNAs have counts above 100. The distribution of the counts over the sgRNAs is a bit skewed ranging from 0 to 10,000. (See ‘norm\_plotRankCount.jpg’).

The sample for "c0.75-m0\_r3" has barely counts, so we removed this one before normalization. Normalization was done with a relative total size factors (See ‘sizefactors.csv’). After normalization a value of 1 was added to the counts to prevent later Infinitive values when calculating log2FoldChange values. (See ‘counts\_histogram.pdf’).

The clustering of the samples shows two big clusters: one ‘untreated (like)’ cluster with the samples for untreated, and single treatments c0-m1 and c0.25-m0 . The other big cluster is with the samples for the combination treatment ‘c0.25-m1’ and c0.75-m0 samples. (See ‘norm\_heatmap.pdf’).  
  
The correlation between the replicates is good. The R^2 (=explained variance) is in most cases above 0.85. An exception is the combination treatment with a value around 0.75. In dropout screen we see in typically for t0: around 0.9, for untreated around 0.8 and for treated around 0.7. (See ‘norm\_log10\_correlation\_plots.jpg’).  
  
The essential genes drop as expected in all conditons, though a bit stronger in one condition compare to the other. (See ‘prob\_distr\_per\_type\_per\_rep.pdf’).

**Statistical Analysis**  
On the sgRNA level, a differential analysis between pairs of conditions was done, using DESeq2. A sorted list based on the DESeq2 statistic was then put into the Robust Rank algorithm of MAGeCK. To the RRA output file we added per gene the median of the log2FoldChange over all the six sgRNAs. See for example ‘deseq2\_ref-ut\_exp-c0.25-m1\_rra\_n.tsv’). In the name ‘ref’ stands for the reference condition in the comparison, and ‘exp’ for the experimental condition. All calculations are done exp vs ref.

As a threshold for hit selection we used a FDR <= 0.1. In the analysis ‘ref-ut\_exp-c0.25-m1’ there are eight hits: POLA1, CLSPN, ABCG2, PPP2R2A, PKMYT1, POLE2. APEX2,   
MAP2K1. We disregarded APEX2 as a hit as the median of log2FoldChange was around 0. RRA called this a hit because it used only two out of six sgRNAs for this gene for the analysis.

To check that difference in treated vs untreated is caused by a decrease in the treated condition and not an increase in the untreated condition, we generated a heatmap with the log2Fc values compare to T0. (See ‘heatmap\_witht0s\_comb\_treatment\_hits.pdf’). Only for MAP2K1 and ABCG2 a part of the difference is explained by an increase in untreated condition. That ‘tr\_ut’ is not always exact ‘tr\_t0’ minus ‘ut-t0’, is due to the fact that the arithmetic mean was used to calculate the mean over the replicates for the log2FoldChange calculation. Furthermore, the median log2Fc was taken over the sgRNAs per gene.

**Synergy calculation.**

We performed a synergy calculation using the BLISS Independence model. An expected additive value was calculated by taken the sum of the log2foldChange of the single drug treatments: c0-m1 and c0.25-m0. Then this additive value was subtracted from the combination treatment (c0.25-m1) resulting in an excess value. For all hits except ABCG2 there is negative excess value, indicating synergy. (See ‘synergy\_results.csv’ and ‘synergy\_result\_hits.pdf’).

**Sofware used**

R : 3.5.2

DESeq2: 1.22.2

MAGeCK (RRA): 0.5.8.