Supplementary figures

Supplementary Fig 1

Comparison of drop out phenotypes in MKN45, RKO, HT1080 highlighting selected pan-lethal genes. The dropout phenotypes are calculated as the average Z-scores of all 20 reagents per gene, and the average shRNA Z-score is displayed on Y-axis and average CRISPR Z-score on the X-axis. Quadrant III contains genes that score as lethal by both CRISPR and shRNA. A few known essential genes are marked by colored dots and indicated in the legend. Quadrant II (red) contains genes that scored lethal only by CRISPR and quadrant IV (grey) contain genes that scored lethal only by shRNA.

Supplementary Fig 2

The genes that scored as lethal by both RNAi and CRISPR were strongly enriched for known essential genes classes. The most statistically significantly enriched GO biological process categories for genes that scored as lethal in DLD1 and RKO are displayed.

Supplementary Fig 3

To identify likely off-target hits the lethality scores of non-expressed genes were examined, as they are expected not to be required for cell viability. The lethality Z score of each gene in CRISPR screens (X-axis) was graphed against its expression level as determined by RNAseq (Y-axis) for the indicated HT1080, RKO and MKN45 lines.

Supplementary Fig 4

shRNAs directed towards CDK9 do not show robust protein depletion. **A.** Lethality Z-score of 6 individual shRNAs used in our screen. **B.** Western blot demonstrating poor protein depletion, only one shRNA had a viability effect and the only one showing moderate protein depletion.

Supplementary fig 5

Additional methods measuring the proliferation effects of individual sgRNA/shRNAs to validate the impact that targeting selected genetic dependencies have on cell viability. A, C, Proliferation of the cell

lines DLD-1 and HT1080 were measured using live cell time-lapse imaging (Incucyte) and the percent confluence at day 6 was graphed. B, D, proliferation was also measured using cell titer glow after 6 day in culture. Similar to the results in our screens, both methodologies showed that depletion of b-catenin (CTNNB1) selectively impaired the proliferation of DLD1 (APC mutant) but not HT1080 cells. Conversely, depletion of NRAS selectively impacted the growth of HT1080 but not DLD1 cells. By contrast, depletion of the mitotic kinase PLK1, which is a broadly essential gene, impacted the proliferation of both cell lines when CRISPR was used but only showed a modest effect by shRNA. Each bar in the graph represent the mean ± SEM (n=2) (*p<0.05, **p<0.01 and p<0.001)

Supplementary Fig 6

Correlation analysis displaying features that correlated most significantly with sgRNA potency.
This analysis was restricted to genes whose activity were required for cell growth (Avg.Z scores < -0.4). Genomic location within conserved PFAM protein domains was the feature that correlated most strongly with sgRNA performance, as well as the extent of sequence conservation across vertebrate species (p<<0.001) regardless of whether or not the region was annotated as a conserved Pfam domain. Doench-Root score (Score based on the nucleotide features of the sgRNA(1)), entropy (Shannon entropy of the 20mer guide sequence), Longest H run (Longest homopolymer nucleotide run in the 20mer guide sequence), phastCons vertebrate conservation (PhastCons(1, 2) conservation score based on the multiple alignments of 100 vertebrate genomes, including human), ExonSize (size in bp of the exon)

Supplementary Fig 7

Effect of relative position within a gene on sgRNA viability effects. Red line represents all sgRNA's with an Average Z score < -0.4. **A,** sgRNAs targeting Pfam domains as well as regions outside Pfam domains. **B,** sgRNA's targeting only PFAM domains.

Supplementary Fig 8

Non-scoring sgRNA in conserved Pfam domains have a reduced editing efficiency compared to guides with strong viability effects. **A.** Example of how sgRNA performance is influence by gene position in the b-catenin locus. . Each dot represents the score of an independent sgRNA, with grey dots indicating sgRNA's targeting regions outside of a Pfam domain, and orange dots indicate sgRNA's targeting PFAM domains. The black line indicates the average dropout score for the neighboring 10 sgRNAs. The protein domain structure of the respective genes is displayed on top with key Pfam domains labeled. Highlighted sgRNAs were chosen for next generation sequencing only 1 out of the 4 sgRNAs chosen had a lethality score < -1. **B.** Viability was measured by CellTiter glow, only 1 out of 4sgRNAs had a robust decrease in viability. **C.** Frequency of frameshift (aqua) and in-frame mutations (green) at the CTNNB1 locus introduced by the sgRNAs described in A and B.

Supplementary Fig 9

Multiple genomic cuts result in DNA damage induced G2/M cell cycle arrest. **A.** western blot for phospho-H2AX in response to sgRNA mediated cleavage of CN amplified region (MMP7, MMP20 and ANGPTL5) or multiple off-target cutting (sgRNAs labeled VEGFA and OR4F5). **B.** cell cycle analysis in response to sgRNA mediated cleavage of CN amplified region (MMP7, MMP20 and ANGPTL5) or by multiple off-target cutting (VEGFA and OR4F5). **C-D,** shows the quantitative analysis of the number of cells at G2/M or subG1 following sgRNA mediated cleavage of CN amplified region (MMP7, MMP20 and ANGPTL5) or after multiple off-target cutting (VEGFA and OR4F5). Each bar in the graph indicates the mean ± SEM (n = 3).

Supplementary Fig 10

Multiple genomic cuts lead to an increase in cell death. **A.** contour map demonstrating an increase in apoptotic cells following sgRNA mediated cleavage of CN amplified region (MMP7, MMP20 and ANGPTL5) or after multiple off-target cutting (VEGFA and OR4F5). **B.** shows the quantitative analysis of the number of cells undergoing apoptosis following sgRNA mediated cleavage of CN amplified region

(MMP7, MMP20 and ANGPTL5) or after multiple off-target cutting (VEGFA and OR4F5). Each bar in the graph indicates the mean ± SEM (n = 2). (**p<0.01 and p<0.001).

Supplementary Fig 11

Pie chart demonstrating that the overall contribution of copy number effects in determining essential genes in aneuploid lines is relatively minor. A Z score cutoff of -1 was used to delineate genes that are cell essential.

Supplementary table 1

Correlation analysis matrix displaying features that correlate most strongly with sgRNA's having off target effects. 12mer matches to genome (Number of sequence sites in any gene with an exact match to at least 12 contiguous bases in the guide), 20/20 matches to genome (Number of sequence sites in the genome with an exact match to the 20 base pair guide sequence. This value will always be at least one for the intended location), 19/20 matches to genome (Number of sequence sites in the genome with 1 mismatch to the 20 base pair guide sequence), 20/20 matches to genes (Number of sequence sites in any gene with an exact match to the 20 base pair guide sequence. This value will always be at least one for the intended location), phastCons vertebrate conservation (PhastCons(1, 2) conservation score based on the multiple alignments of 100 vertebrate genomes ,including human) and Doench-Root_score (Score based on the nucleotide features of the sgRNA(1))

Supplementary table 2

Annotations for the sgRNA libraries containing sequences, target gene information, Zscore and RNAseq data for each cell line screened

Supplementary table 3

Annotations for the sgRNA/shRNA libraries containing average Zscore and copy number information

Supplementary table 4

Annotations for the tiling library containing sequences and target gene information

Supplementary references

1. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research. 2005;15:1034-50.

2. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Human mutation. 2011;32:894-9.