Appendix A

Datasets that were too large to be included in the main thesis have been provided in a digital format and have been referenced throughout. The contents of each file are detailed below.

A.1 Oligonucleotide sequences

Sequences and details are provided for all oligonucleotides, primers, and gRNAs used for the work described in this thesis. They are grouped into different sheets based on the section of work that they relate to.

A.2 Mass spectrometry data

The $\log_2$(fold-change) in protein abundance for each KO line (compared to the parental BOB line) is shown for all targeted proteins.

A.3 NeoR-IRES library backbone

The plasmid map and full sequence of the final neoR-IRES library backbone is provided.

A.4 gRNA library read counts

The raw gRNA read counts from HiSeq2500 analysis of the neoR-IRES, Yusa v1, and Yusa v1.1 gRNA libraries are provided.

A.5 Bayes Factors for iPSC screens

The gene-level BFs obtained from BAGELR analysis of each screen are shown.

A.6 Scaled Bayes Factors for iPSC screens

Gene-level BFs were scaled based on an FDR of 0.05, with any value greater than 0 representing a significant hit. Scaled BFs for all screens are shown.

A.7 MAGeCK depletion values for iPSC screens

The gene-level depletion (negative FDR) values obtained from MAGeCK analysis of each screen are shown.

A.8 Biological replicate overlaps

Genes that were significantly depleted in the biological replicates of the parental BOB and *TP53* KO line screens are shown.

A.9 KO-specific depleted genes

This file shows genes that were significantly depleted in each KO line (based on BAGELR analysis) but not in the parental BOB line screens. Results are provided for three filtering strategies. Genes with a scaled BF > 0 (highlighted in red) were significantly depleted.

A.10 MAGeCK enrichment values for iPSC screens

The gene-level enrichment (positive FDR) values obtained from MAGeCK analysis of each screen are shown.

A.11 PBAF/BAF mutant cancer cell lines

Of the cell lines screened by the Sanger[109] and Broad,[261] those lines that harbour LOF (nonsense or frameshift indel) mutations in *ARID1A, ARID1B, ARID2* or *PBRM1* are listed. These were used for the association analyses described in Chapter 4.

A.12 Sanger Institute screen data

Scaled BFs (FDR 0.05) are provided for all cancer cell line screens screened by Sanger, as published by Behan *et al.* (2019).[109]

A.13 Broad Institute screen data

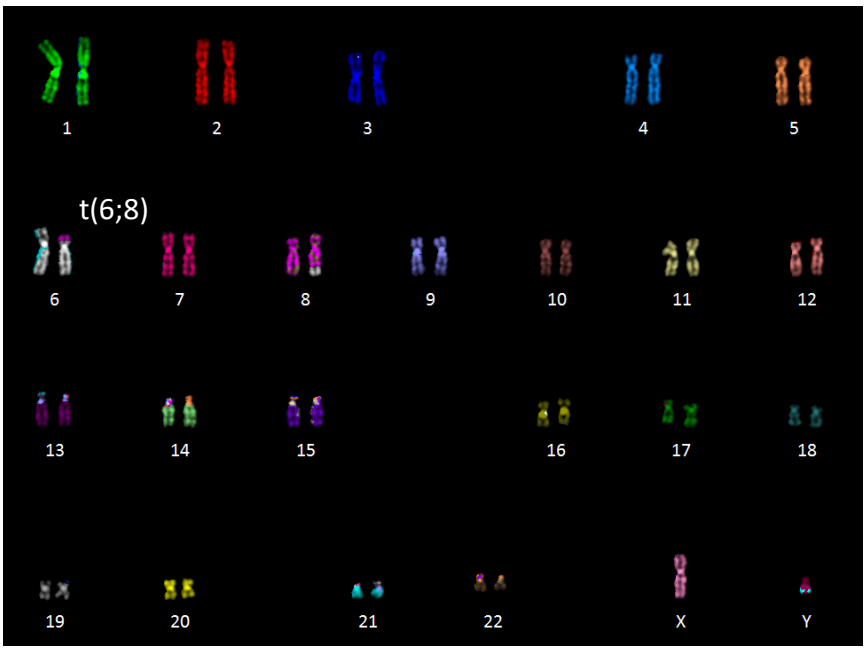
Raw gRNA counts from the Broad cancer cell line screens[261] were re-processed using the pipeline applied by Behan *et al.*[109] (by Clare Pacini, a postdoctoral fellow in Francesco Iorio's lab). Scaled BFs (FDR 0.05) obtained from BAGELR analysis are shown for all lines.

A.14 Candidate PBAF-BAF gene SLIs

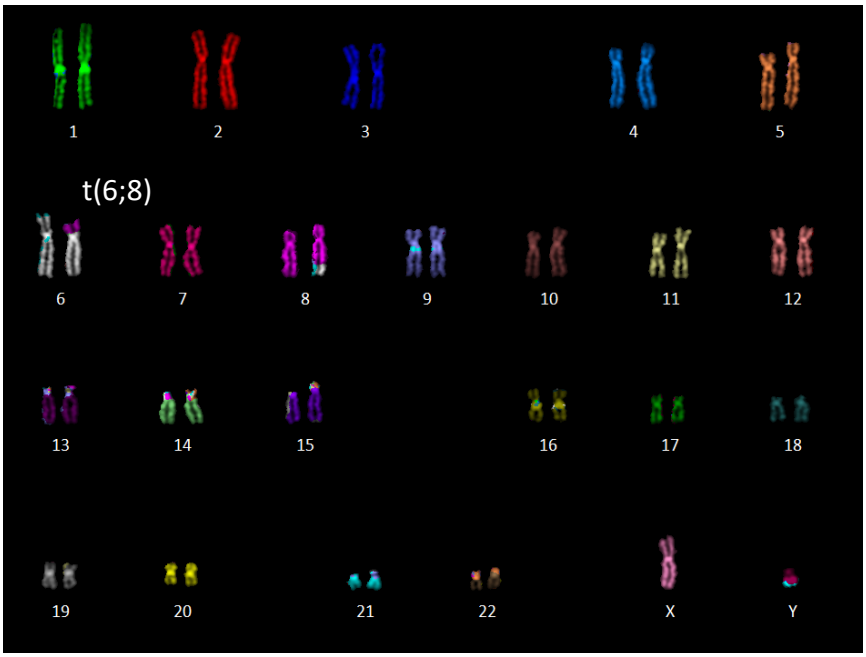
Genes that were identified after filtering for hits specific to PBAF/BAF gene KO lines (as described in Section 4.3) are provided, with the scaled BFs shown for all screens that were used in the filtering. Their status at each stage of the filtering is indicated in the final columns.

Appendix B

a.



b.



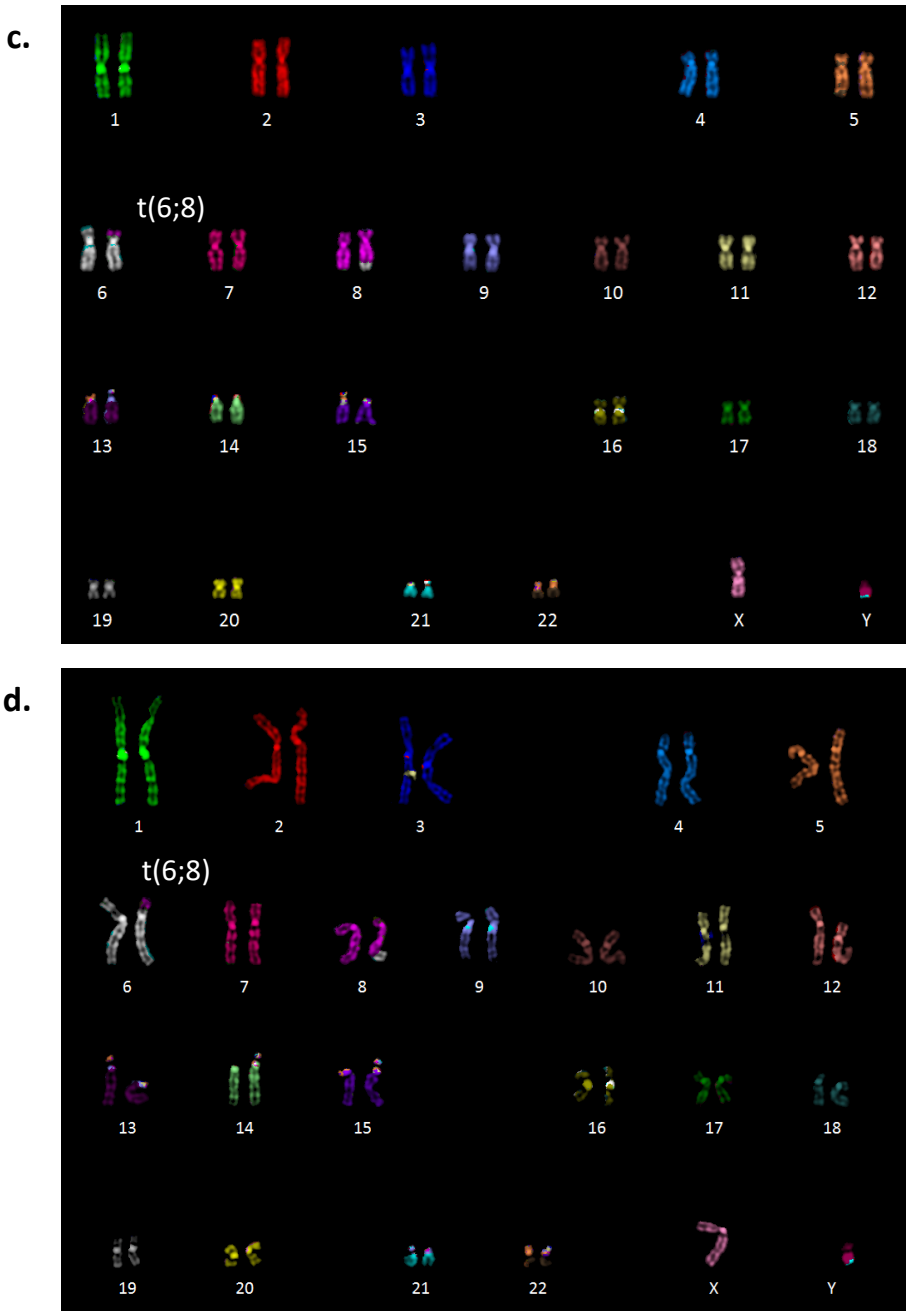


Figure B.1. Karyotype of KO BOB-Cas9 cells. M-FISH of ARID1A_C09-Cas9 **(a)**, ARID1B_C03-as9 **(b)**, ARID2_C11-Cas9 **(c)**, and PBRM1_F09-Cas9 **(d)** cell lines. Ten randomly selected metaphases were analysed for each line; a representative karyotype is shown for each. A balanced translocation between chromosomes 6 and 8 is indicated.

Table B.1. Essential genes identified in parental BOB screens and published hESC screens.

Gene
CDIPT
CYC1
FANCA
FDXR
H2AFZ
MBTPS1
MRPL50
RPL32
RPL34
RPS25
SCAP
SDE2
SETD1B
SNRNP40
TBCB
WDR1
CTDSPL2