Peck et al

3D functional genomics screens identify CREBBP as a targetable driver in aggressive triple negative breast cancer

Supplementary Methods

Cell-based assays

Short-term drug exposure assays were performed in 96 well plates for the days indicated. Cells were plated at a density of 5000 cells/well in ultra-low attachment plates (cls7007, Corning), with drug added at the indicated concentrations 24 hours later. Cell viability was measured using CellTiter-Glo (Promega). Surviving fractions (SF) were calculated relative to DMSO-treated wells from which sensitivity values were calculated. Doxycycline and Palbociclib were purchased from SigmaAldrich (UK). Abemaciclib and Vorinostat, Tricostatin A, Ribociclib and SNS-032 were purchased from Calbiochem (UK). For *in vivo* studies Palbociclib was obtained from Pfizer as a gift. At least three biological replicates were performed for each assay.

RNA extraction and RT-PCR

RNA was extracted using Trizol (Life Technologies) according to the manufacturer's instructions. Nucleic acids were quantified using the Qubit Fluorometer assay (Life Technologies). Gene expression was quantified using Quantitect primers according to the manufacturer's instructions. Relative gene expression was calculated by the $\Delta\Delta$Ct method. Assays were performed in triplicate.

Western blot

HAP1 spheroids were lysed with lysis buffer (50mM Tris pH 7.5, 150mM NaCl, 1mM EDTA, 5mM MgCl2, 10% glycerol and 0.1% NP-40) supplemented with phosphatase and protease inhibitors for at least 10 minutes on ice. NuPAGE® LDS sample buffer (Thermo Fisher

Scientific) supplemented with β-mercaptoethanol (Sigma-Aldrich) was added to the samples, followed by incubation at 95°C for 5 minutes. Samples were loaded on 7-14% NuPAGE® Bis-Tris gels (Invitrogen), run at 120V for 90 minutes with MES-buffer (Invitrogen), and transferred to PVDF membranes (Merck Millipore). Membranes were then incubated in TBS supplemented with 0.1% Tween 20 (TBS-T) (v/v) and 5% skimmed milk powder for 30-60 mins at room temperature, followed by overnight incubation with primary antibodies at +4°C or 1 hour at room temperature. After washing with TBS-T, membranes were incubated with secondary antibodies for 1 hour at room temperature, washed with TBS-T and exposed to ECL reagent (Thermo Scientific). Where indicated, membranes were stripped by incubation at +50 °C in SDS-based stripping buffer (62.5mM Tris HCl pH 6.8, 0.5% SDS, 0.7% β mercapthoethanol) for 30 mins.

Immunohistochemistry of spheroid cultures

Spheroids were grown for 7 days and fixed in 3.8% formaldehyde for 30 min, washed with PBS three times and stored at 4 °C. Spheroids were then pooled, dehydrated, embedded in paraffin and sectioned. The spheroid sections were then de-paraffinised with xylene, rehydrated, microwaved and then incubated overnight with primary antibodies against Ki67 (Dako, MIB-1 F726801-8, pressure cooker, pH6 citrate 2 min at 127 ^{0}C, diluted 1/300), CREBBP (SigmaAldrich, HPA055861, RRID:AB_2682948, pressure cooker, pH6, 1:200 dilution), CAIX (Abcam, ab15086, RRID:AB_2066533, microwave, pH6, diluted 1:1000), FOXM1 (Santa Cruz, RRID:AB_11150135, pressure cooker, pH6, 1:500 dilution). Staining was visualised using 3,3'-diaminobenzidine (DAB) secondary antibody. IHC was quantified by consultant pathologist (IR); Ki67 % positive tumour cells; CREBBP and FOXM1, modified Allred score as detailed in the main methods.

Proteomic profiling

Cell pellets were lysed in 5% SDS/ 100mM TEAB buffer with probe sonication and heating at 95°C for 10min and 300 µg of protein were taken for each sample. Proteins were reduced with

TCEP and alkylated by iodoacetamide followed by TCA (trichloroacetic acid) precipitation. Trypsin (Thermo-Fisher) was added at 1:25 (trypsin:proteins) for overnight digestion at 37^0C. 100 µg - 200 µg of peptides per sample were TMT labelled as instructed by the manufacturer. The TMT labelled peptide mixture was fractionated on a BEH XBridge C18 column (2.1 mm i.d. x 150 mm) with a 35 min gradient from 5 – 35% CH3CN/NH4OH at pH 10. A total of 36 fractions were collected and subjected to IMAC enrichment using the Phos-Select resin (Sigma) and analysed on an Orbitrap Fusion Lumos coupled with an Ultimate 3000 RSLCnano System. The enriched phosphopeptides (from both sets) and the IMAC flow through (used for whole proteome analysis) were analysed on an Orbitrap Fusion Lumos coupled with an Ultimate 3000 RSLCnano System. Samples were loaded on a nanotrap (100 µm id x 2 cm) (PepMap C18, 5 µ) at 10 µL/min with 0.1% formic acid (FA) and then separated on an analytical column (75 µm id x 50 cm) (PepMap C18, 2µ) over a 90 min gradient of 5 – 32% CH3CN/0.1% FA at 300 nL/min. The Orbitrap Fusion was operated in the Top Speed mode at 3 s per cycle. The survey scans (m/z 375-1500) were acquired in the Orbitrap at a resolution of 120K (AGC 4x105 and maximum injection time 50 ms). A lockmass at m/z 445.12003 was also set. The multiply charged ions (2-7) were subjected to HCD fragmentation with a collision energy at 38% and isolation width 0.7 Th. MS/MS spectra were acquired in the Orbitrap (AGC 1x105 and maximum injection time 105 ms) with 50K resolution. Dynamic exclusion width was set at ± 7 ppm for 40 s.

All raw files were processed in Proteome Discoverer 2.2 (Thermo Fisher, RRID:SCR_014477) using the SequestHT and Mascot (V2.3) search engines for phosphopeptide analysis, or SequestHT only for total proteome analysis. Spectra were searched against a FASTA file containing reviewed Uniprot Homo sapiens entries (January 2018). Search parameters were: trypsin with 2 maximum miss-cleavage sites, mass tolerances at 25 ppm for Precursor, and 0.1 Da for fragment ions, dynamic modifications of Deamidated (N, Q), Oxidation (M) and Phospho (S, T, Y), and static modifications of Carbamidomethyl (C) and TMT6plex (peptide N-terminus and K). Peptides were validated by Percolator (RRID:SCR_000287) with q-value

set at 0.01 for the decoy database search. Protein FDR was set at 0.01. Phosphorylation site localization probabilities were computed by the ptmRS node. The TMT10plex reporter ion quantifier included 20 ppm integration tolerance on the most confident centroid peak at the MS3 level. Only unique peptides were used for quantification. The co-Isolation threshold was set at 50%. Peptides with average reported S/N>3 were used for protein quantification. Only master proteins were reported. For full proteome analysis TMT6plex (K) and acetylation (K) were set as dynamic modifications.

Phosphosite motifs were identified using PhosphositePlus (18) with a sequence window of +/- 7 amino acidsand phosphosites with mutant/wild type fold change higher than 1.5 were then analysed by Motif-All algorithm (19). A significance threshold of 1x10-6 and a support threshold of 0.05 were used for the analysis. The motif frequency plot of up-regulated phosphosites against the dataset was generated using TwoSampleLogo software (http://www.twosamplelogo.org/index.html) with the default amino acid colour scheme. CDK4/6 motifs were identified to be enriched based on previous studies (20,21).

Single-cell RNA-sequencing

Single-cell 3' RNA-sequencing was performed on the isogenic HAP1 WT and CREBBPmutant spheroids using the 10X Chromium Genomics platform using standard protocols and sequenced on the Illumina Nextseq. Briefly, 4000 cells in total were loaded to the 10X Chromium™ Chip A as per the manufacturers protocol. cDNA amplification and clean up was performed on all samples using the Chromium™ Single Cell 3' Library & Gel Bead Kit v2, followed by quality control using the Agilent 2200 Tapestation (Agilent Technologies) and D5000 screentape to assess cDNA size, and the Qubit® 2.0 spectrophotometer with the dsDNA High Sensitivity Reagent Kit (Life Technologies) for quantification. 35µL of each cDNA sample was taken forward for library preparation, with 12 cycles of index PCR. Library quality control was carried out using the D1000 screentape and Qubit® dsDNA High Sensitivity reagent kit. Sample libraries were combined in equimolar amounts into a single pool. The final

library pool was loaded to the NextSeq500 at 1.1pM and sequencing carried out with 26bp for Read 1, and 98bp for Read 2 using the Illumina NextSeq®500 High-Output 150-cycle kit.

Raw bcl data from the NextSeq was processed with Cellranger (v.2.2) software (10x Genomics). Data was demultiplexed with Illumina's bcl2fastq tool using the cellranger mkfastq command. Following this, alignment to the hg19 human transcriptome, filtering, cell barcode and UMI (unique molecular identifier) counting was performed using the cellranger count command. The Chromium cellular barcodes are used to perform sample clustering and gene expression analysis, and results were visualised using the Loupe Cell Browser software (10X Genomics). In total 1324 WT and 1398 CREBBPmutant were sequenced cell. Single cell transcriptomes were imported into R and processed with monocle 2.6.4 (22). Single cell expression profiles were then normalised using the median ratio method implemented in the 'estimateSizeFactors' function. Lowly expressed genes were filtered keeping those with a mean normalised count of 0.1 across cells resulting in 8779 and 9014 genes in WT and MT respectively. These genes were retained for dimensionality reduction where cells were ordered in pseudotime according to differentiation trajectories and cell state transitions were reconstructed with the reverse graph embedding algorithm 'DDRTree'. Single cell data has been deposited in the European Nucelotide Archive (ENA) with accession number PRJEB32846.

Calculation of transcriptional heterogeneity

Normalised bulk RNA sequencing data from TCGA was quantified at the exon, junction, isoform and gene levels, which were obtained from (https://gdac.broadinstitute.org/, RNASeqv2, Level 3). Junction and exon level data were normalised using the median ratio method (23). Shannons's equability was calculated for each level compared between CREBBPaltered and WT groups with a Welch's t-test. The Shannon's equitability is a normalized version of Shannon's index, in which ''0'' represents no heterogeneity and ''1'' represents the highest heterogeneity. Shannon's equitability was chosen as the total number of population (genes) may vary for different samples (24).

NanoString gene expression profiling

Gene expression profiling of CREBBP was performed by hybridising 80-430ng RNA with a custom NanoString gene expression panel for 23 hours, followed by processing with the NanoString SPRINT Profiler (NanoString Technologies, USA) as per manufacturer's instructions. For the Belgrade cohort, raw NanoString data was pre-processed and normalised using R package NanoStringNorm (v1.2.1) (RRID:SCR_003382) (25). We performed grid search over NanoStringNorm's parameter space resulting in the choice of optimal parameters that minimised coefficient of variation in positive controls as well as housekeeping genes. These parameter settings were: 'sum' of positive controls, 'mean' of negative controls and 'total.sum' of top genes. Normalised data was $\log_2$ transformed for subsequent analysis.

***Ex Vivo* Assessment of Palbociclib Efficacy**

Ex vivo assessment of Palbociclib sensitivity in patient derived models CTG-0869 and CTG-2055 (Champions Oncology) were implanted into 8-week old female nu/nu nude mice and allowed to form palpable tumours. Once palpable, tumours were harvested and dissociated using the gentleMACS™ Dissociator (Miltenyi Biotec) and cells plated into 96 well plates containing Matrigel. 24 hours post seeding, Palbociclib was added in increasing doses with 6 replicates per dose. Media and drug were replaced on day 6 and 11. Viability was measured using Cell Titre Glo after 14 days treatment. 10% DMSO was used as a positive killing control.

WHIM02 and WHIM30 PDX models were obtained through collaboration with Prof. Ellis & Dr. Li, Washington University in St. Louis (26) and were passaged orthotopically in female NSG hosts. PDXO cultures were established in collaboration with OcellO B.V (https://ocello.nl/). Briefly, tumour samples were coarsely minced with scalpels and then dissociated using tumour digestion media and the GentleMACS dissociator (Miltenyi). The resulting cell suspension was mechanically disrupted, filtered and contaminating mouse stromal cells were removed using Mouse Cell Depletion Kit (Miltenyi Biotec). Resulting cell pellets were plated into 3D culture at approximately 1e3 to 2e3 cells per ul hydrogel. Optimal hydrogel & media was determined for each PDX line on the basis of conditions that favoured the most efficient

growth in initial early-passage cultures. For subsequent passaging, maintenance in culture and analyses, the most favourable media condition was used. All cultures were maintained in humidified incubators at 37°C, 5% CO2.

WHIM PDXO lines were used between passage 10 and 25 for drug response assessment. Organoids were gently dissociated to near–single cell populations using TrypLExpress (Life Technologies) prior to re-suspension within a 1:2 dilution of hydrogel and medium. Cell suspensions were then dispensed into 384 well plates using a liquid handling system (Hamilton Robotics). To ensure optimal seeding densities across all assays, cell numbers were optimised depending on the growth rate of each organoid line, within the region of 200 cells/ µL – 400 cells/ µL. A total volume of 35 µL of media was then dispensed within each well. For Palbociclib assays, a dose range was added to each well 48 hours after seeding, with DMSO used as a negative control at a final concentration of 0.1%. Organoids were exposed to Palbociblib for a total of thirteen days, with media-containing drug replenished every four days. For endpoint assays, CellTiter-Glo 3D (Promega) reagent was applied as per the manufacturer's guidelines. Relative Luminescence values were obtained on a Perkin Elmer Victor X5 plate reader. Data are presented as % Survival normalized to DMSO Control. HAP1 CREBBPmut cells were included as a positive control in the same assay run.

Statistical Analyses

Data from The Cancer Genome Atlas were interrogated using cBioportal (accessed January 2019)(27,28). The search term "CREBBP: MUT HOMDEL HETLOSS", were used to interrogate patients with loss of function alterations in CREBBP (CREBBPaltered) for enrichment analysis (Supplementary Table S5). Statistical analyses and graphical presentation were performed with GraphPad Prism 7.0 (RRID:SCR_002798). The results are presented as the mean ± SD or the mean ± SEM and evaluated using an unpaired Student t-test (heteroscedastic, two-tailed; $p<0.05$ was considered to be significant, with * depicting

p<0.05 and ** p<0.001). Associations with CREBBP alterations and survival were evaluated for breast cancer in METABRIC utilising probe ILMN2293692 that mapped to the gene body, dichotomising gene expression into tertiles. Survival curves were analysed by the method of Kaplan-Meier, with a Wald test p-value <0.05 being considered significant. METABRIC's survival data (disease specific survival) was censored at ten years. Multivariable survival analysis was performed by taking into account *CREBBP* expression status along with age, tumour size, node status and tumour grade for the METABRIC cohort. Similarly, for the Belgrade cohort, multivariable survival analysis (Distance metastasis-free survival) was performed taking into account *CREBBP* expression (separately for IHC and mRNA data) status along with age, TNM stage and tumour grade. In the Belgrade cohort, the size of *CREBBP* mRNA-derived high- and low-risk groups was determined by matching these to the proportion of *CREBBP* IHC positive and negative groups. Differential protein acetylation and phosphorylation were calculated comparing average peptide counts between CREBBP mutant and wild-type HAP1 spheroids and multiple correction adjusted for using the FDR method. Pathway analysis was performed using ConsensusPathDb (RRID:SCR_002231) (29). FOXM1 transcriptional activity in patient samples was assessed using the DoRothEA (Discriminant Regulon Expression Analysis) algorithm as previously described using the RNA-seq data from TCGA (30).

References

18. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res **2015**;43:D512-20
19. He Z, Yang C, Guo G, Li N, Yu W. Motif-All: discovering all phosphorylation motifs. BMC Bioinformatics **2011**;12 Suppl 1:S22
20. Anders L, Ke N, Hydbring P, Choi YJ, Widlund HR, Chick JM, et al. A systematic screen for CDK4/6 substrates links FOXM1 phosphorylation to senescence suppression in cancer cells. Cancer Cell **2011**;20:620-34
21. Kitagawa M, Higashi H, Jung HK, Suzuki-Takahashi I, Ikeda M, Tamai K, et al. The consensus motif for phosphorylation by cyclin D1-Cdk4 is different from that for phosphorylation by cyclin A/E-Cdk2. EMBO J **1996**;15:7060-9
22. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods **2017**;14:979-82
23. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol **2010**;11:R106

24. Hinohara K, Wu HJ, Vigneau S, McDonald TO, Igarashi KJ, Yamamoto KN, et al. KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance. Cancer Cell **2018**;34:939-53 e9
25. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. Bioinformatics **2012**;28:1546-8
26. Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, et al. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Rep **2013**;4:1116-30
27. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal **2013**;6:pl1
28. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov **2012**;2:401-4
29. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB--a database for integrating human functional interaction networks. Nucleic Acids Res **2009**;37:D623-8
30. Garcia-Alonso L, Iorio F, Matchan A, Fonseca N, Jaaks P, Peat G, et al. Transcription Factor Activities Enhance Markers of Drug Sensitivity in Cancer. Cancer Res **2018**;78:769-80