Supplementary Information

A common core of Wnt-regulated gene expression changes are more robust in orthotopic xenografts

The shared response of the Wnt-addicted pancreatic and colorectal cancers led us to evaluate the importance of the stromal microenvironment in regulating the response to Wnt inhibition. We compared the effects of ETC-159 on the transcriptome of HPAF-II cells implanted as orthotopic xenografts, subcutaneous xenografts or cultured in vitro. The response to Wnt inhibitor treatment in orthotopic xenografts was correlated to the response in both the subcutaneous model ($r^2 = 0.64$, Figure 3C) and the in vitro model ($r^2 = 0.46$, Figure 3D). However, the magnitude of the responses varied widely between the models.

Differential expression analysis of the three HPAF-II models identified 4,409 genes whose response to ETC-159 was significantly different (interaction test, FDR < 10%) between models. These genes were enriched for processes including cell cycle, ribosome biogenesis (Figure S3B). 51% of all *Wnt-activated* genes responded differently to ETC-159 across the three systems, with many more genes showing robust fold changes in the orthotopic model compared with the other systems (Figure 3E). In particular, cell-cycle associated genes changed most robustly in the orthotopic model (Figures 3C-D and F). For example, *AURKA* decreased by 3.7 fold at 56 hours after the start of PORCN inhibition in the orthotopic system, 2.1 fold in the subcutaneous model and only 1.7 fold in vitro. *Cyclin E1* did not change in vitro but decreased by 1.7 fold and 3.8 fold in the subcutaneous and orthotopic model, respectively (Figure 3F). Negative regulators of the cell cycle such as *CDKN2B* also exhibited differences in their magnitude of response to PORCN inhibition. In addition to cell cycle associated genes we identified several other genes that did not respond to Wnt inhibition in vitro but behaved as WNT targets in vivo (i.e. *EPHB3* and *TGFBI*, Figure S3C). Thus, the overall response to Wnt inhibition was reduced in the subcutaneous model and further blunted in vitro (Figures 3C-F).

Further highlighting the importance of the orthotopic model was the considerable difference in baseline gene expression between models ~30% of all expressed genes (4,992 genes) were differentially expressed even before treatment when comparing in vitro and orthotopic xenografts. This included greater than two-fold increases (moving

from cultured cells to the orthotopic model) in Wnt pathway genes *WNT11*, *WNT2B*, *FZD8*, *FZD4*, and *LRP5*, as well as the Wnt target genes *NKD1*, *SP5*, *LGR5* and *AXIN2* (Table S4). Additionally, 3,515 genes were differentially expressed at baseline when comparing subcutaneous xenografts with the orthotopic model (absolute fold change > 1.5, FDR < 10%) (Figures S3D-E). Not surprisingly, a number of genes more highly expressed in the orthotopic xenografts compared to in vitro were associated with multicellularity, innate immunity, extracellular matrix organization and cell adhesion (Figure S3D).

Supplemental Material and Methods

Tumor growth and mice treatment

HPAF-II cells were obtained from ATCC. Mouse xenograft models were established by orthotopic injection of HPAF-II cells in NOD scid gamma mice. $1x10^6$ HPAF-II cells stably expressing luciferase, resuspended in 50% matrigel were used for orthotopic injection into the pancreas of NOD-scid-gamma (NSG) mice. The growth of the orthotopic tumors was monitored weekly by injecting the mice with a 150 mg/kg D-Luciferin (Caliper # 122796) and imaging mice using an IVIS imaging system (Caliper Lifesciences). Mice were treated with ETC-159 after establishment of tumors. ETC-159 was formulated in 50% PEG 400 (vol/vol) in water and administered by oral gavage at a dosing volume of 10 μL/g body weight. At sacrifice, tumors were resected, weighed and snap frozen in liquid nitrogen or fixed in 10% neutral buffered formalin. MSCV Myc IRES GFP plasmid was a gift from Scott Lowe, MSKCC, NY, NY (Addgene plasmid # 18770). The T58A site in MSCV Myc IRES GFP plasmid was mutated using site directed mutagenesis. These plasmids were used for generating HPAF-II cells stably overexpressing MYC or MYC T58A under the control of CMV promoter.

Western Blot analysis

Tumors were homogenized in 4% SDS buffer using a polytron homogenizer. Equal amount of proteins were resolved on 10% SDS-polyacrylamide gel and transferred to PVDF membranes. Western blots were performed according to standard methods. In some cases the actin load control was examined on a parallel gel due to similar molecular weights of the antigens. MYC antibody (cat# ab32072) was obtained from Abcam and anti-rabbit IgG-HRP (cat# P0448) antibody was obtained from Cell Signaling Technology (Danvers, MA). The blots were developed using SuperSignal West femto substrate (Thermo Scientific; Rockford, IL). The images were captured digitally using the LAS-3000 Life Science Imager (Fujifilm; Tokyo, Japan). For stripping and reprobing the blots "Restore western blot stripping buffer" from Thermo Fisher was used.

Immunohistochemistry and AgNOR staining

Tumors were fixed in 10% neutral buffered formalin and embedded in paraffin. Tissue sections were deparaffinized in xylene and rehydrated in a series of ethanol gradients. The sections were then stained with hematoxylin and eosin using standard

protocol. For Ki67 staining, after antigen retrieval with sodium citrate buffer pH 6.0 for 30 min, the endogenous peroxidase activity was blocked by incubation with H_2O_2. The sections were then incubated overnight with Ki67 antibody (Leica Novocastra, Cat #: NCL-Ki67-MM1) followed by incubation with an HRP conjugated secondary antibody for 1 hour. Incubation with 3,3'-diaminobenzidine chromogen substrate resulted in brown staining of Ki67 positive cells and the nuclei were counterstained with Mayer's hematoxylin. Brightfield images were acquired on a Nikon Eclipse Ni-E microscope. Total percentage of Ki67 positive cells in each sample were quantified using the analysis software from Nikon.

For staining the nucleolar organizer regions, deparaffinized tissue sections were reduced using freshly prepared 1% dithiothreitol solution for 15 minutes. The tissue sections were incubated for 25 minutes in dark with 50% silver nitrate solution prepared in 2% gelatin and 1% formic acid. After staining, the slides were thoroughly washed with distilled water and incubated in 5% sodium thiosulfate solution for 5 min followed by several washes in water. The sections were dehydrated and mounted using DPX (1). Images were acquired using Nikon E microscope.

RNA isolation

Tumors were homogenized in RLT buffer using a polytron homogenizer and total RNA was isolated using RNAeasy kit (Qiagen) according to manufacturer's protocol. The RNA-seq libraries were prepared using the Illumina TruSeq stranded Total RNA protocol with subsequent PolyA enrichment.

RNA-seq analysis

Data processing and QC. Sequences were assessed for quality and reads originating from mouse (mm10) were removed using Xenome (2). The remaining reads were aligned against hg38 (Ensembl version 79) using STAR v2.5.2 (3) and quantified using RSEM v1.2.31 (4). Reads mapping to chrM or annotated as rRNA, snoRNA and snRNA were removed. Genes which had less than 10 reads mapping on average over all samples were removed. We also required that genes had more than 10 reads mapping to them in at least 2 replicates across at least 3 conditions. Differential expression analysis was performed using DESeq2 (5). Independent filtering was not used in this analysis. Pairwise comparisons were performed using a Wald test. To control for false positives due to multiple comparisons in the genome-wide differential expression analysis, we used the false discovery rate (FDR) that was computed using the Benjamini-Hochberg

procedure. Previously published CRC PDX RNA-seq data (6) (GSE69687) were re-analyzed using the same processing pipeline (above). Expression levels of genes were calculated by converting to counts to TPM (Transcripts per Million).

Comparison of transcriptional response across different models: The transcriptional response to PORCN inhibition across the different models was assessed by investigating the expression of protein-coding genes (after removing histone genes) and their response across three different datasets; orthotopic (0 h and 56 h), subcutaneous (0 h and 56 h) and in vitro (48 h ETC and 48 h Veh). Reads originating from mouse (mm10) were removed using Xenome for both the orthotopic and subcutaneous dataset. The reads originating from human aligned against hg38 (Ensembl version 79) using STAR v2.5.2 and quantified using RSEM v1.2.31. Reads mapping to chrM were removed. Genes which had less than 10 reads mapping on average over all samples were removed. DESeq2 was used to identify genes that responded differently to treatment depending on the model, i.e. we performed an interaction test in which the the response is modeled as follows: y ~ model + condition + model:condition, where model is orthotopic, subcutaneous or in vivo and condition is untreated or treated with ETC-159. Pairwise comparisons to examine baseline differences between models were performed using a Wald test. The false discovery rate (FDR) was computed using the Benjamini-Hochberg procedure.

Identification of Myc-dependent and Myc-independent genes: Different models were compared by investigating the expression genes and their response across time in the three different conditions; WT, MYC OE and MYC T58A. Reads originating from mouse (mm10) were removed using Xenome for both the orthotopic and subcutaneous dataset. The reads originating from human aligned against hg38 (Ensembl version 79) using STAR v2.5.2 and quantified using RSEM v1.2.31. Reads mapping to chrM or annotated as rRNA, snoRNA and snRNA were removed. Genes which had less than 10 reads mapping on average over all samples were removed. DESeq2 was used to to identify genes that responded differently to treatment depending on MYC status, i.e. we performed an interaction test in which the response is modeled as follows: y ~ model + timepoint + model:timepoint, where model is WT, MYC OE or MYC T58A and timepoint is 0h, 8 h, 56 h. Pairwise comparisons to examine baseline differences between models were performed using a Wald test. P-values were corrected for multiple testing using the q-value method (7), with genes with an interaction q-value < 0.1 being classified as MYC-dependent.

Time series clustering

Gene-level counts were transformed using a variance-stabilizing transformation and converted to z-scores. Time was transformed using a square root transformation. All genes differentially expressed over time (DESeq2, false discovery rate (FDR) < 10%) were clustered using GPClust (8) using the Matern32 kernel with a concentration (alpha) parameter of 0.001 and a length scale of 6.5. Genes were assigned to a cluster which they had the highest probability of being a member of. Several parameter settings were used to assess the stability and robustness of clusters generated, with GPClust found to generate stable clustering across different thresholds and runs. Clustering was performed 100 times for a specified set of parameters, with the best clustering taken as the one with the highest variance of information compared to the other clusterings, i.e. the most representative.

Each cluster produced by GPClust is associated with a defined mean and covariance function, which can be discretized to a multivariate normal distribution. In order to generate superclusters, the mean function was hierarchically clustered using correlation distance, and stability was assessed using pvclust (9) with a 10,000 iterations. Groups of clusters found to be stable (approximately unbiased p-value > 0.95) were identified and defined as superclusters. Similar results were obtained using symmetric Kullback Liebler-divergence (also called relative entropy) as a distance metric. TI_{50} (time to 50% inhibition) for each cluster was calculated using the estimated mean function, by calculating the time taken to reach 50% of the initial estimated starting value.

Functional enrichment analysis

Gene Ontology (GO) enrichments performed using GOStats (10) using all genes differentially expressed (FDR < 10%) as background. Terms with an FDR < 5% were defined as significantly enriched. Results from GO enrichments were simplified for presentation purposes by filtering terms with a high semantic similarity (Yu et al., 2010).

Transcription factor binding sites (TFBS) analysis

TFBS motifs were obtained from the JASPAR2016 database (11). Due to the impact of background sequence selection on the downstream results of motif identification (12, 13), promoters (defined as +/- 500bp from the Ensembl annotated transcription start site (TSS)) were classified into different sets based on their GC content. Each set of promoters were then scanned using FIMO (14) with a p-value threshold of $5x10^{-4}$

against the appropriate background. Enrichment of motifs in the promoter of genes in a specific cluster was calculated using a hypergeometric test and p-values were corrected using FDR after removing those motifs that were not present in the cluster of interest.

Comparison with transcriptional responses upon TCF7L2 KD

Microarray data was obtained from GEO:GSE48367 (15) and processed using LIMMA (Linear Models for Microarray and RNA-Seq Data) (Ritchie et al., 2015). Normalized expression values were $\log_2$-transformed. Microarray probes that did not map to an annotated gene in Ensembl version 79 were removed from the analysis, and in the case of multiple probes mapping to one gene, the probe with the largest interquartile range was retained. Genes were defined as significantly downregulated using an FDR threshold of 10% and a fold-change of 0.5. Significance of the overlap between genes downregulated in SH-SY5Y following TCFKD and *Wnt-activated* genes was calculated using a hypergeometric with the number of genes expressed in both the HPAF-II time course and in the microarray experiment as background. Bonferroni correction was used to correct for multiple testing and control false positives.

Analysis of existing TCF7L2 ChIP-seq datasets and TCGA-data

TCF7L2 peaks for various cell lines were obtained from ReMap 2018 (16-20). The distance from a TCF7L2 peaks were assigned to their nearest gene promoter, defined as the Ensembl defined TSS, were calculated. Tables of FPKMs for colon adenocarcinomas (COAD) and pancreatic ductal adenocarcinomas (PDAC) sequenced as part of the TCGA project (21-23) were downloaded and converted to TPMs. The box and whisker plot shows 5-95%ile (vertical line), 25-75%ile (box) and median (horizontal line).

Proteomics

Tumors were homogenized on dry ice and solubilized with 8M urea and 20 mM HEPES. After reduction and alkylation, proteins were digested with trypsin overnight. Peptides in 0.1% TFA were separated using an Ultimate 3000 RSLC nano liquid chromatography system coupled to a Q-Exactive mass spectrometer (both Thermo Scientific) via an EASY-Spray source. Technical duplicates of 500 ng were loaded onto a trap column (Acclaim PepMap 100 C18, 100 µm, 2 cm) at 8 µl/min in 2 % acetonitrile, 0.1 % TFA. Peptides were eluted on-line to an analytical column (PepMap C18, 75 µm × 25 cm) and separated using a ramped 120 minute gradient, 4-45 % of

buffer B (80 % acetonitrile, 0.1 % formic acid). Following data-dependent acquisition, raw data files were loaded into Progenesis QI (PQI v3, Non-linear Dynamics), which performed peak detection, chromatographic run alignment, combined peak list creation (Mascot mgf file) and protein quantitation. Search results (filtered to approximately 1% FDR) were imported back into Progenesis QI and relative protein quantitation was performed using non-conflicting peptides only. Protein normalization was performed by Progenesis QI (24).

Supplemental References

1. Trerè D. AgNOR staining and quantification. *Micron* 2000;31(2):127–131.

2. Conway T et al. Xenome--a tool for classifying reads from xenograft samples. *Bioinformatics* 2012;28(12):i172–8.

3. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29(1):15–21.

4. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011;12(1):323.

5. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15(12):550.

6. Madan B et al. Wnt addiction of genetically defined cancers reversed by PORCN inhibition. *Oncogene* 2016;35(17):2197–2207.

7. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 2003;100(16):9440–9445.

8. Hensman J, Rattray M, Lawrence ND. Fast Nonparametric Clustering of Structured Time-Series. *IEEE Trans Pattern Anal Mach Intell* 2015;37(2):383–393.

9. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 2006;22(12):1540–1542.

10. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics* 2007;23(2):257–258.

11. Mathelier A et al. JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2016;44(D1):D110–5.

12. Roider HG, Lenhard B, Kanhere A, Haas SA, Vingron M. CpG-depleted promoters harbor tissue-specific transcription factor binding signals--implications for motif overrepresentation analyses. *Nucleic Acids Res* 2009;37(19):6305–6315.

13. Worsley-Hunt R, Mathelier A, Del Peso L, Wasserman WW. Improving analysis of transcription factor binding sites within ChIP-Seq data based on topological motif enrichment. *BMC Genomics* 2014;15(1):472.

14. Bailey TL et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 2009;37(Web Server issue):W202–8.

15. Forrest MP, Waite AJ, Martin-Rendon E, Blake DJ. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS ONE* 2013;8(8):e73169.

16. Hazelett DJ et al. Comprehensive functional annotation of 77 prostate cancer risk loci. *PLoS Genet* 2014;10(1):e1004102.

17. Ni M et al. Amplitude modulation of androgen signaling by c-MYC. *Genes Dev* 2013;27(7):734–748.

18. Frietze S et al. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome Biol* 2012;13(9):R52.

19. Trompouki E et al. Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell* 2011;147(3):577–589.

20. Chèneby J, Gheorghe M, Artufel M, Mathelier A, Ballester B. ReMap 2018: an updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. *Nucleic Acids Res* 2018;46(D1):D267–D275.

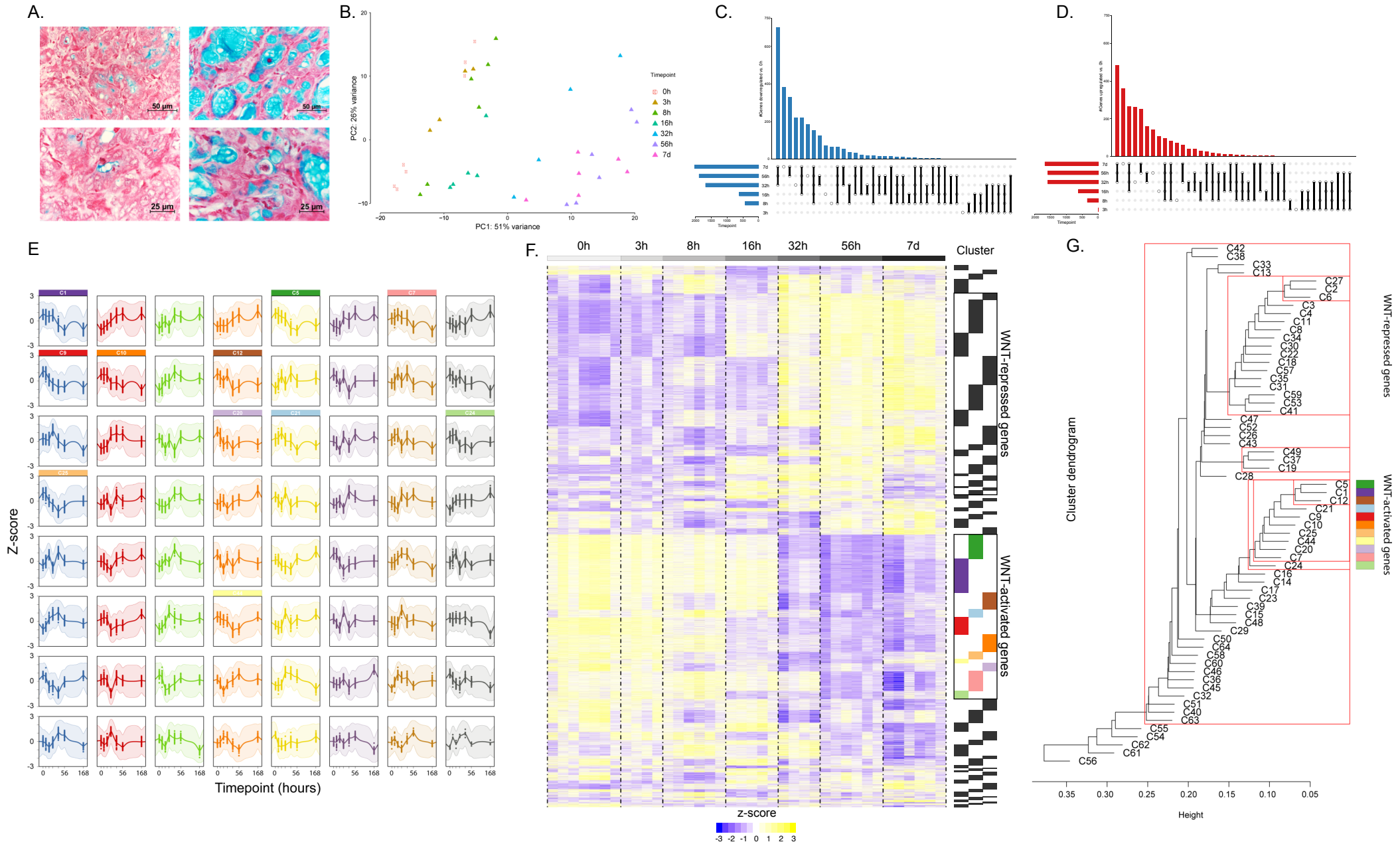
21. Muzny DM et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487(7407):330–337.

22. Cancer Genome Atlas Research Network. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 2017;32(2):185–203.e13.

23. Grossman RL et al. Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med* 2016;375(12):1109–1112.

24. Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43(7):e47–e47.

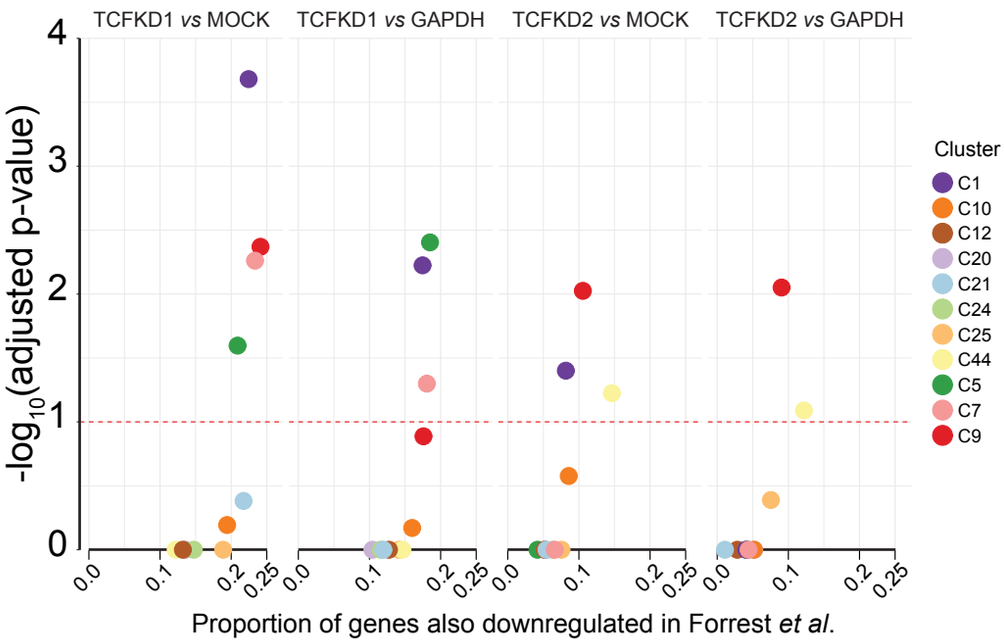
Supplemental Figure 1



Supplemental Figure 1: Time series transcriptomic analysis identifies clusters of genes with distinct dynamics in response to PORCN inhibition (accompanying Figure 1 and 2)

A. Alcian blue staining reveals acinar lumens with abundant mucus, highlighting the differentiated state of HPAF-II tumors following treatment with ETC-159 for 7 days.
B. Principal components analysis of HPAF-II orthotopic time course RNA-seq reveals coherent gene expression changes in response to ETC-159 over time.
C. Upset plot of downregulated genes reveals groups of genes that are downregulated across different sets of timepoints compared to 0 h (FDR < 10%, fold change < 0.5)
D. Upset plot of upregulated genes reveals groups of genes that are upregulated across different sets of timepoints compared to 0 h in response to PORCN inhibition (FDR < 10%, fold change > 1.5)
E. Mean and covariance of all clusters generated using GPclust. Clusters of genes defined as *Wnt-activated* are highlighted. Clusters are ordered by their size.
F. Heatmap displaying all genes that are differentially expressed over time (FDR < 10%), grouped into 64 distinct clusters based on their dynamics in response to PORCN inhibition.
G. Hierarchical clustering of the mean function of each cluster. Distinct stable groupings (i.e. super-clusters) were identified using pvclust with 1,000 iterations.

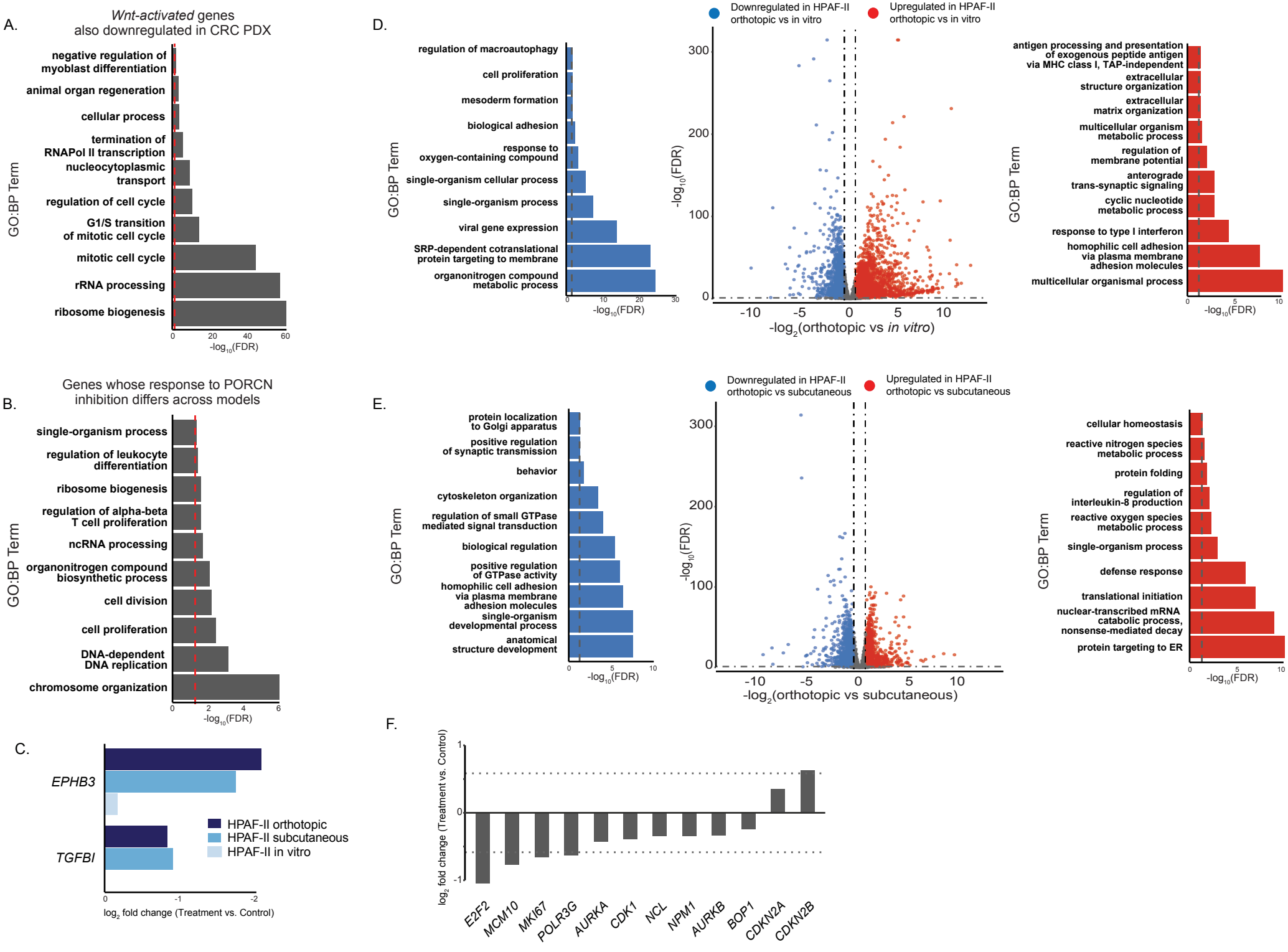
Supplemental Figure 2



Supplemental Figure 2: Clusters of WNT-activated genes are enriched for WNT targets identified following TCF KD in SH-SY5Y cells (accompanying Figure 2).

The Forrest et al. dataset (Forrest et al., 2013) was reanalyzed (Methods) and the TCF-dependent genes identified. Examining the set of genes expressed in both datasets, we determined how many of our *Wnt-activated* genes were also downregulated in Forrest et al. (FDR < 10%, fold change < 0.5). The x-axis shows the proportion of genes in each Wnt-activated cluster that were also found to be downregulated following TCF knockdown. In no case did the cell line data contain more than 25% of the genes identified in our orthotopic model.

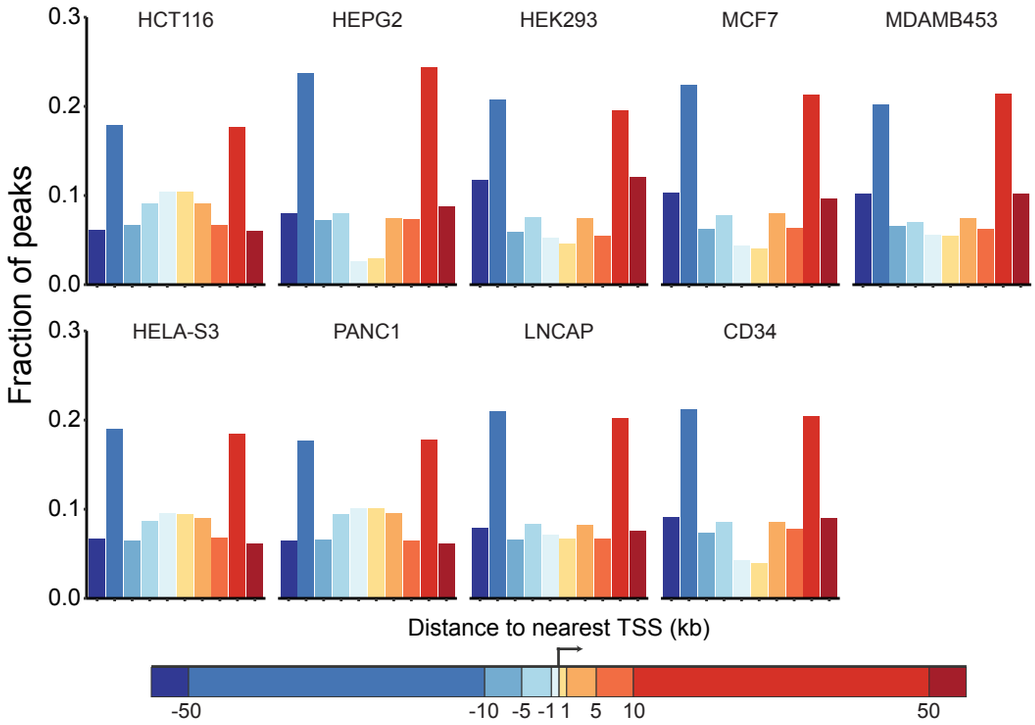
Supplemental Figure 3



Supplemental Figure 3: Comparison of models of Wnt-addicted cancer reveals genes and processes that differ in baseline expression and/or response to ETC-159 (accompanying Figure 3).

A. *Wnt-activated* genes that are downregulated (FDR < 10%) in both orthotopic HPAF-II and CRC PDX models are enriched for cell cycle and ribosome biogenesis genes as assessed by GO:BP enrichment (FDR < 10%, dotted red line).
B. Genes whose response to PORCN inhibition differs across models (interaction test, FDR < 10%) are enriched for cell cycle processes as assessed by GO:BP enrichment (FDR < 5%, dotted red line).
C. *TGBI* and *EPHB3* do not respond in vitro but are WNT targets in both HPAF-II in vivo models.
D. Volcano plot and GO enrichments show that the genes expressed even before treatment differ widely between orthotopic and in vitro models (FDR < 10%). These genes are associated with distinct biological processes as assessed by GO:BP enrichment (FDR < 5%, dotted line).
E. Volcano plot and GO enrichments show that the genes expressed before treatment also differ between orthotopic and subcutaneous models (FDR < 10%). These genes are associated with distinct biological processes as assessed by GO:BP enrichment (FDR < 5%, dotted line).
F. The expression of important representative cell cycle and ribosome biogenesis genes was reduced in AsPC-1 orthotopic tumours following PORCN inhibition (56 h).

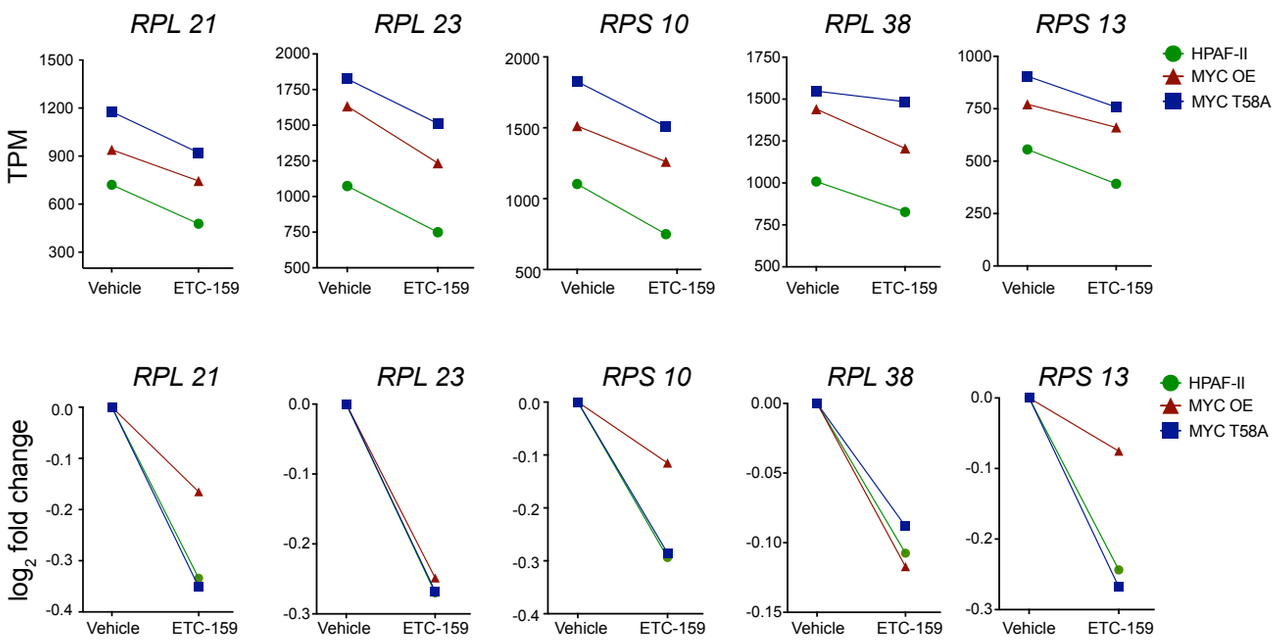
Supplemental Figure 4



Supplemental Figure 4: The majority of TCF7L2 peaks are not located near promoters (accompanying Figure 4).

Analysis of existing TCF7L2 ChIP-seq data reveals that the majority of TCF7L2 binding peaks are located away (<5kb) from the promoters of genes.

Supplementary Figure 5



Supplemental Figure 5: The response of RPLs and RPSs to PORCN inhibition is Myc-independent (accompanying Figure 9).

Expression of representative RPL and RPS genes is enhanced by MYC and is dependent on Wnt signaling.

Supplementary Tables

Table S1: *Differential expression analysis for the HPAF-II orthotopic time-course.*

Table S2*: Time-series clustering of HPAF-II orthotopic time-course.*

Table S3*: Reanalysis of CRC PDX RNA-seq from Madan et al. 2016.*

Table S4*: Results of differential expression for response to ETC-159 both in vitro and subcutaneous models and comparing between models.*

Table S5: *Results of differential expression analysis comparing the response to ETC-159 in WT, MYC OE and MYC T58A orthotopic tumors.*

Table S6: *Results of differential abundance proteomic analysis for HPAF-II orthotopic tumors.*