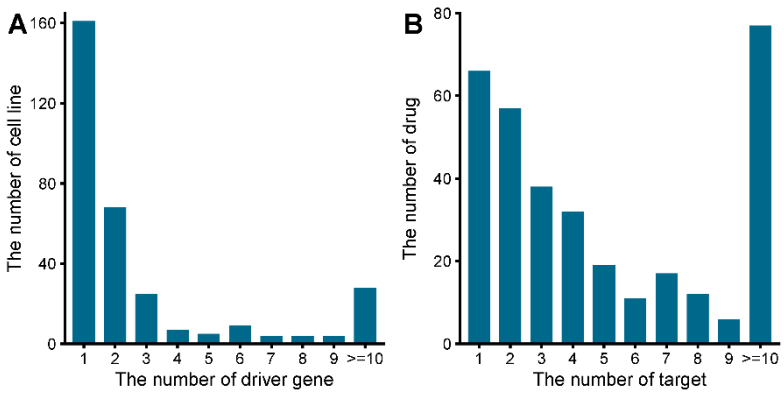
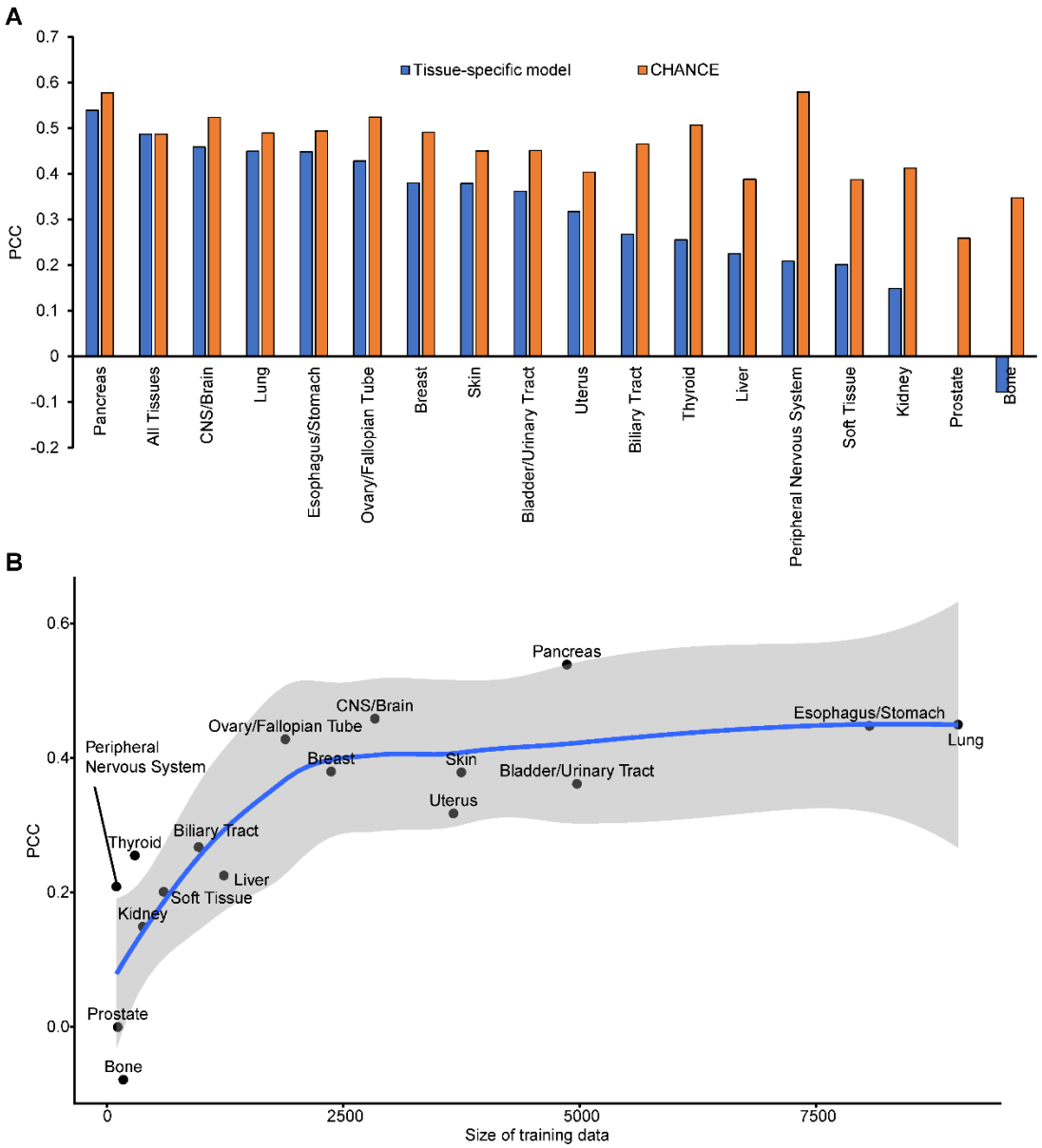
Supplementary Information for

Personalized prediction of anticancer potential of nononcology drugs through learning from genome derived molecular pathways

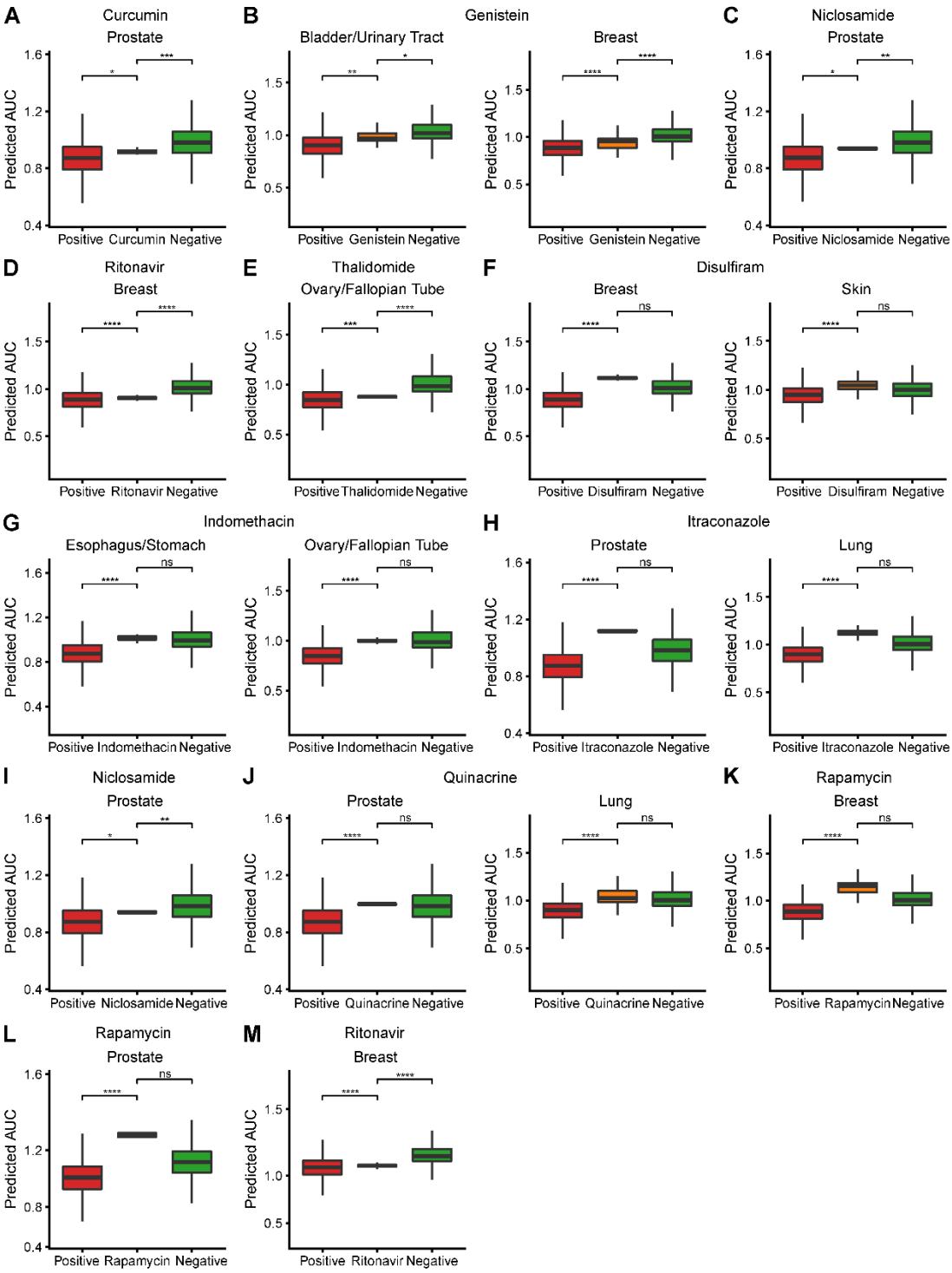
This file includes:
Supplementary Figures 1 to 7



Supplementary Figure 1. Number of driver genes and drug targets. (A) Distribution of the number of driver genes in each cell line. (B) Distribution of the number of targets for each drug.

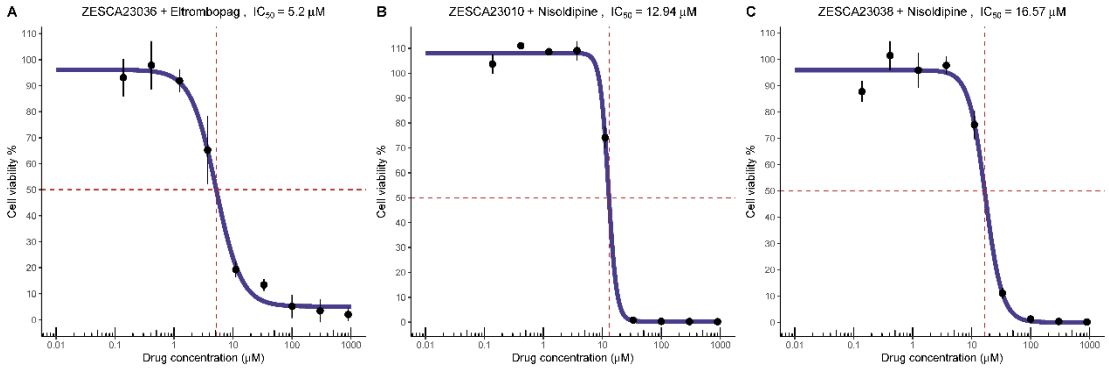


Supplementary Figure 2. Comparing the performances of CHANCE and tissue-specific model. We tested the performances of 17 tissue-specific models by training each of them with only drug-cell line pairs from the corresponding tissue type. The performance of each model was evaluated by 10-fold cross validation. For comparison, we also extracted prediction results for each tissue type in 10-fold cross validation of CHANCE. (A) The barplots show the PCCs between predicted AUC and experimental AUC of tissue-specific models (blue) and CHANCE in different tissue types (orange). (B) The performance of the tissue-specific model improves as the size of the training set increases.

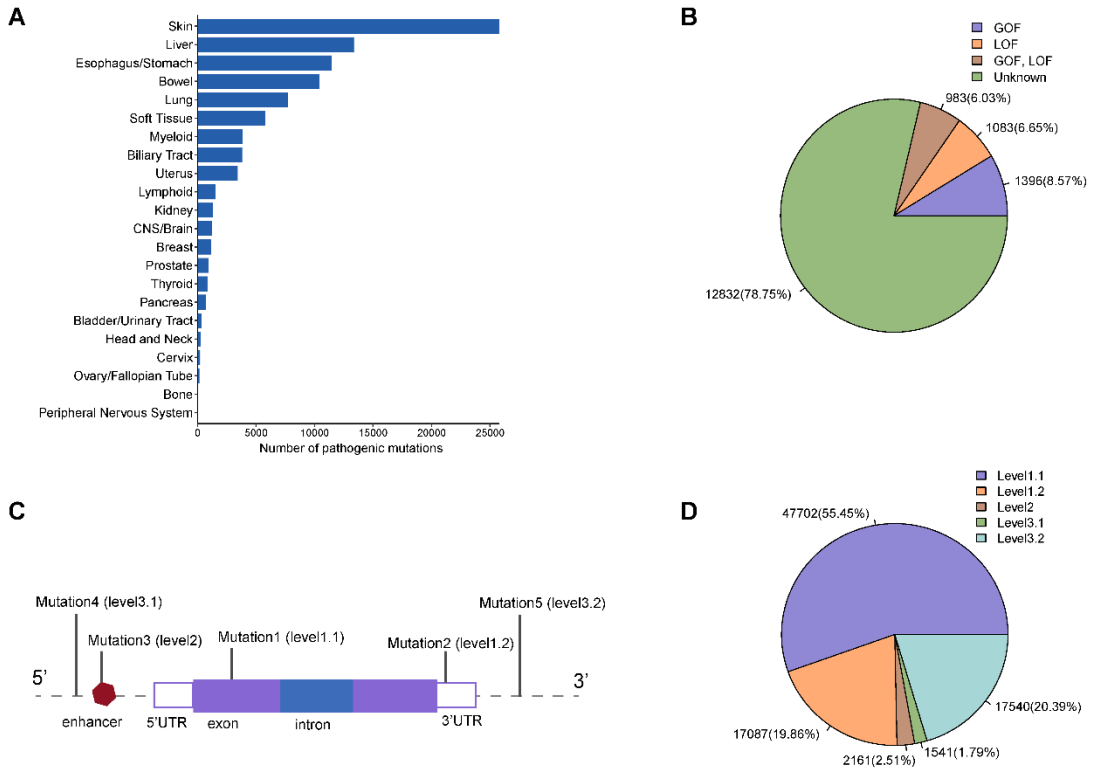


Supplementary Figure 3. Predictive performances of CHANE on all reported drug repurposing cases. Distribution of predicted AUCs of non-oncology drugs in samples from TCGA/ICGC is shown**.** If the average AUC of the reported drug had no significant difference from the average AUC of positive control drugs or even smaller than the average AUC of positive control drugs (one-sided Wilcoxon test), we classified it as a

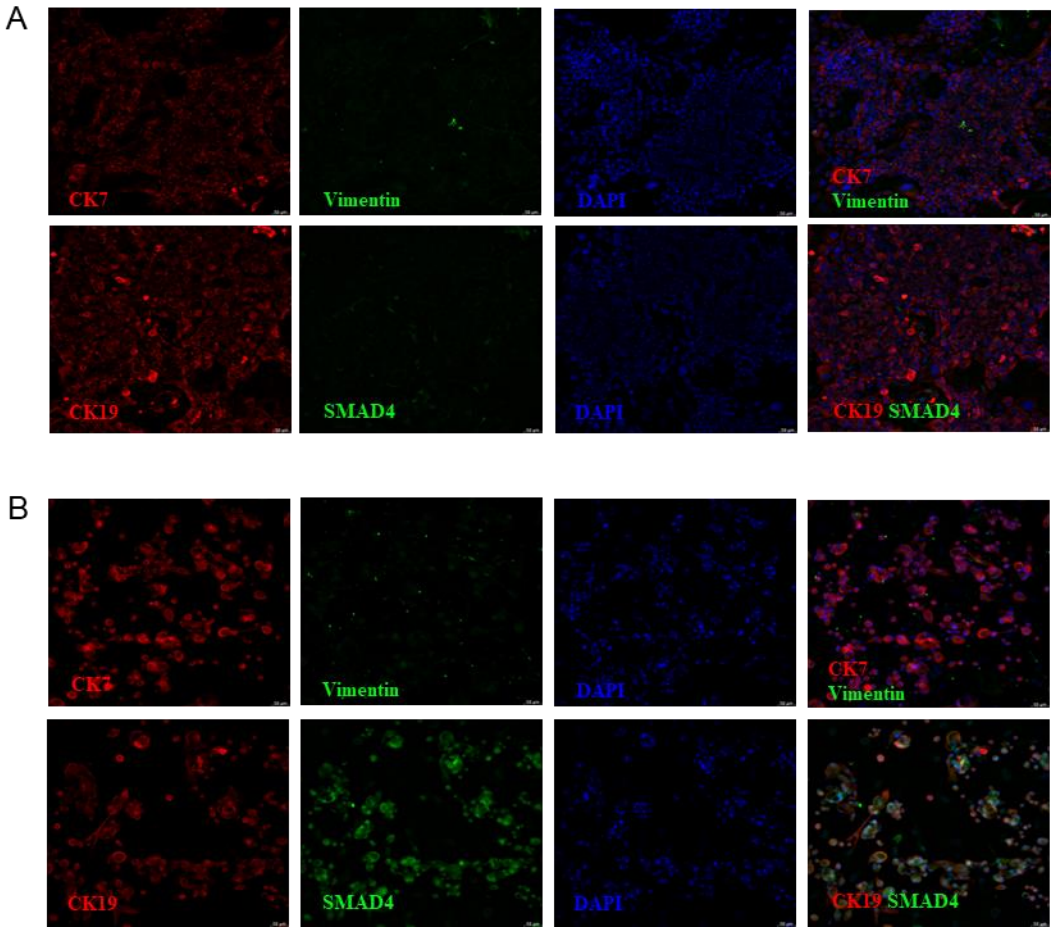
Matched case. If the average AUC of the reported drug fell between the values of positive and negative control drugs (i.e. $AUC_{positive\_control} < AUC_{reported\_drug} < AUC_{negative\_control}$), we classified this case as *Partially Matched*. And if the average AUC of the reported drug did not show a significantly smaller value when compared with negative control drugs, we classified this case as *Mismatched*.



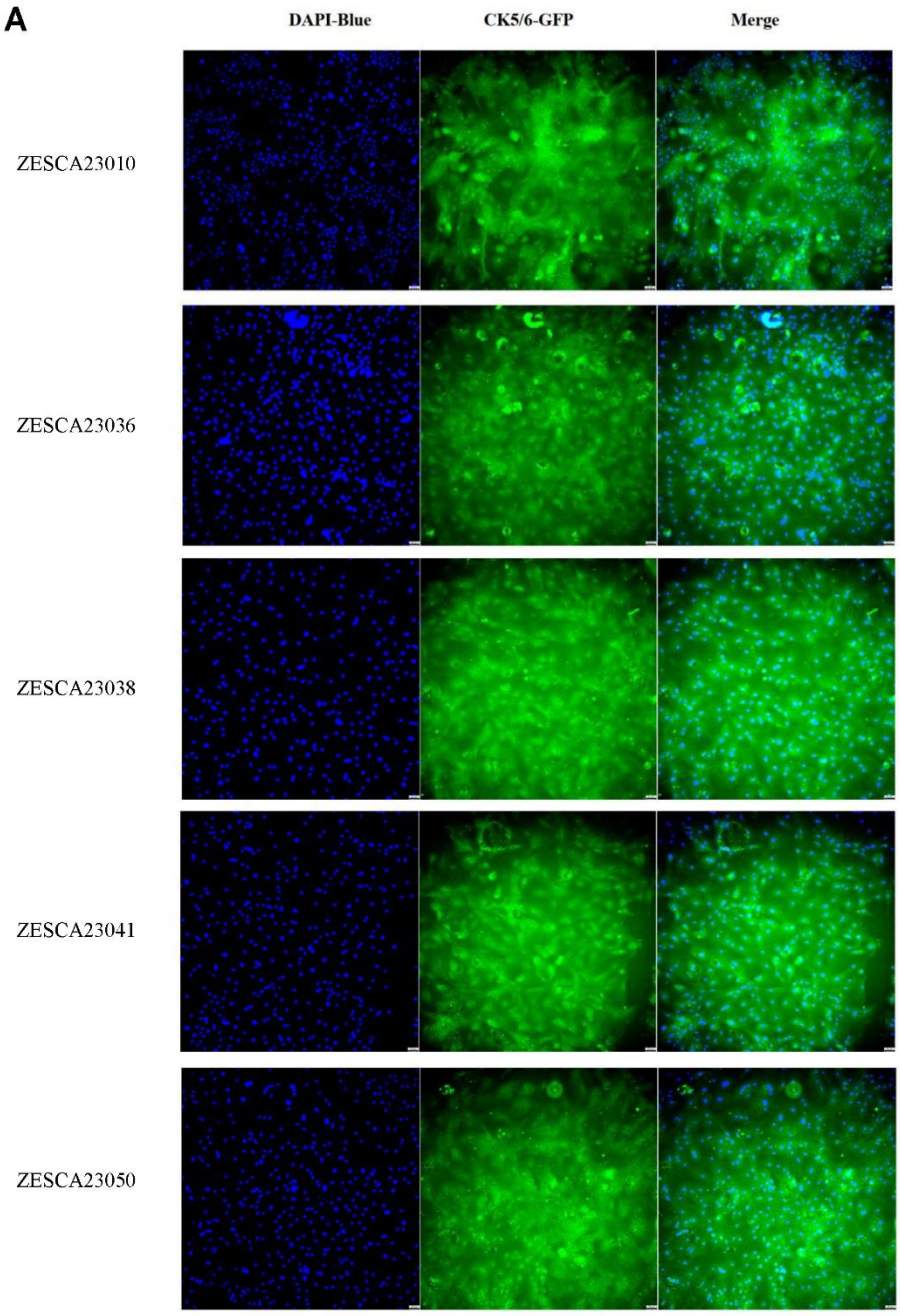
Supplementary Figure 4. **Drug response curves of validated drugs treated on samples of esophageal cancer patients.** The drug name and ID of the corresponding patient are shown at the top of each subplot. The vertical dotted lines mark the points corresponding to the IC_{50} values. Nine different concentrations were used for each drug. The experiments were repeated three times under each concentration. Error bars show the standard deviation of three replicates.



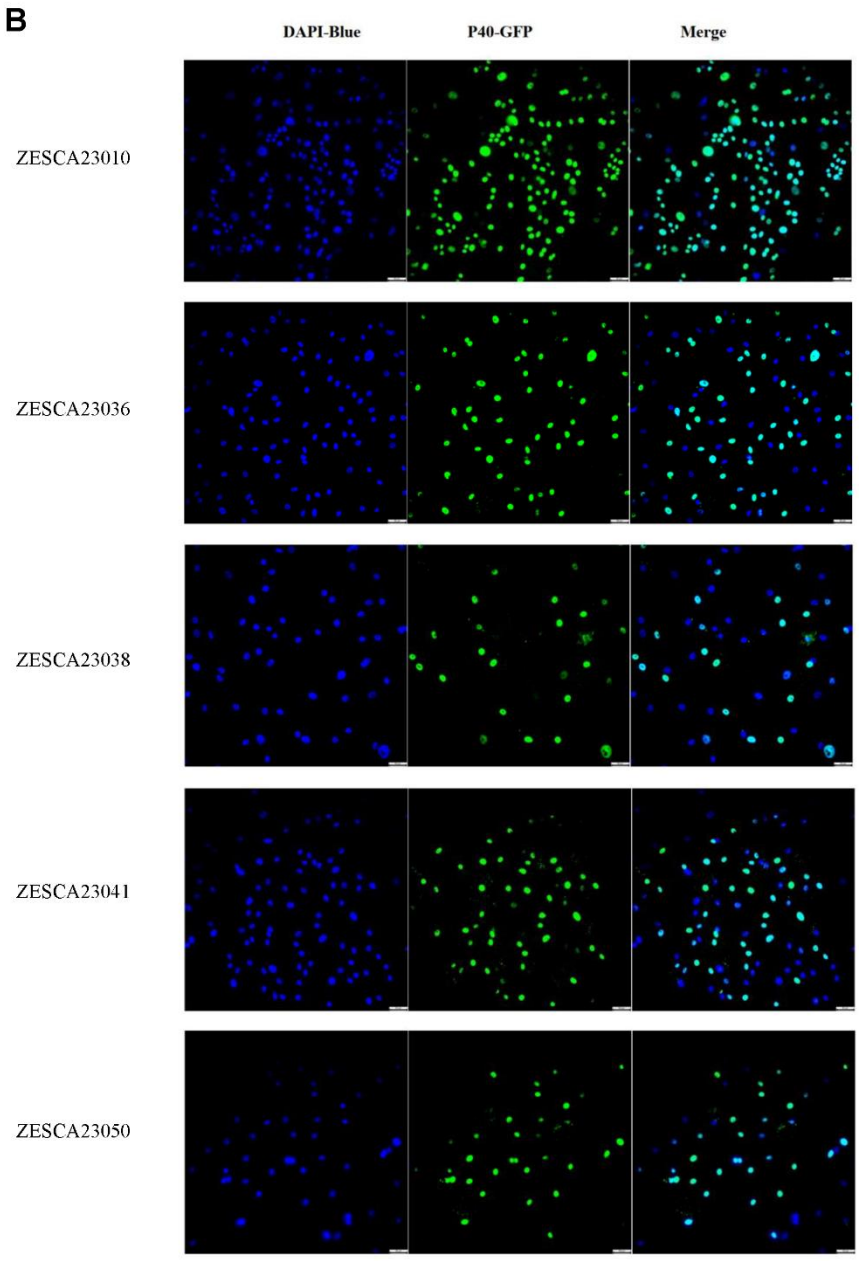
Supplementary Figure 5. Summary of compiled pathogenic mutations from ICGC. (A) The distribution of compiled pathogenic mutations in cancers with different tissue of origin. (B) The proportion of pathogenic mutations with different functional effects. Note that, for some pathogenic mutations that could not be associated with any known cancer genes, we marked their effects as unknown. (C) The association between a pathogenic mutation and a gene was classified into different levels for reflecting its strength. (D) The distribution of pathogenic mutations in different levels.



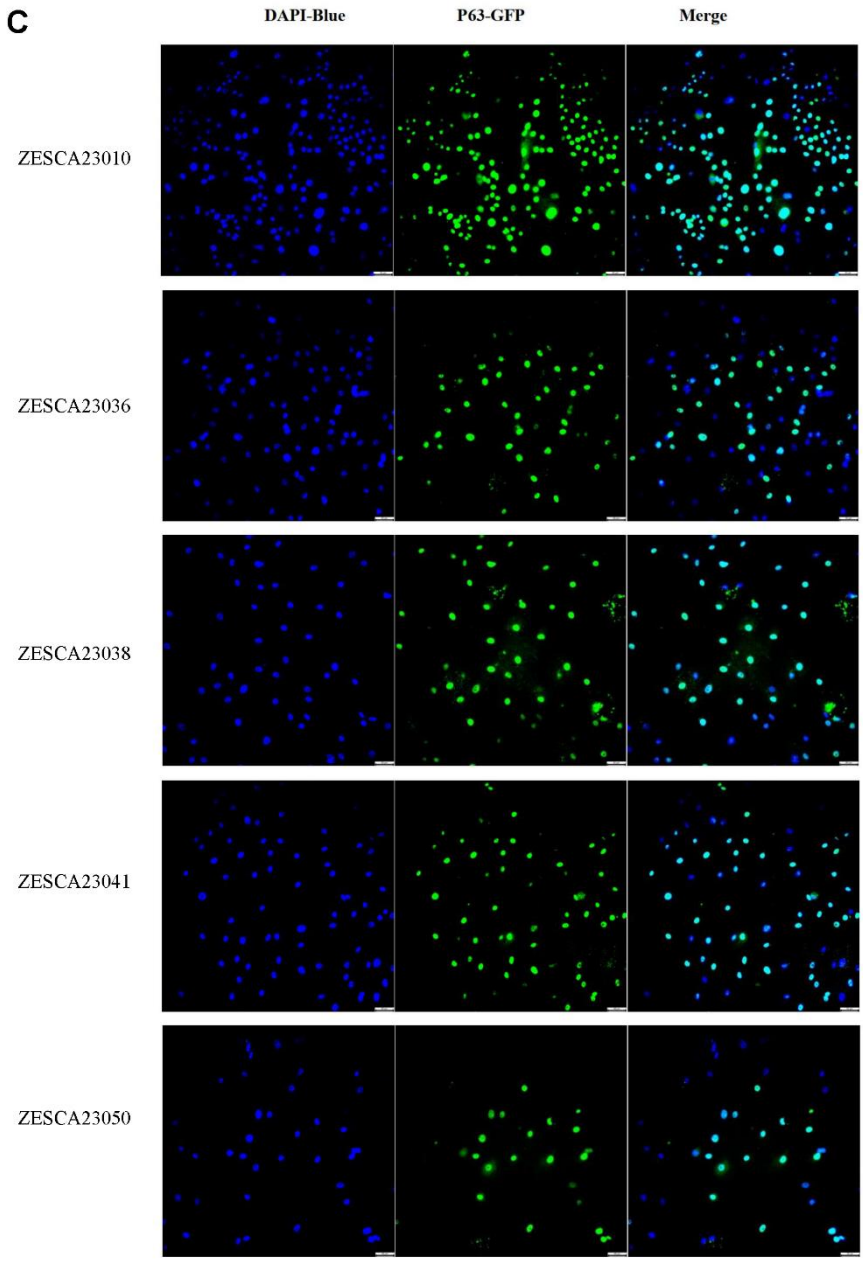
Supplementary Figure 6. **Immunofluorescence assay of patient-derived tumor cells (PDCs) from patients with pancreatic cancer.** CK7 and CK19 are markers of many normal epithelia and epithelial tumors and ubiquitously expressed in pancreatic ductal adenocarcinoma. Vimentin is a marker of mesenchymal cells. Immunofluorescence assay showed that the PDCs from the two advanced pancreatic cancer patients P052 (A) and P080 (B) exhibited CK7+/CK19+/vimentin- expression patterns, indicating most cells are from pancreatic epithelial cells and there is no contamination from mesenchymal cells such as fibroblasts. Furthermore, we stained another tumor suppressor protein SMAD4, which is lost in 60%-90% of PDAC. We confirmed its negative expression in PDCs from patient P052, corroborating the whole genome sequencing result, in which a nonsense mutation SMAD4:p.R445X was detected.



Supplementary Figure 7. Immunofluorescence assay of patient-derived tumor cells (PDCs) from patients with esophageal cancer. CK5/6, p40 and p63 are usually expressed in tumor cells of esophageal squamous cell carcinoma. Immunofluorescence assay showed that these markers were positive in most PDCs from the five patients with esophageal cancer. (A): CK5/6, (B): p40, (C): p63.



Supplementary Figure 7. Immunofluorescence assay of patient-derived tumor cells (PDCs) from patients with esophageal cancer. (continued)



Supplementary Figure 7. Immunofluorescence assay of patient-derived tumor cells (PDCs) from patients with esophageal cancer. (continued)