

Supplement Table 1. Characteristics of GEO datasets used in this study.

Dataset ID	Disease	Sample Size (Case/Control)	Tissue Source	Platform (GPL)	Clinical Information	Normalization Method	Batch Effect
GSE 15471	PDAC	36 / 36	Pancreatic tissue (tumor vs adjacent normal)	GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array)	Limited (tumor vs normal status available)	RMA + log2 transformation	Not required
GSE 20966	T2DM	10 / 10	Human pancreatic islets	GPL1352 (Affymetrix Human X3P Array)	Limited (diabetes status available)	RMA + log2 transformation	Not required