

	TCGA	PCAWG	ONCO-Cancer	TCGA $\cap$ PCAWG	TCGA $\cap$ ONCO-Cancer	PCAWG $\cap$ ONCO-Cancer	TCGA $\cap$ PCAWG $\cap$ ONCO-Cancer
GTEX (18,979 isoforms)	42	6	121	0	32	0	0
ONCOBOX ANTE (79,111 isoforms)	139	123	0	6	0	0	0
GTEX $\cap$ ONCOBOX (10,261 isoforms)	0	4	0	0	0	0	0

Supplementary Table S1 Pancancer Transcript Isoform Counts in Cancer Datasets Contrasted with Normal Datasets. The table provides a detailed comparison of transcript isoform counts between three major cancer datasets (TCGA, PCAWG, ONCO-Cancer) and their intersections while contrasting normal tissue datasets (GTEX, ONCOBOX ANTE) and their intersection. The inclusion criteria for isoforms are: 1.) at least 90% of cancer tissue samples must have TPM value greater than one, and 2.) 100% of normal tissue samples must have a TPM value less than one to allow for heterogeneity in the cancer samples.

	TCGA-BRCA	PCAWG Breast	ONCO-Cancer Breast	TCGA-LUAD	TCGA-LUSC	PCAWG Lung	ONCO-Cancer Lung
GTEX (18,979 isoforms)	81	17	125	81	94	19	380
ONCOBOX ANTE (79,111 isoforms)	432	239	0	335	507	209	0
GTEX $\cap$ ONCOBOX (10,261 isoforms)	2	6	0	0	2	5	0

Supplementary Table S2 Pancancer Transcript Isoform Counts in Breast and Lung Cancer Datasets Contrasted with Normal Datasets. The table provides a detailed comparison of breast- and lung-specific transcript isoform counts between three major cancer datasets (TCGA, PCAWG, ONCO-Cancer) while contrasting the normal tissue datasets (GTEX, ONCOBOX ANTE) and their intersection. The inclusion criteria for isoforms are: 1.) at least 90% of cancer tissue samples must have TPM value greater than one, and 2.) 100% of normal tissue samples must have a TPM value less than one to allow for heterogeneity in the cancer samples.

	Breast Cancer: TCGA-BRCA $\cap$ PCAWG-breast $\cap$ ONCO-breast	Lung Cancer: TCGA-LUAD $\cap$ TCGA-LUSC $\cap$ PCAWG-lung $\cap$ ONCO-lung	Breast Cancer: TCGA-BRCA $\cap$ PCAWG-breast	Lung Cancer: TCGA-LUAD $\cap$ TCGA-LUSC $\cap$ PCAWG-lung
GTEX (18,979 isoforms)	3	6	5	9
ONCOBOX ANTE (79,111 isoforms)	0	0	16	11
GTEX $\cap$ ONCOBOX (10,261 isoforms)	0	0	0	0

Supplementary Table S3 Pancancer Transcript Isoform Counts in Breast and Lung Cancer Datasets Contrasted with Normal Datasets. The table provides a detailed comparison of breast- and lung-specific transcript isoform counts between the intersections of three major cancer datasets (TCGA, PCAWG, ONCO-Cancer) while contrasting the normal tissue datasets (GTEX, ONCOBOX ANTE) and their intersection. The inclusion criteria for isoforms are: 1.) at least 90% of cancer tissue samples must have TPM value greater than one, and 2.) 100% of normal tissue samples must have a TPM value less than one to allow for heterogeneity in the cancer samples.