Supplementary Methods and Results

Reproducible transcriptomic and sensitivity analyses of intraductal papillary mucinous neoplasms. Original GSVA sample scores and fgsea pathway results were recovered from the author-provided R workspace; secondary specimen-level analyses are reported separately.

Supplementary Methods

Original GSE210351 sample scores were retrieved from the saved GSVA output for 41 specimens. The R command history confirms use of ssgseaParam for the three core Hallmark sets, followed by z-score standardization of each pathway across all 41 specimens and equal-weight aggregation. The restricted pathological analysis comprised 24 IPMN specimens. Grade comparisons used two-sided exact Wilcoxon tests. Linear models included the original grade-plus-subtype model, a grade-by-subtype interaction term, and a grade-plus-subtype model adjusted for a six-marker epithelial expression proxy.

For epithelial-marker sensitivity analysis, EPCAM, KRT7, KRT8, KRT18, KRT19, and CDH1 were individually standardized across the 24 IPMN specimens, then averaged. This expression-derived proxy was not interpreted as an estimated percentage of tumour cells. AUCs were apparent estimates within the same 24 cases and do not represent validated clinical prediction.

GSE26647 contains 28 processed, laser-capture microdissected pathological foci from 14 patients. Its GPL5175 transcript-cluster expression matrix was mapped to unambiguous gene symbols, with a single transcript cluster per gene selected by expression interquartile range. Weighted rank-based sample enrichment scores with exponent 0.25 were standardized within the 28 foci and averaged. Patient-clustered grade regression used a t distribution with 13 degrees of freedom. The noninvasive analysis excluded seven invasive carcinoma foci and retained 21 foci from 14 patients. This implementation was a separate sensitivity analysis; GSE210351 original GSVA scores were never replaced by these proxy scores.

The original GSE63104 and GSE26647 fgsea results are presented in full as 50 gene-set tables, with the original gene-set membership version MSigDB Hallmark 7.5.1. The original leading-edge genes were extracted directly from the saved fgsea result objects.

Supplementary Table S1. Cohorts and pathological group definitions

GEO cohort	Platform	Groups	Analysis n	Analytical role
GSE63104	GPL15048	6 low-risk; 17 high-risk	23 specimens	Discovery fgsea
GSE26647	GPL5175	10 low; 5 moderate; 6 high; 7 invasive	28 foci from 14 patients	Replication and clustered sensitivity
GSE210351	GPL24676	6 gastric low; 6 gastric high; 6 intestinal low; 6 intestinal high	24 IPMN of 41 profiled	Original GSVA grade scoring

Supplementary Table S2. All 50 Hallmark pathways in GSE63104

Hallmark pathway	Genes	ES	NES	P	FDR
HALLMARK_E2F_TARGETS	193	0.610	2.902	9.9216e-27	4.9608e-25
HALLMARK_G2M_CHECKPOINT	189	0.587	2.782	1.4462e-23	3.6155e-22
HALLMARK_TNFA_SIGNALING_VIA_NFKB	196	-0.475	-2.398	4.1906e-13	6.9844e-12
HALLMARK_MITOTIC_SPINDLE	197	0.407	1.933	1.2747e-07	1.5933e-06
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	194	0.391	1.862	1.5806e-06	1.5806e-05
HALLMARK_MYC_TARGETS_V1	191	0.375	1.783	1.3272e-05	0.0001106
HALLMARK_ESTROGEN_RESPONSE_EARLY	192	-0.334	-1.676	0.00013622	0.00097302
HALLMARK_MTORC1_SIGNALING	194	0.343	1.629	0.00034902	0.0021814
HALLMARK_GLYCOLYSIS	196	0.333	1.580	0.0006987	0.0038817
HALLMARK_ESTROGEN_RESPONSE_LATE	191	-0.301	-1.510	0.0016816	0.0084082
HALLMARK_INTERFERON_ALPHA_RESPONSE	92	0.383	1.628	0.0029136	0.013244

HALLMARK_SPERMATOGENESIS	132	0.342	1.542	0.0053771	0.022405
HALLMARK_KRAS_SIGNALING_DN	186	-0.290	-1.450	0.0080063	0.030793
HALLMARK_APOPTOSIS	157	-0.287	-1.399	0.010393	0.037119
HALLMARK_PROTEIN_SECRETION	95	0.332	1.425	0.023841	0.074504
HALLMARK_MYOGENESIS	198	-0.263	-1.328	0.023504	0.074504
HALLMARK_P53_PATHWAY	188	-0.260	-1.300	0.03953	0.11627
HALLMARK_APICAL_JUNCTION	192	0.269	1.282	0.042386	0.11774
HALLMARK_PI3K_AKT_MTOR_SIGNALING	104	0.305	1.329	0.05137	0.13518
HALLMARK_DNA_REPAIR	136	0.279	1.264	0.05802	0.14505
HALLMARK_FATTY_ACID_METABOLISM	152	-0.257	-1.247	0.064267	0.14606
HALLMARK_PEROXISOME	103	-0.289	-1.320	0.062099	0.14606
HALLMARK_UV_RESPONSE_UP	151	-0.256	-1.235	0.067302	0.14631
HALLMARK_BILE_ACID_METABOLISM	112	-0.270	-1.249	0.085784	0.17872
HALLMARK_OXIDATIVE_PHOSPHORYLATION	182	0.251	1.189	0.12579	0.25157
HALLMARK_IL6_JAK_STAT3_SIGNALING	87	0.283	1.190	0.1436	0.25643
HALLMARK_COMPLEMENT	195	0.245	1.163	0.14082	0.25643
HALLMARK_INTERFERON_GAMMA_RESPONSE	195	0.245	1.166	0.13608	0.25643
HALLMARK_ALLOGRAFT_REJECTION	193	0.243	1.155	0.16535	0.28509
HALLMARK_ADIPOGENESIS	186	-0.222	-1.113	0.19894	0.33156
HALLMARK_UV_RESPONSE_DN	138	-0.228	-1.091	0.24939	0.40225
HALLMARK_COAGULATION	136	0.240	1.088	0.28328	0.43643
HALLMARK_XENOBIOTIC_METABOLISM	194	-0.210	-1.060	0.28804	0.43643
HALLMARK_HYPOXIA	190	0.224	1.061	0.29984	0.44094
HALLMARK_TGF_BETA_SIGNALING	54	-0.270	-1.073	0.31807	0.45439
HALLMARK_MYC_TARGETS_V2	58	-0.254	-1.025	0.42005	0.5834
HALLMARK_ANDROGEN_RESPONSE	96	-0.217	-0.974	0.4963	0.65302
HALLMARK_KRAS_SIGNALING_UP	192	0.208	0.989	0.49451	0.65302
HALLMARK_WNT_BETA_CATENIN_SIGNALING	42	0.261	0.948	0.5208	0.66769
HALLMARK_ANGIOGENESIS	36	-0.251	-0.911	0.59165	0.73956
HALLMARK_NOTCH_SIGNALING	32	0.259	0.894	0.62176	0.75825
HALLMARK_HEDGEHOG_SIGNALING	35	-0.241	-0.866	0.70901	0.84406
HALLMARK_APICAL_SURFACE	43	0.224	0.817	0.77899	0.87925
HALLMARK_CHOLESTEROL_HOMEOSTASIS	71	-0.203	-0.855	0.76683	0.87925
HALLMARK_IL2_STAT5_SIGNALING	194	0.181	0.862	0.8265	0.87925
HALLMARK_INFLAMMATORY_RESPONSE	197	-0.173	-0.874	0.81183	0.87925
HALLMARK_PANCREAS_BETA_CELLS	40	-0.219	-0.813	0.79724	0.87925
HALLMARK_HEME_METABOLISM	190	-0.156	-0.785	0.96	0.97959
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	106	-0.170	-0.780	0.94581	0.97959
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	45	0.153	0.567	0.99462	0.99462

The estimates above are the saved original fgsea results. Positive NES indicates enrichment in the high-risk pathological group. Raw P values and FDR values are distinct from individual-gene differential expression tests.

Supplementary Table S3. All 50 Hallmark pathways in GSE26647

Hallmark pathway	Genes	ES	NES	P	FDR
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	171	0.641	3.538	8.0258e-35	4.0129e-33
HALLMARK_MYC_TARGETS_V1	145	0.641	3.482	1.2452e-29	3.1131e-28
HALLMARK_MTORC1_SIGNALING	170	0.594	3.285	1.5575e-28	2.5959e-27
HALLMARK_E2F_TARGETS	158	0.607	3.309	1.2321e-27	1.5401e-26
HALLMARK_G2M_CHECKPOINT	162	0.537	2.939	4.3948e-20	4.3948e-19
HALLMARK_TNFA_SIGNALING_VIA_NFKB	175	0.519	2.870	8.708e-20	7.2567e-19
HALLMARK_OXIDATIVE_PHOSPHORYLATION	149	0.527	2.868	1.4829e-17	1.0592e-16
HALLMARK_MITOTIC_SPINDLE	168	0.490	2.691	1.8117e-15	1.1323e-14
HALLMARK_HYPOXIA	173	0.450	2.481	1.3241e-13	7.356e-13
HALLMARK_PROTEIN_SECRETION	84	0.581	2.827	1.3005e-12	6.5023e-12
HALLMARK_ADIPOGENESIS	168	0.432	2.370	4.3646e-11	1.9839e-10
HALLMARK_APOPTOSIS	145	0.449	2.439	7.6307e-11	3.1794e-10
HALLMARK_INTERFERON_GAMMA_RESPONSE	171	0.413	2.279	4.9568e-10	1.9065e-09
HALLMARK_P53_PATHWAY	168	0.414	2.274	6.4781e-10	2.3136e-09
HALLMARK_INTERFERON_ALPHA_RESPONSE	81	0.512	2.501	4.4051e-09	1.4684e-08
HALLMARK_GLYCOLYSIS	175	0.378	2.092	1.057e-08	3.3032e-08
HALLMARK_UV_RESPONSE_DN	123	0.428	2.227	1.3204e-08	3.8836e-08
HALLMARK_APICAL_JUNCTION	174	0.382	2.107	1.6142e-08	4.4838e-08
HALLMARK_CHOLESTEROL_HOMEOSTASIS	63	0.541	2.507	2.1995e-08	5.7881e-08
HALLMARK_KRAS_SIGNALING_UP	168	0.378	2.073	9.4364e-08	2.3591e-07
HALLMARK_COMPLEMENT	166	0.360	1.971	6.9685e-07	1.6592e-06
HALLMARK_ANDROGEN_RESPONSE	83	0.445	2.172	8.65e-07	1.9659e-06
HALLMARK_ANGIOGENESIS	30	0.641	2.466	1.1837e-06	2.5733e-06

HALLMARK_KRAS_SIGNALING_DN	158	-0.391	-1.983	1.6634e-06	3.4655e-06
HALLMARK_COAGULATION	118	0.390	1.996	2.5002e-06	5.0005e-06
HALLMARK_INFLAMMATORY_RESPONSE	197	0.314	1.764	8.2762e-06	1.5916e-05
HALLMARK_TGF_BETA_SIGNALING	52	0.494	2.189	1.0255e-05	1.8991e-05
HALLMARK_PI3K_AKT_MTOR_SIGNALING	90	0.411	2.047	1.5076e-05	2.6921e-05
HALLMARK_IL2_STAT5_SIGNALING	171	0.333	1.836	1.5978e-05	2.7549e-05
HALLMARK_DNA_REPAIR	118	0.365	1.866	2.9509e-05	4.9182e-05
HALLMARK_UV_RESPONSE_UP	136	0.345	1.836	4.026e-05	6.4935e-05
HALLMARK_ESTROGEN_RESPONSE_LATE	173	0.293	1.618	0.00066579	0.0010403
HALLMARK_HEDGEHOG_SIGNALING	32	0.493	1.917	0.00071663	0.0010858
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	40	0.450	1.860	0.0013605	0.0020007
HALLMARK_ESTROGEN_RESPONSE_EARLY	174	0.280	1.544	0.0014668	0.0020954
HALLMARK_IL6_JAK_STAT3_SIGNALING	76	0.332	1.606	0.0060455	0.0083965
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	90	0.310	1.542	0.0081849	0.011061
HALLMARK_MYOGENESIS	171	0.258	1.425	0.0092251	0.012138
HALLMARK_FATTY_ACID_METABOLISM	130	0.262	1.378	0.011928	0.015292
HALLMARK_XENOBIOTIC_METABOLISM	169	0.243	1.334	0.014528	0.01816
HALLMARK_WNT_BETA_CATENIN_SIGNALING	35	0.363	1.461	0.046784	0.057053
HALLMARK_MYC_TARGETS_V2	49	0.321	1.397	0.068306	0.081317
HALLMARK_SPERMATOGENESIS	118	-0.260	-1.272	0.093023	0.10817
HALLMARK_BILE_ACID_METABOLISM	110	-0.260	-1.249	0.11033	0.12537
HALLMARK_HEME_METABOLISM	162	0.220	1.205	0.12319	0.13688
HALLMARK_APICAL_SURFACE	38	0.302	1.235	0.17663	0.19199
HALLMARK_NOTCH_SIGNALING	25	0.329	1.220	0.18797	0.19997
HALLMARK_PEROXISOME	92	0.228	1.142	0.22257	0.23184
HALLMARK_ALLOGRAFT_REJECTION	169	0.148	0.811	0.91822	0.92553
HALLMARK_PANCREAS_BETA_CELLS	36	0.168	0.681	0.92553	0.92553

The estimates above are the saved original fgsea results. Positive NES indicates enrichment in the high-risk pathological group. Raw P values and FDR values are distinct from individual-gene differential expression tests.

Supplementary Table S4. Patient-clustered GSE26647 sensitivity contrasts

Comparison	Score	High	Lower	Beta	95% CI low	95% CI high	P
All foci	E2F	13	15	1.028	0.390	1.666	0.0040532
All foci	G2M	13	15	1.014	0.355	1.674	0.0054972
All foci	Spindle	13	15	0.945	0.230	1.660	0.013543
All foci	Composite	13	15	0.996	0.391	1.600	0.0035099
Noninvasive only	E2F	6	15	1.449	0.733	2.166	0.00075768
Noninvasive only	G2M	6	15	1.479	0.655	2.303	0.0019049
Noninvasive only	Spindle	6	15	0.967	0.190	1.744	0.018574
Noninvasive only	Composite	6	15	1.298	0.601	1.996	0.0014542

Patient clusters were the 14 source-study patients. High-risk includes high-grade and invasive foci in the full comparison; the noninvasive sensitivity analysis excludes all seven invasive foci.

Supplementary Table S5. Equal, reweighted, and individual pathway comparisons in GSE210351

Score definition	E2F weight	G2M weight	Spindle weight	Grade beta	Exact rank P	Apparent AUC
equal	0.333	0.333	0.333	1.038	0.002914	0.847
E2F_half	0.500	0.250	0.250	1.049	0.0036366	0.840
G2M_half	0.250	0.500	0.250	1.102	0.0023164	0.854
Spindle_half	0.250	0.250	0.500	0.964	0.0014326	0.868
E2F_only	1.000	0.000	0.000	1.080	0.017271	0.785
G2M_only	0.000	1.000	0.000	1.294	0.002914	0.847
Spindle_only	0.000	0.000	1.000	0.740	0.0045131	0.833

The weight sensitivity analysis uses the original archived pathway z-scores standardized across all 41 expression profiles. All P values are based on the original 24 IPMN lesions. AUC is an in-sample descriptive quantity.

Supplementary Table S6. Original-score linear models in 24 IPMN lesions

Model	Term	Beta	95% CI low	95% CI high	P
additive_OLS	Intercept	-0.921	-1.380	-0.461	0.00043576
additive_OLS	gradeH	1.038	0.508	1.569	0.00055007
additive_OLS	subtypeI	0.655	0.124	1.185	0.017966
interaction_OLS	Intercept	-0.998	-1.539	-0.458	0.00099436
interaction_OLS	gradeH	1.193	0.429	1.957	0.0039571
interaction_OLS	subtypeI	0.810	0.045	1.574	0.038949
interaction_OLS	gradeH:subtypeI	-0.310	-1.391	0.771	0.55681
epithelial_OLS	Intercept	-0.906	-1.369	-0.443	0.00058095
epithelial_OLS	gradeH	1.068	0.530	1.605	0.00050174
epithelial_OLS	subtypeI	0.596	0.048	1.145	0.034652
epithelial_OLS	EpithelialProxy	0.201	-0.242	0.644	0.35492

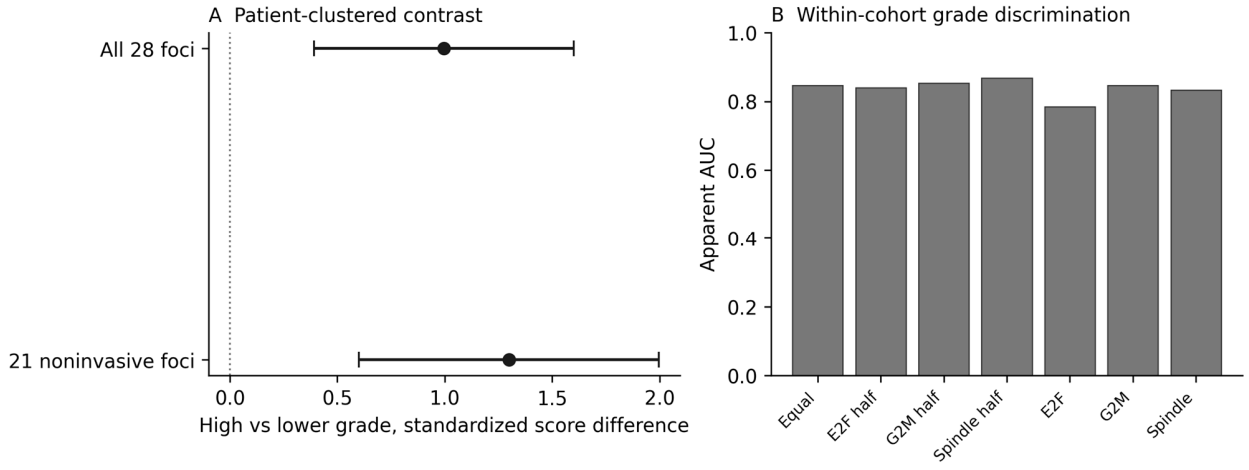
The original additive analysis uses high-grade and gastric subtype as the reference groups. The supplementary high-grade indicator model has the opposite sign for the grade coefficient but the same effect magnitude. EpithelialProxy is a mean standardized epithelial-marker expression score, not a histological cellularity percentage.

Supplementary Table S7. Original cross-cohort leading-edge gene recurrence

Genes present in the leading-edge lists of all three core Hallmark pathways in both original fgsea cohorts: AURKA, BIRC5, CCNB2, CENPE, PLK1.

Membership overlap in Hallmark 7.5.1: E2F and G2M, 73 shared genes; E2F and mitotic spindle, 19; G2M and mitotic spindle, 39. Within the 24 IPMN specimens, original score correlations were 0.932, 0.540, and 0.661, respectively, using Spearman correlation.

Figure S1. Original-score sensitivity summary



Panel A shows patient-clustered noninvasive and all-focus comparisons in GSE26647 with 95% t-based confidence intervals. Panel B shows apparent within-cohort AUC values from the original GSE210351 score for seven prespecified weighting choices.