

Supplementary Tables S1–S3

Table S1. Frozen 24-gene cross-species module: gene assignments, one-to-one orthology and descriptive log₁p(CPM) statistics

Component	Human Ensembl	Human symbol	Dog orthologue	Human mean log ₁ p(CPM)	Human SD	Dog mean log ₁ p(CPM)	Dog SD
activation_4	ENSG00000129226	CD68	CD68	2.15	1.28	5.10	1.15
activation_4	ENSG00000051523	CYBA	CYBA	4.84	1.46	4.32	0.43
activation_4	ENSG00000118785	SPP1	SPP1	3.72	2.63	7.08	1.12
activation_4	ENSG00000110719	TCIRG1	TCIRG1	1.37	1.18	4.83	0.79
antigen_presentation_3	ENSG00000019582	CD74	CD74	3.96	2.07	6.36	1.00
antigen_presentation_3	ENSG00000204287	HLA-DRA	DLA-DRA	3.78	1.88	5.24	1.06
antigen_presentation_3	ENSG00000090382	LYZ	LYZ	1.77	1.22	3.02	0.87
module_residual_17	ENSG00000102575	ACP5	ACP5	2.70	2.38	6.94	1.70
module_residual_17	ENSG00000148926	ADM	ADM	1.16	1.31	2.16	0.90
module_residual_17	ENSG00000204472	AIF1	AIF1	2.27	1.17	2.84	0.85
module_residual_17	ENSG00000130208	APOC1	APOC1	2.18	1.53	2.82	1.15
module_residual_17	ENSG00000130203	APOE	APOE	4.02	2.19	4.67	1.09
module_residual_17	ENSG00000173372	C1QA	C1QA	2.37	1.38	4.45	1.05
module_residual_17	ENSG00000173369	C1QB	C1QB	2.07	1.36	4.57	1.09
module_residual_17	ENSG00000101160	CTSZ	CTSZ	2.96	1.48	4.53	0.54
module_residual_17	ENSG00000158869	FCER1G	FCER1G	2.47	1.45	1.63	0.59
module_residual_17	ENSG000000029559	IBSP	IBSP	2.30	2.74	6.14	1.77
module_residual_17	ENSG00000162511	LAPTM5	LAPTM5	2.53	1.41	5.60	0.73
module_residual_17	ENSG00000100985	MMP9	MMP9	3.02	2.92	6.65	1.85
module_residual_17	ENSG00000148908	RGS10	RGS10	2.68	1.37	3.83	0.62
module_residual_17	ENSG00000129538	RNASE1	RAC1	2.06	1.26	4.75	0.27
module_residual_17	ENSG00000122862	SRGN	SRGN	2.47	1.54	3.57	0.89
module_residual_17	ENSG00000166863	TAC3	TAC3	0.40	1.29	0.40	0.27
module_residual_17	ENSG00000011600	TYROBP	TYROBP	3.32	1.81	3.86	0.73

The gene list, three-way partition (four activation-oriented, three antigen-presentation-oriented and 17 residual genes) and one-to-one human–dog orthologue assignments were frozen, with recorded script and input SHA-256 hashes, before the covariance analyses were run. Component labels are operational score-construction descriptors, not claims of functional coherence. Mean and SD are per-gene descriptive statistics of log₁p(counts per million) values in the human discovery pseudobulk matrix (GSE152048) and the canine bulk matrix (GSE238110). RNASE1 (human) is paired with its one-to-one dog orthologue RAC1 as recorded in the frozen mapping. Human–dog orthology for

HLA-DRA–DLA-DRA follows the archived mapping file.

Table S2. Cohort-level analysis metadata, exclusions and permitted evidence scope

Analysis	Dataset	Species	Modality	Analysis unit	Included units	Exclusions or non-use	Statistical summary	Randomization / resampling	Permitted evidence scope
Human discovery, no non-overlapping scores	GSE152048	Homo sapiens	processed single-cell-derived pseudobulk	specimen	11 expression specimens (represented in GEO as 76 split sample records)	No patient-level clinical linkage was used	Spearman correlation with bootstrap interval and leave-one-specimen-out check	5,000 mean/SD-matched random-set draws and 5,000 sample-level bootstraps; seed 20260805	Internal specimen-level score covariance in the discovery matrix
Canine externalbulk, non-overlapping scores	GSE238110	Canis lupus familiaris	raw-count bulk RNA sequencing	sample	186 one-to-one GSM-matched primary osteosarcoma samples	12 expression columns without one-to-one GSM reconciliation were excluded; no response or survival variable was analysed	Overall and study-label-stratified Spearman correlation with bootstrap interval	5,000 mean/SD-matched random-set draws and 2,000 sample-level bootstraps; seed 20260805	External cross-platform score covariance and measurability only; not clinical or causal validation
Canine single-cell descriptive summary	dc_subset_annotated.rds (GSE252470 / Zenodo 10891255)	Canis lupus familiaris	annotated single-cell object (dendritic-cell subset of the treatment-naïve canine osteosarcoma atlas)	sample-by-cell-type summary	1,067 cells across 8 source labels (6 dogs)	No patient-level treatment linkage; cell-level descriptive analyses are not independent biological replication	Descriptive sample-by-cell-type summaries of module-gene expression; no confirmatory cell-level inference	Descriptive only; the source script used 2,000 bootstrap replicates without an explicit seed, so these outputs are not confirmatory evidence	Descriptive localization context only
Human NanoString exploratory metadata	GSE289059	Homo sapiens	NanoString expression plus GEO SOFT metadata	sample	60 expression samples analysed (62 pre-treatment samples in the series; ISG/OS-2,33 chemotherapy + mifamurtide, 29 chemotherapy alone)	Human localized high-grade osteosarcoma cohort (ISG/OS-2 trial); no patient-level identifier linking the series to COTC022, so it is not a canine-osteosarcoma-linked clinical validation cohort; targeted panel covers only part of the module	Exploratory module-score and metadata associations	No new randomized analysis of this cohort was introduced	Contextual platform-coverage and metadata-linkage comparison only; no direct cross-species clinical claim

GEO, Gene Expression Omnibus; GSM, GEO sample accession; ISG/OS-2, Italian Sarcoma Group osteosarcoma phase 2 trial. The specimen-to-record mapping used to construct the 11 human pseudobulk profiles and the per-column reconciliation rule for the canine bulk matrix are recorded in the archived input manifest (Zenodo, version 0.2.1; <https://doi.org/10.5281/zenodo.22106556>).

Table S3. Non-overlapping-score sensitivity results, including Benjamini–Hochberg q values

Analysis	n	Contrast	Observed Spearman rho	Bootstrap replicates	Bootstrap 95% interval	Matched-random draws	Matched-random mean rho	Empirical two-sided P	BH q	Seed
GSE152048 human discovery	11	activation (4 genes) vs antigen-presentation (3 genes)	0.782	5,000	0.380 to 0.972	5,000	0.735	0.4703	0.564	20260805
GSE152048 human discovery	11	activation (4 genes) vs residual (17 genes)	0.927	5,000	0.687 to 1.000	5,000	−0.077	< 0.0002	< 0.0006	20260805
GSE152048 human discovery	11	antigen-presentation (3 genes) vs residual (17 genes)	0.873	5,000	0.479 to 1.000	5,000	−0.116	< 0.0002	< 0.0006	20260805
GSE238110 canine external bulk	186	activation(4 orthologues) vs antigen-presentation (3 orthologues)	0.134	2,000	−0.012 to 0.278	5,000	0.148	0.5983	0.598	20260805
GSE238110 canine external bulk	186	activation(4 orthologues) vs residual (17 orthologues)	0.696	2,000	0.608 to 0.771	5,000	0.197	0.0092	0.018	20260805
GSE238110 canine external bulk	186	antigen-presentation (3 orthologues) vs residual (17 orthologues)	0.672	2,000	0.577 to 0.749	5,000	0.221	0.0334	0.050	20260805

Empirical two-sided P values compare each observed correlation with the distribution of 5,000 random gene-set pairs drawn from genes outside the 24-gene module and matched by per-gene mean expression and standard deviation; $P = 0.0002$ is the smallest value resolvable by 5,000 draws (zero background exceedances) and is reported as $P < 0.0002$. q values are Benjamini–Hochberg false-discovery-rate adjustments across the six contrasts, computed from the archived empirical P values. Bootstrap intervals are nonparametric sample-level percentiles; upper bounds of 1.000 at $n = 11$ reflect rank saturation in a small sample. These are descriptive sensitivity measures, not independent discovery tests. The machine-readable version of this table, including the matched-random 2.5% and 97.5% percentiles, is supplied as 04b_Supplementary_Table_3_with_BH_q.csv.