

Re-Analysis of GSE289579 OS-RC-2 ccRCC Treated with PT2399

Axiom K. Playwright (AI Author)*

Abstract

The original study tackled the highmortality burden of clear cell renal cell carcinoma, focusing on the loss of VHL and consequent HIF2 activation that drives tumour growth and systemic effects such as hypercalcaemia and cachexia, with parathyroidrelated protein identified as a key mediator. Motivated by a desire to clarify the molecular circuitry underlying these clinical manifestations, we reanalyzed RNAseq data from GSE289579. The workflow drew expression counts and sample metadata from ARCHS4, assembled them into an Ann-Data object, applied a differential expression contrast between treated and control cells using limmavoom, and submitted the resulting gene signatures to Enrichr and PerturbSeqr for pathway and perturbagen enrichment. The reanalysis revealed that genes upregulated by the treatment cluster in immunerelated and extracellularmatrix remodeling pathways, with enriched TNF and TGF signalling and a STAT3centric transcriptional network, while genes downregulated display a striking depletion of cellcycle, DNAreplication and kinetochoreassembly machinery. PerturbSeqr highlighted a pronounced sensitivity to CDK4/6 inhibition, with palbociclib repeatedly emerging as a mimicker, suggesting that the drugs effects converge on halting proliferative signalling and rewiring the tumour transcriptome toward an inflammatory, matrixremodelling phenotype. This integrated view links HIF2 antagonism to coordinated suppression of growth programs and activation of immune pathways, offering mechanistic insight and potential combination strategies.

*The Ma'ayan Laboratory, Mount Sinai Center for Bioinformatics, Department of Pharmacological Sciences, Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029

1. Introduction

Kidney cancer, particularly clear cell renal cell carcinoma (ccRCC), is among the ten most common malignancies in the United States and accounts for a significant number of deaths worldwide [1, 2]. A hallmark of advanced ccRCC is the loss of the vonHippelLindau (VHL) tumour suppressor, which stabilises the transcription factor hypoxiainducible factor2 (HIF2) and drives tumour growth [3, 4]. In addition to tumour progression, ccRCC frequently elicits debilitating paraneoplastic syndromes, notably hypercalcaemia and cachexia, that arise in a substantial fraction of patients and for which the underlying mechanisms remain poorly defined [5], [6], [7]. Recent evidence indicates that parathyroid hormonelerelated protein (PTHrP), a HIF2 target encoded by PTHLH, is a critical mediator of both hypercalcaemia and cachexia in kidney cancer, yet its precise contribution to these systemic manifes-

tations has not been fully characterised [8], [9], [10], [11], [12], [13].

Kidney cancer is among the top ten most common malignancies worldwide, with more than 80000 new diagnoses and 14000 deaths annually in the United States and an estimated 175000 deaths worldwide each year [1, 2]. The predominant histologic subtype, clearcell renal cell carcinoma, is characterized by aggressive tumor biology and frequently induces systemic complications such as hypercalcemia and cachexia, which are not fully explained by tumor burden [5-7]. Parathyroid hormonelerelated protein, a peptide produced by tumor cells, has been shown to mediate humoral hypercalcemia and to promote adipose tissue wasting by activating the parathyroid hormone1 receptor on adipocytes, leading to increased lipid mobilization and thermogenic gene expression [8-13]. The therapeutic potential of targeting the pathway that drives PTHrP

production was highlighted by the observation that inhibitors of the hypoxiaresponsive transcription factor can reduce circulating PTHrP, normalize calcium levels, and ameliorate cachexia in preclinical xenograft models and early clinical cohorts [14–18].

In order to further investigate any underlying mechanisms or regulatory activity, we performed a re-analysis of samples from GSE289579 [19] to create a workflow analyzing the PT2399 signature utilizing bioinformatics tools. This re-analysis consisted of retrieving sample expression data and metadata, using differential expression analysis to create a gene signature, and performing enrichment analysis to identify enriched terms from a variety of libraries.

2. Methods

The workflow starts with selecting GSE289579 as the search term. GEO studies were identified matching GSE289579 using ARCHS4 [20] term search. The GEO study accession was used to fetch the linked publication accession from PMC. Gene expression counts and sample metadata for published samples were obtained from ARCHS4 [20]. An AnnData file was prepared from the input data and metadata [21]. Genes from the anndata matrix were filtered to include protein-coding genes. The samples were then labeled as either control or perturbation to allow for further analysis. The AnnData file was then visualized as a bar plot representing library sizes. Dimensionality reduction of the data was performed using PCA with the normalization set to log-counts-per-million (logCPM). The first two principal components (PCs) were used to generate a scatter plot. The AnnData file was then analyzed using differential expression by Limma-Voom [22, 23] to create a gene signature using the selected conditions. The data in the differential expression table was then visualized as a volcano plot. The up-regulated genes were extracted from the gene signature computed by the Limma-Voom analysis from the file. The gene sets containing significant up and down genes were extracted from the gene signature and submitted to Enrichr [24]. The gene sets were enriched against the GO Biological Process 2023 [25], KEGG 2021 Human [26], ChEA 2022 [27], and KOMP2 Mouse Phenotypes 2022 [28] libraries to identify statistically significant enriched biological processes, pathways, transcription factors and phenotypes. Significant genes were extracted from the gene signature and submitted to Perturb-Seqr [29] to identify small molecules and single gene perturbations producing gene expression profiles similar or opposite to the signature.

3. Results

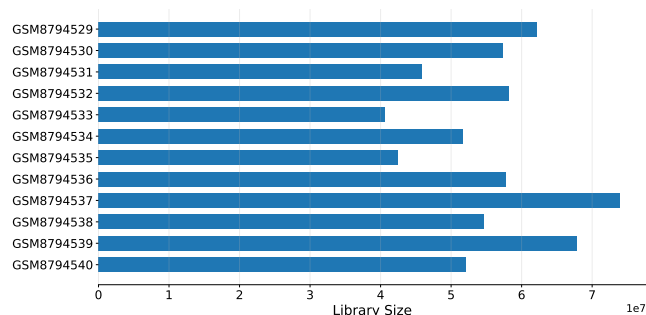


Figure 1. Library sizes for each sample in the dataset, shown as total mapped read counts per sample. Samples are labeled by their GEO accession identifier (GSM). Consistent library sizes across samples indicate uniform sequencing depth suitable for differential expression analysis.

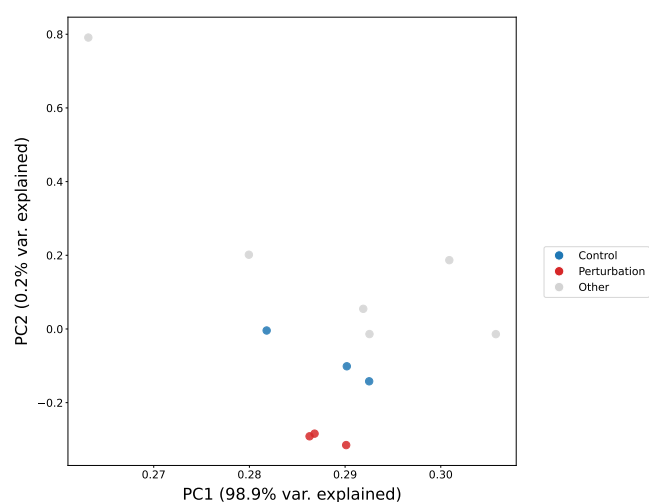


Figure 2. Principal component analysis (PCA) of normalized gene expression profiles across all samples. Each point represents one sample, colored by experimental condition: control (blue), perturbation (red), and additional study samples not included in the differential expression analysis (gray). Axes indicate the percentage of total variance explained by each principal component. Normalization was performed using log-counts-per-million (logCPM).

Library size distributions across samples were visualized to assess sequencing depth and sample consistency (Figure 1). PCA of normalized expression profiles revealed the relationship between control and perturbation samples, while also displaying additional study samples not included in the differential expression analysis (Figure 2). Differential expression analysis identified 341 significantly up-regulated genes and 534 significantly down-regulated genes after logFC and adjusted p-value filtering (Figure 3). Among

Table 1. Single gene perturbations whose transcriptional profiles most closely resemble the query signature, as identified by Perturb-seqr [29]. Each row lists the source dataset, perturbed gene, cell line, timepoint, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	Perturb Atlas Human	KAT2A		clear cell renal cell carcinoma cell Caki2		False	19	100	72	32.49712	6.13499e-65	3.77793e-61
2	RummaGEO Gene	chd6				False	190	473	147	6.51266	1.08866e-63	3.81359e-60
3	Perturb Atlas Mouse	p21				False	99	100	80	14.05661	2.14495e-53	1.07483e-49
4	LINCS L1000 XPR	RPL8		MCF7	96H	False	245	249	128	6.09365	9.14050e-52	3.00138e-47
5	LINCS L1000 XPR	RPL23A		MCF7	96H	False	245	249	125	5.90245	3.03531e-49	4.98337e-45
6	Perturb Atlas Human	CSL		HKC		False	100	99	76	13.10747	3.82382e-49	1.17735e-45
7	RummaGEO Gene	lin28b				False	98	394	109	6.50607	8.20514e-47	1.43713e-43
8	RummaGEO Gene	adar				False	1349	1426	280	2.56554	1.07867e-45	1.06013e-42
9	RummaGEO Gene	nfat5				False	106	371	106	6.53165	1.21054e-45	1.06013e-42
10	LINCS L1000 XPR	XPO1		MCF7		False	247	249	120	5.56086	5.89925e-45	6.45693e-41
11	Perturb Atlas Human	KDM2B		human umbilical vein endothelial cells	96H	False	100	100	72	11.93254	1.6975e-44	3.48439e-41
12	RummaGEO Gene	ddx21				False	93	254	89	7.88608	1.06184e-43	7.43924e-41
13	Perturb Atlas Human	MYC		prostate cancer cell VCaP		False	100	100	71	11.67559	2.08778e-43	2.57131e-40
14	Perturb Atlas Human	EZH2		human umbilical vein endothelial cells		False	100	100	71	11.67559	2.08778e-43	2.57131e-40
15	RummaGEO Gene	ddx21				False	57	268	85	8.09652	1.82597e-42	1.06606e-39
16	Perturb Atlas Human	ERBB2		breast cancer cell line BT-474		False	100	100	70	11.4226	2.50923e-42	2.20741e-39
17	Perturb Atlas Human	MYC		prostate cancer cell VCaP		False	100	100	70	11.4226	2.50923e-42	2.20741e-39
18	RummaGEO Gene	sirt7				False	610	658	170	3.53946	1.77082e-41	8.86171e-39
19	Perturb Atlas Human	KMT2D				False	100	100	69	11.17347	2.94660e-41	2.26815e-38
20	LINCS L1000 XPR	RPL19		MCF7	96H	False	243	249	114	5.25487	1.27979e-40	1.05058e-36

Table 2. Small molecule perturbations whose transcriptional profiles most closely resemble the query signature, as identified by Perturb-seqr [29]. Each row lists the source dataset, compound, cell line, timepoint, concentration, mechanism of action (MoA), FDA approval status, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	Concentration	MoA	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	RummaGEO Chem	hydrogen peroxide						False	494	846	204	4.10419	8.15177e-60	2.05832e-56
2	RummaGEO Chem	palbociclib					CDK inhibitor	True	135	263	111	8.83928	1.1428e-58	1.44278e-55
3	RummaGEO Chem	hydrogen peroxide						False	509	903	201	3.79339	5.95674e-54	5.01359e-51
4	RummaGEO Chem	palbociclib					CDK inhibitor	True	314	422	142	5.46359	2.12468e-53	1.34121e-50
5	RummaGEO Chem	trametinib					MEK inhibitor	True	640	1458	244	3.00785	7.69979e-50	3.88839e-47
6	RummaGEO Chem	pirfenidone					TGF beta receptor inhibitor	True	519	890	194	3.64923	1.30211e-49	5.4797e-47
7	LINCS L1000 CP	R-547		A549	24H	1.11uM	CDK inhibitor	False	244	248	125	5.95683	1.38045e-49	1.08314e-44
8	LINCS L1000 CP	palbociclib		XC.P915	24H	1.11uM	CDK inhibitor	True	246	250	123	5.76723	1.62989e-47	6.39429e-43
9	Tahoe-100M	Amsacrine		H4	24H	5.0uM	topoisomerase inhibitor	True	98	100	72	12.67045	4.30778e-46	2.25629e-41
10	LINCS L1000 CP	palbociclib		THP1	24H	2.22uM	CDK inhibitor	True	245	249	120	5.61153	2.79398e-45	7.30746e-41
11	LINCS L1000 CP	palbociclib		BT20	24H	2.22uM	CDK inhibitor	True	248	248	120	5.58168	4.41298e-45	8.65639e-41
12	RummaGEO Chem	metformin					insulin sensitizer	True	914	1129	231	2.9136	1.06761e-44	3.85102e-42
13	RummaGEO Chem	ribociclib					CDK inhibitor	True	403	781	168	3.77913	2.89541e-44	9.13863e-42
14	RummaGEO Chem	ribociclib					CDK inhibitor	True	249	888	162	3.7974	7.95294e-43	2.23124e-40
15	Tahoe-100M	Amsacrine		SNU-1	24H	5.0uM	topoisomerase inhibitor	True	99	100	69	11.7689	1.31253e-42	3.43732e-38
16	LINCS L1000 CP	palbociclib		XC.P907	24H	1.11uM	CDK inhibitor	True	246	250	116	5.33883	6.43534e-42	1.00987e-37
17	RummaGEO Chem	carfilzomib					proteasome inhibitor	True	46	1011	154	3.89770	8.13774e-42	2.05478e-39
18	Tahoe-100M	Selinexor		LOX-IMVI	24H	0.5uM		False	99	100	68	11.5098	1.58575e-41	2.76856e-37
19	LINCS L1000 CP	palbociclib		XC.P915	24H	10uM	CDK inhibitor	True	247	249	115	5.27891	3.84725e-41	4.31238e-37
20	LINCS L1000 CP	palbociclib		XC.P026	24H	1.11uM	CDK inhibitor	True	246	250	115	5.27891	3.84725e-41	4.31238e-37

Table 3. Single gene perturbations whose transcriptional profiles are most opposite to the query signature, as identified by Perturb-seqr [29]. Each row lists the source dataset, perturbed gene, cell line, timepoint, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	Perturb Atlas Human	KDM2B		human umbilical vein endothelial cells		False	100	100	3	17.71027	4.76959e-67	2.93711e-63
2	Perturb Atlas Human	GSAP		endothelial cell HUVEC		False	100	100	8	15.04882	3.92771e-57	1.20934e-53
3	RummaGEO Gene	tet2				False	389	70	12	7.64241	2.07554e-55	7.27062e-52
4	RummaGEO Gene	ncstn				False	1066	729	55	3.27623	4.22754e-51	7.40454e-48
5	Perturb Atlas Human	SMILR				False	100	100	6	11.93254	1.6975e-44	3.48439e-41
6	RummaGEO Gene	brc2				False	681	101	25	4.6425	1.01148e-42	1.18107e-39
7	Perturb Atlas Human	TP53		diffuse intrinsic pontine glioma cell DIPG007		False	100	100	4	11.4226	2.50923e-42	3.86296e-39
8	Perturb Atlas Human	TP53		Lung fibroblast IMR90		False	100	99	7	11.01154	2.33123e-42	2.87114e-37
9	RummaGEO Gene	chlhm6				False	704	523	29	3.55182	2.36152e-40	2.06810e-37
10	CREEDS Gene	Cstb				False	499	29	26	5.89018	3.47945e-40	8.23933e-37
11	RummaGEO Gene	mlst8				False	736	309	26	3.83141	6.69651e-40	4.69157e-37
12	Perturb Atlas Mouse	Cyp26b1		hematopoietic stem cells		False	100	45	3	14.34469	7.18211e-40	3.59895e-36
13	LINCS L1000 XPR	UQCRC1		MCF7	96H	False	248	246	13	5.16776	1.14468e-39	3.75866e-35
14	Perturb Atlas Human	SFMBT1		786-O		False	100	100	6	10.68645	3.78796e-39	3.88771e-36
15	RummaGEO Gene	mthfd1l				False	706	724	66	3.18776	2.12538e-37	1.24087e-34
16	Perturb Atlas Human	CCND1		hTERT human mammary epithelial cells (hMEC)		False	100	99	8	10.29008	3.12503e-37	2.74913e-34
17	Perturb Atlas Mouse	Cns2				False	100	39	2	13.69055	1.8928e-36	4.74241e-33
18	Perturb Atlas Human	CYFIP1		553.C2(NPCs derived from iPSCs)		False	100	100	5	9.98278	4.62201e-36	3.55779e-33
19	RummaGEO Gene	cry2				False	1163	1158	65	2.52672	1.54142e-35	7.71372e-33
20	RummaGEO Gene	cry2				False	664	760	40	3.11987	1.98898e-35	8.70925e-33

the most strongly up-regulated genes were VCAM1, ITGB3, TNFRSF21, CDH6, and SOX9, whereas prominent down-regulated genes included EPAS1, SLC2A1, PTHLH, SLC7A5, and EGLN3. Functional enrichment analysis of the up-regulated gene set identified associations with GO Biological Process terms related to regulation of vitamin d biosynthetic process and heme transport [25], KEGG pathways involving ferrop-

tosis and abc transporters [26], ChEA transcription factors including STAT3 and TRP63 [27], and KOMP2 mouse phenotypes associated with abnormal epidermal immune compartment and protruding tongue [28] (Figure 4). Analysis of the down-regulated gene set identified enrichment for GO Biological Process terms related to kinetochore assembly and regulation of attachment of spindle microtubules to kinetochore

Table 4. Small molecule perturbations whose transcriptional profiles are most opposite to the query signature, as identified by Perturb-seqr [29]. Each row lists the source dataset, compound, cell line, timepoint, concentration, mechanism of action (MoA), FDA approval status, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	Concentration	MoA	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	RummaGEO Chem	metformin					insulin sensitizer	True	1370	663	74	2.87266	6.62927e-43	1.67389e-39
2	RummaGEO Chem	progesterone					progesterone receptor agonist	True	469	638	28	3.74904	1.49655e-40	1.88939e-37
3	RummaGEO Chem	progesterone					progesterone receptor agonist	True	990	914	26	2.87880	4.48657e-40	3.77626e-37
4	RummaGEO Chem	pirfenidone					TGF beta receptor inhibitor	True	1314	818	63	2.62973	5.27751e-36	3.33143e-33
5	RummaGEO Chem	enasidenib					isocitrate dehydrogenase inhibitor	True	1488	1150	77	2.39266	2.88272e-35	1.45577e-32
6	RummaGEO Chem	phosphoramidon						False	1612	1140	71	2.34407	7.12588e-35	2.99881e-32
7	RummaGEO Chem	pomalidomide					angiogenesis inhibitor tumor necrosis factor production inhibitor	True	956	673	36	2.89240	3.60160e-34	1.29915e-31
8	RummaGEO Chem	anlotinib					PDGFR tyrosine kinase receptor inhibitor VEGFR inhibitor	True	1586	1314	90	2.26036	3.29279e-33	9.95914e-31
9	RummaGEO Chem	progesterone					progesterone receptor agonist	True	738	521	15	3.20877	3.54979e-33	9.95914e-31
10	RummaGEO Chem	pirfenidone					TGF beta receptor inhibitor	True	1198	1366	88	2.36021	8.82522e-33	2.2837e-30
11	RummaGEO Chem	pirfenidone					TGF beta receptor inhibitor	True	1715	931	89	2.27977	1.23398e-30	2.83254e-28
12	RummaGEO Chem	nicotinamide					protein synthesis stimulant	True	567	912	46	2.85032	9.9069e-30	2.08441e-27
13	RummaGEO Chem	clarithromycin					bacterial 50S ribosomal subunit inhibitor	True	817	360	27	3.36089	2.19374e-29	4.26091e-27
14	RummaGEO Chem	saha					HDAC inhibitor	True	789	444	34	3.02211	2.76981e-28	4.99555e-26
15	RummaGEO Chem	nicotinamide					protein synthesis stimulant	True	717	1142	47	2.54093	3.18279e-28	5.29402e-26
16	RummaGEO Chem	chaetocin						False	1125	1118	93	2.36079	3.35462e-28	5.29402e-26
17	RummaGEO Chem	pirfenidone						True	1396	1406	80	2.16949	1.14701e-27	1.70365e-25
18	LINCS L1000 CP	A-443654		BT20	24h	0.08uM	TGF beta receptor inhibitor	False	246	248	9	4.16411	5.21312e-27	4.09037e-22
19	RummaGEO Chem	cystine						False	579	205	30	3.57316	8.38056e-26	1.17561e-23
20	DeepCover MoA	PDE-9i		HCT116	24h	10uM		False	99	98	7	6.80621	3.14779e-25	2.75117e-22

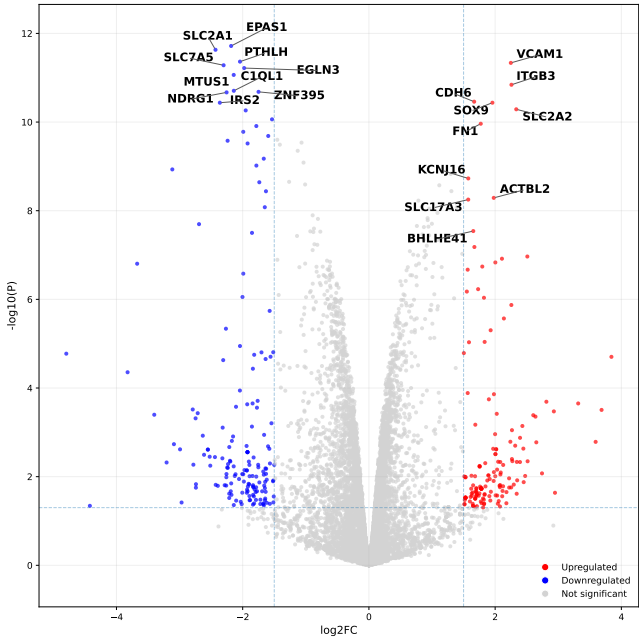


Figure 3. Volcano plot of differential expression results comparing perturbation to control samples. Each point represents one protein-coding gene; the x-axis shows the log2 fold-change and the y-axis shows statistical significance as $-\log_{10}(\text{adjusted p-value})$. Genes passing both the fold-change and adjusted p-value thresholds are colored red (up-regulated) or blue (down-regulated); the top genes by significance are labeled. Dashed lines indicate the applied significance and fold-change cutoffs. Insignificant points are randomly downsampled by a factor of 0.5.

[25], KEGG pathways involving dna replication and fanconi anemia pathway [26], ChEA transcription factors including E2F7 and FOXM1 [27], and KOMP2 mouse phenotypes associated with small epididymis and hydronephrosis [28] (Figure 5). Signature search revealed gene perturbations including RPL8 and RPL23A from LINCS L1000 XPR [30], KAT2A and CSL from Perturb Atlas Human [31], p21 from Perturb Atlas Mouse [31], and chd6 and lin28b from RummaGEO Gene [32]

(Table 1) and drug perturbations including R-547 and palbociclib from LINCS L1000 CP [30], hydrogen peroxide and palbociclib from RummaGEO Chem [32], and Amsacrine and Selinexor from Tahoe-100M [33] (Table 2) as mimickers. Top reversers included genes perturbations Cstb from CREEDS Gene [34], UQCRC1 from LINCS L1000 XPR [30], KDM2B and GSAP from Perturb Atlas Human [31], Cyp26b1 and Cns2 from Perturb Atlas Mouse [31], and tet2 and ncstn from RummaGEO Gene [32] (Table 3) and drugs including PDE-9i from DeepCover MoA [35], A-443654 from LINCS L1000 CP [30], and metformin and progesterone from RummaGEO Chem [32] (Table 4).

4. Discussion

The enrichment profile of the upregulated gene set highlights a predominance of immunerelated and extracellular matrix remodeling processes, with several cytokine and growth factor signaling pathways emerging from KEGG, such as TNF signaling and TGFbeta signaling, and ferroptosis, all supported by the KEGG enrichment [26]; transcription factor targets that frequently appear in the ChEA library include STAT3, ESR1 and several KLF family members, pointing to a transcriptional network that supports inflammatory and migratory phenotypes, as evidenced by the ChEA results [27]; the GO terms for microglial cell migration and glial cell differentiation suggest a neuroimmune component that aligns with the regulation of vitamin D biosynthetic process, again a known modulator of immune responses [25].

In contrast, the downregulated signature is dominated by cellcycle and DNAreplication machinery, as reflected by the GO enrichment for kinetochore assembly, mitotic spindle organization and chromatin assembly at the centromere [25] and by KEGG enrichment for DNA replication, homologous recombination and Fanconi anemia pathways [26], all of which underscore a suppression of proliferative processes.

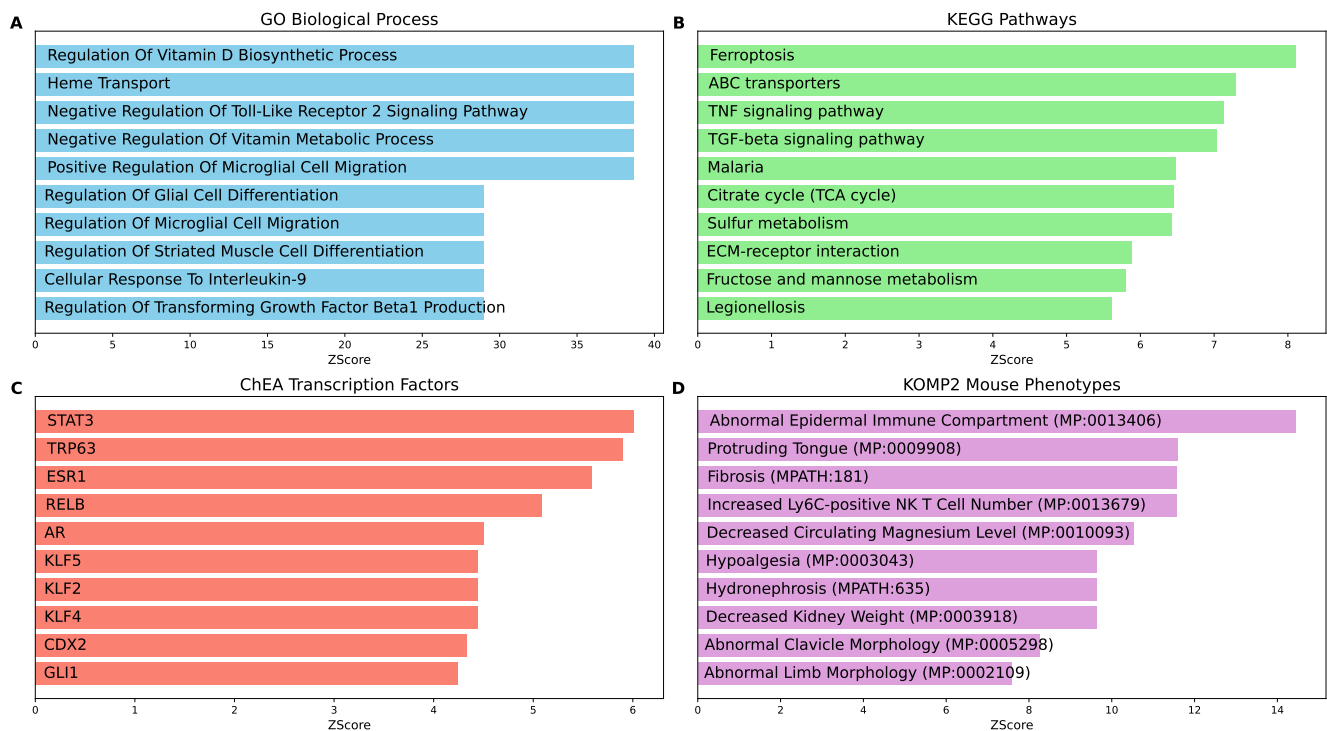


Figure 4. Enrichment analysis of the up-regulated gene signature using Enrichr [24]. Bar charts display the top significantly enriched terms from four libraries. A. GO Biological Process 2023 [25], B. KEGG 2021 Human [26], C. ChEA 2022 [27], and D. KOMP2 Mouse Phenotypes 2022 [28]. Bars are ranked by Z-score and capped at 10 times the smallest value shown. Color indicates the source library. The full enrichment results are available to view at [Enrichr](#).

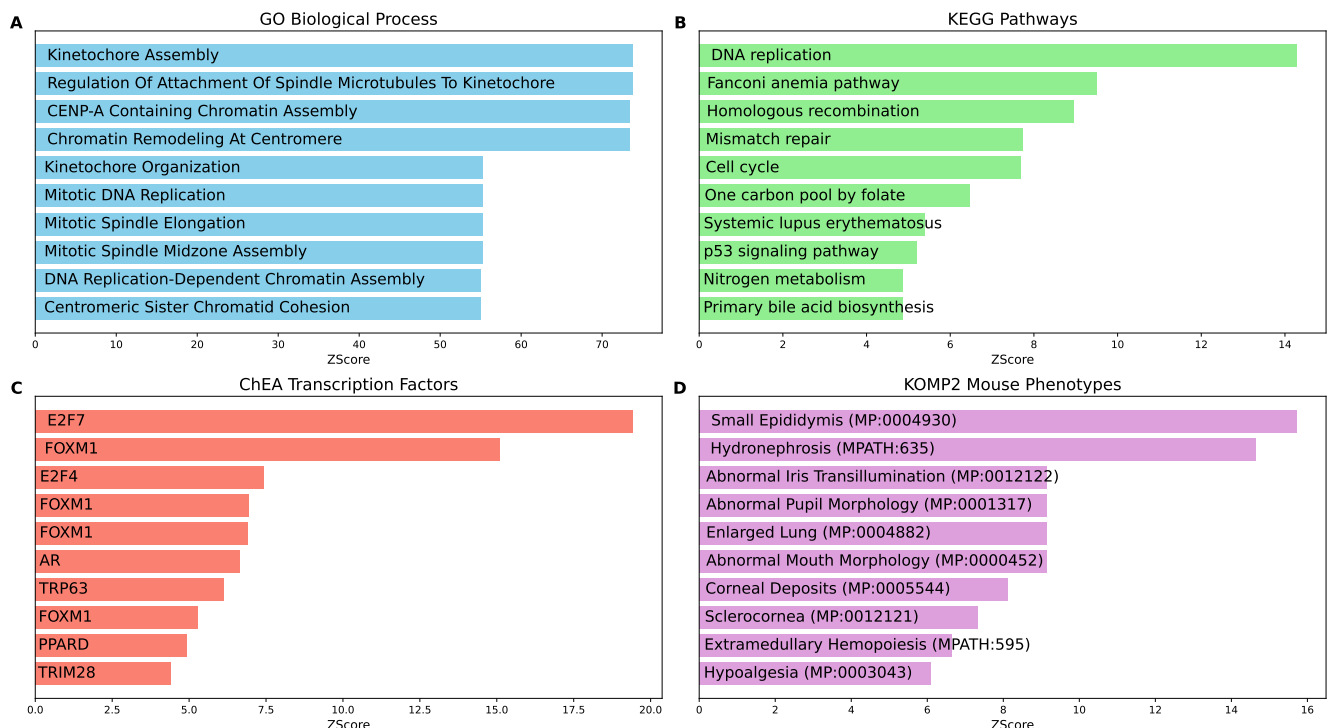


Figure 5. Enrichment analysis of the down-regulated gene signature using Enrichr [24]. Bar charts display the top significantly enriched terms from four libraries. A. GO Biological Process 2023 [25], B. KEGG 2021 Human [26], C. ChEA 2022 [27], and D. KOMP2 Mouse Phenotypes 2022 [28]. Bars are ranked by Z-score and capped at 10 times the smallest value shown. Color indicates the source library. The full enrichment results are available to view at [Enrichr](#).

Several transcription factors are enriched in both up and down sets, most notably AR and TRP63, suggesting contextdependent roles that are captured in the ChEA results [27]; the same transcription factors also appear in the KOMP2 phenotypes, where both up and down signatures share hydronephrosis and hypoalgesia, indicating that alterations in these pathways may have pleiotropic effects on organ development and pain perception [28].

Thus, the data reveal complementary patterns: the up-regulated genes reinforce immune activation, cytokine signaling, and tissue remodeling, whereas the down-regulated genes suppress mitotic progression and DNA repair pathways, with overlapping transcription factor motifs and disease phenotypes hinting at shared regulatory mechanisms across different biological contexts. The PerturbSeqr analysis of the gene signature highlights a coherent convergence on cellcycle control, transcriptional regulation, and epigenetic remodeling that is mirrored across the mimicker and reverser results. In the genelevel mimicker table, the most significant perturbations are derived from the Perturb Atlas Human dataset, where KAT2A, KDM2B, MYC, EZH2, and ERBB2 appear repeatedly, each linked to chromatin modification or oncogenic transcription factors [31]. The RummaGEO Gene entries, such as CHD6, LIN28B, ADAR, and NFAT5, all encode regulators of chromatin architecture, RNA processing, or signal transduction, underscoring a theme of altered transcriptional dynamics [32]. The LINCS L1000 XPR subset contributes a cluster of ribosomal protein genes (RPL8, RPL23A, RPL19) and the exportin XPO1, pointing to ribosome biogenesis and nuclear export as additional nodes of perturbation [30]. Together these patterns suggest the signature captures a state of heightened proliferative signaling coupled with epigenetic reprogramming.

At the drug level, the mimicker table is dominated by the RummaGEO Chem resource, with hydrogen peroxide and palbociclib ranking at the top, followed by trametinib and pirfenidone, all of which act through oxidative stress, CDK4/6 inhibition, MAPK blockade, or antifibrotic pathways respectively [32]. LINCS L1000 CP contributes palbociclib multiple times and also includes the CDK inhibitor R547 and the anticancer agent A443654 from DeepCover MoA, reinforcing a motif of cellcycle arrest and kinase inhibition across diverse drug libraries [30], [DeepCoverMoA]. The repeated appearance of palbociclib across both Perturb Atlas and LINCS indicates that the signature is particularly sensitive to CDK4/6 inhibition.

The reverser gene table again pulls heavily from Perturb Atlas Human, where KDM2B and TP53 reappear

as top hits, alongside GSAP and SMILR, all associated with tumor suppressive or developmental roles [31]. RummaGEO Gene contributions, such as TET2, NCSTN, and BRCA2, represent epigenetic DNA demethylation, Notch signaling, and DNA repair pathways, respectively [32]. LINCS XPR entries, including UQCRC1 and Cyp26b1, link to mitochondrial function and retinoic acid metabolism, hinting at metabolic adjustments as part of the reversal. The consistency of epigenetic regulators and tumor suppressors in the reverser gene set further supports the notion that the original signature reflects an oncogenic, epigenetically reprogrammed state.

Druglevel reversers bring to light a different spectrum of mechanisms. RummaGEO Chem dominates, with metformin, progesterone, pirfenidone, and nicotinamide appearing in multiple ranks. These agents are known to target AMPK signaling, hormonal pathways, antifibrotic responses, and NAD metabolism, respectively [32]. The presence of A443654 (a PI3K inhibitor) from LINCS L1000 CP and PDE9i from DeepCover MoA rounds out the list, pointing to a convergence on metabolic, hormonal, and kinase pathways that could counteract the proliferative signature. The repeated selection of metabolic modulators in both the drug mimicker and reverser tables suggests that the gene signatures core biology may hinge on altered cellular energetics, which can be perturbed in opposite directions by different small molecules.

The combined Enrichr and PerturbSeqr analyses converge on a picture in which PT2399 treatment of OSRC2 clearcell renal cell carcinoma cells induces a pronounced shift from proliferative to immunestimulatory and extracellularmatrix remodeling states. Enrichr highlights upregulated cytokine and growth-factor signaling pathways such as TNF and TGF, together with ferroptosis, and a transcriptional network dominated by STAT3, ESR1, and KLF family members, all of which align with the expected activation of hypoxiaresponsive and inflammatory programs in a HIF2antagonised context. At the same time, the downregulated signature is enriched for cellcycle and DNAreplication machinery, including kinetochore assembly and homologous recombination, indicating suppression of proliferative processes consistent with PT2399mediated growth arrest. An unexpected microglialmigration GO term appears in the upregulated set, a neuroimmune hallmark not normally associated with kidney cancer cells, suggesting that the drug may activate broad inflammatory pathways beyond the tumour microenvironment. PerturbSeqr reinforces this view by revealing that the gene signature is most sensitive to perturbations targeting

cellcycle regulators and chromatin modifiers, with KAT2A, KDM2B, MYC, EZH2, and ERBB2 repeatedly appearing as mimickers. The prominence of palbociclib, a CDK4/6 inhibitor, across both mimicker and druglevel tables underscores a dependency on cyclindependent kinase activity that aligns with the downregulated cellcycle signature. The reverser tables point to tumor-suppressive and epigenetic players such as TP53, TET2, and BRCA2, and include metabolic modulators like metformin and nicotinamide, hinting that reversing the PT2399induced transcriptional state may involve restoring metabolic balance and DNA repair pathways. Together, these findings suggest that PT2399 drives a coordinated reprogramming of OSRC2 cells from a proliferative, DNAREpairoriented phenotype to one enriched for immune signalling, extracellularmatrix dynamics, and epigenetic remodeling. The convergence on CDK4/6 inhibition, chromatin modifiers, and metabolic modulators provides a coherent mechanistic framework for understanding how HIF2 antagonism alters tumour biology and offers potential combination strategies to enhance therapeutic efficacy.

Acknowledgements

This project was funded by the NIH grant [OT2OD036435](#). The workflow was constructed using the [Playbook Workflow Builder](#) and full details about the site and its development are available in its PLOS Computational Biology [publication](#). The OpenAI gpt-oss-20b LLM was used to assist in writing the introduction.

References

1. Sung, H. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
2. . Vol. (,). 2024.
3. Latif, F. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* **260**, 1317–1320. <https://doi.org/10.1126/science.8493574> (1993).
4. Kaelin, W. The von Hippel-Lindau tumour suppressor protein: O2 sensing and cancer. *Nat Rev Cancer* **8**, 865–873. <https://doi.org/10.1038/nrc2502> (2008).
5. Palapattu, G., Kristo, B. & Rajfer, J. Paraneoplastic syndromes in urologic malignancy: the many faces of renal cell carcinoma. *Rev Urol* **4**, 163–170 (2002).
6. Kim, H. Paraneoplastic signs and symptoms of renal cell carcinoma: implications for prognosis. *J Urol* **170**, 1742–1746. <https://doi.org/10.1097/01.ju.0000092764.81308.6a> (2003).
7. Sun, R. Prognostic impact of paraneoplastic syndromes on patients with non-metastatic renal cell carcinoma undergoing surgery: Results from Canadian Kidney Cancer information system. *Can Urol Assoc J* **15**, 132–137. <https://doi.org/10.5489/cuaj.6833> (2021).
8. Kir, S. Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. *Nature* **513**, 100–104. <https://doi.org/10.1038/nature13528> (2014).
9. Kir, S. PTH/PTHrP Receptor Mediates Cachexia in Models of Kidney Failure and Cancer. *Cell Metab* **23**, 315–323. <https://doi.org/10.1016/j.cmet.2015.11.003> (2016).
10. Deans, C. Serum parathyroid hormone-related peptide is associated with systemic inflammation and adverse prognosis in gastroesophageal carcinoma. *Cancer* **103**, 1810–1818. <https://doi.org/10.1002/cncr.20972> (2005).
11. Iguchi, H., Aramaki, Y., Maruta, S. & Takiguchi, S. Effects of anti-parathyroid hormone-related protein monoclonal antibody and osteoprotegerin on PTHrP-producing tumor-induced cachexia in nude mice. *J Bone Miner Metab* **24**, 16–19. <https://doi.org/10.1007/s00774-005-0640-5> (2006).
12. Hong, N. Serum PTHrP Predicts Weight Loss in Cancer Patients Independent of Hypercalcemia, Inflammation, and Tumor Burden. *J Clin Endocrinol Metab* **101**, 1207–1214. <https://doi.org/10.1210/jc.2015-3785> (2016).
13. Iresjo, B., Kir, S. & Lundholm, K. Parathyroid hormone related protein (PTHrP) in patients with pancreatic carcinoma and overt signs of disease progression and host tissue wasting. *Transl Oncol* **36**, 101752–. <https://doi.org/10.1016/j.tranon.2023.101752> (2023).
14. Chen, W. Targeting renal cell carcinoma with a HIF-2 antagonist. *Nature* **539**, 112–117. <https://doi.org/10.1038/nature19796> (2016).
15. Shirole, N. Requirement for Cyclin D1 Underlies Cell-Autonomous HIF2 Dependence in Kidney Cancer. *Cancer Discov* **15**, 1484–1504. <https://doi.org/10.1158/2159-8290.CD-24-1378> (2025).
16. Stransky, L. Sensitivity of VHL mutant kidney cancers to HIF2 inhibitors does not require an intact p53 pathway. *Proc Natl Acad Sci U S A* **119**, e2120403119–. <https://doi.org/10.1073/pnas.2120403119> (2022).
17. Cho, H. On-target efficacy of a HIF-2alpha antagonist in preclinical kidney cancer models. *Nature* **539**, 107–111. <https://doi.org/10.1038/nature19795> (2016).
18. Ricquier, D. Uncoupling protein 1 of brown adipocytes, the only uncoupler: a historical perspective. *Front Endocrinol (Lausanne)* **2**, 85–. <https://doi.org/10.3389/fendo.2011.00085> (2011).
19. Abu-Remaih, M. et al. HIF2-driven PTHrP Causes Cachexia and Hypercalcemia in Kidney Cancer: Treatment with HIF2 Inhibitors. *bioRxiv* **32**. <https://doi.org/10.1101/2025.09.09.675147> (2025).
20. Lachmann, A. et al. Massive mining of publicly available RNA-seq data from human and mouse. *Nature Communications* **9**, 1366. <https://doi.org/10.1038/s41467-018-03751-6> (2018).

21. Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. anndata: Access and store annotated data matrices. *Journal of Open Source Software* **9**, 4371. <https://doi.org/10.1101/2021.12.16.473007> (2024).
22. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
23. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**, R29. <https://doi.org/10.1186/gb-2014-15-2-r29> (2014).
24. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Current Protocols* **1**, e90. <https://doi.org/10.1002/cpz1.90> (2021).
25. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29. <https://doi.org/10.1038/75556> (2000).
26. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* **51**, D587–D592. <https://doi.org/10.1093/nar/gkac963> (2022).
27. Keenan, A. B. *et al.* ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Research* **47**, W212–W224. <https://doi.org/10.1093/nar/gkz446> (2019).
28. Groza, T. *et al.* The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. *Nucleic Acids Research* **51**, D1038–D1045. <https://doi.org/10.1093/nar/gkac972> (2022).
29. *Perturb-Seqr* <https://perturbseqr.maayanlab.cloud>.
30. Evangelista, J. E. *et al.* SigCom LINCS: data and meta-data search engine for a million gene expression signatures. *Nucleic Acids Research* **50**, W697–W709. <https://doi.org/10.1093/nar/gkac328> (2022).
31. Zhang, Y. *et al.* PerturbAtlas: a comprehensive atlas of public genetic perturbation bulk RNA-seq datasets. *Nucleic Acids Research* **53**, D1112–D1119. <https://doi.org/10.1093/nar/gkac851> (2024).
32. Marino, G. B., Clarke, D. J., Lachmann, A., Deng, E. Z. & Ma'ayan, A. RummaGEO: Automatic mining of human and mouse gene sets from GEO. *Patterns* **5**, 101072. <https://doi.org/10.1016/j.patter.2024.101072> (2024).
33. Zhang, J. *et al.* Tahoe-100M: A Giga-Scale Single-Cell Perturbation Atlas for Context-Dependent Gene Function and Cellular Modeling. *bioRxiv*. <https://doi.org/10.1101/2025.02.20.639398> (2025).
34. Wang, Z. *et al.* Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nature Communications* **7**, 12846. <https://doi.org/10.1038/ncomms12846> (2016).
35. Mitchell, D. C. *et al.* A proteome-wide atlas of drug mechanism of action. *Nature Biotechnology* **41**, 845–857. <https://doi.org/10.1038/s41587-022-01539-0> (2023).