

Age- and Sex-Stratified Transcriptomic Profiling of Colon Adenocarcinoma Reveals Autosomal and Functionally Relevant Sex-Chromosome Signatures

Syed Arslan Haider, Prince Danan Biniyam, Nimra Mazhar, BiBi Ayesha, Hamna Tariq and Isfa Sarfraz

The following supplementary tables provide complete statistical data for all differentially expressed genes (DEGs) identified across cohorts. These tables support the condensed findings reported in the main manuscript Results section.

Supplementary Table S1: Top 30 most significantly downregulated genes in young-onset colon adenocarcinoma (TCGA cohort; male vs. female comparison).

Genes are ranked by adjusted p-value. Negative log2 fold change values indicate female-enriched expression relative to male patients. Analysis performed using DESeq2 (padj < 0.05, |log2FC| > 1.0).

Gene Symbol	Log2 Fold Change	p-value
XIST	-10.1625	1.54E-110
IGF2	-5.8928	8.76E-16
MAGEA3	-9.5404	1.13E-14
MTCO1P40	-4.4152	2.03E-12
TSIX	-2.8133	1.07E-09
SOX1	-4.2271	1.64E-08
SLC26A2	-2.5216	1.73E-08
MT-TM	-3.7705	5.36E-07
DEFA5	-3.6704	9.79E-07
UMODL1	-2.2440	1.60E-06
LINC01602	-3.1628	1.70E-06
PUDP	-0.8153	1.78E-06
MAGEA12	-5.0030	3.49E-06
MTND1P23	-2.9986	4.43E-06
AQP8	-3.0127	5.59E-06
KDM5C	-0.6287	5.73E-06
PDE6A	-2.1629	8.92E-06
CSAG1	-4.5331	1.07E-05
CLCA1	-3.2085	1.27E-05
SNHG25	-2.7453	1.33E-05
PRSS1	-2.6966	1.41E-05
IGFALS	-2.5833	1.48E-05
UICLM	-3.4870	1.83E-05
PNPLA4	-1.0041	1.96E-05
NTS	-2.7550	1.98E-05
MAEL	-2.5220	2.05E-05
LINC01819	-3.0381	3.05E-05
MAGEA6	-4.6133	3.85E-05
ARSD	-0.8275	3.87E-05
RPL23AP65	-1.8068	5.06E-05

Supplementary Table S2: Top 30 most significantly upregulated genes in young-onset colon adenocarcinoma (TCGA cohort; male vs. female comparison).

Positive log2 fold change values indicate male-enriched expression.

Gene Symbol	log2FC	lfcSE	Stat	p-value	padj
ZFY	8.0805	0.5348	15.1101	1.39E-51	1.77E-47
LINC00278	6.9381	0.5057	13.7206	7.64E-43	6.47E-39
DDX3Y	6.9524	0.5317	13.0762	4.50E-39	2.86E-35
ZNF736P9Y	8.5294	0.6635	12.8549	8.07E-38	4.10E-34
TMSB4Y	7.9560	0.6197	12.8395	9.85E-38	4.17E-34
USP9Y	7.0436	0.5807	12.1299	7.34E-34	2.66E-30
NLGN4Y	7.0457	0.5865	12.0136	3.01E-33	9.58E-30
ANOS2P	7.1625	0.5970	11.9972	3.68E-33	1.04E-29
EIF1AY	7.5405	0.6358	11.8603	1.90E-32	4.83E-29
GYG2P1	7.5504	0.6407	11.7845	4.69E-32	1.08E-28
RPS4Y1	7.1484	0.6080	11.7563	6.56E-32	1.39E-28
KDM5D	6.6366	0.6007	11.0487	2.22E-28	4.35E-25
TXLNGY	6.9255	0.6479	10.6891	1.15E-26	2.08E-23
PRKY	6.0511	0.5683	10.6468	1.80E-26	3.06E-23
TTY14	6.4710	0.6117	10.5782	3.76E-26	5.98E-23
BCORP1	5.5922	0.5469	10.2253	1.53E-24	2.28E-21
UTY	5.9558	0.7009	8.4978	1.93E-17	2.73E-14
PRORY	6.4982	0.9021	7.2033	5.88E-13	7.11E-10
ARSDP1	4.3385	0.6073	7.1436	9.09E-13	1.05E-09
MUC5AC	4.2620	0.6575	6.4824	9.03E-11	9.56E-08
PLAAT5	3.2459	0.5383	6.0298	1.64E-09	1.60E-06
ZFY-AS1	3.9514	0.7446	5.3069	1.12E-07	9.77E-05
TTY10	3.4766	0.6602	5.2657	1.40E-07	1.18E-04
SEPTIN3	1.7338	0.3561	4.8688	1.12E-06	8.65E-04
OFD1P16Y	4.4915	0.9563	4.6966	2.64E-06	0.001816
MX2	1.2878	0.2768	4.6523	3.28E-06	0.002195
RNASE1	1.1437	0.2495	4.5841	4.56E-06	0.002826
IFIT2	1.1753	0.2654	4.4293	9.45E-06	0.005337
OLFM1	1.7711	0.4007	4.4203	9.86E-06	0.005445
RBMV2EP	4.8761	1.1251	4.3339	1.46E-05	0.007236

Supplementary Table S3: Top 30 most significantly downregulated genes in late-onset colon adenocarcinoma (TCGA cohort; ≥ 50 years, male vs. female comparison).

Genes enriched in female late-onset tumors.

Gene	BaseMean	log2FC	lfcSE	Stat	p-value	padj
XIST	1930.50	-9.1374	0.4222	-21.6429	7.09E-104	9.42E-100
TSIX	4.67	-2.9031	0.3627	-8.0033	1.21E-15	1.46E-12
ZFX	821.69	-0.5374	0.0836	-6.4288	1.29E-10	1.42E-07
KDM5C	5549.69	-0.4663	0.0800	-5.8271	5.64E-09	5.86E-06
CST7	210.90	-2.0283	0.3512	-5.7749	7.70E-09	7.58E-06
CTXND1	11.47	-2.1618	0.4033	-5.3596	8.34E-08	7.15E-05
ZRSR2	474.35	-0.4962	0.0938	-5.2892	1.23E-07	9.89E-05
CALB1	80.26	-2.7085	0.5195	-5.2135	1.85E-07	0.000145
SEMG2	4.10	-4.5564	0.8964	-5.0827	3.72E-07	0.000275
ZNF542P	55.61	-1.1869	0.2372	-5.0046	5.60E-07	0.000372
IGHV7-4-1	83.52	-3.2541	0.6563	-4.9585	7.10E-07	0.000460
NALF1	52.72	-1.8379	0.3728	-4.9298	8.23E-07	0.000521
ANPEP	3241.63	-2.0134	0.4106	-4.9040	9.39E-07	0.000580
IGKV2-29	20.01	-3.0534	0.6269	-4.8706	1.11E-06	0.000672
KDM6A	891.29	-0.5038	0.1044	-4.8267	1.39E-06	0.000819
CPA2	22.09	-2.4018	0.5011	-4.7932	1.64E-06	0.000928
ZNF385B	7.90	-1.9461	0.4127	-4.7156	2.41E-06	0.001307
FUT7	19.86	-1.4329	0.3059	-4.6839	2.81E-06	0.001496
MT3	65.59	-2.0078	0.4307	-4.6617	3.14E-06	0.001634
NHIP	3.38	-1.8452	0.3974	-4.6433	3.43E-06	0.001752
PRKX	1203.06	-0.6927	0.1514	-4.5767	4.72E-06	0.002282
CSPG5	31.09	-1.0844	0.2401	-4.5165	6.29E-06	0.002984
EPHA8	3.65	-2.8530	0.6342	-4.4983	6.85E-06	0.003193
HLA-DRB6	495.94	-1.5826	0.3526	-4.4890	7.16E-06	0.003278
ZNF711	444.99	-0.6442	0.1448	-4.4483	8.66E-06	0.003898
KRT14	4.11	-2.4164	0.5472	-4.4157	1.01E-05	0.004460
SCT	22.89	-1.5018	0.3410	-4.4038	1.06E-05	0.004633
SP140	86.40	-1.3981	0.3211	-4.3533	1.34E-05	0.005747
HLA-DOA	300.49	-1.4088	0.3243	-4.3442	1.40E-05	0.005896
PUDP	1197.76	-0.5231	0.1209	-4.3259	1.52E-05	0.006210

Supplementary Table S4: Top 30 most significantly upregulated genes in late-onset colon adenocarcinoma (TCGA cohort; ≥ 50 years, male vs. female comparison).

Genes enriched in male late-onset tumors, including Y-chromosome epigenetic regulators and Wnt pathway modulators.

Gene Symbol	log2FC	lfcSE	Stat	p-value	padj
USP9Y	8.2097	0.3790	21.6603	4.86E-104	9.42E-100
ZFY	9.1752	0.4277	21.4523	4.34E-102	3.85E-98
PRKY	7.1606	0.3453	20.7379	1.58E-95	1.05E-91
UTY	8.2618	0.4064	20.3277	7.31E-92	3.89E-88
KDM5D	7.9461	0.3939	20.1708	1.77E-90	7.83E-87
TXLNGY	8.4330	0.4198	20.0870	9.59E-90	3.64E-86
EIF1AY	8.7719	0.4381	20.0223	3.52E-89	1.17E-85
DDX3Y	7.8223	0.4120	18.9848	2.28E-80	6.73E-77
NLGN4Y	6.7630	0.3607	18.7491	1.97E-78	5.23E-75
RPS4Y1	8.0275	0.4580	17.5273	8.87E-69	2.14E-65
TMSB4Y	8.6304	0.5344	16.1501	1.13E-58	2.51E-55
GYG2P1	7.7477	0.5234	14.8033	1.40E-49	2.85E-46
ANOS2P	7.2346	0.4935	14.6584	1.19E-48	2.26E-45
LINC00278	7.1760	0.5222	13.7406	5.80E-43	1.03E-39
ZNF736P9Y	8.5396	0.6805	12.5500	3.98E-36	6.61E-33
BCORP1	5.6193	0.4521	12.4296	1.80E-35	2.82E-32
TTY14	5.8037	0.4732	12.2641	1.41E-34	2.08E-31
ARSDP1	5.1101	0.5310	9.6231	6.39E-22	8.93E-19
PRORY	6.7508	0.7504	8.9967	2.33E-19	3.09E-16
TTY10	3.8144	0.4251	8.9721	2.91E-19	3.68E-16
RBMV2EP	5.4709	0.8152	6.7107	1.94E-11	2.24E-08
RBMV2TP	4.7311	0.8123	5.8244	5.73E-09	5.86E-06
M1AP	2.4932	0.4390	5.6791	1.35E-08	1.28E-05
OFD1P16Y	4.4644	0.7909	5.6446	1.66E-08	1.52E-05
PRSS41	1.9650	0.3664	5.3624	8.21E-08	7.15E-05
ZFY-AS1	3.3339	0.6233	5.3486	8.86E-08	7.36E-05
GABRB3	2.3441	0.4570	5.1290	2.91E-07	0.000221
TBL1Y	5.1579	1.0186	5.0635	4.12E-07	0.000296
GRM1	2.5307	0.5054	5.0075	5.51E-07	0.000372
AFAP1-AS1	1.9011	0.3794	5.0113	5.41E-07	0.000372

Supplementary Table S5: Differentially expressed genes in early-onset vs. late-onset colon cancer — Korean population cohort (GSE213092)Analysis performed using GEO2R/limma ($|\log_2FC| > 1.5$, FDR-adjusted $p < 0.05$). Score refers to the limma moderated t-statistic.

Gene Symbol	Log2FC	Score	Direction	Functional Relevance
CPS1	4.144	7.022	Upregulated	Metabolic enzyme; nitrogen metabolism; CMS3 subtype
CXCL5	3.125	3.752	Upregulated	Inflammatory chemokine; tumor microenvironment regulation
VENTX	3.063	5.232	Upregulated	Homeobox factor; cancer cell proliferation
C20orf204	2.535	-	Upregulated	Function not fully characterized
GLDC	2.366	-	Upregulated	Glycine decarboxylase; metabolic reprogramming
IGF2	-2.742	1.504	Downregulated	IGF-1R/PI3K/Akt pathway; growth factor signaling
TCN1	-2.327	1.398	Downregulated	Vitamin B12-binding protein; GI cancer biomarker
JSRP1	-2.173	3.123	Downregulated	Cellular signaling; possible tumor progression