

SUPPLEMENTARY MATERIALS: Statistical Methods Applied to Gene Expression Data to Explore Cancer Features

HANNAH SIMPSON-CLANCY*

Project advisor: A. Smith †

Table SM1

Benjamini-Hochberg adjusted p-values for Wilcoxon Rank-Sum Test differentially expressed genes when comparing clear cell and normal ovary samples.

Gene Symbol	Gene Title	BH P-value
AGR2	anterior gradient 2 homolog (<i>Xenopus laevis</i>)	0.036943
C7	complement component 7	0.036943
TFF3	trefoil factor 3 (intestinal)	0.036943
HGD	homogentisate 1,2-dioxygenase	0.036943
STAR	steroidogenic acute regulatory protein	0.036943
HGD	homogentisate 1,2-dioxygenase	0.036943
PEG3	paternally expressed 3	0.036943
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	0.036943
HS3ST1	heparan sulfate (glucosamine) 3-O-sulfotransferase 1	0.036943
PEG3	paternally expressed 3	0.036943
GAD1	glutamate decarboxylase 1 (brain, 67kDa)	0.036943
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.036943
GABRP	gamma-aminobutyric acid (GABA) A receptor, pi	0.036943
BAMBI	BMP and activin membrane-bound inhibitor homolog (<i>Xenopus laevis</i>)	0.036943
SORBS2	sorbin and SH3 domain containing 2	0.036943
MFAP3L	microfibrillar-associated protein 3-like	0.036943
SLC7A5	solute carrier family 7 (amino acid transporter light chain, L system), member 5	0.036943
FBLN1	fibulin 1	0.036943
BCL2	B-cell CLL/lymphoma 2	0.036943
PCCA	propionyl CoA carboxylase, alpha polypeptide	0.036943

* (h.sc@hotmail.co.uk, University of Huddersfield, UK).

† (a.smith@hud.ac.uk, University of Huddersfield, UK)

Table SM2

Benjamini-Hochberg adjusted p-values for Wilcoxon Rank-Sum Test differentially expressed genes when comparing endometrioid and normal ovary samples.

Gene Symbol	Gene Title	BH P-value
STAR	steroidogenic acute regulatory protein	0.018677
PEG3	paternally expressed 3	0.018677
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	0.018677
SPDEF	SAM pointed domain containing ets transcription factor	0.018677
GAD1	glutamate decarboxylase 1 (brain, 67kDa)	0.018677
LEF1	lymphoid enhancer-binding factor 1	0.018677
TSPAN13	tetraspanin 13	0.018677
EPS8	epidermal growth factor receptor pathway substrate 8	0.018677
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	0.018677
VLDLR	very low density lipoprotein receptor	0.018677
LRRC17	leucine rich repeat containing 17	0.018677
SPDEF	SAM pointed domain containing ets transcription factor	0.018677
C4BPB	complement component 4 binding protein, beta	0.018677
GATM	glycine amidinotransferase (L-arginine:glycine amidinotransferase)	0.018677
SEC14L1	SEC14-like 1 (<i>S. cerevisiae</i>)	0.018677
SPDEF	SAM pointed domain containing ets transcription factor	0.018677
MTUS1	microtubule associated tumor suppressor 1	0.018677
GALNT10	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 (GalNAc-T10)	0.018677
HSDL2	hydroxysteroid dehydrogenase like 2	0.018677
MTUS1	microtubule associated tumor suppressor 1	0.018677

Table SM3

Benjamini-Hochberg adjusted p-values for Wilcoxon Rank-Sum Test differentially expressed genes when comparing mucinous and normal ovary samples.

Gene Symbol	Gene Title	BH P-value
AGR2	anterior gradient 2 homolog (<i>Xenopus laevis</i>)	0.028331
TFF3	trefoil factor 3 (intestinal)	0.028331
HGD	homogentisate 1,2-dioxygenase	0.028331
CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	0.028331
SPDEF	SAM pointed domain containing ets transcription factor	0.028331
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.028331
GABRP	gamma-aminobutyric acid (GABA) A receptor, pi	0.028331
BAMBI	BMP and activin membrane-bound inhibitor homolog (<i>Xenopus laevis</i>)	0.028331
FOXA2	forkhead box A2	0.028331
TMC5	transmembrane channel-like 5	0.028331
PPAP2C	phosphatidic acid phosphatase type 2C	0.028331
TSPAN13	tetraspanin 13	0.028331
TGFA	transforming growth factor, alpha	0.028331
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	0.028331
TRIM2	tripartite motif containing 2	0.028331
VLDLR	very low density lipoprotein receptor	0.028331
LRR17	leucine rich repeat containing 17	0.028331
RASL11B	RAS-like, family 11, member B	0.028331
SPDEF	SAM pointed domain containing ets transcription factor	0.028331
ATP1B1	ATPase, Na ⁺ /K ⁺ transporting, beta 1 polypeptide	0.028331

Table SM4

Benjamini-Hochberg adjusted p-values for Wilcoxon Rank-Sum Test differentially expressed genes when comparing serous and normal ovary samples.

Gene Symbol	Gene Title	BH p-value
MCPH1	microcephalin 1	0.00800
GLP1R	glucagon-like peptide 1 receptor	0.010688
C7	complement component 7	0.017239
STAR	steroidogenic acute regulatory protein	0.017239
GREB1	growth regulation by estrogen in breast cancer 1	0.017239
PEG3	paternally expressed 3	0.017239
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	0.017239
SPDEF	SAM pointed domain containing ets transcription factor	0.017239
HS3ST1	heparan sulfate (glucosamine) 3-O-sulfotransferase 1	0.017239
PEG3	paternally expressed 3	0.017239
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.017239
BAMBI	BMP and activin membrane-bound inhibitor homolog (<i>Xenopus laevis</i>)	0.017239
PLK1S1	polo-like kinase 1 substrate 1	0.017239
BCL2	B-cell CLL/lymphoma 2	0.017239
PCCA	propionyl CoA carboxylase, alpha polypeptide	0.017239
TSPAN5	tetraspanin 5	0.017239
PGR	progesterone receptor	0.017239
TSPAN13	tetraspanin 13	0.017239
CLDN5	claudin 5	0.017239
EPS8	epidermal growth factor receptor pathway substrate 8	0.017239

Table SM5*Fold change values for up-regulated genes between clear cell and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
MUC1	mucin 1, cell surface associated	345.547375
CP	ceruloplasmin (ferroxidase)	220.785119
SCNN1A	sodium channel, non-voltage-gated 1 alpha subunit	194.387111
CLDN3	claudin 3	124.101924
PAX8	paired box 8	113.379109
ABP1	amiloride binding protein 1 (amine oxidase (copper-containing))	111.791398
CD24	CD24 molecule	108.385016
CD24	CD24 molecule	96.098873
LBP	lipopolysaccharide binding protein	78.163984
CD24	CD24 molecule	76.590354
CD24	CD24 molecule	74.750921
CD24	CD24 molecule	73.868422
C1orf186 /// LOC100505650	chromosome 1 open reading frame 186 /// uncharacterized LOC100505650	73.420049
LCN2	lipocalin 2	71.455738
TMPRSS4	transmembrane protease, serine 4	68.531493
HGD	homogentisate 1,2-dioxygenase	66.711632
CLDN3	claudin 3	64.864276
SCGB2A1	secretoglobin, family 2A, member 1	58.424752
SLPI	secretory leukocyte peptidase inhibitor	49.634608
KRT7	keratin 7	48.502661

Table SM6*Fold change values for down-regulated genes between clear cell and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
MAOB	monoamine oxidase B	0.021368
OGN	osteoglycin	0.024556
GAS1	growth arrest-specific 1	0.025406
FOSB	FBJ murine osteosarcoma viral oncogene homolog B	0.027921
GULP1	GULP, engulfment adaptor PTB domain containing 1	0.028215
ITM2A	integral membrane protein 2A	0.028345
ADIRF	adipogenesis regulatory factor	0.035391
ZFPM2	zinc finger protein, FOG family member 2	0.038154
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.039438
STAR	steroidogenic acute regulatory protein	0.042516
RGS2	regulator of G-protein signaling 2, 24kDa	0.043330
NR4A2	nuclear receptor subfamily 4, group A, member 2	0.044453
RNASE4	ribonuclease, RNase A family, 4	0.044537
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	0.045151
PEG10	paternally expressed 10	0.045856
PLCL2	phospholipase C-like 2	0.045899
ANG	angiogenin, ribonuclease, RNase A family, 5	0.046887
LRRC17	leucine rich repeat containing 17	0.047032
DCN	decorin	0.047452
PELI2	pellino E3 ubiquitin protein ligase family member 2	0.048919

Table SM7

Fold change values for up-regulated genes between endometrioid and normal ovary samples.

Gene Symbol	Gene Title	Fold Change
MUC1	mucin 1, cell surface associated	128.292396
SCGB2A1	secretoglobin, family 2A, member 1	117.569576
SCNN1A	sodium channel, non-voltage-gated 1 alpha subunit	85.925992
CLDN3	claudin 3	84.878084
CD24	CD24 molecule	69.593745
CD24	CD24 molecule	63.928613
CD24	CD24 molecule	49.128451
CD24	CD24 molecule	45.717463
CLDN3	claudin 3	42.308865
CD24	CD24 molecule	40.628392
LCN2	lipocalin 2	40.466903
CP	ceruloplasmin (ferroxidase)	40.246521
EPCAM	epithelial cell adhesion molecule	37.710964
SCGB1D2	secretoglobin, family 1D, member 2	36.489278
ESRP1	epithelial splicing regulatory protein 1	34.343929
C1orf186 /// LOC100505650	chromosome 1 open reading frame 186 /// uncharacterized LOC100505650	31.767728
PAX8	paired box 8	30.825799
WFDC2	WAP four-disulfide core domain 2	29.944227
SPINT1	serine peptidase inhibitor, Kunitz type 1	29.849428
CLDN4 /// LOC100996451	claudin 4 /// uncharacterized LOC100996451	29.001255

Table SM8*Fold change values for down-regulated genes between endometrioid and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
STAR	steroidogenic acute regulatory protein	0.026555
FOSB	FBJ murine osteosarcoma viral oncogene homolog B	0.035875
OGN	osteoglycin	0.037653
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	0.040841
ITM2A	integral membrane protein 2A	0.047842
ZFPM2	zinc finger protein, FOG family member 2	0.048785
C4BPB	complement component 4 binding protein, beta	0.053106
MYH11	myosin, heavy chain 11, smooth muscle	0.055826
TCF21	transcription factor 21	0.055952
NR4A3	nuclear receptor subfamily 4, group A, member 3	0.057630
HLF	hepatic leukemia factor	0.059804
ADAMTS5	ADAM metalloproteinase with thrombospondin type 1 motif, 5	0.060998
EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	0.061294
RNASE4	ribonuclease, RNase A family, 4	0.062585
ADIRF	adipogenesis regulatory factor	0.063528
DUSP1	dual specificity phosphatase 1	0.064079
PLCL2	phospholipase C-like 2	0.064239
DCN	decorin	0.065443
MEF2C	myocyte enhancer factor 2C	0.0658099
WISP2	WNT1 inducible signaling pathway protein 2	0.065863

Table SM9*Fold change values for up-regulated genes between mucinous and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
MUC1	mucin 1, cell surface associated	278.478216
AGR2	anterior gradient 2 homolog (<i>Xenopus laevis</i>)	199.397012
SCNN1A	sodium channel, non-voltage-gated 1 alpha subunit	95.387518
CD24	CD24 molecule	78.977298
CD24	CD24 molecule	71.912455
CEACAM6	carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen)	69.527673
LCN2	lipocalin 2	67.150053
TMPRSS4	transmembrane protease, serine 4	57.929675
CD24	CD24 molecule	55.486734
CD24	CD24 molecule	52.413173
SPINT1	serine peptidase inhibitor, Kunitz type 1	49.2678696
CLDN4 /// LOC100996451	claudin 4 /// uncharacterized LOC100996451	45.134678
FXYD3	FXYD domain containing ion transport regulator 3	41.730092
CD24	CD24 molecule	39.911115
EPCAM	epithelial cell adhesion molecule	39.189038
TFF2	trefoil factor 2	38.668977
ST14	suppression of tumorigenicity 14 (colon carcinoma)	38.382493
SPINK1	serine peptidase inhibitor, Kazal type 1	38.095799
SLC44A4	solute carrier family 44, member 4	37.228535
S100P	S100 calcium binding protein P	31.6953699

Table SM10*Fold change values for down-regulated genes between mucinous and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	0.050075
TRO	trophinin	0.050915
FOSB	FBJ murine osteosarcoma viral oncogene homolog B	0.051288
HLF	hepatic leukemia factor	0.056626
ZFPM2	zinc finger protein, FOG family member 2	0.057666
EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	0.058127
ITM2A	integral membrane protein 2A	0.059541
PRELP	proline/arginine-rich end leucine-rich repeat protein	0.061077
PEG10	paternally expressed 10	0.062009
DCN	decorin	0.062124
TCEAL2	transcription elongation factor A (SII)-like 2	0.063304
WT1	Wilms tumor 1	0.065719
NR4A3	nuclear receptor subfamily 4, group A, member 3	0.066079
ADAMTS5	ADAM metallopeptidase with thrombospondin type 1 motif, 5	0.068675
STAR	steroidogenic acute regulatory protein	0.070345
SPRR2B	small proline-rich protein 2B	0.070688
DUSP1	dual specificity phosphatase 1	0.071461
SLC4A3	solute carrier family 4, anion exchanger, member 3	0.073312
CYBRD1	cytochrome b reductase 1	0.073491
AOX1	aldehyde oxidase 1	0.075156

Table SM11*Fold change values for up-regulated genes between serous and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
MUC1	mucin 1, cell surface associated	224.447706
SCNN1A	sodium channel, non-voltage-gated 1 alpha subunit	124.103270
CLDN3	claudin 3	112.848532
MSLN	mesothelin	93.853565
CP	ceruloplasmin (ferroxidase)	80.656258
CD24	CD24 molecule	73.084404
KLK8	kallikrein-related peptidase 8	68.1235298
CD24	CD24 molecule	67.713169
CLDN3	claudin 3	54.096829
CD24	CD24 molecule	51.758286
PAX8	paired box 8	49.358683
FOLR1	folate receptor 1 (adult)	46.635936
CD24	CD24 molecule	44.934352
SCGB2A1	secretoglobin, family 2A, member 1	44.242098
ESRP1	epithelial splicing regulatory protein 1	43.056590
CD24	CD24 molecule	42.849422
CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	39.706740
MUC16	mucin 16, cell surface associated	37.139702
C1orf186 /// LOC100505650	chromosome 1 open reading frame 186 /// uncharacterized LOC100505650	33.526147
EPCAM	epithelial cell adhesion molecule	33.063335

Table SM12*Fold change values for down-regulated genes between serous and normal ovary samples.*

Gene Symbol	Gene Title	Fold Change
STAR	steroidogenic acute regulatory protein	0.0087998
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	0.022513
C4BPB	complement component 4 binding protein, beta	0.030523
OGN	osteoglycin	0.030549
FOSB	FBJ murine osteosarcoma viral oncogene homolog B	0.033441
C7	complement component 7	0.033568
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.037384
TCF21	transcription factor 21	0.039547
PEG3	paternally expressed 3	0.040997
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	0.046214
ITM2A	integral membrane protein 2A	0.046700
MYH11	myosin, heavy chain 11, smooth muscle	0.048582
HLF	hepatic leukemia factor	0.051466
MAOB	monoamine oxidase B	0.051643
MEF2C	myocyte enhancer factor 2C	0.052421
SPRR2B	small proline-rich protein 2B	0.053793
PTPRN2	protein tyrosine phosphatase, receptor type, N polypeptide 2	0.055693
RNASE4	ribonuclease, RNase A family, 4	0.0562195
GATM	glycine amidinotransferase (L-arginine:glycine amidinotransferase)	0.057704
LRRC17	leucine rich repeat containing 17	0.058612

Table SM13

Benjamini-Hochberg adjusted p-values for Kruskal-Wallis test comparing all ovarian cancer subtypes and normal ovaries.

Gene Symbol	Gene Title	BH p-value
EPS8	epidermal growth factor receptor pathway substrate 8	5.39845e – 07
AGR2	anterior gradient 2 homolog (<i>Xenopus laevis</i>)	7.34088e – 07
TSPAN1	tetraspanin 1	7.34088e – 07
PEA15	phosphoprotein enriched in astrocytes 15	7.34088e – 07
TFF3	trefoil factor 3 (intestinal)	9.62937e – 07
PEA15	phosphoprotein enriched in astrocytes 15	9.62937e – 07
CLDN16	claudin 16	9.62937e – 07
TTC38	tetratricopeptide repeat domain 38	1.01412e – 06
ACADS	acyl-CoA dehydrogenase, C-2 to C-3 short chain	1.15680e – 06
SLC3A1	solute carrier family 3 (cystine, dibasic and neutral amino acid transporters, activator of cystine, dibasic and neutral amino acid transport), member 1	1.16461e – 06
KLK5	kallikrein-related peptidase 5	1.16461e – 06
GAS1	growth arrest-specific 1	1.16461e – 06
PPAP2A	phosphatidic acid phosphatase type 2A	1.22751e – 06
PNOC	prepronociceptin	1.22751e – 06
CRABP2	cellular retinoic acid binding protein 2	1.37406e – 06
IQGAP2	IQ motif containing GTPase activating protein 2	1.38829e – 06
BCAM	basal cell adhesion molecule (Lutheran blood group)	1.38829e – 06
ARL1	ADP-ribosylation factor-like 15	1.40938e – 06
USH1C	Usher syndrome 1C (autosomal recessive, severe)	1.40938e – 06
CYP2C9	cytochrome P450, family 2, subfamily C, polypeptide 9	1.40938e – 06

Table SM14*Mean expression levels for Kruskal-Wallis top 20 genes*

Number in Dendrogram	Gene Name	Clear Cell	Endometrioid	Mucinous	Serous	Control
1	EPS8	11.380134	10.653133	12.052507	9.768083	12.396728
2	AGR2	12.084012	9.817472	14.327403	7.909813	6.687903
3	TSPAN1	13.475372	11.465159	13.345123	10.902714	9.528675
4	PEA15	10.405219	11.452433	10.834590	12.278874	10.357018
5	TFF3	7.758476	10.102102	13.450931	8.350407	9.295564
6	PEA15	11.388933	12.460031	11.862639	13.083689	11.955041
7	CLDN16	5.661044	6.581883	5.643856	9.499926	6.074910
8	TTC38	10.221083	9.490952	10.389877	8.956700	10.077230
9	ACADS	9.839720	9.167791	9.660317	8.898366	8.496776
10	SLC3A1	12.680717	8.432246	11.419436	8.249907	8.901485
11	KLK5	8.556565	9.285578	8.473027	11.095777	8.290785
12	GAS1	6.808101	8.837596	8.518631	11.273543	12.106772
13	PPAP2A	10.993033	9.869926	10.779095	9.605166	10.303855
14	PNOC	8.980833	8.965828	8.423116	10.817950	8.024756
15	CRABP2	8.611928	10.948103	9.055343	12.203687	8.651737
16	IQGAP2	8.802383	9.742999	11.415167	9.087379	10.503079
17	BCAM	9.963625	9.198137	8.861498	10.717812	7.763364
18	ARL1	11.575901	11.088639	11.473833	10.809561	12.079162
19	USH1C	10.285553	7.099842	10.213607	7.030654	6.135021
20	CYP2C9	11.279622	8.945123	10.354090	8.842076	8.502452