

Supplementary Information

Gene expression in the urinary bladder: a common carcinoma in-situ gene expression signature exists disregarding histopathological classification

Lars Dyrskjøt, Mogens Kruhøffer, Thomas Thykjaer, Niels Marcussen, Jens L. Jensen, Klaus Møller and Torben F. Ørntoft

Contents:

Table 1 – The best 100 markers.....	2
Table 2 – The 16 gene molecular classifier of CIS	8
Figure 1 – Array normalization	9
Figure 2 – Cross-validation performance using all samples	10
Figure 3 – Cross-validation performance using half of the samples.....	11

Table 1 – The best 100 markers

Feature	Class	T-test	Log fold change	Perm 1%	Perm 5%	Perm 10%	Gene Symbol	Description
221204_s_at	no_CIS	3.74	-2.4	5.12	4.61	4.33	CRTAC1	cartilage acidic protein 1
205927_s_at	no_CIS	3.67	-2.5	4.53	3.98	3.73	CTSE	cathepsin E
210143_at	no_CIS	3.35	-2.8	4.03	3.73	3.45	ANXA10	annexin A10
204540_at	no_CIS	3.15	-1.5	3.87	3.51	3.32	EEF1A2	eukaryotic translation elongation factor 1 alpha 2
214599_at	no_CIS	3.02	-1.9	3.75	3.37	3.14	IVL	involucrin
203649_s_at	no_CIS	2.84	-2.0	3.63	3.20	3.00	PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)
203980_at	no_CIS	2.74	-2.0	3.47	3.12	2.89	FABP4	fatty acid binding protein 4, adipocyte
209270_at	no_CIS	2.39	-1.4	3.38	3.10	2.85	LAMB3	laminin, beta 3
206658_at	no_CIS	2.35	-1.5	3.37	3.05	2.78	UPK3B	uroplakin 3b
220779_at	no_CIS	2.35	-1.6	3.33	2.97	2.73	PADI3	peptidyl arginine deiminase, type III
216971_s_at	no_CIS	2.28	-1.1	3.29	2.91	2.71	PLEC1	plectin 1, intermediate filament binding protein 500kDa
206191_at	no_CIS	2.25	-1.4	3.24	2.86	2.68	ENTPD3	ectonucleoside triphosphate diphosphohydrolase 3
218484_at	no_CIS	2.18	-1.2	3.20	2.81	2.62	LOC56901	NADH:ubiquinone oxidoreductase MLRQ subunit homolog
221854_at	no_CIS	2.1	-1.1	3.19	2.80	2.60	---	ESTs, Highly similar to plakophilin 1; Plakophilin-1 [Homo sapiens] [H.sapiens]
203792_x_at	no_CIS	2.02	-1.2	3.16	2.74	2.55	ZNF144	zinc finger protein 144 (Mel-18)
207862_at	no_CIS	2.01	-1.2	3.16	2.72	2.52	UPK2	uroplakin 2
218960_at	no_CIS	1.93	-1.3	3.14	2.65	2.47	TMPRSS4	transmembrane protease, serine 4
203009_at	no_CIS	1.93	-1.0	3.12	2.62	2.45	LU	Lutheran blood group (Auberger b antigen included)
204508_s_at	no_CIS	1.88	-1.0	3.10	2.60	2.42	FLJ20151	hypothetical protein FLJ20151
211692_s_at	no_CIS	1.87	-1.0	3.06	2.58	2.39	BBC3	BCL2 binding component 3
206465_at	no_CIS	1.86	-0.8	3.04	2.54	2.38	BG1	lipidosin
206122_at	no_CIS	1.85	-1.0	2.92	2.52	2.36	SOX15	SRY (sex determining region Y)-box 15
206393_at	no_CIS	1.83	-1.0	2.89	2.49	2.33	TNNI2	troponin I, skeletal, fast

214639_s_at	no_CIS	1.79	-0.9	2.87	2.49	2.30	HOXA1	homeo box A1
214630_at	no_CIS	1.79	-0.9	2.84	2.44	2.28	CYP11B2	cytochrome P450, subfamily XIB (steroid 11-beta-hydroxylase), polypeptide 2
204465_s_at	no_CIS	1.77	-1.1	2.81	2.42	2.27	INA	internexin neuronal intermediate filament protein, alpha
204990_s_at	no_CIS	1.76	-1.2	2.79	2.41	2.24	ITGB4	integrin, beta 4
205453_at	no_CIS	1.75	-1.0	2.77	2.39	2.22	HOXB2	homeo box B2
215812_s_at	no_CIS	1.74	-1.1	2.77	2.37	2.20	---	ESTs, Highly similar to S6AA_HUMAN Sodium- and chloride-dependent creatine transporter 2 (CT2) [H.sapiens]
217040_x_at	no_CIS	1.74	-1.0	2.75	2.36	2.18	SOX15	SRY (sex determining region Y)-box 15
203759_at	no_CIS	1.73	-0.9	2.75	2.34	2.17	SIAT4C	sialyltransferase 4C (beta-galactoside alpha-2,3-sialyltransferase)
211002_s_at	no_CIS	1.73	-1.0	2.74	2.33	2.17	TRIM29	tripartite motif-containing 29
216641_s_at	no_CIS	1.73	-1.1	2.73	2.31	2.15	LAD1	ladinin 1
221660_at	no_CIS	1.71	-0.9	2.67	2.30	2.13	LOC93408	hypothetical protein BC002778
220026_at	no_CIS	1.71	-1.1	2.66	2.28	2.13	CLCA4	chloride channel, calcium activated, family member 4
209591_s_at	no_CIS	1.69	-0.8	2.63	2.28	2.11	BMP7	bone morphogenetic protein 7 (osteogenic protein 1)
219922_s_at	no_CIS	1.68	-0.9	2.61	2.26	2.08	LTBP3	latent transforming growth factor beta binding protein 3
201641_at	no_CIS	1.67	-1.1	2.61	2.26	2.07	BST2	bone marrow stromal cell antigen 2
204952_at	no_CIS	1.66	-1.1	2.59	2.24	2.07	C4.4A	GPI-anchored metastasis-associated protein homolog
204487_s_at	no_CIS	1.65	-0.9	2.59	2.23	2.06	KCNQ1	potassium voltage-gated channel, KQT-like subfamily, member 1
210761_s_at	no_CIS	1.64	-0.9	2.59	2.23	2.05	GRB7	growth factor receptor-bound protein 7
217626_at	no_CIS	1.63	-0.9	2.58	2.21	2.04	---	ESTs, Highly similar to A53436 3-alpha-hydroxysteroid/dihydrodiol dehydrogenase (EC 1.1.1.-) - human [H.sapiens]

204380_s_at	no_CIS	1.62	-0.9	2.58	2.19	2.03	FGFR3	fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism)
205455_at	no_CIS	1.61	-0.9	2.58	2.17	2.02	MST1R	macrophage stimulating 1 receptor (c-met-related tyrosine kinase)
205073_at	no_CIS	1.61	-0.9	2.58	2.17	2.01	CYP2J2	cytochrome P450, subfamily IJ (arachidonic acid epoxygenase)
203287_at	no_CIS	1.61	-1.0	2.58	2.16	2.00	LAD1	polypeptide 2 ladinin 1
210735_s_at	no_CIS	1.58	-1.0	2.55	2.15	1.99	CA12	carbonic anhydrase XII
203842_s_at	no_CIS	1.57	-0.8	2.54	2.15	1.97	MAPRE3	microtubule-associated protein, RP/EB family, member 3
206561_s_at	no_CIS	1.57	-1.0	2.53	2.14	1.96	AKR1B10	aldo-keto reductase family 1, member B10 (aldose reductase)
214752_x_at	no_CIS	1.56	-0.9	2.52	2.13	1.95	FLNA	filamin A, alpha (actin binding protein 280)
217028_at	CIS	4.87	2.4	5.17	4.67	4.40	CXCR4	chemokine (C-X-C motif) receptor 4
213975_s_at	CIS	4.65	2.6	4.43	4.01	3.76	LYZ	lysozyme (renal amyloidosis)
201859_at	CIS	4.59	2.7	4.15	3.70	3.45	PRG1	proteoglycan 1, secretory granule
219410_at	CIS	4.49	2.2	3.98	3.49	3.29	FLJ10134	hypothetical protein FLJ10134
207173_x_at	CIS	4.37	2.4	3.88	3.33	3.11	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)
214651_s_at	CIS	4.14	1.9	3.83	3.22	2.99	HOXA9	homeo box A9
201858_s_at	CIS	4.06	2.5	3.78	3.09	2.91	PRG1	proteoglycan 1, secretory granule
211430_s_at	CIS	4.03	3.3	3.63	3.05	2.83	IGHG3	immunoglobulin heavy constant gamma 3 (G3m marker)
213891_s_at	CIS	3.86	1.9	3.63	3.02	2.77	---	Homo sapiens cDNA FLJ11918 fis, clone HEMBB1000272, mRNA sequence
221872_at	CIS	3.82	2.4	3.52	2.89	2.73	RARRES1	retinoic acid receptor responder (tazarotene induced) 1
212386_at	CIS	3.77	2.1	3.50	2.87	2.69	---	Homo sapiens cDNA FLJ11918 fis, clone

								HEMBB1000272, mRNA sequence
211161_s_at	CIS	3.76	2.1	3.42	2.84	2.65	---	PRO3121 [Homo sapiens], mRNA sequence
214669_x_at	CIS	3.55	2.3	3.36	2.80	2.62	IGKC	immunoglobulin kappa constant
217388_s_at	CIS	3.44	1.8	3.31	2.79	2.58	KYNU	kynureninase (L- kynurenine hydrolase)
203477_at	CIS	3.36	1.8	3.28	2.75	2.56	COL15A1	collagen, type XV, alpha 1
204688_at	CIS	3.35	1.7	3.26	2.74	2.52	SGCE	sarcoglycan, epsilon
218718_at	CIS	3.35	1.7	3.22	2.70	2.48	PDGFC	platelet derived growth factor C
215176_x_at	CIS	3.32	2.3	3.14	2.67	2.45	IGKC	immunoglobulin kappa constant
201842_s_at	CIS	3.31	2.0	3.11	2.65	2.44	EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1
212667_at	CIS	3.3	1.9	3.11	2.63	2.42	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)
209340_at	CIS	3.27	1.8	3.10	2.61	2.39	UAP1	UDP-N- acetylglucosamine pyrophosphorylase 1
215379_x_at	CIS	3.26	2.2	3.10	2.59	2.39	IGL@	immunoglobulin lambda locus
200762_at	CIS	3.25	1.8	3.05	2.56	2.34	DPYSL2	dihydropyrimidinase- like 2
211896_s_at	CIS	3.21	2.3	3.05	2.53	2.32	DCN	decorin
204141_at	CIS	3.19	1.7	3.05	2.53	2.28	TUBB	tubulin, beta polypeptide
201744_s_at	CIS	3.18	2.0	3.03	2.50	2.27	LUM	lumican
209138_x_at	CIS	3.17	2.4	3.03	2.47	2.24	IGLJ3	immunoglobulin lambda joining 3
214677_x_at	CIS	3.14	2.7	3.02	2.47	2.23	IGLJ3	immunoglobulin lambda joining 3
212077_at	CIS	3.11	1.9	2.99	2.46	2.21	---	Homo sapiens, clone IMAGE:4296901, mRNA, mRNA sequence
206392_s_at	CIS	3.11	2.0	2.98	2.43	2.20	RARRES1	retinoic acid receptor responder (tazarotene induced) 1
212998_x_at	CIS	3.09	1.6	2.94	2.40	2.19	HLA-DQB1	major histocompatibility complex, class II, DQ beta 1
201616_s_at	CIS	3.08	1.5	2.93	2.38	2.18	CALD1	caldesmon 1
205382_s_at	CIS	3.07	1.7	2.88	2.37	2.15	DF	D component of complement (adipsin)
212671_s_at	CIS	3.07	1.8	2.85	2.35	2.14	HLA-DQA1	major histocompatibility complex, class II, DQ

								alpha 1
215121_x_at	CIS	3.06	2.1	2.84	2.34	2.13	IGLJ3	immunoglobulin lambda joining 3
200600_at	CIS	3.05	1.7	2.83	2.33	2.11	MSN	moesin
202746_at	CIS	3.03	1.5	2.80	2.32	2.10	ITM2A	integral membrane protein 2A
202917_s_at	CIS	3	2.5	2.79	2.31	2.08	S100A8	S100 calcium binding protein A8 (calgranulin A)
201560_at	CIS	3	1.6	2.79	2.30	2.08	CLIC4	chloride intracellular channel 4
218918_at	CIS	2.99	1.7	2.77	2.29	2.06	MAN1C1	mannosidase, alpha, class 1C, member 1
218656_s_at	CIS	2.99	1.4	2.76	2.27	2.06	LHFP	lipoma HMGIC fusion partner
201088_at	CIS	2.99	1.8	2.76	2.26	2.04	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)
201291_s_at	CIS	2.97	1.4	2.75	2.25	2.04	TOP2A	topoisomerase (DNA) II alpha 170kDa
215076_s_at	CIS	2.95	2.1	2.72	2.24	2.03	COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
212195_at	CIS	2.94	1.7	2.71	2.22	2.02	---	Homo sapiens mRNA; cDNA DKFZp564F053 (from clone DKFZp564F053), mRNA sequence
209732_at	CIS	2.94	1.6	2.68	2.22	2.00	CLECSF2	C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 2 (activation-induced)
212192_at	CIS	2.94	1.6	2.67	2.22	1.99	LOC115207	hypothetical protein BC013764
221671_x_at	CIS	2.92	2.4	2.67	2.20	1.98	IGKC	immunoglobulin kappa constant
211671_s_at	CIS	2.91	1.4	2.66	2.20	1.98	NR3C1	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
214352_s_at	CIS	2.88	1.8	2.66	2.19	1.97	KRAS2	v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog

Feature: Probe-set on U133A GeneChip

Class: The group in which the marker is up-regulated

T-test: The t-test value

Log fold change: The log₂ to the ratio of the median expression in the groups (CIS/non-CIS).

Perm 1%: The 1% permutation level

Perm 5%: The 5% permutation level

Perm 10%: The 10% permutation level

Table 2 – The 16 gene molecular classifier of CIS

Feature	Class	t-test	Log fold change	Perm 1%	Perm 5%	Perm 10%	Gene Symbol	Description
213633_at	no_CIS	1.51	-0.8	2.46	2.04	1.85	SH3BP1	SH3-domain binding protein 1
212784_at	no_CIS	1.36	-0.8	2.27	1.86	1.70	CIC	capicua homolog (Drosophila)
209241_x_at	no_CIS	1.13	-0.6	1.78	1.48	1.33	MINK	misshapen/NIK-related kinase
217941_s_at	CIS	2.3	1.3	1.96	1.66	1.47	ERBB2IP	erb2 interacting protein
201877_s_at	CIS	2.27	1.3	1.90	1.62	1.45	PPP2R5C	protein phosphatase 2, regulatory subunit B (B56), gamma isoform
209630_s_at	CIS	1.97	1.1	1.54	1.31	1.15	---	Homo sapiens cDNA FLJ38058 fis, clone CTONG2014898, mRNA sequence
202777_at	CIS	1.93	1.1	1.51	1.29	1.12	SHOC2	soc-2 suppressor of clear homolog (C. elegans)
200958_s_at	CIS	1.92	1.3	1.49	1.28	1.11	SDCBP	syndecan binding protein (syntenin)
209579_s_at	CIS	1.79	1.0	1.36	1.16	1.01	MBD4	methyl-CpG binding domain protein 4
209004_s_at	CIS	1.63	0.9	1.21	1.00	0.89	FBXL5	F-box and leucine-rich repeat protein 5
218150_at	CIS	1.6	1.0	1.18	0.98	0.86	ARL5	ADP-ribosylation factor-like 5
202076_at	CIS	1.53	0.9	1.12	0.92	0.82	BIRC2	baculoviral IAP repeat-containing 2
204640_s_at	CIS	1.45	0.9	1.03	0.83	0.75	SPOP	speckle-type POZ protein
201887_at	CIS	1.32	0.8	0.92	0.74	0.66	IL13RA1	interleukin 13 receptor, alpha 1
212802_s_at	CIS	1.31	0.7	0.91	0.72	0.65	DKFZP434C212	DKFZP434C212 protein
212899_at	CIS	1.29	0.7	0.89	0.71	0.64	KIAA1028	KIAA1028 protein

Feature: Probe-set on U133A GeneChip

Class: The group in which the marker is up-regulated

T-test: The t-test value

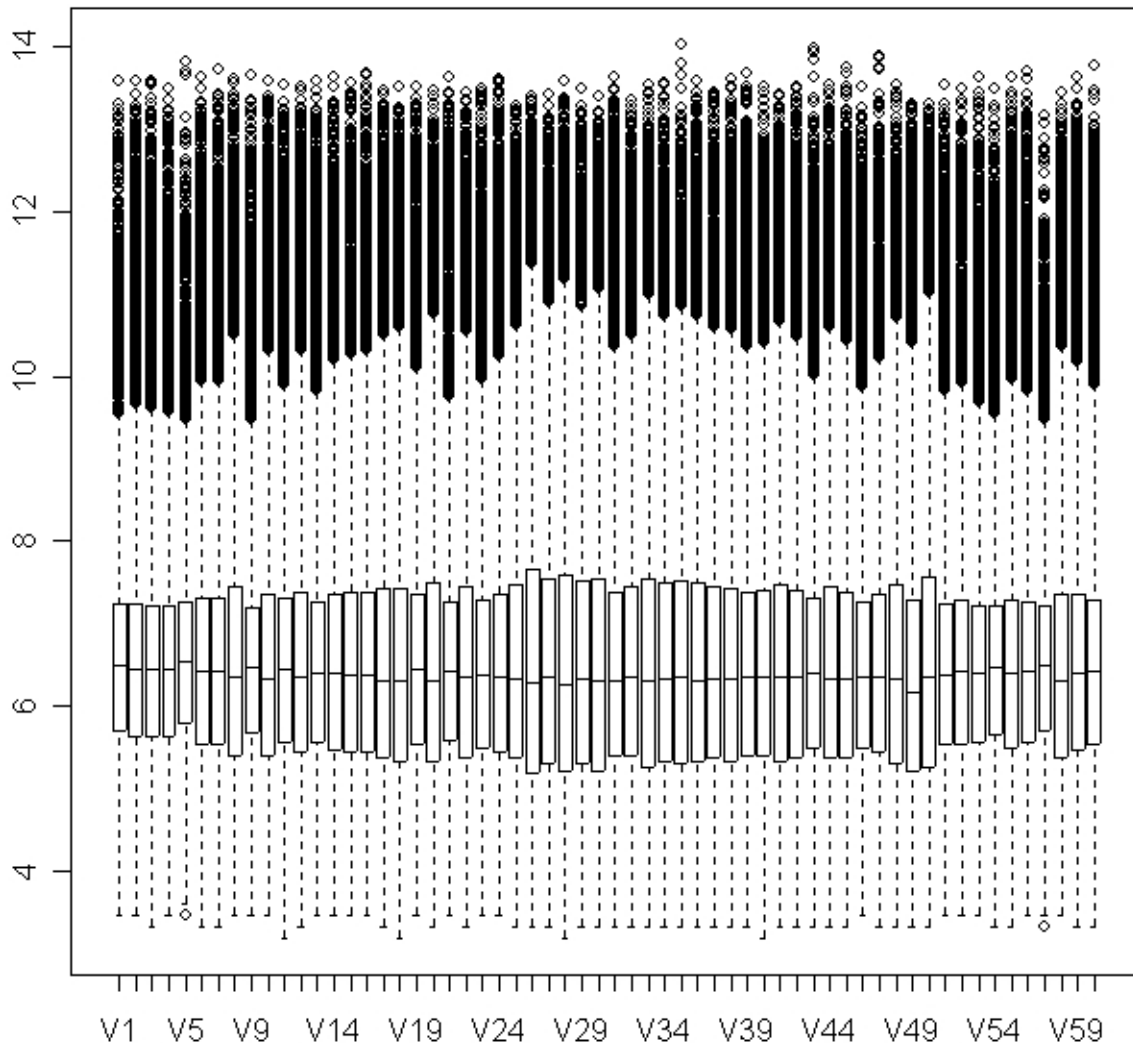
Log fold change: The log₂ to the ratio of the median expression in the groups (CIS/non-CIS).

Perm 1%: The 1% permutation level

Perm 5%: The 5% permutation level

Perm 10%: The 10% permutation level

Figure 1 – Array normalization



Box plot of the log₂ normalised microarray results. The plot illustrates the successful normalisation of the microarray experiments. The median values (line inside box) are very similar and the distribution of the data is acceptable. V1-V9 represents the normal samples, V10-V25 represents the sTCC-CIS samples, V26-V39 represents the sTCC+CIS samples, V40-V49 represents the cystectomy biopsies, and V50-V60 represents the mTCC samples.

Figure 2 – Cross-validation performance using all samples

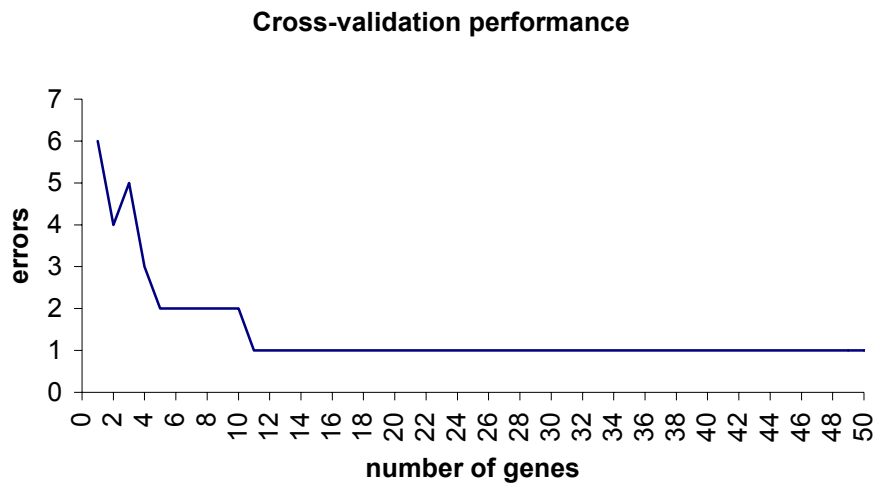


Figure 3 – Cross-validation performance using half of the samples

