Bioinformatics identification of C6orf105 as a putative protein that co-regulates with TFPI

C6orf105 was identified as potentially co-expressing with *TFPI* by a novel bioinformatics approach that entails a global meta-analysis (GAMMA) of microarray data[1], which was also validated in a recent publication[2]. A total of 3,551 human 2-color microarray datasets were downloaded for analysis from the GEO repository.

GAMMA first identified genes co-expressed with *TFPI* and ranked them by how consistently and frequently they were co-expressed in the same direction as *TFPI* (i.e., up-regulated and down-regulated together), weighing the relevance of each gene-gene association with a combination of mutual information (i.e., to account for the differences in the relative frequency by which genes are differentially expressed), the total number of times each gene pair was co-expressed and the "purity" of the pattern by which they were co-expressed (i.e., if a gene pair always change their expression in the same direction, then the "purity" of their pattern is 100%). Table S1 presents the GAMMA results for *TFPI*. Notably, 63 out of the top 100 genes that are predicted by GAMMA to be co-expressed with *TFPI* have Literature Strength (see Table S1 caption). Based on the relative frequency of each term (gene) in MEDLINE, this association was significant ($P < .01$).

Genes already known in the literature to be associated with *TFPI,* and which also appeared in the GAMMA prediction, served as positive controls: trypsin, thrombin, fibrinogen, caveolin-1, vasorin, TFPI-2, PAI-1 and -2, filamin C, glypican 3, adrenomedullin, u-PA, t-PA, thrombospondin, to name a few.

From among the five top co-expressed genes, we chose for experimentation the two genes that were both unknown and had predicted membrane location, namely *KIAA 1644* (putative protein Q3SXP7) and *C6orf105*. We gave priority to the latter because the prediction of it being an androgen-inducible gene coincided with the effects we observed experimentally for the enhancement of TFPI expression and function by androgen.

Once *C6orf105* was chosen, GAMMA was repeated to verify its co-expression patterns and try to identify possible associations. Of note, 23 out of the top 50 genes co-expressed with *TFPI* (Table S1, blue-colored genes) were also co-expressed with *C6orf105*.

After identifying and ranking the genes co-expressed with *C6orf105*, the top 20 were sent for a literature-based “guilt-by-association” analysis of the published associations they had in common, to predict their functions, phenotypes and disease relevance. This was achieved with the IRIDESCENT software, which uses a term thesaurus to recognize when biological “objects” (e.g., genes, diseases, phenotypes, chemicals, etc.) co-occur within a MEDLINE abstract[3-5]. The thesaurus is constructed using popular and freely available sources for object names (e.g. Entrez gene, OMIM, ChemID database, etc.)[4]. At the time of the study, the IRIDESCENT database of associations was constructed using 18,438,436 MEDLINE records, 55% of which had abstracts with the rest only having titles.

This analysis revealed that 8 out of the top 20 genes consistently co-expressed with C6orf105 have also been associated with TFPI in the literature (Table S2).

Table S3 presents predicted associations for TFPI based on co-expression commonalities, with the GAMMA predictions that coincide with known associations (Literature Strength) colored blue, and indicating where same associations are also predicted for C6orf105 (Yes). Most predicted associations relate to haemostasis, fibrinolysis, cancer, sepsis, lipid metabolism, and acute-phase response and inflammation. One out of three predicted associations are common for both TFPI and C6orf105.

Predicted sequences and domains of ADTRP

Fig. S1 displays the predicted sequences of the two potential alternatively spliced forms of ADTRP, of 230 aa and 248 aa, with molecular weights of 27, respectively 29 kDa. The second isoform contains an additional stretch of aminoacids between positions 51 and 68 (orange).

Several combinations of transmembrane domains could be predicted, depending on the software employed, and they are marked on the sequences.

Two predicted palmitoylation sites are possible at Cys_7 and Cys_{62} or Cys_{79} (red). There are no N-glycosylation sites, but several O-glycosylation sites possible.

Phosphorylation could occur at $Ser_{148/165}$ and $Ser_{151/168}$, as well as $Thr_{70/87}$ and $Thr_{80/97}$ and $Tyr_{201/218}$ (blue).

Androgen effect on ADTRP and TFPI mRNA expression

Data mining of GEO array profiles revealed significant up-regulation of both ADTRP and TFPI mRNA expression in the LNCaP prostate cancer cell line after incubation with DHT (Fig. S2).

REFERENCES

1. Wren JD. A global meta-analysis of microarray expression data to predict unknown gene functions and estimate the literature-data divide. *Bioinformatics*. 2009;25(13):1694-1701.
2. Daum JR, Wren JD, Daniel JJ, et al. Ska3 is required for spindle checkpoint silencing and the maintenance of chromosome cohesion in mitosis. *Curr Biol*. 2009;19(17):1467-1472.
3. Wren JD. Extending the mutual information measure to rank inferred literature relationships. *BMC Bioinformatics*. 2004;5:145.
4. Wren JD, Bekeredjian R, Stewart JA, Shohet RV, Garner HR. Knowledge discovery by automated identification and ranking of implicit relationships. *Bioinformatics*. 2004;20(3):389-398.
5. Wren JD, Garner HR. Shared relationship analysis: ranking set cohesion and commonalities within a literature-derived relationship network. *Bioinformatics*. 2004;20(2):191-198.

Table S1. GAMMA results for TFPI (Gene 1)

Blue: genes co-expressed with TFPI that are also co-expressed with ADTRP

Purity: The fraction of times gene 1 and gene 2 are co-expressed in the same direction (100% = always)
 # CoX: # of times gene 1 (TFPI) is seen co-expressed with gene 2 in microarray experiments
Lit Str (Literature strength): Cumulative tally of the number of times TFPI was mentioned with the term in MEDLINE (0.5 for every abstract co-occurrence, 0.8 per sentence); **highlighted** for the top 100 genes co-expressed with TFPI
dMIM: Mutual information (measure of specificity, >2 is significant)
Score: Prioritization of genes that correlates with the probability they share similar or identical functions as gene 1 (TFPI)
Behavior: The direction of co-expression (para=parallel, or in the same direction, anti=anti-parallel or opposite directions)

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
TFPI	Tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	Parallel	100%	207	30	0	15.72	15.72	Extracellular, Cytoplasm, Plasma membrane	SP, KUNITZ, KUNITZ, KUNITZ
KIAA1644	KIAA1644 protein	Parallel	100%	57	2	0	6.90	6.90	?	SP, TM
NCOA2	Nuclear receptor coactivator 2	Parallel	100%	56	2	1	6.84	6.84	Nucleus	HLH, PAS
CBR1	Carbonyl reductase 1	Parallel	96%	68	7	1	6.39	5.58	Cytoplasm	?
PRSS1	Protease, serine, 1 (trypsin 1)	Parallel	100%	56	3	117	5.30	5.30	Extracellular	TRYPSIN, SP
C6orf105	Chromosome 6 open reading frame 105	Parallel	99%	97	5	0	5.12	4.96	?	TM, TM, TM, TM, TM, TM
KLRC1	Killer cell lectin-like receptor subfamily C, member 1	Parallel	100%	85	7	0	4.93	4.93	Plasma membrane	TM, LECTIN_C
ZIC1	Zic family member 1 (odd-paired homolog, Drosophila)	Parallel	100%	59	5	1	4.86	4.86	Nucleus	ZNFC2, ZNFC2, ZNFC2, ZNFC2, ZNFC
TUBB6	Tubulin, beta 6	Parallel	98%	62	6	12	4.86	4.63	?	TUBULIN
ITIH3	Inter-alpha (globulin) inhibitor H3	Parallel	95%	106	9	1	5.09	4.40	Extracellular	SP, VIT, VWA
ACSM3	Acyl-CoA synthetase medium-chain family member 3	Parallel	95%	63	6	0	5.05	4.36	Mitochondrion	?
NUDT11	Nudix (nucleoside diphosphate linked moiety X)-type motif 11	Parallel	96%	73	8	2	4.91	4.33	Cytoplasm	?
UTS2	Urotensin 2	Parallel	98%	104	6	1	4.47	4.21	Plasma membrane	SP
KLRC2	Killer cell lectin-like receptor subfamily C, member 2	Parallel	99%	79	7	0	4.33	4.17	Plasma membrane	LECTIN_C, TM
CADM1	Cell adhesion molecule 1	Parallel	95%	62	10	1	4.80	4.14	Plasma membrane, Nucleus	SP, IG, Ig_LIKE, IGC2, Band 4.1
SEMA3A	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	Parallel	96%	85	10	0	4.55	4.09	Extracellular	SP, SEMA, PSI, IG
TTR	Transthyretin (prealbumin, amyloidosis type I)	Parallel	100%	111	7	2	4.04	4.04	Extracellular	TRANSTHYR, SP
PRKCB1	Protein kinase C, beta 1	Parallel	98%	59	5	0	4.26	4.04	Cytoplasm, Plasma membrane, Nucleus	C1, C1, C2, STkinase
RELN	Reelin	Parallel	98%	64	5	5	4.19	3.99	Extracellular	SP, EGF, EGF, EGF, EGF, EGF, EGF
CPVL	Carboxypeptidase, vitellogenic-like	Parallel	97%	92	10	1	4.41	3.99	Extracellular	SP
IRF4	Interferon regulatory factor 4	Parallel	97%	59	4	0	4.36	3.93	Nucleus, Cytoplasm	IRF
LOC493869	Similar to RIKEN cDNA 2310016C16	Parallel	100%	57	5		3.92	3.92	Integral to membrane	TM
KIAA1576	KIAA1576 protein	Parallel	99%	82	7	0	4.06	3.92	?	?
FLNC	Filamin C, gamma (actin binding protein 280)	Parallel	98%	84	8	1	4.17	3.88	Cytoplasm	CH, CH, IGFLMN, IGFLMN, IGFLMN, I IGFLMN,
ASGR1	Asialoglycoprotein receptor 1	Parallel	100%	105	9	2	3.84	3.84	Plasma membrane, Endoplasmic reticulum	TM, LECTIN_C
UCA1	Urothelial cancer associated 1	Parallel	97%	63	5		4.22	3.83	n/a	n/a
LEFTY2	Left-right determination factor 2	Parallel	95%	62	7	0	4.34	3.74	Extracellular	TGFB, SP
SLC7A3	Solute carrier family 7 (cationic amino acid	Parallel	98%	59	3	1	3.93	3.73	Plasma membrane,	TM, TM, TM, TM, TM, TM, TM

Gene2	**Gene2Name**	**Behavior**	**Purity**	**# CoX**	**# datasets**	**Lit. Str.**	**dMIM**	**Score**	**Cell Location**	**Domains**
	transporter, y+ system), member 3								Nucleus	
MET	Met proto-oncogene (hepatocyte growth factor receptor)	Parallel	98%	64	5	**59**	3.86	3.68	Plasma membrane	SEMA, PSI, IPT, IPT, IPT, IPT
DFNA5	Deafness, autosomal dominant 5	Parallel	94%	63	9	**1**	4.46	3.66	?	?
CPB2	Carboxypeptidase B2 (plasma)	Parallel	98%	55	7	**23.6**	3.84	3.64	?	SP, Zn_pept
PRSS2	Protease, serine, 2 (trypsin 2)	Parallel	100%	61	4	**99**	3.61	3.61	Extracellular	TRYPSIN, SP
RASIP1	Ras interacting protein 1	Parallel	96%	75	6	**2**	4.07	3.60	Cytoplasm, Golgi apparatus, Cytoplasmic vesicle	RA
ROR1	Receptor tyrosine kinase-like orphan receptor 1	Parallel	99%	104	10	**1**	3.69	3.58	Plasma membrane, Cytoplasm	IGC2, KRINGLE, Tyr_Kinase,
SERPINE1	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	Parallel	99%	77	10	**86**	3.71	3.57	Extracellular	SERPIN, SP
EML1	Echinoderm microtubule associated protein like 1	Parallel	100%	55	7	**0**	3.49	3.49	Cytoplasm	CC, WD40, WD40, WD40,
MAGEC1	Melanoma antigen family C, 1	Parallel	100%	62	4	**0**	3.47	3.47	Cytoplasm, Nucleus	MAGE
CRLF1	Cytokine receptor-like factor 1	Parallel	96%	83	8	**0**	3.83	3.43	Extracellular	FN3, FN3, SP, WSXWS
VASN	Vasorin	Parallel	100%	58	6	**2**	3.42	3.42	Plasma membrane, Extracellular	LRR, LRR, LRR, LRR, LRR,
ACF	Apobec-1 complementation factor	Parallel	95%	64	4	**0**	3.85	3.34	Nucleus, Cytoplasm	RRM, RRM, RRM
LAMB1	Laminin, beta 1	Parallel	98%	63	11	**32**	3.50	3.33	Extracellular	LAMNT, EGFLAM, EGFLAM
CFI	Complement factor I	Parallel	95%	126	12	**1**	3.85	3.33	Extracellular	SR, LDLA, LDLA, TRYPSIN, SP
NMU	Neuromedin U	Parallel	97%	95	9	**1**	3.64	3.30	Extracellular	SP
TDGF1	Teratocarcinoma-derived growth factor 1	Parallel	97%	60	4	**2**	3.62	3.27	Extracellular	EGF, SP
NPTX1	Neuronal pentraxin I	Parallel	97%	125	11	**2**	3.59	3.26	Extracellular, Plasma membrane	SP, PTX
SOX2	SRY (sex determining region Y)-box 2	Parallel	100%	57	5	**3**	3.26	3.26	Nucleus	HMG
GDF3	Growth differentiation factor 3	Parallel	92%	63	6	**2**	4.14	3.23	Extracellular	TGFB, SP
F2	Coagulation factor II (thrombin)	Parallel	97%	109	11	**366**	3.50	3.22	Extracellular	GLA, KRINGLE, KRINGLE,
PLA2G4A	Phospholipase A2, group IVA (cytosolic, calcium-dependent)	Parallel	95%	81	10	**0**	3.75	3.22	Cytoplasm, Nucleus	C2
RAI14	Retinoic acid induced 14	Parallel	97%	64	9	**0**	3.50	3.18	Cytoplasm	ANK, ANK, ANK, ANK, ANK
NPTX2	Neuronal pentraxin II	Parallel	89%	64	10	**6**	4.45	3.15	Extracellular, Plasma membrane	SP, CC
PIR	Pirin (iron-binding nuclear protein)	Parallel	97%	59	9	**1**	3.46	3.12	Nuclear matrix	?
ID1	Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein	Parallel	100%	79	9	**6**	3.12	3.12	Nucleus, Centrosome, Cytoplasm	HLH
CP	Ceruloplasmin (ferroxidase)	Parallel	98%	55	9	**9**	3.29	3.12	Extracellular, Endoplasmic reticulum, Golgi apparatus, Plasma membrane	SP
SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	Parallel	98%	82	6	**0**	3.34	3.10	Cytoplasm, Nucleus	SH3, SAM, NLS, NLS, NLS
APOM	Apolipoprotein M	Parallel	100%	79	7	**1**	3.02	3.02	Plasma membrane, Extracellular, Endoplasmic reticulum	TM
DPPA4	Developmental pluripotency associated 4	Parallel	93%	120	9	**0**	3.71	3.01	?	?
ESM1	Endothelial cell-specific molecule 1	Parallel	92%	130	16	**3**	3.90	2.99	Extracellular, Cytoplasm	SP, IB
ASAM	Adipocyte-specific adhesion molecule	Parallel	98%	109	7	**0**	3.13	2.97	Plasma membrane	SP, Ig_LIKE, IGC2, TM
ERRFI1	ERBB receptor feedback inhibitor 1	Parallel	98%	62	7	**1**	3.11	2.96	Cytoplasm	?
IGF2BP3	Insulin-like growth factor 2 mRNA binding protein 3	Parallel	98%	60	4	**0**	3.11	2.96	Cytoplasm, Nucleus	RRM, RRM, KH, KH, KH, KH,

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
CXCL3	Chemokine (C-X-C motif) ligand 3	Parallel	94%	80	8	4	3.57	2.94	Extracellular	SCY, SP, CXC
CTNNAL1	Catenin (cadherin-associated protein), alpha-like 1	Parallel	100%	57	6	29	2.90	2.90	Cytoplasm	?
SILV	Silver homolog (mouse)	Parallel	97%	91	10	1	3.21	2.90	Endosome, Golgi apparatus, Cytoplasm	SP, PKD, TM
CLGN	Calmegin	Parallel	95%	56	7	1	3.41	2.89	Endoplasmic reticulum	SP, TM
CSAG2	CSAG family, member 2	Parallel	100%	56	6	0	2.88	2.88	?	?
CYTL1	Cytokine-like 1	Parallel	92%	60	4	0	3.71	2.86	Extracellular	SP
IGFBP1	Insulin-like growth factor binding protein 1	Parallel	99%	119	10	4	2.93	2.86	Extracellular	SP, IB, THYRO
CAPN2	Calpain 2, (m/II) large subunit	Parallel	92%	59	11	10	3.72	2.85	Cytoplasm, Plasma membrane	CYSPC, CALPAIN_III, EF, EF
FOXD1	Forkhead box D1	Parallel	96%	81	10	1	3.18	2.84	Nucleus	FH
CPS1	Carbamoyl-phosphate synthetase 1, mitochondrial	Parallel	97%	100	12	0	3.11	2.84	Mitochondrion	?
GAL	Galanin	Parallel	100%	98	8	2	2.83	2.83	Extracellular, Secretory granule, Nucleus	SP
ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	Parallel	92%	63	7	0	3.62	2.82	Golgi apparatus	TM
SERPINA7	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7	Parallel	100%	65	6	25	2.81	2.81	Extracellular	SERPIN, SP
PRNP	Prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	Parallel	100%	66	6	0	2.79	2.79	Plasma membrane, Cytoplasm, Golgi apparatus, Nucleus, Extracellular, Endoplasmic reticulum	SP, TM, TM
ANXA1	Annexin A1	Parallel	95%	117	13	3	3.25	2.77	Plasma membrane, Cytoplasm, Nucleus	ANX, ANX, ANX, ANX
CSRP2	Cysteine and glycine-rich protein 2	Parallel	99%	82	8	0	2.86	2.76	Nucleus	LIM, LIM, NLS
FGB	Fibrinogen beta chain	Parallel	97%	63	9	6	3.00	2.72	Extracellular, Endoplasmic reticulum	FBG, SP, CC
MS4A3	Membrane-spanning 4-domains, subfamily A, member 3 (hematopoietic cell-specific)	Parallel	92%	63	4	0	3.45	2.70	Endoplasmic reticulum, Nucleus	TM, TM, TM, TM
G0S2	G0/G1switch 2	Parallel	95%	86	10	0	3.08	2.67	?	TM, CC
NR0B2	Nuclear receptor subfamily 0, group B, member 2	Parallel	98%	55	7	0	2.82	2.67	Nucleus, Cytoplasm	HOLI
POU5F1	POU domain, class 5, transcription factor 1	Parallel	93%	114	9	2	3.31	2.66	Nucleus, Cytoplasm	POU, HOX, CC
MLANA	Melan-A	Parallel	94%	63	6	3	3.24	2.66	Plasma membrane	TM
FAM112B	Family with sequence similarity 112, member B	Parallel	100%	57	2	0	2.65	2.65	?	?
MOXD1	Monooxygenase, DBH-like 1	Parallel	97%	67	7	1	2.89	2.64	Endoplasmic reticulum	SP
CD3D	CD3d molecule, delta (CD3-TCR complex)	Parallel	98%	93	7	1	2.81	2.63	Plasma membrane	SP, ITAM, TM
PCOLCE2	Procollagen C-endopeptidase enhancer 2	Parallel	95%	112	11	1	3.09	2.62	Extracellular	CUB, CUB, C345C, SP
SPTA1	Spectrin, alpha, erythrocytic 1 (elliptocytosis 2)	Parallel	97%	60	6	1	2.89	2.61	Plasma membrane	SPECTRIN, SPECTRIN,EF, EF
BASP1	Brain abundant, membrane attached signal protein 1	Parallel	87%	68	9	1	3.99	2.61	Nucleus, Plasma membrane, Cytoplasm	?
TYRP1	Tyrosinase-related protein 1	Parallel	94%	93	12	2	3.10	2.54	Endosome, Golgi apparatus	TM, SP
AKR1B1	Aldo-keto reductase family 1, member B1 (aldose reductase)	Parallel	98%	56	8	0	2.68	2.54	Cytoplasm, Extracellular	ALDKR
AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	Parallel	99%	92	8	10	2.62	2.53	Extracellular	SERPIN, SP
DKK3	Dickkopf homolog 3 (Xenopus laevis)	Parallel	100%	58	7	0	2.53	2.53	Extracellular, Cytoplasm	SP, CC
LIN28	Lin-28 homolog (C. elegans)	Parallel	97%	102	7	2	2.75	2.52	?	ZnF_C2HC, ZnF_C2HC
NTS	Neurotensin	Parallel	95%	123	11	1	2.92	2.52	Extracellular	SP

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
BCHE	Butyrylcholinesterase	Parallel	97%	62	6	2	2.77	2.51	Extracellular	TM, SP
MYO10	Myosin X	Parallel	98%	63	7	20	2.64	2.51	Cytoplasm, Plasma membrane	MYOSIN, IQ, IQ, IQ, PH, PH, B41, CC
DCT	Dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	Parallel	91%	78	7	2	3.32	2.50	Plasma membrane, Cytoplasm	SP, TM
LTBP1	Latent transforming growth factor beta binding protein 1	Parallel	90%	71	13	5	3.39	2.48	Extracellular	SP, EGF, EGFCA, EGFCA,
ABCC2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	Parallel	95%	59	6	1	2.88	2.46	Plasma membrane	TM, TM, TM, TM, TM, AAA, AAA, AAA
PPAP2B	Phosphatidic acid phosphatase type 2B	Parallel	89%	95	11	0	3.44	2.46	Plasma membrane	TM, TM, TM, TM, TM, TM
DNTT	Deoxynucleotidyltransferase, terminal	Parallel	94%	78	7	0	3.00	2.46	Nucleus	BRCT
DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	Parallel	91%	69	8	0	3.22	2.45	Nucleus, Cytoplasm	DEXDc, HELIC
ASGR2	Asialoglycoprotein receptor 2	Parallel	100%	63	5	0	2.45	2.45	Plasma membrane, Endoplasmic reticulum, Golgi apparatus, Extracellular	TM, LECTIN_C
AKR1C3	Aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II)	Parallel	98%	64	10	0	2.54	2.42	Cytoplasm	ALDKR
PTX3	Pentraxin-related gene, rapidly induced by IL-1 beta	Parallel	98%	111	9	0	2.53	2.40	Extracellular	SP, CC
MGC29506	Hypothetical protein MGC29506	Parallel	100%	59	5	?	2.39	2.39	Cytoplasm, Endoplasmic reticulum, Golgi apparatus	SP
LEPREL1	Leprecan-like 1	Parallel	99%	75	9	0	2.49	2.39	Endoplasmic reticulum, Golgi apparatus	SP, P4HC
DLK1	Delta-like 1 homolog (Drosophila)	Parallel	93%	122	12	0	2.93	2.39	Extracellular, Plasma membrane, Cytoplasm	SP, EGF, EGFCA, EGFCA,TM,
COL9A3	Collagen, type IX, alpha 3	Parallel	91%	93	7	0	3.09	2.36	Extracellular	COLL, COLL, COLL, COLL,
LAMA4	Laminin, alpha 4	Parallel	96%	79	13	0	2.60	2.31	Extracellular, Plasma membrane	EGFLAM, EGFLAM, EGFLAM
LARP6	La ribonucleoprotein domain family, member 6	Parallel	98%	56	6	0	2.43	2.30	?	?
RNF128	Ring finger protein 128	Parallel	96%	67	8	0	2.64	2.30	Late endosome, Endoplasmic reticulum, Golgi apparatus	SP, TM, RING
DDC	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	Parallel	88%	58	7	0	3.37	2.29	Cytoplasm	?
ASS1	Argininosuccinate synthetase 1	Parallel	100%	61	6	0	2.27	2.27	Cytosol, Nucleus, Microsome	?
NTN4	Netrin 4	Parallel	98%	61	10	0	2.38	2.26	Extracellular	SP, LAMNT, EGFLAM, EGFLAM, EGFL
HEMGN	Hemogen	Parallel	88%	65	10	0	3.35	2.26	Nucleus	CC, NLS
TNC	Tenascin C (hexabrachion)	Parallel	91%	89	13	0	2.98	2.24	Extracellular, Plasma membrane	EGF, EGF, EGF, EGF, EGF, FN3, FN3
TPD52L1	Tumor protein D52-like 1	Parallel	91%	69	12	0	2.93	2.23	?	PEST, CC, PEST
IGFBP3	Insulin-like growth factor binding protein 3	Parallel	99%	97	10	0	2.30	2.23	Extracellular, Nucleus, Cytoplasm	SP, IB, THYRO, NLS
IGF2	Insulin-like growth factor 2 (somatomedin A)	Parallel	97%	70	11	0	2.43	2.22	Extracellular	INS, SP
GDA	Guanine deaminase	Parallel	89%	85	9	0	3.10	2.21	Cytoplasm, Plasma membrane	?
FABP1	Fatty acid binding protein 1, liver	Parallel	97%	120	10	3	2.44	2.20	Cytoplasm, Nucleus	?
CPE	Carboxypeptidase E	Parallel	98%	64	9	0	2.27	2.17	Plasma membrane, Cytoplasmic vesicle, Endoplasmic reticulum	SP, Zn_pept

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
CA9	Carbonic anhydrase IX	Parallel	95%	57	8	0	2.53	2.16	Plasma membrane, Cytoplasm	SP, TM
SERPINE2	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	Parallel	92%	106	12	3.6	2.72	2.15	Extracellular	SERPIN, SP
MFAP4	Microfibrillar-associated protein 4	Parallel	96%	82	7	0	2.39	2.13	Extracellular	FBG, RGD, SP
UCHL1	Ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	Parallel	91%	81	11	0	2.79	2.13	Cytoplasm	?
ALDH1A1	Aldehyde dehydrogenase 1 family, member A1	Parallel	91%	105	16	0	2.78	2.13	?	A2M
HBZ	Hemoglobin, zeta	Parallel	94%	109	11	0	2.52	2.12	?	?
JARID1D	Jumonji, AT rich interactive domain 1D	Parallel	91%	58	5	0	2.76	2.11	?	CC, CC, BRIGHT, PHD, JMJC
MAGEA3	Melanoma antigen family A, 3	Parallel	100%	64	6	0	2.10	2.10	Cytoplasm	MAGE
TGFBI	Transforming growth factor, beta-induced, 68kDa	Parallel	93%	86	12	0	2.60	2.09	Extracellular, Cytoplasm, Nucleus	SP
TFPI2	Tissue factor pathway inhibitor 2	Parallel	97%	116	12	37.4	2.24	2.08	Extracellular, Plasma membrane, Endoplasmic reticulum	SP, KUNITZ, KUNITZ, KUNITZ
RPS4Y1	Ribosomal protein S4, Y-linked 1	Parallel	89%	61	9	0	2.98	2.07	Ribosome	S4, KOW
FGL1	Fibrinogen-like 1	Parallel	98%	80	8	0	2.22	2.06	Extracellular, Cytoplasm	SP, FBG
ADAM23	ADAM metallopeptidase domain 23	Parallel	91%	64	10	0	2.70	2.01	Plasma membrane, Extracellular	TM, REP, DISINT, ACR, SP
PLAT	Plasminogen activator, tissue	Parallel	95%	66	11	129.2	2.31	2.01	Extracellular	SP, FN1, EGF, KRINGLE,
DNER	Delta/notch-like EGF repeat containing	Parallel	99%	91	10	0	2.06	1.99	Plasma membrane, Endosome	SP, EGF, EGF, EGFCA, EGF,
PRSS23	Protease, serine, 23	Parallel	92%	63	12	0	2.55	1.99	?	TRYPSIN, SP
FBN1	Fibrillin 1	Parallel	86%	88	10	0	3.08	1.98	Extracellular	EGF, EGFCA, EGFCA, SP
L1TD1	LINE-1 type transposase domain containing 1	Parallel	90%	106	10	0	2.73	1.96	?	CC, CC
AKR1C2	Aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid	Parallel	98%	57	6	0	2.05	1.94	Cytosol	?
FST	Follistatin	Parallel	94%	131	14	0	2.35	1.94	Extracellular	FOLN, KAZAL, FOLN, KAZAL,
SLC7A5	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	Parallel	95%	61	6	0	2.26	1.94	Plasma membrane	TM, TM, TM, TM, TM, TM, TM,
ADA	Adenosine deaminase	Parallel	93%	57	6	0	2.41	1.94	Plasma membrane, Cytoplasm	?
SRPX	Sushi-repeat-containing protein, X-linked	Parallel	94%	64	12	0	2.34	1.93	Plasma membrane	SP, SUSHI, SUSHI, SUSHI
PCSK9	Proprotein convertase subtilisin/kexin type 9	Parallel	87%	69	6	0	2.92	1.92	Extracellular	SP
NNMT	Nicotinamide N-methyltransferase	Parallel	99%	77	13	0	2.00	1.92	Cytoplasm	?
CAV1	Caveolin 1, caveolae protein, 22kDa	Parallel	99%	96	14	2.9	1.97	1.91	Plasma membrane	TM
DKK1	Dickkopf homolog 1 (Xenopus laevis)	Parallel	98%	97	13	0	2.02	1.90	Extracellular	SP
HP	Haptoglobin	Parallel	95%	108	7	0	2.15	1.86	Extracellular	SUSHI, SUSHI, TRYPSIN, SP
TSPAN7	Tetraspanin 7	Parallel	89%	66	9	0	2.61	1.86	Plasma membrane	TM, TM, TM, TM
FGG	Fibrinogen gamma chain	Parallel	97%	61	8	137	2.05	1.86	Extracellular, Golgi apparatus	FBG, SP, CC, CC
NQO1	NAD(P)H dehydrogenase, quinone 1	Parallel	97%	73	9	0	2.02	1.86	Mitochondrion, Microsome, Plasma membrane, Cytoplasm, Nucleus	?
MAL	Mal, T-cell differentiation protein	Parallel	93%	92	12	0	2.26	1.85	Endoplasmic reticulum, Plasma membrane, Golgi apparatus, Mitochondrion, Golgi vesicle, Endosome	TM, TM, TM, TM

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
RHOBTB3	Rho-related BTB domain containing 3	Parallel	97%	58	14	0	2.05	1.84	?	BTB, BTB
BEX1	Brain expressed, X-linked 1	Parallel	95%	96	9	0	2.17	1.84	?	CC
CGA	Glycoprotein hormones, alpha polypeptide	Parallel	96%	55	7	0	2.03	1.82	Extracellular, Endoplasmic reticulum, Golgi apparatus	SP
SNCA	Synuclein, alpha (non A4 component of amyloid precursor)	Parallel	88%	68	7	0	2.65	1.82	Cytoplasm, Nucleus	SYNUC
ITIH2	Inter-alpha (globulin) inhibitor H2	Parallel	96%	110	11	0	2.03	1.81	Extracellular	SP, VIT, VWA
CCL20	Chemokine (C-C motif) ligand 20	Parallel	96%	69	11	0	2.06	1.80	Extracellular, Cytoplasm	SCY, SP
PLAC8	Placenta-specific 8	Parallel	88%	81	12	0	2.67	1.80	Extracellular	SP
TMEM158	Transmembrane protein 158	Parallel	97%	86	11	0	1.97	1.77	?	SP, TM, TM
MGST1	Microsomal glutathione S-transferase 1	Parallel	95%	56	9	0	2.09	1.77	Microsome, Endoplasmic reticulum, Nuclear membrane	SP
APOH	Apolipoprotein H (beta-2-glycoprotein I)	Parallel	98%	123	10	2.6	1.90	1.76	Extracellular, Plasma membrane, Cytoplasm	SUSHI, SUSHI, SUSHI, SUSHI, SP
SERPINB2	Serpin peptidase inhibitor, clade B (ovalbumin), member 2	Parallel	93%	81	9	6.9	2.22	1.76	Cytoplasm, Extracellular, Nucleus	SP, SERPIN
NRN1	Neuritin 1	Parallel	85%	74	8	0	2.84	1.75	Extracellular	SP
COL6A1	Collagen, type VI, alpha 1	Parallel	91%	56	11	0	2.32	1.75	Extracellular	VWA, COLL, COLL, COLL, COLL, VWA
INA	Internexin neuronal intermediate filament protein, alpha	Parallel	93%	58	7	0	2.17	1.75	Cytoplasm	CC, CC
MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	Parallel	93%	83	14	0	2.17	1.73	Extracellular	ZnMc, HX, HX, HX, HX, SP
ITGB7	Integrin, beta 7	Parallel	93%	58	8	0	2.14	1.73	Plasma membrane	SP, PSI, TM, VWA
PLAU	Plasminogen activator, urokinase	Parallel	91%	57	10	11.2	2.26	1.72	Plasma membrane, Cytoplasm, Extracellular	KRINGLE, TRYPSIN, SP
S100B	S100 calcium binding protein B	Parallel	89%	82	9	0	2.42	1.71	Cytoplasm, Extracellular, Nucleus	EF, S_100
MAGEA5	Melanoma antigen family A, 5	Parallel	97%	68	7	0	1.84	1.69	?	MAGE
CNN3	Calponin 3, acidic	Parallel	93%	73	16	0	2.07	1.67	Cytoplasm	CH
AFP	Alpha-fetoprotein	Parallel	96%	119	12	0	1.89	1.66	Extracellular, Cytoplasm	Albumin, Albumin, Albumin, SP
APOA1	Apolipoprotein A-I	Parallel	98%	91	11	1.5	1.77	1.66	Extracellular, Cytoplasm	SP, CC
TSPAN8	Tetraspanin 8	Parallel	83%	98	11	0	2.93	1.65	Plasma membrane	TM, TM, TM, TM
ALAS2	Aminolevulinate, delta-, synthase 2 (sideroblastic/hypochromic anemia)	Parallel	92%	60	5	0	2.13	1.64	Mitochondrion	?
PAGE1	P antigen family, member 1 (prostate associated)	Parallel	100%	58	7	0	1.64	1.64	Cytoplasm	?
SLC2A3	Solute carrier family 2 (facilitated glucose transporter), member 3	Parallel	83%	63	11	0	2.89	1.63	Plasma membrane	TM, TM, TM, TM, TM, TM, TM,
PEG10	Paternally expressed 10	Parallel	91%	100	18	0	2.16	1.63	Nucleus, Cytoplasm	CC, ZNFC2
GAGE7B	G antigen 7B	Parallel	98%	64	4	0	1.70	1.62	?	?
LGALS3	Lectin, galactoside-binding, soluble, 3	Parallel	86%	66	12	0	2.52	1.62	Cytoplasm, Nucleus, Extracellular	GLECT
RHAG	Rh-associated glycoprotein	Parallel	81%	102	11	0	2.95	1.59	Plasma membrane	TM, TM, TM, TM, TM, TM, TM,
PRSS3	Protease, serine, 3 (mesotrypsin)	Parallel	98%	59	4	0	1.67	1.59	?	SP, TRYPSIN
SFRP1	Secreted frizzled-related protein 1	Parallel	88%	76	10	0	2.31	1.59	Extracellular	FZ, C345C, SP
F5	Coagulation factor V (proaccelerin, labile factor)	Parallel	80%	66	11	40.6	3.03	1.57	Extracellular, Plasma membrane, Endoplasmic reticulum, Golgi apparatus	F58C, F58C, SP
CDCA7	Cell division cycle associated 7	Parallel	91%	58	8	0	2.02	1.54	Nucleus	?
NDN	Necdin homolog (mouse)	Parallel	90%	59	7	0	2.11	1.53	Nucleus, Cytoplasm	MAGE

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
AMBP	Alpha-1-microglobulin/bikunin precursor	Parallel	98%	118	10	1	1.60	1.52	Extracellular, Golgi apparatus, Endoplasmic reticulum, Cytoplasm, Plasma membrane	SP, KUNITZ, KUNITZ
FABP7	Fatty acid binding protein 7, brain	Parallel	93%	59	9	0	1.88	1.52	Cytoplasm	?
TM4SF1	Transmembrane 4 L six family member 1	Parallel	93%	60	16	0	1.87	1.52	Plasma membrane	TM, TM, TM, TM
WNT5A	Wingless-type MMTV integration site family, member 5A	Parallel	86%	72	17	0	2.34	1.50	Extracellular	SP, WNT
FN1	Fibronectin 1	Parallel	91%	81	13	0	1.96	1.49	Extracellular, Endoplasmic reticulum	FN1, FN1, FN1, FN1, FN1, FN1, FN2, FN3, FN3, FN3, FN3, FN3, FN3, FN3, FN1, RGD, SP
SCG2	Secretogranin II (chromogranin C)	Parallel	89%	57	10	0	2.06	1.48	Golgi apparatus	SP, CC
GNG11	Guanine nucleotide binding protein (G protein), gamma 11	Parallel	81%	70	12	0	2.71	1.46	?	CC
GPNMB	Glycoprotein (transmembrane) nmb	Parallel	88%	77	13	0	2.10	1.45	Plasma membrane, Endosome	SP, PKD, TM
APOC1	Apolipoprotein C-I	Parallel	90%	58	14	0	1.97	1.42	Extracellular	SP
AHSG	Alpha-2-HS-glycoprotein	Parallel	96%	85	12	0	1.58	1.42	Extracellular	CYSTATIN, CYSTATIN, SP
ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	Parallel	92%	60	8	0	1.83	1.41	Extracellular, Golgi apparatus, Cytoplasmic vesicle	SP, ACR, TSP1, TSP1, TSP1
BDNF	Brain-derived neurotrophic factor	Parallel	74%	69	6	0	3.47	1.40	Extracellular, Secretory granule, Golgi apparatus	?
CDH2	Cadherin 2, type 1, N-cadherin (neuronal)	Parallel	91%	68	11	0	1.85	1.40	Plasma membrane	Cadherin, Cadherin, Cadherin,
APOA2	Apolipoprotein A-II	Parallel	100%	87	9	2.1	1.40	1.40	Extracellular	SP
NR2F2	Nuclear receptor subfamily 2, group F, member 2	Parallel	82%	91	12	0	2.38	1.33	Nucleus	ZnF_C4, HOLI
FSTL1	Follistatin-like 1	Parallel	90%	69	10	0	1.83	1.33	Extracellular	FOLN, KAZAL, SP, CC, EF, EF
PRAME	Preferentially expressed antigen in melanoma	Parallel	87%	85	13	0	1.97	1.30	?	?
HTRA1	HtrA serine peptidase 1	Parallel	97%	60	7	0	1.44	1.30	Extracellular	IB, KAZAL, TRYPSIN, PDZ, SP
GAGE6	G antigen 6	Parallel	100%	58	2	0	1.29	1.29	?	?
LUM	Lumican	Parallel	86%	111	17	0	2.05	1.29	Extracellular	SP, LRR, LRR, LRR, LRR, LRR, LRR
THBS2	Thrombospondin 2	Parallel	81%	85	14	0.5	2.38	1.27	Extracellular	TSPN, VWC, TSP1, TSP1, TSP1, EGF
CXCR4	Chemokine (C-X-C motif) receptor 4	Parallel	78%	55	10	0	2.64	1.26	Plasma membrane, Cytoplasm	TM, TM, TM, TM, TM, TM, TM
ORM2	Orosomucoid 2	Parallel	90%	67	7	0	1.71	1.23	Extracellular	SP
ADM	Adrenomedullin	Parallel	95%	57	7	14.5	1.44	1.23	Extracellular, Cytoplasm	SP
GPC3	Glypican 3	Parallel	93%	94	12	2.8	1.54	1.22	Extracellular, Plasma membrane	SP
FHL1	Four and a half LIM domains 1	Parallel	93%	57	7	0	1.51	1.22	Cytoplasm	LIM, LIM, LIM, LIM
TACSTD2	Tumor-associated calcium signal transducer 2	Parallel	83%	60	8	0	2.10	1.21	Plasma membrane, Cytoplasm	THYRO, TM, SP
GSTM3	Glutathione S-transferase M3 (brain)	Parallel	66%	64	11	0	4.27	1.21	Cytoplasm	?
ANLN	Anillin, actin binding protein	Parallel	96%	68	11	0	1.38	1.20	Nucleus, Nucleolus	CC, CC, PH
APOB	Apolipoprotein B (including Ag(x) antigen)	Parallel	97%	66	10	12.2	1.27	1.15	Extracellular, Endoplasmic reticulum, Golgi apparatus	SP, CC, CC, CC
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	Parallel	77%	79	13	0	2.41	1.11	Plasma membrane, Cytoplasm, Extracellular	TM, Cadherin, Cadherin, SP,
CTGF	Connective tissue growth factor	Parallel	89%	73	18	0.5	1.50	1.06	Extracellular, Golgi apparatus	SP, IB, VWC, CT, TSP1
IGJ	Immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	Parallel	86%	66	13	0	1.64	1.06	Endoplasmic reticulum, Golgi apparatus,	?

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
NPY1R	Neuropeptide Y receptor Y1	Parallel	70%	61	11	0	2.84	1.00	Nucleus, Extracellular Plasma membrane	TM, TM, TM, TM, TM, TM, TM
STAR	Steroidogenic acute regulator	Parallel	67%	58	7	0	3.21	0.98	Mitochondrion	START
MME	Membrane metallo-endopeptidase	Parallel	79%	56	10	0	2.00	0.97	Plasma membrane, Cytoplasm, Endoplasmic reticulum, Golgi apparatus, Endosome, Extracellular matrix	CC, TM, NLS
HBE1	Hemoglobin, epsilon 1	Parallel	83%	71	10	0	1.66	0.96	Cytoplasm, Nucleus	?
PKIB	Protein kinase (cAMP-dependent, catalytic) inhibitor beta	Parallel	64%	78	8	0	3.17	0.84	Nucleus	NES
TACSTD1	Tumor-associated calcium signal transducer 1	Parallel	82%	74	12	0	1.47	0.83	Plasma membrane, Extracellular	SP, TM, THYRO
SPP1	Secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	Parallel	86%	80	14	0	1.23	0.79	Extracellular	SP
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	Parallel	71%	94	14	0	1.88	0.68	Extracellular	EGFCA, EGFCA, EGFCA, EGFL, SP
POSTN	Periostin, osteoblast specific factor	Parallel	77%	61	10	0	1.49	0.68	?	?
COL1A2	Collagen, type I, alpha 2	Parallel	70%	70	12	0	1.65	0.57	Extracellular	COLFI, SP, COLL, COLL, COLL, COLL
PTPRC	Protein tyrosine phosphatase, receptor type, C	Parallel	63%	60	10	0	2.13	0.54	Plasma membrane	TM, FN3, FN3, Tyr_Phos, Tyr_Phos, S
C4orf18	Chromosome 4 open reading frame 18	Parallel	55%	75	11	?	3.24	0.53	?	?
COL1A1	Collagen, type I, alpha 1	Parallel	75%	57	11	0	1.22	0.52	Extracellular	VWC, COLL, COLL, COLL, COLL, COL
FXYD3	FXYD domain containing ion transport regulator 3	Parallel	58%	62	9	0	2.45	0.48	Plasma membrane, Cytoplasm	SP, TM
S100P	S100 calcium binding protein P	Parallel	68%	59	16	0	1.25	0.39	Cytoplasm, Plasma membrane	EF, EF
SUSD3	Sushi domain containing 3	Anti	100%	64	5	0	5.65	5.65	Integral to membrane	SUSHI, TM
PIB5PA	Phosphatidylinositol (4,5) bisphosphate 5-phosphatase, A	Anti	100%	56	4	0	5.25	5.25	Cytoplasm, Plasma membrane	I_Phosphatase
EMG1	EMG1 nucleolar protein homolog (S. cerevisiae)	Anti	100%	57	3	0	5.14	5.14	Nucleus, Nucleolus	?
PARP9	Poly (ADP-ribose) polymerase family, member 9	Anti	100%	57	4	0	5.08	5.08	Nucleus	?
RTN1	Reticulon 1	Anti	100%	59	3	0	4.92	4.92	Endoplasmic reticulum	TM, TM
FLJ20273	RNA-binding protein	Anti	94%	63	6	?	5.95	4.88	?	RRM, RRM, RRM
BATF	Basic leucine zipper transcription factor, ATF-like	Anti	100%	59	5	0	4.78	4.78	Nucleus	BRLZ, CC
CAPN13	Calpain 13	Anti	100%	57	2	0	4.10	4.10	Cytoplasm	CYSPC, CALPAIN_III
ESR1	Estrogen receptor 1	Anti	96%	56	3	2.6	4.52	4.05	Nucleus, Cytoplasm	ZnF_C4, HOLI
COL10A1	Collagen, type X, alpha 1(Schmid metaphyseal chondrodysplasia)	Anti	98%	56	4	0	4.02	3.81	Extracellular	COLL, COLL, COLL, COLL, SP
LOC124220	Similar to common salivary protein 1	Anti	100%	57	4	?	3.71	3.71	?	SP
HDC	Histidine decarboxylase	Anti	100%	65	5	0	3.58	3.58	Cytoplasm	?
ENPP5	Ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative function)	Anti	98%	58	5	0	3.74	3.55	?	TM
HLA-F	Major histocompatibility complex, class I, F	Anti	100%	56	5	0	3.46	3.46	Plasma membrane	SP, IGC, TM
CTSO	Cathepsin O	Anti	93%	58	5	0	4.11	3.32	Lysosome, Extracellular	PPTC1, SP
PIP	Prolactin-induced protein	Anti	95%	58	3	0	3.49	2.98	Extracellular	SP
GBP5	Guanylate binding protein 5	Anti	100%	60	5	0	2.96	2.96	?	CC
DNALI1	Dynein, axonemal, light intermediate chain 1	Anti	96%	56	5	0	3.25	2.91	?	CC
TMEM125	Transmembrane protein 125	Anti	100%	56	2	0	2.87	2.87	?	TM, TM, TM, TM
GJA4	Gap junction protein, alpha 4, 37kDa	Anti	96%	57	5	0	3.09	2.78	Plasma membrane	TM, CNX, TM, TM
TMEM16A	Transmembrane protein 16A	Anti	95%	58	6	0	3.24	2.77	Integral to membrane	TM, TM, TM, TM, TM

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
CD14	CD14 molecule	Anti	97%	63	7	1.3	2.90	2.63	Cell surface, Extracellular, Endoplasmic reticulum, Golgi apparatus, Lysosome, Endosome, Cytoplasm	LRR, SP, LRR, LRR
CYP2B7P1	Cytochrome P450, family 2, subfamily B, polypeptide 7 pseudogene 1	Anti	90%	58	5	0	3.63	2.62	n/a	n/a
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1	Anti	100%	76	6	0	2.61	2.61	Plasma membrane	TM, IGC, SP
VAV1	Vav 1 oncogene	Anti	76%	59	7	0	5.88	2.61	Cytoplasm, Nucleus, Plasma membrane	CH, RHOGEF, PH, C1, SH3, SH2, SH3
PDZK1IP1	PDZK1 interacting protein 1	Anti	96%	55	7	0	2.80	2.51	Plasma membrane	TM, SP
CCL4	Chemokine (C-C motif) ligand 4	Anti	100%	91	9	0	2.35	2.35	Extracellular, Endoplasmic reticulum, Golgi apparatus	SCY, SP
TRPS1	Trichorhinophalangeal syndrome I	Anti	88%	65	6	0	3.37	2.28	Nucleus	ZNFC2, ZNF_GATA, ZNFC2,
EFHD1	EF-hand domain family, member D1	Anti	100%	57	2	0	2.27	2.27	?	EF, EF, CC
HCST	Hematopoietic cell signal transducer	Anti	100%	65	6	0	2.15	2.15	Plasma membrane	TM
VAV3	Vav 3 oncogene	Anti	84%	70	10	0	3.56	2.13	Cytoplasm	CH, RHOGEF, PH, C1, SH3, SH2, SH3
CSF1R	Colony stimulating factor 1 receptor, formerly McDonough feline sarcoma viral (v-fms) oncogene homolog	Anti	97%	71	8	0	2.26	2.07	Plasma membrane	TM, Ig_LIKE, Ig_LIKE, IGC2,
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	Anti	100%	64	6	0	2.04	2.04	Cytoplasm, Nucleolus	TPR, TPR, TPR, TPR, TPR, TPR
SMOC2	SPARC related modular calcium binding 2	Anti	92%	71	7	0	2.63	2.02	Extracellular	KAZAL, THYRO, THYRO, SP
GZMA	Granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)	Anti	98%	62	7	0	2.07	1.97	Extracellular, Cytoplasm	TRYPSIN, SP
CXCL10	Chemokine (C-X-C motif) ligand 10	Anti	93%	68	8	0	2.23	1.77	Extracellular	SP, SCY
CD93	CD93 molecule	Anti	91%	58	8	0	2.23	1.70	Plasma membrane	TM, LECTIN_C, EGFL, EGFL,
HCK	Hemopoietic cell kinase	Anti	98%	57	9	0	1.80	1.70	Cytoplasm, Plasma membrane, Lysosome	SH3, SH2, Tyr_Kinase
ARHGAP25	Rho GTPase activating protein 25	Anti	97%	58	8	0	1.82	1.64	?	CC, PH, RHOGAP
IFI27	Interferon, alpha-inducible protein 27	Anti	99%	79	9	0	1.63	1.57	?	TM, TM
C10orf116	Chromosome 10 open reading frame 116	Anti	90%	62	5	?	1.90	1.40	?	?
HLA-DRB4	Major histocompatibility complex, class II, DR beta 4	Anti	100%	76	5	0	1.31	1.31	Plasma membrane	SP, IGC, TM
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1	Anti	97%	93	10	0	1.41	1.28	Plasma membrane, Endosome, Lysosome	SP, IGC, TM
KIT	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	Anti	75%	60	9	0	3.02	1.27	Plasma membrane	Ig_LIKE, IGC2, Ig_LIKE, Tyr_Kinase, T
TYROBP	TYRO protein tyrosine kinase binding protein	Anti	100%	68	8	0	1.27	1.27	Plasma membrane, Cell surface, Cytoplasm	TM, SP, ITAM
HOXB2	Homeobox B2	Anti	73%	56	8	0	3.22	1.26	Nucleus	HOX
COL11A1	Collagen, type XI, alpha 1	Anti	74%	68	9	0	3.02	1.20	Extracellular	COLL, COLL, COLL, COLL, SP, TSPN
HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1	Anti	99%	70	11	0	1.18	1.13	Plasma membrane	SP, IGC, TM
CD33	CD33 molecule	Anti	56%	55	7	0	6.23	1.12	Plasma membrane	SP, IG, IG, TM
PLEK	Pleckstrin	Anti	91%	64	9	0	1.45	1.08	Plasma membrane	PH, PH, DSH
RGS2	Regulator of G-protein signalling 2, 24kDa	Anti	80%	56	10	0	2.06	1.07	Plasma membrane, Nucleus, Cytoplasm	RGS
F13A1	Coagulation factor XIII, A1 polypeptide	Anti	89%	66	12	0.5	1.49	1.06	Extracellular, Nucleus, Cytoplasm	TRANSGLUT, TGC

Gene2	Gene2Name	Behavior	Purity	# CoX	# datasets	Lit. Str.	dMIM	Score	Cell Location	Domains
GMFG	Glia maturation factor, gamma	Anti	78%	59	8	0	2.10	0.99	?	ADF
HBB	Hemoglobin, beta	Anti	91%	77	12	0	1.30	0.98	Extracellular	?
ZMAT1	Zinc finger, matrin type 1	Anti	78%	59	6	0	1.86	0.88	Nucleus	ZNFC2
SPARCL1	SPARC-like 1 (mast9, hevin)	Anti	95%	55	8	0	0.99	0.84	Extracellular	SP, CC, FOLN, KAZAL, EF
HBD	Hemoglobin, delta	Anti	72%	68	10	0	1.47	0.55	?	?
AIF1	Allograft inflammatory factor 1	Anti	67%	60	10	0	1.84	0.55	Cytoplasm, Nucleus	EF
SRGN	Serglycin	Anti	58%	93	17	0	1.95	0.38	Extracellular, Cytoplasm	SP
MGP	Matrix Gla protein	Anti	66%	67	15	0	1.26	0.36	Extracellular	GLA, SP
CD53	CD53 molecule	Anti	57%	65	12	1.5	1.75	0.32	Plasma membrane	TM, TM, TM, TM
IL1B	Interleukin 1, beta	Anti	51%	73	10	4.1	1.57	0.20	Extracellular, Plasma membrane, Cytoplasm, Nucleus	IL1

Table S2. 8 out of the top 20 genes consistently co-expressed with C6orf105, have also been associated with TFPI in the literature (red)

Gene name	Literature references
C3orf57	
SPANX family, member C	
Alpha-1-microglobulin/bikunin	Macedo-Ribeiro S. PLoS ONE. 2008; 3(2): e1624
Urotensin 2	
SPANX family, member A1	
Sp5 transcription factor	
Asialoglycoprotein receptor 1	Huang Y-CT. Am J Respir Cell Mol Biol. 2009; 41(5): 544
Variable charge, X-linked 2	
RIKEN cDNA 2310016C16	
SH3 domain containing 19	
Vasorin	Katayama H. Genome Med. 2009; 1(4): 47 Chowdhury UR. Invest Ophthalmol Vis Sci. 2010; 51(10): 4921
Hypothetical protein LOC729085	
Paired related homeobox 2	
Fibrinogen gamma chain	137 articles in MEDLINE
Fatty acid binding protein 1, liver	Parent R. Genome Biol. 2008; 9(1): R19 Juhl K. Diabetes. 2008; 57(10): 2755
Transthyretin	Gordon S. Curr Cardiovasc Risk Rep. 2010; 4(1): 1
T cell receptor beta variable 5-4	Ge Y. Blood. 2006 15; 107(4): 1570
Erythrocyte membrane protein band 4.1-like 1	
Coagulation factor II (thrombin)	366 articles in MEDLINE
Ninein-like	

Table S3. Predicted associations for TFPI based on co-expression commonalities

 # Shared Rels: The number of co-expressed genes (out of 20) that are associated.
Obs/Exp: Statistical enrichment ratio between the number of observed commonalities and the number expected by chance alone, given the same number of identically connected genes within a random network of the same size.
Score: A priority score to rank the relative importance of the commonalities found.
Lit Str (Literature strength): Cumulative tally of the number of times TFPI was mentioned with the term in MEDLINE (0.5 for every abstract co-occurrence, 0.8 per sentence). When **blue**, the GAMMA-predicted association is also a published one, indicating that GAMMA predictions for TFPI widely coincide with known associations.
One out of three predicted associations for TFPI is also common for ADTRP (Yes).

Predicted association	# Shared Rels	Obs/Exp	Score	Lit_Str	Is this a predicted association for ADTRP also?
HEPATOCELLULAR CARCINOMA	12	3.81	44	22.8	Yes
GATA-6	4	11.28	44	0	
IL-6	12	3.64	42	16.2	Yes
BFGF	9	4.83	42	23.8	
VEGF	10	4.04	40	20.2	Yes
Preeclampsia	7	5.7	39	19.2	Yes
mitogen-activated protein kinase	11	3.57	38	8.9	Yes
LDL	9	4.31	38	88.3	Yes
Proteoglycans	8	4.67	35	9	
Factor XIa	3	11.8	34	5.6	
factor H	4	8.48	33	0	
NF-kappaB	9	3.63	31	9.8	
neurogenesis	6	5	29	0	
Factor XIIa	3	10.21	29	5.2	
C-Reactive Protein	8	3.65	28	41.6	Yes
MMP-9	7	4.16	28	8.1	
anticoagulant activity	4	7.18	28	89.6	
urokinase-type plasminogen activator receptor	4	7.46	28	3.4	
HNF1	4	7.16	28	0	

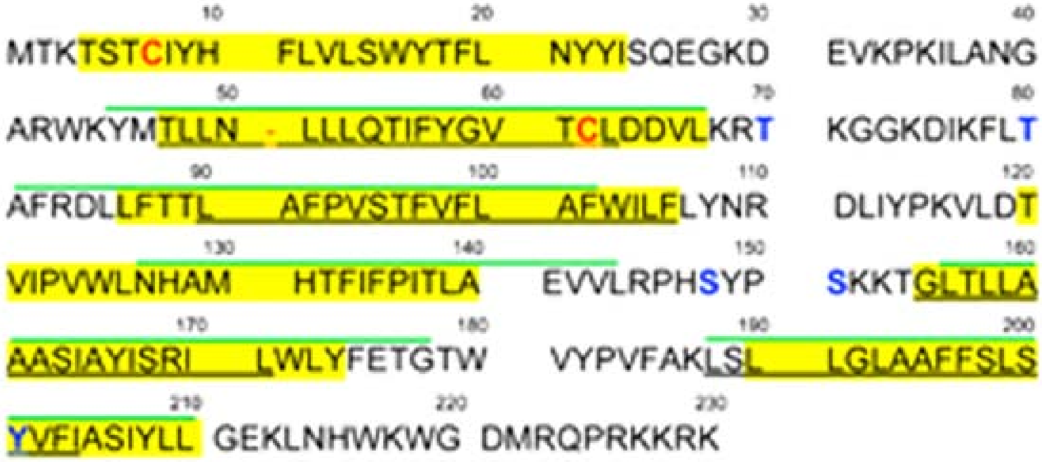
Predicted association	# Shared Rels	Obs/Exp	Score	Lit_Str	Is this a predicted association for ADTRP also?
erythroid differentiation factor	3	9.93	28	0	
phosphorylation	12	2.38	27	13	
TNF-alpha	11	2.54	27	64.1	
gene silencing	6	4.65	27	0.5	
Breast cancer	10	2.67	26	10.7	Yes
fibronectin	8	3.42	26	4.1	Yes
Fibrin	7	3.92	26	126.2	
AFP	6	4.6	26	0	Yes
antithrombin III	5	5.38	26	117	Yes
heparin binding	5	5.32	26	20.7	
adiponectin	5	5.35	26	6.6	Yes
alpha 1-Antitrypsin	5	5.31	26	1.5	
cell surface	10	2.6	25	93.4	
transmembrane	10	2.56	25	7.3	
PAI-1	6	4.41	25	166.6	
Apolipoprotein	6	4.33	25	19.4	
tissue factor	5	5.23	25	1305.7	
chronic kidney disease	5	5.25	25	7.4	
Cripto-1	3	8.84	25	0	
JNK	7	3.65	24	8.8	
plasminogen	6	4.23	24	86.5	
UROKINASE	6	4.18	24	34.3	
meningococcal disease	3	8.19	24	4.5	Yes
brachyury	3	8.08	24	0	Yes
sepsis	8	2.95	23	121.6	Yes
Angiogenesis	8	3.03	23	24.6	
Adipocytes	8	2.97	23	2.1	Yes
inflammatory cytokines	7	3.49	23	7.9	

Predicted association	# Shared Rels	Obs/Exp	Score	Lit_Str	Is this a predicted association for ADTRP also?
Disulfide	7	3.42	23	6.8	
LDL receptor	5	4.68	23	3.1	
Obesity	8	2.92	22	3.4	Yes
leptin	6	3.87	22	2.1	Yes
embryonic stem cells	6	3.86	22	0	Yes
cell proliferation	10	2.22	21	7.9	Yes
glycosylation	7	3.09	21	12.7	Yes
fibrinolysis	5	4.38	21	71.9	
thrombospondin-1	4	5.43	21	2.6	Yes- anti-parallel
TGF-beta1	7	3.04	20	0	Yes
EGFR	7	3.15	20	0	
Serine protease	6	3.52	20	29.7	
Elastase	5	4.09	20	5.4	
thrombomodulin	4	5.26	20	231.8	
Myocardial Infarction	7	2.84	19	29.1	
EGF	7	2.74	19	2.6	
Creatinine	7	2.89	19	2.6	
TGF-BETA	7	2.77	19	1.3	Yes- anti-parallel
low-density lipoprotein	6	3.43	19	19	Yes
Fibroblast growth factor	6	3.42	19	3.1	Yes
ANP	5	3.92	19	0	
cell membrane	9	2.16	18	18	
Immunoglobulin	8	2.29	18	6.9	
Alternative splicing	7	2.73	18	20.1	
protease inhibitors	6	3.04	18	55.2	
chronic renal failure	6	3.15	18	12.7	Yes
APOE	5	3.82	18	3.6	
ovulation	5	3.79	18	1.5	

Predicted association	# Shared Rels	Obs/Exp	Score	Lit_Str	Is this a predicted association for ADTRP also?
HGF	5	3.7	18	0	Yes
SHH	4	4.81	18	0	Yes
oxidative stress	8	2.19	17	8.1	
Antioxidant	8	2.22	17	3.1	
homeostasis	8	2.21	17	2	Yes
IL-1beta	7	2.5	17	15.2	Yes - anti-parallel
cell adhesion	7	2.64	17	5.4	Yes
Prostate cancer	7	2.58	17	3.9	Yes
Fibrinogen	6	2.98	17	106.9	
autocrine	6	2.93	17	9.7	Yes
polysaccharide	6	3.06	17	9	
paracrine	6	3.05	17	8.4	
ERK1	6	2.99	17	7.3	
heat shock	6	3.05	17	0	
Protein C	5	3.43	17	228.1	
clotting	5	3.68	17	132.7	
Plasminogen activator	5	3.61	17	65	
Angiotensin	5	3.55	17	3.6	
factor Xa	4	4.51	17	418.6	
acute phase response	4	4.34	17	3.1	Yes
TLR2	4	4.44	17	0	
signal transduction	8	2.09	16	1	
Thrombin	6	2.72	16	459	
Autoantibodies	6	2.83	16	10.1	
Nephropathy	5	3.45	16	10.3	
hypercholesterolemia	5	3.39	16	4.4	
Non-small cell lung carcinoma	5	3.3	16	0.5	Yes
HBV	5	3.32	16	0	

Predicted association	# Shared Rels	Obs/Exp	Score	Lit_Str	Is this a predicted association for ADTRP also?
serine protease inhibitor	4	4.2	16	19.9	
vitronectin	4	4.15	16	4.9	
ventricular myocytes	4	4.32	16	0	
Dexamethasone	7	2.29	15	0	
IL-8	6	2.69	15	11.7	
Trypsin	6	2.55	15	10	
estrogen receptor	6	2.72	15	2.6	Yes
acute myocardial infarction	5	3.23	15	43.7	
HDL	5	3.15	15	15.1	Yes
glomeruli	5	3.26	15	8.6	
bradykinin	5	3.18	15	6.7	
Osteoarthritis	5	3.23	15	2.3	
PPARgamma	5	3.24	15	0	
Anticoagulant	5	2.8	15	353.7	
FACTOR IX	4	4.06	15	35.5	Yes
proteinase inhibitors	4	4.16	15	25.2	
haemodialysis	4	4.03	15	9.9	

1) 230 aa; Mw ~ 27 kDa

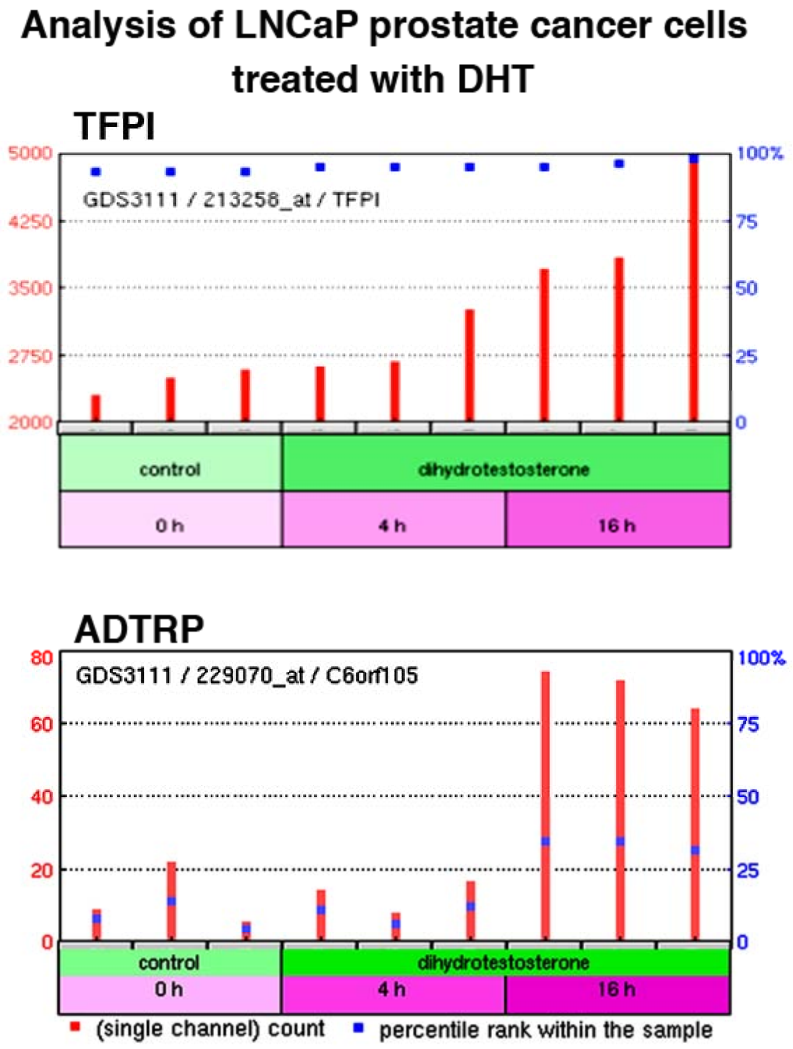


MTKTSTCIYH FLVLSWYTFL NYYISQEGKD EVKPKILANG
ARWKYMTLLN - LLLQTIFYGV TCLDDVLKRT KGGKDIKFLT
AFRDLLFTTL AFPVSTFVFL AFWILFLYNR DLIYPKVLDT
VIPVWLNHAM HTFIFPITLA EVVLRPHSYP SKKTGLTLLA
AASIAYISRI LWLYFETGTW VYPVFAKLSL LGLAAFFSLS
YVFIASIYLL GEKLNHWKWG DMRQPRKKRK

2) 240 aa; Mw ~ 29 kDa

MTKTSTCIYH FLVLSWYTFL NYYISQEGKD EVKPKILANG
ARWKYMTLLN LLKNRTAGFD IYQPGSFRQL LQTIFYGVTC
LDDVLKRTKG GKDIKFLTAF RDLLFTTLAF PVSTFVFLAF
WILFLYNRDL IYPKVLDTVI PVWLNHAMHT FIFPITLAEV
VLRPHSYPSK KTGLTLLAAA SIAYISRILW LYFETGTWVY
PVFAKLSLLG LAAFFSLSYV FIASIYLLGE KLNHWKWGDM
RQPRKKRK

Figure S1. ADTRP isoforms, predicted sequences and domains. According to UniProtKB/Swiss-Prot, the protein encoded by *C6orf105* likely has two variants and 3-6 transmembrane domains. Three different predictions for the transmembrane domains are marked on the sequences: highlighted in yellow, underlined in black and over-lined in green. Possible palmitoylation: **C**; phosphorylation: **T**, **S** and **Y**.



Analysis of LNCaP prostate cancer cells treated with DHT

TFPI

ADTRP

Figure S2. Androgen effect on ADTRP and TFPI mRNA expression

Data mining of GEO array profiles revealed significant up-regulation of both ADTRP and TFPI mRNA expression in the LNCaP prostate cancer cell line after androgen incubation.