

## Supplemental Material

### **Gene Selection For Multiclass Prediction By Weighted Fisher Criterion**

*Jianhua Xuan<sup>1,\*</sup>, Yue Wang<sup>1</sup>, Yibin Dong<sup>1</sup>, Yuanjian Feng<sup>1</sup>, Bin Wang<sup>1</sup>, Javed Khan<sup>2</sup>, Maria Bakay<sup>3</sup>, Zuyi Wang<sup>1,3</sup>, Lauren Pachman<sup>4</sup>, Sara Winokur<sup>5</sup>, Yi-Wen Chen<sup>3</sup>, Robert Clarke<sup>6</sup> and Eric Hoffman<sup>3,\*</sup>*

<sup>1</sup>Dept. of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Arlington, VA 22203, USA, <sup>2</sup>Dept. of Pediatric Oncology, National Cancer Institute, Gaithersburg, MD 20877, USA, <sup>3</sup>Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC 20010, USA, <sup>4</sup>Disease Pathogenesis Program, Children's Memorial Research Center, Chicago, IL 60614, USA, <sup>5</sup>Dept. Of Biological Chemistry, University of California, Irvine, CA 92697, USA, and <sup>6</sup>Lombardi Cancer Center, Georgetown University, Washington, DC 20007, USA

Table S1 lists the top 200 IDGs selected by 1-D wFC from 63 expression profiles of NCI's SRBCTs. As described in the paper, the genes are ranked by their means of 1-D wFC measures using leave-one-out trials on the data set. The table gives a brief description of gene names with gene symbols. Similarly, Table S2 lists the top 100 IDGs selected by 1-D wFC from 39 gene expression profiles of 4 types of muscular dystrophies.

**Table S1.** The top 200 IDGs selected by 1-D wFC from 63 NCI's SRBCT samples.

<b>Image ID</b>	<b>Gene Symbol</b>	<b>Gene Name</b>
461425	MYL4	---
43733	GYG2	glycogenin 2
298417	TFF3	Human secretory protein (P1.B) mRNA, complete cds
866702	PTPN13	protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase)
839736	CRYAB	crystallin, alpha B
307660	FABP4	fatty acid binding protein 4, adipocyte
52076	OLFM1	olfactomedinrelated ER localized protein
345553	TNA	tetranectin (plasminogen-binding protein)
135688	GATA2	GATA-binding protein 2
770394	FCGRT	Fc fragment of IgG, receptor, transporter, alpha
773568	POU4F1	POU domain, class 4, transcription factor 1
769716	NF2	neurofibromin 2 (bilateral acoustic neuroma)
357031	TNFAIP6	tumor necrosis factor, alpha-induced protein 6
809901	COL15A1	collagen, type XV, alpha 1
563673	ALDH7A1	antiquitin 1
364934	DAPK1	death-associated protein kinase 1
365826	GAS1	growth arrest-specific 1
80338	SELENBP1	selenium binding protein 1
113308	KIAA0974	EST
859359	TP53I3	quinone oxidoreductase homolog
502518	LAMB2	laminin, beta 2 (laminin S)
25725	FDFT1	farnesyl-diphosphate farnesyltransferase 1
66552	C20orf194	ESTs
796475	FHL3	ESTs, Moderately similar to skeletal muscle LIM-protein FHL3 [H.sapiens]
345538	CTSL	cathepsin L
784593	ARHE	ESTs
842918	FARP1	chondrocyte-derived ezrin-like protein
784224	FGFR4	fibroblast growth factor receptor 4
754600	NFIX	nuclear factor I/X (CCAAT-binding transcription factor)
244618	FNDC5	ESTs
45542	IGFBP5	Human insulin-like growth factor binding protein 5 (IGFBP5) mRNA
75254	CSRP2	cysteine and glycine-rich protein 2 (LIM domain only, smooth muscle)
745343	REG1A	regenerating islet-derived 1 alpha (pancreatic stone protein, pancreatic thread protein)
1435862	CD99	antigen identified by monoclonal antibodies 12E7, F21 and O13
815239	ARHGEF1	guanine nucleotide exchange factor; 115-kD; mouse Lsc homolog
124605	MINK	cholinergic receptor, nicotinic, epsilon polypeptide
204545	ANTXR1	ESTs
383188	---	recoverin

324494	HSPB2	heat shock 27kD protein 2
142134	ARHGAP8	ESTs
486110	PFN2	profilin 2
293859	TUSC3	Putative prostate cancer tumor suppressor
377461	CAV1	caveolin 1, caveolae protein, 22kD
815284	PEPD	peptidase D
796258	SGCA	sarcoglycan, alpha (50kD dystrophin-associated glycoprotein)
813742	PTK7,XPO6	PTK7 protein tyrosine kinase 7
154472	FGFR1	fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)
295985	CDK6	ESTs
471266	DGCR6L	DiGeorge syndrome critical region 6
132527		ESTs
21652	CTNNA1	catenin (cadherin-associated protein), alpha 1 (102kD)
460487	LTF	lactotransferrin
841641	CCND1	cyclin D1 (PRAD1: parathyroid adenomatosis 1)
142788	SERPINH1	collagen-binding protein 2 (colligen 2)
284001	PRKCM	protein kinase C, mu
897177	---	phosphoglycerate mutase 1 (brain)
770059	---	heparan sulfate proteoglycan 2 (perlecan)
744417	---	carnitine acetyltransferase
486787	CNN3	calponin 3, acidic
809910	IFITM3	interferon-inducible
729964	SMPD1	sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphingomyelinase)
365515	FGF7	fibroblast growth factor 7 (keratinocyte growth factor)
66714	---	Human cell division control-related protein 2b (hcdcrel2b) mRNA, complete cds
784257	KIF3C	kinesin family member 3C
380620	PSEN2	presenilin 2 (Alzheimer disease 4)
841620	DPYSL2	dihydropyrimidinase-like 2
740801	BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)
144932	CDK2AP1	deleted in oral cancer (mouse, homolog) 1
813698	SPRY2	sprouty (Drosophila) homolog 2
377048	MYO1B	Homo sapiens incomplete cDNA for a mutated allele of a myosin class I, myh-1c
308497	HT036	KIAA0467 protein
1048985	FAM12A	---
44563	GAP43	growth associated protein 43
769959	COL4A2	collagen, type IV, alpha 2
782811	HMGA1	high-mobility group (nonhistone chromosomal) protein isoforms I and Y
141768	ERBB2	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (neuro/glioblastoma derived oncogene homolog)
308231	MYO1B	Homo sapiens incomplete cDNA for a mutated allele of a myosin class I, myh-1c
289645	APLP1	amyloid beta (A4) precursor-like protein 1

505491	PTTG1IP	chromosome 21 open reading frame 1
530185	CD83	CD83 antigen (activated B lymphocytes, immunoglobulin superfamily)
245330	IGF2	Human Krueppel-related zinc finger protein (H-plk) mRNA, complete cds
1473131	TLE2	transducin-like enhancer of split 2, homolog of Drosophila E(sp1)
877664	APBB1,COMMD7	ESTs
789253	PSEN2	presenilin 2 (Alzheimer disease 4)
143306	---	lymphocyte-specific protein 1
241412	ELF1	E74-like factor 1 (ets domain transcription factor)
491692	COL4A1	collagen, type IV, alpha 1
767183	HCLS1	hematopoietic cell-specific Lyn substrate 1
782193	LATS2	thioredoxin
203003	NME4	non-metastatic cells 4, protein expressed in
159455	PLD3	similar to vaccinia virus HindIII K4L ORF
752652	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)
68950	---	cyclin E1
308163	YAP1	ESTs
70332	ACP2	acid phosphatase 2, lysosomal
196992	AKR1C1	aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase)
753215	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1
1434905	HOXB7	homeo box B7
810612	S100A11	S100 calcium-binding protein A11 (calgizzarin)
233721	---	insulin-like growth factor binding protein 2 (36kD)
1474955	TAF15	TATA box binding protein (TBP)-associated factor, RNA polymerase II, N, 68kD (RNA-binding protein 56)
813630	PIM1	pim-1 oncogene
45291	DRPLA	dentatorubral-pallidolusian atrophy (atrophin-1)
41591	MN1	meningioma (disrupted in balanced translocation) 1
395708	DPYSL4	
448386	PBX3	
138672		ESTs
212640	ARHGAP4	Rho GTPase activating protein 4
207274	---	Human DNA for insulin-like growth factor II (IGF-2); exon 7 and additional ORF
786084	CBX1	chromobox homolog 1 (Drosophila HP1 beta)
345232	LTA	lymphotoxin alpha (TNF superfamily, member 1)
814526	RNPC1	ESTs
1470048	LY6E	lymphocyte antigen 6 complex, locus E
789204	TLOC1	translocation protein 1
740554	RDX	radixin
760224	XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1
814260	FVT1	follicular lymphoma variant translocation 1
627939	CSRP3	cysteine and glycine-rich protein 3 (cardiac LIM protein)

122159	COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
824704	MPI	mannose phosphate isomerase
950710	PCCA	propionyl Coenzyme A carboxylase, alpha polypeptide
127099	PARD3	ESTs, Moderately similar to atypical PKC specific binding protein [R.norvegicus]
42558	GATM	glycine amidinotransferase (L-arginine:glycine amidinotransferase)
195751	AKAP7	A kinase (PRKA) anchor protein 7
713922	GSTM2	glutathione S-transferase M1
210688	ADCY9	adenylate cyclase 9
82903	TAPBP	TAP binding protein (tapasin)
774502	PTPN12	protein tyrosine phosphatase, non-receptor type 12
811108	TRIP6	thyroid hormone receptor interactor 6
298231	GABBR1	gamma-aminobutyric acid (GABA) B receptor, 1
781047	BUB1	budding uninhibited by benzimidazoles 1 (yeast homolog)
221826	GNA11	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
212542	PBX1	Homo sapiens mRNA; cDNA DKFZp586J2118 (from clone DKFZp586J2118)
755599	IFITM1	interferon-induced protein 17
280837	MGC16824	Human Chromosome 16 BAC clone CIT987SK-A-101F10
280735	TBP	TATA box binding protein
1474684	EFNA1	ephrin-A1
796388	GPC6	general transcription factor IIIA
811000	LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein (galectin 6 binding protein)
40643	PDGFRB	platelet-derived growth factor receptor, beta polypeptide
1469292	PIM2	pim-2 oncogene
451855	PPM1F	'---
823864	TCN2	transcobalamin II; macrocytic anemia
705110	CASP9	caspase 9, apoptosis-related cysteine protease
198982	DNMT2	DNA (cytosine-5-)-methyltransferase 2
1409509	TNNT1	troponin T1, skeletal, slow
1475633	ANXA6	annexin A6
714453	IL4R	interleukin 4 receptor
296448	IGF2	insulin-like growth factor 2 (somatomedin A)
166236	FAM3A	glucose-6-phosphate dehydrogenase
166195	RNH	ribonuclease/angiogenin inhibitor
770868	NAB2,STAT6	NGFI-A binding protein 2 (ERG1 binding protein 2)
111884	TCF7L1	ESTs, Highly similar to HMG box transcription factor [M.musculus]
491565	CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2
346696	TEAD4	TEA domain family member 4
783729	ERBB2	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (neuro/glioblastoma derived oncogene homolog)
379708	CHD3	'---
609663	PRKAR2B	protein kinase, cAMP-dependent, regulatory, type II, beta

208001	CD59	CD59 antigen p18-20 (antigen identified by monoclonal antibodies 16.3A5, EJ16, EJ30, EL32 and G344)
273435	YES1	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1
782503	FADS1	Homo sapiens clone 23716 mRNA sequence
812965	MYC	v-myc avian myelocytomatosis viral oncogene homolog
435953	ITPR3	
823886	MYH10	Smooth muscle myosin heavy chain isoform SMemb [human, umbilical cord, fetal aorta, mRNA Partial, 971 nt]
773246	RING1	ring finger protein 1
825411	HNRPM	N-acetylglucosamine receptor 1 (thyroid)
134748	GCSH	glycine cleavage system protein H (aminomethyl carrier)
1475730	CCT6A	chaperonin containing TCP1, subunit 6A (zeta 1)
358433	RXRG	retinoid X receptor, gamma
753418	---	vasodilator-stimulated phosphoprotein
81518	OCRL	apelin; peptide ligand for APJ receptor
194214		TG-interacting factor (TALE family homeobox)
741831	PLTP	phospholipid transfer protein
39796	HMGCL	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (hydroxymethylglutaricaciduria)
137535	TIF1	transcriptional intermediary factor 1
767495	GLI3	GLI-Kruppel family member GLI3 (Greig cephalopolysyndactyly syndrome)
175123	MFAP4	mitogen-activated protein kinase 7
814080	HDAC1	histone deacetylase 1
80109	HLA-DQA2,HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1
789182	PCNA	proliferating cell nuclear antigen
127509	GBE1	glucan (1,4-alpha-), branching enzyme 1 (glycogen branching enzyme, Andersen disease, glycogen storage disease type IV)
293500	FLJ90440	ESTs
25499	DKFZp564A176	ubiquinol-cytochrome c reductase hinge protein
214572	CDK6	ESTs
810057	CSDA	cold shock domain protein A
665658	MSX2	msh (Drosophila) homeo box homolog 2
810864	CGI-48	ESTs, Highly similar to CGI-48 protein [H.sapiens]
377731	GSTM5	glutathione S-transferase M5
781097	RTN3	neurotrophic tyrosine kinase, receptor-related 1
283315	PGAM2	phosphoglycerate mutase 2 (muscle)
39093	METAP2	methionine aminopeptidase; eIF-2-associated p67
629896	MAP1B	microtubule-associated protein 1B
246722	SERPINE2	CAG repeat domain
45632	GYS1	glycogen synthase 1 (muscle)
416959	NFIB	nuclear factor I/B
323371	APP	amyloid beta (A4) precursor protein (protease nexin-II, Alzheimer disease)
128054	---	ESTs
341588	FBN2,EIF2B3	ESTs, Moderately similar to TRANSLATION INITIATION FACTOR EIF-2B GAMMA SUBUNIT [R.norvegicus]

360047	SMARCD3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3
50941	CDH13	cadherin 13, H-cadherin (heart)

**Table S2.** The top 100 IDGs selected by 1-D wFC from 39 CNMC's MD samples.

Probe Set	Gene Symbol	Gene Name
202966_at	---	---
202965_s_at	CAPN6	calpain 6
202409_at	---	CDNA clone MGC:52263 IMAGE:4123447, complete cds
205422_s_at	ITGBL1	integrin, beta-like 1 (with EGF-like repeat domains)
220968_s_at	PP1057	hypothetical protein PP1057
200665_s_at	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)
202403_s_at	COL1A2	collagen, type I, alpha 2
212320_at	OK/SW-cl.56	beta 5-tubulin
212488_at	COL5A1	collagen, type V, alpha 1
200011_s_at	ARF3	ADP-ribosylation factor 3
200011_s_at	ARF3	ADP-ribosylation factor 3
212489_at	COL5A1	collagen, type V, alpha 1
201527_at	ATP6V1F	ATPase, H <sup>+</sup> transporting, lysosomal 14kDa, V1 subunit F
211161_s_at	COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
204892_x_at	EEF1A1	eukaryotic translation elongation factor 1 alpha 1
202404_s_at	---	---
211734_s_at	FCER1A	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide
215076_s_at	COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
200923_at	LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein
212593_s_at	PDCD4	programmed cell death 4 (neoplastic transformation inhibitor)
213125_at	OLFML2B	olfactomedin-like 2B
212667_at	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)
200678_x_at	---	---
221875_x_at	HLA-F	major histocompatibility complex, class I, F
210794_s_at	MEG3	maternally expressed 3
218144_s_at	C14orf173	chromosome 14 open reading frame 173
37408_at	MRC2	mannose receptor, C type 2
201494_at	PRCP	prolylcarboxypeptidase (angiotensinase C)
206559_x_at	EEF1A1	eukaryotic translation elongation factor 1 alpha 1
208812_x_at	HLA-C	major histocompatibility complex, class I, C
205730_s_at	ABLIM3	actin binding LIM protein family, member 3
213583_x_at	---	MRNA expressed only in placental villi, clone SMAP83.
212082_s_at	MYL6	myosin, light polypeptide 6, alkali, smooth muscle and non-muscle
208834_x_at	RPL23A	ribosomal protein L23a
222280_at	---	CDNA clone IMAGE:6602785, partial cds
208728_s_at	CDC42	cell division cycle 42 (GTP binding protein, 25kDa)
208113_x_at	PABPC1	poly(A) binding protein, cytoplasmic 1
206580_s_at	EFEMP2	EGF-containing fibulin-like extracellular matrix protein 2
200963_x_at	---	---



202310_s_at	COL1A1	collagen, type I, alpha 1
208998_at	UCP2	uncoupling protein 2 (mitochondrial, proton carrier)
215177_s_at	ITGA6	integrin, alpha 6
202760_s_at	AKAP2	A kinase (PRKA) anchor protein 2
201492_s_at	RPL41	ribosomal protein L41
202311_s_at	COL1A1	collagen, type I, alpha 1
202283_at	SERPINF1	serine (or cysteine) proteinase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1
201852_x_at	COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
213477_x_at	EEF1A1	eukaryotic translation elongation factor 1 alpha 1
212094_at	PEG10	paternally expressed 10
221729_at	COL5A2	collagen, type V, alpha 2
218291_at	GALNT14	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 14 (GalNAC-T14)
219213_at	JAM2	junctional adhesion molecule 2
207424_at	MYF5	myogenic factor 5
209320_at	ADCY3	adenylate cyclase 3
202598_at	S100A13	S100 calcium binding protein A13
205381_at	LRRC17	leucine rich repeat containing 17
218192_at	IHPK2	inositol hexaphosphate kinase 2
209888_s_at	MYL1	myosin, light polypeptide 1, alkali; skeletal, fast
200904_at	---	---
213377_x_at	RPS12	ribosomal protein S12
201891_s_at	B2M	beta-2-microglobulin
211991_s_at	HLA-DPA1	major histocompatibility complex, class II, DP alpha 1
201105_at	LGALS1	lectin, galactoside-binding, soluble, 1 (galectin 1)
200959_at	FUS	fusion (involved in t(12;16) in malignant liposarcoma)
201565_s_at	ID2	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
211367_s_at	CASP1	caspase 1, apoptosis-related cysteine protease (interleukin 1, beta, convertase)
204834_at	FGL2	fibrinogen-like 2
209596_at	DKFZp564l1922	adlican
213075_at	OLFML2A	olfactomedin-like 2A
214909_s_at	DDAH2	dimethylarginine dimethylaminohydrolase 2
200953_s_at	CCND2	cyclin D2
203854_at	IF	I factor (complement)
200735_x_at	NACA	nascent-polypeptide-associated complex alpha polypeptide
218678_at	NES	nestin
213080_x_at	RPL5	ribosomal protein L5
211651_s_at	LAMB1	laminin, beta 1
208695_s_at	RPL39	ribosomal protein L39
217025_s_at	DBN1	drebrin 1
201185_at	PRSS11	protease, serine, 11 (IGF binding)
202596_at	ENSA	endosulfine alpha

219049_at	ChGn	chondroitin beta1,4 N-acetylgalactosaminyltransferase
203133_at	SEC61B	Sec61 beta subunit
213422_s_at	MGC3047	hypothetical protein MGC3047
215823_x_at	PABPC1	poly(A) binding protein, cytoplasmic 1
218223_s_at	CKIP-1	CK2 interacting protein 1; HQ0024c protein
203836_s_at	MAP3K5	mitogen-activated protein kinase kinase kinase 5
213048_s_at	---	---
213541_s_at	ERG	v-ets erythroblastosis virus E26 oncogene like (avian)
200911_s_at	TACC1	transforming, acidic coiled-coil containing protein 1
218831_s_at	FCGRT	Fc fragment of IgG, receptor, transporter, alpha
201268_at	NME2	non-metastatic cells 2, protein (NM23B) expressed in
203725_at	GADD45A	growth arrest and DNA-damage-inducible, alpha
211073_x_at	RPL3	ribosomal protein L3
209156_s_at	COL6A2	collagen, type VI, alpha 2
202820_at	AHR	aryl hydrocarbon receptor
209393_s_at	EIF4EL3	eukaryotic translation initiation factor 4E-like 3
218357_s_at	TIMM8B	translocase of inner mitochondrial membrane 8 homolog B (yeast)
200618_at	LASP1	LIM and SH3 protein 1
201217_x_at	RPL3	ribosomal protein L3
208635_x_at	NACA	nascent-polypeptide-associated complex alpha polypeptide
202375_at	SEC24D	SEC24 related gene family, member D (S. cerevisiae)