

# **Supplement Data for: Profiling of Circulating Gene Expression Reveals Molecular Signatures Associated with Intracranial Aneurysm Rupture Risk**

Kerry E. Poppenberg PhD<sup>1,2</sup>, Aichi Chien PhD<sup>5</sup>, Briana A. Santo<sup>1,3</sup>, Lee Chaves PhD<sup>1,2</sup>, Sricharan S. Veeturi<sup>1,4</sup>, Muhammad Waqas MBBS<sup>1,2</sup>, Andre Monteiro MD<sup>1,2</sup>, Adam A. Dmytriw MD MPH<sup>6</sup>, Jan-Karl Burkhardt MD<sup>7</sup>, Maxim Mokin MD-PhD<sup>8</sup>, Kenneth V. Snyder MD-PhD<sup>1,2</sup>, Adnan H. Siddiqui MD-PhD<sup>1,2</sup>, \*Vincent M. Tutino PhD<sup>1,2,3,4</sup>

<sup>1</sup>Canon Stroke and Vascular Research Center; <sup>2</sup>Department of Neurosurgery, <sup>3</sup>Department of Pathology and Anatomical Sciences, <sup>4</sup>Department of Mechanical and Aerospace Engineering, , University at Buffalo, Buffalo, NY, USA

<sup>5</sup>Department of Radiology, University of California Los Angeles, Los Angeles, CA, USA

<sup>6</sup>Neuroendovascular Program, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

<sup>7</sup>Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA, USA

<sup>8</sup>Department of Neurosurgery, University of South Florida, Tampa, FL, USA

## **Email Addresses:**

### **\*Corresponding Author:**

Vincent M. Tutino, PhD

Canon Stroke and Vascular Research Center

Clinical Translational Research Center

875 Ellicott Street

Buffalo, New York 14203 USA

Email: [vincentt@buffalo.edu](mailto:vincentt@buffalo.edu)

Phone: (716) 829-5400

Fax: (716) 854-1850

## SUPPLEMENTAL TABLES

**Supplemental Table 1. Primers used for qPCR.\***

<b>Gene</b>	<b>Primer Sequence</b>	<b>qPCR Prod. Length (bp)</b>
<i>DEFA1</i>	5'-CAAAGCATCCAGGCTCAAGG-3' 5'-CAAGCTCAGCAGCAGAATGC-3'	127
<i>FNI</i>	5'-GGACGCATCACTTGCACTTC-3' 5'-GATGCACTGGAGCAGGTTTC-3'	117
<i>HBA1</i>	5'-GGTCAACTTCAAGCTCCTAAGC-3' 5'-CAAGGGGCAAGAAGCATGG-3'	177
<i>MSLN</i>	5'-ACCTAACATTTCCAGCCTCTCC-3' 5'-GGAATAGCAGCAGGTCCAATG-3'	212
<i>RPL3L</i>	5'-ATTCGGGTCATTGTCCACAC-3' 5'-CCTCACTCTGGCTGAACACG-3'	169
<i>USP9Y</i>	5'-CTCAGCATCTCTTCCAACAGAAC-3' 5'-TCATCCAGGTTTGCCAGCTC-3'	193
<i>HPRT1</i>	5'-TGGCGTCGTGATTAGTGATG-3'	131
<i>Housekeeping</i>	5'-CAAGACGTTCAGTCCTGTCC-3'	
<i>GPI</i>	5'-TCATCATTGCCTCCAAGACC-3'	194
<i>Housekeeping</i>	5'-CACCCAATCCAGAACTCG-3'	

\* Abbreviations: bp=base pair, Prod.=product, qPCR=quantitative polymerase chain reaction

**Supplemental Table 2. Patient and IA Characteristics per Each Risk Stratification.\***

<b>Patient Characteristics</b>	<b>Per Size Stratification</b>		<b>Per PHASES Stratification</b>	
	<i>Low-Risk</i>	<i>High-Risk</i>	<i>Low-Risk</i>	<i>High-Risk</i>
Age (avg. years $\pm$ s.d.)	57.4 $\pm$ 11.6	57.2 $\pm$ 12.6	55.2 $\pm$ 12.6	60.8 $\pm$ 10.3
Female ( <i>n/n<sub>total</sub></i> )	29/37 (78.4%)	21/31 (67.7%)	32/42 (81.0%)	18/26 (69.2%)
Smoking ( <i>n/n<sub>total</sub></i> )	10/37 (27.0%)	9/31 (29.0%)	13/42 (31.0%)	6/26 (23.1%)
Hypertension ( <i>n/n<sub>total</sub></i> )	15/37 (40.5%)	15/31 (48.4%)	13/42 (31.0%)	17/26 (65.4%)
Family history of IA ( <i>n/n<sub>total</sub></i> )	6/37 (16.2%)	3/31 (9.7%)	8/42 (19.1%)	1/26 (3.9%)
<b>Aneurysm Characteristics</b>				
Patients with multiple IAs ( <i>n/n<sub>total</sub></i> )	7/37 (18.9%)	8/31 (25.8%)	8/42 (19.1%)	7/26 (26.9%)
IA size (avg. mm $\pm$ s.d.)	2.63 $\pm$ 0.9	7.04 $\pm$ 3.3	3.52 $\pm$ 1.73	6.15 $\pm$ 4.05
IA location ( <i>n/n<sub>total</sub></i> )				
ACom	5/49 (10.2%)	4/40 (10.0%)	4/52 (7.7%)	5/37 (13.5%)
BA	4/49 (8.2%)	3/40 (7.5%)	3/52 (5.8%)	4/37 (10.8%)
ICA	23/49 (46.9%)	21/40 (52.5%)	35/52 (67.3%)	9/37 (24.3%)
MCA	7/49 (14.3%)	6/40 (15.0%)	5/52 (9.6%)	8/37 (21.6%)
PCom	10/49 (20.4%)	6/40 (15.0%)	5/52 (9.6%)	11/37 (29.7%)

\*These clinical factors were retrieved from patients' medical records. With the exception of age, these datapoints were quantified as binary data points. (Abbreviations: IA=intracranial aneurysm, n=number, avg.=average, s.d.=standard deviation, ACom=anterior communicating artery, BA=basilar artery, ICA=internal carotid artery, MCA=middle cerebral artery, PCom=posterior communicating artery, PHASES=population, hypertension, age, size of aneurysm, earlier SAH from another aneurysm, site of aneurysm)

**Supplemental Table 3. RNA Quality and Sequencing QC Metrics.\***

<b>Sample ID</b>	<b>260/280</b>	<b>RIN</b>	<b>M Seqs</b>	<b>% Ass.</b>	<b>M Ass.</b>	<b>% Aligned</b>
UIA1	2.00	8.46	71.9	22.8%	24.1	96.5%
UIA2	2.00	8.46	71.9	18.1%	20.9	96.3%
UIA3	2.01	9.59	62.9	23.6%	20.6	95.6%
UIA4	1.85	8.25	49.3	20.6%	13.7	95.9%
UIA5	2.03	9.49	50.8	23.7%	18.3	96.5%
UIA6	1.98	9.49	52.2	21.8%	17.3	96.5%
UIA7	1.98	9.59	49.3	27.7%	17.6	95.0%
UIA8	1.66	7.40	36.3	20.2%	10.4	95.8%
UIA9	1.87	6.70	65.6	4.0%	8.0	98.0%
UIA10	1.93	6.50	51.4	22.6%	15.0	96.8%
UIA11	2.04	9.38	54.9	24.0%	18.4	96.7%
UIA12	2.06	8.90	66.3	21.9%	19.3	96.8%
UIA13	2.01	9.80	44.7	24.9%	14.5	96.4%
UIA14	1.92	7.60	44.2	25.3%	15.1	96.2%
UIA15	1.92	6.00	56.5	22.1%	14.5	95.1%
UIA16	1.95	6.80	47.7	28.4%	16.3	96.2%
UIA17	1.97	9.20	49.0	27.7%	19.2	95.5%
UIA18	2.04	8.90	46.2	23.0%	15.1	96.4%
UIA19	2.02	7.20	69.6	24.2%	20.2	96.1%
UIA20	1.95	6.10	70.5	9.3%	15.6	97.4%
UIA21	1.96	8.90	71.3	12.0%	18.3	97.2%
UIA22	1.96	6.10	33.0	23.0%	9.3	96.2%
UIA23	1.96	7.84	59.3	22.4%	15.9	95.4%
UIA24	2.03	9.38	53.3	22.5%	18.0	95.6%
UIA25	2.06	9.00	51.3	25.2%	17.9	96.2%
UIA26	2.03	9.10	66.2	26.8%	25.2	96.4%
UIA27	2.02	9.00	59.9	26.6%	21.5	96.1%
UIA28	1.97	7.90	51.0	24.1%	17.5	96.6%
UIA29	2.04	9.38	69.6	25.2%	24.1	96.3%
UIA30	2.02	9.07	34.0	24.6%	9.7	97.1%
UIA31	1.92	8.66	35.6	26.5%	10.8	97.4%
UIA32	2.03	9.07	38.7	26.3%	11.7	97.4%
UIA33	2.07	8.56	69.4	23.7%	22.9	96.6%
UIA34	2.03	8.77	40.2	25.5%	12.8	97.1%
UIA35	2.06	9.28	35.3	24.6%	10.4	97.1%
UIA36	2.06	9.18	40.6	25.0%	12.3	97.1%
UIA37	1.97	8.15	33.6	23.5%	9.2	97.2%

UIA38	2.03	8.87	40.8	25.7%	12.8	96.8%
UIA39	1.97	9.69	26.1	22.5%	7.3	96.9%
UIA40	2.04	8.97	40.7	25.3%	12.2	96.9%
UIA41	2.05	9.07	20.0	27.0%	6.6	96.9%
UIA42	2.05	8.56	44.2	26.9%	14.2	97.2%
UIA43	1.97	8.66	35.2	25.2%	10.4	97.0%
UIA44	1.99	8.66	37.2	25.8%	11.3	97.0%
UIA45	2.05	8.35	30.3	28.0%	10.0	97.4%
UIA46	2.05	8.77	34.8	25.9%	10.7	97.1%
UIA47	2.06	8.46	33.9	25.4%	10.5	97.0%
UIA48	2.04	8.15	34.2	26.6%	10.5	97.2%
UIA49	2.03	8.15	38.6	26.4%	12.4	96.7%
UIA50	1.77	8.15	32.0	25.3%	10.0	96.9%
UIA51	2.04	9.18	45.4	24.2%	13.7	96.9%
UIA52	1.82	8.04	48.9	28.1%	16.6	96.7%
UIA53	2.04	8.04	46.3	24.7%	14.0	97.1%
UIA54	2.06	8.87	42.9	22.5%	11.9	96.8%
UIA55	2.01	8.97	37.6	25.5%	11.5	97.2%
UIA56	2.09	7.63	32.2	30.7%	12.6	96.6%
UIA57	1.98	8.15	38.4	26.3%	12.4	97.4%
UIA58	2.04	8.97	31.9	27.0%	10.5	96.9%
UIA59	1.95	8.56	51.1	23.2%	14.3	96.7%
UIA60	1.05	7.94	34.0	28.9%	12.6	96.7%
UIA61	2.01	7.94	35.9	23.0%	9.8	97.3%
UIA62	2.05	8.66	37.9	24.6%	11.6	96.5%
UIA63	2.05	8.46	31.2	28.0%	11.6	95.3%
UIA64	1.77	8.87	33.3	27.8%	11.9	96.7%
UIA65	2.05	8.56	34.1	25.7%	10.7	97.0%
UIA66	2.06	7.73	39.0	26.0%	12.1	96.7%
UIA67	1.98	8.10	66.3	24.6%	20.4	96.8%
UIA68	2.01	8.77	66.6	21.6%	19.7	96.0%

\* “Assigned” refers to alignments assigned to genome features, whereas “aligned” refers to aligned reads that overlap genes. Abbreviations: Ass.=assigned, QC=quality control, RIN=RNA integrity number, M=million

**Supplemental Table 4. Differentially Expressed Genes Associated with IA Risk as Defined by IA size or PHASES score.\***

Gene Name	Ensembl ID	Size		PHASES	
		Log2(F-C)	q-value	Log2(F-C)	q-value
<i>A4GALT</i>	ENSG00000128274.15			1.292	0.000
<i>ABCA13</i>	ENSG00000179869.14			-1.080	0.000
<i>ABO</i>	ENSG00000175164.13	1.159	0.002		
<i>ACOT12</i>	ENSG00000172497.8	-1.268	0.001		
<i>ADAMTS2</i>	ENSG00000087116.13			-2.019	0.000
<i>ADORA1</i>	ENSG00000163485.15			-1.273	0.001
<i>ANKRD61</i>	ENSG00000157999.5			-1.583	0.002
<i>AVPR2</i>	ENSG00000126895.13			1.041	0.001
<i>BTNL3</i>	ENSG00000168903.8			-1.447	0.000
<i>C10orf55</i>	ENSG00000222047.8			-1.181	0.022
<i>C1orf53</i>	ENSG00000203724.10			1.203	0.001
<i>C4BPA</i>	ENSG00000123838.10	-1.427	0.000		
<i>C4orf47</i>	ENSG00000205129.8	1.058	0.001		
<i>C9orf129</i>	ENSG00000204352.2			1.261	0.000
<i>CFAP46</i>	ENSG00000171811.12			1.296	0.000
<i>CHRNA7</i>	ENSG00000175344.16			1.050	0.014
<i>CHSY3</i>	ENSG00000198108.3	1.012	0.029		
<i>CLDN22</i>	ENSG00000177300.6	1.426	0.000		
<i>CLGN</i>	ENSG00000153132.12			1.106	0.001
<i>CMBL</i>	ENSG00000164237.8			1.245	0.000
<i>CNTNAP3B</i>	ENSG00000154529.14			1.083	0.008
<i>COL17A1</i>	ENSG00000065618.16			-1.250	0.000
<i>CRISPLD1</i>	ENSG00000121005.8			-1.359	0.001
<i>CSMD1</i>	ENSG00000183117.17	1.564	0.000		
<i>CTNNA2</i>	ENSG00000066032.18			-1.671	0.011
<i>CYP4B1</i>	ENSG00000142973.12	-1.024	0.001		
<i>DDX3Y</i>	ENSG00000067048.16			2.168	0.000
<i>DEFA1</i>	ENSG00000206047.2	-1.087	0.008	-2.221	0.000
<i>DEFA3</i>	ENSG00000239839.5			-1.390	0.000
<i>DEFA4</i>	ENSG00000164821.4			-1.367	0.000
<i>DNER</i>	ENSG00000187957.7			-1.657	0.001
<i>EDIL3</i>	ENSG00000164176.12	-1.038	0.000	1.586	0.000
<i>ELF5</i>	ENSG00000135374.9			1.227	0.000
<i>ELN</i>	ENSG00000049540.16	1.094	0.000		
<i>ESPN</i>	ENSG00000187017.14			1.195	0.000
<i>FAM107A</i>	ENSG00000168309.16			1.062	0.004

<i>FAM181B</i>	ENSG00000182103.4			-1.400	0.000
<i>FAM3B</i>	ENSG00000183844.16			1.639	0.002
<i>FAM83A</i>	ENSG00000147689.16			1.605	0.000
<i>FN1</i>	ENSG00000115414.18	-2.397	0.000	-1.920	0.000
<i>FOXJ1</i>	ENSG00000168269.8			1.293	0.005
<i>GAD1</i>	ENSG00000128683.13	-1.026	0.005		
<i>GDF15</i>	ENSG00000130513.6			1.201	0.000
<i>GLYATL2</i>	ENSG00000156689.6			1.310	0.000
<i>GSTM1</i>	ENSG00000134184.12			1.282	0.003
<i>HBA1</i>	ENSG00000206172.8	1.005	0.018	1.016	0.032
<i>HBA2</i>	ENSG00000188536.12	1.055	0.002	1.097	0.001
<i>HBG1</i>	ENSG00000213934.6			1.743	0.000
<i>HOXB9</i>	ENSG00000170689.9			1.553	0.000
<i>IFI27</i>	ENSG00000165949.12	-2.112	0.000		
<i>IGSF9</i>	ENSG00000085552.16			1.102	0.001
<i>IL2</i>	ENSG00000109471.4			1.212	0.038
<i>KCNJ1</i>	ENSG00000151704.15	1.256	0.000		
<i>KLF14</i>	ENSG00000266265.2	-1.066	0.001		
<i>LITD1</i>	ENSG00000240563.1			1.325	0.000
<i>LOXL4</i>	ENSG00000138131.3			1.111	0.000
<i>LPL</i>	ENSG00000175445.14	-1.249	0.000		
<i>MMP8</i>	ENSG00000118113.11			-1.524	0.000
<i>MMRN2</i>	ENSG00000173269.13	1.255	0.000		
<i>MSLN</i>	ENSG00000102854.14	-1.626	0.000	-1.156	0.000
<i>MYOM2</i>	ENSG00000036448.9	-1.029	0.004		
<i>NAALAD2</i>	ENSG00000077616.10			1.034	0.002
<i>NPAS1</i>	ENSG00000130751.9	-1.347	0.000		
<i>NPC1L1</i>	ENSG00000015520.14			1.338	0.012
<i>NXF3</i>	ENSG00000147206.16	-1.310	0.005		
<i>NYAP2</i>	ENSG00000144460.11	-1.144	0.018		
<i>OLFM4</i>	ENSG00000102837.6			-1.436	0.000
<i>OTOF</i>	ENSG00000115155.16	-1.461	0.000	2.070	0.000
<i>PALM</i>	ENSG00000099864.17			1.055	0.000
<i>PGF</i>	ENSG00000119630.13			1.220	0.000
<i>PKDCC</i>	ENSG00000162878.12			-1.207	0.006
<i>PRSS42</i>	ENSG00000178055.11			1.258	0.000
<i>PRTG</i>	ENSG00000166450.12	-1.167	0.008		
<i>PVRL2</i>	ENSG00000130202.9			-1.282	0.000
<i>RBFOX3</i>	ENSG00000167281.18			1.174	0.001
<i>RIMBP2</i>	ENSG00000060709.13			1.004	0.000

<i>RNF152</i>	ENSG00000176641.10			1.190	0.000
<i>RNF182</i>	ENSG00000180537.12			1.164	0.000
<i>ROBO1</i>	ENSG00000169855.19	-1.015	0.000		
<i>RP11-347C12.1</i>	ENSG00000198064.13			1.218	0.000
<i>RPL3L</i>	ENSG00000140986.7	1.061	0.017	1.413	0.000
<i>RUNDC3A</i>	ENSG00000108309.12			1.520	0.000
<i>SCGB3A1</i>	ENSG00000161055.3	-1.072	0.004		
<i>SFRP2</i>	ENSG00000145423.4	1.002	0.000	1.266	0.000
<i>SLC16A9</i>	ENSG00000165449.11			1.254	0.000
<i>SLC39A5</i>	ENSG00000139540.11			1.085	0.030
<i>SLC51A</i>	ENSG00000163959.9			-1.060	0.000
<i>SMIM6</i>	ENSG00000259120.2			1.537	0.000
<i>SPERT</i>	ENSG00000174015.9	1.283	0.022		
<i>ST8SIA5</i>	ENSG00000101638.13			1.111	0.000
<i>SYTL5</i>	ENSG00000147041.11	-1.106	0.000		
<i>TDRD12</i>	ENSG00000173809.15			-1.201	0.000
<i>TEKT3</i>	ENSG00000125409.12			1.002	0.004
<i>TMPRSS9</i>	ENSG00000178297.12			1.137	0.000
<i>TMSB4Y</i>	ENSG00000154620.5			1.479	0.026
<i>TNFSF18</i>	ENSG00000120337.8	1.260	0.000		
<i>TUSC3</i>	ENSG00000104723.20	1.429	0.000		
<i>USP9Y</i>	ENSG00000114374.12	1.536	0.000	2.261	0.000
<i>UTS2</i>	ENSG00000049247.13			1.436	0.000
<i>UTY</i>	ENSG00000183878.15			2.760	0.000
<i>ZFPM2</i>	ENSG00000169946.13			1.170	0.035
<i>ZFY</i>	ENSG00000067646.11			1.660	0.006

\*Significantly differentially expressed transcripts with q-value<0.05, an absolute fold-change $\geq$ 2, and expression in at least 50% of samples. 36 genes were differentially expressed in IA risk analysis according to aneurysm size, 76 genes were differentially expressed in IA risk analysis according to PHASES score, and 10 genes were identified according to both, which are highlighted in grey. (Abbreviations: F-C=fold-change)



**Supplemental Table 5. Significant IPA Disease and Biological Functions for the IA Risk-Associated Differentially Expressed Genes.\***

<b>Risk Stratification by IA Size</b>				
<b>Categories</b>	<b>Functions</b>	<b>Diseases or Functions Annotation</b>	<b>q-value</b>	<b>Molecules</b>
Cardiovascular System Development and Function	development	Development of vasculature	0.0279	DEFA1, EDIL3, ELN, FN1, LPL, MMRN2, ROBO1, SCGB3A1, SFRP2
Cardiovascular Disease	occlusion	Occlusion of blood vessel	0.0279	ABO, CSMD1, ELN, FN1, HBA1/HBA2, LPL
Cardiovascular Disease, Organismal Injury and Abnormalities	vaso-occlusion	Vaso-occlusion	0.0279	ABO, CSMD1, ELN, FN1, HBA1/HBA2, LPL
Cardiovascular Disease, Organismal Injury and Abnormalities	atherosclerosis	Atherosclerosis	0.0302	CSMD1, ELN, FN1, HBA1/HBA2, LPL
Neurological Disease, Organismal Injury and Abnormalities	cerebral disorder	Cerebral disorder	0.0368	ACOT12, C4BPA, CHSY3, ELN, FN1, GAD1, LPL, MSLN, NPAS1, OTOF, PRTG, ROBO1, RPL3L, SFRP2
Cellular Movement	chemotaxis	Chemotaxis	0.0395	DEFA1, ELN, FN1, ROBO1, TNFSF18
Cardiovascular System Development and Function, Organismal Development	angiogenesis	Angiogenesis	0.0414	DEFA1, EDIL3, ELN, FN1, ROBO1, SCGB3A1, SFRP2
Cellular Movement	cell movement	Cell movement of myeloid cells	0.0461	DEFA1, EDIL3, ELN, FN1, TNFSF18
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking, Inflammatory Response	cell movement	Cell movement of phagocytes	0.0473	DEFA1, EDIL3, ELN, FN1, TNFSF18
Cardiovascular System Development and	vasculogenesis	Vasculogenesis	0.0483	EDIL3, ELN, FN1, ROBO1, SCGB3A1, SFRP2

Function, Organismal Development				
----------------------------------	--	--	--	--

**Risk Stratification by PHASES Score**

<b>Categories</b>	<b>Functions</b>	<b>Diseases or Functions Annotation</b>	<b>B-H p-value</b>	<b>Molecules</b>
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking	cell movement	Cell movement of granulocytes	0.0392	ADORA1, CHRNA7, EDIL3, FN1, GDF15, IL2, MMP8, OLFM4, PGF
Cellular Movement	cell movement	Cell movement of myeloid cells	0.0392	ADORA1, CHRNA7, DEFA1, EDIL3, FN1, GDF15, IL2, MMP8, OLFM4, PGF, UTS2
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking	cell movement	Cell movement of leukocytes	0.0422	ADORA1, CHRNA7, DEFA1, EDIL3, FN1, GDF15, IL2, MMP8, OLFM4, PGF, UTS2
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking, Inflammatory Response	cell movement	Cell movement of phagocytes	0.0422	ADORA1, CHRNA7, DEFA1, EDIL3, FN1, GDF15, IL2, MMP8, PGF, UTS2
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking	cellular infiltration	Cellular infiltration by granulocytes	0.0429	CHRNA7, EDIL3, MMP8, OLFM4, PGF
Cellular Movement, Hematological System Development and Function, Immune Cell	cell movement	Cell movement of neutrophils	0.0459	ADORA1, EDIL3, FN1, GDF15, MMP8, PGF

Trafficking, Inflammatory Response				
Cardiovascular System Development and Function, Organismal Development	angiogenesis	Angiogenesis	0.0469	ADAMTS2, CHRNA7, DEFA1, EDIL3, FN1, GSTM1, IL2, MMP8, PGF, SFRP2, UTS2, ZFPM2
Cardiovascular Disease, Organismal Injury and Abnormalities	atherosclerosis	Atherosclerosis	0.0479	ADORA1, FN1, GDF15, HBA1/HBA2, MMP8, NPC1L1, PGF
Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization	cell-cell contact	Cell-cell contact	0.0479	CHRNA7, CLGN, FN1, GDF15, IGSF9, IL2, NECTIN2, OTOF, PALM, PGF
Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking	cellular infiltration	Cellular infiltration by leukocytes	0.0479	CHRNA7, EDIL3, FN1, IL2, MMP8, OLFM4, PGF
Cellular Assembly and Organization	organization	Organization of cellular membrane	0.0479	A4GALT, AVPR2, CHRNA7, IGSF9, NECTIN2

\*q-value from Benjamini-Hochberg correction. Abbreviations: IPA=Ingenuity Pathway Analysis, IA=intracranial aneurysm, PHASES=population, hypertension, age, size of aneurysm, earlier SAH from another aneurysm, site of aneurysm

**Supplemental Table 6. IPA Networks.**

<b>ID</b>	<b>Molecules in Network</b>	<b>Score</b>	<b>Focus Molecules</b>	
<b>Size</b>				
1	ACOT12, Akt, ANGPTL1, C4BPA, caspase, CD3, DEFA1, EDIL3, ELN, ERK1/2, FN1, GAD1, HBA1/HBA2, Histone h3, IFI27, IFN Beta, Immunoglobulin, Insulin, Jnk, KCNJ1, LPL, MIR320, MSLN, NFkB (complex), NPAS1, NYAP2, PI3K (complex), Pka, Pkc(s), ROBO1, SCGB3A1, SFRP2, TNFSF18, TUSC3, Vegf	50	19	Embryonic Development, Organismal Development, Tissue Development
2	ABO, APRT, AR, ARHGAP35, ARHGEF40, CHSY3, cldn, CLDN22, CSF2RA, CSMD1, ELP3, FOXP3, H4C2, HIPK1, ICAM1, IGF receptor, IGF1R, JAM3, KLF14, Leukocyte Extravasation Signaling, MMP19, MYC, MYOM2, NEU3, NXF3, PAICS, PRTG, PSMD5, RPL24, RPL3L, SYTL5, VIRMA, WAC, WDR44, ZNF148	22	10	Organ Development, Tissue Development, Tissue Morphology
<b>PHASES</b>				
1	A4GALT, Actin, Akt, Alpha catenin, caspase, CD3, Cdc2, COL17A1, CTNNA2, ELF5, ESPN, F Actin, FAM107A, Focal adhesion kinase, HBA1/HBA2, HBG1, hemoglobin, HOXB9, Ige, IgG, IL2, Immunoglobulin, Jnk, Mek, MSLN, NECTIN2, NPC1L1, PI3K (complex), PI3K (family), PKDCC, SFRP2, SRC (family), TCR, Vegf, ZFPM2	34	16	Connective Tissue Disorders, Organismal Development, Organismal Injury and Abnormalities
2	ADAMTS2, ADCY, Ap1, CHRNA7, Collagen type I (complex), Collagen type II, Collagen type III, Collagen type IV, Collagen(s), DEFA1, DEFA4, DNER, EDIL3, elastase, ERK1/2, FAM3B, Fibrin, Fibrinogen, FN1, GDF15, Growth hormone, GSTM1, IL1, Metalloprotease, Mmp, MMP8, PALM, PDGF BB, PGF, Pka, PLC, Pro-inflammatory Cytokine, Rap1, Tgf beta, UTS2	29	14	Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking

\*Abbreviations: IPA=Ingenuity Pathway Analysis, PHASES=population, hypertension, age, size of aneurysm, earlier SAH from another aneurysm, site of aneurysm

**Supplemental Table 7. Code for Running LDA Models.**

<b>Action</b>	<b>Code</b>
Standard Libraries (Numpy, Pandas, Scipy, etc)	<pre>import scipy import pandas as pd import numpy as np from scipy import stats import statsmodels.api as sm import statistics</pre>
Machine Learning Libraries (sklearn)	<pre>from sklearn.model_selection import train_test_split, GridSearchCV,     LeaveOneOut from sklearn.feature_selection import VarianceThreshold, f_classif,     GenericUnivariateSelect, mutual_info_classif, SelectKBest from sklearn.decomposition import PCA from sklearn.linear_model import LogisticRegression from sklearn.neighbors import KNeighborsClassifier from sklearn import tree from sklearn.metrics import accuracy_score, roc_curve, roc_auc_score,     confusion_matrix, classification_report from sklearn.svm import LinearSVC, SVC from sklearn.discriminant_analysis import LinearDiscriminantAnaly-     sis, QuadraticDiscriminantAnalysis</pre>
Post-Processing Li- braries	<pre>import matplotlib import matplotlib.pyplot as plt import seaborn as sns import qgrid %matplotlib inline</pre>
Import and Initialize Data	<pre>df = pd.read_csv('DESTINATION-FILE', delimiter = ',') cases = df['Sample']</pre>
Set the index for fea- tures as the sample names	<pre>features = df.loc[:, 'GENE-FIRST': 'GENE-LAST'] features = features.set_index(df['Sample'])</pre>
Set the index for la- bels as the sample names	<pre>labels = df['Outcome'] labels.index = df['Sample']</pre>
Run models	<pre>for i in range(100): seed=i np.random.seed(seed)  trainX , testX , trainY, testY = train_test_split(features, labels, test_size     = 20, random_state = seed, stratify = labels ) trainX_standardized = trainX testX_standardized = testX</pre>

---

```
clf_lda = LinearDiscriminantAnalysis(solver='eigen', shrinkage='auto',
    priors=None, n_components=None, store_covariance=False,
    tol=0.0001)
solver=['eigen', 'lsqr']
shrinkage=['auto']
priors=[None]
n_components=[None]
store_covariance=[False]
tol=[0.0001]

param_grid = dict(solver=solver, shrinkage=shrinkage, priors=priors ,
    n_components=n_components, store_covariance=store_covariance
    , tol=tol)
loo = LeaveOneOut() ## Use Leave-One-Out Cross-validation
grid_standardized = GridSearchCV(estimator = clf_lda,
    param_grid=param_grid, n_jobs=-2, cv=loo)
grid_result_standardized = grid_standardized.fit(trainX_standardized,
    trainY)
predictions_standardized , predictions_standardized_probs =
    grid_standardized.predict(testX_standardized) , grid_standard-
    ized.predict_proba(trainX_standardized)
test_accuracy_standardized = accuracy_score(testY, predictions_stand-
    ardized)
auc_standardized = roc_auc_score(trainY,predictions_standard-
    ized_probs[:,1])

classification_report=classification_report(testY, predictions_standard-
    ized)

print("Seed #: " + seed)
print('Results using All Features')
print('All features AUC='+ str(auc_standardized))
print('All features Testing Accuracy='+ str(test_accuracy_standard-
    ized))
print(classification_report)
```

---

**Supplemental Table 8: MANCOVA Results.\****IA Size*

<b>Gene</b>	<b>Significant Covariate</b>	<b>q-value</b>
ABO	-	-
ACOT12	-	-
C4BPA	-	-
C4orf47	IA family history	0.014
	Smoking	0.034
CHSY3	-	-
CLDN22	-	-
CSMD1	-	-
CYP4B1	-	-
DEFA1	-	-
EDIL3	-	-
ELN	IA family history	0.029
FN1	-	-
GAD1	-	-
HBA1	-	-
HBA2	-	-
IFI27	-	-
KCNJ1	IA family history	0.012
KLF14	IA family history	0.021
	Hypertension	0.025
	Smoking	0.041
LPL	Hypertension	0.027
MMRN2	IA family history	0.017
MSLN	-	-
MYOM2	-	-
NPAS1	-	-
NXF3	-	-
NYAP2	-	-
OTOF	-	-
PRTG	-	-
ROBO1	-	-
RPL3L	-	-
SCGB3A1	-	-
SFRP2	-	-
SPERT	IA family history	0.047
SYTL5	-	-
TNFSF18	-	-

TUSC3	IA family history	<0.001
USP9Y	-	-

***PHASES***

<b>Gene</b>	<b>Significant Covariate</b>	<b>q-value</b>
A4GALT	Sex	<0.001
ABCA13	-	-
ADAMTS2	-	-
ADORA1	-	-
ANKRD61	-	-
AVPR2	-	-
BTNL3	-	-
C10orf55	-	-
C1orf53	IA family history	0.024
C9orf129	-	-
CFAP46	-	-
CHRNA7	Sex	0.04
CLGN	-	-
CMBL	Sex	<0.001
	Hypertension	0.029
CNTNAP3B	-	-
COL17A1	-	-
CRISPLD1	-	-
CTNNA2	-	-
DDX3Y	Sex	<0.001
DEFA1	-	-
DEFA3	-	-
DEFA4	-	-
DNER	-	-
EDIL3	-	-
ELF5	-	-
ESPN	Sex	0.012
	Smoking	0.031
FAM107A	-	-
FAM181B	-	-
FAM3B	-	-
FAM83A	-	-
FN1	-	-
FOXI1	-	-
GDF15	-	-

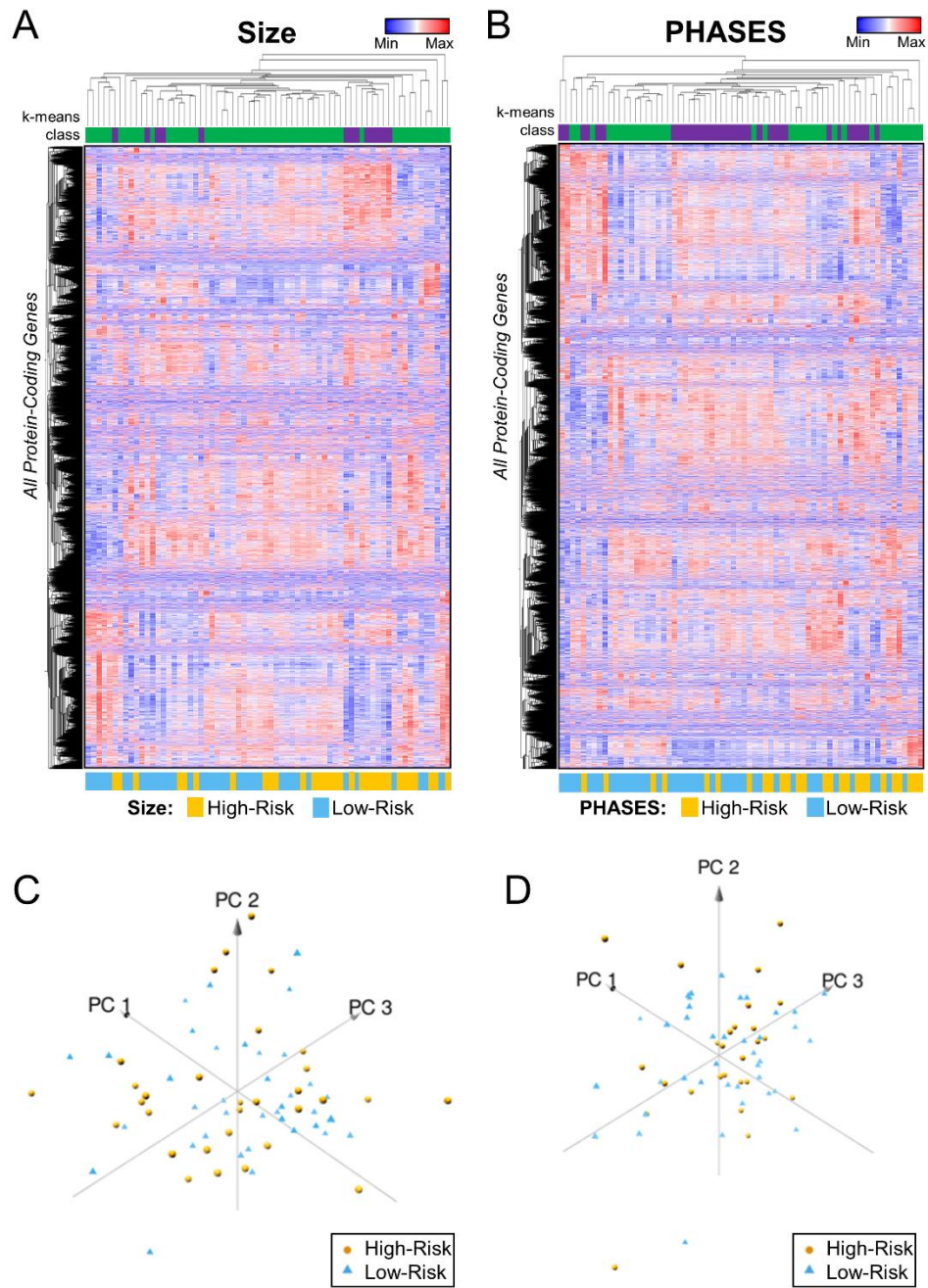


GLYATL2	Smoking	0.011
GSTM1	Sex	0.033
HBA1	-	-
HBA2	-	-
HBG1	-	-
HOXB9	-	-
IGSF9	IA family history	0.042
IL2	-	-
L1TD1	Hypertension	0.012
LOXL4	-	-
MMP8	-	-
MSLN	-	-
NAALAD2	-	-
NPC1L1	-	-
OLFM4	-	-
OTOF	Sex	0.048
PALM	-	-
PGF	Sex	0.005
PKDCC	Smoking	0.006
PRSS42	-	-
PVRL2	-	-
RBFOX3	-	-
RIMBP2	-	-
RNF152	-	-
RNF182	-	-
RP11-347C12.1	-	-
RPL3L	-	-
RUNDC3A	-	-
SFRP2	-	-
SLC16A9	-	-
SLC39A5	-	-
SLC51A	-	-
SMIM6	-	-
ST8SIA5	-	-
TDRD12	-	-
TEKT3	Smoking	0.026
TMPRSS9	Sex	0.005
TMSB4Y	Sex	<0.001
	IA family history	0.007

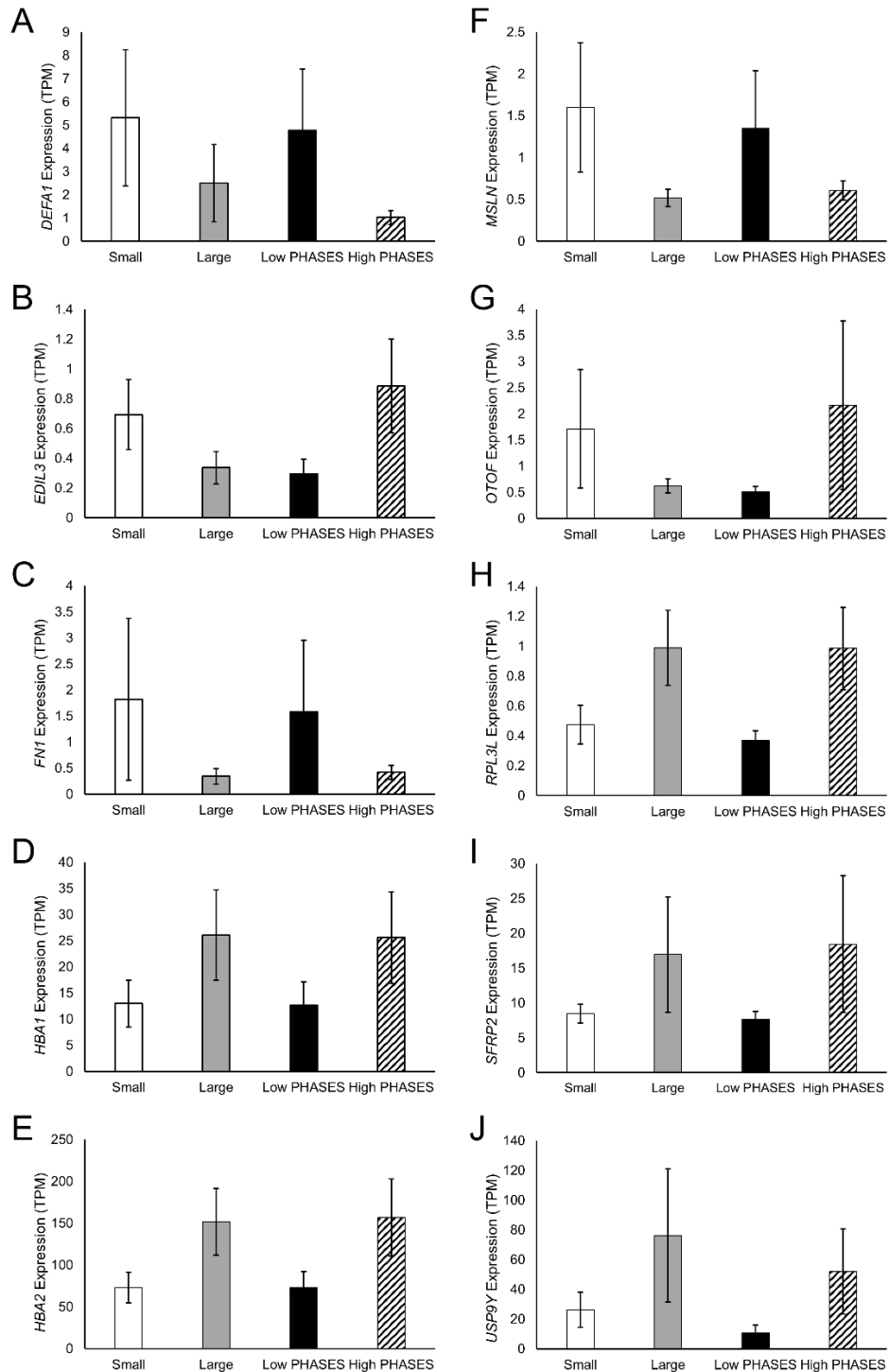
USP9Y	Sex	0.001
	Smoking	0.026
UTS2	Sex	0.001
	IA family history	0.02
UTY	Sex	0.018
ZFPM2	-	-
ZFY	Sex	<0.001
	IA family history	0.033

\* For each gene, potentially confounding patient characteristics were evaluated for significance based on f-MANCOVA group-covariate interaction terms. Correction for multiple testing was also completed, and significance was defined as  $q\text{-value} < 0.05$ . Abbreviations: MANCOVA=multivariate analysis of covariance, DEG=differentially expressed gene, PHASES=population, hypertension, age, size, earlier subarachnoid hemorrhage, site; IA=intracranial aneurysm

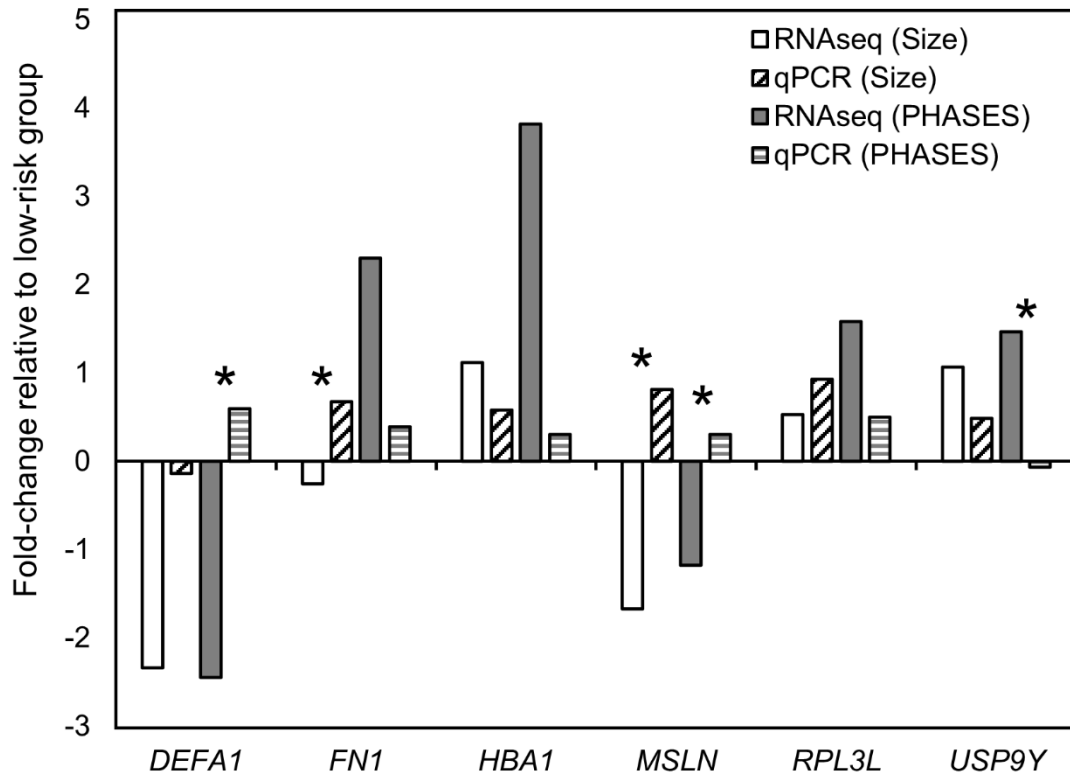
## SUPPLEMENTAL FIGURES



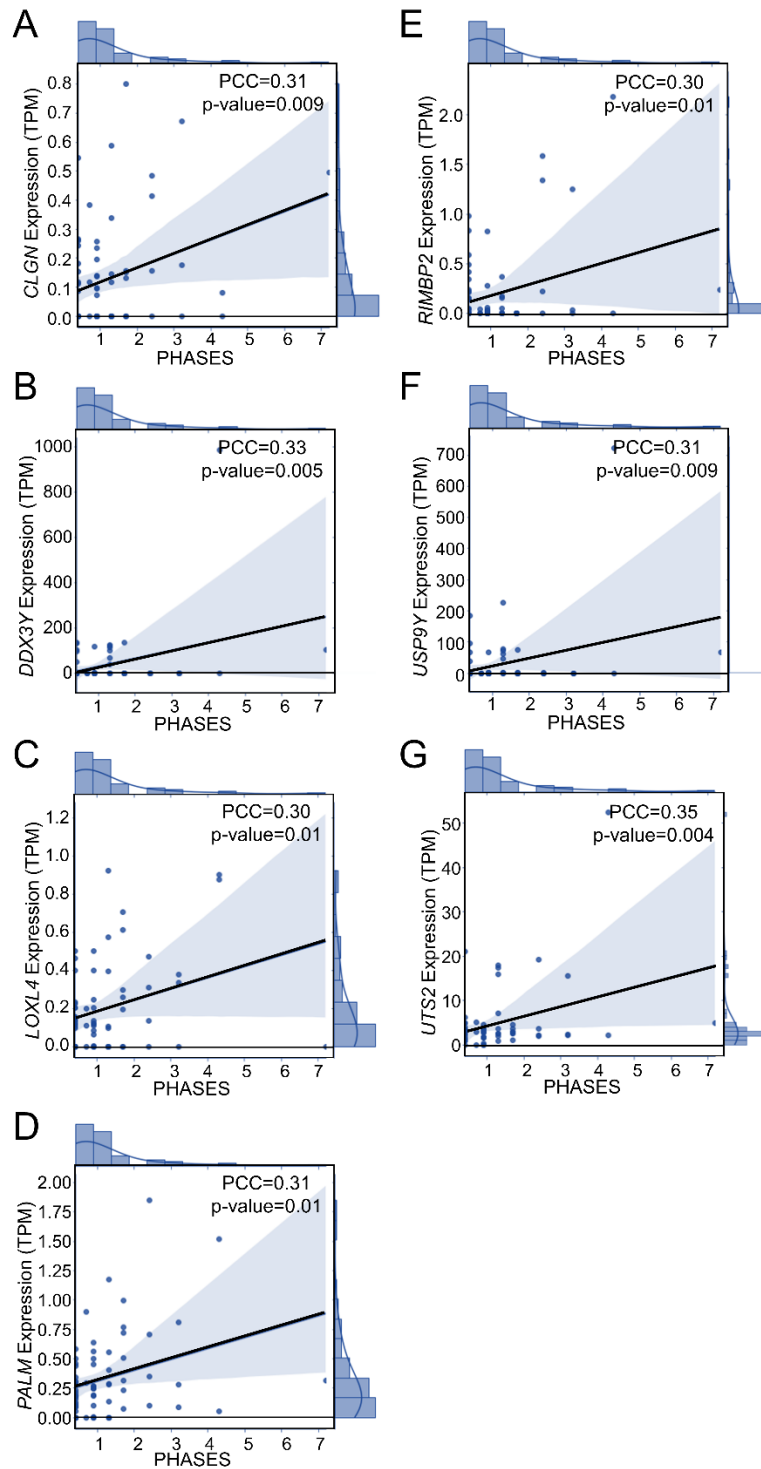
**Supplemental Figure 1. Visualization of Protein-Coding Gene Expression.** Heatmaps generated using TPM expression data of all protein-coding genes show poor grouping of low- and high-risk cases whether risk was classified by **(A)** aneurysm size or **(B)** PHASES score. K means clustering was generally inconclusive. Three-dimensional principal component analysis demonstrates separation between low- and high-risk samples based on both **(C)** aneurysm size and **(D)** PHASES score also showed only limited grouping of low- and high-risk cases. Abbreviations: Min=minimum, Max=maximum, PC=principal component, PHASES=Population, Hypertension, Age, Size, Earlier subarachnoid hemorrhage, Site



**Supplemental Figure 2. Plots of Expression Levels for the 10 Common DEGs.** The average TPM expression level ( $\pm$  standard error) for *DEFA1* (A), *EDIL3* (B), *FN1* (C), *HBA1* (D), *HBA2* (E), *MSLN* (F), *OTOF* (G), *RPL3L* (H), *SFRP2* (I), *USP9Y* (J) in each group (small IA, large IA, low PHASES, and high PHASES).



**Supplemental Figure 3. Verification of RNAseq by qPCR in a Subset of 30 Cases.** For high- vs. low-risk using aneurysm size and PHASES, expression differences in 6 of the 10 shared differentially expressed genes (*DEFA1*, *FN1*, *HBA1*, *MSLN*, *RPL3L*, and *USP9Y*) were generally in the same direction of fold-change and of similar magnitude when comparing RNAseq and qPCR data. The one main exception was *MSLN*, whose expression difference was opposite on qPCR for both comparisons. “\*” indicates differences between RNAseq and qPCR. Abbreviations: qPCR=quantitative polymerase chain reaction, RNAseq=RNA sequencing, PHASES=Population, Hypertension, Age, Size, Earlier subarachnoid hemorrhage, Site



**Supplemental Figure 4. Additional Correlation Plots.** The other 7 genes with significant correlations to PHASES. **A).** *CLGN* has a PCC of 0.31. **B).** *DDX3Y* with a PCC of 0.33. **C).** *LOXL4* achieved a PCC of 0.30. **D).** *PALM* was positively correlated with a PCC of 0.31. **E).** *RIMBP2* also was weakly correlated with a PCC of 0.30 **F).** *USP9Y* had a PCC of 0.31. **G).** *UTS2* with a PCC of 0.35. Abbreviations: PCC=Pearson correlation coefficient, TPM=transcripts per million, PHASES=Population, Hypertension, Age, Size, Earlier subarachnoid hemorrhage, Site