CELL JOURNAL (Yakhteh)

www.celljournal.org

Supplementary Information for

GPX2 and *BMP4* as Significant Molecular Alterations in The Lung Adenocarcinoma Progression: Integrated Bioinformatics Analysis

Mohammad Hossein Derakhshan Nazari, B.Sc.[1#], Rana Askari Dastjerdi, B.Sc.[1#], Parnian Ghaedi Talkhouncheh, B.Sc.[2#], Ahmad Bereimipour, M.Sc.[3], Hamidreza Mollasalehi, Ph.D.[2], Amir Ali Mahshad, B.Sc.[1], Ali Razi, B.Sc.[4], Mohammad Hossein Nazari, M.Sc.[5], Amin Ebrahimi Sadrabadi, Ph.D.[3*], Sara Taleahmad, Ph.D.[3*]

1. Department of Microbiology and Microbial Biotechnology, Faculty of Life Science and Biotechnology, Shahid Beheshti University, Tehran, Iran
2. Department of Cell and Molecular Biology, Faculty of Life Science and Biotechnology, Shahid Beheshti University, Tehran, Iran
3. Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran
4. Biophysics Department, Science Faculty, York University, Toronto, Canada
5. V. Zelman Institute for Medicine and Psychology, Novosibirsk State University, Novosibirsk, Russia

**Corresponding Address: P.O. Box: 16635-148, Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran*

Emails: amin.ebrahimi@royaninstitute.org, s.taleahmad@royan-rc.ac.ir

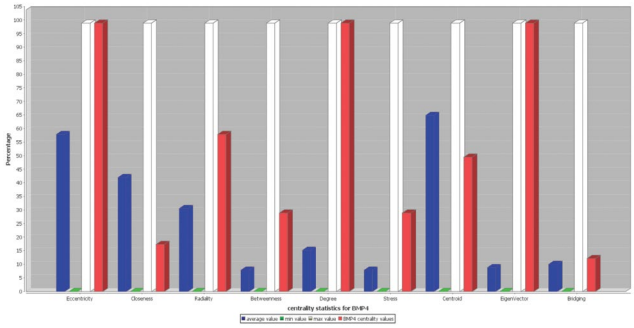

Fig.S1: BMP4 as the critical protein in a network manifested by CentiScape plugin.

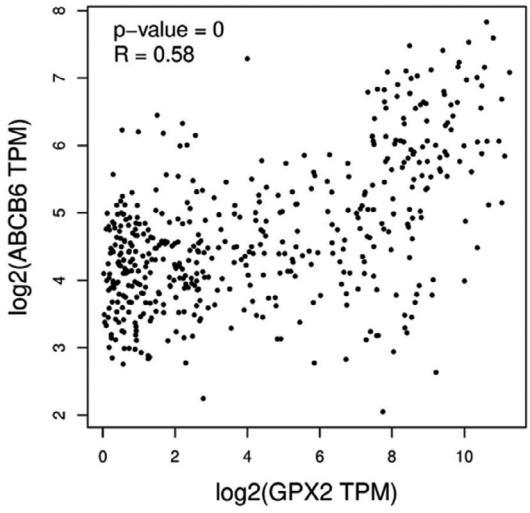

Fig.S2: The positive expression between *GPX2 and ABCB6 has* indicated by Pearson correlation calculation.

Table S1: Selected samples

Stages	Data series						Total samples
	GSE41271		GSE42127		GSE75037		
IA	4 samples	GSM1012836 GSM1012840 GSM1012843 GSM1012848	4 samples	GSM1032885 GSM1032889 GSM1032892 GSM1032895	4 samples	GSM1941127 GSM1941129 GSM1941149 GSM1941153	12
IB	4 samples	GSM1012832 GSM1012833 GSM1012837 GSM1012838	4 samples	GSM1032881 GSM1032882 GSM1032886 GSM1032887	4 samples	GSM1941123 GSM1941125 GSM1941133 GSM1941141	12
IIA	4 samples	GSM1012918 GSM1012900 GSM1012868 GSM1013011	4 samples	GSM1032910 GSM1032932 GSM1032945 GSM1032982	3 samples	GSM1941227 GSM1941247 GSM1941265 -	11
IIB	4 samples	GSM1012847 GSM1012852 GSM1012856 GSM1012869	4 samples	GSM1032898 GSM1032902 GSM1032913 GSM1032914	4 samples	GSM1941121 GSM1941137 GSM1941139 GSM1941155	12
IIIA	4 samples	GSM1012828 GSM1012844 GSM1012845 GSM1012851	4 samples	GSM1032893 GSM1032901 GSM1032907 GSM1033052	4 samples	GSM1941131 GSM1941195 GSM1941199 GSM1941231	12
IIIB	4 samples	GSM1012830 GSM1012835 GSM1012891 GSM1012917	4 samples	GSM1032884 GSM1032957 GSM1033020 GSM1033040	1 sample	GSM1941239 - - -	9
IV	4 samples	GSM1013100 GSM1013060 GSM1013003 GSM1013014	1 sample	GSM1033019 - - -	2 samples	GSM1941135 GSM1941145 - -	7

GSM is the accession code for each chip in the GEO database.

Table S2: Commonly differentially expressed genes (DEG) achieved by the Venn diagram

	Commonly DEGs	
Groups	**Downregulated**	**Upregulated**
Group 1	*ITLN2/STXBP6/CA4/SPTBN1/BMP4/AGER/CDO1/TMPRSS3/ TMEM100/LIPH/TCEAL2/MT3/PROX1/BEX3/CLUL1/CDKN2B/ ARMCX2/KTN1/RHOJ/ARHGAP28/NTNG1/NT5DC1/GPR137C/ LYVE1/CCNG1/CA2/ARHGAP24/CCDC50/PTPRD/AIDA/RCOR2/ ZAK/SHPRH/CSRNP3/LRCH2/ADGRL2/DACH1/SNX14/B3GALNT1/ MPDZ*	*MT1X/MT2A/SNORD46/SNORD55/PTPRE/ TOM1/SNORD56/COL9A2/SLC34A2/VDR/KRT7/ CPM/NOC4L/NAGA/IRF7/DPP9/VWA1/B3GAT3/ GPR132/PYCARD/ZNRD1/RASSF4/SLC52A3/ C1orf229/PTK7/MT1A/TNFAIP2/SNORD33/GPX4*
Group 2	*RAB37/RCC1*	*CLDN10/BEX3/PCSK9*
Group 3	*RPA4/FBXW7/SLC7A2/SFMBT2/NFAT5/UGGT1/PTPRA/NFAT 2IP*	*FKBP2/TRAPPC2B*
Group 4	*NGEF/AK4/CLIP2*	*GFRA3*
Group 5	*MDK/HOXB7/CDH26*	*CDH26*
Group 6	*TMEM27/ZNF136*	*GPX2/NQO1/OSBP2/GAD1/MISP/TRIP13*

Table S3: Heatmap signature genes (|Log FC|>2)

Gene name	Expression level	Progression level	Log FC
GPX2	Upregulated	Stage IIIB to IV	6.596
CLDN10	Upregulated	Stage IB to IIA	3.916
ITLN2	Downregulated	Stage IA to IB	3.83
TMEM27	Downregulated	Stage IIIB to IV	3.51
GAD1	Upregulated	Stage IIIB to IV	3.216
MT1X	Upregulated	Stage IA to IB	3.07
GFRA3	Upregulated	Stage IIB to IIIA	3.02
CA4	Downregulated	Stage IA to IB	2.603
MT2A	Upregulated	Stage IA to IB	2.506
TMPRSS3	Downregulated	Stage IA to IB	2.503
STXBP6	Downregulated	Stage IA to IB	2.44
SLC7A2	Downregulated	Stage IIA to IIB	2.36
AGER	Downregulated	Stage IA to IB	2.313
BMP4	Downregulated	Stage IA to IB	2.306
OSBP2	Upregulated	Stage IIIB to IV	2.273
SNORD46	Upregulated	Stage IA to IB	2.27
NQO1	Upregulated	Stage IIIB to IV	2.206
AK4	Downregulated	Stage IIB to IIIA	2.17
TMEM100	Downregulated	Stage IA to IB	2.05

Table S4: Dysregulated genes play EMT promoter role

Data series	Positive regulation of EMT	Negative regulation of EMT
GSE41271	*BMP4*	
GSE42127	*AXIN2/BMP4/BAMBI/TRIM28/Fgfr1*	*SPRY2/SPRED1/HPN*
GSE75037	*DLG5/FLNA/MDK/MYOCD*	*SFRP2/PTEN/FOXA1/ADIPOR1*

Downregulation of the DEGs contribute to negative regulation of EMT, while the upregulation of DEGs involve in the positive regulation of EMT leads to EMT and metastasis. EMT; Epithelial-to-mesenchymal transition and DEGs; Differentially expressed genes.

Table S5: Key proteins in the protein-protein interaction (PPI) network

Protein	Degree	Betweenness centrality
BMP4	5	0.7
NQO1	4	0.66
RHOJ	3	1
ZNRD1	3	0.66
GPX4	3	0.55