Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Co—expression analysis to identify core regulatory genes of colon cancer stem cells

Yaoqun Wang[1], WeiGuo Zhou[1], Mengyu Sun[1], Chenyu Song[1], Jieping Qiu[1], Feng Zhang[1], Fuchun Huang[1], Yan Liu[2], and Bo Chen[3]

[1]Anhui Medical University
[2]Anhui Medical University School of Public Health
[3]First Affiliated Hospital of Anhui Medical University

March 30, 2022

Abstract

Abstract Objective:Cancer stem cells play an important role in the the drug resistance of malignant tumor and the prognosis, however,the genes that regulate Colon Adenocarcinoma(COAD) stem cells are still unknown. We aim to identify COAD stem cell regulatory genes by co-expression network analysis. Methods:The mRNA expression-based stemness index(mRNAsi)which derives from TCGA database was analyzed based on R language. The Weighted Gene Co-expression Network Analysis(WGCNA) was used to construct co-expression network, establish the correlation between gene modules and modules trait–mRNAsi , and determine hub genes. We use Enrichment analysis of genes function and Protein Interaction(PPI network)to identifying the hub modules and genes. GEO and OnComine databases were used to validate the hub genes. Results: MRNAsi was highly expressed in tumor tissues and the result was statistically significant (p = 7.458E-18), and was positively correlated with the survival rate (p =0.016). 6501 genes were screened(FDR<0.05 and|LogFC|[?]1)for co-expression analysis and 17 modules were obtained, among which the yellow module had the most significant positive correlation with mRNAsi(p=1e-55,correlation coefficient =0.71).We finally selected 9 genes (BUB1, BUB1B, CHEK1, KIF23, MCM10, NCAPG, NCAPH, PLK4, TTK) from the yellow module as the hub genes. The above genes had a significantly high expression in all 13 tumors in the OnComine database. Conclusion: The 9 hub genes play a crucial role in the development of COAD and therapeutic resistance.Finding ways to suppress these genes may be a new target for colon cancer therapy by affecting COAD stem cells.

Hosted file

`Co---expression analysis to identify core regulatory genes of colon cancer stem cells.doc` available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

`figure1.doc` available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

`figure2.doc` available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

`figure3.doc` available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Hosted file

figure4.doc available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

figure5.doc available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

figure6.doc available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

figure7.doc available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Hosted file

figure8.doc available at https://authorea.com/users/471183/articles/562819-co-expression-analysis-to-identify-core-regulatory-genes-of-colon-cancer-stem-cells

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Table 1. GS and MM values of 9 hub genes

ID	GS.mRNAsi	GS.EREG-mRNAsi	MMyellow
NCAPH	0.613353421	0.325467733	0.821312946
CHEK1	0.582452561	0.464886186	0.839509485
NCAPG	0.558220580	0.345243228	0.802558880
BUB1	0.556408265	0.298069420	0.849763135
TTK	0.539711178	0.419365956	0.800986375
PLK4	0.532396805	0.329219409	0.835884468
MCM10	0.531199272	0.315709004	0.826856921
KIF23	0.517140026	0.219189131	0.814318444
BUB1B	0.514994764	0.267168357	0.824199671

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Table 2. Function and pathway enrichment analysis of all genes in yellow module.

Term	Description	Count	pvalue
GO:0007059	chromosome segregation	61	6.76E-41
GO:0006260	DNA replication	58	4.88E-39
GO:0000280	nuclear division	60	5.28E-33
GO:0048285	organelle fission	62	2.46E-32
GO:0000819	sister chromatid segregation	41	1.02E-30
GO:0098687	chromosomal region	68	4.96E-43
GO:0000775	chromosome, centromeric region	47	1.03E-34
GO:0000793	condensed chromosome	47	3.84E-33
GO:0000776	kinetochore	38	1.35E-30
GO:0000779	condensed chromosome, centromeric region	35	7.98E-30
GO:0140097	catalytic activity, acting on DNA	25	1.69E-13
GO:0008094	DNA-dependent ATPase activity	15	4.04E-10
GO:0004386	helicase activity	19	8.71E-10
GO:0016887	ATPase activity	25	8.98E-08
GO:0003678	DNA helicase activity	10	9.06E-08
hsa04110	Cell cycle	32	9.89E-24
hsa03030	DNA replication	14	9.49E-14
hsa03460	Fanconi anemia pathway	12	9.08E-09
hsa03008	Ribosome biogenesis in eukaryotes	16	9.89E-09
hsa03440	Homologous recombination	10	6.28E-08
hsa04115	p53 signaling pathway	11	2.11E-06

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Table 3. Function and pathway enrichment analysis of 9 hub genes.

Term	Description	Count	pvalue
GO:0000819	sister chromatid segregation	6	3.39E-11
GO:0098813	nuclear chromosome segregation	6	2.61E-10
GO:0140014	mitotic nuclear division	6	4.09E-10
GO:0000793	condensed chromosome	5	1.59E-08
GO:0098687	chromosomal region	5	2.20E-07
GO:0000775	chromosome, centromeric region	4	1.13E-06
GO:0004674	protein serine/threonine kinase activity	5	5.70E-07
GO:0035173	histone kinase activity	1	0.007402587
GO:0004712	protein serine/threonine/tyrosine kinase activity	1	0.021021126
KEGG:hsa04110	Cell cycle	4	8.36E-07

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Supplementary Table 1. number of genes contained in each module

Modules	Freq
black	319
blue	937
brown	929
cyan	169
green	467
greenyellow	167
grey	26
grey60	82
lightcyan	83
magenta	204
midnightblue	96
pink	278
purple	345
red	412
salmon	141
turquoise	1318
yellow	528

Posted on Authorea 30 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164865072.22810408/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Supplementary Table 2 .connectivity of nine key genes in protein network

Number	Gene	Degree of connectivity
1	BUB1	36
2	BUB1B	34
2	TTK	23
4	CHEK1	22
5	NCAPG	21
6	KIF23	15
7	MCM10	14
8	NCAPH	8
9	PLK4	7