bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Exploring genes of rectal cancer for new treatments based on protein interaction network

Running Title:protein interaction network of rectal cancer

Wenjing Teng[1] , Chao Zhou[2], , Yan Li[1*]

1 .Department of oncology, Shanghai Municipal Hospital of Traditional Chinese Medicine ,Affiliated to Shanghai University of Traditional Chinese Medicine, Zhijiang road 274 , Shanghai 200041,China.2.Department of oncology, Weifang Traditional Chinese Medicine Hospital,Kuiwen District, Weifang , Shandong Province 261041, China;

***Corresponding author: E-mail：18916767226@163.com.**

Abstract

Objective: To develop a protein-protein interaction network of rectal cancer, which is based on genetic genes as well as to predict biological pathways underlying the molecular complexes in the network. In order to analyze and summarize genetic markers related to diagnosis and prognosis of rectal cancer.

Methods: the genes expression profile was downloaded from OMIM (Online Mendelian Inheritance in Man) database; the protein-protein interaction network of rectal cancer was established by Cytoscape; the molecular complexes in the network were detected by Clusterviz plugin and the pathways enrichment of molecular complexes were performed by DAVID online and Bingo (The Biological Networks Gene Ontology tool).

Results and Discussion: A total of 127 rectal cancer genes were identified to differentially express in OMIM Database. The protein-protein interaction network of rectal cancer was contained 966 nodes (proteins), 3377 edges (interactive relationships) and 7 molecular complexes (score>7.0). Regulatory effects of genes and proteins were focused on cell cycle, transcription regulation and cellular protein metabolic process. Genes of *DDK1*, *sparcl1*, *wisp2*, *cux1*, *pabpc1*, *ptk2* and *htra1* were significant nodes in PPI network. The discovery of featured genes which were probably related to rectal cancer, has a great significance on studying mechanism, distinguishing normal and cancer tissues, and exploring new treatments for rectal cancer.

Keywords: Protein-protein interaction networks; rectal cancer; molecular complexes; pathway ,

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genes

Rectal cancer is one of the common fatal malignant tumors in the world, which is secondly lethal cancer in United States as well as thirdly in Europe. The research has improved systemic treatment, such as micro diagnosis. However, despite advances in detection and care, morbidity and mortality from rectal cancer continues to be high[1].Early detection and diagnosis can be great significance to reduce mortality and improve prognosis, as well as identify those who were at the highest risk duo to improving triage for treatment has the greatest impact on rectal cancers treatment. Studies shown that rectal cancer has a prime paradigm for cancer genetics, which can be prevention by early detection of the pre-disease (neoplastic) state. Therefore, the roles of genetics in rectal cancer have became critical to the missions of disease prevention, early detection and effective treatment.

As report goes, approximately 10% of well defined hereditary rectal cancer has syndromes[2]. Hereditary rectal has some benefits, such as many of it are growing slowly, family members are armed with the knowledge of potential risk of associated carcinomas and empowerment to reduce the disease burden.

A large number of mutation, mismatch, inactivation of tumor suppressor genes and activations of oncogenes are involved in the process of rectal cancer[4].Up to now, genetic marker of rectal cancer can be summarized by the following six aspects：genomic instability, CpG island methylator phenotype, specific microRNAs, histone modification, gene mutation and protein biomarkers. The protein-protein interaction network which is a model of biological molecular interactions can more clearly show the genes, proteins and pathogenesis in the development process of the disease[5]. These genetic traits may partially explain the geographical variance in rectal cancer incidence and mortality as well as the differences between hereditary and sporadic rectal cancer [6].

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 Materials and Methods

Design: Study on enrichment of genomic biological pathways.

Method:

Data acquisition: OMIM(Online Mendelian Inheritance in Man) is a comprehensive,authoritative, daily updated human phenotype database, containing more than 12000 genes of all human genetic diseases, and mainly focusing on hereditary diseases. In addition, text messages, related reference information, sequence records, maps, and related databases are available for each gene[7-8]. This study had started from August 22(2015), searched "rectal cancer"in the OMIM database and obtained human genes associated with rectal cancer information.

The construction of gene/protein interaction networks: rectal cancer associated genes were submitted to Cytoscape 3.2.1 plug-in Agilent Literature Search 2.7.7 (USA Agilent Technologies company) and Pubmed [9]. False positive interaction information was removed from retrieval results. Then, gene/protein interaction relations were read in Cytoscape 2.8.2 and visualized [10].

Network analysis: MODE algorithm in Cytoscape 3.2.1 web analytics plug-in Clusterviz of 1.2 was administrated to make the correlation analysis for the area of the construction of biological networks [11-12]. By analyzing the network structure, proteins were grouped to form molecular compounds in the entire network and were shown in Cytoscape according to the correlation integral value. The areas with integral value higher than 3 were regarded as molecular compounds. The gene/protein names contained in the molecular compounds were submitted to The Database for Annotation, Visualization and Integrated Discovery. [13-14] By retrieving Kyoko Encyclopedia of Genes and Genomes (KEGG) Database, biological pathways involved in chronic myelogenous leukaemia heredity were identified. Then the Biological pathways data were submitted to **Bingo** (in Networks Gene Ontology tool) for enrichment analysis.

Main outcome measures: Protein networks were constructed based on the rectal cancer-related genes, nodes (proteins) and edges (interaction between), molecular complexes in the network and its associated interaction points and nodes (protein) and the edge (interaction between), analyze the biological pathways has involved in the molecular complexes.

2 Results

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.1 Rectal **cancer related genes in OMIM**

According the OMIM database retrieval, it can found that 127 genes were reported to be associated with rectal cancer. After screening and deleting duplicate genes, 127 related genes were identified, which shown in **Table1.**

2.2 Protein interaction networks

Through text mining,127 genetic-related genes shown that there was a network diagram with 638 nodes (proteins) and 1830 edges. As shown in **Figure 1-3**, the diamond is represented OMIM genetic disease related proteins, while the round represented the proteins obtained from text mining.

Figure 1(A+B). Network map of protein interaction (Overall)

As shown in figure 1，the protein-protein interaction network underlining rectal cancer is extremely complex. The edges intersect with each other and several clusters emerge in the figure. The more edges among the genes, the correlation of the genes is more tightening. So the genes formed the round network at the top of figure1 connected with each other more tightly than the genes in the bottom.

Figure 2 the relationship of proteins

After removing the protein molecule in the network，the relationship between proteins in network became more clear, in the centre part of protein-protein network, the relationship between protein complexes emerged is closer, however, at the farther edge of the network, the relationship of it became looser, as shown in **Figure2**.

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3(A): pathway jak2-stat1

Figure 3(B): relationship of gene *Kras*

AS shown in **Figure3,** it can be found some common genes and pathways such as jak2-stat1,Kras, P53,etc as well as genes associated with them.

2.3 Network topology attribute analysis

Network topology attribute analysis shown that the connectivity of nodes in the network (the number of nodes in the network) obeys descending distribution, i.e. with the increase of edges connected to the node, correspondingly the number of nodes decrease , so it can be seen that the gene - protein interaction networks are scale-free networks [15]. We found that the connectivity of nodes in the network greater than or equal to 25 corresponds to a sharp reduction in the number of nodes **Figure 4**. Therefore, we regarded the nodes which the connectivity is greater than / equal to 20 as the key nodes (hub). Key nodes (connectivity score)included ;hif1a(25), cdkn1b(25), rb1(26), plau(26), brca1(27), asns(28), pcna(28),Vegfa(29), stat1(29), cdkn2a(29), egfr(30), tp53(71).

Figure 4. Connectivity degree of each node and betweenness Comparison

(The horizontal axis represents between ness, and the ordinate represents the connectivity degree of protein interaction network. And the graphic in the table represents each node in the network). It can be seen that the connectivity (the number of nodes in the network) of nodes in the network obeys descending distribution, while the connectivity is greater than / equal to 25, the number of nodes corresponding to a sharp decrease.

2.4 The detection of molecular complexes

According MCOMD algorithm analysis, analyse the correlation between genes in network and calculate the score, the number of nodes, the number of edges. There is a total of 74 molecular complexes and 7 of them showed correlation integral values higher than7.0.

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5.

Rank1 score20 Node20 Edges190

Rank2 score20 Node20 Edges190

Rank3 score14 Node14 Edges91

Rank4 score11.662 Node78 Edges449

Rank5 score3.143 Node21 Edges66

Rank6 score9 Node9 Edges36

Rank7 score7.673 Node56 Edges221

Figure 5 Molecular complexes obtained by MCOMD algorithm analysis

The details of the relationship of proteins in molecular complexes 7 was shown in Figure 6

Figure6 the relationship of proteins in molecular complexes 7

The whole network has a total of six parts, including mitogen-activated protein kinase, Specific microRNAs, SRY-related HMG-box, matrix metalloproteinases, cyclinD.

2.5Molecular complex pathway enrichment

Submit the 7 names of protein molecule complexes online to obtain the relevant pathways in **Table 2**.The Protein molecule biological pathways of complexes 1 are not exist,and other complexes contain different pathways.Regulatory effects of genes and proteins mainly focused on cell cycle, transcription regulation, and cellular protein metabolic process.

Table 2 Enrichment of pathways related to molecular complexes

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Binggo results show the gene oncology hierarchical network to the biological processes,**Figure 7**.The size of the node represents the number of the genes, the depth of the node color represents the P values.Diagrams have presented the main biological processes of cluster2,3,5,containing metabolic regulation, transcriptional regulation, biosynthesis, cell differentiation and gene expression regulation and signal transduction, etc.

Figure7

A Bingo results show that the basic structure of the biological cluster1

B Bingo results show that the basic structure of the biological cluster4

C Bingo results show that the basic structure of the biological cluster6

C Bingo results show that the basic structure of the biological cluster7

3 Discussions

Based on a large number of research references, the progression of rectal cancer is a multistep process, containing overcome apoptosis, inhibiting senescence (infinite proliferation), secretion proliferation signal itself, not sensitive to growth signals, angiogenesis and invasion ,which requires a large number of genes and proteins in action[16].Genetic markers of rectal cancer can be summarized from the following six aspects：he first part is genomic instability, close to 12-17% of sporadic rectal cancer cases exist with microsatellite instability (MSI). At present, microsatellite instability - high - H (MSI) has become a positive prognosis of rectal cancer patients' overall survival[17]. 50-85% of Rectal cancer patients has chromosome aberration frequency, chromosome instability (CIN +) usually be associated with colorectal cancer patients' overall survival, progression-free survival (PFS) and poor prognosis after 5 - fluorouracil therapy; the second part is CpG island methylator phenotype. Nearly 29.6% of the rectal cancer patients shown CpG island methylation phenotype - high (CIMP - H) [18], however, its value is still in research. Specific microRNAs and histone modification are another two aspects, which mainly associate with colorectal cancer patients' overall survival and progression-free serial, tumor metastasis, local

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

invasion, tumor volume, tumor staging, treatment results, relapse and drug resistance.

What is more, gene mutations and protein biomarkers have brought a special significance. The APC gene mutant p.D1822V containing homozygous V/V would reduce the risk of rectal cancer[19]. According to the APC (rs565453 y rs1816769) and CTNNB1 (rs229303) gene polymorphisms, the death risk stratification in patients with rectal cancer can be analysed. The APC gene mutant p.I1307K as risk factors of the rectal cancer among Ashkenazi Jewishes has been found[20]. The loss of PTEN gene in 22% of rectal cancer patients has caused no response to the EGFR inhibitors and higher risk of death [21]. And the clinical application of protein molecular markers colorectal cancer mainly included: early diagnosis(hnRNP A1,kininogen-1,adipophilin, Apo AI and C9,OLFM4), Prognosis(SM3,desmin,surviving,hTERT and NM23), Potential therapeutic targets(EB1)[22-23]. Therefore, exploring the protein interaction network of tumors, analysing the characteristics of the protein molecular interaction network, and excavating a variety of signalling pathway and the genes will provide a biological basis for the study of the molecular mechanism of tumor and the further treatment as well.

Bio-molecular network analysis is an important direction of systematic biology research. Large-scale human protein-protein network can provide new insights into protein functions, pathways, molecular machines and functional protein modules [24]. The function of bio-molecular often depends on modularization, the network module is made by a number of nodes in the conjunction of each other and has a stable structure which can often reflect a similar nature between the nodes [25,26]. Analysing the function module is the one of the most common method to analyse biological molecular network. According to 127 genes provided by OMIM, our research has built up protein interaction network of rectal cancer, which contain 996 nodes (proteins), 3377 edges (interaction). Due to the network is very large, the experimental introduced MCOMD algorithm to evaluate the network 's regional integration through the correlation integral. Correlation integral descripted the proteins associated with the degree in the region. Proteins of the same molecular complex generally have the same biological function, therefore we can discover the unknown gene functions or new molecular functional groups, such as cluster 1.

Gene *DDK1*,*sparcl1*,*wisp2*,*cux1*,*pabpc1*,*ptk2* and *htra1* as the composition members of the

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cluster1(score 20) in the centre, have closely relationships with each other as well as other genes. It has reported that the secreted protein acidic and rich in cysteine-like 1 (,*sparcl1*) is expressed in various normal tissues and many types of cancers. Another study has shown that marker *sparcl1* was significantly related to the prognosis and clinical pathological features of the CRC patients [27]. S*parcl1* expression increased with RT and is related to a better prognosis in rectal cancer patients with RT but not in patients without RT. This result may help us to select the patients to the best suited preoperative RT. In the process of rectal cancer, gene WISP2 knockout significantly increased Caco-2 cell invasion and motility. Up-regulation of MMP2, -7 and -9 may indicate that WISP2regulates invasion and motility through MMPs. Regulation of invasion by WISP2 may involve the WNT signalling pathway[28]. Thus, some studies have identified CUX1 as a pan-driver of tumorigenesis and uncover a potential strategy for treating CUX1-mutant tumors. So from cluster1 we predict the effect of gene Cut1 to rectal cancer may be that CUX1 deficiency activates phosphoinositide 3-kinase (PI3K) signalling through direct transcriptional downregulation of the PI3K inhibitor PIK3IP1 (phosphoinositide-3-kinase interacting protein 1), leading to increased tumor growth and susceptibility to PI3K-AKT inhibition. And we also can forecast a few clones existing in colorectal cancer, containing gene mutation of ptk2, htra1 and PABPC1.

Rectal cancer is demonstrated not simply controlled by a particular gene or signaling pathways, but also by the complex process of network system co-ordinately regulated which consisted of a variety of signalling pathways and multiple genes. In the signalling network, it is likely there is some " key regulatory point". At last, a best understanding of chemotherapy molecular targets allowed the identification of genetic markers that can predict the response and/or the toxicity of anti-cancer drugs used in rectal cancers, which could be helpful in the future to propose for each patient a personalized treatment [29-31]. Mutations that can predict the response of new target therapies such as the inhibitors of the JAK kinase inhibitor AG4 90 in colorectal cancer have also been found and will allow the selection of patients who can have benefit from these new therapeutic drugs. The experiment dig out a variety of signalling pathways and genes which can provide reliable directions for molecular mechanism research of treatment, and it need to be further verified.

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Compliance with Ethical Standards

Conflict of interest：We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled exploring genes of rectal cancer for new treatments based on protein interaction network

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Spreadborough P, Doran C.The current thinking on colorectal cancer.J R Nav Med Serv 2015;101(1):47-54.

2. Stoffel EM, Boland CR.Genetics and Genetic Testing in Hereditary Colorectal Cancer. Gastroenterology 2015;pii:S0016-5085(15)01016-1.

3. Schlemper RJ,Dawsey SM,Itabashi M,et al. Differences in diagnostic criteria for esophageal squamous cell carcinoma between Japanese and Western pathologists.Cancer 2000 ;88(5):996-1006.

4. Schlussel AT1, Gagliano RA Jr1, Seto-Donlon S1.The evolution of colorectal cancer genetics-Part 1: from discov ery to practice. J Gastrointest Oncol. 2014 ;5(5):326-35.

5. Zhou C1, Teng WJ, Yang J,et al.Construction of a protein-protein interaction network for chronic myelocytic leukemia and pathway prediction of molecular complexes.Asian Pac J Cancer Prev 2014;15(13):5325-30.

6. Macaron C1, Heald B, Burke CA,et al.Using Genetics to Identify Hereditary Colorectal Polyposis and Cancer Syndromes in Your Patient.Curr Gastroenterol Rep 2015 ;17(10):463.

7. Hamosh A, Scott AF, Amberger JS, et al. Online Mendelian Inheritance in Man (OMIM), a

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

knowledgebase of human genesand genetic disorders. Nucleic Acids Res 2005;33(Database issue):D514-517.

8. Amberger J, Bocchini CA, Scott AF, et al. McKusick's Online Mendelian Inheritance in Man (OMIM). Nucleic Acids Res 2009;37(Database issue):D793-796.

9. Vailaya A, Bluvas P, Kincaid R, et al. An architecture for biological information extraction and representation.Bioinformatics 2005;21 (4):430-438.

10. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498-2504.

11. Chen L, Wang H, Zhang L, et al . Uncovering packaging features of co-regulated modules based on human protein interaction and transcriptional regulatory networks. BMC Bioinformatics, 2010;11:392.

12. Wu BL,Zhou HY,Lv GQ, et al. Protein-protein Interaction network analyses for elucidating the roles of LOXL2-delta72 in esophageal squamous cell carcinoma. Asian Pac J Cancer Prev 2014;15(5): 2345-51.

13. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery[J]. Genome Biol 2003;4(5):P3.

14. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4(1):44-57.

15. Burkard TR, Rix U, Breitwieser FP, Superti-Furga G, Colinge J. A computational approach to analyze the mechanism of action of the kinase inhibitor bafetinib[J].PLoS Comput Biol 2010;6(11):e1001001.

16. Smolle MA1,2, Pichler M1, Haybaeck J2,et al. Genetic markers of recurrence in colorectal cancer. Pharmacogenomics. 2015; 12:1-14.

17. Bacher JW1, Sievers CK2, Albrecht DM3,et al.Improved Detection of Microsatellite Instability in Early Colorectal Lesions.PLoS One 2015 7;10(8):e0132727.

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18. Lee DW, Han SW, Cha Y, Rhee YY, et al.Different prognostic effect of CpG island methylation according to sex in colorectal cancer patients treated with adjuvant FOLFOX. Clin Epigenetics 2015;7(1):63.

19. Guerreiro CS1, Cravo ML, Brito M, Vidal PM,et al .The D1822V APC polym orphism interacts with fat, calcium,and fiber intakes in modulating the risk of colorectalcancer in Portugues e persons.Am J Clin Nutr 2007;85(6):1592-7.

20. Stern HS1, Viertelhausen S, Hunter AG, et al. APC I1307K increases risk of transition from polyp to colorectal carcinoma in Ashkenazi Jews.Gastroenterology 2001;120(2):392-400.

21. Mao C1, Zhou J, Yang Z, Huang Y,et al.KRAS,BRAF and PIK3CA mutations and the l oss of PTEN expression in Chinese patients with colorectal cancer.PLoS One 2012;7(5):e36653.

22. Cappellani A1, Di Vita M, Zanghi A, et al. Biological and clinical markers in colorectal cancer: state of the art.Front Biosci (Schol Ed) 2010;2:422-31.

23. Delektorskaya VV1, Golovkov DA, Kushlinskii NE.Clinical significance of levels of molecular biological markers in zones of invasive front-line of colorectal cancer.Bull Exp Biol Med 2008;146(5):616-9.

24. Ryu JY1, Kim HU, Lee SY. Human genes with a greater number of transcript variants tend to show biological features of housekeeping and essential genes.Mol Biosyst 2015 Aug 17. [Epub ahead of print]

25. Chen L,Wang H,Zhang L,et,al.Uncovering packaging features of co-regulated modules based on human protein interaction and transcriptional regulatory networks[J].BMC Bioinformatics 2010;7(11):392.

26. Kovács IA, Palotai R, Szalay MS,et,al. Community Landscapes: An Integrative Approach to Determine Overlapping Network Module Hierarchy,Identify Key Nodes and Predict Network Dynamics[J].PLoS One 2010;5(9): e12528.

27. Yu SJ1, Yu JK, Ge WT, et al. SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and M

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

APK are prognosis-related in colorectal cancer.World J Gastroenterol 2011;17(15):2028-36.

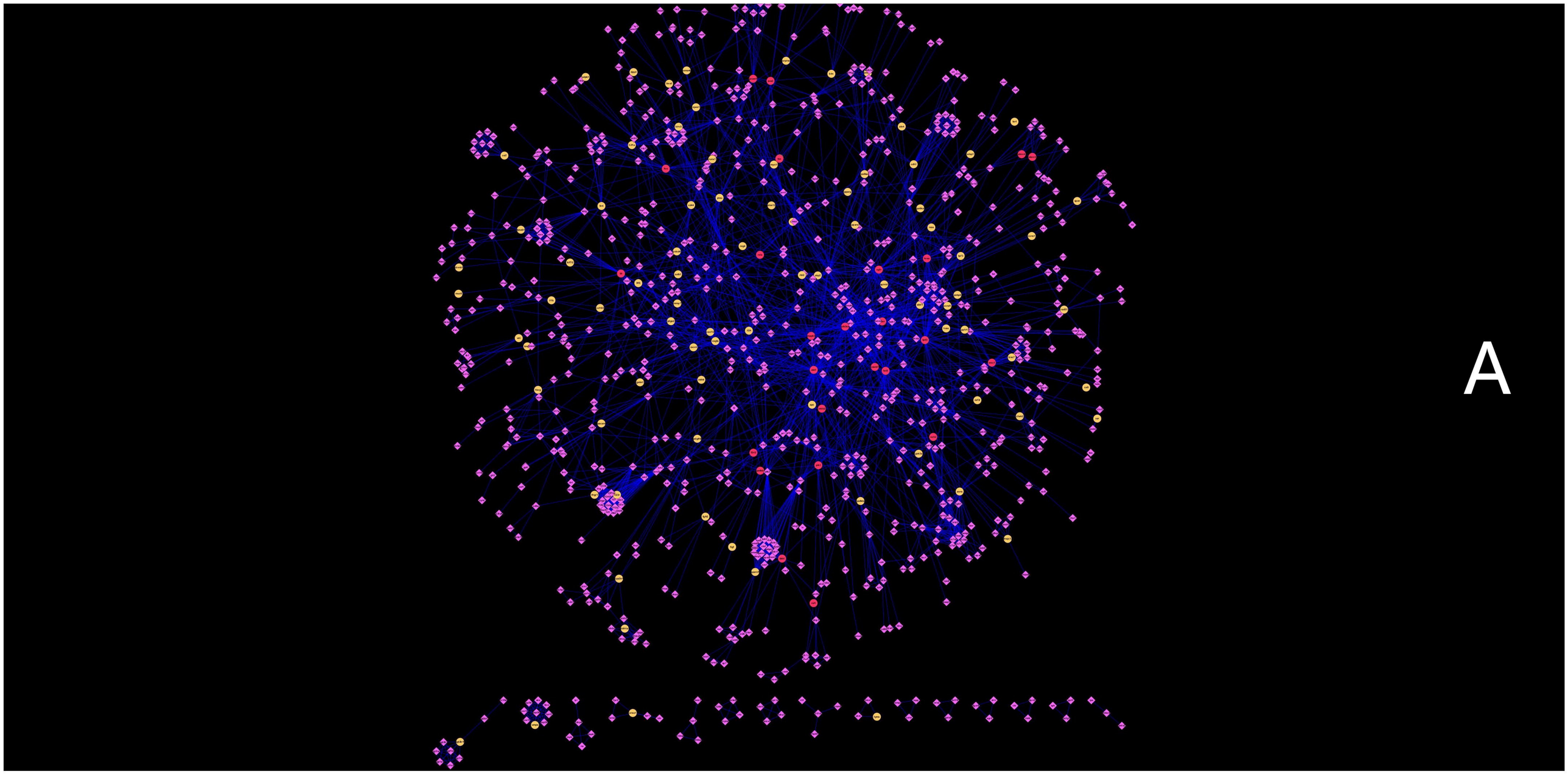
28. Frewer KA1, Sanders AJ, Owen S, et al.A role for WISP2 in colorectal cancer cell invasion and motility.Cancer Genomics Proteomics 2013;10(4):187-96.

29. Gho Kwang-,Cusick M E,Valle D,et al.The human diseases net work. P Natl Acad Sci USA 2007;104(21):8685-90.

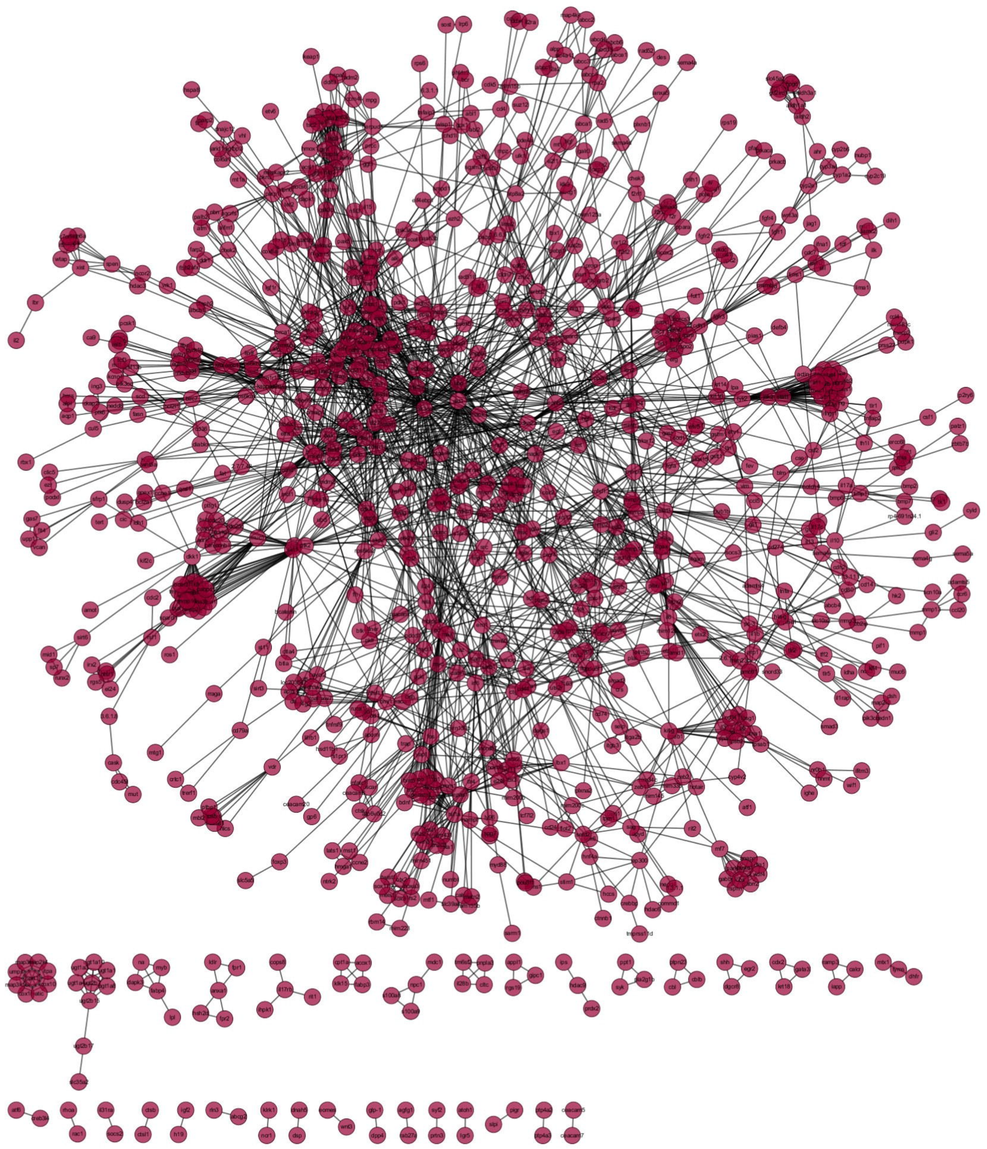
30. Yildirim M A,Goh K I,Cusick M E,et al.Drug-target network. Nat Biotechnol 2007;25(10):1119-26.

31. Hopkins A L.Network pharmacology:the next paradigm in drug discovery. Nat Chen Biol 2008,4(11):682-90.

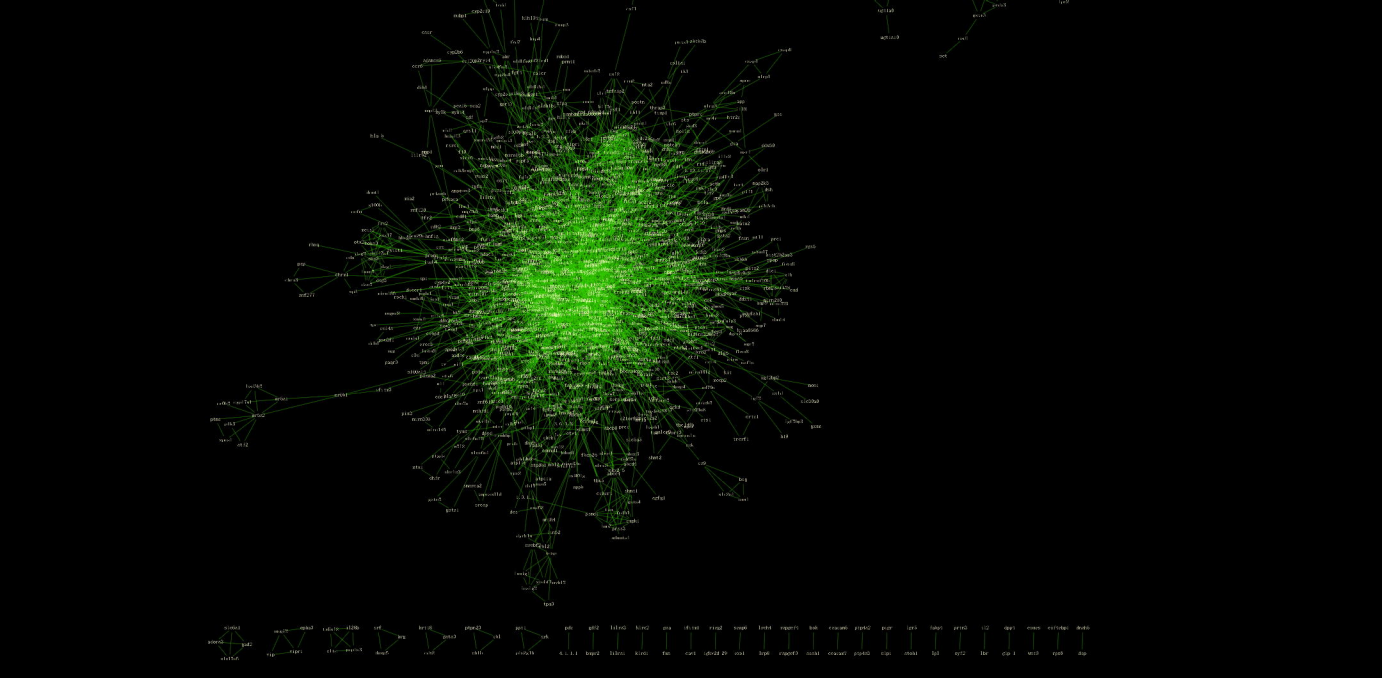
bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

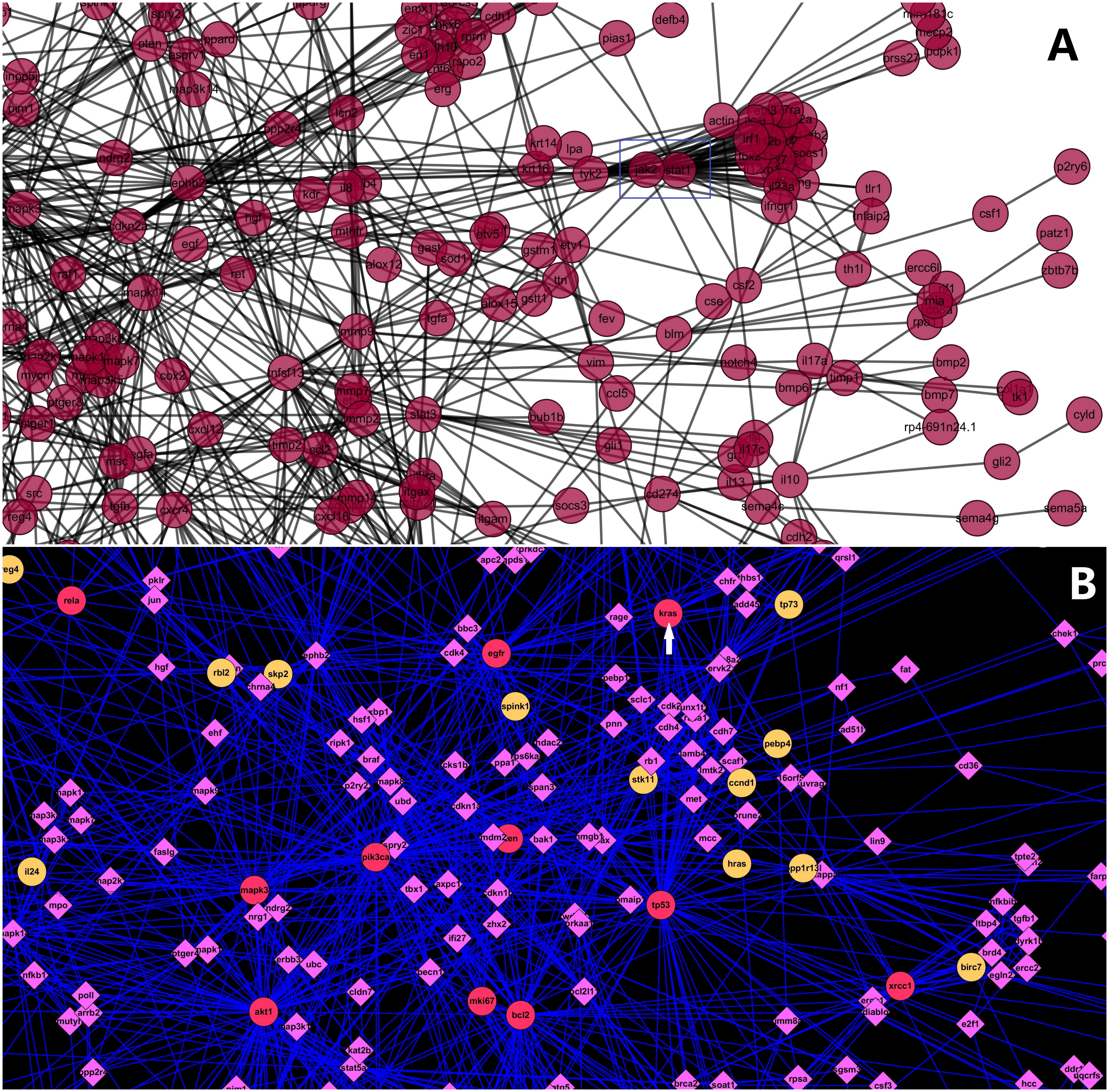


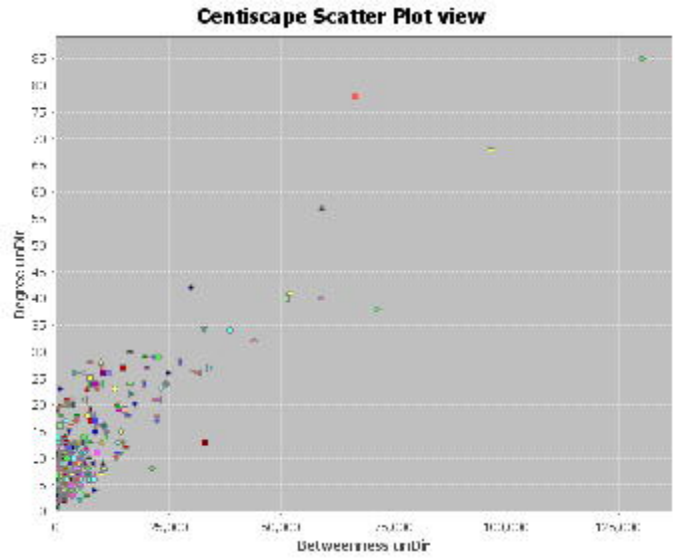
A

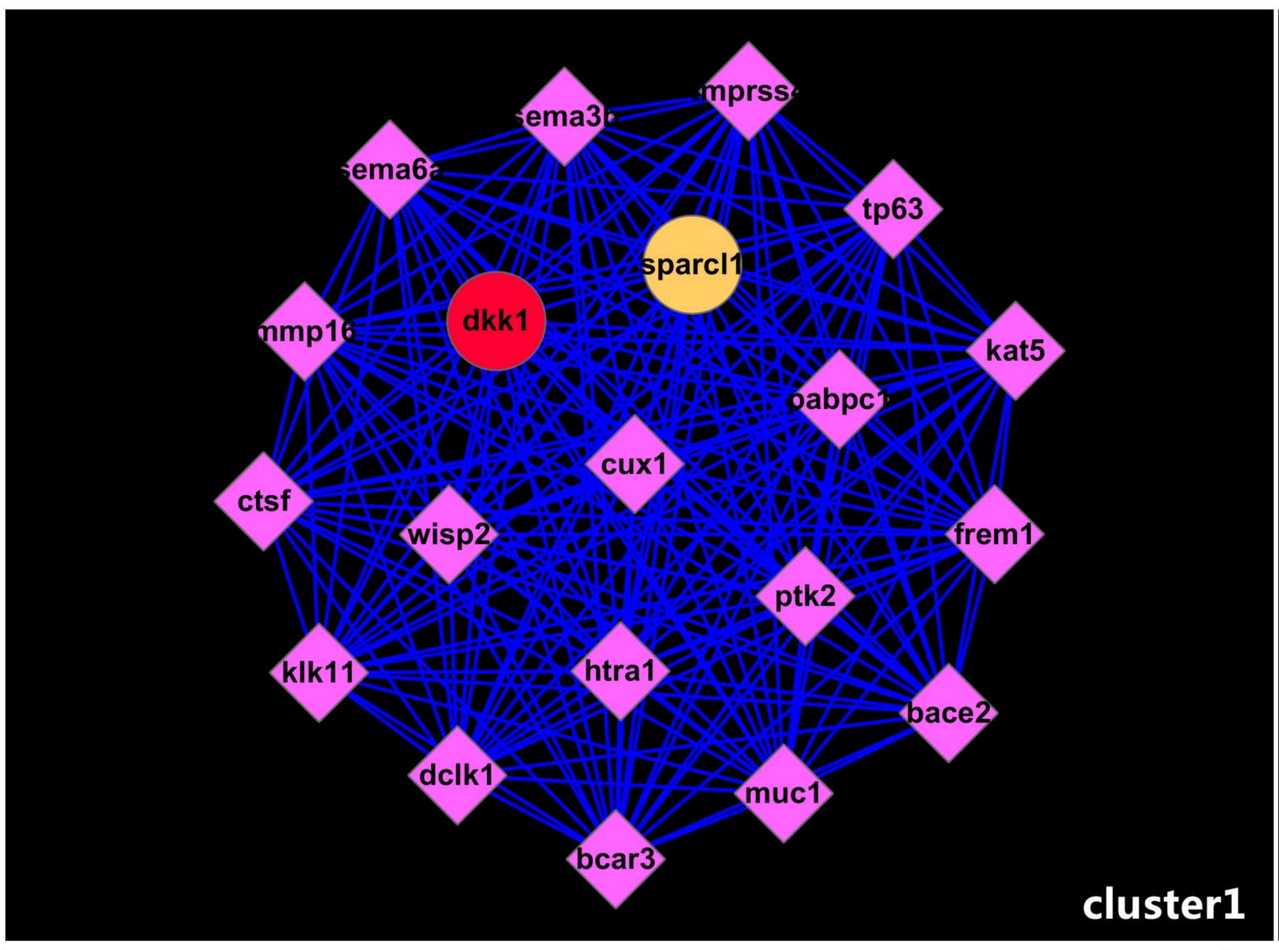


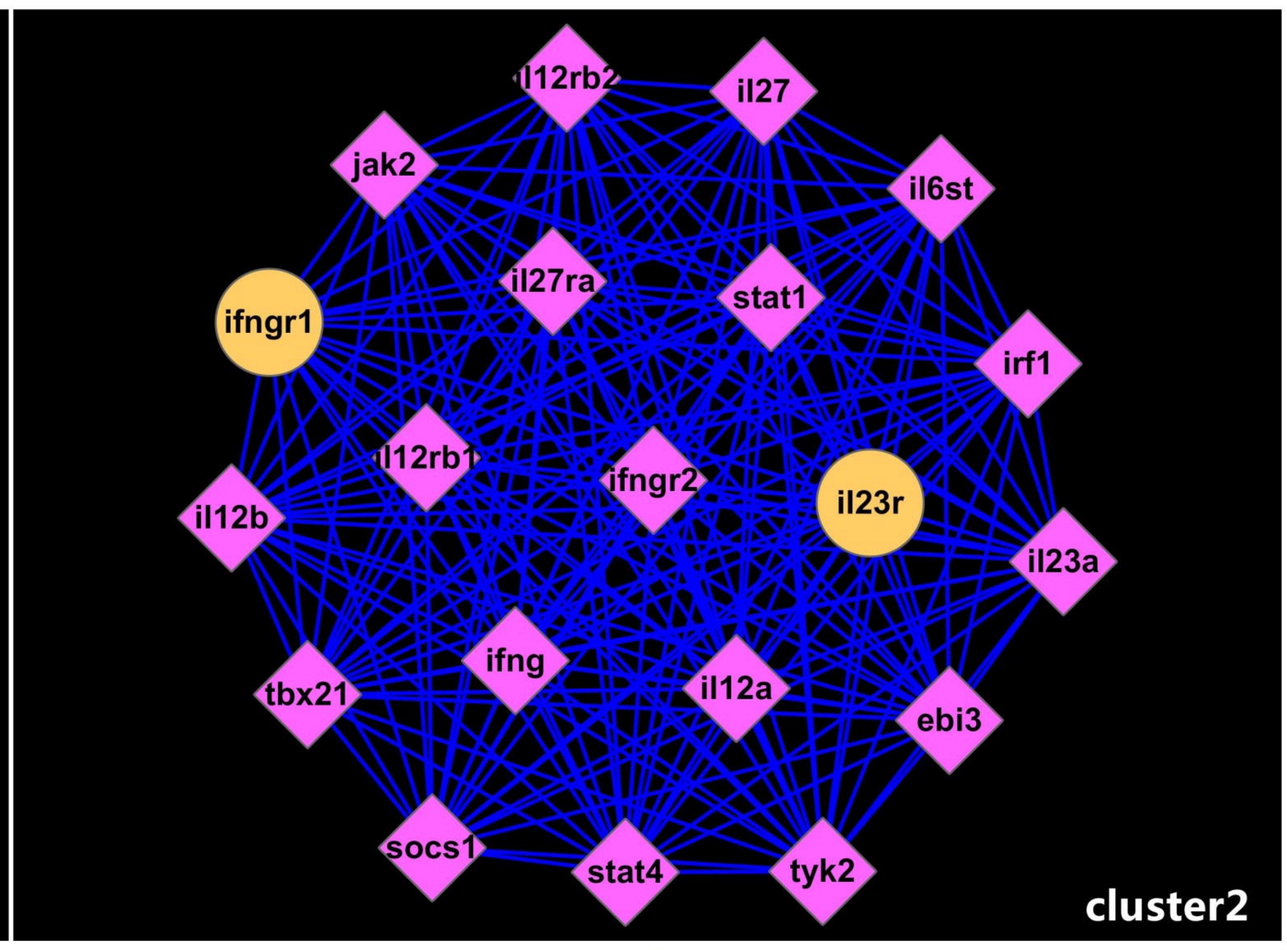
B

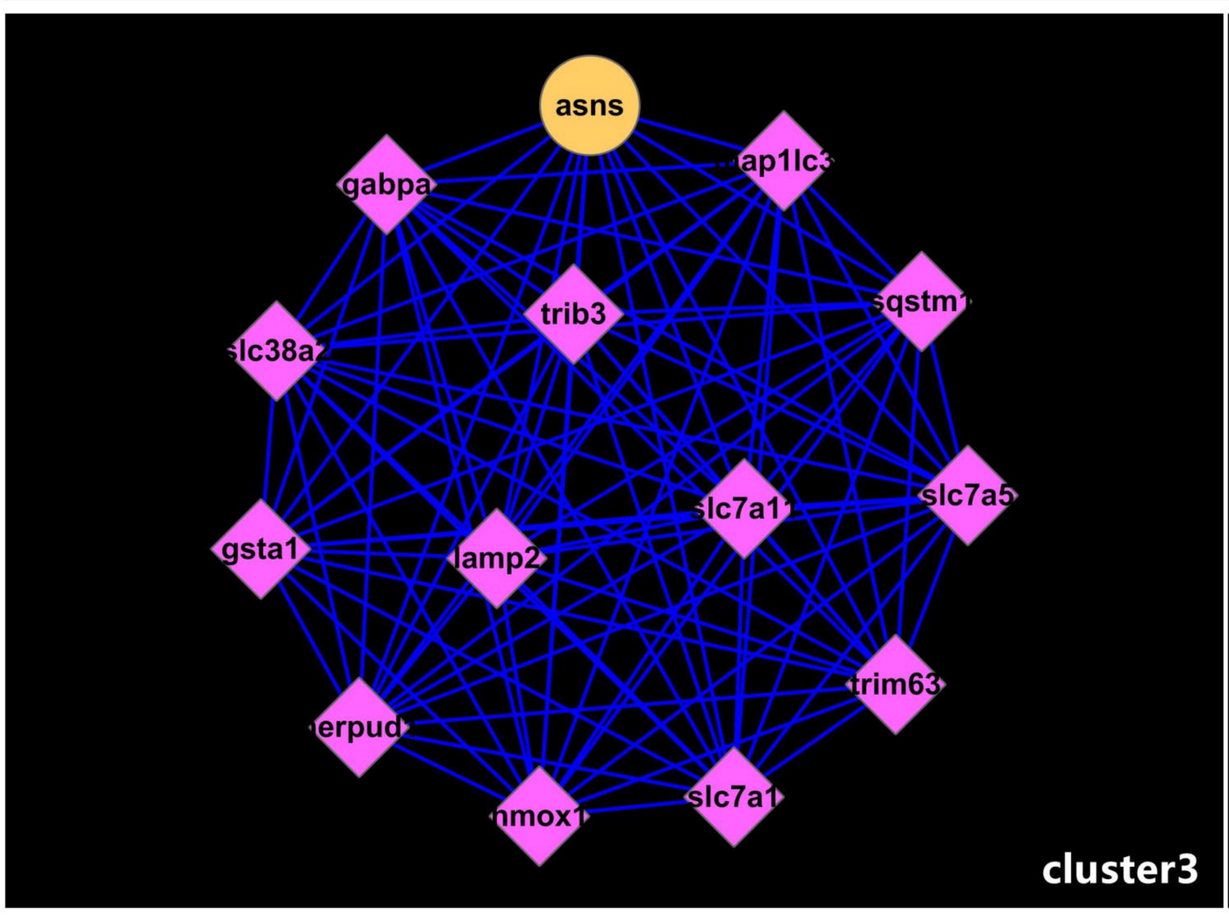


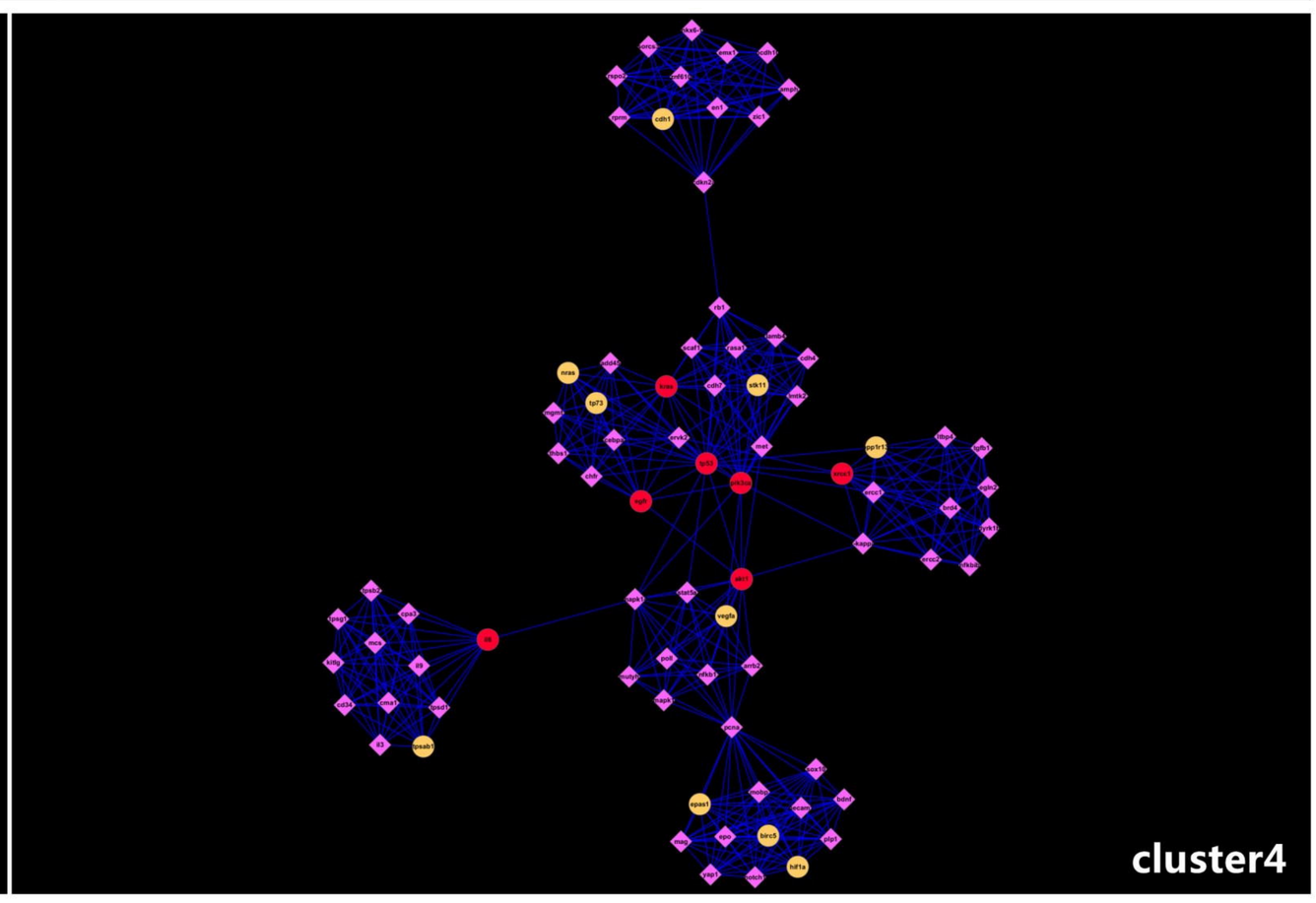

A
spry2
pparg
emx1
zic1
en1
rprm
rspo2
cdh1
erg
defb4
pias1
pten
asprv1
ppard
inpp5j
map3k14
pim1
lcn2
ppp2r4
ndrg2
ephb2
kdr
il8
mapk3
cdkn2a
hgf
mthfr
egf
alox12
ret
raf1
mapk14
mmp9
tgfa
map3k8
mapk1
mapk7
mycn
map3k5
cox2
tnfsf13
mmp7
mmp2
ptger3
ptger1
cxcl12
timp2
ccl2
stat3
msc
vegfa
src
reg4
tgfb
cxcr4
cxcl16
mmp14
itgax
itgam
krt14
lpa
krt16
tyk2
jak2
stat1
actin
irf1
socs1
il23a
ifngr1
tlr1
tnfaip2
gast
sod1
etv5
gstm1
etv1
ttn
alox15
gstt1
fev
cse
csf2
blm
vim
ccl5
notch4
bub1b
gli1
socs3
cd274
il13
il17c
il10
sema4c
cdh2
mirn181c
mecp2
pdpk1
prss27
p2ry6
csf1
patz1
zbtb7b
th1l
ercc6l
mia
rpa1
il17a
timp1
bmp2
bmp6
bmp7
col1a1
tk1
rp4-691n24.1
cyld
gli2
sema4g
sema5a
B
reg4
rela
pklr
jun
apc2
prkdc
chfr
thbs1
qrsl1
tp73
kras
rage
bbc3
cdk4
egfr
hgf
rbl2
skp2
chrna4
ephb2
ervk2
pebp1
fat
nf1
chek1
spink1
sclc1
cdh4
cdh7
hsf1
xbp1
ehf
ripk1
braf
pnn
cks1b
ppa1
hdac2
rb1
lamb4
scaf1
pebp4
cd36
stk11
ccnd1
lmtk2
met
uvrag
mapk9
mapk8
ubd
span3
cdkn1a
prune2
mapk7
faslg
mdm2
bak1
mcc
lin9
hras
il24
pik3ca
spry2
mapk3
tbx1
axpc1
cdkn1b
pmaip1
tp53
tpte2
ndrg2
nrg1
mpo
zhx2
ifi27
ltbp4
tgfb1
nfkbib
dyrk1b
brd4
birc7
ercc2
egln2
erbb3
ubc
ptger4
nfkb1
poll
arrb2
becn1
cldn7
mki67
bcl2
bcl2l11
xrcc1
diablo
akt1
e2f1
kat2b
stat5a
rpsa
sgsm3
brca2
soat1
csf3
hcc


Centiscape Scatter Plot view
Betweenness unDir


sema3
mprss
sema6
tp63
sparcl1
dkk1
mmp16
kat5
cux1
ctsf
wisp2
frem1
ptk2
klk11
htra1
bace2
dclk1
muc1
bcar3
cluster1

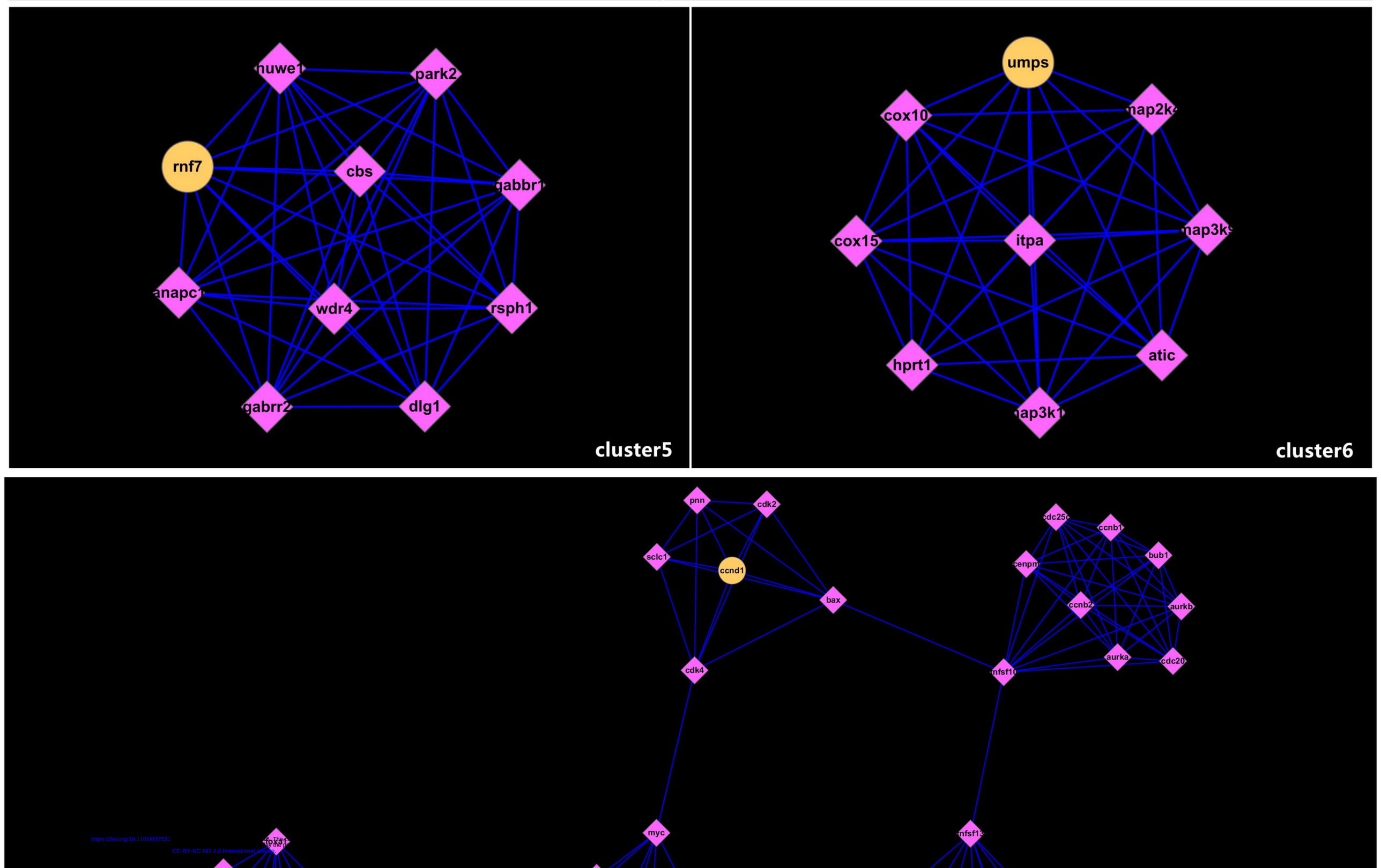

il12rb2
il27
jak2
il6st
il27ra
stat1
ifngr1
irf1
il12rb1
ifngr2
il23r
il12b
il23a
ifng
tbx21
il12a
ebi3
socs1
stat4
tyk2
cluster2

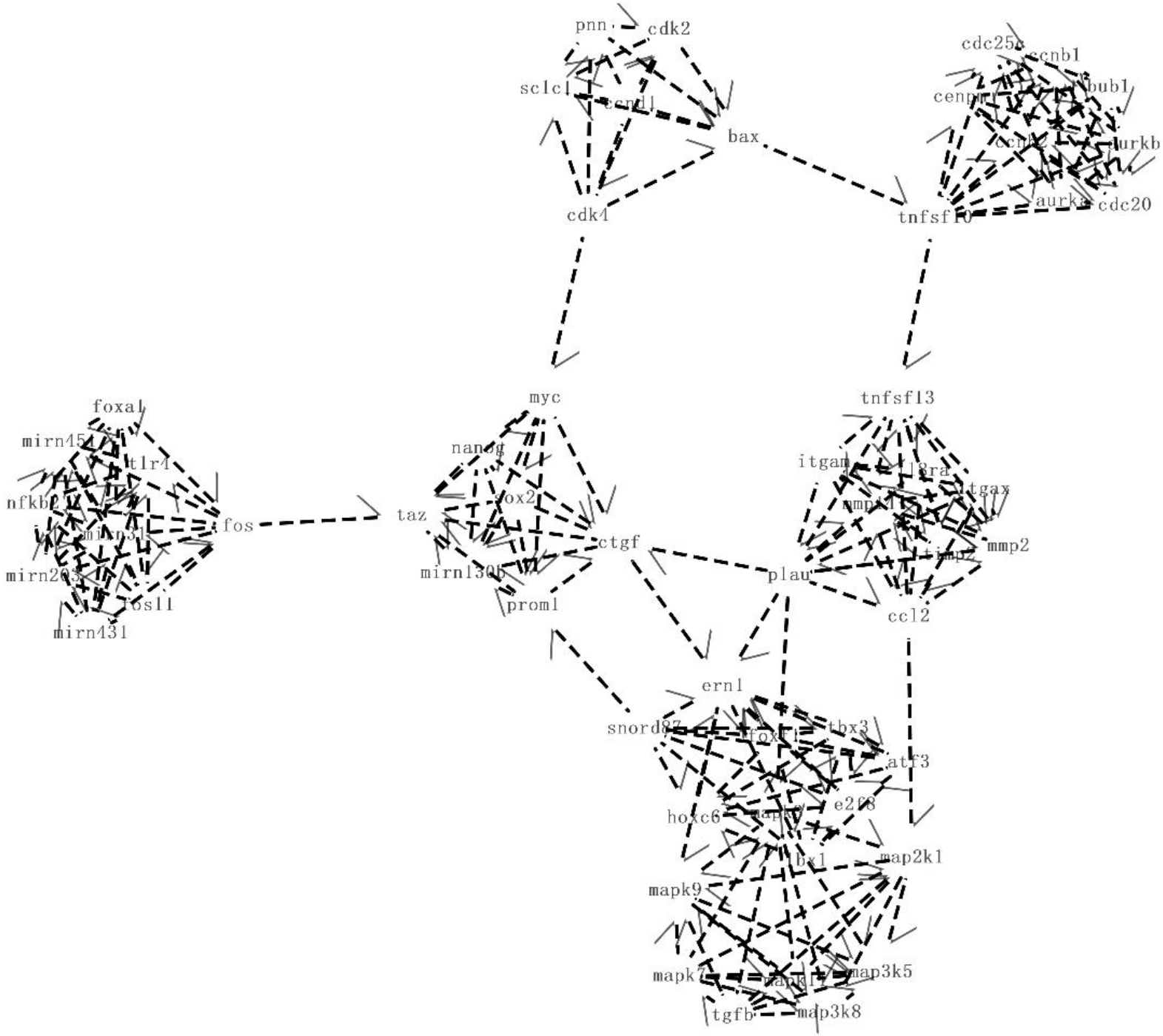

asns
gabpa
trib3
slc38a2
slc7a11
slc7a5
gsta1
lamp2
trim63
slc7a11
nmox1
cluster3

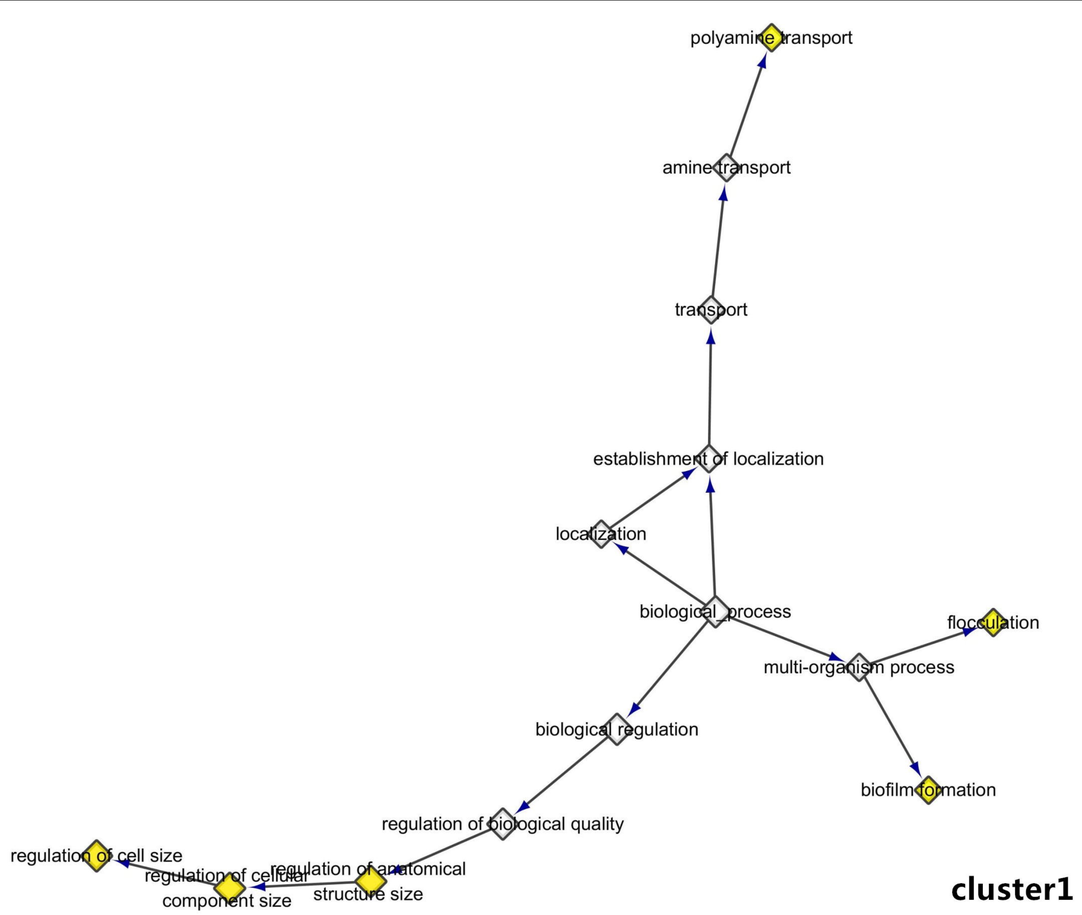

cluster4

huwe1
park2
rnf7
cbs
wdr4
rsph1
dlg1
cluster5

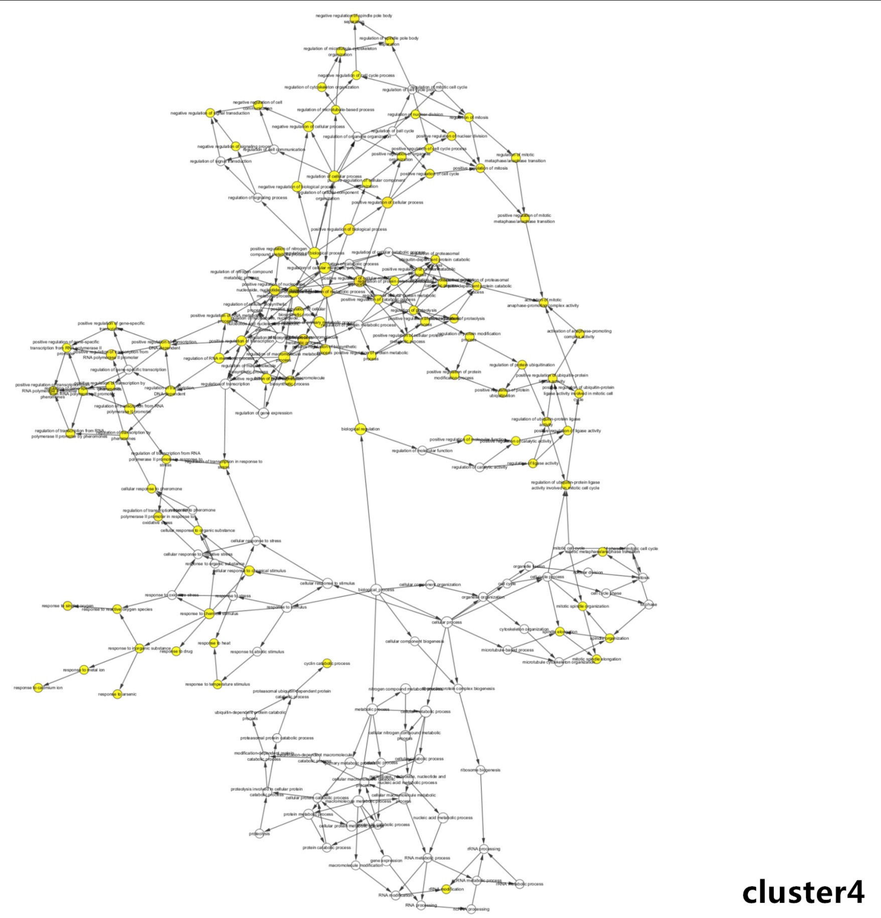
umps
cox10
cox15
itpa
hprt1
atic
cluster6


pnn
cdk2
sclc1
ccnd1
bax
cdk4
cdc25
ccnb1
cenpn
bub1
ccnb2
aurkb
aurka
cdc20
tnfsf10
myc
tnfsf13
nanog
sox2
taz
ctgf
plau
prom1
itgam
il8ra
itgax
mmp14
timp2
mmp2
ccl2
tlr4
nfkb2
fos
mirn31
mirn20
fosl1
mirn43
ern1
nord8
foxf1
tbx3
atf3
e2f8
noxc6
mapk3
lbx1
map2k
mapk9
mapk7
mapk1
map3k
tgfb
map3k
cluster7

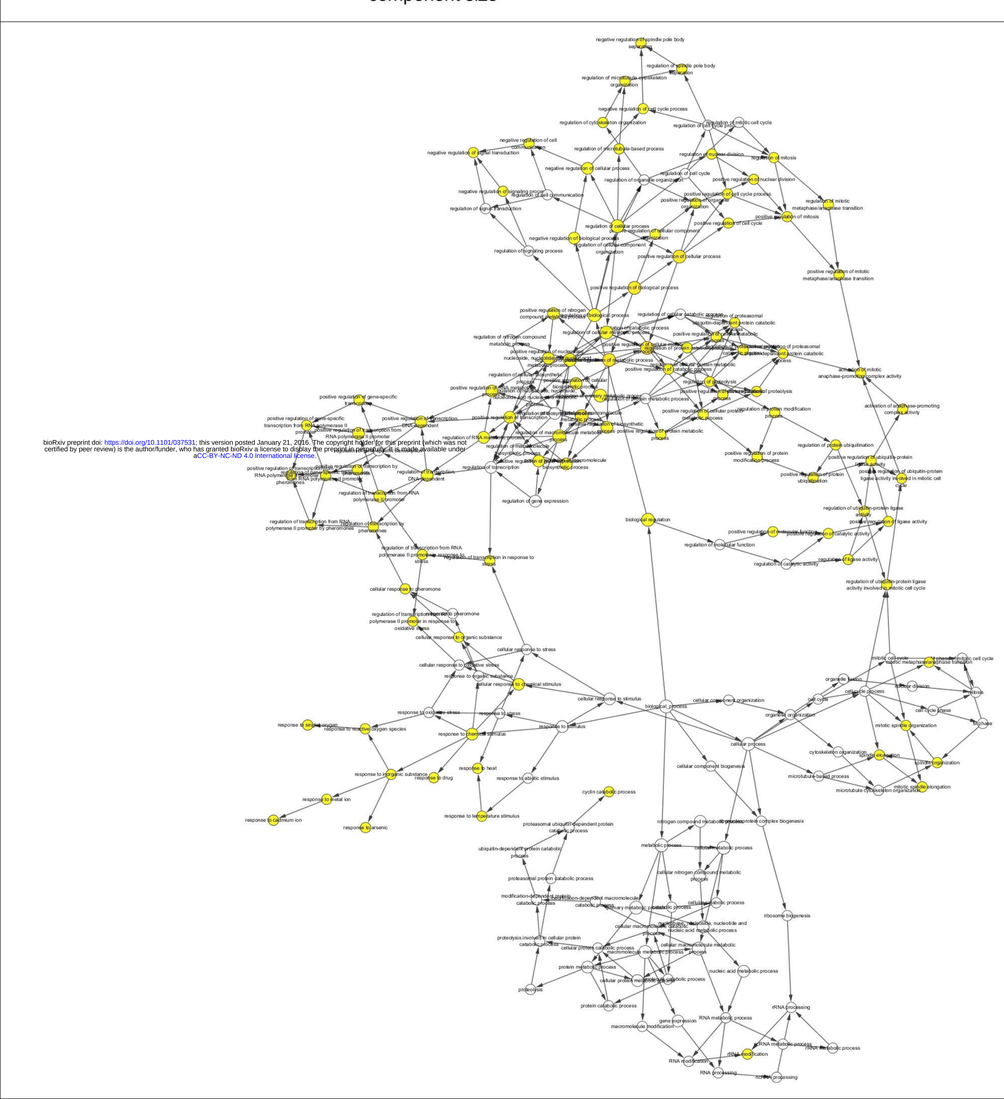

pnn
cdk2
sclc1
ccnd1
bax
cdk4
cdc25c
ccnb1
cenpm
bub1
ccnb2
aurkb
aurka
cdc20
tnfsf10
foxa1
mirn451
tlr4
nfkb2
mirn31
fos
mirn203
fos11
mirn431
myc
nanog
sox2
taz
ctgf
mirn130b
prom1
tnfsf13
itgam
l8ra
itgax
mmp11
mmp2
timp2
plau
ccl2
ern1
snord87
foxf1
tbx3
atf3
e2f8
hoxc6
mapk8
bx1
map2k1
mapk9
mapk7
mapk11
map3k5
tgfb
map3k8



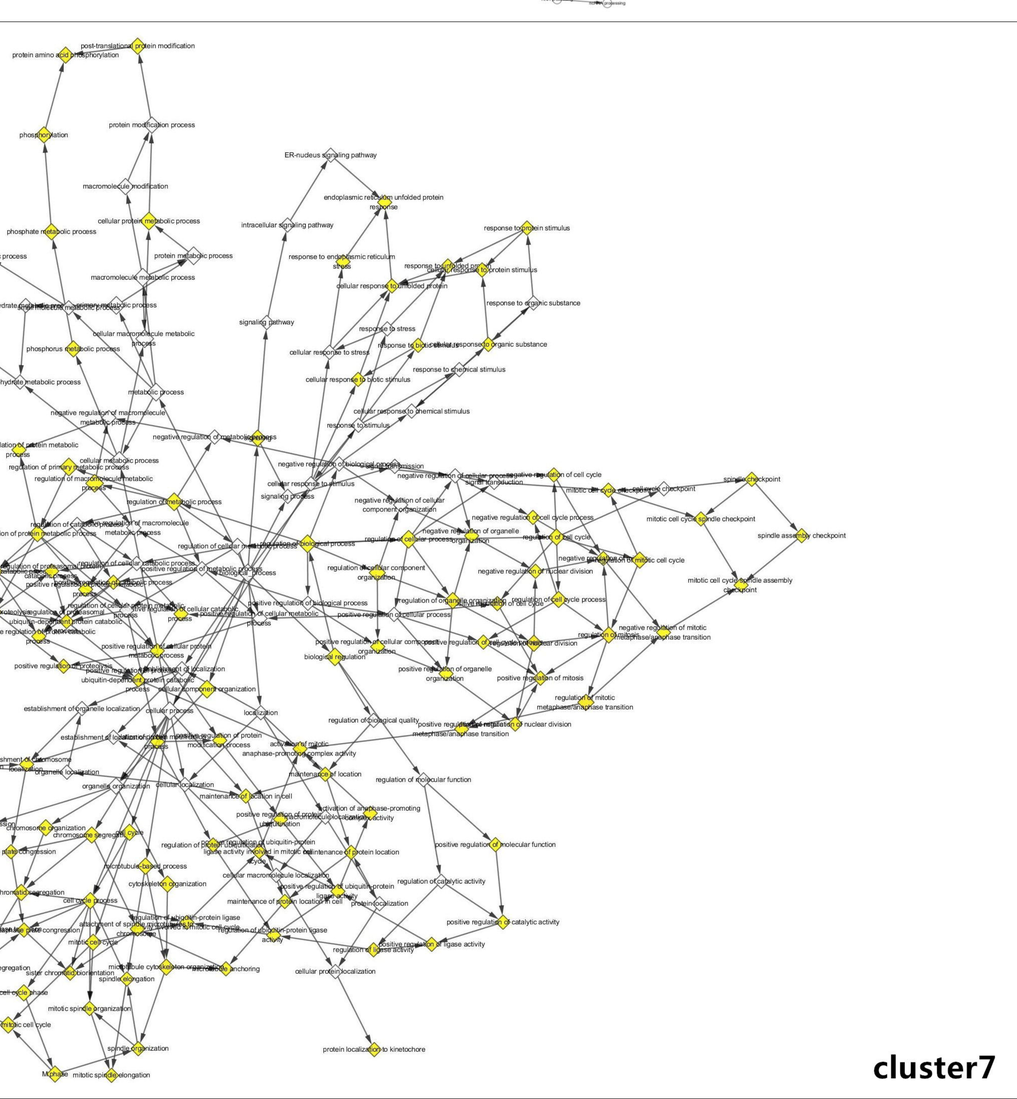
cluster1



cluster4



cluster6



cluster7

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Symbol	Aliases
KRAS	C-K-RAS, CFC2, K-RAS2A, K-RAS2B, K-RAS4A, K-RAS4B, KI-RAS1, KRAS2, NS, NS3,
HIF1A	HIF-1A, HIF-1alpha, HIF1, HIF1-ALPHA, MOP1, PASD8, bHLHe78
TP53	BCC7, LFS1, P53, TRP53
TYMS	HST422, TMS, TS
EGFR	ERBB, ERBB1, HER1, NISBD2, PIG61, mENA
VEGFA	MVCD1, VEGF, VPF
CXCL12	IRH, PBSF, SCYB12, SDF1, TLSF, TPAR1
SLC2A1	CSE, DYT17, DYT18, DYT9, EIG12, GLUT, GLUT-1, GLUT1, GLUT1DS, HTLVR
PROM1	AC133, CD133, CORD12, MCDR2, MSTP061, PROML1, RP41, STGD4
ALDH1A1	ALDC, ALDH-E1, ALDH1, ALDH11, HEL-9, HEL-S-53e, HEL12, PUMB1, RALI
MTHFR	
ERBB2	CD340, HER-2, HER-2/neu, HER2, MLN 19, NEU, NGL, TKR1
CD44	CDW44, CSPG8, ECMR-III, HCELL, HUTCH-I, IN, LHR, MC56, MDU2, MDU3, MIC
EZR	CVIL, CVL, HEL-S-105, VIL2
MMP7	MMP-7, MPSL1, PUMP-1
CEACAM7	CGM2
PEBP4	CORK-1, CORK1, GWTM1933, HEL-S-300, PEBP-4, PRO4408, hPEBP4
PPARG	CIMT1, GLM1, NR1C31, PPARG2, PPARgamma, PPARG
CTNNB1	CTNNB, MRD19, armadillo
SERPINE1	PAI, PAI-1, PAI1, PLANH1

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

BIRC5	API4, EPR-1
XRCC1	RCC
PIK3CA	CLOVE, CWS5, MCAP, MCM, MCMTC, PI3K, p110-alpha
HRAS	C-BAS/HAS, C-H-RAS, C-HA-RAS1, CTLO, H-RASIDX, HAMSV1, RASH1, p21ras
CYP2E1	CPE1, CYP2E, P450-J, P450C2E
EZH2	ENX-1, ENX1, EZH1b, KMT6, KMT6A, WVS, WVS2, EZH2
MKI67	KIA, MIB-, MIB-1, PPP1R105
CA9	CAIX, MN
CEACAM1	BGP, BGP1, BGPI
DPYD	DHP, DHPDHASE, DPD
TGFA	TFGA
CEACAM5	CD66e, CEA
PDPK1	PDK1, PDPK2, PDPK2P, PRO0461
LGR5	FEX, GPR49, GPR67, GRP49, HG38
ETV4	E1A-F, E1AF, PEA3, PEAS3
APOE	AD2, APO-E, LDLCQ5, LPG
IL6	BSF2, HGF, HSF, IFNB2, IL-6
TGFB1	CED, DPD1, LAP, TGFB, TGFbeta
PTGS2	COX-2, COX2, GRIPGHS, PGG/HS, PGHS-2, PHS-2, hCox-2
CRP	PTX1
IL1B	IL-1, IL1-BETA, IL1F2

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AKT1	AKT, CWS6, PKB, PKB-ALPHA, PRKBA, RAC, RAC-ALPHA
VDR	NR1I1, PPP1R163
TLR4	ARMD10, CD284, TLR-4, TOLL
GSTM1	GST1-1, GSTM1a-1a, GSTM1b-1b, GTH4, GTM1, H-B, MU, MU-1, GSTM1
PTEN	10q23del, BZS, CWS1, DEC, GLM2, MHAM, MMAC11, TEP1, PTEN
CDH1	Arc-1, CD324, CDHE, ECAD, LCAM, UVO
TERT	CMM9, DKCA2, DKCB4, EST2, PFBMFT1, TCS1, TP2, TRT, hEST2, hTRT
BCL2	Bcl-2, PPP1R50
CXCR4	CD184, D2S201E, FB22, HM89, HSY3RR, LAP-3, LAP3, LCR1, LESTR, NPY3R, NPYR, NPYF WHIMS
CCND1	BCL1, D11S287E, PRAD1, U21B31
HLA-DQB1	CELIAC1, HLA-DQB, IDDM1
RELA	NFKB3, p65
CASP3	CPP32, CPP32B, SCA-1
HLA-G	MHC-G
MAPK3	ERK-1, ERK1, ERT2, HS44KDAP, HUMKER1A, P44ERK1, P44MAPK, PRKM3, p44-ERK
MMP1	CLG, CLGN
APC	BTPS2, DP2, DP2.5, DP3, GS, PPP1R46
CYP19A1	ARO, ARO1, CPV1, CYAR, CYP19, CYPXIX, P-450AROM
ERCC2	COFS2, EM9, TFIIH, TTD, TTD1, XPD
TCF7L2	TCF-4, TCF4

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

F3	CD142, TF, TFA
TP73	P73
LCN2	24p3, MSFI, NGAL
RUNX1	AML1, AML1-EVI-1, AMLCR1, CBF2alpha, CBFA2, EVI-1, PEBP2aB, PEBP2a
FGFR2	BBDS, BEK, BFR-1, CD332, CEK3, CFD1, ECT1, JWS, K-SAM, KGFR, TK14, T
CHEK2	CDS1, CHK2, HuCds1, LFS2, PP1425, RAD53, hCds1
PLAU	ATF, BDPLT5, QPD, UPA, URK, u-PA
TIMP1	CLGI, EPA, EPO, HCI, TIMP
RAD51	BRCC5, FANCR, HRAD51, HsRad51, HsT16930, MRMV2A, RECA, RAD51
PLK1	PLK, STPK13
POU5F1	OCT3, OCT4, OTF-3, OTF3, OTF4, Oct-3, Oct-4
XRCC3	CMM6
SOX2	ANOP3, MCOPS3
PDGFRB	CD140B, IBGC4, IMF1, JTK12, PDGFR, PDGFR-1, PDGFR1
STK11	LKB1, PJS, hLKB1
DKK1	DKK-1, SK
EPAS1	ECYT4, HIF2A, HLF, MOP2, PASD2, bHLHe73
CFLAR	CASH, CASP8AP1, CLARP, Casper, FLAME, FLAME-1, FLAME1, FLIP, I-FLICE, MRIT, c-FLIP c-FLIPS
FOS	AP-1, C-FOS, p55
SKP2	FBL1, FBXL1, FLB1, p45

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

NRAS	ALPS4, CMNS, N-ras, NCMS1, NS6, NRAS
CCNE1	CCNE, pCCNE1
IL23R	
PPARD	FAAR, NR1C2, NUC1, NUCI, NUCII, PPARB
RPS6KB1	PS6K, S6K, S6K-beta-1, S6K1, STK14A, p70 S6KA, p70(S6K)-alpha, p70-S6K, p7
ANXA1	ANX1, LPC1
PTGS1	COX1, COX3, PCOX1, PES-1, PGG/HS, PGHS-1, PGHS1, PHS1, PTGHS
BLM	BS, RECQ2, RECQL2, RECQL3
CD163	M130, MM130
SPINK1	PCTT, PSTI, Spink3, TATI, TCP
CLDN1	CLD1, ILVASC, SEMP1
PDCD4	H731
IFNGR1	CD119, IFNGR, IMD27A, IMD27B
IL24	C49A, FISP, IL10B, MDA7, MOB5, ST16
ABCC4	MOAT-B, MOATB, MRP4
ALOX12	12-LOX, 12S-LOX, LOG12
BIRC7	KIAP, LIVIN, ML-IAP, MLIAP, RNF50
SATB1	
TFF3	ITF, P1B, TFI
CSF3R	CD114, GCSFR
RBL2	P130, Rb2

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RPS6KA1	HU-1, MAPKAPK1A, RSK, RSK1, p90Rsk
FIGF	VEGF-D, VEGFD
TPSAB1	TPS1, TPS2, TPSB1
GSK3A	
ABCC3	ABC31, EST90757, MLP2, MOAT-D, MRP3, cMOAT2
MIR137	MIRN137, miR-137
TAZ	BTHS, CMD3A, EFE, EFE2, G4.5, LVNCX, Taz1
UGT2B15	HLUG4, UDPGT 2B8, UDPGT2B15, UDPGTH3, UGT2B8
PPP1R13L	IASPP, NKIP1, RAI, RAI4
REG4	GISP, REG-IV, RELP
WISP1	CCN4c, WISP1i, WISP1tc, WISP1
ASNS	ASNSD, TS11
UMPS	OPRT
F2RL2	PAR-3, PAR3
LIMS1	PINCH, PINCH-1, PINCH1
MUC6	MUC-6
RSF1	HBXAP, RSF-1, XAP8, p325
MYCL	L-Myc, LMYC1, bHLHe38, MYCL
RPS6KB2	KLS, P70-beta, P70-beta-1, P70-beta-2, S6K-beta2, S6K2, SRK, STK14B, p70(S6K)-be
RNF7	CKBBP1, ROC2, SAG
RPS6KA2	HU-2, MAPKAPK1C, RSK, RSK3, S6K-alpha, S6K-alpha2, p90-RSK3, pp90RS

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SPARCL1	MAST 9, MAST9, PIG33, SC1
SPEN	HIAA0929, MINT, RBM15C, SHARP
SEMA4C	M-SEMA-F, SEMACL1, SEMAF, SEMAI
DNAJC12	JDP1

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Category	Term	PValue	Genes	FDR
cluster1				
KEGG_PATHWAY	hsa04360:Axon guidance	0.006073822	SEMA6A, PTK2, SEMA3B	3.843791544
EC_NUMBER	3.4.21.-	0.009036531	HTRA1, KLK11, TMPRSS4	4.848129606
PANTHER_PATHWAY	P00004:Alzheimer disease-presenilin pathway	0.019326937	BACE2, MMP16, KAT5	11.44756514
cluster2				
KEGG_PATHWAY	hsa04630:Jak-STAT signaling pathway	3.32E-22	IL23R, IL6ST, SOCS1, STAT1, TYK2, IL12RB2, STAT4, IL23A, IL12RB1, IFNG, IL12A, JAK2, IL12B, IFNGR2, IFNGR1	2.53E-19
BIOCARTA	h_no2il12Pathway:NO2-dependent IL 12 Pathway in NK cells	1.28E-10	IL12RB2, TYK2, STAT4, IL12RB1, IFNG, IL12A, JAK2	1.10E-07
BIOCARTA	h_th1th2Pathway:Th1/Th2 Differentiation	2.81E-09	IL12RB2, IL12RB1, IFNG, IL12A, IL12B, IFNGR2, IFNGR1	2.42E-06
BIOCARTA	h_IL12Pathway:IL12 and Stat4 Dependent Signaling Pathway in Th1 Development	2.81E-09	IL12RB2, TYK2, STAT4, IL12RB1, IFNG, IL12A, JAK2	2.42E-06
KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	3.57E-09	IL12RB2, IL23A, IL12RB1, IL23R, IL6ST, IFNG, IL12A, IL12B, IFNGR2, IFNGR1	2.72E-06
BIOCARTA	h_ifngPathway:IFN gamma signaling pathway	4.16E-08	IFNG, JAK2, STAT1, IFNGR2, IFNGR1	3.57E-05
BIOCARTA	h_nktPathway:Selective expression of chemokine receptors during T-cell polarization	7.64E-07	IL12RB2, IL12RB1, IFNG, IL12A, IFNGR2, IFNGR1	6.57E-04
PANTHER_PATHWAY	P00035:Interferon-gamma signaling pathway	6.85E-06	IFNG, JAK2, STAT1, IFNGR2, IFNGR1	0.004406679
PANTHER_PATHWAY	P00038:JAK/STAT signaling pathway	8.34E-05	STAT4, SOCS1, JAK2, STAT1	0.053656257
BBID	12.IL-6_type_cytok-signal-transduct	8.71E-05	TYK2, IL6ST, SOCS1, JAK2, STAT1	0.07288153
PANTHER_PATHWAY	P00036:Interleukin signaling pathway	9.85E-05	IL12RB2, STAT4, IL23A, IL12RB1, IL6ST, IL12A, STAT1	0.063347144
BIOCARTA	h_tidPathway:Chaper	2.84E-04	IFNG, JAK2, IFNGR2, IFNGR1	0.243475935

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	ones modulate interferon Signaling Pathway			
BBID	48.mice_minus_JAKs_and_STATs	7.11E-04	TYK2, STAT4, JAK2, STAT1	0.593475439
PANTHER_PATHWAY	P00031:Inflammation mediated by chemokine and cytokine signaling pathway	7.32E-04	TYK2, STAT4, IFNG, JAK2, STAT1, IFNGR2, IFNGR1	0.470260756
BBID	75.Stats_activators_of_Apoptosis	0.00193497	IFNG, JAK2, STAT1	1.607237902
KEGG_PATHWAY	hsa05330:Allograft rejection	0.004204138	IFNG, IL12A, IL12B	3.155617774
KEGG_PATHWAY	hsa04940:Type I diabetes mellitus	0.005691714	IFNG, IL12A, IL12B	4.251279397
EC_NUMBER	2.7.10.2	0.009766127	TYK2, JAK2	0.976612696
cluster3				
REACTOME_PATHWAY	REACT_15518:Transmembrane transport of small molecules	1.40E-04	SLC38A2, SLC7A1, SLC7A5, SLC7A11	0.07222524
REACTOME_PATHWAY	REACT_13:Metabolism of amino acids	5.32E-04	SLC38A2, SLC7A1, ASNS, SLC7A5, SLC7A11	0.274131661
cluster4				
KEGG_PATHWAY	hsa05200:Pathways in cancer	8.49E-14	EGFR, CEBPA, IL6, EPAS1, STAT5A, MET, TP53, KITLG, CDH1, BIRC5, NFKB1, RB1, TGFB1, AKT1, MAPK1, NRAS, LAMB4, CDKN2A, HIF1A, KRAS, VEGFA, PIK3CA	9.07E-11
KEGG_PATHWAY	hsa05219:Bladder cancer	6.35E-11	EGFR, NRAS, MAPK1, CDKN2A, KRAS, VEGFA, TP53, CDH1, RB1, THBS1	6.79E-08
KEGG_PATHWAY	hsa05218:Melanoma	4.01E-10	EGFR, AKT1, NRAS, MAPK1, CDKN2A, KRAS, MET, TP53, PIK3CA, CDH1, RB1	4.29E-07
KEGG_PATHWAY	hsa05212:Pancreatic cancer	4.63E-10	EGFR, AKT1, MAPK1, CDKN2A, KRAS, VEGFA, TP53, PIK3CA, NFKB1, RB1, TGFB1	4.94E-07
KEGG_PATHWAY	hsa05220:Chronic myeloid leukemia	7.02E-10	AKT1, NRAS, MAPK1, CDKN2A, KRAS, STAT5A, TP53, PIK3CA, NFKB1, RB1, TGFB1	7.50E-07
KEGG_PATHWAY	hsa05211:Renal cell carcinoma	7.70E-09	AKT1, NRAS, MAPK1, HIF1A, KRAS, EPAS1, MET, VEGFA, PIK3CA, TGFB1	8.23E-06
KEGG_PATHWAY	hsa05223:Non-small cell lung cancer	1.81E-08	EGFR, AKT1, NRAS, MAPK1, CDKN2A, KRAS, TP53, PIK3CA, RB1	1.93E-05

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

KEGG_PATHWAY	hsa05214:Glioma	6.34E-08	EGFR, AKT1, NRAS, MAPK1, CDKN2A, KRAS, TP53, PIK3CA, RB1	6.78E-05
PANTHER_PATHWAY	P04398:p53 pathway feedback loops 2	2.17E-07	AKT1, NRAS, CDKN2A, KRAS, MAPK14, TP53, PIK3CA, RB1, TP73	1.98E-04
KEGG_PATHWAY	hsa05213:Endometrial cancer	3.04E-07	EGFR, AKT1, NRAS, MAPK1, KRAS, TP53, PIK3CA, CDH1	3.25E-04
KEGG_PATHWAY	hsa05210:Colorectal cancer	6.20E-07	EGFR, AKT1, MAPK1, KRAS, MET, TP53, PIK3CA, BIRC5, TGFB1	6.62E-04
KEGG_PATHWAY	hsa05221:Acute myeloid leukemia	6.57E-07	CEBPA, AKT1, NRAS, MAPK1, KRAS, STAT5A, PIK3CA, NFKB1	7.01E-04
KEGG_PATHWAY	hsa05215:Prostate cancer	9.71E-07	EGFR, AKT1, NRAS, MAPK1, KRAS, TP53, PIK3CA, NFKB1, RB1	0.001037757
KEGG_PATHWAY	hsa04722:Neurotrophin signaling pathway	1.17E-06	AKT1, NRAS, MAPK1, BDNF, KRAS, MAPK14, TP53, PIK3CA, NFKB1, TP73	0.00125362
KEGG_PATHWAY	hsa04010:MAPK signaling pathway	2.69E-06	EGFR, AKT1, MAPK1, NRAS, BDNF, KRAS, ARRB2, MAPK14, GADD45G, TP53, NFKB1, TGFB1, RASA1	0.002871532
BIOCARTA	h_erythPathway:Erythrocyte Differentiation Pathway	3.22E-06	IL3, IL6, IL9, KITLG, TGFB1, EPO	0.003887971
REACTOME_PATHWAY	REACT_16888:Signaling by PDGF	9.59E-06	NRAS, MAPK1, KRAS, STAT5A, PIK3CA, THBS1, RASA1	0.007693883
PANTHER_PATHWAY	P00056:VEGF signaling pathway	1.68E-05	AKT1, NRAS, MAPK1, HIF1A, KRAS, MAPK14, VEGFA, PIK3CA	0.015328038
EC_NUMBER	3.4.21.59	4.75E-05	TPSAB1, TPSB2, TPSD1, TPSG1	0.030526793
KEGG_PATHWAY	hsa04370:VEGF signaling pathway	4.94E-05	AKT1, NRAS, MAPK1, KRAS, MAPK14, VEGFA, PIK3CA	0.052708911
REACTOME_PATHWAY	REACT_11061:Signaling by NGF	5.63E-05	AKT1, MAG, NRAS, MAPK1, KRAS, MAPK14, PIK3CA, NFKB1, SORCS3	0.045131702
KEGG_PATHWAY	hsa04664:Fc epsilon RI signaling pathway	6.17E-05	AKT1, NRAS, MAPK1, IL3, KRAS, MAPK14, PIK3CA	0.065876411
KEGG_PATHWAY	hsa04150:mTOR signaling pathway	9.18E-05	AKT1, MAPK1, HIF1A, STK11, VEGFA, PIK3CA	0.097993218
PANTHER_PATHWAY	P00005:Angiogenesis	9.21E-05	AKT1, NRAS, MAPK1, NOTCH1, HIF1A, KRAS, MAPK14, VEGFA, PIK3CA, BIRC5, RASA1	0.084242254
PANTHER_PATHWAY	P00018:EGF receptor signaling pathway	1.11E-04	EGFR, AKT1, NRAS, MAPK1, KRAS, MAPK14, STAT5A, PIK3CA, RASA1	0.101309728
KEGG_PAT	hsa04012:ErbB	1.14E-04	EGFR, AKT1, NRAS, MAPK1, KRAS,	0.121813791

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

HWAY	signaling pathway		STAT5A, PIK3CA	
KEGG_PATHWAY	hsa05216:Thyroid cancer	1.18E-04	NRAS, MAPK1, KRAS, TP53, CDH1	0.126019896
BIOCARTA	h_hcmvPathway:Human Cytomegalovirus and Map Kinase Pathways	1.34E-04	AKT1, MAPK1, MAPK14, NFKB1, RB1	0.161807042
KEGG_PATHWAY	hsa04115:p53 signaling pathway	3.30E-04	CDKN2A, GADD45G, RPRM, TP53, THBS1, TP73	0.352478227
KEGG_PATHWAY	hsa04660:T cell receptor signaling pathway	3.75E-04	AKT1, NRAS, MAPK1, KRAS, MAPK14, PIK3CA, NFKB1	0.399792892
BIOCARTA	h_telPathway:Telomeres, Telomerase, Cellular Aging, and Immortality	3.86E-04	EGFR, AKT1, KRAS, TP53, RB1	0.465899729
PANTHER_PATHWAY	P04393:Ras Pathway	3.87E-04	AKT1, NRAS, MAPK1, CDKN2A, KRAS, MAPK14, PIK3CA	0.353518988
REACTOME_PATHWAY	REACT_9417:Signaling by EGFR	5.16E-04	EGFR, NRAS, MAPK1, KRAS, PIK3CA	0.413199584
KEGG_PATHWAY	hsa04662:B cell receptor signaling pathway	5.22E-04	AKT1, NRAS, MAPK1, KRAS, PIK3CA, NFKB1	0.556225736
KEGG_PATHWAY	hsa05222:Small cell lung cancer	8.79E-04	AKT1, LAMB4, TP53, PIK3CA, NFKB1, RB1	0.934847379
BIOCARTA	h_RacCycDPathway:Influence of Ras and Rho proteins on G1 to S Transition	0.001228749	AKT1, MAPK1, PIK3CA, NFKB1, RB1	1.475202636
KEGG_PATHWAY	hsa04320:Dorso-ventral axis formation	0.001386389	EGFR, MAPK1, NOTCH1, KRAS	1.471198696
REACTOME_PATHWAY	REACT_6900:Signaling in Immune system	0.001592713	MAG, NRAS, MAPK1, KRAS, CD34, PECAM1, PIK3CA, NFKB1, CDH1	1.270531117
PANTHER_PATHWAY	P00059:p53 pathway	0.001789034	AKT1, CDKN2A, GADD45G, TP53, PIK3CA, THBS1, TP73	1.624912242
KEGG_PATHWAY	hsa04510:Focal adhesion	0.001957264	EGFR, AKT1, MAPK1, LAMB4, MET, VEGFA, PIK3CA, THBS1	2.071273278
KEGG_PATHWAY	hsa04620:Toll-like receptor signaling pathway	0.002011203	AKT1, MAPK1, IL6, MAPK14, PIK3CA, NFKB1	2.127799484
KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	0.002108909	EGFR, IL3, IL6, MET, VEGFA, IL9, KITLG, TGFB1, EPO	2.230116541

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PANTHER_PATHWAY	P00010:B cell activation	0.002214285	NRAS, MAPK1, KRAS, MAPK14, PIK3CA, NFKB1	2.007674836
REACTOME_PATHWAY	REACT_604:Hemostasis	0.002330368	MAG, NRAS, KRAS, VEGFA, PECAM1, PIK3CA, THBS1, TGFB1	1.854165541
KEGG_PATHWAY	hsa04630:Jak-STAT signaling pathway	0.002487246	AKT1, IL3, IL6, STAT5A, IL9, PIK3CA, EPO	2.62539532
PANTHER_PATHWAY	P00021:FGF signaling pathway	0.003144809	AKT1, NRAS, MAPK1, KRAS, MAPK14, PIK3CA, RASA1	2.840604013
BIOCARTA	h_il17Pathway:IL 17 Signaling Pathway	0.003157025	IL3, IL6, CD34, KITLG	3.750040719
BIOCARTA	h_stemPathway:Regulation of hematopoiesis by cytokines	0.003157025	IL3, IL6, IL9, EPO	3.750040719
PANTHER_PATHWAY	P04397:p53 pathway by glucose deprivation	0.003458487	AKT1, STK11, TP53, TP73	3.119958155
REACTOME_PATHWAY	REACT_498:Signaling by Insulin receptor	0.003756119	NRAS, MAPK1, KRAS, PIK3CA	2.973670175
KEGG_PATHWAY	hsa04110:Cell cycle	0.005068716	CDKN2A, GADD45G, PCNA, TP53, RB1, TGFB1	5.284031958
BIOCARTA	h_badPathway:Regulation of BAD phosphorylation	0.005426648	AKT1, MAPK1, IL3, KITLG	6.365799458
BIOCARTA	h_crebPathway:Transcription factor CREB and its extracellular signals	0.005426648	AKT1, MAPK1, MAPK14, PIK3CA	6.365799458
KEGG_PATHWAY	hsa04062:Chemokine signaling pathway	0.006278908	AKT1, NRAS, MAPK1, KRAS, ARRB2, PIK3CA, NFKB1	6.507599178
BIOCARTA	h_keratinocytePathway:Keratinocyte Differentiation	0.006958557	EGFR, CEBPA, MAPK1, MAPK14, NFKB1	8.094303481
KEGG_PATHWAY	hsa04914:Progesterone-mediated oocyte maturation	0.007248802	AKT1, MAPK1, KRAS, MAPK14, PIK3CA	7.477850945
KEGG_PATHWAY	hsa04640:Hematopoietic cell lineage	0.007248802	IL3, IL6, CD34, KITLG, EPO	7.477850945
KEGG_PATHWAY	hsa04210:Apoptosis	0.007548465	AKT1, IL3, TP53, PIK3CA, NFKB1	7.7757735
PANTHER_PATHWAY	P00036:Interleukin signaling pathway	0.007614161	AKT1, NRAS, MAPK1, IL6, KRAS, STAT5A, PIK3CA, RASA1	6.753977015
REACTOME	REACT_216:DNA	0.009001	MGMT, PCNA, XRCC1, ERCC1, ERCC2	6.996374809

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

_PATHWAY	Repair			
PANTHER_ PATHWAY	P00053:T cell activation	0.009738791	AKT1, NRAS, MAPK1, KRAS, PIK3CA, NFKB1	8.564577054
cluster5				
KEGG_PAT HWAY	hsa04120:Ubiquitin mediated proteolysis	6.19E-04	ANAPC1, RNF7, HUWE1, PARK2	0.370875716
EC_NUMBE R	6.3.2.-	0.069187942	HUWE1, PARK2	14.94398228
cluster6				
REACTOME _PATHWAY	REACT_1698:Metabli sm of nucleotides	5.07E-04	UMPS, ATIC, HPRT1	NaN
KEGG_PAT HWAY	hsa00983:Drug metabolism	0.001025542	ITPA, UMPS, HPRT1	0.657964401
BIOCARTA	h_mapkPathway:MA PKinase Signaling Pathway	0.00306271	MAP3K9, MAP2K4, MAP3K10	2.110065352
PANTHER_ PATHWAY	P05918:p38 MAPK pathway	0.005195037	MAP3K9, MAP2K4, MAP3K10	3.701051163
KEGG_PAT HWAY	hsa00230:Purine metabolism	0.012459659	ITPA, ATIC, HPRT1	7.749728301
EC_NUMBE R	2.7.11.25	0.030208436	MAP3K9, MAP3K10	16.18260742
KEGG_PAT HWAY	hsa00860:Porphyrin and chlorophyll metabolism	0.038330298	COX10, COX15	22.23317069
PANTHER_ PATHWAY	P00029:Huntington disease	0.04121774	MAP3K9, MAP2K4, MAP3K10	26.27015356
BIOCARTA	h_p38mapkPathway: p38 MAPK Signaling Pathway	0.045417275	MAP3K9, MAP2K4	27.61443797
BBID	100.MAPK_signaling _cascades	0.058659218	MAP2K4, MAP3K10	17.32821753
cluster7				
KEGG_PAT HWAY	hsa04914:Progestero ne-mediated oocyte maturation	2.08E-08	CCNB1, CCNB2, MAP2K1, MAPK3, BUB1, MAPK9, MAPK11, CDC25C, CDK2	2.13E-05
KEGG_PAT HWAY	hsa04114:Oocyte meiosis	1.46E-07	CCNB1, CCNB2, MAP2K1, MAPK3, BUB1, CDC20, AURKA, CDC25C, CDK2	1.50E-04
KEGG_PAT HWAY	hsa04110:Cell cycle	3.96E-07	CCNB1, CCND1, CCNB2, BUB1, CDC20, CDK4, CDC25C, MYC, CDK2	4.05E-04
KEGG_PAT HWAY	hsa05219:Bladder cancer	3.68E-06	CCND1, MAP2K1, MAPK3, CDK4, MYC, MMP2	0.00376453

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

KEGG_PATHWAY	hsa05210:Colorectal cancer	7.34E-06	FOS, CCND1, MAP2K1, BAX, MAPK3, MAPK9, MYC	0.007502731
KEGG_PATHWAY	hsa05200:Pathways in cancer	1.02E-05	FOS, CCND1, MAP2K1, BAX, MAPK3, MAPK9, NFKB2, CDK4, MYC, MMP2, CDK2	0.010452597
KEGG_PATHWAY	hsa04010:MAPK signaling pathway	1.43E-05	FOS, MAP3K5, MAP2K1, MAPK3, MAP3K8, MAPK9, MAPK11, NFKB2, MAPK7, MYC	0.014636534
REACTOME_PATHWAY	REACT_152:Cell Cycle, Mitotic	1.79E-05	CCNB1, CENPM, CCND1, CCNB2, BUB1, CDC20, AURKA, AURKB, CDK4, CDC25C, CDK2	0.012929573
KEGG_PATHWAY	hsa04912:GnRH signaling pathway	1.79E-05	MAP2K1, MAPK3, MAPK9, MAPK11, MAPK7, MMP14, MMP2	0.018347497
EC_NUMBER	2.7.11.24	2.02E-05	MAPK3, MAPK9, MAPK11, MAPK7	0.012135676
KEGG_PATHWAY	hsa04620:Toll-like receptor signaling pathway	2.13E-05	FOS, MAP2K1, MAPK3, MAP3K8, MAPK9, MAPK11, TLR4	0.02182043
PANTHER_PATHWAY	P00006:Apoptosis signaling pathway	2.68E-05	FOS, TNFSF10, MAP3K5, ATF3, BAX, MAPK3, MAPK9, NFKB2, MAPK7	0.02355655
KEGG_PATHWAY	hsa04660:T cell receptor signaling pathway	3.13E-05	FOS, MAP2K1, MAPK3, MAP3K8, MAPK9, MAPK11, CDK4	0.032014891
PANTHER_PATHWAY	P00054:Toll receptor signaling pathway	3.16E-05	MAP2K1, MAPK3, MAP3K8, MAPK9, MAPK11, TLR4, NFKB2	0.02781923
KEGG_PATHWAY	hsa04115:p53 signaling pathway	4.05E-05	CCNB1, CCND1, CCNB2, BAX, CDK4, CDK2	0.041438576
KEGG_PATHWAY	hsa04722:Neurotrophin signaling pathway	6.84E-05	MAP3K5, MAP2K1, BAX, MAPK3, MAPK9, MAPK11, MAPK7	0.069959614
BIOCARTA	h_mapkPathway:MAPKinase Signaling Pathway	3.55E-04	FOS, MAP3K5, MAP2K1, MAPK3, MAP3K8, MAPK9, MAPK11, MAPK7	0.400316292
KEGG_PATHWAY	hsa05216:Thyroid cancer	6.11E-04	CCND1, MAP2K1, MAPK3, MYC	0.622621821
PANTHER_PATHWAY	P00052:TGF-beta signaling pathway	6.33E-04	FOS, FOXF1, FOXA1, MAPK3, MAPK9, MAPK11, MAPK7, FOSL1	0.555596842
KEGG_PATHWAY	hsa05212:Pancreatic cancer	7.67E-04	CCND1, MAP2K1, MAPK3, MAPK9, CDK4	0.781138732
KEGG_PATHWAY	hsa05220:Chronic myeloid leukemia	8.95E-04	CCND1, MAP2K1, MAPK3, CDK4, MYC	0.91105286
PANTHER_PATHWAY	P00010:B cell activation	0.001400729	FOS, MAP2K1, MAPK3, MAPK9, MAPK11, NFKB2	1.225108429
BIOCARTA	h_p53Pathway:p53 Signaling Pathway	0.003024948	CCND1, BAX, CDK4, CDK2	3.364924295

bioRxiv preprint doi: https://doi.org/10.1101/037531; this version posted January 21, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

KEGG_PATHWAY	hsa05213:Endometrial cancer	0.003370738	CCND1, MAP2K1, MAPK3, MYC	3.393288113
KEGG_PATHWAY	hsa05223:Non-small cell lung cancer	0.003753245	CCND1, MAP2K1, MAPK3, CDK4	3.77171584
PANTHER_PATHWAY	P00035:Interferon-gamma signaling pathway	0.004182552	MAPK3, MAPK9, MAPK11, MAPK7	3.618778393
KEGG_PATHWAY	hsa05221:Acute myeloid leukemia	0.004595499	CCND1, MAP2K1, MAPK3, MYC	4.600275355
PANTHER_PATHWAY	P00034:Integrin signalling pathway	0.005122253	MAP3K5, MAP2K1, ITGAX, MAPK3, MAPK9, MAPK11, MAPK7, ITGAM	4.415663351
KEGG_PATHWAY	hsa04621:NOD-like receptor signaling pathway	0.005544037	CCL2, MAPK3, MAPK9, MAPK11	5.525680097
KEGG_PATHWAY	hsa05214:Glioma	0.005798209	CCND1, MAP2K1, MAPK3, CDK4	5.772273682
REACTOME_PATHWAY	REACT_1538:Cell Cycle Checkpoints	0.007137442	CCNB1, CCNB2, CDC20, CDC25C, CDK2	5.054206969
KEGG_PATHWAY	hsa05218:Melanoma	0.008084875	CCND1, MAP2K1, MAPK3, CDK4	7.964786431
BIOCARTA	h_RacCycDPathway: Influence of Ras and Rho proteins on G1 to S Transition	0.008275776	CCND1, MAPK3, CDK4, CDK2	8.961786045
BIOCARTA	h_cellcyclePathway: Cyclins and Cell Cycle Regulation	0.009291865	CCNB1, CCND1, CDK4, CDK2	10.010086