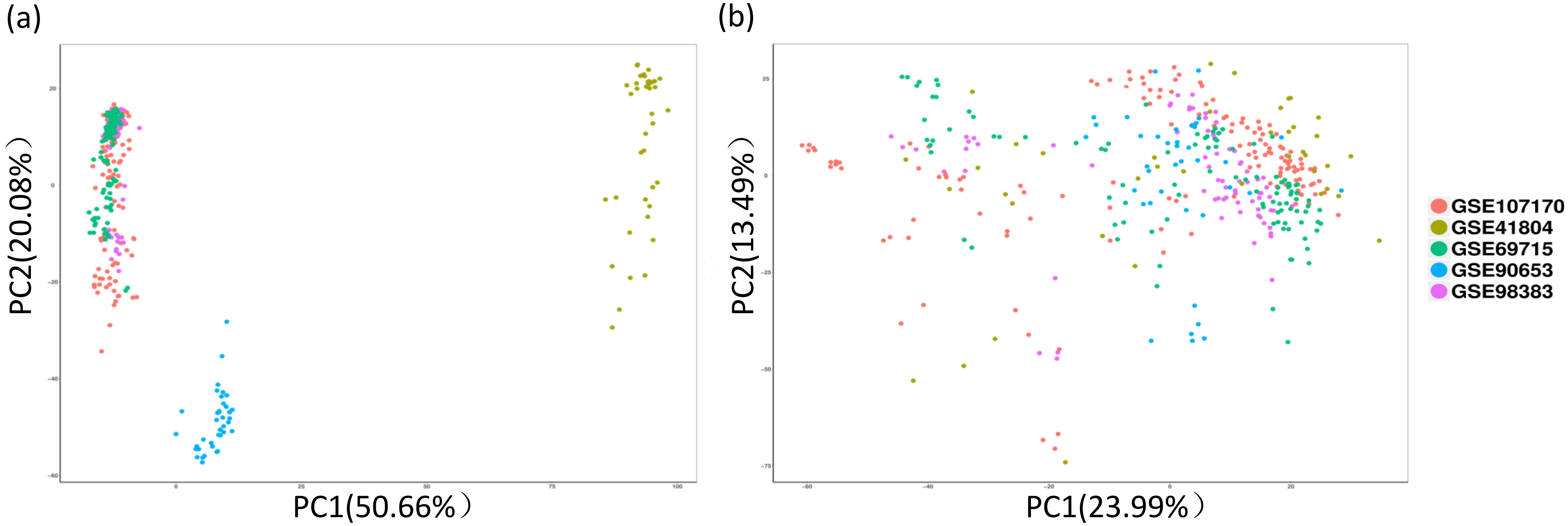
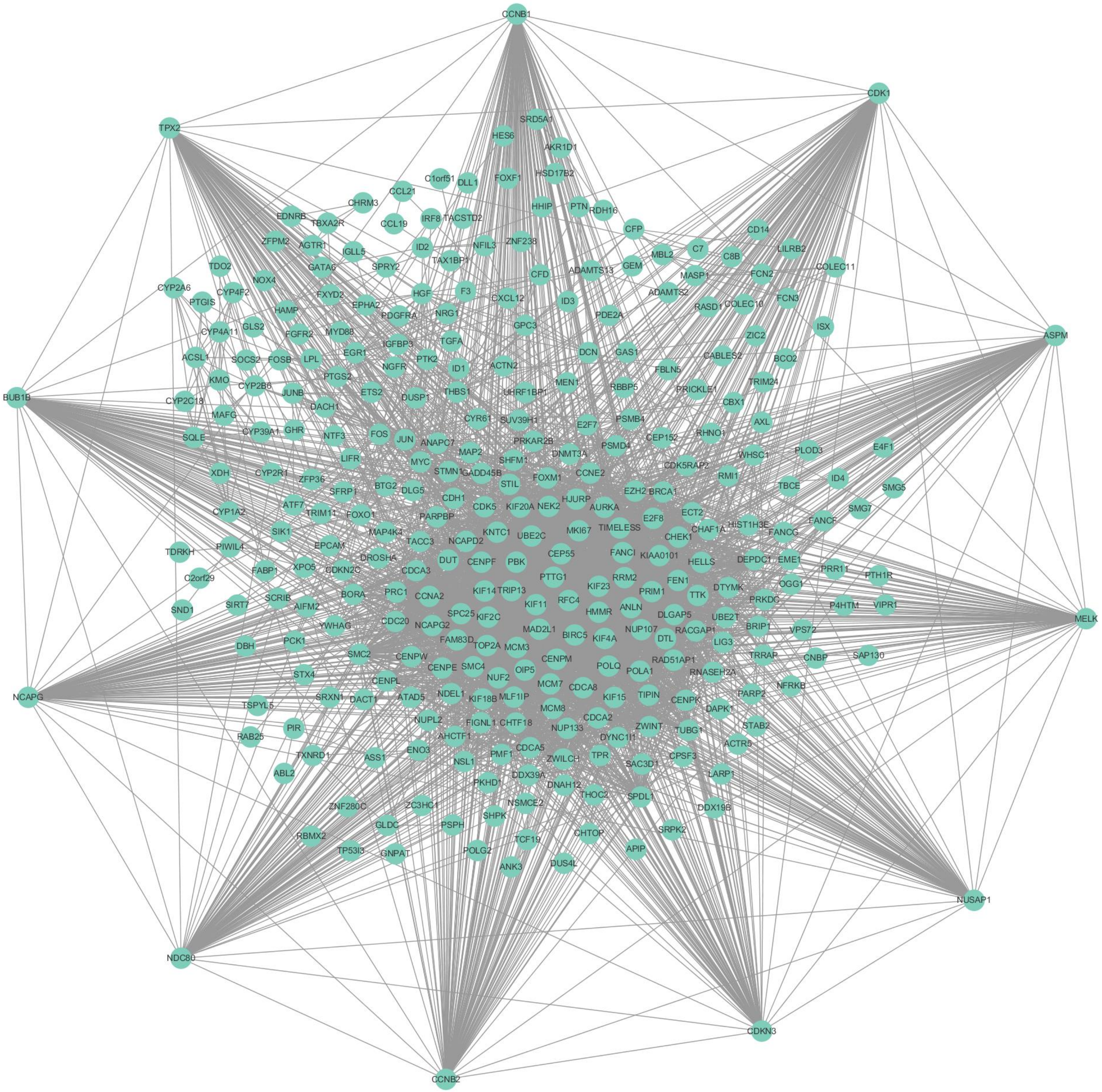
Title: Feature selection with the Fisher score followed by the Maximal Clique Centrality algorithm can accurately identify the hub genes of hepatocellular carcinoma

Authors : Li Chengzhang, Xu Jiucheng


(a)
PC2(20.08%)
20
0
-20
-40
-60
0
25
50
75
100
PC1(50.66%)
(b)
PC2(13.49%)
25
0
-25
-50
-75
-60
-40
-20
0
20
PC1(23.99%)
GSE107170
GSE41804
GSE69715
GSE90653
GSE98383

Supplementary Figure 1. Principal component analysis (PCA) of the integrated gene expression profiles. （a）PC1 and PC2 were the top 1 and 2 components, respectively, before the removal of the batch effect. A clear batch effect was identified prior to the batch effect correction. （b）PC1 and PC2 were the top 1 and 2 components, respectively, after the removal of the batch effect. The batch effect was eliminated after the batch effect correction.


CCNB1
CDK1
TPX2
SRD5A1
HES6
AKR1D1
C1orf51
DLL1
FOXF1
HSD17B2
CCL21
CHRM3
HHIP
PTN
RDH16
EDNRB
TBXA2R
IRF8
TACSTD2
CCL19
CFP
CD14
ZFPM2
AGTR1
IGLL5
ID2
NFIL3
ZNF238
MBL2
C7
C8B
LILRB2
TDO2
GATA6
SPRY2
TAX1BP1
CFD
ADAMTS13
GEM
MASP1
FCN2
COLEC11
CYP2A6
CYP4F2
NOX4
FXYD2
HGF
F3
CXCL12
ID3
ADAMTS2
PTGIS
HAMP
EPHA2
PDGFRA
NRG1
PDE2A
RASD1
COLEC10
FCN3
ISX
CYP4A11
GLS2
FGFR2
MYD88
GPC3
ZIC2
ASPM
ACSL1
SOCS2
FOSB
LPL
EGR1
IGFBP3
TGFA
DCN
GAS1
CABLES2
BCO2
PTK2
KMO
CYP2B6
PTGS2
NGFR
ID1
ACTN2
FBLN5
TRIM24
BUB1B
CYP2C18
JUNB
ETS2
DUSP1
THBS1
UHRF1BP1
MEN1
RBBP5
PRICKLE1
MAFG
DACH1
CBX1
SQLE
CYP39A1
GHR
NTF3
FOS
JUN
CYR61
SUV39H1
E2F7
PSMB4
AXL
RHNO1
ANAPC7
PRKAR2B
CEP152
PSMD4
MAP2
DNMT3A
WHSC1
PLOD3
E4F1
LIFR
MYC
SHFM1
XDH
CYP2R1
STMN1
GADD45B
CCNE2
CDK5RAP2
ZFP36
BTG2
DLG5
STIL
FOXM1
EZH2
BRCA1
RMI1
SFRP1
CDH1
HJURP
TBCE
ID4
SMG5
ATF7
TRIM11
CDK5
KIF20A
NEK2
AURKA
CYP1A2
FOXO1
PARPBP
E2F8
ECT2
SMG7
MAP4K4
KNTC1
UBE2C
MKI67
TIMELESS
CHAF1A
HIST1H3E
FANCF
SIK1
TACC3
CHEK1
FANCG
PIWIL4
EPCAM
NCAPD2
CEP55
FANCI
KIAA0101
TDRKH
DROSHA
CENPF
PBK
HELLS
DEPDC1
EME1
FABP1
XPO5
CDKN2C
CDCA3
DUT
PTTG1
RRM2
FEN1
PRR11
PTH1R
C2orf29
PRC1
KIF14
TRIP13
KIF23
PRIM1
DTYMK
OGG1
SIRT7
SCRIB
BORA
CCNA2
RFC4
TTK
PRKDC
AIFM2
KIF11
ANLN
P4HTM
VIPR1
SND1
YWHAG
CDC20
SPC25
HMMR
DLGAP5
UBE2T
KIF2C
NCAPG2
MAD2L1
NUP107
BRIP1
MELK
DBH
PCK1
FAM83D
BIRC5
KIF4A
RACGAP1
VPS72
TOP2A
MCM3
DTL
LIG3
SMC2
CENPW
CENPM
RAD51AP1
TRRAP
CNBP
SAP130
STX4
CENPE
SMC4
OIP5
POLQ
POLA1
RNASEH2A
NCAPG
CENPL
NUF2
MCM7
CDCA8
NFRKB
PARP2
TSPYL5
SRXN1
DACT1
ATAD5
NDEL1
MLF1IP
KIF18B
KIF15
TIPIN
CENPK
DAPK1
MCM8
NUPL2
FIGNL1
CHTF18
CDCA2
PIR
NUP133
ZWINT
STAB2
RAB25
AHCTF1
DYNC1I1
TUBG1
TXNRD1
ASS1
ENO3
CDCA5
ACTR5
NSL1
PMF1
ZWILCH
TPR
SAC3D1
CPSF3
ABL2
DDX39A
PKHD1
DNAH12
LARP1
ZNF280C
ZC3HC1
THOC2
SPDL1
NSMCE2
DDX19B
GLDC
PSPH
SHPK
CHTOP
RBMX2
TCF19
SRPK2
TP53I3
GNPAT
POLG2
APIP
ANK3
DUS4L
NUSAP1
NDC80
CDKN3
CCNB2

Supplementary Figure 2. Protein-protein interaction network of the feature genes. The network has 374 nodes and 4478 edges. The nodes represent the selected feature genes, and the edges represent the interactions between two genes.