

GEMs

Analyzing omics data in the context of metabolism

NBIS Omics Integration and Systems Biology workshop
Fall 2020, Lund University

Jonathan Robinson

National Bioinformatics Infrastructure Sweden (NBIS)
Science for Life Laboratory (SciLifeLab)
Chalmers University of Technology
jonathan.robinson@scilifelab.se

Information paradox



ACE2

Information paradox

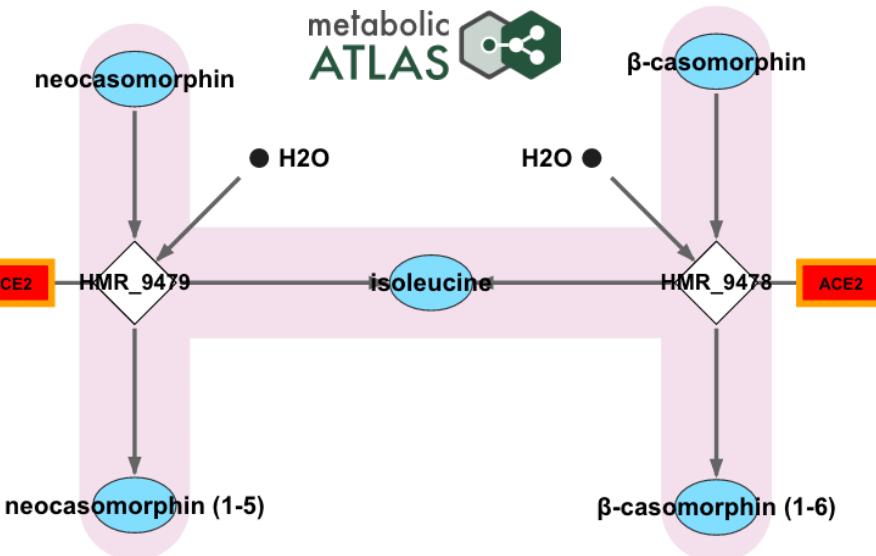

 Article Talk

Angiotensin-converting enzyme 2

From Wikipedia, the free encyclopedia

"ACE2" redirects here. For other uses, see [Ace 2 \(disambiguation\)](#).

Angiotensin-converting enzyme 2 (ACE2)^[5] is an [enzyme](#) attached to the [cell membrane](#) kidney, and intestines.^{[6][7]} ACE2 lowers blood pressure by catalyzing the [hydrolysis](#) of angiotensin (1–7) (a [vasodilator](#)).^{[8][9][10]} ACE2 counters the activity of the related [angiotensin](#) amount of angiotensin-II and increasing Ang(1-7),^[11] making it a promising drug target for tr



[Circulating ACE2 in Cardiovascular and Kidney Diseases.](#)

1 Anguiano L, Riera M, Pascual J, Soler MJ.

 Cite [Curr Med Chem. 2017;24\(30\):3231-3241. doi: 10.2174/0929867324666170414162841.](#)

 PMID: 28413960 [Review](#).

Share Given that ACE2 counterbalances the effects of Ang II, it has been proposed as a biomarker in kidney disease patients. Circulating ACE2 has been studied in human and experimental studies under physiological and pathological conditions and different techniques have b ...



[Angiotensin-converting enzyme 2 \(ACE2\) in disease pathogenesis.](#)

2 Imai Y, Kuba K, Ohto-Nakanishi T, Penninger JM.

 Cite [Circ J. 2010 Mar;74\(3\):405-10. doi: 10.1253/circj.cj-10-0045. Epub 2010 Feb 4.](#)

 PMID: 20134095 [Free article](#). [Review](#).

Share Importantly, ACE2 has been identified as a key SARS-coronavirus receptor and plays a protective role in SARS pathogenesis. Furthermore, the recent explosion of research into the ACE2 homolog, collectrin, has revealed a new physiological function of ACE2 as an ...

[ACE2 - from the renin-angiotensin system to gut microbiota and malnutrition.](#)

3 Perlot T, Penninger JM.

 Cite [Microbes Infect. 2013 Nov;15\(13\):866-73. doi: 10.1016/j.micinf.2013.08.003. Epub 2013 Aug 17.](#)

 PMID: 23962453 [Free PMC article](#). [Review](#).

Summary

Official Symbol ACE2 provided by [HGNC](#)

Official Full Name angiotensin I converting enzyme 2 provided by [HGNC](#)

Primary source HGNC:HGNC:13557

See related [Ensembl:ENSG00000130234](#) [MIM:300335](#)

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as ACEH

Summary The protein encoded by this gene belongs to the angiotensin-converting enzyme family of dipeptidyl carboxypeptidases and has considerable homology to human angiotensin 1 converting enzyme. This secreted protein catalyzes the cleavage of angiotensin I into angiotensin 1-9, and angiotensin II into the vasodilator angiotensin 1-7. ACE2 is known to be expressed in various human organs, and its organ- and cell-specific expression suggests that it may play a role in the regulation of cardiovascular and renal function, as well as fertility. In addition, the encoded protein is a functional receptor for the spike glycoprotein of the human coronavirus HCoV-NL63 and the human severe acute respiratory syndrome coronaviruses, SARS-CoV and SARS-CoV-2, the causative agent of coronavirus disease-2019 (COVID-19). [provided by RefSeq, Aug 2020]

Annotation information Note: This gene has been reviewed for its involvement in coronavirus biology, and is involved in SARS-CoV-2 infection.

Expression Biased expression in small intestine (RPKM 93.7), duodenum (RPKM 69.0) and 5 other tissues [See more](#)

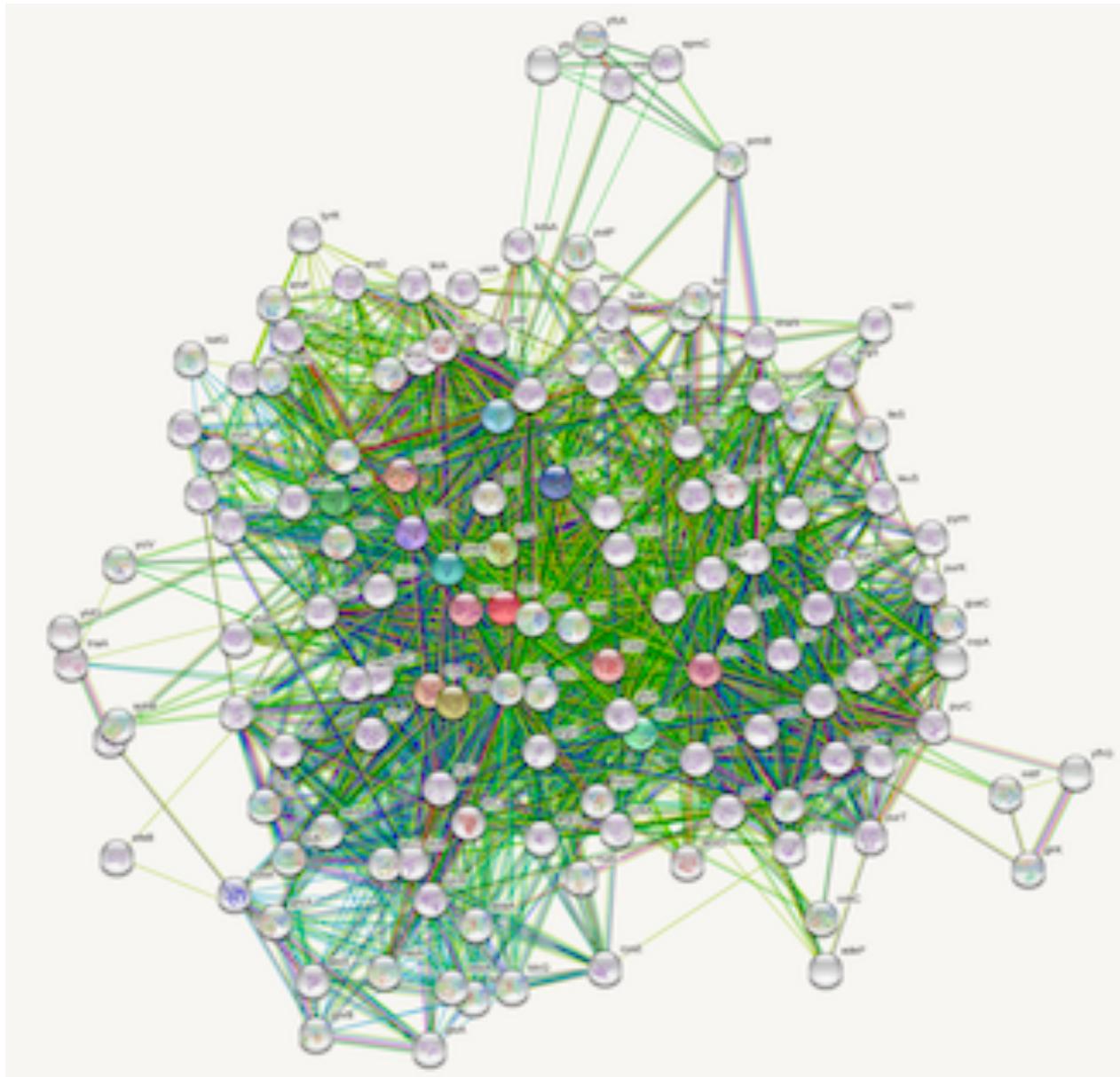


Information paradox



ENOPH1, SLC25A2, GMPPB, SLC1A4, EGFL8, HDC, DNM1L, RHCG, SLC24A4, FABP4, PDSS1, ACSBG1, UGT3A1, TAOK1, CDY1B, PPIE, GCLM, GMPR2, USP9Y, DGKD, ATP2A2, FABP2, ZDHHC7, GLRX2, DPM3, NUP155, GPX1, SLC4A5, POLI, GCH1, APOE, PLPP7, RIPK3, HPSE, PLA2G1B, PYGL, CDK8, MGST1, DGAT2L6, PPP5C, TXNRD3, MACROD2, FABP4, DDR1, USP35, NARS, HERC1, WWP1, PIGF, MACROD1, SUOX, DGKG, PG51, AURKB, CDY1, TKFC, PIM1, SLC1A4, GLDC, AFMID, FAR1, PHYKPL, ATP2B1, PTPDC1, LCT, ASPG, ACOT1, POLR1C, HEMK1, CDA, UPP2, PARP10, SLC15A1, ASNS, TGFB2, POLR2F, CARS, CBL, SLC4A9, INPP4A, SUCLG1, MTAP, PKMYT1, CBL1, MAP2K2, MARS, TECRL, SLC26A3, METTL7B, ME1, AK2, PLA2G12A, DHRS9, EPHA5, GGACT, GGT5, ACYP2, PI4K2A, SLC25A2, USP17L25, TRMT11, DUSP15, PTGR1, PPP3R1, CA1, FN3K, UBE2F, EFL1, TYMS, CNT4, PFKP, RNF126, PRKCI, TAF9, FDPS, SLC19A3, AGK, BCKDHB, CALM2, HK1, UPRT, SRD5A1, SYVN1, GLA, CTDSPL, CYP2J2, ACOT9, PDE8B, AP000812.4, IDH3B, SULT1B1, MAP3K21, HMGCS1, JMD7-PLA2G4B, CA3, SLC25A3, ACAT1, NEK9, NEK6, PPA1, ACP7, IDH1, PIK3R4, KAT6A, ASRGL1, PTPN11, SLC22A15, PTPRR, RPEL1, PRKCSH, DOT1L, TRMT13, GPX6, NLK, PPIE, ALDH1A1, ASA2, PEAK1, PAK6, CTSF, ACE, SDR42E2, IMPA1, SLC2A11, CYLD, PPIP5K2, PRKD3, PTGS1, GCNT1, GC, UGT1A9, HSD11B1L, SLC16A12, USP17L5, BMPR2, SLC39A8, HS3ST1, GRK1, POLB, NCOA6, PKM, PIGN, SLC22A1, EBP, UGT1A8, SPR, PPP3CB, PDE10A, TMEM91, CTSG, NUP210L, SLC38A5, NDUF4, FKBP4, SLC22A11, ATP1A3, GMPPB, MTMR14, SLC32A1, ADSL, PPWD1, PPIA, CMAS, STK10, RENBP, NDUF57, EPHA10, ALAS1, TGM7, TBK1, HSD17B2, MT-ATP8, UQCRCFS1, SDSL, COP1, ATP5F1C, NEDD4L, ALG10B, ATP1B3, DUSP11, STAR5D, MTHFD2L, ABHD14A-ACY1, POLL, RAD18, PCTP, METTL3, AKR1C3, STT3B, B3GNT4, POLD2, SCD, GGACT, UQCRC2, NEK9, KMT5C, QPRT, AKT3, RPUSD4, APOB, DPEP2, CYP11B2, TPTE2, FUCA1, NNMT, GSTM4, RIPK2, PPEF2, SLC52A1, SLC2A1, LALBA, SLC7A3, MYO3B, DUSP23, PYGL, HS3ST2, RNF115, DPYD, ALG3, GLYATL3, MAP3K2, PTGIS, GPCPD1, TOPORS, F13B, GLYATL2, STARD3, USP35, IVD, SLC26A11, PIGN, ALG8, DYRK1A, CYP4A11, SLC16A13, TYW3, GLYATL2, ITPA, UBR3, NDUF4, INPP5D, PNLLPRP3, TFB1M, RNF14, NPC1, MYLK, CPNE7, SLC7A10, PDE4A, CYP2A7, LRAT, ATP8B1, PTPN14, GUCY1B1, PTPN9, NME7, PDK1, SLC28A2, TAF1, PDGFRB, NUP37, PRKCB, TENT2, RPK1, RNF185, PON1, LIPF, MAP4K3, HEPH, PGD, NGLY1, PON1, AL845331.2, FGA, SLC6A9, SLC8A2, CTH, IDH1, TRIP12, PAH, SLC47A1, ST6GALNAC2, CTDSP1, TRIP12, PAK6, MAST2, GMDS, NTRK1, SLC26A11, PIGA, PIGG, DOLPP1, NAGK, MAOA, MARCH6, GCLC, TKT, MGST1, PRKDC, ATP8B1, DUSP8, ADPRH, PDF, PCYT1A, PHOSPHO2, PIGS, SUV39H2, COPB1, STARD10, SLC25A21, BCDIN4, PRPS1L1, LRRK2, SGK2, BIRC6, FLAD1, VRK3, MARK3, ASA2, TMX3, CDS1, NDUFB9, MAPK10, MAGEA2, OXCT1, FOLR1, TPSB2, RNF31, PI4K2A, ALDH8A1, SLC25A28, CHPF2, ZDHHC21, GGACT, MAP3K4, IDS, PGAM1, ASPG, ACOX2, PRDX5, CHST6, MOGAT2, PDXN, GSTA3, TGM7, MGAT4C, DCT, SLC31A1, Zap70, SIAH2, COX8A, SLC37A4, ATP2B4, COX6A2, SLC1A1, PARP16, PCBD1, PRPS1L1, PLA2G4C, NDUF12, MYLK3, NEK5, USP24, GMPR2, BUB1, SLC1G1, JAK1, LALBA, IRAK1, B3GNT2, METTL2B, RPUSD4, ALOX5, PRKCZ, ACOT4, CSAD, PPWD1, ABCC5, MOGAT2, COX7B2, RNF19B, HERC3, ALG8, ABHD14A-ACY1, B4GALNT3, ST6GALNAC6, TRAF7, AK8, NTRK1, ALOX5, USP8, DPYD, STAMB, ST6GALNAC2, ATP5F1B, NCOR2, DUOX1, TPSAB1, ADCY10, LDHA, ACSM2B, UST, HSD17B3, POC1B-GALNT4, MACROD1, GPHN, SLC6A4, AWAT2, CRLS1, FBLL1, SLC15A2, ACAD8, UBA7, THOP1, PTK2B, CDC42BPA, MARS, GSTO2, ARSA, COX7B, XYL72, PLIN3, CDK14, GPL1, MAP3K7, TGDS, PSAP, TAF9, DPEP1, GPAM, TPSB2, UPRT, PARP16, NNT, USP28, CNOT4, PIGT, BMPR2, SLC7A11, GRK1, ENO1, RORA, ABCC4, PGLS, HS2T1, ALG2, RDH8, ST6GALNAC5, KAT5, SYVN1, MAN1A1, QARS, LYPLA1, TTC3, DPM3, GBGT1, HSD17B8, PFKFB4, NAMPT, RNF25, MOCS1, UBEQ2K1, PDE7A, CAMK1G, PJA1, GIF, DMPK, MAP3K2, SLC32A2, MCAT, LCMT1, CLPS, ATP6V0A1, GYG2, HIRP3, GPT2, KIT, INSL3, NDUF13A, G6PC, SLC44A3, AK7, SCD5, PTPRA, USP9Y, DLD, FM05, CDK17, GSTA5, DLX4, HEXA, GPX3, PYCR2, CP, SLC26A3, PTRH1, PLA2G16, FKB10, PDGFR, RPAP2, ATP5MC1, TAOK2, TS1, AC5S1, TAF1, NSDHL, UCP1, SLC27A3, ADH1C, FMO1, DIO2, MAP2K3, SLC9A7, MARCH6, NUDT10, KPNA6, RRM2, EIF2AK3, POLR1B, DHFR2, PDHB, USP29, HK1, SLC03A1, BLVRB, DMPK, PHKB, CHSY3, ADH1A, PASK, USP17L25, AC093512.2, SLC25A15, SLC1A3, GALNT9, SLC25A5, CA1, PRKCQ, CLK2, CDK20, HAGH, NEDD4L, ATP6V1G2, CDK15, QPRT, SLC7A6, AMT, ENPP7, AKR1C4, PTPDC1, PXYLP1, CBX1, RNF123, NUP210L, RFWD3, EYA2, PORCN, GGTL1, HERC1, USP13, SLC34A1, CYP2F1, CLPS, NMRK2, GALNT6, UCP1, CA7, NMRK1, INPP5D, RPS6KA5, SMOX, EPX, SLC33A1, LEENP, GLB1, HUWE1, KMO, TKT1, LRTOMT, CHPF, SIN3A, CAD, ADCY4, MPI, CDK6, ASA2, TXNRD2, GGT2, RHCG, DYRK4, TRIM5, SULT1A1, NUP210, LPCAT4, MAP4K1, ERBB3, STAR, UBE2R2, PHGDH, ALB, BDH2, MKNK2, SLC15A4, AHCYL2, PIGO, CDK2, RNF185, ADO, PDE6A, HERC1, SARS, PPIP5K2, ATP5MPL, SLC16A1, IARS, SULT1E1, CDK10, TRPM6, SRPK1, FHIT, GGACT, PRPS1, MKRN2, NUP62CL, CPT1B, ACAT1, CBR4, PTPRK, MVK, GLYAT, GALNT9, PARP14, GLIPR1, PARP11, NLN, DUSP18, PARP6, MT-ND5, SMG1, MARK1, GLT8D1, FIG4, STRA6, UROC1, NAT1, USP46, SLC25A18, UBR4, SLC27A6, MAP3K14, HARS, PHGDH, DCTPP1, ACAT1, FUT10, SLC16A3, NUAK1, ST3GAL3, UXS1, EIF2AK4, TPTE, METTL21A, AMPD1, ALPP, SLC26A11, IP6K2, GALNT16, HSD17B6, NDUF6B, RPS6KB1, MGMT, PAD11, IMPDH2, PTPRS, MGAT4D, STK32A, PRKAR2B, ASH1L, SRM, STARD6, DNMM1, PIK3CG, ALG9, PPI4, DUSP4, UBA7, AKR1C1, ASPH, B3GNT9, ZDHHC3, MOCS1, ROCK1, PTPRK, SLC7A3, MGAT5B, SLC4A3, POLR2E, QPRT, SRPK2, PTDSS1, SLC35D2, MAT2B, SLC5A7, MAP3K21, SLC25A22, GCAT, NDST2, HSD11B1, MAPKAPK2, CA12, CTDSP1, ATP5F1C, GALNT3, SLK, STARD3, CASK, SLC22A4, LDHAL6A, INPP5A, PARP2, ATP6V1G3, GBGT1, PPIH, B3GLCT, DNMT1, DUSP8, MOCS1, RNF19A, UHMK1, USP12, RPS6KA6, GLS2, NAA20, RNF167, RPE, SLC22A4, SLC1A2, MDP1, SLC30A7, SMURF2, GLYCTK, COMT, NUDT11, SETD7, ACOT13, GDA, NEU2, SLC9A5, B3GAT1, DUSP12, ENO1, CRYZ, PINK1, GPX7, RBP3, SDSL, NPEPPS, ENTPD4, PCYOX1, PPT1, AC0X3, ATP1A2, PIGO, AQP5, SLC2A6, TARS, PIGM, HERC2, CTPS1, ACADS, ZDHHC4, SOD2, CYBRD1, GALNT4, PSKH2, AKR7A3, MCC2C, ESCO2, GNMT, POLR2H, PTPRS, STK32C, TSSK6, INPP4A, AMDHD2, CKB, USP17L29, SLC01B3, AP000812.4, PDHA1, PPIP5K1, DUSP4, GAL3ST2, CSKMT, STK32A, NDUFB1, ATP2A3, NDUF11, NEK8, SLC39A10, SDHB, PJA1, CLK3, CYP39A1, PLCB4, ULK1, SRD5A3, PKN1, CYLD, SUV39H1, GSS, HIRIP3, PRMT2, DHCR24, GFPT2, RBKS, HUWE1, PRKAR1A, RNF103, PAPOLB, DUPD1, COX6C, RFWD3, ATP6V0C, NAA10, METTL14, CDK7, INPP5F, DUSP2, STK17A, EYA3, UGT2B11, DHRS9, FABP3, STK33, ZDHHC11B, SLC17A8, CBLB, USP47, METTL7A, KMT2D, EYA3, CS, SLC6A20, CDK6, LIPE, AK9, ZNRD1, ST3GAL4, ZDHHC15, PGP, XDH, USP40, RNF2, INSRR, STK32B, MGAT4B, SLC25A5, ST6GALNAC3, ACACA, ALG2, FLAD1, NDUF12, HS3ST1, PDE7B, PGAM5, BCAT1, KALRN, PRPS1L1, ATP5MC1, CHMP4A, NDUFAF5, ASMT, SRPK2, USP34, ALDOA, PTPN7, ABCC3, SLC35B4, BRSK2, MAP3K6, SLC6A6, VAPB, EIF4EBP3, NR1D1, PFKFB1, TUT1, NHLRC1, PPI1L2, BMP2K, B3GAT3, DOLP1, FIG4, TRMU, ALDH2, SLC23A1, ALDH5A1, MLYCD, PTPRQ, MT-CYB, PTPRR, NAT1, SLC25A38, PPP1R1A, AHCYL2, SIRT5, ACVR2A, MCAT, UGT2B11, PDHB, ENO3, PRKCA, SLC4A1, SMS, PNLLPRP3, VAPA, EGFR, LTK, GALE, CARS, MAN1B1, ATP2B1, FABP3, DGUOK, PAH, PRKCH, B3GAT1, MKV, ADCY3, PTPRU, DCLK1, PGLYRP1, HSD3B1, SLC13A2, SLC6A20, CIT, UQCRC1, NEK4, UBE2NL, MINPP1, MUSK, TUT7, POLR3GL, CSNK1A1L, DUSP9, GLYCTK, ATP6V1E2, KMO, SLC39A4, NQO2, PTPN13, KAT6A, PANK2, CLK4, CYP4V2, PARP6, PRKACA, GLYATL2, PPP6C, PYURF, UBE2D1, BARD1, SLC17A3, GCSH, RDH5, ITPKA, HSD11B1L, UAP1, ASA2, DYRK3, SMDP2, MAPKAPK3, ABCD3, TGM6, PTPRK, RCHY1, TICRR, PMM2, WEE1, POLR2A, NDUF18, PPT1, PLCL1, CPA6, CTSV, SAPCD1, TPH2, XPO6, EPHA1, MERTK, MPO, ATP5MPL, GGT7, MAGEA2B, TREX1, FHIT, MAGEA2B, SLC35A1, MASTL, WDYHV1, TSSK3, ST20, CHST7, ATP2C2, TMEM54, AANAT, AWAT1, GALNT12, POLR3E, PPP2CB, MAPK4, TGM6, PPCDC, SLC39A10, USP9X, GLT6D1, TXN2, FTCD, PSPH, HS6ST3, TRIM33, ENTPD4, PLCG2, GUCY1B1, GFM1, GSTA4, CTSL, SLC5A11, MMAB, PLCL2, SLC4A1, SLC2A5, GLYAT, UCK1, CTSB, CTSG, AKR1B15, HS3ST3B1, RNF144A, CLK2, RNF31, TRHDE, ME2, GFPT2, OLAH, GALNT13, HSD17B7, GSTM4, DGKD, PIGY, STAR, NDST3, SIRT6, NAAA, SLC5A7, RHBG, IREB2, HS3ST2, ID1, RPEL1, PAFAH1B2, G6PC2, PRKCZ, UGT2B7, PDIA1, DOLP1, TAF9, STAR8, SLC51A1, TLK1, PIN1, MEPCE, PAPOLB, PCTP, ATP5ME, PRODH2, ME3, DUSP4, MIAS, AKR1B15, PIP4P1, AGMAT, PARP3, TRMT12, USP14, PTPRK, CTDSP2, WARS, GSTK1, CDC42BPG, F13B, ENPP2, METTL3, MAP4K4, UBE2Q2, POLR2C, GM2A, ABCC8, FXYD2, DHFR, PPM1G, CA5B, TYRP1, MARS2, CSNK1G1, PARP4, FLT3, SIN3B, HSD3B1, L1CAM, CYP27B1, SLC22A13, TST, CDY1B, NEDD4, MMAB, GK3P, CAMK1, FIG4, GMPS, ATP5MC2, IRAK4, PPP3R1, ZDHHC19, SLC22A7, FHL2, B3GAT3, SLC15A3, SIK3, DPM2, DPEP1, GK, POLD3, PCK2, ELOVL2, PRDX5, RAG1, FKBP14, PRKN, SLC22A3, CYP2C19, AGK, GPAT2, PRDX1, FNTA, GYG1, ATG9B, EGFL8, FMO5, MGAT2, PBK, SIK2, CLK2, NDUF4, FADS1, STK39, GUK1, CT5Z, HS3ST6, SLC31A1, PLK4, PTPN5, STAR3D, CYP2R1, MGST2, IGF1, LGALS13, OBSN1, TOPORS, LIPT2, PDPK1, PPT1, LGSN, PKN3, HS3ST5, MAP3K21, BDH2, SLC5A3, PANK3, ESCO2, PGLS, MARK3, TRMU, CDY1B, PCYT2, CHD9, SLC38A4, PTG1, VRK3, AQP5, RNF144B, CYP7A1, SLC25A27, PDIA5, SLC7A11, CYP2A7, CES1, SHT3A, SYN1, TRMT1, STK33, FKBP10, PNKP, MGAT4D, IL4I1, SLC16A8, HSD17B8, HELZ2, FOSL1, ALG10B, NDUFAF5, PLXNB3, MRPL58, ATP8A1, PDPK1, BCAP31, TGM5, PPP1R3C, PLA2G5, POLR3GL, TXNRD2, ALDH1A3, SPHK1, MRM1, PPP1R2, USP3, FHIT, IMPAD1, CPT1A, SIN3B, AP000812.4, MAPK3, MAP2K6, NME6, NDUF4, PDE6A, ACAT2, COX10, PFKP, ALDH9A1, FXYD2, CTH, MTHFS, GSS, SEPHS1, METTL16, SLC38A1, POLE4, ATP5PD, PPP1R1A, SLC6A1, CYP27B1, MTRR, CAD, SERPINA3, NGLY1, PEAK1, ADCY1, RPK3, GALNT18, CES2, SLC43A1, LPO, CA12, TNFRSF21, ANXA3, PPT1, UBLCP1, SLC35B4, ACADVL, SGMS2, SLC22A3, CDK7, PTPRD, ATP6V1E2, HKDC1, NDUFC2, USP6, SH3GLB1, SLC01A2, UGGT1, TYMS, PHLPP1, SLC2A13, LYZL1, PHLPP2, RING1, SLC25A28, GLT6D1, MYLIP, LPCAT4, KMT2C, CTSC, SLC8A3, RFWD3, NDUFS2, GMPR, BCO1, STAMB, PTPN22, SEPHS2, IRAK2, MPC1, FUT9, RNF103, AQP6, PPIAL4A, USP17L29, CPA1, NUP210L, SLC10A2, PPT2, NDUFS6, CHFR, MYBPP1, PLPPR3, PRMT9, FADS3, RENBP, PAFAH1B1, PLCH1, MRM1, SLC03A1, ALG5, GAD2, EYA1, EEF1A1, BRSK2, SLC27A5, DUSP19, MDH1, NEK9, GMDS, AL160269.1, G6PD, SLC12A3, SLC25A14, MRI1, CHST10, SHPRH, GBA, PLA2G12A, CDK14, TGFBR1, HSPG2, SMYD2, NAA15, CYP4V2, SGPP1, CARNS1, NSUN4, TMEM54, USP21, CLYBL, TPSD1, PLCD4, SLC15A2, NME5, AMY2A, HSD17B11, MYLK2, CNDP2, SPR, SLC7A8, CTDP1, UBE2D1, COASY, PPP2R3A, IMPA1, ETFB, ATP2A2, PRODH, STAR6, INPP5K, POLR1E, PDE1B1, GUCY1A2, PYCR2, MAP3K7, SLC29A2, BMP2K, SULT1B1, DCT, TTL10, DUSP14, S6GALNAC3, GSTA3, TPMT, RNMT, MDH1, NUP85, SLC29A1, PDE6G, HERC1, ETFB, PTPRZ1, CYP11A1, MET, POR, NAA11, GPAA1, ACAT1, IRAK2, PTPN4, PTPN5, IDH3G, HELZ2, SLC7A9, QTRT2, CBLL2, USP13, FKB6, USP38, PASK, PTPRB, ST8S1A1, WNK4, NDST3, CYP4F3, ENTPD1, MGAT5, APOC3, LCT, RPEL1, PADI4, PPM1B, NME2, TPSB2, PADI4, PFKM, ZNRD1, SLC2A3, CLOCK, HRASLS2, USP13, STK24, ACSM2A, IL18BP, SULT1E1, SLC1A6, CHST14, AKT2, POLD2, HDLBP, UGDH, NEK4, MDM2, SIRT2, SARDH, GMPPA, PRKG2, USP10

Information paradox



Information paradox



Often it seems that the more information we have,
the less we can learn from it.

Using techniques such as clustering and enrichment analysis, we can package the information into bite-sized (human-friendly) pieces.

Gene set analysis (GSA)



- Identifies patterns associated with the genes of interest
- Gene sets are defined based on shared properties, functions, interactions, etc. of the genes



Gene Ontology

Biological Process

Actin crosslink formation

mRNA transport

Carbohydrate transport

Regulation of catabolic process

Erythrocyte development

...

Molecular Function

Glucosidase activity

Alpha actinin binding

Cytokine activity

Oxidized DNA binding

Iron ion binding

...

KEGG Pathways

09100 Metabolism

► 09101 Carbohydrate metabolism

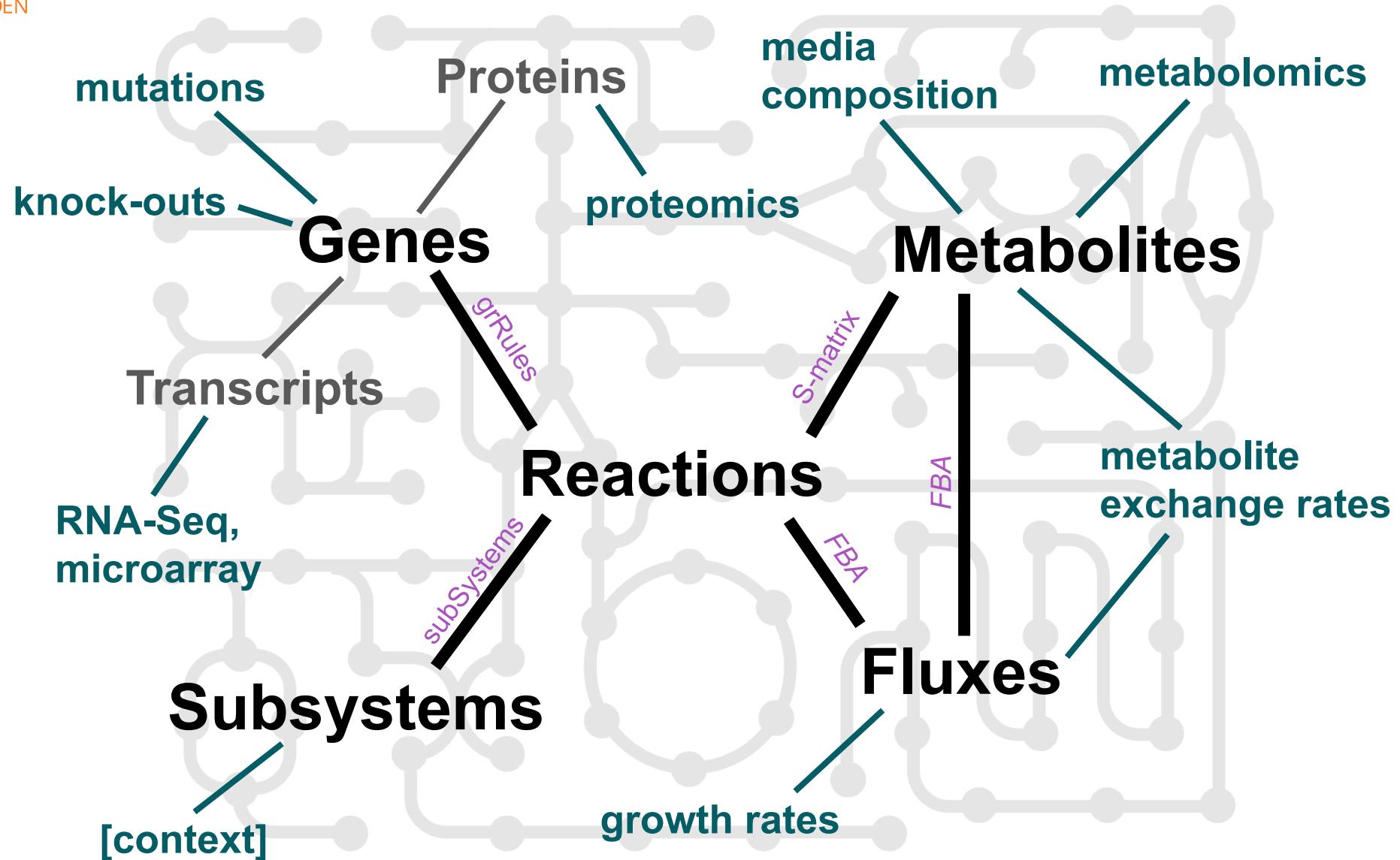
09102 Energy metabolism

► 00190 Oxidative phosphorylation [PATH:[hsa00190](#)]
00195 Photosynthesis
00196 Photosynthesis - antenna proteins
00710 Carbon fixation in photosynthetic organisms
00720 Carbon fixation pathways in prokaryotes
00680 Methane metabolism
► 00910 Nitrogen metabolism [PATH:[hsa00910](#)]
► 00920 Sulfur metabolism [PATH:[hsa00920](#)]

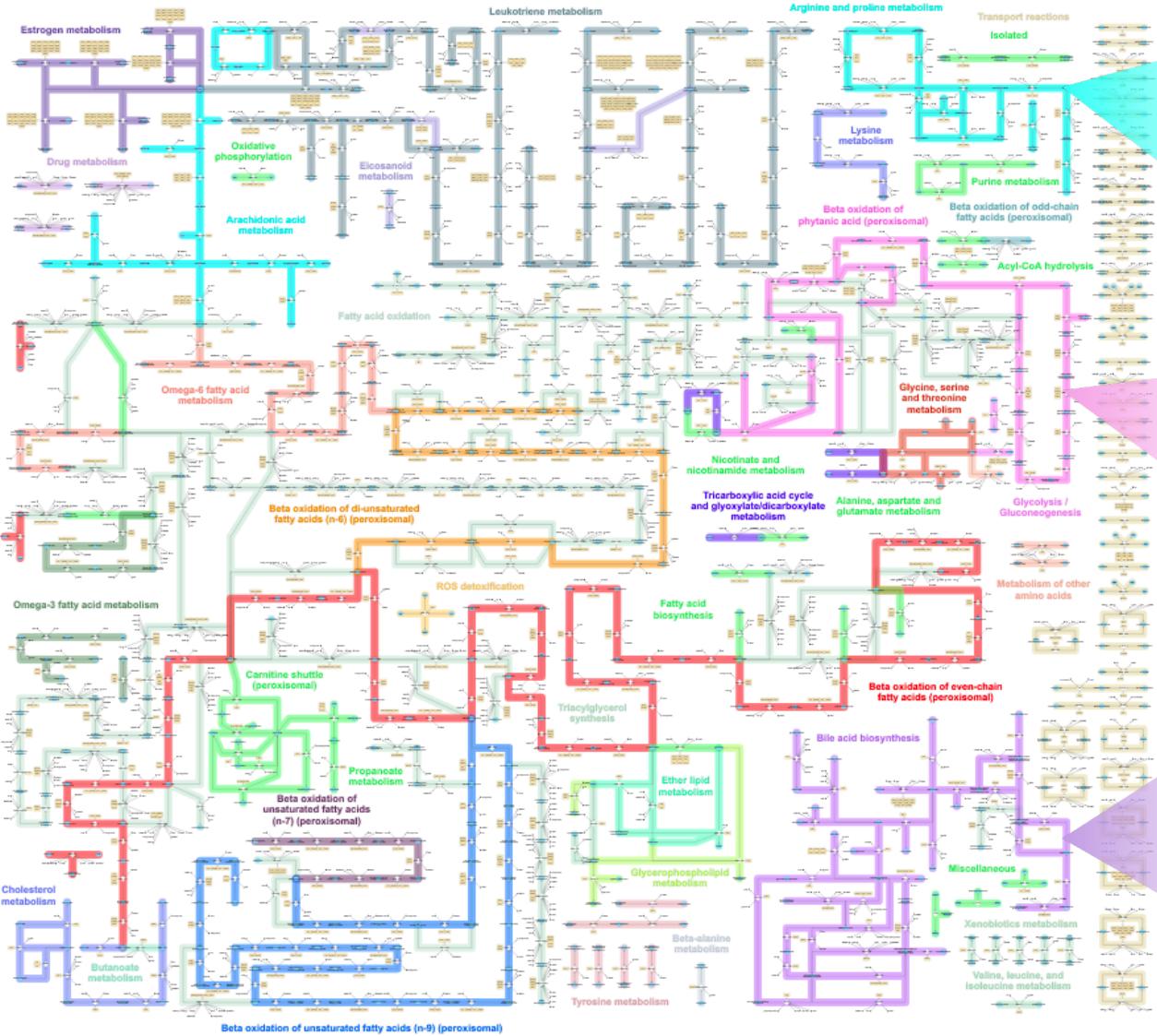
► 09103 Lipid metabolism

► 09104 Nucleotide metabolism

GEM-derived gene sets



Subsystem gene sets



Arginine and proline metabolism

ABHD14A-ACY1	ACY1	AGMAT	ALDH18A1	ALDH1B1	ALDH2	ALDH3A2	ALDH4A1
ALDH7A1	ALDH8A1	ALDH9A1	AMD1	AOC1	AOC2	AOC3	ARG1
CA5A	CA5B	CARNS1	CKB	CKM	CKMT1A	CKMT1B	CKMT2
DHPS	FAR1	FAR2	GAMT	GOT1	GOT2	HOGA1	LEFTY1
						MAOA	MAOB
						MTAP	

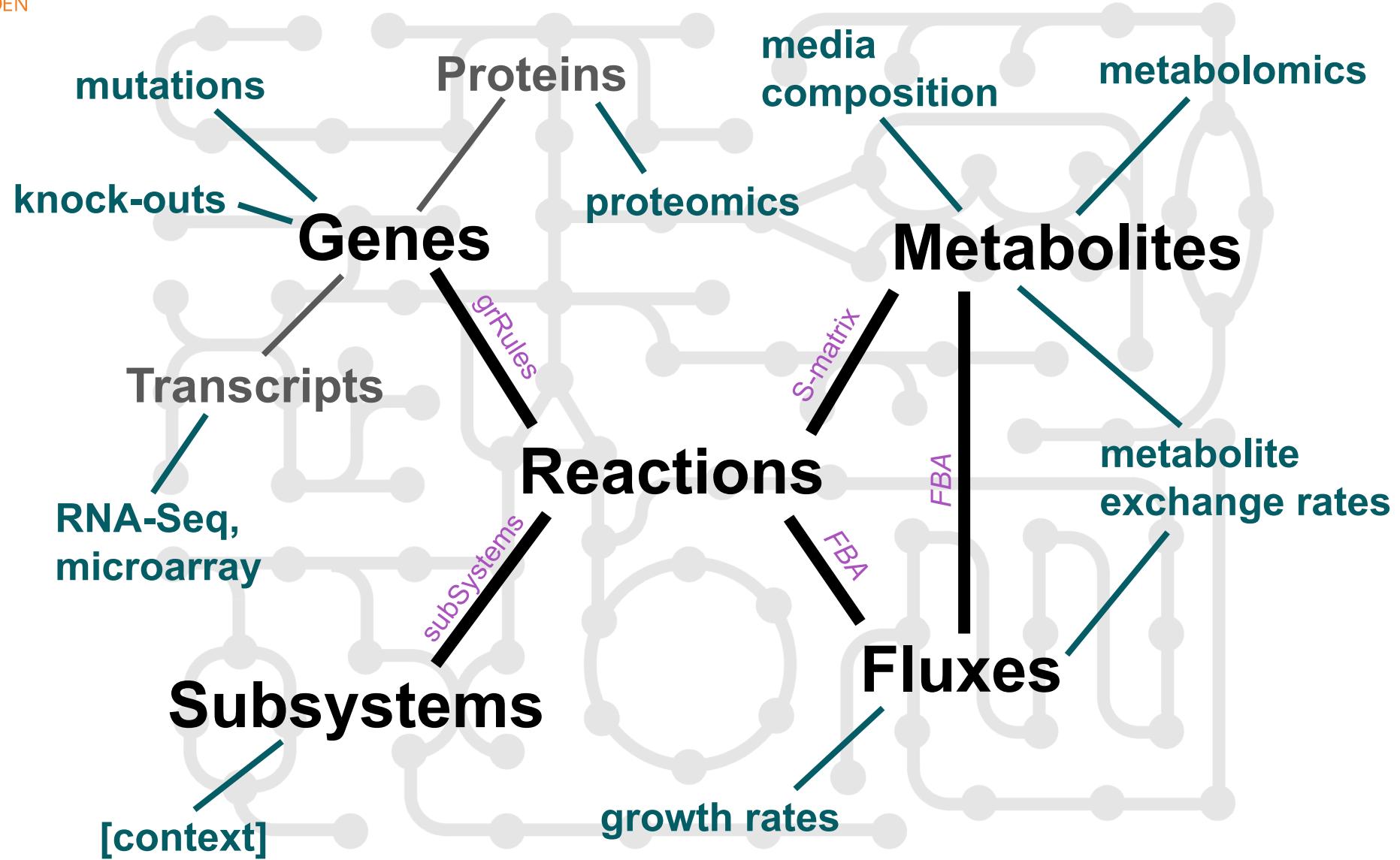
Beta oxidation of phytanic acid

ACAA1	ACOT2	ACOT4	ACOX1	ACOX3	ACSBG1	ACSBG2	ACSL1	ACSL3	ACSL4
ACSL5	ACSL6	AMACR	ECI1	ECI2	EHHADH	HAACL	HADHA	HSD17B4	KRTAP11-1
MEIKIN	MYO5B	PHYH	SLC27A2						

Bile acid biosynthesis

ABCB11	ABCC11	ABCC3	ABCD1	ACAA1	ACAA2	ACOT1	ACOT2	ACOT4	ACOT6
ACOT7	ACOT8	ACOX1	ACOX2	ACOX3	ADH1A	ADH1B	ADH1C	ADH4	ADH5
ADH6	ADH7	ADHFE1	ADO	AKR1B10	AKR1B15	AKR1C1	AKR1C2	AKR1C3	AKR1C4
AKR1D1	ALDH1B1	ALDH2	ALDH3A1	ALDH3A2	ALDH7A1	ALDH9A1	AMACR	BAAT	
BCAP31									

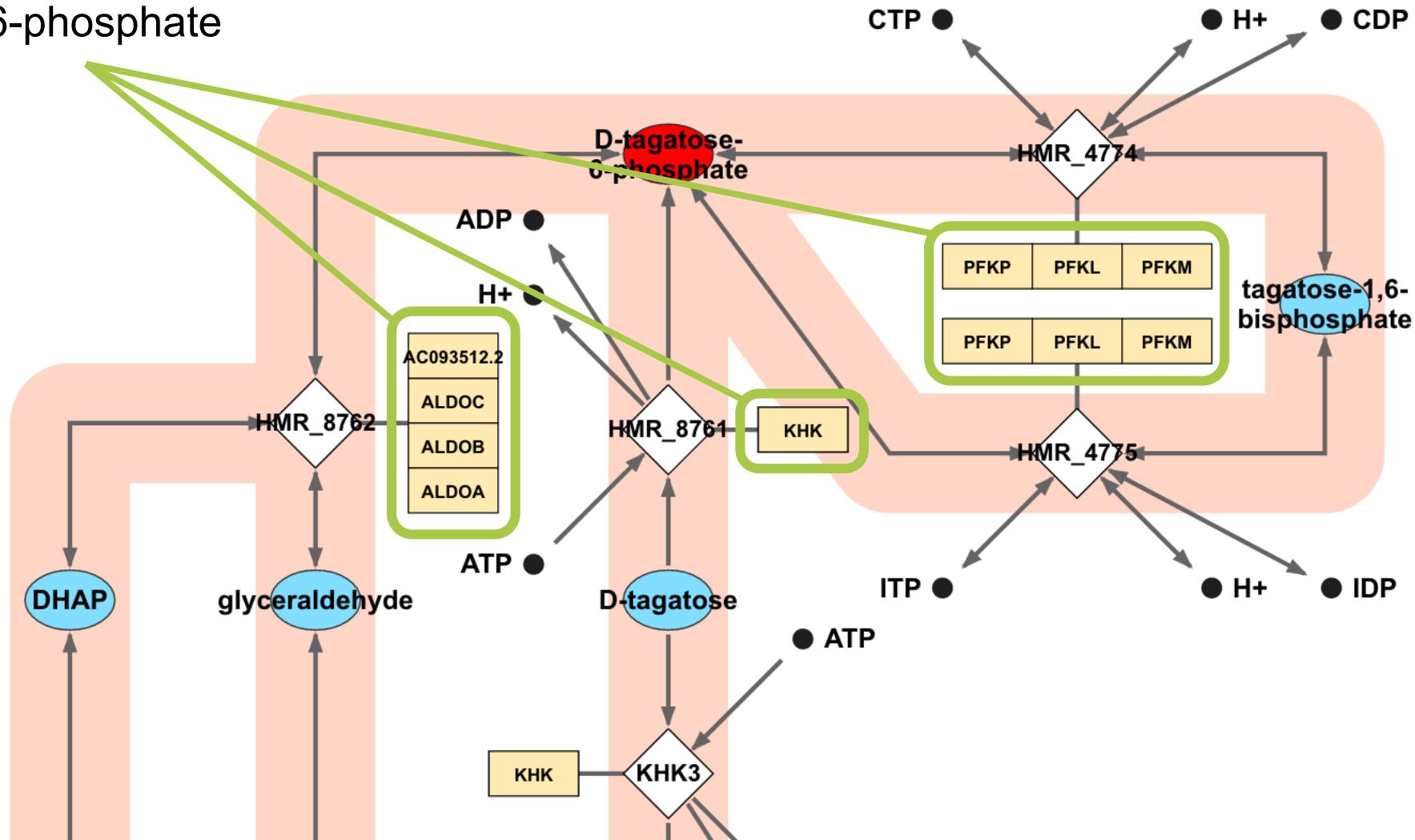
GEM-derived gene sets





Reporter metabolites

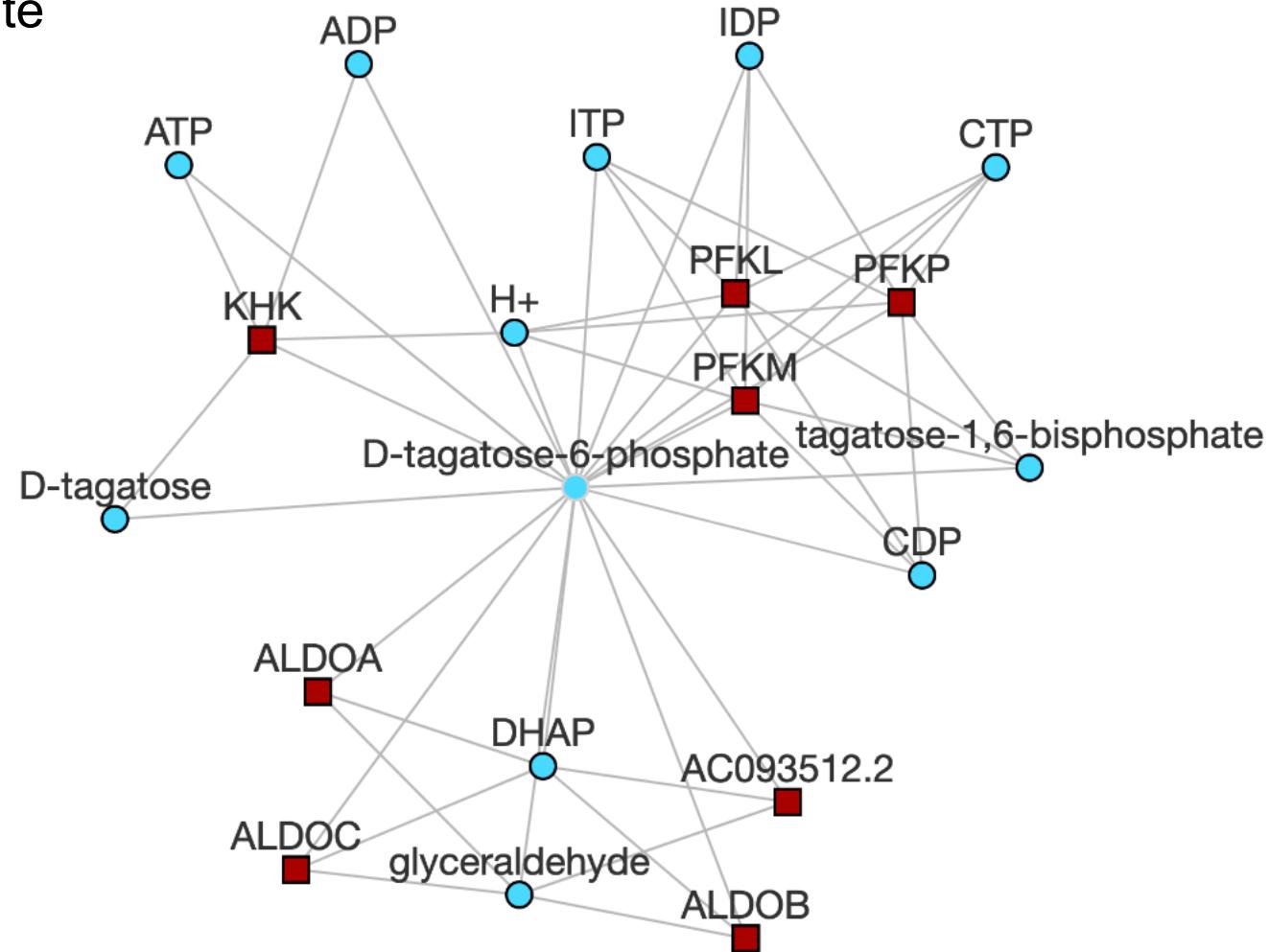
Genes associated with
D-tagatose-6-phosphate



Reporter metabolites



D-tagatose-6-phosphate
interaction partners



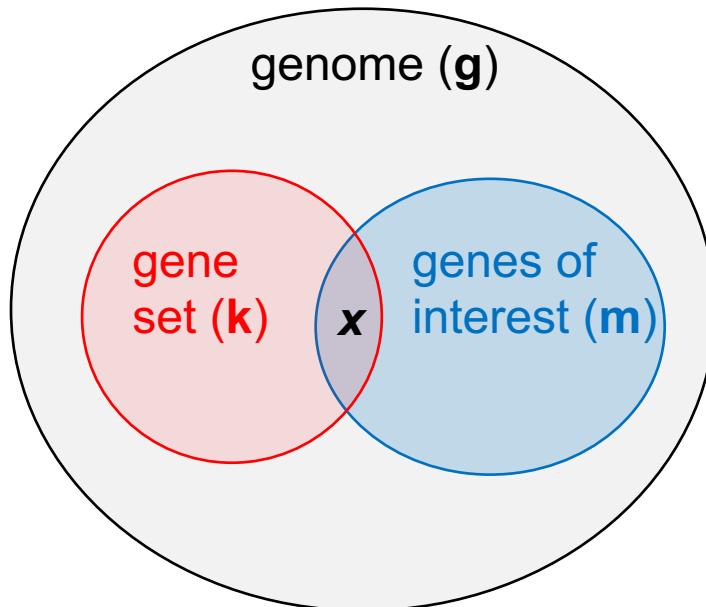
Gene list enrichment



Enrichment or over-representation analysis

Given a list of m genes of interest out of g in the genome and a gene-set of k genes, a statistical enrichment returns the probability that x out of the m genes of interest are in the gene-set.

This is calculated using Fisher's Exact Test (hypergeometric test):



$$p = \frac{\binom{k}{x} \binom{g - k}{m - x}}{\binom{g}{m}}$$

$$\text{note: } \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Gene list enrichment



[Login](#) | [Register](#)

28,099,440 lists analyzed

335,689 terms

168 libraries

Analyze

What's new?

Libraries

Gene search

Term search

About

Help

Input data

Choose an input file to upload. Either in BED format or a list of genes.

Try an example [BED file](#).

[Browse...](#)

No file selected.

Paste a list of valid Entrez gene symbols on each row in the text-box below. [Try a gene set example](#).

```
ALDH3B1
EEF1A1
METTL16
UCKL1
UGT1A4
BCAT2
UGT1A9
UCKL1
HYI
PRODH2
ASNS
GOT1
```

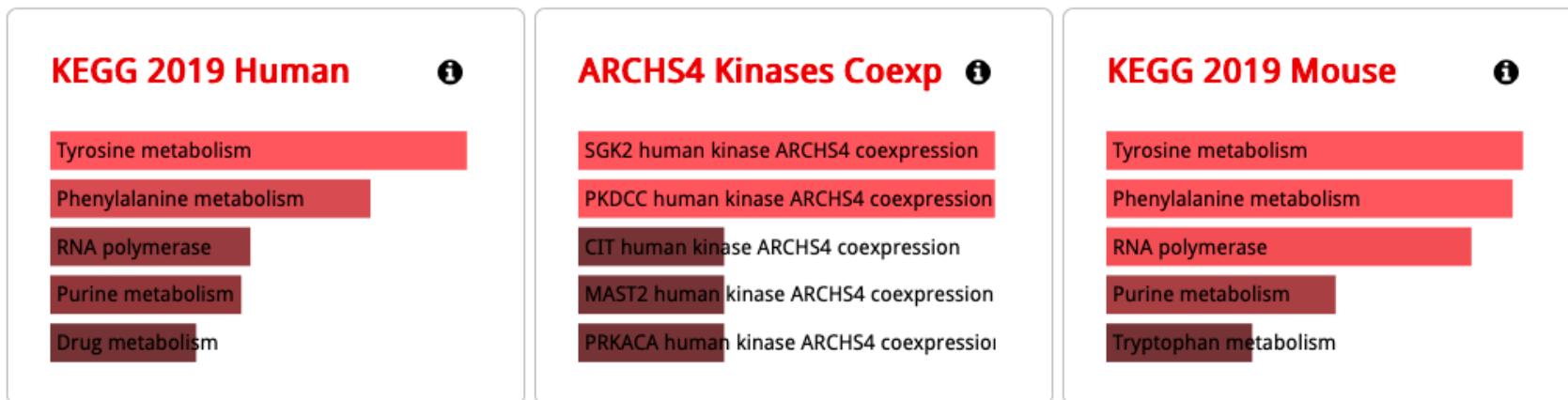
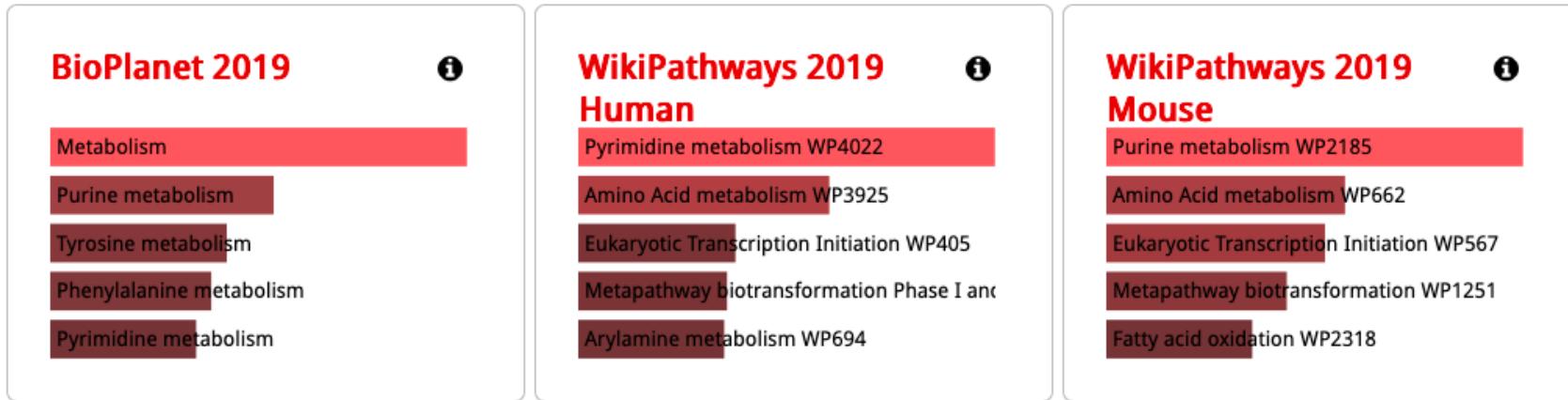
100 gene(s) entered

In order to enable others to search your list please enter a brief description of it.

Contribute your list so it can be searched by others

[Submit](#)

Gene list enrichment

[Login](#) | [Register](#)[Transcription](#) **Pathways** [Ontologies](#) [Diseases/Drugs](#) [Cell Types](#) [Misc](#) [Legacy](#) [Crowd](#)**Description** No description available (89 genes)

Gene list enrichment



Limitations

- Requires arbitrary cutoff to define gene list
- Does not correct for gene-gene correlations (false positives)
- No ranking or relative scoring of genes (gene at the top of the list is identical to bottom)

Gene set analysis



List of genes

ENOPH1
SLC25A2
GMPPB
SLC1A4
EGFL8
HDC

Includes **only** the
genes of interest

Gene-level statistics

0.01	A4GALT
0.89	A4GNT
0.51	AAAS
0.02	AACS
0.33	AADAC
0.08	AADAT
...	...

Includes **ALL** measured/
detected genes

Types of statistics:

- Differential expression p-value
- Differential expression fold-change
- Coefficient or significance of correlation (with phenotype)
- Rank

Gene set analysis

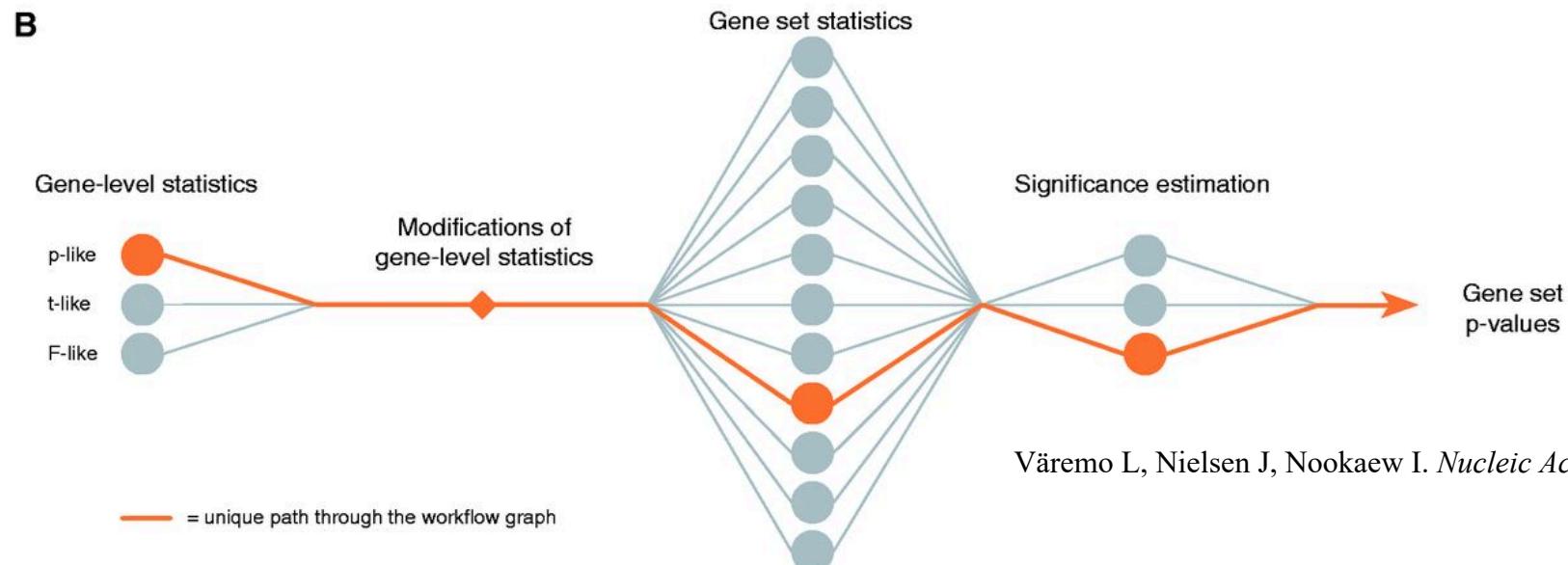
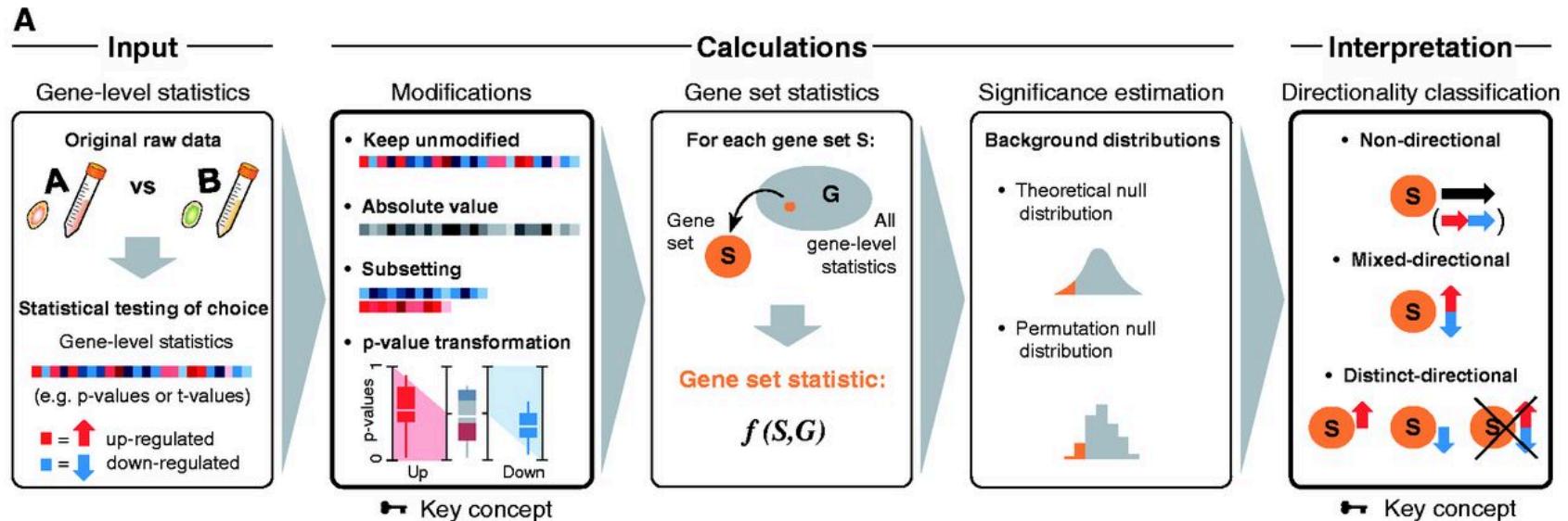


General GSA procedure

1. **Score each gene set** based on the statistics of the genes it contains
2. **Evaluate the significance** of each gene set score based on the score of the null or “background” score distribution

There are *many* methods for both steps 1 and 2

GSA Tools: Piano (R)



Väremo L, Nielsen J, Nookaew I. *Nucleic Acids Res.* 2013;41(8):4378-4391.

GSA Tools: Piano (R)



Gene-level statistics (DE results)

Gene	log2FC	p-value
ENOPH1	-2.4	0.0003
SLC25A2	1.1	0.09
GMPPB	0.3	0.8
SLC1A4	-0.9	0.2
EGFL8	-1.8	0.04
HDC	-6.2	0.0001
A4GALT	3.1	0.0002
...

GSA Tools: Piano (R)



For each gene set, we can calculate 5 different p-values:

Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

Gene set

- Gene 1
- Gene 2
- Gene 3
- Gene 4
- Gene 5
- Gene 6
- Gene 7
- Gene 8

Significantly increased expression



negligible change

Significantly decreased expression

GSA Tools: Piano (R)

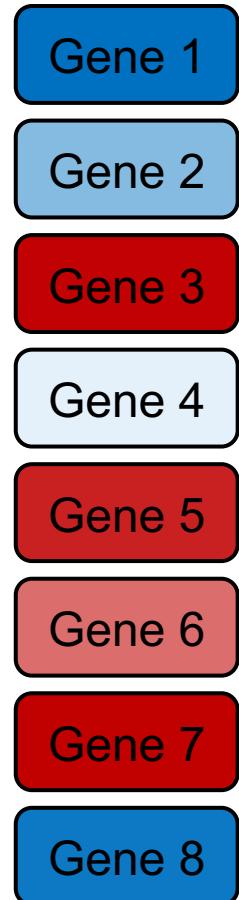


For each gene set, we can calculate 5 different p-values:

Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

Gene set



Significantly increased expression



negligible change

Significantly decreased expression

GSA Tools: Piano (R)



For each gene set, we can calculate 5 different p-values:

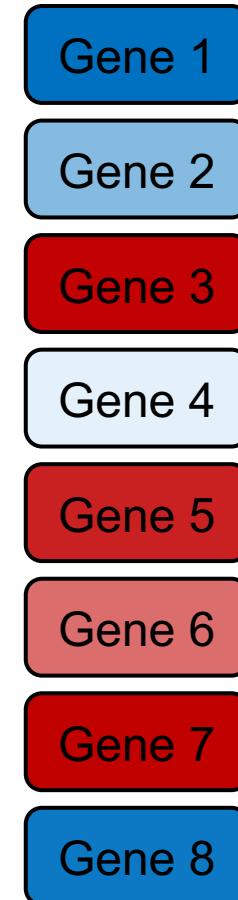
Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

Mixed-directional (down and up):

Test if a *subset* of the gene set is enriched in significantly increased or decreased genes

Gene set



Significantly
increased
expression



negligible
change

Significantly
decreased
expression

GSA Tools: Piano (R)



For each gene set, we can calculate 5 different p-values:

Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

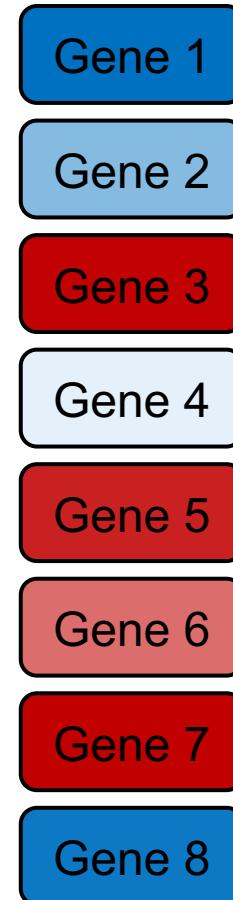
Mixed-directional (down and up):

Test if a *subset* of the gene set is enriched in significantly increased or decreased genes

Distinct-directional (down and up):

Test if the gene set is enriched in significantly increased or decreased genes

Gene set



Significantly increased expression



negligible change

Significantly decreased expression

GSA Tools: Piano (R)



For each gene set, we can calculate 5 different p-values:

✖ Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

✖ Mixed-directional (down and up):

Test if a *subset* of the gene set is enriched in significantly increased or decreased genes

✖ Distinct-directional (down and up):

Test if the gene set is enriched in significantly increased or decreased genes

Gene set

Gene 1
Gene 2
Gene 3
Gene 4
Gene 5
Gene 6
Gene 7
Gene 8

Significantly increased expression



negligible change

Significantly decreased expression

GSA Tools: Piano (R)



For each gene set, we can calculate 5 different p-values:

⑤ Non-directional:

Test for enrichment of significant (low p-value) genes, ignoring fold-change direction.

⑥ Mixed-directional (down and up):

Test if a *subset* of the gene set is enriched in significantly increased or decreased genes

⑦ Distinct-directional (down and up):

Test if the gene set is enriched in significantly increased or decreased genes

Gene set

Gene 1
Gene 2
Gene 3
Gene 4
Gene 5
Gene 6
Gene 7
Gene 8

Significantly
increased
expression



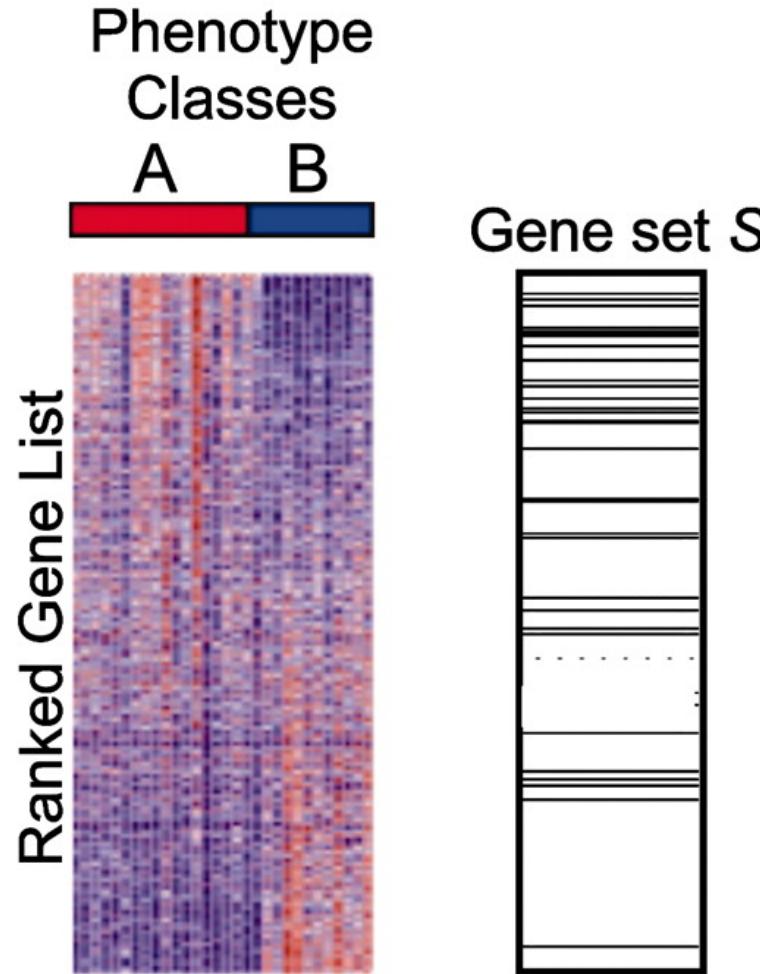
negligible
change

Significantly
decreased
expression

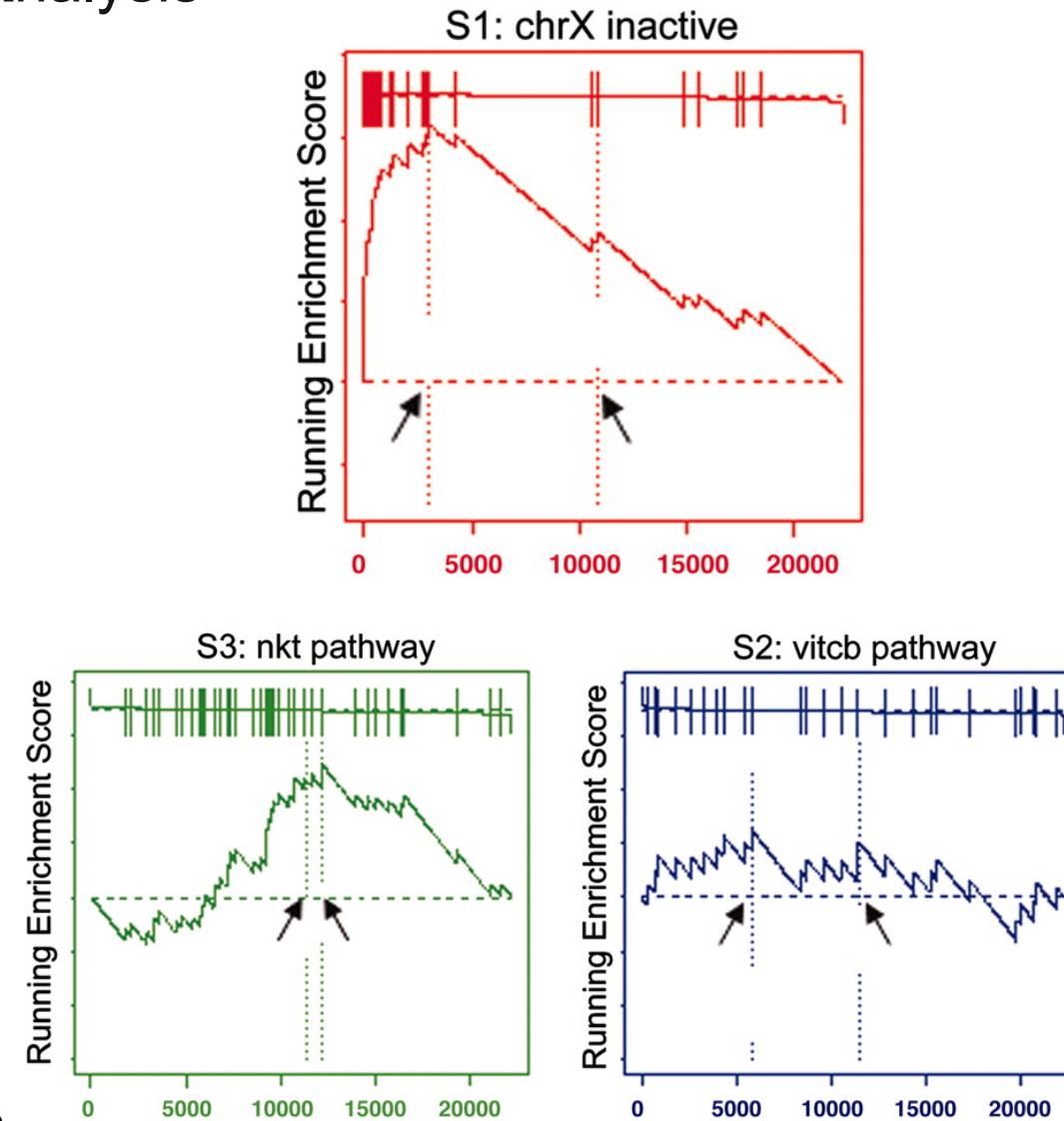


GSA Tools: GSEA (R, python)

Gene Set Enrichment Analysis



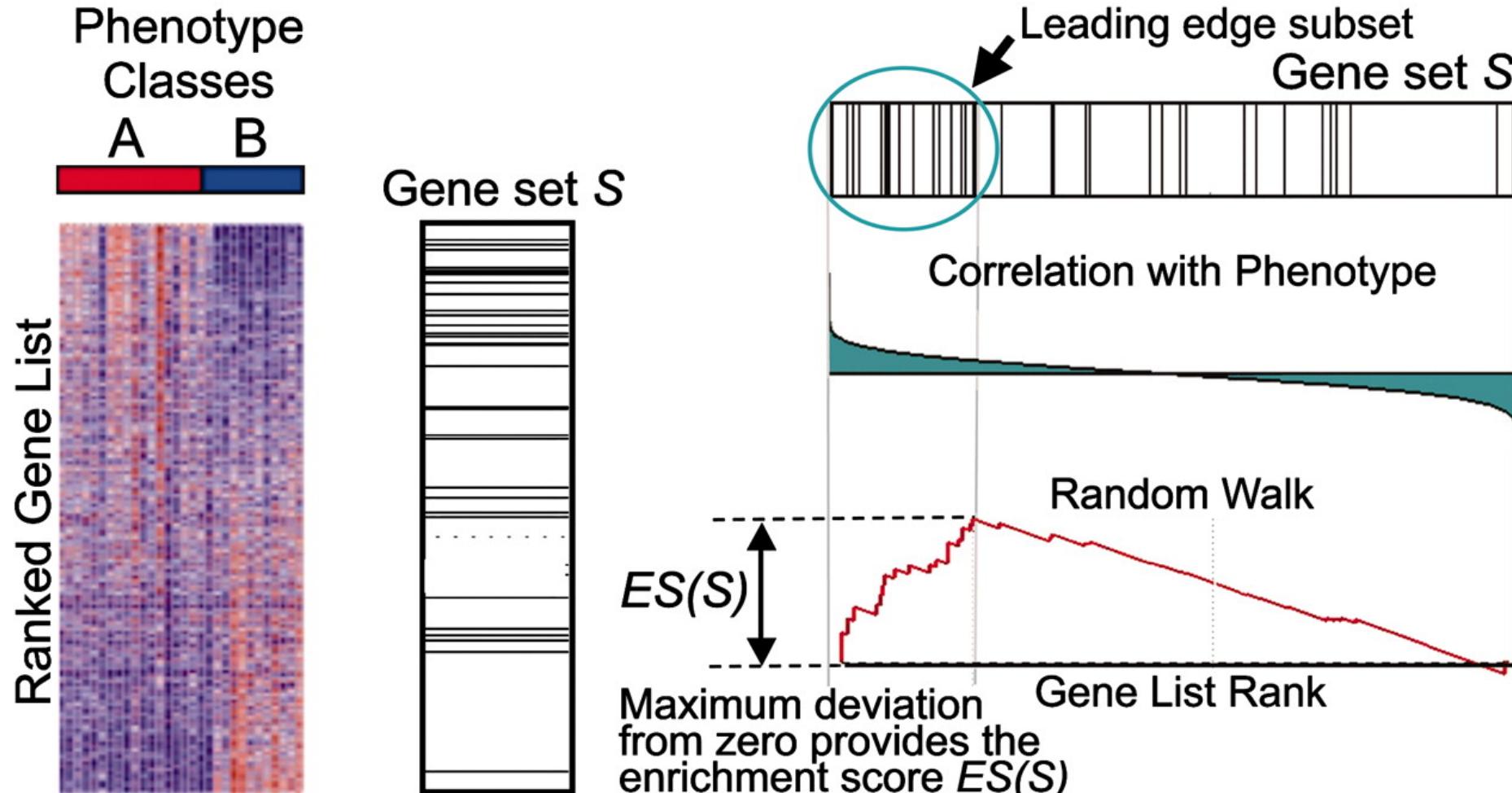
Subramanian, A, et al. PNAS 2005, 102 (43) 15545-15550





GSA Tools: GSEA (R, python)

Gene Set Enrichment Analysis



Context-specific GEMs



A GEM contains all metabolic reactions that are known to occur within an organism

When working with multicellular organisms (e.g., humans), the **“generic” GEM containing all reactions** is not representative of any real cell or tissue type

We can use omics data to **extract a subset** of the generic GEM that is active in our system of interest

This GEM is called an **“extracted” or context-specific GEM**

Context-specific GEMs



There are many methods to generate context-specific GEMs.

For example:

- **iMAT** (Integrative Metabolic Analysis Tool)
- **MBA** (Model Building Algorithm)
- **mCADRE** (metabolic Context-specificity Assessed by Deterministic Reaction Evaluation)
- **tINIT** (Task-driven Integrative Network Inference for Tissues)
- **FASTCORE**

Unfortunately, they were all implemented in MATLAB.

Here are some links to tutorials to using some of the methods:

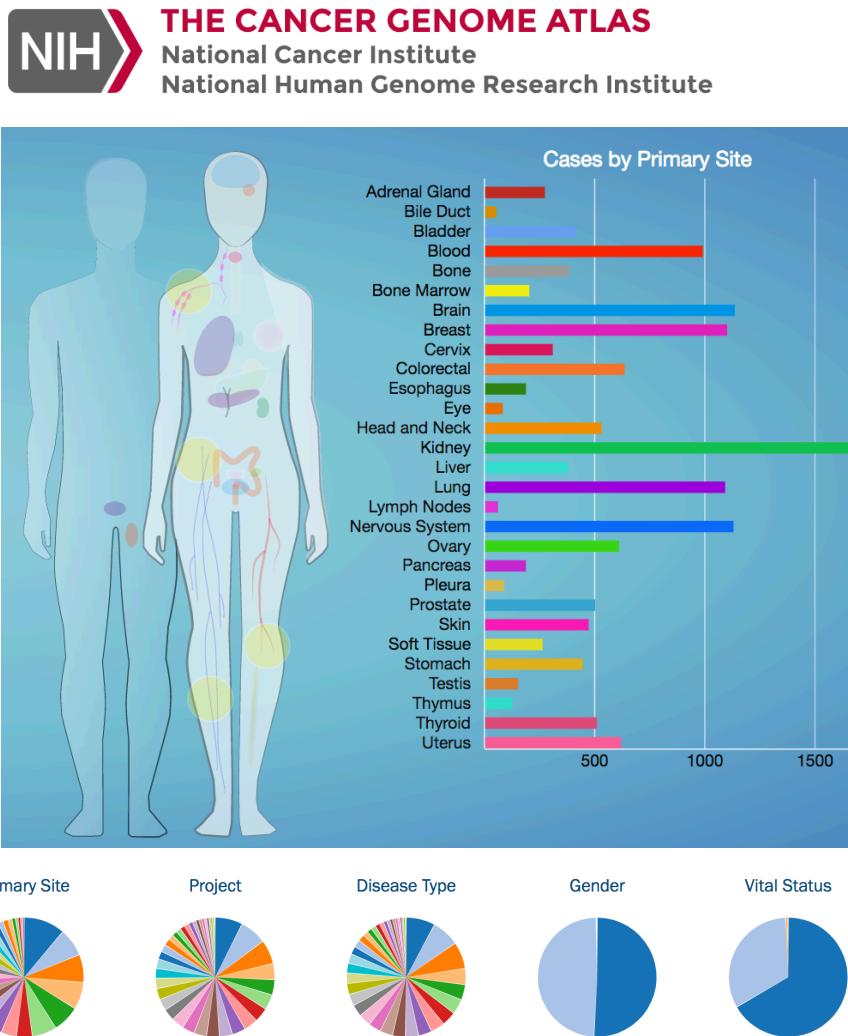
tINIT: https://sysbiochalmers.github.io/Human-GEM-guide/gem_extraction/

iMAT: <https://opencobra.github.io/cobratoolbox/stable/tutorials/tutorialExtractionTranscriptomic.html>

GEM-based comparison of transcriptomes



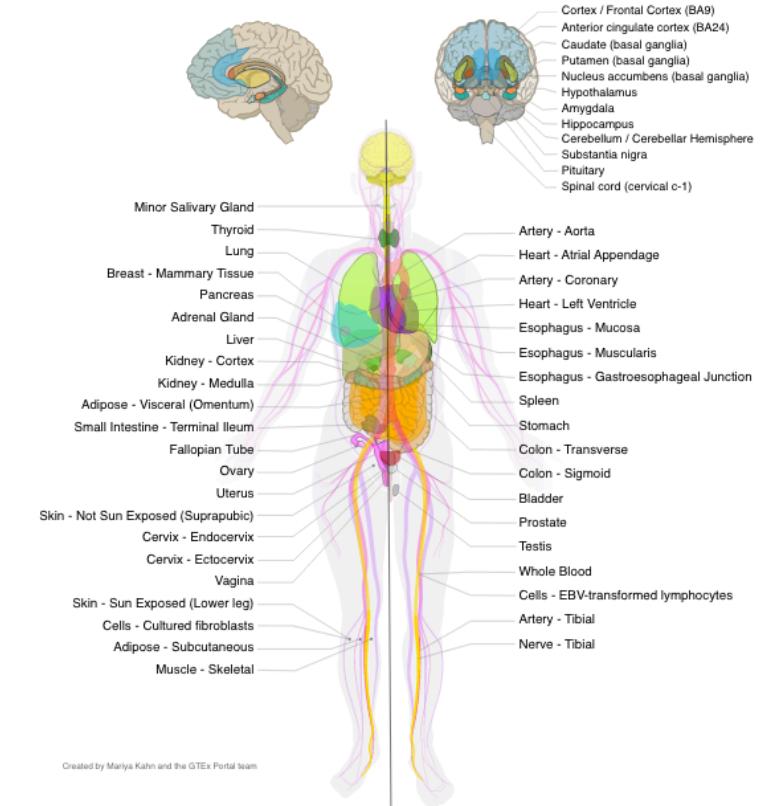
From the study Robinson, et al. An atlas of human metabolism. *Science Signaling* 2020



Tissue Sampling Sites

This page provides a visual representation of the biospecimen source sites (BSSs) for the collection of tissue from postmortem/organ procurement cases for the Genotype-Tissue Expression (GTEx) project.

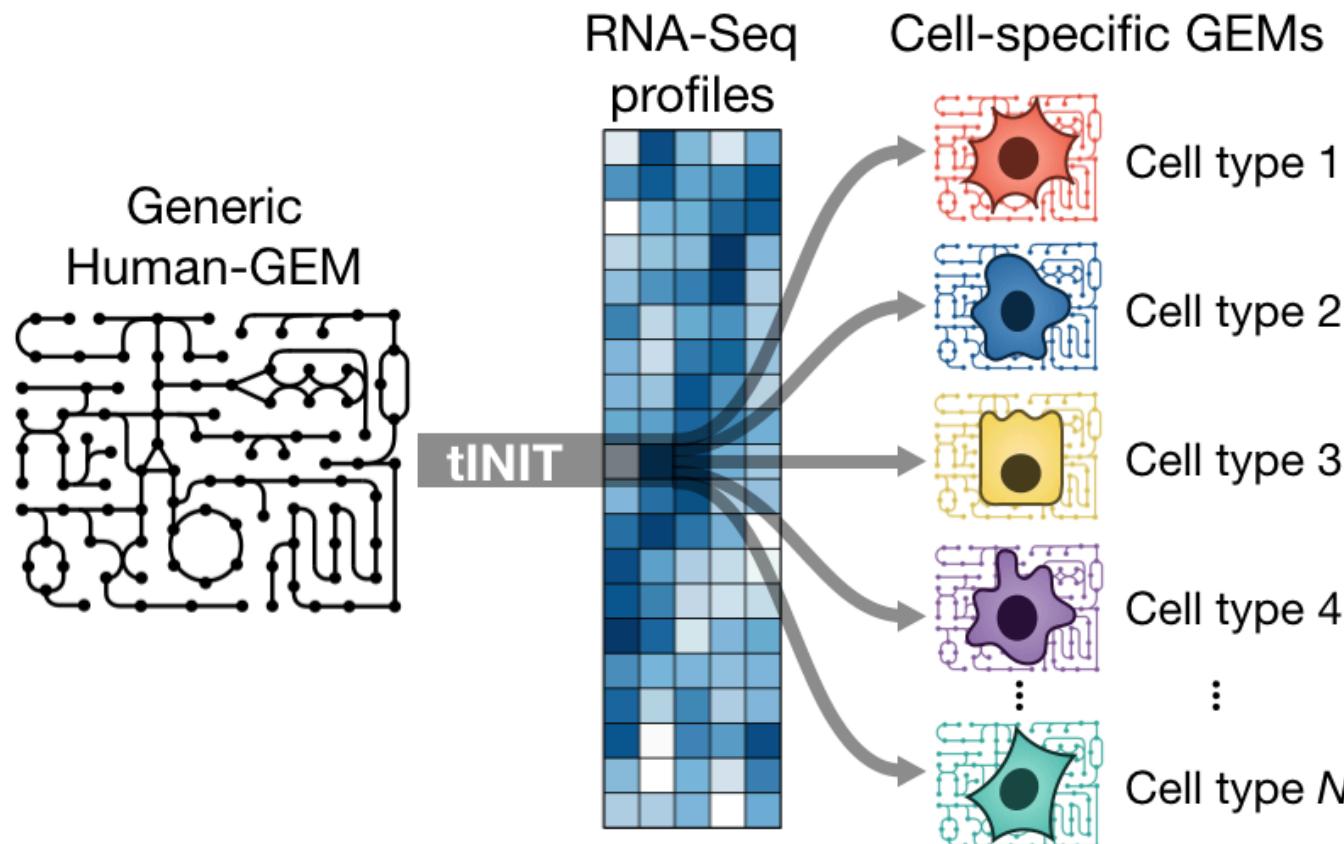
The full documentation on tissue collection procedures can be found on the [GTEx Tissue Harvesting Work Instruction](#).



GEM-based comparison of transcriptomes



Context-specific GEMs were extracted for each of the cancer types and healthy tissue types





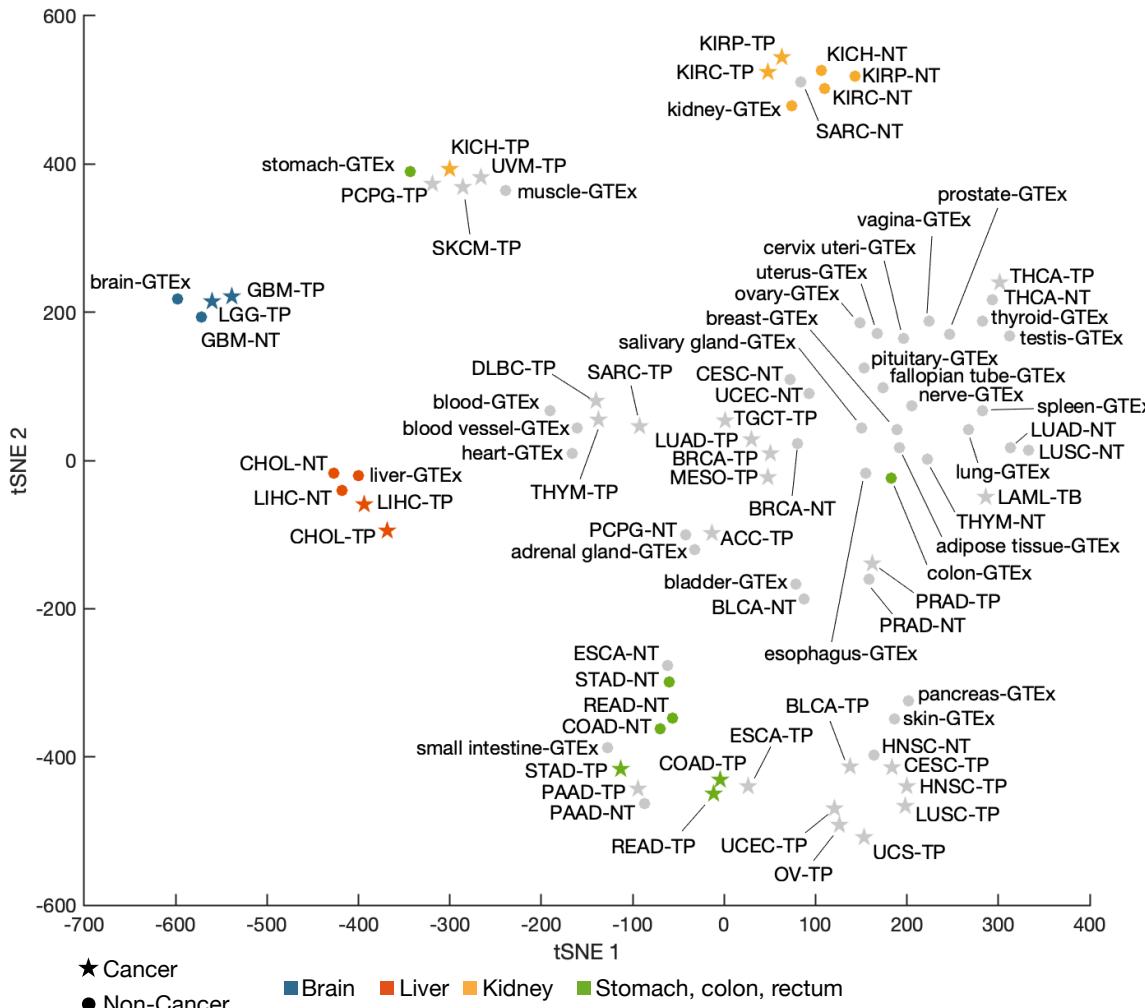
GEM-based comparison of transcriptomes

GEM structure (reaction content) can be represented by a binary vector

Reaction	Lung Tumor	Lung Paired	Lung Healthy	Brain Tumor	Brain Paired	...
rxn1	1	0	1	1	1	 Model contains reaction
rxn2	0	1	1	1	1	
rxn3	0	0	0	0	0	
rxn4	0	1	0	1	0	 Model missing reaction
rxn5	1	1	0	1	1	
rxn6	1	0	0	1	0	
rxn7	0	0	1	1	0	
⋮			⋮⋮			

GEM-based comparison of transcriptomes

Distance (Hamming) between each GEM reaction content vector can be calculated and projected in a tSNE embedding



GEM-based comparison of transcriptomes



If reaction subsystem labels are included, we can look at subsystem-specific differences between GEMs

Subsystem	Reaction	Lung Tumor	Lung Paired	Lung Healthy	Brain Tumor	Brain Paired	...
TCA cycle	rxn1	1	0	1	1	1	
TCA cycle	rxn2	0	1	1	1	1	
Glycolysis	rxn3	0	0	0	0	0	
TCA cycle	rxn4	0	1	0	1	0	
Fatty acid oxidation	rxn5	1	1	0	1	1	
Carnitine shuttle	rxn6	1	0	0	1	0	
Glycolysis	rxn7	0	0	1	1	0	

:

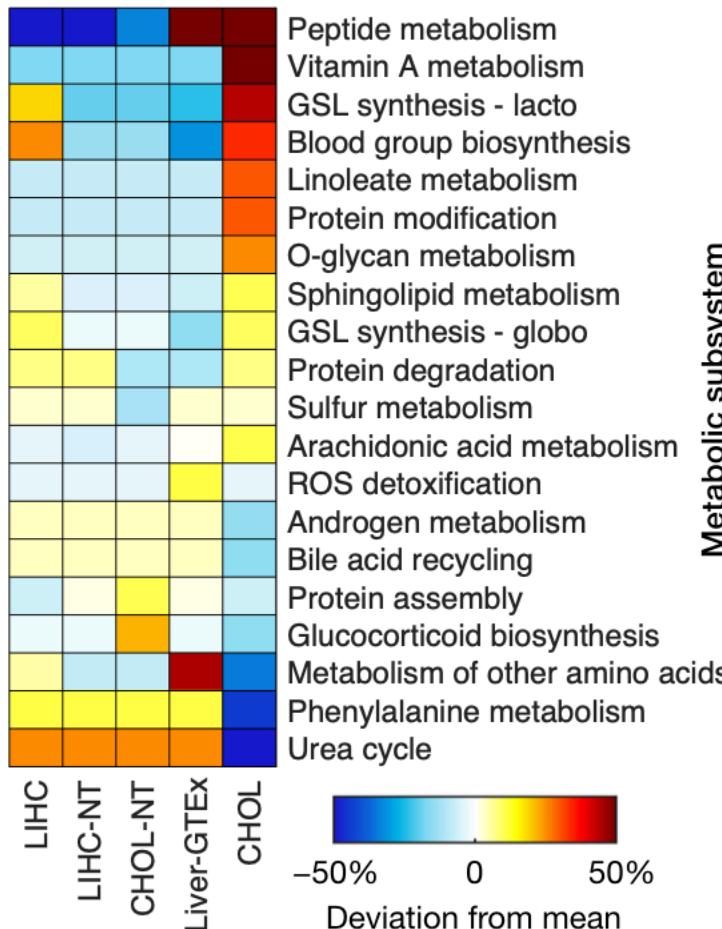
..

GEM-based comparison of transcriptomes



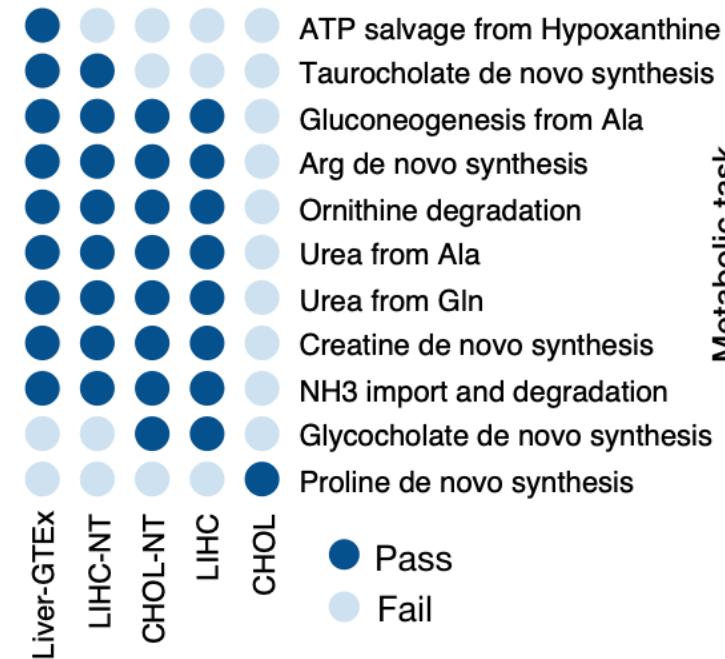
If reaction subsystem labels are included, we can look at subsystem-specific differences between GEMs

Subsystem coverage: Liver



Furthermore, FBA can be used to determine what metabolic functions the GEMs can or cannot perform

Functional comparison: Liver



Exercise: GEM-based GSA



Exercise part 1: (python, short)

Extract metabolite and subsystem gene sets from Human-GEM

Exercise part 2: (R)

Use the GEM-derived gene set collections to evaluate enrichment of differentially expressed genes in different regions of the metabolic network.