Depression: 190 Nodes corresponding to drawings on desktop

Number of Connected Components

('P(X>x) = ', 0.678)

The probability that you will get a randomly selected group of nodes with more connected components than depression is 68%

('The Average number of CCs is:', 127)

('The Median number of CCs is:', 128.0)

**Largest Connected Component**

**('P(X>x) = ', 0.363)**

The probability that the largest connected component will be as big as the largest connected component for depression is 36%.

('The Average biggest CC is:', 31)

('The Median largest CC is:', 29.0)

Number of CC's Excluding Single Nodes

('P(X>x) = ', 0.478)

MDD: 109 Nodes

Number of Connected Components

**('P(X>x) = ', 0.095)**

The probability that you will get a randomly selected group of nodes with more connected components than Major Depressive Disorder is 1%

Implication: there are lots of separate causes that are not linked

Number of CC's Excluding Single Nodes

('P(X>x) = ', 0.897)

Largest Connected Component

('P(X>x) = ', 0.309)

The probability that the largest connected component will be at least as big as the largest connected component for Major Depressive Disorder is 31%.

SHIZO: 157 Nodes

Number of Connected Components

('P(X>x) = ', 0.285)

The probability that you will get a randomly selected group of nodes with more connected components than Schizophrenia is 28%

Largest Connected Component

('P(X>x) = ', 0.653)

The probability that the largest connected component will be as big as the largest connected component for schizophrenia is 65%.

Bipolar Disorder:

Number of Connected Components

('P(X>x) = ', 0.814)

('The Average number of CCs is:', 109)

('The Median number of CCs is:', 110.0)

Largest Connected Component

('P(X>x) = ', 0.452)

('The Average biggest CC is:', 19)

('The Median largest CC is:', 16.0)

MDD Largest Connected Component: 11 Nodes

'ANXA1': This gene encodes a membrane-localized protein that binds phospholipids. This protein inhibits phospholipase A2 and has anti-inflammatory activity. Expression in esophagus, bone marrow, placenta, lung, fat, gallbladder, thyroid

'PLCB1’: The protein encoded by this gene catalyzes the formation of inositol 1,4,5-trisphosphate and diacylglycerol from phosphatidylinositol 4,5-bisphosphate. This reaction uses calcium as a cofactor and plays an important role in the intracellular transduction of many extracellular signals. This gene is activated by two G-protein alpha subunits, alpha-q and alpha-11. Broad expression in the brain and 17 other tissues.

'GRM7': L-glutamate is the major excitatory neurotransmitter in the central nervous system, and it activates both ionotropic and metabotropic glutamate receptors. Glutamatergic neurotransmission is involved in most aspects of normal brain function and can be perturbed in many neuropathologic conditions. The metabotropic glutamate receptors are a family of G protein-coupled receptors that have been divided into three groups on the basis of sequence homology, putative signal transduction mechanisms, and pharmacologic properties. Group I includes GRM1 and GRM5, and these receptors have been shown to activate phospholipase C. Group II includes GRM2 and GRM3, while Group III includes GRM4, GRM6, GRM7 and GRM8. Group II and III receptors are linked to the inhibition of the cyclic AMP cascade but differ in their agonist selectivities. Multiple transcript variants encoding different isoforms have been found for this gene. Expression in brain.

'NCAM1': This gene encodes a cell adhesion protein which is a member of the immunoglobulin superfamily. The encoded protein is involved in cell-to-cell interactions as well as cell-matrix interactions during development and differentiation. The encoded protein has been shown to be involved in development of the nervous system, and for cells involved in the expansion of T cells and dendritic cells which play an important role in immune surveillance. Biased expression in brain, adrenal, and 6 other tissues.

'DGKG': This gene encodes an enzyme that is a member of the type I subfamily of diacylglycerol kinases, which are involved in lipid metabolism. These enzymes generate phosphatidic acid by catalyzing the phosphorylation of diacylglycerol, a fundamental lipid second messenger that activates numerous proteins, including protein kinase C isoforms, Ras guanyl nucleotide-releasing proteins and some transient receptor potential channels. Diacylglycerol kinase gamma has been implicated in cell cycle regulation and in the negative regulation of macrophage differentiation in leukemia cells. Multiple transcript variants encoding different isoforms have been found for this gene. Broad expression in the brain.

'CACNA1C': This gene encodes an alpha-1 subunit of a voltage-dependent calcium channel. Calcium channels mediate the influx of calcium ions into the cell upon membrane polarization. The alpha-1 subunit consists of 24 transmembrane segments and forms the pore through which ions pass into the cell. The calcium channel consists of a complex of alpha-1, alpha-2/delta, beta, and gamma subunits in a 1:1:1:1 ratio. There are multiple isoforms of each of these proteins, either encoded by different genes or the result of alternative splicing of transcripts. The protein encoded by this gene binds to and is inhibited by dihydropyridine. Alternative splicing results in many transcript variants encoding different proteins. Some of the predicted proteins may not produce functional ion channel subunits. Some expression in brain, although strongest expression in endometrium and heart, then brain.

'PPP2R2B': The product of this gene belongs to the phosphatase 2 regulatory subunit B family. Protein phosphatase 2 is one of the four major Ser/Thr phosphatases, and it is implicated in the negative control of cell growth and division. It consists of a common heteromeric core enzyme, which is composed of a catalytic subunit and a constant regulatory subunit, that associates with a variety of regulatory subunits. The B regulatory subunit might modulate substrate selectivity and catalytic activity. This gene encodes a beta isoform of the regulatory subunit B55 subfamily. Defects in this gene cause autosomal dominant spinocerebellar ataxia 12 (SCA12), a disease caused by degeneration of the cerebellum, sometimes involving the brainstem and spinal cord, and in resulting in poor coordination of speech and body movements. Multiple alternatively spliced variants, which encode different isoforms, have been identified for this gene. Broad expression in the brain.

'IMPAD1': This gene encodes a member of the inositol monophosphatase family. The encoded protein is localized to the Golgi apparatus and catalyzes the hydrolysis of phosphoadenosine phosphate (PAP) to adenosine monophosphate (AMP). Mutations in this gene are a cause of GRAPP type chondrodysplasia with joint dislocations, and a pseudogene of this gene is located on the long arm of chromosome 1. Broad expression in fat and adrenal, but also thyroid and brain.

'CACNB2': This gene encodes a subunit of a voltage-dependent calcium channel protein that is a member of the voltage-gated calcium channel superfamily. The gene product was originally identified as an antigen target in Lambert-Eaton myasthenic syndrome, an autoimmune disorder. Mutations in this gene are associated with Brugada syndrome. Alternatively spliced variants encoding different isoforms have been described. Ubiquitous expression in brain (RPKM 4.0), adrenal (RPKM 2.4)

'CACNA2D1': The preproprotein encoded by this gene is cleaved into multiple chains that comprise the alpha-2 and delta subunits of the voltage-dependent calcium channel complex. Calcium channels mediate the influx of calcium ions into the cell upon membrane polarization. Mutations in this gene can cause cardiac deficiencies, including Brugada syndrome and short QT syndrome. Alternate splicing results in multiple transcript variants, some of which may lack the delta subunit portion. Broadly expressed in brain and heart.

'GNAI3’: Guanine nucleotide-binding proteins (G proteins) are involved as modulators or transducers in various transmembrane signaling pathways. G proteins are composed of 3 units: alpha, beta and gamma. This gene encodes an alpha subunit and belongs to the G-alpha family. Mutation in this gene, resulting in a gly40-to-arg substitution, is associated with auriculocondylar syndrome, and shown to affect downstream targets in the G protein-coupled endothelin receptor pathway. Ubiquitous expression in esophagus (RPKM 27.0), bone marrow and 25 other tissues, brain included.

All 11 from MDD are present in this largest connected component:

'DSEL': expression in brain

u'DMD',: This gene spans a genomic range of greater than 2 Mb and encodes a large protein containing an N-terminal actin-binding domain and multiple spectrin repeats. The encoded protein forms a component of the dystrophin-glycoprotein complex (DGC), which bridges the inner cytoskeleton and the extracellular matrix. Deletions, duplications, and point mutations at this gene locus may cause Duchenne muscular dystrophy (DMD), Becker muscular dystrophy (BMD), or cardiomyopathy. Alternative promoter usage and alternative splicing result in numerous distinct transcript variants and protein isoforms for this gene. Expression in heart and fat (and lesser in brain)

u'LAMA1', expression in testis, thyroid, a little in brain

u'ITGA11', Biased expression in endometrium (RPKM 12.1), urinary bladder (RPKM 2.7) and 13 other tissues

35:

**'IMPAD1',**

u'RASAL2', esophagus, fat, brain lesser

**'NCAM1',**

u'NCAN', Restricted expression toward brain

u'CACNB3', ovary, brain

**'CACNB2',**

**'CACNA1C',**

**'GNAI3**',

u'DGKH', This gene encodes a member of the diacylglycerol kinase (DGK) enzyme family. Members of this family are involved in regulating intracellular concentrations of diacylglycerol and phosphatidic acid. **Variation in this gene has been associated with bipolar disorder.**

u'GRM8', brain, testis, other tissues

u'NRG1', thyroid, urinary bladder

u'FGF12', heart and brain

u'NRG3', biased expression in brain

**'GRM7',**

u'DLC1', lung and fat

**'DGKG',**

u'ERBB2', kidney, skin, not so much brain compared to all 24 other tissues

**'PLCB1'**,

**'PPP2R2B'**,

u'PIK3C2A', thyroid, testis

u'TNC', appendix

**'CACNA2D1',**

u'KCNMB2', ovary, adrenal, brain a little yes

u'RGS10', lymph node, appendix, brain a tiny bit

u'ADCY2', brain, prostate

u'SHC4', brain,

u'SST', some brain

u'KIT', thyroid, ovary

u'PTPRG'fat, kidney, some brain

**'ANXA1',**

u'OXTR, ovary, endometrium,some brain

3/3

gwas\_Alzheimer.tsv

3 [u'ITGA8', u'PTK2', u'LPAR3']

gwas\_autism.tsv

3 [u'MMP16', u'ACTB', u'TAF1C']

gwas\_bipolar\_disorder.tsv

3 [u'MKL1', u'ACTG1', u'COL4A2']

gwas\_breast\_cancer.tsv

3 [u'PLCB1', u'MAPK14', u'CTNNA2']

gwas\_crohn.tsv

3 [u'NFATC1', u'CALM1', u'ADCY7']

gwas\_Dementia.tsv

3 [u'MMP16', u'ACTB', u'SF3B1']

gwas\_depression.tsv

3 [u'DMD', u'AGRN', u'APOB']

gwas\_major\_depressive\_disorder.tsv

3 [u'PLCB1', u'MAPK14', u'DSCAM']

gwas\_multiple\_sclerosis.tsv

3 [u'STAT3', u'B3GAT3', u'GPC5']

gwas\_schizophrenia.tsv

3 [u'CACNA1C', u'PPP1CC', u'CENPM']

gwas\_waist\_hip.tsv

3 [u'CEBPA', u'USF2', u'HOXC13']

ulcerative\_colitis.tsv

3 [u'CIITA', u'PAX5', u'TNFRSF6B']

vogelstein.txt

3 [u'AKT1', u'IKBKG', u'CYLD']

Depression:

DONE

3 [u'DMD', u'AGRN', u'APOB']

true

DMD  :  0.000004

node not found:  AGRN – 69 edges between this and others

true

APOB  :  0.000001

AGRN: This gene encodes one of several proteins that are critical in the development of the neuromuscular junction (NMJ), as identified in mouse knock-out studies. The encoded protein contains several laminin G, Kazal type serine protease inhibitor, and epidermal growth factor domains. Additional post-translational modifications occur to add glycosaminoglycans and disulfide bonds. In one family with congenital myasthenic syndrome affecting limb-girdle muscles, a mutation in this gene was found. Alternative splicing results in multiple transcript variants encoding different isoforms.

Longest distance = 15 nodes

Second component longest distance = 5

Farthest two genes from each other: DMD and NRG3, NRG1, PRPRG, RASAL2

One node in between.

Bipolar:

3 [u'MKL1', u'ACTG1', u'COL4A2']

true

MKL1  :  0.000004

node not found:  ACTG1 🡪 291 edges between this node and others

true

COL4A2  :  0.000004

Longest path = 7

Longest second comp path = 5

MLK1 is not central

Actins are highly conserved proteins that are involved in various types of cell motility and in maintenance of the cytoskeleton. Three main groups of actin isoforms have been identified in vertebrate animals: alpha, beta, and gamma. The alpha actins are found in muscle tissues and are a major constituent of the contractile apparatus. The beta and gamma actins co-exist in most cell types as components of the cytoskeleton and as mediators of internal cell motility. Actin gamma 1, encoded by this gene, is a cytoplasmic actin found in all cell types. Mutations in this gene are associated with DFNA20/26, a subtype of autosomal dominant non-syndromic sensorineural progressive hearing loss and also with Baraitser-Winter syndrome. Alternative splicing results in multiple transcript variants. [provided by RefSeq, Jul 2017]

Schizophrenia:

3 [u'CACNA1C', u'PPP1CC', u'CENPM']

true

CACNA1C  :  1E-7

node not found:  PPP1CC 🡪 232 edges between this node and others

true

CENPM  :  3E-8

Longest path: 5

Second component longest path is also 5

CACNA1C is fairly central in the component

The protein encoded by this gene belongs to the protein phosphatase family, PP1 subfamily. PP1 is an ubiquitous serine/threonine phosphatase that regulates many cellular processes, including cell division. It is expressed in mammalian cells as three closely related isoforms, alpha, beta/delta and gamma, which have distinct localization patterns. This gene encodes the gamma isozyme. Alternatively spliced transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Oct 2011]

Depression: Clustering coefficient of individual components

0.179954648526

0.0

DONE

Schizophrenia

0.15641025641

0.4

DONE

Bipolar Disorder:

0.0352941176471

0.474236874237

DONE

VS. Crohn:

0.319114060871

0.0

and Cancer

0.446324843459

0.0

('Average Shortest Path over 1000 trials: ', 3.897)

('Average Diameter of Component 1: ', 6.386)

('Average Diameter of Component 2: ', 3.061)

('Average CC of Component 1: ', 0.24964478689717104)

('Average CC of Component 2: ', 0.4305834467259811)