

Entropia global de rede na expressão gênica no intuito de caracterizar a diferenciação celular em tecidos periféricos de transtornos mentais.

Giovanni Pereira de Andrade, José Cláudio Fonseca Moreira

Departamento de Bioquímica da Universidade Federal do Rio Grande do Sul

Resumo

O diagnóstico complexo e as diferentes etiologias dos transtornos mentais fomenta a busca por marcadores descritivos que simplifiquem a compreensão e a distinção destas patologias. Nesta direção, as pesquisas em amostras de tecido periférico acoplado a técnicas de visualização do perfil proteômico, transcriptômico ou genômico é crescente. O presente estudo teve como base os inúmeros resultados destas pesquisas, ao utilizar bancos de dados de transcriptoma e de interações entre genes e seus produtos, na busca por marcas de mudança no estado de diferenciação celular no sangue periférico em transtornos do espectro autista, transtorno bipolar, esquizofrenia e transtorno depressivo maior. Os resultados apontam variações em subgrupos de esquizofrenia, e avançam no entendimento destes transtornos.

1. Introdução

Por definição, de acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-IV, 1994), o transtorno mental é categorizado como síndrome ou padrão comportamental ou psicológico de importância clínica, sendo correlacionado a sofrimento, incapacitação ou ainda a risco elevado de morte, instabilidade emocional, deficiência ou dano significativo da liberdade emocional. Nesse sentido, diversas doenças são classificadas como transtorno mental, tais como esquizofrenia, desordens de depressão maior, transtornos do espectro autista, entre outros.

No estudo de transtornos mentais, diferentes variáveis já foram correlacionadas, como fatores psicossociais ou hereditários, alterações no sistema imune e desregulações de neurotransmissores. Muitos fatores foram estabelecidos por expressão gênica diferencial, como é o caso de depressão maior e transtornos do espectro autista associados à desregulação do sistema imune em células de sangue periférico (Leday et al, 2018; Gregg et al, 2008). Embora a expressão diferencial de muitos genes seja associada a transtornos mentais, pouco se sabe sobre o impacto que a variação de alguns genes provocaria, mudando a informação das vias de sinalização celular ou aumentando a promiscuidade na sinalização.

Para tentar elucidar o papel de possíveis modificações na sinalização celular, antes é necessário estabelecer um método que permita esta análise. Essa metodologia se utiliza dos conceitos da biologia de sistemas e consiste na integração de dados de expressão e de uma rede global de sinalização celular, onde a probabilidade de interação entre os nós é determinada pelo nível de expressão de cada gene, correspondendo a um nó. A abordagem veio na tentativa de quantificar o modelo de paisagem epigenética de Waddington, e possibilitou a comprovação de que células-tronco possuem maior probabilidade de ativar diferentes vias de sinalização celular, enquanto células teciduais diferenciadas teriam menor probabilidade e, conseqüentemente menor incerteza e menor quantidade total de informação. Essa quantificação, denominada como a "entropia global de rede de sinalização celular", foi desenvolvida por Banerji et al. (2013) e descreveu as células tumorais como atreladas à desdiferenciação celular, tendo maior entropia quando comparada a células de tecido saudável. A entropia de rede foi encontrada associada ao grau do tumor e à menor taxa de sobrevivência em câncer de mama, tanto do tipo positivo quanto negativo para receptor de estrogênio (Banerji et al., 2015).

Desse modo, pode-se questionar o impacto da expressão gênica diferencial na sinalização celular, provocando mudança, aumento ou perda de informação. Assim sendo, este trabalho quantificou e comparou a entropia global de rede de sinalização celular no sangue periférico de indivíduos acometidos de transtornos mentais em relação a indivíduos saudáveis.

2. Resultados

Redes globais e seus parâmetros gerais

Duas redes globais de sinalização celular foram construídas para realizar a análise de entropia de rede, sendo oriundas de diferentes bancos de dados, e seguindo critérios distintos quanto à interação dos nós-genes (ver Metodologia). A primeira rede segue a metodologia de Banerji et al. (2013), aqui

chamada de Rede 1, enquanto a segunda rede foi desenvolvida por metodologia própria deste artigo, chamada Rede 2, onde informações sobre topologia de rede foram extraídas de ambas.

Ao passo que o número de nós das duas redes foi relativamente semelhante, os demais parâmetros apontam composições distintas (Tabela 1). A Rede 1 apresenta maior número de arestas e de diâmetro de rede, além de menor coeficiente de clusterização médio. O maior número de arestas na Rede 1 pode ser justificado pela maior acurácia do Reactome, onde os artigos avaliados são curados manualmente e revisados por pares, dificultando interações que não sejam bem validadas na literatura. O diâmetro ligeiramente maior, e a formação média de clusters menor, também podem ser fatores oriundos desta observação, onde vemos o melhor agrupamento de clusters que podem representar vias na sinalização celular, corroborando a hipótese de que o Reactome seria um banco de dados melhor estruturado.

Parâmetros de nó das redes

Parâmetros de propriedade de nós das duas redes geraram os 1000 genes (top1000) de maior valor de grau de nó, sendo os “genes *hub*” (HG - do inglês, *hub genes*), e os top1000 de *betweenness centrality*, sendo os “genes gargalo” (BG - do inglês, *bottlenecks genes*; informações vide Tabelas 1 e 2, Material Suplementar). Após análise de ontologia gênica dos top1000 de cada rede, a intersecção das 100 funções enriquecidas de maior desdobramento, entre HG ou entre BG, das duas redes, demonstraram diferentes funções biológicas atribuídas a estes parâmetros (Figuras 1 e 2).

Dentre as funções enriquecidas aos BG (Figura 1), ganha destaque as atribuídas a processos do sistema imune, resposta inflamatória e resposta a estímulos químicos ou não, que juntas compreendem 56% das funções gerais, ilustrando a importância da resposta celular e do sistema imune e inflamatório em serem intrinsecamente conectados a diversas vias celulares. Outras funções encontradas em BG mostram locomoção, adesão e estrutura celulares como grandes gargalos entre diferentes vias, representando 20% das funções gerais, além de 4 vias responsáveis pelo desenvolvimento de estruturas anatômicas multicelulares, que representa 16% de termos gerais atribuídos (Tabelas 3 e 4, Material Suplementar).

O enriquecimento de HG revela funções relacionadas à morte celular e regulação da via de sinalização do receptor adrenérgico em maior desdobramento nas duas redes (Figura 2), e a análise dos termos gerais ressalta morte e ciclo celular, ambos processos de decisão de destinos celulares, compreendendo juntos mais de 20% de todos os processos gerais (Tabelas 5 e 6, Material Suplementar). Em análise geral, a resposta a estímulos, podendo ser químicos ou estresse, e a sinalização celular de receptores, agrupando assim todas as vias de sinalização extracelular enriquecidas, somaram 25% de todos os processos. O desenvolvimento de estruturas anatômicas multicelulares novamente ganha destaque, em 10% de todos os processos gerais, e entre as funções específicas de maior desdobramento no enriquecimento. É interessante destacar o aparecimento de vias de resposta a IL2 e IL9, grandes maestros na sinalização imune, além de respostas mediadas pelo receptor Fc, de modulação viral e de seleção, diferenciação e apoptose de células B em alto enriquecimento para as duas redes.

Assim sendo, a avaliação dos parâmetros de nós indica a possível predominância da resposta a estímulos dentre os genes de maior impacto na sinalização celular, e o mecanismo de desenvolvimento de estruturas anatômicas multicelulares intimamente conectado e responsivo a outros processos. A análise também sugere a sinalização do sistema imune e a resposta inflamatória na comunicação dentre diversas vias, de acordo com a definição de gargalo, e genes de alta conexão interna na decisão entre multiplicação ou morte celular.

Hematopoiese

No intuito de validar as redes e suas respectivas metodologias de construção, a busca sistemática por amostras de células saudáveis que demarcasse a hematopoiese celular levou aos dados depositados no GEO (Gene Expression Omnibus) em GSE24759, a partir do estudo de Novershtern et al. (2011), onde os seguintes tecidos foram analisados por entropia de rede: células-tronco hematopoiéticas (HSC), células progenitoras mieloides, células progenitoras de granulócitos e monócitos, células precursoras de linfócitos B iniciais (early B-cell) ou não-iniciais (pro B-cell), todas estas coletadas do sangue de cordão umbilical; células dendríticas (DC) mieloides, monócitos, Natural Killer (NK) maduras, células NKT, células B e células T, estas últimas coletadas de sangue periférico (Figura 3).

O processo de diferenciação celular foi demarcado com eficiência na entropia de rede usando a Rede 1, separando amostras de HSC de células precursoras de B em estágio inicial, de células B e de células T. Além disso, em muitos outros grupos há uma tendência a diferenciar HSC de células precursoras de linfócitos B, de células maduras NK, de células progenitoras mielóides e de monócitos, o que talvez não tenha sido validado estatisticamente apenas por falta de número amostral maior.

A entropia celular global obteve menor força estatística para distinguir grupos celulares de etapas da hematopoiese, onde apenas os grupos de células HSC e células precursoras de linfócitos B iniciais demarcaram a maturação celular, e mesmo que seja vista a mesma tendência de separar os grupos discutidos parágrafo acima, é evidente o maior contraste na Rede 1 ao usar a entropia como artifício para demarcar estados da hematopoiese celular.

Transtorno depressivo maior

A complexa etiologia e as diversas subcategorias fenotípicas que compreendem o transtorno depressivo maior (MDD) exigem igual critério na avaliação amostral de pessoas acometidas desta patologia. Assim sendo, os dois dados de expressão gênica de MDD avaliados compreendem apenas amostras de sangue periférico de adultos, e os dados de expressão são de duas plataformas distintas, sendo o primeiro de microarranjo e o segundo de RNA-Seq. No primeiro, as amostras depositadas no GEO em GSE98793 são importantes por conter dois grupos separados de MDD: do tipo ansioso e sem ansiedade. Em ambos os grupos, submetidos às duas análises de rede, não houve variação significativa na entropia de rede frente ao controle (Figura 4). Para corroborar o resultado, confirmou-se a semelhança no nível de entropia de MDD comparado ao grupo controle saudável, no conjunto de dados de GSE39653, tanto na Rede 1 quanto na Rede 2 (Figura 4).

O primeiro conjunto de dados revelou, no estudo de Leday et al. (2018), maior expressão de genes do sistema imune inato e menor expressão de genes do sistema imune adaptativo, demonstrando um mecanismo de possível alteração da sinalização celular sem interferir na quantidade total de informação. Além disso, o segundo conjunto de dados demonstrou, através da pesquisa de Savitz et al. (2013), alteração na expressão de genes codificadores de proteínas moduladoras na resposta inflamatória mediada pelo Fator Nuclear Kappa-B (NF κ B) e Quinase Regulada por sinal Extracelular (ERK), comparando controles saudáveis a “transtornos do humor”, sendo elas MDD e Transtorno Bipolar. Dessa maneira, ambos os estudos confluem para a alteração da sinalização imunológica e inflamatória, sugerindo modificação na sinalização celular do sistema imune que mantenha a quantidade total de informação.

Transtorno Bipolar

Todas as amostras de pessoas acometidas por Transtorno Bipolar (BD) e seus respectivos controles são de sangue periférico de adultos, e compreendem três estudos depositados no EO: GSE39653, GSE46449 e GSE18312. Conforme citado anteriormente, o primeiro estudo é de RNA-Seq, enquanto os outros dois são de plataformas distintas de microarranjo. Dentre os três conjuntos de dados de expressão, gerados sob três distintas plataformas de quantificação de RNA, apenas o conjunto amostral GSE46449 indicou diferença entre BD e controles saudáveis (Figuras 4 e 5). Contudo, essa diferença foi vista apenas na Rede 2 e apenas nestas amostras, em baixa significância estatística (p-valor entre 0,01 e 0,05). Sobretudo, esse resultado pode estar distante do real estado de maturação celular, considerando o maior poder de diferenciação celular ao usar a Rede 1, conforme descrito anteriormente (Hematopoiese, Resultados), ao passo que a Rede 2, isoladamente, apresenta certas deficiências como descritor sólido desse estado celular.

Migração e ativação de leucócitos, sinalização do receptor Fc, regulação e resposta de receptores de superfície de resposta imune, regulação do sistema imune, resposta a patógenos virais e secreção de citocinas estão entre as funções mais enriquecidas (dados não mostrados) dos genes de maior expressão diferencial (*log Fold Change*) das amostras de GSE46449 no estudo de Clelland et al. (2013), ao passo que a via de apresentação de antígeno se mostrou uma das dez mais enriquecidas no grupo BD, no estudo de Bousman et al. (2010) em GSE18312. Esses dois estudos, conjuntamente ao resultado de Savitz et al. (2013), descrito nos parágrafos anteriores, descrevem possíveis alterações na sinalização celular do sistema imune em sangue periférico de pacientes com BD.

Esquizofrenia

Neste transtorno psiquiátrico, três dados de expressão gênica, de amostras de sangue periférico de adultos, foram coletados do GEO. As três amostras foram avaliadas na Rede 2, enquanto apenas as

amostras GSE18312 e GSE27383 foram avaliadas na Rede 1 (dados de entropia sobre a amostra GSE38484 na Rede 1 foram corrompidos, sendo necessária a repetição da análise). A entropia global de rede da expressão gênica em GSE18312 foi consistente em ambas as redes, além de estar de acordo com a análise dos dados de GSE38484 na Rede 2, sem diferença entre o grupo esquizofrenia (SCZ) e o controle (Figuras 5 e 6). Contudo, o subagrupamento de pessoas acometidas da doença revelou diferenças marcantes na entropia global de rede em GSE27383 na Rede 1, entre os grupos de alta gravidade e de remissão de sintomas, tanto entre eles quanto comparados ao grupo controle, embora nenhuma destas distinções apareçam ao usar a Rede 2.

O estudo de van Beveren et al. (2012), sobre as amostras de GSE27383, encontrou desregulação na via Akt1, assim como redução na expressão de AKT1, ao comparar pessoas saudáveis a pacientes que sofriam de esquizofrenia (SCZ), enquanto não houve variação entre os subgrupos de SCZ com antipsicóticos, sem antipsicóticos, com remissão de sintomas e em estado grave. Além disso, no enriquecimento de vias canônicas dos genes diferencialmente expressos, entre as dez vias de maior poder estatístico foram encontradas vias de sinalização de células NK, de extravasamento de leucócitos, de receptor de células T, de integrinas, de Fc Epsilon RI e de receptor de célula B, processos direta ou indiretamente associados ao sistema imunológico. Esses resultados apresentam concordância ao estudo de Bousman et al. (2010), onde a via de sinalização de CD27 de linfócitos consta entre vias de enriquecimento diferencial em SCZ aos dados GSE18312, apontando fortemente a desregulação do sistema imune como ponto crítico na alteração do sangue de SCZ.

A pesquisa conduzida por Jong et al. (2012) nas amostras em GSE38484 avaliou o perfil de coexpressão dos genes de SCZ, onde a rede de conexões mais fortes destacou, entre as quatro vias mais enriquecidas por análise de vias ingenuínas (IPA), desenvolvimento celular, desenvolvimento e função do sistema hematológico, e hematopoiese. Portanto, a investigação de marcas de diferenciação celular em SCZ merece mais análises para validar a existência de alteração nesses mecanismos, já que vias de diferenciação celular hematológica aparecem na expressão diferencial de SCZ.

Transtornos do Espectro Autista

Seguindo critérios de busca sistemática de amostras de sangue periférico de pessoas acometidas de Transtornos do Espectro Autista (ASD), até o momento desta pesquisa foram encontradas apenas amostras de idade limítrofe em crianças e adolescentes, sendo dados adversos ao critério padrão de seleção dos dados de expressão. Ainda assim, dois conjuntos de dados do GEO foram avaliados: GSE6575 e GSE42133. Para ambas as redes, os resultados se confirmam ao mostrar nenhuma variação entre tecido saudável e de ASD ou ainda de retardo mental e deficiências de desenvolvimento (Figura 7).

A análise de vias ingenuínas (IPA), que expande a análise a *upstream* para incluir reguladores não conectados diretamente aos genes diferencialmente expressos, feita por Gregg et al. (2008), mostrou nas amostras de GSE6575 o enriquecimento de vias de sinalização de células NK e de sinalização de IL-2 dentre as quatro vias enriquecidas por IPA em ASD. A pesquisa de Pramparo et al. (2015), sobre as amostras depositadas em GSE42133 e apenas do sangue periférico de garotos, destacou em genes diferencialmente coexpressos o enriquecimento de funções imunológicas e inflamatórias. Mais uma vez, ambos estudos apontam alterações no sistema imune e inflamatório dentre as principais vias modificadas nos transtornos, enquanto a entropia de sinalização global permanece constante.

3. Discussão

As duas redes de interação proteína-proteína (PPI) direcionam a sinalização celular global como altamente sensível, sob quantificação da entropia global, a perturbações em genes de decisão na multiplicação e na morte celular, de locomoção e adesão celular, de sinalização do sistema imune e resposta inflamatória, de resposta a estímulos químicos, moleculares ou estresse, e de desenvolvimento de estruturas anatômicas multicelulares. Contudo, os bancos de dados de PPI são, comumente, construídos de pesquisas feitas por células *in vitro*, onde ocorre a alta deposição de estudos em que as conexões entre proteínas são estudadas em linhagens e tecidos de câncer, e podem não traduzir interações reais dependendo do tecido. Além disso, como os bancos de dados não são tecido-específicos, espera-se que em geral os genes de mais interações descritas pertençam a vias de manutenção e sinalização basal da célula (Parikshak, Gandal & Geschwind, 2015). Assim sendo, considerando o viés de interações em vias de desfecho celular amplamente identificadas em câncer, deve-se questionar a alta conectividade de genes ligados a processos do ciclo celular, de proliferação, de migração e de adesão, pois são processos drasticamente alterados no câncer. Da mesma forma,

processos basais como a resposta inflamatória e a resposta a estímulos externos precisam ser colocados em dúvida quanto à sua grande conectividade dentro da rede PPI. Ainda assim, a grande representatividade na rede PPI contendo genes de resposta e de sistema imunológico e inflamatório, destacados entre as principais vias de BG e HG, poderia estar próxima de PPI reais na sinalização celular de leucócitos.

Os estudos de transtornos mentais apresentados aqui direcionam as alterações quantitativas dos RNAs mensageiros em sangue periférico a genes presentes em vias de resposta imunológica e inflamatória, localizados entre processos de maior integração entre diferentes clusters funcionais na sinalização celular global. Mesmo assim, na grande maioria dos dados, avaliados diante de duas redes de estruturas peculiares, a quantidade total de informação na rede se mostrou a mesma perante os controles saudáveis. Isso leva à hipótese inicial de que as alterações nestes genes poderiam ser insuficientes para modular novas vias e alterar consideravelmente a sinalização celular. A segunda alternativa permite assimilar que ocorreram mudanças na informação de algumas vias específicas, no entanto o somatório da quantidade total de informação celular seria o mesmo, admitindo o mesmo estado de maturação celular sob comportamentos distintos.

Ao assimilar a hipótese de não alteração no estado celular final, é curioso observar que em dados de SCZ tenham sido enriquecidas funções relacionadas à hematopoiese e diferenciação celular (de Jong et al., 2012), embora a entropia global de sinalização não conseguiu demarcar variações no estado de diferenciação celular entre SCZ e controles. Contudo, a separação de grupos em remissão e de alta gravidade obteve resultados alinhados ao descrito, onde há uma marca de desdiferenciação em SCZ de alta gravidade quando comparado a pacientes em remissão, assim como uma marca de maior diferenciação em SCZ em remissão comparados aos controles, nos dados de GSE27383. Traçando um paralelo entre diferentes tecidos, esses resultados são coerentes com outros estudos na literatura, em que há dano na expressão de marcadores relacionados à diferenciação de oligodendrócitos, células fundamentais para a reconstrução da bainha de mielina, danificada na rede neural de SCZ (Mauney et al., 2015; Radu et al., 2011). A alta penetrância da translocação no locus do gene “Disrupted in Schizophrenia” (DISC1) em casos de SCZ merece destaque, pois o gene DISC1 exerce papel fundamental no neurodesenvolvimento, envolvido em diversas cascatas de sinalização de maturação neuronal e ativação sináptica (Porteous et al., 2011), outro exemplo de falha na diferenciação neuronal em SCZ.

A consistência na similaridade da entropia global entre grupos de MDD e controles saudáveis aponta a ausência de marcas de modificação na maturação de leucócitos. Todavia, é preciso ressaltar as duas hipóteses levantadas acima nessa discussão, sendo inadequado afirmar, de imediato, a não alteração da sinalização celular como um todo. Nesse sentido, o aumento da transcrição de genes relacionados à resposta pró-inflamatória de monócitos, em amostras de MDD, ainda que divergindo diante de idade e de início precoce ou tardio da doença (Grosse et al., 2015), fatores não considerados no agrupamento das amostras avaliadas aqui, é interessante por apontar MDD como alvo de variação nas informações celulares sobre inflamação, sem aumento ou perda da quantidade de informação total e, portanto, entropia global de rede invariável. Contudo, isso implicaria em visualizar, nos grupos de MDD com aumento da ativação pró-inflamatória, o mesmo estado de maturação celular, fator questionável no que tange ao desenvolvimento de diversos tipos de macrófagos e de DC por resposta a citocinas pró-inflamatórias (Kumar et al., 2006). Dessa forma, é essencial agrupar de maneira mais rigorosa as amostras de MDD para avaliar a possibilidade de real alteração no perfil de ativação de linfócitos sob a abordagem de entropia global.

Quanto ao transtorno de BD, a hipótese de mudança no perfil de sinalização global é vista apenas na entropia distinta de BD em GSE46449 na Rede 2, resultado não encontrado em outros dados, ou mesmo no próprio dado, na Rede 1. Enfatizando as melhores distinções no processo hematopoiético obtidas a partir da Rede 1, é necessário cautela ao extrair conclusões precipitadas desse resultado. Nesta patologia, DC derivadas de monócitos coletados de BD revelaram alterações em marcadores de diferenciação celular (Knijff et al., 2006), levantando a possibilidade de marcas similares em monócitos de BD, presentes no sangue periférico. Hiperpolarização de neurônios derivados de células B induzidas pluripotentes (iPSC) de pacientes BD também ilustra a possibilidade de replicar modificações neuronais a partir do tecido periférico (Stern et al., 2017). Contudo, essa distinção, caso exista, aparenta estar distante da detecção por entropia de rede.

Para bebês e crianças diagnosticadas com ASD, não foi observada diferença no estado de diferenciação celular no sangue periférico, e um segundo estudo com as amostras de GSE42133 não obteve coesão em vias ou genes que demarcassem, em escala, o grau de dano ou atraso na linguagem

em crianças com ASD, onde as alterações genéticas não se correlacionam (Lombardo et al., 2018). Porém, já foi demonstrado que em ASD existe marca de “aceleração” da diferenciação microglial no córtex frontal e no córtex temporal (Hanamsagar, et al., 2017), assim como maior quantidade de Fator Neurotrófico Derivado do Cérebro (BDNF) no sangue de crianças ASD (Enstrom et al., 2008), fator secretado por ativação de microglias, células neurais semelhantes aos macrófagos (Kumar et al., 2006). Outro resultado da literatura demonstra que células CD45+ (marcador clássico de leucócitos) apresentam, dentre outras alterações, maiores níveis de Fator Estimulador de Colônia de Granulócitos-Macrófagos (GM-CSF) em crianças e adolescentes de ASD (Ahmad et al., 2020), podendo induzir a conclusão de que ASD teria maior diferenciação de monócitos. Dessa forma, a complexidade de fatores citados e os resultados de entropia obtidos salientam a barreira entre o comportamento celular do tecido periférico e as modificações na morfologia de células do sistema nervoso central deste transtorno.

O presente artigo ressalta a dificuldade em encontrar alterações em apenas uma via de sinalização ou um biomarcador no tecido periférico para caracterizar, de maneira direta, cada transtorno, sendo necessário, pelo menos, a seleção de linhagens celulares específicas e/ou a categorização menos generalista dos transtornos mentais. Por fim, a metodologia aplicada aqui mostra-se eficiente para abrir caminhos a investigações maiores em possíveis marcas de atraso no desenvolvimento no sangue periférico de portadores de SCZ, e questiona a busca por padrões únicos de um transtorno, em vez de estabelecer um critério metodológico mais rígido para compreender estas, se não todas, patologias neurobiológicas e psiquiátricas.

4. Metodologia

Construção das redes globais de sinalização celular

As redes foram construídas tendo os nós como genes, e as arestas foram fundamentadas em interações proteína-proteína, proteína-gene e proteína-RNA mensageiro, bem como conexões que representam correlação direta de vias metabólicas. Toda a síntese foi feita usando o Cytoscape Consortium (Shannon, 2003) e o R Software (Contributors, R).

Para a rede descrita na metodologia de Banerji et al. (2013), chamada neste presente trabalho de Rede 1, as interações foram coletadas a partir da união de três bancos de dados robustos, sendo eles HPRD, IntAct Molecular Interaction Database - EMBL-EBI e Nature PID (Human Protein Reference Database; EMBL-EBI; Pathway Interaction Database), onde os dados foram baixados do Pathway Commons (www.pathwaycommons.org), e filtradas por ontologia gênica usando o Gene Ontology (Ashburner et al., 2000), anotando a sublocalização celular mais externa atribuída aos produtos dos genes e excluindo assim a conexão direta entre genes/nós intracelulares e extracelulares.

No intuito de trazer uma abordagem comparativa, a síntese de outra rede global de sinalização celular foi construída, identificada como Rede 2 e permissiva a maior número de arestas. Sob o critério descrito acima, foram coletadas interações dos bancos de dados HPRD, IntAct e um terceiro banco de dados distinto das origens da Rede 1: Reactome. A Rede 2 foi sintetizada sem etapa de filtro por Ontologia Gênica, aumentando significativamente o número de arestas entre nós.

Análise de topologia de rede

Cálculo dos parâmetros de rede e de nós foram coletados do Cytoscape, onde os dados de topologia global e os genes cujos 1000 maiores valores de degree ou betweenness centrality foram extraídos. Usando as anotações do Gene Ontology, foi aplicado o teste de Fischer usando o PANTHER Overrepresentation Test (Mi et al., 2018) e correção por FDR aos 1000 genes extraídos. A ferramenta de busca AmiGO (Carbon et al., 2008) foi usada para selecionar os termos gerais atribuídos à intersecção das 100 funções mais enriquecidas entre genes hub, ou entre genes gargalos, das duas redes.

Dados de expressão gênica global

Dados de expressão gênica de sangue periférico de humanos acometidos de transtornos mentais, sempre com grupo de controle interno, foram buscados de maneira sistemática na tentativa de obter apenas amostras de adultos, e coletados do repositório Gene Expression Omnibus (GEO), do NCBI (GEO DataSets - NCBI). Foram selecionados apenas estudos onde parâmetros de idade e sexo eram similares entre o grupo controle e grupos de transtornos), e presença de familiares nos grupos de estudos foram critério de exclusão.

Cálculo da entropia global de rede

O valor quantitativo de RNA mensageiro, foi atribuído individualmente a cada nó (gene) da rede. Em suma, valores repetidos para o mesmo gene (oriundos de splicing alternativo) foram aleatoriamente excluídos, obtendo assim um valor de expressão por gene. O cálculo da entropia consiste inicialmente na lei de ação de massas, para mensurar a probabilidade de interação entre dois nós a partir dos valores de expressão individuais, dentro de uma amostra. Tal probabilidade fornece a medida da entropia local, referente à incerteza de um nó perante seus ligantes diretos. A soma ponderada das entropias locais, usando a distribuição estacionária da rede, resulta na entropia global. As entropias globais foram normalizadas pela entropia máxima da rede a fins de representação. Detalhes matemáticos e descritivos são fornecidos no artigo de Banerji et al. (2013).

Análise estatística

Parâmetros estatísticos foram calculados usando o teste não paramétrico de Kruskal-Wallis, seguido de pós teste de Dunns para comparação múltipla entre colunas, para comparar o valor normalizado de entropia de sinalização global em mais de dois grupos no mesmo conjunto de dados de expressão gênica. O teste não paramétrico de soma dos postos de Wilcoxon foi usado para comparar dois grupos, sob o mesmo critério.

5. Referências

Disponível em: <<https://www.psychiatry.org/patients-families/what-is-mental-illness>>

Disponível em: <<http://pid.nci.nih.gov>>

- AHMAD, S. F. et al. Involvement of CD45 cells in the development of autism spectrum disorder through dysregulation of granulocyte-macrophage colony-stimulating factor, key inflammatory cytokines, and transcription factors. **International Immunopharmacology**, v. 83, p. 106466, 2020.
- ASHBURNER, M. et al. Gene Ontology: tool for the unification of biology. **Nature Genetics**, v. 25, n. 1, p. 25–29, 2000.
- BANERJI, C. R. S. et al. Cellular network entropy as the energy potential in Waddington's differentiation landscape. **Scientific Reports**, v. 3, n. 1, 2013.
- BANERJI, C. R. S. et al. Intra-Tumour Signalling Entropy Determines Clinical Outcome in Breast and Lung Cancer. **PLOS Computational Biology**, v. 11, n. 3, 2015.
- BEVEREN, N. J. M. V. et al. Marked Reduction of AKT1 Expression and Deregulation of AKT1-Associated Pathways in Peripheral Blood Mononuclear Cells of Schizophrenia Patients. **PLoS ONE**, v. 7, n. 2, 2012.
- BOUSMAN, C. A. et al. Preliminary evidence of ubiquitin proteasome system dysregulation in schizophrenia and bipolar disorder: Convergent pathway analysis findings from two independent samples. **American Journal of Medical Genetics Part B: Neuropsychiatric Genetics**, v. 153B, n. 2, p. 494–502, 2009.
- CARBON, S. et al. AmiGO: online access to ontology and annotation data. **Bioinformatics**, v. 25, n. 2, p. 288–289, 2008.
- CLELLAND, C. L. et al. Utilization of Never-Medicating Bipolar Disorder Patients towards Development and Validation of a Peripheral Biomarker Profile. **PLoS ONE**, v. 8, n. 6, 2013.
- Contributors**. Disponível em: <<https://www.r-project.org/contributors.html>>
- EMBL-EBI. **The home for big data in biology**. Disponível em: <<http://www.ebi.ac.uk/>>
- ENSTROM, A. et al. Peripheral Blood Leukocyte Production of BDNF following Mitogen Stimulation in Early Onset and Regressive Autism. **American Journal of Biochemistry and Biotechnology**, v. 4, n. 2, p. 121–129, 2008.
- GREGG, J. P. et al. Gene expression changes in children with autism. **Genomics**, v. 91, n. 1, p. 22–29, 2008.
- GROSSE, L. et al. Clinical characteristics of inflammation-associated depression: Monocyte gene expression is age-related in major depressive disorder. **Brain, Behavior, and Immunity**, v. 44, p. 48–56, 2015.
- HANAMSAGAR, R. et al. Generation of a microglial developmental index in mice and in humans reveals a sex difference in maturation and immune reactivity. **Glia**, v. 66, n. 2, p. 460–460, 2017.
- Home - Reactome** . Disponível em: <<https://reactome.org/>>
- Home - GEO DataSets - NCBI**. Disponível em: <<https://www.ncbi.nlm.nih.gov/gds>>
- Home - HPRD**. Disponível em: <<http://www.hprd.org/>>

- JONG, S. D. et al. A Gene Co-Expression Network in Whole Blood of Schizophrenia Patients Is Independent of Antipsychotic-Use and Enriched for Brain-Expressed Genes. **PLoS ONE**, v. 7, n. 6, 2012.
- KNIJFF, E. M. et al. Monocyte-Derived Dendritic Cells in Bipolar Disorder. **Biological Psychiatry**, v. 59, n. 4, p. 317–326, 2006.
- KUMAR, S.; JACK, R. Origin of monocytes and their differentiation to macrophages and dendritic cells. **Journal of Endotoxin Research**, v. 12, n. 5, p. 278–284, 2006.
- LEDAY, G. G. et al. Replicable and Coupled Changes in Innate and Adaptive Immune Gene Expression in Two Case-Control Studies of Blood Microarrays in Major Depressive Disorder. **Biological Psychiatry**, v. 83, n. 1, p. 70–80, 2018.
- LOMBARDO, M. V. et al. Large-scale associations between the leukocyte transcriptome and BOLD responses to speech differ in autism early language outcome subtypes. **Nature Neuroscience**, v. 21, n. 12, p. 1680–1688, 2018.
- MAUNEY, S. A. et al. Differentiation of oligodendrocyte precursors is impaired in the prefrontal cortex in schizophrenia. **Schizophrenia Research**, v. 169, n. 1-3, p. 374–380, 2015.
- MI, H. et al. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. **Nucleic Acids Research**, v. 47, n. D1, 2018.
- NOVERSHTERN, N. et al. Densely Interconnected Transcriptional Circuits Control Cell States in Human Hematopoiesis. **Cell**, v. 144, n. 2, p. 296–309, 2011.
- PARIKSHAK, N. N.; GANDAL, M. J.; GESCHWIND, D. H. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. **Nature Reviews Genetics**, v. 16, n. 8, p. 441–458, 2015.
- Pathway Commons**. Disponível em: <<http://www.pathwaycommons.org/>>
- PORTEOUS, D. J. et al. DISC1 at 10: connecting psychiatric genetics and neuroscience. **Trends in Molecular Medicine**, v. 17, n. 12, p. 699–706, 2011.
- PRAMPARO, T. et al. Prediction of Autism by Translation and Immune/Inflammation Coexpressed Genes in Toddlers From Pediatric Community Practices. **JAMA Psychiatry**, v. 72, n. 4, p. 386, 2015.
- RADU, A. et al. Microarray Database Mining and Cell Differentiation Defects in Schizophrenia. **Advances in Experimental Medicine and Biology Software Tools and Algorithms for Biological Systems**, p. 67–74, 2011.
- RITCHIE, H.; ROSER, M. **Mental Health**. Disponível em: <<https://ourworldindata.org/mental-health>>
- SAVITZ, J. et al. Inflammation and neurological disease-related genes are differentially expressed in depressed patients with mood disorders and correlate with morphometric and functional imaging abnormalities. **Brain, Behavior, and Immunity**, v. 31, p. 161–171, 2013.
- Setembro Amarelo**. Disponível em: <<http://www.ans.gov.br/aans/noticias-ans/consumidor/5140-setembro-amarelo>>
- SHANNON, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. **Genome Research**, v. 13, n. 11, p. 2498–2504, 2003.
- STERN, S. et al. Neurons derived from patients with bipolar disorder divide into intrinsically different sub-populations of neurons, predicting the patients' responsiveness to lithium. **Molecular Psychiatry**, v. 23, n. 6, p. 1453–1465, 2017.

Figuras e Tabelas

Tabela 1: Parâmetros de rede. Valores coletados no Cytoscape Consortium.

	nós	arestas	diâmetro	Coefficiente de clusterização
Rede 1	13990	134215	10	0,136
Rede 2	14151	130034	9	0,147

Bottleneck nodes

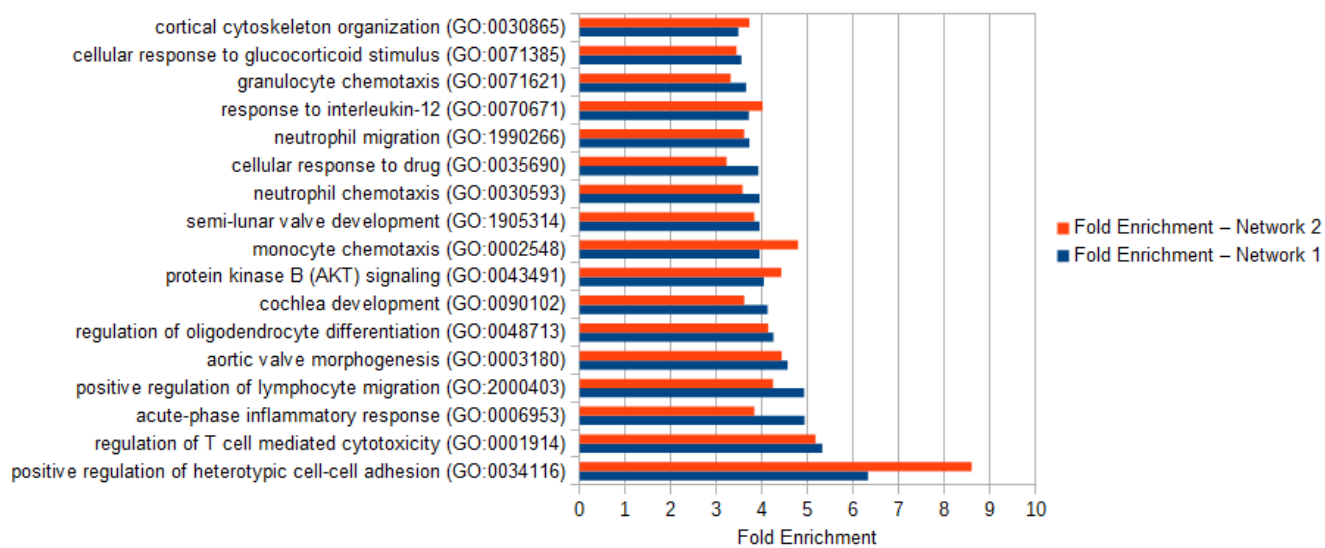


Figura 1. Enriquecimento das principais vias dos genes gargalos. Intersecção das 100 funções enriquecidas dos 1000 nós/genes, de cada rede, com maior valor em *Betweenness Centrality*, definindo os "gargalos" (bottleneck) da rede.

Hub nodes

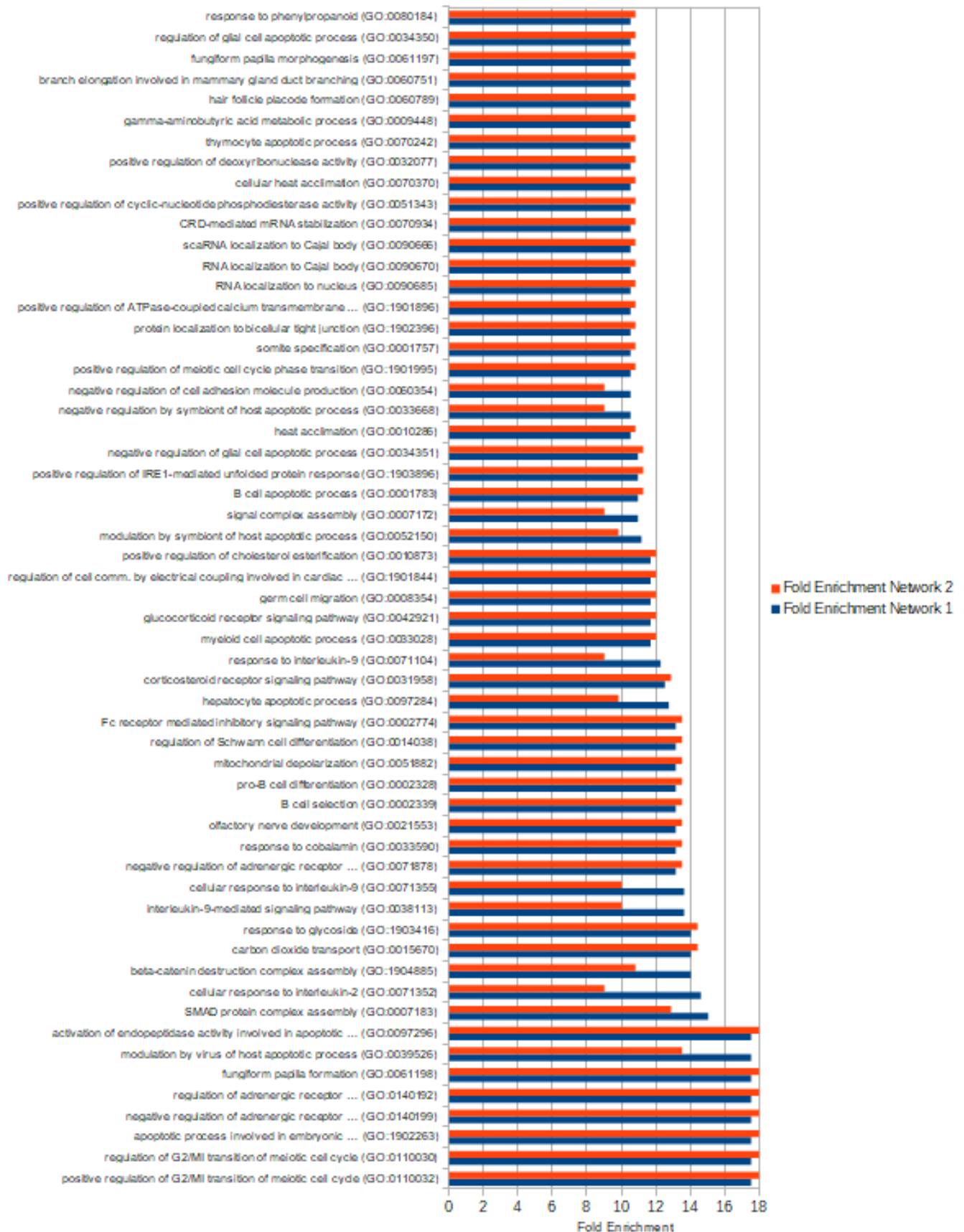


Figura 2. Enriquecimento das principais vias dos genes hub. Intersecção das 100 funções enriquecidas dos 1000 nós/genes, de cada rede, com maior valor em grau de nó, definindo os genes *hubs*.

Hematopoiesis

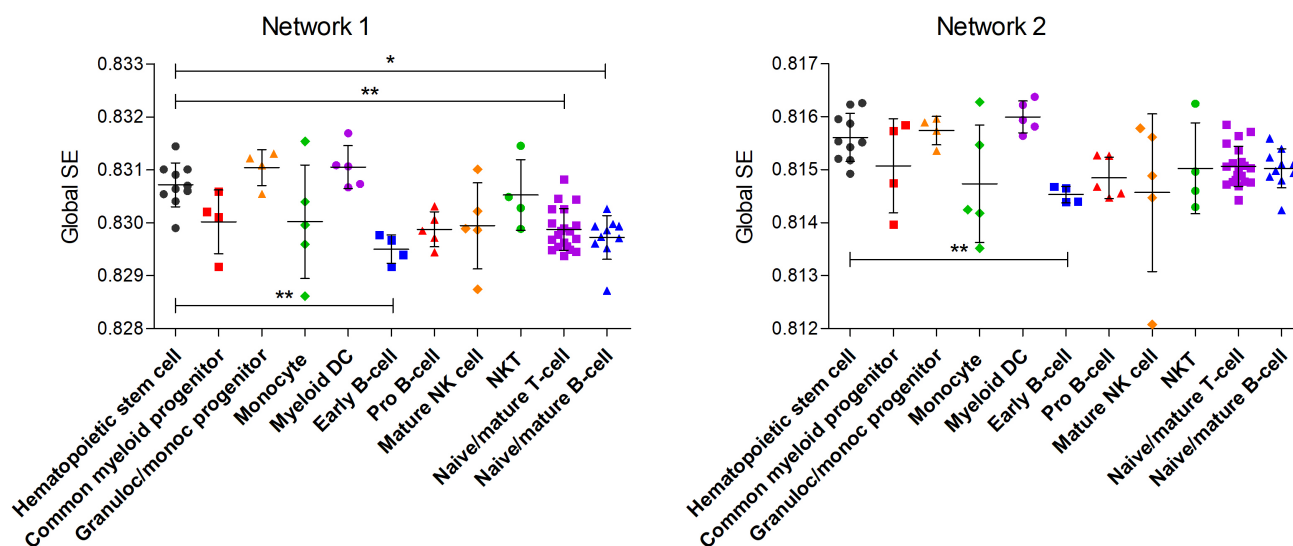
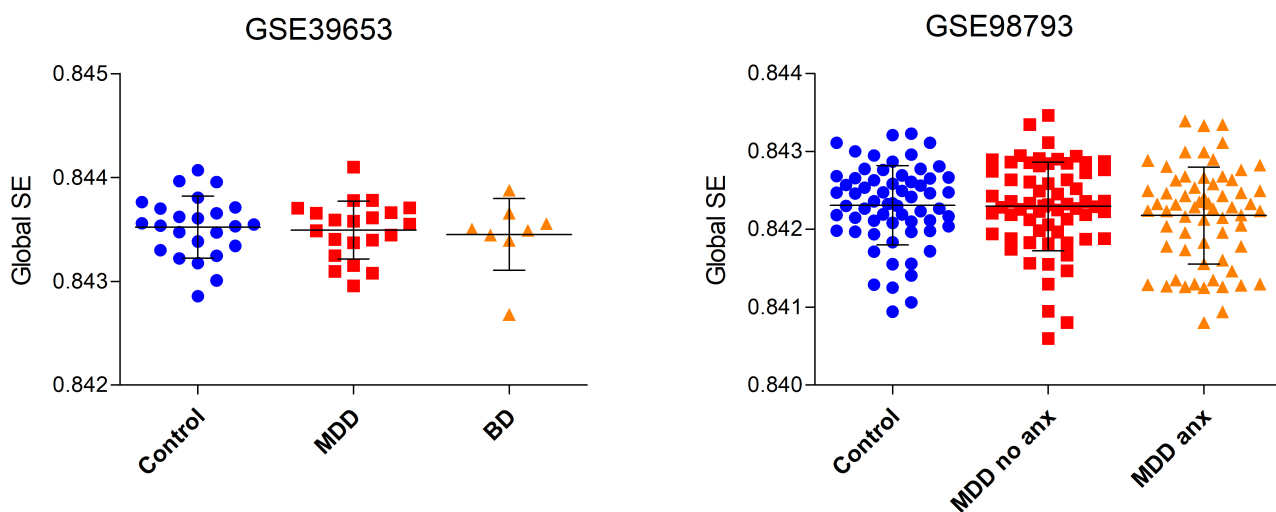


Figura 3. Entropia de rede global de sinalização celular, normalizada pela entropia máxima de rede (Global SE). Análise de amostras do GEO de células-tronco hematopoiéticas (HSC), células progenitoras mieloides, células progenitoras de granulócitos e monócitos, células precursoras de linfócitos B iniciais (early B-cell) ou não-iniciais (pro B-cell), células dendríticas mieloides (DC Myeloid), monócitos, Natural Killer (NK) maduras, células NKT, células B e células T. Diferença estatística representando apenas a comparação entre células progenitoras/tronco e suas linhagens diferenciadas. Além disso, na Rede 1 o grupo DC Myeloid obteve diferença significativa frente Early B-cell ($p-v < 0,01$), Naive/mature B-cell ($p-v < 0,05$) e Naive/mature T-cell ($p-v < 0,01$); enquanto na Rede 2 o grupo DC Myeloid demonstrou variação estatística frente Early B-cell ($p-v < 0,005$), Pro B-cell ($p-v < 0,05$) e Naive/mature T-cell ($p-v < 0,05$).

Network 1



Network 2

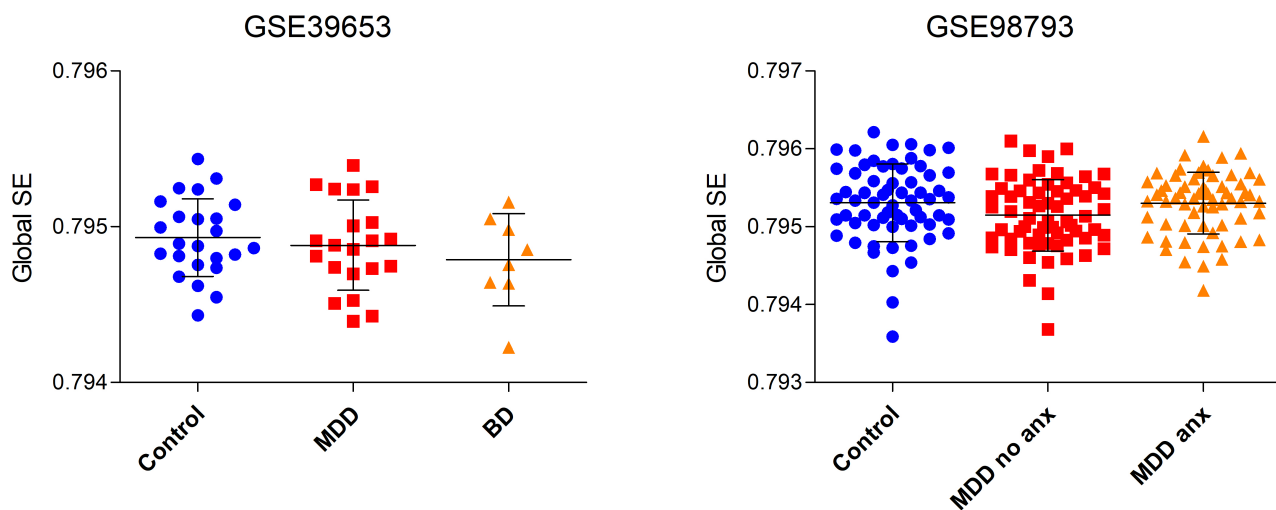
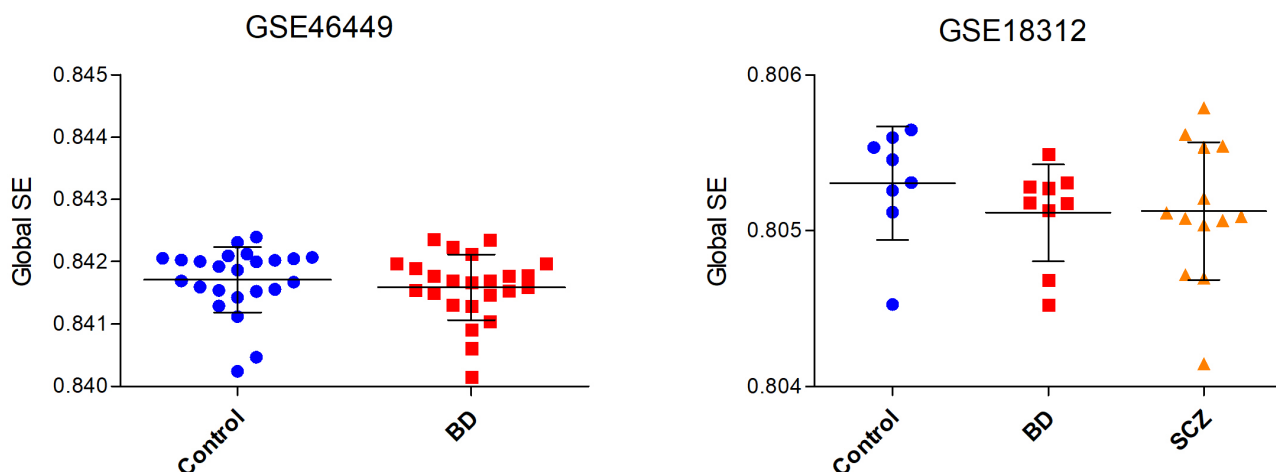


Figura 4. Entropia de rede global de sinalização celular, normalizada pela entropia máxima de rede (Global SE). Análise de amostras do GEO de pessoas saudáveis (Control) e acometidas por transtorno depressivo maior (MDD), do tipo ansioso ou não ansioso (MDD anx, MDD no anx), e transtorno bipolar (BD) por teste de classificação sinalizada de Wilcoxon para comparar estudos com duas amostras, e teste de Kruskal-Wallis seguido de pós teste de Dunns para comparação múltipla entre colunas (* p-v<0,05).

Network 1



Network 2

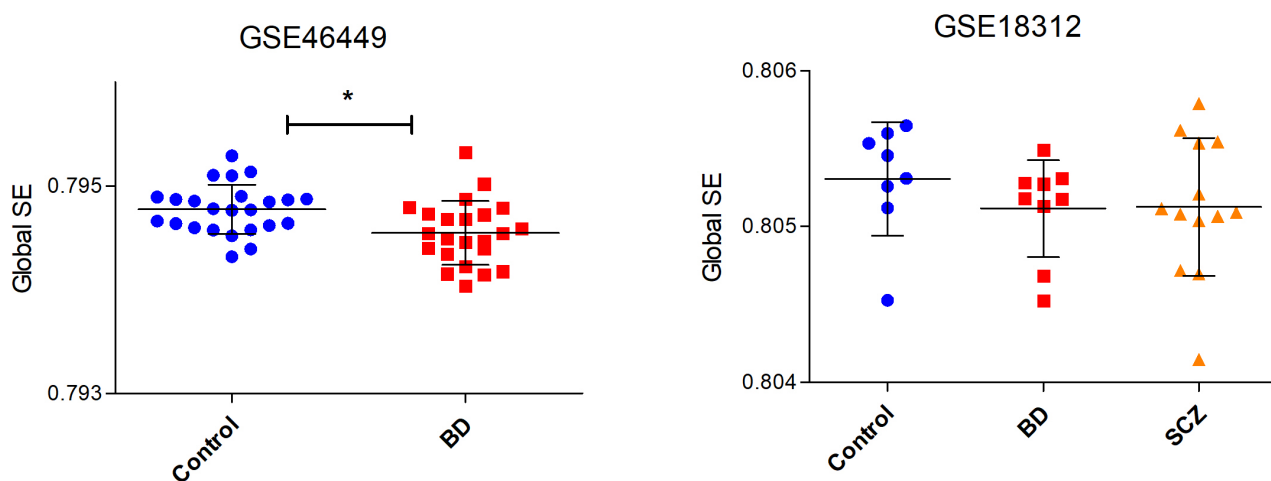
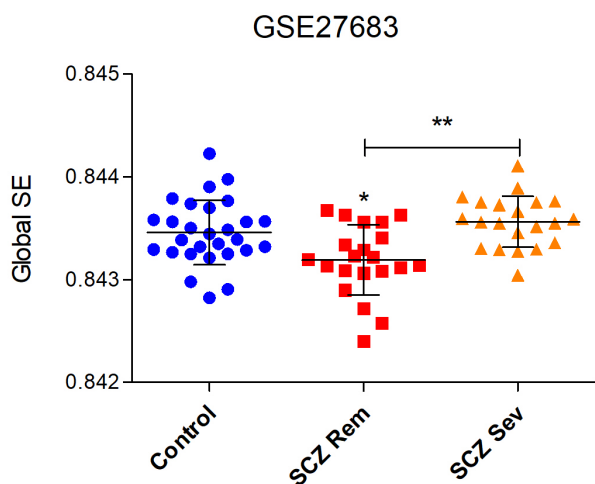


Figura 5. Entropia de rede global de sinalização celular, normalizada pela entropia máxima de rede (Global SE). Análise de amostras do GEO de pessoas saudáveis (Control) e acometidas por transtorno bipolar (BD) ou por esquizofrenia (SCZ), por teste de classificação sinalizada de Wilcoxon para comparar estudos com duas amostras, e teste de Kruskal-Wallis seguido de pós teste de Dunns para comparação múltipla entre colunas (* $p < 0,05$).

Network 1



Network 2

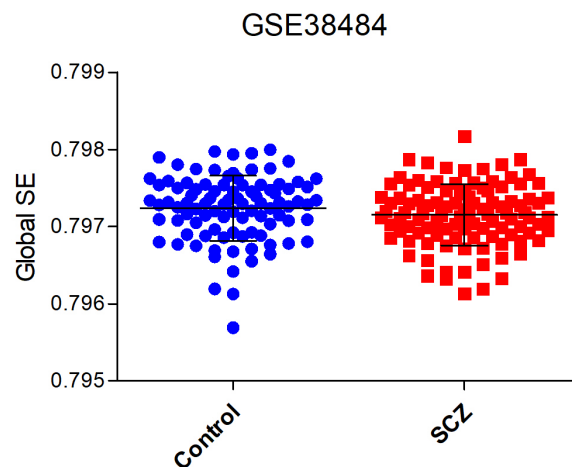
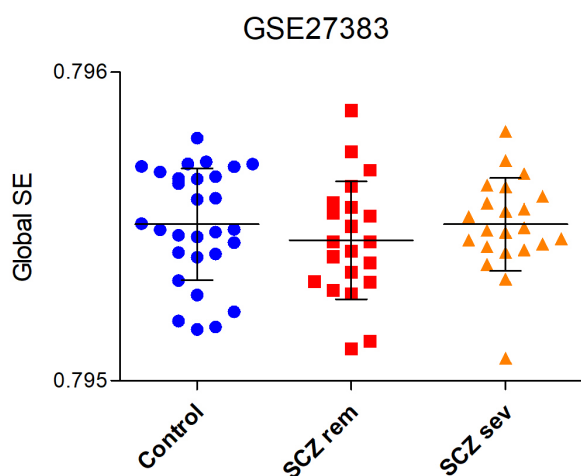
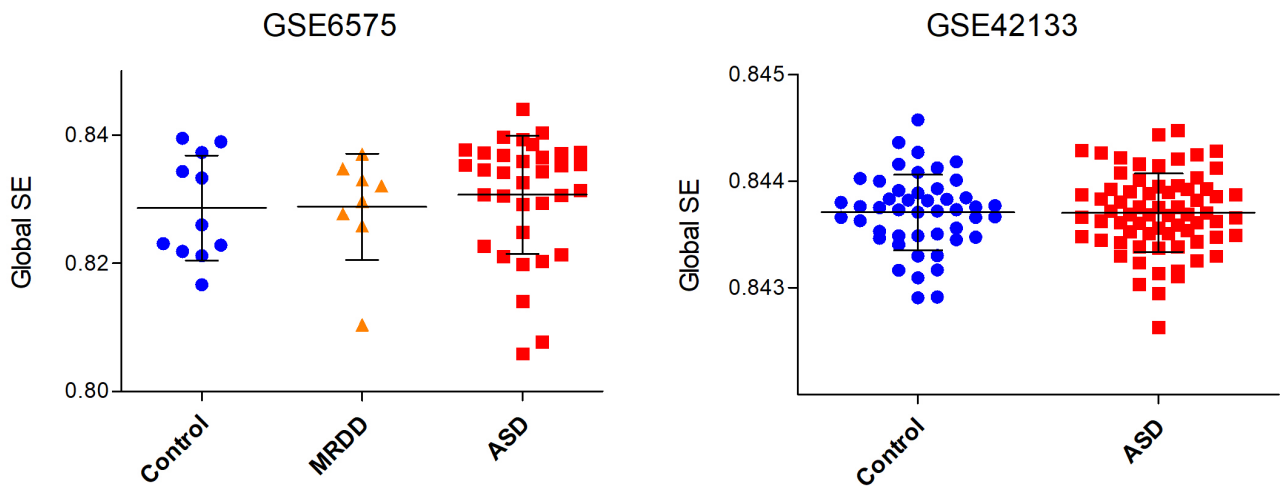


Figura 6 Entropia de rede global de sinalização celular, normalizada pela entropia máxima de rede (Global SE). Análise de amostras do GEO de pessoas saudáveis (Control) e acometidas por esquizofrenia (SCZ), com remissão de sintomas ou em estado agudo da doença (SCZ rem, SCZ sev), por teste de classificação sinalizada de Wilcoxon para comparar estudos com duas amostras, e teste de Kruskal-Wallis seguido de pós teste de Dunns para comparação múltipla entre colunas (* $p < 0,05$).

Network 1



Network 2

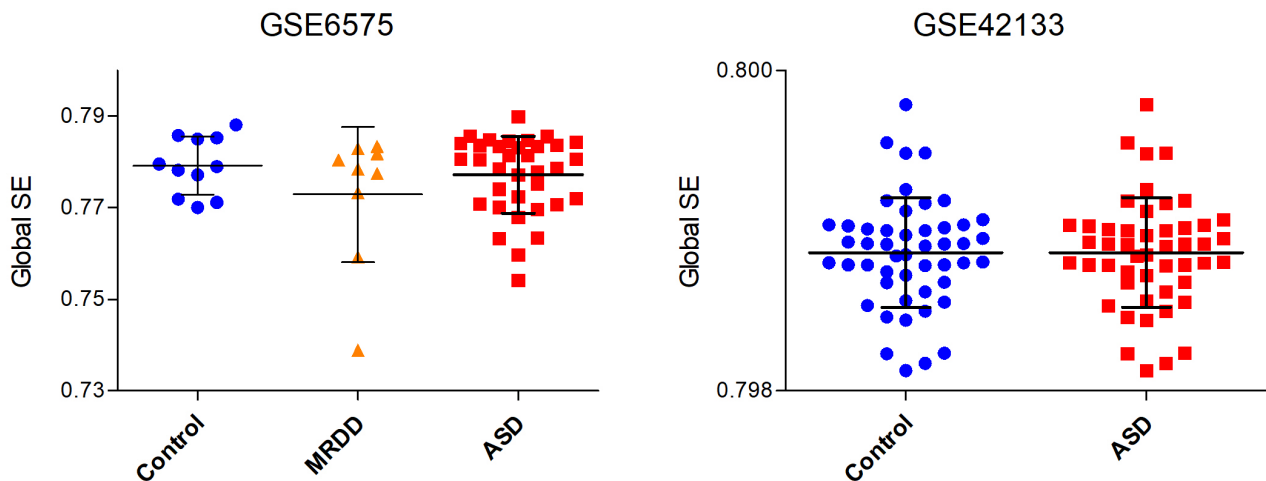


Figura 7. Entropia de rede global de sinalização celular, normalizada pela entropia máxima de rede (Global SE). Análise de amostras do GEO de pessoas saudáveis (Control) e acometidas por transtornos do espectro autista (ASD), ou atraso no desenvolvimento e retardo mental (MRDD), por teste de classificação sinalizada de Wilcoxon para comparar estudos com duas amostras, e teste de Kruskal-Wallis seguido de pós teste de Dunns para comparação múltipla entre colunas (* $p < 0,05$).

Material Suplementar

Somente foram inseridos no arquivo principal as tabelas cujas dimensões eram aceitáveis dentro das proporções deste documento. Em caso de acesso a todos os materiais, por favor entre em contato com o autor do texto.

Tabela Suplementar 3. Desdobramento das 100 funções mais enriquecidas por desdobramento, aos Top1000 genes gargalos. Desdobramento por teste de super-representação PANTHER, usando o Gene Ontology (GO) como base de dados.

GO biological process complete	Fold Enrichment Network 1	Fold Enrichment Network 2	GO general terms (non-redundant)
positive regulation of heterotypic cell-cell adhesion (GO:0034116)	6,32	8,59	cell-cell adesion
regulation of T cell mediated cytotoxicity (GO:0001914)	5,32	5,17	immune system process
acute-phase inflammatory response (GO:0006953)	4,93	3,83	inflammatory response
positive regulation of lymphocyte migration (GO:2000403)	4,92	4,24	immune system process / locomotion
aortic valve morphogenesis (GO:0003180)	4,56	4,43	anatomical structure development
regulation of oligodendrocyte differentiation (GO:0048713)	4,25	4,13	anatomical structure development
cochlea development (GO:0090102)	4,12	3,61	anatomical structure development
protein kinase B (AKT) signaling (GO:0043491)	4,04	4,42	response to stimulus
monocyte chemotaxis (GO:0002548)	3,94	4,79	immune system process / locomotion / response to chemical
semi-lunar valve development (GO:1905314)	3,94	3,83	anatomical structure development
neutrophil chemotaxis (GO:0030593)	3,94	3,57	immune system process / locomotion / response to chemical
cellular response to drug (GO:0035690)	3,92	3,22	response to chemical
neutrophil migration (GO:1990266)	3,72	3,61	immune system process / locomotion
response to interleukin-12 (GO:0070671)	3,71	4,01	response to chemical
granulocyte chemotaxis (GO:0071621)	3,65	3,31	immune system process / locomotion / response to chemical
cellular response to glucocorticoid stimulus (GO:0071385)	3,55	3,44	response to chemical
cortical cytoskeleton organization (GO:0030865)	3,48	3,72	cellular component organization or biogenesis

Tabela Suplementar 4. Abundância dos termos gerais atribuídos aos termos específicos da Tabela Suplementar 3. Análise manual de coleta de termos não redundantes, a partir do Gene Ontology (GO).

General terms	count	percentual proportion
immune system process	6	24
response to chemical	6	24
locomotion	5	20
anatomical structure development	4	16
cell-cell adesion	1	4
cellular component organization or biogenesis	1	4
inflammatory response	1	4
response to stimulus	1	4
total general terms	25	100

Tabela Suplementar 5. Desdobramento das 100 funções mais enriquecidas por desdobramento, aos Top1000 genes hubs. Desdobramento por teste de super-representação PANTHER, usando o Gene Ontology (GO) como base de dados.

GO biological process complete	Fold Enrichment Network 1	Fold Enrichment Network 2	GO general terms (non-redundant)
positive regulation of G2/M1 transition of meiotic cell cycle (GO:0110032)	17,54	18,03	regulation of cell cycle
regulation of G2/M1 transition of meiotic cell cycle (GO:0110030)	17,54	18,03	regulation of cell cycle
apoptotic process involved in embryonic digit morphogenesis (GO:1902263)	17,54	18,03	anatomical structure development / cell death
negative regulation of adenylate cyclase-activating adrenergic receptor signaling pathway involved in heart process (GO:0140199)	17,54	18,03	circulatory system process / regulation of response to stimulus
regulation of adenylate cyclase-activating adrenergic receptor signaling pathway involved in heart process (GO:0140192)	17,54	18,03	circulatory system process / regulation of response to stimulus
fungiform papilla formation (GO:0061198)	17,54	18,03	anatomical structure development
modulation by virus of host apoptotic process (GO:0039526)	17,54	13,53	biological process involved in interspecies interaction between organisms / cell death
activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway (GO:0097296)	17,54	18,03	cell death / metabolic process
SMAD protein complex assembly (GO:0007183)	15,04	12,88	regulation of response to stimulus
cellular response to interleukin-2 (GO:0071352)	14,62	9,02	response to chemical
beta-catenin destruction complex assembly (GO:1904885)	14,03	10,82	cellular component organization or biogenesis
carbon dioxide transport (GO:0015670)	14,03	14,43	transport
response to glycoside (GO:1903416)	14,03	14,43	response to chemical
interleukin-9-mediated signaling pathway (GO:0038113)	13,64	10,02	response to chemical
cellular response to interleukin-9	13,64	10,02	response to chemical

(GO:0071355)			
negative regulation of adenylate cyclase-activating adrenergic receptor signaling pathway (GO:0071878)	13,16	13,53	response to chemical
response to cobalamin (GO:0033590)	13,16	13,53	response to chemical
olfactory nerve development (GO:0021553)	13,16	13,53	anatomical structure development
B cell selection (GO:0002339)	13,16	13,53	anatomical structure development / immune system process
pro-B cell differentiation (GO:0002328)	13,16	13,53	anatomical structure development / immune system process
mitochondrial depolarization (GO:0051882)	13,16	13,53	intracellular organelle
regulation of Schwann cell differentiation (GO:0014038)	13,16	13,53	anatomical structure development
Fc receptor mediated inhibitory signaling pathway (GO:0002774)	13,16	13,53	immune system process
hepatocyte apoptotic process (GO:0097284)	12,76	9,84	cell death
corticosteroid receptor signaling pathway (GO:0031958)	12,53	12,88	response to chemical
response to interleukin-9 (GO:0071104)	12,28	9,02	response to chemical
myeloid cell apoptotic process (GO:0033028)	11,7	12,02	cell death
glucocorticoid receptor signaling pathway (GO:0042921)	11,7	12,02	response to chemical
germ cell migration (GO:0008354)	11,7	12,02	cell motility
regulation of cell communication by electrical coupling involved in cardiac conduction (GO:1901844)	11,7	12,02	multicellular organismal signaling
positive regulation of cholesterol esterification (GO:0010873)	11,7	12,02	metabolic process
modulation by symbiont of host apoptotic process (GO:0052150)	11,16	9,84	biological process involved in interspecies interaction between organisms / cell death
signal complex assembly (GO:0007172)	10,96	9,02	regulation of response to stimulus
B cell apoptotic process (GO:0001783)	10,96	11,27	cell death
positive regulation of IRE1-mediated unfolded protein response	10,96	11,27	regulation of signaling / response to chemical

(GO:1903896)			
negative regulation of glial cell apoptotic process (GO:0034351)	10,96	11,27	negative regulation of cell death
heat acclimation (GO:0010286)	10,53	10,82	response to stress
negative regulation by symbiont of host apoptotic process (GO:0033668)	10,53	9,02	biological process involved in interspecies interaction between organisms / negative regulation of cell death
negative regulation of cell adhesion molecule production (GO:0060354)	10,53	9,02	cell-cell adhesion
positive regulation of meiotic cell cycle phase transition (GO:1901995)	10,53	10,82	regulation of cell cycle
somite specification (GO:0001757)	10,53	10,82	anatomical structure development
protein localization to bicellular tight junction (GO:1902396)	10,53	10,82	macromolecule localization
positive regulation of ATPase-coupled calcium transmembrane transporter activity (GO:1901896)	10,53	10,82	transport
RNA localization to nucleus (GO:0090685)	10,53	10,82	macromolecule localization
RNA localization to Cajal body (GO:0090670)	10,53	10,82	macromolecule localization
scaRNA localization to Cajal body (GO:0090666)	10,53	10,82	macromolecule localization
CRD-mediated mRNA stabilization (GO:0070934)	10,53	10,82	cellular metabolic process
positive regulation of cyclic-nucleotide phosphodiesterase activity (GO:0051343)	10,53	10,82	regulation of molecular function
cellular heat acclimation (GO:0070370)	10,53	10,82	response to stress
positive regulation of deoxyribonuclease activity (GO:0032077)	10,53	10,82	cellular metabolic process
thymocyte apoptotic process (GO:0070242)	10,53	10,82	cell death
gamma-aminobutyric acid metabolic process (GO:0009448)	10,53	10,82	cellular metabolic process
hair follicle placode formation (GO:0060789)	10,53	10,82	anatomical structure development
branch elongation involved in mammary gland duct branching (GO:0060751)	10,53	10,82	anatomical structure development
fungiform papilla	10,53	10,82	anatomical structure development

morphogenesis (GO:0061197)			
regulation of glial cell apoptotic process (GO:0034350)	10,53	10,82	cell death
response to phenylpropanoid (GO:0080184)	10,53	10,82	response to chemical

Tabela Suplementar 6. Abundância dos termos gerais atribuídos aos termos específicos da Tabela Suplementar 5. Análise manual de coleta de termos não redundantes, a partir do Gene Ontology (GO).

General terms	count	percentual proportion
response to chemical	11	16.18
anatomical structure development	10	14.70
cell death	9	13.23
macromolecule localization	4	5.88
regulation of response to stimulus	4	5.88
biological process involved in interspecies interaction between organisms	3	4.41
cellular metabolic process	3	4.41
immune system process	3	4.41
metabolic process	3	4.41
regulation of cell cycle	3	4.41
circulatory system process	2	2.94
negative regulation of cell death	2	2.94
response to stress	2	2.94
transport	2	2.94
cell motility	1	1.47
cell-cell adesion	1	1.47
cellular component organization or biogenesis	1	1.47
intracellular organele	1	1.47
multicellular organismal signaling	1	1.47
regulation of molecular function	1	1.47
regulation of signaling	1	1.47