

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE  
CURSO DE GRADUAÇÃO EM BIOMEDICINA

Natália Cardoso Flores

**APLICAÇÃO DA BIOLOGIA DE SISTEMAS PARA A COMPREENSÃO DE  
FENÓTIPOS DE ADRENOLEUCODISTROFIA LIGADA AO X**

Porto Alegre

2022

Natália Cardoso Flores

**APLICAÇÃO DA BIOLOGIA DE SISTEMAS PARA A COMPREENSÃO DE  
FENÓTIPOS DE ADRENOLEUCODISTROFIA LIGADA AO X**

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel em Biomedicina.

Orientadora: Profa Dra Ursula da Silveira Matte

Coorientadora: Dra. Gerda Cristal Villalba Silva

Porto Alegre  
2022

#### CIP - Catalogação na Publicação

Flores, Natália Cardoso  
Aplicação da biologia de sistemas para a  
compreensão de fenótipos de adrenoleucodistrofia  
ligada ao X / Natália Cardoso Flores. -- 2022.  
54 f.

Orientadora: Ursula da Silveira Matte.

Coorientadora: Gerda Cristal Villalba Silva.

Trabalho de conclusão de curso (Graduação) --  
Universidade Federal do Rio Grande do Sul, Instituto  
de Ciências Básicas da Saúde, Curso de Biomedicina,  
Porto Alegre, BR-RS, 2022.

1. Adrenoleucodistrofia ligada ao X. 2. Biologia de  
sistemas. 3. Expressão gênica. 4. Doenças  
desmielinizantes. 5. Redes de interação  
proteína-proteína. I. da Silveira Matte, Ursula,  
orient. II. Villalba Silva, Gerda Cristal, coorient.  
III. Título.

Natália Cardoso Flores

**APLICAÇÃO DA BIOLOGIA DE SISTEMAS PARA A COMPREENSÃO DE  
FENÓTIPOS DE ADRENOLEUCODISTROFIA LIGADA AO X**

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel em Biomedicina.

Aprovado em: 20 de outubro de 2022.

**BANCA EXAMINADORA**

---

Doutoranda Ana Carolina de Moraes Mello - UFRGS

---

Dr. Jonas Alex Morales Saute - HCPA

---

Prfa. Dra. Ursula da Silveira Matte - UFRGS (orientador)

### **Agradecimentos**

Agradeço a meus pais Lisiane e Neri pelo carinho, apoio e incentivo. Seus esforços para me darem uma boa educação foram essenciais para que eu pudesse realizar essa graduação.

A meu irmão Juliano que sempre esteve presente nos momentos de estresse, me tranquilizando e me direcionando palavras de sabedoria.

A minha avó Norma, que ajudou a me criar e cuidou de mim durante toda a minha vida escolar e parte da minha vida acadêmica me fornecendo as bases para estabelecer meus conhecimentos.

A meu namorado Lucas que foi tão compreensivo quando precisei dedicar meu tempo para minha formação e que além disso doou o próprio tempo para fazer tudo estava ao seu alcance para me ajudar.

A todos os professores que me ensinaram a ser uma pessoa estudiosa e curiosa que sempre procura a verdade acima de tudo.

A meus colegas Thiago e Eduarda, amigos com os quais aprendi muito e recebi oportunidades únicas.

A minha coorientadora Cristal, uma pessoa que mesmo fisicamente longe se mostrou muito presente. Obrigada por toda a dedicação e paciência ao me orientar por conhecimentos e técnicas as quais eu não tinha experiência até então. A tua contribuição nesse trabalho é incalculável.

A minha professora e orientadora Ursula, uma pesquisadora muito comprometida por quem eu nutro enorme respeito e admiração. Obrigada por ter encontrado tempo para me auxiliar nessa pesquisa e me fazer pensar de forma crítica.

Compartilho essa vitória com todas essas pessoas que foram alicerces para a minha formação tanto como profissional quanto como pessoa. Obrigada!

## RESUMO

A Adrenoleucodistrofia ligada ao X (ALD) é um erro inato do metabolismo causado por variantes no gene *ABCD1* que codifica a proteína peroxissomal ALDP. Existem diferentes fenótipos relacionados a essa doença, os quais podem ser classificados de acordo com suas manifestações. O mais severo deles, chamado de adrenoleucodistrofia cerebral (CALD), causa desmielinização da matéria branca do cérebro e apresenta-se ainda na infância em muitos pacientes do sexo masculino, porém, também pode ser desenvolvida mais tardiamente. Curiosamente, pacientes com o mesmo genótipo podem apresentar diferentes fenótipos e não há marcadores genéticos ou metabólicos de prognóstico bem estabelecidos para CALD, bem como também não se conhecem os mecanismos pelos quais é causada. Dessa forma, supõe-se que as lesões à matéria branca do cérebro tenham outras causas que não as puramente genéticas. Nesses casos, tornam-se interessantes abordagens mais amplas, como a biologia de sistemas, pois essa análise pode conectar um grande número de fatores modificadores em potencial e torna possível visualizar as informações mais relevantes para maiores investigações. Esse estudo teve como foco a análise do perfil de expressão de CALD e outro fenótipo de ALD, adrenomieloneuropatia (AMN) em comparação com grupos controle. Os conjuntos de dados foram obtidos de bancos de dados públicos e processados somente com ferramentas de bioinformática para observar-se os genes diferencialmente expressos, sua presença em vias metabólicas e anotações sobre seus processos biológicos além de sua visualização em uma rede de interações. Os resultados mostram alterações na sinalização celular e na expressão de proteínas ribossômicas, confirmando consequências já observadas em outras doenças neurodegenerativas. É sugerido que alguns genes, tais como *RPS2*, *RPS21*, *SRP14* e *FUS*, sejam considerados alvos em potencial para maiores investigações em pesquisas experimentais sobre ALD.

**Palavras-chave:** adrenoleucodistrofia; biologia de sistemas; doenças desmielinizantes; expressão gênica; redes de interações proteína-proteína.

## ABSTRACT

X-linked adrenoleukodystrophy (ALD) is an inborn error of the metabolism caused by variants in *ABCD1* gene that codes for the peroxisomal protein ALDP. There are different phenotypes related to that disease, which can be classified by their manifestations. The severe form, cerebral adrenoleukodystrophy (CALD), causes demyelination of the brain white matter and has an early onset in the childhood of many male patients. Curiously, patients with the same genotype can present different phenotypes and there are no established genetic or metabolic markers of prognosis for CALD and neither a known mechanism by which it is caused. Thus, the damaging of brain white matter is supposed to be caused by reasons not purely genetic. For these purpose, systems biology approaches can be useful to understand the mechanisms of neurological impairment in the different phenotypes. The present study focused on the analysis of the gene expression profile of CALD and another ALD phenotype, adrenomyeloneuropathy (AMN) compared to control groups. The datasets were obtained from public databases and processed only with bioinformatic tools to observe the differential expressed genes, their function in metabolic pathways and biological processes annotations, as well as their interaction in a network. The results showed cell signaling alterations and differential expression of many ribosomal proteins, thus confirming consequences already observed in other neurodegenerative disorders. It is suggested that some of the genes found, like *RPS2*, *RPS21*, *SRP14* and *FUS*, be considered potential targets to further investigation of ALD in experimental research.

**Keywords:** adrenoleukodystrophy; systems biology; demyelinating diseases; gene expression; interaction network, protein

## SUMÁRIO

|  |    |
|--|----|
| <b>1 INTRODUÇÃO</b> .....  | 8  |
| <b>1.1 Erros inatos do metabolismo</b> .....   | 8  |
| <b>1.2 Doenças peroxissomais</b> .....   | 10 |
| <b>1.3 Adrenoleucodistrofia ligada ao X</b> .....                                      | 10 |
| <b>1.4 Biologia de sistemas</b> .....  | 12 |
| <b>1.5 Justificativa</b> .....   | 14 |
| <b>2 OBJETIVOS</b> .....   | 14 |
| <b>2.1 Objetivo geral</b> .....  | 14 |
| <b>2.2 Objetivos específicos</b> .....   | 14 |
| <b>3 ARTIGO CIENTÍFICO</b> .....   | 15 |
| <b>3.1 Introdução</b> .....  | 16 |
| <b>3.2 Metodologia</b> .....   | 16 |
| <b>3.2.1 Análises transcriptômicas</b> .....   | 16 |
| <b>3.2.2 Redes de interação</b> .....  | 18 |
| <b>3.2.3 Análise de enriquecimento</b> .....   | 18 |
| <b>3.3 Resultados</b> .....  | 18 |
| <b>3.3.1 Clusters, genes hub e topografia das redes</b> .....                          | 18 |
| <b>3.3.2 Resultados dos enriquecimentos</b> .....                                      | 22 |
| <b>3.3.3 E-MEXP-3288</b> .....   | 26 |
| <b>3.4 Discussão</b> .....   | 28 |
| <b>3.5 Conclusão</b> .....   | 32 |
| <b>3.6 Referências</b> .....   | 33 |
| <b>3.7 Declarações</b> .....   | 38 |
| <b>3.8 Informações suplementares</b> .....   | 38 |
| <b>4 Conclusões e Perspectivas</b> .....   | 48 |
| <b>Referências gerais</b> .....  | 49 |
| <b>Anexo A - Normas de Publicação da Revista <i>Metabolic Brain Diseases</i></b> ..... | 50 |



# 1 INTRODUÇÃO

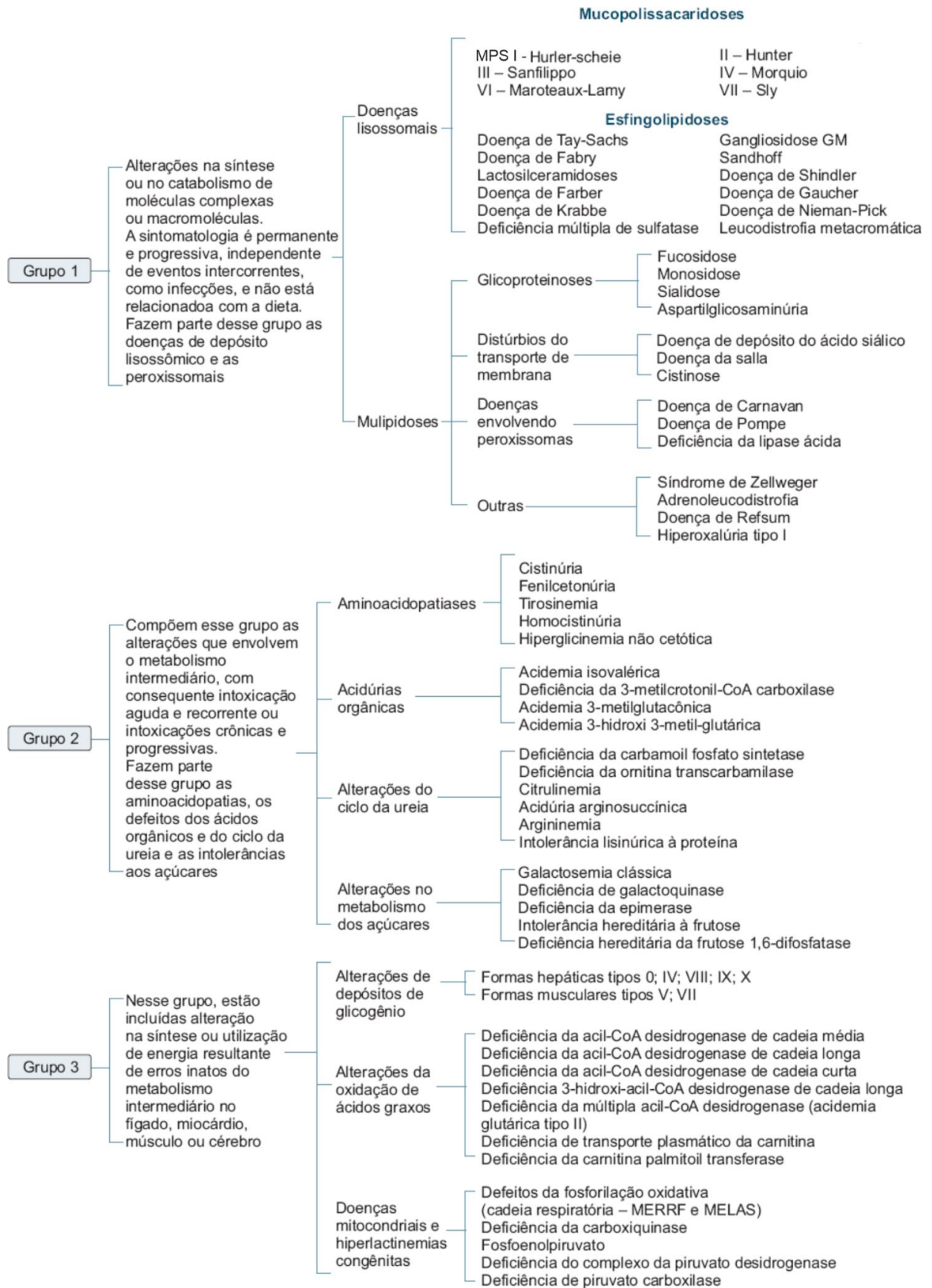
## 1.1 Erros inatos do metabolismo

Os erros inatos do metabolismo (EIM) são um amplo grupo de doenças causadas por mutações em genes que codificam proteínas com funções em vias metabólicas (ARGMANN et al., 2015). Conseqüentemente, esses distúrbios levam a falhas na síntese, degradação, armazenamento e transporte de moléculas tais como carboidratos, ácidos graxos e proteínas. Essas doenças são geneticamente herdadas, sendo a maioria de caráter autossômico recessivo, enquanto algumas apresentam padrão de herança ligado ao cromossomo X. As manifestações patológicas são heterogêneas de forma que muitas vezes não se encontra relação entre a gravidade dos sintomas e a(s) variante(s) genética(s) responsáveis pela doença, sugerindo-se a atuação de fatores modificadores e a semelhança dos EIM com outras doenças complexas (ARGMANN et al., 2015).

As enzimas têm papel chave no desenvolvimento dos EIM, pois são as principais proteínas com função de síntese ou degradação de moléculas orgânicas e normalmente participam de reações em cadeia nas quais os processos são interdependentes. Mutações em genes codificantes para enzimas, bem como em quaisquer outros relacionados com alguma etapa de um processo metabólico, podem causar tanto o acúmulo de produto de uma determinada enzima quanto a falta de substrato para outra. Nesse caso, as substâncias encontram-se respectivamente em concentrações tóxicas ao organismo ou incompatíveis com a manutenção do metabolismo (PINTO, 2017).

Os sintomas encontrados em EIM são inespecíficos e comuns a diversas doenças metabólicas e também a quadros infecciosos, de forma que torna-se difícil o diagnóstico precoce e fácil a confusão com outras disfunções. A classificação das doenças em três grupos organizados de acordo com os fenótipos clínicos pode ser útil para facilitar a sua compreensão (Figura 1; SAUDUBRAY et al., 2006). O primeiro grupo contempla as doenças nas quais os indivíduos afetados apresentam defeitos na síntese ou quebra de moléculas complexas. O segundo grupo envolve os erros de metabolismo intermediário que levam à intoxicação do indivíduo e por fim o terceiro contém as disfunções relacionadas com a deficiência na produção ou uso de energia pelo organismo. Todos esses podem ser subdivididos em categorias mais específicas, destacando-se para essa discussão uma das subclasses do terceiro grupo, as doenças peroxissomais.

Figura 1. Classificação de EIM de acordo com o fenótipo clínico.



Fonte: Adaptado de PINTO, Wagner de Jesus. Bioquímica clínica / Wagner de Jesus Pinto. – 1. ed. – Rio de Janeiro: Guanabara Koogan, 2017.

## 1.2 Doenças peroxissomais

Os peroxissomos são organelas conhecidas por armazenarem enzimas relacionadas ao metabolismo do peróxido de hidrogênio e estão presentes em todas as células, com exceção das hemácias. Dentre suas funções metabólicas, o peroxissomo sintetiza plasmalogênios, importantes para a constituição de membranas e mielina, exerce a beta-oxidação de ácidos graxos de cadeia muito longa (VLCFA, do inglês *very long chain fatty acids*) que não podem ser metabolizados na mitocôndria, realiza alfa-oxidação de ácidos graxos e faz a desintoxicação de glioxilato (POLL-THE, AUBOURG & WANDERS, 2006). Problemas envolvendo uma ou mais funções constituem a subclasse de EIM chamados de doenças peroxissomais.

Esses transtornos são genéticos e apresentam padrão de herança autossômico recessivo, com exceção da adrenoleucodistrofia ligada ao X. As doenças são classificadas em três grupos, sendo o primeiro relacionado a problemas de biogênese do peroxissomo, o segundo referente a deficiência de enzima peroxissomal e o terceiro relativo a deficiência de transporte de substrato peroxissomal (AUBOURG & WANDERS, 2013).

Em geral os pacientes podem apresentar características dismórficas no crânio e esqueleto, anomalias neurológicas e disfunções hepáticas e gastrointestinais. O início dos sintomas é associado normalmente com o atraso do desenvolvimento neural ainda na infância e coincide com a idade escolar. Esses sinais clínicos são observados juntamente com a idade do indivíduo para que seja feito o diagnóstico diferencial, uma vez que não há exame laboratorial que possa testar a presença de todas as doenças peroxissomais de uma vez. Esse é um passo importante para buscar o tratamento mais adequado, pois enquanto que para algumas dessas doenças não há tratamento, para outras existem intervenções capazes de desacelerar a progressão dos danos. Na doença de Refsum, por exemplo, intervenções no consumo de certos tipos de lipídio combinado ou não à plasmaférese têm sido empregadas para deter neuropatias.

## 1.3 Adrenoleucodistrofia ligada ao X

A adrenoleucodistrofia ligada ao X (ALD) é a única constituinte do terceiro grupo de EIM peroxissomais descritas acima e é o distúrbio peroxissomal mais comum com incidência de aproximadamente 1:17000, nos Estados Unidos, considerando os pacientes de ambos os sexos que manifestam sintomas (MOSER et al., 2016). Essa doença se caracteriza pela presença de variantes de perda de função do gene *ABCD1* que se encontra na região 28 do braço longo do cromossomo X. Esse gene codifica uma proteína pertencente à superfamília dos transportadores

de cassete de ATP (ABC), dividida em sete subfamílias, incluindo a subfamília ALD à qual pertence a proteína ALDP. Essa proteína é codificada por *ABCD1* e sua função envolve o transporte de ácidos graxos para o interior do peroxissomo, processo necessário para que seja realizada a oxidação de VLCFA (GeneCards symbol: *ABCD1*).

Consequentemente à perda dessa função ocorre o acúmulo de VLCFA em tecidos e no plasma dos pacientes (RICHMOND et al., 2020). Outros sinais clínicos podem envolver as glândulas adrenais e acometimentos neurológicos, porém os fenótipos são variados e não apresentam relação direta entre a gravidade da doença manifestada com a variante genética do indivíduo. Dentre as formas existentes, destacam-se a adrenomieloneuropatia (AMN) e a adrenoleucodistrofia cerebral (CALD), que muitas vezes são observadas dentro de uma mesma família.

AMN é o fenótipo mais comum e é apresentado por 60% das mulheres heterozigotas após os 40 anos de idade e 60% dos pacientes adultos do sexo masculino por volta dos 20 anos (POLL-THE, AUBOURG & WANDERS, 2006). Nessa manifestação ocorrem danos à medula espinhal, axonopatias distais e, na maioria dos homens, insuficiência das glândulas adrenais. Apesar de incomum, também pode ser observada atrofia do cerebelo e aspecto cinzento da matéria branca em algumas regiões do cérebro, porém sem desmielinização (AUBOURG & WANDERS, 2013).

CALD é considerada a versão mais severa da doença sendo que em 40% dos casos é manifestada ainda na infância por meninos entre 5 e 12 anos no fenótipo conhecido como *childhood cerebral* ALD (CCALD). O início de CCALD é observado com dificuldades escolares, falta de atenção, mudanças de comportamento, dano visual e auditivo que evoluem para quadriplegia, ataxia cerebelar e comumente convulsões, hipertensão intracraniana e acometimentos adrenais. O cérebro tem a matéria branca atingida por desmielinizações, causando lesões muitas vezes simétricas, de caráter inflamatório e rápida degeneração neurológica (POLL-THE, AUBOURG & WANDERS, 2006). Estima-se que 35% dos pacientes do sexo masculino correm risco de desenvolver CCALD e 35% dos pacientes adultos com AMN desenvolvem a forma cerebral da doença, totalizando 65% de chance de um homem com ALD apresentar o fenótipo CALD, enquanto que nenhuma mulher desenvolve a forma cerebral (AUBOURG & WANDERS, 2013). Esses indivíduos também apresentam desmielinização e sintomas neurológicos semelhantes com o mesmo desfecho, porém de forma mais lenta (POLL-THE, AUBOURG & WANDERS, 2006).

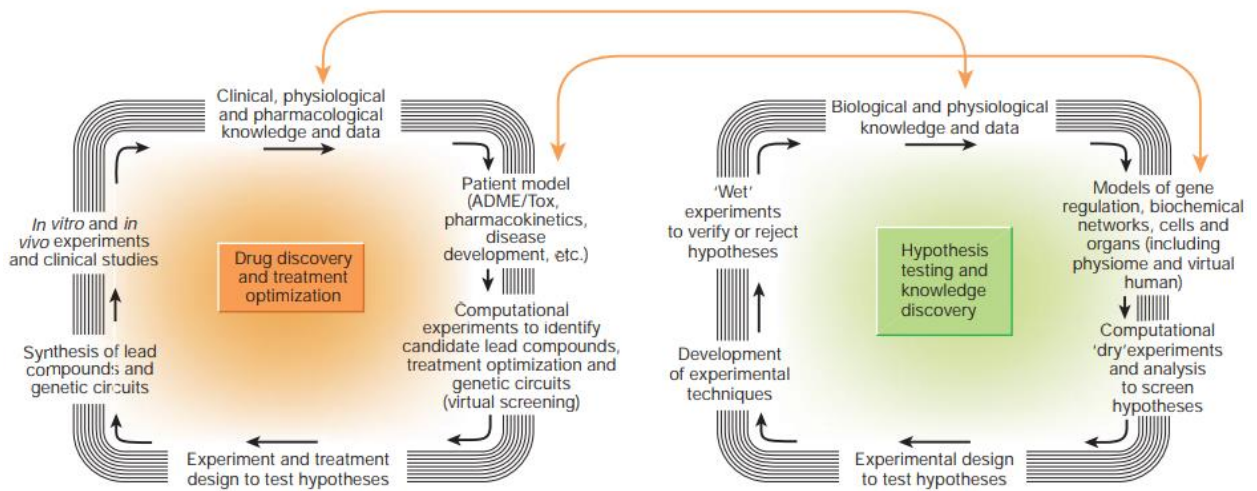
Os melhores métodos para diagnosticar indivíduos de ambos os sexos com ALD incluem a quantificação de lisofosfatidilcolina-C26:0 no sangue por meio de cromatografia líquida

acoplada a espectrometria de massa em tandem e o sequenciamento do gene *ABCD1* (TURK et al., 2020). O diagnóstico precoce permite tratamento com transplante de células tronco hematopoiéticas (TCH) para a desaceleração da desmielinização em casos de CALD, no entanto estima-se que menos de 50% dos pacientes aptos a receber TCH conseguem fazê-lo no período em que o tratamento ainda é eficaz (POLL-THE, AUBOURG & WANDERS, 2006). Outra alternativa é a administração de óleo erúxico, conhecido como Óleo de Lorenzo. Em diferentes estudos no qual esse tratamento foi testado, chegou-se à conclusão que o óleo está associado à normalização dos níveis de VLCFA no sangue, e mostra benefício quando usado por crianças assintomáticas. No entanto, esses resultados podem ser controversos dada a dificuldade de realizar ensaios clínicos com grupos controle (FATOUH, 2022).

#### **1.4 Biologia de sistemas**

Ainda que não exista um consenso sobre sua definição, a biologia de sistemas é, resumidamente, uma abordagem usada para compreender interações em grande escala unindo partes menores que dependem umas das outras. Essas análises são multidisciplinares e têm como bases a bioinformática, a sinalização celular, a genômica e a proteômica. A demanda por essa abordagem surge da complexidade comumente associada aos organismos vivos, nos quais múltiplos fatores atuam para um resultado final. Análises experimentais de forma clássica muitas vezes observam os fatores de forma isolada para identificar qual a possível contribuição de cada um nesse resultado. Com auxílio de ferramentas computacionais é possível combinar diferentes resultados experimentais, filtrá-los, testar sua relevância e por fim sugerir novas hipóteses ou predições *in silico* que podem ser úteis para novos experimentos *in vitro* e *in vivo* (KITANO, 2002). Este é um ciclo de enriquecimento dos conhecimentos sobre biologia e saúde bem como de avanço em tratamentos (Figura 2).

Figura 2. Ciclo de utilização da biologia de sistemas



Fonte: KITANO, Hiroaki. Computational systems biology. Nature, vol.420, 14 nov 2002

Levando em consideração que as proteínas são moléculas importantes para a estrutura e funções metabólicas dos organismos, é natural que elas sejam um foco de análise de abordagens de biologia de sistemas. Para fins de discussão no presente estudo, destaca-se uma técnica em específico que observa as interações proteicas como nós e arestas de uma rede. Nessas redes de interações as proteínas podem ser representadas como nós, enquanto a interação direta entre duas proteínas (interação proteína-proteína) é considerada uma aresta (ALM & ARKIN, 2003). Além dessas interações diretas, a rede permite visualizar nós com maiores números de interações, chamados *hub*, regiões mais densamente ligadas, ou *clusters* (BADER & HOGUE, 2003), bem como propriedades da rede em geral que dizem respeito a sua densidade, ao grau de ligação de seus nós e às demais medidas de centralidade que ligam suas extremidades (SCARDONI, TOSADORI & LAUDANNA, 2014).

Abordagens de biologia de sistemas são sugeridas como boas alternativas para observar fatores modificadores em EIM (ARGMANN et al., 2015). O emprego de redes de interação no estudo de doenças lisossômicas, um outro grupo de EIM, trouxe melhor compreensão sobre os danos neurológicos que ocorrem em mucopolissacaridoses (SILVA & MATTE, 2022). Além disso, já se pesquisou sobre a heterogeneidade dos fenótipos de ALD com abordagens multi-ômicas, também com resultados interessantes (RICHMOND et al., 2020). Pontuando a semelhança de EIMs, pode-se sugerir que a união dessas técnicas seja útil para trazer maior compreensão quanto a mecanismos que causam fenótipos os heterogêneos de ALD.

## **1.5 Justificativa**

ALD é causada por variantes de um único gene, porém a gravidade de seus fenótipos não apresenta correlação com os genótipos, de forma que a doença adquire um caráter heterogêneo. Por esse motivo, se observa sua semelhança com doenças complexas em que há fatores agindo em conjunto, como agentes ambientais, epigenéticos, bem como a contribuição de genes modificadores. Como já observado em outros trabalhos, dois métodos distintos de biologia de sistemas permitiram melhor compreensão tanto de outro EIM como da própria ALD, porém não se conhece uma exploração que incorpore fatores desses dois conceitos. Além disso, não existe consenso na comunidade científica sobre como um fenótipo que não causa desmielinização no cérebro pode passar a fazê-lo e tampouco existem marcadores úteis para o prognóstico de pacientes com a doença. Nessa pesquisa foi então realizada a integração de análises transcriptômicas com o uso de redes de interação na tentativa de explorar novas hipóteses sobre essas questões.

## **2 OBJETIVOS**

### **2.1 Objetivo geral**

O objetivo dessa pesquisa é analisar os perfis de expressão gênica na ALD, comparando amostras de indivíduos com AMN e com CALD.

### **2.2 Objetivos específicos**

- Traçar o perfil de expressão de amostras de AMN, CALD e controles saudáveis.
- Comparar os diferentes conjuntos de dados para obter-se os genes diferencialmente expressos.
- Observar as proteínas codificadas por cada um desses genes em uma rede de interações.
- Procurar por clusters e genes hub que podem ser potenciais agentes modificadores dos fenótipos.
- Relacionar as interações com vias metabólicas em uma análise de enriquecimento.
- Comparar os achados de diferentes conjuntos de dados e procurar por semelhanças e diferenças.

### 3 ARTIGO CIENTÍFICO

Os resultados preliminares obtidos nessa pesquisa foram organizados no formato de artigo científico de acordo com as normas de submissão do periódico “*Metabolic Brain Diseases*”. Neste momento o artigo está em português e será, após avaliação da banca, traduzido para a língua inglesa.

#### **A adrenoleucodistrofia sob a análise de redes de interação**

Natália Flores<sup>1</sup>, Gerda Cristal Villalba Silva<sup>2</sup>, Ursula Matte<sup>2,3</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, Curso de Graduação em Biomedicina, Instituto de Ciências básicas da Saúde, Porto Alegre, RS, Brasil

<sup>2</sup> Núcleo de Bioinformática do Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brasil

<sup>3</sup> Universidade Federal do Rio Grande do Sul, Departamento de Genética, Porto Alegre, RS, Brasil

Autor correspondente: Natália Flores, email: [natyfloreescardoso@gmail.com](mailto:natyflorescardoso@gmail.com)

#### **RESUMO**

A Adrenoleucodistrofia ligada ao X (ALD) é um erro inato do metabolismo causado por variantes no gene *ABCD1*, que codifica a proteína peroxissomal ALDP. Existem diferentes fenótipos relacionados a essa doença, os quais podem ser classificados de acordo com suas manifestações. O mais grave deles, chamado de ALD cerebral (CALD), causa desmielinização da matéria branca do cérebro (MB) e pode iniciar-se ainda na infância. Curiosamente, indivíduos com o mesmo genótipo podem apresentar diferentes fenótipos e não há marcadores genéticos ou metabólicos de prognóstico bem estabelecidos para CALD, bem como também não se conhecem os mecanismos pelos quais é causada. Dessa forma, supõe-se que as lesões à MB tenham outras causas que não as puramente genéticas. Nesses casos, tornam-se interessantes abordagens mais amplas, como a biologia de sistemas, pois essa análise pode conectar um grande número de fatores modificadores em potencial e torna possível visualizar as informações mais relevantes para maiores investigações. Esse estudo teve como foco a análise do perfil de expressão de CALD e outro fenótipo de ALD, adrenomieloneuropatia (AMN) em comparação com grupos controle. Os conjuntos de dados foram obtidos de bancos de dados públicos e processados somente com ferramentas de bioinformática para observar-se os genes diferencialmente expressos, sua presença em vias metabólicas e anotações sobre seus processos biológicos além de sua visualização em uma rede de interações. Os resultados mostram alterações na sinalização celular e na expressão de proteínas ribossômicas, confirmando consequências já observadas em outras doenças neurodegenerativas. É sugerido que alguns genes, tais como *RPS2*, *RPS21*, *SRP14* e *FUS*, sejam considerados alvos em potencial para maiores investigações em pesquisas experimentais sobre ALD.



**Palavras-chave:** adrenoleucodistrofia; biologia de sistemas; doenças desmielinizantes; expressão gênica; redes de interações proteína-proteína.

### 3.1 Introdução

Adrenoleucodistrofia ligada ao X (ALD; OMIM: 300100) é um erro inato do metabolismo (EIM), sendo a doença peroxissomal mais comum, com incidência de 1:17000, nos Estados Unidos, considerando os pacientes de ambos os sexos que manifestam sintomas (Moser et al. 2016). Sua causa está relacionada com variantes no gene *ABCD1*, que codifica a proteína ALDP responsável pelo transporte de ácidos graxos de cadeia muito longa (VLCFA, do inglês *very long chain fatty acids*) para o interior do peroxissomo (Moser 1997), onde esses lipídios sofrem beta-oxidação (Poll-The, Aubourg & Wanders 2006). O comprometimento dessa função provoca o acúmulo de VLCFA nos tecidos e plasma de indivíduos com a doença (O'Neill et al. 1982; Honey et al. 2021).

Uma das formas da doença, conhecida como adrenomieloneuropatia (AMN), é desenvolvida na idade adulta por praticamente todos os pacientes do sexo masculino e 80% das mulheres heterozigotas (Kemp et al. 2016). Nessa manifestação, ocorre neuropatia progressiva da medula espinhal e nervos periféricos, combinada a distúrbios sensoriais (Moser, Mahmood & Raymond 2007). Aproximadamente 35% dos meninos que nascem com ALD e 35% dos homens com AMN desenvolvem uma forma inflamatória mais severa que atinge o sistema nervoso central causando desmielinização da matéria branca e levando a morte em aproximadamente 4 anos após o início dos sintomas. Esse fenótipo é conhecido como ALD cerebral (CALD) (Aubourg & Wanders 2013).

Até dezembro de 2021, foram catalogadas 3401 variantes relacionadas à ALD, dentre elas, 948 prováveis patogênicas, 40 benignas e 249 de significado indeterminado (Mallack et al. 2021). No entanto, não há relação direta entre a severidade dos sintomas apresentados com o genótipo dos pacientes, de forma que mesmo gêmeos monozigóticos podem apresentar fenótipos discordantes (Korenke et al. 1996). Além disso, não se têm marcadores bioquímicos ou genéticos de prognóstico estabelecidos, de forma que não é possível prever quais indivíduos apresentarão acometimentos cerebrais (Honey et al. 2021).

Pacientes CALD diagnosticados logo no início dos sintomas podem ser tratados com transplante de células tronco hematopoiéticas para conter o avanço rápido da doença (Peters et al. 2004). Dessa forma, seria de grande benefício a descoberta de fatores modificadores com potencial para prognóstico. O emprego de biologia de sistemas foi sugerido como alternativa para pesquisas com EIM (Argmann et al. 2015). Nessa linha de pensamento, bons resultados foram obtidos com análises multi-ômicas aplicadas no estudo de ALD (Richmond et al. 2020), bem como no uso de redes de interações para a observação de patologias neurológicas de outros EIM (Silva & Matte 2022).

ALD é uma doença rara e de acometimento do sistema nervoso, o que dificulta a coleta de amostras adequadas para estudos de expressão. Por conta disso, nesse estudo foram incorporados conceitos e amostras de pesquisas anteriores, extraindo-se informações relevantes de forma eficiente. Conjuntos de dados públicos foram submetidos à análise de expressão diferencial e posteriormente os resultados de comparações entre CALD, AMN e controles tiveram suas interações observadas em uma rede.

### 3.2 Metodologia

#### 3.2.1 Análises transcriptômicas

Os conjuntos amostrais utilizados para essa pesquisa foram retirados dos bancos de dados públicos *Gene Expression Omnibus* (GEO), 4 estudos, e *ArrayExpress* (AE), um estudo. Todas as amostras são relativas a análises de transcriptômica de tecidos humanos. Detalhes dos estudos, incluindo tecidos e grupos analisados, estão apresentados na tabela 1. A análise foi feita em *software* Rstudio com os pacotes *limma* v.3.50.3 (Ritchie et al. 2015) para os dados de microarray e *edgeR* v. 3.36.0 (Chen, Lun & Smith 2016) para o de RNA-seq, tendo como objetivo observar a expressão diferencial de CALD vs. AMN e AMN vs. controle, quando possível, e de CALD vs. controle para todas as amostras. Os parâmetros envolveram filtragem por *False Discovery Rate* e *cutoff* de 0,05.

Tabela 1. Resumo sobre as amostras usadas no estudo

| Identificação | Método     | Tecido/Célula   | Fenótipos     | Quantidade | Replicatas |
|---------------|------------|---|---------------|------------|------------|
| GSE108012     | RNA-seq    | iPSCs induzidos para células endoteliais microvasculares do cérebro | Controle      | 3          | 3          |
|               |            |   | CALD          | 3          | 3          |
| GSE117647     | Microarray | fibroblastos  | Controle      | 2          | 8          |
|               |            |   | AMN           | 3          | 8          |
|               |            |   | CALD          | 4          | 8          |
| GSE85804      | Microarray | iPSCs induzidos para oligodendrócitos                               | Controle      | 2          | 3          |
|               |            |   | CALD          | 3          | 3          |
|               |            |   | AMN           | 4          | 3          |
| GSE34309      | Microarray | iPSCs   | Controle      | 3          | 8          |
|               |            |   | CALD          | 2          | 5          |
| GSE34308      | Microarray | Fibroblastos  | Controle      | 3          | 5          |
|               |            |   | CALD          | 2          | 5          |
| E-MEXP-3288   | Microarray | Substância branca cerebral  | Controle      | 13         | 1          |
|               |            |   | CALD infantil | 3          | 1          |
|               |            |   | AMN cerebral  | 8          | 1          |

GEO: *Gene Expression Omnibus*, <https://www.ncbi.nlm.nih.gov/geo/>.

AE: *ArrayExpress*. <https://www.ebi.ac.uk/arrayexpress/>.

### 3.2.2 Redes de interação

Listas contendo os genes diferencialmente expressos obtidos com a análise de transcriptômica foram submetidas ao programa STRING v.11.5 (Szklarczyk et al. 2021). O número de genes foi limitado aos 2000 com menor valor de p, levando em consideração que esse é o número máximo de nós aceitos. As fontes de interação selecionadas foram as obtidas a partir de experimentos, co-expressão, vizinhança e co-ocorrência e o nível de confiança foi imposto como alto (escore mínimo de 0,7). Visualizou-se até 50 interações primárias e 50 secundárias. As redes resultantes foram submetidas então ao software Cytoscape v.3.9.1, onde foram observados os clusters com o aplicativo MCODE v.2.0.2 impondo os parâmetros de *fluff* 0.1 e *node score cutoff* 0.2. com *haircut* (Bader & Hogue 2003).

Também foi realizada uma busca por genes hub em cada rede com o aplicativo cytoHubba v.0.1, na qual observou-se os 75 nós mais importantes de acordo com o parâmetro Maximal Clique Centrality (MCC) (Chin et al., 2014). Por fim, foi empregado também o aplicativo CentiScaPe v.2.2 para a verificação das propriedades *diameter*, *degree*, *closeness*, *stress* e *betweenness* (Scardoni, Tosadori & Laudanna 2014).

### 3.2.3 Análise de Enriquecimento

Visando filtrar os resultados mais relevantes obtidos com a análise de expressão diferencial, as listas contendo os genes diferencialmente expressos e seus respectivos p-valores para cada comparação foram processadas com o pacote pathfindR v.1.6.3. À essa análise foi aplicado o método de correção de Bonferroni, com limiar de 0,05 para o p-valor. Utilizou-se como banco de dados para a pesquisa de anotações o *Kyoto Encyclopedia of Genes and Genomes* (KEGG; Kanehisa & Goto 2000).

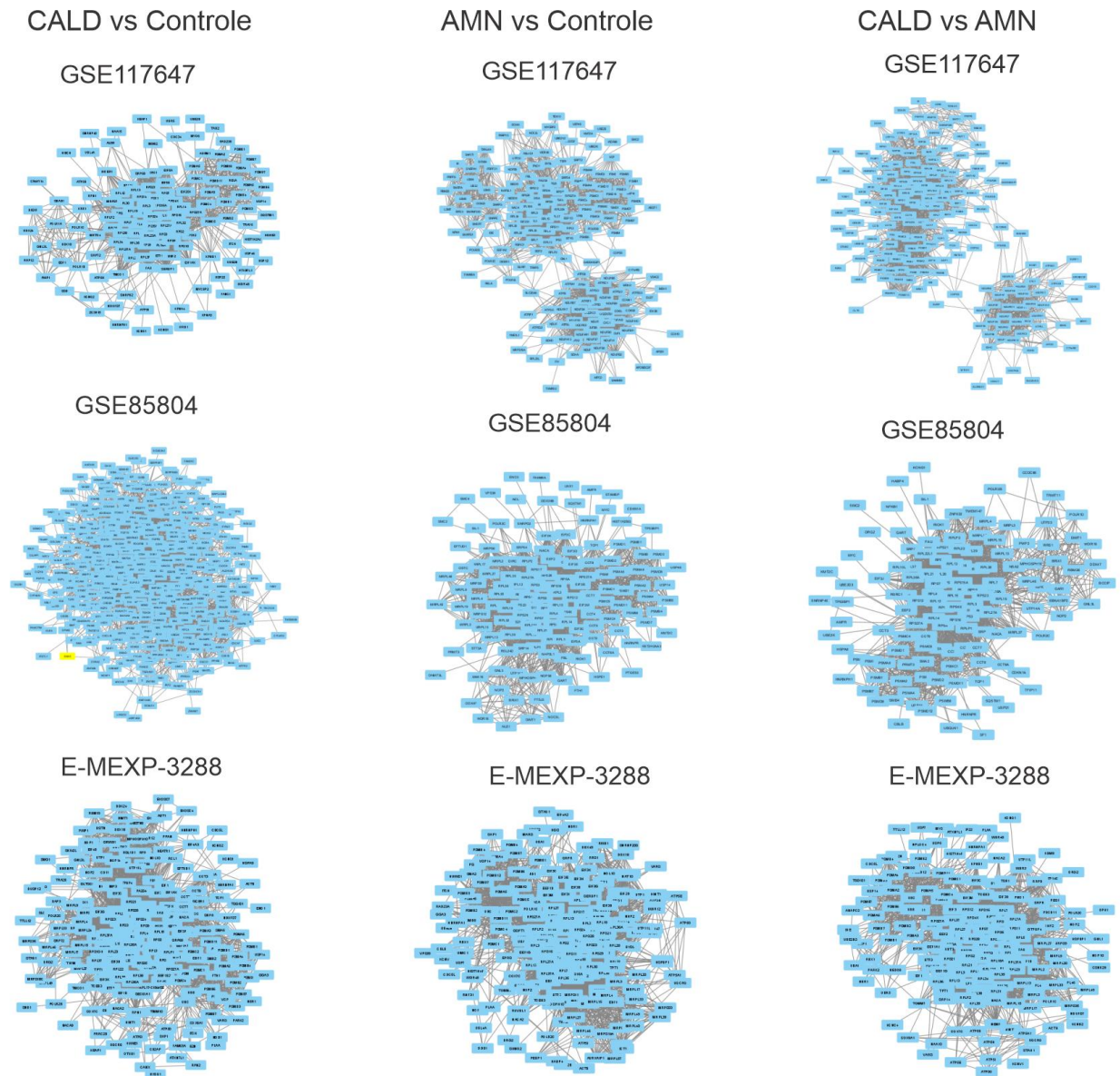
Listas de genes diferencialmente expressos, que constavam dos termos observados com os resultados de pathfindR, foram submetidas ao enriquecimento com o pacote clusterProfiler v.4.2.2 (Wu et al. 2021). Para essa etapa foram procuradas as anotações de processos biológicos fornecidas no banco de dados *Gene Ontology* (GO; Huaiyu et al. 2019).

## 3.3 Resultados

### 3.3.1 Clusters, genes hub e topografia das redes

Cada conjunto de dados foi analisado primeiro individualmente e, após, comparados. Em relação ao número de clusters, os grupos variaram de 46 a 21. A média de clusters foi de 35 (CALD vs Controle), 36 (AMN vs Controle) e 34 (CALD vs AMN) (Tabela suplementar 1). Os clusters são regiões da rede onde os nós encontram-se densamente ligados uns aos outros, indicando possivelmente complexos de moléculas. A representação dos maiores clusters dos grupos com amostras de CALD e AMN são observadas na figura 1. Tanto para GSE117647, quanto para E-MEXP-3288, os clusters de AMN vs controle e CALD vs controle parecem muito semelhantes.

**Fig 1** Os maiores *clusters* dos conjuntos de dados para as comparações de CALD, AMN e controles



Cada nó de uma rede biológica pode ser observado de acordo com as medidas de centralidade, que são parâmetros que ajudam a selecionar quais nós provavelmente são mais influentes para os processos biológicos representados na rede (Scardoni, Tosadori & Laudanna 2014). O grau (*degree*) indica a quantidade de interações que um nó realiza com outros nós vizinhos. Um nó com alto grau é chamado de hub e normalmente sustenta ligações de vizinhos de menor grau. Os resultados encontrados para cada grupo de dados foram comparados e genes em comum foram ordenados de acordo com o parâmetro *degree*. Apenas 4 genes foram comuns a todos os seis grupos de CALD vs controle (*RPL21*, *RPS11*, *RPS15A* e *RPS7*) e mais 26 eram comuns a pelo menos 5 grupos. Em AMN vs controle foram vistos 55 genes comuns a todos os grupos, enquanto que para CALD vs AMN foram observados 76. A fim de encontrar focos mais específicos para análise, procurou-se os genes com *degree* igual ou maior que os 5 primeiros genes dessas listas (tabela suplementar 2). Entre esses resultados mais relevantes foram encontrados três genes, *RPS27A*, *RPS11* e *RPS4X* em comum em todas as comparações (Tabela 2).

Tabela 2. Medidas de centralidade de genes em comum

| Genes           | Betweenness         | Closeness           | Degree   | Stress              |
|-----------------|---------------------|---------------------|----------|---------------------|
| CALD x Controle |                     |                     |          |                     |
| <i>RPS4X</i>    | 6,16E+01 – 9,09E+03 | 4,11E-04 – 1,12E-03 | 59 – 115 | 2,44E+03 – 3,97E+05 |
| <i>RPS27A</i>   | 4,30E+04 – 1,15E+05 | 4,26E-04 – 1,39E-03 | 83 – 118 | 1,76E+05 – 1,96E+06 |
| <i>RPS11</i>    | 6,16E+01 – 1,91E+04 | 1,94E-04 – 1,12E-03 | 50 – 115 | 2,44E+03 – 2,75E+05 |
| <i>RPS9</i>     | 6,16E+01 – 4,21E+03 | 4,06E-04 – 1,12E-03 | 54 – 106 | 2,44E+03 – 2,42E+05 |
| AMN x Controle  |                     |                     |          |                     |
| <i>RPS27A</i>   | 3,97E+04 – 7,04E+04 | 4,51E-04 – 9,75E-04 | 75 – 131 | 3,72E+05 – 1,36E+06 |
| <i>RPS11</i>    | 3,87E+03 – 8,70E+03 | 4,20E-04 – 9,55E-04 | 51 – 115 | 1,39E+05 – 3,33E+05 |
| <i>RPS4X</i>    | 1,74E+03 – 7,11E+03 | 4,19E-04 – 9,47E-04 | 54 – 84  | 7,09E+04 – 2,64E+05 |
| CALD vs. AMN    |                     |                     |          |                     |
| <i>RPS27A</i>   | 3,83E+04 – 7,36E+04 | 5,07E-04 – 9,35E-04 | 80 – 117 | 6,60E+05 – 1,29E+06 |
| <i>RPS11</i>    | 4,49E+03 – 8,66E+03 | 4,67E-04 – 9,18E-04 | 63 – 97  | 1,65E+05 – 2,51E+05 |

Genes em comum entre os cinco primeiros resultados dos diferentes conjuntos de dados ordenadas por *degree*. Para cada parâmetro estão representados respectivamente o menor e o maior valor obtido entre as diferentes amostras.

É possível estimar a relevância desses nós nas redes comparando seus valores com os parâmetros médios (Tabela 3). Dos três genes, *RPS27A* destacou-se em três parâmetros diferentes: *betweenness*, *closeness* e *stress*. Esse gene obteve o maior *degree* em E-MEXP-3288 em CALD vs controle (*degree*: 118), AMN vs controle (*degree*: 131) e CALD vs AMN (*degree*: 117), enquanto os *degrees* médios nesse grupo amostral para essas respectivas comparações foram aproximadamente 16, 17 e 14. Em relação à medida *betweenness*, que indica o quão crucial é um nó intermediário para a interação de outros dois nós, os valores mais altos são considerados mais relevantes. Nesse caso o gene destaca-se, novamente, principalmente na comparação CALD vs controle, já que seu maior valor registrado (1,15E+05) foi acima da média (1,86E+03 para GSE34309). Já o *stress* complementa a análise mostrando quais nós estão mais presentes nos caminhos mais curtos que ligam dois outros nós, sendo novamente os valores maiores os mais significativos. O valor mais alto para *RPS27A* (1,96E+06 em CALD vs controle) pode indicar que esse é um nó intermediário importante se for comparado à média (3,47E+04 para GSE34309). Por fim, o parâmetro de *closeness* diz respeito à proximidade de um nó em relação a todos os outros nós da rede e no caso dos genes observados, todos apresentam valores baixos em relação à média, de forma que pode ser sugerido que esses são nós que possivelmente agem na intersecção dessa rede com outras.

Tabela 3. Médias das medidas de centralidade obtidas com CentiScaPe

|                     |             | Betweenness | Closeness | Degree | Stress   | Nodes | Edges |
|---------------------|-------------|-------------|-----------|--------|----------|-------|-------|
| CALD vs<br>Controle | GSE117647   | 1,84E+03    | 0,13      | 10,40  | 1,99E+04 | 710   | 3702  |
|                     | GSE85804    | 4,54E+03    | 0,00      | 11,10  | 4,54E+04 | 1701  | 9476  |
|                     | GSE34308    | 9,29E+02    | 0,16      | 11,50  | 1,04E+04 | 515   | 2957  |
|                     | GSE34309    | 1,86E+03    | 0,12      | 9,50   | 3,47E+04 | 748   | 3554  |
|                     | GSE108012   | 5,86E+02    | 0,21      | 16,40  | 2,57E+03 | 462   | 3784  |
|                     | E-MEXP-3288 | 2,49E+03    | 0,09      | 15,80  | 5,69E+04 | 905   | 7143  |
| AMN vs<br>Controle  | GSE117647   | 1,74E+03    | 0,15      | 10,80  | 1,90E+04 | 668   | 3612  |
|                     | GSE85804    | 8,78E+02    | 0,19      | 12,60  | 2,78E+04 | 549   | 3464  |
|                     | E-MEXP-3288 | 2,46E+03    | 0,08      | 16,90  | 4,74E+04 | 902   | 7613  |
| CALD vs<br>AMN      | GSE117647   | 2,02E+03    | 0,12      | 8,93   | 3,09E+04 | 739   | 3300  |
|                     | GSE85804    | 9,57E+02    | 0,19      | 11,60  | 2,99E+04 | 537   | 3127  |
|                     | E-MEXP-3288 | 2,13E+03    | 0,10      | 14,40  | 3,82E+04 | 804   | 5801  |

A procura por genes hub foi feita com aplicativo específico para esse propósito e usou o parâmetro MCC, que calcula uma pontuação para estabelecer uma ordem de importância dos genes hub de cada amostra. Na maioria dos conjuntos de dados foram observados muitos genes com pontuações igualmente altas, sendo que GSE108012 foi o que demonstrou a maior quantidade de itens com máxima pontuação, mais precisamente 74 genes. Dessa forma, foi estabelecido que o aplicativo retornasse uma lista com 75 genes hub para cada conjunto de dados. Ao comparar os resultados entre si foram observados 4 genes em comum para todos os grupos de CALD vs controle, 26 para AMN vs controle e 24 para CALD vs AMN (Tabela suplementar 3).

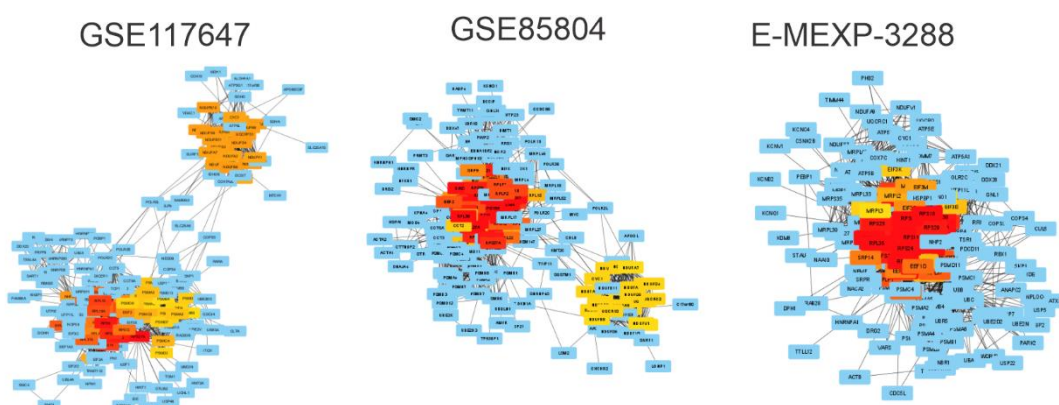
A maioria desses genes codifica proteínas ribossômicas (PRs), com exceção de três genes, *EEFIG*, *EEF2* e *SRP14*, os quais também codificam proteínas relacionadas à síntese proteica. Dada a semelhança de quase todos esses hubs tanto em questões de função quanto em pontuação (por MCC), é preferível sua análise como um conjunto. No entanto, ainda assim pode ser ressaltado que entre as diferentes comparações há três hubs comuns: *RPS11*, *RPS15A* e *RPS7*. Sugere-se, então, que esses sejam genes importantes para os dois fenótipos de ALD observados nesse estudo. Outros 3 genes foram vistos em comum exclusivamente nas comparações de CALD vs AMN: *RPS2*, *RPS21* e *SRP14*. Outro gene *RPL21* foi observado em comum somente entre os grupos de CALD vs controle.

Todos os genes em comum entre todos os grupos amostrais para cada comparação encontrados na análise anterior com CentiScaPe foram observados da mesma forma nessa análise. Os genes *RPS11*, *RPS15A*, *RPS7*, *RPS2*, *RPS21*, *SRP14* e *RPS27A* foram encontrados no maior cluster da análise com MCODE para todos os grupos

de CALD vs AMN. Todos esses resultados em conjunto complementam-se e aumentam a confiança sobre a inferência da importância desses hubs.

A visualização desses genes em redes permite observar que alguns formam dois sub clusters, onde os genes mais relevantes encontram-se normalmente concentrados em apenas um. Os conjuntos de dados com quantidades maiores de genes com pontuação máxima mostraram todos os genes em um mesmo cluster (Figura 2).

**Fig 2** Clusters contendo genes hub para os conjuntos amostrais de CALD vs AMN



Exemplos dos resultados com Cytohubba no qual são vistos subclusters em GSE117647 e GSE85804 e apenas um cluster em E-MEXP-3288. Cada nó nas redes é a representação de um gene. Aqueles considerados hubs encontram-se na cor vermelha (maior pontuação), laranja (pontuação intermediária) e amarela (menor pontuação).

### 3.3.2 Resultados dos enriquecimentos

Todas as amostras submetidas à análise de enriquecimento com pathfindR geraram como resultado uma lista de vias metabólicas e genes, com exceção das amostras de GSE34308. Dentre os genes e vias que mais apareceram nessas análises estão aqueles relacionados à inflamação, ao estresse oxidativo e a outras doenças neurodegenerativas. Algumas vias destacam-se por serem observadas de forma frequente dentre os três termos mais relevantes, de acordo com o p-valor, tanto em AMN como em CALD, como é o exemplo das vias hsa03010 (Ribossomo), hsa05171 (Coronavirus disease) e hsa03050 (Proteossomo).

Observando as comparações de AMN e CALD em relação aos controles e CALD em relação a AMN, foi feita uma intersecção dos resultados, porém somente poucas vias ou nenhuma eram comuns a todos. Nesse caso é importante ressaltar que o número de resultados encontrados pela análise de enriquecimento variou muito entre os conjuntos de dados, dificultando a coleta de mais vias relevantes a partir da intersecção desses resultados (Tabela 4).

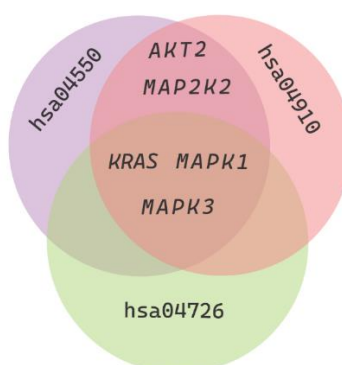
Tabela 4. Quantidades de vias encontradas com análise de enriquecimento para cada conjunto de dados em cada comparação

|             | AMNxControle | CALDxControle | CALDxAMN |
|-------------|--------------|---------------|----------|
| GSE117647   | 2            | 6             | 188      |
| GSE85804    | 23           | 38            | 4        |
| GSE34308    | -            | 0             | -        |
| GSE34309    | -            | 1             | -        |
| GSE108012   | -            | 229           | -        |
| E-MEXP-3288 | 202          | 219           | 212      |

Entre as comparações AMN vs. controle há 21 vias em comum entre ao menos dois grupos. Em CALD vs. controle, foram encontradas ao todo 36 vias comuns entre três grupos e apenas uma comum a quatro dos cinco grupos que obtiveram resultados para essa análise. Essa via é identificada pela ID hsa04612 e representa a via de processamento e apresentação de antígenos. Tanto para AMN vs. controle quanto para CALD vs. controle, não foram encontradas vias comuns a todos os grupos. A comparação dos grupos CALD vs. AMN gerou um resultado de 175 vias comuns a dois grupos e 3 vias presentes em todos os conjuntos: via de sinalização de pluripotência de células tronco (hsa04550), sinapse serotoninérgica (hsa04726) e via de sinalização da insulina (hsa04910). Dentre essas 175 vias foram observadas ainda vias metabólicas de glicólise (hsa00010), ciclo do citrato (hsa00020), fosforilação oxidativa (hsa00190), ribossomo (hsa03010), sinalização de TGF-beta (hsa04350) e sinalização do receptor de células T (hsa04660), de forma semelhante à descrita em outro estudo (Schlüter et al. 2012).

Realizou-se uma intersecção entre os genes presentes em cada uma das vias comuns aos três entre os três grupos de CALD vs AMN: hsa04550, hsa04726 e hsa04910. Em nenhuma foi visto um gene em comum aos três grupos, porém todas as vias contaram com alguns genes comuns a pelo menos dois grupos, sendo um total de 12 genes em hsa04550 (via de sinalização de pluripotência de células tronco), 10 em hsa04726 (sinapse serotoninérgica) e 14 em hsa04910 (via de sinalização da insulina). Na união desses 26 genes, ao menos 3 eram observados mais de uma vez, sendo então comuns entre as três vias, além de também serem comuns a ao menos dois conjuntos de dados (Figura 3).

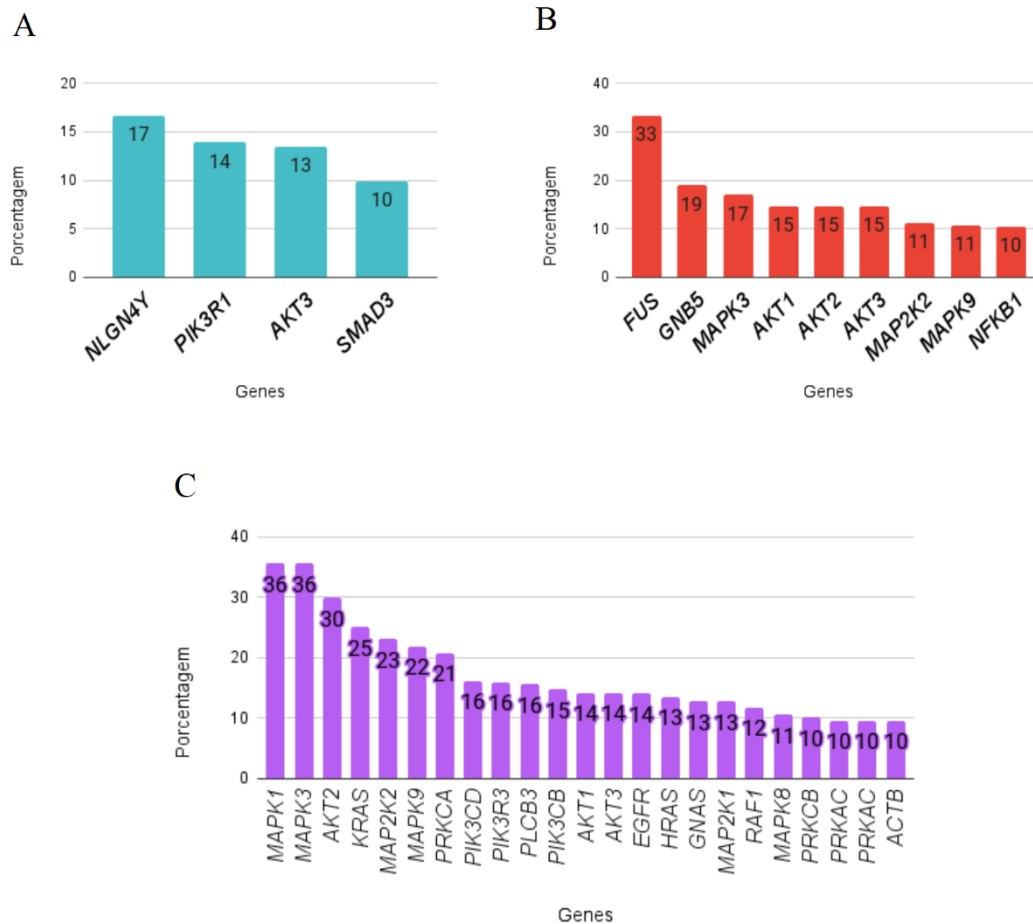
**Fig 3** Genes em comum entre as vias de CALD vs AMN presentes nos três conjuntos de dados





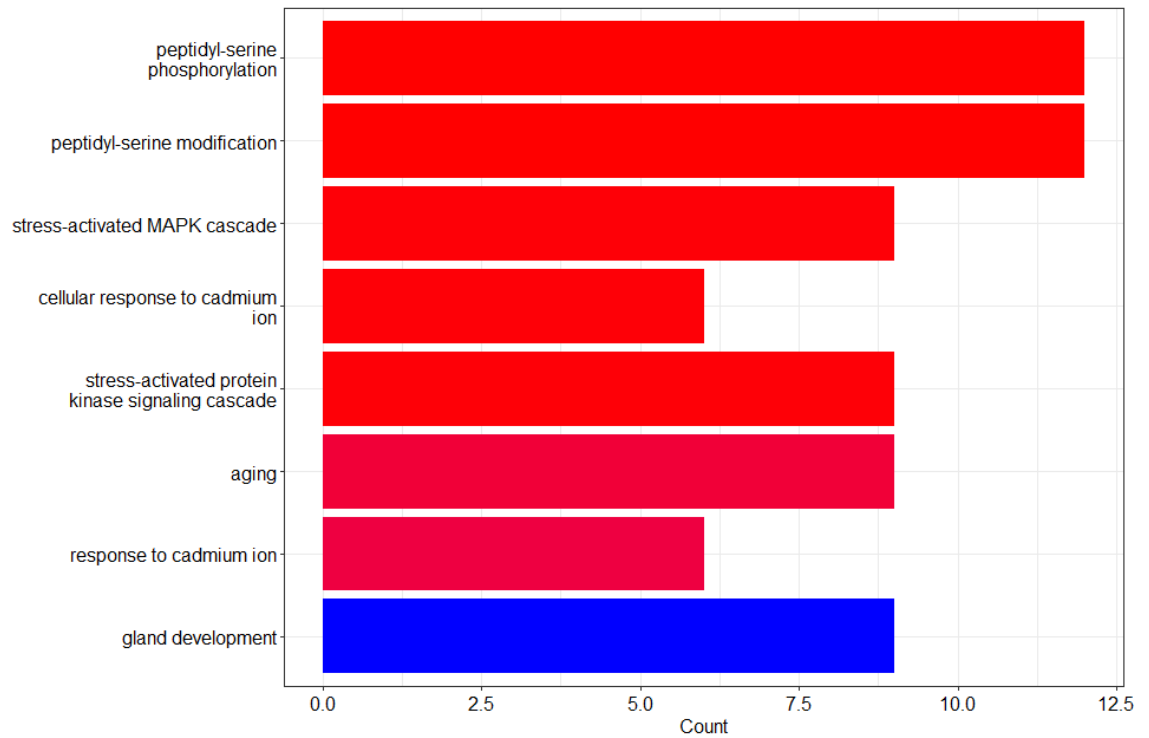
Para identificar quais os genes mais relevantes para os sintomas de ALD em cada uma das 175 vias, observamos quais os genes estavam presentes no maior número de vias. Inicialmente foi calculado o percentual dos genes presentes em todas as vias para cada conjunto de dados e depois observou-se a média de todos os conjuntos para cada gene. A Figura 4 mostra os genes presentes em ao menos 10% das vias. (Figura 4).

**Fig 4** Genes presentes em ao menos 10% das vias de todos os conjuntos amostrais para cada comparação. A) CALD vs. controle. B) AMN vs. controle. C) CALD vs. AMN.



Uma análise de enriquecimento desses genes com anotações de processos biológicos do banco de dados GO foi realizada para CALD vs AMN (Figura 5). Nesse caso não se repetem anotações relacionadas a processos exclusivamente associados a outras doenças como câncer ou infecções como o que foi observado ao procurar pela participação desses genes em vias metabólicas. Aqui destacam-se processos envolvidos com a sinalização celular por meio de fosforilação, como pode ser observado nas três primeiras anotações mais frequentes.

**Fig 5** Processos biológicos mais frequentes entre os grupos amostrais de CALD vs AMN



Também procurou-se observar se esses genes se encontram expressos nas mesmas vias entre os grupos. Em GSE117647 tanto *MAPK1* quanto *MAPK3* foram encontrados em 106 vias. Na lista de vias da análise de enriquecimento ordenada por p-valor, foram procuradas as 10 primeiras vias onde os genes *MAPK1* e *MAPK3* aparecem. O mesmo foi feito para os resultados de E-MEXP-3288 (Tabela 5).

Tabela 5. 10 Primeiras vias na qual *MAPK1* e *MAPK3* aparecem em CALD vs AMN

| GSE117647 |   | E-MEXP-3288 |                                |
|-----------|---|-------------|--------------------------------|
| ID        | Descrição   | ID          | Via                            |
| hsa05020  | Prion disease                                     | hsa05171    | Coronavirus disease - COVID-19 |
| hsa05205  | Proteoglycans in cancer                           | hsa05205    | Proteoglycans in cancer        |
| hsa04012  | ErbB signaling pathway                            | hsa04014    | Ras signaling pathway          |
| hsa05208  | Chemical carcinogenesis - reactive oxygen species | hsa05219    | Bladder cancer                 |
| hsa05213  | Endometrial cancer                                | hsa04010    | MAPK signaling pathway         |
| hsa04210  | Apoptosis   | hsa05215    | Prostate cancer                |
| hsa05230  | Central carbon metabolism in cancer               | hsa04012    | ErbB signaling pathway         |
| hsa04014  | Ras signaling pathway                             | hsa04510    | Focal adhesion                 |
| hsa05203  | Viral carcinogenesis                              | hsa04350    | TGF-beta signaling pathway     |
| hsa04666  | Fc gamma R-mediated phagocytosis                  | hsa05160    | Hepatitis C                    |

Em GSE117647, em todas as 10 vias os genes *MAPK1*, *MAPK3*, *PIK3CD* e *PIK3R3* estão presentes. *HRAS* e *KRAS* também encontram-se frequentemente juntos e não participam das duas mesmas vias: hsa05020 e hsa04666. Em E-MEXP-3288, além de *MAPK1* e *MAPK3*, todas as vias com exceção de hsa04350 e hsa05171 apresentaram também os genes *EGFR*, *MAP2K1* e *RAF1*. Em GSE85804 nenhum desses genes estava presente em nenhuma das vias.

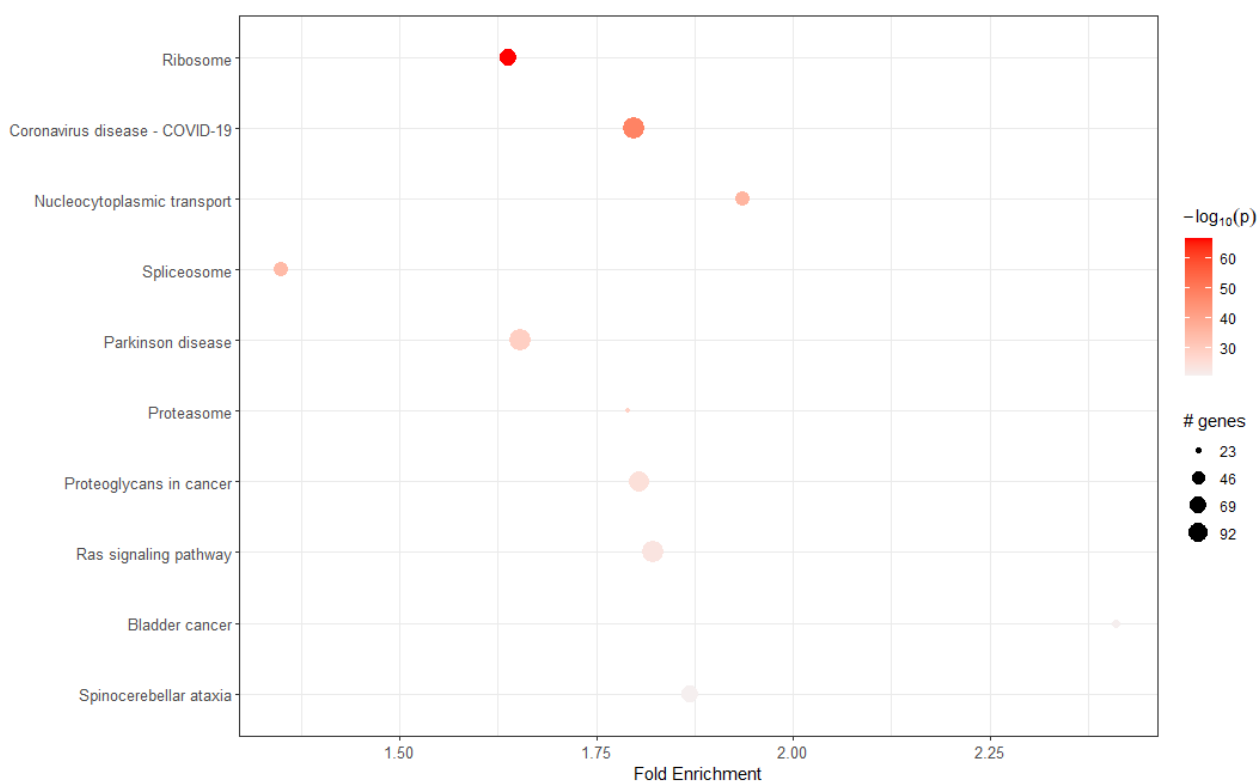
### 3.3.3 E-MEXP-3288

O conjunto de dados E-MEXP-3288 contém amostras de matéria branca de crianças que desenvolveram desmielinização ainda na infância e de adultos com AMN que passaram a apresentar também os sintomas neurais. Dessa forma, julgou-se interessante analisar esses dados de forma isolada, principalmente em relação a comparação de CALD vs AMN, visto que os dois grupos de comparação desse conjunto amostral poderiam se encaixar no fenótipo CALD.

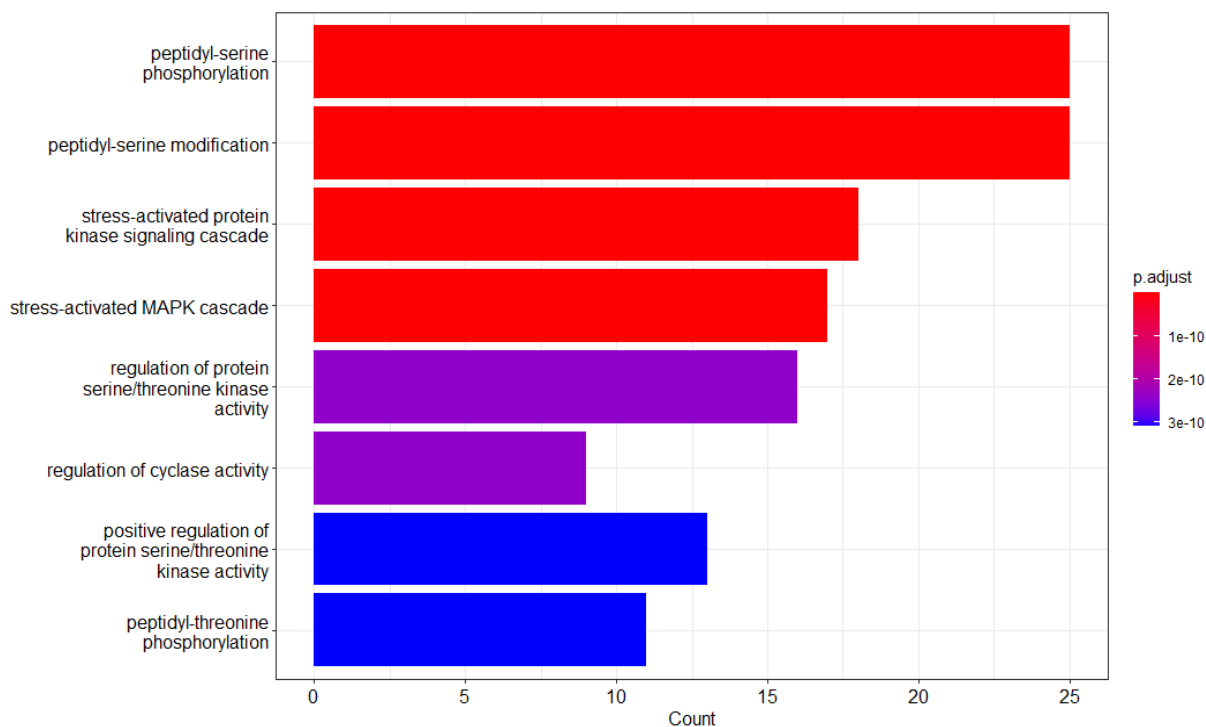
Levando em consideração que o gene com o maior *degree* nesse grupo era de *RPS27A* com *degree* de 117, procurou-se observar os genes que tivessem obtido ao menos metade desse valor (*degree* >57). Dentre esses genes, a maioria é composta de PRs, com sete exceções: *TPT1*, *EIF3M*, *NDUFAB1*, *EEF1G*, *EEF2*, *EEF1A1* e *FAU*. Dos 75 genes hub retornados pela análise com cytohubba, 31 encontram-se na primeira posição de importância e todos codificam PRs, de forma semelhante ao que foi achado em outros grupos de amostras e também nas análises com outros métodos. O gene *RPL21* é um hub comum a todos os conjuntos de dados nos resultados de CALD vs controle, porém só é visto como hub em CALD vs AMN nas amostras de E-MEXP-3288.

Dentre as dez vias mais relevantes na análise de enriquecimento, encontra-se novamente o ribossomo, bem como outras doenças neurológicas, como Parkinson e ataxia espinocerebelar (Figura 6). Uma união de todos os genes diferencialmente expressos envolvidos nessas vias para CALD vs AMN mostrou os genes *RPS6* e *UBA52* em comum.

**Fig 6** 10 vias metabólicas mais relevantes para CALD vs AMN nas amostras de E-MEXP-3288



A união de todos os genes presentes em todas as vias resultou em 75 genes presentes em pelo menos 10% de todas as vias. Esses termos foram então enriquecidos de acordo com anotações de GO sobre seus processos biológicos. De um total de oito termos observados nessa análise e na anterior, que contemplava os outros grupos amostrais de CALD vs AMN, ocorreram quatro anotações comuns. Nos dois casos, mas principalmente aqui, destacam-se processos relacionados à fosforilação de serinas e treoninas, sugerindo envolvimento de proteínas quinases.

**Fig 7** Processos biológicos frequentes em CALD vs AMN para as amostras de E-MEXP-3288

### 3.4 Discussão

A análise por biologia de sistemas de um conjunto de amostras de pacientes de ALD com CALD e AMN revelou o papel central de genes que codificam proteínas ribossômicas (PRs), que atuam como hubs nos diferentes conjuntos de dados. A alteração consistente entre os diferentes grupos de amostras na expressão desses genes pode indicar desequilíbrios na síntese proteica. Já foram descritas alterações em vias ribossômicas em outras doenças neurodegenerativas como Alzheimer, onde foram observadas PRs superexpressas no córtex frontal de pacientes com desenvolvimento rápido da doença (Garcia-Esparcia et al. 2017).

Alterações na síntese proteica são comuns também no envelhecimento, que foi um dos termos encontrados entre as anotações de processos biológicos para genes de CALD vs AMN. O envelhecimento é um processo natural no qual se observa a diminuição de oligodendrócitos e, conseqüentemente, da matéria branca, provavelmente por conta de comprometimentos à barreira hematoencefálica que torna o tecido suscetível à inflamação (Skariah & Todd 2021). No entanto, nesses casos, a síntese proteica encontra-se reduzida, bem como a expressão de PRs, de modo contrário ao que foi encontrado nas amostras de ALD do presente estudo. As PRs encontram-se superexpressas tanto nas amostras de ALD em relação aos controles, quanto em CALD vs AMN. Se por um lado esse resultado confere com o que é observado em outras doenças neurodegenerativas, não se pode dizer o mesmo do envelhecimento mesmo havendo semelhança entre a diminuição gradual da matéria branca. Dessa forma, não se pode concluir uma correlação óbvia estabelecida entre diminuição ou aumento da expressão dessas proteínas em conjunto e a desmielinização.

Torna-se interessante, portanto, observar as PRs fora do contexto estrutural do ribossomo e explorar suas funções além da síntese proteica. Essas proteínas podem atuar na regulação de sua própria expressão gênica, bem como na regulação da expressão de outros genes e também na reparação do DNA mediante interação com sinalizadores que podem desacelerar o ciclo celular e provocar apoptose (Bhavsar, Makley & Tsonis 2010).

*RPS15A*, *RPS7* e *RPS27A* foram hubs comuns encontrados tanto ao comparar ALD com controles quanto ao comparar CALD e AMN entre si. Esses genes codificam, respectivamente, as proteínas S15a, S7 e S27a da subunidade 40S do ribossomo (GeneCards Symbols: *RPS15A*, *RPS7*, *RPS27A*).

Em modelo animal *zebrafish*, mutações nos dois primeiros genes (*RPS15A*, *RPS7*) podem levar a tumores malignos na bainha de nervos periféricos (Amsterdam et al. 2004). Em humanos, *RPS7* é considerado um alvo de oncogenes, por estar relacionado com inibição de crescimento (Pandey & Kumar 2015), enquanto *RPS15A* é considerado um oncogene em si (Ning et al. 2020). Já o gene *RPS27A*, foi descrito recentemente como um gene *switch* que atua sobre outro gene, *IL-18*, na micróglia, servindo de gatilho para neurodegeneração em cinco doenças diferentes (Khayer et al. 2020). Ainda que *IL-18* não tenha sido verificado de forma relevante e nem em comum entre as amostras do presente estudo, não se pode descartar a atuação de *RPS27A* sobre a expressão de outros genes possivelmente relevantes para ALD, dados seus parâmetros de centralidade em diversos dos grupos amostrais.

*RPL21* codifica a única PR da subunidade maior do ribossomo encontrada de forma relevante nesse estudo. O gene foi observado especificamente em CALD vs controle e nas amostras de E-MEXP-3288, incluindo CALD vs AMN. Uma pesquisa observou esse gene superexpresso ao suprimir *KRAS*, um oncogene, em células de câncer pancreático (Li et al. 2012). Os mesmos pesquisadores silenciaram, então, *RPL21* e observaram alterações na expressão de um fator de transcrição importante para a proliferação celular tanto em células de câncer quanto em células pancreáticas saudáveis (Li et al. 2020). No entanto, não foram esclarecidos os mecanismos pelos quais *RPL21* encontrava-se superexpresso, sugerindo uma regulação indireta de *KRAS* sobre *RPL21*. Conforme discutido anteriormente, é possível que existam alterações na sinalização celular comuns tanto em câncer quanto em ALD que podem estar relacionadas com a elevação de *KRAS*, encontrado em 25% das vias de CALD vs AMN, bem como de *RPL21*. Além disso, *RPL21* poderia apresentar no contexto de CALD uma maior proliferação celular.

*RPS2*, *RPS21* e *SRP14* foram encontrados como genes hub, exclusivamente entre a comparação CALD vs AMN. Não foram encontradas na literatura relações já descritas entre esses genes e algum dos fenótipos de ALD. No entanto, a expressão de *RPS2*, bem como *RPS27A*, encontrou-se alterada em camundongos modelo e fibroblastos de indivíduos com Ataxia de Friedreich (McMackin et al. 2019). Em outro estudo, no qual camundongos modelo para glaucoma injetados com anticorpos alfa-sinucleína, apresentaram níveis mais altos da proteína codificada por *SRP14*, os autores atribuem a alteração a uma anormalidade na função de exportação de proteínas (Teister et al. 2017).

Em relação às vias comuns e relevantes para ALD, encontram-se, além das relacionadas à síntese proteica (hsa03010 - Ribossomo), vias relacionadas a infecções (hsa04612 - via de processamento e apresentação de antígenos, hsa05171- COVID-19), à sinalização celular (hsa04012 - Via de sinalização da ErbB) e ao estresse oxidativo (hsa05208 - carcinogênese química - espécies reativas de oxigênio).

Um estudo semelhante, no qual foi analisado o perfil de expressão de camundongos modelo para ALD em diferentes meses de idade encontrou vias que também foram observadas em algumas amostras da presente pesquisa, como via do ribossomo (hsa03010), via metabólica da glicólise (hsa00010), ciclo do citrato (hsa00020), fosforilação oxidativa (hsa00190) e via do proteossomo (hsa03050). Os genes envolvidos nesses processos encontravam-se superexpressos em camundongos jovens (5 meses) antes da apresentação dos sintomas neurológicos e pouco expressas com o avanço da idade (12 meses). Os autores sugerem que essas alterações estão relacionadas com mecanismos compensatórios insustentáveis a longo prazo (Schlüter et al. 2012). Não é possível dizer se esse padrão de expressão se repete nesse estudo, pois não há como comparar as amostras ao longo de um

período. No entanto, as semelhanças observadas podem ser interpretadas como evidências em favor dos achados anteriores.

Outro processo metabólico, a via de sinalização da insulina (hsa04910), foi descrita no mesmo estudo como uma assinatura de ALD, por apresentar-se desregulada em todas as idades nos camundongos, bem como em pacientes infantis e adultos com ALD. Além disso, os autores ainda demonstraram uma relação de causalidade entre o acúmulo de VLCFA e a dessensibilização da insulina, por conta de danos causados a proteínas como os receptores de insulina. Essa informação converge com o achado do presente estudo de que a via hsa04910 encontra-se alterada de forma relevante entre todos os grupos amostrais da comparação CALD vs AMN e também entre alguns dos grupos de CALD vs controle e AMN vs controle.

Duas outras vias observadas nesse estudo, hsa03050 (Proteossomo) hsa05208 (carcinogênese química - espécies reativas de oxigênio), podem guardar uma relação entre si. Os VLCFA são considerados causadores de distúrbios nas mitocôndrias e geram espécies reativas de oxigênio (ROS, do inglês *reactive oxygen species*) em astrócitos de ratos modelo para ALD (Kruska et al. 2015). Além disso, pacientes humanos são suscetíveis a lesões oxidativas decorrentes da peroxidação lipídica. Esses acometimentos são descritos tanto no córtex das adrenais quanto em células neurais como astrócitos e micróglia (Powers et al. 2005). Uma das consequências da peroxidação é o dano às proteínas da membrana (Deon et al. 2016). Outro estudo observou em amostras de ALD a presença de alterações na expressão de genes relacionados com a maquinaria do proteossomo, um complexo proteico que se apresenta em resposta a proteínas danificadas e descreveu a migração desses complexos para a mitocôndria quando as células eram expostas a altas concentrações de VLCFA (Launay et al. 2013). Dessa forma, justifica-se a alteração da via do peroxissomo no presente estudo como uma resposta à toxicidade dos VLCFA e consequentes ROS. Ensaios *in vitro* observando a ação de antioxidantes em pacientes com ALD já foram realizados e mostraram bons resultados sobre o desequilíbrio oxidativo (Marchetti et al. 2017), de forma que esse pode ser uma terapia auxiliar aos tratamentos já estabelecidos.

Vias relacionadas a infecções como COVID-19 (hsa05171), doença do Prion (hsa05020), carcinogênese viral (hsa05203), via de processamento e apresentação de antígenos (hsa04612) e hepatite C (hsa05160) foram apontadas como relevantes pelas análises empregadas, bem como vias associadas a cânceres ou carcinogênese (hsa05215 - câncer de próstata, hsa05213 - câncer endometrial, hsa05230 - metabolismo central do carbono em câncer). No entanto, sugere-se que esses resultados não estão relacionados com nenhuma dessas doenças em si, e sim com a sinalização de resposta do corpo a infecções e inflamações. Um indício que torna essa suposição plausível é o resultado obtido ao analisar os genes mais frequentes de CALD vs AMN de acordo com sua participação em processos biológicos. Esses genes não apresentam em sua maioria nenhuma função específica em casos de infecções ou câncer, mas sim na regulação de sinalização de diversas funções celulares que podem ser comuns a diferentes processos patogênicos.

Ao menos 6 genes codificantes de MAP quinases (MAPK) foram encontrados em mais de 10% das vias de CALD vs AMN. As proteínas dessa família se caracterizam por serem quinases do tipo serina/treonina e integram diversas vias de sinalização, estando envolvidas em processos como proliferação, diferenciação e apoptose. Dentre as MAPK há ERKs, que compõem as sinalizações estimuladas por fatores de crescimento celulares, MAP2Ks, quinases que fosforilam outras MAPK, e JNKs, que participam de respostas ao estresse (Yue & López, 2020). *MAPK3* e *MAPK1* codificam proteínas da subfamília ERK1/2 (GeneCards Symbols: MAPK1, MAPK3), *MAP2K2* e *MAP2K1* codificam quinases MAP2K, também chamadas de MEK (GeneCards Symbols:

MAP2K2, MAP2K1) e *MAPK8* e *MAPK9*, por sua vez, são consideradas JNKs (GeneCards Symbols: MAPK9, MAPK8). Além desses, foram também encontrados genes que codificam as proteínas RAS (GeneCards Symbols: HRAS, KRAS), RAF (GeneCards Symbols: RAF1) AKT (GeneCards Symbols: AKT1, AKT2, AKT3), PI3K (GeneCards Symbols: PIK3CD, PIK3CB, PIK3R3) e receptor de fator de crescimento epidermal (EGFR; GeneCards Symbol: EGFR).

A proteína RAS se ativa por meio de receptores tirosina quinase na membrana das células e responde a neurotrofina no sistema nervoso central ativando diversas vias, incluindo PI3K-AKT e MEK-ERK (Rai et al. 2019). RAS ativa RAF, que por sua vez ativa MEK, que fosforila e ativa ERK. É conhecido que a cascata RAS-RAF-MEK-ERK tem um papel importante para a proliferação celular em cânceres e já foi constatada a hiperativação de ERK nessas doenças. As JNKs, podem ser ativadas por diversos tipos de estresse celular, como por exemplo o oxidativo e também encontram-se hiperativas em cânceres (Kim & Choi 2015). RAS também pode ativar PI3K, que, por sua vez, ativa AKT. A cascata RAS-PI3K-AKT induz a sobrevivência em neurônios, talvez pelo bloqueio de JNK-p53-bax, que é a via apoptótica primária em células neurais (Rai et al. 2019).

As ERK também parecem desempenhar função importante na patogênese de diversas doenças neurodegenerativas, porém de forma controversa. Essas quinases ativam-se na presença de ROS e exercem efeito neuroprotetor contra o estresse oxidativo, no entanto sua atividade de forma crônica promove a morte celular (Rai et al. 2019). Shah e colegas (2017), também sugeriram uma relação entre as vias das MAPK e a resposta a proteínas mal dobradas, desencadeada pelo estresse crônico do retículo endoplasmático (RE) na doença do Prion. A associação de todos esses genes com vias relacionadas a cânceres e à doença do Prion, provavelmente ocorre por conta dessas alterações de sinalização comuns, nas quais são vistas consequências semelhantes causadas por diferentes agentes etiológicos.

Outros três genes observados entre os mais frequentes de CALD vs AMN podem estar relacionados a uma função neuroprotetora: *PRKCA* e *PRKCB* (GeneCards Symbols: PRKCA, PRKCB), que codificam proteínas quinases C (PKC, do inglês *protein kinase C*), e *PLCB3* (GeneCards Symbol: PLCB3), que codifica uma fosfolipase C (PLC, do inglês *phospholipase c*). Estudos recentes sugerem que a via PKA-PLC-PKC-MEK-ERK está envolvida na transdução de sinal de diferenciação em modelo neural estimulado por octadecaneuropeptídeo, substância essa que exerce efeito neuroprotetor em situações de estresse oxidativo (Masmoudi-Kouki et al. 2018; Namsí et al. 2019). Tanto *PRKCA* quanto *PRKCB* encontram-se com expressão diminuída em 4 dos 6 grupos de CALD vs controle. Mesmo sem relevância estatística após correção de múltiplos testes, esse pode ser um indício de que a manifestação de CALD é impulsionada em parte por falhas em mecanismos neuroprotetivos.

O gene mais frequentemente observado nas vias de AMN vs controle foi *FUS* (GeneCards Symbol: FUS; do inglês *fused in sarcoma*) que codifica uma proteína relacionada com regulação de expressão gênica, manutenção da integridade genômica e processamento de micro RNAs e RNAs mensageiros. Curiosamente, esse gene não foi frequentemente visto nas outras comparações. Mutações patogênicas em *FUS* são uma das principais causas de esclerose amiotrófica lateral e estudos sobre essa doença têm mostrado que esse gene apresenta função importante para a resposta de reparo de DNA danificado (Naumann et al. 2018). Levando em consideração o estresse oxidativo descrito em ALD (Kruska et al. 2015) e que esse processo pode ser responsável por danos ao DNA (Deon et al. 2016), seria possível supor que a superexpressão desse gene em AMN exerça uma função protetiva, que talvez seja menos intensa em CALD. Um indício para essa hipótese é de que a expressão de *FUS* se encontra levemente aumentada em dois dos três grupos de AMN vs controle, com significância estatística.



Dentre os genes observados de acordo com as medidas de centralidade da rede de CALD vs AMN nas amostras de tecido nervoso, encontrava-se *EIF3M* (*degree* 59 – *degree* médio: 14,4). Esse gene codifica a subunidade M do fator de iniciação de tradução eucariótico 3 (eIF-3; GeneCards symbol: *EIF3M*). Mutações nos genes que codificam as subunidades de outro fator de iniciação de tradução, eIF2b, causam uma doença conhecida como leucoencefalopatia com desaparecimento da matéria branca. Esse distúrbio se caracteriza pela neurodegeneração principalmente da matéria branca do cérebro após a ação de algum estressor, pois o defeito genético leva a uma hiper supressão da tradução numa situação de stress agudo, o que impede a recuperação do tecido (Moon & Parker 2018).

O complexo eIF-3, por outro lado, parece ter um papel mais importante em situações de estresse crônico e age como regulador da iniciação da tradução de RNAs mensageiros relacionados a resposta integrada ao estresse, que tem início, por sua vez, no estresse do RE (Guan et al. 2017). Estresse do RE é um processo já descrito em amostras de fibroblastos de pacientes com ALD e provavelmente ocorre por conta do acúmulo de VLCFA (van de Beek 2017). Não é possível afirmar que esta alteração de eIF-3 seja a causa da desmielinização nos indivíduos CALD, porém esse complexo pode ter uma influência em cadeia, justamente por ser central para a tradução de outras proteínas que lidam com o estresse crônico causado pelos VLCFA. Outra subunidade de eIF3, eIF3B, foi relacionada com a proliferação celular em adenocarcinoma de pulmão por meio da regulação da via EGFR-ERK (Zhong et al. 2016). Dada as semelhanças entre as vias de sinalização observadas nas amostras dessa pesquisa e diversos cânceres, é possível inferir uma interação entre eIF3, as cascatas de sinalização discutidas anteriormente e a proliferação celular.

O estudo apresenta grandes limitações em relação a diversos aspectos das amostras, que foram retiradas de bancos de dados públicos. Em geral os conjuntos apresentam origens diferentes quanto a tecidos e células coletados, pequeno N amostral e variam quanto ao número de replicatas, tornando difícil estabelecer o significado estatístico das descobertas. Também ocorrem divergências entre os fenótipos considerados em diferentes grupos, de forma que apenas três continham amostras de AMN além de CALD. Desses três, E-MEXP-3288 pode ser interpretado de maneira ambígua, já que suas amostras de AMN são na verdade coletadas de pacientes com AMN que passaram a apresentar os sintomas de CALD. Além disso, não foram realizadas análises experimentais para complementar os principais levantamentos dessa pesquisa, o que prejudica a validação das sugestões discutidas.

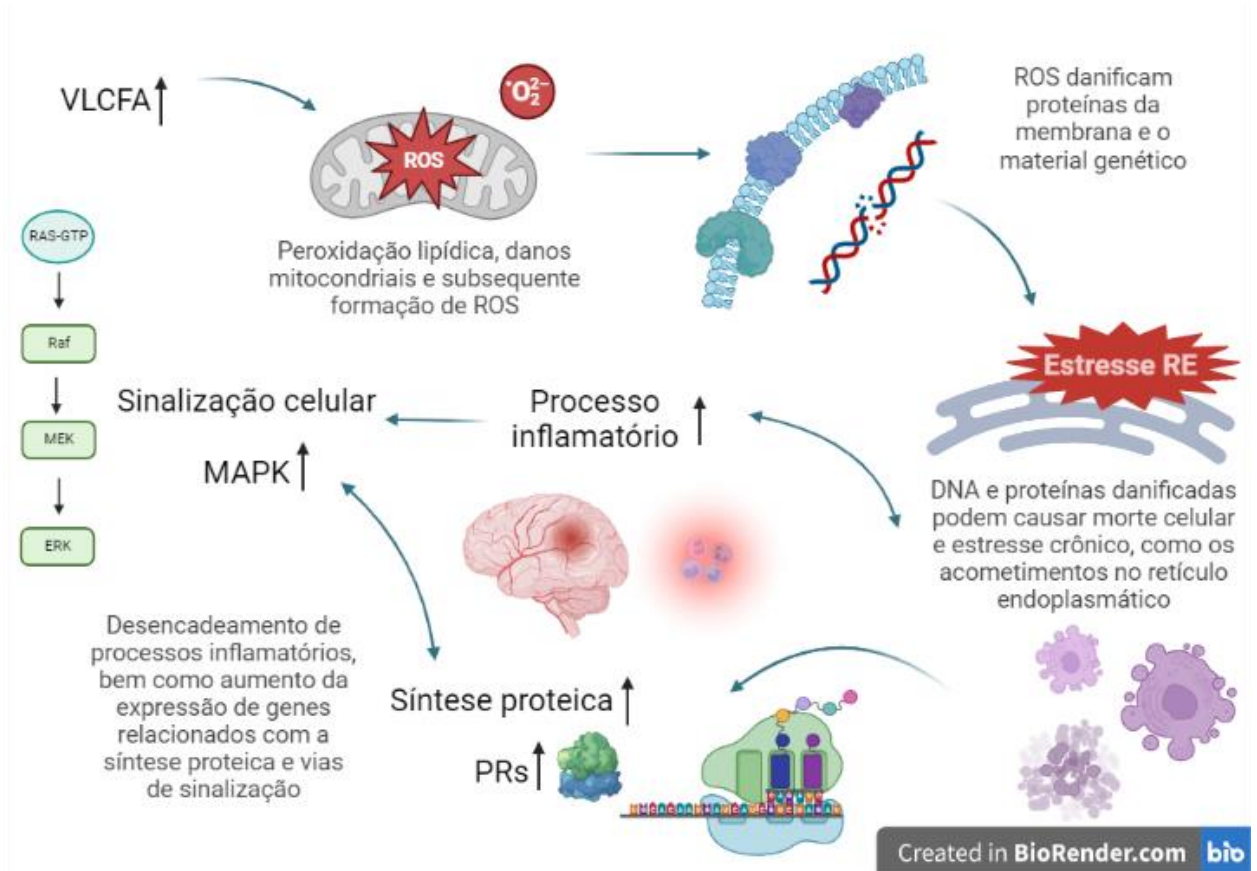
### 3.5 Conclusão

ALD é uma doença heterogênea na qual, provavelmente, a ação de muitos fatores contribui para o desenvolvimento de diferentes fenótipos. Esse estudo explorou os perfis de expressão de dois dos fenótipos mais recorrentes, CALD e AMN, encontrando semelhanças compartilhadas entre ambos. Muitas das alterações observadas em ALD são comuns a outras doenças neurodegenerativas, bem como a cânceres e provavelmente doenças virais, levando a crer que não foram encontrados mecanismos exclusivos de patogenicidade para os fenótipos de ALD estudados. No entanto, nossos achados estão de acordo com alterações metabólicas descritas em outras pesquisas com pacientes e modelos animais de ALD.

Sugere-se que os VLCFA têm um papel importante na patogênese de diferentes fenótipos de ALD, pois podem desencadear o estresse oxidativo por meio da peroxidação lipídica e distúrbios mitocondriais com subsequente formação de ROS. Esses fatores levam a danos estruturais em proteínas e no DNA, que podem causar a morte celular. Alterações nos níveis de expressão de genes que codificam PRs podem representar uma resposta celular a esse estresse. Ao mesmo tempo, a morte celular desencadeia processos inflamatórios que resultam no

aumento da expressão de genes ligados a vias de sinalização celular. Por outro lado, a expressão diferencial de genes neuroprotetores ou ligados ao reparo de DNA podem proteger pacientes com AMN do dano cerebral (Figura 8).

**Fig 8** Esquema contendo um resumo das descobertas e conclusões encontradas nessa pesquisa



Ainda que esse estudo apresente limitações importantes, os resultados sugerem alvos interessantes para pesquisas futuras sobre os mecanismos de doença cerebral em ALD. Além disso, dados obtidos por biologia de sistemas podem contribuir para o desenvolvimento de novas estratégias terapêuticas. De acordo com os principais achados relativos ao estresse oxidativo em ALD, sugere-se que antioxidantes podem ser estudados como terapia complementar de tratamentos convencionais.

### 3.6 Referências

Amsterdam A, Sadler KC, Lai K, Farrington S, Bronson RT, Lees JA, Hopkins N (2004) Many ribosomal protein genes are cancer genes in zebrafish. *PLoS biology*, 2(5). <https://doi.org/10.1371/journal.pbio.0020139>

Argmann CA, Houten SM, Zhu J, Schadt EE (2015) A next generation view of inborn errors of metabolism. *Cell Metab.* 23, 13–26. <https://doi.org/10.1016/j.cmet.2015.11.012>

Aubourg P, Wanders R (2013) Peroxisomal disorders. In: Dulac O, Lasseigne M, Sarnat HB (ed) *Handbook of Clinical Neurology*, Elsevier, Vol. 113, pp 1593-1609

Bader GD, Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics*, 4, 2. <https://doi.org/10.1186/1471-2105-4-2>

Bhavsar RB, Makley LN, Tsonis PA (2010) The other lives of ribosomal proteins. *Hum Genomics* 4(327) <https://doi.org/10.1186/1479-7364-4-5-327>

Chen Y, Lun AAT, Smyth GK (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research*, 5, 1438. doi: 10.12688/f1000research.8987.2.

Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014) cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 8 (Suppl 4), S11. <https://doi.org/10.1186/1752-0509-8-S4-S11>

Deon M, Marchetti DP, Donida B, Wajner M, Vargas M (2016) Oxidative Stress in Patients with X-Linked Adrenoleukodystrophy. *Cell Mol Neurobiol* 36:497–512. <https://doi.org/10.1007/s10571-015-0234-2>

Garcia-Esparcia P, Sideris-Lampretsas G, Hernandez-Ortega K, Grau-Rivera O, Sklaviadis T, Gelpi E, Ferrer I (2017) Altered mechanisms of protein synthesis in frontal cortex in Alzheimer disease and a mouse model. *Am J Neurodegener Dis.*15;6(2):15-25

Guan BJ, van Hoef V, Jobava R, Elroy-Stein O, Valasek LS, Cargnello M, Gao XH, Krokowski D, Merrick WC, Kimball SR, Komar AA, Koromilas AE, Wynshaw-Boris A, Topisirovic I, Larsson O, Hatzoglou M (2017) A Unique ISR Program Determines Cellular Responses to Chronic Stress. *Molecular cell*, 68(5): 885-900.e6. <https://doi.org/10.1016/j.molcel.2017.11.007>

Honey M, Jaspers Y, Engelen M, Kemp S, Huffnagel I C (2021) Molecular Biomarkers for Adrenoleukodystrophy: An Unmet Need. *Cells* 10(12):3427. <https://doi.org/10.3390/cells10123427>

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD (2019) PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic acids research*, 47(D1): D419–D426. <https://doi.org/10.1093/nar/gky1038>

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1): 27–30. <https://doi.org/10.1093/nar/28.1.27>

Kemp S, Huffnagel I C, Linthorst G E, Wanders R J, Engelen M (2016). Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. *Nature reviews. Endocrinology* 12(10):606–615. <https://doi.org/10.1038/nrendo.2016.90>

Khayer N, Mirzaie M, Marashi SA, Jalessi M (2020) Rps27a might act as a controller of microglia activation in triggering neurodegenerative diseases. *PLOS ONE* 15(9): e0239219. <https://doi.org/10.1371/journal.pone.0239219>

Kim EK, Choi EJ (2015) Compromised MAPK signaling in human diseases: an update. *Arch Toxicol* 89: 867–882. <https://doi.org/10.1007/s00204-015-1472-2>

Korenke, GC, Fuchs S, Krasemann E, Doerr HG, Wilichowski E, Hunneman DH, Hanefeld F (1996) Cerebral adrenoleukodystrophy (ALD) in only one of monozygotic twins with an identical ALD genotype. *Annals of neurology* 40(2):254–257. <https://doi.org/10.1002/ana.410400221>

Kruska N, Schoenfeld P, Pujol A, Reiser, G (2015). Astrocytes and mitochondria from adrenoleukodystrophy protein (ABCD1)-deficient mice reveal that the adrenoleukodystrophy-associated very long-chain fatty acids target several cellular energy-dependent functions. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1852. 10.1016/j.bbadis.2015.01.005.

Launay N, Ruiz M, Fourcade S, Schlüter A, Guilera C, Ferrer I, Knecht E, Pujol A (2013) Oxidative stress regulates the ubiquitin–proteasome system and immunoproteasome functioning in a mouse model of X-adrenoleukodystrophy, *Brain*, 136 (3) pp 891–904. <https://doi.org/10.1093/brain/aws370>

Li C, Ge M, Chen D, Sun T, Jiang H, Xie Y, Lu H, Zhang B, Han L, Chen J, Zhu J (2020) RPL21 siRNA blocks proliferation in pancreatic cancer cells by inhibiting DNA replication and inducing G1 arrest and apoptosis. *Front. Oncol.*, 10. <https://doi.org/10.3389/fonc.2020.01730>

Li C, Ge M, Yin Y, Luo M, Chen D (2012) Silencing expression of ribosomal protein L26 and L29 by RNA interfering inhibits proliferation of human pancreatic cancer PANC-1 cells. *Mol Cell Biochem* 370:127–139. <https://doi.org/10.1007/s11010-012-1404-x>

Mallack EJ, Gao K, Engelen M, Kemp S (2022) Structure and Function of the ABCD1 Variant Database: 20 Years, 940 Pathogenic Variants, and 3400 Cases of Adrenoleukodystrophy. *Cells* 11(2):283. <https://doi.org/10.3390/cells11020283>

Marchetti DP, Donida B, Deon M, Jacques CE, Jardim LB, Vargas CR (2017) In vitro effect of N-acetyl-L-cysteine on glutathione and sulfhydryl levels in X-linked adrenoleukodystrophy patients. *Clin Biomed Res* ,37(1).

Masmoudi-Kouki O, Hamdi Y, Ghouili I, Bahdoudi S, Kaddour H, Leprince J, Castel H, Vaudry H, Amri M, Vaudry D, Tonon MC (2018) Neuroprotection with the Endozepine Octadecaneuropeptide, ODN. *Current pharmaceutical design*, 24(33):3918–3925. <https://doi.org/10.2174/1381612824666181112111746>

McMackin MZ, Durbin-Johnson B, Napierala M, Napierala JS, Ruiz L, Napoli E, Perlman S, Giulivi C, Cortopassi GA (2019) Potential biomarker identification for Friedreich’s ataxia using overlapping gene expression patterns in patient cells and mouse dorsal root ganglion. *PLOS ONE* 14(10): e0223209. <https://doi.org/10.1371/journal.pone.0223209>

Moon SL, Parker R (2018) EIF2B2 mutations in vanishing white matter disease hypersuppress translation and delay recovery during the integrated stress response. *RNA*, 24(6): 841–852. <https://doi.org/10.1261/rna.066563.118>

- Moser H W (1997). Adrenoleukodystrophy: phenotype, genetics, pathogenesis and therapy. *Brain : a journal of neurology*, 120 ( Pt 8), 1485–1508. <https://doi.org/10.1093/brain/120.8.1485>
- Moser A B, Jones R O, Hubbard W C, Tortorelli S, Orsini J J, Caggana M, Vogel B H, Raymond G V (2016). Newborn screening for X-Linked adrenoleukodystrophy. *International Journal of Neonatal Screening* 2(4), [15]. <https://doi.org/10.3390/ijns2040015>
- Moser HW, Mahmood A, Raymond GV (2007). X-linked adrenoleukodystrophy. *Nat Clin Pract Neurol* 3: 140–151. <https://doi.org/10.1038/ncpneuro0421>
- Namsi A, Nury T, Khan AS, Leprince J, Vaudry D, Caccia C, Leoni V, Atanasov AG, Tonon MC, Masmoudi-Kouki O, Lizard G (2019) Octadecaneuropeptide (ODN) induces N2a cells differentiation through a PKA/PLC/PKC/MEK/ERK-dependent pathway: Incidence on peroxisome, mitochondria, and lipid profiles. *Molecules*, 24(18), [3310]. <https://doi.org/10.3390/molecules24183310>
- Naumann M, Pal A, Goswami A, Lojewski X, Japtok J, Vehlow A, Naujock M, Günther R, Jin M, Stanslowsky N, Reinhardt P, Sternecker J, Frickenhaus M, Pan-Montojo F, Storkebaum E, Poser I, Freischmidt A, Weishaupt JH, Holzmann K, Troost D et al. (2018) Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nature communications*, 9(1): 335. <https://doi.org/10.1038/s41467-017-02299-1>
- Ning, Q, Pang, Y, Shao, S, Luo M, Zhao L, Hu T, Zhao X (2020) MicroRNA-147b suppresses the proliferation and invasion of non-small-cell lung cancer cells through downregulation of Wnt/ $\beta$ -catenin signalling via targeting of RPS15A. *Clin Exp Pharmacol Physiol*. 47: 449– 458. <https://doi.org/10.1111/1440-1681.13203>
- O'Neill, B P, Moser, H W, & Marmion, L C (1982). Adrenoleukodystrophy: elevated C26 fatty acid in cultured skin fibroblasts and correlation with disease expression in three generations of a kindred. *Neurology* 32(5):540–542. <https://doi.org/10.1212/wnl.32.5.540>
- Pandey V, Kumar V (2015) Stabilization of SIRT7 deacetylase by viral oncoprotein HBx leads to inhibition of growth restrictive RPS7 gene and facilitates cellular transformation. *Sci Rep* 5, 14806. <https://doi.org/10.1038/srep14806>
- Peters C, Charnas LR, Tan Y, Ziegler RS, Shapiro EG, DeFor T, Grewal SS, Orchard PJ, Abel SL, Goldman AI, Ramsay NKC, Dusenbery KE, Loes DJ, Lockman LA, Kato S, Aubourg PR, Moser HW, Krivit W (2004) Cerebral X-linked adrenoleukodystrophy: the international hematopoietic cell transplantation experience from 1982 to 1999. *Blood* 104 (3): 881–888. doi: <https://doi.org/10.1182/blood-2003-10-3402>
- Poll-The B T, Aubourg P, Wanders R (2006) Peroxisomal disorders. In: *Inborn Metabolic Diseases*, 4th edn. Springer, pp 509-521

- Powers JM, Pei Z, Heinzer AK, Deering R, Moser AB, Moser HW, Watkins PA, Smith KD (2005). Adrenoleukodystrophy: oxidative stress of mice and men. *Journal of neuropathology and experimental neurology*, 64(12): 1067–1079. <https://doi.org/10.1097/01.jnen.0000190064.28559.a4>
- Rai SN, Dilnashin H, Birla H, Singh SS, Zahra W, Rathore AS, Singh BK, Singh SP (2019) The Role of PI3K/Akt and ERK in Neurodegenerative Disorders. *Neurotox Res* 35: 775–795. <https://doi.org/10.1007/s12640-019-0003-y>
- Richmond PA, van der Kloet F, Vaz FM, Lin D, Uzozie A, Graham E, Kobor M, Mostafavi S, Moerland PM, Lange PF, van Kampen AHC, Wasserman WW, Engelen M, Kemp S, van Karnebeek CDM. *Front. Cell Dev. Biol.* <https://doi.org/10.3389/fcell.2020.00520>
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47. doi: 10.1093/nar/gkv007.
- Scardoni G, Tosadori G, Faizan M, Spoto F, Fabbri F, Laudanna C (2014). Biological network analysis with CentiScaPe: centralities and experimental dataset integration. *F1000Research*, 3, 139. <https://doi.org/10.12688/f1000research.4477.2>
- Schlüter A, Espinosa L, Fourcade S, Galino J, López E, Ilieva E, Morató L, Asheuer M, Cook T, McLaren A, Reid J, Kelly F, Bates S, Aubourg P, Galea E, Pujol A (2012). Functional genomic analysis unravels a metabolic-inflammatory interplay in adrenoleukodystrophy. *Human molecular genetics*, 21(5):1062–1077. <https://doi.org/10.1093/hmg/ddr536>
- Shah S, Zhao D, Hussain T, Yang L (2017). The Role of Unfolded Protein Response and Mitogen-Activated Protein Kinase Signaling in Neurodegenerative Diseases with Special Focus on Prion Diseases. *Frontiers in aging neuroscience*, 9, 120. <https://doi.org/10.3389/fnagi.2017.00120>
- Silva GCV, Matte US (2022) Neuronetworks: Analysis of brain pathology in mucopolisaccharidoses - a systems biology approach. *Neuroscience Informatics* 2(1): 2772-5286. <https://doi.org/10.1016/j.neuri.2021.100036>
- Skariah G, Todd PK (2021) Translational control in aging and neurodegeneration. *WIREs RNA*. 12:e1628. <https://doi.org/10.1002/wrna.1628>
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ, von Mering C (2021) The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(D1): D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
- Teister J, Anders F, Beck S, Funke S, von Pein H, Prokosh V, Pfeifer N, Grus F (2017) Decelerated neurodegeneration after intravitreal injection of  $\alpha$ -synuclein antibodies in a glaucoma animal model. *Sci Rep* 7, 6260, pp. 1 – 16. <https://doi.org/10.1038/s41598-017-06702-1>

van de Beek MC, Ofman R, Dijkstra I, Wijburg F, Engelen M, Wanders R, Kemp S (2017). Lipid-induced endoplasmic reticulum stress in X-linked adrenoleukodystrophy. *Biochimica et biophysica acta. Molecular basis of disease*, 1863(9): 2255–2265. <https://doi.org/10.1016/j.bbadis.2017.06.003>

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu x, Liu S, Bo X, Yu G (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2(3), 100141. <https://doi.org/10.1016/j.xinn.2021.100141>.

Yue J, López JM (2020) Understanding MAPK Signaling Pathways in Apoptosis. *International Journal of Molecular Sciences*. 21(7):2346. <https://doi.org/10.3390/ijms21072346>

Zhong B, Zhang C, Chen Y, Tang J, Shen J (2016) EIF3B silencing inhibits cell proliferation via regulating EGFR/ERK pathway in lung adenocarcinoma. *Int J Clin Exp Pathol*, 9(11):11179-11187

### 3.7 Declarações

#### Financiamento

Nenhum financiamento, patrocínio ou remuneração foi recebido pelo autores para a elaboração dessa pesquisa.

#### Conflito de interesses

Os autores não têm nenhum conflito de interesses a declarar.

#### Disponibilidade dos dados

As amostras analisadas nesse estudo podem ser acessadas mediante seus códigos nos bancos de dados *Gene Expression Omnibus*, <https://www.ncbi.nlm.nih.gov/geo/> e *ArrayExpress*, <https://www.ebi.ac.uk/arrayexpress/>

### 3.8 Informações Suplementares

Tabela suplementar 1. Total de clusters encontrados com MCODE

|             | CALD vs Controle | AMN vs Controle | CALD vs AMN |
|-------------|------------------|-----------------|-------------|
| GSE117647   | 28               | 32              | 44          |
| GSE85804    | 43               | 31              | 29          |
| GSE34308    | 29               | -               | -           |
| GSE34309    | 44               | -               | -           |
| GSE108012   | 21               | -               | -           |
| E-MEXP-3288 | 46               | 46              | 29          |
| Média       | 35,16667         | 36,33333        | 34          |

Tabela suplementar 2. Medidas de centralidade dos genes comuns entre diferentes conjuntos de dados com os maiores *degrees*

|             | Genes         | Betweenness | Closeness | Degree | Stress   |
|-------------|---------------|-------------|-----------|--------|----------|
| E-MEXP-3288 | <i>RPS27A</i> | 6,99E+04    | 0,00043   | 118    | 1,56E+06 |
|             | <i>RPS11</i>  | 7,98E+03    | 0,00042   | 115    | 4,26E+05 |
|             | <i>RPS4X</i>  | 9,09E+03    | 0,00041   | 115    | 3,97E+05 |
|             | <i>RPS9</i>   | 3,49E+03    | 0,00041   | 106    | 2,42E+05 |
|             | <i>RPS15A</i> | 5,40E+03    | 0,0004    | 106    | 2,57E+05 |
| GSE85804    | <i>RPS11</i>  | 1,91E+04    | 0,00019   | 50     | 2,28E+05 |
|             | <i>RPS15A</i> | 1,22E+04    | 0,0002    | 40     | 1,92E+05 |
|             | <i>RPS7</i>   | 9,40E+03    | 0,0002    | 39     | 1,44E+05 |
|             | <i>RPL21</i>  | 5,47E+03    | 0,00019   | 38     | 8,75E+04 |
|             | <i>RPL34</i>  | 5,00E+03    | 0,00019   | 34     | 7,49E+04 |
| GSE108012   | <i>RPS27A</i> | 4,30E+04    | 0,00139   | 95     | 1,76E+05 |
|             | <i>RPS16</i>  | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS11</i>  | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS23</i>  | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS9</i>   | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS15A</i> | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS7</i>   | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
|             | <i>RPS4X</i>  | 6,16E+01    | 0,00112   | 81     | 2,44E+03 |
| GSE834308   | <i>RPS4X</i>  | 1,23E+03    | 0,00087   | 75     | 3,85E+04 |
|             | <i>RPS11</i>  | 2,08E+03    | 0,00089   | 73     | 4,95E+04 |
|             | <i>RPS16</i>  | 7,67E+02    | 0,00086   | 71     | 2,67E+04 |
|             | <i>RPL4</i>   | 9,63E+02    | 0,00087   | 71     | 3,06E+04 |
|             | <i>RPL15</i>  | 8,69E+02    | 0,00086   | 70     | 2,69E+04 |
|             | <i>RPS9</i>   | 1,59E+03    | 0,00087   | 70     | 4,00E+04 |
| GSE34309    | <i>RPS27A</i> | 1,15E+05    | 0,00056   | 83     | 1,96E+06 |
|             | <i>RPS4X</i>  | 3,53E+03    | 0,00048   | 59     | 1,18E+05 |
|             | <i>RPS11</i>  | 1,38E+04    | 0,00049   | 57     | 2,75E+05 |



|                |               |          |         |     |          |
|----------------|---------------|----------|---------|-----|----------|
|                | <i>RPS16</i>  | 1,47E+03 | 0,00048 | 54  | 7,75E+04 |
|                | <i>RPS9</i>   | 4,21E+03 | 0,00049 | 54  | 1,31E+05 |
| GSE117647      | <i>RPS27A</i> | 1,14E+05 | 0,00058 | 91  | 1,28E+06 |
|                | <i>RPS4X</i>  | 1,03E+03 | 0,0005  | 62  | 2,56E+04 |
|                | <i>RPS2</i>   | 4,55E+03 | 0,00051 | 60  | 7,37E+04 |
|                | <i>RPS3</i>   | 1,27E+04 | 0,00052 | 59  | 1,36E+05 |
|                | <i>RPL27A</i> | 1,84E+03 | 0,0005  | 59  | 5,29E+04 |
| AMN x controle |               |          |         |     |          |
| E-MEXP-3288    | <i>RPS27A</i> | 7,04E+04 | 0,00045 | 131 | 1,36E+06 |
|                | <i>RPS11</i>  | 6,46E+03 | 0,00042 | 115 | 3,33E+05 |
|                | <i>RPS4X</i>  | 7,11E+03 | 0,00042 | 112 | 2,64E+05 |
|                | <i>RPS3</i>   | 1,87E+04 | 0,00043 | 107 | 3,92E+05 |
|                | <i>RPS15A</i> | 3,17E+03 | 0,00041 | 105 | 1,83E+05 |
| GSE85804       | <i>RPS27A</i> | 3,97E+04 | 0,00097 | 87  | 8,53E+05 |
|                | <i>RPS11</i>  | 3,87E+03 | 0,00096 | 85  | 1,39E+05 |
|                | <i>RPS4X</i>  | 1,96E+03 | 0,00095 | 84  | 1,25E+05 |
|                | <i>RPS3</i>   | 1,96E+03 | 0,00092 | 81  | 7,67E+04 |
|                | <i>RPS9</i>   | 1,74E+03 | 0,00094 | 76  | 1,08E+05 |
| GSE117647      | <i>RPS27A</i> | 5,38E+04 | 0,00057 | 75  | 3,72E+05 |
|                | <i>RPS4X</i>  | 1,74E+03 | 0,00052 | 54  | 7,09E+04 |
|                | <i>RPS11</i>  | 8,70E+03 | 0,00054 | 51  | 1,54E+05 |
|                | <i>RPS16</i>  | 9,84E+02 | 0,00052 | 50  | 4,13E+04 |
|                | <i>RPS9</i>   | 1,72E+03 | 0,00052 | 49  | 4,90E+04 |
| CALD x AMN     |               |          |         |     |          |
| E-MEXP-3288    | <i>RPS27A</i> | 7,36E+04 | 0,00051 | 117 | 1,29E+06 |
|                | <i>RPS11</i>  | 4,49E+03 | 0,00047 | 97  | 2,51E+05 |
|                | <i>RPS3</i>   | 6,58E+03 | 0,00047 | 94  | 2,09E+05 |
|                | <i>RPS15A</i> | 3,14E+03 | 0,00046 | 90  | 1,93E+05 |
|                | <i>RPS2</i>   | 2,43E+03 | 0,00046 | 89  | 1,25E+05 |

|           |               |          |         |    |          |
|-----------|---------------|----------|---------|----|----------|
|           | <i>RPS9</i>   | 1,85E+03 | 0,00045 | 89 | 1,30E+05 |
| GSE85804  | <i>RPS27A</i> | 3,83E+04 | 0,00093 | 80 | 7,49E+05 |
|           | <i>RPS11</i>  | 5,01E+03 | 0,00092 | 72 | 1,65E+05 |
|           | <i>RPL15</i>  | 1,84E+03 | 0,00089 | 69 | 1,23E+05 |
|           | <i>RPL18A</i> | 2,69E+03 | 0,0009  | 67 | 1,15E+05 |
|           | <i>RPS2</i>   | 2,13E+03 | 0,0009  | 67 | 1,17E+05 |
| GSE117647 | <i>RPS27A</i> | 7,35E+04 | 0,00053 | 85 | 6,60E+05 |
|           | <i>RPS11</i>  | 8,66E+03 | 0,0005  | 63 | 2,49E+05 |
|           | <i>RPS9</i>   | 3,59E+03 | 0,00049 | 62 | 1,65E+05 |
|           | <i>RPS16</i>  | 1,91E+03 | 0,00048 | 61 | 1,12E+05 |
|           | <i>RPS15A</i> | 3,43E+03 | 0,00048 | 60 | 1,37E+05 |

Genes hub comuns a todos os conjuntos amostrais de cada comparação

| AMN vs controle | CALD vs AMN   | CALD vs controle |
|-----------------|---------------|------------------|
| <i>EEF1G</i>    | <i>EEF2</i>   | <i>RPL21</i>     |
| <i>EEF2</i>     | <i>RPL13</i>  | <i>RPS11</i>     |
| <i>RPL13</i>    | <i>RPL15</i>  | <i>RPS15A</i>    |
| <i>RPL15</i>    | <i>RPL18A</i> | <i>RPS7</i>      |
| <i>RPL18A</i>   | <i>RPL19</i>  |                  |
| <i>RPL19</i>    | <i>RPL35</i>  |                  |
| <i>RPL22</i>    | <i>RPL37</i>  |                  |
| <i>RPL3</i>     | <i>RPL38</i>  |                  |
| <i>RPL35</i>    | <i>RPL4</i>   |                  |
| <i>RPL37</i>    | <i>RPL7</i>   |                  |
| <i>RPL38</i>    | <i>RPL8</i>   |                  |
| <i>RPL4</i>     | <i>RPLP2</i>  |                  |
| <i>RPL7</i>     | <i>RPS11</i>  |                  |
| <i>RPL8</i>     | <i>RPS12</i>  |                  |
| <i>RPLP2</i>    | <i>RPS15A</i> |                  |
| <i>RPS11</i>    | <i>RPS16</i>  |                  |
| <i>RPS12</i>    | <i>RPS2</i>   |                  |
| <i>RPS15A</i>   | <i>RPS21</i>  |                  |
| <i>RPS16</i>    | <i>RPS23</i>  |                  |

|               |               |
|---------------|---------------|
| <i>RPS23</i>  | <i>RPS27A</i> |
| <i>RPS27A</i> | <i>RPS3</i>   |
| <i>RPS28</i>  | <i>RPS7</i>   |
| <i>RPS3</i>   | <i>RPS9</i>   |
| <i>RPS4X</i>  | <i>SRP14</i>  |
| <i>RPS7</i>   |               |
| <i>RPS9</i>   |               |

## Expressão de genes analisados durante a pesquisa para CALD vs controle

| Conjunto amostral | Gene          | P valor ajustado | P valor  | logFC  |
|-------------------|---------------|------------------|----------|--------|
| GSE117647         | <i>RPS4X</i>  | 0,947            | 0,010    | -0,230 |
|                   | <i>RPS27A</i> | 1                | 0,085    | -0,116 |
|                   | <i>RPS11</i>  | 1                | 0,979    | 0,003  |
|                   | <i>RPS9</i>   | 1                | 0,347    | -0,087 |
|                   | <i>RPS15A</i> | 1                | 0,112    | -0,117 |
|                   | <i>RPS7</i>   | 1                | 0,130    | -0,142 |
|                   | <i>RPL21</i>  | 1                | 0,187    | -0,097 |
|                   | <i>NLGN4Y</i> | 1                | 0,266    | 0,094  |
|                   | <i>PIK3R1</i> | 0,947            | 0,011    | -0,347 |
|                   | <i>AKT3</i>   | 1                | 0,125    | -0,137 |
|                   | <i>SMAD3</i>  | 1                | 0,153    | -0,103 |
|                   | <i>PRKCA</i>  | 0,988            | 0,025    | -0,154 |
|                   | <i>PRKCB</i>  | 1                | 0,102    | -0,316 |
| GSE85804          | <i>RPS4X</i>  | 0,894            | 0,432    | 0,130  |
|                   | <i>RPS27A</i> | 0,709            | 0,133    | 0,132  |
|                   | <i>RPS11</i>  | 0,993            | 0,909    | -0,010 |
|                   | <i>RPS9</i>   | 0,818            | 0,257    | 0,174  |
|                   | <i>RPS15A</i> | 0,144            | 0,002    | 0,755  |
|                   | <i>RPS7</i>   | 0,337            | 0,014    | -0,428 |
|                   | <i>RPL21</i>  | 0,151            | 0,002    | -0,845 |
|                   | <i>NLGN4Y</i> | 0,288            | 0,010    | 0,254  |
|                   | <i>PIK3R1</i> | 0,663            | 0,100    | 0,166  |
|                   | <i>AKT3</i>   | 0,831            | 0,277    | 0,089  |
|                   | <i>SMAD3</i>  | 0,007            | 7,52E-06 | -0,780 |
|                   | <i>PRKCA</i>  | 0,893            | 0,428    | -0,090 |
|                   | <i>PRKCB</i>  | 0,462            | 0,030    | -0,297 |
| GSE34308          | <i>RPS4X</i>  | 0,693            | 0,326    | 0,083  |
|                   | <i>RPS27A</i> | 0,423            | 0,042    | 0,198  |
|                   | <i>RPS11</i>  | 0,915            | 0,758    | -0,390 |

|             |               |                      |          |        |
|-------------|---------------|----------------------|----------|--------|
|             | <i>RPS9</i>   | 0,753                | 0,429    | 0,134  |
|             | <i>RPS15A</i> | 0,454                | 0,059    | -0,169 |
|             | <i>RPS7</i>   | 0,747                | 0,418    | -0,091 |
|             | <i>RPL21</i>  | 0,490                | 0,089    | -0,138 |
|             | <i>NLGN4Y</i> | 0,506                | 0,102    | -0,493 |
|             | <i>PIK3R1</i> | 0,729                | 0,385    | 0,393  |
|             | <i>AKT3</i>   | 0,298                | 0,007    | -0,750 |
|             | <i>SMAD3</i>  | 0,159                | 2,35E-04 | 0,772  |
|             | <i>PRKCA</i>  | 0,540                | 0,133    | 0,613  |
|             | <i>PRKCB</i>  | 0,548                | 0,142    | 0,142  |
|             | Gene          | P valor ajustado     | P valor  | logFC  |
| GSE34309    | <i>RPS4X</i>  | 0,467                | 0,042    | 0,173  |
|             | <i>RPS27A</i> | 0,813                | 0,403    | -0,082 |
|             | <i>RPS11</i>  | 0,651                | 0,164    | -0,369 |
|             | <i>RPS9</i>   | 0,782                | 0,343    | 0,062  |
|             | <i>RPS15A</i> | 0,964                | 0,841    | 0,012  |
|             | <i>RPS7</i>   | 0,538                | 0,076    | 0,119  |
|             | <i>RPL21</i>  | 0,879                | 0,558    | 0,031  |
|             | <i>NLGN4Y</i> | 0,040                | 9,41E-06 | -1,587 |
|             | <i>PIK3R1</i> | 0,186                | 0,001    | 0,965  |
|             | <i>AKT3</i>   | 0,523                | 0,068    | -0,177 |
|             | <i>SMAD3</i>  | 0,283                | 0,006    | -0,331 |
|             | <i>PRKCA</i>  | 0,728                | 0,256    | -0,311 |
|             | <i>PRKCB</i>  | 0,325                | 0,009    | -0,624 |
|             | Gene          | False discovery rate | P valor  | logFC  |
| GSE108012   | <i>RPS4X</i>  | 0,331                | 0,039    | -0,226 |
|             | <i>RPS27A</i> | 1                    | 0,876    | 0,020  |
|             | <i>RPS11</i>  | 0,989                | 0,400    | 0,083  |
|             | <i>RPS9</i>   | 1                    | 0,408    | 0,079  |
|             | <i>RPS15A</i> | 1                    | 0,873    | 0,021  |
|             | <i>RPS7</i>   | 1                    | 0,768    | -0,032 |
|             | <i>RPL21</i>  | 1                    | 0,836    | 0,020  |
|             | <i>NLGN4Y</i> | 0,813                | 0,288    | 0,795  |
|             | <i>PIK3R1</i> | 0,006                | 1,92E-04 | 0,645  |
|             | <i>AKT3</i>   | 0,184                | 0,016    | 0,360  |
|             | <i>SMAD3</i>  | 0,395                | 0,060    | -0,226 |
|             | <i>PRKCA</i>  | 0,121                | 0,009    | -0,459 |
|             | <i>PRKCB</i>  | 2,45E-06             | 1,936    | -1,800 |
|             | Gene          | P valor ajustado     | P valor  | logFC  |
| E-MEXP-3288 | <i>RPS4X</i>  | 2,98E-13             | 1,14E-14 | 11,371 |
|             | <i>RPS27A</i> | 2,98E-13             | 4,04E-15 | 10,784 |
|             | <i>RPS11</i>  | 2,98E-13             | 3,82E-14 | 11,777 |
|             | <i>RPS9</i>   | 2,98E-13             | 3,86E-15 | 10,440 |
|             | <i>RPS15A</i> | 3,45E-13             | 1,18E-13 | 11,092 |

|               |          |          |        |
|---------------|----------|----------|--------|
| <i>RPS7</i>   | 3,37E-13 | 1,10E-13 | 10,403 |
| <i>RPL21</i>  | 6,54E-13 | 3,69E-13 | 5,641  |
| <i>NLGN4Y</i> | 5,43E-11 | 5,10E-11 | 5,025  |
| <i>PIK3R1</i> | 3,66E-13 | 1,37E-13 | 5,737  |
| <i>AKT3</i>   | 2,98E-13 | 5,38E-14 | 7,391  |
| <i>SMAD3</i>  | 3,89E-13 | 1,55E-13 | 5,407  |
| <i>PRKCA</i>  | 4,83E-13 | 2,31E-13 | 8,435  |
| <i>PRKCB</i>  | 4,57E-12 | 3,88E-12 | 8,125  |

---

## Expressão de genes analisados durante a pesquisa para AMN vs controle

| Conjunto amostral | Gene          | P valor ajustado | P valor  | logFC  |
|-------------------|---------------|------------------|----------|--------|
| GSE117647         | <i>RPS27A</i> | 0,786            | 0,469    | -0,050 |
|                   | <i>RPS11</i>  | 0,904            | 0,716    | -0,035 |
|                   | <i>RPS4X</i>  | 0,119            | 0,001    | 0,198  |
|                   | <i>FUS</i>    | 0,042            | 2,07E-05 | 0,519  |
|                   | <i>GNB5</i>   | 0,409            | 0,065    | 0,192  |
|                   | <i>MAPK3</i>  | 0,563            | 0,170    | 0,083  |
|                   | <i>AKT1</i>   | 0,570            | 0,177    | -0,075 |
|                   | <i>AKT2</i>   | 0,294            | 0,025    | 0,187  |
|                   | <i>AKT3</i>   | 0,708            | 0,345    | 0,086  |
|                   | <i>MAP2K2</i> | 0,441            | 0,080    | 0,138  |
|                   | <i>MAPK9</i>  | 0,263            | 0,018    | 0,215  |
|                   | <i>NFKB1</i>  | 0,967            | 0,886    | -0,011 |
| GSE85804          | <i>RPS27A</i> | 0,855            | 0,387    | -0,064 |
|                   | <i>RPS11</i>  | 0,838            | 0,343    | 0,112  |
|                   | <i>RPS4X</i>  | 0,313            | 0,004    | -0,710 |
|                   | <i>FUS</i>    | 0,994            | 0,961    | -0,007 |
|                   | <i>GNB5</i>   | 0,021            | 1,96E-05 | 0,750  |
|                   | <i>MAPK3</i>  | 0,677            | 0,068    | 0,360  |
|                   | <i>AKT1</i>   | 0,524            | 0,015    | 0,530  |
|                   | <i>AKT2</i>   | 0,924            | 0,630    | -0,037 |
|                   | <i>AKT3</i>   | 0,778            | 0,193    | 0,136  |
|                   | <i>MAP2K2</i> | 0,904            | 0,557    | 0,194  |
|                   | <i>MAPK9</i>  | 0,358            | 0,005    | -0,310 |
|                   | <i>NFKB1</i>  | 0,769            | 0,176    | -0,257 |
| E-MEXP-3288       | <i>RPS27A</i> | 1,56E-33         | 1,15E-35 | 10,780 |
|                   | <i>RPS11</i>  | 3,95E-33         | 5,83E-35 | 11,568 |
|                   | <i>RPS4X</i>  | 3,31E-35         | 3,42E-38 | 11,270 |
|                   | <i>FUS</i>    | 7,57E-28         | 5,39E-28 | 8,289  |
|                   | <i>GNB5</i>   | 1,78E-31         | 2,98E-32 | 4,958  |
|                   | <i>MAPK3</i>  | 7,08E-33         | 1,70E-34 | 8,380  |
|                   | <i>AKT1</i>   | 1,77E-33         | 1,48E-35 | 7,581  |
|                   | <i>AKT2</i>   | 1,34E-31         | 1,99E-32 | 5,620  |
|                   | <i>AKT3</i>   | 1,28E-31         | 1,87E-32 | 7,239  |
|                   | <i>MAP2K2</i> | 2,99E-31         | 6,08E-32 | 8,169  |
|                   | <i>MAPK9</i>  | 3,48E-30         | 1,32E-30 | 6,595  |
|                   | <i>NFKB1</i>  | 2,30E-31         | 4,27E-32 | 6,987  |

## Expressão de genes analisados durante a pesquisa para CALD vs AMN

| Conjunto amostral | Gene          | P valor ajustado | P valor | logFC  |
|-------------------|---------------|------------------|---------|--------|
| GSE117647         | <i>RPS27A</i> | 0,495            | 0,008   | 0,165  |
|                   | <i>RPS11</i>  | 0,950            | 0,780   | 0,032  |
|                   | <i>RPS15A</i> | 0,832            | 0,428   | -0,070 |

|               |               |                  |         |        |
|---------------|---------------|------------------|---------|--------|
|               | <i>RPS7</i>   | 0,924            | 0,698   | 0,032  |
|               | <i>RPS2</i>   | 0,825            | 0,408   | -0,073 |
|               | <i>RPS21</i>  | 0,954            | 0,792   | 0,026  |
|               | <i>SRP14</i>  | 0,527            | 0,021   | -0,129 |
|               | <i>MAPK1</i>  | 0,546            | 0,037   | -0,216 |
|               | <i>MAPK3</i>  | 0,536            | 0,029   | -0,125 |
|               | <i>AKT2</i>   | 0,497            | 0,009   | -0,198 |
|               | <i>KRAS</i>   | 0,553            | 0,041   | 0,266  |
|               | <i>MAP2K2</i> | 0,500            | 0,010   | -0,152 |
|               | <i>MAPK9</i>  | 0,536            | 0,027   | -0,215 |
|               | <i>PRKCA</i>  | 0,449            | 0,004   | 0,265  |
|               | <i>PIK3CD</i> | 0,532            | 0,024   | -0,205 |
|               | <i>PIK3R3</i> | 0,503            | 0,010   | 0,302  |
|               | <i>PLCB3</i>  | 0,543            | 0,035   | -0,161 |
|               | <i>PIK3CB</i> | 0,693            | 0,197   | 0,057  |
|               | <i>AKT1</i>   | 0,965            | 0,831   | -0,014 |
|               | <i>AKT3</i>   | 0,605            | 0,087   | 0,109  |
|               | <i>EGFR</i>   | 0,509            | 0,013   | 0,211  |
|               | <i>HRAS</i>   | 0,492            | 0,007   | -0,201 |
|               | <i>GNAS</i>   | 0,562            | 0,049   | -0,191 |
|               | <i>MAP2K1</i> | 0,566            | 0,059   | -0,294 |
|               | <i>RAF1</i>   | 0,905            | 0,643   | -0,048 |
|               | <i>MAPK8</i>  | 0,941            | 0,756   | 0,026  |
|               | <i>PRKCB</i>  | 0,536            | 0,027   | -0,704 |
|               | <i>PRKACB</i> | 0,871            | 0,539   | -0,116 |
|               | <i>PRKACG</i> | 0,910            | 0,659   | 0,027  |
|               | <i>ACTB</i>   | 0,527            | 0,021   | -0,270 |
|               | Gene          | P valor ajustado | P valor | logFC  |
| GSE85804      | <i>RPS27A</i> | 0,885            | 0,403   | -0,068 |
|               | <i>RPS11</i>  | 0,894            | 0,436   | -0,102 |
|               | <i>RPS15A</i> | 0,604            | 0,017   | -0,765 |
|               | <i>RPS7</i>   | 0,452            | 0,004   | 0,667  |
|               | <i>RPS2</i>   | 0,884            | 0,399   | 0,371  |
|               | <i>RPS21</i>  | 0,942            | 0,655   | 0,066  |
|               | <i>SRP14</i>  | 0,322            | 0,002   | 0,463  |
|               | <i>MAPK1</i>  | 0,821            | 0,197   | 0,520  |
|               | <i>MAPK3</i>  | 0,710            | 0,047   | 0,318  |
|               | <i>AKT2</i>   | 0,929            | 0,590   | -0,038 |
|               | <i>KRAS</i>   | 0,766            | 0,094   | 0,577  |
|               | <i>MAP2K2</i> | 0,940            | 0,645   | -0,146 |
|               | <i>MAPK9</i>  | 0,446            | 0,004   | 0,310  |
|               | <i>PRKCA</i>  | 0,995            | 0,965   | 0,016  |
|               | <i>PIK3CD</i> | 0,771            | 0,107   | 0,547  |
|               | <i>PIK3R3</i> | 0,967            | 0,795   | -0,027 |
|               | <i>PLCB3</i>  | 0,997            | 0,983   | 0,002  |
| <i>PIK3CB</i> | 0,811         | 0,177            | 0,434   |        |

|             |               |                  |          |        |
|-------------|---------------|------------------|----------|--------|
|             | <i>AKT1</i>   | 0,544            | 0,011    | -0,555 |
|             | <i>AKT3</i>   | 0,770            | 0,099    | 0,135  |
|             | <i>EGFR</i>   | 0,885            | 0,405    | 0,078  |
|             | <i>HRAS</i>   | 0,938            | 0,635    | -0,141 |
|             | <i>GNAS</i>   | 0,882            | 0,389    | 0,148  |
|             | <i>MAP2K1</i> | 0,884            | 0,396    | -0,189 |
|             | <i>RAF1</i>   | 0,867            | 0,339    | 0,290  |
|             | <i>MAPK8</i>  | 0,826            | 0,212    | 0,466  |
|             | <i>PRKCB</i>  | 0,748            | 0,080    | 0,281  |
|             | <i>PRKACB</i> | 0,873            | 0,355    | 0,135  |
|             | <i>PRKACG</i> | 0,991            | 0,936    | 0,008  |
|             | <i>ACTB</i>   | 0,857            | 0,309    | 0,385  |
|             | Gene          | P valor ajustado | P valor  | logFC  |
|             | <i>RPS27A</i> | 1,38E-18         | 2,18E-20 | 10,821 |
|             | <i>RPS11</i>  | 5,81E-18         | 1,54E-18 | 11,579 |
|             | <i>RPS15A</i> | 7,42E-18         | 2,23E-18 | 11,029 |
|             | <i>RPS7</i>   | 1,09E-17         | 3,90E-18 | 10,438 |
|             | <i>RPS2</i>   | 1,38E-18         | 1,40E-20 | 11,762 |
|             | <i>RPS21</i>  | 1,19E-17         | 4,39E-18 | 10,661 |
|             | <i>SRP14</i>  | 1,38E-18         | 6,98E-21 | 10,530 |
|             | <i>MAPK1</i>  | 3,03E-18         | 5,04E-19 | 8,193  |
|             | <i>MAPK3</i>  | 1,42E-18         | 4,33E-20 | 8,253  |
|             | <i>AKT2</i>   | 1,01E-17         | 3,48E-18 | 5,604  |
|             | <i>KRAS</i>   | 5,65E-18         | 1,48E-18 | 3,668  |
|             | <i>MAP2K2</i> | 1,87E-18         | 1,70E-19 | 8,231  |
|             | <i>MAPK9</i>  | 2,58E-18         | 3,58E-19 | 6,606  |
|             | <i>PRKCA</i>  | 1,66E-17         | 6,90E-18 | 4,029  |
|             | <i>PIK3CD</i> | 3,66E-16         | 2,63E-16 | 6,254  |
| E-MEXP-3288 | <i>PIK3R3</i> | 5,08E-16         | 3,76E-16 | 6,814  |
|             | <i>PLCB3</i>  | 5,82E-18         | 1,55E-18 | 6,414  |
|             | <i>PIK3CB</i> | 4,78E-18         | 1,13E-18 | 3,569  |
|             | <i>AKT1</i>   | 1,52E-18         | 8,15E-20 | 7,547  |
|             | <i>AKT3</i>   | 6,52E-18         | 1,83E-18 | 7,392  |
|             | <i>EGFR</i>   | 1,38E-18         | 4,78E-21 | 6,210  |
|             | <i>HRAS</i>   | 7,92E-17         | 4,71E-17 | 7,295  |
|             | <i>GNAS</i>   | 1,38E-18         | 2,45E-20 | 11,362 |
|             | <i>MAP2K1</i> | 5,75E-18         | 1,52E-18 | 8,660  |
|             | <i>RAF1</i>   | 3,66E-18         | 7,22E-19 | 7,910  |
|             | <i>MAPK8</i>  | 9,88E-18         | 3,38E-18 | 6,499  |
|             | <i>PRKCB</i>  | 3,26E-15         | 2,70E-15 | 8,714  |
|             | <i>PRKACB</i> | 1,57E-18         | 9,05E-20 | 10,707 |
|             | <i>PRKACG</i> | 1,51E-18         | 7,38E-20 | 5,919  |
|             | <i>ACTB</i>   | 9,92E-19         | 1,19E-21 | 12,727 |
|             | <i>EIF3M</i>  | 1,73E-18         | 1,34E-19 | 3,897  |



#### 4 CONCLUSÕES E PERSPECTIVAS

A pesquisa demonstrou as alterações na expressão gênica que ocorrem nos fenótipos CALD e AMN, de forma a correlacionar essas informações ao possível quadro fenotípico de indivíduos com ALD. Os resultados para os dois grupos mostram aumento da síntese proteica e grande atividade de cascatas de sinalização envolvidas com a proliferação e diferenciação celular. Sugere-se que essas alterações podem fazer parte de mecanismos compensatórios em resposta aos danos causados pela toxicidade dos VLCFA. De acordo com os achados em outros estudos sobre ALD, bem como em outras doenças, as consequências oxidativas desses ácidos graxos podem envolver danos a proteínas e ao material genético, de forma que o estresse causado pode ativar vias de sinalização que eventualmente estimulam a sobrevivência e o reparo das células.

O estudo apresenta limitações sobre as amostras utilizadas em relação aos fenótipos, idades de pacientes, tecidos coletados, métodos de análise e tamanho de cada grupo. Poucos são os destaques para os fenótipos de forma isolada. Ainda assim, foram observados genes que podem ter envolvimento na patogênese em CALD, como *RPS2*, *RPS21* e *SRP14*, e o gene *FUS*, observado em AMN, que pode ser neuroprotetor. No entanto, não se pode esclarecer quais são os envoltimentos específicos de cada um e se sugere, portanto, que podem ser alvos interessantes para pesquisas futuras, bem como o estudo de antioxidantes para tratamento complementar em ALD.

### Referências Gerais

- ALM E., ARKIN A. P. Biological networks. **Current Opinion in Structural Biology**, Vol. 13, n. 2, p. 193-202, abr. 2003
- ARGMANN, C.A.; HOUTEN, S. M.; ZHU, J.; SCHADT, Eric E. A Next Generation Multiscale View of Inborn Errors of Metabolism. **Cell Metabolism**, vol. 23, n. 1, p. 13-26, 12 jan. 2016
- AUBOURG Patrick, WANDERS Ronald. Peroxisomal disorders. *In*: DULAC, O.; LASSONDE ,M.; SARNAT, H.B. (Ed.). **Handbook of Clinical Neurology**, Vol. 113, Elsevier, 2013. p. 1593-1609
- BADER, G. D.; CHRISTOPHER, W. V. H. An automated method for finding molecular complexes in large protein interaction networks. **BMC bioinformatics**, vol. 4, 13 jan 2003
- FATOUH, M.. Is Lorenzo's oil effective for the treatment of children with X-linked adrenoleukodystrophy? A systematic review of the literature. **Archives of Disease in Childhood**, vol. 107, ago. 2022
- KITANO, H. Computational systems biology. **Nature**, vol.420, 14 nov 2002
- MOSER, A. B.; JONES, R. O.; HUBBARD, W. C. et al. Newborn screening for X-Linked adrenoleukodystrophy. **International Journal of Neonatal Screening**, vol. 2, n. 4, dez. 2016
- PINTO, Wagner de Jesus. **Bioquímica clínica** / Wagner de Jesus Pinto. 1. ed. Rio de Janeiro : Guanabara Koogan, 2017.
- POLL-THE , B.T.; AUBOURG, P.; WANDERS, R.. Peroxisomal disorders. *In*: FERNANDES, J.; SAUDUBRAY, J.M.; van den BERGHE, G.; WALTER, J.H (Ed.). **Inborn Metabolic Diseases**, Springer, 4ª Ed., 2006. p. 509-521
- RICHMOND, P. et al. Multi-omic Approach to Identify Phenotypic Modifiers Underlying Cerebral Demyelination in X-Linked Adrenoleukodystrophy. *Front. Cell Dev. Biol.*, vol.8, 25 Jun 2020
- SAUDUBRAY, Jean-Marie; DESGUERRE, Isabelle; SEDEL, Frédéric; CHARPENTIER, Christiane. Capítulo 1 -A Clinical Approach to Inherited Metabolic Diseases, **Inborn Metabolic Diseases**. Springer, 4ª Ed., 2006
- SCARDONI, G.; TOSADORI, G.; FAIZAN, M. et al. Biological network analysis with CentiScaPe: centralities and experimental dataset integration. *F1000Res*; 1 jul 2014
- SILVA, G.C. V.; MATTE, U. Neuronetworks: Analysis of brain pathology in Mucopolysaccharidoses - A systems biology approach. *Neuroscience Informatics*, vol.2, 1, Mar 2020
- TURK, B. R.; THEDA, C.; FATEMI, Ali; MOSER, A. B. X-linked adrenoleukodystrophy: Pathology, pathophysiology, diagnostic testing, newborn screening and therapies. **Int J Dev Neurosci**, vol. 80, n. 1, p. 52-72, 22 jan. 2020

## Anexo A - Normas de Publicação da Revista *Metabolic Brain Diseases*

### Instructions for Authors

#### Types of papers

Metabolic Brain Disease is committed to high standards of presentation and will consider:

Research Papers: No page limitations but typical length is 6000 words with 60 references and a total of 8 figures and tables.

Short Communications: Usually, 1500 words with no more than 2 tables or figures.

Review Articles: No page limitations but typical length is 6000 words.

Editorials: Approximately 1,000 words with 10 references.

Please contact the Editor to discuss the suitability of topics.

#### Title Page

Please make sure your title page contains the following information.

##### Title

The title should be concise and informative.

##### Author information

- The name(s) of the author(s)
- The affiliation(s) of the author(s), i.e. institution, (department), city, (state), country
- A clear indication and an active e-mail address of the corresponding author
- If available, the 16-digit ORCID of the author(s)

If address information is provided with the affiliation(s) it will also be published.

For authors that are (temporarily) unaffiliated we will only capture their city and country of residence, not their e-mail address unless specifically requested.

#### Abstract

Please provide an abstract of 150 to 250 words. The abstract should not contain any undefined abbreviations or unspecified references.

*For life science journals only (when applicable)*

- Trial registration number and date of registration for prospectively registered trials
- Trial registration number and date of registration, followed by “retrospectively registered”, for retrospectively registered trials

#### Keywords

Please provide 4 to 6 keywords which can be used for indexing purposes.

#### Text

##### Text Formatting

Manuscripts should be submitted in Word.

- Use a normal, plain font (e.g., 10-point Times Roman) for text.
- Use italics for emphasis.
- Use the automatic page numbering function to number the pages.
- Do not use field functions.
- Use tab stops or other commands for indents, not the space bar.
- Use the table function, not spreadsheets, to make tables.
- Use the equation editor or MathType for equations.

- Save your file in docx format (Word 2007 or higher) or doc format (older Word versions).

Manuscripts with mathematical content can also be submitted in LaTeX. We recommend using [Springer Nature's LaTeX template](#).

### **Headings**

Please use no more than three levels of displayed headings.

### **Abbreviations**

Abbreviations should be defined at first mention and used consistently thereafter.

### **Footnotes**

Footnotes can be used to give additional information, which may include the citation of a reference included in the reference list. They should not consist solely of a reference citation, and they should never include the bibliographic details of a reference. They should also not contain any figures or tables.

Footnotes to the text are numbered consecutively; those to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data).

Footnotes to the title or the authors of the article are not given reference symbols.

Always use footnotes instead of endnotes.

### **Acknowledgments**

Acknowledgments of people, grants, funds, etc. should be placed in a separate section on the title page. The names of funding organizations should be written in full.

### **Scientific style**

Generic names of drugs and pesticides are preferred; if trade names are used, the generic name should be given at first mention.

### **References**

#### **Citation**

Cite references in the text by name and year in parentheses. Some examples:

- Negotiation research spans many disciplines (Thompson 1990).
- This result was later contradicted by Becker and Seligman (1996).
- This effect has been widely studied (Abbott 1991; Barakat et al. 1995a, b; Kelso and Smith 1998; Medvec et al. 1999, 2000).

#### **Reference list**

The list of references should only include works that are cited in the text and that have been published or accepted for publication. Personal communications and unpublished works should only be mentioned in the text.

Reference list entries should be alphabetized by the last names of the first author of each work. Please alphabetize according to the following rules: 1) For one author, by name of author, then chronologically; 2) For two authors, by name of author, then name of coauthor, then chronologically; 3) For more than two authors, by name of first author, then chronologically.

If available, please always include DOIs as full DOI links in your reference list (e.g. "https://doi.org/abc").

- Journal article

Gamelin FX, Baquet G, Berthoin S, Thevenet D, Nourry C, Nottin S, Bosquet L (2009) Effect of high intensity intermittent training on heart rate variability in prepubescent children. *Eur J Appl Physiol* 105:731-738.  
<https://doi.org/10.1007/s00421-008-0955-8>

Ideally, the names of all authors should be provided, but the usage of “et al” in long author lists will also be accepted:

Smith J, Jones M Jr, Houghton L et al (1999) Future of health insurance. *N Engl J Med* 965:325–329

- Article by DOI

Slifka MK, Whitton JL (2000) Clinical implications of dysregulated cytokine production. *J Mol Med.* <https://doi.org/10.1007/s001090000086>

- Book

South J, Blass B (2001) *The future of modern genomics.* Blackwell, London

- Book chapter

Brown B, Aaron M (2001) The politics of nature. In: Smith J (ed) *The rise of modern genomics*, 3rd edn. Wiley, New York, pp 230-257

- Online document

Cartwright J (2007) Big stars have weather too. IOP Publishing PhysicsWeb. <http://physicsweb.org/articles/news/11/6/16/1>. Accessed 26 June 2007

- Dissertation

Trent JW (1975) *Experimental acute renal failure.* Dissertation, University of California

Always use the standard abbreviation of a journal’s name according to the ISSN List of Title Word Abbreviations, see

ISSN LTWA

If you are unsure, please use the full journal title.

### Statements & Declarations

The following statements must be included in your submitted manuscript under the heading 'Statements and Declarations'. This should be placed after the References section. Please note that submissions that do not include required statements will be returned as incomplete.

### Funding

Please describe any sources of funding that have supported the work. The statement should include details of any grants received (please give the name of the funding agency and grant number).

Example statements:

*“This work was supported by [...] (Grant numbers [...] and [...]). Author A.B. has received research support from Company A.”*

*“The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.”*

### **Competing Interests**

Authors are required to disclose financial or non-financial interests that are directly or indirectly related to the work submitted for publication. Interests within the last 3 years of beginning the work (conducting the research and preparing the work for submission) should be reported. Interests outside the 3-year time frame must be disclosed if they could reasonably be perceived as influencing the submitted work.

Example statements:

*“Financial interests: Author A and B declare they have no financial interests. Author C has received speaker and consultant honoraria from Company M. Dr. C has received speaker honorarium and research funding from Company M and Company N. Author D has received travel support from Company O. Non-financial interests: Author D has served on advisory boards for Company M and Company N.”*

*“The authors have no relevant financial or non-financial interests to disclose.”*

Please refer to the “Competing Interests” section below for more information on how to complete these sections.

### **Author Contributions**

Authors are encouraged to include a statement that specifies the contribution of every author to the research and preparation of the manuscript.

Example statement:

*“All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [full name], [full name] and [full name]. The first draft of the manuscript was written by [full name] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.”*

Please refer to the “Authorship Principles” section below for more information on how to complete this section.

### **Data Availability**

All original research must include a Data Availability Statement. Data Availability Statements should provide information on where data supporting the results reported in the article can be found. Statements should include, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. When it is not possible to share research data publicly, for instance when individual privacy could be compromised, data availability should still be stated in the manuscript along with any conditions for access.

Example statements:

*“The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT LINK TO DATASETS]”*

*“The datasets generated during and/or analysed during the current study are not publicly available due to [REASON(S) WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.”*

In addition to the above, manuscripts that report the results of studies involving humans and/or animals should include the following declarations:

### **Ethics approval**

Authors of research involving human or animal subjects should include a statement that confirms that the study was approved (or granted exemption) by the appropriate institutional and/or national research ethics committee (including the name of the ethics committee and reference number, if available). For research involving animals, their data or biological material, authors should supply detailed information on the ethical treatment of their animals in their submission. If a study was granted exemption or did not require ethics approval, this should also be detailed in the manuscript.

*“This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of University B (Date.../No...).”*

*“This is an observational study. The XYZ Research Ethics Committee has confirmed that no ethical approval is required.”*

### **Consent to participate**

For all research involving human subjects, freely-given, informed consent to participate in the study must be obtained from participants (or their parent or legal guardian in the case of children under 16) and a statement to this effect should appear in the manuscript.

Example statement:

*“Informed consent was obtained from all individual participants included in the study.”*

*“Written informed consent was obtained from the parents.”*

### **Consent to publish**

Individuals may consent to participate in a study, but object to having their data published in a journal article. If your manuscript contains any individual person’s data in any form (including any individual details, images or videos), consent for publication must be obtained from that person, or in the case of children, their parent or legal guardian. This is in particular applicable to case studies. A statement confirming that consent to publish has been received from all participants should appear in the manuscript.

Example statement:

*“The authors affirm that human research participants provided informed consent for publication of the images in Figure(s) 1a, 1b and 1c.”*

### **Tables**

- All tables are to be numbered using Arabic numerals.
- Tables should always be cited in text in consecutive numerical order.
- For each table, please supply a table caption (title) explaining the components of the table.

- Identify any previously published material by giving the original source in the form of a reference at the end of the table caption.
- Footnotes to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data) and included beneath the table body.

## **Artwork and Illustrations Guidelines**

### **Electronic Figure Submission**

- Supply all figures electronically.
- Indicate what graphics program was used to create the artwork.
- For vector graphics, the preferred format is EPS; for halftones, please use TIFF format. MSOffice files are also acceptable.
- Vector graphics containing fonts must have the fonts embedded in the files.
- Name your figure files with "Fig" and the figure number, e.g., Fig1.eps.
- Definition: Black and white graphic with no shading.
- Do not use faint lines and/or lettering and check that all lines and lettering within the figures are legible at final size.

### **Line Art**

- All lines should be at least 0.1 mm (0.3 pt) wide.
- Scanned line drawings and line drawings in bitmap format should have a minimum resolution of 1200 dpi.
- Vector graphics containing fonts must have the fonts embedded in the files.
- Definition: Photographs, drawings, or paintings with fine shading, etc.
- If any magnification is used in the photographs, indicate this by using scale bars within the figures themselves.
- Halftones should have a minimum resolution of 300 dpi.

### **Combination Art**

- Definition: a combination of halftone and line art, e.g., halftones containing line drawing, extensive lettering, color diagrams, etc.
- Combination artwork should have a minimum resolution of 600 dpi.

### **Color Art**

- Color art is free of charge for online publication.
- If black and white will be shown in the print version, make sure that the main information will still be visible. Many colors are not distinguishable from one another when converted to black and white. A simple way to check this is to make a xerographic copy to see if the necessary distinctions between the different colors are still apparent.
- If the figures will be printed in black and white, do not refer to color in the captions.
- Color illustrations should be submitted as RGB (8 bits per channel).

### **Figure Lettering**

- To add lettering, it is best to use Helvetica or Arial (sans serif fonts).
- Keep lettering consistently sized throughout your final-sized artwork, usually about 2–3 mm (8–12 pt).
- Variance of type size within an illustration should be minimal, e.g., do not use 8-pt type on an axis and 20-pt type for the axis label.



- Avoid effects such as shading, outline letters, etc.
- Do not include titles or captions within your illustrations.

### **Figure Numbering**

- All figures are to be numbered using Arabic numerals.
- Figures should always be cited in text in consecutive numerical order.
- Figure parts should be denoted by lowercase letters (a, b, c, etc.).
- If an appendix appears in your article and it contains one or more figures, continue the consecutive numbering of the main text. Do not number the appendix figures, "A1, A2, A3, etc." Figures in online appendices [Supplementary Information (SI)] should, however, be numbered separately.

### **Figure Captions**

- Each figure should have a concise caption describing accurately what the figure depicts. Include the captions in the text file of the manuscript, not in the figure file.
- Figure captions begin with the term Fig. in bold type, followed by the figure number, also in bold type.
- No punctuation is to be included after the number, nor is any punctuation to be placed at the end of the caption.
- Identify all elements found in the figure in the figure caption; and use boxes, circles, etc., as coordinate points in graphs.
- Identify previously published material by giving the original source in the form of a reference citation at the end of the figure caption.

### **Figure Placement and Size**

- Figures should be submitted within the body of the text. Only if the file size of the manuscript causes problems in uploading it, the large figures should be submitted separately from the text.
- When preparing your figures, size figures to fit in the column width.
- For large-sized journals the figures should be 84 mm (for double-column text areas), or 174 mm (for single-column text areas) wide and not higher than 234 mm.
- For small-sized journals, the figures should be 119 mm wide and not higher than 195 mm.

### **Permissions**

If you include figures that have already been published elsewhere, you must obtain permission from the copyright owner(s) for both the print and online format. Please be aware that some publishers do not grant electronic rights for free and that Springer will not be able to refund any costs that may have occurred to receive these permissions. In such cases, material from other sources should be used.

### **Accessibility**

In order to give people of all abilities and disabilities access to the content of your figures, please make sure that

- All figures have descriptive captions (blind users could then use a text-to-speech software or a text-to-Braille hardware)
- Patterns are used instead of or in addition to colors for conveying information (colorblind users would then be able to distinguish the visual elements)
- Any figure lettering has a contrast ratio of at least 4.5:1

### **Supplementary Information (SI)**

Springer accepts electronic multimedia files (animations, movies, audio, etc.) and other supplementary files to be published online along with an article or a book chapter. This feature can add dimension to the author's article, as certain information cannot be printed or is more convenient in electronic form.

Before submitting research datasets as Supplementary Information, authors should read the journal's Research data policy. We encourage research data to be archived in data repositories wherever possible.

### **Submission**

- Supply all supplementary material in standard file formats.
- Please include in each file the following information: article title, journal name, author names; affiliation and e-mail address of the corresponding author.
- To accommodate user downloads, please keep in mind that larger-sized files may require very long download times and that some users may experience other problems during downloading.
- High resolution (streamable quality) videos can be submitted up to a maximum of 25GB; low resolution videos should not be larger than 5GB.

### **Audio, Video, and Animations**

- Aspect ratio: 16:9 or 4:3
- Maximum file size: 25 GB for high resolution files; 5 GB for low resolution files
- Minimum video duration: 1 sec
- Supported file formats: avi, wmv, mp4, mov, m2p, mp2, mpg, mpeg, flv, mxf, mts, m4v, 3gp

### **Text and Presentations**

- Submit your material in PDF format; .doc or .ppt files are not suitable for long-term viability.
- A collection of figures may also be combined in a PDF file.

### **Spreadsheets**

- Spreadsheets should be submitted as .csv or .xlsx files (MS Excel).

### **Specialized Formats**

- Specialized format such as .pdb (chemical), .wrl (VRML), .nb (Mathematica notebook), and .tex can also be supplied.

### **Collecting Multiple Files**

- It is possible to collect multiple files in a .zip or .gz file.

### **Numbering**

- If supplying any supplementary material, the text must make specific mention of the material as a citation, similar to that of figures and tables.
- Refer to the supplementary files as "Online Resource", e.g., "... as shown in the animation (Online Resource 3)", "... additional data are given in Online Resource 4".
- Name the files consecutively, e.g. "ESM\_3.mpg", "ESM\_4.pdf".

### **Captions**

- For each supplementary material, please supply a concise caption describing the content of the file.

### **Processing of supplementary files**

- Supplementary Information (SI) will be published as received from the author without any conversion, editing, or reformatting.

**Accessibility**

In order to give people of all abilities and disabilities access to the content of your supplementary files, please make sure that

- The manuscript contains a descriptive caption for each supplementary material
- Video files do not contain anything that flashes more than three times per second (so that users prone to seizures caused by such effects are not put at risk)