An optimized pipeline for extracting transcriptomic data from The Human Protein Atlas: insights from normalized MCAD mRNA expression analysis

Clara Beltrão Valente[1,2], Elen Azambuja Lima[3], Letícia Soares de Souza[1], Gustavo da Costa Ferreira[2], Patrícia Fernanda Schuck[2], Hércules Rezende Freitas[1,2]*

[1]Centro Universitário IBMR, Escola de Ciências da Saúde, Av. das Américas, 2603, 22631-002, Rio de Janeiro/RJ, Brasil; [2]Universidade Federal do Rio de Janeiro, Centro de Ciências da Saúde, Laboratório de Erros Inatos do Metabolismo, Av. Carlos Chagas Filho, 373, 21941-599, Rio de Janeiro/RJ, Brasil; [3]Centro Universitário UNA, Escola de Ciências da Saúde, Av. Gov. Valadares, 640, 32510-010, Goiás/GO, Brasil.

ORCID:

Clara Beltrão Valente - 0000-0002-8782-4138

Elen Azambuja Lima - 0009-0008-7719-0558

Letícia Soares de Souza - 0009-0007-3007-742X

Gustavo da Costa Ferreira - 0000-0002-2892-884

Patrícia Fernanda Schuck - 0000-0003-3148-4952

Hércules Rezende Freitas - 0000-0003-1584-9157

Keywords: ACADM, Inborn Errors of Metabolism, Metabolic diseases, MCAD, Neuropathies.

Running title: A pipeline for transcriptomic data extraction.

AABC section: Original article.

Corresponding author: Hércules Rezende Freitas. Universidade Federal do Rio de Janeiro, Centro de Ciências da Saúde, Laboratório de Erros Inatos do Metabolismo, Av. Carlos Chagas Filho, 373, 21941-599, Rio de Janeiro/RJ, Brasil. Phone: +55-21-986122194 / e-mail: hercules.freitas@bioqmed.ufrj.br.

Abstract

Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency is an autosomal recessive disorder classified under Fatty Acid Oxidation Disorders. Individuals with MCAD deficiency exhibit a diminished ability to oxidize medium-chain fatty acids, leading to the accumulation of lipids and their derivatives in blood and various tissues. In this study, we developed an optimized computational pipeline alongside a Shiny application to efficiently extract and analyze mRNA expression data from The Human Protein Atlas. Utilizing this tool, we assessed the normalized mRNA levels of the ACADM gene across multiple tissues and brain structures. The extracted MCAD expression data were systematically collected, categorized, and analyzed, providing comprehensive insights into the gene's expression patterns. Our findings contribute to a deeper understanding of the pathophysiology of MCAD deficiency and demonstrate the effectiveness of the developed pipeline and Shiny app in facilitating transcriptomic data analysis.

Keywords: ACADM, Inborn Errors of Metabolism, Metabolic diseases, MCAD, Neuropathies.

Introduction

Medium-chain acyl-coenzyme A dehydrogenase deficiency (MCADD) is a metabolic disorder that impairs the initial step of beta-oxidation, resulting in reduced energy production. This condition is characterized by the accumulation of medium-chain acyl-CoAs and their corresponding acylcarnitines. MCADD is commonly identified through neonatal screening and selective metabolic diagnosis. The MCAD enzyme is encoded by the ACADM gene, and various pathogenic variants have been described, with an additional number of variants having uncertain significance (Pugliese et al. 2020).

Patients diagnosed with MCADD are susceptible to experiencing acute illnesses or injuries. These individuals are specifically prone to acute metabolic decompensation (AMD) due to the interruption of their fatty acid metabolism. In instances where calorie intake is suspended for a significant period, individuals with MCADD face the risk of developing hypoglycemia. In the context of elective surgeries, for instance, patients must ingest glucose-containing fluids (at least 10% dextrose) orally up to four hours before the operation. Similarly, patients with acute metabolic exacerbation require intravenous fluids containing at least 10% dextrose at no less than 1.5 times the maintenance rate to provide an infusion of 10 mg/kg/min (McGregor et al. 2021).

MCADD is also a potential factor in sudden neonatal death (SND) (Mütze et al. 2022). Non-diagnosed patients in the neonatal period, infancy, and early childhood may undergo AMD and, if not properly treated, suffer from a high mortality rate (Brackett et al. 1994). Previous research has highlighted the importance of early MCADD diagnosis in preventing SND (Catzeflis et al. 1990). Early newborn screening and awareness of clinical

symptoms are crucial for early diagnosis and treatment, which can be life-saving (Lovera et al. 2012). Poor feeding, lack of blood glucose monitoring, and homozygosity of the common gene variant (c.985A > G) are major risk factors for fatal neonatal crisis in MCAD deficiency (MCAD deficiency: to be, or not to be at risk 2014). Post-mortem diagnostic protocols are necessary for correct diagnosis and counseling of the family in cases of unexpected death in the neonatal period (Kazemi et al. 2022).

Unfortunately, delayed or imprecise diagnosis is still a determining factor in MCADD-induced SND (Salim & Ng 2023). Therefore, it is imperative that further efforts are aimed towards identifying key hallmarks of the deficiency, including biochemical and genetic traits that may prove relevant for faster diagnoses. Additionally, the search for new therapeutic options for MCADD is crucial, as current treatments are limited. Experimental drugs like glycerol phenylbutyrate show promise in addressing specific symptoms, yet challenges in trial design hinder comprehensive evaluations. Gene therapy, with its potential for precise targeting of mutations, offers an exciting avenue for treatment development. The urgent need for expanded research, including the incorporation of core outcome measures in clinical trials, underscores the significance of advancing therapeutic options to alleviate the burden of MCAD deficiency on affected individuals and the healthcare system (Mason, Hindmarch & Dunham-Snary 2023).

In this context, the development of an optimized computational pipeline and a user-friendly Shiny application plays a pivotal role in advancing our understanding of MCAD. This study leverages these tools to comprehensively investigate MCAD by evaluating ACADM mRNA expression across diverse human tissues and distinct regions of the human brain using data extracted from The Human Protein Atlas. Furthermore, the

pipeline facilitates the normalization and analysis of transcriptomic data, enabling the exploration of developmental correlations between ACADM mRNA expression and the activity of other crucial human genes. The insights gained from this study not only contribute to a deeper understanding of MCAD deficiency but also hold significant implications for future advancements in treatment and diagnostic strategies. By elucidating the intricate regulatory mechanisms involved in ACADM expression, our findings aim to pave the way for more targeted and effective approaches in the management of MCAD deficiency, ultimately enhancing the prospects for improved patient outcomes.

Materials and methods

Analysis of ACADM gene expression

Expression data for the ACADM gene in both the central nervous system and peripheral tissues were retrieved from The Human Protein Atlas (HPA) (Sjöstedt et al. 2020), with measurements expressed in normalized transcripts per million (nTPM). The independent variables selected for analysis included species, tissue type or specific brain regions, sex, and developmental stage. To ensure reproducibility and efficient data extraction, a custom Python-based pipeline and a R Shiny application were developed.

Python pipeline for mRNA data extraction

This pipeline utilizes the HPA-provided XML data format and incorporates key Python libraries such as Pandas for data manipulation, ElementTree for XML parsing, and Itertools for efficient iteration. The pipeline consists of two primary functions, with the first allowing for the extraction of nTPM values from the ACADM gene across specified brain regions or peripheral tissues based on the assayType attribute (either "humanBrain" or "tissue"), and a second function, which orchestrates the parsing process by iterating over all relevant tissue or brain region entries and consolidating the extracted data into a single Pandas DataFrame. The final dataset is exported to an Excel file, facilitating downstream analysis. The complete Python script, named “HPA_XML_Parser.py”, is publicly available on GitHub (https://github.com/drhrf/XML_HPA_parser) under the GPL-3.0 license, ensuring accessibility and transparency for reproducibility purposes.

To further enhance accessibility and ease of use, a Google Colaboratory[1] notebook was developed to implement the Python-based pipeline, enabling "low code" online utilization of the data extraction process. This notebook incorporates the complete Python script, allowing users to execute the pipeline directly in their web browsers without the need for extensive local setup. By leveraging the interactive environment of Google Colab, researchers can efficiently manipulate and analyze gene expression data with minimal coding effort. Additionally, the Supplementary Figure 1 provides an overview of the workflow, indicating each step from the initial extraction of the XML file to the subsequent data analysis performed using R.

Shiny application

[1] https://colab.research.google.com/github/drhrf/XML_HPA_parser/blob/main/HPA_XML_Parser.ipynb

In addition to the Python pipeline, an interactive R Shiny application was developed to facilitate the exploration, visualization, and download of the processed expression data. The application features a user-friendly interface built with Shiny dashboard, providing a structured and navigable layout. Users can upload XML files and initiate the parsing process through the "Upload & Parse" section. Once parsed, the data is displayed in the "humanBrain Data" and "Tissue Data" tabs as interactive DataTables, allowing for easy browsing and examination of the expression patterns. The "Download" tab enables users to export the consolidated data as an Excel file, segregated into separate worksheets for human brain and tissue data. Additionally, the application includes a "Select Theme" feature powered by Shiny themes, offering dynamic customization of the interface to enhance user experience (Supplementary Figure 2).

The server-side logic handles file uploads, triggers the parsing functions, renders the data tables, and manages the download functionality, ensuring seamless interaction with the data. The complete R script, named “app.R”, is also publicly available on the GitHub repository (https://github.com/drhrf/XML_HPA_parser).

Statistical analysis

Statistical analysis of the expression data was conducted using both graphical and statistical methods to assess data distribution. Histograms and quantile-quantile plots were employed to visually inspect the normality of the data. To quantitatively evaluate normality, the two-sided asymptotic Kolmogorov-Smirnov test and the D’Agostino

skewness test were utilized. These tests revealed significant deviations from normality in the expression variables, as confirmed by both graphical assessments and statistical evaluations. Consequently, nonparametric methods were adopted for subsequent analyses to ensure robustness. The Mann-Whitney U test was applied for comparisons between two independent groups, maintaining an alpha level of 0.05 to control the Type I error rate. For comparisons involving more than two groups, the Kruskal-Wallis test was employed, with Dunn's test for *post-hoc* comparisons, under the same significance threshold.

Results

Expression of ACADM in the cerebral cortex

Initially, we've investigated whether ACADM expression varies due to age (61-94 years, Figure 1a) or cortical hemisphere (Figure 1b). Our findings show that gene expression remains centered in the 20 to 30 nTPM (3 to 3.4 log nTPM) range, regardless of the variable analyzed.

[Figure 1. Cortical ACADM mRNA expression by age and hemisphere.]

When the cortical expression data were segmented by sub-region, on the other hand, areas of consistently higher (> 25) and lower-than-median (< 22) nTPM emerged (Figure 2a, shown as log nTPM). As shown in Figure 2b, the brain regions with higher

expression are the frontotemporal cortex, medial orbitofrontal gyrus, superior orbitofrontal gyrus, and subcallosal gyrus. On the other hand, Figure 2c shows lower ACADM expression in the lateral orbitofrontal gyrus, posterior orbitofrontal gyrus, inferior frontal triangular gyrus, and the subgenual anterior cingulate cortex.

[Figure 2. Normalized ACADM mRNA expression in human cortical areas.]

Expression data from all cortical areas were divided according to sex for further analysis. As shown in Figure 3a, ACADM expression is higher for females in the superior frontal gyrus and the opercular inferior frontal gyrus, while it is higher for males in the ventrolateral prefrontal cortex. No differences were found for overall cortical expression was analyzed (Figure 3b).

[Figure 3. Differential cortical ACADM expression by sex.]

MCAD expression in peripheral tissues

Here, we also investigated whether the variation in ACADM expression across peripheral tissues reflects the metabolic demands and the role of MCAD in fatty acid oxidation. High expression in tissues with high oxidative energy requirements, such as the heart and skeletal muscle, would be consistent with the known function of MCAD in beta-oxidation of fatty acids. Figure 4a indicates a higher expression (> 200 nTPM) of ACADM can be found in the heart muscle, kidney, skeletal muscle, liver, and tongue (only

one data point). Only a limited number of tissues had enough data available on both sexes to allow for comparison. Despite the absolute differences seen on Figures 4B and 4C, however, none of the tests allowed for the rejection of the null hypothesis that both samples derive from the same distribution.

[Figure 4. ACADM expression by tissue and sex.]

Discussion

Cortical expression of the ACADM gene is similar across age groups and cortical hemispheres

As the brain ages, mitochondrial function becomes increasingly compromised, leading to cognitive decline and contributing to neurodegenerative diseases such as Alzheimer's and Parkinson's (Yin et al. 2014). Research has shown that aging is associated with significant changes in mitochondrial gene expression, oxidative stress levels, and energy metabolism within the brain (Manczak et al. 2005, Reutzel et al. 2020). Studies using mouse models have provided insights into how mitochondrial gene expression evolves over time. For instance, Manczak et al. (2005) examined C57BL6 mice and found that mitochondrial genes associated with the respiratory chain complexes I, III, IV, and V are upregulated in middle-aged mice (12 and 18 months old) compared to younger mice (2 months old). This upregulation suggests a compensatory mechanism aimed at maintaining energy production despite the onset of age-related challenges.

However, in older mice (24 months old), the expression of these genes declines, indicating that the compensatory responses are not sustainable in advanced age (Manczak et al. 2005). This decline is accompanied by increased markers of oxidative damage, such as 8-hydroxyguanosine (8-OHG), and the release of cytochrome c, which is involved in the apoptotic pathway.

Longitudinal studies have further demonstrated that the decline in mitochondrial function correlates with reductions in cognitive performance. Reutzel et al. (2020) conducted a study on NMRI mice over a two-year period and observed significant decreases in brain ATP levels and mitochondrial respiration starting at 18 months of age. These metabolic impairments were paralleled by diminished expression of genes involved in mitochondrial biogenesis and antioxidant defense mechanisms. Notably, genes such as cAMP response element-binding protein 1 (creb-1), peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC1-α), and mitochondrial transcription factor A (Tfam) showed decreased expression, highlighting a reduction in the brain's capacity to generate new mitochondria and combat oxidative stress (Reutzel et al. 2020).

Computational models have been employed to unravel the complex interactions within brain energy metabolism during aging. Shichkova and colleagues developed a model encompassing over 66,000 molecular interactions, which identified that the generation of action potentials is primarily impaired due to reduced expression of the Na^+/K^+-ATPase pump and a decreased supply of ATP. This impairment leads to a loss of metabolic flexibility, rendering the brain less capable of responding to stimuli or repairing molecular damage. The model also suggested that astrocytes might sacrifice their own energy stability to support neurons, indicating a systemic imbalance in energy distribution

within the aging brain (Shichkova et al. 2023). Proteomic analyses of brain microvessels have shed light on the molecular changes occurring at the vascular level during aging. Previous research has reported that key proteins involved in oxidative stress response, such as superoxide dismutase-1 and -2, catalase, and thioredoxin, decrease in abundance with age. Additionally, proteins that stabilize mRNA and prevent degradation were reduced, potentially leading to decreased protein synthesis essential for maintaining mitochondrial function. The reduction in glycolytic enzymes and components of mitochondrial complexes I–V suggests a compromised capacity for ATP production, further exacerbating energy deficits in the aging brain (Chandra et al. 2022).

Despite accumulating evidence of numerous metabolic adaptations in the aging brain, recent research has demonstrated that not all mitochondrial enzymes are equally affected by the senescence. Specifically, enzymes like the very long chain acyl coenzyme A dehydrogenase (VLCAD) and the α-ketoglutarate dehydrogenase maintain stable activity levels despite advancing age (Yarian et al. 2005). In this study, we observed that MCAD mRNA expression remains stable across the aging range of 61 to 94 years (Figure 1a). With advancements in diagnosis and treatment, individuals with MCADD are anticipated to live longer and healthier lives. Our data indicate that, at least for unaffected individuals, ACADM expression does not significantly change in later life. However, it is important to acknowledge that previous research has demonstrated that factors such as prenatal protein malnutrition can reduce MCAD mRNA and protein expression in aging rats (Erhuma et al. 2007). These factors should also be considered when evaluating late-life risks for MCADD patients.

Lateralization of gene expression between brain hemispheres is an important evolutionary feature, involving dozens of candidate genes with potential control over the cortical asymmetry (Sun et al. 2006). Despite the knowledge of developmentally important lateralized genes, only recent large-scale studies have been able to identify the subtle transcriptional variations with a functional role in the adult cortex (Francks 2015). One important asymmetry in the brain is metabolism. A positron emission tomography (PET) study has shown that cerebral glucose metabolism is greater in the left medial frontal gyrus, posterior thalamus, lingual gyrus, cuneus and superior cingulate in healthy adults (males and females, 20-69 years), with opposite laterality in the mesio-anterior cerebellum, and lateral frontal and temporal regions (Willis et al. 2002). Our data, as shown in Figure 1b, indicate that ACADM expression is similar in both cortical hemispheres. Additionally, when the nTPM counts of each cortical region was compared by hemisphere, no differences were detected (Supplementary Figure 3).

ACADM is highly expressed in cortical regions associated with cognitive processing

When the cortical expression data were segmented by sub-region, on the other hand, areas of consistently higher (> 25) and lower-than-median (< 22) nTPM emerged (Figure 2a). As shown in Figure 2b, the brain regions with higher expression are the frontotemporal cortex, medial orbitofrontal gyrus, superior orbitofrontal gyrus, and subcallosal gyrus, which are strongly linked to cognitive, emotional, and decision-making processes (Fettes et al. 2017). The frontotemporal cortex is associated with cognitive and emotional processing. Evidence suggests that frontotemporal alterations manifest in

psychotic episodes of patients diagnosed with schizophrenia (Gutiérrez-Galve et al. 2010). Both orbitofrontal gyri are associated with functional deficits involving neurodegeneration, such as amyotrophic lateral sclerosis (Ma et al. 2020). The subcallosal gyrus, in turn, acts as a suppressive regulator of amygdala responses to anxiety-promoting environmental stimuli (Vermetten & Lanius 2012).

On the other hand, Figure 2c shows lower ACADM expression in the lateral orbitofrontal gyrus, posterior orbitofrontal gyrus, inferior frontal triangular gyrus, and the subgenual anterior cingulate cortex. The lateral segment of the orbitofrontal cortex appears to be metabolically related to inhibitory control over addictive behaviors (Goldstein et al. 2005). The posterior orbitofrontal segment, in turn, is associated with the processing of threat anticipation and reward (Stawicka et al. 2022). The inferior frontal triangular gyrus, or pars triangularis, holds functional importance in Broca's area, being linked to the comprehension of the propositional language (Foundas et al. 1996). Finally, the subgenual anterior cingulate cortex is a significant regulator of mood; lesions in this area are correlated with depressive manifestations and anhedonia (Rudebeck et al. 2014).

Sex is a determining variable for ACADM expression in the cerebral cortex

Sex-related differences in gene expression can emerge due to hormonal influences, differential gene regulation, or chromosomal differences. For instance, estrogen has been shown to influence gene expression and could potentially affect ACADM (Fu et al. 2009). Recent work has shown that the activation of G protein-coupled

estrogen receptors drives GCN5L1 expression in human cardiomyocytes, which increases the enzymatic activity and acetylation of MCAD (Manning et al. 2022). Also, a recent study showed that ACADM expression was two-fold higher in the skeletal muscle of female mice after an endurance exercise capacity test (Holcomb et al. 2022).

Here, we further analyzed ACADM expression data from the cerebral cortex (Figure 2) to investigate whether sex is a relevant variable for MCAD. As shown in the Supplementary Figure 4, expression is similar in those areas with nTPM counts at around median and lower-than-median values. Cortical regions with higher ACADM expression, on the other hand, show a more heterogenous behavior between males and females. As shown in Figure 3a, females have higher nTPM counts in the superior frontal gyrus and the opercular inferior frontal gyrus, while the ventrolateral prefrontal cortex expression in greater for males. While all areas with detectable differences in nTPM count were high ACDM expression regions, no overall sex differences in cortical expression were detected in our data (Figure 3b).

Energy-demanding tissues express higher levels of MCAD mRNA

Proper MCAD function is fundamental to the oxidative metabolism of highly active tissues. Figure 4a shows that the heart muscle, kidney, skeletal muscle, and liver express more ACADM than any other tissue analyzed (> 200 nTPM). Although the tongue is indicated as the highest expressing sample, only one data point was available from HPA at the time of collection. Here, we were unable to stablish sex as a significant variable for ACADM expression in peripheral tissues since, as shown in Figure 4b, only a small part

of the data could be segmented by sex and further analyzed (with no detectable differences). Finally, females have overall higher nTPM count for peripheral tissues, but not to a meaningful extent (Figure 4c).

Peroxisome proliferator-activated receptor α (PPARα) knockout mice show approximately 60% decrease in heart MCAD protein levels and, for older mice, an almost two-fold increase in glucose transporter 4 (GLUT4) expression (Dodd et al. 2024), indicating the need for compensatory mechanisms in senescence. MCAD is also an important component of heart repair after myocardial infarction, as shown in an ischemia reperfusion rat model (Lei et al. 2021). Rhabdomyolysis, muscle ache, and weakness are a known feature of fatty acid oxidation disorders, including MCADD and VLCAD deficiency (Mason, Hindmarch & Dunham-Snary 2023). Severe abdominal pain, vomiting, muscle ache, and poor oral intake were key symptoms presented by a 17-year-old female diagnosed with MCADD, as recently discussed by Yusuf and colleagues in a case report (Yusuf et al. 2023). The data presented in Figure 4 corroborates the current view that MCAD is a highly expressed, highly active component of muscle tissues, and that a reduction in its activity is the basis of the muscle-related clinical features of MCADD.

MCAD is also a key regulator of the oxidative metabolism in the liver and kidneys. Wang and colleagues have shown that empagliflozin, an antidiabetic medication, upregulated ACADM expression in the liver by activating the AMPK/FOXA2 signaling pathway, thus reducing lipid deposition in nonalcoholic steatohepatitis patients (Wang et al. 2022). Also, fenofibrate, a dyslipidemia medication, was recently shown to reduce triglyceride and lipid accumulation in a diabetic kidney disease directly through the AMPK/FOXA2/MCAD pathway (Tang et al. 2023).

On the other hand, female rats receiving a high-fat-high-sucrose diet showed increased liver oxidative gene expression, including MCAD, thus reinforcing the enzyme's role in the fine-tunning of liver fatty acid metabolism (Jouenne et al. 2023). Finally, a recent work involving liver-specific MCAD knockout in mice has shown that this enzyme is an appetite regulator for medium-chain fatty acids (Maruyama et al. 2024). Taken together, data from our analysis and these works help highlight the importance of MCAD as a highly expressed regulator of eating behavior and liver metabolism, strengthening the need for specific dietary guidance for MCADD patients.

Conclusion

Our study reveals that the expression of the ACADM gene remains stable across different age groups and between cortical hemispheres in the human brain. This stability persists despite the general decline in mitochondrial function and gene expression associated with aging, suggesting that MCAD may play a critical role in maintaining essential metabolic processes in the cortex throughout the lifespan. The consistent expression of ACADM implies that the enzyme's activity is preserved to support the energy demands of cortical regions, particularly those involved in cognitive and emotional processing. Moreover, we observed that ACADM is highly expressed in specific cortical areas such as the frontotemporal cortex, medial orbitofrontal gyrus, superior orbitofrontal gyrus, and subcallosal gyrus. These regions are closely linked to cognitive functions, emotional regulation, and decision-making processes. The elevated expression of

ACADM in these areas underscores its potential importance in facilitating the high metabolic demands associated with complex neural activities.

Sex-specific differences in ACADM expression were also identified in certain cortical regions with higher expression levels. Females exhibited higher expression in the superior frontal gyrus and opercular inferior frontal gyrus, while males showed greater expression in the ventrolateral prefrontal cortex. These findings suggest that hormonal influences may modulate ACADM expression, although no overall sex differences were detected across the cortex. Understanding these nuances is crucial, as it may inform personalized approaches in treating conditions like MCAD deficiency. Additionally, our data confirm that energy-demanding peripheral tissues express significantly higher levels of ACADM. This aligns with the enzyme's established role in fatty acid oxidation and energy metabolism in tissues with high metabolic rates. The prominent expression of ACADM in these organs highlights its systemic importance and corroborates clinical observations of muscle-related symptoms in MCAD deficiency.

Taken together, these findings enhance our understanding of MCAD's role in both central and peripheral energy metabolism. The stable expression of ACADM in the aging cortex, alongside its significant presence in key cognitive regions and energy-demanding tissues, underscores its potential as a therapeutic target. For individuals with MCAD deficiency, these insights emphasize the need to consider factors such as age, sex, and specific brain regions when developing management strategies. Future research should explore the mechanisms underlying the regulation of ACADM expression and activity, which could lead to improved interventions for metabolic and neurodegenerative disorders.

Finally, the development of a “low code” pipeline, as well as a fully interactive application, for extraction of data from the Human Protein Atlas demonstrates their potential for broader use. By enabling researchers to efficiently extract and analyze mRNA expression data, these tools can be applied to study several other genes of special importance to human health, ultimately aiding in the development of targeted therapeutic strategies and personalized medicine approaches.

Acknowledgments

This research was funded by the National Council for Scientific and Technological Development (CNPq, Brazil), grant number 152071/2020-2, the Tess Research Foundation (TRF, USA), Early-Career Investigator Research Grant 2022/2023.

Author contributions

CBV: Writing – Original Draft; Writing – Review & Editing.

EAL: Writing – Original Draft; Writing – Review & Editing.

LSS: Writing – Original Draft; Writing – Review & Editing.

GCF: Funding Acquisition; Supervision; Validation; Writing – Review & Editing.

PFS: Funding Acquisition; Supervision; Validation; Writing – Review & Editing.

HRF: Writing – Original Draft; Writing – Review & Editing; Funding Acquisition; Conceptualization; Data Curation; Supervision; Project Administration; Software; Visualization; Methodology; Investigation; Formal Analysis.

References

BRACKETT JC, SIMS HF, STEINER RD, NUNGE M, ZIMMERMAN EM, DEMARTINVILLE B, RINALDO P, SLAUGH R & STRAUSS AW. 1994. A novel mutation in medium chain acyl-CoA dehydrogenase causes sudden neonatal death. J Clin Invest 94: 1477–1483.

CATZEFLIS C, BACHMANN C, HALE DE, COATES PM, WIESMANN U, COLOMBO JP, JORIS F & DÉLÈZE G. 1990. Early diagnosis and treatment of neonatal medium-chain acyl-CoA dehydrogenase deficiency: Report of two siblings. Eur J Pediatr 149: 577–581.

CHANDRA PK, CIKIC S, RUTKAI I, GUIDRY JJ, KATAKAM PVG, MOSTANY R & BUSIJA DW. 2022. Effects of aging on protein expression in mice brain microvessels: ROS scavengers, mRNA/protein stability, glycolytic enzymes, mitochondrial complexes, and basement membrane components. GeroScience 44: 371–388.

DODD MS, AMBROSE L, BALL V, CLARKE K, CARR CA & TYLER DJ. 2024. The age-dependent development of abnormal cardiac metabolism in the peroxisome proliferator-activated receptor α-knockout mouse. Atherosclerosis 118599.

ERHUMA A, SALTER AM, SCULLEY DV, LANGLEY-EVANS SC & BENNETT AJ. 2007. Prenatal exposure to a low-protein diet programs disordered regulation of lipid metabolism in the aging rat. Am J Physiol Endocrinol Metab 292: E1702-1714.

FETTES P, SCHULZE L & DOWNAR J. 2017. Cortico-Striatal-Thalamic Loop Circuits of the Orbitofrontal Cortex: Promising Therapeutic Targets in Psychiatric Illness. Frontiers in Systems Neuroscience 11.

FOUNDAS AL, LEONARD CM, GILMORE RL, FENNELL EB & HEILMAN KM. 1996. Pars triangularis asymmetry and language dominance. Proc Natl Acad Sci U S A 93: 719–722.

FRANCKS C. 2015. Exploring human brain lateralization with molecular genetics and genomics. Annals of the New York Academy of Sciences 1359: 1–13.

FU MH, MAHER AC, HAMADEH MJ, YE C & TARNOPOLSKY MA. 2009. Exercise, sex, menstrual cycle phase, and 17β-estradiol influence metabolism-related genes in human skeletal muscle. Physiological Genomics 40: 34–47.

GOLDSTEIN RZ, ALIA-KLEIN N, LESKOVJAN AC, FOWLER JS, WANG G-J, GUR RC, HITZEMANN R & VOLKOW ND. 2005. Anger and depression in cocaine addiction: association with the orbitofrontal cortex. Psychiatry Res 138: 13–22.

GUTIÉRREZ-GALVE L, WHEELER-KINGSHOTT CAM, ALTMANN DR, PRICE G, CHU EM, LEESON VC, LOBO A, BARKER GJ, BARNES TRE, JOYCE EM & RON MA. 2010. Changes in the Frontotemporal Cortex and Cognitive Correlates in First-Episode Psychosis. Biological Psychiatry 68: 51–60.

HOLCOMB LE, ROWE P, O'NEILL CC, DEWITT EA & KOLWICZ JR. SC. 2022. Sex differences in endurance exercise capacity and skeletal muscle lipid metabolism in mice. Physiological Reports 10: e15174.

JOUENNE A, HAMICI K, VARLET I, SOURDON J, DAUDÉ P, LAN C, KOBER F, LANDRIER JF, BERNARD M & DESROIS M. 2023. Relationship of cardiac remodeling and perfusion alteration with hepatic lipid metabolism in a prediabetic high fat high sucrose diet female rat model. Biochem Biophys Res Commun 682: 207–215.

KAZEMI T, FIRGAU E, BUNCH D & KAHWASH SB. 2022. Medium-chain acyl-CoA dehydrogenase deficiency (MCADD) precipitating unexpected death in an infant: Report of a case and a brief review of literature. Malays J Pathol 44: 523–526.

LEI I, TIAN S, GAO W, LIU L, GUO Y, TANG P, CHEN E & WANG Z. 2021. Acetyl-CoA production by specific metabolites promotes cardiac repair after myocardial infarction via histone acetylation EMOTO N & BARTON M (Eds). eLife 10: e60311.

LOVERA C, PORTA F, CACIOTTI A, CATARZI S, CASSANELLO M, CARUSO U, GALLINA MR, MORRONE A & SPADA M. 2012. Sudden unexpected infant death (SUDI) in a newborn due to medium chain acyl CoA dehydrogenase (MCAD) deficiency with an unusual severe genotype. Ital J Pediatr 38: 59.

MA X, LU F, CHEN H, HU C, WANG J, ZHANG S, ZHANG S, YANG G & ZHANG J. 2020. Static and dynamic alterations in the amplitude of low-frequency fluctuation in patients with amyotrophic lateral sclerosis. PeerJ 8: e10052.

MANCZAK M, JUNG Y, PARK BS, PARTOVI D & REDDY PH. 2005. Time-course of mitochondrial gene expressions in mice brains: implications for mitochondrial dysfunction, oxidative damage, and cytochrome c in aging. J Neurochem 92: 494–504.

MANNING JR, THAPA D, ZHANG M, STONER MW, SEMBRAT JC, ROJAS M & SCOTT I. 2022. GPER-dependent estrogen signaling increases cardiac GCN5L1 expression. American Journal of Physiology - Heart and Circulatory Physiology 322: H762.

MARUYAMA T, MATSUI S, KOBAYASHI R, HORII T, OGURI Y, TSUZUKI S, HORIE T, ONO K, HATADA I & SASAKI T. 2024. Medium-chain triglyceride-specific appetite is regulated by the β-oxidation of medium-chain fatty acids in the liver. American Journal of Physiology-Endocrinology and Metabolism 326: E735–E746.

MASON E, HINDMARCH CCT & DUNHAM-SNARY KJ. 2023. Medium-chain Acyl-COA dehydrogenase deficiency: Pathogenesis, diagnosis, and treatment. Endocrino Diabet & Metabol 6: e385.

MASON E, HINDMARCH CCT & DUNHAM-SNARY KJ. 2023. Medium-chain Acyl-COA dehydrogenase deficiency: Pathogenesis, diagnosis, and treatment. Endocrinology, Diabetes & Metabolism 6: e385.

MCAD deficiency: to be, or not to be at risk. 2014. Groningen: Rijksuniversiteit Groningen.

MCGREGOR TL, BERRY SA, DIPPLE KM, HAMID R, COUNCIL ON GENETICS, CHEN E, TROTTER TL, BURKE LW, GELESKE TA, HOPKIN RJ, INTRONE WJ, LYONS MJ, SCHEUERLE AE, STOLER JM, FREEDENBERG DL & JONES MC. 2021. Management Principles for Acute Illness in Patients With Medium-Chain Acyl-Coenzyme A Dehydrogenase Deficiency. Pediatrics 147: e2020040303.

MÜTZE U, NENNSTIEL U, ODENWALD B, HAASE C, CEGLAREK U, JANZEN N, GARBADE SF, HOFFMANN GF, KÖLKER S & HAAS D. 2022. Sudden neonatal death in individuals with medium-chain acyl-coenzyme A dehydrogenase deficiency: limit of newborn screening. Eur J Pediatr 181: 2415–2422.

PUGLIESE M et al. 2020. Outcomes in pediatric studies of medium-chain acyl-coA dehydrogenase (MCAD) deficiency and phenylketonuria (PKU): a review. Orphanet J Rare Dis 15: 12.

REUTZEL M, GREWAL R, DILBERGER B, SILAIDOS C, JOPPE A & ECKERT GP. 2020. Cerebral Mitochondrial Function and Cognitive Performance during Aging: A Longitudinal Study in NMRI Mice. Oxidative Medicine and Cellular Longevity 2020: 4060769.

RUDEBECK PH, PUTNAM PT, DANIELS TE, YANG T, MITZ AR, RHODES SEV & MURRAY EA. 2014. A role for primate subgenual cingulate cortex in sustaining autonomic arousal. Proceedings of the National Academy of Sciences 111: 5391–5396.

SALIM J & NG J. 2023. Unexpected neonatal death due to medium-chain acyl-CoA deficiency. Pathology 55: S43.

SHICHKOVA P, COGGAN JS, KANARI L, BOCI E, FAVREAU C, ANTONEL SM, KELLER D & MARKRAM H. 2023. Breakdown and repair of the aging brain metabolic system. 2023.08.30.555341.

SJÖSTEDT E et al. 2020. An atlas of the protein-coding genes in the human, pig, and mouse brain. Science 367: eaay5947.

STAWICKA ZM, MASSOUDI R, OIKONOMIDIS L, MCIVER L, MULVIHILL K, QUAH SKL, COCKCROFT GJ, CLARKE HF, HORST NK, WOOD CM & ROBERTS AC. 2022. Differential Effects of the Inactivation of Anterior and Posterior Orbitofrontal Cortex on Affective Responses to Proximal and Distal Threat, and Reward Anticipation in the Common Marmoset. Cerebral Cortex 32: 1319–1336.

SUN T, COLLURA RV, RUVOLO M & WALSH CA. 2006. Genomic and Evolutionary Analyses of Asymmetrically Expressed Genes in Human Fetal Left and Right Cerebral Cortex. Cerebral Cortex 16: i18–i25.

TANG C, DENG X, QU J, MIAO Y, TIAN L, ZHANG M, LI X, SUN B & CHEN L. 2023. Fenofibrate Attenuates Renal Tubular Cell Apoptosis by Up-Regulating MCAD in Diabetic Kidney Disease. Drug Design, Development and Therapy 17: 1503–1514.

VERMETTEN E & LANIUS RA. 2012. Chapter 18 - Biological and clinical framework for posttraumatic stress disorder. In: AMINOFF MJ et al. (Eds), Handbook of Clinical Neurology, Elsevier, p.291–342.

WANG Y, SHEN Q-L, XIN Q, SUN B, ZHANG S, FANG Q-H, SHI Y-X, NIU W-Y, LIN J-N & LI C-J. 2022. MCAD activation by empagliflozin promotes fatty acid oxidation and reduces lipid deposition in NASH. Journal of Molecular Endocrinology 69: 415–430.

WILLIS MW, KETTER TA, KIMBRELL TA, GEORGE MS, HERSCOVITCH P, DANIELSON AL, BENSON BE & POST RM. 2002. Age, sex and laterality effects on cerebral glucose metabolism in healthy adults. Psychiatry Research: Neuroimaging 114: 23–37.

YARIAN CS, REBRIN I & SOHAL RS. 2005. Aconitase and ATP synthase are targets of malondialdehyde modification and undergo an age-related decrease in activity in mouse heart mitochondria. Biochemical and Biophysical Research Communications 330: 151–156.

YIN F, BOVERIS A & CADENAS E. 2014. Mitochondrial Energy Metabolism and Redox Signaling in Brain Aging and Neurodegeneration. Antioxidants & Redox Signaling 20: 353–371.

YUSUF IQ, VENKATESAN A, OKAFOR FC, YASIN A & OYIBO SO. 2023. A Young Female With Medium-Chain Acyl-CoA Dehydrogenase Deficiency (MCADD): A Case Report. Cureus 15: e36018.

Figure legends

Figure 1. Cortical ACADM mRNA expression by age and hemisphere.

(a) ACADM gene expression in the brain cortex, measured as log nTPM, through several age groups, ranging from 61 to 94 years of age. (b) Overall cortical expression of ACADM by brain hemisphere. Boxplots are filled in light blue with a red dot marking the mean log nTPM values.

Figure 2. Normalized ACADM mRNA expression in human cortical areas.

(a) ACADM gene expression in different areas of the brain cortex. The red dotted lines indicate lower and higher bounds around the median nTPM values for overall cortical ACADM expression. (b) Higher and (c) lower-than-median nTPM cortical areas. Boxplots are filled in light blue with a red dot marking the mean nTPM values. Boxplots are filled in light blue with a red dot marking the mean nTPM values.

Figure 3. Differential cortical ACADM expression by sex.

(a) Cortical areas with ACADM expression differences between sexes. Data is presented as mean ± standard error. (b) Overall cortical expression of ACADM by sex. Pink and blue bars indicate female and male samples, respectively.

Figure 4. ACADM expression by tissue and sex.

(a) ACADM expression across different peripheral tissues. (b) ACADM expression in tissues where the data was available for both sexes ($p > 0.05$ for all comparisons). (c) Overall tissue expression of ACADM by sex. Pink and blue bars indicate female and male samples, respectively.

Supplementary figure 1. Workflow overview for extraction and analysis of HPA data.

Steps for the analysis of The Human Protein Atlas (HPA) expression data. First, (1) the XML file was collected following the directions available on the HPA website. Then, (2)

the XML file was uploaded to a Google Colaboratory notebook for processing through a Python pipeline. The output, a (3) structured dataset in .XLSX format is then imported into (4) R (or RStudio) for (5) data analysis.

Supplementary figure 2. Interactive R Shiny application for “no code” processing of HPA files.

Overview of the R Shiny application home page. The application complements the Python data pipeline by providing an interactive and user-friendly interface for exploring, visualizing, and downloading processed expression data. Built using the Shiny dashboard framework, the app offers a structured layout with several key features. Users can upload XML files in the "Upload & Parse" section, triggering the data parsing process. The processed data is then displayed in interactive DataTables within the "humanBrain Data" and "Tissue Data" tabs, facilitating the examination of expression patterns. In the "Download" tab, users can export the consolidated data as an Excel file, organized into distinct worksheets for human brain and tissue data. Additionally, a "Select Theme" feature, utilizing Shiny themes, allows for real-time customization of the app interface, enhancing the overall user experience.

Supplementary figure 3. Comparison of nTPM counts across cortical regions by hemisphere.

Analysis of normalized transcript per million (nTPM) counts across different cortical regions, comparing left (pink bars) and right (blue bars) hemispheres. The data show a small number of interhemispheric differences in nTPM counts for the assessed cortical regions, indicating similar gene expression levels regardless of laterality for most of the brain cortex.

Supplementary figure 4. Sex-based analysis of ACADM expression in the cerebral cortex.

Further analysis of ACADM expression in the cerebral cortex to evaluate the impact of sex on MCAD. Expression levels, measured as nTPM, are shown across different cortical regions. In areas where ACADM expression is at or below median nTPM values, no significant differences were observed between males (blue bars) and females (pink bars). However, cortical regions with higher ACADM expression exhibit greater variability, suggesting potential sex-dependent differences in expression patterns.