Application of rMTA-EA Pipeline to Multi-Gene Metabolic Optimization in Neurodegenerative Diseases

Catarina Gomes[1], Bruno Sá[1], and Miguel Rocha[1,2⋆]

[1] Centre of Biological Engineering, University of Minho, Braga, Portugal
[2] LABBELS – Associate Laboratory, Braga/Guimarães, Portugal

Abstract. The Robust Metabolic Transformation Algorithm (rMTA) enhances constraint-based modeling by identifying metabolic interventions that shift a biological system from a pathological to a healthier state. Integrated with Evolutionary Algorithms (EAs), it systematically explores combinatorial gene knockouts, using Robust Transformation Score (rTS) from rMTA as the optimization criterion. We applied the rMTA-EAs pipeline to neurodegenerative diseases, focusing on Alzheimer's Disease (AD). Using transcriptomic data from the Gene Expression Omnibus (GEO) dataset `GSE203206`, which profiles sporadic AD patients and non-demented controls across early- and late-onset cases, our approach captures diverse AD endotypes and reveals metabolic perturbations beyond age-dependent effects. Based on rMTA insights, we used EAs to explore multi-gene interventions, encoding knockouts as binary vectors and optimizing combinations via rTS. By moving beyond single-gene targets, our rMTA-EAs framework enables a broader search for metabolic interventions that revert disease phenotypes. While demonstrated in AD, this scalable approach is applicable to other complex diseases, offering a powerful tool for systems biology.

Keywords: Alzheimer's disease · Robust Metabolic Transformation Algorithm · Evolutionary algorithms · Constraint-based modeling.

1 Introduction

With the increase of life expectancy worldwide, AD has become an exponentially growing public health concern that affects patients and generates substantial socioeconomic impacts for families and caregivers [1]. As the most common neurodegenerative disease, representing approximately 80% of all cases, the number of patients affected worldwide is estimated to continue to rise [2] [3].

Traditional approaches usually focus on single-gene modifications and fail to capture the entangled network of genetic and metabolic interactions underlying AD [4]. In contrast, systems biology methods, particularly Genome-scale Metabolic Models (GEMs), provide a powerful framework for analyzing these

⋆ Corresponding author.

complex interactions by using stoichiometric relationships to predict metabolic flux distributions, and integrating omics data to refine these predictions [5]. Using these models, Metabolic Transformation Algorithm (MTA) systematically identifies metabolic perturbations that can shift a system from a pathological to a healthy state through integrated constraint-based modeling and differential gene expression [6].

A more robust approach, the rMTA, extends the MTA by incorporating worst-case scenario analyses and the Minimization Of Metabolic Adjustment (MOMA) method, thus improving the reliability of the prediction of metabolic intervention. Despite rMTA's effectiveness in identifying promising single-gene knockouts, the complexity of AD demands a broader approach, as single genetic modifications often trigger compensatory metabolic mechanisms, limiting therapeutic efficacy [7].

To address this limitation, this study integrates EAs with rMTA to systematically explore multigene knockout strategies tailored to AD. By encoding genetic perturbations as binary vectors and employing a fitness function based on the rTS, derived from rMTA, EAs efficiently navigate the vast landscape of combinatorial genetic modifications [8].

This approach aims to identify interactive genetic interventions capable of restoring metabolic homeostasis in AD. Through application of AD-specific transcriptomic data, this study seeks to uncover new metabolic vulnerabilities as potential therapeutic targets, advancing precision medicine approaches for other neurodegenerative and metabolic disorders [9].

2 Objectives

The objectives of this project focus on using the rMTA-EAs framework to identify potential sets of target genes associated with metabolic alterations in AD. More specifically, we aim to select a specific AD transcriptomic dataset as an appropriate case study; integrate the disease-specific RNA-sequencing (RNA-seq) data into a human GEMs; implement the rMTA-EAs framework to evaluate multigenic knockouts; and validate promising knockout strategies within their biological context through a comprehensive literature review.

3 State of the Art

3.1 Alzheimer's Disease

AD is a progressive neurodegenerative disorder characterized by the accumulation of Amyloid-Beta (Aβ) plaques, aggregates of misfolded peptides that disrupt neuronal function and hyperphosphorylated Tau protein forming Neurofibrillary Tangles (NFTs) [10]. These pathological hallmarks compromise neuronal integrity through distinct but interconnected mechanisms: Aβ primarily disrupts synaptic transmission and plasticity, while Tau tangles impair axonal transport and the intracellular movement of essential molecules [2].

As a complex, multifactorial, and multigenic disease affecting various metabolic pathways, Alzheimer's disease manifests in two main forms: familial (early-onset) and sporadic (late-onset) [11]. Early-onset AD is less common and results from autosomal dominant mutations in specific genes—particularly Amyloid Precursor Protein (APP), Presenilin 1 (PSEN1), and Presenilin 2 (PSEN2)—leading to an aggressive disease progression [12]. In contrast, late-onset AD is far more prevalent and arises from complex interactions between genetic and environmental factors [12]. The primary genetic risk factor for late-onset AD is the ε_4 allele of Apolipoprotein E (APOE), which influences multiple aspects of Aβ metabolism, including deposition, aggregation, and clearance [13].

Traditional therapeutic strategies targeting isolated molecular pathways have yielded limited clinical efficacy [14]. The extensive crosstalk between metabolic networks establishes robust compensatory mechanisms that circumvent single-point interventions. These networks exhibit high connectivity and redundancy, where perturbations at individual nodes frequently fail to propagate systemically due to alternative flux distributions [15]. Furthermore, AD demonstrates significant metabolic heterogeneity among patients, undermining the effectiveness of standardized single-target approaches and highlighting the need for personalized therapeutic strategies [16].

3.2 Genome-scale Metabolic Models

GEMs are sophisticated computational frameworks designed to comprehensively map an organism's metabolic capabilities by integrating multi-omics data, including genomic, transcriptomic, proteomic, metabolomic, and metabolic flux information [17] [18]. The integration of these multi-layered data enhances the accuracy and applicability of GEMs, enabling personalized disease modeling and facilitating the identification of more precise therapeutic targets [19]. At their core, GEMs are based on a stoichiometric matrix, which mathematically represents the relationships between metabolites and reactions in a biological system (Equation 1).

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mn} \end{bmatrix} \tag{1}$$

The application of constraints defines a solution space of feasible metabolic flux distributions. Flux Balance Analysis (FBA) is a widely used method that identifies optimal flux distributions based on an assumed cellular objective, such as growth or ATP production [20].

Furthermore, MOMA predicts fluxes following gene knockouts by minimizing deviation from the wild-type state, which often yields more biologically plausible outcomes when full adaptation is not feasible [21].

To improve the biological realism of GEMs, context-specific modeling incorporates transcriptomic data. Integrative Metabolic Analysis tool (iMAT) stratifies gene expression into tiers and formulates a Mixed Integer Linear Program-

ming (MILP) that aligns flux activity with expression levels, without relying on a predefined objective [22]. This strategy provides context-specific GEMs without enforcing a universal biological objective, making it especially suitable for modeling human metabolism.

3.3 Metabolic Transformation Algorithm/ Robust Metabolic Transformation Algorithm

The MTA represents a reliable computational method that integrates GEMs with differential gene expression data to identify targeted metabolic interventions that can change biological systems from undesired (disease) to desired (healthy) state [23] [20]. Unlike traditional approaches, which often prioritize the elimination of cells, MTA offers an innovative strategy to identify drug targets that promote metabolic reprogramming [4]. The rMTA emerges as an improvement on the original MTA by incorporating uncertainty analysis and worst-case scenario modeling, substantially reducing false positives in metabolic intervention predictions [7].

A key aspect of MTA is its mathematical formulation, which relies on Mixed Integer Quadratic Programming (MIQP) to determine the optimal set of metabolic transformations. This approach ensures that the selected interventions adhere to biological constraints while maximizing the therapeutic potential of metabolic reprogramming [21]. Additionally, the refinement introduced in rMTA leverages the MOMA algorithm to enhance the biological feasibility of the predicted metabolic shifts [24].

3.4 Evolutionary Algorithms

EAs are bio-inspired optimization techniques that simulate natural selection processes, including truncation/replacement, to iteratively refine solutions to complex problems [25]. They are particularly effective in exploring large search spaces and identifying optimal or near-optimal solutions where traditional optimization methods struggle [26] [27]. In the context of metabolic engineering, EAs can be strategically integrated with rMTA to handle the combinatorial complexity inherent in optimizing multigenic knockouts. This integration enhances research efficiency by enabling the discovery of optimal multigenic interventions, even in vast metabolic networks. Moreover, EAs are particularly valuable in identifying synergistic combinations of gene knockouts that bypass compensatory mechanisms, which often limit the effectiveness of single-gene interventions [28]. By continuously refining metabolic intervention strategies through simulated evolutionary processes, EAs provide a powerful framework for robust optimization in metabolic reprogramming.

4 Methods

4.1 Integration of Omics Data

For this EAs-rMTA approach, we use the GSE203206 dataset from GEO, which contains transcriptomic profiles of sporadic AD patients and non-demented healthy controls [29]. This dataset captures diverse AD endotypes across early and late-onset cases, revealing that disease mechanisms may not strictly correlate with age of occurrence [3]. Leveraging transcriptomic signatures, our computational framework aims to identify metabolic interventions that target pathological mechanisms rather than solely classifying the disease by onset age.

To explore metabolic dysregulations in AD, we employ the Recon3D model [30], which integrates detailed metabolic pathway reconstructions with three-dimensional structural data. By mapping RNA-seq data onto this network, we infer functional shifts in metabolism and prioritize therapeutic targets [31]. The structural constraints in Recon3D enhance biological relevance, refining our understanding of metabolic perturbations in neurodegenerative disease progression.

4.2 Characterization of the Disease Metabolic State

Calculation of Reference Flux in the Disease State To define the disease-specific metabolic profile, we employ iMAT, a computational approach that derives context-specific metabolic models by integrating omics data [22]. iMAT classifies metabolic reactions as highly expressed, lowly expressed, or ambiguous and optimizes the metabolic model to maximize consistency with these classifications while ensuring feasibility. This generates the context-specific metabolic model, from which flux distributions are sampled to obtain Reference Flux Distribution (v^{ref}), serving as a baseline for identifying metabolic dysregulations.

Calculation of Differential Gene Expression We analyze differential gene expression by comparing transcriptomic data between healthy and diseased states. This process involves statistical methods such as fold-change calculations and significance testing using DESeq2 or edgeR. The resulting differentially expressed genes are mapped onto metabolic reactions within the model, categorizing reactions as increased, decreased, or unchanged in flux.

4.3 Metabolic Transformation Algorithm and Robust Metabolic Transformation Algorithm

MTA is a structured computational pipeline that identifies genetic or environmental modifications capable of reverting a diseased metabolic state to a healthier one. The process begins with the integration of transcriptomic data, where gene expression data is incorporated into GEMs, such as Recon3D, to construct a context-specific metabolic network. Next, a reference flux distribution is established using flux sampling techniques to compute v^{ref}, characterizing the disease state. Differential gene expression analysis is first conducted to categorize

metabolic reactions according to their anticipated flux variations. Next, a systematic perturbation approach using MIQP is employed, where each reaction is computationally altered to evaluate its influence on metabolic state transitions. Finally, a Transformation Score (TS) is computed, prioritizing therapeutic targets based on their ability to effectively restore metabolic equilibrium.

Robust Metabolic Transformation Algorithm (rMTA) rMTA extends MTA by introducing a worst-case scenario analysis and MOMA for improved robustness in target selection. A key aspect of this extension is the Worst-case Transformation Score (wTS), which, unlike MTA, which considers only the Best-case Transformation Score (bTS), evaluates the risks of reinforcing disease states. Additionally, rMTA integrates MOMA to model metabolic adjustments after gene knockouts, enhancing biological accuracy. To further refine the ranking of potential therapeutic targets, the robust Transformation Score (rTS) is introduced, balancing bTS, wTS, and the derived MOMA Transformation Score (mTS). By integrating these refinements, rMTA provides a more reliable framework for identifying metabolic interventions with therapeutic potential.

4.4 Evolutionary Algorithms for Multi-Gene Knockout Optimization

To address the complexity of metabolic systems where single-gene interventions may be insufficient, we integrate Evolutionary Algorithms (EAs) into the rMTA framework. Using the Distributed Evolutionary Algorithms in Python (DEAP) library, we explore the combinatorial space of gene deletions to identify synergistic knockouts that optimize rTS [32].

Evolutionary Algorithms Workflow and Optimization Strategy The EAs workflow iteratively refines knockout candidates through a series of genetic operators designed to optimize metabolic interventions. To maintain biological feasibility, a repair mechanism limits the number of knockouts to experimentally viable levels. Truncation/replacement is a population management strategy where the lowest-performing individuals are systematically removed from the population, allowing the best-performing solutions to survive and propagate their genetic characteristics to the next generation. This approach naturally retains stronger solutions and continuously integrates new genetic material, helping to balance diversity with selective pressure in the population. Additionally, mutation, selection, and crossover strategies further refine the population by selectively discarding suboptimal solutions and integrating new ones, ensuring continuous improvement throughout the optimization process[33].

The combination of rMTA and EAs provides a robust framework for metabolic intervention, facilitating the discovery of novel therapeutic targets beyond single-gene perturbations [33].

References

1. Philip Scheltens, Bart De Strooper, Miia Kivipelto, Henne Holstege, Gael Chételat, Charlotte E. Teunissen, Jeffrey Cummings, and Wiesje M. van der Flier. Alzheimer's disease. *The Lancet*, 397(10284):1577–1590, April 2021. Publisher: Elsevier.
2. Jesús Andrade-Guerrero, Alberto Santiago-Balmaseda, Paola Jeronimo-Aguilar, Isaac Vargas-Rodríguez, Ana Ruth Cadena-Suárez, Carlos Sánchez-Garibay, Glustein Pozo-Molina, Claudia Fabiola Méndez-Catalá, Maria-del-Carmen Cardenas-Aguayo, Sofía Diaz-Cintra, Mar Pacheco-Herrero, José Luna-Muñoz, and Luis O. Soto-Rojas. Alzheimer's Disease: An Updated Overview of Its Genetics. *International Journal of Molecular Sciences*, 24(4):3754, January 2023. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
3. Kayalvizhi Rajendran and Uma Maheswari Krishnan. Biomarkers in Alzheimer's disease. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 562:119857, August 2024.
4. Shiri Stempler, Keren Yizhak, and Eytan Ruppin. Integrating transcriptomics with metabolic modeling predicts biomarkers and drug targets for Alzheimer's disease. *PloS One*, 9(8):e105383, 2014.
5. Current status and applications of genome-scale metabolic models | Genome Biology | Full Text.
6. Noam Auslander, Chelsea E. Cunningham, Behzad M. Toosi, Emily J. McEwen, Keren Yizhak, Frederick S. Vizeacoumar, Sreejit Parameswaran, Nir Gonen, Tanya Freywald, Kalpana K. Bhanumathy, Andrew Freywald, Franco J. Vizeacoumar, and Eytan Ruppin. An integrated computational and experimental study uncovers FUT9 as a metabolic driver of colorectal cancer. *Molecular Systems Biology*, December 2017.
7. Luis V. Valcárcel, Verónica Torrano, Luis Tobalina, Arkaitz Carracedo, and Francisco J. Planes. rMTA: robust metabolic transformation analysis. *Bioinformatics (Oxford, England)*, 35(21):4350–4355, November 2019.
8. Fuxiang Ren, Shiyin Li, Zihao Wen, Yidi Liu, and Deyu Tang. The Spherical Evolutionary Multi-Objective (SEMO) Algorithm for Identifying Disease Multi-Locus SNP Interactions. *Genes*, 15(1):11, December 2023.
9. Bruno Araújo Gomes Sá. Implementing metabolic transformation algorithms and their application in ageing-related research. Master's thesis, Escola de Engenharia, Universidade do Minho, Braga, Portugal, December 2024.
10. Ana R. Monteiro, Daniel J. Barbosa, Fernando Remião, and Renata Silva. Alzheimer's disease: Insights and new prospects in disease pathophysiology, biomarkers and disease-modifying drugs. *Biochemical Pharmacology*, 211:115522, May 2023.
11. Justin M. Long and David M. Holtzman. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell*, 179(2):312–339, October 2019. Publisher: Elsevier.
12. Marcos Vinícius Ferreira Silva, Cristina de Mello Gomide Loures, Luan Carlos Vieira Alves, Leonardo Cruz de Souza, Karina Braga Gomes Borges, and Maria das Graças Carvalho. Alzheimer's disease: risk factors and potentially protective measures. *Journal of Biomedical Science*, 26(1):33, May 2019.
13. Alberto Serrano-Pozo, Sudeshna Das, and Bradley T. Hyman. APOE and Alzheimer's disease: advances in genetics, pathophysiology, and therapeutic approaches. *The Lancet. Neurology*, 20(1):68–80, January 2021.

14. Jibon Kumar Paul, Abbeha Malik, Mahir Azmal, Tooba Gulzar, Muhammad Talal Rahim Afghan, Omar Faruk Talukder, Samar Shahzadi, and Ajit Ghosh. Advancing Alzheimer's Therapy: Computational strategies and treatment innovations. *IBRO Neuroscience Reports*, 18:270–282, June 2025.
15. Zeinab Breijyeh and Rafik Karaman. Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules (Basel, Switzerland)*, 25(24):5789, December 2020.
16. 2023 Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 19(4):1598–1695, April 2023.
17. Matthew A. Oberhardt, Keren Yizhak, and Eytan Ruppin. Metabolically remodeling the drug pipeline. *Current Opinion in Pharmacology*, 13(5):778–785, October 2013.
18. David B. Bernstein, Snorre Sulheim, Eivind Almaas, and Daniel Segrè. Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. *Genome Biology*, 22(1):64, February 2021.
19. Partho Sen and Matej Orešič. Integrating omics data in genome-scale metabolic modeling: A methodological perspective for precision medicine. *Metabolites*, 13(7), 2023.
20. Chao Ye, Xinyu Wei, Tianqiong Shi, Xiaoman Sun, Nan Xu, Cong Gao, and Wei Zou. Genome-scale metabolic network models: from first-generation to next-generation. *Applied Microbiology and Biotechnology*, 106(13-16):4907–4920, August 2022.
21. Daniel Segrè, Dennis Vitkup, and George M. Church. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences*, 99(23):15112–15117, November 2002. Publisher: Proceedings of the National Academy of Sciences.
22. Hadas Zur, Eytan Ruppin, and Tomer Shlomi. iMAT: an integrative metabolic analysis tool. *Bioinformatics*, 26(24):3140–3142, December 2010. Publisher: Oxford Academic.
23. Keren Yizhak, Orshay Gabay, Haim Cohen, and Eytan Ruppin. Model-based identification of drug targets that revert disrupted metabolism and its application to ageing. *Nature Communications*, 4(1):2632, October 2013. Publisher: Nature Publishing Group.
24. Tomer Shlomi, Omer Berkman, and Eytan Ruppin. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proceedings of the National Academy of Sciences*, 102(21):7695–7700, May 2005. Publisher: Proceedings of the National Academy of Sciences.
25. Pradnya Vikhar. *Evolutionary algorithms: A critical review and its future prospects.* December 2016. Pages: 265.
26. Jasper A. Vrugt and Bruce A. Robinson. Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Sciences*, 104(3):708–711, January 2007. Publisher: Proceedings of the National Academy of Sciences.
27. Zahra Beheshti and Siti Mariyam Shamsuddin. A review of population-based meta-heuristic algorithm. *International Journal of Advances in Soft Computing and Its Applications*, 5:1–35, March 2013.
28. Govind Nair, Christian Jungreuthmayer, and Jürgen Zanghellini. Optimal knockout strategies in genome-scale metabolic networks using particle swarm optimization. *BMC bioinformatics*, 18(1):78, February 2017.

29. Andrew B. Caldwell, Balaji G. Anantharaman, Srinivasan Ramachandran, Phuong Nguyen, Qing Liu, Ivy Trinh, Douglas R. Galasko, Paula A. Desplats, Steven L. Wagner, and Shankar Subramaniam. Transcriptomic profiling of sporadic Alzheimer's disease patients. *Molecular Brain*, 15(1):83, October 2022.
30. Elizabeth Brunk, Swagatika Sahoo, Daniel C. Zielinski, Ali Altunkaya, Andreas Dräger, Nathan Mih, Francesco Gatto, Avlant Nilsson, German Andres Preciat Gonzalez, Maike Kathrin Aurich, Andreas Prlić, Anand Sastry, Anna D. Danielsdottir, Almut Heinken, Alberto Noronha, Peter W. Rose, Stephen K. Burley, Ronan M. T. Fleming, Jens Nielsen, Ines Thiele, and Bernhard O. Palsson. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology*, 36(3):272–281, March 2018. Publisher: Nature Publishing Group.
31. Edward J. O'Brien, Jonathan M. Monk, and Bernhard O. Palsson. Using Genome-scale Models to Predict Biological Capabilities. *Cell*, 161(5):971–987, May 2015.
32. Félix-Antoine Fortin, François-Marie Rainville, Marc-André Gardner, Marc Parizeau, and Christian Gagné. DEAP: Evolutionary Algorithms Made Easy. *Journal of Machine Learning Research*, 13:2171–2175, 2012.
33. K. R. Patil, I. Rocha, J. Förster, and J. Nielsen. Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics*, 6:308, 2005.