Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Gene expression in amyotrophic lateral sclerosis: an integrative narrative review of microarray findings, pathogenic mechanisms, and technical challenges

Expressão gênica na esclerose lateral amiotrófica: uma revisão narrativa integrativa de achados de microarray, mecanismos patogênicos e desafios técnicos

F. M. H. Jorge[a]*

[a]Universidade de São Paulo – USP, Faculdade de Medicina, Hospital das Clínicas, Departamento de Neurologia, São Paulo, SP, Brasil

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a devastating and progressive neurodegenerative disorder with a complex and multifactorial pathogenesis characterized by the loss of upper and lower motor neurons. Despite advances in molecular biology, its pathogenesis remains incompletely understood. This narrative review synthesizes findings from key microarray studies, focusing on dysregulated pathways such as ribosomal function, protein synthesis, and metabolic processes within motor nerves and Schwann cells. We further explore the emerging role of competitive endogenous RNA (ceRNA) networks, particularly those involving long non-coding RNAs (lncRNAs), as revealed by advanced bioinformatics in Alzheimer's disease research. Drawing on this conceptual framework, we discuss how similar regulatory mechanisms might operate in ALS, identifying this as a promising direction for future investigation rather than an established finding. A critical analysis of the technical limitations of microarrays—including small sample sizes, the curse of dimensionality, and challenges in validation—is presented. Finally, we discuss future perspectives, advocating for the integration of microarray data with multi-omics approaches, the application of sophisticated machine learning and deep learning techniques, and the use of distributed computing to overcome current analytical bottlenecks. This integrative approach aims to bridge molecular discoveries with methodological advancements, fostering a deeper understanding of ALS pathogenesis and highlighting novel therapeutic targets.

Keywords: amyotrophic lateral sclerosis, gene expression, neurodegeneration.

Resumo

A Esclerose Lateral Amiotrófica (ELA) é um distúrbio neurodegenerativo devastador e progressivo com patogênese complexa e multifatorial, caracterizada pela perda de neurônios motores superiores e inferiores. Apesar dos avanços na biologia molecular, sua patogênese permanece ainda não completamente compreendida. Esta revisão narrativa sintetiza descobertas de estudos-chave de expressão gênica (microarray), com foco em vias desreguladas, como função ribossomal, síntese proteica e processos metabólicos em nervos motores e células de Schwann. Explorado ainda o papel emergente das redes de RNA endógeno competitivo (ceRNA), particularmente aquelas envolvendo RNAs longos não codificantes (lncRNAs), conforme revelado por bioinformática avançada em estudos da doença de Alzheimer. Com base nesse arcabouço conceitual, discutido como mecanismos regulatórios semelhantes podem operar na ELA, identificando esta como uma direção promissora para investigações futuras, e não como um achado estabelecido. Uma análise crítica das limitações técnicas dos microarrays — incluindo tamanhos pequenos de amostra, a maldição da dimensionalidade e desafios na validação — é apresentada. Finalmente, discutimos perspectivas futuras, defendendo a integração de dados de microarray com abordagens multiômicas, a aplicação de técnicas sofisticadas de aprendizado de máquina e o uso de computação distribuída para superar os atuais gargalos analíticos. Esta abordagem integrativa visa unir descobertas moleculares com avanços metodológicos, promovendo uma compreensão mais profunda da patogênese da ELA e destacando possíveis novos alvos terapêuticos.

Palavras-chave: esclerose lateral amiotrófica, expressão gênica, neurodegeneração.

1. Introduction

Amyotrophic Lateral Sclerosis (ALS) is a fatal neurodegenerative disorder characterized by the selective and progressive loss of upper and lower motor neurons, leading to muscle atrophy, paralysis, and typically death within 3–5 years of symptom onset (Brown and Al-Chalabi, 2017). While approximately 10% of cases are familial (fALS),

*e-mail: fredericomhjorge@gmail.com
Received: September 8, 2025 – Accepted: April 8, 2026
Editor: Marcelo A.M. Esquisatto

the vast majority are sporadic (sALS), with an etiology involving a complex interplay of genetic susceptibility, environmental factors, and dysregulated molecular pathways (Renton et al., 2014).

The quest to understand ALS pathogenesis has been profoundly shaped by genomic technologies. For over two decades, DNA microarrays have served as a cornerstone for global gene expression profiling, enabling the simultaneous measurement of thousands of transcripts in a single experiment (Schena et al., 1995). Although newer technologies like RNA-seq offer advantages in detecting novel transcripts and providing a broader dynamic range, microarrays remain a robust, cost-effective, and standardized platform for hypothesis-driven research, especially in studies with larger sample sizes or well-annotated genomes (Bolón-Canedo et al., 2019; Zhao et al., 2014).

This narrative review aims to provide an integrative analysis of the role of microarray technology in elucidating the molecular basis of ALS. Drawing on three foundational sources—a primary research doctoral dissertation detailing microarray and qPCR analyses of motor nerve biopsies from ALS patients (Jorge, 2018); a scoping review on lncRNA-associated ceRNA networks in Alzheimer's disease (Sabaie et al., 2021), which provided methodological and analytical frameworks for understanding RNA regulatory mechanisms; and a technical chapter on microarray analysis challenges and future trends (Bolón-Canedo et al., 2019), which contextualized the analytical limitations and innovations in the field—we will: (i) synthesize key findings from microarray studies in ALS, with a focus on sporadic cases and the involvement of non-neuronal cells; (ii) explore pathogenic mechanisms revealed by these studies, including ribosomal dysfunction, metabolic impairment, and the emerging role of RNA regulatory networks; (iii) critically examine the technical and analytical challenges inherent to microarray research; and (iv) discuss future directions for integrating microarray data with other omics approaches and advanced computational methods.

2. Methods

This narrative review was conducted with the aim of synthesizing and critically analyzing findings from gene expression microarray studies in amyotrophic lateral sclerosis (ALS), with an emphasis on pathogenic pathways, technical challenges, and future methodological perspectives.

2.1. Search strategy for published literature

The identification of peer-reviewed articles was performed through searches in the following electronic databases: PubMed/MEDLINE and Scopus. The search strategy was developed by combining terms related to the disease, the technology, and the tissue type of interest. The following descriptors and Boolean operators were used:

Disease terms: ("amyotrophic lateral sclerosis" OR "neurodegenerative diseases")

Technology terms: ("microarray" OR "gene expression profiling" OR "transcriptome" OR "expression array")

Tissue terms: ("nerve" OR "biopsy" OR "motor nerve" OR "Schwann cell")

The search was limited to articles published between January 2000 and December 2023, covering the period of consolidation of microarray technology as a research tool in neuroscience. Only articles published in English were considered.

2.2. Inclusion of grey literature (theses and dissertations)

Recognizing that relevant and original data may be available in sources not indexed in traditional databases—especially in the case of studies involving rare biological samples, such as motor nerve biopsies in ALS—this review also included grey literature. Specifically, the doctoral dissertation by Jorge (2018), deposited in the institutional repository of the University of São Paulo (Biblioteca Digital de Teses e Dissertações - USP), was incorporated. This thesis was identified through:

Search in the institutional repository using the terms "amyotrophic lateral sclerosis" and "gene expression"

Consultation of the CAPES database (Catalog of Theses and Dissertations)

Recommendation by experts in the field, given the unique nature of the sample set analyzed (motor nerve biopsies from patients with sporadic ALS)

2.3. Selection criteria

The following were considered eligible to comprise the analytical basis of this review:

For peer-reviewed articles:

Original studies that used microarray technology for gene expression analysis

Biological samples derived from human nervous tissue (biopsy or post-mortem tissue)

Inclusion of patients with confirmed diagnosis of ALS

Presence of a control group

For grey literature:

Theses or dissertations containing original gene expression data in ALS

Data not previously published in peer-reviewed article format

Availability of sufficient metadata for extraction and analysis (sample size, validation methodology, clinical data)

2.4. Rationale for selection of the three sources

This review is based on three main sources, selected for their complementarity:

Source Type Rationale for Inclusion

Jorge (2018) Doctoral dissertation Presents primary gene expression data from motor nerve biopsies of patients with sporadic ALS—a rare and difficult-to-obtain type of sample, the data from which were not fully published in articles

Sabaie et al. (2021) Scoping review (article) Provides the conceptual and methodological framework for the analysis of ceRNA networks in neurodegenerative diseases, with a focus on Alzheimer's disease, serving as a reference for transposing these mechanisms to other neurodegenerative diseases, such as ALS

Bolón-Canedo et al. (2019) Technical chapter (article) Offers a systematic analysis of the computational and statistical challenges in microarray studies, as well as

proposing solutions based on machine learning and distributed computing

The choice of these three sources reflects an intentional strategy to integrate: (i) primary empirical data obtained from human tissue of direct pathogenic relevance (Jorge, 2018); (ii) conceptual models and emerging regulatory mechanisms in neurodegeneration (Sabaie et al., 2021); and (iii) a critical analysis of the tools and methodological challenges for gene expression data analysis (Bolón-Canedo et al., 2019).

2.5. Data extraction and analysis

Data were extracted on differentially expressed genes (DEGs) in bulbar and spinal ALS; validation methods used (qPCR, independent validation); methodological challenges reported (sample size, dimensionality, overfitting); enriched pathways and gene ontology categories; co-expression networks and centrality metrics; phenotype-specific gene expression patterns. Findings were synthesized into thematic categories: molecular markers, pathway enrichment, and network topology.

3. Results and Discussion

3.1. Microarray studies in ALS: key findings and pathogenic insights

3.1.1. Gene expression profiling in motor nerves and schwann cells

A pivotal study by Jorge (2018) profiled the extensor hallucis brevis nerve in sALS patients and identified 138 differentially expressed genes (DEGs) compared to controls. This research was particularly insightful as it focused on a tissue directly involved in the disease process—the peripheral motor nerve—and further stratified patients into bulbar and spinal onset subtypes.

Bioinformatic enrichment analyses using DAVID, KEGG, and Gene Ontology revealed two central pathways consistently dysregulated:

Ribosomal Pathway and Protein Synthesis: A significant number of upregulated genes encoded ribosomal proteins (e.g., RPL29, RPL23, RPL37, RPLP0, FAU, UBA52). This suggests a profound disruption in protein synthesis homeostasis, which could represent a compensatory stress response to protein misfolding, an attempt to maintain neuronal integrity, or, paradoxically, a source of toxic protein aggregates (Jorge, 2018; Warner and McIntosh, 2009).

Purine Metabolism: Genes involved in purine metabolism (e.g., AK1, APRT, GUK1) were also altered, highlighting critical deficits in cellular energy metabolism and nucleotide synthesis—processes essential for the high energy demands of motor neurons and Schwann cells (Jorge, 2018).

The study underscored the dual role of Schwann cells. The altered gene expression profile suggests these glial cells are not passive bystanders but active participants in the disease process, potentially exerting both neuroprotective and neurotoxic effects through paracrine signaling. For instance, the upregulation of EPS8, a gene involved in actin remodeling and synaptic plasticity, may indicate an attempt to stabilize damaged neuronal connections (Jorge, 2018; Menna et al., 2013).

3.1.2. Phenotype-specific expression patterns

Comparative analysis between bulbar and spinal ALS revealed distinct molecular signatures. Genes such as FAU, UBA52, UXT, and AK1 showed bulbar-specific upregulation, while ETV3 and LRPPRC were downregulated in the bulbar phenotype. In contrast, genes including HOXA5, HOXA6, HOXA7, RPL12, RPL23, RPL29, and RPL37 were consistently upregulated in ALS patients compared to controls but did not differ significantly between phenotypes, suggesting they represent general ALS markers rather than subtype-specific ones (Jorge, 2018).

3.1.3. The expanding transcriptomic universe: ceRNA networks in neurodegeneration

While ALS microarray studies have traditionally focused on protein-coding genes, the transcriptome is far more complex. Evidence from Alzheimer's disease research, as synthesized by Sabaie et al. (2021), illustrates a powerful regulatory mechanism—the competing endogenous RNA (ceRNA) network—that may have relevance to ALS. It is important to note, however, that direct evidence for such mechanisms in ALS remains limited. The following discussion therefore serves primarily to introduce a conceptual framework that may guide future investigations, rather than to assert established ALS-specific findings.

3.1.4. Established findings from ALS microarray studies

Direct evidence from ALS microarray studies, derived from the work of Jorge (2018), has primarily identified alterations in protein-coding genes involved in:

Ribosomal function and protein synthesis: Upregulation of ribosomal protein genes (RPL29, RPL23, RPL37, RPLP0, FAU, UBA52)

Purine metabolism: Altered expression of AK1, APRT, and GUK1

Cytoskeletal and synaptic plasticity: Upregulation of EPS8

These findings represent established ALS-specific molecular alterations confirmed through microarray and qPCR validation in human motor nerve tissue. To date, no ALS-specific microarray study has systematically profiled lncRNAs and miRNAs alongside mRNAs in the same tissue samples to validate ceRNA interactions.

3.1.5. Conceptual frameworks from Alzheimer's Disease Research

The ceRNA hypothesis, first articulated by Salmena et al. (2011), posits that RNA transcripts—including mRNAs, long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs)—can communicate with and regulate each other by competing for shared microRNAs (miRNAs). For example, an lncRNA acting as a ceRNA can "sponge" a miRNA, preventing it from suppressing its target mRNA.

In Alzheimer's disease, Sabaie et al. (2021) identified validated ceRNA axes involving lncRNAs such as:

BACE1-AS: Sponges miRNAs like miR-485-5p, leading to increased expression of BACE1, a key enzyme in amyloid-beta production (Faghihi et al., 2010).

NEAT1 and SOX21-AS1: Both sequester miR-107, which has known roles in regulating synaptic function and amyloid precursor protein (APP) expression (Sabaie et al., 2021; Wang et al., 2008).

These findings demonstrate that ceRNA networks operate in Alzheimer's disease and represent a well-established regulatory mechanism in neurodegeneration.

3.1.6. Implications and hypotheses for future ALS research

It is plausible that similar regulatory mechanisms may exist in ALS, though this remains to be empirically confirmed. Dysregulated lncRNAs, if identified in future ALS studies, could potentially act as ceRNAs, influencing the stability and translation of mRNAs critical for motor neuron survival, axonal transport, and inflammatory responses. As Sabaie et al. (2021) emphasize, understanding ceRNA networks represents a promising avenue that may, with further validation, provide new molecular targets for therapeutic intervention.

3.1.7. Current limitations and research gaps

It is important to acknowledge that direct evidence for ceRNA networks in ALS remains sparse. To date, no ALS-specific microarray study has systematically profiled lncRNAs and miRNAs alongside mRNAs in the same tissue samples to validate ceRNA interactions. Furthermore, the findings from Alzheimer's disease—while mechanistically instructive—cannot be directly extrapolated to ALS without empirical confirmation. The preceding discussion is therefore intended to highlight a promising area for future investigation rather than to assert established knowledge. Future studies integrating multi-omic profiling (mRNA, lncRNA, miRNA) with experimental validation will be essential to determine whether ceRNA networks contribute to ALS pathogenesis.

3.2. Technical challenges in microarray analysis and future solutions

The application of microarrays, as in the study by Jorge (2018) (n=19 patients, 5 controls), faces several well-documented challenges that can impact the validity and generalizability of findings. Bolón-Canedo et al. (2019) provide a comprehensive analysis of these challenges and propose future directions to address them.

3.2.1. The "Small n, Large p" problem and overfitting

A fundamental issue is the "curse of dimensionality" (Jain and Zongker, 1997). Microarray datasets typically have a very small number of samples (n) relative to a very large number of features (p) (thousands of genes). This creates a high risk of overfitting, where a model learns the noise in the training data rather than the underlying biological signal, failing to perform well on independent data (Bolón-Canedo et al., 2019; Michiels et al., 2005). Dougherty (2001) famously highlighted that error estimation in such a scenario is greatly impacted, leading to unsubstantiated hypotheses.

Solutions: As discussed by Bolón-Canedo et al. (2019), robust validation methods are essential. These include:

Correct Validation Schemes: Using strict separation of training and test sets, and employing cross-validation appropriately.

Statistical Techniques: Utilizing methods like subsampling (Politis and Romano, 1994) to generate reliable confidence intervals without the computational burden of full bootstrapping on large datasets.

Combining Datasets: Meta-analysis of multiple public microarray datasets to increase sample size and statistical power.

3.2.2. Feature selection and interpretability

Identifying the most relevant genes from thousands of candidates is crucial. Traditional methods often relied on filter approaches (e.g., selecting genes based on fold-change or t-test p-values) due to their low computational cost. However, these methods ignore interactions between features.

Solutions: Bolón-Canedo et al. (2019) note that with decreasing costs and potential for larger sample sizes, more sophisticated methods become feasible:

Wrapper and Embedded Methods: These use machine learning models (e.g., support vector machines, random forests) to select gene subsets that optimize predictive performance, capturing gene-gene interactions (Guyon et al., 2006).

High-Performance Computing: Distributing feature selection algorithms using platforms like Apache Spark or leveraging GPU acceleration can make these computationally intensive methods practical for large genomic datasets (Ramírez-Gallego et al., 2017; Ramírez-Gallego et al., 2018).

Covariate Significance Testing: In a non-parametric context, combining significance tests with false discovery rate (FDR) control methods (e.g., Benjamini-Hochberg procedure) can robustly identify important features while controlling for multiple comparisons (Benjamini and Hochberg, 1995).

3.2.3. The rise of deep learning and integrative omics

As discussed by Bolón-Canedo et al. (2019), the field is moving towards deep learning models that can automatically learn hierarchical representations from raw data.

Application to Microarrays: While deep learning requires large datasets, the anticipated increase in sample sizes for microarray studies could make this approach viable. Deep neural networks could identify complex, non-linear patterns in gene expression that are missed by traditional methods.

Multimodal Integration: A major strength of deep learning is its ability to integrate heterogeneous data types. Future research should aim to combine microarray data with other omics data (e.g., miRNA expression, proteomics, methylation arrays) and clinical variables to build a more complete model of ALS pathogenesis (Bolón-Canedo et al., 2019; Fakoor et al., 2013).

4. Conclusion and Future Perspectives

Microarray analysis has been instrumental in identifying key molecular pathways involved in ALS, such

as ribosomal dysfunction, metabolic impairment, and neuroinflammation, often highlighting the active role of glial cells like Schwann cells (Jorge, 2018). These findings represent established contributions derived directly from primary empirical data in human motor nerve tissue.

As Bolón-Canedo et al. (2019) emphasize, the technology remains relevant due to its cost-effectiveness, reproducibility, and well-established analytical pipelines, especially for targeted research questions.

The future of microarray research in ALS lies not in isolation, but in integration and advanced analytics:

Larger, Collaborative Studies: Multi-center collaborations are needed to generate datasets with sufficient statistical power to apply robust machine learning and deep learning techniques (Bolón-Canedo et al., 2019).

Exploring the Non-Coding Genome: As highlighted by Sabaie et al. (2021), ceRNA networks represent a powerful regulatory mechanism in neurodegeneration. While direct evidence in ALS remains limited, future studies must intentionally profile miRNAs and lncRNAs alongside mRNAs to map out potential ceRNA networks that could reveal novel regulatory mechanisms in ALS, similar to those emerging in Alzheimer's research. This represents an important direction for hypothesis-driven investigation rather than an established finding.

Data Integration: Combining microarray data with genomic, proteomic, and clinical data will be essential for developing holistic models of the disease (Bolón-Canedo et al., 2019; Fakoor et al., 2013).

Advanced Bioinformatics: Embracing distributed computing, sophisticated feature selection, and deep learning will be key to extracting the maximum biological insight from existing and future microarray datasets (Bolón-Canedo et al., 2019).

By overcoming its technical challenges through methodological innovation and integrative analysis, microarray technology will continue to be a valuable tool in the ongoing effort to understand and ultimately treat Amyotrophic Lateral Sclerosis. However, it is important to recognize that many of the mechanisms discussed—particularly those involving non-coding RNA regulatory networks—remain hypothetical until validated by dedicated ALS-specific studies.

Data Availability Statement

All data analyzed in this study are available in the University of São Paulo (USP) Digital Library of Theses and Dissertations (https://www.teses.usp.br/?lang=en) and PubMed. They can be accessed via the following DOIs and/or hyperlinks: https://www.teses.usp.br/teses/disponiveis/5/5138/tde-02082018-113658/pt-br.php; doi: 10.3389/fnagi.2021.742242; https://doi.org/10.1007/978-1-4939-9442-7_14.

References

BENJAMINI, Y. and HOCHBERG, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, vol. 57, no. 1, pp. 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

BOLÓN-CANEDO, V., ALONSO-BETANZOS, A., LÓPEZ-DE-ULLIBARRI, I. and CAO, R. Challenges and future trends for microarray analysis. In: V. BOLÓN-CANEDO and A. ALONSO-BETANZOS, eds. *Microarray bioinformatics*. New York: Humana, 2019, pp. 283-293. https://doi.org/10.1007/978-1-4939-9442-7_14.

BROWN, R.H. and AL-CHALABI, A., 2017. Amyotrophic lateral sclerosis. *The New England Journal of Medicine*, vol. 377, no. 2, pp. 162-172. https://doi.org/10.1056/NEJMra1603471. PMid:28700839.

DOUGHERTY, E.R., 2001. Small sample issues for microarray-based classification. *Comparative and Functional Genomics*, vol. 2, no. 1, pp. 28-34. https://doi.org/10.1002/cfg.62. PMid:18628896.

FAGHIHI, M.A., ZHANG, M., HUANG, J., MODARRESI, F., VAN DER BRUG, M.P., NALLS, M.A., COOKSON, M.R., ST GEORGE-HYSLOP, P. and WAHLESTEDT, C., 2010. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biology*, vol. 11, no. 5, pp. R56. https://doi.org/10.1186/gb-2010-11-5-r56. PMid:20507594.

FAKOOR, R., LADHACK, S., NAZI, A. and HUBER, M., 2013. Using deep learning to enhance cancer diagnosis and classification. In: *Proceedings of the 30th International Conference on Machine Learning (ICML 2013)*, 2013, Atlanta, GA, USA. Atlanta: JMLR:W&CP.

GUYON, I., GUNN, S., NIKRAVESH, M. and ZADEH, L.A., 2006. *Feature extraction: foundations and applications*. Berlin: Springer. https://doi.org/10.1007/978-3-540-35488-8.

JAIN, A. and ZONGKER, D., 1997. Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153-158. https://doi.org/10.1109/34.574797.

JORGE, F.M.H., 2018. *Correlação clínico-molecular na esclerose lateral amiotrófica fundamentada pelos achados da expressão gênica no nervo extensor curto do hálux*. São Paulo: Universidade de São Paulo. Tese de Doutorado em Neurologia.

MENNA, E., ZAMBETTI, S., MORINI, R., DONZELLI, A., DISANZA, A., CALVIGIONI, D., BRAIDA, D., NICOLINI, C., ORLANDO, M., FOSSATI, G., REGONDI, M.C., PATTINI, L., FRASSONI, C., FRANCOLINI, M., SCITA, G., SALA, M., FAHNESTOCK, M. and MATTEOLI, M., 2013. Eps8 controls dendritic spine density and synaptic plasticity through its actin-capping activity. *The EMBO Journal*, vol. 32, no. 12, pp. 1730-1744. https://doi.org/10.1038/emboj.2013.107. PMid:23685357.

MICHIELS, S., KOSCIELNY, S. and HILL, C., 2005. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet*, vol. 365, no. 9458, pp. 488-492. https://doi.org/10.1016/S0140-6736(05)17866-0. PMid:15705458.

POLITIS, D.N. and ROMANO, J.P., 1994. Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*, vol. 22, no. 4, pp. 2031-2050. https://doi.org/10.1214/aos/1176325770.

RAMÍREZ-GALLEGO, S., LASTRA, I., MARTÍNEZ-REGO, D., BOLÓN-CANEDO, V., BENÍTEZ, J.M., HERRERA, F. and ALONSO-BETANZOS, A., 2017. Fast-mRMR: fast minimum redundancy maximum relevance algorithm for high-dimensional big data. *International Journal of Intelligent Systems*, vol. 32, no. 2, pp. 134-152. https://doi.org/10.1002/int.21833.

RAMÍREZ-GALLEGO, S., MOURIÑO-TALÍN, H., MARTÍNEZ-REGO, D., BOLÓN-CANEDO, V., BENÍTEZ, J.M., ALONSO-BETANZOS, A. and HERRERA, F., 2018. An information theory-based feature selection framework for big data under Apache Spark. *IEEE Transactions on Systems, Man, and Cybernetics.*

Systems, vol. 48, no. 9, pp. 1441-1453. https://doi.org/10.1109/TSMC.2017.2670926.

RENTON, A.E., CHIÒ, A. and TRAYNOR, B.J., 2014. A genome-wide association study of ALS identifies risk genes and pathways. *Nature Neuroscience*, vol. 17, no. 1, pp. 17-23. https://doi.org/10.1038/nn.3584. PMid:24369373.

SABAIE, H., AMIRINEJAD, N., ASADI, M.R., JALALI, A., DANESHMANDPOUR, Y., REZAEI, O., TAHERI, M. and REZAZADEH, M., 2021. Molecular insight into the therapeutic potential of long non-coding RNA-Associated competing endogenous RNA Axes in Alzheimer's Disease: a systematic scoping review. *Frontiers in Aging Neuroscience*, vol. 13, pp. 742242. https://doi.org/10.3389/fnagi.2021.742242. PMid:34899268.

SALMENA, L., POLISENO, L., TAY, Y., KATS, L. and PANDOLFI, P.P., 2011. A ceRNA hypothesis: the Rosetta Stone of a Hidden RNA language? *Cell*, vol. 146, no. 3, pp. 353-358. https://doi.org/10.1016/j.cell.2011.07.014. PMid:21802130.

SCHENA, M., SHALON, D., DAVIS, R.W. and BROWN, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, vol. 270, no. 5235, pp. 467-470. https://doi.org/10.1126/science.270.5235.467. PMid:7569999.

WANG, W.X., RAJEEV, B.W., STROMBERG, A.J., REN, N., TANG, G., HUANG, Q., RIGOUTSOS, I. and NELSON, P.T., 2008. The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, vol. 28, no. 5, pp. 1213-1223. https://doi.org/10.1523/JNEUROSCI.5065-07.2008. PMid:18234899.

WARNER, J.R. and MCINTOSH, K.B., 2009. How common are extraribosomal functions of ribosomal proteins? *Molecular Cell*, vol. 34, no. 1, pp. 3-11. https://doi.org/10.1016/j.molcel.2009.03.006. PMid:19362532.

ZHAO, S., FUNG-LEUNG, W.P., BITTNER, A., NGO, K. and LIU, X., 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One*, vol. 9, no. 1, pp. e78644. https://doi.org/10.1371/journal.pone.0078644. PMid:24454679.