Brief Report

Molecular biomarker discovery targeting neurodevelopmental disorders and cognitive mechanisms

Konstantinos Panagiotopoulos, Marco Agostino Deriu*

PoliTo[BIO]Med Lab, Department of Aerospace and Mechanical Engineering, Politecnico di Torino, Turin, I-10129 Italy

***Correspondence**: marco.deriu@polito.it
DOI: https://doi.org/10.56280/1703020960

Received: 15 June 2025 Accepted: 29 June 2025 Online Published: 30 June 2025

Abstract

Brain disorders, which encompass neurodevelopmental conditions and age-related neurodegenerative diseases, are becoming more prevalent due to population growth, increased life expectancy, and enhanced diagnostic capabilities. Despite clinical differences, these diseases share complex molecular alterations across various omics layers, including dysregulation of gene expression, epigenetic modifications, gene mutations, and alterations in protein networks. Today, we can directly quantify those alterations; however, determining their relationship to brain function and identifying characteristic disease patterns through the integration of these biomarkers remains a significant challenge. Bioinformatics has become crucial in understanding these signals, enabling us to interpret what they reveal about the mechanisms that regulate our physiology and its pathological changes through the integration of multi-omics data. By analyzing the interplay between genetic, transcriptomic, and epigenetic factors, we can reconstruct disease-specific networks and find potential links between early stages of brain development and degenerative processes. Given the vast amounts of molecular data available today and the significant influx of artificial intelligence techniques, it is increasingly important to develop methods that can explore these datasets deeply and extract the hidden information within them. This review, which focuses specifically on neuroscience-associated molecular biomarkers and the main methods used to acquire and analyze them, highlights how computational approaches are broadening our understanding of brain disorders, opening new avenues for early diagnosis, the development of personalized therapies, and a novel perspective on brain health.

Keywords: Bioinformatics, brain complexity, neurodegeneration, multi-omics, neurodevelopmental disorders

1. Introduction

The human brain is the most sophisticated organ in the body, exhibiting extraordinary complexity in both its structure and function. Neurons, one of the main cell types in the brain, are interconnected in a vast network that enables signal cascading and communication between them, facilitating one of the most remarkable examples of emergent properties in nature. This means that the brain's collective activity reveals effects far greater than the sum of its parts, with one of the most unexpected effects being the cognitive processes and states of awareness. Additionally, the human brain undergoes significant changes from the embryonic to the fetal stage, during which cell proliferation and differentiation, as well as the formation and migration of neurons, occur. Changes continue after birth with post-birth proliferation and myelination throughout childhood until reaching the adult brain (Sultana et al., 2024).

However, this extraordinary complexity, while enabling remarkable capabilities, also leaves the brain susceptible to disruptions in its delicate structure and function. The elongated human life span comes with a high risk for neurodegenerative diseases in a wide spectrum of disorders that profoundly impair cognitive processes, motor control, sensory perception, and states of awareness. Moreover, developmental disorders that can affect 10-20% of children and adolescents, according to a recent study, can interfere with the typical trajectory of growth, notably impacting structural, functional, and cognitive maturation processes (Yang et al., 2022). For neurodevelopmental disorders, the causes are often genetic mutations, prenatal or perinatal insults, and environmental or nutritional factors that disrupt brain maturation. In contrast, neurodegenerative disorders are primarily

Journal of Multiscale Neuroscience

caused by genetic factors, protein misfolding, oxidative stress, and age-related neuronal decline.

Brain-related disorders pose serious challenges for individuals, their families, and society at large. They not only reduce the patient's quality of life but also often lead to their dependency, placing significant personal and financial burdens on family members and professional caregivers. Thus, it is essential to address this escalating global burden of neurological disorders with substantial and constant investment in research that focuses on the early diagnosis or prognosis of those conditions.

Over the last few decades, breakthroughs in biology have enabled high-throughput methods that can quantify thousands of molecular targets in parallel, shifting from hypothesis-driven approaches to data-driven ones. Through these methods, novel markers have been identified and utilized as fingerprints for specific diseases, disease states, or drug targets. These biomarkers are measurable substances or molecules within an organism, and their presence or abundance indicates the presence of specific diseases, infections, or environmental exposures. In biology, it is also known that there is an extreme complexity between molecular components of different types that create multi-directional relationships. Thus, the investigation of any biological system and especially the brain must be done at multiple levels, which can be complementary. This multi-level analysis comes intrinsically from the different types of biomolecules that comprise the samples, each amenable to specific high-throughput 'omics' investigations. For instance, genomics focuses on DNA to understand genetic variation and predisposition; epigenomics links the genome to the environment by studying heritable changes in gene expression not caused by DNA mutations; transcriptomics assesses RNA levels to elucidate gene expression patterns; proteomics investigate proteins to understand cellular function and regulation; and metabolomics, profiles metabolites to capture the dynamic state of cellular metabolism. The integration of these distinct molecular layers provides a holistic view necessary for unraveling complex biological systems.

Due to this complexity, solid biopsies in the brain are invasive procedures that can vary from minimal to extended surgeries with the corresponding risk. Although these procedures are crucial for identifying the type of lesions in the brain, such as tumors, the method most used for neurological impairments is the analysis of liquid biopsies from blood plasma, serum, or cerebrospinal fluid (CSF). These methods are far less invasive and safer, while providing important information for neurological conditions. It has been proven that they can be used as sufficient alternatives in disease prognosis and diagnosis. Nevertheless, the blood-brain barrier physically limits the passage of biomolecules; however, advancements in technology have made it possible to detect low concentrations of biomarkers.

Besides direct biological sampling, imaging methods are used extensively in clinical practice to monitor the brain, particularly for brain damage resulting from neurological conditions. In neurodevelopmental disorders, for example, functional and structural magnetic resonance imaging (fMRI & sMRI) is used in early stages of development to assess the brain volume, the white matter volume, lesions in the brain, and the brain's activity and anatomy (Backhausen et al., 2022; Seyffert & Silva, 2005). While in neurodegenerative diseases, computed tomography (CT) scans, positron emission tomography (PET) and MRIs are used to track the progress of the induced damage (Schwarz, 2021). In both cases, imaging is used as a progression tracking method after alterations have started in the brain, and thus it is an extremely useful tool for monitoring progression or diagnosis, rather than a prognostic tool.

2. **High-throughput methods of profiling**

High-throughput molecular profiling enables the discovery of biomarkers across multiple biological layers, ranging from DNA sequence variation and epigenetic marks, such as DNA methylation, to downstream molecular products, including coding and non-coding RNA transcripts, and ultimately the proteins they encode. Each of these omics layers presents its level of complexity, stemming from intricate intra-layer interactions, crosstalk with other molecular levels, and various biochemical modifications that dynamically regulate biological function.

Moreover, these layers exhibit distinct temporal dynamics and molecular stabilities, and for instance, the genome is relatively static, whereas transcriptomes and proteomes have progressively more transient and context-dependent profiles reflecting cellular state and environmental responses (**Figure 1**). Furthermore, each omics domain provides unique insights and requires specialized analytical strategies and technologies to comprehensively characterize the underlying molecular landscape. Integrating these layers through

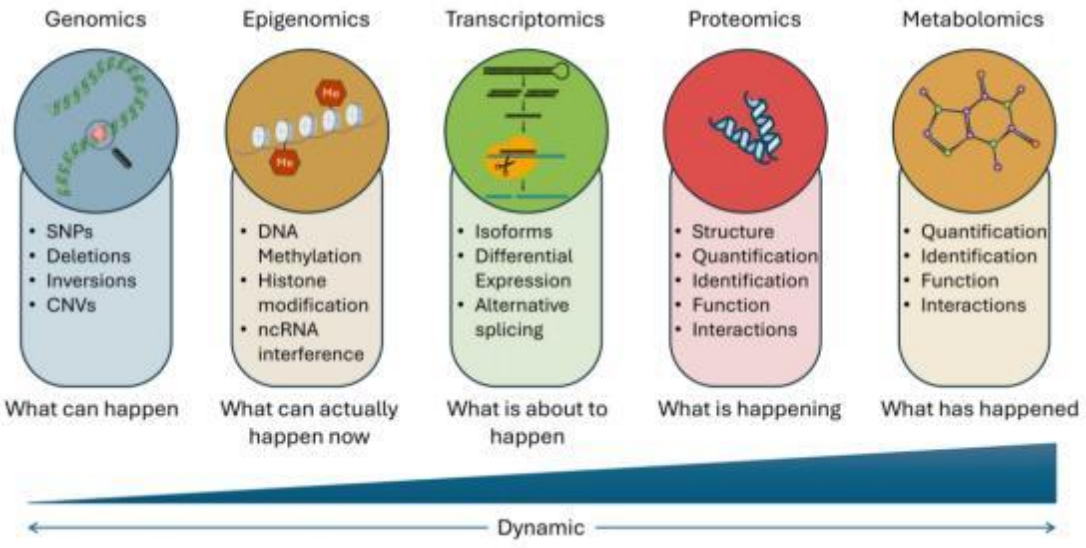

Figure 1: Ma*jor omics approaches investigate distinct classes of biomolecules, providing complementary insights into the biological mechanisms underlying health and disease.*

multi-omics approaches provide a more comprehensive view of biological systems, enhancing our understanding of disease mechanisms and facilitating the development of targeted diagnostic and therapeutic interventions.

2.1 Genomics

Starting with genomics, which encompasses the genes (DNA) of the organism, since it serves as the blueprint for all other products, one can analyze mutations in an individual's genome, which can occur at single base pairs or on more extensive regions, indicating predisposition to develop a specific disorder. Methods that examine mutations include genome-wide association studies (GWAS), where the entire genome is sequenced and point mutations or single-nucleotide polymorphisms (SNPs) are identified. Other methods study copy number variations (CNVs), where regions of the genome are tested to identify duplications, inversions, deletions, and other modifications. Another method of genome studying is whole-exome sequencing (WES), which focuses on the transcribable part of the genome associated with known disease-related variants (Bartha & Győrffy, 2019) (**Table 1**).

Although methods like GWAS do not necessarily identify causative variants, they highlight loci associated with disease risk due to linkage disequilibrium. By analysing the genomic proximity of these variations, it is possible to pinpoint associated regions and potential candidate genes. Examples of biomarker identification with this method include the point mutations in the APP gene that encodes for the Amyloid Beta Precursor Protein and the genes PSEN1 and PSEN2, all known for their role in the manifestation of the different types of Alzheimer's disease (Valdes et al., 2025). Another example is the finding of a strong association between the DRB1*1501 allele and multiple sclerosis, especially for Caucasian, Asian and African American populations, showcasing that those with this allele can have a predisposition to develop the disease (The International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2, 2011). Similar examples exist for neurodevelopmental disorders such as William's syndrome, which is the result of a region deletion in chromosome 7q11.23 (Kozel et al., 2021).

In the case of Huntington's disease (HD), which is a rare inherited genetic neurological condition characterized by neuropsychiatric symptoms, progressive cognitive impairments and movement disorders, the diagnosis is usually confirmed through a genetic test. The test identifies a CAG trinucleotide repeat in the huntingtin gene (HTT), where fewer than 26 repetitions are considered normal, and more than 36 is pathogenic or indicative of an increased risk for asymptomatic carriers to develop HD (Stoker et al., 2022). The complexity of the human brain is also expressed through numerous conditions that mimic HD

Table 1. *Basic methodsfor the different omics layers and their common use*

Omics Layer	Typical dynamics	Most Important Methods	Most frequent use
Genomics	Low	WGS, WES, SNP arrays, CGH arrays	SNPs CNVs Structural Variations Mutations
Epigenomics	Low – moderate	ChIP-seq,ATAC-seq, Methylation arrays	DNA Methylation Profiling Histone Modification Mapping ncRNA Interference
Transcriptomics	Moderate	Microarrays, RNA-seq, scRNA-seq, Long-read RNA-seq	Gene Expression Profiling Alternative Splicing ncRNA Analysis scRNA Analysis
Proteomics	High	LC-MS/MS, Electrophoresis, Western blots, Affinity purification MS	Gene Expression Protein Identification Protein Quantification Post-translational Modification analysis PPI mapping
Metabolomics	High	LC-MS, GC-MS, NMR	Metabolite Identification Metabolite Quantification Metabolic Signatures

in terms of phenotype, and are caused by rare genetic mutations on different genes, like in the case of C9orf72-SMCR8 Complex Subunit (C9orf72) gene that has been associated with a HD-like syndrome, frontotemporal dementia and with amyotrophic lateral sclerosis (ALS) (Mori et al., 2013).

2.2 Transcriptomics

Beyond genetic mutations, transcriptomic profiling is crucial for providing insights into gene expression and identifying differences between neurological conditions and the control population. At the transcriptomics level, in addition to messenger RNA (mRNA), which encodes proteins, non-coding RNA (ncRNA) plays a significant role in regulating gene expression.

In transcriptomics, different high-throughput technologies have been developed since the early '90s, with the emergence of microarrays and later with RNA-sequencing (RNA-Seq) (Lowe et al., 2017). The principle of microarrays is nucleic acid hybridization, where complementary sequences of known probes (short sequences representing specific genes) and the target samples bind together. Upon binding a fluorescent label, which is attached to the target, that emits light, which is read by a scanner to create an expression value for each transcript (Govindarajan et al., 2012) (**Table 1**). In contrast, next-generation sequencing (NGS) methods, such as RNA-seq, read the nucleotides of the target molecule one by one, allowing for the detection of novel transcripts and indications of alternative splicing. There are various methods for RNA-seq, including paired- and single-end reads, different read lengths, and varying read depths. Both microarrays and NGS have variations that work with non-coding transcripts, which is very helpful in studying the co-expression of different types of transcripts, revealing the biological regulatory dynamics between them. Additionally, single-cell RNA-seq technologies allow the study of individual cells, offering a much deeper understanding of type-specific gene expression and spatial insights in the tissue of interest (D'Agostino et al., 2022).

The analysis of high-throughput transcriptomic data, particularly through NGS technologies, enables the identification of molecular signatures associated with neurological disorders. By comparing gene expression profiles across different populations, such as patients and healthy controls, researchers can uncover statistically significant alterations in transcript abundance that may serve as disease-associated biomarkers. In the context of neurological disorders, this approach offers valuable insights into dysregulated

gene networks and pathogenic mechanisms. Furthermore, transcriptomic analyses can reveal previously unannotated or low-abundance transcripts, including non-coding RNAs and novel splice variants, which may play regulatory roles in disease pathophysiology. These novel transcripts hold potential not only as diagnostic or prognostic biomarkers but also as therapeutic targets for precision medicine approaches.

The analysis of transcriptomics data enables the exploration of changes in gene activity between different conditions, such as healthy and diseased states. By identifying groups of genes that show increased or decreased expression, we can highlight biological processes that maybe disrupted in a specific condition. These groups often reflect functional roles, for example, genes involved in inflammation, metabolism, or neural signaling, and help us understand what maybe going wrong at the molecular level.

In addition to studying individual genes, transcriptomics allows us to build co-expression networks. These networks are based on patterns of gene expression that are consistently observed across samples. Genes that show similar expression levels across individuals are likely to be working together or regulated in a similar way. By connecting genes with similar patterns, we can detect clusters, or modules, of co-expressed genes that may be involved in the same biological function.

2.3 Proteomics

Moving beyond transcripts to their protein products, proteins are the most structurally and functionally diverse biomolecules in the body. This complexity arises from their ability to interact with various other molecules, form complexes, exhibit enzymatic activity, and adopt dynamic conformations. Proteomics, the large-scale study of the entire set of proteins expressed in a cell or tissue, primarily focuses on protein identification, quantification, structural analysis, functional characterization, and interaction mapping.

Two main approaches are commonly used in proteomics: experimental and computational. X-ray crystallography is a powerful experimental method for determining the three-dimensional structure of proteins, providing crucial insight into their function, as structure largely dictates biological role. In recent years, computational methods, such as AlphaFold, have made significant progress in accurately predicting protein 3D structures, offering a faster and more accessible complement to experimental techniques.

For protein quantification, methods like liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) are widely used. These approaches not only identify but also quantify proteins with high sensitivity, even at low abundance. In addition, protein quantification assays (e.g., ELISA, Western blotting) remain essential tools in profiling protein expression across various biological samples (**Table 1**).

As a result of those technologies, proteomics analysis has led to the identification of multiple biomarkers also in liquid biopsies, which are minimally invasive, such as CSF and blood samples, for neurological conditions. Some of the most significant findings pertain to disease monitoring and facilitating early diagnosis. Among the most informative CSF biomarkers are total tau (t-tau), phosphorylated tau (p-tau), the amyloid-beta ratio (Aβ42/Aβ40), and neurofilament light chain (NfL). Particularly, in Alzheimer's disease, increased concentrations of p-tau and t-tau are associated with the accumulation of tau proteins and the extent of neuronal injury, while a decrease in CSF Aβ42 levels is indicative of amyloid plaque buildup within the brain (Blennow & Zetterberg, 2018). Another example is NfL, a key indicator of axonal injury, that shows elevated levels across several neurodegenerative disorders, both in blood in lower concentrations and in the CSF (Arrambide et al., 2016; Gaetani et al., 2019; Lu et al., 2015). Measuring NfL in blood provides a valuable, minimally invasive approach for tracking disease advancement and evaluating treatment effectiveness (Khalil et al., 2018).

A commonly used tool derived from proteomics analysis is the creation of protein-protein interaction networks (PPI), which provides insight into the functions of individual proteins, their relationships with other proteins, and their involvement in complex biological processes. These networks can be analyzed using graph theory methods to create clusters or identify important genes within the network, making it a suitable tool for identifying therapeutic targets.

2.4 Epigenomics

Another important factor is the changes in gene expression that occur without DNA alterations, defining how easily accessible genes are at different developmental stages or in response to environmental stimuli. This layer refers to the epigenome, which

encompasses modifications on histones that package the DNA and form chromatin, the methylation of DNA that alters gene expression, and the influence of non-coding RNAs (ncRNAs) that have multiple roles in gene regulation. These changes happen either naturally and are timed based on the developmental stage of the cells or can be the effects of the environment. Tools for analyzing these types of modifications include, among others, the ChIP-seq for finding histone modification and DNA binding proteins, Methyl-seq that determines the methylation patterns in the genome, and ATAC-seq, which profiles chromatin accessibility (**Table 1**).

Part of the epigenome comprises non-coding transcripts that regulate gene expression in various ways, depending on their nature, and they represent a broad category of transcripts that do not translate into proteins. They can be categorized into two broad groups: short or small non-coding transcripts (sncRNAs) and long non-coding transcripts (lncRNAs), with various subcategories and functions. For example, sncRNAs and especially microRNAs (miRNAs), which are the most well-studied subgroup, have mechanisms of action where they can attach to messenger RNA (mRNA), prohibiting its translation and promoting its degradation.

Evidence shows that several miRNAs are dysregulated in ALS in the hippocampus, and other studies suggest that they might play a role in various other neurological and psychiatric conditions by interfering with inflammation signalling pathways (Martinez & Peplow, 2022; Rashidiet al., 2023). Other studies have shown that the presence of miR-101, miR-20a, and miR-17 inhibits the expression of the APP gene, thereby reducing the abundance of the Aβ peptide fragment and demonstrating a protective role against AD. Moreover, the expression of beta-secretase 1 (BACE1) and Aβ aggregation reduced with the presence of miR-149, miR-34a-5p, miR-16, miR-29c and miR-124 (An et al., 2017; Liu et al., 2022).

Long non-coding RNAs (lncRNAs) can have the role of a “miRNA sponge” in the competitive endogenous RNA (ceRNA) hypothesis, where they compete with other transcripts for the shared microRNAs (miRNAs), or they can modulate gene expression by interacting with chromatin-modifying enzymes (Statello et al., 2021). Example of lncRNAs neurodegenerative diseases such as AD, PD, HD and ALS are the nuclear paraspeckle assembly transcript 1 (NEAT1) and the metastasis-associated lung adenocarcinoma transcript 1 (MALAT1). NEAT1 has been shown to be significantly upregulated and is linked to the modulation of significant proteins for these diseases through miRNA sponging. Its role is diverse, as it has been found to promote or delay the progressive loss of neurons, depending on its multiple downstream targets. MALAT1 is important for normal brain development, and its dysregulation is associated with α-synuclein aggregation in AD and PD patients (Jiang & Xu, 2023). Referring to the example of miRNAs with the BACE1 gene, studies have shown that in AD, its antisense gene BACE1-AS has a stabilizing role, and thereby it increases the abundance of the disease-associated BACE1 protein content and the formation of Aβ42 (Riva et al., 2016).

2.5 Metabolomics

The most dynamic layer of the omics field considered in this work is metabolomics, which encompasses the set of small-molecule metabolites in a biological sample and is influenced by both genetic factors and environmental factors. Monitoring changes in brain metabolism enables the prediction of a condition's state. While both proteomics and metabolomics rely on MS for data acquisition due to its high sensitivity and ability to identify and quantify molecules, they often pair MS with different separation techniques beforehand. For instance, both fields frequently utilize LC-MS to separate complex mixtures before MS analysis, with LC being highly versatile for a wide range of compounds. However, metabolomics also commonly employs gas chromatography-mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR) spectroscopy (**Table 1**). GC-MS is used for volatile or semi-volatile metabolites, while NMR provides direct structural identification and quantification of metabolites in a non-destructive manner and without prior separation.

Metabolites are not as clear as genetic mutations in the identification of specific conditions because they are small common molecules that can indicate a disease based on disruptive level patterns rather than single biomarkers. Metabolism plays a highly important role in the brain, as it provides the necessary energy and substrates for biosynthesis, cellular activities, and signaling. In neurodevelopment, dysregulation of energy metabolism or oxidative stress can disrupt neural proliferation and block synaptogenesis. For instance, a deficiency in mitochondrial function can lead to mitochondrial encephalopathies and Leigh syndrome (Daretal., 2023). Mitochondrial dysfunction has been observed in patients with autism spectrum

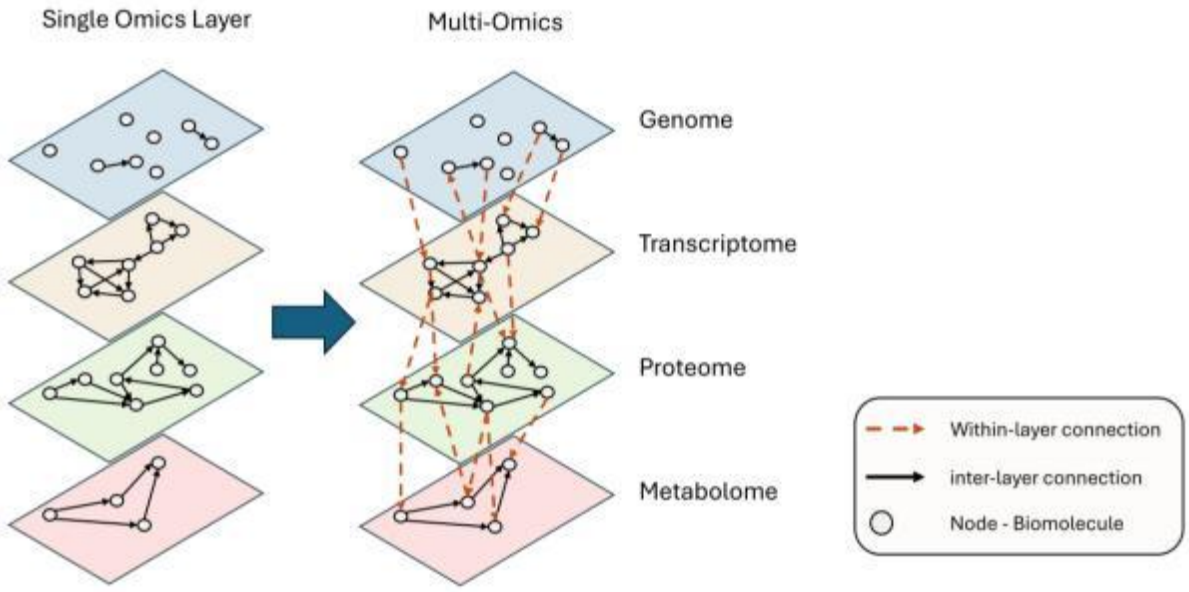

Figure 1. *The single omics layers (left) can be analyzed individually for biomarker discovery and intra-connections between biomolecule. In a multi-layer framework (right) it is possible to uncover underlying connections between biomolecules of different omics levels.*

disorders (ASDs) and other comorbidities such as seizures, fatigue and motor delay (Frye, 2020).

Metabolic abnormalities are also present in neurodegenerative diseases, and particularly, it has been found that glucose hypometabolism and oxidative stress are commonly dysregulated among patients with AD and PD from the analysis of CSF, although these metabolites are non-discriminatory (Quintero Escobar et al., 2021). Additionally, plasma cortisol is found to be associated with cerebral hypometabolism in AD patients, which is mainly observed in posterior cortical regions and might also be related to lower gray matter volume in the hippocampus (Wirth et al., 2019).

3. Network analysis and multi-omics

The complexity of the human brain and the heterogeneity of neurological disorders cannot be fully understood by examining individual biological layers in isolation. As it becomes evident from the different methods used for data acquisition and the nature of this information, the various omics layers are vastly different from one another, making inter-layer comparison impossible and irrelevant. Each omics dimension offers valuable but partial information about cellular states and molecular interactions. To capture the dynamic interplay among these layers, integrative multi-omics computational tools have emerged for the simultaneous analysis of molecular perturbations across multiple regulatory levels. These tools offer insights into disease mechanisms and information cascades that remain hidden when examining a single omics dimension alone, providing a more holistic view of pathophysiological processes.

In this context, network-based approaches offer a powerful conceptual and computational framework for organizing and interpreting complex multi-omics data, considering heterogeneity at various levels. Through graphs, it is possible to represent biological systems as networks, where nodes correspond to molecules—such as genes, proteins, or metabolites—and edges denote their physical, regulatory, or functional interactions. Such systems-level modeling has proven beneficial in the study of systems biology, facilitating the identification of dysregulated pathways, key driver genes, and potential therapeutic targets (Deng et al., 2025; Zenere et al., 2021). Recreated biological networks project the results of individual layer analyses in a unified space and, through computational methods, create an interconnected map of the interactions between biomolecules (**Figure 2**). For instance, co-expression networks derived from transcriptomic profiles can highlight gene modules correlated with disease severity, while integrated protein–protein interaction (PPI) networks enriched with genetic and proteomic data can pinpoint hub proteins implicated in neurodegenerative cascades. Moreover, these analyses can be conducted within the same population across

multiple omics layers or synthesized as meta-analyses combining independent datasets, each approach offering unique advantages in robustness and generalizability.

Modern computational tools, including weighted gene co-expression network analysis (WGCNA), Bayesian inference, and machine learning-based integrative methods, enable researchers to extract meaningful patterns from high-dimensional datasets. This supports more precise patient classification and sheds light on the interplay between genetic factors, molecular disruptions, and clinical features. Through multi-omics integration (transcriptomic, proteomic and epigenomic), Nativio et al. (2020) discovered the importance of epigenetics in early-stage AD and identified transcription- and chromatin–gene feedback loops driving the dysregulation. Examples like this exist for other neurological diseases as well, demonstrating the power of multi-omics analysis over individual layer studies.

As an example, it has been shown that variations in the MECP2 gene are linked to Rett syndrome, a rare postnatal progressive neurodevelopmental disorder affecting mostly females. The mechanism of progression involves the MeCP2 protein, a regulatory protein that binds to methylated DNA and is highly expressed in the brain. There have been found many mutations of the gene that lead to Rett syndrome and Encephalopathy with different genotypes of the MECP2 gene affecting the patient's phenotype (Gold et al., 2024).

4. Discussion and conclusions

In this work, we have comprehensively discussed the application of high-throughput omics profiling methodologies for investigating neurological disorders. We presented key findings and illustrative examples that highlight how different omics layers serve as invaluable biomarkers for assessing, diagnosing, and predicting the prognosis of these complex conditions. Although imaging methods are routinely used in clinical practice for neurological disorders, they can diagnose different conditions based on visual markers at a later stage. Thus, molecular analysis is crucial for prognosis and early diagnosis, even in presymptomatic stages, offering the prospect of early interventions with the proper therapy.

While this work primarily focused on established and widely utilized omics layers, it is crucial to acknowledge the vast and rapidly expanding landscape of additional omics modalities. Emerging fields such as lipidomics, fluxomics, and microbiomics offer unique and complementary insights into various biological processes. Each of these disciplines contributes distinct pieces to the puzzle, and their integration promises a more complete, systems-level understanding of brain function and pathology. For instance, lipidomics can illuminate alterations in membrane composition and signaling, and microbiomics can explore the influence of the gut-brain axis on neurological health.

The challenge of integrating and interpreting the immense volume and heterogeneity of multi-omics data is substantial. While classical bioinformatics approaches have been indispensable for initial data processing, quality control, and the identification of individual biomarkers, the true power of multi-omics lies in its integrative analysis. The integration of the course comes with its own challenges, which lie between heterogeneity and the sample-to-feature ratio, also known as the curse of dimensionality. This is where advanced computational methods, particularly machine learning (ML) and artificial intelligence (AI), become pivotal. For the integration of data, there are three main categories, namely early, intermediate and late, which refer to when the integration takes place and affects the downstream analysis (Zitnik et al., 2019). Early stages integration merges the data before any type of analysis is applied on the distinct layers, intermediate stage uses special algorithms (i.e. multiple kernel learning or deep neural networks) to combine insights from multiple datasets during analysis, and the late stage integration involves building separate models for each dataset and then combining their individual predictions with a second-level model or meta-predictor (Zitnik et al., 2019).

Considering the role of AI and ML in multi-omics analysis, these computational methods have proven beneficial in integrating biological and clinical features, and offer a holistic view of disease pathogenesis and biomarkers discovery for diseases (Catanese et al., 2023; Pammi et al., 2023). Although AI is still in its infancy in multi-omics analysis, it helps identify biomarkers and as a promising means for unraveling the complexity of brain pathologies and improving patients' quality of life.

Neurological disorders – whether it is neurodevelopmental, psychiatric, or neurodegenerative conditions - represent profound dysregulations of typical human brain function. This vulnerability is, to some extent, a consequence of the brain's immense complexity and interconnection. Perturbations across

any of the various omics layers (e.g., genomics, transcriptomics, proteomics, metabolomics, epigenomics) can significantly influence the severity and phenotypic expression of neurological diseases, and these disruptions in such a sophisticated organ can impact amultitude of human bodily functions, ranging from motor control and speech to learning ability and cognition.

Despite the complexity of the brain, advancements in computational biology and sophisticated data and sample acquisition methods offer a foundation for unravelling its intricacies. Recent progress in medicine and technology has empowered clinicians and researchers to discover an expanding array of biomarkers from diverse tissue and fluid samples, allowing for faster, more accurate, and less invasive diagnoses, which in turn improve patient safety and comfort.

While significant improvements in analytical tools for individual omics layers have been achieved, the inherent complexity of the brain and neurological diseases continues to pose a challenge to fully comprehending their underlying mechanisms and progression when relying solely on single-omics analyses. It is an undeniable truth that these individual omics approaches have proven invaluable for diagnostics, prognostics, and drug discovery. However, it is increasingly clear that integrating data across multiple molecular layers is essential for constructing a holistic view necessary for a deeper understanding of these conditions. This need is particularly evident in both neurodegenerative diseases (e.g., Alzheimer's, Parkinson's, and Huntington's disease) and neurodevelopmental disorders (e.g., autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), and intellectual disabilities), where disease severity, clinical presentation, and trajectory are shaped by intricate biological networks spanning multiple regulatory levels.

Therefore, integrative multi-omics approaches, which systematically combine and analyze data across diverse biological layers, are recognized as vital due to the complexity of these conditions. Such approaches are crucial for generating a more comprehensive and holistic understanding of disease pathophysiology, identifying robust and validated biomarkers, and ultimately developing more targeted and effective therapeutic strategies that align with the principles of precision medicine. Continued advancements in data integration, computational modelling, and collaborative research efforts will further unlock the potential of multi-omics to transform our understanding and management of neurological disorders.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Funding Statement

The present research was carried out as part of the PARENT and GALATEA projects. PARENT received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Innovative Training Network, Grant Agreement No. 956394 (https://parenth2020.com/). The GALATEA project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101183057 (https://galateahe.eu/)

References

An, F., Gong, G., Wang, Y., Bian, M., Yu, L. & Wei, C. (2017) MiR-124 acts as a target for Alzheimer's disease by regulating BACE1. *Oncotarget* **8**, 114065–114071.

Arrambide, G., Espejo, C., Eixarch, H., Villar, L.M., Alvarez-Cermeño, J.C. *et al.* (2016) Neurofilament light chain level is a weak risk factor for the development of MS. *Neurology* **87**, 1076–1084.

Backhausen, L.L., Herting, M.M., Tamnes, C.K. & Vetter, N.C. (2022) Best practices in structural neuroimaging of neurodevelopmental disorders. *Neuropsychology Review* **32**, 400–418.

Bartha, Á. & Győrffy, B. (2019) Comprehensive outline of whole exome sequencing data analysis tools available in clinical oncology. *Cancers* **11**, 1725.

Blennow, K. & Zetterberg, H. (2018) Biomarkers for Alzheimer's disease: current status and prospects for the future. *Journal of Internal Medicine* **284**, 643–663.

Catanese,A., Rajkumar, S., Sommer, D., Masrori, P., Hersmus, N., Van Damme, P., Witzel, S., Ludolph, A., Ho, R., Boeckers, T.M. & Mulaw, M. (2023) Multiomics and machine-learning identify novel transcriptional and mutational signatures in amyotrophic lateral sclerosis. *Brain* **146**, 3770–3782.

D'Agostino, N., Li, W. & Wang, D. (2022) High-throughput transcriptomics. *Scientific Reports* **12**, 20313.

Dar,G.M., Ahmad, E., Ali, A., Mahajan, B., Ashraf,G.M. & Saluja, S.S. (2023) Genetic aberration analysis ofmitochondrial respiratory complex I implications in the development of neurological disorders and their clinical significance. *Ageing Research Reviews* **87**, 101906.

Deng, Z., Wu, J., Chen, X., Li, G., Liu, J., Hu, Z., Li, R., Deng, W., 2025. MNMO: discover driver genes from a multi-omics data based-multi-layer network. *Bioinformatics* **41**, btaf134.

Frye, R.E. (2020) Mitochondrial dysfunction in autism spectrum disorder: unique abnormalities and targeted treatments. *Seminars in Pediatric Neurology* **35**, 100829.

Gaetani, L., Blennow, K., Calabresi, P., Di Filippo, M., Parnetti, L. & Zetterberg, H. (2019) Neurofilament light chain as a biomarker in neurological disorders. *Journal of Neurology, Neurosurgery and Psychiatry* **90**, 870–881.

Gold, W.A., Percy, A.K., Neul, J.L., Cobb, S.R., Pozzo-Miller, L., Issar, J.K., Ben-Zeev, B., Vignoli, A. & Kaufmann, W.E. (2024) Rett syndrome. *Nature Reviews Disease Primers* **10**, 84.

Govindarajan, R., Duraiyan, J., Kaliyappan, K. & Palanisamy, M. (2012) Microarray and its applications. *Journal of Pharmacy and Bioallied Sciences* **4**, S310–312.

Jiang, Y. & Xu, N. (2023) The emerging role of autophagy-associated lncRNAs in the pathogenesis of neurodegenerative diseases. *International Journal of Molecular Sciences* **24**, 9686.

Khalil, M., Teunissen, C.E., Otto, M., Piehl, F., Sormani, M.P., Gattringer, T., Barro, C., Kappos, L., Comabella, M., Fazekas, F., Petzold, A., Blennow, K., Zetterberg, H. & Kuhle, J. (2018) Neurofilaments as biomarkers in neurological disorders. *Nature Reviews Neurology* **14,** 577–589.

Kozel, B.A., Barak, B., Kim, C.A., Mervis, C.B., Osborne, L.R., Porter, M. & Pober, B.R. (2021) Williams syndrome. *Nature Reviews Disease Primers* **7**, 42.

Liu, S., Fan, M., Zheng, Q., Hao, S., Yang, L., Xia,Q., Qi, C. & Ge, J. (2022) MicroRNAs in Alzheimer's disease: potential diagnostic markers and therapeutic targets. *Biomedicine & Pharmacotherapy* **148**, 112681.

Lowe,R., Shirley, N., Bleackley, M., Dolan, S. & Shafee, T. (2017) Transcriptomics technologies. *PLoS Computational Biology* **13**, e1005457.

Lu, C.-H., Macdonald-Wallis, C., Gray, E., Pearce,N., Petzold, A., Norgren, N., Giovannoni, G., Fratta, P., Sidle, K., Fish, M., Orrell, R., Howard, R., Talbot, K., Greensmith, L., Kuhle, J., Turner, M.R. & Malaspina, A. (2015) Neurofilament light chain: a prognostic biomarker in amyotrophic lateral sclerosis. *Neurology* **84**, 2247–2257.

Martinez, B. & Peplow, P. (2022) MicroRNA expression in animal models of amyotrophic lateral sclerosis and potential therapeutic approaches. *Neural Regeneration Research* **17**, 728.

Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., Haass, C. & Edbauer, D. (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335–1338.

Nativio, R., Lan, Y., Donahue, G., Sidoli, S., Berson, A., Srinivasan, A.R., Shcherbakova, O., Amlie-Wolf, A., Nie, J., Cui, X., He, C., Wang, L.-S., Garcia, B.A., Trojanowski, J.Q., Bonini, N.M. & Berger, S.L. (2020) An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nature Genetics* **52**, 1024–1035.

Pammi, M., Aghaeepour,N. & Neu, J. (2023) Multiomics, artificial intelligence, and precision medicine in perinatology. *Pediatric Research* **93**, 308–315.

Quintero Escobar, M., Pontes, J.G.D.M. & Tasic, L. (2021) Metabolomics in degenerative brain diseases. *Brain Research* **1773**, 147704.

Rashidi, S.K., Kalirad, A., Rafie, S., Behzad, E. & Dezfouli, M.A. (2023) The role of microRNAs in neurobiology and pathophysiology of the hippocampus. *Frontiers in Molecular. Neuroscience* **16,** 1226413.

Riva, P., Ratti, A. & Venturin, M. (2016) The long non-coding RNAs in neurodegenerative diseases: novel mechanisms of pathogenesis. *Current Alzheimer Research* **13**, 1219–1231.

Schwarz, C.G. (2021) Uses of human MR and PET imaging in research of neurodegenerative brain diseases. *Neurotherapeutics* **18**, 661–672.

Seyffert, M. & Silva R. (2005) FMRI in pediatric neurodevelopmental disorders. *Current Pediatric Reviews* **1**, 17–24.

Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology* **22**, 96–118.

Stoker, T.B., Mason, S.L., Greenland, J.C., Holden, S.T., Santini, H. & Barker, R.A. (2022) Huntington's disease: diagnosis and management. *Practical Neurology* **22**, 32–41.

Sultana, O.F., Bandaru, M., Islam, M.A. & Reddy, P.H. (2024) Unravelling the complexity of human brain: structure, function in healthy and disease states. *Ageing Research Reviews* **100**, 102414.

The International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2, 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219.

Valdes, P., Caldwell, A.B., Liu, Q., Fitzgerald, M.Q., Ramachandran, S., Karch, C.M. *et al*. (2025) Integrativemultiomics reveals common endotypes across PSEN1, PSEN2, and APP mutations in familial Alzheimer's disease. *Alzheimer's Research & Therapy* **17**, 5.

Wirth, M., Lange, C. & Huijbers, W. (2019) Plasma cortisol is associated with cerebral hypometabolism across the Alzheimer's disease spectrum. *Neurobiology of Aging* **84**, 80–89.

Yang, Y., Zhao, S., Zhang, M., Xiang, M., Zhao, J., Chen, S., Wang, H., Han, L. & Ran, J. (2022) Prevalence of neurodevelopmental disorders among US children and adolescents in 2019 and 2020. *Frontiers in Psycholo*gy **13**, 997648.

Zenere, A., Rundquist, O., Gustafsson, M. & Altafini, C. (2021) Using high-throughput multi-omics data to investigate structural balance in elementary gene regulatory network motifs. *Bioinformatics* **38**, 173–178.

Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A. & Hoffman, M.M. (2019) Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Information Fusion* **50**, 71–91.