University of Catania
Department of Biomedical and Biotechnological Sciences

INTERNATIONAL PhD PROGRAM IN NEUROSCIENCE
XXXIII° cycle

INTEGRATIVE MULTI-OMICS PROFILING IN AMYOTROPHIC LATERAL SCLEROSIS: TOWARD PRECISION MEDICINE

Giovanna Maria, Alessandra Morello

PhD Thesis

Coordinator:

Prof Claudio Bucolo

Supervisors:

Prof. Salvatore Salomone

Dr. Sebastiano Cavallaro

Leave this world a little better than you found it.

Robert Baden-Powell

Table of contents

ABSTRACT

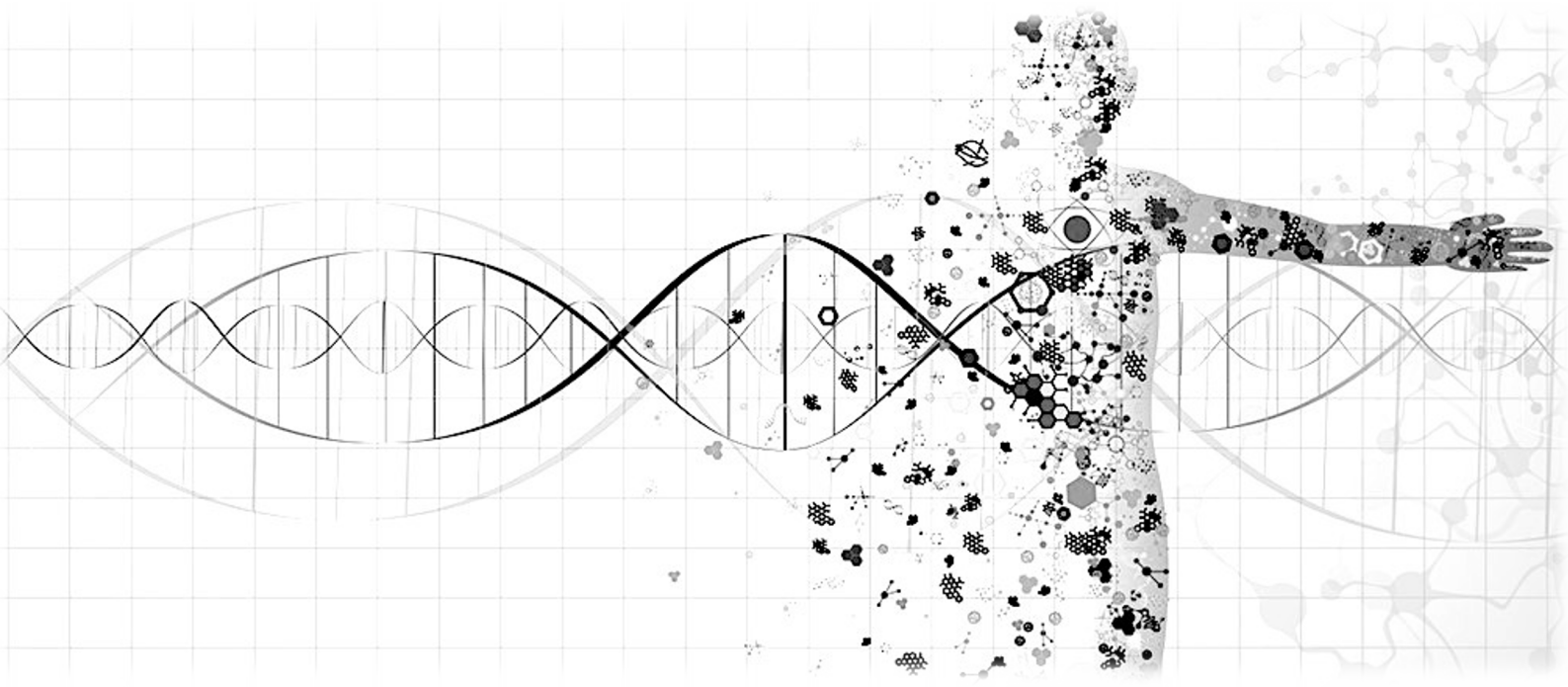
Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disorder, caused by the degeneration of upper and lower motor neurons for which there is no truly effective cure. The lack of successful treatments can be well explained by the complex and heterogeneous nature of ALS, with patients displaying widely distinct clinical features and progression patterns, and distinct molecular mechanisms underlying the phenotypic heterogeneity. Thus, stratifying ALS patients into consistent and clinically-relevant subgroups can be of great value for the development of new precision diagnostics and targeted therapeutics for ALS patients.

In the last years, the use and integration of high-throughput 'omics' approaches (i.e., genomics, transcriptomics) have dramatically changed our thinking about ALS, offering exciting opportunities to decipher the molecular landscape of ALS and underlying heterogeneity of the disease, which will undoubtedly open new doors to therapeutics.

The present PhD thesis aims to apply integrated and comprehensive analyses of multi-omics layers from ALS patients, in order to better characterize the biological heterogeneity underlying ALS and the complex molecular networks that may drive its pathogenesis, providing a rational foundation for exploring new candidate therapeutic targets and biomarkers for the establishment of a more efficacious and personalized genome-guided medicine for ALS.

CHAPTER 1

General Introduction & Aims



AMYOTROPHIC LATERAL SCLEROSIS

Amyotrophic lateral sclerosis (ALS) is a devastating untreatable neurodegenerative disease characterized by the progressive muscular paralysis reflecting degeneration of motor neurons in the brain and spinal cord, leading to paralysis and death, usually from respiratory failure, after 3-5 years [1]. ALS is the commonest of the motor unit diseases in Europe and North America and its incidence is reported to be around 5 cases for 100,000 population per year worldwide [2,3,4]. Approximately 10% of patients newly diagnosed with ALS report a positive family history and are classified as FALS, often with an autosomal dominant pattern of inheritance, and for which several causal genes have been identified, including *SOD1, ALS2, SETX, SPG11, FUS, VAPB, ANG, TARDBP, FIG4, OPTN, ATXN2, UBQLN2, PGRN, PFN1, DCTN1* and *C9ORF72* [5]. The remaining 90%-95% of ALS cases are considered as sporadic (SALS) and appears to be associated to a complex polygenic and multifactorial etiology, involving multiple biological processes, such as oxidative stress, protein aggregation, mitochondrial dysfunction, excitotoxicity and impaired axonal transport [6]. (**Figure 1**). However, this distinction is increasingly recognized to be artificial; FALS and SALS are, in fact, phenotypically indistinguishable and seem to show similar patterns of selective MN degeneration and vulnerability, and many mutations in one or more known FALS-associated genes have been found in SALS patients, suggesting the existence of common molecular mechanisms between these two disease forms [3,6,7].

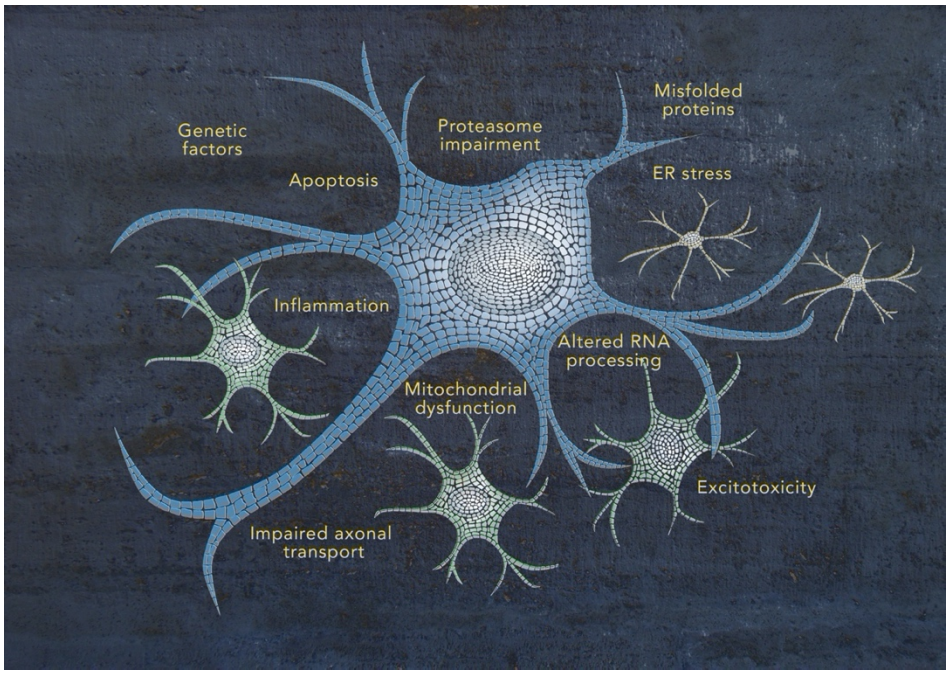


Figure 1. *Schematic representation of the complex mosaic of ALS pathogenesis. From (Morello G. et al., 2017)*

Currently, no cure for ALS exists and only a small number of treatments (e.g., riluzole, edaravone, non-invasive ventilation) appear to mildly slow disease progression and only in some patients [8–10]. One of the reasons for this failure is the use of the lack of adequate knowledge about molecular and cellular players involved in the neurodegenerative progression of ALS and the inappropriate recruitment of patients in clinical studies, without taking into account the medical and molecular heterogeneity of ALS. At present, in fact, the diagnosis is mainly based on clinical, electrophysiological or neuropathologic examination such as evidence for progressive involvement of both upper and lower motor neuron (UMN and LMN) and exclusion of diseases mimicking ALS (i.e., cervical spondylotic myelopathy, multifocal motor neuropathy, Kennedy's disease) as set out in the Revised El Escorial Criteria (REEC), Airlie House criteria (AHC) and Awaji criteria [1,11]. However, it is known that the disease starts with hidden symptoms initiating long before the appearance of any pathological sign and this delay between disease onset and symptoms means missed opportunities for early therapeutic

intervention for ALS patients. Moreover, current diagnostic classification criteria are inadequate to characterize the complex and heterogeneous nature of ALS, as well as the use of a single compound to treat the patient population as a whole may hinder the identification of an effective therapy. Therefore, thus achieving better care for ALS patients.

'OMIC PROFILING FOR MOLECULAR CLASSIFICATION AND BIOMARKER/ TARGET DISCOVERY IN ALS

Advances in "omics" technologies (genomics for DNA, transcriptomics for RNA, epigenomics for histone or DNA modifications, etc.) and their correlation with the clinical phenotypes of the individual patient, are enabling medicine to move from a "one-size-fits-all" approach toward a "personalized" model, helping to clarify the molecular mechanisms underlying pathophysiological aspects of complex human diseases and for uncovering suitable biomarkers and pharmacological targets for a more detailed patient stratification and personalized medicine [10,12–14] (**Figure 2**). To this regard, oncology represents the first and the most well-characterized example of the application of this molecular-based taxonomy paradigm to daily practice, with a multitude of successful individualized or customized therapeutic strategies and immunotherapies available for a wide range of cancers. Based on these promising results, in the last years, this new molecular reclassification has been extended to other polygenic and multifactorial human disorders, including cardiovascular, rheumatic and neuromuscular diseases. Genome-wide association studies, whole-genome/exome sequencing combined with functional genomics, transcriptomics, proteomics, metabolomics, and other omics have already enabled researchers to obtain comprehensive snapshots of biological systems at real time, and with single molecule resolution, thus identifying multiple genomic

alterations related to ALS phenotypes, and providing biological insights to decipher the molecular underpinnings of the disease [15,16].

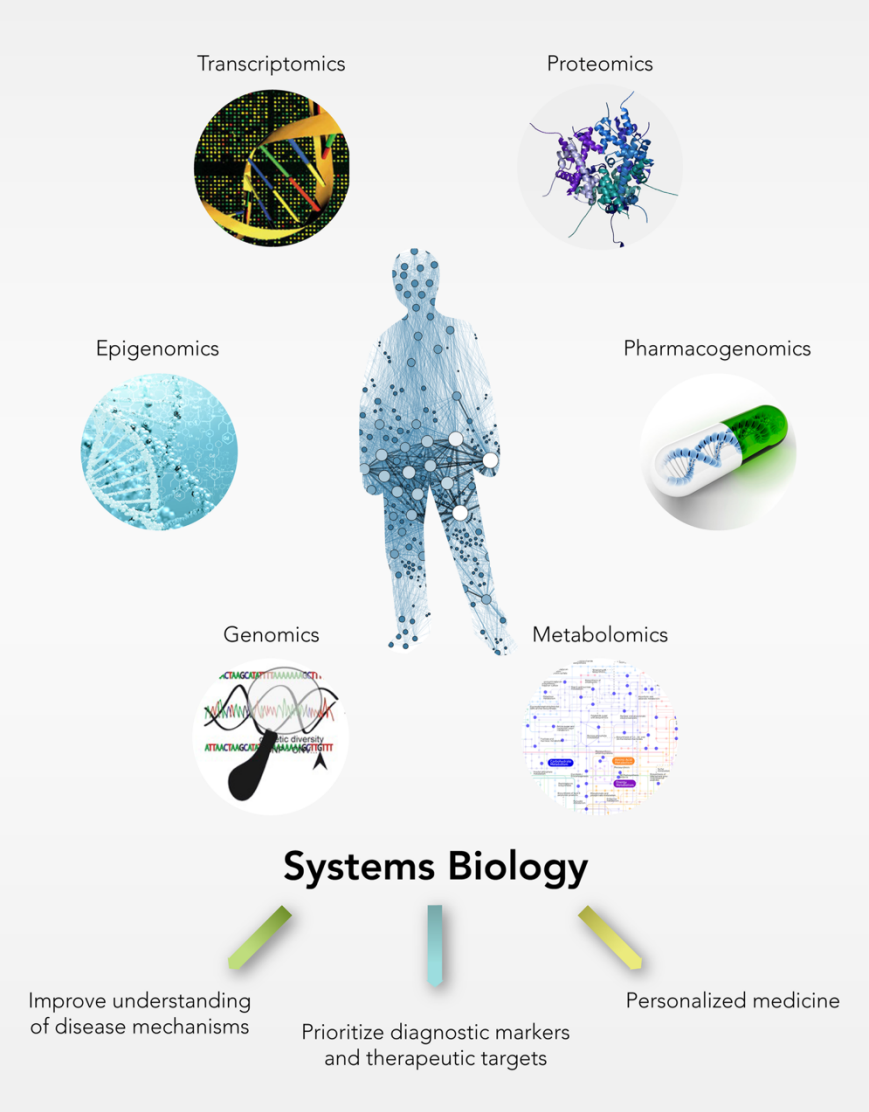
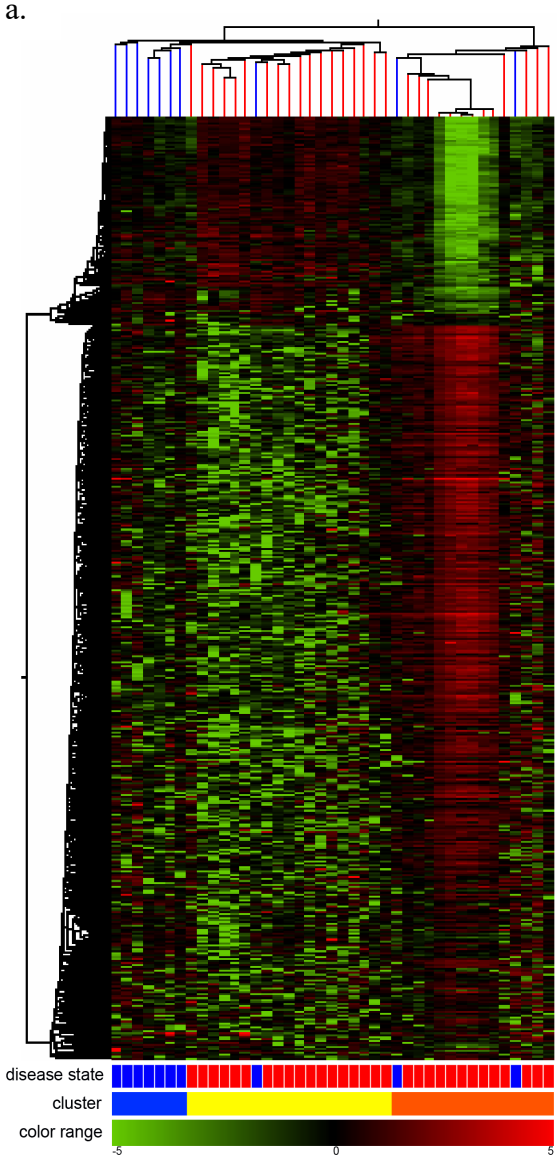


Figure 2. *The systems biology approach: from integration of large-scale "omics" data to personalized medicine practice. From (Morello G. et al., 2017).*

However, to date, lack of progress remains in the understanding of detailed molecular mechanisms underlying ALS and other neurodegenerative disorders, as well as in the patient-specific molecular diagnosis and the consequent translation of this into tailored clinical trials and specific treatments, mostly because of their heterogeneous and multifactorial etiology and the poor availability and complex architecture of nervous tissue.

During the last ten years, our research group has taken important steps towards the development of a new approach to the classification of neurodegenerative diseases, highlighting for the first time the existence of a biological and molecular heterogeneity of ALS [12–14]. In particular, taking advantages of an extensive database of whole genome expression profiles in the motor cortex of sporadic ALS patients, we performed a comprehensive transcriptional analysis of motor cortex samples from control and sporadic ALS (SALS) patients, grouping these on the basis of their similarities measured over the most "hypervariable genes" (9.646 genes with a standard deviation > 1.5). Unsupervised hierarchical clustering analysis allowed to discriminate controls from SALS patients and clearly distinguished two greatly divergent SALS subtypes, each associated with differentially expressed genes and pathways (**Figure 3**). In particular, we identified antigen presentation/processing and extracellular matrix organization as the most representative subgroup-specific pathways in SALS1, while deregulated genes in SALS2 were associated with axonal guidance, oxidative stress and inflammatory intracellular signaling cascades. On one hand, our pathway-based analysis has confirmed the importance of molecular mechanisms that are known to be implicated in ALS pathogenesis. On the other hand, and most importantly, it has been suggested for the first time the differential involvement of these processes in specific subsets of ALS patients.

a.



b.

Cellular processes	Sub-groups	SALS 1	SALS 2
Apoptosis and survival	Intrinsic/extrinsic apoptosis signaling		
	Caspase cascade		
	NF-kB signaling		
Cell adhesion	Remodeling of extracellular matrix		
	Chemokines and adhesion		
	Cell contacts by junctional mechanisms		
	Gap junctions		
	Integrin inside-out signaling		
	Tight junctions		
Cell cycle	Role of 14-3-3 proteins in cell cycle regulation		
	Cell-cycle control system (e.g., cyclins, CDKs)		
	S-phase/ DNA replication		
	ChK signaling/cell cycle arrest		
	Spindle assembly and chromosome separation		
Cytoskeleton remodeling and axonal transport	Cytoskeleton remodeling		
	Keratin filaments		
	Neurofilaments		
	Axonal transport		
	Ephrin B signaling		
	Neuronal growth		
	RhoA regulation pathway		
	CDK5 signaling		
Immune response	Antigen presentation		
	Antigen processing/Proteosome system		
	Cytokines release		
Energy metabolism	Glycolysis		
	Mitochondrial oxidative phosphorylation		
Signal transduction	Activation of PKC via G-Protein coupled receptor		
	Angiotensin signaling		
	Calcium homeostasis		
	cAMP signaling		
	EGF signaling		
	FGF-family signaling		
	GABA-A receptor life cycle		
	IGF-1 receptor signaling		
	Leptin signaling via PI3K-dependent pathway		
	Regulation of eNOS activity		
	Adiponectin signaling		
	Calcium/calmodulin-dependent kinase (CaMK) activity		
	Thyroliberin signaling		
	BDNF signaling		
	Glutamate signaling		
	Acetylcholine-muscarinic receptor signaling		
	Endothelin signaling		
	HGF signaling		
	Shh signaling		
	VEGF-family signaling		

Figure 3. Panel a*. Unsupervised hierarchical clustering of control and SALS patients and correlation between patient stratification and technical variation (array hybridization) or patient demographic (gender, age at onset, age at death, survival time from date of onset, post-mortem intervals (PMI)). Unsupervised hierarchical clustering (similarity measure: Pearson centered; linkage rule: average) was used to cluster control and SALS patients on the basis of their similarities measured over the most informative genes expressed in motor cortex (9646 genes with a standard deviation > 1.5). Similarly, the same genes were clustered on the basis of their similarities measured over the motor cortex of control and SALS patients. In this two-dimensional presentation, each row represents a single gene and each column a motor cortex from control or SALS patients. In the "microarray slide" bar, samples hybridized on the same array are shown with the same color. As shown in the color bar, red indicates up-regulation, green down-regulation, black no change. White squares indicate n/a values. In the dendrograms shown (left and top), the length and the subdivision of the branches display the relatedness of the expression of the genes (left) and the motor cortex (top). Although SALS patients could be clearly distinguished on the basis of their motor cortex gene expression patterns, no significant association was found between their clinical characteristics and cluster assignment.* ***Panel b.*** *Pathways differentially deregulated in cluster SALS patients. Red boxes represent cellular processes mainly up-regulated, green bars down-regulated, red/green bars indicate signal pathways both up- and down-regulated, grey bars indicate no significant change when compared to controls. From (Aronica E. et al., 2015).*

Interestingly, the same accurate transcriptome-based classification was confirmed by utilizing just a restricted subset of genes extensively implicated in monogenic forms of ALS, suggesting that sporadic and monogenic forms of ALS share common etiopathogenic mechanisms and confirming the existence of a molecular heterogeneity in ALS. Moreover, a similar stratification of ALS patients into different molecular subtypes has been confirmed in other recent studies [15,16,17]. In particular, *Tam and co-authors* were able to stratify the transcriptomes of two independent cohorts of ALS postmortem cortex samples into three distinct molecular subtypes, two of which overlapped with the molecular signatures observed in our ALS patient samples [16]. The existence of a molecular sub-classification of SALS patients is not completely surprising if we consider that clinical studies on edaravone, a free radical scavenger recently approved by FDA for ALS treatment showed effectivess only in specific sub-cohorts of patients, supporting the need for clinical trials to take individual variability and genotypic features into account. To this regard, beyond refining ALS molecular architecture, our analyses also revealed a number of potential biomarkers and therapeutic targets that were differentially deregulated in specific subsets of ALS patients [18–21]. Of note, some of these selected target genes exhibit conserved expression patterns in mouse and human ALS, thus providing a rationale to ensure their preclinical trial success. Besides identifying new potential pharmacological targets, our analysis also provided a rational approach for "drug repositioning" for ALS. Under this perspective, many known drugs that were abandoned at clinical stages because of their low efficacy and/or toxicity might be re-evaluated in light of the emerging molecular taxonomy of ALS patients.

Overall, our findings may hold profound implications both for identification of etiopathogenic mechanisms that were not put in evidence by considering ALS pathology

as a single entity, but also for the individualization of diagnostic and therapeutic approaches, bringing us a step closer to the establishment of a more efficacious and personalized genome-guided medicine for ALS. However, single-omics analyses cannot provide a systemic understanding of the precise picture of the composite molecular machinery involved in motor neuron degeneration, making difficult to discriminate genes critical to disease pathogenesis (driver genes) from non-relevant genes (passenger genes). In this context, an integrated and comprehensive view of multiple genomic data types (such as genome and transcriptome) provides powerful opportunities for defining disease subgroups and their molecular drivers, allowing for an overall understanding of the complex molecular networks that drive ALS pathogenesis at yet another level of systemic complexity. In turn, a detailed understanding of how the interaction of multiple pathways drives disease progression may offer a rational foundation for new potential therapeutic targets and biomarkers that will assist in the rapid diagnosis and prognosis of the disease, and for the stratification of patients into different subgroups with the aim of personalized therapeutic strategies.

AIMS OF THE PhD WORK

The aim of the present PhD thesis is to evaluate how the application of an integrative multi-omics approach in ALS may offer the possibility not only to understand how the interaction of multiple altered pathways may determine motor neuron degeneration in ALS (thus providing a powerful means for better defining molecular signatures of the disease), but mostly to identify key molecular drivers that may represent potential biomarkers and therapeutic targets for the development of a patient-oriented diagnosis and therapy. The specific aims discussed in next pages can be summarized as follows:

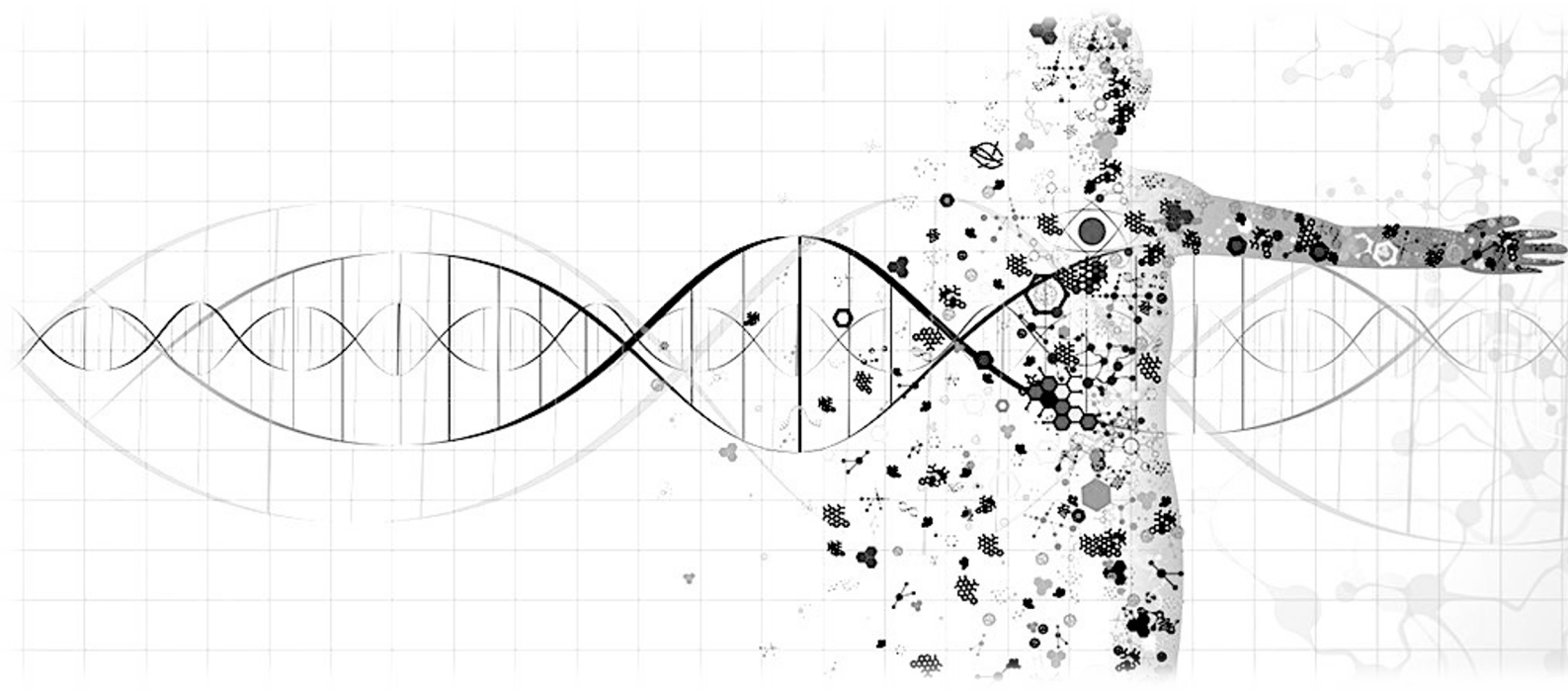
- To analyze genomic structural aberrations (in particular copy number variants-CNVs) occurring in the motor cortex of same set of SALS samples clinically and

transcriptomically characterized in our previous works, to determine whether transcriptional alterations in SALS may be related to genomic DNA alterations, offering the potential for a molecular-based stratification of SALS patients at yet another level of systemic complexity. The integrative analysis of copy number profiles with their associated transcriptomic data revealed distinct genomic signatures and candidate driver genes positively associated with the two previously characterized transcriptome-based sALS subgroups, suggesting a strong interaction between genomic and transcriptomic events in ALS.

- To evaluate the most significant contributions of omics approaches and their integration in unraveling the biological complexity of ALS, highlight how holistic systems biology approaches and multi-omics data integration are ideal to provide a comprehensive characterization of patient-specific molecular signatures that could potentially guide therapeutic decisions.
- To provide a comprehensive multi-omics characterization of some of the most promising molecular targets and signaling mechanisms that have emerged as dysregulated in our works, focusing on those that have already shown results both in *in vitro* and *in vivo* models of ALS.
- To employ a systems biology approach for investigating the complex genetic factors and pathogenic mechanisms that may contribute to motor neuron dysfunctions in a large SCA1 family characterized by the co-existence of a clinically-definite ALS individual bearing an intermediate *ATXN1* expansion and SCA1 patients with a full expansion, some of which manifesting signs of lower motor neuron involvement. Our analysis identified both common and distinctive genomic signatures that may explain the phenotypic heterogeneity observed in this family.

CHAPTER 2

Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS



Morello. G., et al. Scientific Reports (2019) 9:9968. doi:10.1038/s41598-019-46355-w

Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS

Giovanna Morello[a], Maria Guarnaccia[a], Antonio Gianmaria Spampinato[a], Salvatore Salomone[b], Velia D'Agata[c], Francesca Luisa Conforti[f], Eleonora Aronica[d,e] and Sebastiano Cavallaro[a,#]

[a]*Institute of Neurological Sciences, Italian National Research Council, Catania, Italy; Department of Biomedical and Biotechnological Sciences, Section of* [b]*Pharmacology and* [c]*Human Anatomy and Histology, University of Catania, Catania, Italy;* [d]*Department of NeuroPathology, Academic Medical Center, Amsterdam, the Netherlands;* [e]*Swammerdam Institute for Life Sciences, Center for Neuroscience; University of Amsterdam, Amsterdam, The Netherlands;* [f]*Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende (Cosenza), Italy*

#*Corresponding author*: Sebastiano Cavallaro, M.D., Ph.D., Institute of Neurological Sciences (ISN), CNR, Via Paolo Gaifami, 18, 95125 Catania, Italy. Phone: +39.095.7338111; Fax: +39.095.7338110; E-mail: sebastiano.cavallaro@cnr.it

Short title: Molecular Taxonomy of ALS based on multi-omics data.

Keywords: Amyotrophic lateral sclerosis, Genomics, Transcriptome, NeuroArray, Multiomics, Personalized medicine

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is an incurable and fatal neurodegenerative disease. Increasing the chances of success for future clinical strategies requires more in-depth knowledge of the molecular basis underlying disease heterogeneity. We recently laid the foundation for a molecular taxonomy of ALS by whole-genome expression profiling of motor cortex from sporadic ALS (SALS) patients.

Here, we analyzed copy number variants (CNVs) occurring in the same patients, by using a customized exon-centered comparative genomic hybridization array (aCGH) covering a large panel of ALS-related genes. A large number of novel and known disease-associated CNVs were detected in SALS samples, including several subgroup-specific loci, suggestive of a great divergence of two subgroups at the molecular level. Integrative analysis of copy number profiles with their associated transcriptomic data revealed subtype-specific genomic perturbations and candidate driver genes positively correlated with transcriptional signatures, suggesting a strong interaction between genomic and transcriptomic events in ALS pathogenesis. The functional analysis confirmed our previous pathway-based characterization of SALS subtypes and identified 24 potential candidates for genomic-based patient stratification.

To our knowledge, this is the first comprehensive "omics" analysis of molecular events characterizing SALS pathology, providing a road map to facilitate genome-guided personalized diagnosis and treatments for this devastating disease.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder, characterized by progressive loss and degeneration of motor neurons in both the motor cortex, brainstem and spinal cord, and is usually fatal due to respiratory failure within 3–5 years of onset[1]. The disease has an incidence of 2.6 per 100,000 individuals-years and prevalence rates of around 6–7/100,000 in Europe, making it the most common adult-onset motor neuron disease[2]. About 5 -10 % of ALS cases show a family history (FALS), while the remainder of cases are classified as sporadic (SALS), and are probably associated to a polygenic and multifactorial etiology[3-5].

The remarkable advances in genome technologies over the last years have led to a huge progress in deciphering the genes and pathways involved in ALS pathogenesis. From the discovery of the first ALS-associated gene *SOD1*, several candidate-gene or genome-wide association studies (GWAS) have identified multiple single-nucleotide polymorphisms (SNPs) affecting potentially ALS-associated genes, including *C9orf72, TDP43, FUS, MATR3, UBQLN2, VCP* and *OPTN*[6-9]. In this context, a recent large-scale genome-wide association study identified a common missense variant and several rare loss-of-function (LOF) mutations within the microtubule motor protein-encoding gene, KIF5A, as candidate ALS risk factors, further supporting perturbations in cytoskeletal function play an important role in ALS and offering a potential target for drug development[10,11].

In addition to the contribution of SNPs, which account for only a limited number of familial and sporadic ALS cases, evidence suggests that other genomic variants, such as copy-number variations (CNVs), that changes gene dose rather than gene function, may exert a more pronounced effect on the onset and rate of disease progression[5,12,13]. In particular, the involvement of CNVs in ALS susceptibility has been clearly

highlighted in two ALS genome-wide association studies, where multiple rare CNVs were shown to represent a more important risk factor for SALS than common CNVs[14,15].

The complexity of its molecular architecture has completely transformed the way we think about ALS, leading us to reconsider the traditional classification and therapeutic systems. In fact, despite intensive research efforts, the precise causes of ALS remain unknown and there is no cure for this devastating disease. The absence of effective treatments can be due in part by the complex and heterogeneous clinical, biochemical and molecular features of ALS, which is also supported by clinical studies on Edaravone (MCI-186), a free radical scavenger recently approved by FDA for ALS treatment showing effectiveness only in specific sub-cohorts of patients[16]. Developing a robust molecular disease portrait that can explain the heterogeneity of ALS is thus fundamental to improve our understanding of the precise molecular mechanisms underlying disease pathogenesis and develop effective treatments for patients.

Our research group has recently characterized the transcriptional profiles of motor cortex samples from control and SALS patients, grouping these on the basis of their similarities measured over the most "hypervariable genes" (9.646 genes with a standard deviation > 1.5). Unsupervised hierarchical clustering analysis allowed to discriminate controls from SALS patients and clearly distinguished two greatly divergent SALS subtypes, each associated with differentially expressed genes (DEGs) and biological pathways[5,17,18]. In particular, the most representative functional processes deregulated in SALS1 were annotated as involved in the regulation of chemotaxis, immunity, and cell adhesion and communication. Deregulated genes in SALS2, in turn, were selectively associated with cytoskeleton organization, regulation of transport and mitochondrial oxidative phosphorylation[5,17]. While these findings are consistent with previous evidence about the crucial role of these pathogenetic

mechanisms in ALS[17,19-24], they suggest for the first time the differential involvement of these mechanisms in specific subsets of ALS patients, offering a useful starting point for the further development of personalized diagnostics and targeted therapies.

While our work lays the foundation for a molecular taxonomy of ALS, very little information is so far available from the single-omic analysis, which makes difficult to discriminate genes critical to ALS pathogenesis (driver genes) from non-relevant genes (passenger genes). An integrated and comprehensive view of multiple genomic data types (such as genome and transcriptome) may provide a powerful potential for defining disease subgroups and their molecular drivers, allowing for an overall understanding of the complex molecular networks that drive ALS pathogenesis at yet another level of systemic complexity.

In this study, we applied the customized exon-centric comparative genomic hybridization array (aCGH) *NeuroArray* platform, designed to target genes associated with ALS as well as genes associated with other neurological disorders [25], to analyze copy number variants (CNVs) in 40 motor cortex samples of control (10) and SALS (30) patients, clinically and transcriptomically characterized in our previous work[17,25]. Next, we provided the first comprehensive integrative analysis of genomic aberrations with expression data derived from the same patients to identify specific chromosomal regions and genes with concordant alterations in DNA and RNA profiles that may represent promising key molecular candidates for SALS. Finally, functional pathway and network analyses were carried out to gain further insights into the molecular complexity of ALS and reveal novel and yet unrecognized biomarkers and therapeutic targets, potentially useful for the development of personalized medicine in ALS.

RESULTS

Transcriptomically distinct SALS patient subgroups show specific copy number alterations

The customized exon-centric *NeuroArray* aCGH platform was used to identify DNA copy number alterations in 30 SALS patients and 10 controls. A total of 1472 CNVs were detected in SALS, including 780 losses and 692 gains (**Figure 1a**). The chromosomal distribution of all CNVs across the 30 SALS genomes tested is plotted in **Figure 1b**. To reduce individual heterogeneities and identify ALS-related significant CNVs, we focused on those that occurred in at least 10% SALS samples. Accordingly, a total of 488 significant CNVs ranging in size from 7 bp to 5.9 Mb were identified in SALS patients, including 271 losses and 217 gains (**Figure 1c, Supplementary Table 1**). Recurrent CNVs were dispersed in the chromosome 1 to 22, with the most frequent amplifications (76.7%) found in chromosome 14, followed by amplifications in chromosome 17 (70%) (**Supplementary Table 1**). The most common linkage rate in SALS mapped to chromosome 20 with a frequency of 80%, followed by chromosome 1 with a frequency of 76.6% (**Supplementary Table 1**).

To identify subgroup-specific genomic signatures, we analyzed CNV events taking into account the previously characterized transcriptome-based stratification of SALS patients in the two subgroups, SALS1 and SALS2. Overall, 813 aberrant regions were associated with SALS1 and 659 with SALS2 patients (**Figure 1a**). Among these, 335 CNVs (218 losses and 117 gains) were detected as frequently altered in SALS1 patients, while 308 (135 losses and 173 gains) were frequently associated with the SALS2 subgroup (**Figure 1c, Table 1, and Supplementary Tables 2-3**).

Interestingly, a large number of these recurrent amplifications and deletions were detected exclusively in SALS patients (absent in the control samples) (**Table 2**).

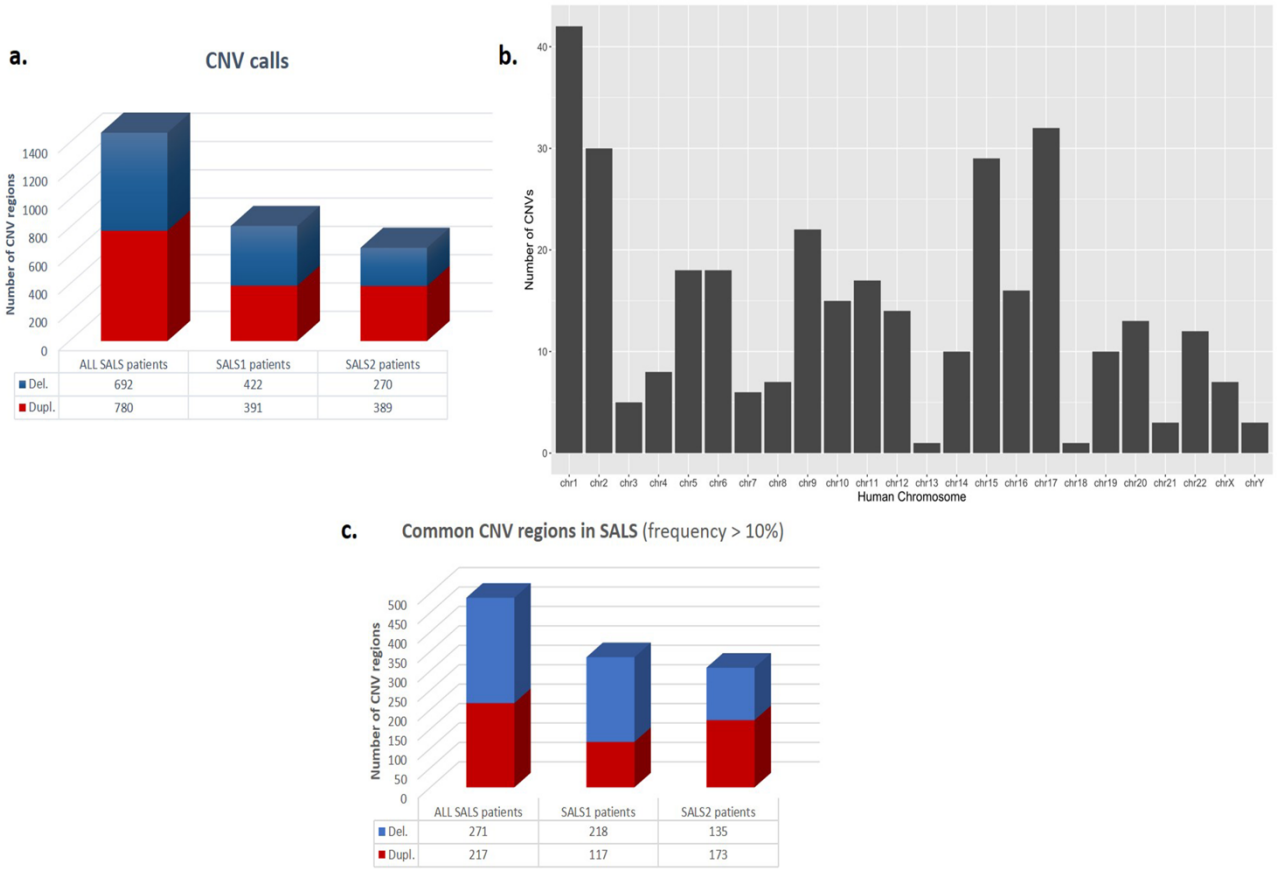


Figure 1. *Significant copy number regions in SALS patient subgroups.* ***(a)*** *Graphical overview of CNV regions detected in SALS patients by NeuroArray platform. The bars represents the number of regions that may be involved in CNV detected in SALS patients (ADM-2 aberration filter: threshold=6; Log2 ratio≥0.5 and 3 consecutive interval probes), both with and without the assignment into the hierarchically determined two sets (SALS1 and SALS2). The blue bar represents the number of deleted regions and the red bar represents the number of amplified regions.* ***(b)*** *Chromosome distribution of CNVs detected with high-resolution custom exon-centered NeuroArray aCGH from SALS patients. The horizontal axis represents different chromosomes and the vertical axis represents the number of regions of each chromosome that may be involved in CNV.* ***(c)*** *Graphical overview of recurrent gains or losses (occurred in at least 10% of the SALS patients), both with and without the assignment into the hierarchically determined two sets (SALS1 and SALS2). The blue bar represents the number of deleted regions and the red bar represents the number of amplified regions that occurred at a high frequency (≥ 10 %) in our cohort of SALS patients.*

Table 1. The top most frequent copy number gain and loss in both SALS patient subgroups.

	SALS1 patients						SALS2 patients					
	Chr	Start	Stop	Size (bps)	Freq. (%)	NeuroArray genes	Chr	Start	Stop	Size (bps)	Freq. (%)	NeuroArray genes
DUPLICATIONS	14	31552632	31552690	59	76.47	AP4S1	14	31552632	31552690	59	76.92	AP4S1
	17	17716576	17720711	4136	70.59	SREBF1	17	17716576	17720711	4136	69.23	SREBF1
	X	122318451	122336599	18149	70.59	GRIA3	22	24376158	24384300	8143	69.23	GSTT1
	X	122318291	122318451	161	64.71	GRIA3	17	17720711	17726812	6102	61.54	SREBF1
	7	100493729	100493862	134	58.82	ACHE	9	129272014	129458220	186207	53.85	LMX1B
	X	122318031	122318291	261	58.82	GRIA3	9	131388073	131394672	6600	53.85	-
	X	122336599	122459975	123377	58.82	GRIA3	17	17715816	17716576	761	53.85	SREBF1
	17	17720711	17726812	6102	52.94	SREBF1	1	55331123	55527185	196063	46.15	DHCR24, PCSK9
	7	100488043	100490289	2247	47.06	ACHE	2	127807997	128439169	631173	46.15	BIN1, LIMS2
	7	100493373	100493729	357	47.06	ACHE	2	241657356	241728764	71409	46.15	KIF1A
	9	131388073	131394672	6600	47.06	-	5	176853852	176869527	15676	46.15	GRK6
DELETIONS	1	47716828	47775972	59145	88.24	STIL	3	155547586	155560259	12674	69.23	SLC33A1
	20	33986975	35569474	1582500	88.24	UQCC, NFS1, PHF20, EPB41L1, DLGAP4, NDRG3, TLDC2, SAMHD1	20	33986975	35575306	1588332	69.23	UQCC, NFS1, PHF20, EPB41L1, DLGAP4, NDRG3, TLDC2, SAMHD1
	20	35569474	35575306	5833	82.35	SAMHD1	1	47767175	47770585	3411	61.54	STIL
	1	47435653	47716828	281176	76.47	STIL	18	9117815	9134199	16385	61.54	NDUFV2
	2	32339724	32429834	90111	76.47	**SPAST**, SLC30A6	3	155560259	155560361	103	53.85	SLC33A1
	3	155551254	155560202	8949	76.47	SLC33A1	12	111956174	111993684	37511	53.85	**ATXN2**
	9	128001127	128001733	607	76.47	HSPA5	1	47435653	47767175	331523	46.15	STIL
	1	47775972	47776133	162	70.59	STIL	1	47770585	47770755	171	46.15	STIL
	2	32323849	32339724	15876	70.59	**SPAST**, SLC30A6	2	32314495	32409410	94916	46.15	**SPAST**, SLC30A6
	2	32429834	32432100	2267	70.59	**SPAST**, SLC30A6	10	70432579	70441196	8618	46.15	TET1
	3	155547586	155551254	3669	70.59	SLC33A1	14	92527804	92562372	34569	46.15	ATXN3

The table lists the gains and losses that occurred in at least 10% of the two previously characterized transcriptome-based SALS subgroups. The chromosomal regions, including the start and end positions, aberration size, frequency in SALS patients and CNV embedded *NeuroArray* genes are listed. Chromosomal positions are referred to the human reference sequence hg19 assembly. Genes previously identified as potential risk factors in ALS are in bold.

Table 2. Chromosomal distribution of the most frequent CNVs exclusively detected in SALS patients

Duplications				
Start	**Stop**	**Lenght (bps)**	**SALS patients**	***NeuroArray*-related genes**
19229182	19826022	596841	6	ALDH4A1, UBR4
22216964	22222489	5526	5	CLCNKA, CLCNKB, **FBXO42,** RCC2, IGSF21, HTR6, VWA5B1, CDA, PINK1, DDOST, EIF4G3, ECE1, USP48, LDLRAD2, HSPG2, APL
22222489	22222904	416	4	CLCNKA, CLCNKB, **FBXO42,** RCC2, IGSF21, HTR6, VWA5B1, CDA, PINK1, DDOST, EIF4G3, ECE1, USP48, LDLRAD2, HSPG2, APL, CDC42
22222904	22379326	156423	3	CLCNKA, CLCNKB, **FBXO42,** RCC2, IGSF21, HTR6, VWA5B1, CDA, PINK1, DDOST, EIF4G3, ECE1, USP48, LDLRAD2, HSPG2, APL
110280992	110467801	186810	3	GSTM3, CSF1
165377539	165378926	1388	5	LMX1A, RXRG
165378926	165406335	27410	3	LMX1A, RXRG
127805579	127806098	520	3	BIN1, LIMS2
152954807	152955141	335	3	CACNB4
37824072	37834861	10790	4	GDNF
135810367	136390729	580363	3	TSC1, RXRA, EDF1, TRAF2, ABCA2, MAN1B1, GRIN1
117263667	117265845	2179	3	CEP164
8790922	8791519	598	3	PIK3R5
8791519	8794231	2713	4	PIK3R5
34198604	34415754	217151	3 (only SALS1)	CCL5, CCL3
56349278	56350271	994	3	MPO
50364607	50364748	142	5	PNKP
18907140	18918487	11348	3 (only SALS1)	PRODH
62886007	63005325	119319	3	ARHGEF9
Deletions				
Start	**Stop**	**Size**	**Ratio SALS/Control**	***NeuroArray*-related genes**
173797400	173826691	29292	5	DARS2
173827623	174553313	725691	3	ZBTB37, RABGAP1L
207791343	207791566	224	3	CR1
207793150	207795258	2109	6	CR1
207795258	207815117	19860	7	CR1
155560361	155572231	11871	4	SLC33A1
155572231	156645173	1072943	3	SLC33A1, KCNAB1, TIPARP, LEKR1
156645399	156660486	15088	3	LEKR1
1810306	3434075	1623770	4	FGFR3, POLN, ZFYVE28, FAM193A, NOP14, RGS12
31625489	31777500	152012	3	HSPA1L
31777500	31783348	5849	4	HSPA1L, HSPA1A
31783348	31797880	14533	6	HSPA1A, HSPA1B
74354308	74530628	176321	3 (only SALS2)	SLC17A5
94767910	94777676	9767	4	TMEM67
94805417	94830376	24960	4	TMEM67
39140211	41979303	2839093	3	CNTNAP3, KGFLP2
125947382	126690362	742981	5	STRBP, DENND2A
101934013	101938026	4014	3	ERLIN1
108124607	108155047	30441	3	ATM
63579805	63673002	93198	3	APH1B, CA12
63673002	64226370	553369	4	HERC1, DAPK2
70551635	70553577	1943	3	COG4
44496313	44496400	88	3	CBS
42373034	42373060	27	4	SEPT3
40460110	41599792	1139683	3	ATP6AP2, MED14, DDX3X, CASK
119673121	119676960	3840	3	CUL4B

The table lists the duplications and deletions that occurred in at least 10% of SALS patients and absent in the control samples. The chromosomal regions, including the start and end positions, number of SALS patients and CNV embedded *NeuroArray* genes are listed. Chromosomal positions are referred to the human reference sequence hg19 assembly. Genes previously identified as potential risk factors in ALS are in bold.

To investigate the reliability of our results and further confirm the potential functional implications of the detected CNVs in ALS pathogenesis, we interrogated our data for overlap with genomic aberrations previously associated to ALS cases available in publicly available databases (i.e., CNVD) and published PubMed literature. Highly similar genomic altered patterns were observed, supporting the functional importance of these regions in disease etiopathogenesis (**Table 3**).

Table 3. Characteristics of the most frequent CNV regions detected in SALS patients and previously associated to ALS from different database and/or published literature.

Chr	Start	Stop	CNV type	Overlapped Genes*	SALS (n=30)	Control (n=10)	Previously reported aberrations	Reference/Database
5	70320678	71554990	Loss	**SMN1, SMN2**, **NAIP,** BDP1, MRPS27	7 (6 SALS1 and 1 SALS2)	2	Loss n=27/167 ALS n=6/167 controls	Corcia P. et al., *Annals of neurology* 2002
8	144635580	145024441	Gain	**GSDMD, EEF1D,** PLEC1	5 (2 SALS1 and 3 SALS2)	1	Gain n=12/781 ALS n=0//621 controls	Wain L.V. et al., *PlosOne*, 2009; Cronin S. et al., *Hum Mol Gen*, 2008
11	424565	792284	Gain	ANO9, DRD4, **DEAF1,** SLC25A22, **OR4A5, OR4C12**	3 (SALS2)	1	Gain n=11/575 ALS n=0/621 controls	Wain L.V. et al., *PlosOne*, 2009
11	4056596	7324468	Loss	STIM1, HBB, TRIM5, CCKBR, APBB1, RRP8, TPP1, **SYT9**	4 (1 SALS1 and 3 SALS2)	1	Loss n=5/575 ALS n=0/621 controls	Wain L.V. et al., *PlosOne*, 2009
15	20575646	23060821	Loss and Gain	**NIPA1**, NIPA2	10 (6 SALS1 and 4 SALS2)	3	Loss and Gain n=15/4434 ALS n=8/14618 controls	CNVD; Blauw HM et al., *Hum Mol Gen*, 2010
15	27017550	27018935	Loss and Gain	GABRB3	3 (1 SALS1 and 2 SALS2)	1	Loss and Gain n=15/4434 ALS n=8/14618 controls	CNVD; Blauw HM et al., *Hum Mol Gen*, 2010
15	41535920	42703427	Loss and Gain	CHP1, NUSAP1, NDUFAF1, MGA, PLA2G4E, CAPN3	5 (2 SALS1 and 3 SALS2)	3	Loss and Gain n=15/4434 ALS n=8/14618 controls	CNVD; Blauw HM et al., *Hum Mol Gen*, 2010
15	63579805	64226370	Loss	APH1B, CA12, HERC1, DAPK2, **DPP8, PTPLAD1, C15orf44, SLC24A1**	4 (3 SALS1 and 1 SALS2)	0	Loss n=1/1875 ALS n=0/8731 controls	CNVD; Blauw HM et al., *Hum Mol Gen*, 2010
17	2580007	4605227	Loss	**OR1D5, OR1D4, OR1D2, OR1G1, OR1A2, OR1A1, SPNS2, ALOX15, PELP1, ARRB2, MED11, CXCL16, ZMYND15,** PAFAH1B1, RAP1GAP2, ITGAE	9 (5 SALS1 and 4 SALS2)	4	Loss and Gain n=1/1875 ALS n=0/8731 controls	CNVD; Blauw HM et al., *Hum Mol Gen*, 2010
17	75277604	78092622	Gain	SEPT9, TNRC6C, DNAH17, CYTH1, **TIMP2,** C1QTNF1-AS1, C1QTNF1, RBFOX3, TBC1D16, **CHMP6, AATK, BAIAP2**	7 (4 SALS1 and 4 SALS2)	2	Gain n=1/12 ALS n=0/24 controls	CNVD, Pamphlett R. et al., *Journal of Neuroscience*, 2011
X	153127628	153602907	Gain	L1CAM, MECP2, FLNA	9 (4 SALS1 and 5 SALS2)	2	Loss	Schoichet S.A. et al., *Amy Lat Scl*, 2009

The table shows the most frequent CNV loci and relative *NeuroArray* genes that partially or completely overlap with genomic aberrations previously associated to ALS cases and reported in publicly available databases (i.e., CNVD) and/or published PubMed literature. The chromosomal regions, including the start and end positions, aberration type and CNV embedded *NeuroArray* genes are listed. In addition, the number of controls and ALS cases from both our experiment and previous scientific reports of these CNV loci was also shown. Chromosomal positions are referred to the human reference sequence hg19 assembly. Genes that may be reasonable ALS candidates are in bold. Chr: Chromosome

Identifying CNV signature genes in SALS

To identify ALS driver genes from aberrant regions, the recurrent CNVs in SALS were annotated and filtered out for genes previously linked to ALS and other neurological diseases as causative and/or susceptibility factors and included in the NeuroArray design. A total of 406 significant CNV genes were obtained, including 251 duplications and 161 deletions (**Figure 2a, Supplementary Table 4**). Among these, 36 were previously identified as ALS-linked genes (**Supplementary Table 4**). The same analysis was also performed on the most frequent subgroup-specific CNV regions, revealing 310 genes as the most significantly altered CNV-genes in SALS1 (137 duplicated and 174 deleted) and 454 genes in SALS2 patients (320 duplicated and 140 deleted). Among these, 28/310 SALS1 and 39/454 SALS2 CNV genes were already associated to ALS (**Figure 2a, Supplementary Tables 5-6**).

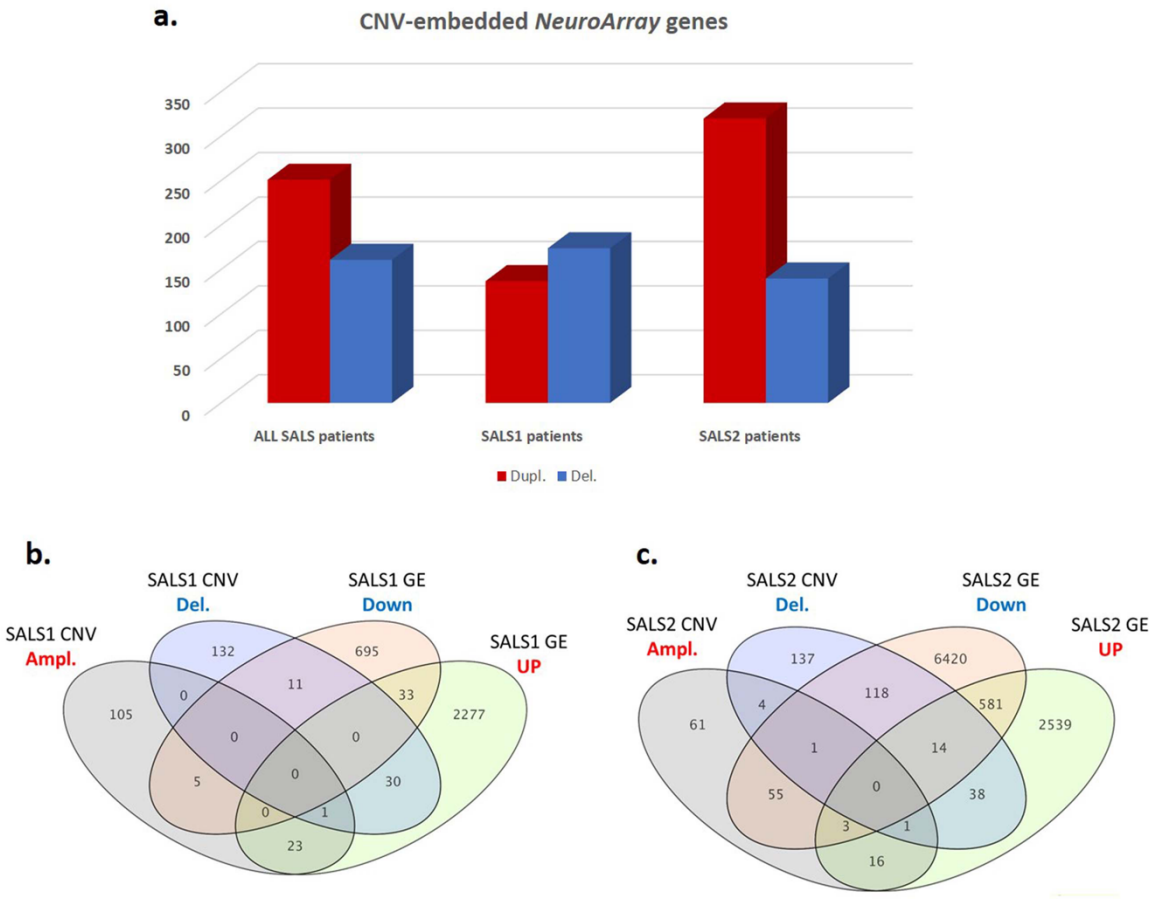


Figure 2. *Integrative analysis of DNA copy number and expression variation in SALS patients reveals a good number of overlapping ALS candidate genes.* ***(a)*** *Graphical overview of the most frequent (≥ 10%) CNV-embedded genes detected in SALS patients, both with and without the assignment into the hierarchically determined two sets (SALS1 and SALS2). (b-c) Venn diagrams compare the number of protein-coding genes obtained from CNV analyses with the genes found to be differentially expressed in SALS1* ***(b)*** *and SALS2* ***(c)*** *patients.*

Integrated analysis of CNVs and gene expression profiling identify candidate ALS-driver genes

To determine whether genomic aberrations contribute to global gene expression patterns in SALS, the identified CNV genes were checked for overlap with the DEGs previously detected in the same patient cohort[17]. We identified 70 overlapping CNV genes (29 duplications and 42 deletions) that were also differentially expressed in SALS1 patients and 246 CNV-driven DEGs in SALS2 patients (173 duplicated and 76 deleted) (**Figure 2b-c, Supplementary Tables 7-8**). Among these, 35 CNV-driven

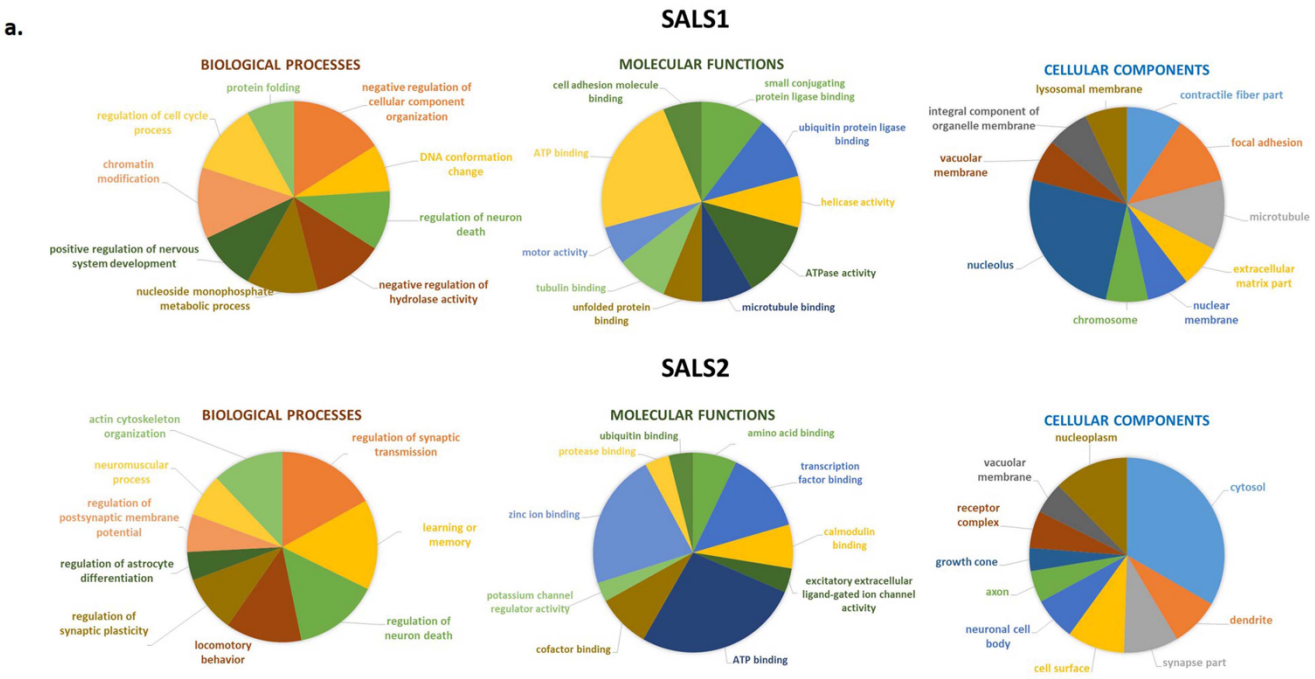
genes (50%) in SALS1 and 112 CNV-driven genes (45%) in SALS2 showed a positive association between gene expression and DNA copy number changes, including 77 up-regulated genes (24 in SALS1 and 53 in SALS2) and 70 down-regulated genes (11 in SALS1 and 59 in SALS2) (**Figure 2b-c, Supplementary Tables 7-8**). Interestingly, several CNV-driven genes were SALS-patient specific (not detected in 10 controls) and most of them were previously linked to ALS.

To demonstrate that the correlation found in our work is meaningful, we also performed a "control experiment" in which we evaluated the overlap between the genomic and transcriptomic data between two random groups within the total SALS samples (different from SALS1 and SALS2). We observed a very low overlap between CNV genes and differentially expressed genes in these randomized disease-related subgroups, confirming the appropriateness/accuracy of our analysis (**Supplementary Figure 2 and Supplementary Table 14**).

Computational systems biology analysis identified distinct drivers and pathways in SALS molecular subtypes

To gain further insights into the biological role of identified CNV-driven DEGs in SALS, functional annotation and pathway enrichment analyses were performed by using specialized bioinformatics tools and databases (i.e., Enrichr, IPA, Metacore). According to GO analysis, the CNV-driven DEGs in SALS1 were significantly enriched in biological processes such as *regulation of cellular component organization* (GO:0051129, P value=0.0002), *DNA conformation change* (GO:0071103, P value=0.0004) and *regulation of neuron death* (GO:1901214, P value=0.0012), whereas *regulation of synaptic transmission* (GO:0050804, P value=3.45E-12) and *learning and memory* (GO:0007611, P value= 5.26E-12) were the most enriched in SALS2 (**Figure 3a,**

Supplementary Table 9). On the basis of molecular function, the CNV-driven DEGs in SALS1 prominently accumulated in *small conjugating protein ligase binding* (GO:0044389, P value=0.0004) and *ubiquitin protein ligase binding* (GO:0031625, P value=0.0004), while the CNV-driven DEGs in SALS2 were enriched in *amino acid binding* (GO:0016597, P value=7.00126E-06) and *transcription factor binding* (GO:0008134, P value=6.6232E-05) (**Figure 3a, Supplementary Table 9**).



b.

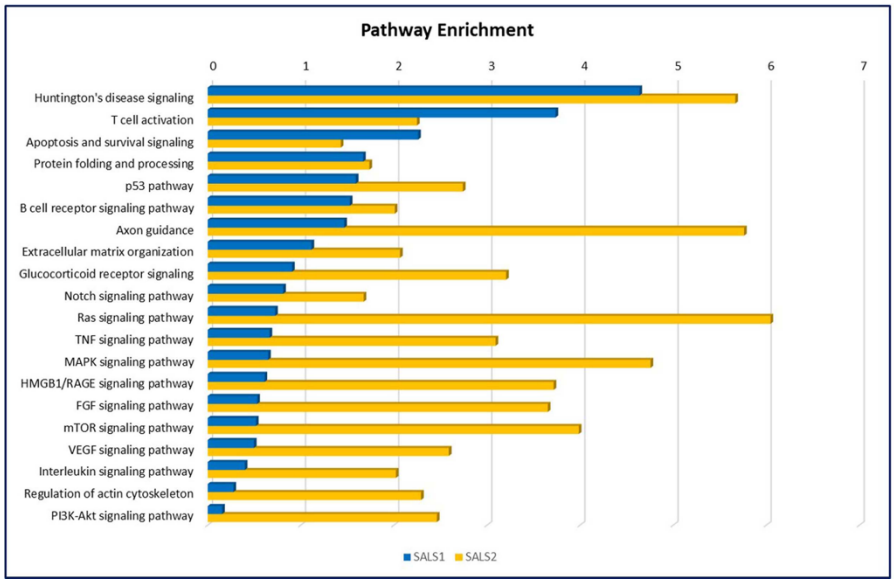


Figure 3. *Functional enrichment analysis for GO and pathway map ontologies revealed significant biological processes associated with the candidate CNV-driven genes in SALS.* ***(a)*** *Pie charts represent the top 10 enriched (P<0.05) GO terms for the 70 CNV-encompassed DEGs in SALS1 and SALS2 patients. The GO terms were subdivided into three GO categories: biological processes, molecular functions and cellular components. Enrichment analyses were performed using the Enrichment Analysis tool in Enrichr. For each category, GO terms or biological features represented in CNV-driven differently expressed genes are indicated.* ***(b)*** *Representation of the top 20 most significantly enriched (P value <0.05) canonical pathway maps associated with the candidate CNV-driven genes in SALS1 and SALS2 patients. A histogram of statistical significance (-log P value) is shown: the list is arranged in descending order with the most significant pathways at the top. The analysis was performed using the MetaCore™ pathway analysis suite.*

Further pathways enrichment analysis identified an immune/inflammatory response, cytoskeleton remodeling and apoptotic processes as the major deregulated processes in SALS (**Figure 3b, Supplementary Table 10**). In particular, *Huntington's disease signaling, T cell activation, Apoptosis and survival signaling* were the most significantly enriched pathways in both SALS patient subgroups, with *AKT1, NFKB1* and *SOS* as the major key involved genes. In addition, *ubiquitin-proteasome pathway, immune-cell mediated inflammatory response* and *apoptosis* were significantly up-regulated in SALS1, while *axonal guidance, oxidative stress* and *inflammatory intracellular signaling cascades* were mainly up-regulated in SALS2 (**Figure 4a**). Notably, the majority of the enriched pathway for CNV-driven DEGs in SALS patients represented processes already associated with ALS pathogenesis, such as immune response, cell adhesion and cell communication (**Figure 4a**).

To better understand the interactions of the CNV-driven genes and identify the best candidate genes in SALS, a protein-protein interaction (PPI) network analysis of their encoding products was performed, revealing a highly interconnected functional network, also including a greater number of ALS-associated genes (**Figure 4b**). The PPI network consisted of 147 nodes and 2787 edges, including 46 CNV-driven DEGs in SALS. Node degree ≥10 was selected as the threshold. *UBA52, RPS27A* and *HIST2H3A* were selected as the hub genes.

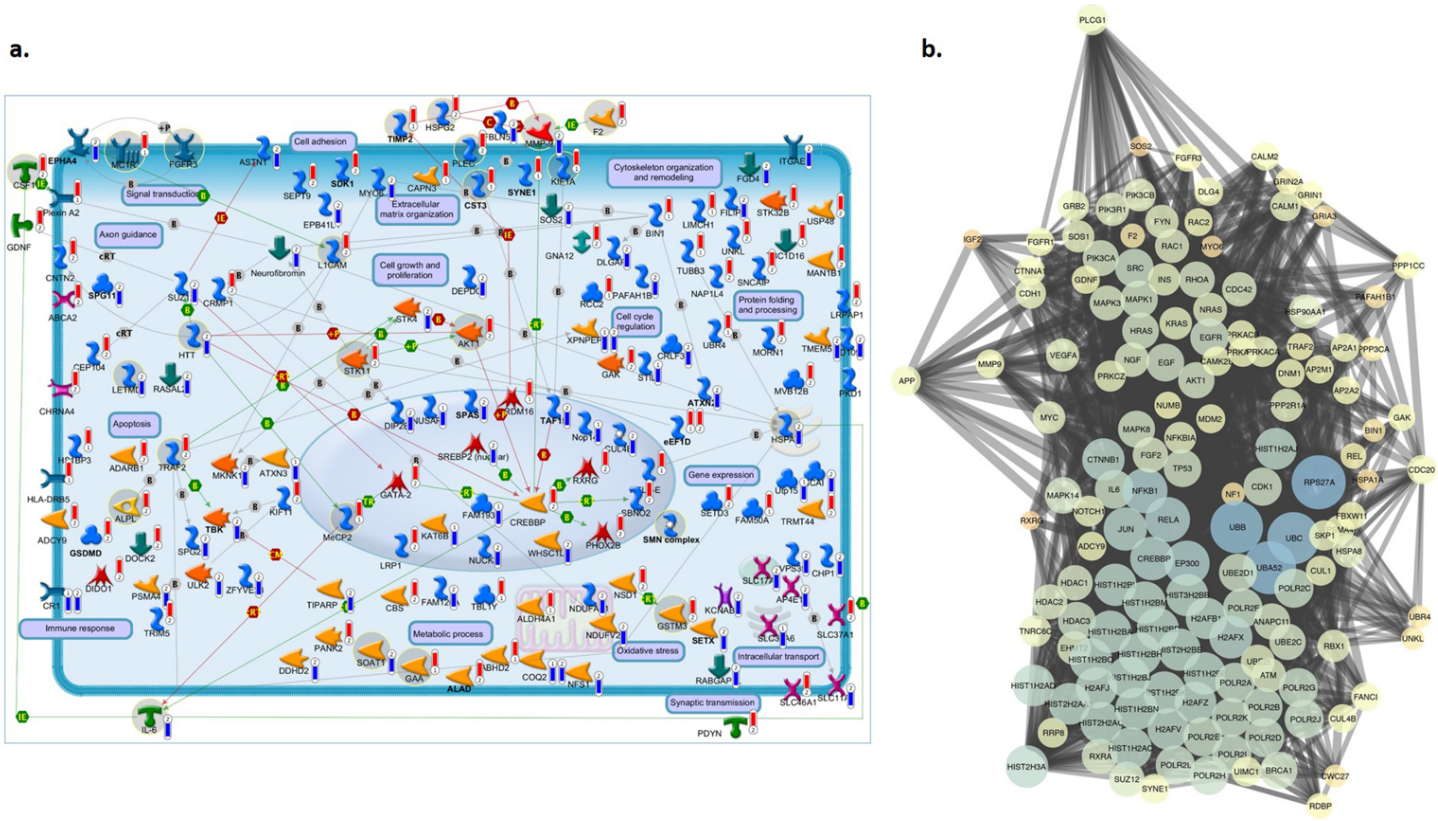


Figure 4. *Definition of SALS subtype-specific genomic signature using pathway and network analyses.* ***(a)*** *A representative illustration showing the functional correlation between SALS associated CNV-driven genes and their biological processes. Interaction map represents the most promising candidate genes showing a positive correlation between gene expression and underlying genomic changes, grouped on the basis of the main biological processes associated with them. The map was created using the MetaCore Pathway Map Creator tool (GeneGo). Gene expression and CNV values are presented on the map as 'thermometer-like' figures with SALS1 patients data represented as thermometer #1 and SALS2 patients #2. Genes associated with overexpression and CNV gain regions are labeled with red dots while genes associated with downregulated expression and homozygous or heterozygous deleted CNVs are labeled with blue dots. A detailed legend for the network objects is shown in the* ***Supplementary Figure 1****.* ***(b)*** *Functional network of known and predicted interactions of the most promising candidate CNV-driven genes. The network was produced by the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) v10 (http://string-db.org/) using default settings. Proteins are represented by spheres. Lines linking proteins indicate evidence for interactions: a red line indicates the presence of gene fusion (genes that are sometimes fused into single open reading frames); a green line – gene neighborhood (genes that reside within 300 bp on the same strand in the genome); a blue line – co-occurrence (gene families whose occurrence patterns across genomes show similarities); a purple line - experimental evidence (interaction extracted from protein-protein interaction databases); a yellow line – text mining (interaction extracted from scientific literature); a light blue line - database (interaction extracted from curated databases); a black line – co-expression (proteins whose genes are co-expressed in the same or in other species).*

To assess the value of CNV-driven genes as potential biomarkers for patient-specific diagnosis and prognosis, we reviewed and analyzed the literature on the CNV-driven genes exhibiting the same expression tendencies. A total of 24 candidate gene markers were selected, including 6 up-regulated CNV-driven DEGs in SALS1 and 18 deregulated (10 up-regulated and 8 down-regulated) CNV-driven genes in SALS2. Interestingly, some of these candidate genes (*TIMP2, AKT1, MMP9, CST3, SMN1* and

SMN2) were previously associated with susceptibility to ALS while the remaining 18 genes (*GAA, KIF1A, MC1R, MECP2, ALPL, HSPG2, L1CAM, PLEC, STK11, CSF1, F2, GSTM3, TRAF2, HSPA5, HTT, IL6, LETMD1, SOAT1*) represent novel candidate mediators for disease progression. The set of CNV-driver genes also included many patient-specific 'druggable' genes that may represent good candidates for the development of personalized, molecularly targeted therapies for SALS patients (**Table 4**).

Table 4. The most promising candidate CNV-driven genes and their utility as potential biomarkers/targets for SALS.

SALS1-specific Biomarkers/Targets									
Symbol	**Gene Name**	**Location**	**Family**	**Molecular aberration**	**Tissue/ Cells**	**Biological Fluid detectability**	**Biomarker Application**	**Diseases**	**Drug**
CST3	cystatin C	Extracellular Space	other	Gain/UP	Brain, Cerebral Cortex	Blood, Plasma/Serum, Bronchoalveolar Lavage Fluid, Cerebral Spinal Fluid, Saliva, Sputum, Synovium/ Synovial Fluid, Tears, Urine	diagnosis, efficacy, prognosis, safety	Immunological /Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-
GAA	glucosidase alpha, acid	Cytoplasm	enzyme	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Bronchoalveolar Lavage Fluid, Urine	diagnosis	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Voglibose, Miglustat
KIF1A	kinesin family member 1A	Cytoplasm	other	Gain/UP	Cerebral Cortex	Blood, Plasma/Serum	diagnosis, prognosis	Immunological Disease, Neurological Disease, et al	-
MC1R	melanocortin 1 receptor	Plasma Membrane	G-protein coupled receptor	Gain/UP	Brain	Blood	diagnosis	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders	-
MECP2	methyl-CpG binding protein 2	Nucleus	transcriptio n regulator	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Not detected in biofluid	unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-
TIMP2	TIMP metallopeptidas e inhibitor 2	Extracellular Space	other	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Blood, Plasma/Serum, Cerebral Spinal Fluid, Urine	efficacy	Immunological/ Inflammatory Disease, Skeletal and Muscular Disorders, et al	Pravastatin , ABT751 (inhibitors)

SALS2-specific Biomarkers/Targets									
Symbol	**Entrez Gene Name**	**Location**	**Family**	**Molecular aberration**	**Tissue/ Cells**	**Biological Fluid detectability**	**Biomarker Application**	**Diseases**	**Drug**
AKT1	AKT serine/threonine kinase 1	Cytoplasm	kinase	Gain/UP	Brain, Cerebral Cortex	Blood	diagnosis, efficacy, response to therapy	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Thalidomide, XL418, ABT100 (inhibitors)
ALPL	alkaline phosphatase, liver/bone/kidney	Plasma Membrane	phosphatase	Gain/UP	Spinal Cord	Blood, Plasma/Serum, Bronchoalveolar Lavage Fluid, Urine	efficacy, safety	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Alendronic acid, Zoledronic acid, Dexamethasone, Zaprinast (inhibitors)
HSPG2	heparan sulfate proteoglycan 2	Extracellular Space	enzyme	Gain/UP	Spinal Cord	Blood, Plasma/Serum, Cerebral Spinal Fluid, Bronchoalveolar Lavage Fluid Tears, Urine	unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-
L1CAM	L1 cell adhesion molecule	Plasma Membrane	other	Gain/UP	Brain, Cerebral Cortex	Blood, Plasma/Serum, Synovium/Synovial Fluid, Urine	unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Ethanol, Rivanicline, Anabaseine, Ketamine, Mecamylamine, Bupropion (inhibitors)
PLEC	plectin	Cytoplasm	other	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Blood, Plasma/Serum	unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-
STK11	serine/threonine kinase 11	Cytoplasm	kinase	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Blood	efficacy	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-
CSF1	colony stimulating factor 1	Extracellular Space	cytokine	Gain/UP	Brain, Cerebral Cortex, Spinal Cord	Blood, Plasma/Serum, Cerebral Spinal Fluid, Bronchoalveolar Lavage Fluid, Synovium/Synovial Fluid, Urine	diagnosis	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	-

F2	coagulation factor II, thrombin	Extracellular Space	peptidase	Gain/UP	Brain, Cerebral Cortex	Blood, Plasma/Serum, Cerebral Spinal Fluid, Bronchoalveolar Lavage Fluid, Tears, Urine	diagnosis, unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	enoxaparin, desirudin, dabigatran etexilate, Fibrinogen, ximelagatran, thrombin inhibitor, antithrombin alfa, aspirin/dabigatran etexilate, dabigatran, ulinastatin, aspirin/ bivalirudin, argatroban, bivalirudin, lepirudin
GSTM3	glutathione S-transferase mu 3	Cytoplasm	enzyme	Gain/UP	Cerebral Cortex	Blood, Plasma/Serum, Urine	diagnosis, prognosis	Neurological Disease, et al	-
TRAF2	TNF receptor associated factor 2	Cytoplasm	enzyme	Gain/UP	Brain	Blood, Plasma/Serum	diagnosis	Immunological/ Inflammatory Disease, et al	-
HSPA5	heat shock protein family A (Hsp70) member 5	Cytoplasm	enzyme	Loss/DOWN	Amygdala, Brain, Cerebellum, Cerebral Cortex	Blood, Plasma/Serum, Synovium/Synovial Fluid, Tears, Urine	efficacy, unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Mefloquine (activation)
HTT	huntingtin	Cytoplasm	transcription regulator	Loss/DOWN	Brain, Cerebral Cortex	Blood, Plasma/Serum	unspecified application	Immunological/Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	
IL6	interleukin 6	Extracellular Space	cytokine	Loss/DOWN	Cortical neurons, Brain, Cerebral Cortex, Spinal Cord	Blood, Plasma/Serum, Synovium/Synovial Fluid, Tears, Urine	diagnosis, disease progression, efficacy, prognosis, response to therapy, safety, unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	Tocilizumab, Siltuximab, Mifamurtide, Mycophenolate mofetil, Acetaminophen, Rifampicin, Prostaglandin E2, Morphine, Hyaluronic acid (activators)
LETMD1	LETM1 domain containing 1	Plasma Membrane	other	Loss/DOWN	Brain, Cerebral Cortex	Blood, Plasma/Serum	unspecified application	Neurological Disease, et al	
MMP9	matrix metallopeptidase 9	Extracellular Space	peptidase	Loss/DOWN	Brain, Cerebral Cortex	Blood, Plasma/Serum, Bronchoalveolar Lavage Fluid, Saliva, Sputum, Synovium/Synovial Fluid, Tears, Urine	diagnosis, disease progression, efficacy, prognosis, unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	GS-5745, Rebimastat, Marimastat, Prinomastat, Glucosamine, Ciprofloxacin, Aclarubicin, Prostaglandin E2, Phorbol 12-myristate 13-acetate (activators)
SMN1/ SMN2	survival of motor neuron 1, telomeric	Nucleus	other	Loss/DOWN	Brain, Cerebral Cortex, Spinal Cord	Blood	diagnosis	Neurological Disease, Skeletal and Muscular Disorders, et al	

SOAT1	sterol O-acyltransferase 1	Cytoplasm	enzyme	Loss/DOWN	Pituitary Gland	Blood	unspecified application	Immunological/ Inflammatory Disease, Neurological Disease, Skeletal and Muscular Disorders, et al	pactimibe, ezetimibe/ fluvastatin, atorvastatin/ezetimibe, ezetimibe/rosuvastatin, ezetimibe/fenofibrate, ezetimibe/simvastatin, ezetimibe, hesperetin

The table lists CNV-driven genes showing the same expressional tendencies between DNA copy number and mRNA expression in SALS subgroups as potential candidate biomarkers and therapeutic targets for ALS. Target and biomarker assessment was performed by using dedicated tools in IPA and MetaCore.

UP: upregulation

DOWN: downregulation

DISCUSSION

In this work, we reported the first fully integrated analysis of CNVs and gene expression profiling derived from the same SALS patients to provide a more comprehensive genomic framework for dissecting molecular heterogeneity of ALS and identify the DEGs with alterations in genomic segments that may represent novel potential markers and/or therapeutic targets.

Taking advantage of the custom-made *NeuroArray* platform, designed to uncover CNVs in clinically relevant genes for ALS and other neurological diseases 25, we performed an exon-focused evaluation of structural imbalances occurring in motor cortex samples from 30 SALS and 10 control patients trascriptomically characterized in our previous work[17]. A large number of aberrations were detected in over 10% of SALS patients, with the highest number of gains documented at chromosomes 14 and 17, and the majority of losses found on the p arm of chromosome 1 and at cytoband 20q11.22-q11.23 (**Figure 1c** and **Supplementary Table 1**). Notably, at 17p region the *SREBF1*, a gene encoding a lipogenic transcription factor whose expression levels were increased in the spinal cords of FALS and SALS patients as well as in ALS animal models, and

whose direct causative role in excitotoxicity-induced neuronal cell death, has been extensively established[26-28].

Our analysis also revealed distinct genomic signatures associated with two previously characterized transcriptome-based SALS subgroups. In particular, the loss of chromosome 1p was the most common chromosomal aberration in SALS1 (~90%), while 60-70% of the SALS2 patients showed simultaneous deletion of precise regions of chromosomes 3q and 18p (**Figure 1c, Table 1, Supplementary Table 1**)[29]. It is interesting to note that genes in these CNV regions showed the same subgroup-specific expressional tendencies. In fact, the 1p33 deletion includes *STIL*, a gene involved in neural protection and survival and whose expression was down-regulated in SALS1, suggesting that a "loss-of-function" of this gene may contribute to render motor neurons vulnerable to excitotoxic insults in these patients[30]. Likewise, SALS2 patients carried the deletion of 18p11.22 that encompassed *NDUFV2*, one of the many components of the mitochondrial oxidative phosphorylation pathway, whose expression levels were decreased in the same patient subgroup. Defects in this subunit have been associated with altered energy production, mitochondrial dysfunction and oxidative stress, representing a risk factor for several neuronal diseases, including ALS[31].

Compared with previous CNV reports[32], some genomic aberrations identified in our study partially or totally overlap with those previously associated with ALS, further supporting them as disease susceptibility variants (**Table 3**). Among these, of particular interest is the deletion of chromosomal region 15q22.2-q22.31 that was detected exclusively in SALS patients and not altered in any controls in our study nor listed in the Database of Genomic Variants – DGV.

Besides strengthening previously reported results, we identified new potential susceptibility loci that were overrepresented in SALS patients and absent in controls

(**Table 2**). Among these, we distinguished the deletion at the chromosome 1q32.2, which encompassed *CR1* (also known as *CD35*), a member of the human regulator of complement activation gene cluster[33,34]. This gene encodes one of the major immune adherence receptors and plays an important role in immune complex processing and clearance via reducing activation of classical and alternative complement cascade activity. Deletion of *CR1* is consistent with the lower mRNA expression level of this receptor in SALS patients, providing support for aberrant complement regulation as a part of ALS process and highlighting the potential use of complement molecules as disease biomarkers[35,36]. Another interesting association is represented by a 162.362-kb duplicated region in 1p36.12 spanning several neuronal genes, including the *FBXO42* encoding an important member of the F-box protein family involved in the ubiquitin-proteasome system and already known to be associated with ALS (**Table 2**)[37]. Of particular interest, three CNV regions (17q12, 22q11.21; 6q13) were selectively detected in specific SALS patient subgroups and absent in all control individuals, providing the basis for a CNV-based molecular classification of the disease (**Table 2**). In particular, the deletion at the 6q13 region harbors the *SLC17A5* gene encoding sialin, a vesicular excitatory amino acid transporter, whose loss-of-function leads to defect in myelin structure and function, contributing to the disruption of axonal integrity and the motor phenotype[38,39]. The increase in gene copy numbers for the 17q12 region encompassing several chemokines (i.e., *CCL5* and *CCL3*) is consistent with the observation that high mRNA expression levels of these chemokines may increase activation of the inflammatory system and changes in blood-brain barrier permeability, two key mechanisms implicated in, and possibly aggravating, motor neuron damage[40,41].

By integrating the analysis of CNVs and gene expression profiling in the same patients and tissue sample, we found that 71.2% of CNV genes were differentially

expressed in SALS patients in comparison to controls, the majority of them were cluster specific, further suggestive of the great divergence of two SALS subgroups at the molecular level (**Figure 2b-c, Supplementary Tables 7-8**). Among these genes, 49% showed a positive association between CNV and mRNA expression, suggesting them as potential driver genes in ALS. These included some genes (*SYNE1, SDK1, EEF1D, GSDMD, TIMP2, CST3, ALAD, AKT1, EPHA4, SPAST, SMN1, SETX, ATXN2, TBK1, SPG11, TAF15, MMP9*) previously reported as potential risk factors in ALS, as well as novel candidates whose association with ALS was previously unappreciated. These genes may represent overlapping genetic signatures among different neurological condition, providing additional features for exploring ALS pathogenesis. To name a few of those detected exclusively in SALS patients (i.e., absent from controls) are *ALDH4A1*, a component of the mitochondrial matrix contributing to protect cells from oxidative stress[42]; *BIN1*, the most significant late-onset susceptibility locus for Alzheimer's disease whose alterations in expression levels and splicing seem to induce muscle weakness and T tubule alterations[43], and *RABGAP1L*, a GTPase-activating protein whose loss-of-function exacerbates neuronal loss[44].

On the other hand, our analysis identified a number of CNV genes, including known ALS-related genes (i.e., *NIPA1, NAIP, VPS54* and *GRIN3B*), which do not appear to exert any apparent influence on expression levels, suggesting that the expression of these genes may be not gene-dose dependent and that they are likely to represent secondary 'passenger' events in ALS pathogenesis. Genes differently regulated at the transcriptomic and genomic level (i.e., showing no or negative correlation), in turn, could result from other, non-CNV-related regulatory mechanisms, such as those associated with gene mutation, promoter methylation, and non-coding RNA regulation (**Figure 2b-c, Supplementary Tables 7-8**). An example is *KIF5A*, a

gene encoding a neuronal kinesin heavy chain that acts as a molecular motor and whose genetic alteration was recently associated with ALS susceptibility10. SALS2 patients showed a selective downregulation of *KIF5A* not overlapping with any CNV, suggesting reduced mRNA expression of this gene is probably due to defective splicing events and/or loss-of-function variants rather than to deletion within its genomic region. It is interesting to note that in the same subgroup we also observed defects in axonal transport as well as a general downregulation of genes involved in mitochondrial oxidative phosphorylation machinery, suggesting a link between loss-of-function events in *KIF5A* and impaired transport and dysfunction of mitochondria in ALS17,45.

Integrated analysis of CNVs and corresponding expression data represents an effective approach to elucidate mediators and mechanisms involved in ALS. However, it has become clear that the process of motor neuronal degeneration is complex and requires many genomic alterations acting in concert. This also emerged by our analysis of PPI network constructed for CNV-driven DEGs that revealed a great number of close interconnections and identified three hub genes (*UBA52, RPS27A* and *HIST2H3A*) involved in inflammatory and immune responses (**Figure 4b**)[46]. We also put forward a 'systems biology' analysis to identify biological processes and signaling pathways that were overly represented by CNV-driven genes in both SALS patient subgroups. Functional enrichment analysis of these candidate genes showed that they were mainly involved in immune/inflammatory signaling, neuronal migration, differentiation and survival, and neurite outgrowth, supporting the concept that these pathways may crosstalk with each other to drive the disease pathogenesis (**Fig 4a**). In particular, Huntington's disease (HD) signaling and Protein folding and processing were among the top canonical pathways for both SALS patient subgroups, supporting the possibility of a causal link between protein aggregation, neurotoxicity and disease severity in ALS

and other neurological diseases, like HD[47,48] (**Fig 4a**). Among the most important candidate genes involved in these processes is *AKT1*, a gene encoding a cAMP-dependent serine-threonine kinase that was amplified in both SALS subgroup and overexpressed in SALS2. Our results are supported by previous evidence that abnormal AKT activation is implicated in several cellular mechanisms involved in ALS, such as the altered elimination of toxic protein aggregates, increasing oxidative stress and rendering cells susceptible to ROS-triggered cell death[49]. Our PPI network analysis identified *AKT* as a highly interconnected node, suggesting that alterations in this protein are not mere passenger events, but may have a great impact on one or more signaling pathways that are recurrently involved in ALS. Pharmacological inhibition of AKT and its downstream pathways has already demonstrated neuroprotective effects by modulating the activation state of microglial cells during neuroinflammation, and promoting cellular clearance in neurodegenerative storage diseases, suggesting a potential role of AKT inhibitors in ALS treatment[50,51].

It is interesting to note that although partially complementary and convergent, the two SALS patient clusters showed different significantly deregulated processes and mediators. In particular, antigen processing and presentation, and *extracellular matrix organization* were the most significantly enriched pathways for the CNV-driven genes in SALS1, while the pathways of highest significance in SALS2 were associated with *axonal guidance, oxidative stress* and *inflammatory intracellular signaling cascades* (**Fig 4a**). Therefore, a careful monitoring of these signaling cascades may help to better diagnose the specific subtype of ALS and optimize treatment strategies. To this regard, the convergent functional analysis of CNV-driven genes also pinpointed known and novel candidate therapeutic targets and biomarkers for early diagnosis, molecular subtyping and targeted therapy in SALS (**Table 4**). In particular, among CNV-driven

genes in SALS1 showing the same expressional tendencies between DNA copy number and mRNA expression, we distinguished some genes (*KIF1A*, *MC1R* and *MECP2*) that were not previously implicated in ALS, representing new candidates for molecularly guided diagnosis and treatments. Within the cluster of CNV-driven genes in SALS2, we found a large number of new candidate ALS genes mainly implicated in oxidative and inflammatory signaling cascades[52-57]. Among these, identification of gain and overexpression of *TRAF2* was in accordance with previous evidence that correlated elevated expression levels of this gene with inflammatory processes in PD and other neurological disorders[58,59]. On the contrary, copy number loss and reduced expression levels *HSPA5* in SALS2 may reflect the suppression of neuroprotective role of this molecule against ER stress-associated cell death, leading to oxidative stress and alterations in calcium homeostasis, and rendering neurons vulnerable to degeneration 60-64. Concordantly, pharmacological activation of HSPA5 and its co-chaperones has demonstrated to exert neuroprotective effects on motor neurons of ALS by reducing ER stress-mediated cell death, supporting a translational potential for HSPA5 induction as a therapy against ALS and other neurologic disorders[65].

Overall, our study provided the first comprehensive and integrated map of genomic and transcriptional events characterizing different SALS subtypes, revealing key drivers and etiopathogenic mechanisms that may have been masked by considering SALS pathology as a single entity. Despite the most obvious limitation in this research was that of a small and unbalanced number of samples, our brain tissue samples represent a vanishingly rare resource for investigating molecular mechanisms underlying neurological disorders. The importance of an integrative analysis such as the one presented here, emerges from recent published data that highlight the presence of somatic changes in brain tissue of patients affected by neurological diseases[66,67].

Moreover, the custom-designed platform used in this study has some disadvantages, including limited use for discovering novel genes or gene features; inability to detect nucleotide repeat expansions or balanced structural chromosomal abnormalities. However, the possibility to simultaneously detect multiple genes involved in neurological disorders may allow for differential diagnosis between common neurological disorders, refine the genotype-phenotype correlations and explore the potential genetic overlapping signatures among different neurological conditions.

In conclusion, the present study proposes the use of a multi-omics analysis as a promising approach for the identification of somatic alteration and candidate drivers in ALS for defining disease subtype and directing molecular targeted clinical trials that more accurately reflect inter-individual differences among patients. Future more in-depth functional and integrative omics studies will be necessary to verify our findings and explore the impact of candidate genes on the outcome of the disease.

MATERIAL AND METHODS

The analysis workflow is shown in **Figure 5** and described below.

Subject cohorts and sample preparation

All samples were provided by the Department of Pathology of the Academic Medical Center (University of Amsterdam). This cohort included motor cortex samples from 30 patients with clear SALS diagnosis and 10 control individuals collected as previously described[17], and whose clinico-pathological parameters are detailed in the **Supplementary Table 12**. Informed consent was obtained for the use of brain tissue and for access to medical records for research purposes, and approval was obtained from the relevant local ethical committees for medical research. All experiments were performed in accordance with relevant guidelines and regulations of both institutions. Genomic DNA was extracted from 10µm-thick sections using the QIAamp Fast DNA Tissue Kit according to the manufacturer's instructions (QIAGEN, Hilden, Germany). The extracted genomic DNA was quantified by using the NanoDrop ND-1000

spectrophotometer (Thermo Fisher Scientific, MA, USA), and assessed for quality by microcapillary electrophoresis on 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

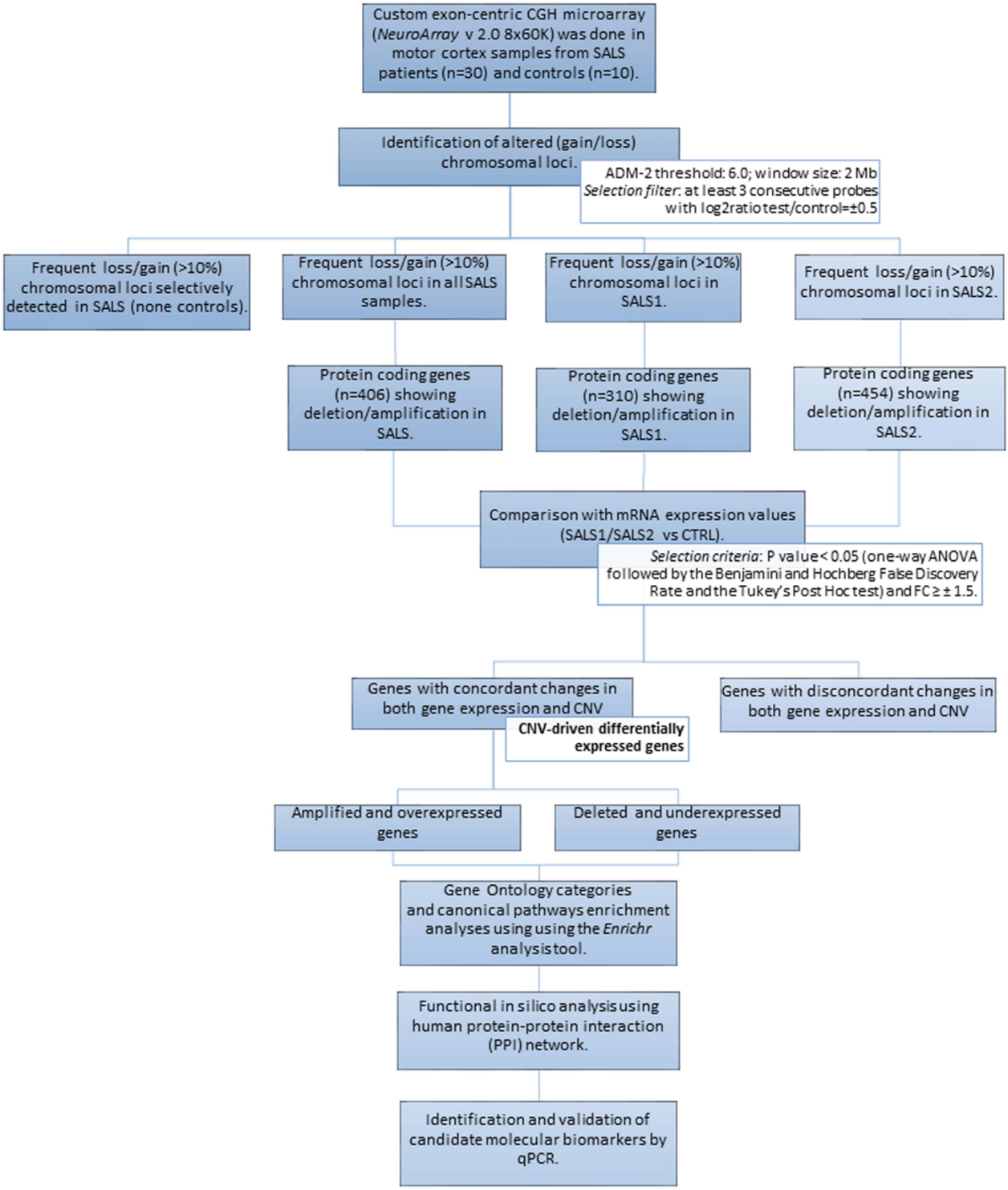


Figure 5. *Experimental workflow of multi-omics analysis for characterization of CNV-driven differentially expressed genes in SALS patients. The workflow depicts the steps performed in this study, from data acquisition to the visualization, validation and export of results in various output formats. See Materials and Methods section for details.*

NeuroArray aCGH processing and data analysis

High-resolution exon-centered analysis of CNVs was done using an 8x60K custom exon-centric *NeuroArray* platform v.2.0 (Agilent Technologies, Santa Clara, CA), tailored to detect single/multi-exon deletions and duplications in a large panel of ALS-related genes (n=154) and to others additional neurological disorders (**Supplementary Table 13**)[25]. Details concerning the *NeuroArray* aCGH platform can be found in **Supplementary Materials**. DNA labeling and hybridization on NeuroArray were performed according to the manufacturer's protocol (Agilent Technologies, Santa Clara, CA). Briefly, DNA test and a reference of the same sex (Euro Reference, Agilent Technologies, Santa Clara, CA), both at the concentration of 500 ng, were double digested with RsaI and AluI for 2 hours at 37°C. After heat inactivation of the enzymes at 65°C for 20 min, each digested sample was labeled by random priming by using the genomic DNA Enzymatic Labelling Kit (Agilent Technologies, Santa Clara, CA) for 2 hours using Cy5-dUTP for patient DNAs and Cy3-dUTP for reference DNAs. Labeled products were column purified by using the SureTag DNA Labeling Kit Purification Columns (Agilent Technologies, Santa Clara, CA). After probe denaturation and pre-annealing with Cot-1 DNA, hybridization was performed at 65°C with rotation for 24 hr. After two washing steps, arrays were scanned at 3 μm resolution using an Agilent G4900DA SureScan Microarray Scanner System and aCGH image data were processed using Agilent's Feature Extraction software to assess the array spot quality as well as check signal and background intensity statistics in the default setting. Feature extracted raw data was normalized, analyzed and visualized using Agilent CytoGenomics v. 4.0.3.12 and Genomic Workbench v. 7.0.4.0 software (Agilent Technologies, Santa Clara, CA, USA). Briefly, after filtering for saturated and non-

uniform probes, data were normalized by GC correction with a window size of 2 kb and Diploid Peak Centralization. The Centralization Normalization Algorithm with a threshold of 6.0 and a bin size of 10 was also used for detecting aberrant regions or regions of constant CNVs. Aberrations were detected by the Aberration Detection Method II algorithm (ADM-2), with a sensitivity threshold of 6.0 and moving average window of 2Mb, which permits to identify all aberrant intervals in a given sample with consistently high or low log ratios based on the statistical score. An aberration filter was applied for identifying copy number alterations; changes were considered as true positive events with a minimum log2ratio test/control of ±0.5 and a minimum of 3 consecutive probes. Positive statistical score meant amplification, while a negative score indicated deletion. Human reference sequence hg19 assembly was used to define the genomic coordinates of detected CNVs. Raw data of the microarrays are available at NCBI's Gene Expression Omnibus (GEO) with the accession number *GSE107375*.

Identification of significantly altered genomic regions and CNV-encompassed genes

For statistical analysis, the ALS samples were divided into two groups (SALS1 and SALS2) based on their previously characterized gene expression profiles[17]. The chromosomal distribution and the frequency of the copy number gains and losses in both SALS subgroups were also investigated. Using ADM-2 generated interval-based amplification and deletion data, penetrance analysis was performed to find the percentage of samples that share aberrations in a particular genomic region among multiple samples. A recurrent CNV was called when the gains or losses occurred in at least 10% of the SALS patients, both with and without the assignment into the hierarchically determined two sets. Multiple amplifications and deletions were counted

as separate events. Aberrant intervals were also filtered taking into account those occurring in at least 10% of the cases and absent in individual controls. In addition, to assess the effective relations between the detected CNVs and ALS pathogenesis, we compared these aberrant regions with those previously associated with ALS via screening of publicly available databases (i.e., CNVD) and published literature.

Frequent amplifications and deletions observed in both SALS patient subgroups were reviewed and annotated to the human hg19 reference genome and then were screened out for only genes included in the *NeuroArray* aCGH design. In addition, significant probe signals were clustered for pathologies according to their location on causative or susceptibility genes through a homemade script on R-platform, in order to search for CNVs in candidate genes of ALS disease[68].

Source of gene expression data

Gene expression data set *E-MTAB-2325* was downloaded from EBI ArrayExpress database, which was annotated using the platform of GPL6480 (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F). A total of 40 samples were selected out, including 10 normal and 30 SALS motor cortex samples. Raw signal values were thresholded to 1, log2 transformed, normalized to the 50th percentile, and baselined to the median of all samples using GeneSpringGX v.14.5 (Agilent Technologies, Italy). Fold change (FC) values were calculated between SALS patients and individual controls. Positive FC meant over-expression, whereas negative FC indicated under-expression. Probes not corresponding to an ENTREZ ID were removed. In cases where several probes corresponded to one ENTREZ ID, the probe showing the highest variance over all samples was chosen for further analysis. Genes that showed a significant P value < 0.05 (one-way ANOVA followed by the Benjamini and Hochberg False Discovery Rate

and the Tukey's Post Hoc test) and FC ≥ ± 1.5 were considered differentially expressed and were taken for further analysis.

Integration of the aCGH and gene expression data

To assess the contribution of genomic aberrations to global gene expression pattern changes in SALS and identify CNV-associated DEGs, we performed an integrated analysis of differential expression values and the corresponding DNA copy number changes through a gene-by-gene approach. In particular, each gene expression measurement was assigned to the corresponding copy number probe interrogating the same named gene. A CNV-driven gene was defined when the gene expression trend was consistent with the copy number change (i.e., up-regulated gene transcript with a chromosomal amplification and down-regulated gene transcript with a chromosomal deletion).

Functional enrichment and biological network analysis

The function of CNV-associated DEGs in SALS patients was annotated and analyzed according to the three organizing principles of Gene Ontology (BP: Biological Process, MF: Molecular Function, CC: Cellular Component) by using the enrichment analysis tool Enrichr[69]. To interpret the biological significance of CNV-driven genes in the context of known biological pathways, we used QIAGEN's Ingenuity Pathway Analysis (IPA®; http://www.ingenuity.com/) and MetaCore repository (Clarivate Analytics, Philadelphia, United States)[70]. Both these programs identify significantly enriched biological pathways and signaling cascades that are associated with a given list of genes by calculating the hypergeometric distribution. P-value <0.05 was set as the threshold to filter out significant terms. In addition, to increase the statistical power of our analysis, we compared our results with three other pathway enrichment analysis tools

and databases (KEGG; Reactome; Panther) and selected signaling pathways that were identified as significantly deregulated by two or more platforms.

To better understand the interactions of the CNV-driven genes and identify the best candidate genes in SALS, an extended protein-protein interaction (PPI) network of their encoding products was predicted by using the STRING database[71] and visualized with the Cytoscape v.3.5.0 software[72]. The extended network was constructed by using the CNV-driven genes as seed molecules and setting a high level of confidence between molecular interactions (high confidence score of at least 0.8) and a maximum number of interactions to 100. In order to identify the "Hub" nodes, a network topology analysis was performed by using the Cytoscape plug-in NetworkAnalyzer based on topological parameters 73. The relative importance of the genes in each network, meaning their ability to hold together the communicating nodes in a biological network, was determined based on the node centrality measure setting the topological parameter "node degree" ≥10. Nodes with high degree (hub genes) represented the genes having important biological functions: the higher the value, the higher the relevance of the gene in connecting regulatory molecules. Likewise, values of edge betweenness were mapped with the edge size: high values of this parameter correspond to a large edge size. The final PPI network was visualized based on node degree and edge betweenness parameters.

Finally, target and biomarker assessment tools in both IPA and Metacore were used to screen candidate CNV-driven genes in order to identify potential candidate biomarkers and therapeutic targets for ALS.

Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR) validation

To confirm the reliability of our data, we validated the *NeuroArray* CGH results performing ad hoc real-time quantitative polymerase chain reaction (RT-qPCR) assays. Briefly, we used DNA extracted from the motor cortex samples of 15 donors assayed by NeuroArray, including 5 controls, 4 SALS1 and 6 SALS2 (**Supplementary Table 11**). From the list of CNV-driven DEGs genes, we selected 5 candidates (*GAA, KIF1A, CSF1, TRAF2, HSPA5*) on the basis of their potential clinical relevance as patient-specific biomarkers and therapeutic targets (**Supplementary Table 11**). The primer sets flanking the putative exonic imbalances were designed using the PrimerBlast tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) as described 74 and were available in **Supplementary Table 11**. RT-qPCR was performed in triplicate using the LightCycler 1.5 (Roche Diagnostics, Germany). Cycling conditions were 95°C for 15 s followed by 40 cycles of 95°C (5 s), 60°C (15 s) and one cycle of 95°C (15 s), 60°C (60 s), 95°C (15 s). The data were analyzed by the 2-ΔΔCt method that requires a healthy control sample (diploid) as a calibrator in all amplifications 75. 2-ΔΔCt$\geq$ 1.4 or $\leq$ 0.6 was defined as copy number gain or loss, respectively, whereas 2-ΔΔCt values included from 0.8 to 1.2 were considered as normal diploid. As calibrator control, we used the same DNA reference hybridized in *NeuroArray* experiments. The specific PCR products were confirmed by the results of melting curve analysis and agarose gel electrophoresis.

DECLARATIONS

Ethics approval and consent to participate

Experiments involving human participants have been approved by an ethical committee (Ethics Committee of the Amsterdam Academic Medical Center, approved protocol:

W11_073) for medical research and have been performed in accordance with ethical standards. Informed consent Informed consent was obtained from all individual participants included in the study.

Availability of data and materials

All data generated during this study are included in this published article and the additional files. Transcriptional data are available at EBI ArrayExpress database with the accession number E-MTAB-2325. Raw data from NeuroArray aCGH analysis are available at NCBI's Gene Expression Omnibus (GEO) with the accession number GSE107375.

Funding

This work was supported by the Italian Ministry of Education, Universities and Research through grant CTN01_00177_817708 and the international PhD programs in Neuroscience and Computer Science of the University of Catania.

Acknowledgements

Authors gratefully acknowledge Valentina La Cognata for her significant contribution during the array design, and Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo, Francesco Marino, Francesco Rabboni for their administrative and technical support.

Author contributions

G.M. and M.G. performed experiments, analyzed and interpreted data. G.M. wrote the manuscript. A.G.S. performed bioinformatics analyses. EA conducted tissue sample preparation. S.S., V.D. and F.L.C. participated in revising the manuscripts. S.C. conceived, directed, and supervised the project. All authors have read and approved the

final version of this manuscript and agreed to be accountable for all aspects of the work. Correspondence and requests for materials should be addressed to the author at sebastiano.cavallaro@cnr.it.

Competing interests

The authors declare no competing interests.

Supplementary Material can be made available in the online version of this article.

REFERENCES

1 Wijesekera, L. C. & Leigh, P. N. Amyotrophic lateral sclerosis. *Orphanet journal of rare diseases* **4**, 3, doi:10.1186/1750-1172-4-3 (2009).

2 Costa, J. & de Carvalho, M. Emerging molecular biomarker targets for amyotrophic lateral sclerosis. Clinica chimica acta; international journal of clinical chemistry 455, 7-14, doi:10.1016/j.cca.2016.01.011 (2016).

3 Simpson, C. L. & Al-Chalabi, A. Amyotrophic lateral sclerosis as a complex genetic disease. *Biochimica et biophysica acta* **1762**, 973-985, doi:10.1016/j.bbadis.2006.08.001 (2006).

4 Bruijn, L. I., Miller, T. M. & Cleveland, D. W. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annual review of neuroscience* **27**, 723-749, doi:10.1146/annurev.neuro.27.070203.144244 (2004).

5 Morello, G., Spampinato, A. G. & Cavallaro, S. Molecular Taxonomy of Sporadic Amyotrophic Lateral Sclerosis Using Disease-Associated Genes. *Frontiers in neurology* **8**, 152, doi:10.3389/fneur.2017.00152 (2017).

6 Volk, A. E., Weishaupt, J. H., Andersen, P. M., Ludolph, A. C. & Kubisch, C. Current knowledge and recent insights into the genetic basis of amyotrophic lateral sclerosis. *Medizinische Genetik : Mitteilungsblatt des Berufsverbandes Medizinische Genetik e.V* **30**, 252-258, doi:10.1007/s11825-018-0185-3 (2018).

7 Al-Chalabi, A., van den Berg, L. H. & Veldink, J. Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nature reviews. Neurology* **13**, 96-104, doi:10.1038/nrneurol.2016.182 (2017).

8 Pampalakis, G. *et al.* New molecular diagnostic trends and biomarkers for amyotrophic lateral sclerosis. *Human mutation*, doi:10.1002/humu.23697 (2018).

9 Caroline, E. The genetic basis of amyotrophic lateral sclerosis: recent breakthroughs. *Advances in Genomics and Genetics.* **5**, 327—345, doi:https://doi.org/10.2147/AGG.S57397 (2015).

10 Nicolas, A. *et al.* Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* **97**, 1268-1283 e1266, doi:10.1016/j.neuron.2018.02.027 (2018).

11 Brenner, D. *et al.* Hot-spot KIF5A mutations cause familial ALS. *Brain : a journal of neurology* **141**, 688-697, doi:10.1093/brain/awx370 (2018).

12 Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444-454, doi:10.1038/nature05329 (2006).

13 Lupski, J. R. Genomic rearrangements and sporadic disease. *Nature genetics* **39**, S43-47, doi:10.1038/ng2084 (2007).

14 Blauw, H. M. *et al.* Copy-number variation in sporadic amyotrophic lateral sclerosis: a genome-wide screen. *The Lancet. Neurology* **7**, 319-326, doi:10.1016/S1474-4422(08)70048-6 (2008).

15 Wain, L. V. *et al.* The role of copy number variation in susceptibility to amyotrophic lateral sclerosis: genome-

wide association study and comparison with published loci. *PloS one* **4**, e8175, doi:10.1371/journal.pone.0008175 (2009).

16 Katyal, N. & Govindarajan, R. Shortcomings in the Current Amyotrophic Lateral Sclerosis Trials and Potential Solutions for Improvement. *Frontiers in Neurology* **8**, doi:10.3389/fneur.2017.00521 (2017).

17 Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiology of disease* **74**, 359-376, doi:10.1016/j.nbd.2014.12.002 (2015).

18 Morello, G. & Cavallaro, S. Transcriptional analysis reveals distinct subtypes in amyotrophic lateral sclerosis: implications for personalized therapy. *Future medicinal chemistry* **7**, 1335-1359, doi:10.4155/fmc.15.60 (2015).

19 Lederer, C. W., Torrisi, A., Pantelidou, M., Santama, N. & Cavallaro, S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC genomics* **8**, 26, doi:10.1186/1471-2164-8-26 (2007).

20 Carter, B. J., Anklesaria, P., Choi, S. & Engelhardt, J. F. Redox modifier genes and pathways in amyotrophic lateral sclerosis. *Antioxidants & redox signaling* **11**, 1569-1586, doi:10.1089/ARS.2008.2414 (2009).

21 Billaud, M. *et al.* Regulation of cellular communication by signaling microdomains in the blood vessel wall. *Pharmacological reviews* **66**, 513-569, doi:10.1124/pr.112.007351 (2014).

22 Zhao, W., Beers, D. R. & Appel, S. H. Immune-mediated mechanisms in the pathoprogression of amyotrophic lateral sclerosis. *Journal of neuroimmune pharmacology : the official journal of the Society on NeuroImmune Pharmacology* **8**, 888-899, doi:10.1007/s11481-013-9489-x (2013).

23 Shao, H., He, Y., Li, K. C. & Zhou, X. A system mathematical model of a cell-cell communication network in amyotrophic lateral sclerosis. *Molecular bioSystems* **9**, 398-406, doi:10.1039/c2mb25370d (2013).

24 Bendotti, C. *et al.* Inter- and intracellular signaling in amyotrophic lateral sclerosis: role of p38 mitogen-activated protein kinase. *Neuro-degenerative diseases* **2**, 128-134, doi:10.1159/000089617 (2005).

25 La Cognata, V. *et al.* A customized high-resolution array-comparative genomic hybridization to explore copy number variations in Parkinson's disease. *Neurogenetics* **17**, 233-244, doi:10.1007/s10048-016-0494-0 (2016).

26 Taghibiglou, C., Lu, J., Mackenzie, I. R., Wang, Y. T. & Cashman, N. R. Sterol regulatory element binding protein-1 (SREBP1) activation in motor neurons in excitotoxicity and amyotrophic lateral sclerosis (ALS): Indip, a potential therapeutic peptide. *Biochemical and biophysical research communications* **413**, 159-163, doi:10.1016/j.bbrc.2011.08.011 (2011).

27 Porstmann, T. *et al.* SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth. *Cell metabolism* **8**, 224-236, doi:10.1016/j.cmet.2008.07.007 (2008).

28 Saxena, S. *et al.* Neuroprotection through excitability and mTOR required in ALS motoneurons to delay disease and extend survival. *Neuron* **80**, 80-96, doi:10.1016/j.neuron.2013.07.027 (2013).

29 Kaminsky, E. B. *et al.* An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genetics in medicine : official journal of the American College of Medical Genetics* **13**, 777-784, doi:10.1097/GIM.0b013e31822c79f9 (2011).

30 Collaboration, A. *et al.* Jet energy resolution in proton-proton collisions at [Formula: see text] recorded in 2010 with the ATLAS detector. *The European physical journal. C, Particles and fields* **73**, 2306, doi:10.1140/epjc/s10052-013-2306-0 (2013).

31 Nishioka, K. *et al.* Genetic variation of the mitochondrial complex I subunit NDUFV2 and Parkinson's disease. *Parkinsonism & related disorders* **16**, 686-687, doi:10.1016/j.parkreldis.2010.09.007 (2010).

32 Morello, G. *et al.* Copy Number Variations in Amyotrophic Lateral Sclerosis: Piecing the Mosaic Tiles Together through a Systems Biology Approach. *Molecular neurobiology*, doi:10.1007/s12035-017-0393-x (2017).

33 Renton, A. E., Chio, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nature neuroscience* **17**, 17-23, doi:10.1038/nn.3584 (2014).

34 Brouwers, N. *et al.* Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. *Molecular psychiatry* **17**, 223-233, doi:10.1038/mp.2011.24 (2012).

35 Bahia El Idrissi, N. *et al.* Complement activation at the motor end-plates in amyotrophic lateral sclerosis. *Journal of neuroinflammation* **13**, 72, doi:10.1186/s12974-016-0538-2 (2016).

36 Goldknopf, I. L. *et al.* Complement C3c and related protein biomarkers in amyotrophic lateral sclerosis and Parkinson's disease. *Biochemical and biophysical research communications* **342**, 1034-1039, doi:10.1016/j.bbrc.2006.02.051 (2006).

37 Leger, B. *et al.* Human skeletal muscle atrophy in amyotrophic lateral sclerosis reveals a reduction in Akt and

an increase in atrogin-1. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **20**, 583-585, doi:10.1096/fj.05-5249fje (2006).

38 Prolo, L. M., Vogel, H. & Reimer, R. J. The lysosomal sialic acid transporter sialin is required for normal CNS myelination. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 15355-15365, doi:10.1523/JNEUROSCI.3005-09.2009 (2009).

39 Scekic-Zahirovic, J. *et al.* Motor neuron intrinsic and extrinsic mechanisms contribute to the pathogenesis of FUS-associated amyotrophic lateral sclerosis. *Acta neuropathologica* **133**, 887-906, doi:10.1007/s00401-017-1687-9 (2017).

40 Pharoah, D. S. *et al.* Expression of the inflammatory chemokines CCL5, CCL3 and CXCL10 in juvenile idiopathic arthritis, and demonstration of CCL5 production by an atypical subset of CD8+ T cells. *Arthritis research & therapy* **8**, R50, doi:10.1186/ar1913 (2006).

41 Garbuzova-Davis, S. & Sanberg, P. R. Blood-CNS Barrier Impairment in ALS patients versus an animal model. *Frontiers in cellular neuroscience* **8**, 21, doi:10.3389/fncel.2014.00021 (2014).

42 Marchitti, S. A., Brocker, C., Stagos, D. & Vasiliou, V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert opinion on drug metabolism & toxicology* **4**, 697-720, doi:10.1517/17425255.4.6.697 (2008).

43 Chapuis, J. *et al.* Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. *Molecular psychiatry* **18**, 1225-1234, doi:10.1038/mp.2013.1 (2013).

44 Yao, Y. *et al.* A striatal-enriched intronic GPCR modulates huntingtin levels and toxicity. *eLife* **4**, doi:10.7554/eLife.05449 (2015).

45 Campbell, P. D. *et al.* Unique function of Kinesin Kif5A in localization of mitochondria in axons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **34**, 14717-14732, doi:10.1523/JNEUROSCI.2770-14.2014 (2014).

46 Hegde, A. N. & Upadhya, S. C. Role of ubiquitin-proteasome-mediated proteolysis in nervous system disease. *Biochimica et biophysica acta* **1809**, 128-140, doi:10.1016/j.bbagrm.2010.07.006 (2011).

47 Chhetri, S. K., Dayanandan, R., Bindman, D., Craufurd, D. & Majeed, T. Amyotrophic lateral sclerosis and Huntington's disease: neurodegenerative link or coincidence? *Amyotrophic lateral sclerosis & frontotemporal degeneration* **15**, 145-147, doi:10.3109/21678421.2013.817586 (2014).

48 Buck, E. *et al.* Comparison of Sirtuin 3 Levels in ALS and Huntington's Disease-Differential Effects in Human Tissue Samples vs. Transgenic Mouse Models. *Frontiers in molecular neuroscience* **10**, 156, doi:10.3389/fnmol.2017.00156 (2017).

49 Los, M., Maddika, S., Erb, B. & Schulze-Osthoff, K. Switching Akt: from survival signaling to deadly response. *BioEssays : news and reviews in molecular, cellular and developmental biology* **31**, 492-495, doi:10.1002/bies.200900005 (2009).

50 de Oliveira, A. C. *et al.* Pharmacological inhibition of Akt and downstream pathways modulates the expression of COX-2 and mPGES-1 in activated microglia. *Journal of neuroinflammation* **9**, 2, doi:10.1186/1742-2094-9-2 (2012).

51 Palmieri, M. *et al.* mTORC1-independent TFEB activation via Akt inhibition promotes cellular clearance in neurodegenerative storage diseases. *Nature communications* **8**, 14338, doi:10.1038/ncomms14338 (2017).

52 Satoh, J., Asahina, N., Kitano, S. & Kino, Y. A Comprehensive Profile of ChIP-Seq-Based Olig2 Target Genes in Motor Neuron Progenitor Cells Suggests the Possible Involvement of Olig2 in the Pathogenesis of Amyotrophic Lateral Sclerosis. *Journal of central nervous system disease* **7**, 1-14, doi:10.4137/JCNSD.S23210 (2015).

53 Nishimura, A. L. *et al.* A novel locus for late onset amyotrophic lateral sclerosis/motor neurone disease variant at 20q13. *Journal of medical genetics* **41**, 315-320 (2004).

54 Martinez-Muriana, A. *et al.* CSF1R blockade slows the progression of amyotrophic lateral sclerosis by reducing microgliosis and invasion of macrophages into peripheral nerves. *Scientific reports* **6**, 25663, doi:10.1038/srep25663 (2016).

55 de Oliveira, G. P., Alves, C. J. & Chadi, G. Early gene expression changes in spinal cord from SOD1(G93A) Amyotrophic Lateral Sclerosis animal model. *Frontiers in cellular neuroscience* **7**, 216, doi:10.3389/fncel.2013.00216 (2013).

56 Lee, J. K. *et al.* MST1 functions as a key modulator of neurodegeneration in a mouse model of ALS. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12066-12071, doi:10.1073/pnas.1300894110 (2013).

57 Ferraiuolo, L. *et al.* Transcriptional response of the neuromuscular system to exercise training and potential implications for ALS. *Journal of neurochemistry* **109**, 1714-1724, doi:10.1111/j.1471-4159.2009.06080.x (2009).

58 Chung, J. Y. *et al.* Elevated TRAF2/6 expression in Parkinson's disease is caused by the loss of Parkin E3 ligase activity. *Laboratory investigation; a journal of technical*

methods and pathology **93**, 663-676, doi:10.1038/labinvest.2013.60 (2013).

59 Alves, C. J., Maximino, J. R. & Chadi, G. Dysregulated expression of death, stress and mitochondrion related genes in the sciatic nerve of presymptomatic SOD1(G93A) mouse model of Amyotrophic Lateral Sclerosis. *Frontiers in cellular neuroscience* **9**, 332, doi:10.3389/fncel.2015.00332 (2015).

60 Kanekura, K., Suzuki, H., Aiso, S. & Matsuoka, M. ER stress and unfolded protein response in amyotrophic lateral sclerosis. *Molecular neurobiology* **39**, 81-89, doi:10.1007/s12035-009-8054-3 (2009).

61 Cai, Y. *et al.* Interplay of endoplasmic reticulum stress and autophagy in neurodegenerative disorders. *Autophagy* **12**, 225-244, doi:10.1080/15548627.2015.1121360 (2016).

62 Shiihashi, G. *et al.* Mislocated FUS is sufficient for gain-of-toxic-function amyotrophic lateral sclerosis phenotypes in mice. *Brain : a journal of neurology* **139**, 2380-2394, doi:10.1093/brain/aww161 (2016).

63 Booth, L. *et al.* GRP78/BiP/HSPA5/Dna K is a universal therapeutic target for human disease. *Journal of cellular physiology* **230**, 1661-1676, doi:10.1002/jcp.24919 (2015).

64 Casas, C. GRP78 at the Centre of the Stage in Cancer and Neuroprotection. *Frontiers in neuroscience* **11**, 177, doi:10.3389/fnins.2017.00177 (2017).

65 Wang, M., Wey, S., Zhang, Y., Ye, R. & Lee, A. S. Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders. *Antioxidants & redox signaling* **11**, 2307-2316, doi:10.1089/ARS.2009.2485 (2009).

66 Nishioka, M., Bundo, M., Iwamoto, K. & Kato, T. Somatic mutations in the human brain: implications for psychiatric research. *Molecular psychiatry*, doi:10.1038/s41380-018-0129-y (2018).

67 Fullard, J. F. *et al.* Assessment of somatic single-nucleotide variation in brain tissue of cases with schizophrenia. *Translational psychiatry* **9**, 21, doi:10.1038/s41398-018-0342-0 (2019).

68 Team, R. C. (ISBN 3-900051-07-0, 2014).

69 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).

70 Nikolsky, Y., Ekins, S., Nikolskaya, T. & Bugrim, A. A novel method for generation of signature networks as biomarkers from complex high throughput data. *Toxicol.Lett.* **158**, 20-29 (2005).

71 Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research* **43**, D447-452, doi:10.1093/nar/gku1003 (2015).

72 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

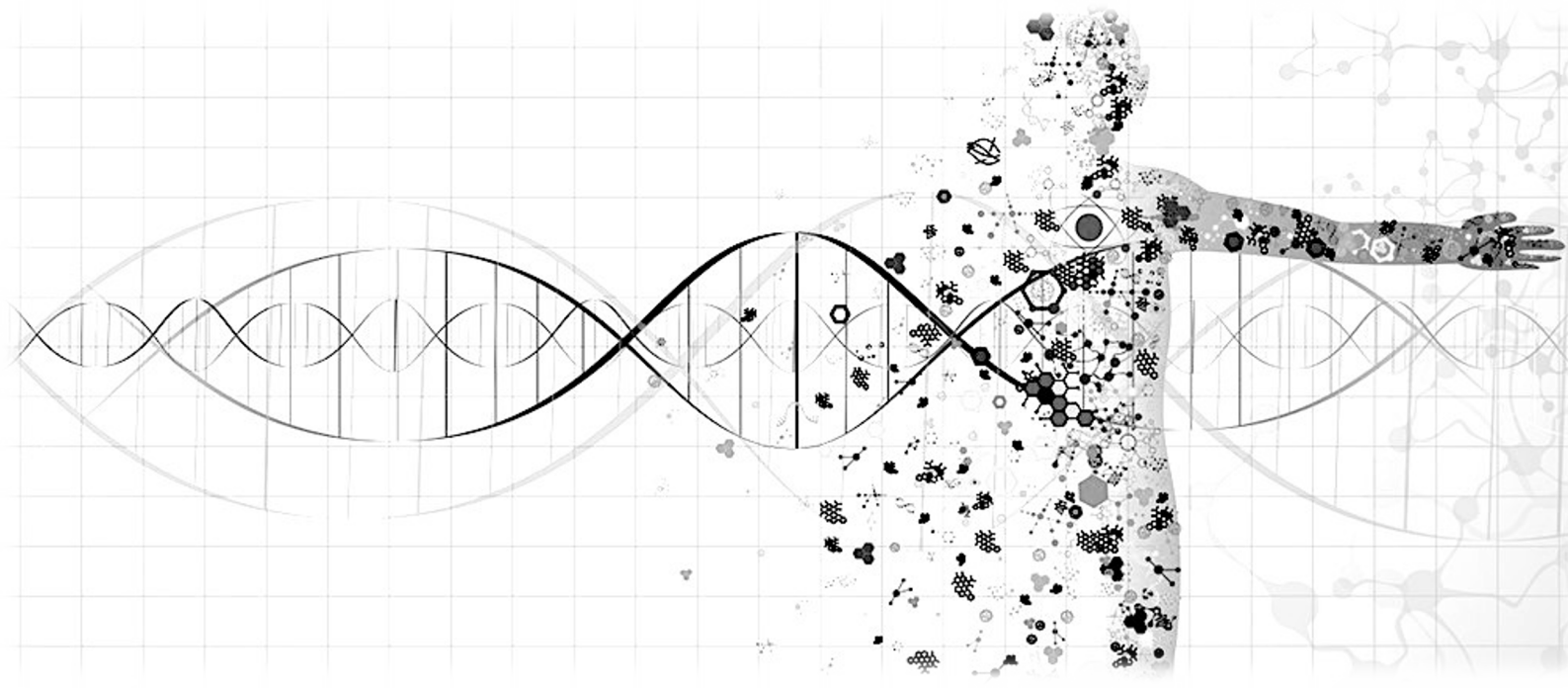
73 Assenov, Y., Ramirez, F., Schelhorn, S. E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **24**, 282-284, doi:10.1093/bioinformatics/btm554 (2008).

74 D'haene, B., Vandesompele, J. & Hellemans, J. Accurate and objective copy number profiling using real-time quantitative PCR. *Methods* **50**, 262-270, doi:http://dx.doi.org/10.1016/j.ymeth.2009.12.007 (2010).

75 Patitucci, A. *et al.* Comparison of different techniques for detecting 17p12 duplication in CMT1A. *Neuromuscular Disorders* **15**, 488-492, doi:http://dx.doi.org/10.1016/j.nmd.2005.04.006 (2005).

CHAPTER 3

From multi-omics approaches to precision medicine in Amyotrophic Lateral Sclerosis



Morello. G., et al. Front. Neurosci. (2020) in press

From multi-omics approaches to precision medicine in Amyotrophic Lateral Sclerosis

Giovanna Morello[a,b], Salvatore Salomone[b], Velia D'Agata[c], Francesca Luisa Conforti[d], and Sebastiano Cavallaro[a,#]

[a]*Institute for Research and Biomedical Innovation (IRIB), Italian National Research Council (CNR), Catania, Italy; Department of Biomedical and Biotechnological Sciences, Section of* [b]*Pharmacology and* [c]*Human Anatomy and Histology, University of Catania, Catania, Italy Catania, Catania, Italy;* [d]*Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende (Cosenza), Italy.*

[#] *Corresponding author*: Sebastiano Cavallaro, M.D., Ph.D., Institute for Research and Biomedical Innovation (IRIB), CNR, Via Paolo Gaifami, 18, 95125 Catania, Italy. Phone: +39.095.7338111, Fax: +39.095.7338110, Email: sebastiano.cavallaro@cnr.it

Running title: Multi-Omics signature in ALS

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disorder, caused by the degeneration of upper and lower motor neurons for which there is no truly effective cure. The lack of successful treatments can be well explained by the complex and heterogeneous nature of ALS, with patients displaying widely distinct clinical features and progression patterns, and distinct molecular mechanisms underlying the phenotypic heterogeneity. Thus, stratifying ALS patients into consistent and clinically-relevant subgroups can be of great value for the development of new precision diagnostics and targeted therapeutics for ALS patients. In the last years, the use and integration of high-throughput 'omics' approaches have dramatically changed our thinking about ALS, improving our understanding of the complex molecular architecture of ALS, distinguishing distinct patient subtypes and providing a rational foundation for the discovery of biomarkers and new individualized treatments. In this review, we discuss the most significant contributions of omics technologies in unraveling the biological heterogeneity of ALS, highlighting how these approaches are revealing diagnostic, prognostic and therapeutic targets for future personalized interventions.

Keywords: Amyotrophic lateral sclerosis, ALS-FTD, personalized medicine, molecular taxonomy, multi-omics, systems biology

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disease, characterized by the progressive deterioration of cortical and spinal motor neurons (MNs), leading invariably to progressive muscle weakness and paralysis. Death, often resulting from respiratory failure due to respiratory muscle weakness, generally occurs after 3–5 years from symptom onset, with only 5–10% of patients' survival beyond 10 years[1]. ALS is the most common adult motor neuron disease with a worldwide annual incidence of about 2 per 100,000 persons and with an estimated prevalence of 5.4 per 100,000 individuals[2]. In most cases, mean age at onset is 50-60 years, while juvenile (before 25 years of age) and "young-onset" ALS cases (before 45 years), represent ~1% and ~10% of all patients, respectively[3]. No disease-modifying strategies are available so far, and therapies that can effectively stop or reverse the disease progression are urgently needed. The mainstay of treatment for ALS is mainly based on symptom management and respiratory support, with only two Food and Drug Administration (FDA)-approved treatments, riluzole and edaravone, that appear to mildly slow disease progression and only in some patients[4–6]. The paucity of effective treatments has been attributed in part to the absence of complete knowledge of ALS pathogenesis, and in part to its heterogeneity with patients displaying widely distinct clinical features and progression patterns, together with a plurality of associated genes.

Over the last few years, the complexity of ALS has led to the concept of a spectrum of different disorders with different pathogenic mechanisms rather than a single disease. From a clinical point of view, in addition to typical or classic ALS (characterized by the simultaneous involvement of upper and lower motor neuron (UMN and LMN) at disease onset), several different phenotypic subtypes can be recognized based on the rate of progression, survival, age of onset, site of onset (bulbar vs spinal)

and prevalence of UMN or LMN motor signs[1]. Additionally, while ALS was historically judged as a pure motor neuron disease, it is now recognized that it represents a multi-systemic disorder affecting other brain regions, including frontotemporal, oculomotor, cerebellar, and/or sensory systems, and more rarely the basal ganglia and autonomic nervous system [7,8]. To this regard, the most common alternative deficit observed in ALS patients is behavioral dysfunction and/or subtle cognitive impairment, which is also comorbid to ALS in about half of ALS individuals, and where a subset of ~15% of patients receive the concomitant diagnosis of ALS with a frontotemporal dementia (FTD) syndrome (referred to as ALS-FTD or FTD-ALS patients)[9–12]. The ALS-FTD relationship has been confirmed through genetic studies, suggesting these conditions can be viewed as divergent ends of the spectrum of a single clinically and etiologically heterogeneous condition[9].

Different clinical profiles are likely to reflect molecular heterogeneity in ALS. In fact, for example, the majority (~90%) of ALS cases are sporadic (SALS), with unknown cause, while ~10% of ALS patients show familiarity for the disease, usually transmitted according to an autosomal dominant inheritance[13]. However, this distinction is increasingly recognized to be artificial; FALS and SALS are, in fact, phenotypically indistinguishable and seem to show similar patterns of selective MN degeneration and vulnerability, and many mutations in one or more known FALS-associated genes have been found in SALS patients, suggesting the existence of common molecular mechanisms between these two disease forms[14–16]. The complexity and heterogeneity of ALS also emerged from a pathophysiologic point of view, with a series of several biological and molecular pathways differently contributing to its development and progression. Despite the understanding of disease pathogenesis is far from exhaustive, numerous genetic and epidemiological risk factors have been identified, as

well as various mechanisms have been suggested, including inflammatory and immune abnormalities, oxidative stress, mitochondrial dysfunction, glutamate excitotoxicity, proteasomal/autophagic impairment, defects in axonal transport and RNA metabolism [14]. With this in mind, it is clear that the current diagnostic classification criteria of ALS, primarily based on person's signs and symptoms, are inadequate to characterize the complex and heterogeneous nature of ALS, as well as the use of a single compound to treat the patient population as a whole may hinder the identification of an effective therapy. Defining and stratification of ALS patients into disease subtypes can not only provide important insights for diagnosis and prognosis but also for clinical trial planning and interpretation, thus achieving better care for ALS patients.

Advances in "omics" technologies (e.g., genome, transcriptome, proteome, epigenome, metabolome) and their correlation with the clinical phenotypes of the individual patient, are enabling medicine to move from a "one-size-fits-all" approach toward a "personalized" model, helping to clarify the molecular mechanisms underlying human disease and to provide both potential biomarkers and pharmacological targets for a more detailed patient stratification and personalized treatments (**Figure 1**). In this review, we discuss advances in the application of "-omics" to further our understanding of ALS, outline the evolving landscape of molecular classifications, and discuss how these techniques are contributing to reveal diagnostic and prognostic biomarkers and molecular targets for future personalized therapeutic interventions.

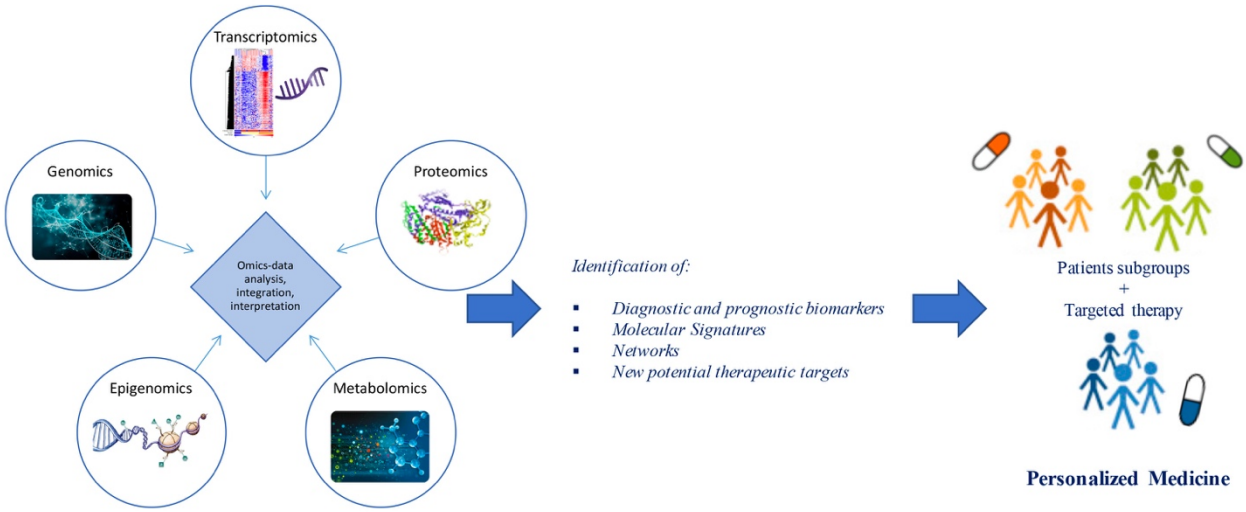


Figure 1. *From omics technology to precision medicine in ALS. Multi-omics (e.g., genomics, transcriptomics, proteomics, epigenomics, metabolomics) data analysis and integration may allow patient stratification and targeted therapies. Through a "systems biology" approach, these technologies may move medicine from a "one-size-fits-all" toward a "personalized" model.*

APPLICATION OF OMICS: A STEP TOWARDS A BETTER UNDERSTANDING OF ALS PATHOGENESIS

Applications of omics platforms range from the detection of genes (genomics), mRNA (transcriptomics), proteins (proteomics), epigenomic factors (epigenomics) and metabolites (metabolomics). Thanks to omics technologies, it is now possible to quantify the amount of particular molecules (genes, mRNA, protein levels, and metabolites) of a biological system, and observe massive interactomes describing their complex interconnections. For complex and multifactorial pathologies such as ALS, the analysis and integration of different omics layers are crucial for the full knowledge of the disease, opening the way to the development of personalized diagnostic and therapeutic tools. Several omics studies have suggested multiple pathologic mechanisms associated to ALS, providing new insights into molecular signatures/markers and moving toward molecular-based classifications and tailored interventions.

Genomics

The genomic landscape of ALS has been extensively surveyed, contributing to our understanding of ALS biological and clinical complexity. Analysis at this level requires not only the study of DNA sequence variations, including single nucleotide polymorphisms (SNPs) or mutations, but also genomic alterations and chromosomal changes, with consequent protein dysfunction or differences in concentration levels. Detailed information regarding ALS-related genes is available via the Amyotrophic Lateral Sclerosis Online Database (**ALSOD**; https://alsod.ac.uk). After the identification of mutations in the *SOD1* gene in 1993 (Rosen and others 1993), more than 30 genes have been involved in the pathology, with the most common disease-causing variants in *C9orf72, SOD1, FUS, and TARDBP*. However, monogenic forms explain only a fraction of the diagnosed cases, suggesting ALS as a polygenic disease[17,18].

Thanks to the development of genome-wide association studies (GWAS) as well as the advances in massive parallel sequencing approaches**,** including whole-genome sequencing (WGS) and whole-exome sequencing (WES), enormous progress has been made in understanding genomics of ALS[19–26]. A growing number of causative and susceptibility genes have been identified so far in both familial and sporadic cases, the majority of which encode proteins implicated in cytoskeleton remodeling and axonal transport, mitochondrial metabolism and turnover, autophagy and proteostasis, membrane trafficking, RNA processing and DNA repair (**Table 1**)[18,24,27–30]. These genetic findings may guide patient stratification into different subgroups depending on which combination of pathways is deregulated, improving their recruitment for translational research and clinical trials [31,32].

Table 1. Summary of the most known genes linked to ALS, their clinical phenotypes and affected pathway

Gene symbol	Gene name	Associated phenotype	Oxidative stress	Mitochondria	Cytoskeleton and axonal dynamics	Protein Trafficking and degradation	Autophagy	Vescicle trafficking	DNA repair	RNA processing	Innate Immunity and Neuroinflammation
SOD1	Superoxide Dismutase 1	ALS, PMA, juvenile ALS	X	X		X	X				
DAO	D-amino acid oxidase	ALS	X								
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	ALS	X	X							
OPTN	Optineurin	ALS, FTD		X			X	X			
CHCHD10	Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 10	ALS, ALS-FTD, FTD, cerebellar ataxia, myophathy	X	X		X					
NEK1	NIMA Related Kinase 1	ALS, ALS-FTD	X	X	X				X		
KIF5A	Kinesin family member 5A	ALS			X						
NEFH	Neurofilament heavy subunit	ALS			X						
TUBA4A	Tubulin Alpha 4a	ALS			X						
DCTN1	Dynactin subunit 1	ALS, ALS-FTD			X			X			
PFN1	Profilin 1	ALS			X	X					
ELP3	Elongator protein 3	ALS, ALS-FTD			X					X	
EPHA4	EPH Receptor A4	ALS			X						
C9orf72	Chromosome 9 open reading frame 72	ALS, ALS-FTD, FTD				X	X	X		X	
PRPH	Peripherin	ALS			X						
CHMP2B	Charged Multivesicular Body Protein 2B	ALS, FTD				X	X	X			
VCP	Valosin containing protein	ALS, ALS-FTD, FTD, IBM, PDB				X	X	X			
FIG4	Phosphoinositide 5-Phosphatase	ALS, PLS, CMT					X	X			
VAPB	Vesicle-associated membrane protein-associated protein B/C	ALS, PMA				X		X			
UBQLN2	Ubiquilin 2	ALS, ALS-FTD, juvenile ALS				X	X				
TBK1	TANK Binding Kinase 1	ALS, FTD				X	X				X
SQSTM1	Sequestosome 1	ALS, ALS-FTD, FTD, IBM, PDB				X	X				
CCNF	Cyclin F	ALS, ALS-FTD				X					
TARDBP	TAR DNA binding protein	ALS, ALS-FTD, FTD							X		
hnRNPA1	Heterogeneous Nuclear Ribonucleoprotein A1	ALS, ALS-FTD, FTD, IBM, PDB				X				X	

hnRNPA2B1	Heterogeneous Nuclear Ribonucleoprotein A2/B1	ALS, ALS-FTD, FTD, IBM, PDB				X				X	
ALS2	Alsin	Juvenile ALS, infantile HSP						X			
SPG11	Spatacsin vescicle trafficking associated	Juvenile ALS, HSP			X	X		X	X		
SIGMAR1	Sigma Non-Opioid Intracellular Receptor 1	Juvenile ALS, dHMN					X				
C21orf2	Cilia- and flagella-associated protein 410	ALS							X		
SETX	Senataxin	Juvenile ALS, AOA2, dHMN							X	X	
FUS	Fused in sarcoma	ALS, ALS-FTD, FTD							X	X	
ATXN2	Ataxin 2	ALS, SCA2						X		X	X
ANG	Angiogenin	ALS, ALS-FTD								X	
MATR3	Matrin 3	ALS, ALS-FTD, distal myopathy								X	
EWSR1	EWS RNA Binding Protein 1	ALS								X	
TAF15	TATA-Box Binding Protein Associated Factor 15	ALS								X	

The table lists genes thought to be causative or risk factors for ALS sorted on the basis of their functional similarity. ALS=Amyotrophic lateral sclerosis; FTD=Frontotemporal dementia; PMA=Progressive muscular atrophy; IBM= Inclusion-body myositis; PDB=Paget disease of bone; HSP=Hereditary Spastic Paraplegia; dHMN=Distal Hereditary Motor Neuropathy; AOA2=Ataxia with oculomotor apraxia type 2; SCA2=Spinocerebellar ataxia type 2

Another important factor increasing the complexity of phenotype-genotype correlations in ALS is the observation of a clinical pleiotropy for ALS genes. Although some mutations associate with very specific ALS clinical profiles (e.g., patients with the Ala4Val mutation in *SOD1* usually have an aggressive form of ALS, whereas those with the homozygous Asp91Ala mutation tend to have a very slowly progressive disease with a generally ascending upper motor neuron phenotype), the majority of disease-causing genes show a high degree of phenotypic heterogeneity, with mutations in the same gene giving rise to different clinical entities, supporting a genetic basis for the observed clinical heterogeneity in ALS. A striking example of pleiotropy is due to *C9orf72* hexanucleotide repeat expansion mutation, which is clearly linked to ALS and FTD but pathogenic expansions have been also observed in a small percentage of patients affected by Alzheimer's (<1%), Huntington's (1-5%), and Parkinson's diseases (1%), as well as atypical parkinsonian syndromes, such as progressive supranuclear palsy (1-8%), corticobasal degeneration (3%), and Lewy body dementia (2%)[37]. Another

interesting example regards a newly identified ALS gene, *KIF5A*. In fact, missense mutations in the N-terminal motor domain of this gene are known to cause hereditary spastic paraplegia and Charcot–Marie–Tooth disease type 2, while ALS-associated mutations are predominantly located at the C-terminal tail domain[38,39]. The possible existence of a common genetic background in neurodegeneration is also supported by the observation that mutations in *ATXN2, SPAST, FIG4, SETX, DCTN1, MATR3, CHCHD10, SQSTM1, VAPB, HNRNPA1, VCP, APOE* and *OPTN* have been reported both ALS and other multisystem disorders, including FTD, spinocerebellar ataxias, parkinsonism and schizophrenia. Among these, *APOE*, the most prevalent genetic risk factor of AD, has been also studied both as a risk factor for ALS and as a modifier of various phenotypic aspects, including age at onset, site of onset, and duration of the disease. As already found for AD**,** inheritance of *APOE* alleles is associated with differences in the clinical course of ALS (with a protective role of E2 allele and a deleterious role of E4 allele) suggesting a potential implication of *APOE* genotype as a biomarker to discriminate clinical efficacy in ALS clinical trials[40–43]. Another genetic determinant of ALS is the trinucleotide repeat expansion occurring in the *ATXN2* gene, with long-expanded repeats that are found to cause spinocerebellar ataxia 2 while intermediate-length polyQ expansion seems to increase the risk of developing ALS, significantly correlate to a spinal phenotype, and associate with shorter survival[44–48]. As for mutant *C9orf72* and other pathological repeats, *ATXN2*-mediated toxicity seems to involve the creation of small toxic homopolymeric proteins, called dipeptide repeats (DPRs), through a process known as repeat-associated non-ATG-initiated (RAN) translation, leading to an impairment of ribosomal biogenesis, nucleocytoplasmic transport, RNA metabolism and protein sequestration, that can cause neurodegeneration and behavioral deficits[49–51]. Disease-modifying therapies designed

or formulated to specifically target the *ATXN2* gene, including the use of antisense oligonucleotides, are currently being studied as a promising therapeutic approach for ALS[51–53].

Besides clinical diagnosis and identification of risk variants and disease modifiers, the genomic analysis may be helpful for explaining the considerable differences in prognostic profiles of ALS patients, thus providing valuable information for designing new therapeutic strategies[54–58]. In particular, mutations in *SOD1*, *EPHA4, KIFAP3* and *UNC13A* seem to affect the progression of ALS disease or the survival of ALS patients. Loss-of-function mutations in *EPHA4* results in significantly longer survival of ALS patients and pharmacological inhibition of EPHA4 signaling has demonstrated to improve functional performance and motor neuron survival in ALS animal models[59,60]. Other genetic variants associated with ALS survival include Asp91Ala, one of the most common mutations in *SOD1* that is associated with a long survival when the locus had homozygous genotype, while that of affected heterozygotes varies; and the rs12608932 located in intron 21 of the *UNC13A* gene that is associated with an increased risk and shorter survival of ALS patients [61-68].

In addition to genetic mutations, the screening of submicroscopic chromosomal changes, known as copy-number variations (CNVs), is potentially informative of genomic alterations related to disease phenotype through the modulation of the expression and function of genes. Several studies have investigated the involvement of these variants in ALS, demonstrating their involvement as risk factors, with multiple rare CNVs more important than common ones[68–74]. In particular, a large number of rare and novel ALS-specific CNV loci were identified in ALS patients, with the majority of these variants exerting a role in biochemical pathways relevant to ALS pathogenesis, including regulation of synaptic transmission and neuronal action

potential, immune response and inflammation, cell adhesion, ion transport, transcriptional regulation and mRNA processing[70–72]. One of the most interesting example is represented by the survival motor neuron (*SMN*) genes, whose copy number alterations seems to increase risk of developing SALS as well as other neurodegenerative disorders, including progressive muscular atrophy (PMA)[19,70,75,76]. However, other studies have not found any significant association between the deletion of either *SMN1* or *SMN2* in ALS, suggesting these conflicting results may be due, in part, to the existence of heterogeneous subgroups of ALS patients. The same ambiguous results are found for copy number changes affecting mitochondrial DNA (mtDNA), with some ALS patients characterized by an accumulation of deletions and other cases showing increased mtDNA copy numbers[70,77,78]. Other examples are heterozygous deletions of *EPHA3*, which seem to confer a protective role against the risk of developing ALS, and deletions in *NEFL* associated with a delayed disease onset and slowed disease progression[68,70].

Notwithstanding the increased knowledge of ALS from a genomic perspective, substantial dilemmas remain from a clinical perspective and large-scale NGS and GWAS projects are currently underway to fully unravel the underlying causes. Among these, of note is Project MinE, an international, large-scale research initiative devoted to discovering genetic causes of ALS by performing whole-genome sequencing of at least 15,000 ALS patients and 7,500 controls, resulting in an open-source genome database, in conjunction with the collection of skin samples to make patient induced pluripotent stem cell lines (iPSCs) [79,80]. Future follow-up studies will be necessary to shed light on the biological drivers of disease and evaluate the direct effect of newly discovered genes on disease diagnosis and management, also determining if they could form candidates for novel gene therapies.

Transcriptomic

Changes in gene expression are widespread in ALS, as revealed by a large body of work on gene expression profiling of RNA samples from peripheral cells or postmortem nervous tissue of ALS patients and animal models. These signature patterns of gene expression have started to provide a more detailed picture of molecular events implicated in ALS pathobiology[81,82,91–97,83–90].

The advent of systems biology and development of high-throughput technologies, including RNA sequencing and high-density microarray platforms, is enabling us not only to discover and define mechanisms of pathogenesis in ALS, but also to differentiate ALS from the "ALS mimic syndromes" and healthy controls and stratify ALS patient into subgroups, facilitating the discovery of biomarkers and new individualized treatments for patients[94,98–101]. In this regard, our research group, in the last years, has taken important steps towards the characterization of a biological and molecular heterogeneity of ALS based on transcriptional profiles. In particular, unsupervised hierarchical clustering of genome-wide transcriptomic profiles generated from post-mortem motor cortex samples from SALS patients has led to separate healthy controls and SALS patients and identify two distinct patient groups (SALS1 and SALS2) depending on the combinations of genes and pathways that were deregulated [102]. In particular, we observed that cell death, antigen processing and presentation and regulation of chemotaxis were the most representative subgroup-specific pathways in SALS1, while deregulated genes in SALS2 were associated with axonal guidance, oxidative and proteotoxic stress (**Figure 2**)[102–104].

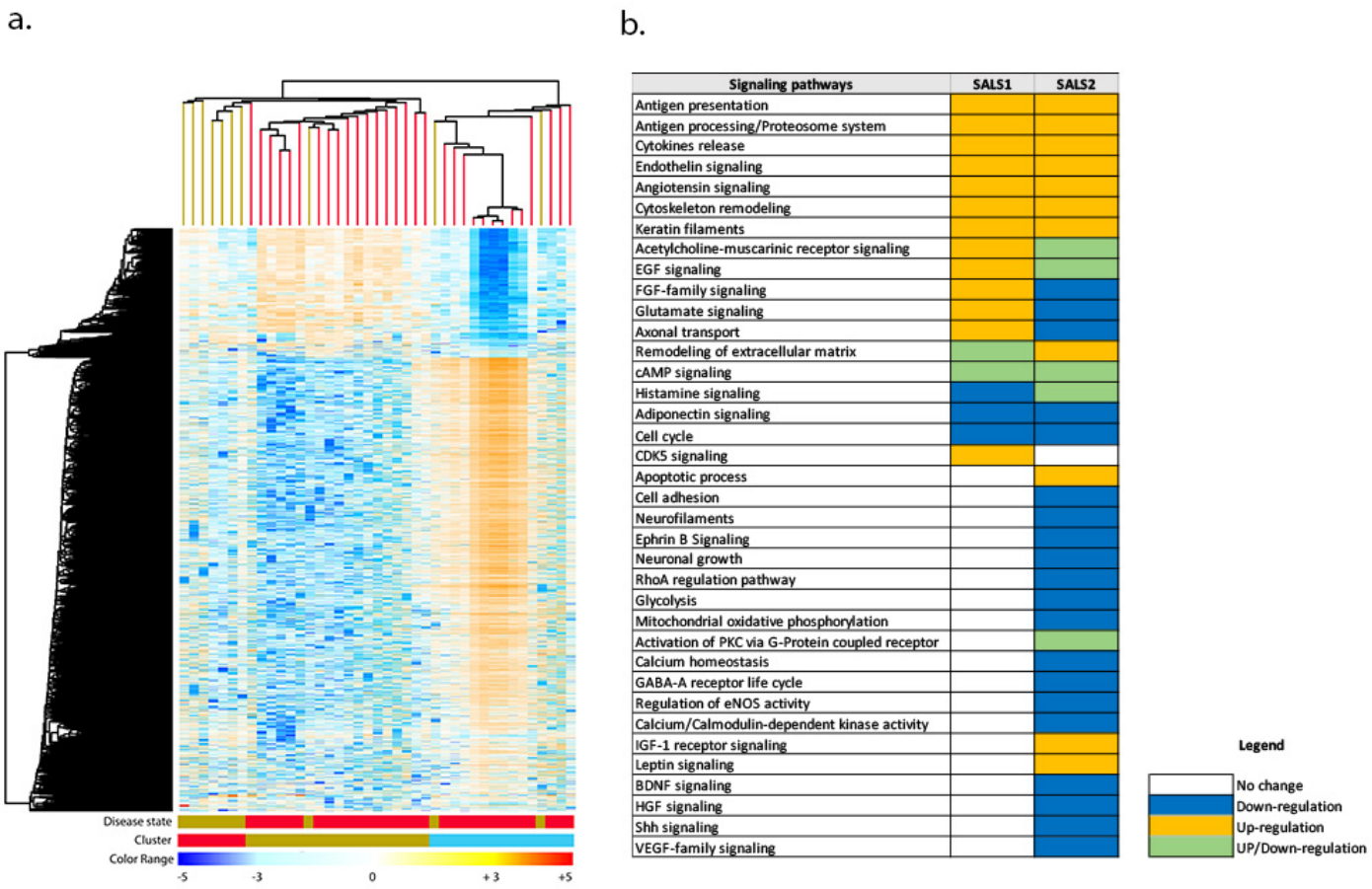


Figure 2. *Molecular-based classification of SALS. (a) Unsupervised hierarchical clustering (similarity measure: Pearson centred; linkage rule: average) was used to cluster control (10 fresh-frozen motor cortex samples from non-neurological patients) and SALS patients (31 fresh-frozen motor cortex samples) on the basis of the similarity in expression profiles of the most "hypervariable genes" (9.646 genes with a standard deviation > 1.5). The same clustering method clearly distinguished two SALS subgroups (SALS1 and SALS2), each associated to differentially expressed genes and pathways. In this two-dimensional presentation, each row represents a single gene and each column a motor cortex from control or SALS patients. As shown in the color bar, highly expressed genes are shown in orange, down-regulated genes in blue, no change in white. In the dendrograms shown (left and top), the length and the subdivision of the branches display the relatedness of the expression of the probes and the motor cortex (top). The disease state is marked as follows: controls patients are indicated by brown rectangles and SALS patients by red rectangles. In the cluster panel, red rectangle refers to control patients, brown to SALS1 and blue to SALS2 patients. For further details, the reader is referred to (Aronica et al., 2015b). (b) Functional pathways deregulated in clustered SALS patients. Orange boxes represent signalling pathways significantly up-regulated, blue bars down-regulated, green bars both up- and down-regulated, white bars indicate no significant change. Figure adapted from Morello et al. (Morello and Cavallaro, 2015).*

Our analysis also showed that some of the deregulated genes in SALS patients were previously associated with FALS, further supporting the existence of common pathological events between two disease forms. Interestingly, we found the differential expression of a substantial number of genes encoding splicing factors in the motor cortex and spinal cord of the same SALS cohort[105]. In particular, we observed transcriptional deregulation across the tissue types and/or disease states (SALS1, SALS2, controls),

with expression changes that were more pronounced for the motor cortex regions than the spinal cord and revealing a significant trend of overexpression for the SALS1 group and a decreased trend in expression for SALS2[105]. Despite, taken together, our results provided a powerful means for revealing etiopathogenetic mechanisms that were not emerged by considering SALS as a single pathology, it is clear that to successfully translate this knowledge to the real-world clinical contexts, the number of biomarkers should be limited. For this purpose, we next asked if the transcriptome-based classification can be reproduced by utilizing just a list of 203 genes highly associated with an increased ALS susceptibility[106,107]. Our results showed that this restricted gene panel was sufficiently representative to separate control from SALS patients, reproducing our previous classification of these patients into molecularly-defined and biologically meaningful subtypes and, consequently, facilitating the identification of promising cluster-specific biomarkers. Further studies will be necessary to investigate if peripheral tissues or easily accessible biological fluids (e.g., peripheral blood monocytes, cerebrospinal fluid, or muscle) can reproduce specific molecular patterns observed in brain regions of ALS patients, allowing for an effective mechanism-based selection of patients for clinical trials of molecular-targeted therapies. Emerging molecular heterogeneity of ALS lays the foundations for developing new therapeutic strategies, targeting disease pathogenesis as a complex system rather than at the level of the single protein molecule and that may have greater relevance to distinct sets of patients. In this regard, altered biological pathways emerged from our analysis provided a good number of potential subgroup-specific biomarkers and therapeutic targets, opening the way to the implementation of genomics-based personalized medicine[108–110]. Of note, some of these target genes exhibit expression profiles similar to those observed

in animal models of ALS, thus providing a rationale to ensure their preclinical trial success[111–113].

Recently, a good number of studies investigated and confirmed the existence of distinct molecular-based clusters of ALS patients, calling attention to the need for better understanding their mechanistic underpinnings and developing treatments based on specific forms of ALS[113–116]. In particular, *Tam et al. (2019)* were able to stratify the transcriptomes by RNA-seq of a largely sporadic set of ALS patients' motor cortex samples into three distinct molecular subgroups, two of which overlapped the molecular signatures observed in our ALS patient samples[116]. Another study compared brain transcriptome profiles in SALS cases carrying and not carrying the *C9orf72* repeat expansion, revealing both shared and distinct transcriptome changes and pathways associated with these two subsets of ALS cases[117]. A further interesting aspect is the possibility of separating rapid and slow ALS in earlier phases of drug development. To this regard, whole-genome expression analysis conducted by *Nardo G. et al. (2013)* in ALS animal models identified specific key genes and molecular pathways associated with fast or slow disease progression, highlighting their role as putative molecular targets for future therapeutic strategies[117].

The majority of the above-described studies assessed RNA samples from postmortem brain tissues. Although they provide essential elements in the pathophysiology of ALS that cannot be otherwise obtained through other approaches used in living patients, these studies reveal end-stage pathogenic mechanisms and do not clarify whether transcriptional differences that separate patient subtypes are a cause or a consequence of the disease process. In that context, the use of iPSC derived from patients suffering from ALS has provided important insights into disease pathophysiology, enabling researchers to explore molecular heterogeneity of ALS and

follow the course of degeneration in the dish[118,119,128–133,120–127]. In addition, transcriptome studies on whole tissue (i.e., motor cortex and spinal cord) fail to capture dynamic changes and the complex heterogeneity of the nervous system, making it difficult to determine how gene expression changes disrupt functional interaction between motor neurons and non-neuronal cells (e.g., microglia, oligodendroglia and astroglia) implicated in ALS pathology. Promising approaches, such as laser capture microdissection (LCM) coupled with RNA sequencing, offer a previously unavailable view of disease progression in ALS, enabling us to explore cell type-specific changes involved in the disease at a particular time point[134,135]. In a recent paper, *Maniatis et al. (2019)* used new RNA-seq based technologies, which they called "spatial transcriptomics", for mapping gene expression changes occurring at different disease stages and in different regions of murine models of ALS and human postmortem spinal cords samples, providing important clues for identifying disease-associated pathways and establishing the key steps in motor neuron degeneration observed in ALS[136].

Proteomics

Detection of specific protein changes in affected brain tissue samples, cell cultures or body fluids such as CSF represents an important pillar in ALS. The discovery of protein biomarkers for ALS, in fact, may aid earlier diagnosis, measure disease progression, exclude other ALS-mimicking syndromes, discriminate between subtypes of ALS that may theoretically respond to different therapeutic strategies and monitoring drug efficacy during clinical trials[137–142]. It is well established that the key neuropathological hallmark of the disease is the accumulation of misfolded cytoplasmic proteins in degenerating motor neurons and their non-neuronal neighbor[143–150]. One of the main protein components of these protein aggregates is TDP-43, a nuclear RNA binding

protein that under stress conditions or when mutated translocates to the cytoplasm where it is hyperphosphorylated and forms insoluble ubiquitin-positive aggregates[143,151]. Such aggregates are present in almost all cases of ALS, including SALS and FALS patients with pathogenic variants of *C9ORF72*[152,153], as well as in other neurodegenerative disorders, including FTD, Parkinson's and Alzheimer's disease[154,155]. It is interesting to note that ALS and FTD have different forms of TDP-43 pathology, suggesting its utility for designing novel diagnostic procedures that could discriminate against these two diseases[156]. In addition, although controversial, several results reported that TDP-43 aggregates occur in the vast majority of *SOD1*- and *FUS*-negative FALS patients, but not in *SOD1/FUS* mutation carriers, suggesting that mutant *TDP-43* may cause ALS through specific pathways of inclusion formation that are distinct from those that underlie SALS or other FALS-associated mutations, opening the way to the development of specific therapeutic approaches that take into account these selective modifications[157,158].

Due to the complex and heterogeneous nature of ALS, it is plausible that a single biomarker could not detect or differentiate between disease subgroups and/or control subjects, sustaining the importance of developing biomarker panels for specific and sensitive diagnostic tests. Recent development of high-throughput Mass Spectrometry-based proteomic (MS) technologies has allowed the simultaneous analysis of multiple proteins, allowing for the definition of comprehensive lists of possible candidate ALS biomarkers[159–162]. In this regard, due to its proximity to the central motor system, the cerebrospinal fluid (CSF) may most probably reflect disease-related alterations, including changes in protein expression, post-translational modification or biochemical turnover than in other body fluids (i.e., blood or urine)[162–166,166,167]. Analyses of the CSF proteome of ALS patients revealed a panel of candidate biomarkers

implicated in synaptic activity, extracellular matrix, inflammatory processes, glial response, axonal damage and apoptosis[161,162,170–173,163–166,166–169]. It is interesting to note that many candidate ALS protein biomarkers show subgroup-specific differential mRNA expression in SALS patients, suggesting their utility in patient stratification and personalized medicine (**Table 2).** Among the most extensively studied fluid biomarkers correlating with the survival of ALS patients, higher levels of neurofilament light chain (NF-L) and the phosphorylated form of neurofilament heavy chain (pNFH) in CSF and plasma samples, as well as their accumulation in brain tissue, have been correlated to shorter life expectancy and a more rapid disease progression and have demonstrated high sensitivity and specificity for separating ALS from ALS-mimic disorders [171,174–178]. Recent works also demonstrated the diagnostic utility of CSF pNFH levels in *C9ORF72*-ALS patients, revealing higher pNFH levels in ALS or ALS/FTD patients carrying *C9ORF72* expansion compared with controls and other ALS or ALS/FTD patients [179–181]. Several other proteins in CSF of ALS patients have demonstrated elevated sensitivity and specificity in distinguishing between ALS patients and neurological disease controls, including IL-10, IL-6, GM-CSF, IL-2, and IL-15 [182]. Proteomic profiling of CSF also identified proteins with a potential prognostic value in ALS, including MIP-1α, wrCRP, HMGB, creatine kinase, granzyme B, and IL-8, whose increased levels have been correlated with more rapidly progressive disease; cystatin C protein levels were positively correlated with survival; increase in GPNMB and UCHL1 were specific for ALS patients showing a short survival time; bFGF increased in ALS patients with longer survival, whereas VGF levels correlated with progressing muscle weakness [183,184].

Table 2. Putative protein biomarkers and their differential expression in distinct SALS patient subgroups.

Biomarker Symbol	Biomarker Name	CSF/Serum/ Plasma	Prognostic/ Diagnostic Value	References	Gene Expression in SALS motor cortex*	
					SALS1	SALS2
Neuron specific						
MAPT	microtubule-associated protein tau	CSF	Disease progression	(164)	↑	↓
NEFH	neurofilament, heavy polypeptide	CSF	Diagnosis and progression	(Rosengren et al., 2002; Brettschneider et al., 2006)	-	↓
NEFM	neurofilament, medium polypeptide	CSF	Diagnosis and progression	(Rosengren et al., 2002)	-	↓
NEFL	neurofilament, light polypeptide	CSF	Diagnosis and progression	(Rosengren et al., 2002; Zetterberg et al., 2007)	-	↓
Hormones and growth factors						
VEGFA	vascular endothelial growth factor A	CSF	Diagnosis and progression	(Moreau et al., 2006; Pasinetti et al., 2006b; Zhao et al., 2008)	-	↓
GDNF	glial cell-line derived neurotrophic factor	CSF	Diagnosis	(Tanaka et al., 2006)	↓	↑
IGFBP-2	insulin-like growth factor binding protein 2	Plasma, Serum	Diagnosis and progression	(Hosback et al., 2007)	-	↓
IGFBP-3	insulin-like growth factor binding protein 3	Plasma, Serum	Diagnosis and progression	(Hosback et al., 2007)	↑	↑
IGFBP-5	insulin-like growth factor binding protein 5	Plasma, Serum	Diagnosis and progression	(Hosback et al., 2007)	↑	↓
FGF-2	fibroblast growth factor 2	CSF, Serum	Diagnosis	(Johansson et al., 2003)	-	↓
HGF	hepatocyte growth factor	CSF	Diagnosis	(Tsuboi et al., 2002)	-	↓
Inflammatory system related						
IL2	interleukin 2	CSF	Diagnosis	(Mitchell et al., 2009)	-	↑
IL4	interleukin 4	CSF, Plasma	Diagnosis and progression	(Furukawa et al., 2015)	-	↑
IL5	interleukin 5 (colony-stimulating factor, eosinophil)	Plasma	Diagnosis	(Lu et al., 2016)	-	↑
IL6	interleukin 6 (interferon, beta 2)	CSF, Plasma	Diagnosis and progression	(Bilic et al., 2006; Mitchell et al., 2009)	-	↓
IL-10	interleukin 10	CSF, Plasma	Diagnosis and progression	(Mitchell et al., 2009; Furukawa et al., 2015; Andrés-Benito et al., 2017)	-	↓
IL-13	interleukin 13	Plasma	Diagnosis and progression	(Shi et al., 2007; Lu et al., 2016)	-	↑
IL-15	interleukin 15	CSF, Plasma	Diagnosis	(Mitchell et al., 2009)	-	↓
TNF	tumor necrosis factor-alpha	CSF, Plasma	Diagnosis	(Andrés-Benito et al., 2017)	↓	-
TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	Serum, Plasma	Diagnosis	(Andrés-Benito et al., 2017)	-	↓

IFNG	interferon, gamma	CSF, Plasma	Diagnosis and progression	(Guo et al., 2017a)	↓	↑
TGFB1	transforming growth factor beta 1	Plasma	Disease progression	(Duque et al., 2020)	-	↑
GFAP	glial fibrillary acidic protein	CSF	Diagnosis	(Benninger et al., 2016)	↑	-
CXCL10	chemokine (C-X-C motif) ligand 10	CSF	Diagnosis and progression	(Tateishi et al., 2010)	↓	-
Enzymes and enzyme inhibitors						
CST3	cystatin C	CSF	Diagnosis	(Ranganathan et al., 2005)	↑	-
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	CSF, Plasma	Diagnosis	(Niebroj-Dobosz et al., 2010)	-	↑
MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	CSF, Serum, Plasma	Diagnosis	(Beuche et al., 2000; Lorenzl et al., 2002)	-	↓
TIMP1	TIMP metallopeptidase inhibitor 1	CSF, Serum, Plasma	Diagnosis	(Lorenzl et al., 2002; Niebroj-Dobosz et al., 2010)	↑	↑
SOD1	superoxide dismutase 1, soluble	CSF, Plasma	Diagnosis	(Jacobsson et al., 2001)	-	↓
CHIT1	chitinase 1 (chitotriosidase)	CSF	Diagnosis and progression	(Thompson et al., 2018)	-	↑
Others						
TARDBP	TAR DNA binding protein	CSF	Diagnosis	(Majumder et al., 2018; Kasai et al., 2019)	-	↓
S100B	S100 calcium binding protein B	CSF	Disease progression	(Süssmuth et al., 2003)	-	↓

↑ = Concentration increased in SALS patients compared to controls, ↓ = Concentration decreased in SALS patients compared to controls.

*Aronica et al., 2015

As for genomics studies, systems biology-oriented approaches in proteomics play a crucial role to reveal relevant biological knowledge on pathological mechanisms that trigger the onset and progression of ALS, providing a mechanistic rationale for stratification of ALS patients based on unique molecular profiles, and identification of disease biomarkers and targets for drug efficacy measurements. In this scenario, the analysis of protein-protein interaction (PPI) networks provides the possibility to group proteins that are interacting with each other's in functional complexes and pathways, resulting critically important in helping us to comprehend complex processes, like ALS,

and identify key signaling cascades, upstream regulatory components, interactome domains, and novel disease-associated protein candidates suitable for therapeutic intervention[185–189]. In this regard, an interesting example is represented by a recent study investigating modules of co-expressed genes or proteins altered in postmortem cortex samples from patients affected by ALS, FTD, ALS/FTD, and healthy disease controls. In this work, *Umoh et al. (2018)* identified co-expression modules (i.e., RNA binding proteins, synaptic transmission, inflammation) differing across the ALS-FTD disease spectrum that may be useful for identifying genes associated with different clinical phenotypes along the ALS-FTD disease spectrum[155].

Other omics (metabolomics, epigenomics, miRNomics)

In addition to genomics, transcriptomics and proteomics, the exponential advances in technologies and informatics tools have stimulated an exponential growth of other areas of biomedical science (metabolomics, epigenomics, spliceomics), offering exciting new possibilities for ALS research. In this context, metabolomics, the scientific study of chemical processes involving metabolites (e.g., sugars, lipids, amino acids, organic acids), represents the downstream of systems biology that links the genome, transcriptome and proteome to patient phenotype, providing an important key tool for discovering potential markers in health or disease[190–194]. In the last years, thanks to the development of high-throughput technologies (such as Mass Spectrometry Combined with Liquid and Gas Chromatography), metabolomics studies identified specific metabolic markers and signatures that can discriminate ALS from controls and non-ALS cases, as well as identify distinct subgroups of SALS patients, moving research towards the development of novel targeted personalized treatments[195,196]. In particular, *Gross et al.* (2018) recently identified two subgroups of SALS case fibroblasts displaying

distinct metabotropic patterns that were also observed in plasma samples from the same patients, thus providing a basis for stratify SALS patients for appropriate targeted therapies[196]. Other metabolite profiling-based studies revealed significantly different metabolic profiles among FALS, SALS and ALS patients carrying different mutations in disease-causing genes (i.e., *C9ORF72, SOD1, TARDBP*, and *FUS*), suggesting the existence of distinct neurodegenerative processes associated with different subtypes of ALS[197–199]. It is interesting to note that changes in the metabolome as well as alterations in energy metabolism, such as an increase in resting energy expenditure, often precede the development of motor symptoms in ALS and correlates to disease progression. For instance, a lipid-specific metabolic abnormality is present at the pre-symptomatic stage of ALS animal models while increased serum levels of total cholesterol, LDL, LDL/HDL ratio, and triglycerides were associated with longer survival and slower disease progression in ALS patients and animal models [200–202]. However, the relationship between lipid levels and ALS is still rather controversial and poorly understood, and some follow-up observational studies of ALS did not observe any association between dyslipidemia and the incidence of ALS[203–205]. These conflicting results may be partly due to the relatively small sample sizes often employed in these observational studies and to the fact that lipid changes can be affected by a myriad of confounding factors, including genetic, nutritional, physical and pathological factors.

Analysis of metabolite profiles can be also used to identify metabolites and biochemical pathways in ALS patients that are modified before or after treatment exposure, giving rise to a new field called pharmacometabolomics[192,206,207]. To this regard, an interesting example is represented by a study that analyzed changes in metabolites and lipids composition in the plasma of ALS patients enrolled in a phase III clinical trial for assessing the effects of TRO19622 (olesoxime), a compound with

neuroprotective and neurodegenerative properties[208]. This study has permitted not only to identify distinct metabolic changes that can distinguish the placebo from the olexosime group but also to reveal metabolic pathways specifically altered after treatment with olesoxime and riluzole in combination in comparison to riluzole therapy alone, supporting the value of blood metabolomic profiles as biomarkers for evaluating the individual response to drug treatments and their side effects[208].

Another layer of complexity to the understanding of complex interactions between the genome and the environment is represented by epigenetic modifications, including DNA methylation, histone post-translational modifications, ATP-dependent chromatin remodeling and RNA-dependent gene silencing[209–216]. Several lines of evidence associate epigenome modifications to ALS development, with alterations in DNA methylation and DNA-(cytosine-5)- methyltransferase (DNMT) enzyme activity, as well as alterations to the balance between histone acetylation and deacetylation observed in blood and post-mortem neural tissue from patients with ALS and in different experimental models[217–219]. Of note, variations in epigenetic marks and modifier enzymes, and alterations in the methylation status of some ALS-related genes promoters were also determined, including hypomethylation of *OPTN*, hypermethylation of *C9orf72* expansion CpG islands in the blood of FTD/ALS patients, whereas mutant *SOD1, FUS* and *TDP43* contribute to global epigenome alteration by inducing alterations in histone post-translational modifications and DNA methylation [220]. While several high-density microarrays or sequencing-based epigenomic technologies are available, particular attention should be paid to EpiSwitch™, a high-resolution platform, recently developed by Oxford BioDynamics, for analyzing structural-functional epigenetic changes in genomic architecture associated with pathological phenotypes called “chromosome conformation signatures”. Using this

innovative technological platform, Salter and colleagues performed a comparative interrogation of the genomic architecture from healthy and ALS-patient blood samples revealing unique chromosomal conformation signatures with the ability to discern between diseased subjects and healthy controls, predict faster versus slower progressing patients at baseline and stratify responsive and non-responsive patients, representing a crucial step towards personalized medicine in ALS[221–223].

MicroRNAs (miRNAs), small non-coding molecules of about 20–22 nucleotides, represent an additional layer of epigenetic regulation that, thanks to their capability to be highly stable in human body fluids, are considered promising biomarkers for neurodegenerative diseases, including ALS[224,225]. Over the last few years, several whole-genome miRNA profiling studies have identified a panel of a dozen miRNAs that can distinguish ALS from controls with high accuracy in blood cells, serum and CSF, and may be altered in pre-symptomatic ALS mutation carriers even years before the estimated disease onset, representing potentially useful biomarkers of early-stage ALS in coming years[226–228]**.** Despite the heterogeneous nature of ALS may prevent a significant correlation of miRNA levels with clinical disease parameters, down-regulation of two miRNAs, miR-1234-3p and miR-1825, not only is specific for ALS, at least when compared with cohorts of Alzheimer's and Huntington's disease, but also significantly correlated with disease characteristics like age of onset, disease severity and duration[229,230]. In particular, while the downregulation of miR-1825 is a general early feature in both FALS and SALS, miR-1234-3p is significantly downregulated only in SALS patients. Of note, a large proportion of SALS patients showed miRNA signatures resembling those of FALS patients and mutation carriers, suggesting alteration of common pathways and a high contribution of genetic factors also in SALS[229,230]. Other examples include down-regulation of miR-206, a specific modulator

of skeletal muscle growth involved in nerve regeneration after injury, which accelerates disease progression in ALS mice, whereas up-regulation of miR-208B and miR-499 is found in the skeletal muscles of patients with slower disease progression, suggesting the potential utility of these microRNAs as promising candidate biomarkers and targets for this motor neuron disease [231–234].

FROM SINGLE LEVEL TO MULTI-OMICS INTEGRATIVE ANALYSES: TOWARD PRECISION MEDICINE IN ALS

As detailed in the previous paragraphs, omics technologies have been used to identify and / or provide functional supporting information for deciphering important players and pathways involved in ALS pathogenesis and identifying a panel of candidate therapeutic targets and biomarkers that will assist in the rapid diagnosis and prognosis assessment of the disease, and in the stratification of patients into different subgroups for specific targeted therapies. However, if considered individually, these technologies are insufficient to clarify the intricate disease mechanisms implicated in ALS. Taking a holistic molecular approach, based on the integration of multiple types of omics data with existing biological knowledge, has the potential role in improving the knowledge of the molecular basis underlying complex and heterogeneous diseases, establishing different molecular subtypes and patient stratification, thus providing a rational foundation for designing new studies to identify novel targets and clinical trials (**Figure 1**)[235–238]. Numerous studies have demonstrated the utility of whole- and multi-omics strategies for deciphering the molecular landscape of neurodegenerative diseases, including ALS, providing a feasible opportunity to develop an efficient and effective personalized diagnostics and patient-guided therapies[31,239–246].

An interesting example of applying integrated omics approaches to define an individual's molecular profile useful for the development and application of personalized medicine in ALS, is represented by recent studies carried out by our research groups. As previously described, transcriptional profiling of post-mortem motor cortex samples from SALS patients has allowed to differentiate two distinct patient subgroups characterized by different deregulated genes and pathways[102,104,107,109,110]. To investigate whether these transcriptional alterations may be related to genomic DNA alterations, and thus represent potential markers for a molecular-based stratification of SALS patients, we integrated gene expression profiling with the analysis of genomic structural aberrations occurring in the motor cortex of the same set of SALS samples[247]. This comprehensive molecular characterization at the genomic and transcriptomic levels revealed subtype-specific genomic alterations positively correlating with transcriptional signature profiles, further confirming the existence of molecular and functional heterogeneity in SALS and suggesting that genomic and transcriptomic events complement each other in driving disease pathogenesis[247]. Beyond refining ALS molecular architecture, our results also pinpointed candidate driver genes potentially useful as therapeutic targets and biomarkers for genomic-based patient stratification and individualized treatment[109–111,247]. Among these, numerous genes involved in histamine receptors, metabolism, transport, secretion and signal transduction, were differentially expressed in the motor cortex as well as in the spinal cord of two molecular-based subgroups of SALS patients and, of note, some of these genes are located within genomic regions disrupted by DNA copy number occurring in SALS patients [112]. By integrating our data with the known pathogenic variants of ALS-related gene reported in the ALSOD database, we identified a good number of coding variants in these genes, supporting the hypothesis that histamine-related genes might represent

candidate biomarkers and targets for patient-oriented ALS care[114]. In this regard, pharmacological modulation of the histamine-related pathway has already proved broad efficacy in ameliorating ALS features, improving motor performance and survival in ALS mice and increasing motor neurons survival *in vivo* and *in vitro* ALS models[112,113,248].

CONCLUSIONS

In the past decade, advanced omics technologies have fostered our understanding of the complex molecular architecture of ALS, contributing in part to explain its clinical heterogeneity, and providing a basis for a molecular taxonomy that may radically change our medical approach to ALS. The identification of relevant classifiers and subgroup-specific diagnostic, prognostic and predictive biomarkers is in fact urgently needed for accelerating the development of effective and personalized treatment approaches in ALS. In this review, we discuss the most significant contributions of omics approaches in unraveling the biological complexity of ALS, highlight how holistic systems biology approaches and multi-omics data integration are ideal to provide a comprehensive characterization of patient-specific molecular signatures that could potentially guide therapeutic decisions. We strongly believe that the future research in ALS, as well as in other neurodegenerative diseases, calls a multidisciplinary holistic approach, integrating multi-layer omics data with multimodal neuroimaging and clinical data. This approach will provide a clear understanding of disease prognosis and progression and accelerate development of innovative, effective and personalized strategies for ALS.

ACKNOWLEDGEMENTS

Authors gratefully acknowledge Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo, Francesco Marino for their administrative and technical support. This work was supported by the Italian Ministry of Education, Universities and Research through grant CTN01_00177_817708 and the international Ph.D. programs in *Neuroscience* of the University of Catania.

DECLARATIONS OF INTEREST

The authors declare no competing interests.

ABBREVIATIONS

ALS = amyotrophic lateral sclerosis; MNs= motor neurons; FDA= Food and Drug Administration; UMN= upper motor neuron; LMN= lower motor neuron; FTD= frontotemporal dementia; FALS= familial ALS; SALS= sporadic ALS; SNPs= single nucleotide polymorphisms; SOD1= Superoxide dismutase 1 [Cu-Zn]; C9orf72= chromosome 9 open reading frame 72; FUS= Fused in Sarcoma RNA binding protein; TDP-43/TARDBP= TAR DNA binding protein; GWAS= genome-wide association studies; WGS = whole-genome sequencing; WES= whole-exome sequencing; KIF5A= kinesin family member 5A; ATXN2=ataxin 2; SPAST= spastin; FIG4= FIG4 phosphoinositide 5-phosphatase; SETX= senataxin; DCTN1= dynactin subunit 1; MATR3= matrin 3; CHCHD10= coiled-coil-helix-coiled-coil-helix domain containing 10; SQSTM1= sequestosome 1; VAPB= VAMP associated protein B and C; HNRNPA1= heterogeneous nuclear ribonucleoprotein A1; VCP= valosin containing protein; OPTN= optineurin; EPHA4= Ephrin type-A receptor 4; KIFAP3= Kinesin Associated Protein 3; UNC13A= Unc-13 Homolog A; CNVs= copy-number variations; SMN=survival motor neuron; PMA= progressive muscular atrophy; mitochondrial DNA (mtDNA); EPHA3=

Ephrin type-A receptor 3; iPSC= induced pluripotent stem cells; LCM= laser capture microdissection; MS= Mass Spectrometry; CSF= cerebrospinal fluid; NF-L= neurofilament light chain; pNFH= phosphorylated neurofilament heavy chain; IL-10= interleukin 10; IL-6=interleukin 6; IL-2= interleukin 2; IL-15= interleukin 15; IL-8= interleukin 8; GM-CSF= Granulocyte-Macrophage Colony-Stimulating Factor; MIP-1α= Macrophage Inflammatory Proteins 1-alpha; wrCRP= wide-range C-reactive protein; HMGB= High Mobility Group Box 1; GPNMB= glycoprotein NMB; UCHL1= ubiquitin C-terminal hydrolase L1; bFGF= basic fibroblast growth factor; VGF= VGF Nerve Growth Factor Inducible; PPI= protein-protein interaction; LDL= low-density lipoprotein; HDL= high- density lipoprotein; DNMT= DNA-(cytosine-5)-methyltransferase; miRNA= MicroRNA.

REFERENCES

1. Brown, R. H. & Al-Chalabi, A. Amyotrophic lateral sclerosis. *New England Journal of Medicine* vol. 377 162–172 (2017).
2. Chiò, A. *et al.* Global epidemiology of amyotrophic lateral sclerosis: A systematic review of the published literature. *Neuroepidemiology* vol. 41 118–130 (2013).
3. Artemiadis, A. K., Peppas, C., Giannopoulos, S., Zouvelou, V. & Triantafyllou, N. Case of Young-Onset Sporadic Amyotrophic Lateral Sclerosis. *Journal of Clinical Neuromuscular Disease* vol. 17 220–222 (2016).
4. Bhandari, R., Kuhad, A. & Kuhad, A. Edaravone: A new hope for deadly amyotrophic lateral sclerosis. *Drugs of Today* **54**, 349–360 (2018).
5. Dash, R. P., Babu, R. J. & Srinivas, N. R. Two Decades-Long Journey from Riluzole to Edaravone: Revisiting the Clinical Pharmacokinetics of the Only Two Amyotrophic Lateral Sclerosis Therapeutics. *Clinical Pharmacokinetics* vol. 57 1385–1398 (2018).
6. Jaiswal, M. K. Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. *Medicinal Research Reviews* vol. 39 733–748 (2019).
7. Abrahams, S., Newton, J., Niven, E., Foley, J. & Bak, T. H. Screening for cognition and behaviour changes in ALS. *Amyotroph. Lateral Scler. Front. Degener.* **15**, 9–14 (2014).
8. Fang, T., Jozsa, F. & Al-Chalabi, A. Nonmotor Symptoms in Amyotrophic Lateral Sclerosis: A Systematic Review. in *International Review of Neurobiology* vol. 134 1409–1441 (Academic Press Inc., 2017).
9. Ferrari, R., Kapogiannis, D., Huey, E. D. & Momeni, P. FTD and ALS: a tale of two diseases. *Curr. Alzheimer Res.* **8**, 273–94 (2011).
10. Achi, E. Y. & Rudnicki, S. A. ALS and frontotemporal dysfunction: A review. *Neurology Research International* vol. 2012 (2012).
11. Zucchi, E., Ticozzi, N. & Mandrioli, J. Psychiatric Symptoms in Amyotrophic Lateral Sclerosis: Beyond a Motor Neuron Disorder. *Front. Neurosci.* **13**, 175 (2019).
12. Chiò, A. *et al.* Cognitive impairment across ALS clinical stages in a population-based cohort.

Neurology **93**, E984–E994 (2019).

13. Ryan, M. *et al.* Determining the incidence of familiality in ALS. *Neurol. Genet.* **4**, (2018).

14. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: From genes to mechanism. *Nature* vol. 539 197–206 (2016).

15. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nature Neuroscience* vol. 17 17–23 (2014).

16. Kirby, J., Al Sultan, A., Waller, R. & Heath, P. The genetics of amyotrophic lateral sclerosis: current insights. *Degener. Neurol. Neuromuscul. Dis.* **6**, 49 (2016).

17. McCann, E. P. *et al.* The genotype-phenotype landscape of familial amyotrophic lateral sclerosis in Australia. *Clin. Genet.* **92**, 259–266 (2017).

18. Mejzini, R. *et al.* ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Frontiers in Neuroscience* vol. 13 (2019).

19. Butchbach, M. E. R. Copy number variations in the survival motor neuron genes: Implications for spinal muscular atrophy and other neurodegenerative diseases. *Frontiers in Molecular Biosciences* vol. 3 (2016).

20. Cirulli, E. T. *et al.* Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science (80-.).* **347**, 1436–1441 (2015).

21. Naruse, H. *et al.* Burden of rare variants in causative genes for amyotrophic lateral sclerosis (ALS) accelerates age at onset of ALS. *J. Neurol. Neurosurg. Psychiatry* **90**, 537–542 (2019).

22. He, J., Mangelsdorf, M., Fan, D., Bartlett, P. & Brown, M. A. Amyotrophic Lateral Sclerosis Genetic Studies: From Genome-wide Association Mapping to Genome Sequencing. *Neuroscientist* vol. 21 599–615 (2015).

23. Little, J. *et al.* Genetic variation associated with the occurrence and progression of neurological disorders. *Neurotoxicology* **61**, 243–264 (2017).

24. Ramanan, V. K. & Saykin, A. J. Pathways to neurodegeneration: Mechanistic insights from GWAS in Alzheimer's disease, Parkinson's disease, and related disorders. *American Journal of Neurodegenerative Diseases* vol. 2 145–175 (2013).

25. Bean, D. M., Al-Chalabi, A., Dobson, R. J. B. & Iacoangeli, A. A Knowledge-Based Machine Learning Approach to Gene Prioritisation in Amyotrophic Lateral Sclerosis. *Genes (Basel).* **11**, (2020).

26. Van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).

27. Robberecht, W. & Eykens, C. The genetic basis of amyotrophic lateral sclerosis: recent breakthroughs. *Adv. Genomics Genet.* **5**, 327 (2015).

28. Cook, C. & Petrucelli, L. Genetic Convergence Brings Clarity to the Enigmatic Red Line in ALS. *Neuron* vol. 101 1057–1069 (2019).

29. Gall, L. Le *et al.* Molecular and Cellular Mechanisms Affected in ALS. *J. Pers. Med. 2020, Vol. 10, Page 101* **10**, 101 (2020).

30. Maurel, C. *et al.* Causative Genes in Amyotrophic Lateral Sclerosis and Protein Degradation Pathways: a Link to Neurodegeneration. *Molecular Neurobiology* vol. 55 6480–6499 (2018).

31. Vijayakumar, U. G. *et al.* A systematic review of suggested molecular strata, biomarkers and their tissue sources in ALS. *Front. Neurol.* **10**, (2019).

32. Volonté, C. *et al.* Omics-based exploration and functional validation of neurotrophic factors and histamine as therapeutic targets in ALS. *Ageing Research Reviews* vol. 62 (2020).

33. Foxe, D. *et al.* Intrafamilial phenotypic variability in the C9orf72 gene expansion: 2 case studies. *Front. Psychol.* **9**, (2018).

34. Al-Chalabi, A., Van Den Berg, L. H. & Veldink, J. Gene discovery in amyotrophic lateral sclerosis: Implications for clinical management. *Nature Reviews Neurology* vol. 13 96–104 (2017).

35. van Blitterswijk, M. *et al.* Genetic modifiers in carriers of repeat expansions in the C9ORF72 gene. *Mol. Neurodegener.* **9**, 38 (2014).

36. Balendra, R. & Isaacs, A. M. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nature Reviews Neurology* vol. 14 544–558 (2018).

37. Bourinaris, T. & Houlden, H. C9orf72 and its Relevance in Parkinsonism and Movement Disorders: A Comprehensive Review of the Literature. *Mov. Disord. Clin. Pract.* **5**, 575–585 (2018).

38. Brenner, D. *et al.* Hot-spot KIF5A mutations cause familial ALS. *Brain* **141**, 688–697 (2018).

39. Nicolas, A. *et al.* Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* **97**, 1268-1283.e6 (2018).

40. Moulard, B., Sefiani, A., Laamri, A., Malafosse, A. & Camu, W. Apolipoprotein E genotyping in sporadic amyotrophic lateral sclerosis, evidence for a major influence on the clinical presentation and prognosis. in *Journal of the Neurological Sciences* vol. 139 34–37 (Elsevier B.V., 1996).

41. Lacomblez, L. *et al.* APOE: A potential marker of disease progression in ALS. *Neurology* **58**, 1112–1114 (2002).

42. Li, Y. J. *et al.* Apolipoprotein E is associated with age at onset of amyotrophic lateral sclerosis. *Neurogenetics* **5**, 209–213 (2004).

43. Lacomblez, L. *et al.* APOE: A potential marker of disease progression in ALS. *Neurology* **58**, 1112–1114 (2002).

44. Sproviero, W. *et al.* ATXN2 trinucleotide repeat length correlates with risk of ALS. *Neurobiol. Aging* **51**, 178.e1-178.e9 (2017).

45. Chiò, A. *et al.* ATXN2 polyQ intermediate repeats are a modifier of ALS survival. *Neurology* **84**, 251–258 (2015).

46. Laffita-Mesa, J. M. *et al.* De Novo Mutations in Ataxin-2 Gene and ALS Risk. *PLoS One* **8**, e70560 (2013).

47. Borghero, G. *et al.* ATXN2 is a modifier of phenotype in ALS patients of Sardinian ancestry. *Neurobiol. Aging* **36**, 2906.e1-2906.e5 (2015).

48. van Blitterswijk, M. *et al.* Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. *Neurobiol. Aging* **35**, 2421.e13-2421.e17 (2014).

49. Barker, H. V., Niblock, M., Lee, Y. B., Shaw, C. E. & Gallo, J. M. RNA misprocessing in C9orf72-linked neurodegeneration. *Frontiers in Cellular Neuroscience* vol. 11 195 (2017).

50. Hutten, S. RAN translation down. doi:10.1038/s41593-019-0482-4.

51. Hergesheimer, R. *et al.* Advances in disease-modifying pharmacotherapies for the treatment of amyotrophic lateral sclerosis. *Expert Opinion on Pharmacotherapy* vol. 21 1103–1110 (2020).

52. Van Den Heuvel, D. M. A., Harschnitz, O., Van Den Berg, L. H. & Pasterkamp, R. J. Taking a risk: a therapeutic focus on ataxin-2 in amyotrophic lateral sclerosis? *Trends Mol. Med.* **20**, 25–35 (2014).

53. Scoles, D. R. & Pulst, S. M. Oligonucleotide therapeutics in neurodegenerative diseases. *RNA Biology* vol. 15 707–714 (2018).

54. Chiò, A., Mazzini, L. & Mora, G. Disease-modifying therapies in amyotrophic lateral sclerosis. *Neuropharmacology* vol. 167 107986 (2020).

55. Cappella, M., Ciotti, C., Cohen-Tannoudji, M. & Biferi, M. G. Gene therapy for ALS-A perspective. *International Journal of Molecular Sciences* vol. 20 (2019).

56. Tanaka, F. *et al.* Personal genome analysis in amyotrophic lateral sclerosis. *Brain and Nerve* **65**, 257–265 (2013).

57. Su, X. W., Broach, J. R., Connor, J. R., Gerhard, G. S. & Simmons, Z. Genetic heterogeneity of amyotrophic lateral sclerosis: Implications for clinical practice and research. *Muscle and Nerve* **49**, 786–803 (2014).

58. Geyer, F. C., Decker, T. & Reis-Filho, J. S. Genomweite Expressionsprofile als klinische Entscheidungshilfe : Bereit für die Praxis? *Pathologe* **30**, 141–146 (2009).

59. Van Hoecke, A. *et al.* EPHA4 is a disease modifier of amyotrophic lateral sclerosis in animal models and in humans. *Nat. Med.* **18**, 1418–1422 (2012).

60. Rué, L. *et al.* Reducing EphA4 before disease onset does not affect survival in a mouse model of

Amyotrophic Lateral Sclerosis. *Sci. Rep.* **9**, 1–9 (2019).

61. Cady, J. *et al.* Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. *Ann. Neurol.* **77**, 100–113 (2015).

62. Harms, M. B. & Baloh, R. H. Clinical Neurogenetics: Amyotrophic Lateral Sclerosis. *Neurologic Clinics* vol. 31 929–950 (2013).

63. Landers, J. E. *et al.* Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 9004–9009 (2009).

64. Gaastra, B. *et al.* Rare genetic variation in UNC13A may modify survival in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Front. Degener.* **17**, 593–599 (2016).

65. Daoud, H. *et al.* Analysis of the UNC13A gene as a risk factor for sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **67**, 516–517 (2010).

66. Yang, B. *et al.* UNC13A variant rs12608932 is associated with increased risk of amyotrophic lateral sclerosis and reduced patient survival: a meta-analysis. *Neurol. Sci.* **40**, 2293–2302 (2019).

67. Diekstra, F. P. *et al.* UNC13A is a modifier of survival in amyotrophic lateral sclerosis. *Neurobiol. Aging* **33**, 630.e3-630.e8 (2012).

68. Uyan, Ö. *et al.* Genome-Wide Copy Number Variation in Sporadic Amyotrophic Lateral Sclerosis in the Turkish Population: Deletion of EPHA3 Is a Possible Protective Factor. *PLoS One* **8**, e72381 (2013).

69. Butchbach, M. E. R. Copy number variations in the survival motor neuron genes: Implications for spinal muscular atrophy and other neurodegenerative diseases. *Frontiers in Molecular Biosciences* vol. 3 7 (2016).

70. Morello, G. *et al.* Copy Number Variations in Amyotrophic Lateral Sclerosis: Piecing the Mosaic Tiles Together through a Systems Biology Approach. *Molecular Neurobiology* vol. 55 1299–1322 (2018).

71. Blauw, H. M. *et al.* A large genome scan for rare CNVs in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **19**, 4091–4099 (2010).

72. Wain, L. V. *et al.* The role of copy number variation in susceptibility to amyotrophic lateral sclerosis: Genome-wide association study and comparison with published loci. *PLoS One* **4**, (2009).

73. Blauw, H. M. *et al.* Copy-number variation in sporadic amyotrophic lateral sclerosis: a genome-wide screen. *Lancet Neurol.* **7**, 319–326 (2008).

74. Vadgama, N. *et al.* De novo single-nucleotide and copy number variation in discordant monozygotic twins reveals disease-related genes. *Eur. J. Hum. Genet.* **27**, 1121–1133 (2019).

75. Sangare, M. *et al.* Does the survival motor neuron copy number variation play a role in the onset and severity of sporadic amyotrophic lateral sclerosis in Malians? *eNeurologicalSci* **3**, 17–20 (2016).

76. Blauw, H. M. *et al.* SMN1 gene duplications are associated with sporadic ALS. *Neurology* **78**, 776–780 (2012).

77. Keeney, P. M. & Bennett, J. P. ALS spinal neurons show varied and reduced mtDNA gene copy numbers and increased mtDNA gene deletions. *Mol. Neurodegener.* **5**, (2010).

78. Mawrin, C. *et al.* Single-cell analysis of mtDNA deletion levels in sporadic amyotrophic lateral sclerosis. *Neuroreport* **15**, 939–943 (2004).

79. van der Spek, R. A. A. *et al.* The project MinE databrowser: bringing large-scale whole-genome sequencing in ALS to researchers and the public. *Amyotroph. Lateral Scler. Front. Degener.* **20**, 432–440 (2019).

80. Van Rheenen, W. *et al.* Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. *Eur. J. Hum. Genet.* **26**, 1537–1546 (2018).

81. Kirby, J. *et al.* Mutant SOD1 alters the motor neuronal transcriptome: implications for familial ALS. *Brain* **128**, 1686–1706 (2005).

82. de Oliveira, G. P., Alves, C. J. & Chadi, G. Early gene expression changes in spinal cord from

SOD1G93A amyotrophic lateral sclerosis animal model. *Front. Cell. Neurosci.* **6**, (2013).

83. Malaspina, A., Kaushik, N. & De Belleroche, J. Differential expression of 14 genes in amyotrophic lateral sclerosis spinal cord detected using gridded cDNA arrays. *J. Neurochem.* **77**, 132–145 (2008).

84. Jiang, Y. M. *et al.* Gene expression profile of spinal motor neurons in sporadic amyotrophic lateral sclerosis. *Ann. Neurol.* **57**, 236–251 (2005).

85. Dangond, F. *et al.* Molecular signature of late-stage human ALS revealed by expression profiling of postmortem spinal cord gray matter. *Physiol. Genomics* **16**, 229–239 (2004).

86. Malaspina, A. & de Belleroche, J. Spinal cord molecular profiling provides a better understanding of amyotrophic lateral sclerosis pathogenesis. *Brain Res. Brain Res. Rev.* **45**, 213–29 (2004).

87. Raman, R. *et al.* Gene expression signatures in motor neurone disease fibroblasts reveal dysregulation of metabolism, hypoxia-response and RNA processing functions. *Neuropathol. Appl. Neurobiol.* **41**, 201–226 (2015).

88. Wang, X. S., Simmons, Z., Liu, W., Boyer, P. & Connor, J. Differential expression of genes in amyotrophic lateral sclerosis revealed by profiling the post mortem cortex. *Amyotroph. Lateral Scler.* **7**, 201–216 (2006).

89. Rahman, M. R., Islam, T., Huq, F., Quinn, J. M. W. & Moni, M. A. Identification of molecular signatures and pathways common to blood cells and brain tissue of amyotrophic lateral sclerosis patients. *Informatics Med. Unlocked* **16**, 100193 (2019).

90. Recabarren-Leiva, D. & Alarcón, M. New insights into the gene expression associated to amyotrophic lateral sclerosis. *Life Sciences* vol. 193 110–123 (2018).

91. Dickson, D. W. *et al.* Extensive transcriptomic study emphasizes importance of vesicular transport in C9orf72 expansion carriers. *Acta Neuropathol. Commun.* **7**, 150 (2019).

92. Maria D'erchia, A. *et al.* Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS OPEN. doi:10.1038/s41598-017-10488-7.

93. Recabarren-Leiva, D. & Alarcón, M. New insights into the gene expression associated to amyotrophic lateral sclerosis. *Life Sci.* **193**, 110–123 (2018).

94. Krokidis, M. G. & Vlamos, P. Transcriptomics in amyotrophic lateral sclerosis. *Front. Biosci. - Elit.* **10**, 103–121 (2018).

95. Lederer, C. W., Torrisi, A., Pantelidou, M., Santama, N. & Cavallaro, S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics* **8**, (2007).

96. Pasinelli, P. & Brown, R. H. Molecular biology of amyotrophic lateral sclerosis: Insights from genetics. *Nature Reviews Neuroscience* vol. 7 710–723 (2006).

97. Saris, C. G. J., Groen, E. J. N., Koekkoek, J. A. F., Veldink, J. H. & Van Den Berg, L. H. Meta-analysis of gene expression profiling in amyotrophic lateral sclerosis: A comparison between transgenic mouse models and human patients. *Amyotroph. Lateral Scler. Front. Degener.* **14**, 177–189 (2013).

98. Cooper-Knock, J. *et al.* Gene expression profiling in human neurodegenerative disease. *Nature Reviews Neurology* vol. 8 518–530 (2012).

99. Heath, P. R., Kirby, J. & Shaw, P. J. Investigating cell death mechanisms in amyotrophic lateral sclerosis using transcriptomics. *Frontiers in Cellular Neuroscience* vol. 7 (2013).

100. Recabarren-Leiva, D. & Alarcón, M. New insights into the gene expression associated to amyotrophic lateral sclerosis. *Life Sciences* vol. 193 110–123 (2018).

101. Krokidis, M. G. Transcriptomics and Metabolomics in Amyotrophic Lateral Sclerosis. in *Advances in Experimental Medicine and Biology* vol. 1195 205–212 (Springer, 2020).

102. Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol. Dis.* **74**, (2015).

103. Morello, G., Spampinato, A. G. & Cavallaro, S.

Molecular taxonomy of sporadic amyotrophic lateral sclerosis using disease-associated genes. *Front. Neurol.* **8**, (2017).

104. Morello, G., Spampinato, A. G. & Cavallaro, S. Neuroinflammation and ALS: Transcriptomic Insights into Molecular Disease Mechanisms and Therapeutic Targets. *Mediators Inflamm.* **2017**, (2017).

105. La Cognata, V., Gentile, G., Aronica, E. & Cavallaro, S. Splicing Players Are Differently Expressed in Sporadic Amyotrophic Lateral Sclerosis Molecular Clusters and Brain Regions. *Cells* **9**, 159 (2020).

106. Morello, G., Spampinato, A. G., Conforti, F. L. & Cavallaro, S. Taxonomy Meets Neurology, the Case of Amyotrophic Lateral Sclerosis. *Front. Neurosci.* **12**, (2018).

107. Morello, G., Spampinato, A. G. & Cavallaro, S. Molecular Taxonomy of Sporadic Amyotrophic Lateral Sclerosis Using Disease-Associated Genes. *Front. Neurol.* **8**, 152 (2017).

108. Morello, G., Conforti, F. L., Parenti, R., D'Agata, V. & Cavallaro, S. Selection of potential pharmacological targets in ALS based on whole-genome expression profiling. *Curr. Med. Chem.* **22**, (2015).

109. Morello, G., Spampinato, A. G., Conforti, F. L., D'Agata, V. & Cavallaro, S. Selection and Prioritization of Candidate Drug Targets for Amyotrophic Lateral Sclerosis Through a Meta-Analysis Approach. *J. Mol. Neurosci.* **61**, (2017).

110. Morello, G. & Cavallaro, S. Transcriptional analysis reveals distinct subtypes in amyotrophic lateral sclerosis: Implications for personalized therapy. *Future Med. Chem.* **7**, (2015).

111. Maugeri, G. *et al.* Molecular mechanisms involved in the protective effect of pituitary adenylate cyclase-activating polypeptide in an in vitro model of amyotrophic lateral sclerosis. *J. Cell. Physiol.* **234**, 5203–5214 (2019).

112. Apolloni S, Amadio S, Fabbrizio P, Morello G, Spampinato AG, Latagliata EC, Salvatori I, Proietti D, Ferri A, Madaro L, Puglisi-Allegra S, Cavallaro S, V. C. Histaminergic transmission slows progression of amyotrophic lateral sclerosis. *J. Chachexia, Sarcopenia Muscle* (2019) doi:10.1002/jcsm.12422.

113. Apolloni, S. *et al.* Histamine regulates the inflammatory profile of SOD1-G93A microglia and the histaminergic system is dysregulated in amyotrophic lateral sclerosis. *Front. Immunol.* **8**, (2017).

114. Apolloni, S. *et al.* Histaminergic transmission slows progression of amyotrophic lateral sclerosis. *J. Cachexia. Sarcopenia Muscle* jcsm.12422 (2019) doi:10.1002/jcsm.12422.

115. Jones, A. R. *et al.* Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes. *Neurobiol. Aging* **36**, 2006.e1-2006.e9 (2015).

116. Tam, O. H. *et al.* Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *CellReports* **29**, 1164-1177.e5 (2019).

117. Prudencio, M. *et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175–1182 (2015).

118. Lee, J. H., Liu, J. W., Lin, S. Z., Harn, H. J. & Chiou, T. W. Advances in Patient-Specific Induced Pluripotent Stem Cells Shed Light on Drug Discovery for Amyotrophic Lateral Sclerosis. *Cell Transplantation* vol. 27 1301–1312 (2018).

119. Fujimori, K. *et al.* Modeling sporadic ALS in iPSC-derived motor neurons identifies a potential therapeutic agent. *Nat. Med.* **24**, 1579–1589 (2018).

120. Hawrot, J., Imhof, S. & Wainger, B. J. Modeling cell-autonomous motor neuron phenotypes in ALS using iPSCs. *Neurobiology of Disease* vol. 134 104680 (2020).

121. Coatti, G. C. *et al.* Stem cells for amyotrophic lateral sclerosis modeling and therapy: Myth or fact? *Cytometry Part A* vol. 87 197–211 (2015).

122. Bohl, D., Pochet, R., Mitrecic, D. & Nicaise, C. Modelling and Treating Amyotrophic Lateral Sclerosis through Induced- Pluripotent Stem Cells Technology. *Curr. Stem Cell Res. Ther.* **11**, 301–312 (2016).

123. Myszczynska, M. & Ferraiuolo, L. New in vitro

models to study amyotrophic lateral sclerosis. in *Brain Pathology* vol. 26 258–265 (Blackwell Publishing Ltd, 2016).

124. Hedges, E. C., Mehler, V. J. & Nishimura, A. L. The Use of Stem Cells to Model Amyotrophic Lateral Sclerosis and Frontotemporal Dementia: From Basic Research to Regenerative Medicine. *Stem Cells International* vol. 2016 (2016).

125. Sances, S. *et al.* Modeling ALS with motor neurons derived from human induced pluripotent stem cells. *Nature Neuroscience* vol. 19 542–553 (2016).

126. Selvaraj, B. T., Livesey, M. R. & Chandran, S. Modeling the C9ORF72 repeat expansion mutation using human induced pluripotent stem cells. *Brain Pathol.* **27**, 518–524 (2017).

127. Guo, W., Fumagalli, L., Prior, R. & van den Bosch, L. Current advances and limitations in modeling ALS/FTD in a dish using induced pluripotent stem cells. *Frontiers in Neuroscience* vol. 11 (2017).

128. Csobonyeiova, M., Polak, S., Nicodemou, A. & Danisovic, L. Induced pluripotent stem cells in modeling and cell-based therapy of amyotrophic lateral sclerosis. *Journal of Physiology and Pharmacology* vol. 68 649–657 (2017).

129. Ghaffari, L. T., Starr, A., Nelson, A. T. & Sattler, R. Representing diversity in the dish: Using patient-derived in vitro models to recreate the heterogeneity of neurological disease. *Frontiers in Neuroscience* vol. 12 (2018).

130. Centeno, E. G. Z., Cimarosti, H. & Bithell, A. 2D versus 3D human induced pluripotent stem cell-derived cultures for neurodegenerative disease modelling. *Molecular Neurodegeneration* vol. 13 (2018).

131. Ziff, O. J. & Patani, R. Harnessing cellular aging in human stem cell models of amyotrophic lateral sclerosis. *Aging Cell* vol. 18 (2019).

132. Halpern, M., Brennand, K. J. & Gregory, J. Examining the relationship between astrocyte dysfunction and neurodegeneration in ALS using hiPSCs. *Neurobiology of Disease* vol. 132 (2019).

133. Chang, C. Y. *et al.* Induced pluripotent stem cell (iPSC)-based neurodegenerative disease models for phenotype recapitulation and drug screening. *Molecules* vol. 25 (2020).

134. Namboori, S. C. *et al.* Single cell transcriptomics identifies master regulators of dysfunctional pathways in SOD1 ALS motor neurons. *bioRxiv* 593129 (2019) doi:10.1101/593129.

135. Liu, W. *et al.* Single-cell RNA-seq analysis of the brainstem of mutant SOD1 mice reveals perturbed cell types and pathways of amyotrophic lateral sclerosis. *Neurobiol. Dis.* **141**, 104877 (2020).

136. Maniatis, S. *et al.* Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science (80-.).* **364**, 89–93 (2019).

137. Chipika, R. H., Finegan, E., Li Hi Shing, S., Hardiman, O. & Bede, P. Tracking a Fast-Moving Disease: Longitudinal Markers, Monitoring, and Clinical Trial Endpoints in ALS. *Front. Neurol.* **10**, (2019).

138. Webster, C. P., Smith, E. F., Shaw, P. J. & De Vos, K. J. Protein homeostasis in amyotrophic lateral sclerosis: Therapeutic opportunities? *Frontiers in Molecular Neuroscience* vol. 10 (2017).

139. Ruegsegger, C. & Saxena, S. Proteostasis impairment in ALS. *Brain Research* vol. 1648 571–579 (2016).

140. Barschke, P., Oeckl, P., Steinacker, P., Ludolph, A. & Otto, M. Proteomic studies in the discovery of cerebrospinal fluid biomarkers for amyotrophic lateral sclerosis. *Expert Review of Proteomics* vol. 14 769–777 (2017).

141. Hedl, T. J. *et al.* Proteomics approaches for biomarker and drug target discovery in als and ftd. *Frontiers in Neuroscience* vol. 13 (2019).

142. Yerbury, J. J., Farrawell, N. E. & McAlary, L. Proteome Homeostasis Dysfunction: A Unifying Principle in ALS Pathogenesis. *Trends in Neurosciences* vol. 43 274–284 (2020).

143. Baloh, R. H. TDP-43: The relationship between protein aggregation and neurodegeneration in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *FEBS Journal* vol. 278 3539–3549 (2011).

144. Parakh, S. & Atkin, J. D. Protein folding alterations in amyotrophic lateral sclerosis. *Brain Research* vol. 1648 633–649 (2016).

145. Nguyen, D. K. H., Thombre, R. & Wang, J. Autophagy as a common pathway in amyotrophic lateral sclerosis. *Neuroscience Letters* vol. 697 34–48 (2019).

146. Malik, R. & Wiedau, M. Therapeutic Approaches Targeting Protein Aggregation in Amyotrophic Lateral Sclerosis. *Frontiers in Molecular Neuroscience* vol. 13 (2020).

147. Perlson, E., Maday, S., Fu, M. meng, Moughamian, A. J. & Holzbaur, E. L. F. Retrograde axonal transport: Pathways to cell death? *Trends in Neurosciences* vol. 33 335–344 (2010).

148. Prell, T., Lautenschläger, J. & Grosskreutz, J. Calcium-dependent protein folding in amyotrophic lateral sclerosis. *Cell Calcium* vol. 54 132–143 (2013).

149. Tang, A. Y. RNA-binding proteins associated molecular mechanisms of motor neuron degeneration pathogenesis. *Molecular Biotechnology* vol. 56 779–786 (2014).

150. Navone, F., Genevini, P. & Borgese, N. Autophagy and Neurodegeneration: Insights from a Cultured Cell Model of ALS. *Cells* **4**, 354–386 (2015).

151. Weskamp, K. & Barmada, S. J. TDP43 and RNA instability in amyotrophic lateral sclerosis. *Brain Research* vol. 1693 67–74 (2018).

152. Chew, J. *et al.* C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science (80-.).* **348**, 1151–1154 (2015).

153. Mackenzie, I. R. A. *et al.* Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Ann. Neurol.* **61**, 427–434 (2007).

154. Amador-Ortiz, C. *et al.* TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Ann. Neurol.* **61**, 435–445 (2007).

155. Umoh, M. E. *et al.* A proteomic network approach across the ALS - FTD disease spectrum resolves clinical phenotypes and genetic vulnerability in human brain . *EMBO Mol. Med.* **10**, 48–62 (2018).

156. Mackenzie, I. R. A. & Rademakers, R. The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Current Opinion in Neurology* vol. 21 693–700 (2008).

157. Farrawell, N. E. *et al.* Distinct partitioning of ALS associated TDP-43, FUS and SOD1 mutants into cellular inclusions. *Sci. Rep.* **5**, (2015).

158. Jeon, G. S. *et al.* Pathological Modification of TDP-43 in Amyotrophic Lateral Sclerosis with SOD1 Mutations. *Mol. Neurobiol.* **56**, 2007–2021 (2019).

159. Hedl, T. J. *et al.* Proteomics approaches for biomarker and drug target discovery in als and ftd. *Frontiers in Neuroscience* vol. 13 (2019).

160. Ekegren, T., Hanrieder, J. & Bergquist, J. Clinical perspectives of high-resolution mass spectrometry-based proteomics in neuroscience: exemplified in amyotrophic lateral sclerosis biomarker discovery research. *J. Mass Spectrom.* **43**, 559–71 (2008).

161. Collins, M. A., An, J., Hood, B. L., Conrads, T. P. & Bowser, R. P. Label-free LC-MS/MS proteomic analysis of cerebrospinal fluid identifies protein/pathway alterations and candidate biomarkers for amyotrophic lateral sclerosis. *J. Proteome Res.* **14**, 4486–4501 (2015).

162. Krüger, T., Lautenschläger, J., Grosskreutz, J. & Rhode, H. Proteome analysis of body fluids for amyotrophic lateral sclerosis biomarker discovery. *PROTEOMICS - Clin. Appl.* **7**, 123–135 (2013).

163. Pasinetti, G. M. *et al.* Identification of potential CSF biomarkers in ALS. *Neurology* **66**, 1218–22 (2006).

164. Bowser, R., Cudkowicz, M. & Kaddurah-Daouk, R. Biomarkers for amyotrophic lateral sclerosis. *Expert Rev. Mol. Diagn.* **6**, 387–98 (2006).

165. Ryberg, H. & Bowser, R. Protein biomarkers for amyotrophic lateral sclerosis. *Expert Review of Proteomics* vol. 5 249–262 (2008).

166. von Neuhoff, N. *et al.* Monitoring CSF Proteome Alterations in Amyotrophic Lateral Sclerosis: Obstacles and Perspectives in Translating a Novel

Marker Panel to the Clinic. *PLoS One* **7**, (2012).

167. Lehmer, C. *et al.* Poly- <scp>GP</scp> in cerebrospinal fluid links *C9orf72* -associated dipeptide repeat expression to the asymptomatic phase of <scp>ALS</scp> / <scp>FTD</scp>. *EMBO Mol. Med.* **9**, 859–868 (2017).

168. Oeckl, P. *et al.* Proteomics in cerebrospinal fluid and spinal cord suggests UCHL1, MAP2 and GPNMB as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis. *Acta Neuropathol.* **139**, 119–134 (2020).

169. Zubiri, I. *et al.* Tissue-enhanced plasma proteomic analysis for disease stratification in amyotrophic lateral sclerosis. *Mol. Neurodegener.* **13**, (2018).

170. Collins, M. A. Identification of Amyotrophic Lateral Sclerosis Disease Mechanisms by Cerebrospinal Fluid Proteomic Profiling. (2016).

171. Benatar, M., Wuu, J., Andersen, P. M., Lombardi, V. & Malaspina, A. Neurofilament light: A candidate biomarker of presymptomatic amyotrophic lateral sclerosis and phenoconversion. *Ann. Neurol.* **84**, 130–139 (2018).

172. Zhu, S. *et al.* Targeted Multiple Reaction Monitoring Analysis of CSF Identifies UCHL1 and GPNMB as Candidate Biomarkers for ALS. *J. Mol. Neurosci.* **69**, 643–657 (2019).

173. Online, R. *et al.* Proteomics Approaches for Biomarker and Drug Target Discovery in ALS and FTD Publication Details. doi:10.3389/fnins.2019.00548.

174. Verde, F. *et al.* Neurofilament light chain in serum for the diagnosis of amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **90**, 157–164 (2019).

175. De Schaepdryver, M. *et al.* Serum neurofilament heavy chains as early marker of motor neuron degeneration. *Ann. Clin. Transl. Neurol.* **6**, 1971–1979 (2019).

176. De Schaepdryver, M. *et al.* Comparison of elevated phosphorylated neurofilament heavy chains in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **89**, 367–373 (2018).

177. Gaiottino, J. *et al.* Increased Neurofilament Light Chain Blood Levels in Neurodegenerative Neurological Diseases. *PLoS One* **8**, (2013).

178. Lu, C. H. *et al.* Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis. *Neurology* **84**, 2247–2257 (2015).

179. Gendron, T. F. *et al.* Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72-associated amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, (2017).

180. Balendra, R., Moens, T. G. & Isaacs, A. M. Specific biomarkers for C9orf72 FTD / ALS could expedite the journey towards effective therapies . *EMBO Mol. Med.* **9**, 853–855 (2017).

181. Floeter, M. K. & Gendron, T. F. Biomarkers for Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Associated With Hexanucleotide Expansion Mutations in C9orf72. *Front. Neurol.* **9**, (2018).

182. Mitchell, R. M. *et al.* A CSF biomarker panel for identification of patients with amyotrophic lateral sclerosis. *Neurology* **72**, 14–19 (2009).

183. Barschke, P., Oeckl, P., Steinacker, P., Ludolph, A. & Otto, M. Proteomic studies in the discovery of cerebrospinal fluid biomarkers for amyotrophic lateral sclerosis. *Expert Review of Proteomics* vol. 14 769–777 (2017).

184. Ranganathan, S. *et al.* Proteomic profiling of cerebrospinal fluid identifies biomarkers for amyotrophic lateral sclerosis. *J. Neurochem.* **95**, 1461–1471 (2005).

185. Shurte, L. Determining Protein-Protein Interactions of ALS-Associated SOD1. *Brows. all Theses Diss.* (2016).

186. Vella, D., Zoppis, I., Mauri, G., Mauri, P. & Di Silvestre, D. From protein-protein interactions to protein co-expression networks: a new perspective to evaluate large-scale proteomic data. *Eurasip Journal on Bioinformatics and Systems Biology* vol. 2017 1–16 (2017).

187. Snider, J. *et al.* Fundamentals of protein interaction network mapping. *Mol. Syst. Biol.* **11**, 848 (2015).

188. Mao, Y., Kuo, S.-W., Chen, L., Heckman, C. J. & Jiang, M. C. The essential and downstream common proteins of amyotrophic lateral sclerosis: A protein-protein interaction network analysis. *PLoS One* **12**, e0172246 (2017).

189. Rao, V. S., Srinivas, K., Sujini, G. N. & Kumar, G. N. S. Protein-Protein Interaction Detection: Methods and Analysis. *Int. J. Proteomics* **2014**, (2014).

190. Kumar, A., Ghosh, D. & Singh, R. L. Amyotrophic Lateral Sclerosis and Metabolomics: Clinical Implication and Therapeutic Approach. *J. Biomarkers* **2013**, 1–15 (2013).

191. Blasco, H. *et al.* Metabolomics in amyotrophic lateral sclerosis: How far can it take us? *European Journal of Neurology* vol. 23 447–454 (2016).

192. Lanznaster, D., de Assis, D. R., Corcia, P., Pradat, P. F. & Blasco, H. Metabolomics biomarkers: A strategy toward therapeutics improvement in ALS. *Frontiers in Neurology* vol. 9 (2018).

193. Germeys, C., Vandoorne, T., Bercier, V. & Van Den Bosch, L. Existing and emerging metabolomic tools for ALS research. *Genes* vol. 10 (2019).

194. Lanznaster, D. *et al.* Metabolomics: A tool to understand the impact of genetic mutations in amyotrophic lateral sclerosis. *Genes* vol. 11 (2020).

195. Lanznaster, D., de Assis, D. R., Corcia, P., Pradat, P.-F. & Blasco, H. Metabolomics Biomarkers: A Strategy Toward Therapeutics Improvement in ALS. *Front. Neurol.* **9**, 1126 (2018).

196. Gross, S. *et al.* Identification of a Distinct Metabolomic Subtype of Sporadic ALS Patients. *bioRxiv* 416396 (2018) doi:10.1101/416396.

197. Jääskeläinen, O. *et al.* Low Serum High-Density Lipoprotein Cholesterol Levels Associate with the C9orf72 Repeat Expansion in Frontotemporal Lobar Degeneration Patients. *J. Alzheimer's Dis.* **72**, 127–137 (2019).

198. Wuolikainen, A., Andersen, P. M., Moritz, T., Marklund, S. L. & Antti, H. ALS patients with mutations in the SOD1 gene have an unique metabolomic profile in the cerebrospinal fluid compared with ALS patients without mutations. *Mol. Genet. Metab.* **105**, 472–478 (2012).

199. Lanznaster, D. *et al.* Metabolomics: A Tool to Understand the Impact of Genetic Mutations in Amyotrophic Lateral Sclerosis. *Genes* vol. 11 (2020).

200. Germeys, C., Vandoorne, T., Bercier, V. & Van Den Bosch, L. Existing and emerging metabolomic tools for ALS research. *Genes* vol. 10 (2019).

201. Dorst, J. *et al.* Patients with elevated triglyceride and cholesterol serum levels have a prolonged survival in amyotrophic lateral sclerosis. *J. Neurol.* **258**, 613–617 (2011).

202. De Aguilar, J. L. G. Lipid biomarkers for amyotrophic lateral sclerosis. *Front. Neurol.* **10**, (2019).

203. Zufiría, M. *et al.* ALS: A bucket of genes, environment, metabolism and unknown ingredients. *Progress in Neurobiology* vol. 142 104–129 (2016).

204. De Aguilar, J. L. G. Lipid biomarkers for amyotrophic lateral sclerosis. *Front. Neurol.* **10**, 284 (2019).

205. Zeng, P. & Zhou, X. Causal effects of blood lipids on amyotrophic lateral sclerosis: A Mendelian randomization study. *Hum. Mol. Genet.* **28**, 688–697 (2019).

206. Rattray, N. J. W. & Daouk, R. K. Pharmacometabolomics and Precision Medicine Special Issue Editorial. *Metabolomics* vol. 13 1–4 (2017).

207. Blasco, H. *et al.* A pharmaco-metabolomics approach in a clinical trial of ALS: Identification of predictive markers of progression. *PLoS One* **13**, (2018).

208. Blasco, H. *et al.* A pharmaco-metabolomics approach in a clinical trial of ALS: Identification of predictive markers of progression. *PLoS One* **13**, (2018).

209. Jirtle, R. L. Epigenome: the program for human health and disease. *Epigenomics* **1**, 13–16 (2009).

210. Dolinar, A., Ravnik-Glavač, M. & Glavač, D. Epigenetic mechanisms in amyotrophic lateral sclerosis: A short review. *Mechanisms of Ageing and Development* vol. 174 103–110 (2018).

211. Ebbert, M. T. W., Lank, R. J. & Belzil, V. V. An epigenetic spin to ALS and FTD. in *Advances in Neurobiology* vol. 20 1–29 (Springer New York LLC, 2018).

212. Bennett, S. A., Tanaz, R., Cobos, S. N. & Torrente, M. P. Epigenetics in amyotrophic lateral sclerosis: a role for histone post-translational modifications in neurodegenerative disease. *Translational Research* vol. 204 19–30 (2019).

213. Douglas, A. G. L. Non-coding RNA in C9orf72-related amyotrophic lateral sclerosis and frontotemporal dementia: A perfect storm of dysfunction. *Non-coding RNA Research* vol. 3 178–187 (2018).

214. Wang, T., Zhang, J. & Xu, Y. Epigenetic basis of lead-induced neurological disorders. *International Journal of Environmental Research and Public Health* vol. 17 1–23 (2020).

215. Klingl, Y. E., Pakravan, D. & Van Den Bosch, L. Opportunities for histone deacetylase inhibition in amyotrophic lateral sclerosis. *British Journal of Pharmacology* (2020) doi:10.1111/bph.15217.

216. Calió, M. L. *et al.* Mitochondrial Dysfunction, Neurogenesis, and Epigenetics: Putative Implications for Amyotrophic Lateral Sclerosis Neurodegeneration and Treatment. *Frontiers in Neuroscience* vol. 14 (2020).

217. Bennett, S. A., Tanaz, R., Cobos, S. N. & Torrente, M. P. Epigenetics in amyotrophic lateral sclerosis: a role for histone post-translational modifications in neurodegenerative disease. *Translational Research* vol. 204 19–30 (2019).

218. Paez-Colasante, X., Figueroa-Romero, C., Sakowski, S. A., Goutman, S. A. & Feldman, E. L. Amyotrophic lateral sclerosis: Mechanisms and therapeutics in the epigenomic era. *Nature Reviews Neurology* vol. 11 266–279 (2015).

219. Dolinar, A., Ravnik-Glavač, M. & Glavač, D. Epigenetic mechanisms in amyotrophic lateral sclerosis: A short review. *Mechanisms of Ageing and Development* vol. 174 103–110 (2018).

220. Masala, A. *et al.* Epigenetic Changes Associated with the Expression of Amyotrophic Lateral Sclerosis (ALS) Causing Genes. *Neuroscience* **390**, 1–11 (2018).

221. Poesen, K. The Chromosomal Conformation Signature: A New Kid on the Block in ALS Biomarker Research? *EBioMedicine* vol. 33 6–7 (2018).

222. Salter, M. *et al.* Initial Identification of a Blood-Based Chromosome Conformation Signature for Aiding in the Diagnosis of Amyotrophic Lateral Sclerosis. *EBioMedicine* **33**, 169–184 (2018).

223. Poesen, K. The Chromosomal Conformation Signature: A New Kid on the Block in ALS Biomarker Research? *EBioMedicine* **33**, 6–7 (2018).

224. Sharma, S. & Lu, H. C. microRNAs in Neurodegeneration: Current Findings and Potential Impacts. *J. Alzheimer's Dis. Park.* **08**, (2018).

225. Ricci, C., Marzocchi, C. & Battistini, S. MicroRNAs as Biomarkers in Amyotrophic Lateral Sclerosis. *Cells* **7**, 219 (2018).

226. Rizzuti, M. *et al.* MicroRNA expression analysis identifies a subset of downregulated miRNAs in ALS motor neuron progenitors. *Sci. Rep.* **8**, 1–12 (2018).

227. Joilin, G., Leigh, P. N., Newbury, S. F. & Hafezparast, M. An Overview of MicroRNAs as Biomarkers of ALS. *Front. Neurol.* **10**, 186 (2019).

228. Figueroa-Romero, C. *et al.* Expression of microRNAs in human post-mortem amyotrophic lateral sclerosis spinal cords provides insight into disease mechanisms. *Mol. Cell. Neurosci.* **71**, 34–45 (2016).

229. Freischmidt, A. *et al.* Serum microRNAs in sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* **36**, 2660.e15-2660.e20 (2015).

230. Takahashi, I. *et al.* Identification of plasma microRNAs as a biomarker of sporadic Amyotrophic Lateral Sclerosis. *Mol. Brain* **8**, 67 (2015).

231. Toivonen, J. M. *et al.* MicroRNA-206: A Potential Circulating Biomarker Candidate for Amyotrophic Lateral Sclerosis. *PLoS One* **9**, e89065 (2014).

232. De Luna, N. *et al.* Downregulation of miR-335-5P in Amyotrophic Lateral Sclerosis Can Contribute to Neuronal Mitochondrial Dysfunction and Apoptosis. *Sci. Rep.* **10**, 1–12 (2020).

233. Ma, G. *et al.* MiR-206, a key modulator of skeletal muscle development and disease. *Int. J. Biol. Sci.* **11**, 345–352 (2015).

234. DI Pietro, L. *et al.* Potential therapeutic targets for ALS: MIR206, MIR208b and MIR499 are modulated during disease progression in the skeletal muscle of patients. *Sci. Rep.* **7**, 1–11 (2017).

235. Mitropoulos, K., Katsila, T., Patrinos, G. P. & Pampalakis, G. Multi-Omics for Biomarker Discovery and Target Validation in Biofluids for Amyotrophic Lateral Sclerosis Diagnosis. *Omi. A J. Integr. Biol.* **22**, 52–64 (2018).

236. Mirza, B. *et al.* Machine learning and integrative analysis of biomedical big data. *Genes* vol. 10 (2019).

237. Yu, X. T. & Zeng, T. Integrative analysis of omics big data. in *Methods in Molecular Biology* vol. 1754 109–135 (Humana Press Inc., 2018).

238. Nguyen, N. D. & Wang, D. Multiview learning for understanding functional multiomics. *PLoS Computational Biology* vol. 16 e1007677 (2020).

239. Olivier, M., Asmis, R., Hawkins, G. A., Howard, T. D. & Cox, L. A. The need for multi-omics biomarker signatures in precision medicine. *International Journal of Molecular Sciences* vol. 20 (2019).

240. Mitropoulos, K., Katsila, T., Patrinos, G. P. & Pampalakis, G. Multi-omics for biomarker discovery and target validation in biofluids for amyotrophic lateral sclerosis diagnosis. *Omi. A J. Integr. Biol.* **22**, 52–64 (2018).

241. Lam, S. *et al.* A systems biology approach for studying neurodegenerative diseases. *Drug Discov. Today* (2020) doi:10.1016/j.drudis.2020.05.010.

242. Santiago, J. A., Bottero, V. & Potashkin, J. A. Dissecting the molecular mechanisms of neurodegenerative diseases through network biology. *Frontiers in Aging Neuroscience* vol. 9 (2017).

243. Bu, L. L. *et al.* Toward precision medicine in Parkinson's disease. *Ann. Transl. Med.* **4**, (2016).

244. Hampel, H. *et al.* Precision pharmacology for Alzheimer's disease. *Pharmacological Research* vol. 130 331–365 (2018).

245. Castrillo, J. I., Lista, S., Hampel, H. & Ritchie, C. W. Systems biology methods for Alzheimer's disease research toward molecular signatures, subtypes, and stages and precision medicine: Application in cohort studies and trials. in *Methods in Molecular Biology* vol. 1750 31–66 (Humana Press Inc., 2018).

246. Hampel, H. *et al.* Revolution of Alzheimer Precision Neurology. Passageway of Systems Biology and Neurophysiology. *Journal of Alzheimer's Disease* vol. 64 S47–S105 (2018).

247. Morello, G. *et al.* Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS. *Sci. Rep.* **9**, 9968 (2019).

248. Apolloni, S. *et al.* Histamine Regulates the Inflammatory Profile of SOD1-G93A Microglia and the Histaminergic System Is Dysregulated in Amyotrophic Lateral Sclerosis. *Front. Immunol.* **8**, 1689 (2017).

249. Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol. Dis.* (2015) doi:10.1016/j.nbd.2014.12.002.

250. Rosengren, L. E., Karlsson, J.-E., Karlsson, J.-O., Persson, L. I. & Wikkelsø, C. Patients with Amyotrophic Lateral Sclerosis and Other Neurodegenerative Diseases Have Increased Levels of Neurofilament Protein in CSF. *J. Neurochem.* **67**, 2013–2018 (2002).

251. Süssmuth, S. D., Tumani, H., Ecker, D. & Ludolph, A. C. Amyotrophic lateral sclerosis: Disease stage related changes of tau protein and S100 beta in cerebrospinal fluid and creatine kinase in serum. *Neurosci. Lett.* **353**, 57–60 (2003).

252. Brettschneider, J. *et al.* Neurofilament heavy-chain NfHSMI35 in cerebrospinal fluid supports the differential diagnosis of Parkinsonian syndromes. *Mov. Disord.* **21**, 2224–2227 (2006).

253. Zetterberg, H., Jacobsson, J., Rosengren, L., Blennow, K. & Andersen, P. M. Cerebrospinal fluid neurofilament light levels in amyotrophic lateral sclerosis: Impact of SOD1 genotype. *Eur. J. Neurol.* **14**, 1329–1333 (2007).

254. Moreau, C. *et al.* Paradoxical response of VEGF

expression to hypoxia in CSF of patients with ALS. *J. Neurol. Neurosurg. Psychiatry* **77**, 255–257 (2006).

255. Zhao, Z. *et al.* Vgf is a novel biomarker associated with muscle weakness in amyotrophic lateral sclerosis (ALS), with a potential role in disease pathogenesis. *Int. J. Med. Sci.* **5**, 92–99 (2008).

256. Pasinetti, G. M. *et al.* Identification of potential CSF biomarkers in ALS. *Neurology* **66**, 1218–1222 (2006).

257. Tanaka, M. *et al.* Intrathecal upregulation of granulocyte colony stimulating factor and its neuroprotective actions on motor neurons in amyotrophic lateral sclerosis. *J. Neuropathol. Exp. Neurol.* **65**, 816–825 (2006).

258. Hosback, S. *et al.* Circulating insulin-like growth factors and related binding proteins are selectively altered in amyotrophic lateral sclerosis and multiple sclerosis. *Growth Horm. IGF Res.* **17**, 472–479 (2007).

259. Johansson, A., Larsson, A., Nygren, I., Blennow, K. & Askmark, H. Increased serum and cerebrospinal fluid FGF-2 levels in amyotrophic lateral sclerosis. *Neuroreport* **14**, 1867–1869 (2003).

260. Tsuboi, Y., Kakimoto, K., Akatsu, H., Daikuhara, Y. & Yamada, T. Hepatocyte growth factor in cerebrospinal fluid in neurologic disease. *Acta Neurol. Scand.* **106**, 99–103 (2002).

261. Furukawa, T. *et al.* CSF cytokine profile distinguishes multifocal motor neuropathy from progressive muscular atrophy. *Neurol. Neuroimmunol. NeuroInflammation* **2**, e138 (2015).

262. Lu, C.-H. *et al.* Systemic inflammatory response and neuromuscular involvement in amyotrophic lateral sclerosis. *Neurol. - Neuroimmunol. Neuroinflammation* **3**, e244 (2016).

263. Bilic, E. *et al.* Comparison of the growth hormone, IGF-1 and insulin in cerebrospinal fluid and serum between patients with motor neuron disease and healthy controls. *Eur. J. Neurol.* **13**, 1340–1345 (2006).

264. Andrés-Benito, P. *et al.* Inflammatory gene expression in whole peripheral blood at early stages of sporadic amyotrophic lateral sclerosis. *Front. Neurol.* **8**, (2017).

265. Shi, N. *et al.* Increased IL-13-producing T cells in ALS: Positive correlations with disease severity and progression rate. *J. Neuroimmunol.* **182**, 232–235 (2007).

266. Guo, J., Yang, X., Gao, L. & Zang, D. Evaluating the levels of CSF and serum factors in ALS. *Brain Behav.* **7**, (2017).

267. Duque, T., Gromicho, M., Pronto-Laborinho, A. C. & de Carvalho, M. Transforming growth factor-ß plasma levels and its role in amyotrophic lateral sclerosis. *Med. Hypotheses* **139**, (2020).

268. Benninger, F., Glat, M. J., Offen, D. & Steiner, I. Glial fibrillary acidic protein as a marker of astrocytic activation in the cerebrospinal fluid of patients with amyotrophic lateral sclerosis. *J. Clin. Neurosci.* **26**, 75–78 (2016).

269. Tateishi, T. *et al.* CSF chemokine alterations related to the clinical course of amyotrophic lateral sclerosis. *J. Neuroimmunol.* **222**, 76–81 (2010).

270. Niebroj-Dobosz, I., Janik, P., Sokołowska, B. & Kwiecinski, H. Matrix metalloproteinases and their tissue inhibitors in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. *Eur. J. Neurol.* **17**, 226–231 (2010).

271. Beuche, W. *et al.* Matrix metalloproteinase-9 is elevated in serum of patients with amyotrophic lateral sclerosis. *Neuroreport* **11**, 3419–3422 (2000).

272. Lorenzl, S., Albers, D. S., Narr, S., Chirichigno, J. & Beal, M. F. Expression of MMP-2, MMP-9, and MMP-1 and their endogenous counterregulators TIMP-1 and TIMP-2 in postmortem brain tissue of Parkinson's disease. *Exp. Neurol.* **178**, 13–20 (2002).

273. Jacobsson, J., Jonsson, P. A., Andersen, P. M., Forsgren, L. & Marklund, S. L. Superoxide dismutase in CSF from amyotrophic lateral sclerosis patients with and without CuZn-superoxide dismutase mutations. *Brain* **124**, 1461–1466 (2001).

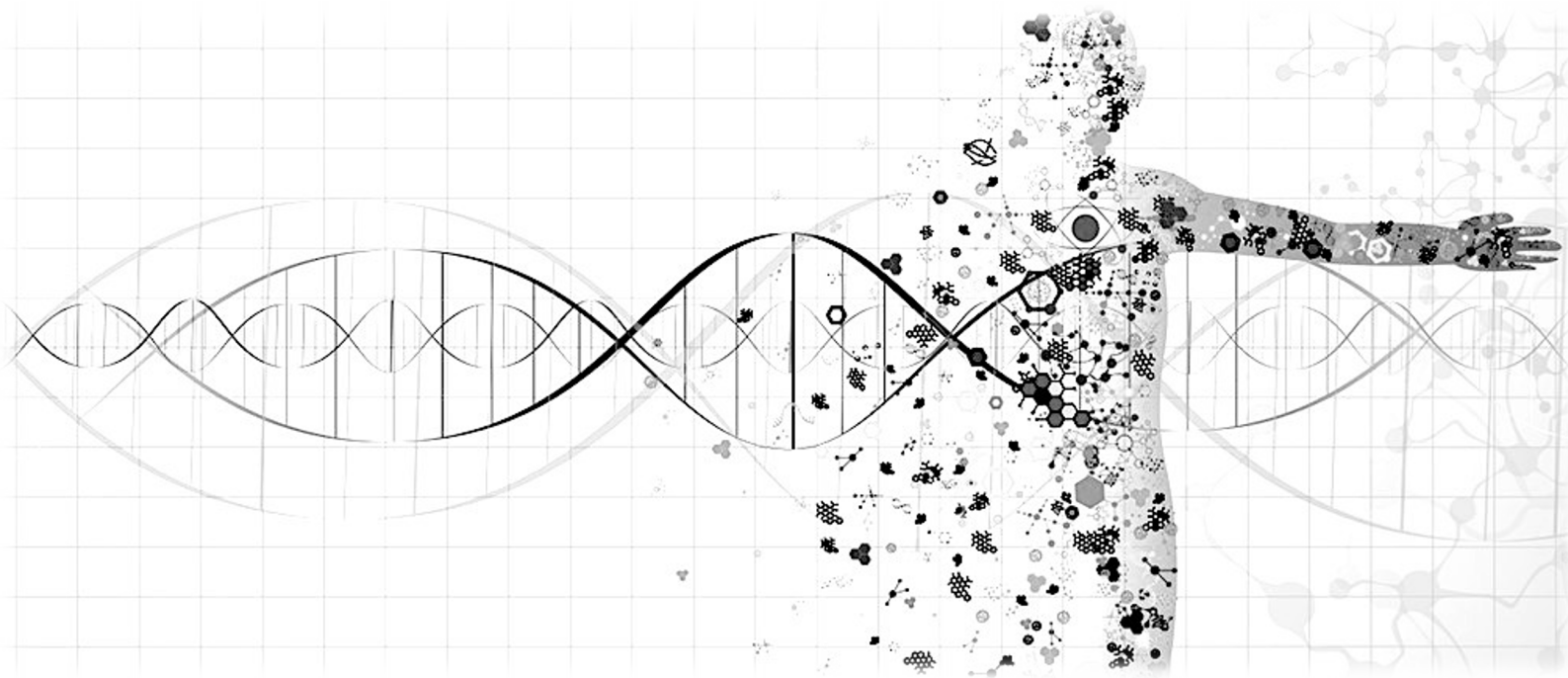
274. Thompson, A. G. *et al.* Cerebrospinal fluid macrophage biomarkers in amyotrophic lateral sclerosis. *Ann. Neurol.* **83**, 258–268 (2018).

275. Kasai, T. *et al.* Combined use of CSF NfL and CSF TDP-43 improves diagnostic performance in ALS. *Ann. Clin. Transl. Neurol.* **6**, 2489–2502 (2019).

276. Majumder, V., Gregory, J. M., Barria, M. A., Green, A. & Pal, S. TDP-43 as a potential biomarker for amyotrophic lateral sclerosis: A systematic review and meta-analysis. *BMC Neurol.* **18**, (2018).

CHAPTER 4

Omics-based exploration and functional validation of neurotrophic factors and histamine as therapeutic targets in ALS



Volontè, C.,et al. Ageing Research Reviews. (2020) 62: 10112 DOI: 10.1016/j.arr.2020.101121

Omics-based exploration and functional validation of neurotrophic factors and histamine as therapeutic targets in ALS

Cinzia Volonté[a,b] *, Giovanna Morello[c], Antonio Gianmaria Spampinato[c], Susanna Amadio[b], Savina Apolloni [b], Velia D'Agata[d], and Sebastiano Cavallaro[c]

[a] CNR, Institute for Systems Analysis and Computer Science, Via Dei Taurini 19, 00185 Rome, Italy; [b] IRCCS Fondazione Santa Lucia, Rome, Italy, Via Del Fosso di Fiorano 65, 00143 Rome, Italy; [c] CNR, Institute for Research and Biomedical Innovation, Via Paolo Gaifami, 18, 95126, Catania, Italy; [d] Section of Human Anatomy and Histology, Department Biomedical and Biotechnological Sciences, University of Catania, Via Santa Sofia, 87, 95100 Catania, Italy.

***Correspondence to**: *Cinzia Volonté*, E-mail: cinzia.volonte@cnr.it. Phone: +3906 501703084. Fax: +3906 501703321

ABSTRACT

A plethora of genetic and molecular mechanisms have been implicated in the pathophysiology of the heterogeneous and multifactorial amyotrophic lateral sclerosis (ALS) disease, and hence the conventional "one target-one drug" paradigm has failed so far to provide effective therapeutic solutions precisely because of this complex nature of ALS. This review intends to highlight how the integration of emerging "omics" approaches may provide a rational foundation for the comprehensive exploration of molecular pathways and dynamic interactions involved in ALS, the identification of candidate targets and biomarkers that assist in the rapid diagnosis and prognosis, finally in the stratification of patients into different subgroups for personalized therapeutic strategies. To this purpose, particular emphasis will be placed on some potential therapeutic targets, including neurotrophic factors and histamine signaling mediators that have emerged as dysregulated at different omics levels in specific subgroups of ALS patients, and have already shown promising results both in *in vitro* and *in vivo* models of ALS. Finally, we will discuss about the utility of using integrated omics coupled with network-based approaches to provide additional guidance for personalization of medicine applications in ALS.

Keywords: ALS, drug targets, histamine, multiomics, network pharmacology, neurotrophins

INTRODUCTION

Amyotrophic Lateral Sclerosis (ALS) is a rare progressive and chronically debilitating motor neuron disease with a new case diagnosed as fast as every 90 minutes. It is estimated that around 450,000 people worldwide are living with ALS and the number of cases is projected to increase as the population ages[1]. The etiology is not defined yet for most ALS patients. In particular, approximately 10%-15% of individuals with ALS have an inherited form of the disease thus classified as familial ALS (fALS)[2], while the cause of the remaining percent of cases, known as sporadic ALS (sALS), is still unknown. Currently, no cure exists and only a small number of treatments (e.g., riluzole, edaravone, non-invasive ventilation) delay death[3].

One of the reasons for this failure is the use of inappropriate recruitment of patients in clinical studies, without taking into account the medical and molecular heterogeneity of ALS. At present, the diagnosis is mainly based on clinical examination, but it is known that the disease starts with hidden symptoms initiating long before the appearance of any pathological sign, i.e. when an initial and peripheral disconnection of motor axons from the muscles gradually evolves into a loss of approximately one-third of the total motor neurons. This occurs long before any muscular weakness and/or atrophy is perceived. Because of this lag between disease onset and symptoms, ALS patients have to face with a first lost opportunity for therapeutic intervention. Due to a further underestimation of early symptoms by both patients and clinicians, and to a diagnosis often based on the exclusion of other possible diseases, another temporal delay of more than one year usually elapses before a definite ALS diagnosis can be formulated. Regrettably, this represents a second lost opportunity for initiating therapy.

Despite tremendous progress has been made during the past decade, there is yet lack of knowledge about ALS pathogenesis, but many hypotheses. It is now recognized that ALS is heterogeneous and multifactorial, characterized by the complex interplay between multiple genetic and environmental factors. Most consolidated pathogenic mechanisms include growth factor deprivation and glutamate toxicity, accompanied by oxidative stress, aberrant free radical handling, mitochondrial failing, metabolic alteration, proteasome dysfunction, protein aggregation, cytoskeletal impairment, and deranged axonal transport[4,5]. This complicated etiopathogenesis is obviously the summation of many pathological events and the main reason why ALS is undefeated today. Thus, far more research is necessary to shed light on the mechanisms underlying ALS, and to find effective and personalized treatments for patients.

Recent advances in high-throughput functional omics technologies and analytical tools for data integration, offer exciting opportunities to decipher the molecular landscape of ALS and the dynamics of disease manifestation. This review intends to highlight the perspectives of integrative multi-omics approaches, with a particular emphasis on some candidate genes and pathways that might be useful for the development of a personalized genome-guided medicine for ALS.

INTEGRATIVE "OMICS" APPROACHES IN THE SEARCH FOR THERAPEUTIC TARGETS IN ALS

The pressing need for treating ALS has not yet met due to the complexity of disease mechanisms and multiple unknown genes and pathways. In the past decade, numerous high-throughput "omics" studies have been conducted on ALS to better understand the biology behind the disease, which will undoubtedly open new doors to therapeutics. Genome-wide association, whole genome/exome sequencing combined with functional

genomics, transcriptomics, proteomics, metabolomics, and other omics have already enabled researchers to obtain comprehensive snapshots of biological systems at real time, and with single molecule resolution, thus identifying multiple genomic alterations related to ALS phenotypes, and providing biological insights to decipher the molecular underpinnings of the disease[6,7]. For instance, a recent study by *Maniatis and coworkers* has reported the use of RNA sequencing to define transcriptomic changes of mouse spinal cords over the course of the disease, as well as of postmortem tissues from ALS patients, identifying disease-associated pathways and establishing the key steps of motor neuron degeneration observed in ALS[8]. Moreover, the growing number of ALS-linked genes has allowed researchers to identify a number of shared intracellular processes contributing to motor neuron degeneration in ALS, including oxidative stress, mitochondrial dysfunction, apoptotic mechanisms, disordered axonal transport, neuroinflammation, and protein aggregation[9,10]. In particular, some mutated RNA-binding proteins, including *TDP43, FUS, SOD1, UBQLN1* and *OPTN*, have been found in aggregates in the cytoplasm of motor neurons in ALS patients, leading either to a broad range of deficits in RNA metabolism or to misfolding and mislocalizations[11]. Other interesting studies have shown that phosphorylated neurofilament heavy subunit protein concentrations in the cerebrospinal fluid (CSF) and blood of ALS patients represent candidate diagnostic and prognostic biomarkers, allowing to: (i) discriminate between patients with ALS and healthy individuals, and among different subtypes of motor neuron patients in the diagnostic assessment at an early stage, when clinical signs are localized and subtle; (ii) provide monitoring during clinical trials to ensure target engagement [12,13,14].

Although each layer of the omics profile offers exciting opportunities to identify important players in ALS pathogenesis, it is not sufficient by itself to acquire the precise picture of the composite molecular machinery involved in motor neuron degeneration.

In this direction, integration of multi-omics measures with existing biological knowledge is essential for understanding how the interaction of multiple pathways is driving disease progression, providing a rational foundation for new potential therapeutic targets and biomarkers that will assist in the rapid diagnosis and prognosis of the disease, and for the stratification of patients into different subgroups with the aim of specific therapeutic strategies. A good example for such an approach is seen in previous works published by our research groups, in which we have established the foundation for a functional molecular taxonomy of ALS. In particular, we have firstly analyzed transcriptional profiles of motor cortex samples from sporadic ALS (sALS) patients and identified two specific subgroups of patients (sALS1 and sALS2) depending on the combinations of genes and pathways that were deregulated[15, 16, 17]. We have identified antigen processing and presentation, and extracellular matrix organization as the most representative subgroup-specific pathways in sALS1, while deregulated genes in sALS2 were associated with axonal guidance, oxidative stress and inflammatory intracellular signaling cascades. On one hand, our pathway-based analysis has confirmed the importance of molecular mechanisms that are known to be implicated in ALS pathogenesis. On the other hand and most importantly, it has been suggested for the first time the differential involvement of these processes in specific subsets of ALS patients. It is interesting to note that a similar stratification of ALS patients into different molecular subtypes has been confirmed in other recent studies [18–20]. In particular, *Tam O.H. et al. (2019)* were able to stratify the transcriptomes of two independent cohorts of ALS postmortem cortex samples into three distinct molecular subtypes, two of which overlapped the molecular signatures observed in our ALS patient samples [18].

To further elucidate the genomic events characterizing sALS pathology, we have recently characterized copy number variants (CNVs) occurring in the same patients [21]. Our findings have revealed distinct genomic signatures associated with the two previously characterized transcriptome-based sALS subgroups, suggesting a strong interaction between genomic and transcriptomic events in ALS. Beyond refining ALS molecular architecture, our goal was to identify and validate potential candidates for genomic-based patient stratification and individualized treatment. In other words, our aim was to identify genes that functionally drive or protect from ALS, and interrogate about their respective signaling pathways and therapeutic targets[22,23,24]. Therefore, genes and pathways identified in the characterization of ALS subtypes were explored by using different pharmacological databases (i.e., PharmGkb, DrugBank), in order to identify compounds that inhibit or induce specific genes or proteins expression, or that block or stimulate specific pathways [25,26]. Our analysis has revealed a good number of potential biomarkers and therapeutic targets that are differentially deregulated in specific subsets of ALS patients, bringing us a step closer to the establishment of a more efficacious and personalized genome-guided medicine for ALS. Besides identifying new potential pharmacological targets, our analysis might also provide a rational approach for "drug repositioning" for ALS. Under this perspective, many known drugs that were abandoned at clinical stages because of their low efficacy and/or toxicity might be re-evaluated in light of the emerging molecular taxonomy of ALS patients. Of note, some of the selected target genes that have emerged from our analysis exhibit conserved expression patterns in mouse and human ALS, thus providing a rationale to ensure their preclinical trial success [22].

Overall, a more complete analysis and evaluation of the molecular characteristics of the disease in each single patient is essential not only to reveal etiopathogenic

mechanisms that were not clear by considering sALS pathology as a single entity (thus providing a powerful means for defining molecular signatures for this disease), but mostly to design more efficacious and individualized therapeutic interventions. In the following sections, we will provide a comprehensive multi-omics characterization of some of the most promising molecular targets and signaling mechanisms that have emerged as dysregulated in our works, focusing on those that have already shown results both in *in vitro* and *in vivo* models of ALS[27-31].

OMICS-BASED EXPLORATION AND FUNCTIONAL VALIDATION OF NEUROTROPHIC FACTORS AND NEUROPEPTIDES AS THERAPEUTIC TARGETS IN ALS

As potential targets for therapy, neurotrophic factors (NTFs) have been exploited in ALS, because of their large spectrum of biological effects, and because they regulate several physiological processes in the CNS, for instance neuronal differentiation and survival, axonal outgrowth and synapses maintenance, proliferation and differentiation of neural stem cells. Interestingly, several studies have already demonstrated the crucial role of NTFs to promote survival, and to be in part protective in both *in vitro* and *in vivo* models of motor neuron degeneration, thus representing a promising therapeutic strategy to treat ALS [32,33]. In agreement with these works, our integrative omics analysis has identified a good number of potent neurotrophic and neuroprotective molecules as statistically deregulated in the same direction both in ALS mouse models and human patients (**Table 1**) [15, 16, 17]. These include, among others, Ciliary Neurotrophic Factor (CNTF), Brain-derived Neurotrophic Factor (BDNF), Pituitary Adenylate cyclase-activating Polypeptide (PACAP), Epidermal growth factor (EGF) and Endothelin-1 (ET-1), Fibroblast Growth Factors (FGFs), Insulin-like Growth Factor 1

(IGF-1), Vascular Endothelial Growth Factor (VEGF), and Granulocyte-Colony Stimulating Factor (G-CSF), together with their corresponding receptors.

Table 1. Neurotrophins and histamine signaling pathways differentially deregulated in cluster SALS patients

SALS1			
Signaling pathway	**Gene target**	**Expression**	**Therapeutic proposal**
EGF signaling	*ErbB2*	**Up-regulated**	EGFR inhibitors (i.e., PKI166, BMS690514, Canertinib, Masoprocol, Gefitinib, Suramin, PD153035, Genistein, Erlotinib, Everolimus)
FGF-family signaling	*FGF1*	**Up-regulated**	FGFR agonist (i.e., PF-05231023)
	FGFR2	**Down-regulated**	
Histamine signaling	*H1R*	**Up-regulated**	H1-H4R modulators (i.e., Clozapine, Orphenadrine, JNJ7777120)
	H4R	**Down-regulated**	

SALS2			
Signaling pathway	**Gene target**	**Expression**	**Therapeutic proposal**
BDNF signaling	*TrkB*	**Up-regulated**	TrkB inhibitors (i.e., Cyclotraxin-B and Ana-12)
PACAP signaling	*ADCYAP1*	**Down-regulated**	PACAP receptor agonists (i.e., Maxadilan)
	ADCYAP1R1		
EGF signaling	*EGF*	**Up-regulated**	EGFR/ErbB inhibitors (i.e., PKI166, BMS690514, Canertinib, Masoprocol, Gefitinib, Suramin, PD153035, Genistein, Erlotinib, Everolimus)
	EGFR		
	ErbB2		
EDN1 signaling	*EDN1*	**Up-regulated**	EDNR-B antagonists (i.e., BQ-788, Bosentan, IRL-2500)
	EDNRB		
FGF-family signaling	*FGF1*	**Up-regulated**	FGFR inhibitors (Orantinib, Brivanib, Dovitinib, Suramin, Pentosan polysulfate)
	FGFR1		
	FGFR2		
	FGFR3		
IGF-1 receptor signaling	*IGF1R*	**Up-regulated**	IGF1R inhibitors (ie., Masoprocol, BMS-754807 and Linsitinib)
VEGF-family signaling	*VEGFA*	**Down-regulated**	VEGF agonists (i.e., SB-509, Celecoxib)
G-CSF signaling	*CSF1*	**Down-regulated**	G-CSF modulators (i.e., pegfilgrastim, JNJ-40346527)
	CSF2RA	**Up-regulated**	
Histamine signaling	*H1R*	**Up-regulated**	Histamine receptor modulators (i.e., Orphenadrine, Ranitidine, Thioperamide, JNJ7777120)
	H2R	**Down-regulated**	
	H3R	**Down-regulated**	
	H4R	**Up-regulated**	

HDC	**Up-regulated**
HNMT	**Up-regulated**
DAO	**Up-regulated**

Ciliary Neurotrophic Factor

CNTF, a polypeptide hormone promoting neurotransmitter synthesis and neurite outgrowth, is one of the first NTFs to be investigated as potential drug target in ALS models, showing to improve motor function and survival and decreasing neuronal degeneration and muscle atrophy, when injected intraperitoneally or subcutaneously in the pmn/pmn and wobbler mice models of motor neuron disease [34,35]. Altered levels of *CNTF* expression are found in the brain and spinal cord of both ALS patients and animal models, and *CNTF* knockout mice develop atrophy and loss of motor neurons with aging, suggesting *CNTF* as a possible modifier gene for ALS [36–38]. In addition, genetic alterations in the *CNTF* gene are also reported in individuals affected by ALS, including for example null mutations affecting splice sites [38]. Despite these premises, in the 90's, two clinical phase I/II and phase III results obtained by subcutaneous administration of recombinant human CNTF (rHCNTF) in ALS patients, have demonstrated no statistically difference between rHCNTF-treated patients and placebo-treated patients, and even adverse effects at high doses[39]. The subsequent, phase I clinical trial with intrathecal pump delivery has suggested intrathecal administration as the potential preferred route for administrating CNTF but, also in this case, no significant clinical benefits were obtained for ALS patients [41,40]. Despite clinical trials in ALS patients have failed to show significant effects of CNTF on disease progression, promising results *in vivo* were obtained when CNTF was co-administrated with other NTFs, such as BDNF, suggesting that a synergistic action of these molecules

(or their derivatives) may overcome side effects and increase the chances of success in arresting or reducing disease progression [41-42].

Brain-derived Neurotrophic Factor

BDNF is a neurotrophic factor that by binding two different receptors, respectively the low affinity p75NTR and the high affinity Tyrosine Receptor Kinase B.T1 (TrkB.T1), modulates glutamate receptor activity, synapse stability, dopaminergic, cholinergic, serotonergic, and GABAergic signaling, synaptogenesis, and dendritogenesis[43]. Increased expression of *BDNF* and its *TrkB* receptor was found in the motor cortex of sALS patients, in accordance with the observation that a prolonged TrkB activation may render the motor neurons more vulnerable to pathophysiological insults, perhaps contributing to ALS [15, 16, 17]. In fact, early synaptic hyper-excitability of motor neurons in ALS apparently enhances BDNF-mediated signaling, thereby causing glutamate excitotoxicity, and motor neuron death. In addition, the decrease of p75NTR expression, as well as the deletion of *TrkB.t1,* correlated with the delay of impairment and mortality in a mouse model of ALS [44-45]. Despite the search for pharmacological compounds interacting with TrkB has been difficult, selective TrkB inhibitors (i.e., Cyclotraxin-B and Ana-12) have now shown promising effects by protecting motor neurons and decreasing neurotoxicity. The manipulation of BDNF/TrkB may thus give rise to neuroprotective therapeutic strategies in the treatment of diseases such as ALS [46,47].

Pituitary Adenylate Cyclase-activating Polypeptide

Another interesting neurotrophic factor is PACAP, belonging to the vasoactive intestinal polypeptide/secretin/glucagon superfamily and involved in a wide range of physiological processes, including cell survival, stress response and cell division [48-49]. PACAP can initiate multiple signaling pathways by binding to three seven-

transmembrane G protein-coupled receptors, PAC1R, VPAC1 and VPAC2, characterized by different ligand-binding specificities. Several *in vitro* and *in vivo* studies have revealed numerous biological activities of PACAP in the peripheral and central nervous system and its protective effects in neurodegenerative disorders including ALS [50-56]. For instance, neuroprotective activities of PACAP are already demonstrated against glutamate-induced excitotoxicity *in vitro* and *in vivo* in the most exploited animal model of ALS that recapitulates many features of the disease, the SOD1-G93A mice [56–59]. In our studies, PACAP and its receptor PAC1R exhibit a completely different mRNA and protein expression profile in the two sALS subgroups, perhaps due to different turnover rates or translation efficacy. While in sALS2 both PACAP and PAC1R are downregulated, in sALS1 they show opposite expression levels [15, 16, 17]. To validate PACAP as therapeutic target in ALS, we have then investigated the potential contribution of PACAP/PAC1R axis in motor neuron survival by using two well-characterized *in vitro* models of ALS, the motor neuron-like hybrid cell line NSC-34 expressing human SOD1-G93A, and human induced Pluripotent Stem Cells (iPSC)-derived motor neurons [60-61]. Our data have demonstrated that PACAP is able to rescue neuronal cells from apoptosis following neurodegenerative stimuli induced by growth factors deprivation, suggesting the involvement of the PACAP-PAC1R pathway in ALS pathology, and its role as a potential drug target to enhance motor neuron viability.

Epidermal Growth Factor

The mechanism underlying neuroprotective properties of PACAP could involve EGF pathway by activating its receptor EGFR and stimulating survival, proliferation, maturation, and migration of different cell types [61-63]. Despite its neuroprotective role, several lines of evidence suggest that the activation of EGFR signaling pathway may

trigger quiescent astrocytes to become reactive astrocytes and, consequently, to play a role in the pathophysiology of many neurodegenerative disease including ALS [63,64]. Consistent with the view that EGFR activation may be a marker of reactive astrogliosis surrounding degenerating motor neurons, we have observed increased expression levels of *EGF* and *EGFR* both in sALS2 patients and SOD1-G93A mouse models, making EGFR signaling pathway and PACAP/EGFR axis as attractive candidates for further pre-clinical studies [15, 16, 17]. To this regard, pharmacological inhibition of EGFR signaling cascade has been already successful in ALS preclinical studies by enhancing axon regeneration, and providing a significant delay in the onset of multiple behavioral measures of disease progression [65-70].

Endothelin-1

Another neuropeptide that is able to interfere with EGFR signaling in astrocytes is ET-1 encoded by the *EDN1* gene that, together with its G-protein coupled receptor B (EDNR-B), is abundantly expressed in sALS motor cortex and in reactive astrocytes in the spinal cord of SOD1-G93A mice [15, 16, 17, 71,72]. EDN1 exerts toxic effects on motor neurons by activating several processes implicated in ALS pathogenesis, such as axonal degeneration, alteration of water homeostasis, increased sensitivity to oxidative stress and excitotoxic damage [73-74]. Moreover, recent studies have shown that EDN1 overexpression may be a direct consequence of reduced C9ORF72 levels, one of the most common known cause of ALS [74,75]. Consistently with these data, we have demonstrated that pharmacological treatments aimed at lowering EDN1 levels or antagonizing its effects may represent interesting therapeutic strategies in ALS, by protecting motor neurons from oxidative stress, inflammation and axonal damage both in ALS animal models and in mixed spinal cord cultures enriched with reactive astrocytes [74,76].

Fibroblast Growth Factors

FGFs constitute a family of multifunctional proteins expressed in the brain where mediate diverse physiologic functions, including cell differentiation, migration and survival, playing an important role in brain development and neuroprotection [77]. The aberrant activity of FGFs and their receptors, either through gain or loss of function, has been involved in different neuropathological conditions, including ALS [78,79]. Differential expression of multiple genes encoding FGFs and their receptors FGFRs has been reported in sALS patients and in SOD1-G93A transgenic mice, starting at a pre-symptomatic stage and progressing with the spread of disease [15, 16, 17]. If, on the one hand, disruption of FGFs signaling induces demyelination and axonal damage in ALS, on the other one, SOD1 aberrant function seems to mediate oxidative-stress-induced damage by inducing the release of FGF-1 from motor neurons that, in turn, activates spinal cord astrocytes and initiates motor neuron apoptosis in ALS [80]. Of note, the selective inhibition of FGFRs has shown promising results in ALS preclinical studies demonstrating to prevent motor neuron death by reducing astrocyte activation and oxidative damage. Similarly, molecular and pharmacological modulation of FGFs signaling may represent a therapeutic approach for ALS due to effects not only on neurogenesis, but also on synaptic formation, neuron-glia interactions and inflammation [80–82].

Insulin-like Growth Factor 1

In addition to the FGFs signaling system, also IGFs are part of a complex system used by cells to communicate with their physiologic environment and regulate normal physiology as well as a number of pathological states. It has been reported that the production of IGF-1 is impaired in skeletal muscle fibers of ALS patients [83].

Administration of IGF-1 is protective in the transgenic rodent model of ALS [84], suggesting the possibility to use IGFs signaling as a potential strategy for the treatment of ALS. While intra-parenchymal spinal cord delivery of IGF in SOD1-G93A mice has shown higher expression of IGF-1 accompanied only by partial rescue of pathological features, a stereotaxic injection into the deep cerebellar nuclei significantly extends mice lifespan [85,86]. Recently, the injection of self-complementary adeno-associated viral vector 9 (scAAV9), a more efficient transducing agent for IGF-1, has extended survival, and motor performance of SOD1-G93A mice when injected either intramuscularly or intravenously [87,88]. Despite these premises, the safety and efficacy identified in animal studies was not translated to human trials. In fact, two randomized double-blind placebo-controlled clinical trials administering recombinant human IGF to patients with ALS have demonstrated limited or no effect on disease progression, while a phase II clinical trial is under investigation but results are not available yet [89,90]. The differential expression of IGF and its receptors between the two subgroups of sALS patients, as demonstrated previously [15], may contribute to explain this clinical failure, suggesting that further clinical investigations should take into consideration genomic heterogeneity of ALS for a more accurate enrollment of patients into clinical trials.

Vascular Endothelial Growth Factor

Trophic factors possess a short time frame for protection of motor neurons once the noxious process is triggered, and this is probably due to the rate at which motor neurons die during the time course of the disease. In murine models of familial ALS, the administration of VEGF before the beginning of symptoms provides a significantly better protection, as proven by delayed symptoms progression and increased lifespan, as compared to that obtained when the trophic factors are administered at the

symptoms onset [91,92]. VEGF is part of the system that restores the oxygen supply to tissues when blood circulation is inadequate such as in hypoxic conditions. VEGF's normal function is to create new blood vessels during embryonic development, after injury, in muscle following exercise, and new vessels to bypass blocked vessels. Decreased expression levels of *VEGF-A* are found in a specific subgroup of sALS patients (sALS2), enforcing the view that reduced VEGF signaling may play a role in the pathogenesis of ALS [15]. Activation of VEGF receptor 2 triggers the phosphorylation of intracellular pathways driven by phosphatidyl-inositol-3-kinase (PI3-K), Akt, phospholipase C-ß, and mitogen-activated protein kinase that promote the inhibition of pro-apoptotic factors like Bad and caspases 9 and 3. The activation of these intracellular signaling pathways has been extensively studied in the CNS, and VEGF-dependent activation of PI3-K/Akt is sufficient to prevent motor neuronal death in familial models of ALS *in vitro* [93,94]. In addition, some drugs activating VEGF (i.e., SB-509, Celecoxib) have been tested in phase II clinical trials for ALS, showing encouraging results [95,96].

Granulocyte-Colony Stimulating Factor

G-CSF, a cytokine/hormone produced by the endothelium, macrophages, and immune cells, stimulates the bone marrow to produce granulocytes and hematopoietic stem cells, and release them into the bloodstream stem. Protective properties with reduction of disease progression and increased survival and rescue of motor neurons are demonstrated in SOD1-G93A mice when delivered continuously of G-CSF [97]. Similar results are also obtained with subcutaneous injection of a more stable analog of G-CSF, pegfilgrastim [98]. It is interesting to note that also another secreted cytokine belonging to the same family, the colony stimulating factor 1 (CSF1), has been associated with inflammation in the central and peripheral nervous system in ALS, particularly by

inducing microglial cell proliferation and neuronal damage in disease [98]. Differential expression of CSF1 is found in the motor cortex of sALS patients as well as in the spinal cord of SOD1-G93A mice [15, 16, 17]. Pharmacological administration of selective inhibitors of CSF1-CSF1R signaling has demonstrated neuroprotective and anti-inflammatory effects in ALS, by reducing microglial cell proliferation and protecting skeletal muscle from denervation, consequently slowing disease progression, attenuating motor neuron cell death and extending survival of SOD1-G93A mice [98-99].

Interactions with Glutamate

Due to their wide spectrum of functions in the CNS, neurotrophic factors have been exploited for decades as therapeutic hypothesis in ALS. However, some intrinsic drawbacks have emerged between preclinical success and clinical failure of trophic factors in ALS therapy, perhaps due to routes of administration, CNS penetrance, safe dosages, treatment start time or long-term efficacy, and need of synergistic drug association. Notwithstanding these limitations, trophic factors remain essential for motor neuron maintenance and survival and are still considered as potential candidate molecules for the treatment of patients with ALS. To this regard, the use of new delivery systems to improve bioavailability, including prodrugs, nanocarriers and small molecules that can mimic the effects of these molecules, may overcome NTFs limits and represent a useful therapeutic strategy for ALS patients.

In addition to their trophic roles, neurotrophic factors are known to interact with glutamate and modify glutamate signaling directly, by modulating the expression of glutamate receptor subunits and calcium-regulating proteins, or also indirectly, by controlling the production of antioxidant enzymes, energy-regulating proteins, and anti-apoptotic or autophagy protein family members. Physiological glutamate has been

known for several years to regulate neurogenesis, neurite outgrowth, synaptogenesis, and neuron survival in the developing and adult mammalian nervous system, other than possessing its canonical role as neurotransmitter at the synapses. The trophic effect of glutamate receptor activation is developmental stage-dependent, and determinant to the selective survival of neurons that have to make proper connections. During this sensitive developmental stage, any interference with the glutamate receptor functioning may generate widespread neuronal loss. This often involves neuron-glia interactions also through glutamate-induced release of trophic factors from glia. For instance, glutamate can stimulate the production of BDNF, which, in turn, modifies neuronal glutamate sensitivity, calcium homeostasis, and plasticity. However, depending on the strength of the stimulus, glutamate receptors can mediate biphasic effects, with excessive stimulation becoming neurotoxic. In particular, under conditions of oxidative and metabolic stress, excessive activation of glutamate receptors may contribute to neuronal dysfunction and degeneration in diseases ranging from stroke to Alzheimer's disease to psychiatric disorders [100]. Attention must therefore be paid to these features, when therapeutic manipulation of excitatory amino acid receptors is considered in the clinical setting.

OMICS-BASED EXPLORATION AND FUNCTIONAL VALIDATION OF HISTAMINE SIGNALING AS THERAPEUTIC TARGET IN ALS

Not only neuropeptides and neurotrophic factors, but also several other neurotransmitters and neuromodulators can interfere with the glutamatergic signaling in the nervous system. Among these, histamine (HA) modulates glutamate receptor activity, increases the excitatory post-synaptic potentials, and facilitates the induction of long-term potentiation in the hippocampus, also favoring the direct release of

glutamate evoked for instance by depolarization in hippocampal synaptosomes and in cultured astrocytes [101,102].

Since its discovery back in 1910 and further acceptance as a neurotransmitter in 1984, HA has indeed gained always more attention in health and disease, although its role in CNS dysfunctions still remains under investigation. Histaminergic neurons are located in the tuberomamillary nucleus of the posterior hypothalamus from where they project their axons all over the CNS, thus granting to HA a well-distinct and important pleiotropic role. As a neuroimmune modulator acting both *in vivo* and *in vitro*, of course HA has a central role in several neurological functions, for instance regulating the sleep-wake cycle, nociception, motor circuits, satiety signaling, energy balance, learning, and memory [103–105]. HA signaling in neuronal and non-neuronal cells including mast cells, is mediated by different classes of proteins abundantly distributed throughout the CNS that are: i) G protein-coupled receptors named H1R, H2R, H3R, H4R; ii) intracellular and extracellular enzymes accountable for HA synthesis (histidine decarboxylase, HDC) and degradation (histamine N-methyltransferase, HNMT, and diamine oxidase, DAO or AOC); iii) transporters among which primarily the VMAT2 that is responsible for HA vesicular uptake in the CNS [106–108]. These receptors, enzymes and transporters are thus the fundamental partners in the intercellular communication system mediated by HA in the CNS and involved in several neuropathological conditions [109].

Recent studies by our group have shown that HA behaves as a bioactive molecule also in ALS [110-112]. First of all, H1R-H4R, HDC, HNMT and DAO proteins are abundantly expressed in primary microglia [29] and motor neurons [27-29] isolated from SOD1-G93A mice. Of note, HA receptors and enzymes are expressed also in microglia and motor neurons localized in lumbar spinal cord of SOD1-G93A mice at symptomatic phase of the disease and, most importantly, they are dysregulated during disease

progression. For example, H1R is down regulated in lumbar spinal cord at pre-symptomatic and symptomatic phases, but up regulated in cortex and hypothalamus at symptomatic and end stage of the disease. H2R is down regulated in spinal cord at end stage and in cortex at symptomatic phase. H3R is increased of about two-fold in the hypothalamus at end stage, while it's not apparently modulated as a function of disease progression in spinal cord and in motor cortex. H4R is increased in spinal cord at symptomatic phase, while in cortex at pre-symptomatic phase. HDC in spinal cord is significantly down regulated at pre-symptomatic phase, but up regulated at end stage of disease, and over expressed in cortex and hypothalamus at symptomatic phase. HNMT is augmented in both spinal cord and cortex at symptomatic phase, and in hypothalamus at symptomatic and end stage of disease. Finally, DAO is up-regulated in lumbar spinal cord and hypothalamus at symptomatic and end stage, but it's not affected during the disease in cortex [29].

In addition to the protein expression data, whole-genome expression profiles of motor cortex and spinal cord from healthy subjects and sporadic patients have demonstrated that numerous genes involved in HA receptors, metabolism, transport, secretion and signal transduction, are deregulated in the two transcriptome-based subgroups of patients sALS1 and sALS2 [27-29,113], segregated by unsupervised hierarchical clustering [15]. In particular, the cortical mRNA expression of H1R, HDC, HNMT and DAO is increased in sALS2, while H2R and H3R are selectively reduced. H4R is differently modulated in sALS1 (down-regulated) versus sALS2 (up-regulated), with respect to healthy individuals. However, spinal cord analysis has demonstrated that H1R is reduced and H3R is increased in sALS1 and sALS2 patients [113]. Despite some peculiarities and exceptions, the HA axis in both sALS1 and sALS2 subgroups of patients demonstrates a substantial up-regulation of H1R in cortex (similarly to what

reported in SOD1-G93A mice) and down-regulation in spinal cord (as found in pre-symptomatic and symptomatic SOD1-G93A mice), in addition to the increased expression of H3R in spinal cord. These results clearly indicate that the histaminergic system is affected during ALS, within relevant tissue- and disease phase-specific deregulations of expression. Furthermore, the vast majority of HA-related genes that appear deregulated in SOD1-G93A mice at terminal stage are also deregulated in at least one of sALS patient subgroups (e.g. HRH1, HRH3). This evidence offers a good rationale for the selection and prioritization of HA genes as potential biomarkers and targets for patient-oriented ALS care.

Under this perspective, a large-scale meta-analysis of genome-wide association studies performed on patients affected by sALS has moreover identified that the Ile105 polymorphism on the Thr105Ile allele in the *HNMT* gene (generating about 60% decrease of HA-degrading activity), causes a trend in delaying the onset of ALS symptom by about three years. This suggests that the Thr105Ile allele in the *HNMT* gene could potentially become an important therapeutic target and protective modifier for the treatment of ALS [115].

By combining gene expression profiles, copy number variants and single nucleotide polymorphisms of ALS patients, further studies have then adopted a multiomics approach for integrating transcriptomic and genomic data with the ALS-linked pathogenic variants obtained from the ALSdb database. This approach has allowed to capture HA pathway associations in ALS [27-29, 113]. The genes coding for HA receptors, enzymes and transporters have shown numerous pathological variants. In particular, genome-wide analysis of multiple genomic aberrations occurring in sALS patients has identified some HA-related genes that are copy number variants-affected (*H2R* and *DAO* show amplification, while *H3R* duplication), and that also show a positive

correlation with transcriptomic changes in one or both of sALS patient subgroups. Among these, we find for instance the genes encoding *ADCYAP1*, *CCKBR* and *H3R*. Moreover, single nucleotide polymorphisms occur in *H1R*, *H2R*, *H3R*, *H4R*, *HDC*, *HNMT* and *DAO* genes [113]. In detail, the histamine H3R presents genetic, transcription as well as protein variations correlated with ALS (**Figure 1**). The genome profiling from sALS patients' data indicates that the *HR3R* gene associates with CNV duplication regions and missense SNV. Moreover, the profiling from post-mortem spinal cords from the two molecularly distinct sALS patients subgroups, proves that the transcription of *HR3R* is increased between 3.2- and 3.3-fold in comparison with individual controls. Finally, protein expression data demonstrates that the receptor is increased at least 2-fold in the hypothalamus of SOD1-G93A mice at end stage of disease. In other words, a clear homogeneity exists between the HA-related HR3R gene driven by protein, gene expression, CNV, and SNP data. This knowledge validates the hypothesis that HA-related genes might indeed represent candidate drivers in ALS pathogenesis, and that the HA axis might become a potential target for therapy.

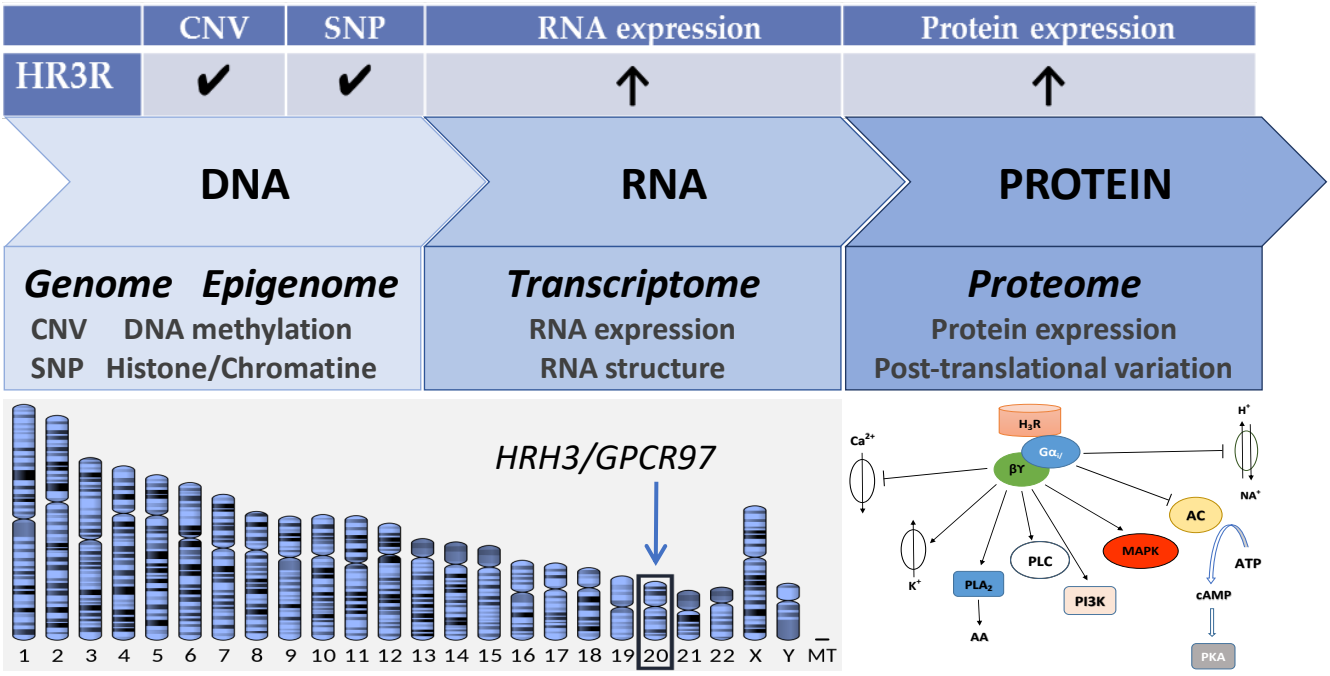


Figure 1. *Multiomic involvement of histamine H3 receptor in ALS. The picture shows that the histamine H3 receptor presents genetic, transcription and protein variations correlated with ALS. In particular from sALS patients' data, the genome profiling indicates that the HR3R gene associates with copy number variants (CNV) duplication regions and missense single nucleotide variants (SNV). Moreover, the profiling from post-mortem spinal cords from two molecularly distinct sALS patients subgroups, shows that the transcription of HR3R is increased between 3.2- and 3.3-fold in comparison with control healthy subject. Finally, protein expression data demonstrates that the receptor is increased at least 2-fold in the hypothalamus of SOD1-G93A mice at end stage of disease. The data are reported from references [27-29, 113].*

In support of this, current research indeed validates this hypothesis by demonstrating that chronic administrations of the HA precursor histidine to SOD1-G93A mice from disease onset up to the end stage of the disease, increases the histamine content in spinal cord, most importantly giving raise to general ameliorative effects. In particular, histidine improves behavioral features of ALS, retards disease progression, recovers motor performance, increases life-span, reduces motor neuron loss and neuroinflammation in spinal cord, and finally ameliorates neuromuscular junction integrity and muscle atrophy [27]. Besides the established neuroprotective effects in motor neurons and anti-inflammatory actions described in microglia, the induction of the heat shock and autophagy response by HA observed in *in vitro* and *in vivo* ALS systems surely contributes to the rescue of motor neurons and spine density loss that occurs in the motor cortex of treated ALS mice [28]. By proposing that the histaminergic modulation can indeed interfere at different levels and within different time frames with the ALS course, we trust that HA might assume a new translational benefit in the development of more effective therapeutics against the disease [116-117].

SYSTEMS PHARMACOLOGY: A NEW PARADIGM FOR DRUG DEVELOPMENT AND PERSONALIZED MEDICINE IN ALS

Given the multifactorial etiopathogenesis of ALS, it is not surprising that the conventional 'one target-one drug" paradigm fails to provide effective treatments

against the disease. Indeed, complex diseases like ALS, are rarely caused by a single gene abnormality, but rather by the perturbation of multiple series of intracellular and intercellular interactions between molecular entities, including protein-protein binding, gene co-expression, RNA interactions, and many other types of molecular functional association [118–121]. On this basis, it would be advisable to shift the ALS drug discovery research towards new alternative approaches aimed at the simultaneous targeting of multiple proteins (and therefore etiologies) involved in disease onset and progression, and at the identification of synergistic drug effects, thus offering the possibility to reach dysfunctional processes that may be impractical to resolve with a single drug therapy. This approach, known as poly-pharmacology, will not only facilitate the discovery of new and interesting drug-target associations, but also will provide a comprehensive understanding of the drug's mechanisms of action and off-target effects, thus offering new hope for the design of more effective and safer treatments for patients [122,123]. To this regard, the application of network-based approaches, particularly protein-protein interaction networks, represents a useful tool in the drug-discovery pipeline, because each protein-target is not working alone, but in a framework containing its connectivity with other proteins, allowing for a comprehensive understanding of the molecular basis of the disease and identification of important alternative targets for its treatment, by using multi-omics datasets [124–127]. Despite in the last years a diverse range of works have been reported in the literature highlighting the power of multi-target approaches and network pharmacology for neurodegenerative diseases, including Charcot-Marie-Tooth disease type I and Alzheimer's disease, the application of polypharmacology to ALS has been minimal [128-130]. To provide an example of how the use of the molecular interactome could substantially support drug discovery for personalized medicine in ALS, we have built two drug-target networks genes encoding neurotrophins and HA

signaling mediators that we have found differentially deregulated in the two sALS patient subgroups (sALS1 and sALS2), together with their pharmacological modulators (**Figures 2-3**). In particular, we downloaded 373 human genes related to neurotrophins (*CNTF, BDNF, PACAP, EGF, ET-1, GFG, IGF-1, VEGF, G-CSF*) and histamine signaling mediators from the Gene Ontology database AmiGO (http://amigo.geneontology.org/amigo), and then evaluated their expression in the two sALS patient subgroups. Next, we investigated drug-gene interactions and potential druggability of these neurotrophins/HA-related differentially expressed genes, by using the Drug Gene Interaction Database (DGIdb, http://dgidb.org). Finally, the two SALS clusters-associated drug-target networks were visualized and analyzed by Cytoscape software (v3.2.0, http://www.cytoscape.org/), in order to identify the most highly-interconnected drugs/targets, known as "hub nods". Personalized drug-target interactions not only provide a comprehensive mapping of the different involvement of neurotrophins and HA signaling cascades in two sALS subgroups, but also identify selective candidate driver genes and drugs, thus providing guidance for precision medicine in the future. For example, both networks highlighted a significant role of EGFR/ERBB signaling in ALS pathogenesis, supporting its relative pharmacological modulators as promising therapeutic approaches. In particular, *ERBB2* and one of its modulators, everolimus, were the most interconnected target/drug in the sALS1-related network and among the top five hub nods in the sALS2-related network (**Figures 2 and 3, Table 2**). Despite not already tested in ALS, the rapamycin analogue everolimus has demonstrated to reduce degeneration of neurons and ameliorate early declines in motor performance in preclinical models of multiple neurodegenerative diseases including Huntington's and Alzheimer's diseases, suggesting its repositioning in ALS [134–137]. In addition, the sALS1-related network has also revealed the selective alteration of tumor

necrosis factor-α (TNF), suggesting that its pharmacological modulation may represent a therapeutic strategy to block or slow disease progression in specific subgroups of ALS patients (**Figure 2, Table 2**).

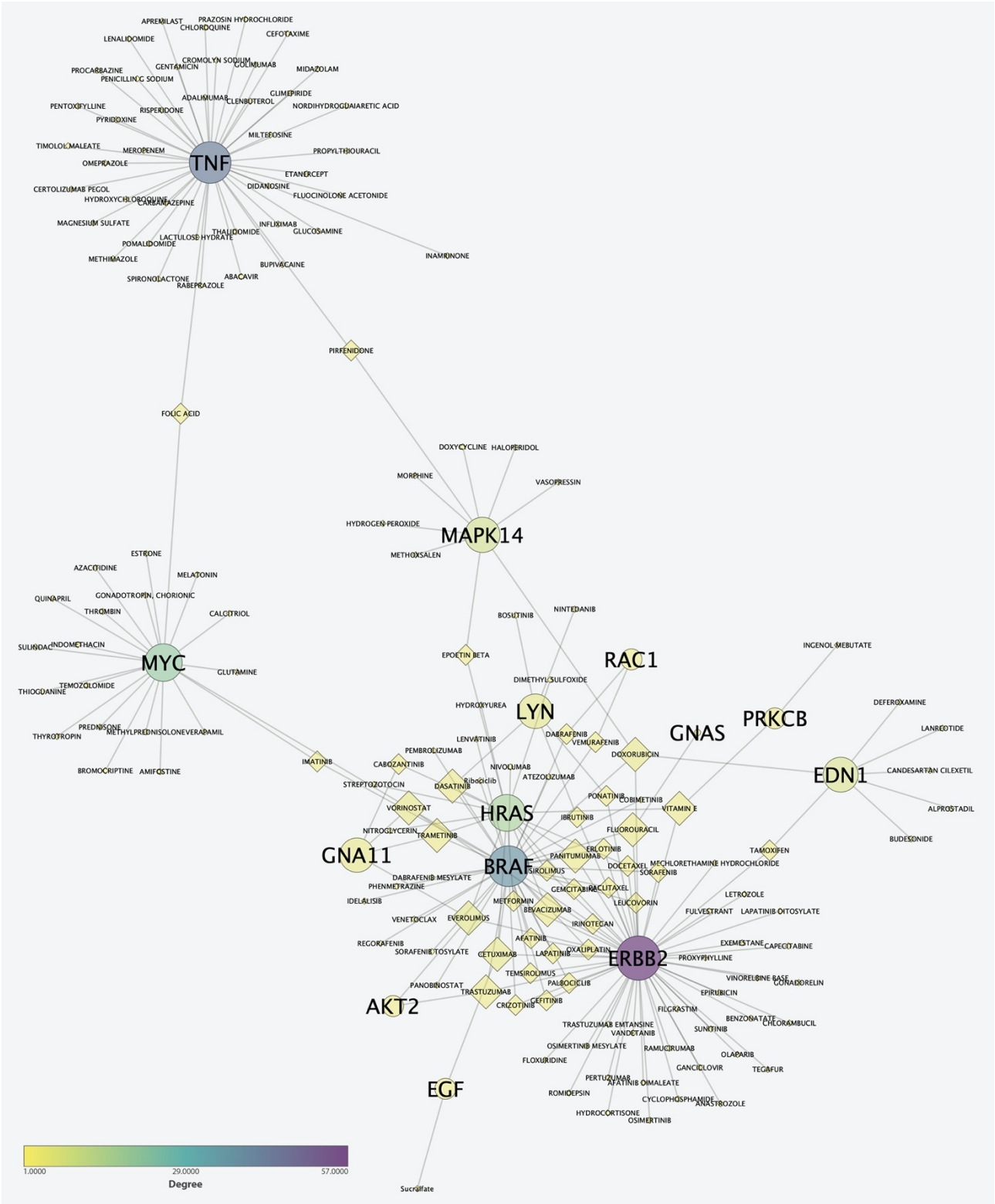


Figure 2. *Drug–Gene network in sALS1. The DG-network is generated between the neurotrophins and histamine related target genes in sALS1 and their known pharmacological modulators. Circles and rhombus correspond to target genes and drugs, respectively. Each node (target/drug) is colored along a color gradient on the basis of its degree of connectivity (the number of connections with other nodes) and the edges represent interactions between drugs and targets as well as between two genes. ERBB2 and everolimus were the hub protein and drug, respectively, in the network. Network properties for the 5 most connected "hub nodes" (drugs and targets) are detailed in Table 2.*

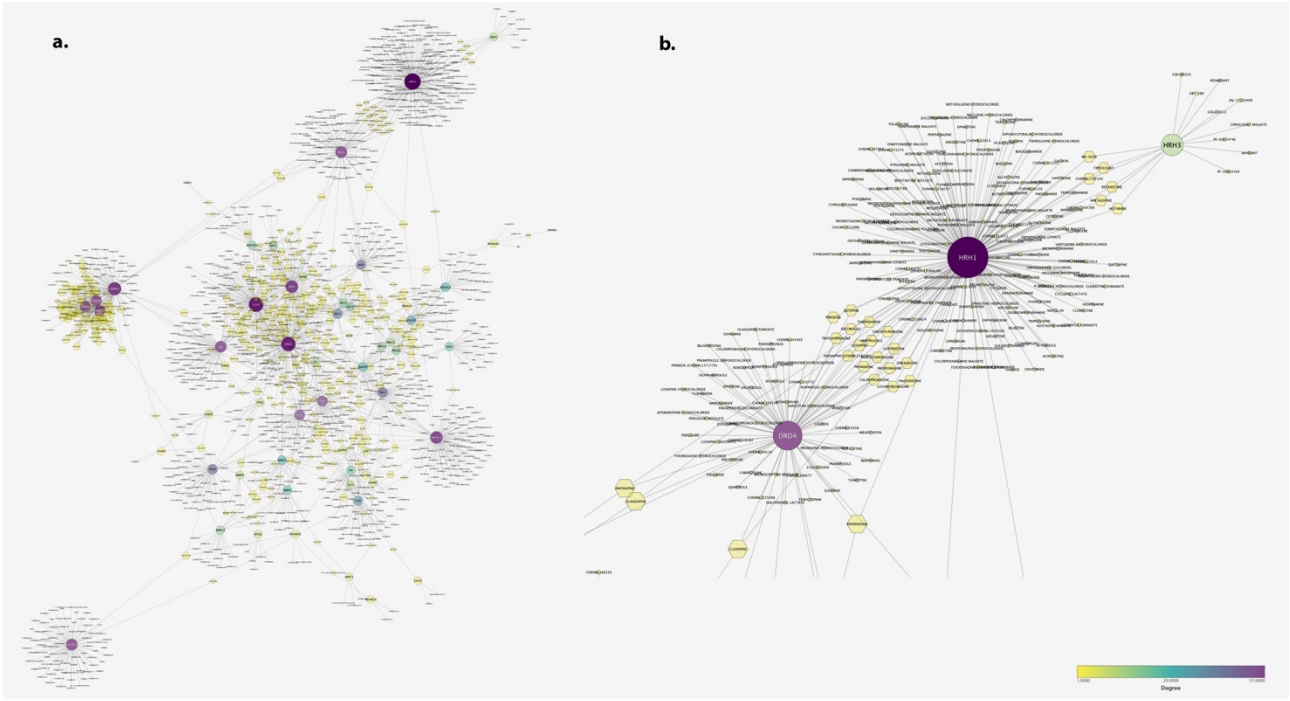


Figure 3*. Drug–Gene network in sALS2. The DG-network (a) is generated between the neurotrophins and histamine related target genes in sALS2 and their known pharmacological modulators. Circles and rhombus correspond to target genes and drugs, respectively. Each node (target/drug) is colored along a color gradient on the basis of its degree of connectivity (the number of connections with other nodes) and the edges represent interactions between drugs and targets as well as between two genes. HRH1 and paclitaxel were the hub protein and drug, respectively, in the network (b). Network properties for the 5 most connected "hub nodes" (drugs and targets) are detailed in Table 2.*

Table 2. The 5 top nodes (Targets/Drugs) of SALS1 and SALS2-related networks based on degree values.

SALS1			
Target	**Degree**	**Drug**	**Degree**
ERBB2	57	EVEROLIMUS	5
TNF	43	DOXORUBICIN	4
BRAF	38	VORINOSTAT	4
MYC	21	CETUXIMAB	4
HRAS	17	TRASTUZUMAB	3

SALS2			
Target	**Degree**	**Drug**	**Degree**
HRH1	186	PACLITAXEL	9
EGFR	146	GSK-690693	8
ERBB2	139	EVEROLIMUS	7
GABRA1	113	DOXORUBICIN	7
BRAF	100	ENZASTAURIN	6

To this regard, lenalidomide is a potent immunomodulatory agent that, by reducing the TNF expression, inactivates downstream effector caspases, extending survival in transgenic mouse models of ALS [138,139]. On the other hand, the complex sALS2-related drug-target network has highlighted several targets involved in inflammatory, apoptotic, and survival gene signaling (**Figure 3 a-b**). In particular, we have observed a significant and selective contribution of GABAergic and histaminergic signaling with *HRH1* as the most interconnected target in the network (**Figure 3 a-b, Table 2**). In the context of a multi-targeted therapy for ALS, an interesting example is represented by clopazine, an HRH1 antagonist that induces changes in GABA release in brain regions and also modulates DRD4, exerting neuroprotective effects in a variety of neurological disorders [140,141] (**Figure 3a-b, Tables 1-2**). Moreover, sALS2-related network analysis has supported the potential repurposing of multiple microtubule-binding anticancer drugs (i.e., paclitaxel, docetaxel) for the genomics-driven therapy in ALS (**Figure 3 a-b, Table 2**). Despite the clinical use of these drugs is often limited by their neurotoxicity, they have shown to restore lost nerve signals in neurodegenerative diseases like Alzheimer's and their safety analogues have shown remarkable neuroprotective properties for motor neurons both in cell culture and in rodents, by stimulating neuronal survival and axonal sprouting[142,143].

CONCLUSIONS

While ALS has been known for over 200 hundred years, we are still in an initial stage for its comprehension, with most research carried out in the last thirty years or so, thanks also to the advent of high-throughput omics data analysis for the identification of molecular alterations, candidate gene drivers and their impact on the outcome of the disease. In the last years, our research group has highlighted for the first time the

existence of a biological and molecular heterogeneity in ALS postmortem cortex samples. Although the use of postmortem brain tissues does not allow to clarify whether investigated signaling cascades are a cause or consequence of the disease process, it represents a vanishingly rare resource for investigating molecular mechanisms, underlying neurological disorders, and providing essential features that cannot be obtained by using other approaches or living patients.

With this in mind, the most important goal for the next decade of ALS research should include the (i) understanding of the origins of motor neuron death as hallmark of the disease; (ii) recognition of the overall contribution of genes, protein targets and mechanisms involved in the pathogenesis and progression of the disease; (iii) validation of new biomarkers to help early and precise diagnosis; (iv) rational and effective patient stratification and system therapy efforts; (v) not least, repurposing or discovering new treatments including drugs, antibodies, stem cell and gene therapies. Overall, as highlighted by the several examples provided in the present work, the potential of network pharmacology and molecular subtyping in ALS may not only provide new important clues to the causes of the disease but, mostly, provide a clear understanding of disease prognosis and progression and thus guide personalized treatments.

ACKNOWLEDGEMENTS

The work from our laboratories described in this work was supported by Italian Ministry for Education, University and Research through grant Flagship Project NanoMAX (grant number B81J13000310005) (C.V.), grant number CTN01_00177_817708 (S.C.), the international PhD programs in *Neuroscience* and *Computer Science* of the University of Catania and moreover by the Italian Ministry of Health through Ricerca Corrente to Fondazione Santa Lucia.

DECLARATIONS OF INTEREST

None

ABBREVIATIONS

ALS, Amyotrophic Lateral Sclerosis; BDNF, Brain-Derived Neurotrophic Factor; CNS, Central Nervous System; CNTF, Ciliary Neurotrophic Factor; CNV, copy number variants; CSF, cerebrospinal fluid; DAO, Diamine oxidase; EGF, epidermal growth factor; ET-1, endothelin-1; G-CSF, Granulocyte-Colony Stimulating Factor; GDNF, Glial-Derived Neurotrophic Factor; HA, Histamine; HDC, Histidine Decarboxylase; FGFs, Fibroblast growth factors; HNMT, Histamine N-methyltransferase; IGF-1, Insulin-like Growth Factor 1; iPSC, induced Pluripotent Stem Cells; NTFs, Neurotrophic Factors; PACAP, pituitary adenylate cyclase-activating polypeptide; sALS, sporadic ALS; SNV, single nucleotide variations; SOD1, Superoxide Dismutase 1; TrkB.T1, Tyrosine Receptor Kinase B.T1; VEGF, Vascular Endothelial Growth Factor.

REFERENCES

1. Arthur, K. C. *et al.* Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat. Commun.* **7**, (2016).
2. Chiò, A., Mazzini, L. & Mora, G. Disease-modifying therapies in amyotrophic lateral sclerosis. *Neuropharmacology* vol. 167 107986 (2020).
3. Mehta, P. *et al.* Prevalence of Amyotrophic Lateral Sclerosis - United States, 2014. *MMWR. Morb. Mortal. Wkly. Rep.* **67**, 216–218 (2018).
4. Chia, R., Chiò, A. & Traynor, B. J. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *The Lancet Neurology* vol. 17 94–102 (2018).
5. Cook, C. & Petrucelli, L. Genetic Convergence Brings Clarity to the Enigmatic Red Line in ALS. *Neuron* vol. 101 1057–1069 (2019).
6. Caballero-Hernandez, D. *et al.* The 'Omics' of Amyotrophic Lateral Sclerosis. *Trends Mol. Med.* **22**, 53–67 (2016).
7. Mitropoulos, K., Katsila, T., Patrinos, G. P. & Pampalakis, G. Multi-Omics for Biomarker Discovery and Target Validation in Biofluids for Amyotrophic Lateral Sclerosis Diagnosis. *Omi. A J. Integr. Biol.* **22**, 52–64 (2018).
8. Maniatis, S. *et al.* Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science* **364**, 89–93 (2019).
9. Nowicka, N., Juranek, J., Juranek, J. K. & Wojtkiewicz, J. Risk Factors and Emerging Therapies in Amyotrophic Lateral Sclerosis. *Int. J. Mol. Sci.* **20**, (2019).
10. Babić Leko, M. *et al.* Molecular Mechanisms of

Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion. *Behav. Neurol.* **2019**, 2909168 (2019).

11. Blokhuis, A. M., Groen, E. J. N., Koppers, M., van den Berg, L. H. & Pasterkamp, R. J. Protein aggregation in amyotrophic lateral sclerosis. *Acta Neuropathol.* **125**, 777–94 (2013).
12. Wilke, C. *et al.* Correlations between serum and CSF pNfH levels in ALS, FTD and controls: a comparison of three analytical approaches. *Clin. Chem. Lab. Med.* **57**, 1556–1564 (2019).
13. Oeckl, P. *et al.* Multicenter validation of CSF neurofilaments as diagnostic biomarkers for ALS. *Amyotroph. Lateral Scler. Front. Degener.* **17**, 404–413 (2016).
14. La Cognata, V. *et al. NeuroArray*: A Customized aCGH for the Analysis of Copy Number Variations in Neurological Disorders. *Curr. Genomics* **19**, 431–443 (2018).
15. Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol. Dis.* **74**, (2015).
16. Morello, G., Spampinato, A. G. & Cavallaro, S. Neuroinflammation and ALS: Transcriptomic Insights into Molecular Disease Mechanisms and Therapeutic Targets. *Mediators Inflamm.* **2017**, (2017).
17. Morello, G., Spampinato, A. G. & Cavallaro, S. Molecular taxonomy of sporadic amyotrophic lateral sclerosis using disease-associated genes. *Front. Neurol.* **8**, (2017).
18. Tam, O. H. *et al.* Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *CellReports* **29**, 1164-1177.e5 (2019).
19. Jones, A. R. *et al.* Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes. *Neurobiol. Aging* **36**, 2006.e1-2006.e9 (2015).
20. Vijayakumar, U. G. *et al.* A systematic review of suggested molecular strata, biomarkers and their tissue sources in ALS. *Front. Neurol.* **10**, 400 (2019).
21. Morello, G. *et al.* Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS. *Sci. Rep.* **9**, 9968 (2019).
22. Morello, G., Spampinato, A. G., Conforti, F. L., D'Agata, V. & Cavallaro, S. Selection and Prioritization of Candidate Drug Targets for Amyotrophic Lateral Sclerosis Through a Meta-Analysis Approach. *J. Mol. Neurosci.* **61**, (2017).
23. Morello, G., Conforti, F. L., Parenti, R., D'Agata, V. & Cavallaro, S. Selection of potential pharmacological targets in ALS based on whole-genome expression profiling. *Curr. Med. Chem.* **22**, (2015).
24. Morello, G. & Cavallaro, S. Transcriptional analysis reveals distinct subtypes in amyotrophic lateral sclerosis: Implications for personalized therapy. *Future Med. Chem.* **7**, (2015).
25. Thorn, C. F., Klein, T. E. & Altman, R. B. PharmGKB: the Pharmacogenomics Knowledge Base. *Methods Mol. Biol.* **1015**, 311–20 (2013).
26. Wishart, D. S. *et al.* DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **36**, D901–D906 (2008).
27. Apolloni, S. *et al.* Histaminergic transmission slows progression of amyotrophic lateral sclerosis. *J. Cachexia. Sarcopenia Muscle* jcsm.12422 (2019) doi:10.1002/jcsm.12422.
28. Apolloni S, Caputi F, Pignataro A, Amadio S, Fabbrizio P, Ammassari-Teule M, V. C. Histamine Is an Inducer of the Heat Shock Response in SOD1-G93A Models of ALS. *Int J Mol Sci* **20**, E3793 (2019).
29. Apolloni, S. *et al.* Histamine regulates the inflammatory profile of SOD1-G93A microglia and the histaminergic system is dysregulated in amyotrophic lateral sclerosis. *Front. Immunol.* **8**, (2017).
30. Maugeri, G. *et al.* Molecular mechanisms involved in the protective effect of pituitary adenylate cyclase-activating polypeptide in an in vitro model of amyotrophic lateral sclerosis. *J. Cell. Physiol.* **234**, (2019).
31. Bonaventura, G. *et al.* PACAP and PAC1R are differentially expressed in motor cortex of amyotrophic lateral sclerosis patients and support survival of iPSC-derived motor neurons. *J. Cell. Physiol.* **233**, 3343–3351 (2018).
32. Gouel, F., Rolland, A.-S., Devedjian, J.-C., Burnouf, T. & Devos, D. Past and Future of Neurotrophic Growth Factors Therapies in ALS: From Single Neurotrophic Growth Factor to Stem Cells and Human Platelet Lysates. *Front. Neurol.* **10**, 835 (2019).
33. Tovar-y-Romo, L. B., Ramírez-Jarquín, U. N., Lazo-Gómez, R. & Tapia, R. Trophic factors as modulators

of motor neuron physiology and survival: implications for ALS therapy. *Front. Cell. Neurosci.* **8**, (2014).

34. Stahl, N. *et al.* Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science (80-.).* **263**, 92–95 (1994).

35. Sendtner, M. *et al.* Ciliary neurotrophic factor prevents degeneration of motor neurons in mouse mutant progressive motor neuronopathy. *Nature* **358**, 502–504 (1992).

36. Laaksovirta, H., Soinila, S., Hukkanen, V., Röyttä, M. & Soilu-Hänninen, M. Serum level of CNTF is elevated in patients with amyotrophic lateral sclerosis and correlates with site of disease onset. *Eur. J. Neurol.* **15**, 355–359 (2008).

37. Giess, R. *et al.* Early Onset of Severe Familial Amyotrophic Lateral Sclerosis with a SOD-1 Mutation: Potential Impact of CNTF as a Candidate Modifier Gene. *Am. J. Hum. Genet.* **70**, 1277–1286 (2002).

38. Gros-Louis, F., Gaspar, C. & Rouleau, G. A. Genetics of familial and sporadic amyotrophic lateral sclerosis. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1762**, 956–972 (2006).

39. Bongioanni, P., Reali, C. & Sogos, V. Ciliary neurotrophic factor (CNTF) for amyotrophic lateral sclerosis or motor neuron disease. *Cochrane Database Syst. Rev.* (2004) doi:10.1002/14651858.CD004302.pub2.

40. Group, A. C. T. S. A double-blind placebo-controlled clinical trial of subcutaneous recombinant human ciliary neurotrophic factor (rHCNTF) in amyotrophic lateral sclerosis. ALS CNTF Treatment Study Group. *Neurology* **46**, 1244–9 (1996).

41. Penn, R. D., Kroin, J. S., York, M. M. & Cedarbaum, J. M. Intrathecal Ciliary Neurotrophic Factor Delivery for Treatment of Amyotrophic Lateral Sclerosis (Phase I Trial). *Neurosurgery* **40**, 94–100 (1997).

42. Henriques, A., Pitzer, C. & Schneider, A. Neurotrophic growth factors for the treatment of amyotrophic lateral sclerosis: where do we stand ? *Front. Neurosci.* **4**, 32–32 (2010).

43. Bathina, S. & Das, U. N. Brain-derived neurotrophic factor and its clinical implications. *Arch. Med. Sci.* **11**, 1164–78 (2015).

44. Yanpallewar, S. U., Barrick, C. A., Buckley, H., Becker, J. & Tessarollo, L. Deletion of the BDNF truncated receptor TrkB.T1 delays disease onset in a mouse model of amyotrophic lateral sclerosis. *PLoS One* **7**, e39946 (2012).

45. Turner, B. A. *et al.* TrkB/BDNF Signaling Regulates Photoreceptor Progenitor Cell Fate Decisions. *Dev. Biol.* **299**, 455 (2006).

46. Carriedo, S. G. *et al.* AMPA exposures induce mitochondrial Ca(2+) overload and ROS generation in spinal motor neurons in vitro. *J. Neurosci.* **20**, 240–50 (2000).

47. Pradhan, J., Noakes, P. G. & Bellingham, M. C. The Role of Altered BDNF/TrkB Signaling in Amyotrophic Lateral Sclerosis. *Front. Cell. Neurosci.* **13**, 368–368 (2019).

48. Sherwood, N. M., Krueckl, S. L. & McRory, J. E. The Origin and Function of the Pituitary Adenylate Cyclase-Activating Polypeptide (PACAP)/Glucagon Superfamily*. *Endocr. Rev.* **21**, 619–670 (2000).

49. ARIMURA, A. *et al.* PACAP Functions as a Neurotrophic Factor. *Ann. N. Y. Acad. Sci.* **739**, 228–243 (1994).

50. Chen, X.-Y., Du, Y.-F. & Chen, L. Neuropeptides Exert Neuroprotective Effects in Alzheimer's Disease. *Front. Mol. Neurosci.* **11**, 493 (2019).

51. Maasz, G. *et al.* Pituitary adenylate cyclase-activating polypeptide (PACAP) has a neuroprotective function in dopamine-based neurodegeneration in rat and snail parkinsonian models. *Dis. Model. Mech.* **10**, 127–139 (2017).

52. Reglodi, D., Kiss, P., Lubics, A. & Tamas, A. Review on the Protective Effects of PACAP in Models of Neurodegenerative Diseases In Vitro and In Vivo. *Curr. Pharm. Des.* **17**, 962–972 (2011).

53. Lee, E. H. & Seo, S. R. Neuroprotective roles of pituitary adenylate cyclase-activating polypeptide in neurodegenerative diseases. *BMB Rep.* **47**, 369–75 (2014).

54. Yang, R. *et al.* Therapeutic potential of PACAP for neurodegenerative diseases. *Cell. Mol. Biol. Lett.* **20**, (2015).

55. Maugeri, G. *et al.* PACAP and VIP Inhibit HIF-1α-Mediated VEGF Expression in a Model of Diabetic Macular Edema. *J. Cell. Physiol.* **232**, 1209–1215 (2017).

56. Tomimatsu, N. & Arakawa, Y. Survival-promoting activity of pituitary adenylate cyclase-activating polypeptide in the presence of phosphodiesterase

inhibitors on rat motoneurons in culture: cAMP-protein kinase A-mediated survival. *J. Neurochem.* **107**, 628–635 (2008).

57. Waschek, J. VIP and PACAP: neuropeptide modulators of CNS inflammation, injury, and repair. *Br. J. Pharmacol.* **169**, 512–523 (2013).

58. Ringer, C. *et al.* PACAP signaling exerts opposing effects on neuroprotection and neuroinflammation during disease progression in the SOD1(G93A) mouse model of amyotrophic lateral sclerosis. *Neurobiol. Dis.* **54**, 32–42 (2013).

59. Weihe, E., Ringer, C. & Schütz, B. Emerging roles for neuropeptides in neuro-immune regulation in amyotrophic lateral sclerosis: Lessons from the superoxide dismutase 1-G93A mouse model. *J. Neuroimmunol.* **275**, 177 (2014).

60. Maugeri, G. *et al.* Molecular mechanisms involved in the protective effect of pituitary adenylate cyclase-activating polypeptide in an in vitro model of amyotrophic lateral sclerosis. *J. Cell. Physiol.* **234**, 5203–5214 (2019).

61. Junier, M. P., Hill, D. F., Costa, M. E., Felder, S. & Ojeda, S. R. Hypothalamic lesions that induce female precocious puberty activate glial expression of the epidermal growth factor receptor gene: differential regulation of alternatively spliced transcripts. *J. Neurosci.* **13**, 703–13 (1993).

62. Ayuso-Sacido, A. *et al.* Activated EGFR signaling increases proliferation, survival, and migration and blocks neuronal differentiation in post-natal neural stem cells. *J. Neurooncol.* **97**, 323–337 (2010).

63. Liu, B. & Neufeld, A. H. Activation of epidermal growth factor receptors in astrocytes: From development to neural injury. *J. Neurosci. Res.* **85**, 3523–3529 (2007).

64. Liu, B., Chen, H., Johns, T. G. & Neufeld, A. H. Epidermal Growth Factor Receptor Activation: An Upstream Signal for Transition of Quiescent Astrocytes into Reactive Astrocytes after Neural Injury. *J. Neurosci.* **26**, 7532–7540 (2006).

65. Li, Z.-W. *et al.* Epidermal growth factor receptor inhibitor ameliorates excessive astrogliosis and improves the regeneration microenvironment and functional recovery in adult rats following spinal cord injury. *J. Neuroinflammation* **11**, 71 (2014).

66. Koprivica, V. *et al.* EGFR Activation Mediates Inhibition of Axon Regeneration by Myelin and Chondroitin Sulfate Proteoglycans. *Science (80-.).* **310**, 106–110 (2005).

67. Yang, Q. *et al.* Blocking epidermal growth factor receptor attenuates reactive astrogliosis through inhibiting cell cycle progression and protects against ischemic brain injury in rats. *J. Neurochem.* **119**, 644–653 (2011).

68. Le Pichon, C. E. *et al.* EGFR Inhibitor Erlotinib Delays Disease Progression but Does Not Extend Survival in the SOD1 Mouse Model of ALS. *PLoS One* **8**, e62342 (2013).

69. Trieu, V. N. & Uckun, F. M. Genistein Is Neuroprotective in Murine Models of Familial Amyotrophic Lateral Sclerosis and Stroke. *Biochem. Biophys. Res. Commun.* **258**, 685–688 (1999).

70. Zhao, Z., Fu, J., Li, S. & Li, Z. Neuroprotective Effects of Genistein in a SOD1-G93A Transgenic Mouse Model of Amyotrophic Lateral Sclerosis. *J. Neuroimmune Pharmacol.* (2019) doi:10.1007/s11481-019-09866-x.

71. Ranno, E. *et al.* Endothelin-1 is over-expressed in amyotrophic lateral sclerosis and induces motor neuron cell death. *Neurobiol. Dis.* **65**, 160–171 (2014).

72. Moody, T. W. *et al.* Endothelin causes transactivation of the EGFR and HER2 in non-small cell lung cancer cells. *Peptides* **90**, 90–99 (2017).

73. Lederer, C. W., Torrisi, A., Pantelidou, M., Santama, N. & Cavallaro, S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics* **8**, 26 (2007).

74. Fomin, V. *et al.* The C9ORF72 Gene, Implicated in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia, Encodes a Protein That Functions in Control of Endothelin and Glutamate Signaling. *Mol. Cell. Biol.* **38**, (2018).

75. Kotni, M. K., Zhao, M. & Wei, D.-Q. Gene expression profiles and protein-protein interaction networks in amyotrophic lateral sclerosis patients with C9orf72 mutation. *Orphanet J. Rare Dis.* **11**, 148 (2016).

76. D'Antoni, S., Ranno, E., Spatuzza, M., Cavallaro, S. & Catania, M. V. Endothelin-1 Induces Degeneration of Cultured Motor Neurons Through a Mechanism Mediated by Nitric Oxide and PI3K/Akt Pathway. *Neurotox. Res.* **32**, 58–70 (2017).

77. Henriques, A., Pitzer, C. & Schneider, A. Neurotrophic growth factors for the treatment of amyotrophic lateral sclerosis: where do we stand ?

Front. Neurosci. **4**, 32 (2010).

78. Cassina, P. *et al.* Astrocyte activation by fibroblast growth factor-1 and motor neuron apoptosis: implications for amyotrophic lateral sclerosis. *J. Neurochem.* **93**, 38–46 (2005).

79. Vargas, M. R. *et al.* Fibroblast Growth Factor-1 Induces Heme Oxygenase-1 via Nuclear Factor Erythroid 2-related Factor 2 (Nrf2) in Spinal Cord Astrocytes. *J. Biol. Chem.* **280**, 25571–25579 (2005).

80. Woodbury, M. E. & Ikezu, T. Fibroblast Growth Factor-2 Signaling in Neurogenesis and Neurodegeneration. *J. Neuroimmune Pharmacol.* **9**, 92–101 (2014).

81. Li, S., Bock, E. & Berezin, V. Neuritogenic and Neuroprotective Properties of Peptide Agonists of the Fibroblast Growth Factor Receptor. *Int. J. Mol. Sci.* **11**, 2291–2305 (2010).

82. Ohta, Y. *et al.* Intrathecal injection of epidermal growth factor and fibroblast growth factor 2 promotes proliferation of neural precursor cells in the spinal cords of mice with mutant human SOD1 gene. *J. Neurosci. Res.* **84**, 980–992 (2006).

83. Lunetta, C. *et al.* Impaired expression of insulin-like growth factor-1 system in skeletal muscle of amyotrophic lateral sclerosis patients. *Muscle Nerve* **45**, 200–208 (2012).

84. Kaspar BK, Lladó J, Sherkat N, R. J. G. F. Retrograde viral delivery of IGF-1 prolongs survival in a mouse ALS model. *Science (80-.).* **301**, 839–842 (2003).

85. Dodge, J. C. *et al.* Delivery of AAV-IGF-1 to the CNS Extends Survival in ALS Mice Through Modification of Aberrant Glial Cell Activity. *Mol. Ther.* **16**, 1056–1064 (2008).

86. Lepore, A. C. *et al.* Intraparenchymal spinal cord delivery of adeno-associated virus IGF-1 is protective in the SOD1G93A model of ALS. *Brain Res.* **1185**, 256–265 (2007).

87. Wang, W. *et al.* Systemic administration of scAAV9-IGF1 extends survival in SOD1G93A ALS mice via inhibiting p38 MAPK and the JNK-mediated apoptosis pathway. *Brain Res. Bull.* **139**, 203–210 (2018).

88. Lin, H. *et al.* Intramuscular Delivery of scAAV9-hIGF1 Prolongs Survival in the hSOD1G93A ALS Mouse Model via Upregulation of D-Amino Acid Oxidase. *Mol. Neurobiol.* **55**, 682–695 (2018).

89. Borasio, G. D. *et al.* A placebo-controlled trial of insulin-like growth factor-I in amyotrophic lateral sclerosis. *Neurology* **51**, 583–586 (1998).

90. Sakowski, S. A., Schuyler, A. D. & Feldman, E. L. Insulin-like growth factor-I for the treatment of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **10**, 63–73 (2009).

91. Storkebaum, E. *et al.* Treatment of motoneuron degeneration by intracerebroventricular delivery of VEGF in a rat model of ALS. *Nat. Neurosci.* **8**, 85–92 (2005).

92. Azzouz, M. *et al.* VEGF delivery with retrogradely transported lentivector prolongs survival in a mouse ALS model. *Nature* **429**, 413–417 (2004).

93. Li, B., Xu, W., Luo, C., Gozal, D. & Liu, R. VEGF-induced activation of the PI3-K/Akt pathway reduces mutant SOD1-mediated motor neuron cell death. *Mol. Brain Res.* **111**, 155–164 (2003).

94. Tolosa, L., Mir, M., Asensio, V. J., Olmos, G. & Lladó, J. Vascular endothelial growth factor protects spinal cord motoneurons against glutamate-induced excitotoxicity via phosphatidylinositol 3-kinase. *J. Neurochem.* **105**, 1080–1090 (2008).

95. Pronto-Laborinho, A. C., Pinto, S. & de Carvalho, M. Roles of vascular endothelial growth factor in amyotrophic lateral sclerosis. *Biomed Res. Int.* **2014**, 947513 (2014).

96. Cudkowicz, M. E. *et al.* Trial of celecoxib in amyotrophic lateral sclerosis. *Ann. Neurol.* **60**, 22–31 (2006).

97. Pitzer, C. *et al.* Granulocyte-colony stimulating factor improves outcome in a mouse model of amyotrophic lateral sclerosis. *Brain* **131**, 3335–3347 (2008).

98. Rando, A. *et al.* Granulocyte Colony-Stimulating Factor Ameliorates Skeletal Muscle Dysfunction in Amyotrophic Lateral Sclerosis Mice and Improves Proliferation of SOD1-G93A Myoblasts in vitro. *Neurodegener. Dis.* **17**, 1–13 (2017).

99. Chitu, V., Gokhan, Ş., Nandi, S., Mehler, M. F. & Stanley, E. R. Emerging Roles for CSF-1 Receptor and its Ligands in the Nervous System. *Trends Neurosci.* **39**, 378–393 (2016).

100. Martínez-Muriana, A. *et al.* CSF1R blockade slows the progression of amyotrophic lateral sclerosis by reducing microgliosis and invasion of macrophages into peripheral nerves. *Sci. Rep.* **6**, 25663 (2016).

101. Kárpáti, A. *et al.* Histamine elicits glutamate

release from cultured astrocytes. *J. Pharmacol. Sci.* **137**, 122–128 (2018).

102. Rodriguez, F. J., Lluch, M., Dot, J., Blanco, I. & Rodriguez-Alvarez, J. Histamine modulation of glutamate release from hippocampal synaptosomes. *Eur. J. Pharmacol.* **323**, 283–286 (1997).

103. Cacabelos, R., Torrellas, C., Fernández-Novoa, L. & Aliev, G. Neuroimmune Crosstalk in CNS Disorders: The Histamine Connection. *Curr. Pharm. Des.* **22**, 819–848 (2016).

104. Wada, H., Inagaki, N., Itowi, N. & Yamatodani, A. Histaminergic neuron system in the brain: Distribution and possible functions. *Brain Res. Bull.* **27**, 367–370 (1991).

105. Fukui, H. *et al.* Histamine H1 Receptor Gene Expression and Drug Action of Antihistamines. in 161–169 (2016). doi:10.1007/164_2016_14.

106. Haas, H. & Panula, P. The role of histamine and the tuberomamillary nucleus in the nervous system. *Nat. Rev. Neurosci.* **4**, 121–130 (2003).

107. Passani, M. B. & Blandina, P. Histamine receptors in the CNS as targets for therapeutic intervention. *Trends Pharmacol. Sci.* **32**, 242–249 (2011).

108. Peter, D. *et al.* Differential expression of two vesicular monoamine transporters. *J. Neurosci.* **15**, 6179–6188 (1995).

109. Hu, W. & Chen, Z. The roles of histamine and its receptor ligands in central nervous system disorders: An update. *Pharmacol. Ther.* **175**, 116–132 (2017).

110. Volonté, C., Parisi, C. & Apolloni, S. New kid on the block: does histamine get along with inflammation in amyotrophic lateral sclerosis? *CNS Neurol. Disord. Drug Targets* (2015) doi:10.2174/1871527314666150225143921.

111. Volonté, C., Apolloni, S. & Sabatelli, M. Histamine beyond its effects on allergy: Potential therapeutic benefits for the treatment of Amyotrophic Lateral Sclerosis (ALS). *Pharmacology and Therapeutics* (2019) doi:10.1016/j.pharmthera.2019.06.006.

112. Volonté, C., Amadio, S., Fabbrizio, P. & Apolloni, S. Functional microglia neurotransmitters in amyotrophic lateral sclerosis. *Seminars in Cell and Developmental Biology* (2019) doi:10.1016/j.semcdb.2019.04.014.

113. Apolloni S, Amadio S, Fabbrizio P, Morello G, Spampinato AG, Latagliata EC, Salvatori I, Proietti D, Ferri A, Madaro L, Puglisi-Allegra S, Cavallaro S, V. C. Histaminergic transmission slows progression of amyotrophic lateral sclerosis. *J. Chachexia, Sarcopenia Muscle* (2019) doi:10.1002/jcsm.12422.

114. Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol. Dis.* (2015) doi:10.1016/j.nbd.2014.12.002.

115. Chen, H., Kankel, M. W., Su, S. C., Han, S. W. S. & Ofengeim, D. Exploring the genetics and non-cell autonomous mechanisms underlying ALS/FTLD. *Cell Death and Differentiation* (2018) doi:10.1038/s41418-018-0060-4.

116. Apolloni, S., Fabbrizio, P., Parisi, C., Amadio, S. & Volonté, C. Clemastine Confers Neuroprotection and Induces an Anti-Inflammatory Phenotype in SOD1G93A Mouse Model of Amyotrophic Lateral Sclerosis. *Mol. Neurobiol.* **53**, 518–531 (2016).

117. Apolloni, S., Fabbrizio, P., Amadio, S. & Volonté, C. Actions of the antihistaminergic clemastine on presymptomatic SOD1-G93A mice ameliorate ALS disease progression. *J. Neuroinflammation* **13**, 191 (2016).

118. Ramsay, R. R., Majekova, M., Medina, M. & Valoti, M. Key Targets for Multi-Target Ligands Designed to Combat Neurodegeneration. *Front. Neurosci.* **10**, 375 (2016).

119. Geldenhuys, W. J. & Van der Schyf, C. J. Designing drugs with multi-target activity: the next step in the treatment of neurodegenerative disorders. *Expert Opin. Drug Discov.* **8**, 115–129 (2013).

120. Benchekroun, M. & Maramai, S. Multitarget-directed ligands for neurodegenerative diseases: real opportunity or blurry mirage? *Future Med. Chem.* **11**, 261–263 (2019).

121. Ramsay, R. R., Popovic-Nikolic, M. R., Nikolic, K., Uliassi, E. & Bolognesi, M. L. A perspective on multi-target drug discovery and design for complex diseases. *Clin. Transl. Med.* **7**, 3 (2018).

122. Bolognesi, M. L. & Cavalli, A. Multitarget Drug Discovery and Polypharmacology. *ChemMedChem* **11**, 1190–1192 (2016).

123. Villaveces, J. M., Koti, P. & Habermann, B. H. Tools for visualization and analysis of molecular networks, pathways, and -omics data. *Adv. Appl. Bioinform. Chem.* **8**, 11–22 (2015).

124. Murakami, Y., Tripathi, L. P. & Prathipati, P.

Network analysis and in silico prediction of protein–protein interactions with applications in drug discovery. *Curr. Opin. Struct. Biol.* **44**, 134–142 (2017).

125. Athanasios, A., Charalampos, V., Vasileios, T. & Ashraf, G. Recent updates on the application of protein-protein interaction network in drug discovery. *Curr. Drug Metab.* **17**, 1–1 (2016).

126. Chakraborty, C., Doss C, G. P., Chen, L. & Zhu, H. Evaluating protein-protein interaction (PPI) networks for diseases pathway, target discovery, and drug-design using 'in silico pharmacology'. *Curr. Protein Pept. Sci.* **15**, 561–71 (2014).

127. Casas, A. I. *et al.* From single drug targets to synergistic network pharmacology in ischemic stroke. *Proc. Natl. Acad. Sci.* **116**, 7129–7136 (2019).

128. Chumakov, I. *et al.* Polytherapy with a combination of three repurposed drugs (PXT3003) down-regulates Pmp22 over-expression and improves myelination, axonal and functional parameters in models of CMT1A neuropathy. *Orphanet J. Rare Dis.* **9**, 201 (2014).

129. Hughes, R. E., Nikolic, K. & Ramsay, R. R. One for All? Hitting Multiple Alzheimer's Disease Targets with One Drug. *Front. Neurosci.* **10**, 177 (2016).

130. Liu, H. *et al.* AlzPlatform: An Alzheimer's Disease Domain-Specific Chemogenomics Knowledgebase for Polypharmacology and Target Identification Research. *J. Chem. Inf. Model.* **54**, 1050–1060 (2014).

131. Prati, F., Uliassi, E. & Bolognesi, M. L. Two diseases, one approach: multitarget drug discovery in Alzheimer's and neglected tropical diseases. *Med. Chem. Commun.* **5**, 853–861 (2014).

132. Vaz, F. N. C. *et al.* Polypharmacy in Alzheimer's disease patients in Brazil: Guidance for pharmaceutical assistance. *F1000Research* **6**, 2068 (2017).

133. Morgan, S., Duguez, S. & Duddy, W. Personalized Medicine and Molecular Interaction Networks in Amyotrophic Lateral Sclerosis (ALS): Current Knowledge. *J. Pers. Med.* **8**, (2018).

134. Cassano, T. *et al.* Early intrathecal infusion of everolimus restores cognitive function and mood in a murine model of Alzheimer's disease. *Exp. Neurol.* **311**, 88–105 (2019).

135. Bové, J., Martínez-Vicente, M. & Vila, M. Fighting neurodegeneration with rapamycin: mechanistic insights. *Nat. Rev. Neurosci.* **12**, 437–452 (2011).

136. Fox, J. H. *et al.* The mTOR kinase inhibitor Everolimus decreases S6 kinase phosphorylation but fails to reduce mutant huntingtin levels in brain and is not neuroprotective in the R6/2 mouse model of Huntington's disease. *Mol. Neurodegener.* **5**, 26 (2010).

137. Talboom, J. S., Velazquez, R. & Oddo, S. The mammalian target of rapamycin at the crossroad between cognitive aging and Alzheimer's disease. *npj Aging Mech. Dis.* **1**, 15008 (2015).

138. Neymotin, A. *et al.* Lenalidomide (Revlimid®) administration at symptom onset is neuroprotective in a mouse model of amyotrophic lateral sclerosis. *Exp. Neurol.* **220**, 191–197 (2009).

139. Kiaei, M. *et al.* Thalidomide and lenalidomide extend survival in a transgenic mouse model of amyotrophic lateral sclerosis. *J. Neurosci.* **26**, 2467–2473 (2006).

140. Factor, S. A. & Friedman, J. H. The emerging role of clozapine in the treatment of improvement disorders. *Movement Disorders* vol. 12 483–496 (1997).

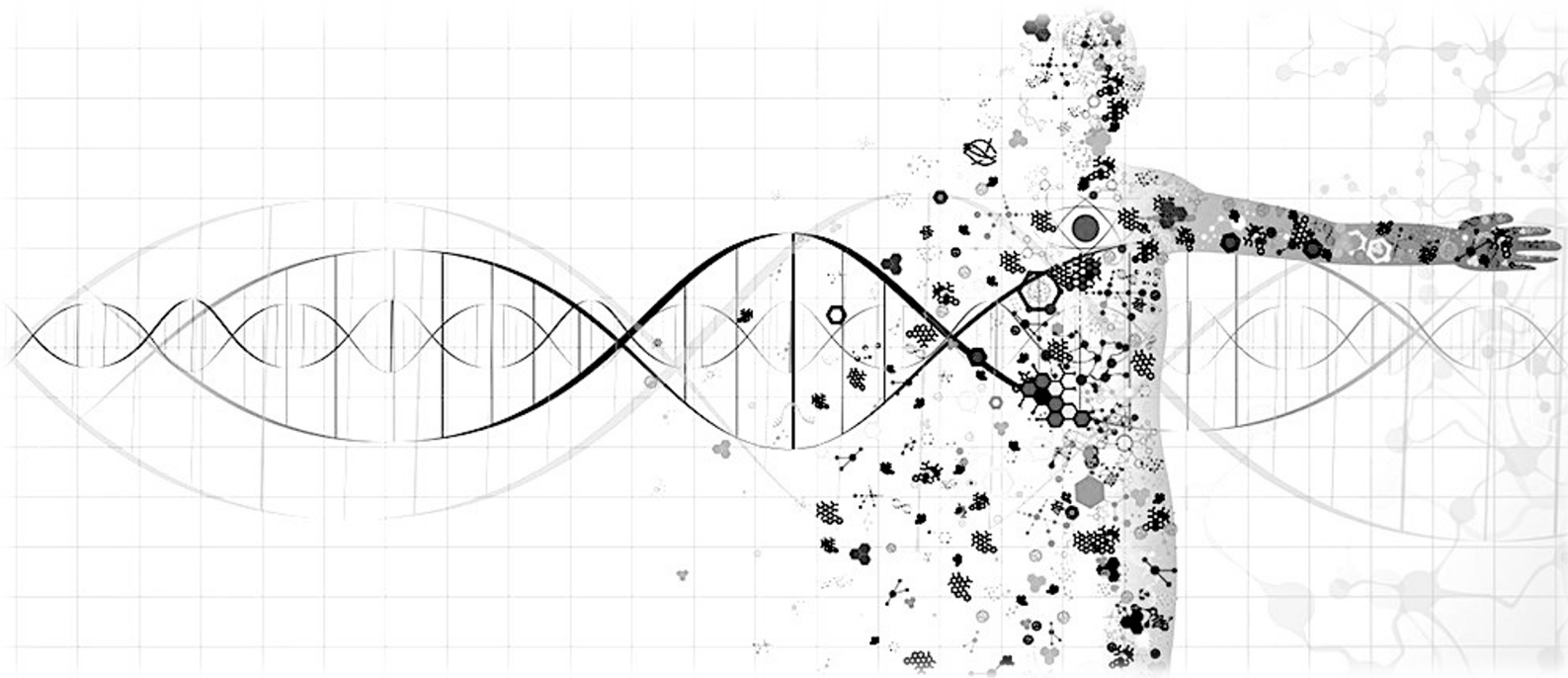
141. Turner, B. J., Rembach, A., Spark, R., Lopes, E. C. & Cheema, S. S. Opposing effects of low and high-dose clozapine on survival of transgenic amyotrophic lateral sclerosis mice. *J. Neurosci. Res.* **74**, 605–613 (2003).

142. Sunyach, C. *et al.* Olesoxime delays muscle denervation, astrogliosis, microglial activation and motoneuron death in an ALS mouse model. *Neuropharmacology* **62**, 2346–2353 (2012).

143. Shemesh, O. A. & Spira, M. E. Rescue of neurons from undergoing hallmark tau-induced Alzheimer's disease cell pathologies by the antimitotic drug paclitaxel. *Neurobiol. Dis.* **43**, 163–175 (2011).

CHAPTER 5

Genomic portrait of a sporadic amyotrophic lateral sclerosis case in a large spinocerebellar ataxia type 1 family



Morello. G., et al. Genomics. (2020) submitted

Genomic portrait of a sporadic amyotrophic lateral sclerosis case in a large spinocerebellar ataxia type 1 family

Giovanna Morello[a,b], Giulia Gentile[a], Rossella Spataro[c], Antonio Gianmaria Spampinato[a,d], Maria Guarnaccia[a], Salvatore Salomone[b], Vincenzo La Bella[c,], Francesca Luisa Conforti[e,#,*] and Sebastiano Cavallaro[a, #,*]

[a]*Institute for Research and Biomedical Innovation (IRIB), Italian National Research Council (CNR), Catania, Italy;* [b]*Department of Biomedical and Biotechnological Sciences, Section of Pharmacology, University of Catania, Catania, Italy;* [c] *ALS Clinical Research Center and Neurochemistry Laboratory, BioNeC, University of Palermo, Italy;* [d]*Department of Mathematics and Computer Science, University of Catania, Catania, Italy;* [e]*Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende (Cosenza), Italy*

[#]Corresponding authors:

Sebastiano Cavallaro, M.D., Ph.D., Institute for Research and Biomedical Innovation (IRIB), CNR, Via Paolo Gaifami, 18, 95125 Catania, Italy. Phone: +39.095.7338111, Fax: +39.095.7338110, Email: sebastiano.cavallaro@cnr.it

Francesca Luisa Conforti, Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Arcavacata di Rende - 87036 Rende (CS), Italy. Phone: +39.0984493155; Fax: +39. 0984493271, Email: francescaluisa.conforti@unical.it

*FLC and SC are co-last authors on this work.

Short title: Genomic portrait of ALS in SCA1 family

ABSTRACT

An expanded repeat region in the spinocerebellar ataxia type 1 (SCA1) gene *ATXN1* increases the risk for amyotrophic lateral sclerosis (ALS), supporting the existence of a relationship between these two neurodegenerative disorders. We recently reported the co-existence, in a large SCA1 family, of a clinically-definite ALS individual bearing an intermediate *ATXN1* expansion and SCA1 patients with a full expansion, some of which manifesting signs of lower motor neuron involvement (here referred as SCA1-MN). In this study, to fully investigate the complex genetic factors and pathogenic mechanisms that may contribute to motor neuron dysfunctions, we employed a systems biology approach that integrated multiple genomic analysis (sequence and copy number variations) of the ALS patient and some SCA1 family members (with or without MN phenotype). Our analysis identified candidate genes/variants that, in addition to or in combination with ATXN1, may contribute to motor neuron degeneration phenotype. Functional and pathway enrichment analysis of these variants revealed both common and distinctive biological processes significantly altered in ALS and SCA1-MN patients. Among alterations detected exclusively in the ALS patient (absent in SCA1), we distinguished likely pathogenic genetic variants in *TAF15* and *C9ORF72*, two ALS-linked genes involved in the regulation of RNA metabolism, similarly to *ATXN1*, suggesting a selective role for this pathway in ALS pathogenesis.

Keywords: Spinocerebellar ataxia, Amyotrophic Lateral Sclerosis, SCA1-MN, NGS, customized aCGH, Multi-omics, Pathway, Network.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder characterized by a severe muscle weakness with atrophy caused by the loss of motor neurons (MNs) in the motor cortex, brainstem, and spinal cord [1]. The genetic and molecular architecture of ALS is complex as the disease is associated with a multitude of causative genes and biological pathways. A limited number of genes, including Cu/Zn superoxide dismutase (*SOD1*), fused in sarcoma/translocated in liposarcoma (*FUS/TLS* or *FUS*), transactive response DNA binding protein 43 kDa (*TDP-43*) and chromosome 9 open reading frame 72 (*C9ORF72*), are responsible for a significant percentage of both familial (FALS) and sporadic ALS (SALS) cases [2,3,4.] In addition to multiple disease-associated genetic variants, there is evidence about putatively associated variants with a moderate or small effect size that may act as predisposing factors or modifiers of the disease phenotype [5–7]. Among the different genetic risk factors for ALS is ataxin-1 (*ATXN1*), a gene involved in transcriptional regulation that normally contains a segment of 22-23 CAG trinucleotide repeats, encoding for a polyglutamine (polyQ) tract. Intermediate-length (~29–33 CAG) repeats are consistently associated with increased risk for ALS, while high poly-Q repeat expansions (>34 CAG) cause spinocerebellar ataxia type 1 (SCA1), an adult-onset autosomal dominant neurodegenerative disease that is characterized by progressive cerebellar degeneration causing loss of motor coordination and balance [8,9].

In our previous work, we described a large SCA1 family, in which one non-SCA1 member, bearing an intermediate *ATXN1* poly-Q expansion, was instead affected by ALS [10] (**Figure 1**). The coexistence of ALS and SCA1 in the same family is very rare, and supports a role for *ATXN1* in the pathogenesis of ALS [10]. Traditional genetic testing for the ALS patient did not identify mutations in the ALS-causing genes *SOD1,*

C9ORF72, FUS, TARDPB, and *ANG*. The phenotypic variability of this family is further complicated by the presence of a "central branch" in the genealogical tree (termed as MN-branch), including SCA1 patients showing early signs and symptoms of lower MN involvement, reinforcing a putative pathogenic link between SCA1 and other degenerative MN diseases, including ALS (**Figure 1**).

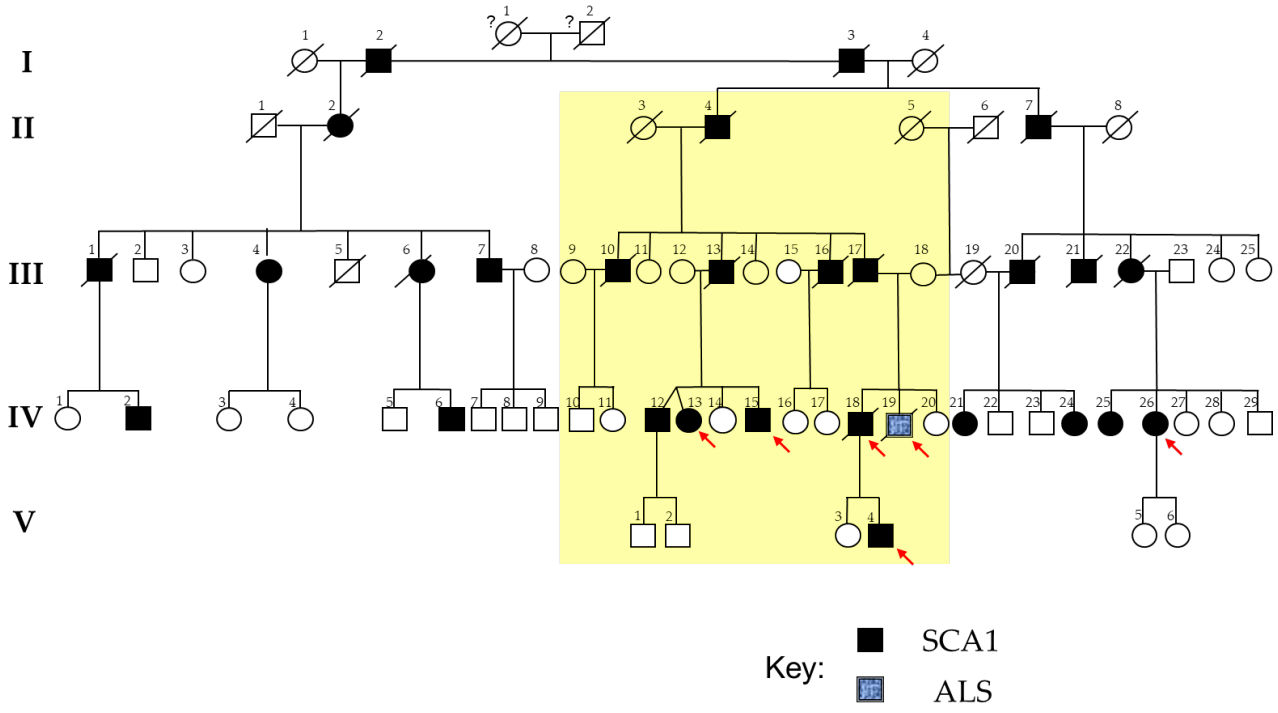


Figure 1. *SCA1-ALS family pedigree. Square indicates male; circle female; slash deceased; black symbols indicate patients affected by SCA1; blue symbol indicates the patient affected by ALS. The red arrows indicate the patients in whom genomic analyses were performed. The "central branch" of the genealogical tree, presenting patients showing ataxic-spastic phenotype with lower MN signs or symptoms is highlighted in yellow.*

In this study, we employed a systems biology approach that integrated multiple genomic data (sequence and copy number variations, CNV) from the ALS patient along with some SCA1 family members (with or without MN phenotype), in order to fully investigate the complex genetic factors and pathogenic mechanisms that may contribute to motor neuron dysfunctions.

RESULTS

Identification of potentially disease-causing sequence variants in ALS associated genes

We utilized a 39-ALS gene panel NGS-based targeted sequencing [11] in the ALS patient and some SCA1 individuals (with or without lower MN signs and symptoms), in order to search for genetic variants that, in addition to or in combination with *ATXN1*, may influence ALS phenotype or susceptibility. Our analysis identified a total of 19 non-synonymous variants in the ALS patient (**Supplementary Table 1**). After applying the filtering criteria (MAF < 5% and *in silico* pathogenicity prediction), 6 gene variants in *ALS2, CCNF, NEFH, NEK1, SETX, TAF15* genes were selected (**Table 1**). Among these, we distinguished three novel coding variants (previously not catalogued in dbSNP), including two missense SNVs in *ALS2* and *NEFH* genes and a frameshift insertion in *TAF15* exon 15 that was selectively found in the ALS patients (absent in SCA1 individuals) (**Table 1**). Of note, the majority of variants (4/6) detected in the ALS proband were shared with SCA1 family memers belonging to the MN-branch but not with the "pure" SCA1 patient (IV-26), suggesting these variants may contribute to induce motor neuron dysfunctions (**Table 1**).

Table 1. Potential pathogenic variants identified in the ALS patient by targeted panel sequencing

										Genotype		
Chr	**Pos**	**Ref**	**Gene**	**Variant class**	**Exon**	**AAChange (refGene)**	**ExAC_ Freq**	**dbSNP**	**Polyphen/ SIFT**	**#IV-19 (ALS)**	**#IV-15 (SCA1-MN)**	**#V-4 (SCA1-MN)**	**#V-26 (SCA1)**
2	202626479	T	ALS2	nonsynonymous SNV	4	p.Ser80Arg (NM_00113574)	.	.	Damaging/ Pathogenic	T/G	T/G	-	-
16	2498978	G	CCNF	nonsynonymous SNV	10	p.Arg406Gln (NM_001323538)	0,00	rs146438723	Damaging/ Pathogenic	G/A	G/A	-	G/A
22	29885908	A	NEFH	nonsynonymous SNV	4	p.Asp760Gly (NM_021076)	.	.	Damaging/ Pathogenic	A/G	A/G	-	-
4	170428901	C	NEK1	nonsynonymous SNV	22	p.Ala626Thr (NM_001199397.1)	0,05	rs33933790	Damaging/ Pathogenic	C/T	-	C/T	-
9	135204010	T	SETX	nonsynonymous SNV	10	p.Lys992Arg (NM_001351527)	0,02	rs61742937	Damaging/ Pathogenic	T/C	-	T/C	-
17	34171599	-	TAF15	frameshift insertion	15	p.Ser433fs (NM_139215)	.	.	Damaging/ Pathogenic	C/CG	-	-	-

Chromosome coordinates are given according to hg19 assembly (UCSC genome browser https://genome.ucsc.edu/).

Identification of copy number variants related to ALS phenotype

In addition to point mutations in ALS driver genes, we searched for numerical chromosomal aberrations in the ALS patient and SCA1 patients of the MN-branch. Using a customized exon-centric high-resolution aCGH platform "*NeuroArray v. 1.0*" [12,13], we identified 16 significant CNVs in the ALS patient, five of which encompass known ALS genes (*VPS54, SCN7A, CHMP2B, LPA, C9ORF72*) (**Table 2)**. According to ACMG guidelines [14], 9 variants were classified as likely pathogenic, 4 as likely benign and 3 of uncertain clinical significance (**Table 2).** Most of the observed CNVs were rare deletions, including 3 novel deletions affecting parts of *C9orf72*, *SCN1A*, and *WRN* genes that did not overlap with any CNVs previously described in the DGV (**Table 2).** The large majority of these alterations were not detected in SCA1 patients, suggesting their specific relevance in ALS etiopathogenesis (**Table 2, Supplementary Tables 2-4)**. Four of the 15 deletions in the ALS patient were also detected in SCA1 family

members belonging to MN-branch, suggesting these alterations may play a role in motor neuron dysfunction (**Table 2, Supplementary Tables 2-4)**. In particular, deletion of the *NSF* gene was identified and validated in the ALS patient as well as in all SCA1 patients of the MN-branch (IV-18, IV-13, IV-15, V-4), while it was absent in patients with a 'pure' SCA1 phenotype (IV-26 and IV-2).

Table 2. CNVs identified in the ALS proband by *NeuroArray* aCGH

Chr	Start	Stop	#Probes	Log2 ratio (Test/Ctrl)	pval	Gene	Common CNV (DGV frequency)	qPCR validation	Clinical interpretation	Detected in SCA1-MN family members (patient code)
1	98,164,881	98,187,177	6	-0.63967	1.54E-10	*DPYD*	Yes (0.005-0.04%)		Likely pathogenic	No
2	64,146,992	64,211,176	25	-0.340603	5.01E-13	***VPS54***	Yes (0.003%)	X	Likely pathogenic	Yes (#18*, #13,#15)
2	166,852,501	166,870,328	14	-0.532297	4.49E-14	*SCN1A*	Not		Likely pathogenic	No
2	166,911,120	166,913,035	5	-0.675035	5.51E-11	*SCN1A*	Yes (0.1%)		Likely pathogenic	No
2	167,328,904	167,334,011	6	-0.754815	4.24E-11	***SCN7A***	Yes (0.005%-1%)	X	Likely pathogenic	No
2	179,536,740	179,540,750	9	-0.573806	1.35E-10	*TTN*	Yes (0.4%)		Likely pathogenic	No
3	87,299,007	89,814,870	10	-0.452305	1.51E-10	***CHMP2B***, *EPHA3*	Yes (0.003-1%)	X	Uncertain clinical significance	No
3	93,772,085	113,652,487	20	-0.285366	4.68E-11	*ARL13B*	Yes (0.006-0.03%)		Uncertain clinical significance	No
6	161,026,135	161,067,305	17	0.477799	2.80E-21	***LPA***	Yes (>70%)		Likely benign	No
7	17,362,101	17,375,411	12	-0.648473	3.43E-22	*AHR*	Yes (0.0034-0.04%)		Likely pathogenic	Yes (#13)
8	30,947,985	30,999,316	23	-0.307664	1.60E-11	*WRN*	Not		Likely pathogenic	No
9	27,558,554	27,573,862	13	-0.399277	2.08E-11	***C9orf72***	Not	X	Likely pathogenic	No
10	70,892,631	70,931,418	15	-0.364046	4.32E-10	*VPS26A*	Yes (>10%)		Likely benign	No
17	44,301,037	44,771,900	16	-0.612643	4.50E-29	*NSF*	Yes	X	Likely benign	Yes (#18, #13,#15)
21	38,791,571	38,865,493	15	-0.358089	8.53E-11	*DYRK1A*	Yes (0.003-5%)		Uncertain clinical significance	Yes (#13)
X	108,902,635	108,906,573	6	-0.732362	5.03E-10	*ACSL4*	Yes (2%)		Likely benign	No

Chromosome coordinates are given according to hg19 assembly (UCSC genome browser https://genome.ucsc.edu/). ALS genes inside CNVs are depicted in bold. Log2 ratio (test/ctrl)= duplications (red), deletions (blue). DGV frequency indicates the population frequency of respective CNV in the Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home). Clinical interpretation was manually assessed and classified into different categories, according to the American College of Medical Genetics and Genomics (ACMG) guidelines for CNVs. Of note, some of alterations reported here were previously described to validate the *NeuroArray* v.1 platform as a genomic profiling assay for ALS [12]. **Sample #18 reported a duplication in the same genomic region.*

Functional and pathway enrichment analysis of CNV-driven genes identified disease-specific molecular signatures

To explore the overall contribution of CNV-driven dysregulated genes to motor neuron dysfunctions, we performed a functional enrichment analysis to characterize their aberrant functions in the ALS patient and SCA1 patients with MN signs. Gene Ontology and pathway enrichment analyses revealed both common and distinctive biological processes and signaling cascades significantly altered in ALS and SCA1-MN patients (**Table 3, Supplementary Table 5**). In particular, the regulation of synaptic transmission and membrane trafficking were overrepresented in all patients, whereas endocytosis, regulation of growth rate and cytoskeleton organization were specifically affected in the ALS patient (**Table 3, Supplementary Table 5**).

Table 3. Functional enrichment analysis of CNV-associated gene sets in ALS and SCA1-MN patients

GO biological processes	ALS	SCA IV-18	SCA IV-13	SCA IV-15
growth				
vacuolar transport				
circadian rhythm				
peptidyl-tyrosine modification, phosphorylation				
regulation of hydrolase activity				
cell morphogenesis				
regulation of endocytosis				
regulation of growth rate				
lysosomal transport				
membrane depolarization during action potential				
regulation of microtubule cytoskeleton organization				
neuronal action potential				
regulation of microtubule-based process				
endosomal transport				
Golgi vesicle transport				
autophagy				
regulation of microtubule cytoskeleton organization				

Pathway name	ALS	SCA IV-18	SCA IV-13	SCA IV-15
Interaction between L1 and Ankyrins				
Phase 0 - rapid depolarisation				
Retrograde transport at the Trans-Golgi-Network				
Muscle contraction				
L1CAM interactions				
Cardiac conduction				
Intra-Golgi and retrograde Golgi-to-ER traffic				
Axon guidance				
Membrane Trafficking				
Fatty acid, triacylglycerol, and ketone body metabolism				
Endocytosis				
Vesicle-mediated transport				

Green boxes represent processes significantly enriched in ALS and SCA1-MN patients; grey bars indicate no significant change.

WES-based mutation profile confirms the functional impact of axon guidance, cell adhesion and immune response in ALS

As both targeted panel sequencing and *NeuroArray* aCGH focuses on a limited set of known disease-associated genes, the ALS patient's specific genomic profile was further investigated by performing WES analysis in order to looking for novel disease causing genes and mechanisms potentially associated with ALS pathogenesis (**Supplementary Table 6)**. Due to the complexity to identify likely damaged genes among the large number of genetic variants discovered by WES and considering that genes do not play independent roles but form biological function and pathway networks through their intricate interactions, we applied a systems biology approach to prioritize genes with variants and investigate their potential functional impact on ALS. Interestingly, this functional analysis confirmed our previous CNV-based molecular characterization of the ALS patient, highlighting *cytoskeleton organization, transmembrane transport, axon guidance and cell adhesion* as the most significantly enriched GO terms and *focal adhesion, extracellular matrix organization and autophagy-lysosome* as the most overrepresented pathways within the set of mutated genes in the ALS patient (**Figure 2a**, **Supplementary Table 7**).

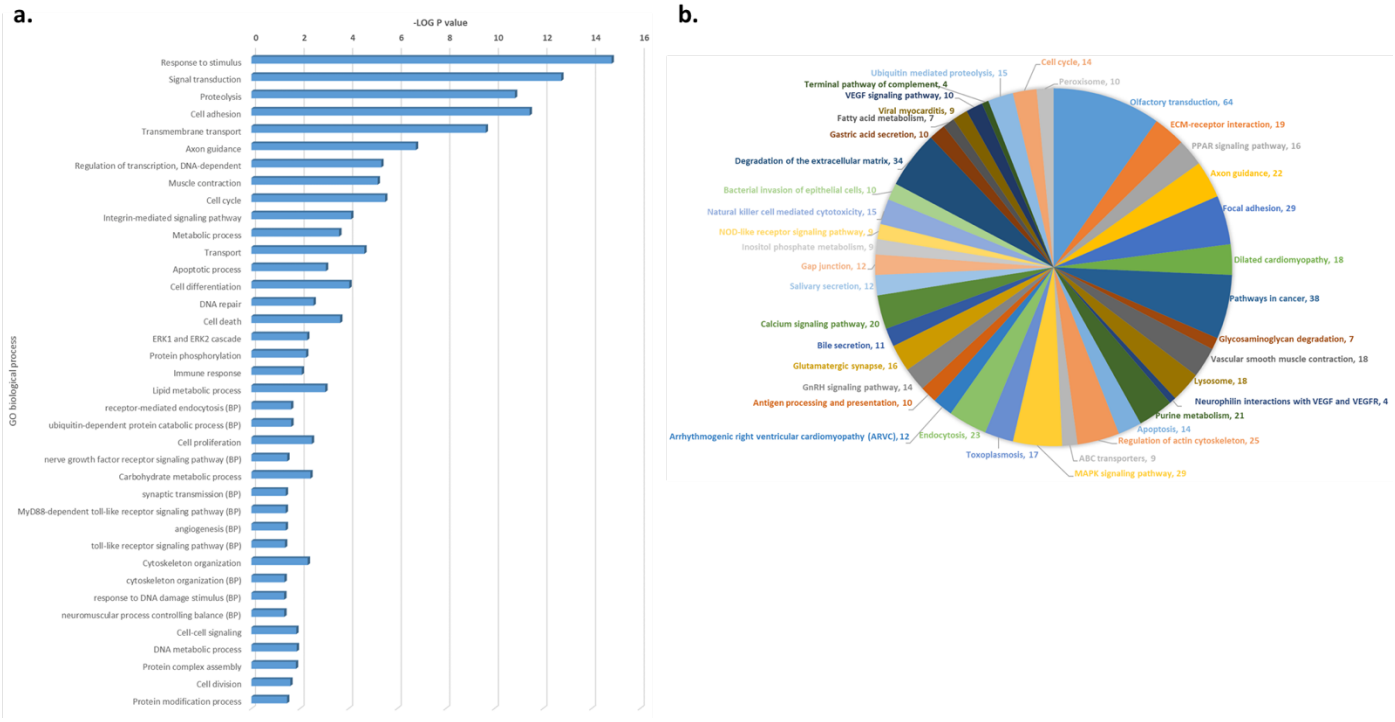


Figure 2. *Functional enrichment analysis of the SNP mutant genes identifiedin the ALS patient by WES. The most representative* ***(a)*** *GO and* ***(b)*** *pathways terms that exhibit statistically significant differences are shown in the graphic (Fisher's exact test, FDR; P < 0.05).*

Genes affected by rare and/or potentially pathogenic variants identified in the ALS patient were mapped to the PPI network in order to further investigate how these genes could jointly confer ALS susceptibility (**Supplementary Table 8, Supplementary Figure 1**). In addition to support the pathogenic role of genes already known to be associated with ALS (i.e., *CHMP2B MAPT, DYNC1H1, ERBB4, GRN, OPTN, SQSTM1, TBK1, TUBA4A, VCP, VEGFA*), our WES-based network analysis identified new potential causal genes, including *RPS27A*, *UBA52*, *UBC* and *UBB* that were identified as the most significant bottleneck proteins connecting different complexes or pathways in the network (**Figure 3a**). Of note, 25 mutated genes in the ALS patient showed a direct interaction with *ATXN1* **(Figure 3b**). Between them, of note, we distinguished some SCA genes (*ATXN3, ATXN7, ATXN2L*) as well as two genes already known to be associated with ALS (*SETX, VCP*) (**Figure 3b**).

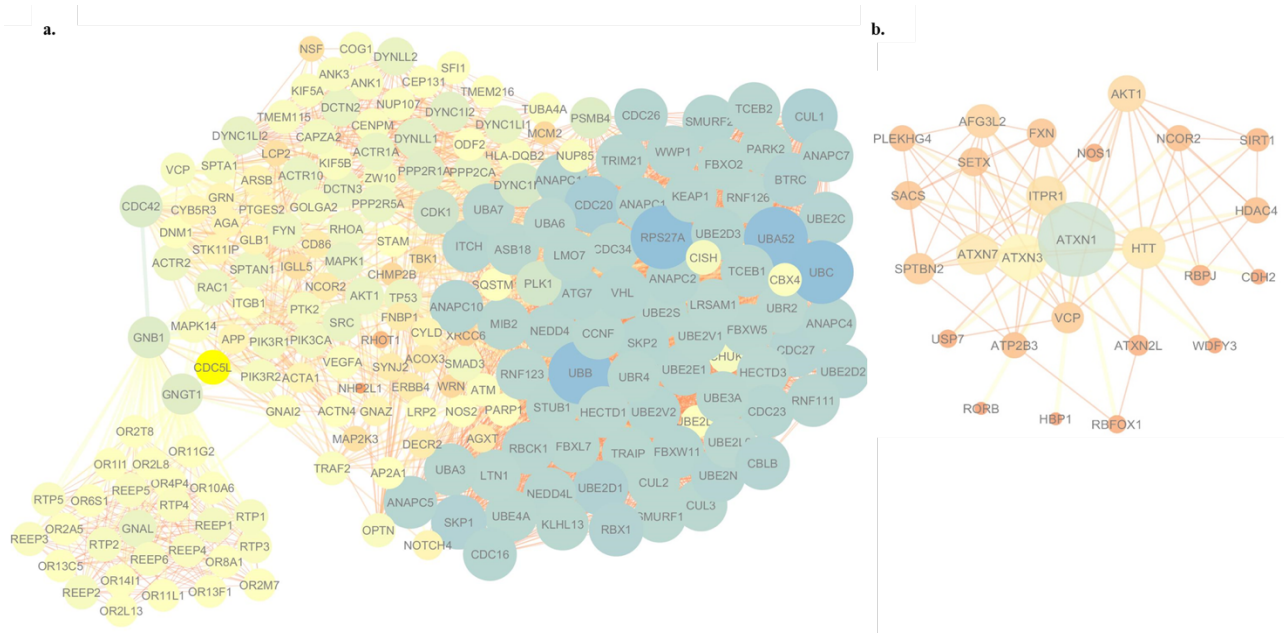


Figure 3. *Integrative network analysis of the genetic variants in the ALS patient. (**a**) Protein-protein interaction (PPI) network of hub genes (degree ≥ 10) affected by rare missense mutations in the ALS patient. The node size was proportional to the degree and the edge width was proportional to the combined score based on the STRING database. Genes with higher degree values have a stronger capacity of modulating adjacent genes. (**b**) ATXN-1 gene-centered subnetwork. The subnetwork includes most of the ATXN-1 -related genes involved in the Notch-specific miRNA-TF regulatory network. The node size was proportional to the degree and the edge width was proportional to the combined score based on the STRING database.*

DISCUSSION

The aim of this study was to provide a comprehensive genomic profiling of a clinically-definite ALS individual, bearing an intermediate *ATXN1* poly-Q expansion and belonging to a large SCA1 family with a "central branch" of SCA1 patients showing early signs and symptoms of lower MN involvement [10]. In particular, we used a set of parallel high-throughput genomic approaches, including (i) a NGS-based targeted mutational analysis focused on a restricted number of ALS genes (exons and flanking regions) and characterized by high coverage, (ii) an high-resolution exon-targeted CNVs analysis of ALS-related genes expanded to those concerning other neurological disorders, and (iii) GO and pathway-based analyses of genetic variants identified in the ALS patient by *NeuroArray* aCGH and WES.

With regard to targeted NGS panel analysis of the ALS patient, we identified six potentially deleterious exonic variants in genes traditionally associated with ALS (**Table 1**). Among these, we distinguished three novel heterozygous exonic variants (c.238A>C in ALS2, c.2279A>G in *NEFH* and c.1296_1297in in *TAF15*) showing consistent results for their pathogenicity between the *in-silico* predictors used (**Table 1, Supplementary Table 1**). In particular, the frameshift mutation in exon 15 of the *TAF15* gene was detected exclusively in the ALS patient (absent in all SCA1 samples), suggesting a selective role for this variant in disease etiopathogenesis (**Table 1**). Similar to other ALS-linked RNA-binding proteins, including ATXN1, TAF15, when mutated, is more aggregation prone *in vitro*, supporting a key role for RNA metabolism defects in ALS and suggests that this class of proteins might contribute very broadly to ALS pathogenesis [15,16].

Using the customized *NeuroArray* aCGH platform we identified a list of chromosomal aberrations in genes previously implicated in a wide range of neurological disorders, that could be responsible for motor neuron degeneration phenotypes observed in this family (**Table 2**). In particular, our analysis identified nine likely pathogenic deletions, some of which encompassing ALS genes, including a novel deletion in *C9ORF72* gene that was detected exclusively in the ALS patient and absent in other SCA1 family members (**Table 2**). Apart from *C9ORF72* hexanucleotide (GGGGCC) repeat expansion, representing the most common genetic cause of both familial and sporadic ALS, decreased *C9ORF72* mRNA levels were found in patient-derived cells and tissue, as well as deletion of this gene locus leads to disruption in endosomal trafficking, synaptic vesicle function, regulation of the actin cytoskeleton and formation of autophagosome, resulting in MN degeneration (19,20). Moreover, previous results indicated that loss-of-function of C9ORF72 not only induces a deleterious effect on

neuron survival by itself, but synergizes with polyQ-Ataxin-2 toxicity to induce motor neuron dysfunction and neuronal cell death [19]. Considering that ataxin-1 has similarities with ataxin-2, our data prompted us to explore the potential contribution of a multiple-hit pathological mechanism in ALS involving C9ORF72 haploinsufficiency and neuronal toxicity resulting from mutant *ATXN1*.

In the context of a systems biology view, the assessment of the functional impact of CNVs detected in ALS and SCA1-MN patients, revealed both specific and common disease-associated molecular signatures (**Table 3**). In particular, the *regulation of synaptic transmission* and *vesicular trafficking to lysosomes* were significantly enriched both in the ALS patient and SCA1-MN family members, suggesting these mechanisms may be involved in motor neuron degenerative processes characterizing these individuals (**Table 3**). To this regard, the deletion of *NSF*, a gene encoding an enzyme that plays an important role in synaptic vesicle release, was found both in ALS and SCA1-MN patients (but absent in patients with 'pure' SCA1 phenotype) (**Table 2**). Although further studies are needed to deepen the role of *NSF* in MN degenerative processes, decreased expression levels of *NSF* were previously found in the motor cortex of SALS patients and its depletion is known to attenuate calcium-dependent delivery of adherens junction proteins to intercellular junctions, thereby producing defects of protein exocytosis and resulting in severe neuronal damage and eventually neuronal death [20–22]. Functional analysis also identified *endosomal transport* and *regulation of cytoskeleton organization* as signaling pathways selectively deregulated in the ALS patient, supporting previous reports indicating dysregulated autophagy and microtubule dynamic instability as biological processes implicated in familial and sporadic insults in ALS [23,24]. Of note, the *NeuroArray*-based genomic signature of the ALS proband was reproduced by functional enrichment analysis of WES data generated

for this patient, further supporting the role of cytoskeletal defects in axons and aberrant transmembrane transport as well as immune response and regulation of ubiquitin-mediated proteolysis in disease pathogenesis (**Figure 2**, **Supplementary Table 7**).

The involvement of dysregulated immune system, proteasome activity and altered cytoskeleton remodeling in driving the ALS phenotype also emerged in the PPI network analysis that highlights the central role of multiple ubiquitin coding genes *(i.e.*, *UBA52, RPS27A, UBC, UBB*) identified as the most interconnected nodes in the network (**Figure 3a**). Substantial contribution by these molecules to ALS pathogenesis was previously described [25] and their abnormal processing and assembling were demonstrated to confer cytotoxic effects, contributing to motor neuronal damage [26–28]. Interestingly, network analysis also revealed a strong interaction between some mutated genes in the ALS patient and *ATXN1* (Figure 3b). Among these, we distinguished SNV and indel variations in some SCA genes (*ATXN3, ATXN7, ATXN2L*) as well as known ALS-linked genes (i.e., *SETX, VCP*), suggesting that genetic variations in these genes may play an additive role in potentiating the neuropathological effects mediated by *ATXN1* repeat expansion (**Figure 3b, Supplementary Table 6**).

Our integrative genomic analysis allowed us to comprehensively investigate genetic variations and molecular mechanisms occurring in ALS and SCA1-MN patients of a large SCA1 family, revealing specific and genomic signatures that may explain progressive degeneration of motor neurons observed in these patients. Overall, our study supports the utility of an individual genomics approach in identifying genetic modifiers for complex disorders characterized by different molecular mechanisms and extreme phenotypic heterogeneity.

MATERIALS AND METHODS

The SCA1 family with a member affected by ALS

Figure 1 shows the large pedigree of the SCA1 family, spanning five generations [10]. The founder could not be identified. All patients belonging the fourth-generation of this large family underwent an accurate clinical evaluation, which confirmed in all the presence of cerebellar ataxia [10]. Interestingly, patients belonging to a branch of the family, all descendants from the patient II-4, showed early signs and symptoms of lower MN involvement (this branch, termed as "*MN-branch*", is highlighted in yellow in **Figure 1**). The ALS patient, bearing an *ATXN1* intermediate expansion, was in this branch too. None of the other fourth-generation patients, belonging to the other branches of this family, showed signs or symptoms of lower MN degeneration, even after years-long disease duration. After a psychological and genetic counselling, we obtained blood from four of the five SCA1 individuals of the MN-branch (i.e., IV-13, IV-15, IV-18, V-4), from the ALS patient (i.e., IV-19) and from a SCA1 patient (IV-26) without lower MN signs and symptoms. All these subjects underwent a comprehensive clinical evaluation.

The ALS member of this family was a 47-year-old worker affected by an upper-limb onset disease while having an *ATXN1* intermediate CAG expansion. His phenotype has been already fully described [10]. Shortly, he had a history of progressive atrophy and weakness of the right hand, which quickly spread to the contralateral limb. Genetic screening for the major ALS-related genes (*SOD1, C9ORF72, FUS, TARDPB, ANG*) was negative. Analysis of *ATXN1* showed an intermediate CAG expansion in both alleles (33/33) with no CAT interruptions. He had a rapidly evolving disease ($\Delta FS > 3.33$). Ten months after diagnosis, he died due to myocardial infarction. IV-18 was the ALS

patient's brother (**Figure 1**). When he was 30 years old, unsteadiness when walking, truncal titubation and slurred speech occurred. No symptoms or signs of lower MN degeneration were reported at onset. Genetic testing revealed an expanded CAG repeat of 33/54 in *ATXN1*. As expected, cerebellar ataxia worsened over time. By the age of 45, a full-blown ALS phenotype occurred, with rapidly progressive distal, and then proximal, muscle atrophy, marked weakness, brisk reflexes in all four limbs, anarthria, dysphagia and atrophy of the tongue with fasciculations. He died at the age of 46 years because of respiratory failure.

The other three SCA1 members of the family, i.e., IV-13, IV-15, and V-4, showed, respectively, a clinical onset at age 37, 40 and 22 with an ataxic-spastic phenotype, and with early lower bulbar-related MN signs or symptoms (i.e., mixed ataxic and flaccid dysarthria, dysphagia and atrophy of the tongue with abundant fasciculations). Genetic testing revealed expanded *ATXN1* CAG repeats of 30/49 for IV-13, 30/49 for IV-15 and 29/57 for V-4, but not mutations in the four ALS-related genes.

In the other SCA1 member of the family (i.e., IV-26), belonging to a branch not showing early lower MN signs or symptoms (**Figure 1**), the clinical onset was at the age of 25 with a prominent ataxic-spastic phenotype and slow progression. At the time of the clinical examination, she was 43-years old, wheel-chair-bound and with a severe cerebellar ataxia. Genetic testing revealed expanded *ATXN1* CAG repeats of 30/45.

All subjects described in this study signed informed consent. All samples were collected and all experiments were performed in accordance with the World MedicalAssociation Declaration of Helsinki. This study was approved by the Ethics Committees of the University of Palermo (document 04/2019, April 29, 2019).

DNA extraction

Blood samples were obtained from all subjects. Genomic DNA was isolated from peripheral blood leukocytes using the salting out method, quantified by using the NanoDrop ND-1000 spectrophotometer and assessed for quality by microcapillary electrophoresis on a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

Targeted Next-Generation Sequencing and Data Processing

A custom targeted NGS-based panel, encompassing 39 ALS-related genes and their 25 bp flanking regions, was used on an Ion Torrent™ Personal Genome Machine™ (PGM) sequencer (ThermoFisher Scientific), as previously described [11]. Briefly, genomic DNA (50 ng) from the ALS proband and 3 SCA1 patients (IV-15, IV-26, V-4), two of which are SCA1-MN affected (IV-15, V-4), was used for library preparation with the Ion AmpliSeq™ Library Kit 2.0. Libraries were then quantified using the Invitrogen™ Qubit™ Fluorometer to determine the dilution factor resulting in a concentration of ~ 100 pM. The template preparation was performed with the Ion PGM™ Hi-Q™ View OT2 Kit on the ION OT2 instrument (ThermoFisher Scientific) using an emulsion polymerase chain reaction (PCR) method. The enriched libraries were purified using the Ion OneTouch™ ES, according to the manufacturer's protocol. The Ion Sphere Particles were loaded onto an Ion 316 chip and sequenced with the Ion PGM™ Hi-Q™ View Sequencing Kit using the ION PGM machine. Sequencing was performed by running 10 samples on a ION 316 chip. For more information about data related to the run, please refer to our previous work [11].

After sequencing, the raw data were processed by the Torrent Suite Software v5.10 (Thermo Fisher Scientific, Inc.) using the standard pipeline parameters. Read alignment and variant identification were carried out with the Torrent Mapping

Program (TMAP) v3.4.1 and Torrent Variant Caller (TVC) v5.0 software. The readings were mapped to the human reference sequence build GRCh37/Hg19 (Genome Reference Consortium Human Build 37, https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/), by limiting to the regions of target genes. The Coverage Analysis plugin was applied to all data and used to assess amplicon coverage for regions of interest. Initial variant calling from the Ion AmpliSeq™ sequencing data was generated using Torrent Suite and Ion Reporter Software with the plug-in 'variant caller' program. To eliminate erroneous base calling, two filtering steps were used to generate the final variant calls. For basic filtering, raw variants were selected using the following parameters: Phred quality score > 20, an average depth of total coverage > 20, each variant coverage > 5, and P<0.0001. The second filter was employed by filtering out possible strand-specific errors (i.e., a mutation was detected only in one, but not both, strands of DNA).

Variant prioritization and assessment of pathogenicity

Potential pathogenic variants were sequentially assessed by following the guidelines for the interpretation of sequence variants proposed by the ACMG [14]. In particular, we filtered only non-synonymous exonic variants that were either absent or had a minor allele frequency (MAF) ≤0.05 in ExAC Browser (http://exac.broadinstitute.org/). Moreover, the functional impact of the missense mutations was estimated using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (https://sift.bii.a-star.edu.sg/), Mutation Taster (http://www.mutationtaster.org/) and PhyloP (http://compgen.cshl.edu/phast/background.php) tools. Retained variants were revealed to be damaging by at least one of the prediction programs.

NeuroArray aCGH processing and data analysis

High-resolution exon-centered analysis of CNVs was done using an 8x60K custom exon-centric *NeuroArray* platform v.1.0 (Agilent Technologies, Santa Clara, CA), tailored to detect single/multi-exon deletions and duplications in a large panel of genes associated to several neurological disorders, including ALS (n=154) and SCA (n=52) [12]. DNA labeling and hybridization on *NeuroArray* were performed according to the manufacturer's protocol (Agilent Technologies, Santa Clara, CA). Briefly, DNA test from the ALS proband and some SCA1-MN affected family members (IV-18, IV-13, IV15), together with a reference of the same sex (Euro Reference, Agilent Technologies, Santa Clara, CA), at the concentration of 500 ng, were double digested with RsaI and AluI for 2 hours at 37°C. After heat inactivation of the enzymes at 65°C for 20 min, each digested sample was labeled by random priming by using the genomic DNA Enzymatic Labelling Kit (Agilent Technologies, Santa Clara, CA) for 2 hours using Cy5-dUTP for patient DNAs and Cy3-dUTP for reference DNAs. Labeled products were column purified by using the SureTag DNA Labeling Kit Purification Columns (Agilent Technologies, Santa Clara, CA). After probe denaturation and pre-annealing with Cot-1 DNA, hybridization was performed at 65°C with rotation for 24 hr. After two washing steps, arrays were scanned at 3 μm resolution using an Agilent G4900DA SureScan Microarray Scanner System, and aCGH image data were processed using Agilent's Feature Extraction software to assess the array spot quality as well as check signal and background intensity statistics in the default setting.

Feature extracted raw data were normalized, analyzed and visualized using Agilent CytoGenomics v. 4.0.3.12 software (Agilent Technologies, Santa Clara, CA, USA). Briefly, after filtering for saturated and non-uniform probes, data were normalized by GC correction with a window size of 2 kb and Diploid Peak

Centralization. The Centralization Normalization Algorithm with a threshold of 6.0 and a bin size of 10 was also used for detecting aberrant regions or regions of constant CNVs. Aberrations were detected by the Aberration Detection Method II algorithm (ADM-2), with a sensitivity threshold of 6.0 and moving an average window of 2 Mb, which permits to identify all aberrant intervals in a given sample with consistently high or low log-ratios based on the statistical score. An aberration filter was applied for identifying copy number alterations; changes were considered as true positive events with a minimum log2 ratio test/control of ± 0.25 and a minimum of 3 consecutive probes. A positive statistical score meant an amplification, while a negative score indicated a deletion.

Human reference sequence hg19 assembly was used to define the genomic coordinates of detected CNVs. To assess the effective relations between the detected CNVs and ALS pathogenesis, we compared identified aberrant regions with those previously associated with ALS via screening of publicly available databases and published literature. Once identified, aberrations were manually assessed and classified into different categories (pathogenic, benign, likely benign, likely pathogenic, uncertain clinical significance), according to the American College of Medical Genetics and Genomics (ACMG) guidelines for CNVs [14]. In addition, all CNVs that are absent both from DGV or that are reported in very low frequency (<1%) were considered as rare.

CNV validation

Ad hoc quantitative real-time polymerase chain reaction (qPCR) assays were performed to validate genomic imbalances detected by the *NeuroArray*. Briefly, we used DNA extracted from peripheral blood samples of 3 patients (IV-13, IV-18, IV-19), assayed by *NeuroArray*, and additional 3 samples including a SCA1-MN patient (V-4) and two

"pure" SCA1 family members (IV-2, IV-26). Primers flanking the putative exonic imbalances were designed using the PrimerBlast tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) as described [41]. RT-qPCR was performed in triplicate using the LightCycler 1.5 (Roche Diagnostics, Germany). Cycling conditions were 95°C for 15 s followed by 40 cycles of 95°C (5 s), 60°C (15 s) and one cycle of 95°C (15 s), 60°C (60 s), 95°C (15 s). The relative quantification was measured using the ΔΔCt method that requires a healthy control sample (diploid) as a calibrator in all amplifications [42]. As calibrator control, we used the same DNA reference hybridized in the *NeuroArray* experiments. A control gene, checked as normal double-copies on NeuroArray, was used as a reference for normalization. 2-ΔΔCt≥ 1.4 or ≤ 0.6 was defined as copy number gain or loss, respectively, whereas 2-ΔΔCt values from 0.8 to 1.2 were considered as normal diploid. PCR products were visualized by agarose gel electrophoresis.

Functional enrichment analysis of the CNV-associated gene sets

To analyze and visualize functional and biological shared- or specific-features with respect to ALS and/or MN phenotypes, genes located in the identified CNV regions, both in ALS patients and its SCA1-MN relatives, were used for ontology and pathway enrichment analyses. In particular, the CNV-associated gene sets for each patient were submitted to the bioinformatics resource ToppGene Suite (https://toppgene.cchmc.org/) and ToppCluster (https://toppcluster.cchmc.org/), which allow to perform a gene list functional enrichment based on Gene Ontology, KEGG, Reactome, and Panther pathway [29,30]. The extent of statistical enrichment for each functional group was determined by applying a Fisher's Exact Test then corrected by Benjamini-Hochberg

False Discovery Rate (FDR) procedure and the number of enriched genes >2 and P<0.05 were selected as cut-off criteria to consider statistically significant differences.

Genomic scale profiling of ALS by whole-exome sequencing

The ALS patient's specific genomic profile was further investigated by performing a pathway enrichment analysis of whole-exome sequencing (WES) data, in order to verify also if a mutational signature extracted from WES data confirmed our previous pathway-based CNV characterization of the disease phenotype. Specifically, genomic DNA of the ALS patient was sequenced using the Ion Chef and Ion S5 Next Generation Sequencing platform (Thermo Fisher Scientific), following the manufacturer's recommended protocol. In brief, 100 ng of DNA was used as starting quantity for library preparation by using the Ion Ampliseq Exome RDY Kit 1 × 8 that permits to obtain a coverage >97% of the Consensus Coding Sequences (CCDS), >90% base on-target, and >90% coverage uniformity. After barcode ligation using an Ion Xpress Barcode Adapters kit (Thermo Fisher Scientific), library samples were purified using Agencourt AMPure XP reagent (Beckman Coulter, Brea, CA) and subsequently inspected for quality by using Bioanalyzer 2100 instrument and DNA High Sensitivity kit (Agilent Technologies). The final sequencing libraries were quantified using a Qubit 4.0 Fluorometer (Thermo Fisher Scientific) and subsequently diluted to a concentration of 100 pM before being clonally amplified on the Ion Chef System using the Ion 540 Chip kit-Chef prior to loading on an Ion 540 Chip with an additional sample for sequencing with the Ion S5 platform (Thermo Fisher Scientific).

Data were processed using Ion Torrent platform-specific pipeline software, Torrent Suite Software v5.10 (Thermo Fisher Scientific, Inc.) using Germ Line - Low Stringency parameters to generate sequence reads, trim adapter sequences, filter, and

remove poor signal-profile reads. The alignment was done against the human reference sequence build GRCh37/Hg19 (Genome Reference Consortium Human Build 37, https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/). Initial variant calling from the Ion AmpliSeq™ sequencing data was generated using Torrent Suite and Ion Reporter Software with a plug-in 'variant caller' program. To eliminate erroneous base calling, two filtering steps were used to generate the final variant calling. For basic filtering, raw variants were selected using the following parameters: Phred quality score > 20, average depth of total coverage > 20, each variant coverage > 5, and P<0.0001. The second filter was employed by filtering out possible strand-specific errors (i.e., a mutation was detected only in one, but not both, strands of DNA).

Exome variants were further filtered for rare non-coding variants, nonsynonymous SNVs, frameshift INDELs, MNVs and SNVs/INDELs affecting stop codons and splice sites. The KEGG pathway enrichment analysis of the genomic regions affected by the remaining variants was carried out using the ToppFun tool (https://toppgene.cchmc.org/enrichment.jsp) to identify the most relevant functional pathways plausibly involved in driving the ALS phenotype. Pathway enrichment analysis was performed by using the human genome as background and applying a corrected Fisher's exact test with a P<0.05 as a threshold to consider statistically significant differences.

Visualization of the protein-protein interaction network

To further investigate the interaction and correlation between genes harboring rare non-coding and/or potentially pathogenic coding variants we constructed an extended protein-protein interaction (PPI) network of their encoding products by using the STRING database and visualized with the Cytoscape v.3.7.1 software. The extended

network was constructed by using the candidate genes as seed molecules and setting a high level of confidence between molecular interactions (high confidence score of at least 0.8) and a maximum number of interactions to 100. In order to identify the "Hub" nodes, a network topology analysis was performed by using the Cytoscape plug-in NetworkAnalyzer based on topological parameters. The relative importance of the genes in each network, meaning their ability to hold together the communicating nodes in a biological network, was determined based on the node centrality measure setting the topological parameter "node degree" ≥10. Nodes with high degree (hub genes) represented genes having important biological functions: the higher the value, the higher the relevance of the gene in connecting regulatory molecules. Likewise, values of edge betweenness were mapped with the edge size: high values of this parameter correspond to large edge size. After removing the nodes with a score of 0, the final PPI network was visualized based on node degree and edge betweenness parameters. Moreover, an additional PPI network was constructed in order to assess possible interactions between *ATXN1* and other SNP- and CNV-driven genes in the ALS proband. This ATXN1-centered network was constructed by setting a moderate level of confidence between molecular interactions (confidence score of at least 0.4) and a maximum number of interactions to 100.

ACKNOWLEDGMENTS

Authors gratefully acknowledge Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo, Francesco Marino for their administrative and technical support.

AUTHOR CONTRIBUTIONS

G.G. performed microarray experiments. G.G. and G.M. analyzed and interpreted microarray data. G.M. and M.G. performed whole-exome sequencing experiments,

analyzed and interpreted data. G.G. performed targeted panel NGS experiments. G.M. and G.G. analyzed and interpreted targeted panel NGS data. A.G.S. performed bioinformatics analyses. G.M. wrote the manuscript. R.S. identified family members, performed clinical evaluations, visits and cared for the patient's recruitment. V.L.B. provided patient samples, clinical and demographic details. S.S., F.L.C. and S.C. participated in revising the manuscripts. F.L.C. conceived the project. F.L.C. and S.C. directed and supervised the project. All authors have read and approved the final version of this manuscript and agreed to be accountable for all aspects of the work. Correspondence and requests for materials should be addressed to the authors at sebastiano.cavallaro@cnr.it and francescaluisa.conforti@unical.it .

CONFLICT OF INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Experiments involving human participants have been approved by the Ethical Committee Palermo 1 (document 04/2019) and have been performed in accordance with the World Medical Association Declaration of Helsinki Helsinki. Informed consent was obtained from all subjects included in the study.

FUNDING

This work was supported by the Italian Ministry of Education, Universities and Research through grant CTN01_00177_817708 and the international Ph.D. programs in *Neuroscience* and *Computer Science* of the University of Catania.

ABBREVIATIONS

ALS = amyotrophic lateral sclerosis; ATXN-1 = ataxin-1; ATXN-2 = ataxin-2; SCA1 = spinocerebellar ataxia type I; SCA2 = spinocerebellar ataxia type II; aCGH =

comparative genomic hybridization array; WES = whole-exome sequencing; NGS = next-generation sequencing; SOD1 = Cu/Zn superoxide dismutase-1; FUS= fused in sarcoma/translocated in liposarcoma; TARDBP/TDP-43= transactive response DNA binding protein 43 kDa; C9ORF72= chromosome 9 open reading frame 72; SNV = single-nucleotide variant; CNV = copy number variant; MN = motor neuron; CSF = cerebrospinal fluid; MRI = magnetic resonance imaging; EMG = electromyography; MUPs = motor unit potentials; PSWs = positive sharp waves; DSF = disease-free survival; Cy5 = cyanine 5; Cy3 = cyanine 3; dUTP= deoxyuridine triphosphate; ADM-2 = aberration detection method II algorithm; CNVD = Copy Number Variation in Disease; ACMG = American College of Medical Genetics; DGV = Database of Genomic Variants; UTR = untranslated region; ncRNA = non coding RNA; LoF = loss-of-function; MAF = minor allele frequency; EVS = Exome Variant Server; HGMD = Human Gene Mutation Database; LOVD = Leiden Open Variation Database; OMIM = Online Mendelian Inheritance in Man; GO = Gene Ontology; BP = biological process; MF = molecular function; CC = Cellular Component; PPI = protein-protein interaction; SCN1A = sodium voltage-gated channel alpha subunit 1; WRN = Werner syndrome ATP-dependent helicase; LPA = Lipoprotein(A); SCN7A = sodium voltage-gated channel alpha subunit 7; EPHA3 = Ephrin type-A receptor 3; EPHA1 == Ephrin type-A receptor 1; EPHA2 = Ephrin type-A receptor 2; EPHA4 == Ephrin type-A receptor 4; EPHA5 = Ephrin type-A receptor 5; EPHA6 = Ephrin type-A receptor 6; EPHB1 == Ephrin type-B receptor 1; EPHX1 = Epoxide Hydrolase 1; CHMP2B = Charged Multivesicular Body Protein 2B; VPS54 = Vacuolar Protein Sorting-Associated Protein 54; TTN = Titin; AHR = Aryl Hydrocarbon Receptor; NSF = N-Ethylmaleimide Sensitive Factor, Vesicle Fusing ATPase; ARL13B = ADP Ribosylation Factor Like GTPase 13B; MNVs = Multiple nucleotide variants; INDEL = Insertion/Deletion; CAPN2 = Calpain 2; DCC =

DCC netrin 1 receptor; LRRK1 = Leucine Rich Repeat Kinase 1; MMRN1 = Multimerin 1; NOS2 = Nitric Oxide Synthase 2; ARSB = Arylsulfatase B; TMEM216 = Transmembrane Protein 216; LRP2 = LDL Receptor Related Protein 2; NOTCH4 = Notch Receptor 4; PCK1 = Phosphoenolpyruvate Carboxykinase 1; TNK1 = Tyrosine Kinase Non Receptor 1; CNTN4 = Contactin 4; DIAPH3 = Diaphanous Related Formin 3; CDH13 = Cadherin 13; DYNC1H1 = Dynein Cytoplasmic 1 Heavy Chain 1; CPZ = Carboxypeptidase Z, NEFH = Neurofilament Heavy; TRPM7 = Transient Receptor Potential Cation Channel Subfamily M Member 7; VPS26A = Vacuolar Protein Sorting-Associated Protein 26, Retromer Complex Component A; KIF1A = Kinesin Family Member 1A; CYP2D6 = Cytochrome P450 Family 2 Subfamily D Member 6; GLE1 = GLE1 RNA Export Mediator; SETX = Senataxin; ALS2 = Alsin Rho Guanine Nucleotide Exchange Factor ALS2; CCNF = Cyclin F; PFN1 = Profilin 1; CEP112 = Centrosomal Protein 112; VAPB = Vesicle-Associated Membrane Protein Associated Protein B And C; AVS=ALS variant server; DPYD = Dihydropyrimidine Dehydrogenase; SNARE = soluble NSF attachment protein receptor; MATR3 = Matrin 3; ERBB4 = Erb-B2 Receptor Tyrosine Kinase 4; NEK1 = NIMA Related Kinase 1; OPTN = Optineurin; ANO2 = Anoctamin 2; CHCHD10 = Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 10; ELAVL4 = Embryonic Lethal, Abnormal Vision, Drosophila (ELAV)-Like 4; GRIN1 = Glutamate Ionotropic Receptor NMDA Type Subunit 1; NMDA = N-methyl-D-aspartate; ZNF280A = Zinc Finger Protein 280A; C22orf46 = Chromosome 22 Open Reading Frame 46; APEH = Acylaminoacyl-Peptide Hydrolase; MST1 = Macrophage Stimulating 1; DCHS2 = Dachsous Cadherin-Related 2; FAT1 = FAT Atypical Cadherin 1; BACE1 = Beta-Secretase 1.

Supplementary Material can be made available in the online version of this article.

REFERENCES

1. Mathis, S., Goizet, C., Soulages, A., Vallat, J.-M. & Masson, G. Le. Genetics of amyotrophic lateral sclerosis: A review. *J. Neurol. Sci.* **399**, 217–226 (2019).
2. Volk, A. E., Weishaupt, J. H., Andersen, P. M., Ludolph, A. C. & Kubisch, C. Current knowledge and recent insights into the genetic basis of amyotrophic lateral sclerosis. *Medizinische Genetik* vol. 30 252–258 (2018).
3. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nature Neuroscience* vol. 17 17–23 (2014).
4. Sabatelli, M. *et al.* New ALS-Related Genes Expand the *Spectrum Paradigm* of Amyotrophic Lateral Sclerosis. *Brain Pathol.* **26**, 266–275 (2016).
5. Keller, M. F. *et al.* Genome-wide analysis of the heritability of amyotrophic lateral sclerosis. *JAMA Neurol.* **71**, 1123–1133 (2014).
6. Gibson, S. B. *et al.* The evolving genetic risk for sporadic ALS. *Neurology* **89**, 226–233 (2017).
7. Krüger, S. *et al.* Rare Variants in Neurodegeneration Associated Genes Revealed by Targeted Panel Sequencing in a German ALS Cohort. *Front. Mol. Neurosci.* **9**, (2016).
8. Conforti, F. L. *et al.* Ataxin-1 and ataxin-2 intermediate-length PolyQ expansions in amyotrophic lateral sclerosis. *Neurology* **79**, 2315–2320 (2012).
9. Rüb, U. *et al.* Spinocerebellar ataxia type 1 (SCA1): New pathoanatomical and clinico-pathological insights. *Neuropathol. Appl. Neurobiol.* **38**, 665–680 (2012).
10. Spataro, R. & La Bella, V. A case of amyotrophic lateral sclerosis with intermediate ATXN-1 CAG repeat expansion in a large family with spinocerebellar ataxia type 1. *J. Neurol.* **261**, 1442–1443 (2014).
11. Ungaro, C. *et al.* ALS and CHARGE syndrome: a clinical and genetic study. *Acta Neurol. Belg.* **118**, 629–635 (2018).
12. La Cognata, V. *et al. NeuroArray* : A Customized aCGH for the Analysis of Copy Number Variations in Neurological Disorders. *Curr. Genomics* **19**, 431–443 (2018).
13. La Cognata, V. *et al.* A customized high-resolution array-comparative genomic hybridization to explore copy number variations in Parkinson's disease. *Neurogenetics* **17**, (2016).
14. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–423 (2015).
15. Yue, S., Serra, H. G., Zoghbi, H. Y. & Orr, H. T. The spinocerebellar ataxia type 1 protein, ataxin-1, has RNA-binding activity that is inversely affected by the length of its polyglutamine tract. *Hum. Mol. Genet.* **10**, 25–30 (2001).
16. Ravanidis, S., Kattan, F. G. & Doxakis, E. Unraveling the pathways to neuronal homeostasis and disease: Mechanistic insights into the role of RNA-binding proteins and associated factors. *International Journal of Molecular Sciences* vol. 19 (2018).
17. Therrien, M., Rouleau, G. A., Dion, P. A. & Parker, J. A. Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in C. elegans. *PLoS One* **8**, e83450 (2013).
18. Babić Leko, M. *et al.* Molecular Mechanisms of Neurodegeneration Related to *C9orf72* Hexanucleotide Repeat Expansion. *Behav. Neurol.* **2019**, 1–18 (2019).
19. Sellier, C. *et al.* Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* **35**, 1276–1297 (2016).
20. Kawahara, Y. & Kwak, S. Excitotoxicity and ALS: What is unique about the AMPA receptors expressed on spinal motor neurons? *Amyotroph. Lateral Scler.* **6**, 131–144 (2005).
21. Lederer, C. W., Torrisi, A., Pantelidou, M., Santama, N. & Cavallaro, S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral

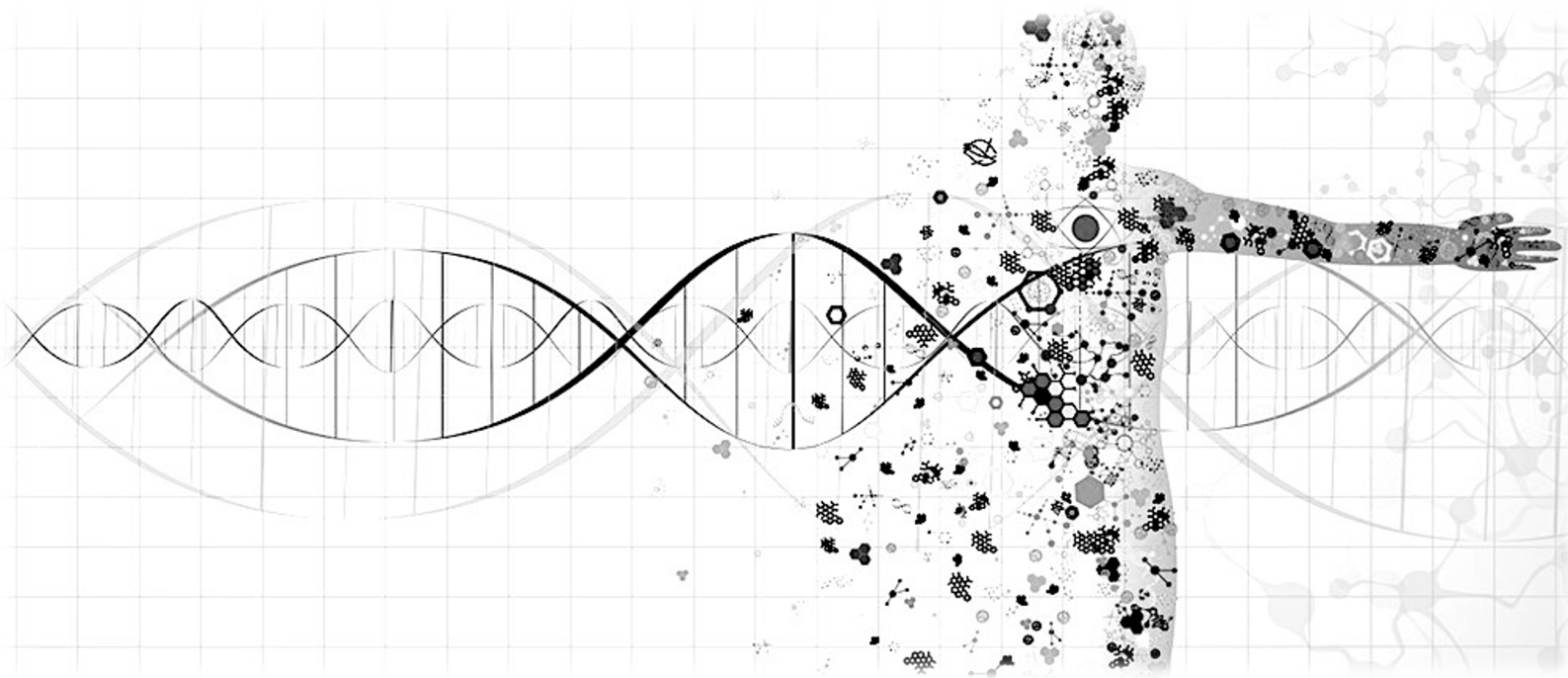
sclerosis. *BMC Genomics* **8**, 26 (2007).

22. Yuan, D., Liu, C. & Hu, B. Dysfunction of Membrane Trafficking Leads to Ischemia-Reperfusion Injury After Transient Cerebral Ischemia. *Transl. Stroke Res.* **9**, 215–222 (2018).
23. Clark, J. A., Yeaman, E. J., Blizzard, C. A., Chuckowree, J. A. & Dickson, T. C. A case for microtubule vulnerability in amyotrophic lateral sclerosis: Altered dynamics during disease. *Frontiers in Cellular Neuroscience* vol. 10 (2016).
24. Burk, K. & Pasterkamp, R. J. Disrupted neuronal trafficking in amyotrophic lateral sclerosis. *Acta Neuropathologica* vol. 137 859–877 (2019).
25. Morello, G. *et al.* Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS. *Sci. Rep.* **9**, 9968 (2019).
26. Bax, M. *et al.* The Ubiquitin Proteasome System Is a Key Regulator of Pluripotent Stem Cell Survival and Motor Neuron Differentiation. *Cells* **8**, (2019).
27. Shahheydari, H. *et al.* Protein Quality Control and the Amyotrophic Lateral Sclerosis/Frontotemporal Dementia Continuum. *Front. Mol. Neurosci.* **10**, 119 (2017).
28. Webster, C. P., Smith, E. F., Shaw, P. J. & De Vos, K. J. Protein Homeostasis in Amyotrophic Lateral Sclerosis: Therapeutic Opportunities? *Front. Mol. Neurosci.* **10**, 123 (2017).
29. Kaimal, V., Bardes, E. E., Tabar, S. C., Jegga, A. G. & Aronow, B. J. ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and networkbased dissection of biological systems. *Nucleic Acids Res.* **38**, (2010).
30. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, (2009)

.

CHAPTER 6

General discussion & Concluding remarks



GENERAL DISCUSSION

Advances in high-throughput genomic, transcriptomic, proteomic, and other diagnostic techniques in the past decades has provided a framework for developing a new, more accurate, and refined "molecular taxonomy" of human diseases which implies the use of molecular data (i.e., gene expression, copy number variants, single nucleotide polymorphisms, and haplotype analysis) to classify patients into distinct subgroups with differing diagnostic, prognostic, or therapeutic implications [22–25]. This new disease classification has profound implications not only providing new insights into studying mechanisms and environmental causes underpinning diseases but also facilitating the development of a more precise diagnosis and individualized treatment for optimal therapeutic efficacy. While incredible successes have been achieved for several multifactorial diseases such as cancer, determination of specific molecular signatures and biomarkers for mechanism-based classification and tailored therapeutic interventions remains an aspirational objective when dealing with multifaceted heterogeneous clinical states such as neurodegenerative disorders, which may reflect combined effects of various genes and their interaction with environmental factors.

Within this context, a perfect example is represented by the complex and heterogeneous nature of ALS, characterized by distinct clinical features and progression patterns, together with a plurality of associated genes. Recently, works from others and our laboratory have started to characterize and explore the hidden biological and molecular heterogeneity of SALS, enabling the identification of etiopathogenic mechanisms and potential therapeutic targets that were not put in evidence by considering SALS pathology as a single entity [12-17]. In particular, by using an unsupervised hierarchical clustering of genome-wide transcriptomic profiles generated from post-mortem motor cortex samples from SALS patients we were able to separate

healthy controls and SALS patients and identify two distinct patient groups (SALS1 and SALS2) depending on the combinations of genes and pathways that were deregulated [14]. Of note, a similar stratification of ALS patients has been reproduced in a recent study conducted by *Tam O.H. et al. (2019)* and other transcriptome profiling-based studies confirmed the existence of distinct molecular subtypes of ALS [15,16,17]. Despite these works lay the foundation for a molecular taxonomy of ALS, very little information is so far available from the single-omic studies, that if considered individually, are insufficient to clarify the intricate disease mechanisms implicated in ALS. Therefore, taking a holistic molecular approach, based on the integration of multiple types of omics data with existing biological knowledge, may offer the opportunity to improve the knowledge of the molecular basis underlying ALS at yet another level of systemic complexity, providing a feasible opportunity to better defining disease subgroups and their molecular drivers in order to develop an efficient and effective personalized diagnostics and patient-guided therapies [22–25].

With these premises in mind, in *Chapter 2* of this PhD thesis we explored the ALS heterogeneity at different levels of omics data, by integrating gene expression profiling with the analysis of genomic structural aberrations occurring in the motor cortex of the same set of SALS samples, in order to investigate whether transcriptional alterations may be related to genomic DNA alterations, and thus represent potential markers for a molecular-based stratification of SALS patients. Taking advantage of the custom-made *NeuroArray* platform, designed to uncover CNVs in clinically relevant genes for ALS and other neurological diseases, we were able to reveale distinct genomic signatures associated with two previously characterized transcriptome-based SALS subgroups, suggesting that genomic and transcriptomic events complement each other in driving disease pathogenesis and further confirming our previous pathway-based

characterization of SALS subtypes. In fact, although partially complementary and convergent, the two SALS patient clusters showed different significantly deregulated processes and mediators. In particular, antigen processing and presentation, and extracellular matrix organization were the most significantly enriched pathways for the CNV-driven genes in SALS1, while the pathways of highest significance in SALS2 were associated with axonal guidance, oxidative stress and inflammatory intracellular signaling cascades. Therefore, a careful monitoring of these signaling cascades may help to better diagnose the specific subtype of ALS.

Emerging molecular heterogeneity of ALS lays the foundations for developing new therapeutic strategies, targeting disease pathogenesis as a complex system rather than at the level of the single protein molecule and that may have greater relevance to distinct sets of patients. In this regard, altered biological pathways emerged from our analyses provided a good number of potential subgroup-specific biomarkers and therapeutic targets, opening the way to the implementation of genomics-based personalized medicine [19–21]. In *Chapter 3 and 4*, we emphasized how the analysis and integration of different omics layers are crucial for the full knowledge of the biological heterogeneity of ALS and for proving a comprehensive characterization of patient-specific molecular signatures that could potentially guide therapeutic decisions. To this purpose, we performed a comprehensive multi-omics characterization of some of the most promising molecular targets and signaling mechanisms, including neurotrophic factors and histamine signaling mediators, that have emerged as dysregulated at different omics levels in specific subgroups of ALS patients in our analysis, focusing on those that have already shown results both in *in vitro* and *in vivo* models of ALS[26–28]. Besides identifying new potential pharmacological targets, our analysis also provided a rational approach for "drug repositioning" for ALS. Under this perspective, many known

drugs that were abandoned at clinical stages because of their low efficacy and/or toxicity might be re-evaluated in light of the emerging molecular taxonomy of ALS patients.

Finally, in *Chapter 5*, we employed a systems biology approach to comprehensively investigate genetic variations and molecular mechanisms occurring in in a large SCA1 family, in which one non-SCA1 member was a clinically-definite ALS individual bearing an intermediate *ATXN1* expansion and SCA1 patients with a full expansion, some of which manifesting signs of lower motor neuron involvement. In particular, we used a set of parallel high-throughput genomic approaches, including (i) a NGS-based targeted mutational analysis focused on a restricted number of ALS genes (exons and flanking regions) and characterized by high coverage, (ii) an high-resolution exon-targeted CNVs analysis of ALS-related genes expanded to those concerning other neurological disorders, and (iii) GO and pathway-based analyses of genetic variants identified in the ALS patient by *NeuroArray* aCGH and WES. Our integrative genomic study allowed us to identify both common and distinctive candidate genes/variants and related biological processes that, in addition to or in combination with *ATXN1*, may contribute to motor neuron degeneration phenotype observed in this family. Among these, we distinguished the involvement of synaptic vesicle trafficking in driving motor neuron degenerative processes characterizing both the ALS patient and SCA1-MN family members, while defects in the regulation of RNA metabolism and ubiquitin system seems to be selectively implicated in the ALS pathogenesis.

Overall, the data presented in this PhD thesis provided evidences regarding the utility of an individual genomics approach in identifying genetic modifiers for a complex disorder like ALS, characterized by different molecular mechanisms and extreme phenotypic heterogeneity, also offering a rational foundation for exploring new candidate therapeutic targets and biomarkers for a genomic-based patient stratification

and personalized diagnostic and treatment approaches. Future more in-depth functional and integrative omics studies will be necessary to verify our findings and explore the impact of candidate genes on the outcome of the disease.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The significant advancement of high-throughput technologies (such as genomics, transcriptomics, proteomics, metabolomics), and diagnostic techniques throughout the past decades outlines a transformative approach to defining, studying and treating complex, multifactorial and heterogenous disorders, such as ALS. Rather than approaching ALS as a single entity, we advocate targeting therapies to distinct "clusters" of patients based on their specific genomic and molecular features.

Although still in its early stages and despite some existing drawbacks, the potential of molecular taxonomy in neurology is clear and the hurdles for the employment of this powerful new tool are rapidly being overcome. The intersection of omics and precision medicine in the future will rely on several components, including advanced bioinformatics, the merger of multi-omics technologies, big health data, and deep machine learning to extract novel insights. Looking into the future, we envision that prodigious advances in genomics, proteomics and biomarker development for ALS and other complex disorders may combine with the standard classification system, bringing us a step closer to the establishment of a precision medicine approach in neurology practice.

GENERAL LIST OF REFERENCES

1. Brown, R. H. & Al-Chalabi, A. Amyotrophic lateral sclerosis. *New England Journal of Medicine* vol. 377 162–172 (2017).
2. Chiò, A. *et al.* Global epidemiology of amyotrophic lateral sclerosis: A systematic review of the published literature. *Neuroepidemiology* vol. 41 118–130 (2013).
3. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nature Neuroscience* vol. 17 17–23 (2014).
4. Mejzini, R. *et al.* ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Frontiers in Neuroscience* vol. 13 1310 (2019).
5. Ryan, M. *et al.* Determining the incidence of familiality in ALS. *Neurol. Genet.* **4**, (2018).
6. Taylor, J. P., Brown, R. H. & Cleveland, D. W. Decoding ALS: From genes to mechanism. *Nature* vol. 539 197–206 (2016).
7. Kirby, J., Al Sultan, A., Waller, R. & Heath, P. The genetics of amyotrophic lateral sclerosis: current insights. *Degener. Neurol. Neuromuscul. Dis.* **6**, 49 (2016).
8. Bhandari, R., Kuhad, A. & Kuhad, A. Edaravone: A new hope for deadly amyotrophic lateral sclerosis. *Drugs of Today* **54**, 349–360 (2018).
9. Dash, R. P., Babu, R. J. & Srinivas, N. R. Two Decades-Long Journey from Riluzole to Edaravone: Revisiting the Clinical Pharmacokinetics of the Only Two Amyotrophic Lateral Sclerosis Therapeutics. *Clinical Pharmacokinetics* vol. 57 1385–1398 (2018).
10. Jaiswal, M. K. Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. *Medicinal Research Reviews* vol. 39 733–748 (2019).
11. Traynor, B. J. *et al.* Clinical features of amyotrophic lateral sclerosis according to the El Escorial and Airlie House diagnostic criteria: A population-based study. *Arch. Neurol.* **57**, 1171–1176 (2000).
12. Morello, G., Spampinato, A. G. & Cavallaro, S. Molecular Taxonomy of Sporadic Amyotrophic Lateral Sclerosis Using Disease-Associated Genes. *Front. Neurol.* **8**, 152 (2017).
13. Morello, G., Spampinato, A. G., Conforti, F. L. & Cavallaro, S. Taxonomy Meets Neurology, the Case of Amyotrophic Lateral Sclerosis. *Front. Neurosci.* **12**, (2018).
14. Aronica, E. *et al.* Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol. Dis.* **74**, (2015).
15. Vijayakumar, U. G. *et al.* A systematic review of suggested molecular strata, biomarkers and their tissue sources in ALS. *Front. Neurol.* **10**, (2019).
16. Tam, O. H. *et al.* Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *CellReports* **29**, 1164-1177.e5 (2019).
17. Jones, A. R. *et al.* Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes. *Neurobiol. Aging* **36**, 2006.e1-2006.e9 (2015).
18. Morello, G., Spampinato, A. G. & Cavallaro, S. Neuroinflammation and ALS: Transcriptomic Insights into Molecular Disease Mechanisms and Therapeutic Targets. *Mediators Inflamm.* **2017**, (2017).
19. Morello, G., Spampinato, A. G., Conforti, F. L., D'Agata, V. & Cavallaro, S. Selection and Prioritization of Candidate Drug Targets for Amyotrophic Lateral Sclerosis Through a Meta-Analysis Approach. *J. Mol. Neurosci.* **61**, (2017).
20. Morello, G. & Cavallaro, S. Transcriptional analysis reveals distinct subtypes in amyotrophic lateral sclerosis: Implications for personalized therapy. *Future Med. Chem.* **7**, (2015).
21. Morello, G., Conforti, F. L., Parenti, R., D'Agata, V. & Cavallaro, S. Selection of potential pharmacological targets in ALS based on whole-genome expression profiling. *Curr. Med. Chem.* **22**, (2015).
22. Mitropoulos, K., Katsila, T., Patrinos, G. P. & Pampalakis, G. Multi-Omics for Biomarker Discovery and Target Validation in Biofluids for Amyotrophic Lateral Sclerosis Diagnosis. *Omi. A J. Integr. Biol.* **22**, 52–64 (2018).
23. Mirza, B. *et al.* Machine learning and integrative analysis of biomedical big data. *Genes* vol. 10 (2019).
24. Yu, X. T. & Zeng, T. Integrative analysis of omics big data. in *Methods in Molecular Biology* vol. 1754 109–135 (Humana Press Inc., 2018).
25. Nguyen, N. D. & Wang, D. Multiview learning for understanding functional multiomics. *PLoS Computational Biology* vol. 16 e1007677 (2020).
26. Maugeri, G. *et al.* Molecular mechanisms involved in the protective effect of pituitary adenylate cyclase-activating polypeptide in an in vitro model of amyotrophic lateral sclerosis. *J. Cell. Physiol.* **234**, (2019).
27. Apolloni, S. *et al.* Histaminergic transmission slows progression of amyotrophic lateral sclerosis. *J. Cachexia. Sarcopenia Muscle* **10**, 872–893 (2019).
28. Apolloni, S. *et al.* Histamine Regulates the Inflammatory Profile of SOD1-G93A Microglia and the Histaminergic System Is Dysregulated in Amyotrophic Lateral Sclerosis. *Front. Immunol.* **8**, 1689 (2017).

LIST OF PUBLICATIONS AND SCIENTIFIC CONTRIBUTIONS

1. "*From multi-omics approaches to precision medicine in Amyotrophic Lateral Sclerosis*". **Morello G**., Salomone S., D'Agata V., Conforti F.L. and Cavallaro S. *Front Neurosci.* (2020). *under review*.

   I.F. (2019): 3.707

2. "*Multi-level omics data to support stratified medicine in neurodegenerative diseases*". La Cognata V.*, **Morello G.*,** Cavallaro S. *Front Neurosci.* (2020). *under review*.

   I.F. (2019): 3.707

3. "*Genomic portrait of a sporadic amyotrophic lateral sclerosis case in a large spinocerebellar ataxia type 1 family*". **Morello G**., Gentile G., Spataro R., Spampinato A.G., Guarnaccia M., Salomone S., La Bella V., Conforti F.L., Cavallaro S. Genomics (2020). *submitted*

   I.F. (2019): 6.205

4. "*Genetic investigation of amyotrophic lateral sclerosis patients in south Italy: a two-decade analysis*". Ungaro C., Sprovieri T., **Morello G**., Perrone B., Spampinato A.G., Simone I.L., Trojsi F., Monsurrò M.R., Spataro R., La Bella V., Andò S., Cavallaro S., Conforti F.L. *Neurobiol Aging* (2020)

   I.F. (2019): 4.347

5. "*Differential Vulnerability of Oculomotor Versus Hypoglossal Nucleus During ALS: Involvement of PACAP*". Maugeri G, D'Amico AG, **Morello G**, Reglodi D, Cavallaro S, D'Agata V. *Front Neurosci*. (2020)

   I.F. (2019): 3.707

6. "*Omics-based exploration and functional validation of neurotrophic factors and histamine as therapeutic targets in ALS*". Volonté C, **Morello G**, Spampinato AG, Amadio S, Apolloni S, D'Agata V, Cavallaro *S. Ageing Res Rev*. (2020).
   I.F. (2019): 10.616
7. "*VDAC1 deletion knocks mitochondria down triggering metabolic rewiring in yeast*" Magrì A, Di Rosa MC, Orlandi I, Guarino F, Reina S, Guarnaccia M, **Morello G**, Spampinato A, Cavallaro S, Messina A, Vai M, De Pinto V. *Cellular and Molecular Life Sciences* (2019)
   I.F. (2018): 7.014
8. "*Medicina Genomica: tecnologie di indagine ad alta risoluzione per indirizzare diagnosi e trattamento delle malattie rare" per il manuale "Come orientarsi nel complesso mondo delle malattie rare: per un approccio pratico alla diagnosi e alla gestione delle malattie rare*". Chapter book. La Cognata V., **Morello G.,** CNR edizioni, ISBN 9788880803034 (2019).
9. "*Integrative multi-omic analysis identify new drivers and pathways in molecularly distinct subtypes of ALS*". **Morello G.,** Guarnaccia M., Spampinato A.G., Salomone S., D'Agata V., Conforti F.C., Aronica E., and Cavallaro S., Scientific Reports, *Scientific Reports*, (2019).
   I.F. (2018): 4.011
10. "*Histaminergic transmission slows progression of amyotrophic lateral sclerosis*". Apolloni S., Amadio S.; Fabbrizio P., **Morello G.,** Spampinato A.G., Latagliata E.C., Salvatori I., Proietti D., Ferri A., Madaro L., Puglisi-Allegra S., Cavallaro S., and Volonté C., *Journal of Cachexia, Sarcopenia and Muscle*. (2019).
    I.F. (2018): 10.754

11. "*Cx26 partial loss causes accelerated presbycusis by redox imbalance and dysregulation of Nfr2 pathway*". Fetoni AR, Zorzi V, Paciello F, Ziraldo G, Peres C, Raspa M, Scavizzi F, Salvatore AM, Crispino G, Tognola G, Gentile G, Spampinato AG, Cuccaro D, Guarnaccia M, **Morello G**., Van Camp G, Fransen E, Brumat M, Girotto G, Paludetti G, Gasparini P, Cavallaro S, Mammano F., *Redox Biology*, (2018).

    I.F. (2018): 7.793.

12. "*Taxonomy meets neurology, the case of amyotrophic lateral sclerosis*". **Morello G.,** Spampinato A.G., Conforti F.C., Cavallaro S., *Frontiers in Neuroscience*, (2018).

    I.F. (2018): 3.648.

13. "*Molecular mechanisms involved in the protective effect of PACAP in an in vitro model of ALS*". Maugeri G., D'Amico A.G., Rasà D.M., Federico C., Saccone S., **Morello G.,** La Cognata V., Cavallaro S. and D'Agata V., *Journal of Cellular Physiology*, (2018).

    I.F. (2017): 3.923.

14. "*NeuroArray: A Customized aCGH for the Analysis of Copy Number Variations in Neurological Disorders*". La Cognata V., **Morello G.,** Gentile G., Cavalcanti F., Cittadella R., Conforti F.L., De Marco E.V., Magariello A., Muglia M., Patitucci A., Spadafora P., D'Agata V., Ruggieri M., Cavallaro S. *Current Genomics* (2018).

    I.F. (2017): 2.172.

15. "*Clinical features and genetic characterization of two dizygotic twins with C9orf72 expansion*". Conforti FL., Tortelli R., **Morello G.,** Capozzo R., Barulli MR., Cavallaro S., Logroscino G. *Neurobiol Aging*. (2018).

    I.F. (2017): 4.454.

16. "*Histamine Regulates the Inflammatory Profile of SOD1-G93A Microglia and the Histaminergic System Is Dysregulated in Amyotrophic Lateral Sclerosis*". Apolloni S., Fabbrizio P., Amadio S.; Napoli G., **Morello G.,** Iemmolo R.; Aronica E., Cavallaro S, Volonte C. *Front Immunol*. (2017).

    I.F. (2017): 5.511. Numero di citazioni: 4 (Fonte: Web of Science)

17. "*Neuroinflammation and ALS: Transcriptomic Insights into Molecular Disease Mechanisms and Therapeutic Targets*". **Morello G**., Spampinato A.G., Cavallaro S. *Mediators of Inflammation* (2017).

    I.F. (2017): 3.549; Numero di citazioni: 4 (Fonte: Google Scholar 2018).

18. "*Molecular taxonomy of sporadic amyotrophic lateral sclerosis using disease–associated genes*". **Morello G**., Spampinato A.G., Cavallaro S. *Front. Neurol.* (2017).

    I.F. (2017): 3.508; Numero di citazioni: 3 (Fonte: WoS 2018 2018).

19. "*Copy Number Variations in Amyotrophic Lateral Sclerosis: Piecing the Mosaic Tiles Together through a Systems Biology Approach*". **Morello G.**, Guarnaccia M., Spampinato A.G., La Cognata V., D'Agata V., Cavallaro S. *Mol Neurobiol*. (2017).

    I.F. (2017): 5.076; Numero di citazioni: 4 (Fonte: Google Scholar 2018).

20. "*Selection and Prioritization of Candidate Drug Targets for Amyotrophic Lateral Sclerosis Through a Meta-Analysis Approach*". **Morello G**, Spampinato AG, Conforti FL, D'Agata V, Cavallaro S. *J Mol Neurosci*. (2017).

    I.F. (2017): 2.454; Numero di citazioni: 5 (Fonte: Google Scholar 2018).

21. "*Copy number variability in Parkinson's disease: assembling the puzzle through a systems biology approach*". La Cognata V., **Morello G.**, D'Agata V. and Cavallaro S. (2017). *Hum Genet* 136(1): 13-37.

    I.F. (2017): 3.390; Numero di citazioni: 9 (Fonte: Google Scholar 2018).

ACKNOWLEDGEMENTS

First of all, I would like to thank my mentor, Dr. Sebastiano Cavallaro, for his constant guidance, support and help over the years, allowing my personal and scientific growth.

A special thanks to supervisor Prof. Salvatore Salomone, for his excellent supervision, constructive comments and discussion and for his continuous support of my PhD study and related research.

I would like to express my gratitude to Prof. Filippo Drago, Prof. Salvatore Salomone and Prof. Claudio Bucolo for giving me the opportunity to take part to this prestigious PhD program that represented an important moment for my scientific growth.

I am also thankful to all the past and present members, colleagues and friends of the Institute for Biomedical Research and Innovation, with whom I shared time, energies, smiles and efforts. A special thanks to Antonio, my "desk" Valentina, Mariella, Rosario and Giulia, who shared with me all the good and bad of my PhD path.

I thank my PhD colleagues and all the members of the PhD Program with which I shared this experience pleasantly, particularly, Giorgia, Giuliana, Agostino, Cristina, for their friendship and emotional support through all these last years.

At least, I owe special gratitude to my father and my whole family, for their continuous and unconditional love and support.