Molecular Genetics and Metabolism Reports 35 (2023) 100974

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism Reports

journal homepage: www.elsevier.com/locate/ymgmr

An *in silico* approach to identify early damage biomarker candidates in metachromatic leukodystrophy

Jessica Gómez [a,*], Laura Artigas [a], Raquel Valls [a], Javier Gervas-Arruga [b]

[a] Anaxomics Biotech SL, Barcelona, Spain
[b] Takeda Development Center Americas, Inc., Lexington, MA, USA

ARTICLE INFO

Keywords:
Metachromatic leukodystrophy
Systems biology
In silico
Modeling
Demyelination
Biomarker

ABSTRACT

Metachromatic leukodystrophy (MLD) is a rare, autosomal recessive lysosomal storage disease. Deficient activity of arylsulfatase A causes sulfatides to accumulate in cells of different tissues, including those in the central and peripheral nervous systems, leading to progressive demyelination and neurodegeneration. Although there is some association between specific arylsulfatase A alleles and disease severity, genotype–phenotype correlations are not fully understood. We aimed to identify biomarker candidates of early tissue damage in MLD using a modeling approach based on systems biology. A review of the literature was performed in an initial disease characterization step, allowing identification of pathophysiological processes involved in MLD and proteins relating to these processes. Three mathematical models were generated to simulate different stages of MLD at the molecular level: an early pro-inflammatory stage model (including only processes considered to be active in the early stages of disease), a pre-demyelination stage model (including additional processes that are active after some disease progression), and a demyelination stage model (in which all pathophysiological processes are active). The models evaluated 3457 proteins of interest, individually and by pairs through data mining techniques, applying five filters to prioritize biomarkers that could differentiate between the models. Sixteen potential biomarkers were identified, including effectors relating to mitochondrial dysfunction, remyelination, and neurodegeneration. The findings were corroborated in a gene expression data set from T lymphocytes of patients with MLD; all candidates formed combinations that were able to distinguish patients with MLD from controls, and all but one candidate distinguished late-infantile MLD from juvenile MLD as part of a combinatorial biomarker pair. In particular, pro-neuregulin-1 appeared as differential on all comparisons (patients with MLD vs controls and within clinical subtypes); casein kinase II subunit alpha was detected as a potential individual marker within clinical subtypes. These findings provide a panel of biomarker candidates suitable for experimental validation and highlight the utility of mathematical models to identify biomarker candidates of early tissue damage in MLD with a high degree of accuracy and sensitivity.

1. Introduction

Metachromatic leukodystrophy (MLD; OMIM 250100 and 249900) is a rare, autosomal recessive lysosomal storage disease caused by pathogenic variants in the arylsulfatase A gene (*ARSA*) or prosaposin gene [1]. Deficient activity of arylsulfatase A (ASA; EC 3.1.6.8) causes sulfatides to accumulate in the cells of different tissues, including those in the central and peripheral nervous systems, leading to progressive demyelination and deterioration of gross motor function [1,2].

MLD is generally classified into three subtypes depending on the age of onset of the first signs and symptoms [3]: late-infantile (age at onset: ≤2.5 years), juvenile (age at onset: 2.5–<16 years), and adult (age at onset: ≥16 years). The juvenile subtype has been further divided into early-juvenile MLD (age at onset: 2.5–<6 years) and late-juvenile MLD (age at onset: 6–<16 years) [4]. Late-infantile MLD accounts for 50–60% of cases and is the most rapidly progressing subtype; death typically

Abbreviations: ARSA, arylsulfatase A gene; ASA, arylsulfatase A; BDNF, brain-derived neurotrophic factor; CK2 alpha, casein kinase II subunit alpha; CSF, cerebrospinal fluid; FABP5, fatty acid-binding protein 5; MLD, metachromatic leukodystrophy; MRI, magnetic resonance imaging; MS, multiple sclerosis; NfL, neurofilament light chain; NRG1, pro-neuregulin-1; PBMC, peripheral blood mononuclear cell; RRMS, relapsing–remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; TPMS, Therapeutic Performance Mapping System.

* Corresponding author at: Clinical Division, Anaxomics Biotech, Carrer de la Diputació 237, 1° 1ª, 08007 Barcelona, Spain.
E-mail address: jessica.gomez@anaxomics.com (J. Gómez).

https://doi.org/10.1016/j.ymgmr.2023.100974
Received 6 October 2022; Received in revised form 6 April 2023; Accepted 7 April 2023
Available online 15 May 2023

occurs in the 5 years from symptom onset [5].

In the absence of specific treatment, the therapeutic approach to date has been based on the treatment of comorbidities generated by the disease. Allogeneic hematopoietic stem cell transplantation has been performed in some patients with MLD, where guidelines exist for determining candidates for the juvenile and adult subtypes [6]. However, the clinical benefit is unclear for patients with late-infantile MLD or those patients who already present with a more advanced stage of the disease [5,7]. Autologous hematopoietic stem cell gene therapy with atidarsagene autotemcel (Libmeldy, Orchard Therapeutics) has been approved in some regions for the treatment of presymptomatic patients with late-infantile or early-juvenile MLD, or early-symptomatic patients with early-juvenile MLD [8,9]. However, successful diagnosis of MLD in presymptomatic patients relies on newborn screening programs, which, despite having been shown to be feasible in real-world scenarios [10,11], still face significant implementation challenges, with timelines spanning up to 12 years in some countries [12]. In practice, although sulfatide accumulation begins at birth and neurological signs of brain tissue damage are evident before symptom onset [13], patients typically meet early developmental milestones but then enter a phase of stagnation before the presentation of motor symptoms [14,15]. Therefore, delays in diagnosis until after symptom onset could restrict treatment options for patients [8,16,17]. There are no approved treatments for late-infantile MLD in symptomatic patients, but intrathecally delivered recombinant human ASA is being investigated for the treatment of symptomatic patients with late-infantile MLD in a phase 2 study (NCT03771898). Therefore, early diagnosis and prediction of likely disease course for MLD are imperative [18].

With advancing therapeutic options, there is an increased need to identify easily accessible biomarkers that reflect the disease progression and monitor the treatment course [19]. Although there is some association between specific *ARSA* alleles and disease severity [1,14,20,21], genotype–phenotype correlations are not fully established. For instance, although ASA values below 1% of residual activity and biallelic protein-truncating *ARSA* variants are highly predictive of an early disease onset and rapid progression, high residual enzyme activity does not exclude an early onset, and *ARSA* variants with the same residual enzyme activity may have different effects [21]. Current biomarkers for MLD include white matter hyperintensities on T2-weighted magnetic resonance imaging (MRI) [22] and urine sulfatide levels [23]. These are used diagnostically [5], but their utility as markers of disease progression or phenotype is less clear: reports of T2-pseudonormalization complicate the use of MRI to monitor disease progression [24], and there is little evidence to suggest that urine sulfatide levels can distinguish between MLD subtypes [25]. Cerebrospinal fluid (CSF) sulfatide levels have also failed to show correlations with clinical signs of decreasing motor function [26]. Levels of cytokines in the CSF and urine *N*-acetylaspartate levels have also been shown to differentiate between patients with MLD and controls [19,27], but their utility for differentiating subtypes is uncertain given the small sample sizes employed in these studies. High levels of neurofilament light chain (NfL) at diagnosis have been associated with rapid disease progression in late-infantile and early-juvenile MLD [28]; however, the ability of NfL to predict disease course early is unknown. In addition, disease-specific scales for monitoring motor function decline in MLD, such as the Gross Motor Function Classification in MLD, do not typically allow for monitoring of more subtle changes in disease progression within their predefined levels; moreover, they are based on the ability to walk, so they are only applicable from the age of 18 months [29]. This leaves a gap in the ability to monitor the disease before symptom onset, or to monitor the very early stages of disease. Together, these factors present challenges for the early prediction of likely disease course and, consequently, for appropriate disease management.

Systems biology methods are being implemented to help to understand the molecular identifiers of certain disease states and to aid drug discovery [30–33]. The Therapeutic Performance Mapping System (TPMS) technology (Anaxomics Biotech, Barcelona, Spain) [34,35] uses pattern recognition and artificial intelligence techniques to integrate available biological, pharmacological, and medical knowledge to create mathematical models that simulate disease phenotypes *in silico* [32]. The models build a complex network of proteins and their known biochemical relationships to map how changes in the system (e.g., a disease, or treatment with a drug) can result in signal alterations, protein activity changes, and differential outcomes [34,35]. Therefore, *in silico* modeling offers an approach that may help to identify new protein biomarker candidates indicative of early tissue damage in MLD. TPMS has been used in both biomarker [30,33] and drug discovery [31,32] in a number of disease areas, with some validation in clinical settings [33,36].

We aimed to create systems biology-based MLD models to allow for the proposal of a panel of potential biomarkers that may assist in the prediction of early tissue damage in MLD and subsequent follow-up. These biomarkers would be considered candidates for future clinical validation in presymptomatic patients or in different clinical subtypes.

2. Material and methods

TPMS is a tool that creates mathematical models of the protein pathways underlying a drug mechanism of action or a disease process to explain a clinical outcome or phenotype [34,35]. Models were built using TPMS technology for different pathophysiological stages of MLD, associating demyelination processes with the most advanced model stages.

2.1. Disease characterization of MLD

Scientific literature searches on the molecular pathogenesis and pathophysiology of MLD were performed to identify the main pathophysiological processes involved during tissue damage in MLD (Supplementary Table S1). Publications were also reviewed to identify protein or gene candidates with activity (or lack thereof) functionally associated with a particular pathophysiological MLD process; these protein or gene candidates were termed ‘condition effectors’.

A data compilation step was attempted whereby gene expression data from patients with MLD, patients with other leukodystrophies with a similar age at symptom onset (<30 months; Supplementary Table S2), and patients with multiple sclerosis (MS) were compiled. One data set from patients with MLD and four data sets from patients with MS were found through a search of publicly available repositories [37–39]. No data on patients with other leukodystrophies were available. However, information was insufficient to define each disease stage at the molecular level. Therefore, these five data sets from patients with MLD and patients with MS were instead used later in the process to corroborate the biomarker candidates identified through the mathematical models.

2.2. Generation of mathematical models

After disease characterization, a human protein–protein interaction network was created to generate mathematical models [35]. Then, a selection of known pathophysiological input–output signals (namely drug–indication and drug–adverse drug reaction relationships) were collated as a training set and used to train the models (Supplementary Table S3). The training set was based on a collection of different biological and clinical databases, which provided the information for the input–output relationships. The biological effectors database was used to extract information that related biological processes to their molecular effectors. Information gathered during the molecular disease characterization step through scientific literature searches was also added to the training set.

TPMS was then used to model pathophysiological input-output relationships, considering their protein-based definition. TPMS applied a multilayer perceptron-like and sampling strategy to find all biologically plausible solutions and to prioritize those more relevant mathematically, generating a universe of possible mechanistic solutions (called virtual patients). TPMS generated a population of 1500 virtual patients with

MLD. The top 20 MLD effectors that trigger the modulation of the rest of the MLD effectors were labeled as input, and all identified effectors of MLD (except those from the 'deficiency of ASA' process) as output. Using information obtained during disease characterization, three stages of early pathological tissue changes in MLD were defined by whether processes involved in MLD were activated or inhibited. Three models were generated to represent these three stages: early pro-inflammatory, pre-demyelination, and demyelination (Fig. 1). The predicted protein activity was used to obtain the TSignal (the average signal arriving at the protein effectors) [35], a model-derived measure of protein set definitions (e.g., MLD pathophysiological processes). This TSignal was used to evaluate the models for each MLD process for each of the 1500 virtual patients, providing estimates of interindividual and inter-process variability. The final models had to be able to reproduce a TSignal >0 between input and output, and to reach correctly half of the effector proteins that made up the output. The population of virtual patients was divided into quartiles according to the TSignal obtained for each MLD process. This division was used to set activation and inhibition thresholds for each process, with activated processes defined as those with a TSignal in the top quartile, and inhibited processes defined as those with a signal in the bottom quartile. After setting the activation and inhibition thresholds for each process of interest, a compendium of 250 mathematical solutions was generated for each stage model using the same input–output settings but restricting the activated/inhibited processes to be within the TSignals thresholds. These models provided a view of optimized paths between input and output protein sets in each stage model.

The mathematical models were not generated on specific data on clinical subtypes or disease over time, but represent a theoretical framework of MLD based on the available literature output from the disease characterization step. As such, they do not simulate different phases of the disease over time or the separate clinical subtypes of MLD, but rather the order of appearance of each group of molecular pathophysiological processes involved during early (presymptomatic) tissue damage in MLD as defined in the literature.

2.3. Identification of biomarker candidates

The protein activity levels predicted by the models were explored using data mining techniques, as previously described [35], to identify biomarker candidates of early tissue damage. The following comparisons were made: early pro-inflammatory stage versus pre-demyelination stage; pre-demyelination stage versus demyelination stage; and early pro-inflammatory stage versus demyelination stage. Potential classifiers based on one or two proteins were evaluated.

The resulting biomarker candidates generated from the mathematical models were then narrowed based on four filters, taking into account:

1. statistical significance (according to cross-validated p values ≤ 0.05 by the 10 K-fold) to discern the generated mathematical model solutions with cross-validated accuracy >70% (filter 1)
2. cross-validated sensitivity ≥70%, or cross-validated sensitivity ≥65%, and cross-validated specificity ≥65% (filter 2)
3. biological significance; candidates with no biological significance were discarded (filter 3)
4. ease of measurement; only those located on the cell membrane, secreted in the blood and/or urine, or previously detected as potential markers in peripheral blood mononuclear cells (PBMCs) for diseases involving the brain [40–42] were retained (filter 4).

A final filter was added after the biomarker candidate corroboration process was conducted (see Section 2.4). In particular, biomarker candidates whose expression levels showed statistically significant differences in gene expression data from patients were retained.

2.4. Biomarker candidate corroboration

A gene expression data set from T lymphocytes of patients with MLD from the National Center for Biotechnology Information (NCBI) Gene

Fig. 1. Progressive activation of early pathological processes identified in the disease characterization step, and how these processes relate to the three mathematical models.
ASA, arylsulfatase A; ER, endoplasmic reticulum.

Expression Omnibus [43] (accession number GSE23350; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23350; Supplementary Table S4) was used in this study to corroborate the model-derived candidates. This data set contained T-lymphocyte samples from 24 patients with MLD; details of their clinical characteristics can be found in Supplementary Table S5. Given that MLD is a rare genetic disease, there are limited data upon which to build mathematical models and to corroborate any potential classifiers. As such, we conducted a complementary corroboration analysis on data from patients with MS, a demyelinating condition that partially overlaps in disease pathology with MLD [44]. In total, four data sets from patients with MS were used: three of these were from the ArrayExpress functional genomics database [45,46] (https://www.ebi.ac.uk/arrayexpress/; E-MTAB-2374, E-MTAB-4890, E-MTAB-5151; Supplementary Table S4) and one used a gene expression data set from the NCBI Gene Expression Omnibus [47] (accession number GSE136411; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136411; Supplementary Table S4).

Data sets from patients with MS and the data set from patients with MLD were analyzed independently and used to: 1) determine whether the biomarker candidates identified through the mathematical models were differentially expressed in the patients' data sets; 2) check if the biomarker candidates could also be detected as both individual biomarkers and combinatorial pairs by applying data mining techniques to the patients' data; and 3) identify new potential candidates that were not highlighted in the analyses of the models after applying the prioritization filters and that, combined with the model-derived candidates, had a good classification potential. Patients with MLD were compared with controls, healthy individuals matched for age and sex. Then, patients with late-infantile MLD were compared with two other patient groups: the rest of the MLD population and patients with juvenile MLD. Patients with relapsing–remitting MS (RRMS) were compared with patients with secondary progressive MS (SPMS).

We also used our models to evaluate the behavior of NfL, a protein that has previously been explored as a potential biomarker in MLD.

A schematic of the machine learning process is shown in Fig. 2.

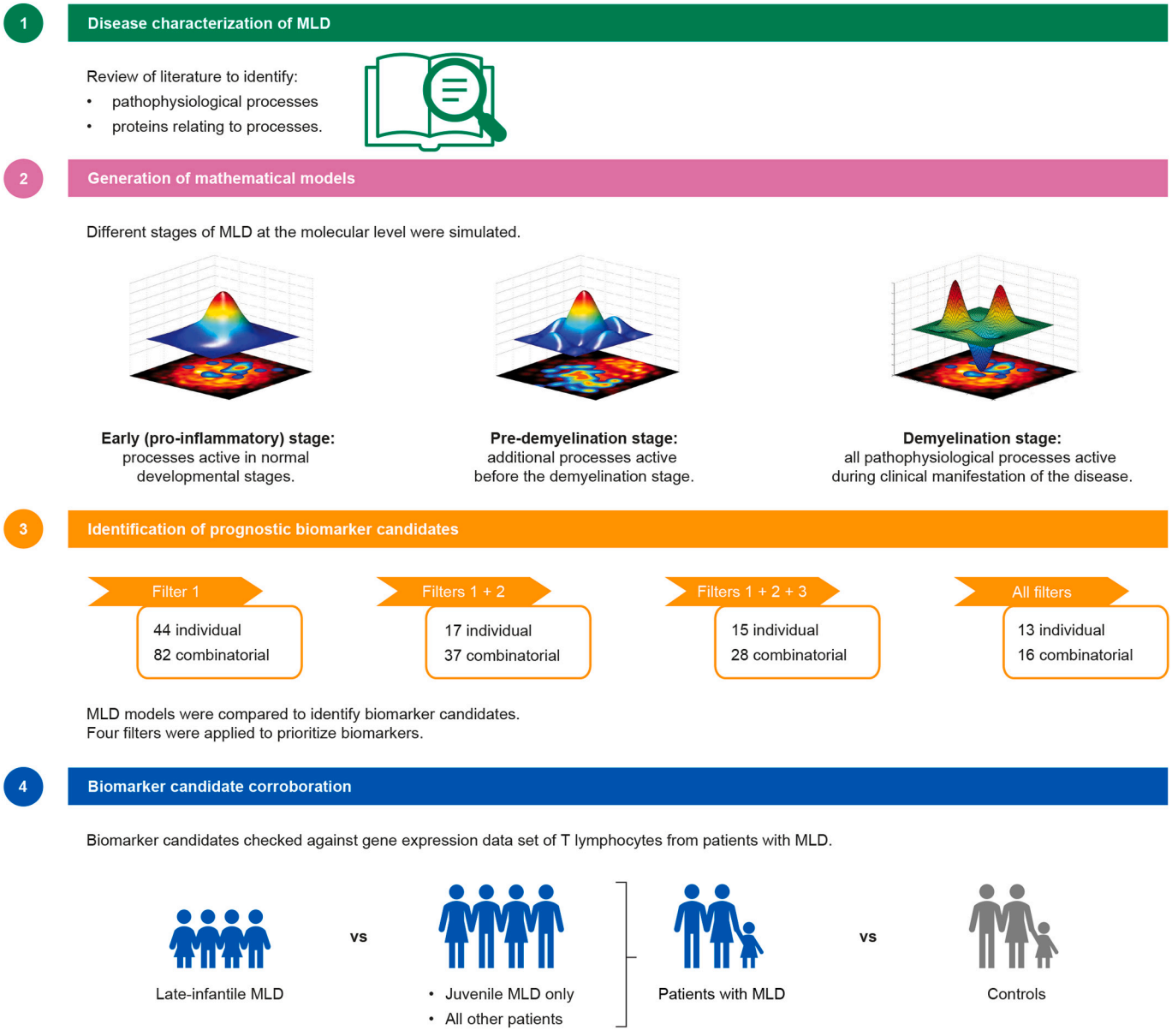

Fig. 2. Steps involved in the identification of biomarker candidates for poor prognosis in MLD.
MLD, metachromatic leukodystrophy.

3. Results

3.1. Disease characterization of MLD

The pathophysiological processes related to MLD and the proteins related to these processes were identified in a disease characterization process through a manual review of the scientific literature. Thirteen pathophysiological processes were identified during this step (Table 1), with a total of 198 unique condition effectors identified.

3.2. Generation of mathematical models and identification of biomarker candidates

The 13 processes identified in the disease characterization step were used to generate three models that simulated the pathophysiology of MLD at different pathophysiological stages (Fig. 1). Each process identified in the disease characterization was classified into these models: processes 2–8 were classified as active during the early pro-inflammatory stage, processes 9 and 10 were additionally active during the pre-demyelination stage, and all 13 processes were active during the demyelination stage.

The mathematical models identified 3457 proteins of interest. These proteins were evaluated both individually and as combinatorial pairs to identify the most promising biomarker candidates. After applying the five filters, 13 individual biomarkers and 16 combinatorial biomarker pairs were identified as being potentially involved in early tissue damage in MLD (see Table 2 for cross-validated accuracy, sensitivity, specificity, and p values). Notably, none of the individual biomarkers were able to differentiate between the pre-demyelination stage and the demyelination stage models (Supplementary Table S6).

All 16 biomarker candidates were found to be linked directly to at least three MLD effectors according to the human protein–protein interaction network. A graphical representation of this network connecting the MLD effectors and the 16 biomarker candidates is given in Fig. 3. Furthermore, 10 of the 16 biomarker candidates were identified as MLD effectors based on the disease characterization: five candidates were effectors of mitochondrial dysfunction, four were effectors of remyelination, and one was an effector of demyelination and neurodegeneration (Supplementary Table S6). An overview of the biomarker candidates, their function, and their potential relationships with neurodegenerative diseases is provided in Supplementary Table S7.

We also found that most biomarkers were expected to be easily measurable via relatively noninvasive procedures: two biomarkers were secreted in blood, four were secreted in urine, and the remaining biomarkers, with the exception of KLF2, HCFC1, and BCL2L11, were measurable in blood or PBMCs, or were located in the cell membrane (Supplementary Table S6).

NfL did not show high biomarker capabilities in the modeling analysis in relation to MLD (Table 3). In particular, cross-validated sensitivity values were low. In comparison, biomarker candidates identified by means of mathematical models all showed cross-validated sensitivity values >65% and cross-validated accuracy values >70% (as defined by prioritization filters; Table 2), indicating overall better classification values than NfL.

3.3. Corroboration of biomarker candidates in the T-lymphocyte data set from patients with MLD

A corroboration analysis of the biomarker candidates identified through the mathematical models was conducted using the GSE23350 T-lymphocyte data set from patients with MLD, and six out of the 16 biomarker candidates were differentially expressed in the data set considering the following comparisons: patients with MLD versus controls, late-infantile MLD versus the rest of the MLD population, and late-infantile MLD versus juvenile MLD (Table 4); pro-neuregulin-1 (NRG1) appears as differential on all three comparisons. As part of the corroboration step, we performed a classification analysis using the expression data to check the existing 16 candidates and to identify new potential biomarker candidates. At the individual level, none of the candidates presented classification potential with a cross-validated accuracy >70% and a significant adjusted p value for the comparison between patients with MLD and controls. However, for the comparisons of clinical subtypes, the corroboration identified one of the 16 proteins, casein kinase II subunit alpha (CK2 alpha), as a potential individual biomarker able to distinguish late-infantile MLD from both juvenile MLD and the rest of the MLD patient population in the data set.

We performed a further data mining analysis on the T-lymphocyte data set to identify combinatorial pairs of biomarker candidates able to distinguish: 1) patients with MLD from controls and 2) late-infantile MLD from juvenile MLD and the rest of the MLD population,

Table 1
MLD processes, their effectors, and their order within the mathematical models.

Process name	Number of effectors	Process order
Deficiency of ASA	2	1
Inflammatory response	22	2
Perturbation of calcium signaling	12	2
Impaired autophagic function	27	2
ER stress and unfolded protein response	12	2
Alterations in myelin structure	2	2
Alterations in lipid trafficking	11	2
Metallothioneins overexpression	4	2
Mitochondrial dysfunction	31	3
Loss of barrier integrity	11	3
Oxidative stress	14	4
Demyelination and neurodegeneration	19	4
Remyelination	42	4

Early pro-inflammatory stage
Pre-demyelination stage
Demyelination stage

The processes are numbered according to their order of appearance in the course of tissue damage. The mathematical models do not simulate the disease over time or the clinical subtypes of the disease. Note: 'deficiency of ASA' was not considered in the modeling owing to difficulties in relating it mechanistically to the other MLD processes.
ASA, arylsulfatase A; ER, endoplasmic reticulum; MLD, metachromatic leukodystrophy.

Table 2
The evaluation of the 16 biomarker candidates for the identification of early tissue damage in MLD (as identified by mathematical models).

Protein name	Cross-validated accuracy (%)	Cross-validated sensitivity (%)	Cross-validated specificity (%)	Cross-validated *p* value
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	78.8	82.8	74.8	< 0.001
Toll-interacting protein	78.8	74.0	83.6	< 0.001
Krüppel-like factor 2	76.2	79.6	72.8	< 0.001
Mitofusin-2	75.6	82.4	68.8	< 0.001
Transcription factor 4	73.2	71.2	75.2	< 0.001
Brain-derived neurotrophic factor	70.2	74.8	65.6	< 0.001
Host cell factor C1	70.2	72.4	68.0	< 0.001
UDP-*N*-acetylglucosamine-peptide *N*-acetylglucosaminyltransferase 110 kDa subunit	74.4	84.4	64.4	< 0.001
Fatty acid-binding protein 5	73.4	68.4	78.4	< 0.001
Bcl-2-like protein 11	77.8	67.6	88.0	< 0.001
BCL2/adenovirus E1B 19 kDa protein-interacting protein 3	74.6	69.2	80.0	< 0.001
GTP-binding protein Rheb	72.8	66.8	78.8	< 0.001
Casein kinase II subunit alpha	70.6	74.8	66.4	< 0.001
Neurogenic locus notch homolog protein 1	79.8	84.4	75.2	< 0.001
Pro-neuregulin-1, membrane-bound isoform	80.6	85.2	76.0	< 0.001
Cytokine receptor-like factor 2	76.6	77.2	76.0	< 0.001

In the case of biomarkers found in different comparisons, the table shows the classification values for the highest-performing biomarker identified.
GTP, guanosine triphosphate; MLD, metachromatic leukodystrophy; UDP, uridine diphosphate.

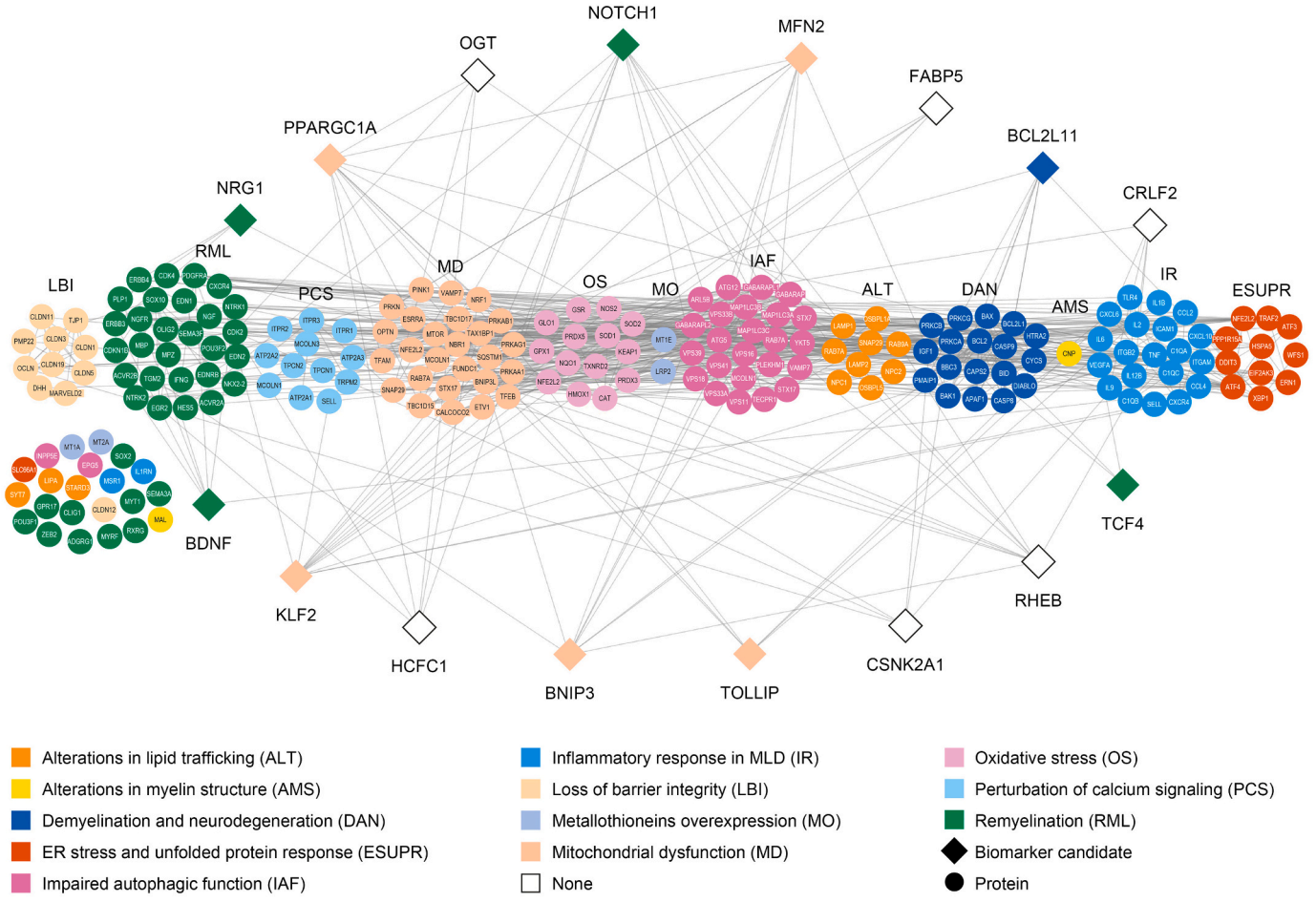

Fig. 3. Graphical representation of the protein network connecting MLD effectors and the 16 biomarker candidates, considering direct interactions of the human protein network.
Built using Cytoscape software. Diamonds denote biomarker candidates; circles denote effector proteins; the fill color indicates the MLD process to which the effectors belong.
Standard gene symbols have been used and full protein names can be found at https://www.uniprot.org/.
ER, endoplasmic reticulum; MLD, metachromatic leukodystrophy.

Table 3
Evaluation of NfL as an individual biomarker of early tissue damage in MLD according to mathematical models.

	Model comparison	Cross-validated accuracy (%)	Cross-validated sensitivity (%)	Cross-validated specificity (%)	Cross-validated *p* value
NfL	Pre-demyelination vs early pro-inflammatory	54.60	36.80	72.40	0.018
	Demyelination vs early pro-inflammatory	60.20	43.20	77.20	8.66×10^{-7}
	Demyelination vs pre-demyelination	59.00	27.60	90.40	1.38×10^{-7}

MLD, metachromatic leukodystrophy; NfL, neurofilament light chain.

considering at least one of the 16 biomarker candidates as part of the combination. For comparison 1), all candidates were able to distinguish patients with MLD from controls as part of combinatorial biomarker pairs, with 88 combinatorial pairs identified in total. All 88 identified pairs were formed by one of the 16 candidates together with one of seven additional proteins (listed in Table 5). For comparison 2), 15 of the 16 biomarker candidates were able to distinguish late-infantile MLD from juvenile MLD in the data set as part of combinatorial biomarker pairs, with 33 combinatorial pairs identified in total. Only one biomarker (brain-derived neurotrophic factor [BDNF]) was not corroborated. All 33 identified pairs were formed by one of the 15 candidates together with one of six additional proteins (listed in Table 6). Furthermore, 13 of these combinatorial biomarker pairs exclusively distinguished late-infantile MLD from juvenile MLD in the T-lymphocyte data set.

3.4. Corroboration of biomarker candidates in data sets from patients with MS

Results from the complementary corroboration of the 16 biomarker candidates in data sets from patients with MS found differential expression in PBMCs between patients with RRMS and those with SPMS in one data set (E-MTAB-5151) for five of the 16 biomarkers, although none were replicated in more than one data set (Table 4). Five of the 16 biomarkers (OGT, MFN2, HCFC1, BCL2L11, and RHEB) were identified as individual biomarkers able to distinguish patients with RRMS from those with SPMS, according to data from the E-MTAB-5151 data set. However, none of the 16 biomarkers were identified as individual markers able to distinguish between SPMS and RRMS in the other three data sets.

As for the biomarker candidate corroboration with the MLD data set, a data mining analysis was performed on the MS data sets to identify combinatorial pairs of biomarker candidates able to distinguish between SPMS and RRMS considering at least one of the 16 biomarker candidates as part of the combination. Across the four data sets, 232 combinatorial biomarker pairs were identified as able to distinguish between SPMS and RRMS. All 16 biomarker candidates identified by mathematical models were present as combinatorial biomarkers in each MS data set, except for GSE136411, in which five of the 16 biomarkers were not found to be part of any combinatorial biomarker pair. However, only 10 of the 232 combinatorial pairs were present in at least two different data sets.

4. Discussion

We aimed to propose a panel of potential biomarker candidates that could be validated to assist in the prediction of early damage in MLD, using systems biology-based modeling. Our mathematical models identified 16 biomarker candidates; six of these candidates were differentially expressed in at least one of the MLD data set comparisons,

Table 4
Biomarker candidates involved in the identification of early tissue damage in MLD (as identified by mathematical models) and whether they were significantly differentially expressed in patients with different subtypes of MLD and MS, respectively.

Protein name	Gene name	Differential in data from patients with MLD			Differential in data from patients with MS			
		GSE23350 LI vs rest of MLD (T lymphocytes)	GSE23350 LI vs J MLD (T lymphocytes)	GSE23350 MLD vs control samples (T lymphocytes)	E-MTAB-2374 SPMS vs RRMS (CSF cells)	E-MTAB-4890 SPMS vs RRMS (PBMCs)	E-MTAB-5151 SPMS vs RRMS (PBMCs)	GSE136411 SPMS vs RRMS (PBMCs)
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	*PPARGC1A*	×	×	×	×	×	×	–
Toll-interacting protein	*TOLLIP*	×	✓	×	×	×	×	×
Krüppel-like factor 2	*KLF2*	×	×	×	×	×	×	✓
Mitofusin-2	*MFN2*	×	×	×	×	×	✓	×
Transcription factor 4	*TCF4*	×	×	×	×	×	×	×
Brain-derived neurotrophic factor	*BDNF*	×	×	×	×	×	×	–
Host cell factor C1	*HCFC1*	×	×	×	×	×	✓	×
UDP-*N*-acetylglucosamine-peptide *N*-acetylglucosaminyltransferase 110 kDa subunit	*OGT*	×	×	×	×	×	✓	×
Fatty acid-binding protein 5	*FABP5*	×	×	✓	×	–	×	–
Bcl-2-like protein 11	*BCL2L11*	×	×	×	×	×	✓	–
BCL2/adenovirus E1B 19 kDa protein-interacting protein 3	*BNIP3*	×	×	×	×	×	×	×
GTP-binding protein Rheb	*RHEB*	✓	×	×	×	×	✓	×
Casein kinase II subunit alpha	*CSNK2A1*	✓	✓	×	×	×	×	×
Neurogenic locus notch homolog protein 1	*NOTCH1*	✓	×	×	×	×	×	×
Pro-neuregulin-1, membrane-bound isoform	*NRG1*	✓	✓	✓	×	×	×	×
Cytokine receptor-like factor 2	*CRLF2*	×	×	×	×	×	×	–

Ticks indicate significant differences in marker expression between the cohorts. Crosses indicate no significant difference in marker expression between the cohorts. Dashes indicate that no value for this protein was obtained either by the platform or by the protocol used to obtain the gene expression data.
CSF, cerebrospinal fluid; GTP, guanosine triphosphate; J, juvenile; LI, late-infantile; MLD, metachromatic leukodystrophy; MS, multiple sclerosis; PBMC, peripheral blood mononuclear cell; RRMS, relapsing–remitting MS; SPMS, secondary progressive MS; UDP, uridine diphosphate.

Table 5
Combinatorial biomarker pairs (consisting of model-identified candidates combined with an additional protein) that distinguish patients with MLD from controls in the T-lymphocyte data set.

Protein name	Gene name	Additional protein (gene name)
Bcl-2-like protein 11	*BCL2L11*	*AGK*
		F2RL1
		DDX41
		IL9
		MS4A4A
		F2RL1
		DDX41
		AGK
		MS4A4A
		IL9
Brain-derived neurotrophic factor	*BDNF*	*ADGRG7*
		AGK
		F2RL1
		DDX41
		IL9
		PLEKHG5
		MS4A4A
BCL2/adenovirus E1B 19 kDa protein-interacting protein 3	*BNIP3*	*ADGRG7*
		AGK
		F2RL1
		IL9
		DDX41
		MS4A4A
Cytokine receptor-like factor 2	*CRLF2*	*ADGRG7*
		F2RL1
		AGK
		IL9
		DDX41
		MS4A4A
Casein kinase II subunit alpha	*CSNK2A1*	*ADGRG7*
		AGK
		F2RL1
		DDX41
		ADGRG7
		MS4A4A
		IL9
Fatty acid-binding protein 5	*FABP5*	*PLEKHG5*
		AGK
		F2RL1
		DDX41
		MS4A4A
Host cell factor C1	*HCFC1*	*IL9*
		F2RL1
		AGK
		DDX41
		IL9
		MS4A4A
		PLEKHG5
Krüppel-like factor 2	*KLF2*	*ADGRG7*
		AGK
		F2RL1
		DDX41
		IL9
Mitofusin-2	*MFN2*	*MS4A4A*
		AGK
		F2RL1
		IL9
		DDX41
Neurogenic locus notch homolog protein 1	*NOTCH1*	*MS4A4A*
		F2RL1
		AGK
		IL9
		DDX41
Pro-neuregulin-1, membrane-bound isoform	*NRG1*	*MS4A4A*
		F2RL1
		AGK
		DDX41
		IL9
UDP-*N*-acetylglucosamine-peptide *N*-acetylglucosaminyltransferase 110 kDa subunit	*OGT*	*MS4A4A*
		AGK
		F2RL1
		DDX41
		IL9
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	*PPARGC1A*	*MS4A4A*
		AGK
		IL9
		F2RL1
		DDX41
GTP-binding protein Rheb	*RHEB*	*ADGRG7*
		F2RL1
		AGK
		IL9
Transcription factor 4	*TCF4*	*DDX41*
		F2RL1
		AGK
		DDX41
		IL9
Toll-interacting protein	*TOLLIP*	*MS4A4A*

Table 6
The 33 combinatorial biomarker pairs that distinguish late-infantile MLD from juvenile MLD in the T-lymphocyte patient data set. The table shows the 15 biomarker candidates identified in the mathematical modeling, their corresponding gene name, and which additional proteins were found to make up the combinatorial pair.

Model biomarker	Gene name	Additional protein (gene name)
Bcl-2-like protein 11	*BCL2L11*	*SUN2*
		ABCA6
		GMEB2
BCL2/adenovirus E1B 19 kDa protein-interacting protein 3	*BNIP3*	*SUN2*
		ABCA6
		FAM120A
Cytokine receptor-like factor 2	*CRLF2*	*SUN2*
		ABCA6
Casein kinase II subunit alpha	*CSNK2A1*	*ABCA6*
		CELF6
		EEF1A1
Fatty acid-binding protein 5	*FABP5*	*ABCA6*
		CELF6
Host cell factor C1	*HCFC1*	*ABCA6*
		FAM120A
Krüppel-like factor 2	*KLF2*	*ABCA6*
Mitofusin-2	*MFN2*	*ABCA6*
		CELF6
		EEF1A1
Neurogenic locus notch homolog protein 1	*NOTCH1*	*ABCA6*
		EEF1A1
		CELF6
Pro-neuregulin-1, membrane-bound isoform	*NRG1*	*SUN2*
		ABCA6
		CELF6
UDP-*N*-acetylglucosamine-peptide *N*-acetylglucosaminyltransferase 110 kDa subunit	*OGT*	*ABCA6*
		FAM120A
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	*PPARGC1A*	*ABCA6*
GTP-binding protein Rheb	*RHEB*	*SUN2*
		ABCA6
Transcription factor 4	*TCF4*	*ABCA6*
Toll-interacting protein	*TOLLIP*	*GMEB2*
		EEF1A1

GTP, guanosine triphosphate; MLD, metachromatic leukodystrophy; UDP, uridine diphosphate.

with NRG1 appearing as differential on all three comparisons. Furthermore, all biomarkers, with the exception of BDNF, were able to distinguish late-infantile MLD from juvenile MLD in a T-lymphocyte gene expression data set as part of a combinatorial pair. CK2 alpha was identified as an individual biomarker able to distinguish late-infantile MLD from both juvenile MLD and the rest of the MLD population in the T-lymphocyte gene expression data set.

Several different protein types were identified as biomarker candidates, including transcriptional regulators, growth factors, and enzyme subunits. The relevance of each of the biomarker candidates to MLD pathophysiology is highlighted by the fact that all biomarker candidates obtained from the modeling analysis were identified as either MLD effectors or as having a direct link to processes implicated in MLD. The dysregulation of some of these candidates has also been implicated in a range of neurodegenerative disorders such as Parkinson's disease [48] and Alzheimer's disease [49], as well as other rare lysosomal storage diseases such as Krabbe's disease [50].

In particular, NRG1 appeared as differential in the three comparisons performed in the MLD data sets (patients with MLD vs controls; late-infantile MLD vs the rest of MLD population; and late-infantile MLD vs juvenile MLD). NRG1 is an important factor in regulating the remyelination process in the early phases of nerve injury [51] and also signals oligodendrocyte progenitor cell proliferation and differentiation [52]. To our knowledge, NRG1 has not been directly linked to MLD development previously; however, our findings suggest that altered expression of NRG1 may play a role in MLD etiology. The characterization of biomarkers that are independent from disease stages and could predict symptom onset in patients would be beneficial to investigate further. This is particularly important to progress the adoption of newborn screening for MLD: the current lack of clinical biomarkers that can reliably identify when symptoms may begin to manifest poses ethical challenges for review boards and hinders implementation [19,21,53].

CK2 alpha, encoded by the gene *CSNK2A1*, was identified as a possible individual marker for early tissue damage in MLD using the mathematical models and was corroborated in the analyses performed on the T-lymphocyte gene expression data set in the comparison of clinical subtypes. CK2 alpha is the alpha subunit of casein kinase II, a serine/threonine kinase [54], and is implicated in a broad spectrum of biological processes [55]. It is expressed more in the brain than in any other tissue [56], upregulated in nerve injury [57], and involved in the signaling pathways that regulate inflammatory responses [54]. Altered expression of CK2 has been documented in the brain of patients with Alzheimer's disease [58]. Further, as a dominant regulator of the Th17/Treg equilibrium, which is disturbed in MS [54], it is possible that this CK2 may have relevance in autoimmune inflammatory disorders [54]. However, it is important to note that neither CK2 alpha nor NRG1 are unique to MLD, so their utility as disease-specific markers may be limited [54,56].

NfL is released into the CSF and blood upon neuroaxonal injury [28]. It has shown promise as a noninvasive biomarker for demyelinating diseases [59], including as a marker of clinical phenotype in MLD and in assessing treatment response [28]. Given its potential significance for MLD, we investigated NfL by means of mathematical models, but found that other proteins identified via our models demonstrated better classification values than NfL, particularly in terms of sensitivity. In addition, NfL was not identified as a potential candidate in the corroboration step with the T-lymphocyte data set. It is important to exercise caution when comparing biomarkers found *in vivo* with those identified via *in silico* modeling. In fact, the influence of NfL in MLD may exist at the mechanistic level rather than at the gene expression level, and it is possible that NfL expression may not be altered in the lymphocytes of patients with MLD: in this case, NfL would not have been identified as a potential candidate in our modeling. Furthermore, although NfL has a well-established role in neurodegeneration, its utility as a diagnostic or prognostic marker in MLD requires further investigation. A study found some evidence of higher NfL levels in presymptomatic patients with MLD than in healthy controls but, after symptom onset, levels tended to decline over time in this study [28]. Further work is needed to understand the differential patterns of NfL among patient groups and throughout the disease course.

The corroboration of the 16 biomarker candidates identified by mathematical models in a T-lymphocyte data set from controls and patients with MLD supports the model findings, with only BDNF not corroborated in the comparison between MLD subtypes. T cells play a central role in neuroinflammation, and inflammation in the central nervous system has been found to alter T-cell function and identity [60]. Therefore, these data provide a suitable platform to investigate potentially altered proteins in MLD. Our findings align with previous research supporting the potential relationship between the central nervous system and the immune system in demyelinating diseases. For example, subpopulations of T-helper cells and related cytokines have been found to modulate the inflammatory responses in MS [61]. Overexpression of metallothionein genes has also been reported in both the T lymphocytes and brain tissue of patients with MLD, which further correlated with disease progression [43]. This is thought to reflect a response to oxidative and inflammatory processes in MLD [43]. Nevertheless, the T-lymphocyte data set was the only data set available for biomarker candidate corroboration, which may not be fully representative of the systemic disease. Further, the average age at sampling of patients who provided the T-lymphocyte data suggests that these patients were likely symptomatic and, as such, any hypotheses on early-stage biomarkers should be considered tentative. Validation in patient samples, particularly those early-symptomatic, is imperative and will allow us to determine to what extent these biomarkers translate to clinical practice and their value in predicting prognosis. Validation in patient data is likely to be more valuable than that in animal models, given that ASA-deficient mouse models do not fully reproduce aspects of MLD pathophysiology observed in humans, such as widespread demyelination [62].

The corroboration analyses conducted with data sets of patients with MS yielded inconclusive results. Although some of the biomarker candidates identified by mathematical models were able to distinguish between RRMS and SPMS, this finding was not replicated in all available MS data sets. The heterogeneity among the MS data sets means that we should interpret these results with caution. Additionally, the biomarkers corroborated in the MS data sets were not the same biomarkers corroborated in the MLD data set. This may reflect the differing disease pathogenesis between MS and MLD. The biomarker candidates identified here may perform better in distinguishing among subtypes of other leukodystrophies (with closer pathophysiology to MLD) than among subtypes of MS. For example, biomarkers associated with globoid cell leukodystrophy (Krabbe's disease), a disease also characterized by myelin loss in the central and peripheral nervous systems due to a deficiency in the galactosylceramidase enzyme, may better reflect pathophysiological processes also associated with MLD than biomarkers associated with MS. Processes associated with glial cell death are known to occur in Krabbe's disease, and overexpression of fatty acid-binding protein 5 (FABP5; a candidate emerging from our *in silico* models) in particular has been shown to accelerate this process [50]. Data from Krabbe's disease or other leukodystrophies were unavailable at the time of this analysis; however, these would provide useful comparators for future corroboration of our findings. Furthermore, the MLD gene expression data set contained data from only two adult patients, so more work is needed to understand the value of these biomarkers to distinguish adult MLD from other subtypes.

In silico modeling has been beneficial in a number of disease areas to identify new candidates for investigation [30]. In addition, given that *in silico* modeling can predict properties that might not be inferable from observation, it could be particularly beneficial for diseases like MLD, for which data are limited owing to the rarity of the disease. An advantage of the biomarker candidates identified here is that most are expected to be easily measured through relatively noninvasive procedures for

patients. Specifically, BDNF and NRG1 are secreted in the blood, TOLLIP, MFN2, FABP5, and NOTCH1 are secreted in urine, and several others are measurable in PBMCs or located on the cell membrane. This is an important consideration if candidates were to be validated for use in clinical practice, and studies have demonstrated that biomarker levels measured in the CSF or serum can be analyzed with sufficient sensitivity to indicate the presence or progression of neurological diseases [63,64]. However, it is important to note that the levels of proteins in the blood, urine, or PBMCs may not always be a direct reflection of their levels or activity in the central nervous system.

Although these results provide promising candidates for future validation in clinical cohorts, they must be viewed in the context of the modeling methodology and the limitations associated with it. The TPMS models are mechanistic and, as such, specific variants relating to a particular protein or the accumulation of deposits cannot be considered in the model. However, the modeling can identify processes that appear because of such variants, and these have been characterized in the models. The TMPS technology also relies on publicly available data, including high-throughput data and the current state of knowledge around the molecular mechanisms involved in the disease of interest. The availability of high-throughput data from patients with MLD was limited, so the inclusion of such data to define disease stage models (which is part of the model generation) was not feasible. Additionally, there is a lack of knowledge on MLD disease pathophysiology at the molecular level, which made it difficult to obtain a literature-based characterization. For example, the first process, 'deficiency of ASA', was not considered as an input in the modeling process owing to difficulties in relating it mechanistically to the other MLD processes. Therefore, the mathematical models must be viewed as simplified simulations of real disease states and not as fully representative of complex pathophysiology. However, the *in silico* models considered the whole human protein network and utilized several known pathophysiological signals as a training set. Additionally, all biomarkers provided by the modeling analysis (13 individual biomarkers and 16 combinatorial biomarker pairs) presented cross-validation accuracy levels >70% and were corroborated on the available expression data, supporting their role in the disease and differences among disease subtypes. As such, the biomarkers identified here can be seen as promising candidates for further validation in patients with MLD.

5. Conclusions

Overall, this study provided a suitable panel of biomarker candidates for the prediction of early tissue damage in MLD and subsequent follow-up. These candidates were corroborated using T-lymphocyte data from patients with MLD and most are expected to be easily measurable, making them viable candidates for clinical validation.

Funding

This study was funded by Takeda. Under the direction of the authors, medical writing support was provided by Emma Davies PhD of Oxford PharmaGenesis, Oxford, UK and was funded by Takeda Development Center Americas, Inc.

Author contributions

J.G.-A. was involved in study conception and design; J.G., L.A., and R.V. designed and performed the mathematical modeling analysis. All authors were involved in data interpretation. All authors were involved in the drafting and revision of the manuscript critically for intellectually important content. All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work.

Declaration of Competing Interest

J.G., L.A., and R.V. are employees of Anaxomics Biotech, SL. J.G.-A. is an employee of Takeda and a Takeda stockholder.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

The authors would like to thank Cristina Segú-Vergés of Anaxomics Biotech SL for her critical appraisal of the scientific content of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymgmr.2023.100974.

References

[1] A.A. Shaimardanova, D.S. Chulpanova, V.V. Solovyeva, A.I. Mullagulova, K. V. Kitaeva, C. Allegrucci, A.A. Rizvanov, Metachromatic leukodystrophy: diagnosis, modeling, and treatment approaches, Front. Med. 7 (2020) 576221, https://doi.org/10.3389/fmed.2020.576221.

[2] V. Gieselmann, Metachromatic leukodystrophy: genetics, pathogenesis and therapeutic options, Acta Paediatr. 97 (2008) 15–21, https://doi.org/10.1111/j.1651-2227.2008.00648.x.

[3] V. Gieselmann, I. Krägeloh-Mann, Metachromatic leukodystrophy. The online metabolic and molecular bases of inherited disease, 2014 [Accessed 11 May 2021].

[4] C. Kehrer, S. Elgün, C. Raabe, J. Böhringer, S. Beck-Wödl, A. Bevot, N. Kaiser, L. Schöls, I. Krägeloh-Mann, S. Groeschel, Association of age at onset and first symptoms with disease progression in patients with metachromatic leukodystrophy, Neurology 96 (2021) e255–e266, https://doi.org/10.1212/wnl.0000000000011047.

[5] N. Gomez-Ospina, Arylsulfatase A deficiency, in: M.P. Adam, G.M. Mirzaa, R. A. Pagon, et al. (Eds.), GeneReviews®, University of Washington, Seattle, 2020, pp. 5–6 [Accessed 16 July 2018].

[6] K.M. Page, E.O. Stenger, J.A. Connelly, D. Shyr, T. West, S. Wood, L. Case, M. Kester, S. Shim, L. Hammond, M. Hammond, C. Webb, A. Biffi, B. Bambach, A. Fatemi, J. Kurtzberg, Hematopoietic stem cell transplantation to treat leukodystrophies: clinical practice guidelines from the Hunter's Hope Leukodystrophy Care Network, Biol. Blood Marrow Transplant. 25 (2019) e363–e374, https://doi.org/10.1016/j.bbmt.2019.09.003.

[7] J. Beschle, M. Döring, C. Kehrer, C. Raabe, U. Bayha, M. Strölin, J. Böhringer, A. Bevot, N. Kaiser, B. Bender, A. Grimm, P. Lang, I. Müller, I. Krägeloh-Mann, S. Groeschel, Early clinical course after hematopoietic stem cell transplantation in children with juvenile metachromatic leukodystrophy, Mol. Cell. Pediatr. 7 (2020) 12, https://doi.org/10.1186/s40348-020-00103-7.

[8] European Medicines Agency. Libmeldy summary of product information, 2022. https://www.ema.europa.eu/en/medicines/human/EPAR/libmeldy [Accessed 13 April 2023].

[9] F. Fumagalli, V. Calbi, M.G. Natali Sora, M. Sessa, C. Baldoli, P.M.V. Rancoita, F. Ciotti, M. Sarzana, M. Fraschini, A.A. Zambon, S. Acquati, D. Redaelli, V. Attanasio, S. Miglietta, F. De Mattia, F. Barzaghi, F. Ferrua, M. Migliavacca, F. Tucci, V. Gallo, U. Del Carro, S. Canale, I. Spiga, L. Lorioli, S. Recupero, E. S. Fratini, F. Morena, P. Silvani, M.R. Calvi, M. Facchini, S. Locatelli, A. Corti, S. Zancan, G. Antonioli, G. Farinelli, M. Gabaldo, J. Garcia-Segovia, L.C. Schwab, G.F. Downey, M. Filippi, M.P. Cicalese, S. Martino, C. Di Serio, F. Ciceri, M. E. Bernardo, L. Naldini, A. Biffi, A. Aiuti, Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access, Lancet 399 (2022) 372–383, https://doi.org/10.1016/s0140-6736(21)02017-1.

[10] G. Morton, S. Thomas, P. Roberts, V. Clark, J. Imrie, A. Morrison, The importance of early diagnosis and views on newborn screening in metachromatic leukodystrophy: results of a caregiver survey in the UK and Republic of Ireland, Orphanet J. Rare Dis. 17 (2022) 403, https://doi.org/10.1186/s13023-022-02550-z.

[11] X. Hong, J. Daiker, M. Sadilek, N. Ruiz-Schultz, A.B. Kumar, S. Norcross, W. Dansithong, T. Suhr, M.L. Escolar, C. Ronald Scott, A. Rohrwasser, M.H. Gelb, Toward newborn screening of metachromatic leukodystrophy: results from analysis of over 27,000 newborn dried blood spots, Genet. Med. 23 (2021) 555–561, https://doi.org/10.1038/s41436-020-01017-5.

[12] S.A. Jones, D. Cheillan, A. Chakrapani, H.J. Church, S. Heales, T.H.Y. Wu, G. Morton, P. Roberts, E.F. Sluys, A. Burlina, Application of a novel algorithm for expanding newborn screening for inherited metabolic disorders across Europe, Int. J. Neonatal Screen. 8 (2022) 20, https://doi.org/10.3390/ijns8010020.

[13] S. Groeschel, C. Kehrer, C. Engel, A. Bley, R. Steinfeld, W. Grodd, I. Krageloh-Mann, Metachromatic leukodystrophy: natural course of cerebral MRI changes in relation to clinical course, J. Inherit. Metab. Dis. 34 (2011) 1095–1102, https://doi.org/10.1007/s10545-011-9361-1.
[14] D.F. van Rappard, J.J. Boelens, N.I. Wolf, Metachromatic leukodystrophy: disease spectrum and approaches for treatment, Best Pract. Res. Clin. Endocrinol. Metab. 29 (2015) 261–273, https://doi.org/10.1016/j.beem.2014.10.001.
[15] F. Eichler, C. Sevin, M. Barth, F. Pang, K. Howie, M. Walz, A. Wilds, C. Calcagni, C. Chanson, L. Campbell, Understanding caregiver descriptions of initial signs and symptoms to improve diagnosis of metachromatic leukodystrophy, Orphanet J. Rare Dis. 17 (2022) 370, https://doi.org/10.1186/s13023-022-02518-z.
[16] A.A. Boucher, W. Miller, R. Shanley, R. Ziegler, T. Lund, G. Raymond, P.J. Orchard, Long-term outcomes after allogeneic hematopoietic stem cell transplantation for metachromatic leukodystrophy: the largest single-institution cohort report, Orphanet J. Rare Dis. 10 (2015) 94, https://doi.org/10.1186/s13023-015-0313-y.
[17] A. Biffi, E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, C. Baldoli, S. Martino, A. Calabria, S. Canale, F. Benedicenti, G. Vallanti, L. Biasco, S. Leo, N. Kabbara, G. Zanetti, W.B. Rizzo, N.A.L. Mehta, M.P. Cicalese, M. Casiraghi, J. J. Boelens, U. Del Carro, D.J. Dow, M. Schmidt, A. Assanelli, V. Neduva, C. Di Serio, E. Stupka, J. Gardner, C. Von Kalle, C. Bordignon, F. Ciceri, A. Rovelli, M. G. Roncarolo, A. Aiuti, M. Sessa, L. Naldini, Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy, Science 341 (2013) 1233158, https://doi.org/10.1126/science.1233158.
[18] D.A. Wenger, S. Coppola, S.-L. Liu, Insights into the diagnosis and treatment of lysosomal storage diseases, Arch. Neurol. 60 (2003) 322, https://doi.org/10.1001/archneur.60.3.322.
[19] L. Laugwitz, L. Zizmare, V. Santhanakumaran, C. Cannet, J. Böhringer, J.G. Okun, M. Spraul, I. Krägeloh-Mann, S. Groeschel, C. Trautwein, Identification of neurodegeneration indicators and disease progression in metachromatic leukodystrophy using quantitative NMR−based urinary metabolomics, JIMD Rep. 63 (2022) 168–180, https://doi.org/10.1002/jmd2.12273.
[20] V. Gieselmann, I. Krägeloh-Mann, Metachromatic leukodystrophy – an update, Neuropediatrics. 41 (2010) 1–6, https://doi.org/10.1055/s-0030-1253412.
[21] V. Santhanakumaran, S. Groeschel, K. Harzer, C. Kehrer, S. Elgün, S. Beck-Wödl, H. Hengel, L. Schöls, T.B. Haack, I. Krägeloh-Mann, L. Laugwitz, Predicting clinical phenotypes of metachromatic leukodystrophy based on the arylsulfatase A activity and the *ARSA* genotype? – Chances and challenges, Mol. Genet. Metab. 137 (2022) 273–282, https://doi.org/10.1016/j.ymgme.2022.09.009.
[22] J.-E. Cheon, I.-O. Kim, Y.S. Hwang, K.J. Kim, K.-C. Wang, B.-K. Cho, J.G. Chi, C. J. Kim, W.S. Kim, K.M. Yeon, Leukodystrophy in children: a pictorial review of MR imaging features, Radiographics. 22 (2002) 461–476, https://doi.org/10.1148/radiographics.22.3.g02ma01461.
[23] M.R. Natowicz, E.M. Prence, P. Chaturvedi, D.S. Newburg, Urine sulfatides and the diagnosis of metachromatic leukodystrophy, Clin. Chem. 42 (1996) 232–238.
[24] P. Martin, G.E. Hagberg, T. Schultz, K. Harzer, U. Klose, B. Bender, T. Nägele, K. Scheffler, I. Krägeloh-Mann, S. Groeschel, T2-Pseudonormalization and microstructural characterization in advanced stages of late-infantile metachromatic leukodystrophy, Clin. Neuroradiol. 31 (2020) 969–980, https://doi.org/10.1007/s00062-020-00975-2.
[25] M.A.F. Tan, M. Fuller, Z.A.M.H. Zabidi-Hussin, J.J. Hopwood, P.J. Meikle, Biochemical profiling to predict disease severity in metachromatic leukodystrophy, Mol. Genet. Metab. 99 (2010) 142–148, https://doi.org/10.1016/j.ymgme.2009.09.006.
[26] C.Í. Dali, N.W. Barton, M.H. Farah, M. Moldovan, J.E. Månsson, N. Nair, M. Dunø, L. Risom, H. Cao, L. Pan, M. Sellos-Moura, A.M. Corse, C. Krarup, Sulfatide levels correlate with severity of neuropathy in metachromatic leukodystrophy, Ann. Clin. Transl. Neurol. 2 (2015) 518–533, https://doi.org/10.1002/acn3.193.
[27] K.A. Thibert, G.V. Raymond, J. Tolar, W.P. Miller, P.J. Orchard, T.C. Lund, Cerebral spinal fluid levels of cytokines are elevated in patients with metachromatic leukodystrophy, Sci. Rep. 6 (2016) 24579, https://doi.org/10.1038/srep24579.
[28] S. Beerepoot, H. Heijst, B. Roos, M.M.C. Wamelink, J.J. Boelens, C.A. Lindemans, P. M. Van Hasselt, E.H. Jacobs, M.S. Van Der Knaap, C.E. Teunissen, N.I. Wolf, Neurofilament light chain and glial fibrillary acidic protein levels in metachromatic leukodystrophy, Brain 145 (2021) 105–118, https://doi.org/10.1093/brain/awab304.
[29] C. Kehrer, G. Blumenstock, C. Raabe, I. Krägeloh-Mann, Development and reliability of a classification system for gross motor function in children with metachromatic leucodystrophy, Dev. Med. Child Neurol. 53 (2011) 156–160, https://doi.org/10.1111/j.1469-8749.2010.03821.x.
[30] V. Lorén, A. Garcia-Jaraquemada, J.E. Naves, X. Carmona, M. Mañosa, A. M. Aransay, J.L. Lavin, I. Sánchez, E. Cabré, J. Manyé, E. Domènech, ANP32E, a protein involved in steroid-refractoriness in ulcerative colitis, identified by a systems biology approach, J. Crohn's Colitis 13 (2019) 351–361, https://doi.org/10.1093/ecco-jcc/jjy171.
[31] N. Gimenez, R. Tripathi, A. Giró, L. Rosich, M. López-Guerra, I. López-Oreja, H. Playa-Albinyana, F. Arenas, J.M. Mas, P. Pérez-Galán, J. Delgado, E. Campo, J. Farrés, D. Colomer, Systems biology drug screening identifies statins as enhancers of current therapies in chronic lymphocytic leukemia, Sci. Rep. 10 (2020) 22153, https://doi.org/10.1038/s41598-020-78315-0.
[32] D. Romeo-Guitart, J. Forés, M. Herrando-Grabulosa, R. Valls, T. Leiva-Rodríguez, E. Galea, F. González-Pérez, X. Navarro, V. Petegnief, A. Bosch, M. Coma, J.M. Mas, C. Casas, Neuroprotective drug for nerve trauma revealed using artificial intelligence, Sci. Rep. 8 (2018) 1879, https://doi.org/10.1038/s41598-018-19767-3.
[33] G. Moncunill, A. Scholzen, M. Mpina, A. Nhabomba, A.B. Hounkpatin, L. Osaba, R. Valls, J.J. Campo, H. Sanz, C. Jairoce, N.A. Williams, E.M. Pasini, D. Arteta, J. Maynou, L. Palacios, M. Duran-Frigola, J.J. Aponte, C.H.M. Kocken, S. T. Agnandji, J.M. Mas, B. Mordmüller, C. Daubenberger, R. Sauerwein, C. Dobaño, Antigen-stimulated PBMC transcriptional protective signatures for malaria immunization, Sci. Transl. Med 12 (2020) eaay8924, https://doi.org/10.1126/scitranslmed.aay8924.
[34] L. Artigas, M. Coma, P. Matos-Filipe, J. Aguirre-Plans, J. Farrés, R. Valls, N. Fernandez-Fuentes, J. De La Haba-Rodriguez, A. Olvera, J. Barbera, R. Morales, B. Oliva, J.M. Mas, In-silico drug repurposing study predicts the combination of pirfenidone and melatonin as a promising candidate therapy to reduce SARS-CoV-2 infection progression and respiratory distress caused by cytokine storm, PLoS One 15 (2020), e0240149, https://doi.org/10.1371/journal.pone.0240149.
[35] G. Jorba, J. Aguirre-Plans, V. Junet, C. Segú-Vergés, J.L. Ruiz, A. Pujol, N. Fernández-Fuentes, J.M. Mas, B. Oliva, In-silico simulated prototype-patients using TPMS technology to study a potential adverse effect of sacubitril and valsartan, PLoS One 15 (2020), e0228926, https://doi.org/10.1371/journal.pone.0228926.
[36] A. Bayes-Genis, O. Iborra-Egea, G. Spitaleri, M. Domingo, E. Revuelta-López, P. Codina, G. Cediel, E. Santiago-Vacas, A. Cserkóová, D. Pascual-Figal, J. Núñez, J. Lupón, Decoding empagliflozin's molecular mechanism of action in heart failure with preserved ejection fraction using artificial intelligence, Sci. Rep. 11 (2021) 12025, https://doi.org/10.1038/s41598-021-91546-z.
[37] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K. A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets—update, Nucleic Acids Res. 41 (2012) D991–D995, https://doi.org/10.1093/nar/gks1193.
[38] A. Athar, A. Füllgrabe, N. George, H. Iqbal, L. Huerta, A. Ali, C. Snow, N. A. Fonseca, R. Petryszak, I. Papatheodorou, U. Sarkans, A. Brazma, ArrayExpress update – from bulk to single-cell expression data, Nucleic Acids Res. 47 (2018) D711–D715, https://doi.org/10.1093/nar/gky964.
[39] Y. Perez-Riverol, M. Bai, Leprevost F. Da Veiga, S. Squizzato, Y.M. Park, K. Haug, A.J. Carroll, D. Spalding, J. Paschall, M. Wang, N. Del-Toro, T. Ternent, P. Zhang, N. Buso, N. Bandeira, E.W. Deutsch, D.S. Campbell, R.C. Beavis, R.M. Salek, U. Sarkans, R. Petryszak, M. Keays, E. Fahy, M. Sud, S. Subramaniam, A. Barbera, R.C. Jiménez, A.I. Nesvizhskii, S.-A. Sansone, C. Steinbeck, R. Lopez, J.A. Vizcaíno, P. Ping, H. Hermjakob, Discovering and linking public omics data sets using the omics discovery index, Nat. Biotechnol. 35 (2017) 406–409, https://doi.org/10.1038/nbt.3790.
[40] K.M. Ryan, I. Patterson, D.M. McLoughlin, Peroxisome proliferator-activated receptor gamma co-activator-1 alpha in depression and the response to electroconvulsive therapy, Psychol. Med. 49 (2019) 1859–1868, https://doi.org/10.1017/s0033291718002556.
[41] L. Lo Iacono, S. Bussone, D. Andolina, R. Tambelli, A. Troisi, V. Carola, Dissecting major depression: the role of blood biomarkers and adverse childhood experiences in distinguishing clinical subgroups, J. Affect. Disord. 276 (2020) 351–360, https://doi.org/10.1016/j.jad.2020.07.034.
[42] F. Alizadeh, J. Tavakkoly-Bazzaz, A. Bozorgmehr, A.A. Azarnezhad, M. Tabrizi, Ananloo E. Shahsavand, Association of transcription factor 4 (*TCF4*) gene mRNA level with schizophrenia, its psychopathology, intelligence and cognitive impairments, J. Neurogenet. 31 (2017) 344–351, https://doi.org/10.1080/01677063.2017.1396330.
[43] M. Cesani, E. Cavalca, R. Macco, G. Leoncini, M.R. Terreni, L. Lorioli, R. Furlan, G. Comi, C. Doglioni, D. Zacchetti, M. Sessa, C.R. Scherzer, A. Biffi, Metallothioneins as dynamic markers for brain disease in lysosomal disorders, Ann. Neurol. 75 (2014) 127–137, https://doi.org/10.1002/ana.24053.
[44] T. Williams, H. Houlden, E. Murphy, N. John, N.C. Fox, J.M. Schott, M. Adams, I. Davagananam, J. Chataway, D.S. Lynch, How to diagnose difficult white matter disorders, Pract. Neurol. 20 (2020) 280–286, https://doi.org/10.1136/practneurol-2020-002530.
[45] A.M. Müller, E. Jun, H. Conlon, S.A. Sadiq, Cerebrospinal hepatocyte growth factor levels correlate negatively with disease activity in multiple sclerosis, J. Neuroimmunol. 251 (2012) 80–86, https://doi.org/10.1016/j.jneuroim.2012.06.008.
[46] S. Srinivasan, M. Severa, F. Rizzo, R. Menon, E. Brini, R. Mechelli, V. Martinelli, P. Hertzog, M. Salvetti, R. Furlan, G. Martino, G. Comi, E.M. Coccia, C. Farina, Transcriptional dysregulation of interferome in experimental and human multiple sclerosis, Sci. Rep. 7 (2017), https://doi.org/10.1038/s41598-017-09286-y.
[47] M. Acquaviva, R. Menon, M. Di Dario, G. Dalla Costa, M. Romeo, F. Sangalli, B. Colombo, L. Moiola, V. Martinelli, G. Comi, C. Farina, Inferring multiple sclerosis stages from the blood transcriptome via machine learning, Cell Rep. Med. 1 (2020), 100053, https://doi.org/10.1016/j.xcrm.2020.100053.
[48] T. Zhang, L. Xue, L. Li, C. Tang, Z. Wan, R. Wang, J. Tan, Y. Tan, H. Han, R. Tian, T. R. Billiar, W.A. Tao, Z. Zhang, BNIP3 protein suppresses PINK1 kinase proteolytic cleavage to promote mitophagy, J. Biol. Chem. 291 (2016) 21616–21629, https://doi.org/10.1074/jbc.M116.733410.
[49] D.H. Cribbs, N.C. Berchtold, V. Perreau, P.D. Coleman, J. Rogers, A.J. Tenner, C. W. Cotman, Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study, J. Neuroinflammation 9 (2012) 179, https://doi.org/10.1186/1742-2094-9-179.
[50] A. Cheng, I. Kawahata, K. Fukunaga, Fatty acid binding protein 5 mediates cell death by Psychosine exposure through mitochondrial macropores formation in oligodendrocytes, Biomedicines 8 (2020) 635, https://doi.org/10.3390/biomedicines8120635.

[51] F.R. Fricker, A. Antunes-Martins, J. Galino, R. Paramsothy, F. La Russa, J. Perkins, R. Goldberg, J. Brelstaff, N. Zhu, S.B. McMahon, C. Orengo, A.N. Garratt, C. Birchmeier, D.L.H. Bennett, Axonal neuregulin 1 is a rate limiting but not essential factor for nerve remyelination, Brain 136 (2013) 2279–2297, https://doi.org/10.1093/brain/awt148.

[52] H. Colognato, S. Ramachandrappa, I.M. Olsen, C. ffrench-Constant, Integrins direct Src family kinases to regulate distinct phases of oligodendrocyte development, J. Cell Biol. 167 (2004) 365–375, https://doi.org/10.1083/jcb.200404076.

[53] M.C. Cornel, T. Rigter, M.E. Jansen, L. Henneman, Neonatal and carrier screening for rare diseases: how innovation challenges screening criteria worldwide, J. Community Genet. 12 (2021) 257–265, https://doi.org/10.1007/s12687-020-00488-y.

[54] S.A. Gibson, E.N. Benveniste, Protein kinase CK2: an emerging regulator of immunity, Trends Immunol. 39 (2018) 82–85, https://doi.org/10.1016/j.it.2017.12.002.

[55] T. Nuñez De Villavicencio-Diaz, A. Rabalski, Protein kinase CK2: intricate relationships within regulatory cellular networks, Pharmaceuticals 10 (2017) 27, https://doi.org/10.3390/ph10010027.

[56] P.R. Blanquet, Casein kinase 2 as a potentially important enzyme in the nervous system, Prog. Neurobiol. 60 (2000) 211–246, https://doi.org/10.1016/s0301-0082(99)00026-x.

[57] S.-R. Chen, H.-Y. Zhou, H.S. Byun, H. Chen, H.-L. Pan, Casein kinase II regulates *N*-methyl-D-aspartate receptor activity in spinal cords and pain hypersensitivity induced by nerve injury, J. Pharmacol. Exp. Ther. 350 (2014) 301–312, https://doi.org/10.1124/jpet.114.215855.

[58] C.A. Marshall, J.D. McBride, L. Changolkar, D.M. Riddle, J.Q. Trojanowski, V.M. Y. Lee, Inhibition of CK2 mitigates Alzheimer's tau pathology by preventing NR2B synaptic mislocalization, Acta Neuropathol. Commun. 10 (2022) 30, https://doi.org/10.1186/s40478-022-01331-w.

[59] S. Thebault, D.R. Tessier, H. Lee, M. Bowman, A. Bar-Or, D.L. Arnold, H.L. Atkins, V. Tabard-Cossa, M.S. Freedman, High serum neurofilament light chain normalizes after hematopoietic stem cell transplantation for MS, Neurol. Neuroimmunol. Neuroinflam. 6 (2019), e598, https://doi.org/10.1212/nxi.0000000000000598.

[60] N. Benallegue, H. Kebir, J.I. Alvarez, Neuroinflammation: extinguishing a blaze of T cells, Immunol. Rev. 311 (2022) 151–176, https://doi.org/10.1111/imr.13122.

[61] M. Kunkl, S. Frascolla, C. Amormino, E. Volpe, L. Tuosto, T helper cells: the modulators of inflammation in multiple sclerosis, Cells 9 (2020) 482, https://doi.org/10.3390/cells9020482.

[62] B. Hess, P. Saftig, D. Hartmann, R. Coenen, R. Lüllmann-Rauch, H.H. Goebel, M. Evers, K. von Figura, R. D'Hooge, G. Nagels, P. De Deyn, C. Peters, V. Gieselmann, Phenotype of arylsulfatase A-deficient mice: relationship to human metachromatic leukodystrophy, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14821–14826, https://doi.org/10.1073/pnas.93.25.14821.

[63] A. Brodovitch, J. Boucraut, E. Delmont, A. Parlanti, A.-M. Grapperon, S. Attarian, A. Verschueren, Combination of serum and CSF neurofilament-light and neuroinflammatory biomarkers to evaluate ALS, Sci. Rep. 11 (2021) 703, https://doi.org/10.1038/s41598-020-80370-6.

[64] A. Leuzy, N.C. Cullen, N. Mattsson-Carlgren, O. Hansson, Current advances in plasma and cerebrospinal fluid biomarkers in Alzheimer's disease, Curr. Opin. Neurol. 34 (2021) 266–274, https://doi.org/10.1097/wco.0000000000000904.