Computers in Biology and Medicine 176 (2024) 108588

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Paired plasma lipidomics and proteomics analysis in the conversion from mild cognitive impairment to Alzheimer's disease

Alicia Gómez-Pascual [a,b,**,+], Talel Naccache [c,+], Jin Xu [d], Kourosh Hooshmand [b], Asger Wretlind [b], Martina Gabrielli [e], Marta Tiffany Lombardo [e,f], Liu Shi [g], Noel J. Buckley [h,i], Betty M. Tijms [j], Stephanie J.B. Vos [k], Mara ten Kate [j], Sebastiaan Engelborghs [l,m], Kristel Sleegers [n,o], Giovanni B. Frisoni [p,q], Anders Wallin [r], Alberto Lleó [s], Julius Popp [t,u], Pablo Martinez-Lage [v], Johannes Streffer [w], Frederik Barkhof [x,y], Henrik Zetterberg [z,aa,ab,ac], Pieter Jelle Visser [j,k], Simon Lovestone [h,ad], Lars Bertram [ae,af], Alejo J. Nevado-Holgado [h], Alice Gualerzi [ag], Silvia Picciolini [ag], Petroula Proitsi [ah], Claudia Verderio [e], Juan A. Botía [a,1], Cristina Legido-Quigley [b,d,*,1]

[a] Department of Information and Communications Engineering Faculty of Informatics, University of Murcia, Murcia, Spain
[b] Steno Diabetes Center Copenhagen, Herlev, Denmark
[c] Department of Data Science, City University of London, United Kingdom
[d] Institute of Pharmaceutical Science, King's College London, London, United Kingdom
[e] CNR Institute of Neuroscience, 20854, Vedano al Lambro, Italy
[f] School of Medicine and Surgery, University of Milano-Bicocca, 20126, Italy
[g] Novo Nordisk Research Centre Oxford (NNRCO), Oxford, United Kingdom
[h] Department of Psychiatry, University of Oxford, United Kingdom
[i] Kavli Institute for Nanoscience Discovery, Denmark
[j] Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands
[k] Department of Psychiatry and Neuropsychology, Alzheimer Centrum Limburg, Maastricht University, Maastricht, the Netherlands
[l] Reference Center for Biological Markers of Dementia (BIODEM), Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium
[m] Department of Neurology and Bru-BRAIN, UZ Brussel and Center for Neurosciences (C4N), Vrije Universiteit Brussel, Brussels, Belgium
[n] Complex Genetics Group, VIB Center for Molecular Neurology, VIB, Antwerp, Belgium
[o] Institute Born-Bunge, Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium
[p] University of Geneva, Geneva, Switzerland
[q] IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
[r] Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
[s] Neurology Department, Hospital Sant Pau, Barcelona, Spain, Centro de Investigación en Red en enfermedades neurodegenerativas (CIBERNED)
[t] Old age psychiatry, University Hospital of Lausanne, University of Lausanne, Switzerland
[u] Department of Geriatric Psychiatry, University Hospital of Psychiatry Zürich, University of Zürich, Switzerland
[v] CITA-Alzheimer Foundation, San Sebastian, Spain
[w] AC Immune SA, Lausanne, Switzerland, formerly Janssen R&D, LLC. Beerse, Belgium at the time of study conduct
[x] Department of Radiology and Nuclear Medicine, Amsterdam UMC, Vrije Universiteit, the Netherlands
[y] Queen Square Institute of Neurology and Centre for Medical Image Computing, University College London, United Kingdom
[z] Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
[aa] Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden
[ab] UK Dementia Research Institute at UCL, London, United Kingdom
[ac] Department of Neurodegenerative Disease, UCL Institute of Neurology, London, United Kingdom
[ad] Janssen Medical (UK), High Wycombe, United Kingdom
[ae] Lübeck Interdisciplinary Platform for Genome Analytics, University of Lübeck, Lübeck, Germany
[af] Department of Psychology, University of Oslo, Oslo, Norway
[ag] IRCCS Fondazione Don Carlo Gnocchi ONLUS in Milan, Italy
[ah] Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

* Corresponding author. Steno Diabetes Center Copenhagen, Herlev, Denmark.
** Corresponding author. Steno Diabetes Center Copenhagen, Herlev, Denmark.
E-mail address: cristina.legido_quigley@kcl.ac.uk (C. Legido-Quigley).
+ Contributed equally.
[1] Senior authors with equal contribution.

https://doi.org/10.1016/j.compbiomed.2024.108588

Available online 13 May 2024

ARTICLE INFO

Keywords:
Alzheimer's disease
Integrative omics
Machine learning
Metabolomics
Mild cognitive impairment
Proteomics

ABSTRACT

Background: Alzheimer's disease (AD) is a neurodegenerative condition for which there is currently no available medication that can stop its progression. Previous studies suggest that mild cognitive impairment (MCI) is a phase that precedes the disease. Therefore, a better understanding of the molecular mechanisms behind MCI conversion to AD is needed.

Method: Here, we propose a machine learning-based approach to detect the key metabolites and proteins involved in MCI progression to AD using data from the European Medical Information Framework for Alzheimer's Disease Multimodal Biomarker Discovery Study. Proteins and metabolites were evaluated separately in multiclass models (controls, MCI and AD) and together in MCI conversion models (MCI stable vs converter). Only features selected as relevant by 3/4 algorithms proposed were kept for downstream analysis.

Results: Multiclass models of metabolites highlighted nine features further validated in an independent cohort (0.726 mean balanced accuracy). Among these features, one metabolite, oleamide, was selected by all the algorithms. Further *in-vitro* experiments in rodents showed that disease-associated microglia excreted oleamide in vesicles. Multiclass models of proteins stood out with nine features, validated in an independent cohort (0.720 mean balanced accuracy). However, none of the proteins was selected by all the algorithms. Besides, to distinguish between MCI stable and converters, 14 key features were selected (0.872 AUC), including tTau, alpha-synuclein (SNCA), junctophilin-3 (JPH3), properdin (CFP) and peptidase inhibitor 15 (PI15) among others.

Conclusions: This omics integration approach highlighted a set of molecules associated with MCI conversion important in neuronal and glia inflammation pathways.

1. Introduction

Mild cognitive impairment (MCI) is defined as the symptomatic predementia stage characterized by objective impairment in cognition that does not interfere notably with activities of daily life [1]. It is estimated that over 15% of community dwellers have MCI. The prevalence of MCI increases with age and decreases with education and it is a heterogeneous and unstable condition [2]. MCI individuals who eventually progress to Alzheimer's Disease (AD) diagnosis are classified as MCI converters (cMCI) while those who remain stable or improve are classified as MCI stable (sMCI). It has been reported that approximately 38% of individuals with prevalent or incident MCI will revert back to a normal cognition diagnosis while 29% will progress to dementia [3]. A better characterization of the molecular mechanisms behind MCI conversion to AD is necessary for an early detection of AD. Previous studies on MCI progression to AD analyzed data from different natures, specially cognition tests, demographic information, neuroimaging and cerebrospinal fluid biomarkers [4–9] using traditional statistics methods such as statistical tests, correlation and similar metrics [9] as well as models such as Cox regression [5,9] or Bayesian networks [10]. With the advancement of high-throughput sequencing technologies, it is possible the generation of large-scale data such as the genome, transcriptome, proteome, metabolome and lipidome, which can provide new insights into complex diseases such AD [11]. These measures are taken from biofluids such as plasma and cerebrospinal fluid, as structural and functional changes in the brain can be reflected in these fluids [12]. Recent studies on MCI progression are starting to incorporate omics data, including genomics [4,13], transcriptomics [13], proteomics [10,14] and metabolomics [15,16] data. This omics data have been analyzed in combination with imaging and clinical data through machine learning (ML) algorithms such as lasso regression [4,10,17], random forest [18], support vector machines (SVM) [4,17] as well as deep learning algorithms [4,6,8]. The main advantage of multi-omics data is that it can be integrated to study the complex interplay between different biological layers (i.e. transcriptomics, metabolomics, proteomics). To this end, different strategies for multi-omics data integration has been proposed, including (1) early concatenation, commonly used because of its simplicity, it consists of concatenation of every datasets into a single large matrix; (2) intermediate integration, where datasets are projected into a latent space transformed or mapped to reduce their complexity, either independently or jointly; (3) late integration, where each omic is analyzed separately and most relevant features are integrated afterward [19]. Most of the studies of MCI progression to AD analyzed a single type of omic. Therefore, the integration of multi-omics data is poorly explored in the field of study of MCI progression to AD [13].

Here, we analyzed plasma proteomics and metabolomics data from healthy controls, MCI and AD participants from the European Medical Information Framework for Alzheimer's Disease Multimodal Biomarker Discovery Study (EMIF-AD MBD) [20] in order to identify key pathways involved in the MCI conversion to AD. To the best of our knowledge, this is the second study where both, proteomics and metabolomics data is analyzed to investigate the progression from MCI to AD. The first study, published by François, M. et al. (2022) [21], was carried out on a small cohort (NC = 40, MCI = 20, AD = 20) with only untargeted proteomics and untargeted metabolomics data. The present research focuses on the EMIF-AD study cohort made up of 800 participants with targeted and untargeted metabolomic and proteomic data from plasma samples. In our study, two research questions were proposed. First, we wanted to identify the most reliable proteins and metabolites that differentiate between controls, MCI and AD subjects. To this end, proteins and metabolites were analyzed separately with four ML algorithms and repeated features were kept for downstream analysis. Secondly, we proposed a more specific research question: which proteins and metabolites are most relevant in differentiating between sMCI and cMCI donors? To this end, paired proteomics and metabolomics data from MCI conversion participants with follow-up information were concatenated, which allowed us to capture inter-omics interactions not detected when evaluating each omic separately. To the best of our knowledge, this is the first study where paired proteomics and metabolomics data from MCI participants with follow-up information is analyzed. Although previous studies have investigated MCI progression to AD in MCI participants with follow-up information, they analyzed omics different from proteomics and metabolomics [13].

Multiclass models of metabolites selected oleamide, an endocannabinoid, as a key metabolite to distinguish between controls, MCI and AD participants. Since another endocannabinoid, anandamide, was previously found in microglia cells [22], we investigated oleamide in the same cell type with *in-vitro* experiments and demonstrated that disease-associated microglia excreted oleamide in vesicles. Multiclass protein models did not show any proteins selected by the four algorithms. Our MCI conversion models showed that Alpha-synuclein (SNCA), Properdin (CFP), Peptidase inhibitor 15 (PI15), pancreatic hormone (PPY), phospholipase A2 (PLA2G1B) and testis-expressed sequence 29 protein (TEX 29), not selected with multiclass models,

were extracted as the most confident features to distinguish between sMCI and cMCI. In summary, our results suggested pathways in MCI conversion to be associated with inflammation, sedation and neural degradation.

2. Methods

2.1. Participants from the EMIF-AD Multimodal biomarker discovery study

This study employed data from the European Medical Information Framework for Alzheimer's Disease Multimodal Biomarker Discovery Study (EMIF-AD MBD) [20]. It is a cross-cohort study consisting of collated data from 11 European cohorts that aims to discover novel diagnostic and prognostic markers for AD-type dementia by performing analyses in multiple biomarker modalities. The 11 cohorts included three multicenter studies and eight single centers (for more information, see Supplementary Table 1). For this study, a total of 230 participants with normal cognition (NC), 184 participants with AD-type dementia and 386 participants with MCI were included. The criteria to diagnose NC subjects was a normal performance on neuropsychological assessment (within 1.5 SD of the average for age, gender and education). MCI subjects were diagnosed according to the criteria of Petersen [23] in nine cohorts and the criteria of Winblad et al. [24] in the other two cohorts. Finally, AD-type dementia was diagnosed according to the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria (NINCDS-ADRDA) [25]. Of 386 MCI participants, 100 were later diagnosed with AD-type dementia (defined as cMCI), 219 remained as MCI (defined as sMCI) and 67 participants do not have this information available. The average clinical follow-up time for MCI participants was 2.2 (SD 1.3) years. From all cohorts, available data on demographics, clinical information, neuropsychological testing, cognition and Aβ status data were gathered. In addition, targeted and untargeted metabolomic and proteomic analyses were performed on plasma.

2.2. Metabolomics data acquisition and preprocessing

Metabolomics data was acquired by Metabolon Inc. (Morrisville, NC, USA). The relative levels of 883 plasma metabolites were measured in fasting blood samples using three different mass spectrometry methods previously described by Kim M. et al. (2019) [26]. Area counts for each metabolite in each sample were extracted from the raw data. The extracted area counts were then normalized to correct for variation resulting from instrument inter-day tuning differences. After removing metabolites whose levels fall above 3 standard deviations of the mean value as well as metabolites with more than 50% missing data, 540 metabolites were kept for downstream analysis. Remaining missing values were imputed using the k-nearest neighbour algorithm with k = 24 through the preProcess' function (method = "knnImpute") from the caret R package (6.0–94 version) [27]. Subsequently, the metabolomics data were $\log_{10}$ transformed.

2.3. Proteomics data acquisition and preprocessing

Plasma protein levels were assessed in plasma using the SOMAscan assay platform (SomaLogic Inc.). This aptamer-based assay enabled the simultaneous measurement of up to 3630 proteins [28]. Samples were grouped and measured separately. Proteins with more than 20% of missing values were not detected in these datasets, therefore, all proteins were kept for downstream analysis. Remaining missing values were imputed using the k-nearest neighbour algorithm with k = 60 through the preProcess function (method = "knnImpute") from the caret 6.0–94 R package. Subsequently, the proteomics data were $\log_{10}$ transformed. Batch effects were corrected separately for the validation cohort (the EDAR cohort) and the rest of the cohorts using the 'ComBat' function from the sva R package (3.46 version) [29].

2.4. Number of subjects for multiclass models and MCI conversion models

Our aim was to discover the most relevant clinical characteristics, proteins and metabolites involved in two different tasks: (1) classifying samples into NC, MCI and AD donors and (2) distinguishing between sMCI and cMCI. In both approaches, clinical covariates were included in the models. Although 26 clinical covariates were selected, 14 of them were discarded because of a high percentage of missing values (>20%), including CSF markers, neuroimaging markers and cognition tests. The clinical variables maintained for downstream analysis included: years of education (Eduy), age, gender, amyloid status (CSF $A\beta_{42/40} < 0.061$ determines abnormality), Aβ z-score, APOE dich, local p-Tau, local t-Tau, mini-mental state examination (MMSE), priority attention z-score, priority language z-score and priority memory immediate z-score.

First, to identify the most relevant proteins and metabolites implicated in the classification of NC, MCI and AD donors, proteins and metabolites were treated separately. On the one hand, multiclass models of proteins were created using the levels of 3630 proteins and 12 clinical covariates as predictors. To create multiclass models of proteins, data from the EDAR cohort was reserved for validation purposes (NC = 21, MCI = 44, AD = 65) while data from the other cohorts was randomly split into train (NC = 147, MCI = 240, AD = 84) and test (NC = 62, MCI = 102, AD = 35) in 70/30 proportion. On the other hand, multiclass models for metabolites were created using as predictors the levels of 540 metabolites and the same 12 clinical covariates. To create those models, the EDAR cohort was reserved for validation purposes (NC = 29, MCI = 18, AD = 32) and the remaining data was randomly split into train (NC = 119, MCI = 181, AD = 73) and test (NC = 51, MCI = 77, AD = 31) in 70/30 proportion. Statistics of the clinical characteristics of this dataset have been previously published by Shi et al. [30]. Although subjects used for proteomics and metabolomics models are not identical, most of the samples overlap. Specifically, 508 samples have paired data, 103 samples only have metabolomics and 292 only have proteomics.

Then, MCI conversion models were created to specifically identify the key molecules to distinguish between sMCI and cMCI. To create MCI conversion models, 386 MCI participants were initially available in the EMIF-AD dataset. First, only MCI participants with follow-up information were kept to create the models, leading to 319 MCI participants (cMCI = 100, sMCI = 219). Additionally, in the next step, only participants with both proteomics and metabolomics data available were kept to create MCI conversion models, leading to 194 MCI participants with follow-up information and both proteomics and metabolomics data available (cMCI = 91, sMCI = 103). The descriptive statistics of clinical characteristics for this dataset are reported in Supplementary Table 2. To build MCI conversion models, clinical features (n = 12), proteomics (n = 3630) and metabolomics (n = 540) were concatenated in the same matrix. In this scenario, subjects from the EDAR cohort were not reserved for validation purposes since it did not include enough samples to this end (sMCI = 5, cMCI = 6). Instead, all MCI participants with follow-up information and paired metabolomics and proteomics data were used for the hyperparameter tuning of the models (sMCI = 103, cMCI = 91).

2.5. Machine learning-based pipeline

A machine learning-based pipeline was developed to create both multiclass models of proteins, multiclass models of metabolites and MCI conversion models, slightly adapted to account for each problem's particularities. The caret R package functions (6.0–94 version) were used to deal with different ML algorithms and the typical ML tasks, including data preprocessing, model tuning, and performance evaluation. Firstly, control parameters to train the models were defined in the 'trainControl' function from the caret R package. For multiclass models, a 10-folds stratified cross-validation repeated three times was applied.

Instead, for MCI conversion models, a 3-folds stratified cross-validation repeated 10 times was applied since a low number of folds implies larger validation sets and lower variance across estimates for classifier performance. Then, for multiclass models, imbalanced class distribution within cross-validation was dealt with a downsampling approach on the majority class. For MCI conversion models, classes were already quite balanced (sMCI = 103, cMCI = 91), therefore, class balancing was not of application. To create the models, four ML algorithms belonging to four different algorithm families were proposed: logistic regression (LR), random forest (RF), support vector machines (SVM) and multi-layer perceptron network (MLP). These algorithms are implemented within caret with the glmnet (glmnet library 4.1–8), rf (randomForest library 4.7–11), svmLinear (kernlab library 0.9–32) and mlpWeightDecay (RSNNS library 0.4–17) packages, respectively. For each algorithm, hyperparameter tuning and model selection was performed using training samples and producing performance estimates based on cross-validation for each hyperparameter value. First, a hyperparameter grid was designed to explore a wide area of the hyperparameter space with large distances between successive values for each hyperparameter. It helps in focusing on a specific area of that space. After this area was selected, a more specific grid was designed only to cover that area. Balanced accuracy was used to select the best model within multiclass models to avoid biases introduced by imbalance between the different classes and the AUC (the Area Under the ROC curve) was used for MCI conversion models as it is best suited for binary classification tasks. To assess whether the MCI conversion binary classifier was actually learning, we used a binomial test to check whether the accuracy (AC) is higher than the no information rate (NIR) i.e., the proportion of elements from the majority class. The test is implemented in the confusionMatrix method from caret. In addition, to evaluate classifiers performance for each class, precision recall curves were also created, appropriate for imbalance classes. These curves were built through the 'pr_curve' function from the yardstick R package (1.3.0 version) [31].

2.6. Feature importance

All the four algorithm families contributed to the final multiclass models of proteins and those of metabolites with the 20 most relevant features selected by their corresponding best model. Such relevance was assessed for all four algorithms with a permutation-based approach [32,33]. It is based on the assumption that if a variable is important, after permuting the values of this variable, the model's performance will worsen. The larger the change in model performance, the more important the variable will be. Thanks to being a model-agnostic method, it can be used within all algorithms and compare predictors' relevance across models. Since the permutation feature importance relies on measurements of the model error, permutation was applied on unseen test data. The pipeline uses the 'explain' function followed by 'model_parts' function both from the DALEX R package (2.4.3 version) to estimate variable importance on the test set as a change in loss function (loss cross entropy in this case) after variable permutations. Each feature is permuted 10 times by default. After performing the permutation analysis within each technique, no bias towards uncorrelated features was detected in all the four cases.

For MCI conversion models, the permutation-based approach was not of application since all of the MCI participants with follow-up information and both metabolomics and proteomics data available were used for the hyperparameter tuning and permutation feature importance requires estimates of model error produced on unseen examples. Therefore, model-specific methods were used instead through the 'varImp' function from the caret R package. Finally, once detected the top 20 most relevant features for each algorithm, the features selected in at least 3/4 algorithms were kept for downstream analysis. To demonstrate the reliability of the selected features, a logistic regression model is trained with selected features as predictors and tested with a validation cohort.

2.7. Univariate analysis

Selected molecules level distribution was pairwise compared between diagnostic groups with Wilcoxon test (two-side test) through the wilcox.test function from the stats R package (4.2.2. version). Differences in proteins or metabolites levels adjusting for age and sex covariates were tested with an analysis of covariance (ANCOVA) using the 'aov' function from the stats R package (4.2.2. version).

2.8. Correlation network

The features selected as relevant by at least three of the four algorithms from MCI conversion models were included in a correlation network with MCI conversion as the target. Correlations were estimated using the 'cor_auto' function from the 'qgraph' R package (1.9.8 version) [34]. The correlation network was represented using the 'qgraph' function from this same package.

2.9. Further studies for selected proteins

The metabolites and proteins repeatedly selected were further studied. The metabolite selected by all the algorithms was further studied with in-vitro experiments (see Microglia in-vitro experiments section). Instead, proteins selected by all algorithms were further investigated in the literature to see if they had been previously associated with AD-phenotype. To this end, two different resources were used: 1) a systematic review from Kiddle, S. J. et al. (2014) [35], where a list of 21 published discovery or panel-based blood proteomics studies of AD was reviewed and 2) Agora database, a web application that hosts high-dimensional human transcriptomic, proteomic, and metabolomic evidence for whether or not genes are associated with Alzheimer's disease (https://agora.adknowledgeportal.org/).

2.10. Microglia in-vitro experiments

2.10.1. Rodent microglial culture preparation and extracellular vesicles isolation

Pure murine and rat primary microglia were obtained as established in Gabrielli et al. [36]. Independent experiments were performed as independent cell preparations and from those several replicates were collected. Microglia have been stimulated with 1:20 Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF) from murine GM–CSF–transfected X63 cells [37]. GM-CSF is a member of the colony-stimulating factor superfamily that induces microglial proliferation, migration and upregulation of surface markers [38]. Supernatants have been cleared from cell debris before storing. Cells have been scraped in physiological solution, pelleted and stored in methanol. Extracellular vesicles (EVs) have been isolated through differential centrifugation from the cell supernatant upon 30 min ATP stimulation at 110,000×g [39] and stored at −80C°.

2.10.2. Isolation of EVs from the plasma of AD patient and control group

Plasma samples were collected from five subjects with a diagnosis of Alzheimer's disease (75.6 ± 2.7 years, 2/3 males/females ratio) and five healthy controls (65.6 ± 3.7 years, 3/2 males/females ratio). All participants or their representatives provided informed written consent following the protocol approved by the Ethical Committee of Fondazione Don Carlo Gnocchi, according to the declaration of Helsinki. Inclusion criteria for AD considered: diagnosis of AD according to NIA-AA criteria; mild dementia stage as documented with a Clinical and Dementia Rating (CDR) scale score between 0.5 and 1; absence of psychiatric or systemic illness. To have a global index of cognitive functioning, all subjects performed cognitive evaluation using the Montreal Cognitive Assessment test (MoCA). Human peripheral blood samples of each subject were collected in EDTA-treated tubes (BD Vacutainer, Becton Dickinson). To isolate the plasma, samples were

centrifuged at centrifugation at 1300*g* × 10 min for removing cells and then at 1800g × 10 min for the depletion of platelets. Plasma aliquots were anonymized, aliquoted and stored at −80 °C in the biorepository of the Laboratory of Nanomedicine and Clinical Biophotonics of Fondazione Don Carlo Gnocchi (Milan, Italy) until further use. 500 μl of plasma samples were then thawed and centrifuged at 10,000 g for 10 min and then used for EVs isolation by size exclusion chromatography (SEC; qEV, Izon, Christchurch, New Zealand), following manufacturer's instructions. Eluted fractions from 6 to 8 containing EVs in PBS were retained, added with 0.1% DMSO and stored at −20 °C until further analysis.

3. Results

3.1. Machine learning models performance per data modality and algorithm

Multiclass models were created to identify the key molecules that best distinguish amongst NC, MCI and AD individuals. To this end, four different ML algorithms were applied on plasma samples in two different scenarios: (1) metabolites and clinical covariates and (2) proteins and clinical covariates. For multiclass models of metabolites and clinical data, LR and RF showed the highest performance on the test set (0.819 and 0.821 mean balanced accuracy, respectively) (see Supplementary Table 3). Then, SVM reported a slightly lower performance followed by MLP (0.773 and 0.747 mean balanced accuracy, respectively) (Supplementary Table 3). Interestingly, a very similar trend was observed for multiclass models of proteins, where the number of variables is more than 6 times the number of variables in multiclass models of metabolites (3630 vs 540 features). LR reached the highest performance on the test (0.819 mean balanced accuracy) followed by RF (0.784), SVM (0.765) and finally MLP (0.718) (see Supplementary Table 3). Note that the LR is a linear model with L1 regularization and RF is naturally suited for high-dimensional problems due to its random choice of predictors when splitting trees within the forest. MLP is not specially prepared for this type of problems when the number of samples is low. This is not the case for SVM which, besides, it deals with high-dimensional problems in a natural way. However, the variability induced by the choice of kernel function may inject a high variability in its behaviour.

For both metabolomics and proteomics approaches, experiments reported the AD class as showing the lowest precision rate for SVM and MLP algorithms (see Supplementary Fig. 1). This comes from a high proportion of false positives, mostly MCI participants classified as AD subjects. When comparing multiclass models performance using metabolomics with those using proteomics data, LR experiments reported identical performance for the test set (0.819 and 0.819 test mean balanced accuracy for metabolites and proteins, respectively). On the other hand, RF performs better with metabolomics than with proteomics (0.821 vs 0.784 test mean balanced accuracy, respectively).

The MCI conversion problem was also addressed by creating models from clinical and paired metabolomics and proteomics data from MCI participants with follow-up information. These models reported a mean AUC of 0.64 across the different algorithms (see Supplementary Table 4). Among them, RF stood out with the highest performance within cross-validation with a 0.66 AUC (*P-value* [Acc > NIR] < 0.0225). Nevertheless, all of the algorithms showed a very similar performance. The final hyperparameter grid used for each algorithm and data modality is available in Supplementary Table 5.

3.2. Key molecules to classify NC, MCI and AD participants

For multiclass models of metabolites and clinical variables, the top 20 most relevant features were extracted per algorithm with a permutation-based approach (Supplementary Table 6). Then, the overlap between the top 20 most predictive clinical and molecular features of each algorithm was represented in a Venn Diagram for metabolites models (Fig. 1A). Only features selected by 3/4 algorithms were kept for downstream analysis since we considered them universal (i.e. not exclusive to a particular algorithm), suggesting more biological relevance. Features selected by all algorithms in multiclass models of metabolites included four clinical features (MMSE, priority attention z-score, priority language z-score and priority memory immediate z-score) and one metabolite, oleamide. Additionally, one clinical covariate and three metabolites were selected by 3/4 algorithms: Eduy (years of education), aspartate, serotonin and linoleamide. It is not surprising that oleamide and linoleamide were correlated since both molecules belong to the fatty acid primary amides, of which oleamide is the most explored member. At molecular level, both oleamide and linoleamide were classified as modulators of intracellular Ca^{2+} homeostasis via regulation of SERCA activity [40]. Additionally, both molecules are considered endogenous sleep-inducing lipids [41,42]. In the dataset of study, oleamide levels were detected to be increased in sMCI subjects compared to NC (Wilcoxon test $P < 8.721 \cdot 10^{-10}$, W = 7793) and further increased in cMCI participants compared to sMCI (Wilcoxon test $P < 2.696 \cdot 10^{-4}$, W = 4548) (Fig. 1C). However, oleamide levels decreased in AD donors compared to cMCI donors (Wilcoxon test $P < 1.593 \cdot 10^{-8}$, W = 9640.5) and AD donors reached similar oleamide levels compared to NC donors. Changes in oleamide levels were significantly associated with diagnosis status after adjusting for age and sex covariates (ANCOVA $P < 2 \cdot 10^{-16}$). These same models showed that sex had no influence on oleamide levels while age had a significant effect (ANCOVA $P < 1.22 \cdot 10^{-8}$).

To validate the molecules selected by 3/4 algorithms (five clinical covariates and four metabolites), these features were used as predictors in a final model, whose experimentation reported a very good performance in both test and validation sets (0.845 and 0.726 mean balanced accuracy, respectively) (Table 1). The good mean balanced accuracy obtained in validation suggests these molecules are effective in distinguishing NC, MCI and AD participants within independent cohorts to those used in training. Finally, to demonstrate that metabolites provide additional value in the model in comparison with clinical variables, a second model was created using as predictors only clinical covariates selected (five covariates). This model reported a lower performance for test and validation sets (0.746 and 0.688 mean balanced accuracy, respectively) compared with the model including both selected clinical covariates and selected metabolites (0.845 and 0.726 mean balanced accuracy, respectively) (Table 1). Results showed that clinical covariates selected reported a good performance to classify NC, MCI and AD but in combination with selected metabolites, a significant improvement in the performance is observed ($P_{adj} < 6.74 \cdot 10^{-5}$). From this point, only metabolites selected by all the algorithms were further characterized with additional experiments.

For multiclass models created with proteins, the same approach was repeated. For each algorithm, the top 20 features were extracted. Then, the overlap between the features selected by each algorithm was represented in a Venn Diagram (Fig. 1B). In this case, five clinical covariates were selected by all the algorithms, including Aβ z-score, amyloid status, MMSE, priority attention z-score and priority language z-score. However, none of the proteins were selected by all the algorithms. Additionally, three other clinical variables and one protein were selected by 3/4 algorithms, including local p-Tau, local t-Tau, priority memory immediate z-score and lysosomal alpha-glucosidase protein (GAA, P10253). The features selected by 3/4 algorithms, eight clinical features and one protein, were further validated in an independent cohort. The model trained with these features as predictors reported a very good performance on test and validation sets (0.777 and 0.720 mean balanced accuracy, respectively), which shows the confidence of these features to classify NC, MCI and AD participants (see Table 1). To show that the protein selected, GAA, provides an additional value in this model, a second model was trained only with the clinical covariates selected (eight covariates). This model reported a lower performance on test and validation cohorts (0.757 and 0.703 mean balanced accuracy, respectively) compared with the model including also GAA protein as predictor

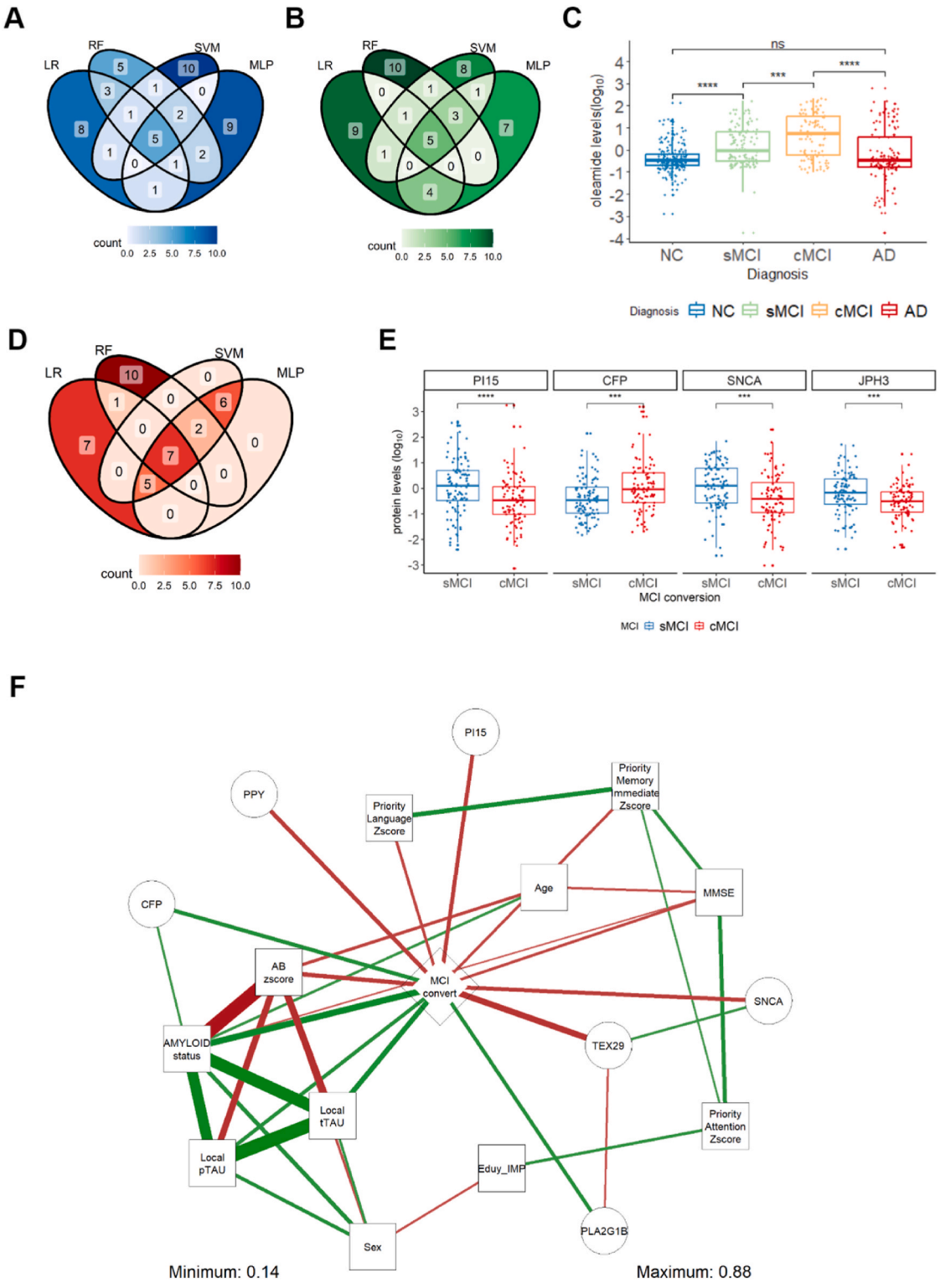

Fig. 1. Most relevant clinical features, proteins and metabolites extracted with multiclass models and MCI conversion models. Venn diagram shows the overlap of the top 20 most relevant predictors of each algorithm for **(A)** multiclass models of proteins, **(B)** multiclass models of metabolites and **(C)** MCI conversion models of proteins and metabolites; **(D)** Oleamide level distribution is represented for NC (n = 203), sMCI (n = 128), cMCI (n = 99) and AD (n = 136) donors. Differences in oleamide levels between pairs of groups were estimated using a Wilcoxon test. **(E)** Differences in the protein levels between sMCI (n = 291) and cMCI (n = 100) were estimated using a Wilcoxon test. Only relevant proteins that differentiate MCI stable and converters that are expressed mainly in the brain or in bone marrow lymphoid tissues were included (based on Human Protein Atlas dataset). **(F)** Correlation network including the variables selected as relevant by at least three of the four algorithms proposed in all the approaches: multiclass models for proteins, multiclass models for metabolites and MCI conversion models. Pearson correlations were estimated using the sMCI (n = 103) and cMCI (n = 91) donors with paired data (proteomics and metabolomics). Only significant correlations ($P < 0.05$) with a magnitude over $|0.2|$ are represented. Positive correlations were represented with green color, negative correlations with red. The color saturation and the width of the edges corresponds to the absolute weight and scale relative to the strongest weight in the graph. Features were grouped into target, clinical, proteins or metabolites. ns, non-significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Table 1
Validation of selected features from multiclass models of proteins and multiclass models of metabolites. Features selected by at least 3/4 algorithms of multiclass models of metabolomics were validated in a final model. The performance of this model was compared with another model created only with selected clinical features. The same procedure was repeated for proteomics data.

Omics data	Selected features	Cross-validation (mean balanced accuracy ± SD)	Test (mean balanced accuracy)	Validation (mean balanced accuracy)
Metabolites	Clinical covariates	0.763 (0.066)	0.746	0.688
	Clinical covariates and metabolites	0.810 (0.052)	0.845	0.726
Proteins	Clinical covariates	0.803 (0.036)	0.757	0.703
	Clinical covariates and proteins	0.819 (0.036)	0.777	0.720

(0.777 and 0.720 mean balanced accuracy, respectively) (Table 1). GAA significantly improves the model performance ($P_{adj} < 0.0442$), providing an additional value in the models.

3.3. Key molecules to distinguish between MCI stable and converter

The overlap of the top 20 most relevant features of the four models for the MCI converter scenario was represented with a Venn Diagram (Fig. 1D–Supplementary Table 6). Seven features were selected as relevant in all models including: local tTau as clinical variable and Alpha-synuclein (SNCA), Properdin (CFP), Peptidase inhibitor 15 (PI15), pancreatic hormone (PPY), phospholipase A2 (PLA2G1B) and testis-expressed sequence 29 protein (TEX 29). In addition, proteins selected by three algorithms included local pTau as clinical feature Junctophilin-3 (JPH3) among other proteins (for more information see Supplementary Table 6). Interestingly, none of these clinical features or proteins were also selected as relevant by the multiclass models. Among the proteins selected in at least 3 of the 4 algorithms, the ones expressed mainly in the brain or in bone marrow lymphoid tissues were highlighted in Fig. 1E, which includes SNCA, CFP, PI15 and JPH3. Specifically, in our small cohort of MCI participants with follow-up information and paired metabolomics and proteomics data, CFP was increased in cMCI participants compared to sMCI donors ($P < 6.99 \cdot 10^{-4}$, W = 6010) while JPH3 ($P < 6.072 \cdot 10^{-4}$, W = 3348), PI15 ($P < 9.625 \cdot 10^{-5}$, W = 3164) and SNCA ($P < 1.672^{-4}$, W = 3217) were decreased in plasma (Fig. 1E). After adjusting for age and sex covariates, significant differences between cMCI and sMCI participants were still present for PI15 (ANCOVA $P < 5.96 \cdot 10^{-4}$), SNCA ($P < 7.08 \cdot 10^{-4}$), JPH3 ($P < 9.91 \cdot 10^{-4}$) and CFP ($P < 6.39 \cdot 10^{-4}$) proteins. Sex had no influence on the levels of these proteins while age showed a significant effect on SNCA protein levels ($P < 0.0305$).

To investigate the prediction ability of the features selected repeatedly by 3/4 algorithms of MCI conversion models, a final model was created with these features as predictors. This model showed a mean AUC of 0.872 within the cross-validation (Table 2). To demonstrate that molecules selected provide an additional value in this model compared to the clinical covariates, a second model only with selected covariates was created. Significant differences in the performance of these two models were observed ($P_{adj} < 3.77 \cdot 10^{-15}$). Specifically, the model created with only clinical covariates showed a mean AUC of 0.649 and the model created with both selected clinical covariates and proteins showed a mean AUC of 0.872, which suggests that selected molecules, proteins in this case, contribute significantly to this model (Table 2). Additionally, proteins selected as relevant by all the algorithms for MCI conversion, except PLA2G1B, have been previously associated with AD-related phenotype. In more detail, these proteins were found to have brain eQTL and RNA expression changes in the AD brain (Supplementary Table 7).

Table 2
MCI conversion models performance using selected clinical features and molecules.

Omics data	Selected features	Cross-validation (ROC ± SD)
Clinical features, metabolites and proteins	Clinical covariates	0.649 (0.062)
	Clinical covariates and molecules	0.872 (0.036)

Finally, the relationship between MCI participants' clinical features and selected proteins were represented in a correlation network, with MCI conversion as the target (Fig. 1F). All the proteins correlated with MCI conversion with a very similar strength (0.33 mean correlation in absolute value). As observed, SNCA, TEX29, PPY and PI15 were decreased in cMCI (negative correlation shown in red) while PLA2G1B and CFP were increased in cMCI (positive correlation shown in green). Additionally, the results showed that cMCI had relatively higher correlation with brain biomarkers such as amyloid status (0.430 correlation) and local t-Tau (0.372 correlation) compared to sMCI while MMSE score was lower for cMCI individuals compared to sMCI individuals (−0.261 correlation).

3.4. Oleamide detected within microglia vesicles in-vitro

Oleamide metabolism was flagged as interesting in the multiclass models. Since another endocannabinoid, anandamide, had been detected within microglia and their vesicles [19–21] we were motivated to follow up this result. Hence, oleamide on microglia cultures was measured and in their secreted EVs (supplementary methods, Supplementary Table 5). Mice microglia cultures showed that supernatant microglia, both for unstimulated and activated microglia, had very low oleamide concentrations (0.006 and 0.201 μg, respectively) (Supplementary Fig. 2A). Then, oleamide concentration was increased in unstimulated microglia and further increased in activated microglia (1.191 and 2.24 μg, respectively). The highest concentration of oleamide was found in activated microglia EVs (12.62 μg). Once demonstrated that oleamide was synthesized and secreted in mice microglial EVs, oleamide was further quantified in microglial EVs from other organisms. A slightly lower amount of oleamide was observed in rat microglia EVs (5.733 μg) compared to mice (8.986 μg) (Supplementary Fig. 2B). In addition, in a small pilot study in individuals with AD and control subjects, higher oleamide concentration was measured in plasma EVs from AD participants (4.688 μg) compared with control subjects (0.529 μg) (Wilcoxon test $P < 0.05$) (Supplementary Fig. 2C).

4. Discussion

The use of omics data to uncover disease-specific pathways has gained great interest in recent years. Omics data represent complex molecular interactions, many following non-linear relationships, hence machine learning can help understanding these relationships. In this study, we used the EMIF-AD MBD dataset and applied four different ML algorithms to identify key molecules involved in the conversion of MCI to AD. Further molecular exploration was applied on the molecules consistently selected by four algorithms.

4.1. Models performance

Multiclass models showed a pretty good performance to distinguish between NC, MCI and AD participants using both metabolomics (0.790 mean balanced accuracy on test) and proteomics data (0.772 on test) (Table 1). For metabolomics data, RF stood out from the rest of the algorithms with the highest performance on the test set (0.821 mean

balanced accuracy), followed by LR (0.819). A similar trend was observed for proteomics data, where LR showed the highest performance on the test set (0.819 mean balanced accuracy) followed by RF (0.784). The model's performance was in line with previous work of Stamate et al. [43], where a cohort with fewer individuals and groups (115 AD donors and 242 NC donors) was used to test the performance of plasma metabolites to categorize AD when compared to CSF biomarkers. To this end, different machine learning algorithms were used, including deep learning, which produced an AUC value of 0.85, extreme gradient boosting, which reported a 0.88 and the RF model, which resulted in a 0.85 AUC value.

4.2. Multiclass models selected molecules and clinical features

~~A~~lthough MMSE, priority language z-score and priority attention z-score were selected as key clinical covariates to classify NC, MCI and AD participants, these variables did not play an important role to distinguish between MCI stable and converters. Instead, Tau-related measures (p-Tau and t-Tau) were shown to be the most relevant for MCI conversion. Additionally, known risk factors for AD such as sex, age or years of education were not chosen consistently by the models (see Supplementary Table 6). Regarding the metabolites, oleamide was selected as a relevant predictor by all algorithms for the classification of NC, MCI and AD participants. Oleamide is a molecule thought to be synthesized in the brain to aid with sleep [44] and is a potent endogenous sedative endocannabinoid [41]. In a previous study with fewer participants, oleamide was associated with elevated amyloid levels in MCI participants [26]. Xie et al. [45] showed that sleep helps with molecular clearance in the brain, hence we could hypothesise that if oleamide aids with sleep in early AD, oleamide synthesis could be a coping mechanism to help clear misfolded protein. Furthermore, multiclass models selected other three metabolites by three of the four algorithms: linoleamide, a relatively unknown fatty amide when compared to oleamide, and aspartate and serotonin, both neurotransmitters with excitatory and inhibitory functions, respectively. Previous studies have shown that these neurotransmitters were decreased in the brain and were associated with Alzheimer's disease [46]. Finally, multiclass models of proteins highlighted one protein, lysosomal alpha-glucosidase (GAA) by 3/4 algorithms. GAA is commonly associated with hyperglycemia, which is a risk factor for Alzheimer's disease [47].

4.3. MCI conversion models selected proteins

Four key proteins were selected by all the algorithms in MCI conversion models: local tTau as clinical variable and Alpha-synuclein (SNCA), Properdin (CFP), Peptidase inhibitor 15 (PI15), pancreatic hormone (PPY), phospholipase A2 (PLA2G1B) and testis-expressed sequence 29 protein (TEX 29). Decreased levels of PI15, SNCA and JPH3 were found in the plasma of cMCI participants compared to sMCI individuals (Fig. 1E) while CFP followed the opposite trend. The JPH3 protein is particularly interesting because it is a neuron specific protein. JPH3 controls electrical excitability of neurons in different brain regions and is involved in the regulation of intracellular calcium signaling. JPH3 has been previously linked to Huntington-like disease-2 [48]. CFP belongs to the complement system, a well-established pathway of inflammation. This supports the hypothesis that inflammation in the brain exacerbates the progression of MCI. The last protein selected by all algorithms, PI15, is an inhibitor against trypsin and, although unexplored, it may play a role in protein degradation in the central nervous system.

4.4. Further experiments with oleamide in microglia and EVs

Microglia are known to release a variety of signaling molecules that impact synaptic transmission in response to injury or inflammation, playing a crucial role in maintaining balance in neuronal networks [49]. Stella N [50]. demonstrated that microglia produce *in-vitro* 20-fold higher amounts of endocannabinoids (ECs) than neurons or astrocytes, likely representing the main source of ECs in the inflamed brain. In this regard, Gabrielli, M. et al. [51] demonstrated that ECs are secreted by microglia through extracellular membrane vesicles~~.~~ In addition, ECs have been linked to learning, memory and long-term plasticity [52–54] and can be neuromodulator lipids [55]. All this evidence motivated us to carry out *in-vitro* experiments to quantify oleamide concentration within rodent microglia. Results showed that oleamide was present in activated microglia and enriched in EVs released in the pericellular space. In a small pilot study with five individuals with AD and five controls, we observed that EVs from blood contained oleamide in their cargo and oleamide was more abundant in persons with AD.

5. Limitations

This study had the advantage of having MCI longitudinal data that allowed us to investigate the molecular pathways involved in MCI progression to AD. Although two different criteria were applied to diagnose MCI participants across the different cohorts, both criteria were considered equivalent since the scheme that Petersen proposed to diagnose MCI subtypes was later adapted in the Winblad et al. report. Additionally, this dataset also has CSF biomarkers data as well as extensive clinical covariates that helped with the interpretation of the results. Nevertheless, cMCI participants showed higher CSF tau and amyloid concentrations compared to sMCI participants, which can lead to biases in the molecules selection process. In addition, only 194 MCI participants with follow-up information and both proteomics and metabolomics data were available to create MCI conversion models. Because of the low number of samples available and the high heterogeneity of MCI participants to convert to AD, all of the samples were used for the hyperparameter tuning, giving us more power to find the best molecules. In a future release of the EMIF-AD with more MCI participants data available, a subset of samples would be left out of the training and model selection process for testing the model in order to detect any bias in the training data. Finally, another limitation of this study is the collinearity between the variables, which is a common problem when dealing with omics data. To address this problem, algorithms from four different families were applied, each one dealing with correlated variables in a different way. Then, only features selected as relevant by all the algorithms were kept for downstream analysis since they were considered more universal. Selected features were validated in an independent cohort to demonstrate their reliability. In this regard, oleamide was one of the molecules selected by all the algorithms. Further biological characterization of this molecule was carried out in microglia primary culture, which demonstrated that oleamide was secreted in EVs from microglia, the first time this mechanism is shown *in-vitro*. However, *in-vitro* experiments are deemed an artificial condition compared with microglia *in-vivo*.

6. Conclusions

In this study, we identified key molecules involved in MCI progression to AD with a machine learning-based pipeline. One the one hand, four different algorithms identified oleamide as a key molecule to classify NC, MCI and AD participants. This molecule is linked to sleep and memory. *In-vitro* experiments showed that oleamide is secreted by microglia via EVs. Additionally, a pilot study in humans showed a higher plasma concentration of oleamide in people with AD compared with healthy controls. Moreover, key proteins to predict MCI conversion to AD included phospholipase A2 (PLA2G1B), properdin (CFP), alpha-synuclein (SNCA) and junctophilin-3 (JPH3), all of them expressed mainly in the brain or in bone marrow lymphoid tissues. In relation to AD, neural protein JPH3 is a novel potential target together. The ML-based pipeline also confirmed established proteins such as synuclein (SNCA) and protein activators in the complement cascade (CFP).

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by Aristotle University of Thessaloniki Medical School Ethics Committee; Ethics Committee of the Medical Faculty Mannheim, University of Heidelberg; Ethic and Clinical Research Committee Donostia; Ethics Committee Inserm and Aix Marseille University; The Healthcare Ethics Committee of the Hospital Clínic; Central Clinical Research and Clinical Trials Unit (UICEC Sant Pau); INSERM Ethical Committee; Ethic Committee of the IRCCS San Giovanni di Dio FBF; Comitato Etico IRCCS Pascale - Napoli; Ethics Committee at Karolinska Institutet; Ethische commissie onderzoek UZ/KU Leuven; Research Ethics Committee Lausanne University Hospital; Medical Ethical Committee Maastricht University Medical Center; Committee on Health Research Ethics, Region of Denmark; Ethics Committee of Mediterranean University; University of Lille Ethics Committee; Ethical Committee at the Medical Faculty, Leipzig University; Ethical Committee at the Medical Faculty, University Hospital Essen; Ethics Committee University of Antwerp; Ethical Committee of University of Genoa; Ethics Committee, University of Gothenburg; Human Ethics Committee of the University of Perugia; and Medical Ethics Committee VU Medical Center. The patients/participants provided their written informed consent to participate in this study.

Code availability

The code used to preprocess the EMIF-AD proteomic and metabolomic data as well as the code to create the machine learning models and represent the results are available on the ML-multiomics repository on GitHub (https://github.com/aliciagp/ML-multiomics).

Data availability

The averaged protein and metabolites plasma levels for each diagnosis group (NC, sMCI, cMCI and AD) have been deposited on GitHub (https://github.com/aliciagp/ML-multiomics). Patient data are available upon request by contacting the EMIF-AD data hub steering committee via the academic EMIF-AD lead, Prof. Pieter Jelle Visser, and data access coordinator Dr. Stephanie Vos (s.vos@maasticuniversity.nl) since EMIF-AD is considered sensitive patient data under GDPR. Requests should include the following information, name and contact details of the person requesting the data, aims and objectives, study design and methods, requested molecular data and clinical variables, outcomes, timelines and cohorts of interest in the EMIF-AD catalogue (https://emif-catalogue.eu/login). Requests will be subject to consideration by the steering committees of the cohorts and the management board. Time frame for a response will be within 4 months. Data requests under agreement will be subject to appropriate confidentiality obligations and restrictions.

CRediT authorship contribution statement

Alicia Gómez-Pascual: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Talel Naccache:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Jin Xu:** Writing – review & editing. **Kourosh Hooshmand:** Writing – review & editing, Investigation. **Asger Wretlind:** Writing – review & editing, Visualization. **Martina Gabrielli:** Writing – review & editing, Validation, Resources, Project administration, Investigation. **Marta Tiffany Lombardo:** Writing – review & editing, Validation, Resources, Project administration, Investigation. **Liu Shi:** Writing – review & editing. **Noel J. Buckley:** Writing – review & editing. **Betty M. Tijms:** Writing – review & editing. **Stephanie J.B. Vos:** Writing – review & editing. **Mara ten Kate:** Writing – review & editing. **Sebastiaan Engelborghs:** Writing – review & editing. **Kristel Sleegers:** Writing – review & editing. **Giovanni B. Frisoni:** Writing – review & editing. **Anders Wallin:** Writing – review & editing. **Alberto Lleó:** Writing – review & editing. **Julius Popp:** Writing – review & editing. **Pablo Martinez-Lage:** Writing – review & editing. **Johannes Streffer:** Writing – review & editing. **Frederik Barkhof:** Writing – review & editing. **Henrik Zetterberg:** Writing – review & editing. **Pieter Jelle Visser:** Writing – review & editing. **Simon Lovestone:** Writing – review & editing. **Lars Bertram:** Writing – review & editing. **Alejo J. Nevado-Holgado:** Writing – review & editing. **Alice Gualerzi:** Writing – review & editing, Validation, Resources, Project administration, Investigation. **Silvia Picciolini:** Writing – review & editing, Validation, Resources, Project administration, Investigation. **Petroula Proitsi:** Writing – review & editing. **Claudia Verderio:** Writing – review & editing, Validation, Resources, Project administration, Investigation. **Juan A. Botía:** Writing – original draft, Supervision, Conceptualization. **Cristina Legido-Quigley:** Writing – original draft, Supervision, Conceptualization.

Declaration of competing interest

SL is named as an inventor on biomarker intellectual property protected by Proteome Sciences and Kings College London unrelated to the current study and within the past five years has advised for Optum labs, Merck, SomaLogic and been the recipient of funding from AstraZeneca and other companies via the IMI funding scheme. HZ has served at scientific advisory boards and/or as a consultant for Abbvie, Alector, ALZPath, Annexon, Apellis, Artery Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen, Novo Nordisk, Pinteon Therapeutics, Red Abbey Labs, reMYND, Passage Bio, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in symposia sponsored by Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work). AL has served at scientific advisory boards of Fujirebio Europe, Eli Lilly, Novartis, Nutricia and Otsuka and is the inventor of a patent on synaptic markers in CSF (all unrelated to this study). JP has served at scientific advisory boards of Fujirebio Europe, Eli Lilly and Nestle Institute of Health Sciences, all unrelated to this study. SE has received unrestricted research grants from Janssen Pharmaceutica and ADx Neurosciences and has served at scientific advisory boards of Biogen, Eisai, Novartis, Nutricia/Danone, all unrelated to this study. FB is a steering committee or iDMC member for Biogen, Merck, Roche, EISAI and Prothena. Consultant for Roche, Biogen, Merck, IXICO, Jansen, Combinostics. Research agreements with Merck, Biogen, GE Healthcare, Roche. Co-founder and shareholder of Queen Square Analytics LTD, all unrelated to this study. CLQ has received funding related to this study from Pfizer and via the IMI funding scheme, in the past five years has been the recipient of funding from Novo Nordisk, unrelated to this study.

Acknowledgements

This research was conducted as part of the EMIF-AD MBD project which has received support from the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement [no 115372], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme [FP7/2007–2013] and EFPIA companies' in-kind contribution. The DESCRIPA study was funded by the European Commission within the 5th framework program [QLRT-2001-2455]. The EDAR study was funded by the European Commission within the 5th framework program [contract # 37670]. The Leuven cohort was funded by the Stichting voor Alzheimer Onderzoek [grant numbers #11020, #13007 and #15005]. RV is a senior clinical investigator of the Flemish Research Foundation (FWO). The San Sebastian GAP study is partially funded by the Department of Health of the Basque Government [allocation 17.0.1.December 08, 0000.2.454.01.41142.001.H]. We acknowledge the contribution of the personnel of the Genomic Service Facility at the VIB-U Antwerp Center for Molecular Neurology. The research at VIB-CMN

is funded in part by the University of Antwerp Research Fund. HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council [#2018–02532], the European Research Council [#681712], Swedish State Support for Clinical Research [#ALFGBG-720931], the Alzheimer Drug DiscoveryFoundation (ADDF), USA [#201809–2016862], and the UK Dementia Research Institute at UCL. FB is supported by the NIHR biomedical research centre at UCLH. LS is funded by the Virtual Brain Cloud from European commission [grant no. H2020-SC1-DTH-2018-1]. R.G. was supported by the National Institute for Health Research [NIHR] Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. JX, AW and CLQ were supported by Lundbeck Foundation, Denmark [grant R344-2020-989]. Finally, this publication has also been made possible by the support of the Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia (Spain), which finances the PhD of Alicia Gómez-Pascual [21259/FPI/19].

The results published here are in whole or in part based on data obtained from Agora, a platform initially developed by the NIA-funded AMP-AD consortium that shares evidence in support of AD target discovery. Agora is available at: https://doi.org/10.57718/agora-adknowledgeportal.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108588.

References

[1] S. Gauthier, B. Reisberg, M. Zaudig, R.C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, J.L. Cummings, M. de Leon, H. Feldman, M. Ganguli, H. Hampel, P. Scheltens, M.C. Tierney, P. Whitehouse, B. Winblad, Mild cognitive impairment, Lancet 367 (2006) 1262–1270, https://doi.org/10.1016/S0140-6736(06)68542-5.

[2] W. Bai, P. Chen, H. Cai, Q. Zhang, Z. Su, T. Cheung, T. Jackson, S. Sha, Y.-T. Xiang, Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies, Age Ageing 51 (2022), https://doi.org/10.1093/ageing/afac173 afac173.

[3] R.O. Roberts, D.S. Knopman, M.M. Mielke, R.H. Cha, V.S. Pankratz, T.J. H. Christianson, Y.E. Geda, B.F. Boeve, R.J. Ivnik, E.G. Tangalos, W.A. Rocca, R. C. Petersen, Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal, Neurology 82 (2014) 317–325, https://doi.org/10.1212/WNL.0000000000000055.

[4] Y. Varatharajah, V.K. Ramanan, R. Iyer, P. Vemuri, Alzheimer's disease neuroimaging Initiative, predicting Short-term MCI-to-AD progression using imaging, CSF, Genetic factors, cognitive Resilience, and demographics, Sci. Rep. 9 (2019) 2235, https://doi.org/10.1038/s41598-019-38793-3.

[5] D. Heister, J.B. Brewer, S. Magda, K. Blennow, L.K. McEvoy, F. the, A.D. N. Initiative, Predicting MCI outcome with clinically available MRI and CSF biomarkers, Neurology 77 (2011) 1619–1628, https://doi.org/10.1212/WNL.0b013e3182343314.

[6] C. Hinrichs, V. Singh, G. Xu, S.C. Johnson, Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population, Neuroimage 55 (2011) 574–589, https://doi.org/10.1016/j.neuroimage.2010.10.081.

[7] K. Huang, Y. Lin, L. Yang, Y. Wang, S. Cai, L. Pang, X. Wu, L. Huang, A multipredictor model to predict the conversion of mild cognitive impairment to Alzheimer's disease by using a predictive nomogram, Neuropsychopharmacology 45 (2020) 358–366, https://doi.org/10.1038/s41386-019-0551-0.

[8] G. Lee, K. Nho, B. Kang, K.-A. Sohn, D. Kim, Predicting Alzheimer's disease progression using multi-modal deep learning approach, Sci. Rep. 9 (2019) 1952, https://doi.org/10.1038/s41598-018-37769-z.

[9] E.E. Smith, S. Egorova, D. Blacker, R.J. Killiany, A. Muzikansky, B.C. Dickerson, R. E. Tanzi, M.S. Albert, S.M. Greenberg, C.R.G. Guttmann, Magnetic Resonance imaging white Matter Hyperintensities and brain volume in the prediction of mild cognitive impairment and dementia, Arch. Neurol. 65 (2008) 94–100, https://doi.org/10.1001/archneurol.2007.23.

[10] H. Liu, X. Zhou, H. Jiang, H. He, X. Liu, A semi-mechanism approach based on MRI and proteomics for prediction of conversion from mild cognitive impairment to Alzheimer's disease, Sci. Rep. 6 (2016) 26712, https://doi.org/10.1038/srep26712.

[11] C. Manzoni, D.A. Kia, J. Vandrovcova, J. Hardy, N.W. Wood, P.A. Lewis, R. Ferrari, Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences, Brief. Bioinform. 19 (2018) 286–302, https://doi.org/10.1093/bib/bbw114.

[12] S. Janelidze, N. Mattsson, S. Palmqvist, R. Smith, T.G. Beach, G.E. Serrano, X. Chai, N.K. Proctor, U. Eichenlaub, H. Zetterberg, K. Blennow, E.M. Reiman, E. Stomrud, J.L. Dage, O. Hansson, Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia, Nat. Med. 26 (2020) 379–386, https://doi.org/10.1038/s41591-020-0755-1.

[13] H.-T. Li, S.-X. Yuan, J.-S. Wu, Y. Gu, X. Sun, Predicting conversion from MCI to AD combining multi-modality data and based on molecular subtype, Brain Sci. 11 (2021) 674, https://doi.org/10.3390/brainsci11060674.

[14] I.O. Korolev, L.L. Symonds, A.C. Bozoki, A.D.N. Initiative, Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via Probabilistic Pattern classification, PLoS One 11 (2016) e0138866, https://doi.org/10.1371/journal.pone.0138866.

[15] M. Orešič, T. Hyötyläinen, S.-K. Herukka, M. Sysi-Aho, I. Mattila, T. Seppänan-Laakso, V. Julkunen, P.V. Gopalacharyulu, M. Hallikainen, J. Koikkalainen, M. Kivipelto, S. Helisalmi, J. Lötjönen, H. Soininen, Metabolome in progression to Alzheimer's disease, Transl. Psychiatry 1 (2011) e57, https://doi.org/10.1038/tp.2011.55.

[16] Y.-L. Huang, C.-H. Lin, T.-H. Tsai, C.-H. Huang, J.-L. Li, L.-K. Chen, C.-H. Li, T.-F. Tsai, P.-N. Wang, Discovery of a metabolic Signature Predisposing high risk patients with mild cognitive impairment to converting to Alzheimer's disease, Int. J. Mol. Sci. 22 (2021) 10903, https://doi.org/10.3390/ijms222010903.

[17] L. Shi, S. Westwood, A.L. Baird, L. Winchester, V. Dobricic, F. Kilpert, S. Hong, A. Franke, A. Hye, N.J. Ashton, A.R. Morgan, I. Bos, S.J.B. Vos, N.J. Buckley, M. T. Kate, P. Scheltens, R. Vandenberghe, S. Gabel, K. Meersmans, S. Engelborghs, E. E. De Roeck, K. Sleegers, G.B. Frisoni, O. Blin, J.C. Richardson, R. Bordet, J. L. Molinuevo, L. Rami, A. Wallin, P. Kettunen, M. Tsolaki, F. Verhey, A. Lleó, D. Alcolea, J. Popp, G. Peyratout, P. Martinez-Lage, M. Tainta, P. Johannsen, C. E. Teunissen, Y. Freund-Levi, L. Frölich, C. Legido-Quigley, F. Barkhof, K. Blennow, H. Zetterberg, S. Baker, B.P. Morgan, J. Streffer, P.J. Visser, L. Bertram, S. Lovestone, A.J. Nevado-Holgado, Discovery and validation of plasma proteomic biomarkers relating to brain amyloid burden by SOMAscan assay, Alzheimers Dement. J. Alzheimers Assoc. 15 (2019) 1478–1488, https://doi.org/10.1016/j.jalz.2019.06.4951.

[18] P. Proitsi, M. Kim, L. Whiley, A. Simmons, M. Sattlecker, L. Velayudhan, M. K. Lupton, H. Soininen, I. Kloszewska, P. Mecocci, M. Tsolaki, B. Vellas, S. Lovestone, J.F. Powell, R.J.B. Dobson, C. Legido-Quigley, Association of blood lipids with Alzheimer's disease: a comprehensive lipidomics analysis, Alzheimers Dement, J. Alzheimers Assoc. 13 (2017) 140–151, https://doi.org/10.1016/j.jalz.2016.08.003.

[19] M. Picard, M.-P. Scott-Boyer, A. Bodein, O. Périn, A. Droit, Integration strategies of multi-omics data for machine learning analysis, Comput. Struct. Biotechnol. J. 19 (2021) 3735–3746, https://doi.org/10.1016/j.csbj.2021.06.030.

[20] I. Bos, S. Vos, R. Vandenberghe, P. Scheltens, S. Engelborghs, G. Frisoni, J. L. Molinuevo, A. Wallin, A. Lleó, J. Popp, P. Martinez-Lage, A. Baird, R. Dobson, C. Legido-Quigley, K. Sleegers, C. Van Broeckhoven, L. Bertram, M. Ten Kate, F. Barkhof, H. Zetterberg, S. Lovestone, J. Streffer, P.J. Visser, The EMIF-AD Multimodal Biomarker Discovery study: design, methods and cohort characteristics, Alzheimer's Res. Ther. 10 (2018) 64, https://doi.org/10.1186/s13195-018-0396-5.

[21] M. François, A.V. Karpe, J.-W. Liu, D.J. Beale, M. Hor, J. Hecker, J. Faunt, J. Maddison, S. Johns, J.D. Doecke, S. Rose, W.R. Leifert, Multi-omics, an integrated approach to identify novel blood biomarkers of Alzheimer's disease, Metabolites 12 (2022) 949, https://doi.org/10.3390/metabo12100949.

[22] E. Eljaschewitsch, A. Witting, C. Mawrin, T. Lee, P.M. Schmidt, S. Wolf, H. Hoertnagl, C.S. Raine, R. Schneider-Stock, R. Nitsch, O. Ullrich, The endocannabinoid anandamide protects neurons during CNS inflammation by induction of MKP-1 in microglial cells, Neuron 49 (2006) 67–79, https://doi.org/10.1016/j.neuron.2005.11.027.

[23] R.C. Petersen, Mild cognitive impairment as a diagnostic entity, J. Intern. Med. 256 (2004) 183–194, https://doi.org/10.1111/j.1365-2796.2004.01388.x.

[24] B. Winblad, K. Palmer, M. Kivipelto, V. Jelic, L. Fratiglioni, L.-O. Wahlund, A. Nordberg, L. Bäckman, M. Albert, O. Almkvist, H. Arai, H. Basun, K. Blennow, M. de Leon, C. DeCarli, T. Erkinjuntti, E. Giacobini, C. Graff, J. Hardy, C. Jack, A. Jorm, K. Ritchie, C. van Duijn, P. Visser, R.C. Petersen, Mild cognitive impairment–beyond controversies, towards a consensus: report of the International working group on mild cognitive impairment, J. Intern. Med. 256 (2004) 240–246, https://doi.org/10.1111/j.1365-2796.2004.01380.x.

[25] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, E.M. Stadlan, Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and human Services task Force on Alzheimer's disease, Neurology 34 (1984) 939–944, https://doi.org/10.1212/wnl.34.7.939.

[26] M. Kim, S. Snowden, T. Suvitaival, A. Ali, D.J. Merkler, T. Ahmad, S. Westwood, A. Baird, P. Proitsi, A. Nevado-Holgado, A. Hye, I. Bos, S. Vos, R. Vandenberghe, C. Teunissen, M. ten Kate, P. Scheltens, S. Gabel, K. Meersmans, O. Blin, J. Richardson, E. De Roeck, K. Sleegers, R. Bordet, L. Rami, P. Kettunen, M. Tsolaki, F. Verhey, I. Sala, A. Lléo, G. Peyratout, M. Tainta, P. Johannsen, Y. Freund-Levi, L. Frölich, V. Dobricic, S. Engelborghs, G.B. Frisoni, J.L. Molinuevo, A. Wallin, J. Popp, P. Martinez-Lage, L. Bertram, F. Barkhof, N. Ashton, K. Blennow, H. Zetterberg, J. Streffer, P.J. Visser, S. Lovestone, C. Legido-Quigley, Primary fatty amides in plasma associated with brain amyloid burden, hippocampal volume, and

memory in the European Medical Information Framework for Alzheimer's Disease biomarker discovery cohort, Alzheimers Dement. J. Alzheimers Assoc. 15 (2019) 817–827, https://doi.org/10.1016/j.jalz.2019.03.004.
[27] M. Kuhn, Building predictive models in R using the caret package, J. Stat. Software 28 (2008) 1–26, https://doi.org/10.18637/jss.v028.i05.
[28] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E.N. Brody, J. Carter, A.B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T.R. Keeney, N. Kim, T.H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S.K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R.M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D.I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J.D. Vaught, M. Vrkljan, J.J. Walker, M. Watrobka, S. Waugh, A. Weiss, S.K. Wilcox, A. Wolfson, S.K. Wolk, C. Zhang, D. Zichi, Aptamer-based Multiplexed proteomic Technology for biomarker discovery, PLoS One 5 (2010) e15004, https://doi.org/10.1371/journal.pone.0015004.
[29] J.T. Leek, W.E. Johnson, H.S. Parker, A.E. Jaffe, J.D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments, Bioinformatics 28 (2012) 882–883, https://doi.org/10.1093/bioinformatics/bts034.
[30] L. Shi, J. Xu, R. Green, A. Wretlind, J. Homann, N.J. Buckley, B.M. Tijms, S.J. B. Vos, C.M. Lill, M.T. Kate, S. Engelborghs, K. Sleegers, G.B. Frisoni, A. Wallin, A. Lleó, J. Popp, P. Martinez-Lage, J. Streffer, F. Barkhof, H. Zetterberg, P.J. Visser, S. Lovestone, L. Bertram, A.J. Nevado-Holgado, P. Proitsi, C. Legido-Quigley, Multiomics profiling of human plasma and cerebrospinal fluid reveals ATN-derived networks and highlights causal links in Alzheimer's disease, Alzheimers Dement, J. Alzheimers Assoc. 19 (2023) 3350–3364, https://doi.org/10.1002/alz.12961.
[31] M. Kuhn, D. Vaughan, E. Hvitfeldt, Yardstick: Tidy Characterizations of Model Performance, 2024. https://github.com/tidymodels/yardstick.
[32] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[33] L. Breiman, Manual on Setting up, using, and understanding random forests V3.1. https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf, 2002.
[34] S. Epskamp, A.O.J. Cramer, L.J. Waldorp, V.D. Schmittmann, D. Borsboom, Qgraph: network visualizations of relationships in Psychometric data, J. Stat. Software 48 (2012) 1–18, https://doi.org/10.18637/jss.v048.i04.
[35] S.J. Kiddle, M. Sattlecker, P. Proitsi, A. Simmons, E. Westman, C. Bazenet, S. K. Nelson, S. Williams, A. Hodges, C. Johnston, H. Soininen, I. Kłoszewska, P. Mecocci, M. Tsolaki, B. Vellas, S. Newhouse, S. Lovestone, R.J.B. Dobson, Candidate blood proteome markers of Alzheimer's disease Onset and progression: a systematic review and replication study, J. Alzheimers Dis. 38 (2014) 515–531, https://doi.org/10.3233/JAD-130380.
[36] M. Gabrielli, I. Prada, P. Joshi, C. Falcicchia, G. D'Arrigo, G. Rutigliano, E. Battocchio, R. Zenatelli, F. Tozzi, A. Radeghieri, O. Arancio, N. Origlia, C. Verderio, Microglial large extracellular vesicles propagate early synaptic dysfunction in Alzheimer's disease, Brain J. Neurol. 145 (2022) 2849–2868, https://doi.org/10.1093/brain/awac083.
[37] T. Zal, A. Volkmann, B. Stockinger, Mechanisms of tolerance induction in major histocompatibility complex class II-restricted T cells specific for a blood-borne self-antigen, J. Exp. Med. 180 (1994) 2089–2099, https://doi.org/10.1084/jem.180.6.2089.
[38] H.O. Dikmen, M. Hemmerich, A. Lewen, J.-O. Hollnagel, B. Chausse, O. Kann, GM-CSF induces noninflammatory proliferation of microglia and disturbs electrical neuronal network rhythms in situ, J. Neuroinflammation 17 (2020) 235, https://doi.org/10.1186/s12974-020-01903-4.
[39] I. Prada, M. Gabrielli, E. Turola, A. Iorio, G. D'Arrigo, R. Parolisi, M. De Luca, M. Pacifici, M. Bastoni, M. Lombardi, G. Legname, D. Cojoc, A. Buffo, R. Furlan, F. Peruzzi, C. Verderio, Glia-to-neuron transfer of miRNAs via extracellular vesicles: a new mechanism underlying inflammation-induced synaptic alterations, Acta Neuropathol. 135 (2018) 529–550, https://doi.org/10.1007/s00401-017-1803-x.
[40] S. Yamamoto, M. Takehara, M. Ushimaru, Inhibitory action of linoleamide and oleamide toward sarco/endoplasmic reticulum Ca2+-ATPase, Biochim. Biophys. Acta Gen. Subj. 1861 (2017) 3399–3405, https://doi.org/10.1016/j.bbagen.2016.09.001.
[41] C.J. Fowler, Oleamide: a member of the endocannabinoid family? Br. J. Pharmacol. 141 (2004) 195–196, https://doi.org/10.1038/sj.bjp.0705608.
[42] J.K. Huang, C.R. Jan, Linoleamide, a brain lipid that induces sleep, increases cytosolic Ca2+ levels in MDCK renal tubular cells, Life Sci. 68 (2001) 997–1004, https://doi.org/10.1016/s0024-3205(00)01002-x.
[43] D. Stamate, M. Kim, P. Proitsi, S. Westwood, A. Baird, A. Nevado-Holgado, A. Hye, I. Bos, S.J.B. Vos, R. Vandenberghe, C.E. Teunissen, M.T. Kate, P. Scheltens, S. Gabel, K. Meersmans, O. Blin, J. Richardson, E. De Roeck, S. Engelborghs, K. Sleegers, R. Bordet, L. Ramit, P. Kettunen, M. Tsolaki, F. Verhey, D. Alcolea, A. Lléo, G. Peyratout, M. Tainta, P. Johannsen, Y. Freund-Levi, L. Frölich, V. Dobricic, G.B. Frisoni, J.L. Molinuevo, A. Wallin, J. Popp, P. Martinez-Lage, L. Bertram, K. Blennow, H. Zetterberg, J. Streffer, P.J. Visser, S. Lovestone, C. Legido-Quigley, A metabolite-based machine learning approach to diagnose Alzheimer-type dementia in blood: results from the European Medical Information Framework for Alzheimer disease biomarker discovery cohort, Alzheimers Dement. N. Y. N 5 (2019) 933–938, https://doi.org/10.1016/j.trci.2019.11.001.
[44] B.F. Cravatt, O. Prospero-Garcia, G. Siuzdak, N.B. Gilula, S.J. Henriksen, D. L. Boger, R.A. Lerner, Chemical characterization of a family of brain lipids that induce sleep, Science 268 (1995) 1506–1509, https://doi.org/10.1126/science.7770779.
[45] L. Xie, H. Kang, Q. Xu, M.J. Chen, Y. Liao, M. Thiyagarajan, J. O'Donnell, D. J. Christensen, C. Nicholson, J.J. Iliff, T. Takano, R. Deane, M. Nedergaard, Sleep Drives metabolite clearance from the Adult brain, Science 342 (2013) 373–377, https://doi.org/10.1126/science.1241224.
[46] S.G. Snowden, A.A. Ebshiana, A. Hye, O. Pletnikova, R. O'Brien, A. Yang, J. Troncoso, C. Legido-Quigley, M. Thambisetty, Neurotransmitter imbalance in the brain and Alzheimer's disease Pathology, J. Alzheimers Dis. JAD 72 (2019) 35–43, https://doi.org/10.3233/JAD-190577.
[47] Y. An, V.R. Varma, S. Varma, R. Casanova, E. Dammer, O. Pletnikova, C.W. Chia, J. M. Egan, L. Ferrucci, J. Troncoso, A.I. Levey, J. Lah, N.T. Seyfried, C. Legido-Quigley, R. O'Brien, M. Thambisetty, Evidence for brain glucose dysregulation in Alzheimer's disease, Alzheimers Dement, J. Alzheimers Assoc. 14 (2018) 318–329, https://doi.org/10.1016/j.jalz.2017.09.011.
[48] G. Stevanin, H. Fujigasaki, A.-S. Lebre, A. Camuzat, C. Jeannequin, C. Dode, J. Takahashi, C. San, R. Bellance, A. Brice, A. Durr, Huntington's disease-like phenotype due to trinucleotide repeat expansions in the TBP and JPH3 genes, Brain J. Neurol. 126 (2003) 1599–1603, https://doi.org/10.1093/brain/awg155.
[49] H. Kettenmann, F. Kirchhoff, A. Verkhratsky, Microglia: new roles for the synaptic Stripper, Neuron 77 (2013) 10–18, https://doi.org/10.1016/j.neuron.2012.12.023.
[50] N. Stella, Endocannabinoid signaling in microglial cells, Neuropharmacology 56 (Suppl 1) (2009) 244–253, https://doi.org/10.1016/j.neuropharm.2008.07.037.
[51] M. Gabrielli, N. Battista, L. Riganti, I. Prada, F. Antonucci, L. Cantone, M. Matteoli, M. Maccarrone, C. Verderio, Active endocannabinoids are secreted on extracellular membrane vesicles, EMBO Rep. 16 (2015) 213–220, https://doi.org/10.15252/embr.201439668.
[52] B.D. Heifets, P.E. Castillo, Endocannabinoid signaling and long-term synaptic plasticity, Annu. Rev. Physiol. 71 (2009) 283–306, https://doi.org/10.1146/annurev.physiol.010908.163149.
[53] M. Kano, T. Ohno-Shosaku, Y. Hashimotodani, M. Uchigashima, M. Watanabe, Endocannabinoid-mediated control of synaptic transmission, Physiol. Rev. 89 (2009) 309–380, https://doi.org/10.1152/physrev.00019.2008.
[54] T. Harkany, K. Mackie, P. Doherty, Wiring and firing neuronal networks: endocannabinoids take center stage, Curr. Opin. Neurobiol. 18 (2008) 338–345, https://doi.org/10.1016/j.conb.2008.08.007.
[55] E.K. Farrell, D.J. Merkler, Biosynthesis, degradation, and pharmacological importance of the fatty acid amides, Drug Discov. Today 13 (2008) 558–568, https://doi.org/10.1016/j.drudis.2008.02.006.