



OPEN

## A new strategy to uncover fragile X proteomic biomarkers using the nascent proteome of peripheral blood mononuclear cells (PBMCs)

Olivier Dionne<sup>✉</sup> & François Corbin<sup>✉</sup>

Fragile X syndrome (FXS) is the most prevalent inherited cause of intellectual disabilities and autism spectrum disorders. FXS result from the loss of expression of the FMRP protein, an RNA-binding protein that regulates the expression of key synaptic effectors. FXS is also characterized by a wide array of behavioural, cognitive and metabolic impairments. The severity and penetrance of those comorbidities are extremely variable, meaning that a considerable phenotypic heterogeneity is found among fragile X individuals. Unfortunately, clinicians currently have no tools at their disposal to assay a patient prognosis upon diagnosis. Since the absence of FMRP was repeatedly associated with an aberrant protein synthesis, we decided to study the nascent proteome in order to screen for potential proteomic biomarkers of FXS. We used a BONCAT (Biorthogonal Non-canonical Amino Acids Tagging) method coupled to label-free mass spectrometry to purify and quantify nascent proteins of peripheral blood mononuclear cells (PBMCs) from 7 fragile X male patients and 7 age-matched controls. The proteomic analysis identified several proteins which were either up or downregulated in PBMCs from FXS individuals. Eleven of those proteins were considered as potential biomarkers, of which 5 were further validated by Western blot. The gene ontology enrichment analysis highlighted molecular pathways that may contribute to FXS physiopathology. Our results suggest that the nascent proteome of PBMCs is well suited for the discovery of FXS biomarkers.

Fragile X syndrome (FXS) is an X-linked neurodevelopmental disorder which represents the most prevalent inherited cause of intellectual disabilities (ID) and autism spectrum disorders (ASD). Several other comorbidities are also associated with FXS, including behavioural issues such as anxiety, aggressivity and hyperactivity as well as physical and metabolic anomalies. The penetrance of those problems is highly variable, resulting in a large phenotypical heterogeneity among individuals with Fragile X (FX)<sup>1-3</sup>. FXS typically originate from the expansion of CGG trinucleotide repeats found in the 5' UTR of the *FMR1* gene. The full mutation (n > 200 CGG) is associated with the epigenetic silencing of *FMR1*, and consequently, with the loss of expression of the fragile X mental retardation protein (FMRP)<sup>4,5</sup>.

FMRP remains the most valuable biomarker for FXS, since it can be used to identify fragile X individuals and to predict, to some extent, their cognitive functions<sup>6-8</sup>. Other biochemical biomarkers, including intracellular signalling pathways, polymorphisms, level of protein synthesis and the quantification of specific proteins and metabolites, have also been proposed<sup>9,10</sup>. From those, only a handful have been tested in clinical contexts<sup>11-17</sup>. Interestingly, it has been shown that the circulating level of MMP9 (Matrix Metalloproteinase 9), Ras, Hk1 (Hexokinase 1), APP (amyloid- $\beta$  precursor protein) and ERK (Extracellular Regulated Kinase) phosphorylation in blood platelets can be modulated following specific treatment, thus corroborating their potential use as objective outcome measures for FXS clinical trials<sup>11,12,14,16</sup>. However, clinicians still have no prognostic tools at their disposal, meaning that the announcement of a FX diagnosis is accompanied by a great deal of uncertainty regarding the severity of the comorbidities presented by the affected individual. Those facts clearly illustrate one of the actual shortcomings in the management of individuals with FXS and consequently, the need for the discovery of new biomarkers to assess a patient's prognosis upon diagnosis.

The discovery of biomarkers using shotgun proteomic approaches faces two major limitations. First, the dynamic expression range of proteins within the proteome makes identification and quantification of low

Department of Biochemistry and Functional Genomic, Faculty of Medicine and Health Sciences, Université de Sherbrooke and Centre de Recherche du CHUS, CIUSSS de l'Estrie-CHUS, Sherbrooke, QC, Canada. ✉email: olivier.dionne@usherbrooke.ca; francois.corbin@usherbrooke.ca

abundance proteins challenging, especially in data-dependent acquisition method<sup>18</sup>. This fact becomes even more important when considering that weakly expressed proteins are usually critical for the characterization of physiopathological mechanisms underlying complex pathologies, such as FXS. Secondly, differentially expression analysis of large proteome dataset highlights, most of the time, a plethora of statistically significant dysregulated proteins between samples. However, not all those findings bear significance from a biological standpoint. It is therefore crucial to use an approach that promotes the identification of proteins that are both dysregulated and biologically relevant. Different workflow can be utilized to achieve such outcomes. One of the most straightforward and widely used method consists of the purification of a specific sub-proteome, either by sub-cellular compartments fractionation or by affinity chromatography. Such approaches ultimately produce proteomic samples that are both smaller, which enhance coverage of the proteome, and enriched in biologically relevant proteins<sup>19</sup>. Targeting the sub-proteome that best depicts the diseases-induced defects therefore constitutes the foundation of an efficient screening strategy.

FMRP is a multifunctional RNA binding protein involved in several molecular processes including regulation of histone modifications, micro-RNA interference, alternative splicing and RNA editing<sup>20–22</sup>. However, FMRP is mainly recognized as a translational regulator, either as a repressor or enhancer<sup>22,23,24</sup>. Many mRNAs subjected to FMRP translational regulation play a key role in neurodevelopment and synaptic transmission<sup>25</sup>. In fact, several alterations found in FXS, including the elevated number of immature dendritic spines and improper synaptic plasticity, seem related to the absence of FMRP's translational control<sup>26,27</sup>. Furthermore, we recently show that the rates of protein synthesis are altered in FX peripheral blood mononuclear cells (PBMCs)<sup>28</sup>. These observations prompted us to use the nascent proteome, defined as proteins synthesized in a defined timeframe, for the screening of proteomic biomarkers. We hypothesize that this sub-proteome will help us to overcome problems typically encountered when performing shotgun proteomic experiments for biomarkers discovery. Indeed, we postulate that the nascent proteome, due to the translational defects encountered in FXS, will provide a proteomic signature that is well suited for such a task. Furthermore, PBMCs constitute a non-invasive model which transcriptome is known to moderately correlate with that of neuronal tissues<sup>29</sup> and that can be repeatedly collected, making it ideal for our purpose.

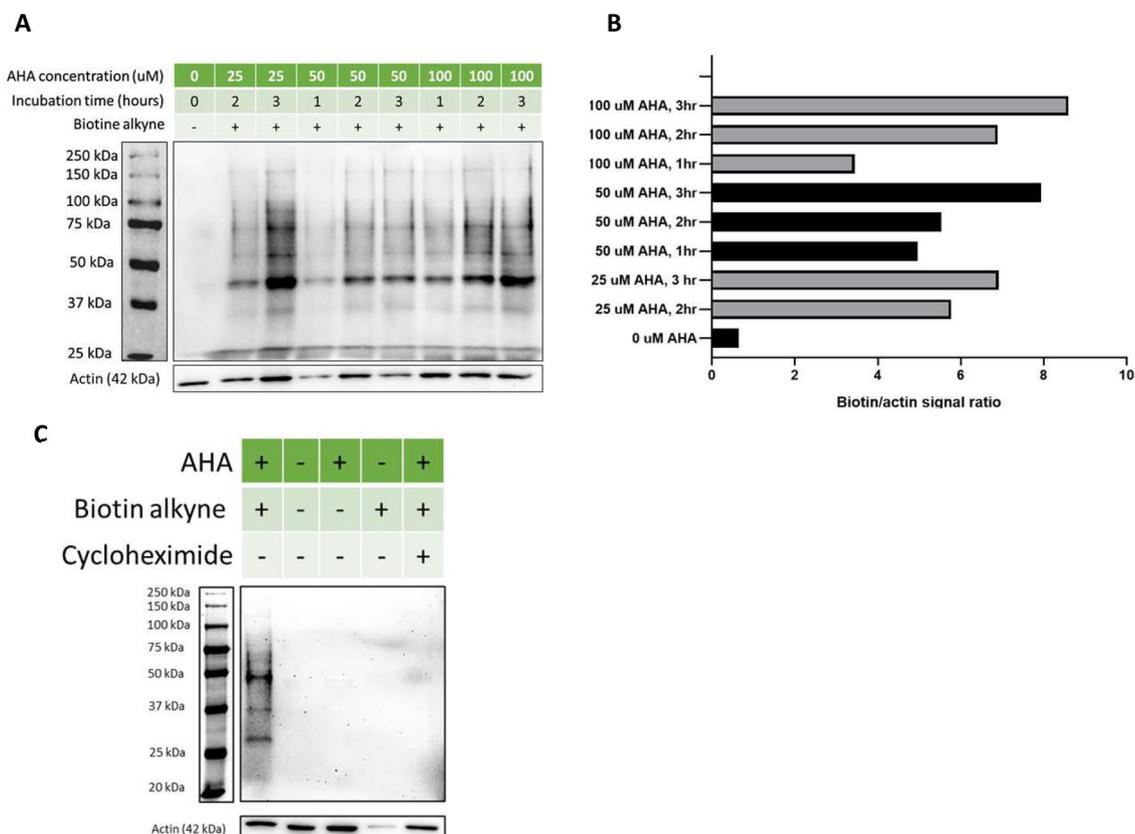
In the present report, we used a BONCAT method coupled to label free mass spectrometry-based proteomic to purify and quantify nascent proteins produced by PBMCs. Indeed, BONCAT (Bioorthogonal Noncanonical Amino Acid Tagging) is a proteomic technique making use of the methionine surrogate L-azidohomoalanine (AHA) to enable the labelling and the purification of newly synthesized proteins<sup>30</sup>. It has been thoroughly used in a variety of blood cells, including PBMCs, to investigate the nascent proteome under different experimental conditions<sup>31–34</sup>. An adaptation of this method has also been used in the hippocampus of Fmr1 KO mice to screen for potential FXS biomarkers. Some of the candidates identified with this approach were also shown to be deregulated in the plasma of FX individuals, thus highlighting the relevance of using the nascent proteome for biomarker discovery in FXS<sup>35</sup>. Consequently, we used BONCAT to compare the nascent proteomes of 7 FXS males and 7 age and sex-matched controls and successfully identified differentially expressed proteins that represents potent biomarker candidates for FXS. To our knowledge, this report also constitutes the first proteomic screening for the discovery of biomarkers in native human samples of FX individuals.

## Methods

**Reagents and antibodies.** Acid citric dextrose (ACD) tubes were from BD Vacutainer. Ficoll-Paque was purchased from GE Healthcare. Azidohomoalanine (AHA), bicinchoninic acid assay kit, Click-it protein reaction kit, biotin alkyne probe, dithiothreitol (DTT), Triton-X100m, C18 tips, goat polyclonal anti-biotin and goat anti-rabbit Alexa FluorVR 680 IgG antibodies were from ThermoFisher. RPMI 1640 Met, P8340 protease inhibitor cocktail, magnetics streptavidin beads, ammonium bicarbonate (ABC), iodoacetamide (IAA), formic acid (FA), mouse monoclonal anti-actin (clone AC-15), Donkey anti-goat IgG, and goat anti-rabbit IgG HRP-conjugated antibodies were bought from MilliporeSigma. Trypsin/lys-C and the enhanced chemiluminescence kit (ECL) were from Promega and PerkinElmer respectively. Rabbit monoclonal anti-ILK (EP1593Y), rabbit monoclonal anti-ANXA2 (ERP13052B) and mouse monoclonal anti-FERMT3 (3D6) were from Abcam. Mouse monoclonal anti-ATP2A3 (PL/IM430) was from Santa Cruz Biotechnology. Recombinant chicken anti-VCL antibody was from Immune Biosolutions. Goat anti-mouse IgG HRP-conjugated was from Jackson IR. Goat anti-mouse IRDyeVR 800CW IgG was from LI-COR Biosciences.

**Study population and ethic declarations.** The study population included 7 FX patients and 7 healthy controls. Participants were all males and were matched for age. The recruitment was performed through the Fragile X Clinic, at the *CIUSSS de l'Estrie-CHUS* (Sherbrooke, Québec, Canada). Informed written consent was obtained from healthy controls and from a caregiver for FXS participants. All participants had blood draw in the morning to decrease potential diurnal variation. All experimental protocols described in this study were approved by the Ethics Review Board of the *CIUSSS de l'Estrie-CHUS* and carried out in line with the Declaration of Helsinki.

**PBMCs isolation.** PBMCs isolation was carried out using Ficoll-Paque following the manufacturer instructions with some minor modifications. Briefly, blood samples were collected by venipuncture into 8 mL ACD tubes and centrifuged at 300g for 10 min to allow plasma collection. A volume of PBS, equal to the volume of plasma collected, was then added to each tube and resulting blood samples were placed onto a layer of Ficoll-Paque (blood/Ficoll-Paque ratio of 4:3). Afterwards, samples were centrifuged at 500g for 30 min. PBMCs were subsequently collected, washed two times with PBS and counted on a flow cytometer (DXH-9000 hematology analyzer, Beckman Coulter<sup>®</sup>).



**Figure 1.** Determination of the optimal AHA labelling condition. **(A)** Anti-biotin Western Blot of protein extracts obtained from PBMCs under different AHA labelling conditions. **(B)** Quantification of the biotin signal obtained by Western blot. Biotin signal intensities were normalized with the corresponding actin signal. **(C)** Specificity of the AHA labelling and conjugation of the biotin probe towards newly synthesized proteins. Full-length blots for biotin and cropped blots for actin are displayed. Full-length actin blots are displayed in Supplementary Figure 3. Multiple exposures of blots displayed in **(A)** and **(C)** can be found in Supplementary Figures 1 and 2 respectively.

**AHA labelling.** Freshly extracted PBMCs were first resuspended into warm (37 °C) RPMI 1640 Met<sup>-</sup> (supplemented with 2 mM L-glutamine) and incubated for 30 min at 37 °C under gentle agitation to deplete the intracellular reserve of methionine. PBMCs were then diluted to a concentration of 3 million cells/mL and labelled with 100 µM AHA for two hours. Cells were pelleted and stored at -80 °C after labelling. This protocol was adapted from the one previously described by another group which also perform BONCAT in PBMCs with further optimization of the AHA concentration and labelling time<sup>32</sup>. The optimized labelling conditions as well as the specificity of the AHA labelling and subsequent Click reaction toward nascent proteins were determined by Western blot using a specific anti-biotin antibody (Fig. 1).

**PBMCs protein extract.** PBMCs were lysed in ice-cold 50 mM Tris pH 8.0, 1% NaDoc, 1% P8340 protease inhibitor cocktail and 250 U/mL Benzonase nuclease. After 15 min on ice, cells were homogenized with a 28G needle and proteins were pelleted for 25 min at 20,000g (4 °C). Protein quantification in the supernatant was determined by bicinchoninic acid assay.

**Isolation of nascent proteins and preparation for MS analysis.** Two experimental groups were formed (control and FX) by combining an equal amount of protein from each participant. The conjugation of the biotin probe was performed on 150 µg of pooled proteins using the Click-it protein reaction kit, following the manufacturer instructions and using a biotin alkyne probe. Nascent biotinylated proteins were purified by an overnight incubation with magnetic streptavidin beads at 4 °C. The beads were washed 4 times with PBS containing 0.1% Triton X-100 and 5 times with a 20 mM pH 8.0 ABC (ammonium bicarbonate) buffer. Cysteine reduction was realized by incubating the beads with 10 mM DTT at 60 °C for 30 min and subsequent alkylation with 15 mM iodoacetamide for 30 min at room temperature in the dark. Proteins digestion into peptide was performed overnight at 37 °C with 1 µg of trypsin/lys-C. The digestion was stopped by adding formic acid to a final concentration of 1%. The supernatant was transferred to a clean tube and the beads were washed twice with 60% acetonitrile in 0.1% formic acid. All supernatants were combined and evaporated. The peptide samples were reconstituted in 0.1% trifluoroacetic acid, desalted with a C18 tip and dried. Nascent proteins were suspended in

1% formic acid and kept at  $-20\text{ }^{\circ}\text{C}$ . Four technical replicates were made for each experimental group, and each replicate was injected twice into the mass spectrometer (intra-assay replicate).

**Preparation of PBMCs total proteome for MS analysis.** The same experimental groups used in the previous section were also used for the preparation of the total proteome. Briefly,  $5\text{ }\mu\text{g}$  of pooled protein extracts were solubilized with  $8\text{ M}$  urea in  $10\text{ mM}$  HEPES pH 8.0 (protein/urea ratio (p/v) of 1:1). Reduction of proteins was performed with  $10\text{ mM}$  DTT at  $60\text{ }^{\circ}\text{C}$  for 30 min and alkylation with  $15\text{ mM}$  IAA at room temperature in the dark for 30 min. Urea concentration was lowered below  $1\text{ M}$  by adding  $50\text{ mM}$  pH 8.0 ABC buffer. Proteins digestion into peptide was performed overnight at  $37\text{ }^{\circ}\text{C}$  with  $0.25\text{ }\mu\text{g}$  (ratio trypsin/protein of 1:20) of trypsin/lys-C and stopped by adding 1% formic acid. Peptide samples were evaporated, desalted with a C18 tip and evaporated. The total proteome was suspended in 1% formic acid and kept at  $-20\text{ }^{\circ}\text{C}$ . Four technical replicates were made for each experimental group, and each replicate was injected twice into the mass spectrometer (intra-assay replicate).

**LC-MS/MS analysis.** Peptides were injected into an HPLC (nanoElute, Bruker Daltonics) and loaded onto a trap column with a constant flow of  $4\text{ }\mu\text{L}/\text{min}$  (Acclaim PepMap100 C18 column,  $0.3\text{ mm id}\times 5\text{ mm}$ , Dionex Corporation) and eluted onto an analytical C18 Column ( $1.9\text{ }\mu\text{m}$  beads size,  $75\text{ }\mu\text{m}\times 25\text{ cm}$ , PepSep). Peptides were eluted over a 2-h gradient of acetonitrile (5–37%) in 0.1% formic acid at  $400\text{ nL}/\text{min}$  while being injected into a TimsTOF Pro Mass Spectrometer equipped with a Captive Spray nano-electrospray source (Bruker Daltonics). Data were acquired using data-dependent auto-MS/MS with a 100–1700  $m/z$  mass range, with PASEF enabled with a number of PASEF scans set at 10 (1.27 s duty cycle) and a dynamic exclusion of 0.4 min,  $m/z$  dependent isolation window and collision energy of 42.0 eV. The target intensity was set to 20,000, with an intensity threshold of 2500.

**Mass spectrometry data analysis.** Proteins identification from the raw data was accomplished with the MaxQuant software<sup>36</sup> (version 1.6.10.0). Peaks list was searched against the Uniprot human database (09/2019)<sup>37</sup>. Trypsin was set as digestion enzyme with specificity for arginine and lysine (but not before proline). A maximum of two miss cleavages was tolerated. Oxidation of methionine and acetylation of proteins N-terminus were set as variable modifications, while carbamidomethylation of cysteine was set as a fixed modification. Carbamylation of lysine was set as variable modification only for the analysis of the total proteome. The false discovery rate of peptides (minimum of 7 amino acids) and proteins were set to 0.05 using a reverse database. The mass tolerance was set to 7 ppm for precursor ions and 20 ppm for fragment ions. The MaxQuant label-free quantification (LFQ) was used, with a minimum ratio count of two, for accurate intensity based quantification of proteins between samples<sup>38</sup>.

**Proteomic data processing.** Processing of proteomics data was carried out with the Perseus software<sup>39</sup> (version 1.6.7.0.). Proteins identified as "potential contaminants", "only identified by site" or "reverse" by MaxQuant were excluded. Data were normalized (average intensities of each sample equal within each group) before further analyzes. Proteins only identified in at least 50% of the samples from each group with a minimum of 1 unique peptide and 2 total peptides were kept. The statistical significance of protein expression between the two groups was evaluated by a two-tailed Student *t* test, carried out in the Perseus software ( $p < 0.05$  for the nascent proteome;  $p < 0.01$  after correction with a permutation-based FDR, for the total proteome analysis). We choose to be more stringent for the differential expression analysis carried out for the total proteome to lower the numbers of statistical hits. The corresponding volcano plots were drawn using the ggplot2 package in R<sup>40,41</sup>.

For the nascent proteome analysis, a "negative control" was produced to eliminate contaminating and endogenous biotinylated proteins. To achieve this, an unlabelled protein extract from a control PBMCs was subjected to all the steps of the workflow describe above (Click reaction, streptavidin pull down, on beads proteins preparation for MS analysis, etc.). Proteins identified with a fold change (experimental sample/negative control) superior to 1.1 were kept for further analysis. Proteins that do not fulfill this criterion where excluded.

**Bioinformatic analysis.** Functional annotation enrichment analyses were carried out with the Panther classification system (<http://www.pantherdb.org>) using the Gene Ontology, Panther protein class and REACTOME pathway annotations sets<sup>42–44</sup>. Protein–protein interaction networks (PPI) were obtained from the web-based LENS tool (Lens for Enrichment and Network Studies of Proteins) at the website: <http://severus.dbmi.pitt.edu/LENS><sup>45</sup>.

**Western blots.** Optimal AHA labelling conditions and specificity towards newly synthesized proteins were confirmed by an anti-biotin Western blot (Fig. 1). Briefly,  $10\text{ }\mu\text{g}$  of protein samples were resolved on a 10% SDS-PAGE, transferred onto a nitrocellulose membrane, block with 5% non-fat dry milk and incubated with the following antibodies: anti-biotin (1:1000) and anti-actin (1:5000). Blots were revealed using an enhanced chemiluminescence (ECL) kit and imaged with ChemiDoc (Bio-Rad).

Western blots were also used to validate 5 proteins found deregulated in fragile X PBMCs by the proteomic screening (ILK, ATP2A3, ANXA2, FERMT3 and VCL). Briefly,  $15\text{ }\mu\text{g}$  of proteins from each participant were resolved on a 9% (ILK, ATP2A3 and VCL) or a 12% (ANXA2 and FERMT3) SDS-PAGE, transferred onto a nitrocellulose membrane, block with 5% non-fat dry milk and incubated with the following antibodies: anti-ILK (1:2000), anti-ANXA2 (1:2000), anti-ATP2A3 (1:250), anti-FERMT3 (1:2000), anti-VCL (1:1000) and anti-actin (1:5000). Anti-goat IgG (1:10,000), anti-mouse IgG (1:10,000) and anti-rabbit IgG (1:10,000) HRP-conjugated

ID	Age	Sex	Medication	Fragile X diagnostic
<b>Fragile X cohort</b>				
X1	25	Male		Full mutation
X2	40	Male	Levothyroxine, antipsychotic (Olanzapine)	Full mutation
X3	17	Male		Full mutation
X4	29	Male		Full mutation
X5	43	Male		Full mutation
X6	30	Male	Anti-diabetic (Metformin)	Full mutation
X7	25	Male		Full mutation
<b>Control cohort</b>				
C1	24	Male		NA
C2	23	Male		NA
C3	41	Male	Antiandrogen (Finasteride)	NA
C4	26	Male		NA
C5	18	Male		NA
C6	34	Male	Levothyroxine	NA
C7	35	Male		NA

**Table 1.** Individual characteristics of each participant.

secondary antibodies were used for ECL revelation, while anti-mouse IRDyeVR 800CW IgG (1:10,000) and anti-rabbit Alexa FluorVR 680 IgG (1:10,000) were used for fluorescence-based immunostaining. All immunoblots were analyzed with the Image-J software (NIH)<sup>46</sup>. Immunoblots were revealed either by ECL (imaged with ChemiDoc, BioRad) or by fluorescence (imaged with the Odyssey Infrared Imaging System LI-COR Biosciences).

**Statistical analyses.** Fisher exact tests were used with R studio to perform Gene Ontology enrichment analysis. A p-value inferior to 0.05 (corrected with a permutation-based FDR) was considered significant. For the Western blot analysis, statistically significant difference was determined by a two-tailed Student *t* test calculated in GraphPad Prism (version 8.3.0 for Windows, GraphPad Software, San Diego, California USA, <http://www.graphpad.com>). A p-value inferior to 0.05 was considered significant.

## Results

**Population characteristics.** The study population included 7 fully mutated males with fragile X (mean age of  $28.7 \pm 9.0$ ) and 7 healthy males (mean age of  $27.7 \pm 8.1$ ). Individual characteristics of each participant, including age, medication and fragile X diagnostic are listed in Table 1.

### The nascent proteins identified are enriched in functional annotations associated with RNA metabolism.

As shown in Fig. 2A, a smaller number of nascent proteins were identified. Moreover, 7 of the 109 newly synthesized proteins identified were not found among the 1770 proteins within the total proteome. Intriguingly, 5 of those 7 proteins are found deregulated in fragile X PBMCs (Table 2). Moreover, these nascent proteins display an enrichment for functional annotations related to nucleic acid, nucleosides, and protein binding as well as in cellular component organization or biogenesis (Fig. 2B–D). Several of these annotations can be associated with RNA metabolism, suggesting that the nascent proteins identified in this study may be involved in the same molecular processes as FMRP. Taken together, these results support our hypothesis that the isolation of nascent proteins will promote the identification of dysregulated proteins between FX patients and control individuals that are also relevant from a biological standpoint.

### Differentially expressed proteins are found in PBMCs nascent and total proteome.

Analysis of the total proteome successfully identified 1770 distinct proteins. More precisely, 1752 and 1768 were identified in fragile X and control individuals respectively (Supplementary Table S1). Of these proteins, 200 were found to be deregulated in FXS samples (Fig. 3A). Indeed, 135 were found to be upregulated while 65 were found downregulated (Supplementary Table S2). On the other hand, the BONCAT approach made it possible to identify a total of 109 nascent proteins, 105 from the FXS patients and 106 from control samples (Supplementary Table S3). The total number of nascent proteins identified was comparable to that of a previous study which also used BONCAT in PBMCs with similar labelling conditions<sup>32</sup>. Thirty nascent proteins were found deregulated in FX PBMCs (Fig. 3B), of which 17 were upregulated and 13 downregulated (Table 2).

Moreover, 11 dysregulated proteins were identified using both approaches. The trend of the expression of 7 of those 11 proteins was shown to be constant between the nascent and total proteomes. More specifically, TLN1, FERMT3, HIST1H4A, ILK, MPO and VCL were constantly upregulated, while AHNK was downregulated. However, ATP2A3, PDIA3, PF4 and ANXA2 presented opposite trend of perturbation between the nascent and the total proteome (Fig. 3C and Table 2).

**Figure 2.** Proteins identified in the nascent and total proteome show distinct features. Far less nascent proteins were identified. (A) Venn diagrams representing the total number of proteins identified (FX and CTL combined) in the proteomic screening. When compared to the proteins identified in the total proteome, nascent proteins identified in this study presents an enrichment for annotations with specific Gene Ontology terms. A fisher exact test was used to determine enriched annotations between the two proteomes. FDR inferior to 0.05 was considered significant. Pie charts representing the distribution of different annotations between the two proteomes: PANTHER protein class (B), Gene Ontology molecular function (C) and biological process (D). The exposed part represents enriched annotations in the corresponding proteome. The Venn diagram was made using the Venn Diagram Plotter software (version 2.0, <https://omics.pnl.gov/software/venn-diagram-plotter>).

**Bioinformatic analysis of the differentially expressed proteins.** Functional annotation enrichment analysis of the differentially expressed proteins found in the total proteome of FX PBMCs was carried out with PANTHER. We used all the 1770 proteins identified in the total proteome as reference. Results showed that dys-regulated proteins are involved in cellular adhesion, platelet aggregation and degranulation, MAPK2 and MAPK activation and in signalling pathways related to the MAPK and integrins (Fig. 4).

A PPI network of the 11 proteins found differentially expressed in FX nascent and total proteome (Table 2) is illustrated in Fig. 5A. This network shows that these proteins have common interaction partners, strongly suggesting a potential interaction between them and their involvement in the same biological pathways. A second PPI network was generated using the same 11 proteins set as candidates and with the FMR1 gene set as target. As illustrated in Fig. 5B, all proteins were found to be enriched in this network, showing that the 11 candidates possess common interaction partners with FMRP. This result suggests the implication of those 11 proteins in molecular mechanism in which FMRP is involved, and consequently, the involvement of their deregulation in the mechanism underlying FXS pathophysiology. It is also noteworthy that two proteins (NCL and PRPF40A) differentially expressed in the nascent proteome are present in the interaction network containing FMRP and the 11 candidate proteins.

All those observations are supported by the statistics associated with the LENS analysis (Supplementary Table S4), in which three values are used to describe the network connectivity. These values include "minimum shortest path length", "average shortest path length" and the number of "disconnected nodes". Those results support the fact that the two networks are well enriched. Indeed, all three values are less important in "candidates" when only the 11 candidates are given as well as in "candidate to target" when compared to the values associated with "candidates to random" and "target to random".

**Validation of candidate proteins by western blot.** Five of the 11 proteins differentially expressed in both nascent and total proteome (ILK, ATP2A3, ANXA2, FERMT3 and VCL) from pooled samples of FX PBMCs were chosen to be validated by Western blot (Fig. 6). The trend of perturbation of ILK, ATP2A3, ANXA2 and VCL measured by Western blot was consistent with that measured by mass spectrometry. Furthermore, the difference measured for ILK ( $p=0.0250$ ) and ANXA2 (0.0394) were found to be statistically significant. Full-length blots are available in Supplementary Figure 4.

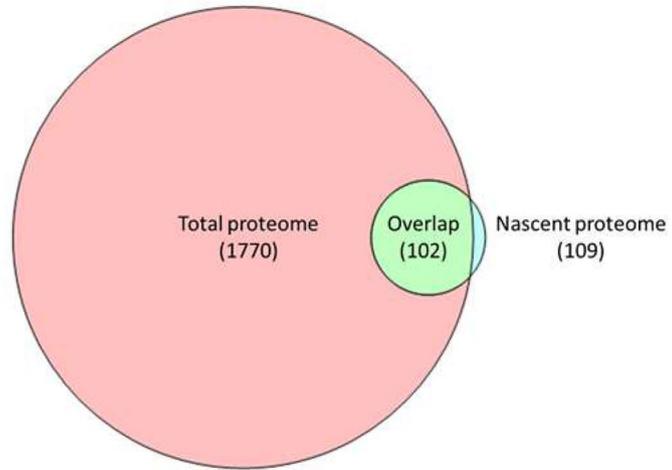
## Discussion

The discovery of biomarkers using shotgun proteomic workflow recurrently led to the identification of a high number of protein candidates. Most of them might ultimately be irrelevant, meaning that a lot of time and resources are spent on selection and validation of newly identified candidates before using them in a clinical setting. This is particularly true when screening experiments are conducted on human biological samples. Indeed, disease-induced heterogeneity and interindividual variability inherent to the use of such samples both add an extra level of intricacy to this already complex equation. One way to limit these drawbacks involves the elaboration of a comprehensive targeted strategy and a rigorous experimental procedure. Indeed, a careful selection of affected individuals and the use of a subproteome that specifically aims at the biological defects induced by the studied pathology can lead to the identification of a higher yield of valuable biomarkers.

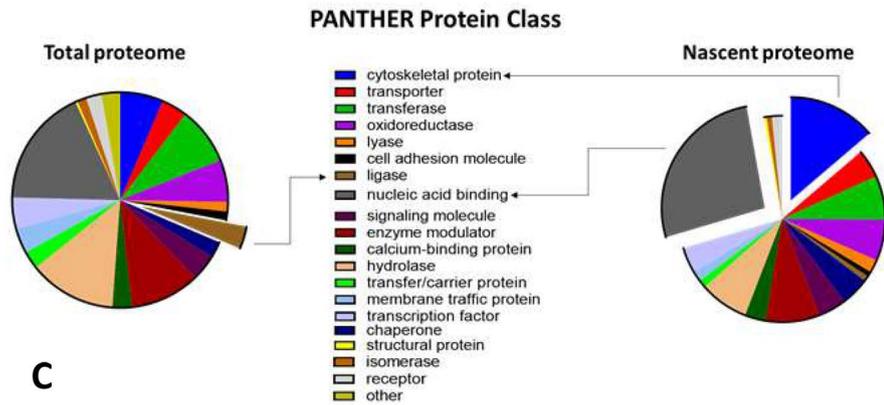
In the current study, we deliberately selected 7 fully mutated FXS males in order to restrain the etiologic heterogeneity within our FX cohort. The inclusion of mosaic or female patients (which both expressed variable level of FMRP) would have greatly increased the variability of our proteomic analysis<sup>6,7</sup>. Furthermore, we chose to study the nascent proteome to screen for proteomic biomarkers, since it is well established that the absence of FMRP leads an aberrant protein synthesis. Moreover, as shown by the functional annotation enrichment analysis (Fig. 2C,D), the proteins identified in the nascent proteome might be well adapted to portray the consequences of FMRP's absence on the various molecular processes in which it is involved, especially those associated with RNA (regulation of translation, alternative splicing, micro-RNA interference and RNA editing)<sup>20,21,22</sup>. Indeed, newly synthesis proteins identified in this study present enrichment for annotations that can be related to RNA metabolism, such as: nucleic acid binding, nucleoside binding and organic cyclic compounds binding<sup>39,40</sup>. Our strategy successfully highlighted 30 dysregulated proteins among the 109 nascent proteins identified from FX individuals PBMCs. This subset approximately represents 28% of all nascent proteins identified in this study (Table 2). Moreover, the mRNA of 20 of those 30 proteins were shown to be bound by FMRP<sup>47,48</sup>. Considered as a whole, these results withstand PBMC's nascent proteome as a promising way to identify potential biomarkers for FXS.

It is well known in the proteomic field that different strategy leads inevitably to the identification of different sets of candidates. Remarkably, more than 35% of proteins found deregulated in the nascent proteome of FX were also shown to be deregulated in our total proteome analysis elevating these 11 proteins as biomarker

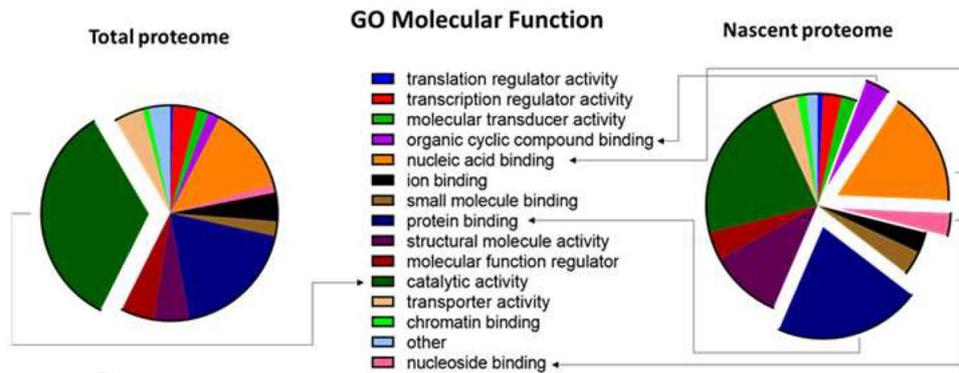
**A**



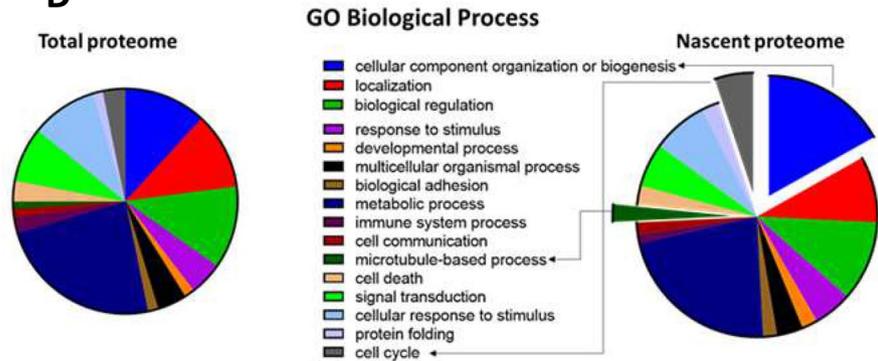
**B**



**C**



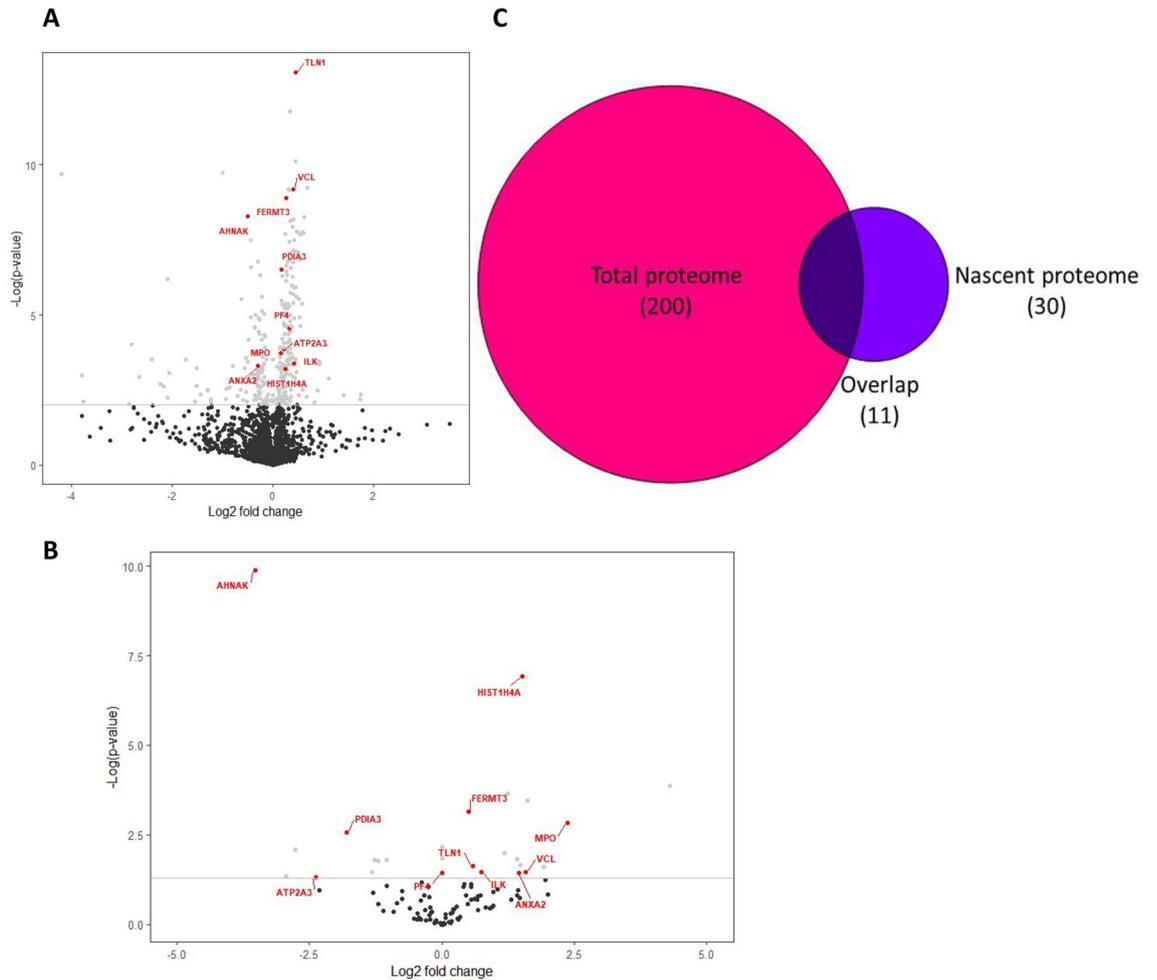
**D**



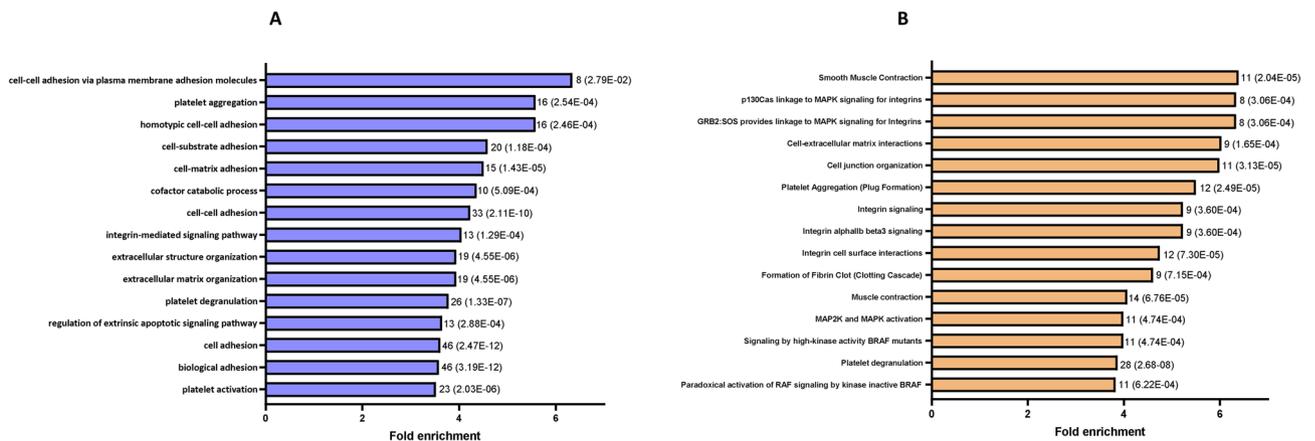
Gene names	Protein names	Fold change nascent proteome	p-value nascent proteome	Fold change total proteome	q-value total proteome
AHNAK	Neuroblast differentiation-associated protein AHNAK	0.087	<0.0001	0.704	<0.0001
ATP2A3	Sarcoplasmic/endoplasmic reticulum calcium ATPase 3	0.193	0.048	1.168	<0.0001
PDIA3	Protein disulfide-isomerase A3	0.288	0.003	1.121	<0.0001
PF4	Platelet factor 4	Detected only in the control group	0.037	1.252	<0.0001
TLN1	Talin-1	1.497	0.024	1.369	<0.0001
FERMT3	Fermitin family homolog 3	1.412	0.001	1.202	<0.0001
HIST1H4A	Histone H4	2.861	<0.0001	1.188	<0.0001
ILK	Integrin-linked protein kinase	1.684	0.034	1.342	<0.0001
MPO	Myeloperoxidase	5.157	0.001	1.115	<0.0001
ANXA2	Annexin A2	2.735	0.037	0.818	<0.0001
VCL	Vinculin	3.013	0.034	1.330	<0.0001
SPP2	Secreted phosphoprotein 24	0.131	0.045	Not detected	NA
HIST1H2AC; HIST3H2A	Histone H2A type 1-C; Histone H2A type 3; Histone H2A type 1-B/E	2.361	<0.0001	Not detected	NA
PRPF40A	Pre-mRNA-processing factor 40 homolog A	3.052	<0.0001	Not detected	NA
PRPF38B	Pre-mRNA-splicing factor 38B	19.741	<0.0001	Not detected	NA
AP3D1	AP-3 complex subunit delta-1	Detected only in the fragile X group	0.028	Not detected	NA
GNAI2	Guanine nucleotide-binding protein G(i) subunit alpha-2	0.487	0.016	0.993	ns
NCL	Nucleolin	0.402	0.035	1.045	ns
DEK	Protein DEK	0.414	0.015	0.853	ns
H1FX	Histone H1x	0.436	0.017	1.186	ns
RAB1B; RAB1C	Ras-related protein Rab-1B; Putative Ras-related protein Rab-1C	Detected only in the control group	0.007	1.161	ns
SNRPD3	Small nuclear ribonucleoprotein Sm D3	Detected only in the control group	0.014	1.009	ns
SNRPE	Small nuclear ribonucleoprotein E	0.147	0.008	0.833	ns
RPL30	60S ribosomal protein L30	Detected only in the control group	0.035	0.828	ns
RPL7	60S ribosomal protein L7	2.799	0.022	0.972	ns
H2AFY	Core histone macro-H2A.1	2.257	0.010	1.007	ns
NOP58	Nucleolar protein 58	2.680	0.015	1.049	ns
RPS18	40S ribosomal protein S18	3.767	0.024	0.955	ns
PCBP1	Poly(rC)-binding protein 1	Detected only in the fragile X group	0.046	1.330	ns
TOP1	DNA topoisomerase 1	Detected only in the fragile X group	0.019	0.547	ns

**Table 2.** Differentially expressed proteins in the fragile X nascent proteome. Thirty proteins are found differentially expressed ( $p < 0.05$ ) in the nascent proteome of the fragile X group. From those proteins, 11 are also found deregulated in the total proteome ( $q < 0.01$ ) of the fragile X group and 5 were only detected in the nascent proteome.

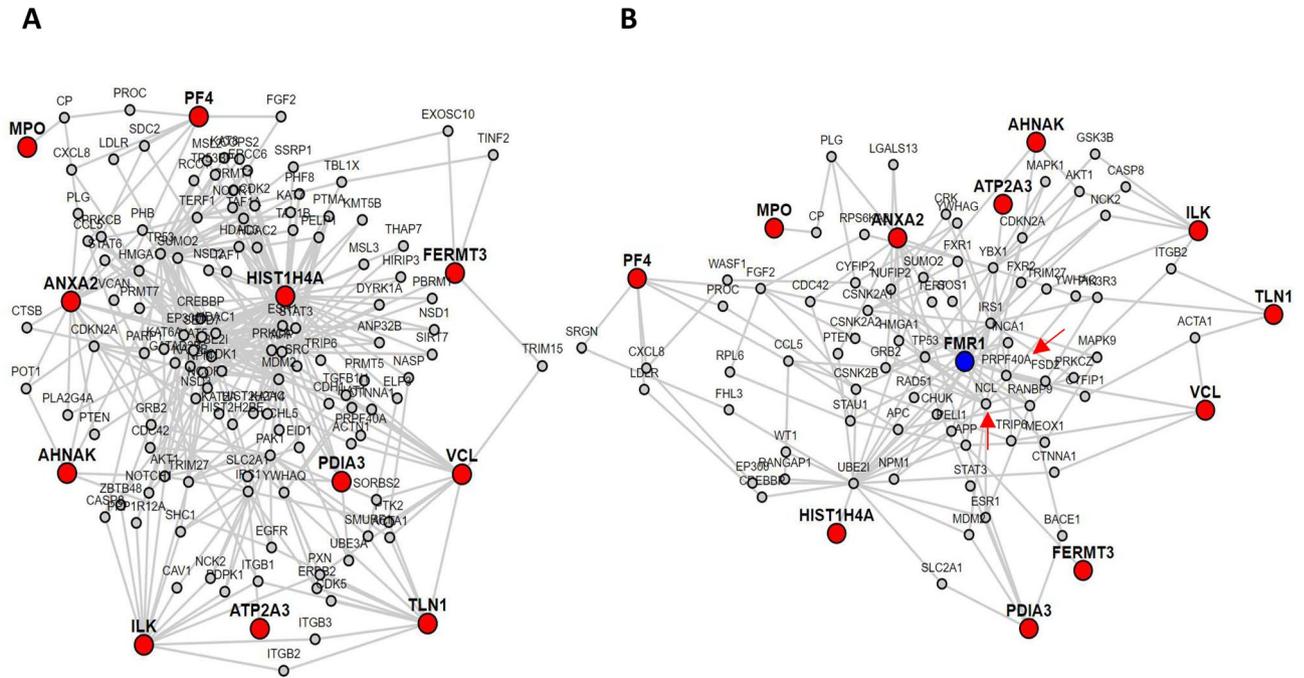
candidates (Fig. 2A and Table 2). Interestingly, 4 of these 11 proteins showed the opposite trend of perturbation between the nascent and total proteomes. These differences could arise from an alteration of their turnover rates, which in turn may result from FMRP's absence. As such, the perturbation observed in the nascent proteome can be indicative of the level of synthesis of a specific protein, while the trend observed in the total proteome may reflect the level of catabolism. In this sense, ANXA2 (which was upregulated in the nascent proteome but downregulated in the total proteome), may present both higher levels of synthesis and degradation in PBMCS of FX individuals. An altered protein turnover rate may also explain the substantial difference regarding the fold changes of the 30 dysregulated proteins in the nascent proteome of FX PBMCS. Moreover, these opposite trends of perturbation between the two proteomes could also be the result of a compensatory mechanism, in which the level of synthesis of a protein is adjusted to compensate for its abnormal expression in the total proteome.



**Figure 3.** Differentially expressed proteins are found in fragile X PBMCs. **(A)** Volcano plot of the proteins identified in the total proteome of PBMCs. The horizontal bar represents a p-value of 0.01, which was considered significant. **(B)** Volcano plot of the proteins identified in the nascent proteome of PBMCs. The horizontal bar represents a p-value of 0.05, which was considered significant. The proteins marked in red are found to be differentially expressed in both proteomes. **(C)** Venn diagram of the number of differentially expressed proteins found in both fragile X nascent and total proteome. This graph was made using the Venn Diagram Plotter software (version 2.0, <https://omics.pnl.gov/software/venn-diagram-plotter>).



**Figure 4.** Functional annotation enrichment analysis of the differentially expressed proteins found in fragile X PBMCs total proteome. The number of proteins associated and corresponding FDR for each annotation are represented beside the bar. **(A)** The top 15 gene ontology biological process annotations ranked by the fold enrichment. **(B)** The top 15 REACTOME annotations ranked by the fold enrichment.

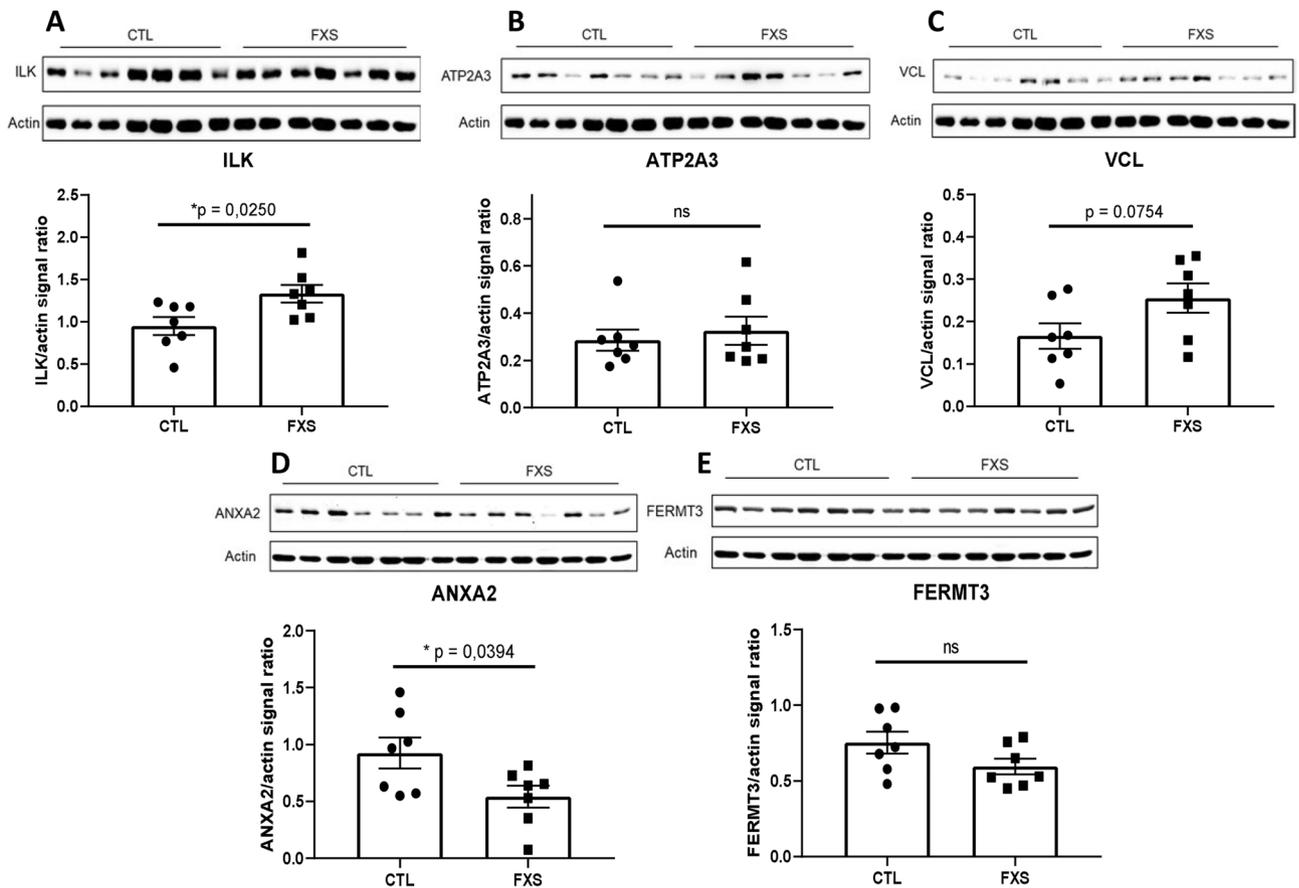


**Figure 5.** Protein–protein interaction networks of 11 proteins found differentially expressed in both nascent and total proteome of fragile X PBMCs. **(A)** Interaction network corresponding only of the 11 candidates (red). **(B)** Interaction network generated with the 11 candidate proteins (red) and FMRP (blue) set as target. The NCL and PRPF40A proteins, which are found deregulated in FXS nascent proteome, are also present in this network (red arrows).

We further validated the dysregulated expression of 5 of those 11 candidates (ILK, ATP2A3, ANXA2, FERMT3 and VCL) by Western blot using the total proteome of PBMCs. Furthermore, as supported by the PPI network (Fig. 5B), many of those 11 proteins can be associated, because of their function and interaction, to the biological process involved in FXS physiopathology.

One of those validated candidates is the ANXA2 protein. The latter inhibits the degradation of LDL receptors mediated by the Proprotein convertase subtilisin/kexin type 9 (PCSK9). Such function is achievable through the binding of ANXA2 to PCSK9, which can occur in both an intra and extracellular manners<sup>49–51</sup>. We have previously found a high rate of hypocholesterolaemia in the FX population<sup>3</sup>. Furthermore, there was no correlation between the plasma level of PCSK9 and LDL cholesterol in FX individuals, a phenomenon which is, however, observed in healthy controls<sup>3,52</sup>. Here we report a dysregulation of ANXA2 expression in FXS which provides new insights of the underlying mechanism of hypocholesterolaemia reported in this population. Further studies are warranted to validate the interaction between FMRP, ANXA2, PCSK9 and cholesterol levels. ANXA2 is also known to interact with S100A10 and AHNAK (another candidate of this study) to form a complex that increase cell surface expression of L-type Voltage-gated calcium channels in mouse brain<sup>53</sup>. AHNAK also promotes the activity of the Raf/MEK/ERK signalling cascade, a pathway hyperactivated in the brain of KO mice and in FX patients blood cells<sup>11,54–57</sup>. Interestingly, a recent study has found deregulation in the expression of AHNAK in PBMCs of children with idiopathic ASD, suggesting the deregulation of AHNAK as a shared pathophysiological mechanism of both conditions<sup>58</sup>.

Three others validated candidates (FERMT3, ILK and VCL), along with TLN1 (another candidate of this study), are involved in integrin activation and/or subsequent signal transduction<sup>59–61</sup>. Furthermore, the functional annotation enrichment analysis (Fig. 4) shown that the proteins deregulated in FX PBMCs are associated with biological processes related to cell adhesion and, consequently, in integrins signaling and underlying signalling cascades. Taken together, those observations suggest an alteration of integrin mediated signalling in FXS. Integrins activation leads to a multitude of cellular processes which are transduced by many signalling pathways, such as the Raf/MEK/ERK and PI3K/AKT/mTOR cascades<sup>61–66</sup>, both of which are known to be upregulated in KO mice and in cells derived from FX individuals<sup>11,49–52,62–64,67</sup>. Integrins are also involved in many neuronal processes, including neurite outgrowth, synapse formation and synaptic transmission<sup>70–72</sup>. Impairing integrin binding to the extracellular matrix (ECM) with RGD peptides (a sequence recognize by integrin extracellular domains) reduced synaptic strength by promoting a decrease in post-synaptic AMPA receptor expression<sup>65,66,68–70</sup>. In neuronal tissue, RGD peptides can be released by the enzymatic activity of the metalloproteinase 9 (MMP9) towards the components of the ECM<sup>71–75</sup>. The overexpression of MMP9 found in the brain of Fmr1 KO mice and in FXS patients<sup>12,71,76–78</sup> can therefore be associated with the impaired synaptic plasticity observed in animal models of the disorder<sup>79–82</sup>. These observations, along with dysregulation of FERMT3, ILK, TLN1 and VCL in fragile X PBMCs suggest that deregulation of integrin activity plays a role in FXS molecular physiopathology and that peripheral blood cells, such as PBMCs or platelets, are relevant models to address this hypothesis<sup>54,79,83–85</sup>.



**Figure 6.** Validation of 5 candidate proteins by western blot. (A) Integrin-linked protein kinase (ILK). (B) Sarcoplasmic/endoplasmic reticulum calcium ATPase 3 (ATP2A3). (C) Vinculin (VCL). (D) Annexin A2 (ANXA2). (E) Fermitin family homolog 3 (FERMT3). (F) Signals of all proteins were normalized to the actin signal to account for loading disparity. A two-tailed Student *t* test was used to calculate the significance of the difference measured between the two groups. Cropped blots are displayed. For full-length blots, refer to Supplementary Figure 4.

Characterization of the nascent proteome has been previously carried out in the mouse model of FXS. Indeed, Bowling and al utilized an adaptation of the BONCAT technique to identify newly synthesis proteins in *Fmr1* KO and WT mice hippocampus under basal and stimulated conditions. They identified several candidates, of which 3 were also shown to be deregulated in the plasma of FXS individuals<sup>35</sup>. A subsequent study validated 2 of those previously identified candidates (along with MMP9) during a clinical trial, showing that their modulation is correlated with treatment efficacy, thus establishing the relevance of the nascent proteome of mice brain to screen for FXS biomarkers<sup>16</sup>. Our present report supports the use of the nascent proteome for biomarker discovery using *ex vivo* native cells from FX individuals. The use of a human model makes it possible to take into consideration the polymorphisms specific to each patient and consequently reflect the phenotypic heterogeneity found within the FXS population. Such benefit cannot be achieved with animal models of FXS, since KO animals share a highly similar genetic background between each other<sup>86</sup>. Furthermore, PBMCs can be repeatedly collected and be used to monitor the effect of disease-modifying drugs on the nascent proteome during clinical trials. However, it should be noted that despite the employment of a similar experimental strategy, only one differentially expressed nascent protein (ANXA2) identified in this study was also dysregulated in the nascent proteome of KO mice hippocampus<sup>35</sup>.

The principal limitation of our study consists of the somewhat low number of proteins identified, especially in the nascent proteome. One possible explanation for this drawback resides in the variations regarding turnover rates between different proteins. Some of the most abundant proteins (ex: histones, cytoskeletal, heat shock proteins, etc.) presents very high turnover rate<sup>87</sup>, which means that they can hinder the identification of low abundant proteins in the nascent proteome. The two hours timeframe used to label nascent proteins may also have contributed to reducing the number of proteins identified since it favours only proteins with high turnover rate. The concentration of AHA used in this study may also hamper our ability to identify nascent proteins. Indeed, previous studies reported deeper coverage of the nascent proteome by using higher concentration of AHA<sup>33,88</sup>. Despite these potential limitations, the number of nascent proteins identified in PBMCs is similar to the one reported by another study using similar experimental conditions<sup>32</sup>. Our experimental workflow, which consisted of the processing of pooled samples, can also have limited the scope of our analysis by limiting our capacity to individually assess each participant's unique proteome. The quantitative aspect of this study is also limited, as

it is the case for all label-free based shotgun proteomic experiments. As such, all the candidates identified in this study will have to be extensively validated in a larger population before being confirmed as FX biomarkers.

## Conclusion

We took advantage of the known translational defects caused by the absence of FMRP to identify several potential biomarkers in FXS. Furthermore, our strategy allows for a minimal recruitment of patients, which limited the inter-individual variation within our FX cohorts, a known caveat of human samples. Obviously, further validation of those candidates in a larger FX cohort in relation to clinical profile is warranted. Nevertheless, the strategy put forth in the present study clearly indicates the feasibility, even for rarer disease, to uncover biomarkers using pathology-driven sub-proteomic strategies with limited human samples.

## Data availability

All data generated or analyzed during this study are included in this article. If any additional information is required, it may be obtained by request from the corresponding author.

Received: 29 January 2021; Accepted: 5 July 2021

Published online: 26 July 2021

## References

- Ciaccio, C. *et al.* Fragile X syndrome: A review of clinical and molecular diagnoses. *Ital. J. Pediatr.* **43**, 39 (2017).
- Kidd, S. A. *et al.* Fragile X syndrome: A review of associated medical problems. *Pediatrics* **134**, 995–1005 (2014).
- Çaku, A. *et al.* New insights of altered lipid profile in fragile X syndrome. *PLoS One* **12**, e0174301 (2017).
- Pieretti, M. *et al.* Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66**, 817–822 (1991).
- Verkerk, A. J. M. H. *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 (1991).
- Lessard, M., Chouiali, A., Drouin, R., Sébire, G. & Corbin, F. Quantitative measurement of FMRP in blood platelets as a new screening test for fragile X syndrome. *Clin. Genet.* **82**, 472–477 (2012).
- Kim, K. *et al.* Association between IQ and FMR1 protein (FMRP) across the spectrum of CGG repeat expansions. *PLoS One* **14**, e0226811 (2019).
- Budimirovic, D. B. *et al.* A genotype-phenotype study of high-resolution FMR1 nucleic acid and protein analyses in fragile X patients with neurobehavioral assessments. *Brain Sci.* **10**, 694 (2020).
- Zafarullah, M. & Tassone, F. Molecular biomarkers in fragile X syndrome. *Brain Sci.* **9**, 96 (2019).
- Pal, R. & Bhattacharya, A. Modelling protein synthesis as a biomarker in fragile X syndrome patient-derived cells. *Brain Sci.* **9**, 59 (2019).
- Pellerin, D. *et al.* Lovastatin corrects ERK pathway hyperactivation in fragile X syndrome: Potential of platelet's signaling cascades as new outcome measures in clinical trials. *Biomarkers* **21**, 497–508 (2016).
- Dziembowska, M. *et al.* High MMP-9 activity levels in fragile X syndrome are lowered by minocycline. *Am. J. Med. Genet. A* **161**, 1897–1903 (2013).
- Erickson, C. A. *et al.* Impact of acamprosate on behavior and brain-derived neurotrophic factor: An open-label study in youth with fragile X syndrome. *Psychopharmacology* **228**, 75–84 (2013).
- Erickson, C. A. *et al.* Impact of acamprosate on plasma amyloid- $\beta$  precursor protein in youth: A pilot analysis in fragile X syndrome-associated and idiopathic autism spectrum disorder suggests a pharmacodynamic protein marker. *J. Psychiatr. Res.* **59**, 220–228 (2014).
- AlOlaby, R. R. *et al.* Molecular biomarkers predictive of sertraline treatment response in young children with fragile X syndrome. *Brain Dev.* **39**, 483–492 (2017).
- Jasoliya, M. *et al.* Blood-based biomarkers predictive of metformin target engagement in fragile X syndrome. *Brain Sci.* **10**, 361 (2020).
- Thurman, A. J. *et al.* Controlled trial of lovastatin combined with an open-label treatment of a parent-implemented language intervention in youth with fragile X syndrome. *J. Neurodev. Disord.* **12**, 12 (2020).
- Ludwig, C. *et al.* Data-independent acquisition-based SWATH-MS for quantitative proteomics: A tutorial. *Mol. Syst. Biol.* **14**, e8126 (2018).
- Cox, B. & Emili, A. Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nat. Protoc.* **1**, 1872–1878 (2006).
- Richter, J. D. & Zhao, X. The molecular biology of FMRP: New insights into fragile X syndrome. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/s41583-021-00432-0> (2021).
- Dockendorff, T. C. & Labrador, M. The fragile X protein and genome function. *Mol. Neurobiol.* **56**, 711–721 (2019).
- Dionne, O. & Corbin, F. An “Omic” overview of fragile X syndrome. *Biology* **10**, 433 (2021).
- Bagni, C. & Klann, E. Molecular functions of the mammalian fragile X mental retardation protein: Insights into mental retardation and synaptic plasticity. *The Autisms: Molecules to Model Systems* 126–146 (2012).
- Maurin, T. & Bardoni, B. Fragile X mental retardation protein: To be or not to be a translational enhancer. *Front. Mol. Biosci.* **5**, 113 (2018).
- Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
- Lombroso, P. J. Genetics of childhood disorders: XLVIII. Learning and memory, Part 1: Fragile X syndrome update. *J. Am. Acad. Child Adolesc. Psychiatry* **42**, 372–375 (2003).
- Bear, M. F., Huber, K. M. & Warren, S. T. The mGluR theory of fragile X mental retardation. *Trends Neurosci.* **27**, 370–377 (2004).
- Dionne, O., Lortie, A., Gagnon, F. & Corbin, F. Rates of protein synthesis are reduced in peripheral blood mononuclear cells (PBMCs) from fragile X individuals. *PLoS One* **16**, e0251367 (2021).
- Sullivan, P. F., Fan, C. & Perou, C. M. Evaluating the comparability of gene expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **141B**, 261–268 (2006).
- Dieterich, D. C., Link, A. J., Graumann, J., Tirrell, D. A. & Schuman, E. M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci.* **103**, 9482–9487 (2006).
- Howden, A. J. M. *et al.* QuaNAT: Quantitating proteome dynamics in primary cells. *Nat. Methods* **10**, 343–346 (2013).
- Bian, F. *et al.* Nascent proteomes in peripheral blood mononuclear cells as a novel source for biomarker discovery in human stroke. *Stroke* **45**, 1177–1179 (2014).
- Bowling, H. L. *et al.* Optimization of protocols for detection of de novo protein synthesis in whole blood samples via azide-alkyne cycloaddition. *J. Proteome Res.* **19**, 3856–3866 (2020).

34. Choi, K.-Y., Lippert, D. N. D., Ezzatti, P. & Mookherjee, N. Defining TNF- $\alpha$  and IL-1 $\beta$  induced nascent proteins: Combining bio-orthogonal non-canonical amino acid tagging and proteomics. *J. Immunol. Methods* **382**, 189–195 (2012).
35. Bowling, H. *et al.* Altered steady state and activity-dependent de novo protein expression in fragile X syndrome. *Nat. Commun.* **10**, 1710 (2019).
36. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
37. The UniProt Consortium. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
38. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13**, 2513–2526 (2014).
39. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
40. R Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing) (2020).
41. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009). <https://doi.org/10.1007/978-0-387-98141-3>.
42. Fabregat, A. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
43. Resource, T. G. O. 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
44. Mi, H. *et al.* Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **14**, 703–721 (2019).
45. Handen, A. & Ganapathiraju, M. K. LENS: Web-based lens for enrichment and network studies of human proteins. *BMC Med. Genomics* **8**, S2 (2015).
46. Rueden, C. T. *et al.* Image J2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* **18**, 529 (2017).
47. Ascano, M. *et al.* FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**, 382–386 (2012).
48. Maurin, T. *et al.* HITS-CLIP in various brain areas reveals new targets and new modalities of RNA binding by fragile X mental retardation protein. *Nucleic Acids Res.* **46**, 6344–6355 (2018).
49. Mayer, G., Poirier, S. & Seidah, N. G. Annexin A2 is a C-terminal PCSK9-binding protein that regulates endogenous low density lipoprotein receptor levels. *J. Biol. Chem.* **283**, 31791–31801 (2008).
50. Ly, K. *et al.* Annexin A2 reduces PCSK9 protein levels via a translational mechanism and interacts with the M1 and M2 domains of PCSK9. *J. Biol. Chem.* **289**, 17732–17746 (2014).
51. Seidah, N. G. *et al.* Annexin A2 is a natural extrahepatic inhibitor of the PCSK9-induced LDL receptor degradation. *PLoS One* **7**, e41865 (2012).
52. Ben Djoudi Ouadda, A. *et al.* Ser-Phosphorylation of PCSK9 (Proprotein Convertase Subtilisin-Kexin 9) by Fam20C (Family With Sequence Similarity 20, Member C) Kinase Enhances Its Ability to Degrade the LDLR (Low-Density Lipoprotein Receptor). *ATVB* **39**, 1996–2013 (2019).
53. Jin, J. *et al.* Ahnak scaffolds p11/Anxa2 complex and L-type voltage-gated calcium channel and modulates depressive behavior. *Mol. Psychiatry* <https://doi.org/10.1038/s41380-019-0371-y> (2019).
54. Michalon, A. *et al.* Chronic pharmacological mGlu5 inhibition corrects fragile X in adult mice. *Neuron* **74**, 49–56 (2012).
55. Weng, N., Weiler, I. J., Sumis, A., Berry-Kravis, E. & Greenough, W. T. Early-phase ERK activation as a biomarker for metabolic status in fragile X syndrome. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **147B**, 1253–1257 (2008).
56. Bhattacharya, A. *et al.* Genetic removal of p70 S6 kinase 1 corrects molecular, synaptic, and behavioral phenotypes in fragile X syndrome mice. *Neuron* **76**, 325–337 (2012).
57. Wang, X. *et al.* Activation of the extracellular signal-regulated kinase pathway contributes to the behavioral deficit of fragile X-syndrome. *J. Neurochem.* **121**, 672–679 (2012).
58. Shen, L. *et al.* Proteomics study of peripheral blood mononuclear cells (PBMCs) in autistic children. *Front. Cell. Neurosci.* **13**, 105 (2019).
59. Legate, K. R., Montañez, E., Kudlacek, O. & Füssler, R. ILK, PINCH and parvin: The tIPP of integrin signalling. *Nat. Rev. Mol. Cell Biol.* **7**, 20–31 (2006).
60. Lai-Cheong, J. E., Parsons, M. & McGrath, J. A. The role of kindlins in cell biology and relevance to human disease. *Int. J. Biochem. Cell Biol.* **42**, 595–603 (2010).
61. Legate, K. R., Wickström, S. A. & Fässler, R. Genetic and cell biological analysis of integrin outside-in signaling. *Genes Dev.* **23**, 397–418 (2009).
62. Walker, J. L., Fournier, A. K. & Assoian, R. K. Regulation of growth factor signaling and cell cycle progression by cell adhesion and adhesion-dependent changes in cellular tension. *Cytokine Growth Factor Rev.* **16**, 395–405 (2005).
63. Chen, H.-C., Appeddu, P. A., Isoda, H. & Guan, J.-L. Phosphorylation of tyrosine 397 in focal adhesion kinase is required for binding phosphatidylinositol 3-kinase. *J. Biol. Chem.* **271**, 26329–26334 (1996).
64. Webb, D. J. *et al.* FAK–Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. *Nat. Cell Biol.* **6**, 154–161 (2004).
65. Pankov, R. *et al.* Specific beta1 integrin site selectively regulates Akt/protein kinase B signaling via local activation of protein phosphatase 2A. *J. Biol. Chem.* **278**, 18671–18681 (2003).
66. Ivaska, J. *et al.* Integrin alpha 2 beta 1 promotes activation of protein phosphatase 2A and dephosphorylation of Akt and glycogen synthase kinase 3 beta. *Mol. Cell. Biol.* **22**, 1352–1359 (2002).
67. Kumari, D. *et al.* Identification of fragile X syndrome specific molecular markers in human fibroblasts: A useful model to test the efficacy of therapeutic drugs. *Hum. Mutat.* **35**, 1485–1494 (2014).
68. Gross, C. & Bassell, G. J. Excess protein synthesis in FXS patient lymphoblastoid cells can be rescued with a p110 $\beta$ -selective inhibitor. *Mol. Med.* **18**, 336–345 (2012).
69. Gross, C. *et al.* Excess phosphoinositide 3-kinase subunit synthesis and activity as a novel therapeutic target in fragile X syndrome. *J. Neurosci.* **30**, 10624–10638 (2010).
70. McGeachie, A. B., Cingolani, L. A. & Goda, Y. A stabilising influence: Integrins in regulation of synaptic plasticity. *Neurosci. Res.* **70**, 24–29 (2011).
71. Shi, Y. & Ethell, I. M. Integrins control dendritic spine plasticity in hippocampal neurons through NMDA receptor and Ca<sup>2+</sup>/calmodulin-dependent protein kinase II-mediated actin reorganization. *J. Neurosci.* **26**, 1813–1822 (2006).
72. Chan, C.-S. *et al.*  $\beta$ 1-Integrins are required for hippocampal AMPA receptor-dependent synaptic transmission, synaptic plasticity, and working memory. *J. Neurosci.* **26**, 223–232 (2006).
73. Sidhu, H., Dansie, L. E., Hickmott, P. W., Ethell, D. W. & Ethell, I. M. Genetic removal of matrix metalloproteinase 9 rescues the symptoms of fragile X syndrome in a mouse model. *J. Neurosci.* **34**, 9867–9879 (2014).
74. Janusz, A. *et al.* The fragile X mental retardation protein regulates matrix metalloproteinase 9 mRNA at synapses. *J. Neurosci.* **33**, 18234–18241 (2013).
75. Bilousova, T. V. *et al.* Minocycline promotes dendritic spine maturation and improves behavioural performance in the fragile X mouse model. *J. Med. Genet.* **46**, 94–102 (2009).
76. Park, Y. K. & Goda, Y. Integrins in synapse regulation. *Nat. Rev. Neurosci.* **17**, 745–756 (2016).
77. Ferrer-Ferrer, M. & Dityatev, A. Shaping synapses by the neural extracellular matrix. *Front. Neuroanat.* **12**, 40 (2018).
78. Cingolani, L. A. *et al.* Activity-dependent regulation of synaptic AMPA receptor composition and abundance by beta3 integrins. *Neuron* **58**, 749–762 (2008).

79. Sidorov, M. S., Auerbach, B. D. & Bear, M. F. Fragile X mental retardation protein and synaptic plasticity. *Mol. Brain* **6**, 15 (2013).
80. Shang, Y. *et al.* Fragile X mental retardation protein is required for chemically-induced long-term potentiation of the hippocampus in adult mice. *J. Neurochem.* **111**, 635–646 (2009).
81. Bostrom, C. A. *et al.* Rescue of NMDAR-dependent synaptic plasticity in Fmr1 knock-out mice. *Cereb Cortex* **25**, 271–279 (2015).
82. Suvrathan, A., Hoefler, C. A., Wong, H., Klann, E. & Chattarji, S. Characterization and reversal of synaptic defects in the amygdala in a mouse model of fragile X syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11591–11596 (2010).
83. Gkogkas, C. G. *et al.* Pharmacogenetic inhibition of eIF4E-dependent Mmp9 mRNA translation reverses fragile X syndrome-like phenotypes. *Cell Rep.* **9**, 1742–1755 (2014).
84. Hynes, R. O. Integrins: Bidirectional, allosteric signaling machines. *Cell* **110**, 673–687 (2002).
85. Pellerin, D., Lortie, A. & Corbin, F. Platelets as a surrogate disease model of neurodevelopmental disorders: Insights from fragile X syndrome. *Platelets* **29**, 113–124 (2018).
86. Dahlhaus, R. Of men and mice: Modeling the fragile X syndrome. *Front. Mol. Neurosci.* **11**, 41 (2018).
87. Boisvert, F.-M. *et al.* A quantitative spatial proteomics analysis of proteome turnover in human cells. *Mol. Cell. Proteomics* **11**, M111.011429 (2012).
88. Zhang, G. *et al.* In-depth quantitative proteomic analysis of de novo protein synthesis induced by brain-derived neurotrophic factor. *J. Proteome Res.* **13**, 5707–5714 (2014).

## Acknowledgements

We warmly thank all participants and their family for their participation in this study. We are grateful to Dr. Artuela Çaku, Pr. Jean-François Lepage and Armita Abolghasemi for their critical review of the manuscript.

## Author contributions

F.C. and O.D. designed the study. F.C. requested funding, supervised the study, recruited fragile X individuals and revised the manuscript. O.D. performed experiments, data analysis, wrote the original draft and recruited control individuals. All authors read and approved the final manuscript.

## Funding

This work was supported by the Centre de Recherche du CHUS and the Foundation of the Stars. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. OD holds a Frederick Banting and Charles Best Canada Graduate Scholarship Master Award from the Canadian Institutes of Health Research (CIHR) and a Master Scholarship from the Fonds de Recherche du Québec—Santé (FRQS).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94027-5>.

**Correspondence** and requests for materials should be addressed to O.D. or F.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)