Toward high spatially resolved proteomics using expansion microscopy

Lauranne Drelich, Soulaimane Aboulouard, Julien Franck, Michel Salzet, Isabelle Fournier, Maxence Wisztorski

Towards High Spatially-Resolved Proteomics using Expansion Microscopy

Lauranne Drelich[1], Soulaimane Aboulouard[1], Julien Franck[1], Michel Salzet[1,2], Isabelle Fournier[1,2]*, Maxence Wisztorski[1]*

[1]Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire Spectrométrie de Masse - PRISM, Lille, 59000, France.

[2]Institut Universitaire de France (IUF), Paris, 75000, France

*Corresponding authors: maxence.wisztorski@univ-lille.fr / isabelle.fournier@univ-lille.fr

ABSTRACT: Expansion Microscopy (EM) is an emerging approach for morphological examination of biological specimens at nanoscale resolution using conventional optical microscopy. To achieve physical separation of cell structures, tissues are embedded in a swellable polymer and expanded several folds in an isotropic manner. This work shows the development and optimization of physical tissue expansion as a new method for spatially resolved large-scale proteomics. Herein, we established a novel method to enlarge the tissue section to be compatible with manual microdissection on regions of interest and MS-based proteomic analysis. A major issue in Expansion Microscopy is the loss of protein information during the mechanical homogenization phase due to the use of Proteinase K. For isotropic expansion, different homogenization agents were investigated, both to maximize protein identification and to minimize protein diffusion. Best results were obtained with SDS for homogenization. Using our modified protocol, we were able to enlarge a tissue section more than 3-fold and identified up to 655 proteins from 1mm in size after expansion, equivalent to 330 µm in their real size corresponding thus to an average of 260 cells. This approach can be performed easily without any expensive sampling instrument. We demonstrated the compatibility of sample preparation for expansion microscopy and proteomic study in a spatial context.

Tissue and tumor heterogeneity have been shown to be a critical issue in oncology, especially with regards to patient care. Besides the spatial organization of this cellular heterogeneity is fundamental to the understanding of the physiopathological mechanisms, promoting the need for cell/tissue molecular mapping[1]. So far, this has been essentially approached by genomic and transcriptomic strategies[1–3]. Indeed, spatially-resolved proteomics remains a challenge due to the limited tools available. Traditionally, organ and tissue analysis relies on a chemical homogenization of the cellular population; leading to an unavoidable loss of spatial making impossible to assess the cellular heterogeneity and the architecture of the tissue.

Mass spectrometry imaging (MSI) is a powerful tool to visualize cellular heterogeneity within a tissue and in different samples[4,5]. MSI information can thus serve to guide spatially resolved omics analysis and identify the molecules contained within regions of interest (ROI) discriminated by a specific molecular profile. To achieve spatially-resolved proteomics, several micro-sampling approaches have been developed including e.g., Laser Capture Microdissection (LCM)[6,7], Liquid MicroJunction (LMJ) micro-extraction[8–10], and hydrogel discs containing trypsin[11,12]. Still, their spatial accuracy is inevitably limited by the sample size that can be obtained, typically hundreds of micrometers. Yet, increasing spatial resolution of spatially-resolved proteomic strategies remains a challenge. Looking in that direction an interesting lead is to explore physically modified tissue section to sample smaller ROIs. It was previously shown that spatial resolution could be increased in MALDI-MSI by stretching the sample. In this method, samples are prepared by adhering tissue section to glass beads array fixed onto a stretchable membrane[13]. Then, the stretching of the membrane separates the tissue section into thousands of cell-sized pieces of tissue which are then further analyzed by MALDI-MSI.

Here we have developed a new strategy to perform spatially resolved proteomics. The originality of this approach is to expand equally in the 3 dimensions, a piece of tissue or a tissue section, to easily select ROIs for shotgun proteomics. The expansion workflow developed is divided into several phases including the expansion, the tissue micro-sampling and the proteomics analysis (**Figure 1A**).

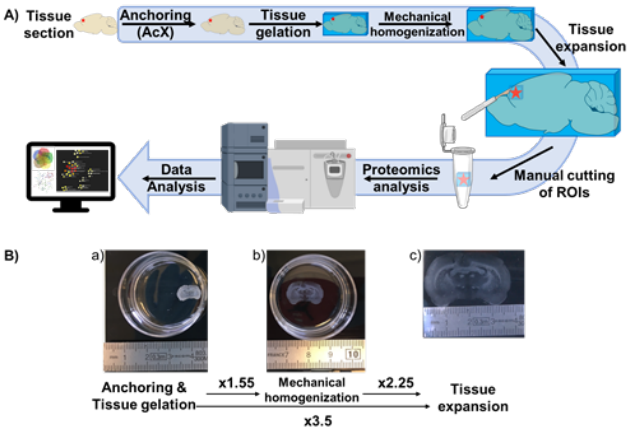

Figure 1. General workflow of the spatially-resolved proteomics based

on tissue expansion A). After tissue expansion, spatially-resolved proteomic analysis is performed by cutting a ROI directly in the gel with a scalpel blade. Gel-embedded tissue pieces are then subjected to enzymatic digestion and peptides are extracted to be analyzed by mass spectrometry. B). Rat brain tissue section prior and after expansion (a-c). Tissue original size is 1.2 cm (a) and is magnified by a factor 1.55 after protein anchoring and tissue gelation (b) to reach an expansion factor of about 3 after the mechanical homogenization and expansion (c) (post-expansion size of 3.8 cm).

Interestingly, this approach can be achieved without any sophisticated or cost expensive instrument. We adapted the protocol originally developed by Boyden et al. for expansion microscopy in which a tissue section can be magnified by 3- to 100-fold[14–16]. This technique initially was developed to image samples at the nanoscale with conventional microscopes. To obtain this expansion, a dense and uniform mesh of swellable polymer is introduced in the tissue. Briefly, tissue sections are deparaffinized and then cover by a solution of AcX. AcX is a linker that binds the primary amine groups on proteins and is then incorporated into the swellable hydrogel polymer during the process of gelation. Thus, the binding created by AcX on the proteins is tethered to the hydrogel polymer chains allowing the biomolecules to retain their spatial organization relative to one other. Thereafter, the mechanical homogenization step consists of suppressing mechanical properties in the sample to keep structural integrity and sample organization during expansion. Disruption of the sample is generally realized by using enzymatic digestion. Proteins are anchored to the polymer network and the resulting hydrogel-tissue hybrid expands after immersion in water, resulting in a physical magnification of the tissue (**Figure 1B**). After some adjustments, we demonstrated that it is possible to cut off a specific expanded region and process the sample to perform MS-based proteomics analysis. Herein, we report a protocol to achieve a 3-fold volumetric expansion and obtain the identification of over 650 proteins for a region of 460 µm original diameter. By doing so, we have overcome some limitations of the existing spatially resolved proteomic methods enabling small ROIs on tissue sections containing a limited quantity of material/cells to be analyzed.

MATERIAL & METHODS

The complete protocols are detailed in supporting information

Reagents and Chemicals

For the different experiments, high purity chemicals were purchased from various suppliers (complete list in the supporting information) and used without further modification.

Tissue expansion

Tissue preparation

Adult Wistar male rats (University of Lille) were sacrificed according to animal use protocols approved by the University of Lille Animal Ethics committee. FFPE and Fresh frozen brain sections of 12 µm thickness were used.

Tissue gelation

Tissue expansion was performed following the protocol published by Tillberg[14]. The complete workflow is presented in **Figure 1A**. Briefly, succinimidyl ester of 6-((acryloyl)amino) hexanoic acid (acryloyl-X (AcX), 0.1mg/mL in PBS) is deposited on tissue and incubated overnight in a humid chamber at room temperature. The gelling solution is freshly prepared. The final concentrations of chemicals in PBS (1x) are 0.01% (w/v) 4-hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl (4-Hydroxy TEMPO), 2M NaCl, 8.6% (w/v) sodium acrylate, 30%(v/v) acrylamide/bis-acrylamide (30% solution; 37.5:1), 0.2% (w/v) TEMED, and 0.2% (w/v) APS. This solution is deposited on the tissue and spread inside a gelation chamber with a glass cover and placed in a humid chamber at 37°C for 2 hours.

Mechanical homogenization

Original protocol used proteinase K (8 units/mL) diluted in the homogenization buffer (50 mM Tris/HCl (pH8), 1 mM EDTA, 0.5% Triton X-100, 1 M NaCl). The gel-containing tissue is submerged in 2mL of proteinase K for 3 hours at 60°C. Several homogenization processes were tested to replace proteinase K. Different enzymes were used but to minimize the total amount of used enzyme, the homogenization solution was deposited directly on the tissue section instead of complete immersion. Our first experiment was conducted using a proteinase K solution diluted to 4 units/mL. In a second step, we tested several other enzymes and homogenization agents. LysC (20µg/mL in 6M Urea, 1mM EDTA) was deposited with a sufficient volume to cover the surface of the tissue section (typically 500 µL) and placed at 37°C overnight. Then, 1mL of Trypsin-EDTA solution conventionally used for cell culture dissociation (0.025% trypsin and 0.01% EDTA in PBS) was deposited and placed at 37°C for 20 minutes. Finally, we tested SDS as homogenization agent. First, we tested 5% SDS in water with 30 minutes at 95°C incubation and then we tested an incubation at a decrease temperature of 58°C in a humidity chamber overnight.

Expansion and cutting

Tissue sections embedded in the gel were separated from the glass slide by submersion in HPLC water for one hour. The water was replaced 3 times every 15 minutes. Once maximal expansion size was reached, square areas of 1x1mm^2 or 5x5mm^2 were cut manually with a scalpel blade and transferred into a microcentrifuge tube. A control region, located outside the tissue, was also cut to evaluate the possible diffusion of proteins and peptides in the gel. The expansion factor, determined from the pre-expansion measurement (**Figure 1Ba**) divided by the post-expansion measurement (**Figure 1Bc**), is usually between 2.9- and 3.4-fold. This means that an area of 5x5 mm^2 post-expansion corresponds to 1.6x1.6 mm^2 real size and about 1x1mm^2 is obtained from a 330x330 µm^2 original area. Thereafter, results will be presented considering the region size post-expansion: 5x5 mm^2 and 1x1 mm^2. A biopsy punch with a diameter of 1.5 mm was tested for sampling regions after tissue expansion. For imaging-like experiments, 15 consecutive regions of 1x1 mm^2 were cut, each along a line through the tissue, each piece of gel being assimilated as an image pixel.

Proteomics analysis

Pieces of gel in the microcentrifuge tube were covered with 50µL of NH_4HCO_3 (50mM). The reduction process was achieved by incubation of the pieces by adding DTT solution (45mM in NH_4HCO_3 50mM) 15min at 50°C. For the alkylation, gel pieces were incubated with IAA (100mM in NH_4HCO_3 50mM) 15min at room temperature in obscurity.

Finally, 20µL of trypsin solution (20µg/mL in NH_4HCO_3 50mM) were added in each sample and incubated overnight at 37°C. Digestion is stopped by adding TFA 1% final volume.

Liquid Microjunction experiment

To compare to another spatially-resolved proteomic method, LMJ based-microextraction was performed according to the previously published protocol[17,18]. Briefly, the ROI was first digested using trypsin solution deposited by a piezoelectric microspotter Chemical Inkjet Printer (CHIP-1000, Shimadzu, CO, Kyoto, Japan), and peptides were extracted using the TriVersa Nanomate platform (Advion Biosciences Inc., Ithaca, NY, USA) with a Liquid Extraction Surface Analysis (LESA) option.

NanoLC-MS & MS/MS analysis

Desalted samples were analyzed by nanoLC-MS/MS (Q Exactive mass spectrometer, Thermo Scientific) (see Support Information for complete description). Identification of proteins and label-free quantification (LFQ) were performed with MaxQuant [19,20] (Version 1.6.1). The search was done against a database containing reviewed proteome for Rattus norvegicus from UniProtKB/Swiss-Prot (8,168 sequences, July 2019). False discovery rates lower than 1% were set at peptides and proteins level. Relative LFQ of proteins was conducted into MaxQuant using the MaxLFQ algorithm[21] with default parameters. The match between run (MBR) feature, with a match time, window of 0.7 min and an alignment window of 20 min, was activated to increase peptide/protein identification.

Analysis of identified proteins was performed using Perseus software (http://www.perseus-framework.org/) (version 1.6.0.7). The file containing the information from the identification was used and hits from the reverse database, proteins with only modified peptides and potential contaminants were removed. The different methods were evaluated in terms of overlap of protein identifications (Venn diagrams) and expressed as Pearson correlation coefficients (dot plots and r value). Visualization in Venn Diagram was performed with BioVenn (http://www.biovenn.nl/) [22]. For quantification-based mass spectrometry imaging, LFQ values of proteins were used to construct images with TIGR Multiexperiment viewer (MEV v 4.9).

The datasets used for analysis were deposited at the ProteomeXchange Consortium[23] (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[24] with the dataset identifier PXD021919 (for review: Username: reviewer_pxd021919@ebi.ac.uk Password: J7TFjxds).

RESULTS AND DISCUSSION

Proteinase K homogenization prevents proteins identification

To achieve expansion-based spatially-resolved proteomics, we need to magnify a tissue section using a swellable polymer allowing, thus, for an isotropic tissue expansion. In the first instance, we based our workflow of tissue expansion on the published protocol from Tillberg et al[14] for the protein-retention expansion microscopy (proEXM) (**Figure 1A)**. Conventionally, tissue sections are covered with the AcX solution. Then the mechanical homogenization step is generally realized by enzymatic digestion with the proteinase K, and the gel is finally submerged in water until maximal expansion **(Figure 1B)**.

We followed this protocol from a FFPE tissue section after deparaffinization and we proceeded with a scalpel blade to a manual cutting post-expansion of small squares of 5x5 mm^2 size on the embedded tissue and outside the tissue (hydrogel alone) as a control to assess possible molecule diffusion within the gel. To facilitate the handling and the manual sampling of the gel, we had to modify the original protocol and use slightly higher concentration of acrylamide/bisacrylamide. This led us to obtain an expansion of 3.1x which is lower than the 4x expansion obtained in the original work and might be attributed to the inter-protein crosslinks already present in the tissue that will not be possibly dissociated by proteinase K. The gel pieces were then submitted to a conventional proteomics digestion protocol and retrieved peptides were analyzed by LC-MS/MS. 14 proteins were identified from the gel piece containing tissue and 9 proteins from the control region **(Figure S1)**. The number of proteins identified is very low compared to what is expected for such a tissue surface. Moreover, an almost equivalent number of proteins is identified outside the tissue section from the control region. The detection of peptides outside the tissue tends to indicate a loss of peptides produced during the homogenization step using the proteinase K. Despite a FFPE tissue was used in this experiment, which normally prevent diffusion because the proteins are cross-linked by paraformaldehyde, the diffusion was not avoided here. Indeed, the AcX creates a link between proteins and hydrogel. Since proteinase K preferentially cleaves at aliphatic of aromatic amino acid residues with low specificity, it leads to the formation of small peptides that are non-anchored to the gel. These small peptides can diffuse through the reticulation of the hydrogel during the different processing steps of expansion or proteomics analysis. This diffusion causes a loss of protein localization and significantly reduce protein identification.

Nevertheless, some proteins could be identified from expanded tissue. These results incite us to investigate the replacement of the proteinase K during the mechanical homogenization step to reduce the protein losses.

Comparative analysis of different alternatives for Proteinase K

As a replacement of proteinase K, we used two enzymes with specific cleavage sites, *i.e.,* trypsin and LysC. We also tested an anionic detergent, the SDS, which allows disrupting protein bonds inducing protein linearization. Using the trypsin or LysC, the expansion factor was estimated to be 2.1-fold; 2.3-fold with SDS, and 2.4-fold with proteinase K **(Figure 2A, B, C, and D)**. The correlation between the size of the sampled regions and their real size according to the expansion factor is given in **Table 1** with an estimation of the corresponding number of cells. For trypsin and LysC, an incomplete homogenization is observed with a lot of cracks and tissue distortion appearing during the expansion process **(Figure 2B and C)**. Using SDS, the expansion is more homogeneous **(Figure 2D),** and the expansion factor is quite similar to the one obtained with the proteinase K. After proteomic analysis, the same low number of proteins were identified using Proteinase K. However, for the other

conditions, the number of identified proteins was significantly increased **(Table 1 and DataS1)**. In the 5x5 mm² regions, up to 536 proteins were quantified using trypsin, 488 for SDS, and 495 for LysC. For the 1x1mm² regions, 35 proteins were quantified for trypsin, 32 for SDS, and 65 for LysC respectively. In the control sample, only 2 proteins were quantified for LysC, 1 for trypsin and none for SDS, revealing no or only very minimal diffusion within the gel.

Table 1. Expansion factor and number of proteins identified according to the homogenizing agent.

	Homogenization	*Expansion Factor*	*Real size (estimated number of cells)*	*Number of proteins ID (control)*
5x5 mm²	Proteinase K	2.4	2.08x2.08 mm² (~6,750)	2 (2)
	Trypsin	2.1	2.38x2.38 mm² (~8,840)	536 (1)
	LysC	2.1	2.38x2.38 mm² (~8,840)	495 (0)
	SDS	2.3	2.17x2.17 mm² (~7,350)	488 (0)
1x1 mm²	Proteinase K	2.4	416x416 µm² (~270)	0 (0)
	Trypsin	2.1	476x476 µm² (~350)	35 (1)
	LysC	2.1	476x476 µm² (~350)	65 (2)
	SDS	2.3	435x435 µm² (~295)	32 (0)

However, as the homogenization with trypsin and LysC was not homogeneous, it is difficult to precisely estimate the real size of the digested region **(Figure 2B and C)**. Based on the measured expansion factor, a variation of about 20% in the number of analyzed cells was observed between the different tests **(Table 1)**. This number of cells is also difficult to estimate when the expansion is not perfectly isotropic as for trypsin and LysC. In these cases, a high number of cracks and apparent deformations were observed, which means that some groups of cells retain their original size, resulting in a higher-than-expected cell density in the sampled region **(Figure 2B and C; Table 1)**. This variability was also observed in each condition of the 1x1 mm² region **(Figure 2E)**. However, for the 5x5 mm² region, a high protein identification overlap (approximately 73%) was observed between Trypsin, Lys C, and SDS **(Figure 2 F).**

Considering these results, we decided to focus our development on SDS to replace proteinase K in the homogenization process. As a homogenization agent, SDS allows lipid removal, keeps protein integrity, and affects only their conformation and charge without disturbing their anchoring to the hydrogel [14,25].

Method optimization and reproducibility for proteomics using SDS homogenization

As a first step, we increased the homogenization time up to one night in a humid chamber with a temperature maintained to 58°C instead of 95°C to avoid liquid evaporation and hydrogel degradation. These improvements result in a greater expansion factor up to 3 times the initial tissue size and an increased number of identified proteins.

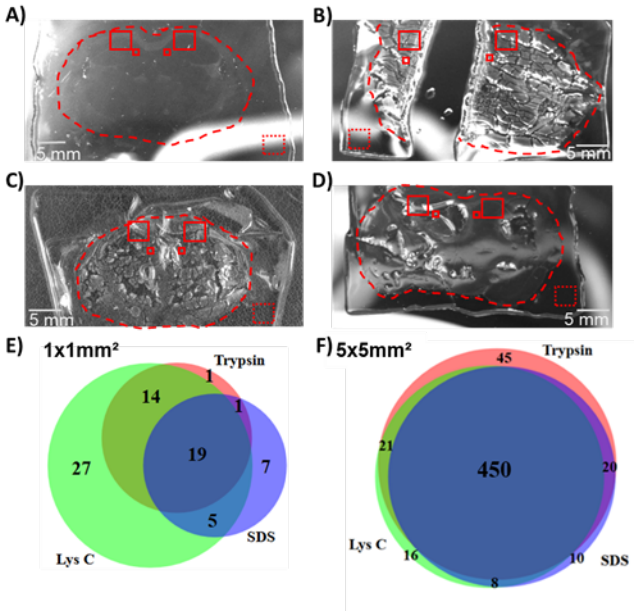

Figure 2. Comparative analysis of different alternatives for Proteinase K. Post-expansion rat brain sections with different homogenization processes using A) Proteinase K, B). Trypsin, C) LysC and D) SDS. The red dotted line delimits the contour of the tissue section. Squares correspond to the different sampled regions: 5x5 mm² and 1x1 mm² inside the tissue and a control area (5x5 mm²) in the surrounding gel only. E) Venn diagrams representation of proteins identified using the different homogenizing agents for E) 1x1 mm² square and F) 5x5 mm².

Reproducibility was assessed by comparing triplicate sampled regions from consecutive FFPE tissue sections (S1, S2, and S3) of the midbrain **(Figure 3A)**. Only 1 protein was identified in the control area from S3, no protein was identified in S1 and S2 control areas which means that proteins keep their location within the tissue despite the longer preparation time **(DataS2)**. Overlapping protein identification (Venn diagrams) was used to evaluate qualitative reproducibility and Pearson's correlation coefficients (dot plots and r-value) for quantitative reproducibility. The number of proteins identified in all replicates, regardless of their abundance, is presented in the Venn diagrams, while the quantitative values (LFQ values) were used to calculate the scatter plots.

For the 5x5 mm² regions, 1259 different proteins were identified with 898, 945, and 994 proteins respectively for replicate 1, 2, and 3 **(Table 2; DataS2)**. Considering the expansion factor, the real analyzed area is around 1.6x1.6 mm². Assuming that cells are round-shaped and 15 µm diameter in average, and that a 12 µm tissue thickness corresponds to a cell monolayer, we can estimate that around 4,000 cells have been analyzed based on 130,000 cells/mm³ updated estimations in neural tissues. The optimization allows doubling the number of protein identifications compared to previous experiments. In **Figure 3B**, the Venn diagram shows that 675 proteins are shared by all replicates corresponding to 53.6% of common proteins with a small

individual variation (about 5%). Looking at the direct side-by-side comparison **(Figure S2A)**, replicates S2 and S3 are closed with high overlap in protein identification (approximately 80%). The protein content of S1 is slightly different from that of the two other replicates (around 60% of overlapping). But for the quantification, the Pearson's correlation coefficient values show that the identified proteins gave the same quantification value between the replicates (mean of 0.956) **(Figure 3C)**.

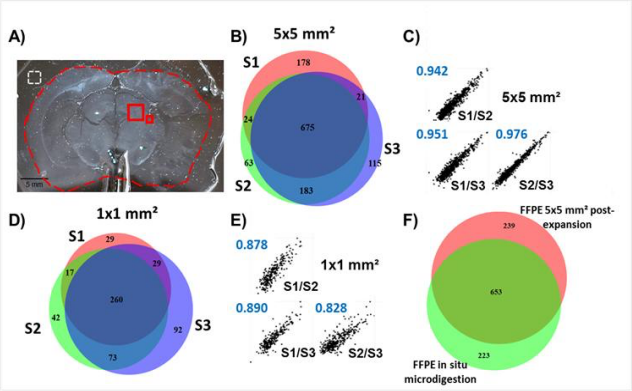

Figure 3. Analysis of proteomic reproducibility using the SDS homogenization protocol. A) Post-expansion widefield image of rat brain section after SDS homogenization, red square stand for the analyzed region and the white square stand for the control region. Experiments were performed in triplicates on 3 consecutive tissue sections. Venn diagram representation of the number of proteins identified in the extraction of B) 5x5 mm^2 in size and D) 1x1 mm^2. Pearson correlation of proteins quantified in the extraction of (C) 5x5 mm^2 in size and (E) 1x1 mm^2. F). Comparison of proteins identified in expansion and with digestion in situ and liquid micro junction extraction.

Reduction of the sampling region to 1x1 mm^2 after expansion corresponds to a real analyzed region of 330x330 μm^2. This surface corresponds to approximately 600 cells from which 335 groups of proteins were identified in section 1; 392 in section 2 and 454 in section 3 for a total of 542 distinct proteins **(Table 2; DataS2)**. A higher variability than for the 5x5vmm^2 region was observed with 260 proteins shared by all replicates representing 47.97% of common proteins **(Figure 3D)** (around 60% for pairwise comparison **(Figure S2B))**. The same observations can be made for the quantitative reproducibility (variation of Pearson's correlation coefficient from 0.83 to 0.89 **(Figure 3E)**. We observed an average of 50% of the proteins identified by the MBR feature with 5x5 mm^2 samples. It should be noted that more than 98% of the proteins are also identified in the 5x5mm^2 regions, suggesting that using a reference sample a larger region containing more cells helps to identify more proteins from smaller samples using the MBR.

Good reproducibility both qualitatively and quantitatively is obtained even though variations exist. The variations can be explained by the evolution of the histological features from the consecutive tissue sections and the difficulties of precisely sampling the same region due to the transparency of the expanded tissue. Reproducibility on smaller regions is lower than on larger ones, probably due to manual cutting and expansion factor variation.

Comparison with fresh frozen tissues, on tissue micro-digestion and micro-extraction

Formalin fixation is advantageous for preservation and conservation of cellular and architectural morphologic detail in tissue sections but results in formaldehyde cross-links limiting that could limit the detection of certain proteins especially for tissues stored over long periods[26]. For this purpose, we also applied the developed strategy to a fresh frozen tissue section with encouraging results but poor reproducibility and great variability. Indeed, we identified a total of 795 proteins in the 5x5 mm^2 samples in all replicates with respectively 136 proteins in section 1, 788 in section 2, and 78 in section 3 **(Table 3 DataS2)**.

Table 2. Expansion factor, estimation of the number of cells and number of proteins identified after homogenization with SDS.

	Section (1,2cm)	*Expansion Factor*	*Real size (estimated number of cells)*	*Number of proteins ID (control)*
5x5mm²	S1	3.08	1.62x1.62 mm^2 (~4,090)	898 (0)
	S2	3.04	1.64x1.64 mm^2 (~4,195)	945 (0)
	S3	3.04	1.64x1.64 mm^2 (~4,195)	994 (1)
1x1mm²	S1	3.08	325x325 μm^2 (~165)	335 (0)
	S2	3.04	329x329 μm^2 (~170)	392 (0)
	S3	3.04	329x329 μm^2 (~170)	454 (1)
punch biopsy	S1	3.08	Ø 454 µm (~250)	582 (0)
	S2	3.04	Ø 460 µm (~260)	682 (0)
	S3	3.04	Ø 460 µm (~260)	700 (1)

As the tissue was not fixed, we observed a significant diffusion of proteins within the gel. This results in a wide variation in the number of identified proteins and a loss of the actual localization (*i.e.,* 10 proteins, 26 and 0 respectively in the control region of sections 1, 2, and 3 - **Table 3**). Interestingly, the mean expansion factor is 3.26 for fresh tissue versus 3.05 for FFPE highlighting the effect of protein cross-linking in FFPE on the expansion. Nonetheless, the expansion factor is lower than observed in conventional protocols (expansion factor of 4), the replacement of proteinase K leading to incomplete homogenization and is a compromise to avoid protein losses. Incomplete homogenization might result in small distortions during the expansion of the tissue though not affecting proteomic because of the scale at which sampling is performed.

We then performed a comparison with another direct surface sampling method using in situ digestion and extraction by liquid microextraction. This microextraction strategy was previously demonstrated in different studies for tissue microenvironment characterization[5,8,10,18,27]. The results obtained after tissue expansion were compared to the one

obtained with optimized expansion-based proteomic workflow. On-tissue enzymatic digestion combined with liquid microextraction performed on a FFPE tissue section allows the identification of 876 proteins from a region of 900 µm in diameter. Compared to the identified proteins in the replicate of the same region from a region of 5x5mm^2 post-expansion FFPE tissue section, we observed 653 commons proteins **(Figure 3F)**. An equivalent number of proteins are identified in the two techniques with 223 proteins unique to the *in situ* micro-digestion/ liquid microextraction and 239 to the expansion proteomics **(DataS3)**.

Use of a biopsy punch for small regions sampling

We observed that sampling the same region in consecutive sections with a scalpel blade results in differences in gel size and thus affect protein identification. In this sense, an instrument capable of cutting a specific region with reproducible size was subsequently used. To improve reproducibility, a biopsy punch was used to cut and remove a tissue disc. This allowed to precisely cut a 1400 µm diameter gel disc (**Figure 4A**) corresponding to ~460 µm diameter region of tissue before expansion (around 260 cells) **(Table 2)**. From three consecutive regions, we obtained a punch diameter of 1400±37 µm. A total of 808 distinct groups of protein were identified in the replicates (582, 682, and 700 in replicate S1, S2, and S3 respectively (**Figure 4B; DataS2**).

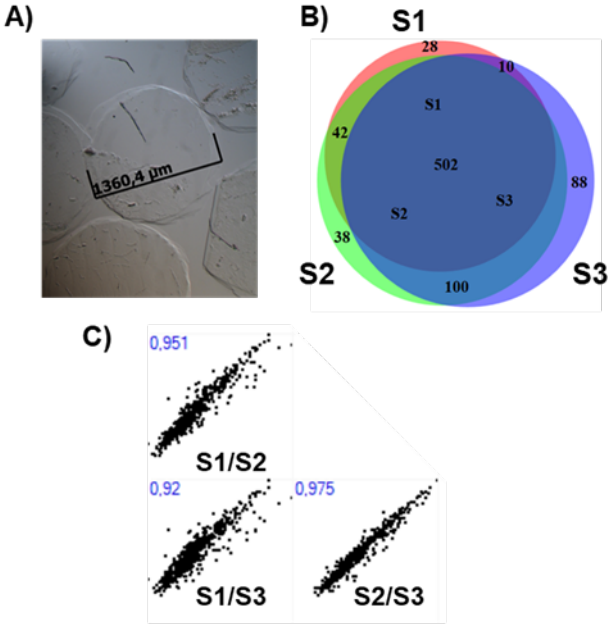

Figure 4. The utilization of punch biopsy improves sampling reproducibility. A) Disk of gel using a punch biopsy. B) Venn diagram representation of the number of proteins identified using punch biopsy extraction. Experiments were performed in triplicates on 3 consecutive tissue sections C) Pearson correlation of LFQ in each replicate.

The Venn diagram shows a good reproducibility between the three replicates with 62.13% of common protein. The quantitative reproducibility is high with an average Pearson's correlation of 0.949 and attains 0.975 between S2 and S3 (**Figure 4C)**. For individual variability between the replicates, the same observations as for manual sampling are made with higher differences between replicates S1 and S3 **(Figure S2C)**.

These results highlight that variation observed in the proteomics content is mainly due to the manual cut of the ROI. Indeed, it is difficult to sample gel pieces of the same size manually. Reproducibility is significantly improved with the use of a punch biopsy. This circular blade allows cutting regions of the same size easily and with precision. Consequently, we observed a reduction of individual variation and a considerable increase of common protein and quantitative reproducibility. A variation is still observed certainly due to minor variation in the position of the sampling due to the transparency of the tissue section. Evolution of the histological features between the tissue sections or variation of the expansion factor could also be a source of variations. Similar results were recently obtained using enzyme delivering hydrogels as a tool for localized analyte extraction directly on the tissue sample[11,12], allowing the identification of about 700 proteins with a hydrogel of 357 µm diameter[12].

Quantification-based mass spectrometry profiling using tissue expansion

The next step consisted of analyzing consecutive adjacent points to perform quantification-based profiling, as we have done in other publications[9,28,29]. To achieve this, we sampled 15 consecutive regions of 1x1mm^2 each along a line through a rat brain tissue section to assimilate each piece of gel to an image pixel **(Figure 5 A)**. It is then possible to use MS-based quantification data for an image reconstruction[28]. All extracts were digested and analyzed by LC-MS/MS. A total of 511 proteins were identified. Images were constructed and each square represents a piece of gel with its position on the tissue. The differences in expressions are color-coded, red represents protein overexpression whereas dark blue represents a low detection **(Figure 5B and C, Table S1)**. Despite a spatial resolution of 330µm, expressions of proteins are different depending on the localization in the tissue.

Expression of proteins specific to each of the region of the rat brain was then examined. The validation of this spatial mapping is confirmed by the overexpression of the Creatine kinase and the Calcium/calmodulin-dependent protein kinase in the cortex and the fimbria of the hippocampus which are known to contain these proteins[30,31]. In these regions, overexpression of Alpha-internexin relative to the thalamus is also obbserved. This protein is an intermediate filament involved in neuron morphogenesis and neurite outgrowth[32]. Cofilin-1 appears to be preferentially expressed in the hippocampus region. This specific protein, as a regulator of actin dynamics, may contribute to degenerative processes[33]. Preferential distribution of GFAP is observed in the thalamus, which is consistent with previous publications[28,30]. In particular, MBP quantification is in line with immunohistology data already published[9]. Proteins from housekeeping genes such as Tubulin beta-4 chain and 14-3-3 protein epsilon were also identified. These proteins are known to present limited variation in expression in tissues and here the calculated coefficient of variation is 2% and 1% respectively for these 2 proteins **(Figure 5C)**.

This strategy provides indirect molecular imaging based on the identification and subsequent quantification of proteins with a spatial resolution of less than 330 µm. It represents an interesting new feature for MSI as it enables to map many high and low abundance proteins. Recently, methods have been developed to increase the number of proteins

identified from a small number of cells. For example, the nanoPOTS[34] method improves the identification of thousands of proteins from a reduced number of cells. The combination of the tissue expansion with this type of method would certainly allow an increase in the number of proteins identified from small regions.

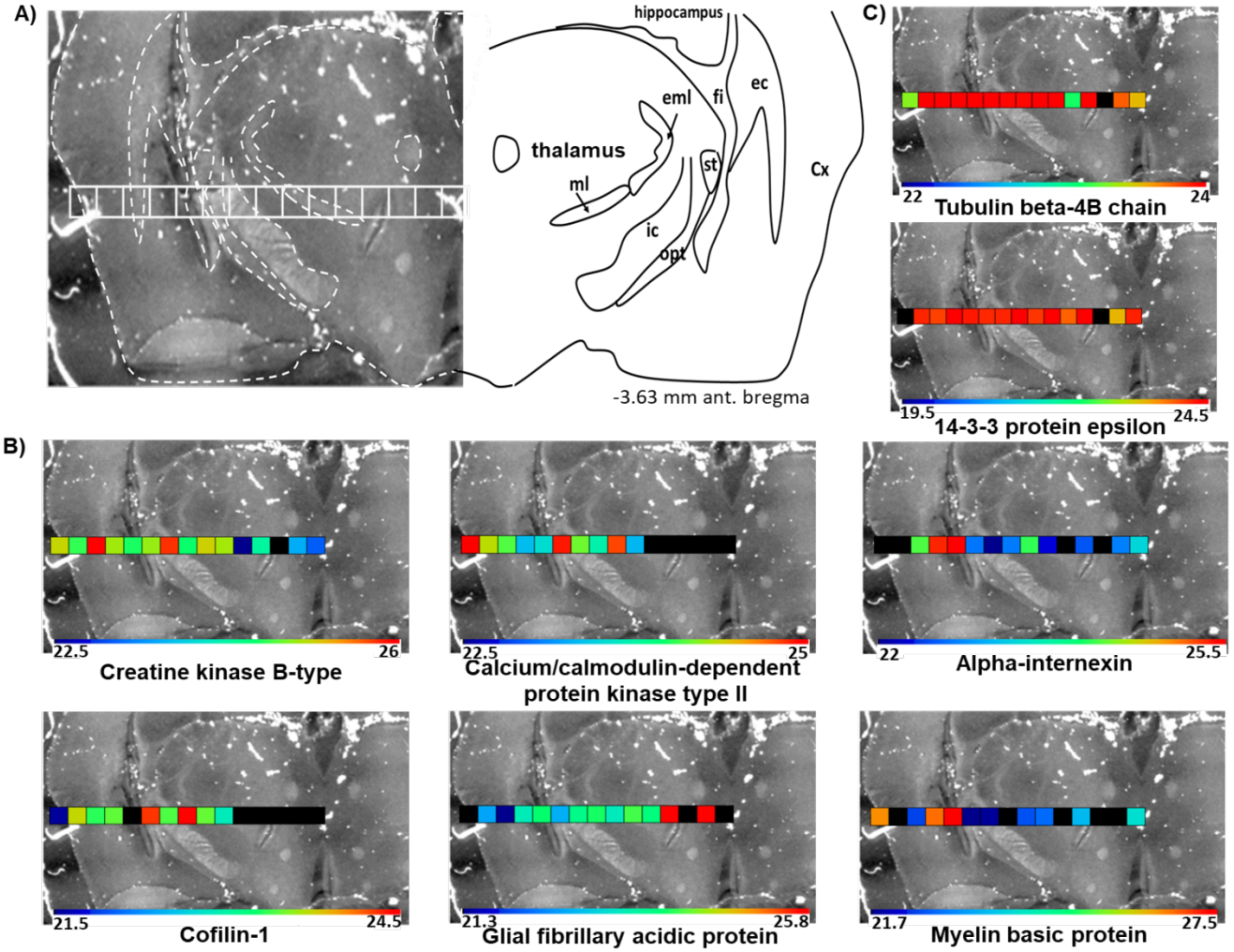

Figure 5. Quantification-based mass spectrometry profiling using tissue expansion A) Image of post-expansion rat brain tissue section, and histological annotations (ec: external capsule; fi: fimbria of the hippocampus; ic: internal capsule; Cx: cerebral cortex; st: stria terminalis; opt: optic tract; eml: external medullary lamina; ml: medial lemniscus). Squares represent extractions for micro proteomic. B) Reconstructed distribution of representative proteins and C) housekeeping genes based on label-free quantification values

CONCLUSION

In conclusion, we demonstrated for the first time that expansion microscopy can be made compatible with conventional large-scale mass spectrometry-based proteomics. Substituting the proteinase K used in the mechanical homogenization step by SDS reduces protein loss due to the generation of small peptides that diffuse through the hydrogel. Using our protocol, a physical expansion of an FFPE tissue section by more than 3-fold was achieved, makes it possible to withdraw a well-controlled area easily and manually with an original size down to 330 µm side. We successfully used this protocol to perform spatially resolved proteomics analysis of regions of interest and to identify more than 655 proteins for around 260 cells. Identification of proteins from expanded tissue showed good qualitative and quantitative reproducibility. This strategy is particularly useful for mapping the protein content of closely related regions on a tissue section which has been challenging in previous spatially resolved proteomics studies. We performed a quantification-based mass spectrometry imaging of more than 500 proteins with a lateral resolution close to 330 µm. This resolution can be further reduced by investigating some new expansion microscopy protocols like the x10 Expansion Microscopy[35], ZOOM (Zoom by hydrOgel cOnversion Microscopy)[36], or iterative expansion microscopy (iExM) [37] to expand the biological sample up to 20- to 100-fold. Interestingly as our protocol is compatible with conventional expansion microscopy, it will be possible to target a protein using fluorescent antibodies[14] and get the proteome from the regions where this protein is expressed. Tissue expansion can also be combined with ExFISH which involves the use of a linker that enables RNA to be covalently attached to the gel[38,39]. Then, fluorescent in situ hybridization (FISH) imaging of mRNA can be done[38] and the region containing fluorescence signal can be delimited, cut, and analyzed to get the corresponding translated protein. Thus, this strategy opens a new way towards simple single-cell proteomics from

tissues and reveal variations in protein expression in cell population for understanding specific disease mechanisms.

Table 3. Expansion factor and number of proteins identified in fresh frozen tissue after homogenization with SDS in each triplicate.

	Section (1,5cm)	*Expansion Factor*	*Real size (estimated number of cells)*	*Number of proteins ID (control)*
5x5mm²	S1	3.26	1.53x1.53mm² (~3,650)	136 (10)
	S2	3.20	1.56x1.56mm² (~3,800)	788 (0)
	S3	3.33	1.50x1.50mm² (~3,510)	78 (26)

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supporting information.docx: contains 2 supporting figures, 1 supporting table, and the supporting methods.
DataS1.xlsx: contains the list of quantified proteins for a post-expansion region of 5x5mm² or 1x1mm² and off tissue using different homogenization agents.
DataS2.xlsx: contains the list of quantified proteins post-expansion in reproducibility experiments using SDS as a homogenization agent.
DataS3.xlsx: contains the list of quantified proteins corresponding to the Venn diagram in post-expansion tissue proteomics vs in situ micro-digestion/liquid microjunction extraction and the quantification data from the quantification-based MS profiling.

AUTHOR INFORMATION

Co-corresponding Authors

* maxence.wisztorski@univ-lille.fr / isabelle.fournier@univ-lille.fr

Author Contributions

All authors made contribution to write the manuscript. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

We acknowledge M. Denfer and P-G. Ferron for their technical assistance. This work was funded by University of Lille, Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation and Inserm.

REFERENCES

(1) Moor, A. E.; Itzkovitz, S. Spatial Transcriptomics: Paving the Way for Tissue-Level Systems Biology. *Curr. Opin. Biotechnol.* **2017**, *46*, 126–133. https://doi.org/10.1016/j.copbio.2017.02.004.

(2) Burrell, R. A.; McGranahan, N.; Bartek, J.; Swanton, C. The Causes and Consequences of Genetic Heterogeneity in Cancer Evolution. *Nature* **2013**, *501* (7467), 338–345. https://doi.org/10.1038/nature12625.

(3) Graham, T. A.; Sottoriva, A. Measuring Cancer Evolution from the Genome. *J. Pathol.* **2017**, *241* (2), 183–191. https://doi.org/10.1002/path.4821.

(4) Quanico, J.; Franck, J.; Wisztorski, M.; Salzet, M.; Fournier, I. Progress and Potential of Imaging Mass Spectrometry Applied to Biomarker Discovery. In *Methods in Molecular Biology*; Methods Mol Biol, 2017; Vol. 1598, pp 21–43. https://doi.org/10.1007/978-1-4939-6952-4_2.

(5) Le Rhun, E.; Duhamel, M.; Wisztorski, M.; Gimeno, J.-P.; Zairi, F.; Escande, F.; Reyns, N.; Kobeissy, F.; Maurage, C.-A.; Salzet, M.; Fournier, I. Evaluation of Non-Supervised MALDI Mass Spectrometry Imaging Combined with Microproteomics for Glioma Grade III Classification. *Biochim. Biophys. Acta - Proteins Proteomics* **2017**, *1865* (7), 875–890. https://doi.org/10.1016/j.bbapap.2016.11.012.

(6) De Marchi, T.; Braakman, R. B. H.; Stingl, C.; van Duijn, M. M.; Smid, M.; Foekens, J. A.; Luider, T. M.; Martens, J. W. M.; Umar, A. The Advantage of Laser-Capture Microdissection over Whole Tissue Analysis in Proteomic Profiling Studies. *Proteomics* **2016**. https://doi.org/10.1002/pmic.201600004.

(7) Longuespée, R.; Baiwir, D.; Mazzucchelli, G.; Smargiasso, N.; De Pauw, E. Laser Microdissection-Based Microproteomics of Formalin-Fixed and Paraffin-Embedded (FFPE) Tissues. In *Methods in Molecular Biology*; 2018; Vol. 1723, pp 19–31. https://doi.org/10.1007/978-1-4939-7558-7_2.

(8) Quanico, J.; Franck, J.; Dauly, C.; Strupat, K.; Dupuy, J.; Day, R.; Salzet, M.; Fournier, I.; Wisztorski, M. Development of Liquid Microjunction Extraction Strategy for Improving Protein Identification from Tissue Sections. *J. Proteomics* **2013**, *79*, 200–218. https://doi.org/10.1016/j.jprot.2012.11.025.

(9) Wisztorski, M.; Desmons, A.; Quanico, J.; Fatou, B.; Gimeno, J.-P.; Franck, J.; Salzet, M.; Fournier, I. Spatially-Resolved Protein Surface Microsampling from Tissue Sections Using Liquid Extraction Surface Analysis. *Proteomics* **2016**, *16* (11–12), 1622–1632. https://doi.org/10.1002/pmic.201500508.

(10) Delcourt, V.; Franck, J.; Leblanc, E.; Narducci, F.; Robin, Y. M.; Gimeno, J. P.; Quanico, J.; Wisztorski, M.; Kobeissy, F.; Jacques, J. F.; Roucou, X.; Salzet, M.; Fournier, I. Combined Mass Spectrometry Imaging and Top-down Microproteomics Reveals Evidence of a Hidden Proteome in Ovarian Cancer. *EBioMedicine* **2017**, *21*, 55–64. https://doi.org/10.1016/j.ebiom.2017.06.001.

(11) Harris, G. a; Nicklay, J. J.; Caprioli, R. M. Localized in Situ Hydrogel-Mediated Protein Digestion and Extraction Technique for on-Tissue Analysis. *Anal. Chem.* **2013**, *85* (5), 2717–2723. https://doi.org/10.1021/ac3031493.

(12) Rizzo, D. G.; Prentice, B. M.; Moore, J. L.; Norris, J. L.; Caprioli, R. M. Enhanced Spatially Resolved Proteomics Using On-Tissue Hydrogel-Mediated Protein Digestion. *Anal. Chem.* **2017**, *89* (5), 2948–2955. https://doi.org/10.1021/acs.analchem.6b04395.

(13) Monroe, E. B.; Jurchen, J. C.; Koszczuk, B. A.; Losh, J. L.; Rubakhin, S. S.; Sweedler, J. V. Massively Parallel Sample Preparation for the MALDI MS Analyses of Tissues. *Anal. Chem.* **2006**, *78* (19), 6826–6832. https://doi.org/10.1021/ac060652r.

(14) Tillberg, P. W.; Chen, F.; Piatkevich, K. D.; Zhao, Y.; Yu, C.-C.; English, B. P.; Gao, L.; Martorell, A.; Suk, H.-J.; Yoshida, F.; DeGennaro, E. M.; Roossien, D. H.; Gong, G.; Seneviratne, U.; Tannenbaum, S. R.; Desimone, R.; Cai, D.; Boyden, E. S. Protein-Retention Expansion Microscopy of Cells and Tissues Labeled Using Standard Fluorescent Proteins and Antibodies. *Nat. Biotechnol.* **2016**, *34* (9), 987–992. https://doi.org/10.1038/nbt.3625.

(15) Chen, F.; Tillberg, P. W.; Boyden, E. S. Expansion Microscopy. *Science (80-.).* **2015**, *347* (6221), 543–548. https://doi.org/10.1126/science.1260088.

(16) Zhao, Y.; Bucur, O.; Irshad, H.; Chen, F.; Weins, A.; Stancu, A. L.; Oh, E.-Y.; DiStasio, M.; Torous, V.; Glass, B.; Stillman, I. E.; Schnitt, S. J.; Beck, A. H.; Boyden, E. S. Nanoscale Imaging of Clinical Specimens Using Pathology-Optimized Expansion Microscopy. *Nat. Biotechnol.* **2017**, *35* (8), 757–764. https://doi.org/10.1038/nbt.3892.

(17) Wisztorski, M.; Quanico, J.; Franck, J.; Fatou, B.; Salzet, M.; Fournier, I. Droplet-Based Liquid Extraction for Spatially-Resolved Microproteomics Analysis of Tissue Sections. In *Methods in molecular biology (Clifton, N.J.)*; Humana Press, New York, NY, 2017; Vol. 1618, pp 49–63. https://doi.org/10.1007/978-1-4939-7051-3_6.

(18) Wisztorski, M.; Fatou, B.; Franck, J.; Desmons, A.; Farré, I.; Leblanc, E.; Fournier, I.; Salzet, M. Microproteomics by Liquid Extraction Surface Analysis: Application to FFPE Tissue to Study the Fimbria Region of Tubo-Ovarian Cancer. *PROTEOMICS - Clin. Appl.* **2013**, *7* (3–4), 234–240. https://doi.org/10.1002/prca.201200070.

(19) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. https://doi.org/10.1038/nbt.1511.

(20) Tyanova, S.; Temu, T.; Carlson, A.; Sinitcyn, P.; Mann, M.; Cox, J. Visualization of LC-MS/MS Proteomics Data in MaxQuant. *Proteomics* **2015**, *15* (8), 1453–1456. https://doi.org/10.1002/pmic.201400449.

(21) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **2014**, *13* (9), 2513–2526. https://doi.org/10.1074/mcp.M113.031591.

(22) Hulsen, T.; de Vlieg, J.; Alkema, W. BioVenn – a Web Application for the Comparison and Visualization of Biological Lists Using Area-Proportional Venn Diagrams. *BMC Genomics* **2008**, *9* (1), 488. https://doi.org/10.1186/1471-2164-9-488.

(23) Vizcaíno, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Ríos, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P.-A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H.-J.; Albar, J. P.; Martinez-Bartolomé, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange Provides Globally Coordinated Proteomics Data Submission and Dissemination. *Nat. Biotechnol.* **2014**, *32* (3), 223–226. https://doi.org/10.1038/nbt.2839.

(24) Vizcaíno, J. A.; Côté, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; O'Kelly, G.; Schoenegger, A.; Ovelleiro, D.; Pérez-Riverol, Y.; Reisinger, F.; Ríos, D.; Wang, R.; Hermjakob, H. The PRoteomics IDEntifications (PRIDE) Database and Associated Tools: Status in 2013. *Nucleic Acids Res.* **2013**, *41* (Database issue), D1063-9. https://doi.org/10.1093/nar/gks1262.

(25) Ku, T.; Swaney, J.; Park, J.-Y.; Albanese, A.; Murray, E.; Cho, J. H.; Park, Y.-G.; Mangena, V.; Chen, J.; Chung, K. Multiplexed and Scalable Super-Resolution Imaging of Three-Dimensional Protein Localization in Size-Adjustable Tissues. *Nat. Biotechnol.* **2016**, *34* (9), 973–981. https://doi.org/10.1038/nbt.3641.

(26) Wiśniewski, J. R. Proteomic Sample Preparation from Formalin Fixed and Paraffin Embedded Tissue. *J. Vis. Exp.* **2013**, No. 79. https://doi.org/10.3791/50589.

(27) Delcourt, V.; Franck, J.; Quanico, J.; Gimeno, J.-P.; Wisztorski, M.; Raffo-Romero, A.; Kobeissy, F.; Roucou, X.; Salzet, M.; Fournier, I. Spatially-Resolved Top-down Proteomics Bridged to MALDI MS Imaging Reveals the Molecular Physiome of Brain Regions. *Mol. Cell. Proteomics* **2018**, *17* (2), 357–372. https://doi.org/10.1074/mcp.M116.065755.

(28) Franck, J.; Quanico, J.; Wisztorski, M.; Day, R.; Salzet, M.; Fournier, I. Quantification-Based Mass Spectrometry Imaging of Proteins by Parafilm Assisted Microdissection. *Anal. Chem.* **2013**, *85* (17), 8127–8134. https://doi.org/10.1021/ac4009397.

(29) Quanico, J.; Franck, J.; Gimeno, J. P.; Sabbagh, R.; Salzet, M.; Day, R.; Fournier, I. Parafilm-Assisted Microdissection: A Sampling Method for Mass Spectrometry-Based Identification of Differentially Expressed Prostate Cancer Protein Biomarkers. *Chem. Commun.* **2015**, *51* (22), 4564–4567. https://doi.org/10.1039/c4cc08331h.

(30) Katagiri, T.; Hatano, N.; Aihara, M.; Kawano, H.; Okamoto, M.; Liu, Y.; Izumi, T.; Maekawa, T.; Nakamura, S.; Ishihara, T.; Shirai, M.; Mizukami, Y. Proteomic Analysis of Proteins Expressing in Regions of Rat Brain by a Combination of SDS-PAGE with Nano-Liquid Chromatography-Quadrupole-Time of Flight Tandem Mass Spectrometry. *Proteome Sci.* **2010**, *8* (1), 41. https://doi.org/10.1186/1477-5956-8-41.

(31) Erondu, N.; Kennedy, M. Regional Distribution of Type II Ca2+/Calmodulin-Dependent Protein Kinase in Rat Brain. *J. Neurosci.* **1985**, *5* (12), 3270–3277. https://doi.org/10.1523/JNEUROSCI.05-12-03270.1985.

(32) Bott, C. J.; Winckler, B. Intermediate Filaments in Developing Neurons: Beyond Structure. *Cytoskeleton* **2020**, *77* (3–4), 110–128. https://doi.org/10.1002/cm.21597.

(33) Bamburg, J. R.; Bernstein, B. W.; Davis, R. .; Flynn, K. C.; Goldsbury, C.; Jensen, J. R.; Maloney, M. T.; Marsden, I. T.; Minamide, L. S.; Pak, C. W.; Shaw, A. E.; Whiteman, I.; Wiggan, O. ADF/Cofilin-Actin Rods in Neurodegenerative Diseases. *Curr. Alzheimer Res.* **2010**, *7* (3), 241–250. https://doi.org/10.2174/156720510791050902.

(34) Piehowski, P. D.; Zhu, Y.; Bramer, L. M.; Stratton, K. G.; Zhao, R.; Orton, D. J.; Moore, R. J.; Yuan, J.; Mitchell, H. D.; Gao, Y.; Webb-Robertson, B.-J. M.; Dey, S. K.; Kelly, R. T.; Burnum-Johnson, K. E. Automated Mass Spectrometry Imaging of over 2000 Proteins from Tissue Sections at 100-Mm Spatial Resolution. *Nat. Commun.* **2020**, *11* (1), 8. https://doi.org/10.1038/s41467-019-13858-z.

(35) Truckenbrodt, S.; Maidorn, M.; Crzan, D.; Wildhagen, H.; Kabatas, S.; Rizzoli, S. O. X10 Expansion Microscopy Enables 25-nm Resolution on Conventional Microscopes. *EMBO Rep.* **2018**, *19* (9). https://doi.org/10.15252/embr.201845836.

(36) Park, H.; Choi, D.; Park, J. S.; Sim, C.; Park, S.; Kang, S.; Yim, H.; Lee, M.; Kim, J.; Pac, J.; Rhee, K.; Lee, J.; Lee, Y.; Lee, Y.; Kim, S. Scalable and Isotropic Expansion of Tissues with Simply Tunable Expansion Ratio. *Adv. Sci.* **2019**, *6* (22), 1901673. https://doi.org/10.1002/advs.201901673.

(37) Chang, J.-B.; Chen, F.; Yoon, Y.-G.; Jung, E. E.; Babcock, H.; Kang, J. S.; Asano, S.; Suk, H.-J.; Pak, N.; Tillberg, P. W.; Wassie, A. T.; Cai, D.; Boyden, E. S. Iterative Expansion Microscopy. *Nat. Methods* **2017**, *14* (6), 593–599. https://doi.org/10.1038/nmeth.4261.

(38) Chen, F.; Wassie, A. T.; Cote, A. J.; Sinha, A.; Alon, S.; Asano, S.; Daugharthy, E. R.; Chang, J.-B.; Marblestone, A.; Church, G. M.; Raj, A.; Boyden, E. S. Nanoscale Imaging of RNA with Expansion Microscopy. *Nat. Methods* **2016**, *13* (8), 679–684. https://doi.org/10.1038/nmeth.3899.

(39) Asano, S. M.; Gao, R.; Wassie, A. T.; Tillberg, P. W.; Chen, F.; Boyden, E. S. Expansion Microscopy: Protocols for Imaging Proteins and RNA in Cells and Tissues. *Curr. Protoc. Cell Biol.* **2018**, *80* (1), e56. https://doi.org/10.1002/cpcb.56.

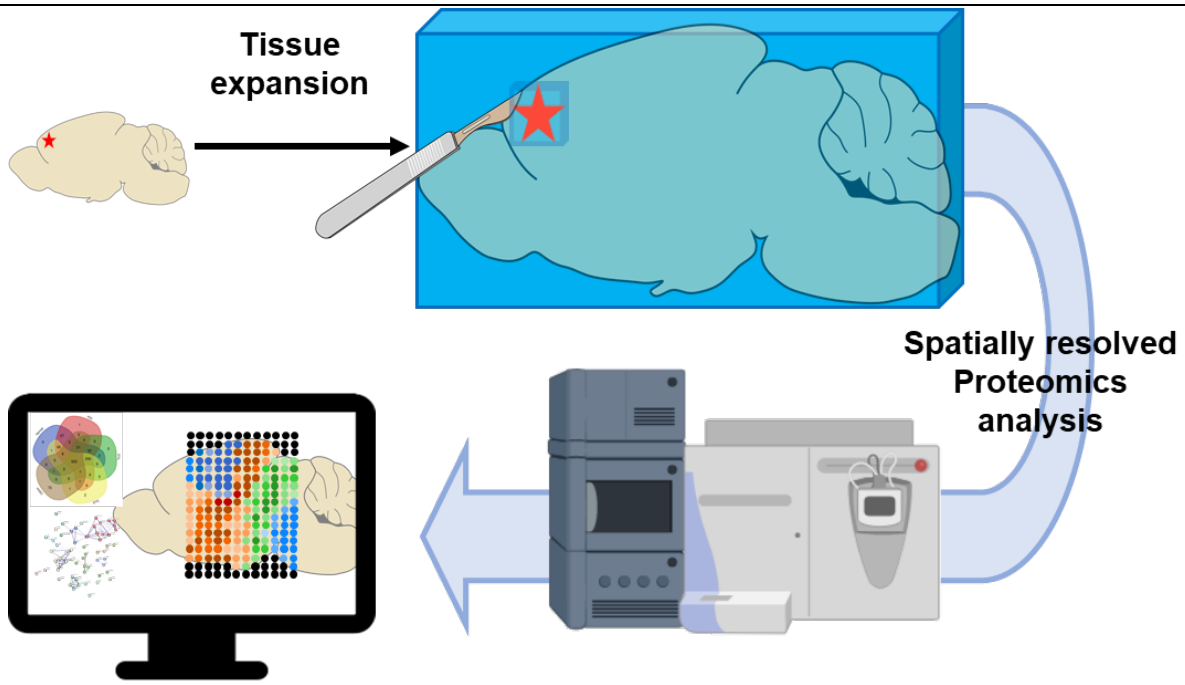

Table of Contents artwork