Received: 6 July 2021 | Accepted: 23 August 2021

DOI: 10.1002/bab.2249

ORIGINAL ARTICLE

Selymatra: A web application for protein-profiling analysis of mass spectra

Davide Nardone[1] | **Angelo Ciaramella**[1] | **Mariangela Cerreta**[2] | **Salvatore Pulcrano**[3] | **Gian C. Bellenchi**[3] | **Linda Leone**[4] | **Giuseppe Manco**[2,§] | **Ferdinando Febbraio**[2,§]

[1] Department of Science and Technology, University of Naples Parthenope, Naples, Italy

[2] Institute of Biochemistry and Cellular Biology, National Research Council, Naples 80131, Italy

[3] Institute of Genetic and Biophysics "A. Buzzati Traverso", National Research Council, Naples, Italy

[4] Department of Chemical Sciences, University of Naples "Federico II", Naples 80126, Italy

Correspondence
Ferdinando Febbraio, Institute of Biochemistry and Cellular Biology, National Research Council, Naples 80131, Italy.
Email: ferdinando.febbraio@cnr.it;
Angelo Ciaramella, Department of Science and Technology, University of Naples Parthenope, Naples, Italy.
Email: angelo.ciaramella@uniparthenope.it

[§]these authors contribute equally to the work (co-last)

Abstract

Surface enhanced laser desorption/ionization-time of flight (SELDI-TOF) mass spectrometry is a variant of the matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry. It is used in many cases especially for the analysis of protein profiling and for preliminary screening of biomarkers in complex samples. Unfortunately, these analyses are time consuming and protein identification is generally strictly limited. SELDI-TOF analysis of mass spectra (SELYMATRA) is a web application (WA) developed to reduce these limitations by (i) automating the identification processes and (ii) introducing the possibility to predict proteins in complex mixtures from cells and tissues. The WA architectural pattern is the model-view-controller, commonly used in software development. The WA compares the mass value between two mass spectra (sample vs. control) to extract differences, and, according to the set parameters, it queries a local database to predict most likely proteins based on their masses and different expression amplification. The WA was validated in a cellular model overexpressing a tagged NURR1 receptor, being able to recognize the tagged protein in the profiling of transformed cells. A help page, including a description of parameters for WA use, is available on the website.

KEYWORDS
MALDI-TOF, protein prediction, protein profiling, SELDI-TOF, SELYMATRA

1 | INTRODUCTION

Comparisons of the protein profiling incellular systems could give important information about the changes in the protein expression after cell alterations such as under oxidative stress or in cancer,[1–4] in order to better understand the cell answer to these events, and for the discovery of disease signals to be used as diagnostic markers or targets for drugs.[5–7] In the last

Abbreviations: Avg, average; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; MSMS, tandem mass spectrometry; MVC, model-view-controller; MW, molecular weight; NURR1, nuclear receptor related 1; pI, isoelectric point; Q, quadrupole; SELDI, Surface enhanced laser desorption/ionization; TET, tetracycline; TOF, time of flight; WA, web application.

decades, several mass spectrometry (MS) instruments and methodologies have been developed in order to improve the results of protein profiling. Surface enhanced laser desorption/ionization-time of flight (SELDI-TOF) mass spectrometry is a variant of the matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry, especially used for the high-throughput analysis of protein profiles in cell samples[8–10] as well as for the characterization of protein-based nanodevices.[11,12] This technique exploits active matrices to select/sort proteins on the basis of their chemical–physical characteristics.[13] Despite the high potential of this technique, the lack of protein identification and the difficulty to analyze complex spectra reduce the fields of SELDI-TOF application.[14] In particular, a large amount of data must be accumulated in order to identify putative biomarkers, and often pretreatment steps of samples become necessary to enhance differences and reduce the background noise in the MS spectra. In addition, several other techniques are later necessary for the identification of biomarkers and protein–protein interactions, such as sequencing by tandem quadrupole-time of flight mass spectrometry (MSMS-QTOF), immunoprecipitation, and Western blotting.[9,15–18] These limits made it necessary to develop new strategies in order to improve these methodologies.

The SELDI-TOF MS analysis process can be summarized in the following phases: (i) spectra acquisition; (ii) data preprocessing (baseline subtraction, normalization); (iii) clusterization; (iv) statistical analysis. Bioinformatic tools are crucial in the last phase, for the comparison of protein peak patterns obtained from different samples. One of the most common algorithms used for this purpose is the classification and regression tree.[19] Decision trees are easy to use and interpret, but their application is not suited to complex sets of data, where multidimensional relationships between SELDI peaks are expected. Alternatively, to methods based on decision trees, different software have been employed for biostatistical cluster analysis, based either on the classic linear regression algorithm, such as the ProPeak Software package[20] or the logistic algorithm.[21] Typically, these methods require a deep knowledge of mathematics and statistics to be efficiently used, limiting their accessibility to more experienced users.

In this work, we present a new web application (WA), named SELYMATRA, for the SELDI-TOF analysis of mass spectra (http://140.164.61.25:5055/SELYMATRA). The WA was developed with the aim of significantly reducing the analysis time on large amounts of MS data, and at the same time to obtain information on proteomic profiles so far not obtainable. This WA is able to predict clusters of proteins starting from the comparison among proteome profiles obtained by SELDI-TOF MS measurements. The architectural pattern used to develop the WA is the model-view-controller (MVC), extremely used in the development of software systems (Figure 1). The WA expects a user to upload a Microsoft Excel spreadsheet file format, which is the only format whereby the application works, usually generated by means of proprietary MS software, such as SELDI-TOF software. The WA was validated using immortalized neuronal cells overexpressing a tagged NURR1 receptor. Using our WA, we were able to recognize the tagged protein in the profiling of transformed cells with respect to the profiling of neuronal cells not expressing it.

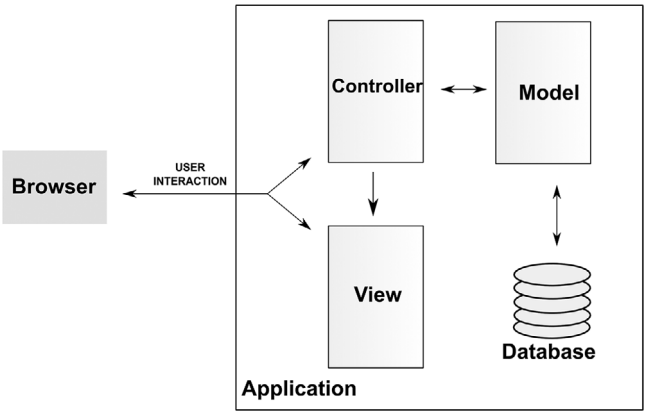

FIGURE 1 Scheme of model-view-controller used for WA design. The MVC model describes how the application represents the data, including the database management (model), how the information is displayed at the end-user (view), and how the interface is managed and controlled by the user (controller)

2 | RESULTS AND DISCUSSION

The type of data treated by the application is strictly related to the SELDI-TOF or MALDI-TOF MS technology and their own software. Samples from different biological sources can be used, such as plasma, serum, blood, etc. The WA aims to carry out a particular pattern recognition analysis, presently not included in any proprietary SELDI software package. This is accomplished by looking for specific similarities between two mass spectra under different washing and binding conditions, and in doing that, it yields to a deeper data analysis, which significantly reduces the duration of troublesome proteomic experiments (Figure 2).

Once a file is uploaded, the WA applies a method of data analysis and processing, as follows:

1. Mass spectra comparison process;
2. Recognition process;
3. Protein ID matching process.

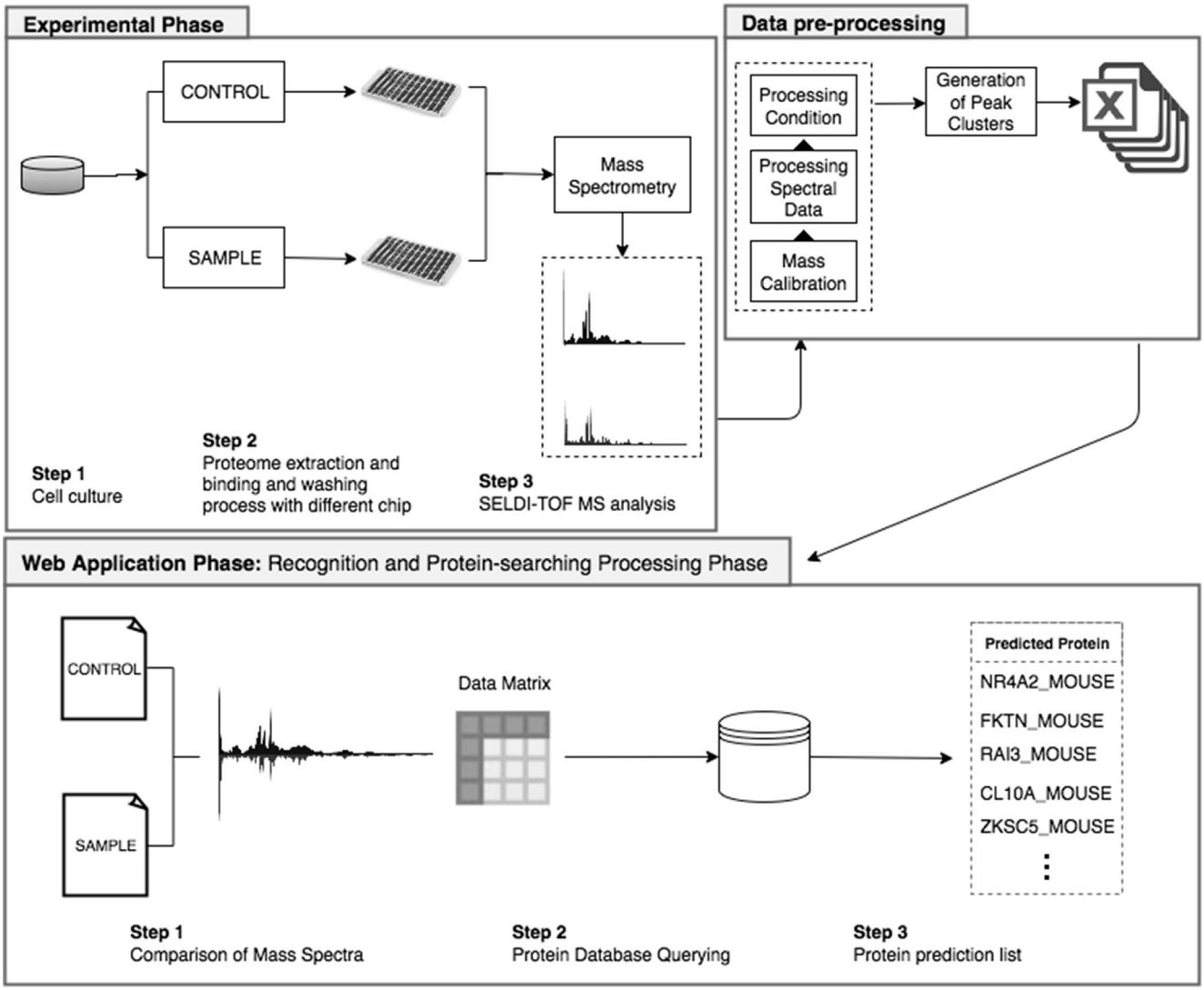

FIGURE 2 Work-flow of experimental data production, preprocessing, and data elaboration. The process was described by three phases: an experimental phase for the production of MS spectra; followed by a preprocessing of data for the generation of peak clusters; and a final phase involving the WA, which compares the MS spectra clusters and performs the protein database searching and prediction

Through these three actions, the WA elaborates the data to perform a key feature extraction among two mass spectra and give differences as well as up- and downregulation in protein expressions, respectively. According to these results and to the characteristics of the experiment, it queries a local database for the prediction of the proteins that are most likely related to the previous results (Figure 2).

2.1 | Mass spectra comparison and recognition processes

The mass spectra comparison and recognition processes are fully tunable by the choice and values of some parameters. Based on the user parameter selection, different results may be obtained. Basically, each procedure is separately developed, but the underlying idea is quite similar. That is to say, a math inequality is used for both modeling procedure, by which a result set of values is generated:

$$\left(A_{m/z_{\text{avg}}} Am/z_{\text{avg}} - \sigma_{m/z}\right) * \delta < B_{m/z} Bm/z < \left(A_{m/z_{\text{avg}}} A\frac{m}{z}_{\text{avg}} + \sigma_{m/z}\right) * \delta, \quad (1)$$

where $A_{m/z\text{avg}}$ and $\sigma_{m/z}$ are the mean value and the standard deviation of the sample (A), respectively, δ is a correction factor used for increasing or diminishing the search range, and $B_{m/z}$ is the current m/z value to find in the control sample (B).

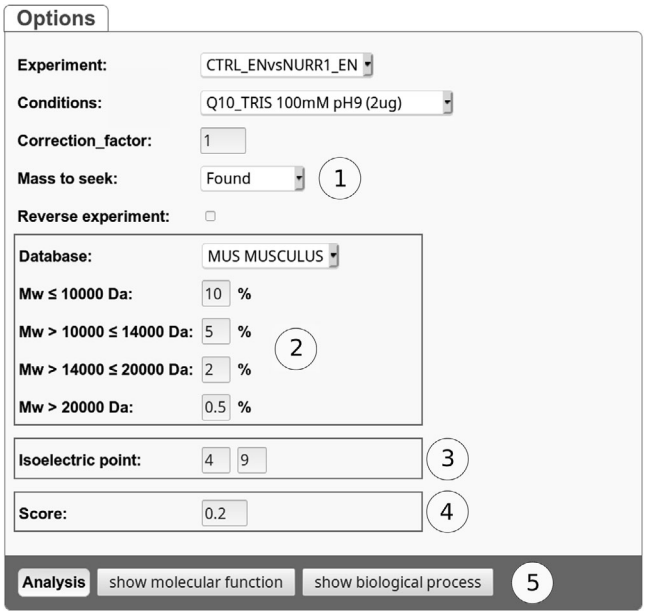

FIGURE 3 WA option panel. This panel contains all the parameters used in SELYMATRA to recognize and predict proteins. Almost all parameters are editable by users (for their complete description, see Supplementary Table 2)

According to the user needs, four types of m/z can be identified for any experiment (Figure 3 point 1):

- *Found*: A set of m/z_{avg} satisfying Equation (1);
- *Missing*: A set of m/z_{avg} not satisfying Equation (1) or rather the complement ensemble to the found (A) masses against all the (B) masses, formally expressed as

$$A = B - A = \left\{ m/z_{\text{avg}} \in B \wedge m/z_{\text{avg}} \notin A \right\}. \quad (2)$$

- *Upgrade or downgrade*: A set of m/z_{avg} satisfying Equation (1) and the median intensity of mass variation constraint, expressed as

$$\frac{B_{m/z}Bm/z - A_{m/z_{\text{avg}}}Am/z_{\text{avg}}}{A_{m/z_{\text{avg}}}Am/z_{\text{avg}}} * 100 \leq \lambda, \quad (3)$$

where $B_{m/z\text{avg}}$ and $A_{m/z\text{avg}}$ are the m/z_{avg} of the control sample (B) and the running sample (A), respectively (which both satisfy Equation (1)) and λ is an index of variation of the intensity median for distinguishing the affinity among the relative intensities of two mass peaks.

2.2 | Protein ID matching process

The set of the identified masses alone, as obtained by the mass spectra comparison and recognition process, is not sufficient to predict the protein identities. Therefore, we have built up a process to infer more information on such masses, which helps in predicting the most likely protein corresponding to a specific m/z value. In order to make this possible, two further features were considered for each mass of the result set:

- a dynamic amplification factor (depending on the m/z value), used for the assessment of a molecular mass (MW) error-range (Figure 3 point 2);
- a type of chip and binding/washing condition, for the assessment of an isoelectric point (pI) range (Figure 3 point 3).

Both these features help in defining a confidence interval that limits the protein matching process. ID matching process is possible through the building, maintaining, and tuning of a database of proteins, whose data have been retrieved through java API provided by Uniprot database.[22]

A score parameter (Figure 3 point 4) was added in order to set the threshold for the range of confidence used for the identification/prediction of proteins. Lower values correspond to a strict rule for the correspondence of masses to assigned proteins. All these parameters are editable by the user in order to allow the software to adapt to his/her own experiments. The analysis button starts the processing of data resulting in an output listing the predicted proteins. In addition, two other buttons, activating a graphical report of the predicted proteins ordered according to their functions or their biological role, were implemented (Figure 3 point 5). This tool produces a histogram representation of the data, allowing for a fast comparison between different measurements of the same sample (Figure 4).

2.3 | Validation of SELYMATRA

The WA should be able to recognize rough patterns of proteins related to the phenomena represented in the target sample. An example was the identification, in the nuclei of the A1 cell line, of the transcription factor 3 × FLAG-Nurr1,[23] overexpressed in the expression cassette TET-ON. We analyzed two different sample conditions for each sample (overexpressing or not the transcription factor 3 × FLAG-Nurr1). The comparison of clustered files generated by the ProteinChip Data Manager software was only able to indicate a single MS peak in the range of masses (~71 to 72 KDa) of the transcription factor 3 × FLAG-Nurr1 (Figure 5).

As expected, SELYMATRA associated the single MS peak at about 72 KDa to the 3 × FLAG-Nurr1 ID, which was present in the local database (Figure 5), demonstrating good confidence in the data analysis.

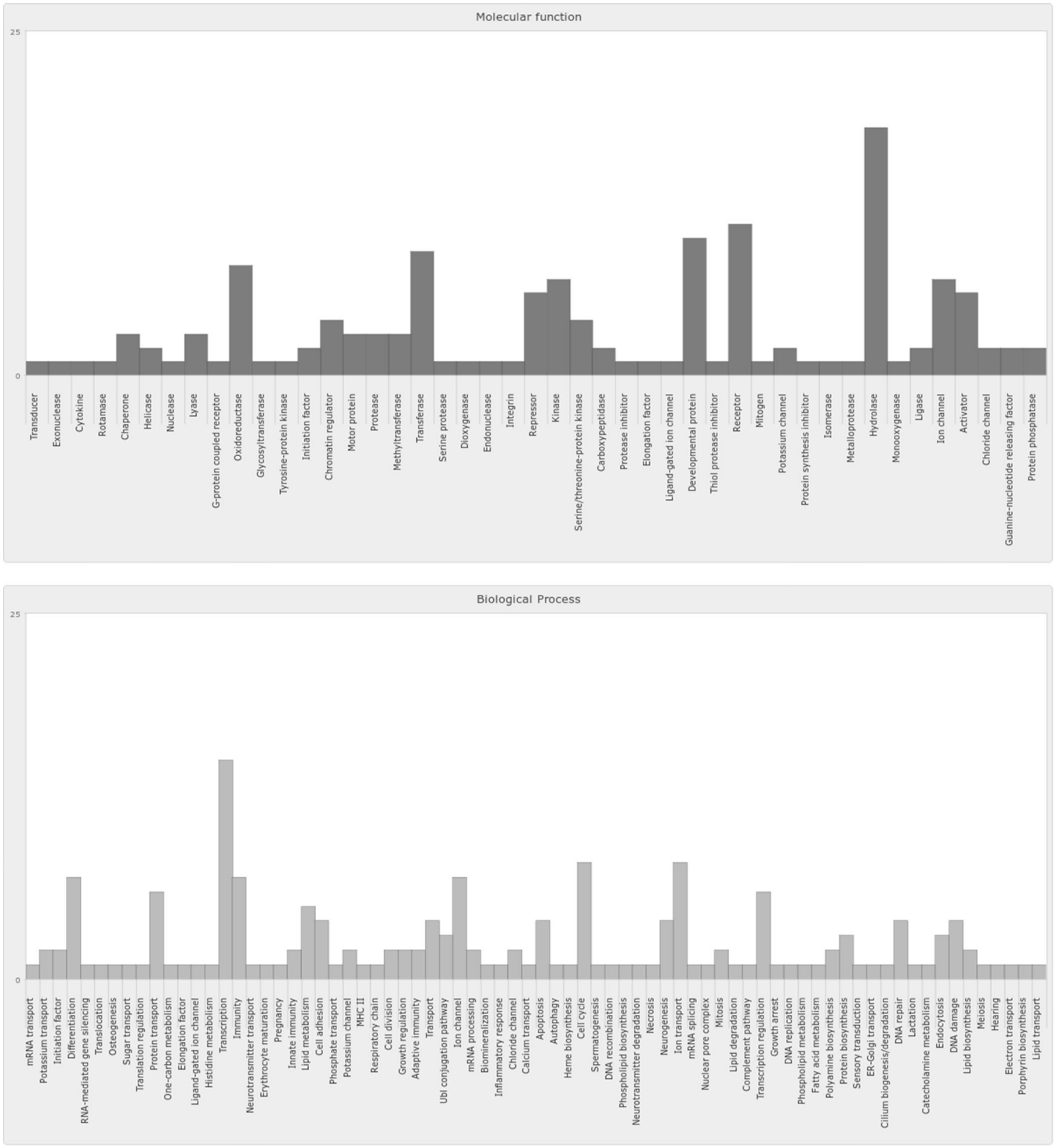

FIGURE 4 Graphical reports of predicted proteins. The predicted proteins can be grouped on the basis of their molecular function and/or their involvement in the biological processes

3 | CONCLUSION

SELYMATRA web application provides the scientific community with a new SELDI-TOF analysis toolset that is able to deliver further insight into complex clustered mass data. Particularly, through ad hoc modeling procedures, it is possible to obtain key features identifying potential outliers or features discerning the similarity between two mass spectra. Based on this result, it has also been possible to predict sets of proteins, simply interfacing the WA to a proteome database. As a consequence, SELYMATRA significantly speeds up the data analysis and the processing of the

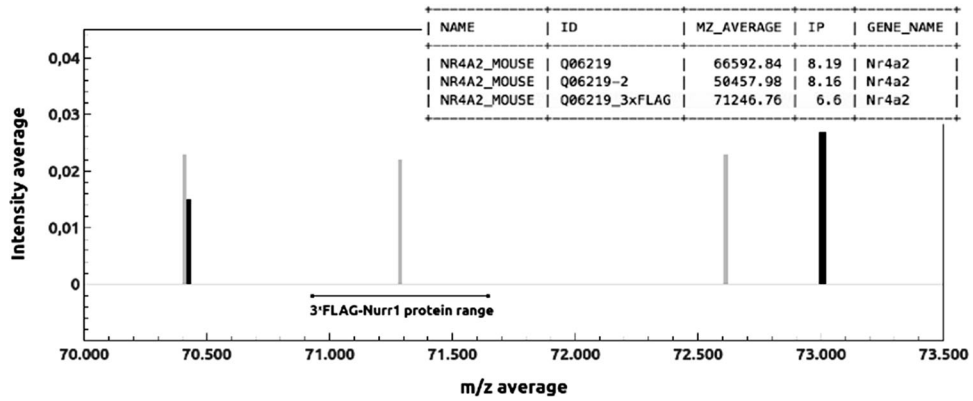

FIGURE 5 SELYMATRA WA validation using the transcription factor 3 × FLAG-Nurr1. Representation of cluster of peaks from the preprocessing analysis of A1 cell line expressing the transcription factor 3 × FLAG-Nurr1 (in gray), compared to a standard A1 cell line (in black). In the insert table, the assignment of the single MS peak at about 72 kD to the 3 × FLAG-Nurr1 by the SELYMATRA WA is reported

experimental results of protein expression in the differentiation of dopaminergic cell line A1 driven by 3 × FLAG-Nurr1 overexpression. In conclusion, this WA looks like a worthy tool of interest for these applications in the biological field.

4 | MATERIALS AND METHODS

4.1 | Cell culture and NURR1 gene overexpression

The dopaminergic cell line A1[24] was infected with a couple of lentiviruses overexpressing the cDNA for the transcription factor NURR1 (LV-NURR1) under the control of an inducible (Tet-On) promoter and the rtTA transactivator (rtTA-LV). This procedure allows controlling the expression of Nurr1 by the addition of doxycycline (Supplementary Figure 1). Cells expressing NURR1 have been plated in MEMF12 10% FBS and infected the following day with LV-NURR1 and rtTA-LV or with rtTA-LV alone, as a control, and growth in presence of doxycycline for 48 h.

4.2 | Protein extraction

After 48 h of induction (approximately 72 h of culture), the cells have been collected in ice-cold phosphate-buffered saline in the presence of protease inhibitors, to avoid further protein modification, like dephosphorylation, proteolysis, and so on. The cells used as a control were grown in the absence of doxycycline in the medium. The cells were lysed according to the Andrews et al. protocol in order to separate the cytoplasm and the nuclear fractions for the analysis to obtain separately the cytoplasm and the nuclear extract.[25] The protein concentration of nuclear and cytoplasmic samples was determined by the Bradford method,[26] using bovine serum albumin as standard.

4.3 | ProteinChip preparation

The ProteinChip arrays were purchased from Bio-Rad (BioRad Laboratories, Hercules, CA, USA). We performed our analysis on the anion exchange (Q10) array with 100 mM Tris–HCl buffer pH 9.0 (low stringency condition) and with 100 mM Tris–HCl buffer pH 7.0 (high stringency condition); and on the cation exchange (CM10) array with 100 mM sodium acetate buffer pH 4 (low stringency condition) and Hepes 50 mM pH 7 (high stringency condition). Samples were loaded in quadruplicate onto preequilibrated spot surfaces of arrays; each sample corresponds to 5 μg of total proteins. After 30 min incubation at room temperature, the unbound proteins were washed out in two steps with the corresponding low and high stringency condition buffer, and a final washing step using HPLC-grade water has been applied once to remove salts. Finally, 1 μL of saturated sinapinic acid (BioRad Laboratories) in 50% acetonitrile (Aldrich, Sigma-Aldrich) and 0.5% trifluoroacetic acid (Fluka, Sigma-Aldrich) was added twice to each spot on the chip and allowed to dry.

4.4 | ProteinChip acquisition

The ProteinChips were analyzed using a linear TOF mass spectrometer SELDI-TOF (BioRad Laboratories, USA), with the following protocols. Protocol A: laser power 3000 nJ, matrix attenuation 2700, focus mass 12,500, sample rate 600 and 50% spot surface fired for ion profiling between the mass/charge (m/z) range of 0–25,000. Protocol B: laser power 5000 nJ, matrix attenuation 2500,

focus mass 30,000, sample rate 600 and 50% spot surface fired for ion profiling between the mass/charge (m/z) range of 25,000–50,000. We have used the ProteinChip Data Manager 3.5 software (BioRad Laboratories), and the spectra were externally calibrated with ProteinChip all-in-one peptide standard (BioRad Laboratories). We have processed the spectra through baseline subtraction and smoothing. The mass has been aligned and normalized using the total ion current within the m/z range of 3–30 kDa for the low mass proteins and 25–50 kDa for the high mass proteins.

4.5 | Data preprocessing

The software ProteinChip Data Manager identifies the peaks in the spectra and groups them into clusters. The comparison between the clusters of peaks highlights, by a scatter plot, the differences in intensity between them. Furthermore, it is able to graphically depict any differences only between the cluster of peaks present in the two samples but is not able to classify the cluster of peaks with different intensity values and to detect the presence/absence of peaks between two samples. So cluster data were exported in data-sheet format for importing in WA.

4.6 | WA principles

SELYMATRA has been developed with the purpose of helping researchers in the study of protein profiling in complex mixtures from cells and tissues. Starting from the comparison among two protein profiles, the SELYMATRA WA detects the presence/absence of peaks between the two samples and assigns the m/z values obtained from MS spectra to the most probable corresponding protein. The protein prediction was carried out using the whole integrated database of protein sequences from the mouse. In this way, the WA makes it possible to easily study the changes in protein expression in altered cell and tissue samples. Moreover, it expects that the file given as input is not just a single mass spectrum but that it is generated as an elaborated mass spectrum (got from two or more mass spectra) obtained by a clustering analysis which may be carried out through the ProteinChip Data Manager (BioRad Laboratories) as well or by other alternative MS software. Of course, we strongly suggest not skipping the clustering analysis step because it is done with the purpose to minimize the bias introduced by the experiment steps. If so, for each given sample (more than one mass spectra) it will generate a data sheet representing a spectrum cluster. Regardless of how the clustered file is generated, it must contain some specific fields, such as the type of chip and conditions used for the sample (Supplementary Table 1). It is important to point out that the real numeric fields taken into account by the software are essentially three: m/z average, m/z STD, and intensity median of the spectral clustering. Therefore, theoretically, the application could work also with the lack of the other fields, but, practically, for now, the file format to give in input to the application has to be structured with all fields listed in Supplementary Table 1.

4.7 | WA Implementation

The SELYMATRA web application takes the m/z values obtained by SELDI or MALDI-TOF from two separate experiments as input and returns both the identified (found or missing) m/z value and numeric scores and the list of predicted corresponding proteins (Supplementary Figure 2). The application has been realized to work with the Microsoft Excel spreadsheet file format such as xls, xlsx, etc. We used the MVC design, in which the communications between the client and the database are mediated by the web application server (Figure 1). The server is implemented using a combination of hypertext markup language pages (HTML), and Java Server Pages (JSP) as a rich component-based user interface. The mySQL software is used for the relational database. Graphical output of predicted proteins based on their function and a biological pathway were also given. The Apache Tomcat, an open-source implementation of the Java Servlet and JavaServer Pages, is used for reporting the results in graphical output.

ACKNOWLEDGMENTS

The research leading to these results has received funding from MIUR-FIRB: Medical Research in Italy (MERIT). Grant no. 2: ref. 0017153, granted to GM. We thank Mr. Fabio Concilio for the informatics support, and Ms. Liliana Parlato for her technical assistance.

Open Access Funding provided by Consiglio Nazionale delle Ricerche within the CRUI-CARE Agreement.

[Correction added on May 16, after first online publication: CRUI-CARE funding statement has been added.]

CONFLICTS OF INTEREST

The authors declare no competing interests in this article.

AUTHORS' CONTRIBUTION

DN carried out the WA development and programming, produced the pictures and helped to draft the manuscript. AC participated in the design of the study, in the WA development and helped to draft the manuscript. MC carried out the MS measurements and the sample analysis. SP carried out the cell grow, NURR1 overexpression and protein

extraction. GCB supervised the cell biology experiments and NURR1 overexpression. LL participated in the design of the study and helped to draft the manuscript. GM participated in the design of the study and helped to draft the manuscript. FF conceived of the study, and participated in its design and coordination and wrote the manuscript. All authors read and approved the final manuscript

ORCID

Davide Nardone https://orcid.org/0000-0003-0486-1791
Angelo Ciaramella https://orcid.org/0000-0001-5592-7995
Salvatore Pulcrano https://orcid.org/0000-0002-5324-8689
Gian C. Bellenchi https://orcid.org/0000-0002-9302-0021
Linda Leone https://orcid.org/0000-0001-7293-1814
Giuseppe Manco https://orcid.org/0000-0001-6191-3952
Ferdinando Febbraio https://orcid.org/0000-0002-3653-9847

REFERENCES

1. Liu H, Hou D, Wu D, Yin H, Wu X. Serum pattern profiling for analyzing different types of stress by protein chip technology: a preliminary study. Eur J Mass Spectrom. 2010;16(5):619–23. https://doi.org/10.1255/ejms.1086
2. Salzano AM, Febbraio F, Farias T, Cetrangolo GP, Nucci R, Scaloni A, Manco G. Redox stress proteins are involved in adaptation response of the hyperthermoacidophilic archaeon *Sulfolobus solfataricus* to nickel challenge. Microb Cell Fact. 2007;6(25):1–11. https://doi.org/10.1186/1475-2859-6-25
3. Liu X, Mo F, Zeng H, Zhu S, Ma X. Quantitative proteomic analysis of cerebrospinal fluid from patients with diffuse large Bâ€'cell lymphoma with central nervous system involvement: a novel approach to diagnosis. Biomed Rep. 2019;11(2):70–78. https://doi.org/10.3892/br.2019.1222
4. Fakhar HBZ, Zali H, Rezaie-Tavirani M, Darkhaneh RF, Babaabas B. Proteome profiling of low grade serous ovarian cancer. J of Ovarian Res. 2019;12(64):1–14. https://doi.org/10.1186/s13048-019-0535-z
5. Malik G, Rojahn E, Ward MD, Gretzer MB, Partin AW, Semmes OJ. Veltri RW. SELDI protein profiling of dunning R-3327 derived cell lines: identification of molecular markers of prostate cancer progression. The Prostate. 2007;67(14):1565–75. https://doi.org/10.1002/pros.20646
6. Anyaiwe DEO, Singh GB, Wilson GD, Geddes TJ. Computational convolution of SELDI data for the diagnosis of Alzheimer's disease. High-Throughput. 2018;7(2):1–14. https://doi.org/10.3390/ht7020014
7. Hsu C-W, Chang K-P, Huang Y, Liu H-P, Hsueh P-C, Gu P-W, Yen W-C, Wu C-C. Proteomic profiling of paired interstitial fluids reveals dysregulated pathways and salivary NID1 as a biomarker of oral cavity squamous cell carcinoma*[S]. Mol Cell Proteomics. 2019;18(10):1939–49. https://doi.org/10.1074/mcp.ra119.001654
8. Wibom C, Protein expression profiling of brain tumor tissue using SELDI-MS. Methods Mol Biol. 2013;1023:191–202. https://doi.org/10.1007/978-1-4614-7209-4_13
9. Chung L, Moore K, Phillips L, Boyle FM, Marsh DJ, Baxter RC. Novel serum protein biomarker panel revealed by mass spectrometry and its prognostic value in breast cancer. Breast Cancer Res. 2014;16(R63):1–12. https://doi.org/10.1186/bcr3676
10. Mu AK-W, Lim B-K, Aminudin N, Hashim OH, Shuib AS. Application of SELDI-TOF inN-glycopeptides profiling of the urine from patients with endometrial, ovarian and cervical cancer. Arch Physiol Biochem. 2016;122(3):111–6. https://doi.org/10.3109/13813455.2016.1151441
11. Lin X-C, Wang X-N, Liu L, Wen Q, Yu R-Q, Jiang J-H. Surface enhanced laser desorption ionization of phospholipids on gold nanoparticles for mass spectrometric immunoassay. Anal Chem. 2016;88(20):9881–4. https://doi.org/10.1021/acs.analchem.6b02733
12. Zambrano G, Ruggiero E, Malafronte A, Chino M, Maglio O, Pavone V, Nastri F, Lombardi A. Artificial heme enzymes for the construction of gold-based biomaterials. Int J Mol Sci. 2018;19(10):2896. https://doi.org/10.3390/ijms19102896
13. Seibert V, Wiesner A, Buschmann T, Meuer J. Surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip® technology in proteomics research. Pathology - Research and Practice. 2004;200(2):83–94. https://doi.org/10.1016/j.prp.2004.01.010
14. Poon TCW. Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. Expert Rev of Proteomics. 2007;4(1):51–65. https://doi.org/10.1586/14789450.4.1.51
15. Cimini D, Corte KD, Finamore R, Andreozzi L, Stellavato A, Pirozzi AVA, Ferrara F, Formisano R, De Rosa M, Chino M, Lista L, Lombardi A, Pavone V, Schiraldi C. Production of human pro-relaxin H2 in the yeast Pichia pastoris. BMC Biotech. 2017;17(1):4. https://doi.org/10.1186/s12896-016-0319-0
16. Jayachandran G, Roth JA, Ji L. In: editors. Analysis of protein-protein interaction using proteinchip array-based SELDI-TOF mass spectrometry. Methods Mol Biol. 2012;818:217–26. https://doi.org/10.1007/978-1-61779-418-6_15
17. Itri F, Monti DM, Chino M, Vinciguerra R, Altucci C, Lombardi A, Piccoli R, Birolo L, Arciello A. Identification of novel direct protein-protein interactions by irradiating living cells with femtosecond UV laser pulses. Biochem and Biophys Res Comm. 2017;492(1):67–73. https://doi.org/10.1016/j.bbrc.2017.08.037
18. Lombardi A, Andreozzi C, Pavone V, Triglione V, Angiolini L, Caccia P. Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations. Electrophoresis. 2013;34(16):2394–406. https://doi.org/10.1002/elps.201300045
19. Huang Y-J, Xuan C, Zhang B-B, Liao M, Deng K-F, He M, Zhao J-M. SELDI-TOF MS profiling of serum for detection of nasopharyngeal carcinoma. J Exp Clin Cancer Res. 2009;28(85):1–7. https://doi.org/10.1186/1756-9966-28-85
20. Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clinic Chem. 2002;48(8):1296–304. https://doi.org/10.1093/clinchem/48.8.1296

21. Poon TCW, Sung JJY, Chow SM, Ng EKW, Yu ACW, Chu ESH, Hui AMY, Leung WK. Diagnosis of gastric cancer by serum proteomic fingerprinting. Gastroenterology. 2006;130(6):1858–64. https://doi.org/10.1053/j.gastro.2006.02.011
22. Magrane M, Consortium UniProt knowledgebase: a hub of integrated protein data. Database. 2011;2011(bar009):1–13. https://doi.org/10.1093/database/bar009
23. Volpicelli F, De Gregorio R, Pulcrano S, Perrone-Capano C, di Porzio U, Bellenchi GC. Correction: direct regulation of Pitx3 expression by nurr1 in culture and in developing mouse midbrain. PLoS One. 2020;15(5):e0233918. https://doi.org/10.1371/journal.pone.0233918
24. Colucci-D'Amato GL, Tino A, Pernas-Alonso R, Ffrench-Mullen JMH, Di Porzio U. Neuronal and glial properties coexist in a novel mouse CNS immortalized cell line. Exp Cell Res. 1999;252(2):383–91. https://doi.org/10.1006/excr.1999.4636
25. Andrews NC, Faller DV. A rapid micrqpreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. 1991;19(9):2499. https://doi.org/10.1093/nar/19.9.2499
26. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72(1-2):248–54. https://doi.org/10.1016/0003-2697(76)90527-3

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Nardone D, Ciaramella A, Cerreta M, Pulcrano S, Bellenchi GC, Leone L, Manco G, Febbraio F. Selymatra: A web application for protein-profiling analysis of mass spectra. *Biotechnology and Applied Biochemistry*. 2022;**69:**1821–1829. https://doi.org/10.1002/bab.2249