Essential Features and Use Cases of the Cerebrospinal Fluid Proteome Resource (CSF-PR)

Astrid Guldbrandsen[1,2], *Yehia Farag*[1,2], *Ragnhild Reehorst Lereim*[1,2], *Frode S. Berven*[1,*], *Harald Barsnes*[1,2]

[1]Proteomics Unit, Department of Biomedicine, University of Bergen, Bergen, Norway.

[2]Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway.

*Corresponding author, e-mail: frode.berven@biomed.uib.no

i. Running head: Essential Features and Use Cases of CSF-PR

ii. Summary/Abstract:

Every year, a large number of published studies present biomarkers for various neurological disorders. Many of these studies are based on mass spectrometry proteomics data and describe comparison of the abundance of proteins in cerebrospinal fluid between two or more disease groups. As the number of such studies is growing, it is no longer straightforward to obtain an overview of which specific proteins are increased or decreased between the numerous relevant diseases and their many subcategories, or to see the larger picture or trends between related diseases. To alleviate this situation we therefore mined the literature for mass spectrometry-based proteomics studies including quantitative protein data from cerebrospinal fluid of patients with multiple sclerosis, Alzheimer's disease and Parkinson's disease and organized the extracted data in the Cerebrospinal Fluid Proteome Resource (CSF-PR). CSF-PR is freely available online at http://probe.uib.no/csf-pr, is highly interactive and allows for

easy navigation, visualization and export of the published scientific data. This chapter will guide the user through some of the most important features of the tool and show examples of the suggested use cases.

iii. Key Words: mass spectrometry, cerebrospinal fluid, neurodegenerative/neurological disorders/diseases, biomarkers, data mining, data visualization, multiple sclerosis, Parkinson's disease, Alzheimer's disease

1. Introduction

Cerebrospinal fluid (CSF) surrounds the brain and spinal cord, and is believed to be a promising source for biomarkers for diseases affecting the central nervous system, as the detected proteins may reflect ongoing neurological processes [1-4]. Mass spectrometry-based proteomics is a method of choice for such studies and can measure thousands of proteins and peptides from a single sample.

During the last 10-15 years numerous papers describing mass spectrometry results from CSF proteome mapping and biomarker studies have been published [5-9]. These types of studies often result in large data matrices, but usually only a small portion of the results are highlighted and discussed in the paper. Furthermore, while some authors provide all the underlying raw data (both identification and quantification), this is unfortunately not the default.

We therefore saw the need to extract all protein and peptide related data across the relevant mass spectrometry studies and organize it in a common format for easy navigation, and make this information freely available to the scientific community. In 2014, we created the first

version of the CSF Proteome Resource (CSF-PR), mainly as a repository of the mass spectrometry data associated with our comprehensive CSF proteome mapping study [10]. The initial goal was to organize our large-scale CSF data in a structured manner, but also to share protein and peptide lists (with additional details) with other researchers.

Next, we switched the focus to quantitative CSF data, resulting in an expanded and redesigned version of CSF-PR, to also include quantitative results, both from our own group and from other research groups. To achieve this, we mined the available literature for studies presenting protein abundance comparisons in CSF from patients with multiple sclerosis (MS), Alzheimer's disease (AD) and Parkinson's disease (PD) , focusing on studies using mass spectrometry, and organized the extracted data in a revised version of CSF-PR, referred to as CSF-PR 2.0 [9].

When choosing the data to include, we applied specific quality control filters for the literature searches related to technology, methodological set-up and sample size. Since the initial release of CSF-PR 2.0, an additional 12 datasets from 8 recent publications have been added [11-18], resulting in CSF-PR now containing a total of 97 datasets from 25 publications. New relevant studies fulfilling the criteria are continuously being added.

Several useful features for browsing, sorting, filtering and investigating the included data are available. This makes CSF-PR ideal for *e.g.* finding which proteins that are identified and quantified in CSF, investigating relevant literature for potential biomarker proteins, comparing protein abundances between diseases or selecting promising proteotypic peptides to use in targeted analyses [19-22].

2. Methods

A description of CSF-PR features and a step-by-step guide through the resource is here presented, starting from the welcome page. The guide is divided into several sections for easy navigation and mainly focuses on the quantitative data.

2.1 The CSF-PR welcome page

When entering the CSF-PR web page, an overview of the resource status and sections is displayed (**Figure 1**). The user can choose to navigate to either (*i*) the manually curated quantitative data extracted from the publications [9], (*ii*) the identification data from the above mentioned CSF mapping paper [10], (*iii*) make a search across all resource data, or (*iv*) compare their own protein data to the data in the resource.

2.2 Quantification data

The quantification data tab takes you to an overview of the disease categories for which there is quantitative data in CSF-PR (**Figure 2**), and the user can choose an individual disease category or inspect all diseases at the same time. After selection, an overview of the selected data is presented as a heat map (**Figure 3**), displaying the different disease sub-categories for which data are available, plus the number of available datasets for each comparison.

2.2.1 Filtering, customizing and investigating the quantitative data

Several options for filtering, customizing and investigating the CSF-PR data are available. The user can either select specific comparisons from the heat map, or click *Select all disease group comparisons* in the bottom right corner to selects all datasets. Selection of cells will result in yellow edges around the cell. (*See* **Note 1 and 2)**

By navigating to the top left corner in the heat map, the *Dataset filters* window can be opened (**Figure 4**). This view provides an overview of the distribution of various data properties, related to publication year, study type, technology, analytical approach *etc*., and can be used to filter for specific data properties. Upon each selected filter, the pie charts are updated to display the current distribution, providing the user full control over the data being displayed.

Given that different papers often use highly specific disease subgroups and because the categorization of the various disease subgroups is a subject of debate, we decided to use the same sub-group names in the heat map as is provided in the paper. However, we added the possibility to rename and combine existing groups. *e.g*. to reduce the number of disease sub-group comparisons in the heat map and plots (**Figure 5A**, *See* **Note 3**). There is also a possibility to reorder the disease subgroups or exclude certain groups, to better organize and optimize the views and plots in the resource (**Figure 5B).** These features are available through the *Recombine disease groups* and *Reorder and select disease groups* icons in the top left corner of the heat map (*See* **Figure 3**).

Whenever a selection is made from the heat map, new tabs/levels become visible to the right. The second tab includes a bubble plot providing an overview of the direction of regulation between the disease sub-groups for the currently selected proteins (**Figure 6**), with the bubble size representing the number of proteins changed in the specific direction between the disease sub-groups. One, several or all bubbles can be selected. By default, all are selected.

The next tab provides a table with all the comparison- and protein-specific data. Plots showing the direction of regulation between all the disease group comparisons are shown for each protein. The table can be sorted on protein accession number or name, or on a specific value (direction of regulation) in a specific comparison (**Figure 7**) (*See* **Note 4** and **5**).

To view more details for a specific protein, the protein can be selected from the table, upon which a new *Protein Details* tab will become available below the *Protein Table* in the side panel. This tab presents all the results related to the selected protein from the relevant publications under the current filtration. The graph at the top summarizes the protein comparison results from each dataset in a single triangle as either *Increased*, *Decreased* or *Equal* (**Figure 8**) (*See* **Note 6, 7, 8** and **9**). Below the graph is a table containing the available peptide details from the datasets illustrated in the graph. This information can be useful in the selection of proteotypic peptides as is described in more details below.

2.3 Searching for proteins or peptides

The search tab in the home page allows the user to search for specific proteins (or peptides) across all or selected datasets in CSF-PR. This is useful to figure out if certain proteins are identified and/or quantified in CSF (*See* **Note 10**). It can also be used to inspect protein regulation across disease groups and categories. Multiple search keys are supported.

2.3.1 Searching for a specific protein/peptide of interest - Basic

1. Navigate to the Search tab in the CSF-PR welcome page.
2. Type either protein name(s), accession number(s) or peptide sequence(s) and click the appropriate search term as well as the disease (if restricting the search to a specific disease).
3. Load the results, which will then be visible in the same framework as described above.
4. Observe the protein abundance status across various disease group comparisons in the plot. Expand to view a labelled x-axis. This plot shows the average results for all datasets quantifying this protein between the two specific disease groups.

5. Sort or filter the plot, drop comparisons from the view, flip disease groups A and B, or export to Excel using the icons in the bottom right corner. Click on any data point to view specific dataset details, as is described above. Click the accession number to view related UniProt (https://www.uniprot.org/) information about the specific protein.
6. To view more detailed and dataset specific protein results, navigate to the *Protein Details* tab. Here, it becomes clear how many datasets that support the summarized results showed in the *Protein Table.*

2.3.2 Evaluating/selecting biomarker candidates

As part of a biomarker discovery project, it is necessary to evaluate the list of candidate proteins, in terms of their potential as markers of one or several categories (diagnostic, prognostic etc.). For this purpose, CSF-PR is a very useful tool, as it can easily give the user an overview of what other researchers have already found regarding your potential markers, in relation to MS, AD and PD. This can tell you something about each marker's potential sensitivity and selectivity, and could also provide some indications as to which processes each marker can monitor.

1. Navigate to the search section and insert the list of biomarker candidates to evaluate. Observe which one are found in the database. The proteins not found in CSF-PR might be difficult to identify or quantify using standard proteomics methods.
2. Load the search results and inspect the candidate proteins abundance across the various disease group comparisons.
3. By sorting the table in various ways, and perhaps merging and re-ordering the patient groups, interesting trends across disease categories and sub-groups can be revealed. Some questions that might be relevant in this setting: (*i*) Which studies have found the protein and what methods did they use? (*ii*) Is the protein only differentially expressed

in your disease of interest? (*iii*) Does the protein seem to be affected by treatment? (*iv*) If the protein is similarly affected across different diseases, which processes are common for these diseases? (*v*) What is the planned control group for your follow-up/verification experiments? Is it suitable?

4. By navigating to the heat map tab and clicking the dataset filters icon, you can observe the various features of the data, related to approaches (discovery/targeted) and specific quantification methods used, e.g. selected/parallel reaction monitoring (SRM/PRM), tandem mass tag (TMT) or label-free. If you, for example, are interested in using SRM or PRM for your follow up experiments, it may be a good idea to apply the filter for SRM/PRM or targeted and mass spectrometry, as you will then only see proteins that have previously been quantified by SRM/PRM.
5. Based on the information obtained from CSF-PR one can hopefully arrive at a better selection of candidate markers to procced with in further experiments.

2.3.3 Investigating peptide information for targeted assays (SRM/PRM)

CSF-PR is a great tool to use if you are in the process of verifying potential biomarker candidates. Biomarker discovery experiments often result in a number of potential candidates, as outlined above. The selection of which peptides to use as surrogates for your proteins of interest is not always straightforward. Some guideline papers have been published on the selection of such peptides [23,24], but it can also be a good idea to observe which peptides have worked well for other researchers. To investigate peptide details for a protein of interest, follow these steps:

1. Navigate to the *Search* panel and search for your protein of interest. You will see your results appearing if there are any hits for this protein in the resource. Hits can be either

in the identification data [10] or in the quantification data [9]. You can choose either to load only data for a specific disease category (MS, AD or PD) or to load all data.

2. To get all available information on the protein, load all data. Your protein will appear alone in the protein table (*See* **Note 11**). Click in the table to open the next level tab, which is the *Protein Details*.
3. In the *Protein Detail*s tab, all relevant peptide data for your protein of interest can be found in the table below the overview plot (*See* **Figure 8**).
4. To view detailed dataset and protein specific information, click any of the triangles or the comparison in the table.
5. To see the publication from which the specific dataset has been extracted, click the publication link, taking you to the publication in PubMed.
6. The *Protein Coverage* section in the peptide overview gives information about which specific peptides have been used in each dataset to arrive at the final quantitative conclusion for the protein (*See* **Figure 8**). The coverage across the full protein sequence is here shown, and when available, also the quantitative results for each peptide, where abundance (increase, decrease, equal) and significance status is illustrated by colour coding. (*See* **Note 12)**
7. Hover over the peptide to see the sequence plus the start and end indexes.
8. Click a peptide to view detailed peptide specific data (e.g. sequence, fold change, *p*-value, comments) in a pop-up window.

You will find that not all peptides are identified for most proteins, or even has the same abundance or significance status, and there may be many reasons for this such as post-translational modifications. This information is crucial in the selection of appropriate peptides to represent a protein, and can save much time in peptide suitability testing.

2.4 Compare to your own data

The compare section of CSF-PR makes it possible to insert protein results from your own experiments as the search key, and reveal how your protein data compares to what is previously known in the field.

2.4.1 Comparison of your own data to the data in CSF-PR

1. Navigate to the *Compare* tab in the CSF-PR welcome page.
2. Select or enter the disease groups that best describe your specific comparison.
3. Insert the list of protein accession numbers into the appropriate fields, based on their observed increased, decreased on equal abundance between the two disease groups used in your study.
4. Click compare. The *Results Overview* will appear giving an overview of which proteins are found in the search, in which disease categories the proteins are found, plus the total number of hits. Make selections from the pie chart if desired, or load all data.
5. After loading the data, the protein data will be available in the same framework (heat map, plots and tables) as described above, but the inserted user data is now integrated with the resource data (**Figure 9**).
6. The data can now be inspected, sorted and filtered as described above, with the user data always visible and directly comparable with the resource data.

4. Notes

Note 1: By default, only the CSF datasets are visible in the heat map, and pooled samples are included. Some serum datasets have also been included to the resource, if they represented verification of initial CSF studies. If the user wants to also display the serum datasets, this must be specifically selected by clicking the red drop symbol in the upper left corner of the heat map. In this area is also the possibility to remove datasets from studies that have employed pooling of samples.

Note 2: In the lower right corner, other tabs are also available, such as the possibility to allow for multiple selections from the heat map, to clear filters and to export heat map data. It is also possible to view all the datasets and publications from which the datasets are collected. This overview will appear in a pop-up window, and clicking each dataset shows meta data for the dataset, related to e.g. experimental details, technology, protein and peptide numbers and patient features. A link to the original papers is also provided in this window.

Note 3: The recombination of disease groups feature can be useful for advanced users, but should be used with caution. We recommend looking into the original publication for patient details.

Note 4: Detailed sorting and filtering options for the protein table is available by clicking the *Sort or filter comparison* icon in the bottom right corner. Here, specific comparisons can also be removed from the table. The possibility to switch/flip disease groups is also available through the *Switch disease groups* icon, as is the opportunity to *Clear all applied filters*.

Note 5: Data can be exported to Excel at this level and all other levels in CSF-PR. This is available through the Excel icon in the lower right corner.

Note 6: By default, all individual datasets are represented by individual data points, but the user can choose to see the summarized trend for all datasets combined by clicking the *Show/hide individual datasets* icon on the right. By selecting this view, only one averaged data point will be shown for each comparison.

Note 7: Hover over each data point to see the details related to the disease groups compared, the number of patients and the *p*-value (if available), or click on a data point to get a pop-up with even more details related to the protein and the dataset.

Note 8: The size of the triangle represents the number of patients in the specific study (the larger the triangle, the more patients included in the study). Furthermore, the colour and placement indicate the increased, decreased or equal abundance of the protein in the specific A vs. B comparison.

Note 9: The graph can be ordered by trend by clicking the *Order dataset by trend* icon in the bottom right corner.

Note 10: Please keep in mind that CSF-PR does not represent a complete overview of all proteins and peptides found in CSF, as not all types of publications are included. Some were excluded based on our specific criteria related to *e.g.* methods, patient number, disease groups or study type [9].

Note 11: To be able to see the disease group comparison, either hover over the specific data point in the plot or click the icon in the upper right corner of the plot window to get a labelled x-axis.

Note 12: To see only the significantly changed peptides click the *Show/hide not significant and stable peptides* icon at the bottom right.

5. Acknowledgements

A.G. and H.B. is supported by the Research Council of Norway. H.B., Y.F. and R.R.L. is supported by Bergen Research Foundation. FB is supported by the Western Norway Regional Health Authority.

6. References

1. Reiber H (2001) Dynamics of brain-derived proteins in cerebrospinal fluid. Clin Chim Acta 310 (2):173-186
2. Shevchenko G, Konzer A, Musunuri S, Bergquist J (2015) Neuroproteomics tools in clinical practice. Biochim Biophys Acta 1854 (7):705-717. doi:10.1016/j.bbapap.2015.01.016
3. Johanson CE, Duncan JA, 3rd, Klinge PM, Brinker T, Stopa EG, Silverberg GD (2008) Multiplicity of cerebrospinal fluid functions: New challenges in health and disease. Cerebrospinal Fluid Res 5:10. doi:10.1186/1743-8454-5-10
4. Carlyle BC, Trombetta BA, Arnold SE (2018) Proteomic Approaches for the Discovery of Biofluid Biomarkers of Neurodegenerative Dementias. Proteomes 6 (3). doi:10.3390/proteomes6030032
5. Kroksveen AC, Opsahl JA, Guldbrandsen A, Myhr KM, Oveland E, Torkildsen O, Berven FS (2015) Cerebrospinal fluid proteomics in multiple sclerosis. Biochim Biophys Acta 1854 (7):746-756. doi:10.1016/j.bbapap.2014.12.013
6. Kroksveen AC, Opsahl JA, Aye TT, Ulvik RJ, Berven FS (2011) Proteomics of human cerebrospinal fluid: discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. J Proteomics 74 (4):371-388. doi:10.1016/j.jprot.2010.11.010
7. Portelius E, Brinkmalm G, Pannee J, Zetterberg H, Blennow K, Dahlen R, Brinkmalm A, Gobom J (2017) Proteomic studies of cerebrospinal fluid biomarkers of Alzheimer's disease: an update. Expert Rev Proteomics 14 (11):1007-1020. doi:10.1080/14789450.2017.1384697
8. Roche S, Gabelle A, Lehmann S (2008) Clinical proteomics of the cerebrospinal fluid: Towards the discovery of new biomarkers. Proteomics Clin Appl 2 (3):428-436. doi:10.1002/prca.200780040
9. Guldbrandsen A, Farag Y, Kroksveen AC, Oveland E, Lereim RR, Opsahl JA, Myhr KM, Berven FS, Barsnes H (2017) CSF-PR 2.0: An Interactive Literature Guide to Quantitative Cerebrospinal Fluid Mass Spectrometry Data from Neurodegenerative Disorders. Mol Cell Proteomics 16 (2):300-309. doi:10.1074/mcp.O116.064477
10. Guldbrandsen A, Vethe H, Farag Y, Oveland E, Garberg H, Berle M, Myhr KM, Opsahl JA, Barsnes H, Berven FS (2014) In-depth characterization of the cerebrospinal fluid (CSF) proteome displayed through the CSF proteome resource (CSF-PR). Mol Cell Proteomics 13 (11):3152-3163. doi:10.1074/mcp.M114.038554
11. Begcevic I, Brinc D, Brown M, Martinez-Morillo E, Goldhardt O, Grimmer T, Magdolen V, Batruch I, Diamandis EP (2018) Brain-related proteins as potential CSF biomarkers of Alzheimer's disease: A targeted mass spectrometry approach. J Proteomics 182:12-20. doi:10.1016/j.jprot.2018.04.027
12. Begcevic I, Brinc D, Dukic L, Simundic AM, Zavoreo I, Basic Kes V, Martinez-Morillo E, Batruch I, Drabovich AP, Diamandis EP (2018) Targeted Mass Spectrometry-Based Assays for Relative Quantification of 30 Brain-Related Proteins and Their Clinical Applications. J Proteome Res 17 (7):2282-2292. doi:10.1021/acs.jproteome.7b00768
13. Pavelek Z, Vysata O, Tambor V, Pimkova K, Vu DL, Kuca K, Stourac P, Valis M (2016) Proteomic analysis of cerebrospinal fluid for relapsing-remitting multiple sclerosis and clinically isolated syndrome. Biomed Rep 5 (1):35-40. doi:10.3892/br.2016.668
14. Brinkmalm G, Sjodin S, Simonsen AH, Hasselbalch SG, Zetterberg H, Brinkmalm A, Blennow K (2018) A Parallel Reaction Monitoring Mass Spectrometric Method for Analysis of Potential CSF Biomarkers for Alzheimer's Disease. Proteomics Clin Appl 12 (1). doi:10.1002/prca.201700131
15. Kroksveen AC, Guldbrandsen A, Vaudel M, Lereim RR, Barsnes H, Myhr KM, Torkildsen O, Berven FS (2017) In-Depth Cerebrospinal Fluid Quantitative Proteome and Deglycoproteome Analysis: Presenting a Comprehensive Picture of Pathways and Processes Affected by Multiple Sclerosis. J Proteome Res 16 (1):179-194. doi:10.1021/acs.jproteome.6b00659
16. Paterson RW, Heywood WE, Heslegrave AJ, Magdalinou NK, Andreasson U, Sirka E, Bliss E, Slattery CF, Toombs J, Svensson J, Johansson P, Fox NC, Zetterberg H, Mills K, Schott JM (2016) A targeted proteomic multiplex CSF assay identifies increased malate dehydrogenase and other

neurodegenerative biomarkers in individuals with Alzheimer's disease pathology. Transl Psychiatry 6 (11):e952. doi:10.1038/tp.2016.194
17. Stoop MP, Runia TF, Stingl C, van der Vuurst de Vries RM, Luider TM, Hintzen RQ (2017) Decreased Neuro-Axonal Proteins in CSF at First Attack of Suspected Multiple Sclerosis. Proteomics Clin Appl 11 (11-12). doi:10.1002/prca.201700005
18. Timirci-Kahraman O, Karaaslan Z, Tuzun E, Kurtuncu M, Baykal AT, Gunduz T, Tuzuner MB, Akgun E, Gurel B, Eraksoy M, Kucukali CI (2018) Identification of candidate biomarkers in converting and non-converting clinically isolated syndrome by proteomics analysis of cerebrospinal fluid. Acta Neurol Belg. doi:10.1007/s13760-018-0954-4
19. Macron C, Lane L, Nunez Galindo A, Dayon L (2018) Deep Dive on the Proteome of Human Cerebrospinal Fluid: A Valuable Data Resource for Biomarker Discovery and Missing Protein Identification. J Proteome Res. doi:10.1021/acs.jproteome.8b00300
20. Barkovits K, Linden A, Galozzi S, Schilde L, Pacharra S, Mollenhauer B, Stoepel N, Steinbach S, May C, Uszkoreit J, Eisenacher M, Marcus K (2018) Characterization of Cerebrospinal Fluid via Data-Independent Acquisition Mass Spectrometry. J Proteome Res 17 (10):3418-3430. doi:10.1021/acs.jproteome.8b00308
21. Sathe G, Na CH, Renuse S, Madugundu A, Albert M, Moghekar A, Pandey A (2018) Phosphotyrosine profiling of human cerebrospinal fluid. Clin Proteomics 15:29. doi:10.1186/s12014-018-9205-1
22. Martin NA, Nawrocki A, Molnar V, Elkjaer ML, Thygesen EK, Palkovits M, Acs P, Sejbaek T, Nielsen HH, Hegedus Z, Sellebjerg F, Molnar T, Barbosa EGV, Alcaraz N, Gallyas F, Jr., Svenningsen AF, Baumbach J, Lassmann H, Larsen MR, Illes Z (2018) Orthologous proteins of experimental de- and remyelination are differentially regulated in the CSF proteome of multiple sclerosis subtypes. PLoS One 13 (8):e0202530. doi:10.1371/journal.pone.0202530
23. Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 4:222. doi:10.1038/msb.2008.61
24. Bollinger JG, Stergachis AB, Johnson RS, Egertson JD, MacCoss MJ (2016) Selecting Optimal Peptides for Targeted Proteomic Experiments in Human Plasma Using In Vitro Synthesized Proteins as Analytical Standards. Methods Mol Biol 1410:207-221. doi:10.1007/978-1-4939-3524-6_12

Figures

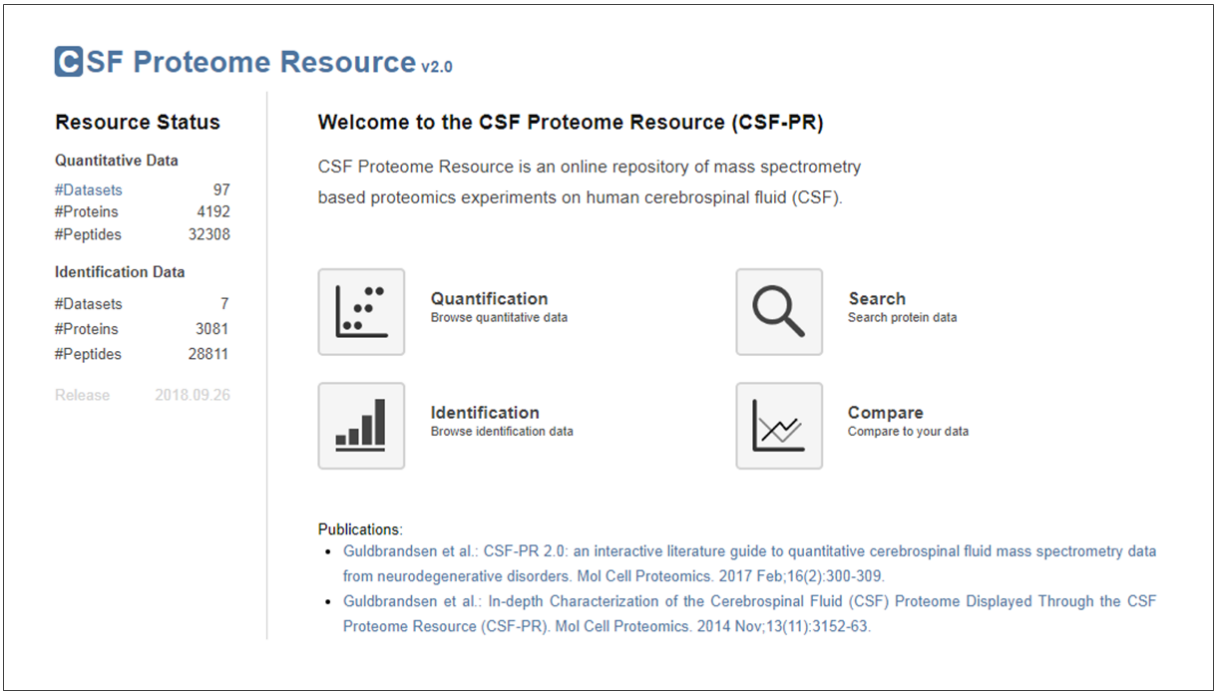


Figure 1: Overview of the CSF-PR welcome page, where the different sections of the webpage can be accessed. A status of the data in the resource is shown in the upper left corner, along with release notes.

Figure 2: Overview of the available disease categories in CSF-PR. The user can click to view a selected disease category or to load data for all disease categories. Additionally, the number of datasets in each disease category is indicated as a fraction of the total number of datasets in the repository (n=97).

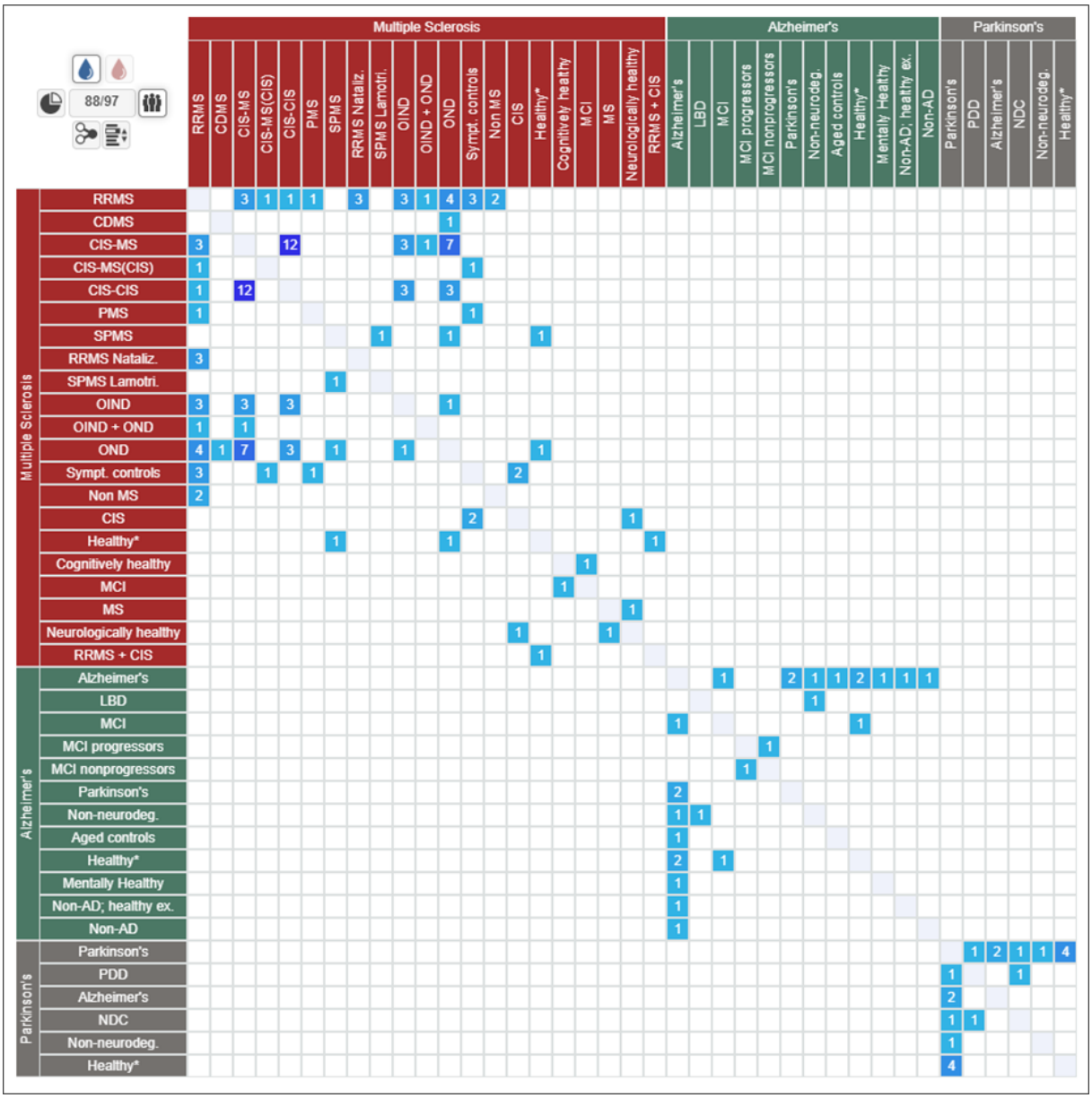


Figure 3: Heat map displaying some of the available disease categories and disease sub-group comparisons. The number inside each cell shows the number of datasets available for this specific comparison, the darker the colour, the more datasets are available.

Figure 4: The Dataset Explorer allows the user to explore properties of the currently selected data, and filter for specific properties. Numbers in the charts indicate the number of data sets the given filter will apply to. Multiple selections are supported and the charts adapts as filters are applied.

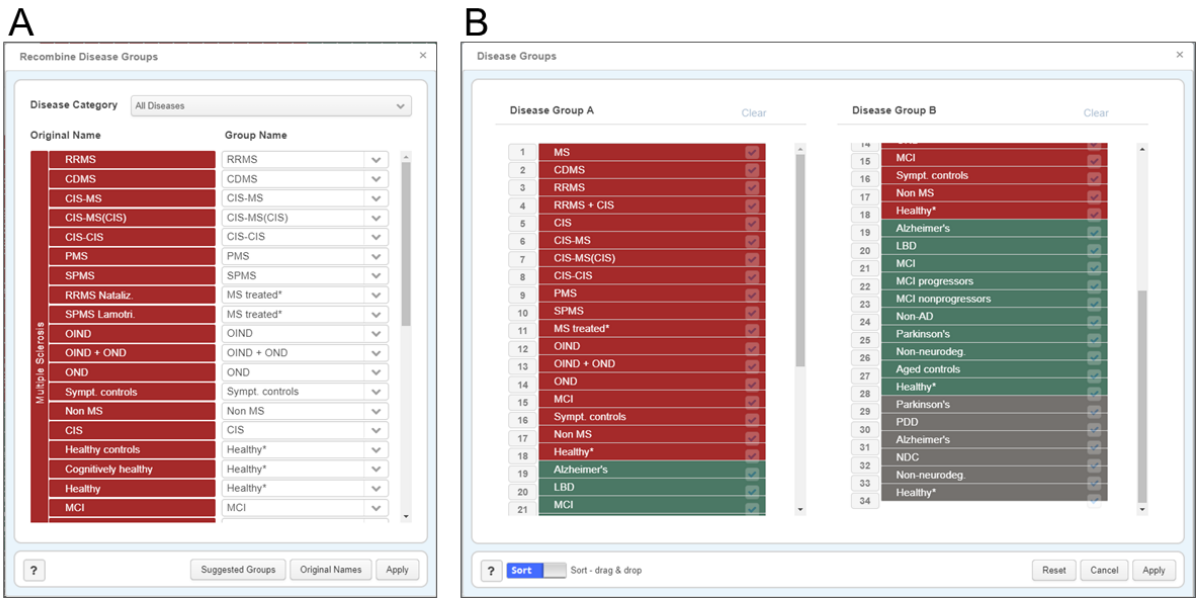


Figure 5: Pop-up dialogs where the user can (A) perform recombination or renaming of disease groups, or (B) select and sort the disease groups.

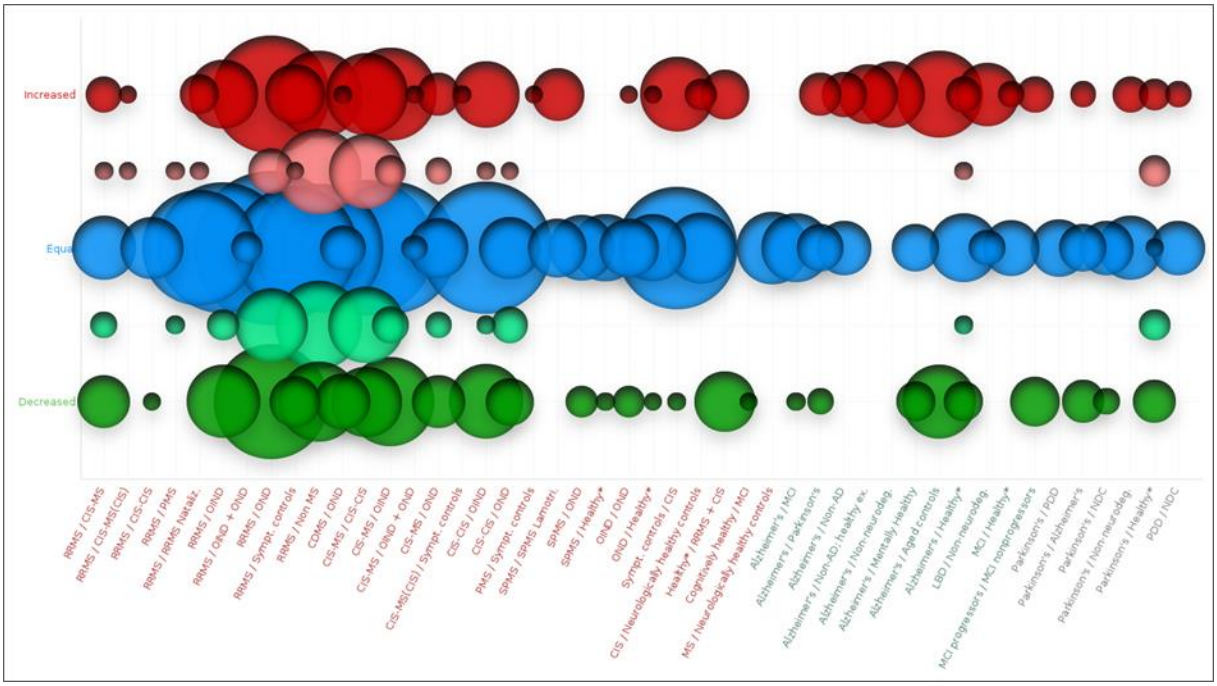


Figure 6: Bubble plot giving an overview of the number of proteins found increased, decreased or equal between the various disease subgroups. Selection can be made by clicking the bubbles.

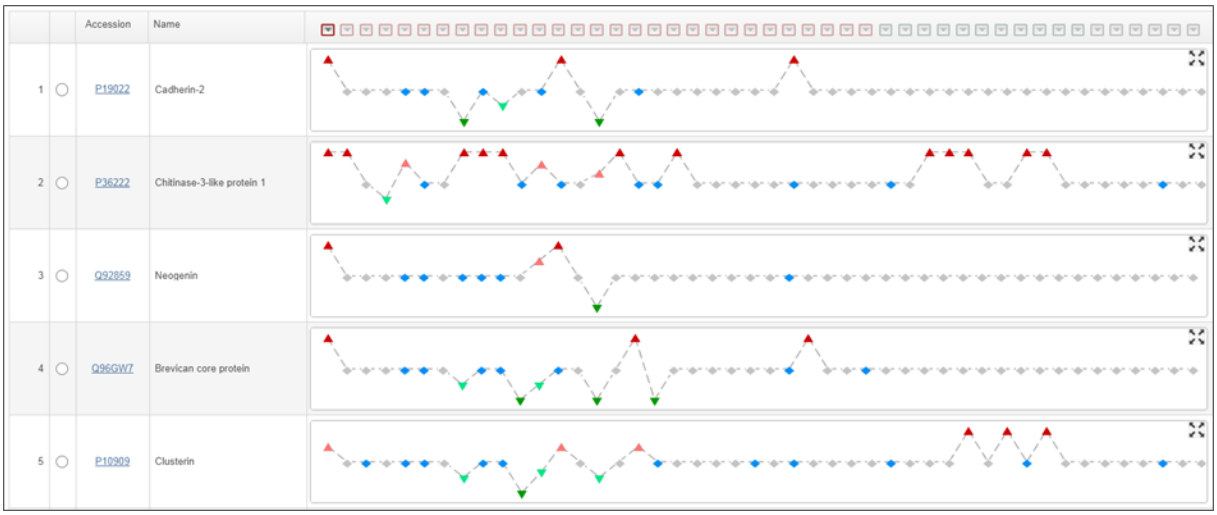


Figure 7: The protein table lists all the proteins under the current selection, and shows a plot illustrating how the proteins change across the various disease group comparisons.

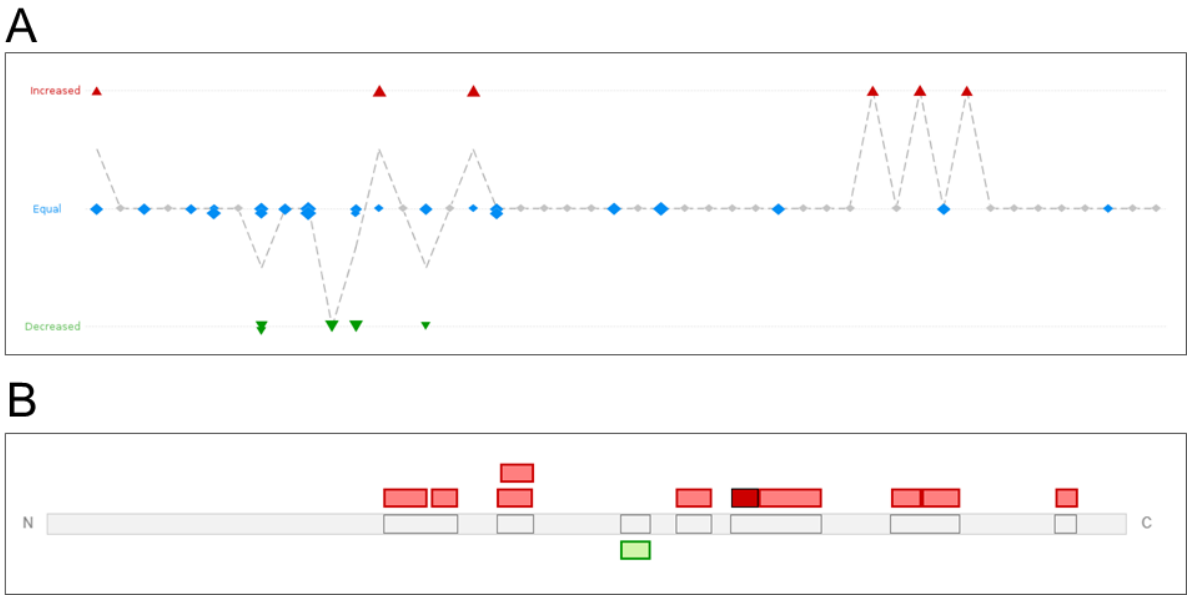
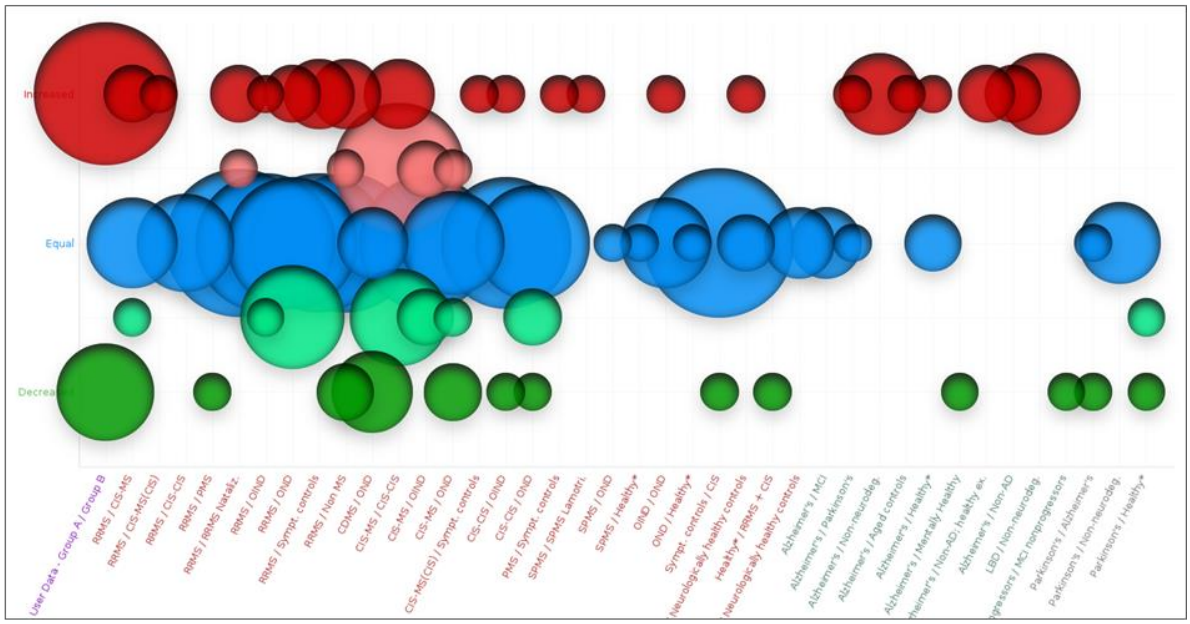
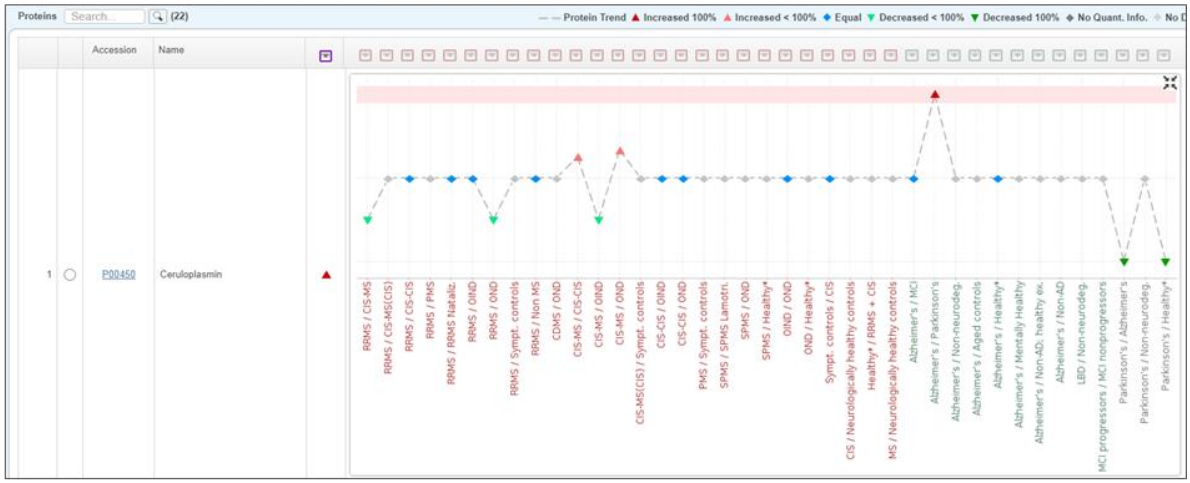


Figure 8: **A**: Overview of the datasets giving evidence for increased (red triangle), decreased (green triangle) or equal (blue square) abundance of the selected protein across the various comparisons (x-axis, not labelled in this view). Each data point represents a specific protein fold change (disease groups A vs. B) in a specific dataset. The specific comparison can be found by hovering over the symbol. **B**: Protein sequence coverage overview showing the peptides used for quantification in the specific selected dataset. Red indicates increased abundance and green indicated decreased abundance. Dark red and dark green indicates statistically significant changes in disease group A vs. B.

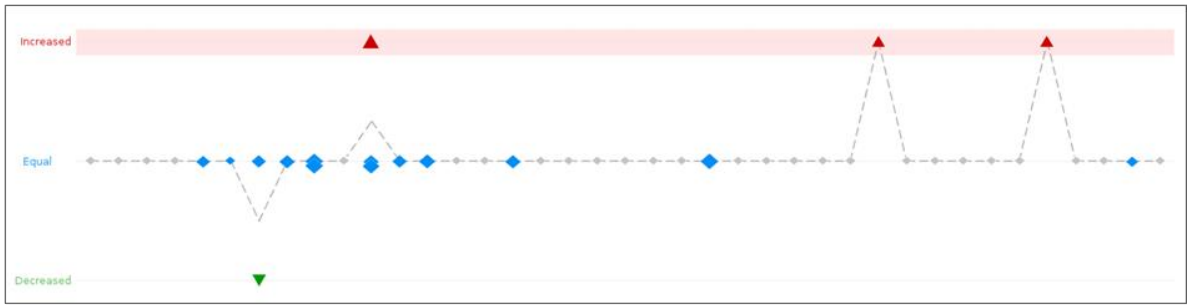
A


Increased
Equal
Decreased
User Data - Group A / Group B
RRMS / CIS-MS
RRMS / CIS-MS(CIS)
RRMS / CIS-CIS
RRMS / PMS
RRMS / RRMS Natali.
RRMS / OIND
RRMS / OND
RRMS / Sympt. controls
RRMS / Non MS
CDMS / OND
CIS-MS / CIS-CIS
CIS-MS / OIND
CIS-MS / OND
CIS-MS(CIS) / Sympt. controls
CIS-CIS / OIND
CIS-CIS / OND
PMS / Sympt. controls
SPMS / SPMS Lamotri.
SPMS / OND
SPMS / Healthy*
OIND / OND
OND / Healthy*
Sympt. controls / CIS
Neurologically healthy controls / CIS
Healthy* / RRMS + CIS
Neurologically healthy controls
Alzheimer's / MCI
Alzheimer's / Parkinson's
Alzheimer's / Non-neurodeg.
Alzheimer's / Aged controls
Alzheimer's / Healthy*
Alzheimer's / Mentally Healthy
Alzheimer's / Non-AD; healthy ex.
Alzheimer's / Non-AD
LBD / Non-neurodeg.
MCI progressors / MCI nonprogressors
Parkinson's / Alzheimer's
Parkinson's / Non-neurodeg.
Parkinson's / Healthy*

B


Proteins
Search...
(22)
Protein Trend
Increased 100%
Increased < 100%
Equal
Decreased < 100%
Decreased 100%
No Quant. Info.
Accession
Name
1
P00450
Ceruloplasmin

C


Increased
Equal
Decreased

D

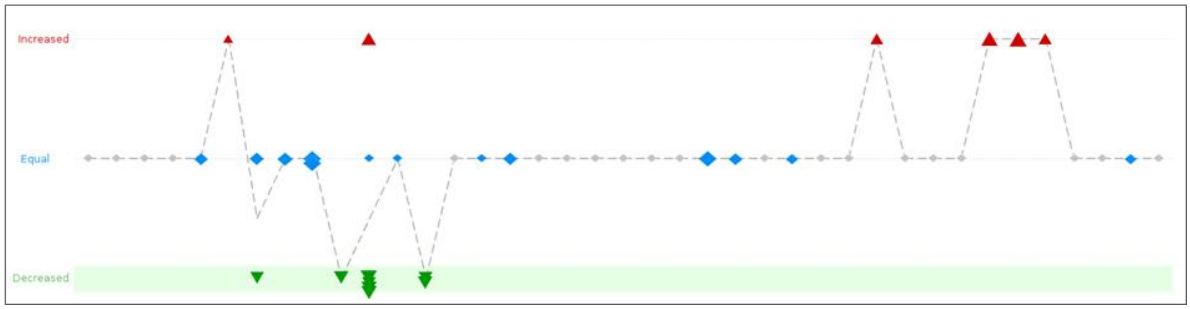

Increased
Equal
Decreased

Figure 9: **A**: The bubble plot with integrated user data (far left), **B**: The protein table (expanded view to show the comparisons) with the user data results (increased) integrated and shown as a light red horizontal line. **C**: The protein details plot shown with the user data results (increased) integrated and shown as a light red horizontal line. **D**: The protein details plot shown with the user data results (decreased) integrated and shown as a light green horizontal line.