ASMS 2022 Minneapolis



# Visualization of Mass Spectrometry Related Data

#### ASMS Bioinformatics Interest Group

Arzu Tuğçe Güler, Amsterdam University Medical Centers, Amsterdam Claire O'Donovan, European Bioinformatics Institute, Cambridge

# Agenda

- ✤ Overview
- Visualization tools
- Metabolomics study examples
- ✤ Interactive discussion

#### **Two different worlds?**





Figure made by Arzu Tugce Guler created with Biorender.com

#### A picture is worth a thousand words

> datasets::anscombe

Anscombe's Quartet

| x1 x2 x3    | x4 y1     | y2 y3       | y4        |           |           |           |           |           |
|-------------|-----------|-------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1 10 10 10  | 8 8.04    | 9.14 7.46   | 6.58      |           |           |           |           |           |
| 2 8 8 8     | 8 6.95    | 8.14 6.77   | 5.76      |           |           |           |           |           |
| 3 13 13 13  | 8 7.58    | 8.74 12.74  | 7.71      |           |           |           |           |           |
| 4 9 9 9     | 8 8.81    | 8.77 7.11   | 8.84      |           |           |           |           |           |
| 5 11 11 11  | 8 8.33    | 9.26 7.81   | 8.47      |           |           |           |           |           |
| 6 14 14 14  | 8 9.96    | 8.10 8.84   | 7.04      |           |           |           |           |           |
| 7 6 6 6     | 8 7.24    | 6.13 6.08   | 5.25      |           |           |           |           |           |
| 8 4 4 4     | 19 4.26   | 3.10 5.39   | 12.50     |           |           |           |           |           |
| 9 12 12 12  | 8 10.84   | 9.13 8.15   | 5.56      |           |           |           |           |           |
| 10 7 7 7    | 8 4.82    | 7.26 6.42   | 7.91      |           |           |           |           |           |
| 11 5 5 5    | 8 5.68    | 4.74 5.73   | 6.89      |           |           |           |           |           |
| > fBasics:: | basicStat | s(anscombe) |           |           |           |           |           |           |
|             | x         | 1 x2        | x3        | x4        | y1        | y2        | у3        | y4        |
| nobs        | 11.00000  | 0 11.000000 | 11.000000 | 11.000000 | 11.000000 | 11.000000 | 11.000000 | 11.000000 |
| NAS         | 0.00000   | 0.000000    | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  |
| Minimum     | 4.00000   | 0 4.000000  | 4.000000  | 8.000000  | 4.260000  | 3.100000  | 5.390000  | 5.250000  |
| Maximum     | 14.00000  | 0 14.000000 | 14.000000 | 19.000000 | 10.840000 | 9.260000  | 12.740000 | 12.500000 |
| 1. Quartile | 6.50000   | 0 6.500000  | 6.500000  | 8.000000  | 6.315000  | 6.695000  | 6.250000  | 6.170000  |
| 3. Quartile | 11.50000  | 0 11.500000 | 11.500000 | 8.000000  | 8.570000  | 8.950000  | 7.980000  | 8.190000  |
| Mean        | 9.00000   | 9.000000    | 9.000000  | 9.000000  | 7.500909  | 7.500909  | 7.500000  | 7.500909  |
| Median      | 9.00000   | 9.000000    | 9.000000  | 8.000000  | 7.580000  | 8.140000  | 7.110000  | 7.040000  |
| Sum         | 99.00000  | 0 99.000000 | 99.000000 | 99.000000 | 82.510000 | 82.510000 | 82.500000 | 82.510000 |
| SE Mean     | 1.00000   | 0 1.000000  | 1.000000  | 1.000000  | 0.612541  | 0.612568  | 0.612196  | 0.612242  |
| LCL Mean    | 6.77186   | 1 6.771861  | 6.771861  | 6.771861  | 6.136083  | 6.136024  | 6.135943  | 6.136748  |
| UCL Mean    | 11.22813  | 9 11.228139 | 11.228139 | 11.228139 | 8.865735  | 8.865795  | 8.864057  | 8.865070  |
| Variance    | 11.00000  | 0 11.000000 | 11.000000 | 11.000000 | 4.127269  | 4.127629  | 4.122620  | 4.123249  |
| Stdev       | 3.31662   | 5 3.316625  | 3.316625  | 3.316625  | 2.031568  | 2.031657  | 2.030424  | 2.030579  |
| Skewness    | 0.00000   | 0.000000    | 0.000000  | 2.466911  | -0.048374 | -0.978693 | 1.380120  | 1.120774  |
| Kurtosis    | -1.52892  | 5 -1.528926 | -1.528926 | 4.520661  | -1.199123 | -0.514319 | 1.240044  | 0.628751  |



Take home lesson: never trust summary statistics alone; always visualize your data!

#### Datasaurus (aka Anscombosaurus)



Take home lesson: "never trust summary statistics alone; always visualize your data!"

Albert Cairo: http://www.thefunctionalart.com/2016/08/download-datasaurus-never-trust-summary.html

#### Visualization of mass spectrometry data



1.Doana. *Metabolic Phenotyping in Personalized and Public Healthcare* 2016 2.<u>https://phosphopedia.gs.washington.edu/PhosphoproteomicsAssay/index.xhtml</u> 3.Saeed et. al IE*EE/ACM Trans Comput Biol Bionform* 2018



#### The -omics





Figure made by Arzu Tugce Guler created with Biorender.com

# Visualizations for mass spectrometry-based proteomics data

- Numeric data (with or without labels)
   e.g. volcano plots, heatmaps, boxplots, PCAs etc.
- Trees / networks

e.g. ontologies, gene ontology enrichments, protein interaction networks etc.

• Images

e.g. experiment workflows, anatomical visualizations, western blots etc.

#### Volcano plots



## Heatmap







## Boxplot



## PCA



(Median) PCA of all samples

PC 3

#### Western blots / gels, microscope images



## Tools for visualizing numeric data

|    | A B                  | С        | D        | E        | F               | G                     | н      | 1   | J | К | L | М      | N     | 0     | Ρ       | Q       | R      | S      | т | U | V |
|----|----------------------|----------|----------|----------|-----------------|-----------------------|--------|-----|---|---|---|--------|-------|-------|---------|---------|--------|--------|---|---|---|
| 1  | probability spectrum | expect   | ions     | peptide  | protein         | calc_neutral_pep_mass | ppm    |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 2  | 0,9996 calibrated    | 0,00013  | 3 okt-22 | K.EALGIP | PA sp P31939    | 1173,6394             | 7,9633 | o   |   |   |   |        |       |       | O       |         |        |        |   |   |   |
| 3  | 0,9957 calibrated    | 0,0181   | l sep-22 | K.QITQS  | AL sp P78527    | 1299,7147             | 7,3962 |     |   |   |   | Calibr | ated: | mme 5 | 5. inte | rnal ta | raet 0 | .5 ppm | ו |   |   |
| 4  | 1 calibrated         | 0,000161 | l sep-68 | K.EKAPSI | IF sp   P17980  | 1959,0565             | 6,9717 | 600 |   |   |   |        |       |       | ,       |         | 0      |        |   |   |   |
| 5  | 0,9985 calibrated    | 0,000429 | ep-20    | K.DAYQV  | IL sp Q9BZZ     | 1219,6449             | 6,9061 | 000 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 6  | 1 calibrated         | 9,29E-05 | 5 nov-20 | K.HLSVN  | D sp P30048     | 1205,6517             | 6,5533 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 7  | 0,9996 calibrated    | 1,23E-06 | 5 aug-18 | K.LASDLI | E sp P12814     | 1214,6659             | 6,5425 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 8  | 1 calibrated         | 0,000526 | 5 aug-30 | R.IGNVE  | IS sp P15927    | 1710,004              | 6,224  | 500 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 9  | 0,9977 calibrated    | 0,00076  | 5 okt-18 | R.DLIVAT | Fl/ sp P31939   | 1041,6434             | 6,1518 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 10 | 1 calibrated         | 3,50E-06 | 5 dec-84 | K.SGYHQ  | (Si sp   P10768 | 2307,1244             | 6,1046 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 11 | 0,9917 calibrated    | 0,0084   | 1 jul-28 | R.TINLY  | PL sp   O4380   | 1744,9036             | 6,1018 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 12 | 1 calibrated         | 8,88E-05 | 5 15/34  | K.VSHVS  | TC sp P00558    | 1739,9054             | 6,0417 | 400 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 13 | 0,996 calibrated     | 0,000578 | 8 nov-24 | K.ASSAA  | Gl sp Q1456     | 1172,6513             | 5,9881 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 14 | 0,9988 calibrated    | 0,00306  | 5 okt-20 | K.VDATA  | E sp P13667     | 1132,5612             | 5,952  |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 15 | 0,997 calibrated     | 5,15     | 5 2/240  | K.GPRGS  | Q sp Q6ZN1      | 3417,465              | 5,7625 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 16 | 1 calibrated         | 0,000106 | 5 okt-24 | R.VTWD   | SS sp Q9ULV     | 1537,6984             | 5,6546 | 300 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 17 | 0,9981 calibrated    | 0,364    | 4 okt-60 | K.EGKPT  | IV sp 07593     | 1844,9044             | 5,6252 | ¢   |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 18 | 1 calibrated         | 0,00664  | 1 nov-72 | R.KPLVLO | C[ sp P27695    | 2190,1467             | 5,5782 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 19 | 0,9942 calibrated    | 0,00695  | jun-26   | K.VWQL   | Q[ sp Q9NYU     | 1661,8526             | 5,3964 | 200 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 20 | 1 calibrated         | 2,21E-05 | 6 dec-22 | K.TGQAA  | E sp 00023      | 1156,6452             | 5,3292 | 200 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 21 | 1 calibrated         | 0,00201  | 15/52    | K.AHGGY  | 'S' sp P06576   | 1405,6739             | 5,3035 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 22 | 0,9976 calibrated    | 0,0109   | 9 nov-18 | K.ALGDL  | IS sp Q9Y49     | 987,5601              | 5,2979 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 23 | 1 calibrated         | 5,52E-06 | 5 13/56  | K.TVAGO  | D sp   P30043   | 1511,8671             | 5,2174 | 100 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 24 | 0,9966 calibrated    | 0,000705 | 5 sep-18 | K.VLDIIA | TI sp P07686    | 1098,6649             | 5,2127 | 100 |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 25 | 1 calibrated         | 6,88E-05 | 5 sep-80 | K.VIAVY  | DL sp P38646    | 2250,2148             | 5,2031 |     |   |   |   |        |       |       |         |         |        |        |   |   |   |
| 26 | 0,9991 calibrated    | 0,265    | 5 sep-44 | R.FIFEDC | QII sp P02730   | 1562,7841             | 5,1363 |     |   |   |   |        | _     |       |         |         |        |        |   |   |   |
| 27 | 0,9971 calibrated    | 0,000224 | sep-18   | K.VNTPT  | T sp P26639     | 1150,5982             | 5,1182 | 0   |   |   |   |        |       |       |         |         |        |        |   |   |   |

#### Excel\*

- Quick and adaptable
- Especially good for internal sanity checks

#### RStudio

#### File Edit Code View Plots Session Build Debug Profile Tools Help

🔍 🔹 🧠 🚽 🔚 🗐 😓 🛛 🍌 Go to file/function 🔤 🗐 👻 Addins 👻

| - |                   |
|---|-------------------|
| K | Project: (None) * |

| le Terminal × Jobs ×  | Environment History Connect | actions Tutorial                          |                                 |                |
|---|-----------------------------|---|---------------------------------|----------------|
| ers/arzutugceguler/Dropbox (Personal)/wdir/cortex/ 🧇  | 🧹 😅 🔒 🐨 Import Dataset 👻    | 1   |                                 | 🗏 List 👻       |
| }   | Global Environment *        |   |                                 | 0              |
| else if( (fold_change[i]>min_log2_fold_change) && (pval[i,1] <max_pval) )<="" td=""><td>no up</td><td>33</td><td></td><td></td></max_pval)> | no up                       | 33  |                                 |                |
| i noints(fold_channe[i]_log10(nval[i_1])_nch=19col="forestoreen").  | non na intensities          | Large numeric (42030 elements, 4.7 MB)    |                                 |                |
| <pre>#text(fold_change[i]+0.20,-log10(pva[i])+0.20, labels[i,1], col="forestgreen");</pre>  | protein                     | chr [1:4011] "Cullin-4B" "Brefeldin A-    | inhibited quanine nucleotide-ex | change protein |
| changes[i] <- "+"   | 0140_1                      | 36871                                     |                                 |                |
| <pre>print(c("up", labels[i,1]))</pre>  | 0140 2                      | 3495L                                     |                                 |                |
| no_up <- no_up + 1  | 0140 3                      | 3608L                                     |                                 |                |
| 3   | Q140_4                      | 3642L                                     |                                 |                |
| "up" "Agrn"   | Q20_1                       | 3491L                                     |                                 |                |
| "up" "Kalrn"  | Q20_2                       | 3416L                                     |                                 |                |
| down Iomms<br>"un" "Knna4"  | Q20_3                       | 3518L                                     |                                 |                |
| up Lamn"  | Q20_4                       | 3374L                                     |                                 |                |
| 'up" "Pde4a"  | Files Dista Destance Hab    | 10-10-10-10-10-10-10-10-10-10-10-10-10-1  |                                 |                |
| up" "Srp54"<br>'deum" "Covide"  | Files Plots Packages Hel    | Viewer                                    |                                 |                |
| down coxre<br>//own "Prma"  | 🗢 🖘 🎤 Zoom 🖓 🚈 Export       | - 0 🥖                                     |                                 | 🧐 Publish      |
| up "Lops"<br>up" "Abcel"<br>up" "kjfap3"<br>"up" "kjfap3"<br>up "sol2"  | 1                           | Cortex mutant vs wi                       | dtype                           |                |
| υρ" "cul4a"<br>υρ" "Hp1bp3"<br>υρ" "ττl12"<br>down" "cdk13"<br>υρ" "Unc13a"   | ω –                         |   |                                 |                |
| up" "Raly"<br>down" "Cox8a"<br>up" "Srm"<br>"cbarp"<br>"cbarp"  | ω –                         | ° °,                                      |                                 |                |
| ລ້ອກກີ "Gprin3"<br>ມາງ "UbaG"<br>ມາງ "Trc19"<br>ມາງ "Pitrm1"  | 0(p-value                   | ၀ ၀<br>၀ဆိုရာ ၀<br>ရ ရ စ                  | •                               |                |
| Jown" "Ccdc9"<br>p" "bcaf11"<br>p" "tbX15"  |                             |   | •                               |                |
| up" "Man2c1"  |                             | • ***********                             | 63° 1                           |                |
| up segn   | (4 -                        |   |                                 |                |
| p" "Pacsin3"  |                             | 2 10 10 10 10 10 10 10 10 10 10 10 10 10  |                                 |                |
|   |                             | 8 (B) |                                 |                |
| ip" "Acs13"   |                             |   | ဆီႏိုင်ပီ ဖ                     |                |
| p" "Ttpal"<br>"Lamtor?"   |                             |   | §~~ @                           |                |
| p" "Cvth1"  |                             |   | 0                               |                |
| lown" "Go1ga5"  | ° -                         |   |                                 |                |
| p" "Mtmr1"  |                             |   |                                 | 1              |
| ne(v=min_log2_fold_change, lty=3);<br>ne(v=-min_log2_fold_change, lty=3);   | -6                          | -4 -2 0                                   | 2 4                             | 6              |
| <pre>ine(h=-log10(max_fdr), lty=3);</pre>   |                             | log2(fold change)                         |                                 |                |
|   | *                           |   |                                 |                |

# Useful R libraries for more advanced, publication quality visualizations

There are currently 161 Proteomics and 118 Mass Spectrometry packages in Bioconductor version 3.15.

- ggplot2
- lattice
- shiny\*
- plotly\*
- MSstats
- protViz



#### Perseus



## Tools for icons / anatomical visualizations



## Tools for icons / anatomical visualizations



\* commercial

## Tools for icons / anatomical visualizations



COMICS



## Tools for network visualizations





#### Tools for network visualizations



## Interactive plots



• FiglinQ\*



\* commercial / paid subscription

#### 

#### ▶ ♥ M <>> ♥ <> ● ♥ <> ●

<u>↑</u> Import

atguler 🗸



| * |   |        |              |              |              |             |              |              |              |              |              |             |              |     |
|---|---|--------|--------------|--------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|-----|
|   | А | В      | С            | D            | Е            | F           | G            | Н            | I.           | J            | К            | L           | М            |     |
| 1 | 1 | Nampt  | 0.215198204  | 0.298555106  | 0.452983439  | 0.839390337 | 0.328380436  | -1.574401259 | -1.55439043  | -1.675691485 | 0.492631763  | 0.949865162 | 0.964351356  | 0.2 |
| 2 | 2 | Actr3  | 0.052884884  | -0.432632983 | 0.443350881  | 1.353461981 | 0.107286558  | -1.084760666 | -1.380514383 | -1.896587133 | 0.531504989  | 0.719460905 | 0.667079151  | 0.9 |
| 3 | 3 | Add1   | 0.198700041  | -0.016361643 | 0.040212072  | 0.758406579 | -0.233907223 | -1.302159309 | -1.38063097  | -1.088271618 | -0.533631742 | 1.187776089 | 1.9056741    | 0.4 |
| 4 | 4 | Cdk5   | 0.03537906   | -0.776342154 | 0.374365509  | 0.962719321 | -1.003786564 | -0.605718374 | -1.678633332 | -0.268917114 | -0.013393712 | 0.293967843 | 2.172554255  | 0.5 |
| 5 | 5 | RpI9   | 0.727042794  | -0.953475177 | -0.081508435 | 1.052151918 | -0.345692635 | -1.121615767 | -1.85233593  | -0.614684105 | 0.554352283  | 0.243497774 | 1.326607227  | 1.0 |
| 6 | 6 | Erh    | 0.376109749  | -0.505896509 | -0.168753669 | 0.848609746 | -0.409587651 | -2.162545919 | -1.431761503 | 0.322642356  | 1.076853275  | 0.27465266  | 0.753981233  | 1.0 |
| 7 | 7 | Mvp    | -0.933950245 | -0.813658237 | 0.580756485  | 1.300670505 | -0.854547679 | -0.713124335 | -1.241092205 | -0.788911879 | 0.886659861  | 1.467129707 | 0.064919449  | 1.0 |
| 8 | 8 | Cacnb4 | -0.247023597 | 0.264265597  | 1.223759294  | 0.792339027 | -0.476846278 | -1.638639212 | -1.205763578 | -0.688996553 | 0.807108283  | 1.749877453 | -0.187789887 | -0. |
|   |   |        |              |              |              |             |              |              |              |              |              |             |              |     |

#### cerebellum\_insol\_data × +



## Searching for (visualization) tools



• Essential scientific and technical information about software tools, databases and services for bioinformatics and the life sciences.

| > C ☵   VPN 🖨 bio.tools/t   | C            | 1 🖄 🖸 🥪 🕽           | ⊳ ♡  ♦       | × 🗠 🕅    |  |
|---|--------------|---------------------|--------------|----------|--|
| Everything: visualization tools × Search bio.tools  | 660 tools    | Explore +           | Login        | Sign-up  |  |
| GenBank taxonomy tools   👌 🗯  |              |                     |              |          |  |
| Utilities for manipulations and visualization tasks on GenBank taxonomic information.   |              |                     |              |          |  |
| Taxonomy       Suite     MIT       GenBank taxonomy tools   |              |                     |              |          |  |
| Gitools   👌 🏭 🏟   |              |                     |              |          |  |
| A framework for analysis and visualization of multidimensional genomic data using interactive heat-maps.  |              |                     |              |          |  |
| Genomics<br>Visualisation<br>Desktop application  |              |                     |              |          |  |
| Repitools   👌 🗮 🗯   |              |                     |              |          |  |
| Tools for the analysis of enrichment-based epigenomic data. Features include summarization and visualization of epigenomic data across promoters according to gene emethylation/binding, BayMeth for quantifying methylation etc. | xpression co | ontext, finding reg | ions of diff | erential |  |
| Epigenomics Differentially-methylated region identification Transcriptional regulatory element prediction   |              |                     |              |          |  |
| Command-line tool Library GPL-2.0 BioConductor  |              |                     |              |          |  |



### Colors & ambiguity



0140\_1 0140\_3 0140\_2 0140\_2 - Q50\_1 - Q20\_3 - Q50\_2 • Q50\_3 - Q20\_2 • 020\_1 ·Q50\_4

log2(fold change)

## Colors & ambiguity







# Mass spectrometry data in MetaboLights

the home for metabolomics experiments and derived information

Claire O'Donovan EMBL-EBI Metabolomics Team www.ebi.ac.uk





# FAIR Guiding Principles

Provide a guideline for data producers and publishers to enhance the reusability of scientific data

Findable Accessible Interoperable Re-useable Re-useable Findable Interoperable Re-useable Accessible Interoperable Re-useable Interoperable Re-useable

Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.



#### Metabolomics: data management





Can metabolomics learn from the other omic techniques and not repeat many of the same mistakes that have been made in which a lot of data were collected but little information was gleaned?

#### Metabolomics: the final frontier?

Timothy D Veenstra *Genome Medicine* 2012 **4**:40 <u>https://doi.org/10.1186/gm339</u> © BioMed Central Ltd. 2012



# The Metabolomics Standards Initiative (MSI)

The MSI is an academic policy provider, to support the development of open data and metadata formats for metabolomics.

- community-agreed reporting standards
- minimal information (MI) checklists
- data exchange formats to support the MIs reporting standards, allow data to be efficiently applied, shared and reused







#### Open Source Study Repository & Metabolite Knowledgebase

| MetaboL  | ights   | Examples: Alanine, Homo s   | Search   |
|--|---|---|--|
| Home Browse Studies Browse   | Compounds Browse Species Download Help  | Give us feedback About  | Submit Study Login   |
| MetaboLights is a database for Metal<br>technique and covers metabolite stru<br>concentrations, and experimental dat<br>for a number of leading journals.<br>More about us | bolomics experiments and derived information. The ctures and their reference spectra as well as their bia from metabolic experiments. MetaboLights is the r | database is cross-species, cross-<br>ological roles, locations and<br>recommended Metabolomics repository | Tweets by @MetaboLights  MetaboLights  MetaboLights  MTBLS706: Gluconeogenesis using glycer as a substrate in bloodstream ebi.ac.uk/metabolights/M  Dec 11, 20  MetaboLights  MetaboLights |
|  |   | 0   | MTBLS658: A robust intracellular metabol<br>extraction protocol for hu<br>ebi.ac.uk/metabolights/M   |
| Study  | Compound Library  | Training  | MetaboLights<br>@MetaboLights<br>MTBLS538: 13C metabolomics reveals  |
| BROWSE   | COMPOUNDS   | TRAINING ONLINE   | widespread change in carbon fate du<br>ebi.ac.uk/metabolights/M  |
| ORCID SEARCH   | SPECIES   | QUICK TOUR  | ○ Dec 5, 20           MetaboLights           @MetaboLights   |

https://www.ebi.ac.uk/metabolights



# MetaboLights user base





#### **MetaboLights**



Geographical distribution of submitted studies (Top 10: China: 507, USA: 362, UK: 356, Germany: 235, France: 94, Japan: 68, Spain: 66, Italy: 64, Australia: 53, Netherlands: 39)





MetaboLights study submission rates per year







## **MetaboLights Database**



- Cross species, cross technique
- Comprehensive information collection
- Open access, discoverable through various portals



# MetaboLights: Ontology mapping

**Ontologies: Controlled vocabulary** 





## MetaboLights: Browse or search studies

| Home Browne Studies   | Browso Compo | Inde Browno Species                | Download Hol                       | n Čivo us foodback | About           |                   | Submit Study   | Login    |
|-----------------------|--------------|------------------------------------|------------------------------------|--------------------|-----------------|-------------------|----------------|----------|
| Home Browse Studies   | Browse Compo | ands Browse Species                | Download Hei                       | p Give us leedback | About           |                   | Submit Study   | Login    |
| MetaboLights / Search |              |                                    |                                    |                    |                 |                   |                |          |
| Filter your results   |              | 145 results , showing              | g 1 to 10                          |                    |                 | 6 C               | + Show mo      | ore data |
| Туре                  | ۲            |                                    |                                    |                    |                 | _                 |                |          |
| ✓ study □ compound    |              | Utilization of M<br>Carcinoma in P | etabolomics to<br>atients with Liv | Identify Serum     | Biomarkers fo   | r Hepatocellula   | ar             | 0        |
| Technology            | ٩            | 🖞 Study Identifier                 | MTBLS17                            |                    | Organism        | Homo sapiens      |                |          |
| Organism              | ۲            | G Study Size                       | 27.58GB                            |                    | Study Factors   | Disease, Injectio | on, Experiment |          |
| Organism Part         | ۲            | Submitted by                       | Habtom Ressom                      | 🖂 Jinlian Wang 🖂   |                 |                   |                |          |
|                       |              |                                    |                                    |                    |                 |                   |                |          |
|                       |              | A lipidomic hun                    | an nonulation                      | and translation    | al feeding stud | v of benatic      |                | 0        |
|                       |              | steatosis and d                    | e novo lipogen                     | esis (Human pla    | asma assay)     | y of nepatic      |                |          |
|                       |              | 🗄 Study Identifier                 | MTB                                | _S616              | Organism        | Homo              | sapiens        |          |
|                       |              | 🕀 Study Size                       | 2.450                              | ЗB                 | Study Factors   | Samp              | ling time      |          |
|                       |              |                                    | luloo                              | Griffin 🖂          |                 |                   |                |          |

MTBLS72

32.22GB

Robert Padilla 🖂

Organism

Study

Factors

sapiens

Cognitive Status

BAO:reference compound, blank, Homo

Study Identifier

Study Size

Submitted by



| MBL-EBI 🚯 About us 🔥 Training 🔀 Research 🔍 Service:  | 5   |  | EMBL-EBI   |
|--|---|--|--|
| 🍌 MetaboLight  | S   | Examples: Alanine, Homo sapier   | Search<br>ns, Urine, MTBLS1  |
| Home         Browse Studies         Browse Compounds           Status         Public         Release Date         2019-09-19   | Browse Species Download Help  | Give us feedback About   | Submit Study Login   |
| MTBLS749: Alterations in the tyrosine a<br>fluid of Huntington's disease subjects  | and phenylalanine pathway   | rs revealed by biochemical p   | profiling in cerebrospinal   |
| Huntington's disease (HD) is a severe neurological dise<br>expansion in the huntingtin (HTT) gene, but how this tr<br>analyzed the metabolome of cerebrospinal fluid (CSF)<br>tyrosine metabolism, including tyrosine, thyroxine, L-Dd<br>demonstrated moderate to strong associations to meas<br>premanifest HD subjects. The phenylalanine and the pu<br>lumichrome were commonly found in mutated HTT car<br>findings demonstrates that the CSF metabolome can b<br>novel HD therapies.<br>Homo saplens<br>PUBLICATIONS<br>Alterations in the tyrosine and<br>phenylalanine pathways revealed by<br>biochemical profiling i | ase leading to psychiatric symptoms,<br>anslates into the clinical phenotype of<br>from premanifest and manifest HD sul<br>DPA and dopamine, was significantly a<br>sures of disease severity and symptor<br>inter metabolisms were also significan<br>firers and the levels correlated with the<br>e used to characterize molecular path | motor impairment and cognitive decili<br>f HD remains elusive. Using liquid chro<br>ojects as well as control subjects. Inter<br>altered in manifest compared with pren<br>ns. Thyroxine and dopamine also corre<br>tity altered, but associated less to dise.<br>five year risk of disease onset in pren<br>togenesis occurring in HD, which may the<br>openesis occurring in HD, which may the openesis occurring in HD, which may the<br>openesis occurring in HD, which may the openesis occurring in HD, which may the opene | ne. The disease is caused by a CAG<br>matography mass spectrometry, we<br>-group differences revealed that the<br>manifest HD. These metabolites<br>lated with the five year risk of onset in<br>ase severity. Decreased levels of<br>nanifest carriers. These blochemical<br>be essential for future development of |
| Herman Stephanie, Niemelä Valter, Emami     Khoonsari Payam, Sun      O  | Matabalitas Eilas   |  |  |
| Descriptors Protocols Samples Assays   | Hierabolites Files  |  |  |
| Huntingtin (Huntington Disease) Untargeted Me  | atabolites Liquid Chromatography Ma   | iss Spectrometry   |  |
| Factors  |   |  |  |
| Phenotype : Phenotype<br>Replicate : Replicate<br>MS Format : Format   |   |  |  |





#### Release Date 2019-09-19

MTBLS749: Alterations in the tyrosine and phenylalanine pathways revealed by biochemical profiling in cerebrospinal fluid of Huntington's disease subjects

Kim Kultima, Stephanie Herman

Status Public

Huntington's disease (HD) is a severe neurological disease leading to psychiatric symptoms, motor impairment and cognitive decline. The disease is caused by a CAG expansion in the huntingtin (HTT) gene, but how this translates into the clinical phenotype of HD remains elusive. Using liquid chromatography mass spectrometry, we analyzed the metabolome of cerebrospinal fluid (CSF) from premanifest and manifest HD subjects as well as control subjects. Inter-group differences revealed that the tyrosine metabolism, including tyrosine, thyroxine, L-DOPA and dopamine, was significantly altered in manifest compared with premanifest HD. These metabolites demonstrated moderate to strong associations to measures of disease severity and symptoms. Thyroxine and dopamine also correlated with the five year risk of onset in premanifest HD subjects. The phenylalanine and the purine metabolisms were also significantly altered, but associated less to disease severity. Decreased levels of lumichrome were commonly found in mutated HTT carriers and the levels correlated with the five year risk of disease onset in premanifest carriers. These biochemical findings demonstrates that the CSF metabolome can be used to characterize molecular pathogenesis occurring in HD, which may be essential for future development of novel HD therapies.

#### Homo sapiens

PUBLICATIONS

Alterations in the tyrosine and phenylalanine pathways revealed by biochemical profiling i...

🚓 Herman Stephanie, Niemelä Valter, Emami Khoonsari Payam, Sun...

| Ð |  |
|---|--|
|   |  |

Factors Phenotype : Phenotype Replicate : Replicate





| <complex-block></complex-block>   | 🗌 EMBL-EBI 🔮 About us 🔥 Training 🕺 Research 🔌 Services  |   | EMBL-EBI  | · •    |
|---|---|---|---|--------|
| Comparison of the constrained of the constrain | Metabol ights   |   | Search  |        |
| Note:   | Wietabelighto   |   |   |        |
| Image: Second Secon       | Home Browse Studies Browse Compounds Browse Species Download Help   | Give us feedback About  | Submit Study Login  | $\leq$ |
| Build Strategies   Minist of Hundington's disease subjects Minist of Hundington's disease subjects Minist of Hundington's disease leading to psychiatic synctoms, motor impairment and cognitive decline. The disease is caused by a CAS expansion in the hundington (HTI) gene, luch how this translates into the clinical phenotype of HO remains dusc. Using liquid chromatography mass spectrometry, we analyzed the metaboline of covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They cover and covershoping huld (CS) from pre-mainfest HO readings. They   | Status Public Release Date 2019-09-19   |   |   |        |
| Kin Kultima, Stephanie Herman     Huntington's disease (HD) is a severe neurological disease leading to psychiatric symotoms, motor impairment and cognitive decline. The disease is caused by a OAD expension in the huntingtin (HTT) gene, but how this translates into the clinical phonotype of HD remained usables. Using leadid chromatography mass spectrometry, on analyzed the metabolices of expension in the huntingtin (HTT) gene, but how this translates into the clinical phonotype of HD remained usables. Using leadid chromatography mass spectrometry, on analyzed that the metabolices of expension in the huntingtin (HTT) gene, but how this translates and dispersive. Why and subjects and usables. Using leadid chromatography mass spectrometry of expension in the huntingtin (HTT) gene, but how the translate of dispersive eventy hard significantly altered in the huntingtin the the specification of the seventy and dispersive. The dispersive eventy hard significantly altered in the human term of the dispersive eventy hard significantly altered in the huntingtin (HTT) gene, but how the interview event hard significantly altered in the huntingtin (HTT) gene (HT                                  | MTBLS749: Alterations in the tyrosine and phenylalanine pathways<br>fluid of Huntington's disease subjects  | ; revealed by biochemical profi   | iling in cerebrospinal  |        |
| Hurrington's disease (HD) is a severe neurological disease leading to psychiatric symptoms, motore, motor Using fluid chores. The disease is caused by A CAG evaluation in the Nuringtin (HTT) gene, but now list transities to the clinical phones evaluation. Using fluid chores revealed that the through presser, through        | Kim Kultima, Stephanie Herman   |   |   |        |
| PUBLICATIONS     Publications in the broncine and     escriptors        Protocols     Samples     Assays     Metabolites     Files     Huntingtin (Huntington Disease)     Untargeted Metabolites   Liquid Chromatography Mass Spectrometry   Factors     Phenotype:   Phenotype:   Replicate:   Replicate:   MS Format:  | Huntington's disease (HD) is a severe neurological disease leading to psychiatric symptoms, r<br>expansion in the huntingtin (HTT) gene, but how this translates into the clinical phenotype of 1<br>analyzed the metabolome of cerebrospinal fluid (CSF) from premanifest and manifest HD subj<br>tyrosine metabolism, including tyrosine, thyroxine, L-DDPA and dopamine, was significantly al<br>demonstrated moderate to strong associations to measures of disease severity and symptom<br>premanifest HD subjects. The phenylalanine and the purine metabolisms were also significant<br>lumichrome were commonly found in mutated HTT carriers and the levels correlated with the<br>findings demonstrates that the CSF metabolome can be used to characterize molecular pathor<br>novel HD therapies. | notor impairment and cognitive decline. Th<br>HD remains elusive. Using liquid chromator<br>ects as well as control subjects. Inter-grou<br>tered in manifest compared with premanife<br>s. Thyroxine and dopamine also correlated<br>ly altered, but associated less to disease s<br>five year risk of disease onset in premanife<br>ggenesis occurring in HD, which may be es | he disease is caused by a CAG<br>graphy mass spectrometry, we<br>up differences revealed that the<br>est HD. These metabolites<br>I with the five year risk of onset in<br>everity. Decreased levels of<br>ast carriers. These biochemical<br>sential for future development of |        |
| PUBLICATIONS     Attractions in the turnstine and     escriptors        Protocols     Samples     Assays     Metabolites     Files     Factors     Phenotype:     Phenotype:     Phenotype:     Phenotype:     Missionationationationationationationationat   | Homo sapiens  |   |   |        |
| Attractions in the turneling and     Protocols     Samples     Assays   Metabolites     Files     KEYWORDS     Huntingtin (Huntington Disease)   Untargeted Metabolites   Liquid Chromatography Mass Spectrometry     Factors     Phenotype:   Phenotype:   Phenotype:   Phenotype:   Replicate:   Replicate:   Replicate:   Format::   | PUBLICATIONS  |   |   |        |
| Rescriptors Protocols Samples Assays Metabolites Files     KEYWORDS     Huntingtin (Huntington Disease) Untargeted Metabolites Liquid Chromatography Mass Spectrometry     Factors     Phenotype   Phenotype   Phenotype   Replicate   Replicate   Format   | Alterations in the turnsine and   |   |   |        |
| KEYWORDS         Huntingtin (Huntington Disease)       Untargeted Metabolites       Liquid Chromatography Mass Spectrometry         Factors         Phenotype: Phenotype         Replicate: Replicate         MS Format: Format   | escriptors Protocols Samples Assays Metabolites Files   |   |   |        |
| Huntingtin (Huntington Disease)       Untargeted Metabolites       Liquid Chromatography Mass Spectrometry         Factors  | KEYWORDS  |   |   |        |
| Huntingtin (Huntington Disease)       Untargeted Metabolites       Liquid Chromatography Mass Spectrometry         Factors       Factors         Phenotype : Phenotype       Phenotype : Phenotype         Replicate : Replicate : Replicate : MS Format : Format       Format : Format   |   |   |   |        |
| Factors Phenotype Replicate : Replicate MS Format : Format  | Huntingtin (Huntington Disease) Untargeted Metabolites Liquid Chromatography Mass   | Spectrometry  |   |        |
| Phenotype : Phenotype         Replicate : Replicate         MS Format : Format  |   |   |   |        |
| Phenotype         Replicate : Replicate         MS Format : Format  | Factors   |   |   |        |
| Replicate : Replicate MS Format : Format  | Phenotype : Phenotype   |   |   |        |
| MS Format : Format  | Replicate : Replicate   |   |   |        |
|   | MS Format : Format  |   |   |        |



| 🐴 EMBL-EBI 🏾 🔀 About us   | ⊨ 🌡 Training 🔯 Research 🔍 Services   | EMBL-EBI 🎆 Hinxton +  |    |
|---|--|---|----|
|   | VetaboLights   | Search<br>Examples: Alanine, Homo sapiens, Urine, MTBLS1  |    |
| Home  | Browse Studies Browse Compounds Browse Species Download Help   | Give us feedback About Submit Study Login   |    |
| Status Public<br>MTBLS749: A<br>fluid of Hunti  | Release Date 2019-09-19<br>Uterations in the tyrosine and phenylalanine pathways<br>ngton's disease subjects   | revealed by biochemical profiling in cerebrospinal  |    |
| Kim Kultima, Steph  | ianie Herman   |   |    |
| Huntington's disea<br>expansion in the hu<br>analyzed the meta<br>tyrosine metabolisi<br>demonstrated moo<br>premanifest HD su<br>lumichrome were o<br>findings demonstra<br>novel HD therapies | se (HD) is a severe neurological disease leading to psychiatric symptoms, n<br>untingtin (HTT) gene, but how this translates into the clinical phenotype of H<br>bolome of cerebrospinal fluid (CSF) from premanifest and manifest HD subje-<br>m, including tyrosine, thyroxine, L-DOPA and dopamine, was significantly alt<br>derate to strong associations to measures of disease severity and symptoms<br>ubjects. The phenylalanine and the purine metabolisms were also significant<br>commonly found in mutated HTT carriers and the levels correlated with the f<br>ates that the CSF metabolome can be used to characterize molecular patho<br>s. | notor impairment and cognitive decline. The disease is caused by a CAG<br>HD remains elusive. Using liquid chromatography mass spectrometry, we<br>ects as well as control subjects. Inter-group differences revealed that the<br>tered in manifest compared with premanifest HD. These metabolites<br>s. Thyroxine and dopamine also correlated with the five year risk of onset in<br>ly altered, but associated less to disease severity. Decreased levels of<br>five year risk of disease onset in premanifest carriers. These biochemical<br>genesis occurring in HD, which may be essential for future development of |    |
| Homo sapiens  |  |   |    |
| PUBLICATIONS  |  |   |    |
| Alterations in 1  | the twocine and  |   |    |
| criptors Protocols Sample<br>KEYWORDS<br>Huntingtin (Huntington Disease)  | s Assays Metabolites Files Untargeted Metabolites Liquid Chromatography Mass   | Spectrometry  |    |
|   | Phenotype  |   |    |
| Factors<br>Phenotype : Phenotype<br>Replicate : Replicate<br>MS Format : Format   | Ontology Details<br>Name Elo<br>Prefize EFO<br>Short Form EFO_0000651<br>Hith: http://www.Ebi.Ac.Uk/Efo/EFO_0000651<br>Tem Details<br>Dire Obsenvable form Taken By Some Character (Or Group Of Characteri) In An Individual Or An O<br>Database Cons. reference: SNOMEDCT8116006 NChC16977 MO:192 NIFSTDibirnles, 2087 McSi-<br>Tem Editor: Tomasz Adamusiak James Malone Jie Zheng   | rganiem, Excluding Pathology And Disease. The Detectable Outward Manifestations Of A Specific Genotype.<br>±D010641   |    |
|   | -  |   | ОК |
|   | L  |   |    |

Des



| Descriptors Protocols Samples Assays Metabolites Files   |
|--|
| i≣ Colapse all   |
| Sample collection  |
| The lumbar puncture was performed through the L3/L4 or L4/L5 interspace, and CSF was collected in a polypropylene tube that was centrifuged for 5 min at 250 x g at room temperature. The supernatant was transferred, into fresh polypropylene tubes in aliquots of 240 µL and stored at -80 °C until analysis.   |
| Extraction   |
| CSF samples were thawed on ice, and 100 µl was transferred and mixed with 410 µl ice-cold methanol (MeOH) supplemented with an internal standard cocktail (D4-6keto-prostaglandin-F1-alpha, D4-thromboxane-B2, D4-prostaglandin-F2-alpha, D4-prostaglandin-F2, D4-prostaglandin-D2, D4-15-deoxy-delta12, 14-prostaglandin-J2, D4-cortisol and levonorgestrel) at a final concentration of 0.25 µM. The samples were further vortexed for 15 s and incubated at -20 °C for 30 min, followed by centrifugation at 2040 $\sim$ more |
| Post Extraction Derivatization   |
| Chromatography   |
| The liquid chromatography-mass spectrometry analyses were performed using a Thermo Ultimate 3000 HPLC and Thermo Q-Exactive Orbitrap mass spectrometer. 10 µJ sample was injected to a Thermo Accucore aQ RP C18 column (100 x 2.1 mm, 2.6 µm particle size). The analytical gradient was initiated with an isocratic flow for 3 min (0% B), followed by a 2.6 min gradient (0-10% B), 8.3 min (10-100% B) and 3 min (100% B), followed, finally, by re-equilibration and washing of the column for 3 min (0% B), wh $\sim$ more |
| Chromatography Instrument Column Type Column Model   |
| Mass spectrometry  |
| Mass spectrometry data were acquired in profile mode (in positive and negative ionization mode) on a Thermo Q-Exactive Orbitrap mass spectrometer, using a mass range of 70-900 m/z during the first 5 min and 148-900 m/z in the following 15 min (to avoid low mass contaminants) in the positive ionization mode and 70-900 m/z throughout in negative ionization mode. To improve the identification of metabolites, tandem mass spectrometry analyses in positive and negative ionization mode were performed s • more      |
| Scan Polarity Scan M/z Range Instrument Mass Analyzer Ion Source   |
| Data transformation  |
| The acquired raw data was converted to an open source format (mzML). Peaks were centroided by msconvert from ProteoWizard [1] and preprocessed using the following pipeline within the KNIME platform [2]: the peak-picked data was quantified by FeatureFinderMetabo [3] and the resulting features were linked across the samples using featureLinkerUnlabelledQT [4], allowing 10 s retention time tolerance and 5 ppm mass deviation (performed irrespective of charge state across the samples). The non-defaul • more      |
| Metabolite identification  |
| All metabolic features with a 75% coverage across samples were matched against an in house library of characterized metabolites using a 15 ppm mass<br>tolerance and a 20 s time window. Metabolites (identified metabolic features) of interest were manually curated on MS/MS level when available.  |

EMBL-EBI

| escriptors Protocols Samples  | Assays Metabolites Files   |
|---|--|
|   | ≡ Collapse al  |
| Sample collection   |  |
| The lumbar puncture was performed t 250 x g at room temperature. The sup  | through the L3/L4 or L4/L5 interspace, and CSF was collected in a polypropylene tube that was centrifuged for 5 min at<br>sernatant was transferred, into fresh polypropylene tubes in aliquots of 240 μL and stored at -80 °C until analysis.   |
| Extraction  |  |
| CSF samples were thawed in ice,<br>cocktail (D4-6keto-protaglandin-<br>delta12, 14-prostaglandin-J2, D4-<br>at -20 °C for 0 min, followed by c        | Chromatography   |
| Post Extraction Derivatization  | The liquid chromatography-mass spectrometry analyses were performed using a Thermo Ultimate 3000 HPLC and Thermo Q-Exactive Orbitrap mass  |
| Chromatography  | spectrometer. 10 µl sample was injected to a Thermo Accucore aQ RP C18 column (100 x 2.1 mm, 2.6 µm particle size). The analytical gradient was initiated with an isocratic flow for 3 min (0% B), followed by a 2.6 min gradient (0-10% B), 8.3 min (10-100% B) and 3 min (100% B), followed, finally, by re- |
| The liquid chromatography-mass s<br>spectrometer. 10 µl sample was inj<br>with an isocratic flow for 3 min (09<br>equilibration and washing of the co | equilibration and washing of the column for 3 min (0% B), wh - more  |
| Chromatography Instrument Column  | Chromatography Instrument Column Type Column Model   |
| Mass spectrometry   | Mass spectrometry  |
| Mass spectrometry data were acq<br>a mass range of 70-900 m/z durin<br>and 70-900 m/z throughout in neg<br>negative ionization mode were per          | Mass spectrometry data were acquired in profile mode (in positive and negative ionization mode) on a Thermo Q-Exactive Orbitrap mass spectrometer, using   |
| Scan Polarity Scan M/z Range In   | and 70-900 m/z throughout in negative ionization mode. To improve the identification of metabolites, tandem mass spectrometry analyses in positive and   |
| Datatransformation  | negative ionization mode were performed s 👻 more   |
| The acquired raw tata was conver<br>the following pipeline within the Kh<br>across the samples using restureL<br>state across the samples). The no    | Scan Polarity Scan M/z Range Instrument Mass Analyzer Ion Source   |
| Metabolite identification   |  |

All metabolic features with a 75% coverage across samples were matched against an in house library of characterized metabolites using a 15 ppm mass tolerance and a 20 s time window. Metabolites (identified metabolic features) of interest were manually curated on MS/MS level when available.



| Descriptors Protoco  | ls Samples Assa | ys Metabolites | Files        |                     |                         |                |            |
|----------------------|-----------------|----------------|--------------|---------------------|-------------------------|----------------|------------|
| File: s_MTBLS749.txt |                 |                |              | ltems per pag       | e: <u>494</u> ▼ 1 - 494 | 4 of 494 🛛 🛛 🕹 | < > >      |
| Filter               |                 |                |              |                     |                         |                |            |
| Protocol REF         | San             | nple Name      | Organism     | Organism part       | Phenotype               | Replicate      | MS format  |
| Sample collectio     | n 1_Rep1_NEG    |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | 1_Rep1_POS      |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | n 1_Rep2_NEG    |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |
| Sample collectio     | n 1_Rep2_POS    |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |
| Sample collectio     | n 10_Rep1_NEG   |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | n 10_Rep1_POS   |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | n 10_Rep2_NEG   |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |
| Sample collectio     | n 10_Rep2_POS   |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |
| Sample collectio     | n 11_Rep1_NEG   |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | n 11_Rep1_POS   |                | Homo sapiens | cerebrospinal fluid | control                 | 1              | MS1 format |
| Sample collectio     | n 11_Rep2_NEG   |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |
| Sample collectio     | n 11_Rep2_POS   |                | Homo sapiens | cerebrospinal fluid | control                 | 2              | MS1 format |



| Descript                | ors Protocols                        | Samples A  | ssays Met                                    | abolites F  | iles                                   |   |                            |              |                      |  |   |                                    |
|-------------------------|--------------------------------------|--|--|---|--|---|----------------------------|--------------|----------------------|--|---|------------------------------------|
| Assay<br>File:<br>Filte | y Sheet 1 Assay<br>a_MTBLS1987_LC-MS | / Sheet 2 Assa                                       | ay Sheet 3                                   | Assay Sheet 4<br>txt  | 4                                      |   | lten                       | ns per page: | <u>69 ▼</u> 1-       | - 69 of 69                               | I< <                                      | > >                                |
|                         | Protocol REF                         | Parameter Value<br>-<br>Chromatography<br>Instrument | Parameter<br>Value -<br>Autosampler<br>model | Parameter<br>Value -<br>Column<br>model                                       | Parameter<br>Value -<br>Column<br>type | Parameter<br>Value -<br>Guard<br>column | Labeled<br>Extract<br>Name | Label        | Protocol REF         | Parameter<br>Value -<br>Scan<br>polarity | Parameter<br>Value -<br>Scan m/z<br>range | Parameter<br>Value -<br>Instrument |
|                         | Chromatography                       | Waters<br>ACQUITY<br>UPLC H-<br>Class System         |  | ACQUITY<br>UPLC<br>BEH<br>Amide<br>(1.7 µm,<br>2.1 mm x<br>100 mm;<br>Waters) | HILIC                                  |   |                            |              | Mass<br>spectrometry | Positive                                 | 50-1000                                   | Waters<br>Xevo<br>G2-S<br>QTof     |
|                         | Chromatography                       | Waters<br>ACQUITY<br>UPLC H-<br>Class System         |  | ACQUITY<br>UPLC<br>BEH<br>Amide<br>(1.7 µm,<br>2.1 mm x<br>100 mm;<br>Waters) | HILIC                                  |   |                            |              | Mass<br>spectrometry | Positive                                 | 50-1000                                   | Waters<br>Xevo<br>G2-S<br>QTof     |
|                         | Chromatography                       | Waters<br>ACQUITY<br>UPLC H-<br>Class System         |  | ACQUITY<br>UPLC<br>BEH<br>Amide<br>(1.7 µm,<br>2.1 mm x<br>100 mm;<br>Waters) | HILIC                                  |   |                            |              | Mass<br>spectrometry | Positive                                 | 50-1000                                   | Waters<br>Xevo<br>G2-S<br>QTof     |







| LOIS Protocols Sam        | pres Assays metabolités File        | 5                                     |                        |                       |                                    | .peg        | .d .ra     |
|---------------------------|-------------------------------------|---------------------------------------|------------------------|-----------------------|------------------------------------|-------------|------------|
| y Sheet 1 Assay Shee      | Descriptors Protocols Samples       | Assays Metabolites Files              |                        |                       |                                    |             | -          |
| : a_MTBLS1987_LC-MS_posit |                                     |                                       |                        |                       |                                    |             |            |
| or                        | Assay Sheet 1 Assay Sheet 2         | Assay Sheet 3 Assay Sheet 4           |                        |                       |                                    |             | raw        |
|                           | File: a_MTBLS1987_LC-MS_negative_re | everse-phase_metabolite_profiling.txt |                        | Items per pa          | age: 68 💌 1 - 68 of 68 🛛 🕹 🗸 🗸     | > >1        |            |
| MS Assay Name             |                                     |                                       |                        |                       |                                    |             |            |
|                           | Filter                              |                                       |                        |                       |                                    |             |            |
|                           |                                     |                                       |                        |                       |                                    |             | mp .dat    |
| QC_HILICpos01             | MS Assay Name                       | Raw Spectral Data File                | Protocol REF           | Normalization<br>Name | Derived Spectral Data File         | Protocol R  |            |
|                           |                                     |                                       |                        |                       |                                    |             | .nmr<br>ml |
|                           | QC RPLCneg01                        | RPLCneg/QC RPLCneg01.raw              | Data                   |                       | RPLCneg/RPLCnegative MSMS 20eV.mgf | Metabolit   | .mz<br>xml |
| QC_HILICpos02             |                                     |                                       | transformation         |                       |                                    | identificat |            |
|                           |                                     |                                       |                        |                       |                                    |             | d .xr      |
|                           |                                     |                                       | Dete                   |                       |                                    | Matchella   | .mz        |
|                           | QC_RPLCneg02                        | RPLCneg/QC_RPLCneg02.raw              | Data<br>transformation |                       | RPLCneg/RPLCnegative_MSMS_20eV.mgf | identificat | data       |
| QC_HILICPOSUS             |                                     |                                       |                        |                       |                                    |             | .cef       |
|                           |                                     |                                       |                        |                       |                                    |             | cnx        |
|                           | QC RPLCneg03                        | RPLCneg/QC RPLCneg03.raw              | Data                   |                       | RPLCneg/RPLCnegative MSMS 20eV.mgf | Metabolit   |            |
|                           |                                     |                                       | transformation         |                       |                                    | identificat |            |
|                           |                                     |                                       |                        |                       |                                    |             |            |



| scriptors  | Protocols Sam                                   | oles Assays            | Metabo              | ites Files            |   |                           |                |
|------------|---|------------------------|---------------------|-----------------------|---|---------------------------|----------------|
| MAF She    | MAF Sheet 2                                     | MAF Sheet 3            | MAF Sh              | eet 4                 |   |                           |                |
| File: m_M1 | FBLS1987_LC-MS_positiv                          | e_hilic_metabolite     | _profiling_v        | 2_maf.tsv             | Items per page: 13 👻  | 1 – 13 of 13   <          | < > >          |
| Filter     |   |                        |                     |                       |   |                           |                |
|            | Structure                                       | Database<br>identifier | Chemical<br>formula | SMILES                | InChi   | Metabolite identification | Mass to charge |
|            | OH<br>NH  | CHEBI:26271            |                     | OC(=0)C1CCCN1         | InChI=1S/C5H9NO2/c7-5(8)4-<br>2-1-3-6-4/h4,6H,1-3H2,(H,7,8)                   | Proline                   | 116.07029      |
| 1          | H <sub>3</sub> C I<br>H <sub>3</sub> C N 0<br>O | CHEBI:17750            |                     | C[N+](C)(C)CC([O-])=O | InChI=1S/C5H11NO2/c1-<br>6(2,3)4-5(7)8/h4H2,1-3H3                             | Betaine                   | 118.08747      |
| н          | IN NH2 OH                                       | CHEBI:28300            |                     | NC(CCC(N)=0)C(0)=0    | InChI=1S/C5H10N2O3/c6-<br>3(5(9)10)1-2-4(7)8/h3H,1-<br>2,6H2,(H2,7,8)(H,9,10) | Glutamine                 | 147.07607      |
| н          | L <sub>C</sub> S OH                             | CHEBI:16811            |                     | CSCCC(N)C(O)=O        | InChI=1S/C5H11NO2S/c1-9-<br>3-2-4(6)5(7)8/h4H,2-<br>3,6H2,1H3,(H,7,8)         | Methionine                | 150.05818      |



| F Sheet 1 MAF Sheet 2         | MAF Sheet 3            | MAF SI              | neet 4             |       |  |                  |                  |                   | -                          |                        |                        |                      |                        |            |
|-------------------------------|------------------------|---------------------|--------------------|-------|--|------------------|------------------|-------------------|----------------------------|------------------------|------------------------|----------------------|------------------------|------------|
| m_MTBLS1987_LC-MS_posit       | tive_hilic_metabolite  | _profiling_v        | 2_maf.tsv          |       | Items per page: 13                       | ▼ 1-1            | 3 of 13   <      | < > >             |                            |                        |                        |                      |                        |            |
| er                            |                        |                     |                    | Desci | riptors Protocols Sa                     | imples /         | Assays Meta      | bolites Files     | 3                          |                        |                        |                      |                        |            |
| Structure                     | Database<br>identifier | Chemical<br>formula | SMILES             | M     | AF Sheet 1 MAF Sheet                     | 2 MAF S          | Sheet 3 MAF      | Sheet 4           |                            |                        |                        |                      |                        |            |
| OH                            | CHEBI:26271            |                     | OC(=O)C1CCCN1      | File  | e: m_MTBLS1987_LC-MS_posi                | tive_reverse     | -phase_metabolit | e_profiling_v2_ma | af.tsv                     | ltem                   | s per page: 18         | ≠ 1 – 18 of 18       | < <                    | > >        |
|                               |                        |                     |                    |       | Structure                                | QC<br>Cpos07     | QC<br>RPLCpos08  | QC<br>RPLCpos09   | Woman<br>HCW099<br>RPLCpos | Woman Pt019<br>RPLCpos | Woman Pt119<br>RPLCpos | Men Pt074<br>RPLCpos | Woman Pt022<br>RPLCpos | Mer<br>RPI |
| H <sub>3</sub> C              | CHEBI:17750            |                     | C[N+](C)(C)CC([O-] |       | Structure not available                  | 889097           | 5.308643602      | 5.339502163       | 5.609366786                | 5.00663743             | 4.679825794            | 5.33411991           | 4.604604231            | 5.169      |
| н <sup>д</sup> ст <u>о</u> .  |                        |                     |                    |       | Critic Critic                            | <b></b> i811223  | 3.267285197      | 3.271734835       | 3.461711921                | 3.424440921            | 3.424736856            | 3.348127133          | 3.467831182            | 3.369      |
| н, м он                       | CHEBI:28300            |                     | NC(CCC(N)=O)C(C    |       |  |                  |                  |                   |                            |                        |                        |                      |                        |            |
| ŇH <sub>2</sub>               |                        |                     |                    |       | nje Star                                 |                  |                  |                   |                            |                        |                        |                      |                        |            |
| H <sub>I</sub> C <sup>S</sup> | CHEBI:16811            |                     | CSCCC(N)C(O)=O     |       | •Y                                       | 022534           | 4.545411209      | 4.535673448       | 4.778082405                | 4.957507954            | 4.912036478            | 4.890949777          | 4.694373625            | 4.802      |
| Q                             |                        |                     |                    |       | о сн.                                    |                  |                  |                   |                            |                        |                        |                      |                        |            |
|                               |                        |                     |                    |       | R1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | н,<br>н, !670279 | 2.929480823      | 2.802480293       | 3.436838655                | 2.919315525            | 3.004794026            | 2.759335266          | 3.009196088            | 3.02       |



| FTP Download Aspera Download  |                              |
|---|------------------------------|
| ) ISA METADATA  | <ul> <li>Download</li> </ul> |
| Q search meta data  |                              |
| a_MTBLS1987_LC-MS_negative_hilic_metabolite_profiling.txt September 08 2020 08:33:26                    | ٥                            |
| a_MTBLS1987_LC-MS_negative_reverse-phase_metabolite_profiling.txt September 08 2020 08:32:52            | ٥                            |
| a_MTBLS1987_LC-MS_positive_hilic_metabolite_profiling.txt September 08 2020 08:39:47                    | ٥                            |
| a_MTBLS1987_LC-MS_positive_reverse-phase_metabolite_profiling.txt September 08 2020 08:57:54            | ٥                            |
| 🗆 📄 i_Investigation.txt September 08 2020 14:10:11  | ٥                            |
| m_MTBLS1987_LC-MS_negative_hilic_metabolite_profiling_v2_maf.tsv December 06 2020 12:20:17              | ٥                            |
| 🗆 📄 m_MTBLS1987_LC-MS_negative_reverse-phase_metabolite_profiling_v2_maf.tsv September 08 2020 08:50:13 | ٥                            |
| m_MTBLS1987_LC-MS_positive_hilic_metabolite_profiling_v2_maf.tsv September 08 2020 08:50:38             | ٥                            |
| 🗆 📄 m_MTBLS1987_LC-MS_positive_reverse-phase_metabolite_profiling_v2_maf.tsv September 08 2020 09:01:31 | ٥                            |
| s_MTBLS1987.txt         December 06 2020 12:45:46   | ٥                            |
| XAW / DERIVED FILES   |                              |
| Q search raw files  |                              |
| HILICneg September 07 2020 15:54:36   | ٥                            |
| HILICpos September 07 2020 15:53:13   | ٥                            |
| RPLCneg September 16 2020 22:14:29  | ٥                            |
| RPLCpos September 07 2020 16:33:31  | ۵                            |



# Acknowledgements

#### **EMBL-EBI Metabolomics team**

- Claire O'Donovan (PI)
- Pamela Pruski
- Mark Williams
- Felix Amaladoss
- Callum Martin
- Thomas Payne



The Metabolomics team's activities were supported in the past year by the above funding bodies



## **Questions:**

- Do you submit your data to public repositories? MetaboLights/Metabolomics Workbench/MetaboBank (and PRIDE)
- Do you reuse public mass spectrometry studies?
- a) Metadata
- b) Raw data
- c) Both



## **Questions:**

• What data do you think we are missing?

 What tools do you use when manipulating/viewing the data? – either your internal data or the public data

• What tools would you like available with the repositories?

