

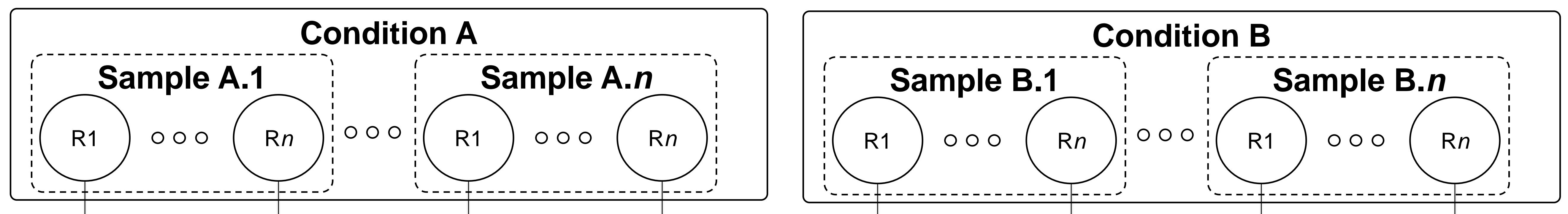
# Using S2P to process emPAI quantitation results obtained from MS/MS Mascot searches

H. López-Fernández<sup>1,2,3</sup>, J.E. Araújo<sup>3</sup>, H.M. Santos<sup>3</sup>, A. Fernández-Lodeiro<sup>3</sup>, C. Fernández-Lodeiro<sup>3</sup>, J. Djafari<sup>3</sup>, J. Coelho<sup>3</sup>, G. Pinto<sup>3</sup>, N. Morais<sup>3</sup>, G. Marcelo<sup>3</sup>, J. Machado<sup>3</sup>, S. Jorge<sup>3</sup>, J. Prates<sup>3</sup>, E. Oliveira<sup>3</sup>, J. Fernández-Lodeiro<sup>3</sup>, C. Lodeiro-Espino<sup>3</sup>, J. L. Capelo-Martínez<sup>3</sup>

<sup>1</sup>SING Research Group, Escuela Superior de Ingeniería Informática, University of Vigo, Edificio Politécnico, Campus Universitario As Lagoas s/n, 32004 Ourense, Spain.

<sup>2</sup>CINBIO - Centro de Investigaciones Biomédicas, University of Vigo, Campus Universitario Lagoas-Marcosende, 36310, Vigo, Spain.

<sup>3</sup>BIOSCOPE Group, UCIBIO-REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Portugal.

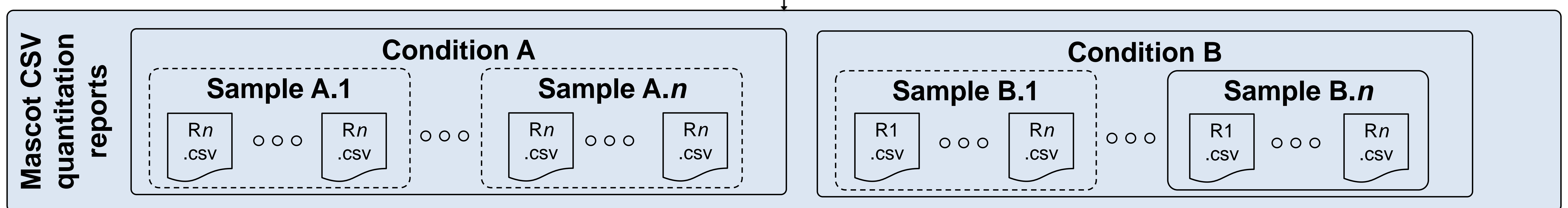


LC-MS/MS Analysis

Raw data

Mascot search engine

The Mascot search engine creates one CSV report file for each experiment replicate, containing the protein identifications along with their corresponding emPAI quantitation values.



S2P is used to process the Mascot CSV reports, allowing users to normalize the emPAI quantitation values, and creating the three different visualizations shown below.

S2P software

Sample	Replicate	Protein	Quantification	Normalized quantification	Protein mass
Ecoli_250ng_1	1	Thiol peroxidase	5.00	0.01	1.48
Ecoli_250ng_1	1	Aspartate-semialdehyde...	1.39	0.00	0.97
Ecoli_250ng_1	1	Phosphoglycerate kinase	2.01	0.00	0.53
Ecoli_250ng_1	1	NADP-specific glutamat...	1.21	0.00	0.32
Ecoli_250ng_1	1	Uncharacterized protein...	2.25	0.00	0.60
Ecoli_250ng_1	1	Alkyl hydroperoxide red...	3.37	0.00	0.89
Ecoli_250ng_1	1	30S ribosomal protein S4	3.99	0.00	1.06
Ecoli_250ng_1	1	Aspartate aminotransfe...	2.32	0.00	0.61
Ecoli_250ng_1	1	Elongation factor G	1.17	0.00	0.31
Ecoli_250ng_1	1	Serine hydroxymethyltra...	1.73	0.00	0.46
Ecoli_250ng_1	1	Cysteine synthase A	0.72	0.00	0.19
Ecoli_250ng_1	1	Metal-binding protein yo...	1.64	0.00	0.43
Ecoli_250ng_1	1	Isocitrate lyase	2.25	0.00	0.60
Ecoli_250ng_1	1	Malate dehydrogenase	2.25	0.00	0.60
Ecoli_250ng_1	1	Malate dehydrogenase	2.25	0.00	0.60
Ecoli_250ng_1	1	Transaldolase B	1.95	0.00	0.52
Ecoli_250ng_1	1	Bifunctional aspartokina...	0.67	0.00	0.18
Ecoli_250ng_1	1	Trigger factor	1.78	0.00	0.47
Ecoli_250ng_1	1	Pyruvate dehydrogenas...	1.04	0.00	0.28
Ecoli_250ng_1	1	6-phosphogluconate de...	1.43	0.00	0.38
Ecoli_250ng_1	1	50S ribosomal protein L10	3.61	0.00	0.96
Ecoli_250ng_1	1	LeuT/Val-binding protein	1.04	0.00	0.28
Ecoli_250ng_1	1	2	1.52	0.00	0.40
Ecoli_250ng_1	1	Aspartate carbamoyltra...	1.75	0.00	0.46
Ecoli_250ng_1	1	Aldehyde-alcohol dehydi...	0.87	0.00	0.23
Ecoli_250ng_1	1	Aderylosuccinate synth...	1.10	0.00	0.29
Ecoli_250ng_1	1	30S ribosomal protein S1	1.12	0.00	0.30
Ecoli_250ng_1	1	50S ribosomal protein L9	9.48	0.01	2.51
Ecoli_250ng_1	1	High-affinity zinc uptake...	1.08	0.00	0.28
Ecoli_250ng_1	1	ATP synthase subunit b...	1.31	0.00	0.35
Ecoli_250ng_1	1	Chaperone protein dnaK	0.76	0.00	0.19

E [mean normalized PQ] (#2)	E [std normalized PQ] (#2)	E [rsd normalized PQ] (#2)	E [mean mass] (#2)	E [std mass] (#2)	E [POP] (#2)	#1 vs. #2 [p-value]	#1 vs. #2 [t-value]
1.64e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0010098079...	0.30698162682...
3.03e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0097665503...	0.43345696923...
4.32e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0122959583...	0.43345696923...
8.21e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0162320782...	0.43345696923...
5.32e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0185041462...	0.43345696923...
1.69e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0262686767...	0.43345696923...
1.03e-02	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0281445951...	0.43345696923...
8.48e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.029356956322...	0.43345696923...
5.19e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0296229302...	0.43345696923...
1.12e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.0347545117...	0.43345696923...
1.34e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.055615679996...	0.43345696923...
1.98e-02	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.05102079333...	0.43345696923...
1.12e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.05443118802...	0.43345696923...
1.12e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.05511848018...	0.43345696923...
1.34e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.056115679996...	0.43345696923...
2.25e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.05907185340...	0.43345696923...
2.46e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.063569566322...	0.43345696923...
2.16e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.06892252573...	0.43345696923...
1.12e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.07199004907...	0.43345696923...
1.12e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.07316846425...	0.43345696923...
1.80e-02	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.07425510507...	0.43345696923...
3.24e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.07567252462...	0.43345696923...
3.03e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.07738552126...	0.43345696923...
5.01e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.08440160388...	0.43345696923...
8.64e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.08544458129...	0.43345696923...
6.09e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.08696727958...	0.43345696923...
1.17e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.08720146591...	0.43345696923...
7.26e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.09136695485...	0.43345696923...
4.37e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.09252307883...	0.43345696923...
2.16e-04	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.09431727425...	0.43345696923...
2.55e-03	0.00e+00	0.00e+00	0.00e+00	0.00e+00	1/1	0.09703991106...	0.43345696923...

**Replicates table:** a table with one row by each single protein quantitation in the whole dataset. This is the raw quantification data arranged as a table.

**Replicates (by protein) table:** a table where replicates can be compared by protein identification.

Protein	Ecoli_250ng_1_1	Ecoli_250ng_1_1 [Normalized]	Ecoli_250ng_1_2	Ecoli_250ng_1_2 [Normalized]	Ecoli_250ng_1_3
FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase	1.70e-01	1.80e-04	1.70e-01	1.80e-04	1.70e-01
ABC transporter arginine-binding protein 1			2.90e-01	3.07e-04	2.90e-01
Bifunctional protein fobP					
Thioredoxin-1	7.40e-01	7.84e-04	7.40e-01	7.84e-04	7.40e-01
Thioredoxin reductase	1.10e-01	1.17e-04			
Aminacylhistidine dipeptidase	1.40e-01	1.48e-04	1.40e-01	1.48e-04	1.40e-01
Cold-shock DEAD box protein A	2.20e-01	2.33e-04	1.40e-01	5.30e-05	1.40e-01
Glutamate-cysteine ligase			6.00e-02	6.36e-05	
Ribose-phosphate pyrophosphokinase	1.10e-01	1.17e-04	1.10e-01	1.17e-04	2.20e-01
Pyridoxalpyridoxamine 5--phosphate oxidase	1.40e-01	1.48e-04	1.40e-01	1.48e-04	1.40e-01
Uncharacterized protein yhcH	8.00e-02	8.48e-05			
Surfactin/CoA ligase [ADP-forming] subunit alpha	2.60e-01	2.75e-04	2.60e-01	2.75e-04	2.60e-01
Bifunctional purine biosynthesis protein purH	2.00e-01	2.12e-04	5.30e-01	5.62e-04	2.00e-01
Uncharacterized protein yggJ					
Stationary-phase-induced ribosome-associated protein	7.90e-01	8.37e-04	7.90e-01	8.37e-04	7.90e-01
Asparagine synthetase B [glutamine-hydrolyzing]	6.00e-02	6.36e-05	6.00e-02	6.36e-05	6.00e-02
Beta-galactosidase	1.50e-01	1.59e-04	3.00e-02	3.18e-05	
Surfactin/CoA ligase [ADP-forming] subunit alpha	2.60e-01	2.75e-04	2.60e-01	2.75e-04	2.60e-01
Thiol peroxidase	5.60e+00	5.93e-03	5.60e+00	5.93e-03	8.63e+00
Protein ychN					3.00e-01
Inner membrane protein yeaI	7.00e-02	7.42e-05			
Outer membrane protein A	7.50e-01	7.95e-04	9.30e-01	9.85e-04	1.12e+00
Transketolase 1	5.50e-01	5.83e-04	5.50e-01	5.83e-04	6.30e-01
Outer membrane protein C	9.00e-02	9.54e-05			
LPF092 membrane protein yajC	3.20e-01	3.39e-04			
2	1.52e+00	1.61e+03	1.24e+00	1.31e+03	1.24e+00
Dihydrodipicolinate synthase	2.50e-01	2.65e-04	3.90e-01	4.13e-04	2.50e-01
Methionyl-tRNA synthetase	1.50e-01	1.59e-04	2.00e-01	2.12e-04	5.00e-02
Isochorismate synthase entC	6.30e-01	6.68e-04	5.00e-01	5.30e-04	5.00e-01
5	2.30e-01	2.44e-04	1.10e-01	1.17e-04	2.30e-01

**Conditions table:** a table where experimental conditions can be compared by protein identification using the quantitation values. In this table, the t-test results of the comparison between each pair of experimental conditions are provided in order to detect differences in the average quantitation values.

## Acknowledgements

H. López-Fernández is supported by a post-doctoral fellowship from Xunta de Galicia. J. E. Araújo acknowledges the financial support given by the Portuguese Foundation for Science and Technology under doctoral grant number SFRH/BD/109201/2015. J. L. Capelo acknowledges Associação Científica ProteoMass for financial support.

## Try S2P now!

S2P is freely available at <http://www.sing-group.org/s2p>