

C-COMPASS: a user-friendly neural network tool profiles cell compartments at protein and lipid levels

In the format provided by the
authors and unedited

1 **Inventory of Supplementary information**

- 2 - Procedural suggestion for generating a custom marker list

3

4 **Procedural suggestion for generating a custom marker list**

5 In certain cases, it may be necessary to generate a custom list of spatial marker proteins,
6 particularly when reference data for specific tissues, organisms, or biological conditions is
7 unavailable. However, it is possible to derive marker proteins that reflect compartment-specific
8 profile patterns directly from the dataset itself. To do so, we recommend the following pre-
9 processing steps on the gradient data: 1) Impute missing values with zero. 2) Apply MinMax
10 scaling to normalize all protein profiles to a range between 0 and 1.

11 Next, perform hierarchical clustering across the protein dimensions (rows) only, without
12 clustering the fraction dimension (columns). After clustering, identify protein clusters that are
13 enriched for annotations associated with specific organelles or compartments. We recommend
14 using Gene Ontology Cellular Component (GOCC) annotations for this purpose. This
15 enrichment can be assessed manually or via algorithmic approaches that identify clusters with
16 the strongest enrichment for the compartments of interest. If performed manually, we
17 recommend prioritizing clusters with distinctive profile patterns over large cluster sizes.

18 The resulting set of proteins can then be compiled into a marker list, with each protein
19 annotated according to the enriched compartment. To be used in C-COMPASS, this list must
20 contain at least two columns: one serving as an identifier (e.g. gene names), and one indicating
21 the assigned compartment annotation. A comprehensive list of marker proteins used in our
22 experiments can be found in the supplementary material.

23 Below, we provide a list of publications in which this strategy has been successfully applied to
24 derive compartment-specific marker proteins:

25 Klingelhuber, F. et al. A spatiotemporal proteomic map of human adipogenesis. *Nat Metab* **6**,
26 861-879 (2024)

27 Krahmer, N. et al. Organellar Proteomics and Phospho-Proteomics Reveal Subcellular
28 Reorganization in Diet-Induced Hepatic Steatosis. *Dev Cell* **47**, 205-221 e207 (2018)