LION/web: a web-based ontology enrichment tool for lipidomic data analysis

Martijn R. Molenaar[§], Aike Jeucken[§], Tsjerk A. Wassenaar[‡], Chris H. A. van de Lest[§], Jos F. Brouwers[§], J. Bernd Helms*[§]

[§] *Department of Biochemistry and Cell Biology, Faculty of Veterinary Medicine, Utrecht University, 3584 CM, Utrecht, The Netherlands*

[‡] *Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands*

* *To whom correspondence should be addressed: J.B.H (J.B.Helms@uu.nl)*

SUMMARY

A major challenge for lipidomic analyses is the handling of the large amounts of data and the translation of results to interpret the involvement of lipids in biological systems. We built a new lipid ontology (LION) that associates over 50,000 lipid species to biophysical, chemical and cell biological features. By making use of enrichment algorithms, we used LION to develop a web-based interface (LION/web, www.lipidontology.com) that allows identification of lipid-associated terms in lipidomes. LION/web was validated by analyzing a lipidomic dataset derived from well-characterized sub-cellular fractions of RAW 264.7 macrophages. Comparison of isolated plasma membranes with the microsomal fraction showed a significant enrichment of relevant LION-terms including 'plasma membrane', 'headgroup with negative charge, 'glycerophosphoserines', 'above average bilayer thickness', and 'below average lateral diffusion'. A second validation was performed by analyzing the membrane fluidity of CHO cells incubated with arachidonic acid. An increase in membrane fluidity was observed both experimentally by using pyrene decanoic acid and by using LION/web, showing significant enrichment of terms associated with high membrane fluidity ('above average', 'very high' and 'high lateral diffusion', and 'below average transition temperature'). The results demonstrate the functionality of LION/web, which is freely accessible in a platform-independent way.

KEYWORDS

lipidomics; lipids; membrane biology; lipid ontology; LION; LION-term enrichment analysis; membrane biology; web-tool; data analysis; LION/web

25 **MAIN TEXT**

26 The comprehensive study of lipids, also termed lipidomics, is gaining momentum.
27 Instrumentation is becoming increasingly more sensitive, precise and fast, and the use of
28 lipidomics to address key questions in membrane biology has become widespread. As a result,
29 datasets are rapidly increasing both in terms of size and complexity. Due to a lack of methods
30 to perform global and in-depth data mining, lipidomic research tends to focus on individual
31 lipid classes or lipid species. A common approach in other 'omics' disciplines to reduce
32 complexity is the use of ontologies *e.g.,* Gene Ontology (Ashburner et al., 2000), Chemical
33 Entities of Biological Interest ontology (Degtyarenko et al., 2008), combined with statistical
34 tools to determine terms of interest.
35 Although lipid structure is closely related to lipid function, it is currently impossible to
36 associate properties of individual lipids with complex lipid mixtures of cellular lipidomes.
37 Examples of biophysical properties that play an important role in membrane biology are
38 numerous and include membrane thickness as driving force in the sub-cellular localization of
39 proteins (Sharpe et al., 2010), membrane fluidity regulating bacterial survival (Inda et al.,
40 2014), membrane heterogeneity in cellular signaling (Sezgin et al., 2017), intrinsic curvature
41 of lipids as key player in lipid droplet biogenesis (Ben M'barek et al., 2017; Thiam et al.,
42 2013) or COPI coat disassembly (Bigay et al., 2003), and net charge of membranes as a
43 determinant in lipid-protein interactions (Enkavi et al., 2017).
44 Here, we aim to provide a lipid ontology database and complementary enrichment analysis
45 tool that (i) contains chemical and biophysical information of lipid species, (ii) is platform
46 independent and compatible with routine mass spectrometry-based lipid analysis, (iii) can be
47 used by researchers without computer programming skills, and (iv) is freely available to the
48 scientific community.

49 We constructed an ontology database called LION that links over 50,000 lipid species with
50 four major branches: 'lipid classification' (Fahy et al., 2009), 'chemical and physical
51 properties' (fatty acid length and unsaturation, headgroup charge, intrinsic curvature,

membrane fluidity, bilayer thickness), 'function', and 'subcellular localization' (as described in literature). The resulting database contains more than 250,000 connections ('edges'), providing a detailed system for in-depth annotation of lipids. An example of all LION-terms associated with a single phosphatidylserine (PS) lipid species, PS(34:2), is depicted in **Figure S1**.

An important feature of LION is the association of lipid species with biophysical properties. We made use of experimental data (Marsh, 2010) and data obtained by coarse-grain molecular dynamics simulation (CG-MD) (Wassenaar et al., 2015), each providing distinct biophysical properties. These data were used to estimate the biophysical properties of all related lipids in the LION-database by multiple regression analysis.

The regression models were validated in two ways. First, we performed leave-one-out cross-validations (LOOCV) of all three models (**Fig. S2 A-C)**, showing satisfactory agreement between determined and predicted values. Second, we compared two properties closely associated with membrane fluidity: 'transition temperature' (from experimental datasets) and 'lateral diffusion' (from the CG-MD datasets) **(Fig. S2 D)**. As expected, lipids with low transition temperatures were predicted to have high lateral diffusion values at a defined simulation temperature and vice versa.

Subsequently, all numerical datapoints for each biophysical property were categorized into five pre-defined groups ('very low', 'low', 'average', 'high', 'very high'). The limits of each group were determined based on the presence of lipid species reported in four lipidomics publications (Andreyev et al., 2010; Haraszti et al., 2016; Köberlin et al., 2015; Lin et al., 2017). These values were subsequently used to categorize all applicable lipid species present in LION (**Fig. S2 E**).

Next, we used LION as a basis to build an ontology enrichment tool that facilitates reduction of lipidome complexities in an unbiased manner. We made use of an adapted version of 'topGO', an R-package designed for enrichment analysis of GO-terms (Alexa and Rahnenfuhrer, 2017). Subsequently, we designed a web-tool with R-package Shiny

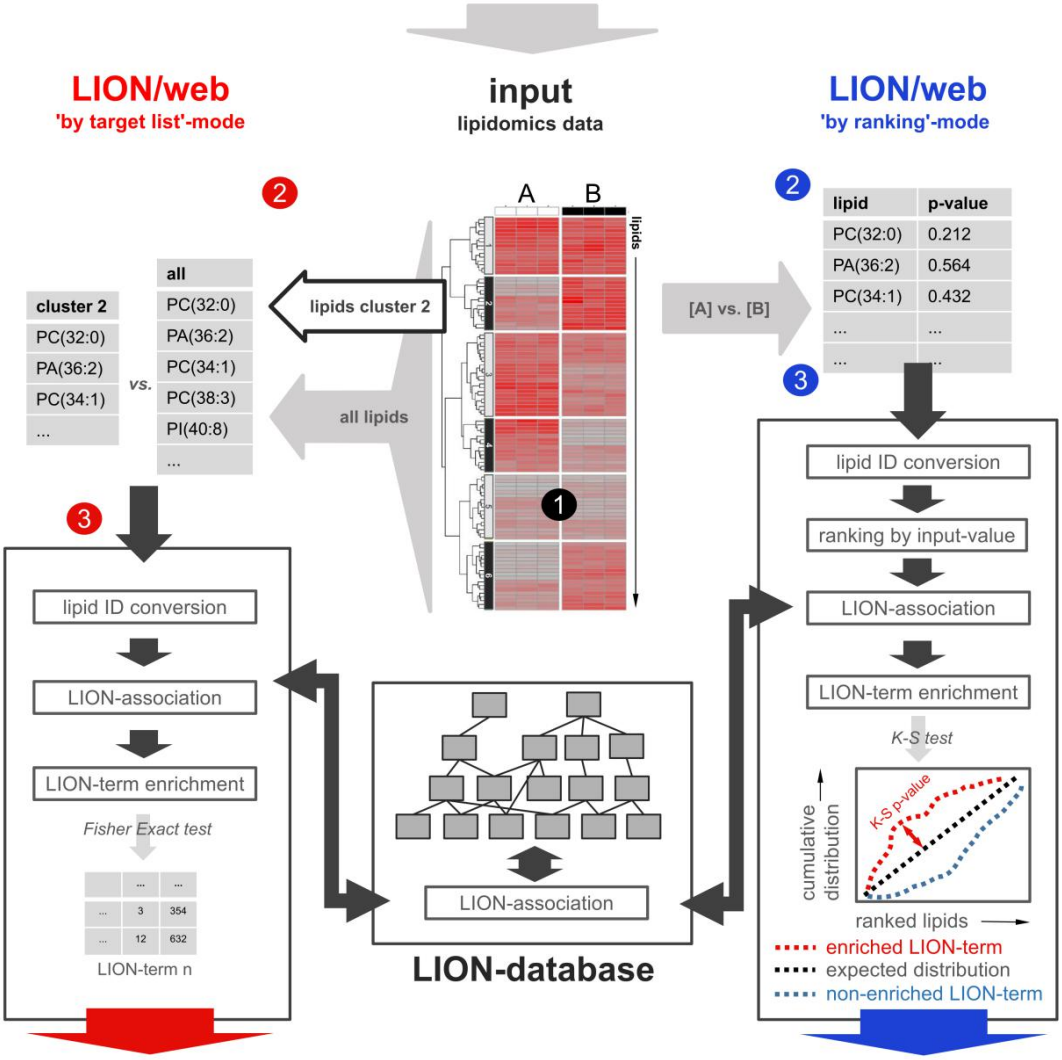
79 ('LION/web', www.lipidontology.com) that offers an intuitive user-interface and supports two

80 major workflows **(Fig. 1 and Note S1)**: enrichment analysis of a subset of lipids of interest

81 ('by target list') and enrichment analysis performed on a complete and ranked list of lipids

82 ('by ranking', referred to as 'SAFE' in the context of genes (Barry et al., 2005)).

83



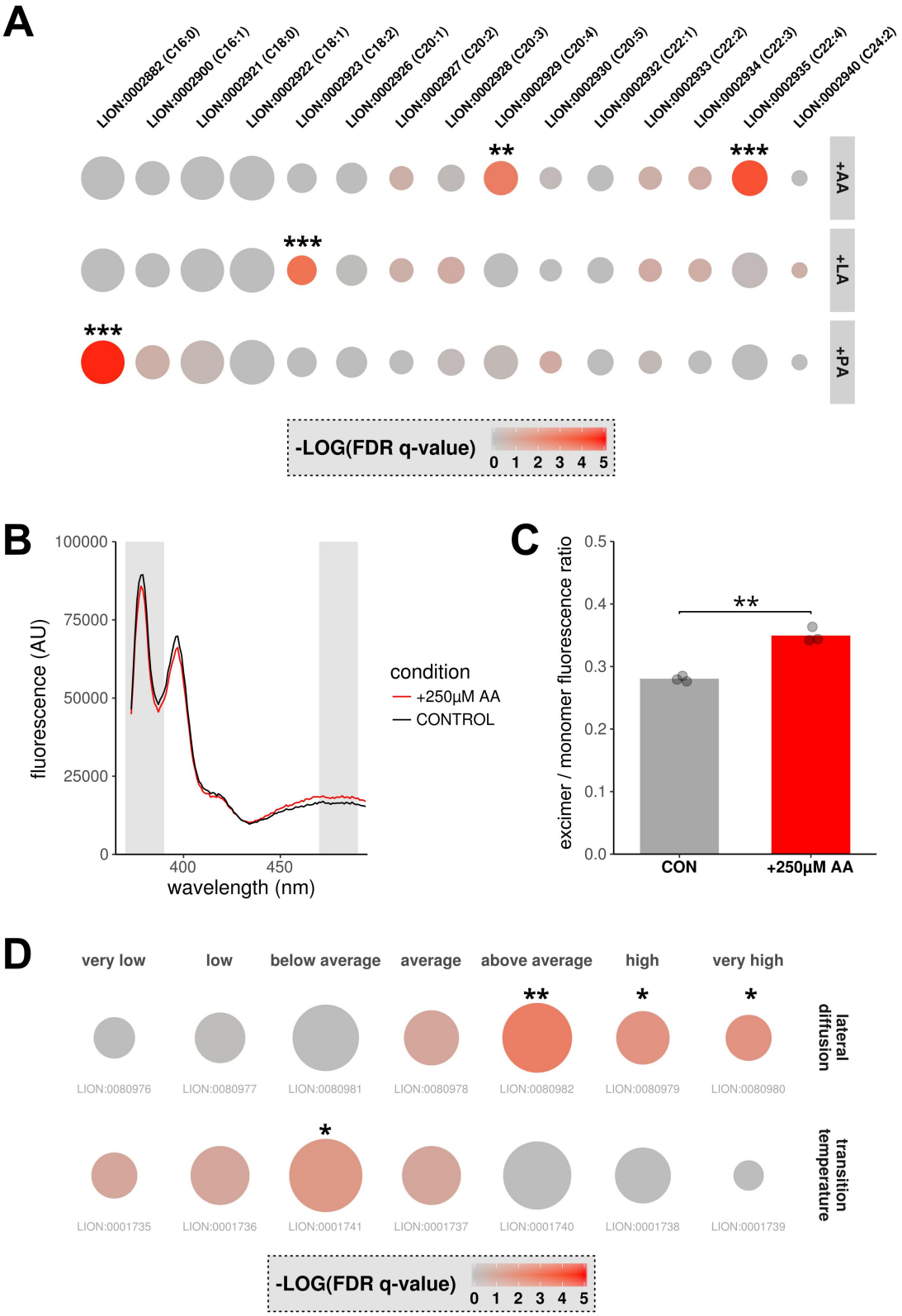
84

85 **Figure 1. Enrichment analysis approaches supported by LION/web.** A lipidomics dataset
86 containing lipid identifiers and abundances derived from two or more conditions (❶) can be
87 processed in two ways by LION/web. In the 'by target list'-mode (left, ❷), a subset of lipids
88 (*e.g.*, derived from thresholding or clustering) is compared to the total set of lipids. After
89 standardization of lipid nomenclature (❸), applicable LION-terms are associated and
90 assessed for enrichment in the subset by Fisher's exact statistics. Alternatively, in the 'by
91 ranking'-mode, input lipids are ranked by the provided values ('local' statistics). By default, *P*
92 values from one-tailed t-tests are used (❷). After ranking, lipid nomenclature is standardized

93 (❸). Applicable LION-terms are subsequently associated to the dataset and distributions are
94 compared to a uniform distribution by 'global' statistics (here, Kolmogorov–Smirnov tests).
95 Calculated *P* values of LION-terms from both approaches are corrected for multiple testing
96 (Benjamini-Hochberg).

97

98 To test the functionality of LION/web, we made use of a previously published and well
99 characterized dataset containing lipidomics data from several sub-cellular fractions of RAW
100 264.7 macrophages (Andreyev et al., 2010). First, we re-normalized the dataset by expressing
101 all lipid species as fraction of the total amount of lipid per sample. LION/web was able to
102 reformat (for information about input conventions, see **Note S2**) and match the vast majority
103 (>97%) of the submitted lipids. Subsequently, we compared the isolated plasma membrane
104 (PM) with the endoplasmic reticulum (ER) fraction from non-stimulated macrophages and
105 assessed all LION-terms for enrichment **(Fig. S3)**. In good agreement with current
106 descriptions of the selected organelles (Holthuis and Menon, 2014; van Meer et al., 2008),
107 significant enriched LION-terms included terms associated with chemical descriptions (*e.g.,*
108 'glycerophosphoserines', 'headgroup with negative charge', 'phosphosphingolipids'),
109 biological features ('plasma membrane') and biophysical properties (*e.g.,* 'above average
110 bilayer thickness', 'below average lateral diffusion', 'very low lateral diffusion', 'very high
111 bilayer thickness', 'neutral intrinsic curvature'). LION/web also reported the significant
112 enrichment of 'very high transition temperature', which is in line with the (very) low lateral
113 diffusion terms (see also **Fig. S2 D**). Also the term 'very low transition temperature' was
114 reported to be significantly enriched. Inspection of the lipid species responsible for the
115 LION-term 'very low transition temperature' revealed the presence of lipids that all contain
116 polyunsaturated fatty acids (PUFAs) with at least four unsaturations. This may be a
117 macrophage-specific phenomenon, related to their involvement in inflammation (Calder,
118 2015).
119 To further validate LION/web, we used two different experimental approaches. First, we
120 investigated the enrichment of LION-terms associated with chemical features that can be



121

122 **Figure 2. LION-term enrichment and membrane fluidity of CHO-k1 cells.** CHO-k1 cells
123 were incubated overnight with PA, LA or AA (100 µM, complexed to BSA) **(A)** or with AA (250
124 µM, complexed to BSA) **(B-D)**. All incubations were performed in triplicate. For control
125 incubations, cells were incubated with fatty-acid free BSA. **(A,D)** After extraction and

126 lipidomics profiling by LC-MS/MS, enrichment analyses of the conditions of interest versus
127 control incubations were performed by LION/web of **(A)** LION-terms indicating the presence of
128 selected fatty acids or **(D)** LION-terms indicating the degree of membrane fluidity. Dot sizes in
129 the dot plots are scaled to the number of associated lipids; colors are scaled to the level of
130 enrichment. **(B,C)** After incubation, fluorescence emission spectra between 370 and 500 nm of
131 lysates containing pyrenedecanoic acid (PDA) were measured in triplicate **(B)**. Fluorescence
132 spectra examples of either control (black) or AA-stimulated lysates (red). Gray shades indicate
133 monomer and excimer fluorescence filters. **(C)** Mean ratios (bar) and individual datapoints
134 (dots) of excimer over monomer fluorescence (representative data of three independent
135 experiments). Statistical significance was determined by Student's two-tailed t-test. **(A,C,D)** * P
136 < 0.05, ** P < 0.01, *** P < 0.001.

137

138 easily incorporated into lipids (*e.g.,* fatty acids as building blocks). To this end, CHO-k1 cells
139 were incubated overnight in the presence of palmitic acid (PA), linoleic acid (LA) or
140 arachidonic acid (AA) complexed to bovine serum albumin (BSA). Subsequently, lipids were
141 analysed by LC-MS/MS **(Data S1** and **Fig. S4)** and submitted to LION/web ('by
142 ranking'-mode). LION/web offers the option to limit analysis to specific terms of interest.
143 After pre-selection of LION-terms that indicate the presence of fatty acids as lipid building
144 blocks, LION/web reported the significant enrichment of the respective fatty acid in the three
145 different conditions **(Fig. 2 A** and **Data S2)**.
146 Second, to investigate the enrichment of biophysical LION-terms, we incubated CHO-k1 cells
147 with arachidonic acid (AA). This procedure is known to increase membrane fluidity (Yang et
148 al., 2011). After incubation, the membrane fluidity properties of the samples were analyzed
149 both experimentally and by LION/web. Membrane fluidity was experimentally assessed using
150 pyrene decanoic acid (PDA). This fluorescent probe can exist as monomer or excimer,
151 resulting in a shift of its emission spectrum. The ratio of excimer over monomer fluorescence
152 is proportional to the degree of membrane fluidity (Eisinger and Scarlata, 1987). To this end,
153 fluorescence spectra of lysates from cells incubated with or without AA were measured **(Fig.
154 2 B)**. As expected, the ratio of excimer/monomer forms of PDA revealed a significant
155 increase in membrane fluidity of lysates in the presence of AA **(Fig. 2 C)**. For parallel
156 LION/web analysis of membrane fluidity properties, lipids were extracted from the same

samples and analysed by LC-MS/MS **(Data S3** and **Fig. S5)**. LION contains two sets of terms associated with membrane fluidity: ‘transition temperature’ and ‘lateral diffusion’. Accordingly, LION/web was set to limit enrichment analyses to these sets, after which the lipidomic data were analyzed (‘by ranking’ mode). In line with the experimentally measured increase in membrane fluidity, terms associated with high membrane fluidity ('above average’, 'very high’ and 'high lateral diffusion’, and 'below average transition temperature’) were significantly enriched in cells that had been treated with AA **(Fig. 2 D** and **Data S4)**. Taken together, we have presented a new ontology called LION that enables flexible annotation of lipid species and that covers most commonly found lipid classes and fatty acid distributions. Furthermore, it combines the well-established lipid class hierarchy from LIPIDMAPS with biophysical data that were not previously available. To explore lipid datasets in an unbiased manner, we built an online web-tool that does not require knowledge of programming languages. We believe that this lipid database and associated web-tool bridges the gap between lipidomics and cell biology by revealing patterns that are of biological interest.

ACKNOWLEDGEMENTS

We thank Xin He, PhD, for providing and supporting the topONTO R-package. We thank Jeroen W.A. Jansen for the excellent technical assistance with the lipidomics experiments.

AUTHOR CONTRIBUTIONS

M.R.M. and J.B.H. conceived the project. M.R.M. developed LION, LION/web and performed the experiments. A.J. tested and suggested improvements for LION/web. C.H.A.v.d.L. and T.A.W. contributed to the regression models and statistical concepts. C.H.A.v.d.L. and J.F.B. contributed to the lipidomics data processing and analysis. M.R.M. and J.B.H. wrote the manuscript.

183

184 **COMPETING FINANCIAL INTERESTS**

185 The authors declare no competing financial interests.

186

187 **REFERENCES**

188 Aimo, L., Liechti, R., Hyka-Nouspikel, N., Niknejad, A., Gleizes, A., Götz, L., Kuznetsov, D.,
189 David, F.P.A., van der Goot, F.G., Riezman, H., et al. (2015). The SwissLipids
190 knowledgebase for lipid biology. Bioinforma. Oxf. Engl. *31*, 2860–6.

191 Alexa, A., and Rahnenfuhrer, J. (2017). Gene set enrichment analysis with topGO.
192 Bioconductor.

193 Andreyev, A.Y., Fahy, E., Guan, Z., Kelly, S., Li, X., McDonald, J.G., Milne, S., Myers, D.,
194 Park, H., Ryan, A., et al. (2010). Subcellular organelle lipidomics in TLR-4-activated
195 macrophages. J. Lipid Res. *51*, 2785–97.

196 Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P.,
197 Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification
198 of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–9.

199 Barry, W.T., Nobel, A.B., and Wright, F.A. (2005). Significance analysis of functional
200 categories in gene expression studies: a structured permutation approach. Bioinforma. Oxf.
201 Engl. *21*, 1943–9.

202 Ben M'barek, K., Ajjaji, D., Chorlay, A., Vanni, S., Forêt, L., and Thiam, A.R. (2017). ER
203 Membrane Phospholipids and Surface Tension Control Cellular Lipid Droplet Formation.
204 Dev. Cell *41*, 591–604.e7.

205 Bigay, J., Gounon, P., Robineau, S., and Antonny, B. (2003). Lipid packing sensed by
206 ArfGAP1 couples COPI coat disassembly to membrane bilayer curvature. Nature *426*, 563–6.

207 Bligh, E.G., and Dyer, W.J. (1959). A rapid method of total lipid extraction and purification.
208 Can. J. Biochem. Physiol. *37*, 911–917.

209 Calder, P.C. (2015). Marine omega-3 fatty acids and inflammatory processes: Effects,
210 mechanisms and clinical relevance. Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids
211 *1851*, 469–484.

212 Degtyarenko, K., de Matos, P., Ennis, M., Hastings, J., Zbinden, M., McNaught, A.,
213 Alcántara, R., Darsow, M., Guedj, M., and Ashburner, M. (2008). ChEBI: a database and
214 ontology for chemical entities of biological interest. Nucleic Acids Res. *36*, D344–50.

215 Eisinger, J., and Scarlata, S.F. (1987). The lateral fluidity of erythrocyte membranes
216 temperature and pressure dependence. Biophys. Chem. *28*, 273–281.

217 Enkavi, G., Mikkolainen, H., Güngör, B., Ikonen, E., and Vattulainen, I. (2017). Concerted
218 regulation of npc2 binding to endosomal/lysosomal membranes by
219 bis(monoacylglycero)phosphate and sphingomyelin. PLOS Comput. Biol. *13*, e1005831.

220 Fahy, E., Subramaniam, S., Murphy, R.C., Nishijima, M., Raetz, C.R.H., Shimizu, T., Spener,
221 F., van Meer, G., Wakelam, M.J.O., and Dennis, E.A. (2009). Update of the LIPID MAPS
222 comprehensive classification system for lipids. J. Lipid Res. *50 Suppl*, S9–14.

223 Haraszti, R.A., Didiot, M.-C., Sapp, E., Leszyk, J., Shaffer, S.A., Rockwell, H.E., Gao, F.,
224 Narain, N.R., DiFiglia, M., Kiebish, M.A., et al. (2016). High-resolution proteomic and
225 lipidomic analysis of exosomes and microvesicles from different cell sources. J. Extracell.
226 Vesicles *5*, 32570.

227 Holthuis, J.C.M., and Menon, A.K. (2014). Lipid landscapes and pipelines in membrane
228 homeostasis. Nature *510*, 48–57.

229 Inda, M.E., Vandenbranden, M., Fernández, A., de Mendoza, D., Ruysschaert, J.-M., and
230 Cybulski, L.E. (2014). A lipid-mediated conformational switch modulates the thermosensing
231 activity of DesK. Proc. Natl. Acad. Sci. U. S. A. *111*, 3579–84.

232 Köberlin, M.S., Snijder, B., Heinz, L.X., Baumann, C.L., Fauster, A., Vladimer, G.I., Gavin,
233 A.C., and Superti-Furga, G. (2015). A Conserved Circular Network of Coregulated Lipids
234 Modulates Innate Immune Responses. Cell *162*, 170–183.

235 de Kroon, A.I.P.M., Rijken, P.J., and De Smet, C.H. (2013). Checks and balances in
236 membrane phospholipid class and acyl chain homeostasis, the yeast perspective. Prog. Lipid
237 Res. *52*, 374–94.

238 Liebisch, G., Vizcaíno, J.A., Köfeler, H., Trötzmüller, M., Griffiths, W.J., Schmitz, G.,
239 Spener, F., and Wakelam, M.J.O. (2013). Shorthand notation for lipid structures derived from
240 mass spectrometry. J. Lipid Res. *54*, 1523–30.

241 Lin, L., Ding, Y., Wang, Y., Wang, Z., Yin, X., Yan, G., Zhang, L., Yang, P., and Shen, H.
242 (2017). Functional lipidomics: Palmitic acid impairs hepatocellular carcinoma development
243 by modulating membrane fluidity and glucose metabolism. Hepatol. Baltim. Md *66*, 432–448.

244 Marrink, S.J., de Vries, A.H., and Mark, A.E. (2004). Coarse Grained Model for
245 Semiquantitative Lipid Simulations. J. Phys. Chem. B *108*, 750–760.

246 Marsh, D. (2010). Structural and thermodynamic determinants of chain-melting transition
247 temperatures for phospholipid and glycolipids membranes. Biochim. Biophys. Acta *1798*,
248 40–51.

249 van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are
250 and how they behave. Nat. Rev. Mol. Cell Biol. *9*, 112–24.

251 Sezgin, E., Levental, I., Mayor, S., and Eggeling, C. (2017). The mystery of membrane
252 organization: composition, regulation and roles of lipid rafts. Nat. Rev. Mol. Cell Biol. *18*,
253 361–374.

254 Sharpe, H.J., Stevens, T.J., and Munro, S. (2010). A comprehensive comparison of
255 transmembrane domains reveals organelle-specific properties. Cell *142*, 158–69.

256 Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS:
257 processing mass spectrometry data for metabolite profiling using nonlinear peak alignment,
258 matching, and identification. Anal. Chem. *78*, 779–787.

259 Thiam, A.R., Farese, R.V., and Walther, T.C. (2013). The biophysics and cell biology of lipid
260 droplets. Nat. Rev. Mol. Cell Biol. *14*, 775–86.

261 Wächter, T., and Schroeder, M. (2010). Semi-automated ontology generation within
262 OBO-Edit. Bioinforma. Oxf. Engl. *26*, i88–96.

263 Wassenaar, T.A., Ingólfsson, H.I., Böckmann, R.A., Tieleman, D.P., and Marrink, S.J. (2015).
264 Computational lipidomics with insane: A versatile tool for generating custom membranes for
265 molecular simulations. J. Chem. Theory Comput. *11*, 2144–2155.

266 Xia, J., Sinelnikov, I.V., Han, B., and Wishart, D.S. (2015). MetaboAnalyst 3.0–making
267 metabolomics more meaningful. Nucleic Acids Res. *43*, W251–7.

268 Yang, X., Sheng, W., Sun, G.Y., and Lee, J.C.M. (2011). Effects of fatty acid unsaturation
269 numbers on membrane fluidity and α-secretase-dependent amyloid precursor protein
270 processing. Neurochem. Int. *58*, 321–9.

271

Methods

Creation of lipid ontology (LION). We built an ontology database that connects over 50,000 lipid species to the following four major branches: 'lipid classification', 'function', 'cellular localization' and 'physical or chemical properties'. For readability, a term is included at the top of each branch to indicates the nature of a LION-branch. These 'category' terms are distinguished from other LION-terms with an ID containing the prefix 'CAT'.

The classification system is based on the LIPIDMAPS classification (Fahy et al., 2009). Downstream, we added an extra level between classes and species to enable mapping of lipid identifiers that lack detailed structural information. This concept is also used in the Swiss Lipids system (Aimo et al., 2015). The branch 'function' comprises three subcategories: 'lipid component' (associated with lipids that are primary regarded as structural component of lipid bilayers), 'lipid-mediated signaling' (lipids that have been implicated in signaling) and 'lipid-storage' (lipids that are associated with storage, primarily in lipid droplets). In the category 'cellular localization', lipid classes that are enriched in particular cellular organelles are linked to their corresponding organelle terms (Holthuis and Menon, 2014; van Meer et al., 2008). The branch 'physical or chemical properties' comprises a number of subcategories. First, a number of chemical descriptions ('contains fatty acid', 'fatty acid unsaturation', 'fatty acid length' and 'type by bond') was inferred from the species names. Second, data about 'intrinsic curvature' (de Kroon et al., 2013; Thiam et al., 2013) were categorized into either negative, neutral or positive curvature. As data on species-level are limited, curvature was assumed to be predominantly headgroup-dependent and fatty acid composition was neglected. The third subcategory, 'charge headgroup', was divided into three groups based on structural data: 'negative', 'positive/zwitter-ion' and 'neutral' (Fahy et al., 2009). This last term comprises also lipids lacking a headgroup. The fourth subcategory in 'physical or chemical properties' is 'chain-melting transition temperature'. This property is derived from a number of sources, comprehensively reviewed by Marsh (Marsh, 2010). This dataset covers a range of lipid classes in both glycerophospholipids (PC, PE, PG, PA, PS) and sphingolipids (SM). We

299 made use of multiple regression analysis with lipid class, fatty acid length and unsaturation as
300 predictors to facilitate data extrapolation to previously unreported lipid species. The obtained
301 model was validated by leave-one-out cross-validation (LOOCV). Briefly, one datapoint from
302 the dataset was taken out, after which the model was rebuilt with the remaining points as
303 training set. Subsequently, the selected datapoint was used as validation sample. This
304 procedure was repeated for all the datapoints (**Fig. S2 C**). Next, values predicted by the
305 obtained model of all applicable lipid species present in LION were divided into quintiles
306 with limits based on four reported lipidomics datasets (Andreyev et al., 2010; Haraszti et al.,
307 2016; Köberlin et al., 2015; Lin et al., 2017) and categorized into five representative classes:
308 'very low', 'low', 'average', 'high' or 'very high' chain-melting transition temperature (a
309 flow-chart of this procedure is depicted in **Fig. S2 E**).
310 In addition to these experimental data sets, we also used data (Wassenaar et al., 2015) that
311 was obtained by coarse grain molecular dynamics simulation (MARTINI force-field (Marrink
312 et al., 2004)) and which includes membrane properties 'bilayer thickness' and 'lateral
313 diffusion'. The dataset contains lipids from five common classes of glycerophospholipids (PC,
314 PS, PG, PA, PE), but lacks sphingolipids and sterols. By definition, coarse-grained lipids
315 represent a range of structures. To be able to use the dataset in the ontology system, names of
316 coarse-grained lipids were translated into their representing counterparts. Subsequently, lipid
317 properties were extrapolated to the entire database by multiple regression analysis models
318 (with lipid class, fatty acid length and unsaturation as predictors) and validated by LOOCV
319 (**Fig. S2 A-B**). We followed the same procedure as used for transition temperatures;
320 extrapolated results for both properties were divided into quintiles (based on values of
321 reported datasets (Andreyev et al., 2010; Haraszti et al., 2016; Köberlin et al., 2015; Lin et al.,
322 2017), predicted by our models) and categorized into five representative classes: 'very low',
323 'low', 'average', 'high' or 'very high'.
324 The initial structure of LION was build with OBOEdit v.2.3.1 (Wächter and Schroeder, 2010)
325 and formatted as OBO-file. Subsequently, custom R-scripts connected specific terms with
326 more general terms based on the described datasets. The entire ontology can be found as **File**

327 **S1.** In addition, a table containing all LION-terms with corresponding LION-identifier is
328 provided in **Data S5**.

329 **Implementation of enrichment analysis tool.** To use LION with existing ontology
330 enrichment tools, we used an adapted and generalized version of Bioconductor R-package
331 'topGO' (Alexa and Rahnenfuhrer, 2017). This version, called 'topOnto', allows users to
332 include ontologies other than those provided with the package. A custom Perl-script was used
333 to convert the ontology file from OBO- to SQLite-format. Apart from this extra feature, the
334 'topOnto' package provides the same functionality as the original version. To perform the
335 enrichment analysis with 'topOnto', two statistical approaches were used. In the 'by target-list
336 mode', Fisher-exact statistics are used to indicate enrichment. In the 'by ranking' mode,
337 Kolmogorov-Smirnov tests are used as 'global' statistics. In both approaches, topGO's classic
338 algorithm was selected. After LION enrichment analysis, raw *P* values were corrected for
339 multiple testing (Benjamini-Hochberg). The R-scripts were used to build the user-friendly
340 web-based tool LION/web **(Note S1)** with R-package 'shiny'. The application has been made
341 available on the shinyapps.io server as a free online tool, accessible through
342 http://www.lipidontology.com/.

343 **Cell culture and preparation of fatty acid-albumin complexes.** CHO-k1 cells were
344 cultured in Ham's F-12 medium (Thermo Fisher Scientific, Waltham, MA, USU)
345 supplemented with 7.5% FBS (Thermo Fisher Scientific, Waltham, MA, USU), 100 units/ml
346 penicillin and 100 µg/ml streptomycin (Thermo Fisher Scientific, Waltham, MA, USU). Cells
347 were grown in a humidified incubator at 37°C containing 5% CO_2 and passaged twice a week.
348 Stocks of 10 mM arachidonic acid, linoleic acid, oleic acid, or palmitic acid (all obtained
349 from Sigma, St. Louis, MO, USA) were complexed to 2 mM fatty-acid free BSA (Sigma, St.
350 Louis, MO, USA), filter-sterilized and stored at −20 °C. All experimental incubations were
351 performed in plastic 6-well culture dishes (Corning, Tewksbury, MA, USA).

352

353 **Measuring membrane fluidity.** After overnight incubation in the absence or presence of
354 fatty-acids (using fatty acid-free BSA or fatty acids coupled to BSA, respectively), cells were
355 washed and scraped in PBS. Cells were subsequently homogenized on ice with 26-gauge
356 needles (BD Bioscience, San Jose, CA, USA). Homogenates or blanks were mixed with
357 pyrenedecanoic acid (PDA) in the manufacturer's supplied dilution buffer (Membrane fluidity
358 kit, Abcam, Cambridge, UK) and transferred into a 96-well plate (black plastic with glass
359 bottom, Greiner Bio-One, Frickenhausen, Germany). After 30 minutes of incubation at 37°C,
360 fluorescence spectra (excitation at 360nm, emission between 375-500 nm, 37°C) were
361 measured with a temperature-controlled fluorescence microplate reader (CLARIOstar, BMG
362 Labtech, Offenburg, Germany). Data were processed in R by expressing monomer (370-390
363 nm) and excimer (470-490 nm) mean fluorescence after blank-subtraction as ratios. Data were
364 expressed as means. Differences were analyzed by two-tailed Student's t-tests. P values <
365 0.05 were considered significant.

366 **Lipidomics by LC-MS/MS.** After incubation, lipids were extracted as described before
367 (Bligh and Dyer, 1959). Subsequently, lipid extracts were dried under nitrogen and dissolved
368 in 100 µL chloroform/methanol (1:1) and injected (10 µL) on a hydrophilic interaction liquid
369 chromatography (HILIC) column (2.6 µm HILIC 100 Å, 50 x 4.6 mm, Phenomenex, Torrance,
370 CA), eluted by an eluens gradient (flow rate of 1 mL/min) from ACN/acetone (9:1, v/v) to
371 ACN/H_2O (7:3, v/v) with 10mM ammonium formate, both containing 0.1% formic acid. The
372 column effluent was connected to a heated electrospray ionization (hESI) source of a mass
373 spectrometer (Fusion, Thermo Scientific, Waltham, MA). The measurements were performed
374 with an orbitrap resolution of 120.000, generating 30 data-dependent MS/MS spectra per
375 second in the linear ion trap.

376 **Lipidomics data analysis.** Acquired raw datafiles were converted to mzXML-files and
377 processed with R-package 'xcms' v2.99.3 (Smith et al., 2006). After deisotoping, annotation
378 of lipids was performed by matching measured MS-1 m/z values with theoretical m/z values.
379 Lipids with the same or similar m/z values - *e.g.*, BMP(38:4) and PG(38:4) - could by

380 distinguished by differences in retention time **(Fig. S4 and S5)**. Lipid annotation containing
381 individual fatty acids as used in **Fig. 2 A** and **Fig. S4** was accomplished by examining MS-2
382 spectra. When MS-2 spectra were available for a given MS-1 peak, the most abundant fatty
383 acid combination was used to annotate the lipid. The resulting experimental datasets, as well
384 as the public RAW 264.7 macrophage dataset (Andreyev et al., 2010), were normalized by
385 expressing all lipids as ratios of the sum of all intensities per sample. MetaboAnalyst 3.0 (Xia
386 et al., 2015) was used to replace missing values (of the RAW 264.7 dataset) by half of the
387 minimum positive value in the original data, and to perform Principal Component Analysis
388 (with Pareto scaling).

389 **Software and R-packages.** All R-scripts were run with RStudio v1.0.153 (R v3.4.4) with the
390 following packages: 'shiny', 'visNetwork', 'data.table', 'GMD', 'igraph', 'reshape2',
391 'ggplot2', 'ggthemes', 'shinyTree', 'shinyWidgets', 'shinythemes', 'RSQLite', 'topOnto' and
392 'xcms' (Smith et al., 2006). Perl-scripts were run with Perl v5.26.0. All figures were build in
393 R and processed in Cytoscape v3.5.1 or Inkscape v0.92.2.

394 **Data and code availability.** The raw lipidomics data are available as Supplemental Data. The
395 public RAW 264.7 macrophages dataset (Andreyev et al., 2010) is available on the journal's
396 website. R-package 'topOnto' is available at https://github.com/hxin/topOnto, the associated
397 R-package containing the LION database at
398 https://github.com/martijnmolenaar/topOnto.LION2.db.
399