Spatial proteomics reveals lipid droplet reorganization in symbiotic *Paramecium bursaria* cells

Yan-Jun Chen[1], Kamal Md Mostafa[1]*, Chuan-Chih Hsu[2], and Jun-Yi Leu[1,3]*

[1] Institute of Molecular Biology, Academia Sinica, Taipei, 11529, Taiwan

[2] Institute of Plant and Microbial Biology, Academia Sinica, Taipei, 11529, Taiwan

[3] Lead Contact

* Correspondence:

Kamal Md Mostafa (mostafa.kamal42@gmail.com)

Jun-Yi Leu (jleu@imb.sinica.edu.tw)

Running title: A spatial proteome of the *Paramecium bursaria*

Keywords: ciliate, endosymbiosis, protein subcellular localization, lipid droplet, proteomics

Abstract

Endosymbiosis is an important adaptive mechanism allowing organisms to exploit novel niches. The flexible relationship between the ciliate *Paramecium bursaria* and green algae represents a model system for studying early endosymbiosis evolution. However, the mechanisms underlying how *P. bursaria* cells maintain this endosymbiotic relationship remain unclear. Here, we present a spatial proteomics atlas of *P. bursaria* cells with or without endosymbionts. We reveal that lipid droplets are reorganized in symbiotic hosts, with many proteins undergoing subcellular translocation. Consistent with our proteomics data, we show that the morphology and localization of lipid droplets changes near endosymbionts in host cells. Perturbing symbiotic cells with chemical inhibitors of lipid metabolism reduces endosymbiotic algal stability, revealing an involvement of lipid droplets in the host-endosymbiont interaction. Our study provides a comprehensive resource on protein localization and translocation in *P. bursaria* cells and elucidates how those cells remodel an existing organelle to maintain endosymbiosis.

Introduction

It is now well accepted that the first eukaryotic cell on Earth originated from endosymbiosis between bacterial and archaeal cells (Brueckner and Martin, 2020; Rivera et al., 1998). The establishment of endosymbiosis, representing a close relationship by which one organism (the endosymbiont) lives within another (a host), is often driven by complementation between the abilities of the endosymbiont and host. Consequently, the interacting partners can explore new niches or resources inaccessible to them when living alone (Sudakaran et al., 2017; Wein et al., 2019). Endosymbiosis has played a crucial role during the evolution of eukaryotic cells. In addition to during the formation of early eukaryotes, another endosymbiosis event generated the photosynthetic eukaryotic cell leading to the evolution of algae and plants (de Alda et al., 2014; Douglas, 1998). Since those early events, endosymbiosis has occurred continuously and influenced the evolution of various eukaryotic lineages (Zilber-Rosenberg and Rosenberg, 2008). However, little is known about how endosymbiosis is established initially and maintained.

The rapid increase in genomic data in recent years has facilitated deciphering pathways potentially involved in host-endosymbiont interactions (Hamada et al., 2018; Ishikawa et al., 2016). Nonetheless, only a few of them have been validated experimentally. Ciliates have long been known to harbor various endosymbionts and it has been speculated that such associations enable ciliate populations to dominate in certain marine and freshwater habitats (Esteban et al., 2010). Among ciliates, *Paramecium bursaria* is well known for its endosymbiotic relationship with the green algae *Chlorella spp.*, which has fascinated microbiologists for over a century (Fujishima, 2009). In the wild, most *P. bursaria* cells harbor several hundred algal cells in the cytoplasm, and the endosymbionts can be stably transmitted to daughter cells during cell division (Kodama and Fujishima, 2008). However, symbiotic *P. bursaria*

68 cells (green cells) can be artificially induced to lose their algal cells and form
69 aposymbiotic cells (white cells) (Ayala and Weis, 1987). When aposymbiotic *P.*
70 *bursaria* cells are mixed with free-living chlorella cells, they can quickly reestablish
71 endosymbiosis (Meier and Wiessner, 1988). Such flexibility indicates that the
72 endosymbiotic relationship is still at an early evolutionary stage.

73 Although *Paramecium* species have long served as model systems for cell
74 biology and studies of endosymbiosis (Long et al., 2023), only a small proportion of
75 their protein-coding genes have been functionally characterized (Gao et al., 2024;
76 Hauser et al., 2000; Jenkins et al., 2021). Even by integrating information on protein
77 domains and sequence homology, potential functions can only be assigned to
78 approximately half of the 15000 protein-coding genes in the *P. bursaria* genome (Cheng
79 et al., 2020). Thus, a systematic approach is required to understand the sheer complexity
80 of ciliate cells.

81 The emergence of new organelles and the subsequent expansion of the associated
82 protein repertoire is a key feature of eukaryote evolution. Ciliated protozoa represent a
83 lineage characterized by an abundance of unique organelles, and they also possess some
84 of the largest cells and gene families among all existing eukaryotes (Lynch et al., 2022).
85 In *P. bursaria* cells, many organelles have been reported as involved in the process of
86 endosymbiosis, with the cytoplasmic space being extensively rearranged to facilitate
87 hundreds of endosymbiotic algal cells (Fujishima and Kodama, 2012).

88 Many biological processes are executed in specific organelles or subcellular
89 localizations. Controlling protein abundance and localization enables a cell to regulate
90 these processes accurately. Moreover, cells can remodel or reorganize their organelles
91 to fine-tune their physiology in response to developmental or environmental cues.
92 Dissecting protein abundance and localization under different conditions can facilitate
93 our understanding of a protein's potential function and how cells adjust their physiology.

Recent advances in mass spectrometry-based proteomics have substantially improved our understanding of the protein landscape in various cell types (Aebersold and Mann, 2016; Mund et al., 2022). To gain insights into how *P. bursaria* cells reorganize their subcellular structure to facilitate hundreds of endosymbiotic algae, we generated spatial proteomics maps of aposymbiotic and endosymbiotic *P. bursaria* cells. Our data reveal coordinated remodeling of the signaling transduction and lipid transfer pathways at the level of protein abundance and localization in endosymbiotic cells. Microscopic analysis further showed that the morphology and localization of lipid droplets changes close to the endosymbiotic algae. Inhibiting lipid metabolism reduced endosymbiont stability. Our study not only offers a comprehensive resource of protein localization and translocation data between two types of ciliate cells, but also reveals the involvement of lipid droplets in endosymbiosis.

Results and Discussion

A comprehensive spatial proteome map of *Paramecium bursaria* cells with or without endosymbiotic algae.

Although the cell biology of *P. bursaria* has been studied for decades, its genome was only decoded recently (Cheng et al., 2020; He et al., 2019). Nonetheless, a considerable proportion of its encoded proteins are of unknown function. We investigated the spatial distribution of the *P. bursaria* proteome given that a protein's subcellular localization can provide information about its function (Su et al., 2007). Moreover, many organelles are involved in the process of endosymbiosis (Kodama and Fujishima, 2022; Kodama and Sumita, 2022; Song et al., 2017). Thus, by comparing proteome maps of aposymbiotic and symbiotic cells, we anticipated gaining insights into how those organelles are reorganized to facilitate endosymbiosis.

We utilized protein correlation profiling to establish organellar protein

localization (Itzhak et al., 2019). In brief, *P. bursaria* cells were lysed by nitrogen cavitation, and the cellular organelles were separated by differential centrifugation. We collected a total of eleven fractions, and the distinct protein patterns displayed by the individual fractions indicated that our protocol could successfully separate different cellular compartments (Supplementary Fig. 1a). The collected fractions were subsequently analyzed by liquid chromatography-mass spectrometry (LC-MS) through the data-independent acquisition (DIA) mode, and the resulting data were processed with Spectronaut for label-free quantitation (LFQ) analysis (Fig. 1a, see Methods for details). We analyzed three biological replicates for both symbiotic and aposymbiotic cells. Approximately 95% of identified proteins were shared between replicates, and most proteins presented a low coefficient of variation (CV) between replicates, indicative of high data quality (Fig. 1b and Supplementary Fig. 1b).

To construct spatial proteomics maps, we needed marker proteins of different cellular compartments for correlation profiling. Consequently, we selected a total of 279 marker proteins based on ortholog localization information from previous studies and machine-learning algorithms (Supplementary Data 1, see Methods for details) (Licknack, 2022; Pedregosa et al., 2011). These markers define 13 cellular compartments, including the cytosol, proteasomes, ribosomes, peroxisomes, endoplasmic reticulum (ER), Golgi, plasma membrane, nucleus, lipid droplets, basal bodies, mitochondria, phagolysosomes, and trichocysts. We detected distinct fractionation profiles for the different cell compartment markers (Fig. 1c and Supplementary Fig. 2 and 3), indicating that the different compartments had been well separated.

Next, we used the total proteins (n = 10255) that overlapped between aposymbiotic and symbiotic cells for spatial proteome characterization. Our mass spectrometry data were globally normalized and then analyzed using support vector

146 machine (SVM)-based supervised learning, which computes the probability of a protein
147 belonging to one of our pre-defined subcellular compartments based on how their
148 localization correlates with that of the compartment markers (Itzhak et al., 2016). Our
149 analysis allowed us to classify the subcellular localizations of 5872 and 4840 proteins
150 in the symbiotic and aposymbiotic cells, respectively, representing about half of the
151 identified *P. bursaria* proteome (Fig. 2a and 2b and Supplementary Fig. 4).

152 When we projected the 33-dimension protein profiling dataset (from 33 fractions
153 of three replicates, Supplementary Fig. 5 and Supplementary Data 2) of classified
154 proteins into two dimensions by t-distributed stochastic neighbor embedding (t-SNE),
155 we found that the proteins of different cellular compartments formed discrete clusters
156 (Fig. 2c), indicating that our analysis can clearly distinguish proteins displaying
157 differential localization patterns. In addition to helping decipher the possible function
158 of unknown proteins, our comprehensive localization maps allowed us to investigate
159 the physiological changes that occur during the process of endosymbiosis.

160

161 **Comparison between our spatial proteome maps and Deeploc prediction.**

162 Since *P. bursaria* cells cannot be manipulated genetically and only a few proteins have
163 been experimentally characterized based on their cellular localization, it is almost
164 impossible to experimentally validate our proteomics data systematically. Nonetheless,
165 a pre-trained protein language model, DeepLoc 2.0, has been demonstrated to predict
166 protein subcellular localization in several organisms with reasonable accuracy
167 (Thumuluri et al., 2022). We compared our localization data with the results predicted
168 by DeepLoc (Supplementary Data 3). A large proportion of the proteins assigned to the
169 ER and mitochondria in our experiments were consistently predicted by DeepLoc to
170 localize to those organelles, respectively (Fig. 3a). The majority of nuclear proteins
171 predicted by DeepLoc were also found in a group mapping to the nucleus. It is worth

172 noting that the majority of the training data used in developing DeepLoc was obtained
173 from model eukaryotic organisms evolutionarily distant from *P. bursaria*, which may
174 influence the accuracy of our DeepLoc predictions. Moreover, approximately one-third
175 of the *P. bursaria* proteome does not have any reasonable orthologs in databases of
176 other model organisms.

177 Interestingly, Deeploc predicted many trichocyst- and phagolysosome-associated
178 proteins as being extracellular. Trichocysts are secretory organelles found in ciliates and
179 in some flagellate protists (Pease, 1947). They comprise a cavity and long threads of
180 trichocyst proteins that can be released extracellularly in response to external stimuli,
181 such as upon detection of predators (Plattner, 2017). Phagolysosomes are formed by the
182 fusion of phagosomes with lysosomes after phagocytosis. Although phagolysosomes
183 are primarily involved in degrading food and debris, they may also be used to degrade
184 secreted proteins or exosomes engulfed by phagocytosis or endocytosis (Gurung et al.,
185 2021). Further studies of the "extracellular" nature of trichocyst and phagolysosome
186 proteins are needed to understand this finding.

187

188 **Lipid droplets are reorganized in symbiotic cells.**

189 Given that symbiotic *P. bursaria* cells reorganize their cytoplasm to accommodate
190 hundreds of endosymbiotic algae, we expected that the localization of some proteins
191 may be altered for this purpose. Comparing the spatial proteome maps between
192 symbiotic and aposymbiotic cells, we observed a dramatic increase in proteins
193 localizing to lipid droplets in the symbiotic cells (Fig. 2b). To confirm this observation,
194 we performed a highly stringent test using movement and reproducibility (MR) scores
195 (Itzhak et al., 2019), which compared the fractionation profiles from all replicates
196 between the two types of cells (see Methods for details). Our MR analysis identified
197 240 proteins for which fractionation profiles and subcellular localization were

198 significantly altered (false discovery rate (FDR) < 5%) (Supplementary Data 4). Among
199 them, proteins moving to lipid droplets represented the largest group of translocated
200 proteins in symbiotic cells (Supplementary Fig. 6), excluding proteins without a
201 specific localization (i.e., the unknown group). This result is consistent with our spatial
202 proteome map (Fig. 2), which shows that lipid droplets are significantly reorganized.
203 Lipid droplets regulate intracellular lipid storage, transport, and metabolism
204 (Walther and Farese, 2012). They are highly dynamic structures, interacting with
205 various organelles (including ER, Golgi, mitochondria, lysosomes, and peroxisomes)
206 to exchange proteins and lipids, and acting as vital hubs of cellular metabolism
207 (Olzmann and Carvalho, 2019). A gene ontology (GO) analysis of lipid droplet proteins
208 specifically in symbiotic cells revealed enrichment for many pathways involved in
209 vesicle trafficking, membrane organization, ion and protein transport, and lipid
210 metabolism (Fig. 3b and Supplementary Data 5). These pathways are all related to
211 various functions of lipid droplets, indicating that the functions of lipid droplets may
212 shift in symbiotic cells.
213
214 **Compromising lipid metabolism reduces endosymbiosis stability.**
215 Our spatial proteomics data revealed that lipid droplets undergo a dramatic change in
216 composition in symbiotic cells. Consequently, we stained the cells with the lipid droplet
217 dye LipidSpot™ to examine them by confocal microscopy. We observed that, in
218 individual symbiotic cells, hundreds of tiny lipid droplets seemed to interact with the
219 endosymbiotic algae (Fig. 4a). Next, we applied transmission electron microscopy
220 (TEM) to examine the relationship between lipid droplets and endosymbiotic algae
221 more closely. Indeed, the lipid droplets were localized near most algae, with some of
222 them being directly attached to the perialgal vacuole (PV) membrane surrounding
223 individual algal cells (Fig. 4b). Together, these observations indicate that the

224 reorganization of lipid droplets in the symbiotic cells may play a role in the host cells'
225 interactions with their endosymbionts.

226 We investigated if lipid droplets are crucial for maintaining endosymbionts by
227 applying two types of lipid metabolism inhibitors, H89 and T863. H89 blocks protein
228 kinase A and other kinases, leading to unbalanced lipolysis and lipogenesis (Londos et
229 al., 2005; Rogne and Taskén, 2014). T863 can reduce triglyceride synthesis and lipid
230 droplet formation by inhibiting diacylglycerol acyltransferase 1 (DGAT1) (Yen et al.,
231 2008). We treated symbiotic cells with a mild dose of H89 or T863 that did not influence
232 host survival. Application of these drugs reduced the numbers of lipid droplets in
233 symbiotic cells (Fig. 5a), and the treatments negatively influenced the stability of
234 endosymbionts, with the number of algae in host cells decreasing significantly (Fig. 5b
235 and 5c).

236 Previous studies have shown that some parasites induce lipid droplet formation
237 in the host cell to obtain nutrients or evade immune attack (Roingeard and Melo, 2017).
238 For example, mammalian cells infected by the protozoan *Toxoplasma* display increased
239 numbers of lipid droplets (Mota et al., 2014). These newly formed lipid droplets attach
240 to the parasitophorous vacuole containing *Toxoplasma* cells, allowing the parasites to
241 utilize the nutrients for their own reproduction (Nolan et al., 2017). Similarly, the lipid
242 droplets in *P. bursaria* cells may facilitate nutrient uptake from the host by the
243 endosymbiotic algae. Alternatively, lipid droplets may provide essential components
244 for generating or maintaining PV membranes. Inside *P. bursaria*, a PV membrane
245 surrounds each algal cell, and this membranous structure needs to divide and reorganize
246 each time the algal cells reproduce. If PV structures become compromised, it would
247 likely result in unstable endosymbiosis. Lastly, lipid droplets may serve as a signal
248 exchange hub between the host and endosymbionts, with their disruption meaning
249 impaired communication between the two interacting partners, potentially prompting

250 digestion of the endosymbionts.

251 Lipid droplets have been observed to play a crucial role in host-pathogen
252 interactions. Pathogens, including viruses, bacteria, and eukaryotic parasites, can
253 manipulate host lipid metabolism and exploit lipid droplets as a nutrient source or as a
254 platform for viral replication (Roingeard and Melo, 2017). Moreover, lipid droplets
255 participate in immune responses by regulating the production of lipid-derived signaling
256 molecules, such as eicosanoids, which influence inflammation and pathogen defense
257 mechanisms (Roingeard and Melo, 2017). Whether hijacked lipid droplets are
258 completely remodeled in parasitic relationships, as observed herein for symbiotic *P.*
259 *bursaria* cells, remains to be established. Our spatial proteomics data have also shown
260 that proteins involved in the nitric oxide and cGMP-protein kinase G (PKG) signaling
261 pathways localized to lipid droplets and increased in abundance in symbiotic cells
262 (Supplementary Fig. 7). These signaling cascades can regulate many downstream
263 pathways, influencing intracellular concentrations of cations, lipid transport, and
264 vesicle trafficking (Kissmehl et al., 1996; Pirayesh et al., 2021). Thus, the role of lipid
265 droplets is likely more complex in *P. bursaria* cells since their interacting partners form
266 a mutualistic relationship, and not a pathogen-host interaction.

267

268 # Methods

269 **Cell culture**

270 *P. bursaria* cells were cultured in CGM medium (12 mM sodium acetate, 0.5% yeast
271 extract) with Dryl's buffer (2 mM sodium citrate, 1 mM NaH_2PO_4/Na_2HPO_4 and 1.5
272 mM $CaCl_2$) and fed with *Chlorogonium* cells every five days. The cell cultures were
273 kept at 23 °C with a light-to-dark cycle of 12 h:12 h. To obtain aposymbiotic *P. bursaria*
274 strains, green symbiotic cells were treated with cycloheximide (100 μg/ml) for about
275 two weeks, as described previously (Kodama et al., 2011). The aposymbiotic cells were

276 maintained under low-light conditions.

277 For the inhibitor treatment experiments, symbiotic *P. bursaria* cells were treated
278 with different concentrations of H89 (Cat. # HY-15979, MedChemExpress, NJ, USA)
279 or T863 (Cat. # HY-32219, MedChemExpress) to test their viability and growth. We
280 used 1 μM of H89 and 10 μM of T863 in the final experiment since these drug
281 concentrations did not have any apparent deleterious effects on *P. bursaria* cells.

282

283 **Organelle fractionation**

284 *P. bursaria* cells were harvested from a 10 μm-diameter nylon mesh sieve after
285 removing cell debris using cheesecloth. The cells were washed on the nylon mesh with
286 Dryl's solution, before being decanted into 15 ml tubes. Then, the cells were
287 resuspended in lysis buffer (10 mM HEPES-KOH (pH7.5), 2 mM EDTA-KOH (pH8.0),
288 2 mM magnesium, 1 mM/2 mM PMSF/SHAM, 250 mM sucrose and 1X protease
289 inhibitor cocktail) and lysed using a nitrogen cavitation bomb. The cavitation bomb was
290 pressurized to 300 psi and incubated for 20 min before slowly releasing the pressure.

291 The lysate was then differentially centrifuged at 300 x *g* for 3 min, 1000 x *g* for
292 10 min, 3000 x *g* for 10 min, 5000 x *g* for 10 min, 9000 x *g* for 15 min, 12000 x *g* for
293 15 min, 15000 x *g* for 15 min, 30000 x *g* for 20 min, 79000 x *g* for 43 min, and 120000
294 x *g* for 45 min. After each centrifugation, the pellet was collected and resuspended in
295 UTU buffer (6 M urea, 2 M Thiourea, 100 mM Tris, 2.5 mM DTT, 5 mM EDTA-KOH
296 (pH8.0), 1 mM/2 mM PMSF/SHAM, and 1X protease inhibitor cocktail) and stored at
297 -80 ^{o}C. After the 12000 x *g* centrifugation step, we added acetone to the supernatant to
298 precipitate the protein overnight at -20 ^{o}C. Protein concentration was determined by
299 Bradford assay.

300

301 **Sample preparation and LC-MS analysis**

302 Protein digestion in the S-Trap microcolumn was performed according to a previous
303 report (Chen et al., 2023). In brief, 5 μg of protein in lysis buffer was reduced and
304 alkylated using 10 mM Tris(2-carboxyethyl) phosphine hydrochloride and 40 mM 2-
305 chloroacetamide at 45 °C for 15 min. A final concentration of 5.5% (v/v) phosphoric
306 acid (PA) followed by a six-fold volume of binding buffer (90%, v/v, MeOH in 100 mM
307 TEAB) was next added to the protein solution. After gentle vortexing, the solution was
308 loaded into an S-Trap microcolumn. The solution was removed by spinning the column
309 at 4,000 x *g* for 1 min. The column was washed three times with 150 μL binding buffer.
310 Finally, 20 μL of digestion solution containing 200 ng Lys-C and trypsin in 50 mM
311 TEAB was added to the column and incubated at 47 °C for 2 h. Each digested peptide
312 was eluted using 40 μL of three buffers consecutively: (1) 50 mM triethylammonium
313 bicarbonate; (2) 0.2% (v/v) formic acid in H_2O; and (3) 50% (v/v) acetonitrile (standard
314 elution). Elution solutions were collected in a tube and dried under a vacuum. A total
315 of 500 ng peptides was loaded into an Evotip and then the Evotip was placed into an
316 Evosep One system (Evosep Biosystems, Odense, Denmark) coupled to a timsTOF HT
317 mass spectrometer (Bruker Daltonics, MA, USA).

318 Indexed retention time peptides (iRT) (KI-3002-2, Biognosys, Schlieren,
319 Switzerland) were dissolved in 50 μl of 0.1% (v/v) formic acid (FA) with 3% (v/v) ACN
320 to achieve a 10× stock solution. The stock solution was then diluted 10-fold. Dried
321 peptides were reconstituted in 200 μl of 0.1% (v/v) TFA. Then, 20 μl of samples and 10
322 μl of 1× iRT solution were loaded onto EvoTips and analyzed using the same Evosep
323 One system.

324 Eluting peptides were separated analytically using the 30 SPD method on an
325 Evosep Performance column (15 cm x 150 μm ID, 1.5 μm, EV1137; Evosep) at 40 °C
326 within a Captive Spray source (Bruker Daltonics), equipped with a 20 μm emitter (ZDV
327 Sprayer 20, Bruker Daltonics), followed by dia-PASEF mode analysis. For dia-PASEF,

328 the instrument firmware was adapted to perform data-independent isolation of multiple
329 precursor windows within a single TIMS frame. An optimized dia-PASEF scheme was
330 applied that targeted +2 and +3 ions in a two-window method. Sixteen of these scans
331 (resulting in 32 windows, each of 25-Da window size) covered an m/z range of 400 to
332 1200. Each dia-PASEF cycle comprised an MS1 survey frame, followed by 16 dia-
333 PASEF frames, with this setup resulting in a total cycle time of 1.8 s.
334

335 **Data-independent acquisition (DIA) analysis**

336 DIA data signal extraction and quantification were performed with an analysis pipeline
337 in Spectronaut (Biognosys, v18.7, https://biognosys.com/software/spectronaut/) using
338 a standard setting with Carbamidomethylation (Cys) specified as a fixed modification,
339 oxidation (Met), and N-terminal acetylations specified as variable modifications. We
340 selected dynamic retention time prediction with local regression calibration.
341 Interference correction for the MS and MS2 levels was enabled. The FDR was set to
342 1% at peptide precursor and protein levels using scrambled decoy generation and a
343 dynamic library size fraction set at 0.1. MS2-based quantification was used, enabling
344 local cross-run normalization. Peptide precursor abundances were measured by the sum
345 of fragment ion peak areas and peptide grouping was performed based on modified
346 peptides. Protein-level quantification was performed against the spectrum library
347 generated *in silico* by Spectronaut with standard settings and a stripped peptide
348 sequence area obtained from the mean precursor area.
349

350 **Data processing and assessment of proteomics depth**

351 For label-free quantitation (LFQ) maps, Spectronaut already globally normalizes
352 intensities, hence no further normalization was required. Proteins not identified in all
353 replicates were removed. Three biological replicates were performed for each cell type.

354 Proteomic differences were evaluated for statistical significance ($P<0.05$) by Student's
355 t-tests and corrected for multiple testing using the Benjamini-Hochberg correction.
356 Normalization (0-1) of each fractionation profile was performed by summing the
357 intensities across all fractions and dividing each intensity by the profile sum.
358 Proteomics depth was assessed by counting protein groups or proteins identified
359 in one or all replicates. Venn diagrams and UpSet plots were generated to evaluate
360 protein overlap.
361
362 **Marker protein curation and support vector machine (SVM) analysis**
363 To generate the subcellular compartment marker protein list, first we examined the
364 marker proteins identified in *Paramecium tetraurelia* by Timothy J. Licknack
365 (Licknack, 2022) by manually reviewing the literature and InterProScan annotation.
366 Next, we used OrthoFinder to find the orthologous markers in the *P. bursaria* genome
367 (Emms and Kelly, 2015). These marker proteins were inspected for their protein
368 abundance distribution profiles and checked for their deviation from the cluster using a
369 5-fold cross-validation function provided in Scikit-learn (Pedregosa et al., 2011). Only
370 the marker proteins present in all compared datasets were included, and identical SVM
371 parameters were used. Classification of unlabeled proteins was performed using SVM.
372 Based on SVM, we achieved a prediction accuracy (F1-score) of 99.2% for DK2g
373 (symbiotic) and 97.8% for DK2w (aposymbiotic) cells (Supplementary Fig. 3a). Finally,
374 we established a list of manually curated proteins defining 13 subcellular compartments
375 (Cytosol, Proteasome, Ribosome, Peroxisome, ER, Golgi, Plasma membrane, Nucleus,
376 Lipid droplet, Basal body, Mitochondrion, Phagolysosome, and Trichocyst)
377 (Supplementary Data 1). To further evaluate the performance of our organelle maps,
378 their power to predict protein localization was assessed using quality-filtered and
379 normalized (0-1) data with full replicate coverage.

380 For supervised classification, the marker proteins covering 13 subcellular
381 localizations were used by SVM to assign all other proteins to the organellar cluster.
382 Machine learning was conducted using the functionality-based command provided in
383 Scikit-learn. SVM enables selection of maker classes, definition of a hold-out test set,
384 automated hyperparameter optimization and, finally, SVM training and prediction.
385 During training, the parameters C and gamma of the radial basis function were
386 optimized via an iterative grid search, employing 5-fold cross-validation. At each grid
387 point, SVMs were run on each loaded dataset, before identifying the optimum for
388 summed accuracy across datasets. Finally, we confirmed analytical performance using
389 Q-Sep (Geladaki et al., 2019), which quantifies the separation between a pair of
390 predicted protein sets by comparing intra-group and inter-group distances
391 (Supplementary Fig. 3b and Supplementary Data 6).
392 For prediction, the SVM was fitted using the training set and 5-fold cross-
393 validation to convert the raw SVM scores into probabilities. Proteins were then assigned
394 to the best-fitting organelle model and divided into confidence classes based on the
395 probability: >0.9, very high confidence; >0.8, high confidence; >0.5, medium
396 confidence; >0.4, low confidence; <0.4, best guess assignment (Schessner et al., 2023).
397 For the hold-out test set, a misclassification matrix was derived to calculate global
398 marker prediction recall (proportion of correctly predicted markers to the total number
399 of markers), organelle-specific recall (proportion of markers correctly assigned to the
400 cluster), and organelle-specific precision (ratio of markers correctly assigned to the
401 number of all markers assigned to the cluster). The harmonic mean of recall and
402 precision, the F1 score, was used as the primary readout for SVM performance.
403
404 **Statistics and reproducibility**
405 Each of the final 33 fractions per cell type was comprised of three concatenated 11-

406 fraction iterations of the experimental procedure. Proteins with a missing value in any
407 of the 33 data points per cell type were excluded from further analysis. Protein features
408 were examined according to unsupervised clusters and subcellular compartments using
409 enrichment analysis. No statistical methods were used to predetermine the sample size.
410 The experiments were not randomized and the investigators were not blinded to
411 allocation during experiments and outcome assessment.
412
413 **t-SNE (t-distributed stochastic neighbor embedding) visualization**
414 t-SNE was used to reduce dimensionality and visualize data. For t-SNE, the
415 preprocessed data were embedded into two dimensions using a perplexity of 60 and
416 extract gradient calculation for a maximum of 1000 iterations. The computed
417 coordinates were recorded and used to obtain the two-dimensional data projection
418 presented herein.
419
420 **Protein feature annotation and gene ontology (GO) enrichment analysis**
421 For subcellular localization, protein feature annotation was predicted using DeepLoc
422 2.0 (Thumuluri et al., 2022). GO terms were predicted by InterProScan (Blum et al.,
423 2021), the String database (Szklarczyk et al., 2023), DeepFRI (Gligorijevic et al., 2021),
424 and Pannzer 2 (Törönen et al., 2018). Fisher's exact tests were applied to the annotation
425 terms. The resulting p-values were corrected for multiple hypotheses using Bonferroni
426 correction, and a cutoff of 5% FDR was applied (Klingelhuber et al., 2024).
427
428 **Protein subcellular localization shift analysis.**
429 To evaluate the reproducibility of our (0-1)-normalized profiles, we calculated
430 the inter-profile scatter across replicates. The absolute distance (Manhattan Distance)
431 for each protein of each replicate profile to the mean profile from all replicates was

calculated, and these distances were averaged to obtain that protein's profile scatter (Supplementary Data 7). Global profile scatter was then plotted as a density function for proteins common to all compared maps. The distribution for all profiled proteins could then be displayed as an additional output, regardless of overlap with the other examined maps. The greater the proportion of proteins with low scatter, the better the between-map reproducibility. As a numeric readout, we determined the scatter at a specified quantile of each distribution (Itzhak et al., 2016; Schessner et al., 2023).

We compared organelle maps from DK2w and DK2g cells. Replicates were numbered 1-3. Each normalized protein biodistribution profile was obtained across the six maps. Next, delta profiles were calculated within each cognate pair of untreated and treated maps by subtracting the (0-1)-normalized profiles from a treated map from the profiles of its matching control map. For each of the three obtained sets of delta profiles, a multidimensional outlier test based on the robust Mahalanobis distance was performed using Scikit-learn (Supplementary Data 8). These distances follow a chi2 distribution and the resulting p-values reflect the probability of observing a given profile shift (or a greater shift) by chance. The three p-values from the three replicates were combined for each protein using the Fisher method. The combined p-values were corrected for multiple testing using the Benjamini-Hochberg false-discovery-rate (FDR) approach and it was then $-\log(10)$ transformed to obtain a Movement score (M score). We chose an FDR of 0.05 as our initial stringent cut-off, corresponding to an M score of 1.3. We also included the requirement that all of the replicate p-values needed to be <0.1 to filter out proteins with single very small p-values in only one replicate.

We also calculated a reproducibility score (R score) as a further stringency filter. For each protein, the Pearson correlation of all pairs of delta profiles was calculated (Rep 1 vs. 2, 1 vs. 3, 2 vs. 3), and the median correlation was designated as the R score (Supplementary Data 9). Profile shifts with $M > 1.3$ and $R > 0.75$ were considered

458 significant and reproducible. We also calculated the cosine correlation of each protein
459 between any replicates and set the criteria of cosine correlation > 0.75 (Supplementary
460 Data 10) (Itzhak et al., 2016; Schessner et al., 2023).
461
462 **Confocal microscopy**
463 For lipid droplet staining, *P. bursaria* cells were first fixed with 4% formaldehyde for
464 10 min and resuspended in 1000X dilution lipid droplet dye LipidSpot™ (Cat. # 70065-
465 T, Biotium, CA, USA) for 30 mins at room temperature. Imaging was performed using
466 a ZEISS LSM 780 Upright Confocal Microscope (ZEISS, Oberkochen, Germany).
467
468 **Transmission electron microscopy (TEM) imaging**
469 *P. bursaria* cells were collected and fixed in a mixed aldehyde solution containing 2.5%
470 glutaraldehyde (GA) and 4% paraformaldehyde (PFA) in 100 mM sodium cacodylate
471 buffer (pH 7.2). Fixation was performed using a microwave. Samples were irradiated
472 at 500W for 1 min, followed by a 30-s pause, and an additional 1-min irradiation under
473 vacuum (20 in. Hg). This cycle was repeated eight times. After the final microwave
474 cycle, samples were shaken for 30 min at room temperature, followed by an additional
475 2 h of shaking before overnight storage at 4 °C.
476 Following fixation, samples were rinsed in 100 mM sodium cacodylate buffer
477 (pH 7.2) under microwave conditions of 150W for 1 min, repeated four times, with no
478 vacuum applied. Post-fixation was performed using 1% osmium tetroxide (OsO_4) in
479 100 mM sodium cacodylate buffer (pH 7.2). The post-fixation process followed similar
480 microwave conditions: 450W for 1 min, followed by a 30-s pause, and another 1-min
481 irradiation under a 20-in. Hg vacuum. This cycle was repeated eight times, with samples
482 shaken at room temperature for 30 min after each cycle, followed by a final 2-h shaking
483 period before overnight storage at 4 °C.

484 After post-fixation, samples were rinsed with distilled water (ddH_2O) under
485 microwave conditions of 150W for 1 min, repeated four times, without a vacuum. En-
486 bloc staining was performed using 2% uranyl acetate (UA). The staining process was
487 carried out under microwave irradiation at 450W for 1 min, followed by a 30-s pause,
488 and another 1-min of irradiation with a vacuum of 20-in. Hg. After staining, samples
489 were rinsed with ddH_2O, with four cycles of microwave irradiation at 150W for 1 min
490 per cycle, without a vacuum. Samples were then dehydrated through a graded ethanol
491 series (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%). Each ethanol
492 step involved microwave irradiation at 200W for 10 s, followed by a pause of 20 s, and
493 another 10-s irradiation. This process was repeated twice for each ethanol concentration,
494 without a vacuum.

495 Following dehydration, resin infiltration was carried out using spurr's resin
496 without an accelerator. Initially, 5% spurr's resin was applied for 2 days at room
497 temperature. Samples were then infiltrated through a graded resin series (10% to 100%)
498 under a vacuum, with shaking every 8 hours. Two changes of 100% spurr's resin
499 without accelerator were applied at 8-h intervals under a vacuum, followed by three
500 changes of 100% spurr's resin with the accelerator under the same conditions. Finally,
501 resin-infiltrated samples were embedded in molds and polymerized in a pre-warmed
502 oven at 70 °C for 24 hours.

503

504 **Data availability.**

505 The mass spectrometry proteomics data have been deposited in the ProteomeXchange
506 Consortium (Deutsch et al., 2023) . The mass
507 spectrometry proteomics data analyzed in this study are provided as Supplementary
508 Data. The Source Data file provides other data generated or analyzed in this study. All
509 supplementary files and Source Data are provided with this article.

510

511 ## Acknowledgments

512 We thank members of the Leu lab for helpful discussion and comments on the
513 manuscript. We also thank John O'Brien for manuscript editing, the IMB Imaging Core
514 for imaging experiments, and the IPMB Proteomics Core for proteomics experiments.
515 This work was supported by Academia Sinica of Taiwan (grant no. AS-IA-110-L01 and
516 AS-GCS-113-L03) and the National Science and Technology Council of Taiwan
517 (NSTC 113-2326-B-001-002). KMM was supported by an NSTC postdoctoral
518 fellowship (NSTC 113-2811-B-001-065).

519

520 ## Author contributions

521 JYL conceived the study. YJC, KMM and JYL designed analyses and interpreted
522 results. YJC and CCH performed the experiments. YJC, KMM and CCH performed
523 data analysis. YJC and JYL wrote the paper. All authors read and approved the final
524 manuscript.

525

526 ## Declaration of interests

527 The authors declare no competing interests.

528

529 ## Figure legends

530 **Figure 1. Application of differential centrifugation to separate subcellular**
531 **compartments of *P. bursaria* cells.**

532 (a) Workflow of *P. bursaria* subcellular localization mapping by differential
533 centrifugation. *P. bursaria* cells with (DK2g) or without (DK2w) endosymbionts were
534 lysed by nitrogen cavitation and the organelles were separated by differential
535 centrifugation. Organelle fractions were collected and analyzed by DIA-LC/MS.

536 Protein profiles were generated and used to predict protein localizations using SVM
537 machine learning. (b) Venn diagrams showing numbers of proteins identified in three
538 replicate experiments. The final datasets contained 10622 proteins for DK2g cells and
539 10491 proteins for DK2w cells. (c) Schematic representation of subcellular
540 compartments in *P. bursaria* and the distribution profile of marker proteins. The
541 distribution profiles of all of the marker proteins from three replicates were averaged
542 and are shown for each compartment. The numbers in brackets represent the numbers
543 of marker proteins for each compartment. Individual profiles are shown in
544 Supplementary Fig. 2. ER, endoplasmic reticulum.
545
546 **Figure 2. Subcellular localization maps of symbiotic and aposymbiotic *P. bursaria***
547 **cells.**
548 (a) Proportions of identified proteins in different subcellular compartments. Protein
549 localization was assigned using SVM-based supervised learning. Proteins that could
550 not be assigned to a specific compartment were classified into the "unknown" group.
551 (b) Numbers of identified proteins that were assigned to different subcellular
552 compartments. (c) t-SNE projections of the proteins assigned to different subcellular
553 compartments in DK2g and DK2w cells. Proteins are colored according to their
554 subcellular localization.
555
556 **Figure 3. Comparative analysis of spatial proteomes and gene ontology (GO)**
557 **analysis of symbiotic cell-specific lipid proteins.**
558 (a) Comparison between the results predicted by the pre-trained protein language model,
559 DeepLoc 2.0, and our subcellular localization maps. Color coding represents different
560 predicted subcellular localizations, as determined by DeepLoc 2.0. (b) GO analysis of
561 symbiotic cell-specific lipid droplet proteins. Only GO terms with an enrichment score

562 > 2 and an adjusted p-value < 0.05 are shown. All of the enriched GO terms belong to
563 five major categories.
564
565 **Figure 4. The morphology and localization of lipid droplets are altered in**
566 **symbiotic cells.**
567 (a) Lipid droplets in aposymbiotic and symbiotic *P. bursaria* cells. Lipid droplets were
568 stained using LipidSpot^{TM}. The Cy5 channel was used to detect the chlorophyll of
569 endosymbiotic algal cells. Scale bar, 10 μm. (b) Many lipid droplets are attached to the
570 perialgal vacuole (PV) membrane of the endosymbionts, as revealed by TEM imaging.
571 The arrow indicates the PV membrane surrounding individual algal cells. Scale bar, 1
572 μm; LD, lipid droplet.
573
574 **Figure 5. Treating symbiotic *P. bursaria* cells with lipid metabolism inhibitors**
575 **reduces endosymbiont stability.**
576 (a) Treating symbiotic *P. bursaria* cells with mild doses of lipid metabolism inhibitors
577 H89 (1 μM) or T863 (10 μM) reduces the number of lipid droplets. Lipid droplets (in
578 green) were stained using LipidSpot^{TM}. (b) H89 or T863 treatment affects the stability
579 of endosymbiotic algae in *P. bursaria* cells. Scale bar, 10 μm. (c) Numbers of
580 endosymbiotic algae are significantly reduced after two weeks of treatment. ***, p-
581 value < 0.001, Mann-Whitney U test.
582

583 ## Supplementary Figure Legends

584 **Supplementary Figure 1. Differential centrifugation consistently separates**
585 **different populations of proteins.**
586 (a) SDS-PAGE of subcellular fractions collected from differential centrifugation. (b)
587 Proteins identified from three replicates exhibit small coefficients of variation (CV) in

588 terms of their abundance. Only proteins identified in both symbiotic and aposymbiotic
589 cells were analyzed (n = 10255). The top panel shows the distribution of protein
590 abundance CV, and the bottom panel shows the cumulative percentage of total protein
591 abundance CV. The medians of DK2g and DK2w cells are 0.25 and 0.21, respectively.
592
593 **Supplementary Figure 2. Fractionation distribution profiles of 279 subcellular**
594 **compartment marker proteins.**
595 Profiles of each marker protein averaged from three replicates of symbiotic (a) and
596 aposymbiotic (b) cells are shown.
597
598 **Supplementary Figure 3. Quality examination of the marker proteins.**
599 (a) Prediction accuracy for marker proteins, as indicated by the F1 score for DK2g and
600 DK2w cells. Organelle-specific recall indicates the proportion of markers correctly
601 assigned to the cluster, and organelle-specific precision indicates the ratio of markers
602 correctly assigned to the number of all markers assigned to the cluster. The F1 score is
603 derived from the harmonic mean of recall and precision, representing the primary
604 readout for SVM performance. (b) Q-Sep analysis shows that proteins within the same
605 subcellular compartment have shorter Q-Sep distances than those between different
606 compartments.
607
608 **Supplementary Figure 4. Proteins from the same subcellular compartment show**
609 **similar fractionation distribution profiles.**
610 Profiles of each protein averaged from three replicates are shown. The color intensity
611 indicates the percentage of an individual protein contained in each fraction.
612
613 **Supplementary Figure 5. Fractionation distribution profiles of all proteins with**

614 **predicted subcellular localizations.**

615 Concatenated profiles of each protein from three replicates.

616

617 **Supplementary Figure 6. Sankey diagram showing how proteins are translocated**
618 **to different subcellular compartments under conditions of aposymbiosis (DK2w**
619 **cells) and symbiosis (DK2g cells).**

620 The analysis was performed by calculating Movement-Reproducibility (M-R) scores.
621 Significantly altered map shift (FDR<5%) corresponds to an M score cut-off at 1.3. The
622 M score is calculated as -log10 of the Benjamini-Hochberg-corrected combination of
623 p-values derived from three independent chi2 distributions of robust Mahalanobis
624 distance (see Methods for details). In addition, an R score cut-off of 0.75 was applied,
625 and only proteins with a p-value <0.1 in three replicates were considered. For this
626 stringent analysis, only 240 proteins were considered to have a significant profile shift
627 and altered subcellular compartment localization under the condition of symbiosis.
628 Among them, lipid droplets represent the largest known group.

629

630 **Supplementary Figure 7. Proteins involved in nitric oxide signaling and cGMP-**
631 **protein kinase G signaling pathways are localized to lipid droplets and increased**
632 **in abundance in symbiotic cells.**

633 (a) Protein abundance of several key components in the nitric oxide and cGMP-PKG
634 signaling pathways are increased in symbiotic cells. *, p-value < 0.05, unpaired
635 Student's t-test. (b) Model of signal transduction and lipid transport in symbiotic *P.*
636 *bursaria* cells. Nitric oxide generated by nitric oxide synthase activates guanylate
637 cyclase to catalyze GTP into cGMP and regulate the expression of lipocalin, a protein
638 involved in lipid transport. cGMP then activates protein kinase G to perform
639 downstream phosphorylation signaling. In addition, cGMP can activate the CNG

640 channel to regulate the concentration of cations inside and outside cells. Interestingly,
641 we observed that the phosphodiesterase enzyme in the cGMP signaling pathway was
642 located in the ER of aposymbiotic cells, but was upregulated and translocated to lipid
643 droplets in symbiotic cells, indicating functional adjustment of the lipid droplets.
644

645 Supplementary Data

646 **Supplementary Data 1: Marker protein list.**

647 **Supplementary Data 2: Normalized fractionation profiles.**

648 **Supplementary Data 3: DeepLoc-predicted protein localizations.**

649 **Supplementary Data 4: Translocated proteins identified using the MR analysis.**

650 **Supplementary Data 5: GO analysis of symbiotic cell-specific lipid droplet**
651 **proteins.**

652 **Supplementary Data 6: Q-Sep distances between all pairs of compartments in**
653 **DK2g and DK2w cells.**

654 **Supplementary Data 7: Manhattan distance of each replicate profile.**

655 **Supplementary Data 8: Mahalanobis distances of individual protein profiles.**

656 **Supplementary Data 9: Pearson correlation of all pairs of delta profiles.**

657 **Supplementary Data 10: Cosine correlation of each protein among replicates.**

658

659 References

660 Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome
661 structure and function. Nature *537*, 347-355.
662 Ayala, A., and Weis, D.S. (1987). The Effect of Puromycin and Cycloheximide on the
663 Infection of Algae-Free Paramecium-Bursaria by Symbiotic Chlorellae. Journal of
664 Protozoology *34*, 377-381.
665 Blum, M., Chang, H.Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A.,
666 Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S.*, et al.* (2021). The InterPro protein
667 families and domains database: 20 years on. Nucleic Acids Res *49*, D344-D354.
668 Brueckner, J., and Martin, W.F. (2020). Bacterial Genes Outnumber Archaeal Genes in

669 Eukaryotic Genomes. Genome Biol Evol *12*, 282-292.
670 Chen, C.W., Tsai, C.F., Lin, M.H., Lin, S.Y., and Hsu, C.C. (2023). Suspension
671 Trapping-Based Sample Preparation Workflow for In-Depth Plant Phosphoproteomics.
672 Anal Chem *95*, 12232-12239.
673 Cheng, Y.H., Liu, C.F., Yu, Y.H., Jhou, Y.T., Fujishima, M., Tsai, I.J., and Leu, J.Y.
674 (2020). Population genomics reveals genome plasticity in Paramecium bursaria.
675 submitted.
676 de Alda, J.A.G.O., Esteban, R., Diago, M.L., and Houmard, J. (2014). The plastid
677 ancestor originated among one of the major cyanobacterial lineages. Nature
678 Communications *5*.
679 Deutsch, E.W., Bandeira, N., Perez-Riverol, Y., Sharma, V., Carver, J.J., Mendoza, L.,
680 Kundu, D.J., Wang, S.B., Bandla, C., Kamatchinathan, S.*, et al.* (2023). The
681 ProteomeXchange consortium at 10 years: 2023 update. Nucleic Acids Research *51*,
682 D1539-D1548.
683 Douglas, S.E. (1998). Plastid evolution: origins, diversity, trends. Curr Opin Genet Dev
684 *8*, 655-661.
685 Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole
686 genome comparisons dramatically improves orthogroup inference accuracy. Genome
687 Biology *16*.
688 Esteban, G.F., Fenchel, T., and Finlay, B.J. (2010). Mixotrophy in ciliates. Protist *161*,
689 621-641.
690 Fujishima, M. (2009). Endosymbionts in Paramecium, Vol 12 (Springer).
691 Fujishima, M., and Kodama, Y. (2012). Endosymbionts in Paramecium. Eur J Protistol
692 *48*, 124-137.
693 Gao, Y., Solberg, T., Wang, R., Yu, Y., Al-Rasheid, K.A.S., and Gao, F. (2024).
694 Application of RNA interference and protein localization to investigate housekeeping
695 and developmentally regulated genes in the emerging model protozoan Paramecium
696 caudatum. Commun Biol *7*, 204.
697 Geladaki, A., Britovsek, N.K., Breckels, L.M., Smith, T.S., Vennard, O.L., Mulvey,
698 C.M., Crook, O.M., Gatto, L., and Lilley, K.S. (2019). Combining LOPIT with
699 differential ultracentrifugation for high-resolution spatial proteomics. Nature
700 Communications *10*.
701 Gligorijevic, V., Renfrew, P.D., Kosciolek, T., Leman, J.K., Berenberg, D., Vatanen, T.,
702 Chandler, C., Taylor, B.C., Fisk, I.M., Vlamakis, H.*, et al.* (2021). Structure-based
703 protein function prediction using graph convolutional networks. Nat Commun *12*, 3168.
704 Gurung, S., Perocheau, D., Touramanidou, L., and Baruteau, J. (2021). The exosome
705 journey: from biogenesis to uptake and intracellular signalling. Cell Communication
706 and Signaling *19*.

Hamada, M., Schroder, K., Bathia, J., Kurn, U., Fraune, S., Khalturina, M., Khalturin, K., Shinzato, C., Satoh, N., and Bosch, T.C.G. (2018). Metabolic co-dependence drives the evolutionarily ancient Hydra-Chlorella symbiosis. Elife *7*.

Hauser, K., Haynes, W.J., Kung, C., Plattner, H., and Kissmehl, R. (2000). Expression of the green fluorescent protein in Paramecium tetraurelia. Eur J Cell Biol *79*, 144-149.

He, M., Wang, J.F., Fan, X.P., Liu, X.H., Shi, W.Y., Huang, N., Zhao, F.Q., and Miao, M. (2019). Genetic basis for the establishment of endosymbiosis in Paramecium. Isme J *13*, 1360-1369.

Ishikawa, M., Yuyama, I., Shimizu, H., Nozawa, M., Ikeo, K., and Gojobori, T. (2016). Different Endosymbiotic Interactions in Two Hydra Species Reflect the Evolutionary History of Endosymbiosis. Genome Biol Evol *8*, 2155-2163.

Itzhak, D.N., Schessner, J.P., and Borner, G.H.H. (2019). Dynamic Organellar Maps for Spatial Proteomics. Curr Protoc Cell Biol *83*, e81.

Itzhak, D.N., Tyanova, S., Cox, J., and Borner, G.H. (2016). Global, quantitative and dynamic mapping of protein subcellular localization. Elife *5*.

Jenkins, B.H., Maguire, F., Leonard, G., Eaton, J.D., West, S., Housden, B.E., Milner, D.S., and Richards, T.A. (2021). Emergent RNA-RNA interactions can promote stability in a facultative phototrophic endosymbiosis. P Natl Acad Sci USA *118*.

Kissmehl, R., Treptau, T., Hofer, H.W., and Plattner, H. (1996). Protein phosphatase and kinase activities possibly involved in exocytosis regulation in Paramecium tetraurelia. Biochemical Journal *317*, 65-76.

Klingelhuber, F., Frendo-Cumbo, S., Omar-Hmeadi, M., Massier, L., Kakimoto, P., Taylor, A.J., Couchet, M., Ribicic, S., Wabitsch, M., Messias, A.C.*, et al.* (2024). A spatiotemporal proteomic map of human adipogenesis. Nat Metab *6*, 861-879.

Kodama, Y., and Fujishima, M. (2008). Cycloheximide induces synchronous swelling of perialgal vacuoles enclosing symbiotic Chlorella vulgaris and digestion of the algae in the ciliate Paramecium bursaria. Protist *159*, 483-494.

Kodama, Y., and Fujishima, M. (2022). Endosymbiotic Chlorella variabilis reduces mitochondrial number in the ciliate Paramecium bursaria. Sci Rep *12*, 8216.

Kodama, Y., Inouye, I., and Fujishima, M. (2011). Symbiotic Chlorella vulgaris of the ciliate Paramecium bursaria plays an important role in maintaining perialgal vacuole membrane functions. Protist *162*, 288-303.

Kodama, Y., and Sumita, H. (2022). The ciliate allows budding of symbiotic cells singly from the digestive vacuole membrane into the cytoplasm during algal reinfection. Protoplasma *259*, 117-125.

Licknack, T.J. (2022). A spatial proteome of Paramecium tetraurelia (ASU Library: Arizona State University).

Londos, C., Sztalryd, C., Tansey, J.T., and Kimmel, A.R. (2005). Role of PAT proteins

in lipid metabolism. Biochimie *87*, 45-49.
Long, H.G., Johri, P., Gout, J.F., Ni, J.H., Hao, Y., Licknack, T., Wang, Y.H., Pan, J., Jiménez-Marín, B., and Lynch, M. (2023). Genetics, Genomics, and Evolution. Annual Review of Genetics *57*, 391-410.
Lynch, M., Schavemaker, P.E., Licknack, T.J., Hao, Y., and Pezzano, A. (2022). Evolutionary bioenergetics of ciliates. J Eukaryot Microbiol *69*.
Meier, R., and Wiessner, W. (1988). Infection of algae-free Paramecium bursaria with symbiotic Chlorella sp. Isolated from green paramecia: I. Effect of the incubation period. Eur J Protistol *24*, 69-74.
Mota, L.A.M., Roberto-Neto, J., Monteiro, V.G., Lobato, C.S.S., de Oliveira, M.A., da Cunha, M., D'Avila, H., Seabra, S.H., Bozza, P.T., and DaMatta, R.A. (2014). Culture of mouse peritoneal macrophages with mouse serum induces lipid bodies that associate with the parasitophorous vacuole and decrease their microbicidal capacity against. Mem I Oswaldo Cruz *109*, 767-774.
Mund, A., Brunner, A.D., and Mann, M. (2022). Unbiased spatial proteomics with single-cell resolution in tissues. Molecular Cell *82*, 2335-2349.
Nolan, S.J., Romano, J.D., and Coppens, I. (2017). Host lipid droplets: An important source of lipids salvaged by the intracellular parasite Toxoplasma gondii. PLoS Pathog *13*, e1006362.
Olzmann, J.A., and Carvalho, P. (2019). Dynamics and functions of lipid droplets. Nat Rev Mol Cell Bio *20*, 137-155.
Pease, D.C. (1947). The structure of trichocysts revealed by the electron microscope. J Cell Comp Physiol *29*, 91-94.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.*, et al.* (2011). Scikit-learn: Machine Learning in Python. J Mach Learn Res *12*, 2825-2830.
Pirayesh, N., Giridhar, M., Ben Khedher, A., Vothknecht, U.C., and Chigri, F. (2021). Organellar calcium signaling in plants: An update. Biochim Biophys Acta Mol Cell Res *1868*, 118948.
Plattner, H. (2017). Trichocysts-Paramecium's Projectile-like Secretory Organelles Reappraisal of their Biogenesis, Composition, Intracellular Transport, and Possible Functions. J Eukaryot Microbiol *64*, 106-133.
Rivera, M.C., Jain, R., Moore, J.E., and Lake, J.A. (1998). Genomic evidence for two functionally distinct gene classes. Proc Natl Acad Sci U S A *95*, 6239-6244.
Rogne, M., and Taskén, K. (2014). Compartmentalization of cAMP Signaling in Adipogenesis, Lipogenesis, and Lipolysis. Hormone and Metabolic Research *46*, 833-840.
Roingeard, P., and Melo, R.C. (2017). Lipid droplet hijacking by intracellular pathogens.

783 Cell Microbiol *19*.
784 Schessner, J.P., Albrecht, V., Davies, A.K., Sinitcyn, P., and Borner, G.H.H. (2023).
785 Deep and fast label-free Dynamic Organellar Mapping. Nat Commun *14*, 5252.
786 Song, C.H., Murata, K., and Suzaki, T. (2017). Intracellular symbiosis of algae with
787 possible involvement of mitochondrial dynamics. Scientific Reports *7*.
788 Su, E.C., Chiu, H.S., Lo, A., Hwang, J.K., Sung, T.Y., and Hsu, W.L. (2007). Protein
789 subcellular localization prediction based on compartment-specific features and
790 structure conservation. BMC Bioinformatics *8*, 330.
791 Sudakaran, S., Kost, C., and Kaltenpoth, M. (2017). Symbiont Acquisition and
792 Replacement as a Source of Ecological Innovation. Trends Microbiol *25*, 375-390.
793 Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable,
794 A.L., Fang, T., Doncheva, N.T., Pyysalo, S.*, et al.* (2023). The STRING database in
795 2023: protein-protein association networks and functional enrichment analyses for any
796 sequenced genome of interest. Nucleic Acids Research *51*, D638-D646.
797 Thumuluri, V., Armenteros, J.J.A., Johansen, A.R., Nielsen, H., and Winther, O. (2022).
798 DeepLoc 2.0: multi-label subcellular localization prediction using protein language
799 models. Nucleic Acids Research *50*, W228-W234.
800 Törönen, P., Medlar, A., and Holm, L. (2018). PANNZER2: a rapid functional
801 annotation web server. Nucleic Acids Research *46*, W84-W88.
802 Walther, T.C., and Farese, R.V., Jr. (2012). Lipid droplets and cellular lipid metabolism.
803 Annu Rev Biochem *81*, 687-714.
804 Wein, T., Romero Picazo, D., Blow, F., Woehle, C., Jami, E., Reusch, T.B.H., Martin,
805 W.F., and Dagan, T. (2019). Currency, Exchange, and Inheritance in the Evolution of
806 Symbiosis. Trends Microbiol *27*, 836-849.
807 Yen, C.L., Stone, S.J., Koliwad, S., Harris, C., and Farese, R.V., Jr. (2008). Thematic
808 review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. J Lipid
809 Res *49*, 2283-2301.
810 Zilber-Rosenberg, I., and Rosenberg, E. (2008). Role of microorganisms in the
811 evolution of animals and plants: the hologenome theory of evolution. FEMS Microbiol
812 Rev *32*, 723-735.

813

Figure 1

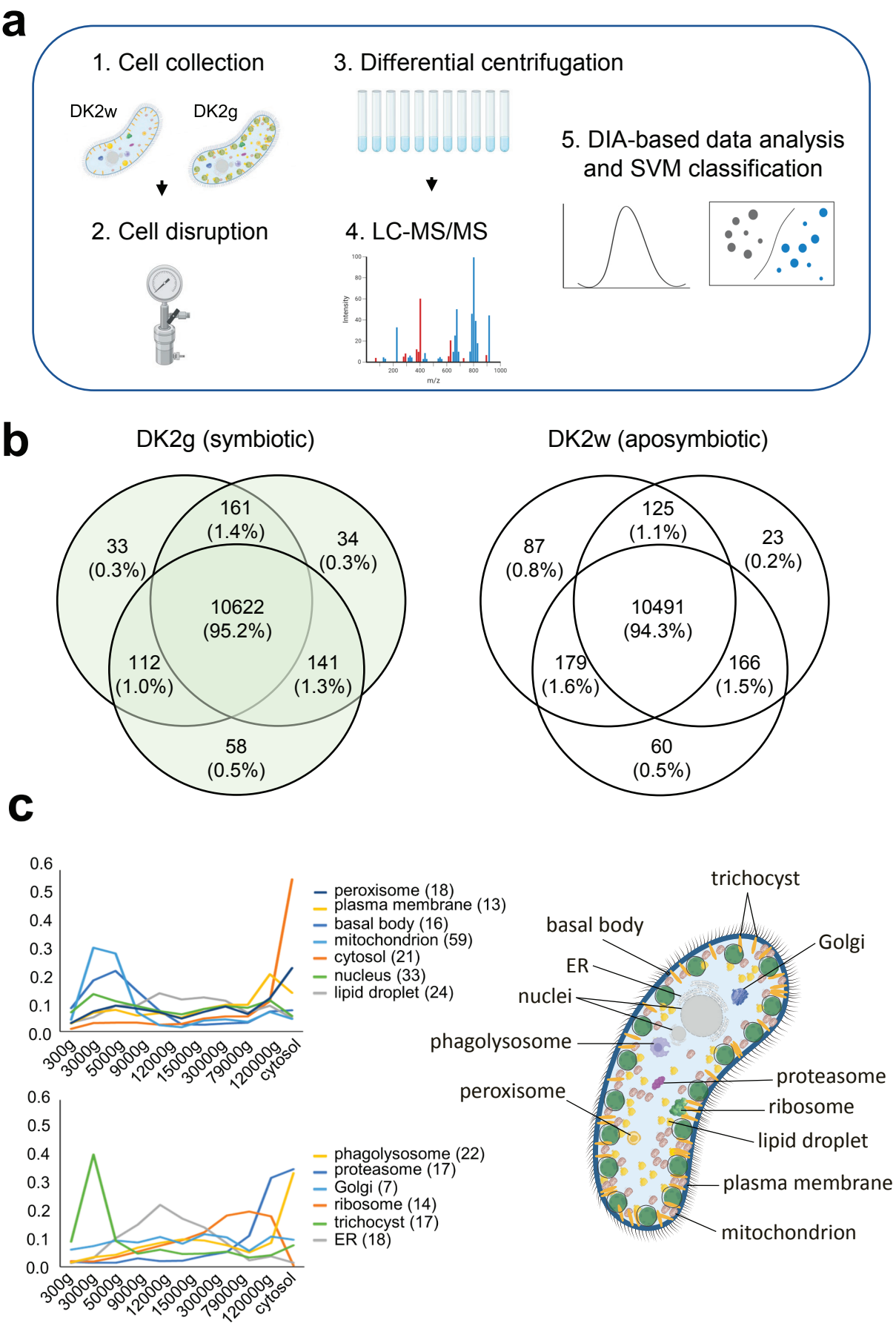

a
1. Cell collection
DK2w
DK2g
2. Cell disruption
3. Differential centrifugation
4. LC-MS/MS
Intensity
m/z
5. DIA-based data analysis and SVM classification
b
DK2g (symbiotic)
33 (0.3%)
161 (1.4%)
34 (0.3%)
10622 (95.2%)
112 (1.0%)
141 (1.3%)
58 (0.5%)
DK2w (aposymbiotic)
87 (0.8%)
125 (1.1%)
23 (0.2%)
10491 (94.3%)
179 (1.6%)
166 (1.5%)
60 (0.5%)
c
peroxisome (18)
plasma membrane (13)
basal body (16)
mitochondrion (59)
cytosol (21)
nucleus (33)
lipid droplet (24)
phagolysosome (22)
proteasome (17)
Golgi (7)
ribosome (14)
trichocyst (17)
ER (18)
300g
3000g
5000g
9000g
12000g
15000g
30000g
79000g
120000g
cytosol
trichocyst
basal body
Golgi
ER
nuclei
phagolysosome
proteasome
peroxisome
ribosome
lipid droplet
plasma membrane
mitochondrion

Figure 2

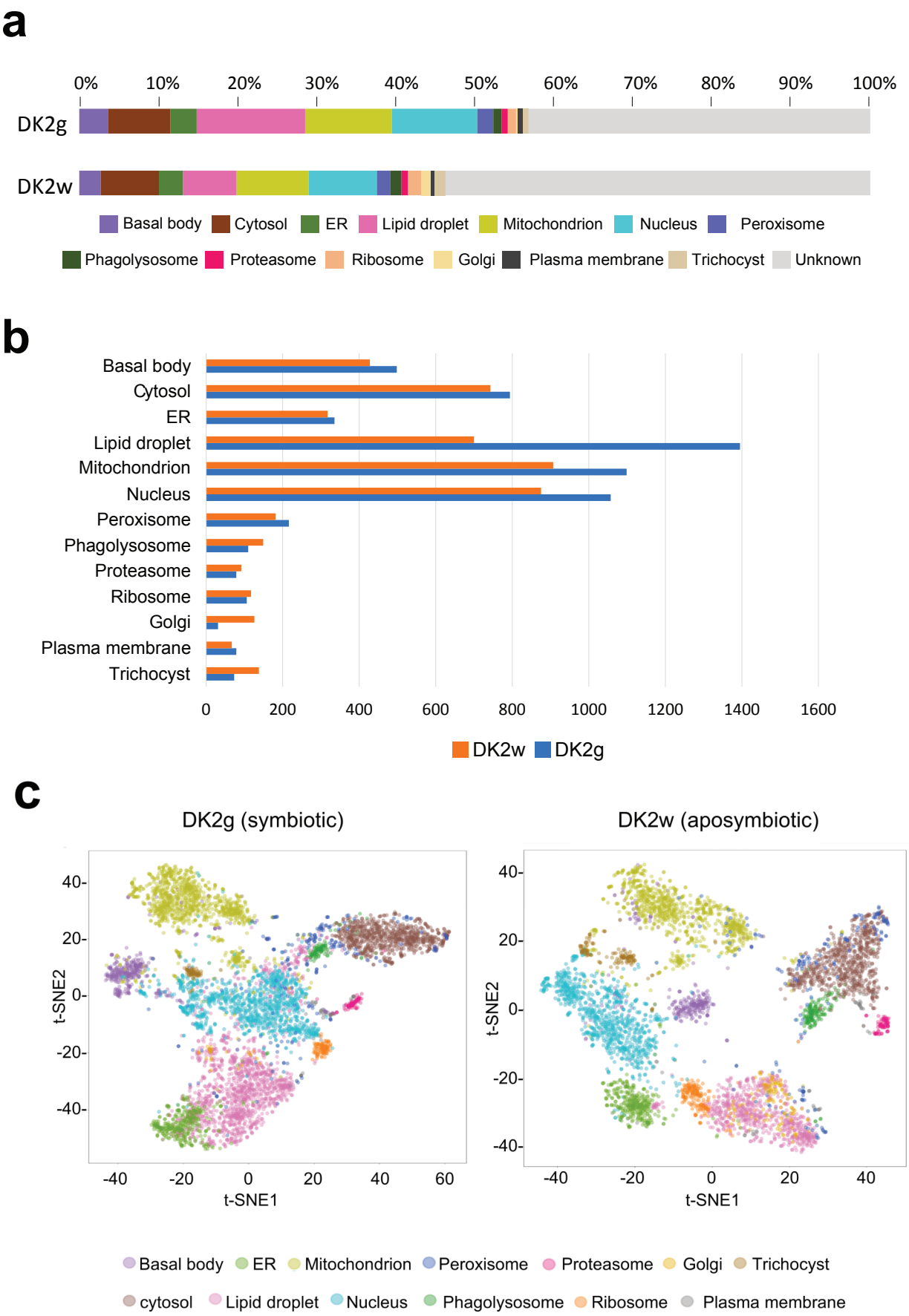

a
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
DK2g
DK2w
Basal body
Cytosol
ER
Lipid droplet
Mitochondrion
Nucleus
Peroxisome
Phagolysosome
Proteasome
Ribosome
Golgi
Plasma membrane
Trichocyst
Unknown
b
Basal body
Cytosol
ER
Lipid droplet
Mitochondrion
Nucleus
Peroxisome
Phagolysosome
Proteasome
Ribosome
Golgi
Plasma membrane
Trichocyst
0
200
400
600
800
1000
1200
1400
1600
DK2w
DK2g
c
DK2g (symbiotic)
DK2w (aposymbiotic)
t-SNE2
t-SNE1
40
20
0
-20
-40
-40
-20
0
20
40
60
Basal body
ER
Mitochondrion
Peroxisome
Proteasome
Golgi
Trichocyst
cytosol
Lipid droplet
Nucleus
Phagolysosome
Ribosome
Plasma membrane

Figure 3

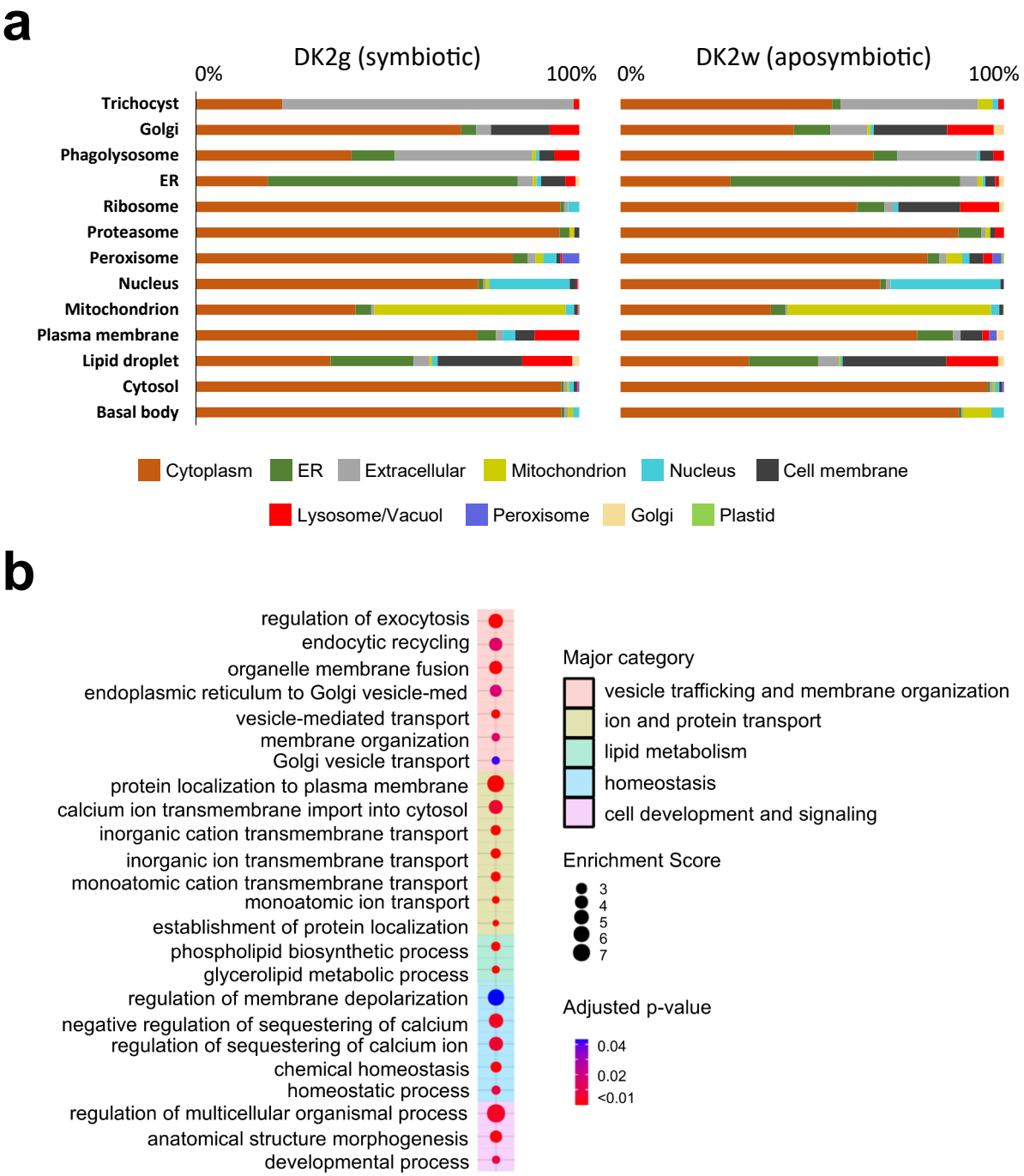

a
DK2g (symbiotic)
DK2w (aposymbiotic)
0%
100%
0%
100%
Trichocyst
Golgi
Phagolysosome
ER
Ribosome
Proteasome
Peroxisome
Nucleus
Mitochondrion
Plasma membrane
Lipid droplet
Cytosol
Basal body
Cytoplasm
ER
Extracellular
Mitochondrion
Nucleus
Cell membrane
Lysosome/Vacuol
Peroxisome
Golgi
Plastid
b
regulation of exocytosis
endocytic recycling
organelle membrane fusion
endoplasmic reticulum to Golgi vesicle-med
vesicle-mediated transport
membrane organization
Golgi vesicle transport
protein localization to plasma membrane
calcium ion transmembrane import into cytosol
inorganic cation transmembrane transport
inorganic ion transmembrane transport
monoatomic cation transmembrane transport
monoatomic ion transport
establishment of protein localization
phospholipid biosynthetic process
glycerolipid metabolic process
regulation of membrane depolarization
negative regulation of sequestering of calcium
regulation of sequestering of calcium ion
chemical homeostasis
homeostatic process
regulation of multicellular organismal process
anatomical structure morphogenesis
developmental process
Major category
vesicle trafficking and membrane organization
ion and protein transport
lipid metabolism
homeostasis
cell development and signaling
Enrichment Score
3
4
5
6
7
Adjusted p-value
0.04
0.02
<0.01

Figure 4

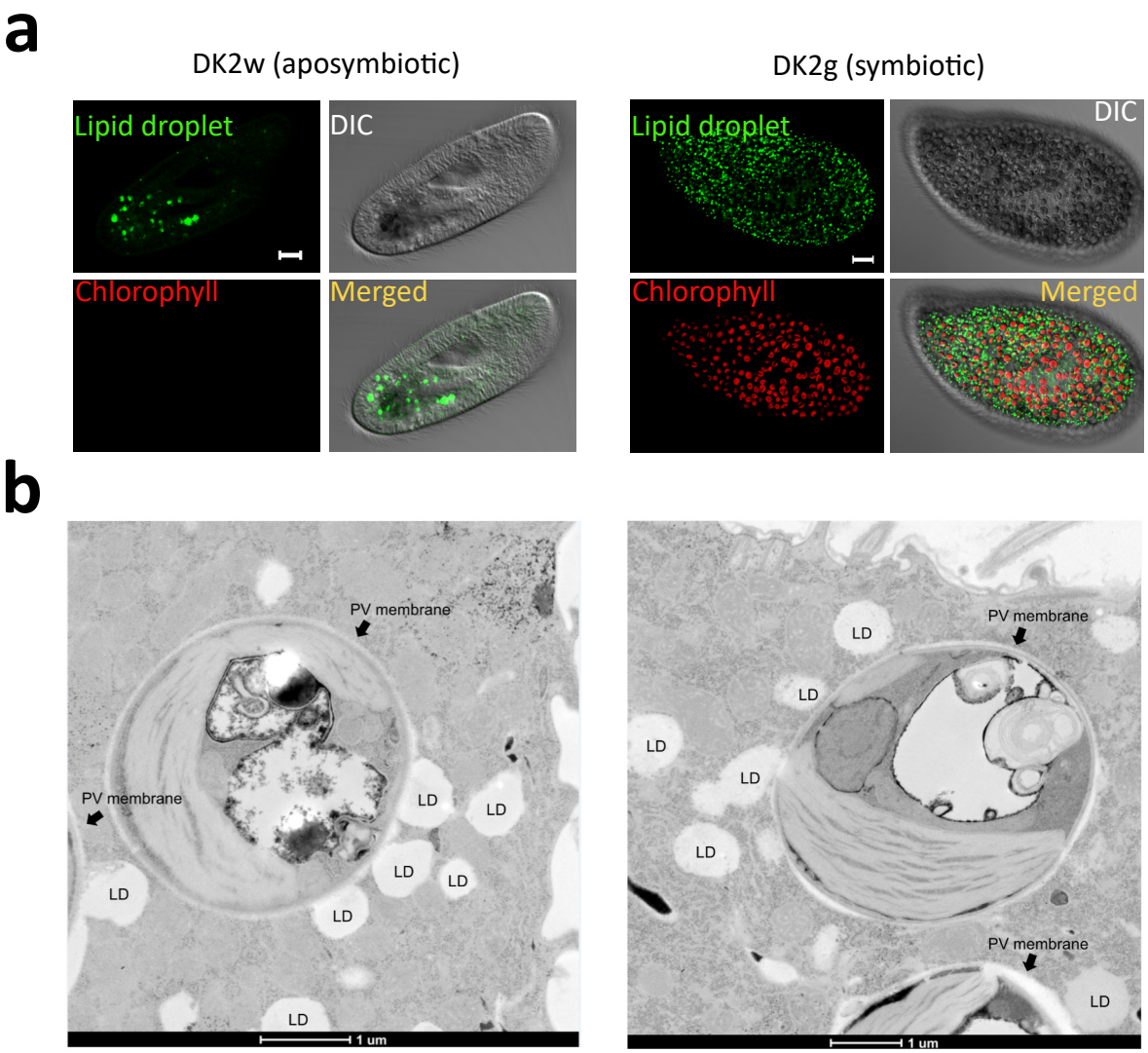
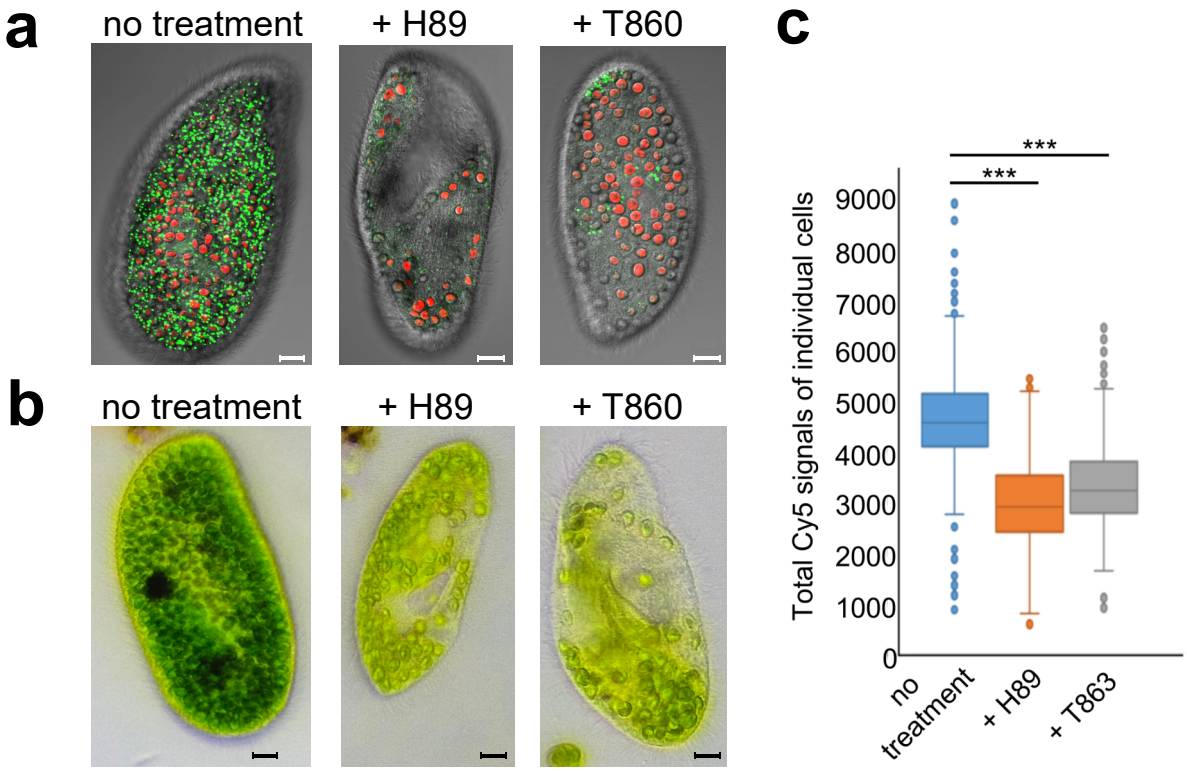
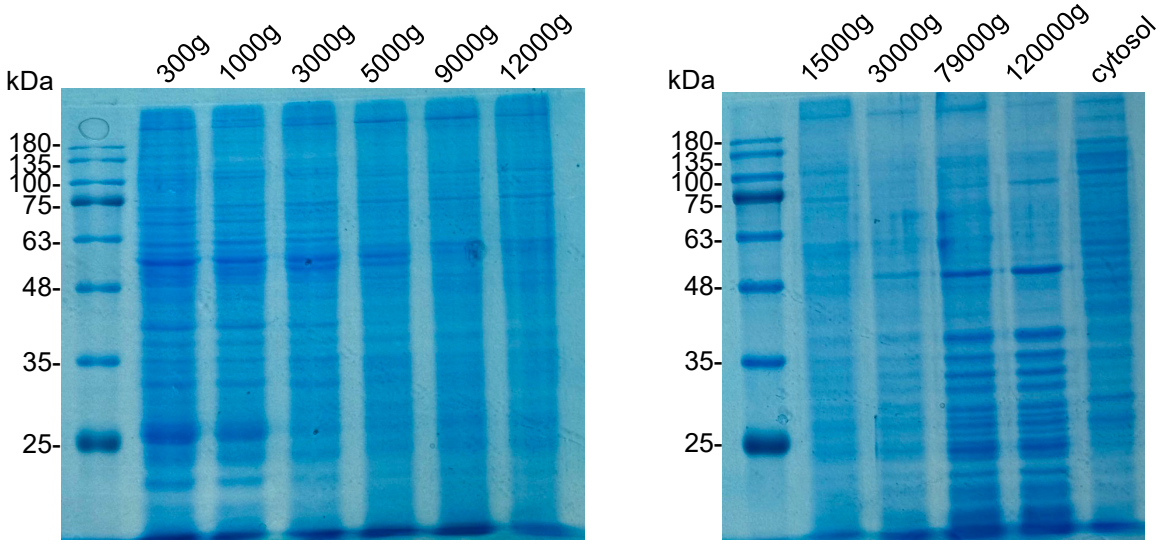

a
DK2w (aposymbiotic)
DK2g (symbiotic)
Lipid droplet
DIC
Chlorophyll
Merged
Lipid droplet
DIC
Chlorophyll
Merged
b
PV membrane
PV membrane
LD
1 um
PV membrane
PV membrane
LD
1 um

Figure 5

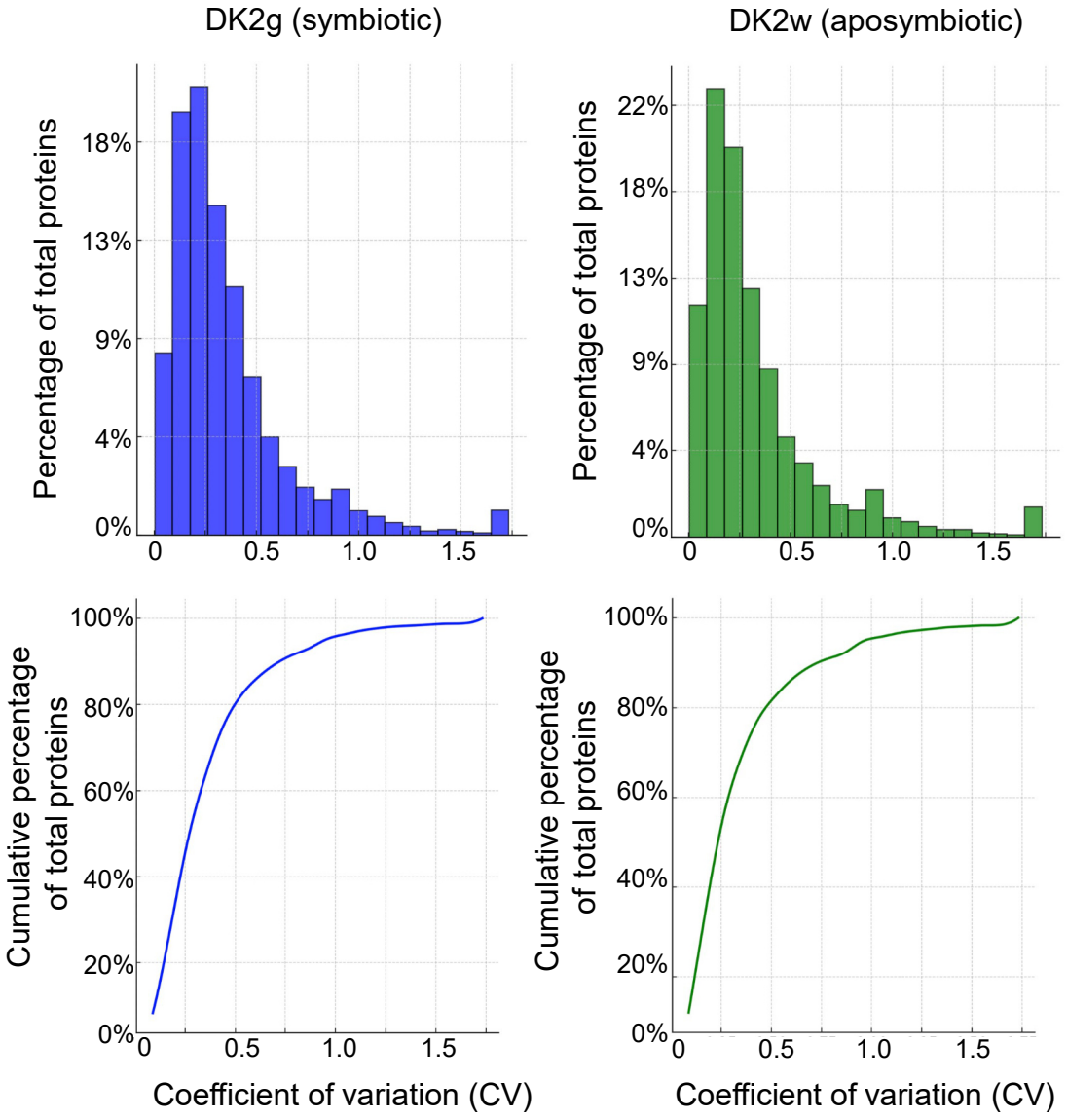

a
no treatment
+ H89
+ T860
b
no treatment
+ H89
+ T860
c
***
***
Total Cy5 signals of individual cells
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
no treatment
+ H89
+ T863

Supplementary Figure 1

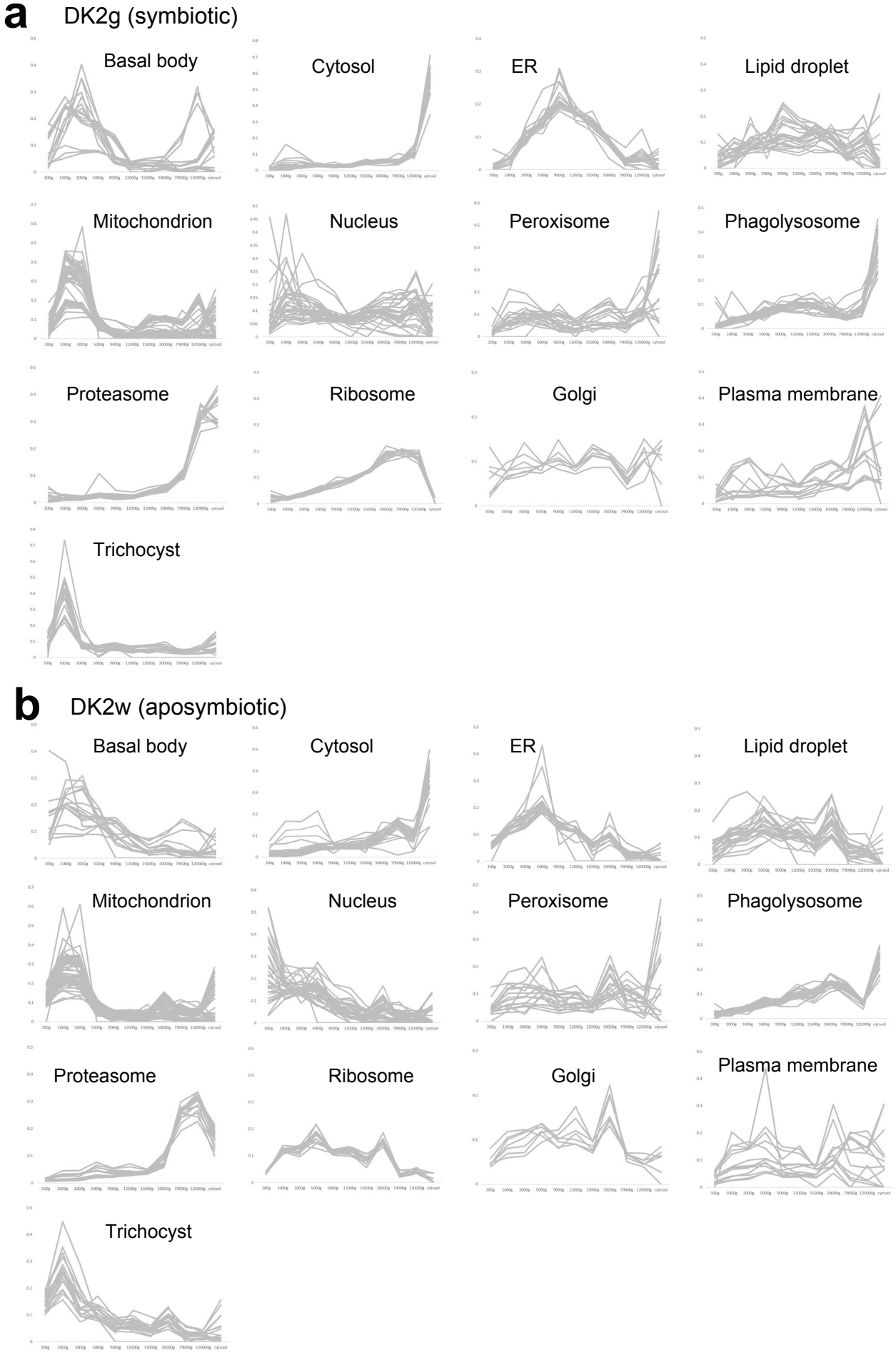
a



b



Supplementary Figure 2



Supplementary Figure 3

a

DK2g (symbiotic)

Total Accuracy on Training Data: 0.992831541218638
Classification Report for Each Marker:

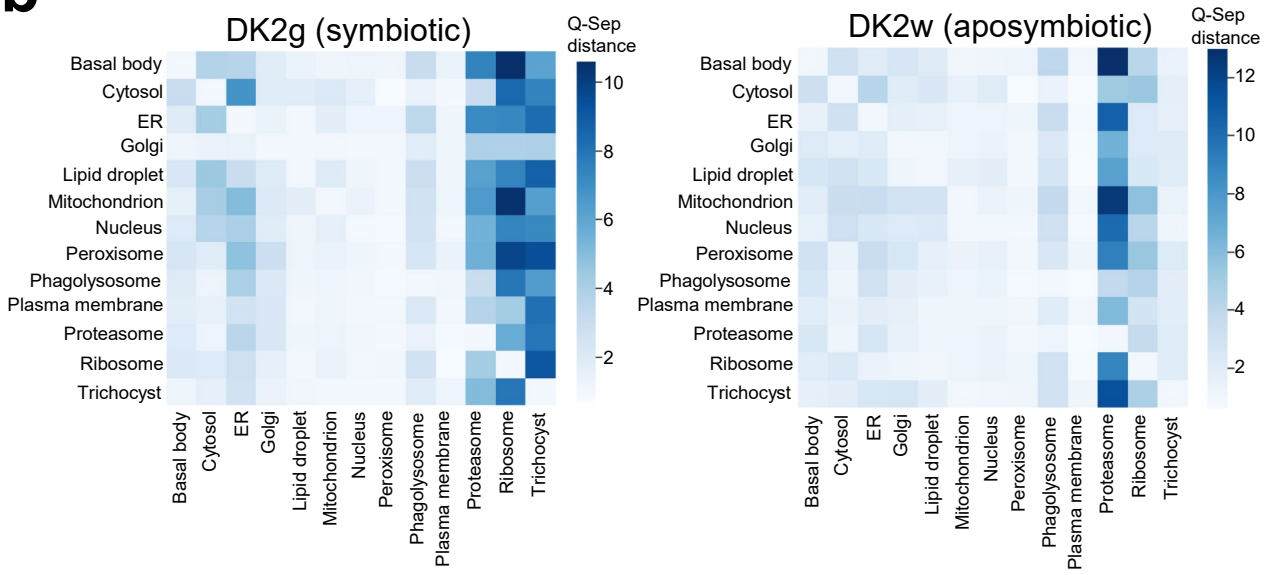
	precision	recall	f1-score	support
basal body rootlet	1.00	1.00	1.00	16
cytosol	1.00	1.00	1.00	21
lipid droplet	1.00	0.92	0.96	24
membrane Trafficking: Vacuole (Trafficking to PM)	1.00	1.00	1.00	13
mitochondria	1.00	1.00	1.00	59
nuclear	0.97	1.00	0.99	33
peroxisome	1.00	1.00	1.00	18
proteasome	1.00	1.00	1.00	17
ribosome	1.00	1.00	1.00	14
trafficking: Vacuole (ER Chaperones)	0.95	1.00	0.97	18
trafficking: Vacuole (Phagolysosomal)	1.00	1.00	1.00	22
trafficking: Vacuole(Trafficking to CV; Golgi)	1.00	1.00	1.00	7
trichocyst matrix	1.00	1.00	1.00	17
accuracy			0.99	279
macro avg	0.99	0.99	0.99	279
weighted avg	0.99	0.99	0.99	279

DK2w (aposymbiotic)

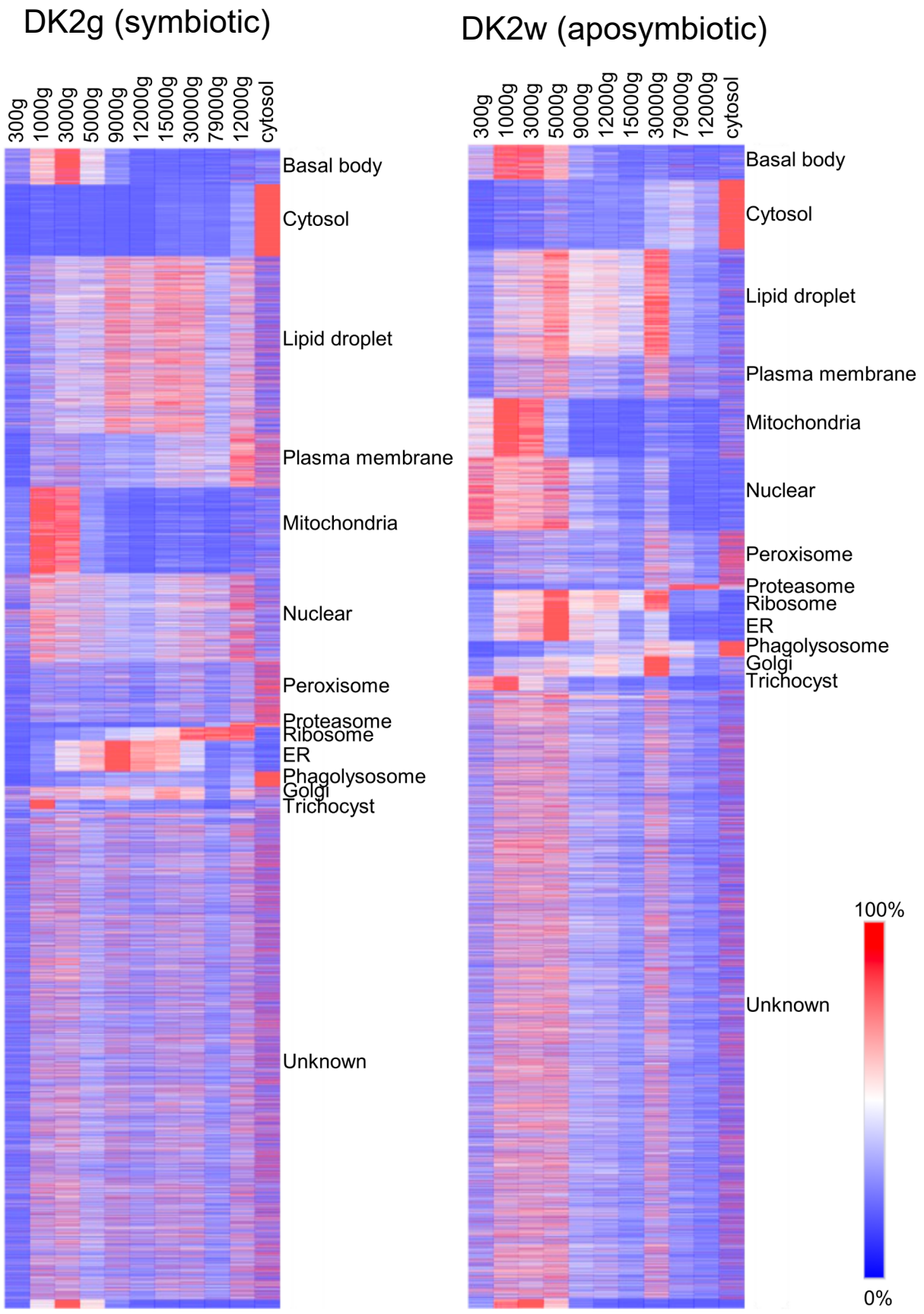
Total Accuracy on Training Data: 0.978494623655914
Classification Report for Each Marker:

	precision	recall	f1-score	support
basal body rootlet	1.00	1.00	1.00	16
cytosol	1.00	1.00	1.00	21
lipid droplet	0.88	0.88	0.88	24
membrane Trafficking: Vacuole (Trafficking to PM)	1.00	1.00	1.00	13
mitochondria	1.00	1.00	1.00	59
nuclear	1.00	1.00	1.00	33
peroxisome	0.95	1.00	0.97	18
proteasome	1.00	1.00	1.00	17
ribosome	1.00	1.00	1.00	14
trafficking: Vacuole (ER Chaperones)	0.95	1.00	0.97	18
trafficking: Vacuole (Phagolysosomal)	1.00	1.00	1.00	22
trafficking: Vacuole(Trafficking to CV; Golgi)	0.80	0.57	0.67	7
trichocyst matrix	1.00	1.00	1.00	17
accuracy			0.98	279
macro avg	0.97	0.96	0.96	279
weighted avg	0.98	0.98	0.98	279

b

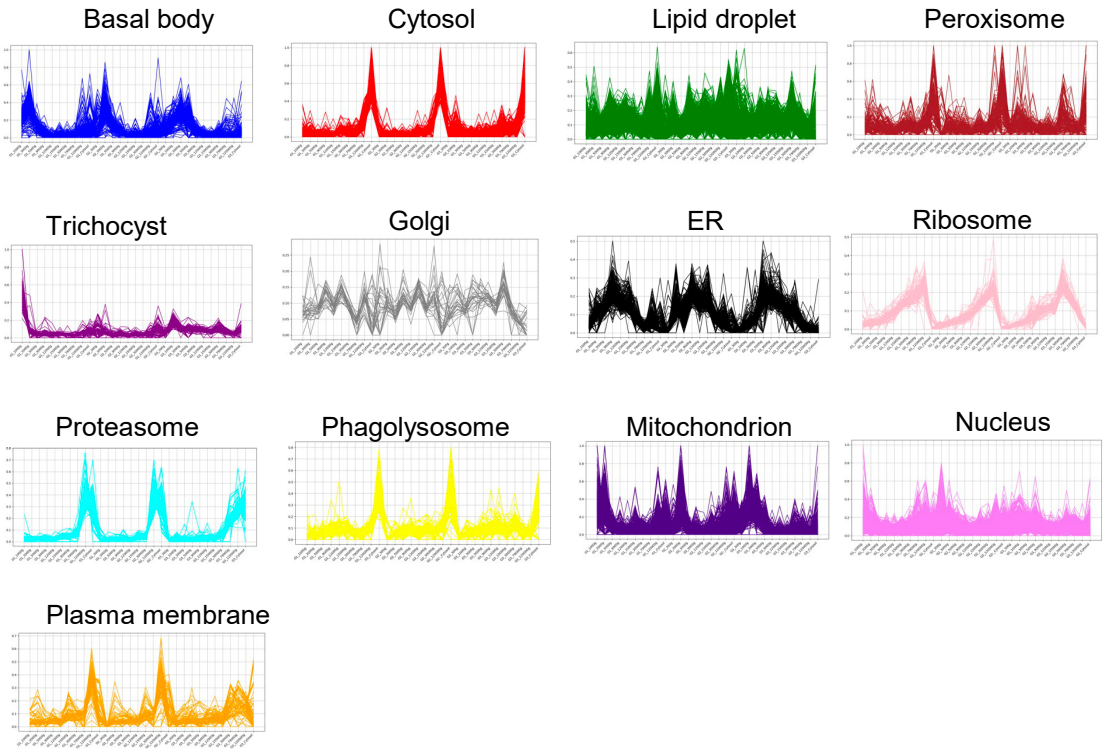


Supplementary Figure 4

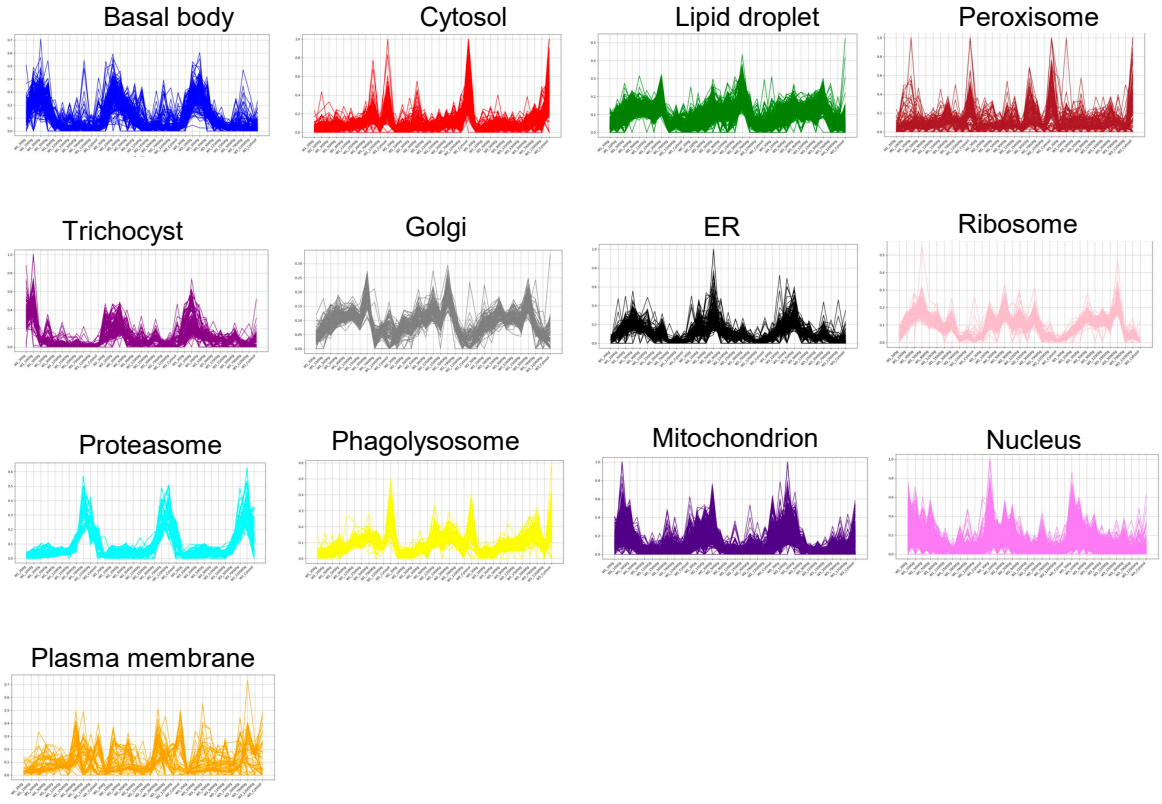


Supplementary Figure 5

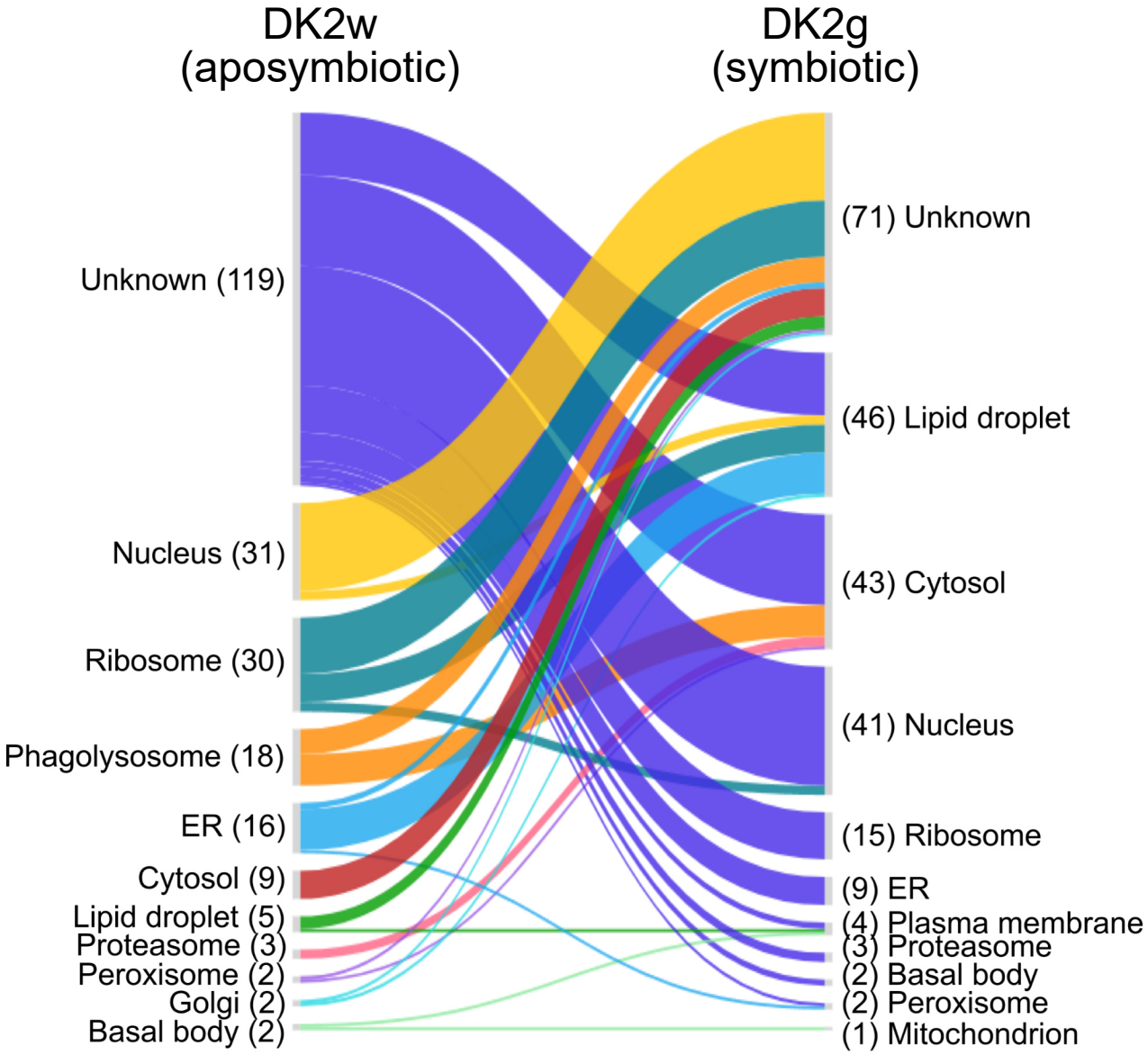
DK2g (symbiotic)



DK2w (aposymbiotic)

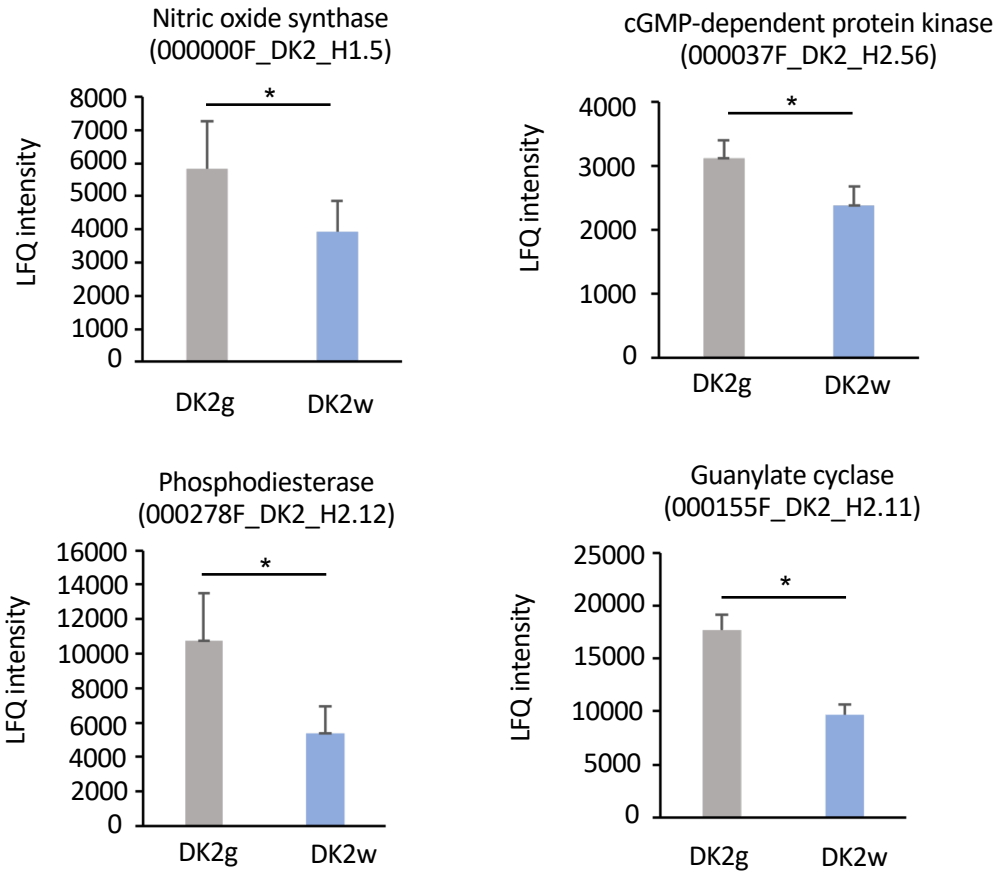
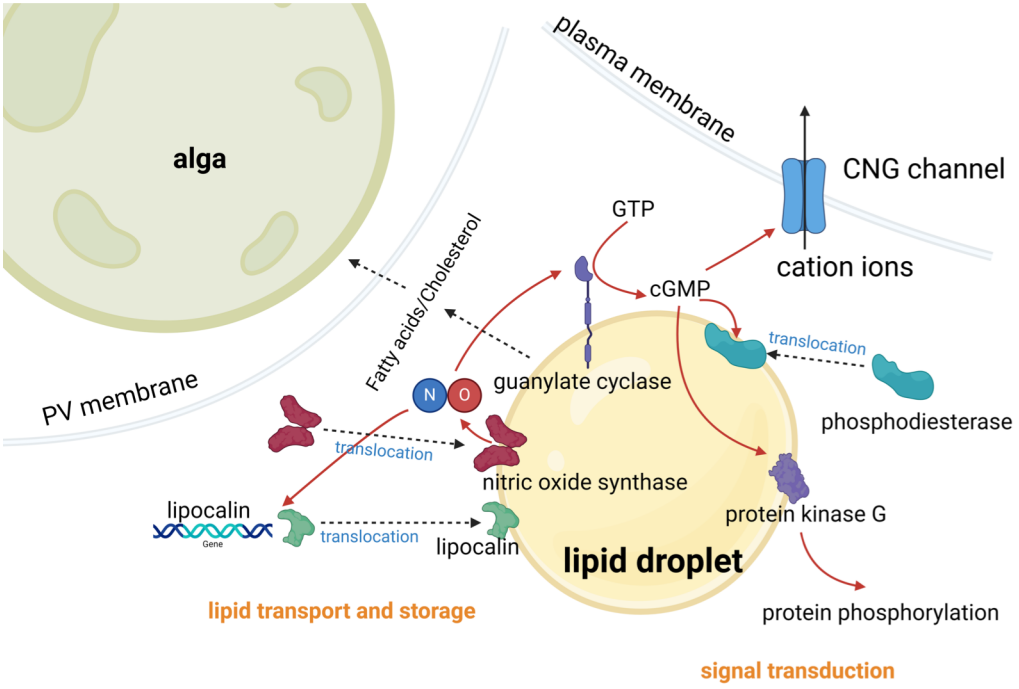


Supplementary Figure 6



Supplementary Figure 7

a



b