



The Puzzle of Metabolite Exchange and Identification of Putative Octotrico Peptide Repeat Expression Regulators in the Nascent Photosynthetic Organelles of *Paulinella chromatophora*

Linda Oberleitner¹, Gereon Poschmann², Luis Macorano¹, Stephan Schott-Verdugo^{3,4}, Holger Gohlke^{3,5}, Kai Stühler^{2,6} and Eva C. M. Nowack^{1*}

OPEN ACCESS

Edited by:

Feng Gao,
Tianjin University, China

Reviewed by:

Hwan Su Yoon,
Sungkyunkwan University,
South Korea
Gennaro Agrimi,
University of Bari Aldo Moro, Italy

*Correspondence:

Eva C. M. Nowack
e.nowack@uni-duesseldorf.de;
e.nowack@hhu.de

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 16 September 2020

Accepted: 05 November 2020

Published: 27 November 2020

Citation:

Oberleitner L, Poschmann G,
Macorano L, Schott-Verdugo S,
Gohlke H, Stühler K and
Nowack ECM (2020) The Puzzle
of Metabolite Exchange
and Identification of Putative Octotrico
Peptide Repeat Expression
Regulators in the Nascent
Photosynthetic Organelles
of *Paulinella chromatophora*.
Front. Microbiol. 11:607182.
doi: 10.3389/fmicb.2020.607182

¹ Department of Biology, Institute of Microbial Cell Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, ² Medical Faculty, Institute for Molecular Medicine, Proteome Research, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, ³ Department of Pharmacy, Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, ⁴ Faculty of Engineering, Centro de Bioinformática y Simulación Molecular, Universidad de Talca, Talca, Chile, ⁵ Jülich Supercomputing Centre, John von Neumann Institute for Computing, Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich GmbH, Jülich, Germany, ⁶ Molecular Proteomics Laboratory, Biologisch-Medizinisches Forschungszentrum, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

The endosymbiotic acquisition of mitochondria and plastids more than one billion years ago was central for the evolution of eukaryotic life. However, owing to their ancient origin, these organelles provide only limited insights into the initial stages of organellogenesis. The cercozoan amoeba *Paulinella chromatophora* contains photosynthetic organelles—termed chromatophores—that evolved from a cyanobacterium ~100 million years ago, independently from plastids in plants and algae. Despite the more recent origin of the chromatophore, it shows tight integration into the host cell. It imports hundreds of nucleus-encoded proteins, and diverse metabolites are continuously exchanged across the two chromatophore envelope membranes. However, the limited set of chromatophore-encoded solute transporters appears insufficient for supporting metabolic connectivity or protein import. Furthermore, chromatophore-localized biosynthetic pathways as well as multiprotein complexes include proteins of dual genetic origin, suggesting that mechanisms evolved that coordinate gene expression levels between chromatophore and nucleus. These findings imply that similar to the situation in mitochondria and plastids, also in *P. chromatophora* nuclear factors evolved that control metabolite exchange and gene expression in the chromatophore. Here we show by mass spectrometric analyses of enriched insoluble protein fractions that, unexpectedly, nucleus-encoded transporters are not inserted into the chromatophore inner envelope membrane. Thus, despite the apparent maintenance of its barrier function, canonical metabolite transporters are missing in this membrane. Instead we identified several expanded groups of short chromatophore-targeted orphan proteins. Members of one of these groups are characterized by a single transmembrane

helix, and others contain amphipathic helices. We hypothesize that these proteins are involved in modulating membrane permeability. Thus, the mechanism generating metabolic connectivity of the chromatophore fundamentally differs from the one for mitochondria and plastids, but likely rather resembles the poorly understood mechanism in various bacterial endosymbionts in plants and insects. Furthermore, our mass spectrometric analysis revealed an expanded family of chromatophore-targeted helical repeat proteins. These proteins show similar domain architectures as known organelle-targeted expression regulators of the octotricopeptide repeat type in algae and plants. Apparently these chromatophore-targeted proteins evolved convergently to plastid-targeted expression regulators and are likely involved in gene expression control in the chromatophore.

Keywords: organellogenesis, metabolite transport, proteome, evolution, cyanobacteria, cercozoa, Rhizaria, envelope membranes

INTRODUCTION

Endosymbiosis has been a major driver for the evolution of cellular complexity in eukaryotes. During organellogenesis, linkage of the previously independent biological networks of the former host and endosymbiont resulted in a homeostatic and synergistic association. Two critical factors during this dauntingly complex process appear to be the establishment of metabolic connectivity between the symbiotic partners, and the evolution of nuclear control over protein expression levels within the organelle.

Besides mitochondria and primary plastids that evolved via endosymbiosis more than one billion years ago, recently, a third organelle of primary endosymbiotic origin has been identified (Nowack, 2014; Gabr et al., 2020). The photosynthetically active “chromatophore” of cercozoan amoeba of the genus *Paulinella* evolved around 100 million years ago from a cyanobacterium (Marin et al., 2005; Delaye et al., 2016). Hence, scrutiny of photosynthetic *Paulinella* species can help to determine the common rules and degrees of freedom in the integration process of a eukaryotic organelle. A method for the genetic manipulation of *P. aulinella* has not been established yet, but genomic, transcriptomic, and proteomic data as well as protein biochemical experimentation already allowed fascinating insights into the relationship between host cell and chromatophore. Similar to the evolution of mitochondria and plastids, also in the chromatophore, reductive genome evolution resulted in the loss of many metabolic functions (Nowack et al., 2008; Reyes-Prieto et al., 2010), around 70 genes were transferred from the chromatophore to the nucleus of the host cell (Nowack et al., 2011, 2016; Zhang et al., 2017), and functions lost from the chromatophore genome are compensated by import of nucleus-encoded proteins (Nowack and Grossman, 2012; Singer et al., 2017). In a previous study, we identified by protein mass spectrometry (MS) around 200 nucleus-encoded, chromatophore-targeted proteins in *Paulinella chromatophora* (Singer et al., 2017) that we refer to as import candidates. These proteins fall into two classes: short import candidates [<90 amino acids (aa)] that lack obvious targeting signals, and long import candidates (>250 aa) that carry a conserved N-terminal

sequence extension—likely a targeting signal—that is referred to as “chromatophore transit peptide” (crTP). Bioinformatic identification of crTPs in a large dataset of translated nuclear transcripts from *P. chromatophora* allowed to extend the catalog of likely chromatophore-targeted proteins to >400 import candidates (Singer et al., 2017).

Metabolic capacities of chromatophore and host cell are highly complementary resulting in the need for extensive exchange of metabolites such as sugars, amino acids, and cofactors across the two envelope membranes that surround the chromatophore (Nowack et al., 2008; Singer et al., 2017; Valadez-Cano et al., 2017). Furthermore, substrates for carbon, sulfur, and nitrogen assimilation (e.g., HCO_3^- , SO_4^{2-} , NH_4^+) and metal ions (e.g., Mg^{2+} , Cu^{2+} , Mn^{2+} , and Co^{2+}) that serve as cofactors of chromatophore-localized proteins have to be imported into the chromatophore. Whereas the chromatophore inner membrane (IM) clearly derives from the cyanobacterial plasma membrane, the outer membrane (OM) has been interpreted as being host-derived (Kies, 1974; Sato et al., 2020). The nature of the transporters underlying the deduced solute (and protein) transport processes across this membrane system is unknown.

In plants and algae, transport across the plastid IM is mediated by a large set of multi-spanning transmembrane (TM) proteins that are highly specific for their substrates. These transporters contain usually four or more TM α -helices (TMHs) and are of the single subunit secondary active or channel type (Facchinelli and Weber, 2011). This set of transporters apparently evolved mainly via the retargeting of existing host proteins to the plastid IM rather than the repurposing of endosymbiont proteins (Facchinelli and Weber, 2011; Fischer, 2011; Karkar et al., 2015). Transport across the plastid OM is enabled largely by (semi-)selective pores formed by nucleus-encoded β -barrel proteins (Breuers et al., 2011).

Another important issue during organellogenesis is the establishment of nuclear control over organellar gene expression supporting (i) adjustment of the organelle to the physiological state of the host cell, and (ii) assembly of organelle-localized protein complexes composed of subunits encoded in either the organellar or nuclear genome in stoichiometric amounts (Woodson and Chory, 2008; Hammani et al., 2014). Also in

P. chromatophora, the import of nucleus-encoded proteins resulted in protein complexes of dual genetic origin (e.g., photosystem I; Nowack and Grossman, 2012). The difference in copy numbers between chromatophore and nuclear genome (~100 vs. one or two copies, Nowack et al., 2016) calls for coordination of gene expression between nucleus and chromatophore.

To test the hypotheses that nuclear factors were recruited to establish (i) metabolic connectivity between chromatophore and host cell and (ii) control over gene expression levels within the chromatophore, here we analyzed the previously obtained proteomic dataset derived from isolated chromatophores and a newly generated proteomic dataset derived from enriched insoluble chromatophore proteins with a focus on chromatophore-targeted TM proteins and putative expression regulators.

MATERIALS AND METHODS

Cultivation of *P. chromatophora* and Chromatophore Isolation

P. chromatophora CCAC0185 (axenic version; Nowack et al., 2016) was grown (Nowack and Grossman, 2012) and chromatophores isolated as described previously (Singer et al., 2017). In brief, *P. chromatophora* cells were washed three times with isolation buffer (50 mM HEPES pH 7.5, 2 mM EGTA, 2 mM MgCl₂, 250 mM sucrose, and 125 mM NaCl) and depleted of dead cells on a discontinuous 20–80% Percoll gradient. The resulting pellet of intact cells was resuspended in isolation buffer, cells were broken in a cell disruptor (Constant Systems) at 0.5 kbar, and intact chromatophores were isolated on another discontinuous 20–80% Percoll gradient. To increase purity, isolated chromatophores were re-isolated from a third Percoll gradient (prepared as before). Recovered chromatophores were washed three times in isolation buffer, supplemented with protease inhibitor cocktail (Roche cOmplete), frozen in liquid nitrogen, and stored at –80°C until further use.

Transmission Electron Microscopy (TEM)

Isolated chromatophores were fixed in isolation buffer containing 1.25% glutaraldehyde for 45 min on ice followed by 30 min post-fixation in 1% OsO₄ in isolation buffer at room temperature. Fixed chromatophores were washed, mixed with 14.5% (w/v) BSA, pelleted, and the pellet was fixed with 2.5% glutaraldehyde for 20 min at room temperature. The fixed pellet was dehydrated in rising concentrations of ethanol (from 60 to 100% at –20°C) and then infiltrated with Epon using propylene oxide as a transition solvent. Epon was polymerized at 60°C for 24 h. 70 nm ultrathin sections were prepared and contrasted with uranyl acetate and lead citrate according to (Reynolds, 1963). A Hitachi H7100 TEM (Hitachi, Tokyo, Japan) with Morada camera (EMSIS GmbH, Münster, Germany) operated at 100 kV was used for TEM analyses. Essentially the same protocol was used for intact *P. chromatophora* cells, however, the isolation buffer was replaced by growth medium (WARIS-H, McFadden and Melkonian, 1986; supplemented with 1.5 mM Na₂SiO₃).

Protein Fractionation

CM and PM Samples

Isolated chromatophores or *P. chromatophora* cells were washed with Buffer I (50 mM HEPES pH 7.5, 125 mM NaCl, 0.5 mM EDTA) at 20,000 × g or 200 × g, respectively. Pellets were resuspended in Buffer I and broken by two passages in a cell disruptor at 2.4 kbar. Lysates were supplemented with 500 mM NaCl (final concentration) and passed five times through a 0.6 mm cannula. Cell debris was removed by two successive centrifugation steps at 15,500 × g. The supernatant was subjected to ultracentrifugation for 1 h at 150,000 × g (Beckmann L-80XL optima ultracentrifuge, Rotor 70.1 Ti at 50,000 rpm). Pellets were resuspended in 100 mM Na₂CO₃ pH > 11 and incubated for 1 h intermitted by 15 passes through a 0.6 mm cannula. Then, insoluble proteins were collected by ultracentrifugation (as before), and subsequently washed with Buffer II (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA) by passage through a 0.6 mm cannula until no particles were visible. Finally, the insoluble fraction was pelleted by ultracentrifugation and solubilized at 36°C in Buffer II supplemented with 1% TritonX-100, 1% Na-deoxycholate, and 0.1% SDS.

CL Samples

Protein was extracted from intact isolated chromatophores by precipitation with 10% trichloroacetic acid for 30 min on ice and pelleted at 21,000 × g for 20 min. Pellets were washed twice with ice cold acetone for 10 min and finally resuspended in Buffer II plus detergents.

Protein concentration was determined in a Neuhaus assay (Neuhaus et al., 1979). Aliquots were supplemented with SDS sample buffer (final conc. 35 mM Tris-HCl pH 7.0, 7.5% Glycerol, 3% SDS, 150 mM DTT, Bromophenol blue), frozen in liquid nitrogen, and stored at –80°C until MS-analysis. All steps were performed at 4°C, protease inhibitor cocktail (Roche cOmplete) was added to all buffers used.

MS Analysis and Protein Identification

Sample preparation and subsequent MS/MS analysis of three independent preparations of CM, PM, and CL samples was essentially carried out as described (Singer et al., 2017). Briefly, proteins were in-gel digested in (per sample) 0.1 µg trypsin in 10 mM ammonium hydrogen carbonate overnight at 37°C and resulting peptides resuspended in 0.1% trifluoroacetic acid. Two independent MS analyses were performed. In MS experiment 1, 500 ng protein per sample, and in MS experiment 2, 500 ng protein per lysate and 1.5 µg protein per membrane sample was analyzed. Peptides were separated on C18 material by liquid chromatography (LC), injected into a QExactive plus mass spectrometer, and the mass spectrometer was operated as described (Singer et al., 2017). Raw files were further processed with MaxQuant (MPI for Biochemistry, Planegg, Germany) for protein identification and quantification using standard parameters. MaxQuant 1.6.2.10 was used for the MS experiment 1 analysis and MaxQuant 1.6.3.4 for MS experiment 2. Searches were carried out using 60,108 sequences translated from a *P. chromatophora* transcriptome and the 867 translated genes predicted on the chromatophore genome (Singer et al., 2017).

Peptides and proteins were accepted at a false discovery rate of 1%. Proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD021087.

Protein Enrichment Analysis

Intensities of individual proteins were normalized by division of individual intensities in each replicate by the sum of intensities of all proteins identified with ≥ 2 peptides in the same replicate. Each protein was assigned an intensity level representing its $\log_{10}$ transformed mean normalized intensity from three replicates in either fraction added 7 ($\log_{10}(\text{normInt}) + 7$), enabling a simple ranking of intensities in a logarithmic range from 0 to 6.

The enrichment factor for each protein in CM as compared to PM or CL samples ($E_{CM/PM}$ or $E_{CM/CL}$, respectively) was calculated as $E_{CM/PM} = \frac{\text{normInt}_{CM}}{\text{normInt}_{PM}}$ or $E_{CM/CL} = \frac{\text{normInt}_{CM}}{\text{normInt}_{CL}}$ [Supplementary Table S1; missing values (intensity = 0) were excluded from the calculation of means]. Proteins identified with at least three spectral counts (SpC) in the chromatophore (i.e., CM + CL fractions) and either $E_{CM/PM} > 1.5$ in at least one of two MS experiments or $0.5 < E_{CM/PM} < 1.5$ in both MS experiments were considered as enriched in chromatophores (see Supplementary Figure S1). Correspondingly, $E_{CM/CL} > 1$ indicate protein enrichment, $E_{CM/CL} < 1$ depletion in CM samples.

Furthermore, a statistic approach was applied to visualize differences between proteins enriched or exclusively found in a certain fraction. In pairwise comparisons, only proteins were considered showing valid *normInt* values in all three replicates of at least one of the samples being compared. *NormInt* values were $\log_2$ transformed and missing values imputed by values from a down shifted normal distribution (width 0.3 SD, down shift 1.8 SD) followed by a pairwise sample comparison based on Student's *t*-tests and the significance analysis of microarrays algorithm ($S_0 = 0.8$, FDR 5%) (Tusher et al., 2001). Differences between individual proteins in CM vs. PM or CM vs. CL samples were calculated as $\log_2(\text{normInt}_{CM}) - \log_2(\text{normInt}_{PM})$ or $\log_2(\text{normInt}_{CM}) - \log_2(\text{normInt}_{CL})$, respectively.

Sequence and Structural Bioinformatics Analyses

TMHs were predicted with TMHMM 2.0 (Krogh et al., 2001) and CCTOP (Dobson et al., 2015), pore-lining residues in TMHs were predicted with MEMSAT-SVM-pore (Nugent and Jones, 2012), and AMP peptides were predicted with AmpGram (Burdukiewicz et al., 2020). Sequence motifs were discovered using MEME 5.0.5 algorithm (Bailey and Elkan, 1994) in classic mode and visualized with WebLogo (Crooks et al., 2004), number and position of motifs in protein sequences were determined with MAST 5.0.5 using default settings (Bailey and Gribskov, 1998). The *P. chromatophora* transcriptome was screened for (i) conserved motifs shown in Figure 4A groups 2 and 3 and (ii) the degenerate 38 aa motif shown in Figure 6C using FIMO 5.0.5 with default settings (Grant et al., 2011). Proteins that contain at least 5 repeats of the 38 aa motif with a *p*-value $< e^{-10}$ and/or at

least 1 repeat with a *p*-value $< e^{-20}$ were considered candidate OPR-proteins. α -helices in repetitive elements or AMP-like proteins were predicted with Jpred4 (Drozdetskiy et al., 2015) and NetSurfP-2.0 (Schantz Klausen et al., 2019), respectively. Helical wheel projections were created with HeliQuest (Gautier et al., 2008). Functional protein domains were found with DELTA-BLAST (Boratyn et al., 2012). Targeting signals were predicted with PredAlgo (Tardif et al., 2012) for CrRAP and CrTab1, and TargetP 2.0 (Almagro Armenteros et al., 2019), WoLFPSORT (Horton et al., 2007), and Predotar (Small et al., 2004) for *P. chromatophora* proteins. Tertiary structure predictions were obtained using Phyre2 (Kelley et al., 2015) in normal mode. Area-proportional Venn diagrams were calculated with eulerAPE (Micallef and Rodgers, 2014).

Transporters were classified according to the Transporter Classification Database (Saier et al., 2016). Complete lists of the transporters depicted in Figures 1, 2D and methods for their identification and classification are provided in Supplementary Table S2. No OM porins could be identified in the chromatophore genome based on sequence similarity or topology predictions using MCMBB (Bagos et al., 2004).

RESULTS

Paucity of Chromatophore-Encoded Solute Transporters

Although diverse metabolites have to be exchanged constantly between the chromatophore and cytoplasm, we identified genes for only 25 solute transporters on the chromatophore genome (Nowack et al., 2008; Singer et al., 2017; Valadez-Cano et al., 2017). As judged from the localization of their cyanobacterial orthologs, only 19 of these transporters putatively localize to the chromatophore IM, whereas three likely localize to thylakoids and for the remaining three the localization could

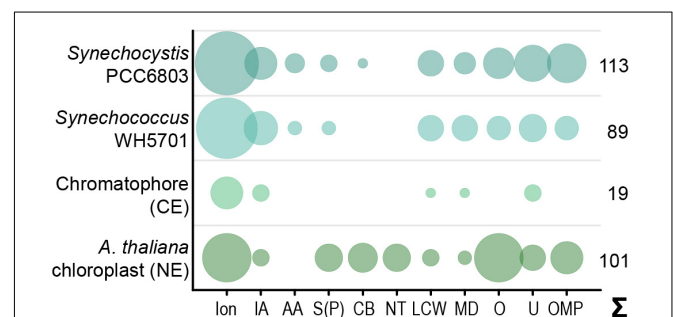


FIGURE 1 | Predicted solute transport capacities of the chromatophore, *Synechococcus* sp. WH5701, *Synechocystis* sp. PCC6803, and the *Arabidopsis thaliana* chloroplast. Only transport systems for which experimental evidence suggests localization to the plasma membrane or the organellar envelope are shown. CE, chromatophore-encoded; NE, nucleus-encoded; Ion, ions/metals; IA, inorganic anions (phosphate, sulfate, nitrate, bicarbonate); AA, amino acid; S(P), sugars (hexoses, oligosaccharides) or sugar-phosphates; CB, mono-/di-/tricarboxylates; NT, nucleotides; LCW, lipid and lipopolysaccharide; MD, multidrug; O, other; U, unknown; OMP, outer membrane pores; Σ, total predicted transporters.

not be determined (**Figure 1** and **Supplementary Table S2**). In comparison, in *Synechococcus* sp. WH5701, a free-living relative of the chromatophore (Marin et al., 2007), and the model cyanobacterium *Synechocystis* sp. PCC6803, genes for ~89 and >100 putative envelope transporters were identified, respectively (**Figure 1** and **Supplementary Table S2**; Paulsen et al., 2000). Substrates of most of the chromatophore IM transporters are—according to annotation—restricted to inorganic ions (e.g., Na^+ , K^+ , Fe^{2+} , Mg^{2+} , PO_4^{2-} , HCO_3^-). Notably, cyanobacterial uptake systems for nitrogen and sulfur compounds such as nitrate (Omata et al., 1993), ammonium (Montesinos et al., 1998), urea (Valladares et al., 2002), amino acids (Quintero et al., 2001), or sulfate (Laudenbach and Grossman, 1991) are missing. Only one transporter of the DME-family (10 TMS Drug/Metabolite Exporter; PCC0734) could potentially be involved in metabolite export, and one transporter of the DASS-family (Divalent Anion: Na^+ Symporter; PCC0664) could facilitate import of either di-/tricarboxylates or sulfate via Na^+ symport. However, due to the multitude of substrates transported by members of both families (Jack et al., 2001; Markovich, 2012), precise substrate specificities cannot be predicted. Chromatophore-encoded β -barrel OM pores could not be identified.

In contrast, in plants and algae, a combination of bioinformatic and proteomic studies identified 100–150 putative solute transporters in the plastid IM; 37 of these transporters have been confidently assigned functions and many of them transport metabolites (Weber et al., 2005; Mehrshahi et al., 2013; Karkar et al., 2015; Marchand et al., 2018; **Figure 1** and **Supplementary Table S2**). Several porins are known to permit passage of solutes across the chloroplast OM (Breuers et al., 2011; Wang et al., 2013; Goetze et al., 2015; Harsman et al., 2016). Almost all of these transport systems are encoded in the nucleus and post-translationally inserted into the plastid envelope membranes.

Enrichment of Insoluble Protein Fractions and Proteomic Analysis

The scarcity of chromatophore-encoded solute transporters suggested that in *P. chromatophora*, as in plastids, nucleus-encoded transport systems establish metabolic connectivity of the chromatophore. However, among 432 previously identified import candidates (Singer et al., 2017), only 3 proteins contained more than one predicted TMH (**Table 1**). One of these proteins (identified by *in silico* prediction, i.e., bioinformatic identification of the crTP) contains two TMHs, only one of which is predicted with high confidence. Of the other two proteins (identified by MS), one is short and contains two predicted TMHs; the other contains eight predicted TMHs. However, this latter protein was identified with one peptide only and shows no BlastP hits against the NCBI nr database, whereas an alternative ORF (in the reverse complement) shows similarity to an NAD-dependent epimerase/dehydratase. Therefore, this latter protein likely represents a false positive (a false discovery rate of 1% was accepted in this analysis).

TABLE 1 | Previously identified import candidates do not comprise nucleus-encoded solute transporters.

	0 TMH	1 TMH	>1 TMH
<i>In silico</i> predicted (crTP)	289	3	1
LC-MS/MS identified	194	11	2
Total	416	13	3

The table lists numbers of proteins previously identified to be imported into the chromatophore [by *in silico* prediction (based on presence of a crTP), liquid chromatography coupled to tandem MS (LC-MS/MS), and total] (Singer et al., 2017) sorted by the number of predicted TMHs (outside of the crTP).

The absence of multi-spanning TM proteins among import candidates could have two reasons. (i) Similar to the mTP-independent insertion of many nucleus-encoded carriers into the mitochondrial IM (Ferramosca and Zara, 2013), these proteins might use a crTP-independent import route, impairing their prediction as import candidates. (ii) TM proteins are often underrepresented in LC-MS analyses owing to low abundance levels as well as unfavorable retention and ionization properties. In fact, our previous MS analysis identified 47% of the soluble but only 21% of TMH-containing chromatophore-encoded proteins (**Figure 2C**).

Thus, to enhance identification of TM proteins, we enriched TM proteins by collecting the insoluble fractions from isolated chromatophores (CM samples) and intact *P. chromatophora* cells (PM samples). Electron microscopic analysis of isolated chromatophores suggested that the chromatophore OM is lost during chromatophore isolation (**Figure 2A**, compare also Kies, 1974; Sato et al., 2020). Comparison of CM and PM samples to chromatophore lysates (CL samples) by SDS-PAGE revealed distinct banding patterns between the three samples and high reproducibility between three biological replicates (**Figure 2B**). Further enrichment of membrane proteins or separation of IM, OM, and thylakoids was not feasible owing to the slow growth of *P. chromatophora* (~ one cell division per week), low yield of chromatophore isolations, and the loss of the OM. Two consecutive, independent MS analyses of three replicates of each, CM, PM, and CL samples led to the identification of 1,886 nucleus- and 555 chromatophore-encoded proteins over all fractions (**Table 2** and **Supplementary Table S1**). Although most chromatophore-localized TM proteins were also identified in our analyses in CL samples (**Table 2**), individual TM proteins were clearly enriched in CM compared to CL samples (**Supplementary Figure S2**).

In CM samples, 46% (or 98 of 213) of the chromatophore-encoded TM proteins were identified, representing a gain of 118% compared to our previous analysis (**Figure 2C**); in particular, of the 25 chromatophore-encoded solute transport systems, 72% (or 18 proteins) were identified with at least one subunit, and 60% (or 15 proteins) were identified with their TM subunit (**Figure 2D**) while our previous study identified only three of these transporters. Highest intensities (representing a rough estimation for protein abundances) were found in CM samples for an ABC-transporter annotated as multidrug importer of the P-FAT family (levels 4–5; see section “Materials and Methods” and **Figure 2D**, placing the

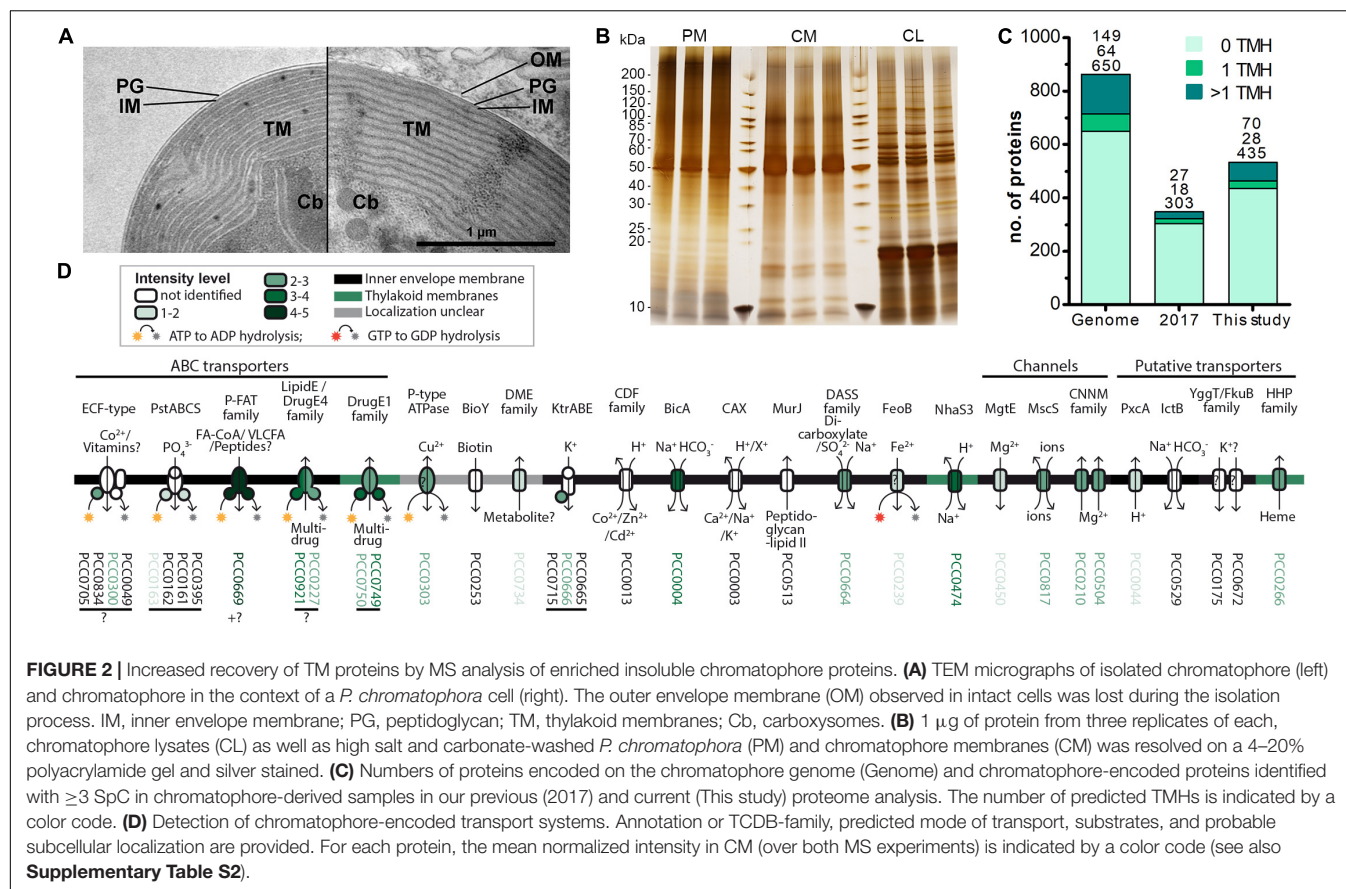


TABLE 2 | Proteins identified in this study by LC-MS/MS.

	All proteins		1 TMH		> 1 TMH	
	CE	NE	CE	NE	CE	NE
CM	533 (506)	297 (236)	28 (24)	20 (13)	70 (67)	5 (2)
CL	551	354	28	25	67	7
Chromatophore total	555	361	28	27	70	7
PM	385	1691	24	209	50	175
Total	555	1886	28	218	70	179

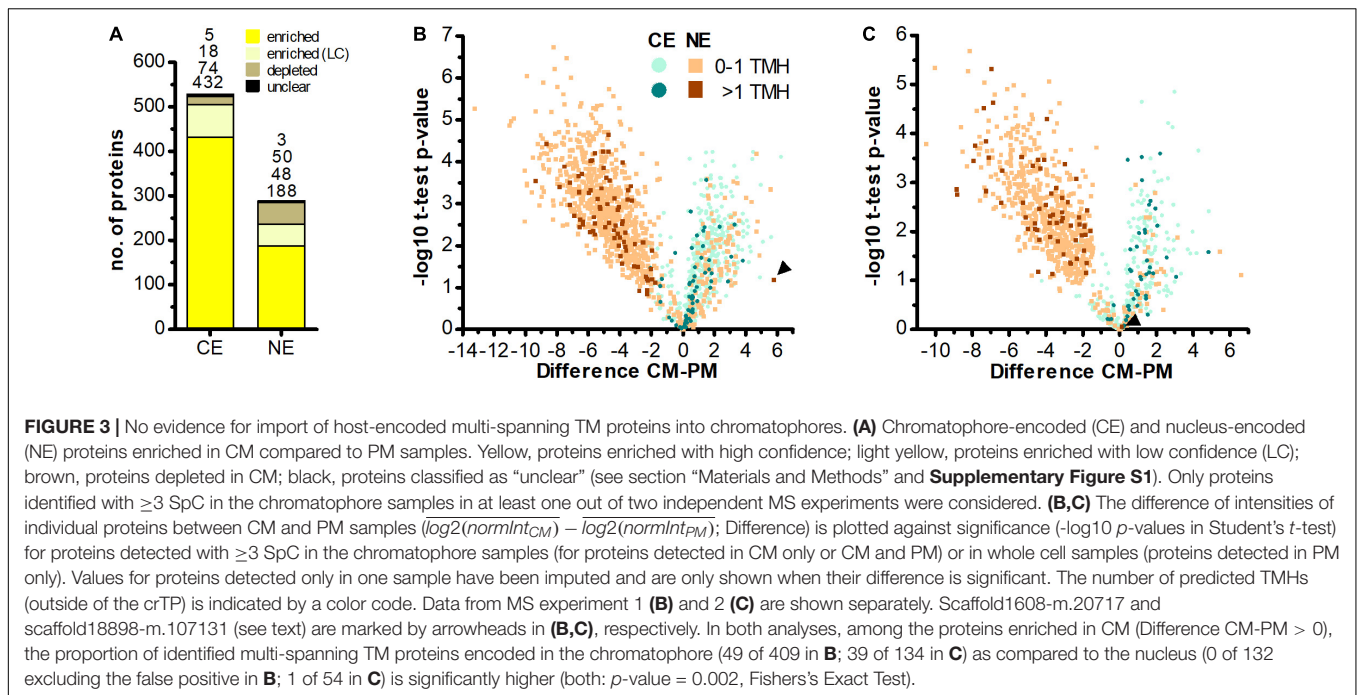
Numbers of chromatophore-encoded (CE) and nucleus-encoded (NE) proteins identified in at least one out of two independent MS experiments with ≥3 spectral counts (SpC) in chromatophore-derived samples (i.e., CM + CL) or whole cell membranes (PM). The number of predicted TMHs (outside of the crTP) is indicated. For proteins identified in CM samples, total number of proteins and number of proteins enriched in CM as compared to PM samples (in brackets) is indicated separately.

transporter among the 10% most abundant proteins in CM). Also the bicarbonate transporter BicA, two multidrug efflux ABC-transporters, and an NhaS3 proton/sodium antiporter were found in the upper tiers of abundance levels (levels 3–4, placing them among the 30% most abundant proteins in CM). The remaining transporters showed moderate to low abundance levels (Figure 2D).

No Multi-Spanning TM Proteins Appear to Be Imported Into the Chromatophore

Determination of nucleus-encoded proteins enriched in CM compared to PM samples led to the identification of 188

high confidence (HC) [and further 48 low confidence (LC); see section “Materials and Methods” and **Supplementary Figure S1**] import candidates (Figure 3A and **Supplementary Table S3**). Nucleus-encoded multi-spanning TM proteins appeared invariably depleted in chromatophores (Figures 3B,C). Only two of 236 import candidates were multi-spanning TM proteins (Table 2). However, one of these (with 7 predicted TMHs, scaffold1608-m.20717, arrowhead in Figure 3B) was identified by only one hepta-peptide and shows no similarity to other proteins in the NCBI nr database whereas an overlapping ORF (in another reading frame) encodes a peroxidase that was MS-identified in Singer et al. (2017) likely classifying the protein as a false positive. For the other import candidate



(scaffold18898-m.107131; with an enrichment level close to 0; arrowhead in **Figure 3C**) a full-length transcript sequence is missing precluding determination of the correct start codon. Thus, this protein might represent in fact a short import candidate with a single TMH. Of the three nucleus-encoded multi-spanning TM proteins that were present but appeared depleted in CM compared to PM samples (**Table 2**), two were annotated as mitochondrial NAD(P) transhydrogenase and mitochondrial ATP/ADP translocase, suggesting a mild contamination of CM samples with mitochondrial membrane material.

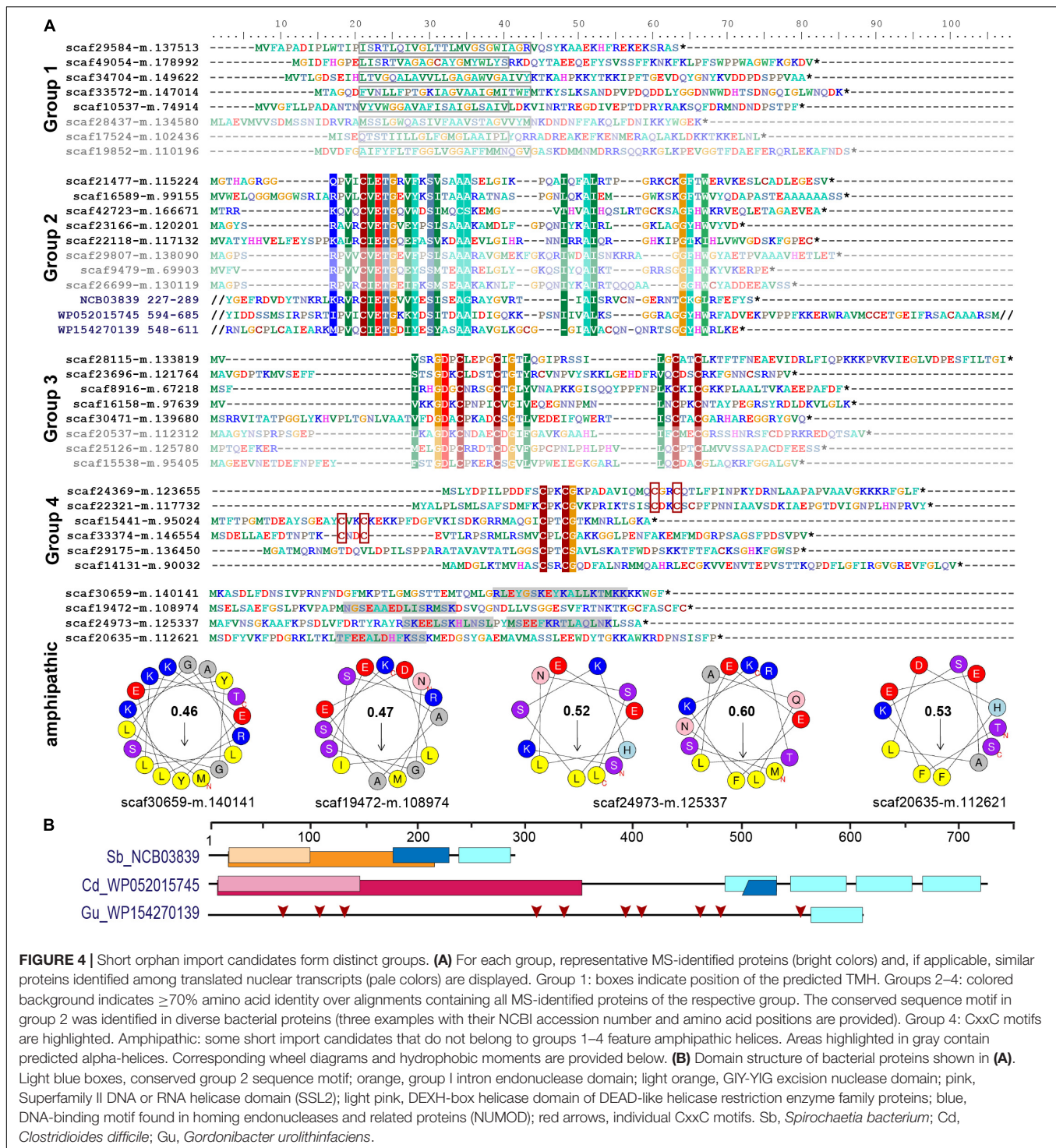
In comparison, 70 chromatophore-encoded multi-spanning TM proteins were identified in CM samples, and 67 of these appeared enriched in CM samples. In PM samples, 50 chromatophore- and 175 nucleus-encoded multi-spanning TM proteins were found (**Table 2**).

To test for the robustness of TMH predictions obtained by TMHMM, import candidates were re-analyzed with a second TMH prediction tool [the Consensus Constrained TOPology prediction (CCTOP); **Supplementary Table S3**]. Although the exact positions or lengths of individual helices were slightly altered in many cases, overall the predictions were largely congruent between the two prediction tools. For 480 out of 508 import candidates, predicted numbers of TMHs were essentially identical between TMHMM and CCTOP; CCTOP predicted 23 additional import candidates with a single TMH, and four additional import candidates with two or three TMHs outside of the crTPs (with three out of four proteins showing a rather low reliability score of the prediction of < 65). Importantly, also CCTOP results did not yield any evidence for the insertion of classical nucleus-encoded transporters (i.e., proteins with ≥ 4 TMHs) into the chromatophore IM. The remaining text refers to TMHMM predictions.

Targeting of Single-Spanning TM Proteins and Antimicrobial Peptide-Like Proteins to the Chromatophore

In contrast to the striking lack of multi-spanning TM proteins, there were 13 (5 HC and 8 LC) single-spanning TM proteins (containing one TMH outside of the crTP) among the identified import candidates (**Table 2**). Three of these proteins contain a TMH close to their C-terminus and likely represent tail-anchored proteins. One of these proteins is long and annotated as low-density lipoprotein receptor-related protein 2-like, the other two (with N-terminal sequence information missing) as polyubiquitin. However, most import candidates with one TMH (10 proteins) represent short proteins. These short import candidates included two high light-inducible proteins (i.e., thylakoid-localized cyanobacterial proteins involved in light acclimation of the cell; Zhang et al., 2017). The remaining eight proteins are orphan proteins lacking detectable homologs in other species (BlastP against NCBI nr database, cutoff e^{-03}); all of these contain a TMH with a large percentage of small amino acids (26–45% Gly, Ala, Ser) close to their negatively charged N-terminus (**Figure 4A**).

In our previous proteome analysis, short orphan proteins represented the largest group of MS-identified import candidates (1/3 of total). However, most of these proteins did not possess predicted TMHs. Based on the occurrence of specific Cys motifs (CxxC, CxxxxC) and stretches of positively charged amino acids these short proteins were described as antimicrobial peptide (AMP)-like proteins (Singer et al., 2017). Including the eight TMH-containing proteins (see above), the current study identified further 19 short orphan import candidates (or—only few proteins—showing similarity to hypothetical proteins in other species). Scrutiny of all 88 short orphan import candidates



(resulting from both studies together) revealed that besides the TMH-containing proteins (group 1, 10 proteins), these short import candidates form at least three further distinct groups (Figure 4A). Members of group 2 (12 proteins) contain a conserved motif of unknown function that occurs also in bacterial proteins that often possess domains with functions related to DNA processing (Figures 4A,B). Members of group 3 (10

proteins) contain another conserved motif of unknown function that encompasses two Cys-motifs (CxxxxC and CxxC). Members of group 4 (30 proteins) show either one or two CxxC mini motifs (one of these is often CPxCG) but no further sequence conservation. The remaining 26 short orphan import candidates have no obvious common characteristics but several appear to have a propensity to form amphipathic helices (Figure 4A).

Screening a large nuclear *P. chromatophora* transcriptome dataset (Nowack et al., 2016) revealed additional putative members of groups 1–3 (Figure 4A and Supplementary Figure S3): further 53 translated transcripts represent short proteins with a predicted TMH in the N-terminal 2/3 of the sequence that is rich (>20%) in small amino acids and have an N-terminus with a net charge ≤ 0 . Notably, the TMHs of >90% of all group 1 proteins comprise at least one (small)xxx(small) motif (where “small” stands for Gly, Ala or Ser and “x” for any amino acid) which can promote oligomerization of single-spanning TM proteins (Teese and Langosch, 2015). Furthermore, many of these putative group 1 short import candidates are predicted to have antimicrobial activity and/or pore-lining residues (Supplementary Table S4) together suggesting their possible function as oligomeric pores or channels. Further 192 and 28 translated transcripts contain the conserved motifs of group 2 or 3, respectively. Importantly, all MS-identified members of these extended protein groups were identified in chromatophore-derived samples in this and our previous analysis.

An Expanded Family of Octotrico Peptide Repeat Putative Expression Regulators Is Targeted to the Chromatophore

Of the 235 import candidates (excluding the false positive, see above) identified in this study (Figure 3A), 159 were known import candidates (Singer et al., 2017; Figure 5A, Supplementary Table S3), with 46 proteins now experimentally confirming previously only *in silico* predicted import candidates. 76 proteins represent new import candidates, mostly lacking N-terminal sequence information (42 proteins) or representing short import candidates (22 proteins). A particularly large number of newly MS-identified import candidates (24 proteins) fall into the category “genetic information processing” (Figure 5B). Among these proteins an expanded group of 10 RNA-binding or RAP domain-containing proteins (where RAP stands for RNA binding domain abundant in apicomplexans, Lee and Hong, 2004) stood out.

These RNA-binding proteins encompass, in addition to the crTP, from N- to C-terminus a variable region of 0–320 aa followed by a ~105 aa long conserved region (CR1), 2–13 repeats of a degenerate 38 aa motif with the most conserved residues being xxxPxxxxLxxxxxxxxxxxxxxFxxQxxxxxLNAxAKL, often followed by a 110 aa long conserved region (CR2), and the 60 aa long RAP domain (Figure 6). This domain organization resembles the one of organelle-targeted octotrico peptide repeat (OPR; i.e., 38 aa peptide repeat) gene expression regulators in green algae and plants (Figures 6B,D) and repeat-containing T3SS effector proteins described from symbiotic or pathogenic bacteria (Figures 6B,E,F). The repeat motifs in all of these proteins share the prediction to form two antiparallel α -helices. Homology-based 3D-structure prediction of *Paulinella* OPR proteins suggests folding of the α -helical repeats into a super helix (or α -solenoid) structure (Figure 6G) as described for OPR proteins in the Viridiplantae.

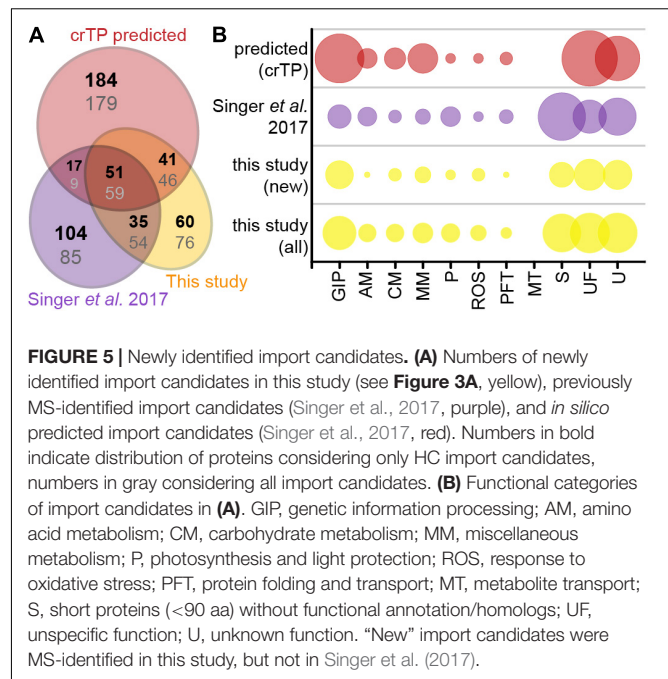


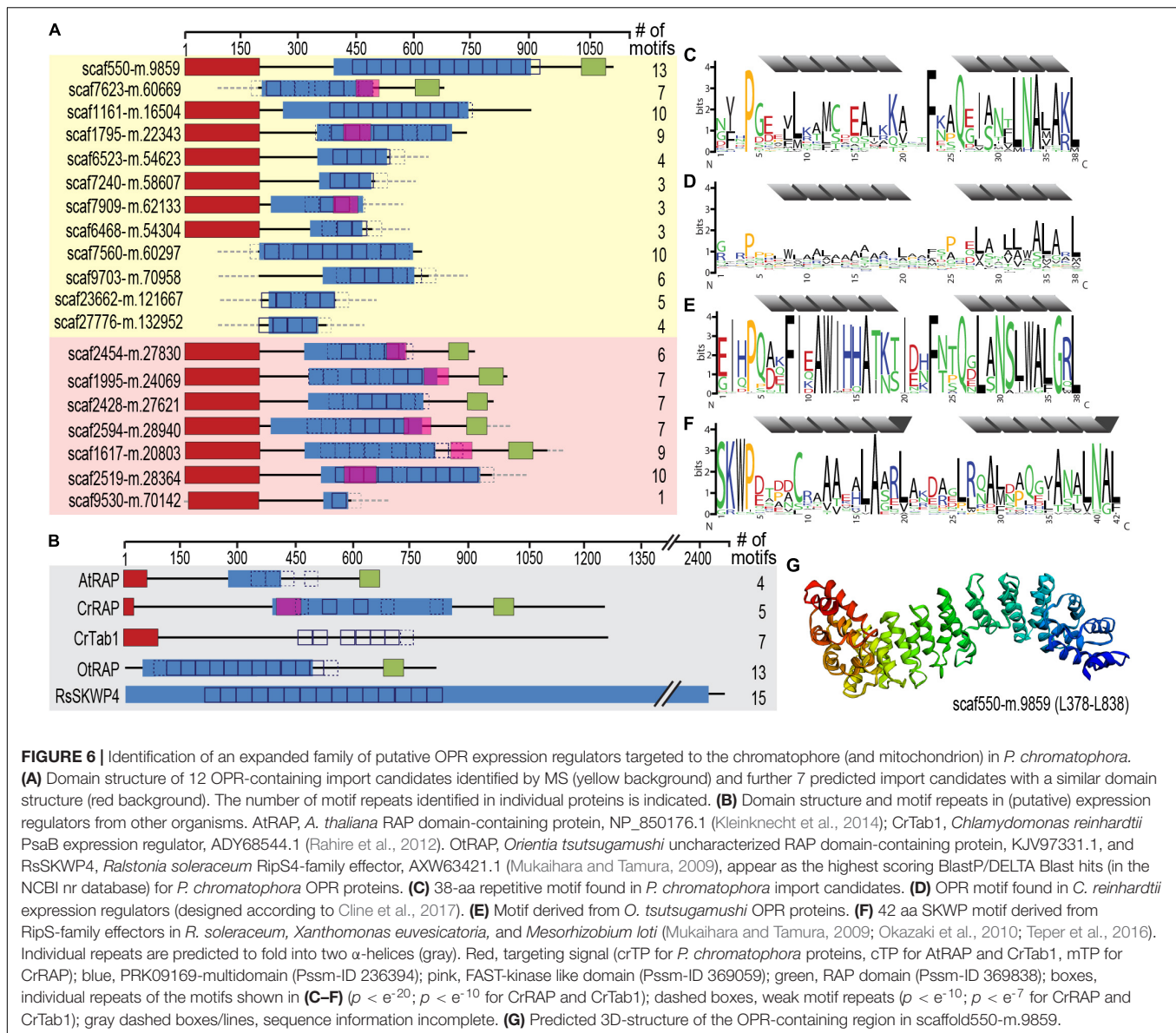
FIGURE 5 | Newly identified import candidates. **(A)** Numbers of newly identified import candidates in this study (see Figure 3A, yellow), previously MS-identified import candidates (Singer et al., 2017, purple), and *in silico* predicted import candidates (Singer et al., 2017, red). Numbers in bold indicate distribution of proteins considering only HC import candidates, numbers in gray considering all import candidates. **(B)** Functional categories of import candidates in **(A)**. GIP, genetic information processing; AM, amino acid metabolism; CM, carbohydrate metabolism; MM, miscellaneous metabolism; P, photosynthesis and light protection; ROS, response to oxidative stress; PFT, protein folding and transport; MT, metabolite transport; S, short proteins (<90 aa) without functional annotation/homologs; UF, unspecific function; U, unknown function. “New” import candidates were MS-identified in this study, but not in Singer et al. (2017).

Screening the complete *P. chromatophora* transcriptome identified OPR proteins as part of an expanded protein family containing at least 101 members with 1–13 individual OPR motifs (Supplementary Table S5). Besides the 12 chromatophore-localized OPR proteins identified by MS (Figure 6A), of the further 12 OPR proteins identified only in the transcriptome for which full-length N-terminal sequence information was available, seven proteins contained a crTP (Figure 6A), the remaining five a mitochondrial targeting signal.

DISCUSSION

Metabolite Transport

Despite the obvious need for extensive metabolite exchange between the chromatophore and cytoplasm (Valadez-Cano et al., 2017), the chromatophore likely lost on the order of 70 solute transporters following symbiosis establishment (Figure 1). The remaining transport systems do not appear apt to establish metabolic connectivity (Figure 2D). Solely two systems, a DME family and a DASS family transporter, might be involved in metabolite transport. Furthermore, there are three ABC-transporters for which substrate specificity is unknown. However, the high energy costs associated with their ATP-consuming primary active mode of transport appears to be incongruous with high-throughput metabolite shuttling. Some of these ABC-transporters might have become specialized for protein import instead. In line with this idea, the ABC-half transporter PCC0669 that showed highest ion intensities among all chromatophore-encoded transporters (Figure 2D), possesses 33% similarity to BclA of *Bradyrhizobium* sp., a nitrogen-fixing bacterium harbored by *Aeschynomene* legumes. BclA functions as an importer for nodule-specific cysteine-rich (NCR) peptides



produced by the host plants symbiotic nodule cells (Guefrachi et al., 2015). However, since other transporters in the same family are involved in peroxisomal transport of fatty acids or fatty acyl-CoA (Linka and Esser, 2012), similar substrates could also be transported by PCC0669.

In plants, insertion of nucleus-encoded transporters into the plastid IM is crucial for metabolic connectivity; these are mostly native host proteins but also include products of horizontally acquired genes (Facchinelli and Weber, 2011; Fischer, 2011; Karkar et al., 2015). Also in more recently established endosymbiotic associations, such as plant sap-feeding insects with nutritional bacterial endosymbionts, multiplication of host transporters followed by their recruitment to the host/endosymbiont interface apparently was involved in establishing metabolic connectivity (Price et al., 2011; Duncan et al., 2014). However, these transporters localize to

the symbiosomal membrane, a host membrane that surrounds bacterial endosymbionts. The mechanism enabling metabolite transport across the symbionts' IM and OM, with symbiont-encoded transport systems being scarce, is a longstanding, unanswered question (Mergaert et al., 2017).

Despite the import of hundreds of soluble proteins into the chromatophore, our work provided no evidence for the insertion of nucleus-encoded transporters (nor any other multi-spanning TM proteins) into the chromatophore IM (or thylakoids). The possibility that such proteins escaped detection for technical reasons appears improbable because: (i) 72% of the chromatophore-encoded transporters were identified in CM samples. Assuming comparable abundances for nucleus-encoded chromatophore-targeted transporters, a large percentage of these proteins should have been detected, too. (ii) More than 100 nucleus-encoded transporters or transporter components were

detected in comparable amounts of PM samples showing that our method is feasible to detect this group of proteins. (iii) IM transporters were repeatedly identified in comparable analyses of cyanobacterial (Pisareva et al., 2011; Plohnke et al., 2015; Liberton et al., 2016; Baers et al., 2019; Choi et al., 2020) or plastidial membrane fractions (Bräutigam et al., 2008; Simm et al., 2013; Bouchnak et al., 2019). Thus, a general mechanism to insert nucleus-encoded multi-spanning TM proteins into chromatophore IM and thylakoids likely has not evolved (yet) in *P. chromatophora* (although a few such proteins might insert spontaneously based on their individual physicochemical properties). Post-translational migration of highly hydrophobic membrane proteins through the aqueous cytoplasm might be a challenging task. A cell would either have to develop factors that prevent hydrophobic proteins from aggregation or mistargeting to the endoplasmic reticulum or introduce mutations that reduce overall hydrophobicity in transmembrane regions (Popot and Devitry, 1990; Adams and Palmer, 2003; Oh and Hwang, 2015). Thus, import of soluble proteins might be more straight-forward to evolve and establish at an earlier stage of organellogenesis than import of hydrophobic proteins.

The protein composition of the chromatophore OM is currently unclear. However, its putative host origin and the notion that proteins traffic into the chromatophore likely via the Golgi (Nowack and Grossman, 2012) suggest that nucleus-encoded transporters can be targeted to the OM by vesicle fusion. Nonetheless, our findings spotlight the puzzling absence of suitable transporters that would allow metabolite exchange across the chromatophore IM. The conservation of active and secondary active IM transporters on the chromatophore genome (**Figure 2D**) strongly implies that the chromatophore IM kept its barrier function and there is an electrochemical gradient across this lipid bilayer.

In contrast to the absence of multi-spanning TM proteins, we identified numerous short single-spanning TM and AMP-like orphan proteins among chromatophore-targeted proteins. These short import candidates fall into at least four expanded groups, suggesting some degree of functional specialization. Interestingly, expanded arsenals of symbiont-targeted polypeptides convergently evolved in many taxonomically unrelated symbiotic associations and thus seem to represent a powerful strategy to establish host control over bacterial endosymbionts (Mergaert, 2018). It has been suggested that these “symbiotic AMPs” have the ability to self-translocate across or self-insert into endosymbiont membranes and mediate control over various biological processes in the symbionts including translation, septum formation or modulation of membrane permeability and metabolite exchange (Mergaert et al., 2006, 2017; van de Velde et al., 2010; Login et al., 2011; Farkas et al., 2014; Carro et al., 2015; Mergaert, 2018). For example, the AMP Ag5 is produced in root nodules of the Alder tree that house the nitrogen-fixing endosymbiont *Frankia alni*. When *Frankia* cells are treated *in vitro* with Ag5 concentrations <1 μ M, the release of specific amino acids is triggered, whereas higher concentrations harm and ultimately kill the bacterium (Carro et al., 2015).

The discovery of TMH-containing group 1 proteins appears to be of particular interest in the context of metabolite exchange. The frequent occurrence of (small)xxx(small) motifs might indicate the potential of these proteins to oligomerize by allowing for close proximity between interacting TMHs. Such associations are known to be stabilized by interfacial van der Waals interactions and/or hydrogen bonding resulting from the excellent geometric fit between the interacting TMHs (Moore et al., 2008; Teese and Langosch, 2015). The predicted pore-lining residues (**Supplementary Table S4**) in the TMHs of many of these proteins further suggest that they could form homo- or hetero-oligomeric channels. It has been previously reported that AMPs can arrange in channel-like assemblies which facilitate diffusion along concentration gradients (Rahaman and Lazaridis, 2014; Wang et al., 2016), though the lifetime and selectivity of such arrangements requires further investigation. Given the size of the metabolites to be transported, they would be required to form multimer arrangements in barrel-stave (**Supplementary Figure S4**) or shortly lived toroidal pores, while maintaining the overall impermeability of the membrane. The formation of such pores still begs the question of how they could maintain a selective metabolite transport. An interesting example in that respect is the VDAC channel of the mitochondrial OM which has been described to follow a stochastic gating mechanism, in which only bigger and, hence, slowly diffusing molecules would be allowed to permeate (Berezkhovskii and Bezrukov, 2018).

An alternative mode of action involves soluble, short import candidates which could interact with the chromatophore envelope membranes via stretches of positively charged amino acids and amphipathic helices (**Figure 4A**), and putatively modulate membrane permeability (Mergaert et al., 2017) in what is known as carpet model (Wimley, 2010). The mechanism by which such an interaction could cause a transient permeabilization is still a matter of debate, although the asymmetric distribution of peptides on the membrane bilayer has been pointed out as plausible reason (Guha et al., 2019). This asymmetric distribution creates an imbalance of mass, charge, surface tension, and lateral pressure. A combination of these factors is hypothesized to lead to stochastic local dissipation events relieving asymmetry by peptide, and possibly lipid, translocation and concomitantly inducing transient permeability to polar molecules. Further experimental work with the identified proteins could shed light on the potential transport mechanism.

Other short import candidates might also attack targets inside of the chromatophore (e.g., DNA, specific RNA species, the replication or translation machineries). The group 2 sequence motif is found also in hypothetical bacterial proteins which include domains related to DNA processing functions (**Figure 4B**). Thus, group 2 proteins might provide the host with control over aspects of genetic information processing in the chromatophore. The presence of dozens to hundreds of similar proteins in the various groups, points to a functional interdependence or reciprocal control of individual peptides. In insects, co-occurring AMPs have been shown to synergize, e.g.,

some AMPs permeabilize membranes to enable entry of other AMPs that have intracellular targets (Rahnamaeian et al., 2015).

Nuclear Control Over Expression of Chromatophore-Encoded Proteins

Besides the establishment of metabolic connectivity, our analyses illuminated another cornerstone in organogenesis, the evolution of nuclear control over organellar gene expression. Previously, we identified a large number of proteins annotated as transcription factors among chromatophore-targeted proteins (Singer et al., 2017). Here we described a novel class of chromatophore-targeted helical repeat proteins. Helical repeat proteins appear to represent ubiquitous nuclear factors involved in regulation of organellar gene expression (Hammani et al., 2014). These proteins are generally characterized by the presence of degenerate 30–40 aa repeat motifs, each of them containing two antiparallel α -helices. The succession of motifs underpins the formation of a super helix that enables sequence specific binding to nucleic acids.

The *P. chromatophora* nuclear genome encodes at least 101 OPR helical repeat proteins (**Figure 6C**). OPR proteins have mostly been studied in the green alga *C. reinhardtii*, where 44 OPR genes were identified in the nuclear genome. Almost all of these OPR proteins are predicted to localize to organelles (Eberhard et al., 2011; **Figure 6D**) and five have been shown experimentally to be involved in post-transcriptional steps of chloroplast gene expression. The only known *A. thaliana* OPR protein is AtRAP (Kleinknecht et al., 2014; **Figure 6B**), a factor promoting chloroplast rRNA maturation. With around 450 members, pentatricopeptide repeat (PPR, repeats of 35 aa) proteins represent the most prominent family of organelle-targeted helical repeat proteins with functions in gene expression regulation in land plants (Lurin et al., 2004; Colcombet et al., 2013). The *C. reinhardtii* genome encodes only 14 PPR proteins (Tourasse et al., 2013), indicating that different families of organelle-targeted helical repeat proteins have expanded in different phyla to fulfill similar purposes.

Also the *Paulinella* OPR proteins seem to be mostly organelle-targeted. Many *Paulinella* OPR proteins possess, in addition to the OPR stretches, a Fas-activated serine/threonine (FAST) kinase-like domain (Tian et al., 1995) and a C-terminal RAP domain (**Figure 6A**). This domain combination is also present in some of the *C. reinhardtii* OPR proteins (e.g., CrRAP in **Figure 6B**), the *A. thaliana* AtRAP protein (**Figure 6B**), and the FASTK family of vertebrate nucleus-encoded regulators of mitochondrial gene expression (Boehm et al., 2016). Additionally, some bacterial T3SS effector proteins (**Figure 6B**) show similar domain architectures. However, the exact molecular functions of FAST kinase-like and RAP domains as well as the two conserved regions in *Paulinella* OPR proteins (CR1 and CR2) that share no similarity with known domains remain unknown.

In conclusion, in parallel to the evolution of mitochondria and plastids, also during chromatophore evolution an expanded family of chromatophore-targeted helical repeat proteins evolved. Based on the similarity of their domain architecture to known organelle-targeted expression regulators, the OPR proteins in *P. chromatophora* likely serve as nuclear factors modulating chromatophore gene expression by direct binding to specific target RNAs. Probably chromatophore-targeted OPR proteins evolved from pre-existing mitochondrial expression regulators and were recruited to the chromatophore by crTP acquisition. However, the RNA-binding ability of *Paulinella* OPR proteins, their specific target sequences as well as their ability to modulate expression of chromatophore-encoded proteins remain to be tested experimentally.

DATA AVAILABILITY STATEMENT

The names of the repository/repositories and accession number(s) can be found below: PRIDE Archive (<https://www.ebi.ac.uk/pride/archive/>); accession number: PXD021087.

AUTHOR CONTRIBUTIONS

EN conceived the study, analyzed the data, and wrote the manuscript. LO conceived the study, performed most of the experimental work, analyzed the data, and wrote the manuscript. GP and KS performed MS analyses. LM performed TEM analyses. SS-V and HG generated and analyzed oligomeric pore models. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Deutsche Forschungsgemeinschaft CRC 1208 project B09 (to EN), A03 (to HG), and project Z01 (to KS); and Deutsche Forschungsgemeinschaft grant NO 1090/1-1 (to EN).

ACKNOWLEDGMENTS

We thank the core facility Elektronenmikroskopie UKD (HHU Düsseldorf) for their help with thin sectioning and TEM analyses. This manuscript has been released as a pre-print at bioRxiv (Oberleitner et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.607182/full#supplementary-material>

REFERENCES

- Adams, K. L., and Palmer, J. D. (2003). Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol. Phylogenet. Evol.* 29, 380–395. doi: 10.1016/s1055-7903(03)00194-5
- Almagro Armenteros, J. J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., et al. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance*. 2:e201900429. doi: 10.26508/lsa.201900429
- Baers, L. L., Breckels, L. M., Mills, L. A., Gatto, L., Deery, M. J., Stevens, T. J., et al. (2019). Proteome mapping of a cyanobacterium reveals distinct compartment organization and cell-dispersed metabolism. *Plant Physiol.* 181, 1721–1738. doi: 10.1104/pp.19.00897
- Bagos, P. G., Liakopoulos, T. D., and Hamodrakas, S. J. (2004). Finding beta-barrel outer membrane proteins with a markov chain model. *WSEAS Trans. Biol. Biomed.* 2, 186–189.
- Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Systems Mol. Biol.* 2, 28–36.
- Bailey, T. L., and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* 14, 48–54. doi: 10.1093/bioinformatics/14.1.48
- Berezhevskii, A. M., and Bezrukov, S. M. (2018). Stochastic gating as a novel mechanism for channel selectivity. *Biophys. J.* 114, 1026–1029. doi: 10.1016/j.bpj.2018.01.007
- Boehm, E., Zornoza, M., Jourdain, A. A., Magdalena, A. D., García-Consuegra, I., Merino, R. T., et al. (2016). Role of FAST kinase domains 3 (FASTKD3) in post-transcriptional regulation of mitochondrial gene expression. *J. Biol. Chem.* 291, 25877–25887. doi: 10.1074/jbc.m116.730291
- Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., and Madden, T. L. (2012). Domain enhanced lookup time accelerated BLAST. *Biol. Direct.* 7:12. doi: 10.1186/1745-6150-7-12
- Bouchnak, I., Brugière, S., Moyet, L., Le Gall, S., Salvi, D., Kuntz, M., et al. (2019). Unraveling hidden components of the chloroplast envelope proteome: opportunities and limits of better MS sensitivity. *Mol. Cell. Proteomics* 18, 1285–1306. doi: 10.1074/mcp.ra118.000988
- Bräutigam, A., Hoffmann-Benning, S., and Weber, A. P. M. (2008). Comparative proteomics of chloroplast envelopes from C3 and C4 plants reveals specific adaptations of the plastid envelope to C4 photosynthesis and candidate proteins required for maintaining C4 metabolite fluxes. *Plant Physiol.* 148, 568–579. doi: 10.1104/pp.108.121012
- Breuers, F. K. H., Bräutigam, A., and Weber, A. P. M. (2011). The plastid outer envelope - a highly dynamic interface between plastid and cytoplasm. *Front. Plant Sci.* 2:97. doi: 10.3389/fpls.2011.00097
- Burdukiewicz, M., Sidorczuk, K., Rafacz, D., Pietluch, F., Chilimoniuk, J., Rödiger, S., et al. (2020). Proteomic screening for prediction and design of antimicrobial peptides with AmpGram. *Int. J. Mol. Sci.* 21:4310. doi: 10.3390/ijms21124310
- Carro, L., Pujic, P., Alloisio, N., Fournier, P., Boubakri, H., Hay, A. E., et al. (2015). Alnus peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing *Frankia*. *ISME J.* 9, 1723–1733. doi: 10.1038/ismej.2014.257
- Choi, J.-S., Park, Y. H., Oh, J. H., Kim, S., Kwon, J., and Choi, Y.-E. (2020). Efficient profiling of detergent-assisted membrane proteome in cyanobacteria. *J. Appl. Phycol.* 32, 1177–1184. doi: 10.1007/s10811-019-01986-4
- Cline, S. G., Laughbaum, I. A., and Hamel, P. P. (2017). CCS2, an octatricopeptide-repeat protein, is required for plastid cytochrome c assembly in the green Alga *Chlamydomonas reinhardtii*. *Front. Plant Sci.* 8:1306. doi: 10.3389/fpls.2017.01306
- Colombet, J., Lopez-Obando, M., Heurtevin, L., Bernard, C., Martin, K., Berthomé, R., et al. (2013). Systematic study of subcellular localization of *Arabidopsis* PPR proteins confirms a massive targeting to organelles. *RNA Biol.* 10, 1557–1575. doi: 10.4161/rna.26128
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Delaye, L., Valadez Cano, C., and Pérez Zamorano, B. (2016). How really ancient is *Paulinella chromatophora*? *PLoS Curr.* 8:ecurrents.tol.e68a099364bb1a1e129a17b4e06b0c6b. doi: 10.1371/currents.tol.e68a099364bb1a1e129a17b4e06b0c6b
- Dobson, L., Reményi, I., and Tusnády, G. E. (2015). CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res.* 43, W408–W412.
- Drozdzetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 43, W389–W394.
- Duncan, R. P., Husnik, F., Van Leuven, J. T., Gilbert, D. G., Dávalos, L. M., McCutcheon, J. P., et al. (2014). Dynamic recruitment of amino acid transporters to the insect/symbiont interface. *Mol. Ecol.* 23, 1608–1623. doi: 10.1111/mec.12627
- Eberhard, S., Loisel, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., et al. (2011). Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of atpA transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Plant J.* 67, 1055–1066. doi: 10.1111/j.1365-3113.2011.04657.x
- Facchinelli, F., and Weber, A. P. M. (2011). The metabolite transporters of the plastid envelope: an update. *Front. Plant Sci.* 2:50. doi: 10.3389/fpls.2011.00050
- Farkas, A., Maróti, G., Dürgo, H., Györgypál, Z., Lima, R. M., Medzihradský, K. F., et al. (2014). *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5183–5188. doi: 10.1073/pnas.1404169111
- Ferramosca, A., and Zara, V. (2013). Biogenesis of mitochondrial carrier proteins: molecular mechanisms of import into mitochondria. *Biochim. Biophys. Acta-Mol. Cell Res.* 1833, 494–502. doi: 10.1016/j.bbamcr.2012.11.014
- Fischer, K. (2011). The import and export business in plastids: transport processes across the inner envelope membrane. *Plant Physiol.* 155, 1511–1519. doi: 10.1104/pp.110.170241
- Gabr, A., Grossman, A. R., and Bhattacharya, D. (2020). *Paulinella*, a model for understanding plastid primary endosymbiosis. *J. Phycol.* 56, 837–843. doi: 10.1111/jpy.13003
- Gautier, R., Douget, D., Antonny, B., and Drin, G. (2008). HELIQUEST: a web server to screen sequences with specific α -helical properties. *Bioinformatics* 24, 2101–2102. doi: 10.1093/bioinformatics/btn392
- Goetze, T. A., Patil, M., Jeshen, I., Bölter, B., Grahl, S., and Soll, J. (2015). Oep23 forms an ion channel in the chloroplast outer envelope. *BMC Plant Biol.* 15:47. doi: 10.1186/s12870-015-0445-1
- Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. doi: 10.1093/bioinformatics/btr064
- Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., et al. (2015). *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol. Plant-Microbe Interact.* 28, 1155–1166. doi: 10.1094/mpmi-04-15-0094-r
- Guha, S., Ghimire, J., Wu, E., and Wimley, W. C. (2019). Mechanistic landscape of membrane-permeabilizing peptides. *Chem. Rev.* 119, 6040–6085. doi: 10.1021/acs.chemrev.8b00520
- Hammani, K., Bonnard, G., Bouchoucha, A., Gobert, A., Pinker, F., Salinas, T., et al. (2014). Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* 100, 141–150. doi: 10.1016/j.biochi.2013.08.031
- Harsman, A., Schock, A., Hemmis, B., Wahl, V., Jeshen, I., Bartsch, P., et al. (2016). OEP40, a regulated glucose-permeable β -barrel solute channel in the chloroplast outer envelope membrane. *J. Biol. Chem.* 291, 17848–17860. doi: 10.1074/jbc.m115.712398
- Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587.
- Jack, D. L., Yang, N. M., and Saier, M. H. (2001). The drug/metabolite transporter superfamily. *Eur. J. Biochem.* 268, 3620–3639. doi: 10.1046/j.1432-1327.2001.02265.x
- Karkar, S., Facchinelli, F., Price, D. C., Weber, A. P. M., and Bhattacharya, D. (2015). Metabolic connectivity as a driver of host and endosymbiont integration. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10208–10215. doi: 10.1073/pnas.1421375112
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053
- Kies, L. (1974). Electron microscopical investigations on *Paulinella chromatophora* Lauterborn, a thecamoeba containing blue-green endosymbionts (cyanelles). *Protoplasma* 80, 69–89.
- Kleinknecht, L., Wang, F., Stübe, R., Philipp, K., Nickelsen, J., and Bohne, A. V. (2014). RAP, the sole octotricopeptide repeat protein in *Arabidopsis*, is required

- for chloroplast 16S rRNA maturation. *Plant Cell* 26, 777–787. doi: 10.1105/tpc.114.122853
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Laudenbach, D. E., and Grossman, A. R. (1991). Characterization and mutagenesis of sulfur-regulated genes in a cyanobacterium - evidence for function in sulfate transport. *J. Bacteriol.* 173, 2739–2750. doi: 10.1128/jb.173.9.2739-2750.1991
- Lee, I., and Hong, W. (2004). RAP - A putative RNA-binding domain. *Trends Biochem. Sci.* 29, 567–570. doi: 10.1016/j.tibs.2004.09.005 +
- Liberton, M., Saha, R., Jacobs, J. M., Nguyen, A. Y., Gritsenko, M. A., Smith, R. D., et al. (2016). Global proteomic analysis reveals an exclusive role of thylakoid membranes in bioenergetics of a model cyanobacterium. *Mol. Cell. Proteomics* 15, 2021–2032. doi: 10.1074/mcp.m115.057240
- Linka, N., and Esser, C. (2012). Transport proteins regulate the flux of metabolites and cofactors across the membrane of plant peroxisomes. *Front. Plant Sci.* 3:3. doi: 10.3389/fpls.2012.00003
- Login, F. H., Balmann, S., Vallier, A., Vincent-Monégat, C., Vigneron, A., Weiss-Gayet, M., et al. (2011). Antimicrobial peptides keep insect endosymbionts under control. *Science* 334, 362–365. doi: 10.1126/science.1209728
- Lurin, C., Andres, C., Aubourg, S., Bellou, M., Bitton, F., Bruyère, C., et al. (2004). Genome-wide analysis of *Arabidopsis pentatricopeptide* repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16, 2089–2103. doi: 10.1105/tpc.104.022236
- Marchand, J., Heydarizadeh, P., Schoefs, B., and Spetea, C. (2018). Ion and metabolite transport in the chloroplast of algae: lessons from land plants. *Cell Mol. Life Sci.* 75, 2153–2176. doi: 10.1007/s00018-018-2793-0
- Marin, B., Nowack, E. C. M., Glöckner, G., and Melkonian, M. (2007). The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like gamma-proteobacterium. *BMC Evol. Biol.* 7:85. doi: 10.1186/1471-2148-7-85
- Marin, B., Nowack, E. C. M., and Melkonian, M. (2005). A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156, 425–432. doi: 10.1016/j.protis.2005.09.001
- Markovich, D. (2012). Sodium-sulfate/carboxylate cotransporters (SLC13). *Curr. Top. Membr.* 70, 239–256. doi: 10.1016/b978-0-12-394316-3.00007-7
- McFadden, G. I., and Melkonian, M. (1986). Use of hepes buffer for microalgal culture media and fixation for electron-microscopy. *Phycologia* 25, 551–557. doi: 10.2216/i0031-8884-25-4-551.1
- Mehrshahi, P., Stefano, G., Andaloro, J. M., Brandizzi, F., Froehlich, J. E., and DellaPenna, D. (2013). Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proc. Natl. Acad. Sci. U S A.* 110, 12126–12131. doi: 10.1073/pnas.1306331110
- Mergaert, P. (2018). Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Nat. Prod. Rep.* 35, 336–356. doi: 10.1039/c7np00056a
- Mergaert, P., Kikuchi, Y., Shigenobu, S., and Nowack, E. C. M. (2017). Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol.* 25, 703–712. doi: 10.1016/j.tim.2017.04.007
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium-legume* symbiosis. *Proc. Natl. Acad. Sci. U S A.* 103, 5230–5235.
- Micallef, L., and Rodgers, P. (2014). eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses. *PLoS One* 9:e101717. doi: 10.1371/journal.pone.0101717
- Montesinos, M. L., Muro-Pastor, A. M., Herrero, A., and Flores, E. (1998). Ammonium/Methylammonium permeases of a cyanobacterium-Identification and analysis of three nitrogen-regulated amt genes in *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 273, 31463–31470. doi: 10.1074/jbc.273.47.31463
- Moore, D. T., Berger, B. W., and DeGrado, W. F. (2008). Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Structure* 16, 991–1001. doi: 10.1016/j.str.2008.05.007
- Mukaihara, T., and Tamura, N. (2009). Identification of novel *Ralstonia solanacearum* type III effector proteins through translocation analysis of hrpB-regulated gene products. *Microbiology-(UK)* 155, 2235–2244. doi: 10.1099/mic.0.027763-0
- Neuhoff, V., Philipp, K., Zimmer, H. G., and Mesecke, S. (1979). A simple, versatile, sensitive and volume-independent method for quantitative protein determination which is independent of other external influences. *Hoppe-Seylers Zeitschrift Für Physiologische Chemie.* 360, 1657–1670.
- Nowack, E. C. M. (2014). *Paulinella chromatophora* - rethinking the transition from endosymbiont to organelle. *Acta Soc. Botanicorum Poloniae* 83, 387–397.
- Nowack, E. C. M., and Grossman, A. R. (2012). Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U S A.* 109, 5340–5345. doi: 10.1073/pnas.1118800109
- Nowack, E. C. M., Melkonian, M., and Glöckner, G. (2008). Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr. Biol.* 18, 410–418. doi: 10.1016/j.cub.2008.02.051
- Nowack, E. C. M., Price, D. C., Bhattacharya, D., Singer, A., Melkonian, M., and Grossman, A. R. (2016). Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U S A.* 113, 12214–12219. doi: 10.1073/pnas.1608016113
- Nowack, E. C. M., Vogel, H., Groth, M., Grossman, A. R., Melkonian, M., and Glöckner, G. (2011). Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. *Mol. Biol. Evol.* 28, 407–422. doi: 10.1093/molbev/msq209
- Nugent, T., and Jones, D. T. (2012). Detecting pore-lining regions in transmembrane protein sequences. *BMC Bioinform.* 13:169. doi: 10.1186/1471-2105-13-169
- Oberleitner, L., Poschmann, G., Macorano, L., Schott-Verdugo, S., Gohlke, H., Stühler, K., et al. (2020). The puzzle of metabolite exchange and identification of putative octotrico peptide repeat expression regulators in the nascent photosynthetic organelles of *Paulinella chromatophora*. *bioRxiv* [preprint] doi: 10.1101/2020.08.26.269498.
- Oh, Y. J., and Hwang, I. (2015). Targeting and biogenesis of transporters and channels in chloroplast envelope membranes: unsolved questions. *Cell Calcium* 58, 122–130. doi: 10.1016/j.ceca.2014.10.012
- Okazaki, S., Okabe, S., Higashi, M., Shimoda, Y., Sato, S., Tabata, S., et al. (2010). Identification and functional analysis of type III effector proteins in *Mesorhizobium loti*. *Mol. Plant-Microbe Interact.* 23, 223–234. doi: 10.1094/mpmi-23-2-0223
- Omata, T., Andriess, X., and Hirano, A. (1993). Identification and characterization of a gene-cluster involved in nitrate transport in the cyanobacterium *Synechococcus* sp. PCC7942. *Mol. General Genet.* 236, 193–202. doi: 10.1007/bf00277112
- Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R., and Saier, M. H. Jr. (2000). Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J. Mol. Biol.* 301, 75–100. doi: 10.1006/jmbi.2000.3961
- Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450.
- Pisareva, T., Kwon, J., Oh, J., Kim, S., Ge, C., Wieslander, A., et al. (2011). Model for membrane organization and protein sorting in the cyanobacterium *Synechocystis* sp. PCC 6803 inferred from proteomics and multivariate sequence analyses. *J. Proteome Res.* 10, 3617–3631. doi: 10.1021/pr200268r
- Plöhnke, N., Seidel, T., Kahmann, U., Rögner, M., Schneider, D., and Rexroth, S. (2015). The proteome and lipidome of *Synechocystis* sp. PCC 6803 cells grown under light-activated heterotrophic conditions. *Mol. Cell. Proteom.* 14, 572–584. doi: 10.1074/mcp.m114.042382
- Popot, J. L., and Devitry, C. (1990). On the microassembly of integral membrane-proteins. *Ann. Rev. Biophys. Biophys. Chem.* 19, 369–403. doi: 10.1146/annurev.bb.19.060190.002101
- Price, D. R. G., Duncan, R. P., Shigenobu, S., and Wilson, A. C. C. (2011). Genome expansion and differential expression of amino acid transporters at the aphid/Buchnera symbiotic interface. *Mol. Biol. Evol.* 28, 3113–3126. doi: 10.1093/molbev/msr140
- Quintero, M. J., Montesinos, M. L., Herrero, A., and Flores, E. (2001). Identification of genes encoding amino acid permeases by inactivation of selected ORFs from the *Synechocystis* genomic sequence. *Genome Res.* 11, 2034–2040. doi: 10.1101/gr.196301
- Rahaman, A., and Lazaridis, T. (2014). A thermodynamic approach to alamethicin pore formation. *Biochim. et Biophys. Acta - Biomembranes* 1838, 98–105. doi: 10.1016/j.bbame.2013.09.012
- Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J. D. (2012). Identification of an OPR protein involved in the translation initiation of the PsaB subunit

- of photosystem I. *Plant J.* 72, 652–661. doi: 10.1111/j.1365-313x.2012.05111.x
- Rahnamaeian, M., Cytrynska, M., Zdybicka-Barabas, A., Dobszlaff, K., Wiesner, J., Twyman, R. M., et al. (2015). Insect antimicrobial peptides show potentiating functional interactions against Gram-negative bacteria. *Proc. R. Soc. B-Biol. Sci.* 282:20150293. doi: 10.1098/rspb.2015.0293
- Reyes-Prieto, A., Yoon, H. S., Moustafa, A., Yang, E. C., Andersen, R. A., Boo, S. M., et al. (2010). Differential gene retention in plastids of common recent origin. *Mol. Biol. Evol.* 27, 1530–1537. doi: 10.1093/molbev/msq032
- Reynolds, E. S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J. Cell Biol.* 17, 208–212. doi: 10.1083/jcb.17.1.208
- Saier, M. H., Reddy, V. S., Tsu, B. V., Ahmed, M. S., Li, C., and Moreno-Hagelsieb, G. (2016). The Transporter Classification Database (TCDB): recent advances. *Nucleic Acids Res.* 44, D372–D379.
- Sato, N., Yoshitomi, T., and Mori-Moriyama, N. (2020). Characterization and biosynthesis of lipids in *Paulinella micropora* MYN1: evidence for efficient integration of chromatophores into cellular lipid metabolism. *Plant Cell Physiol.* 61, 869–881. doi: 10.1093/pcp/pcaa011
- Schantz Klausen, M., Jespersen, M. C., Nielsen, H., Jensen, K. K., Jurtz, V. I., Sonderby, C. K., et al. (2019). NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning. *Proteins-Structure Funct. Bioinform.* 87, 520–527. doi: 10.1002/prot.25674
- Simm, S., Papasotiriou, D. G., Ibrahim, M., Leisegang, M. S., Mueller, B., Schorge, T., et al. (2013). Defining the core proteome of the chloroplast envelope membranes. *Front. Plant Sci.* 4:11. doi: 10.3389/fpls.2013.00011
- Singer, A., Poschmann, G., Mühlich, C., Valadez-Cano, C., Hänsch, S., Rensing, S. A., et al. (2017). Massive protein import into the early evolutionary stage photosynthetic organelle of the amoeba *Paulinella chromatophora*. *Curr. Biol.* 27, 2763–2773. doi: 10.1016/j.cub.2017.08.010
- Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4, 1581–1590. doi: 10.1002/pmic.200300776
- Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugiere, S., et al. (2012). PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol. Biol. Evol.* 29, 3625–3639. doi: 10.1093/molbev/mss178
- Teese, M. G., and Langosch, D. (2015). Role of GxxxG motifs in transmembrane domain interactions. *Biochemistry* 54, 5125–5135. doi: 10.1021/acs.biochem.5b00495
- Teper, D., Burstein, D., Salomon, D., Gershovitz, M., Pupko, T., and Sessa, G. (2016). Identification of novel *Xanthomonas euvesicatoria* type III effector proteins by a machine-learning approach. *Mol. Plant Pathol.* 17, 398–411.
- Tian, Q. S., Taupin, J. L., Elledge, S., Robertson, M., and Anderson, P. (1995). Fas-activated serine threonine kinase (FAST) phosphorylates TIA-1 during Fas-mediated apoptosis. *J. Exp. Med.* 182, 865–874. doi: 10.1084/jem.182.3.865
- Tourasse, N. J., Choquet, Y., and Vallon, O. (2013). PPR proteins of green algae. *RNA Biol.* 10, 1526–1542. doi: 10.4161/rna.26127
- Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U S A.* 98, 5116–5121. doi: 10.1073/pnas.091062498
- Valadez-Cano, C., Olivares-Hernandez, R., Resendis-Antonio, O., DeLuna, A., and Delaye, L. (2017). Natural selection drove metabolic specialization of the chromatophore in *Paulinella chromatophora*. *BMC Evol. Biol.* 17:99. doi: 10.1186/s12862-017-0947-6
- Valladares, A., Montesinos, M. L., Herrero, A., and Flores, E. (2002). An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol. Microbiol.* 43, 703–715. doi: 10.1046/j.1365-2958.2002.02778.x
- van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126. doi: 10.1126/science.1184057
- Wang, Y., Chen, C. H., Hu, D., Ulmschneider, M. B., and Ulmschneider, J. P. (2016). Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide. *Nat. Commun.* 7:13535.
- Wang, Z., Anderson, N. S., and Benning, C. (2013). The phosphatidic acid binding site of the *Arabidopsis* trigalactosyldiacylglycerol 4 (TGD4) protein required for lipid import into chloroplasts. *J. Biol. Chem.* 288, 4763–4771. doi: 10.1074/jbc.m112.438986
- Weber, A. P. M., Schwacke, R., and Flügge, U. I. (2005). Solute transporters of the plastid envelope membrane. *Annu. Rev. Plant Biol.* 56, 133–164. doi: 10.1146/annurev.arplant.56.032604.144228
- Wimley, W. C. (2010). Describing the mechanism of antimicrobial peptide action with the interfacial activity model. *ACS Chem. Biol.* 5, 905–917. doi: 10.1021/cb1001558
- Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395. doi: 10.1038/nrg2348
- Zhang, R., Nowack, E. C. M., Price, D. C., Bhattacharya, D., and Grossman, A. R. (2017). Impact of light intensity and quality on chromatophore and nuclear gene expression in *Paulinella chromatophora*, an amoeba with nascent photosynthetic organelles. *Plant J.* 90, 221–234. doi: 10.1111/tpj.13488

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Oberleitner, Poschmann, Macorano, Schott-Verdugo, Gohlke, Stühler and Nowack. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.