



HAL
open science

Phylogenetic and functional diversity of aldehyde-alcohol dehydrogenases in microalgae

Robert van Lis, Yohann Couté, Sabine Brugière, Nicolas Tourasse, Benoist
Laurent, Wolfgang Nitschke, Olivier Vallon, Ariane Atteia

► **To cite this version:**

Robert van Lis, Yohann Couté, Sabine Brugière, Nicolas Tourasse, Benoist Laurent, et al.. Phylogenetic and functional diversity of aldehyde-alcohol dehydrogenases in microalgae. *Plant Molecular Biology*, 2021, 105 (4-5), pp.497-511. 10.1007/s11103-020-01105-9 . hal-03404235

HAL Id: hal-03404235

<https://hal.science/hal-03404235v1>

Submitted on 3 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6 *Phylogenetic and functional diversity of aldehyde-alcohol dehydrogenases in microalgae*
7

8
9 Robert van Lis^{1,2}, Yohann Couté³, Sabine Brugière³, Nicolas J. Tourasse⁴, Benoist Laurent⁵, Wolfgang
10 Nitschke¹, Olivier Vallon⁴ and Ariane Atteia^{1,6,*}
11
12
13

14 ¹Aix Marseille Université, CNRS, BIP UMR 7281, Marseille, France

15 ²LBE, Univ Montpellier, INRAE, Narbonne, France

16 ³Univ Grenoble Alpes, CEA, INSERM, IRIG, BGE, Grenoble, France

17 ⁴UMR7141 CNRS-Sorbonne Université, Institut de Biologie Physico-Chimique, Paris, France

18 ⁵FR 550 CNRS, Institut de Biologie Physico-Chimique, Paris, France

19 ⁶MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France.
20
21
22
23
24

25 *Corresponding author: Ariane Atteia

26 E-mail address: ariane.atteia@cnrs.fr

27 Present address: MARBEC, Station Ifremer, Avenue Jean Monnet, Sète, France
28
29
30
31
32
33
34
35
36
37
38

39 **Abstract**

40
41 Aldehyde-alcohol dehydrogenases (ADHEs) are large metalloenzymes that typically perform the two-step
42 reduction of acetyl-CoA into ethanol. These enzymes consist of an N-terminal acetylating aldehyde
43 dehydrogenase domain (ALDH) and a C-terminal alcohol dehydrogenase (ADH) domain. ADHEs are present in
44 various bacterial phyla as well as in some unicellular eukaryotes. Here we focus on ADHEs in microalgae, a
45 diverse and polyphyletic group of plastid-bearing unicellular eukaryotes. Genome survey shows the uneven
46 distribution of the *ADHE* gene among free-living algae, and the presence of two distinct genes in various species.
47 We show that the non-photosynthetic Chlorophyte alga *Polytomella* sp. SAG 198.80 harbors two genes for
48 ADHE-like enzymes with divergent C-terminal ADH domains. Immunoblots indicate that both ADHEs
49 accumulate in *Polytomella* cells growing aerobically on acetate or ethanol. ADHE1 of ~105-kDa is found in
50 particulate fractions, whereas ADHE2 of ~95-kDa is mostly soluble. The study of the recombinant enzymes
51 revealed that ADHE1 has both the ALDH and ADH activities, while ADHE2 has only the ALDH activity.
52 Phylogeny shows that the divergence occurred close to the root of the *Polytomella* genus within a clade formed
53 by the majority of the Chlorophyte ADHE sequences, next to the cyanobacterial clade. The potential
54 diversification of function in *Polytomella* spp. unveiled here likely took place after the loss of photosynthesis.
55 Overall, our study provides a glimpse at the complex evolutionary history of the ADHE in microalgae which
56 includes *i*) acquisition via different gene donors, *ii*) gene duplication and *iii*) independent evolution of one of the
57 two enzymatic domains.

58
59
60 **Keywords:** Alcohol dehydrogenase (ADH), bifunctional enzyme, gene duplication, *Polytomella*, spirosome

61
62 **Declarations**

63 **Funding**

64 This work was supported by the Centre National de la Recherche Scientifique (CNRS), Aix-Marseille Université
65 (AMU) and the LABEX Dynamo (ANR-11-LABX-0011-01). Proteomic experiments were partly supported by
66 the Agence Nationale de la Recherche (ANR-10-INBS-08-01 ProFI grant).

67
68 **Conflicts of interest**

69 The authors declare that they have no conflicts of interest with the contents of this article.

70
71 **Ethics approval** (include appropriate approvals or waivers) N/A

72
73 **Consent to participate** (include appropriate statements) N/A

74
75 **Consent for publication** (include appropriate statements) N/A

76
77 **Availability of data and material** (data transparency)

78 The ADHE1 cDNA sequence can be found in the DDBJ/EMBL/GenBank databases under the accession number
79 PRJEB33479

80
81 **Code availability** (software application or custom code) N/A

82
83 **Authors' contributions**

84 AA and RvL conceived the study, designed and conducted the experiments. OV, NJT, BL and AA assembled and
85 annotated genomes. YC and SB performed the proteomic studies. AA, RvL, WN and OV performed the
86 bioinformatics analyses. AA, RvL and OV wrote the manuscript and generated the figures.

87 All authors participated and approved the final version of the manuscript.

88 **Introduction**

89

90 Microalgae represent a diverse and polyphyletic group of plastid-bearing unicellular eukaryotes. At the base of
 91 aquatic ecosystems, microalgae also contribute significantly to carbon cycling and storage. Deciphering algal
 92 genome sequences has shed light on unforeseen metabolic functions and identified targets for genetic
 93 manipulations, both for laboratory reference algae such as *Thalassiosira pseudonana* (Armbrust et al. 2004) and
 94 *Chlamydomonas reinhardtii* (Merchant et al. 2007) and for algae with a biotechnological potential such as
 95 *Chlorella* spp. (Guarnieri et al. 2018; Wu et al. 2019) and *Nannochloropsis* spp. (Radakovits et al. 2012; Wang et
 96 al. 2014). The availability of genomes of less studied algal species (Kim et al. 2014; Nelson et al. 2019; Blaby-
 97 Haas and Merchant 2019) have greatly helped our understanding of *i*) the molecular mechanisms underlying
 98 adaptation to adverse ecological niches and *ii*) the evolutionary history of the different lineages. Among the
 99 surprises revealed by algal genome sequencing is the extended repertoire of fermentative enzymes in *C.*
 100 *reinhardtii* (Atteia et al. 2006; Mus et al. 2007). Such a repertoire was long thought to be restricted to prokaryotes
 101 and anaerobic protists (Rosenthal et al. 1997; Müller et al. 2012). Multidisciplinary studies have lifted the veil on
 102 the significance of these enzymes and their integration in the global metabolism of the photosynthetic alga (Mus
 103 et al. 2007; Hemschemeier et al. 2008; Catalanotti et al. 2012; Magneschi et al. 2012; van Lis et al. 2013; van Lis
 104 et al. 2017; van Lis et al. 2020).

105 The aldehyde-alcohol dehydrogenase (ADHE) is one of the enzymes discovered in the green alga *C.*
 106 *reinhardtii* that was thought to be lacking in aerobic eukaryotes. In bacteria and amitochondriate eukaryotes,
 107 ADHE is a crucial fermentative enzyme that interconverts acetyl-CoA to ethanol with acetaldehyde as an
 108 intermediate (Kessler et al. 1991; Bruchhaus and Tannich 1994; Sánchez 1998). ADHE is a fascinating example
 109 of a non-cytoskeleton enzyme that can self-assemble into long filaments. The ability of the ADHE to auto-
 110 assemble *in vivo* into large right-handed helices, also named spiroosomes, was first described in *Lactobacillus*
 111 (Kawata et al. 1975), and later in *Escherichia coli* (Kessler et al. 1992) and in the anaerobic protist *Entamoeba*
 112 *histolytica* (Bruchhaus and Tannich 1994). Recent cryo-electron microscopy studies of *E. coli* ADHE revealed
 113 the mechanism of spiroosome formation and highlighted residues critical for oligomerization (Kim et al. 2019;
 114 Pony et al. 2020).

115 The ADHE in the green alga *C. reinhardtii* resembles its bacterial and eukaryotic counterparts in several
 116 aspects: *i*) it has an *N*-terminal acetylating aldehyde dehydrogenase domain and a *C*-terminal alcohol
 117 dehydrogenase domain; *ii*) it catalyzes the sequential NADH-dependent reductions of acetyl-CoA to acetaldehyde
 118 and then to ethanol and *iii*) has the ability to form spiroosomes (van Lis et al. 2017). However, the algal enzyme
 119 differs from previously described ADHEs in its intracellular localization, *i.e.* it is a chloroplast protein, and in its
 120 regulation, in particular the fact that it is expressed at ambient oxygen levels and the correlation to starch
 121 accumulation (van Lis et al. 2017). In the unicellular alga *Polytomella* sp. SAG 198.80 (*Polytomella* sp.), a close
 122 relative of *C. reinhardtii* that lacks photosynthetic pigments and a plastid genome (Pringsheim 1955; Smith and
 123 Lee 2014), an ADHE-like protein was also identified (Atteia et al. 2003). However, the enzymatic activities of
 124 the protein, which resides in the mitochondrion (Atteia et al. 2003), have never been investigated. Since the
 125 discovery of *ADHE* genes in the Chlorophytes *Polytomella* and *C. reinhardtii*, homologous genes have been found
 126 in other lineages, including in algae harboring secondary plastids such as the Rhizarian *Bigeloviella natans* and
 127 the Cryptophyte *Guillardia theta* (Atteia et al. 2013). As in *C. reinhardtii* and *Polytomella* sp., the *ADHE* gene in
 128 *B. natans* and *G. theta* is expressed at ambient oxygen levels (Gould et al. 2019), unlike in bacteria.

129

130 In the course of sequencing the genome of *Polytomella* sp., we found as expected the *ADHE* gene whose
 131 product has been described earlier as a major mitochondrial protein (Atteia et al. 2003), but also a second gene
 132 for a putative *ADHE*. In this study, we validate the presence of two distinct *ADHE*-like proteins in the non-
 133 photosynthetic alga, and we provide a comparison of their biochemical and enzymatic characteristics. A genomic
 134 and phylogenetic study of *ADHE*s in unicellular algae reveals novel aspects of the origin and evolution of *ADHE*s
 135 in mitochondriate eukaryotes.

136

137

138 **Results**

139

140 **Structural features of *Polytomella* sp. ADHEs inferred by their primary sequences**

141

142 The genome sequence of *Polytomella* sp. revealed, in addition to the previously described *ADHE* gene (Atteia et al. 2003), another gene for a putative ADHE. For reasons detailed below, we call the newly uncovered gene
 143 *ADHE1*, while the *ADHE* gene firstly described is renamed *ADHE2*. The overall primary sequence comparison
 144 indicates that *PsADHE1* is 104-aa larger than *PsADHE2*, but also significantly less related to *PsADHE2* than to
 145 its counterpart in the photosynthetic alga *C. reinhardtii* (*CrADHE1*), sharing respectively 60% and 72% amino
 146 acid identity. To bring insights into the differences between the ADHEs in the colorless alga, their amino acid
 147 sequences were aligned with that of *E. coli* (*EcADHE*), to date the best studied bacterial ADHE, and that of *C.*
 148 *reinhardtii*, the only algal ADHE characterized so far. The most relevant diverging features are depicted in Fig.
 149 1A. As compared to *PsADHE2*, *PsADHE1* possesses extended *N*- and *C*-termini. The *C*-terminal extension of
 150 *PsADHE1* (55-aa) is characterized by a high content in proline and hydroxylated residues (Supplementary Fig.
 151 S1). Compared to their homologues in *E. coli* (and other bacteria), the algal ADHEs have an *N*-terminal extension,
 152 indicative of intracellular targeting. As shown earlier, the 27-aa *N*-terminus in *PsADHE2* serves as a targeting
 153 sequence to the mitochondrion (Atteia et al. 2003). This localization is predicted with high confidence by the
 154 public software programs MitoFate, TargetP and Predotar (Table 1). In contrast, a plastid localization is predicted
 155 by the software Predalgo, designed to predict the intracellular targeting of precursor proteins of *C. reinhardtii* and
 156 other Chlorophytes (Tardif et al. 2012). For *PsADHE1*, the programs agreed to various extents on a plastid
 157 targeting prediction (Table 1). As no adequate protocol to obtain pure fractions of the nonphotosynthetic plastids
 158 and mitochondria of the colorless alga is available at present, the plastid localization of *ADHE1* remains
 159 hypothetical.
 160

161 The multiple sequence alignment indicates that the ADHEs in *Polytomella* diverge most in the alcohol
 162 dehydrogenase (ADH) domain (Supplementary Fig. S1). While *PsADHE1* harbors the two Fe-ADH signature
 163 patterns (Prosite ADH_iron_1, PS00913; ADH_iron_2, PS00060), *PsADHE2* has a poorly conserved
 164 ADH_iron_2 signature, lacking in particular the two histidine residues required for iron coordination (Montella
 165 et al. 2005) (Conserved Domains database cd08178). Furthermore, in contrast to *PsADHE1*, *PsADHE2* lacks a
 166 few residues identified by cryo-EM as critical for *E. coli* ADHE auto-assembly (Kim et al. 2019). In particular,
 167 the aromatic residue between the two Fe-ADHE signatures (F670 in *EcADHE*, Y735 in *PsADHE1*), suggested to
 168 be critical for spiroosome formation, is replaced by an arginine in *PsADHE2* (R699) (Supplementary Fig. S1). The
 169 distinct features revealed by sequence analysis warranted a detailed biochemical and enzymatic comparative study
 170 of the two *Polytomella* ADHEs.
 171

171

172 **Kinetic analysis of recombinant *PsADHE1* and *PsADHE2***

173

174 ADHE typically catalyzes the reversible conversion of acetyl-CoA into ethanol with an acetaldehyde intermediate
 175 (Fig. 1C). Four distinct reactions can be assessed *in vitro* by spectrophotometry measuring the oxidation of NADH
 176 or the reduction of NAD⁺, as shown earlier with the *C. reinhardtii* ADHE (van Lis et al. 2017). To get insights
 177 into the respective enzymatic properties of the two *Polytomella* ADHEs the proteins were produced
 178 heterologously in *E. coli*. Coding sequences lacking the putative targeting sequences were cloned into the pET24a
 179 expression vector (primers sequences are given in Supplementary Table S1). A truncated version of *PsADHE1*
 180 lacking the last 41 residues (Supplementary Fig. S1) was also produced to investigate the significance of the *C*-
 181 terminal extension. The recombinant ADHEs were purified via affinity chromatography (Fig. 1B), taking
 182 advantage of the hexahistidine (6xHis)-tag introduced at their *C*-terminus by the vector. The ability of each of the
 183 three enzymes to catalyze the different reactions (V_{max}) was assessed. The acetaldehyde formed in reaction #1 is
 184 in direct proximity to the ADH active site and reacts away to ethanol immediately (reaction #2) since both
 185 reactions use NADH, producing NAD in a 1:1 stoichiometry. Therefore, it can be considered that the V_{max}
 186 measured in presence of acetyl-CoA and NAD represents the combined activities of reaction #1+2, with half of
 187 this value corresponding to the actual V_{max} of reaction #1. Measurement of reaction #1 alone may be approached

188 using EDTA to chelate the Fe of the ADH-moiety and inhibit reaction #2. However, this approach was not
 189 successful in previous studies on the CrADHE and was not attempted here (van Lis et al. 2017).
 190 As shown in Table 2, purified recombinant ADHE1 (rADHE1) catalyzed the NADH-dependent ALDH and ADH
 191 reactions (#1, 2) as well as the reverse reactions (#3, 4). Thus, *Ps*ADHE1 appears as a genuine bifunctional ADHE.
 192 The study of the truncated ADHE1 (rADHE1_t) did not reveal significant differences with respect to rADHE1
 193 for reactions #1, 2 and 3 (Table 2). However, the activity of rADHE1_t for reaction#4 was found to be 2.5 to 3
 194 times higher than that of rADHE1, suggesting a regulatory role in the conversion of acetaldehyde to acetyl-CoA
 195 for the C-terminal extension. Recombinant ADHE2 (rADHE2) catalyzed the acetaldehyde formation from acetyl-
 196 CoA and NADH (reaction#1) but failed to form ethanol from acetaldehyde and NADH (reaction#2). The rate of
 197 conversion of acetyl-CoA to acetaldehyde was slightly lower for rADHE2 than rADHE1, with 1.52 and 1.88
 198 U/mg, respectively. In contrast, the CoA-dependent acetaldehyde oxidation activity (reaction#4) for rADHE2, of
 199 23 U/mg, was significantly higher than that for rADHE1 of 0.27 U/mg. The high activity measured for the
 200 conversion of acetaldehyde into acetyl-CoA by rADHE2 might be related to a specific physiological role (see
 201 below). The fact that rADHE2 has no ADH activity was not so surprising in view of the absence of key residues
 202 for metal coordination (Fig. 1A).

203

204 **Blue native (BN) PAGE analysis of recombinant ADHE2**

205

206 BN-PAGE is a powerful method to analyze the oligomeric states and associated activities of soluble as well as
 207 membrane-bound enzymes (Schägger and von Jagow 1991; Schägger et al. 1994). By applying soluble fractions
 208 of *Polytomella* crude mitochondria on BN-PAGE, the native ADHE2 was previously identified as a dimer of 180
 209 kDa (Atteia et al. 2003; van Lis et al. 2005). We used this method here to analyze the recombinant ADHE2 of
 210 *Polytomella* with respect to that of *C. reinhardtii* (*Cr*-rADHE1), which was shown to separate in two predominant
 211 oligomeric forms, *i.e.* a dimeric and a tetrameric form, with a NAD-dependent ADH activity (van Lis et al. 2017).
 212 As shown in Fig. 2, the migration of rADHE2 is fuzzy with the prevalence of a monomeric and a dimeric form.
 213 Unlike with *Cr*-rADHE1, none of these rADHE2 forms has the ability to oxidize ethanol (Fig. 2, “ADH activity”),
 214 which corroborates the *in vitro* activity measurements (Table 2). In the light of the recently reported cryo-EM
 215 structures (Kim et al. 2019; Pony et al. 2020), it may be possible that the oligomeric ADHE forms identified on
 216 BN-PAGE for *C. reinhardtii* correspond to the basic unit (dimer) and the next immediate unit (tetramer) in the
 217 spirosome formation. While the dimer formation occurs in a head-to-head arm-crossing fashion independent of
 218 the ADH domain, the tetramer results from the hydrophobic interactions between the ADH domains from a dimer
 219 in a tail-to-tail manner (Kim et al. 2019). The seeming lack of a tetramer for rADHE2 might be related to the
 220 replacement of several residues in the ADH domain that are key for the interactions, in particular the aromatic
 221 residue (F670, *E. coli* numbering) and the residues which form the hydrophobic pocket (Fig. 1A; Supplemental
 222 Fig. S1).

223

224 **Accumulation of ADHEs in *Polytomella* sp. cells**

225

226 To gain insights into the growth conditions leading to the accumulation of ADHE1 and ADHE2 in the non-
 227 photosynthetic alga, we carried out immunoblot experiments using a polyclonal antibody against bacterial and
 228 algal ADHEs (anti-ALDH/ADH). As shown in Fig 5A, this serum detects with comparable signals the two
 229 recombinant ADHEs, thereby allowing a relative quantification of the proteins in total cell extracts. We first
 230 probed cells grown on ethanol at pH 6.0, in which ADHE2 was shown earlier to be abundant (Atteia et al. 2003).
 231 In these cells, immunoblotting experiments identified two distinct proteins (Fig. 3A). Based on the relative
 232 electrophoretic migration of the overexpressed proteins, we assigned the larger protein of ~105-kDa to ADHE1,
 233 and the smaller protein of ~95-kDa to ADHE2. Looking at the relative intensities of their signals, it can be
 234 concluded that ADHE1 is significantly less abundant than ADHE2. In *Polytomella* sp. cells grown on acetate at
 235 pH 6.0, the relative abundance between ADHE1 and ADHE2 was found to be comparable to that in ethanol-grown
 236 cells (Fig. 3B, lane 2), suggesting that the carbon source has no direct effect on the accumulation of the ADHEs.
 237 In extracts from acetate-grown cells, mass spectrometry-based proteomics identified both ADHEs with a large

238 number of peptides (Supplementary Table S2). The estimation of the ADHE relative abundances using intensity-
 239 based absolute quantification (iBAQ) (Schwanhäusser et al. 2011) is in agreement with ADHE2 being
 240 significantly more abundant than ADHE1, with an extracted iBAQ value for ADHE2 more than five times higher
 241 than for ADHE1 (Supplementary Table S3). As shown in Fig. 3B, *Cr*ADHE1 migrates slightly faster than
 242 *Ps*ADHE1. This could indicate that the C-terminal extension of *Ps*ADHE1 (Fig. 1A) is present, at least partially,
 243 in the mature protein. The effect of anoxia on the respective accumulation levels of both ADHE was assessed
 244 following a protocol described earlier for *C. reinhardtii* (van Lis et al. 2017). Data indicate that anoxia has no
 245 significant effect on the levels of ADHEs (Supplementary Figure S2). Similar findings were reported for the
 246 ADHE of *C. reinhardtii* (van Lis et al. 2017).

247 It was reported earlier that *Ps*ADHE2 is a soluble protein (Atteia et al. 2003). Fractionation of
 248 *Polytomella* cells after sonication and ultracentrifugation confirmed that ADHE2 is predominantly a soluble
 249 protein (Fig. 3B, lane 4). In contrast, ADHE1 is found in the particulate fraction (Fig. 3B, lane 3), where it could
 250 be associated with membranes or assembled in spiroosomes.

251

252

253 **Overview of the distribution of the ADHE genes among microalgae**

254

255 The growing availability of sequenced genomes and transcriptomes allowed reassessment of the distribution and
 256 evolution of aldehyde-alcohol dehydrogenase genes among microalgae. The unicellular algae surveyed here
 257 belong to the supergroup of Archaeplastida, which harbor a primary plastid, as well as to supergroups or clades
 258 bearing secondary plastids (SAR, Haptista and Cryptista clades). As shown in Supplementary Table 4, less than
 259 half of the species analyzed (20/49) have a gene for a *bona fide* ADHE, defined here by the presence of two
 260 consecutive domains with similarities to ALDH and ADH (Fig. 1A). This is clearly not due to incomplete genome
 261 annotation, but reflects the gene loss in most algal clades. As observed in Supplementary Table S4, the majority
 262 of the algae harboring an *ADHE* gene fall within the Chlorophyte division (*i.e.* green algae) of Archaeplastida, as
 263 do *Polytomella* and *C. reinhardtii*. In Chlorophyceae and Trebouxiophyceae, the *ADHE* gene is common although
 264 some species seem to lack it. Amid the Mamiellophyceae, the tiniest Chlorophytes known to date (Piganeau et al.
 265 2011), no *ADHE* gene could be found. In the algae with secondary plastids, a scattered distribution of the gene is
 266 observed (Supplementary Table S4). To date, *B. natans* is the only species with a secondary plastid deriving from
 267 a green alga for which the genome sequence is publicly available. In this alga, a genuine *ADHE* gene is present,
 268 as reported earlier (Atteia et al. 2013). As shown in Supplementary Table S4, several species with a plastid
 269 deriving from a red alga contain an *ADHE* gene. This is unexpected as of today the gene for the bifunctional
 270 enzyme has not been described in red algae (Rhodophytes) (Supplementary Table S4). Among the eight
 271 Stramenopiles surveyed here, the marine diatom *Fistulifera solaris* is the only one with an *ADHE* gene (in fact it
 272 has two highly related genes). In the genomes of the two free-living chromerids (Alveolates) sequenced to date,
 273 *Chromera velia* and *Vitrella brassicaformis*, genuine *ADHE* genes were found. Finally, the survey indicates the
 274 absence of an *ADHE* gene in the Haptista *Emiliana huxleyi* but its presence in the Cryptista *G. theta*. If we
 275 consider the environment in which the microalgae thrive, it seems that the *ADHE* gene is more common in
 276 freshwater species than in marine species (Supplementary Table S4).

277 Based on the multiple alignment of 105 eukaryotic and bacterial protein sequences (Supplementary Table
 278 S5) and the reconstruction of an unrooted Maximum Likelihood phylogenetic tree (Fig. 4), we found that ADHEs
 279 from the most evolved Chlorophyte clades (Trebouxiophyceae, Chlorophyceae) group together, close to the clade
 280 of Cyanobacteria. This proximity suggests a vertical inheritance of the *ADHE* gene in this algal division, including
 281 for *Polytomella* sp. SAG 198.80. The remaining green algal ADHEs sequences, which belong to the other clades
 282 (Chlorodendrophyceae, Palmophyllophyceae, Prasinophyceae) are found within a composite eukaryotic group
 283 made of sequences from amitochondriate eukaryotes and from the algae *B. natans* and *G. theta*, described earlier
 284 (van Lis et al. 2017). Within this same group are also found the sequences of the Alveolates *C. velia* and *V.*
 285 *brassicaformis*. The position of the two *ADHE* sequences of *F. solaris* in a clade formed by the proteobacterial
 286 sequences might indicate gene acquisition by lateral transfer or a contaminant bacterial DNA.

287 Our gene survey revealed that half of the algal species (10/19) harbor two distinct ADHE genes
 288 (Supplemental Table S4). In subsections of the tree containing algal ADHEs, the paralogs of a given species
 289 usually group together, suggesting recent gene duplication or possibly allelic divergence as for *Picochlorum*
 290 NBRC 102739 or *Pyramimonas amyliifera*. In *C. reinhardtii*, two pseudogenes were found next to one another on
 291 chromosome 16, highly similar to *CrADHE1*, testifying of the tendency of ADHE to duplicate. Chlorellales were
 292 different in that the divergence of the two paralogs occurred early, *i.e.* before the divergence of *Micractinium* and
 293 *Chlorella*. An ancestral duplication appears very likely for the *Polytomella* genus. Indeed, the tree shown in Fig.
 294 4, the ADHEs of the colorless algae are found close to each other surrounded by most green algal ADHEs. This
 295 suggests that ADHE1 and ADHE2 proteins have evolved by duplication from a cyanobacterial-type ancestral
 296 gene. The long branch length of *PsADHE2* (Fig. 4, sequence 1) strongly suggests an unusually rapid sequence
 297 divergence in *Polytomella* sp. (which translated into unusual biochemical properties as discussed above) in clear
 298 contrast with the situation in the photosynthetic algal species.

299 Illumina sequencing of genomic DNA is available for four known species of the *Polytomella* genus, *P.*
 300 *capuana*, *P. magna*, *P. parva*, and *P. piriformis* (Smith and Lee 2014), which are distinguished by the structure
 301 of their mitochondrial DNA (Smith et al. 2013). The genomic data of these species were searched for the presence
 302 of *ADHE* genes. In all four species, two *ADHE* genes were found. While one gene codes for a genuine ADHE
 303 (named ADHE1), the other codes for an atypical ADHE showing divergence in the C-terminal part of the protein
 304 (named ADHE2). The predicted ADHE2 sequences were analyzed for the deriving features revealed in this study,
 305 *i.e.* the lack of the Fe-coordinating residues and of the aromatic residue required for auto-assembly. As depicted
 306 in Fig. 5, the ADHE2 sequences in *P. piriformis* and *P. parva* are comparable to *PsADHE2* (*Polsp_ADHE2*). In
 307 contrast, the sequences of the other two species show only one of the two distinctive features. In *P. magna*, the
 308 ADHE2 lacks the aromatic residue (a valine instead) necessary for spiroosome formation but has residues required
 309 for iron-binding. In *P. capuana*, the opposite is observed.

310 The comparison of the C-terminal ADH domain of ADHE1 paralogs in algal species indicates a high
 311 degree of conservation, including between *CrADHE1* and one of the pseudogenes (- Ψ 1, Fig. 5). It is however
 312 interesting to note that all but one *Polytomella* species (*P. magna*) have a longer C-terminus.

313

314

315 Discussion

316

317 The non-photosynthetic alga *Polytomella* sp. harbors two enzymes related to bacterial ADHEs

318

319 Aldehyde-alcohol dehydrogenases are bifunctional enzymes long thought to be restricted to Bacteria and
 320 eukaryotes devoid of mitochondria, where they play a crucial role in fermentative metabolism (Kessler et al. 1991;
 321 Bruchhaus and Tannich 1994; Boxma et al. 2004). The occurrence of ADHE-like proteins in microalgae, which
 322 are mitochondriate eukaryotes, was interesting in several respects, especially for their integration in an aerobic
 323 metabolic context and their evolutionary origin. We have shown recently that the photosynthetic alga *C.*
 324 *reinhardtii* harbors an ADHE that resembles its bacterial counterparts in *i)* the enzymatic activities catalyzed and
 325 *ii)* the ability to oligomerize into spiroosomes. The phylogeny pointed to the vertical inheritance of the algal enzyme
 326 from a cyanobacterial ancestor (van Lis et al. 2017). With the study of the ADHE-related enzymes in *Polytomella*
 327 sp. SAG 198.80 and other *Polytomella* species, we disclosed fascinating facets of the evolution of the so-called
 328 bifunctional enzyme.

329 Significant molecular and biochemical data on the oxidative phosphorylation complexes of *Polytomella*
 330 sp. have been obtained over the years that provided further evidence for the phylogenetic relationship of the
 331 colorless alga with the green alga *C. reinhardtii* (Atteia et al. 1997; Funes et al. 2002; van Lis et al. 2005; van Lis
 332 et al. 2007). The metabolic routes existing in *Polytomella* have in contrast been poorly investigated and
 333 consequently, the knowledge on their mechanisms underlying highly versatile heterotrophic capacities is scarce
 334 (Wise 1955; Wise 1959; Atteia et al. 2000). Through this study we provide insights into the aldehyde/alcohol
 335 metabolism unexplored so far in *Polytomella* species. Based on its sequence similarity to *C. reinhardtii* ADHE1
 336 (Fig. 1A, Supplementary Fig. S1) and enzymatic activities (Table 2), it seems plausible that *Polytomella* sp.

337 ADHE1 functions in the same way, that is in the ethanolic fermentation fueled by starch glycolysis (Magneschi
 338 et al. 2012; van Lis et al. 2017). Although *Polytomella* sp. ADHE2 shares a number of features with *CrADHE1*
 339 and bacterial ADHEs, our experimental data clearly show that it is not a typical aldehyde-alcohol dehydrogenase.
 340 First of all, *PsADHE2* is impaired in ethanol dehydrogenase activity (Table 2; Fig. 2). Furthermore, *PsADHE2* is
 341 a soluble protein occurring as a dimer (Atteia et al. 2003; van Lis et al. 2005) which likely follows from the lack
 342 of the key residues required for spiroosome formation (Fig. 1A; Supplemental Fig. S1). The absence of an ADHE2
 343 in *C. reinhardtii* illustrates the fact that making predictions on the metabolic enzymes in *Polytomella* sp. from *C.*
 344 *reinhardtii* studies may have its limits, and emphasizes the importance de develop studies dedicated to the
 345 colorless alga. The ADHE gene duplication and divergence (of function and possibly of cellular localization) in
 346 the non-photosynthetic alga might be seen as beneficial for the adaptation to a new lifestyle (Zimorski and Martin
 347 2014; Copley 2020).

348 To our knowledge, this work is the first report of a full-length ADHE that performs only one of the two
 349 typical enzymatic reactions. The presence of such an enzyme in the colorless alga is puzzling and raises questions
 350 about its metabolic integration. Indeed, in nature there exist acetylating aldehyde dehydrogenases (*AcAldDH*)
 351 (EC 1.2.1.10) which catalyze the same overall reaction as *PsADHE2*, that is the interconversion of acetaldehyde
 352 and acetyl-CoA using NADH as a co-factor (Burton and Stadtman 1953). These mono-enzymatic proteins of ~45
 353 kDa can catalyze the formation of acyl-CoAs, such as the conversion of acetaldehyde to acetyl-CoA in the
 354 utilization of ethanolamine by *Salmonella typhimurium* (Roof and Roth 1992), or the formation of aldehydes in
 355 alcohol-producing bacteria such as *Clostridium beijerinckii* (Toth et al. 1999). The reason for keeping a full-length
 356 ADHE with a non-functional alcohol dehydrogenase domain is at present unclear, but it is tempting to propose
 357 that ADHE2 bears an as yet unknown additional function. It is worth mentioning here that the ADHE in *E. coli*
 358 was found to act as an ATP-dependent RNA helicase through the associate with the small subunit head domain
 359 of the ribosome (Shasmal et al. 2016; Kurylo et al. 2018).

360 In *Polytomella*, we found that the relative abundance of the ADHEs is comparable in cells growing on
 361 ethanol and on acetate (pH 6.0) (Fig. 3), which is not in favor of the involvement of ADHE2 (and ADHE1) in
 362 ethanol assimilation. Earlier results on ethanol-growing *Polytomella* sp. indicated that the pH of the medium had
 363 an impact on the expression and accumulation levels of ADHE2, with high levels at pH 6.0 but much lower at pH
 364 3.7 (Atteia et al. 2003). The external pH as a regulatory factor may provide hints as to the connection between
 365 ADHE2 and mitochondrial respiration, since it was found that *i*) ADHE2 steady-state levels were inversely
 366 regulated to those of respiratory complexes I, III, IV and to the levels of cytochrome *c*, and *ii*) the respiratory rates
 367 were 20-25% lower in isolated mitochondria from cells grown at pH 6.0 ('high' ADHE2) than from cells grown
 368 at pH 3.7 ('low' ADHE2) (Atteia et al. 2000). It may be proposed that a downregulation of OXPHOS activity by
 369 the external pH alters TCA cycle activity necessitating a metabolic rerouting involving reductant and CoA pools,
 370 in which ADHE2 has a role to play.

371 Assuming that native ADHE2 only catalyzes the interconversion of acetyl-CoA and acetaldehyde as is
 372 the case for the recombinant enzyme (Table 2), the question arises as to the directionality of the reaction catalyzed
 373 *in vivo* by this mitochondrial protein. At pH 6.0 on acetate or on ethanol, *PsADHE2* is very abundant. When
 374 functioning in the forward direction (Fig. 1C, reaction #1), its activity would lead to the production of
 375 acetaldehyde from acetyl-CoA. Acetaldehyde is highly toxic for the mitochondrial function, and must be
 376 metabolized quickly to prevent damage of proteins and DNA (Cederbaum et al. 1973; Cederbaum et al. 1974). In
 377 the case of *E. coli* ADHE, it is proposed that the acetaldehyde remains sequestered inside the spiroosome structure
 378 and is converted into ethanol at the inner surface without being released into the cell's cytoplasm (Kim et al.
 379 2019). In mitochondria, acetaldehyde detoxification typically occurs via CoA-independent aldehyde
 380 dehydrogenase, and results in the production of NADH and acetate. In the *Polytomella* sp. genome, seven distinct
 381 mon-enzyme ALDH genes were identified, and two of them code for enzymes predicted to reside in the
 382 mitochondrion. Setting up tools to follow the accumulation and activity of ALDH enzymes in isolated organelles
 383 is required to clarify their role in relation to ADHE2. In the proteomic-based analysis of the acetate-grown cells,
 384 no ALDH was identified at levels comparable to those of ADHE2 (Supplementary Table S3). When functioning
 385 in the reverse direction (Fig. 1C, reaction #4), ADHE2 would use acetaldehyde to produce acetyl-CoA, a central
 386 metabolite (Pietrocola et al. 2015). It has been shown earlier that *Polytomella caeca* can use acetaldehyde as

387 growth substrate (Wise 1968; Wise 1970), but the assimilation pathway was not described. In bacteria,
 388 acetaldehyde is involved in metabolic pathways involving various carbohydrate, lipid, and aromatic compounds
 389 (Lees and Jago 1978). In *Polytomella* sp. ADHE2 may be used to assimilate other C-sources, but a source of
 390 acetaldehyde in the mitochondrial matrix must be envisioned. In *Polytomella* cells grown on acetate or on ethanol,
 391 acetaldehyde could be produced in the matrix from ethanol, as in the yeast *Saccharomyces cerevisiae*, by one of
 392 at least nine mono-enzymatic ADH predicted from the *Polytomella* sp. genome. Ethanol is indeed a good carbon
 393 source for growth of *Polytomella* (Atteia et al. 2000). However, we showed here that ADHE2 is also highly
 394 expressed in cells growing on acetate (Fig. 3). Acetate assimilation occurs as in any acetoflagellate via acetyl-
 395 CoA synthase likely in the cytosol. Without a better knowledge of the metabolic enzymes and routes existing in
 396 the colorless alga, it is difficult to propose a pathway from cytosolic acetyl-CoA to mitochondrial acetaldehyde.
 397 In any case, the abundance of ADHE2 prompts further examination of acetaldehyde as important metabolite in
 398 *Polytomella*. Elucidation of the metabolic integration of the ADHE-like enzyme in the colorless alga will have to
 399 rely on the use of large-scale analyses including proteomics and metabolic fluxes but also on biochemical
 400 approaches.

401

402 **Diversity of ADHEs in unicellular plastid-bearing eukaryotes**

403

404 The genome survey carried out in this work shows that ADHE is a common enzyme among microalgae. Our
 405 phylogenetic tree does not allow a full reconstitution of the deep phylogeny of ADHEs, but does reveal a complex
 406 history within Chlorophytes (Archaeplastida supergroup). As illustrated here, there seems to be a tendency for
 407 gene duplication and maintenance of *bona fide* ADHE duplicates among Chlorophytes, especially in the
 408 Chlorophyceae and Trebouxiophyceae (Fig. 4, Supplementary Table S4,). Sometimes, diverging evolution of one
 409 of the ADHE duplicates can occur as exemplified here with the nonphotosynthetic *Polytomella* species.
 410 Interestingly in these species, most mutations seem to have occurred in the C-terminal with two distinct and non-
 411 exclusive outcomes: *i*) the loss of ADH activity (mutations of the iron-coordinating residues) and *ii*) the loss of
 412 spiroosome formation ability (mutation of the aromatic residue). While the ADHE2 in *P. magna*, the most ancestral
 413 of the *Polytomella* species (Smith et al. 2013), retains the iron-binding and active site residues of the ancestor, it
 414 has lost the key aromatic residue for spiroosome formation. We therefore predict that ADHE2 in *P. magna*, while
 415 retaining its capacity for ethanol production/oxidation, will not be found in the particulate fraction but rather as a
 416 soluble protein. This raises the question of the potential toxicity of the acetaldehyde produced in the absence of
 417 spiroosome sequestration. Conversely, *P. capuana* ADHE2, which conserves the aromatic residue but has lost the
 418 residues necessary for ADH activity, is likely to be able to self-assemble. Without the residues for iron-
 419 coordination and the aromatic residue for the spiroosome formation, the ADHE2 in *Polytomella* sp. SAG 198.80
 420 and in the species *P. piriformis* and *P. parva* represent the most divergent duplicates.

421 Since *Polytomella* sp. is the only alga known to date to harbor a full-length ADHE with a catalytically-
 422 defective ADH domain, it is tempting to correlate this evolution to the loss of photosynthetic activity in some
 423 way. However, it is clearly not a prerequisite for loss of photosynthesis, as it has not occurred in *P. magna*. The
 424 different *Polytomella* species represent different stages of functional adaptation of one of the duplicates, but the
 425 link between acquisition of mitochondrial targeting sequence and the loss of ADH activity or ability to auto-
 426 assemble, and the generalized metabolic consequences of the ADHE2 evolution, remains to be studied.

427

428 **Concluding remarks**

429

430 With the study of ADHEs in the algae *C. reinhardtii* and *Polytomella* we just begin to understand the significance
 431 of this fascinating multifunctional enzyme amid eukaryotes that thrive in oxic conditions, and to decipher the
 432 evolutionary steps since gene acquisition in the Archaeplastida supergroup. The original properties of *Polytomella*
 433 ADHE2 unveiled by this work are likely to serve as an inspiration for enzymatic engineering, since ADHEs are
 434 promising protein targets in biotechnological applications in ethanol/butanol production, especially in *Clostridium*
 435 spp.

436

437
438

439 **Materials and Methods**

440
441
442

441 **Algal strain and growth conditions**

443 *Polytomella* strain SAG 198.80 (*Polytomella* sp.) was obtained from the SAG algal culture collection of the
444 University of Göttingen (<http://sagdb.uni-goettingen.de/>) and axenized using ampicilline. *Polytomella* sp. was
445 cultivated either on acetate (MAP medium) or ethanol (MEP medium) in low light at room temperature (24°C)
446 without shaking. Both MAP and MEP culture media contain 35 mM MES (Sigma Aldrich), 0.8 mM K₂HPO₄, 0.2
447 mM KH₂PO₄, 7.5 mM NH₄Cl, 0.3 mM CaCl₂, 0.5 mM MgSO₄. Oligoelements are present at 1.39 µM ZnSO₄, 0.8
448 µM H₃BO₃, 2.65 µM MnSO₄, 0.74 µM FeCl₃, 0.83 µM NaMoO₄, 0.16 µM CuSO₄, 0.6 µM KI. The pH was
449 adjusted to 5.8 with KOH. Thiamine and cyanocobalamin are added to a final concentration of 20 µg/L and 1
450 µg/L, respectively. In both media, carbon source is added to a final concentration of 30 mM. One-liter Erlenmeyer
451 flasks containing 0.2 L of culture medium were inoculated with 4 mL of preculture in exponential growth phase.
452 Cells were harvested in their late log phase, after ~3 days (Atteia et al. 2000).

453

454 **Cell fractionation and preparation of cell extracts for mass spectrometry**

455

456 Acetate-grown cells were harvested (3,000 x g for 10 min) at 4°C and resuspended in HEPES 50 mM (pH 7.5)
457 supplemented with the protein inhibitors 1mM amino caproic acid, 0.5 mM benzamidine). Cells were broken by
458 sonication and briefly further centrifuged at 1,500 rpm to remove insoluble material (starch) and unbroken cells.
459 The resulting supernatant was either analyzed directly by mass-spectrometry to identify and evaluate the relative
460 abundance of the algal ADHEs. Alternatively, the supernatant was ultracentrifuged at 4°C (130,000 x g for one
461 hour) to separate the soluble and particulate fractions. The resulting pellet was washed once in HEPES medium
462 (with protease inhibitors).

463

464 **PAGE analysis, immunodetection of ADHEs, and in-gel activity**

465

466 Proteins in whole cells and cell extracts were precipitated by chloroform/methanol (Wessel and Flügge 1984), and
467 then resuspended in 2% SDS (w/v). Protein concentrations were determined using the bicinchoninic acid Pierce™
468 Protein assay kit (ThermoFisher) using bovine serum albumin as standard. For SDS-PAGE analysis, protein
469 samples (40 µg) were resuspended in 2% SDS (w/v) and 2% beta-mercaptoethanol, and boiled for 90 sec.
470 Insoluble material was removed via a 2 min-centrifugation at 13,000 x g. Apparent molecular masses were
471 estimated using Prestained protein ladder Plus (Euromedex). After electrophoresis, proteins were stained with
472 Coomassie Brilliant Blue R250 or transferred onto nitrocellulose membranes (GE Healthcare,) and stained with
473 Ponceau S. Immunoblotting was carried out as described (van Lis et al. 2017). The anti-ADH/ALDH (Agrisera)
474 was used at 1/1,000.

475

476 Blue Native-PAGE was carried out using a 3-12% linear polyacrylamide gradient as described (van Lis et al.
477 2017). The molecular mass of the rADHE oligomers were estimated using NativeMark™Unstained Protein
478 Standard (Thermofisher Scientific). Recombinant CrADHE, used as control protein, was prepared as described
479 previously (van Lis et al. 2017). For ADH activity, BN-gel was washed twice in 100 mM Tris (pH 8.8) and then
480 incubated in 25 ml of 100 mM Tris-Cl (pH 8.8) containing 20 mg of NAD, 5 mg of nitroblue tetrazolium, 1 mg
481 of phenazine methosulfate, and 25 µl of absolute ethanol.

482

483 **Mass spectrometry-based proteomics**

484

485 Protein extract from acetate-grown cells was stacked in the top of a 10% SDS-PAGE gel and stained with
486 Coomassie Brilliant blue. The band was excised and submitted to in-gel digestion using modified trypsin

487 (sequencing grade) (Promega) as previously described (Salveti et al. 2016). Resulting peptides were analyzed by
 488 online nanoliquid chromatography coupled to tandem MS (UltiMate 3000 and LTQ-Orbitrap Velos Pro, Thermo
 489 Scientific). Peptides were sampled on a 300 μm x 5 mm PepMap C18 precolumn and separated on a 75 μm x 250
 490 mm PepMap C18 column (Thermo Scientific) using a 25-min gradient for individual gel bands and a 80-min
 491 gradient for mitochondrial soluble fraction. MS and MS/MS data were acquired using Xcalibur (Thermo
 492 Scientific). Peptides and proteins were identified using Mascot (version 2.6) through concomitant searches against
 493 homemade *Polytomella* protein sequence database, classical contaminants sequence database and the
 494 corresponding reversed databases. The Proline software (<http://proline.profi-proteomics.fr>) was used to filter the
 495 results: conservation of rank 1 peptides, peptide length ≥ 7 amino acids, false discovery rate of peptide-spectrum
 496 matches $< 1\%$ as calculated on scores by employing the reverse database strategy, minimum peptide-spectrum
 497 match score of 25, and minimum of 1 specific peptide per identified protein group. Proline was then used to
 498 perform a MS1-based label-free quantification of the protein groups identified in the sample.

499

500 ***Polytomella* sp. ADHE gene cloning, expression and purification of recombinant enzymes**

501 Via rapid cDNA amplification-based techniques, both full-length coding sequences were cloned. In brief, total
 502 RNA was isolated from acetate-grown cells using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany).
 503 Complementary DNA (cDNA) amplifications were done using the FirstChoice®RLM-RACE Kit (Thermo Fisher
 504 Scientific) and specific primers designed based on the *Polytomella* draft genome sequence. Using the 3' RACE
 505 DNAs as templates, the full-length sequences coding *Polytomella* ADHEs were amplified by PCR using *Pfu*
 506 polymerase (Promega). The primers used to amplify ADHE1 cDNA (2973 bp) and ADHE2 cDNA (2717 bp) are
 507 listed in Supplementary Table S1. Sequencing of *PsADHE2* cDNAs revealed a few differences with the nucleotide
 508 sequence published earlier (Atteia et al. 2003), which are believed to result from an inappropriate assembly of
 509 partial cDNA sequences.

510

511 The full-length cDNAs were used as templates to amplify the coding sequence for predicted mature proteins using
 512 the primers listed in Supplementary Table S1. Amplified fragments were cloned into the pET-24a expression
 513 vector (Merck Novagen) using the NdeI and XhoI restriction site within the multiple cloning site to incorporate a
 514 6xHis tag at the C-terminus of the proteins. Constructs with confirmed nucleotide sequences (at GATC Biotech)
 515 were used to transform *Escherichia coli* BL21 (DE3). Transformed cells from plates were inoculated in the auto-
 516 induction medium ZYM5052 (Studier 2005) supplemented with kanamycin (30 $\mu\text{g}/\text{ml}$), and grown overnight
 517 either at 25°C for rADHE1, or at 37°C for rADHE2. The cells were harvested by centrifugation, washed once in
 518 Buffer 2 (300 mM NaCl, 50 mM sodium phosphate, pH 8.0) and flash-frozen and stored at -20°C until use. Protein
 519 purification was carried out at 4°C as follows. Cell pellets were resuspended in Buffer 1 supplemented with 0.5
 520 mg/ml lysozyme, 0.5 mg/ml DNase and 2 mM MgCl_2 , and protease inhibitors (PMSF, amino caproic acid,
 521 benzamidine) and subjected to three cycles of freezing (liquid nitrogen) and thawing. The lysates were clarified
 522 via centrifugation (10,000 x g for 15 min, at 4°C) and the supernatant was gently mixed with HIS-Select Nickel
 523 Affinity gel (Sigma-Aldrich) equilibrated with Buffer 2. The mixture was poured into an empty column, and the
 524 gel was extensively washed with 20 mM imidazole in Buffer 2. His-tagged ADHEs were eluted with 100 mM
 525 imidazole in Buffer 2 rapidly desalted on PD10 Desalting columns (GE Healthcare) pre-equilibrated with 50 mM
 526 Tricine, pH 8.0, and finally concentrated on a Vivaspinn 50 (Sartorius) to reach a final concentration of 5-7 mg
 527 protein/ml.

528

529 **Enzyme assays**

530

531 Freshly purified recombinant ADHEs were assessed for the different enzymatic activities at 25°C. Standard assay
 532 medium for NADH-dependent reactions consisted of 50 mM potassium phosphate (pH 7.0) and 0.42 mM NADH.
 533 Acetyl-CoA reductase activity was determined in the presence of 0.21 mM acetyl-CoA, acetaldehyde reductase
 534 activity with 1 mM acetaldehyde and 0.21 mM CoASH. The rates of both reductase reactions were monitored
 535 spectrophotometrically following the disappearance of NADH at 340 nm. Standard assay medium for NAD^+ -
 536 dependent reactions contained 1.5 mM NAD^+ in 50 mM Glycine/NaOH buffer pH 9.0. Ethanol dehydrogenase

537 activity was assayed in 170 mM ethanol (1% v/v); acetaldehyde dehydrogenase activity was assayed in 1 mM
538 acetaldehyde and 0.2 mM CoASH. The rates of both dehydrogenases were monitored by the formation of NADH
539 at 340 nm. All reactions were initiated by the addition of purified recombinant enzyme (25-40 µg). In all the
540 assays, one unit of enzyme activity (U) was defined as 1 µmol of NADH (or NAD) converted per min.

541

542 **Bioinformatic analysis**

543

544 The genome of *Polytomella* sp. SAG 198.80 was assembled from paired-end Illumina and PacBio sequences,
545 using a combination of the programs ABYSS (1.9.0), CERULEAN (0.1), PBJELLY (15.2.20) and BESST (2.1).
546 Details will be published elsewhere. *Ab initio* annotation was performed with AUGUSTUS, using available
547 transcriptome assemblies to generate the species models. Protein sequences were aligned using MAFFT (7.130)
548 or Clustal Omega 1.2.4. (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Enzyme domain searches were done at
549 <https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>. The iron-containing alcohol dehydrogenases signatures
550 (ADH_iron_1, PS00913; ADH_iron_2, PS00060) were obtained at Prosite (<https://prosite.expasy.org/>) (Bairoch
551 1992). Intracellular targeting predictions were made using TargetP (Plant network v2.0) (Emanuelsson et al.
552 2000), Mitofates (plant) v1 (Fukasawa et al. 2015), Predotar (Small et al. 2004) and Predalgo (Tardif et al. 2012).

553

554 **Phylogenetic dataset construction and analysis**

555

556 ADHE proteins sequences were searched among microalgae via BLASTP using the *C. reinhardtii* ADHE
557 sequence (UniprotKB entry Q1RS84_CHLRE) as a query against the proteomes available at the US National
558 Center for Biotechnology and Information (NCBI) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), at Algoblast
559 (https://giavap-genomes.ibpc.fr/cgi-bin/AlgoBLAST/algoBlast_mainpage.php), at
560 CryptoDB (<https://cryptodb.org/cryptodb/>), at the Joint Genome Institute (JGI) genome Portal
561 (<https://genome.jgi.doe.gov/portal/>), and at Phytozome from the JGI (<https://phytozome.jgi.doe.gov/>). All the
562 retrieved sequences were checked manually for the length and the presence of the key residues and domains of
563 ADHE. A total of 49 algal genome sequences were surveyed. Phylogenetic analyses were conducted using
564 MEGAX and MEGA7 software (<http://www.megasoftware.net>), and phylogenetic relationships were inferred
565 using the maximum likelihood (ML) method applied with the JTT amino-acid substitution matrix-based model
566 (Jones et al. 1992).

567

568 **Nucleotide sequence accession numbers**

569 The ADHE1 cDNA sequence can be found in the DDBJ/EMBL/GenBank databases under the accession number
570 PRJEB33479. The sequence of ADHE2 cDNA (UniProtKB entry- Q70YJ9) has been updated.

Acknowledgements

This work was supported by the Centre National de la Recherche Scientifique (CNRS), Aix-Marseille Université (AMU) and the LABEX Dynamo (ANR-11-LABX-0011-01). Proteomic experiments were partly supported by the Agence Nationale de la Recherche (ANR-10-INBS-08-01 ProFI grant).

We thank Dr. David R. Smith (Dalhousie University, Canada) for sharing genome assemblies of *P. magna*, *P. capuana*, *P. piriformis* and *P. parva*.

Reference list

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov V V, Kröger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamtracoln K, Valentin K, Vardi

- A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306:79–86 . <https://doi.org/10.1126/science.1101156>
- Atteia A, Dreyfus G, González-Halphen D (1997) Characterization of the alpha and beta-subunits of the F0F1-ATPase from the alga *Polytomella* spp., a colorless relative of *Chlamydomonas reinhardtii*. *Biochim Biophys Acta* 1320:275–284 . [https://doi.org/10.1016/s0005-2728\(97\)00031-5](https://doi.org/10.1016/s0005-2728(97)00031-5)
- Atteia A, van Lis R, Gelius-Dietrich G, Adrait A, Garin J, Joyard J, Rolland N, Martin W (2006) Pyruvate formate-lyase and a novel route of eukaryotic ATP synthesis in *Chlamydomonas* mitochondria. *J Biol Chem* 281:9909–9918 . <https://doi.org/10.1074/jbc.M507862200>
- Atteia A, van Lis R, Mendoza-Hernández G, Henze K, Martin W, Riveros-Rosas H, González-Halphen D (2003) Bifunctional aldehyde/alcohol dehydrogenase (ADHE) in chlorophyte algal mitochondria. *Plant Mol Biol* 53:175–188 . <https://doi.org/10.1023/B:PLAN.0000009274.19340.36>
- Atteia A, van Lis R, Ramírez J, González-Halphen D (2000) *Polytomella* spp. growth on ethanol. Extracellular pH affects the accumulation of mitochondrial cytochrome c550. *Eur J Biochem* 267:2850–2858 . <https://doi.org/10.1046/j.1432-1327.2000.01288.x>
- Atteia A, van Lis R, Tielens AGM, Martin WF (2013) Anaerobic energy metabolism in unicellular photosynthetic eukaryotes. *Biochim Biophys Acta - Bioenerg* 1827:210–223 . <https://doi.org/10.1016/j.bbabi.2012.08.002>
- Bairoch A (1992) Prosite: a dictionary of sites and patterns in proteins. *Nucleic Acids Res* 20:2013–2018 . <https://doi.org/10.1093/nar/20.suppl.2013>
- Blaby-Haas CE, Merchant SS (2019) Comparative and Functional Algal Genomics. *Annu Rev Plant Biol* 70:605–638 . <https://doi.org/10.1146/annurev-arplant-050718-095841>
- Boxma B, Voncken F, Jannink S, van Alen T, Akhmanova A, van Weelden SWH, van Hellemond JJ, Ricard G, Huynen M, Tielens AGM, Hackstein JHP (2004) The anaerobic chytridiomycete fungus *Piromyces* sp. E2 produces ethanol via pyruvate:formate lyase and an alcohol dehydrogenase E. *Mol Microbiol* 51:1389–1399 . <https://doi.org/10.1046/j.1365-2958.2003.03912.x>
- Bruchhaus I, Tannich E (1994) Purification and molecular characterization of the NAD(+)-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica*. *Biochem J* 303 (Pt 3):743–748 . <https://doi.org/10.1042/bj3030743>
- Burton RM, Stadtman ER (1953) The oxidation of acetaldehyde to acetyl coenzyme A. *J Biol Chem* 202:873–890
- Catalanotti C, Dubini A, Subramanian V, Yang W, Magneschi L, Mus F, Seibert M, Posewitz MC, Grossman AR (2012) Altered fermentative metabolism in *Chlamydomonas reinhardtii* mutants lacking pyruvate formate lyase and both pyruvate formate lyase and alcohol dehydrogenase. *Plant Cell* 24:692–707 . <https://doi.org/10.1105/tpc.111.093146>
- Cederbaum AI, Lieber CS, Rubin E (1973) Effect of acetaldehyde on activity of shuttles for the transport of reducing equivalents into the mitochondria. *FEBS Lett* 37:89–92 . [https://doi.org/10.1016/0014-5793\(73\)80432-6](https://doi.org/10.1016/0014-5793(73)80432-6)
- Cederbaum AI, Lieber CS, Rubin E (1974) The effect of acetaldehyde on mitochondrial function. *Arch Biochem Biophys* 161:26–39 . [https://doi.org/https://doi.org/10.1016/0003-9861\(74\)90231-8](https://doi.org/https://doi.org/10.1016/0003-9861(74)90231-8)
- Copley SD (2020) Evolution of new enzymes by gene duplication and divergence. *FEBS J* 287:1262–1283 . <https://doi.org/10.1111/febs.15299>
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005–1016 . <https://doi.org/10.1006/jmbi.2000.3903>
- Fukasawa Y, Tsuji J, Fu S-C, Tomii K, Horton P, Imai K (2015) MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol Cell Proteomics* 14:1113–1126 . <https://doi.org/10.1074/mcp.M114.043083>
- Funes S, Davidson E, Gonzalo Claros M, van Lis R, Pérez-Martínez X, Vázquez-Acevedo M, King MP, González-Halphen D (2002) The typically mitochondrial DNA-encoded ATP6 subunit of the F1F0-

- ATPase is encoded by a nuclear gene in *Chlamydomonas reinhardtii*. *J Biol Chem* 277:6051–6058 .
<https://doi.org/10.1074/jbc.M109993200>
- Gould SB, Garg SG, Handrich M, Nelson-Sathi S, Gruenheit N, Tielens AGM, Martin WF (2019) Adaptation to life on land at high O₂ via transition from ferredoxin-to NADH-dependent redox balance. *Proceedings Biol Sci* 286:20191491 . <https://doi.org/10.1098/rspb.2019.1491>
- Guarnieri MT, Levering J, Henard CA, Boore JL, Betenbaugh MJ, Zengler K, Knoshaug EP (2018) Genome Sequence of the Oleaginous Green Alga, *Chlorella vulgaris* UTEX 395. *Front Bioeng Biotechnol* 6:37 .
<https://doi.org/10.3389/fbioe.2018.00037>
- Hemschemeier A, Jacobs J, Happe T (2008) Biochemical and Physiological Characterization of the Pyruvate Formate-Lyase Pfl1 of *Chlamydomonas reinhardtii* , a Typically Bacterial Enzyme in a Eukaryotic Alga . *PLoS One* 3:e2526 . <https://doi.org/10.1186/1546-2218-3-2526>
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8:275–282 . <https://doi.org/10.1093/bioinformatics/8.3.275>
- Kawata T, Masuda K, Yoshino K (1975) Presence of fine spirals (spirosomes) in *Lactobacillus fermenti* and *Lactobacillus casei*. *Jpn J Microbiol* 19:225–227 . <https://doi.org/10.1111/j.1348-0421.1975.tb00872.x>
- Kessler D, Herth W, Knappe J (1992) Ultrastructure and pyruvate formate-lyase radical quenching property of the multienzymic AdhE protein of *Escherichia coli*. *J Biol Chem* 267:18073–18079
- Kessler D, Leibrecht I, Knappe J (1991) Pyruvate-formate-lyase-deactivase and acetyl-CoA reductase activities of *Escherichia coli* reside on a polymeric protein particle encoded by adhE. *FEBS Lett* 281:59–63 .
[https://doi.org/10.1016/0014-5793\(91\)80358-a](https://doi.org/10.1016/0014-5793(91)80358-a)
- Kim G, Azmi L, Jang S, Jung T, Hebert H, Roe AJ, Byron O, Song J-J (2019) Aldehyde-alcohol dehydrogenase forms a high-order spiroosome architecture critical for its activity. *Nat Commun* 10:4527 .
<https://doi.org/10.1038/s41467-019-12427-8>
- Kim KM, Park J-H, Bhattacharya D, Yoon HS (2014) Applications of next-generation sequencing to unravelling the evolutionary history of algae. *Int J Syst Evol Microbiol* 64:333–345 .
<https://doi.org/10.1099/ijs.0.054221-0>
- Kurylo CM, Parks MM, Juette MF, Zinshteyn B, Altman RB, Thibado JK, Vincent CT, Blanchard SC (2018) Endogenous rRNA Sequence Variation Can Regulate Stress Response Gene Expression and Phenotype. *Cell Rep* 25:236-248.e6 . <https://doi.org/10.1016/j.celrep.2018.08.093>
- Lees GJ, Jago GR (1978) Role of Acetaldehyde in Metabolism: A Review 1. Enzymes Catalyzing Reactions Involving Acetaldehyde. *J Dairy Sci* 61:1205–1215 . [https://doi.org/https://doi.org/10.3168/jds.S0022-0302\(78\)83708-4](https://doi.org/https://doi.org/10.3168/jds.S0022-0302(78)83708-4)
- Magneschi L, Catalanotti C, Subramanian V, Dubini A, Yang W, Mus F, Posewitz MC, Seibert M, Perata P, Grossman AR (2012) A mutant in the ADH1 gene of *Chlamydomonas reinhardtii* elicits metabolic restructuring during anaerobiosis. *Plant Physiol* 158:1293–1305 . <https://doi.org/10.1104/pp.111.191569>
- Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L, Marshall WF, Qu LH, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov V V., Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, Grigoriev I V., Rokhsar DS, Grossman AR, Cardol P, Cerutti H, Chanfreau G, Chen CL, Cognat V, Croft MT, Dent R, Dutcher S, Fernández E, Fukuzawa H, González-Ballester D, González-Halphen D, Hallmann A, Hanikenne M, Hippler M, Inwood W, Jabbari K, Kalanon M, Kuras R, Lefebvre PA, Lemaire SD, Lobanov A V., Lohr M, Manuell A, Meier I, Mets L, Mittag M, Mittelmeier T, Moroney J V., Moseley J, Napoli C, Nedelcu AM, Niyogi K, Novoselov S V., Paulsen IT, Pazour G, Purton S, Ral JP, Riaño-Pachón DM, Riekhof W, Rymarquis L, Schroda M, Stern D, Umen J, Willows R, Wilson N, Zimmer SL, Allmer J, Balk J, Bisova K, Chen CJ, Elias M, Gendler K, Hauser C, Lamb MR, Ledford H, Long JC, Minagawa J, Page MD, Pan J, Pootakham W, Roje S, Rose A, Stahlberg E, Terauchi AM, Yang P, Ball S, Bowler C, Dieckmann CL, Gladyshev VN, Green P, Jorgensen R, Mayfield S, Mueller-Roeber B, Rajamani S, Sayre RT, Brokstein P, Dubchak I, Goodstein D, Hornick L, Huang YW, Jhaveri J, Luo Y, Martínez D, Ngau WCA, Otiillar B, Poliakov A, Porter A, Szajkowski L, Werner G, Zhou K (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* (80-) 318:245–251 .
<https://doi.org/10.1126/science.1143609>

- Montella C, Bellolell L, Pérez-Luque R, Badía J, Baldoma L, Coll M, Aguilar J (2005) Crystal structure of an iron-dependent group III dehydrogenase that interconverts L-lactaldehyde and L-1,2-propanediol in *Escherichia coli*. *J Bacteriol* 187:4957–4966 . <https://doi.org/10.1128/JB.187.14.4957-4966.2005>
- Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu R-Y, van der Giezen M, Tielens AGM, Martin WF (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev* 76:444–495 . <https://doi.org/10.1128/MMBR.05024-11>
- Mus F, Dubini A, Seibert M, Posewitz MC, Grossman AR (2007) Anaerobic acclimation in *Chlamydomonas reinhardtii*: anoxic gene expression, hydrogenase induction, and metabolic pathways. *J Biol Chem* 282:25475–25486 . <https://doi.org/10.1074/jbc.M701415200>
- Nelson DR, Chaiboonchoe A, Fu W, Hazzouri KM, Huang Z, Jaiswal A, Daakour S, Mystikou A, Arnoux M, Sultana M, Salehi-Ashtiani K (2019) Potential for Heightened Sulfur-Metabolic Capacity in Coastal Subtropical Microalgae. *iScience* 11:450–465 . <https://doi.org/10.1016/j.isci.2018.12.035>
- Pietrocola F, Galluzzi L, Bravo-San Pedro JM, Madeo F, Kroemer G (2015) Acetyl coenzyme A: a central metabolite and second messenger. *Cell Metab* 21:805–821 . <https://doi.org/10.1016/j.cmet.2015.05.014>
- Piganeau G, Grimsley N, Moreau H (2011) Genome diversity in the smallest marine photosynthetic eukaryotes. *Res Microbiol* 162:570–577 . <https://doi.org/10.1016/j.resmic.2011.04.005>
- Pony P, Rapisarda C, Terradot L, Marza E, Fronzes R (2020) Filamentation of the bacterial bi-functional alcohol/aldehyde dehydrogenase AdhE is essential for substrate channeling and enzymatic regulation. *Nat Commun* 11:1426 . <https://doi.org/10.1038/s41467-020-15214-y>
- Pringsheim EG (1955) The Genus *Polytomella**. *J Protozool* 2:137–145 . <https://doi.org/10.1111/j.1550-7408.1955.tb02413.x>
- Radakovits R, Jinkerson RE, Fuerstenberg SI, Tae H, Settlege RE, Boore JL, Posewitz MC (2012) Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat Commun* 3:686 . <https://doi.org/10.1038/ncomms1688>
- Roof DM, Roth JR (1992) Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in *Salmonella typhimurium*. *J Bacteriol* 174:6634–6643 . <https://doi.org/10.1128/jb.174.20.6634-6643.1992>
- Rosenthal B, Mai Z, Caplivski D, Ghosh S, de la Vega H, Graf T, Samuelson J (1997) Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *J Bacteriol* 179:3736–3745 . <https://doi.org/10.1128/jb.179.11.3736-3745.1997>
- Salvetti A, Couté Y, Epstein A, Arata L, Kraut A, Navratil V, Bouvet P, Greco A (2016) Nuclear Functions of Nucleolin through Global Proteomics and Interactomic Approaches. *J Proteome Res* 15:1659–1669 . <https://doi.org/10.1021/acs.jproteome.6b00126>
- Sánchez LB (1998) Aldehyde dehydrogenase (CoA-acetylating) and the mechanism of ethanol formation in the amitochondriate protist, *Giardia lamblia*. *Arch Biochem Biophys* 354:57–64 . <https://doi.org/10.1006/abbi.1998.0664>
- Schägger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem* 217:220–230 . <https://doi.org/10.1006/abio.1994.1112>
- Schägger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* 199:223–231 . [https://doi.org/10.1016/0003-2697\(91\)90094-a](https://doi.org/10.1016/0003-2697(91)90094-a)
- Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* 473:337–342 . <https://doi.org/10.1038/nature10098>
- Shasmal M, Dey S, Shaikh TR, Bhakta S, Sengupta J (2016) *E. coli* metabolic protein aldehyde-alcohol dehydrogenase-E binds to the ribosome: a unique moonlighting action revealed. *Sci Rep* 6:19936 . <https://doi.org/10.1038/srep19936>
- Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4:1581–1590 . <https://doi.org/10.1002/pmic.200300776>

- Smith DR, Hua J, Archibald JM, Lee RW (2013) Palindromic genes in the linear mitochondrial genome of the nonphotosynthetic green alga *Polytomella magna*. *Genome Biol Evol* 5:1661–1667 . <https://doi.org/10.1093/gbe/evt122>
- Smith DR, Lee RW (2014) A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiol* 164:1812–1819 . <https://doi.org/10.1104/pp.113.233718>
- Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41:207–234 . <https://doi.org/10.1016/j.pep.2005.01.016>
- Tardif M, Atteia A, Specht M, Cogne G, Rolland N, Brugière S, Hippler M, Ferro M, Bruley C, Peltier G, Vallon O, Cournac L (2012) PredAlgo: A New Subcellular Localization Prediction Tool Dedicated to Green Algae. *Mol Biol Evol* 29:3625–3639 . <https://doi.org/10.1093/molbev/mss178>
- Toth J, Ismaiel AA, Chen JS (1999) The ald gene, encoding a coenzyme A-acylating aldehyde dehydrogenase, distinguishes *Clostridium beijerinckii* and two other solvent-producing clostridia from *Clostridium acetobutylicum*. *Appl Environ Microbiol* 65:4973–4980
- van Lis R, Baffert C, Couté Y, Nitschke W, Atteia A (2013) *Chlamydomonas reinhardtii* chloroplasts contain a homodimeric pyruvate: Ferredoxin oxidoreductase that functions with FDX1. *Plant Physiol* 161:57–71 . <https://doi.org/10.1104/pp.112.208181>
- van Lis R, Brugière S, Baffert C, Couté Y, Nitschke W, Atteia A (2020) Hybrid cluster proteins in a photosynthetic microalga. *FEBS J* 287:721–735 . <https://doi.org/10.1111/febs.15025>
- van Lis R, González-Halphen D, Atteia A (2005) Divergence of the mitochondrial electron transport chains from the green alga *Chlamydomonas reinhardtii* and its colorless close relative *Polytomella* sp. *Biochim Biophys Acta* 1708:23–34 . <https://doi.org/10.1016/j.bbabi.2004.12.010>
- van Lis R, Mendoza-Hernández G, Groth G, Atteia A (2007) New insights into the unique structure of the F0F1-ATP synthase from the chlamydomonad algae *Polytomella* sp. and *Chlamydomonas reinhardtii*. *Plant Physiol* 144:1190–1199 . <https://doi.org/10.1104/pp.106.094060>
- van Lis R, Popek M, Couté Y, Kosta A, Drapier D, Nitschke W, Atteia A (2017) Concerted up-regulation of aldehyde/alcohol dehydrogenase (ADHE) and starch in *Chlamydomonas reinhardtii* increases survival under dark anoxia. *J Biol Chem* 292:2395–2410 . <https://doi.org/10.1074/jbc.M116.766048>
- Wang D, Ning K, Li J, Hu J, Han D, Wang H, Zeng X, Jing X, Zhou Q, Su X, Chang X, Wang A, Wang W, Jia J, Wei L, Xin Y, Qiao Y, Huang R, Chen J, Han B, Yoon K, Hill RT, Zohar Y, Chen F, Hu Q, Xu J (2014) Nannochloropsis genomes reveal evolution of microalgal oleaginous traits. *PLoS Genet* 10:e1004094 . <https://doi.org/10.1371/journal.pgen.1004094>
- Wessel D, Flüggé UI (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138:141–143 . [https://doi.org/10.1016/0003-2697\(84\)90782-6](https://doi.org/10.1016/0003-2697(84)90782-6)
- Wise DL (1955) Carbon Sources for *Polytomella caeca*. *J Protozool* 2:156–158 . <https://doi.org/10.1111/j.1550-7408.1955.tb02416.x>
- Wise DL (1959) Carbon Nutrition and Metabolism of *Polytomella caeca*. *J Protozool* 6:19–23 . <https://doi.org/10.1111/j.1550-7408.1959.tb03921.x>
- Wise DL (1968) Effects of Acetaldehyde on Growth and Biosynthesis in an Algal Flagellate *Polytomella caeca*. *J Protozool* 15:528–531 . <https://doi.org/10.1111/j.1550-7408.1968.tb02169.x>
- Wise DL (1970) Effect of Acetaldehyde on Growth in Succinate Media and Labeling RNA with ¹⁴C Succinate in *Polytomella caeca*. *J Protozool* 17:1970
- Wu T, Li L, Jiang X, Yang Y, Song Y, Chen L, Xu X, Shen Y, Gu Y (2019) Sequencing and comparative analysis of three *Chlorella* genomes provide insights into strain-specific adaptation to wastewater. *Sci Rep* 9:9514 . <https://doi.org/10.1038/s41598-019-45511-6>
- Zimorski V, Martin WF (2014) Subcellular targeting of proteins and pathways during evolution. *New Phytol* 201:1–2 . <https://doi.org/10.1111/nph.12566>

Table 1. Targeting predictions for *Polytomella* and *C. reinhardtii* ADHEs.

Predictions were made using four on-line localization predictors.

	MitoFate		TargetP			Predotar v1.04			PredAlgo		
	M probability	Cleavage site	M score	C score	Loc - RC	M score	C score	prediction	M score	C score	Target
PsADHE1	0.047	nd	0.047	0.882	C - 2	0.07	0.34	possibly P	0.00861	0.890	C
PsADHE2	0.991	25-26	0.784	0.134	M - 2	0.7	0.06	M	0.0911	1.824	C
CrADHE1	0.017	nd	0.184	0.597	C - 3	0.49	0.49	possiby M	0.0299	3.946	C

Table 2. Enzymatic assays monitoring different activities of *Polytomella* sp. recombinant ADHEs.

The different reactions assessed are depicted in Figure 2. Enzymatic activities were determined spectrophotometrically following the production or consumption of NADH. Results are presented as the mean +/- standard deviation of triplicate samples measured from three independent protein preparations. Note that the assays for reaction 1 were performed in absence of an iron chelator. Since acetaldehyde formed in reaction #1 reacts away to ethanol immediately, and reactions #1 and #2 both produce NAD in a 1:1 stoichiometry, the real Vmax for reaction 1 can be extrapolated, corresponding to 50% to the determined Vmax.

Reaction	Substrates	Products	Vmax ($\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$)		
			rADHE1	rADHE1_t	rADHE2
#1	acetyl-CoA	acetaldehyde, CoASH (ethanol)	1.88 \pm 0.0085	2.221 \pm 0.096	1.52 \pm 0.087
#2	acetaldehyde	ethanol	2.095 \pm 0.171	2.246 \pm 0.080	0.0034 \pm 0.002
#3	ethanol	acetaldehyde	2.887 \pm 0.257	2.552 \pm 0.075	0.0087 \pm 0.001
#4	acetaldehyde, CoASH	acetyl-CoA	0.271 \pm 0.0103	0.619 \pm 0.073	23.45 \pm 1.104