A deeply conserved protease, acylamino acid-releasing enzyme (AARE), acts in ageing in Physcomitrella and Arabidopsis

Sebastian N.W. Hoernstein[1], Buğra Özdemir[1,5], Nico van Gessel[1], Alessandra A. Miniera[1], Bruno Rogalla von Bieberstein[1,6], Lars Nilges[1], Joana Schweikert Farinha[1,7], Ramona Komoll[1,8], Stella Glauz[1], Tim Weckerle[1,9], Friedrich Scherzinger[1,10], Marta Rodriguez-Franco[2], Stefanie J. Müller-Schüssele[3], Ralf Reski[1,4,*]

[1] Plant Biotechnology, Faculty of Biology, University of Freiburg, Schaenzlestrasse 1, 79104 Freiburg, Germany

[2] Cell Biology, Faculty of Biology, University of Freiburg, Schaenzlestrasse 1, 79104 Freiburg, Germany

[3] Molecular Botany, Department of Biology, Technical University of Kaiserslautern, Erwin-Schrödinger-Strasse 70, 67663 Kaiserslautern, Germany

[4] Signalling Research Centres BIOSS and CIBSS, Schaenzlestrasse 18, 79104 Freiburg, Germany

[5] Euro-BioImaging Bio-Hub, EMBL Heidelberg, Meyerhofstraße 1, 69117 Heidelberg, Germany

[6] Department of Anesthesiology, University Hospital Würzburg, Oberduerrbacher Strasse 6, 97072 Würzburg, Germany

[7] Institute for Molecular Biosciences, Goethe University Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt, Germany

[8] Heraeus Medical GmbH, Philipp-Reis-Straße 8-13. 61273 Wehrheim, Germany

[9] Eurofins GeneScan GmbH, Engesserstraße 4, 79108 Freiburg, Germany

[10] Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, 04103 Leipzig, Germany

* Corresponding author: ralf.reski@biologie.uni-freiburg.de

ORCID-IDs S.N.W.H.: 0000-0002-2095-689X, B.Ö.: 0000-0001-9823-0581, N.v.G.: 0000-0002-0606-246X, A.A.M.: 0000-0003-1638-5438, B.R.v.B.: 0000-0002-9650-4064, L.N.: 0000-0003-

0716-7236, J.S.F.: 0000-0002-5602-3901, R.K.: 0000-0002-3142-0863, S.G.: 0000-0002-2305-4525, T.W.: 0000-0002-9584-5449, F.S.: 0000-0001-5825-6621, M.R.F.: 0000-0003-1183-2075, S.J.M.S.: 0000-0003-4061-1175, R.R.: 0000-0002-5496-6711

Key words aging, Physcomitrium, plant ageing, protein oxidation, ROS

Abstract

Reactive oxygen species (ROS) are constant by-products of aerobic life. In excess, ROS lead to cytotoxic protein aggregates, which are a hallmark of ageing in animals and linked to age-related pathologies in humans. Acylamino acid-releasing enzymes (AARE) are bifunctional serine proteases, acting on oxidized proteins. AARE are found in all domains of life, albeit under different names, such as acylpeptide hydrolase (APEH/ACPH), acylaminoacyl peptidase (AAP), or oxidized protein hydrolase (OPH). In humans, AARE malfunction is associated with age-related pathologies, while their function in plants is less clear. Here, we provide a detailed analysis of *AARE* genes in the plant lineage and an in-depth analysis of AARE localization and function in the moss Physcomitrella and the angiosperm Arabidopsis. *AARE* loss-of-function mutants have not been described for any organism so far. We generated and analysed such mutants and describe a connection between AARE function, aggregation of oxidized proteins and plant ageing, including accelerated developmental progression and reduced life span. Our findings complement similar findings in animals and humans, and suggest a unified concept of ageing may exist in different life forms.

Introduction

Reactive oxygen species (ROS) are by-products of O_2 metabolism and represent a challenge to all aerobic life. ROS play a dual role as potentially lethal oxidants and as signalling molecules[1]. Therefore, aerobic organisms possess sophisticated redox systems to scavenge and detoxify excess ROS. The major sources of ROS in plant cells are the electron transport chains in mitochondria and plastids, but also in peroxisomes and at the plasma membrane[2]. Environmental stresses such as heat, drought or intense light are factors for increasing ROS production to detrimental levels. Plants possess a repertoire of detoxifying enzymes such as catalases, superoxide dismutases, ascorbate peroxidases, glutathione peroxidase-like proteins and peroxiredoxins. Electrons for reduction are largely provided *via* non-enzymatic components such as ascorbic acid, glutathione and NADPH[3,4,5,6,7]. In addition, a range of heat-shock proteins assist in disaggregation or refolding of damaged proteins[8,9,10].

Despite conversion into non-toxic derivatives, the continuous exposure to ROS results in oxidation of DNA, lipids, and proteins[5]. On the protein level, ROS lead to irreversible cysteine oxidation, advanced glycation end-products, adducts with ROS-generated reactive aldehydes, and carbonylation of amino-acid side-chains[11,12]. If not cleared *via* proteolysis, an excess of oxidized proteins accumulates to cytotoxic protein aggregates. Plant antioxidant systems and the role of ROS as signalling molecules in abiotic stress responses are well studied[13]. Yet, factors involved in a plant cell's last line of defence, such as the proteolytic systems for the clearance of irreversibly oxidized proteins, are still underexplored.

A class of serine proteases is evolutionary deeply conserved as their activity is found in bacteria, archaea, animals and plants, and can degrade irreversibly oxidized proteins[14,15,16,17]. These proteases have different names in different organisms, e.g. acylamino acid-releasing enzyme (AARE), acylpeptide hydrolase (APEH/ACPH), acylaminoacyl peptidase (AAP), or oxidized protein hydrolase (OPH)[18] but are collectively addressed as AARE here. AARE acts in multimeric complexes[14,17] as a bifunctional protease as it cleaves N^{α}-acetylated amino acids from oligopeptides *via* an exopeptidase mode, but also cleaves oxidized proteins *via* an endopeptidase mode[16,19,20,21,22]. AARE isoforms from various organisms show different specificities towards N^{α}-acetylated amino-acids, with bacterial and archaeal enzymes preferring AcLeu and AcPhe substrates[15,23,24,25] and plant and animal isoforms preferring AcAla, AcMet or AcGly substrates[16,17,26,27]. Besides substrate specificities, their subcellular localization appears to be conserved among eukaryotes, as human (HsACPH) and Arabidopsis (AtAARE) AAREs are reported as cytosolic enzymes[20,21].

In humans, AARE malfunction is linked to different types of cancer[26,28,29] and sarcoma cell viability[30]. Moreover, AARE and proteasomal activity correlate and cooperatively prevent cytotoxic aggregate formation[31,32,33]. Several selective inhibitors have been identified[34,35] and blocking of AARE function is considered as anti-cancer treatment[28]. Despite an increasing number of studies on AARE functionality in humans, AAREs in plants are far less characterized.

AARE from *Arabidopsis thaliana* (AtAARE) and from cucumber have endo- and exopeptidase functions[17], and silencing of *AtAARE* increased the levels of oxidized proteins[21]. AtAARE activity was also detected in plastid stroma fractions, although a fusion with a fluorescent reporter did not co-localize with chloroplasts[21]. Suppression of *AtAARE via* RNAi resulted in an enhanced accumulation of oxidized proteins in roots and enhanced electrolyte leakage in leaves, but a further impact on plant physiology was not described[21].

Moreover, the complete loss of function of this protease has not yet been reported for any organism, neither bacterium, archaeon, animal, or plant. Thus, although the deep evolutionary conservation of AARE suggests its pivotal role in all major life forms, experimental evidence is far from optimal.

In a proteomics study on protein arginylation we identified an AARE homolog from the moss Physcomitrella[36]. Our further analysis revealed altogether three Physcomitrella AARE homologs (PpAARE1-3). Here, we analysed the subcellular localization of these PpAARE isoforms and of their homolog from the angiosperm Arabidopsis (AtAARE). We show that an alternative splicing event is targeting PpAARE1 to chloroplasts, mitochondria and the cytosol. We provide evidence that an alternative translation initiation is sufficient to localize AtAARE to the same three subcellular compartments. Bioinformatic analyses of several genomes suggest that the localization of AARE in chloroplasts and mitochondria is conserved across the plant lineage. Employing combinatorial gene knockouts and protein co-immunoprecipitation we found distinct interactions between these three isoforms and their concerted action on progressive ageing in Physcomitrella. Likewise, an Arabidopsis *AARE* loss-of-function mutant exhibits enhanced levels of oxidized proteins and accelerated bolting, as a hallmark of plant ageing.

Results

AARE gene family expansion and splice variants

Previously, PpAARE1 (Pp1s619_3V6.1) was identified as the most prominent target for N-terminal arginylation in Physcomitrella[36,37]. N-terminal arginylation mediates poly-ubiquitination *via* the N-degron pathway, thus triggering subsequent proteasomal degradation[38]. Simultaneously, two homologs (PpAARE2: Pp1s108_134V6.1 and PpAARE3: Pp1s97_68V6.1) were identified, although those were not proven arginylation targets[36]. Meanwhile, a new Physcomitrella genome version with chromosome assembly and updated gene versions was released[39]. Consequently, the gene accessions used here are *PpAARE1* (Pp3c2_30720V3.1), *PpAARE2* (Pp3c12_21080V3.1), and *PpAARE3* (Pp3c7_25140V3.1). According to *OrthoMCL*[40,41], all three proteins are homologs of the *Arabidopsis thaliana* acylamino acid-releasing enzyme (AtAARE: AT4G14570). According to publicly available data (https://peatmoss.online.uni-marburg.de)[42] *PpAARE1-3* are expressed in all major Physcomitrella tissues and developmental stages, although at varying levels (Fig. S1a). Except for leaves (phylloids) and spores, *PpAARE1* is the most strongly expressed gene of this family (between 4 and 20 times). In contrast, *PpAARE2* is expressed considerably stronger than *PpAARE1* and *PpAARE3* in spores (Fig. S1a). Likewise, *AtAARE* is expressed in all major Arabidopsis tissues (Fig. S1b, data utilized from Mergner et al. (2020)[43] and downloaded from http://athena.proteomics.wzw.tum.de/). These data indicate a strong, positive correlation between transcript level and protein abundance across all tissues (Fig. S1b). Stress conditions decrease *AARE* expression in Arabidopsis shoots and in Physcomitrella protonemata (Fig. S1c-e).

To investigate whether other plants also possess multiple AARE homologs and to infer their phylogenetic relation, we conducted *BLASTP*[44] searches against protein models from selected species using the protein sequence of AtAARE as a query. We selected the alga *Chlamydomonas reinhardtii*[45], the liverwort *Marchantia polymorpha*[46], the peat moss *Sphagnum fallax* (*Sphagnum fallax* v1.1, DOE-JGI, http://phytozome.jgi.doe.gov/), the lycophyte *Selaginella moellendorffii*[47], the monocot *Oryza sativa*[48] and the dicot *Populus trichocarpa*[49], all available at the Phytozome12 database (https://phytozome.jgi.doe.gov). Additionally, we performed a NCBI *BLASTP* search against the charophyte *Klebsormidium nitens* proteome[50], and identified a single homolog (GAQ80280.1) in this species. We also included proteins of *Funaria hygrometrica*[51], a close relative to Physcomitrella from the Funariacea family[52], in our search. Finally, the AtAARE sequence was again searched against the Physcomitrella[39] and Arabidopsis[53] proteomes.

Homology of the resulting *BLAST* hits was confirmed if the reciprocal best *BLAST* hit against *A. thaliana* was again AtAARE.

In *P. trichocarpa*, we identified a single homolog for which three distinct splice variants are annotated (Potri.008G160400.1, Potri.008G160400.2, Potri.008G160400.3). These encode different protein isoforms, but two variants seem to encode non-catalytic proteins. AARE enzymes are prolyl-oligopeptidases with a conserved catalytic triad (Ser/Asp/His) in the C-terminal peptidase domain[54,55]. In splice variant 2 (Fig. S2, Potri.008G160400.2) alternative splicing results in the deletion of the catalytic Asp whereas the whole catalytic triad is lacking in splice variant 3 (Potri.008G160400.3). Hence, we consider these splice variants as non-active and disregard them from further discussion.

In rice, we identified two homologs (LOC_Os10g28020.3, LOC_Os10g28030.1), with an additional splice variant (LOC_Os10g28020.1) at one locus which encodes an N-terminal extension.

In *C. reinhardtii, M. polymorpha*, *S. fallax* and *S. moellendorffii*, we identified a single ortholog each. In *M. polymorpha*, three distinct splice variants are annotated (Mapoly0111s0036.1, Mapoly0111s0036.2, Mapoly0111s0036.3). The latter two are UTR (untranslated region) splice variants, thus resulting in the same protein sequence, whereas Mapoly0111s0036.1 encodes an N-terminal extension of 97 aa compared to the other two variants. In *F. hygrometrica* we identified three distinct isoforms.

Finally, our *BLASTP* searches using the latest Physcomitrella protein models[39] confirmed three homologs of AtAARE (Pp3c2_30720V3.1, Pp3c12_21080V3.1, Pp3c7_25140V3.1). Additionally, this search revealed another hit, Pp3c1_2590V3.1. This gene is composed of a single exon and encodes a 131 aa protein which harbours the AARE N-terminal domain (PF19283). However, it lacks a catalytic peptidase domain and is hardly expressed across different culture conditions and tissues[56]. We also did not find any proteomics evidence across several Physcomitrella analyses[36,57,58,59] for this protein. Therefore, we excluded this gene from further analysis.

We then used phylogenetic reconstruction to investigate the origin of gene duplications within the gene family. As an outgroup, we included the well-characterized rat[16] and human[26] AARE and two isoforms of the Antarctic icefish *Chionodraco hamatus*[27]. Physcomitrella and *F. hygrometrica* share three distinct pairs of orthologs hinting at an expansion in the common ancestor of the two species. Our phylogenetic analysis did not resolve AARE subfamilies across kingdoms (Fig. 1a)

174 and we conclude that the gene family expansions observed in rice and in the Funariaceae are
175 lineage-specific events.

176 In addition, this analysis reveals a closer relationship between *PpAARE1* and *PpAARE3*, which
177 presumably originate from a more recent gene duplication event, compared to *PpAARE2*. This is
178 supported by the fact that the open reading frames (ORFs) of *PpAARE1* and *PpAARE3* are
179 represented by a single exon whereas the ORF of *PpAARE2* is split across 17 exons, similar to
180 *AtAARE* (Fig. 1b). This is in congruence with a more recent emergence of intron-poor genes in
181 intron-rich families linked to stress response and developmental processes[60] and in line with
182 intron-less orphan Physcomitrella genes as earliest responders to abiotic stress[61].

183 For the three *PpAARE* genes, several splice variants are annotated, but only two splice variants
184 of *PpAARE1* give rise to distinct protein isoforms (Fig. 1b; PpAARE1_1, PpAARE1_2). Both splice
185 variants are present in Physcomitrella protonema (Fig. 1b, c), according to RT-PCR with splice
186 variant-specific primers (Supplementary Table S1). Likewise, for *AtAARE* two different ORF
187 definitions exist. With Araport11[62], a new version of the gene model was introduced exhibiting a
188 longer ORF at the 5' end (Fig. 1b). We detected the full-length transcript *via* reverse transcription
189 polymerase chain reaction (RT-PCR, Fig. 1c).

190 For *PpAARE1*, alternative splicing in the 5' end results in an N-terminal truncated variant whereas
191 the longer non-spliced variant encodes an N-terminal plastid transit peptide (cTP) according to
192 *TargetP2.0*[63]. A cleavage of the transit peptide at the predicted cleavage site (Ala^{72}-M^{73},
193 Supplementary Table S2) of PpAARE1 would release exactly the protein encoded by the short
194 splice variant. In contrast, PpAARE2 and PpAARE3 do not harbour any predicted N-terminal
195 targeting signals. Moreover, PpAARE3 is also lacking the WD40 domain that is present in
196 PpAARE1 and PpAARE2 (Fig. 1d).

197 The extension of the originally annotated ORF of *AtAARE* also encodes a plastid transit peptide
198 (Fig. 1d). To our knowledge, the longer variant of AtAARE has not yet been investigated, whereas
199 the short variant of AtAARE localizes to the nucleus and the cytosol[21]. In agreement with the latter
200 findings, we could predict a nuclear localization sequence (NLS, KKKK) with *LOCALIZER*[64]. Thus,
201 targeting of AtAARE to the cytosol and the nucleus, but also to plastids could be enabled by
202 alternative translation initiation. Likewise, PtAARE harbours a plastid transit peptide and a
203 potential alternative translation initiation site downstream of the predicted cTP cleavage site.

204 Accordingly, we checked for NLS in PpAARE isoforms and found one (KRRP, Supplementary
205 Table S2) in PpAARE1 and PpAARE3, whereas PpAARE2 has none, further supporting our

hypothesis that PpAARE1 and PpAARE3 originate from a relatively recent gene duplication event. Accordingly, alternative splicing also generates two distinct transcripts for AARE1 in rice (OsAARE1, Supplementary Table S2), where one variant encodes a potential plastid transit peptide.

For all other plant species, no plastid targeting sequence was predicted, while the *C. reinhardtii* AARE harbours a mitochondrial targeting sequence (Supplementary Table S2).

PpAARE1 and AtAARE in mitochondria, chloroplasts and cytoplasm

Organellar targeting of AARE has not yet been reported, although AARE activity was observed in plastid-enriched fractions of cucumber[17]. However, chloroplasts, peroxisomes and mitochondria are major hubs of ROS generation[65], and thus are likely organelles with elevated levels of oxidized proteins. Thus, we investigated whether PpAARE1 and AtAARE would localize to chloroplasts *in vivo*.

We generated fusion constructs of the PpAARE isoforms and of AtAARE with eGFP for transient expression in Physcomitrella protoplasts. Due to the presence of a predicted plastid targeting peptide for PpAARE1, eGFP was fused in frame to the 3' end of all coding sequences (CDS). Since also peroxisomes are ROS-producing organelles, we used *PlantPredPTS1*[66,67] to check for the presence of C-terminal positioned peroxisomal targeting signals. None of the selected AARE isoforms were predicted to localize to peroxisomes (Supplementary Table S2). Although AtAARE has a C-terminal CKL tripeptide, which is experimentally verified to mediate peroxisomal targeting[66,68], the properties of its other C-terminal amino acids most likely prevent peroxisomal targeting. A more recent prediction approach for PTS1-mediated targeting for Arabidopsis proteins[69] further supports this conclusion. Peroxisomal targeting can also be mediated *via* N-terminal nona-peptides[70,71], but these motifs are also not present within the first 100 aa in any of the selected AARE sequences. In agreement with these predictions eGFP was fused to the 3' end of the CDSs.

For PpAARE1 three different fusion constructs were assembled. Among these, a fusion of the CDS of the short splice variant (Pp3c2_30720V3.1) and eGFP was cloned, as well as a fusion of the CDS of the longer splice variant (Pp3c2_30720V3.2) and eGFP. Additionally, we cloned a fusion of eGFP and the sequence part in which both variants differ (M^1-A^{72}, Pp3c2_30720V3.2). This part harbours the plastid transit peptide predicted by *TargetP2.0*. All fusion constructs were expressed under the control of the Physcomitrella Actin5 promoter[57,72] in a pMAV4 plasmid backbone[73].

238 The PpAARE1 isoform derived from the short splicing variant (PpAARE1_1, Fig. 1b) clearly
239 localized to the cytoplasm (Fig. 2). The same localization was observed for PpAARE2 and
240 PpAARE3 (Fig. 2). Despite a predicted NLS, we did not observe a nuclear localization, either for
241 PpAARE1 or for PpAARE3.

242 The isoform encoded by the longer splice variant of *PpAARE1* (PpAARE1_2, Fig. 1b) localized to
243 chloroplasts and surprisingly also to mitochondria (Fig. 2). In contrast to the diffuse cytosolic
244 distribution of PpAARE1, specific foci were observed in chloroplasts. To investigate whether the
245 N-terminal sequence differing between the two PpAARE1 variants (M^1-A^{72}, Pp3c2_30720V3.2) is
246 sufficient to confer dual targeting, we fused this N-terminal sequence 5' to eGFP and observed
247 again dual localization (Fig. 2, PpAARE1_Nt). Full-length PpAARE1 was necessary to localize
248 eGFP to foci within chloroplasts, as the PpAARE1_Nt:eGFP fusion led to a uniform distribution.
249 However, full-length PpAARE1 was also homogeneously distributed throughout the cytoplasm.
250 This indicates the presence of interactors that recruit PpAARE1 to specific sites or complexes
251 within the chloroplasts. Further, we conclude that the N-terminal extension of PpAARE1_2
252 encodes an ambiguous targeting signal for import into chloroplasts and mitochondria as it is
253 capable of directing the fusion protein simultaneously to both organelles.

254 Simultaneous localization of proteins to chloroplasts and mitochondria can be mediated *via*
255 ambiguous targeting signals which are recognized by both translocation machineries. We
256 evaluated whether *ATP2*[75], a tool for the prediction of ambiguous targeting, would recognize
257 PpAARE1 but this was not predicted to be dually targeted. In contrast, AtAARE was predicted to
258 be dually targeted *via* an ambiguous signal. Thus, we cloned the analogous three fusion
259 constructs for AtAARE and investigated the subcellular localization of their encoded proteins
260 accordingly.

261 The AtAARE isoform translated from the shorter ORF (AtAARE, Fig. 1b) localized to the
262 cytoplasm (AtAARE_SV, Fig. 2), as observed for PpAARE1_1. This result is partially in agreement
263 with Nakai et al. (2012)[21] since we could not observe nuclear localization. Using the fusion
264 construct of the longer AtAARE variant, we observed clear dual targeting of the protein to
265 chloroplasts and mitochondria (AtAARE_LV, Fig. 2), as observed for PpAARE1_2. Here, the
266 eGFP signal was distributed homogeneously in the chloroplasts, in contrast to the foci of
267 PpAARE1:eGFP.

268 Next, we cloned only the N-terminal sequence differing between both variants (M^1-A^{55}, longer
269 ORF definition, Fig. 1b) and fused it to eGFP. In order to investigate whether the exact N-terminal

difference between the two AtAARE variants would be sufficient for targeting, the M^{56} (same as M^{1} in the shorter variant), which is the P1 aa at the predicted cleavage site, was deleted. Using this construct, the eGFP signal localized to chloroplasts and mitochondria (AtAARE_Nt, Fig. 2). The signal within chloroplasts was homogeneously distributed, similar to the longer AtAARE variant. Thus, we conclude that the N-terminal extension of both long PpAARE and AtAARE variants is sufficient for dual targeting of proteins *in vivo*. Intriguingly, the longer variant of AtAARE localized exclusively to chloroplasts and mitochondria although alternative translation initiation should be possible. This is interesting as alternative translation initiation is also possible in the longer splice variant of *PpAARE1* (PpAARE1_2). In the latter also, the fusion protein localizes exclusively to chloroplasts and mitochondria, which excludes the possibility of an alternative translation initiation, at least in protoplasts. There are numerous transcripts in mammals where translation of an upstream positioned ORF suppresses the translation of the downstream main ORF[76]. A similar scenario is conceivable in Physcomitrella. However, it remains unclear how and if translation from the internal start codons is controlled. It is also possible that factors controlling alternative translation initiation of AtAARE are absent in Physcomitrella, at least in a spatiotemporal manner, or they might only be triggered in specific physiological situations. According to our data, the translation of the two variants of PpAARE1 is mainly controlled by alternative splicing and not by alternative translation initiation.

In summary, PpAARE1 and AtAARE localize to three subcellular compartments *via* an ambiguous targeting signal. In contrast, PpAARE2 and PpAARE3 localize solely to the cytoplasm.

Double knockout of *PpAARE1/2* reduces lifespan

Null mutants of *AARE* have not been described in any organism, and the biological role of this protease apart from its catalytic function remained unclear. Hence, we created *AARE* knockouts in Physcomitrella by deleting the full CDS of each gene *via* gene targeting (Fig. 3a) according to[77]. To enable subsequent knockout (KO) of other AARE genes, different selection markers were used. Since three distinct AARE genes exist in Physcomitrella which result in cytosolic proteases, we generated all possible combinations of double KOs and triple KOs to avoid potential compensation of the loss of function.

Plants surviving the selection procedure were screened *via* PCR for the absence of the respective CDS and correct integration of the KO construct in the target locus (Fig. S4a). Additionally, the number of genomic integrations of the KO construct was measured *via* quantitative PCR (qPCR)

302 as described[78] (Fig. S4b). At least three independent lines were identified for all single, double
303 and triple KOs (Fig. S4c-i) and a line with only a single integration in the genome was detected
304 for each KO (j-p, line numbers with stars). Further, haploidy of all lines was confirmed *via* flow
305 cytometry (Fig. S5a-c) as described[79]. These precautions were made as the transformation
306 procedure may generate plants with multiple integrations[80], possibly leading to off-target effects.
307 Further, the transformation procedure may lead to diploid plants with altered gene expression[81].
308 We used haploid lines with a single integration of the KO construct (Fig. S4) for subsequent
309 experiments.

310 Typically, AARE exopeptidase activity is assayed *via* N^{α}-acetylated amino acids like AcAla or
311 AcMet coupled to a reporter such as para-nitro-anilide (pNA) or 7-amido-4-methylcoumarin
312 (AMC). From these, AcAla-pNA was tested for several eukaryotic AAREs[22,26], including
313 AtAARE[17]. Here, we analysed the impact of AARE loss of function on the activity towards AcAla-
314 pNA and AcLeu-pNA. The latter is a substrate of bacterial and archaeal AARE isoforms but also
315 eukaryotic isoforms exhibit cleavage activity on this substrate[16,22,23,24,25]. On the single KO level,
316 the exopeptidase activity on both substrates was significantly reduced in the ∆PpAARE1 mutant
317 whereas the single KO of the other isoforms did not affect the activity (Fig. 3b). This strong impact
318 of PpAARE1 on the exopeptidase activity was consistent across all transgenic mutant lines (Fig.
319 S6).

320 One important step in Physcomitrella development is the transition from protonema to
321 gametophores, a developmental progression regulated among others by plant age and the
322 nutrient status[82]. The single KOs were phenotypically inconspicuous on the gametophore level
323 (Fig. 3c). In contrast, gametophores of ∆*PpAARE1/2* and ∆*PpAARE1/3/2* were severely stunted
324 and colonies were denser compared to wild type (WT) or the other mutants (Fig. 3c). This growth
325 effect is restricted to gametophores since protonema growth on solid medium did not differ
326 between WT and the KOs (Fig. 3d, e). Intriguingly, ∆*PpAARE1/2* and ∆*PpAARE1/3/2* mutants
327 showed accelerated developmental transition, as they developed gametophores from protonema
328 earlier (Fig. 3d). Since other KO lines did not show this effect, we attributed this to the double KO
329 of *PpAARE1/2* and performed a quantitative comparison with WT. Here, gametophores were
330 already observed after 6 days in the double KO of *PpAARE1/2*, while in WT a similar number of
331 gametophores per colony was observed only after 13 days (Fig. 3f). Consequently, the double
332 KO of *PpAARE1/2* causes accelerated developmental progression but gametophores remained
333 ultimately smaller than in WT (Fig. 3c). These effects are not linked to AARE exopeptidase activity

since the exopeptidase activity in gametophores was significantly reduced in all lines with a KO of PpAARE1 (Fig. 3g), which mimics the activity profile in protonema (Fig. S6).

To analyse AARE endopeptidase activity, we assessed the total levels of oxidized proteins in gametophores. In this assay, protein carbonyl groups derived from oxidation are derivatized with 2,4-dinitrophenylhydrazine (DNPH) to 2,4-dinitrophenylhydrazone (DNP). This irreversible modification is then recognized on Western blots by a primary anti-DNP antibody. Since DNPH can also react with oxidation states of cysteine side chains[83], this assay detects not only protein carbonylation but general protein oxidation.

With this assay we found that PpAARE2 had the strongest impact on the level of oxidized proteins in gametophores (Fig. 3h) and thus is not linked to exopeptidase activity. This was consistently observed in three independent analyses (Fig. 3h, Fig. S7). Apparently, PpAARE3 does not have any impact on exopeptidase and endopeptidase activity in gametophores under standard conditions (Fig. 3g, h).

Taken together, PpAARE1 predominantly acts as exopeptidase, while PpAARE2 predominantly acts as endopeptidase, and only the simultaneous loss of both activities in the double knockout mutants has the severest phenotypical consequences.

We found another remarkable difference between WT and mutants with a double KO of *PpAARE1/2* in older plants. After 5 months of cultivation, Δ*PpAARE1/2* and Δ*PpAARE1/3/2* were only viable at the tip of the gametophores (Fig. 4a), whilst most of the colony was already dead. In contrast, gametophores of WT and the other KOs were fully viable. After 8 months, Δ*PpAARE1/2* and Δ*PpAARE1/3/2* were already dead, in contrast to WT and the respective parental lines, which only showed some dead gametophores (Fig. 4b, c).

In summary, mutants with a double KO of *PpAARE1/2* exhibit accelerated developmental transition from protonema to gametophore (Fig. 3f), while size and life span of gametophores is strikingly reduced (Fig. 3c, Fig. 4a-c). In contrast, these effects are not visible in Δ*PpAARE1/3*. Therefore, these ageing phenotypes are linked to the concurrent loss of major AARE endopeptidase and exopeptidase activity.

Distinct *in vivo* interactions of PpAARE isoforms

In different organisms, AARE forms different homomeric complexes such as dimers[14], tetramers[17], or hexamers[24]. Thus, we analysed whether the PpAARE isoforms can interact with each other. Previously, all three isoforms were identified, although the protein modification used

for pulldown (N-terminal arginylation) was only identified on PpAARE1[36]. This gave rise to two hypotheses: First, PpAARE2 and PpAARE3 are also targets for N-terminal arginylation, but modified peptides were not identified for these isoforms. Second, the isoforms interact in complexes which were pulled down due to the N-terminal arginylation of PpAARE1. We generated Citrine fusion lines for each isoform *via* in-frame tagging at the native locus (knock-in, Fig. 5a). The original stop codons of the respective PpAARE CDS were deleted. The Citrine-tag was intended for two different analyses: First, it should enable *in vivo* monitoring of PpAARE isoforms expressed from the native promoter, and second, it is a valid tag for co-immunoprecipitation (Co-IP) *via* commercially available trap-beads.

In plants with a detectable fusion transcript (Fig. S8a-c) the presence of the target protein was checked *via* IP and subsequent MS analysis. For *PpAARE3:Citrine* lines, we detected transcripts in one line and obtained only insufficient coverage and intensity at the MS across several Co-IPs. Thus, these Physcomitrella lines were excluded from further analysis. The co-precipitation of other PpAARE isoforms with the respective bait isoforms was observed in test IPs (Fig. S8d-e) confirming previous MS-Data[36]. All plants harbouring the Citrine fusion were phenotypically inconspicuous (Fig. S8f) and haploid (Fig. S8g). Although the fusion proteins were detected in two independent lines for each of the two isoforms (PpAARE1, PpAARE2), we could not observe any Citrine signal within Physcomitrella protonemata or gametophores, probably due to the low abundance of the PpAARE isoforms. Nevertheless, MS intensities and sequence coverage enabled quantitative Co-IPs. The MS data have been deposited in PRIDE with the accession codes PXD033854 and PXD038742[84,85].

When targeting PpAARE1:Citrine, both other isoforms appeared as significant interacting partners (Fig. 5b, $p < 0.01$, FDR = 0.01). In a reciprocal Co-IP targeting PpAARE2:Citrine only, PpAARE1 appeared as significant interacting partner. PpAARE3 was not detected in this pulldown. Although lacking a reciprocal IP targeting PpAARE3:Citrine, the data show that PpAARE1 interacts with PpAARE2 and PpAARE3, whereas PpAARE2 only interacts with PpAARE1. Consequently, there are distinct interactions of PpAARE1 with PpAARE2 and PpAARE3 *in vivo*, possibly resulting in cytosolic heteromeric AARE complexes in Physcomitrella.

AARE affects bolting time in Arabidopsis

In Physcomitrella three *AARE* genes exist, and the concerted action and interaction of the enzymes affect plant ageing. To evaluate if this is an evolutionary conserved function, we analysed the situation in Arabidopsis. Here, it was known that silencing of the single *AARE* gene

397 leads to an accumulation of oxidized proteins, whereas overexpression did not affect their levels[21].
398 To gain deeper insights, we screened for available Arabidopsis T-DNA mutants.

399 We identified two T-DNA insertion lines (SALK_071125 and GK-456A10) in the T-DNA Express
400 database at the SIGnAL website (http://signal.salk.edu). SALK_071125 (s68) has a T-DNA
401 insertion in an intron at the 5´UTR, whereas the T-DNA insertion in GK-456A10 (GK) maps to an
402 intron in the region encoding the catalytic domain (Fig. 6a). We obtained these mutants, identified
403 homozygous populations of GK-456A410 containing the T-DNA insertion by their resistance to
404 sulfadiazine, and confirmed their genotype by PCR. In the case of the Salk line (s68), homozygous
405 plants had lost their resistance to kanamycin, but we confirmed their genotype by PCR.

406 Homozygous plants of both genotypes were back-crossed with WT Col-0, and brought to
407 homozygosis for subsequent experiments. Primers for screening and validation of both T-DNA
408 lines are listed in Supplementary Table S3. Additionally, we analysed *AARE* gene expression *via*
409 RT-PCR. The transcript was still detectable in both lines (Fig. 6a), although very reduced in s68,
410 while the protein was not detectable *via* Western blot (Fig. 6b). Surprisingly, in WT AARE was
411 detected at around 100 kDa, although the estimated molecular weight is approximately 84 kDa
412 (90 kDa for the longer ORF variant without cleavage of the plastid targeting peptide) which
413 indicates the presence of posttranslational modifications. Phenotypically, neither seedlings nor
414 adult mutant plants showed obvious deviations from WT (Col, Fig. S9)

415 Next, we assayed AARE exopeptidase function in Arabidopsis WT and the two mutants. The
416 exopeptidase activity on AcAla-pNA was significantly reduced in both T-DNA mutants. In contrast,
417 the activity on AcLeu-pNA did not change significantly, although a slight reduction was observed
418 in the GK line (Fig. 6c). Thus, in agreement with Western blot and activity data, the remaining
419 transcripts detected *via* RT-PCR (Fig. 6a) are not translated to functional enzymes, suggesting
420 complete loss of AARE function in the mutants. Based on this characterization, we concentrated
421 on T-DNA mutant s68.

422 Subsequently, we assayed the AARE endopeptidase activity in Arabidopsis. The most striking
423 feature reported were increased levels of oxidized proteins in AARE-silenced Arabidopsis plants[21]
424 which is in line with our findings in Physcomitrella AARE mutants. To corroborate this, we
425 investigated levels of oxidized proteins in the Arabidopsis T-DNA mutant (s68) in comparison to
426 WT (Col) cultivated under short day conditions. In both genotypes, the distribution of oxidized
427 proteins was different to Physcomitrella gametophores (Fig. 6d). Using the same extraction
428 protocol, oxidized proteins in Arabidopsis were mainly of lower molecular weight, whereas in

Physcomitrella they were mainly at higher molecular weight (Fig. 3h). Despite these size differences, the level of oxidized proteins was higher in the Arabidopsis mutant than in WT (Fig. 6d), mimicking the situation in Physcomitrella. Together, we found AtAARE exopeptidase and endopeptidase activity to be reduced in our experiments.

Oxidized proteins accumulate during vegetative growth of Arabidopsis and are depleted at the transition to bolting[86]. It was not clear from this study whether the level of oxidized proteins is a signal for bolting, or if the reset of protein oxidation at the bolting transition is a side-effect of enhanced protein turnover during this process. To address this, we assessed the bolting behaviour of Arabidopsis WT and mutant plants and found that bolting in mutant plants differed significantly from WT. In WT, bolting started at 24±1 rosette leaves whereas it started at 21±2 rosette leaves in the mutants (Fig. 6e). Consequently, this accelerated developmental transition in the mutants correlates with enhanced levels of oxidized proteins.

Discussion

A universal definition of ageing is difficult due to strongly differing characteristics between and within the domains of life. In humans, the term ageing is inherently linked to degenerative diseases whereas plant ageing is considered as the progression of developmental transitions from seeds to senescence. Nevertheless, a common feature of ageing is the execution of a genetic program that controls growth, development and maturation. In turn, the progression of this genetic program depends on the metabolic state and environmental factors. Among the molecules responsible for ageing are ROS. Their involvement in ageing and diseases was first postulated by Harman[87] and was extended by considering mitochondria as central sources[88]. In humans, mitochondrial malfunction in line with increased ROS levels is a central determinant of ageing and associated pathologies[89]. In plants, mitochondrial ROS increase during ageing of seeds[90] and are major determinants of germination capacity[91]. In photosynthetic tissues chloroplasts are the major source of ROS and their levels increase during ageing[92,93,94]. Plants cannot escape from stress situations that increase ROS production to detrimental levels, and despite several layers of ROS defence, oxidized proteins constitute the major share of modified molecules under stress. Consequently, protein oxidation and subsequent aggregate deposition are hallmarks of ageing[95,96]. The degradation of artificially oxidized proteins by AARE has been demonstrated repeatedly, and silencing of *AtAARE* in turn increased the levels of oxidized proteins[19,20,21,22].

460 However, the contribution of AARE to the progress of ageing remained elusive, although several
461 studies associate AARE-function to age-related diseases[97,98,99]. Here, we identified three *AARE*
462 genes in Physcomitrella. This gene family expansion is an outstanding feature of the moss family
463 Funariaceae, as most organisms we analysed contain only a single *AARE* gene, and only a few
464 have two. We analysed these three isoforms in Physcomitrella and compared selected features
465 to the single isoform in Arabidopsis. Our data reveal specific functions in age-related
466 developmental transitions and life span determination.

467 Surprisingly, we found triple localization of AtAARE and PpAARE1 to chloroplasts, mitochondria
468 and the cytosol *in vivo*, suggesting a functional role for AARE in these cellular compartments. In
469 Physcomitrella, the triple localization is mediated *via* alternative splice variants and in Arabidopsis
470 likely *via* alternative translation initiation. Although there have been indications of AtAARE being
471 associated to chloroplasts[17,21], there was no clear evidence for organellar targeting of this
472 protease until now. It is remarkable that the triple localization of AARE is evolutionary conserved
473 between Arabidopsis and Physcomitrella, although likely executed *via* different molecular
474 mechanisms. This suggests an essential and evolutionary conserved function of AARE activity in
475 the cytoplasm, chloroplasts and mitochondria. As mosses and seed plants diverged more than
476 500 million years ago[100], this is a deep evolutionary conservation.

477 Previously, AARE exopeptidase activity was observed in cucumber chloroplasts[17], and AARE
478 peptides were found in proteomes of Arabidopsis chloroplasts[101] and mitochondria[102]. In contrast,
479 AARE was not identified in plastid or mitochondrial proteomes of Physcomitrella[57,59]. We found
480 AARE exopeptidase activity in chloroplasts and mitochondria of Physcomitrella. Chloroplasts are
481 a major source of ROS in photosynthetic tissues exposed to light, whereas mitochondria are the
482 major source of ROS in non-photosynthetic tissues or in the dark[93,94,103,104]. Until now, it remained
483 unresolved how plants deplete oxidized proteins from these organelles.

484 In yeast and mammals, the ATP-dependent LON and AAA proteases are involved in clearance
485 of misfolded and oxidized mitochondrial proteins. Intriguingly, mutants of plant homologs of LON
486 proteases did not show clear effects on the levels of oxidized proteins, but AAA-type FTSH
487 proteases may play a role[105]. Nevertheless, stressors such as heat, drought or intense light
488 compromise photosynthesis[106,107] and mitochondrial respiration[108], leading to a depletion of ATP
489 and to mitochondrial dysfunction[109]. In turn, energy supply for ATP-dependent defence systems
490 such as heat-shock proteins and AAA-type proteases is severely compromised, leaving the
491 question unanswered how oxidized proteins in chloroplasts and mitochondria can be cleared.
492 Because AARE is an ATP-independent protease, our data suggest that organellar-targeted AARE

493 may act as an ATP-independent defence to prevent or attenuate protein aggregation in the major
494 ROS-producing organelles.

495 Based on localization prediction, 70% of our selected plant species possess one AARE isoform
496 that localizes to either plastids or to mitochondria. Whether dual targeting *via* ambiguous targeting
497 also occurs in these species remains to be experimentally validated. Further, we do not exclude
498 that the remaining species also have organellar AARE isoforms, because our predictions may be
499 compromised by incomplete gene models. Such incomplete AARE gene models without a transit
500 peptide were present in earlier genome annotations of Arabidopsis and Physcomitrella[41,53], whilst
501 complete gene models with transit peptides were only introduced with later versions[39,62].

502 It is not yet clear how translation of both variants in Arabidopsis is regulated, but a recent study
503 highlights the importance of alternative translation initiation in shaping different subcellular
504 localizations or functions of proteoforms[110]. This mechanism is also present in Physcomitrella,
505 where dual targeting of FtsZ isoforms to the cytoplasm and chloroplasts is enabled *via* alternative
506 translation initiation[111]. Thus, localization of PpAARE1 to the cytoplasm is also possible from the
507 longer splice variant. Alternative translation initiation of *AtAARE* is further evidenced by
508 proteomics data (www.peptideatlas.org/builds/arabidopsis/)[112]. Here, we found evidence for N-
509 terminal acetylation of M^{56} which is the initiator methionine of the short variant (PXD012708[113]).

510 A reporter fusion of the shorter *AtAARE* ORF was observed in the cytoplasm and in the nucleus[21].
511 We did not detect a nuclear localization of any AARE isoform. Using *LOCALIZER*[64] we identified
512 an NLS in AtAARE, in PpAARE1 and PpAARE3, but not in PpAARE2. In contrast, the human
513 AARE homolog HsACPH does not have a predictable NLS, but nuclear import is mediated *via*
514 interaction with XRCC1 under stress, where it acts in DNA-damage repair[114]. Similarly, a nuclear
515 localization of AARE might also occur under stress *in planta*.

516 In Physcomitrella, AARE1 is the dominant exopeptidase, whereas AARE2 acts predominantly as
517 endopeptidase, and the operation mode of AARE3 remains unresolved. Crystal structures from
518 bacterial and archaeal AAREs revealed two possible entrances for substrates to the catalytic
519 centres[14,115] but the mode of substrate entry is not fully understood. Although the quaternary
520 arrangements of subunits differ between species[115,116], the secondary structure arrangement is
521 conserved across kingdoms and specific subunit interactions (multimerization) are likely a
522 mechanism to mediate substrate specificity and modulate activity. This could be further used as
523 a switch between endopeptidase and exopeptidase activity and additionally enable access of the

524 catalytic centre *via* structural re-arrangements. Accordingly, the interaction between the distinct
525 PpAARE isoforms may modify substrate specificity and activity.

526 Our data indicate a partial compensation between PpAARE1 and PpAARE2. Both single KO
527 mutants are phenotypically inconspicuous under normal conditions, although a significant
528 reduction in exopeptidase activity was observed in Δ*PpAARE1*. PpAARE2 had the strongest
529 effect on the accumulation of oxidized proteins in gametophores on the single isoform KO level,
530 whereas the double KO of *PpAARE1* and *PpAARE2* showed the most striking effect in all lines
531 as accelerated transition from protonema to gametophores. This phenotype is similar to mutants
532 with disturbed auxin transport[117]. However, colony growth in these mutants was reduced which is
533 different from our double KO mutant (*PpAARE1/2*). Intriguingly, this double KO results in stunted
534 gametophores and a reduced life span. This phenotype partially resembles a loss-of-function
535 mutant of a central component of autophagy (PpATG3) in Physcomitrella[118]. In Δ*PpATG3*,
536 gametophores exhibit a reduced life span and colonies are smaller than in WT. In contrast,
537 gametophore size was not affected. Further, photosynthetic capacity in Δ*PpATG3* was also
538 reduced, an effect which is apparently not caused by AARE depletion[21] and *PpAARE* genes were
539 not differentially expressed in Δ*PpATG3*[118]. We conclude that the reduced life span observed in
540 Δ*PpAARE1/2* and Δ*PpAARE1/3/2* is not due to an impaired autophagy system causing nitrogen
541 starvation. This is in line with data which opposes autophagy at the onset of senescence in
542 Arabidopsis[119].

543 In mammals, elevated levels of oxidized proteins are associated with age-related pathologies,
544 such as Alzheimer´s disease, diabetes and different types of carcinomas[120]. If proteolytic
545 clearance fails, further accumulation of oxidized proteins causes protein aggregation, which is a
546 hallmark of ageing in animals[96,121,122]. A connection between protein oxidation and ageing was
547 less well studied in plants. Plastid ROS levels increase during ageing[92], which is in line with strong
548 oxidation of plastid proteins in ageing leaves[86]. Likewise, protein oxidation marks the
549 developmental transition between vegetative growth and flowering in Arabidopsis[86].
550 Physcomitrella mutants Δ*PpAARE1/2* and Δ*PpAARE1/3/2* showed accelerated developmental
551 transition from protonema to gametophores, reduced life span and increased levels of oxidized
552 proteins as signs of accelerated ageing. This is supported by the fact that gametophore tips, which
553 is younger tissue, are viable longer than the older stems in both mutants. In the Arabidopsis *AARE*
554 T-DNA mutants we found increased levels of oxidized proteins under normal cultivation conditions
555 and an accelerated developmental transition, in this case premature bolting. These findings
556 suggest an evolutionary conserved connection between protein oxidation and ageing.

557 We provide here a detailed analysis of *AARE* genes in the plant lineage and an in-depth analysis
558 of AARE localization and function in the moss Physcomitrella and the annual angiosperm
559 Arabidopsis. *AARE* loss-of-function mutants have not been described for any organism so far.
560 We generated and analysed such mutants and describe a connection between AARE function,
561 aggregation of oxidized proteins and plant ageing, including accelerated developmental
562 progression and reduced life span. Our findings complement similar findings in humans and
563 animals where AARE malfunction is associated with protein aggregation and age-related
564 pathologies.

565 To solidify the role of AARE in ageing in different life forms, particularly in plants with contrasting
566 maximum life spans and in animals of different complexity, loss-of-function mutants should be
567 established and analysed in selected model species. Likewise, a deeper understanding of AARE-
568 function in human diseases is desirable. Together, such analyses may contribute to a unified
569 concept of ageing in different life forms.

570

Methods

Cultivation of Physcomitrella

Physcomitrella WT (new species name: *Physcomitrium patens* (Hedw.) Mitt.[123]); ecotype "Gransden 2004" and AARE KO lines were cultivated in Knop medium[124] supplemented with microelements. Knop medium (pH 5.8) containing 250 mg/L KH_2PO_4, 250 mg/L KCl, 250 mg/L $MgSO_4$ x 7 H_2O, 1,000 mg/L $Ca(NO_3)_2$ x 4 H_2O and 12.5 mg/L $FeSO_4$ x 7 H_2O supplemented with 10 mL per litre of a microelement (ME) stock solution[125,126] (309 mg/L H_3BO_3, 845 mg/L $MnSO_4$ x 1 H_2O, 431 mg/L $ZnSO_4$ x 7 H_2O, 41.5 mg/L KI, 12.1 mg/L Na_2MoO_4 x 2 H_2O, 1.25 mg/L $CoSO_4$ x 5 H_2O, 1.46 mg/L $Co(NO_3)_2$ x 6 H_2O). For cultivation on solid medium, 12 g/L Agar was added to the KnopME medium. Moss suspension cultures were disrupted weekly with an ULTRA-TURRAX (IKA) at 18,000 rpm for 90 s. If not indicated otherwise, moss was grown under standard light conditions (55 µmol photons/m^2s) at 22°C in a 16h/8h light/dark cycle.

Hydroponic Physcomitrella gametophore cultures were assembled as described[36,127]. Here, a thin layer of protonema from suspension was distributed on gauze mesh (PP, 250 m mesh, 215 m thread, Zitt Thoma GmbH, Freiburg, Germany) capped on a glass ring. The glass rings with protonema-covered mesh gauze were placed in Magenta®Vessels (Sigma-Aldrich, St. Louis, USA) and KnopME medium was added until the protonema-covered gauze mesh was moist. The medium was exchanged every 4 weeks. Gametophores were harvested after 12 weeks.

Gametophore colonies on Agar plates (KnopME) were generated by transplanting single gametophores to new plates. Plates were sealed with Parafilm®.

Generation of AARE knockout lines

Knockout constructs for each *PpAARE* gene were generated by amplifying genomic regions as homologous flanks. The PCR products were fused to a neomycin (*PpAARE1*), hygromycin (*PpAARE2*) and zeocin (*PpAARE3*) resistance cassettes, respectively, employing triple template PCR as described[128] with primer sequences listed in Supplementary Table S1 using the *Q5* polymerase (New England Biolabs, Ipswich, USA). The knockout constructs were released from their vector backbones with XhoI (*PpAARE1*), BglII (*PpAARE2*) and DraI (*PpAARE3*), respectively. Enzymes were purchased from Thermo Fisher Scientific. Digested plasmids were precipitated and sterilized prior to transfection using standard ethanol precipitation method[129]. Transfection of Physcomitrella WT or KO (for consecutive knockouts) protoplasts was conducted via PEG-mediated procedure[77,130]. The WT strain as well as the identified *PpAARE* KO lines are

602 accessible via the International Moss Stock Center (IMSC, www.moss-stock-center.org). IMSC
603 accession numbers for the mutant and WT are available in Supplementary Table S4.

604 Screening of plants surviving antibiotics selection was done by PCR. KO mutant plants surviving
605 the antibiotics selection were checked for the presence of a genomic region which should be
606 removed upon homologous recombination-based integration of the knockout construct (Fig. S4a).
607 In case of the absence of this WT signal, plants were further checked for correct 5' and 3'
608 integration of the respective knockout construct using primers listed in Supplementary Table S1.

609 Protonema growth and gametophore induction

610 Suspension cultures were started at the same day and disrupted weekly as described. Dry weight
611 was measured in triplicates and suspension density was adjusted to 440 dry weight (DW) per litre
612 (mg DW/L) as described[131]. Droplets of 15 µL were distributed on solid medium (Knop ME, 0.02%
613 glucose). Sealed plates were cultivated as described above. Three droplets each per line were
614 distributed on one plate and all lines were grown on the same plate. 3 plates (12 colonies) were
615 used per assay. Colony areas were measured with *ImageJ*. White pixels counted from binarized
616 colony images were used as quantitative values. Gametophores were counted upon visibility of
617 the first leafy structure on buds.

618 Generation of PpAARE-Citrine knock-in lines

619 Knock-in constructs to fuse the coding sequence of Citrine to the C-terminus of PpAARE isoforms
620 via an 18 bp linker[128] at the endogenous genomic locus were cloned via Gibson assembly[132]. All
621 necessary parts were amplified using primers listed in Supplementary Table S1. Additionally, XhoI
622 (*PpAARE1*), SalI (*PpAARE2*) and BamHI (*PpAARE3*) restriction sites were added to the 5' and 3'
623 ends of the respective knock-in construct. All parts were assembled into pJet1.2 vector backbone
624 (Thermo Scientific) using the Gibson Assembly®Cloning Kit from New England Biolabs (Ipswich,
625 Massachusetts, USA). Transfection of Physcomitrella WT protoplasts were conducted as
626 described[77,130] by co-transfecting a plasmid containing a neomycin phosphotransferase
627 resistance (nptII) cassette as transient selection marker (pBSNNNEV, Mueller et al., 2014)[57]. The
628 linearized plasmid and the co-transfection vector were purified and sterilized *via* ethanol
629 precipitation[129] prior to transfection.

630 The presence of Citrine was checked with primers listed in Supplementary Table S1 and resulting
631 positive lines were further checked for correct 5' and 3' integration by PCR using the Plant Phire
632 Kit with primers listed in Supplementary Table S1. All identified fusion lines are available *via* the

633 International Moss Stock Center (IMSC, www.moss-stock-center.org). IMSC accessions are
634 listed in Supplementary Table S4.

635 qPCR analysis

636 The copy number of the integrated KO constructs was determined using a qPCR-based method[78].
637 Genomic DNA was extracted from 100 mg frozen protonema using the GeneJEt Plant Genomic
638 DNA Purification Kit (Thermo Scientific, Waltham, USA). DNA concentrations were adjusted to 3
639 ng/µl for each sample and qPCR was performed with primers for the 5' and 3' flanks as well as
640 with primers for the corresponding selection cassette. Additionally, primers for the single copy
641 gene *CLF* (Pp3c22_2940V3) were used as internal control. Reference lines were WT as well as
642 in-house lines with known single integrations of the used selection cassettes. Primers are listed
643 in Supplementary Table S5. PCR reactions were done using 2x SensiFAST Mix (Bioline, London,
644 UK) and analysed in a Lightcycler 480 II (Roche, Basel, Schweiz).

645 cDNA preparation

646 RNA was extracted using the innuPREP Plant RNA Kit (Analytik Jena, Jena, Germany). The
647 extracted RNA was treated with DNAse I (Thermo Scientific) and subsequently reverse
648 transcribed into cDNA using Superscript III Reverse Transcriptase (Life Technologies, Carlsbad,
649 USA).

650 Fusion constructs for subcellular localization

651 All constructs were generated using Gibson assembly[132] and integrated into a
652 PpAct5:Linker:eGFP-MAV4 vector backbone[73]. Integration of the different coding sequences was
653 done in frame in front of the Linker:eGFP. All parts for the Gibson assembly (inserts and
654 corresponding vector backbones) were amplified either with Phusion™ polymerase (Thermo
655 Fisher Scientific) or with HiFi polymerase (PCR Biosystems Ltd) according to the manufacturer´s
656 instructions. The primers were designed to have a 25 bp overlap to the corresponding fragment
657 to be fused with. All primers and combinations are listed in Supplementary Table S6. In the case
658 of the N-terminal difference of AtAARE (M^1-A^{55} of AT4G14570.1, gene model provided by TAIR
659 (https://www.arabidopsis.org/)) the Actin5 promoter was replaced by the CaMV35S promoter[133]
660 previously validated in Physcomitrella[134].

661 Cloned plasmids were purified using the PureYield™ Plasmid Midiprep kit (Promega, Wisconsin,
662 USA) according to the manufacturer´s instructions. The plasmids were purified and sterilized *via*
663 ethanol precipitation[129].

Confocal microscopy

Confocal imaging was performed on transiently transfected live protoplasts using Leica TCS SP8 (Leica Microsystems, Wetzlar, Germany). Immediately before microscopy, MitoTracker™ Orange CMTMRos (ThermoFisher Scientific) was added to protoplast suspensions to a final concentration of 100 nM. For all imaging experiments, an HC PL APO CS2 63x/1.40 OIL objective was used with a zoom factor of 4. The pinhole was set to 35.4 µm. For excitation, a WLL laser (70%) was used. In a sequential acquisition setup, eGFP and chlorophyll were excited with the same laser beam (488 nm, 2%) and their signals were detected simultaneously, whereas MitoTracker™ was excited with a different laser beam (550 nm, 2%) and its signal was detected separately. The detection ranges were specified as 502 nm - 546 nm for eGFP, 662 nm - 732 nm for chlorophyll and 597 nm - 636 nm for MitoTracker™. The images were acquired as z-stacks with the number of optical sections varying with the protoplast size. The voxel sizes of the resulting z-stacks were 0.0903, 0.0903, 0.239 µm in the x-y-z order. For visual representation and analysis, single slices with the best signal-to-noise ratio were selected and extracted from each z-stack using FIJI software.

Cultivation of Arabidopsis

Seeds were surface-sterilized for 4 min. in 80% ethanol and subsequently for 1 min. in 100% ethanol. Seeds were placed on plates containing ½ MS supplemented with 1% (D+) sucrose and 0.8% Agar. Alternatively, seeds were directly placed on sterilized soil. Seeds were incubated at 8°C for 2-4 days in the dark for stratification before placing them in growth chambers. Plants were further cultivated under short day conditions at 22°C and 70 µmol photons/m^2s in an 8h/16h light/dark cycle. For bolting assays, plants were placed in a phyto-chamber at 22°C and 70 µmol photons/m^2s in a 16h/8h light/dark cycle (long day condition). Rosette leaves and days since sowing were counted upon appearance of the shoot in the middle of the rosette.

Screening of Arabidopsis AARE mutants

Arabidopsis thaliana lines with T-DNA insertions in the At4G14570 locus were identified from the public T-DNA Express database at the SIGnAL website (Salk Institute Genomic Analysis Laboratory). Lines GK-456A10, SALK_080653C, SALK_071125C, and SALK_205137C, were obtained from the Nottingham Arabidopsis Stock Centre. Homozygous mutant alleles were verified by PCR using the following primers: forward LB GK 5´-ATATTGACCATCATACTCATTGC-3´ and reverse GK-456 5´-CTTCAAAGAAACACCAATCAG-3´for the GK-456A10 line, and forward LB-pROK 5´-GCGTGGACCGCTTGCTGCAACT-3´ and

696 reverse Salk_53 5´-TCTTTAGCCGAATCAGTTCCAGA-3´for the SALK_080653C,
697 SALK_071125C, and SALK_205137C lines. Identified homozygous mutant plants were back-
698 crossed with Arabidopsis Col-0. The F2 generation was screened for homozygous mutant plants
699 using the above listed primer-sets to identify the mutant allele or substituting the forward primers
700 with forward WT-GK-456 5´-AAGATGCTTTGCAGTCTCTA-3´ and forward WT-Salk 5´-
701 ACTGCCTTATGATCCATTGTCTC-3´, to identify the GK and SALK lines WT alleles, respectively.
702 RT-PCR was additionally performed to check for the presence of AtAARE transcripts using Taq
703 polymerase on cDNA prepared as described above with primers At1-At4. All primer combinations
704 are listed in Supplementary Table S4.

705 AARE exopeptidase activity

706 The enzyme activity assay according to [17] was modified. Here, tissue (80-100 mg) was
707 homogenized in liquid nitrogen and dissolved in 1 mL extraction buffer (50 mM PBS, 1 mM EDTA,
708 2 mM DTT). After centrifugation at 20,000 x g for 20 min at 4°C, 300 µl supernatant was mixed
709 with 700 µL reaction buffer (50 mM HEPES-KOH, pH 7.5 containing 1 mM Ac-Ala-pNA (Bachem,
710 Bubendorf, Switzerland) or 50 mM HEPES-KOH, pH 7.5, 10% DMSO containing 1mM AcLeu-
711 pNA (Bachem) and incubated at 37°C for 30-120 min. The reaction was stopped by the addition
712 of 500 µL 30% acetic acid. Absorbance was measured at 410 nm in a photospectrometer. This
713 approach was used to generate the data of Fig. 3b and Fig. S6a-d. Later the method was modified
714 (Fig. 3g, Fig. S6e-g). Here, tissue was homogenized in liquid nitrogen and dissolved in 100 µL
715 extraction buffer (50 mM PBS, 1 mM EDTA, 2 mM DTT) per 10 mg FW. After centrifugation at
716 20,000 x g for 20 min at 4°C, 5 µL supernatant was mixed with 195 µL reaction buffer (50 mM
717 HEPES-KOH, pH 7.5 containing 1 mM Ac-Ala-pNA (Bachem, Bubendorf, Switzerland)) in a 96
718 well micro titer plate and incubated at 37°C for 30-120 min. In the case of Ac-Leu-pNA as
719 substrate, 50 mM HEPES-KOH, pH 7.5 with 10% DMSO containing 1mM AcLeu-pNA (Bachem)
720 was used as reaction buffer. Every biological sample was measured in three technical replicates.
721 Absorbance was measured at 410 nm. Activity was calculated using a molar absorbance
722 coefficient[21] of 8.8 $mM*cm^{-1}$ and represents the release of pNA [µmol] per minute normalized to
723 the total protein concentration of the sample. The protein concentration was determined using the
724 A_{280} method of a NanoDrop™ (Thermo Fisher Scientific) or with a Bradford assay[135].

725 Western Blots

726 Western Blots were performed as described[74] using the *ECL Advance detection kit* (GE
727 Healthcare). The primary antibody against *A. thaliana* AARE was kindly provided by Dr. Yasuo

728 Yamauchi[17]. The primary antibody was diluted 1:10,000 in TBST-Buffer with 2% Blocking (GE
729 Healthcare) and incubated on the membrane for 2 h. As secondary antibody anti-Guinea pig
730 antibody, coupled to a horseradish peroxidase (Agrisera, AS 10 1496) diluted 1:10,000 in 2%
731 TBST with 2% Blocking (GE Healthcare), was applied for 1 h.

732 Detection of oxidized proteins

733 Plant tissues were homogenized in liquid nitrogen and proteins were extracted in 50 mM PBS, 50
734 mM DTT, 1 mM EDTA. 3-6 µg total protein was derivatized with DNPH and subsequently detected
735 with an anti-DNP antibody according to the manufacturer's instruction of the *OxyBlot Protein*
736 *Oxidation Detectio*n Kit (S7150, Sigma-Aldrich). Equal amounts of the derivatized protein samples
737 were employed as loading control on a separate SDS-gel and stained with Coomassie or silver
738 staining.

739 Flow cytometry

740 Flow cytometry analysis was performed as described[79]. Here, protonemata were chopped with a
741 razor blade in a small petri dish (6 cm diameter) in 2 mL of DAPI-buffer containing 0.01 mg/L 4′,6-
742 Diamidin-2-phenylindol (DAPI), 1.07 g/L $MgCl_2$ x 6 H_2O, 5 g/L NaCl, 21.11g/L Tris, 0.1% Triton,
743 pH 7. The solution was filtered using 30 µm pore size filters and the fluorescence intensity was
744 measured using a Cyflow®Space flow cytometry system (Partec, Munich, Germany).

745 Computational predictions

746 Predictions for the presence of cleavable targeting peptides were performed with *TargetP2.0*[63].
747 Additional predictions of subcellular localizations were performed with *LOCALIZER*[64]. The
748 presence of peroxisomal targeting signals was predicted with *PredPlantPTS1*[66,67]. Prediction of
749 protein domains was performed using *InterProScan*[135] and protein domain annotations according
750 to PFAM[137] were used.

751 Co-Immunoprecipitation

752 Co-immunoprecipitation was performed using GFP-Trap Magnetic Particles M-270 (Chromotek,
753 Planegg-Martinsried, Germany) as recommended by the manufacturer with modifications. 300
754 mg protonema was homogenized in a 2 mL reaction tube using a tungsten and a glass bead. For
755 each line three biological replicates were realized. The extraction buffer was chosen according to
756 the manufacturer's recommendations for plant samples and contained 25 mM HEPES-KOH,
757 pH7.5, 2 mM EDTA, 100 mM NaCl, 200 nM DTT, 0.5% Triton X-100, 1% plant protease inhibitor
758 cocktail (PPI, P9599, Sigma Aldrich). Ground plant material was dissolved in a final volume of 2

759 mL ice-cold extraction buffer and incubated for 10 min in a water quench ultrasonic device.
760 Samples were centrifuged at 4°C at 20,000xg for 30 min. For each sample 25 µL magnetic particle
761 slurry was washed with 500 µL extraction buffer. The sample supernatant was transferred to the
762 cleaned beads and incubated, rotating for 1 h at 6°C. Subsequently, beads were washed with 1
763 mL extraction buffer without Triton and PPI and again with 500 µL. Beads were then dissolved in
764 500 µL wash buffer (10 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.5 mM EDTA), transferred to a new
765 reaction tube and washed again in 500 µL wash buffer. A RapiGest solution (0.2% in 50 mM Tris-
766 HCl, pH7.5; RapiGest SF Surfactant (Waters, Milford, MA, USA) was mixed 3:1 with 5 mM DTT
767 in 50 mM Tris-HCl, pH7.5. 30 µL of the resulting mixture was applied to each sample. Samples
768 were incubated at 95°C for 5 min under continuous shaking. Samples were cooled down to RT
769 and 5 µL of a trypsin (V5117, P) solution (0.025 µg/µL in 50 mM Tris-HCl, pH 7.5) were added to
770 each sample. Digestion on the beads was performed for 30 min at 32°C under continuous
771 shaking. Supernatants were transferred to new reaction tubes and the remaining beads were
772 washed twice with 50 µL 5 mM Iodoacetamide solution (in 50 mM Tris-HCl, pH7.5). The wash
773 supernatants were combined with the trypsin-containing supernatant and incubated over night at
774 30°C under continuous shaking. Acid-catalysed cleavage of the RapiGest surfactant was
775 performed as recommended by the manufacturer. Samples were purified using C18-STAGE-Tips
776 as described[138] and eluted from the Tip in 30% ACN in 0.1% FA.

777 Mass spectrometry measurement and data analysis

778 MS analysis was performed on an Orbitrap Q-Exactive Plus instrument (Thermo Fisher Scientific)
779 coupled to an UltiMate 3000 RSLCnano system (Dionex LC Packings/Thermo Fisher Scientific)
780 as described[139]. Database search and label-free quantitation was performed using *MaxQuant*
781 software V 1.6.0.16[36140]. For each Co-IP a specific database was employed containing all V3.3
782 proteins of Physcomitrella[39] as well as the sequence of the respective fusion protein. Additionally,
783 the contaminant list provided by the software was included. Decoys were generated on the fly by
784 reverting the given protein sequences. Variable modifications were formation of pyro Glu (N term
785 Q, −17.026549 Da), oxidation (M, +15.994915 Da), acetylation (N-term, +42.010565 Da) and
786 deamidation (N, +0.9 84016 Da). Carbamidomethylation (C, +57. 021464 Da) was specified as
787 fixed modification. Enzymatic specificity was set to tryptic with semi-specific free N-terminus. An
788 FDR of 0.01 was set for protein identification. LFQ values[141] were used as quantitative values.
789 Interaction analysis was performed in Perseus V 1.6.12.0[142]. Missing values were imputed from
790 a normal distribution with a down-shift of 1.8 and distribution width of 0.3. Interaction partners
791 were accepted at an FDR of 0.01 and a p-value less than 0.01.

Raw files of the test-IP measurements (Fig. S8e) were processed using Mascot Distiller V2.7.10 and searched against all Physcomitrella protein models V3.3[39] using Mascot Server V2.7.0 (Matrix Science).Processed mgf files from immunoprecipitation experiments targeting N-terminal arginylation (PXD003232[36,143]) and the test-IPs were searched again against all Physcomitrella protein models V3.3[39] using Mascot Server V2.7.0 (Matrix Science). The precursor mass tolerance was 5 ppm and the fragment mass tolerance was 0.02 Da. Variable modifications were formation of pyro Glu (N term Q, −17.026549 Da), oxidation (M, +15.994915 Da), acetylation (N-term, +42.010565 Da) and deamidation (N, +0.9 84016 Da). Carbamidomethylation (C, +57. 021464 Da) was specified as fixed modification. Enzymatic specificity was set to tryptic with semi-specific free N-terminus. Search results were loaded in Scaffold™ 5 (V5.0.1, Proteome Software) and proteins were accepted at an FDR = 1 and peptides at an FDR = 0.5. A table of identified proteins is accessible in Supplementary Table S7.

Multiple sequence alignment and phylogenetic reconstruction

Homologous protein sequences were aligned with UPP[144] (version 4.4.0) using default parameters and subsequently translated into a codon-aware CDS alignment with PAL2NAL[145] (version 1.4). Based on this multiple coding sequence alignment we reconstructed a maximum likelihood tree with RAxML[146] (version 8.2.12) using the GTRCAT model with 1000 rapid bootstrap samples. The tree was rooted at the split between animal and plant sequences and plotted in R[147] using the packages ape[148] and ggtree[149].

Statistics and reproducibility

Statistical differences in datasets were analysed with one-way Anova with subsequent post-hoc test if different lines at same conditions were compared. Two-way Anova with subsequent post-hoc test was performed to analyse differences between lines at different conditions. Sample sizes of biological replicates are specified in the figure legends. Anova and post-hoc analysis was performed in R[147]. Statistical significance was accepted at $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***).

Data availability

The authors confirm that all relevant data supporting the findings of this study are available within the paper and its supplementary files. The mass spectrometry proteomics data have been deposited at the ProteomeXchange Consortium via the PRIDE partner repository[84,85] with the dataset identifier PXD033854 and 10.6019/PXD038742. Plant lines used in this study are

available upon reasonable request from the corresponding author (R.R.) or *via* the International Moss Stock Center (IMSC, www.moss-stock-center.org). IMSC accessions are listed in Supplementary Table S4. Full Blot images for Fig. 6b and Fig. S3b are available in Fig. S10. Supplementary Table S8 contains all numeric source data used to generate the graphs and charts in this study. Plasmids generated in this study are available from the International Moss Stock Center IMSC (https://www.moss-stock-center.org) with the accession numbers P1519 (*PpAARE1* KO), P1841 (*PpAARE2* KO (pJet)), P1520 (*PpAARE3* KO (pJet)), P1655 (*PpAARE1:Citrine* KI (pJet)), P1813 (*PpAARE2:Citrine* KI (pJet)), P1814 (*PpAARE3:Citrine* KI), P1833 (*PpAARE1_1:eGFP* (pMAV4)), P1834 (*PpAARE1_2:eGFP* (pMAV4)), P1853 (*PpAARE1_Nt:eGFP* (pMAV4)), P1855 (*PpAARE3:eGFP* (pMAV4)), P1856 (*PpAARE3:eGFP* (pMAV4)), P1854 (*AtAARE_SV:eGFP* (pMAV4)), P1881 (*AtAARE_LV:eGFP* (pMAV4)) and P1862 (*AtAARE_Nt:eGFP* (pMAV4)).

Acknowledgements

We thank Christine Glockner, Agnes Novakovic and Eija Schulze for expert technical assistance and Anne Katrin Prowse for language editing. Support with the Arabidopsis assays from Dr. Philipp Schwenk is gratefully acknowledged. We thank Prof. Dr. Bettina Warscheid for the possibility to use the QExactive Plus instrument and Prof. Dr. Yasuo Yamauchi for the primary antibody against *A. thaliana* AARE. We gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2189 (CIBSS to R.R.) and by the Wissenschaftliche Gesellschaft Freiburg. Open Access funding enabled and organized by Projekt DEAL.

Author contributions

S.N.W.H. designed research, performed experiments, analysed data, and wrote the manuscript. B.Ö. and N.v.G. analysed data and helped writing the manuscript. A.A.M. analysed data. B.R.v.B., L.N., J.S.F., R.K., S.G., T.W., and F.S. performed experiments. M.R.F. and S.J.M.S. designed research and helped writing the manuscript. R.R. designed and supervised research, acquired funding, and wrote the manuscript. All authors approved the final version of the manuscript.

Competing interests

All authors declare to have no competing interests.

References

1. Mittler, R., Zandalinas, S. I., Fichman, Y. & Van Breusegem, F. Reactive oxygen species signalling in plant stress responses. *Nature Reviews Molecular Cell Biology* **23**, 663-679 (2022).
2. Tola, A. J., Jaballi, A. & Missihoun, T. D. Protein carbonylation: Emerging roles in plant redox biology and future prospects. *Plants* **10**, 1451 (2021).
3. Gratão, P. L., Polle, A., Lea, P. J. & Azevedo, R. A. Making the life of heavy metal-stressed plants a little easier. *Functional Plant Biology* **32**, 481-494 (2005).
4. Smirnoff, N. Ascorbic acid: metabolism and functions of a multi-facetted molecule. *Current Opinion in Plant Biology* **3**, 229-235 (2000).
5. Sharma, P., Jha, A. B., Dubey, R. S. & Pessarakli, M. Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *Journal of Botany* **2012**: 217037 (2012).
6. Müller-Schüssele, S. J., et al. Chloroplasts require glutathione reductase to balance reactive oxygen species and maintain efficient photosynthesis. *Plant Journal* **103**, 1140-1154 (2020).
7. Soares, C., Carvalho, M. E., Azevedo, R. A. & Fidalgo, F. Plants facing oxidative challenges—A little help from the antioxidant networks. *Environmental and Experimental Botany* **161**, 4-25 (2019)
8. Parsell, D. A., Kowal, A. S., Singer, M. A. &, Lindquist, S. Protein disaggregation mediated by heat-shock protein Hspl04. *Nature* **372**, 475-478 (1994).
9. Queitsch, C., Hong, S. W., Vierling, E. & Lindquist, S. Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis. *Plant Cell* **12**, 479-492 (2000).
10. Mogk, A., Bukau, B. &, Kampinga, H. H. Cellular handling of protein aggregates by disaggregation machines. *Molecular Cell* **69**, 214-226 (2018).
11. Møller, I. M., Jensen, P. E. & Hansson, A. Oxidative modifications to cellular components in plants. *Annual Review of Plant Biology* **58**, 459-481 (2007).
12. Ciacka, K., Tymiński, M., Gniazdowska, A. & Krasuska, U. Carbonylation of proteins—an element of plant ageing. *Planta* **252**, 12 (2020).
13. Farooq, M. A., et al. Acquiring control: The evolution of ROS-Induced oxidative stress and redox signaling pathways in plant stress responses. *Plant Physiology and Biochemistry* **141**, 353-369 (2019).
14. Bartlam, M., et al. Crystal structure of an acylpeptide hydrolase/esterase from *Aeropyrum pernix K1*. *Structure* **12**, 1481-1488 (2004).

15. Brunialti, E. A., Gatti-Lafranconi, P. & Lotti, M. Promiscuity, stability and cold adaptation of a newly isolated acylaminoacyl peptidase. *Biochimie* **93**, 1543-1554 (2011).
16. Tsunasawa, S., Narita, K. & Ogata, K. Purification and properties of acylamino acid-releasing enzyme from rat liver. *Journal of Biochemistry* **77**, 89-102 (1975).
17. Yamauchi, Y., Ejiri, Y., Toyoda, Y. & Tanaka, K. Identification and biochemical characterization of plant acylamino acid–releasing enzyme. *Journal of Biochemistry* **134**, 251-257 (2003).
18. Rawlings, N. D., et al. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research* **46**, D624-D632 (2018).
19. Fujino, T., et al. Identification of the cleavage sites of oxidized protein that are susceptible to oxidized protein hydrolase (OPH) in the primary and tertiary structures of the protein. *Journal of Biochemistry* **127**, 1087-1093 (2000).
20. Shimizu, K., et al. Overexpression of oxidized protein hydrolase protects COS-7 cells from oxidative stress-induced inhibition of cell growth and survival. *Biochemical and Biophysical Research Communications* **304**, 766-771 (2003).
21. Nakai, A., Yamauchi, Y., Sumi, S. & Tanaka, K. Role of acylamino acid-releasing enzyme/oxidized protein hydrolase in sustaining homeostasis of the cytoplasmic antioxidative system. *Planta* **236**, 427–436 (2012).
22. Gogliettino, M., et al. A novel class of bifunctional acylpeptide hydrolases–potential role in the antioxidant defense systems of the Antarctic fish *Trematomus bernacchii*. *FEBS Journal* **281**, 401-415 (2014).
23. Kiss, A. L., et al. The acylaminoacyl peptidase from *Aeropyrum pernix* K1 thought to be an exopeptidase displays endopeptidase activity. *Journal of Molecular Biology* **368**, 509-520 (2007).
24. Szeltner, Z., et al. Characterization of a novel acylaminoacyl peptidase with hexameric structure and endopeptidase activity. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1794**, 1204-1210 (2009).
25. Gogliettino, M., et al. Identification and characterisation of a novel acylpeptide hydrolase from *Sulfolobus solfataricus*: structural and functional insights. *PLOS ONE* **7**, e37921 (2012).
26. Scaloni, A., et al. Deficiency of acylpeptide hydrolase in small-cell lung carcinoma cell lines. *Journal of Laboratory and Clinical Medicine* **120**, 546-552 (1992).

27. Riccio, A., et al. A new APEH cluster with antioxidant functions in the antarctic hemoglobinless icefish *Chionodraco hamatus*. *PLOS ONE* **10**, e0125594 (2015).
28. Gogliettino, M., et al. Selective inhibition of acylpeptide hydrolase in SAOS-2 osteosarcoma cells: is this enzyme a viable anticancer target? *Molecular Biology Reports* **48**, 1505-1519 (2021).
29. Tangri, A., et al. Deubiquitinase UCHL1 maintains protein homeostasis through the PSMA7–APEH–Proteasome axis in high-grade serous ovarian carcinoma. *Molecular Cancer Research* **19**, 1168-1181 (2021).
30. Palumbo, R., et al. APEH inhibition affects osteosarcoma cell viability via downregulation of the proteasome. *International Journal of Molecular Sciences* **17**, 1614 (2016).
31. Shimizu, K., Kiuchi, Y., Ando, K., Hayakawa, M. & Kikugawa, K. Coordination of oxidized protein hydrolase and the proteasome in the clearance of cytotoxic denatured proteins. *Biochemical and Biophysical Research Communications* **324**, 140–146 (2004).
32. Palmieri, G., et al. Acylpeptide hydrolase inhibition as targeted strategy to induce proteasomal down-regulation. *PLOS ONE* **6**, e25888 (2011).
33. Bergamo, P., et al. RedOx status, proteasome and APEH: Insights into anticancer mechanisms of t10, c12-conjugated linoleic acid isomer on A375 melanoma cells. *PLOS ONE* **8**, e80900 (2013).
34. Yamin, R., Zhao, C., O'Connor, P. B., McKee, A. C. & Abraham, C. R. Acyl peptide hydrolase degrades monomeric and oligomeric amyloid-beta peptide. *Molecular Neurodegeneration* **4**, 33 (2009).
35. Bachovchin, D. A., et al. A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nature Chemical Biology* **10**, 656-663 (2014).
36. Hoernstein, S. N. W., et al. Identification of targets and interaction partners of arginyl-tRNA protein transferase in the moss *Physcomitrella patens*. *Molecular & Cellular Proteomics* **15**, 1808-1822 (2016).
37. Schuessele, C., et al. Spatio-temporal patterning of arginyl-tRNA protein transferase (ATE) contributes to gametophytic development in a moss. *New Phytologist* **209**, 1014–1027 (2016).
38. Varshavsky, A. N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences U.S.A.* **116**, 358-366 (2019).
39. Lang, D., et al. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant Journal* **93**, 515-533 (2018).

40. Chen, F., Mackey, A. J., Stoeckert Jr, C. J. & Roos DS OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Research* **34**, D363-D368 (2006).
41. Zimmer, A. D., et al. Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. *BMC Genomics* **14**, 498 (2013).
42. Fernandez-Pozo, N., et al. PEATmoss (Physcomitrella Expression Atlas Tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *Plant Journal* **102**, 165-177 (2020).
43. Mergner, J., et al. Mass-spectrometry-based draft of the *Arabidopsis* proteome. *Nature* **579**, 409-414 (2020).
44. Altschul, S. F., et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402 (1997).
45. Merchant, S. S., et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245-250 (2007).
46. Bowman, J. L., et al. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **171**, 287-304 (2017).
47. Banks, J. A., et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**, 960-963 (2011).
48. Ouyang, S., et al. The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Research* **35**, D883-D887 (2007).
49. Tuskan, G. A., et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596-1604 (2006).
50. Hori, K., et al. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* **5**, 3978 (2014).
51. Kirbis, A., et al. Transcriptional landscapes of divergent sporophyte development in two mosses, *Physcomitrium* (Physcomitrella) *patens* and *Funaria hygrometrica*. *Frontiers in Plant Science* **11**, 747 (2020).
52. Ostendorf, A. K., et al. Polyploidization within the Funariaceae – a key principle behind speciation, sporophyte reduction and the high variance of spore diameters? *Bryophyte Diversity and Evolution* **43**, 164-179 (2021).
53. Lamesch, P., et al. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* **40**, D1202-D1210 (2012).

54. Rawlings, N. D., Polgar, L. & Barrett, A. J. A new family of serine-type peptidases related to prolyl oligopeptidase. *Biochemical Journal* **279**, 907-908 (1991).
55. Polgar, L. The prolyl oligopeptidase family. *Cellular and Molecular Life Sciences* **59**, 349-362 (2002).
56. Perroud, P. F., et al. The *Physcomitrella patens* gene atlas project: large-scale RNA-seq based expression data. *Plant Journal* **95**, 168-182 (2018).
57. Mueller, S. J., et al. Quantitative analysis of the mitochondrial and plastid proteomes of the moss *Physcomitrella patens* reveals protein macrocompartmentation and microcompartmentation. *Plant Physiology* **164**, 2081-2095 (2014).
58. Fesenko, I. A., et al. Specific pools of endogenous peptides are present in gametophore, protonema, and protoplast cells of the moss *Physcomitrella patens*. *BMC Plant Biology* **15**, 87 (2015).
59. Fesenko, I., et al. The *Physcomitrella patens* chloroplast proteome changes in response to protoplastation. *Frontiers in Plant Science* **7**, 1661 (2016).
60. Liu, H., et al. The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families. *Plant Journal* **105**, 1072-1082 (2021).
61. Beike, A. K. et al. Insights from the cold transcriptome of *Physcomitrella patens*: global specialization pattern of conserved transcriptional regulators and identification of orphan genes involved in cold acclimation. *New Phytologist* **205**, 869-881 (2015).
62. Cheng, C. Y., et al. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant Journal* **89**, 789-804 (2017).
63. Armenteros, J. J. A., et al. Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance* **2**, e201900429 (2019).
64. Sperschneider, J., et al. LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports* **7**, 44598 (2017).
65. Waszczak, C., Carmody, M. & Kangasjärvi, J. Reactive oxygen species in plant signaling. *Annual Review of Plant Biology* **69**, 209-236 (2018).
66. Lingner, T., et al. Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and *in vivo* subcellular targeting analyses. *Plant Cell* **23**, 1556-1572 (2011).
67. Reumann, S., Buchwald, D. & Lingner, T. PredPlantPTS1: a web server for the prediction of plant peroxisomal proteins. *Frontiers in Plant Science* **3**, 194 (2012).
68. Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J. & Subramani, S. A conserved tripeptide sorts proteins to peroxisomes. *Journal of Cell Biology* **108**, 1657-1664 (1989).

69. Wang, J., Wang, Y., Gao, C., Jiang, L. & Guo, D. PPero, a computational model for plant PTS1 type peroxisomal protein prediction. *PLOS ONE* **12**, e0168912 (2017).
70. Petriv, I., Tang, L., Titorenko, V. I. & Rachubinski, R. A. A new definition for the consensus sequence of the peroxisome targeting signal type 2. *Journal of Molecular Biology* **341**, 119-134 (2004).
71. Reumann, S. Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiology* **135**, 783-800 (2004).
72. Weise, A., et al. Use of *Physcomitrella patens actin* 5′ regions for high transgene expression: importance of 5′ introns*. Applied Microbiology and Biotechnology* **70**, 337-345 (2006).
73. Özdemir, B., et al. Cytological analysis and structural quantification of FtsZ1-2 and FtsZ2-1 network characteristics in *Physcomitrella patens*. *Scientific Reports* **8**, 11165 (2018).
74. Lang, E. G. E., et al. Simultaneous isolation of pure and intact chloroplasts and mitochondria from moss as the basis for sub-cellular proteomics. *Plant Cell Reports* **30**, 205–215 (2011).
75. Fuss, J., Liegmann, O., Krause, K. & Rensing, S. A. Green targeting predictor and ambiguous targeting predictor 2: the pitfalls of plant protein targeting prediction and of transient protein expression in heterologous systems. *New Phytologist* **200**, 1022-1033 (2013).
76. Orr, M. W., Mao, Y., Storz, G. & Qian, S. B. Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Research* **48**, 1029-1042 (2020).
77. Hohe, A., et al. An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, *Physcomitrella patens*. *Current Genetics* **44**, 339-347 (2004).
78. Noy-Malka, et al. A single CMT methyltransferase homolog is involved in CHG DNA methylation and development of *Physcomitrella patens*. *Plant Molecular Biology* **84**, 719-735 (2014).
79. Heck, M. A., et al. Axenic *in vitro* cultivation of 19 peat moss (*Sphagnum* L.) species as a resource for basic biology, biotechnology, and paludiculture. *New Phytologist* **229**, 861-876 (2021).
80. Kamisugi, Y., et al. The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic Acids Research* **34**, 6205-6214 (2006).

81. Rempfer, C., et al. Autopolyploidization affects transcript patterns and gene targeting frequencies in Physcomitrella. *Plant Cell Reports* **41**, 153-173 (2022).
82. Lorenz, S., Tintelnot, S., Reski, R. & Decker, E. L. Cyclin D-knockout uncouples developmental progression from sugar availability. *Plant Molecular Biology* **53**, 227-236 (2003).
83. Dalle-Donne, I., et al. Protein carbonylation: 2, 4-dinitrophenylhydrazine reacts with both aldehydes/ketones and sulfenic acids. *Free Radical Biology and Medicine* **46**, 1411-1419 (2009).
84. Deutsch, E. W., et al. The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. *Nucleic Acids Research* **48**, D1145-D1152 (2020).
85. Perez-Riverol, Y., et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research* **50**, D543-D552 (2022).
86. Johansson, E., Olsson, O. & Nyström, T. Progression and specificity of protein oxidation in the life cycle of *Arabidopsis thaliana*. *Journal of Biological Chemistry* **279**, 22204-22208 (2004).
87. Harman, D. Aging: A theory based on free radical and radiation chemistry. *Journal of Gerontology* **11**, 298–300 (1956).
88. Harman, D. The biologic clock: the mitochondria? *Journal of the American Geriatrics Society* **20**, 145-147 (1972).
89. Giorgi, C., et al. Mitochondria and reactive oxygen species in aging and age-related diseases. *International Review of Cell and Molecular Biology* **340**, 209-344 (2018).
90. Mao, C., et al. Nitric oxide regulates seedling growth and mitochondrial responses in aged oat seeds. *International Journal of Molecular Sciences* **19**, 1052 (2018).
91. Ratajczak, E., Małecka, A., Ciereszko, I. & Staszak, A. M. Mitochondria are important determinants of the aging of seeds. *International Journal of Molecular Sciences* **20**, 1568 (2019).
92. Munné-Bosch, S. & Alegre, L. Plant aging increases oxidative stress in chloroplasts. *Planta* **214**, 608-615 (2002).
93. Noctor, G., Veljovic-Jovanovic, S., Driscoll, S., Novitskaya, L. & Foyer, C. H. Drought and oxidative load in the leaves of C3 plants: a predominant role for photorespiration? *Annals of Botany* **89**, 841-850 (2002).

94. Foyer, C. H. & Noctor, G. Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. *Physiologia Plantarum* **119**, 355-364 (2003).
95. Sohal, R. S. & Weindruch, R. Oxidative stress, caloric restriction, and aging. *Science*, **273**, 59-63 (1996).
96. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194-1217 (2013).
97. Senthilkumar, R., Reddy, P. N. & Sharma, K. K. Studies on trypsin-modified bovine and human lens acylpeptide hydrolase. *Experimental Eye Research* **72**, 301-310 (2001).
98. Palmieri, G., et al. Low erythrocyte levels of proteasome and acyl-peptide hydrolase (APEH) activities in Alzheimer's disease: a sign of defective proteostasis? *Journal of Alzheimer's Disease* **60**, 1097-1106 (2017).
99. Marshall, I., et al. Analyzing the activity and expression of acyl peptide enzyme hydrolase (APEH) in the blood serum of patients with type II diabetes. *Bios* **90**, 70-78 (2019).
100. Resemann, H. C., et al. Convergence of sphingolipid desaturation across over 500 million years of plant evolution. *Nature Plants* **7**, 219-232 (2021).
101. Zybailov, B., et al. Sorting signals, N-terminal modifications, and abundance of the chloroplast proteome. *PLOS ONE* **3**, e1994 (2008).
102. Rugen, N., et al. Protein interaction patterns in *Arabidopsis thaliana* leaf mitochondria change in dependence to light. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1862**, 148443 (2021).
103. Maxwell, D. P., Wang, Y. & McIntosh, L. The alternative oxidase lowers mitochondrial reactive oxygen production in plant cells. *Proceedings of the National Academy of Sciences U.S.A.* **96**, 8271-8276 (1999).
104. Møller, I. M. Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annual Review of Plant Biology* **52**, 561-591 (2001).
105. Smakowska, E., Czarna, M. & Janska, H. Mitochondrial ATP-dependent proteases in protection against accumulation of carbonylated proteins. *Mitochondrion* **19**, 245-251 (2014).
106. Al-Khatib, K. & Paulsen, G. M. Enhancement of thermal injury to photosynthesis in wheat plants and thylakoids by high light intensity. *Plant Physiology* **90**, 1041-1048 (1989).

107. Wang, Q. L., Chen, J. H., He, N. Y. & Guo, F. Q. Metabolic reprogramming in chloroplasts under heat stress in plants. *International Journal of Molecular Sciences* **19**, 849 (2018).

108. Sweetlove, L. J., et al. The impact of oxidative stress on Arabidopsis mitochondria. *Plant Journal* **32**, 891-904 (2002).

109. Tiwari, B. S., Belenghi, B. & Levine, A. Oxidative stress increased respiration and generation of reactive oxygen species, resulting in ATP depletion, opening of mitochondrial permeability transition, and programmed cell death. *Plant Physiology* **128**, 1271-1281 (2002).

110. Willems, P., Ndah, E., Jonckheere, V., van Breusegem, F. & van Damme, P. To new beginnings: Riboproteogenomics discovery of N-terminal proteoforms in *Arabidopsis thaliana*. *Frontiers in Plant Science* **12**, 778804 (2021).

111. Kiessling, J., et al. Dual targeting of plastid division protein FtsZ to chloroplasts and the cytoplasm. *EMBO Reports* **5**, 889-894 (2004).

112. van Wijk, K. J., et al. The Arabidopsis PeptideAtlas: Harnessing worldwide proteomics data to create a comprehensive community proteomics resource. *Plant Cell* **33**, 3421-3453 (2021).

113. Zhang, H., et al. Arabidopsis proteome and the mass spectral assay library. *Scientific Data* **6**, 278 (2019).

114. Zeng, Z., et al. Acylpeptide hydrolase is a component of the cellular response to DNA damage. *DNA Repair* **58**, 52-61 (2017).

115. Brocca, S., et al. A bacterial acyl aminoacyl peptidase couples flexibility and stability as a result of cold adaptation. *FEBS Journal* **283**, 4310-4324 (2016).

116. Menyhárd, D. K., et al. A self-compartmentalizing hexamer serine protease from *Pyrococcus horikoshii*: substrate selection achieved through multimerization. *Journal of Biological Chemistry* **288**, 17884-17894 (2013).

117. Viaene, T., et al. Directional auxin transport mechanisms in early diverging land plants. *Current Biology* **24**, 2786-2791 (2014).

118. Chen, Z., et al. Comprehensive analysis of the *Ppatg3* mutant reveals that autophagy plays important roles in gametophore senescence in *Physcomitrella patens*. *BMC Plant Biology* **20**, 440 (2020).

119. Jia, M., et al. Noncanonical ATG8-ABS3 interaction controls senescence in plants. *Nature Plants* **5**, 212-224 (2019).

120. Dalle-Donne, I., Giustarini, D., Colombo, R., Rossi, R. & Milzani, A. Protein carbonylation in human diseases. *Trends in Molecular Medicine* **9**, 169-176 (2003).

121. Hipp, M. S., Kasturi, P. & Hartl, F. U. The proteostasis network and its decline in ageing. *Nature Reviews Molecular Cell Biology* **20**, 421-435 (2019).

122. Chen, Y. R., et al. Tissue-specific landscape of protein aggregation and quality control in an aging vertebrate. Preprint at: https://doi.org/10.1101/2022.02.26.482120 (2022).

123. Medina, R., et al. Phylogenomic delineation of *Physcomitrium* (Bryophyta: Funariaceae) based on targeted sequencing of nuclear exons and their flanking regions rejects the retention of *Physcomitrella*, *Physcomitridium* and *Aphanorrhegma*. *Journal of Systematics and Evolution* **57**, 404-417 (2019).

124. Reski, R. & Abel, W. O. Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta* **165**, 354-358 (1985).

125. Egener, T., et al. High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. *BMC Plant Biology* **2**, 6 (2002).

126. Schween, G., Hohe, A., Koprivova, A. & Reski, R. Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *Journal of Plant Physiology* **160**, 209–212 (2003).

127. Erxleben, A., Gessler, A., Vervliet-Scheebaum, M. & Reski, R. Metabolite profiling of the moss *Physcomitrella patens* reveals evolutionary conservation of osmoprotective substances. *Plant Cell Reports* **31**, 427-436 (2012).

128. Tian, G. W., et al. High-throughput fluorescent tagging of full-length *Arabidopsis* gene products in planta. *Plant Physiology* **135**, 25-38 (2004).

129. Sambrook, J. & Russell, D. W. Standard ethanol precipitation of DNA in microcentrifuge tubes. *Cold Spring Harbor Protocols* **2006**, pdb-prot4456 (2006).

130. Hohe, A. & Reski, R. Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Science* **163**, 69-74 (2002).

131. Decker, E. L., et al. Strigolactone biosynthesis is evolutionarily conserved, regulated by phosphate starvation and contributes to resistance against phytopathogenic fungi in a moss, *Physcomitrella patens*. *New Phytologist* **216**, 455-468 (2017).

132. Gibson, D. G., et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).

133. Amack, S. C. & Antunes, M. S. CaMV35S promoter–a plant biology and biotechnology workhorse in the era of synthetic biology. *Current Plant Biology* **24**, 100179 (2020).

134. Horstmann, V., Huether, C. M., Jost, W., Reski, R. & Decker, E. L. Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude. *BMC Biotechnology* **4**, 13 (2004).

135. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254 (1976).

136. Blum, M., et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research* **49**, D344-D354 (2021).

137. Mistry, J., et al. Pfam: The protein families database in 2021. *Nucleic Acids Research* **49**, D412-D419 (2021).

138. Hoernstein, S. N. W., et al. Host cell proteome of *Physcomitrella patens* harbors proteases and protease inhibitors under bioproduction conditions. *Journal of Proteome Research* **17**, 3749-3760 (2018).

139. Top, O., et al. Expression of a human cDNA in moss results in spliced mRNAs and fragmentary protein isoforms. *Communications Biology* **4**, 964 (2021).

140. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367-1372 (2008).

141. Cox, J., et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics* **13**, 2513-2526 (2014).

142. Tyanova, S., et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740 (2016).

143. Hoernstein, S. N. W., et al. Identification of targets and interaction partners of arginyl-tRNA protein transferase in the moss *Physcomitrella patens*. http://dx.doi.org/10.6019/PXD003232 (2016).

144. Nguyen, N. D., Mirarab, S., Kumar, K. & Warnor, T. Ultra-large alignments using phylogeny-aware profiles. *Genome Biology* **16**, 124 (2015).

145. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* **34**, W609–W612 (2006).

146. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

147. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2022).

148. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528 (2018).

149. Yu, G., Smith, D., Zhu, H., Guan, Y. & Lam, T. T. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28-36 (2017).

Supplementary Information

Figure S1 Expression levels of PpAARE genes and AtAARE in different tissues and under stress conditions.

Figure S2 Sequence alignment of AtAARE and PtAARE isoforms.

Figure S3 AARE activity in Physcomitrella plastids and mitochondria.

Figure S4 Identification of PpAARE single, double and triple KO mutants by PCR on genomic DNA.

Figure S5 FCM analysis of all confirmed PpAARE KO lines to analyse the ploidy.

Figure S6 AARE exopeptidase activity protonema of KO mutants.

Figure S7 Plate overview and additional OxyBlots from Physcomitrella gametophore colonies with loading controls.

Figure S8 Screening for haploid Physcomitrella lines with a citrine fusion transcript and overview on identification of the three PpAARE isoforms in different IP experiments.

Figure S9 Phenotypes of seedlings and adult AtAARE T-DNA mutant lines (GK, s68) and WT (Col).

Supplementary Table S1 List of primers used to assemble KO and KI constructs and primers used to screen for transgenic plants in Physcomitrella.

Supplementary Table S2 List of gene accession numbers for all sequences used for phylogenetic analysis and corresponding subcellular localization predictions.

Supplementary Table S3 List of primers used to screen possible T-DNA mutant plants in Arabidopsis.

Supplementary Table S4 List of accession numbers for the International Moss Stock Center (IMSC) of all transgenic Physcomitrella lines.

Supplementary Table S5 List of primers used for qPCR on genomic DNA to determine copy numbers of KO constructs in transgenic Physcomitrella lines.

Supplementary Table S6 List of primers used to assemble the constructs for transient localization analysis in Physcomitrella protoplasts.

Supplementary Table S7 Spectrum report of the revised database search of the anti-$Arg^{(Nt)}$ IP published in[36].

Supplementary Table S8 This table contains all numeric source data used to generate the graphs in Fig. 3b, Fig. 3e-g, Fig. 5b, Fig. 6c, Fig. 6e, Fig. S1, Fig. S3c, Fig. S4j-p, Fig. S5, Fig. S6, Fig. S8d, Fig. S8e and Fig. S8g.

Fig. 1 Phylogenetic tree for AARE isoforms from selected organisms, gene structures and proteins domains of *P. patens* and *A. thaliana* AARE isoforms. (a) Maximum likelihood tree based on an alignment of multiple *AARE* coding sequences. Duplication events in several species appear to be lineage-specific. Node values reflect percentage confidence values based on 1000 bootstrap replicates. Species abbreviations: At: *Arabidopsis thaliana*; Pt: *Populus trichocarpa*; Os: *Oryza sativa*; Pp: *Physcomitrium patens*; Fh: *Funaria hygrometrica*; Sf: *Sphagnum fallax*; Mp: *Marchantia polymorpha*; Sm: *Selaginella moellendorffii*; Kn: *Klebsormidium nitens*; Cr: *Chlamydomonas reinhardtii*; Rn: *Rattus norvegicus*; Hs: *Homo sapiens*; Ch: *Chionodraco hamatus*. (b) Gene structure of *PpAARE1-3* and *AtAARE*. For *PpAARE1* two splice variants exist. For *AtAARE* two different 5' UTR definitions are present (upper: Araport11[62]; lower: TAIR10[53]). (c) Left: Both splice variants of *PpAARE1* were amplified from complementary DNA (cDNA; +: with cDNA template; -: without reverse transcriptase). Expected amplicon size: *PpAARE1_1*: 2512 bp; *PpAARE1_2*: 2596 bp. Primers were designed to be specific for the splice variants (Table S1). Right: The longer open reading frame of *AtAARE* was amplified from cDNA. Expected amplicon size: 2457 bp (Table S3). (d) Protein structures showing PFAM-domains for PpAARE1-3 and AtAARE. All isoforms contain an AARE N-terminal domain (PF19283) and a catalytic Peptidase S9 domain (PF00326). PpAARE1, PpAARE2 and AtAARE additionally contain a WD40 domain (PF07676). The long splice variant of PpAARE1 and the longer open reading frame of AtAARE encode a predicted N-terminal plastid transit peptide (cTP). AA: amino acid. Cleavable N-terminal sequences were predicted by *TargetP2.0*[63].

Fig. 2 Confocal microscopy images showing the localization of PpAARE isoforms and AtAARE with C-terminal fused eGFP in Physcomitrella protoplasts. Fusion proteins with domain structures are depicted. PpAARE1_1, PpAARE2 and PpAARE3 localize to the cytoplasm. PpAARE1_2 localizes to specific foci in plastids (white arrows) and to mitochondria. The N-terminal extension of PpAARE1_2 encoding a predicted plastid transit peptide (PpAARE1_Nt) directs eGFP to plastids and mitochondria. The short variant of AtAARE (SV) localizes to the cytoplasm. The long variant (LV) localizes to plastids and mitochondria. The N-terminal extension of the long variant of AtAARE localizes to plastids and mitochondria. Left image: chlorophyll autofluorescence (red) and eGFP (green). Middle image: chlorophyll autofluorescence (red) and MitoTracker™ (magenta). Right image: chlorophyll autofluorescence (red), eGFP (green), Mitotracker™ (magenta) and co-localization of eGFP and MitoTracker™ (white). Bars = 7 µm.

Based on this data, PpAARE1 is targeted to three organelles: chloroplasts, mitochondria and the cytosol. To independently scrutinize these findings, we measured AARE enzyme activity in the

organelles. Using organelle fractionation as previously described[74], we detected AARE activity in chloroplasts, the cytosol, and mitochondria, although to a lesser extent in the latter (Fig. S3). Thus, *in vivo* localization of fusion proteins and enzyme measurements after cell fractionation independently confirm the predicted triple localization of AARE in Physcomitrella.

Fig. 3 Phenotype analyses, AARE activity and level of oxidized proteins in Physcomitrella WT and KO mutants. (a) Each isoform was knocked out using a different selection marker. The selection markers confer resistance against G418 (npt), hygromycin (hpt) and zeocin (zeo), respectively. Homologous regions (HR) were chosen in order to remove the full CDS of the corresponding gene upon integration of the selection marker. (b) Mean AARE activity on AcAla-pNA in Physcomitrella protonema with standard deviation (n= three biological replicates). Significance levels are based on a one-way Anova and subsequent post-hoc test (***: $p<0.001$) and indicate differences compared to WT. (c) Phenotypes of gametophores of WT and the different KOs cultivated on KnopME for 4 months. All colonies were grown on the same plate. Upper panel: Bar = 0.5 mm; lower panel: Bar = 2mm. (d) Protonema colonies grown on KnopME with 0.02% glucose taken from suspension culture. Bar = 1 mm. Images were taken 9 days after spotting. (e) Box plot showing the distribution of colony areas (n = 12 colonies per boxplot; colony images taken 7 days after spotting). No significant difference of colony areas between WT and the KOs was observed (tested via one-way Anova). (f) Box plot showing the number of gametophores per colony and day of WT and Δ*PpAARE1/2*. Counting was performed at indicated days after spotting protonema suspension culture on solid medium (KnopME with 0.02% glucose). Indicated significance levels are based on a two-way Anova with subsequent post-hoc test (***: $p<0.001$). The boxplots (e, f) depict the mean (horizontal bold line) of the data, the interquartile range (box) and the 1.5x interquartile range (whiskers). (g) Mean AARE exopeptidase activity on AcAla-pNA in gametophores. Depicted is the mean of three independent colonies per line with standard deviations. All colonies were cultivated on the same plate (Fig. S7a). Significance levels are based on one-way Anova and indicate differences compared to WT (***: $p<0.001$). (h) Levels of oxidized proteins in gametophores of WT and the different KOs (Coomassie-stained loading control is shown in Fig. S7b). The analysis was repeated three times (additional Blots available in Fig. S7c, d). Detection of oxidized proteins was performed with the OxyBlot™ *Protein Oxidation Detection Kit* (Merck).

Fig. 4 Physcomitrella gametophore colonies of varying age. (a) Colonies of all representative knockout mutants after 5 months on solid medium (KnopME). Bar = 5 mm. Gametophores of Δ*PpAARE1/2* and Δ*PpAARE1/3/2* are only viable at the tip, whereas plant material at the base

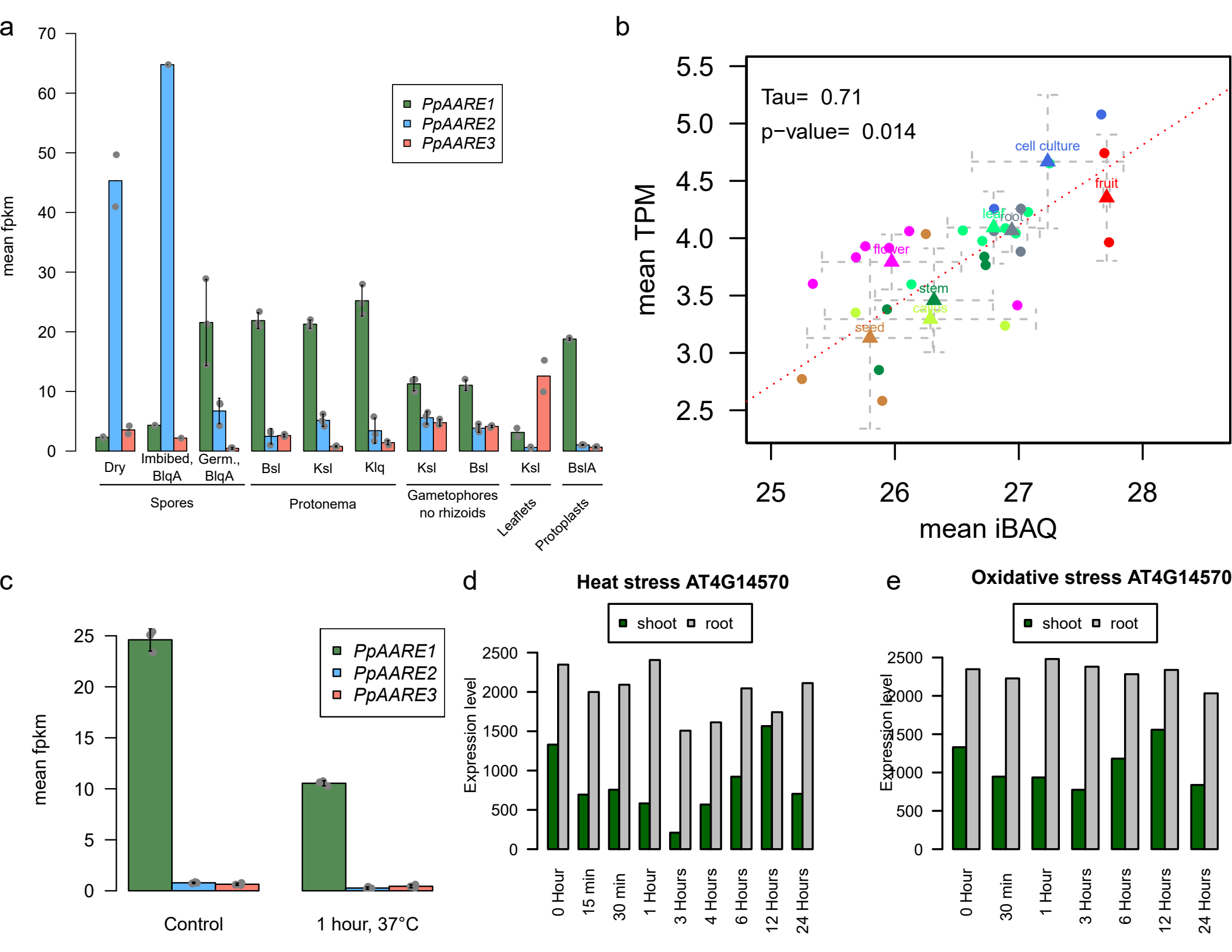
(red arrow) is dead. (b) Colonies after 8 months on solid medium. Δ*PpAARE1* is the parental line for Δ*PpAARE1/2*. Gametophores of WT and Δ*PpAARE1* are still viable whereas Δ*PpAARE1/2* gametophores are mostly dead. Bar = 2 mm. (c) Colonies after 8 months on solid medium. Δ*PpAARE1* is the parental line for Δ*PpAARE1/3* and Δ*PpAARE1/3* is the parental line for Δ*PpAARE1/3/2*. Colonies of WT, Δ*PpAARE1* and Δ*PpAARE1/3* still have viable gametophores, whereas Δ*PpAARE1/3/2* gametophores are dead. Bar = 2 mm.

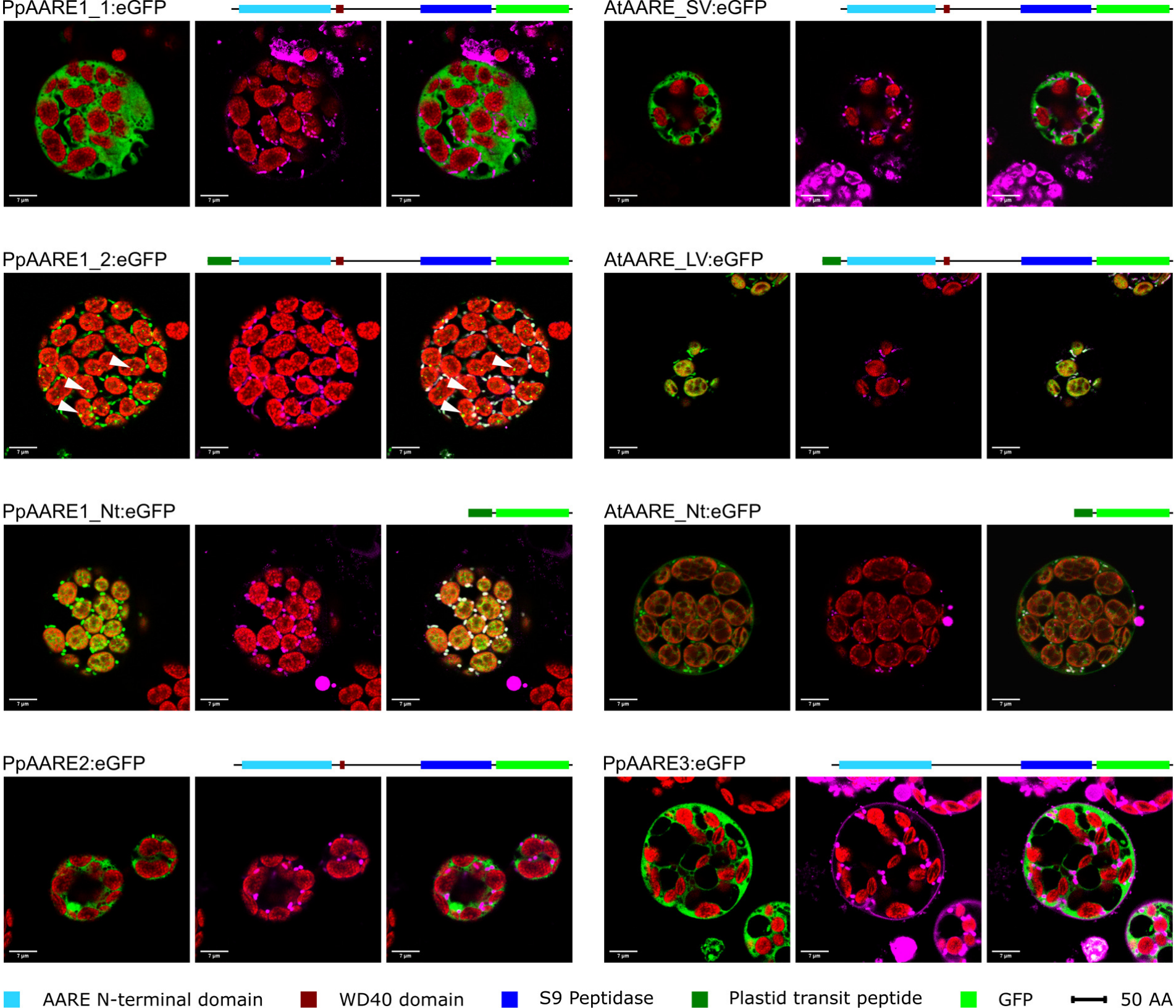
Fig. 5 Reciprocal co-immunoprecipitation (Co-IP) with Citrine-tagged PpAARE isoforms. (a) Tagging of PpAARE isoforms was realized by in-frame fusion with a linker-*Citrine* CDS at the respective native locus *via* homologous recombination. Original stop codons were deleted. (b) Volcano plots showing the result of the Co-IPs against each of the PpAARE:Citrine fusion proteins. Left panel: Pulldown of PpAARE1:Citrine. Right panel: Pulldown of PpAARE2:Citrine. Co-IP was performed with GFP-Trap Magnetic Particles M-270 (Chromotek) and protonema from suspension culture. Log_2 ratios of normalized label-free quantitation values (LFQ) are plotted against $-log_{10}$ of adjusted p-values. Proteins significantly enriched in the Citrine tagged pulldown are shown in blue circles ($p < 0.01$, FDR = 0.01). Significantly enriched PpAARE isoforms are depicted as red crosses.

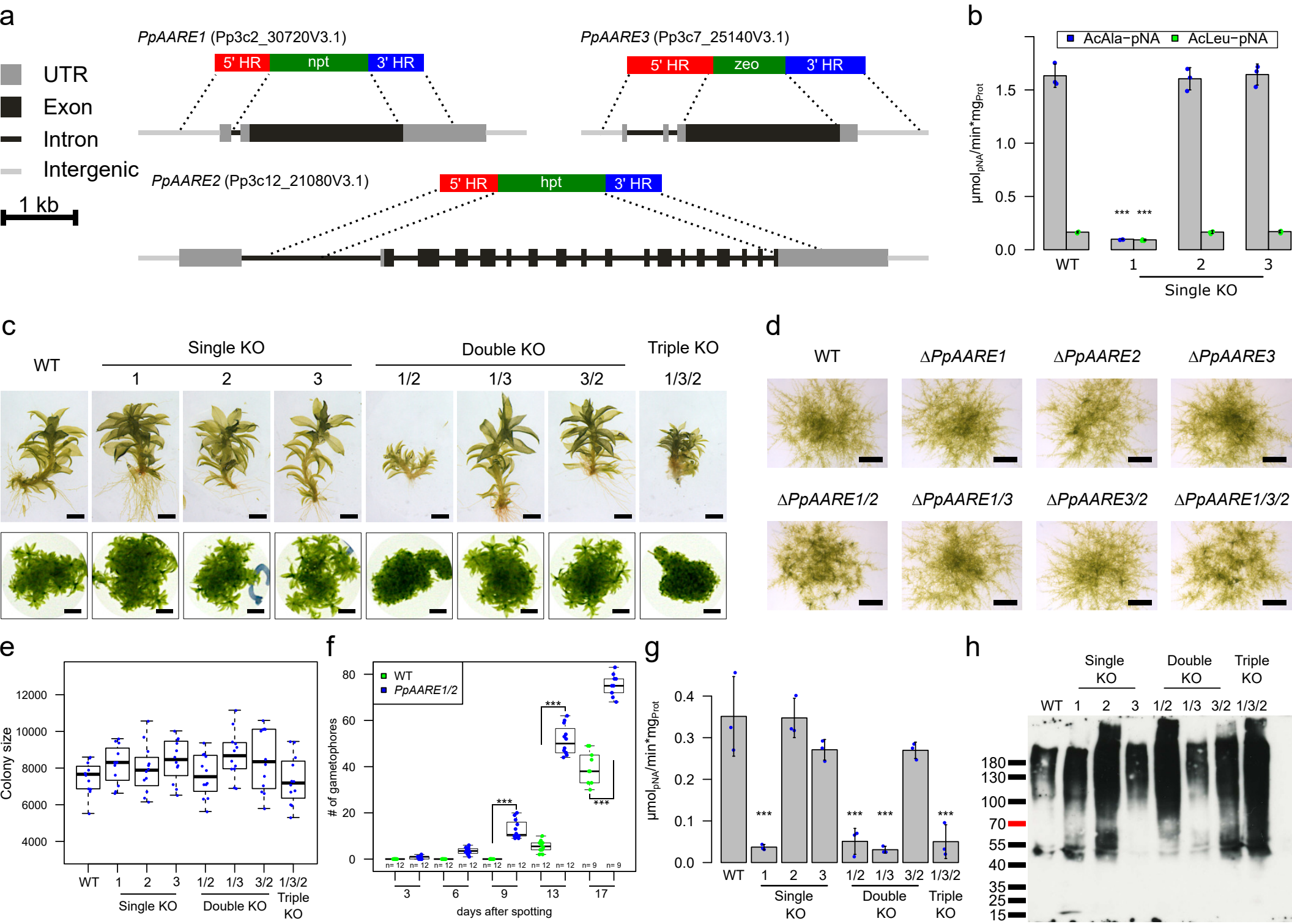
Fig. 6 Molecular and phenotypical characterization of *A. thaliana* T-DNA mutant lines. (a) Structure of the AtAARE gene and positions of T-DNA insertions (red arrows indicate positions of T-DNA insertion, s68 (SALK_071125) and GK (GK-456A10)) and primers (At1, At2) used for RT-PCR (below). Transcription analysis of *AtAARE* was performed by RT-PCR in WT (Col), s68, and GK plants. Negative controls without reverse transcriptase (-), a water control (H_2O) and genomic DNA control (gDNA) are indicated. Expected amplicon sizes: *AtAARE*: cDNA 739 bp, gDNA: 1204 bp; *PP2AA3* (reference gene): cDNA: 517 bp, gDNA: 911 bp. Primers are listed in Table S3. (b) Western blot analysis on extracts of the two T-DNA mutants and WT probed with a polyclonal @AARE antibody. In both T-DNA lines AtAARE is not detectable. Full blot images are available in Fig. S10. (c) Mean exopeptidase activity in *A. thaliana* extracts on AcAla-pNA and AcLeu-pNA with standard deviation (n = 3 biological replicates). Significance levels are based on one-way Anova and subsequent post-hoc test (***: $p<0.001$) and indicate differences compared to WT (Col). (d) Detection of oxidized proteins in WT (Col) and *AtAARE* mutant (s68). From three independent plants one exterior rosette leaf was taken after cultivation for 8 weeks in short-day conditions (8h light/16h dark) Bars correspond to 2 cm. Protein loading control by Coomassie staining is depicted below. (e) Boxplot of the bolting analysis of WT (Col, n = 14) and *AtAARE* mutant plants (s68, n=13). The boxplot (e, f) depicts the mean (horizontal bold line) of the data,

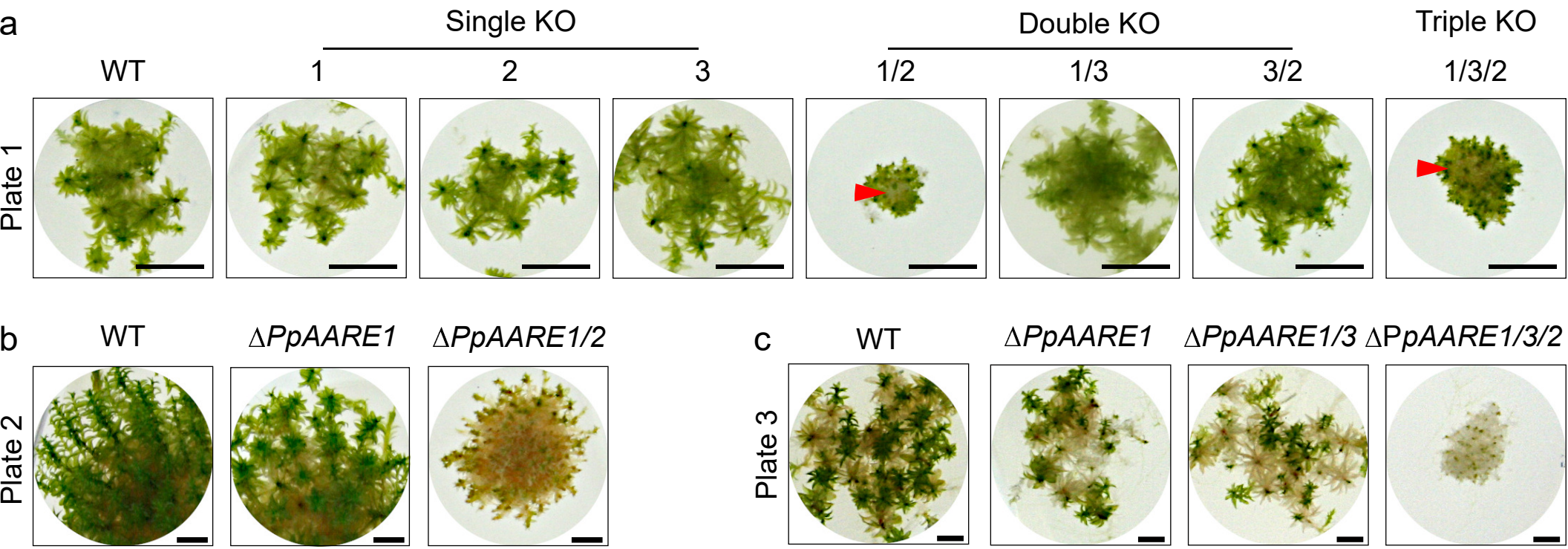
1361 the interquartile range (box) and the 1.5x interquartile range (whiskers). Outliers are depicted as
1362 white dots. Significance level is based on a one-way Anova and subsequent post-hoc test (***:
1363 $p<0.001$). Exemplary plants (45-day-old plants grown in long-day conditions, 16h light/8h dark)
1364 are depicted below the boxplot. Bars correspond to 1 cm.

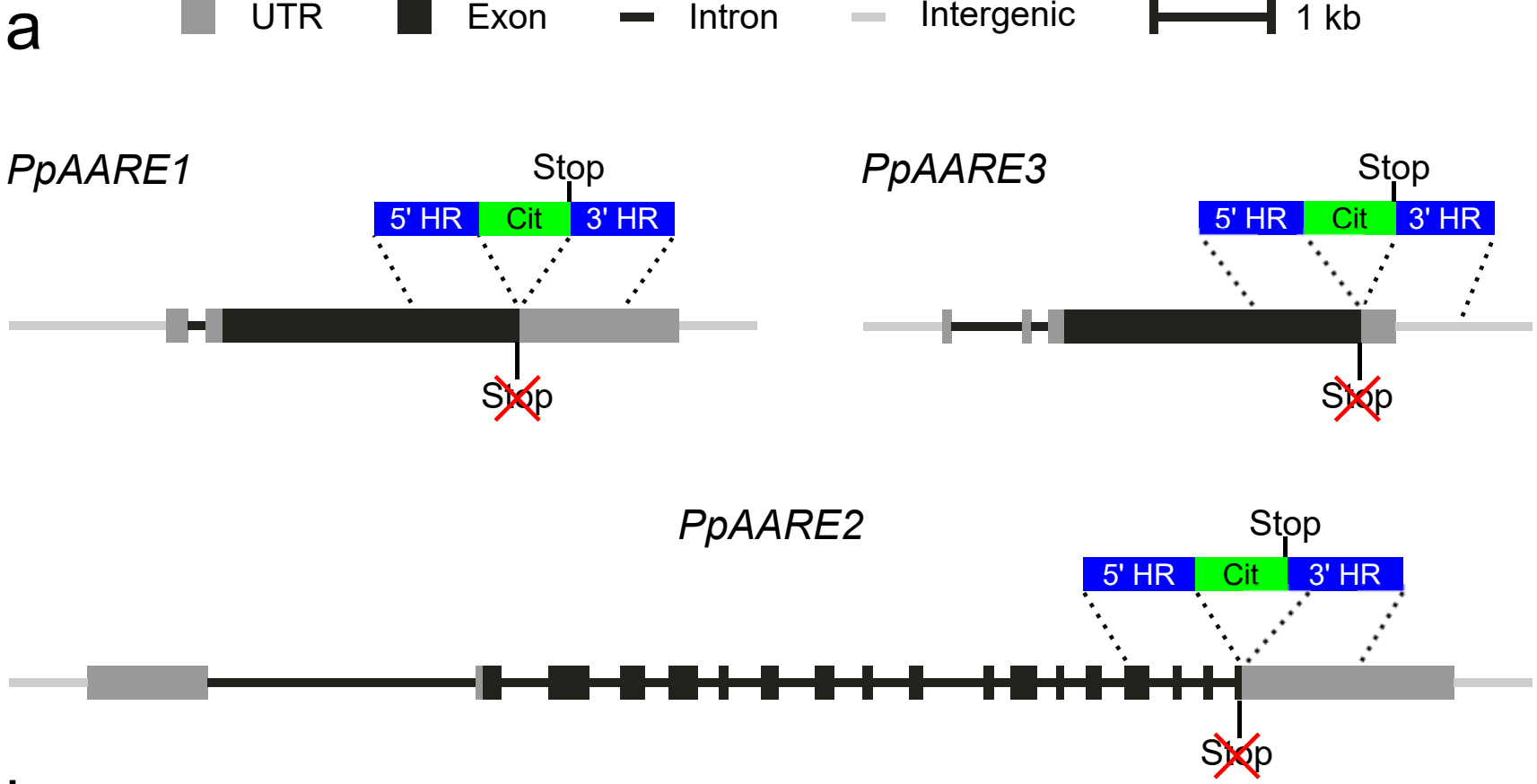
1365

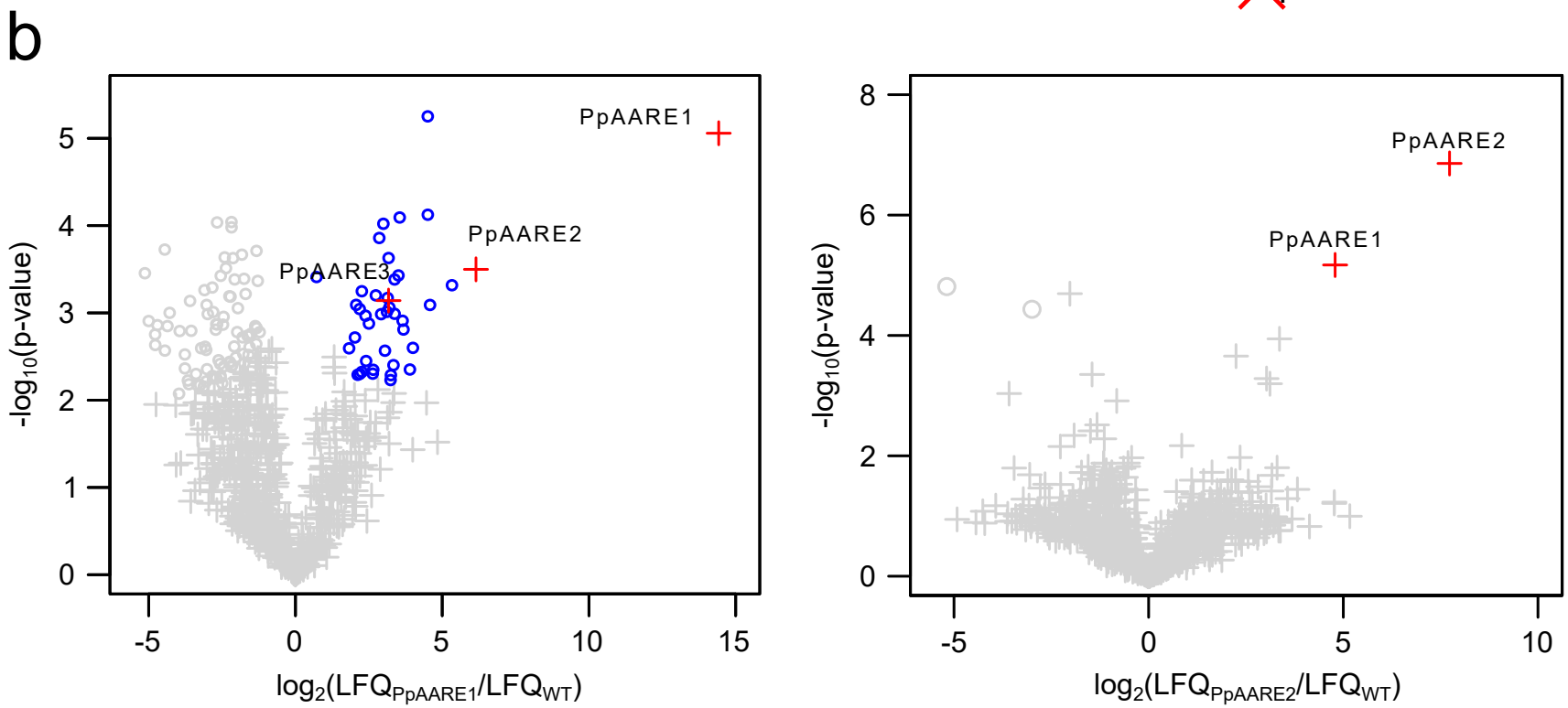

a
mean fpkm
PpAARE1
PpAARE2
PpAARE3
Dry
Imbibed, BlqA
Germ., BlqA
Bsl
Ksl
Klq
Ksl
Bsl
Ksl
BslA
Spores
Protonema
Gametophores no rhizoids
Leaflets
Protoplasts
b
Tau= 0.71
p−value= 0.014
mean TPM
mean iBAQ
cell culture
fruit
leaf
root
flower
stem
callus
seed
c
mean fpkm
PpAARE1
PpAARE2
PpAARE3
Control
1 hour, 37°C
d
Heat stress AT4G14570
shoot
root
Expression level
0 Hour
15 min
30 min
1 Hour
3 Hours
4 Hours
6 Hours
12 Hours
24 Hours
e
Oxidative stress AT4G14570
shoot
root
Expression level
0 Hour
30 min
1 Hour
3 Hours
6 Hours
12 Hours
24 Hours

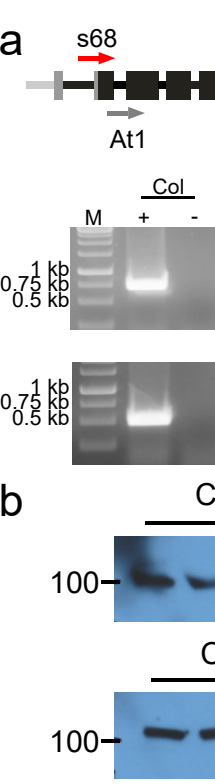

PpAARE1_1:eGFP
AtAARE_SV:eGFP
PpAARE1_2:eGFP
AtAARE_LV:eGFP
PpAARE1_Nt:eGFP
AtAARE_Nt:eGFP
PpAARE2:eGFP
PpAARE3:eGFP
AARE N-terminal domain
WD40 domain
S9 Peptidase
Plastid transit peptide
GFP
50 AA

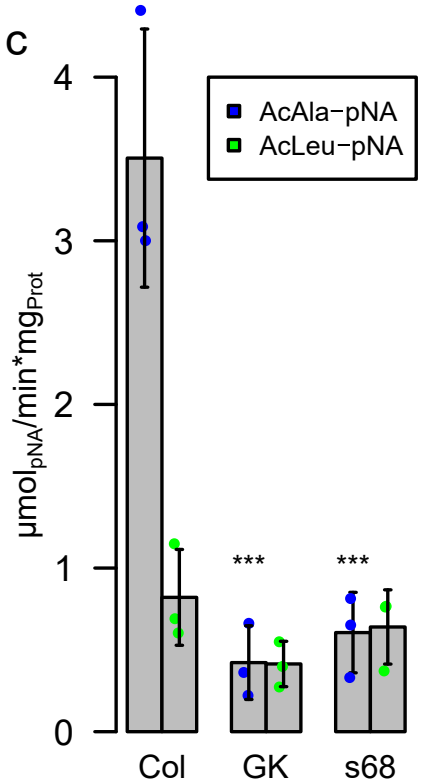

a
UTR
Exon
Intron
Intergenic
1 kb
PpAARE1 (Pp3c2_30720V3.1)
5' HR
npt
3' HR
PpAARE3 (Pp3c7_25140V3.1)
5' HR
zeo
3' HR
PpAARE2 (Pp3c12_21080V3.1)
5' HR
hpt
3' HR
b
AcAla-pNA
AcLeu-pNA
µmol pNA/min*mg Prot
1.5
1.0
0.5
0.0
***
***
WT
1
2
3
Single KO
c
WT
Single KO
1
2
3
Double KO
1/2
1/3
3/2
Triple KO
1/3/2
d
WT
ΔPpAARE1
ΔPpAARE2
ΔPpAARE3
ΔPpAARE1/2
ΔPpAARE1/3
ΔPpAARE3/2
ΔPpAARE1/3/2
e
Colony size
12000
10000
8000
6000
4000
WT
1
2
3
1/2
1/3
3/2
1/3/2
Single KO
Double KO
Triple KO
f
WT
PpAARE1/2
# of gametophores
80
60
40
20
0
***
***
***
n= 12
n= 9
3
6
9
13
17
days after spotting
g
µmol pNA/min*mg Prot
0.4
0.3
0.2
0.1
0.0
***
***
***
***
WT
1
2
3
1/2
1/3
3/2
1/3/2
Single KO
Double KO
Triple KO
h
Single KO
Double KO
Triple KO
WT
1
2
3
1/2
1/3
3/2
1/3/2
180
130
100
70
55
40
35
25
15

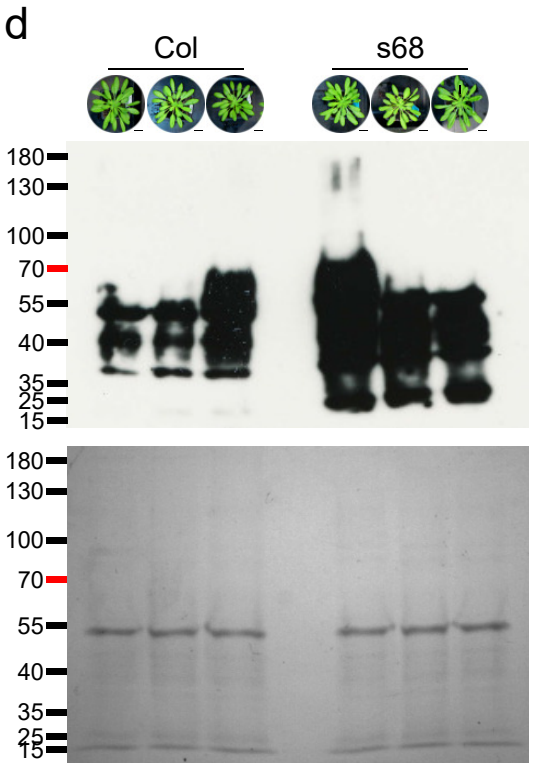

a
Single KO
Double KO
Triple KO
WT
1
2
3
1/2
1/3
3/2
1/3/2
Plate 1
b
WT
ΔPpAARE1
ΔPpAARE1/2
Plate 2
c
WT
ΔPpAARE1
ΔPpAARE1/3
ΔPpAARE1/3/2
Plate 3

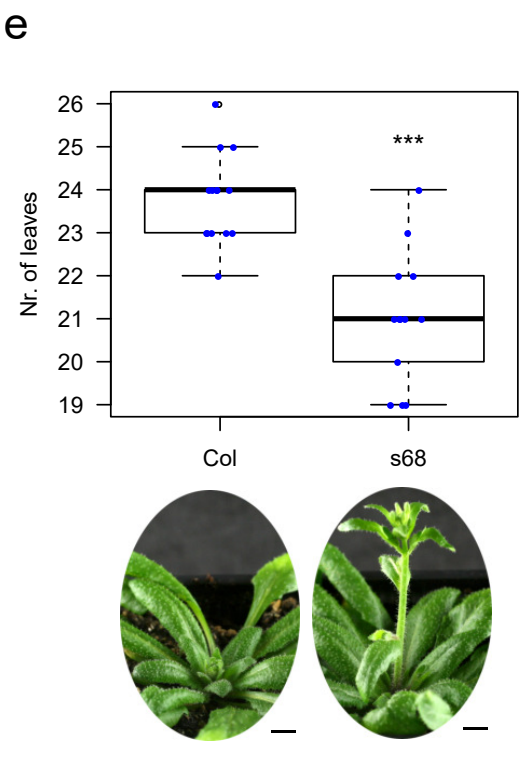

a
UTR
Exon
Intron
Intergenic
1 kb
PpAARE1
Stop
5' HR
Cit
3' HR
Stop
PpAARE3
Stop
5' HR
Cit
3' HR
Stop
PpAARE2
Stop
5' HR
Cit
3' HR
Stop


b
PpAARE1
PpAARE2
PpAARE3
-log10(p-value)
log2(LFQPpAARE1/LFQWT)
PpAARE2
PpAARE1
-log10(p-value)
log2(LFQPpAARE2/LFQWT)


a
s68
GK
At1
At2
Peptidase S9
Col
GK
s68
H2O
Col gDNA
M
+
-
1 kb
0.75 kb
0.5 kb
AtAARE
PP2AA3
b
Col
GK
100
Col
s68
100
αAtAARE


c
AcAla−pNA
AcLeu−pNA
μmol_pNA/min*mg_Prot
0
1
2
3
4
***
***
Col
GK
s68


d
Col
s68
180
130
100
70
55
40
35
25
15


e
Nr. of leaves
26
25
24
23
22
21
20
19
***
Col
s68