Patron: Her Majesty The Queen

ROTHAMSTED RESEARCH

Rothamsted Research
Harpenden, Herts, AL5 2JQ

Telephone: +44 (0)1582 763133
Web: http://www.rothamsted.ac.uk/

Rothamsted Repository Download

A - Papers appearing in refereed journals

Whitcomb, S. J., Rakpenthai, A., Bruckner, F., Fischer, A., Parmar, S., Erban, A., Kopka, J., Hawkesford, M. J. and Hoefgen, R. 2020. Cysteine and methionine biosynthetic enzymes have distinct effects on seed nutritional quality and on molecular phenotypes associated with accumulation of a methionine-rich seed storage protein in rice. *Frontiers in Plant Science.* 11, p. 118.

The publisher's version can be accessed at:

- https://dx.doi.org/10.3389/fpls.2020.01118

The output can be accessed at: https://repository.rothamsted.ac.uk/item/9813x/cysteine-and-methionine-biosynthetic-enzymes-have-distinct-effects-on-seed-nutritional-quality-and-on-molecular-phenotypes-associated-with-accumulation-of-a-methionine-rich-seed-storage-protein-in.

© 22 July 2020, Please contact library@rothamsted.ac.uk for copyright queries.

22/07/2020 12:23 repository.rothamsted.ac.uk library@rothamsted.ac.uk

Rothamsted Research is a Company Limited by Guarantee
Registered Office: as above. Registered in England No. 2393175.
Registered Charity No. 802038. VAT No. 197 4201 51.
Founded in 1843 by John Bennet Lawes.

Cysteine and methionine biosynthetic enzymes have distinct effects on seed nutritional quality and on molecular phenotypes associated with accumulation of a methionine-rich seed storage protein in rice.

1 **Sarah J Whitcomb[1], Apidet Rakpenthai[1], Franziska Brückner[1], Axel Fischer[2], Saroj Parmar[3],**
2 **Alexander Erban[4], Joachim Kopka[4], Malcolm J Hawkesford[3], Rainer Hoefgen[1*]**

3 [1]Laboratory of Amino Acid and Sulfur Metabolism, Department of Molecular Physiology, Max
4 Planck Institute of Molecular Plant Physiology, Potsdam, Germany
5 [2]Bioinformatics Infrastructure Group, Max Planck Institute of Molecular Plant Physiology, Potsdam,
6 Germany
7 [3]Plant Sciences Department, Rothamsted Research, Harpenden, United Kingdom
8 [4]Applied Metabolome Analysis Infrastructure Group, Max Planck Institute of Molecular Plant
9 Physiology, Potsdam, Germany
10

11 *** Correspondence:**
12 Rainer Hoefgen
13 hoefgen@mpimp-golm.mpg.de

14 **Life science identifier for *Oryza sativa* ssp. *japonica*:** urn:lsid:ipni.org:names:60471378-2

15 **Keywords: Methionine (Met), Cysteine (Cys), nutritional quality, Seed Storage Protein (SSPs),**
16 **serine acetyltransferase, cystathionine gamma-synthase, endoplasmic reticulum (ER stress),**
17 **Oryza sativa (rice).**

18 **Abstract**
19 Staple crops in human and livestock diets suffer from deficiencies in certain "essential" amino acids
20 including methionine. With the goal of increasing methionine in rice seed, we generated a pair of
21 "PushxPull" double transgenic lines, each containing a methionine-dense seed storage protein (2S
22 albumin from sunflower, HaSSA) and an exogenous enzyme for either methionine (feedback
23 desensitized cystathionine gamma synthase from Arabidopsis, AtD-CGS) or cysteine (serine
24 acetyltransferase from *E. coli*, EcSAT) biosynthesis. In both double transgenic lines, the total seed
25 methionine content was approximately 50% higher than in their untransformed parental line, *Oryza*
26 *sativa* ssp. *japonica* cv. Taipei 309. HaSSA-containing rice seeds were reported to display an altered
27 seed protein profile, speculatively due to insufficient sulfur amino acid content. However, here we
28 present data suggesting that this may result from an overloaded protein folding machinery in the
29 endoplasmic reticulum rather than primarily from redistribution of limited methionine from
30 endogenous seed proteins to HaSSA. We hypothesize that HaSSA-associated endoplasmic reticulum
31 stress results in redox perturbations that negatively impact sulfate reduction to cysteine, and we
32 speculate that this is mitigated by EcSAT-associated increased sulfur import into the seed, which
33 facilitates additional synthesis of cysteine and glutathione. The data presented here reveal challenges
34 associated with increasing the methionine content in rice seed, including what may be relatively low
35 protein folding capacity in the endoplasmic reticulum and an insufficient pool of sulfate available for
36 additional cysteine and methionine synthesis. We propose that future approaches to further improve
37 the methionine content in rice should focus on increasing seed sulfur loading and avoiding the
38 accumulation of unfolded proteins in the endoplasmic reticulum.

39 **1 Introduction**

40 Unlike plants, all animals lack the enzymatic machinery to synthesize *de novo* some of the 20
41 proteinaceous amino acids. These so called "essential amino acids" must be consumed in their diet.
42 Further, for optimal growth, these essential amino acids must be consumed in the right balance for
43 the animal's metabolic needs. Amino acids that are in excess of amount defined by the first limiting
44 amino acid will be catabolized, and the effective protein content of the feed will be reduced.

45 Methionine (Met) is one such essential amino acid but the amount present in plant-based animal
46 feed blends is typically insufficient optimal livestock growth and health. Cysteine (Cys) is not strictly
47 considered an essential amino acid in the diets of animals because it can be synthesized from
48 methionine, but in dietary situations where methionine is limited, cysteine becomes conditionally
49 essential. Increasing the methionine and cysteine content of commodity cereals and grain legumes
50 would benefit farmers by elevating the value of their crop, and would be of benefit to livestock
51 rearing by reducing the need for synthetic amino acid supplementation of animal feed.
52 Crops differ dramatically in how they store methionine. For example, in potato tubers 90% of the
53 methionine is soluble (Dancs et al., 2008), while in alfalfa leaves (Amira et al., 2005) and cereal
54 grains (Amir et al., 2018) almost all methionine is incorporated into protein. In tissues such as seeds
55 that store Met predominantly in protein, a relatively direct approach to elevate the Met content is to
56 increase the protein sink strength fraction by introducing genes for methionine-rich seed storage
57 proteins (SSP). This approach assumes that the sink strength for methionine in endogenous seed
58 proteins is relatively low and this limits methionine accumulation in the seed. As seeds contain only
59 low levels of free amino acids (Amir et al., 2018), this approach also assumes that the synthesis
60 and/or metabolism of methionine is sensitive to signaled demand from SSPs. In order to achieve
61 meaningful increases in protein-incorporated Met, the transgene needs to be highly expressed and the
62 peptide/protein stable in the targeted tissue. Typically accumulation of foreign proteins is enhanced
63 by targeting to the endoplasmic reticulum (ER) (Twyman et al., 2003). However, this can have the
64 undesired consequence of overloading the protein folding and processing capacity of the ER (Oono et
65 al., 2010; De Wilde et al., 2013).

66 Metabolic engineering of methionine biosynthesis is an alternative approach to increase
67 methionine in the seed. The choice of which enzyme(s) to modify in which tissue(s) is complex and
68 based on species-specific knowledge (or assumptions), such as where the methionine used in seed
69 tissue for protein translation is synthesized, and in species capable of *de novo* Met synthesis in seeds,
70 whether all steps of the sulfur assimilation pathway are also active in seeds as opposed to a pathway
71 intermediate being transported into the seed. Additionally, one must consider if biosynthetic and
72 transport pathways in other tissues can compensate for bottlenecks and limitations in the seed. The
73 main assumption behind this metabolic engineering approach is that the pool size of free methionine
74 and/or metabolic flux to methionine in the seed influence the profile of proteins synthesized.

75 Over the past several decades the general methods described above have been successfully
76 used to substantially increase the methionine content of maize, soybean, and several other grain
77 legumes (Molvig et al., 1997; Chiaiese et al., 2004; Song et al., 2013; Kim et al., 2014; Xiang et al.,
78 2017; Amir et al., 2019). Although the quantity of rice used in livestock feed is currently dwarfed by
79 other commodity crops (Global Rice Science Partnership, 2013), it is important to address
80 improvement of rice protein quality. Among the major cereals, rice has the highest net protein
81 utilization by livestock (Juliano, 1992). Furthermore, the regions of the world that cultivate maize
82 and rice are globally distinct, and the development of rice varieties with increased methionine content
83 would allow livestock farmers in rice-focused regions to reduce their reliance on blending higher
84 methionine maize with soybeans for their livestock feed.

85 The sunflower seed albumin 2A, HaSSA, has attributes that make it an attractive choice for
86 transgenic methods to study the effect of increased protein sink strength for sulfur amino acids (S-
87 AA): it is remarkably dense in methionine and cysteine (16% and 8% by length, respectively) (Kortt

This is a provisional file, not the final typeset article

88 et al., 1991), and it has been shown to be processed correctly in the seeds of several crops, which is
89 important for transgene protein accumulation. Furthermore, it is rumen stable (Spencer et al., 1988;
90 McNabb et al., 1994), making it suitable for blended sheep feeds, which need to be particularly rich
91 in methionine for wool growth (Qi and Lupton, 1994). Based on positive results in lupin seed
92 (Molvig et al., 1997), the *HaSSA* gene was introduced into rice under the control of a wheat glutelin
93 (SSP gene) promoter and targeted to the endoplasmic reticulum (ER), where endogenous SSPs are
94 post-translationally modified by disulfide bond formation and glycosylation (Hagan et al., 2003). By
95 expressing HaSSA the goal was to "pull" additional S-AA, Met and Cys, into the seed protein
96 fraction. Hagan and colleagues achieved high levels of HaSSA accumulation (to approximately 7%
97 of total seed protein) in their transformed rice line, hereafter referred to as SSA (Hagan et al., 2003).
98 At this level of accumulation, if the methionine incorporated into HaSSA was simply additive to
99 endogenous seed proteins at parental accumulation levels, then a 40% increase in total seed
100 methionine would be expected in SSA seeds. However, only a 25% increase in total seed methionine
101 was observed, and due to high intra-line variation this increase was not considered statistically
102 significant (Hagan et al., 2003). In addition to the gap between predicted and observed change in total
103 seed methionine, the profile of expressed proteins in SSA seeds was clearly different from the
104 parental rice cultivar, Taipei 309 (Hagan et al., 2003). Several of the major changes were consistent
105 with the hypothesis that free methionine became strongly limiting in the SSA seeds and the
106 heterologous protein diverted free methionine away from endogenous protein translation. These
107 results suggested that sink strength is not the primary factor limiting methionine content in rice seeds
108 and that subsequent efforts in rice should focus on increasing the supply of methionine in the seed.
109 To address the assumed deficiency in methionine supply in rice, two enzymes in the
110 methionine biosynthetic pathway have been targeted for modification: cystathionine-gamma-synthase
111 (CGS) catalyzes the first of three enzymatic steps to synthesize methionine from cysteine (Hesse et
112 al., 2004); serine acetyl-transferase (SAT) catalyzes the formation of O-acetyl serine (OAS) which
113 provides the carbon backbone for cysteine, the thiol precursor for methionine (Watanabe et al.,
114 2015). In rice it is not known to what extent methionine for protein synthesis is synthesized *de novo*
115 in the seed and to what extent it (or a precursor) is imported from another tissue such as leaves. With
116 this in mind, a ubiquitin promoter was chosen to drive expression of heterologous transgenes coding
117 for these key enzymes in S-AA biosynthesis (Nguyen et al., 2012; Whitcomb et al., 2018).
118 In the case of CGS, a feedback-desensitized version of the enzyme from Arabidopsis (AtD-
119 CGS) was chosen. Heterologous expression *AtD-CGS* in tobacco (Hacham et al., 2008; Matityahu et
120 al., 2013), soybean (Song et al., 2013), and azuki bean (Hanafy et al., 2013) resulted in large
121 increases in free methionine in vegetative tissues and in seeds. However, heterologous expression of
122 *AtD-CGS* in rice did not result in increased methionine in leaves or seeds despite persistently
123 elevated CGS activity in leaves (enzymatic activity in seeds was not tested) (Whitcomb et al., 2018).
124 We suggested that flux to methionine was increased in the *AtD-CGS* transgenic lines (hereafter
125 referred to simply as CGS), but in rice the concentration of free methionine may be homeostatically
126 regulated and the additional synthesized methionine catabolized, similar to results for lysine (Karchi
127 et al., 1994).
128 Ubiquitin promoter-driven expression of *cysE* from *E. coli* (*EcSAT*) was found to be more
129 successful in rice than ubiquitin promoter-driven expression of *AtD-CGS* (Nguyen et al., 2012;
130 Whitcomb et al., 2018). While the concentration of free methionine remained unchanged in the seeds
131 of the *EcSAT* transgenic lines (hereafter referred to simply as SAT), total methionine was
132 significantly increased in some of the generated lines. These results suggested that methionine
133 accumulation in SSA seeds (Hagan et al., 2003) may not be limited primarily by insufficient
134 methionine synthesis but rather by insufficient cysteine synthesis.
135 Here we generated double transgenic lines, containing both increased sink strength ("pull")
136 for methionine and increased cystine/methionine biosynthetic enzyme activity ("push"), to test

137 whether combining these traits would result in a synergistic increase in rice seed methionine. This
138 combinatorial approach has proved successful in maize, narbon bean, and potato (Demidov et al.,
139 2003; Dancs et al., 2008; Planta et al., 2017), but to our knowledge, this is the first report to combine
140 increased sink and source strength for methionine in rice. Specifically, we crossed SSA transgenic
141 plants with the SAT transgenic line or the CGS transgenic line that we deemed most promising.
142 Among the SAT rice lines generated by Nguyen et al. (2012), all had similarly high free cysteine
143 (Cys) levels in their seeds. SAT47 was chosen for our "PushxPull" study because it showed the
144 greatest increase in protein-incorporated methionine. Among the CGS lines generated by Whitcomb
145 et al. (2018), CgSx4 was chosen because it had the highest measured CGS activity and marginally
146 higher free methionine in seeds.

147 **2 Results**

148 **2.1 Expression of *AtD-CGS*, *EcSAT*, and *HaSSA* transcripts and accumulation of HaSSA**
149 **protein in seeds**

150 As a first step to characterize the seeds of the five transgenic lines used in this study we used
151 RNA-seq data to determine transcript abundances of the *AtD-CGS* transgene and endogenous *CGS*
152 genes, the *EcSAT* transgene and endogenous *SAT* genes, and the *HaSSA* transgene. The *AtD-CGS* and
153 *EcSAT* transgenes are both driven by a maize ubiquitin promoter, and their transcripts were highly
154 abundant in milkripe seeds, even after at least five generations post-transformation (Fig. 1A, B).
155 Considering only expressed genes in milkripe seeds, the *AtD-CGS* and *EcSAT* transgenes were
156 expressed in the 97th and 99th percentile, respectively, in this tissue. We identified one expressed
157 endogenous *CGS* transcript (Os03g0376100, 76th percentile) and three expressed endogenous *SAT*
158 transcripts (Os03g0133900, Os03g0196600, and Os05g0533500, 81st, 21st, and 24th percentile,
159 respectively) in milkripe seeds. The endogenous *CGS* transcript level was not significantly affected
160 in any of the transgenic lines, but we did observe moderate upregulation of one of the endogenous
161 *SAT* isoforms (Os05g0533500) in SSA seeds (1.6-fold relative to Taipei, padj < 0.001) and
162 CGSxSSA seeds (1.4-fold relative to Taipei, padj < 0.02). We did not find any compensatory
163 reduction in endogenous *CGS* or *SAT* transcripts in seeds expressing the *AtD-CGS* or *EcSAT*
164 transgenes. These data show that at least one endogenous isoform of *CGS* and *SAT* is highly
165 expressed in milkripe seeds. Therefore, transgene expression is likely to increase total enzymatic
166 activity of the relevant methionine or cysteine biosynthetic step rather than to bring a new
167 biosynthetic "trait" into the seeds.

168 Transcripts of the *HaSSA* transgene were also highly abundant in milkripe seeds (>99th
169 percentile) (Fig. 1C), but as a sink for methionine and cysteine, HaSSA protein levels are of greater
170 importance. Equal quantities of the SDS-soluble protein fraction from mature seeds were separated
171 by SDS-PAGE and Coomassie stained (Fig. 1D). Hagan et al. 2003 reported HaSSA accumulation
172 equivalent to approximately 7 % of salt extracted seed protein (Hagan et al., 2003), but we were not
173 able to clearly identify the HaSSA protein band in SDS-solubilized seed protein by Coomassie
174 staining alone. Identification of the HaSSA protein band achieved by western blotting with a
175 polyclonal HaSSA antibody. We observed a similar level of HaSSA protein accumulation in the
176 CGSxSSA and SATxSSA seeds, but it did not appear to be higher than in the parental SSA line (Fig.
177 1D), despite the increased expression of the S-amino acid biosynthetic enzyme transgenes (Fig. 1A,
178 B).
179

180 **2.2 Seed protein profiles**

181 The protein profile of SSA seeds is different from that of the parental Taipei as the abundance
182 of several major SSPs is altered: reduction in glutelin acidic and basic subunits, reduction in alpha-
183 globulin, and accumulation of prolamin 7/14 (Hagan et al., 2003; Islam et al., 2005). Since total S-
184 AA and protein content of seeds was not significantly changed in SSA seeds, the authors concluded

This is a provisional file, not the final typeset article

185 that production of HaSSA protein diverted the limited supply of free cysteine and methionine from
186 endogenous relatively S-AA rich SSPs.
187 Based on these data, we expected *EcSAT* and/or *AtD-CGS* to suppress the altered protein
188 profile phenotype of *HaSSA* in double transgenic seeds. However, the protein profiles of CGSxSSA
189 and SATxSSA seeds look very similar to SSA (Fig. 2). In contrast to SSA, the 1D SDS-PAGE
190 protein profiles of SAT and CGS appeared indistinguishable from Taipei (Fig. 2). These data suggest
191 that either additional S-AA production in EcSAT- and AtD-CGS-containing lines is insignificant or
192 other factors besides S-AA limitation contribute to the altered seed protein profile of HaSSA-
193 containing seeds.
194
195 **2.3 Expression of endoplasmic reticulum quality control genes**
196 Other examples in rice seeds of highly expressed foreign proteins targeted to the secretory
197 pathway have also resulted in significant changes in endogenous protein accumulation that partially
198 overlap with those we observe in HaSSA-containing seeds (Takagi et al., 2005; Yasuda et al., 2005;
199 Oono et al., 2010). Notably, the foreign proteins in these studies were not particularly rich in
200 methionine or cysteine. More generally, strong expression of secretory pathway-targeted transgenes
201 puts a heavy demand on the quality control machinery of the ER and can induce the unfolded protein
202 response (UPR) in an attempt to maintain protein folding and modification fidelity as well as timely
203 transport of proteins out of the ER to other destinations in the secretory pathway (Oono et al., 2010;
204 De Wilde et al., 2013; Liu and Howell, 2016). The accumulation of glutelin precursors in the
205 HaSSA-containing seeds (Fig. 2) could indicate retention of proglutelin in the ER due to misfolding
206 or insufficient trafficking of proglutelin out of the ER via COPII vesicles to the Golgi and eventually
207 to protein storage vacuoles where it is proteolytically cleaved into acidic and basic subunits (Satoh-
208 Cruz et al., 2010; Tian et al., 2013; Qian et al., 2015). Therefore, we investigated whether the altered
209 protein profile in HaSSA-containing seeds may be associated with activation of an unfolded protein
210 response (UPR) due to an overloaded ER protein processing machinery.
211 Specific changes characteristic of the UPR include increases in ER resident chaperone, co-
212 chaperone, and protein disulfide isomerase expression (Liu and Howell, 2016). The most abundant
213 chaperone system in the ER lumen is composed of ATP-regulated Hsp70 family chaperones (BiP),
214 ATP-independent co-chaperones of the Hsp40 family (DnaJ) and a nucleotide exchange factor.
215 Additional important chaperones include the Hsp90 family (GRP94) and the calnexin/calreticulin
216 (CNX/CRT) proteins. Protein disulfide isomerases (PDI) mediate disulfide bond formation,
217 dissolution, and reformation and are critical for protein folding. Differential expression analysis of
218 these chaperones, co-chaperones, and protein disulfide isomerases revealed a general upregulation in
219 seeds of the three HaSSA-containing lines and only minor expression changes in seeds of SAT and
220 CGS lines (Fig. 3). Further, the major isoform in milkripe seeds of each enzyme family (BiP1
221 Os02g0115900, GRP94 Os06g0716700, CNX Os04g0402100, CRT Os07g0246200, and PDIL1-1
222 Os11g00199200) is strongly and significantly upregulated specifically in the HaSSA-containing
223 lines. Minor isoforms BiP4 Os05g0428600 and PDIL2-3 Os09g0451500, the nucleotide exchange
224 factor for Hsp70 chaperones (NEF Os09g0512700), and the electron acceptor for PDI-mediated
225 disulfide bond formation (ERO1 Os03g0733800) were also found to be upregulated, further
226 supporting a functional increase in ER protein folding capacity in HaSSA-containing seeds. Notably,
227 we observed little if any attenuation of UPR gene upregulation in the CGSxSSA and SATxSSA
228 seeds, as might be expected if S-AA limitation for HaSSA translation were the cause of the increased
229 chaperone expression.
230
231 **2.4 Methionine, cysteine, and glutathione levels in mature seeds**
232 Published results from our group and others suggest that methionine content in rice is limited by
233 both sink strength (the proportion of methionine codons in SSP transcripts) and by source strength

234 (synthetic flux to S-AA) (Hagan et al., 2003; Lee et al., 2003; Nguyen et al., 2012; Whitcomb et al.,
235 2018). Total seed methionine in SSA seeds was not significantly different from parental Taipei
236 despite the strong accumulation of HaSSA protein (Hagan et al., 2003). We measured a comparable
237 increase (25% compared to Hagan's 30%), but due to much lower intra-line variation in our data, the
238 calculated p-value (< 0.005) was found to be highly significant (Fig. 4A). Our selected SAT line also
239 showed increased total methionine relative to parental Taipei, but this increase was lower than
240 reported by Nguyen et al. (2012) (1.45 fold compared to 4.75 fold) (Fig. 4A). This disparity may be
241 due to a different technical method for determining total amino acid concentration in the seed. We
242 chose here a standard method used in food quality control and employed a company to run the
243 respective analyses. Consistent with Whitcomb et al. (2018), ubiquitin promoter-driven expression of
244 feedback-desensitized *AtD-CGS* did not result in an appreciable increase in total seed methionine
245 (Fig. 4A).
246 The primary question of our study was whether push and pull traits would synergistically interact
247 when combined in the same rice plant to allow increased accumulation of methionine in the seed. The
248 data give a different answer to this question depending on which "push" trait is used: increased
249 activity in the cysteine biosynthesis pathway, or farther downstream in the methionine pathway
250 specifically. Combining *EcSAT* and *HaSSA* in seeds, both of which successfully increased total Met
251 individually, did not lead to a further increase. However, co-expression of *AtD-CGS* with *HaSSA* in
252 the seed was found to synergistically affect total seed methionine: CGSxSSA seeds had higher total
253 methionine than either parent and higher than the sum of total methionine level in each parent, but
254 accumulation was not higher than in SAT and SATxSSA seeds. Three transgenic lines (SAT,
255 SATxSSA and CGSxSSA) had total methionine levels comparable to those in IR64, a "high protein"
256 variety, reflecting an increase of approximately 50% over the parental Taipei line (Fig. 4A).
257 In our transgenic lines, the magnitude of changes in total cysteine in the seed were found to be
258 smaller than those for methionine (Fig. 4B). The HaSSA protein is extremely methionine-rich (16%
259 by length), but it also has a higher proportion of cysteine residues (8%) than most rice seed proteins.
260 Despite this characteristic, total cysteine was actually lower in SSA seeds than in Taipei; the
261 observed increases in total methionine seemed to come at the expense of total cysteine in this line.
262 Hagan et al. 2003 also observed a decrease in total cysteine in SSA seeds. SAT catalyzes the
263 formation of OAS, the carbon backbone for cysteine synthesis, and as expected, expression of *EcSAT*
264 resulted in an increase in total cysteine in the seed (Hagan et al., 2003). Combining *HaSSA* with
265 *EcSAT* relieved the decrease in total Cys found in SSA seeds but did not lead to a further increase
266 above *EcSAT* alone (as is true for total Met), and although cysteine is a substrate for AtD-CGS, no
267 decrease in total cysteine was observed in CGS seeds.
268 The observed increase in S-AA content in SATxSSA and CGSxSSA seeds (Fig. 4C) could be
269 explained by especially high HaSSA levels, but based on the SSA western blot (Fig. 1D), HaSSA
270 protein accumulation in the double transgenic lines was if anything slightly lower than in SSA alone.
271 The total S-AA data (Fig. 4C) taken together with the seed protein profiles (Fig. 2) and ER chaperone
272 gene expression patterns (Fig. 3) argue for an ER-stress based mechanism for seed protein profile
273 change in HaSSA-containing seeds, rather than a redistribution of dramatically limited S-AA supply.
274 In rice seeds the proportion of S-amino acids that remain free rather than incorporated into
275 polypeptides is extremely small. In the case of methionine, the average total pool size in Taipei was
276 measured to be 2 g/kg dry weight (DW), which is equivalent to approximately 13000 pmol/mg DW,
277 while the average measured free methionine concentration was only approximately 1.4 pmol/mg
278 DW. Highly significant differences in free methionine concentration were observed, but given how
279 small the pool of free methionine is relative to the total pool size it is unclear whether these
280 differences have any biological relevance (Fig. 4D). While the total quantity of methionine and
281 cysteine stored in the rice seed are very similar (both approximately 2 g/kg DW) (Fig. 4D, E), the
282 proportion that remains soluble is approximately 10-fold higher for Cys than for Met. We found a

This is a provisional file, not the final typeset article

283 strong accumulation of free Cys in EcSAT-containing lines, as expected from literature. However,
284 the level of accumulation was not attenuated by the presence of the strong sink for S-AA in the form
285 of HaSSA protein (Fig. 4E).
286 Glutathione (GSH), a tripeptide of cysteine, glutamic acid, and glycine, is a major metabolic
287 product of cysteine, and is at least an order of magnitude more abundant in mature rice seeds than
288 free cysteine. The two transgenic lines with the highest total cysteine and free cysteine, SAT and
289 SATxSSA, also had the highest GSH levels (Fig. 4F). Interestingly, GSH also accumulated
290 somewhat in AtD-CGS-containing lines, despite no increase in cysteine concentration, suggesting
291 that increased cysteine synthesis occurs in AtD-CGS-containing seeds, but that it is diverted to
292 downstream products such as GSH.
293
294 **2.5 Sulfate levels and sulfur pools**
295 Unexpectedly, the sulfate concentration in mature seeds was found to be approximately four-
296 fold higher in EcSAT-containing seeds than in all other lines (Fig. 5A). SAT and SATxSSA plants
297 also had approximately four to seven-fold higher sulfate levels in their leaves (Fig. 5B, C).
298 Our goal was to generate rice seeds enriched in S-AA. This either requires the seed to import
299 more sulfur from other parts of the plant for assimilation into cysteine and methionine or to
300 redistribute the existing sulfur pools to S-AA, or a combination of these two mechanisms. To assess
301 whether redistribution of sulfur or additional import was driving the significantly elevated S-AA
302 levels in SAT, SATxSSA, and CGSxSSA seeds, we measured the total sulfur content. SAT and
303 SATxSSA seeds showed large increases in total sulfur (56% and 41% respectively). Although more
304 subtle, CGSxSSA seeds also contained elevated total sulfur, and the size of that increase (13%) was
305 comparable in size to the increase in S-AA in this line (Fig. 5D). These data suggest that increased
306 import of sulfur into the seed, rather than redistribution, underlies the increase in total Cys and Met in
307 the seed of SAT, CGSxSSA, and SATxSSA plants. In the case of EcSAT-containing seeds, the
308 increase in total sulfur is substantially greater than the increase in S-AA, indicating that there are
309 other sulfur-compound pools in these seeds that experience an increase. Although SAT and
310 SATxSSA seeds had large relative increases in sulfate and GSH, both major transport forms of
311 sulfur, the magnitude of these pool size changes do not account for the large increase in total sulfur in
312 these seeds. In fact, the *EcSAT* transgene had a larger effect on the size of the "sulfur-other" pool
313 (sulfur not contained in the sulfate, GSH, or S-AA pools) than on sulfate, GSH, or S-AA (Fig. 5E).
314 Resolution of the identity of the differentially accumulating sulfur-containing metabolites in this pool
315 was beyond the scope of our study and remains an interesting open question.
316
317 **2.6 Relative transcript abundance of sulfate transporters, S-assimilatory enzymes and S-**
318 **methylmethionine degradation enzymes**
319 Various S-compounds are known to be transported between tissues in different plant species
320 (Gigolashvili and Kopriva, 2014). One of the most mobile S-compounds is sulfate. Sulfate
321 transporters (SULTR) are a family of proteins that can be classified into four groups based on
322 specifics of their function and localization (Takahashi et al., 2011). The rice genome contains 13
323 SULTRs, the expression of which were determined in milkripe rice seeds in order to assess whether
324 transcriptional regulation of these genes might provide evidence that increased sulfate import
325 contributes to the observed differences in total sulfur in mature seeds (Fig. 6A). Among the 13 rice
326 SULTRs, we found the highest expressed were group 3 and group 4 transporters, which are
327 responsible for intracellular sulfate transport (into plastids and out of vacuoles, respectively). High-
328 affinity, plasma membrane-localized transporters in the SULTR group 1, which could be expected to
329 contribute to sulfate import from source tissues, were notably expressed at much lower levels, with
330 SULTR1;1 Os03g0195800 and SULTR1;2 Os03g0196000 close to the detection limit. This may
331 explain the remarkably low level of sulfate in rice seeds. Three SULTRs were found to be

332 differentially expressed in transgenic lines: SULTR4;1 Os09g0240500, SULTR1;3 Os08g0406400
333 and SULTR1;1 Os03g0195800. While the two group 1 SULTRs are likely plasma membrane-
334 localized and were strongly and significantly induced, their absolute expression remained very low,
335 and neither were upregulated in EcSAT-containing seeds. This suggests that additional S in these
336 seeds is not due to increased transcription of plasma membrane localized transporters. However,
337 these data do not exclude the possibility that SULTR activity may be differentially regulated post-
338 transcriptionally or that sulfate flux into the seed may vary based on the concentration of sulfate in
339 the phloem. In fact, the measured sulfate concentration in leaves from SAT and SATxSSA plants was
340 found to be 6 to 7-fold higher than Taipei (Fig. 5B, C). Notably, we observed strong upregulation of
341 SULTR4;1, but only in SSA and CGSxSSA (Fig. 6A). SULTR4;1 has been shown to be localized to
342 the tonoplast and is likely involved in sulfate efflux from this intracellular storage compartment
343 (Kataoka et al., 2004; Zuber et al., 2010). Sulfate, whether imported across the plasma membrane or
344 released from vacuolar storage, is reduced and assimilated into cysteine or adenosine 3'-phospho 5'-
345 phosphosulfate (PAPS) in several enzymatic steps. Mirroring the differential expression pattern of
346 SULTR4;1, at least one isoform of each enzyme in the sulfate assimilation pathway is
347 transcriptionally upregulated in both SSA and CGSxSSA milkripe seeds: ATP sulfurylase (ATPS)
348 Os03g0743900, adenosine 5'-phosphosulfate reductase 1 (APRL1) Os07g0509800, sulfite reductase
349 (SiR) Os05g0503300, O-acetylserine (thiol) lyase (OASTL) Os12g0625000 (Fig. 6C).
350 Transcriptional upregulation of these enzymes and the vacuolar SULTR in SSA and CGSxSSA seeds
351 could be the result of especially high unmet demand for reduced sulfate in these lines.
352 Another major inter-tissue transport form of sulfur is GSH (Kuzuhara et al., 2000). Two GSH
353 transporters have been identified thus far in rice, OsOPT3 and OsOPT6 (Zhang et al., 2004;
354 Wongkaew et al., 2018). OsOPT6 transcripts were below the detection limit in milkripe seeds, and
355 OsOPT3 transcripts were expressed at an extremely low level (data not shown). Unless their
356 expression strongly increases later in seed maturation, GSH is unlikely to be more than a very minor
357 form of sulfur transported into the seed and cannot explain the differential accumulation of sulfur.
358 Two methionine-derived sulfur-containing metabolites, S-methylmethionine (SMM) and S-
359 adenosylmethionine (SAM), are also potential forms of sulfur imported into the seed. SMM is readily
360 detected in rice seedlings (Menegus et al., 2004), and the importance of SMM in inter-tissue sulfur
361 transport has long been suggested (Bourgis et al., 1999; Tan et al., 2010; Cohen et al., 2017). Thus
362 far there are no transmembrane transporters identified for SMM in plants (Gigolashvili and Kopriva,
363 2014), and therefore their expression pattern in seeds cannot be checked. But if increased SMM
364 import underlies the differential accumulation of sulfur, then one might expect enzymes involved in
365 SMM metabolism to be differentially regulated. Methionine can be generated from SMM and
366 homocysteine in one step via the action of homocysteine S-methyltransferase (HMT). Of the four
367 annotated HMTs in rice, two are strongly upregulated in SSA and CGSxSSA seeds, but are expressed
368 at much lower levels than the major HMT isoform in seeds, HMT2, which showed only very minor
369 expression changes in all our transgenic lines (Fig. 6C). Whether upregulation of the minor isoforms,
370 HMT1 and HMT3, results in higher flux of SMM to methionine in these seeds remains unclear, and
371 the transcript data provide no evidence that differential SMM transport into the seed is relevant for
372 the observed sulfur accumulation. SAM transmembrane transporters are represented by a family of
373 seven genes in rice (GO:0000095). However, transcripts for all seven were expressed at only low
374 levels in milkripe seeds, and none were deemed differentially regulated relative to Taipei (data not
375 shown).
376 Taken together, expression patterns of S-compound transporters do not point to a specific S-
377 compound that is increasingly imported into seeds with elevated S-AA contents. Of course, it
378 remains possible that the import activity of specific S-compound transporters is increased in the
379 absence of mRNA accumulation. Given the very high levels of sulfate in SAT and SATxSSA leaves,

This is a provisional file, not the final typeset article

380 these lines may have higher phloem concentrations of sulfate and thereby increased import of sulfate
381 into the seed.
382
383 **2.7 OAS and OAS-cluster gene expression in milkripe seeds**
384 OAS is often considered an indicator of sulfur deficiency or, alternatively, a signaling
385 molecule responsive to sulfur deficiency. In the seeds of SSA and CGSxSSA lines exists an
386 interesting situation where OAS (Fig. 7A) and many transcriptional markers of sulfur deficiency
387 (Fig. 6) accumulate while the sulfate and total sulfur concentration in these seeds is not reduced
388 relative to Taipei (Fig. 5A, D). Since EcSAT catalyzes the formation of OAS, it was initially
389 surprising that OAS does not accumulate in SAT seeds and that *EcSAT* expression suppresses
390 HaSSA-associated OAS accumulation in the double transgenic seeds.
391 Hubberten et al. 2012 identified a cluster of six genes in Arabidopsis whose expression is
392 highly correlated with OAS accumulation in a variety of conditions, including sulfur-deficiency and
393 oxidative stress (Hubberten et al., 2012). The so-called OAS-cluster in Arabidopsis is composed of
394 APR3, SDI1, SDI2, LSU1, SHM7/MSA1, and ChaC-like. With the exception of SDI1 and SDI2,
395 multiple orthologs of each Arabidopsis gene are annotated in the rice genome. At least one ortholog
396 of each OAS-cluster member was found in our data to be strongly and significantly upregulated in
397 SSA and CGSxSSA seeds but not in SAT or SATxSSA (Fig. 7B), i.e. specifically in those lines with
398 elevated OAS levels (Fig. 7A).
399
400 **3 Discussion**
401 **3.1 The HaSSA-associated seed protein profile may result from insufficient ER folding capacity**
402 **rather than insufficient S-AA provision**
403 We generated so-called "PushxPull" lines by combining a strong S-AA sink with cysteine or
404 methionine biosynthesis transgenes with the idea that increased production of S-AA would allow for
405 high levels of HaSSA accumulation without downregulation of endogenous relatively S-rich SSP.
406 However, the seed protein profile phenotype of SSA was not reverted to control composition by co-
407 expression with either *AtD-CGS* or *EcSAT* (Fig. 2), despite higher total S-AA levels in both double
408 transgenic lines (Fig. 4C). Specifically, total methionine was determined to be 20% higher in
409 CGSxSSA and SATxSSA seeds than in SSA (Fig. 4A), and total cysteine was found to be 30% and
410 60% higher, in CGSxSSA and SATxSSA, respectively, than in SSA (Fig. 4B). These data suggest
411 that the seed protein profile phenotype observed by Hagan et al. (2003) and Islam et al. (2005) in
412 SSA rice seeds was not necessarily the result of insufficient S-AA supply and preferential
413 incorporation into HaSSA.
414 Perhaps the most noticeable change in the SSA seed protein profile is to the glutelin
415 component. Glutelins are referred to in Islam et al. (2005) as being relatively S-AA rich and their
416 decrease in SSA seeds is used as evidence of S-AA redistribution. But with only 0.1-2.4% S-AA, we
417 would instead characterize glutelins as having merely low or intermediate S-AA density, falling on
418 the left side of the S-AA density distribution of SSPs and proteins coded for by the top 1% expressed
419 transcripts in rice seeds (Supplemental Fig. 1). Furthermore, it is not clear from data presented in
420 Hagan et al. (2003) or Islam et al. (2005) whether the total glutelin content (proglutelin plus
421 proteolytically cleaved glutelin subunits) of SSA seeds is altered. Given the lack of clear reduction in
422 total glutelin levels and its low to intermediate S-AA density, we considered the possibility that
423 accumulation of proglutelins and the reduction in glutelin subunits were symptoms of insufficient
424 protein processing in the ER rather than a consequence of S-AA limitation. Direct reduction of the
425 folding capacity of the rice endosperm ER has been achieved by deficiency in a primary chaperone in
426 the ER lumen, PDIL1-1. In PDIL1-1-deficient seeds, proglutelin strongly accumulated and processed
427 glutelin subunits were strongly reduced (Satoh-Cruz et al., 2010). Proglutelin also accumulated in

428 seeds with a genetic block to secretory protein transport out of the ER to the Golgi via COPII vesicles
429 (Tian et al., 2013; Qian et al., 2015).
430 In addition to differential accumulation of proglutelin and glutelin subunits, alpha-globulin levels
431 are reduced in HaSSA-containing seeds. Unlike glutelins, alpha-globulin is S-AA dense, with 5.9%
432 methionine and 4.3% cysteine by length (Supplemental Fig. 1). But reduction in this S-AA rich SSP
433 is also consistent with overloaded ER protein folding capacity due to high *HaSSA* expression. High
434 expression of other foreign proteins and peptides targeted to the ER in rice seeds can produce seed
435 protein profiles very similar to that of SSA, irrespective of the S-AA content of the highly expressed
436 foreign protein. For example, transgenic rice seeds expressing dimers of 42 amino acid beta-amyloid
437 peptides (Ab1-42) have strong accumulation in proglutelins and reductions in alpha-globulin levels
438 despite the transgenic peptide containing zero cysteine residues and having only intermediate
439 methionine density (2.4% by length) (Oono et al., 2010). Transgenic rice expressing seven tandem
440 human T-cell epitopes to Japanese Cedar pollen (7Crp) also show increased proglutelins and
441 decreased alpha-globulin (Takagi et al., 2005). The transgenic peptide contains zero methionines and
442 is 3% cysteine by length. Importantly, the seed protein profile phenotype in 7Crp seeds depends upon
443 the expression strength of the 7Crp peptide. At low expression levels, the seed protein profile looks
444 very similar to untransformed seeds (Takaiwa et al., 2007), while at five-times higher expression,
445 proglutelin accumulated and alpha-globulin was clearly reduced (Takagi et al., 2005). Not only were
446 the seed protein profiles of the beta-amyloid and 7Crp rice seeds similar to SSA, CGSxSSA, and
447 SATxSSA, but a similar set of ER resident chaperones and co-chaperones are upregulated, including
448 BiP1 Os02g0115900, BiP4 Os05g0428600, Hsp70 Os02g0710900, NEF Os09g0512700, GRP96
449 Os06g0716700, CNX Os04g0402100 and CRT Os03g0832200, presumably as a response to
450 increased demand for folding capacity in the ER. Based on the data above, we propose that the seed
451 protein profile phenotype in HaSSA-containing seeds is primarily a non-specific effect of ER protein
452 folding capacity overload, rather than S-AA limitation.
453 Lupin seeds have also been targeted for high HaSSA accumulation (Molvig et al., 1997), and
454 despite similar relative levels of exogenous protein accumulation (5% of total protein in lupin, 7% in
455 rice), the authors did not observe a large change in the seed protein profile. This raises the question,
456 is rice particularly susceptible to ER overloading? Rice has the lowest grain protein content amongst
457 major cereals (Juliano, 1992) which perhaps limits the seed's capacity to synthesize foreign proteins
458 without producing stress in the ER. Rice seeds also have extremely low levels of available sulfur (in
459 the form of sulfate) in comparison to some other crops (Molvig et al., 1997; Chiaiese et al., 2004;
460 Shinmachi et al., 2010) which may prevent meaningful increases in GSH synthesis when there is
461 greater than normal protein folding load in the ER. The precise role of GSH in the ER is still under
462 investigation, but there are strong indications that GSH may be critical for reduction and PDI-
463 mediated isomerization of non-native disulfide bonds (Chakravarthi et al., 2006). Misfolded proteins
464 containing non-native disulfide bonds would be expected to accumulate in situations where the ER
465 protein-folding machinery is under heavy load due to strong expression of ER targeted foreign
466 transgenes such as *HaSSA*. Furthermore, some level of oxidative stress due to H_2O_2 production is
467 intrinsic to disulfide bond formation in the ER (Harding et al., 2003). An overloaded ER that is not
468 able to sufficiently correct and resolve misfolded proteins can result in H_2O_2 production significant
469 enough to alter the glutathione redox state (Ozgur et al., 2014; Dietz et al., 2016) and to influence the
470 cytosolic redox potential (Lai et al., 2018). Therefore, in the context of an overloaded ER protein
471 folding machinery, it might be adaptive to specifically downregulate expression of SSPs with very
472 high S-AA density, such as S-rich 10kDa prolamin (Hagan et al., 2003), so that cysteine could then
473 be repurposed for GSH synthesis.
474
475 **3.2 HaSSA provokes a sulfur deficiency-like molecular phenotype in rice seeds that is**
476 **suppressed by co-expression with *EcSAT* but not *AtD-CGS***

This is a provisional file, not the final typeset article

477 In addition to the seed protein profile phenotype described above, we also found an
478 association between the *HaSSA* transgene and a S-deficiency-like molecular phenotype (Fig. 6, Fig.
479 7), despite no reduction in total sulfur (Fig. 5D). The strongest gene expression changes include
480 upregulation of S-assimilation pathway genes SULTR1;1 Os03g0195800, SULTR4;1
481 Os09g0240500, ATPS 0s03g0743900, APRL1 Os07g0509800, SAT Os05g0533500, and OASTL
482 Os12g0625000. OAS and the OAS-cluster genes (Hubberten et al., 2012) also accumulated
483 specifically in the same set of lines. OAS has previously been shown to accumulate in chickpea seeds
484 expressing high levels of HaSSA protein (Chiaiese et al., 2004). The HaSSA-associated S-deficiency
485 molecular phenotype described here is presumably the result of signaling of especially high S-AA
486 demand for seed protein translation or of ER protein folding overload and moderate oxidative stress,
487 both of which generate a higher need for cysteine.
488 Unlike the seed protein profile phenotype, the main features of the HaSSA-associated S-
489 deficiency molecular phenotype are suppressed by co-expression of *EcSAT* (Fig. 6, Fig. 7). Notably,
490 OAS, the product of SAT enzymatic activity, does not accumulate in the seeds of EcSAT-containing
491 plants (Fig. 7A) despite very strong expression of the transgene and no compensatory reduction in
492 endogenous SAT transcript levels in this tissue (Fig. 1B). While initially surprising, these results are
493 very similar to other published studies that also used a Cys-sensitive isoform of SAT as their
494 transgene. Transgenic Arabidopsis that highly expressed a Cys-sensitive SAT from watermelon also
495 had no OAS accumulation despite high measured SAT activity and increased GSH concentration,
496 suggesting that additional OAS was synthetized but it was readily converted to Cys, which was then
497 incorporated into GSH (Noji and Saito, 2002). Similarly, Hopkins et al. measured OAS levels in
498 *EcSAT* (Cys-sensitive) transgenic potato that were similar to WT (Hopkins et al., 2005). Our EcSAT
499 isoform is cysteine-inhibited and indeed the concentration of free cysteine is significantly higher in
500 SAT and SATxSSA seeds. This suggests that OAS production is essentially buffered by negative
501 feedback at the level of EcSAT activity and additional OAS synthesized by EcSAT is quantitatively
502 consumed by OASTL activity and converted to cysteine. Such complete conversion to cysteine of the
503 additional OAS produced by the same *EcSAT* transgene has already been shown *in vitro* (Harms et
504 al., 2000). OASTL Os030747800, the highest expressed OASTL isoform in milkripe seeds, is
505 moderately yet significantly (padj < 0.0001) upregulated, specifically in SAT-containing seeds,
506 perhaps facilitating the conversion of the additional OAS produced in these transgenic lines.
507 OAS is expected to accumulate in situations where the supply of sulfide is insufficient for the
508 OAS available for formation of cysteine by OASTL. The production of sulfide from sulfate requires
509 the contribution of eight electrons from reductants. As shown in Figure 3, seeds from all three of our
510 HaSSA-containing lines showed signs of ER stress in the form of strong upregulation of UPR genes.
511 Conditions that result in an overloaded protein folding machinery in the ER have been linked to
512 perturbed redox status (Ozgur et al., 2014), and OAS is known to accumulate in situations of
513 oxidative stress, such as menadione treatment (Lehmann et al., 2009). We propose that strong *HaSSA*
514 expression may result in reduced availability of reductant (GSH and ferredoxin) for APR and SiR, a
515 lower sulfide : OAS ratio than in non-HaSSA-containing seeds, and as a consequence, OAS
516 accumulation. Furthermore, we suggest that the *EcSAT* transgene results in a large enough increase in
517 sulfur import into the seed (Fig. 5D) for the sulfate reduction pathway to produce enough sulfide to
518 keep OAS from accumulating in SATxSSA seeds (Fig. 7A). Expression of the *AtD-CGS* transgene
519 also increases S-import into the seed, but not enough to keep the flux to sulfide high enough for the
520 additional S-AA demanded by HaSSA translation and folding. Flux from sulfate to sulfide may also
521 be increased in SATxSSA seeds compared to SSA and CGSxSSA due to additional synthesis of GSH
522 (the electron donor for APR) (Bick et al., 1998) (Fig. 4F) and ferredoxin (the electron donor for SiR)
523 (Krueger and Siegel, 1982) made possible by a larger pool of synthesized cysteine.
524 Given the importance of thiol-based antioxidant systems in mitigating oxidative stress (Ulrich
525 and Jakob, 2019), it is reasonable that sulfur assimilation, and cysteine synthesis in particular, would

526 be enhanced under such conditions. In fact, the activity of both APR and SAT have been shown to
527 increase during oxidative stress. The activity of the Arabidopsis APR1 isoform increases under
528 oxidative conditions, probably via disulfide bond formation (Bick et al., 2001). Soybean plastidic
529 SAT has been shown to be phosphorylated under oxidative stress, resulting in a loss of feedback
530 inhibition by cysteine and increased activity (Liu et al., 2006). Although there is no evidence of SAT
531 regulation by phosphorylation in Arabidopsis, the chloroplast cyclophilin CYP20-3 participates in
532 folding of AtSAT1 in a redox-dependent fashion, to enhance the activity of AtSAT1 and to affect
533 thiol contents (Dominguez-Solis et al., 2008). Additionally, it was recently shown that the
534 availability of (reduced) GSH in Arabidopsis and tobacco positively affects S-assimilation
535 downstream of APR activity, including the steady state concentration of cysteine and methionine
536 (Cohen et al., 2020). The data presented in this study support that redox regulation of sulfate
537 assimilation and cysteine synthesis also occurs in rice.
538
539 **3.3 Improving nutritional quality of seeds**

540 Using a set of S-AA "push", "pull" and "PushxPull" lines, we show that multiple paths can lead
541 to higher seed methionine in rice. First, ubiquitin promoter-driven *EcSAT* is sufficient on its own to
542 produce rice seeds with 50% higher total methionine content (Fig. 4A). This demonstrated that there
543 is an accessible path in rice seeds to higher methionine that does not require increasing the protein
544 sink for methionine. This transgenic line was produced to "push" cysteine synthesis, but
545 unexpectedly, we found that the transgene had the effect of increasing sulfur content in the seed (Fig.
546 5D). As indicated above, rice seeds have very little available sulfate relative to some other cereals
547 (Shinmachi et al., 2010) and grain legumes (Molvig et al., 1997; Chiaiese et al., 2004), and this
548 characteristic may influence the capacity for increased accumulation of S-AA and GSH in the tissue.
549 The relative contribution of increased SAT activity and increased sulfur supply to the resultant
550 improvement in seed protein quality in SAT transgenic seeds remains to be investigated.

551 Second, while ubiquitin promoter-driven *AtD-CGS* and glutenin promoter-driven *HaSSA* are
552 ineffective on their own, in combination they synergistically interact to result in seeds with higher
553 seed methionine (Fig. 4A). Like *EcSAT*, but to a lesser extent, the *AtD-CGS* transgene results in
554 increased seed sulfur content, some of which is incorporated into cysteine (Fig. 5D). AtD-CGS is
555 ineffective on its own because the cysteine available as a substrate for CGS might be tightly
556 regulated in the plastid and flux to GSH might be prioritized over methionine (Fig. 4F). Translation
557 of the large number of *HaSSA* transcripts produced in the developing seed has the potential to "pull"
558 significant quantities of additional methionine into the seed. In fact, the *HaSSA* transgene has been
559 successfully used to increase the methionine content of lupin seeds (Molvig et al., 1997) and
560 chickpeas (Chiaiese et al., 2004). In both cases, total seed sulfur remained nearly constant, but the
561 concentration of sulfate in the seed dropped significantly, suggesting that the additional methionine
562 demanded by HaSSA protein accumulation required increased assimilation of the existing sulfate
563 pool. In rice seed, the concentration of sulfate relative to total sulfur is remarkably low relative to
564 both lupin and chickpea (approximately 1%, 30%, and 20% of total S in rice, lupin, and chickpea,
565 respectively)(Molvig et al., 1997; Chiaiese et al., 2004). Introduction of the new sink HaSSA seems
566 insufficient to induce S-AA biosynthesis in the seed or to trigger sulfate uptake. Further, the load
567 HaSSA puts on the ER for its folding appears to generate significant stress, including oxidative
568 stress. In particular, the high number of Cys residues in HaSSA may amplify the folding load since
569 formation of incorrect disulfide bridges (which need to be resolved and the correct bridges formed)
570 compounds with each additional cysteine in the peptide. Because oxidizing environments promote
571 disulfide bond formation this effect may be especially strong when the ER is experiencing stress
572 since the stress produces ROS. Dissolution of incorrect disulfide bonds oxidizes GSH to GSSG
573 (Chakravarthi et al., 2006), reducing the pool of reductants available for sulfate to sulfide reduction
574 for eventual Cys and Met amino acid synthesis. *AtD-CGS* and *HaSSA* are effective in combination

This is a provisional file, not the final typeset article

575 because they relieve deficiencies in the other. *AtD-CGS*-mediated increase in S-import results in
576 additional flux to Cys, which is further pulled into Met synthesis due to strong sink in the form in
577 HaSSA (Fig. 4A). In the context of this strong sink for Met, the relative prioritization of synthesized
578 Cys for GSH synthesis is weaker. HaSSA-associated decrease in reductant supply for sulfate
579 reduction is complemented by increased supply of S for assimilation to Cys and Met.
580 In contrast to *AtD-CGS*, when *EcSAT* is co-expressed with *HaSSA* it does not synergistically
581 increase methionine content in the seed. In fact, the increase in total methionine and cysteine are
582 nearly equivalent in both SAT and SATxSSA. However, the seed protein profiles in these two lines
583 look very different from each other, with the SATxSSA seeds having the ER stress protein profile
584 phenotype and SAT seeds having a profile very similar to untransformed Taipei (Fig. 2). The
585 additional S-AA in SAT seeds are likely distributed into many seed proteins, or into a few proteins
586 with very high S-AA content that are expressed at low levels, both of which would make perceiving
587 differences by 1D SDS-PAGE difficult. Taken together these results suggest that the total S-AA level
588 is not a reliable determinant of the seed protein profile in rice.
589
590 **3.4 Limited seed sulfate availability/intake/loading limits S-AA related nutritional quality in**
591 **rice**
592 A common feature of all three transgenic lines in this study with increased S-AA content is
593 elevated total sulfur in the seed. This indicates that sulfur loading is a factor that can limit grain
594 cysteine and methionine content in rice and that the seed is able to assimilate additional sulfur into
595 organic molecules including S-AA. However, we were unable to determine what form or forms of
596 sulfur are differentially imported into high S-AA lines. Transcript data from milkripe seeds did not
597 reveal clear candidates based on annotated transporter expression nor expression of enzymes that
598 would be responsible for metabolizing specific imported S-species into major branch point
599 metabolites, such as cysteine. Some indirect evidence points to sulfate as being the major
600 differentially imported form of sulfur. In EcSAT-containing lines sulfate levels were considerably
601 higher than in Taipei in all tissues tested, i.e. leaves from plants 34 and 70 days after germination,
602 and mature seeds (Fig. 5A-C). And it was previously shown that total CGS activity (sum of activity
603 from endogenous isoforms and *AtD-CGS* transgene) was positively correlated with sulfate content in
604 leaves (Whitcomb et al., 2018). Therefore, phloem loading with sulfate may be increased in our S-
605 AA "push" lines. Higher levels of sulfate in the phloem would likely translate into greater sulfate
606 unloading into seed tissues, even in the absence of transcriptional upregulation of well-expressed
607 plasma membrane-localized SULTRs in the seed. There is also evidence that SULTR activity can be
608 regulated both post-transcriptionally and post-translationally (Hopkins et al., 2005; Rouached et al.,
609 2005; Yoshimoto et al., 2007; Shibagaki and Grossman, 2010). OASTL has been shown to inhibit
610 sulfate transport across the plasma membrane by direct interaction with the STAS domain of
611 SULTR1;2 (Shibagaki and Grossman, 2010). OASTL also interacts with SAT, forming the cysteine
612 synthase complex (Hell and Wirtz, 2011), and one could speculate that high EcSAT accumulation
613 could bind-up enough OASTL to reduce OASTL binding with SULTR1;2 and thereby reduce
614 OASTL-mediated inhibition of sulfate transport via SULTR1;2. In future studies, flux analysis with
615 labeled S would be required to confidently identify the S-species that are differentially imported into
616 EcSAT, and to a lesser extent AtD-CGS, containing seeds.
617 In conclusion, the set of five transgenic lines presented here provide insight into the factors
618 that limit cysteine and methionine accumulation in rice seed and suggest different approaches to
619 produce even greater increases than those achieved here. First, seed sulfur loading is strongly
620 implicated as a critical factor, although which transport form of sulfur is differentially imported is
621 still unknown. Additionally, experiments to separate the effect of EcSAT's serine acetyl transferase
622 activity from its effect on seed sulfur loading should be performed in order to design the next
623 iteration of rice transgenic lines. Notably, both *EcSAT* and *AtD-CGS* expression in our transgenic

624 lines are driven by ubiquitin promoters. It would be helpful to know whether seed-specific or leaf – 625 specific expression of these enzymes would be sufficient to increase import of S into the seed. 626 Second, the very high expression level and S-AA density of HaSSA may actually inhibit its 627 effectiveness as a method to "pull" S-AA in to rice seeds. The data here suggest that rice seeds 628 possess insufficient protein folding capacity in the ER for the very high levels of HaSSA 629 accumulation, which results in oxidative stress. Translation of *HaSSA* transcripts may capture enough 630 synthesized cysteine (for translation and for conversion into methionine for translation) to 631 shortchange other uses of cysteine, such as synthesis of GSH and other antioxidants that are 632 particularly important in the context of accumulated unfolded proteins in the ER and altered redox 633 state in other cellular compartments. Therefore, a SSP with more moderate S-AA density and/or 634 expressed at more moderate levels may be more effective for increasing total methionine in rice 635 seeds. For example, 2S albumin from sesame (S2SA) has approximately 3-fold lower S-AA density 636 than HaSSA, but low copy number transformation resulted in a large increase in both total cysteine 637 and methionine in rice seed and no major changes in the endogenous seed protein profile were 638 observed (Lee et al., 2003). We interpret this as a sign that the seeds are experiencing less ER 639 unfolded protein stress than the SSA rice line used here.

640 Improving the nutritive quality of commodity crops remains an important goal. In this study 641 we provided a deeper understanding of the processes underlying rice seed protein accumulation and 642 outlined a basis for overcoming hurdles to further increase protein-bound S-AA content to meet 643 nutritional needs.

644 **4 Materials and methods**

645

646 **4.1 Rice lines**

647

648 The line Taipei, *Oryza sativa* ssp. *japonica* cv. Taipei 309, IRGC accession 42576, served as the 649 untransformed control line for the single and double transgenic lines in this study. The line IR64, 650 *Oryza sativa* ssp. *indica* cv. IR64, served as a high-protein reference line. Generation of the line SSA 651 was previously described in Hagan et al., (2003). Briefly, Taipei was transformed by microparticle 652 bombardment with a transgene construct containing the coding region of the 2S albumin SFA8 from 653 sunflower, *HaSSA*, and the KDEL ER retention sequence driven by Bx17 wheat high molecular 654 weight glutenin promoter. Generation of the line CGS was previously described in Whitcomb et al. 655 (2018), and the line CgSx4 from that study is simply referred to here as line CGS. Briefly, Taipei was 656 transformed by *Agrobacterium tumefaciens* with a transgene construct containing the coding region 657 for a feedback desensitized variant of *Arabidopsis thaliana* cystathionine-gamma-synthase, *AtD-CGS* 658 and the chloroplast-targeting transit peptide from *Pisum sativum*, driven by the maize ubiquitin 1 659 promoter. Generation of the line SAT was previously described in Nguyen et al. (2012), and the line 660 SAT47 from that study is simply referred to here as line SAT. Briefly, Taipei was transformed by 661 *Agrobacterium tumefaciens* with a transgene construct containing the *cysE* gene from *Escherichia* 662 *coli*, *EcSAT*, and the Arabidopsis rbcS 5‘ signal sequence for chloroplastic targeting, driven by the 663 maize ubiquitin 1 promoter. The CGSxSSA and SATxSSA lines were generated by crossing 664 homozygous T2 SAT and CGS plants with the SSA line. In these crosses, SSA served as the pollen 665 donor. F1 plants were screened by PCR for the presence of the *HaSSA* transgene. Double transgenic 666 plants in the F2 generation were screened by PCR for both *AtD-CGS* and *HaSSA* or *EcSAT* and 667 *HaSSA*, and two segregants from each cross were chosen for further analysis. Single seed decent was 668 used to propagate the lines to the F3, F4, or F5 generation depending on the assay (see 4.3 Plant 669 material sampling). Data from the two lines derived from each of the initial PushxPull crosses were 670 analyzed separately and found to have substantively similar molecular and metabolic phenotypes. For 671 visualization and significance testing, data from these lines were combined and presented as simply

This is a provisional file, not the final typeset article

672 CGSxSSA or SATxSSA. PCR primers for screening: AtD-CGS_F 5' agg atc cgt ccg tca gct gag cat
673 taa agc and AtD-CGS_R 5' aaa gct tga tgg ctt cga gag ctt gaa g; EcSAT_F 5' gac gct act caa gca cga
674 aa and EcSAT_R 5' ccc atc ccc ata ctc aaa tg; HaSSA_F 5' atg gca agg ttt tcg atc gt and HaSSA_R 5'
675 att tgg cat ggt tgg gac at.
676
677 **4.2 Growth conditions**
678
679 Seeds were germinated at 28°C in the dark on tap water-soaked paper towels. After 10 days,
680 seedlings were transferred to a growth chamber with the following conditions: 12 h day length with a
681 photon flux density set at 600 µE m-2 s-1 (Lamps: Iwasaki Eye MT 400 DL/BH E40, DHL Licht,
682 Wülfrath, Germany); 26 °C in the light and 22 °C in the dark; relative humidity of 75 % in the light
683 and 70% in the dark. Single seedlings were transplanted into individual pots (18 cm deep, 10 cm
684 diameter) of waterlogged soil after seven days of acclimation to the growth chamber. Soil was a
685 mixture of two parts potting soil (70 % white moss peat, 30 % clay) and one-part sand. One gram of
686 slow-release fertilizer (Plantacote Depot 4 M; Lanxess, Langenfeld, Germany) and 0.1 g Fetrilon
687 Combi (Compo, Münster, Germany) were mixed into the soil in each pot. Pots were kept partially
688 submerged in water until seeds ripened, then all water was withdrawn and the plants allowed to dry
689 out.
690
691 **4.3 Plant material sampling**
692
693 Milkripe and mature seeds were harvested 10 days and 21 days after anthesis, respectively. Unfilled
694 seeds and those with fungus-infected hulls were excluded. Milkripe seeds were flash frozen in liquid
695 nitrogen, and mature seed samples were dried at 50 °C for 2 days prior to storage at -80 °C. Seed
696 samples from double transgenic plants that were used for RNAseq and sulfur-containing metabolite
697 analyses were of the F5 generation and the F4 generation, respectively. The 2nd and 3rd mature leaf
698 on the main tiller were harvested at 34 and 70 days after transplantation and flash frozen. Leaf
699 samples from double transgenic plants were from the F3 generation.
700
701 **4.4 Seed protein extraction and western blotting**
702
703 100 mg finely ground mature seed samples were thawed on ice in 600 µL of protein extraction buffer
704 [62.5 mM Tris-HCL pH 6.8, 2% SDS, 10% glycerol]. The protein concentration of the supernatant
705 after centrifugation was determined by Bradford assay with BSA as a standard. Western blotting for
706 HaSSA was performed with an anti-HaSSA polyclonal antibody kindly donated by Linda Tabe,
707 CSIRO and an anti-sheep HRP-conjugated antibody from Thermo Scientific (Pierce).
708
709 **4.5 Total sulfur content**
710
711 Whole mature seed samples with husks were milled with a Retsch Ultra-Centrifugal Mill ZM200 and
712 dried overnight at 80 °C. 250 mg aliquots of milled and dried sample were digested in 5 mL of trace
713 analysis grade 65% nitric acid:perchloric acid (15:85, v/v) for 8.5 hours at increasing temperatures [2
714 hours at room temperature, 3 hours at 60 °C, 1 hour at 100 °C, 1 hour at 120 °C, 1.5 hours at 175 °C].
715 After cooling, 4 mL of 25% (v/v) nitric acid was added and the tubes were reheated to 80 °C for 1
716 hour. 13 mL of ultra-pure water was added, the solutions mixed well, and then heated again at 80 °C
717 for 30 minutes. After cooling, these solutions were brought up to 20 mL with 5% nitric acid (v/v) and
718 filtered (Whatman, no. 42 from GE Healthcare). Solutions were analyzed by inductively coupled
719 plasma optical emission spectroscopy (ICP-OES) on an Optima 7300 DV ICP-OES (Perkin Elmer

720 LAS Ltd., Seer Green, UK) using appropriate quality control checks, e.g. calibration verification
721 standards, in house standards, and certified reference materials.
722
723 **4.6 Total methionine and cysteine content**
724
725 Dry, mature seed samples of approximately 1 g were were oxidized with ice-cooled performic acid.
726 Oxidation reactions were stopped with hydrobromic acid and then dried in a vacuum rotary
727 evaporator. Residues were resuspended in 6N hydrochloric acid and hydrolysis allowed to proceed
728 for 24 hours at 110 °C. After drying the samples in a vacuum rotary evaporator, the hydrolyzed
729 residue was resuspended in demineralized water. The HPLC method for derivatization, separation,
730 and quantification of methionine and cysteine content was performed essentially as previously
731 described (Algermissen et al., 1989). Briefly, immediately prior to injection into the HPLC, samples
732 were derivatized with ortho-phthaldehyde (OPA) in methanol with addition of borate buffer pH 9.5
733 and mercaptoethanol. HPLC was performed using a Merck LiChrospher® 5 µm RP-18 100 Å, 250 x
734 3 mm ID column connected to an HPLC system from Agilent in combination with a fluorescence
735 detector from Shimadzu. Acetate buffer solutions with different methanol concentrations were used
736 as HPLC eluents for the separation of amino acids by a gradient method.
737
738 **4.7 Glutathione and free cysteine concentration**
739
740 50 mg aliquots of frozen, finely ground mature seed tissue were suspended in 500 µL 100 mM HCl,
741 vortexed, and cleared by centrifugation. 60 µL of supernatant was then reduced by incubation at
742 room temperature with 35 µL of 10 mM DTT in 100 µL of 0.25 M CHES, pH 9.4. The samples
743 where then derivatized for 15 minutes in the dark with 5 µL of 25 mM monobromobimane in
744 acetonitrile. The derivatization was stopped by adding 110 µL of 100 mM methanesulfonic acid and
745 the major soluble thiols were quantified by HPLC, as previously described (Kreft et al., 2003). HPLC
746 was performed using a Eurospher 100-5 C18, Column 250 x 4 mm connected to a Dionex HPLC
747 system, and data were collected and processed with Chromeleon software version 6.8 from Dionex.
748
749 **4.8 Free methionine determination**
750
751 50 mg aliquots of frozen, finely ground mature seed tissue were resuspended in 300 µl ice cold
752 methanol followed by 15 min of shaking at 950 rpm at 70 °C. As a second extraction step, 167 µL of
753 chloroform was added and samples were shaken at 950 rpm for 5 minutes at 37 °C. 333 µL of water
754 was vortexed into the samples, and then they were centrifuged for 5 min at 20,800 g. 100 µL aliquots
755 of the upper polar phase were dried overnight in a vacuum concentrator. The dried polar phase was
756 resuspended in 70 µL of 5 mM sodium phosphate buffer pH 6.2 and subjected to HPLC analysis with
757 pre-column, online ortho-phthaldehyde (OPA) derivatization in combination with fluorescence
758 detection (Lindroth and Mopper, 1979). HPLC was performed using a Phenomenex HyperClone™ 3
759 µm ODS (C18) 120 Å, LC column 150 x 4.6 mm connected to a Dionex HPLC system, and data
760 were processed with Chromeleon software version 6.8 from Dionex.
761
762 **4.9 Sulfate concentration determination**
763
764 Dried polar phase aliquots from methanol: chloroform extraction (see above, free methionine
765 determination) were resuspended in 1 mL ULC/MS grade de-ionized water, vortexed and spun at
766 4 °C for 15 min at full speed in a tabletop microcentrifuge. Sulfate concentration was determined by
767 anion exchange chromatography using an ICS-3000 from Dionex. Data were collected and processed
768 with Chromeleon software version 6.8 from Dionex.

This is a provisional file, not the final typeset article

769
770 **4.10 O-acetyl serine determination**
771
772 Metabolite profiling of milkripe seed tissue by gas chromatography-mass spectrometry (GC-MS) was
773 performed essentially as previously described (Erban et al., 2007; Erban et al., 2019). To enrich for
774 polar primary metabolites and small secondary products, 60 mg samples were extracted with
775 methanol and chloroform as described above for free methionine determination but with the
776 following changes for GC-MS analysis: extraction volumes were scaled to the larger initial sample
777 mass, $^{13}C_6$-sorbitol at 17 mg/L was added to the methanol extraction step to allow later correction of
778 analytical variance, and 160 µL, rather than 100 µL, polar phase aliquots were taken for analysis.
779 Dried polar phase aliquots were chemically derivatized by methoxyamination and trimethylsilylation.
780 A mixture of n-alkanes was added to the derivatized samples to serve as retention index standards.
781 Gas chromatography coupled to electron impact ionization time-of-flight mass spectrometry was
782 performed using an Agilent 6890N24 gas chromatograph and a Pegasus III mass spectrometer from
783 LECO Instruments. ChromaTOF software (LECO, St. Joseph, MI, USA) was used to process the
784 chromatograms. For metabolite identification TagFinder (Luedemann et al., 2008), NIST mass
785 spectral library and search software (https://www.nist.gov/srd/nist-standard-reference-database-1a-
786 v17), and the mass spectral and retention time index reference collection of the Golm Metabolome
787 Database (http://gmd.mpimp-golm.mpg.de/) were used. Data were annotated and curated manually.
788 Mass spectral intensity was normalized to sample fresh weight and $^{13}C_6$-sorbitol. The twice silylated
789 (2TMS) derivative of O-acetylserine (MPIMP ID, A141001; http://gmd.mpimp-
790 golm.mpg.de/Analytes/b4a42a07-e58c-4fb3-83e5-21fc5dc3330a.aspx) was chosen for relative
791 quantification of OAS levels in milkripe seeds.
792
793 **4.11 Differential gene expression analysis**
794
795 Milkripe seed samples were harvested 10 days after flowering on an individual panicle basis. Total
796 RNA was prepared from approximately 100 mg of frozen, finely ground tissue with Ribospin™
797 Seed/Fruit kit (GeneAll Biotechnology, Korea) followed by two rounds of DNase treatment (first
798 round with Riboclear™ plus! from GeneAll, second round with TURBO DNA-*free* Kit from
799 Invitrogen) and PCR was performed to assess possible remaining genomic DNA contamination.
800 Only those total RNA samples with RIN quality scores greater than 7.3 by Bioanalyzer were
801 submitted for RNA sequencing. Library preparation and sequencing were performed at the Max
802 Planck Genome Center, Cologne, Germany (https://mpgc.mpipz.mpg.de/home). rRNA depletion was
803 performed on 1000 ng total RNA samples with the Ribo-Zero™ rRNA Removal Kit (seed/root) from
804 Illumina. Subsequent library preparation was performed with NEBNext® Ultra™ Directional RNA
805 Library Prep Kit for Illumina® (New England Biolabs) according to the manufacturer's instructions.
806 Quality and quantity were assessed at all steps via capillary electrophoresis (TapeStation, Agilent
807 Technologies) and fluorometry (Qubit, Thermo Fisher Scientific). Libraries were immobilized and
808 processed onto a flow cell with cBot (Illumina) and subsequently sequenced on HiSeq3000 system
809 (Illumina) with approximately 50 x 10^6 strand-specific 150 bp single-end reads for each milkripe
810 seed sample. The quality of the raw fastq sequence data was assessed by FastQC
811 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw data were mapped first against the
812 rice genome and then against a set of transgenes. In detail, sequence data were mapped using STAR
813 (Dobin et al., 2013; Dobin and Gingeras, 2015) version 2.7.1a with the parameters --quantMode
814 TranscriptomeSAM GeneCounts, --outSAMtype BAM Unsorted, and --outReadsUnmapped Fastx.
815 Ensembl version 43 (IRGSP 1.0) genome reference in FASTA format
816 (ftp://ftp.ensemblgenomes.org/pub/release-43/plants/fasta/oryza_sativa/dna/Oryza_sativa.IRGSP-
817 1.0.dna.toplevel.fa.gz) and Ensembl version 43 cDNA Annotation in GTF format

818 (ftp://ftp.ensemblgenomes.org/pub/release-43/plants/gtf/oryza_sativa/Oryza_sativa.IRGSP-
819 1.0.43.gtf.gz) were used for genome indexing (--sjdbOverhang 149 --genomeSAindexNbases 13). In
820 order to determine RNA read counts for the *AtD-CGS*, *EcSAT*, and *HaSSA* transgenes, reads
821 unmapped to the rice genome were used as input for mapping against a fasta file containing the
822 sequences X56686.1, M15745.1, and NM_110977.3, downloaded from NCBI), which was indexed
823 with bwa version 7.17 in "is" mode. Reads were mapped against it in backtrack ("aln") mode (Li and
824 Durbin, 2009) and converted to sam format with bwa samse. Afterwards samtools version 1.8 was
825 used to convert sam to bam format (samtools view -bS), sort (samtools sort), and index (samtools
826 index) the resulting bam alignment files in order to create read counts per input sequence (samtools
827 idxstats) (Li et al., 2009). Anti-strand read counts from the ReadsPerGene files of all 22 samples
828 were merged in order to perform differential expression analysis. Reads that could be unambiguously
829 mapped to either a rice gene or transgene were retained for further analysis in R version 3.6.1 with
830 the package DESeq2 (Love et al., 2014) version 1.24.0. Differences in sequencing depth between
831 samples were accounted for in DESeq2 with median-of-ratio normalization. Differential expression
832 analysis was performed using dispersion and log fold change estimates after shrinkage with the
833 DESeq2 *maximum a posteriori* option. To determine the significance level of estimated Log_2-fold
834 change, Wald tests were performed, and genes were considered differentially expressed relative to
835 Taipei if the Benjamini-Hochberg adjusted p-value was less than 0.01. Genes were considered
836 “expressed” in milkripe seeds if their average normalized counts were greater than 100. The
837 empirical cumulative distribution function ecdf from the R stats package version 3.6.1 was used to
838 calculate the expression strength percentile among all “expressed” genes in milkripe seeds.
839
840 **4.12 Data visualization**
841
842 Quantitative data visualizations were prepared in R version 3.6.1. Univariate scatter plots and the
843 stacked bar plot were generated with the R package ggplot2 version 3.2.1. Significance testing results
844 were added to the plots with the R package ggpubr version 0.2.3. Heatmaps were prepared with the R
845 packages ComplexHeatmap version 2.0.0 and circlize version 0.4.7.
846

847 **5 Conflict of Interest**
848 *The authors declare that the research was conducted in the absence of any commercial or financial*
849 *relationships that could be construed as a potential conflict of interest.*

850 **6 Author Contributions**
851 RH and SW contributed to the conception and design of the study. SW, AR, and FB grew all plants,
852 harvested samples, and processed them for analytical measurements. FB performed the crosses and
853 selected transgenic lines. SP and MH performed the ICP-OES analysis. AE, and JK performed
854 analysis of the GC-MS data. SW and AF performed the RNA-seq analysis. SW performed the
855 statistical analysis and prepared all visualizations. SW took the lead in interpreting the data and wrote
856 the first draft of the manuscript. All authors contributed to manuscript revision and approved the
857 submitted version.
858

859 **7 Funding**
860 The primary funding sources for this study were the Max Planck Society (Germany) and the EU
861 RTN BIONUT (BIOchemical and genetic dissection of control of plant mineral NUTrition) FP7:
862 264296. Rothamsted Research receives grant-aided support from the Biotechnology and Biological

This is a provisional file, not the final typeset article

Sciences Research Council (BBSRC) through the Designing Future Wheat programme [BB/P016855/1].

8 Acknowledgments

We are grateful to Ines Fehrle in the Applied Metabolome Analysis Infrastructure Group at the Max Planck Institute for Molecular Plant Physiology (Germany) for performing the GC-MS measurements. We thank Linda Tabe at CSIRO (Australia) for kindly providing Taipei and SSA seeds and the anti-HaSSA antibody used in this study. We thank the Max Planck-Genome-centre Cologne (http://mpgc.mpipz.mpg.de/home/) for performing the RNA sequencing in this study and the Institut für Getreideverarbeitung GmbH, Nuthetal, Germany for performing the total methionine and cysteine determination.

9 Supplementary Material

Supplementary Figure 1: Distribution of methionine, cysteine, and sulfur amino acid density in seed proteins.
Supplementary Table 1: Significance testing results for genes presented in figures 1, 3, 6, and 7.

10 Data Availability Statement

The RNA-seq dataset generated for this study has been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO series accession number GSE149252 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE149252)

884 **Figure Legends**

885

886 **Figure 1: Expression of CGS, SAT, and SSA transcripts and accumulation of SSA protein in**
887 **seeds**
888 Normalized RNAseq read counts in milkripe seeds for **(A)** the *AtD-CGS* transgene and rice *CGS*
889 isoforms, **(B)** the *EcSAT* transgene and rice *SAT* isoforms, and **(C)** the *HaSSA* transgene are plotted
890 on a $\log_{10}$ y-axis to facilitate comparison of expression levels between transgenes and endogenous
891 isoforms. Each symbol in the plotting area **(A-C)** represents the reads from the milkripe stage seeds
892 of an individual plant. Biological replicates: Taipei n = 4, SSA n = 3, CGS n = 5, SAT n = 3,
893 CGSxSSA n = 4, SATxSSA n = 3. Significance testing is presented in Supplemental Table 1. **(D)**
894 The SDS-soluble protein fraction from mature seeds was subjected to 15% acrylamide SDS-PAGE
895 followed by western blotting with a polyclonal anti-SSA antibody. Each lane contains 20µg of
896 protein extracted from the seeds of an individual plant. The western blot was performed at least 3
897 times. The data shown are representative.

898

899 **Figure 2: Seed protein profiles of mature seeds**
900 The SDS-soluble protein fraction from mature seeds was subjected to 13% acrylamide SDS-PAGE
901 followed by Coomassie staining. Each lane contains 10 µg of protein extracted from the seeds of an
902 individual plant. Protein band labeling is based on previously published data (Hagan et al., 2003;
903 Yasuda et al., 2009). The protein extraction and SDS-PAGE were performed 3 times. The data
904 shown are representative.

905

906 **Figure 3: Differential expression of ER chaperones, co-chaperones, and protein disulfide**
907 **isomerases in milkripe seeds**
908 Differential expression data are presented in a heatmap matrix of Log_2-fold change (Log2FC) relative
909 to Taipei. The mean normalized read count in Taipei for each gene is presented in a bar graph
910 annotation of the heatmap. Those ER resident chaperones (BiP, Hsp70, GRP94, CNX, CRT), co-
911 chaperones (NEF, DjB, DjC), and protein disulfide isomerase (PDIL, ERO1) genes with a mean
912 normalized read count > 1000 in at least one line are shown. Biological replicates: Taipei n = 4, SSA
913 n = 3, CGS n = 5, SAT n = 3, CGSxSSA n = 4, SATxSSA n = 3. Significance testing is presented in
914 Supplemental Table 1.

915

916 **Figure 4: Methionine, cysteine, and glutathione levels in mature seeds**
917 The concentrations of the sulfur-containing amino acids methionine **(A, D)** and cysteine **(B, E)** as
918 well as the cysteine-containing tripeptide glutathione **(F)** were determined in mature seeds. The
919 concentration of total sulfur-amino acids (S-AA) was determined by adding the total concentration of
920 methionine and cysteine in the same seed sample **(C)**. Each black dot indicates the measured
921 concentration in the seeds from an individual plant. The mean concentration in each line is indicated
922 by a red diamond. The data presented are from two independent experiments. Wilcoxon rank-sum
923 test was performed to compare the median concentration in each line to the median of Taipei. The
924 significance level of each test is designated: ns (not significant) for $p > 0.05$, * for $p \leq 0.05$, ** for p
925 ≤ 0.01, *** for $p \leq 0.001$, and **** for $p \leq 0.0001$

926

927 **Figure 5: Sulfur pools**
928 The concentration of sulfate was determined in **(A)** mature seeds and leaves from plants **(B)** 70 days
929 (d70) after transplantation and **(C)** 34 days (d34) after transplantation. The total sulfur (S) in mature
930 seeds was also determined **(D)**. Each black dot indicates the measured concentration in the specified
931 tissue from an individual plant. The mean concentration in each line is indicated by a red diamond.
932 The data presented are from two independent experiments. Wilcoxon rank-sum test was performed to

This is a provisional file, not the final typeset article

933 compare the median concentration in each line to the median of Taipei. The significance level of each
934 test is designated: ns (not significant) for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, *** for $p \leq 0.001$,
935 and **** for $p \leq 0.0001$. For those mature seed samples for which total S, sulfate, GSH, total Met,
936 and total Cys were determined, the mean concentration of each of these sulfur pools was calculated
937 for each line. The S-other pool is the remainder of the sulfate, GSH, total Met, and total Cys
938 concentrations subtracted from the total S concentration in each sample. The absolute and relative
939 size of sulfur pools in mature seeds are visualized in a stacked bar graph **(E)**. Biological replicates for
940 the S pools analyses: n = 12 for all lines except SATxSSA, for which n = 28.
941
942 **Figure 6: Differential expression of genes for sulfate transporters, S-assimilatory enzymes, and**
943 **SMM degradation enzymes in milkripe seeds**
944 Differential expression data are presented as heatmap matrices of Log_2-fold change (Log2FC)
945 relative to Taipei. The mean normalized read count in Taipei for each gene is presented in a bar graph
946 annotation of the heatmap. All 13 genes annotated as sulfate transporters (SULTR) in rice **(A)** and all
947 4 genes annotated as homocysteine methyltransferases (HMT) in rice **(C)** are shown. Annotated ATP
948 sulfurylase, APR reductase, APR kinase, sulfite reductase, and OAS-thiol-lyase genes with a mean
949 normalized read count > 500 in at least one line are shown **(B)**. Biological replicates: Taipei n = 4,
950 SSA n = 3, CGS n = 5, SAT n = 3, CGSxSSA n = 4, SATxSSA n = 3. Significance testing is
951 presented in Supplemental Table 1.
952
953 **Figure 7: OAS accumulation and differential expression of OAS-cluster genes in milkripe seeds**
954 **(A)** The relative abundance (arbitrary units) of O-acetylserine (OAS) in milkripe seeds is presented.
955 Each black dot indicates the determined relative abundance in the seeds of independent individual
956 plants (n = 4). The mean abundance in each line is indicated by a red diamond. Wilcoxon rank-sum
957 test was performed to compare the median accumulation in each line to the median of Taipei. The
958 significance level of each test is designated: ns (not significant) for $p > 0.05$, * for $p \leq 0.05$. **(B)**
959 Differential expression data for the 15 putative orthologs in rice of the 7 so-called "OAS-cluster"
960 genes in *Arabidopsis thaliana* are presented as a heatmap of Log_2-fold change (Log2FC) relative to
961 Taipei with genes clustered hierarchically using Euclidean distance metric and average linkage.
962 Biological replicates: Taipei n = 4, SSA n = 3, CGS n = 5, SAT n = 3, CGSxSSA n = 4, SATxSSA n
963 = 3. Significance testing is presented in Supplemental Table 1.
964
965
966 **References**
967 Algermissen, B., Nündel, M., and Riedel, E. (1989). Analytik von Aminosäuren mit Fluoreszenz-
968 HPLC. *GIT Labor-Fachzeitschrift* 33(9)**,** 783-790.
969 Amir, R., Cohen, H., and Hacham, Y. (2019). Revisiting the attempts to fortify methionine contents
970 in plant seeds. *J Exp Bot*. doi: 10.1093/jxb/erz134.
971 Amir, R., Galili, G., and Cohen, H. (2018). The metabolic roles of free amino acids during seed
972 development. *Plant Sci* 275**,** 11-18. doi: 10.1016/j.plantsci.2018.06.011.
973 Amira, G., Ifat, M., Tal, A., Hana, B., Shmuel, G., and Rachel, A. (2005). Soluble methionine
974 enhances accumulation of a 15 kDa zein, a methionine-rich storage protein, in transgenic
975 alfalfa but not in transgenic tobacco plants. *J Exp Bot* 56(419)**,** 2443-2452. doi:
976 10.1093/jxb/eri237.
977 Bick, J.A., Aslund, F., Chen, Y., and Leustek, T. (1998). Glutaredoxin function for the carboxyl-
978 terminal domain of the plant-type 5'-adenylylsulfate reductase. *Proc Natl Acad Sci U S A*
979 95(14)**,** 8404-8409. doi: 10.1073/pnas.95.14.8404.

980 Bick, J.A., Setterdahl, A.T., Knaff, D.B., Chen, Y., Pitcher, L.H., Zilinskas, B.A., et al. (2001).
981 Regulation of the plant-type 5'-adenylyl sulfate reductase by oxidative stress. *Biochemistry*
982 40(30)**,** 9040-9048. doi: 10.1021/bi010518v.
983 Bourgis, F., Roje, S., Nuccio, M.L., Fisher, D.B., Tarczynski, M.C., Li, C., et al. (1999). S-
984 methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel
985 type of methyltransferase. *Plant Cell* 11(8)**,** 1485-1498. doi: 10.1105/tpc.11.8.1485
986 Chakravarthi, S., Jessop, C.E., and Bulleid, N.J. (2006). The role of glutathione in disulphide bond
987 formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep* 7(3)**,** 271-275.
988 doi: 10.1038/sj.embor.7400645.
989 Chiaiese, P., Ohkama-Ohtsu, N., Molvig, L., Godfree, R., Dove, H., Hocart, C., et al. (2004). Sulphur
990 and nitrogen nutrition influence the response of chickpea seeds to an added, transgenic sink
991 for organic sulphur. *J Exp Bot* 55(404)**,** 1889-1901. doi: 10.1093/jxb/erh198.
992 Cohen, A., Hacham, Y., Welfe, Y., Khatib, S., Avice, J.-C., and Amir, R. (2020). Evidence of a
993 significant role of glutathione reductase in the sulfur assimilation pathway. *The Plant Journal*.
994 doi: doi: 10.1111/tpj.14621.
995 Cohen, H., Salmon, A., Tietel, Z., Hacham, Y., and Amir, R. (2017). The relative contribution of
996 genes operating in the S-methylmethionine cycle to methionine metabolism in Arabidopsis
997 seeds. *Plant Cell Rep* 36(5)**,** 731-743. doi: 10.1007/s00299-017-2124-1.
998 Dancs, G., Kondrak, M., and Banfalvi, Z. (2008). The effects of enhanced methionine synthesis on
999 amino acid and anthocyanin content of potato tubers. *Bmc Plant Biology* 8. doi: Artn 65
1000 10.1186/1471-2229-8-65.
1001 De Wilde, K., De Buck, S., Vanneste, K., and Depicker, A. (2013). Recombinant antibody
1002 production in Arabidopsis seeds triggers an unfolded protein response. *Plant Physiol* 161(2)**,**
1003 1021-1033. doi: 10.1104/pp.112.209718.
1004 Demidov, D., Horstmann, C., Meixner, M., Pickardt, T., Saalbach, I., Galili, G., et al. (2003).
1005 Additive effects of the feed-back insensitive bacterial aspartate kinase and the Brazil nut 2S
1006 albumin on the methionine content of transgenic narbon bean (Vicia narbonensis L.).
1007 *Molecular Breeding* 11(3)**,** 187-201. doi: 10.1023/a:1022814506153.
1008 Dietz, K.J., Turkan, I., and Krieger-Liszkay, A. (2016). Redox- and Reactive Oxygen Species-
1009 Dependent Signaling into and out of the Photosynthesizing Chloroplast. *Plant Physiol* 171(3)**,**
1010 1541-1550. doi: 10.1104/pp.16.00375.
1011 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR:
1012 ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1)**,** 15-21. doi:
1013 10.1093/bioinformatics/bts635.
1014 Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. *Curr Protoc*
1015 *Bioinformatics* 51**,** 11 14 11-19. doi: 10.1002/0471250953.bi1114s51.
1016 Dominguez-Solis, J.R., He, Z., Lima, A., Ting, J., Buchanan, B.B., and Luan, S. (2008). A
1017 cyclophilin links redox and light signals to cysteine biosynthesis and stress responses in
1018 chloroplasts. *Proc Natl Acad Sci U S A* 105(42)**,** 16386-16391. doi:
1019 10.1073/pnas.0808204105.
1020 Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene
1021 expression and hybridization array data repository. *Nucleic Acids Res* 30(1)**,** 207-210. doi:
1022 10.1093/nar/30.1.207.
1023 Erban, A., Fehrle, I., Martinez-Seidel, F., Brigante, F., Mas, A.L., Baroni, V., et al. (2019). Discovery
1024 of food identity markers by metabolomics and machine learning technology. *Sci Rep* 9(1)**,**
1025 9697. doi: 10.1038/s41598-019-46113-y.
1026 Erban, A., Schauer, N., Fernie, A.R., and Kopka, J. (2007). Nonsupervised construction and
1027 application of mass spectral and retention time index libraries from time-of-flight gas

This is a provisional file, not the final typeset article

1028 chromatography-mass spectrometry metabolite profiles. *Methods Mol Biol* 358**,** 19-38. doi:
1029 10.1007/978-1-59745-244-1_2.
1030 Gigolashvili, T., and Kopriva, S. (2014). Transporters in plant sulfur metabolism. *Front Plant Sci* 5**,**
1031 442. doi: 10.3389/fpls.2014.00442.
1032 Global Rice Science Partnership (2013). "Rice Almanac". 4th ed. (Los Baños, Philippines:
1033 International Rice Research Institute).
1034 Hacham, Y., Matityahu, I., Schuster, G., and Amir, R. (2008). Overexpression of mutated forms of
1035 aspartate kinase and cystathionine gamma-synthase in tobacco leaves resulted in the high
1036 accumulation of methionine and threonine. *Plant Journal* 54(2)**,** 260-271. doi:
1037 10.1111/j.1365-313X.2008.03415.x.
1038 Hagan, N.D., Upadhyaya, N., Tabe, L.M., and Higgins, T.J. (2003). The redistribution of protein
1039 sulfur in transgenic rice expressing a gene for a foreign, sulfur-rich protein. *Plant Journal*
1040 34(1)**,** 1-11. doi: 10.1046/j.1365-313X.2003.01699.x.
1041 Hanafy, M.S., Rahman, S.M., Nakamoto, Y., Fujiwara, T., Naito, S., Wakasa, K., et al. (2013).
1042 Differential response of methionine metabolism in two grain legumes, soybean and azuki
1043 bean, expressing a mutated form of Arabidopsis cystathionine gamma-synthase. *J Plant*
1044 *Physiol* 170(3)**,** 338-345. doi: 10.1016/j.jplph.2012.10.018.
1045 Harding, H.P., Zhang, Y., Zeng, H., Novoa, I., Lu, P.D., Calfon, M., et al. (2003). An integrated
1046 stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell*
1047 11(3)**,** 619-633. doi: 10.1016/s1097-2765(03)00105-9.
1048 Harms, K., von Ballmoos, P., Brunold, C., Hofgen, R., and Hesse, H. (2000). Expression of a
1049 bacterial serine acetyltransferase in transgenic potato plants leads to increased levels of
1050 cysteine and glutathione. *Plant Journal* 22(4)**,** 335-343. doi: 10.1046/j.1365-
1051 313x.2000.00743.x.
1052 Hell, R., and Wirtz, M. (2011). Molecular Biology, Biochemistry and Cellular Physiology of
1053 Cysteine Metabolism in Arabidopsis thaliana. *Arabidopsis Book* 9**,** e0154. doi:
1054 10.1199/tab.0154.
1055 Hesse, H., Kreft, O., Maimann, S., Zeh, M., and Hoefgen, R. (2004). Current understanding of the
1056 regulation of methionine biosynthesis in plants. *J Exp Bot* 55(404)**,** 1799-1808. doi:
1057 10.1093/jxb/erh139.
1058 Hopkins, L., Parmar, S., Blaszczyk, A., Hesse, H., Hoefgen, R., and Hawkesford, M.J. (2005). O-
1059 acetylserine and the regulation of expression of genes encoding components for sulfate uptake
1060 and assimilation in potato. *Plant Physiol* 138(1)**,** 433-440. doi: 10.1104/pp.104.057521.
1061 Hubberten, H.M., Klie, S., Caldana, C., Degenkolbe, T., Willmitzer, L., and Hoefgen, R. (2012).
1062 Additional role of O-acetylserine as a sulfur status-independent regulator during plant growth.
1063 *Plant Journal* 70(4)**,** 666-677. doi: 10.1111/j.1365-313X.2012.04905.x.
1064 Islam, N., Upadhyaya, N.M., Campbell, P.M., Akhurst, R., Hagan, N., and Higgins, T.J. (2005).
1065 Decreased accumulation of glutelin types in rice grains constitutively expressing a sunflower
1066 seed albumin gene. *Phytochemistry* 66(21)**,** 2534-2539. doi:
1067 10.1016/j.phytochem.2005.09.002.
1068 Juliano, B.O. (1992). Structure, chemistry, and function of the rice grain and its fractions. *Cereal*
1069 *Foods World* 37**,** 772-774.
1070 Karchi, H., Shaul, O., and Galili, G. (1994). Lysine synthesis and catabolism are coordinately
1071 regulated during tobacco seed development. *Proc Natl Acad Sci U S A* 91(7)**,** 2577-2581.
1072 Kataoka, T., Watanabe-Takahashi, A., Hayashi, N., Ohnishi, M., Mimura, T., Buchner, P., et al.
1073 (2004). Vacuolar sulfate transporters are essential determinants controlling internal
1074 distribution of sulfate in Arabidopsis. *Plant Cell* 16(10)**,** 2693-2704. doi:
1075 10.1105/tpc.104.023960.

1076 Kim, W.S., Jez, J.M., and Krishnan, H.B. (2014). Effects of proteome rebalancing and sulfur
1077 nutrition on the accumulation of methionine rich delta-zein in transgenic soybeans. *Front*
1078 *Plant Sci* 5**,** 633. doi: 10.3389/fpls.2014.00633.
1079 Kortt, A.A., Caldwell, J.B., Lilley, G.G., and Higgins, T.J. (1991). Amino acid and cDNA sequences
1080 of a methionine-rich 2S protein from sunflower seed (Helianthus annuus L.). *Eur J Biochem*
1081 195(2)**,** 329-334. doi: 10.1111/j.1432-1033.1991.tb15710.x.
1082 Kreft, O., Hoefgen, R., and Hesse, H. (2003). Functional analysis of cystathionine gamma-synthase
1083 in genetically engineered potato plants. *Plant Physiology* 131(4)**,** 1843-1854. doi:
1084 10.1104/pp.102.015933.
1085 Krueger, R.J., and Siegel, L.M. (1982). Spinach siroheme enzymes: Isolation and characterization of
1086 ferredoxin-sulfite reductase and comparison of properties with ferredoxin-nitrite reductase.
1087 *Biochemistry* 21(12)**,** 2892-2904. doi: 10.1021/bi00541a014.
1088 Kuzuhara, Y., Isobe, A., Awazuhara, M., Fujiwara, T., and Hayashi, H. (2000). Glutathione levels in
1089 phloem sap of rice plants under sulfur deficient conditions. *Soil Science and Plant Nutrition*
1090 46(1)**,** 265-270. doi: Doi 10.1080/00380768.2000.10408782.
1091 Lai, Y.S., Renna, L., Yarema, J., Ruberti, C., He, S.Y., and Brandizzi, F. (2018). Salicylic acid-
1092 independent role of NPR1 is required for protection from proteotoxic stress in the plant
1093 endoplasmic reticulum. *Proc Natl Acad Sci U S A* 115(22)**,** E5203-E5212. doi:
1094 10.1073/pnas.1802254115.
1095 Lee, T.T., Wang, M.M., Hou, R.C., Chen, L.J., Su, R.C., Wang, C.S., et al. (2003). Enhanced
1096 methionine and cysteine levels in transgenic rice seeds by the accumulation of sesame 2S
1097 albumin. *Biosci Biotechnol Biochem* 67(8)**,** 1699-1705. doi: 10.1271/bbb.67.1699.
1098 Lehmann, M., Schwarzlander, M., Obata, T., Sirikantaramas, S., Burow, M., Olsen, C.E., et al.
1099 (2009). The metabolic response of Arabidopsis roots to oxidative stress is distinct from that of
1100 heterotrophic cells in culture and highlights a complex relationship between the levels of
1101 transcripts, metabolites, and flux. *Mol Plant* 2(3)**,** 390-406. doi: 10.1093/mp/ssn080.
1102 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
1103 transform. *Bioinformatics* 25(14)**,** 1754-1760. doi: 10.1093/bioinformatics/btp324.
1104 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence
1105 Alignment/Map format and SAMtools. *Bioinformatics* 25(16)**,** 2078-2079. doi:
1106 10.1093/bioinformatics/btp352.
1107 Lindroth, P., and Mopper, K. (1979). High-Performance Liquid-Chromatographic Determination of
1108 Subpicomole Amounts of Amino-Acids by Precolumn Fluorescence Derivatization with
1109 Ortho-Phthaldialdehyde. *Analytical Chemistry* 51(11)**,** 1667-1674. doi:
1110 10.1021/ac50047a019.
1111 Liu, F., Yoo, B.C., Lee, J.Y., Pan, W., and Harmon, A.C. (2006). Calcium-regulated phosphorylation
1112 of soybean serine acetyltransferase in response to oxidative stress. *J Biol Chem* 281(37)**,**
1113 27405-27415. doi: 10.1074/jbc.M604548200.
1114 Liu, J.X., and Howell, S.H. (2016). Managing the protein folding demands in the endoplasmic
1115 reticulum of plants. *New Phytol* 211(2)**,** 418-428. doi: 10.1111/nph.13915.
1116 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion
1117 for RNA-seq data with DESeq2. *Genome Biol* 15(12)**,** 550. doi: 10.1186/s13059-014-0550-8.
1118 Luedemann, A., Strassburg, K., Erban, A., and Kopka, J. (2008). TagFinder for the quantitative
1119 analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling
1120 experiments. *Bioinformatics* 24(5)**,** 732-737. doi: 10.1093/bioinformatics/btn023.
1121 Matityahu, I., Godo, I., Hacham, Y., and Amir, R. (2013). Tobacco seeds expressing feedback-
1122 insensitive cystathionine gamma-synthase exhibit elevated content of methionine and altered
1123 primary metabolic profile. *Bmc Plant Biology* 13. doi: Artn 206
1124 10.1186/1471-2229-13-206.

This is a provisional file, not the final typeset article

1125 McNabb, W.C., Spencer, D., Higgins, T.J., and Barry, T.N. (1994). In-vitro rates of rumen
1126 proteolysis of ribulose-1,5-bisphosphate carboxylase (rubisco) from lucerne leaves, and of
1127 ovalbumin, vicilin and sunflower albumin 8 storage proteins. *Journal of the Science of Food*
1128 *and Agriculture* 64(1)**,** 53-61. doi: 10.1002/jsfa.2740640109.
1129 Menegus, F., Lilliu, I., Brambilla, I., Bonfa, M., and Scaglioni, L. (2004). Unusual accumulation of
1130 S-methylmethionine in aerobic-etiolated and in anoxic rice seedlings: an 1H-NMR study. *J*
1131 *Plant Physiol* 161(6)**,** 725-732. doi: 10.1078/0176-1617-01081.
1132 Molvig, L., Tabe, L.M., Eggum, B.O., Moore, A.E., Craig, S., Spencer, D., et al. (1997). Enhanced
1133 methionine levels and increased nutritive value of seeds of transgenic lupins (Lupinus
1134 angustifolius L.) expressing a sunflower seed albumin gene. *PNAS* 94(16)**,** 8393-8398. doi:
1135 10.1073/pnas.94.16.8393.
1136 Nguyen, H.C., Hoefgen, R., and Hesse, H. (2012). Improving the nutritive value of rice seeds:
1137 elevation of cysteine and methionine contents in rice plants by ectopic expression of a
1138 bacterial serine acetyltransferase. *Journal of Experimental Botany* 63(16)**,** 5991-6001. doi:
1139 10.1093/jxb/ers253.
1140 Noji, M., and Saito, K. (2002). Molecular and biochemical analysis of serine acetyltransferase and
1141 cysteine synthase towards sulfur metabolic engineering in plants. *Amino Acids* 22(3)**,** 231-
1142 243. doi: 10.1007/s007260200011.
1143 Oono, Y., Wakasa, Y., Hirose, S., Yang, L., Sakuta, C., and Takaiwa, F. (2010). Analysis of ER
1144 stress in developing rice endosperm accumulating beta-amyloid peptide. *Plant Biotechnol J*
1145 8(6)**,** 691-718. doi: 10.1111/j.1467-7652.2010.00502.x.
1146 Ozgur, R., Turkan, I., Uzilday, B., and Sekmen, A.H. (2014). Endoplasmic reticulum stress triggers
1147 ROS signalling, changes the redox state, and regulates the antioxidant defence of Arabidopsis
1148 thaliana. *J Exp Bot* 65(5)**,** 1377-1390. doi: 10.1093/jxb/eru034.
1149 Planta, J., Xiang, X., Leustek, T., and Messing, J. (2017). Engineering sulfur storage in maize seed
1150 proteins without apparent yield loss. *Proc Natl Acad Sci U S A* 114(43)**,** 11386-11391. doi:
1151 10.1073/pnas.1714805114.
1152 Qi, K., and Lupton, C.J. (1994). A Review of the Effects of Sulfur Nutrition on Wool Production and
1153 Quality. *Sheep & Goat Research Journal* 10(2)**,** 133-140.
1154 Qian, D., Tian, L., and Qu, L. (2015). Proteomic analysis of endoplasmic reticulum stress responses
1155 in rice seeds. *Sci Rep* 5**,** 14255. doi: 10.1038/srep14255.
1156 Rouached, H., Berthomieu, P., El Kassis, E., Cathala, N., Catherinot, V., Labesse, G., et al. (2005).
1157 Structural and functional analysis of the C-terminal STAS (sulfate transporter and anti-sigma
1158 antagonist) domain of the Arabidopsis thaliana sulfate transporter SULTR1.2. *J Biol Chem*
1159 280(16)**,** 15976-15983. doi: 10.1074/jbc.M501635200.
1160 Satoh-Cruz, M., Crofts, A.J., Takemoto-Kuno, Y., Sugino, A., Washida, H., Crofts, N., et al. (2010).
1161 Protein disulfide isomerase like 1-1 participates in the maturation of proglutelin within the
1162 endoplasmic reticulum in rice endosperm. *Plant Cell Physiol* 51(9)**,** 1581-1593. doi:
1163 10.1093/pcp/pcq098.
1164 Shibagaki, N., and Grossman, A.R. (2010). Binding of cysteine synthase to the STAS domain of
1165 sulfate transporter and its regulatory consequences. *J Biol Chem* 285(32)**,** 25094-25102. doi:
1166 10.1074/jbc.M110.126888.
1167 Shinmachi, F., Buchner, P., Stroud, J.L., Parmar, S., Zhao, F.J., McGrath, S.P., et al. (2010).
1168 Influence of sulfur deficiency on the expression of specific sulfate transporters and the
1169 distribution of sulfur, selenium, and molybdenum in wheat. *Plant Physiol* 153(1)**,** 327-336.
1170 doi: 10.1104/pp.110.153759.
1171 Song, S.K., Hou, W.S., Godo, I., Wu, C.X., Yu, Y., Matityahu, I., et al. (2013). Soybean seeds
1172 expressing feedback-insensitive cystathionine -synthase exhibit a higher content of
1173 methionine. *Journal of Experimental Botany* 64(7)**,** 1917-1926. doi: 10.1093/jxb/ert053.

1174 Spencer, D., Higgins, T.J., Freer, M., Dove, H., and Coombe, J.B. (1988). Monitoring the fate of
1175 dietary proteins in rumen fluid using gel electrophoresis. *Br J Nutr* 60(2)**,** 241-247. doi:
1176 10.1079/bjn19880096.
1177 Takagi, H., Saito, S., Yang, L., Nagasaka, S., Nishizawa, N., and Takaiwa, F. (2005). Oral
1178 immunotherapy against a pollen allergy using a seed-based peptide vaccine. *Plant Biotechnol*
1179 *J* 3(5)**,** 521-533. doi: 10.1111/j.1467-7652.2005.00143.x.
1180 Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in
1181 photosynthetic organisms: molecular functions and regulations of transporters and
1182 assimilatory enzymes. *Annu Rev Plant Biol* 62**,** 157-184. doi: 10.1146/annurev-arplant-
1183 042110-103921.
1184 Takaiwa, F., Takagi, H., Hirose, S., and Wakasa, Y. (2007). Endosperm tissue is good production
1185 platform for artificial recombinant proteins in transgenic rice. *Plant Biotechnol J* 5(1)**,** 84-92.
1186 doi: 10.1111/j.1467-7652.2006.00220.x.
1187 Tan, Q., Zhang, L., Grant, J., Cooper, P., and Tegeder, M. (2010). Increased phloem transport of S-
1188 methylmethionine positively affects sulfur and nitrogen metabolism and seed development in
1189 pea plants. *Plant Physiology* 154(4)**,** 1886-1896. doi: 10.1104/pp.110.166389.
1190 Tian, L., Dai, L.L., Yin, Z.J., Fukuda, M., Kumamaru, T., Dong, X.B., et al. (2013). Small GTPase
1191 Sar1 is crucial for proglutelin and alpha-globulin export from the endoplasmic reticulum in
1192 rice endosperm. *J Exp Bot* 64(10)**,** 2831-2845. doi: 10.1093/jxb/ert128.
1193 Twyman, R.M., Stoger, E., Schillberg, S., Christou, P., and Fischer, R. (2003). Molecular farming in
1194 plants: host systems and expression technology. *Trends Biotechnol* 21(12)**,** 570-578. doi:
1195 10.1016/j.tibtech.2003.10.002.
1196 Ulrich, K., and Jakob, U. (2019). The role of thiols in antioxidant systems. *Free Radic Biol Med* 140**,**
1197 14-27. doi: 10.1016/j.freeradbiomed.2019.05.035.
1198 Watanabe, M., Hubberten, H.M., Saito, K., and Hoefgen, R. (2015). "Serine Acetyltransferase," in
1199 *Amino Acids in HIgher Plants,* ed. J.P.F. D'Mello. CAB International), 195-218.
1200 Whitcomb, S.J., Nguyen, H.C., Bruckner, F., Hesse, H., and Hoefgen, R. (2018). CYSTATHIONINE
1201 GAMMA-SYNTHASE activity in rice is developmentally regulated and strongly correlated
1202 with sulfate. *Plant Sci* 270**,** 234-244. doi: 10.1016/j.plantsci.2018.02.016.
1203 Wongkaew, A., Asayama, K., Kitaiwa, T., Nakamura, S.I., Kojima, K., Stacey, G., et al. (2018).
1204 AtOPT6 Protein Functions in Long-Distance Transport of Glutathione in Arabidopsis
1205 thaliana. *Plant Cell Physiol* 59(7)**,** 1443-1451. doi: 10.1093/pcp/pcy074.
1206 Xiang, X., Wu, Y., Planta, J., Messing, J., and Leustek, T. (2017). Overexpression of Serine
1207 Acetyltransferase in Maize Leaves Increases Seed-Specific Methionine-Rich Zeins. *Plant*
1208 *Biotechnol J*. doi: 10.1111/pbi.12851.
1209 Yasuda, H., Hirose, S., Kawakatsu, T., Wakasa, Y., and Takaiwa, F. (2009). Overexpression of BiP
1210 has inhibitory effects on the accumulation of seed storage proteins in endosperm cells of rice.
1211 *Plant Cell Physiol* 50(8)**,** 1532-1543. doi: 10.1093/pcp/pcp098.
1212 Yasuda, H., Tada, Y., Hayashi, Y., Jomori, T., and Takaiwa, F. (2005). Expression of the small
1213 peptide GLP-1 in transgenic plants. *Transgenic Res* 14(5)**,** 677-684. doi: 10.1007/s11248-
1214 005-6631-4.
1215 Yoshimoto, N., Inoue, E., Watanabe-Takahashi, A., Saito, K., and Takahashi, H. (2007).
1216 Posttranscriptional regulation of high-affinity sulfate transporters in Arabidopsis by sulfur
1217 nutrition. *Plant Physiol* 145(2)**,** 378-388. doi: 10.1104/pp.107.105742.
1218 Zhang, M.Y., Bourbouloux, A., Cagnac, O., Srikanth, C.V., Rentsch, D., Bachhawat, A.K., et al.
1219 (2004). A novel family of transporters mediating the transport of glutathione derivatives in
1220 plants. *Plant Physiol* 134(1)**,** 482-491. doi: 10.1104/pp.103.030940.
1221 Zuber, H., Davidian, J.C., Wirtz, M., Hell, R., Belghazi, M., Thompson, R., et al. (2010). Sultra4;1
1222 mutant seeds of Arabidopsis have an enhanced sulphate content and modified proteome

This is a provisional file, not the final typeset article

1223 suggesting metabolic adaptations to altered sulphate compartmentalization. *Bmc Plant*
1224 *Biology* 10. doi: Artn 78
1225 10.1186/1471-2229-10-78.
1226