1 STOMAGEN-like Regulates Stomatal Generation Causing Increased Water Use

2 Efficiency during C₄ Evolution

3 Yong-Yao Zhao^{1,2}, Amy Mingju Lyu, Genyun Chen, Xin-Guang Zhu^{1,2}

4 1. Center of Excellence for Molecular Plant Sciences, Chinese Academy of Sciences

5 2. State Key Laboratory for Plant Molecular Genetics

6

7 Abstract

8 Compared with C₃ plants, C₄ plants maintain lower stomatal conductance (g_s),
 9 attributed to decreased maximal stomatal conductance (g_{smax}), without losing
 10 photosynthetic CO₂ uptake rate. Which stomatal characteristics caused the decreased
 11 g_{smax} and how did the characteristics change along C₄ evolution as well as the
 12 molecular mechanism underlying this change remains unknown.

13 Stomatal patternings were examined in *Flaveria* genus, which contains species at
 14 different evolutionary stages from C₃ to C₄ photosynthesis. We further used
 15 comparative transcriptomics, transgenic experiments and semi-*in-vitro* analysis to
 16 identify the gene underlying the decreased g_{smax} in C₄ species.

17 Results were as follows: the evolution from C₃ to C₄ species was accompanied by
 18 a stepwise decrease in stomatal density (SD) and dramatic decreased SD occurred
 19 between C₃-C₄ species and C₄-like species; *FSTOMAGEN* gene positively controls SD
 20 and its changed expression underlies the decreased SD during C₄ evolution;
 21 Furthermore, this mechanism is shared between monocotyledons and dicotyledons,
 22 indicated by the lower expression of *FSTOMAGEN* homologs in maize than rice.

23 This study suggests that lower SD is another major feature of C₄ evolution, besides
 24 C₄ enzymes and Kranz anatomy. *FSTOMAGEN* can be used to engineer lowered SD ,
 25 a C₄ feature required in the current effort of C₄ crop engineering.

26

27 **Key words:**

28 Water use efficiency, Stomatal density, C₄ photosynthesis, STOMAGEN, Stomatal
29 conductance

30

31 **Introduction**

32

33 Compared with C₃ plants, C₄ plants have higher light use efficiency (LUE) as a
34 result of a CO₂ concentrating mechanism (1). At the same time, C₄ plants also have
35 higher water use efficiency (WUE), as a result of both lower g_s and higher
36 photosynthetic CO₂ uptake rate (A) due to the CO₂ concentrating mechanism (2, 3).
37 C₄ photosynthesis therefore optimizes water and carbon relations by reducing
38 transpiration without compromising carbon assimilation (4). Improvements in WUE
39 enables C₄ plants to better adapt to saline lands, hot and drought subtropical
40 environments and open habitats compared with their close C₃ relatives, which raises
41 the possibility that C₄ photosynthesis might have been selected as a water-conserving
42 mechanism (4) and that water limitation might be the primary driver for the initial
43 ecological advantage of C₄ over C₃ (5-7). Given that C₄ plants have higher
44 photosynthetic energy conversion efficiency and higher potential for biomass
45 productivity, their lower g_s is critical for the survival and fitness of C₄ plants in the
46 field, especially in a world where water shortage is predicted to occur more often.

47 C₃ plants, *e.g.* eucalyptus and arabidopsis, can gain improved WUE through
48 reducing transpiration under long-term changed environmental conditions by
49 decreasing g_{smax} (8, 9), which implies that the decreasing g_{smax} may be an option that
50 plants use to increase WUE. Indeed, phylogenetically informed comparison shows
51 that C₄ plants usually have a lower g_{smax} compared with their close C₃ relatives (10),
52 indicating that the lower g_{smax} contribute to the higher WUE of C₄ land plants.

Therefore, low g_{smax} , same as the C₄ metabolite shuttle and Kranz anatomy, may be a required feature for C₄ land plants. But so far, in comparison with the extensive studies on the C₄ metabolite shuttle and Kranz anatomy, relatively little effort has been invested into studying the mechanism underlying the decreased g_{smax} in C₄ plants.

In principle, g_{smax} positively correlates to operating g_s (8) and the g_{smax} is determined by stomatal density (SD) and stomatal size (SS). In Eucalyptus and Arabidopsis, SD positively controls g_{smax} and operating g_s (8, 11-13). Higher SD also appears to have been selected as a strategy to gain higher g_{smax} to cope with low atmospheric CO₂ levels over the last 400 million years (14). In contrast, when grown under drought, potato plants severely reduce SS to lower operating g_s and increase WUE (15). Under ABA treatment, a sufficiently small SS can lower g_{smax} to achieve higher WUE in *Tradescantia virginiana* (16). Therefore, in different species or under different environmental conditions, either SD or SS is used as a primary determinant for g_{smax} in plants. In the case of C₄ photosynthesis, it also remains unknown whether SD or SS is the dominant characteristic responsible for the lower g_{smax} in C₄ relative to C₃ plants. For example lower g_{smax} of different species in one C₄ clade were attributed to either smaller SS or lower SD (10), which may be a result of a compounded phylogenetic effect, i.e. species used in the study were evolutionarily distant from each other and hence genetic factors unrelated to photosynthesis may influenced stomatal patterning (10).

Stomatal development is controlled by a well-studied developmental patterning program (17, 18). Stomatal patterning is controlled by a series of ligands and the transmembrane receptors, leading to downstream changes of MAPKK, MAPKs and transcription factors, ultimately influencing stomatal development (17, 18). TOO MANY MOUTHS (TMM) and ERECTA-family (ERfs: ER, ERL1, ERL2) are receptor-like kinases that negatively regulate stomatal development (19, 20), and the TMM interacts with ERfs to regulate stomatal differentiation (19, 21). The ligands that binds to ER and TMM receptors are mainly EPIDERMAL PATTERNING

82 FACTOR/EPIDERMAL PATTERNING FACTOR-LIKE family (EPF/EPFL-family)
 83 peptides, including EPF1, EPF2, and *STOMAGEN*/EPFL9 (22). The *EPF/EPFL*-family
 84 genes encode cysteine-rich peptides, which can fold into a three-dimensional structure
 85 through disulfide bond to drive stomatal patterning (23). Overexpression of *EPF1*
 86 decreases the number of stomata, whose loss of function mutation increases the
 87 number of stomata and results in stomatal cluster (24). *EPF2*, which acts earlier in
 88 stomatal development relative to *EPF1*, has a similar genetic effect as *EPF1* (25).
 89 *STOMAGEN*, in contrast to either *EPF1* or *EPF2*, positively regulates *SD* (26-28).
 90 *STOMAGEN* can compete with EPF1 and EPF2 for ERfs (29), and then prevent the
 91 phosphorylation of downstream receptors, which further results in the changed
 92 expression of three transcription factors of basic Helix-loop-Helix (bHLH) paralogs,
 93 SPCH, MUTE and FAMA, ultimately influencing stomatal development (30).

94 To study whether *SD* or *SS* underlies the decreased g_{smax} in C_4 species, here we use
 95 the *Flaveria* genus, a representative of dicotyledons which contains closely related
 96 species at different evolutionary stages from C_3 to C_4 photosynthesis. Using a genus
 97 with multiple intermediate species will enable examination of changes in stomatal
 98 patterning during C_4 evolution. In addition, rice and maize, as representatives of
 99 monocotyledons, are used to test whether the identified mechanism is conserved in
 100 monocotyledons as well. This study aims to answer the following questions: 1) Which
 101 stomatal patterning parameter, i.e. *SD* and *SS*, underlies the decreased g_{smax} in C_4
 102 evolution and how did it change during the evolution? 2) What's the molecular
 103 mechanism underlying the change of this stomatal patterning parameter? 3) Is this
 104 molecular mechanism conserved between monocotyledons and dicotyledons?

105

106 **Result**

107 **C_4 species reduce stomatal densities during C_4 evolution**

108 The maximum value of operational g_s (marked with circles on figure) at saturated
 109 light intensity of C₃ *Flaveria* species is higher than C₄ species, indicating that the C₃
 110 *Flaveria* species tends to have a higher g_s (Fig 1a). On the whole, the operational g_s at
 111 the saturated light of C₃ species and C₃-C₄ species in *Flaveria* are higher than C₄
 112 species and C₄-like species (Fig 1b). We then compared the *SD* of different *Flaveria*
 113 species grown either outdoors or in greenhouse on both sides of leaves. We found that,
 114 in comparison to C₄ and C₄-like photosynthesis, the *SD* of C₃ and C₃-C₄
 115 photosynthesis on the abaxial leaf surface of *Flaveria* were nearly 150% higher (Fig.
 116 1d and e). With the progression from C₃ to C₄ species, there was a distinctly gradual
 117 decrease in *SD* (Fig. 1f and g). The difference in *SD* between C₃ and C₃-C₄ (type I)
 118 species wasn't as great as that between C₃ and C₄ species; and the *SD* of C₃-C₄ (type II)
 119 species fell in between C₃-C₄ (type I) and C₄-like species (Fig. 1d, e, f and g). The *SD*
 120 of C₄ and C₄-like species were similar, although C₄-like is slightly smaller than C₄ (Fig.
 121 1d, e, f and g). The g_s at saturated light of *F. sonorensis* (son) (C₃-C₄ (type I)) is
 122 extremely high (marked with circle on figure), matching with its highest *SD* within
 123 *Flaveria* genus (Fig 1a, f). The observed reductions in both g_s and *SD* of C₄ and
 124 C₄-like *Flaveria* species compared to C₃ and C₃-C₄ species suggest that the decrease in
 125 *SD* contributes most of the decrease in g_s during C₄ evolution (Fig 1h). For the adaxial
 126 leaf surface, the *SD* of C₃ species was again higher than that of the C₄ species,
 127 although the difference between *SD* in the intermediate species was less significant
 128 (Supplementary fig. 2a, b, c, d).

129 We next chose three *Flaveria* species, i.e. *F. robusta* (*rob*), *F. ramosissima* (*ram*)
 130 and *F. bidentis* (*bid*) to represent C₃, C₃-C₄ and C₄ species, respectively, to conduct
 131 detailed analysis of gas exchange and guard cell morphological features. Similar to
 132 the saturated light condition, the WUE (A/g_s) of *rob* measured under unsaturated light
 133 intensity was also higher than *bid* (Fig.1c). Furthermore, under unsaturated light, the
 134 operational g_s of C₃ *Flaveria* species was 71% higher than that of C₄ species (Fig. 1c),
 135 consistent with the higher *SD* of C₃ species, implying that operational g_s under
 136 unsaturated light is also associated with *SD*. For plants grown in outdoor conditions,

SS which for this study was taken as guard cell length (l) showed the opposite trend. The l of *bid* was the longest (Fig. 1j). This inverse relationship between SD and SS was also observed earlier (8, 10, 14). In greenhouse, l of *ram* was the longest, however, l of *bid* and *rob* were almost the same, even though the l of *bid* was slightly longer (Fig. 1j). The ratio between the number of guard cells and the number of epidermal cells, i.e. the stomatal index (SI), in *rob* were both nearly 50% higher than *bid* and *ram* in plants grown in either greenhouse or outdoors, suggesting that the proportion of stomata occupying the epidermis, i.e. SI , contributed to the highest SD in *rob* (Fig. 1i and k). In addition to the above analysis of *Flaveria* plants grown from cuttings, we also planted *Flaveria* germinated from seeds. Results showed that the SD of C_3 species was again higher than C_4 species, consistent with the cuttings (Supplementary fig. 3).

149

150 ***FSTOMAGEN* positively controls stomatal density (SD) and may be a factor** 151 **contributing to the decreased SD in C_4 *Flaveria* species**

The signaling pathway for stomatal development has been studied extensively, which includes a signal transduction, MAPK cascade and activation/inactivation of basic Helix-loop-Helix (bHLH) transcription factors (17). We first identified amino acid sequence encoded by homologs of stomatal development related genes in *Flaveria*. Amino acid sequence alignment analysis shows that homologs of epidermal patterning factor/epidermal patterning factor-Like gene family (*EPF/EPFL*-family) was conserved at the C-terminal part of amino acid sequence (Fig 2A, Supplementary fig 4). Each member of *EPF/EPFL*-family genes encodes a secretory protein containing a signal peptides at the N-terminus and a functional domain at the C-terminus (24, 25, 31). Analysis showed hydrophobicity of the homologs of *EPFs/EPFL* in *Flaveria* was similar to the *EPF/EPFL*, implying the homologs are also the secretory proteins and may have the same physiological function as *EPF/EPFL* (Supplementary fig 5). We hypothesized that either the amino acid sequence or

165 expression levels of stomatal developmental related genes changed during C₃ to C₄
166 evolution. Analysis results revealed no differences within the functional region of
167 amino acid sequence for STOMAGEN, EPF1 and EPF2 homologs between C₃ and C₄
168 species (Fig. 3d and supplementary fig. 4a, b), indicating the change of amino acid
169 sequence is unlikely to be a cause of the changed *SD* during C₄ evolution.

170 As an increased transcript abundance has been a major feature for many genes
171 related to C₄ cycle or transporters during C₃ to C₄ evolution in *Flaveria* (32, 33), we
172 further compared the transcript abundance of these stomatal developmental related
173 genes between C₃ and C₄ species in *Flaveria* genus. The expression patterns of known
174 C₄ related genes in our data analysis were same as the multiple previous works, which
175 indicates that our data collection and analysis procedure were reliable (Supplementary
176 fig. 10) (32, 33). First, we found that most homologs of stomatal developmental
177 related genes showed much higher transcript abundance in developing tissues as
178 compared to developed tissues (Fig. 3b), which is consistent with their function of
179 regulating stomatal differentiation in developing leaf (24, 26, 28, 34). *STOMAGEN*,
180 *EPF1*, *EPF2* and *TMM* showed higher expression levels in C₃ species than in those in
181 C₄ species (Fig. 3b, Supplementary fig. 7). RT-qPCR quantification of the expression
182 levels of *STOMAGEN*, *EPF1*, *EPF2* homologs between C₃ and C₄ species also
183 confirmed the results from RNA-SEQ (Fig. 3c). Considering that the *EPF1*, *EPF2*
184 and *TMM* are negative regulators for stomatal development (22, 24, 25, 29) and the
185 lower *SD* of C₄ species, these genes should not be the factors responsible for the
186 decreased *SD* in C₄ species. *STOMAGEN* is a positive regulator of *SD* (22, 26, 27),
187 whose lower expression in C₄ species was consistent with the decreased *SD* in C₄
188 species (Fig. 1d,e,f,g, Supplementary fig 3), and the expression of the reduced parts of
189 *STOMAGEN* was higher than that of *EPF1* and *EPF2*. In this study, we renamed
190 *STOMAGEN* in *Flaveria* as *FSTOMAGEN*.

191 There are considerable species differences in amino acid sequence of signal
192 peptide of *STOMAGEN* homologs, especially between species with large evolutionary
193 distance (Fig. 2a). The functional domain of *STOMAGEN* is however highly

conserved, not only between different *Flaveria* species (Fig. 3d), but also between species that have a large evolutionary distance (Fig. 2a). In *Arabidopsis*, the increased expression of *STOMAGEN* leads to increased *SD*, and when applied exogenously the signal peptide is not required for synthetic *STOMAGEN* to induce the increased in *SD* (26, 27, 31, 34), indicating the replacement of signal peptide will not affect the *STOMAGEN*'s function increasing *SD*. A phylogenic tree constructed for *STOMAGEN* from different plants can be divided into two clades, i.e. dicotyledon (blue clades) and monocotyledon (red clades) (Fig. 2b). Previous studies showed that overexpressing either of the two rice homologs of *STOMAGEN* in rice can cause an increased *SD* in *Arabidopsis* (35, 36), implying strong conservation of *STOMAGEN* function across monocots and dicots. When plants are grown under environmental conditions that promote or inhibit stomatal development e.g. at high or low light intensity, the expression of *STOMAGEN* is known to increase or decrease, respectively (12, 34, 37). A similar modulation of *STOMAGEN* expression levels was observed for rob when grown at different light conditions (Supplementary fig 8).

In order to confirm that *FSTOMAGEN* can induce stomatal generation, we overexpressed the *FSTOMAGEN* in *Arabidopsis*. We recombined the signal peptide of *STOMAGEN* from *Arabidopsis* to the functional region of *FSTOMAGEN* and named the hybrid gene as AFSTO (Fig. 2c). Overexpressing either the AFSTO or the intact *FSTOMAGEN* (FSTO) from *F.robusta* into *Arabidopsis* resulted in increased *SD* (Fig. 2d, e, f); Furthermore, overexpression of *FSTOMAGEN* also induced clustered stomata (Fig. 2e, f), as also reported for other *STOMAGEN* homologs (26). To further confirm that *FSTOMAGEN* can increase *SD* in *Flaveria*, we synthesized a peptide representing the *FSTOMAGEN* functional region, consisting of 45 amino acids (Fig. 2a). *In vitro* application of the *FSTOMAGEN* peptide increased the *SD* and caused stomatal cluster in *F. bidentis*, and the amount of *FSTOMAGEN* is positively correlated with the number of stomata as previously reported (26, 27, 31, 34, 35) (Fig. 2g, h, i, supplementary fig. 12). These results demonstrate that, in addition to its *in vivo* function, *in vitro* application of *FSTOMAGEN* can also directly positively

223 regulate *SD* in *Flaveria*. This is consistent with the notion that *STOMAGEN* is
224 expressed in mesophyll cells and is secreted out of cells to influence the stomatal
225 development in the epidermal layer in *Arabidopsis* (26, 27).

226

227 **The low transcript abundance of *FSTOMAGEN* underlie the decreased stomatal**
228 **density of *Flaveria* species during *C*₄ evolution**

229 We examined the copy number of *STOMAGEN* in *Flaveria* species using blastn
230 with the coding sequence of *FSTOMAGEN* as query sequence against all the
231 assembled contigs in each *Flaveria* species. Results show that there was only one hit
232 sequence with e-value less than 10^{-5} in each species, suggesting that there is a single
233 copy of *FSTOMAGEN* in each *Flaveria* species (Supplementary fig. 11). Therefore
234 we can ignore the impact of other *FSTOMAGEN* paralogs on stomatal development,
235 and there will be no quantitative errors in the gene expression levels caused by other
236 paralogs not being considered. Further examination of the RNA-seq data of *Flaveria*
237 showed that the expression levels of *FSTOMAGEN* in *C*₃ species were markedly
238 higher than those in *C*₄ species along different developmental stages of leaves (Fig.
239 3e). During leaf growth and development, the expression of *FSTOMAGEN* in both *bid*
240 and *rob* increased first, reaching the highest level at leaf development stage 3 or 4,
241 and then declined (Fig. 3e). These similar patterns of *STOMAGEN* expression in these
242 two *flaveria* species indicated that it would be valid to compare the *STOMAGEN*
243 expression levels in leaves at the same development stage between different *Flaveria*
244 species. The expression levels of *FSTOMAGEN* in *C*₄ and *C*₄-like species were lower
245 than those in *C*₃ and *C*₃-*C*₄ species in *Flaveria* based on either RNA-seq or RT-qPCR
246 quantification (Fig. 3b, f, g), consistent with the different *SD* among different
247 photosynthetic types (Fig. 1d, e). Expression levels of *FSTOMAGEN* increased with
248 the increase of *SD* when the data from a large number of *Flaveria* species with
249 different photosynthetic type were combined (Fig. 3h). Together these results suggest

250 that a gradual change in the *FSTOMAGEN* expression levels underlied the stepwise
251 change in *SD* during C₄ evolution.

252

253 **The role of *FSTOMAGEN* homologs in regulating *SD* between rice and maize.**

254 Species of the *Flaveria* genus are dicotyledonous plants. To gain further insight into
255 whether *STOMAGEN* might be a generic factor leading to the difference in *SD*
256 between C₃ and C₄ species across higher plants, we tested whether *STOMAGEN*
257 homologs are differentially expressed between C₃ and C₄ species in monocotyledon
258 species. We used rice and maize in this study. The abaxial leaf *SD* of C₃ rice were 240%
259 higher than that of C₄ maize, consistent with a 574% higher *g_s* in rice compared to
260 maize (Fig. 4a, b). The stomatal length (*l*) in maize was 133% longer, and stomatal
261 index (*SI*) showed significant difference (Fig. 4a, b). *SD* showed a positive
262 relationship with *STOMAGEN* transcript levels in Arabidopsis, rice and maize (Fig. 4e,
263 f), i.e. with the increase in the transcript abundance of *STOMAGEN* homologs, the *SD*
264 gradually increases. We did not observe a significant difference in the efficacy of
265 *STOMAGEN* homologs from Arabidopsis and either rice or maize in terms of their
266 effectiveness in controlling *SD* (Fig. 4e,f), suggesting that the function of
267 *STOMAGEN* homologs is highly conserved between dicotyledonous and
268 monocotyledonous plants. This similarity in efficacy allowed us to compare the
269 expression levels of *STOMAGEN* homologs from different species, even though their
270 amino acid sequences differ. To test whether the difference in *SD* between rice and
271 maize is due to changed expression levels of *STOMAGEN*, we combined the
272 transcript abundance of their different paralogs, since there are two and three paralogs
273 of *STOMAGEN* in rice and maize (35), respectively. We found that *STOMAGEN*
274 homologs in rice had much higher transcript abundance than that in maize (Fig. 4d),
275 while the transcript abundance of other stomatal development related genes did not
276 differ to the same extent between rice and maize (Fig. 4c). Together our observations

277 suggest that the lower transcript abundance of *STOMAGEN* homologs in maize also
278 underlies its lower *SD* compared with rice.

279

280 **Method**

281 **Plant materials, sample condition and data sources**

282 All *Flaveria* species used in this study were kindly provided by Prof. Rowan F.
283 Sage (University of Toronto). Plants used for stomatal phenotyping were re-grown
284 from cuttings. Placed in an environment humidified with humidifiers to generate roots.
285 After plants generated roots, they were transferred outdoors (Institute of Plant
286 Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China), and
287 transplanted into 4-L pots filled with topsoil in April. In July, the number of stomata
288 on abaxial and adaxial surfaces of leaves were counted. Another batch of *Flaveria*,
289 which were also grown from cuttings, were transferred to the phytotron (Institute of
290 Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China).
291 Growth temperature of the phytotron was maintained at 25-28°C, photoperiod was
292 maintained at 16h light / 8h dark, relative humidity (RH) was at 60-65% and
293 photosynthetic photon flux density (PPFD) was maintained at about 180 $\mu\text{mol m}^{-2} \text{s}^{-1}$.
294 Two months later, the numbers of stomata on the abaxial and adaxial surfaces of
295 leaves were counted. RNA was extracted from 1.5~3.0 cm long young leaves on the
296 axial shoot which were still expanding. Three hours before sampling, we provided
297 sufficient water to ensure that plants were not water limited, since drought and water
298 status can affect expression levels of *STOMAGEN* and *SD* (12, 38). Gene expression
299 was further examined under normal growth conditions by RT-qPCR for the *F.robusta*,
300 *F.rammosissima* and *F.bidentist*, which are representative of C₃, C₃-C₄, C₄ species in
301 the *Flaveria* genus. These three species were grown in the same phytotron with the
302 same environmental conditions, except that we did not water plants before sampling.
303 To test whether plants grown from cuttings or from seeds show difference in the *SD*,
304 we further grew *Flaveria* species from seeds in phytotron and examined stomatal

properties. Same environmental conditions were maintained for phytotron as above.
Two months after germination of seeds, we counted stomatal number.

All plants were well watered to avoid drought stress during their growth. We supplied the same amount of commercial nutrient solution (Peters Professional, SCOTTS, N:P:K = 20:20:20+TE, at recommended dose for application) to the different species to avoid nutrient deficiency.

Rice (*Oryza Sativa Nipponbare*) and maize (*Zea maize B73*) were grown in the greenhouse (Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China) under a PPFD of 550 $\mu\text{mol}/\text{m}^2/\text{sec}$, a photoperiod of 12h light / 12h dark, and a daytime temperature of around 31°C and a night-time temperature of around 22°C. The RH was maintained at around 50-55%. The plants were regularly watered to avoid drought stress.

In addition to the expression data collected by in this study, we also used public data for comparison of transcript abundance. Considering that stomatal development related genes are mainly expressed in young tissues, in each of these used public data source, only samples for young leaves or young tissue of leaves were included. The transcript levels of stomatal development related genes in *Flaveria* were obtained from three studies. The first data source is the One Thousand Plants (1KP) Consortium; we subsequently analyzed the data according to the methods described in previous work (39, 40). The second and third data sources were Mallman et al (2014) (33) and Kumari Billakurthi et al (2018) (41), respectively. For the second data source, the second and fourth visible leaves were sampled; the second leaf was not completely mature, and stomatal developmental related genes were still actively expressed. Expression levels for stomatal development related genes in developing leaves for rice and maize were obtained from Wang et al. (42).

330

331 Conservation and hydrophobicity analysis of proteins

Amino acid sequence of the homologs of *STOMAGEN*, EPF1 and EPF2 were aligned by clustalx, and graphs for the multiple alignments were drawn with ESPrit 3.0 (43) (<http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). Hydrophobicity of proteins was determined by ProtScale (44) (<https://web.expasy.org/protscale/>).

336

337 **Construction of the phylogenetic tree of *STOMAGEN***

To construct the phylogenetic tree of *STOMAGEN*, we queried the orthologs of *STOMAGEN* in representative species in Viridiplantae. Protein sequence of *STOMAGEN* from *A. thaliana* (AT4G12970) was used as query sequence. We selected 29 representative species which were chosen according to two criteria: 1) one species to represent each main lineage, e.g., *M.pusilla* represents *M.pusilla* chlorophyte, and *M.polymorpha* represents the ancestral type of Embryophyte; 2) one lineage to represent monocot and one lineages to represent dicot. We used the brassicaceae family (including *Arabidopsis thaliana*) to represent dicot and used the grass family (including *Oryza sativa*) to represent monocot. The protein sequences of representative species were downloaded from Phytozome (<https://phytozome.jgi.doe.gov/>). Protein sequences for the *Flaveria* species were predicted using OrfPredictor (45) based on denovo assembled transcript as described in (46). *STOMAGEN* orthologs in *Flaveria* and other species (except for *A. thaliana*) were predicted by using blastp from BLAST (V 2.2.31+) (47) with a E-value threshold of 1e-5, where the protein sequence of *STOMAGEN* from *A. thaliana* was used as query. Protein sequences of *STOMAGEN* orthologs were then aligned using MUSCLE (48). Phylogenetic tree was constructed by applying Raxml (49) based on protein sequences using PROTGAMMAILG (General time reversible amino acid substitution model with assumption that variations in sites follow gamma distribution) model. Bootstrap scores were obtained based on 100,000 bootstraps.

358

359 **Measurement of stomatal conductance (g_s)**

360 The LI-6400 Portable Photosynthesis System (Li-Cor, Inc., Lincoln, NE, USA) was
 361 used to measure g_s on the youngest fully expanded leaves from approximately
 362 three-months-old *Flaveria* plants grown from cuttings. During the measurements, the
 363 cuvette temperature was maintained at 25°C, the PPFD in the cuvette was maintained
 364 at ~500 $\mu\text{mol}/\text{m}^2\text{s}$ and the CO_2 concentration was maintained at around 400 ppm.
 365 Chamber humidity was matched to ambient air humidity. Leaf-to-air vapor pressure
 366 difference (VPD) was maintained around 1.1 kPa. Before the g_s measurements,
 367 *Flaveria* plants were adequately watered, maintained under constant light with a
 368 PPFD of ~500 $\mu\text{mol}/\text{m}^2\text{s}$ until g_s reached a steady state. Six different plants were used
 369 to measure g_s for every *Flaveria* species. Eight different plants were used to measure
 370 g_s for both rice and maize. External environments during the g_s measurements for rice
 371 and maize were controlled to be their growth conditions. We used the third or fourth
 372 leaves from the apex for g_s measurements. In maize and rice, the 16-day-old third
 373 leaves, which were the youngest mature leaves, were used for gas exchange
 374 measurements (42).

375

376 **Phenotyping stomatal traits**

377 The youngest fully expanded leaves (usually the third or fourth leaves from the
 378 apex) at the main stem were used to phenotype stomata related traits. In *Flaveria* and
 379 *maize*, stomata related traits were observed using both nail polish method and
 380 epidermis directly peeled from leaves (6, 11). For the first method, nail polish was
 381 painted on the both leaf surfaces to produce the impressions of epidermis. After
 382 drying of the nail polish, the nail polish was peeled off the leaf and placed on a
 383 microscope slide for examination. For the second method, leaf epidermis was peeled
 384 off manually and transferred onto microscope slide. For the measurement of SI in rice,
 385 a small piece of rice blade was soaked in 9:1 (V/V) ethanol:acetic acid overnight,
 386 washed with water and cleared in chloral hydrate solution (glycerol:chloral

hydrate:water, 1:8:1) overnight, and then placed on a microscope slide for imaging. Light microscope (smarte; CNOPTec), camera and image processed software (OPTPRO; CNOPTec) were used to observe and image epidermis at x100 magnification. Center of a leaf was chosen to count and measure stomatal density and stomatal length respectively. Five stomata randomly chosen along a diagonal of the varnish peels or epidermis were measured for stomatal length.

393

394 **RNA isolation and quantitative real-time PCR**

395 The leaf samples taken from main stem were immediately frozen in liquid nitrogen
396 and used directly for RNA extraction. Total RNA was extract from the young leaves
397 using the PureLink RNA Mini kit (Thermo Fisher Scientific). First-strand cDNA was
398 synthesized by the TransScript One-step gDNA Removal and cDNA Synthesis
399 SuperMix (TRANSGEN). Real-time PCR was performed on the CFX connect^{TM}
400 (Bio-Rad) using the UNICON^{TM} qPCR SYBR ®Green Master Mix (YEASEN)
401 according to the manufacturer's instructions. The calculation of expression according
402 to the $2^{-\Delta\Delta CT}$.

403 The expression stability of internal reference genes was assessed by the software:
404 geNORM and NormFinder (50, 51). Specifically, we first calculated the coefficient of
405 variations (CV) of all genes in different *Flaveria* species. After that, all genes were
406 sorted according to the calculated CV values. 150 genes with the lowest CV were
407 selected as potential candidate reference genes. Among these, those with mean
408 expression level >300 were chosen as candidate reference genes. We also included
409 those commonly used house-keeping genes, which include *UBC9* (AT4G27960),
410 *UBQ11* (AT4G05050), *ACT7* (AT5G09810), *GAPDH* (AT1G13440), *EF1a*
411 (AT5G60390), *CACMS* (AT5G46630) (Supplementary fig. 6a, b, d), as the internal
412 reference genes in this study. The specificity of the primers (Supplementary table 2)
413 for all these candidate reference genes were identified by agarose electrophoresis
414 (Supplementary fig. 6c). With Genorm (Supplementary fig. 6e) and Normfinder

(Supplementary fig. 6f), we concluded that *ACT7*, *EF1a* and *HEME2* showed stable expression levels across rob, ram and bid. *EF1a* and *ACT7* are classic reference genes, and they showed minor variations under different conditions. The geometrical means of the expression level of *ACT7* and *EF1a* were finally used as normalization factors. All primers used for RT-qPCR are listed in Supplementary table 2. For each gene, three technical replicates and at least three biological replicates were performed.

421

422 **Vector construction and transformation**

A list of primer sequences were used for plasmid construction (Supplementary table 2). The *STOMAGEN* and *FSTOMAGEN* DNA sequences, named ATSTO-cDNA and FloSTO-cDNA respectively, were PCR-amplified from cDNA in Arabidopsis and *Frobusta*. To obtain the AFSTO fragment, signal peptide region of *STOMAGEN* was cloned from ATSTO-cDNA by PCR-amplification, and functional region of *FSTOMAGEN* was PCR-amplified from FloSTO-cDNA. The signal peptide and functional region were combined by fusion PCR. To obtain an intact *FSTOMAGEN* (FSTO) gene, the whole gene in *Frobusta*, including signal peptide, propeptide and *STOMAGEN* functional region, was amplified from the cDNA of *Frobusta*. To over-express AFSTO and FSTO, the AFSTO and FSTO were cloned separately to a pcambia-3300 binary vector which has a CaMV 35S promoter. This construct was introduced into *A. tumefaciens* strain GV3101, transformed into Arabidopsis using a floral dipping procedure. Transgenic lines were selected on soil that had been treated with a diluted solution of phosphinothricin (diluted 3000 times) over night. Phenotype was observed using three individual leaves for each line, and we used three lines of T1 plant in total.

439

440 **Refolding of synthetic *FSTOMAGEN***

Chemically synthesized peptide (Synthesized peptide sequence: IGSVKPTCTY
NECRGCRSRCRAEQVPVEGNDPINSAYHYRCICHR, Sangon Biotech, Shanghai,
China) was dissolved in 20 mM Tris-HCl, pH 8.8, and 50 mM NaCl and then dialysed
(Sangon Biotech, 0.5 KD-1 KD) for 1 day at 4°C against 0.5 mM glutathione (reduced
and oxidized forms, Macklin) and 200 mM L-arginine (meilunbio) at pH 8.0, which
were further dialysed three times against 20 mM Tris-HCl, pH 8.8, and 50 mM NaCl
for 1.5 days to remove glutathione.

Stomata induction assay

When the first true leaf or cotyledon appeared in *F. bidentis* plants that had
germinated in 1/2 MS sterilized solid medium, FSTOMAGEN peptide was applied on
the plants. After the *F. bidentis* plants were further incubated in 1/2 MS liquid medium
at 22°C for 5 days, stomatal number on the abaxial surfaces for the first true leaf was
counted under a differential interference contrast microscope (DIC).

Plotting and statistical analysis

Unpaired two-tailed t-test ($\alpha=0.05$) were performed by GraphPad Prism 7. The
linear regressions ($\alpha=0.05$) were performed by GraphPad Prism 7. In Fig 4, the F test
for the intercept and slope from the linear regression analysis were performed by
GraphPad Prism 7. Differences between means of multi-groups were assessed with
one-way analysis of variance (ANOVA). Duncan's multiple range test was performed
using the agricolae package of the R (<https://www.r-project.org/>) to detect differences
between two groups. The A and g_s at the saturated light intensity were extracted from
(2, 52) and the data describing the correlation between SD and expression level of
STOMAGEN homologs in Arabidopsis, rice and maize (Fig. 4e,f) was digitized from
(35, 53) using the OriginPro 2018.

467

468 Discussion

469 C₄ plants have higher water use efficiency and light use efficiency, as result of the
 470 combined effect of decreased operating g_s , g_{smax} and a CO₂ concentrating mechanism
 471 (Fig 1b, c) (3, 4, 10). Comparative studies using plants from different C₄ clades show
 472 that lower g_{smax} is due to either a smaller stomatal size (SS) for some C₄ clades, or
 473 lower stomatal density (SD) for other clades (10). We collected and analysed SD data
 474 from literature for a large number of C₃ and C₄ species (10, 54-60). The average SD
 475 on both sides leaf surface of C₃ species was significantly higher than that of C₄ species
 476 (Supplementary fig. 1). Recent studies have shown that when the physiological and
 477 ecological differences between C₃ and C₄ plants are compared, the phylogenic effect
 478 must be considered (3, 61), i.e. comparing either stomatal density or size without
 479 considering the phylogeny inevitably has the compounded effect of other genetic
 480 factors other than photosynthetic types. So here we used the *Flaveria* genus which
 481 contains species with close evolutionary distance and intermediate photosynthetic
 482 types to examine changes in SD and SS during C₄ evolution. We showed that,
 483 compared to the SD of C₃ species, the SD of C₄ species decreased (Fig. 1d, e). This
 484 decrease in SD is due to less stomata generation rather than augmented stomata and
 485 epidermal cells (Fig. 1k). In two other closely related Cleomaecae
 486 species, *Gynandropsis gynandra* (C₄) also has a lower SD than *Tarenaya hassleriana*
 487 (C₃), consistent with the difference in their g_s (62). In contrast, the SS as represented
 488 by stomatal length was longer in C₄ species compared to C₃ species in *Flaveria* (Fig.
 489 1e, 4a), consistent with previous reports that SD negatively correlates with SS (8, 14),
 490 suggesting that the SS is not an important contributor to the decreased g_{smax} and
 491 operational g_s in C₄ species.

492 The lower SD in phylogenetically closely related C₄ species compared to C₃
 493 species suggests that, during evolution from C₃ to C₄ species, the SD may gradually
 494 decrease. This is indeed shown in (Fig 1d, e, f, g,) where with the gradual transition of

495 photosynthetic type from C₃ to C₃-C₄ (type I), to C₃-C₄ (type II), to C₄-like and to C₄,
 496 there is a gradually decrease in *SD*. Usually stomata conductance is significantly
 497 higher on the abaxial surface, as shown in maize, broad bean, and wild tomato species
 498 (63-65). Here we found that there is a highly significant difference in *SD* on the
 499 abaxial surface between the C₃ species and the C₄ species (Fig. 1f,g); by comparison,
 500 on the adaxial surface, though the difference between *SD* in the C₃ and C₄ species was
 501 smaller, C₄ species still had lower *SD* (Supplemental fig 2c, d). All these show that
 502 lower *SD* is the determinant of the commonly observed lower *g_{smax}* and
 503 correspondingly higher water use efficiency in C₄ compared to C₃ plants (2, 66).
 504 Therefore, lower stomatal density could be another major feature of C₄ evolution,
 505 besides C₄ enzymes and Kranz anatomy.

506 The molecular program controlling stomata differentiation has been extensively
 507 studied and the core signal transduction pathway is known (17, 19, 21, 29). In theory,
 508 since *SD* is controlled by a signaling network composed of many interacting
 509 components, many options can be used to alter *SD* by transgenic manipulation, as
 510 demonstrated previously, e.g. increasing expression of homologs of *EPF1* to reduce
 511 *SD* (67-70), increasing the expression of *EPF2* to reduce *SD* (25), **and**
 512 **over-expressing** *STOMAGEN* to increase *SD* (26-28). In fact, these options have been
 513 used to improve WUE and drought resistance (11, 67-70) or improve photosynthetic
 514 efficiency (13). Out of these different options to reduce *SD*, here we present results
 515 which indicate that a downregulation of the expression of *FSTOMAGEN*,
 516 *STOMAGEN*-like genes, have been selected as a molecular switch to reduce *SD*
 517 during C₄ evolution in *Flaveria*, which is also the case in monocotyledon (maize and
 518 rice) (Fig 4d). The evidence from this study and previous reports show that the role of
 519 *STOMAGEN* in controlling *SD* is well conserved between dicotyledons and
 520 monocotyledon (35, 71). When synthetic *FSTOMAGEN* peptide was applied on young
 521 developing leaves, we found that similar increases in *SD* were observed between
 522 *Arabidopsis* and *Flaveria* species (Fig. 3i). Furthermore, when the correlations
 523 between the expression level of *STOMAGEN* and *SD* was compared between

524 Arabidopsis and rice, or between Arabidopsis and maize, there is non-significant
525 difference for the slope of the correlations (Fig. 4e, f). All these suggest that the
526 declining expression of *STOMAGEN* homologs underlies the observed decrease in *SD*
527 during the evolution of C₄ photosynthesis.

528 Many earlier studies suggest that water limitation might have been a primary
529 selective pressure for evolution of C₄ photosynthesis (5-7). In *F. angustifolia*, a type I
530 C₃-C₄ intermediate *Flaveria* species, when grown under outdoor natural light
531 conditions, the *STOMAGEN* expression level, *SD* and *g_s* were still similar to those in
532 two C₃ *Flaveria* species, i.e. *F. cronquistii* and *F. pringlei* (2) (Fig. 1d, 3b), as also the
533 case in *Heliotropin* (72). In contrast, the gamma star of *F. angustifolia* was
534 dramatically decreased (73) compared to their C₃ ancestors, as a result of an operating
535 CO₂ concentrating mechanism (CCM) of C₂ (74). This suggests that during the
536 evolution from C₃ to C₃-C₄ intermediate species, where the movement of GDC from
537 mesophyll cell to bundle sheath cell occurred, the primary evolutionary driving force
538 should be to elevate CO₂ levels around Rubisco, rather than H₂O saving (75). After
539 this step, reducing water loss might become crucial. The emergence of C₂ mechanisms
540 can increase leaf photosynthetic CO₂ uptake rate and biomass production, which can
541 in principle create a higher water demand and increased the danger of embolism,
542 especially in the arid and semi-arid areas, in the intermediate species (4). This will
543 create an even greater challenge once the complete C₄ photosynthetic metabolism
544 emerged where photosynthetic CO₂ uptake rate and biomass production are further
545 increased. Therefore, decreased *SD* and further increased water use efficiency
546 observed in this study and earlier studies (2, 52) should be a required change during
547 the transition from the C₃-C₄ intermediate species and later C₄ species.

548 Many genes control the guard cell developing, i.e. *STOMAGEN*, *EPFs/EPFLs*, *ER*
549 (Fig 3A). In principle, any of these genes controlling stomatal patterning can be
550 potentially selected to alter *SD* by regulation of expression during the emergence of
551 C₄ photosynthesis. In fact, these genes have innate mechanisms to alter their
552 expression under different environments, such as low CO₂ (76). However, not only in

the *Flaveria* genus, but also in the grass family, *STOMAGEN* was shown to be the regulator selected to reduce *SD*. The two specialities of *STOMAGEN* probably have accounted for the phenomenon. First, *STOMAGEN* can regulate the entire molecular process of stomatal development, rather than *EPF1*, *EPF2*, *TMM* and the members of *ERECTA*, which individually only regulate particular phases of the stomatal development (18, 19, 21, 22, 24, 25). Second, which is potentially more important, *STOMAGEN* is expressed in the mesophyll cell, while all other regulators controlling stomatal development are expressed in the cell of stomatal lineage (22, 24-27). Having a genetic regulator controlling stomatal density in the mesophyll cell, where many of the C_4 related genes are expressed, can enable an easier coordination between stomatal development and C_4 photosynthesis. Future work is still needed to elucidate molecular level coordinating the expression of *STOMAGEN* and C_4 photosynthetic related genes.

566

567 Reference

- 568 1. O. V. Sedelnikova, T. E. Hughes, J. A. Langdale, Understanding the Genetic Basis of C_4 Kranz
569 Anatomy with a View to Engineering C_3 Crops. *Annu Rev Genet* **52**, 249-270 (2018).
- 570 2. P. J. Vogan, R. F. Sage, Water-use efficiency and nitrogen-use efficiency of C_3 - C_4 intermediate
571 species of *Flaveria* Juss. (Asteraceae). *Plant, cell & environment* **34**, 1415-1430 (2011).
- 572 3. S. H. H. Taylor, S. P. Rees, M. Ripley, B. S. Woodward, F. I. Osborne, C. P., Ecophysiological traits
573 in C_3 and C_4 grasses: a phylogenetically controlled screening experiment. *New Phytol* **185**,
574 780-791 (2010).
- 575 4. C. P. Osborne, L. Sack, Evolution of C_4 plants: a new hypothesis for an interaction of CO_2 and
576 water relations mediated by plant hydraulics. *Philos Trans R Soc Lond B Biol Sci* **367**, 583-600
577 (2012).
- 578 5. H. Zhou, B. R. Helliker, M. Huber, A. Dicks, E. Akcay, C_4 photosynthesis and climate through
579 the lens of optimality. *Proc Natl Acad Sci U S A* **115**, 12057-12062 (2018).
- 580 6. G. Reeves *et al.*, Natural Variation within a Species for Traits Underpinning C_4 Photosynthesis.
581 *Plant physiology* **177**, 504-512 (2018).

- 582 7. B. P. Williams, I. G. Johnston, S. Covshoff, J. M. Hibberd, Phenotypic landscape inference
583 reveals multiple evolutionary paths to C₄ photosynthesis. *Elife* **2**, e00961 (2013).
- 584 8. P. J. Franks, P. L. Drake, D. J. Beerling, Plasticity in maximum stomatal conductance
585 constrained by negative correlation between stomatal size and density: an analysis using
586 *Eucalyptus globulus*. *Plant, cell & environment* **32**, 1737-1748 (2009).
- 587 9. T. Doheny-Adams, L. Hunt, P. J. Franks, D. J. Beerling, J. E. Gray, Genetic manipulation of
588 stomatal density influences stomatal size, plant growth and tolerance to restricted water
589 supply across a growth carbon dioxide gradient. *Philos Trans R Soc Lond B Biol Sci* **367**,
590 547-555 (2012).
- 591 10. S. H. Taylor *et al.*, Photosynthetic pathway and ecological adaptation explain stomatal trait
592 diversity amongst grasses. *New Phytol* **193**, 387-396 (2012).
- 593 11. P. J. Franks, W. Doheny-Adams T, Z. J. Britton-Harper, J. E. Gray, Increasing water-use
594 efficiency directly through genetic manipulation of stomatal density. *New Phytol* **207**,
595 188-195 (2015).
- 596 12. E. T. Hamanishi, B. R. Thomas, M. M. Campbell, Drought induces alterations in the stomatal
597 development program in *Populus*. *J Exp Bot* **63**, 4959-4971 (2012).
- 598 13. Y. Tanaka, S. S. Sugano, T. Shimada, I. Hara-Nishimura, Enhancement of leaf photosynthetic
599 capacity through increased stomatal density in *Arabidopsis*. *New Phytologist* **198**, 757-764
600 (2013).
- 601 14. P. J. F. a. D. J. Beerling, Maximum leaf conductance driven by CO₂ effects on stomatal size and
602 density over geologic time. *PNAS* **106**, 10343–10347 (2009).
- 603 15. S. h. Zhang, X. f. Xu, Y. m. Sun, J. l. Zhang, C. z. Li, Influence of drought hardening on the
604 resistance physiology of potato seedlings under drought stress. *Journal of Integrative*
605 *Agriculture* **17**, 336-347 (2018).
- 606 16. Peter J. Franks, G. D. Farquhar, The Effect of Exogenous Absciscic Acid on Stomatal
607 Development, Stomatal Mechanics, and Leaf Gas Exchange in *Tradescantia virginiana*. *Plant*
608 *physiology* **125**, 935–942 (2001).
- 609 17. L. J. Pillitteri, K. U. Torii, Mechanisms of stomatal development. *Annu Rev Plant Biol* **63**,
610 591-614 (2012).
- 611 18. N. Zoulias, E. L. Harrison, S. A. Casson, J. E. Gray, Molecular control of stomatal development.
612 *Biochem J* **475**, 441-454 (2018).

- 613 19. G. Lin *et al.*, A receptor-like protein acts as a specificity switch for the regulation of stomatal
614 development. *Genes Dev* **31**, 927-938 (2017).
- 615 20. E. D. Shpak, J. M. McAbee, L. J. Pillitteri, K. U. Torii, Stomatal patterning and differentiation by
616 synergistic interactions of receptor kinases. *Science* **309**, 290-293 (2005).
- 617 21. J. S. Lee *et al.*, Direct interaction of ligand-receptor pairs specifying stomatal patterning.
618 *Genes Dev* **26**, 126-136 (2012).
- 619 22. T. Shimada, S. S. Sugano, I. Hara-Nishimura, Positive and negative peptide signals control
620 stomatal density. *Cell Mol Life Sci* **68**, 2081-2088 (2011).
- 621 23. A. L. Rychel, K. M. Peterson, K. U. Torii, Plant twitter: ligands under 140 amino acids enforcing
622 stomatal patterning. *J Plant Res* **123**, 275-280 (2010).
- 623 24. K. Hara, R. Kajita, K. U. Torii, D. C. Bergmann, T. Kakimoto, The secretory peptide gene EPF1
624 enforces the stomatal one-cell-spacing rule. *Genes Dev* **21**, 1720-1725 (2007).
- 625 25. L. Hunt, J. E. Gray, The signaling peptide EPF2 controls asymmetric cell divisions during
626 stomatal development. *Curr Biol* **19**, 864-869 (2009).
- 627 26. S. S. Sugano *et al.*, Stomagen positively regulates stomatal density in Arabidopsis. *Nature* **463**,
628 241-244 (2010).
- 629 27. T. Kondo *et al.*, Stomatal density is controlled by a mesophyll-derived signaling molecule.
630 *Plant Cell Physiol* **51**, 1-8 (2010).
- 631 28. L. Hunt, K. J. Bailey, J. E. Gray, The signalling peptide EPFL9 is a positive regulator of stomatal
632 development. *New Phytol* **186**, 609-614 (2010).
- 633 29. J. S. Lee *et al.*, Competitive binding of antagonistic peptides fine-tunes stomatal patterning.
634 *Nature* **522**, 439-443 (2015).
- 635 30. L. J. Pillitteri, K. U. Torii, Breaking the silence: three bHLH proteins direct cell-fate decisions
636 during stomatal development. *Bioessays* **29**, 861-870 (2007).
- 637 31. S. Ohki, M. Takeuchi, M. Mori, The NMR structure of stomagen reveals the basis of stomatal
638 density regulation by plant peptide hormones. *Nat Commun* **2**, 512 (2011).
- 639 32. U. Gowik, A. Bräutigam, K. L. Weber, A. P. M. Weber, P. Westhoff, Evolution of C₄
640 Photosynthesis in the Genus Flaveria: How Many and Which Genes Does It Take to Make C₄?
641 *The Plant Cell* **23**, 2087-2105 (2011).
- 642 33. J. Mallmann *et al.*, The role of photorespiration during the evolution of C₄ photosynthesis in
643 the genus Flaveria. *Elife* **3**, e02478 (2014).

- 644 34. M. Hronkova *et al.*, Light-induced STOMAGEN-mediated stomatal development in
645 Arabidopsis leaves. *J Exp Bot* **66**, 4621-4630 (2015).
- 646 35. J. Lu *et al.*, Homologous genes of epidermal patterning factor regulate stomatal development
647 in rice. *J Plant Physiol* **234-235**, 18-27 (2019).
- 648 36. L. Xuxin, The Effects for Stomata Developmental on Arabidopsis by Overexpression of Rice
649 EPFL Family Genes. *Molecular Plant Breeding* **15**, 2042- 2047 (2017).
- 650 37. J. Y. Zhang, S. B. He, L. Li, H. Q. Yang, Auxin inhibits stomatal development through
651 MONOPTEROS repression of a mobile peptide gene STOMAGEN in mesophyll. *Proc Natl Acad*
652 *Sci U S A* **111**, E3015-3023 (2014).
- 653 38. Z. Xu, G. Zhou, Responses of leaf stomatal density to water status and its relationship with
654 photosynthesis in a grass. *J Exp Bot* **59**, 3317-3325 (2008).
- 655 39. M. J. Lyu *et al.*, RNA-Seq based phylogeny recapitulates previous phylogeny of the genus
656 Flaveria (Asteraceae) with some modifications. *BMC Evol Biol* **15**, 116 (2015).
- 657 40. A. L. Ming-Ju *et al.*, The Coordination and Jumps along C₄ Photosynthesis Evolution in the
658 Genus Flaveria. *bioRxiv*, 460287 (2018).
- 659 41. K. Billakurthi *et al.*, Transcriptome dynamics in developing leaves from C₃ and C₄ Flaveria
660 species reveal determinants of Kranz anatomy. *bioRxiv*, (2018).
- 661 42. L. Wang *et al.*, Comparative analyses of C₄ and C₃ photosynthesis in developing leaves of
662 maize and rice. *Nat Biotechnol* **32**, 1158-1165 (2014).
- 663 43. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript
664 server. *Nucleic Acids Res* **42**, W320-324 (2014).
- 665 44. D. R. F. Kyte J, A simple method for displaying the hydropathic character of a protein. *Journal*
666 *of Molecular Biology* **157**, 105-132 (1982).
- 667 45. X. J. Min, G. Butler, R. Storms, A. Tsang, OrfPredictor: predicting protein-coding regions in
668 EST-derived sequences. *Nucleic acids research* **33**, W677-680 (2005).
- 669 46. M.-J. A. Lyu *et al.*, The Coordination and Jumps along C₄ Photosynthesis Evolution in the
670 Genus Flaveria. *bioRxiv*, 460287 (2018).
- 671 47. S. A. Shiryev, J. S. Papadopoulos, A. A. Schaffer, R. Agarwala, Improved BLAST searches using
672 longer words for protein seeding. *Bioinformatics* **23**, 2949-2951 (2007).
- 673 48. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput.
674 *Nucleic acids research* **32**, 1792-1797 (2004).

- 675 49. A. Stamatakis, RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with
676 thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).
- 677 50. J. L. J. a. T. F. Ø. Claus Lindbjerg Andersen, Normalization of real-time quantitative reverse
678 transcription-PCR data a model-based variance estimation approach to identify genes suited
679 for normalization, applied to bladder and colon cancer data sets. *Cancer Res* **64**, 5245-5250
680 (2004).
- 681 51. D. P. K. Vandesompele J, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F., Accurate
682 normalization of real-time quantitative RT-PCR data by geometric averaging of multiple
683 internal control genes. *Genome Biology* **3**, research0034.0031–0034.0011 (2002).
- 684 52. D. A. Way, G. G. Katul, S. Manzoni, G. Vico, Increasing water use efficiency along the C₃ to C₄
685 evolutionary pathway: a stomatal optimization perspective. *J Exp Bot* **65**, 3683-3693 (2014).
- 686 53. Q. Chen, Functional analysis of STOMAGEN-like gene in regulating stomatal development in
687 maize. *Master*, (2017).
- 688 54. V. S. R. D. a. M. Santakumari, Stomatal characteristics of some dicotyledonous plants in
689 relation to the C₄ and C₃ pathways of photosynthesis. *Plant and cell physiology* **18**, 935-938
690 (1977).
- 691 55. M. E. K. a. S. A. M. Marc D. Abrams, Relating Wet and Dry Year Ecophysiology to Leaf
692 Structure in Contrasting Temperate Tree Species. *Ecology* **75**, 123-133 (1994).
- 693 56. C. T. a. D. R. Murray, Effects of Elevated Atmospheric CO₂ Concentration on Leaf Anatomy and
694 Morphology in Panicum Species Representing Different Photosynthetic Modes. *International*
695 *Journal of Plant Sciences* **160**, 1063-1073 (1999).
- 696 57. A. S. Soares *et al.*, Adaxial/abaxial specification in the regulation of photosynthesis and
697 stomatal opening with respect to light orientation and growth with CO₂ enrichment in the C₄
698 species Paspalum dilatatum. *New Phytologist* **177**, 186-198 (2007).
- 699 58. I. Gindell, Stomatal Number and Size as Related to Soil Moisture in Tree Xerophytes in Israel
700 *Ecological Society of America* **50**, 263-267 (1968).
- 701 59. V. S. Pathare, N. Koteyeva, A. B. Cousins, Increased adaxial stomatal density is associated with
702 greater mesophyll surface area exposed to intercellular air spaces and mesophyll
703 conductance in diverse C₄ grasses. *The New phytologist* **225**, 169-182 (2019).
- 704 60. J. A. Cardoso, M. Pineda, L. Jimenez Jde, M. F. Vergara, I. M. Rao, Contrasting strategies to
705 cope with drought conditions by two tropical forage C₄ grasses. *AoB Plants* **7**, plv107 (2015).

- 706 61. E. J. Edwards, C. J. Still, M. J. Donoghue, The relevance of phylogeny to studies of global
707 change. *Trends Ecol Evol* **22**, 243-249 (2007).
- 708 62. S. Aubry *et al.*, A Specific Transcriptome Signature for Guard Cells from the C₄ Plant
709 *Gynandropsis gynandra*. *Plant physiology* **170**, 1345-1357 (2016).
- 710 63. S. P. Driscoll, A. Prins, E. Olmos, K. J. Kunert, C. H. Foyer, Specification of adaxial and abaxial
711 stomata, epidermal structure and photosynthesis to CO₂ enrichment in maize leaves. *J Exp*
712 *Bot* **57**, 381-390 (2006).
- 713 64. W.-H. W. Xi-Qing Wang, and Sarah M. Assmann, Differential Responses of Abaxial and Adaxial
714 Guard Cells of Broad Bean to Absciscic Acid and Calcium. *Plant physiology* **118**, 1421–1429
715 (1998).
- 716 65. C. D. Muir, M. À. Conesa, J. Galmés, Independent evolution of ab- and adaxial stomatal
717 density enables adaptation. *bioRxiv*, (2015).
- 718 66. R. J. Webster *et al.*, High C₃ photosynthetic capacity and high intrinsic water use efficiency
719 underlies the high productivity of the bioenergy grass *Arundo donax*. *Sci Rep* **6**, 20694 (2016).
- 720 67. R. S. Caine *et al.*, Rice with reduced stomatal density conserves water and has improved
721 drought tolerance under future climate conditions. *The New phytologist* **221**, 371–384
722 (2018).
- 723 68. U. Mohammed *et al.*, Rice plants overexpressing OsEPF1 show reduced stomatal density and
724 increased root cortical aerenchyma formation. *Scientific reports* **9**, 5584 (2019).
- 725 69. J. Dunn *et al.*, Reduced stomatal density in bread wheat leads to increased water-use
726 efficiency. *J Exp Bot* **70**, 4737-4748 (2019).
- 727 70. J. Hughes *et al.*, Reducing Stomatal Density in Barley Improves Drought Tolerance without
728 Impacting on Yield. *Plant physiology* **174**, 776-787 (2017).
- 729 71. X. Yin *et al.*, CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental
730 gene EPFL9 in rice. *Plant Cell Rep* **36**, 745-757 (2017).
- 731 72. P. J. Vogan, M. W. Frohlich, R. F. Sage, The functional significance of C₃-C₄ intermediate traits
732 in *Heliotropium* L. (Boraginaceae): gas exchange perspectives. *Plant, cell & environment* **30**,
733 1337-1345 (2007).
- 734 73. J. W. Maurice S. B. Ku, Ziyu Dai, Rick A. Scott, Chun Chu, and Gerald E. Edwards,
735 Photosynthetic and Photorespiratory Characteristics of *Flaveria* species. *Plant physiology* **96**,
736 518-528 (1991).

- 737 74. O. Keerberg, T. Parnik, H. Ivanova, B. Bassuner, H. Bauwe, C₂ photosynthesis generates about
738 3-fold elevated leaf CO₂ levels in the C₃-C₄ intermediate species *Flaveria pubescens*. *J Exp Bot*
739 **65**, 3649-3656 (2014).
- 740 75. R. F. Sage, T. L. Sage, F. Kocacinar, Photorespiration and the evolution of C₄ photosynthesis.
741 *Annu Rev Plant Biol* **63**, 19-47 (2012).
- 742 76. C. B. Engineer *et al.*, Carbonic anhydrases, EPF2 and a novel protease mediate CO₂ control of
743 stomatal development. *Nature* **513**, 246-250 (2014).

744

745

746 **Acknowledgements**

747 The work was funded by Strategic Priority Research Program of the Chinese
748 Academy of Sciences (grant number: XDB27020105) and the general program of
749 National Science Foundation of China (31870214). We thank Prof. Rowan F. Sage
750 (University of Toronto) for providing the *Flaveria* cuttings.

751

752 **Author contributions**

753 YYZ and XGZ designed the study. YYZ observed and measured the stomatal
754 characteristic. YYZ carried out the molecular experiment and the treatment of
755 synthetic peptides in vivo. GYC guided the treatment of peptides in vivo. Amy MJL
756 analyzed the one kp RNA-seq data. YYZ and XGZ interpreted results and wrote the
757 article.

758 **Competing interests**

759 The authors declare no competing financial interests.

760

761 Figure Legend

762 Fig 1: Stomatal conductance, density (SD) and its patterning.

763 **a**, Stomatal conductance (g_s) responses to net photosynthesis rate (A) under saturated
764 light. The circles on the graph indicate larger g_s values in different photosynthetic
765 types (Date from (2, 52)). **b**, The average operating g_s and A at saturated light. Error
766 bar indicates s.e.m. **c**, G_s responses to A at the growth light in two *Flaveria* species (*F.*
767 *Robusta*: rob, *F. bidentis*: bid). **d**, **e**, The difference in SD of abaxial surface between
768 species with different photosynthetic types in *Flaveria* genus grown outdoors (**d**) and
769 in greenhouse (**e**). The type I represents C_3 - C_4 (type I), and the type II represents
770 C_3 - C_4 (type II). **f**, **g**, The change of SD of abaxial surface from C_3 to C_4 in *Flaveria*
771 species grown outdoors (**f**) and in greenhouse (**g**). (n=4, biologically independent
772 replicates). **h**, Correlation between SD and g_s under the saturated light. **i**, The
773 photographs of stomatal patterning on abaxial surface in rob, ram and bid. The
774 photographs from the left to right are for rob, ram and bid grown outdoors (top) and in
775 greenhouse (bottom). **j**, **k**, Comparison of guard cell length(l) (**j**) and stomatal index
776 (SI) (**k**) between rob, ram, and bid. Different colours represent plants grown outdoors
777 (blue) or in greenhouse (green). l , (n=20, biologically independent replicates). SI ,
778 (n=3, biologically independent replicates). Error bar indicates s.d. Different letters
779 represent statistically significant differences (one-way ANOVA, Duncan's multiple
780 range test ($\alpha = 0.05$)). The asterisks represent statistically significant differences
781 ($P < 0.05$, t-test, two-tailed distribution, unpaired)

782

783 Fig 2: The Function of *FSTOMAGEN* in controlling the SD

784 **a**, Protein conservative analysis of the homologs of *FSTOMAGEN* in different
785 species. The amino acid regions marked by green line and red line represent the signal
786 peptide and the functional domain of *FSTOMAGEN*, respectively. Red letters
787 represent conserved amino acids in most species, white letters with red backgrounds

788 represent amino acid conserved across all species. **b**, Phylogenetic tree for
 789 FSTOMAGEN. blue represents dicotyledon and red represents monocotyledon. **c**, The
 790 structure of recombinant genes used to transform Arabidopsis. AFSTO represents the
 791 recombination of the signal peptide of STOMAGEN from Arabidopsis and the
 792 functional domain of FSTOMAGEN from *F.robusta*. FSTO represents the intact
 793 STOMAGEN from *F.robusta*. **d**, Photographs showing stomata on a leaf for the wild
 794 type Arabidopsis. Scar bar represents 10 μ m. **e**, Photographs of stomata in a leaf from
 795 Arabidopsis with AFSTO over-expressed. Scar bar represents 10 μ m. The bar graph
 796 shows the difference in *SD* between WT and AFSTO overexpression in Arabidopsis. **f**,
 797 Photographs of stomata in a leaf from Arabidopsis with FSTO over-expressed. Scar
 798 bar represents 10 μ m. The bar graph shows the difference in *SD* between WT and
 799 FSTO overexpression in Arabidopsis. **g**, Photographs of leaf surface taken differential
 800 interference contrast microscope showing the increased stomata density in *F. bidentis*
 801 (*bid*) leaves with FSTOMAGEN applied. Scale bar represents 20 μ m. **h**, The stomatal
 802 density for *bid* leaves with FSTOMAGEN applied at different concentrations (n=3,
 803 biologically independent replicates). **i**, The *SD* in leaves of *bid* and Arabidopsis
 804 treated with FSTOMAGEN and STOMAGEN respectively at different concentration.
 805 The percentage indicates the percentage increase of *SD* in the treated group compared
 806 to that in the untreated group. Error bar indicates s.d. The asterisks represent
 807 statistically significant differences ($P<0.05$, t-test, two-tailed distribution, unpaired).

808

809 Fig 3: Transcript abundance of stomatal developmental related genes between
 810 species with different photosynthetic types in the *Flaveria* genus.

811 **a**, A model of the signaling pathway for the stomatal development in Arabidopsis. A
 812 protodermal cell (PDC, white) becomes a meristemoid mother cell (MMC, orange),
 813 which divides asymmetrically into two daughter cells, with one being a meristemoid
 814 (M, red) and the other being a stomatal-lineage ground cell (SLGC). The M cell
 815 develops into a guard mother cell (GMC, yellow); the SLGC develops into a

816 pavement cell (PC). The GMC conducts an symmetrical division to produce two
817 equal-sized guard cells (GC, green). EPF1, EPF2 and STOMAGEN competitively
818 bind to ER and TMM, which can deliver a signal to the YDA MAPK cascade. SPCH,
819 MUTE and FAMA ultimately are inactivated through phosphorylation by the YDA
820 MAPK cascade. **b**, Comparison of the transcript abundance of stomatal development
821 related genes between C₃, C₃-C₄, C₄-like and C₄ specie in *Flaveria* genus. J represents
822 samples from the juvenile leaves, M represents samples from mature leaves. C₃, C₃-C₄,
823 C₄-like and C₄ were arranged from the left to right sequentially. **c**, Relative expression
824 levels determined by RT-qPCR and expression level determined by RNA-seq of
825 FSTO (FSTOMAGEN), FEPP1 and FEPP2 in *rob* and *bid* (n=4, biologically
826 independent replicates). **d**, Protein conservative analysis of the homologs of
827 FSTOMAGEN between different species in the *Flaveria* genus. The amino acid
828 regions marked by green line and red line represent the signal peptide and the
829 functional domain of STOMAGEN, respectively. Red letters represent conserved
830 amino acids in most species; white letters with red background represent amino acids
831 conserved across all species. **e**, Expression levels of *FSTOMAGEN* at the different
832 developmental stages of leaves in *rob* and *bid*. **f**, Comparison of the expression levels
833 of *FSTOMAGEN* between different species with different photosynthetic types in
834 *Flaveria* genus. **g**, Comparison of relative expression levels of *FSTOMAGEN* between
835 *rob*, *ram* and *bid*. (n=3, biologically independent replicates). **h**, Combined data show
836 the relation between *SD* (n=4, biologically independent replicates) and expression
837 levels of *FSTOMAGEN* (n≥3, biologically independent replicates). The relative
838 expression in figure (**c**), (**f**), (**h**) indicate the measurements with RT-qPCR, and the
839 others was measured with RNA-seq. The geometric means of *EF1a* and *ACT7* were
840 used as the reference in the calculation of expression levels for other genes. Error bars
841 indicate s.d. The asterisks represent statistically significant differences ($P<0.05$, t-test,
842 two-tailed distribution, unpaired)

843

844 Fig 4, Comparison of stomatal pattern and expression levels of stomatal development
845 related genes between rice (C₃) and maize (C₄).

846 **a**, Stomatal characteristics for rice and maize. *SD*, (n=8, biologically independent
847 replicates); *l*, (n=25, biologically independent replicates); *SI*, (n=3, biologically
848 independent replicates); *g_s*, (n=8, biologically independent replicates). Error bar
849 indicates s.d. The asterisks represent statistically significant differences ($P<0.05$, t-test,
850 two-tailed distribution, unpaired). **b**, Photographs of stomatal pattern in rice and
851 maize. Scale bars represent 50 μ m. **c**, Comparison of transcript abundance of stomatal
852 development related genes between rice and maize. The developmental stages of
853 leaves were divided into 11 and 15 parts in rice and maize, respectively. R represents
854 the leaves of rice, M represents the leaves of maize. **d**, Comparison of the expression
855 levels of *STOMAGEN* homologs between rice and maize. There are two and three
856 paralogs of *STOMAGEN* in rice and maize, respectively. We summed up their
857 expression levels to calculate the overall *STOMAGEN* homologs expression levels in
858 maize and rice. **e,f**, Comparison of the differences in correlation between *SD* and
859 relative mRNA levels of the homologs of *STOMAGEN* for Arabidopsis and rice (**e**),
860 Arabidopsis and maize (**f**). *OsSTOMAGEN* represents the homologs of *STOMAGEN*
861 in rice. *ZmSTOMAGEN* represents the homologs of *STOMAGEN* in maize.

862

863 Legends for Supplemental Figures

864 Supplementary fig 1. Comparison of *SD* between C₃ and C₄ species. **a**, Comparison of
865 *SD* on abaxial surface of leaves between 70 C₃ and 47 C₄ species, with data collected
866 from literature (see details in the discussion section). **b**, Comparison of *SD* on adaxial
867 surface of leaves between 26 C₃ and 27 C₄ species with data collected from literature
868 (see details in the discussion section). Error bar indicates s.d. The asterisks represent
869 statistically significant differences ($P<0.05$, t-test, two-tailed distribution, unpaired).

870 Supplementary fig 2: *SD* on the adaxial surface of leaves in *Flaveria*. **a, b**, The
 871 changes of *SD* on the adaxial leaf surface from *C*₃ to *C*₄ species in the *Flaveria* genus
 872 grown outdoors (**a**) and in greenhouse (**b**). (n=3, biologically independent replicates).
 873 **c, d**, The difference in *SD* in species with different photosynthetic types grown
 874 outdoors (**c**) or in greenhouse (**d**). Different letters represent statistically significant
 875 differences (one-way ANOVA, Duncan's multiple range test ($\alpha = 0.05$)).

876 Supplementary fig 3: *SD* in *rob* (*C*₃), *ram* (*C*₃-*C*₄) and *bid* (*C*₄) grown from seeds. The
 877 *SD* on abaxial (ab) (n=5, biologically independent replicates) and adaxial (ad) (n=4,
 878 biologically independent replicates) surfaces were observed in *rob*, *ram* and *bid*
 879 grown in greenhouse. Different colours represent different photosynthetic types.
 880 Different letters represent statistically significant differences (one-way ANOVA,
 881 Duncan's multiple range test ($\alpha = 0.05$)).

882 Supplementary fig 4: Conservative analysis of amino acid sequences for EPF1 and
 883 EPF2. **a**. Comparison of amino acid sequences of homologs of EPF1 between
 884 different *Flaveria* species and *Arabidopsis*. **b**. Comparison of amino acid sequences of
 885 homologs of EPF2 between *Flaveria* species and *Arabidopsis*. White letters with red
 886 background represent amino acids conserved across all species. Red letters represent
 887 amino acids with similar biochemical property.

888 Supplementary fig 5: Hydrophobicity analysis of proteins. **a**, the protein of stomagen
 889 (left) and the protein of FSTOMAGEN (right). **b**, the protein of EPF1 (left) and the
 890 protein of FEPEF1 (right). **c**, the protein of EPF2 (left) and the protein of FEPEF2
 891 (right).

892 Supplementary fig 6: Assessment of the stability of expression levels of reference
 893 genes. **a,b**, Comparison of the stability of expression levels of candidate reference
 894 genes in the *Flaveria* species by RNA-seq. **c**, Gel electrophoresis showing the
 895 specificity of the primers for candidate reference genes. From left to right, *ACT7*,
 896 *EF1a*, *J3*, *PNdO*, *CPN20*, *UBC9*, *GAPDH*, *F2N1.5*, *PSRP4*, *CACMS*, *HEME2*, and
 897 *UBQ11*. **d**, The coefficient of variations (CV) for the candidate reference genes. **e**,

898 The expression stability for candidate reference genes calculated by Genorm. **f**, The
899 expression stability for candidate reference genes calculated by Normfinder.

900 Supplementary fig 7: The expression levels of stomatal development related genes in
901 *Flaveria* by RNA-seq. Expression levels of stomatal development related genes in
902 species with different photosynthetic types in the *Flaveria* genus.

903 Supplementary fig 8: Relative expression levels of FSTOMAGEN and *SD* for rob (C_3)
904 under different light conditions. (A) Relative expression levels for FSTOMAGEN
905 under two different light conditions. (B) *SD* under two different light conditions.
906 Natural light: the maximal photosynthetic photon flux density in the greenhouse was
907 about 1300~1400 $\text{mmol m}^{-2} \text{s}^{-1}$. Light in the phytotron: 20~100 $\text{mmol m}^{-2} \text{s}^{-1}$.
908 Reference gene used to calculate the expression level was *ACT7*. The asterisks
909 represent statistically significant differences ($P < 0.05$, t-test, two-tailed distribution,
910 unpaired).

911 Supplementary fig 9: Relative expression levels of stomata development related genes
912 determined with different reference genes and different primers. **a**, **b**, **c**,
913 Quantification with different primers and the same *EFla* reference gene. Relative
914 expressions shown in **a and b** were determined using the same primer pairs while that
915 in **c** used other primer pairs (See details of the primers used in Table S2). **d**, Relative
916 expression levels determined with the gene *J3* used as the reference gene.

917 Supplementary fig 10: Transcript abundance of genes in the C_4 cycle and C_4 related
918 transporters. J represents juvenile leaves and M represents mature leaves. Different
919 colours represent species with different photosynthetic types in the *Flaveria* genus.
920 Red colour represents C_3 species, grey colour represents C_3 - C_4 species, green colour
921 represents C_4 -like species, and blue colour represents C_4 species.

922 Supplementary fig 11: Alignment of the functional domain of *STOMAGEN* with
923 transcripts from different *Flaveria* species by Blast. From the top to bottom, *F*.

924 *robusta*, *F. pringlei*, *F. angustifolia*, *F. ramosissima*, *F. vaginata*, *F. australasica*, *F.*
925 *bidentis*.

926 Supplementary fig 12: The increase of *SD* in cotyledon of *F. bidentis* (*bid*) with in
927 vitro application of FSTOMAGEN. Photographs of the leaf surface in the untreated
928 (control) and treated bid plants. scale bar, 20 mm.

929

930

931

932

933 Supplementary table 1: Statistics for the mapping of the RNA-seq data of the juvenile
934 and mature leaves in different species from the *Flaveria* genus.

935 Supplementary table 2: primers used in this study

936