Relationships of Leaf Net Photosynthesis, Stomatal Conductance, and Mesophyll Conductance to Primary Metabolism: A Multispecies Meta-Analysis Approach[1]

Jorge Gago*, Danilo de Menezes Daloso, Carlos María Figueroa, Jaume Flexas, Alisdair Robert Fernie, and Zoran Nikoloski

Research Group on Plant Biology under Mediterranean Conditions, Departament de Biologia, Universitat de les Illes Balears, 07122 Palma de Mallorca, Illes Balears, Spain (J.G., J.F.); Central Metabolism Group, Molecular Physiology Department, Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14476 Golm, Germany (J.G., D.d.M.D., A.R.F.); System Regulation Group, Metabolic Networks Department, Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14476 Golm, Germany (C.M.F.); and Systems Biology and Mathematical Modeling Group, Molecular Physiology Department, Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14476 Golm, Germany (Z.N.)

ORCID IDs: 0000-0003-4047-0480 (C.M.F.); 0000-0003-2671-6763 (Z.N.).

Plant metabolism drives plant development and plant-environment responses, and data readouts from this cellular level could provide insights in the underlying molecular processes. Existing studies have already related key in vivo leaf gas-exchange parameters with structural traits and nutrient components across multiple species. However, insights in the relationships of leaf gas-exchange with leaf primary metabolism are still limited. We investigated these relationships through a multispecies meta-analysis approach based on data sets from 17 published studies describing net photosynthesis (A) and stomatal (g_s) and mesophyll (g_m) conductances, alongside the 53 data profiles from primary metabolism of 14 species grown in different experiments. Modeling results highlighted the conserved patterns between the different species. Consideration of species-specific effects increased the explanatory power of the models for some metabolites, including Glc-6-P, Fru-6-P, malate, fumarate, Xyl, and ribose. Significant relationships of A with sugars and phosphorylated intermediates were observed. While g_s was related to sugars, organic acids, *myo*-inositol, and shikimate, g_m showed a more complex pattern in comparison to the two other traits. Some metabolites, such as malate and Man, appeared in the models for both conductances, suggesting a metabolic coregulation between g_s and g_m. The resulting statistical models provide the first hints for coregulation patterns involving primary metabolism plus leaf water and carbon balances that are conserved across plant species, as well as species-specific trends that can be used to determine new biotechnological targets for crop improvement.

Plant metabolism drives plant development alongside a considerable number of plant-environment interactions (Brunetti et al., 2013). This is largely accomplished through the joint modulation of enzyme and metabolite levels. Characterization of metabolites can be carried out by different profiling technologies, comprising the field of metabolomics (Brunetti et al., 2013). Recent advances in metabolic profiling technologies have led to the identification and quantification of 1,000 to 2,000 metabolites, while the number for the whole plant kingdom is estimated to be ~200,000, with a single plant species including up to a few thousand metabolites (e.g. ~5,000 for Arabidopsis [*Arabidopsis thaliana*]; Obata and Fernie, 2012). A metabolic profile comprises data about (changes in) the levels of metabolites in a given system. The data gathered via metabolic profiling may provide insights in the regulation of metabolic networks. Additionally, they can be used in elucidating gene functions, designing crop breeding strategies, understanding plant growth and plasticity, and characterizing plant-community responses to biotic and abiotic stresses (Meyer et al., 2007; Obata and Fernie, 2012; Brunetti et al., 2013).

In addition, metabolic profiles obtained from these technologies have been used to build statistical models of complex phenotypic traits. For instance, statistical models of good performance have been developed for biomass, one of the agronomically important complex plant traits (Meyer et al., 2007; Sulpice et al., 2013). The

[1] This work was supported by the Max Planck Society (A.R.F., C.M.F., J.G., Z.N.), by a CNPq (National Council for Technological and Scientific Development, Brazil) scholarship (D.d.M.D.), and by Plan Nacional, Spain (contracts BFU2011-23294 and CTM2014-53902-C2-1-P to J. F. and J.G.).

* Address correspondence to xurxogago@gmail.com.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Jorge Gago (xurxogago@gmail.com).

J.G., A.R.F, and Z.N. conceived the project and research plans; J.G. performed most of the data compilation; J.G and Z.N. performed the statistical modeling; J.G., D.d.M.D., C.M.F., J.F., A.R.F., and Z.N. performed analysis, discussed results, and wrote the article.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.01660

success of the modeling approach ultimately depends on the ability to translate the statistical relationships in future crop breeding purposes (Fernie, 2012). While the first attempts of modeling complex agronomically relevant traits based on metabolic profiles were carried out in the model plant Arabidopsis (Timm et al., 2012; Sulpice et al., 2013), recent approaches have expanded these efforts to other plant species, including important crops (Araújo et al., 2011b; Witt et al., 2012; Riedelsheimer et al., 2013).

Biomass depends on the underlying metabolic profile of the plant system, while this profile may reflect different physiological, structural, and compositional traits related to biomass. For instance, advances in the study of leaf photosynthesis has resulted in multispecies and biome data sets that allow us to prove the relationships between leaf mass area, specific leaf area, and nutrient content (e.g. leaf nitrogen and phosphorus concentrations; Reich et al., 1999; Wright et al., 2004; Kattge et al., 2011; Walker et al., 2014). These studies identified several clear trends, namely, maximum photosynthesis (A_{max}) correlates positively with nitrogen and phosphorous leaf content and negatively with leaf lifespan and leaf mass area (Reich et al., 1999), independently of biome, growth form, or plant functional type (Wright et al., 2004). Recently, it was reported that the maximum carboxylation velocity of Rubisco (V_{cmax}) was positively correlated with nitrogen content, while phosphorous increases the sensitivity of the relationship (Walker et al., 2014).

Despite these advances, there are only few studies that have examined the relationships between metabolic profiles and plant in vivo photosynthetic rates and conductances (Aranjuelo et al., 2011; Warren, 2011; Warren et al., 2011, 2012). Elucidating relations between metabolites and conductances is particularly relevant, since conductances are the diffusive physiological determinants that mainly drive the final plant productivity and plant's relation with the environment. Identifying statistical relationships between metabolism and photosynthesis can provide important insights into how changes in particular metabolites, potentially achievable via metabolic engineering, may modulate photosynthesis and, in turn, improve plant growth, water use efficiency, and stress tolerance.

Carbon and water fluxes inside the leaf follow Fick's First Law of diffusion for CO_2 assimilation rates (A) and transpiration (E):

$$A = g_{sc} \cdot (C_a - C_i)$$

$$E = g_{sw} \cdot (W_i - W_a)$$

where g_{sc} and g_{sw} are the stomatal conductance to CO_2 and water vapor; C_a and C_i are the concentrations of CO_2 in the atmosphere and the substomatal cavity; and W_i and W_a are the concentrations of water and CO_2 vapor in the substomatal cavity and the atmosphere. Clearly, stomata are the common structures for carbon and water exchange ($g_{sc} = g_{sw}/1.6$) between the inner part of the leaf and the atmosphere, processes that ultimately drive photosynthesis. In addition to carbon and water exchange, there are other limiting processes inside the leaf. One of these involves the different consecutive barriers to CO_2 flux from the substomatal cavities to the carboxylation sites in the chloroplast stroma and is characterized by the mesophyll conductance (g_m; Flexas et al., 2013). Concerning g_m at steady state conditions Equation 1 can be rearranged as follows:

$$A = g_m \cdot (C_i - C_c)$$

where C_c is the CO_2 concentration at the carboxylation site of Rubisco. Taking into account the previous statements, C_c highly depends on the CO_2 supply, limited by the diffusive parameters (g_s and g_m), but also on the CO_2 consumption by Rubisco, driven by the maximum carboxylation rates (V_{cmax}) and the photosynthetic metabolism (Gago et al., 2014). Indeed, the widely used photosynthetic model (Farquhar et al., 1980) considers that photosynthesis is limited by the biochemical relations of the maximum carboxylation capacity of Rubisco (V_{cmax}), maximum rate of electron transport (J_{max}), and the use of triose-phosphates (Diaz-Espejo et al., 2012). The first limitation (operating mostly under CO_2 limiting conditions) is related to V_{cmax} and depends on the kinetic properties of Rubisco (at saturating ribulose-bisphosphate concentrations) as well as O_2 and CO_2 concentrations. The second limitation (which is more important at CO_2 saturating conditions) is based on the assumption that photosynthesis depends on the rates of photosynthetic electron transport to produce NADPH and ATP, related to J_{max}. The third limitation can be considered as a product of feedback regulation and is related to the availability of inorganic phosphate needed for ribulose-bisphosphate regeneration, which depends on the relative rates of triose-phosphates production and inorganic phosphate release during Suc and starch synthesis (Diaz-Espejo et al., 2012).

Metabolite profiling has only recently been linked to in vivo leaf gas exchange (Nunes-Nesi et al., 2011; Warren et al., 2011, 2012; Araújo et al., 2011b; Aranjuelo et al., 2013; Medeiros et al., 2015). The stomatal pore is the first diffusional barrier for the influx of CO_2 for photosynthesis. The magnitude of the stomatal pore can directly affect net photosynthetic rate by limiting the amount of substrate (CO_2) for fixation. Conversely, the reduction of the internal CO_2 concentration (C_i) by photosynthesis and the transport and accumulation of photosynthesis-derived metabolites in the surrounding guard cells is known to regulate stomatal movements (Lawson and Blatt, 2014). Given this interplay between stomatal movements and photosynthesis, it is not surprising that both processes are tightly coregulated. However, it is important to note that guard cells are known to respond to a wide range of both endogenous and environmental signals (Kim et al., 2010). As a result, photosynthesis is not the only process that influences stomatal movements. It is also important to

consider the complex interaction with mesophyll and biochemical limitations, which may explain why the relationship between A and g_s is not always unique (Messinger et al., 2006). Due to the high expression of ATP-dependent ion transporters (H^+-ATPase) found in guard cells (Bauer et al., 2013) and their high metabolic activity, guard cells have a relatively large energetic demand (Zeiger et al., 2002; Medeiros et al., 2015). However, even at high respiratory rates (Vani and Raghavendra, 1994; Araújo et al., 2011b), the amount of energy required for the regulation of stomatal movements seems to be greater than the synthesis capacity of guard cells. This fact connects the metabolism of mesophyll cells with stomatal movements since the former have been shown to support the metabolism of guard cells by providing metabolites and ATP (Zeiger et al., 2002; Wang et al., 2014). Furthermore, there are several pieces of evidence showing that accumulation of photosynthesis-derived metabolites, especially sugars and organic acids, in the apoplast space of guard cells can regulate stomatal movements (Lu et al., 1995; Kang et al., 2007; Lee et al., 2008; Araújo et al., 2011b; Fujita et al., 2013). Nevertheless, the mechanisms underlying the modulation of g_s, g_m, and A by sugars and organic acids are still unknown. Moreover, it is not clear whether these processes are conserved across plant species.

Mesophyll conductance, g_m, a limiting factor for CO_2 diffusion to the carboxylation sites in the stroma, is usually tightly coregulated with the stomatal conductance, g_s (Flexas et al., 2013). However, in vivo measurements of g_m are not as straightforward as those of A and g_s. The reasons for this lack of comparability is that g_m can be estimated based on three distinct models employing different theoretical approaches, methodologies, and assumptions (Pons et al., 2009; Flexas et al., 2012). These approaches comprise (1) the "Harley" method, combining chlorophyll fluorescence and gas-exchange data (Harley et al., 1992); (2) the "Ethier" curve-fitting method (Ethier and Livingston, 2004); and (3) the "Evans" method, employing an online carbon isotope discrimination methodology (Evans et al., 1986). There is another approach to estimate g_m in a manner based on leaf anatomical properties, which agrees with in vivo approaches in several angiosperms as well as in several fern species (Tosens et al., 2012; Tomás et al., 2013; Carriquí et al., 2015). However, several concerns were recently published about the reliability and uncertainties of the different in vivo methodologies (Tholen et al., 2012; Sun et al., 2014). Moreover, a wide range of parameters can largely affect g_m since, in addition to the leaf anatomical properties (leaf and cell wall thickness and chloroplast surface exposure to mesophyll air spaces), biochemical processes that are not yet fully understood are likely involved (e.g. related to aquaporins and carbonic anhydrase activities; Flexas et al., 2015). Furthermore, the usual tight coregulation between g_m and the other physiological traits (A and g_s) leads to an extremely complex scenario for the assessment of this trait (Flexas et al., 2013; Gago et al., 2014).

Therefore, photosynthesis is tightly linked to the rest of metabolism through coregulation. The challenge is then to rely on the existing publicly available data sets from metabolomic and ecophysiological studies to identify and further dissect the interactions between A, g_s, and g_m. In this context, we carried out a multispecies meta-analysis aiming to identify and interpret statistical relationships conserved or species-specific through multivariate regression modeling (instead of the classical pairwise correlation approach traditionally employed in plant physiology literature) between leaf primary metabolites and A, g_s, and g_m, using data on these parameters alongside metabolic profiles from several species grown under different conditions. The resulting models can be used to determine new biotechnological targets for crop improvement and can increase our understanding about metabolic network interactions and responses in leaf water and carbon balances conserved across plant species.

RESULTS

Relationship between Net Photosynthesis and Conductances

In our compiled data set, which includes 14 species grown under optimal and stress conditions (water stress and nutrient deficiency) and/or transgenic lines, each of the three simultaneously assessed physiological traits varied in a wide range: net photosynthetic rate, A, from 0.8 to 30.6 μmol CO_2 m^{-2} s^{-1}; stomatal conductance, g_s, from 0.01 to 0.62 mol CO_2 m^{-2} s^{-1}; and mesophyll conductance, g_m, from 0.07 to 0.42 mol CO_2 m^{-2} s^{-1} (Supplemental Data Set S1). A statistically significant positive correlation between A and g_s (P value < 0.001) was observed after pooling all data (Fig. 1A). The correlation between A and g_m was not as strong when pooling all data; instead, only data from control treatments showed a significant correlation (Fig. 1B). Moreover, the relationship between g_s and g_m was very disperse and nonsignificant (Fig. 1C), suggesting uncoupling of the two traits that may depend on the experimental conditions, in the case of several *Acacia* and *Eucalyptus* species (Warren, 2011; Warren et al., 2011).

Photosynthesis and Conductances Driven by Primary Metabolism

Our assembled data set on net photosynthesis and both types of conductances contained the profiles of 53 metabolites. These data allowed us to build statistical models for the metabolites, as responses, in terms of the net photosynthesis and conductances as predictors. To this end, for each metabolite, we tested a linear model with A, g_s, and g_m as independent variables without grouping for the genetic background (i.e. species); this first modeling strategy coincides with classical multivariate regression. We also considered the six expanded models that, in addition to the fixed effects for the three

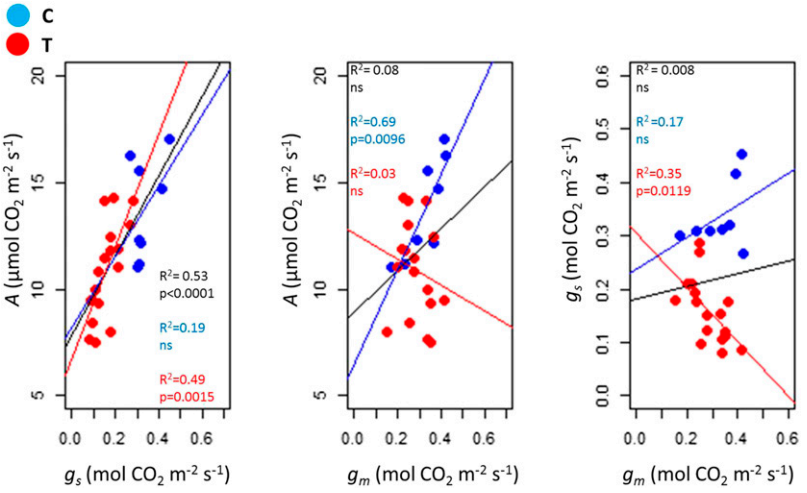


Figure 1. Relationships between net photosynthesis (A) and stomatal conductance (g_s; A), net photosynthesis (A) and mesophyll conductance (g_m; B), and both conductances (g_s and g_m; C) in a multispecies data set (Supplemental Data Set S1). Data were pooled for different species under "control" (blue dots) or "treatment" (red dots) conditions. Species analyzed in the plots include *Acacia maidenii*, *Acacia mangium*, *Eucalyptus arenacea*, *Eucalyptus delegatensis*, *Eucalypus globulus*, *Eucalyptus nitens*, *Eucalyptus regnans*, and *Eucalyptus socialis*. Data were extracted from Warren (2011) and Warren et al. (2011).

independent variables, included either only random intercepts or both random intercept and random slope for each of the independent variables (i.e. A, g_s, and g_m) to determine if the relationships are conserved or species specific between them and each metabolite (see Fig. 2 for an illustration of the procedure). We then used ANOVA for model selection, allowing us to check if random-effect models for each metabolite have larger explanatory power in comparison to the model without random effect. For each of the selected models, we determined its (marginal) coefficient of determination and performed additional tests for the fixed effects to aid the interpretation (Supplemental Data Set S2).

Inspection of the distribution of the coefficients of determination (i.e. percentage of variance in the response is explained by the predictors) for the models without random effects (conserved pattern) indicated that they were in the range between 0.42% (Fru-6-P [F6P]) and 45.01% (malate) with mean of 11.00% and median of 10.02% (Supplemental Fig. S1). Models for 26 metabolites exhibited coefficients of determination with values of at least 10%, while 47 metabolites were at least 20%.

A conserved pattern independent of species-specific behavior was observed for some metabolites for each of the three physiological traits. Interestingly, A exhibited a significant positive coefficient (0.03) only in the model for Fru (P value = 0.037, F-test; Fig. 3; Supplemental Data Set S2). On the other hand, g_s appeared with positive significant coefficients in the models for malate and Man (P values = 0.001 and 0.011, respectively, F-test), but with negative significant coefficients in the models for citrate, shikimate, and inositol (P values = 0.018, 0.038, and 0.009, respectively, F-test; Fig. 3; Supplemental Data Set S2). In the case of g_m, significant coefficients (P values < 0.05) were observed in models for many more metabolites than in the models for the previous traits (19 in total), including Arg, γ-aminobutyric acid (GABA), Gln, His, Phe, Pro, Tyr, Val, dehydroascorbate, hydroxybenzoic acid, malate, maleate, arabinose, Gal, Glc, maltose, Man, raffinose, and mannitol (all negative, except for maltose; Fig. 3; Supplemental Data Set S2).

The lack of good explanatory power for these models suggested that the genetic background may play a role, which is next examined by means of mixed-effects models. Models with random effects for both the intercept and the slope of A were selected for nine of the 53 analyzed metabolites (previously only one showed a correlation with this trait). These metabolites were His, Ile, Leu, Lys, Ser, gluconate, Xyl, Glc-6-P (G6P), and F6P. Including random effects considerably increased the marginal coefficients of determination in comparison to the simple models with fixed factors described above: the minimum was 23.44% (Leu) and the maximum was 89.69% (Ile), followed by G6P and F6P (>89%) with mean of 66.71% and median of 70.75% (Fig. 3; Supplemental Data Set S2), indicating that these relationships tend to be modulated in a species-specific manner. The largest variances in the random intercept and slope (higher species-specific role) for A were observed in the models of His, Lys, and gluconate. The only negative significant fixed effects were found in the models for Ile (g_s, P value = 0.0005; g_m, P value = 0.016), while positive significant fixed effects were found for G6P (g_s, P value = 0.0006; g_m, P value = 0.019) and F6P (g_m, P value = 0.002). Interestingly, A did not show any significant fixed effect in the models with random slope and intercept for this variable (Supplemental Data Set S2).

Models with random effects for both the intercept and slope of g_s were selected for 12 metabolites (just five were previously selected with fixed effects): Asp, Cys, GABA, Glu, Phe, Tyr, hydroxybenzoic acid, gluconate, malate, maleate, Man, and mannitol. However, the marginal coefficients of determination did not increase considerably, unlike the case of random effects for A, in comparison to the simple models considered above: The minimum was 0.07% (mannitol) and the maximum was 5.45% (malate), followed by Asp and hydroxybenzoic acid (>4.20%) with mean of 2.11% and median of 1.28% (Fig. 3; Supplemental Data Set S2). The largest variances in the random intercept and slope (for g_s) were observed in the models of gluconate, Cys, Phe, and Tyr, which are highly species specific. Since all are

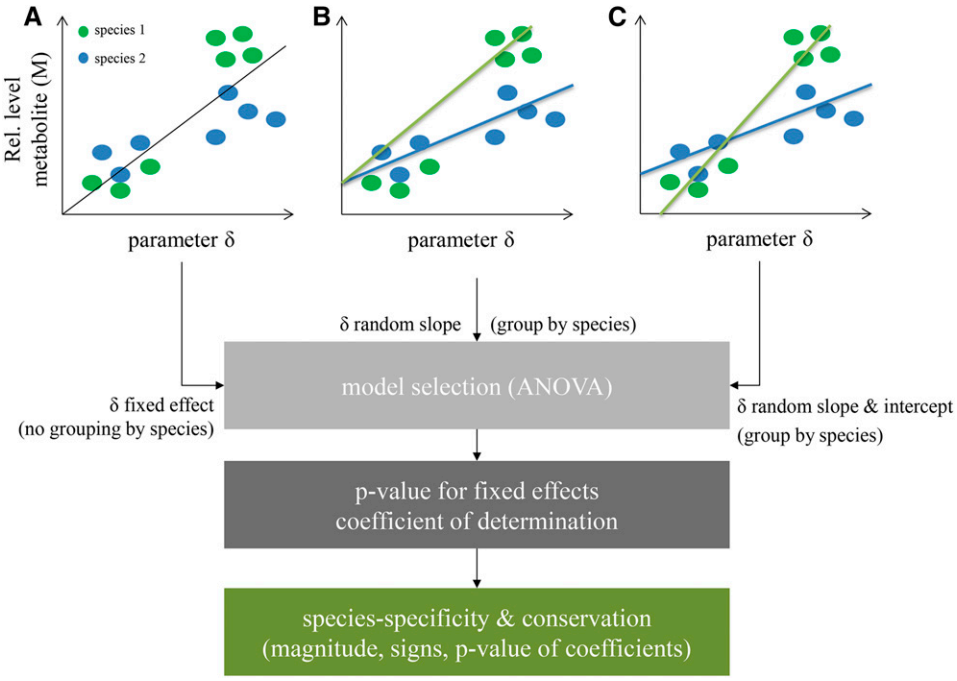


Figure 2. Illustration of the mixed-effects modeling strategy. Hypothetical model with a single predictor (parameter δ), on the *x* axis, for the relative level of metabolite (M), on the *y* axis. The data are gathered from two species, S1 and S2, denoted in green and blue, respectively. A, Model with fixed effect only; this model amounts to regressing the relative metabolite levels of M on parameter δ without grouping for species; the regression line is shown in black. B, Model with fixed intercept, but random slopes for the parameter δ for each of the species S1 and S2; the two regression lines (green and blue, for species S1 and S2, respectively) have the same intercept but different slopes. C, Model with random intercept and random slopes for the parameter δ for each of the species S1 and S2. In our study, the models are multivariate and include A, g_s, and g_m. This results in different models based on the combinations of fixed and random effects for the three parameters used. Models are selected based on ANOVA, and *P* values and coefficient of determination are calculated. Interpretation is then done with respect to the magnitude, sign, and *P* value of the fixed effect as well as the species specificity (due to the consideration of random effects due to the grouping by species).

closely related to oxidative stress protectants it seems reasonable to interpret that this finding reflects the different strategies for protection against reactive oxygen species in the studied plants. In all models, g_m also had a significant fixed effect (except for Asp, gluconate, and mannitol), which was not the case for A and g_s. This effect was negative in all discussed models, except for gluconate and mannitol (Supplemental Data Set S2).

Models with random effects for both the intercept and the slope of g_m were selected for 28 metabolites (Arg, Cys, GABA, Glu, Gln, His, Ile, Leu, Lys, Pro, Ser, Thr, Tyr, Val, and fumarate), with random slopes for A and g_s for most of these metabolites, but showed significant fixed effects for g_m. However, the marginal coefficients of determination were the smallest (smaller than 0.07%) in comparison to those with random slope for A and g_s. The largest variances in the random intercept and slope (for g_m) were observed in the models of Arg, Cys, His, and Thr. The significant fixed effect differed between metabolites. For instance, there was no significant fixed effect of A; fixed effect of g_s was significant for malate, while g_m exhibited significant fixed effect for Arg, GABA, Gln, His, Pro, Tyr, Val, hydroxybenzoic acid, malate, maleate, arabinose, Glc, Man, and mannitol (Supplemental Data Set S2).

Altogether, the analysis based on linear mixed-effects models indicated that the inclusion of random slope for A considerably improved the explanatory models for a limited number of amino acids and sugars. In addition, g_m exhibited significant fixed effects in the majority of the models only upon inclusion of random intercepts and slopes for A and g_s. Furthermore, almost all of the metabolites whose models included a random slope for A or g_s could also be modeled with random slope for g_m, although at expense of explanatory power. This points to the interplay between A, g_s, and g_m; however, models with interactions could not be considered due to lack of data.

Photosynthesis and Conductances Driven by Primary Metabolism

In the following analysis, we examined to what extent the profiles of A, g_s, and g_m as multiple responses (dependent variables) can be explained by the metabolites as predictors (i.e. independent variables). To this end, we used partial least squares (PLS) for building the statistical models. The model of good (root) mean squared error of prediction [(R)MSEP] and coefficient of multiple determination (R^2) bias-corrected via

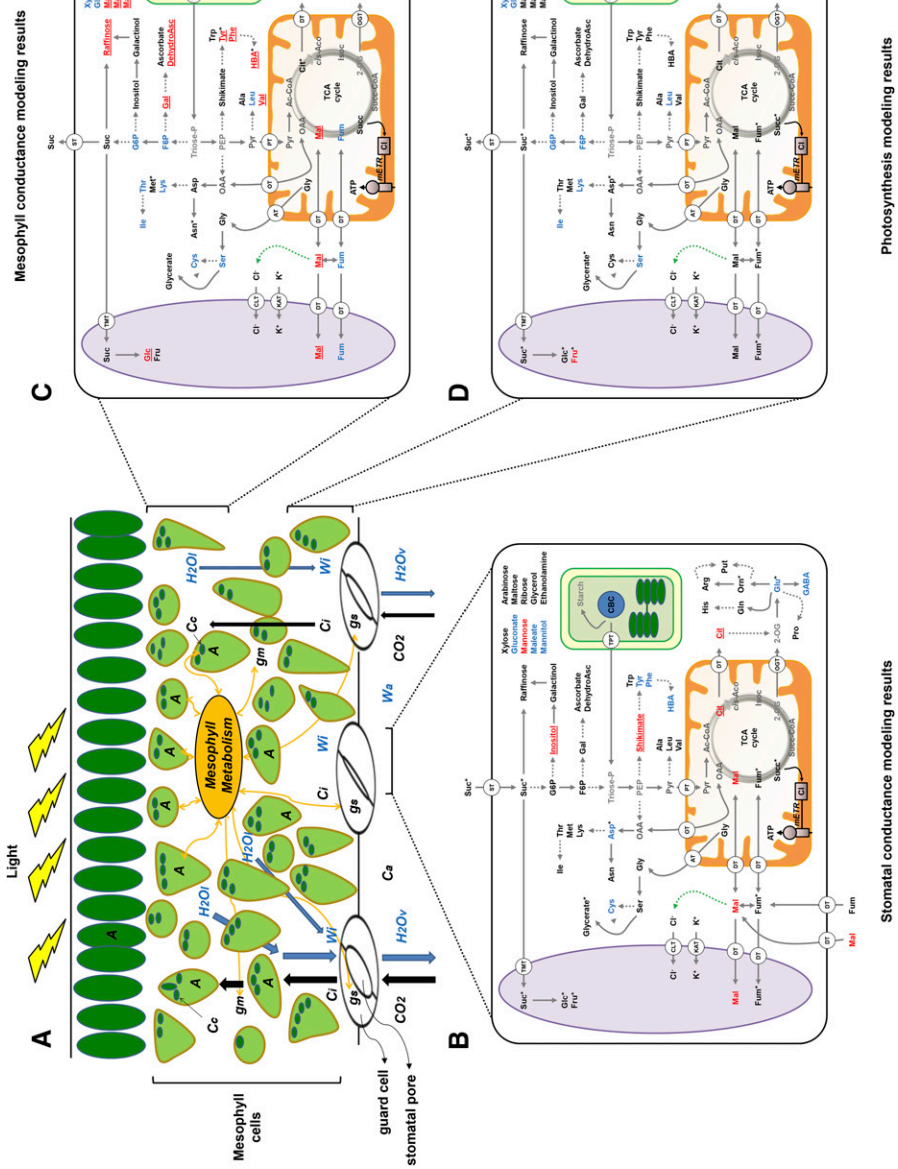


Figure 3. A, Schematic diagram for CO_2 and H_2O diffusion pathways along a leaf cross section. H_2O_l and H_2O_v, water at liquid and gaseous state; C_a, ambient CO_2 concentration; C_i, CO_2 substomatal cavity concentration; C_c, CO_2 concentration at the site of Rubisco carboxylation within the chloroplastic stroma; g_s, stomatal conductance; g_m, mesophyll conductance; W_l and W_a, concentration of water vapor at the substomatal cavity and the atmosphere. B, Relevant metabolic pathways within a guard cell and metabolites significantly related to g_s. C and D, Relevant metabolic pathways within a mesophyll cell and metabolites significantly related to g_m and A, respectively. Metabolites in red and blue were significantly related to physiological traits by our linear mixed-effects models with fixed and random effects, respectively. Metabolites with significant negative relationships are underlined. Metabolites found with both approaches were highlighted according to the model with fixed effects. Metabolites significantly related to physiological traits in the PLS modeling approach are marked with an asterisk. Metabolites in black were present in our data set but were not significantly related to any trait, and those in gray were absent from our data set. Dashed arrows indicate steps involving several reactions. CBC, Calvin-Benson cycle; TPT, triose-phosphate transporter; ST, Suc transporter; TMT, Suc transporter; PT, pyruvate transporter; OT, putative oxaloacetate transporter; AT, amino acid transporter; DT, dicarboxylate transporter; OGT, 2-oxoglutarate transporter; CLT, Cl^- transporter; KAT, K^+ transporter; PEP, phosphoenolpyruvate; Pyr, pyruvate; HBA, hydroxybenzoic acid; OAA, oxaloacetate; Ac-CoA; acetyl coenzyme A; Cit, citrate; cis-Aco, cis-aconitate; Isoc, isocitrate; 2-OG, 2-oxoglutarate; Succ-CoA, succinyl coenzyme A; Succ, succinate; Fum, fumarate; Mal, malate; Put, putrescine; DehydroAsc, dehydroascorbate. Amino acids were abbreviated using the standard three-letter code. The green arrow represents the effect of malate on the vacuolar Cl^- transporter.

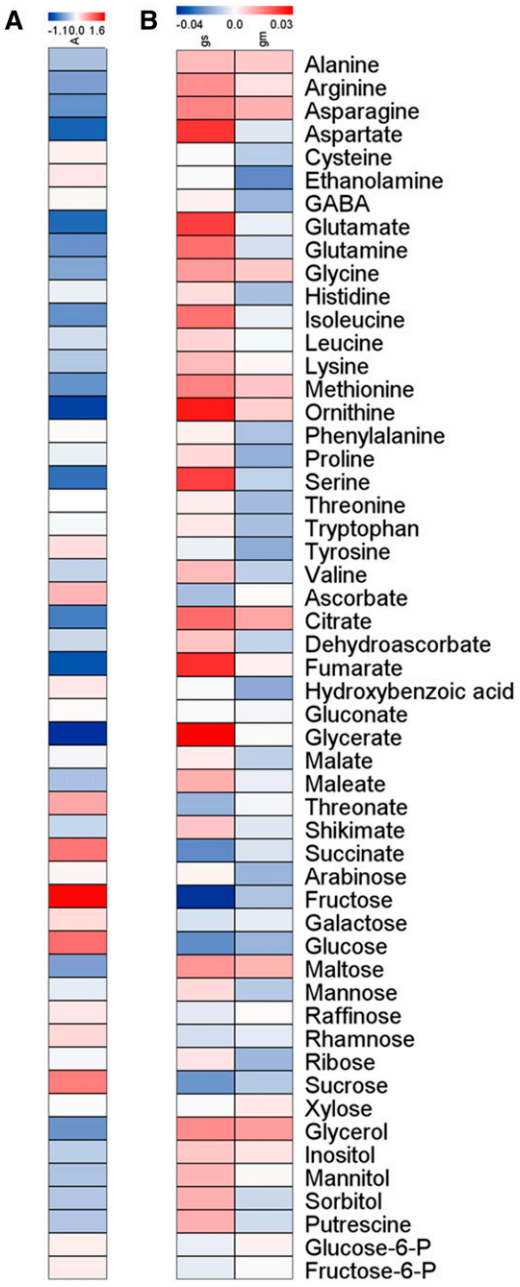


Figure 4. Heat map of the coefficients in the models obtained from PLS analysis for each trait and metabolite. Color scales used in A and B are different.

cross-validation scheme can be helpful in detecting which metabolites may affect the responses via various feedbacks (Supplemental Fig. S2). Inspection of the (R) MSEP indicated that the smallest value across the three dependent variables was obtained when using two components (Supplemental Fig. S2A). Therefore, we next inspected the coefficients of the PLS model built from two components explaining 19.74% of the variance in g_m, having a good compromise with R^2 (Supplemental Fig. S2B). The coefficients of the models for A and g_s were highly negatively correlated (−0.99) but weakly correlated to the coefficients in the model for g_m (0.39 with g_s and −0.47 with A; Supplemental Fig. S3). This implies that the variables of high (small) influence on A have a negligible (large) effect on g_s (Supplemental Data Set S2). This finding implies that metabolites that have a positive effect on A tend to have a negative effect on g_s, a fact that must be considered when developing metabolic engineering strategies to improve photosynthesis. This is particularly interesting since, from Figure 1, A and g_s are positively correlated, but the metabolites used to explain them do not have coefficients of the same sign in the respective models.

For completeness, Figure 4 illustrates a heat map based on the values of the coefficients of individual metabolites in the models for each of the three traits. Interestingly, most of the sugars (Glc, Fru, Suc, Gal, and raffinose) and hexose-phosphates (G6P and F6P) had a positive effect on A; however, many amino acids (e.g. Asn, Gln, His, Ile, Leu, and Met) and organic acids (e.g. citrate, fumarate, glycerate, and maleate) entered the models with negative coefficients (Fig. 4A). In the case of g_s and g_m, different trends can be observed, namely, several amino acids had positive coefficients in the models for g_s and g_m, but some organic acids (e.g. malate and shikimate) had coefficients of opposite sign in the models of g_s and g_m (Fig. 4B). In addition, it is interesting to note that both showed the same negative relationship with the main sugars (Glc, Fru, and Suc; Fig. 4B). The largest (positive) coefficients in the model of g_m included glycerol, citrate, asparagine, maltose, and Met, while the smallest (negative) coefficients were associated with ethanolamine, hydroxybenzoic acid, Tyr, and Pro (Fig. 4B; Supplemental Data Set S2).

DISCUSSION

Our goal was to study the relationships between metabolite profiles from primary metabolic pathways and in vivo leaf photosynthesis and conductances. For this purpose, we collected published data from 17 articles, which cover 14 species analyzed at different developmental stages and environmental conditions. The range of the in vivo physiological traits was in agreement with previous photosynthetic multispecies meta-analysis (Wright et al., 2004; Flexas et al., 2013; Gago et al., 2014; Walker et al., 2014). Nevertheless, the amount of data available for modeling the relationship between g_m and metabolite profiles is limited. Analysis of these parameters showed that A is positively correlated with both g_s and g_m (Fig. 1, A and B), in agreement with the fact that both conductances determine the flux of CO_2 that reaches the Rubisco carboxylation sites in the chloroplast stroma (Flexas et al., 2013). However, the relationship between g_s and g_m was largely scattered and nonsignificant (Fig. 1C), suggesting a certain degree of "uncoupling" as shown, for instance, during drought acclimation and rewatering in different environments in herbaceous and woody plants (Galle et al., 2009, 2011). This is a similar case for most of the g_m data

in our experimental data set, gathered from the drought recovery experiment of Warren et al. (2011) in *Eucalyptus* and *Acacia* sp. While the positive relationships of A with g_s and g_m suggest a direct positive relationship between both conductances, C_i and C_c are not just regulated by the conductances themselves. Moreover, these CO_2 concentrations are also regulated by C_a, the photosynthetic activity, light, leaf temperature, and Rubisco amounts and kinetics. For the same reasons, the relationship between A and g_m can be variable along different environments, particularly when considering different species (Farquhar et al., 1980; Flexas et al., 2013). In any case, the "supply" of CO_2 (conductances) drives the final CO_2 assimilation, as it was previously demonstrated in multispecies data sets; in fact, this is one of the main targets for the improvement of water use efficiency (i.e. the ratio of carboxylation versus water losses in transpiration) and the bottleneck for productivity in semiarid areas, considering water scarcity and predictions of climate change (Flexas et al., 2013; Gago et al., 2014).

To the best of our knowledge, this is the first approach for multispecies primary metabolic profiling in relation to photosynthesis-limiting physiological traits, with overlap in the metabolic profiles between species and studies ranging from 33.9 to 98.1%. Previous multispecies meta-analysis only studied photosynthesis and/or its biochemical parameters in relation to the content of essential nutrients such as nitrogen and phosphorous, establishing the "leaf economics spectrum" through a global view approach (Wright et al., 2004; Walker et al., 2014). In addition, several studies used metabolomics to predict traits such as growth and/or biomass (which can be considered the final, composite outputs of metabolism; Meyer et al., 2007; Sulpice et al., 2013), while others targeted the role of some intermediates of the tricarboxylic acid (TCA) cycle in driving stomatal conductance and thus photosynthesis (Araújo et al., 2011b). In this work, we developed a more complex approach to leaf water and carbon balances in a multispecies data set, considering the feedback effects between the actors (in vivo gas-exchange and primary metabolic networks), in addition to the analysis of conserved and/or species-specific patterns captured in a variety of statistical models beyond the classical pairwise correlation approaches.

Photosynthesis and Primary Metabolism: Major Sugars and Succinate as Markers of High Photosynthetic Activity

In the modeling approaches employed in this work, soluble sugars (Glc, Fru, and Suc) were proportionally the most important parameters related to photosynthetic activity (Figs. 3 and 4; Supplemental Data Set S2), taking place mainly in the mesophyll cells. These statistical relationships could be easily explained given that the Calvin-Benson cycle produces triose-phosphates that are either used to replenish the CO_2 acceptor (ribulose-bisphosphate) or to synthesize the major photosynthetic products, starch and Suc (Stitt et al., 2010).

When a conserved strategy was applied to model the metabolic response using A as input, it had significant effect (P value = 0.037) only for Fru, while the relationship was not significant for Suc or Glc (Fig. 3; Supplemental Data Set S2). Conversely, when metabolites were used to predict the photosynthetic response by means of PLS modeling, we found a positive effect of all major soluble sugars (Suc, Glc, and Fru) and succinate (Fig. 4; Supplemental Data Set S2). These results strongly agree with previously reported data showing that accumulation of Suc is related to the photosynthetic capacity (Stitt et al., 1983; Kruckeberg et al., 1989; Neuhaus et al., 1990). Considering our data set, these relationships probably respond to the wide range covered by A among different species (0.8–30.6 μmol CO_2 m^{-2} s^{-1}). Thus, it seems likely that sugar metabolism can be used as a marker of the photosynthetic activity. Indeed, we found several metabolites of the Suc synthesis pathway that were positively related to A. When linear mixed modeling was performed considering random effects (i.e. species specific), we found the variance of nine metabolites to be explained by A with positive effect including G6P and F6P (Fig. 3; Supplemental Data Set S2). It has been previously shown that the levels of these metabolites are related with A (Stitt et al., 1983; Gerhardt et al., 1987). Even though these metabolites rise marginally with increases in light intensity, the flux to Suc increases at a higher rate (Neuhaus et al., 1990). This is accomplished by relieving inhibition of the cytosolic Fru-1,6-bisphosphatase (by lowering Fru-2,6-bisphosphate levels) and activating Suc-6-P synthase (by reducing its phosphorylation). This feed-forward mechanism provides an effective way to increase carbon flux into Suc with minor changes in steady state levels of metabolites (Neuhaus et al., 1990) and explains the positive correlation observed for hexose-phosphates and Suc with A.

In addition to soluble sugars, succinate was also positively related with A (Fig. 4; Supplemental Data Set S2). It has been reported that reduction of TCA cycle intermediates (malate and fumarate) in Arabidopsis plants overexpressing a heterologous NADP-malic enzyme led to reduction in leaf thickness, photosynthesis, and biomass. Malate and fumarate are accumulated during the day and used in the following dark period to fuel respiration. Therefore, lower levels of these organic acids in mutant plants promoted the use of reduced substrates for respiration, such as fatty acids and proteins (Zell et al., 2010). Whether succinate plays a similar role than malate and fumarate remains an open question. Recently, $^{13}CO_2$ labeling of Arabidopsis whole rosettes demonstrated that biosynthesis of succinate and amino acids (such as Ile, Leu, and Lys) are directly related to carbon assimilation (Ishihara et al., 2015). Even though labeling was incomplete for these metabolites, a considerable portion of newly fixed carbon was directed to the synthesis of succinate and these amino acids during light hours. Our results are in close agreement with these observations, given that these molecules were positively related to A in the modeling approaches employed in our work. Little information is

available concerning the biotechnological potential of manipulating amino acids metabolism. Conversely, it is known that antisense inhibition of the succinate dehydrogenase iron-sulfur subunit enhances A and g_s, in a mechanism that might be mediated by malate accumulation (Araújo et al., 2011b). Therefore, succinate metabolism may be an effective target for photosynthesis improvement.

Stomatal Conductance and Primary Metabolism: The Role of Organic Acids and Other Putative Regulators

It is known that the magnitude of stomatal opening can range according to the accumulation of metabolites from mesophyll in the symplast and apoplast of guard cells. These metabolites can induce stomatal opening or closure, depending on the concentration and the combination of stimuli (Lawson and Blatt, 2014). However, precise information about the mechanisms by which mesophyll cells influence stomatal movements is still lacking. Interestingly, the models used here were able to identify two groups of metabolites, namely, sugars and organic acids, which are known to regulate stomatal movement. While soluble sugars (Suc, Fru, and Glc) and succinate entered the model of A with positive coefficients, they had opposite signs in the models for g_s. The production of sugars in the mesophyll cells and their transport to guard cells represent one of the main points of connection between mesophyll and guard cells and may present a central role in the tight regulation of A-g_s (Lawson and Blatt, 2014). This hypothesis is supported by the fact that in periods of high photosynthetic rate, sugars produced in mesophyll cells are transported to the guard cell apoplast space by the transpiration stream, where they accumulate and induce stomatal closure by a process that remains not well understood (Lu et al., 1995; Kang et al., 2007). Alternatively, Suc, and probably Glc and Fru (given the presence of a presumed hexose carrier in guard cells; Ritte et al., 1999), can enter guard cells and induce stomatal closure in a mechanism dependent of hexokinase and mediated by abscisic acid (ABA; Kelly et al., 2013). This explains the opposite correlation we found for sugars with A and g_s and provides further evidence for a role of Suc and its derivatives, Glc and Fru, in connecting mesophyll and guard cells. Interestingly, our results showed that Man also is positively related with g_s. There is evidence showing that this sugar can act as an osmolyte and its accumulation in leaves inhibits the uptake of Suc, Fru, and Glc, while decreasing starch levels (Lucas and Wilson, 1987). It seems likely that Man may positively regulate stomatal opening, acting as an osmolyte in guard cells and/or by negatively regulating the accumulation of sugars and starch, which has been quantitatively related to stomatal aperture (Outlaw and Manchester, 1979; Lasceve et al., 1997; Kelly et al., 2013; Prasch et al., 2015). Taken together, these results reveal a complex regulatory network involving Suc and Man, and possibly starch, in regulating stomatal movements, a process where these metabolites seem to play a signaling role (apart from their osmotic potential).

Besides the role of the above-mentioned sugars, it is also known that organic acid accumulation can regulate stomatal movements. Malate, which showed a positive coefficient for g_s in the conserved modeling, is the main organic acid related to stomatal movements (Fernie and Martinoia, 2009). It seems that the flux of malate between vacuole, cytosol, and the apoplastic space of guard cells is pivotal to regulate stomatal movements (Chen et al., 2012; Hills et al., 2012). Indeed, it has been shown that this metabolite can act as a counter-ion of K^+ in the vacuole of guard cells or as a signaling molecule to activate guard cell vacuolar chloride channels during stomatal opening (Hedrich and Marten, 1993; De Angeli et al., 2013). Further evidence for the importance of malate and their products and precursors for guard cell metabolism and stomatal movements is demonstrated by the higher g_s found in transgenic plants for enzymes related to organic acid metabolism (Laporte et al., 2002; Nunes-Nesi et al., 2007; Araújo et al., 2011b; Penfield et al., 2012) or guard cell malate transporters (Meyer et al., 2010; Medeiros et al., 2015). The findings of these works are reflected across our entire data compilation, which revealed a strong and positive correlation between the levels of malate and g_s (Fig. 3; Supplemental Data Set S2).

It is well known that the TCA cycle intermediates are important for stomatal movements (Fernie and Martinoia, 2009; Araújo et al., 2011b). Besides malate (discussed above), positive correlations between g_s and fumarate, Asp, Glu, and Orn were also observed. Furthermore, succinate and citrate were negatively correlated with g_s following the PLS and linear mixed effect models, respectively. Altogether, these results further support the idea that metabolites associated with the TCA cycle are involved in the regulation of stomatal movements. Given the similarity in the chemical structure of fumarate in comparison to malate, its pattern of accumulation throughout the day, and that both seem to be involved in the photosynthetic fluxes distribution during light-induced stomatal opening, fumarate is believed to exert the same function as malate in guard cells (Araújo et al., 2011a; Daloso et al., 2015). Indeed, plants that have higher accumulation of malate and fumarate show higher g_s (Nunes-Nesi et al., 2007; Araújo et al., 2011b; Medeiros et al., 2015). In contrast, transgenic tobacco plants that accumulate less succinate and 2-oxoglutarate in leaves and less succinate and citrate in their guard cells also have higher g_s (Daloso et al., 2016). These findings further support our results and suggest that the different parts of the TCA cycle, which is supposed to operate in a noncyclic mode in illuminated leaves (Cheung et al., 2014), have differential contribution to stomatal movements. It seems likely that malate and fumarate would be important osmolytes and signaling molecules in the A-g_s relationship, while the roles of citrate and succinate may be related to maintenance of the energy status and

homeostasis. For instance, the degradation of succinate may be a mechanism to induce ATP production, while citrate could act as a substrate for the synthesis of Glu and Orn, which showed a positive correlation with g_s. However, the role of Glu and Orn for guard cell regulation has yet to be investigated.

Beyond sugars and organic acids, our modeling approach also determined that *myo*-inositol and shikimate entered the models of g_s with negative coefficients. The polyol *myo*-inositol is synthesized from G6P, and it is known that its derivative *myo*-inositol-3-phosphate (IP_3) has an important role in the signaling pathway of ABA-induced stomatal closure (Schroeder et al., 2001). The binding of ABA in G-receptors present in guard cell membranes (Liu et al., 2007; Wan and Liu, 2008) induces changes in phospholipase-C proteins, which in turn induce production of IP_3 (Hunt et al., 2003). IP_3 accumulated in the cytosol activates the efflux of Ca^{2+} from the vacuole and endoplasmic reticulum, which induce the efflux of ions from guard cells and consequently stomatal closure (Fan et al., 2004). Therefore, the negative correlation of *myo*-inositol with g_s found in our data could perhaps be anticipated from the literature. The other metabolite that was also negatively related to g_s was shikimate, a known precursor of flavonols and phenylpropanoids. It was recently described that flavonols develop an important role buffering the ABA-induced oxidative response in the guard cells. Arabidopsis flavonol-less mutants showed a quicker stomatal response, demonstrating that they can act as dampeners regulating the oxidative burst mediated by ABA and the subsequent stomatal closure by the guard cells (Watkins et al., 2014). Overall, these results suggest that shikimate and flavonol/phenylpropanoid metabolism are also involved in regulating the guard cell metabolic network, and this mechanism could be highly conserved between different plant species.

Mesophyll Conductance and Primary Metabolism: A Complex Interplay

Results from our modeling approaches displayed a certain degree of agreement with the intrinsic complexity of g_m and/or the concerns and assumptions of the different methodologies to estimate it, i.e. an important biological complexity driving CO_2 diffusion through different scales, from enzymatic activities (aquaporins and carbonic anhydrases) to the leaf structure and anatomical properties, like cell wall thickness and chloroplast surface exposed to the air spaces of the mesophyll (see "Introduction"). For instance, only one and five metabolites were significantly associated with A and g_s, respectively, when only fixed effects were considered; however, A and g_s had significant effect in the models for nine and 12 metabolites, respectively, in the species-specific modeling approach (i.e. when random effects were considered). Conversely, g_m was found to have significant effect on 19 and 28 metabolites in the case of conserved and species-specific models, respectively (Fig. 3; Supplemental Data Set S2). Interestingly, there were two main groups of metabolites (amino acids and sugars) in which g_m was significant for the models with fixed effects, each of them representing 42% of the total significant metabolites found with our modeling approaches. Amino acids accounted for half of metabolites with significant effects, followed by organic acids. The latter included malate and fumarate, both intermediates of the TCA cycle, which were previously related with the regulation of guard cell turgor and stomatal opening (Araújo et al., 2011b; Daloso et al., 2015), as it was previously indicated in the models for g_s (Figs. 3 and 4). In a certain sense, g_m models showed an intermediate position between A and g_s (Fig. 3). These results fit well with the high degree of coupling between both conductances traditionally observed in wider multispecies data sets (Flexas et al., 2013; Gago et al., 2014) and reported elsewhere; however, under certain conditions, such as the recovery from a water stress period, both conductances can be uncoupled (Fig. 1C).

Surprisingly, some sugars with significant effect for g_m did not have significant effect for A or g_s. These metabolites include Gal, arabinose, and raffinose for the conserved pattern, in addition to Xyl for the species-specific modeling (Fig. 3; Supplemental Data Set S2). Interestingly, these sugars and the hydroxybenzoic acid derivatives could be mostly related with the synthesis and maintenance of the cell wall structure and its pectin matrix (Schnitzler et al., 1992; Alonso et al., 2010; Chen et al., 2013; Franková and Fry, 2013). Cell wall cannot be considered as a fixed element in cell metabolism. Plants invest an important amount of resources in the synthesis and turnover of the cell wall (mostly sugar polymers). For example, it has been described that its glycosidic components can account for 18% of the total dry weight biomass (Williams et al., 2008; Masakapalli et al., 2010; Chen et al., 2013; Franková and Fry, 2013). The role of sugars in relation with the cell wall can be highly dynamic. For instance, Arabidopsis *mur1* mutants with L-Fuc deficiency showed a severe, stunted phenotype; however, when Fuc was exogenously applied, the plants recovered the wild-type phenotype (O'Neill et al., 2001).

In addition, different cell wall composition (including the pectin matrix) can affect the final thickness, structure, and porosity, thus determining the size of the molecules that can pass through it, as well as the tortuosity of the diffusion pathway (Carpita et al., 1979; Baron-Epel et al., 1988; Franková and Fry, 2013; Flexas et al., 2015). In our data set, we analyzed 14 different species, from trees (such as *Eucalyptus* and *Acacia* species) to herbs (like Arabidopsis, tomato [*Solanum lycopersicum*], and rice [*Oryza sativa*]; Supplemental Data Set S1). It is well known that cell wall composition can strongly differ between plant species and groups, especially in those from the Graminaceae family, ferns, and equisetums (Fry et al., 2008; Franková and Fry, 2013). Additionally, it has been described pooling together 61 different species that the cell wall thickness shows a negative significant relationship with g_m

(Tosens et al., 2012; Tomás et al., 2013; Carriquí et al., 2015).

It is important to note that the composition of the plasma and chloroplastic membranes, as well as the chloroplast exposure to mesophyll air spaces, can largely affect g_m. It has been suggested that molecules from the raffinose family can help in stabilizing membranes and proteins of the thylakoid membrane under freezing temperatures, thus preserving chloroplast integrity. Additionally, these sugars might prevent cellular dehydration by deep insertion between lipid layers (Van den Ende, 2013). Ethanolamine is a known precursor of phosphatidylethanolamine and phosphatidylcholine, both major phospholipids in eukaryotic membranes (Kwon et al., 2012). Membrane properties and composition directly affect their permeability to solutes, water, and protons, potentially affecting protein membrane activities (Burgos et al., 2011). Moreover, lipidomes have previously been demonstrated to exhibit high predictive power with regard to agronomic traits in maize (*Zea mays*; Riedelsheimer et al., 2013). Considering previous multispecies leaf anatomical analysis, cell and chloroplast membrane limitations for CO_2 diffusion are not so important as cell wall thickness and chloroplast exposure but certainly impose a significant limitation to g_m (Tomás et al., 2013; Flexas et al., 2014). However, their role and relationship with this trait is currently unknown, even considering their importance affected by environmental factors, as temperature that drives important permeability changes in the membranes that finally affect photosynthesis (Evans and von Caemmerer, 2013).

Finally, further efforts must be made to deeply explore and validate the knowledge extracted from the empirical modeling methodology employed in this study. The relationships observed can be confirmed by conducting an independent experiment, gathering the same metabolic profiles and gas-exchange measurements as in this study, and calculating the error in the prediction and/or through an experiment-based approach using metabolic engineering to obtain mutants (in comparison to the wild type) that target the modeled traits (e.g. increase in the level of a metabolite with positive coefficient and decrease in the metabolite entering the model with a negative coefficient). Subsequently, it must be tested if these mutants have affected photosynthesis, growth, water use efficiency, and stress tolerance in comparison to their wild types and the predicted response, in accordance with the built models.

CONCLUSION

In this work, we assessed the mutual relationships between photosynthesis and stomatal and mesophyll conductances with metabolite profiles from pathways of primary plant metabolism. Our modeling results from the multispecies meta-analysis showed some conserved patterns from herbaceous to woody plants for these traits. When a species-specific strategy was employed, models increased their explanatory power, suggesting certain patterns depend exclusively on each analyzed species. Photosynthesis was widely associated with sugars, phosphorylated intermediates, and organic and amino acids that are linked to biosynthetic processes during the day. Our results for stomatal conductance together with evidence from past studies implied that the balance in the level of sugars, organic acids, *myo*-inositol, and shikimate may regulate stomatal movements through a complex metabolic network and, consequently, photosynthesis. What remains unclear, and thus deserves further exploration, is the identity of the transporters involved in importing these metabolites to guard cell. In addition, it is important to understand when (in which period of the day, which period of stress, and in response to which stimulus) and where (apoplast, symplast, vacuole, chloroplast, and mitochondrion) these metabolites would be accumulated to regulate stomatal movements. In the case of g_m, the models showed a greater complexity when compared to the other traits. Interestingly, some metabolites (like malate) were selected for both g_s and g_m, suggesting a tight coregulation between these parameters, as it has been suggested in the literature. In addition, the significance of some sugars mostly related with cell wall composition and structure (such as arabinose, Xyl, and Gal) opens new biotechnological opportunities to regulate this trait and thus improve water use efficiency. More work is needed to extend this analysis to other species from different life forms and climate origin, in order to decipher the relationships between metabolism and in vivo carbon and water leaf exchanges. Linking ecophysiology and metabolomics offers a new perspective to improve our understanding of leaf water and carbon balances, providing opportunities to better define biotechnological targets related to water use efficiency, productivity, and plant stress tolerance.

MATERIALS AND METHODS

Data Set Compilation

Publically available data of A, g_s, and g_m combined with metabolic profiles, collected in parallel in the same samples, were compiled from 17 published articles describing results for 14 different species. The data set comprised a maximum of 71 data points of A, while compilation of g_s and g_m was reduced to 43 and 21 data points, respectively. Metabolic profiles included data for 53 metabolites, covering amino acids, organic acids, sugars, sugar-alcohols, and phosphorylated intermediates, from the 14 genetic backgrounds (species) under transgenic or environmental (water stress and nutrient deficiency) perturbations. Metabolite overlap between all works ranged from 33.9 to 98.1%. The species included in the data set belong to herbs and woody plants (five herbs; eight trees). Plants were grown under different environments (greenhouses and growth chambers) at optimal environmental conditions (maximum photosynthetic levels were recorded) and under different treatments (such as drought, nutrient deficiency, and transgenic lines with some affected trait). The employed data are detailed in Supplemental Data Set S1. In the following sections, we describe the general methodologies employed in the reviewed papers for acquisition of gas-exchange, fluorescence, and metabolite data.

Gas-Exchange and Fluorescence Measurements

In vivo leaf water and carbon balances and fluorescence measurements were done by clamping fully expanded leaves into the chamber of an open infrared gas analyzer system. All gas-exchange measurements compiled in this data set were performed between 22 and 25°C, at saturating light intensity (250–2,000

μmol photon m^{-2} s^{-1}), atmospheric CO_2 concentrations (380–420 ppm), and nonstressed conditions of vapor deficit pressure (1.5–2.5). When mesophyll conductance was calculated, authors followed the methods of variable J and isotopic discrimination (Harley et al., 1992; Warren, 2011).

Metabolite Analysis

Generally, metabolite analysis was performed in the same leaf type used for gas-exchange measurements. The samples were collected avoiding the midrib, at midday (to prevent the diurnal variations of metabolites), and immediately frozen in liquid nitrogen for subsequent analysis. Extraction procedures were developed by quick grinding of the tissues under liquid nitrogen and subsequent addition of the appropriate extraction buffer. Polar phase was dried, and subsequently the samples were derivatized and the relative levels of metabolites were established by gas chromatography-mass spectrometry following previously described methodologies for data acquisition, annotation, and interpretation (Lisec et al., 2006). More information about the methodologies followed for photosynthetic characterization and metabolite extraction and quantification can be obtained from each article described in Supplemental Data Set S1.

Modeling Approach

Relationships between A, g_s, g_m, and the metabolic profile were analyzed with two different statistical approaches. In the first approach, we modeled the relationship of metabolic readouts, as individual responses, to A, g_s, and g_m as independent variables with linear models of different complexity. Since the differences due to the genetic background are expected to propagate and affect the remaining cellular levels, we used linear mixed-effects modeling approach. To this end, we compared models in which A, g_s, and g_m are fixed effects with those in which the intercept is a random effect and each of A, g_s, and g_m is treated as a random effect. In other words, if A, g_s, and g_m are treated as fixed effects, we are in the classical regression setting. If we allow that any of these regressors have slopes (i.e. regression coefficients), which can vary across the species, then we are in the mixed-effects setting. This strategy allowed us to address the issue of how robust the identified relationships are across species. The mixed-effects linear model approach is robust to missing values and does not suffer from unbalanced experimental design (see below), both present in our setting.

In the second approach, we modeled the relationship of A, g_s, and g_m, as responses, to the metabolic readouts as independent variables, to investigate any relationships between metabolism and the three physiological responses. Since metabolic profiles tend to be correlated due to the structure and regulation of the underlying metabolic networks, and as there are more variables than observations, we used PLS regression.

Since the available metabolite levels are relative (i.e. are expressed as ratios between wild-type/transgenic or ambient condition/water stress), the absolute measurements for the responses were first transformed into relative levels. To provide more robust estimates, the missing values were imputed using a recent random forest imputation method that outperforms other existing alternatives (Stekhoven and Bühlmann, 2012; Waljee et al., 2013; Gromski et al., 2014). Altogether, this resulted in a data matrix with 48 observations, nonuniformly distributed among 14 species, for each data profile.

Linear Mixed-Effects Models

For each metabolite, we tested several models of the type:

$$y = X\beta + Zu + \varepsilon$$

where denotes the given vector with the relative levels of the metabolite across the considered species and experimental set-ups, β stands for the vector of fixed effect to be determined, for the vector of random effects to be determined, ε for the observation error vector, under the assumption of normality for u and ε and the assumption that $Cov(u,\varepsilon)=0$, and X and Z are appropriate design matrices, as detailed in the following. We considered one linear model with A, g_s, and g_m as independent variables without grouping for the genetic background (*i.e.*, species). This is equivalent to classical multivariate regression in which A, g_s, and g_m are fixed effects whose regression coefficients we determined. We also considered the six expanded models which, in addition to the fixed effects for the independent variables, included either only random intercepts or both random intercept and random slope for each of the independent variables (*i.e.*, A, g_s, and g_m). The random effects were with respect to the grouping for the considered species. Since we end up with multiple models for each metabolite based on the choice of fixed and random effects, we need to select the model for final inspection. For each metabolite, comparison based on ANOVA was conducted for model selection. For each of the selected models we determined its (marginal) coefficient of determination and performed additional tests for the fixed effects to aid the interpretation. The models were implemented in R by using the lme4 package (Nakagawa and Schielzeth, 2013; Bates et al., 2015).

PLS Regression

For PLS, we used A, g_s, and g_m as multiple responses (dependent variables) and the metabolite profiles as predictors (i.e. independent variables). The model of (R)MSEP and coefficient of multiple determination (R^2) bias-corrected via cross-validation scheme can be helpful in detecting which metabolites may affect the responses via various feedbacks. We used the R function *plsr* to obtain and analyze these models.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Distribution of the number of linear mixed models of the three independent variables and the coefficient of determination for each metabolite.

Supplemental Figure S2. (R)MSEP and R^2 bias-corrected via cross-validation scheme of the models with different numbers of components through PLS modeling of the physiological traits.

Supplemental Figure S3. Coefficients of the metabolites in the models with one and two components for the three dependent variables.

Supplemental Data Set S1. Literature survey of photosynthetic characteristics in 17 different species.

Supplemental Data Set S2. Results obtained using different modeling approaches.

ACKNOWLEDGMENTS

We thank to Dr. Eva Miedes for critical reading of the manuscript and CFE, IRFS, and AFM for keeping our inspiration alive.

Received October 26, 2015; accepted March 10, 2016; published March 14, 2016.

LITERATURE CITED

Alonso AP, Piasecki RJ, Wang Y, LaClair RW, Shachar-Hill Y (2010) Quantifying the labeling and the levels of plant cell wall precursors using ion chromatography tandem mass spectrometry. Plant Physiol **153:** 915–924

Aranjuelo I, Cabrera-Bosquet L, Morcuende R, Avice JC, Nogués S, Araus JL, Martínez-Carrasco R, Pérez P (2011) Does ear C sink strength contribute to overcoming photosynthetic acclimation of wheat plants exposed to elevated CO_2? J Exp Bot **62:** 3957–3969

Aranjuelo I, Tcherkez G, Molero G, Gilard F, Avice J-C, Nogués S (2013) Concerted changes in N and C primary metabolism in alfalfa (*Medicago sativa*) under water restriction. J Exp Bot **64:** 885–897

Araújo WL, Nunes-Nesi A, Fernie AR (2011a) Fumarate: Multiple functions of a simple metabolite. Phytochemistry **72:** 838–843

Araújo WL, Nunes-Nesi A, Osorio S, Usadel B, Fuentes D, Nagy R, Balbo I, Lehmann M, Studart-Witkowski C, Tohge T, et al (2011b) Antisense inhibition of the iron-sulphur subunit of succinate dehydrogenase enhances photosynthesis and growth in tomato via an organic acid-mediated effect on stomatal aperture. Plant Cell **23:** 600–627

Baron-Epel O, Gharyal PK, Schindler M (1988) Pectins as mediators of wall porosity in soybean cells. Planta **175:** 389–395

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw **67:** 10.18637/jss.v067.i01

Bauer H, Ache P, Lautner S, Fromm J, Hartung W, Al-Rasheid KA, Sonnewald S, Sonnewald U, Kneitz S, Lachmann N, et al (2013) The stomatal response to reduced relative humidity requires guard cell-autonomous ABA synthesis. Curr Biol **23:** 53–57

Brunetti C, George RM, Tattini M, Field K, Davey MP (2013) Metabolomics in plant environmental physiology. J Exp Bot **64:** 4011–4020

Burgos A, Szymanski J, Seiwert B, Degenkolbe T, Hannah MA, Giavalisco P, Willmitzer L (2011) Analysis of short-term changes in the *Arabidopsis thaliana* glycerolipidome in response to temperature and light. Plant J **66:** 656–668

Carpita N, Sabularse D, Montezinos D, Delmer DP (1979) Determination of the pore size of cell walls of living plant cells. Science **205:** 1144–1147

Carriquí M, Cabrera HM, Conesa MÀ, Coopman RE, Douthe C, Gago J, Gallé A, Galmés J, Ribas-Carbo M, Tomás M, Flexas J (2015) Diffusional limitations explain the lower photosynthetic capacity of ferns as compared with angiosperms in a common garden study. Plant Cell Environ **38:** 448–460

Chen X, Alonso AP, Shachar-Hill Y (2013) Dynamic metabolic flux analysis of plant cell wall synthesis. Metab Eng **18:** 78–85

Chen Z-H, Hills A, Bätz U, Amtmann A, Lew VL, Blatt MR (2012) Systems dynamic modeling of the stomatal guard cell predicts emergent behaviors in transport, signaling, and volume control. Plant Physiol **159:** 1235–1251

Cheung CYM, Poolman MG, Fell DA, Ratcliffe RG, Sweetlove LJ (2014) A diel flux balance model captures interactions between light and dark metabolism during day-night cycles in C_3 and crassulacean acid metabolism leaves. Plant Physiol **165:** 917–929

Daloso DM, Antunes WC, Pinheiro DP, Waquim JP, Araújo WL, Loureiro ME, Fernie AR, Williams TCR (2015) Tobacco guard cells fix CO_2 by both Rubisco and PEPcase while sucrose acts as a substrate during light-induced stomatal opening. Plant Cell Environ **38:** 2353–2371

Daloso DM, Williams TCR, Antunes WC, Pinheiro DP, Müller C, Loureiro ME, Fernie AR (2016) Guard cell-specific upregulation of sucrose synthase 3 reveals that the role of sucrose in stomatal function is primarily energetic. New Phytol **209:** 1470–1483

De Angeli A, Zhang J, Meyer S, Martinoia E (2013) AtALMT9 is a malate-activated vacuolar chloride channel required for stomatal opening in Arabidopsis. Nat Commun **4:** 1804

Diaz-Espejo A, Bernacchi CJ, Collatz GJ, Sharkey TD (2012) Models of photosynthesis. *In* J Flexas, F Loreto, H Medrano, eds, Terr. Photosynth. a Chang. Environ. Cambridge University Press, Cambridge, UK, pp 98–112

Ethier GJ, Livingston NJ (2004) On the need to incorporate sensitivity to CO_2 transfer conductance into the Farquhar-von Caemmerer-Berry leaf photosynthesis model. Plant Cell Environ **27:** 137–153

Evans JR, Sharkey TD, Berry JA, Farquhar GD (1986) Carbon isotope discrimination measured concurrently with gas exchange to investigate CO_2 diffusion in leaves of higher plants. Aust J Plant Physiol **13:** 281–292

Evans JR, von Caemmerer S (2013) Temperature response of carbon isotope discrimination and mesophyll conductance in tobacco. Plant Cell Environ **36:** 745–756

Fan LM, Zhao Z, Assmann SM (2004) Guard cells: a dynamic signaling model. Curr Opin Plant Biol **7:** 537–546

Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta **149:** 78–90

Fernie AR (2012) Grand challenges in plant systems biology: Closing the circle(s). Front Plant Sci **3:** 35

Fernie AR, Martinoia E (2009) Malate. Jack of all trades or master of a few? Phytochemistry **70:** 828–832

Flexas J, Barbour MM, Brendel O, Cabrera HM, Carriquí M, Díaz-Espejo A, Douthe C, Dreyer E, Ferrio JP, Gago J, et al (2012) Mesophyll diffusion conductance to CO_2: an unappreciated central player in photosynthesis. Plant Sci **193-194:** 70–84

Flexas J, Carriquí M, Coopman RE, Gago J, Galmés J, Martorell S, Morales F, Diaz-Espejo A (2014) Stomatal and mesophyll conductances to CO_2 in different plant groups: Underrated factors for predicting leaf photosynthesis responses to climate change? Plant Sci **226:** 41–48

Flexas J, Díaz-Espejo A, Conesa MA, Coopman RE, Douthe C, Gago J, Gallé A, Galmés J, et al (2015) Mesophyll conductance to CO_2 and Rubisco as targets for improving intrinsic water use efficiency in C_3 plants. Plant Cell Environ

Flexas J, Niinemets U, Gallé A, Barbour MM, Centritto M, Diaz-Espejo A, Douthe C, Galmés J, Ribas-Carbo M, Rodriguez PL, et al (2013) Diffusional conductances to CO_2 as a target for increasing photosynthesis and photosynthetic water-use efficiency. Photosynth Res **117:** 45–59

Franková L, Fry SC (2013) Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. J Exp Bot **64:** 3519–3550

Fry SC, Nesselrode BH, Miller JG, Mewburn BR (2008) Mixed-linkage (1–>3,1–>4)-beta-D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. New Phytol **179:** 104–115

Fujita Y, Yoshida T, Yamaguchi-Shinozaki K (2013) Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. Physiol Plant **147:** 15–27

Gago J, Douthe C, Florez-Sarasa I, Escalona JM, Galmes J, Fernie AR, Flexas J, Medrano H (2014) Opportunities for improving leaf water use efficiency under climate change conditions. Plant Sci **226:** 108–119

Galle A, Florez-Sarasa I, Aououad HE, Flexas J (2011) The Mediterranean evergreen *Quercus ilex* and the semi-deciduous *Cistus albidus* differ in their leaf gas exchange regulation and acclimation to repeated drought and re-watering cycles. J Exp Bot **62:** 5207–5216

Galle A, Florez-Sarasa I, Tomas M, Pou A, Medrano H, Ribas-Carbo M, Flexas J (2009) The role of mesophyll conductance during water stress and recovery in tobacco (*Nicotiana sylvestris*): acclimation or limitation? J Exp Bot **60:** 2379–2390

Gerhardt R, Stitt M, Heldt HW (1987) Subcellular metabolite levels in spinach leaves: Regulation of sucrose synthesis during diurnal alterations in photosynthetic partitioning. Plant Physiol **83:** 399–407

Gromski PS, Xu Y, Kotze HL, Correa E, Ellis DI, Armitage EG, Turner ML, Goodacre R (2014) Influence of missing values substitutes on multivariate analysis of metabolomics data. Metabolites **4:** 433–452

Harley PC, Loreto F, Di Marco G, Sharkey TD (1992) Theoretical considerations when estimating the mesophyll conductance to CO_2 flux by analysis of the response of photosynthesis to CO_2. Plant Physiol **98:** 1429–1436

Hedrich R, Marten I (1993) Malate-induced feedback regulation of plasma membrane anion channels could provide a CO_2 sensor to guard cells. EMBO J **12:** 897–901

Hills A, Chen Z-H, Amtmann A, Blatt MR, Lew VL (2012) OnGuard, a computational platform for quantitative kinetic modeling of guard cell physiology. Plant Physiol **159:** 1026–1042

Hunt L, Mills LN, Pical C, Leckie CP, Aitken FL, Kopka J, Mueller-Roeber B, McAinsh MR, Hetherington AM, Gray JE (2003) Phospholipase C is required for the control of stomatal aperture by ABA. Plant J **34:** 47–55

Ishihara H, Obata T, Sulpice R, Fernie AR, Stitt M (2015) Quantifying protein synthesis and degradation in arabidopsis by dynamic $^{13}CO_2$ labeling and analysis of enrichment in individual amino acids in their free pools and in protein. Plant Physiol **168:** 74–93

Kang Y, Outlaw WH, Jr., Andersen PC, Fiore GB (2007) Guard-cell apoplastic sucrose concentration–a link between leaf photosynthesis and stomatal aperture size in the apoplastic phloem loader *Vicia faba* L. Plant Cell Environ **30:** 551–558

Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G, Garnier E, Westoby M, et al (2011) TRY - a global database of plant traits. Glob Change Biol **17:** 2905–2935

Kelly G, Moshelion M, David-Schwartz R, Halperin O, Wallach R, Attia Z, Belausov E, Granot D (2013) Hexokinase mediates stomatal closure. Plant J **75:** 977–988

Kim T-H, Böhmer M, Hu H, Nishimura N, Schroeder JI (2010) Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca^{2+} signaling. Annu Rev Plant Biol **61:** 561–591

Kruckeberg AL, Neuhaus HE, Feil R, Gottlieb LD, Stitt M (1989) Decreased-activity mutants of phosphoglucose isomerase in the cytosol and chloroplast of *Clarkia xantiana*. Impact on mass-action ratios and fluxes to sucrose and starch, and estimation of Flux Control Coefficients and Elasticity Coefficients. Biochem J **261:** 457–467

Kwon Y, Yu S-I, Lee H, Yim JH, Zhu J-K, Lee B-H (2012) Arabidopsis serine decarboxylase mutants implicate the roles of ethanolamine in plant growth and development. Int J Mol Sci **13:** 3176–3188

Laporte MM, Shen B, Tarczynski MC (2002) Engineering for drought avoidance: expression of maize NADP-malic enzyme in tobacco results in altered stomatal function. J Exp Bot **53:** 699–705

Lasceve G, Leymarie J, VavasseurA (1997) Alterations in light-induced stomatal opening in a starch-deficient mutant of *Arabidopsis thaliana* L. deficient in chloroplast phosphoglucomutase activity. Environment **20:** 350–358

Lawson T, Blatt MR (2014) Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. Plant Physiol **164:** 1556–1570

Lee M, Choi Y, Burla B, Kim Y-Y, Jeon B, Maeshima M, Yoo J-Y, Martinoia E, Lee Y (2008) The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO_2. Nat Cell Biol **10:** 1217–1223

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc **1:** 387–396

Liu X, Yue Y, Li B, Nie Y, Li W, Wu W-H, Ma L (2007) A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. Science **315:** 1712–1716

Lu P, Zhang SQ, Outlaw WH Jr, Riddle KA (1995) Sucrose: a solute that accumulates in the guard-cell apoplast and guard-cell symplast of open stomata. FEBS Lett **362:** 180–184

Lucas WJ, Wilson C (1987) Influence of mannose on the apoplasmic retrieval systems of source leaves. Plant Physiol **85:** 423–429

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG (2010) Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis cell suspension using steady-state stable isotope labeling. Plant Physiol **152:** 602–619

Medeiros DB, Daloso DM, Fernie AR, Nikoloski Z, Araújo WL (2015) Utilizing systems biology to unravel stomatal function and the hierarchies underpinning its control. Plant Cell Environ **38:** 1457–1470

Messinger SM, Buckley TN, Mott KA (2006) Evidence for involvement of photosynthetic processes in the stomatal response to CO_2. Plant Physiol **140:** 771–778

Meyer RC, Steinfath M, Lisec J, Becher M, Witucka-Wall H, Törjék O, Fiehn O, Eckardt A, Willmitzer L, Selbig J, Altmann T (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **104:** 4759–4764

Meyer S, Mumm P, Imes D, Endler A, Weder B, Al-Rasheid KAS, Geiger D, Marten I, Martinoia E, Hedrich R (2010) AtALMT12 represents an R-type anion channel required for stomatal movement in Arabidopsis guard cells. Plant J **63:** 1054–1062

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol Evol **4:** 133–142

Neuhaus HE, Quick WP, Siegl G, Stitt M (1990) Control of photosynthate partitioning in spinach leaves : Analysis of the interaction between feedforward and feedback regulation of sucrose synthesis. Planta **181:** 583–592

Nunes-Nesi A, Araújo WL, Fernie AR (2011) Targeting mitochondrial metabolism and machinery as a means to enhance photosynthesis. Plant Physiol **155:** 101–107

Nunes-Nesi A, Carrari F, Gibon Y, Sulpice R, Lytovchenko A, Fisahn J, Graham J, Ratcliffe RG, Sweetlove LJ, Fernie AR (2007) Deficiency of mitochondrial fumarase activity in tomato plants impairs photosynthesis via an effect on stomatal function. Plant J **50:** 1093–1106

Obata T, Fernie AR (2012) The use of metabolomics to dissect plant responses to abiotic stresses. Cell Mol Life Sci **69:** 3225–3243

O'Neill MA, Eberhard S, Albersheim P, Darvill AG (2001) Requirement of borate cross-linking of cell wall rhamnogalacturonan II for Arabidopsis growth. Science **294:** 846–849

Outlaw WH, Manchester J (1979) Guard cell starch concentration quantitatively related to stomatal aperture. Plant Physiol **64:** 79–82

Penfield S, Clements S, Bailey KJ, Gilday AD, Leegood RC, Gray JE, Graham IA (2012) Expression and manipulation of phosphoenolpyruvate carboxykinase 1 identifies a role for malate metabolism in stomatal closure. Plant J **69:** 679–688

Pons TL, Flexas J, von Caemmerer S, Evans JR, Genty B, Ribas-Carbo M, Brugnoli E (2009) Estimating mesophyll conductance to CO_2: methodology, potential errors, and recommendations. J Exp Bot **60:** 2217–2234

Prasch CM, Ott KV, Bauer H, Ache P, Hedrich R, Sonnewald U (2015) β-amylase1 mutant Arabidopsis plants show improved drought tolerance due to reduced starch breakdown in guard cells. J Exp Bot **66:** 6059–6067

Reich PB, Ellsworth DS, Walters MB, Vose JM, Gresham C, Volin JC, Bowman WD (1999) Generality of leaf trait relationships: A test across six biomes. Ecology **80:** 1955–1969

Riedelsheimer C, Brotman Y, Méret M, Melchinger AE, Willmitzer L (2013) The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. Sci Rep **3:** 2479

Ritte G, Rosenfeld J, Rohrig K, Raschke K (1999) Rates of sugar uptake by guard cell protoplasts of *pisum sativum* L. Related To the solute requirement for stomatal opening. Plant Physiol **121:** 647–656

Schnitzler JP, Madlung J, Rose A, Ulrich Seitz H (1992) Biosynthesis of p-hydroxybenzoic acid in elicitor-treated carrot cell cultures. Planta **188:** 594–600

Schroeder JI, Kwak JM, Allen GJ (2001) Guard cell abscisic acid signalling and engineering drought hardiness in plants. Nature **410:** 327–330

Stekhoven DJ, Bühlmann P (2012) MissForest–non-parametric missing value imputation for mixed-type data. Bioinformatics **28:** 112–118

Stitt M, Lunn J, Usadel B (2010) Arabidopsis and primary photosynthetic metabolism - more than the icing on the cake. Plant J **61:** 1067–1091

Stitt M, Wirtz W, Heldt HW (1983) Regulation of sucrose synthesis by cytoplasmic fructosebisphosphatase and sucrose phosphate synthase during photosynthesis in varying light and carbon dioxide. Plant Physiol **72:** 767–774

Sulpice R, Nikoloski Z, Tschoep H, Antonio C, Kleessen S, Larhlimi A, Selbig J, Ishihara H, Gibon Y, Fernie AR, Stitt M (2013) Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. Plant Physiol **162:** 347–363

Sun Y, Gu L, Dickinson RE, Pallardy SG, Baker J, Cao Y, DaMatta FM, Dong X, Ellsworth D, Van Goethem D, et al (2014) Asymmetrical effects of mesophyll conductance on fundamental photosynthetic parameters and their relationships estimated from leaf gas exchange measurements. Plant Cell Environ **37:** 978–994

Tholen D, Ethier G, Genty B, Pepin S, Zhu X-G (2012) Variable mesophyll conductance revisited: theoretical background and experimental implications. Plant Cell Environ **35:** 2087–2103

Timm S, Mielewczik M, Florian A, Frankenbach S, Dreissen A, Hocken N, Fernie AR, Walter A, Bauwe H (2012) High-to-low CO_2 acclimation reveals plasticity of the photorespiratory pathway and indicates regulatory links to cellular metabolism of Arabidopsis. PLoS One **7:** e42809

Tomás M, Flexas J, Copolovici L, Galmés J, Hallik L, Medrano H, Ribas-Carbó M, Tosens T, Vislap V, Niinemets Ü (2013) Importance of leaf anatomy in determining mesophyll diffusion conductance to CO_2 across species: quantitative limitations and scaling up by models. J Exp Bot **64:** 2269–2281

Tosens T, Niinemets U, Vislap V, Eichelmann H, Castro Díez P (2012) Developmental changes in mesophyll diffusion conductance and photosynthetic capacity under different light and water availabilities in *Populus tremula*: how structure constrains function. Plant Cell Environ **35:** 839–856

Van den Ende W (2013) Multifunctional fructans and raffinose family oligosaccharides. Front Plant Sci **4:** 247

Vani T, Raghavendra AS (1994) High mitochondrial activity but incomplete engagement of the cyanide-resistant alternative pathway in guard cell protoplasts of pea. Plant Physiol **105:** 1263–1268

Waljee AK, Mukherjee A, Singal AG, Zhang Y, Warren J, Balis U, Marrero J, Zhu J, Higgins PD (2013) Comparison of imputation methods for missing laboratory data in medicine. BMJ Open **3:** 1–7

Walker AP, Beckerman AP, Gu L, Kattge J, Cernusak LA, Domingues TF, Scales JC, Wohlfahrt G, Wullschleger SD, Woodward FI (2014) The relationship of leaf photosynthetic traits - V_{cmax} and J_{max} - to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. Ecol Evol **4:** 3218–3235

Wan X-Y, Liu J-Y (2008) Comparative proteomics analysis reveals an intimate protein network provoked by hydrogen peroxide stress in rice seedling leaves. Mol Cell Proteomics **7:** 1469–1488

Wang S-W, Li Y, Zhang X-L, Yang H-Q, Han X-F, Liu Z-H, Shang Z-L, Asano T, Yoshioka Y, Zhang C-G, Chen YL (2014) Lacking chloroplasts in guard cells of crumpled leaf attenuates stomatal opening: both guard cell chloroplasts and mesophyll contribute to guard cell ATP levels. Plant Cell Environ **37:** 2201–2210

Warren CR (2011) How does P affect photosynthesis and metabolite profiles of *Eucalyptus globulus*? Tree Physiol **31:** 727–739

Warren CR, Aranda I, Cano FJ (2011) Responses to water stress of gas exchange and metabolites in *Eucalyptus* and *Acacia* spp. Plant Cell Environ **34:** 1609–1629

Warren CR, Aranda I, Cano FJ (2012) Metabolomics demonstrates divergent responses of two *Eucalyptus* species to water stress. Metabolomics **8:** 186–200

Watkins JM, Hechler PJ, Muday GK (2014) Ethylene-induced flavonol accumulation in guard cells suppresses reactive oxygen species and moderates stomatal aperture. Plant Physiol **164:** 1707–1717

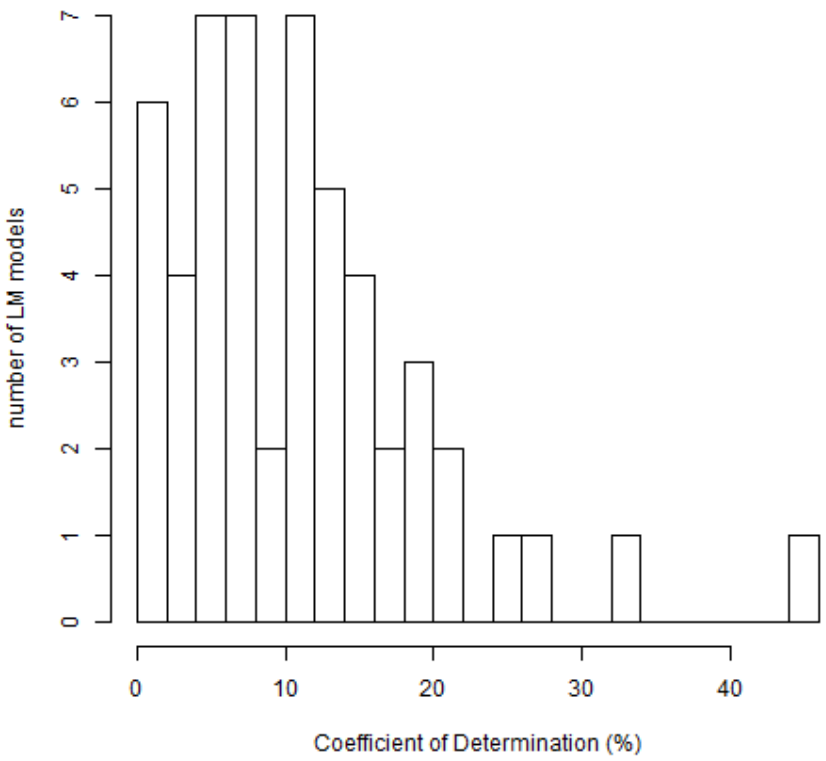
Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG (2008) Metabolic network fluxes in heterotrophic Arabidopsis cells: stability of the flux distribution under different oxygenation conditions. Plant Physiol **148:** 704–718

Witt S, Galicia L, Lisec J, Cairns J, Tiessen A, Araus JL, Palacios-Rojas N, Fernie AR (2012) Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. Mol Plant **5:** 401–417

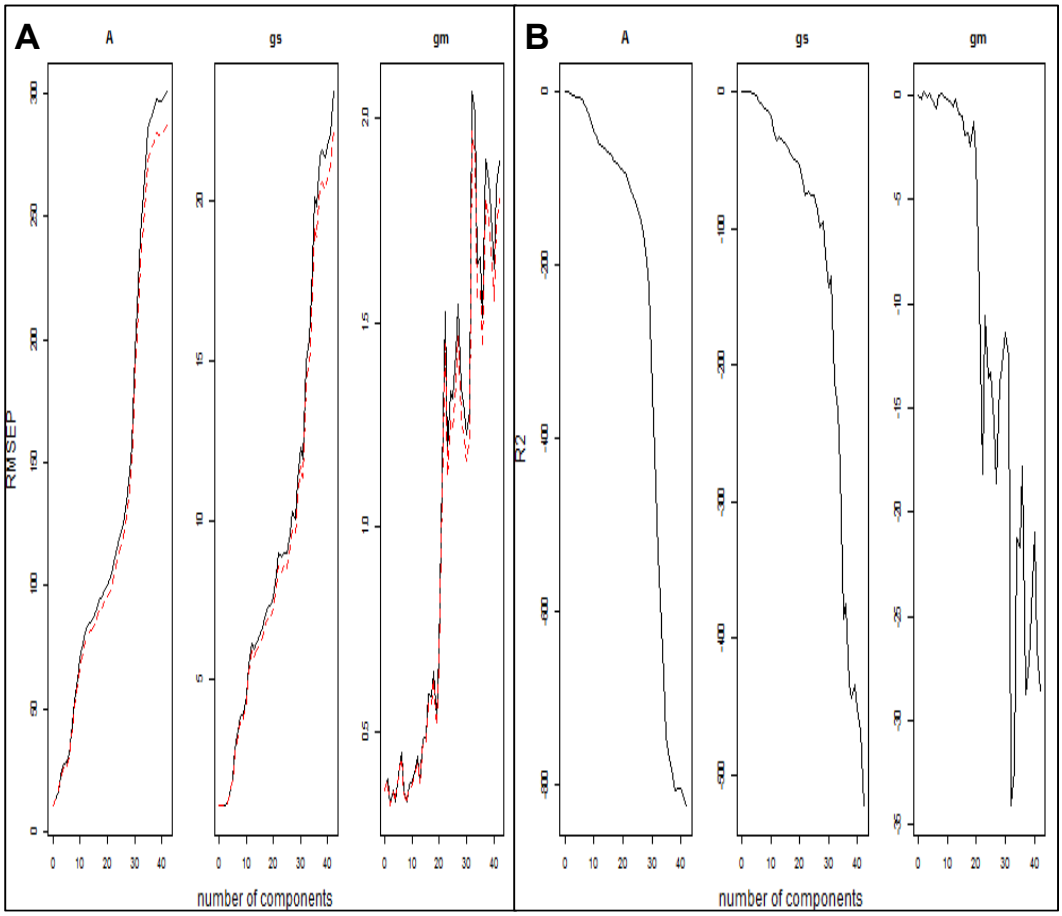
Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M, et al (2004) The worldwide leaf economics spectrum. Nature **428:** 821–827

Zeiger E, Talbott LD, Frechilla S, Srivastava A, Zhu J (2002) The guard cell chloroplast: A perspective for the twenty-first century. New Phytol **153:** 415–424

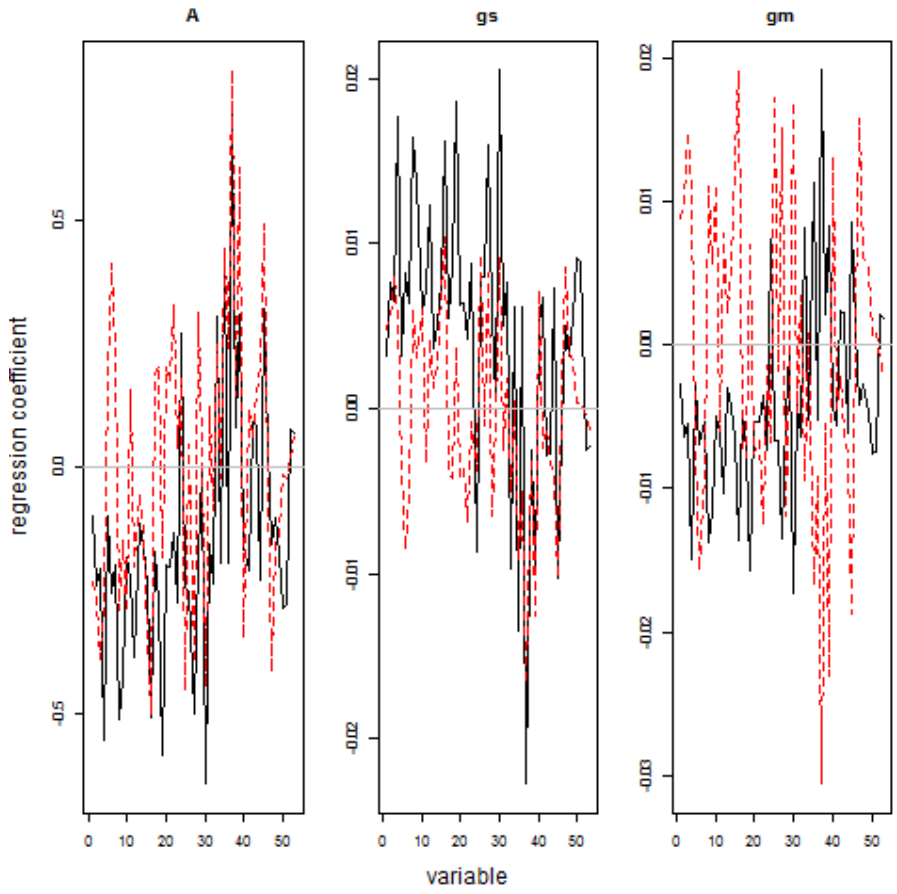
Zell MB, Fahnenstich H, Maier A, Saigo M, Voznesenskaya EV, Edwards GE, Andreo C, Schleifenbaum F, Zell C, Drincovich MF, Maurino VG (2010) Analysis of Arabidopsis with highly reduced levels of malate and fumarate sheds light on the role of these organic acids as storage carbon molecules. Plant Physiol **152:** 1251–1262



Supplemental Figure S1. Distribution of the number of linear mixed models (without random effect) of the three independent variables (A, g_s, and g_m) and the coefficient of determination for each metabolite.



Supplemental Figure S2. (A) Root mean squared error of prediction [(R)MSEP] and (B) coefficient of multiple determination (R^2) bias-corrected via cross-validation scheme of the models with different number of components (metabolites as predictors) through PLS modeling of the physiological traits (*A*, g_s, and g_m as dependent variables).



Supplemental Figure S3. Coefficients of the metabolites (x-axis, naming not shown for simplicity) in the models with one component (black) and two components (red) for the three dependent variables (A, g_s, and g_m).