frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 April 2018
doi: 10.3389/fpls.2018.00448

Phenotypic Trait Identification Using a Multimodel Bayesian Method: A Case Study Using Photosynthesis in *Brassica rapa* Genotypes

Jonathan R. Pleban[1]*, *D. Scott Mackay*[1], *Timothy L. Aston*[2], *Brent E. Ewers*[2,3] and *Cynthia Weinig*[2,3,4]

[1] Department of Geography, University at Buffalo, Buffalo, NY, United States, [2] Department of Botany, University of Wyoming, Laramie, WY, United States, [3] Program in Ecology, University of Wyoming, Laramie, WY, United States, [4] Department of Molecular Biology, University of Wyoming, Laramie, WY, United States

Agronomists have used statistical crop models to predict yield on a genotype-by-genotype basis. Mechanistic models, based on fundamental physiological processes common across plant taxa, will ultimately enable yield prediction applicable to diverse genotypes and crops. Here, genotypic information is combined with multiple mechanistically based models to characterize photosynthetic trait differentiation among genotypes of *Brassica rapa*. Infrared leaf gas exchange and chlorophyll fluorescence observations are analyzed using Bayesian methods. Three advantages of Bayesian approaches are employed: a hierarchical model structure, the testing of parameter estimates with posterior predictive checks and a multimodel complexity analysis. In all, eight models of photosynthesis are compared for fit to data and penalized for complexity using deviance information criteria (DIC) at the genotype scale. The multimodel evaluation improves the credibility of trait estimates using posterior distributions. Traits with important implications for yield in crops, including maximum rate of carboxylation (V_{cmax}) and maximum rate of electron transport (J_{max}) show genotypic differentiation. *B. rapa* shows phenotypic diversity in causal traits with the potential for genetic enhancement of photosynthesis. This multimodel screening represents a statistically rigorous method for characterizing genotypic differences in traits with clear biophysical consequences to growth and productivity within large crop breeding populations with application across plant processes.

Keywords: A/C_i curves, Bayesian models, *Brassica rapa*, chlorophyll fluorescence, multimodel analysis, phenotyping, photosynthesis

OPEN ACCESS

Edited by:
Andreia Michelle Smith-Moritz, University of California, Davis, United States

Reviewed by:
Robert Brian O'Hara, Norwegian University of Science and Technology, Norway
Tom De Swaef, Institute for Agricultural and Fisheries Research (ILVO), Belgium

***Correspondence:**
Jonathan R. Pleban
jrpleban@buffalo.edu

Specialty section:
This article was submitted to Plant Biophysics and Modeling, a section of the journal Frontiers in Plant Science

Received: *27 April 2017*
Accepted: *22 March 2018*
Published: *17 April 2018*

Citation:
Pleban JR, Mackay DS, Aston TL, Ewers BE and Weinig C (2018) Phenotypic Trait Identification Using a Multimodel Bayesian Method: A Case Study Using Photosynthesis in Brassica rapa Genotypes. Front. Plant Sci. 9:448. doi: 10.3389/fpls.2018.00448

INTRODUCTION

Maintaining food security for the world's rapidly growing population is a paramount challenge for science. Classic and modern genomic breeding programs represent one of the major tools for increasing food supply to counter this Malthusian dilemma. The success of breeding programs in part depends on the ability to quickly identify beneficial phenotypic traits in breeding populations (Sadras et al., 2013). Experimentation and modeling assists trait identification while producing insights into plant physiology and crop productivity (Sinclair and Seligman, 1996;

Hammer et al., 2006). Mechanistic modeling using known or theorized ecological, biochemical and biophysical principles further advance understanding through connecting yield to causal traits (Laisk and Nedbal, 2009; Tardieu, 2010). These mechanistic models of plant physiology use statistical tools to estimate trait variation by organizing phenomenological data into meaningful mathematical representations of enzymatic and protein activity responsible for plant processes (DeWitt, 1965; von Caemmerer, 2000; Patrick et al., 2009; McDowell et al., 2013). In this way models can estimate valuable phenotypic information through specifying physiologically meaningful trait values from data.

Photosynthesis is a primary target for selective enhancement in crops (Long et al., 2006; Singh et al., 2014; Furbank et al., 2015) and the mechanisms of photosynthesis are well-studied with models used to characterize assimilatory strategies across taxa (Wullschleger, 1993; Patrick et al., 2009; Gu et al., 2012). For these reasons, photosynthesis was chosen as a target for multimodel phenotyping. Modeling of photosynthesis evolves as theory tests data in attempts to replicate the pathway of enzymatic and protein responses responsible for carbon fixation, light harvesting, and electron transfer (DeWitt, 1965; Farquhar et al., 1980, 2001; von Caemmerer, 2000; Yin et al., 2009). At each stage of model development simplifying assumptions are made regarding the behavior of these pathways. For example, assumptions are made in regard to the degree of trait response to temperature fluctuation or regarding the relative resistance change in the CO_2 pathway from the atmosphere to the site of carboxylation (Medlyn et al., 2002a; Pons et al., 2009). Such assumptions affect modeling efforts in three key ways. First, assumptions accommodate unknowns and can shift model emphases from mechanistic processes to empirical relationships. Second, assumptions impact model complexity, forcing modelers to assess the performance of models ranging in complexity (Knorr and Heimann, 2001; Martre et al., 2015). Third, assumptions can influence the uncertainty of trait estimates (Mackay et al., 2012). Uncertainty quantification is therefore necessary when making claims of trait differentiation. More broadly, uncertainty quantification can inform the cyclical process of model improvement by repeatedly testing updated theory against data (Box, 2001).

A number of theoretical developments with empirical support have identified the critical factors limiting leaf-level CO_2 assimilation (A). Generally, the limiting factors are divided into two major classes: diffusional and biochemical. Diffusional limits can be further subdivided into a stomatal limitation that is imposed by guard cell control over stomatal conductance and a mesophyll limitation regulating CO_2 and H_2O transport between the intracellular space and the site of carboxylation (C_c) (Ethier and Livingston, 2004; Grassi and Magnani, 2005; Niinemets et al., 2009a; Damour et al., 2010). Limits on photosynthesis imposed by mesophyll conductance (g_m) have been shown to be of a similar magnitude to g_s limitation (Grassi and Magnani, 2005). However, uncertainty remains in understanding the limits g_m imposes across taxa and environments as all methods of estimating g_m rely on models sensitive to parameterization (Pons et al., 2009; Gu and Sun, 2014; Théroux-Rancourt and Gilbert, 2017). The leaf biochemical limitations controlling A are summarized by two primary factors, Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) limited A (A_c) and regeneration of ribulose biphosphate (RuBP) limited A (A_J). A_c follows the Michaelis–Menten enzymatic kinetics for RuBisCO. This requires amongst other parameters the estimation of the maximum rate of carboxylation (V_{cmax}). A_J is coordinated by the electron transport rate (ETR) across photosystems II and I (PSII, PSI), which produces ATP and NADPH needed for the Calvin carboxylation cycle (von Caemmerer, 2000). Measurements of chlorophyll fluorescence have been used as proxies for ETR limited A_J (Genty et al., 1989; Baker, 2008). A carbon metabolism limitation or triose phosphate utilization (TPU) limitation has also been identified (Sharkey et al., 1985). Beyond the major diffusional and biochemical limitation, studies have sought to better understand the influence temperature has on photosynthetic performance (Bernacchi et al., 2001; Medlyn et al., 2002a; Patrick et al., 2009). Temperature influence can be modeled using an activation energy (E_i) parameter following an Arrhenius function (von Caemmerer, 2000). Both diffusional and biochemical limitations are temperature-responsive using these modeling methods (Bernacchi et al., 2001, 2002; Leuning, 2002; Kattge and Knorr, 2007). In total, the inclusion or absence of these limitations, constraints or assumptions regarding leaf biochemistry and biophysics results in models of varying complexity.

Progressively, each incremental change in model form represents an alternative view of how the photosynthetic machinery behaves. A multimodel framework can test for the strengths and weaknesses of these alternative views. Information criteria such as Akaike Information Criteria and Bayesian correlate Deviance Information Criterion (DIC) provide metrics for evaluating model adequacy through combining terms of both the goodness of fit and model complexity (Akaike, 1998; Spiegelhalter et al., 2002). A recent statistical argument suggests that multimodel analyses are a superior method, compared to null hypothesis approaches, to test mechanisms against data (McElreath, 2016). Climate models leverage the multimodel approach for generating estimates of regional temperature change (Tebaldi et al., 2004) and these climate model ensembles have been used in crop prediction models (Ruane et al., 2017). Hydrological work also embraces Bayesian multimodel approaches to both inform groundwater estimates and sampling schema (Xue et al., 2014).

The current state of the art in photosynthesis models seeks to capture known and theoretical biophysical processes of the diffusional limitations and the light and light-independent reactions responsible for observed variation across environments and taxa (van der Tol et al., 2009; Yin et al., 2009). Concurrently modelers aim to parameterize at finer evolutionary scales (Yin et al., 2001; Patrick et al., 2009; Yamori et al., 2014). The broad division of models between C-3 and C-4 plants represents a critical improvement in this context (von Caemmerer, 2000). The C-3/C-4 evolutionary shift resulted in mechanistic differences requiring unique modeling frameworks for successfully understanding these two broad assimilatory systems. But approaches ignore model structural differences

when considering variation among closely related individuals. Here we argue that multimodel analysis can assist in testing for variation across evolutionary context, as the structure, distribution, abundance, and therefore behavior of critical enzymes and proteins should be conserved in closely related populations relative to evolutionarily distant ones. Of relevance to crop breeding and highly outcrossing wild species, allelic combinations from two parents may lead to physiological responses beyond the range expressed by the parents due to transgressive segregation (Rieseberg et al., 1999). In these cases, identifying genetic controls over biophysical traits and processes across the species as a whole may be more complicated. Even so, instances of transgressive segregation explore phenotypic space that may be acted upon by natural or artificial selection (Rieseberg et al., 2003). In such cases genotypic differences may require genotype-specific behaviors and allow for reduced model complexity by eliminating the need for particular parameters. For example, it may be unnecessary to account for g_m-limitation if all experimental genotypes have uniformly high g_m relative to g_s. Therefore, when probing for trait variation within a genetically variable species, it is important to explicitly test if genotype- or species-level model forms are preferred. Overall, our aim was a robust complexity analysis of multiple leaf photosynthesis models for evaluation of photosynthetic differences of different allelic combinations.

MATERIALS AND METHODS

Overview

Model evaluation should not only considerer fit because increasing complexity often improves fit but may not increase predictive power. Therefore, checking models for both fit and parsimony should occur iteratively (Box, 1976; Spiegelhalter et al., 2002). To avoid unnecessarily high dimensionality, statistical measures have been developed that penalize complexity as part of a model selection strategy (Akaike, 1998; Spiegelhalter et al., 2002; Plummer, 2008). In the screening tool developed here, the complexity associated with three leaf level modifications of photosynthesis models was quantified: temperature constraints, g_m limitation, and derivational form of ETR.

Bayesian methodologies are well-suited for parameterization, uncertainty analysis and multimodel evaluation (Gelman et al., 2004; Kruschke, 2010). These methods have wide application in plant ecology and physiology (Ogle and Barber, 2008; Patrick et al., 2009; Mackay et al., 2012; Gou et al., 2017). While many Bayesian studies have focused on just one or two of these beneficial features; often parameterization and uncertainty analysis (Zhu et al., 2011; Gou et al., 2017), we sought to leverage all three features. Additionally, Bayesian methodologies can be tested in hierarchical structures whereby multiple datasets can be combined to inform parameterization at different levels, such as among taxa or genotypes, spatial scales or other known sampling characteristics (Gelman et al., 2004). Here, a simple two level hierarchical structure is adopted, where individual and genotypic level traits are estimated. This analysis is summarized in **Figure 1** as a five-part trait-screening methodology; four of these five parts are presented in detail here. First, experimental data, described later in the Plant Physiological Measurements section, is collected across genotypes and/or within treatments; the data set presented here is across genotypes under unstressed growth conditions. Second, alternative models are constructed in an effort to challenge commonly held yet non-definitive assumptions regarding the process of interest; here multiple photosynthesis models are established using curves of CO_2 assimilation (A) vs. intercellular CO_2 concentrations (C_i) (A/C_i). Third, observation data is passed into competing models using a sampling scheme designed to produce Bayesian posteriors. Fourth, the posterior distributions of each model are examined, in order to identify traits with genotypic differentiation and to evaluate genotypic differences in light of model complexity. Fifth, findings are summarized to develop future experimental-modeling iterations. In sum, potentially beneficial traits are identified, areas for model improvement are considered, and resultant posterior distributions are used to update trait priors for subsequent evaluation. At the same time, new experimentation can be considered based on findings and genetic analysis of trait estimates may support classic or molecular breeding based on candidate genes. Ultimately, this multimodel analysis aims at critically evaluating trait differences within the population under investigation and represents an important step in linking phenome to genome, here demonstrated on the vital processes of carbon assimilation and light harvesting.

Study Site and Genotypes

Data were obtained from experiments undertaken in the summers of 2012 and 2013 at the University of Wyoming Research and Extension Center Field Complex (41.32 N, 105.56 W) in Laramie WY, USA. Details of the experiment are given in Aston et al. (in prep). We included six genotypes of *B. rapa*: four crop accessions [oilseed, subsp. *Pusa Kalyani* cgn06834 (*oil*); turnip, subsp. *Maiskaja* cgn06710 (*tur*); Chinese cabbage, subsp. *Pekinensis* cgn13942 (*cab*); broccoletto subsp. *Quarantina* cgn06825 (*bro*)] and two recombinant inbred lines (RILs) (*r46* and *r301*). Crop accession seeds were obtained from the Wageningen University and Research Center for Genetic Resources. The RILs, *r46* and *r301*, are the F8 offspring of a cross between the IMB211 genotype derived from the Wisconsin Fast Plant™ population and the R500 genotype, an oilseed long cultivated in India. The two RILs, full siblings, were selected based on the expression of transgressive segregation for intrinsic water use efficiency (*WUE*) identified in earlier research (Edwards et al., 2011, 2012). Plants were germinated in a greenhouse and after two weeks transplanted to a rain shelter in the field where they were grown under well-watered conditions. All measurements were taken on days 25–28 after planting.

Plant Physiological Measurements

Infrared leaf gas exchange (IRGA) measurements (Li-6400XT, Li-Cor, Lincoln, NE, USA) were taken to measure A/C_i response with a constant irradiance of 2,000 μmol m^{-2} s^{-1} for CO_2 concentrations of ∼50, 100, 200, 300, 400, 500, 600, 800, 1,000, 1,250, 1,500, and 2,000 μmol mol^{-1}. IRGA measurements monitor fluxes of both CO_2 and H_2O allowing for direct measurement of A and transpiration (E) with indirect means

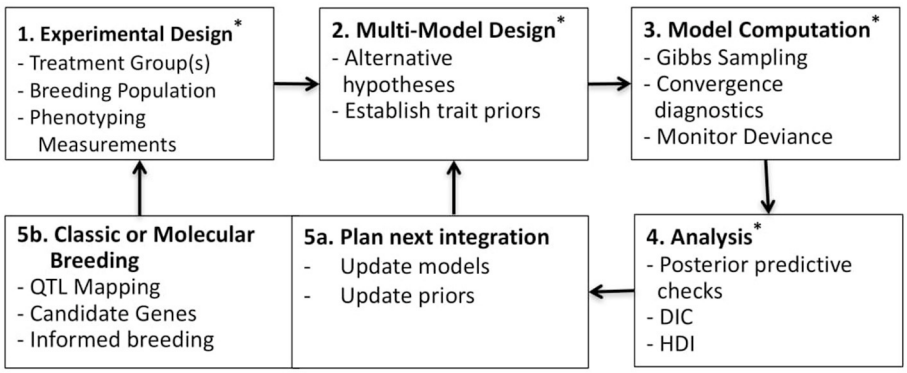

FIGURE 1 | Workflow for multimodel Bayesian phenotyping. Box 1: Experimental design across varietals and/or treatments; here focus was on A/C_i curves from six *B. rapa* genotypes. Box 2: Alternative models constructed for testing assumptions regarding the process of interest; here eight photosynthesis models with trait priors established from literature (**Tables 1**–**4**). Box 3: Bayesian sampling for generating posterior distributions across all models X individuals, here executed in rjags (Plummer, 2014). Box 4: Posterior distributions evaluated using posterior predictive checks (**Figures 2**–**4**), genotypic model complexity compared using DIC (**Table 5**), and posterior trait distributions of each model scrutinized using boxplots (**Figures 5**–**8**) and high-density interval (HDI) analysis (**Table 6**). Box 5a: Model evaluation and updating along with updating of prior distributions for population. Box 5b: Genomic analysis of identified traits and classic or molecular breeding for establishing next test population (**Figure 9**). * indicates steps competed in this study.

of assessing g_s and C_i. CO_2 response curves were measured between 10:00 and 16:00 h on fully expanded, mature leaves with leaf temperature maintained near 22°C and relative humidity maintained within 10% of ambient. A steady state was achieved at each CO_2 increment prior to each gas exchange observation. Temperature and vapor pressure deficit averaged 21.1°C (±3.0) and 2.01 kPa (±0.4), respectively. Chlorophyll fluorescence measurements were taken in conjunction with each gas exchange observation. Chlorophyll fluorescence observations measure the re-radiated near infrared light by the leaf. Fluorescence serves as one means, albeit small, of dissipating excess light energy (Maxwell and Johnson, 2000). Fluorescence yield is used to quantify the amount of light energy transferred from excited photosystem II to primary quinone acceptors, such as plastoquinone, driving downstream photosynthetic light reactions (Baker, 2008). This is done using a saturating flash followed by a dark pulse to measure $F_m' - F_s/F_m'$, where Fs is steady state fluorescence yield and F_m' is maximum light-adapted fluorescence yield. $F_m' - F_s/F_m'$ is commonly referred to as effective quantum yield of PSII (ϕ_{PSII}) (Genty et al., 1989). A simple method using ϕ_{PSII} to estimate total flux of the electron transport chain based on fluorescence (J_f) (μmol m^{-2} s^{-1}) is described in Equation (3.5) of **Table 2**. Equation (3.5) makes assumptions regarding the partitioning of light between photosystems (f) (0.5) and the fractional value of light absorptance by leaf photosynthetic pigments (α_{leaf}) (0.85). In total 31 individual *A/Ci* curves were tested in the analysis (six *r301*, six *r46*, five *bro*, five *cab*, three *oil,* and six *tur*).

Bayesian Modeling Approach

A Bayesian model is comprised of three sets of probability statements (Gelman et al., 2004; Kruschke, 2010; Ogle and Barber, 2011). First, prior statements represent a statistically sound and repeatable method of summarizing known information, in this case regarding plant photosynthesis physiological traits (Ogle and Barber, 2008). The second probability statement is the likelihood function(s) expressing the probability that a given model could have generated a particular set of data. The final probabilistic statements are the posterior distributions describing the strength of a model once data have been tested as well as the degree of uncertainty in parameterization. Here 31 *A/Ci* curves were independently evaluated using a suite of eight models following this approach. The data, parameters, and predictions made by these eight models are described in **Table 1**, with model equations identified in **Table 2**. Each model has a coded name based on the assumptions therein (**Table 3**). The model and implementation codes used for analysis are provided at https://github.com/jrpleban/Bayes_Farquhar_Models_2_level_Hierarchy. Priors on parameters are shown in **Table 4**. We have chosen to estimate some parameters often set as constants (K_c, K_o) to evaluate a given model's ability to discern traits expected to be conserved in this population. A literature survey for each parameter was used to provide statistical distributions for parameter priors. Many parameters have ample data across taxa, such as for J_{max25}, allowing a normal distribution of priors (Wullschleger, 1993). When a trait's distribution was more uncertain, broad priors were used, such as for ϕ_J.

All models assumed that observations of A (A_n) (μmol m^{-2} s^{-1}) followed a normal distribution:

$$A_n \sim N(A_{exp}, \tau) \quad (1)$$

where A_{exp} is the expected photosynthetic rate, τ is precision ($1/\sigma^2$) describing the variability in measurement error. A hierarchical design nested individual plant parameters within the genotypic populations. This nested design was used for all

TABLE 1 | List of abbreviations used for observations, predictions, and parameters of eight photosynthesis models.

Abbreviation	Definition	Units	Models using
OBSERVATIONAL DATA			
A_n	CO_2 assimilation rate observed	μmol m^{-2} s^{-1}	All
C_i	Intercellular CO_2 partial pressure observed	Pa	CiCc Models
C_a	Ambient CO_2 partial pressure observed	Pa	CaCc Models
T	Leaf temperature observed	°C	Temp Models
P	Pressure observed	Pa	All
g_s	Conductance to CO_2 from atmosphere to intercellular space observed	μmol m^{-2} s^{-1}	CaCc Models
O	Ambient O_2 (assumed 21% atmosphere)	Pa	All
Q	Photosynthetically active radiation observed	μmol m^{-2} s^{-1}	All
ϕ_{PSII}	Quantum yield of photosystem II based on Chlorophyll fluorescence	e^- $photon^{-1}$	Jf Models
J_f	Electron transport rate from Chlorophyll fluorescence (Equation 3.5)	μmol m^{-2} s^{-1}	Jf Models
PROCESS MODEL PREDICTIONS			
A_{exp}	Expected CO_2 assimilation rate	μmol m^{-2} s^{-1}	All
A_c	Rubisco limited rate of CO_2 assimilation	μmol m^{-2} s^{-1}	All
A_j	Electron transport limited rate of CO_2 assimilation	μmol m^{-2} s^{-1}	All
J_m	Rate of electron transport following Equation (3.6)	μmol m^{-2} s^{-1}	Jm models
PROCESS MODEL CONSTANTS			
R	Universal gas constant (8.314 J K^{-1} mol^{-1})	J K^{-1} mol^{-1}	All
α_{leaf}	Absorptance of leaf photosynthetic pigments (0.85)	unitless	All
f	Partitioning of energy between PSII and PSI (0.5)	unitless	Jf Models
PROCESS MODEL PARAMETERS IDENTIFIED BY BAYESIAN ESTIMATION			
R_d (R_{d25})	Respiration rate in the dark (standardized to 25°C)	μmol m^{-2} s^{-1}	All
Γ^* (Γ^*25)	CO_2 photocompensation point (standardized to 25°C)	Pa	All
K_c (K_{c25})	Michaelis-Menten constant for Rubisco for CO_2 (standardized to 25°C)	Pa	All
K_o (K_{o25})	Michaelis-Menten constant for Rubisco for O_2 (standardized to 25°C)	kPa	All
$Ei's$ ($_{Kc}$, $_{Ko}$, $_{Rd}$, $_{Vcmax}$, $_{\Gamma^*}$, $_{Jmax}$, $_{gm}$)	Activation energy used in Arrhenius function	KJ mol^{-1}	Temp Models
g_m (g_{m25})	Mesophyll conductance to CO_2 (standardized to 25°C)	μmol m^{-2} $s^{-1}$$Pa^{-1}$	CiCc Models
V_{cmax} (V_{cmax25})	Maximum rate of carboxylation (standardized to 25°C)	μmol m^{-2} s^{-1}	All
J_{max} (J_{max25})	Maximum rate of electron transport (standardized to 25°C)	μmol m^{-2} s^{-1}	Jm models
ϕ_J	Quantum yield estimate using Equation (3.6)	e^- $photon^{-1}$	Jm models
θ_J	Curvature factor photosynthetic light response curve	unitless	Jm models

Equations described in ***Table 2****. Model coding described in* ***Table 3****.*

eight photosynthesis models with each model and genotype run independently. Parameters undergoing individual and genotypic estimation employed normal distributions following:

$$\mu Y_i \sim N(\mu Y_{geno}, \tau_{\mathrm{Ygeno}}) \quad (2)$$

where μY_i is individual level parameters means, μY_{geno} are genotypic parameters means and $\tau_{\mathrm{Y}geno}$ is the genotypic precision ($1/\sigma^2$). **Table 4** shows the parameters estimated at both the individual and genotypic level as well as those only estimated genotypically. The choice of parameters estimated at genotypic level considered both evolutionary constraints for K_c and K_o (Galmes et al., 2005) and an analysis of trait variance using a suite of non- hierarchical models for $E_{i'}$s (data not shown). The variance priors for individual level parameters used weakly informed gamma distributions. A weakly informed Folded-Cauchy distribution (implemented as a truncated t-distribution with one degree of freedom) was used to describe prior distribution for process model variance structure (τ) (Gelman, 2006). This was centered at 0, set at the range of [0,∞) with a standard deviation of 2.5.

Photosynthesis Models

Within plant physiology and earth system science the Farquhar, von Caemmerer and Berry model of photosynthesis (FM) stands out for its mechanistic, principally biophysical/biochemical, basis for modeling C3 photosynthesis (Farquhar et al., 1980; von Caemmerer, 2000). FM, developed at the leaf scale, originally proposed two rate limiting factors controlling A by finding the minimum of RuBisCO limited A, A_c, and RuBP regeneration limited A, A_J (Farquhar et al., 1980), with a g_m limitation added subsequently (Ethier and Livingston, 2004). A triose phosphate utilization limitation (*TPU*) of A (Sharkey et al., 1985) was considered using a similar quadratic structure. Results showed *TPU* affecting A at greater than 93.6 μmol m^{-2} s^{-1}, above the maximum A_n (70.3 μmol m^{-2} s^{-1}) found in our data, and therefore a *TPU* limitation was not included.

TABLE 2 | List of equations used in eight photosynthesis models.

Equation No.	Equation	Description (models using equation)
3.1	$A_{exp} = min(A_c, A_j)$	Expected CO_2 assimilation rate as minimum of 2 limiting factors (All)
3.2	$A_i = \frac{-b+\sqrt{b^2-4ac}}{2a}$	General quadratic form for solving A_c, A_j (All)
3.3*a*	$a = \frac{-1}{g_m}$ $b = \frac{V_{cmax}-R_d}{g_m} + C_i + K_c(\frac{1+O}{K_o})$ $c = R_d(C_i + K_c(\frac{1+O}{K_o})$	A_C solution using intercellular CO_2 (C_i) and mesophyll conductance (g_m) (CiCc_Jf, CiCc_Jm, CiCc_Jf_Temp, CiCc_Jm_Temp)
3.3*b*	$a = \frac{-1}{g_s}$ $b = \frac{V_{cmax}-R_d}{g_s} + C_a + K_c(\frac{1+O}{K_o})$ $c = R_d(C_a + K_c(\frac{1+O}{K_o})$	A_C solution using ambient CO_2 (C_a) and stomatal conductance (g_s) (CaCc_Jf, CaCc_Jm, CaCc_Jf_Temp, CaCc_Jm_Temp)
3.4*a*	$a = \frac{-1}{g_m}$ $b = \frac{\frac{J_j}{4}-R_d}{g_m} + C_i + 2\Gamma^*$ $c = R_d(C_i + 2\Gamma^*) - \frac{J_j}{4}(C_i - \Gamma^*)$	A_J solution using C_i and g_m (CiCc_Jf, CiCc_Jm, CiCc_Jf_Temp, CiCc_Jm_Temp)
3.4*b*	$a = \frac{-1}{g_s}$ $b = \frac{\frac{J_j}{4}-R_d}{g_s} + C_a + 2\Gamma^*$ $c = R_d(C_{iobs} + 2\Gamma^*) - \frac{J_j}{4}(C_a - \Gamma^*)$	A_J solution using C_a and g_s (CaCc_Jf, CaCc_Jm, CaCc_Jf_Temp, CaCc_Jm_Temp)
3.5	$J_f = \frac{F'_m-F_s}{F'_m} fQ\alpha_{leaf}$	Chlorophyll fluorescence derivation of electron transport rate (ETR) (J_f) (CiCc_Jf, CaCc_Jf, CiCc_Jf_Temp, CaCc_Jf_Temp)
3.6	$a = \theta_J$ $b = -(Q * \phi_J * \alpha_{leaf}) - Jmax$ $c = Q\phi_J * Jmax$	Quadratic roots for whole chain ETR (J_m) (CiCc_Jm, CaCc_Jm, CaCc_Jm_Temp, CiCc_Jm_Temp)
3.7	$P = P_{25} exp[\frac{E_i(T-298)}{298RT}]$	Arrhenius temperature response for parameter (P) (CiCc_Jf_Temp, CiCc_Jm_Temp, CaCc_Jf_Temp, CaCc_Jm_Temp)

Modeling equations are presented in **Table 2**, and all approaches used the general quadratic form of FM, Equation (3.2). All models predicted both A_c and A_J but varied in their inclusion of three components: (1) the use of a temperature constraint on model parameters, (2) the inclusion or absence of a g_m limitation, and (3) the derivation for estimating A_J. This 2^3 design leads to eight modeling formulations when accounting for all combinations. First a subset of models included a temperature constraint on parameters, others differed in the inclusion or absence (assuming infinite) of a g_m limitation (Equations 3.3 a or b and 3.4 a or b), and finally a subset used two alternative characterizations of ETR, following either Equation (3.5) or (3.6) (**Table 2**). Models were developed on all combinations of these three assumptions (**Table 3**). Models with a g_m limitation were coded CiCc, while infinite g_m models were coded CaCc. Models using Equation (3.5), fluorescence derived ETR, were coded with a Jf, and models using Equation (3.6) to derive ETR were coded with a Jm. Models with a temperature constraint on parameters had an added Temp in model identifier.

Mesophyll conductance has been demonstrated to limit A and has been integrated into the FM (Ethier and Livingston, 2004; Flexas et al., 2008). Increasingly, g_m is being shown to impact photosynthesis under stressful conditions (Flexas et al., 2008; Niinemets et al., 2009b; Tomás et al., 2014). It is, however, still common practice to assume an infinite g_m (Thornton et al., 2005; Kattge et al., 2009) due to limited knowledge of its interspecific variation and dynamics as well as the challenges and costs associated with some estimation techniques (Niinemets et al., 2009a; Gu and Sun, 2014; Hanson et al., 2016). Here we integrated a g_m limitation in four models using a curve fitting approach

TABLE 3 | Coding of eight photosynthesis models based on three contrasting assumptions with the total number of structural parameters in each model.

Model	Temperature constraint on parameters	g_m limitation estimated	ETR derived using Equation (3.5)	ETR derived using Equation (3.6)	Number of structural parameters
CiCc_Jf		X	X		6
CiCc_Jm		X		X	9
CaCc_Jf			X		5
CaCc_Jm				X	8
CiCc_Jf_Temp	X	X	X		12
CiCc_Jm_Temp	X	X		X	16
CaCc_Jf_Temp	X		X		10
CaCc_Jm_Temp	X			X	14

TABLE 4 | Prior probability distributions of parameters used in eight photosynthesis models.

Parameter	Prior distribution dnorm (mean, precision)	Prior type	Citation(s)
E_{gm}	Dnorm (49.6, 0.1)	Broadly informed (geno)	Ethier and Livingston, 2004; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
E_{Jmax}	Dnorm (46.1, 0.01)	Broadly informed (geno)	Leuning, 2002; Medlyn et al., 2002a; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
E_{Kc}	Dnorm (70.4, 0.5)	Broadly informed (geno)	von Caemmerer, 2000; Ethier and Livingston, 2004; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
E_{Ko}	Dnorm (36.0, 0.5)	Broadly informed (geno)	von Caemmerer, 2000; Ethier and Livingston, 2004; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
E_{Rd}	Dnorm (63.9, 0.1)	Broadly informed (geno)	Bernacchi et al., 2001; Ethier and Livingston, 2004; Sharkey et al., 2007; Zhu et al., 2011
E_{Vcmax}	Dnorm (65.4, 0.5)	Broadly informed (geno)	Leuning, 2002; Medlyn et al., 2002a; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
E_{Γ^*}	Dnorm (26.8, 0.5)	Broadly informed (geno)	von Caemmerer, 2000; Ethier and Livingston, 2004; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
$g_{m(25)}$	Dnorm (2.5, 0.025)	Broadly informed (ind)	Ethier and Livingston, 2004; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
$J_{max(25)}$	Dnorm (171,0.000308)	Well-informed, C3 crops (ind)	Wullschleger, 1993
$K_{c(25)}$	Dnorm (27.24, 0.5)	Broadly informed (geno)	von Caemmerer, 2000; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
$K_{o(25)}$	Dnorm (30400,1.0 $\times$ 10^{-5})	Broadly informed (geno)	von Caemmerer, 2000; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
$R_{d(25)}$	Dnorm (1.17, 1)	Broadly informed (ind)	Zhu et al., 2011
$V_{cmax(25)}$	Dnorm (90, 0.000625)	Well-informed, C3 crops (ind)	Wullschleger, 1993
$\Gamma^*_{(25)}$	Dnorm (3.86,10)	Broadly informed (ind)	von Caemmerer, 2000; Sharkey et al., 2007; Patrick et al., 2009; Zhu et al., 2011
θ_J	Dnorm (0.8, 10)	Broadly informed (ind)	Lambers et al., 2008
ϕ_j	Dnorm (0.4, 10)	Broadly informed (ind)	Lambers et al., 2008

Descriptions of parameters are given in ***Table 1****. Prior distributions are based on summaries of literature listed in citations. Distributions are described as broadly informed from data on C3 species, well-informed from data on C3 crops. Prior type indicates the traits hierarchical level as genotypic estimation only (geno) or genotype and individual plant level (ind). Prior distributions for individual level traits describe genotype level mean.*

following Equations (3.3a) and (3.4a) (Pons et al., 2009). In a subset of these g_m itself was given a temperature dependency (Bernacchi et al., 2002).

Chlorophyll fluorescence measurements are widely used in plant physiological investigations, including for the quantification of photosystem II (PSII) operating efficiency and fluorescence derived ϕ_{PSII} (Genty et al., 1989; Maxwell and Johnson, 2000). There remains considerable uncertainty in using the fluorescence derivation of ETR (Equation 3.5) for describing A_J (Maxwell and Johnson, 2000; Baker, 2008). Here we tested the utility of fluorometry calculated ETR based on two assumptions. First, leaf absorptance (α_{leaf}), assuming α_{blue} and α_{red} of 0.92 and 0.87 respectively, was used to establish absorbed photosynthetically active photon flux density (Q) (He et al., 2007). Second, the partitioning of energy between photosystem I (PSI) and PSII (f) was assumed equal at 0.5. Because of the limitations in using fluorometry derived ETR when no alternate electron routes are included, we also considered a classically derived empirical model to estimate ETR, following Equation (3.6) (von Caemmerer, 2000). This formulation required the parameterization of J_{max}, ϕ_J and a light response curvature parameter (θ_J).

Temperature dependencies have been developed and demonstrated for RuBisCO activity, mediated by the Michaelis-Menten enzymatic constants K_c and K_o, as well as V_{cmax}, R_d, Γ^*, g_m, and J_{max} (Bernacchi et al., 2002; Medlyn et al., 2002a). For models that assume a temperature constraint, we used an Arrhenius style temperature response function, Equation

(3.7) (Bernacchi et al., 2002; Medlyn et al., 2002b; Patrick et al., 2009). This simple model required the estimation of one temperature response parameter (E_i) representing the activation energy. Estimates of temperature dependency were made for six parameters in total (V_{cmax}, J_{max}, R_d, Γ^*, g_m, K_c, *and* K_o) with E_{Vcmax}, E_{Rd}, E_{Γ^*}, E_{Kc}, *and* E_{Ko} estimated in all Temp models, E_{Jmax}, estimated in Jm_Temp models, and E_{gm} estimated CiCc_Temp models.

Model Computation

We used Gibbs Sampling, a Markov Chain Monte Carlo (MCMC) method, to generate the posterior distributions of parameters (θ) and errors (Gelman and Rubin, 1992; Kruschke, 2010). Sampling was conducted with rjags within the R Foundation for Statistical Computing (Plummer, 2014; R Development Core Team, 2014). After a burn in period of 200,000 iterations, four independent MCMC chains were run for 250,000 iterations for each model by genotype. Each chain was sampled every 20th frame yielding 50,000 samples per model per genotype. Across models, a univariate potential scale reduction factor ($\sqrt{\hat{R}}$) provided a convergence diagnostic for each parameter. A multivariate potential scale reduction factor ($\sqrt{\hat{R}_M}$) provided a single convergence metric for the entire model (Brooks and Gelman, 1998). In all we evaluated eight alternative model structures on 31 individuals from six genotypes of *B. rapa* resulting in 248 unique parameterizations of *A/Ci* response. Diagnostics of the four-chain convergence were conducted using visual inspection of trace and density plots demonstrating chain convergence in similar sample space. Chain convergence diagnostic tools $\sqrt{\hat{R}}$ and $\sqrt{\hat{R}_M}$ did not exceed the recommended maximum of 1.2, with maximums across all individuals and models of 1.06 and 1.01 for the univariate and multivariate convergence statistics respectively.

Model Scoring Metric

To quantitatively compare model results for each individual, genotype, and species, we used the Deviance Information Criterion (DIC), a Bayesian analog to Akaike Information Criterion (Spiegelhalter et al., 2002). DIC considers all models to be conceptually equal and acts purely as a model scoring metric not an analytical evaluation of model functional form (Gelman et al., 2004). DIC is calculated by combining a deviance term and a complexity penalty term (Spiegelhalter et al., 2002). The Bayesian model deviance ($D(\theta)$) is based on the residuals between the model and the data, computed with

$$D(\theta) = -2\log\left[p(Y|\theta)\right] + 2\log\left[f(Y)\right] \quad (3)$$

where Y is observed data, θ represents all parameters for the model, $p(Y|\theta)$ is the likelihood function defined by the model, and $f(Y)$ is a standardizing term remaining constant for all models and therefore having no influence model comparison. The Bayesian deviance alone is not a strong model discrimination metric as higher dimensional models could be favorably biased. DIC attempts to account for this bias with a parameterization penalty (Spiegelhalter et al., 2002; Plummer, 2008). The penalty, plug-in deviance (pD), from the Spiegelhalter derivation is $pD = \overline{D(\theta)} - D(\overline{\theta})$, where $D(\theta)$ is posterior mean of the deviance using all parameters samples of sequence and $D(\theta)$ is deviance evaluated at the posterior mean of the all parameters. In the calculation of DIC the posterior distribution of $D(\theta)$ is used to express mean deviance following

$$DIC = \overline{D(\theta)} + 2pD. \quad (4)$$

$\overline{D(\theta)}$ and $\overline{D(\theta)}$ are easily calculated from MCMC output through the monitoring of $D(\theta)$ of all simulated values. ΔDIC is calculated as the difference between model DIC score and the genotype minimum DIC score. No significant ΔDIC has been universally accepted, however differences of ten are often employed (Spiegelhalter et al., 2003) and used here.

Parameter Variability

We evaluated full model sets of trait posteriors through the development of multiple posterior predictive checks; using the posterior trait distributions to simulate A, A_{exp}, while considering the uncertainty in posterior distributions (Kruschke, 2013). Two methods were then used to compare posterior trait distributions. First, boxplots of posterior parameter distributions were compared across genotypes and against the prior probability distributions. Second, high-density intervals (HDIs) were used at eight percentiles (50, 60, 70, 80, 85, 90, 95, and 99%). HDI is a Bayesian posterior comparison metric identifying portions of the posterior distributions having a higher probability density than regions outside that interval (Kruschke, 2010). To describe the relative credibility of trait variance the differences in the posterior mean distributions were taken for all traits within a given model (Kruschke, 2013). This difference was then evaluated for intersection with zero at the eight percentiles listed above. The maximum HDI percentile of the posterior trait differences, which did not intersect with zero, was used to describe degree of credible trait variance between two genotypes. This HDI differencing test was conducted across all genotypes and traits within each model.

Model Sensitivity Analysis

To evaluate consistency in model performance a sensitivity analysis was conducted. The entire analysis was rerun after adding Gaussian noise with a mean of 0.0 and standard deviation of 2.0 μmol m^{-2} s^{-1} to the A_n data. This was chosen to mimic instrumentation error of IRGA observations. All statistical analysis was conducted in the in R software environment (R Development Core Team, 2014).

RESULTS

Model Performance

Posterior parameter distributions were used to predict A_c and A_J for each model to compute A_{exp} at individual and genotypic levels. The genotype level mean standard error of A_{exp} for all models was 2.75 μmol m^{-2} s^{-1} with a minimum of 0.9 μmol m^{-2} s^{-1} for *r46* in the CiCc_Jm model and a maximum of 7.29 μmol m^{-2} s^{-1} for *tur* in the CiCc_Temp_Jf model. The *A/Ci* observations (A_n vs. C_i or C_a) of the modeled data are shown along with 95% genotypic credible intervals

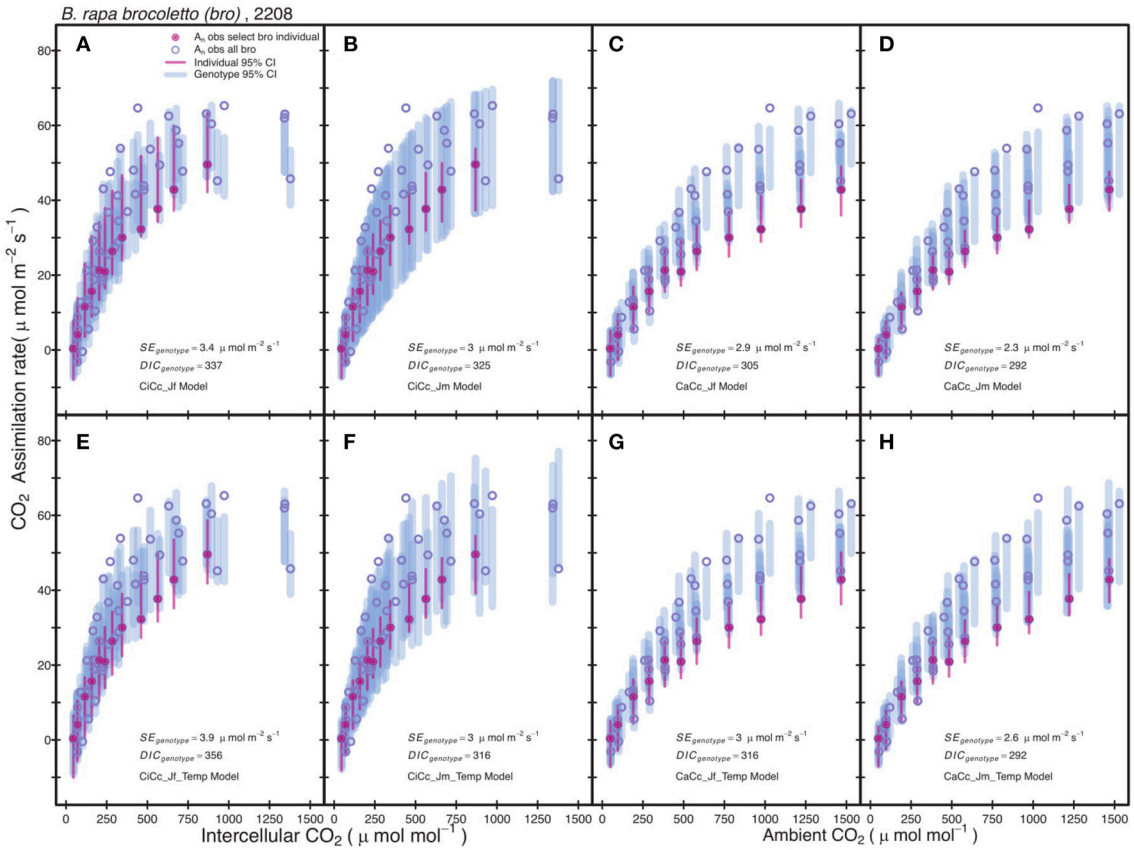

FIGURE 2 | Comparison between observed CO_2 assimilation (A_n) and estimated assimilation for eight photosynthesis models of *B. rapa,* var. *bro*, including a single individual level estimate. Each plot shows A_n for a select *bro* individual (filled-circle) and A_n for all other *bro* data (open-circle). The Bayesian 95% credible intervals (vertical lines) are shown at individual level (thin-line) and genotypic (thick-line). Models differ in three assumptions: **(A–D)** have no temperature constraint on parameters while **(E–H)** use an Arrhenius style temperature constraint on model parameters, **(A,C,E,G)** use an estimate of ETR derived from Equation (3.5) based on chlorophyll fluorescence, **(B,D,F,H)** estimate ETR from Equation (3.6), **(A,B,E,F)** predict mesophyll conductance limitations using intercellular CO_2 observations while **(C,D,G,H)** assume infinite mesophyll conductance using ambient CO_2 observations.

(CI) for *bro* along with 95% individual CI for selected *bro* individual (**Figure 2**). Individual level CI's fell mostly within genotypic CI in **Figure 2**, with the exception of CaCc Jf models where individual CI's fell below genotypic at high CO_2 availability.

Genotypic posterior trait distributions were used to construct 95% CIs on A_{exp}. **Figure 3** shows the *A/Ci* observations (A_n vs. C_i or C_a) for all *r46* and *r301* individuals with the 95% CI for each model. A narrower range in 95% CIs was found in models assuming infinite g_m for *r46* and *r301* (**Figure 3**) as well as crop accessions (data not shown). Models estimating ETR using fluorescence (Jf models) showed lower overall 95% CIs on *A* than Jm models in all genotypes except *r46*. This is seen in the larger number of points beyond the upper CI limit for *r301* in **Figures 3A,B**; this same result was also found across crop accessions (data not shown). Finally, to evaluate genotypic vs. species level parameterization, an accession level parameterization was developed and used to predict the RILs A/C_i response (**Figure 4**). Data from both RILs, most noticeably *r301*, fell outside the 95% accession based CI (**Figure 4**).

Model Structural Comparison

Genotypic model DIC scores were used to compute ΔDIC along with genotype pD's (**Table 5**). For the species, the CaCc_Jm_Temp and the CaCc_Jm models were top tier models for four of the six genotypes, with the exceptions being *cab* and *r301*. CiCc_Jm, CaCc_Jf, and CaCc_Jf_Temp were each included in one genotypes' top-tier, *r301*, *oil* and *oil,* respectively. CiCc_Jf, CiCc_Jf_Temp, and CiCc_Jm_Temp failed to have a ΔDIC of less than 10 across genotypes. The pD's were consistently highest in models deriving ETR using Eqn 3.6 relative to fluorescence derived ETR (Equation 3.5). pD's were also consistently higher in models assuming a g_m limitation relative to infinite g_m.

Parameter Differentiation

Genotypic posterior trait distributions were compared using posterior boxplots and an analysis of HDI's. Trait distributions showed some parameters with high probability of genotypic variation, including J_{max}, V_{cmax}, Γ^*, and E_{Vcmax}, and traits with limited probability of variation, including K_o, ϕ_J and θ_J (**Figures 5–8**). V_{cmax}, Γ^*, R_d, K_c, and K_o were the five traits

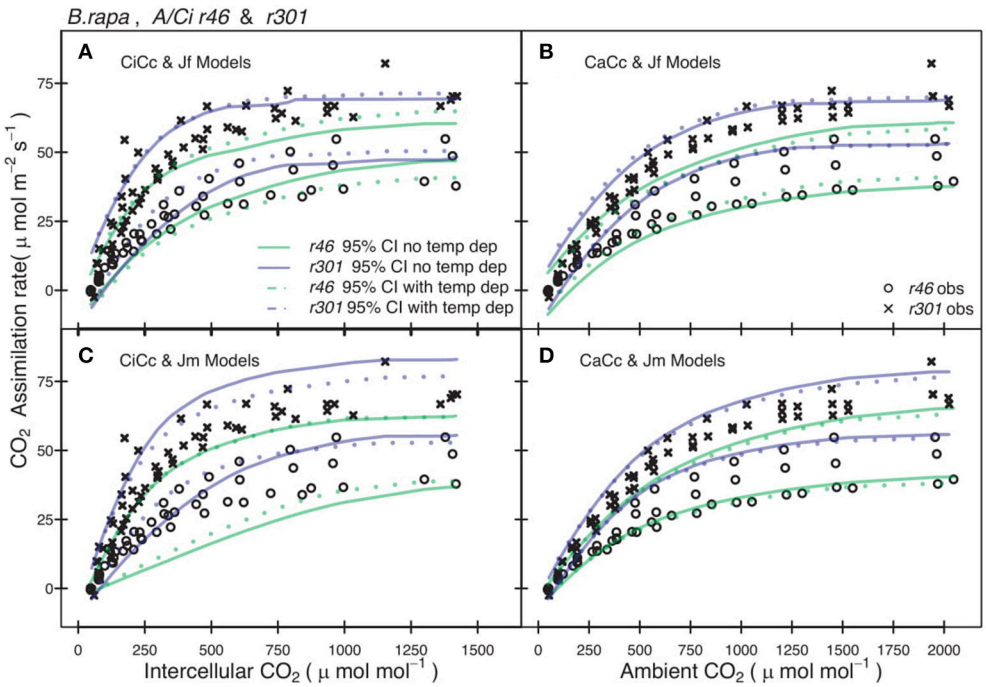

FIGURE 3 | Comparison between observed CO_2 assimilation (A_n) and predicted 95% credible interval (CI) for eight models of two *B. rapa* genotypes (*r46*, *r301*). Each plot shows A_n for both *r46* (open circles) and *r301* (closed circles) and Bayesian 95% CI for two models one with temperature constraint (dotted lines) the other without (solid lines), using genotype level posterior trait distributions. **(A)** Models shown incorporated an estimate of mesophyll conductance (g_m) and used chlorophyll fluorescence (Equation 3.5, **Table 2**) to characterize electron transport rate (ETR). **(B)** Models assumed an infinite g_m and used Equation (3.5) to characterize ETR. **(C)** Models shown incorporated an estimate of g_m and used Equation (3.6) to characterize ETR. **(D)** Models assumed an infinite g_m and used Equation (3.6) to characterize ETR. Smoothed CI's using Loess fit to present multiple models together.

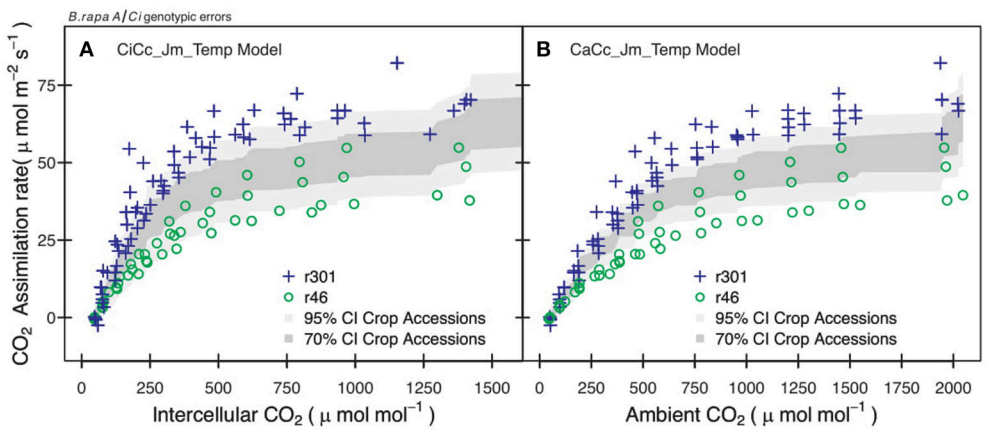

FIGURE 4 | Examples of errors from non-genotypic parameterization for two *B.rapa* genotypes. **(A)** Observation points of *r46* (open circles) and *r301* (plus) with 95% (light gray) and 70% (dark gray) credible interval for CiCc_Jm_Temp model using combined parameter distributions from agricultural accessions (*bro*, *cab, oil, tur*). **(B)** Observation points of *r46* and *r301* with 95% and 70% credible intervals for CaCc_Jm_ Temp model using combined posterior parameter distributions of accessions.

estimated in all eight models. V_{cmax} showed genotypic variance in all models with *r46* notably lower than other genotypes in all Jf models (**Figures 5**, **7**). Γ^* also showed genotypic variance across models with lower estimates in g_m limited models (**Figures 7**, **8**) than infinite g_m models (**Figures 5**, **6**). R_d showed genotypic variance in five of the eight models, most pronounced in models CaCc_Jf and CaCc_Jf _Temp (**Figure 5**), while the only variance seen in K_c was in the CaCc_Jm model, with *r46* differing from *oil* (**Figure 6**). The temperature activation energies (*Ei*'s) showed limited probability of genotypic variance with two exceptions.

TABLE 5 | Genotype DIC increment with respect to genotype minimum (ΔDIC) with mean effective number of parameters (pD) for eight models.

	Genotype ΔDIC (pD)					
Model	***r301***	***r46***	***bro***	***cab***	***oil***	***tur***
CiCc_Jf	165.0 (11.5)	129.6 (9.7)	45.1 (13.4)	115.6 (12.0)	50.5 (5.7)	113.4 (8.9)
CiCc_Jm	**0** (20.8)	68.4 (28.7)	32.9 (12.7)	17.8 (16.0)	27.3 (8.9)	21.3 (53.0)
CaCc_Jf	138.5 (9.1)	85.1 (6.5)	13.2 (5.4)	81.1 (5.2)	**0** (7.9)	96.8 (8.1)
CaCc_Jm	68.1 (16.4)	**0** (12.4)	**0** (12.9)	**0** (9.0)	23.3 (11.5)	31.4 (20.1)
CiCc_Jf_Temp	210.2 (7.0)	129.1 (5.6)	64.5 (11.5)	156.0 (8.5)	45.2 (10.7)	182.7 (6.5)
CiCc_Jm_Temp	17.6 (17.0)	54.8 (15.8)	24.4 (11.4)	39.1 (29.3)	21.2 (13.4)	15.1 (24)
CaCc_Jf_Temp	134.9 (10.1)	98.7 (5.5)	17.0 (8.2)	81.3 (12.9)	**5.8** (9.4)	70.0 (6.3)
CaCc_Jm_Temp	60.4 (16.6)	**9.9** (13.8)	**0.6** (8.8)	15.3 (19)	18.3 (7.5)	**0** (14.3)

Bolded are models with top-tier ΔDIC scores for respective genotypes using ΔDIC threshold of 10.

E_{Vcmax} for *r46* was lower relative to other genotypes in the CiCc_Jf_Temp model and estimates in *r46* and *cab* were also lower in the CiCc_Jm_Temp models (**Figure 8**) E_{Jmax} also showed variance in models CiCc_Jf_Temp and CiCc_Jm_Temp (**Figure 8**). Amongst the ETR traits modeled using Equation (3.6) only J_{max} showed genotypic variation, this was found across all Jm based models (**Figures 6**, **8**), the variation was dominated by higher estimates for *r301*.

To describe the magnitude of genotypic trait variance, the differences in posterior parameter distributions among genotypes were computed for each model. These differences were then evaluated at eight HDI percentiles for overlap with zero; the maximum HDI interval not overlapping with zero was selected as the probability of variance. J_{max}, V_{cmax}, Γ^*, and E_{Vcmax} were found with a probability of variance at 95% HDI (**Figure 9**). At 80% E_{Jcmax} and g_m show differences, at 70% R_d showed differences and at 50% HDI K_c emerges as variable (**Figure 9**). At 50% HDI half of the 16 traits estimated showed variance. To summarize the posterior trait distributions across models **Table 6** lists maximum HDI percentile of traits differences. Of note in **Table 6**, J_{max} is classified as highly variable, differences found at <90% HDI, in all Jm based models. The variability in J_{max} is dominated by the contrast between the two RILS, with *r46*'s median posterior between 50-100 μmol m^{-2} s^{-1} <*r301*'s (**Figures 6**, **8**). Variance in V_{cmax} is dominated by differences in *r46* relative to other genotypes; most notable in the CiCc_Jf, CaCc_Jf models (**Figures 5**, **7**).

Sensitivity Analysis Results

Gaussian noise with mean of 0.0 μmol m^{-2} s^{-1} and standard deviation of 2.0 μmol m^{-2} s^{-1} was added to the A_n data, followed by a re-analysis. The resultant posterior parameter distributions were wider in some cases and some shifts in median estimates were seen, but no systematic trends were identified in these shifts. For example, in traits that play critical roles in the A/C_i response, the J_{max} noisy genotypic level median estimate was 2.2 μmol m^{-2} s^{-1} greater compared to the original analysis and for V_{cmax} the noisy median estimate was 4.4 μmol m^{-2} s^{-1} less than original analysis in the CiCc_Jm_Temp model. This is illustrated in comparing **Figures 10A,C** with **Figures 10B,C**.

DISCUSSION

Multimodel Approach

We show here that a multimodel based approach improves phenotypic information discovery in three critical ways. First, our trait analysis using a set of models identified potential genotypic differences requiring further investigation and revealed model components needing reevaluation (**Table 5**). Second, the Bayesian parameterization scheme revealed an expected trait hierarchy (**Table 6**). Finally, the multimodel approach provided greater confidence in estimates of trait variation among genotypes (**Figure 9**).

Performance of Models Based on Assumptions

The complexity analysis assessed the influence of factors not addressed experimentally (i.e., temperature) and of physiological mechanisms (i.e., g_m and ETR derivation) not yet characterized in the population under study. While a single preferred model structure was not identified using ΔDIC, we were able to evaluate the relative performance of model assumptions employed. First, DIC strongly favored derivation of ETR, and therefore A_J, from Equation (3.6) (**Table 5**), as Jm based models were in the top tier in seven cases, while Jf models were rated as top-tier only two times and only for *oil*. This confirms previously identified limitations and illustrates the need to consider alternate e^- paths when using fluorometry to characterize A_J (Baker, 2008; Yin et al., 2009). Fluorometry estimates all PSII e^- excitation at the beginning of the e^- transport chain; using this to estimate the assimilatory outcome of e^- transport does not distinguish e^-'s used for photosynthetic linear electron flow and the alternative pathways of e^- transport (Miyake, 2010). The biological relevance of these alternate pathways lies in the reduction of photooxidative stress (Foyer and Shigeoka, 2011), specifically the protection of PSII from heat and light stress (Miyake, 2010). Interestingly *oil* showed a preference for CaCc_Jf based models (**Table 5**). The divergence of *oil* from the other genotypes may be due to diminished flow to alternate pathways or a unique *f*. The parameter differentiation of *oil* reflects the allelic composition of that genotype, and warrants further investigation. An expanded genotypic sample may enable model

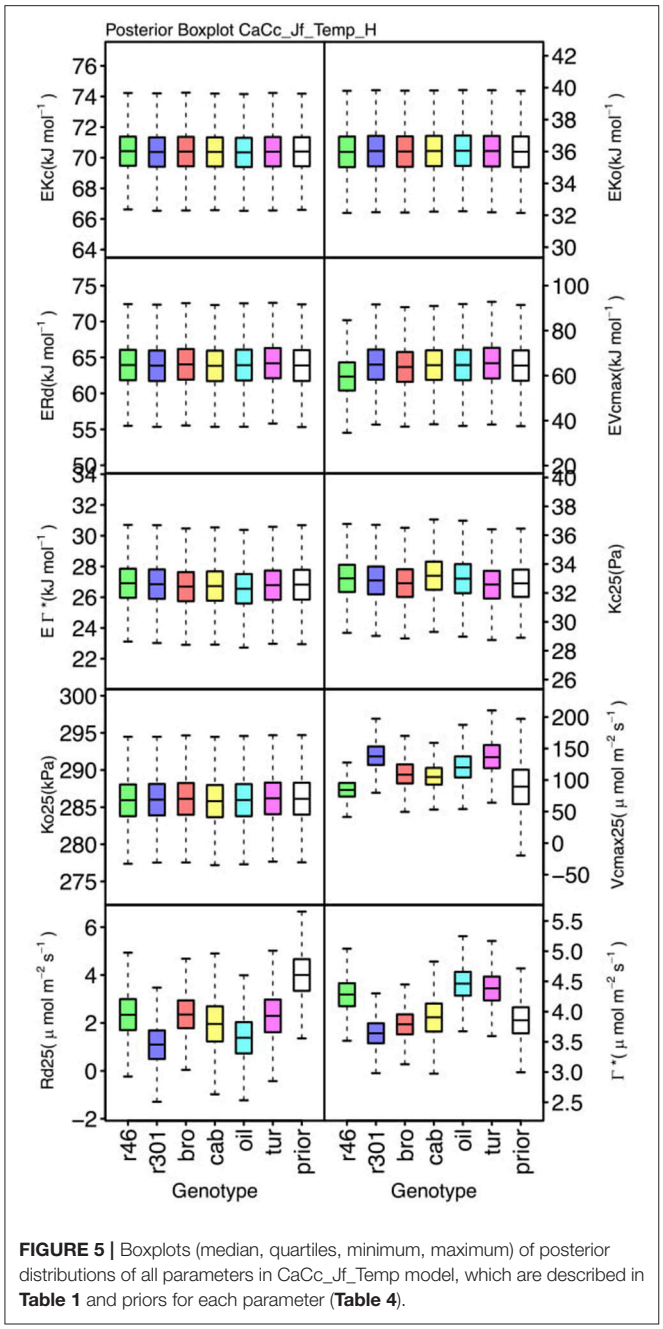

FIGURE 5 | Boxplots (median, quartiles, minimum, maximum) of posterior distributions of all parameters in CaCc_Jf_Temp model, which are described in **Table 1** and priors for each parameter (**Table 4**).

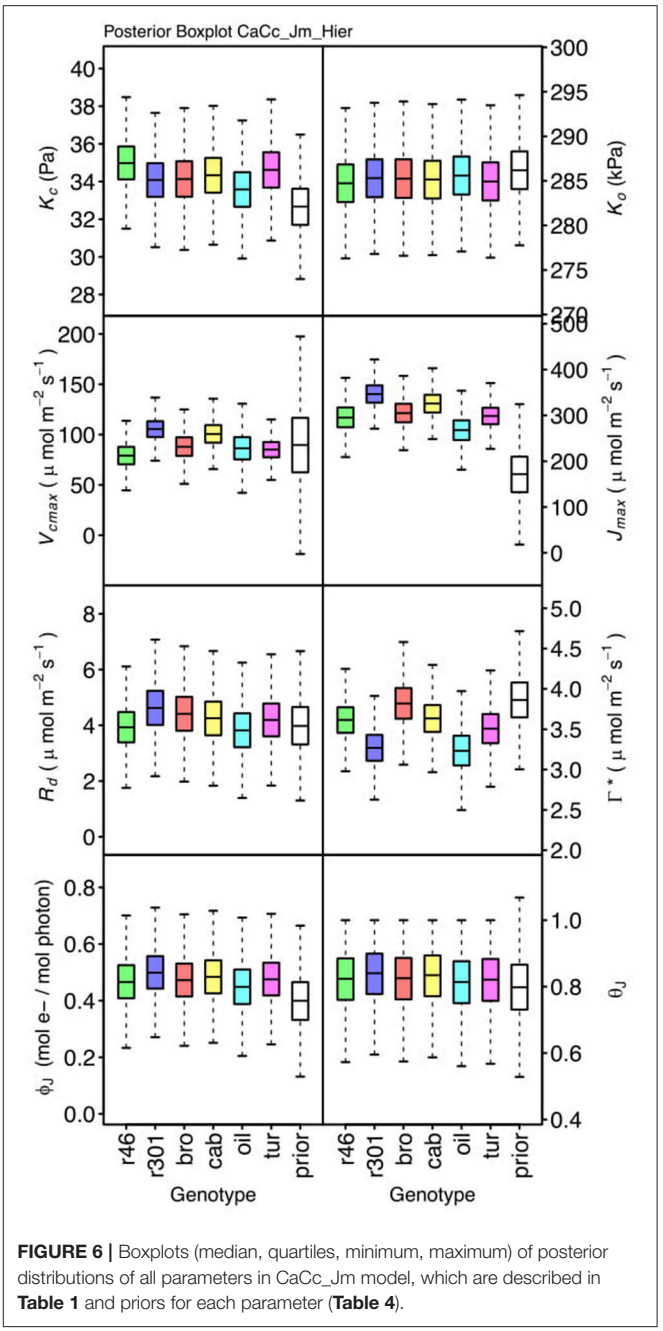

FIGURE 6 | Boxplots (median, quartiles, minimum, maximum) of posterior distributions of all parameters in CaCc_Jm model, which are described in **Table 1** and priors for each parameter (**Table 4**).

modification for investigating alternate e- flow and f (Laisk and Loreto, 1996; Yin et al., 2009; Livingston et al., 2010).

Second, the combination of chlorophyll florescence derived ETR and g_m-limitation (CiCc_Jf and CiCc_Jf_Temp) was not selected as a top tier model by any of the genotypes (**Table 5**), this shows the overall preference for both Equation (3.6) derivation of ETR and infinite g_m. Interestingly, the preference for infinite g_m models based on ΔDIC emerged even though g_m as a trait was shown to vary in this population (**Figure 8** and **Table 6**). A debate persists on the response of g_m to environmental conditions (Flexas et al., 2007; Tazoe et al., 2011) with possible mechanisms governing g_m behavior including anatomical components, biochemical changes such as aquaporin expression, and chloroplast surface area adjustment (Flexas et al., 2006; Chaumont and Tyerman, 2014; Tomás et al., 2014). Each of these may be variable within plant populations, and while both limited and ∞ g_m were viewed favorably here, further modeling work should aim for integration of g_m limitation, particularly in plants under stress and in those with intrinsically low g_m. The addition of g_m limitation increased model complexity relative to ∞ g_m counterparts, pD's in **Table 5**, in most cases, with *cab* and *oil* showing exceptions with slightly reduced or

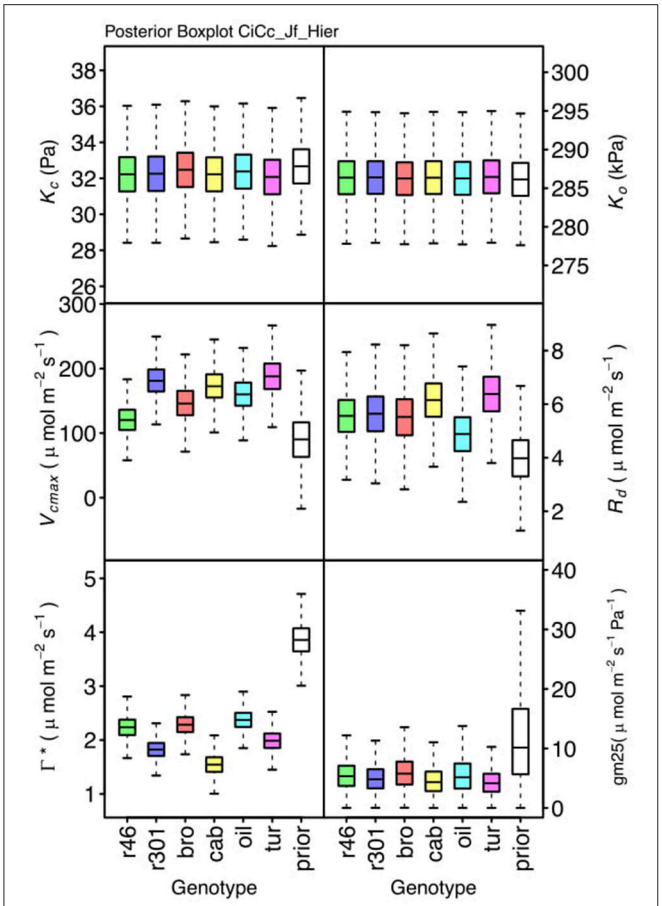

FIGURE 7 | Boxplots (median, quartiles, minimum, maximum) of posterior distributions of all parameters in CiCc_Jf model, which are described in **Table 1** and priors for each parameter (**Table 4**).

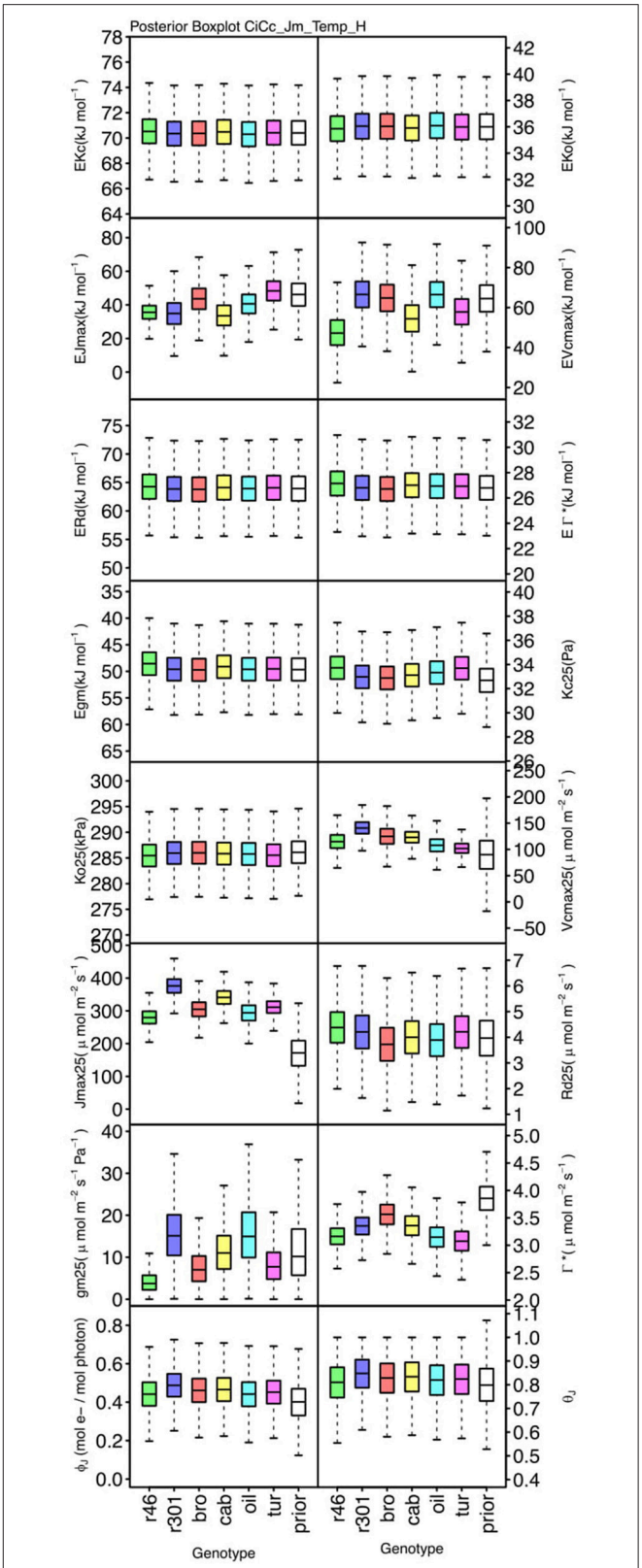

FIGURE 8 | Boxplots (median, quartiles, minimum, maximum) of posterior distributions of all parameters in CiCc_Jm_Temp model, which are described in **Table 1** and priors for each parameter (**Table 4**).

similar pD's in ∞ g_m models. The failure of g_m-limitation to improve model performance in all cases may have been expected given the lack of environmental stress and an attendant lack of strict plant regulation of g_m. If water, heat and/or salinity stress were imposed, then the increased model complexity associated with dynamic g_m may in fact have been necessary to accurately represent the A/C_i response (Grassi and Magnani, 2005; Niinemets et al., 2009b; Tomás et al., 2014).

Third, the addition of temperature constraints had limited influence on model performance as in most cases temperature limited models and their counterparts were not discriminated by ΔDIC; four of six genotypes had both temperature constrained and the unconstrained alterative in there top-tier (**Table 5**). From an empirical perspective, this is promising as it indicates that the instrumentation and methodology used distinguished trait differences among the genotypes despite any temperature differences among trials or throughout the A/C_i measurement period. Temperature constraints have been universally advocated and biochemically justified for informing parameterization (Berry and Bjorkman, 1980; von Caemmerer, 2000; Bernacchi et al., 2001; Yamori et al., 2014). Trait evaluation over greater

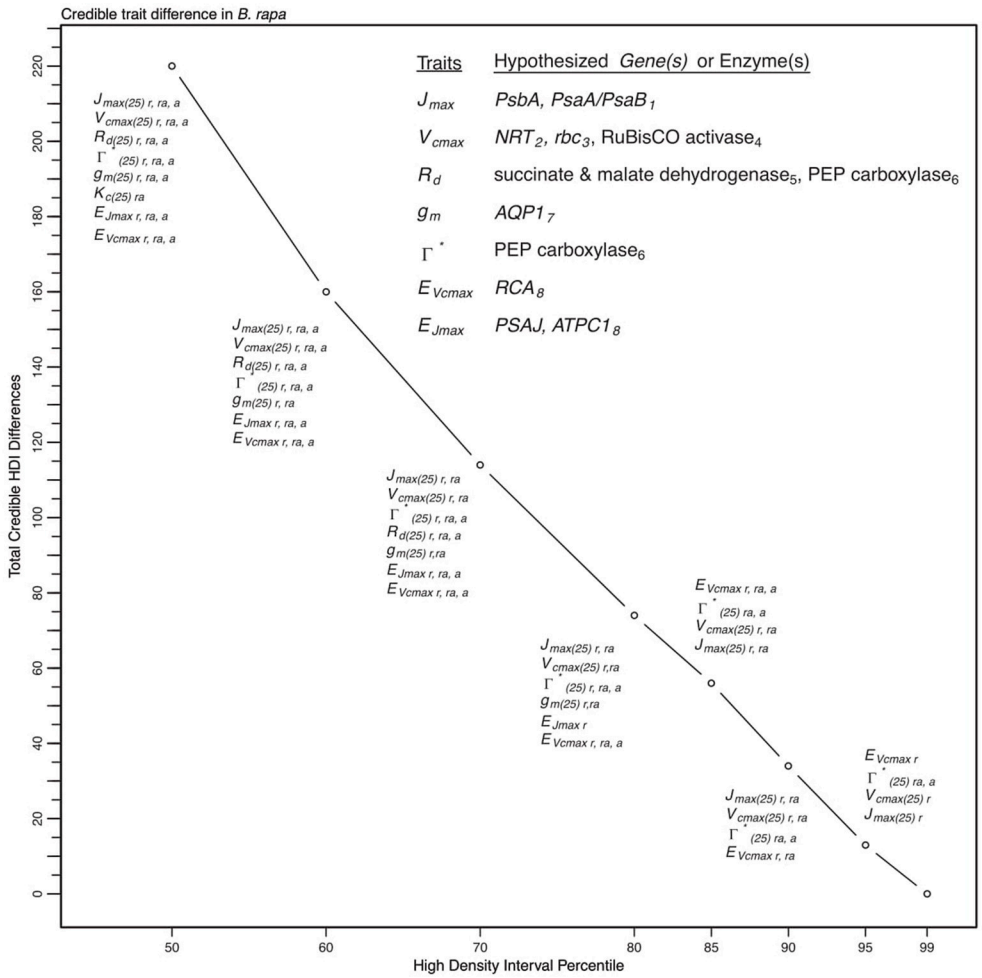

FIGURE 9 | Illustration of trait potential for variation in represented *B. rapa* population along with prospective mechanistic underpinnings. The total number of instances where HDI percentile difference for a trait did not intersect with zero across eight models at eight HDI percentiles. At each percentile, the parameters identified with credible interval differences are listed as well as indication if the difference was between RILs (*r46* and *r301*) (r), between RILs and an agricultural accession (ra) or between two agricultural accessions (a). Zero credible interval differences were observed at 99% HDI. References for hypothesized genes (italics) and enzymes in subscript as follows: 1, Foyer et al. (2012); 2, Masclaux-Daubresse et al. (2010); 3, Hauser et al. (2015); 4, Yamori et al. (2012); 5, Araujo et al. (2012); 6, Häusler et al. (1999); 7, Hanba et al. (2004); and 8, Song et al. (2014).

temperature ranges may identify where these two model classes (temperature constrained vs. unconstrained) differ in suitability.

Genotype Level Parameterization

We found differing degrees of genotypic trait variation based on evaluation of posterior distributions revealing a hierarchical structure of photosynthetic trait variation (**Figures 5–8** and **Tables 6**). Using an HDI percentile analysis $K_{o(25)}$, ϕ_J, θ_J, and E_i's, other than E_{Vcmax} and E_{Jmax} did not show genotypic variability (**Table 6**). Lack of variability in $K_{o(25)}$ reflects the limited mutational landscape for RuBisCO proteins (Studer et al., 2014) even while selection promotes diversification of other traits. The emergence of $K_{c(25)}$ as variable in two models was surprising for this reason and points for the need to reconsider the prior distributions of this trait in future analysis Non-variable results also support trait conservation for temperature dependencies with the possible exception of E_{Vcmax} and E_{Jmax} (Sharkey et al., 2007). For these temperature dependencies, Medlyn et al. (2002a) used *A/Ci* curves at different temperatures to establish E_i's; such an approach could confirm results found here. The non-variable results for estimates of ϕ_J and θ_J can be explained potentially by the lack of light variation in the *A/Ci* dataset. At saturating light conditions variation in J_{max} would be expected while the light conditions would not serve as strong drivers of ETR response for low light traits ϕ_J and θ_J (**Figures 6**, **8**). An analysis using a combined *A/Ci* and light response (LR) curve approach (Patrick et al., 2009) should inform estimates of ϕ_J and θ_J (Evans et al., 1993). Better integration of chlorophyll fluorescence data may also improve the models ability to identify genotypic variation in ETR traits. The degree of variation found in this population for J_{max} and V_{cmax} was striking, particularly as J_{max} showed variation beyond the data provided as prior (**Figure 10**) (Wullschleger,

TABLE 6 | Summary of genotypic trait variability using differencing of high density intervals (HDI).

Traits \ Models	CiCc_Jf	CiCc_Jm	CaCc_Jf	CaCc_Jm	CiCc_Jf_Temp	CiCc_Jm_Temp	CaCc_Jf_Temp	CaCc_Jm_Temp
E_{gm}								
E_{Jmax}						70		80
E_{Kc}								
E_{Ko}								
E_{Rd}								
E_{Vcmax}						80		95
$E_{\Gamma*}$								
$g_{m(25)}$		50				80		
$J_{max(25)}$		95		90		95		90
$K_{c(25)}$				50				
$K_{o(25)}$								
$R_{d(25)}$	70	70	80		50		70	
$V_{cmax(25)}$	90	80	90	85	95	90	95	95
$\Gamma^*_{(25)}$	95	95	90	85	90	80	95	80
θ_J								
ϕ_J								

Dot-pattern indicates traits not in a given model. White indicated no emergence of HDI difference at 50% HDI. Number indicates at which HDI percentile did the difference in posterior distributions no longer overlap with zero. Color coding from dark for largest HDI percentile (95) to light for smallest (50).

1993). Variability in Γ^* puts in question the continued use of constants for describing Γ^*, and supports the observation that complex diffusion pathways and potential environmental feedbacks complicate the estimation of Γ^* (Hanson et al., 2016).

Based on our analysis of this population, J_{max}, V_{cmax}, and Γ^* have a high probability of variation as multiple models described them as variable at high HDI percentiles (**Figure 9** and **Table 6**). Although we sampled the range of extreme crop phenotypes found in *B. rapa* (including cabbages with high leaf allocation, turnips with dramatic root allocation, and brocoletto and oilseed types with predominant reproductive allocation), trait distinction was highest between the two RILs, which were full siblings (**Figures 5–8**). While the parents differ in key photosynthetic traits (Edwards et al., 2011), the even greater phenotypic difference expressed between these two RILs must arise from transgressive segregation in *WUE* (Edwards et al., 2012) and reflects either novel additive effects of allelic combinations or novel epistasis (Rieseberg et al., 1999, 2003). In contrast to the highly differentiated RIL traits, crop accessions may vary more in biomass partitioning than photosynthetic traits, reflecting the targets of selection during domestication and diversification (Edwards et al., 2016; Yarkhunova et al., 2016). The fact that phenotypes are more readily distinguished between RILs (full siblings) than among crops highlights the opportunity for genetic characterization of these traits in experimental genotypes.

Our results illustrate the need for a genotypic parameterization scheme (**Figure 3**) while offering targets (**Figure 9**) for further genetic dissection. We therefore propose genomic and transcriptomic analysis to further understanding of the factors controlling observed trait distinction among identified targets. The differentiation between RILs for many traits suggests that the existing RIL population derived from a cross between an oilseed (R500) and a rapid cycling genotype (IMB211) would be an effective one in which to begin the genetic dissection of traits underlying variation in *A*. RIL populations developed from crosses between R500 × turnip, R500 × cabbage, and R500 × brocoletto that are under development will provide additional segregating lines for the genetic dissection of *A* within this crop following a process similar to one in *Zea mays* (Dell'Acqua et al., 2015). Breeding efforts focused on the mechanisms underlying variations in J_{max25} and V_{cmax25} constitute the best targets for

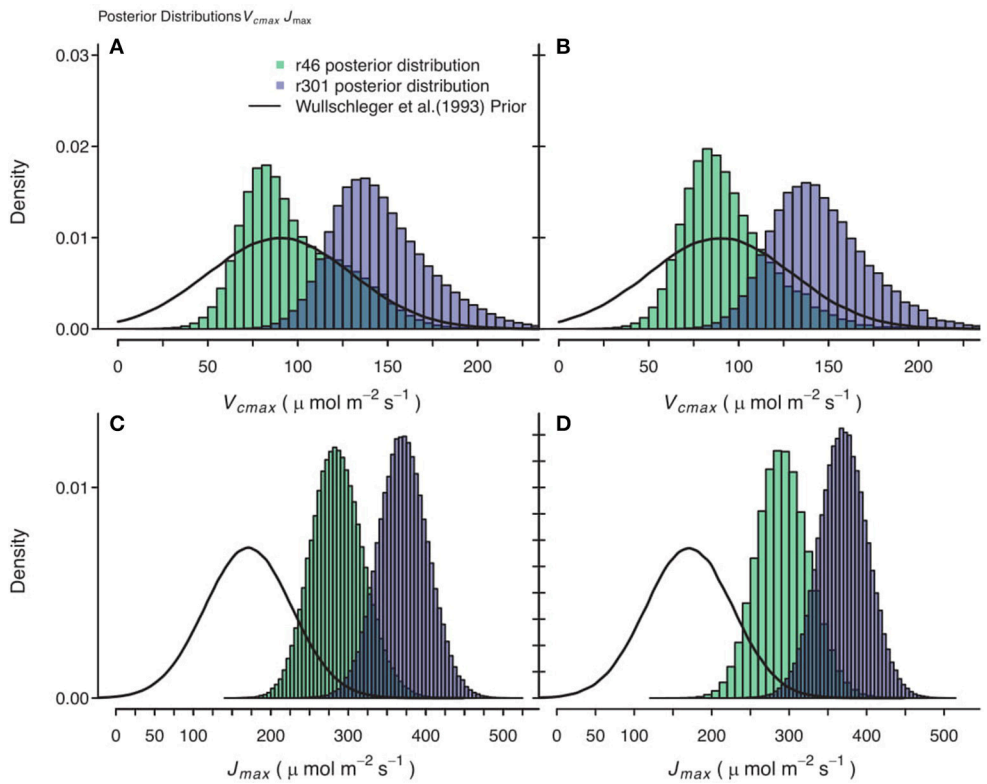

FIGURE 10 | Comparison of posterior parameter distributions of two *B. rapa* genotypes (*r46, r301*) for two photosynthetic traits, maximum rate of carboxylation (V_{cmax}) and maximum rate of electron transport Umax). **(A,C)** Show combined posterior distributions of all models for V_{cmax} and J_{max} respectively using observational data and prior on parameters based on Wullschleger (1993) for C3 crops. **(B,D)** Show combined posterior distributions of all individuals and models for V_{cmax} and J_{max}, respectively using observational data with 2.0 μmol m^{-2} s^{-1} random noise added to assimilation data and using same Wullschleger (1993) prior.

increasing A and thereby yield in agricultural crops (Long et al., 2006).

Method Limitations

The curve-fitting method of g_m estimation does not have the benefit of using alternative g_m measurement techniques based on other data types (Pons et al., 2009; Tazoe et al., 2011; Hanson et al., 2016). Estimation of g_m based on combined fluorometry/gas exchange methods should consider the consequences of alternative electron pathways for A_J, because differences between linear electron transport and total electron transport may not be entirely accounted for through g_m alone (Yin et al., 2009). State of the art methods propose a dynamic g_m responding to variations in both CO_2 partial pressure and light using variable ETR rates from chlorophyll fluorescence and/or online discrimination methods (Tazoe et al., 2011; Gu and Sun, 2014). Introduced here is a methodological approach that addresses uncertainty and enables rapid screening. Dynamic g_m models could be incorporated into the screening tool given appropriate data to address uncertainty associated with online discrimination techniques. A fully integrated photosynthesis model using linear electron flow and total electron flow from gas exchange and fluorometry observations coupled to online discrimination data may help resolve concerns related to g_m estimation (Pons et al., 2009; Tazoe et al., 2011; Gu and Sun, 2014).

Partitioning of energy between PSI and PSII (f) was assumed 0.5, an assumption that does not hold in all cases (Laisk and Loreto, 1996). This assumption complicates the understanding of variation in A_J; if the assumption is valid, then mechanics of PSII light harvesting appear to be different in *oil* relative to others, however if invalid, then *oil* may in fact have different photosystem partitioning relative to other genotypes. f could have been made a parameterized trait but lacking meaningful data we choose to set f constant. Finally, we lack independent validation of parameters. Many parameters (R_d, Γ^*, ϕ_J) can be estimated independently through alternative gas-exchange methodologies (Laisk et al., 2002, 2007; Hanson et al., 2016), while others (K_c, K_o) can be evaluated using *in-vitro* methodologies (von Caemmerer et al., 1994). Such parameter validation would provide an alternate means of assessing model suitability and could be integrated into a Bayesian framework. The practical implications of proposed trait validation, including g_m mentioned above, on large populations remain problematic monetarily and logistically.

Finally, alternatives to DIC could be considered in future multimodel comparison studies, such as the widely Applicable Information Criteria (Watanabe, 2010). DIC relies heavily on the

mean of the posterior distribution presenting some bias against posteriors with skewed distribution, potentially a problem for posteriors here including g_m, but these alternate scoring metrics are not ready implemented in rjags currently (Watanabe, 2010; Gelman et al., 2014).

Implications

Our modeling of phenotypic variation helps clarify how allelic variation impacts the expression of biophysical traits (**Figure 1**). Three improvements along the pipeline from large breeding populations to selection of genotypes with enhanced yield and stress response were identified. Two of these improvements support the use of multiple model approaches for discovering important information content not available in single model analysis. First comparison of multiple models was critical in determining differentiation of traits among genotypes. For example, the CiCc_Jm_Temp model, which is similar to a commonly utilized method (Sharkey et al., 2007), did not identify R_d as variable based HDI analysis, yet six of the remaining seven models did. Further, CiCc_Jm_Temp found g_{m25} to be variable at 80% HDI, the highest g_m variation found. Given finite resources for further investigation, our approach supports quantifying the genetic architecture of R_d within this population (**Figure 9**) while an approach solely relying on CiCc_Jm_Temp would support scrutiny of g_m. The demonstrated uncertainty in trait estimates also supports focused model improvement and/or modified experimentation. Second, potential genotypic differences were revealed using complexity analysis, which would not have been observed in single-model analysis. Specifically, complexity analysis demonstrated ETR differences in this population as some genotypes wholly selected Jm based models while others *oil* selected an alternative ETR derivation (**Table 5**; **Figures 5–8**). Pitting competing models against one another allowed specific genotypic responses to emerge and identified model components in need of revision. Moreover, testing competing mechanistic models is superior to null hypothesis testing using frequentist statistical approaches (McElreath, 2016). Finally, the posterior trait distributions represent knowledge to be preserved as one expands models. The Bayesian updating procedures of the sensitivity analysis provides a way to codify this knowledge. Information preservation further informs our understanding of plant physiology and should embolden modelers attempting to link traits relevant to plant productivity to genes. V_{cmax}, J_{max}, and g_m are hypothesized to underline genetic variation in A for 13 lines of *Ozyzo sativa* (Gu et al., 2012), as was similarly shown here for V_{cmax} & J_{max}. Indeed, both genotypic and evolutionarily conserved parameters have been advocated for crop models (Yin et al., 2004; Bertin et al., 2009; Gu et al., 2014). We can think of these as having a hierarchical organization in which genotypic parameters are clearly distinguished from conserved parameters. This hierarchy should be continually informed by both modeling output such as those provided here and through phylogenetic analysis of genomes when possible (Galmes et al., 2005).

CONCLUSIONS

The integration of data from six genotypes into eight photosynthesis models allowed for a comprehensive exploration of trait space occupied by this population. We found considerable variability in key photosynthetic traits of a globally important agricultural crop while revealing a hierarchical structure of trait variation. Because photosynthesis represents one of the major processes governing plant growth and development, the genotype level screening described here using competing mechanistic models can inform our understanding of the links between observed variances and genetic controls. Bayesian methodology, emerging as a powerful tool in plant sciences, permits the explicit incorporation of prior information, propagation of uncertainty from measurements to models and offers a way to improve phenotyping methods while incorporating new data and theory.

AUTHOR CONTRIBUTIONS

TA, BE, and CW: planned and designed experiments; TA, JP, and BE: conducted fieldwork; JP and DM: developed models and analyzed data; JP, DM, TA, BE, and CW: wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank Carmela Guadagno, Daniel Potts, and Xiaonan Tai for comments on an earlier version of the manuscript. This research was supported by National Science Foundation (Nos. 0654305, 1102998, 1025965, 1444571, 1547796). The views expressed here are those of the authors and do not necessarily represent the views of the NSF. We are especially grateful for the insightful comments made by the reviewers and the Associate Editor.

REFERENCES

Akaike, H. (1998). "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike,* eds E. Parzen, K. Tanabe, G. Kitagawa (New York, NY: Springer), 199–213.

Araujo, W. L., Nunes-Nesi, A., Nikoloski, Z., Sweetlove, L. J., and Fernie, A. R. (2012). Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. *Plant Cell Environ.* 35, 1–21. doi: 10.1111/j.1365-3040.2011.02332.x

Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis *in vivo*. *Annu. Rev. Plant Biol.* 59, 89–113. doi: 10.1146/annurev.arplant.59.032607.092759

Bernacchi, C. J., Portis, A. R., Nakano, H., von Caemmerer, S., and Long, S. P. (2002). Temperature response of mesophyll conductance. Implications for the determination of Rubisco enzyme kinetics and for limitations to photosynthesis *in vivo*. *Plant Physiol.* 130, 1992–1998. doi: 10.1104/pp.008250

Bernacchi, C., Singsaas, E., Pimentel, C., Portis, A. Jr., and Long, S. (2001). Improved temperature response functions for models of Rubisco limited photosynthesis. *Plant Cell Environ.* 24, 253–259. doi: 10.1111/j.1365-3040.2001.00668.x

Berry, J., and Bjorkman, O. (1980). Photosynthetic response and adaptation to temperature in higher plants. *Ann. Rev. Plant Physiol.* 31, 491–543. doi: 10.1146/annurev.pp.31.060180.002423

Bertin, N., Martre, P., Génard, M., Quilot, B., and Salon, C. (2009). Under what circumstances can process-based simulation models link genotype to phenotype for complex traits? Case-study of fruit and grain quality traits. *J. Exp. Bot.* 61, 955–967. doi: 10.1093/jxb/erp377

Box, G. (2001). Statistics for discovery. *J. Appl. Stat.* 28, 285–299. doi: 10.1080/02664760120034036

Box, G. E. P. (1976). Science and statistics. *J. Am. Stat. Assoc.* 71, 791–799. doi: 10.1080/01621459.1976.10480949

Brooks, S. P., and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7, 434–455.

Chaumont, F., and Tyerman, S. D. (2014). Aquaporins: highly regulated channels controlling plant water relations. *Plant Physiol.* 164, 1600–1618. doi: 10.1104/pp.113.233791

Damour, G., Simonneau, T., Cochard, H., and Urban, L. (2010). An overview of models of stomatal conductance at the leaf level. *Plant Cell Environ.* 33, 1419–1438. doi: 10.1111/j.1365-3040.2010.02181.x

Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., et al. (2015). Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in Zea mays. *Genome Biol.* 16:167. doi: 10.1186/s13059-015-0716-z

DeWitt, C. (1965). *Photosynthesis of Leaf Canopies.* Mededeling 274, Instituut voor Biologisch en Scheikundig Onderzoek van Landbouwgewassen, Wageningen. Neth.

Edwards, C. E., Ewers, B. E., McClung, C. R., Lou, P., and Weinig, C. (2012). Quantitative variation in water-use efficiency across water regimes and its relationship with circadian, vegetative, reproductive, and leaf gas-exchange traits. *Mol. Plant* 5, 653–668. doi: 10.1093/mp/sss004

Edwards, C. E., Ewers, B. E., and Weinig, C. (2016). Genotypic variation in biomass allocation in response to field drought has a greater affect on yield than gas exchange or phenology. *BMC Plant Biol.* 16:185. doi: 10.1186/s12870-016-0876-3

Edwards, C. E., Ewers, B. E., Williams, D. G., Xie, Q., Lou, P., Xu, X., et al. (2011). The genetic architecture of ecophysiological and circadian traits in Brassica rapa. *Genetics* 189, 375–390. doi: 10.1534/genetics.110.125112

Ethier, G. J., and Livingston, N. J. (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. doi: 10.1111/j.1365-3040.2004.01140.x

Evans, J. R., Jakobsen, I., and Ögren, E. (1993). Photosynthetic light-response curves. *Planta* 189, 191–200. doi: 10.1007/BF00195076

Farquhar, G., von Caemmerer, S., and Berry, J. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta* 149, 78–90. doi: 10.1007/BF00386231

Farquhar, G. D., von Caemmerer, S., and Berry, J. A. (2001). Models of Photosynthesis. *Plant Physiol.* 125, 42–45. doi: 10.1104/pp.125.1.42

Flexas, J., Diaz-Espejo, A., Galmes, J., Kaldenhoff, R., Medrano, H., and Ribas-Carbo, M. (2007). Rapid variations of mesophyll conductance in response to changes in CO_2 concentration around leaves. *Plant Cell Environ.* 30, 1284–1298. doi: 10.1111/j.1365-3040.2007.01700.x

Flexas, J., Ribas-Carbo, M., Diaz-Espejo, A., Galmés Almés, J, and Medrano, H. (2008). Mesophyll conductance to CO_2: current knowledge and future prospects. *Plant Cell Environ.* 31, 602–621. doi: 10.1111/j.1365-3040.2007.01757.x

Flexas, J., Ribas-Carbo, M., Hanson, D. T., Bota, J., Otto, B., Cifre, J., et al. (2006). Tobacco aquaporin NtAQP1 is involved in mesophyll conductance to CO_2 *in vivo*. *Plant J.* 48, 427–439. doi: 10.1111/j.1365-313X.2006.02879.x

Foyer, C. H., Neukermans, J., Queval, G., Noctor, G., and Harbinson, J. (2012). Photosynthetic control of electron transport and the regulation of gene expression. *J. Exp. Bot.* 63, 1637–1661. doi: 10.1093/jxb/ers013

Foyer, C. H., and Shigeoka, S. (2011). Understanding Oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol.* 155, 93–100. doi: 10.1104/pp.110.166181

Furbank, R. T., Quick, W. P., and Sirault, X. R. (2015). Improving photosynthesis and yield potential in cereal crops by targeted genetic manipulation: Prospects, progress and challenges. *Field Crops Res.* 182, 19–29. doi: 10.1016/j.fcr.2015.04.009

Galmes, J., Flexas, J., Keys, A. J., Cifre, J., Mitchell, R. A. C., Madgwick, P. J., et al. (2005). Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves. *Plant Cell Environ.* 28, 571–579. doi: 10.1111/j.1365-3040.2005.01300.x

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1, 515–534. doi: 10.1214/06-BA117A

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., (2004). *Bayesian Data Analysis. Texts in Statistical Science Series*. Boca Raton, FL: Chapman & Hall/CRC.

Gelman, A., Hwang, J., and Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.* 24, 997–1016. doi: 10.1007/s11222-013-9416-2

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist.Sci.* 7, 457–472. doi: 10.1214/ss/1177011136

Genty, B., Briantais, J.-M., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochim. Biophys. Acta* 990, 87–92. doi: 10.1016/S0304-4165(89)80016-9

Gou, F., van Ittersum, M. K., and van der Werf, W. (2017). Simulating potential growth in a relay-strip intercropping system: model description, calibration and testing. *Field Crops Res.* 200, 122–142. doi: 10.1016/j.fcr.2016.09.015

Grassi, G., and Magnani, F. (2005). Stomatal, mesophyll conductance and biochemical limitations to photosynthesis as affected by drought and leaf ontogeny in ash and oak trees. *Plant Cell Environ.* 28, 834–849. doi: 10.1111/j.1365-3040.2005.01333.x

Gu, J., Yin, X., Stomph, T. J., and Struik, P. C. (2014). Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. *Plant Cell Environ.* 37, 22–34. doi: 10.1111/pce.12173

Gu, J., Yin, X., Stomph, T.-J., Wang, H., and Struik, P. C. (2012). Physiological basis of genetic variation in leaf photosynthesis among rice (*Oryza sativa* L.) introgression lines under drought and well-watered conditions. *J. Exp. Bot.* 63, 5137–5153. doi: 10.1093/jxb/ers170

Gu, L., and Sun, Y. (2014). Artefactual responses of mesophyll conductance to CO_2 and irradiance estimated with the variable J and online isotope discrimination methods. *Plant Cell Environ.* 37, 1231–1249. doi: 10.1111/pce.12232

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., et al. (2006). Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593. doi: 10.1016/j.tplants.2006.10.006

Hanba, Y. T., Shibasaka, M., Hayashi, Y., Hayakawa, T., Kasamo, K., Terashima, I., et al. (2004). Overexpression of the barley aquaporin HvPIP2; 1 increases internal CO2 conductance and CO_2 assimilation in the leaves of transgenic rice plants. *Plant Cell Physiol.* 45, 521–529. doi: 10.1093/pcp/pch070

Hanson, D. T., Stutz, S. S., and Boyer, J. S. (2016). Why small fluxes matter: the case and approaches for improving measurements of photosynthesis and (photo) respiration. *J. Exp. Bot.* 67, 3027–3039. doi: 10.1093/jxb/erw139

Hauser, T., Popilka, L., Hartl, F. U., and Hayer-Hartl, M. (2015). Role of auxiliary proteins in Rubisco biogenesis and function. *Nat. Plants* 1:15065. doi: 10.1038/nplants.2015.65

Häusler, R. E., Kleines, M., Uhrig, H., Hirsch, H. J., and Smets, H. (1999). Overexpression of phospho enol pyruvate carboxylase from *Corynebacterium glutamicum* lowers the CO_2 compensation point (Γ^*) and enhances dark and light respiration in transgenic potato. *J. Exp. Bot.* 50, 1231–1242.

He, C., Shen, G., Pasapula, V., Luo, J., Venkataramani, S., Qiu, X., et al. (2007). Ectopic expression of AtNHX1 in cotton (*Gossypium hirsutum* L.) increases proline content and enhances photosynthesis under salt stress conditions. *J. Cotton Sci.* 11, 266–274.

Kattge, J., and Knorr, W. (2007). Temperature acclimation in a biochemical model of photosynthesis: a reanalysis of data from 36 species. *Plant Cell Environ.* 30, 1176–1190. doi: 10.1111/j.1365-3040.2007.01690.x

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C. (2009). Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob. Chang. Biol.* 15, 976–991. doi: 10.1111/j.1365-2486.2008.01744.x

Knorr, W., and Heimann, M. (2001). Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new

photosynthesis and energy balance scheme. *Glob. Biogeochem. Cyc.* 15, 207–225. doi: 10.1029/1998GB001059

Kruschke, J. (2010). *Doing Bayesian data analysis: A tutorial introduction with R*. New York, NY: Academic Press.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *J. Exp. Psychol.* 142:573. doi: 10.1037/a0029146

Laisk, A., Eichelmann, H., Oja, V., Talts, E., and Scheibe, R. (2007). Rates and roles of cyclic and alternative electron flow in potato leaves. *Plant Cell Physiol.* 48, 1575–1588. doi: 10.1093/pcp/pcm129

Laisk, A., and Loreto, F. (1996). Determining photosynthetic parameters from leaf CO_2 exchange and chlorophyll fluorescence (ribulose-1, 5-bisphosphate carboxylase/oxygenase specificity factor, dark respiration in the light, excitation distribution between photosystems, alternative electron transport rate, and mesophyll diffusion resistance. *Plant Physiol.* 110, 903–912. doi: 10.1104/pp.110.3.903

Laisk, A., and Nedbal, L. (2009). *Photosynthesis in silico: Understanding Complexity From Molecules To Ecosystems*. Dordrecht: Springer Science & Business Media.

Laisk, A., Oja, V., Rasulov, B., Rämma, H., Eichelmann, H., Kasparova, I., et al. (2002). A computer-operated routine of gas exchange and optical measurements to diagnose photosynthetic apparatus in leaves. *Plant Cell Environ.* 25, 923–943. doi: 10.1046/j.1365-3040.2002.00873.x

Lambers, H., Chapin, F.S., and Pons, T. L. (2008). *Plant Physiological Ecology*. New York, NY: Springer Science.

Leuning, R. (2002). Temperature dependence of two parameters in a photosynthesis model. *Plant Cell Environ.* 25, 1205–1210. doi: 10.1046/j.1365-3040.2002.00898.x

Livingston, A. K., Cruz, J. A., Kohzuma, K., Dhingra, A., and Kramer, D. M. (2010). An arabidopsis mutant with high cyclic electron flow around photosystem I (hcef) involving the NADPH dehydrogenase complex. *Plant Cell* 22, 221–233. doi: 10.1105/tpc.109.071084

Long, S. P., Zhu, X. G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant Cell Environ.* 29, 315–330. doi: 10.1111/j.1365-3040.2005.01493.x

Mackay, D. S., Ewers, B. E., Loranty, M. M., Kruger, E. L., and Samanta, S. (2012). Bayesian analysis of canopy transpiration models: a test of posterior parameter means against measurements. *J. Hydrol.* 432–433, 75–83. doi: 10.1016/j.jhydrol.2012.02.019

Martre, P., Quilot-Turion, B., Luquet, D., Ould-Sidi Memmah, M.-M., Chenu, K., and Debaeke, P. (2015). "Model-assisted phenotyping and ideotype design," in *Crop Physiology*, ed V. O. Sadras (San Diego, CA: Elsevier), 349–373.

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., and Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Ann. Bot.* 105, 1141–1157. doi: 10.1093/aob/mcq028

Maxwell, K., and Johnson, G. N. (2000). Chlorophyll fluorescence—a practical guide. *J. Exp. Bot.* 51, 659–668. doi: 10.1093/jexbot/51.345.659

McDowell, N. G., Fisher, R. A., Xu, C., Domec, J. C., Hölttä, T., Mackay, D. S., et al. (2013). Evaluating theories of drought-induced vegetation mortality using a multimodel–experiment framework. *New Phytol.* 200, 304–321. doi: 10.1111/nph.12465

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*. New York, NY: CRC Press.

Medlyn, B. E., Dreyer, E., Ellsworth, D., Forstreuter, M., Harley, P., Kirschbaum, M., et al. (2002a). Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. *Plant Cell Environ.* 25, 1167–1179. doi: 10.1046/j.1365-3040.2002.00891.x

Medlyn, B. E., Loustau, D., and Delzon, S. (2002b). Temperature response of parameters of a biochemically based model of photosynthesis. I. Seasonal changes in mature maritime pine (*Pinus pinaster Ait.*). *Plant Cell Environ.* 25, 1155–1165. doi: 10.1046/j.1365-3040.2002.00890.x

Miyake, C. (2010). Alternative electron flows (water–water cycle and cyclic electron flow around PSI) in Photosynthesis: molecular mechanisms and physiological functions. *Plant Cell Physiol.* 51, 1951–1963. doi: 10.1093/pcp/pcq173

Niinemets, Ü., Díaz-Espejo, A., Flexas, J., Galmés, J., and Warren, C. R. (2009a). Importance of mesophyll diffusion conductance in estimation of plant photosynthesis in the field. *J. Exp. Bot.* 60, 2271–2282. doi: 10.1093/jxb/erp063

Niinemets, Ü., Díaz-Espejo, A., Flexas, J., Galmés, J., and Warren, C. R. (2009b). Role of mesophyll diffusion conductance in constraining potential photosynthetic productivity in the field. *J. Exp. Bot.* 60, 2249–2270. doi: 10.1093/jxb/erp036

Ogle, K., and Barber, J. (2008). "Bayesian data—model integration in plant physiological and ecosystem ecology," in *Progress in Botany*, eds U. Lüttge, W. Beyschlag, and J. Murata (Berlin; Heidelberg: Springer), 281–311.

Ogle, K., and Barber, J. (2011). "Bayesian statistics," in *Encyclopedia of Theoretical Ecology*, eds A. Hastings and L. Gross (Berkeley, CA: University of California Press), 307–316.

Patrick, L. D., Ogle, K., and Tissue, D. T. (2009). A hierarchical Bayesian approach for estimation of photosynthetic parameters of C3 plants. *Plant Cell Environ.* 32, 1695–1709. doi: 10.1111/j.1365-3040.2009.02029.x

Plummer, M. (2008). Penalized loss functions for Bayesian model comparison. *Biostatistics* 9, 523–539. doi: 10.1093/biostatistics/kxm049

Plummer, M (2014). *rjags: Bayesian Graphical Models Using MCMC.* Available online at: http://mcmc-jags.sourceforge.net/

Pons, T. L., Flexas, J., Von Caemmerer, S., Evans, J. R., Genty, B., Ribas-Carbo, M., et al. (2009). Estimating mesophyll conductance to CO_2: methodology, potential errors, and recommendations. *J. Exp. Bot.* 60, 2217–2234. doi: 10.1093/jxb/erp,081

R Development Core Team (2014). *R: a Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing, 2012. Available online at: http://cran.r-project.org

Rieseberg, L. H., Archer, M. A., and Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity* 83, 363–372. doi: 10.1038/sj.hdy.6886170

Rieseberg, L. H., Widmer, A., Arntz, A. M., and Burke, B. (2003). The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 1141–1147. doi: 10.1098/rstb.2003.1283

Ruane, A. C., Rosenzweig, C., Asseng, S., Boote, K. J., Elliott, J., Ewert, F., et al. (2017). An AgMIP framework for improved agricultural representation in integrated assessment models. *Environ. Res. Lett.* 12:125003. doi: 10.1088/1748-9326/aa8da6

Sadras, V., Rebetzke, G., and Edmeades, G. (2013). The phenotype and the components of phenotypic variance of crop traits. *Field Crops Res.* 154, 255–259. doi: 10.1016/j.fcr.2013.10.001

Sharkey, T. D., Bernacchi, C. J., Farquhar, G. D., and Singsaas, E. L. (2007). Fitting photosynthetic carbon dioxide response curves for C3 leaves. *Plant Cell Environ.* 30, 1035–1040. doi: 10.1111/j.1365-3040.2007.01710.x

Sharkey, T. D., Berry, J. A., and Raschke, K. (1985). Starch and sucrose synthesis in Phaseolus vulgaris as affected by light, CO_2, and abscisic acid. *Plant Physiol.* 77, 617–620. doi: 10.1104/pp.77.3.617

Sinclair, T. R., and Seligman, N. A. G (1996). Crop modeling: from infancy to maturity. *Agron. J.* 88, 698–704. doi: 10.2134/agronj1996.00021962008800050004x

Singh, J., Pandey, P., James, D., Chandrasekhar, K., Achary, V. M. M., Kaul, T., et al. (2014). Enhancing C3 photosynthesis: an outlook on feasible interventions for crop improvement. *Plant Biotechnol. J.* 12, 1217–1230. doi: 10.1111/pbi.12246

Song, Y., Chen, Q., Ci, D., Shao, X., and Zhang, D. (2014). Effects of high temperature on photosynthesis and related gene expression in poplar. *BMC Plant Biol.* 14:111. doi: 10.1186/1471-2229-14-111

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 64, 583–639. doi: 10.1111/1467-9868.00353

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003). *WinBUGS User Manual: Version.* Cambridge, UK.

Studer, R. A., Christin, P.-A., Williams, M. A., and Orengo, C. A. (2014). Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2223–2228. doi: 10.1073/pnas.1310811111

Tardieu, F. (2010). Why work and discuss the basic principles of plant modelling 50 years after the first plant models? *J. Exp. Bot.* 61, 2039–2041. doi: 10.1093/jxb/erq135

Tazoe, Y., Von Caemmerer, S., Estavillo, G. M., and Evans, J. R. (2011). Using tunable diode laser spectroscopy to measure carbon isotope discrimination and mesophyll conductance to CO_2 diffusion dynamically

at different CO_2 concentrations. *Plant Cell Environ.* 34, 580–591. doi: 10.1111/j.1365-3040.2010.02264.x

Tebaldi, C., Mearns, L. O., Nychka, D., and Smith, R. L. (2004). Regional probabilities of precipitation change: a Bayesian analysis of multimodel simulations. *Geophys. Res. Lett.* 31:24. doi: 10.1029/2004GL021276

Théroux-Rancourt, G., and Gilbert, M. E. (2017). The light response of mesophyll conductance is controlled by structure across leaf profiles. *Plant Cell Environ.* 40, 726–740. doi: 10.1111/pce.12890

Thornton, P. E., Running, S. W., and Hunt, E. (2005). *Biome-BGC: Terrestrial Ecosystem Process Model, Version 4.1. 1. Model Product.* Oak Ridge, TN: Oak Ridge National Laboratory Distributed Active Archive Center. Available online at: https://daac.ornl.gov/

Tomás, M., Medrano, H., Brugnoli, E., Escalona, J., Martorell, S., Pou, A., et al. (2014). Variability of mesophyll conductance in grapevine cultivars under water stress conditions in relation to leaf anatomy and water use efficiency. *Aust. J. Grape Wine Res.* 20, 272–280. doi: 10.1111/ajgw.12069

van der Tol, C., Verhoef, W., and Rosema, A. (2009). A model for chlorophyll fluorescence and photosynthesis at leaf scale. *Agric. For. Meteorol.* 149, 96–105. doi: 10.1016/j.agrformet.2008.07.007

von Caemmerer, S., Evans, J. R., Hudson, G. S., and Andrews, T. J. (1994). The kinetics of ribulose-1, 5-bisphosphate carboxylase/oxygenase *in vivo* inferred from measurements of photosynthesis in leaves of transgenic tobacco. *Planta* 195, 88–97. doi: 10.1007/BF00206296

von Caemmerer, S. S. (2000). *Biochemical Models of Leaf Photosynthesis*. Collingwood, VIC: CSIRO.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* 11, 3571–3594.

Wullschleger, S. D. (1993). Biochemical limitations to carbon assimilation in C3 plants—a retrospective analysis of the A/Ci curves from 109 species. *J. Exp. Bot.* 44, 907–920. doi: 10.1093/jxb/44.5.907

Xue, L., Zhang, D., Guadagnini, A., and Neuman, S. P. (2014). Multimodel Bayesian analysis of groundwater data worth. *Water Resour. Res.* 50, 8481–8496. doi: 10.1002/2014WR015503

Yamori, W., Hikosaka, K., and Way, D. A. (2014). Temperature response of photosynthesis in C3, C4, and CAM plants: temperature acclimation and temperature adaptation. *Photosyn. Res.* 119, 101–117. doi: 10.1007/s11120-013-9874-6

Yamori, W., Masumoto, C., Fukayama, H., and Makino, A. (2012). Rubisco activase is a key regulator of non-steady-state photosynthesis at any leaf temperature and, to a lesser extent, of steady-state photosynthesis at high temperature. *Plant J.* 71, 871–880. doi: 10.1111/j.1365-313X.2012.05041.x

Yarkhunova, Y., Edwards, C. E., Ewers, B. E., Baker, R. L., Aston, T. L., McClung, C. R., et al. (2016). Selection during crop diversification involves correlated evolution of the circadian clock and ecophysiological traits in *Brassica rapa*. *New Phytol.* 210, 133–144. doi: 10.1111/nph.13758

Yin, X. Y., Kropff, M. J., Goudriaan, J., and Stam, P. (2001). A model analysis of yield differences among recombinant inbred lines in barley. *Agron. J.* 92, 114–120. doi: 10.2134/agronj2000.921114x

Yin, X., Struik, P. C., and Kropff, M. J. (2004). Role of crop physiology in predicting gene-to-phenotype relationships. *Trends Plant Sci.* 9, 426–432. doi: 10.1016/j.tplants.2004.07.007

Yin, X., Struik, P. C., Romero, P., Harbinson, J., Evers, J. B., Van Der Putten, P. E., et al. (2009). Using combined measurements of gas exchange and chlorophyll fluorescence to estimate parameters of a biochemical C3 photosynthesis model: a critical appraisal and a new integrated approach applied to leaves in a wheat (*Triticum aestivum*) canopy. *Plant Cell Environ.* 32, 448–464. doi: 10.1111/j.1365-3040.2009.01934.x

Zhu, G.-F., Li, X., Su, Y.-H., Lu, L., and Huang, C.-L. (2011). Seasonal fluctuations and temperature dependence in photosynthetic parameters and stomatal conductance at the leaf scale of Populus euphratica Oliv. *Tree Physiol.* 31, 178–195. doi: 10.1093/treephys/tpr005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.