




RESEARCH PAPER

Linking photosynthesis and yield reveals a strategy to improve light use efficiency in a climbing bean breeding population

Beat Keller^{1,2} , Jonatan Soto³, Angelina Steier⁴, Ana Elisabeth Portilla-Benavides³, Bodo Raatz^{3,†}, Bruno Studer², Achim Walter¹, Onno Muller⁴ , and Milan O. Urban^{3,*} 

¹ Crop Science, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland

² Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland

³ Bean Program, Crops for nutrition and health, International Center for Tropical Agriculture (CIAT), Cali, Colombia

⁴ Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, Jülich, Germany

[†] Present address: Limagrain, la Ménitère, France.

* Correspondence: m.urban@cgiar.org or dehydrins@gmail.com

Received 21 June 2023; Editorial decision 9 October 2023; Accepted 21 October 2023

Editor: Johannes Kromdijk, University of Cambridge, UK

Abstract

Photosynthesis drives plant physiology, biomass accumulation, and yield. Photosynthetic efficiency, specifically the operating efficiency of PSII (F_q'/F_m'), is highly responsive to actual growth conditions, especially to fluctuating photosynthetic photon fluence rate (PPFR). Under field conditions, plants constantly balance energy uptake to optimize growth. The dynamic regulation complicates the quantification of cumulative photochemical energy uptake based on the intercepted solar energy, its transduction into biomass, and the identification of efficient breeding lines. Here, we show significant effects on biomass related to genetic variation in photosynthetic efficiency of 178 climbing bean (*Phaseolus vulgaris* L.) lines. Under fluctuating conditions, the F_q'/F_m' was monitored throughout the growing period using hand-held and automated chlorophyll fluorescence phenotyping. The seasonal response of F_q'/F_m' to PPFR ($\text{Response}_{G;PPFR}$) achieved significant correlations with biomass and yield, ranging from 0.33 to 0.35 and from 0.22 to 0.31 in two glasshouse and three field trials, respectively. Phenomic yield prediction outperformed genomic predictions for new environments in four trials under different growing conditions. Investigating genetic control over photosynthesis, one single nucleotide polymorphism (Chr09_37766289_13052) on chromosome 9 was significantly associated with $\text{Response}_{G;PPFR}$ in proximity to a candidate gene controlling chloroplast thylakoid formation. In conclusion, photosynthetic screening facilitates and accelerates selection for high yield potential.

Keywords: Breeding, bean, light use efficiency, photosynthesis, phenotyping, selection.

Abbreviations: A_{max} , maximum photosynthetic rate; $\text{Response}_{G;PPFR}$, response of F_q'/F_m' to PPFR; ϵ_{ei} , energy conversion efficiency of intercepted PPFR to photochemical energy; g_s , stomatal conductance; ChlF, chlorophyll fluorescence; ET, electron transport; F_q'/F_m' , operating efficiency of PSII; FRRF, fast repetition rate flash; G×E, genotype by environment interaction; GEBV, genomic estimated breeding values; G×Ec, interaction between genotype and the covariate; GWAS, genome-wide association study; LIFT, light-induced fluorescence transient; LUE, light use efficiency; MTCI, MERIS terrestrial chlorophyll index; NDVI, normalized difference vegetation index; NPQ, non-photochemical quenching; PPFR, photosynthetic photon fluence rate; PRI, photochemical reflectance index; SdFe, seed iron concentration.

Introduction

Photosynthesis is highly responsive to fluctuating environmental conditions, constantly adjusting solar energy uptake to optimize growth (Demmig-Adams *et al.*, 2012). Therefore, photosynthetic performance reflects growth conditions and the plant's ability to cope with these conditions (Murchie *et al.*, 2018). The highly dynamic response of photosynthesis to changing environmental conditions challenges the true estimation of photosynthetic performance in temporal and spatial dimensions (i.e. diurnally up to the growth period and within the individual plant canopy up to heterogeneities in the field, respectively) (Kaiser *et al.*, 2018). Nevertheless, photosynthetic parameters have successfully predicted biomass and yield in crops and breeding lines (Fischer *et al.*, 1998; Lopez *et al.*, 2019; Keller *et al.*, 2022b). Unfortunately, these parameters have rarely been exploited. Is photosynthetic performance a limiting factor for yield? Can photosynthetic traits be used as a breeding target to increase yield (Gifford and Evans, 1981; Sinclair *et al.*, 2019)?

Photosynthesis for yield improvement

Photosynthetic efficiency continues to be a focus for crop yield improvement research because conventional traits, such as harvest index, are already highly exploited (Reynolds *et al.*, 2012; Long *et al.*, 2015). Although photosynthetic performance has often improved in high-yielding lines, remaining genetic variation suggests further potential for improvement (Driever *et al.*, 2014; Theeuwes *et al.*, 2022). When targeting photosynthetic traits for crop improvement, there are two main strategies: (i) optimizing the photosynthetic pathway through genetic engineering, or (ii) selecting breeding lines with higher photosynthetic efficiency under apparent growing conditions (Parry *et al.*, 2011). The former strategy, aimed at improving the photosynthetic processes, has been successful in field trials in tobacco, rice, and soybean (Kromdijk *et al.*, 2016; South *et al.*, 2019; López-Calcano *et al.*, 2020; De Souza *et al.*, 2022). The latter strategy, which aims to increase photosynthetic efficiency by selecting adapted lines to growing conditions, produced some striking results. For example, Fischer *et al.* (1998) reported a linear yield increase comparing eight wheat (*Triticum aestivum* L.) varieties released between 1962 and 1988. The yield increase (+27%) was associated with 23% and 63% increases in the maximum photosynthetic rate (A_{\max}) and stomatal conductance (g_s), respectively. Another study reported a correlation of $r=0.57$ between steady-state spikelet photosynthesis and biomass using 12 contrasting wheat breeding lines (Molero and Reynolds, 2020). In soybean, photosynthesis and grain yield were strongly correlated ($r=0.8$) in 383 lines (Lopez *et al.*, 2019). Strong correlations between biomass and photosynthesis measured by gas exchange have also been shown in earlier studies in wheat and soybean (Ashley and Boerma, 1989; Gutiérrez-Rodríguez *et al.*, 2000; Carmo-Silva *et al.*, 2017). In

many other studies reviewed by Zelitch (1982), a relationship between photosynthesis and yield was not observed because the photosynthetic performance was quantified over a short period that was not representative of the entire fluctuating growing season.

Photosynthesis under fluctuating conditions

The energy transduction from incident solar radiation to biomass involves the interception of photosynthetically active radiation by the leaf canopy, conversion of radiant energy to photochemical energy, and transduction into biomass (Zhu *et al.*, 2010). The light energy intercepted by plant leaves is transmitted, reflected, and absorbed depending on the pigment composition (Porcar-Castell *et al.*, 2014). The absorbed solar energy is then dynamically partitioned between three different pathways (Butler, 1978). Photosynthesis and dissipation of excess energy as heat through non-photochemical quenching (NPQ) are two pathways that are physiologically regulated to optimize plant growth and development (Butler, 1978; Bilger and Björkman, 1990). Approximately 2–4% of the absorbed energy is re-emitted as chlorophyll fluorescence (ChlF) via the third pathway (Porcar-Castell *et al.*, 2014). Significant losses occur through NPQ (up to 50% over the growing season), preventing damage and determining whole-canopy photosynthesis (Ishida *et al.*, 2014; Murchie and Ruban, 2020). The changing light intensity during the day together with the NPQ response causes the typical diurnal pattern of photosynthesis (Pieruschka *et al.*, 2010; Ishida *et al.*, 2014; Keller *et al.*, 2019a). This highly dynamic response of photosynthesis requires long-term monitoring to determine photosynthetic efficiency throughout the season. Photosynthesis measurements via CO_2 assimilation over longer growing periods are time consuming—even when the photochemical uptake is approximated by variables such as A_{\max} , which are usually derived under high light intensities around midday. In addition, the photosynthetic capacity may vary in different leaf types, as shown for wheat (Salter *et al.*, 2020). This may explain why such screenings have not been further implemented in breeding programs when phenotypic information from thousands of lines is required. In contrast to gas exchange measurements, the operating efficiency of PSII (F_q'/F_m') can be obtained in seconds and is widely used in plant physiology as a proxy for growth performance (Kalaji *et al.*, 2016; Murchie *et al.*, 2018). It is determined by the modulated ChlF signal, which describes the proportion of energy quantum used for electron transport (ET) from the absorbed photosynthetic photon fluence rate (PPFR) (Schreiber *et al.*, 1986; Baker, 2008). The ET at PSII is linearly related to CO_2 assimilation in the absence of photorespiration or cyclic ET (Genty *et al.*, 1989; Yamori *et al.*, 2012). Photorespiration and cyclic

ET are increased, for example, under heat and drought stress conditions (Golding *et al.*, 2004; Zhang and Sharkey, 2009; Hermida-Carrera *et al.*, 2016; Yamori and Shikanai, 2016). This method is suitable for rapid, non-invasive, long-term observation of linear ET, approximating the energy conversion efficiency of intercepted PPFR to photochemical energy (ϵ_e), as well as the competing NPQ (Demmig-Adams *et al.*, 2012; Keller *et al.*, 2022b).

Early genomic and phenomic selection

Genetic gain depends on high selection intensity on a heritable trait in short breeding cycles. The selection of a trait can be aided by genomic or phenomic estimates. Genomic predictions allow calculation of a genomic estimated breeding value (GEBV) from seed DNA samples, enabling selection before planting (Meuwissen *et al.*, 2001). However, many breeding programs require cheap and rapid selection methods at early stages of a selection process because genotyping of thousands of breeding lines is expensive (Furbank *et al.*, 2019). Phenomic selection uses high-throughput measurements, such as the spectral or ChlF signal, to predict a target trait. This also favors adaptation to new and future environments, accounting for genotype by environment interaction ($G \times E$). The $G \times E$ can then be described by the interaction between genotype and the covariate ($G \times E_c$), which is measured for each line in each environment (Eeuwijk *et al.*, 2016). Therefore, if measurements between lines are sufficiently correlated with the target trait, and are cheap and precise, phenomic selection is an efficient alternative to genomic selection.

Recently, ChlF measurements were shown to estimate biomass in 12 maize (*Zea mays* L.) and nine soybean [*Glycine max* (L.) Merr.] lines under field and naturally fluctuating conditions using automated high-throughput phenotyping (Keller *et al.*, 2022b). This approach could be modified to screen large numbers of early breeding lines for genetic variability in photosynthetic performance. This is especially useful in crops with longer growth cycles, such as the climber types of common bean (*Phaseolus vulgaris* L.). Photosynthetic performance is especially interesting in legumes, because they counteract protein dilution when assimilating more CO_2 by increasing biological nitrogen fixation for protein synthesis (Evans and Clarke, 2019).

To study the relationship between photosynthetic traits and biomass production in climbing bean breeding lines, we investigated four hypotheses: (i) photosynthetic performance over the growing season is correlated with biomass production; (ii) selection based on the ϵ_e can increase the selection effectiveness of higher grain yield lines; (iii) photosynthesis-related traits (determined under the specific growing conditions) are more accurate than genomic prediction to predict yield of new lines in new environments; and (iv) genes contributing in the control of photosynthesis can be identified using genome-wide association study (GWAS).

Materials and methods

Plant material

This study used 178 breeding lines out of the complete climbing bean panel (VEC, total 290 lines) genotyped by Keller *et al.* (2022a). The panel represents the genetic variation of the climbing bean breeding program at the International Center for Tropical Agriculture (CIAT). Most of the lines belong to the Andean gene pool. A few lines are admixtures between the Andean and Mesoamerican gene pools.

Growth conditions

Glasshouse

Plants were grown in 2 liter black pots (square opening 12×2 cm) in an unheated glasshouse located at the agricultural research station Campus Klein-Altendorf of the University of Bonn, Germany ($50^\circ 36' 50.7''\text{N}$, $6^\circ 59' 38.9''\text{E}$, altitude of 185 m a.s.l.). The trials, CKA20D and CKA20E, were sown on the first of September and October in 2020, respectively. The pots were filled with ~ 2 liters of turf-clay substrate (ED73, Einheitserdewerke, Sinntal-Altengronau, Germany). The plants were grown in six rows with 1.5 m row spacing and 22 pots per row with 14 cm spacing (Supplementary Fig. S1). The first and last pot in a row were excluded as a border. No fertilizers or plant protection agents were used. Four seeds per pot were seeded and thinned after ~ 2 weeks in order to grow two plants per pot. Bamboo stakes (~ 1 m) were inserted into each pot as support for the climbing beans. Drip irrigation was used twice a week to ensure well-watered growth conditions. A total of 136 lines were evaluated in two glasshouse trials (CKA20D, $n=86$ and CKA20E, $n=85$ lines), in a partially replicated design with four check lines in six replicates (Fig. 1A, B). The PPFR and temperature were measured every minute by a LI-COR sensor (LI-COR, Inc., Lincoln, NE, USA) and a HMP110 temperature sensor (Vaisala, Helsinki, Finland), respectively, at ~ 0.5 m above the pots. No artificial heating or lighting was used in the glasshouse.

Field

Plants were grown in the field in Darién, Colombia ($3^\circ 53' 31''\text{N}$, $76^\circ 31' 00''\text{W}$, altitude of 1491 m a.s.l.) and in Palmira ($3^\circ 30' 03.0''\text{N}$, $76^\circ 21' 03.5''\text{W}$, altitude of 965 m a.s.l.) as described by Keller *et al.* (2022a). Briefly, the trials in Darién, Dar18B ($n=120$ lines) and Dar19B ($n=126$ lines), were sown in the second season of the year in 2018 and 2019, respectively. The trial in Palmira (Pal19D, $n=117$ lines) was sown in the fourth season of the year in 2019. The soil types were an Inceptisol and Mollisol in Darién and in Palmira, respectively. The plants were treated and irrigated according to the standard CIAT protocol and empirical knowledge of the field. The PPFR values were recorded by the MultispeQ during every measurement. A total of 127 lines were evaluated in the field in 1–3 replicates (Fig. 1C, D).

Phenotyping

The light-induced fluorescence transient (LIFT) and the MultispeQ devices measured photosynthetic performance under incident sunlight in the glasshouse and field, respectively. Both devices measure the F_q'/F_m' by probing the ChlF signal. Additionally, the LIFT device measures the full reflectance spectrum on the measured target while the MultispeQ is limited to only several spectral bands. Finally, biomass and grain yield were measured by destructive phenotyping at harvest.

Light-induced fluorescence transient

The ChlF and spectra were measured using the LIFT-REM device (Soliense Inc., New York, NY, USA), mounted on an automated platform

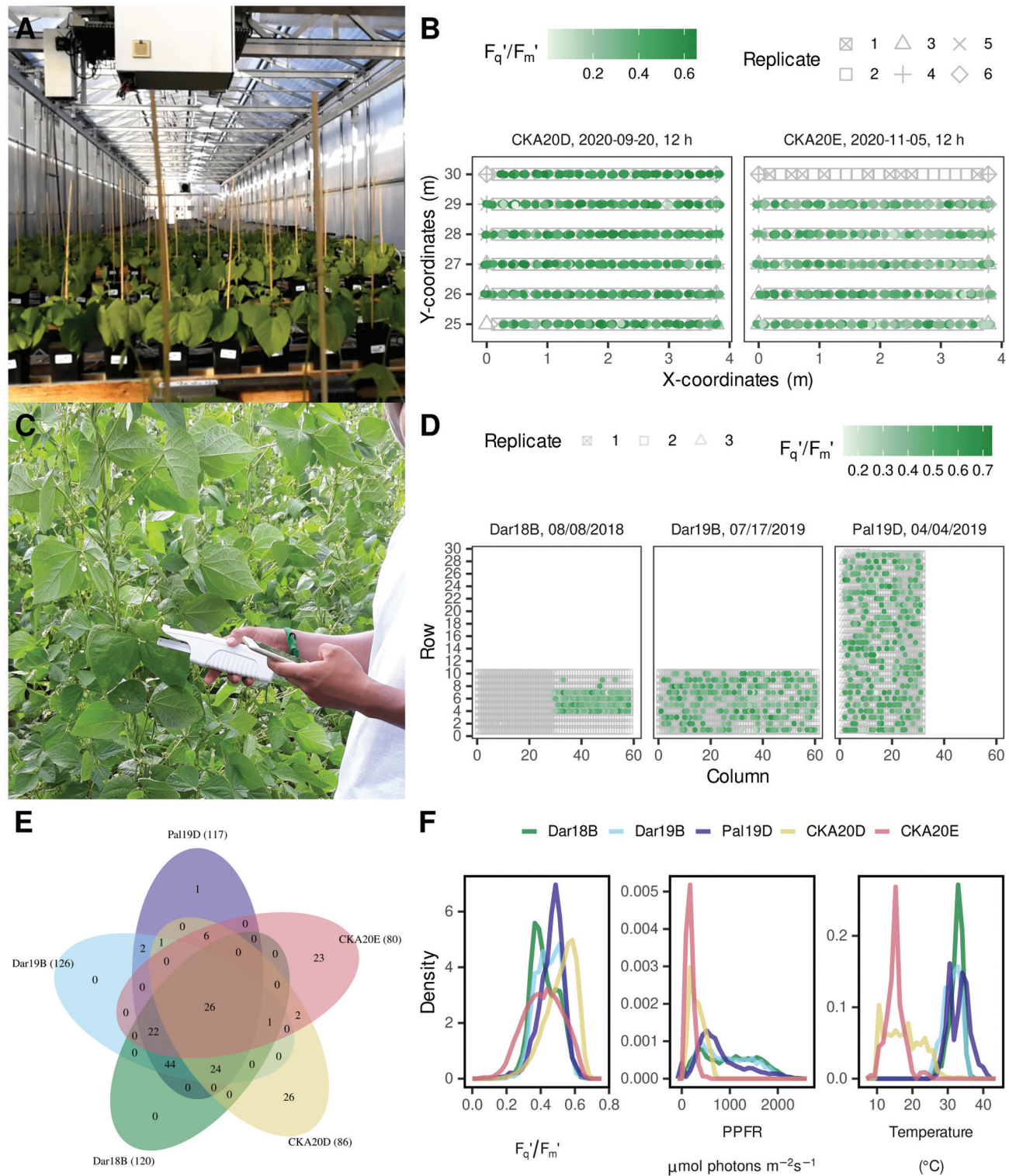


Fig. 1. Photosynthesis measurements based on probed chlorophyll fluorescence (ChlF) of climbing bean lines in the glasshouse and field. (A) Scans in the glasshouse were carried out by an automated light-induced fluorescence transient (LIFT) system under fluctuating conditions. (B) Climbing bean lines were arranged in pots (gray symbols) and the operating efficiency of PSII (F_q'/F_m') was measured (green) every hour of a measurement day over two growing seasons. Examples of a single measurement run at 12.00 h of the two glasshouse trials are shown ($n=688$ respective 700 measurements). (C) The F_q'/F_m' was measured by hand-held MultispeQ devices in the field on 14 d in three growing seasons. (D) Examples of a single measurement day of the three field trials are shown ($n=1879$, 1359, and 1203 measurements, respectively). (E) In total, 178 climbing bean lines were phenotyped and partially replicated in the five trials. The two glasshouse trials (CKA20D and CKA20E) were carried out at Campus Klein-Altendorf in 2020 in Germany whereas

the three field trials (Dar18B, Dar19B, and Pal19D) took place in Darién (Dar) and Palmira (Pal) in Colombia in 2018 and 2019. The number of lines for each trial is given in parentheses. (F) Probability density diagram of the acquired F_q'/F_m' and its associated photosynthetic photon fluence rate (PPFR) and temperature values for the five trials ($n=77\ 780$ measurements).

that scans the pots from a distance of 1.5 m in the glasshouse (Fig. 1A). Measurements were automated by scanning the plant canopy at a speed of 3 cm s^{-1} , as described by Keller *et al.* (2019a). Briefly, ChlF measurements were acquired by using a fast repetition rate flash (FRRF) every 1.5 s. The FRRF generates $\sim 40\ 000\ \mu\text{mol photons m}^{-2}\text{ s}^{-1}$ of blue light (445 nm) with an excitation circle of 7 cm^2 at a distance of 60 cm. The focus was then adjusted and optimized to a measuring distance of 1.25 m with a dynamic range of $\sim 20\text{ cm}$, optimizing the measurements for leaves between 15 cm and 35 cm above the ground. The minimum ChlF yield was determined by the first flashlet, and the maximum ChlF yield (F_m') was the average of the 301st and 302nd flashlet (Keller *et al.*, 2019b). The difference between the two ChlF yields represents the variable ChlF (F_q'), which is used to calculate F_q'/F_m' (Baker, 2008). Spectral measurements were acquired between the FRRFs through the LIFT lens using the LIFT's integrated STS-VIS spectrometer (Ocean Insight, Orlando, FL, USA). Wavelengths were recorded between 400 nm and 800 nm with a resolution of 0.46 nm and an acquisition time of 1 s. This resulted in ~ 3 measurements per pot and hour of a measurement day. The MERIS terrestrial chlorophyll index (MTCI) was calculated as:

$$\text{MTCI} = (R_{754} - R_{710}) / (R_{710} + R_{680}) \quad (1)$$

where R indicates the wavelength. The photochemical reflectance index (PRI) and normalized difference vegetation index (NDVI) were calculated as described in Keller *et al.* (2022b).

MultispeQ

The hand-held MultispeQ v.1 (PhotosynQ, USA) was used to acquire F_q'/F_m' in the field (Kuhlgert *et al.*, 2016). The adaxial leaf side of upper, fully irradiated, youngest, but already developed, leaves was measured in the early mornings (09.00 h–11.00 h) and early afternoons (13.00 h–15.00 h) to capture all possible phenotypic variability and to get information on genotype acclimation. Measurements were taken from the central part of the trifoliate, avoiding the central vein. The measured height was set at $\sim 1.5\text{ m}$, according to the growth stage and development of each line. The previous measurements showed that plant height from where the measurements were taken did not significantly affect the values (comparing fully developed leaves on the outer canopy; data not shown). The relative chlorophyll content was estimated by the transmitted red (650 nm) and infrared (940 nm) light using a series of measurements under increasing light intensities (Markwell *et al.*, 1995; Kuhlgert *et al.*, 2016). The spectral bands were calibrated using color cards provided by the manufacturer. All three trials were measured using the Photosynthesis RIDES no open/close protocol (photosynq.org/protocols).

Grain yield and agronomic traits

Agronomic traits, especially grain yield, for the VEC lines are publicly available (Keller *et al.*, 2022a). To determine above-ground biomass grown in the glasshouse, plants per pot were cut at the soil level, dried at $75\text{ }^\circ\text{C}$ for 48 h, and weighed.

Data preparation

Grain yield and biomass

For the glasshouse trials, each biomass phenotype (y_{ijkl}) was modeled with an effect for the experiment (E_i), the line (L_j), the $G \times E$ (LE_{ij}), the row (row_k), and the column ($column_l$) of the pots:

$$y_{ijkl} = E_i + L_j + LE_{ij} + row_k + column_l + \varepsilon_{ijkl} \quad (2)$$

The adjusted means for biomass were then extracted over both trials as well as for each separate trial, for each breeding line.

Regarding the three field trials, the spatially corrected adjusted means for grain yield and seed iron concentration (SdFe) overall and for the field trials separately were taken from Keller *et al.* (2022a).

Operating efficiency of PSII (F_q'/F_m')

Measurements under low light conditions ($>100\ \mu\text{mol photons m}^{-2}\text{ s}^{-1}$) and, for the LIFT, with a signal-to-noise ratio <30 were discarded. All variables were cleaned for outliers which were more distant from the median than 2.5 times the interquartile range. The LIFT measurements were associated with the corresponding pot (breeding line) according to the position of the platform recorded every 30 cm (Fig. 1). The measurements were then associated with the corresponding PPFR value recorded in the same minute.

Regarding the field trials, the MultispeQ device was used to measure F_q'/F_m' and the environmental variables.

Modeling of F_q'/F_m'

Basic line model

Measured F_q'/F_m' values (y_{jm}) were modeled in each experiment separately using fixed effects for each breeding line j (L_j), for the square root of PPFR at time point m (P_m), and for the interaction between line and the square root of PPFR (LP_{jm}), as well as an error term (ε_{jm}):

$$y_{jm} = \mu + L_j + P_m + LP_{jm} + \varepsilon_{jm} \quad (3)$$

For the field trials, L_j was replaced by a plot effect to account for spatial variability in the field (even between the same lines).

Line model with interaction for the glasshouse and field

To obtain adjusted means and responses for the glasshouse and field, model (3) included a fixed effect for each experiment (E_i) and an interaction between experiment, line, and PPFR at time point m (ELP_{ijm}). To account for the absorbance of the incoming light, an effect was added for the spectral index MTCI and the relative chlorophyll content (A_m) at the time point m for the LIFT and MultispeQ device, respectively. To account for light heterogeneity at the top of the canopy, the reflectance (B_m) at time point m was used to model F_q'/F_m' derived from LIFT canopy scans in the glasshouse. In the field, the F_q'/F_m' values measured at the leaf level were corrected for an effect for each of the MultispeQ devices. This resulted in the following model for F_q'/F_m' values (y_{ijm}) with the associated error term (ε_{ijm}):

$$y_{ijm} = \mu + E_i + L_j + P_m + ELP_{ijm} + A_m + B_m + \varepsilon_{ijm} \quad (4)$$

Adjusted means and responses for F_q'/F_m' were separately calculated for the glasshouse and the field experiments. F_q'/F_m' data points were classified as measurement error and discarded when they were a Cook's distance farther from the median than 50 times the upper quartile range. Breeding lines with <30 available measurements were discarded from the

analysis to ensure robust regression results. Previously, the sum of squares for each variable was calculated by adding the explanatory variables PRI, NDVI, relative chlorophyll content, and air temperature as additive effects for each line at time point m , as: $C_m = \sum_{q=1}^Q c_{mq} * \gamma_q$, where Q is the number of additional explanatory variables, c_{mq} is the measured value of the q th variable at time point m , and γ_q is the main effect (coefficient) for each variable. This extension of model (4) resulted in:

$$y_{ijm} = \mu + E_i + L_j + P_m + ELP_{ijm} + A_m + B_m + C_m + \varepsilon_{ijm} \quad (5)$$

The variables constituting C_m were dropped in model (4) because they explained <1.5% of the variance.

Response to PPFR

From models (3) and (4), the adjusted genotypic means for F_q'/F_m' and the response of F_q'/F_m' to PPFR ($\text{Response}_{G,PPFR}$) were extracted over all trials and separately for the glasshouse and field trials using the emmeans R package, as described by Keller et al. (2022b). Briefly, the $\text{Response}_{G,PPFR}$ refers to the coefficients b_{ELP} corresponding to the ELP term as $ELP = \{ELP_{ijm}\} = b_{ELP} \times Z_{LE} \times X_p$ where Z_{LE} is a design matrix for the lines in each experiment and X_p is a vector of the PPFR values. In other words, b_{ELP} describes the slope of the F_q'/F_m' values with increasing PPFR for each breeding line.

Statistical analysis

Correlation analysis

Adjusted means for biomass and grain yield were correlated to the $\text{Response}_{G,PPFR}$ of each breeding line separately for the glasshouse and the field trials based on model (4). Pearson correlation coefficient (r) was then calculated.

Post-hoc significance analysis

The Tukey HSD test was used to perform post-hoc significance analysis between the lines in the upper and lower 20% percentiles of the $\text{Response}_{G,PPFR}$ as well as the lines in between (intermediate), forming three different groups in terms of photosynthetic energy conversion efficiency.

Heritability calculation

Heritability (H^2) of the derived $\text{Response}_{G,PPFR}$ was calculated between the different experiments based on model (4). Heritability of the biomass was calculated based on model (2). In both cases, the following equation was used:

$$H^2 = \frac{\sigma_L}{(\sigma_L + \sigma_\varepsilon \sqrt{n})} \quad (6)$$

where σ_L was the variance of the breeding lines, and σ_ε the error variance, which was divided by the averaged number of replicates (\bar{n}) to account for the unbalanced design.

Genomic predictions

The GEBV for biomass and grain yield were calculated for three different models using the Bayesian generalized linear regression (BGLR) R package (Pérez and de los Campos, 2014). A Gaussian prior was assumed for the random effects. The Gibbs sampler generated 16 000 iterations of which the first 4000 were burned-in and the remainder were thinned by

factor 5. Genomic information for 14 913 single nucleotide polymorphism (SNP) markers was already available for the climbing bean lines (Keller et al., 2022a).

Basic genotype model

In the genotype model, the phenotype (y_{ij}) was described as the sum of a fixed effect for the i th experiment (E_i), a random effect for the j th genotype (g_j) and the error ε_{ij} :

$$y_{ij} = E_i + g_j + \varepsilon_{ij} \quad (7)$$

where $g \sim N(0, K\sigma_g^2)$ with K as the kinship matrix. The kinship matrix based on the SNP matrix was calculated using the rrBLUP package.

G×E model

In the G×E model, an interaction effect between the i th location and the j th genotype (gE_{ij}) was added:

$$y_{ij} = E_i + g_j + gE_{ij} + \varepsilon_{ij} \quad (8)$$

where $gE \sim N(0, I \otimes K\sigma_{gE}^2)$ with I as the identity matrix for the experiments and \otimes denotes the Kronecker product.

G×Ec model

In the G×Ec model, the location effect was replaced (parameterized) by measured values, namely the $\text{Response}_{G,PPFR}$ (v_i), in the i th environment. Additionally, an interaction effect between the i th environmental covariates—in this case the physiological covariates ($V = v_i'$)—and the j th genotype (gV_{ij}) was added:

$$y_{ij} = E_i + v_j + g_j + gV_{ij} + \varepsilon_{ij} \quad (9)$$

with $v \sim N(0, V\sigma_v^2)$, $gV \sim N(0, V \otimes K\sigma_{gV}^2)$, and $\varepsilon \sim N(0, I\sigma_\varepsilon^2)$.

Cross-validation

Cross-validation was carried out to predict new lines in an observed environment (CV1) and new lines in a new environment (CV2). CV1 gives information about the prediction accuracy within one trial while CV2 represents a more practical situation where new lines are predicted for a new season (i.e. environment). CV1 and CV2 were calculated for all three models presented. In the case of model (9), the ChIF data would be known for all lines; that is the $\text{Response}_{G,PPFR}$ would also be available for new lines and in a new environment. The cross-validation was 60-fold where the dataset was randomly divided into three parts 10 times and each part was validated based on the training of the remaining two-thirds of the dataset.

Genome-wide association study

GWAS was carried out using the Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) algorithm from GAPIT (Huang et al., 2019). The first three principal components of the SNP matrix were used to correct for population structure as described for this climbing bean panel in Keller et al. (2022a). SNP markers with a minor allele frequency <5% were excluded. The effect of significant markers was visualized using boxplots for the trials of the current study and for six previous trials carried out by Keller et al. (2022a) and two trials by Barbosa et al. (2018). Candidate genes were identified within a region with a maximal distance of 100 000 bp from an identified significant marker-trait association and based on their relationship to the

photosynthetic apparatus using gene annotations from Phytozome v13 (Goodstein *et al.*, 2012).

Data visualization

Plots were done in R using ggplot2 with ggh4x extension (Wickham, 2016). The magick R package was used for images (Ooms, 2023).

Results

The F_q'/F_m' of climbing bean lines was assessed throughout the growing season by taking ~700 measurements per hourly scan in the glasshouse (Fig. 1A, B) and ~1500 measurements per day in the field (including morning and afternoon measurements; Fig. 1C, D). Between 80 and 126 lines per trial were phenotyped, resulting in 178 lines of which 26 were measured in all five trials (Fig. 1E). A total of 79 645 measurements were acquired over the growing seasons in 57 d under fluctuating conditions (Fig. 1F). The associated weather data represented the growing conditions, which varied across the five trials. PPFR and temperature in particular varied between the glasshouse and field trials. The PPFR and temperature values

were highest in the Palmira field trial (Pal19D) and lowest in the CKA20E glasshouse trial. The PPFR explained 10.3% and 69.9% of the variance for F_q'/F_m' in the glasshouse and field trials using model (5), respectively (Tables 1, 2). Additionally, ~1% was explained by the square root of PPFR. Air temperature as well as the measurement dates accounted for <1% of the variation for F_q'/F_m' . Reflectance and MTCI explained 7.0% and 4.6% of the variation for F_q'/F_m' , respectively, in the glasshouse canopy scans (Table 1), while relative chlorophyll content explained 1.7% in the field (Table 2). The breeding lines (i.e. the genotypes) explained 10.0% and 1.5% of the variance for F_q'/F_m' in the glasshouse and field trials, respectively.

Light use efficiency and yield

After filtering, F_q'/F_m' was modeled using 61 252 and 11 201 ChlF measurements taken in the glasshouse and field, respectively (Fig. 2A). The general response of F_q'/F_m' to PPFR was more pronounced in low-yielding lines, indicating lower light use efficiency (LUE), especially at high PPFR (Fig. 2B). The $\text{Response}_{G:\text{PPFR}}$, calculated according to the basic model (3), showed significant correlations with yield in four trials (Supplementary Fig. S2). This highlights the importance of the response of F_q'/F_m' to PPFR for biomass production, although the $\text{Response}_{G:\text{PPFR}}$ (i.e. the Line:PPFR interaction) explained <1% of the variation for F_q'/F_m' (Tables 1, 2). Model (4) calculated the $\text{Response}_{G:\text{PPFR}}$ with additional effects for chlorophyll content, reflectance, and device ID. This model showed slightly higher correlations between $\text{Response}_{G:\text{PPFR}}$ and yield, ranging

Table 1. Explanatory variables of the operating efficiency of PSII (F_q'/F_m') in two climbing bean glasshouse trials

Explanatory variable	df	Sum of squares	Explained variance (%)
Trial	1	163.757	22.253
PPFR	1	75.915	10.316
Line	130	73.529	9.992
Reflectance	1	51.271	6.967
MTCI	1	33.618	4.568
Trial:Line	34	12.463	1.694
Trial:MTCI	1	11.428	1.553
sqrt(PPFR)	1	7.662	1.041
PPFR:Line	130	7.6	1.033
PRI	1	5.502	0.748
NDVI	1	5.268	0.716
Temperature	1	2.647	0.36
Date	1	2.504	0.34
MTCI:NDVI	1	2.455	0.334
Trial:MTCI:NDVI	1	0.868	0.118
Trial:PPFR:Line	34	0.57	0.077
Trial:NDVI	1	0.425	0.058
Trial:Reflectance	1	0.093	0.013
Trial:PPFR	1	0.024	0.003
Residuals	62 741	278.302	37.818

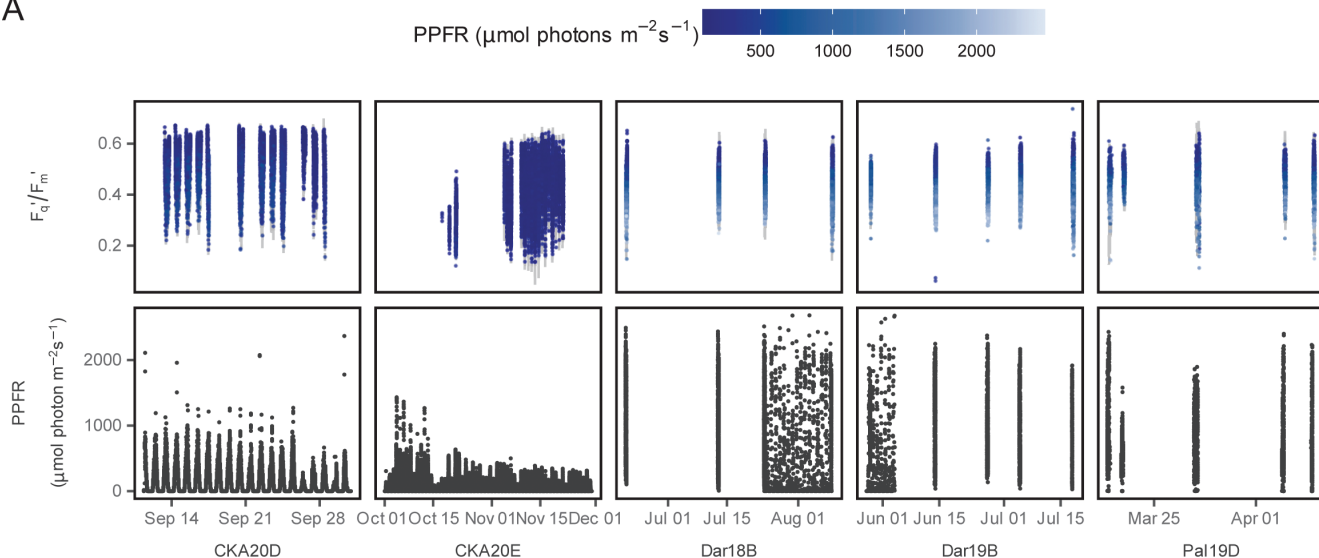
The degrees of freedom (df), associated sum of squares, explained variance, and the residual are calculated using linear regression according to model (5). The explanatory variables were trial, photosynthetic photon fluence rate (PPFR), line, MERIS terrestrial chlorophyll index (MTCI), normalized difference vegetation index (NDVI), photochemical reflectance index (PRI), reflectance, the square root (sqrt) of PPFR, temperature, and date, as well as the interactions between variables denoted by the ':' symbol. In total, 66 550 data points of 131 breeding lines were acquired by automated chlorophyll fluorescence (ChlF) measurements.

Table 2. Explanatory variables of the operating efficiency of PSII (F_q'/F_m') in three climbing bean field trials

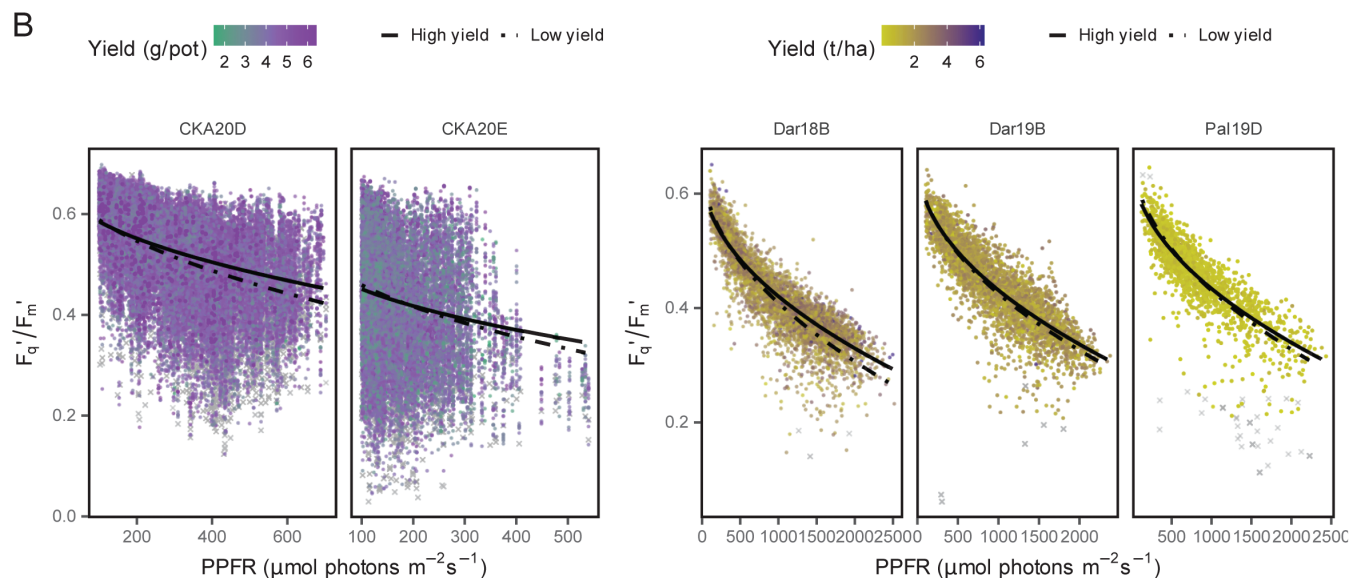
Explanatory variable	df	Sum of squares	Explained variance (%)
PPFR	1	54.525	69.917
Trial	2	3.019	3.872
Chlorophyll	1	1.314	1.686
Line	126	1.176	1.508
Trial:Line	234	0.851	1.091
Trial:Device.ID	13	0.845	1.083
Trial:PPFR:Line	234	0.685	0.879
Date	13	0.649	0.832
sqrt(PPFR)	1	0.563	0.722
PPFR:Line	126	0.49	0.629
Temperature	1	0.007	0.009
Trial:PPFR	2	0.001	0.001
Residuals	12 340	13.859	17.771

The degrees of freedom (df), associated sum of squares, explained variance, and the residual are calculated using linear regression according to model (5). The explanatory variables were photosynthetic photon fluence rate (PPFR), trial, line, relative chlorophyll content, date, and temperature, as well as the interactions between variables denoted by the ':' symbol. In total, 13 095 data points of 127 breeding lines were acquired by hand-held chlorophyll fluorescence (ChlF) measurements.

A



B



C

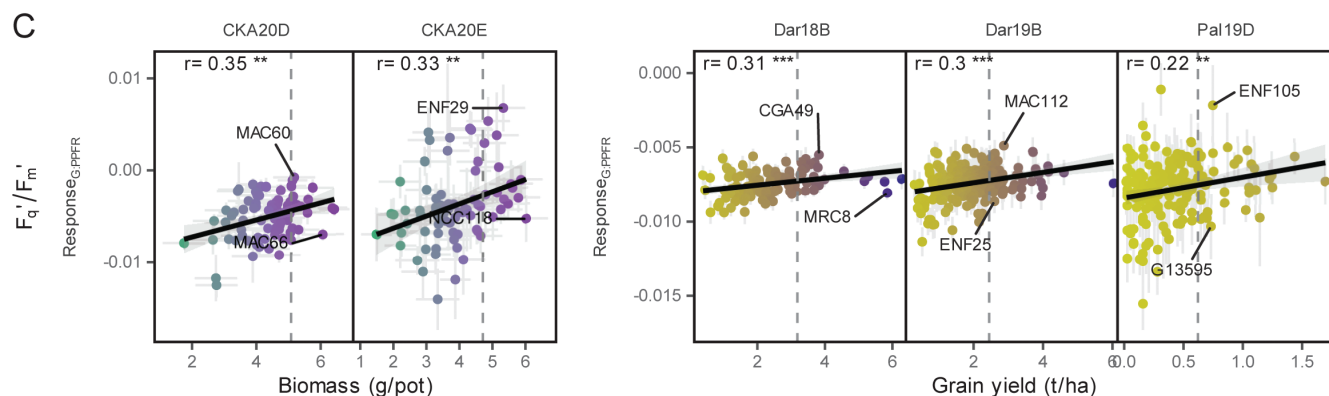


Fig. 2. Dynamic photosynthesis is linked to biomass and yield. (A) Operating efficiency of PSII (F_q'/F_m') was measured under a fluctuating photosynthetic photon fluence rate (PPFR) over the season in five trials ($n=72\,453$ selected in colors; discarded data points in gray; total $n=77\,780$). (B) The F_q'/F_m' was related to PPFR modeled with a square root term for the 20% percentile of high-yielding (solid line) and low-yielding lines (dotted line). (C) The response of F_q'/F_m' to PPFR ($\text{Response}_{G,PPFR}$) was modeled according to model (4) and correlated to biomass and yield. Light gray areas show the 95% confidence

interval of the regression line. Contrasting breeding lines for $\text{Response}_{\text{G,PPFR}}$ within the 20% percentile of high-yielding lines (dashed line) are labeled. Gray error bars show the SEM. The significance of the Pearson correlation coefficient (r) is indicated as *** $P < 0.001$; ** $P < 0.01$. The two glasshouse trials (CKA20D and CKA20E) were carried out at Campus Klein-Altendorf in 2020 in Germany whereas the three field trials (Dar18B, Dar19B, and Pal19D) took place in Darién (Dar) and Palmira (Pal) in Colombia in 2018 and 2019.

from $r = 0.22$ to 0.35 (Fig. 2C). The diagnostic plots showed that model residuals were independent over time, had constant variance, and no significant deviation from a normal distribution, indicating reasonable model fits (Supplementary Fig. S3). In contrast to $\text{Response}_{\text{G,PPFR}}$, the adjusted mean of F_q'/F_m' was not shown to be a reliable predictor of biomass (Supplementary Fig. S4). A significant effect on yield was observed when lines were grouped according to the calculated $\text{Response}_{\text{G,PPFR}}$. The breeding lines in the lower 20% quantile of the $\text{Response}_{\text{G,PPFR}}$, corresponding to a high ϵ_e , showed significantly higher yield compared with the lines in the upper 20% quantile in four out of five trials (Supplementary Fig. S5). This result confirms that the trait $\text{Response}_{\text{G,PPFR}}$ can be used as a selection target to identify lines with increased yield potential.

Correlation of traits and heritability

The $\text{Response}_{\text{G,PPFR}}$ values of the three field trials were positively correlated (Supplementary Fig. S6). An analysis of this trait across the field trials resulted in a heritability of 0.21 (Supplementary Table S1). In contrast, the $\text{Response}_{\text{G,PPFR}}$ values of the two glasshouse trials were negatively correlated and, therefore, had a heritability of 0.15. These results indicate strong G×E effects, reflecting different growing conditions. The heritabilities of biomass in the glasshouse and grain yield in the field were higher, 0.36 and 0.53, respectively. The overall $\text{Response}_{\text{G,PPFR}}$ between the three field trials and the two glasshouse trials was uncorrelated, as was biomass to grain yield (Fig. 3). However, both the overall $\text{Response}_{\text{G,PPFR}}$ of the glasshouse and field trials were significantly correlated with biomass and grain yield ($r = 0.23$ and 0.26), respectively, demonstrating prediction stability under similar growth conditions. Interestingly, above-ground biomass from the glasshouse trials correlated positively with SdFe evaluated in the field trials ($r = 0.19$).

Genomic assisted predictions

The genomic predictions of biomass and grain yield according to models (7) to (9) resulted in different prediction accuracies. The interaction terms in the G×E and G×Ec models, generally increased the prediction accuracy in the five trials (Fig. 4A). The variance explained by the genotype [i.e. the molecular markers as described in the basic genotype model (7)] ranged from 26.4% to 50.6% (Fig. 4B). Comparing the variance components of the G×E model (8) and the G×Ec model (9), the effects of the interaction term were similar (Fig. 4B). This result indicates that the $\text{Response}_{\text{G,PPFR}}$ as covariate

was almost sufficient to describe all the G×E effects on the yield. With early-stage phenotypic information available (the $\text{Response}_{\text{G,PPFR}}$ in our study), the G×E model (8) performed best when predicting new lines in an observed environment (CV1) and new lines in a new environment (CV2) (Fig. 4C). The G×Ec model (9) achieved averaged prediction accuracies between 0.35 and 0.38, an increase in CV2 of 53.1% compared with the G×E model (8). When predicting new lines in a new environment (CV2), the prediction accuracy of the G×E model (8) was similar to the basic genotype model (7), indicating strong year-to-year effects. Using only genomic data, the highest accuracies were obtained in both Darién trials. However, only one of these trials (Dar19B) exceeded previous phenomic prediction accuracy (without any genomic data, Fig. 2).

Genome-wide association study for photosynthetic response

Investigating genetic control over photosynthesis, GWAS revealed a significant association between $\text{Response}_{\text{G,PPFR}}$ and the SNP marker (Chr09_37766289_13052) on Chr 9 at 37.77 Mbp in the Dar18B trial (Fig. 5A, B). Although not significant in the remaining trials, the effect of this marker was also observed in the CKA20D glasshouse and the Pal19D field trials (Fig. 5C). Furthermore, this marker caused pleiotropic effects on yield and SdFe in the complete VEC population including 290 lines evaluated in eight and four trials (Barbosa *et al.*, 2018; Keller *et al.*, 2022a), respectively (Supplementary Fig. S7). However, the marker was significant neither for yield nor for SdFe (Supplementary Fig. S8), suggesting that the $\text{Response}_{\text{G,PPFR}}$ was a rather stable physiological component of those more complex traits. In fact, this marker (Chr09_37766289_13052) showed a positive effect on yield in six out of the eight field trials. That was more stable than several markers for yield identified in these trials (Supplementary Fig. S9). A candidate gene (Phvul.009G258500) expressing a curvature thylakoid 1 protein was identified in proximity at 37.84 Mbp, which could be responsible for the improved photosynthetic response.

Discussion

Photosynthesis is a promising target for improving crop production, because the energy conversion efficiency of sunlight to biomass continues to show high genetic variation even in high-yielding breeding lines (Zhu *et al.*, 2010; Driever *et al.*,

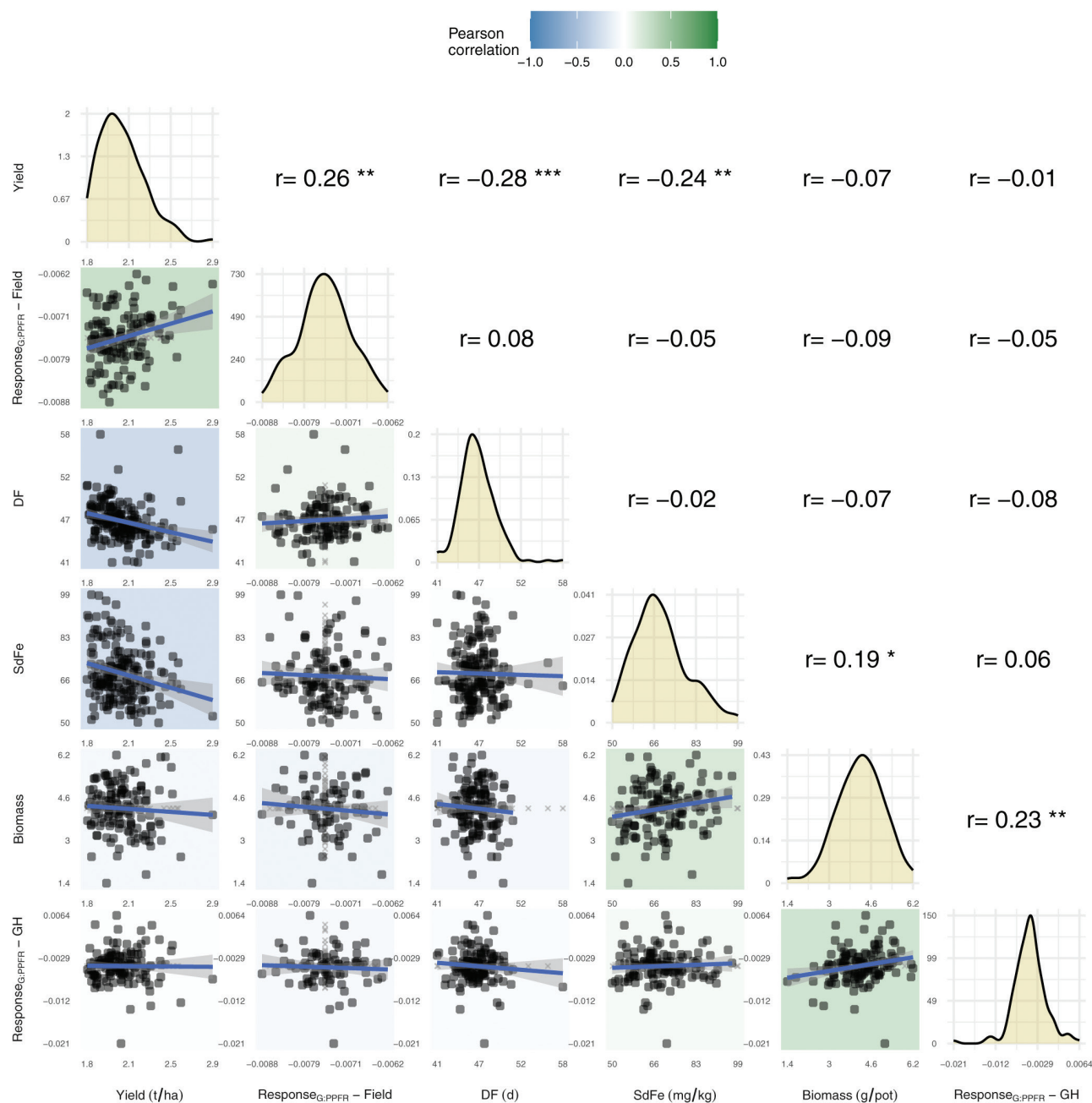


Fig. 3. Correlations of overall adjusted means between photosynthetic and agronomic traits. The traits were grain yield, response of F_q'/F_m' to PPFR ($\text{Response}_{G,PPFR}$) of the field trials, days to flowering (DF), seed iron concentration (SdFe), biomass, and $\text{Response}_{G,PPFR}$ of the glasshouse (GH) trials. The diagonal shows a density diagram with the estimated distribution of the phenotypes ($n=131$ lines in the greenhouse and $n=127$ lines in the field; total $n=178$ lines). The significance of the Pearson correlation coefficient (r) from pairwise observations (with non-missing values) is indicated as *** $P<0.001$; ** $P<0.01$; * $P<0.05$. Missing values were imputed using the population mean (shown in gray) and were only used to display complete pairwise observations.

2014; Keller et al., 2022b). Our study highlights the potential of automated high-throughput ChlF measurements for the selection of thousands of breeding lines in early generations. This approach is valid in the new environments and when there is a possibility to use similar high-throughput techniques, especially

where genomic selection is still very expensive. The positive relationship between $\text{Response}_{G,PPFR}$ and biomass production was shown using single pot experiments (potentially single plants; we used two plants per pot) in the glasshouse. The data trends were confirmed with hand-held instruments in field conditions.

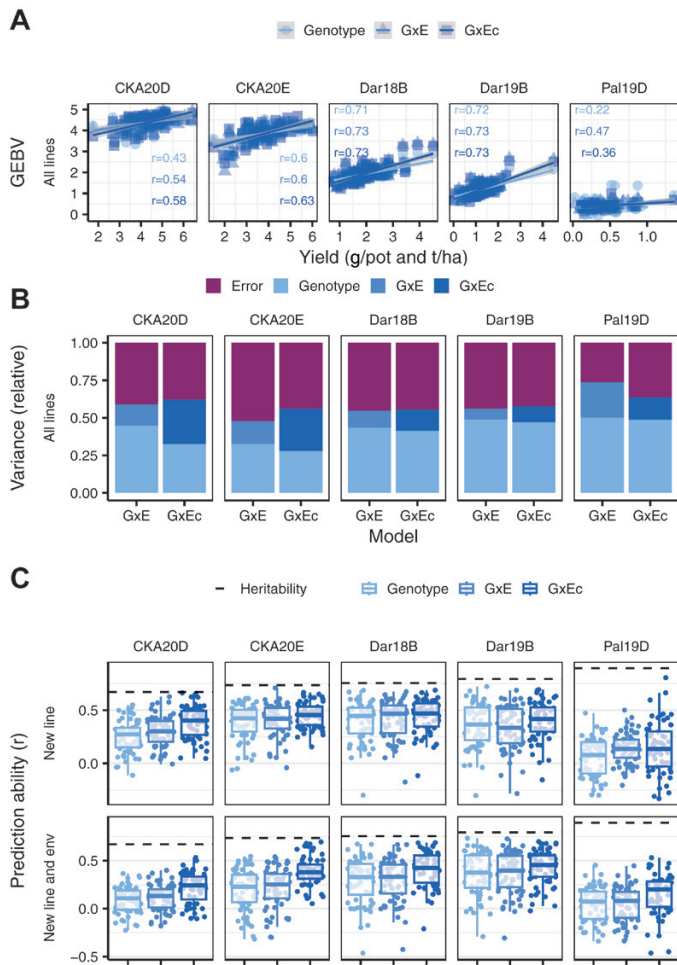


Fig. 4. Predictions for biomass and grain yield in the two glasshouse and three field trials, respectively, using genomic and chlorophyll fluorescence (ChlF) data. (A) The genomic estimated breeding values (GEBVs) were calculated based on genotypic data, genotype by environment interaction (G×E), and interaction between genotype and the covariate (G×Ec) as described in models (7) to (9). The response of F_q'/F_m' to PPFR ($\text{Response}_{G,PPFR}$) was used as covariate. The Pearson correlation coefficient (r) between measured and predicted value represents the prediction accuracy either of new lines in known environments or new lines in new environments. The two glasshouse trials (CKA20D, $n=86$ and CKA20E, $n=80$) were carried out at Campus Klein-Altendorf in 2020 in Germany, whereas the three field trials (Dar18B, $n=120$; Dar19B, $n=126$; and Pal19D, $n=117$) took place in Darién (Dar) and Palmira (Pal) in Colombia in 2018 and 2019. (B) The variance components for each model are shown in relative size. (C) Boxplot for biomass and yield prediction accuracy using two different 60-fold cross-validation scenarios and the three different models. The dashed line shows the square root of the spatial corrected heritability.

Solar energy conversion and yield

The relationship between F_q'/F_m' and PPFR tended towards a linear response, a known result in field plants grown under fluctuating light conditions, and its slope was a robust yield predictor (Meacham *et al.*, 2017; Violet-Chabrand *et al.*, 2017; Keller *et al.*, 2022b). Breeding lines which tolerate or benefit from high light

intensities—showing a low $\text{Response}_{G,PPFR}$ and a high ϵ_c —have potential for high LUE and high yield (Keller *et al.*, 2022b). Identified lines with low $\text{Response}_{G,PPFR}$ values and high yield (Fig. 2C) indicated (i) high transduction efficiency of photochemical energy to biomass, which may result from high root and stem conductivity and efficient carbon and mineral partitioning and (ii) the potential for higher yields of those lines when improving photosynthetic efficiency (Driever *et al.*, 2014). In the climbing bean lines studied, additional variation for symbiotic nitrogen fixation was reported by Barbosa *et al.* (2018), which may also contribute to differences in ϵ_c . In grain legumes, the costs of nitrogen fixation increase sink strength, emphasizing conversion efficiency as a breeding target. These findings highlight the potential of the source–sink ratio as well as photosynthesis as breeding targets (Evans and Clarke, 2019; Wu *et al.*, 2023).

Canopy and leaf photosynthesis

All trials and conditions in the current study showed a positive relationship between plant photosynthetic performance and yield. This result confirmed that $\text{Response}_{G,PPFR}$, as suggested by Keller *et al.* (2022b), is a robust yield indicator, regardless of whether the bean leaves were measured from the top of the canopy via automated measurements, or directly at the leaf level via handheld devices. However, the residuals were much higher in the glasshouse at the canopy level (37.8%, Table 1) than in the field at the leaf level (17.8%, Table 2). This finding probably represents the very different conditions in the field and glasshouse, where the lower PPFR affected the distribution of the residuals. It is likely that the spectral indices in the presented model (4) accounted partially for canopy structure and differences in leaf angles rather than in leaf pigment composition (Keller *et al.*, 2019a). Relative chlorophyll content showed a minor influence on F_q'/F_m' , because it determines light absorption at the leaf level and further distribution into the canopy (Acebron *et al.*, 2023).

The top leaves of the canopy are the most important part to measure because ~70% of solar energy is absorbed by the outer canopy (Song *et al.*, 2013). In soybean, Roth *et al.* (2022) showed that varieties with early canopy closure have higher yield, highlighting the importance of high light interception efficiency. However, quantifying the influence of the inner canopy structure, in terms of photosynthesis and nutrient re-assimilation, may further improve overall phenotyping precision and generate deeper knowledge on the effect of photosynthesis–yield relationship.

Photosynthesis and yield

Lopez *et al.* (2019) reported correlations between photosynthesis and grain yield of up to 0.8 in soybean. In wheat, Gutiérrez-Rodríguez *et al.* (2000) and Fischer *et al.* (1998) reported correlations of up to 0.79 between wheat yield and g_s when measured under light-saturated conditions (between $1200 \mu\text{mol m}^{-2} \text{s}^{-1}$ and $2000 \mu\text{mol m}^{-2} \text{s}^{-1}$). Our correlations were much lower, up

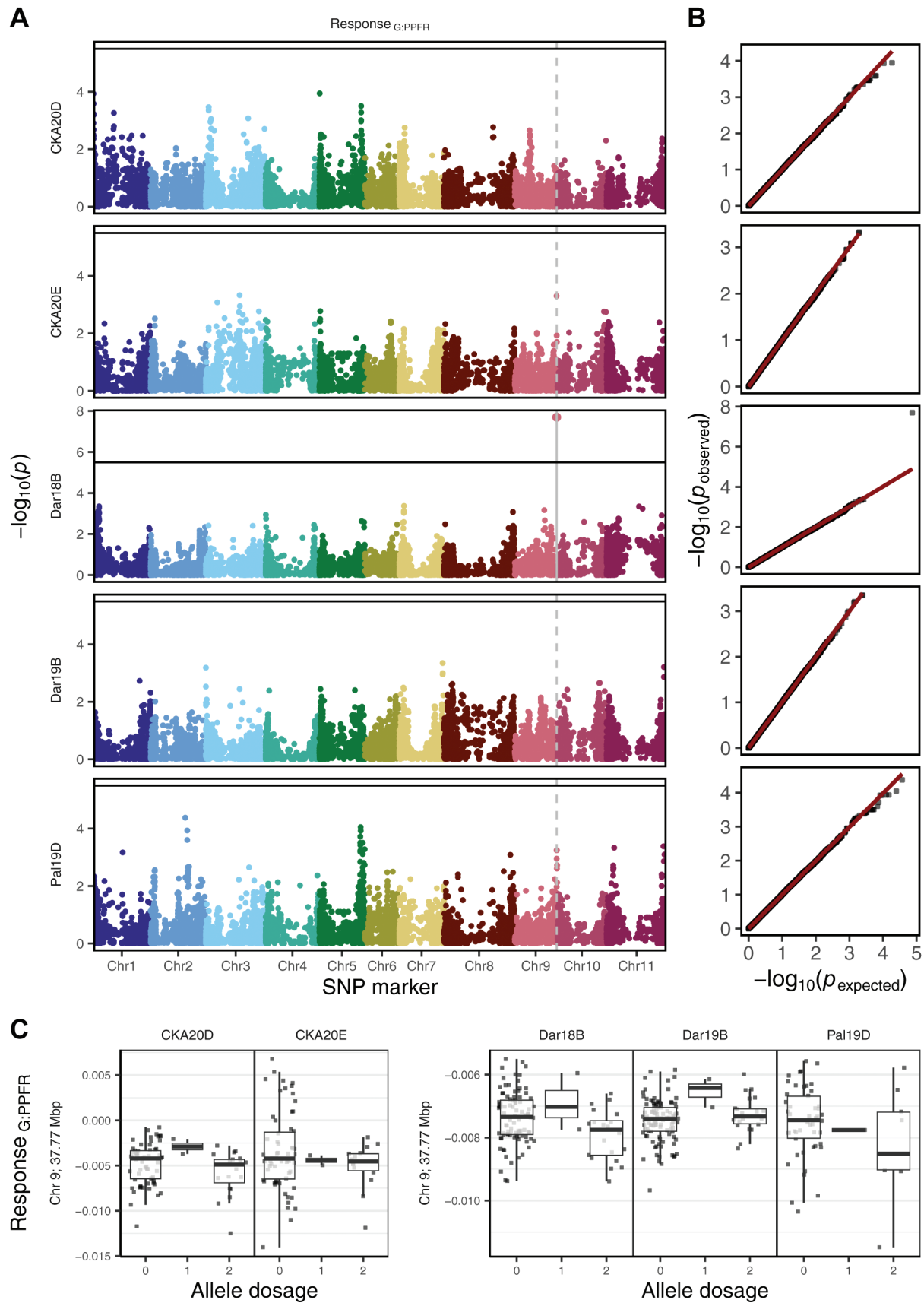


Fig. 5. The genome-wide association study (GWAS) for response of F_q'/F_m' to PPFR (Response_{G:PPFR}) of the five trials. (A) Manhattan plots showing the log-transformed P -value for every marker–trait association. The SNP markers are displayed according to their physical position on the genome.

The Bonferroni-corrected significance level is indicated at 5% with a solid line. (B) The quantile–quantile plot of expected and observed P -values for every marker. (C) Boxplot of $\text{Response}_{\text{G:PPFR}}$ according to allele dosage (0, 1, or 2 alternative alleles) of the significantly associated marker (Chr09_37766289_13052) for all five trials. The two glasshouse trials (CKA20D, $n=86$; and CKA20E, $n=80$) were carried out at Campus Klein-Altendorf in 2020 in Germany whereas the three field trials (Dar18B, $n=120$; Dar19B, $n=126$; and Pal19D, $n=117$) took place in Darién (Dar) and Palmira (Pal) in Colombia in 2018 and 2019.

to 0.43. The main reason for this difference is that we derived ChlF differently. Our method measured ET, which is subject to greater losses for biomass production than CO_2 assimilation, which occurs in between these two processes. Importantly, this means that when breeding lines are selected using ChlF, both high transduction efficiency and low photorespiration are probably included as ‘hidden’ selection traits. The lower prediction accuracy in our study using ChlF is sufficiently compensated by the possibility of automation, which allows significantly higher throughput than gasometric measurements. Additionally, the difference in measurements between ET and CO_2 assimilation facilitates selection of lines with less photorespiration. However, this point needs to be verified by testing the hypothesis that photorespiration is genetically diverse and varies in different environments, as recently reported for wheat lines (Coast *et al.*, 2019, 2021). Nevertheless, the current study showed a significant increase in yield when selecting for low $\text{Response}_{\text{G:PPFR}}$, namely high ϵ_e and LUE (Supplementary Fig. S5). Overall, the robust correlations between photosynthetic performance and yield are promising targets for increasing selection intensity in breeding programs.

Genetic control of photosynthetic response

One of the field trials identified a significant marker– $\text{Response}_{\text{G:PPFR}}$ association on chromosome 9 at 37.77 Mbp (Fig. 5). This association was also visible (although not significant) in another field trial as well as in one glasshouse trial. Our results suggest some common genetic control under these different conditions, despite the lack of phenotypic correlation for $\text{Response}_{\text{G:PPFR}}$ between glasshouse and field conditions. Furthermore, the identified marker showed a positive effect on grain yield and a negative effect on SdFe in previous field trials described by Keller *et al.* (2022a), suggesting that SdFe was likely diluted by increased photosynthates (Supplementary Fig. S7). Importantly, the marker was not detected with GWAS for grain yield, indicating $\text{Response}_{\text{G:PPFR}}$ as a distinct, but physiological and causal component of yield. The genomic region of the marker revealed a candidate gene (Phvul.009G258500 at 37.84 Mbp, a curvature thylakoid 1 protein) controlling the formation of chloroplasts (Sandoval-Ibáñez *et al.*, 2021). Interestingly, the *GmFtsH25* gene was recently identified in soybean, where it increases photosynthesis under high light intensities and alters chloroplast structure, allowing a denser stacking of the thylakoids (Wang *et al.*, 2023). Our quantitative trait locus (QTL), however, needs validation in further studies, especially to clarify under which conditions it is effective. In a much larger population including bush beans, Keller *et al.*

(2022a) identified another QTL for days to flowering in close proximity (<60 000 bp away) to our photosynthetic QTL. On chromosome 9, at 25.29 Mbp, Leitão *et al.* (2021) reported a QTL for g_s under glasshouse conditions using 158 common bean accessions. In conclusion, the $\text{Response}_{\text{G:PPFR}}$ probably contributed to yield as one of several physiological components which is rather stable across environments.

Phenomic or genomic selection?

Genomic data were used to predict grain yield and agronomic traits in 178 climbing bean lines. Our prediction accuracy was lower than that reported by Keller *et al.* (2022a), most probably due to the smaller pool of climbing bean lines in the training set. In bush beans, Keller *et al.* (2020) showed that the training set needed >200 lines to achieve the highest genomic prediction accuracy in an elite breeding panel. Additionally, phenotypic data collected at early stages in early-generation breeding plots have been shown to improve yield prediction (Rutkoski *et al.*, 2016). The $\text{Response}_{\text{G:PPFR}}$ used in our study showed potential to add information about the particular environment in which new lines are to be predicted. The low correlation between $\text{Response}_{\text{G:PPFR}}$ across trials, resulting in low broad-sense heritability, underlined that photosynthesis generally reflects the specific growing conditions in each trial (Keller *et al.*, 2019a), which is useful information for site-specific predictions. In agreement with this statement, the $\text{Response}_{\text{G:PPFR}}$ almost completely described different environments, comparing the variance components of the $\text{G} \times \text{E}$ model (8) and the $\text{G} \times \text{Ec}$ model (9). Furthermore, phenomic predictions using $\text{Response}_{\text{G:PPFR}}$ achieved even better prediction accuracy than the genomic prediction model (8) for new environments, except for one of the Darién trials (compare Fig. 2C with Fig. 4C). Similarly, in soybean, phenomic prediction outperformed genomic predictions in terms of accuracy (Lopez *et al.*, 2019). Combining phenomic and genomic data, as in the $\text{G} \times \text{Ec}$ model (9), increased the prediction accuracy in general. However, having phenotypic information of lines available in the newly used environment is not a realistic CV2 scenario for any breeding program. Acquiring phenotypic information, however, is very useful when genotypic information is not yet available and yield is difficult to assess, as is the case for selection in early generation breeding lines.

Conclusion

Photosynthetic performance has been shown to be a critical yield component. Using automated photosynthesis phenotyping, biomass was predicted by the $\text{Response}_{\text{G:PPFR}}$ in >130

common bean breeding lines in single pots. The phenomic prediction method also proved to be effective in the field using hand-held devices. Compared with genomic prediction based on molecular markers, phenomic prediction was more accurate in four out of five trials. Importantly, this suggests that phenomic prediction can be used as an early selection target, for example in F_3 or F_4 individuals in progeny rows, to screen large numbers of plants. Since the phenotyping accounts for $G \times E$, breeding programs can improve selection and better understand adaptation, especially in new environments. Similarly, the identified significant association between marker (Chr09_37766289_13052) and the $\text{Response}_{G:PPFR}$ trait may be used for marker-assisted selection on a physiological yield component where markers directly linked to yield are not stable throughout the environments. In contrast, genomic prediction was more accurate in a known location, confirming its potential to select more advanced (or phenomically pre-selected) breeding lines based on training data from the same or a similar location.

Supplementary data

The following supplementary data are available at [JXB online](#).

Table S1. Heritability (H^2) by trait and conditions.

Fig. S1. Climbing bean lines at 43 days after sowing.

Fig. S2. Adjusted means of operating efficiency of PSII (F_q'/F_m') and the response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$) were modeled according to the basic model (3) and correlated to biomass and yield.

Fig. S3. Diagnostic plots for the residuals from operating efficiency of PSII (F_q'/F_m') values using model (4) for all trials.

Fig. S4. Adjusted means of operating efficiency of PSII (F_q'/F_m') and the response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$) were modeled according to the basic model (4) and correlated to biomass and yield.

Fig. S5. Breeding lines grouped for the response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$) differ in yield in all five trials.

Fig. S6. Correlations of adjusted means between traits and trials. The traits were yield, biomass, and response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$) of the five trials.

Fig. S7. Allele dosage effect for days to flowering (DF), seed iron concentration (SdFe), and grain yield of the molecular marker (Chr09 37766289 13052) on chromosome 9 at 37.77 Mbp significantly linked to response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$).

Fig. S8. The genome-wide association study (GWAS) for biomass, grain yield, and seed iron concentration (SdFe) of all available climbing bean (VEC) trials.

Fig. S9. Allele dosage effect for response of F_q'/F_m' to PPFR ($\text{Response}_{G:PPFR}$) and yield of significantly associated molecular marker identified in trials with the climbing bean population (VEC).

Dataset S1. Chlorophyll fluorescence (ChlF) data from the field trials acquired by the MultispeQ device.

Dataset S2. Chlorophyll fluorescence (ChlF) data from the glasshouse trials acquired the light-induced fluorescence transient (LIFT) device.

Dataset S3. Biomass data from the glasshouse trials.

Acknowledgements

We thank the Field Lab Campus Klein-Altendorf of the University of Bonn for managing the plants in the glasshouse, especially Sandra Markwitz and Laura Berger. Special thanks to Seider, Ester, Marta, Herika, Jarden, and others for the MultispeQ measurements in the field in Darién and Palmira in Colombia.

Author contributions

BK, OM, BR, BS, and MU: conceptualization; JS, AS, AP-B, OM, and MU: conducting the experiments and providing the phenotypic data; BK, JS, AW, and MU: performing modeling, data analysis, and interpretation; BK: drafting the manuscript, which was improved by AW, AP-B, OM, BS, JS, and MU. All authors approved the manuscript before submission.

Conflict of interest

No conflict of interest declared.

Funding

The authors would like to thank the COOP Research Program on 'Sustainability in Food Value Chains' of the ETH Zurich World Food System Center and the ETH Zurich Foundation for supporting this project. The COOP Research Program is supported by the COOP Sustainability Fund. MU is grateful for the GIZ, PIAF (Germany) long-lasting support. Part of this work was enabled by the German-Plant-Phenotyping Network which is funded by the Federal Ministry of Education and Research (Germany) (project identification number: 031A053).

Data availability

Grain yield and agronomic traits are available from [Keller *et al.* \(2022a\)](#). The ChlF and biomass data from the glasshouse are provided in [Supplementary Datasets S1–S3](#). The MultispeQ data are also available on the PhotosynQ data base after creating an account (Darién 2018: <https://photosynq.org/projects/climbers-in-darien-2018>; Darién 2019: <https://photosynq.org/projects/19-04-climbers-darien>; Palmira 2019: <https://photosynq.org/projects/19-03-climbers-palmira>).

References

Acebron K, Salvatori N, Alberti G, Muller O, Peressotti A, Rascher U, Matsubara S. 2023. Elucidating the photosynthetic responses in chlorophyll-deficient soybean (*Glycine max*, L) leaf. *Journal of Photochemistry and Photobiology* **13**, 100152.

- Ashley DA, Boerma HR.** 1989. Canopy photosynthesis and its association with seed yield in advanced generations of a soybean cross. *Crop Science* **29**, 1042–1045.
- Baker NR.** 2008. Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annual Review of Plant Biology* **59**, 89–113.
- Barbosa N, Portilla E, Buendia HF, Raatz B, Beebe S, Rao I.** 2018. Genotypic differences in symbiotic nitrogen fixation ability and seed yield of climbing bean. *Plant and Soil* **428**, 223–239.
- Bilger W, Björkman O.** 1990. Role of the xanthophyll cycle in photo-protection elucidated by measurements of light-induced absorbance changes, fluorescence and photosynthesis in leaves of *Hedera canariensis*. *Photosynthesis Research* **25**, 173–185.
- Butler W.** 1978. Energy distribution in the photochemical apparatus of photosynthesis. *Annual Review of Plant Physiology* **29**, 345–378.
- Carmo-Silva E, Andralojc PJ, Scales JC, Driever SM, Mead A, Lawson T, Raines CA, Parry MAJ.** 2017. Phenotyping of field-grown wheat in the UK highlights contribution of light response of photosynthesis and flag leaf longevity to grain yield. *Journal of Experimental Botany* **68**, 3473–3486.
- Coast O, Posch BC, Bramley H, Gaju O, Richards RA, Lu M, Ruan Y-L, Trethowan R, Atkin OK.** 2021. Acclimation of leaf photosynthesis and respiration to warming in field-grown wheat. *Plant, Cell & Environment* **44**, 2331–2346.
- Coast O, Shah S, Ivakov A, et al.** 2019. Predicting dark respiration rates of wheat leaves from hyperspectral reflectance. *Plant, Cell & Environment* **42**, 2133–2150.
- Demmig-Adams B, Cohu CM, Muller O, Adams WW III.** 2012. Modulation of photosynthetic energy conversion efficiency in nature: from seconds to seasons. *Photosynthesis Research* **113**, 75–88.
- De Souza AP, Burgess SJ, Doran L, Hansen J, Manukyan L, Maryn N, Gotarkar D, Leonelli L, Niyogi KK, Long SP.** 2022. Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science* **377**, 851–854.
- Driever SM, Lawson T, Andralojc PJ, Raines CA, Parry MAJ.** 2014. Natural variation in photosynthetic capacity, growth, and yield in 64 field-grown wheat genotypes. *Journal of Experimental Botany* **65**, 4959–4973.
- Eeuwijk FA, Bustos-Korts DV, Malosetti M.** 2016. What should students in plant breeding know about the statistical aspects of genotype × environment interactions? *Crop Science* **56**, 2119–2140.
- Evans JR, Clarke VC.** 2019. The nitrogen cost of photosynthesis. *Journal of Experimental Botany* **70**, 7–15.
- Fischer RA, Rees D, Sayre KD, Lu Z-M, Condon AG, Saavedra AL.** 1998. Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Science* **38**, 1467–1475.
- Furbank RT, Jimenez-Berni JA, George-Jaeggli B, Potgieter AB, Deery DM.** 2019. Field crop phenomics: enabling breeding for radiation use efficiency and biomass in cereal crops. *New Phytologist* **223**, 1714–1727.
- Genty B, Briantais J-M, Baker NR.** 1989. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta* **990**, 87–92.
- Gifford RM, Evans LT.** 1981. Photosynthesis, carbon partitioning, and yield. *Annual Review of Plant Physiology* **32**, 485–509.
- Golding AJ, Finazzi G, Johnson GN.** 2004. Reduction of the thylakoid electron transport chain by stromal reductants—evidence for activation of cyclic electron transport upon dark adaptation or under drought. *Planta* **220**, 356–363.
- Goodstein DM, Shu S, Howson R, et al.** 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**, D1178–D1186.
- Gutiérrez-Rodríguez M, Reynolds MP, Larqué-Saavedra A.** 2000. Photosynthesis of wheat in a warm, irrigated environment: II. Traits associated with genetic gains in yield. *Field Crops Research* **66**, 51–62.
- Hermida-Carrera C, Kapralov MV, Galmés J.** 2016. Rubisco catalytic properties and temperature response in crops. *Plant Physiology* **171**, 2549–2561.
- Huang M, Liu X, Zhou Y, Summers RM, Zhang Z.** 2019. BLINK: a package for the next level of genomewide association studies with both individuals and markers in the millions. *GigaScience* **8**, giy154.
- Ishida S, Uebayashi N, Tazoe Y, Ikeuchi M, Homma K, Sato F, Endo T.** 2014. Diurnal and developmental changes in energy allocation of absorbed light at PSII in field-grown rice. *Plant and Cell Physiology* **55**, 171–182.
- Kaiser E, Morales A, Harbinson J.** 2018. Fluctuating light takes crop photosynthesis on a rollercoaster ride. *Plant Physiology* **176**, 977–989.
- Kalaji HM, Jajoo A, Oukarroum A, Brestic M, Zivcak M, Samborska IA, Cetner MD, Łukasik I, Goltsev V, Ladle RJ.** 2016. Chlorophyll a fluorescence as a tool to monitor physiological status of plants under abiotic stress conditions. *Acta Physiologica Plantarum* **38**, 102.
- Keller B, Ariza-Suarez D, de la Hoz J, Aparicio JS, Portilla-Benavides AE, Buendia HF, Mayor VM, Studer B, Raatz B.** 2020. Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Frontiers in Plant Science* **11**, 1001.
- Keller B, Ariza-Suarez D, Portilla-Benavides AE, et al.** 2022a. Improving association studies and genomic predictions for climbing beans with data from bush bean populations. *Frontiers in Plant Science* **13**, 830896.
- Keller B, Matsubara S, Rascher U, Pieruschka R, Steier A, Kraska T, and Muller O.** 2019a. Genotype specific photosynthesis × environment interactions captured by automated fluorescence canopy scans over two fluctuating growing seasons. *Frontiers in Plant Science* **10**, 1482.
- Keller B, Vass I, Matsubara S, Paul K, Jedmowski C, Pieruschka R, Nedbal L, Rascher U, Muller O.** 2019b. Maximum fluorescence and electron transport kinetics determined by light-induced fluorescence transients (LIFT) for photosynthesis phenotyping. *Photosynthesis Research* **140**, 221–233.
- Keller B, Zimmermann L, Rascher U, Matsubara S, Steier A, Muller O.** 2022b. Toward predicting photosynthetic efficiency and biomass gain in crop genotypes over a field season. *Plant Physiology* **188**, 301–317.
- Kromdijk J, Iowacka KG, Leonelli L, Gabilly ST, Iwai M, Niyogi KK, Long SP.** 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857.
- Kuhlgert S, Austic G, Zegarac R, et al.** 2016. MultispeQ Beta: a tool for large-scale plant phenotyping connected to the open PhotosynQ network. *Royal Society Open Science* **3**, 160592.
- Leitão ST, Bicho MC, Pereira P, Paulo MJ, Malosetti M, Araújo SS, van Eeuwijk F, Vaz Patto MC.** 2021. Common bean SNP alleles and candidate genes affecting photosynthesis under contrasting water regimes. *Horticulture Research* **8**, 1–14.
- Long SP, Marshall-Colon A, Zhu X-G.** 2015. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66.
- Lopez MA, Xavier A, Rainey KM.** 2019. Phenotypic variation and genetic architecture for photosynthesis and water use efficiency in soybean (*Glycine max* L Merr). *Frontiers in Plant Science* **10**, 680.
- López-Calcano PE, Brown KL, Simkin AJ, Fisk SJ, Viallet-Chabrand S, Lawson T, Raines CA.** 2020. Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants* **6**, 1054–1063.
- Markwell J, Osterman JC, Mitchell JL.** 1995. Calibration of the Minolta SPAD-502 leaf chlorophyllmeter. *Photosynthesis Research* **46**, 467–472.
- Meacham K, Sirault X, Quick WP, von Caemmerer S, Furbank R.** 2017. Diurnal solar energy conversion and photoprotection in rice canopies. *Plant Physiology* **173**, 495–508.
- Meuwissen TH, Hayes BJ, Goddard ME.** 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829.
- Molero G, Reynolds MP.** 2020. Spike photosynthesis measured at high throughput indicates genetic variation independent of flag leaf photosynthesis. *Field Crops Research* **255**, 107866.
- Murchie EH, Kefauver S, Araus JL, Muller O, Rascher U, Flood PJ, Lawson T.** 2018. Measuring the dynamic photosyntheticome. *Annals of Botany* **122**, 207–220.

- Murchie EH, Ruban AV.** 2020. Dynamic non-photochemical quenching in plants: from molecular mechanism to productivity. *The Plant Journal* **101**, 885–896.
- Ooms J.** 2023. magick: Advanced Graphics and Image-Processing in R. <https://cran.r-project.org/web/packages/magick/vignettes/intro.html>
- Parry MAJ, Reynolds M, Salvucci ME, Raines C, Andralojc PJ, Zhu X-G, Price GD, Condon AG, Furbank RT.** 2011. Raising yield potential of wheat II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany* **62**, 453–467.
- Pérez P, de los Campos G.** 2014. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**, 483–495.
- Pieruschka R, Klimov D, Kolber ZS, Berry JA.** 2010. Monitoring of cold and light stress impact on photosynthesis by using the laser induced fluorescence transient (LIFT) approach. *Functional Plant Biology* **37**, 395–402.
- Porcar-Castell A, Tyystjärvi E, Atherton J, van der Tol C, Flexas J, Pfundel EE, Moreno J, Frankenberg C, Berry JA.** 2014. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *Journal of Experimental Botany* **65**, 4065–4095.
- Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, Parry M, Slafer G.** 2012. Achieving yield gains in wheat. *Plant, Cell & Environment* **35**, 1799–1823.
- Roth L, Barendregt C, Bétrix C-A, Hund A, Walter A.** 2022. High-throughput field phenotyping of soybean: spotting an ideotype. *Remote Sensing of Environment* **269**, 112797.
- Rutkoski J, Poland J, Mondal S, Autrique E, Pérez LG, Crossa J, Reynolds M, Singh R.** 2016. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3* **6**, 2799–2808.
- Salter WT, Merchant A, Trethowan RM, Richards RA, Buckley TN.** 2020. Wide variation in the suboptimal distribution of photosynthetic capacity in relation to light across genotypes of wheat. *AoB Plants* **12**, plaa039.
- Sandoval-Ibáñez O, Sharma A, Bykowski M, et al.** 2021. Curvature thylakoid 1 proteins modulate prolamellar body morphology and promote organized thylakoid biogenesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **118**, e2113934118.
- Schreiber U, Schliwa U, Bilger W.** 1986. Continuous recording of photochemical and non-photochemical chlorophyll fluorescence quenching with a new type of modulation fluorometer. *Photosynthesis Research* **10**, 51–62.
- Sinclair TR, Rufty TW, Lewis RS.** 2019. Increasing photosynthesis: unlikely solution for world food problem. *Trends in Plant Science* **24**, 1032–1039.
- Song Q, Zhang G, Zhu X-G, Song Q, Zhang G, Zhu X-G.** 2013. Optimal crop canopy architecture to maximise canopy photosynthetic CO₂ uptake under elevated CO₂—a theoretical study using a mechanistic model of canopy photosynthesis. *Functional Plant Biology* **40**, 108–124.
- South PF, Cavanagh AP, Liu HW, Ort DR.** 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* **363**, eaat9077.
- Theeuwes TPJM, Logie LL, Harbinson J, Aarts MGM.** 2022. Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany* **73**, 3122–3137.
- Violet-Chabrand S, Matthews JS, Simkin AJ, Raines CA, Lawson T.** 2017. Importance of fluctuations in light on plant photosynthetic acclimation. *Plant Physiology* **173**, 2163–2179.
- Wang L, Yang Y, Yang Z, et al.** 2023. GmFtsH25 overexpression increases soybean seed yield by enhancing photosynthesis and photosynthates. *Journal of Integrative Plant Biology* **65**, 1026–1040.
- Wickham H.** 2016. ggplot2: elegant graphics for data analysis, 2nd edn. New York: Springer.
- Wu A, Brider J, Busch FA, et al.** 2023. A cross-scale analysis to understand and quantify the effects of photosynthetic enhancement on crop growth and yield across environments. *Plant, Cell & Environment* **46**, 23–44.
- Yamori W, Masumoto C, Fukayama H, Makino A.** 2012. Rubisco activase is a key regulator of nonsteady-state photosynthesis at any leaf temperature and, to a lesser extent, of steady-state photosynthesis at high temperature. *The Plant Journal* **71**, 871–880.
- Yamori W, Shikanai T.** 2016. Physiological functions of cyclic electron transport around photosystem I in sustaining photosynthesis and plant growth. *Annual Review of Plant Biology* **67**, 81–106.
- Zelitch I.** 1982. The close relationship between net photosynthesis and crop yield. *Bioscience* **32**, 796–802.
- Zhang R, Sharkey TD.** 2009. Photosynthetic electron transport and proton flux under moderate heat stress. *Photosynthesis Research* **100**, 29–43.
- Zhu XG, Long SP, Ort DR.** 2010. Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biology* **61**, 235–261.