

# Soybean Nitrogen Uptake and Utilization in Argentina and United States Cultivars

José L. Rotundo,<sup>★</sup> Lucas Borrás, Jason De Bruin and Palle Pedersen

## ABSTRACT

Crop production increase needed to satisfy a growing world population depends, at least partially, on increasing current genetic gain in yield. Theory proposes that increased genetic gain can be attained using diverse high-yielding parents. Physiological traits, compared to molecular or morphological markers, are hypothesized to better estimate parental diversity. A trait-based hybridization approach will require assessing diversity for physiological traits. Here, we phenotyped several traits using two physiological frameworks to assess diversity in 25 and 65 elite cultivars from Argentina (ARG) and the United States (USA), respectively. First, we identified the highest-yielding cultivar clusters across two environments in each country. These cultivars had the highest N uptake at both ARG and USA. Therefore, there was no genotypic diversity for total N uptake within each cluster. For other traits, the highest yielding clusters did not show the highest values. There was residual diversity within ARG and USA highest yielding clusters in the temporal pattern of N uptake, N use efficiency, and N harvest index. Stacking these traits in one cultivar could potentially increase yield by 13%. The possibility of such stacking, however, depends on the nature of the phenotypic correlation among traits. We demonstrated that several trade-off correlations between phenotyped traits, thought to hinder stacking, are actually not biologically based.

J.L. Rotundo and L. Borrás, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario/CONICET. Campo Experimental Villarino S/N, Zavalla S2125ZAA, Santa Fe, Argentina; J. De Bruin and P. Pedersen, Agronomy Dep., Iowa State Univ., 110 Curtiss Hall, Ames, IA 50011; J. De Bruin, current address, DuPont Pioneer, 21888 N 950th Rd., Adair, IL 61411; P. Pedersen, current address, Syngenta Crop Protection, P.O. Box 18300, Greensboro, NC 27419. Received 16 Sept. 2013. <sup>★</sup>Corresponding author (jrotundo@unr.edu.ar).

**Abbreviations:** Agro2009, 2009 crop season, Agronomy Farm, near Ames, IA; ARG, Argentina; conc\_vbR7, vegetative biomass nitrogen concentration at physiological maturity; HI, harvest index; John2009, 2009 crop season, Johnson Farm, near Ames, IA;  $N_{veg}$ , total nitrogen uptake between planting and flowering;  $N_{remob}$ , apparent nitrogen remobilization; NHI, nitrogen harvest index for biomass production; NUE, nitrogen use efficiency for biomass production; seed N, seed nitrogen concentration at physiological maturity; syR7, seed yield at physiological maturity; USA, United States; vbR7, vegetative (non-seed) biomass at physiological maturity; Zav2010, 2009 to 2010 crop season, Zavalla, Santa Fe, Argentina; Zav2011, 2010 to 2011 crop season, Zavalla, Santa Fe, Argentina.

**S**OYBEAN [*Glycine max* (L.) merr.] seed yield has been increasing between 0.5 and 1% per year during the last 20 years in the United States (Specht et al., 1999). Approximately 50% of this increase is due to genetic improvement while the other 50% is attributed to better management practices (Duvick, 2005; Egli, 2008; Specht et al., 1999). Future yield increases must continue to come, at least partially, from genetic improvement.

A more rapid rate of genetic gain could be achieved through increased interaction between crop physiologists and plant breeders (Cooper and Hammer, 2005; Duvick, 2005; Hammer et al., 2006). Traditional breeding of autogamous species is an empirical

Published in Crop Sci. 54:1153–1165 (2014).

doi: 10.2135/cropsci2013.09.0618

© Crop Science Society of America | 5585 Guilford Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

activity based on hybridization of high-yielding cultivars, population development, progeny evaluation, and selection for yield (Fehr, 1998). Multidisciplinary approaches are needed to fulfill the objective of increasing future genetic gains (Wollenweber et al., 2005). One aspect where crop physiology can assist breeding is parental selection to develop populations with higher chances of producing transgressive highest-yielding progeny (Reynolds et al., 2011). The probability of obtaining transgressive high-yielding progeny is expected to be highest as parental genetic distance increases (Falconer and Mackay, 1960). Genetic distance has been explored using parental coefficients (Gizlice et al., 1993), molecular markers (Brown-Guedira et al., 2000), and/or simple morphological traits (Perry, 1991). Empirical evidence relating genetic distance and transgressive segregation have yielded mixed results (Bohn et al., 1999; Burkhamer et al., 1998; Carpentieri-Pípolo et al., 2000; Gizlice et al., 1993; Kisha et al., 1997). We hypothesize that using physiological traits to assess genetic distance may have more predictive value than other characteristics less, or not directly, related with yield (e.g., molecular markers, morphological traits, or pedigree information; Reynolds et al., 2011). A required step for using physiological traits to assess genetic distance is defining a physiological framework to identify environmentally-stable traits related to seed yield.

Seed yield is highly dependent on N uptake (Sinclair and Horie, 1989). There are different theoretical frameworks linking N uptake and N use with yield (Cregan and Vanberkum, 1984; Xu et al., 2012). Total N uptake at physiological maturity (R7) can be characterized according to the timing of N uptake:

$$\text{Total N at R7 (kg N ha}^{-1}\text{)} = N_{\text{veg}} + N_{\text{R1-R5}} + N_{\text{R5-R7}} \quad [1]$$

where  $N_{\text{veg}}$  is uptake between planting and R1,  $N_{\text{R1-R5}}$  is uptake during the seed number determination period, and  $N_{\text{R5-R7}}$  is uptake during the seed-filling period. Seed yield can then be explained, as a function of total uptake, using this framework:

$$\text{Seed yield (kg ha}^{-1}\text{)} = \text{Total N at R7} \times \text{NUE} \times \text{HI} \quad [2]$$

where NUE is physiological N use efficiency for biomass production (kg total biomass kg N uptake<sup>-1</sup>) (Novoa and Loomis, 1981) and HI is harvest index (kg seed kg total shoot biomass<sup>-1</sup>). Another framework frequently used to analyze seed yield variation in relation to N is:

$$\text{Seed yield (kg ha}^{-1}\text{)} = \text{Total N at R7} \times \text{NHI} \times \text{Seed N}^{-1} \quad [3]$$

where NHI is the N harvest index (kg seed N kg N uptake<sup>-1</sup>) and seed N is the N concentration in the seed.

Another trait characterizing soybean N use, not included explicitly in the above mentioned functions, is apparent N remobilization (%). This is defined as the proportion of seed N coming from vegetative storage at R5.

Stacking some of these physiological traits into a single cultivar requires understanding the nature and sign of the correlations among them. Correlation analysis among some of these traits (or other traits affecting yield) have been routinely assessed in the literature. In most cases, these correlation analyses violate assumptions of independence between the correlating variables due to shared terms (Brett, 2004). For example, correlating NUE and HI could produce spurious negative correlations because total biomass is the NUE numerator and the HI denominator. Also, a negative correlation is expected between total N uptake and NHI since the latter has total N uptake in the denominator. The problems associated with correlation analysis of traits sharing common terms have been analyzed by Donald and Hamblin (1976) when correlating total biomass with HI. However, straightforward correlations between traits sharing common terms (e.g., total N vs. NUE or total biomass vs. HI) are commonly reported in the literature. In our study we assessed whether correlations between meaningful physiological traits that share a common variable are genuine or spurious. This is a critical step to determine if a tradeoff between two physiological processes would hinder future trait stacking.

The present study has four objectives: (i) identify high-yielding cultivar clusters from the midwestern United States and central Argentinean Pampas and quantify physiological differences among clusters within regions; (ii) evaluate the importance of total N uptake in explaining yield differences among cultivars; (iii) quantify genotypic diversity in N uptake and N use traits within the highest-yielding clusters from USA and ARG to identify cultivars with high yield through different N mechanisms; (iv) analyze trait correlations, critical for defining trait-based hybridization strategies (Reynolds et al., 2011). In a previous paper, we demonstrated cultivar differences in strategies to attain high yield using a model explaining seed number determination (Rotundo et al., 2012). Here, a similar approach was followed using an N uptake and N use framework.

## MATERIALS AND METHODS

### Experimental Details

In the 2009 to 2010 and 2010 to 2011 crop seasons, 25 soybean cultivars (Supplemental Table S1a) from maturity groups IV to V were evaluated in a Vertic Argiudoll, Roldán series located in Zavalla (33°1' S, 60°53' W), Santa Fe, Argentina (Zav2010 and Zav2011, respectively). Planting dates were 27 Nov. 2009 and 16 Dec 2010. Plots had four rows, 5.5 m long, with 0.52 m row spacing. Stand density was 38 plants m<sup>-2</sup>. Weeds and pests were successfully controlled and no irrigation was applied. Precipitation during the growing seasons was 637 mm and 424 mm for Zav2010 and Zav2011, respectively.

In 2009, 65 cultivars (Supplemental Table S1b) from maturity groups II to III were evaluated at two locations, Agronomy Farm (42°1' N, 93°46' W) and Johnson Farm (41°58' N, 93°38' W), near Ames, Iowa, United States (Agro2009 and John2009, respectively). Soils were Clarion fine loam (Fine-loamy, mixed, mesic Typic Haplaquolls) at both sites. Planting date was 15 May 2009. The two sites had different crop rotations. At John2009 the previous crop was oats (*Avena sativa* L.) and at Agro2009 it was corn (*Zea mays* L.). Plots were similar to ARG but with 0.76-m row spacing. Stand density was 40 plants m<sup>-2</sup>. Weeds and pests were successfully controlled and no irrigation was applied. Precipitation during the growing season was 396 mm and was similar for both sites due to proximity. We never inoculated seeds with *Bradyrhizobium japonicum* since both ARG and USA regions have a long history of soybean production and residual bacteria (De Bruin et al., 2010). Nodulation was observed at both regions.

Soybean phenological stages were estimated three times per week, on a per-plot basis, scoring 20 consecutive individual plants each time (Fehr and Caviness, 1977). At both environments (ARG and USA) all measurements were taken in the two central rows.

## Direct Measurement Variables

A total of ten variables were directly measured. The abbreviation and description of these variables are presented in Table 1.

Total biomass at R1 (tbR1) was hand clipped from 0.5 m of two central rows (0.52 m<sup>-2</sup> in ARG, and 0.76 m<sup>-2</sup> in USA). Samples were dried at 60°C for at least 96 h, and weighed. Nitrogen concentration of this biomass (conc\_tbR1) was obtained by the Dumas method for total N on a ground subsample (Jung et al., 2003). In all biomass sampling there were 0.5-m border among clipping areas.

Biomass at R5 was also hand clipped as in R1. The entire sample was separated into pod (plus seeds) and non-pod fractions to measure reproductive (rbR5) and vegetative biomass (vbR5), respectively (Table 1). Nitrogen concentration was measured using Dumas method on each fraction (conc\_rbR5 and conc\_vbR5).

At R7, one meter of the two central rows was sampled (1.04 m<sup>-2</sup> in ARG, and 1.52 m<sup>-2</sup> in USA). Samples were dried at 60°C for at least 96 h, weighed and threshed. Non-seed biomass was saved. Seed biomass was subtracted from the whole sample to measure seed yield (syR7) and vegetative biomass (vbR7) at R7. Seeds and vegetative biomass were ground separately and analyzed for total N (seed N and conc\_vbR7). All variables, including yield, were expressed on a dry weight base.

## Calculated Physiological Traits

A total of eight physiological traits associated with N uptake and use were calculated from the variables described before (Table 1). Seed N was another physiological trait, but it was directly measured.

### Nitrogen Uptake Traits

Total N uptake at R7 was calculated as:

$$\text{Total N at R7 (kg ha}^{-1}\text{)} = \text{syR7} \left( \frac{\text{seed N}}{100} \right) + \text{vbR7} \left( \frac{\text{conc}_{\text{vbR7}}}{100} \right) \quad [4]$$

where syR7 is seed yield and vbR7 is vegetative biomass, both at physiological maturity. Seed N and conc\_vbR7 are seed and

**Table 1. Description of variables directly measured and the ones calculated from the former.**

Abbreviation	Description	Units
Variables measured R1, R5 and R7 <sup>†</sup>		
tbR1	Total biomass at beginning flowering	kg ha <sup>-1</sup>
conc_tbR1	Total nitrogen concentration at beginning flowering	0.1 g kg <sup>-1</sup>
rbR5	Reproductive (pod + seed) biomass at beginning seed filling	kg ha <sup>-1</sup>
vbR5	Vegetative (non-pod) biomass at beginning seed filling	kg ha <sup>-1</sup>
conc_rbR5	Reproductive nitrogen concentration at beginning seed filling	0.1 g kg <sup>-1</sup>
conc_vbR5	Vegetative biomass nitrogen concentration at beginning seed filling	0.1 g kg <sup>-1</sup>
syR7	Seed yield at physiological maturity	kg ha <sup>-1</sup>
vbR7	Vegetative (non-seed) biomass at physiological maturity	kg ha <sup>-1</sup>
seed N	Seed nitrogen concentration at physiological maturity	0.1 g kg <sup>-1</sup>
conc_vbR7	Vegetative biomass nitrogen concentration at physiological maturity	0.1 g kg <sup>-1</sup>
Calculated physiological traits		
Total N at R7	Total nitrogen uptake at physiological maturity	kg ha <sup>-1</sup>
N <sub>veg</sub>	Total nitrogen uptake between planting and flowering	kg ha <sup>-1</sup>
N <sub>R1-R5</sub>	Total nitrogen uptake between flowering and beginning seed filling	kg ha <sup>-1</sup>
N <sub>R5-R7</sub>	Total nitrogen uptake between beginning seed filling and maturity	kg ha <sup>-1</sup>
NUE	Nitrogen use efficiency for biomass production	kg kg <sup>-1</sup>
HI	Harvest index	kg kg <sup>-1</sup>
NHI	Nitrogen harvest index	kg kg <sup>-1</sup>
N <sub>remob</sub>	Apparent nitrogen remobilization	kg ha <sup>-1</sup>

<sup>†</sup>Fehr and Caviness (1977).

vegetative biomass N concentrations, respectively, at physiological maturity.

Nitrogen uptake during the vegetative growth stage was calculated as the N captured until beginning of flowering (R1) stage:

$$N_{\text{veg}} \text{ (kg ha}^{-1}\text{)} = \text{tbR1} \left( \frac{\text{conc}_{\text{tbR1}}}{100} \right) \quad [5]$$

where tbR1 is total biomass measured at R1 and conc\_tbR1 is the N concentration of this biomass.

Nitrogen uptake during the period for seed number determination (R1–R5) was estimated as the difference between total N at R5 minus total N at R1:

$$N_{\text{R1-R5}} \text{ (kg ha}^{-1}\text{)} = \left[ \text{vbR5} \left( \frac{\text{conc}_{\text{vbR5}}}{100} \right) + \text{rbR5} \left( \frac{\text{conc}_{\text{rbR5}}}{100} \right) \right] - \text{tbR1} \left( \frac{\text{conc}_{\text{tbR1}}}{100} \right) \quad [6]$$

where vbR5 and rbR5 are vegetative (non-pod) and reproductive (pods + seeds) biomass at R5, respectively. Conc\_vbR5 and conc\_rbR5 are the N concentrations of those fractions.

Nitrogen uptake during the seed-filling period was calculated as the difference between total N at physiological maturity (R7) minus total N at the beginning of seed filling (R5):

$$N_{R5-R7} \text{ (kg ha}^{-1}\text{)} = \left[ \text{vbR7} \left( \frac{\text{conc\_vbR7}}{100} \right) + \text{syR7} \left( \frac{\text{seed N}}{100} \right) \right] - \left[ \text{vbR5} \left( \frac{\text{conc\_vbR5}}{100} \right) + \text{rbR5} \left( \frac{\text{conc\_rbR5}}{100} \right) \right] \quad [7]$$

where syR7 is seed yield and vbR7 is vegetative biomass, both at physiological maturity. Seed N and conc\_vbR7 are seed and vegetative biomass N concentrations, respectively, at physiological maturity. VbR5 and rbR5 are vegetative (non-pod) and reproductive (pod + seed) biomass at R5. Conc\_vbR5 and conc\_rbR5 are the N concentrations of those fractions.

### Nitrogen Use Traits

Nitrogen use efficiency was defined as the ratio between total shoot biomass (kg ha<sup>-1</sup>) and total N (kg ha<sup>-1</sup>), both at R7:

$$\text{NUE (kg kg}^{-1}\text{)} = \frac{\text{syR7} + \text{vbR7}}{\text{syR7} \left( \frac{\text{seed N}}{100} \right) + \text{vbR7} \left( \frac{\text{conc\_vbR7}}{100} \right)} \quad [8]$$

where syR7 is seed yield and vbR7 is vegetative biomass, both at physiological maturity. Seed N and conc\_vbR7 are seed and vegetative biomass N concentrations, respectively, at physiological maturity.

Harvest index is not an N-use trait, but is relevant to understand yield differences in conceptual frameworks using N as the main driver for yield. It was estimated as the ratio between seed yield and total biomass shoot at R7:

$$\text{HI (kg kg}^{-1}\text{)} = \left( \frac{\text{syR7}}{\text{syR7} + \text{vbR7}} \right) \quad [9]$$

where syR7 is seed yield and vbR7 is vegetative biomass, both at physiological maturity.

Nitrogen harvest index was calculated as the ratio of seed N (kg ha<sup>-1</sup>) to total shoot N (kg ha<sup>-1</sup>) at R7:

$$\text{NHI (kg kg}^{-1}\text{)} = \frac{\text{syR7} \left( \frac{\text{seed N}}{100} \right)}{\text{syR7} \left( \frac{\text{seed N}}{100} \right) + \text{vbR7} \left( \frac{\text{conc\_vbR7}}{100} \right)} \quad [10]$$

Apparent N remobilization (N<sub>remob</sub> [kg N ha<sup>-1</sup>]) during the seed-filling period was calculated as vegetative non-pod N at R5 minus vegetative non-seed N at R7 (Zeihner et al., 1982):

$$N_{\text{remob}} \text{ (kg ha}^{-1}\text{)} = \left[ \text{vbR5} \left( \frac{\text{conc\_vbR5}}{100} \right) \right] - \left[ \text{vbR7} \left( \frac{\text{conc\_vbR7}}{100} \right) \right] \quad [11]$$

where vbR5 is vegetative non-pod biomass and conc\_vbR5 is the N concentration of that fraction; vbR7 is vegetative non-seed biomass at

physiological maturity and conc\_vbR7 is the N concentration of that fraction.

### Experimental Design and Statistical Analyses

We used a randomized complete block design with three replications in all trials. Sources of variation were environments (two planting dates in ARG, and two locations in USA), blocks, and cultivars.

Cluster analysis for seed yield was analyzed similar to Rotundo et al. (2012), following De la Vega et al. (2001) and Fox and Rosielle (1982). Criteria for cluster definition were at least 0.76 root square in all cases. Once clusters were defined, analysis of variance for seed yield, total biomass at R7, total N at R7, and other physiological traits was conducted including cluster category as another variation source. The final model included environment, cluster, and cultivars nested within clusters as fixed factors. Blocks (nested within environment) and all the interactions including blocks were considered random factors. After evaluating differences among clusters, we were specifically interested in residual cultivar variation within the highest-yielding clusters from ARG and USA. Therefore, we focused only on those cultivars for further analysis. In this case, the model included environment and cultivars as fixed factors and blocks (nested within environment) and all the interactions including blocks were considered random factors. For both analyses, we used the MIXED procedure of SAS (SAS Institute, 1999).

Pearson correlation analysis was conducted to analyze trait correlation within highest-yielding cultivars. Correlations among traits were only tested for those physiological traits showing significant differences within the highest-yielding cultivars (Table 5). A Monte Carlo simulation approach was used to test for spurious correlations between physiological traits sharing common direct measurement variables (Brett, 2004). This approach compares actual data from Pearson's coefficient against a null model generated using randomly-obtained variables instead of comparing the actual correlation with the R = 0 model (Brett, 2004). This approach has been successfully applied to understand other types of agricultural correlations, such as the tolerance to herbivory vs. potential growth in grasses (Rotundo and Aguiar, 2008). We generated random values for the directly-measured variables described in Table 1. The condition for these random values was to have the same number of observations, mean, and standard deviation as the real data set.

An example with the correlation between NUE (Eq. [8]) and HI (Eq. [9]) for 10 cultivars is described. First, we calculated the actual average data and standard deviation for syR7, vbR7, conc\_vbR7, and seed N. Second, because 10 cultivars are being correlated, 10 random values of these four variables were sampled from a normal distribution with mean and standard deviation similar to the actual data. With these variables, NUE and HI are then calculated for the 10 observations (cultivars). Finally, Pearson's correlation is conducted for these 10 randomly-generated observations and the Pearson's coefficient stored. This procedure is repeated 10,000 times. The 10th and 90th percentile interval of these Pearson's coefficients is reported as the expected value under random association between variables. The correlation is assumed to be spurious (originated from variables sharing common terms) if the correlation calculated with real data is within the 10th and 90th randomly generated

**Table 2. Analysis of variance for seed yield, total aboveground biomass, and nitrogen at physiological maturity for 25 cultivars in Argentina (ARG) and 65 cultivars in the United States (USA) evaluated in two environments within each region. Clusters were defined as a function of cultivar responses to the environment.**

Source	Seed yield	Total biomass at R7	Total nitrogen at R7
	kg ha <sup>-1</sup>	kg ha <sup>-1</sup>	kg N ha <sup>-1</sup>
<b>ARG</b>			
Environment	NS	NS	NS
Cluster	**	*	*
Env × Cluster	**	**	**
Cultivar <sub>(Cluster)</sub>	NS	NS	NS
Env × Cultivar <sub>(Cluster)</sub>	NS	NS	NS
<b>USA</b>			
Environment	*	**	**
Cluster	***	***	***
Env × Cluster	*	**	*
Cultivar <sub>(Cluster)</sub>	NS	*	NS
Env × Cultivar <sub>(Cluster)</sub>	NS	NS	NS

\* Significant at the 0.05 probability level.

\*\* Significant at the 0.01 probability level.

\*\*\* Significant at the 0.001 probability level.

interval (Brett, 2004). Monte Carlo simulations were conducted using R (R Development Core Team, 2011).

## RESULTS

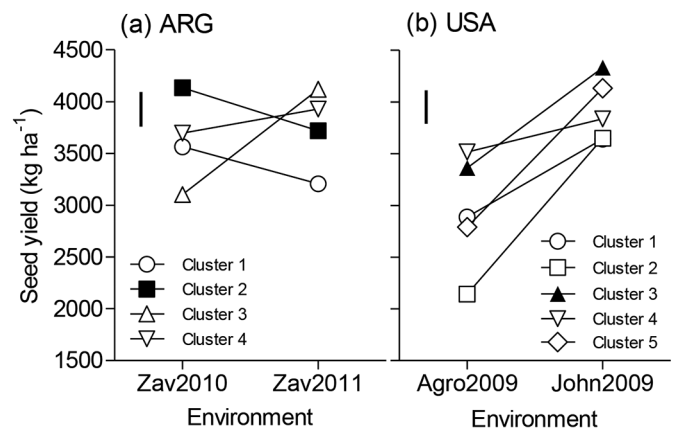
### Identification of Highest-yielding Clusters and Physiological Differences Among Clusters (Objective 1)

Average yields (dry weight basis) in ARG were 3470 and 3718 kg ha<sup>-1</sup> for Zav2009 and Zav2010, respectively. In USA, average yields were 2927 and 3884 kg ha<sup>-1</sup> for Agro2009 and John2009, respectively. Yield difference between USA environments was probably related to the different previous crop.

Cluster analysis of cultivar environmental responses identified four groups in ARG and five groups in USA with similar responses (data not shown). An  $R^2$  value of 0.76 was chosen as criterion to define clusters, so 76% of the original variation was retained by including cluster identification (instead of cultivars) as a source of variation. Analysis of variance for yield showed no significant cultivar effects or cultivar by environment interactions within clusters (Table 2), indicating clustering was appropriate.

There was a significant environment by cluster interaction for ARG and USA for yield (Fig. 1). In ARG, post-hoc comparison indicated Cluster 2 (7 cultivars) was the most successful in terms of seed yield across environments (Fig. 1a). In USA, Cluster 3 (12 cultivars) was identified as the most successful (Fig. 1b).

The four clusters identified in ARG differed significantly in  $N_{veg}$ , HI, seed N, and  $N_{remob}$  (Table 3). Clusters 2,



**Figure 1. Yield for (a) four groups of soybean cultivars from Argentina (ARG) and (b) five groups of soybean cultivars from the United States (USA). Cultivar groups were identified by cluster analysis in two environments. Zav2010 and Zav2011 stand for the two environments (years) evaluated in Zavalla, Argentina. Agro2009 and John2009 stand for the two environments (locations) evaluated in the USA. Bar shows the least significant difference (LSD;  $p < 0.05$ ) for comparison.**

3, and 4 had similar  $N_{veg}$ , which were larger than the value registered for Cluster 1. Harvest index and  $N_{remob}$  was highest for Cluster 2 (the highest-yielding cluster) compared with Clusters 3 and 1, respectively (Table 3). Seed N was among the lowest values for Cluster 2 compared with Clusters 1 and 4 (Table 3). In USA (Table 4), clusters differed significantly for all the traits. Cluster 3 (the highest-yielding cluster) had intermediate  $N_{veg}$ ,  $N_{R1-R5}$ , NUE, and  $N_{remob}$  compared with other clusters (Table 4). Cluster 3 presented the lowest seed N among the five clusters for USA.

### Total N Uptake vs. Total Biomass Explaining Yield Differences (Objective 2)

For both ARG and USA, total N uptake and biomass were highly correlated with seed yield (Fig. 2). The highest-yielding cultivars were always the ones accumulating more biomass and showing higher N uptake at maturity. There were differences in the yield variation that total N uptake and total biomass explained. For ARG, total biomass explained 68% yield variation while total N uptake explained 82% (Fig. 2a,b). In USA, biomass explained 87% yield variation while total N uptake explained 93% (Fig. 2c,d). At both ARG and USA, cultivar and environmental variation in yield was explained more by total N uptake than by total biomass.

### Genotypic Variation in N Uptake and N-use Traits for the Highest-yielding Cultivars (Objective 3)

#### Nitrogen Uptake

There were no differences in total N uptake among the highest-yielding cultivars from ARG (Cluster 2) and USA (Cluster 3, Table 5). Average N uptake for ARG and USA high-yielding clusters was similar (264 and 263 kg ha<sup>-1</sup> at

**Table 3. Average physiological trait values for different cultivar clusters from Argentina having contrasting responses to the environment. Cluster 2 (bold) was identified as the most successful in both environments. Zav2010 and Zav2011 stand for the two environments (years) evaluated in Zavalla, Argentina. Data are presented for the interaction between cluster and environment (whenever this was significant) or for the cluster main effect (if the interaction was not significant).**

Environment	Cluster	$N_{veg}^{\dagger}$	$N_{R1-N5}$	$N_{R5-R7}$	NUE	HI	NHI	Seed N	$N_{remob}$
			kg ha <sup>-1</sup>			kg kg <sup>-1</sup>		0.1 g kg <sup>-1</sup>	kg ha <sup>-1</sup>
Zav2010	Cluster 1			76.4					
	<b>Cluster 2</b>			82.6					
	Cluster 3			43.0					
	Cluster 4			74.3					
Zav2011	Cluster 1			24.8					
	<b>Cluster 2</b>			30.7					
	Cluster 3			50.6					
	Cluster 4			48.2					
Average	Cluster 1	56.2	132.4		32.6	0.444	0.836	5.80	148.5
	<b>Cluster 2</b>	67.9	139.8		32.3	0.468	0.848	5.71	166.6
	Cluster 3	66.5	146.5		32.4	0.439	0.813	5.70	160.8
	Cluster 4	68.5	133.5		31.8	0.461	0.842	5.79	158.4
Environment		***	**	**	*	NS	NS	**	NS
Cluster		***	NS	NS	NS	*	NS	**	*
Env × Cluster		NS	NS	*	NS	NS	NS	NS	NS
LSD <sup>‡</sup> ( <i>P</i> < 0.05)		4.3	–	20	–	0.026	–	0.06	11.8

\* Significant at the 0.05 probability level.

\*\* Significant at the 0.01 probability level.

\*\*\* Significant at the 0.001 probability level.

<sup>†</sup> HI, harvest index;  $N_{remob}$ , apparent nitrogen remobilization;  $N_{R1-R5}$ , total nitrogen uptake between flowering and beginning seed filling;  $N_{R5-R7}$ , total nitrogen uptake between beginning seed filling and maturity;  $N_{veg}$ , total nitrogen uptake between planting and flowering; NHI, nitrogen harvest index; NUE, nitrogen use efficiency for biomass production; Seed N, seed nitrogen concentration at physiological maturity.

<sup>‡</sup> LSD for significant Cluster or Environment by Cluster effects.

**Table 4. Average physiological trait values for different U.S. cultivar clusters having contrasting responses to the environment. Cluster 3 (bold) was identified as the most successful in both environments. Agro2009 and John2009 stand for the two environments (locations) evaluated in the USA. Data are presented for the interaction between cluster and environment or for the cluster main effect.**

Environment	Cluster	$N_{veg}^{\dagger}$	$N_{R1-N5}$	$N_{R5-R7}$	NUE	HI	NHI	Seed N	$N_{remob}$
			kg ha <sup>-1</sup>			kg kg <sup>-1</sup>			kg ha <sup>-1</sup>
Agro2009	Cluster 1			51.7		0.481	0.861		
	Cluster 2			31.5		0.426	0.817		
	<b>Cluster 3</b>			57.3		0.493	0.873		
	Cluster 4			66.6		0.488	0.867		
	Cluster 5			42.8		0.467	0.853		
John2009	Cluster 1			67.1		0.513	0.884		
	Cluster 2			89.1		0.501	0.867		
	<b>Cluster 3</b>			93.7		0.515	0.886		
	Cluster 4			59.8		0.502	0.879		
	Cluster 5			92.4		0.513	0.882		
Average	Cluster 1	44.7	125.9		28.7			6.11	141.0
	Cluster 2	42.5	109.0		30.2			6.10	115.9
	<b>Cluster 3</b>	47.8	138.5		29.2			5.97	154.5
	Cluster 4	53.1	142.5		29.0			6.08	160.7
	Cluster 5	49.3	130.1		28.8			6.18	146.0
Environment		**	**	**	**	**	**	NS	**
Cluster		***	***	*	*	***	***	***	***
Env × Cluster		NS	NS	**	NS	**	*	NS	NS
LSD <sup>‡</sup> ( <i>P</i> < 0.05)		2.1	9.4	14.9	1.0	0.012	0.014	0.07	8.5

\* Significant at the 0.05 probability level.

\*\* Significant at the 0.01 probability level.

\*\*\* Significant at the 0.001 probability level.

<sup>†</sup> HI, harvest index;  $N_{remob}$ , apparent nitrogen remobilization;  $N_{R1-R5}$ , total nitrogen uptake between flowering and beginning seed filling;  $N_{R5-R7}$ , total nitrogen uptake between beginning seed filling and maturity;  $N_{veg}$ , total nitrogen uptake between planting and flowering; NHI, nitrogen harvest index; NUE, nitrogen use efficiency for biomass production; Seed N, seed nitrogen concentration at physiological maturity.

<sup>‡</sup> LSD for significant Cluster or Environment by Cluster effects.

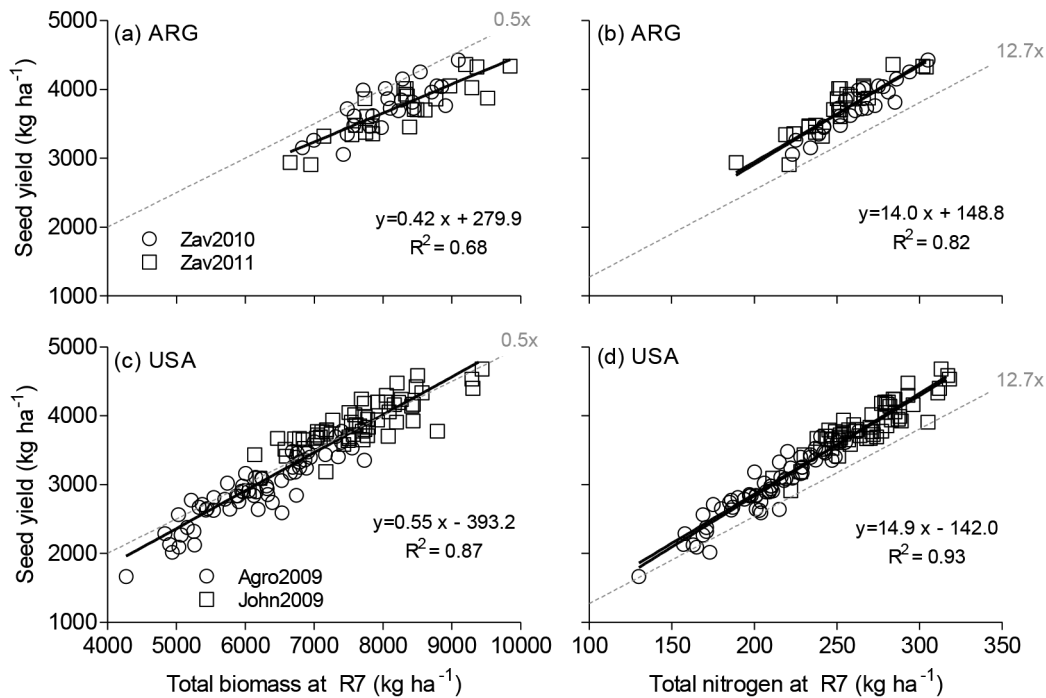


Figure 2. Relationship between seed yield and (a, c) total biomass and (b, d) total nitrogen at physiological maturity for Argentina (ARG) (a, b) and the USA (c, d). Twenty-five cultivars were evaluated in ARG and 65 in the USA. Zav2010 and Zav2011 stand for the two environments (years) evaluated in Zavalla, Argentina. Agro2009 and John2009 stand for the two environments (locations) evaluated in the USA. Each data point is the average of three replications at one environment. Gray dotted lines indicate 0.50 harvest index (a, c) and the average seed yield per total N uptake at R7 (b, d) as calculated by Salvagiotti et al. (2008).

R7 for ARG and USA, respectively). However, there were cultivar differences in the temporal pattern of N uptake among genotypes within the highest-yielding clusters at both ARG and USA (Fig. 3). Figure 3 shows cultivar data averaged across environments in ARG and USA since no significant interaction ( $p > 0.05$ ) between cultivars and environments was detected.

Relating N uptake during different phenological periods and the fraction of time (days) allotted to those periods allowed identification of cultivars having different N uptake per unit time at both ARG and USA. For example, cultivar NA5485 (Fig. 3a, ID no. 1) allocated 47% of the total cycle from planting to R1 and accumulated 47% of total N uptake during this period. However, cultivar DM4210 (Fig. 3a, ID no. 6) had 37% of the total cycle allocated to this period while N uptake was <19% during this period. In contrast to ARG, there were no significant differences in N uptake during the vegetative period for cultivars in USA Cluster 3 (Fig. 3a). On average for USA cultivars, 18% of N uptake occurred during the vegetative period while this period accounted for approximately 43% of the total cycle.

Most N uptake occurred between R1 and R5 (cultivars above the 2:1 line, Fig. 3b). There was significant variation in N uptake per unit duration of the R1 to R5 period. For example, the relative duration of the period from R1 to R5 for cultivar DM4210 (Fig. 3b, ID no. 6) was 25% of the total cycle and the proportion of the total

N uptake during this period was 50%; cultivar DM4970 (Fig. 3b, ID no. 2) had a similar R1 to R5 period duration, but the proportion of total N uptake that was accumulated during this period was more than 60%. Similar differences can be observed for USA when cultivars K285RR (ID no. 17) and ML2666 (ID no. 8) are compared (Fig. 3b).

Cluster 3 USA cultivars allotting 30% of the cycle to seed filling (R5 to R7) captured approximately 20 to 40% of total N, with most cultivars being close to the 1:1 line (Fig. 3.c). Conversely, mean values for the ARG Cluster 2 were clearly beneath the 1:1 line (35% for duration and 21% for N uptake), although some cultivars of this group were closer to this line than others.

### Nitrogen Use

There was significant cultivar variation in NUE within both ARG Cluster 2 and USA Cluster 3 (Table 5). In ARG, only one cultivar (NA5485) differed from all other cultivars. In USA, there was a continuum from 27.5 to 32 kg biomass kg N<sup>-1</sup>. Harvest index showed significant variation within the two highest-yielding clusters from ARG and USA, although the variation was rather low. For ARG, HI ranged from 0.44 to 0.49 kg kg<sup>-1</sup> while for USA variation ranged from 0.48 to 0.53 kg kg<sup>-1</sup>.

Nitrogen harvest index differed among the cultivars from the highest-yielding clusters from ARG and USA, but the variation was rather low. ARG cultivars from Cluster 2 ranged from 0.82 to 0.87 kg N kg N<sup>-1</sup> and

**Table 5. Average values and analysis of variance for yield physiological determinants for most successful cultivars from Argentina (ARG) (Cluster 2) and the United States USA (Cluster 3).**

ID no.) cultivars	Total N at R7	Nitrogen use efficiency	Harvest index	Nitrogen harvest index	N remobilization R5–R7	Seed N concentration
	kg N ha <sup>-1</sup>	kg kg N <sup>-1</sup>	kg kg <sup>-1</sup>	kg N kg N <sup>-1</sup>	kg N ha <sup>-1</sup>	g 0.1 kg <sup>-1</sup>
1) NA5485RG	256	35.0a <sup>†</sup>	0.44c	0.84bc	192.4a	5.74
2) DM4970	267	32.7b	0.45bc	0.83c	182.9ab	5.60
3) A4725RG	253	32.1b	0.48a	0.86ab	167.6ab	5.70
4) DM4250	280	31.4b	0.47ab	0.86ab	167.6abc	5.84
5) DM4670	271	31.1b	0.47ab	0.82c	162.3bc	5.65
6) DM4210	269	31.2b	0.49a	0.87a	153.5c	5.71
7) NK4300	254	32.3b	0.48a	0.86ab	139.8c	5.77
Environment	*	**	*	NS	NS	**
Cultivar <sub>(Cluster2)</sub>	NS	*	*	*	*	NS
E × C	NS	NS	NS	NS	NS	NS
LSD ( <i>P</i> < 0.05)	-	2.6	0.03	0.03	29.4	-
8) ML2666	263	28.7bc	0.51ab	0.87bc	185.3a	5.95bc
9) NKS30J8	257	28.8bc	0.49bc	0.87bc	178.5ab	6.15ab
10) P92M76	266	27.5c	0.52a	0.90a	166.6abc	6.27a
11) K283RR	259	29.4bc	0.50bc	0.87bc	159.7abc	5.93bc
12) K204RR	254	29.0bc	0.51ab	0.89ab	159.1abc	5.93bc
13) NKS29J6	257	27.9c	0.53a	0.90a	154.2abc	6.02abc
14) K275RR	283	30.4ab	0.48c	0.89ab	152.0abc	5.95bc
15) DKB2752	252	30.0ab	0.48c	0.86c	149.1abc	5.85bc
16) Jack	282	28.7bc	0.50bc	0.87bc	142.0c	6.12ab
17) K285RR	257	29.3bc	0.53a	0.90a	140.1c	5.79c
18) S29324	259	28.8bc	0.51ab	0.87bc	133.3c	5.99abc
19) AG2802	259	32.0a	0.48c	0.86c	131.2c	5.72c
Environment	***	*	*	NS	NS	NS
Cultivar <sub>(Cluster3)</sub>	NS	*	**	*	*	*
E × C	NS	NS	NS	NS	NS	NS
LSD ( <i>P</i> < 0.05)	-	2.2	0.03	0.03	35.9	0.31

\* Significant at the 0.05 probability level.

\*\* Significant at the 0.01 probability level.

\*\*\* Significant at the 0.001 probability level.

† Within columns, means followed by the same letter are not significantly different according to LSD (0.05) for cultivar main effect.

cultivars from the USA Cluster 3 ranged from 0.86 to 0.90 kg N kg N<sup>-1</sup> (Table 5).

Nitrogen remobilization during seed filling also differed among the high-yielding cultivars in ARG and USA. In ARG, remobilization ranged from 140 to 192 kg N ha<sup>-1</sup>, while in USA it ranged from 131 to 185 kg N ha<sup>-1</sup>.

Seed N concentration was not different across the highest-yielding cultivars from Cluster 2 in ARG. The average for this cluster was 5.7% grain N (equivalent to 30% protein on 13% moisture basis). However, differences were observed within the highest-yielding cultivars in USA Cluster 3. Seed N concentration for cultivars in USA ranged from 5.7 to 6.3% grain N (30 to 34% protein on 13% moisture basis; Table 5).

### Correlation among Traits within the Highest-Yielding Clusters from ARG and USA (Objective 4)

We also analyzed trait correlations because they are critical for trait-based hybridization strategies. Because many

traits share common measurements, we tested for spurious correlations. All correlations are described in Table 6.

For ARG Cluster 2 cultivars, NUE was negatively correlated with N uptake<sub>R5–R7</sub> and HI according to F-test significance (Table 6). However, Pearson's coefficient of correlation between NUE and HI (−0.89) was within the 10th and 90th percentile of randomly-generated Pearson's coefficients, indicating the detected correlation between those traits was spurious. Nitrogen uptake during the vegetative period (N<sub>veg</sub>) was negatively correlated (F-test) with N<sub>R5–R7</sub> (−0.85) and HI (−0.92). These values of Pearson's coefficient were outside the 10th to 90th random coefficient interval, suggesting genuine negative associations between these traits. N<sub>R5–R7</sub> was inversely correlated with N remobilization, and the last one was negatively correlated with HI. In both cases there was no evidence of spurious correlation. Positive, non-spurious correlations were observed between: (i) NUE vs. N<sub>veg</sub>, (ii) N<sub>veg</sub> vs. N remobilized, and (iii) N<sub>R5–R7</sub> vs. HI (Table 5).



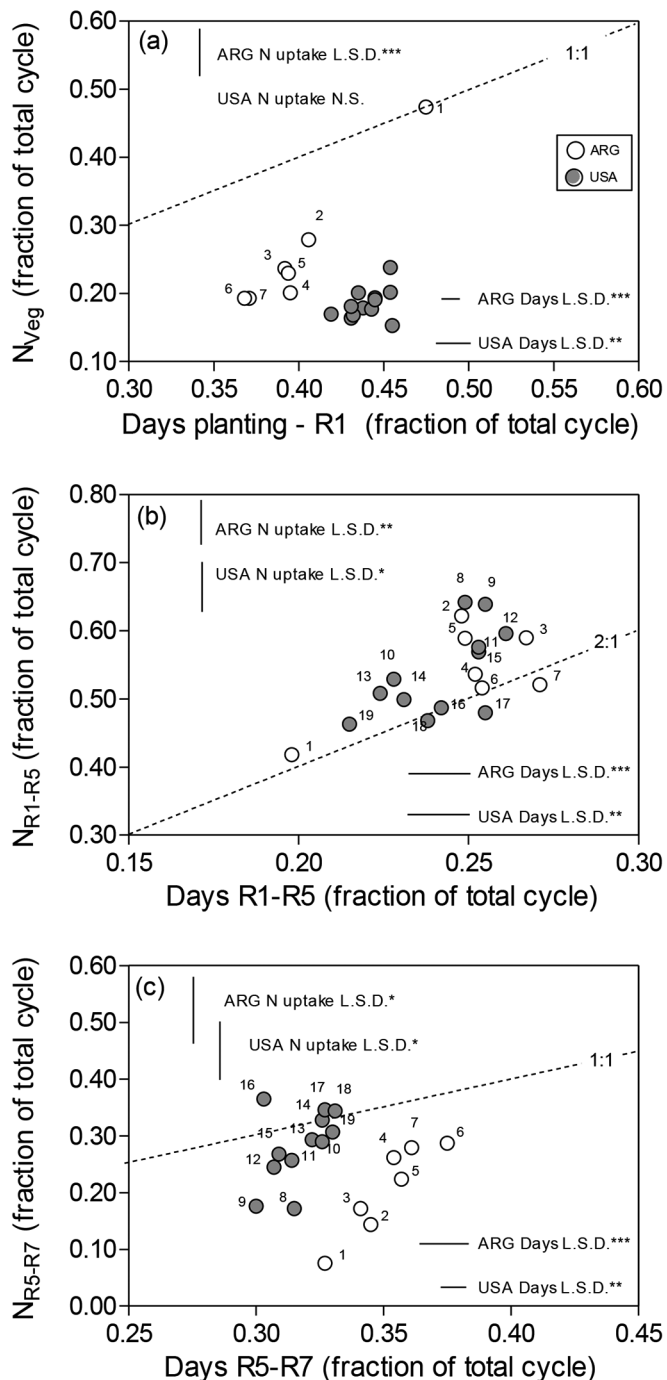


Figure 3. Relationship between the proportion (%) of total aboveground N uptake at R7 captured during the vegetative period ( $N_{veg}$ ), the seed-set period ( $N_{R1-R5}$ ) and the seed-filling period ( $N_{R5-R7}$ ) and the proportion (%) of time each period has relative to the total crop-cycle duration. Empty circles indicate cultivars from Argentina (ARG) Cluster 2 and grey circles indicate the USA cultivars from Cluster 3. Gray lines show 1:1 relation (a, c) or 2:1 relation (b). Vertical and horizontal lines show LSD ( $p < 0.05$ ). Number indicates ID no. as presented in Table 5.

For USA Cluster 3 cultivars, NUE was negatively correlated with HI and seed N (Table 6). As in ARG, Pearson's coefficient of NUE and HI was contained by the randomly-generated interval suggesting a spurious correlation. However, there is no evidence of spurious

correlation between NUE and seed N.  $N_{R1-R5}$  was negatively correlated with  $N_{R5-R7}$ , but this correlation was spurious. Likewise,  $N_{R5-R7}$  was negatively correlated with N remobilized, but Pearson's coefficient was included in the randomly-generated interval.

## DISCUSSION

Total N uptake is calculated after measuring total biomass and N concentration. Therefore, it has two error sources (estimation of biomass and N), unlike total biomass, which has only one. We expected yield variation to be more correlated with total biomass than with total N uptake. However, we verified the opposite result; N uptake was more related to yield across cultivars and environments than biomass in ARG and USA. Similarly, Pazdernik et al. (1997) showed that N uptake explained more yield variation than total plant biomass ( $R^2 = 0.55$  vs. 0.44 for N uptake and total biomass, respectively) when comparing 20 different cultivars from maturity groups 0 and I. Our results help support the hypothesis by Sinclair and Jamieson (2006) stating total N is likely to explain a higher proportion of yield variations than attributes related to carbohydrate availability since the former supports many functions in the plant, such as leaf area expansion, carbon fixation, or storage for translocation.

The slope of the linear relationship between yield and N uptake, when forcing y-intercept to zero, was 14.6 kg seed per kg N uptake for ARG and 14.3 kg seed per kg N uptake for USA. These estimates of NUE at the seed level (Good et al., 2004) are close to values reported by Salvagiotti et al. (2008) after a comprehensive review of published literature. Maximum and minimum N dilution is 18.8 and 8.4 kg seed per kg N uptake, respectively (Salvagiotti et al., 2008). The values for ARG and USA (average 14.5 kg seed per kg N uptake) suggest cultivars in both environments were equally balanced in terms of N and other nutrients (Janssen et al., 1990; Witt et al., 1999).

Cultivar clusters having maximum yield were identified both in ARG and USA. No significant differences in total N uptake were observed among cultivars within both clusters since total N uptake and yield were highly correlated. On average for ARG and USA, maximum yield was attained with  $\sim 260$  kg N  $ha^{-1}$ . Cultivars differed, however, in their temporal pattern of N uptake. Highest rates of N uptake were observed during the period encompassing R1 to R5, in agreement with earlier reports (George and Singleton, 1992; Zapata et al., 1987). Results from our experiment identified cultivars having increased rates of total N uptake during different developmental stages.

Whether differences among cultivars in N uptake or timing are due to better absorption of mineral soil N or different N fixation warrants future investigations. We speculate improved N uptake may be due to improved N

**Table 6. Pearson coefficients between traits explaining yield variation across most-successful cultivars from Argentina (ARG) (Cluster 2) and the United States (USA) (Cluster 3). Data within parentheses is 10th to 90th percentile interval for Pearson's coefficient generated by Monte Carlo random simulations. Only the traits with significant ( $p < 0.05$ ) variation within the highest-yielding cluster were tested for correlation studies. Bold indicates significant (F-test) non-spurious correlations.**

(a) ARG Cluster 2	$N_{veg}^\dagger$	$N_{R1-R5}$	$N_{R5-R7}$	HI	NHI	$N_{remob}$
NUE	<b>0.90<sup>††</sup></b> (-0.56 to +0.55)	-0.64 (-0.55 to +0.54)	<b>-0.83*</b> (-0.69 to +0.38)	-0.89** (-0.95 to -0.59)	-0.22 (-0.66 to +0.41)	0.67 (-0.55 to +0.55)
$N_{veg}$		-0.50 (-0.96 to -0.61)	<b>-0.85*</b> (-0.55 to +0.55)	<b>-0.92**</b> (-0.55 to +0.55)	-0.43 (-0.54 to +0.54)	<b>0.81*</b> (-0.56 to +0.55)
$N_{R1-R5}$			0.20 (-0.84 to +0.02)	0.28 (-0.55 to +0.55)	-0.32 (-0.54 to +0.54)	-0.01 (0.00 to +0.84)
$N_{R5-R7}$				<b>0.89**</b> (-0.46 to +0.62)	0.49 (-0.61 to +0.49)	<b>-0.88**</b> (-0.55 to +0.54)
HI					0.54 (+0.10 to +0.87)	<b>-0.85*</b> (-0.47 to +0.62)
NHI						-0.43 (-0.37 to +0.61)
(b) USA Cluster 3	$N_{R1-R5}$	$N_{R5-R7}$	HI	NHI	$N_{remob}$	Seed N <sup>-1</sup>
NUE	-0.33 (-0.39 to +0.39)	0.29 (-0.40 to +0.39)	-0.67* (-0.93 to -0.67)	-0.55 (-0.80 to -0.24)	-0.48 (-0.50 to +0.29)	<b>-0.78**</b> (-0.76 to -0.18)
$N_{R1-R5}$		-0.80** (-0.85 to -0.41)	-0.19 (-0.39 to +0.40)	-0.16 (-0.40 to -0.40)	0.90*** (+0.61 to +0.91)	0.40 (-0.40 to +0.40)
$N_{R5-R7}$			-0.01 (-0.34 to +0.44)	0.14 (-0.39 to +0.41)	-0.86*** (-0.88 to -0.50)	-0.18 (-0.03 to +0.67)
HI				0.78** (+0.55 to +0.88)	0.06 (-0.20 to +0.57)	0.19 (-0.40 to +0.40)
NHI					0.12 (-0.10 to +0.62)	0.35 (-0.22 to +0.56)
$N_{remob}$						0.49 (-0.39 to +0.40)

\* Significant at the 0.05 probability level for standard Pearson's correlation analysis.

\*\* Significant at the 0.01 probability level for standard Pearson's correlation analysis.

\*\*\* Significant at the 0.001 probability level for standard Pearson's correlation analysis.

† HI, harvest index;  $N_{remob}$ ; apparent nitrogen remobilization;  $N_{R1-R5}$ , total nitrogen uptake between flowering and beginning seed filling;  $N_{R5-R7}$ , total nitrogen uptake between beginning seed filling and maturity;  $N_{veg}$ , total nitrogen uptake between planting and flowering; NHI, nitrogen harvest index; NUE, nitrogen use efficiency for biomass production; Seed N; seed nitrogen concentration at physiological maturity.

fixation since total N uptake has been correlated with N fixation rate (George and Singleton, 1992).

Two different N physiological frameworks were used to explain yield differences (Eq. [2] and [3]; Xu et al., 2012). Both frameworks utilize total N uptake as their main driver for yield variations. However, we found no differences among cultivars within highest-yielding clusters for ARG and USA for this trait, demonstrating the importance of N uptake for maximum yield. One of the proposed frameworks (Eq. [2]) affects total N uptake by NUE and HI. Using this framework, we found cultivars with contrasting strategies attained higher yield; similar yield was produced by different combinations of NUE and HI. Retrospective studies analyzing release year effects on soybean traits showed significant genetic improvement in HI (Jin et al., 2010; Morrison et al., 1999). Even though this trait has been improved, variation to continue improving HI was evident in the highest-yielding clusters from ARG and USA. To the best of our knowledge, studies dealing with the indirect effects of genetic improvement on yield in NUE have never been reported for soybeans.

The second studied framework (Eq. [3]) affects total N uptake by NHI and seed N. For highest-yielding cultivars in ARG, there was no genotypic variation in seed N, while variation in NHI was significant. Within the USA, highest-yielding cluster NHI and seed N varied significantly across cultivars. This variation allowed similar final yields, but it was attained by using contrasting methods in terms of NHI and seed N. Again, a retrospective study showed there has been substantial genetic improvement in NHI (Jin et al., 2011). Our results demonstrate residual variation for NHI, even in highly-productive cultivars. Increased seed N requires higher source-to-sink ratio during seed filling (Rotundo et al., 2009, 2011), and increased source-to-sink ratio is generally attained by reducing sink numbers rather than increasing source area (Rotundo et al., 2011). In cultivars from USA Cluster 3, any potential yield reduction due to increased seed N was counterbalanced by increases in NHI, resulting in no yield changes.

There was substantial genotypic variation within both ARG and USA highest-yielding clusters for apparent N remobilization during seed filling. Nitrogen remobilization

ranged from 30 to 100% depending on cultivars and environmental conditions (Egli et al., 1978, 1983; Zeiher et al., 1982). Even though this trait does not fit in any of the above-mentioned physiological frameworks, it has central importance in relation to the duration of seed filling (Sinclair and de Wit, 1975, 1976). Our results indicated that high-yielding cultivars from both ARG and USA had substantial genotypic diversity for this trait. For example, ARG Cluster 2 had cultivars with remobilization values from 140 to 192 kg N ha<sup>-1</sup>, while USA Cluster 3 had cultivars ranging from 131 to 185 kg N ha<sup>-1</sup>. Our results indicate that there are different ways to accumulate seed N within the highest-yielding genotypes: (i) high remobilization or (ii) high concurrent N capture during seed filling.

Even though variation in other traits besides N uptake timing was significant, the range of variation was, in some cases, lower than expected, as in NUE. One possibility is that we focused only on elite high-yielding cultivars, having reduced genetic diversity as a result of strong breeding efforts (Hyten et al., 2006). Another possibility, related to NUE, is that variation in this trait is expressed at low soil-N conditions, as shown for maize (*Zea mays* ssp. *mays* L.) and sugarcane (*Sacharum officinarum* L.) (Hirel et al., 2007; Robinson et al., 2007). In soybean, the relative independence from soil N levels impedes the generation of low N availability environments.

Rotundo et al. (2012) also found that cultivars with the highest seed number had different ways to achieve it. They had different combinations of seed-set period duration, biomass partitioning and, seed-set efficiency (Rotundo et al., 2012). The present article expands those findings utilizing a physiological framework related to N uptake and use.

Results from the current paper provide valuable information to develop trait-based hybridization strategies (Reynolds et al., 2011). If we consider traits to be independent, combining the average N uptake from the highest-yielding cultivars with the highest observed values for NUE and HI gives a hypothetical cultivar (for both ARG and USA data) having ~4400 kg ha<sup>-1</sup> (dry weight basis). This is the attainable yield for this hypothetical cultivar grown in the rainfed conditions we tested. Clusters 2 (ARG) and 3 (USA) averaged ~3900 kg ha<sup>-1</sup> (dry weight basis). Combining these traits represents a 13% yield improvement compared with current cultivars. In general, using traditional breeding, the historical genetic improvement rate of soybean yield is, on average, 0.5% per year (Jin et al., 2010; Morrison et al., 2000; Wilcox, 2001). Therefore, under the same scenario, this 13% hypothetical yield increase would take ~18 yr to occur. Crossing parents having contrasting physiological strategies may help reduce this time.

The possibility of pyramiding desirable traits will depend on the nature of the correlation between them. Negative correlations (or tradeoff) may arise from three different nonexclusive causes. The worst scenario for trait stacking is biophysical constraints (Weih, 2003). Another possible cause

of the negative correlations is genetic linkages (Weih, 2003). This negative correlation may be eliminated by recombination using large progeny numbers in bi-parental crosses. Finally, a negative correlation may arise just because there is a numerical co-dependence between the variables being correlated (Brett, 2004; Donald and Hamblin, 1976); that is, the variables being correlated share a common term. This is frequent in crop physiology, where simple directly-measured traits are mathematically combined to express meaningful physiological mechanisms or processes. We found that some negative (and significant according to F-tests) correlations are spurious. For example, NUE and HI were negatively correlated both in ARG and USA. However, because they share total biomass as a common term (NUE in the numerator and HI in the denominator), we observed actual coefficients were contained in the randomly-generated interval. Therefore, the correlation was spurious and combining both traits in the same cultivar does not represent a constraint since the correlation is not biological (Brett, 2004).

## CONCLUSIONS

Cultivars with superior yield were identified from ARG and USA. Total N uptake was the primary yield driver shown by the highest-yielding commercial cultivars at both environments. It is clear that attempts to increase yield will need to focus on mechanisms associated with this process.

There were clear developmental differences among the highest-yielding clusters for N uptake both in Argentina and USA. Some high-yielding cultivars had higher N uptake during vegetative (planting to R1) and late reproductive periods (R5 to R7), while others relied mostly on N uptake during the seed-set period (R1–R5).

Highest-yielding cultivars differed in the physiological strategies to attain maximum yield; combining high NUE and HI traits or, alternatively, high NHI and similar seed N can result in major yield increases.

Understanding the nature of the correlations among physiological traits is central to this trait-based hybridization approach. Sorting out whether tradeoffs are physiological, genetic, or spurious is critical to quantifying the chances of trait stacking. For ARG, there were six significant negative correlations between traits. In this case, only one (NUE vs. HI) was spurious. In USA, three out of the four significant negative correlations were spurious. This indicates more chances of trait stacking in USA high-yielding cultivars than in ARG.

## Supplemental Information Available

Supplemental information is included with this article.

## Acknowledgments

This research was supported in part by Agencia Nacional de Investigaciones (Prestamo BID PICT-2011-1292). JLR and LB are members of CONICET, the Argentine National Research Council.

## References

- Bohn, M., H.F. Utz, and A.E. Melchinger. 1999. Genetic similarities among winter wheat cultivars determined on the basis of RFLPS, AFLPS, and SSRs and their use for predicting progeny variance. *Crop Sci.* 39:228–237. doi:10.2135/cropsci1999.0011183X003900010035x
- Brett, M.T. 2004. When is a correlation between non-independent variables “spurious”? *Oikos* 105:647–656. doi:10.1111/j.0030-1299.2004.12777.x
- Brown-Guedira, J.A.T., R.L. Nelson, and M.L. Warburton. 2000. Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Sci.* 40:815–823. doi:10.2135/cropsci2000.403815x
- Burkhamer, R.L., S.P. Lanning, R.J. Martens, J.M. Martin, and L.E. Talbert. 1998. Predicting progeny variance from parental divergence in hard red spring wheat. *Crop Sci.* 38:243–248. doi:10.2135/cropsci1998.0011183X003800010041x
- Carpentieri-Pípolo, V., A.E. Pípolo, F.A. Martins da Silva, and M.R. Petek. 2000. Soybean parent selection based on genetic diversity. *Braz. Arch. Biol. Technol.* 43:295–300. doi:10.1590/S1516-89132000000300008
- Cooper, M., and G.L. Hammer. 2005. Preface to Special Issue: Complex traits and plant breeding—can we understand the complexities of gene-to-phenotype relationships and use such knowledge to enhance plant breeding outcomes? *Aust. J. Agric. Res.* 56:869–872. doi:10.1071/AR05151
- Cregan, P.B., and P. Vanberkum. 1984. Genetics of nitrogen-metabolism and physiological biochemical selection for increased grain crop productivity. *Theor. Appl. Genet.* 67:97–111. doi:10.1007/BF00317013
- De Bruin, J.L., P. Pedersen, S.P. Conley, J.M. Gaska, S.L. Naeve, J.E. Kurle, R.W. Elmore, L.J. Giesler, and L.J. Abendroth. 2010. Probability of yield response to inoculants in fields with a history of soybean. *Crop Sci.* 50:265–272. doi:10.2135/cropsci2009.04.0185
- De la Vega, A.J., S.C. Chapman, and A.J. Hall. 2001. Genotype by environment interaction and indirect selection for yield in sunflower: I. Two-mode pattern analysis of oil and biomass yield across environments in Argentina. *Field Crops Res.* 72:17–38. doi:10.1016/S0378-4290(01)00162-9
- Donald, C.M., and J. Hamblin. 1976. Biological yield and harvest index of cereals as agronomic and plant breeding criteria. *Adv. Agron.* 28:361–405. doi:10.1016/S0065-2113(08)60559-3
- Duvick, D.N. 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). *Adv. Agron.* 86:83–145. doi:10.1016/S0065-2113(05)86002-X
- Egli, D.B. 2008. Comparison of corn and soybean yields in the United States: Historical trends and future prospects. *Agron. J.* 100:S79–S88. doi:10.2134/agronj2006.0286c
- Egli, D.B., J.E. Leggett, and W.G. Duncan. 1978. Influence of N-stress on leaf senescence and N redistribution in soybeans. *Agron. J.* 70:43–47. doi:10.2134/agronj1978.00021962007000010011x
- Egli, D.B., L. Meckel, R.E. Phillips, D. Radcliffe, and J.E. Leggett. 1983. Moisture stress and N-redistribution in soybean. *Agron. J.* 75:1027–1031. doi:10.2134/agronj1983.00021962007500060036x
- Falconer, D.S., and T.F.C. Mackay. 1960. Introduction to quantitative genetics. Longman, Essex, UK.
- Fehr, W.R. 1998. Principles of cultivar development. MacMillan Publishing Company,
- Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. Iowa Agricultural and Home Economics Experiment Station Special Report, Ames Iowa.p. 3–11.
- Fox, P.N., and A.A. Rosielle. 1982. Reducing the influence of environmental main-effects on patterns on pattern analysis of plant breeding environments. *Euphytica* 31:645–656. doi:10.1007/BF00039203
- George, T., and P.W. Singleton. 1992. Nitrogen assimilation traits and dinitrogen fixation in soybean and common bean. *Agron. J.* 84:1020–1028. doi:10.2134/agronj1992.00021962008400060022x
- Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1993. Genetic diversity in North American soybean: II. Prediction of heterosis in F<sub>2</sub> populations of southern founding stock using genetic similarity measures. *Crop Sci.* 33:620–626. doi:10.2135/cropsci1993.0011183X003300030039x
- Good, A.G., A.K. Shrawat, and D.G. Muench. 2004. Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? *Trends Plant Sci.* 9:597–605. doi:10.1016/j.tplants.2004.10.008
- Hammer, G., M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. van Eeuwijk, S. Chapman, and D. Podlich. 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
- Hirel, B., J. Le Gouis, B. Ney, and A. Gallais. 2007. The challenge of improving nitrogen use efficiency in crop plants: Towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 59:2369–2387. doi:10.1093/jxb/erm097
- Hyten, D.L., Q.J. Song, Y.L. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, and P.B. Cregan. 2006. Impacts of genetic bottlenecks on soybean genome diversity. *Proc. Natl. Acad. Sci. USA* 103:16666–16671. doi:10.1073/pnas.0604379103
- Janssen, B.H., F.C.T. Guiking, D. van der Eijk, E.M.A. Smaling, J. Wolf, and H. van Reuler. 1990. A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). *Geoderma* 46:299–318. doi:10.1016/0016-7061(90)90021-Z
- Jin, J., L. Liu, G. Wang, L. Mi, Z. Shen, X. Chen, and S.J. Herbert. 2010. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. *Field Crops Res.* 115:116–123. doi:10.1016/j.fcr.2009.10.016
- Jin, J., X. Liu, G. Wang, J. Liu, L. Mi, X. Chen, and S.J. Herbert. 2011. Leaf nitrogen status as a main contributor to yield improvement of soybean cultivars. *Agron. J.* 103:441–448.
- Jung, S., D.A. Rickert, N.A. Deak, E.D. Aldin, J. Recknor, L.A. Johnson, and P.A. Murphy. 2003. Comparison of Kjeldahl and Dumas methods for determining protein contents of soybeans products. *J. Am. Oil Chem. Soc.* 80:1169–1173. doi:10.1007/s11746-003-0837-3
- Kisha, T.J., C.H. Sneller, and B.W. Diers. 1997. Relationship between genetic distance among parents and genetic variance in populations of soybean. *Crop Sci.* 37:1317–1325. doi:10.2135/cropsci1997.0011183X003700040048x
- Morrison, M.J., H.D. Voldeng, and E.R. Cober. 1999. Physiological changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agron. J.* 91:685–689. doi:10.2134/agronj1999.914685x

- Morrison, M.J., H.D. Voldeng, and E.R. Cober. 2000. Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Crop Sci.* 92:780–784.
- Novoa, R., and R.S. Loomis. 1981. Nitrogen and plant production. *Plant Soil* 58:177–204. doi:10.1007/BF02180053
- Pazdernik, D.L., P.H. Graham, and J.H. Orf. 1997. Variation in the pattern of nitrogen accumulation and distribution in soybean. *Crop Sci.* 37:1482–1486. doi:10.2135/cropsci1997.0011183X003700050011x
- Perry, M.C. 1991. Geographical patterns of variation in the USDA soybean germplasm collection: I. Morphological traits. *Crop Sci.* 31:1350–1355. doi:10.2135/cropsci1991.0011183X00310050054x
- R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reynolds, M., D. Bonnett, S.C. Chapman, R.T. Furbank, Y. Manés, D.E. Mather, and M.A.J. Parry. 2011. Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. *J. Exp. Bot.* 62:439–452. doi:10.1093/jxb/erq311
- Robinson, N., A. Fletcher, A. Whan, C. Critchley, N. Von Wiren, P. Lakshmanan, and S. Schmidt. 2007. Sugarcane genotypes differ in internal nitrogen use efficiency. *Func. Plant Biol.* 34:1122–1129.
- Rotundo, J.L., and M.R. Aguiar. 2008. Herbivory resistance traits in populations of *Poa ligularis* subjected to historically different sheep grazing pressure in Patagonia. *Plant Ecol.* 194:121–133. doi:10.1007/s11258-007-9279-8
- Rotundo, J.L., L. Borrás, M.E. Westgate, and J.H. Orf. 2009. Relationship between assimilate supply per seed during seed filling and soybean seed composition. *Field Crops Res.* 112:90–96. doi:10.1016/j.fcr.2009.02.004
- Rotundo, J.L., L. Borrás, and M.E. Westgate. 2011. Linking assimilate supply and seed developmental processes that determine soybean seed composition. *Eur. J. Agron.* 35:184–191. doi:10.1016/j.eja.2011.05.002
- Rotundo, J.L., L. Borrás, J. De Bruin, and P. Pedersen. 2012. Physiological strategies for seed number determination in soybean: Biomass accumulation, partitioning and seed set efficiency. *Field Crops Res.* 135:58–66. doi:10.1016/j.fcr.2012.06.012
- Salvagiotti, F., K.G. Cassman, J.E. Specht, D.T. Walters, A. Weiss, and A. Dobermann. 2008. Nitrogen uptake, fixation and response to fertilizer N in soybeans: A review. *Field Crops Res.* 108:1–13. doi:10.1016/j.fcr.2008.03.001
- Sinclair, T.R., and C.T. de Wit. 1975. Photosynthate and nitrogen requirements for seed production by various crops. *Science* 189:565–567. doi:10.1126/science.189.4202.565
- Sinclair, T.R., and C.T. de Wit. 1976. Analysis of carbon and nitrogen limitations to soybean yield. *Agron. J.* 68:319–324. doi:10.2134/agronj1976.00021962006800020021x
- Sinclair, T.R., and T. Horie. 1989. Leaf Nitrogen, Photosynthesis, and Crop Radiation Use Efficiency— a Review. *Crop Sci.* 29:90–98. doi:10.2135/cropsci1989.0011183X002900010023x
- Sinclair, T.R., and P.D. Jamieson. 2006. Grain number, wheat yield, and bottling beer: An analysis. *Field Crops Res.* 98:60–67. doi:10.1016/j.fcr.2005.12.006
- Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential—A genetic and physiological perspective. *Crop Sci.* 39:1560–1570. doi:10.2135/cropsci1999.3961560x
- Weih, M. 2003. Trade-offs in plants and the prospects for breeding using modern biotechnology. *New Phytol.* 158:7–9. doi:10.1046/j.1469-8137.2003.00716.x
- Wilcox, J.R. 2001. Sixty years of improvement in publicly developed elite soybean lines. *Crop Sci.* 41:1711–1716. doi:10.2135/cropsci2001.1711
- Witt, C., A. Dobermann, S. Abdulrachman, H.C. Gines, W. Guanghuo, R. Nagarajan, S. Satawatanont, T. Thuc Son, P. Sy Tan, L. Van Tiem, G.C. Simbahan, and D.C. Olk. 1999. Internal nutrient efficiencies of irrigated lowland rice in tropical and subtropical Asia. *Field Crops Res.* 63:113–138. doi:10.1016/S0378-4290(99)00031-3
- Wollenweber, B., J.R. Porter, and T. Lubberstedt. 2005. Need for multidisciplinary research towards a second green revolution—Commentary. *Curr. Opin. Plant Biol.* 8:337–341. doi:10.1016/j.pbi.2005.03.001
- Xu, G., X. Fan, and A.J. Miller. 2012. Plant nitrogen assimilation and use efficiency. *Annu. Rev. Plant Biol.* 65:1–30.
- Zapata, F., S.K.A. Danso, G. Hardarson, and M. Fried. 1987. Time course of nitrogen-fixation in field-grown soybean using N-15 methodology. *Agron. J.* 79:172–176. doi:10.2134/agronj1987.00021962007900010035x
- Zeiger, C., D.B. Egli, J.E. Leggett, and D.A. Reicosky. 1982. Cultivar differences in N-redistribution in soybeans. *Agron. J.* 74:375–379. doi:10.2134/agronj1982.00021962007400020027x