

Reduced soybean photosynthetic nitrogen use efficiency associated with evolutionary genetic bottlenecks

José L. Rotundo^{A,B} and Lucas Borrás^A

^AFacultad de Ciencias Agrarias, Universidad Nacional de Rosario, Campo Exp. Villarino S/N, S2125ZAA, Zavalla, Santa Fe, Argentina.

^BCorresponding author. Email: jrotundo@unr.edu.ar

Abstract. Soybean has a narrow genetic base thought to limit future yield genetic gains. However, there is no evidence whether this reduction in genetic diversity correlates with diversity loss for any yield trait. We tested how photosynthetic nitrogen use efficiency (leaf photosynthesis per unit nitrogen, NUE_p) evolved from the wild relative *Glycine soja* Siebold & Zucc. to the current *Glycine max* (L.) Merr. Five populations resulting from different evolutionary bottlenecks were evaluated under field conditions. Populations were wild ancestors, domesticated Asian landraces, North American ancestors, and modern cultivars. Genotypic differences in photosynthesis and leaf nitrogen were evident, creating a significant 3-fold variation in phenotypic NUE_p . There was a parallel reduction in molecular marker and phenotypic NUE_p diversity after each evolutionary bottleneck. *G. soja* had three times more NUE_p diversity and 25% more average NUE_p compared with the elite modern cultivars. Two strategies for increasing NUE_p were identified: (i) increases in light saturated photosynthesis (P_{max}), and, alternatively, (ii) reductions in leaf nitrogen. A modelling approach showed that NUE_p will increase yield only if based on increased P_{max} . Our study quantified the genetic potential of exotic germplasm available for trait-directed breeding. Results antagonise the concept that elite germplasm is always superior for any relevant yield trait when compared with undomesticated germplasm.

Additional keywords: carbon assimilation, genetic gain, natural genetic variation, phenotypic diversity, trait based hybridisation.

Received 15 January 2016, accepted 3 May 2016, published online 7 June 2016

Introduction

Prospects to increase soybean yields rely on faster rates of genetic improvement (Ainsworth *et al.* 2012). However, many consider a narrow genetic base a major constraint for increasing yield genetic gain in most modern crop species (e.g. Tanksley and McCouch 1997). Commercial soybean cultivars are known in particular for their very narrow genetic base, since modern cultivars were developed from few ecotypes collected in their centre of origin (Mikel *et al.* 2010). Broadening this base requires using exotic material to incorporate useful genetic diversity (Concibido *et al.* 2003). *Glycine soja* Siebold & Zucc. is the closest soybean wild relative, considered the undomesticated ancestor of *G. max* (L.) Merr. *G. soja* and *G. max* cross with one another and give fertile descendants, making *G. soja* a potential source of new alleles for commercial breeding. However, the use of wild relatives to broaden a breeding program is limited due to large yield drags associated with undesirable traits (Concibido *et al.* 2003). When breeders decide to cross among their elite lines for developing new commercial cultivars they are indirectly acknowledging there is enough genetic diversity among these to further improve their germplasm. Most importantly, there is a general belief that elite

germplasm is always more efficient at any yield related trait than undomesticated germplasm (Denison 2009).

Recent prospects to genetically improve crop yields have pointed at increases in leaf photosynthesis through different mechanisms (Long *et al.* 2006; Ainsworth *et al.* 2012; Gu *et al.* 2014). Efforts include bioengineering C_3 crops to express the more efficient C_4 photosynthetic system, improving Rubisco efficiency via directed mutations, or screening for natural photosynthesis variation (Long *et al.* 2006). In soybean, light saturated photosynthesis (P_{max}) is a trait used to model genotypic differences in crop growth and yield (Boote *et al.* 2003); it has been used as a selection criterion to increase soybean yield, without positive results (e.g. Dornhoff and Shibles 1970; Wiebold *et al.* 1981).

P_{max} has a saturation response to specific leaf nitrogen (SLN, $g\ N\ m^{-2}$), and SLN is negatively correlated with total leaf area responsible for light capture (Sinclair and Horie 1989). Increases in P_{max} associated with higher SLN results in lower total crop leaf area and less radiation interception. Improvements in carbon assimilation at the leaf level are counterbalanced by reduced radiation interception at the crop level. This can result in no changes, or even reductions, in total

canopy assimilation. Therefore, P_{\max} must increase without additional increases in SLN to positively impact crop yield; there is a need for more P_{\max} per unit of leaf nitrogen (Boote *et al.* 2003). This proposed decoupling of P_{\max} and SLN is physiologically feasible, resulting in increased photosynthetic nitrogen use efficiency (NUE_p , $\mu\text{g CO}_2 \text{ mg N}^{-1} \text{ s}^{-1}$) (e.g. Quick *et al.* 1991; Zhu *et al.* 2007). There is, however, basically no knowledge on the evolution or available genotypic diversity in NUE_p .

Evolutionary events like domestication, introduction to new areas and intensive commercial breeding act as genetic bottlenecks, and negatively impact genetic diversity. For soybean, Hyten *et al.* (2006) evaluated these genetic bottlenecks for DNA sequence diversity by comparing genotypes from four populations: soybean wild relatives (*G. soja*), domesticated Asian *G. max* landraces, North American ancestors of modern varieties and elite North American cultivars. Results indicated the most severe diversity loss occurred after domestication and introduction; domestication resulted in 50% diversity reduction while introduction accounted for losing 78% of rare alleles. Despite this genetic diversity loss, there is no information on the impact of these genetic bottlenecks on any trait that can have a positive impact on canopy growth and yield.

Our general objective was to test whether going back through the soybean evolutionary pathway, from current cultivars to wild relatives, could reveal novel variation in leaf traits resulting in increased NUE_p . Variation can be lost by random processes like genetic drift or by directional selection at breeding programs. Our specific objectives were: (i) to evaluate NUE_p variability among and within populations, and compare it against a diversity index for DNA sequence data published by Hyten *et al.* (2006); (ii) to evaluate the existence of different physiological strategies to attain high NUE_p ; and (iii) to evaluate the crop level relevance of NUE_p to increase yield using a simulation model.

Materials and methods

Germplasm

We phenotyped the same 111 *Glycine soja* Siebold & Zucc. and *Glycine max* (L.) Merr. genotypes that were evaluated by Hyten *et al.* (2006). These genotypes correspond to four different populations: (i) 22 *G. soja* accessions collected from its natural distributional areas in China, Korea and Japan (from now on Wild relatives), (ii) 50 *G. max* landraces, domesticated in China, Korea, and Japan (from now on Asian landraces), (iii) 14 *G. max* founders of North America breeding programs representing 86% of the initial variation in commercial cultivars (from now on NAm ancestors), and (iv) 25 *G. max* commercial cultivars released in North America between 1980 and 1990 (from now on Elite NAm cultivars). In addition, 10 *G. max* elite commercial cultivars from Argentina released between 2000 and 2010 were included (from now on Elite ARG cultivars). Due to seed quality and availability 9 genotypes were lost from the original list tested by Hyten *et al.* (2006). The complete list of genotypes is available in Table S1, available as Supplementary Material to this paper.

Experimental design and growing conditions

A 2 year field experiment was conducted at the Campo Experimental Villarino, Universidad Nacional de Rosario, located at Zavalla (33°1'S, 60°53'W), Santa Fe, Argentina. Soils are Vertic Argiudoll, Roldán series. Each experiment was a complete randomised block design with five replications. Each replicate consisted of four plants growing 0.25 m apart within the same row and 0.52 m between rows. This low stand density was intended to grow isolated plants fully exposed to sun. Weeds were chemically controlled before soybean planting and hand removed after soybean emergence. Pest and diseases were prevented by insecticide and fungicide applications. Plots were irrigated frequently to avoid any drought stress.

Measurements

From a developmental perspective, individual leaf photosynthesis is maximum (less than 10% variation from absolute maximum) between 18 and 26 days after total leaf expansion (Evans 1993). Fully-expanded leaves were tagged daily after V4 and the date was recorded to ensure selecting leaves having between 18 and 26 days after full expansion at measurement. Within each year, measurements were conducted early in the plant cycle during the vegetative stage (stages V8–V10) of the different genotypes. This early vegetative sampling helped avoid possible interactions with maturity group and senescence due to reproductive growth. Leaf photosynthesis was measured during three consecutive sunny days (blocks 1 and 2 on day 1, blocks 3 and 4 on day 2 and block 5 on day 3) during 3 h around solar noon. Measurements were made using a portable gas exchange system (Li-Cor 6400, Li-Cor). Measured leaves were typically attached to central nodes, leaflets where exposed to sun (sunny leaves), and one central leaflet per replicate was measured. The system was zeroed using anhydrous calcium carbonate to eliminate water, and soda lime to remove carbon dioxide from the air entering the system. Leaf temperature was set at 25°C. Irradiance intensity was set at $1500 \mu\text{mol m}^{-2} \text{ s}^{-1}$ using red-blue light source in an integrated chamber. Flow rate was set at $300 \mu\text{mol s}^{-1}$, and ambient CO_2 concentration of $380 \mu\text{mol mol}^{-1}$. Each measurement was logged after reaching three stability criteria. These criteria were that slopes with time (seconds) were <1 for CO_2 concentration of the sample cell (CO_2S parameter, $\mu\text{mol mol}^{-1}$) and water concentration (H_2OS parameter, mmol mol^{-1}), and <0.5 for g_s (Cond parameter, $\text{mol m}^{-2} \text{ s}^{-1}$).

Immediately after measuring leaf gas exchange, the measured leaflet was removed from the plant and photographed with a digital camera on a white background including a known area standard (2 cm^2) within the image. Pixels for each leaflet and standard were counted using ImageJ software for leaf area estimations. The leaflet was then dried for 96 h at 65°C, weighed and ground.

Leaf traits

Leaf maximum photosynthesis (P_{\max} , $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) was directly obtained from the Li-Cor instrument. Leaf nitrogen concentration (LNC, $\text{g N g}^{-1} \times 100$) was measured with Kjeldahl after grinding and drying the tissue for at least 96 h on an air-forced oven at 65°C. Specific leaf weight (SLW, g cm^{-2})

was calculated as the ratio between leaf dry weight (g) and leaf area (cm²). Specific leaf nitrogen (SLN, g N cm⁻²) was calculated as the product between LNC and SLW. Finally, photosynthetic nitrogen use efficiency (NUE_p, μmol CO₂ g N⁻¹ s⁻¹) was calculated as the ratio between P_{\max} and SLN.

Modelling of NUE_p as a yield related trait

The CROPGRO soybean model, embedded in DSSAT v4 (Jones *et al.* 2003), was used to test the value of NUE_p as a yield related trait. This mechanistic model is extensively utilised to address diverse objectives, like (i) assisting breeders in the evaluation of specific traits to increase seed yield (e.g. Boote *et al.* 2003), (ii) assessing on-farm yield for management purposes (e.g. Mercau *et al.* 2007), and (iii) calculating yield gaps at regional scales for policy makers (e.g. Aramburu Merlos *et al.* 2015). The current modelling exercise falls within the first of the above mentioned categories.

Modelling the growth and development of specific cultivars, in a given soil, weather, and management condition, is governed by a series of genetic coefficients. For our modelling exercise, yield of a default maturity group II cultivar grown in Ames (Iowa, USA) and another default maturity group IV cultivar grown in Zavalla (Santa Fe, Argentina) was simulated using a 15 year weather series for each location. Weather files included daily records of rainfall, maximum and minimum temperatures, and radiation. Genetic coefficients for these cultivars are provided with the DSSAT v4. These coefficients are derived from standard performance trials or more controlled environmental experiments using the crop model in an optimisation mode (Mavromatis *et al.* 2001). Soil series for each site were Clarion (Typic Haplaquolls) and Roldán (Vertic Argiudoll) for Ames and Zavalla, respectively. Simulated planting date was 15 May and 15 November for Ames and Zavalla, respectively.

Cultivars with increased NUE_p were simulated by modifying, one at a time, genetic coefficients associated with leaf nitrogen concentration (reduced LNC while maintaining P_{\max}), specific leaf weight (reduced SLW while maintaining P_{\max}), and P_{\max} (increased P_{\max} while maintaining LNC and SLW) following the approximation used by Boote *et al.* (2003). The CROPGRO soybean model abbreviations for these coefficients are PROLFG, SLAVR, and LFMAX, respectively; initial values were 0.285, 375, and 1.030 respectively. The concept was to compare the simulated yield of a standard genotype versus the yield of a modified genotype having changed one genetic coefficient as a putative yield increase trait. The magnitude of the modifications was determined according to the variation observed in our field experiments. No intention of model validation was pursued since the objective of our exercise was to test yield sensitivity to changes in specific traits related with NUE_p and not to predict yield in a particular environment.

Statistical analysis

Density distribution of genotypic NUE_p within each population (mean of two environments and five replications per genotype) was constructed using the Nonparametric Kernel Smoothing Method of the sm package in R (Bowman and Azzalini 2014).

Phenotypic diversity for each population was calculated as the difference between maximum and minimum NUE_p relative to the average NUE_p (Gu *et al.* 2014).

Differences among and within populations were tested using a mixed model ANOVA. Years, blocks within year, and the interactions including these factors were considered random factors. Populations and genotypes within populations were considered fixed factors. Data were analysed using the MIXED procedure of SAS (SAS Institute). Means were compared with a Fisher-protected least significant difference (l.s.d.) test at the 0.05 probability level. Each entry belonging to the Wild relative population was then tested against the Elite NAM and ARG cultivars having the highest NUE_p to identify genotypes having increased NUE_p relative to these reference cultivars. Those genotypes having significant difference relative to these benchmark cultivars were selected for identifying different physiological strategies to attain high NUE_p.

To identify different physiological strategies in terms of leaf traits leading to high NUE_p *post hoc* comparisons for P_{\max} , SLW, and LNC were conducted only for Wild relative genotypes having significantly higher NUE_p compared with the commercial reference cultivars. Comparing the physiological processes determining NUE_p allowed testing whether different strategies exist for increased NUE_p determination.

Results

Phenotypic variation in NUE_p among and within populations

Soybean Wild relative, *G. soja*, had the highest NUE_p compared with the other populations (Table 1). The lowest NUE_p was observed in Elite NAM and ARG cultivars, whereas the Asian landraces and NAM ancestors had intermediate NUE_p.

Significant population and genotypic differences were also evident for P_{\max} , SLW, and LNC ($P < 0.001$), not only for NUE_p (Table 1). Plotting P_{\max} versus SLN showed scattered data, indicating substantial variation in NUE_p across genotypes (Fig. 1). The specific trait values for each particular genotype are described in Table S1.

The Wild relative *G. soja* population had a wider distribution of NUE_p values compared with the other populations (Fig. 2). The phenotypic diversity index calculated for NUE_p was highest for the Wild relative population (Table 2). This wider distribution and increased diversity in *G. soja* was associated with individual genotypes having NUE_p values above the average of the other four populations. The phenotypic diversity index dropped substantially for the Asian landraces. The NAM ancestors and the Elite NAM cultivars had similar levels of phenotypic diversity, which were lower than the Asian landraces. The lowest diversity was observed for the Elite ARG cultivars (Table 2).

The observed phenotypic diversity values for NUE_p positively correlated with nucleotide diversity as reported by Hyten *et al.* (2006). Both approximations evaluating diversity (molecular and phenotypic) indicated a major diversity loss after the domestication bottleneck. Also, they both show small, if any, difference between NAM ancestors and Elite NAM cultivars.

Table 1. Average photosynthetic N use efficiency (NUE_p) for soybean populations differing in evolutionary history

Different letters indicate significant differences ($P < 0.05$). Genotypes evaluated within each population were 22 Wild relatives, 50 Asian landraces, 14 North American ancestors, 25 Elite cultivars from North America (NAM) and 10 from Argentina (ARG). Descriptions of each individual genotype are available in Table S1, available as Supplementary Material to this paper. Data are the mean of 2 years and five replications ($n = 10$). Different letters indicate significant differences (Fisher's l.s.d. test)

Population	Photosynthetic N use efficiency ($\mu\text{mol CO}_2 \text{ g N}^{-1} \text{ s}^{-1}$)	Specific leaf weight (g cm^{-2})	Leaf N concentration ($\text{g N g}^{-1} \times 100$)	Specific leaf N (g N cm^{-2})	Maximum photosynthesis ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$)
Wild relatives (<i>G. soja</i>)	10.57a	46.9d	4.19b	2.05d	20.69b
Asian landraces	9.09b	60.1b	4.23b	2.39c	21.11b
North Am. ancestors	9.15b	57.4c	4.19b	2.51b	22.41a
Elite NAM cultivars	8.55c	64.5a	4.21b	2.66a	22.32a
Elite ARG cultivars	8.62bc	59.3b	4.47a	2.51b	22.84a
Source of variation	<i>P</i> -value	<i>P</i> -value	<i>P</i> -value	<i>P</i> -value	<i>P</i> -value
Population	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Genotype (population)	<0.0001	<0.0001	<0.0001	<0.0001	0.0053

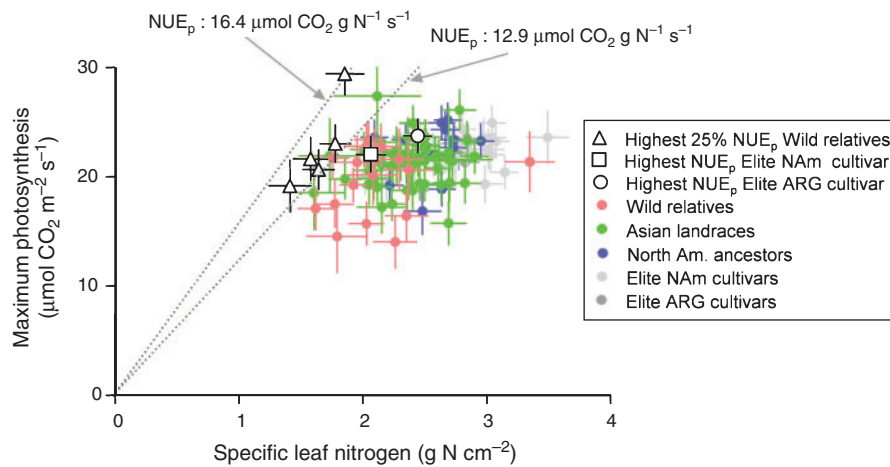


Fig. 1. Scatterplot of maximum photosynthesis (P_{max}) and specific leaf nitrogen (SLN) for 121 soybean genotypes from five populations differing in evolutionary history. Data for each genotype are the average of two years and five replications ($n = 10$). Genotypes evaluated within each population were 22 Wild relatives, 50 Asian landraces, 14 North American ancestors, 25 Elite cultivars from North America (NAM), and 10 from Argentina (ARG). Open triangles indicate the superior 25% Wild relative genotypes in terms of photosynthetic N use efficiency (NUE_p). The open square and open circle indicates the elite cultivar from North America (NAM) and Argentina (ARG), respectively, having the absolute highest NUE_p. The grey dotted lines show the isolines of NUE_p embracing the best selected Wild relatives.

Within the Wild relative population there were five *G. soja* genotypes having maximum NUE_p (Fig. 2). These five genotypes had significantly higher NUE_p compared with the best Elite NAM and ARG cultivars (Table 3). The photosynthetic nitrogen use efficiency for these *G. soja* genotypes ranged from 12.9 and 16.4 $\mu\text{mol CO}_2 \text{ g N}^{-1} \text{ s}^{-1}$ (Table 3; Fig. 2). Photosynthetic NUE for the best Elite NAM and ARG cultivars were 10.6 and 9.8 $\mu\text{mol CO}_2 \text{ g N}^{-1} \text{ s}^{-1}$ respectively (Table 3).

Different physiological strategies to attain high NUE_p

Selected *G. soja* genotypes had between 25 to 60 % more NUE_p compared to the Elite NAM and ARG cultivars having the highest NUE_p within those groups. This NUE_p response was attained via different combinations of specific leaf traits, indicating there were different strategies for attaining high NUE_p

(Table 3). For instance, the *G. soja* genotype PI597459 had the highest NUE_p by combining intermediate SLW and LNC with the highest P_{max} (Table 3). The physiological strategy for PI366120 to excel in NUE_p was combining low SLW with high LNC. In contrast, PI407301 combined a relatively higher SLW with a very low LNC value (Table 3). The remaining two genotypes (PI407275 and PI562559) had intermediate strategies compared with these extremes. In all cases, except for PI597459, the strategies were not based on P_{max} differences since all these genotypes had similar values for P_{max} .

Modelling of NUE_p as a yield-related trait

Compared with the average of PI533602 and A5409RG (Elite NAM and ARG cultivars with the highest NUE_p respectively), *G. soja* genotypes with increased NUE_p had: a 36% reduction

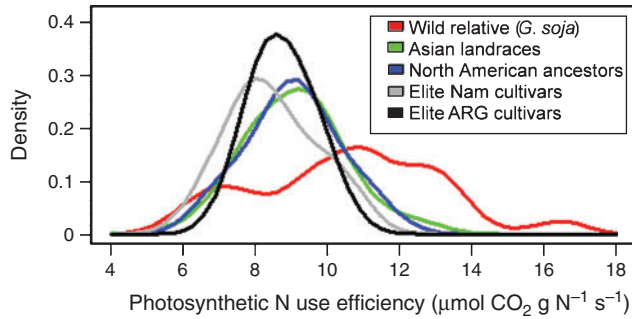


Fig. 2. Density distribution for photosynthetic nitrogen use efficiency across five soybean populations differing in evolutionary history. Data for each genotype are the average of 2 years and five replications ($n=10$). Genotypes evaluated within each population were 22 Wild relative *G. soja*, 50 Asian landraces, 14 North American ancestors, 25 Elite cultivars from North America (NAM) and 10 from Argentina (ARG).

Table 2. Phenotypic diversity index for photosynthetic nitrogen use efficiency (NUE_p) across soybean populations differing in evolutionary history

Genotypes evaluated within each population were 22 Wild relatives, 50 Asian landraces, 14 NAM ancestors, 25 Elite cultivars from North America (NAM) and 10 from Argentina (ARG). Data for each genotype are the average of 2 years and five replications ($n=10$). Nucleotide diversity, as measured by two indicators (expected heterozygosity and polymorphic sites), are data from 102 randomly selected genes from Hyten *et al.* (2006). Abbreviation: na, not available

Population	Phenotypic diversity for NUE_p (%) ^A	Nucleotide diversity ^B	
		Π^D	Θ^D
Wild relatives (<i>G. soja</i>)	93.7	2.17	2.35
Asian landraces	70.8	1.43	1.15
North Am ancestors	48.3	1.14	1.00
Elite NAM cultivars	48.8	1.11	0.83
Elite ARG cultivars	25.5	na	na

^ACalculated as ((maximum – minimum)/mean) \times 100 (Gu *et al.* 2014).

^BData from Hyten *et al.* (2006).

^CExpected heterozygosity per nucleotide site.

^DNumber of polymorphic sites.

in SLW (PI366120), an 8% reduction in LNC (PI407301), and a 22% increase in P_{max} (PI597459) (Table 3). These three genotypes had the lowest SLW, lowest LNC, and highest P_{max} , respectively, and can be referred as three different physiological strategies available for increasing NUE_p . According to these results, we modified the default genetic coefficients for SLW (reduced 36% relative to the default), LNC (reduced 8%), and P_{max} (increased 22%). These changes allowed simulating different strategies to attain increased NUE_p and test their yield effect.

Increasing NUE_p via reduced SLW or reduced LNC caused almost no increase in yield in neither Argentina nor North America simulations (Fig. 3). However, when the increase in NUE_p is obtained via higher P_{max} , without proportional increases in nitrogen investment, there is an average yield increase of 24 and 12% for Argentina and North America respectively (Fig. 3). In both simulations, there was a positive correlation between simulated total aboveground biomass and yield, but no association was observed between harvest index and yield (data not shown). There was a tendency for larger yield benefits in years with higher yields.

Discussion

Soybean has achieved an average of 22 kg ha⁻¹ year⁻¹ yield gain as a result of conventional breeding by hybridisation among elite soybean lines (Rincker *et al.* 2014). This successful strategy of directional selection has contributed to the narrowing of the soybean genetic base currently used by breeders. In addition, involuntary genetic drift associated with domestication and introduction to new areas may be responsible for the loss of useful traits due to chance (Hyten *et al.* 2006). Our study shows that tracing back the evolutionary history of soybean allows for the identification of genotypes with the capacity to improve current commercial cultivars (e.g. Xiao *et al.* 1996; Tanksley and McCouch 1997; Reynolds *et al.* 2007). We hypothesise that focusing on yield rather than on the identification of putative yield-related traits has contributed to the limited number of positive breeding examples with exotic germplasm in the past. In the present study we identified several *G. soja* genotypes that have increased NUE_p . These lines are possible candidates

Table 3. Leaf trait values for the highest 25% Wild relatives *Glycine soja* in terms of high photosynthetic N use efficiency

Selection criterion was that these genotypes significantly differ from the Elite cultivars having the highest NUE_p (PI533602 for North America and A5409 for Argentina). Data are the average of 2 years and five replications ($n=10$). Different letters indicate significant differences (Fisher's l.s.d. test)

Population	Genotype	Photosynthetic N use efficiency ($\mu\text{mol CO}_2 \text{ g N}^{-1} \text{ s}^{-1}$)	Specific leaf weight (g cm^{-2})	Leaf N concentration ($\text{g N g}^{-1} \times 100$)	Specific leaf N (g N cm^{-2})	Maximum photosynthesis ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$)
Wild relatives (<i>G. Soja</i>)	PI366120	13.11b	33.4d	4.43ab	1.47d	19.20b
	PI407275	13.78b	40.6cd	3.95bc	1.58cd	21.57b
	PI407301	12.54b	42.4bc	3.86c	1.64cd	20.73b
	PI562559	13.02b	41.5cd	4.31ab	1.78cd	23.01b
	PI597459	16.04a	45.8bc	4.02bc	1.86c	29.45a
Elite NAM cultivar	PI533602	10.62c	49.6ab	4.07bc	2.07b	22.02b
Elite ARG cultivar	A5409RG	9.80c	55.4a	4.43a	2.45a	23.74b

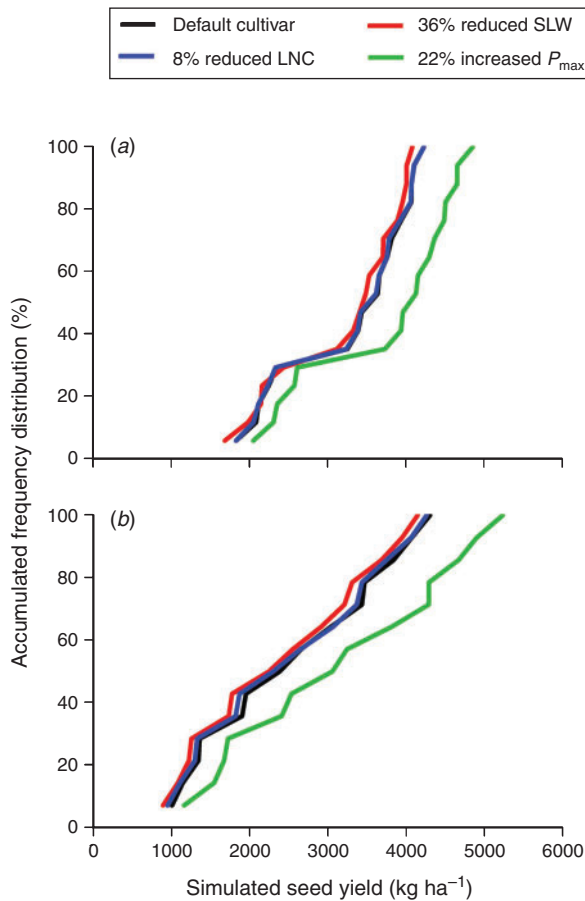


Fig. 3. Frequency distribution of simulated yield for a standard maturity group II genotype grown in Iowa North America (Fig. 3a), and a standard maturity group IV genotype grown in Santa Fe Argentina (Fig. 3b). Simulations were done with the Soybean CROPGRO model using a 15 year weather series. Different lines represented a default standard cultivar (black) and cultivars with increased NUE_p via three different strategies (reduced specific leaf weight, SLW, reduced leaf nitrogen concentration, LNC, or increased maximum photosynthesis, P_{max}).

for wide crosses with elite germplasm that could be used in trait-based hybridisation seeking to combine useful physiological traits into superior progeny, as proposed by Reynolds *et al.* (2011).

Hyten *et al.* (2006) demonstrated that genetic bottlenecks during soybean evolution caused severe diversity losses in molecular marker allele diversity. However, the correlation between DNA marker diversity and phenotypic diversity is controversial (Reed and Frankham 2001). Molecular variation is not strongly associated with phenotypic variation, suggesting quantitative trait variation should be directly assessed (Stupar 2010). Our results showed that within population diversity in NUE_p matched the variation observed in molecular markers as reported by Hyten *et al.* (2006). The positive association between phenotypic and molecular diversity we observed illustrates that at least some of the markers assessed by Hyten *et al.* (2006) were not neutral; variants in those markers altered the expression of different genes. Reed and Frankham (2001) have noted that different selective forces (that should be lower in

neutral molecular markers compared with markers affecting yield related phenotypic traits) are likely the major cause of weak correlations between molecular and phenotypic variation. The major genetic bottleneck due to domestication caused the first reduction in both molecular marker and NUE_p diversity. Further bottlenecks (introduction and commercial breeding) caused paralleled reductions in both diversity estimates. We note that the lowest NUE_p diversity was observed in Elite ARG cultivars, which are known to be derived from NAM cultivars.

Photosynthetic NUE varied between and within populations as a function of P_{max} , SLW and LNC combinations. Comparisons at the population level demonstrated that *G. soja* had the highest NUE_p compared with the Asian landraces, the NAM ancestors and the Elite NAM and ARG cultivars. At the individual genotypic level, there was considerable variation in NUE_p ranging from ~ 5 to $\sim 16 \mu\text{mol CO}_2 \text{ g N}^{-1} \text{ s}^{-1}$, as shown by the lack of correlation between P_{max} and SLN (Fig. 1). From the Wild relative population, five *G. soja* lines were identified as being superior to the highest NAM and ARG Elite cultivars in terms of NUE_p . We noted that this increased NUE_p originated through different combinations of increased P_{max} , and/or reduced SLW and LNC values. Only PI597459 had increased NUE_p due to an increased P_{max} ($29.4 \mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$). This value is higher than the upper limit of known genetic variation for this trait as reported for commercial cultivars by Boote *et al.* (2003). The genotype PI366120 has SLW substantially lower than the range reported by Boote *et al.* (2003), indicating the existence of very thin leaves. Some of these lines are possible candidates for crossing with modern commercial cultivars for testing the value of using exotic germplasm for yield increases, and for molecular marker studies targeting the detection of relevant genomic regions.

The decoupling of P_{max} and SLN necessary to increase NUE_p , as demonstrated for some genotypes, can be explained by at least three physiological mechanisms. First, reducing Rubisco synthesis by antisense technology demonstrated increased Rubisco activation without changes in P_{max} , increasing NUE_p (Quick *et al.* 1991); genotypes having less Rubisco may have increased NUE_p if activation is increased. Second, changes in leaf nitrogen partitioning between Rubisco and cell wall proteins is a mechanism that can also determine genotypic differences in NUE_p (Harrison *et al.* 2009). Finally, using evolutionary algorithms, Zhu *et al.* (2007) demonstrated that the distribution of leaf nitrogen between carbon metabolism enzymes can be optimised, resulting in P_{max} increases investing similar amounts of leaf nitrogen. Fig. 1 describes large genotypic variations between P_{max} and SLN.

Lowered NUE_p in elite cultivars could have been caused by processes other than random drift. A plausible alternative hypothesis is that yield negatively correlates with NUE_p , and that the lower NUE_p in elite cultivars is caused by directional selection. Studies describing soybean genetic progress showed NUE_p is not necessarily associated with increased yield (Jin *et al.* 2011). Results from our modelling exercise showed that for NUE_p to increase yield it has to be determined by increased P_{max} for a given amount of nitrogen invested in the leaf. The strategies associated with increased NUE_p via 'savings' in N investments did not determine any yield advantage. This is in general agreement with Boote *et al.* (2003), showing the

potential benefit of increased P_{\max} only when it is not associated with increases in SLW. Any increase in SLW would positively impact SLN and therefore reduce NUE_p . The average simulated yield increases ranged from 12 to 24%. If these gains are to be achieved via conventional breeding it would represent 12 to 24 years, assuming an average 1% genetic gain per year. The possibility to successfully translate *G. soja* traits to *G. max* would depend on the correlations between LNC, SLW and P_{\max} . Phenotypic correlations among those traits for *G. soja* genotypes were low and not significant (data not shown).

Several studies indicate there is unused genetic potential in exotic germplasm for yield increases (Xiao *et al.* 1996; Gur and Zamir 2004; Reynolds *et al.* 2007). Yield testing of exotic lines may mislead their potential to produce genetic gains because of yield drags associated with pre-domestication traits. A physiological approach, like the one we have proposed, might be necessary. There is a need to develop and use physiological frameworks to identify and screen useful traits that may be hidden in low-yielding un-adapted exotic germplasm. This will lead to a more rational germplasm selection when conducting wide crosses with elite soybean cultivars.

Conclusions

By unravelling the evolutionary soybean history we demonstrated that *G. soja* is a reservoir of useful traits that play out through changes in NUE_p . Photosynthetic nitrogen use efficiency was higher in *G. soja* compared with soybean landraces, ancestors, and even elite germplasm. This increased NUE_p in specific *G. soja* genotypes was attained by increases in P_{\max} or, alternatively, by reductions in LNC. A modelling exercise was pursued to scale-up these leaf traits to canopy level processes. By doing so, we demonstrated that for NUE_p to increase yield it has to be determined by P_{\max} increases and not by leaf nitrogen savings. Physiological frameworks to identify and screen useful traits hidden in low-yielding unadapted exotic germplasm are needed.

Acknowledgements

The project was supported by the Ministerio Nacional de Promoción Científica y Tecnológica (PICT-2011–1292). JL Rotundo and L Borrás are members of CONICET, the Argentine National Research Council. We acknowledge valuable comments by C Ballaré. Authors also wish to thank IFEVA for lending their photosynthesis system.

References

Ainsworth EA, Yendrek CR, Skoneczka JA, Long SP (2012) Accelerating yield potential in soybean: potential targets for biotechnological improvement. *Plant, Cell & Environment* **35**, 38–52. doi:10.1111/j.1365-3040.2011.02378.x

Aramburu Merlos F, Monzon JP, Mercau JL, Taboada M, Andrade FH, Hall AJ, Jobbagy E, Cassman K, Grassini P (2015) Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Research* **184**, 145–154. doi:10.1016/j.fcr.2015.10.001

Boote KJ, Jones JW, Batchelor WD, Nafziger ED, Myers O (2003) Genetic coefficients in the CROPGRO-Soybean model: links to field performance and genomics. *Agronomy Journal* **95**, 32–51. doi:10.2134/agronj2003.0032

Bowman AW, Azzalini A (2014) R package ‘sm’: nonparametric smoothing methods. (ver. 2.2–5.4). Available at <http://www.stats.gla.ac.uk/~adrian/sm> [Verified 10 April 2015]

Concibido VC, La Vallee B, McLaird P, Pineda N, Meyer J, Hummel L, Yang J, Wu K, Delannay X (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theoretical and Applied Genetics* **106**, 575–582.

Denison RF (2009) Darwinian agriculture: real, imaginary and complex trade-offs as constraints and opportunities. In ‘Crop physiology applications for genetic improvement and agronomy’. (Eds VO Sadras, DJ Miralles) pp. 213–231. (Academic Press: New York)

Dornhoff GM, Shibles RM (1970) Varietal differences in net photosynthesis of soybean leaves. *Crop Science* **10**, 42–45. doi:10.2135/cropsci1970.0011183X001000010016x

Evans LT (1993) ‘Crop evolution, adaptation and yield.’ (Cambridge University Press: New York)

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

Gur A, Zamir D (2004) Unused natural variation can lift yield barriers in plant breeding. *PLoS Biology* **2**, e245. doi:10.1371/journal.pbio.0020245

Harrison MT, Edwards EJ, Farquhar GD, Nicotra AB, Evans JR (2009) Nitrogen in cell walls of sclerophyllous leaves accounts for little of the variation in photosynthetic nitrogen-use efficiency. *Plant, Cell & Environment* **32**, 259–270. doi:10.1111/j.1365-3040.2008.01918.x

Hyten DL, Song QJ, Zhu YL, Choi IY, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB (2006) Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16666–16671. doi:10.1073/pnas.0604379103

Jin J, Liu X, Wang G, Liu J, Mi L, Chen X, Herbert SJ (2011) Leaf nitrogen status as a main contributor to yield improvement of soybean cultivars. *Agronomy Journal* **103**, 441–448. doi:10.2134/agronj2010.0344

Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt LA, Wilkens PW, Singh U, Gijsman AJ, Ritchie JT (2003) The DSSAT cropping system model. *European Journal of Agronomy* **18**, 235–265. doi:10.1016/S1161-0301(02)00107-7

Long SP, Zhu XG, Naidu SL, Ort DR (2006) Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment* **29**, 315–330. doi:10.1111/j.1365-3040.2005.01493.x

Mavromatis T, Boote KJ, Jones JW, Irmak A, Shinde D, Hoogenboom G (2001) Developing genetic coefficients for crop simulation models with data from crop performance trials. *Crop Science* **41**, 40–51. doi:10.2135/cropsci2001.41140x

Mercau JL, Dardanelli JL, Collino DJ, Andriani JM, Irigoyen A, Satorre EH (2007) Predicting on-farm soybean yields in the pampas using CROPGRO-soybean. *Field Crops Research* **100**, 200–209. doi:10.1016/j.fcr.2006.07.006

Mikel MA, Diers BW, Nelson RL, Smith HH (2010) Genetic diversity and agronomic improvement of north American soybean germplasm. *Crop Science* **50**, 1219–1228. doi:10.2135/cropsci2009.08.0456

Quick WP, Schurr U, Fichtner K, Schulze ED, Rodermeil SR, Bogorad L, Stitt M (1991) The impact of decreased rubisco on photosynthesis, growth, allocation and storage in tobacco plants which have been transformed with antisense RBCS. *The Plant Journal* **1**, 51–58. doi:10.1111/j.1365-313X.1991.00051.x

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**, 1095–1103. doi:10.1111/j.0014-3820.2001.tb00629.x

- Reynolds M, Dreccer F, Trethowan R (2007) Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany* **58**, 177–186. doi:[10.1093/jxb/erl250](https://doi.org/10.1093/jxb/erl250)
- Reynolds M, Bonnett D, Chapman SC, Furbank RT, Manès Y, Mather DE, Parry MAJ (2011) Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. *Journal of Experimental Botany* **62**, 439–452. doi:[10.1093/jxb/erq311](https://doi.org/10.1093/jxb/erq311)
- Rincker K, Nelson R, Specht J, Slepser D, Cary T, Cianzio SR, Casteel S, Conley S, Chen P, Davis V, Fox C, Graef G, Godsey C, Holshouser D, Jiang GL, Kantartzi SK, Kenworthy W, Lee C, Mian R, McHale L, Naeve S, Orf J, Poysa V, Schapaugh W, Shannon G, Uniatowski R, Wang D, Diers B (2014) Genetic Improvement of US soybean in maturity groups II, III, and IV. *Crop Science* **54**, 1419–1432.
- Sinclair TR, Horie T (1989) Leaf nitrogen, photosynthesis, and crop radiation use efficiency – a review. *Crop Science* **29**, 90–98. doi:[10.2135/cropsci1989.0011183X002900010023x](https://doi.org/10.2135/cropsci1989.0011183X002900010023x)
- Stupar RM (2010) Into the wild: the soybean genome meets undomesticated relative. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21947–21948. doi:[10.1073/pnas.1016809108](https://doi.org/10.1073/pnas.1016809108)
- Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* **277**, 1063–1066. doi:[10.1126/science.277.5329.1063](https://doi.org/10.1126/science.277.5329.1063)
- Wiebold WJ, Shibles R, Green DE (1981) Selection for apparent photosynthesis and related leaf traits in early generations of soybeans. *Crop Science* **21**, 969–973. doi:[10.2135/cropsci1981.0011183X002100060039x](https://doi.org/10.2135/cropsci1981.0011183X002100060039x)
- Xiao J, Grandlillo S, Sang Nag A, McCouch SR, Tanksley SD, Li J, Yuan L (1996) Genes from wild rice improve yield. *Nature* **384**, 223–224. doi:[10.1038/384223a0](https://doi.org/10.1038/384223a0)
- Zhu XG, De Sturler E, Long SP (2007) Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiology* **145**, 513–526. doi:[10.1104/pp.107.103713](https://doi.org/10.1104/pp.107.103713)