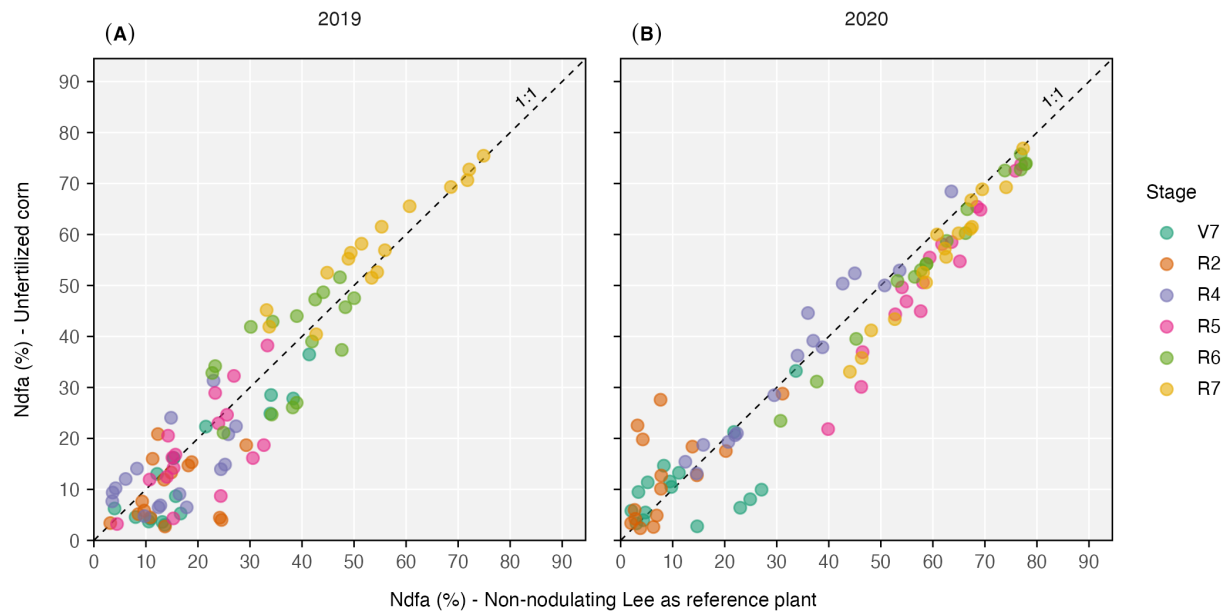
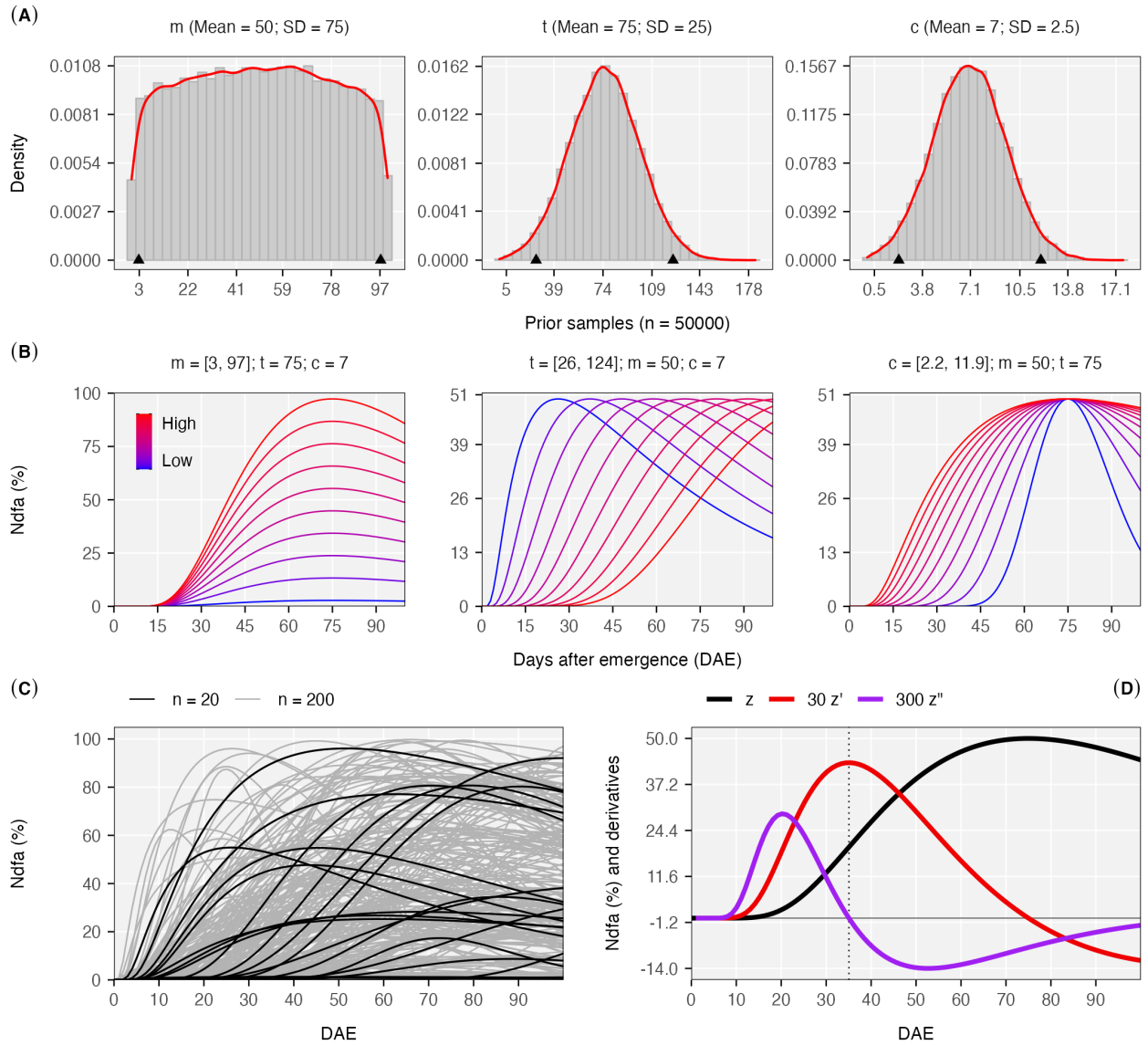


1 **Supplementary material**



2 **Supplementary Figure S1.** Nitrogen derived from the atmosphere (Ndfa, %) estimated using
3 unfertilized corn or non-nodulating soybeans (Lee genotype) as the reference plant. Unfertilized
4 corn was growing in an eight-row strip by the side of the experiments, while the non-nodulating
5 soybeans was randomized along with the four tested genotypes in 2019 (A) and 2020 (B)
6 (Manhattan, KS). Soybean samplings occurred at similar days after emergence in both site-year
7 experiments. Targeted stages were the seventh leaf (V7), full flowering (R2), pod setting (R4),
8 beginning of seed filling (R5), full seed (R6), and physiological maturity (R7).
9



10

11 **Supplementary Figure S2.** Gaussian growth model describing changes on nitrogen derived
 12 from the atmosphere (Ndfa, %) over time (days after emergence). Model parameters are the
 13 maximum predicted value (m), timing of the maximum (t), and a third parameter which controls
 14 growth and decay rates (c). Prior distributions were non-restrictive but introduced biological
 15 knowledge to the model fitting, with black triangles representing the 95% probability intervals
 16 (A). Within the high probability intervals, each individual parameter could produce contrasting
 17 seasonal Ndfa shapes (B). When sampled simultaneously, non-linear parameter priors proved
 18 flexible to describe Ndfa throughout the season (C). Besides the Ndfa true state (z), first (z') and
 19 second (z'') derivatives were used to calculate secondary parameters of interest (D).

Supplementary Table S1. Soybean seed yield, seed size, seed number, protein, and oil concentration for each genotype during the 2019 and 2020 growing seasons (site-years). Values represent the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$) within the Bayesian statistical framework. Uppercase letters compare site-years within genotypes, while lowercase letters compare the genotypes within site-years. Letters were omitted when treatment levels did not differ.

Genotype	Yield Mg ha ⁻¹	Seed size ^a mg seed ⁻¹	Seed number ^b seed m ⁻²	Protein g kg ⁻¹	Oil g kg ⁻¹
2019					
Williams 82	2.90	122 B	2360	435 Aa	206 B
P34T43R2	2.99	132 A	2250	432 Aab	199 B
P35T75X	3.11	124 B	2500	420 Abc	201 B
P37T51PR	2.95	125 B	2320	406 Ac	202 B
2020					
Williams 82	3.24	137 A	2370	410 Ba	228 A
P34T43R2	3.52	142 A	2460	400 Bab	223 A
P35T75X	3.43	137 A	2500	399 Bab	220 A
P37T51PR	3.63	144 A	2500	392 Bb	222 A

^a Soybean seed weight; ^b Seed number was calculated based on seed yield and seed size.

Supplementary Table S2. Logistic non-linear model parameters and means comparison for dry matter (Mg ha^{-1}) accumulation over time (days after emergence). Uppercase letters compare the site-years in 2019 and 2020 (Manhattan, KS), while lowercase letters compare the genotypes. Primary model parameters are the asymptote (m), timing of maximum growth (g), and a third parameter controlling the growth rate (k). Secondary model parameters are the maximum growth rate (r), and the predicted value at physiological maturity (R7 stage) (f). Values represent the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$) within the Bayesian statistical framework.

Genotype	m	g	k	r^a	f^a
2019					
Williams 82	6.98 Aa	65 Aa	0.09 Aa	0.157 Aa	6.57 Ba
P34T43R2	7.74 Aa	71 Aa	0.08 Aa	0.149 Aa	6.78 Ba
P35T75X	8.10 Aa	71 Aa	0.09 Aa	0.192 Aa	7.47 Ba
P37T51PR	7.95 Aa	72 Aa	0.08 Aa	0.170 Aa	7.09 Ba
2020					
Williams 82	8.55 Aa	56 Aa	0.08 Aa	0.173 Aa	8.20 Aa
P34T43R2	9.54 Aa	60 Aa	0.08 Aa	0.202 Aa	9.06 Aa
P35T75X	9.77 Aa	62 Aa	0.08 Aa	0.188 Aa	9.10 Aa
P37T51PR	10.14 Aa	62 Aa	0.07 Aa	0.188 Aa	9.31 Aa

^a Derived quantities based on primary model parameters.

Supplementary Table S3. Logistic non-linear model parameters for total nitrogen (N) uptake, uptake from the soil supply, and from the symbiotic N fixation process (kg ha⁻¹). Uppercase letters compare the site-years (2019 and 2020, Manhattan, KS), while lowercase letters compare the genotypes. Primary model parameters are *m* (asymptote), *g* (timing of maximum growth in days after emergence), and *k* (which controls the growth rate). Secondary parameters are *r* (maximum growth rate), and *f* (predicted value at physiological maturity, R7 stage). Values are the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$) within the Bayesian statistical framework.

Uptake	Genotype	<i>m</i>	<i>g</i>	<i>k</i>	<i>r</i> ^a	<i>f</i> ^a
Total	2019					
	Williams 82	160.3 Aa	55 Aa	0.11 Aa	4.47 Aa	157.7 Aa
	P34T43R2	204.0 Aa	66 Aa	0.07 Aa	3.59 Aa	182.9 Aa
	P35T75X	229.6 Aa	68 Aa	0.08 Aa	4.62 Aa	209.2 Aa
	P37T51PR	181.5 Aa	65 Aa	0.07 Aa	3.43 Aa	165.6 Ba
	2020					
	Williams 82	181.4 Aa	48 Aa	0.11 Aa	5.08 Aa	179.8 Aa
	P34T43R2	221.6 Aa	54 Aa	0.11 Aa	5.99 Aa	218.8 Aa
	P35T75X	224.7 Aa	54 Ba	0.10 Aa	5.33 Aa	220.2 Aa
	P37T51PR	228.0 Aa	58 Aa	0.08 Aa	4.49 Aa	216.3 Aa
	2019					
	Williams 82	96.2 Aa	49 Aa	0.16 Aa	3.80 Aa	95.9 Aa
Soil	P34T43R2	93.3 Aa	49 Aa	0.11 Aa	2.40 Aa	90.8 Aa
	P35T75X	109.8 Aa	57 Aa	0.10 Aa	2.76 Aa	106.1 Aa
	P37T51PR	111.4 Aa	58 Aa	0.07 Aa	2.02 Aa	102.2 Aa
	2020					
	Williams 82	67.1 Ba	32 Ba	0.14 Aa	2.33 Aa	66.5 Ba
	P34T43R2	81.0 Aa	38 Aa	0.11 Aa	2.20 Aa	79.6 Aa
	P35T75X	89.7 Aa	39 Ba	0.12 Aa	2.52 Aa	88.4 Aa
	P37T51PR	86.8 Aa	41 Aa	0.09 Aa	1.82 Aa	83.8 Aa
	2019					
	Williams 82	101.6 Aa	83 Aa	0.08 Ba	2.05 Ba	76.8 Bbc
	P34T43R2	117.9 Aa	82 Aa	0.11 Ba	3.18 Ba	100.3 Bab
	P35T75X	134.3 Aa	83 Aa	0.10 Aa	3.58 Aa	111.6 Aa
Fixation	P37T51PR	76.7 Aa	79 Aa	0.10 Aa	2.04 Aa	66.7 Bc
	2020					
	Williams 82	112.2 Aa	55 Ba	0.23 Aa	6.31 Aab	112.1 Aa
	P34T43R2	133.9 Aa	58 Ba	0.25 Aa	8.20 Aa	133.9 Aa
	P35T75X	120.0 Aa	57 Ba	0.20 Aa	5.87 Aab	119.8 Aa
	P37T51PR	122.9 Aa	61 Ba	0.12 Aa	3.65 Ab	119.9 Aa

^a Derived quantities based on primary model parameters.

Supplementary Table S4. Gaussian growth non-linear model parameters describing seasonal changes on nitrogen derived from the atmosphere (Ndfa, %). Primary model parameters are the maximum predicted value (m), timing of the maximum (t , days after emergence), and a third parameter controlling growth and decay rates (c). Secondary parameters were g (timing of maximum growth), r (maximum growth rate), f (predicted value at physiological maturity, R7 stage), and AUC (area under the curve). Values are the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$) within the Bayesian statistical framework. Uppercase letters compare the site-year from 2019 and 2020 (Manhattan, KS) while lowercase letters compare the four soybean genotypes.

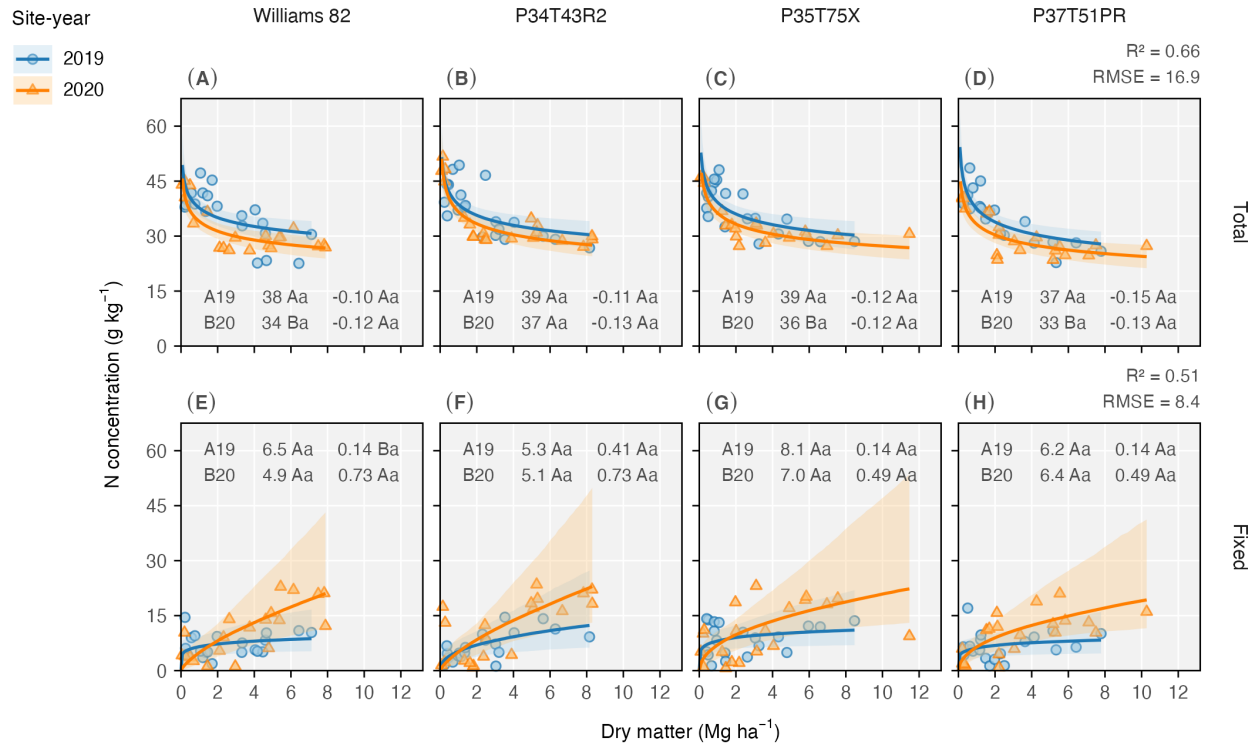
Genotype	m	t	c	g^a	r^a	f^a	AUC ^a
2019							
Williams 82	46 Ba	110 Aa	7.2 Aa	60 Aa	0.96 Ba	44 Ba	15.9 Ba
P34T43R2	57 Aa	116 Aa	6.7 Aa	67 Aa	1.18 Ba	53 Aa	16.3 Ba
P35T75X	51 Aa	110 Aa	7.7 Aa	58 Aa	1.04 Aa	49 Aa	18.2 Ba
P37T51PR	41 Ba	107 Aa	6.8 Aa	61 Aa	0.90 Ba	39 Aa	13.9 Ba
2020							
Williams 82	67 Aa	87 Aa	6.4 Aa	46 Aa	1.74 Aa	64 Aa	31.3 Aa
P34T43R2	67 Aa	88 Ba	5.9 Aa	51 Aa	1.75 Aa	64 Aa	29.0 Aa
P35T75X	62 Aa	87 Aa	6.5 Aa	46 Aa	1.59 Aa	59 Aa	28.5 Aa
P37T51PR	61 Aa	82 Aa	6.9 Aa	41 Aa	1.63 Aa	57 Aa	30.9 Aa

^a Derived quantities based on primary model parameters.

Supplementary Table S5. Gaussian peak non-linear model parameters for the seasonal changes on soil nitrate (NO₃) and ammonia (NH₄). Both soil nitrogen forms were expressed in mg dm⁻³ and assessed over time (days after emergence). Gaussian peak primary parameters are the baseline concentration (*b*), concentration at the peak (*m*), timing of the peak (*t*), and a fourth parameter controlling the growth and decay rates, or width of the peak (*w*). Only the area under the curve (AUC) was calculated as a secondary parameter. Values are the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results (*p* < 0.05) within the Bayesian framework. Uppercase letters compare the site-year from 2019 and 2020 (Manhattan, KS), while lowercase letters compare the four genotypes.

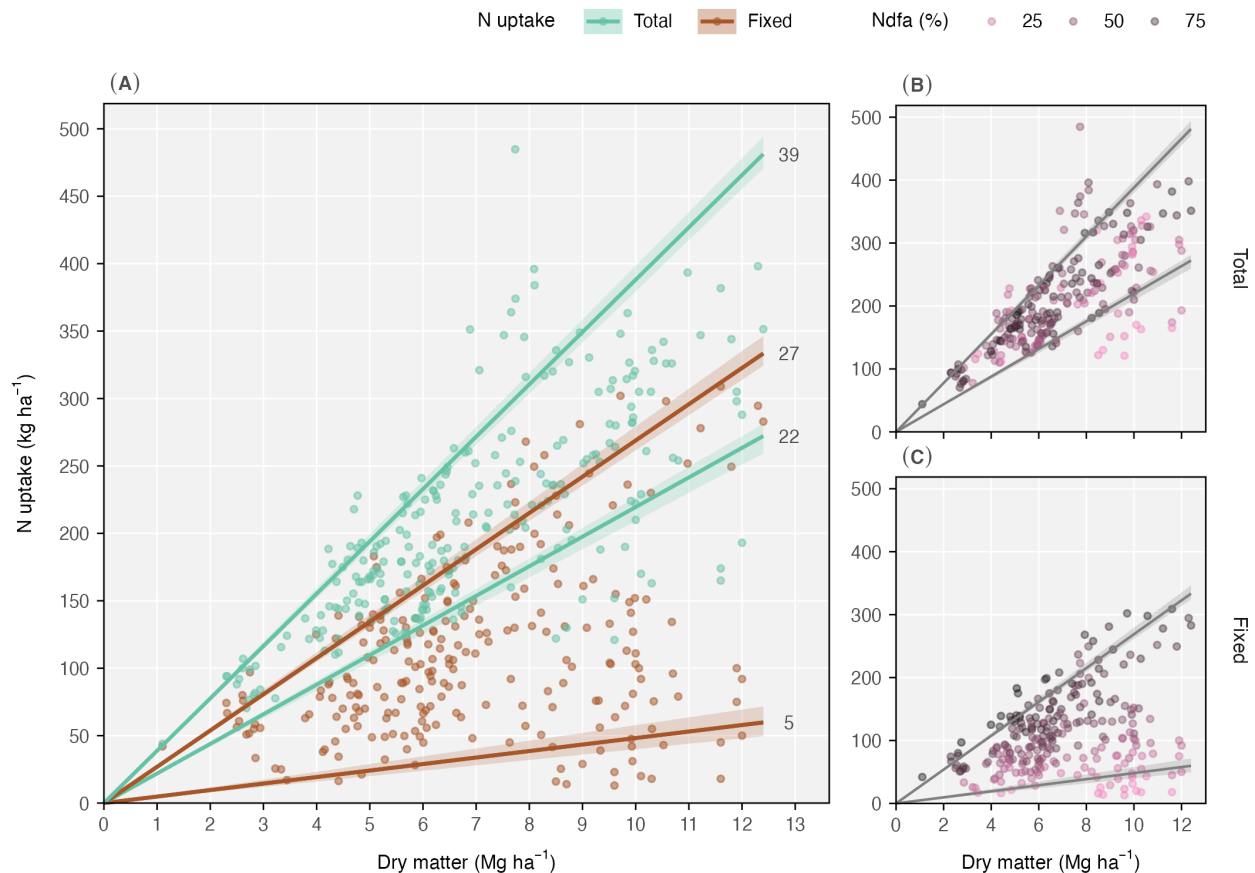
Form	Genotype	<i>b</i>	<i>m</i>	<i>t</i>	<i>w</i>	AUC ^a
NO ₃	2019					
	Williams 82	0.07 Aa	23.81 Aa	50 Aa	18.1 Aa	749 Aa
	P34T43R2	0.07 Aa	14.82 Ab	48 Aa	21.7 Aa	574 Aa
	P35T75X	0.07 Aa	17.13 Aab	48 Aa	20.4 Aa	611 Aa
	P37T51PR	0.07 Aa	19.28 Aab	48 Aa	21.5 Aa	731 Aa
	2020					
	Williams 82	0.07 Aa	13.98 Ba	11 Ba	27.7 Aa	487 Aa
	P34T43R2	0.07 Aa	19.75 Aa	7 Ba	28.0 Aa	635 Aa
	P35T75X	0.07 Aa	22.72 Aa	5 Ba	27.0 Aa	660 Aa
	P37T51PR	0.07 Aa	26.71 Aa	5 Ba	24.9 Aa	700 Aa
	2019					
	Williams 82	6.31 Aa	8.07 Aa	123 Aa	32.2 Aa	693 Aa
NH ₄	P34T43R2	5.88 Aa	10.34 Aa	122 Aa	26.8 Aa	641 Aa
	P35T75X	5.84 Aa	8.70 Aa	128 Aa	31.9 Aa	624 Aa
	P37T51PR	5.89 Aa	6.16 Aa	117 Aa	37.4 Aa	684 Aa
	2020					
	Williams 82	4.37 Ba	7.50 Aa	34 Ba	10.0 Ba	559 Ba
	P34T43R2	4.44 Ba	6.69 Aa	33 Ba	10.4 Ba	557 Ba
	P35T75X	4.22 Ba	6.55 Aa	34 Ba	11.7 Ba	546 Ba
	P37T51PR	4.20 Ba	8.29 Aa	34 Ba	9.0 Ba	544 Ba

^a Derived quantities based on primary model parameters.



65

66 **Supplementary Figure S3.** Dilution curves for total nitrogen (N) (A-D) and fixed N (E-H)
 67 concentration in the aboveground dry matter (shoot). Panel columns represent the four soybean
 68 genotypes and colors represent the site-year from 2019 and 2020 (Manhattan, KS). Exponential
 69 non-linear model parameters are presented and compared within each panel. The first value
 70 represents u (N concentration when dry matter equals to one) and the second value represents
 71 k (which controls exponential growth or decay). Values are the median of estimated marginal
 72 means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$).
 73 Uppercase letters compare site-years and lowercase letters compare genotypes. The last
 74 sampling (at the R7 stage) was removed to avoid potential variation due to leaf senescence.
 75 Coefficient of determination (R^2) and root mean squared errors (RMSE) were presented.



Supplementary Figure S4. Database from Ciampitti and Salvagiotti (2018) showing the relationship between nitrogen (N) uptake (kg ha^{-1}) and aboveground dry matter (Mg ha^{-1}) at physiological maturity (R7 stage). Boundary functions were fitted over the 10th and 90th percentiles for total N uptake and fixed N (A). Nitrogen derived from the atmosphere (Ndfa, %) is presented across different levels of N uptake (B) and fixed N (C). Observations with lower Ndfa are located close to the lower boundary of fixed N and vice-versa, reasonably explaining a wider range of variation for fixed N than for total N uptake across similar dry matter levels.