PNAS

Main Manuscript for

Rewinding evolution in planta: a Rubisco-null platform validates high-performance ancestral enzymes

Vishalsingh R. Chaudhari, Myat T. Lin[1], Kevin M. Hines[2], and Maureen R. Hanson*

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

[1]Present address: C16 Biosciences, New York, NY 10019.

[2]Present address: Foundation Medicine, 400 Summer St, Boston, MA 02210

* Corresponding author- Maureen R. Hanson.

Email: mrh5@cornell.edu

Author Contributions: V.R.C. and M.R.H. designed research. V.R.C., K.M.H, and M.T.L. performed research. V.R.C. and M.R.H. analyzed data and wrote the manuscript. All authors reviewed the manuscript.

Competing Interest Statement: M.R.H. is on the scientific advisory board of Plastomics, Inc. The authors declare no other competing interest.

Classification: Biological Sciences/ Plant Biology

Keywords: Rubisco, CRISPR, chloroplast transformation, ancestral, transgenic plant

This PDF file includes:

Main Text
Figures 1 to 6

Abstract

Improving the photosynthetic enzyme Rubisco is a key target for enhancing C_3 crop productivity, but progress has been hampered by the difficulty of evaluating engineered variants *in planta* without interference from the native enzyme. Here, we report the creation of a Rubisco-null *Nicotiana tabacum* platform by using CRISPR-Cas9 to knock out all 11 nuclear-encoded small subunit (*rbcS*) genes. Knockout was achieved in a line expressing cyanobacterial Rubisco from the plastid genome, allowing the recovery of viable plants. We then developed a chloroplast expression system for co-expressing both large and small subunits from the plastid genome. We expressed two resurrected ancestral Rubiscos from the Solanaceae family. The resulting transgenic plants were phenotypically normal and accumulated Rubisco to wild-type levels. Importantly, kinetic analyses of the purified ancestral enzymes revealed they possessed a 16-20% higher catalytic efficiency ($k_{cat,air}/K_{c,air}$) under ambient conditions, driven by a significantly faster turnover rate ($k_{cat,air}$). We have demonstrated that our system allows robust *in vivo*

assessment of novel Rubiscos and that ancestral reconstruction is a powerful strategy for identifying superior enzymes to improve photosynthesis in C_3 crops.

Significance Statement

The inefficiency of the photosynthetic enzyme Rubisco limits crop productivity. A major challenge in improving Rubisco is testing new variants *in planta* without interference from the native enzyme. We have overcome this problem by creating a Rubisco-null tobacco platform using CRISPR-Cas9 and an optimized chloroplast expression system that enables robust evaluation of novel Rubiscos. Using this platform, we show that resurrected ancestral enzymes assemble efficiently *in planta* and retain faster catalytic properties previously observed in the *E. coli* based system. This work establishes ancestral reconstruction as a promising strategy for discovering compatible and faster Rubiscos, and provides a powerful tool for advancing photosynthetic improvements in crops.

Main Text

Introduction

Global photosynthetic carbon assimilation, estimated at approximately 250 billion tons of CO_2 annually, is primarily mediated by the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (1, 2). Rubisco catalyzes the initial and rate-limiting step in the Calvin-Benson-Bassham (CBB) cycle, fixing inorganic carbon (CO_2) into organic molecules, thus supporting virtually all life on Earth. Despite its pivotal role, Rubisco is notably inefficient, characterized by a low turnover rate (k_{cat}) and a tendency for oxygenation reactions that lead to energetically wasteful photorespiration. In C_3 plants, which make up most terrestrial vegetation, this oxygenation side reaction triggers photorespiration and can reduce net carbon assimilation by up to 20–36% (3).

Rubisco, especially in C_3 plants, has long been a focus for kinetic improvement because of its rate-limiting nature and agronomic importance. (4). Natural Rubisco variants display a wide range of kinetic rates (k_{cat}) and CO2 specificities ($S_{C/O}$) which are typically inversely correlated (5, 6). Among the wide diversity of Rubisco superior outliers exist, suggesting that kinetic properties of C_3 Rubisco could be optimized (5, 7, 8).

Form-I Rubisco, found in plants, cyanobacteria and several other autotrophs, is an L_8S_8 hexadecamer comprising eight large subunits (LSU) and eight small subunits (SSU) (9). In eukaryotes, LSU is encoded by a single chloroplast *rbcL* gene and SSUs are encoded by a family of nuclear *rbcS* genes. In *Nicotiana tabacum* (tobacco), which is an allotetraploid, there are 11 *rbcS* genes, 6 from parent *N. sylvestris* and 5 from *N. tomentosiformis* (10). Rubisco assembly requires several steps and assembly factors (9). Two catalytic sites are formed between two antiparallel LSU strands and a holoenzyme L_8S_8 has 8 catalytic sites (9). Even though SSUs are distant from the active site, it is well known that they play a crucial role in modulating Rubisco's kinetic and structural properties and remain an important target for improvement (11, 12).

Historically, efforts to engineer a faster or more CO_2-specific Rubisco have been thwarted by practical complexities of plant transformation (13, 14). Rubisco requires an array of chloroplast chaperonins and chaperones (such as chaperonins Cpn60/Cpn20 and dedicated chaperones RAF1, RAF2, RbcX, and BSD2) to fold and assemble correctly which are often species-specific, causing incompatibility issues when transgenic Rubiscos are expressed without cognate chaperones (15, 16). Mutagenesis to improve kinetic performance is often restricted by the many necessary interactions between LSUs, assembly factors and SSUs, requiring complementary modifications to these non-catalytic proteins. The breakthrough development of *Arabidopsis thaliana* Rubisco assembly in *Escherichia coli* has significantly accelerated research in this area (17). Subsequently, the process has been optimized for more plant species (8, 11, 18).

Our lab has previously utilized the *E. coli* system to resurrect and characterize ancestral Rubisco enzymes from periods of warmer temperatures and higher atmospheric CO_2 (~30 million years ago), representing the projected environment in this century (19). A significant proportion of these resurrected ancestral Rubiscos exhibited faster catalysis with comparable specificity to the extant enzyme. This crucial finding demonstrates the potential to improve Rubisco kinetics without the canonical catalytic trade-off. Due to rapidly increasing aerial CO_2 concentrations in the postindustrial age, insufficient time has occurred for plants to adapt, and improved kinetic properties may be sought through genetic modifications.

Transferring engineered variants into plants is critical for testing their performance *in vivo* and their broader metabolic integration. A problem in improving Rubisco arises from the presence of robustly expressed native Rubiscos, which can mask the biochemistry of an engineered Rubisco and also potentially result in hybrid Rubiscos (12, 20). Diminishing or removing the expression of native Rubiscos is essential for investigating novel Rubiscos *in planta*.

Here, we report the creation of a Rubisco-null *Nicotiana tabacum* line by CRISPR-Cas9-mediated knockout of all 11 nuclear *rbcS* genes. This provides a "clean slate" for evaluating engineered Rubiscos without background interference from WT enzyme. We demonstrate that co-expressing a single *rbcS* gene with *rbcL* in the chloroplast genome of the knockout line, using our dual-promoter system, restores Rubisco expression to wild-type levels. We evaluated two of our previously characterized ancestral Rubisco combinations in this platform, confirming their efficient assembly and promising kinetic properties. Our system provides a unique opportunity to robustly test performance of compatible Rubisco *in vivo* without background interference, to study individual subtypes, and to make transgenic plants with maternally inherited Rubisco.

Results

NOR-SeLS: Generation of a Rubisco-Null Tobacco Line. To create a platform for evaluating Rubisco variants without interference from the native enzyme, we generated a Rubisco-null *Nicotiana tabacum* line. We began with the markerless transgenic tobacco line SeLSΔaadA, which expresses cyanobacterial Rubisco (*Se_rbcLS*) from *Synechococcus elongatus* PCC7942 and has the wild-type Rubisco LSU gene (*Nt_rbcL*) replaced with the cyanobacterial Rubisco genes in the chloroplast genome (Fig. 1A). SeLSΔaadA plants are able to sustain photoautotrophic growth under elevated CO_2 conditions or with sucrose supplementation (Fig. 1B), enabling disruption of native *rbcS* genes without lethal consequences.

We employed a CRISPR-Cas9 construct with two guide RNAs targeting a conserved region across all *rbcS* genes (*SI Appendix*, Fig. S1). Initial screening of transformants by Sanger sequencing suggested effective knockdown, but subsequent high-throughput sequencing of seeds revealed lower-than-anticipated knockout penetration in the T1 generation, likely due to undetected chimerism (SI Appendix, Fig. S2 and S3). Furthermore, one of the guide RNAs (gRNA #1) proved ineffective.

To enhance knockout efficiency, we replaced the ineffective guide RNA and retransformed the SeLS*-CRISPR #4-C2 line exhibiting the highest number of knockout events and from which the Cas9 cassette had been segregated (*SI Appendix*, Fig. S3). Through several rounds of regeneration and high-throughput screening, we mitigated the chimerism and established a stable knockout line. Analysis of T1 seedlings confirmed complete mutation and frameshift of all *rbcS* genes, with the exception of *rbcS_T2* (*SI Appendix*, Fig. S4). The *rbcS_T2* allele harbored a 6-bp deletion, resulting in the absence of two amino acids and the substitution of another in the predicted protein. However, analysis using *E. coli* expression confirmed that the observed mutation in *rbcS_T2* resulted in a non-functional SSU incapable of assembling into Rubisco holoenzyme and showed no carboxyarabinitol-1,5-bisphosphate (CABP) binding in the extract (*SI Appendix*, Fig. S5). Seeds from Cas9- and kanamycin marker-free plants were collected, and mutations in the T2 generation were confirmed. The final set of mutations is listed in the *SI Appendix*, Fig. S4. These lines, in which all native *rbcS* genes were disrupted and which express only the

cyanobacterial Rubisco large and small subunits, were designated NOR-SeLS (No Other Rubisco except SeLS).

Validating NOR-SeLS Lines. To validate the NOR-SeLS line, we performed chloroplast transformation with the plasmid LSC-NtL, thereby reintroducing the *Nt_rbcL* gene into the chloroplast genome and removing the cyanobacterial Rubisco genes (Fig. 2A, 2B). The resulting plants (NOR-cNtL) exhibited a pale phenotype and slow growth, and were incapable of growing without sucrose supplementation, as expected due to the absence of functional Rubisco in the absence of SSUs (Fig. 2C). We were unable to detect the wild-type LSU, likely due to its instability and degradation in the absence of small subunits (Fig. 2D). We were also unable to detect cyanobacterial LSU indicating its complete removal. In contrast, transformation of wild-type plants with LSC-NtL resulted in normal green plants with no discernible photosynthetic phenotype compared to wild-type (*SI Appendix*, Fig. S6).

Next, nuclear complementation of NOR-cNtL plants was achieved by introducing native *rbcS_S1a* and *rbcS_T1* under the control of their endogenous promoters via Agrobacterium-mediated transformation, generating NOR-cNtL-nS1a and NOR-cNtL-nT1 plants, respectively (Fig. 3A). These transformants displayed a green phenotype (Fig. 3B) and were capable of photosynthesis, overcoming the severe growth defect in the parental NOR-cNtL line. However, they produced lower amounts of Rubisco compared to wild type, likely due to the lower copy number of the nuclear-encoded small subunit genes (Fig. 3C). Interestingly, large subunit levels varied and were proportional to the small subunit expression levels across several nuclear transformants. Furthermore, single insertion transformants segregated into photosynthesis-competent (tT, Tt and TT) and -incompetent seedlings (tt) in an expected Mendelian 3:1 ratio, suggesting that the knockouts are resistant to reversion (*SI Appendix*, Fig. S7). This genetic analysis confirmed that NOR-SeLS and NOR-cNtL lines did not harbor any native *Nt_rbcS* genes.

Chloroplast Expression of Wild-Type and Ancestral Nicotiana Rubiscos Achieves Wild-type-like Growth and Protein Levels. Following the validation of *rbcS* knockout in NOR-SeLS, we proceeded to evaluate chloroplast-based expression of Rubisco enzymes. We designed a chloroplast expression strategy to co-express both the LSU and SSU from the plastid genome. A control construct, LSC-NtLS1, incorporated the native tobacco *Nt_rbcL* gene with its native regulatory elements, followed by a codon-optimized SSU gene (*rbcS_S1a*) and *aadA* selectable marker driven by the *psbA* promoter (Fig. 4A). With biolistic transformation, this construct replaced the cyanobacterial Rubisco module, creating the NOR-cNtLS1 line, which expresses tobacco LSU and SSU exclusively from the chloroplast genome. Similar constructs encoding previously characterized ancestral Rubisco pairs (Sola and Nico) were also integrated into the NOR-SeLS line. Amino acid substitutions distinguishing these ancestral enzymes from wild-type Rubisco are detailed in *SI Appendix*, Fig. S8.

Stable homoplasmic transformants were isolated for all three gene pairs and confirmed by Southern blot analysis (Fig. 4B). All three chloroplast-expressor lines (collectively NOR-c lines) were phenotypically indistinguishable from wild-type (WT) tobacco, displaying robust, healthy growth (Fig. 4C, Fig. 5A). Immunoblot analysis confirmed the accumulation of both plant LSU and SSU at high levels in the NOR-c lines, consistent with the WT control (Fig. 4D), with concurrent loss of reactivity to anti-Se_RbcL, indicating that the cyanobacterial cassette was removed, consistent with the Southern blots. Quantitative growth analyses confirmed that chloroplast transformants exhibited growth rates similar to wild-type plants when cultivated in soil under controlled green room conditions (Fig. 5A, B, *SI Appendix*, Fig. S9). At 67 days post-sowing (dps), NOR-c transformants accumulated fresh weights ranging from ~145 to 159 g, closely paralleling wild-type plants (~141 g). There were no significant differences in the fresh weights of WT and NOR-c rubisco expressing lines at 67 and 74 days. Furthermore, Rubisco content in total soluble protein extracts, assessed through CABP-binding assays of fully activated Rubisco and Bradford protein quantification, was similar or slightly higher in NOR-c transformants (~33–37% Rubisco/TSP) compared to wild-type controls (~31%).

Ancestral Rubiscos Expressed in Transgenic Plants Exhibit Superior Kinetic Properties. To further evaluate the functional consequences of expressing the ancestral enzymes, we first assessed photosynthetic performance at the leaf level using steady-state CO_2 assimilation versus intercellular CO_2 concentration (A/C_i) response curves. Under Rubisco-limiting conditions, the net CO_2 assimilation rates of the NOR-cNico and NOR-cSola lines were similar to those of the NOR-cNtLS1 and wild-type (WT) plants (Fig. 6A). The maximum rate of carboxylation (V_{cmax}) in chloroplast transformants ranged between ~110–127 $\mu mol \cdot m^{-2} \cdot s^{-1}$, closely matching the wild-type rate (~126 $\mu mol \cdot m^{-2} \cdot s^{-1}$). In contrast, homozygous segregants of NOR-cNtL-nS1a line, which have a nuclear-expressed SSU gene, displayed slower aerial photosynthesis (~ 68 $\mu mol \cdot m^{-2} \cdot s^{-1}$) and growth rates, attributed to lower Rubisco accumulation levels relative to wild-type (Fig. 6A, *SI Appendix*, Fig. S9B).

Rubisco activation levels were quantified from leaf extracts used in A/C_i measurements via ^{14}C-CABP - binding assays comparing initial active and fully activated Rubisco content (Fig. 6A). Activation levels in chloroplast transformants ranged from ~50–57%, closely paralleling wild-type plants (~53%).

To directly probe the enzymatic properties underpinning this performance, Rubisco was extracted from wild-type and chloroplast transformants for *in vitro* analysis using ^{14}C-radiometric assays. Rubisco carboxylation turnover rates (k_{cat}) measured by radiometric assays (^{14}C fixation) at 25ºC and ambient oxygen concentration indicated significantly enhanced kinetics in ancestral Rubisco variants. The ancestral Nico and Sola enzymes exhibited higher k_{cat} values of 3.31 s^{-1} and 3.41 s^{-1}, respectively, compared to 2.93 s^{-1} in wild-type plant Rubisco and 2.90 s^{-1} in chloroplast-expressed WT-S1 Rubisco. No significant difference was observed in enzyme affinity ($K_{c,air}$) between ancestral and wild-type Rubiscos. As a result, the overall catalytic efficiency ($k_{cat,air}/K_{c,air}$) of the ancestral Rubiscos was improved by 16–20% over the native WT enzyme. These findings validate that ancestral variants identified through reconstruction experiments can be successfully expressed in tobacco and confer superior kinetic traits.

Discussion

C_3 plants represent a large majority of terrestrial plant species. Many major agroeconomic crops, like wheat, rice, soybean and potato, are C_3 plants. Due to slow catalytic rates of C_3 Rubiscos, C_3 crops have long been a target for increasing carbon assimilation by improving Rubisco characteristics or increasing the enzyme's access to CO_2 (4, 21-24). While recent breakthroughs expressing and evolving Rubisco variants in microbial systems have rapidly identified promising candidates, translating these advances to crops remains a formidable challenge (25). A definitive test of any engineered Rubisco requires its evaluation *in planta*, within the context of the complex metabolic networks within a plant.

A primary obstacle to this goal has been the robust expression of the native Rubisco in host plants. The presence of endogenous large and small subunits can lead to the formation of undesirable hybrid enzymes, confounding kinetic analyses and masking the true performance of the introduced variant. Hence, studying specific Rubiscos *in vivo* and their integration into plant biochemistry requires the removal or suppression of native Rubiscos. Attempts to remove or reduce native Rubiscos have been underway for decades and these efforts have progressively yielded a greater degree of removal of native subunits and provided a wealth of foundational information (26-31). Furthermore, for future strategies involving the introduction of complex structures like pyrenoids or carboxysomes, the complete removal of native subunits is essential to ensure the precise and ordered assembly of the new components (20).

To create our Rubisco-null platform, we performed a complete CRISPR-knockout of all native Rubisco subunits. Using a tobacco line expressing a cyanobacterial Rubisco allowed us to remove the essential, nuclear-encoded *rbcS* genes without causing a lethal phenotype. However, eliminating all 11 copies of the *rbcS* gene family in allotetraploid tobacco presented a significant technical barrier. Unlike previous work targeting only the most highly expressed homologs, our goal was complicated by high sequence similarity across the gene family, which led to frequent gene conversion events during CRISPR-mediated repair.

Although not all the isoforms are expressed at high levels, it was important to knock out all *rbcS* genes since it is unknown how these are regulated and they might serve as a template for reversion. Gene conversion was observed to occur at high frequency in early stages and is especially evident by the similarity of mutations across knockouts. Because the knockout mutations were phenotypically silent in the SeLS background, unlike previous efforts (31), we relied on high-throughput sequencing to uncover and systematically eliminate undetected chimerism through successive rounds of transformation and regeneration. The resulting stable NOR-SeLS line was rigorously validated. The early stage reversions led to final mutations which were mostly identical between alleles and genes, lending a degree of resistance from reversion once the knockout was achieved (*SI Appendix*, Fig. S4B). We showed that no Rubisco was formed when the cyanobacterial LS pair was replaced with the native large subunit (Fig. 2D). Only upon further supplementation with a small subunit did photosynthetic growth occur (Fig. 3). Evidently, large and small subunits are unstable and both are rapidly degraded when not part of the fully assembled Rubisco. There is some evidence that accumulation of chaperone- bound LSU without SSU causes transcriptional inhibition of *rbcL* (32). We have previously seen that unbound SSU is rapidly degraded in chloroplasts (33).

The NOR-cNtL-n lines, with single nuclear insertions, provide a source of rubisco-less NOR-cNtL seeds, which could be useful for generating unique SSU subtypes or combinations (*SI Appendix*, Fig. S7). There is some evidence that the expression of different SSUs varies under different environmental conditions and might provide a form of adaptation (34). It remains unclear whether the L_8S_8 holoenzyme assembles with uniform or mixed SSUs. The NOR lines could be useful for such investigations or for purifying mono-SSU Rubisco.

The NOR-SeLS platform provides an advantage for testing LSU-SSU pairs *in vivo*. We aimed to achieve wild-type Rubisco levels, a goal that had not been met in previous studies using bicistronic *rbcL*-rbcS operons under *rbcL* regulatory sequences. To overcome this limitation, we decoupled *rbcS* expression from *rbcL* and placed it under the strong *psbA* promoter with a customized Shine–Dalgarno sequence. This dual-promoter strategy yielded an important improvement: we were able to express Rubiscos at levels comparable to those in wild type, which may represent either saturating concentrations in C_3 plants (Fig. 4D, 5B) or levels constrained by *rbcL* regulation. Chloroplasts are considered to have ample translational capacity and act as net producers of amino acids and energy for the cell. They have been used to express recombinant proteins reaching several percent of total soluble protein without noticeable effects on growth (35). In our system, total Rubisco abundance was comparable to wild type, indicating that overall protein demand was not increased but rather redistributed from the cytosol to the chloroplast, which we expect to be easily balanced. Consistent with this, no adverse effects on growth were observed. The ability to express both LSU and SSU at WT levels provides a robust and sensitive system to test holistic effects of Rubisco characteristics. *In planta* expression of variant Rubiscos at WT or higher levels, which can saturate the plant metabolic fluxes related to its activity, can also more easily reveal the other tradeoffs or metabolic bottlenecks. Such fluxes would be difficult to attain stably through expression from the nucleus.

We used our platform to test two ancestral Rubiscos, resurrected from a higher CO_2 era and previously shown to have superior kinetics when expressed in *E. coli* (19). As postulated previously, the ancestral enzymes assembled efficiently in tobacco chloroplasts, accumulated and activated to WT-like levels, demonstrating their robust compatibility with assembly mechanisms. The resultant transformants grew with WT-like vigor in air and soil (Fig. 5). Most importantly, the ancestral enzymes exhibited a significantly faster catalytic rate, resulting in a 16-20% improvement in catalytic efficiency (Fig. 6C) similar to trends observed in the *E. coli* system. Interestingly, despite this clear kinetic advantage, the transgenic lines did not show a correspondingly large increase in biomass under controlled growth chamber conditions. This is not entirely unexpected, as photosynthesis in C_3 plants under stable, non-stressful growth chamber conditions may be limited by factors other than Rubisco activity, such as RuBP regeneration (light-

limited). The true benefit of a faster Rubisco may only become apparent under fluctuating field conditions where light, nutrition, temperature, and water-stress can transiently limit or increase photosynthesis.

Our NOR-SeLS Rubisco expression platform provides a clean route not only to introduce better enzymes but also to unravel the next set of metabolic bottlenecks that may constrain carbon assimilation. Our NOR-cSola and NOR-cNico provide an *in planta* validation of our hypothesis that resetting plants to their ancestral Rubiscos from high CO_2 eras can provide a shortcut to naturally optimized and compatible Rubisco variants. By expressing a kinetically superior Rubisco at saturating levels, we can now begin to identify and address the downstream pathways that will need to be optimized to realize the full potential of an improved engine at the heart of photosynthesis. Future work in which transgenic plants are grown in the field will be essential to determine the real-world agronomic impact of these promising ancestral enzymes.

Materials and Methods

Plant Materials and Growth Conditions

Nicotiana tabacum cv. Samsun was utilized as the wild-type (WT) control. The transgenic lines generated in this study were derived from a marker-less line obtained through spontaneous homologous recombination of a direct repeat of the *Nt-TrbcL* terminator from the SeLS transgenic background, which expresses cyanobacterial Rubisco (36). Plants were grown in soil using Lambert LM-111 all-purpose mix within a controlled environment growth chamber (Percival Scientific, USA) maintained at 25ºC, 60% relative humidity, under a 14-h photoperiod supplied by cool white fluorescent lamps (150-200 µmol photons m^{-2} s^{-1}). Plants received full fertilization (21-5-20 NPK) after reaching an approximate height of 10 cm. For specific growth studies, plants were grown on benches under uniform illumination (200 µmol photons m^{-2} s^{-1}; GE MVR400 lamps), received full fertilization, and fresh weight was recorded at designated time points. For sterile tissue culture, seeds were surface-sterilized and germinated on Murashige and Skoog (MS) medium (M5524; Sigma Aldrich) containing 0.8% (w/v) agar (pH 5.8) and 3% (w/v) sucrose in Magenta GA-7 boxes or petri dishes.

Generation and Molecular Analysis of *rbcS* Knockout Lines

Guide RNAs (gRNAs) targeting conserved regions of the tobacco *rbcS* gene family were designed using the CCTOP tool (37). The pAGM4723 binary vector served as the backbone to express Cas9 under the control of the Cauliflower Mosaic Virus (CaMV) 35S promoter and two distinct sgRNAs, each driven by an *Arabidopsis* U6 promoter, as previously described (23). The initial construct was designated pAGM_rbcS_12. To enhance knockout efficiency, a subsequent construct, pAGM_rbcS_23, was generated by replacing gRNA#1 with gRNA#3.

Binary vectors pAGM_rbcS_12 and pAGM_rbcS_23 were introduced into *Agrobacterium tumefaciens* strain GV3101 via electroporation. For tobacco *Agrobacterium*-mediated transformation, leaf explants (approx. 2 cm^2) from sterile 6-week-old SeLSΔaadA plants were briefly immersed in an *Agrobacterium* suspension (resuspended in 1x liquid MS) carrying the desired plasmid and co-cultivated for 2 days on RMOP medium. Following co-cultivation, leaf explants were washed and transferred to selective RMOP medium containing 250 mg/L timentin and 100 mg/L kanamycin. Regenerated shoots were excised and rooted on MS medium supplemented with 100 mg/L kanamycin.

Genomic DNA was isolated from tobacco leaf tissue using the CTAB method. Kanamycin-resistant T_0 plantlets transformed with pAGM_rbcS_12 were screened for *rbcS* mutations by PCR amplification of target loci using gene-specific primers, followed by Sanger sequencing. One T_0 plant, #4 (out of six analyzed) exhibiting the highest knockout efficiency was selected for progeny analysis. High-throughput screening of 88 T_1 progeny from this selected line was performed to assess mutation transmission and chimerism (*SI Appendix*, Fig. S2 and S3). Targeted PCR amplicons covering the *rbcS* loci were sequenced using Illumina MiSeq (2x150 bp paired-end reads) at the Biotechnology Resource Center,

Cornell University. Raw sequence reads were filtered to retain *rbcS* sequences. A custom filtering method employing unique sequence regions was used to classify reads according to specific *rbcS* genes, and the fraction of remaining WT-like sequences within the target regions was calculated to determine CRISPR mutation penetration. Due to observed genetic chimerism in the T_1 generation, a subsequent round of transformation was conducted using pAGM_rbcS_23 on a Cas9-free, segregated T_1 line derived from the initial pAGM_rbcS_12 transformation. Further rounds of regeneration and screening were performed to isolate stable, non-chimeric knockout lines, designated NOR-SeLS. Mutation frequencies and allelic compositions were analyzed using the CRISP-ID tool (38). Progeny of NOR-SeLS (T_1/T_2 generations) were screened by PCR using primers specific for the Cas9 transgene to select Cas9-free and marker-free individuals arising from genetic segregation. The final mutational status of the selected NOR-SeLS lines was confirmed by Sanger sequencing of the target *rbcS* loci (*SI Appendix*, Fig. S4).

E. coli Expression for SSU Functional Assay

The functional impact of the specific mutation identified in the *rbcS_T2* allele was evaluated using an *E. coli* expression system, as previously described (Lin et al., 2020). This system enables the co-expression of the tobacco RbcL subunit, the RbcS variant of interest, and a defined set of chaperones required for plant Rubisco assembly (Cpn60α/β, Cpn20, RbcX, RAF1, RAF2, and BSD2). Soluble protein extracts from these *E. coli* cultures were analyzed for the formation of the Rubisco holoenzyme and its ability to bind the inhibitor CABP.

Chloroplast Transformation

Chloroplast transgenic lines were generated by biolistic transformation of the NOR- SeLS line, replacing the integrated cyanobacterial Rubisco cassette via homologous recombination. The LSC-NtL construct was designed to reintroduce the native tobacco *rbcL* gene (*Nt_rbcL*) and generate NOR-cNtL line. For expression of WT and ancestral Rubisco variants, chloroplast expression vectors were constructed. Coding sequences for the large subunits (LSUs) retained WT codon usage, except for the specific amino acid substitutions defining the ancestral variants. Mature small subunit (SSU) sequences, lacking the native chloroplast transit peptide but incorporating appropriate substitutions for ancestral variants, were codon-optimized for expression in tobacco chloroplasts. Codon optimization was done based on *rbcL* codon usage, similar to Martin-Avila et al (29). The LSC-NtLS1 construct, used to generate NOR-cNtLS1 contained the native *Nt_rbcL* gene under its endogenous regulatory elements, followed downstream by the codon-optimized *rbcS_S1a* gene and the *aadA* selectable marker, both driven by the *psbA* promoter. A g10L 5'UTR sequence with modifications to reduce secondary structures around the Shine-Dalgarno sequence was used for *rbcS*. Analogous constructs for expressing ancestral Rubisco variants #50 Sola-2L1S and #1 Nico-1L1S (LSC-Sola and LSC-Nico for NOR-cSola and NOR-cNico, respectively) were created by substituting the *Nt_rbcL* and *rbcS_S1a* coding sequences with the corresponding mutations (details in *SI Appendix*, Fig. S7) (19). Transformation was performed via particle bombardment following established protocols (39). Selection of transplastomic events utilized RMOP regeneration medium containing 500 mg/L spectinomycin. Successful integration into the chloroplast genome and achievement of homoplasmy were confirmed by Southern blot analysis using a probe hybridizing to the *accD* region.

Nuclear Transformation

For nuclear complementation, genomic fragments containing the native *rbcS_S1a* or *rbcS_T1* genes, including their endogenous 5' and 3' regions, were cloned into a modified pAGM binary vector backbone, replacing the Cas9 and sgRNA expression cassettes. Transformation of NOR-cNtLplants was carried out using the *Agrobacterium*-mediated leaf disc method described above. Transgenic lines were selected on RMOP medium containing 100 mg/L kanamycin. The resulting lines (NOR-cNtL-nS1a and NOR-cNtL-nT1) were assessed for phenotypic rescue and Rubisco protein expression levels. Segregation analysis

of the kanamycin resistance marker in the T_1 generation was performed to infer the transgene copy number and confirm the stability of the *rbcS* knockout background.

Protein Analysis

Total soluble protein was extracted from leaf samples snap-frozen in liquid nitrogen. Tissue was homogenized in an ice-cold extraction buffer (25 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% glycerol, 0.1% Triton X-100, 150 mM NaCl, 1 mM DTT, with protease inhibitors; or 50 mM Bicine-NaOH pH 8.2, 20 mM $MgCl_2$, 1 mM EDTA, with protease inhibitors and DTT/β-mercaptoethanol). Insoluble cellular debris was removed by centrifugation. Protein concentration in the supernatant was determined using the Bradford assay, with Bovine Serum Albumin (BSA) as the standard.

Protein extracts were analyzed by both denaturing SDS-PAGE and non-denaturing (native) PAGE. SDS-PAGE and subsequent immunoblotting were performed using Mini-Protean TGX gels (Bio-Rad) according to manufacturer's protocol. Non-denaturing PAGE was conducted using a Tris-Glycine buffering system at 4ºC according to the manufacturer's protocols. For immunoblotting following SDS-PAGE (Biorad Any kD TGX gels), proteins were transferred to polyvinylidene difluoride (PVDF) membranes. Membranes were blocked using a commercial blocking buffer (Everyblot, Biorad) and incubated with primary antibodies against tobacco LSU (rabbit anti-RbcL, Agrisera, cat# AS03 037), cyanobacterial LSU (rabbit anti-Se_RbcL, kind gift from Dr. Orr, Lancaster University), and tobacco SSU (rabbit anti-RbcS, Agrisera, cat# AS07 259). Detection utilized fluorescently-conjugated secondary antibodies (IRDye 800CW goat anti-rabbit IgG, Thermofisher, cat# 32735). Signals were visualized and quantified using an Odyssey Imager (LI-COR) and associated analysis software (Image Studio, LI-COR).

Rubisco active site concentration was quantified using a [^{14}C]CABP binding assay as described previously (11). Soluble protein extracts were incubated with a molar excess of [^{14}C]CABP. Protein-bound [^{14}C]CABP was separated from free ligand by size-exclusion chromatography (SEC) using Sephadex G50 fine column. Radioactivity associated with the Rubisco-containing fractions was measured by liquid scintillation counting using Ultima Gold cocktail and a Tri-Carb counter (Perkin Elmer). For Rubisco activation status, a protocol described by Choquette et al was followed (40). Briefly, Rubisco from frozen leaf tissue was extracted using extraction buffer (50 mM EPPS-OH, pH 8; 5 mM $MgCl_2$; 2 mM DTT; 1 mM EDTA; 1% PVPP (w/v); 1% (v/v) Plant protease protein inhibitor). 'Initial' number of activated Rubisco was measured by incubating with [^{14}C]CABP for 30 min on ice and then activating using CO_2-Mg buffer (50 mM EPPS-OH, pH 8; 80 mM $NaHCO_3$; 40 mM $MgCl_2$; 1 mM EDTA) in the presence of >100-molar excess of [^{12}C]CABP. The 'total' number of catalytic sites were measured by fully activating Rubisco in CO_2-Mg buffer and binding with [^{14}C]CABP. Both samples were measured by SEC as described earlier.

A/C$_i$ Measurements

Leaf gas exchange was measured using a portable photosynthesis system (LI-COR 6800) equipped with a 2 cm^2 leaf chamber. Measurements were conducted on the youngest fully expanded leaves (typically leaf 5 or 6) of plants of the same age, under controlled environmental conditions: leaf temperature maintained at 25ºC, vapor pressure deficit (VPD) at approximately 1.2 kPa, and chamber flow rate at 300 µmol s^{-1}. The response of net CO_2 assimilation rate (A) to varying intercellular CO_2 concentrations (Ci) (*A/C*$_i$ curves) was determined by systematically altering the reference CO_2 concentration (from ambient down to 50 µmol mol^{-1}, then increasing stepwise up to 1600 µmol mol^{-1}) under saturating irradiance (determined from A/PAR curves, typically 1200 µmol photons m^{-2} s^{-1}). The maximum rate of RuBP carboxylation (V_{cmax}) was estimated by fitting the *A/C*$_i$ response data to established C_3 photosynthesis models using plantecophys R package (41).

In Vitro Rubisco Kinetic Assays

Leaf discs (approx. 2 cm^2) were excised and immediately homogenized in 1 ml of ice-cold extraction buffer (100 mM Bicine-NaOH pH 7.9, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM EGTA, 50 mM 2-mercaptoethanol,

10 mM DTT, 20 mM $NaHCO_3$, Sigma plant protease inhibitor cocktail (1:100 dilution), 1 mM PMSF, 2 mM benzamidine, 5 mM ε-aminocaproic acid) using an ice-cold Dounce homogenizer. The crude extract was clarified by centrifugation at 16,000 *g* for 3 min at 4ºC. The supernatant was desalted using Zeba spin columns (7kD cutoff, Thermofisher) pre-equilibrated with extraction buffer lacking protease inhibitors and PMSF. Desalted extracts were divided into aliquots, snap-frozen in liquid N_2, and stored at -80ºC until analysis. Rubisco carboxylase activity was measured immediately after thawing aliquots using a previously described method (11). Assays were performed at 25ºC in assay buffer (100 mM Bicine-NaOH pH 8.0, 20 mM $MgCl_2$, 0.8 mM RuBP) containing varying concentrations of $NaH^{14}CO_3$. Reactions were initiated by adding 20 µl of the leaf extract and terminated after exactly 1 min by the addition of 200 µl of 1 M formic acid. Samples were dried completely on a heating block at 100ºC. Acid-stable ^{14}C incorporation was quantified by liquid scintillation counting (Ultima Gold, PerkinElmer) after resuspending the residue in 500 µl H_2O. Kinetic parameters (K_c for CO_2 and k_{cat}) were determined by measuring carboxylation rates across a range of dissolved CO_2 concentrations and fitting the data to the Michaelis-Menten equation using non-linear regression analysis.

Data Availability. All relevant data are included in the article, SI appendix or Datasets.

Acknowledgments

This work was funded by the U.S. Department of Energy, Office of Basic Energy Sciences; Chemical Sciences, Geosciences, and Biosciences Division grant DE-SC0020142 (to M.T.L. and M.R.H.); and National Science Foundation MCB-2131582 to M.R.H. We would like to thank David Stern's lab (Boyce Thompson Institute) for lending the Li-6800 instrument for photosynthesis assays. We would like to thank Dr. Laura Gunn (Cornell University) for helpful suggestions and Dr. Douglas Orr (Lancaster University, UK) for kindly providing the anti-Se rbcL antibody.

References

1. C. B. Field, M. J. Behrenfeld, J. T. Randerson, P. Falkowski, Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237-240 (1998).
2. Y. M. Bar-On, R. Milo, The global mass and average rate of rubisco. *Proc Natl Acad Sci U S A* **116**, 4738-4743 (2019).
3. B. J. Walker, A. VanLoocke, C. J. Bernacchi, D. R. Ort, The Costs of Photorespiration to Food Production Now and in the Future. *Annual review of plant biology* **67**, 107-129 (2016).
4. C. Iniguez, P. Aguilo-Nicolau, J. Galmes, Improving photosynthesis through the enhancement of Rubisco carboxylation capacity. *Biochem Soc Trans* **49**, 2007-2019 (2021).
5. A. I. Flamholz *et al.*, Revisiting Trade-offs between Rubisco Kinetic Parameters. *Biochemistry* **58**, 3365-3376 (2019).
6. D. J. Orr *et al.*, Surveying Rubisco Diversity and Temperature Response to Improve Crop Photosynthetic Efficiency. *Plant Physiol* **172**, 707-717 (2016).
7. J. N. Young *et al.*, Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J Exp Bot* **67**, 3445-3456 (2016).
8. Z. G. Oh *et al.*, Unique biogenesis and kinetics of hornwort Rubiscos revealed by synthetic biology systems. *Mol Plant* **17**, 1833-1849 (2024).
9. A. Bracher, S. M. Whitney, F. U. Hartl, M. Hayer-Hartl, Biogenesis and Metabolic Maintenance of Rubisco. *Annual review of plant biology* 10.1146/annurev-arplant-043015-111633 (2017).
10. N. Sierro *et al.*, The tobacco genome sequence and its comparison with those of tomato and potato. *Nat Commun* **5**, 3833 (2014).
11. M. T. Lin, W. D. Stone, V. Chaudhari, M. R. Hanson, Small subunits can determine enzyme kinetics of tobacco Rubisco expressed in Escherichia coli. *Nat Plants* **6**, 1289-1299 (2020).
12. Y. Mao *et al.*, The small subunit of Rubisco and its potential as an engineering target. *J Exp Bot* **74**, 543-561 (2023).
13. M. Gionfriddo, K. Zang, M. Hayer-Hartl, The challenge of engineering Rubisco for improving photosynthesis. *FEBS Lett* **597**, 1679-1680 (2023).

14. S. M. Whitney, R. E. Sharwood, Rubisco Engineering by Plastid Transformation and Protocols for Assessing Expression. *Methods Mol Biol* **2317**, 195-214 (2021).
15. M. T. Lin, M. R. Hanson, Red algal Rubisco fails to accumulate in transplastomic tobacco expressing Griffithsia monilis RbcL and RbcS genes. *Plant Direct* **2**, e00045 (2018).
16. S. M. Whitney, P. Baldet, G. S. Hudson, T. J. Andrews, Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J* **26**, 535-547 (2001).
17. H. Aigner *et al.*, Plant RuBisCo assembly in E. coli with five chloroplast chaperones including BSD2. *Science* **358**, 1272-1278 (2017).
18. J. Archer, M. Kathpalia, B. U. Lee, S. Li, T. Wang, Effects of chaperone selectivity on the assembly of plant Rubisco orthologs in E. coli. *J Exp Bot* 10.1093/jxb/eraf140 (2025).
19. M. T. Lin, H. Salihovic, F. K. Clark, M. R. Hanson, Improving the efficiency of Rubisco by resurrecting its ancestors in the family Solanaceae. *Sci Adv* **8**, eabm6871 (2022).
20. D. J. Orr *et al.*, Hybrid Cyanobacterial-Tobacco Rubisco Supports Autotrophic Growth and Procarboxysomal Aggregation. *Plant Physiol* **182**, 807-818 (2020).
21. M. R. Hanson, M. T. Lin, A. E. Carmo-Silva, M. A. Parry, Towards engineering carboxysomes into C3 plants. *Plant J* **87**, 38-50 (2016).
22. M. Gionfriddo, T. Rhodes, S. M. Whitney, Perspectives on improving crop Rubisco by directed evolution. *Semin Cell Dev Biol* **155**, 37-47 (2024).
23. K. M. Hines, V. Chaudhari, K. N. Edgeworth, T. G. Owens, M. R. Hanson, Absence of carbonic anhydrase in chloroplasts affects C3 plant development but not photosynthesis. *Proc Natl Acad Sci U S A* **118** (2021).
24. B. Forster *et al.*, The Chlamydomonas reinhardtii chloroplast envelope protein LCIA transports bicarbonate in planta. *J Exp Bot* **74**, 3651-3666 (2023).
25. D. H. Loh, L. H. Gunn, A SynBio explosion: a whole new world for Rubisco engineering. *J Exp Bot* **76**, 2593-2597 (2025).
26. G. S. Hudson, J. R. Evans, S. von Caemmerer, Y. B. Arvidsson, T. J. Andrews, Reduction of ribulose-1,5-bisphosphate carboxylase/oxygenase content by antisense RNA reduces photosynthesis in transgenic tobacco plants. *Plant Physiol* **98**, 294-302 (1992).
27. S. R. Rodermel, M. S. Abbott, L. Bogorad, Nuclear-organelle interactions: nuclear antisense gene inhibits ribulose bisphosphate carboxylase enzyme levels in transformed tobacco plants. *Cell* **55**, 673-681 (1988).
28. H. Matsumura *et al.*, Hybrid Rubisco with Complete Replacement of Rice Rubisco Small Subunits by Sorghum Counterparts Confers C(4) Plant-like High Catalytic Activity. *Mol Plant* **13**, 1570-1581 (2020).
29. E. Martin-Avila *et al.*, Modifying Plant Photosynthesis and Growth via Simultaneous Chloroplast Transformation of Rubisco Large and Small Subunits. *Plant Cell* **32**, 2898-2916 (2020).
30. P. Khumsupan *et al.*, Generating and characterizing single- and multigene mutants of the Rubisco small subunit family in Arabidopsis. *J Exp Bot* **71**, 5963-5975 (2020).
31. S. Donovan, Y. Mao, D. J. Orr, E. Carmo-Silva, A. J. McCormick, CRISPR-Cas9-Mediated Mutagenesis of the Rubisco Small Subunit Family in Nicotiana tabacum. *Front Genome Ed* **2**, 605614 (2020).
32. W. Wietrzynski, E. Traverso, F. A. Wollman, K. Wostrikoff, The state of oligomerization of Rubisco controls the rate of synthesis of the Rubisco large subunit in Chlamydomonas reinhardtii. *Plant Cell* **33**, 1706-1727 (2021).
33. M. T. Lin, A. Occhialini, P. J. Andralojc, M. A. J. Parry, M. R. Hanson, A faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**, 547-+ (2014).
34. A. Dedonder, R. Rethy, H. Fredericq, M. Van Montagu, E. Krebbers, Arabidopsis rbcS Genes Are Differentially Regulated by Light. *Plant Physiology* **101**, 801-808 (1993).
35. J. A. Schmidt, J. M. McGrath, M. R. Hanson, S. P. Long, B. A. Ahner, Field-grown tobacco plants maintain robust growth while accumulating large quantities of a bacterial cellulase in chloroplasts. Nat Plants 5, 715-721 (2019).
36. A. Occhialini, M. T. Lin, P. J. Andralojc, M. R. Hanson, M. A. J. Parry, Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO2. *Plant Journal* **85**, 148-160 (2016).
37. M. Stemmer, T. Thumberger, M. Del Sol Keyer, J. Wittbrodt, J. L. Mateo, CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**, e0124633 (2015).

526 38. J. Dehairs, A. Talebi, Y. Cherifi, J. V. Swinnen, CRISP-ID: decoding CRISPR mediated indels by
527 Sanger sequencing. *Sci Rep* **6**, 28973 (2016).
528 39. P. Maliga, T. Tungsuchat-Huang, K. A. Lutz, "Transformation of the Plastid Genome in Tobacco: The
529 Model System for Chloroplast Genome Engineering" in Methods in Molecular Biology. (Springer US,
530 2021), 10.1007/978-1-0716-1472-3_6, pp. 135-153.
531 40. N. E. Choquette, E. A. Ainsworth, W. Bezodis, A. P. Cavanagh, Ozone tolerant maize hybrids maintain
532 Rubisco content and activity during long-term exposure in the field. *Plant, Cell & Environment* **43**,
533 3033-3047 (2020).
534 41. R. A. Duursma, Plantecophys - An R Package for Analysing and Modelling Leaf Gas Exchange Data.
535 *PloS one* **10**, e0143346 (2015).

536

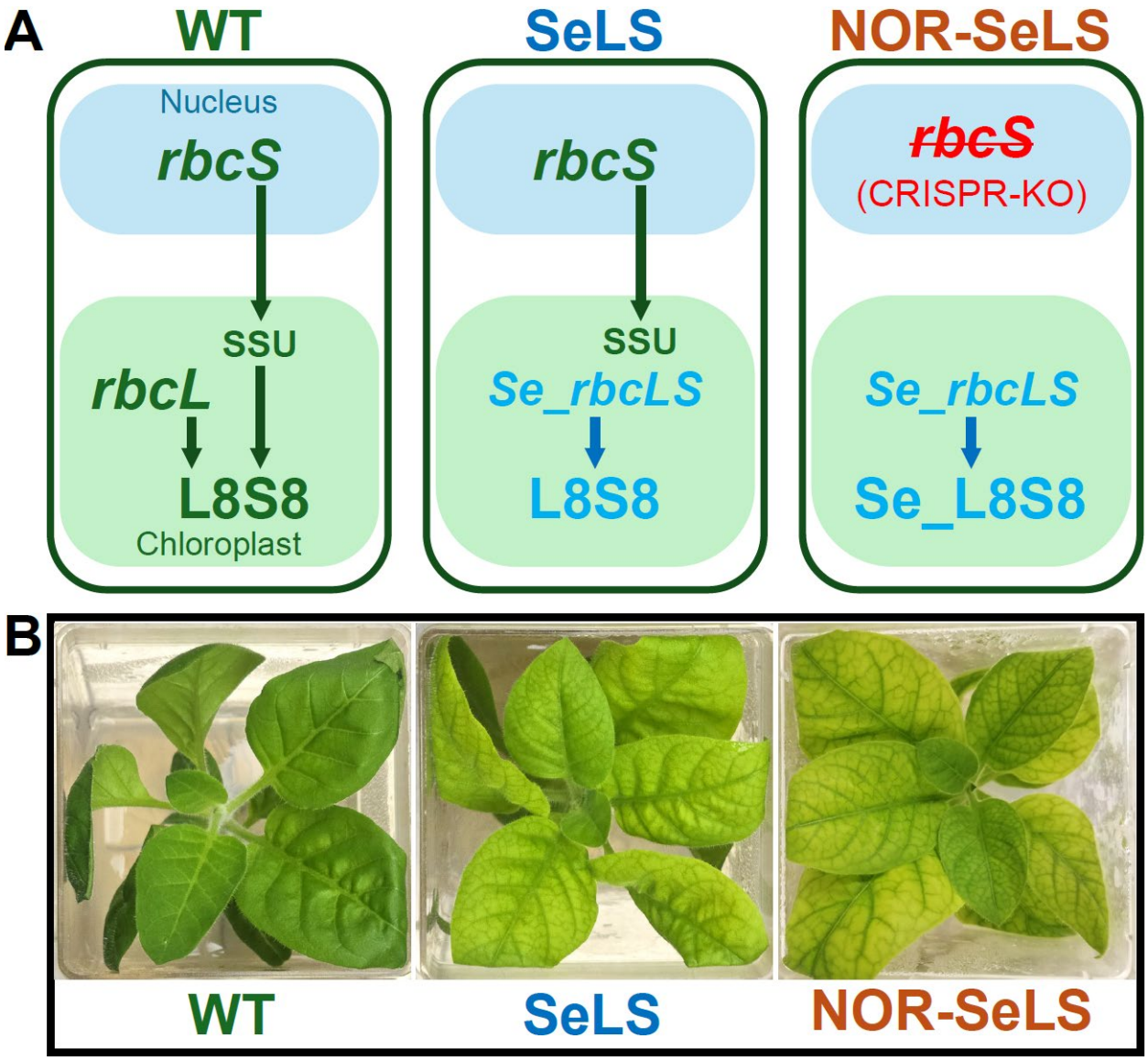


Fig. 1. Generation of the NOR-SeLS *Nicotiana tabacum* line devoid of all native Rubisco sequences.

(*A*) Schematic outlining the genetic background within *N. tabacum* plant cells. The Rubisco large subunit (LSU) is encoded by the chloroplast genome (*rbcL*). Native small subunits (SSUs) coded by nuclear *rbcS* genes are imported into the chloroplast and assemble with LSUs to form the L_8S_8 holoenzyme. A previously developed precursor line, SeLS, was generated by replacing the native *rbcL* with cyanobacterial rubisco genes *Se_rbcL* and *Se_rbcS* in the chloroplast. The NOR-SeLS line (No Other Rubisco except SeLS) was created by CRISPR-Cas9-mediated knockout of all 11 native *rbcS* genes within a marker-less SeLSΔaadA background. (*B*) Comparative phenotype of wild-type (WT), SeLS and NOR-SeLS lines grown photomixotrophically on sucrose medium under ambient air. Similar to SeLS, NOR-SeLS is paler than WT when grown at ambient air in media.

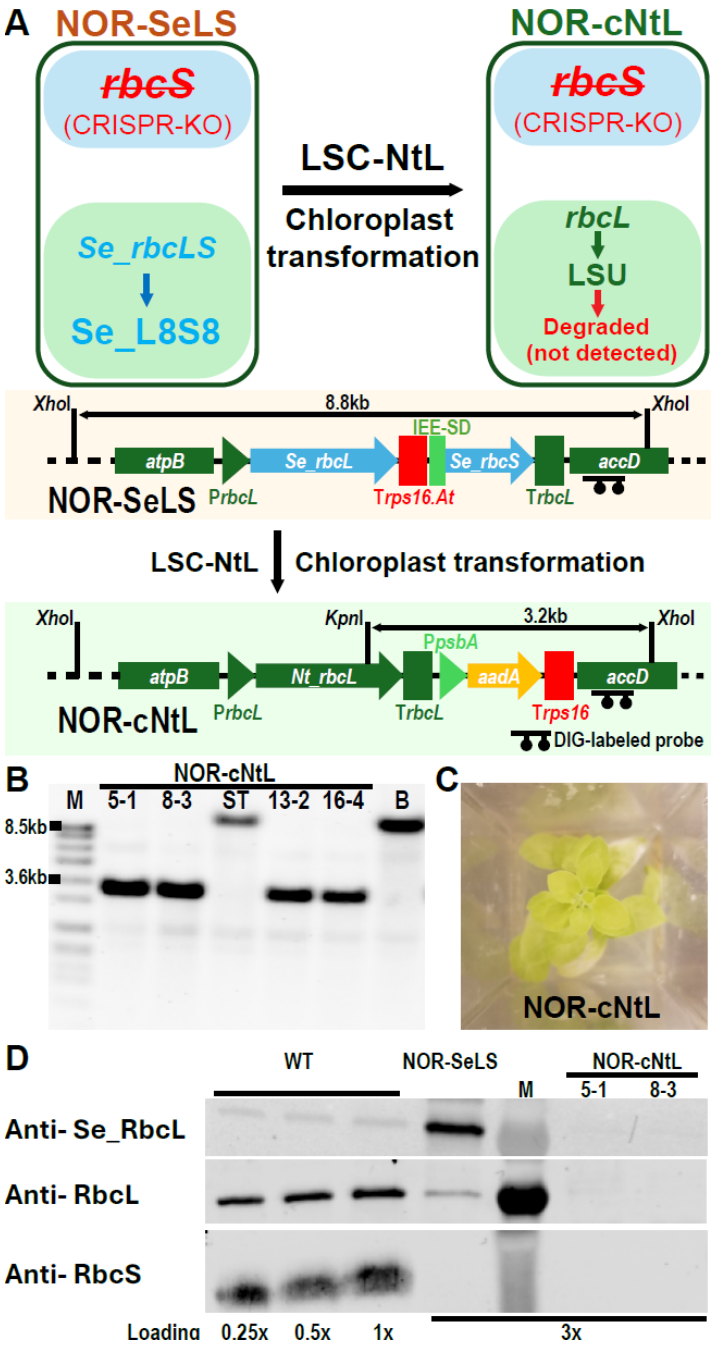


Fig. 2. Validating NOR-SeLS lines. (*A*) The NOR-SeLS line was transformed with the LSC-NtL vector, which replaced the cyanobacterial cassette with the native LSU (*Nt_rbcL*) via homologous recombination. The resulting NOR-cNtL line contains *Nt_rbcL* but lacks any *rbcS* genes, thus preventing functional Rubisco assembly and confirming the knockout phenotype. *atpB* and *accD* mark the flanking region for the *rbcL* locus transformation. P*rbcL* and P*psbA* are native promoters from the corresponding genes. T*rps16*, T*rbcL* and T*rps16.At* are terminators from corresponding genes, with "*At*" marking the terminator from *Arabidopsis thaliana*. IEE-SD is the intergenic region. XhoI and KpnI mark the relevant restriction sites for Southern blot. Probe binds to *accD* region. (*B*) Southern blot confirming the replacement of the 8.8 kb Se_*rbcLS* locus with the 3.2 kb *Nt_rbcL* construct in NOR-cNtL transformants. The parental NOR-SeLS line (Lane B, background) shows an 8.8 kb band corresponding to the *Se-rbcL*S-containing region. Four independent NOR-cNtL transformants (Lanes 2, 3, 5 and 6) exhibit the expected 3.2 kb band, confirming successful integration of the *Nt_rbcL* construct. Lane M contains the DNA size marker and ST is a spontaneous transformant. (*C*) Representative phenotype of NOR-cNtL plants grown on MS medium. (*D*) Immunoblot analysis of total leaf protein extracts. Protein extracts run using SDS-PAGE gel and transferred to PVDF membrane were probed with anti-Se_RbcL antibody specific to cyanobacterial LSU, anti-RbcL specific to plant LSU, and Anti-RbcS specific to plant SSU. The transformant's protein extract was loaded in 3x amounts to allow detection of lower amounts of Rubisco.

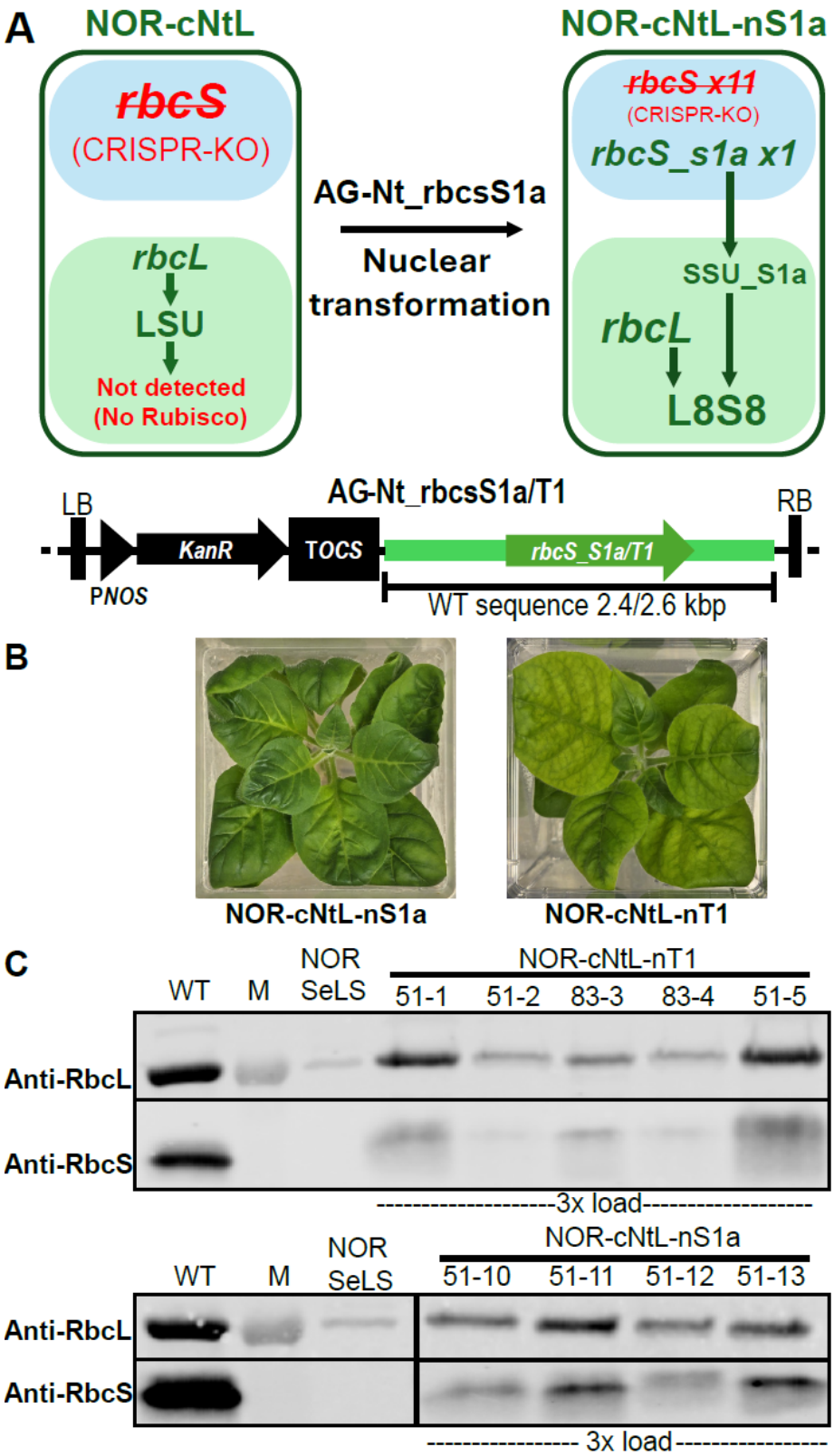


Fig. 3. Nuclear complementation of the Rubisco-deficient NOR-cNtL line. (*A*) Schematic of the plasmid constructs used for *Agrobacterium*-mediated transformation. The constructs encode either the wild-type *rbcS_S1a* or *rbcS_T1* gene under their respective native regulatory sequences. (*B*) Rescue of the NOR-cNtL phenotype by single *rbcS* gene expression. Stable transformants expressing either *rbcS_S1a* (NOR-cNtL-nLS1a) or *rbcS_T1* (NOR-cNtL-nT1) regain green pigmentation and exhibit restored growth on MS medium. (*C*) Immunoblot analysis showing reappearance of the Rubisco in multiple complemented lines. Total soluble protein loaded: 1.6 µg for wild-type (WT) and NOR-SeLS, and 4.2 µg for the complemented lines. M lane has molecular weight markers. Immunoblotting was performed with anti-RbcL and anti-RbcS antibodies.

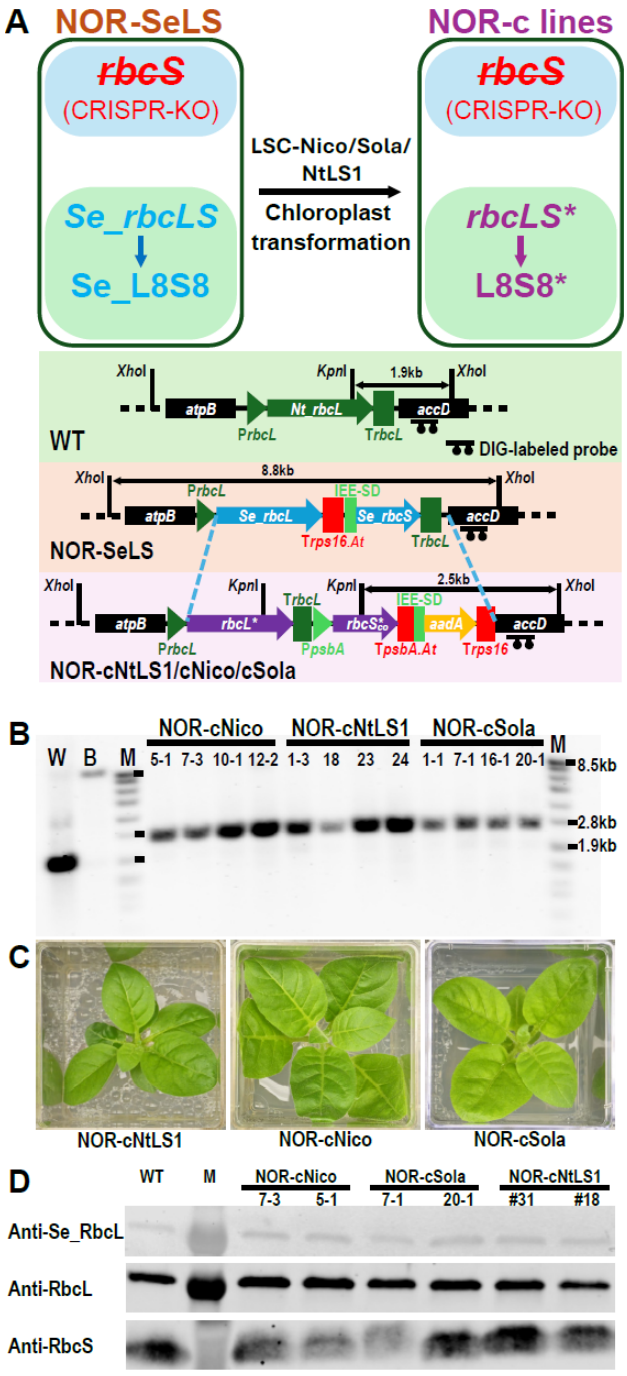


Fig. 4. Engineering and characterization of *N. tabacum* lines expressing Rubisco entirely from the chloroplast genome. (*A*) Schematic representation of the chloroplast *rbcL* locus in WT and chloroplast expressor lines expressing either cyanobacterial Rubisco (NOR-SeLS), wild-type Rubisco (NOR-cNtLS1) or ancestral Rubisco variants (NOR-cNico and NOR-cSola). Chloroplast expressor lines were generated by chloroplast transformation of the NOR-SeLS background, and '*' marks the gene or Rubisco corresponding to the construct used. Blue dashes show homologous recombination for NOR-SeLS transformation. (*B*) Southern blot analysis confirming chloroplast transformation and homoplasmy. Genomic DNA digested with *Xho*I and *Kpn*I was hybridized using a probe for *accD* locus. NOR-SeLS shows a native 8.3 kb band, which is replaced by a 2.5 kb band in the transformed NOR-cNtLS1, NOR-cNico, and NOR-cSola lines, indicating homoplasmic transformation. M, DNA size marker. (*C*) Representative NOR-cNtLS1, NOR-cNico, and NOR-cSola plants grown on MS medium exhibit robust greening and autotrophic growth. (*D*) Immunoblot analysis of Rubisco subunit accumulation. Total soluble protein extracts (0.5 µg) from WT and the chloroplast expressor lines were loaded per lane of the SDS-PAGE gel. Blots were probed with antibodies recognizing the respective LSU and SSU. The engineered lines, NOR-cNtLS1, NOR-cSola and NOR-cNico, accumulate comparable levels of Rubisco compared to WT. M, Protein size marker.

A

B

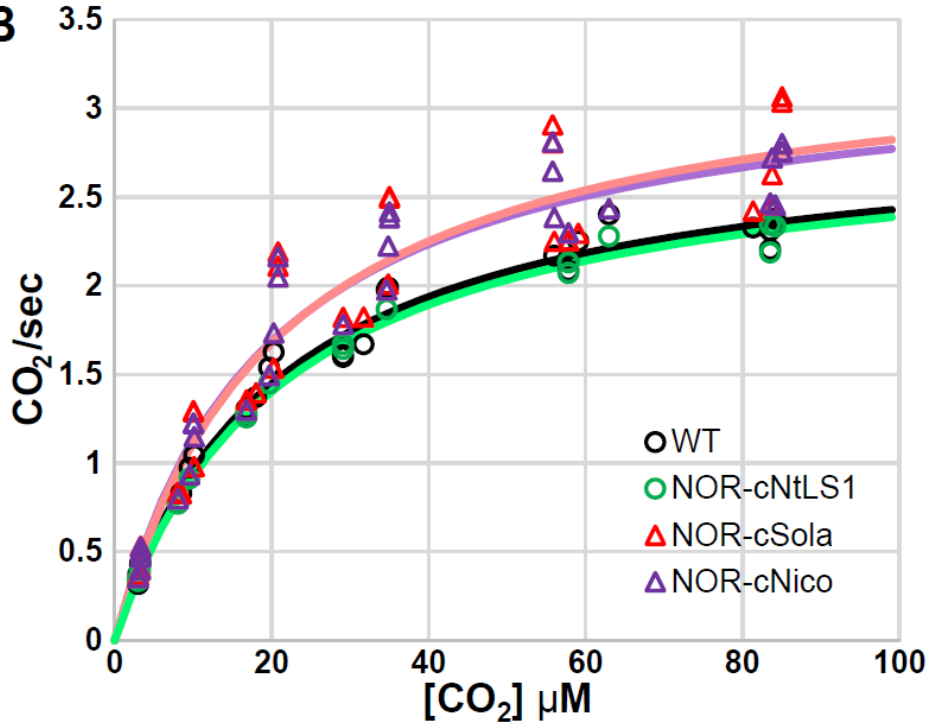
Genotype	Fresh weight (g)		Rubisco/TSP
	67 dps	74 dps	%
WT	141.5 ± 9.9	182.9 ± 12.7	30.7 ± 1.8
NOR-cNtLS1	153.9 ± 4.2	194.4 ± 7.6	33.4 ± 0.8
NOR-cSola	159.3 ± 3.2	207.9 ± 10.6	37.3 ± 1.3
NOR-cNico	145.2 ± 3.3	199.4 ± 6.9	34.7 ± 1.7

Fig. 5. Comparative growth analysis of *N. tabacum* lines expressing native Rubisco exclusively from chloroplasts. Plants were cultivated in soil under controlled environment conditions (ambient CO_2 concentration, ~200 µmol m^{-2} s^{-1} PPFD, 25ºC, with full fertilization). (*A*) Representative phenotypes of wild-type (WT) and the chloroplast expression line NOR-cNtLS1, NOR-cSola and NOR-cNico at 32 days post-sowing (dps) (*B*) Quantitative comparison of biomass accumulation (fresh weight) presented in the table for WT, NOR-C-NtLS1, NOR-C-Sola, and NOR-C-Nico, measured at 67 and 74 dps. Values represent the mean ± SEM for n=6 or more plants per genotype. P-values for one way ANOVA are >0.1

A

Genotype	$V_{cmax,air}$	Activation (%)
WT	126.23 ± 2.7	53.6 ± 2.94
NOR-cNtLS1	110.87 ± 4.4	56.0 ± 2.02
NOR-cSola	129.35 ± 5.3	50.4 ± 1.56
NOR-cNico	127.77 ± 5.8	57.3 ± 1.42
NOR-cNtL-nS1a	68.25** ± 2.4	n.d.

B



C

Assay at Air level [O_2] and 25 °C			
Genotype	$K_{C,air}$ (µM)	$k_{cat,air}$ (s^{-1})	$k_{cat}/K_{C,air}$
WT	20.52 ± 1.05	2.93 ± 0.04	0.145 ± 0.008
NOR-cNtLS1	21.33 ± 0.90	2.90 ± 0.08	0.136 ± 0.004
NOR-cSola	20.63 ± 1.35	3.41 ± 0.14**	0.170 ± 0.018
NOR-cNico	19.27 ± 1.73	3.31 ± 0.05*	0.178 ± 0.018

PNAS

Supporting Information for

Resurrecting Rubisco: Superior kinetics of ancestral enzymes validated *in planta* using a Rubisco knockout tobacco platform

Vishalsingh R. Chaudhari, Myat T. Lin, Kevin M. Hines and Maureen R. Hanson*

* Corresponding author- Maureen R. Hanson.
Email: mrh5@cornell.edu

This PDF file includes:

Supporting text
Figures S1 to S9

Other supporting materials for this manuscript include the following:

Datasets S1 and S2

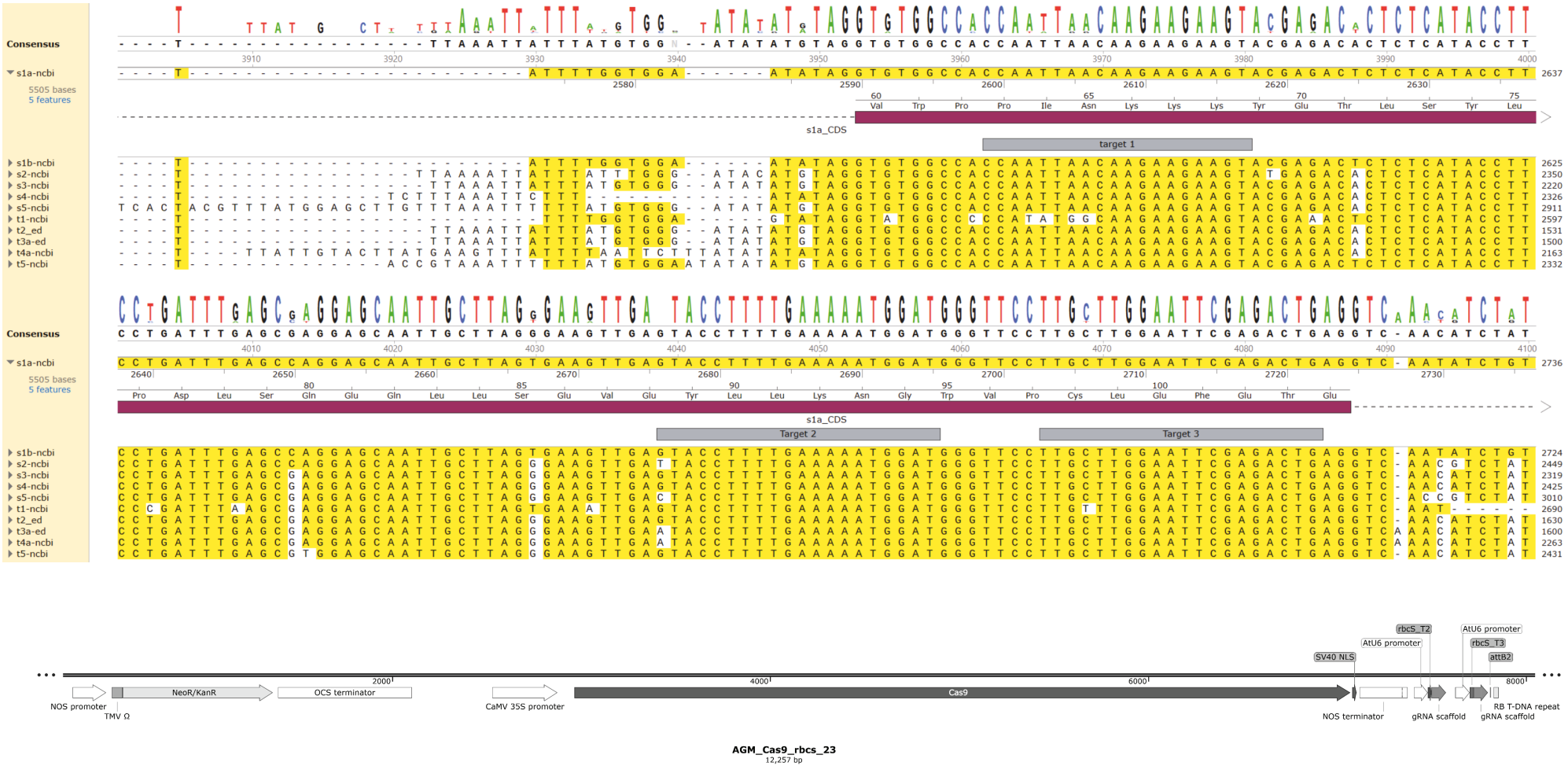


Fig. S1. Target region for CRISPR-mediated mutation. (Top) A nucleotide region highly conserved among *rbcS* gene sequences was targeted using three guide RNAs. The binding sites for guide RNAs T1, T2, and T3 are indicated as targets 1, 2, and 3, respectively. (Bottom) Constructs AGM_Cas9_rbcS_12 and AGM_Cas9_rbcS_23, expressing Cas9 along with pairs of guide RNAs (T1+T2 and T2+T3, respectively), were employed for genome editing.

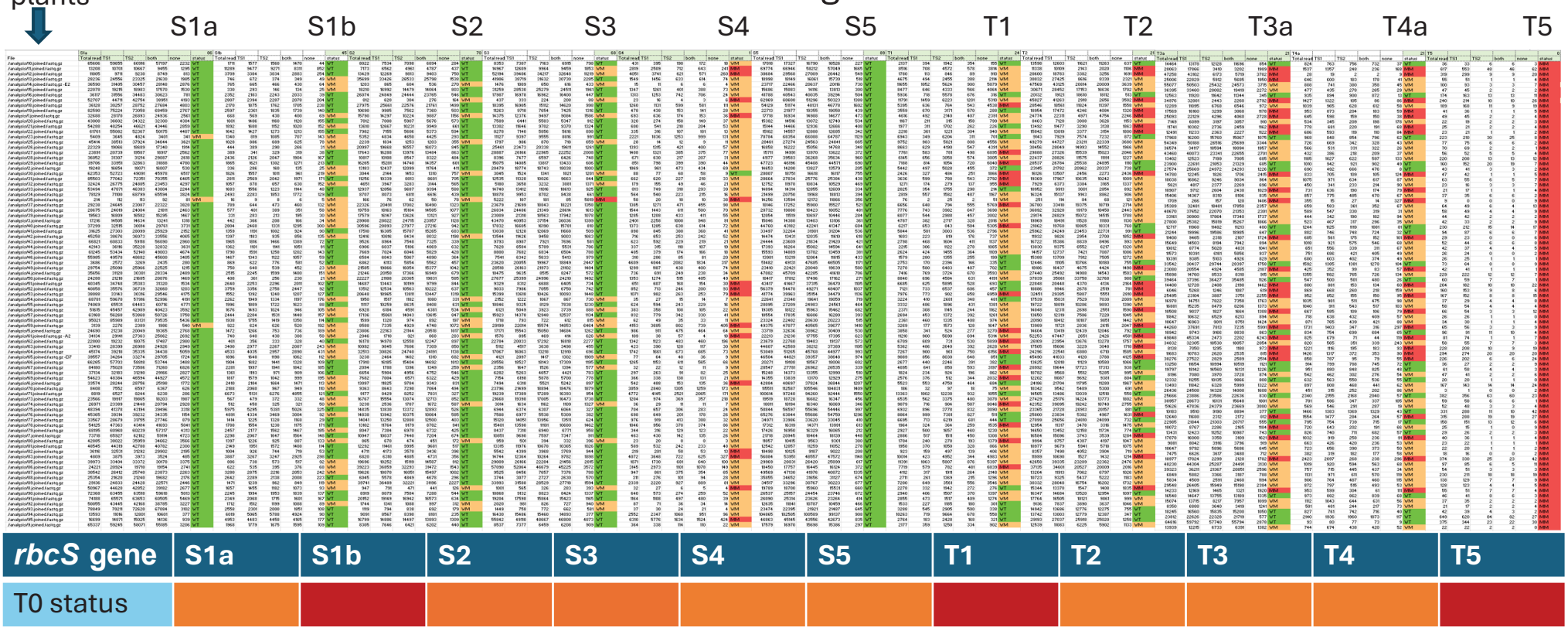


Fig. S2. High-throughput genotyping of *rbcS* mutations in T1 progeny from CRISPR-Cas9 transformant #4. 88 T1 seeds derived from transformant #4 were grown and analyzed for mutations at the 11 targeted *rbcS* loci. Each row in the provided table corresponds to an individual seedling, and columns represent the status of each *rbcS* gene: green = wild-type alleles, orange = heterozygous knockout, and red = biallelic knockout. The mutation detected in the parental T0 plant is shown below for reference. The lower than Mendelian frequency of inherited mutations in the T1 generation reflects somatic chimerism in the T0 plant.

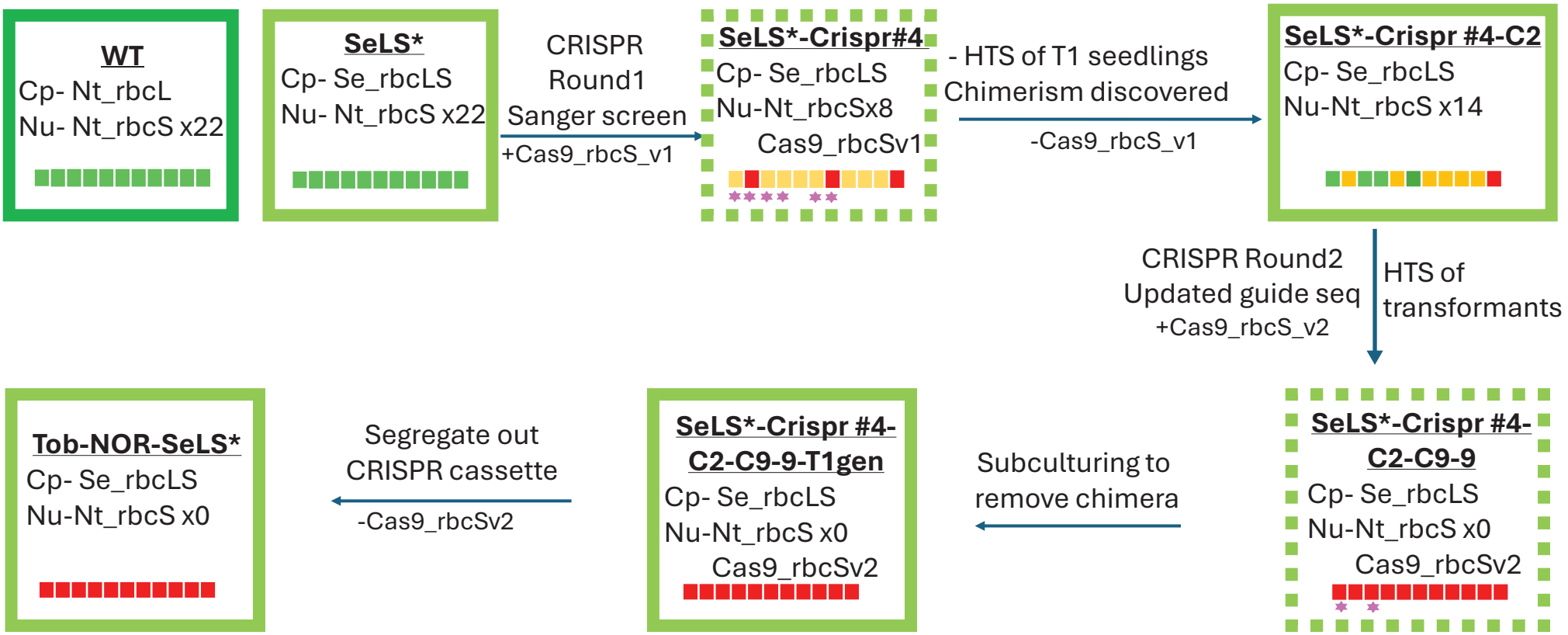


Fig. S3. Workflow for generating the *rbcS*-knockout NOR-SeLS line via iterative CRISPR-Cas9 editing (supports Fig.1).Schematic overview of the multi-step process to eliminate all endogenous nuclear *rbcS* genes in tobacco (shown as 11 green boxes in WT and SeLS). The process started with the SeLSΔaadA line, a marker-free derivative of the SeLS line (Fig. 1) which expresses chloroplast-encoded cyanobacterial Rubisco LSU (*Se_rbcL*) and SSU (*Se_rbcS*). Round 1: The SeLSΔaadA line was transformed with a CRISPR-Cas9 construct expressing guide RNAs (gRNAs) 1 and 2 targeting all *rbcS* loci. Initial Sanger sequencing of *rbcS* loci from six T0 transformants identified plant #4 as carrying most mutations, with heterologous (orange) or biallelic (red) knockouts indicated. However, high-throughput genotyping of 88 T1 progeny derived from #4 revealed incomplete knockout (four *rbcS* loci remained wild-type), low inheritance frequency of mutations, and identified gRNA T1 as ineffective. This discrepancy highlighted significant somatic chimerism in the T0 generation, complicated by the lack of a discernible phenotype for *rbcS* knockouts. Genes behaving chimerically are marked with a purple star. Round 2: A Cas9-free T1-progeny of #4 showing most knockout penetration and segregated Cas9 was selected and re-transformed with a new CRISPR-Cas9 construct expressing gRNAs 2 and 3. High-throughput screening identified transformants with all targeted genes disrupted, which were then repeatedly subcultured to eliminate chimerism. Screening of 10 T1 seeds confirmed a complete knockout across all progeny. Lines free of the CRISPR-Cas9 cassette were selected, and seeds (T2) were harvested and subsequently retested to ensure stable knockouts.

A

rbcS gene	Mutation
S1a	2x 1 bp insertion, frame shift
S1b	2 x 1 bp insertion, frame shift
S2	1 bp insertion, 4 bp insertion + 4(5) bp deletion, frameshift
S3	2x 1 bp insertion, frameshift
S4	2x 1 bp insertion, frameshift
S5	2x 1 bp insertion, frameshift
T1	2x 1 bp insertion, frameshift
T2	6 bp deletion (2 aa deletion, 1 aa substitution), defective SSU
T3	1 bp insertion, frameshift
T4	2x 1 bp insertion, frameshift
T5	3 bp + 4 bp deletion, frameshift

B


s3 mRNA
target 1
Target 2
target3
S4 CDS
s4 mRNA
target 1
target 2
target3
s5 mRNA
target 1
target 2
target3
T1 CDS
target 1
Target 2
Target 3

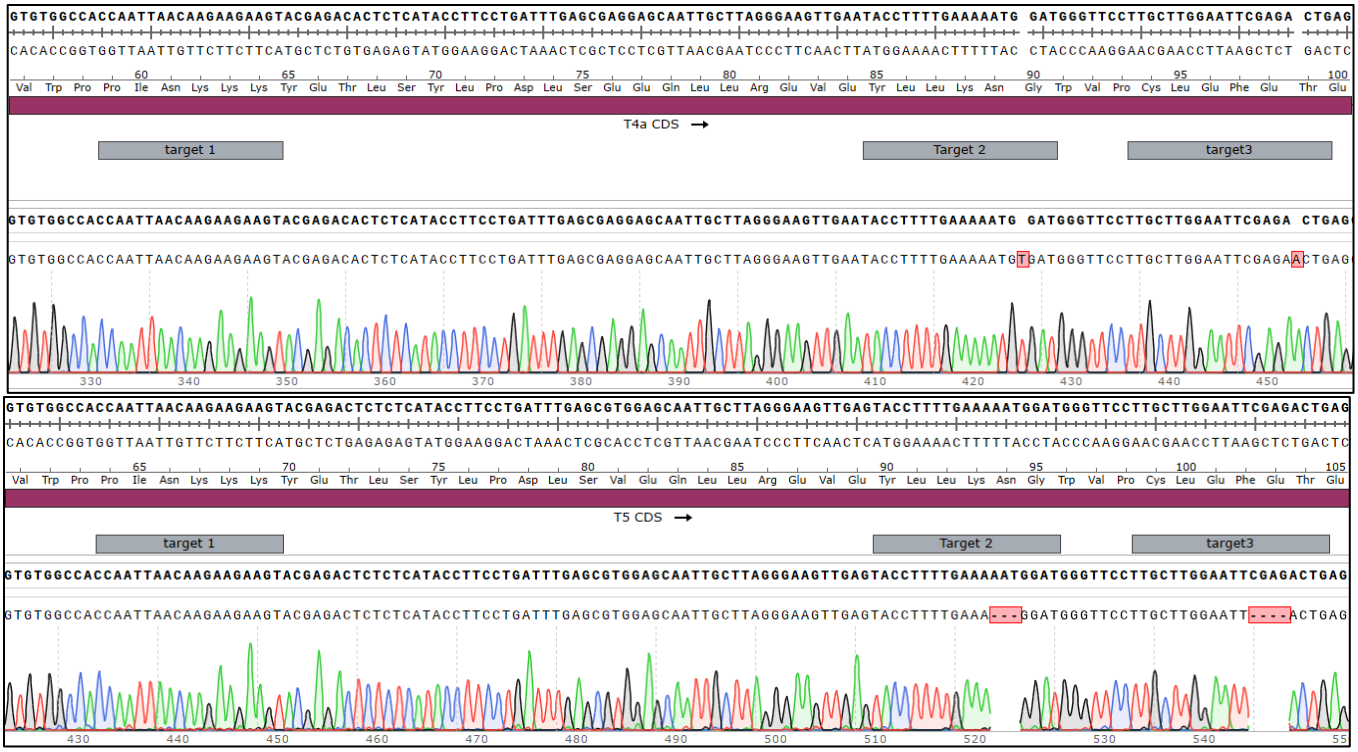


Fig.S4. Characterization of CRISPR-Cas9 induced mutations at targeted *rbcS* loci. (*A*) Table summarizing the specific mutations identified at each targeted *rbcS* locus in the selected NOR-SeLS line #52a. The predominant mutation type across most loci resulted in frameshifts, predicted to cause premature stop codons. A notable exception is the *rbcS_T2*, see Supp. Figure S5. (*B*) Analysis of Sanger sequencing reads covering these loci indicated that the induced mutations were primarily homozygous, with identical mutant alleles detected at most sites. *rbcS_T3* loci could not be amplified in final line and is presumably recombined with other allele.

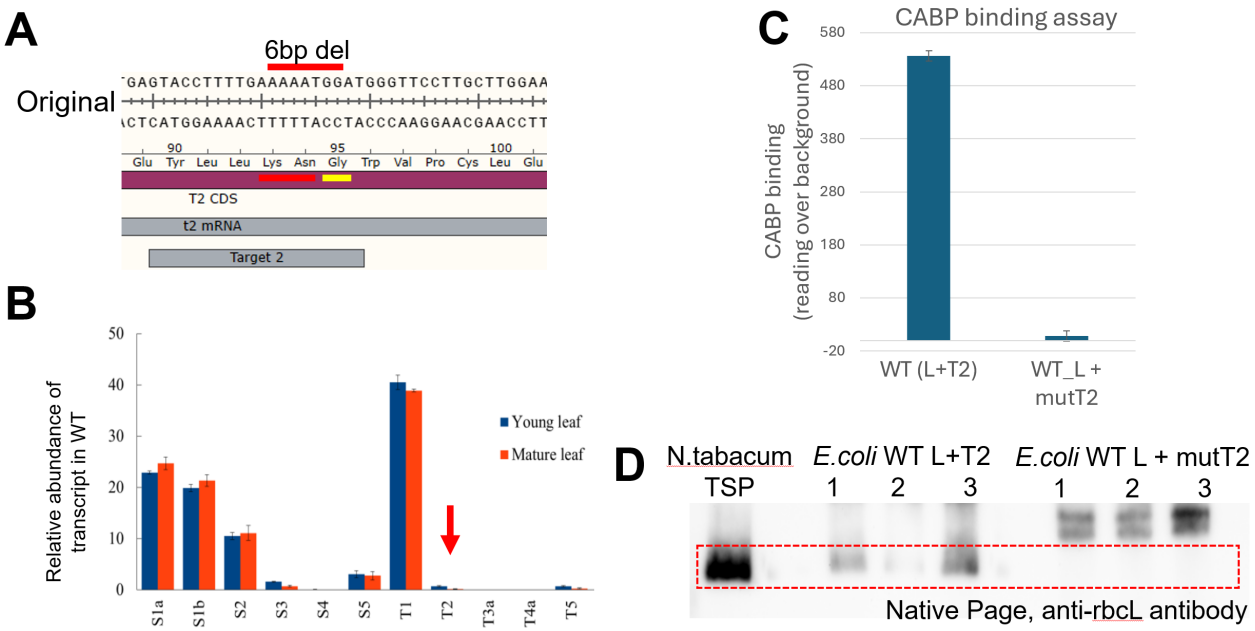


Fig. S5. Functional analysis of the *rbcS_T2* gene product carrying an in-frame deletion. (*A*) Diagram illustrating the 6 bp in-frame deletion identified within the *rbcS_T2* locus. This mutation causes a 2 amino acid deletion and 1 amino acid substitution (Gly95->Arg93) (*B*) Relative expression of endogenous *rbcS* transcript levels in wild-type *Nicotiana tabacum*, indicating low expression of *rbcS_T2*. The reads are plotted from Sequence Read Archive (SRA) data available publicly. (*C*) *In vitro* analysis of Rubisco assembly competence using an *E. coli*-based Rubisco expression system (*Lin et al., 2020 Nat Plants)* expressing LSU alongside either wild-type *rbcS_T2* SSU (T2) or the mutant variant (mutT2). Extracts were assayed for binding to the radio-labeled inhibitor [^{14}C]CABP. Strong CABP binding, indicative of correctly assembled Rubisco active sites, was observed only in extracts containing the WT SSU. (*D*) Rubisco holoenzyme assembly state analyzed by Blue Native PAGE (BN-PAGE) followed by immunoblotting using an anti-RbcL antibody. When co-expressed with WT *rbcS-T2* SSU, LSU predominantly assembled into the L8S8 holoenzyme complex. In contrast, co-expression with the mutated-T2 SSU resulted in impaired L8S8 formation, with LSU accumulating in higher molecular weight forms, potentially representing assembly intermediates complexed with chaperonins.

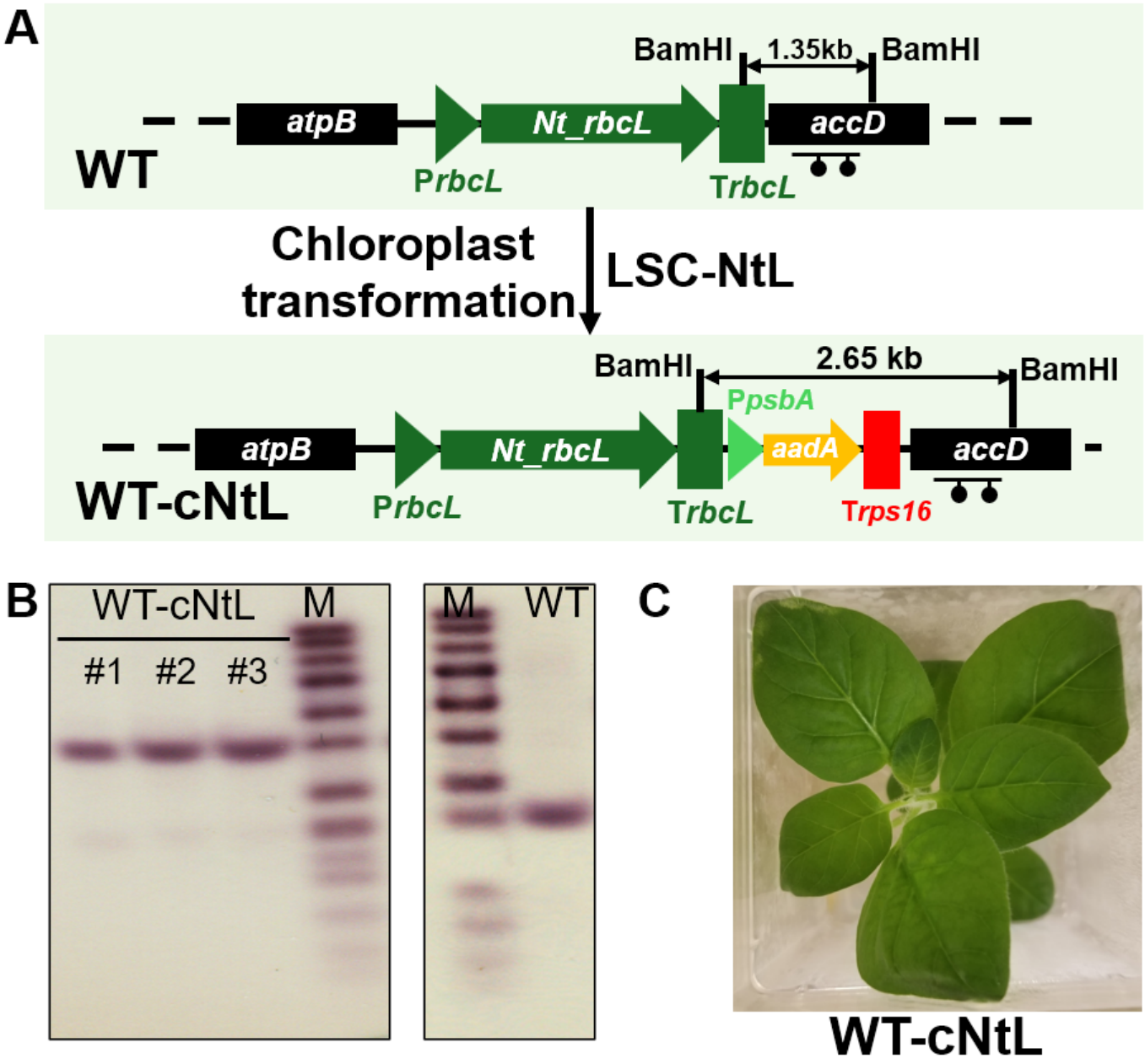


Fig. S6. Control transformation of the LSC-NtL vector into WT *N. tabacum* Samsun for experiment in Figure 2. (*A*) Schematic representation of the WT rbcL locus before and after transformation. *Bam*HI sites are indicated that were used for southern blot validation in (*B*). (*C*) The transformed plants exhibit green leaves and are phenotypically the same as WT.

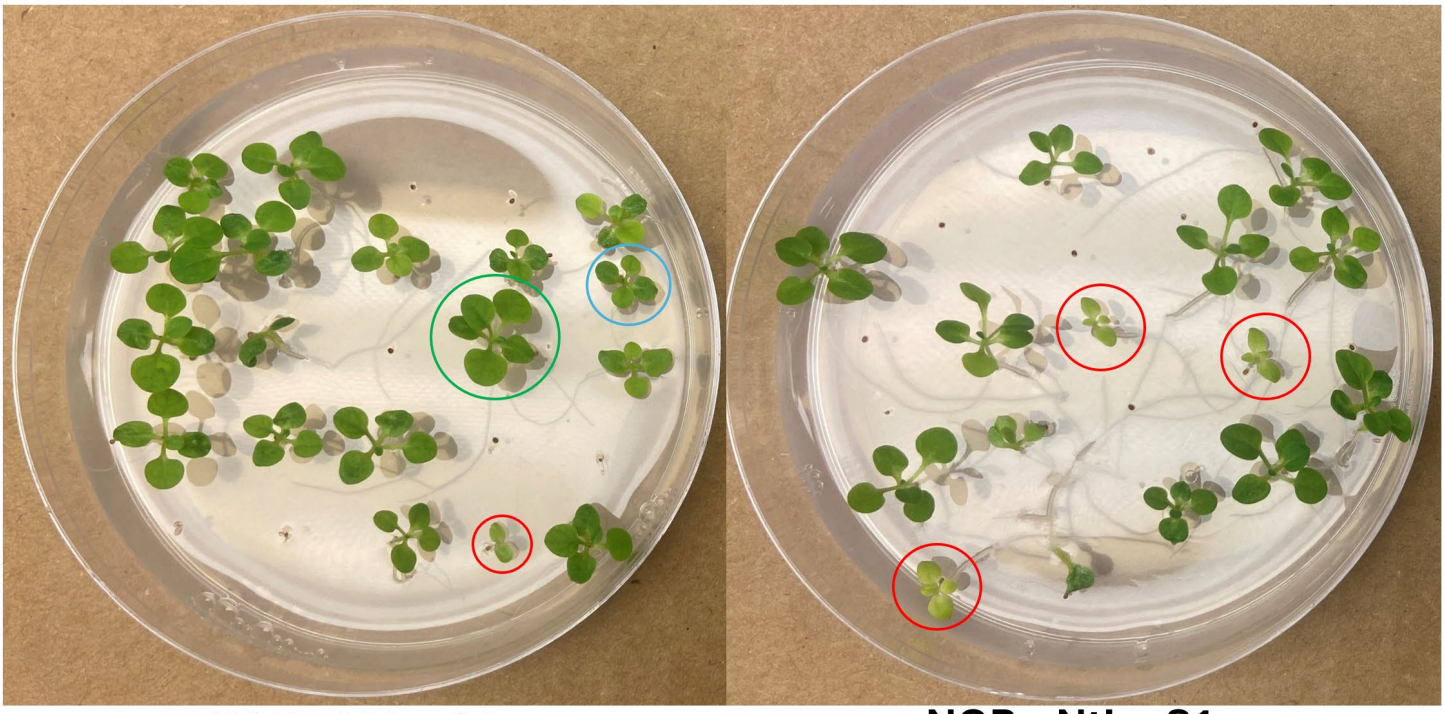


Fig. S7. T1 seeds from the NOR-cNtL-nS1a line were germinated on Murashige and Skoog (MS) medium with and without kanamycin selection. Segregation of the inserted *rbcS* gene is reflected in phenotypic variation: seedlings with green circles indicate heterozygous or homozygous presence of the transgene, while those with red circles represent null segregants that exhibit growth arrest. Notably, *rbcS-null* seedlings continue to grow slowly and produce true leaves on MS medium without kanamycin, suggesting that these lines can serve as a source of viable Rubisco-deficient plants.

RUBISCO	Large subunit mutations	Small subunit Mutations
WT_S1a	reference	reference
Nico_1L1S	L225I, K429Q	N8G, Q23E, S28R, V30I, E88Q
Sola_2L1S	L225I, K429Q	K9M, Q23D, S28K, V30I, E88Q

Fig. S8. A table showing amino acid changes between WT and ancestral Rubisco proteins used in this study.

Fig. S9. Phenotypes of mature wild-type and engineered *N.tabacum* lines expressing Rubisco. Plants were grown in soil under controlled environment conditions (as described for Fig. 5) (*A*) Representative images of wild-type (WT) and chloroplast expressor lines at 76 dps. (*B*) Comparison of representative mature plants photographed at 75 dps: WT, the chloroplast expression NOR-cNtLS1, and the nuclear complemented lines NOR-cNtL-nS1a and NOR-cNtL-nT1.

Dataset S1 (separate file). Raw data and statistical analyses for this study.
Dataset S2 (separate file). Rich text file for plasmid sequences used in the study.