Genetic Variations and Physiological Mechanisms Underlying Photosynthetic Capacity in Soybean (*Glycine max* (L.) Merrill)

2022

MOHAMMAD JAN SHAMIM

Table of Contents

Summary

An increase in the human population to over nine billion people by 2050 will require more than double the current rate of global food production. Using conventional agricultural practices will not assure future crop yield increases as these practices have already been optimized in most modern agriculture. In the green revolution of the 20th century, improvements in crop yield potential were due to the allocation of more biomass to harvesting organs (such as seeds). The culmination of the harvest index (HI) leaves a further increase in yield potential to be bridged by increasing total biomass. Biomass accumulation, however, largely relies on photosynthetic capacity. Compared to HI, photosynthesis has been much less improved. Historically, photosynthetic enhancement was not a part of crop breeding primarily because seed yield was considered sink limited, instruments for measuring photosynthesis were scarce, and photosynthetic phenotyping was costly, labor-intensive, and time-consuming.

Moreover, wild-related germplasms with broader geographical distribution and greater stress tolerance have rarely been used in crop breeding, and neither has their potential been exploited in soybean improvement. These have been some of the limiting factors, especially in exploring beneficial genetic traits for photosynthetic capacity. With high throughput photosynthetic systems, improved breeding techniques, and high-resolution sequencing data, germplasms need to be explored, and valuable materials need to be manipulated/utilized in plant breeding to improve CO_2 fixation. This study used soybean as a testbed for assessing genetic variations in leaf photosynthetic capacity. It is a major commodity crop that is rich in protein and oil. Increasing its yield potential is essential to meet the increasing demand for foodstuff. I conducted this study to evaluate

the photosynthetic capacity of soybean in wild-related progenies and Japanese soybean germplasms and elucidated the physiological mechanisms and genetic architecture underpinning photosynthesis.

Firstly, a hypothesis was assessed to see if the backcross-derived progenies from *Glycine max* (cv. Dwight) and *G. tomentella* Hayata have higher light-saturated photosynthesis (A_{sat}) than their recurrent parent, Dwight. *G. tomentella* Hayata, native to Australia, is a wild-related perennial relative in the tertiary gene pool of soybean. It has diverse morphological, cytological, and genetic nature. It grows in various soil and climatic conditions and is resistant to many diseases and abiotic stresses. An experiment was conducted on two progenies (that showed higher A_{sat} in a preliminary study) and Dwight using a randomized complete block design (RCBD) at Kyoto University. Gas exchange, Rubisco content, nitrogen content, stomatal density, biomass, and yield attributes were measured. As a result, significant increases in A_{sat}, biomass, and seed yield were observed in the progenies relative to Dwight. Physiological mechanism underpinning A_{sat} was also elucidated. No significant differences in stomatal conductance (g_s) between progenies and Dwight, but a higher intercellular CO_2 concentration (C_i) in Dwight suggested some non-stomatal limitations in Dwight. In other words, progenies acquired a much-improved mesophyll activity. Mesophyll activity includes the diffusion of CO_2 from C_i to the site of carboxylation and the fixation of the diffused CO_2 in the Calvin cycle, meaning higher mesophyll activity could be attributed to enhanced mesophyll conductance or the activity, quantity, and specify of enzymes in photosynthetic pathways. It suggests that the progenies might have inherited some advantageous traits from *G. tomentella*.

Secondly, the photosynthetic capacity of the Japanese soybean germplasm was explored, and the mechanisms underpinning A_{sat} were elucidated. Previous studies reported that Japanese soybean cultivars are less productive than the US soybean cultivars.

Studies have also shown that the photosynthetic capacity of the Japanese soybean cultivars is lower than the US soybean cultivars. The overall physiological potential of the Japanese soybean germplasm, however, remains unexplored. A Japanese soybean mini core collection (GmJMC) that represents a significant part of the agro-morphological and geographical trait variations of the Japanese soybean was studied intensively for its A_{sat}. The experiment was conducted for three consecutive years at Kyoto University. In 2019 and 2020, the A_{sat} of 90 and 74 GmJMC accessions was assessed, respectively. A_{sat} was measured three times in 2019 and five times in 2020. A high throughput photosynthesis measurement system, MIC-100, was used in this study. T-distributed stochastic neighbor embedding (t-SNE) was used to reduce the dimensions (8 photosynthetic measurements) of the data to two components (SNE1 and SNE2) which explained 22.1% and 77.9% variations in $A_{sat,}$ respectively. SNE1 was negatively correlated with $A_{sat,}$ whereas SNE2 was positively correlated with the actual measurements. Using the K-means clustering algorithm based on these components, the germplasm was classified into four clusters for their photosynthetic capacity. Cluster2 had highly photosynthesizing accessions, which in many cases had photosynthetic rates higher than the elite cultivars of Fukuyutaka (Fu) and Enrei (En). In 2021, nine representative accessions were selected to identify physiological mechanisms underlying A_{sat}. Leaf gas exchange, biochemical, and morphological parameters were measured three times. Although variations existed, A_{sat} of the newly identified accessions, Fu, and En overlapped. Further analysis revealed that stomatal conductance (g_s) was a major bottleneck in the Japanese soybean. Among Japanese soybean accessions, however, GmJMC47 was identified as the best photosynthesizing accession as it had the highest A_{sat}, stomatal conductance (g_s), stomatal density ($S_{Density}$), electron transfer rate (ETR), and light use efficiency of photosystem II (Fv'/Fm') but the

lowest non-photochemical quenching ($NPQ_{(t)}$). It indicates that GmJMC47 had a higher CO_2 supply and an efficient light-harvesting system.

Lastly, a genome-wide association study and RNA expression analysis were conducted to uncover the genetic inferences of photosynthetic capacity in the Japanese soybean germplasm. The experiment was conducted on 90 and 74 accessions in 2019 and 2020, respectively. A_{sat} was measured eight times in 2019 and 2020. t-SNE was used to reduce the dimensions of the data from eight measurements to two components (SNE1 and SNE2), which were used as phenotypic data representing variations in A_{sat}. Genotypic data was retrieved from a SNP dataset. A genome-wide association study (GWAS) analysis based on SNE components revealed significant associations with a single nucleotide polymorphism (SNP) on chromosomes 14 and 17 for SNE1 and SNE2, respectively. Since SNE2 had a significantly positive correlation with the actual measurements and explained a major part of variations in A_{sat}, it was considered for further genetic analysis. All genes within 250-kbp region spanning both sides of the most significant SNP on chromosome 17 were retrieved through a Phytozome search. Among the candidate genes linked to photosynthesis in the genomic region, variations in the expression of a gene (*Glyma.17G226700*) encoding G protein alpha subunit 1 (GPA1) showed a strong correlation ($r = 0.72$, $p < 0.01$) with that of A_{sat}. Since GPA1 influences stomatal attributes and chloroplast development, it suggests this putative/suggestive gene may be controlling A_{sat} through stomatal and metabolic pathways.

In conclusion, enhancing photosynthesis can potentially improve biomass and seed yield. To do so, inefficiencies in the photosynthetic machinery need to be removed, which will rely upon exploring diverse germplasms for photosynthetic traits and manipulating their genetic architecture underlying photosynthesis. Among several sources of diversity, the use of crop wild relatives is vital for broadening the gene pool of soybean, enhancing

stress tolerance, and boosting physiological and yielding potentials. In this study, enhanced A_{sat} in the progenies reveals that the progenies might have inherited beneficial traits from *G. tomentella*. Similarly, the availability of high throughput systems and high-resolution sequencing data enabled us to assess the photosynthetic capacity of the Japanese soybean germplasm, classify them into four clusters, and identify the genetic architecture and physiological mechanisms underlying photosynthesis. Studies assessing photosynthetic capacity in wild-related soybean progenies and the Japanese soybean germplasm are lacking, making this study one of the first to address photosynthetic capacity in them to this extent. These studies propose that exploring the photosynthetic potential in diverse germplasms could be a promising avenue toward enhancing resource use efficiencies of photosynthesis and biomass productivity in soybean.

Chapter One

General Introduction

1.1. Need for food

The world's population is estimated to grow to over 9 billion by 2050 and will coincide with income growth in nearly all parts of the world. An increase in the human population solely is not a big problem. For instance, a 10% increase in the human population is projected to be responded to by a similar rise in crop consumption and production. The problem, however, begins to appear when per capita income increases. A similar 10% increase in per capita income will result in only one-third of the projected increase in major crops' consumption and production. Agricultural productivity has steadily increased over the past few decades.

Nevertheless, prospects for future yield enhancement seem uncertain (Sands et al., 2014). Despite gains in many parts of the world, the stagnation of seed yield in an estimated 24-27% of the areas where maize, rice, wheat, and soybean are grown underscores the challenge of meeting the rising demands for agriculture (Ray et al., 2012). **Figure 1.1** shows historical and future trends of crop productivity in four major commodity crops. If the yield increase continues at the current pace, it will fall short of the target for 2050.

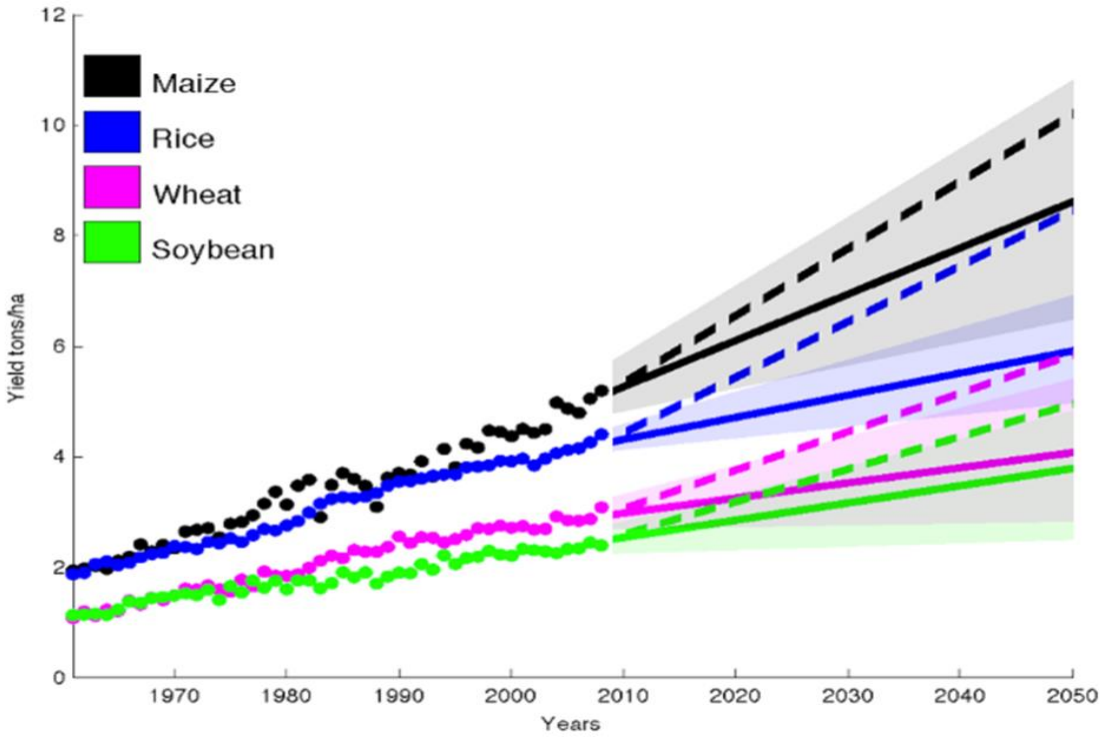


Figure 1.1: Global yield production for four major commodity crops. Observed (closed circles) and predicted (solid lines) area-weighted global yield from 1961 to 2008. Shadings show 90% confidence regions which are derived from 99 bootstraps. Dashed lines represent a trend of ~2.4% annual increase four doubling production by 2050 without bringing more land under cultivation. The base year starts from 2008. Courtesy: (Ray et al., 2012).

Figure 1.1, however, is a deceivingly optimistic prediction because climate change, including rising temperature, drought, flooding, and air pollution, will further aggravate the challenge of food production (Christensen and Hewitson, 2007). These unfortunate reports coincide with studies in which a two-fold increase in crop production is essential to meeting the global demand for food. More specifically, a per capita demand for crops that increases with per capita income will require a 100-110% increase in the global supply of crop yield from 2005 to 2050. The current rate of agriculture intensification in rich countries through more significant inputs and in developing countries through land clearing if from one side adds a billion ha of land to agriculture, it from the other side, results in an enormous amount of gas emission (~3 Gt y^{-1}) and nitrogen use (250 mt y^{-1}). Contrarily, increasing the potential of crops using high-yielding technology will meet the same target with a land clearing of only 0.2 billion ha, one-third of gas emission, and 225

mt y^{-1} use of nitrogen (Tilman et al., 2011). Increasing the potential of crops requires optimizing many physiological and biochemical processes linked to CO_2 acquisition and fixation, water use efficiency (WUE), radiation use efficiency (RUE), and nitrogen assimilation. Several of these processes have their fate selected by photosynthesis.

1.2. Photosynthesis as an avenue to future yield enhancement

To keep a balance between resources and sufficient crop production, the practice of sustainable agriculture becomes imperative (Maurino and Weber, 2013). Photosynthesis is one way to produce enough food, practice sustainability, and evade the exploitation of natural resources (Janssen et al., 2014). Photosynthesis is a process in which plants convert sunlight into biochemical energy that subsequently drives all life on the earth (Evans, 2013). This process accounts for 90- 95% of the plant's aboveground biomass (Da-Yong et al., 2012). Around 40% of plants' dry mass consists of carbon derived from photosynthesis. Photosynthesis nourishes plant growth cycles and is fundamental to the energy demand of all life on earth (Lambers et al., 2008). Increasing plant productivity through enhancing photosynthesis is an enticing approach. The Monteith (1977) formula of yield potential best explains how light conversion efficiency (ε_c) thus photosynthesis is a trait of future yield enhancement.

$$P_n = (S_t . \varepsilon_c . \varepsilon_i)/k$$

$$Y_p = \eta . P_n$$

Here, P_n stands for primary production. S_t denotes cumulative solar radiation in a growing season. ε_i represents the efficiency of light interception by a plant canopy. k is the plant energy content which depends on the oil content of the grain. Y_p is the yield potential, and η is the efficiency through which biomass is allocated to seeds (partitioning efficiency or Harvest Index [HI]). Experiments have documented all parameters except the ε_c

approach at their theoretical peak (Evans, 1997; Long et al., 2006; Zhu et al., 2010). For instance, there is little interest in prolonging plant growth duration to avoid interference with multiple cropping; thus, little chance of increases in S_t is available (Mitchell and Sheehy, 2000). Moreover, the fertilizer application rate has been optimal (Makino, 2011).

Plant morphological characteristics have also been optimized. They have little or no potential for intercepting a large amount of incident radiation (Peng, 2000), leaving ε_c as the only target for crop yield improvement (Mitchell and Sheehy, 2000). Theoretically, the ε_c is 5-6%; however, without some exceptions where ε_c is 3.5%, it is only 1% in the present situations (Mann, 1999) or 0.024% and 0.034% for C_3 and C_4 crops, respectively (Long et al., 2006). This means that, from around 90% of intercepted solar radiation, only 1% is converted into biomass, which indicates the enormous potential plant possesses to produce greater biomass if ε_c is optimized (Da-Yong et al. 2012). Improving photosynthesis, therefore, is one of the fundamental processes that can enhance ε_c.

Interestingly, biochemical, molecular, and microclimate-based empirical photosynthetic models reveal that genetic engineering of photosynthetic machinery and the exploration of variations in plant germplasm may increase crop productivity and ensure agricultural sustainability (Evans, 2013; Gu et al., 2014; Reynolds et al., 2011; Zhu et al., 2010). Theoretical calculations have also revealed that the genetic engineering of RUE could enhance wheat productivity by 50%. Turning this productivity into an agronomic context will require improving crops' structural and reproductive aspects (Reynolds et al., 2002). The improvement of RUE, nevertheless, largely depends on optimizing physiological and biochemical processes that define the fate of CO_2 uptake and its utilization, making photosynthesis a key target (Long et al., 2006).

The controversy of whether seed yield is limited by source or sink has long been on the table and needs clarifications (Evans, 2013). Photosynthetic assimilates that exceed the

need of leaves, load into the phloem, and flow toward the sink. The accumulation of assimilates in the sink collectively determines crop yield. Crop yield, in turn, is affected by environmental and metabolic factors; thus, sink and assimilation (source) should have a balance (Rossi et al., 2015). However, recently developed hybrid rice cultivars show sink limitations when examined in elevated CO_2 studies. Kromdijk and Long (2016) reviewed supporting evidence for the hypothesis that increases in photosynthesis increase yield. In Free-Air CO_2 Enrichment (FACE) studies, the practice of sink limitation can be easily simulated. In these experiments, the open-air CO_2 is elevated to the desired level from sowing to harvest. Both photosynthetic and yield performance of crop species are compared with ambient CO_2 conditions. Theoretically, increasing the concentration of CO_2 from 400 ppm to 550 ppm should result in a 30% increase in seed yield, which became a reality using newly developed hybrid cultivars of rice.

On the other hand, super hybrid rice cultivars have produced lower yields; the main problem was the poor filling of caryopses in their inferior spikelet. This shows that the availability of assimilates is becoming a problem in cultivars with improved sink. Sink that was once a major barrier to higher yield achievement turns out to be no longer rate-limiting, suggesting that genetic improvement for yield potential would require enhancement of source, photosynthesis.

1.3. Problems within photosynthetic machinery

Photosynthesis forms the basis of plant growth, and its enhancement has the potential to improve plant productivity, thus contributing to greater food security. Although plant growth depends on this process, photosynthetic rates do not necessarily interpret variations in growth rate. This is due to water and nutrient supply limitations, biomass acquisition for resource uptake, transpiration, and many more. The exhaustion of conventional yield-

improving parameters has brought the concept of digging into photosynthesis to raise crop yield (Evans, 2013; Lambers et al., 2008).

Improving photosynthesis, however, is a complicated process and requires the removal of inefficiencies inside and outside of this machinery. Upon illumination, for instance, three main phases, such as electron transfer, activation of enzymes, and stomatal opening limit leaf photosynthesis (Yamori et al., 2020; Yamori and Shikanai, 2016). Numerous studies have reported such limitations in various conditions. A heat treatment study in field-grown Pima cotton, for example, indicated that photosynthesis was functionally limited by electron transport responses and Ribulose 1, 5-bisphosphate (RuBP) regeneration during and after heat treatment (Wise et al., 2004). With the current CO_2 levels reaching 400 μmol m^{-2} s^{-1}, light-saturated photosynthesis is rate-limited by electron transport rather than carbon fixation (Sukenik et al., 1987). Many studies have reported ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), which catalyzes carboxylation and oxygenation, as a major limiting factor. Despite being a critical enzyme that transforms atmospheric CO_2 into biomass, Rubisco is a slow and, in some cases, deceiving catalyst that restricts productivity and leads to the waste of resources such as water and nutrients (Long et al., 2006; Whitney et al., 2011). These are a few examples of mounting research that suggests any expectation of seed yield improvement through photosynthesis would first need to remove inefficiencies within this system.

1.4. Targets for improving photosynthesis

So far, several specific targets have been identified to improve photosynthesis. On a canopy level, the adjustment of plant canopy architecture (leaf angle, plant density, number of branches, internode elongation, etc.) to intercept and accumulate more radiation, avoid self-shading, become less prone to biotic stresses, and become easy for foliar

application of chemicals is suggested to be promising. This target is not ideal for crops that have already achieved an optimum canopy architecture but may work for many other crop species (Chavarria et al., 2017). Stay green (SG) is another valuable secondary trait that maintains green leaves and photosynthesis for longer after anthesis, especially under stress. Functional SG, due to the delayed loss of chlorophyll, is different from the green stem disorder (GSD or unfunctional SG), where chlorophyll degradation is a problem, thus making it an essential agronomic trait (Kamal et al., 2019). SG may be necessary if it does not interrupt a multi-cropping system or unfavorable weather does not challenge it. Areas with two or more seasons and areas with unfavorable climatic conditions at the end or during the maturity of crops will have adverse effects.

The morphological characteristics, especially the CO_2 diffusional pathways, have been reported to play a significant role in photosynthetic capacity. In non-steady-state photosynthesis of Arabidopsis, Sakoda et al. (2021) found that g_s imposed the most critical limitations on photosynthesis comparing g_m. The US soybean cultivars showed a higher g_s than the Japanese soybean, which was attributed to the higher stomatal density of the US soybean. It was concluded that higher photosynthetic capacity in the U.S cultivar could be due to the higher g_s potential of these cultivars (Tanaka et al., 2010). Moreover, the manipulation of guard cell transport and metabolism was suggested to balance the CO_2 uptake and water loss (Lawson and Blatt, 2014). Similarly, a study on transgenic tobacco revealed that increased stomatal density of these tobacco lines increased photosynthesis by 30% (Tanaka et al., 2013), making g_s one of the promising targets to be manipulated for enhanced photosynthesis.

On a Calvin cycle level, manipulating some enzymes and substrates can boost photosynthesis. Increasing the amount of sedoheptulose-1,7-bisphosphatase activity in transgenic tobacco stimulated photosynthesis (Lefebvre et al., 2005). Similarly, increasing

the levels of sedoheptulose-1,7-bisphosphate, fructose-1,6-bisphosphate aldolase, and the cyanobacterial putative-inorganic carbon transporter B (*ictB*) improves photosynthetic CO_2 fixation (12-19% in mature plants), leaf area, biomass, and yield in transgenic tobacco (Simkin et al., 2015). Increasing the quantity, specificity, or affinity of CO_2 in Rubisco has also been proven to increase photosynthesis (Parry et al., 2012). Furthermore, the reduction of transient photoinhibitions, increase in RuBP regeneration capacity, and increase in specificity of Rubisco for CO_2 have been identified as ideal targets for the genetic manipulation of light-saturated photosynthesis in C_3 crops (Lawlor et al., 1999).

Introducing foreign materials or the C_4 pathways (carbon concentrating mechanism) to some C_3 crops is another auspicious target as these mechanism enables CO_2 concentrations and limits oxygenation reaction by Rubisco. For instance, transgenic soybean lines developed by introducing the cyanobacterial *ictB* gene showed 9% and 10% higher photosynthetic rates in two independently replicated experiments in the U.S (Hay et al., 2017). Studies on the possibilities of converting C_3 crops to C_4 crops have also concluded that relatively fewer genes control the transition of C_3 to C_4. Although the mechanism of differentiation is flexible between C_3 and C_4 crops, efforts to create a single-cell C_4 system had little success. Therefore, it is suggested that the interaction of the leaf morphology and photosynthesis and the divergent development of mesophyll and bundle sheet need to be elucidated (Zhu et al., 2010).

1.5. Strategies for improving photosynthesis

Both empirical and theoretical studies have suggested many auspicious routes that could ultimately result in enhanced photosynthesis. The use of wild-related species in plant breeding, the exploration of plant germplasm or crop panels for promising photosynthetic traits, understanding physiological mechanisms and genetic architecture underpinning

photosynthesis, and developing high throughput strategies that could harness natural and genetic variations in a time, cost, and labor efficient manner are promising avenues onwards enhancing photosynthesis. Here, I have focused on some realistic strategies that may potentially increase photosynthetic capacity in soybean.

1.5.1. The use of wild-related relatives in soybean breeding as a means of improving photosynthesis

Extreme diversity in morphology, cytology, genomic structure, growth climate, growth habit, and soil conditions has given little chance to utilize genetic diversity in the breeding programs of soybean domestication (Singh et al., 1990). Recently, the use of wild-related species in soybean breeding has attracted attention primarily because of their natural resistance to rust disease and broad geographical distribution. Soybean has 26 wild-related species. Among them, *Glycine tomentella* Hayata is unique for having different cytotypes, wide geographical distribution, and resistance to rust disease (Chung and Singh, 2008; Sing and Nelson, 2015; Schoen et al., 1992).

Biotic and abiotic stress tolerance has already been reported in wild-related species. Little is known about their photosynthetic capacity. A wild rice species (*Oryza australiensis*), for instance, revealed greater photosynthetic capacity compared to temperate-adapted *O. sativa* after being exposed to 45 ^{o}C heat, owing its better photosynthetic performance to vigorous carboxylation capacity and thermotolerance (Phillips et al., 2021). Modern wheat cultivars have been reported to have lower photosynthetic capacity than their wild ancestors (Evans and Dunstone, 1970). Although limited studies have reported advantages in the photosynthetic capacity of wild-related species comparing their domesticated or commercial counterparts, they hint at and lay a foundation for the manipulation of wild-related species in plant breeding

1.5.2. Screening germplasm for the natural variations of photosynthesis

Harnessing natural variations in plant germplasm is vital for photosynthetic enhancement. Historically, the plant efficiency of converting sunlight into biomass has not been improved. Unavailability of instruments to measure photosynthesis, sink-limited rather than assimilate-limited production of major food crops, and the idea that photosynthesis is not a driver for yield improvement are the three major constraints that prevented the optimization of light conversion efficiency. The number of germplasms stored in the gene banks is far greater than those used in plant breeding. The gap between the available and in-use germplasms may hold resource use efficiencies. Therefore, exploring the unexplored germplasm is crucial (Kromdijk and Long 2016). Natural variations of photosynthesis exist among and within species. For instance, nitrogen concentration explains photosynthetic capacity within species. Among species, however, leaf nitrogen and the photosynthetic nitrogen use efficiency (PNUE) define most of the variations (Hikosaka, 2010).

Sakoda et al. (2016) studied natural variations of physiological and morphological factors in diverse soybean genotypes and found significant variations among them. They discovered that a line (line 14) had 11% higher A_{sat} than the reference genotype (*Tachinagaha*) and suggested that the CO_2 fixation was due to the higher content of rubisco. Since rubisco is one of the targets for improving photosynthesis (Long et al., 2006), Sakoda et al. (2016) findings can be an excellent example of how screening plant germplasm for good traits is an essential strategy.

1.5.3. Understanding genetic architecture underpinning photosynthetic variations

Enhancing photosynthesis through genetic manipulation is getting increasing attention recently due to the availability of high throughput field phenotyping and increasingly accessible DNA sequencing technology. Genetic manipulation is essential in the classical optimization of morphology and upscaling processes inside the photosynthetic machinery. Many targets have been identified that could potentially improve photosynthesis. For example, manipulating cytochrome b_6f complex proteins, reducing the size of light-absorbing antenna size, introducing genes from cyanobacteria, the optimization of Rubisco, and many more have been repeatedly reported, which requires the understanding of genetic architecture (Evans, 2013). One of the most common uses of DNA sequences (genotypes) and phenotypes is the genome-wide association study (GWAS), also known as whole-genome regression (WGR). This method analyzes the association of single nucleotide polymorphisms (SNPs) with variations in a quantitative or qualitative trait, finally identifying some essential genes.

Using GWAS, researchers have reported many novel loci and genes that could potentially regulate photosynthesis. A robust genetic correlation (r = 0.80) between seed yield and photosynthesis as well as some novel loci were reported in the U.S. soybean nested association map (NAM) panel (Lopez et al. 2019). Several novel loci related to photosynthesis and growth-related traits were reported in tomatoes that may regulate photosynthesis, primary metabolism, and plant growth (de Oliveira Silva et al., 2018). A gene co-expression network underlying photosynthesis was studied and revealed that one of the highly linked genes (*PtoPsbX1*) significantly improved photosynthesis in *Arabidopsis thaliana* (Xiao et al., 2020). With DNA sequencing becoming cheaper and high throughput field phenotyping technology easily accessible, GWAS studies and

subsequent manipulation of genes would be significant. Currently available studies and gene editing opportunities have enticed many scientists. These studies collectively suggest that understanding the genetic architecture of plant species and their correlation with quantitative traits would hopefully pave the way to remove inefficiencies within the photosynthetic apparatus and increase its capacity.

1.6. Soybean as a testbed

Soybean [*Glycine max* (L.) Merrill] originated from its wild ancestor, *G. soja*, and was first cultivated in China. *G. soja* is an annual, weed-like, climbing soybean ancestor with black seeds in its pods that scatter at maturity and grows wild in China, Taiwan, Japan, and far Eastern Russia (Singh et al., 1990; Singh, 2010). Due to morphological and biological diversity in the wild ancestors of soybean, it is argued that soybean has originated independently in different regions in eastern Asia. Scholars in general, however, agree that soybean has originated from China based on the following reasons. Firstly, *G. soja* has the most intensive presence in China, whereas its presence is sparse in other countries. Secondly, China has the oldest record of soybean cultivation. Thirdly, carbonized soybean was first discovered in China, dating back 3,000 years. And lastly, soybean has been introduced to all major soybean producers directly or indirectly from China. While the origin of soybean is China, scholars have different viewpoints about its domestication (Singh, 2010).

China was the world's largest soybean producer and exporter in the early 20th century. In the 1950s, soybean production developed intensively in the U.S, which is now considered the world's largest soybean producer. In 1970, it was introduced to Brazil, the second major soybean producer. Afterward, this crop was introduced to different countries worldwide (Singh, 2010).

Soybean is typically a short-day, warm-season plant (Garner and Allard, 1920), which is adapted to narrow bands of latitude due to photoperiodic sensitivities (Cober and Morrison, 2010). It is an economically important plant of the Leguminosae family which contains 40 and 20% of protein and oil in its seed mass, respectively (Chung and Singh, 2008; Koester et al., 2014). This crop enriches the soil with nitrogen through biological nitrogen fixation. Soybean has a proliferated top-root system, thus playing a fundamental role in improving soil's physical properties. It is one of the leading crops grown under rain-fed conditions, exploiting the available rainfall pattern (Singh, 2010; Chung and Singh, 2008). Since soybean contributes mainly to the world's economy, its production to meet the food and feed demands of the world's expanding population is vital (Ainsworth et al., 2012). Studying the natural genetic variation of photosynthetic capacity of the wild-related soybean progenies and Japanese soybean germplasm could divulge important information about the genetic potential of the Japanese soybean.

1.7. Objectives

This study is aimed at exploring different avenues to increase leaf photosynthetic capacity. Chapter two aimed to assess if the wild-related backcross-derived progenies have higher photosynthesis than the recurrent parent, Dwight. These wild-related progenies were derived from the backcross of cultivated soybean (cv. Dwight) and *G. tomentella* Hayata (a wild relative of soybean). Chapter three assessed the genotypic variations of photosynthetic capacity in the Japanese soybean germplasm. The objectives were a) to assess the photosynthetic capacity of the Japanese soybean accessions in the Japanese soybean mini core collection (GmJMC) and b) to understand the physiological mechanisms underlying photosynthetic capacity. Chapter four aimed to understand the

genetic architecture of GmJMC underpinning phenotypic variations of A_{sat} to find valuable materials affecting photosynthetic capacity.

1.8. Methodology

Although there is some commonality among chapters, the methodology of the chapters differs based on objectives. In the second chapter, a field experiment with a Randomized Complete Block Design (RCBD) was conducted. Gas exchange parameters, biochemical parameters, morphological parameters, biomass, and yield attributes were measured intensively. In the third chapter, two different experiments were conducted. Gas exchange was measured in both experiments; however, different machines were used. For screening A_{sat} and identifying the genetic architecture, MIC-100 (MASA international corporation, Tokyo), a high throughput system, was used. However, a standard photosynthesis measurement system, LI-6800 (LI-COR, USA), was used to identify physiological mechanisms. Chapter four focuses on Genome-Wide Association Study (GWAS). This chapter shares phenotypic data with Chapter three. Genotypic data and an RNA analysis toolkit were utilized in addition to photosynthetic phenotyping. GWAS analysis was conducted in R. PLINK2 was used for data management, Gaston package was used for fitting the GWAS regression, and finally, the CLC workbench was used to extract haplotypes and create primers.

Chapter Two

Physiological Analysis of Leaf Photosynthesis of Backcross-derived Progenies from Soybean (*Glycine max* (L.) Merrill) and *G. tomentella* Hayata

2.1. Abstract

Enhancing leaf photosynthesis is an enticing aspect of increasing crop seed yield. Using wild-related species in soybean breeding can be a potential source of improving leaf photosynthesis. Two backcross-derived progenies of soybean (*Glycine max*) with *G. tomentella* Hayata were evaluated for photosynthesis, biomass, and seed yield. Gas exchange, Rubisco content, and leaf nitrogen content (N_{cont}) were measured from the flower initiation stage (R_1) up to the seed initiation stage (R_5). Results revealed significant increases in light-saturated photosynthesis (A_{sat}), mesophyll activity (A_{sat} /C_i), seed yield, and total aboveground dry weight (TDW) in progenies relative to Dwight. A significantly higher specific leaf weight (SLW) was observed in progenies, which was strongly correlated with A_{sat} (r=0.86***). There was no significant difference between Dwight and the progenies in stomatal conductance (g_s), but Dwight had higher intercellular CO_2 concentrations (C_i). It indicates that increases in A_{sat} were associated with improved A_{sat} /C_i. These findings suggest the potential use of wild soybean relatives in breeding to enhance soybean leaf photosynthesis.

2.2. Introduction

The enhancement of photosynthetic capacity has been reported in the historic soybean germplasm. It was observed that recently released soybean cultivars had higher daily carbon gain than the older cultivars (Koster et al., 2016). To sustain yield gains, the sunlight conversation efficiency to biomass, which was also reported to be higher in modern soybean cultivars, is proposed as a target to enhance through greater photosynthesis (Evans, 1997; Long et al., 2006; Zhu et al., 2010). One of the avenues to accelerate photosynthetic enhancement in soybean could be the utilization of soybean wild-related species.

Soybean has 26 wild-related species (Singh and Nelson, 2015). Among these species, *G. tomentella* Hayata has four cytotypes, broad geographical distribution, and resistance to leaf rust (Chung and Singh, 2008; Schoen et al., 1992). Singh and Nelson (2015) crossed *G. tomentella* Hayata with *G. max* [cv. Dwight] in a struggle to broaden the soybean gene pool and introduce leaf rust resistance to commercial soybean cultivars. The F_1 generation was backcrossed with Dwight four times until self-fertile progenies were developed. Then, self-fertile progenies were crossed with each other until the F6 generation. The resultant progenies (BC_4F_6) with 1% introgression from *G. tomentella* produced a higher seed yield relative to Dwight in field trials (Begemann, 2015) and a controlled experiment (Akpertey et al., 2018). The genetic material in wild *Glycine* species has been referred to as a "locked treasure." Therefore, I hypothesize that using wild-related soybean species in breeding programs may have resulted in progenies with enhanced photosynthetic capacity and improvements in yield.

The photosynthetic capacity of crop wild relatives has been reported in many studies. For instance, Evans and Dunstone (1970) compared the photosynthetic capacity of modern

wheat cultivars and their wild ancestor and reported lower photosynthetic rates in modern wheat. Similarly, Xue and Gao (2017) studied A_{sat} between wild and cultivated soybean species along with Cadmium (Cd) stress; and reported that without Cd stress, there was no significant difference between them. Their assessment of photosynthetic capacity was the study of two diverse backgrounds; however, studies comparing the photosynthetic capacity of soybean progenies carrying genetic factors from wild-related soybean species have not yet been reported. Therefore, the objectives of this study are 1) to explore the difference and the underlying physiological mechanism in gas exchange parameters between Dwight and progenies and 2) to have an insight into the apparent seed yield and TDW enhancement in progenies.

2.3. Materials and methods

2.3.1. Plant materials and sowing procedure

Two progenies were selected as plant material from several backcrossed-derived progenies (BC_4F_6) of soybean (*G. max*, cv. Dwight) and *G. tomentella* Hayata (Singh and Nelson, 2015). These progenies (LG11-3341 and LG09-12682) were chosen based on higher A_{sat} in a preliminary experiment in Illinois, U.S. **(Figure 2.1)**. LG11-6210 is not a progeny but was cultivated as a checked high-yield genotype, and Dwight is the recurrent soybean parent. Progenies and Dwight are indeterminate in growth. They were sown on Jul 04, 2016, and June 27, 2017, in a randomized complete block design (RCBD) with two replications in an experimental field at the Laboratory of Crop Science, Graduate School of Agriculture, Kyoto University (35º.2" N, 135º. 47" E). The plot size was 1 x 0.75 m and 2.4 × 2.9 m, with a density of 9.52 plants per m^2 in 2016 and 2017, respectively. Seeds were soaked in pesticides before sowing to minimize pest damage. A 3-10-10 gm^{-2} ratio

of N, P_2O_5, and K_2O fertilizers was used during sowing. Three seeds were planted per hill but only one was retained after unifoliate leaves fully expanded. Irrigation and pesticide applications were practiced regularly.

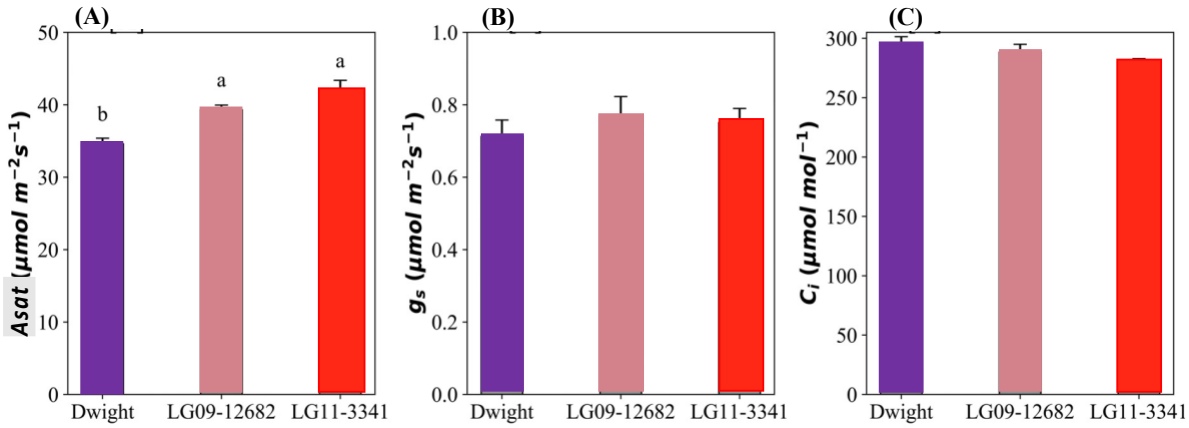


Figure 2.1. Genetic Variation of Gas Exchange traits between Dwight and progenies in preliminary observations in Illinois, U.S.A. Light-saturated photosynthesis **(A)**, stomatal conductance **(B)**, and intercellular CO_2 concentration **(C)**. Bars represent the average value of 4 biological replications, and error bars show the standard error of the population.

2.3.2. Photosynthetic measurements

A_{sat}, g_s, and C_i were measured in the central leaflets of the most recent, fully expanded trifoliolate of six randomly selected plants in each plot. These measurements were conducted from 9:00 AM to 12:30 PM on 37, 44, 51, and 58 days after planting (DAP) in 2017. An LI-6400XT portable photosynthesis system (LI-COR Bioscience; Lincoln, NE, USA) was used for these measurements. The temperature inside the chamber cuvette was set to 33 ^oC. The reference level of carbon dioxide was controlled at 400 µmol mol^{-1} using CO_2 cartridges, with an airflow rate of 500 µmol s^{-1}. Light intensity was set to 1500 µmol m^{-2} s^{-1} with 10% blue light. Humidity was kept within the chamber in a 50-80% range and adjusted to the ambient environment accordingly. Diurnal effects were minimized by a random selection of plots for all genotypes.

2.3.3. Rubisco content and leaf nitrogen content

Leaves were tagged with adhesive tape after the gas exchange measurements. The labeled samples were collected from three of the six plants, where the gas exchange was measured. The sampled leaves were then taken to the laboratory, and four discs of 1.2 cm in diameter were removed using a leaf punching device. The discs were frozen immediately in liquid nitrogen and stored at -80 ^{o}C. We followed the Rubisco procedure after Makino et al. (1986). Leaf tissues were homogenized using chilled mortars and pestles in the presence of an extraction buffer, which consists of 50 mM Hepes-KOH, five mM $MgCl_2$, one mM EDTA, 0.1% (w/v) PVPP, 0.05% (v/v) Triton X-100, 5% glycerol, four mM amino-n-caproic acid, 0.8 mM benzamidine-HCl, and five mM DTT at pH 7.4 with a small amount of quartz sand. The homogenate was centrifuged at 14,500 *g* for 6 minutes. Fifty-eight µl of the supernatant was mixed with 50 µl of 2*sodium dodecyl sulfate (SDS), 4 µl of Bromo phenyl blue (BPB), and 4 µl of mercaptanol in a 1.5 ml sample tube. After the vortex, the sample is heated for five minutes. Ten µl of the mixture is then inserted into the comb holes of the Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and placed in a running buffer. Bovine serum albumin (BSA) is assigned as a standard. The gel is electrophorized for 75 minutes. The gel is then colored overnight and decolored for 8 hours. The bands on the gel are cut and placed in Formamide. After the incubation period of 5 hours at 50 ^{o}C, the final product was read at 595nm using a spectrophotometer. After taking samples for Rubisco content, the leaf area of the remaining leaves was determined using a Leaf Area meter (LI-3100C Area Meter). The leaves were then oven-dried at 80 ^{o}C for 72 hours to assess total leaf dry weight. A Kjeldahl Digestion method was used to measure N_{cont}. A leaf sample of 0.2 g was dissolved in 4 ml of concentrated H_2SO_4. The solution was heated until the acid became colorless. The colorless solution was diluted with water until the total volume reached 40 ml. Twenty

µl aliquots of this solution were separated into glass tubes. The separated solution was mixed with 2480 µl of distilled water, 1000 µl of Indophenol A, and 1500 µl of Indophenol B. They were read at 635nm wavelength using a spectrophotometer. I adopted the N_{cont} measurement method from Vickery (1946).

2.3.4. Stomatal density

A lateral leaflet of the same trifoliate previously used for gas exchange was also sampled and stored in an iced box for measuring stomatal density ($S_{Density}$). The Suzuki Universal Method of Printing (SUMP) was used for printing stomatal maps. A droplet (almost 10µl) of SUMP liquid (amyl acetate) was placed on a SUMP disc. The leaf was then attached to the disc and was placed between two magnets until the liquid became dry. After detaching the leaves, the discs were observed in a 100x magnification light microscope (BH-32; OLYMPUS) with a mounted digital camera (FLOYD; Wraymar, Multi-Interface Digital Camera). Digital pictures were then analyzed using ImageJ (ImageJ 1.50i; Wayne Rasband, National Institute of Health, USA). I followed Tanaka and Shiraiwa (2009) to determine $S_{Density}$.

2.3.5. Yield components

Five and Ten plants per plot were harvested at full maturity (R_8) for 2016 and 2017, respectively. The plants were separated into seeds and stems (including pod shells). Branches and pod shells were oven-dried at 80 ^{o}C for 72 hours, and dry matter was measured. To determine seed yield, seed moisture was calculated and set to 14%, whereas, in TDW, the moisture content of the seed was reduced to 0%.

2.3.6. Statistical analysis

Tow-way analysis of Variance (ANOVA) and Tukey Post Hoc (Tukey-HSD) test of multiple comparisons were performed on seed yield, TDW, and Harvest Index using R's built-in and Agricolae packages (Mendiburu, 2015). A one-way ANOVA test was conducted on gas exchange parameters. The difference was determined using Tukey's Post Hoc test of multiple comparisons (the difference is stated in the text only). Pearson correlation test was performed using R's Psych package version 2.2.3 (Revelle, 2022). These tests were verified using Python's Penguin, Scikit-Learn, Statsmodels, and Scipy modules.

2.4. Results:

2.4.1. Gas exchange measurements

A_{sat} was significantly higher ($p < 0.05$) in progenies relative to Dwight. Its values ranged from 27.55μmol $m^{-2}s^{-1}$ in Dwight on 44 DAP to 35.19μmol $m^{-2}s^{-1}$ in LG11-3341 on 58 DAP. Lower A_{sat} was observed for all genotypes on 44 DAP, followed by a gradual increase up to 58 DAP. LG11-6210 had higher A_{sat} in the early reproductive stages; the difference, however, was insignificant in later growth stages **(Figure 2.2A)**.

g_s ranged from 1.18 mol m^{-2} s^{-1} in LG11-6210 on 32 DAP to 1.57 mol m^{-2} s^{-1} in LG09-12682 on 44 DAP. No significant and consistent differences in g_s were observed between Dwight and the progenies **(Figure 2.2B)**. Unlike A_{sat}, higher C_i values were observed for Dwight. C_i was significantly higher ($p < 0.05$) in Dwight than the progenies, ranging from 294.26μmol mol^{-1} for LG11-3341 on 58 DAP up to 307.17μmol mol^{-1} in Dwight on 44 DAP **(Figure 2.2C)**.

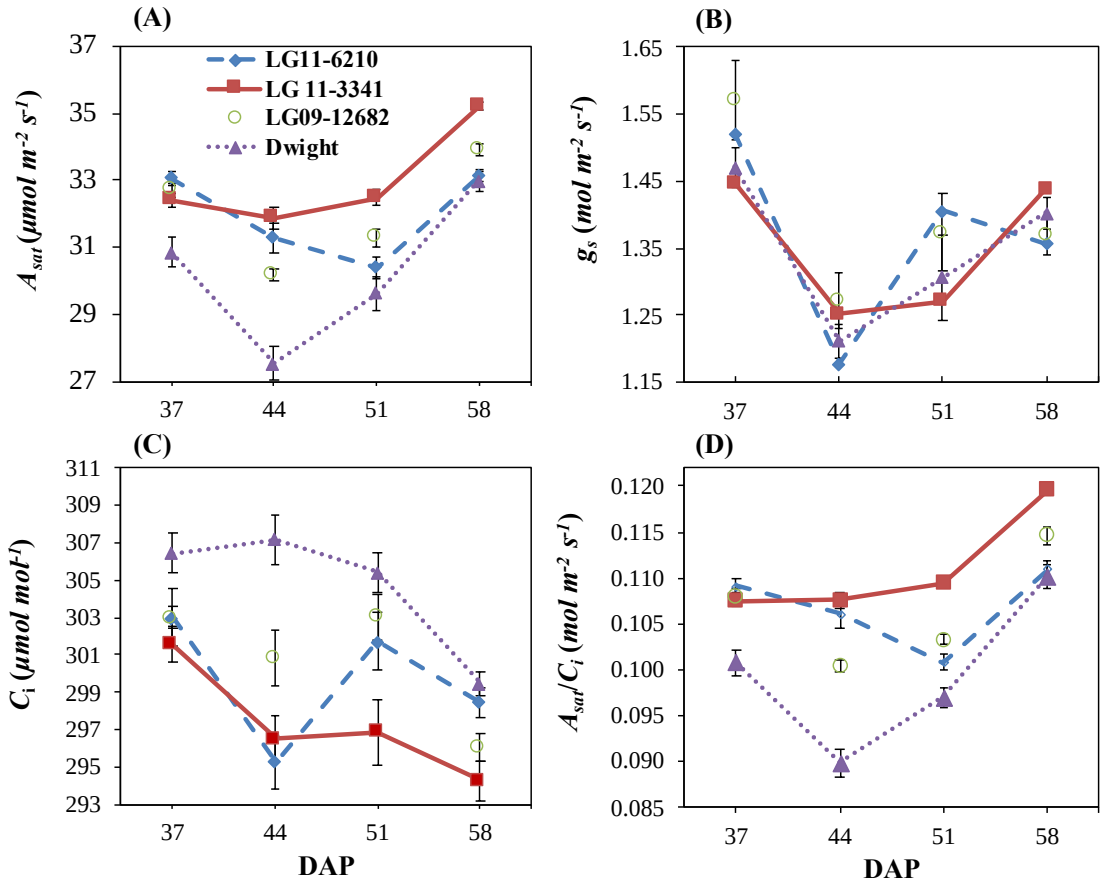


Figure 2.2: Genetic variation of gas exchange parameters between Dwight and progenies: light-saturated photosynthesis **(A)**, stomatal conductance **(B)**, intercellular CO_2 concentration **(C)**, and mesophyll activity **(D)**. Each data point represents the average value of 12 biological replications. Bars show the standard error of the population (n=12).

2.4.2. Biochemical parameters

The range of Rubisco content started from 1.21 gm^{-2} in LG11-3341 on 32 DAP to 2.27 gm^{-2} in LG11-6210 on 58 DAP. LG11-6210 and LG09-12682 showed a linear increase in Rubisco content. Comparing Dwight, the progenies had higher Rubisco content during late flowering (44 DAP), but it decreased significantly in LG11-3341 on 58 DAP. Interestingly, Dwight followed a completely inverse pattern of LG11-3341. Its Rubisco content decreased significantly on 44 DAP, followed by a sharp increase on 58 DAP, surpassing both progenies. The checked genotype, LG11-6210, showed substantially higher Rubisco content than the progenies and Dwight **(Figure 2.3A)**.

N_{cont} ranged from 1.17 gm^{-2} in Dwight on 44 DAP to 2.15 gm^{-2} in LG09-12682 on 58 DAP. It varied among measurements and genotypes. In the early measurements (37-44 DAP), progenies had higher $N_{cont,}$ but the difference between progenies and Dwight was significant only on 44 DAP. However, in the later measurements (51-58 DAP), the difference was not significant between Dwight and progenies **(Figure 2.3B).**

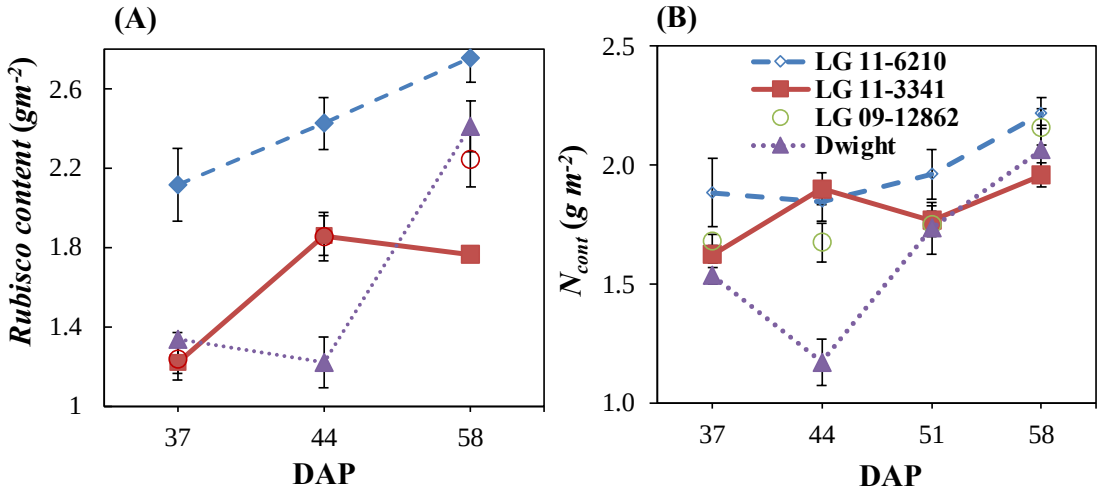


Figure 2.3: Variations of Rubisco and N_{cont} between Dwight and the progenies. Rubisco content **(A)**, and leaf nitrogen content **(B)**. Each data point represents the average of six biological replications. Bars show the standard error of the population (n=6).

2.4.3. Stomatal density

$S_{Density}$ ranged from 249 mm^{-2} in LG11-3341 to 350 mm^{-2} in LG11-6210. Dwight had a linear increase in $S_{Density}$ over time, whereas fluctuations in $S_{Density}$ were observed in both progenies. There was no significant difference in $S_{Density}$ between the progenies and Dwight **(Figure 2.4A)**. Although not the subject of this research, the intrinsic water use efficiency was also higher in progenies compared to Dwight. However, the difference was more pronounced in LG11-3341 in the later three measurements **(Figure 2.4B).**

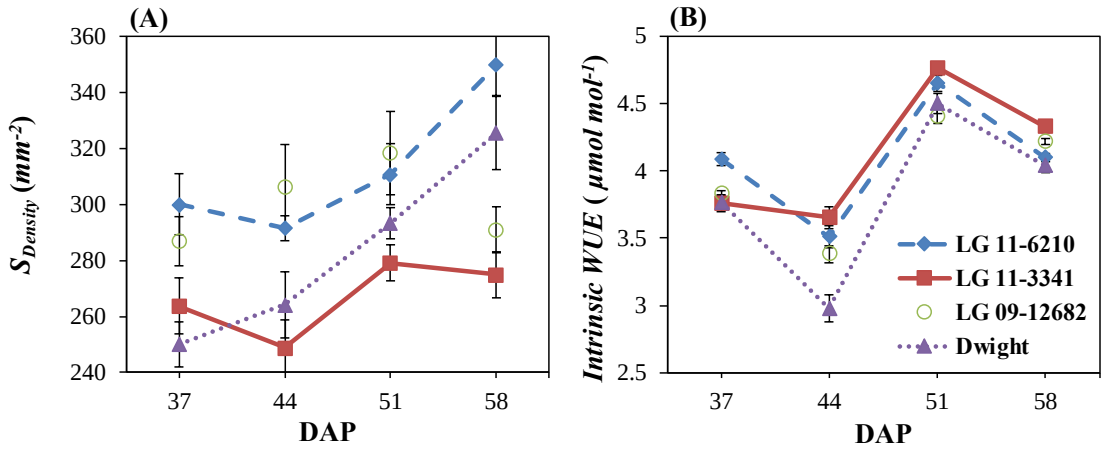


Figure 2.4: Variation of stomatal density and intrinsic WUE between Dwight and progenies. Stomatal density **(A)**, and intrinsic water use efficiency **(B)**. Each data point represents the average value of 6 and 12 biological replications for $S_{Density}$ and intrinsic WUE, respectively. Bars indicate the standard error of population (n=6) and (n=12) for $S_{Density}$ and Intrinsic WUE, respectively. Intrinsic WUE is calculated as A_{sat} divided by the Transpiration of water vapors measured simultaneously during gas exchange.

As **Table 2.1** shows, A_{sat} had significant positive correlations with SLW and N_{cont}, whereas the correlation with C_i was negative. Similarly, C_i had strong negative correlations with SLW, Rubisco, and N_{cont}. g_s had a moderate correlation ($p = 0.49^{\cdot}$) with A_{sat}; however, statistical analysis showed no effect ($p > 0.1$) of $S_{Density}$ on g_s, Ci, and A_{sat}.

Table 2.1. Correlation coefficients between gas exchange and leaf traits in a pooled dataset.

	A_{sat}	g_s	C_i	SLW	Rubisco	A_{sat}/C_i	N_{cont}
g_s	0.49†						
C_i	-0.68**	0.24					
SLW	0.86***	0.44	-0.58*				
Rubisco	0.40	-0.27	-0.62*	0.64*			
A_{sat}/C_i	0.99***	0.37	-0.78***	0.85***	0.46		
N_{cont}	0.72**	0.16	-0.68**	0.86***	0.83***	0.74***	
$S_{Density}$	0.16	0.06	-0.11	0.46	0.76**	0.15	0.58*

A_{sat} represents light-saturated photosynthesis, g_s stands for stomatal conductance, C_i is the intercellular CO_2 concentration, SLW represents specific leaf weight, A_{sat}/C_i shows mesophyll activity, N_{cont} represents leaf nitrogen content, and $S_{Density}$ is the density of stomata. Values are averaged for each measurement (n=16).
*** $p = 0$, ** $p < 0.001$, * $p < 0.05$,
† $p < 0.10$

2.4.4. Seed yield and TDW

Seed yield ranged from 20.60 g plan^{-1} in Dwight to 34.84 g plan^{-1} in LG09-12682 in 2016. In 2017, on the other hand, seed yield ranged from 42.68 g plan^{-1} in Dwight to 58.28 g plan^{-1} in LG11-3341. Progenies produced significantly higher seed yields than Dwight in 2016, whereas the difference was only significant between LG11-3341 and Dwight in 2017 **(Figure 2.5A**). In 2016, TDW was as low as 32.76 g plant^{-1} in Dwight and as high as 54.80 g plant^{-1} in LG11-3341. Both progenies produced significantly higher TDW in 2017. A steep increase in TDW was noticed in 2017 for Dwight and progenies; however, only LG11-3341 had significantly higher TDW than Dwight **(Figure 2.5B)**.

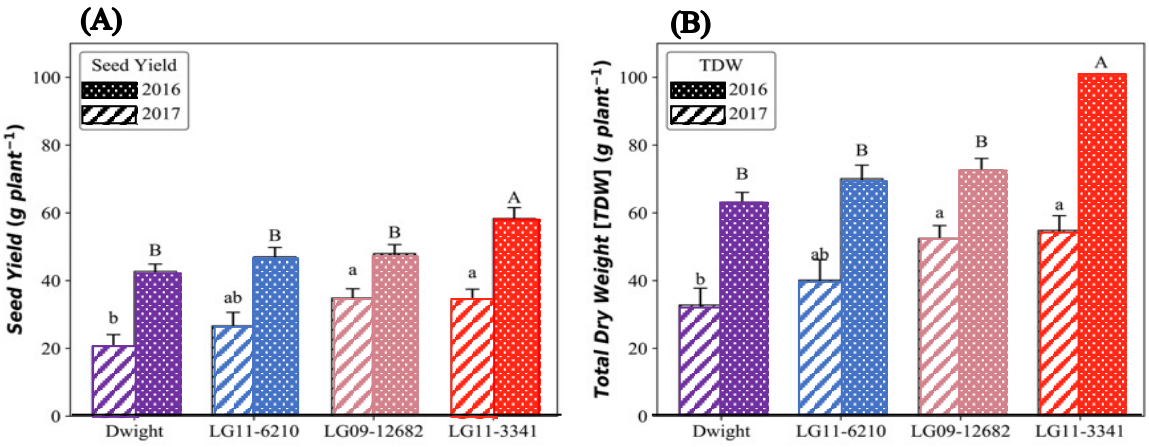


Figure 2.5: Variations of seed yield **(A)**, and total dry weight **(B)** between Dwight and progenies. Bars show seed yield and total dry weight (TDW) for 2016 (n=10) and 2017 (n=20), respectively. Error bars indicate the standard error of the population. Lowercase and uppercase letters show significant differences at $p < 0.05$ for 2016 and 2017, respectively.

Table 2.2 shows the Two-way ANOVA test conducted on seed yield, TDW, and Harvest HI. There was a significant difference between years and among genotypes regarding seed yield and TDW; however, the difference in HI was significant among genotypes and when the interaction of year and genotypes was considered.

Table 2.2: Analysis of Variance on seed yield, TDW, and Harvest Index (HI).

Factor			Seed Yield g plant $^{-1}$	TDW g plant $^{-1}$	HI
Mean	Year(Y)	2016	29.01	45.02	0.59
		2017	48.99	75.42	0.59
	Genotype(G)	Dwight	35.32	52.96	0.60
		LG11-6210	40.23	60.07	0.61
		LG09-12682	43.58	65.84	0.60
		LG11-3341	50.19	82.27	0.56
	ANOVA	Y	0.000^{***}	0.000^{***}	$0.105^{n.s}$
		G	0.000^{***}	0.000^{***}	0.000^{***}
		Y x G	$0.320^{n.s}$	$0.148^{n.s}$	0.025^{*}

Here, TDW represents the total dry weight, and HI stands for harvest index.
*** p =0, * $p < 0.05$, n.s: not significant.

2.5. Discussion

Several studies have noted discrepancies in leaf photosynthetic capacity for crop species varying in genomic or cytoplasmic formation (Broderson et al., 2008; Brown and Bouton, 1993; Carver et al., 1989; Hay et al., 2017; Izhar and Wallace, 1967; Sakoda et al., 2016). For example, modern wheat cultivars have shown lower photosynthetic rates than their wild ancestors (Evans and Dunstone, 1970). No improvements in the photosynthetic rates in rice breeding have been noted in the last 50 years in Venezuela (Pieters et al., 2011). On the other hand, Koester et al. (2016) reported photosynthetic enhancement of recently released soybean cultivars. Similarly, Morrison et al. (2000) reported a 0.52% annual increase in photosynthetic rates of newly released cultivars.

2.5.1. A_{sat} and mechanisms underlying it

This study documents the potential for utilizing wild soybean species in breeding to enhance soybean leaf photosynthesis. The results highlight that soybean progenies had higher A_{sat} than Dwight **(Figure 2.2A)**. Variations of A_{sat} in these progenies were associated with an increase in mesophyll activity not only due to lower C_i per given g_s but also an apparent proxy of mesophyll activity (A_{sat}/C_i), which includes both CO_2 diffusion from intercellular spaces into the chloroplast and its fixation, followed the same pattern as A_{sat}. Since A_{sat}/C_i represents CO_2 diffusion from intercellular air spaces into the chloroplast and its fixation, it is reasonable to assume that progenies have either a higher mesophyll conductance or increased capacity for CO_2 fixation in the Calvin cycle. These findings are consistent with the results of Fischer et al. (1998). They also elucidated the physiological mechanism for increased photosynthetic capacity, stating that 2/3 increases in A_{max} were associated with increases in A_{max}/C_i, and 1/3 increase was

related to increases in C_i. The effect of C_i on A_{max} in their study was linked to higher stomatal conductance in genotypes with higher photosynthetic rates. In this study, however, g_s did not differ noticeably between progenies and Dwight. Therefore, I concluded that the variations in A_{sat} were solely linked to improvements in A_{sat} /C_i.

2.5.2. Biochemical and morphological attributes

There was no significant correlation between A_{sat} and Rubisco content but A_{sat}/C_i and Rubisco content were correlated when all the data were pooled. This was due to the lower Rubisco content of LG11-3341 and LG09-12682 and its association with a higher A_{sat} **(Figure 2.3A)**. It suggests that factors other than Rubisco content participated in enhanced CO_2 fixation in the progenies. One possibility could be the increased activity of Rubisco Activase. Since Rubisco activase is controlled by the ratio of Adenosine triphosphate (ATP) to Adenosine diphosphate (ADP) [ATP: ADP] (Streusand and Portis, 1987), progenies might have developed a more efficient PSII and electron transfer system resulting in better proton harvesting. Another possibility could be the higher activation status of Rubisco content. Simply put, the more significant the portion of Rubisco activated, the more the CO_2 is fixed.

The effect of N_{cont} on leaf photosynthesis has been well documented in different studies (Allison and Williams, 1997; Boussadia et al., 2010; Sakoda et al., 2016). In this study, N_{cont} was also significantly correlated with A_{sat} (r = 0.72**) when the data were pooled for all four measurements. Within each measurement, there was a significant correlation between A_{sat} and N_{cont} on 37-44 DAP, but no correlation was observed in the latter two measurements. Since N_{cont} was not significantly different between progenies and Dwight in the last two measurements (51-58 DAP), suggesting that variations in A_{sat} were not due to variations in N_{cont} **(Figure 2.3B)**.

It is reported that plant biomass and seed yield are tightly linked to transpiration (Sinclair et al., 1984). Condon et al. (2004) emphasized the importance of improved water-use efficiency (WUE) for sustainable agricultural water supply. An outstanding characteristic of the progenies was the higher intrinsic water-use efficiency compared to Dwight (Ares et al., 2000). As **(Figure 2.4 B)** demonstrates, the progenies maintained a higher A_{sat} and intrinsic WUE simultaneously. These results suggest that the progenies could be used in breeding programs to combine higher photosynthetic capacity with increased WUE.

2.5.3. Seed yield and total dry weight

As a result, the progenies were observed to have higher seed yield and TDW relative to Dwight **(Figure 2.5)**. Compared to 2017, a drastically lower seed yield and TDW were observed in 2016. I believe that late sowing did not allow plants to invest enough in vegetative biomass before reaching the reproductive stage, producing poor yield. Koutroubas et al. (1998) reported that the dry matter accumulation and nitrogen at early growth stages are essential in producing higher seed yields. Similarly, Specht et al. (1999) said that increasing seed yield potential through increasing total dry matter plays a role. In this study, yield enhancements of the progenies and the difference between these two years were also related to the enhanced dry matter accumulation. The lower HI of LG11-3341 was not due to its inability to convert biomass to seed yield; instead, this progeny is very tall, and lodging occurs despite controlling measures. These results of seed yield were, as a whole, consistent with the findings of Akpertey et al. (2018) and (Begemann, 2015); however, genotypes, in general, were lower performing in a study conducted by Akpertey et al. (2018). A higher seed yield for all genotypes was observed, but the difference was evident for progenies, especially LG11-3341. In Akpertey et al. (2018), this progeny

produced a lower seed yield than Dwight. This could be related to the lodging that occurred in Akpertey et al. (2018) experiment; we, on the other hand, managed to reduce lodging.

2.6. Conclusion

This is the first study that reports enhanced leaf photosynthetic capacity in backcross-derived progenies from soybean and its wild relative, *G. tomentella* Hayata. This study showed that higher A_{sat}/C_i triggered enhanced leaf photosynthesis in progenies. LG11-3341 and LG09-12682 were characterized by consistently higher A_{sat} than Dwight throughout the growing season. This study demonstrates that using *Glycine tomentella* Hayata in soybean breeding can significantly increase leaf photosynthetic capacity and other vital parameters such as WUE in the resulting progenies.

Chapter Three

Genotypic Variation in Leaf Photosynthetic Capacity of Japanese Soybean and Physiological Mechanisms Underpinning It

3.1. Abstract

Improving biomass relies on radiation use efficiency (RUE), which in turn depends on manipulating photosynthetic processes in plants. Efficient strategies to assess photosynthetic capacity in crops need to be developed to identify suitable targets that have the potential to improve photosynthetic efficiencies. Here, the photosynthetic capacity of the Japanese soybean mini core collection (GmJMC) was evaluated using a newly developed high-throughput photosynthesis measurement system "MIC-100" and analyzed physiological mechanisms underpinning photosynthesis. K-means clustering of light-saturated photosynthesis (A_{sat}) classified GmJMC accessions into four distinct clusters, with Cluster2 comprised of highly photosynthesizing accessions. Among GmJMC accessions, GmJMC47 was characterized by the highest A_{sat}, stomatal conductance (g_s), stomatal density ($S_{Density}$), electron transfer rate (ETR), and light use efficiency of photosystem II (Fv'/Fm') and the lowest non-photochemical quenching ($NPQ_{(t)}$), indicating that GmJMC47 has greater CO_2 supply and efficient light-harvesting systems. These results prove that exploring plant germplasm is a valuable strategy to unlock the potential of resource use efficiencies for photosynthesis.

3.2. Introduction

Although conventional breeding still adds to increases in crop seed yield (Khush, 2001; Stapper and Fischer, 1990), exhaustion of some key yield-improving parameters has been reported (Zhu et al., 2010). Developing novel strategies to sustain sufficient foodstuff is essential (Maurino and Weber, 2013; Ray et al., 2013). Manipulating photosynthetic pathways to produce enough food, practice sustainability, and avoid exploiting natural resources is, therefore, a promising strategy (Janssen et al., 2014).

Photosynthesis, however, contains numerous inefficiencies. Biochemical and diffusional factors majorly limit photosynthesis. Several targets have been proposed to modify photosynthetic capacity genetically. For instance, a large quantity of Rubisco is required for better photosynthetic capacity. Increases in the specificity or affinity of Rubisco to CO_2 can boost photosynthesis (Parry et al., 2012). Lowering the severity of photoinhibition and increasing the regeneration rate of RuBP make ideal targets for the genetic manipulation of light-saturated photosynthesis in C_3 crops (Lawlor et al., 1999).

Similarly, a study on transgenic *Arabidopsis thaliana* revealed that increased stomatal density (maximum 372%) increased the photosynthetic rate by 30% (Tanaka et al., 2013). The genetic manipulation of g_s in *slac1* mutant (lacking stomatal anion channel protein that controls stomatal closure) revealed that a higher g_s enhances the rate of photosynthesis and the ratio of internal to external CO_2 concentrations (Kusumi et al., 2012). Mounting evidence suggests that the genetic manipulation of rate-limiting processes enhances photosynthesis. The problem, however, is that a large part of plant germplasms remains unexplored to look for improved genetic factors. Therefore, harnessing natural variations in photosynthesis is an ideal route toward removing bottlenecks in photosynthesis.

Soybean is an indispensable part of the Japanese diet; it is also said to increase longevity in Japanese people (Gabriel et al., 2018). Japanese soybean cultivars, however,

are less productive than the US cultivars as the latter cultivars have higher radiation use efficiency (RUE), greater photosynthetic capacity, and dry matter production (Kawasaki et al., 2016). GmJMC (JMC, hereinafter) is small but diverse germplasm that represents a major part of trait variations in the Japanese soybean (Kaga et al., 2012). It comprises ninety-six accessions selected from 1603 soybean accessions based on SNP variations. However, JMC has never been evaluated for genotypic variations in leaf photosynthetic capacity. Considering this collection's geographical and agro-morphological characteristics, it was assumed that studying leaf photosynthetic capacity would divulge valuable information and provide suitable targets for resource use efficiencies. This study aimed to assess the photosynthetic capacity of JMC accessions using a newly developed high-throughput photosynthesis measurement system, MIC-100, by analyzing physiological variations and elucidating the mechanism underlying photosynthetic discrepancies as an avenue to improve photosynthetic efficiency.

3.3. Materials and methods

3.3.1. Plant materials

All germplasms were obtained from the National Agricultural and Food Research Organization (NARO; Gene bank, Japan). Two experiments were conducted for this study. The first experiment was aimed at assessing phenotypic variations in A_{sat}. Ninety accessions were selected from the JMC and were sown on June 18, 2019, among which 16 revealed lodging and poor seed quality. The remaining 74 accessions were cultivated again on June 20, 2020, using a complete randomized design (CRD). Ten plants were sown in each plot.

The second experiment was conducted in 2021 to measure various parameters related to photosynthetic capacity in ten representative accessions (indicated by white dots inside markers in **Figure 3.2**) and investigated the underlying mechanisms. Seven candidate accessions (JMC13, JMC16, JMC37, JMC43, JMC47, JMC49, JMC56) were selected from the best photosynthesizing cluster, Cluster2; two accessions (JMC56 and JMC112) from Cluster1; one accession (JMC25) from Cluster3; and a Chinese accession Peking (named 'PE' as a check for having higher A_{sat} and g_s). JMC25 and JMC112 are two photosynthetically elite cultivars named Enrei ('En') and Fukuyutaka ('Fu'), respectively. These accessions were cultivated in a randomized complete block design (RCBD) with two replicates.

All experiments were conducted at the Laboratory of Crop Science, Graduate School of Agriculture, Kyoto University (35°.2" N, 135°. 47" E). The sowing density was 9.52 plants m^{-2}. The space between rows and plants in a row was 0.7 m and 0.15 m, respectively. A 3:10:10 gm^{-2} ratio of $N:P_2O_5:K_2O$ fertilizers were applied during sowing. Three seeds

were planted per hill, of which only one was retained after unifoliate leaves had fully expanded. Irrigation and pesticide application were made regularly.

3.3.2. Photosynthetic phenotyping

I used a high throughput portable close-chamber infrared gas analyzer, MIC-100 (MASA International Cooperation, Japan), to analyze A_{sat} in 2019 and 2020. This system requires less than 30 seconds per sample, whereas a standard LI-6800 achieves stability in at least 45 seconds for the "Fast Survey Measurements". MIC100 explains around 95% variations in A_{sat} measured by a standard LI-6800 system (LI-COR, USA), making it a high throughput efficient system (Tanaka et al., 2021; LI-COR, 2021). Moreover, my observations in the field revealed that MIC-100 could measure A_{sat} up to seven-fold faster than LI-6800. Measurements were conducted for two consecutive years: three in 2019 (05 Aug, 13 Aug, and 18 Aug) and five in 2020 (06 Aug, 14 Aug, 21 Aug, 27 Aug, and 31 Aug). The data was collected from 9:00 AM to 1:30 PM from a central leaflet of a fully expanded, recently matured trifoliolate from flower initiation stage (R_1) to seed initiation stage (R_5). I measured A_{sat} in 360 and 240 leaves on each measurement day for 2019 and 2020, respectively, collectively summing up to 2280 samples. Only one plant was sampled at a time so that A_{sat} is distributed throughout the survey to compensate for the diurnal privilege of accessions. Light intensity inside the cuvette was 1200 μmol $m^{-2}s^{-1}$ which is the maximum capacity of this system. The temperature was automatically controlled and did not exceed 36 ^{o}C. Ambient CO_2 and humidity conditions were used for the experiment, with the initial CO_2 concentration set at (370–390 ppm) and the final concentration set twenty points lower than the initial concentration.

In 2021, an LI-6800 (LI-COR, USA) was used for gas exchange and chlorophyll fluorescence measurements on 22 Jul, 30 Jul, and 11 Aug. These parameters were

measured in four randomly selected plants in each plot from 9:00 AM to 11:00 AM. The temperature inside the chamber cuvette was 33 ºC. Reference level CO_2 was set at 400 µmol mol^{-1} using CO_2 cartridges, with an airflow rate of 500 µmol s^{-1}. Light intensity was set to 1200 µmol $m^{-2}s^{-1}$ with 10% blue light, and relative humidity was set at 60%. Non-photochemical quenching was calculated as $NPQ_{(t)}$ from chlorophyll fluorescence parameters following Tietz et al. (2017).

3.3.3. K-Means clustering analysis

Before performing the K-means clustering analysis, the data was standardized so that each measurement had a mean of zero and a standard deviation of one. T-distributed Stochastic Neighbor Embedding (t-SNE) was used to reduce the dimensionality of the data (A_{sat} was measured eight times throughout 2019 and 2020). The number of components in t-SNE was set to two (SNE1 and SNE2), and one thousand iterations were given to the algorithm to reduce the sum of square differences among clusters. The output data was used in the K-means clustering algorithm with four centroids.

3.3.4. Biochemical parameters

Nitrogen content (N_{cont}), Chlorophyll a (*Chl a*), Chlorophyll b (*Chl b*), and Chlorophyll a+b (*Chl a+b*) were quantified at the same leaves where the gas exchange was measured. For measuring *Chl a*, *Chl b*, and *Chl a+b*, four-leaf discs of 0.5 cm diameter were taken and dissolved in 2ml N, N-Dimethylformamide (DMF) as described by Porra (2002). The samples were then wrapped in aluminum foil and stored at 5 ºC for over 24 hours. Spectroscopic readings of the supernatant were taken at 646.8 nm, 663.8 nm, and 750 nm wavelengths using a U-2910 Spectrophotometer (HITACHI, Japan). The area of the remaining leaves (after removing discs) was measured using a leaf area meter LI-3100C (LI-COR, USA). They were then oven-dried for 72 hours and were used for

quantifying N_{cont} using Kjeldahl digestion. Vickery (1946) was followed for quantifying N_{cont}. The oven-dried leaves were crushed, and a sample of 0.2 g was dissolved and heated in 4 ml of highly concentrated sulfuric acid until the solution turned colorless. The colorless solution was then diluted to 40 ml with distilled water. Finally, 20 µl aliquots of this solution were separated into glass tubes and mixed with 2.48 ml of distilled water, 1 ml of Indophenol A, and 1.5 ml of Indophenol B. The samples were then read at 635 nm wavelength using a U-1100 Spectrophotometer (HITACHI, Japan).

3.3.5. Morphological parameters

Stomatal density ($S_{Density}$), stomatal length (S_{Length}), and stomatal width (S_{Width}) were measured using the same leaves previously used for the gas exchange measurements. I followed Tanaka and Shiraiwa (2009) to determine stomatal parameters. After taking discs for chlorophyll content analysis, the Suzuki Universal Method of Printing (SUMP) was used for printing stomatal maps. A droplet (around ten µl) of SUMP liquid (amyl acetate) was placed on a SUMP disc. The leaf was placed on the disc and left to air-dry for ~20 mins. The discs were later observed at 100X (for $S_{Density}$) and 400X (for S_{Length} and S_{Width}) magnifications under light microscope BH-32 (OLYMPUS, Japan) with a Multi-Interface Digital Camera FLOYD (Wraymar, Japan). ImageJ was used for counting stomata and measuring their length and width.

For specific leaf weight (SLW), the fresh leaf area and the dry weight of the leaves during Ncont measurement were recorded. SLW was derived as the ratio of a leaf's dry weight to its fresh area (dry weight (g)/ leaf area (m^{-2})).

3.3.6. Statistical tests and graphs

Two-way ANOVA and Tukey's Honest Significant Difference (Tukey HSD) analysis were conducted for all parameters using the built-in R functions of 'aov' and 'TukeyHSD'; and the 'HSD.test' function of Agricolae package (de Mendiburu, 2015). Accessions (G) and the measurement dates (D) were used as factors. Due to field homogeneity, the block difference was considered negligible. The 'pearsonr' function of the Scipy library (Virtanen et al., 2020) in Python was used to find correlations among various parameters. Python's matplotlib.pyplot (Hunter, 2007) library was used for creating figures and Sciket Learn library (Pedregosa et al., 2011) for K-means clustering.

3.4. Results

3.4.1. Screening, distribution, and clustering of A_{sat}

As shown in **Figure 3.1**, there were significant variations in A_{sat} among accessions and measurement dates. A_{sat} ranged from 11.17–32.25 μmol $m^{-2}s^{-1}$. In 2019 and 2020, the frequency of accessions with lower A_{sat} was higher at the beginning and the end of August. In contrast, the A_{sat} measurements conducted in mid-August had a narrow range, with most accessions showing greater A_{sat}. Shapiro's test of normality revealed that the data were normally distributed. In general, the A_{sat} of Fu (red triangle in **Figure 3.1**) was poor at the beginning of August, but its A_{sat} gradually increased towards the middle and end of August.

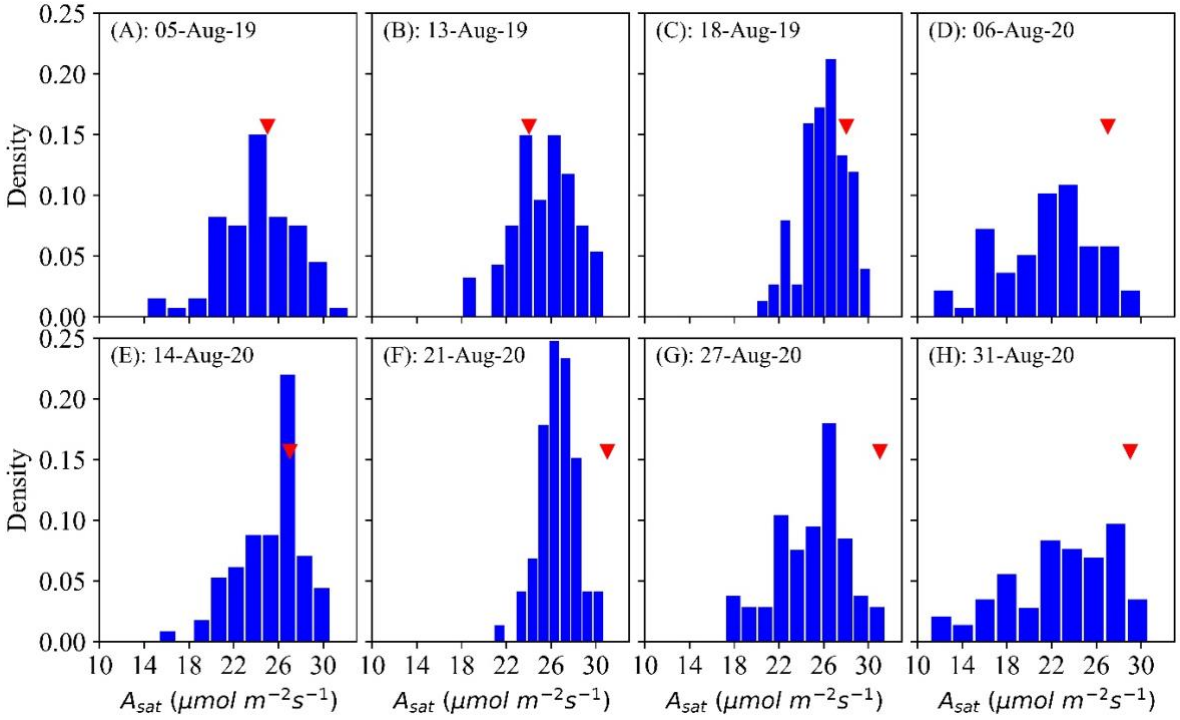


Figure 3.1: Variation of A_{sat} in the GmJMC (JMC) accessions and its comparison with Fukuyutaka (Fu) across measurements. Red triangles show the A_{sat} performance of Fu.

t-SNE compressed variations of A_{sat} into two components. SNE1 explained only 22.9% variations, whereas SNE2 explained 77.0% variations in A_{sat} measurements. SNE1 had a negative correlation with all the A_{sat} measurements and was particularly strong with measurements taken on 05 Aug 2019 and 06 Aug 2020. In contrast, SNE2 was positively correlated with all the A_{sat} measurements (**Taabe 3.1**).

Taabe 3.1: Correlation coefficient between t-SNE components and A_{sat} in each measurement date. SNE1 and SNE2 show first and second t-SNE components, respectively.

Date	SNE1	SNE2
05 Aug, 2019	-0.65	0.55
13 Aug 2019	-0.13	0.52
18 Aug, 2019	0.13	0.57
06 Aug, 2020	-0.75	0.37
14 Aug, 2020	-0.34	0.57
21 Aug, 2020	-0.57	0.65
27 Aug, 2020	-0.13	0.76
31 Aug, 2020	-0.30	0.85

t-SNE classified accessions into four clusters with great differences between photosynthetic capacities (**Figure 3.2**). The computed Silhouette coefficient was 0.76, indicating that the clusters are spherical with only minor overlapping. Significant differences in A_{sat} among the clusters were observed (p <0.001, **Figure 3.3**). On average, Cluster2 had the highest A_{sat} and consisted of some of the best photosynthesizing (A_{sat}) accessions. Several of these accessions showed photosynthetic rates comparable to or higher than the elite cultivars, En in Cluster3 and Fu in Cluster1. Cluster1 and Cluster3 had cross-year variations in their A_{sat} values, whereas Cluster4 was characterized by the lowest A_{sat} values in 2019 and 2020.

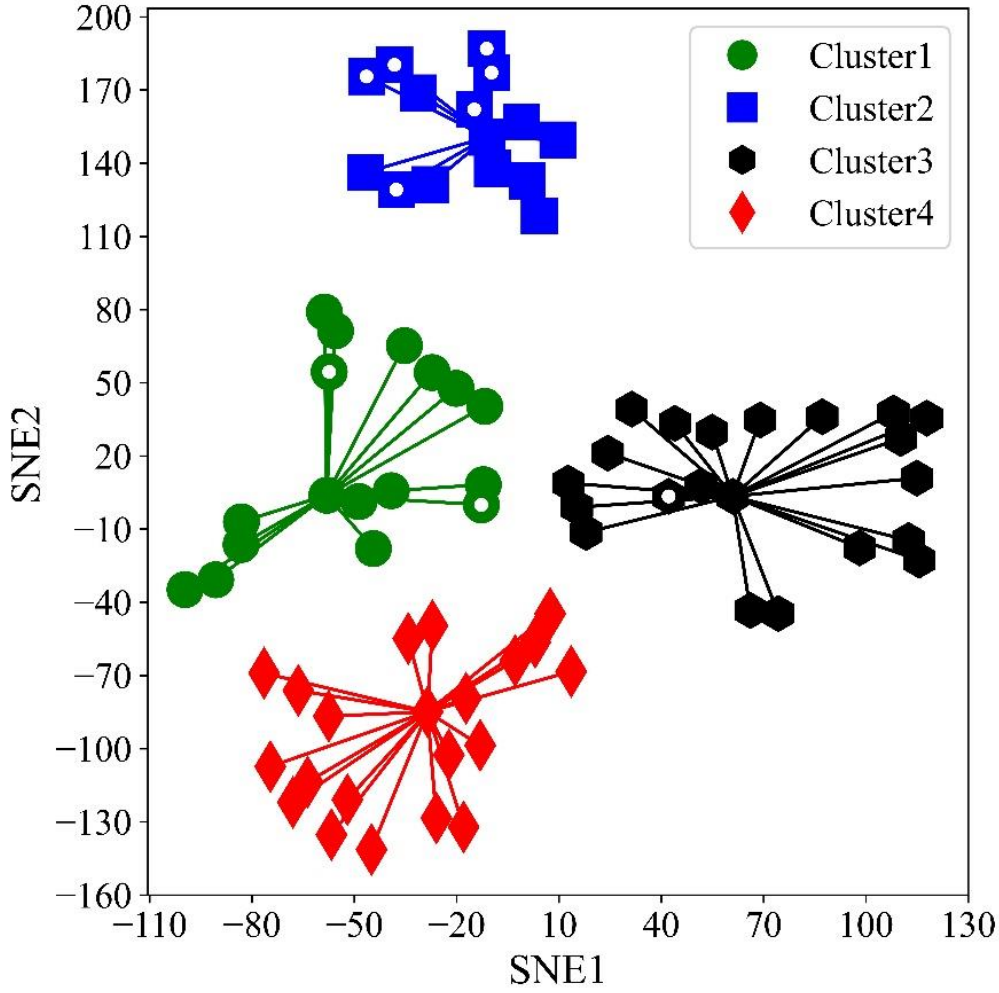


Figure 3.2: Two-dimensional scatter plot of four K-means clusters. White dots inside markers indicate representative accessions used for the second experiment.

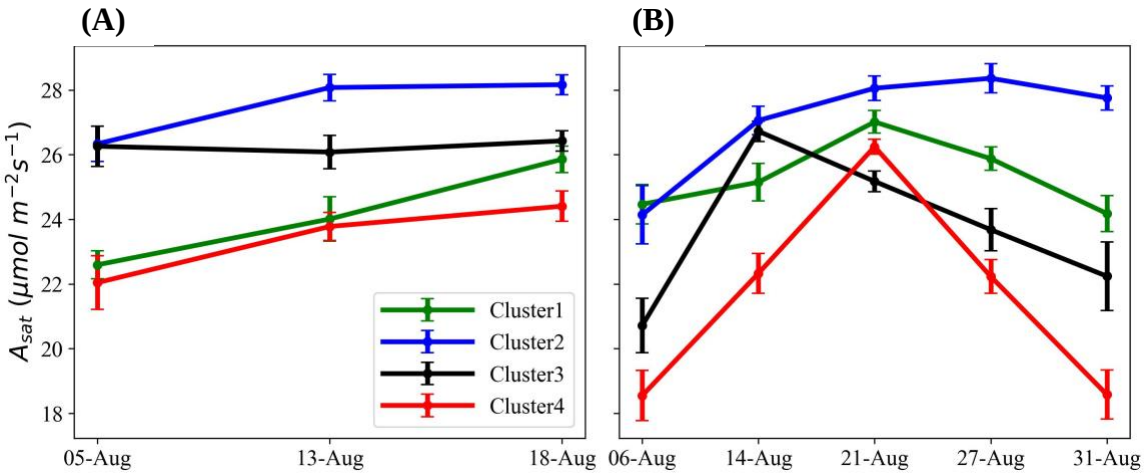


Figure 3.3: Difference of fluctuation in A_{sat} for each cluster across measurements: **(A)** 2019, **(B)** and 2020. Datapoint represents the mean of all members in a cluster. Error bars represent standard error.

3.4.2. Gas exchange and chlorophyll fluorescence

Considerable variations of A_{sat} among ten representative accessions and measurement dates were observed during the second experiment in 2021 (**Figure 3.4A**). A_{sat} ranged from 24.1 $m^{-2}s^{-1}$ in JMC56 on 22 Jul to 41.5 µmol $m^{-2}s^{-1}$ in JMC47 on 30 Jul. The lowest A_{sat} was observed on 22 Jul, which then increased sharply on 30 Jul, followed by a drop on 11 Aug. Among accessions, PE had the highest A_{sat} on 22 Jul, followed by JMC47. PE and JMC47 were significantly higher than En, JMC43, and JMC56 ($p < 0.05$). On 30 Jul, JMC47 showed the highest A_{sat} and was significantly greater than En, JMC13, JMC16, JMC49, and JMC56 ($p < 0.01$). It was also the best performing accession on 11 Aug, but the difference was insignificant among accessions. The g_s measurements showed a similar pattern to A_{sat} but with greater magnitudes (**Figure 3.4B**). It ranged from 0.46 in JMC43 on 22 Jul to 1.69 mol $m^{-2}s^{-1}$ in JMC47 on 30 Jul. The g_s values were exceptionally low on 22 Jul, followed by a sharp increase on 30 Jul and a drop on 11 Aug. PE and JMC47 had the highest g_s values in all measurements. Seasonal change in C_i was also comparable to A_{sat} and g_s. The lowest rates of C_i were recorded on 22 Jul, followed by an increase on 30 Jul and a drop on 11 Aug (**Figure 3.4C**). Among accessions, JMC43 had the lowest C_i of

273.3 µmol mol^{-1} on 22 Jul, whereas JMC56 had a C_i of 313.6 µmol mol^{-1} on 22 Jul. When the apparent mesophyll activity (A/C_i) was calculated, its variations were observed to be strongly correlated with variations in A_{sat}. JMC47 had the highest A/C_i values in all measurements, whereas En and JMC56 had the lowest A/C_i on 22 Jul.

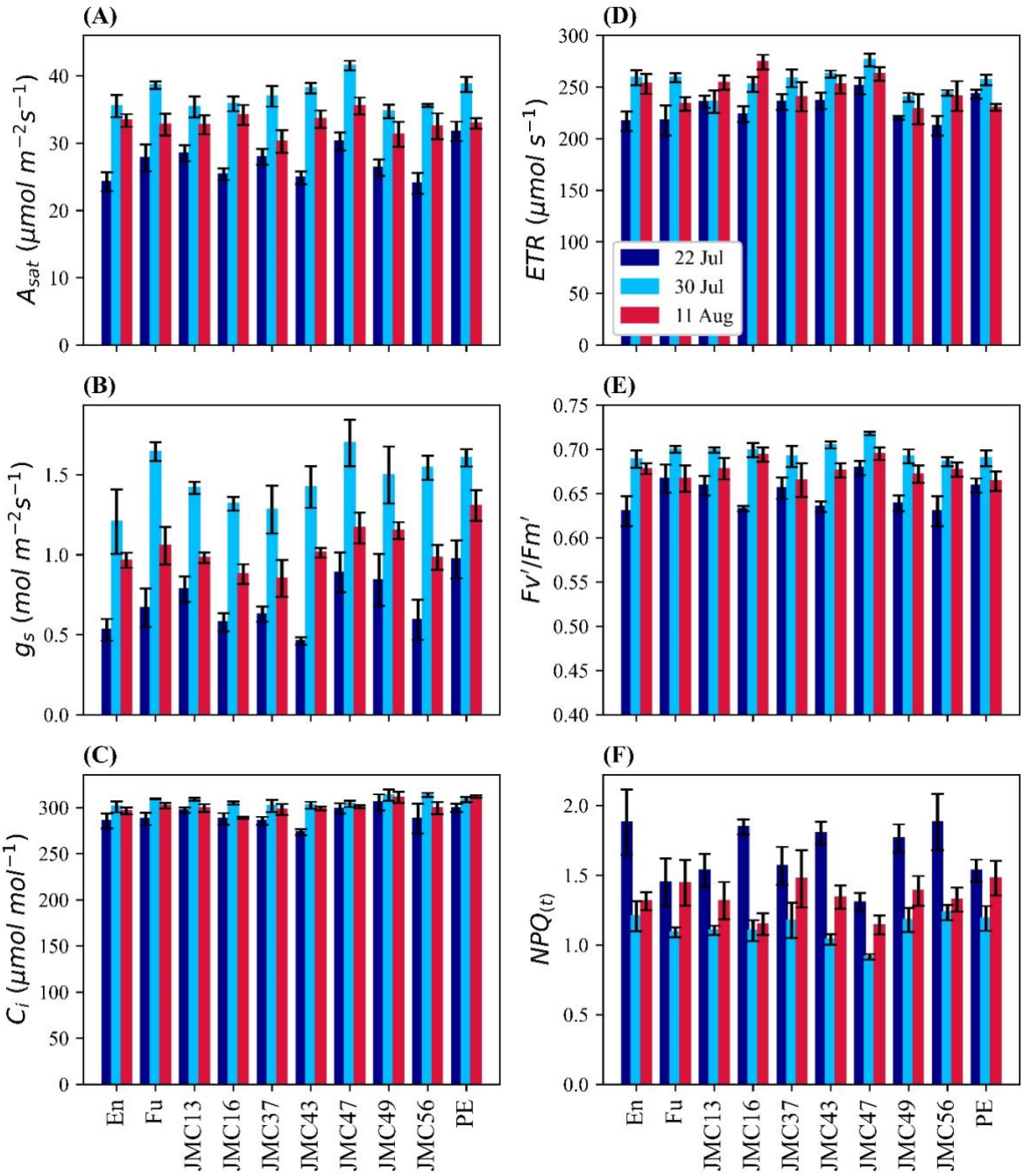


Figure 3.4: Phenotypic variations in gas exchange parameters and chlorophyll fluorescence parameters; Light-saturated photosynthetic rate (**A**), stomatal conductance **(B)**, intercellular CO_2 concentration **(C)**, electron transport rate (**D**), the light utilization efficiency of *PSII* **(E)**, and Non-photochemical quenching (**F**). Datapoint is the mean of four observations. Error bars indicate standard error.

Chlorophyll fluorescence parameters exhibited a similar pattern to gas exchange parameters. On average, ETR was significantly lower on 22 Jul than on 30 Jul and 11 ($p < 0.05$), but the values were not significantly different between 30 Jul and 11 Aug (**Figure 3.4D**). ETR ranged from 212.1 μmol $m^{-2}s^{-1}$ in JMC56 to 276.2 μmol $m^{-2}s^{-1}$ in JMC47. *FV'/Fm'* rates were also the lowest on 22 Jul,y followed by an increase on 30 Jul and a drop on 11 Aug (**Figure 3.4E**). JMC47 exhibited the highest *Fv'/Fm'* among all the measurements. In contrast, $NPQ_{(t)}$ exhibited a reverse gas exchange pattern and fluorescence parameters (**Figure 3.4F**). Its mean values were 1.65, 1.12, and 1.33 on 22 Jul, 30 Jul, and 11 Aug, respectively, and the difference was significant ($p < 0.01$). It ranged from 0.91 in JMC47 on 30 Jul to 1.8 in En on 22 Jul. JMC47 had significantly low $NPQ_{(t)}$ values on 22 Jul and 30 Jul among all the accessions ($p < 0.01$).

Table 3.2: A two-way ANOVA table showing tests among gas exchange and chlorophyll fluorescence parameters.

		Gas Exchange Parameters				Chlorophyl Fluorescence		
		A_{sat}	g_s	C_i	A/C_i	ETR	*Fv'/Fm'*	$NPQ_{(t)}$
Gen (G)	En	30.84^{c}	0.90^{c}	293.98^{bc}	0.105^{bc}	243.02^{bc}	0.666^{a}	1.48^{a}
	Fu	33.10^{abc}	1.13^{abc}	299.50^{abc}	0.110^{abc}	236.82^{bc}	0.678^{ab}	1.32^{ab}
	JMC13	32.16^{bc}	1.07^{abc}	301.99^{abc}	0.106^{bc}	241.93^{bc}	0.679^{ab}	1.32^{ab}
	JMC16	31.59^{bc}	0.93^{bc}	294.20^{bc}	0.107^{bc}	248.22^{abc}	0.676^{b}	1.34^{ab}
	JMC37	31.85^{bc}	0.93^{bc}	294.72^{bc}	0.108^{bc}	244.91^{bc}	0.671^{b}	1.40^{a}
	JMC43	32.20^{bc}	0.97^{bc}	291.62^{c}	0.110^{abc}	250.60^{ab}	0.672^{b}	1.40^{a}
	JMC47	36.13^{a}	1.29^{a}	301.78^{abc}	0.120^{a}	265.78^{a}	0.700^{a}	1.12^{b}
	JMC49	30.74^{c}	1.17^{ab}	309.85^{a}	0.099^{c}	229.57^{c}	0.667^{b}	1.45^{a}
	JMC56	30.56^{c}	1.05^{abc}	300.50^{abc}	0.102^{c}	232.69^{bc}	0.664^{b}	1.49^{a}
	PE	34.61^{ab}	1.29^{a}	306.17^{ab}	0.113^{ab}	243.38^{bc}	0.671^{b}	1.39^{a}
Date (D)	22-Jul	27.12^{c}	0.70^{c}	290.84^{c}	0.093^{c}	229.32^{a}	0.649^{c}	1.65^{a}
	30-Jul	37.11^{a}	1.46^{a}	306.75^{a}	0.121^{a}	254.52^{b}	0.697^{a}	1.12^{c}
	11-Aug	33.06^{b}	1.05^{b}	300.82^{b}	0.110^{b}	247.41^{c}	0.677^{b}	1.33^{b}
ANOVA	G	***	***	***	***	***	***	***
	D	***	***	***	***	***	***	***
	G x D	*	ns	ns	†	*	ns	ns

Here, A_{sat} stands for light-saturated photosynthesis, g_s represents stomatal conductance, C_i represents Intercellular CO_2 concentrations, A_{sat}/C_i shows the apparent mesophyll activity, ETR stands for electron transport rate, *Fv'/Fm'* represents light utilization efficiency of PSII, and $NPQ_{(t)}$ is non-photochemical quenching. Superscripted letters show a significant difference at p =0.05 or below based on Tukey's Honest Significant Difference (HSD) test. ***p = 0, **p < 0.001, *p < 0.01, † p < 0.05, ns: not significant.

A_{sat} was strongly correlated with g_s in all the measurements; however, the correlation was weaker on 11 Aug (**Figure 3.5**). The positive correlation between A_{sat} and C_i was significant on 22 Jul (p <0.01), whereas it became negatively correlated in the last two measurements. Interestingly, the positive correlation between C_i and g_s was constantly observed across all the measurements. ETR, *Fv'/Fm'*, and *PhiPSII* had strong correlations with A_{sat}.On the contrary, $NPQ_{(t)}$ was negatively correlated with A_{sat} and other Chlorophyll fluorescence parameters.

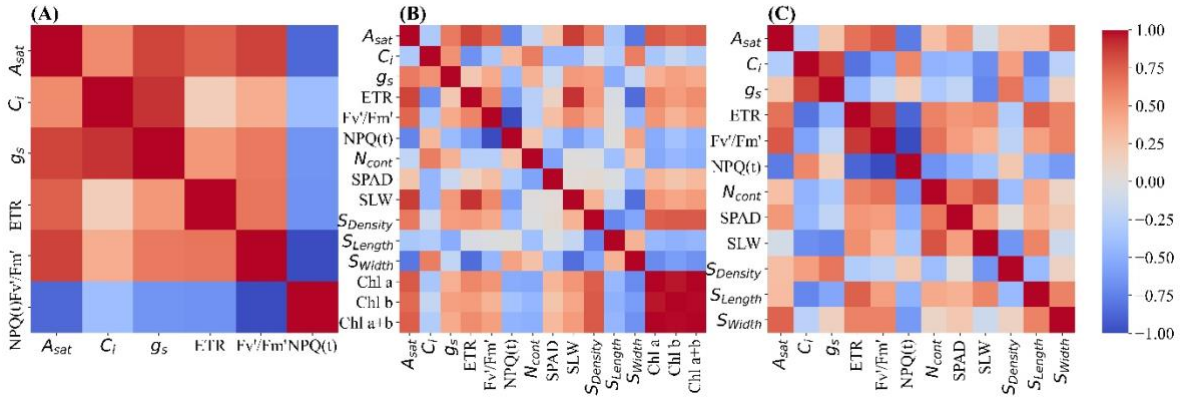


Figure 3.5: Correlation between measured parameters. Panel (**A), (B),** and **(C)** show correlation for 22 Jul, 30 Jul, and 11 Aug, respectively. The strength of the correlation can be guessed from the color bar on the, right where the positive and negative correlations are shown in red and blue colors, respectively.

3.4.3. Chlorophyll and nitrogen content

Variations in *Chl a*, *Chl b*, and *Chl a+b* contents were observed among accessions. En, JMC49, and JMC56 are distinguished with lower *Chl a, Chl b, and Chl a+b* contents. *Chl a* was significantly ($p < 0.05$) higher in PE than in En and JMC56, and *Chl b* was significantly higher in PE than in En, JMC13, and JMC56 ($p < 0.05$, **Figure 3.6A**). JMC47 had the highest *Chl a, Chl b, and Chl a+b* contents among the JMC accessions and was comparable with PE. I observed consistency of N_{cont} variation among accessions between 30 Jul and 11 Aug except for JMC16 and JMC56 (**Figure 3.6B**). The lowest N_{cont} of 1.25 gm^{-2} was seen in JMC37 and JMC43, whereas JMC56 had a higher N_{cont} value of 1.92 g m^{-2} on 30 Jul than most accessions. I did not observe any significant difference in N_{cont} values on 11 Aug. On average, N_{cont} was higher on 30 Jul than on 11 Aug.

The correlation between *Chl a, Chl b, and Chl a+b* and A_{sat} was significant ($p < 0.01$, **Figure 3.5**). However, I did not observe a significant correlation between chlorophyll contents and chlorophyll fluorescence parameters.

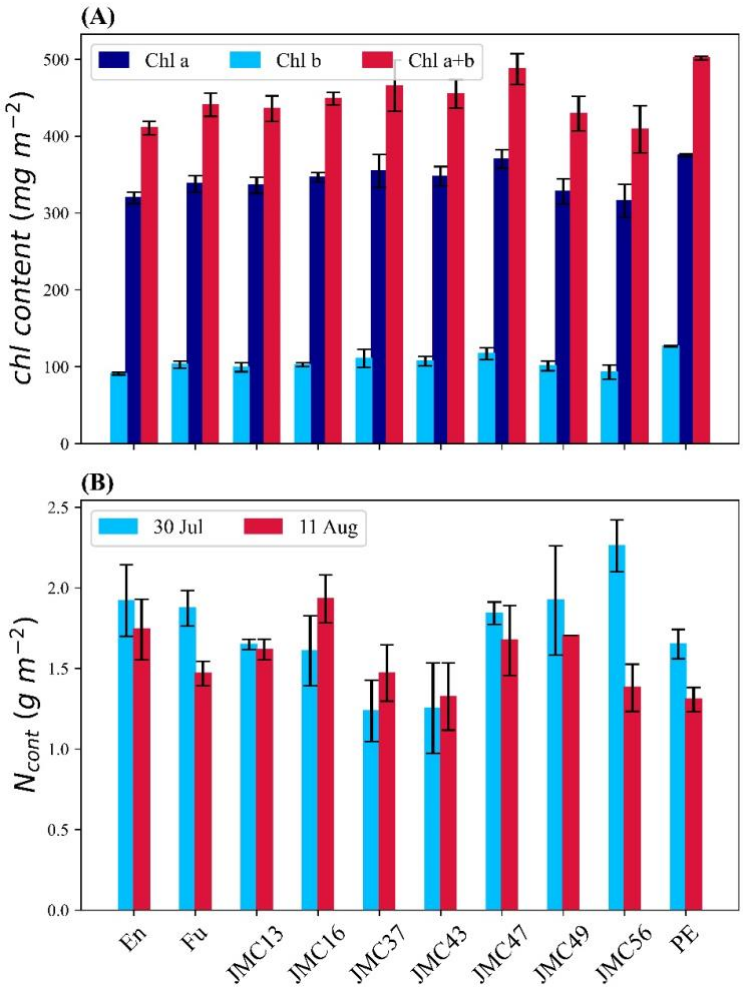


Figure 3.6: Phenotypic variations in biochemical parameters: *Chl a*, *Chl b*, and *Chl a+b* **(A)**, and nitrogen content **(B)**. Each data point is the mean of four observations. Error bars indicate standard errors.

3.4.4. Stomatal attributes and specific leaf weight

JMC47 had the highest $S_{Density}$ among the JMC accessions on 30 Jul (**Figure 3.7A**). PE and JMC47 had significantly more stomata than Fu, JMC13, JMC37, and JMC43 ($p < 0.05$). The $S_{Density}$ difference among accessions was not clear on 11 Aug. Variations of S_{Length} and S_{Width} (**Figure 3.7B**) were also observed among accessions and between measurements. Still, the difference was only significant between the measurement dates ($p < 0.05$) and not among the accessions. S_{Length} was twofold longer than S_{Width} for all the accessions.

En and JMC47 showed distinguishably higher SLW on 30 Jul, but there was no consistency between the values of 30 Jul and 11 Aug (**Figure 3.7C**). JMC13 and JMC49 had smaller SLW values than other accessions on 30 Jul; however, on 11 Aug, PE and JMC56 had the lowest SLW values. Accession (G) to measurement date (D) interaction was observed at $p < 0.05$, but the average SLW values were neither consistently significant among accessions nor between the two measurement dates.

A positive correlation between gas exchange parameter g_s and $S_{Density}$ was observed (**Figure 3.5**), but it was only significant on 11 Aug ($p < 0.05$). A strong negative correlation existed between $S_{Density}$ and $S_{Length,}$ whereas the correlation between $S_{Density}$ and S_{Width} was insignificant. S_{Length} was also negatively correlated with A_{sat}, C_i, and g_s. SLW was not correlated with A_{sat} in pooled data, nor was it in individual measurements.

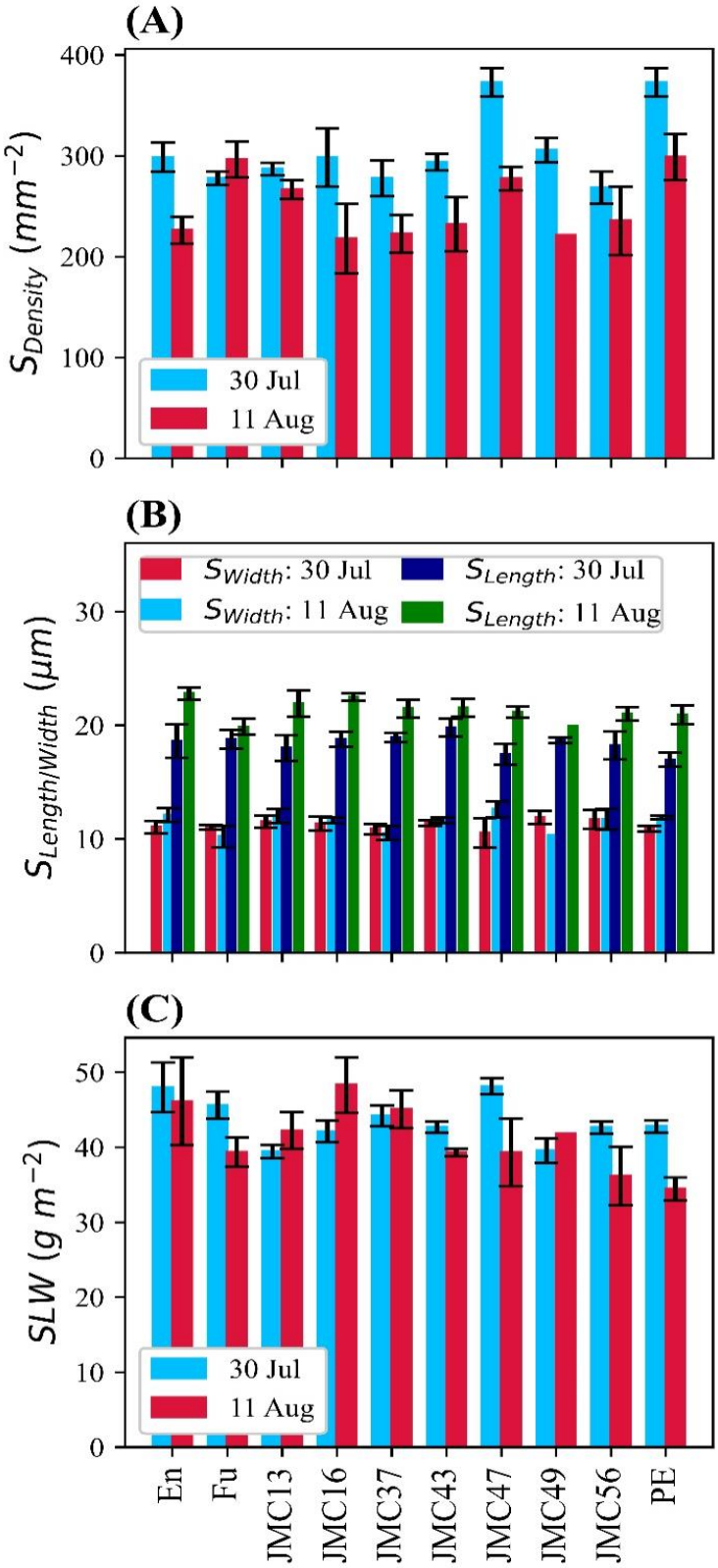


Figure 3.7: Phenotypic variations in morphological parameters: **S**tomatal density **(A),** the length and width of stomata **(B),** and specific leaf weight **(C)**. Datapoint is the mean of four observations. Error bars indicate standard errors

3.5. Discussion

The culmination of conventional yield-improving parameters has widened the margin between food demand and crop yield (Zhu et al., 2008; Zhu et al., 2010), leaving the potential yield productivity to be bridged by the manipulation of photosynthetic processes in plants (Kromdijk et al., 2016; Long et al., 2006). Although exploiting natural variations of photosynthetic apparatus is considered promising for increasing crop productivity (Heckmann et al., 2017), unexplored germplasm is believed to be a major bottleneck (Kromdijk and Long, 2016) towards achieving it. I studied leaf photosynthetic capacity in JMC 1) to harness natural genotypic variations of photosynthetic capacity and 2) to identify physiological traits underpinning photosynthetic variations. To my knowledge, this study is the first to assess JMC for photosynthetic studies on this scale.

3.5.1. Distribution of A_{sat}

Seeking genotypic variations of photosynthetic capacity among JMC accessions over a time-lapse is essential for elucidating the true potential of accessions as they are different maturity groups and show various morphological characteristics. Moreover, photosynthesis is a highly dynamic process that can be affected by multiple factors. Using a high-throughput photosynthetic system, 'MIC-100', cross-time variations in A_{sat} were assessed with greater frequencies. The aggregation and interpretation of extensive data through large-scale phenotyping will be essential to solving the bottleneck of harnessing limited genetic variations in photosynthetic research. Therefore, T-distribution Stochastic Neighboring Embedding (t-SNE) was used for dimensionality reduction. The use of t-SNE has important implications. T-SNE is one of the best dimensionality reduction algorithms for linear and non-linear data. It is also a robust method for visualizing high-dimensional data in two- or three-dimensional panels.

The enhanced photosynthetic capacity of the newly studied JMC accessions relative to Fu (red triangles in **Figure 3.1**, which is one of the elite photosynthesizing Japanese cultivars) has important implications. The low photosynthetic capacity of Fu during early growth stages could be due to its early vegetative stages. Fu is a late maturity cultivar that reaches R_1 when other accessions are around R_3. Although Fu is a highly optimized soybean cultivar, my observations showed that all accessions, including Fu, tend to reduce its photosynthetic capacity at the beginning of September as the weather cools down, implying Fu is not synchronizing its physiological peak with the metrological peak (well suited metrological conditions). It shows that we need to search for the photosynthetic potential of soybean when reproductive stages synchronize with suitable metrological conditions.

3.5.2. K-means clustering

t-SNE analysis revealed some suitable photosynthetic characteristics of the JMC. K-means clustering classified the studied accessions into four distinct clusters with varying A_{sat} values (**Figure 3.2** and **Figure 3.3**). The correlation between actual data and t-SNE components indicates that SNE2 represents the genetic potential of photosynthetic capacity, whereas SNE1 seems to explain seasonal variations in A_{sat} because of their high correlation with measurements at the beginning of Aug in both years (**Table 3.1**). The difference in photosynthetic performance among these clusters could be attributed to some geographical or agro-morphological basis, in addition to the unexplored limitations within photosynthetic machinery. Among four clusters, Cluster2 is particularly interesting as its members showed their high photosynthetic rates equivalent to or higher than the photosynthetically elite cultivars, Fu in Cluster1 and En in Cluster3 (**Figure 3.2**). Higher A_{sat} values among the JMC accessions serve as a source of evidence for the claim that

harnessing natural variations of photosynthesis is an ideal route toward removing bottlenecks in this process and improving seed yield (Heckmann et al., 2017).

3.5.3. A_{sat} and physiological mechanism underpinning photosynthetic capacity

In general, the photosynthetic capacity of the JMC accessions was highly dependent on CO_2 diffusion, mainly due to stomatal limitations (**Figure 3.4**). Low g_s on 22 Jul could be attributed to the early vegetative stages of the accessions. The effects of g_s on photosynthetic capacity are well-established (Cornish et al., 1991; Fay et al., 1995; Taylor and Long, 2017; Yamori et al., 2020). The correlation between g_s and $S_{Density}$ (**Figure 3.5B-C**) and the higher g_s and $S_{Density}$ in JMC47 and PE indicates that higher g_s in JMC47 and PE were due to their leaves allocating more surface to stomata. These results are consistent with the findings of Franks et al. (2009). Tanaka et al. (2010) also reported a strong correlation between potential stomatal conductance and $S_{Density.}$ They argued that it could be underpinning the physiological potential for greater productivity in the U.S. soybean cultivars. In this study, JMC accessions with higher A_{sat} had higher $g_{s,}$ and higher g_s values were usually accompanied by a higher $S_{Density}$, indicating that Japanese soybean may show greater photosynthetic capacity if their stomatal g_s is manipulated (**Figure 3.7**).

The best performance of JMC47 throughout vegetative and reproductive growth stages appears to be due to enhanced light utilization in addition to higher g_s. This accession displayed the highest ETR, *Fv'/Fm'*, *PhiPSII*, and the lowest $NPQ_{(t)}$ values (**Figure 3.4**), indicating that it utilizes sunlight much more efficiently than the elite cultivar of Fu and En. The continuous superiority of ETR and *Fv'/Fm'* in JMC47 across measurement dates is consistent with findings that the application of 24-epibrassinolide increased ETR and the net-photosynthetic rates under water-deficient conditions and suggested that plants

with higher ETR are more efficient in photosynthesis (Pereira et al., 2019). Sheng et al. (2008) studied the effect of *Arbuscular mycorrhiza* (AM) on photosynthesis and water status in maize. They found that enhanced photosynthetic capacity was due to improved *Fv'/Fm'*, *Fv*/*Fm*, and *PhiPSII* and lower non-photochemical quenching (NPQ). Although no treatment was used, the natural superiority of JMC47 in showing greater *Fv'/Fm'* could be utilized as a potential resource in enhancing the light use efficiency of the photosynthetic process.

The relationship between chlorophyll content and photosynthesis has been studied intensively (Buttery and Buzzell, 1977; Holly et al., 2017; Li et al., 2006), and a strong positive correlation between leaf chlorophyll content and photosynthesis has been reported. A positive correlation (**Figure 3.5B**) between chlorophyll content (*Chl a*, *Chl b*, *Chl a+b*) and A_{sat} falls in agreement with previous findings. Higher A_{sat} in JMC47 could also be related to its high chlorophyll content; however, the higher chlorophyll content in PE but lower A_{sat} than that of JMC47 may be due to other unidentified factors. N_{cont} and SLW are correlated with A_{sat} (Allison et al., 1997; Boussadai et al., 2010; Sakoda et al., 2016). In this study, however, I did not observe any correlation. These results indicate that JMC47 had both stomatal and light utilization advantages over other accessions, leading to its highest photosynthetic capacity.

3.6. Conclusion

This study reports detailed information about the photosynthetic capacity of Japanese soybean accessions and the physiological mechanism underpinning it. Here, I classified Japanese soybean germplasm into four clusters, with Cluster2 comprising highly photosynthesizing accessions. Experiment on the ten representative accessions revealed that there were stomatal and non-stomatal limitations on A_{sat} with major limitations due to CO_2 diffusion (lower $S_{Density}$). Among all accessions, JMC47 had the highest $A_{sat,}$ which could be attributed to its high g_s, $S_{Density}$, chlorophyll content, and chlorophyll fluorescence parameters. These results provide solid evidence to suggest that exploiting the genetic variations of plant germplasms is an ideal route to unlocking the resource use efficiency in photosynthesis.

Chapter Four

Analysis of Genetic Architecture for Photosynthetic Capacity of Japanese Soybean Germplasm

4.1. Abstract

Although consensus on the future role of photosynthesis in increasing crop productivity exists, strategies to enhance it are lacking. The availability of high-throughput photosynthetic phenotyping and high-density sequencing has made the exploration of plant germplasms for the genetic architecture of phenotypic variations much easier. Here, the genetic architecture of light-saturated photosynthesis (A_{sat}) was studied in 74 accessions of the Japanese soybean germplasm. Genome-wide association study (GWAS) analysis based on the variation of A_{sat} revealed a significant association with a single nucleotide polymorphism (SNP) on chromosome 17. Among the candidate genes related to photosynthesis in the genomic region, variation in the expression of a gene encoding G protein alpha subunit 1 (GPA1) showed a strong correlation (r = 0.72, p <0.01) with that of A_{sat}. The putative gene may impose photosynthetic regulations through stomatal and chloroplast pathways. This study provides an essential insight into identifying genetic materials for enhancing photosynthesis, suggesting new crop improvement routes.

4.2. Introduction

With the availability of high-throughput photosynthetic phenotyping systems and high-resolution sequencing data, exploring plant germplasms for identifying promising accessions and elucidating the association between genetic architecture and photosynthesis is essential. Studies have assessed the possibility of enhancing photosynthesis through genomics and gene modification. Various strategies have been developed to examine the genetic variations of photosynthetic traits (Lopez et al., 2019; Vieira et al., 2006; Wang et al., 2020). Although novel loci have been identified in these studies, most of them used a single phenotypic event. This approach is subjected to separate statistical penalties, thus, reducing the statistical power of genotype to phenotype association. Reducing the dimensions of highly correlated composite data is proposed as a strategy to overcome this problem. For instance, principal component analysis (PCA) has been increasingly used in GWAS analysis. In PCA, correlated variables are transformed into dimensionally reduced variables called components, which extract useful information and factors from the data that underly trait variations (Reid et al., 2016, Yano et al., 2019).

I conducted a GWAS analysis to identify genomic regions underpinning A_{sat} in the Japanese soybean germplasm using the t-SNE dimensionality reduction method. Complex trait variations of this germplasm make it essential to understand genetic inferences of photosynthetic variations. In this study, I aimed to elucidate the genetic architecture of photosynthesis in the Japanese soybean germplasm to identify useful material for improving photosynthesis.

4.3. Materials and methods

4.3.1. Planting materials

The experiment was conducted in the experimental field at the Laboratory of Crop Science, Graduate School of Agriculture, Kyoto University (35°.2" N, 135°. 47" E). Ninety accessions of JMC were cultivated in a complete randomized design on June 18, 2019. Among accession, GmJMC02 failed to grow. Due to lodging and late maturity, the number of JMC accessions was reduced to 74. These 74 accessions were again sown on 20 Jun 2020. The distance between rows and plants within a row was 70 cm and 15 cm, respectively. The sowing density was 9.52 plants m^{-2}, where ten plants were cultivated per plot. A 3:10:10 gm^{-2} ratio of $N:P_2O_5:K_2O$ fertilizers were used during sowing. Three seeds were planted per hill, of which only one was retained after unifoliate leaves had fully expanded. Irrigation and pesticide application were made regularly.

4.3.2. Photosynthetic phenotyping

Photosynthetic phenotyping used in this experiment is the same as in Chapter three. Briefly, a high-throughput infrared gas analyzer, MIC-100, was used to measure A_{sat} in JMC accessions. Three measurements were conducted in 2019 (5 Aug, 13 Aug, and 18 Aug) and five in 2020 (6 Aug, 14 Aug, 21 Aug, 27 Aug, and 31 Aug). The data was collected from 9:00 AM to 1:30 PM from a central leaflet of a fully expanded and recently matured trifoliolate from flower initiation stage (R_1) to seed initiation stage (R_5). I measured A_{sat} in 360 and 240 leaves on each measurement day for 2019 and 2020, respectively, which collectively sums up to 2280 samples. Although three and four samples were taken for each accession in 2019 and 2020, only one plant was sampled at a time so that A_{sat} is distributed throughout the survey time to compensate for the diurnal

privilege of accessions (please refer to "**3.3.2 Photosynthetic phenotyping"** in Chapter three for further details).

4.3.3. Data preprocessing

Since A_{sat} was measured eight times in 2019 and 2019, dimensionality reduction and the removal of collinearity were necessary before conducting genetic analysis. The data was standardized so that each measurement had a mean of zero and a standard deviation of one. Later, T-distributed Stochastic Neighbor Embedding (t-SNE) was used to reduce the dimensionality of the data from eight variables (eight A_{sat} measurements) to two components (SNE1 and SNE2). Data from SNE1 and SNE2 was then subjected to genetic analysis.

4.3.4. Genome-wide association study (GWAS)

The SNP dataset obtained for 198 soybean accessions of mini-core collection (Kajiya-Kanegae et al., 2021) was used in this study. The SNP dataset with minor allele frequencies (MAF) of more than 5% for the 74 accessions was filtered using PLINK 2.0 (Chang et al., 2015). Association tests between the t-SNE and the SNPs dataset were conducted using a function of 'association.test' in Gaston package ver. 1.5.7 in R (McKenna et al., 2010). First, the mixed linear model (MLM) included five principal components as fixed effects. A genetic relationship matrix (GRM) and p values of the marker-trait associations were calculated using the 'GRM' function and Wald's test, respectively. Manhattan plot and Q-Q plot were drawn using the 'manhattan' and 'qqplot.pvalues' functions, respectively. The genome-wide significant threshold was obtained based on a false discovery rate (FDR) at a 5% level using the 'p.adjust' function in R (Benjamini and Hochberg, 1995).

4.3.5. Gene selection

All genes and their annotations were in a 250 kbp region from both sides of the significant peak detected in GWAS, together with annotations from the reference genome sequence Gmax_275_v2.0 from Phytozome (Phytozome v.12, **Appendix 1**). About 19 genes with photosynthesis-related functions or higher expression in photosynthesis-related plant tissues were selected based on the Phytozome database search. The SNP data and Indels in these genes among the 74 accessions were extracted from Illumina read mapping data (Kajiya-Kanegae et al., 2021) using CLC Genomics Workbench 12 (Qiagen, Germany).

4.3.6. RNA expression

On 24 Aug 2021, A_{sat} was measured in the central leaflets of 14 randomly selected JMC accessions and took about 10 mg of a fresh tissue sample for gene expression analysis immediately after the A_{sat} measurement. Leaf samples were frozen in liquid nitrogen and stored at -60 ºC. Total RNA was extracted by TRI Reagent (Molecular Research Center, USA). First-strand cDNA was synthesized from 500 ng of total RNA by ReverTra ACE qPCR RT Master Mix with gDNA Remover (TOYOBO, Japan). Quantitative RT-PCR was performed according to the manufacturer's instructions by diluting the synthesized cDNA ten folds and using a KAPA SYBR FAST qPCR Kit (KAPA Biosystems, USA) and a ViiA7 real-time PCR system (Thermo Fisher Scientific, USA). The qRT-PCR conditions were as follows: (1) initial denaturation at 95 ºC for 30 s; (2) 40 cycles of denaturation at 95 ºC for 5 s; and (3) final annealing and elongation at 60 ºC for 30 s. Expression levels were standardized using a VPS-like gene (Glyma.09G196600). The primer pairs used for quantitative RT-PCR are listed in **Table 4.1**.

Table 4.1: pairs of primers used for Quantitative RT-PCR

Primer name	Primer sequence (5'-3')
Glyma.17G226100 qRT-PCR F	ATTTAGAGATGTGGATAACGTG
Glyma.17G226100 qRT-PCR R	AAGATTGAGTGCAACAACATTA
Glyma.17G226700 qRT-PCR F	AAACAAGAGATTGAGCATGCA
Glyma.17G226700 qRT-PCR R	ATCTTAAAGACGCGATCTACG
Glyma.17G226900 qRT-PCR F	TCATCCAACCACAGGAAACG
Glyma.17G226900 qRT-PCR R	ACATATTCTAGCATTACAAGTTC
VPS-like qRT-PCR F	AAAGAGTCTCATCCCACAAC
VPS-like qRT-PCR R	CGCATATTCCCAATCTCAGA

4.4. Results

4.4.1. Distribution of A_{sat}

As expected, there were significant variations of A_{sat} among accessions and between years. The range of A_{sat} in JMC accessions was from 11.17 to 32.25. The frequency of accessions with lower photosynthetic capacity was higher at the beginning and end of August, whereas accessions generally had better photosynthetic performance in mid-August. Many newly studied accession were observed with a photosynthetic capacity higher than the photosynthetically elite soybean cultivar of Fukuyutaka (Fu), especially in the first half of August (refer to **Chapter 3 Figure 3.1**). Photosynthetic phenotyping was conducted eight times, however, t-SNE compressed variations of A_{sat} into two components. SNE1 explained only 22.9% variations, whereas SNE2 explained 77.1% variations in A_{sat} measurements. SNE1 had a negative correlation with all the A_{sat} measurements and was particularly strong with measurements taken on 5 Aug, 2019 and 6 Aug, 2020. In contrast, SNE2 was positively correlated with all the A_{sat} measurements (**Table 4.2**).

Table 4.2: Correlation coefficient between t-SNE components and A_{sat} in each measurement date. SNE1 and SNE2 show first and second t-SNE components, respectively.

Date	SNE1	SNE2
05 Aug, 2019	-0.65	0.55
13 Aug, 2019	-0.13	0.52
18 Aug, 2019	0.13	0.57
06 Aug, 2020	-0.75	0.37
14 Aug, 2020	-0.34	0.57
21 Aug, 2020	-0.57	0.65
27 Aug, 2020	-0.13	0.76
31 Aug, 2020	-0.30	0.85

4.4.2. Genetic characterization and RNA expression of the candidate genes

Using A_{sat} measurements directly, I found some association between A_{sat} and genomic regions, but they were neither consistent nor noise-free (**Appendix 2**). Due to limited total variations, I preferred not to use the individual measurement for GWAS. Instead, the GWAS was conducted using SNE components as phenotypes. Based on **Table 4.2**, SNE2 was interpreted to represent the genetic potential of A_{sat}. GWAS using t-SNE components revealed a region on chromosomes 17 and 14 associated with SNE components variations (**Figure 4.1** and **Figure 4.2**). Among SNE components, SNE2 had a strong correlation with A_{sat}. The significance of this correlation dropped beyond -log (p) of 8. SNE1, however, was correlated with $A_{sat,}$ but the strength of the correlation was poor than that of SNE2. As a result, I did not consider SNE1 for further analysis.

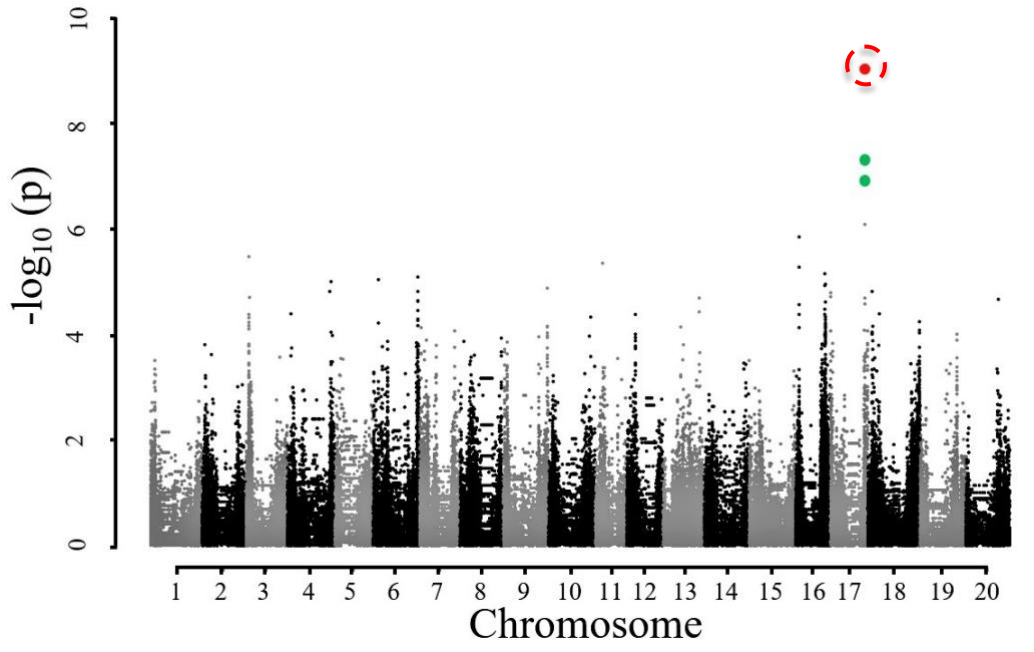


Figure 4.1: Manhattan plot for showing significant association of SNPs with SNE2. Red and green dots on chromosome 17 represent significant and suggestive associations where -log (*p*) value is less than 8 and 5, respectively.

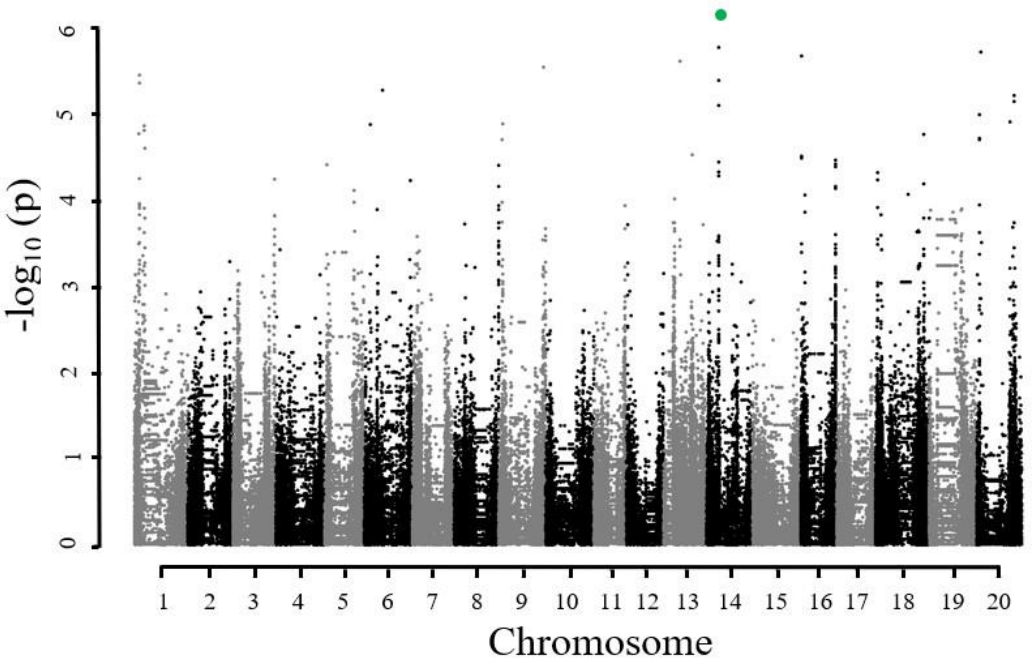


Figure 4.2: Manhattan plot for showing suggestive association of SNPs with SNE1. The green dot on chromosome 14 represents the suggestive association where the – log (*p*) value is less than 5.

In a 500 kbp region spanning both sides of the significant peak detected in GWAS of SNE2, 42 genes are annotated (**Appendix 1**). Among them, 19 genes were proteins

assumed to be involved in photosynthesis-related functions. RNA expression analysis of the three candidate genes, *Glyma.17G226100*, *Glyma.17G226700*, and *Glyma.17G226900*, revealed a strong correlation between the expression levels of *Glyma.17G226700*, and A_{sat} (r = 0.72), g_s (r = 0.69), transpiration or E (r = 0.62), ETR (r =0.56), and leaf temperature or *Tleaf* (r = -0.60); however, there was no correlation between the expression of this gene and *intercellular* CO_2 *concentration* (C_i) (**Figure 4.3**). The correlation of *Glyma.17G226100* and *Glyma.17G226900* with A_{sat} is shown in **Figure 4.4A** and **Figure 4.4C**.

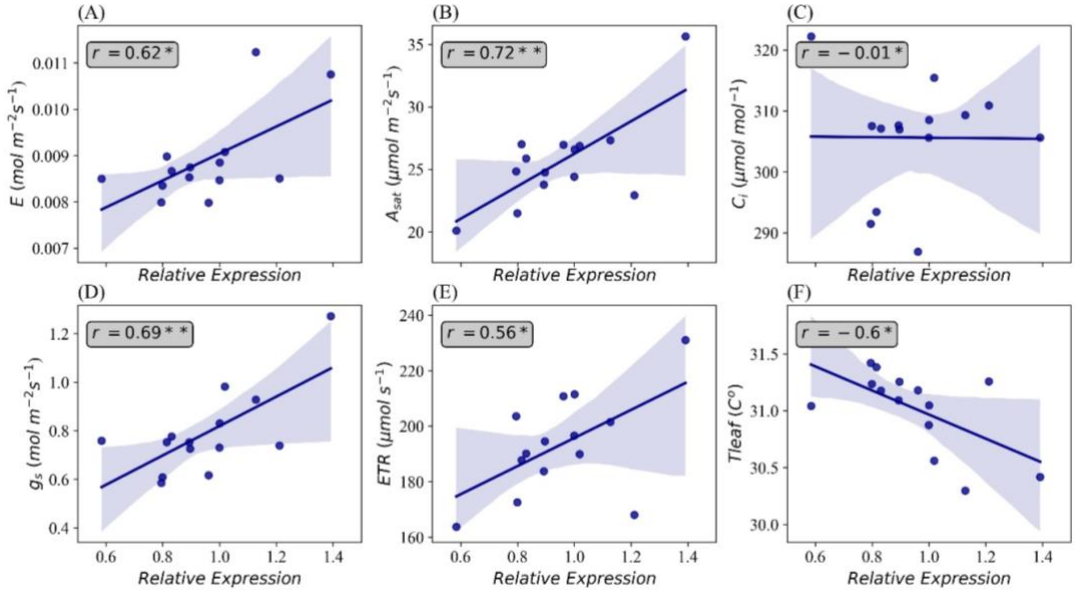


Figure 4.3: Correlations between relative gene expression of *Glyma.17G226700* and transpiration (**A**), light-saturated photosynthesis (**B**), intercellular CO_2 concentration (**C**), stomatal conductance (**D**), electron transfer rate (E), and leaf temperature (**F**). Shaded areas indicate 95% confidence intervals. Annotations inside the plots indicate Pearson's correlation coefficient. Each data point is the average of four observations. *p < 0.05, **p < 0.01.

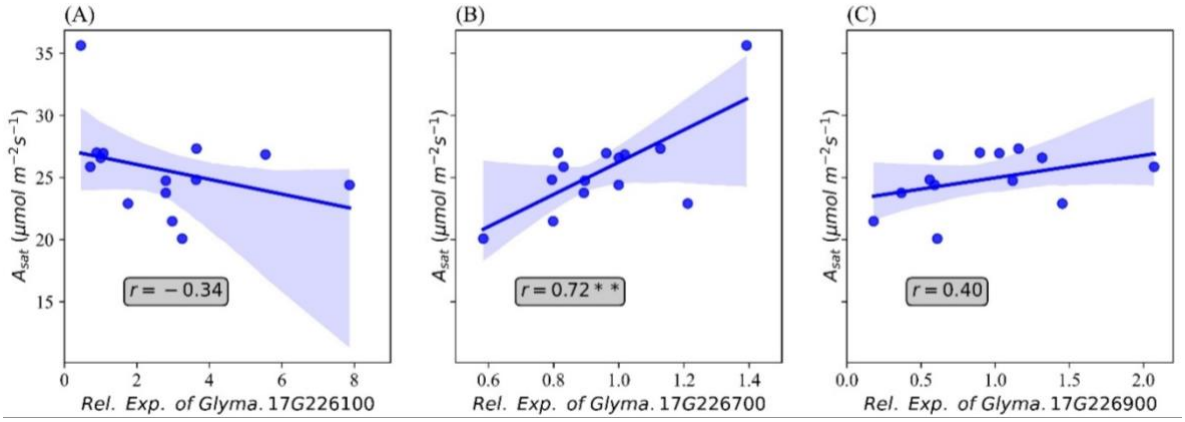


Figure 4.4: Correlations between relative expression of candidate genes and A_{sat}. Panels show scatter plots of A_{sat} and the relative expression of *Glyma.17G226100* (**A**), *Glyma.17G226700* (**B**) and Glyma.17G226900 (**C**). Shaded areas show a 95% confidence interval. Boxes inside the plots show the correlation coefficient. $^{**}p < 0.01$.

4.5. Discussion

The economic importance of soybean due to its higher protein and oil contents in its seed mass and its use in industry and biofuel production have made it an important target for genetic improvement (Homrich et al., 2012). Understanding the genetic architecture of photosynthetic capacity in Japanese soybean germplasm is one of the vital strategies to breed it for higher productivity. The photosynthetic capacity was measured in JMC to identify the underlying genetic inferences of light-saturated photosynthesis (A_{sat}) and find useful material for improving photosynthesis.

4.5.1. A_{sat} distribution

Photosynthesis is a physiological process that can be rate-limited by many factors. Even if the environment is kept optimal, processes within photosynthetic pathways impose limitations on it. Keeping the volatile nature of photosynthesis in mind and the fact that the growth environment and growth stages affect it, measuring it repeatedly is essential to determine the photosynthetic potential of crop species. Screening germplasm that is consisted of accessions with various maturity groups makes single, time-specific

photosynthetic measurement impractical. This is mainly due to the growth-related responsiveness of photosynthetic capacity in various accessions. For instance, some accessions could be in the active seed filling stage with higher photosynthetic capacity while another set of accessions may either be in late seed filling or late vegetative growth stages where photosynthetic capacity is naturally low. Measuring A_{sat} with higher frequencies in this experiment was necessary. It not only accounted for the overall distribution of A_{sat} during the reproductive growth stages in which the demand for assimilates is usually high, but it also harnessed cross-year variations within each accession, making phenotypic variations much more robust and reliable than sticking to a single measurement at a particular growth stage.

4.5.2. t-SNE analysis

Dimensionality reduction combined with GWAS to identify novel loci for different quantitative traits has previously been used (Liu et al., 2021; Liang et al., 2018; Reid et al., 2016; Yano et al., 2019; Yang et al., 2016). Here, I used the t-SNE method for the dimensionality reduction of A_{sat} measured eight times in two consecutive years. The use of t-SNE has some important implications. It is a machine learning technique that seeks to identify patterns in a highly dimensional dataset. Its advantage lies in its ability to preserve local points, meaning points close to each other in the actual data will also tend to be close to each other in the t-SNE components. t-SNE works well for linear and non-linear data. Moreover, it extracts factors underlying quantitative traits and is also a very robust method for visualizing high-dimensional data in two- or three-dimensional panels (Jake, n.d.; Kumar, 2019).

The t-SNE analysis revealed some suitable photosynthetic characteristics of the JMC. The nearest neighboring accuracy was 76%, meaning local patterns were well reflected in

the t-SNE components. The strength and type of correlations between actual data and t-SNE components indicate that SNE2 represents the genetic potential of photosynthetic capacity, whereas SNE1 explains seasonal variations in A_{sat} because of their high correlation with measurements at the beginning of Aug in both years (**Table 4.2**).

4.5.3. Genetic architecture of A_{sat}

The use of SNE2 in GWAS has two critical implications. Firstly, SNE2 explains a significant part of variations in A_{sat} values, and secondly, it had a strong positive correlation with all the measurement dates (**Table 4.2**). GWAS using SNE2 identified a genomic region at chromosome 17 (**Figure 4.1**). The accession number used in the present study was relatively small for GWAS, and the detected peak should be interpreted carefully. Increasing the number of accessions had some practical limitations. For instance, measuring A_{sat} in a large population requires significant time. Provided the dynamic nature of A_{sat}, lengthy measurement in changing diurnal conditions could impose great environmental constraints, thus fading the genetic potential of accessions. Reducing the number of biological replications for each accession to increase the accession number could lead to a biased representation of A_{sat}, meaning the higher the number of biological replications of A_{sat} for a specific accession, the more accurate the A_{sat} will be. Although a high throughput photosynthetic machine was used, it is still labor-intensive and time-consuming, thus inflecting phenotypic constraints. However, the repeated measurement of A_{sat} in 8 environments across two years and subsequent dimensionality reduction enabled us to detect the possible genomic region controlling the photosynthetic capacity.

One of the genes near the identified SNP is *Glyma.17G226700*, encoding G protein alpha subunit 1 (GPA1), which binds to abscisic acid (ABA). GPA1 mutants have been reported to have increased transpiration efficiency by minimizing g_s and stomatal density

(Nilson and Assmann 2010). The loss-of-function mutation of GPA1 led to a decrease in stomatal density in the lower epidermis of *Arabidopsis* cotyledons suggesting that GPA1 positively regulates stomatal density in response to environmental and developmental cues (Zhang et al., 2008). GPA1 has also been reported to regulate chloroplast development and response mechanism to intracellular and intercellular signals (Zhang et al., 2009). In a previous study on soybean, *Glyma.17G226700* was also detected through GWAS of photosynthetic traits concerning phosphorus efficiency (Lü et al., 2018). The positive relationship between the gene expression level of *Glyma.17G226700* and A_{sat} (**Figure 4.3B**) in this study suggests that GPA1 could be an essential factor in determining the natural variation of photosynthetic capacity in soybean.

The selection of genes in a 250 kbp region spanning both sides of the significant peak was based on the average linkage disequilibrium (LD) which is reported by (Kajiya-kanegae et al., 2021). Region-specific LD was not calculated in this experiment, but a magnified Manhattan plot did not show decay in a 250 kbp upstream region, suggesting there could be other potential genes. Nevertheless, further analysis of genes and genes' functions in regulating photosynthesis is needed to utilize genetic resources in future breeding programs. Other recent studies focusing on the genetic architecture of photosynthetic capacity have also reported several loci. None of them, however, overlapped with this study. Lopez et al. (2019) identified loci related to photosynthetic capacity on chromosomes 3 and 15 in the NAM population of the US soybean. Wang et al. (2020) also reported loci affecting the photosynthetic rate on chromosomes 6, 16, 18, and 19 in Chinese soybean germplasm. Yang et al. (2022) identified some loci controlling chlorophyll fluorescence on chromosomes 2, 4, 7, 8, 14, 15, 16, 18, 19, and 20 in Chinese soybean germplasm. Detecting various genomic regions in studies, including this report,

suggests that several genes control photosynthesis. It also implicates the complex genetic regulation and environmental interaction of soybean leaf photosynthesis.

4.6. Conclusion

This study reports detailed information about the genetic architecture underpinning photosynthesis. Here, I conducted a GWAS analysis to identify the genetic architecture of A_{sat} in the Japanese soybean accession. The GWAS revealed a locus on chromosome 17 that was significantly correlated with SNE2. Further genomic studies suggested the involvement of a gene, *Glyma.17G226700*, encoding GPA1. RNA expression analysis of *Glyma.17G226700* had a significant positive correlation with A_{sat}. These results indicate that exploring the genetic architecture of soybean germplasms could be a fast-forward way to enhance photosynthetic capacity. Further study may be needed to elucidate the genetic architecture and the gene functions underpinning A_{sat} in the Japanese soybean germplasm.

Chapter Five

General Discussion

5.1. Overview

Enhancing photosynthetic capacity can increase biomass and seed yield (Evans, 2013; Gu et al., 2014; Reynolds et al., 2011; Zhu et al., 2010). One way of photosynthetic enhancement is to assess and explore natural and induced genetic variations in plant germplasm, hunt promising individuals, study their physiological mechanisms and genetic architecture, and identify useful material that would help boost photosynthesis capacity in plant species. This study dug into the photosynthetic capacity of backcross-derived progenies from a cultivated soybean and its distant wild relative, *Glycine tomentella*, Hayata, and looked at the yield potential. The photosynthetic capacity of the Japanese soybean core collection (GmJMC) was assessed, and some valuable candidates were identified. After reviewing the natural variation of photosynthetic capacity in GmJMC, I studied the genetic architecture of soybean light-saturated photosynthesis, identified a gene, and analyzed the expression of the gene.

5.2. Understanding the diversity of photosynthesis in crop species

With the availability of DNA sequencing, the underlying genetic architecture associated with photosynthetic phenotyping can help breeders use the genetic diversity in crop panels to improve photosynthetic capacity. Genetic architecture underlying

phenotypic variations is increasingly becoming popular but harnessing phenotypic variations is still a significant drawback in the studies toward greater photosynthetic capacity. Photosynthetic capacity under saturating light conditions is crucial as it shows no limitations of available sunlight. However, photosynthesis is limited by several other factors besides sunlight. CO_2 diffusion from the ambient atmosphere into the intercellular leaf spaces, termed stomatal conductance (g_s), can be affected by the boundary layer resistance, stomatal resistance, stomatal number, and the size of intercellular spaces, let alone environmental conditions. The transfer of CO_2 from intercellular spaces to the site of carboxylation (mesophyll conductance, g_m) can be affected by the mesophyll surface exposed to the intercellular spaces, dissolving in H_2O, aquaporins, and many more. CO_2 diffusion is not the only problem to consider. Rubisco, its activity, affinity, and specificity are the other major problems in C_3 crops that pose problems to CO_2 fixation. Even if neither the sunlight nor the Rubisco is a limitation, the substrate, RuBP, that binds to CO_2 to form simple sugars is becoming a limiting factor for the photosynthetic capacity of crops. Other problems can be sourced to nutrients and the biochemical components of photosynthetic machinery such as nitrogen, phosphorus, chlorophyll, etc. Photosynthesis also has structural drawbacks, such as thylakoid composition, stroma, phloem loading and unloading, and source and sink relations. Limitation in photosynthetic machinery varies from plant to plant and time to time, making it difficult to improve it. Considering this complexity, finding materials that exhibit unique traits regarding photosynthesis would help better understand improving efforts. Therefore, there is an urgent need to explore physiological and genetic diversity in soybean, identify promising individuals with better photosynthetic capacity or fewer limitations, study their genetic architecture and then use them in plant breeding to increase potential productivity through photosynthetic enhancement.

5.3. Photosynthetic enhancement in wild-related progenies

The use of wild-related species in plant breeding is vital for broadening the gene pool of crop species as they have diverse genomic constitutions with valuable alleles. The current advancement in genomic technology has allowed us to acquire phylogenetic information about crop species and use them in future breeding (Pour-Aboughadareh et al., 2021). Crop wild relatives (CWR) are adapted to various environments and contain different traits for biotic and abiotic stress. The use of CWR has an annual economic contribution of $267 to $384 million, mainly due to the beneficial traits it possesses and introduced to cultivated species (Seiler et al., 2017). With global warming, the thermotolerance of crop species will soon reach its limit, which necessitates the identification and introduction of tolerant traits (Atwell et al., 2014). Thermotolerance in growth and photosynthetic superiority has been reported in rice CWR comparing cultivated species when exposed to the extreme heat of 45 ^{o}C (Phillips et al., 2021). The genetic information of wild-related soybean species has not been utilized in soybean breeding due to low cross-ability, pre-matured pod abortion, and drastic genetic architecture (Ladizinsky et al., 1979). Singh et al. (1990), on the other hand, were able to make a cross between soybean (*Glycine max*) and *Glycine tomentella* Hayata. *G. tomentella* Hayata (2n = 78) is native to Australia and a perennial relative in the tertiary gene pool of soybean. Progenies derived from the backcross of *G. tomentella* Hayata and *G. max* (cv. Dwight) (Singh and Nelson, 2015) showed higher yield in controlled as well as field experiments (Akpertey et al., 2018; Begemann, 2015). Detailed genetic studies revealed that these genotypes have 1% introgression of *G. tomentella* (Akpertey et al., 2018). Little was known about the photosynthetic potential of these backcross-derived progenies. The photosynthetic and yield superiority in these backcross-derived progenies comparing Dwight reveals that wild CWR carries multiple beneficial traits. If these traits

are incorporated into soybean breeding, they may result in significant yield gains. Increasing stomatal conductance leads to unnecessary water loss and lower water-use efficiency (WUE). Enhanced mesophyll activity is, thus, a good sign in enhancing CO_2 fixation. Enhanced mesophyll activity can also mean better carboxylation activity or substrate regeneration. In this case, CWR could be an effective tool to rump up CO_2 fixation in the Calvin cycle, far beyond the intercellular spaces, indicating that using soybean CWR in soybean breeding can enhance photosynthesis and WUE in addition to biomass and seed yield.

5.4. Harnessing natural variation in crop germplasms

Natural selection for improved plant architecture or growth performance has reduced genetic diversity in many crop species. Reduction in genetic diversity risks the prospects of improving crops. In wheat, for instance, selection for better traits reduced genetic diversity from 1950 to 1989. However, it was increased from 1990 to 1997, which means breeders averted the narrowing of the germplasm base and consequently increased genetic diversity through the introgression of novel traits (Reif et al., 2005). Genetic loss occurs in all crop species, but the extent of loss differs. The most considerable loss of genetic diversity occurred in elite open-pollinated cultivars and inbred lines, whereas the minor loss occurred in wild germplasms, followed by landraces (Rauf et al. 2010). The lack of photosynthetic enhancement is believed to be due to the gap between the germplasm used in crop breeding and the one stored in gene banks. Historic breeding efforts for crop improvement did not include photosynthetic gain because there were no or fewer photosynthetic measurement systems and the belief that seed yield is sink limited rather than source limited. FACE experiments on super hybrid rice cultivars, nevertheless, proved that seed yield in these cultivars is source-limited, necessitating the increase of

photosynthetic capacity. To enhance photosynthetic capacity in any possible way, it is first essential to identify useful information or traits in related crop species that can be later used in plant breeding or genetic manipulation. To do so, we need to use the landraces or local accessions that have undergone only minor losses of genetic diversity in crop breeding.

This study assessed the natural variations of photosynthetic capacity in the Japanese soybean germplasm. This germplasm has extensive geographical and morphological variations from canopy to seed level, but little was known about its photosynthetic potential. The present study gave us the overall idea of photosynthetic performance in the Japanese soybean germplasm. Accessions in Cluster2 can be used in further studies to extract valuable insight for later use. Studying photosynthetic capacity in Japanese soybean germplasm opened the door to further discussion about its photosynthetic potential. Identifying some promising individuals, such as JMC47 in this study, have also proved that plant germplasm and its genetic diversity may possess a "locked treasure of useful traits" that can be advantageous if incorporated into plant improvement strategies.

5.5. Understanding the genetic architecture of phenotypic variations

Natural variations in the genetics of animals and plants are the pivotal point of evolutionary and functional biology. Studying genomic revolutions profoundly affects natural variations, making it possible to understand the complex genetic architecture of phenotypic variations (Borevitz and Nordborg, 2004). Phenotypic variations are central in adaptation, natural selection, and response to diseases of all living beings. Genetic mapping is the primary method for identifying complex traits for phenotypic variations. These traits pose a challenge because of gene-gene interaction, gene-environment

interactions, heterogeneity, and other technical or statistical power (Glazier et al., 2002). The availability of high-resolution sequencing data and its lower prices have allowed us to identify small haplotype blocks that are highly correlated with complex trait variations (Brachi et al., 2011).

Photosynthesis is a complex quantitative trait that is controlled by multiple genes. With the current advancement of sequencing technology, especially Genome-wide association studies (GWAS), the genetic basis of photosynthetic variation is being increasingly explored. GWAS has identified several loci associated with photosynthetic capacity (Lopez et al., 2019; Vieira et al., 2006; Wang et al., 2020). In this study, a genomic region was identified that was highly associated with A_{sat}. However, there was no overlapping of the genomic regions among earlier studies, which could be because several genes control photosynthesis. The significant correlation between A_{sat} and the RNA expression of the identified gene, *Glyma.17G226700*, suggests that the gene may be involved in the A_{sat} performance of the Japanese soybean. The gene in this study was previously documented in soybean (Lü et al., 2018). In their studies, this gene was mainly related to C_i, which could be the phosphorus treatment. In this research, on the other hand, it was associated with A_{sat}. This gene is believed to be coding G protein alpha subunit 1 (GPA1). GPA1 binds to abscisic acid and helps increase transpiration efficiency by minimizing g_s. For having broad geographical and morphological variations, further genetic studies are essential in understanding the genetic architecture of various photosynthetic traits in the Japanese soybean germplasm.

5.6. Concluding remarks

The quest for harnessing photosynthetically elite accessions in plant germplasm is vital to identifying useful material for breeding. Here, three major experiments were conducted and reported new findings. In the first experiment, I studied the photosynthetic capacity of two backcross-derived perogies and their recurrent parent (Dwight) and observed that the progenies had better photosynthetic rates and biomass. Progenies had higher mesophyll activity due to either higher mesophyll conductance or enzymatic advantages in Calvin Cycle. In the second experiment, I screened the Japanese soybean germplasm and clustered them into four clusters based on A_{sat}. Among clusters, Cluster2 had the best photosynthesizing accessions. I observed that Japanese accessions were limited by CO_2 diffusion and light utilization. JMC47 was considered the best performing accession throughout the growing season in 2021, which had higher stomatal conductance, higher stomatal density, higher light utilization, and higher electron transfer rate. In the third experiment, I studied the genetic architecture of the photosynthetic capacity of the Japanese soybean and found a locus on chromosome 17. Further study suggested that a gene, *Glyma.17G226700*, controls the A_{sat} performance of the genotypes. *Glyma.17G226700* encodes G protein alpha subunit 1 (GPA 1), which increases transpiration efficiency by minimizing stomatal conductance and density. RNA expression analysis of the gene also had a significant correlation with A_{sat}. These studies suggest that exploring the photosynthetic potential in crop germplasm could be a promising avenue toward enhancing resource use efficiencies.

References

Ainsworth, E.A., Yendrek, C.R., Skoneczka, J.A., and Long, S.P. (2012). Accelerating yield potential in soybean: Potential targets for biotechnological improvement. *Plant Cell Environ.* 35, 38–52. doi:10.1111/j.1365-3040.2011.02378.x.

Akpertey, A., Singh, R.J., Diers, B.W., Graef, G.L., Mian, M.A.R., Shannon, J.G., Scaboo, A.M., Hudson, M.E., Thurber, C.S., Brown, P.J., and Nelson, R.L. (2018). Genetic introgression from *Glycine tomentella* to soybean to increase seed yield. *Crop Sci.* 58, 1277–1291. doi:10.2135/cropsci2017.07.0445.

Pieters, A.J., Graterol, E., Reyes, E., Alvarez, R. and Gonzalez, A. (2011). Fifty years of genetic improvement of rice in Venezuela: What has been achieved?. *Interciencia*, 36, 943-948.

Available at: https://www.cabdirect.org/cabdirect/abstract/20123096510

[Accessed: June 15, 2022].

Allison, J. C. S., Williams, H. T., and Pammenter, N. W. (1997). Effect of specific leaf nitrogen content on photosynthesis of sugar- cane. *Ann. Applied Biol*., 131, 339–350. doi: org/10.1111/j.1744-7348.1997.tb05160.x.

Ares, A., Fownes, J.H., and Sun, W. (2000). Genetic differentiation of intrinsic water-use efficiency in the Hawaiian *Acacia kao*. *Int. J. Plant Sci.* 161, 909-915.

Available at: https://www.journals.uchicago.edu/doi/epdf/10.1086/317559

[Accessed: June 14, 2022].

Atwell, B.J., Wang, H., and Scafaro, A.P. (2014) Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*?. *Plant Sci*., 215,48–58.

doi: org/10.1016/j.plantsci.2013.10.007

Begemann, S. (2015). AGWEB: Soybean steal the spot-light. Available at: https://www.Agweb.com/mobile/article/soybeans-steal-the-spotlight-naa-sonja-gjerde/ [Accessed: December 09, 2015].

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57, 289-300. doi: org/10.1111/j.2517-6161.1995.tb02031.x.

Borevitz, J.O., and Nordborg, M. (2003). The impact of genomics on the study of natural variation in Arabidopsis. *Plant Physiol*. 132, 718. doi:10.1104/PP.103.023549.

Boussadia, O., Steppe, K., Zgallai, H., Ben El Hadj, S., Braham, M., Lemeur, R., and Van Labeke, M.C. (2010). Effects of nitrogen deficiency on leaf photosynthesis, carbohydrate status, and biomass production in two olive cultivars, "Meski" and "Koroneiki." *Sci. Hortic.* (*Amsterdam*), 123, 336–342.

doi: org/10.1016/j.scienta.2009.09.023.

Brachi, B., Morris, G.P., and Borevitz, J.O. (2011). Genome-wide association studies in plants: The missing heritability is in the field. *Genome Biol*. 12, 1–8. doi:10.1186/GB-2011-12-10-232/FIGURES/2.

Brodersen, C., Lavergne, S., and Molofsky, J. (2008). Genetic variation in photosynthetic characteristics among invasive and native populations of reed canarygrass (*Phalaris arundinacea*). *Biol. Invasions*, 10, 1317–1325.

doi: org/10.1007/s10530-007-9206-x.

Brown, R. H., and Bouton, J. H. (1993). Physiological and genetics of interspecific hybrids between photosynthetic types. *Plant Physiol., Plant Mol. Biol*., 44, 435-456. doi: org/10.1146/annurev.pp.44.060193.002251.

Buttery, B. R., and Buzzell, R. L. (1977). The relationship between chlorophyll content and rate of photosynthesis in soybeans. *Canad. J. Plant Sci.*, 6, 190-197.

doi: 10.4141/cjps81-029.

Buttery, B.R., Buzzell, R.I., and Findlay, W.I. (1981). Relationships among photosynthetic rate, bean yield, and other characters in field-grown cultivars of soybean. *Canad. J. Plant Sci.* 61, 191–198. doi:10.4141/cjps81-029.

Carver, B.F., Hohnson, R.C. and Rayburn, A.L. (1989). Genetic analysis of photosynthetic variation in hexaploid and tetraploid wheat and their interspecific hybrids. *Photosynth. Res*., 20, 105–118. doi: org/10.1007/BF00034120.

Chang C., Chow C., Tellier L., Vattikuti S., Purcell S., and Lee J. (2015). Second-generation Plink: rising to the challenge of larger and richer datasets, *GigaScience* 4,7. doi: 10.1186/s13742-015-0047-8.

Chavarria, G., Carerzan, A., Muller, M., and Rakocevic, M. (2017). Soybean architecture plants: From solar radiation interception to crop protection. Soybean-the basis of yield, biomass, and productivity. *IntechOpen, London*. 15-35. doi: 10.5772/67150.

Christensen, J. H., and Hewitson, B. (2007). Regional climate projections. In climate change 2007: The physical science basis. Contribution of working group 1 to the fourth assessment report of the intergovernmental panel on climate change. 265. *University Press, Cambridge*. IPCC 2007.

Chung, G., and Singh, R. J. (2008). Broadening the genetic base of soybean: A multidisciplinary approach. *Crit. Rev. Plant Sci.,* 27, 295–341.

doi: org/10.1080/ 07352680802333904.

Cober, E.R., and Morrison, M.J. (2010). Regulation of seed yield and agronomic characters by photoperiod sensitivity and growth habit genes in soybean. *Theor. Appl. Genet.* 120, 1005–1012. doi:10.1007/s00122-009-1228-6.

Condon, A.G., Richards, R.A., Rebetzke, G.J., and Farquhar, G.D. (2004). Breeding for high water-use efficiency. *J. Exp. Bot*., 55, 2447–2460. doi: org/10.1093/jxb/erh277.

Cornish, K., Radin, J. W., Turcotte, E. L., Lu, Z., and Zeiger, E. (1991). Enhanced photosynthesis and stomatal conductance of Pima cotton (*Gossypium barbadense* L.) Bred for increased yield. *Plant Physiol.*, 97, 484–489. doi: 10.1104/pp.97.2.484.

Da-Yong, L., Zhi-an, Z., Dian-jun, Z., Li-yan, J., and Yuan-li, W. (2012). Comparison of net photosynthetic rate in leaves of soybean with different yield levels. *J. NE Agri. Univ. (English Edition)*. *Asian Pac. Trop. Med. Press*, 19, 14–19.

doi: 10.1016/S1006-8104(13)60017-3.

de Oliveira Silva, F. M., Lichtenstein, G., Alseekh, S., Rosado-Souza, L., Conte, M., Suguiyama, V. F., Lira, B.S., Fanourakis, D., Usadel, B., Bhering, L.L. and DaMatta, F.M.,. (2018). The genetic architecture of photosynthesis and plant growth-related traits in tomatoes. *Plant Cell Environ*., 41, 327–341. doi: 10.1111/PCE.13084.

Evans, J.R. (2013). Improving Photosynthesis. *Plant Physiol*. 162, 1780–1793. doi:10.1104/pp.113.219006.

Evans, L. T., and Dunstone, R. L. (1970). Some physiological aspects of evolution in wheat. *Aus. J. Biol. Sci.*, 23, 725–741. doi: org/10.1071/ BI9700725.

Evans, L.T. (1997). Adapting and improving crops: The endless task. *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* 352: 901–906. doi: org/10.1098/rstb.1997.0069.

Fay, P.A., Knapp Eay, A. K., and Fay, P.A. (1995) Stomatal and photosynthetic responses to shade hi sorghum, soybean, and eastern gamagrass. *Physiol. Planta*., 94, 613-620, doi: org/10.1111/j.1399-3054.1995.tb00975.x.

Fischer, R.A., Rees, D., Sayre, K.D., Lu, Z.M., Condon, A.G., and Larque Saavedra, A. (1998). Wheat yield progress associated with higher stomatal conductance and photosynthetic rate and cooler canopies. *Crop Sci.*, 38, 1467–1475.

doi: org/10.2135/cropsci1998.0011183X003800060011x.

Franks, P. J., Drake, P. L., and Beerling, D. J. (2009). Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: An analysis using *Eucalyptus globulus*. *Plant, Cell Environ.*, 32, 1737–1748.

doi: 10.1111/j.1365-3040.2009.002031.x.

Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Sci*., 16, 635–644.

doi: 10.1016/j.tplants.2011.09.005.

Gabriel, A. S., Ninomiya, K., and Uneyama, H. (2018). The role of the japanese traditional diet in healthy and sustainable dietary patterns around the world. *Nutrients*, 10, 173.

doi: 10.3390/nu10020173.

Garner, W. W., and Allard, H. A. (1920). Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *J. Agri. Res.* 18, 553–606. doi: 10.1175/1520-0493(1920)48<415b:EOTRLO>2.0.CO;2.

Glazier, A.M., Madeau, J.H., and Aitman, T.J. (2002). Finding genes that underlie complex traits. *Science*, 298, 2345–2349. doi: 10.1126/science.1076641.

Gu, J., Yin, X., Stomph, T.-J., and Struick, P. C. (2014). Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. *Plant Cell Environ.* 37, 22–34. doi: 10.1111/pce.12173.

Hay, W.T., Bihmidine, S., Mutlu, N., Hoang, K. Le, Awada, T., Weeks, D.P., Clemente, T.E., and Long, S.P. (2017). Enhancing soybean photosynthetic CO_2 assimilation using a cyanobacterial membrane protein, *ictB. J. Plant Physiol*. 212, 58–68. doi:10.1016/j.jplph.2017.02.003.

Heckmann, D., Schlüter, U., and Weber, A. P. M. (2017). Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Mol. Plant*, 10, 878–890. doi: 10.1016/j.molp.2017.04.009.

Hikosaka, K. (2010). Mechanisms underlying interspecific variation in photosynthetic capacity across wild plant species. *Plant Biotech*., 27, 223-229.

doi: org/10.5511/plantbiotechnology.27.223.

Holly, C., Jing, M. C., Xiangzhong, L., Paul, B., Bin ,C., and Ralf, M.S. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Wiley Online Library*, 23, 3513-3524. doi: org/10.1111/gcb.13599.

Homrich, M. S., Wiebke-Strohm, B., Weber, R. L., and Bodanese-Zanettini, M. H. (2012). Soybean genetic transformation: A valuable tool for the functional study of genes and the production of agronomically improved plants. *Genet. Mol. biology*, 35 998–1010. doi: org/10.1590/s1415-47572012000600015.

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. *Comp. Sci. Eng*., 9,90-95. doi: 10.1109/MCSE.2007.55.

Izhar, S., and Wallace D. H. (1967). Studies of the physiological basis for yields differences. III. Genetic variation in photosynthetic efficiency of *Phaseolus vulgaris*. *Crop Sci*, 7, 457-460.

doi: org/10.2135/cropsci1967.0011183X000700050015x.

Jake. H., (n.d.). DisplayR: How t-SNE works and dimensionality reduction. Available at: https://www.displayr.com/using-t-sne-to-visualize-data-before-prediction/ [Accessed on: June 18, 2022].

Janssen, P.J.D., Lambreva, M.D., Plumera, N., Bartolucci, C., Antonacci, A., Buonasera, K., Frese, R.N., Scognamiglio, V., and Rea, G. (2014). Photosynthesis at the forefront of a sustainable life. *Front. Chem.* 2, 1–22. doi:10.3389/fchem.2014.00036.

Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., and Katayose, Y. (2012). Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breed. Scie.*, 592, 566–592. doi: 10.1270/jsbbs.61.566.

Kajiya-Kanegae, H., Nagasaki, H., Kaga, A., Hirano, K., Ogiso-Tanaka, E., Matsuoka, M., Ishimori, M., Ishimoto, M., Hashiguchi, M., Tanaka, H. and Akashi, R., 2021. Whole-genome sequence diversity and association analysis of 198 soybean accessions in mini-core collections. *DNA Res*., 28, 1-13.

doi: 10.1093/dnares/dsaa032.

Kamal, N.M., Gorafi, Y.S.A., Abddelrahman, M., Abdellatef, E., and Tsujimoto, H. (2019) 'Stay-Green Trait: A prospective approach for yield potential, and drought and heat stress adaptation in globally important cereals. *Internat. J. Mol. Sci. Multidiscip. Dig. Pub. Inst.*, 20. doi: 10.3390/IJMS20235837.

Kawasaki, Y., Tanaka, Y., Katsura, K., Purcell, L. C., and Shiraiwa, T. (2016). Plant production science yield and dry matter productivity of Japanese and us soybean cultivars. *Plant Prod. Sci*., 19, 257-266. doi: 10.1080/1343943X.2015.1133235.

Khush, G.S. (2001). Green revolution: the way forward. *Nat. Rev. Genet*., 2, 815-822. doi: org/10.1038/35093585.

Koester, R.P., B.M. Nohl, B.W. Diers, and E.A. Ainsworth. (2016). Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. *Plant Cell Environ*., 39,1058–1067. doi: .org/10.1111/pce.12675.

Koester, R.P., Skoneczka, J.A., Cary, T.R., Diers, B.W., and Ainsworth, E.A. (2014). Historical gains in soybean (*Glycine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *J. Exp. Bot*. 65, 3311–3321. doi:10.1093/jxb/eru187.

Koutroubas, S., D.. Papakosta, and A. Gagianas. (1998). The importance of early dry matter and nitrogen accumulation in soybean yield. *Eur. J. Agron*. 9, 1–10. doi: org/10.1016/S1161-0301(97)00067-1.

Kromdijk, J., and Long, S.P. (2016). One crop breeding cycle from starvation? How engineering crop photosynthesis for rising CO_2 and temperature could be one important route to alleviation. *Proc. R. Soc. B* 283: 20152578. doi: org/10.1098/rspb.2015.2578.

Kromdijk, J., Głowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., and Long, S. P. (2016). Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, 354, 857–861. doi:10.1126/science.aai8878.

Kumar, H. (2019). The professional point: Advantages and disadvantages of t-SNE over PCA (PCA vs t-SNE). Available at: http://theprofessionalspoint.blogspot.com/2019/03/advantages-and-disadvantages-of-t-sne.html [Accessed on: June 18, 2022]

Kusumi, K., Hirotsuka, S., Kumamura, T., and Iba, K. (2012). Increased leaf photosynthesis caused by elevated stomatal conductance in a rice mutant deficient in SLAC1, a guard cell anion channel protein. *J. Exp. Bot*., 63, 5635-5644. doi:10.1093/jxb/err313.

Ladizinsky, G., Newell, C.A., and Hymowitz, T. (1979). Wide crosses in soybeans: Prospects and limitations. *Euphytica* 282, 421–423. doi:10.1007/BF00056600.

Lambers, H., F.S. Chapin III, and T.L., Pons. (2008). Plant ecological physiology. 11-99 Springer. Available at: https://link.springer.com/book/10.1007/978-0-387-78341-3 [Accessed: June 14, 2022].

Lawlor, D. W., Paul, M. J., and Foyer, C. H. (1999). Genetic manipulation of photosynthesis. *J. Exp. Bot.*, 50, 4–12.

Available at: http://www.jstor.org/stable/23696000 [Accessed: June 14, 2022].

Lawson, T., and Blatt, M.R. (2014). Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. *Plant Physiol*. 164, 1556–1570. doi:10.1104/pp.114.237107.

Lefebvre, S., Lawson, T., Zakhleniuk, O. V., Lloyd, J.C., and Raines, C.A., (2005). Increased sedoheptulose-1,7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. *Plant Physiol*. 138, 451–460. doi:10.1104/PP.104.055046.

LI-COR Biosciences. (2021). LI-6800 portable photosynthesis system: advancements in gas exchange and fluorescence research.Available at: https://www.licor.com/documents/02057bkvtf900ijr91z9gcwjyzwzic46 [Accessed March 22, 2022].

Li, R.-H., Guo, P.-P., Baumz, M., Grand, S., and Ceccarelli, S. (2006). Evaluation of chlorophyll content and fluorescence parameters as indicators of drought tolerance in barley. *Agri. Sci. in China* 5, 751-757 doi:10.1016/S1671-2927(06)60120-x.

Liang, X., Sha., Q., Rho, Y., and Zhang, S. (2018) A hierarchical clustering method for dimension reduction in join analysis of multiple phenotypes. *Genet. Epidemoil*., 42, 344-353. doi: org/10.1002/gepi.22124.

Liu, F., Zhou, Z., Cai, M., Wen, Y and Zhang., J. (2021) AGNEP: An agglomerative nesting clustering algorithm for phenotypic dimension reduction in join analysis of multiple phenotypes. Front. Genet., 12, 648831. doi: 10.3389/fgene.2021.648831.

Long, S.P., X.G. Zhu, S.L. Naidu, and D.R. Ort. (2006). Can improvement in photosynthesis increase crop yields?. *Plant Cell Environ*. 29, 315–330. doi: org/10.1111/j.1365-3040.2005.01493.x.

Lopez, M. A., Xavier, A., and Rainey, K. M. (2019). Phenotypic variation and genetic architecture for photosynthesis and water use efficiency in soybean (*Glycine max L.* Merr.). *Front. Plant Sci.*, 10, 1–12. doi: 10.3389/fpls.2019.00680.

Lü, H., Yang, Y., Li, H., Liu, Q., Zhang, J., Yin, J., Chu, S., Zhang, X., Yu, K., Lv, L., Chen, X., and Zhang D. (2018). Genome-wide association studies of photosynthetic traits related to phosphorus efficiency in soybean. *Front. Plant Sci.* 9. 1226. doi: 10.3389/fpls.2018.01226.

Makino, A. (2011). Photosynthesis, grain yield, and nitrogen utilization in rice and wheat. *Plant Physiol*. 155, 125–129. doi: 10.1104/pp.110.165076.

Makino, A., Mae, T., and Ohira, K. (1986). Colorimetric measurement of protein stained with coomassie brilliant blue r on sodium dodecyl sulfate-polyacrylamide gel electrophoresis by eluting with formamide. *Agric. Biol. Chem.*, 50, 1911–1912. doi: org/10.1080/00021369.1986.10867672.

Mann, C. C. (1999). Genetic engineers aim to soup up crop photosynthesis. *Science*, 283, 314-316. doi: 10.1126/science.283.5400.314.

Maurino, G., and Weber, P.M. (2013). Engineering photosynthesis in plants and synthetic microorganisms. *J. Exp. Bot*., 64, 743–751. doi: org/10.1093/jxb/ers263.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res., 20, 1297–303. doi: 10.1101/gr.107524.110.

Mendiburu, F.D. (2015). Agricolae: statistical procedures. *Agri. Res. R package, version 1.3-5*. Available at: https://cran.r-project.org/web/packages/agricolae/agricolae.pdf [Accessed: June 14, 2022].

Mitchell, P.L., and J.E. Sheehy. (2000). Performance of a potential C_4 rice: overview from quantum yield to grain yield. *Stud. Plant Sci*. 7, 145–163. doi: org/10.1016/S0928-3420(00)80012-9.

Monteith, J., (1977). Climate and the efficiency of crop production in Britain. *Philos. Trans. R. Soc. Lond*. 281, 277–294. doi: org/10.1098/rstb.1977.0140.

Morrison, M.J., Voldeng, H.D., and Cober, E.R. (2000). Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in *Canad. Agron. J.*, 92, 780–784. doi: org/10.2134/agronj2000.924780x.

Nilson, S.E., and Assmann, S.M. (2010). The α-subunit of the Arabidopsis heterotrimeric G protein, GPA1, is a regulator of transpiration efficiency. *Plant Physiol.*, 152, 2067–2077. doi: 10.1104/PP.109.148262.

Parry, M.A.J, Adnralojc, P.J., Scales, J.C, Salvucci, M.E., Carmo-silva, E., Alonso, H., and Whitney, S.M. (2012). Rubisco activity and regulation as targets for crop improvement. *J. Exp. Bot*. 64, 717-730. doi: 10.1093/jxb/ers336. X.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. *J. Mach Learn. Res*., 12, 2825-2830. Available at: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://githubhelp.com [Accessed: June 15, 2022].

Peng, S. (2000). Single-leaf and canopy photosynthesis of rice. *Stud. Plant Sci*. 7, 213–228. doi: org/10.1016/S0928-3420(00)80017-8.

Pereira, Y., Rodrigues, W., Lima, E., Santos, L., Silva, M., and Lobato, A. (2019). Brassinosteroids increase electron transport and photosynthesis in soybean plants under water deficit. *Photosynthetica*, 57, 181-191 doi: org/10.32615/ps.2019.029.

Phillips, A. L., Scafaro, A. P., and Atwell, B. J. (2021). Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures. *Plant Mol. Biol.* 1, 1-17. doi: 10.1007/s11103-021-01210-3.

Porra, R. J. (2002). The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls a and b. *Photo. Res*. 73, 149-156. doi: 10.1023/A:1020470224740.

Pour-Aboughadareh, A., Kianersi, F., Poczai, P., Moradkhani, H. (2021). Potential of wild relatives of wheat: ideal genetic resources for future breeding programs. *Agron.*, 11, 1656. doi:10.3390/AGRONOMY11081656.

Rauf, S., Silva, J.T., Khan, A.A., and Naveed, A. (2010). Consequences of plant breeding on genetic diversity. *Intern. J. Plant Breed*. 4, 1-21. Available at: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Consequences+of+plant+breeding+on+genetic+diversity&btnG= [Accessed: June 15, 2022].

Ray, D.K., Mueller, N.D., West, P.C., and Foley, J.A. (2013). Yield trends are insufficient to double global crop production by 2050. *Plos One*, 8, 1-8.

doi: 10.1371/journal.pone.0066428.

Ried, J.S., Jeff M, J., Chu, A.Y., Bragg-Gresham, J.L., Van Dongen, J., Huffman, J.E., Ahluwalia, T.S., Cadby, G., Eklund, N., Eriksson, J. and Esko, T. (2016). A principal component meta-analysis on multiple anthropometric traits identifies novel loci for body shape. *Nat. Com*., 23, 1-11. doi: 10.1038/ncomms13357

Reif, J.C., Zhang, P., Dreisigacker, S., Warburton, M.L., Van Ginkel, M., Hoisington, D., Bohn, M., and Melchinger, A.E. (2005) . Wheat genetic diversity trends during domestication and breeding. *Theor. Appl. Genet*. 110, 859–864. doi:10.1007/S00122-004-1881-8/FIGURES/3.

Revelle, W. (2022). Psych: procedures for psychological, psychometric, and personality research. *NW Univ., Evanston, Illinois*. R package, version 2.2.5, Available at: https://personality-project.org/r/psych-manual.pdf [Accessed: June 15, 2022].

Reynolds, M., Bonnett, D., Chapman, S.C., Furbank, R.T., Manè, Y., Mather, D.E., and Parry, M.A.J. (2011). Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. *J. Exp. Bot.* 62, 439–452. doi:10.1093/jxb/erq311.

Reynolds, M.P., van Ginkel, and M., Ribaut, J. (2002). Avenues for genetic modification of radiation use efficiency in wheat. *J. Exp. Bot*., 51, 459-473. doi: 10.1093/jexbot/51.suppl_1.459.

Rossi, M., Bermudez, L., and Carrari, F. (2015). Crop yield: Challenges from a metabolic perspective. *Current Opinion Plant Biol*., 25, 79–89. doi: 10.1016/j.pbi.2015.05.004.

Sakoda, K., Tanaka, Y., Long, S. P., and Shiraiwa, T. (2016). Genetic and physiological diversity in the leaf photosynthetic capacity of soybean. *Crop Sci.*, 56, 2731–2741. doi: Org/10.2135/cropsci2016.02.0122.

Sakoda, K., Yamori, W., Groszmann, M., and Evans, J.R. (2021). Stomatal, mesophyll conductance, and biochemical limitations to photosynthesis during induction. *Plant Physiol*., 185, 146-160. doi:10.1093/plphys/kiaa011.

Sands, R. D., Jones, C. A., and Marshall, E. (2014). Global drivers of agricultural demand and supply. Economic research report, 174. *U.S. Dept. of Agri., Economic Res., Services*. (No. 1477-2016-121072). Available at: https://www.ers.usda.gov/webdocs/publications/45272/49035_err174.pdf?v=0 [Accessed: June 14, 2022].

Schoen, D. J., Burdon, J. J., and Brown, A. H. D. (1992). Resistance of *Glycine tomentella* to soybean leaf rust *Phakopsora pachyrhizi* in relation to ploidy level and geographic distribution. TAG. *Theor. Appl. Genet*., 83, 827–832. doi: org/ 10.1007/BF00226704.

Seiler, G. J., Qi, L.L., and Laura F.M (2017). Utilization of sunflower crop wild relatives for cultivated sunflower improvement. *Crop Sci.*, 57, 1083-1101. doi:10.2135/CROPSCI2016.10.0856.

Sheng, M., Tang, M., Chen, H., Yang, B., Zhang, F., and Huang, Y. (2008). Influence of arbuscular mycorrhizae on photosynthesis and water status of maize plants under salt stress. *Mycorrhiza* 18, 287-296. doi: 10.1007/s00572-008-0180-7.

Silva-Perez, V., Molero, G., Serbin, S.P., Condon, A.G., Reynolds, M.P., Furbank, R.T., Evans, J.R. (2018). Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *J. Exp. Bot*. 69, 483–496. doi:10.1093/jxb/erx421.

Simkin, A.J., Mcausland, L., Headland, L.R., Lawson, T., and Raines, C.A. (2015). Multigene manipulation of photosynthetic carbon assimilation increases CO_2 fixation and biomass yield in tobacco. *J. Exp. Bot.*, 66, 4075–4090.

doi: 10.1093/jxb/erv204.

Sinclair, T.R., Tanner, C.B., and Bennett, J.M. (1984). Water-use efficiency in crop production. *Bioscience*, 34, 36–40. doi: org/10.2307/1309424.

Singh, R. J., and Nelson, R. L. (2015). Intersubgeneric hybridization between *Glycine max* and *G. tomentella*: Production of F_1, amphidiploid, BC_1, BC_2, BC_3, and fertile soybean plants. *Theoret. Appl. Genet*., 128, 1117–1136. doi: org/10. 1007/s00122-015-2494-0.

Singh, R.J., K.P. Kollipara, and T. Hymowitz. (1990). Backcross-derived progeny from soybean and *Glycine tomentella* Hayata intersubgeneric hybrids. *Crop Sci.* 30, 871-874. doi:10.2135/cropsci1990.0011183X003000040021x.

Singh, Guriqbal. (2010). The Soybean botany, production and uses. CABI, 1-40.

Specht, J.E., D.J. Hume, and S. V. Kumudini. (1999). Soybean yield potential - A genetic and physiological perspective. *Crop Sci.*, 39, 1560–1570.

doi: org/10.2135/cropsci1999.3961560x.

Stapper, M., and Fischer R. A. (1990). Genotype, sowing date, and plant spacing influence on high-yielding irrigated wheat in southern New South Wales. 3. Potential yields and optimum flowering dates. *Aust. J. Agric. Res.*, 41, 1043–1056.

doi: https://doi.org/10.1071/AR9900997.

Streusand, V.J., and Portis, AR. (1987). Rubisco Activase mediates ATP-dependent activation of ribulose bisphosphate carboxylase. *Plant Physiol*., 85, 152–4.

doi: org/10.1104/pp.85.1.152.

Sukenik, A., Bennett, J., and Falkowski, P. (1987). Light-saturated photosynthesis — Limitation by electron transport or carbon fixation? *Biochim. Biophys. Acta - Bioenerg.*, 891, 205–215. doi:10.1016/0005-2728(87)90216-7.

Takahara, K., Kasajima, I., Takahashi, H., Hashida, S.N., Itami, T., Onodera, H., oki, S., Yanagisawa, S., Kawai-Yamada, M. and Uchimiya, H. (2010). metabolome and photochemical analysis of rice plants overexpressing Arabidopsis NAD kinase gene 1. *Plant Physiol*., 152, 1863-1873. doi: 10.1104/pp.110.153098.

Tanaka, Y., and Shiraiwa, T. (2009). Stem growth habit affects leaf morphology and gas exchange traits in soybean. *Ann. Bot.*, 104, 1293–1299. doi:10.1093/aob/mcp240.

Tanaka, Y., Fujii, K., and Shiraiwa, T. (2010). Variability of leaf morphology and stomatal conductance in soybean [*Glycine max* (L.) Merr.] cultivars. *Crop Sci.* 50, 2525–2532. doi:10.2135/cropsci2010.02.0058.

Tanaka, Y., Sugano, S. S., Shimada, T., and Hara-Nishimura, I. (2013). Enhancement of leaf photosynthetic capacity through increased stomatal density in *Arabidopsis*. *New Phytologist*, 198, 757–764. doi: 10.1111/nph.12186.

Tanaka, Y., Taniyoshi, K., Imamura, A., Mukai, R., Sukemura, S., Sakoda, K., and Adachi, S. (2021). MIC-100, a new system for high-throughput phenotyping of instantaneous leaf photosynthetic rate in the field. *Funct. Plant Biol*., 49, 496-504. doi: 10.1071/FP21029.

Taylor, S. H., and Long, S. P. (2017). Slow induction of photosynthesis on the shade to sun transitions in wheat may cost at least 21% of productivity. *Philosophical Transactions of the Royal Society of Lond. Series B, Biol. SciencSci.,* 372, 20160543. doi: 10.1098/rstb.2016.0543.

Tietz, S., Hall, C. C., Cruz, J. A., and Kramer, D. M. (2017). $NPQ_{(T)}$: A chlorophyll fluorescence parameter for rapid estimation and imaging of non-photochemical quenching of excitons in photosystem-II-associated antenna complexes. *Plant Cell and Environ*., 40, 1243–1255. doi: 10.1111/pce.12924.

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceed. of the Nat. Acad. Sci. USA*, 108, 20260–20264. doi: 10.1073/pnas.1116437108/suppl_file/pnas.201116437si.pdf.

Vickery, H.B. (1946). The early years of the Kjeldahl Method to determine nitrogen. *J. Biol. Med*., 18, 473-516.

Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2602049/ [Accessed: June 15, 2022].

Vieira, A. J. D., De Oliveira, D. A., Soares, T. C. B., Schuster, I., Piovesan, N. D., Martínez, C. A Barros, E.G.D. and Moreira, M.A. (2006). Use of the QTL approach

to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. *Braz. J. Plant Physiol*. 18, 281–290.

doi: 10.1590/S1677- 04202006000200004.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. and Van Der Walt, S.J. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, 17, 261-272. doi: 10.1038/s41592-019-0686-2.

Wang, L., Yang, Y., Zhang, S., Che, Z., Yuan, W., and Yu, D. (2020). GWAS reveals two novel loci for photosynthesis-related traits in soybean. *Mol. Genet. Gen*. 295, 705–716. doi:10.1007/s00438-020-01661-1.

Whitney, S. M., Houtz, R. L., and Alonso, H. (2011). Update on understanding and manipulating Rubisco biogenesis: advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco. *Plant Physiol*. 155, 27–35. doi: 10.1104/pp.110.164814.

Wise, R. R., Olson, A. J., Schrader, S. M., and Sharkey, T. D. (2004). Electron transport is the functional limitation of photosynthesis in field-grown Pima cotton plants at high temperature. *Plant Cell Environ.*, 27, 717–724.

doi: 10.1111/j.1365-3040.2004.01171.

Xiao, L., Liu, X., Lu, W., Chen, P., Quan, M., Si, J., Du, Q., and Zhang, D. (2020). Genetic dissection of the gene coexpression network underlying photosynthesis in Populus. *Plant Biotech. J.,* 18, 1015–1026. doi: 10.1111/pbi.13270.

Xue, Z., and Gao, H. (2017). The difference of photosynthetic responses to the cadmium stress between a wild soybean (*Glycine soja*) and a cultivated soybean. *Bull. Environ. Contam. Toxicol.,* 99, 405–410. doi: org/10.1007/s00128-017-2143-1.

Yamori, W., and Shikanai, T. (2016). Physiological functions of cyclic electron transport around photosystem in sustaining photosynthesis and plant growth. *Ann. Rev. Plant Biol.*, 67, 81–106. doi: 10.1146/annurev-arplant-043015-112002.

Yamori, W., Kusumi, K., Iba, K., and Terashima, I. (2020). Increased stomatal conductance induces rapid changes to photosynthetic rate in response to naturally fluctuating light conditions in rice. *Plant Cell and Environ.*, 43, 1230–1240. doi: 10.1111/pce.13725.

Yang., J.J., Li, J., Williams, K., and Buu, A., (2019). An efficient genome-wide association test for multivariate phenotypes based on the fisher combination function. *BMC Bioinformatics*, 17, 1-11. doi: 10.1186/s12859-015-0868-6.

Yang, Y., Wang, L., Che, Z., Wang, R., Cui, R., Xu, H., et al. (2022). Novel 1434 target sites for soybean yield enhancement by photosynthesis. *Plant Physiol*. 1435 268:153580. doi: 10.1016/J.JPLPH.2021.153580

Yano, K., Morinaka, Y., Wang, F., Huang, P., Takehara, S., Hirai, T., Ito, A., Koketsu, E., Kawamura, M., Kotake, K. and Yoshida, S.. (2019). GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proc. Nat. Acad. Sci. USA*, 116, 2162–21267. doi: 10.1073/pnas.1904964116.

Zhang, L., Hu, G., Cheng, Y., and Huang, J. (2008). Heterotrimeric G protein α and β subunits antagonistically modulate stomatal density in *Arabidopsis thaliana*. *Dev. Biol.*, *324*, 68–75. doi: 10.1016/j.ydbio.2008.09.008.

Zhang, L., Wei, Q., Wu, W., Cheng, Y., Hu, G., Hu, F., Sun, Y., Zhu, Y., Sakamoto, W. and Huang, J.. (2009). Activation of the heterotrimeric G protein a-subunit GPA1 suppresses the ftsh-mediated inhibition of chloroplast development in Arabidopsis. *Plant J.*, 58, 1041-1053. doi: 10.1111/j.1365-313X.2009.03843.x.

Zhu, G., Kurek, and I., Liu, L. (2010). Engineering photosynthetic enzymes involved in CO_2-assimilation by gene shuffling in the chloroplast. Advances in photosynthesis and respiration. *Springer, Dordecht*, 31, 307–322.

doi: org/10.1007/978-90-481-8531-3_20.

Zhu, X.G., Long, S.P., and Ort, D.R. (2008). What is the maximum efficiency with which photosynthesis can convert solar energy into biomass?. *Curr. Opin. Biotechnol*. 19, 153–159. doi: 10.1016/j.copbio.2008.02.004.

Zhu, X.G., Long, S.P., Ort D.R. (2010). Improving photosynthetic efficiency for greater yield. *Ann. Rev. Plant Biol*., 61, 235–261.

doi: org/10.1146/annurev-arplant-042809-112206.

List of Abbreviations

Symbol/*Abbr.*	Description and Unit
A_{max}	Maximum Photosynthesis (μmol $m^{-2}s^{-1}$)
ABA	Abscisic Acid
ANOVA	Analysis of Variance
A_{sat}	Light-Saturated Photosynthesis (μmol $m^{-2}s^{-1}$)
C_3	C_3 Plants
C_4	C_4 Plants
Chl a	Chlorophyll a
Chl a + b	Chlorophyll a and b
Chl b	Chlorophyll b
C_i	Intercellular CO_2 Concentration (μmol mol^{-1})
CO_2	Carbon Dioxide
CWR	Crop Wild Relatives
DAP	Days After Planting
DMF	N, N-Dimethylformamide
DNA	Deoxyribose Nucleic Acid
ETR	Electron Transfer Rate (μmol m^{-2}-s^{-1})
FACE	Free Air CO_2 Enrichment
Fv'/Fm'	Light Use Efficiency of Photosystem II
g_m	Mesophyll Conductance (mol m^{-2}-s^{-1})
GmJMC	Japanese Soybean Mini Core Collection
GPA1	G Protein Alpha Subunit 1
g_s	Stomatal Conductance (mol m^{-2}-s^{-1})
GSD	Green Stem Disorder
GWAS	Genome Wide Association Studies
HI	Harvest Index
ictB	Cyanobacterial Putative-Inorganic Carbon Transporter B
J_{max}	Maximum Rate of Electron Transfer
JMC	Japanese Soybean Mini Core Collection
K_2O	Potassium Oxide
LMA	Leaf Mass Area (g m^{-2})
N	Nitrogen (fertilizer)
NAM	Nest Association Map (the U.S soybean collection)
N_{cont}	Leaf Nitrogen Content (g m^{-2})
$NPQ_{(t)}$	Non-Photochemical Quenching
P_2O_5	Phosphorus Pentaoxide (Fertilizer)
PAR	Photosynthetically Active Radiation (μmol $m^{-2}s^{-1}$)
PhiPSII	Quantum efficiency of Photosystem II

Symbol/*Abbr.*	Description and Unit
PNUE	Photosynthetic Nitrogen Use Efficiency
PSII	Photosystem II
R_1	Beginning Bloom
R_2	Full Bloom
R_3	Beginning Pod
R_4	Full Pod
R_5	Beginning Seed
R_6	Full Seed
R_7	Beginning Maturity
R_8	Full Maturity
RCBD	Randomized Complete Block Design
RH	Relative Humidity (%)
RuBP	Ribulose 1,5-bisphosphate
Rubsico	Ribulose 1,5-bisphosphate Carboxylation/Oxygenation
RUE	Radiation Use Efficiency (g MJ^{-1})
$S_{Density}$	Stomatal Density (n mm^{-2})
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SG	Stay Green
SLA	Specific Leaf Area (m^2 g^{-1})
S_{Length}	The length of Stomata (n mm^{-2})
SLW	Specific Leaf Weight (g m^{-2})
SNP	Single Nucleotide Polymorphism
S_t	Cumulative Solar Radiation (MJ m^{-2})
SUMP	Suzuki Universal Method of Printing
S_{Width}	The Width of Stomata (n mm^{-2})
TDW	Total Dry Weight (g m^{-2})
t-SNE	T Distribution Stochastic Neighboring Embedding
V_{cmax}	Maximum Rate of Carboxylation
WGR	Whole-Genome Regression
Tleaf	Leaf Temperature (^oC)
E	Transpiration (mol m^{-2}-s^{-1})
WUE	Water Use Efficiency (g MJ^{-1} day^{-1})
Y_p	Yield Potential
$\mathcal{E}_c$	the Light Conversion Efficiency
$\mathcal{E}_i$	the Light Interception Efficiency
η	Harvest Index

List of JMC accessions

ID	Cultivar	Status	Origin	Mother accession*1
GmJMC002*2	WASE KURO DAIZU	Landrace	JAPAN (KUMAMOTO)	29619
GmJMC003*2	NATSU KURAKAKE	Landrace	JAPAN (KUMAMOTO)	29494
GmJMC004	KITAJIRO	Landrace	JAPAN (CHIBA)	28880
GmJMC005	WASEOUSODE (SHIKAOI ITOH)	Landrace	JAPAN (HOKKAIDO)	53347
GmJMC007*2	TOKACHI NAGAHA	Breeders line	JAPAN (HOKKAIDO)	
GmJMC008	KANAGAWA WASE	Landrace	JAPAN (KUMAMOTO)	29621
GmJMC009	SHIZUNAIDAIZU	Breeders line	JAPAN (HOKKAIDO)	53267
GmJMC013	CHIZUKA IBARAKI 1	Breeders line	JAPAN (IBARAKI)	29149
GmJMC016	JUKKOKU	Landrace	JAPAN (SAITAMA)	28507
GmJMC021	OOYACHI 2	Breeders line	JAPAN (HOKKAIDO)	53264
GmJMC023	KUROGOYOU	Landrace	JAPAN (FUKUSHIMA)	28214
GmJMC025	ENREI	Breeders line	JAPAN (NAGANO)	
GmJMC026	ONI HADAKA	Breeders line	JAPAN (TOCHIGI)	37744
GmJMC028	KOITO	Landrace	JAPAN (CHIBA)	
GmJMC030	KURODAIZU(AO HIGUU CHUU)	Landrace	JAPAN (OKINAWA)	30135
GmJMC031	SHIRO MITSU MAME	Landrace	JAPAN (NAGANO)	28600
GmJMC032	NATTOU KOTSUBU	Breeders line	JAPAN (IBARAKI)	29161
GmJMC033	BANSEI HIKARIKURO	Landrace	JAPAN (HOKKAIDO)	27708
GmJMC034	MIYAGI SHIROME	Breeders line	JAPAN (MIYAGI)	27886
GmJMC037	YAKUMO MEAKA	Landrace	JAPAN (HOKKAIDO)	53292
GmJMC039	NATTOUMAME	Landrace	JAPAN (NAGANO)	73172
GmJMC040	KOIBUCHIMURA ZAIRAI	Landrace	JAPAN (IBARAKI)	28922
GmJMC041*2	DATE CHA MAME	Landrace	JAPAN (MIYAGI)	28869
GmJMC043	TAKIYA	Landrace	JAPAN (YAMAGATA)	27922
GmJMC044	SHAKKIN NASHI	Landrace	JAPAN (GUNMA)	28474
GmJMC047	AKITA ANI	Landrace	JAPAN (YAMAGATA)	27920
GmJMC049	HIKU ANDA	Landrace	JAPAN (OKINAWA)	35409
GmJMC050	FUKUI SHIRO	Landrace	JAPAN (FUKUI)	27948
GmJMC051	KURODAIZU(GEIHOKU)	Unknown	JAPAN (HIROSHIMA)	227175
GmJMC052	KISAYA(NATSU)	Landrace	JAPAN (KAGOSHIMA)	29500
GmJMC053	ABURA MAME	Landrace	JAPAN (FUKUSHIMA)	27900
GmJMC054	ZAIRAI 51-6	Landrace	JAPAN (AICHI)	29242
GmJMC055	SAKURAMAME	Landrace	JAPAN (YAMAGATA)	87833

ID	Cultivar	Status	Origin	Mother accession*1
GmJMC056	TAMAHOMARE	Breeders line	JAPAN (NAGANO)	29184
GmJMC057	SHAKUJOU MAME	Landrace	JAPAN (Unspecified)	
GmJMC058	YAHAGI	Landrace	JAPAN (AICHI)	29210
GmJMC059	SOKOSHIN	Landrace	JAPAN (NIIGATA)	28252
GmJMC060	SHIMO HISAKATA DAIZU	Landrace	JAPAN (NAGANO)	28947
GmJMC061	KOMAME	Landrace	JAPAN (GUNMA)	73030
GmJMC062	AZEMAME	Landrace	JAPAN (TOCHIGI)	67658
GmJMC063	AOBAKO	Landrace	JAPAN (YAMAGATA)	76512
GmJMC064	MEGURO 1	Landrace	JAPAN (AOMORI)	28049
GmJMC065	OOJIRO	Landrace	JAPAN (GUNMA)	27974
GmJMC067	HOUJAKU	Landrace	JAPAN (NAGANO)	28770
GmJMC068	ZAIRAI 51-2	Landrace	JAPAN (AICHI)	29253
GmJMC069	CHADAIZU	Landrace	JAPAN (MIYAGI)	27890
GmJMC076*2	IHHON SANGOU	Landrace	JAPAN (IBARAKI)	28914
GmJMC077	HITORIMUSUME	Landrace	JAPAN (YAMAGATA)	76514
GmJMC078	AKASAYA	Landrace	JAPAN (NARA)	29346
GmJMC079	KURUMIMAME	Landrace	JAPAN (MIYAGI)	85549
GmJMC080	HIME DAIZU	Landrace	JAPAN (GIFU)	29213
GmJMC081	AKUDEN SHIRAZU	Landrace	JAPAN (NAGANO)	28320
GmJMC082	AOAKIMAME	Landrace	JAPAN (HYOGO)	67662
GmJMC085	DAIZU	Landrace	JAPAN (WAKAYAMA)	110355
GmJMC088	CHUU TEPPOU	Landrace	JAPAN (GIFU)	29207
GmJMC090	DADACHAMAME	Landrace	JAPAN (YAMAGATA)	87838
GmJMC091	KUROTOME	Landrace	JAPAN (MIYAGI)	27891
GmJMC092	KUROHIRA	Landrace	JAPAN (IWATE)	27863
GmJMC093	ZAI 52-12	Landrace	JAPAN (CHIBA)	35338
GmJMC095	AKASAYA	Unknown	JAPAN (ISHIKAWA)	28264
GmJMC096	NAKAHATA ZAIRAI	Landrace	JAPAN (SHIZUOKA)	29217
GmJMC097	IPPON SUZUNARI	Landrace	JAPAN (AICHI)	29208
GmJMC098	AKA DAIZU	Landrace	JAPAN (TOKUSHIMA)	29389
GmJMC099	AMAGI ZAIRAI 90D	Landrace	JAPAN (FUKUOKA)	76302
GmJMC100	KUROMAME	Landrace	JAPAN (SAITAMA)	73149
GmJMC101	COL/EHIME/1983/UTSUNOMIYA 22	Landrace	JAPAN (EHIME)	29408
GmJMC102	KURAKAKE	Landrace	JAPAN (NIIGATA)	208667
GmJMC104	DAIZU(SHIRO)	Landrace	JAPAN (NARA)	90917
GmJMC105	MAETSUE ZAIRAI 90B	Landrace	JAPAN (OITA)	76321
GmJMC106*2	HIMESHIRAZU	Breeders line	JAPAN (CHIBA)	67990
GmJMC110	COL/TANBA/1989/ODAGAKI 2	Landrace	JAPAN (HYOGO)	
GmJMC111	AMAGI ZAIRAI 90A	Landrace	JAPAN (FUKUOKA)	76310
GmJMC112	FUKUYUTAKA	Breeders line	JAPAN (KUMAMOTO)	

ID	Cultivar	Status	Origin	Mother accession*1
GmJMC114	COL/EHIME/1-2	Unknown	JAPAN (EHIME)	227399
GmJMC116	SHIRATAMA	Landrace	JAPAN (IWATE)	27858
GmJMC117	AKISENGOKU	Breeders line	JAPAN (KUMAMOTO)	
GmJMC121	KOSA MAME	Landrace	JAPAN (TOCHIGI)	28489
GmJMC126	KOKUBU 7	Landrace	JAPAN (HYOGO)	29363
GmJMC128	GIN DAIZU	Landrace	JAPAN (OKAYAMA)	
GmJMC130	HITASHIMAME	Landrace	JAPAN (YAMAGATA)	76560
GmJMC131	COL/EHIME/1983/UTSUNOMIYA 28	Landrace	JAPAN (EHIME)	29413
GmJMC133	DAIZU	Landrace	JAPAN (KOCHI)	81622
GmJMC137	COL/EHIME/1983/UTSUNOMIYA 37	Landrace	JAPAN (EHIME)	29427
GmJMC139	BUNSEI	Landrace	JAPAN (KUMAMOTO)	29542
GmJMC145	SHIMOTSURA	Landrace	JAPAN (KUMAMOTO)	29534
GmJMC149	MOCHI-DAIZU	Landrace	JAPAN (MIE)	70093
GmJMC158	KUMAJI 1	Landrace	JAPAN (KUMAMOTO)	28377
GmJMC161	ITSUKI ZAIRAI 83H	Landrace	JAPAN (KUMAMOTO)	29715
GmJMC167	NANKAN ZAIRAI 83	Landrace	JAPAN (KUMAMOTO)	29696
GmJMC172*2	TSURUSENGOKU	Breeders line	JAPAN (CHIBA)	67989
GmJMC177	HAI MAME	Landrace	JAPAN (YAMANASHI)	28433
GmJMC179	SAGA ZAIRAI	Landrace	JAPAN (SAGA)	29292
GmJMC180	KOMUTA	Landrace	JAPAN (KUMAMOTO)	29493
GmJMC184	BAN KURO DAIZU	Landrace	JAPAN (KUMAMOTO)	29533

*1 The JMC accessions may not be the same as their parent accessions genetically because they were selected as pure lines.

*2 Most seeds of GmJMC172 have a hard seed coat, and scarification of the seed coat is required before sowing. Scarification is recommended for GmJMC002, 003, 007, 041, 076, and 106 containing hard seeds [Retrieved from NARO Genebank – Japanese Soybean Core Collection (affrc.go.jp)

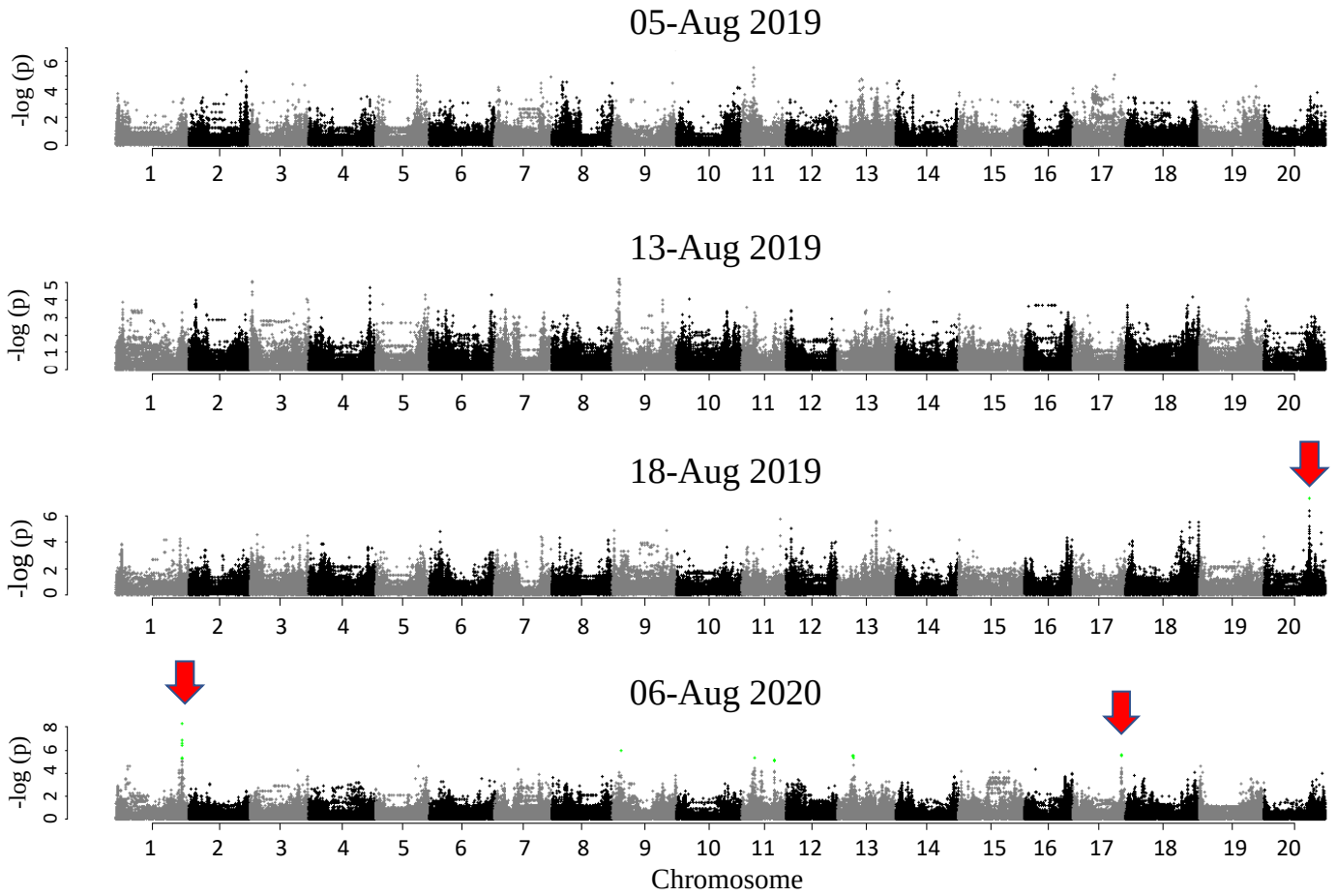
Appendix 1

List of candidate genes in 500-kbps region spanning both sides of the most significant SNP. Refer to the paper for the functions of the genes

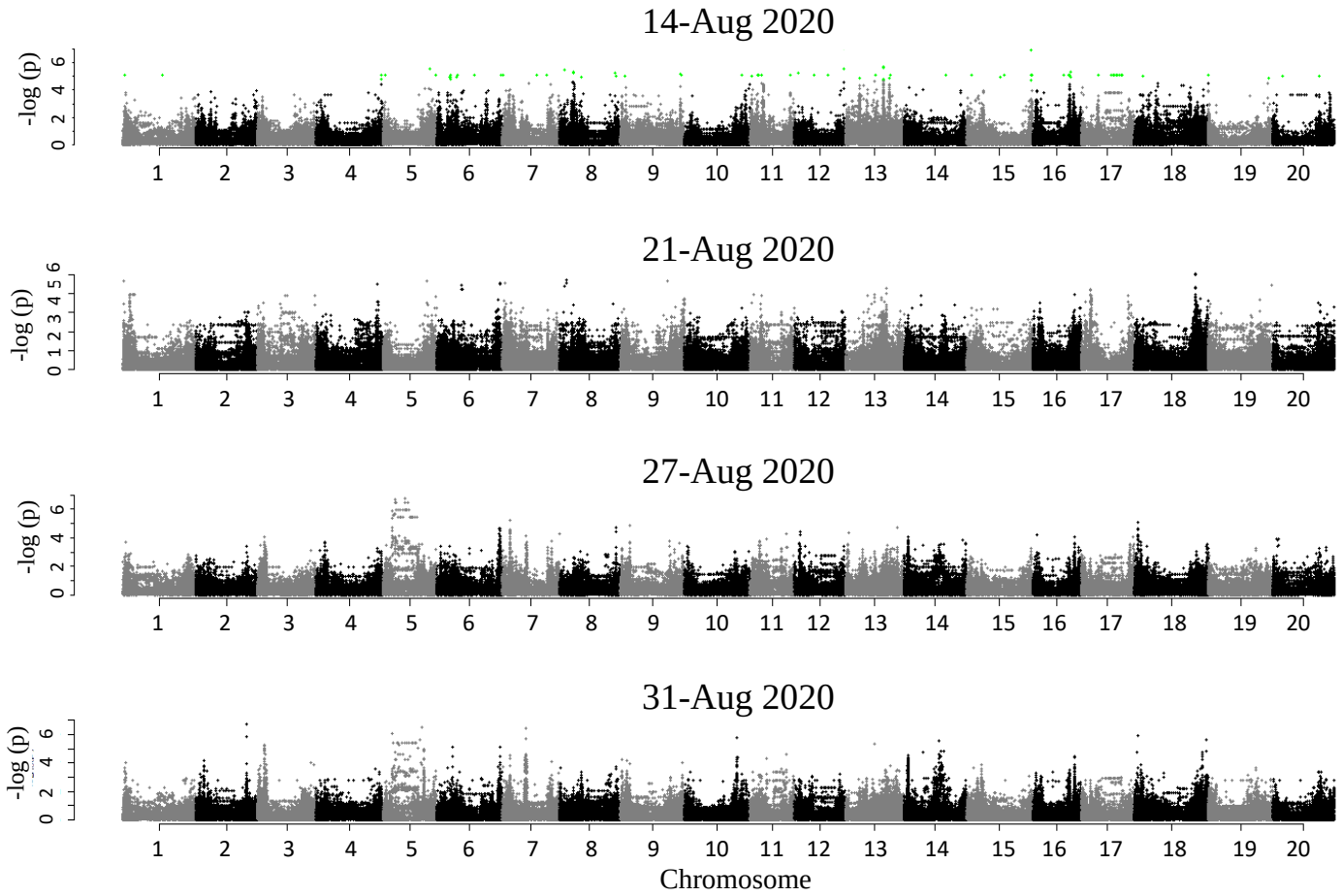
Remarks	Gene	Chr	Start_postion (bp)	End_postion (bp)	TAIR annotation number	Gene name
	Glyma.17G225800	Chr17	38,002,427	38,010,272	AT4G33210	SLOMO
	Glyma.17G225900	Chr17	38,011,908	38,012,862	AT5G21910	
	Glyma.17G226000	Chr17	38,013,862	38,014,832		
RT-PCR	Glyma.17G226100	Chr17	38,032,261	38,040,867	AT2G26330	ER,QRP1
	Glyma.17G226200	Chr17	38,068,206	38,069,105	AT2G28740	HIS4
	Glyma.17G226300	Chr17	38,070,959	38,075,701	AT4G33200	ATXI-I,XI-15,XI-I
	Glyma.17G226400	Chr17	38,075,702	38,090,606	AT4G33200	ATXI-I,XI-15,XI-I
	Glyma.17G226500	Chr17	38,091,013	38,093,234	AT5G21940	
	Glyma.17G226600	Chr17	38,105,929	38,111,302	AT2G26310	
RT-PCR	Glyma.17G226700	Chr17	38,112,713	38,121,429	AT2G26300	ATGPA1,GP ALPHA 1,GPA1
	Glyma.17G226800	Chr17	38,130,275	38,136,095	AT2G26200	
RT-PCR	Glyma.17G226900	Chr17	38,161,706	38,164,919	AT2G26080	AtGLDP2,GLDP2
	Glyma.17G227000	Chr17	38,167,203	38,167,409	AT5G11970	
	Glyma.17G227100	Chr17	38,175,388	38,181,189	AT5G11960	
	Glyma.17G227200	Chr17	38,184,820	38,193,145	AT5G11950	
	Glyma.17G227300	Chr17	38,198,671	38,199,394		
	Glyma.17G227400	Chr17	38,208,471	38,211,994	AT2G26180	IQD6
	Glyma.17G227500	Chr17	38,210,720	38,217,231	AT2G26170	CYP711A1,MAX1
	Glyma.17G227600	Chr17	38,237,933	38,241,566	AT2G26150	ATHSFA2,HSFA2
SNP position with most significant peak			38,255,561			
	Glyma.17G227700	Chr17	38,259,023	38,262,073	AT4G33040	
	Glyma.17G227800	Chr17	38,275,485	38,278,808	AT1G55120	AtcwINV3,ATFRUCT5,FRUCT5
	Glyma.17G227900	Chr17	38,284,476	38,287,851	AT5G11920	AtcwINV6,cwINV6
	Glyma.17G228000	Chr17	38,290,133	38,296,500	AT3G50700	AtIDD2,IDD2
	Glyma.17G228100	Chr17	38,305,771	38,313,230	AT2G26140	ftsh4
	Glyma.17G228200	Chr17	38,314,962	38,315,760		
	Glyma.17G228300	Chr17	38,316,355	38,326,030	AT5G11910	
	Glyma.17G228400	Chr17	38,332,780	38,337,858	AT2G46210	
	Glyma.17G228500	Chr17	38,343,564	38,348,355	AT5G50960	ATNBP35,NBP35
	Glyma.17G228600	Chr17	38,349,687	38,353,286	AT1G60960	ATIRT3,IRT3
	Glyma.17G228700	Chr17	38,364,416	38,367,780	AT4G26150	CGA1,GATA22
	Glyma.17G228800	Chr17	38,383,184	38,389,776	AT2G26080	AtGLDP2,GLDP2
	Glyma.17G228900	Chr17	38,398,341	38,400,442	AT5G60030	
	Glyma.17G229000	Chr17	38,405,283	38,412,601	AT3G25690	CHUP1
	Glyma.17G229100	Chr17	38,419,116	38,429,069	AT5G11850	
	Glyma.17G229200	Chr17	38,425,777	38,427,208		
	Glyma.17G229300	Chr17	38,434,624	38,438,642	AT2G26070	RTE1
	Glyma.17G229400	Chr17	38,439,762	38,442,487	AT5G11810	
	Glyma.17G229500	Chr17	38,443,771	38,448,625	AT4G33000	ATCBL10,CBL10,SCABP8
	Glyma.17G229600	Chr17	38,450,991	38,457,056	AT5G26040	HDA2
	Glyma.17G229700	Chr17	38,460,625	38,472,586	AT5G11800	ATKEA6,KEA6
	Glyma.17G229800	Chr17	38,475,630	38,477,885	AT5G47390	
	Glyma.17G229900	Chr17	38,494,068	38,495,137	AT2G26040	PYL2,RCAR14

Appendix 2

Manhattan plots show the association of SNPs with each measurement. Red vectors and green dots show genomic regions where the association is significant.



Manhattan plots show the association of SNPs with each measurement. Green dots show genomic regions where the association is significant.



Acknowledgments

All praises go to Allah the almighty, the most merciful, and the most companionate. Everything belongs to him, and every success is due to his blessings.

This thesis was written and completed at the Laboratory of Crop Science, Graduate School of Agriculture, Kyoto University. This goal would not have been achieved without the following people and organizations' direct/indirect support and contributions.

Foremost, it is my great pleasure to express my hearty thanks and warmest gratitude to my advisors and mentors, His Excellency Assistant Professor (Dr.) Yu Tanaka and His Excellency Professor (Dr.) Shiraiwa Tatsuhiko for their endless support, patience, approachability, guidance, untiring efforts, sympathy, and encouragement throughout my studies. I am also indebted to the advice, patience, and support of Dr. Akito Kaga for conducting genomic analysis, corresponding to my research and reviewing my second paper.

I extend my gratitude to His Excellency Associate Professor (Dr.) Tomoyuki Tanaka for invaluable guidance, support, precious comments, suggestions, and advice.

I am also grateful to Dr. Yamatani Hiroshi for conducting RNA expression analysis, Dr. Sakoda Kazuma for his input into my research and for instructing me on the procedure of Rubisco analysis., and Dr. Idowu for his kindness in teaching me the know-how-to of leaf nitrogen analysis.

My deepest, heartfelt appreciation goes to the sweetest people in the world, my beloved parents, for going through many hardships to facilitate my way towards who I am today. Their prayers and moral support energized me in every movement. Also, I owe a substantial debt to my siblings for their inspiration and motivations.

I would particularly like to express my most incredible gratitude to the Japan Ministry of Education, Sport, Culture, Science, and Technology (MEXT), Ministry of Agriculture, Forestry, and Fisheries (Smart-breeding system for Innovative Agriculture, BAC1003), Japan International Cooperation Agency (JICA) and Japan International Cooperation Center (JICE) for their valuable and continuous financial support and daily take-care. This achievement would not be possible without their constant hold-up.

I will always be grateful to the incredible advisors, the most helpful lab mates, my family and friends, and to your creative suggestions, precious comments, and kindest attitude; I will always be indebted.

List of Publications

1. Chapter two

Shamim, M.J., Tanaka, Y., Sakoda, K., Shiraiwa, T., Nelson, R.L., 2020. Physiological analysis of leaf photosynthesis of backcross-derived progenies from soybean (*Glycine max* (L.) Merrill) and *G. tomentella* Hayata. *Plant Prod. Sci*. 24, 109–117. Doi:10.1080/1343943X.2020.1807369.

2. Chapter Three and Chapter four

Shamim. M.J., Kaga, A., Tanaka. Y., Hiroshi Y., Shiraiwa, T. (2002). Analysis of physiological variations and genetic architecture for photosynthetic capacity of Japanese soybean germplasm. *Front. Plant Sci*. 13, 910527. doi: 10.3389/fpls.2022.910527.