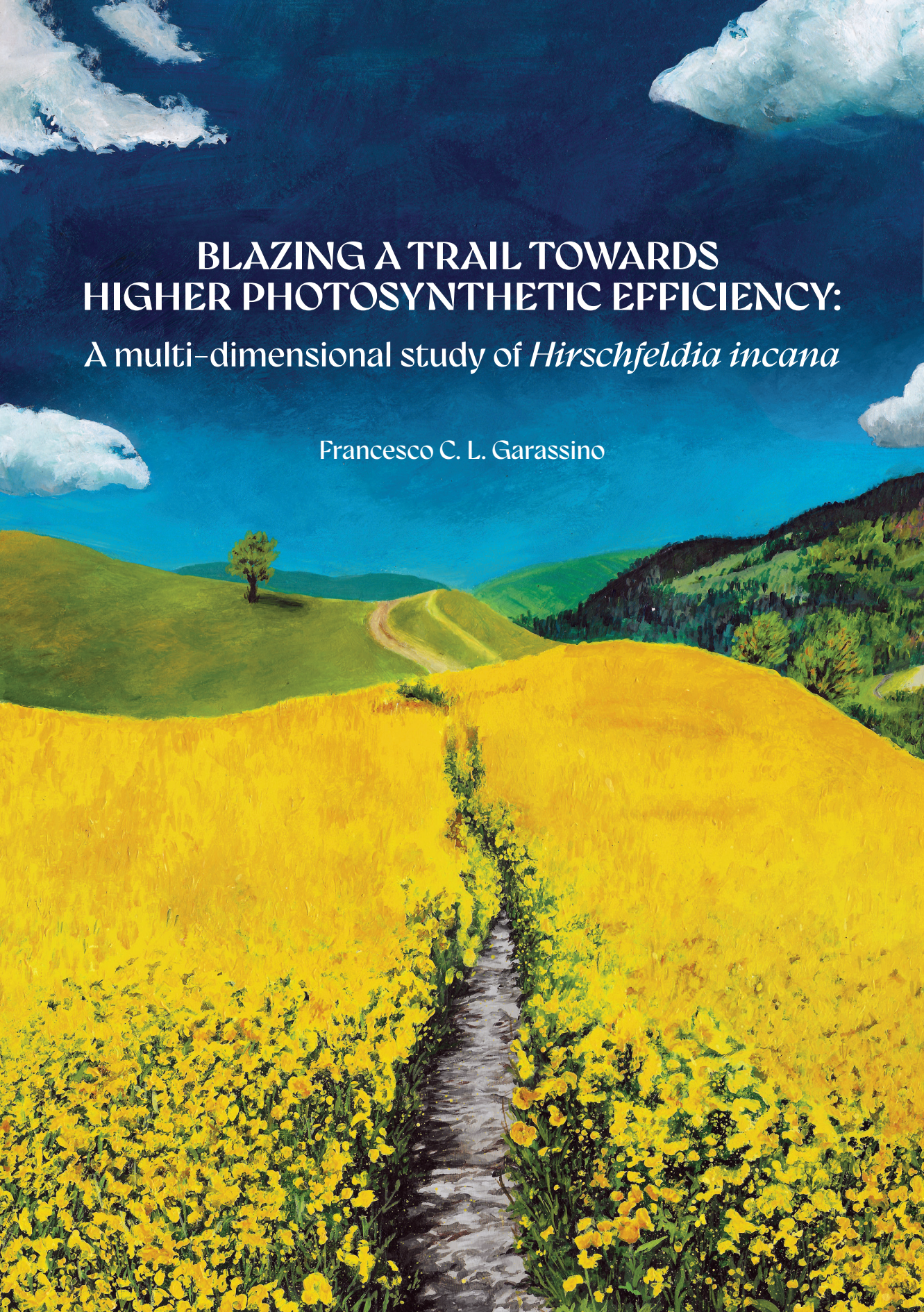

BLAZING A TRAIL TOWARDS HIGHER PHOTOSYNTHETIC EFFICIENCY:

A multi-dimensional study of *Hirschfeldia incana*

Francesco C. L. Garassino

Propositions

1. A less crop-centric perspective in research will allow to achieve higher crop yields.
   (this thesis)
2. Our understanding of photosynthetic light-use efficiency at high irradiance will benefit from a less leaf-centric perspective.
   (this thesis)
3. Brandolini's law applies to current academic publishing.
4. Generative AI challenges traditional academic propositions as demonstrations of independent thought.
5. Only hands-on experience in agricultural production yields valuable opinions on it.
6. Science communication must celebrate the complexity of problems rather than oversimplifying it.

Propositions belonging to the thesis, entitled

"Blazing a trail towards higher photosynthetic efficiency: a multi-dimensional study of *Hirschfeldia incana*"

Francesco C. L. Garassino
Wageningen, 11/12/2023

Blazing a trail towards higher photosynthetic efficiency: A multi-dimensional study of *Hirschfeldia incana*

Francesco C. L. Garassino

Thesis committee

Promotor
Prof. Dr M. G. M. Aarts
Personal chair at the Laboratory of Genetics
Wageningen University & Research

Co-promotors
Dr S. Smit
Assistant professor, Bioinformatics Group
Wageningen University & Research

Dr J. Harbinson
Associate professor, Laboratory of Biophysics
Wageningen University & Research

Other members
Prof. Dr C. S. Testerink, Wageningen University & Research
Prof. Dr A. Weber, Heinrich-Heine-Universität, Düsseldorf, Germany
Prof. Dr D. M. Kramer, Michigan State University, East Lansing, U.S.A.
Dr L. Berke, Genetwister Technologies B.V., Wageningen

This research was conducted under the auspices of the Graduate School Experimental Plant Sciences.

Blazing a trail towards higher photosynthetic efficiency: A multi-dimensional study of *Hirschfeldia incana*

Francesco C. L. Garassino

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A. P. J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 11 December 2023
at 4 p.m. in the Omnia Auditorium.

Francesco C. L. Garassino
Blazing a trail towards higher photosynthetic efficiency: A multi-dimensional study of *Hirschfeldia incana*,
192 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2023).
With references, with summary in English.

ISBN 978-94-6447-916-4
DOI 10.18174/639737

Voglio però ricordarti com'eri
Pensare che ancora vivi
Voglio pensare che ancora mi ascolti
E che come allora sorridi

Francesco Guccini, Canzone per un'amica

In memory of Paolo Borra (1994-2013)

CONTENTS

CHAPTER 1

General introduction

L'unica gioia al mondo è cominciare. È bello vivere perché vivere è cominciare, sempre, ad ogni istante. Quando manca questo senso: -prigione, malattia, abitudine, stupidità-, si vorrebbe morire.

Cesare Pavese, Il mestiere di vivere: Diario 1935-1950

Quoting Sir Patrick Geddes, *"The leaf is the chief product and phenomenon of Life: this is a green world, with animals comparatively few and small, and all dependent upon the leaves. By leaves we live"* (Macdonald, 2020). These words, which resonated over a century ago in an auditorium at Dundee University, beautifully metaphorize how photosynthesis is the biochemical process directly or indirectly sustaining most life on Earth. Photosynthesis indeed provides most organisms with the energy allowing their survival and growth, and the oxygen required to utilize it. However, not all plants are equal when it comes to their photosynthesis, with some performing it in much more efficient ways. What exactly determines the photosynthetic efficiency of plants, i.e. which genetic, biochemical, and physiological factors are important, remains to be described in detail. This thesis focuses on a plant with high photosynthetic light-use efficiency, the little-known *Hirschfeldia incana*, and presents findings on the genetic background of its performance. The work presented in this thesis pushes the boundaries of our understanding of photosynthesis a little further, contributing to the growing body of knowledge on the most important biochemical process on Earth.

1.1 An historical perspective on photosynthesis

The long history of photosynthesis research is made of countless contributions by many researchers over the world that have shaped our understanding of this fundamental process. A number of excellent publications have been dedicated to this fascinating topic and present a very detailed reconstruction of many events in the history of photosynthesis research (e.g., Govindjee *et al.* (2005); Nickelsen (2015a)). While not aiming at a comparable level of detail, this section summarizes the key events leading to the discovery and understanding of photosynthesis.

The history of photosynthesis starts at the very dawn of plant science, with one of the earliest recorded experiments in the field. In the 17th century Jan Baptist van Helmont conducted the famous 'willow tree' experiment, already proposed almost 200 years earlier by Nicholas of Cusa. Searching for the the matter required for tree growth, he followed the growth of a willow seedling in well-weighed soil over a period of five years, providing the growing tree nothing but rain water. Noticing that after five years the weight of the tree had substantially increased by around 75 kg while the weight of the soil had only decreased by grams, van Helmont concluded that the water used for irrigation, instead of the soil, had determined the growth of the willow tree (Van Helmont, 1652; Krikorian and Steward, 1968). In the following decades, Edme Mariotte and Stephen Hales proposed that plants and their leaves obtain nourishment from the air (Pennazio, 2011). Hales also highlighted the roles of leaves in transpiration and gas exchange (Hales, 1727; Loomis, 1960; Hill, 2012), and was the first to hypothesize a role for light in plant growth. More attention was drawn to air and gaseous exchange by Joseph Priestley's experiments. Towards the end of the 18th century, he observed that plants have the ability to "restore" air that had been made unfit for animal life by the burning of candles, and theorized about plants' ability to restore *phlogiston*, the "burning principle", thus hinting at oxygen formation

by plants (Priestley, 1776; Geerdt, 2007). Around the same time, Jan Ingen-Housz was the first to robustly link the phlogiston production by plants to solar radiation. Albeit mostly neglected by the scientific community, Ingen-Housz was therefore the first to study and report on the process of photosynthesis (Ingen-Housz, 1779; Gest, 2000; Geerdt, 2007; Magiels, 2010). Notably, he also observed plant respiration in the dark and abandoned the alchemy-bound phlogiston paradigm for that of modern chemistry proposed by Antoine de Lavoisier in 1789 (de Lavoisier and Cuchet, 1789; Gest, 2000).

The 18th and 19th centuries witnessed further advancements in photosynthesis research. Jean Senebier's observations in 1782 demonstrated that plants absorb carbon dioxide from the air, which contributed to the gradual shift of the scientific community from the phlogiston theory to oxygen chemistry (Senebier, 1782; Hill, 2012). Nicholas-Théodore de Saussure delivered in 1804 a conclusive demonstration that water is directly involved in photosynthesis by showing that growing plants gain more in weight than can be explained by the assimilation of the carbon in the absorbed carbon dioxide (de Saussure, 1804). Furthermore, de Saussure is credited with developing a conceptual framework for the source and supply pathways of every major elemental component of plants, thus largely advancing the understanding of plant nutrition and growth (Hill, 2012). Julius Robert von Mayer, in 1845, described the conversion of light energy into chemical energy by plants by reporting that plants converted "light power" into "chemical power" (von Mayer, 1845; Devlin, 1969). Roughly forty years later, in 1882, Theodor Wilhelm Engelmann's work on a filamentous alga and aerotactic bacteria confirmed that the conversion of light energy to chemical energy occurs in the chloroplasts (Drews, 2005; Hintz, 2021). A few years passed until Charles Barnes proposed the term "photosyntax", or alternatively "photosynthesis", to describe the formation of complex carbon compounds under the influence of light (Barnes, 1893).

In the early 20th century, significant breakthroughs improved our understanding of photosynthesis. Frederick Frost Blackman introduced the concept of limitations in photosynthesis in 1905, based on his findings on the response of photosynthesis to light and temperature (Blackman, 1905). Richard Willstatter and his co-authors' studies on the structure and chemistry of chlorophyll culminated in 1915 with a systematic overview on chlorophylls (Willstätter, 1915) and the award of the Nobel Prize for Chemistry (Nobel Lectures in Chemistry, 1966). Following their experiments with intermittent light, Robert Emerson and William Arnold proposed in 1931 the separation of photosynthesis into light-dependent and light-independent reactions (Emerson and Arnold, 1932). Robert Hill's discovery in 1937 that the "light phase" of photosynthesis can operate independently from the "dark phase" is credited with the start of the "modularization" of photosynthesis research (Hill, 1937; Nickelsen, 2015b,c). On the side of the light reactions, Edward D. McAlister and Jack Myers revealed in 1940 an inverse relationship between chlorophyll fluorescence emission and CO_2 uptake (McAlister and Myers, 1940), while in 1941 Samuel Ruben and colleagues used heavy oxygen (^{18}O) to demonstrate that the oxygen generated during photosynthetic reactions originates from water (Ruben *et al.*, 1941). Furthermore, Daniel Arnon and co-authors in 1954 identified ATP as the product of light-dependent pho-

tosynthetic phosphorylation(Arnon *et al.*, 1954). On the side of the dark reactions, in 1956 Melvin Calvin, Andrew Benson, and James Bassham employed radioactively labeled $^{14}CO_2$ in 1956 to elucidate the pathway of carbon assimilation in photosynthesis (Bassham *et al.*, 1954), leading to Calvin's Nobel Prize-winning work (Calvin, 1962). Later years saw further contributions to the elucidation of light reactions, as Robert Emerson and co-workers published key articles in the late 1950s indicating the presence of two separate photochemical systems in photosynthesis (Emerson *et al.*, 1957; Emerson and Chalmers, 1958; Emerson and Rabinowitch, 1960). Following this, Robin Hill and Fay Bendall proposed the theoretical "Z scheme" model for the photosynthetic light reactions in 1960, providing a framework for understanding the coordination of two photosystems operating in tandem (Hill and Bendall, 1960; Govindjee *et al.*, 2017). In 1966, the alternative photosynthetic pathway of C_4 plants was described for the first time by Marshall D. Hatch and Charles R. Slack (Slack and Hatch, 1967).

Photosynthesis research has been developing from the 1960s to the present date at such a high rate that it would be impossible to cite all key developments in a concise way (as an example, see the 36-fold increase in published field measurements of photosynthesis in 1994 as compared to 1984 reported in Long *et al.* (1996)). Nevertheless, a few fundamental developments must be cited in order to deliver an informative overview of the foundations of modern photosynthesis research. Gas-exchange measurements in photosynthesis gained prominence in the 1970s, with the book published in 1971 by Zdeněk Šesták, Jiri Čatský and Paul G. Jarvis, laying the foundations for the widespread adoption of gas-exchange methods (Šesták *et al.*, 1971). The issue of the limited portability of gas-exchange instrumentation (Koch *et al.*, 1971) started being tackled with the development of the first truly portable gas analysis system in 1972 (Schulze, 1972). Further major progresses in gas-exchange measurements of photosynthesis were achieved in 1980, when Graham D. Farquhar, Susanne von Caemmerer, and Joe A. Berry integrated various aspects of carbon assimilation into the widely known biochemical model for parameterizing gas exchange in leaves (Farquhar *et al.*, 1980). Bernard Genty and colleagues, in 1989, linked the quantum yield of photosynthetic electron transport with chlorophyll fluorescence quenching, opening avenues for measuring photosynthetic light-use efficiency with measurements of chlorophyll fluorescence (Genty *et al.*, 1989). These pivotal discoveries and developments in the measurement of photosynthesis have shaped the research field, providing a foundation for further exploration of this vital biological process. Over the past sixty years, photosynthesis has been studied across scales, from the microscopic to the planetary ones, leading to the vast knowledge base available nowadays (Janssen *et al.*, 2014; Ryu *et al.*, 2019).

1.2 Photosynthesis, crop yield, and their improvement

The works summarized in the previous section, along with countless others accumulating over the hundreds-years history of photosynthesis research, have contributed to our broad understanding of photosynthesis biochemistry and its ecophysiology.

This is apparent from the hundreds of pages dedicated to photosynthesis across plant physiology textbooks (Taiz *et al.*, 2015; Willey, 2016; Lambers and Oliveira, 2019), as well as the vast production of whole books dedicated to photosynthesis itself (Pessarakli, 2016; Becklin, 2021; Blankenship, 2021; Ruban *et al.*, 2022; Sharwood, 2023). However, our knowledge on photosynthesis has not plateaued. If indeed the key molecular, biochemical and physiological components of photosynthesis are nowadays well described, two avenues for future research are wide open: understanding the impact of the changing climate on photosynthetic processes, and enhancing the efficiency of photosynthesis as means for supporting the sustainable intensification of agriculture by increasing the productivity, or yields, of major crops (Murchie *et al.*, 2009).

1.2.1 Improving crop yield via (genetic engineering of) photosynthesis

"Global change" is a holistic term referring to the large-scale changes the biosphere is experiencing due to the interaction of its physico-chemical and biological components with modern human societies (Steffen, 2005; Cuff and Goudie, 2009; Cornell *et al.*, 2012). Among the challenges faced by humanity as a result of global change, the combined effect of population growth and insufficient increase of yields for the most prominent staple crops poses a particularly prominent issue to society, which is increasingly more investigated by the plant sciences community (Ray *et al.*, 2013). Technological innovations and plant breeding had a significant impact on agricultural production over the 20th century, resulting in spectacular yield increases for major crops. However, yield increases appear to have slowed, if not come to a halt, in the 21st century (Ray *et al.*, 2012, 2013; Schauberger *et al.*, 2018). The causes for the limited increase of crop yields, often referred to as yield stagnation, are complex and a comprehensive analysis of them would be beyond the scope of this thesis. Nevertheless, it is important to note that these causes can be divided into two major groups: causes preventing the increase of yield potential and causes impeding the realization of current yield potentials (van Ittersum *et al.*, 2013). Realizing yield potential, i.e. the yield that a crop can attain if optimal management practices are put in place, and without biotic and abiotic stresses, remains an important goal even within technologically advanced areas of the world such as Europe (Schils *et al.*, 2018). Additionally, realizing the yield potential of current crops will likely not be enough to meet the ever-increasing demand for food predicted for the next twenty to thirty years and beyond (Tian *et al.*, 2021). Therefore, an increase in yield potential will be essential in meeting future global demands for food production.

Yield potential for a crop is firstly determined by the amount of incident solar radiation over a cropping season (Monteith, 1977; Zhu *et al.*, 2010), a factor that cannot be substantially modified by human intervention. Two of the three additional determinants of yield potential, namely the efficiency of light interception and the efficiency of partitioning carbohydrates into harvestable organs, have been improved in crop plants to be very close to their theoretical maxima (Zhu *et al.*, 2010; Long *et al.*, 2015). However, the remaining determinant of yield potential, which is photosynthesis, has

not followed a similar pattern of improvement. In fact, the efficiency of photosynthetic energy conversion still has the potential to be improved by more than 50% (Zhu *et al.*, 2010; Foulkes *et al.*, 2022; Roney and Walker, 2023). The concept of increasing crop yield through genetic enhancement of photosynthesis was already proposed in the late 1980s (Austin, 1989). This idea quickly spread within the scientific community (El-Sharkawy and Cock, 1990; Mehta and Sarkar, 1992; Mehta *et al.*, 1992; Austin, 1993; Zeng-ping *et al.*, 1995; Lawlor, 1995; Prasad *et al.*, 1996), and its first practical applications emerged in the early 21st century, albeit with mostly mixed results (Ruan *et al.*, 2012). In the last fifteen years, more promising results have been obtained through a series of studies conducted on multiple crop species, demonstrating increased photosynthesis and, in some cases, improved yield as a result of genetic engineering of single or restricted sets of genes (Rosenthal *et al.*, 2011; Kromdijk *et al.*, 2016; South *et al.*, 2019; Wang *et al.*, 2020; De Souza *et al.*, 2022; López-Calcagno *et al.*, 2020, 2019; Yoon *et al.*, 2020). However, the inconsistency of results across multiple seasons (De Souza *et al.*, 2022), species (Garcia-Molina and Leister, 2020), or growing conditions (Ruiz-Vera *et al.*, 2022) serves as a stark reminder of the challenges in improving the highly complex trait of photosynthesis to achieve significant yield increases (Flexas, 2016; Sinclair *et al.*, 2019; Passioura, 2020; Araus *et al.*, 2021). These challenges are increasingly recognized within the photosynthesis research community, with a more holistic view of the photosynthetic processes being proposed as an approach to address them (Wu *et al.*, 2019; Kohli *et al.*, 2020; Zhu *et al.*, 2022; Harbinson and Yin, 2023; Wu *et al.*, 2023).

1.2.2 A case for natural variation and *Hirschfeldia incana*

One holistic way towards increasing crop yields via photosynthetic improvement would be utilizing natural genetic variation. This approach, based on exploring the genetic diversity present within crop species and their wild relatives, is not a new concept (Frankel, 1977; Poehlman and Quick, 1983; Austin, 1989). However, it was not until the early 2000s that advancements in technology enabled the systematic exploration of natural genetic variation, leading to a surge in interest in this area (Skovmand *et al.*, 2001a,b; Maloof *et al.*, 2001; Ramanatha Rao and Hodgkin, 2002; Gur and Zamir, 2004). The intersection of natural genetic variation and photosynthesis as means to enhance crop yield potential has particularly gained traction over the past decade (Lawson *et al.*, 2012; Huang and Han, 2014; van Bezouw *et al.*, 2019; Huang *et al.*, 2022; Sharwood *et al.*, 2022; Theeuwen *et al.*, 2022).

Many studies have focused on exploring natural variation within crops, model species, and even trees, thereby predominantly examining *intra*-specific variation (Theeuwen *et al.*, 2022; Taylor *et al.*, 2022). However, there has been an increasing emphasis on investigating the potential of crop wild relatives. These wild plant taxa, characterized by their relatively close genetic relationship to cultivated crops and harboring useful traits from a breeding perspective, have gained attention as potential resources for crop breeding efforts (Maxted *et al.*, 2006). In the past two decades, the use of crop wild relatives has already allowed for a better understanding and the

harnessing of the determinants of key crop traits (Brozynska *et al.*, 2016; Sharma *et al.*, 2020; Zhang and Batley, 2020; Quezada-Martinez *et al.*, 2021; Bohra *et al.*, 2022).

Based on the idea of investigating traits beyond the restricted genetic variability of crops, the focus of my research lies on *inter*-specific natural genetic variation for photosynthesis within the Brassicaceae family. This plant family comprises economically important crop species, crop wild relatives (Quezada-Martinez *et al.*, 2021), and the well-known model plant species *Arabidopsis thaliana* (Koornneef and Meinke, 2010), along with several species exhibiting intriguing photosynthetic characteristics. Recent studies have highlighted the presence of C3-C4 intermediate photosynthetic traits in certain Brassicaceae species (Triesch *et al.*, 2022; Guerreiro *et al.*, 2023), as well as variation in photosynthesis rates and dynamics among key *Brassica* crops (Taylor *et al.*, 2020). Furthermore, the exceptional photosynthetic performance under high irradiance of *Hirschfeldia incana*, a member of the Brassicaceae family, was reported over four decades ago (Canvin *et al.*, 1980).

1.3 Key definitions for photosynthesis research

The term “photosynthetic performance”, while descriptive, is too general to be accepted within the scientific community. At a closer look, it would be more appropriate to present *Hirschfeldia incana* as a species with high photosynthetic light-use efficiency under high irradiance (Garassino *et al.*, 2022). Widely accepted definitions for both *photosynthetic light-use efficiency* and *high irradiance* do not exist. Therefore, this section will provide such definitions as they were employed throughout the research project resulting in this thesis.

1.3.1 Defining high irradiance

Irradiance is primarily a radiometric concept, representing the radiant power received by a surface per unit area, measured in W/m^2 in SI units (International Organization for Standardization [ISO], 2022). However, in the context of plant science and photosynthesis research, irradiance is defined as the quantity of photons received by a surface per unit area and unit time, measured in $mol\,m^{-2}s^{-1}$ (Incoll *et al.*, 1981; Salisbury, 1991). The quantity defined in this way encompasses the totality of the electromagnetic spectrum, parts of which are not used by plants to drive their photosynthetic machinery. Therefore, irradiance is commonly defined in plant science in terms of photosynthetically active radiation (PAR), which encompasses the waveband from 400 to 700 nm. The combination of these two definitions results in the most common measure for irradiance in plant science and photosynthesis research, photosynthetic photon flux density (PPFD), measured as the quantity of photons in the 400-700 nm wavelength received by a surface per unit of area and time, and is generally expressed as $\mu mol\,m^{-2}\,s^{-1}$ (Alados *et al.*, 1996). While PPFD is an important measure of incident irradiance, it cannot represent the amount of radiation

available for photosynthesis over time. This has beeen overcome by the introduction of the daily light integral (DLI), a measure widely employed in horticultural research (Faust *et al.*, 2005; Van Iersel *et al.*, 2010; Kjaer *et al.*, 2012; Dou *et al.*, 2018; Gavhane *et al.*, 2023). DLI is defined as the total photosynthetic photon flux density delivered over the course of one day, obtained by multiplying PPFD by the duration of the irradiance exposure period. It is measured in $\mathrm{mol\,m^{-2}\,d^{-1}}$ (Faust and Logan, 2018).

Currently, no universally accepted definition exists for the concept of high irradiance or "high light". A review of studies conducted in highly controlled environments (i.e., growth chambers) reveals that the term "high light" has been associated to irradiance values ranging between 400 and 2000 $\mathrm{\mu mol\,m^{-2}\,s^{-1}}$ (Lichtenthaler and Burkart, 1999; Kouřil *et al.*, 2013; Szymańska *et al.*, 2017; van Rooijen *et al.*, 2018; Balfagón *et al.*, 2019). However, observations conducted in natural environments can be extrapolated to apply to artificial lighting and deliver a coherent definition of high irradiance. DLIs measured in warm-temperate areas during the months with the highest solar irradiance range between 60 and 65 $\mathrm{mol\,m^{-2}\,d^{-1}}$, consistent with year-long observations at the equator (Ritchie, 2010). Therefore, high irradiance can be defined as irradiance levels that enable the achievement of such DLIs or even higher. For typical conditions in highly controlled environments, such as climate-controlled plant growth chambers, with 12 hours of illumination and a "square" irradiance profile (i.e., with irradiance going from zero to its maximum value in the span of seconds), a PPFD of 1504 $\mathrm{\mu mol\,m^{-2}\,s^{-1}}$ would result in a DLI of 65 $\mathrm{mol\,m^{-2}\,d^{-1}}$. Thus, irradiance conditions exceeding 1500 $\mathrm{\mu mol\,m^{-2}\,s^{-1}}$ applied in typical controlled-environment settings can be referred to as high irradiance or equivalently as "high light". The same reasoning can be adapted to define case-specific high irradiance conditions that remain comparable at the DLI level.

1.3.2 Defining photosynthetic efficiency and capacity

Given the multiplicity of scales at which photosynthesis research is conducted, ranging from the molecular to the ecosystem scale, it is not surprising that different definitions for the "efficiency" of photosynthesis co-exist. In agronomic settings, photosynthetic efficiency is usually defined in terms of radiation use efficiency (RUE). RUE is calculated as the ratio between the biomass accumulated and radiating energy absorbed by a plant, and is expressed in $\mathrm{g\,MJ^{-1}}$ (Sinclair and Muchow, 1999). In the field of remote sensing, photosynthetic light-use efficiency (LUE) is generally defined as the ratio of net primary productivity (NPP) to absorbed photosynthetically active radiation. Since NPP is defined as the amount of carbon produced by primary producers per unit area and time, this measure of LUE is expressed in grams of plant-assimilated carbon per absorbed MJ of photosynthetically active radiation (Medlyn, 1998; Gitelson and Gamon, 2015).

In photosynthesis research conducted at the leaf or sub-leaf level, thus mostly focused on biophysical and physiological aspects, the concept of photosynthetic light-use efficiency is often taken to mean the quantum yield (Φ) of photosynthesis. Given

that the quantum yield for any light-dependent process is defined as the rate at which that process occurs relative to the rate of photon absorption by the system (Skillman, 2008), the quantum yield of photosynthesis is usually identified with the quantum yield for photosynthetic CO_2 fixation (ΦCO_2), i.e. the mol of CO_2 assimilated per mol of absorbed photons (Hogewoning *et al.*, 2012). It is worth pointing out that very often literature reports the *maximum* quantum yield of photosynthesis, i.e. the quantum yield measured under light-limited conditions corresponding to low irradiances (Singsaas *et al.*, 2001). Considering that with increasing irradiance factors intrinsic to the photosynthetic process (such as the biochemical characteristics of Ribulose-1,5-bisphosphate carboxylase/oxygenase, or RuBisCo) will reduce the photosynthetic quantum yield (Genty and Harbinson, 2004), photosynthetic LUE at high irradiances expressed in terms of ΦCO_2 might be misinterpreted. This is due to the fact that a high irradiance ΦCO_2 for a given plant will necessarily be lower than the maximum ΦCO_2 for the same plant, i.e. the value most researchers would be used to working with, and this could lead to wrongly categorizing the plant. Another parameter that is often employed as a proxy for photosynthetic quantum yield is the quantum yield of photosystem II (ΦPSII), the first component in the photosynthetic electron transport chain, measured based on chlorophyll fluorescence (Genty *et al.*, 1989). This measurement is affected by the same relationship to increasing irradiance (Yin *et al.*, 2011), and can therefore be prone to the same misinterpretation issues.

A more univocal definition of the photosynthetic LUE of plants can be achieved by relying on their maximum CO_2 assimilation rate (P_{max}), or photosynthetic capacity. P_{max} is defined as the CO_2 assimilation that can be measured when the photosynthetic apparatus of a plant is completely saturated by light (Marshall and Biscoe, 1980). A plant can sustain a high P_{max} only if its photosynthetic quantum yield at high, saturating irradiances is also high. High photosynthetic quantum yield at saturating irradiances results in higher CO_2 assimilation rates at non-saturating irradiances above the range of irradiances which are strictly light-limiting, and thus P_{max} can be a good description of the overall photosynthetic light-use efficiency of plants outside the strictly light-limited zone of irradiances.

1.4 Multidisciplinary approaches to resolve the complexity of photosynthesis

The biochemical and physiological complexity of photosynthesis is well-recognized, and has been addressed in the past forty years in different ways by photosynthesis researchers. While some researchers focused on the simultaneous multi-functional measurements of photosynthetic processes (e.g. Hogewoning *et al.* (2012)) others made use of increasingly detailed mathematical models to try and parametrize what could not be measured (Stirbet *et al.*, 2020). Recently, modelling efforts have advanced from trying to interpret and predict physiological alterations of photosynthesis as a response to differing environmental parameters, to predicting the effect of specific genetic modifications to improve photosynthesis (Zhu *et al.*, 2007), and to

compiling prioritized lists of bottleneck processes and underlying genes for genetic engineering (Long *et al.*, 2006; Ort *et al.*, 2011, 2015). The limited success achieved so far in increasing photosynthesis and crop yields by targeted engineering of single genetic components of the photosynthetic machinery, however, demonstrates that the genetic basis of the process is highly complex (van Bezouw *et al.*, 2019). In addition to the challenges posed by complexity of the genetic determinants of photosynthesis, researchers working to increase crop productivity by means of improved photosynthesis will face challenges due the impact of global climate change on the multi-faceted interaction between photosynthesis and the environment (Tkemaladze and Makhashvili, 2016; Armstrong *et al.*, 2023; Verslues *et al.*, 2023). Indeed, while studies evaluating increases in temperature and drought predicted by current climate change models have shown that those can exert a major disruptive force on photosynthetic dynamics (King *et al.*, 2006; Ruiz-Vera *et al.*, 2013; Gray *et al.*, 2016; Huang *et al.*, 2019), global climate change will also increase the frequency of extreme climatic event that will further threaten the crop productivity by, between other effects, further hampering photosynthetic CO_2 assimilation (Way *et al.*, 2021; Brestic *et al.*, 2021).

The level of complexity of the interactions between photosynthesis, global change, and crop yields is increasingly appreciated within the scientific community, and it is now established that the most promising, if not the only, way to address them and realize the much-needed yield increases with photosynthesis is the result of multidisciplinary research efforts (Araus *et al.*, 2021; Zhu *et al.*, 2022; Zenda *et al.*, 2023). The swift development of high-throughput plant phenotyping technologies combined with the improvement of (DNA) sequencing technologies and bioinformatics tools has made large-scale studies aimed at unraveling the genetic determinants of photosynthetic traits possible (van Bezouw *et al.*, 2019; Theeuwen *et al.*, 2022). Technological development, however, has progressed for other disciplines as well, offering promising tools for further expanding the scope of multidisciplinary photosynthesis research. Advancements in sequencing technology and data analysis have largely benefited the field of transcriptomics, making detailed studies of the transcriptional signature of photosynthetic responses possible in many plant species (Miller *et al.*, 2017; Xiong *et al.*, 2019; Loudya *et al.*, 2021; An *et al.*, 2022; Han *et al.*, 2023), as well as the study of photosynthetic responses to novel agronomic practices (Song *et al.*, 2020; Zhu *et al.*, 2023; Chai *et al.*, 2023). In the context of genomics and transcriptomics, the development of pangenomics approaches aimed at surpassing the compression of large sets of genomic and transcriptomic sequences into single consensus assemblies has opened up new avenues for the identification of novel genetic variants for crop improvement (Tao *et al.*, 2019; Zanini *et al.*, 2022; Li *et al.*, 2022; Chapman *et al.*, 2022) and will allow for the exploration of photosynthesis-related genes and transcriptional patterns from a novel perspective. The development of more accurate and quantitative mass spectrometry technologies and methods has enabled advanced studies of the photosynthetic protein machinery at unprecedented resolutions (Fan *et al.*, 2019; Lande *et al.*, 2020; Wang *et al.*, 2021; Zimmer *et al.*, 2021), as well as the response of photosynthetic proteomes to differences in environmental parameters such as light quality and irradiance (Flannery *et al.*, 2021a,b).

These technologies, combined with newly developed or largely improved ones that allow for the resolution of structures of photosynthetic proteins, highly precise gene editing, and even guided evolution of photosynthetic components (Zhu *et al.*, 2022) are providing the photosynthesis research community with an unprecedented wealth of data. This presents researchers with the possibility to follow multidisciplinary approaches in their quest to unravel the complexity of the relationship between photosynthesis and crop productivity (Araus *et al.*, 2021; Zhu *et al.*, 2022; Zenda *et al.*, 2023).

1.5 Rationale and outline of this thesis

The research culminating in this thesis stems from the unprecedented possibilities for multi-disciplinary research on photosynthesis described above. The project I undertook was the first to start within a larger research initiative on exploring *Hirschfeldia incana* for its high photosynthetic light-use efficiency at high irradiance. Thus, I first focused on establishing genomics, transcriptomics, and phenomics resources that would then serve as the foundation for studies on the genetic and physiological determinants of the species' photosynthetic LUE. I then made use of the established resources to compare *H. incana* to a number of its relatives from the Brassicaceae family, testing the hypothesis that differences in photosynthetic LUE among the species, especially under high irradiances, could be explained in terms of physiological and underlying genetic differences. While naturally far from delivering a full explanation of these differences, this thesis presents the most relevant discoveries I made on the photosynthetic LUE of *H. incana* and its genomic, transcriptomic, physiological, and, ultimately, genetic determinants.

Prior to this work, the remarkable photosynthetic characteristics of *H. incana* had only been succinctly reported in 1980 (Canvin *et al.*, 1980), while the species was marginally recognized by the scientific community for its status as a nuisance weed and metal accumulator (Lee *et al.*, 2004; Auguy *et al.*, 2013). *H. incana*, however, presents a number of characteristics making it the perfect candidate for studies on photosynthetic capacity and efficiency, and a valuable addition to studies on natural genetic variation in photosynthesis. Therefore, in **Chapter 2**, co-authors and I expand the case for greater focus on exploring and exploiting natural genetic variation in photosynthesis presented in this introduction. We discuss how we need to identify which photosynthetic traits to target, first and foremost at the leaf level, if we are to commit significant energies to increasing the photosynthesis of major crops. We propose maximum photosynthetic rate (P_{max}, or photosynthetic capacity) as an excellent trait for the improvement of crop photosynthesis, provided its genetic basis is elucidated. To achieve this ambitious goal, we propose the use of naturally occurring species with high photosynthetic capacity as models for exploring the physiological and genetic basis of high photosynthetic efficiency. After reviewing literature for reports of such species, we identify Brassicaceae species *H. incana* (L.) Lagr.-Foss. as a promising candidate, and proceed to describe the basic biology, evolutionary history, and photosynthetic characteristics of the species. Based on the number of

highlighted positive characteristics, we conclude that *H. incana* has the potential to be an excellent model species for studies aiming at understanding natural genetic variation in photosynthetic light-use efficiency.

In **Chapter 3**, I lay the foundation for the use of *H. incana* as a plant model species for studies on photosynthetic light-use efficiency. After confirming the species' exceptional photosynthetic LUE and capacity by comparing it to a selection of relatives from the Brassicaceae species, co-authors and I present the results of the *H. incana* genome sequencing project, enabling comparative genomics and transcriptomics. We describe how *H. incana* has extensively diversified from those of close relatives *Brassica rapa* and *Brassica nigra* through large chromosomal rearrangements, species-specific transposon activity, and differential retention of duplicated genes. Considering the pivotal role of gene duplication in driving evolution of traits improving plant performance in challenging environments, we investigate whether variation in copy number for a selection of genes involved in photosynthesis and/or photoprotection correlates with their expression in *H. incana*, *B. nigra*, and *B. rapa*, and with the higher photosynthetic capacity of these species compared to that of *Arabidopsis thaliana*. Being able to show that this holds true for six of nine tested genes, and having excluded a general effect of overexpression of photosynthetic genes, we conclude that the role of copy number variation on the expression of genes potentially playing a role in photosynthetic light-use efficiency appears important and should be subjected to a broader investigation. We also point out how such future investigations will benefit from the constructed genome assembly and annotation of *H. incana*.

Following up on the suggestion of a greater focus on transcriptional patterns of genes potentially related to photosynthetic light-use efficiency, in **Chapter 4**, I present the results of a study conducted on the transcriptomes of *H. incana*, *B. nigra*, *B. rapa*, and *A. thaliana* plants grown under contrasting low and high irradiance conditions. After describing how I designed and constructed a growth system capable of consistently and reliably growing plants with "super-natural" irradiance, *i.e.* resulting in DLIs higher than reported anywhere so far, co-authors and I present report on the results of the comparative analysis of gene expression we performed using a panproteome to identify homology relationships between genes of the different species. We report that as expected all four species actively regulate genes associated with the photosynthetic process in response to high irradiance, and I describe unique gene expression patterns specific to *H. incana*. We show how in certain cases *H. incana* exhibits differential expression for specific genes not observed in the other species, while in other cases the species achieves significantly higher transcript abundance, sometimes independent of the irradiance treatment, via the simultaneous expression of multiple gene copies. Therefore, We conclude that *H. incana* has a number of specific gene expression patterns resulting from high irradiance, and that these patterns converge into three strategies: constitutive higher gene expression, higher gene expression determined by multiple gene copies, and canonical differential gene expression as a result of irradiance. Furthermore, We point out how the large dataset generated within my study should be further analysed with focus shifting from strictly photosynthetic genes to genes encoding to physiological processes such as e.g. water and nutrients uptake that must be crucial in supporting high photosynthetic

capacity.

Building on the comparative approach that has been the foundation of the work described in previous chapters, in **Chapter 5** I present the results of a phenomics-enabled exploration of photosynthetic light-use efficiency and a set of potentially underlying traits in a panel of ten Brassicaceae species representing key events in the evolutionary history of the family. This research confirms that *H. incana*'s photosynthetic light-use efficiency, measured as quantum yield of photosystem II or ΦPSII, is higher than that of most species in the panel, while reporting the high photosynthetic LUE of *B. nigra*, *Brassica tournefortii*, and *Zahora ait-atta*. Co-authors and I furthermore describe how an analysis of twenty-one parameters encompassing six classes of physiological and anatomical plant characteristics yielded an unexpected correlation between ΦPSII parameters and stomatal counts, thus implying their role in achieving high photosynthetic LUE. After discussing how this chapter presents a number of novel methodologies to create and explore phenomics datasets, we conclude that the dataset we generated on natural variation for photosynthesis and underlying traits is a fundamental resource providing a foundation for future genomic and transcriptomic dissection.

In the final chapter of this thesis, **Chapter 6**, I integrate all the findings described in the previous chapters following two directions. First, I compile a description of observed natural variation for photosynthesis and its light-use efficiency, and explain how *Hirschfeldia incana* emerges as the current best performer within the family. Then, I provide a unifying view of the results of my genomics and transcriptomics investigations, and summarize the knowledge I have accumulated on the genetic algorithm of high photosynthetic LUE. I then interpret all the findings of my work in the context of the Leaf Economics Spectrum, speculating on additional components of the photosynthetic algorithm of *H. incana*. Expanding on ecophysiological considerations, I argue that key traits enabling high photosynthetic LUE can be identified outside leaves, and make a case for a less leaf-centric perspective in photosynthesis research. I then argue that the field of photosynthesis research, and its attemtps to increase the yield of crops, will benefit from focusing more on non-crop species, such as the one presented in this thesis. Finally, I reflect on why photosynthesis research has not yet met the goal of crop yield improvement, and on what kind of perspective change is required to enable to reach this goal. I conclude the chapter by listing future directions for research stemming from the questions prompted by this work.

References

Alados, I., Foyo-Moreno, I., and Alados-Arboledas, L. (1996). Photosynthetically active radiation: Measurements and modelling. *Agricultural and Forest Meteorology*, **78**(1), 121–131.

An, J., Wei, X., and Huo, H. (2022). Transcriptome analysis reveals the accelerated expression of genes related to photosynthesis and chlorophyll biosynthesis contribution to shade-tolerant in Phoebe bournei. *BMC Plant Biology*, **22**(1), 268.

Araus, J. L., Sanchez-Bragado, R., and Vicente, R. (2021). Improving crop yield and resilience through optimization of photosynthesis: Panacea or pipe dream? *Journal of Experimental Botany*, **72**(11), 3936–3955.

Armstrong, E. M., Larson, E. R., Harper, H., Webb, C. R., Dohleman, F., Araya, Y., Meade, C., Feng, X., Mukoye, B., Levin, M. J., Lacombe, B., Bakirbas, A., Cardoso, A. A., Fleury, D., Gessler, A., Jaiswal, D., Onkokesung, N., Pathare, V. S., Phartyal, S. S., Sevanto, S. A., Wilson, I., and Grierson, C. S. (2023). One hundred important questions facing plant science: An international perspective. *New Phytologist*, **238**(2), 470–481.

Arnon, D. I., Whatley, F. R., and Allen, M. B. (1954). Photosynthesis by Isolated Chloroplasts. II. Photosynthetic Phosphorylation, the Conversion of Light into Phosphate Bond Energy. *Journal of the American Chemical Society*, **76**(24), 6324–6329.

Auguy, F., Fahr, M., Moulin, P., Brugel, A., Laplaze, L., Mzibri, M. E., Filali-Maltouf, A., Doumas, P., and Smouni, A. (2013). Lead Tolerance and Accumulation in Hirschfeldia incana, a Mediterranean Brassicaceae from Metalliferous Mine Spoils. *PLOS ONE*, **8**(5), e61932.

Austin, R. B. (1989). Genetic variation in photosynthesis. *The Journal of Agricultural Science*, **112**(3), 287–294.

Austin, R. B. (1993). Crop Photosynthesis: Can We Improve on Nature? In *International Crop Science I*, chapter 91, pages 697–701. John Wiley & Sons, Ltd.

Balfagón, D., Sengupta, S., Gómez-Cadenas, A., Fritschi, F. B., Azad, R. K., Mittler, R., and Zandalinas, S. I. (2019). Jasmonic Acid Is Required for Plant Acclimation to a Combination of High Light and Heat Stress. *Plant Physiology*, **181**(4), 1668–1682.

Barnes, C. R. (1893). On the food of green plants. *Botanical Gazette*, **18**(11), 403–411.

Bassham, J. A., Benson, A. A., Kay, L. D., Harris, A. Z., Wilson, A. T., and Calvin, M. (1954). The path of carbon in photosynthesis. XXI. The cyclic regeneration of carbon dioxide acceptor1. *Journal of the American chemical society*, **76**(7), 1760–1770.

Becklin, K. M. . (2021). *Photosynthesis, Respiration. and Climate Change*. Advances in Photosynthesis and Respiration, Including Bioenery and Related Processes ; Volume 48. Springer, Cham, Switzerland.

Blackman, F. F. (1905). Optima and limiting factors. *Annals of botany*, **19**(74), 281–295.

Blankenship, R. E. (2021). *Molecular Mechanisms of Photosynthesis*. Wiley, Hoboken, third edition. edition.

Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., and Varshney, R. K. (2022). Reap the crop wild relatives for breeding future crops. *Trends in Biotechnology*, **40**(4), 412–431.

Brestic, M., Yang, X., Li, X., and Allakhverdiev, S. I. (2021). Crop photosynthesis for the twenty-first century. *Photosynthesis Research*, **150**(1), 1–3.

Brozynska, M., Furtado, A., and Henry, R. J. (2016). Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnology Journal*, **14**(4), 1070–1085.

Calvin, M. (1962). The Path of Carbon in Photosynthesis. *Angewandte Chemie International Edition in English*, **1**(2), 65–75.

Canvin, D. T., Berry, J. A., Badger, M. R., Fock, H., and Osmond, C. B. (1980). Oxygen exchange in leaves in the light. *Plant Physiology*, **66**(2), 302–307.

Chai, S.-K., Ooi, S.-E., Ho, C.-L., Ong-Abdullah, M., Chan, K.-L., Fitrianto, A., and Namasivayam, P. (2023). Transcriptomic Analysis Reveals Suppression of Photosynthesis and Chlorophyll Synthesis Following Gibberellic Acid Treatment on Oil Palm (Elaies guineensis). *Journal of Plant Growth Regulation*.

Chapman, M. A., He, Y., and Zhou, M. (2022). Beyond a reference genome: Pangenomes and population genomics of underutilized and orphan crops for future food and nutrition security. *New Phytologist*, **234**(5), 1583–1597.

Cornell, S. ., Prentice, I. C. ., House, J., and Downy, C. (2012). *Understanding the Earth System : Global Change Science for Application*. Cambridge University Press, Cambridge ;.

Cuff, D. J. and Goudie, Andrew. (2009). *The Oxford Companion to Global Change*. Oxford University Press, New York.

de Lavoisier, A. L. and Cuchet, G. J. (1789). *Traité Élémentaire de Chimie: Présenté Dans Un Ordre Nouveau et d'après Les Découvertes Modernes ...* Chez Cuchet.

de Saussure, N. T. (1804). *Recherches Chimiques Sur La Vegetation*. Nyon.

De Souza, A. P., Burgess, S. J., Doran, L., Hansen, J., Manukyan, L., Maryn, N., Gotarkar, D., Leonelli, L., Niyogi, K. K., and Long, S. P. (2022). Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science*, **377**(6608), 851–854.

Devlin, R. M. (1969). *Plant Physiology*. Van Nostrand Reinhold, New York, 2nd ed. edition.

Dou, H., Niu, G., Gu, M., and Masabni, J. G. (2018). Responses of Sweet Basil to Different Daily Light Integrals in Photosynthesis, Morphology, Yield, and Nutritional Quality. *HortScience*, **53**(4), 496–503.

Drews, G. (2005). Contributions of Theodor Wilhelm Engelmann on phototaxis, chemotaxis, and photosynthesis. *Photosynthesis Research*, **83**(1), 25–34.

El-Sharkawy, M. A. and Cock, J. H. (1990). Photosynthesis of Cassava (Manihot esculenta). *Experimental Agriculture*, **26**(3), 325–340.

Emerson, R. and Arnold, W. (1932). A separation of the reactions in photosynthesis by means of intermittent light. *The Journal of General Physiology*, **15**(4), 391–420.

Emerson, R. and Chalmers, R. V. (1958). Speculations concerning the function and phylogenetic significance of the accessory pigments of algae. *Phycol Soc News Bull*, **11**, 51–56.

Emerson, R. and Rabinowitch, E. (1960). Red Drop and Role of Auxiliary Pigments in Photosynthesis. *Plant Physiology*, **35**(4), 477–485.

Emerson, R., Chalmers, R., and Cederstrand, C. (1957). Some factors influencing the long-wave limit of photosynthesis. *Proceedings of the National Academy of Sciences*, **43**(1), 133–143.

Fan, Y., Chen, J., Wang, Z., Tan, T., Li, S., Li, J., Wang, B., Zhang, J., Cheng, Y., Wu, X., Yang, W., and Yang, F. (2019). Soybean (Glycine max L. Merr.) seedlings response to shading: Leaf structure, photosynthesis and proteomic analysis. *BMC Plant Biology*, **19**(1), 34.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta*, **149**(1), 78–90.

Faust, J. E. and Logan, J. (2018). Daily Light Integral: A Research Review and High-resolution Maps of the United States. *HortScience horts*, **53**(9), 1250–1257.

Faust, J. E., Holcombe, V., Rajapakse, N. C., and Layne, D. R. (2005). The Effect of Daily Light Integral on Bedding Plant Growth and Flowering. *HortScience*, **40**(3), 645–649.

Flannery, S. E., Pastorelli, F., Wood, W. H. J., Hunter, C. N., Dickman, M. J., Jackson, P. J., and Johnson, M. P. (2021a). Comparative proteomics of thylakoids from Arabidopsis grown in laboratory and field conditions. *Plant Direct*, **5**(10), e355.

Flannery, S. E., Hepworth, C., Wood, W. H. J., Pastorelli, F., Hunter, C. N., Dickman, M. J., Jackson, P. J., and Johnson, M. P. (2021b). Developmental acclimation of the thylakoid proteome to light intensity in Arabidopsis. *The Plant Journal*, **105**(1), 223–244.

Flexas, J. (2016). Genetic improvement of leaf photosynthesis and intrinsic water use efficiency in C3 plants: Why so much little success? *Plant Science*, **251**, 155–161.

Foulkes, M. J., Molero, G., Griffiths, S., Slafer, G. A., and Reynolds, M. P. (2022). Yield Potential. In M. P. Reynolds and H.-J. Braun, editors, *Wheat Improvement: Food Security in a Changing Climate*, pages 379–396. Springer International Publishing, Cham.

Frankel, O. H. (1977). Natural Variation and Its Conservation. In A. Muhammed, R. Aksel, and R. C. von Borstel, editors, *Genetic Diversity in Plants*, Basic Life Sciences, pages 21–44. Springer US, Boston, MA.

Garassino, F., Wijfjes, R. Y., Boesten, R., Reyes Marquez, F., Becker, F. F. M., Clapero, V., van den Hatert, I., Holmer, R., Schranz, M. E., Harbinson, J., de Ridder, D., Smit, S., and Aarts, M. G. M. (2022). The genome sequence of Hirschfeldia incana, a new Brassicaceae model to improve photosynthetic light-use efficiency. *The Plant Journal*, **112**(1), 1298–1315.

Garcia-Molina, A. and Leister, D. (2020). Accelerated relaxation of photoprotection impairs biomass accumulation in Arabidopsis. *Nature Plants*, **6**(1), 9–12.

Gavhane, K. P., Hasan, M., Singh, D. K., Kumar, S. N., Sahoo, R. N., and Alam, W. (2023). Determination of optimal daily light integral (DLI) for indoor cultivation of iceberg lettuce in an indigenous vertical hydroponic system. *Scientific Reports*, **13**(1), 10923.

Geerdt, M. (2007). Jan IngenHousz, or why don't we know who discovered photosynthesis? In *1st Conference of the European Philosophy of Science Association*, Madrid.

Genty, B. and Harbinson, J. (2004). Regulation of Light Utilization for Photosynthetic Electron Transport. In N. R. Baker, editor, *Photosynthesis and the Environment*, volume 5, pages 67–99. Kluwer Academic Publishers, Dordrecht.

Genty, B., Briantais, J.-M., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA) - General Subjects*, **990**(1), 87–92.

Gest, H. (2000). Bicentenary homage to Dr Jan Ingen-Housz, MD (1730–1799), pioneer of photosynthesis research. *Photosynthesis Research*, **63**(2), 183–190.

Gitelson, A. A. and Gamon, J. A. (2015). The need for a common basis for defining light-use efficiency: Implications for productivity estimation. *Remote Sensing of Environment*, **156**, 196–201.

Govindjee, Beatty, J. T., Gest, H., and Allen, J. F. (2005). *Discoveries in Photosynthesis*, volume 20. Springer.

Govindjee, Shevela, D., and Björn, L. O. (2017). Evolution of the Z-scheme of photosynthesis: A perspective. *Photosynthesis Research*, **133**(1), 5–15.

Gray, S. B., Dermody, O., Klein, S. P., Locke, A. M., McGrath, J. M., Paul, R. E., Rosenthal, D. M., Ruiz-Vera, U. M., Siebers, M. H., Strellner, R., Ainsworth, E. A., Bernacchi, C. J., Long, S. P., Ort, D. R., and Leakey, A. D. B. (2016). Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. *Nature Plants*, **2**(9), 1–8.

Guerreiro, R. F., Bonthala, V. S., Schluter, U., Triesch, S., Weber, A. P., and Stich, B. (2023). A genomic panel for studying C3-C4 intermediate photosynthesis in the Brassiceae tribe. *bioRxiv*, page 2023.01.22.525068.

Gur, A. and Zamir, D. (2004). Unused Natural Variation Can Lift Yield Barriers in Plant Breeding. *PLOS Biology*, **2**(10), e245.

Hales, S. (1727). *Statical Essays: Vegetable Staticks: Or, An Account of Some Statical Experiments On The Sap in Vegetables: Being an Essay towards a Natural History of Vegetation. Also, a Specimen of An Attempt to Analyse the Air, By a Great Variety of Chymio-Statical Experiments; Which Were Read at Several Meetings before the Royal Society. 1*, volume 1. W. and J. Innys; T. Woodward.

Han, S.-Y., Kim, W.-Y., Kim, J. S., and Hwang, I. (2023). Comparative transcriptomics reveals the role of altered energy metabolism in the establishment of single-cell C4 photosynthesis in Bienertia sinuspersici. *Frontiers in Plant Science*, **14**.

Harbinson, J. and Yin, X. (2023). Modelling the impact of improved photosynthetic properties on crop performance in Europe. *Food and Energy Security*, **12**(1), e402.

Hill, J. F. (2012). Early Pioneers of Photosynthesis Research. In J. J. Eaton-Rye, B. C. Tripathy, and T. D. Sharkey, editors, *Photosynthesis: Plastid Biology, Energy Conversion and Carbon Assimilation*, pages 771–800. Springer Netherlands, Dordrecht.

Hill, R. (1937). Oxygen Evolved by Isolated Chloroplasts. *Nature*, **139**(3525), 881–882.

Hill, R. and Bendall, F. (1960). Function of the Two Cytochrome Components in Chloroplasts: A Working Hypothesis. *Nature*, **186**(4719), 136–137.

Hintz, N. H. (2021). Highlighting Theodor W. Engelmann's "Farbe und Assimilation"[Color and Assimilation]. *Limnology and Oceanography Bulletin*, **30**(4), 121–126.

Hogewoning, S. W., Wientjes, E., Douwstra, P., Trouwborst, G., van Ieperen, W., Croce, R., and Harbinson, J. (2012). Photosynthetic Quantum Yield Dynamics: From Photosystems to Leaves. *The Plant Cell*, **24**(5), 1921–1935.

Huang, G., Peng, S., and Li, Y. (2022). Variation of photosynthesis during plant evolution and domestication: Implications for improving crop photosynthesis. *Journal of Experimental Botany*, **73**(14), 4886–4896.

Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T. F., Peng, S., Berry, J. A., Wang, K., Mao, J., Alkama, R., Cescatti, A., Cuntz, M., De Deurwaerder, H., Gao, M., He, Y., Liu, Y., Luo, Y., Myneni, R. B., Niu, S., Shi, X., Yuan, W., Verbeeck, H., Wang, T., Wu, J., and Janssens, I. A. (2019). Air temperature optima of vegetation productivity across global biomes. *Nature Ecology & Evolution*, **3**(5), 772–779.

Huang, X. and Han, B. (2014). Natural Variations and Genome-Wide Association Studies in Crop Plants. *Annual Review of Plant Biology*, **65**(1), 531–551.

Incoll, L. D., Long, S. P., and Ashmore, M. A. (1981). SI units in publications in plant science. *Commentaries in plant science*, **2**, 83–96.

Ingen-Housz, J. (1779). *Experiments upon Vegetables: Discovering Their Great Power of Purifying the Common Air in the Sun-Shine, and of Injuring It in the Shade and at Night. To Which Is Joined, a New Method of Examining the Accurate Degree of Salubrity of the Atmosphere*. P. Elmsly, in the Strand; and H. Payne, in Pall Mall.

International Organization for Standardization [ISO] (2022). ISO 9288:2022.

Janssen, P. J. D., Lambreva, M. D., Plumeré, N., Bartolucci, C., Antonacci, A., Buonasera, K., Frese, R. N., Scognamiglio, V., and Rea, G. (2014). Photosynthesis at the forefront of a sustainable life. *Frontiers in Chemistry*, **2**.

King, A. W., Gunderson, C. A., Post, W. M., Weston, D. J., and Wullschleger, S. D. (2006). Plant Respiration in a Warmer World. *Science*, **312**(5773), 536–537.

Kjaer, K. H., Ottosen, C.-O., and Jørgensen, B. N. (2012). Timing growth and development of Campanula by daily light integral and supplemental light level in a cost-efficient light control system. *Scientia Horticulturae*, **143**, 189–196.

Koch, W., Lange, O. L., and Schulze, E. D. (1971). Ecophysiological investigations on wild and cultivated plants in the Negev Desert. *Oecologia*, **8**(3), 296–309.

Kohli, A., Miro, B., Balié, J., and d'A Hughes, J. (2020). Photosynthesis research: A model to bridge fundamental science, translational products, and socio-economic considerations in agriculture. *Journal of Experimental Botany*, **71**(7), 2281–2298.

Koornneef, M. and Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal*, **61**(6), 909–921.

Kouřil, R., Wientjes, E., Bultema, J. B., Croce, R., and Boekema, E. J. (2013). High-light vs. low-light: Effect of light acclimation on photosystem II composition and organization in Arabidopsis thaliana. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1827**(3), 411–419.

Krikorian, A. D. and Steward, F. C. (1968). Water and Solutes in Plant Nutrition: With Special Reference to van Helmont and Nicholas of Cusa. *BioScience*, **18**(4), 286–292.

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**(6314), 857–861.

Lambers, H. and Oliveira, R. S. (2019). *Plant Physiological Ecology*. Springer, Cham, Switzerland, third edition. edition.

Lande, N. V., Barua, P., Gayen, D., Kumar, S., Chakraborty, S., and Chakraborty, N. (2020). Proteomic dissection of the chloroplast: Moving beyond photosynthesis. *Journal of Proteomics*, **212**, 103542.

Lawlor, D. W. (1995). Photosynthesis, productivity and environment. *Journal of Experimental Botany*, **46**(special_issue), 1449–1461.

Lawson, T., Kramer, D. M., and Raines, C. A. (2012). Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. *Current Opinion in Biotechnology*, **23**(2), 215–220.

Lee, P. L. M., Patel, R. M., Conlan, R. S., Wainwright, S. J., and Hipkin, C. R. (2004). Comparison of Genetic Diversities in Native and Alien Populations of Hoary Mustard (Hirschfeldia incana [L.] Lagreze-Fossat). *International Journal of Plant Sciences*, **165**(5), 833–843.

Li, W., Liu, J., Zhang, H., Liu, Z., Wang, Y., Xing, L., He, Q., and Du, H. (2022). Plant pan-genomics: Recent advances, new challenges, and roads ahead. *Journal of Genetics and Genomics*, **49**(9), 833–846.

Lichtenthaler, H. K. and Burkart, S. (1999). Photosynthesis and high light stress. *Bulg. J. Plant Physiol*, **25**(3-4), 3–16.

Long, S., Farage, P., and Garcia, R. (1996). Measurement of leaf and canopy photosynthetic CO2 exchange in the field. *Journal of Experimental Botany*, **47**(11), 1629–1642.

Long, S. P., Zhu, X.-G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment*, **29**(3), 315–330.

Long, S. P., Marshall-Colon, A., and Zhu, X.-G. (2015). Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell*, **161**(1), 56–66.

Loomis, W. E. (1960). Historical introduction. In A. Pirson, editor, *Die CO2-Assimilation / The Assimilation of Carbon Dioxide: In 2 Teilen / 2 Parts*, pages 85–114. Springer Berlin Heidelberg, Berlin, Heidelberg.

López-Calcagno, P. E., Fisk, S., Brown, K. L., Bull, S. E., South, P. F., and Raines, C. A. (2019). Overexpressing the H-protein of the glycine cleavage system increases biomass yield in glasshouse and field-grown transgenic tobacco plants. *Plant Biotechnology Journal*, **17**(1), 141–151.

López-Calcagno, P. E., Brown, K. L., Simkin, A. J., Fisk, S. J., Vialet-Chabrand, S., Lawson, T., and Raines, C. A. (2020). Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants*, **6**(8), 1054–1063.

Loudya, N., Mishra, P., Takahagi, K., Uehara-Yamaguchi, Y., Inoue, K., Bogre, L., Mochida, K., and López-Juez, E. (2021). Cellular and transcriptomic analyses reveal two-staged chloroplast biogenesis underpinning photosynthesis build-up in the wheat leaf. *Genome Biology*, **22**(1), 151.

Macdonald, M. (2020). *Patrick Geddes's Intellectual Origins*. Edinburgh University Press.

Magiels, G. (2010). *From Sunlight to Insight: Jan IngenHousz, the Discovery of Photosynthesis & Science in the Light of Ecology*. VUBPRESS.

Maloof, J. N., Borevitz, J. O., Dabi, T., Lutes, J., Nehring, R. B., Redfern, J. L., Trainer, G. T., Wilson, J. M., Asami, T., Berry, C. C., Weigel, D., and Chory, J. (2001). Natural variation in light sensitivity of Arabidopsis. *Nature Genetics*, **29**(4), 441–446.

Marshall, B. and Biscoe, P. V. (1980). A Model for C3 Leaves Describing the Dependence of Net Photosynthesis on Irradiance. *Journal of Experimental Botany*, **31**(1), 29–39.

Maxted, N., Ford-Lloyd, B. V., Jury, S., Kell, S., and Scholten, M. (2006). Towards a definition of a crop wild relative. *Biodiversity & Conservation*, **15**(8), 2673–2685.

McAlister, E. D. and Myers, J. (1940). The time course of photosynthesis and fluorescence observed simultaneously. *Smithsonian Miscellaneous Collections*.

Medlyn, B. E. (1998). Physiological basis of the light use efficiency model. *Tree Physiology*, **18**(3), 167–176.

Mehta, H. and Sarkar, K. R. (1992). Heterosis for leaf photosynthesis, grain yield and yield components in maize. *Euphytica*, **61**(2), 161–168.

Mehta, H., Sarkar, K. R., and Sharma, S. K. (1992). Genetic analysis of photosynthesis and productivity in corn. *Theoretical and Applied Genetics*, **84**(1), 242–255.

Miller, M. A. E., O'Cualain, R., Selley, J., Knight, D., Karim, M. F., Hubbard, S. J., and Johnson, G. N. (2017). Dynamic Acclimation to High Light in Arabidopsis thaliana Involves Widespread Reengineering of the Leaf Proteome. *Frontiers in Plant Science*, **8**, 1239.

Monteith, J. L. (1977). Climate and the efficiency of crop production in Britain. *Philosophical transactions of the royal society of London. B, Biological Sciences*, **281**(980), 277–294.

Murchie, E. H., Pinto, M., and Horton, P. (2009). Agriculture and the new challenges for photosynthesis research. *New Phytologist*, **181**(3), 532–552.

Nickelsen, K. (2015a). *Explaining Photosynthesis: Models of Biochemical Mechanisms, 1840-1960*, volume 8. Springer.

Nickelsen, K. (2015b). The Path of Carbon in Photosynthesis (1937–1954). In *Explaining Photosynthesis: Models of Biochemical Mechanisms, 1840-1960*, pages 201–250. Springer Netherlands, Dordrecht.

Nickelsen, K. (2015c). Struggling with the Standard Model (1930–1941). In *Explaining Photosynthesis: Models of Biochemical Mechanisms, 1840-1960*, pages 93–147. Springer Netherlands, Dordrecht.

Nobel Lectures in Chemistry (1966). *Nobel Lectures in Chemistry 1901 - 1921*. Elsevier, Amsterdam.

Ort, D. R., Zhu, X., and Melis, A. (2011). Optimizing Antenna Size to Maximize Photosynthetic Efficiency. *Plant Physiology*, **155**(1), 79–85.

Ort, D. R., Merchant, S. S., Alric, J., Barkan, A., Blankenship, R. E., Bock, R., Croce, R., Hanson, M. R., Hibberd, J. M., Long, S. P., Moore, T. A., Moroney, J., Niyogi, K. K., Parry, M. A. J., Peralta-Yahya, P. P., Prince, R. C., Redding, K. E., Spalding, M. H., van Wijk, K. J., Vermaas, W. F. J., von Caemmerer, S., Weber, A. P. M., Yeates, T. O., Yuan, J. S., and Zhu, X. G. (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proceedings of the National Academy of Sciences*, **112**(28), 8529–8536.

Passioura, J. B. (2020). Translational research in agriculture. Can we do it better? *Crop and Pasture Science*, **71**(6), 517–528.

Pennazio, S. (2011). The Dawn of Photosynthesis. *Theoretical Biology Forum : 104, 2, 2011*.

Pessarakli, M. . (2016). *Handbook of Photosynthesis*. Books in Soils, Plants, and the Environment. CRC Press, Boca Raton, FL, third edition. edition.

Poehlman, J. M. and Quick, J. S. (1983). Crop Breeding in a Hungry World. In *Crop Breeding*, chapter 1, pages A–19. John Wiley & Sons, Ltd.

Prasad, B. T. K., Sheshshayee, M. S., Nataraj, K. N., Shankar, A. G., Udayakumar, M., and Prasad, T. G. (1996). Regulation of carboxylation by RuBisCO content and its efficiency in sunflower and soybean. *Current Science*, **70**(7), 675–680.

Priestley, J. (1776). *Experiments and Observations on Different Kinds of Air*. J. Johnson.

Quezada-Martinez, D., Addo Nyarko, C. P., Schiessl, S. V., and Mason, A. S. (2021). Using wild relatives and related species to build climate resilience in Brassica crops. *Theoretical and Applied Genetics*, **134**(6), 1711–1728.

Ramanatha Rao, V. and Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture*, **68**(1), 1–19.

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., and Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications*, **3**(1), 1293.

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**(6), e66428.

Ritchie, R. J. (2010). Modelling photosynthetic photon flux density and maximum potential gross photosynthesis. *Photosynthetica*, **48**(4), 596–609.

Roney, H. E. and Walker, B. J. (2023). Photon to plate: A holistic view of photosynthetic and anthropogenic energy fluxes. *Food and Energy Security*, **12**(1), e348.

Rosenthal, D. M., Locke, A. M., Khozaei, M., Raines, C. A., Long, S. P., and Ort, D. R. (2011). Over-expressing the C3 photosynthesis cycle enzyme Sedoheptulose-1-7 Bisphosphatase improves photosynthetic carbon gain and yield under fully open air CO2fumigation (FACE). *BMC Plant Biology*, **11**(1), 123.

Ruan, C.-J., Shao, H.-B., and Teixeira da Silva, J. A. (2012). A critical review on the improvement of photosynthetic carbon assimilation in C3 plants using genetic engineering. *Critical Reviews in Biotechnology*, **32**(1), 1–21.

Ruban, A. A. V., Foyer, C. H., and Murchie, E. (2022). *Photosynthesis in Action : Harvesting Light, Generating Electrons, Fixing Carbon*. Academic Press, Amsterdam.

Ruben, S., Randall, M., Kamen, M., and Hyde, J. L. (1941). Heavy Oxygen (O18) as a Tracer in the Study of Photosynthesis. *Journal of the American Chemical Society*, **63**(3), 877–879.

Ruiz-Vera, U. M., Siebers, M., Gray, S. B., Drag, D. W., Rosenthal, D. M., Kimball, B. A., Ort, D. R., and Bernacchi, C. J. (2013). Global Warming Can Negate the Expected CO2 Stimulation in Photosynthesis and Productivity for Soybean Grown in the Midwestern United States. *Plant Physiology*, **162**(1), 410–423.

Ruiz-Vera, U. M., Acevedo-Siaca, L. G., Brown, K. L., Afamefule, C., Gherli, H., Simkin, A. J., Long, S. P., Lawson, T., and Raines, C. A. (2022). Field-grown ictB tobacco transformants show no difference in photosynthetic efficiency for biomass relative to the wild type. *Journal of Experimental Botany*, **73**(14), 4897–4907.

Ryu, Y., Berry, J. A., and Baldocchi, D. D. (2019). What is global photosynthesis? History, uncertainties and opportunities. *Remote Sensing of Environment*, **223**, 95–114.

Salisbury, F. B. (1991). Système Internationale: The Use of SI Units in Plant Physiology. *Journal of Plant Physiology*, **139**(1), 1–7.

Schauberger, B., Ben-Ari, T., Makowski, D., Kato, T., Kato, H., and Ciais, P. (2018). Yield trends, variability and stagnation analysis of major crops in France over more than a century. *Scientific Reports*, **8**(1), 16865.

Schils, R., Olesen, J. E., Kersebaum, K.-C., Rijk, B., Oberforster, M., Kalyada, V., Khitrykau, M., Gobin, A., Kirchev, H., Manolova, V., Manolov, I., Trnka, M., Hlavinka, P., Palosuo, T., Peltonen-Sainio, P., Jauhiainen, L., Lorgeou, J., Marrou, H., Danalatos, N., Archontoulis, S., Fodor, N., Spink, J., Roggero, P. P., Bassu, S., Pulina, A., Seehusen, T., Uhlen, A. K., Żyłowska, K., Nieróbca, A., Kozyra, J., Silva, J. V., Maçãs, B. M., Coutinho, J., Ion, V., Takáč, J., Mínguez, M. I., Eckersten, H., Levy, L., Herrera, J. M., Hiltbrunner, J., Kryvobok, O., Kryvoshein, O., Sylvester-Bradley, R., Kindred, D., Topp, C. F. E., Boogaard, H., de Groot, H., Lesschen, J. P., van Bussel, L., Wolf, J., Zijlstra, M., van Loon, M. P., and van Ittersum, M. K. (2018). Cereal yield gaps across Europe. *European Journal of Agronomy*, **101**, 109–120.

Schulze, E.-D. (1972). A new type of climatized gas exchange chamber for net photosynthesis and transpiration measurements in the field. *Oecologia*, **10**, 243–251.

Senebier, J. (1782). *Mémoires Physico-Chymiques: Sur l'influence de La Lumière Solaire Pour Modifier Les Êtres Des Trois Règnes de La Nature, et Sur Tout Ceux Du Règne Végétal*. Chez Barde, Manget & Compagnie.

Šesták, Z., Čatský, J., and Jarvis, P. G. P. G. (1971). *Plant Photosynthetic Production; Manual of Methods.* Junk, The Hague.

Sharma, M., Punya, and Gupta, B. B. (2020). Role of Wild Relatives for Development of Climate-Resilient Varieties. In R. K. Salgotra and S. M. Zargar, editors, *Rediscovery of Genetic and Genomic Resources for Future Food Security*, pages 303–314. Springer, Singapore.

Sharwood, R. E., Quick, W. P., Sargent, D., Estavillo, G. M., Silva-Perez, V., and Furbank, R. T. (2022). Mining for allelic gold: Finding genetic variation in photosynthetic traits in crops and wild relatives. *Journal of Experimental Botany*, **73**(10), 3085–3108.

Sharwood, R. R. E. (2023). *Understanding and Improving Crop Photosynthesis*. Burleigh Dodds Series in Agricultural Science ; Number 130. Burleigh Dodds Science Publishing, Cambridge, UK.

Sinclair, T. R. and Muchow, R. C. (1999). Radiation Use Efficiency. In D. L. Sparks, editor, *Advances in Agronomy*, volume 65, pages 215–265. Academic Press.

Sinclair, T. R., Rufty, T. W., and Lewis, R. S. (2019). Increasing Photosynthesis: Unlikely Solution For World Food Problem. *Trends in Plant Science*, **24**(11), 1032–1039.

Singsaas, E. L., Ort, D. R., and DeLucia, E. H. (2001). Variation in measured values of photosynthetic quantum yield in ecophysiological studies. *Oecologia*, **128**(1), 15–23.

Skillman, J. B. (2008). Quantum yield variation across the three pathways of photosynthesis: Not yet out of the dark. *Journal of Experimental Botany*, **59**(7), 1647–1661.

Skovmand, B., Reynolds, M. P., and Delacy, I. H. (2001a). Mining Wheat Germplasm Collections for Yield Enhancing Traits. In Z. Bedö and L. Láng, editors, *Wheat in a Global Environment: Proceedings of the 6th International Wheat Conference, 5–9 June 2000, Budapest, Hungary*, Developments in Plant Breeding, pages 761–771. Springer Netherlands, Dordrecht.

Skovmand, B., Reynolds, M. P., and Delacy, I. H. (2001b). Searching Genetic Resources for Physiological Traits with Potential for Increasing Yield. In J. I. Ortiz - Monasterio, A. MacNab, and M. P. Reynolds, editors, *Application of Physiology in Wheat Breeding*. CIMMYT, Mexico.

Slack, CR. and Hatch, MD. (1967). Comparative studies on the activity of carboxylases and other enzymes in relation to the new pathway of photosynthetic carbon dioxide fixation in tropical grasses. *Biochemical Journal*, **103**(3), 660–665.

Song, T., Yang, F., Das, D., Chen, M., Hu, Q., Tian, Y., Cheng, C., Liu, Y., and Zhang, J. (2020). Transcriptomic analysis of photosynthesis-related genes regulated by alternate wetting and drying irrigation in flag leaves of rice. *Food and Energy Security*, **9**(3), e221.

South, P. F., Cavanagh, A. P., Liu, H. W., and Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, **363**(6422), eaat9077.

Steffen, 1947-2023., W. L. W. L. (2005). *Global Change and the Earth System : A Planet under Pressure*. Global Change–the IGBP Series. Springer, Berlin ;.

Stirbet, A., Lazár, D., Guo, Y., and Govindjee, G. (2020). Photosynthesis: Basics, history and modelling. *Annals of Botany*, **126**(4), 511–537.

Szymańska, R., Ślesak, I., Orzechowska, A., and Kruk, J. (2017). Physiological and biochemical responses to high light and temperature stress in plants. *Environmental and Experimental Botany*, **139**, 165–177.

Taiz, L., Zeiger, E., Møller, I., and Murphy, A. (2015). *Plant Physiology and Development*. Sinauer Associates, Incorporated, Publishers.

Tao, Y., Zhao, X., Mace, E., Henry, R., and Jordan, D. (2019). Exploring and exploiting pan-genomics for crop improvement. *Molecular Plant*, **12**(2), 156–169.

Taylor, G., Garassino, F., Aarts, M. G. M., and Harbinson, J. (2022). Improving C3 photosynthesis by exploiting natural genetic variation: Hirschfeldia incana as a model species. *Food and Energy Security*, **n/a**(e420).

Taylor, S. H., Orr, D. J., Carmo-Silva, E., and Long, S. P. (2020). During photosynthetic induction, biochemical and stomatal limitations differ between Brassica crops. *Plant, Cell & Environment*, **43**(11), 2623–2636.

Theeuwen, T. P. J. M., Logie, L. L., Harbinson, J., and Aarts, M. G. M. (2022). Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany*, **73**(10), 3122–3137.

Tian, X., Engel, B. A., Qian, H., Hua, E., Sun, S., and Wang, Y. (2021). Will reaching the maximum achievable yield potential meet future global food demand? *Journal of Cleaner Production*, **294**, 126285.

Tkemaladze, G. S. and Makhashvili, K. A. (2016). Climate changes and photosynthesis. *Annals of Agrarian Science*, **14**(2), 119–126.

Triesch, S., Denton, A. K., Bouvier, J. W., Buchmann, J. P., Reichel-Deland, V., Martins Guerreiro, R. N. F., Busch, N., Schlüter, U., Stich, B., and Kelly, S. (2022). Transposable elements contribute to the establishment of the glycine shuttle in Brassicaceae species. *bioRxiv*, page 2022.12. 06.519256.

van Bezouw, R. F. H. M., Keurentjes, J. J. B., Harbinson, J., and Aarts, M. G. M. (2019). Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. *The Plant Journal*, **97**(1), 112–133.

Van Helmont, J. B. (1652). *Ortus Medicinae, Id Est Initia Physicae Inaudita: Progressus Medicinae Novus, In Morborum Ultionem Ad Vitam Longam*. Sumptibus Joan Ant. Huguetan & Guillielmi Barbier.

Van Iersel, M. W., Dove, S., Kang, J.-G., and Burnett, S. E. (2010). Growth and Water Use of Petunia as Affected by Substrate Water Content and Daily Light Integral. *HortScience*, **45**(2), 277–282.

van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z. (2013). Yield gap analysis with local to global relevance—A review. *Field Crops Research*, **143**, 4–17.

van Rooijen, R., Harbinson, J., and Aarts, M. G. M. (2018). Photosynthetic response to increased irradiance correlates to variation in transcriptional response of lipid-remodeling and heat-shock genes. *Plant Direct*, **2**(7), e00069.

Verslues, P. E., Bailey-Serres, J., Brodersen, C., Buckley, T. N., Conti, L., Christmann, A., Dinneny, J. R., Grill, E., Hayes, S., Heckman, R. W., Hsu, P.-K., Juenger, T. E., Mas, P., Munnik, T., Nelissen, H., Sack, L., Schroeder, J. I., Testerink, C., Tyerman, S. D., Umezawa, T., and Wigge, P. A. (2023). Burning questions for a warming and changing world: 15 unknowns in plant abiotic stress. *The Plant Cell*, **35**(1), 67–108.

von Mayer, J. R. (1845). *Die Organische Bewegung in Ihrem Zusammenhange Mit Dem Stoffwechsel*. C. Drechsler.

Wang, D., Mu, Y., Hu, X., Ma, B., Wang, Z., Zhu, L., Xu, J., Huang, C., and Pan, Y. (2021). Comparative proteomic analysis reveals that the Heterosis of two maize hybrids is related to enhancement of stress response and photosynthesis respectively. *BMC Plant Biology*, **21**(1), 34.

Wang, L.-M., Shen, B.-R., Li, B.-D., Zhang, C.-L., Lin, M., Tong, P.-P., Cui, L.-L., Zhang, Z.-S., and Peng, X.-X. (2020). A synthetic photorespiratory shortcut enhances photosynthesis to boost biomass and grain yield in rice. *Molecular Plant*, **13**(12), 1802–1815.

Way, D. A., Becklin, K., and Ward, J. (2021). Leaf Carbon Flux Responses to Climate Change: Challenges and Opportunities. In K. M. Becklin, J. K. Ward, and D. A. Way, editors, *Photosynthesis, Respiration, and Climate Change*, pages 3–13. Springer International Publishing, Cham.

Willey, N. (2016). *Environmental Plant Physiology*. Garland Science, Taylor & Francis Group, New York.

Willstätter, R. (1915). Chlorophyll. *Journal of the American Chemical Society*, **37**(2), 323–345.

Wu, A., Hammer, G. L., Doherty, A., von Caemmerer, S., and Farquhar, G. D. (2019). Quantifying impacts of enhancing photosynthesis on crop yield. *Nature Plants*, **5**(4), 380–388.

Wu, A., Brider, J., Busch, F. A., Chen, M., Chenu, K., Clarke, V. C., Collins, B., Ermakova, M., Evans, J. R., Farquhar, G. D., Forster, B., Furbank, R. T., Groszmann, M., Hernandez-Prieto, M. A., Long, B. M., Mclean, G., Potgieter, A., Price, G. D., Sharwood, R. E., Stower, M., van Oosterom, E., von Caemmerer, S., Whitney, S. M., and Hammer, G. L. (2023). A cross-scale analysis to understand and quantify the effects of photosynthetic enhancement on crop growth and yield across environments. *Plant, Cell & Environment*, **46**(1), 23–44.

Xiong, B., Qiu, X., Huang, S., Wang, X., Zhang, X., Dong, T., Wang, T., Li, S., Sun, G., Zhu, J., and Wang, Z. (2019). Physiological and transcriptome analyses of photosynthesis and chlorophyll metabolism in variegated Citrus (Shiranuhi and Huangguogan) seedlings. *Scientific Reports*, **9**(1), 15670.

Yin, X., Sun, Z., Struik, P. C., and Gu, J. (2011). Evaluating a new method to estimate the rate of leaf respiration in the light by analysis of combined gas exchange and chlorophyll fluorescence measurements. *Journal of Experimental Botany*, **62**(10), 3489–3499.

Yoon, D.-K., Ishiyama, K., Suganami, M., Tazoe, Y., Watanabe, M., Imaruoka, S., Ogura, M., Ishida, H., Suzuki, Y., Obara, M., Mae, T., and Makino, A. (2020). Transgenic rice overproducing Rubisco exhibits increased yields with improved nitrogen-use efficiency in an experimental paddy field. *Nature Food*, **1**(2), 134–139.

Zanini, S. F., Bayer, P. E., Wells, R., Snowdon, R. J., Batley, J., Varshney, R. K., Nguyen, H. T., Edwards, D., and Golicz, A. A. (2022). Pangenomics in crop improvement—from coding structural variations to finding regulatory variants with pangenome graphs. *The Plant Genome*, **15**(1), e20177.

Zenda, T., Wang, N., Yan, X., Dong, A., Yang, Q., Zhong, Y., and Duan, H. (2023). Opportunities and avenues for achieving crop climate resilience. *Environmental and Experimental Botany*, **213**, 105414.

Zeng-ping, T., Xiu-zhen, L., Wei-juan, C., and Zhao-ying, Y. (1995). Reprobing into rice breeding for high photosynthetic efficiency. *Journal of Integrative Plant Biology*, **37**(8).

Zhang, F. and Batley, J. (2020). Exploring the application of wild species for crop improvement in a changing climate. *Current Opinion in Plant Biology*, **56**, 218–222.

Zhu, S., Xiao, J., Han, S., Li, X., Li, Z., Wei, B., Zhang, D., Wang, R., Li, R., Yang, L., and Li, S. (2023). Transcriptomics Combined with Photosynthetic Physiology and Leaf Structure Analysis Revealed Increased Sugarcane Yield by Fenlong-Ridging. *Agronomy*, **13**(5), 1196.

Zhu, X.-G., de Sturler, E., and Long, S. P. (2007). Optimizing the Distribution of Resources between Enzymes of Carbon Metabolism Can Dramatically Increase Photosynthetic Rate: A Numerical Simulation Using an Evolutionary Algorithm. *Plant Physiology*, **145**(2), 513–526.

Zhu, X.-G., Long, S. P., and Ort, D. R. (2010). Improving Photosynthetic Efficiency for Greater Yield. *Annual Review of Plant Biology*, **61**(1), 235–261.

Zhu, X.-G., Hasanuzzaman, M., Jajoo, A., Lawson, T., Lin, R., Liu, C.-M., Liu, L.-N., Liu, Z., Lu, C., Moustakas, M., Roach, T., Song, Q., Yin, X., and Zhang, W. (2022). Improving photosynthesis through multidisciplinary efforts: The next frontier of photosynthesis research. *Frontiers in Plant Science*, **13**.

Zimmer, D., Swart, C., Graf, A., Arrivault, S., Tillich, M., Proost, S., Nikoloski, Z., Stitt, M., Bock, R., Mühlhaus, T., and Boulouis, A. (2021). Topology of the redox network during induction of photosynthesis as revealed by time-resolved proteomics in tobacco. *Science Advances*, **7**(51), eabi8307.

CHAPTER 2

Improving C_3 photosynthesis by exploiting natural genetic variation: *Hirschfeldia incana* as a model species

Food and Energy Security, 2023, 12, e420.

Graham Taylor[1,4,*], Francesco Garassino[2,*], Mark G.M. Aarts[2], Jeremy Harbinson[3]

(1) Horticulture and Product Physiology Group, Wageningen University and Research, Wageningen (NL)

(2) Laboratory of Genetics, Wageningen University and Research, Wageningen (NL)

(3) Laboratory of Biophysics, Wageningen University and Research, Wageningen (NL)

(4) Current affiliation: Optimal Labs Limited, London (UK)

(*) These authors contributed equally to this work

Abstract

Despite research efforts toward improving crop photosynthetic energy conversion efficiencies over the past 40 years, photosynthetic efficiencies remain far from their theoretical maxima. A major challenge has been that plant photosynthesis is a complex process, controlled by many underlying genetic factors and highly dynamic in response to short-term environmental changes. Recent approaches to improving photosynthesis involved model-based identification of the bottlenecks in photosynthesis followed by their genetic modification (GM). While these approaches were successful and inspirational, their dependency on the use of GM techniques may restrict their implementation in some jurisdictions. We therefore suggest greater research focus on a different, yet complementary, approach to improving photosynthetic efficiency: the exploration and exploitation of natural genetic variation in photosynthesis. A substantial improvement in phenotyping and genotyping technology over the past decade has highlighted natural variation in photosynthetic sub-traits for crop and model species. However, a comprehensive understanding of all the factors responsible for photosynthetic limitations is still lacking. We therefore propose the use of high photosynthetic capacity species as models for the exploration of the physiological and genetic basis of high photosynthetic efficiency. While most high photosynthetic capacity species are not suitable as models due to complex genetics and evolutionary distance from crops, we have identified Brassicaceae species *Hirschfeldia incana* (L.) Lagr.-Foss as a promising candidate. In this perspective paper, we describe and advocate the use of *H. incana* as a model for the exploration of high maximum CO_2 assimilation rates (P_{max}) found in some C_3 species. We describe the basic biology and evolutionary history of the species and report preliminary data on its photosynthetic characteristics. Our findings suggest *H. incana* is an excellent model species for studies aiming at understanding natural genetic variation in photosynthetic efficiency.

2.1 Introduction

Improving crop plant photosynthesis was first proposed about 40 years ago (Austin, 1989) but since then progress in realizing this goal has been limited. A major hurdle to progress has been the genetic and physiological complexity of photosynthesis and the difficulty of phenotyping for photosynthesis on a large scale. High-throughput photosynthetic phenotyping has been a challenge, creating one of the major bottlenecks to progress in breeding for better photosynthesis. Only in the last decade has the development of large-scale, high-throughput photosynthetic phenotyping technologies and advances in genotyping techniques allowed researchers to begin to understand the genetics underlying variation in photosynthetic traits (van Bezouw *et al.*, 2019). Improvements in portable gas analysis systems now allow the routine measurement of assimilation in the field, while high-throughput phenotyping techniques—usually based on chlorophyll fluorescence—are now available for mass measurement. Here, "high-throughput" means taking a measurement of a photosynthetic property on several hundred plants per species multiple times per day rather than in the order of 10 to 100 plants per day which is typical when using portable gas analysis systems. These technological improvements in phenotyping have depended on our enhanced understanding of photosynthetic physiology, particularly in the use of chlorophyll fluorescence-derived parameters. More generally, there has also been an increase in larger scale research into the genetics and scope of natural variation of photosynthesis of the kind needed to exploit natural variation in photosynthesis as a path to crop yield improvement (Dann and Leister, 2017; Flood *et al.*, 2011; Theeuwen *et al.*, 2022). Evidence of natural variation in photosynthetic subtraits has now been reported for major staple crops such as wheat (Driever *et al.*, 2014; Molero and Reynolds, 2020), rice (Acevedo-Siaca *et al.*, 2020a,b; Gu *et al.*, 2014; Lin *et al.*, 2018; Wang *et al.*, 2015), maize (Strigens *et al.*, 2013), soybean (Burgess *et al.*, 2020; Gilbert *et al.*, 2011; Lopez *et al.*, 2019), potato (Prinzenberg *et al.*, 2018; Xu *et al.*, 2021), and sorghum (Ortiz *et al.*, 2017), as well as for the model species *Arabidopsis thaliana* (Prinzenberg *et al.*, 2020; van Rooijen *et al.*, 2015, 2017). Even for tree species, such as *Fagus sylvatica* (Aranda *et al.*, 2014), whose photosynthetic properties are often less intensively studied than those of crop plants and model species, there is growing interest in the natural variation of their photosynthetic properties.

Photosynthesis is a complex process with many potential underlying genetic factors, and much remains to be discovered about the naturally occurring variability of the process (van Bezouw *et al.*, 2019). In particular, photosynthesis has been extensively studied and understood as a steady-state process, while the kinetics of the photosynthetic response to a short-term change in the environment has been largely ignored (with some exceptions, e.g., Harbinson and Woodward (1984); Pearcy (1990)). Recently, it has been more widely recognized that slow, limiting responses to environmental fluctuations over various timescales (typically ¿1 s) can result in sub-optimal CO2 assimilation; this is particularly relevant for photosynthesis in the field, where it is now recognized that photosynthesis may only rarely be at steady-state (Lawson *et al.*, 2012). The desire to increase future crop yield sustainably further complic-

ates the challenge of yield improvement, so in addition to the obvious importance of photosynthetic light-use efficiency (LUE or quantum yield), the efficient use of water (Franks *et al.*, 2015; Leakey *et al.*, 2019; Richards *et al.*, 1993), nitrogen (Hirel *et al.*, 2007; Raun and Johnson, 1999; Swarbreck *et al.*, 2019), phosphorous (Heuer *et al.*, 2017; Veneklaas *et al.*, 2012), and micronutrients (Dimkpa and Bindraban, 2016; Fageria *et al.*, 2008; Monreal *et al.*, 2016; Welch and Graham, 2002) for photosynthesis are increasingly recognized as important and genetically complex traits influencing crop production sustainability. Inevitably, traits connected to the efficiencies of nutrient and water use will come into greater focus as part of improving crop yields.

2.2 Photosynthetic improvement potential

As a process in vascular plants, oxygenic photosynthesis is relatively inefficient (Cardona *et al.*, 2018). This inefficiency is a major factor contributing to a realized yield productivity below the theoretical maximum in crop species (Zhu *et al.*, 2010). The energy conversion efficiency of C_3 and C_4 pathways—including that of bioenergy crops—is only 2.4% (4.6% theoretical maximum) and 3.7% (6% theoretical maximum) based on the total incident solar radiation intercepted by a leaf canopy (Beadle and Long, 1985; Monteith, 1977; Piedade *et al.*, 1991; Zhu *et al.*, 2008). This photosynthetic inefficiency, while highlighting productivity gaps between realized and theoretical limits, also highlights the scope for future research toward improving photosynthesis potential. Given these gaps, it is increasingly argued that one of the best means available to allow the World food supply to meet the projected increases in global food demand is through the improvement of the photosynthetic performances of crop plants (Long *et al.*, 2006; Murchie *et al.*, 2009; Parry *et al.*, 2011; Zhu *et al.*, 2008, 2010). This will aid with meeting increased future demand for food implied by a global population increase alongside economic growth (Tilman *et al.*, 2002), which remains a significant challenge; it is expected that crop yields will need to increase by up to 110% (Tilman *et al.*, 2002; Zhu *et al.*, 2010) and this should be achieved without further expansion of the world's land area dedicated to agricultural production otherwise we risk further loss of biodiversity.

Improving photosynthesis as a means to improve crop yield was first considered several decades ago (Gifford and Evans, 1981; Zelitch, 1975), but at that time the technical and scientific barriers to achieving this goal were insurmountable and the idea essentially lapsed. During the last two decades, however, the idea that improvements to photosynthesis will be our best biological option for significantly improving crop yields has come to the fore (Evans, 2013; Lawson *et al.*, 2012; von Caemmerer and Evans, 2010; Zhu *et al.*, 2010). In 2012, the "Realizing Increased Photosynthetic Efficiency" (ripe.illinois.edu) project initiative was formed, demonstrating an important commitment to this endeavor. While there has been some scientific debate on whether improved leaf-level C3 photosynthesis would even translate into yield increases (Sinclair *et al.*, 2004, 2019), evidence from CO_2 enrichment studies and genetic modification (GM) approaches support the idea that higher photosynthetic

efficiency will translate to better yields (Long *et al.*, 2006; Mitchell and Sheehy, 2006; Sheehy *et al.*, 2008; von Caemmerer and Evans, 2010). Additionally, increased atmospheric CO_2 mole fraction or supplementary lighting can achieve improved photosynthesis in protected cultivation, such as greenhouses, resulting in higher yields. With field-grown crop plants, the scope for environmental modification to increase yield is limited and therefore significant crop yield gains are likely to be more dependent on genetic approaches. This implies that we need to find or create genetic diversity that is coupled with photosynthesis variation.

2.3 How to improve photosynthesis

Recent approaches to improving photosynthesis have been guided by model-based identification of the bottlenecks in photosynthesis (Poolman *et al.*, 2000; South *et al.*, 2019; Zhu *et al.*, 2010). These approaches have been successful and inspirational, but they depend on the use of GM techniques (Driever *et al.*, 2017; Kromdijk *et al.*, 2016; López-Calcagno *et al.*, 2020; Simkin *et al.*, 2015; South *et al.*, 2019). The regulatory framework on the use of GM and similar novel plant breeding techniques (NPBT), however, can be restrictive and the commercial cultivation of NBPT-based cultivars has been highly limited in some jurisdictions such as the European Union. For these jurisdictions, an alternative to NPBT-based crop improvement is needed, such as conventional plant breeding exploiting naturally occurring trait variation. In addition to regulatory issues, GM (and similar) normally depends on an approach involving the intelligent creation, redesign, or modification (i.e., engineering) of photosynthetic traits. In general, this approach requires a thorough knowledge of the physiological or developmental pathway targeted for improvement. The NPBT route to improvement is therefore likely to be limited because many photosynthetic traits are not yet sufficiently understood for NPBT approaches to be able to be applied successfully.

Alongside GM research, we therefore suggest greater research focus on a different, yet complementary, approach to improving photosynthetic efficiency: the exploration and exploitation of natural genetic variation in C_3 photosynthesis. The natural, intraspecies variation for photosynthetic traits in species that can be crossed with crop plants could be used to improve crop photosynthesis via conventional plant breeding routes, resulting in better photosynthetic assimilation rates, and, ultimately, a higher yield potential (Lawson *et al.*, 2012; Zhu *et al.*, 2010). Making use of this variation in conventional breeding is more efficient if the genes or quantitative trait loci (QTLs) underpinning the variation for photosynthetic traits have been identified. Such genes can be further studied to understand the biology explaining the variation and can become a target for marker-assisted breeding, gene editing or GM approaches. On a broader scale, the physiological diversity of photosynthesis within plants and algae represents a collection of templates or models for particular traits, such as a very high maximum CO_2 assimilation rate or exceptional qE (an important photoprotective mechanism acting in Photosystem II, *PSII*). Unraveling the physiological and genetic bases of exceptional traits is a route to building more energy-efficient

photosynthesis, and assists in understanding the physiological and genetic limits of genotypes with phenotypic deficiencies, such as normal maximum CO_2 assimilation rates or normal qE.

2.4 Leaves, photosynthesis, and traits

The leaf is the organizational level of the plant at which the properties, or traits, of photosynthesis, and particularly the environmental dependency of these traits, are commonly assessed. These leaf-level properties arise from an underlying, highly coordinated developmental process that builds and maintains the photosynthetic machine embedded in the leaf mesophyll cells. This process also gives rise to larger-scale leaf-structural properties that provide the diffusive pathways for efficient gas exchange, and the vascular tissues essential for the transport of water, nutrients, and assimilates. Photosynthesis is a highly flexible process that is strongly influenced in the short to long term by the environment in ways and extents that depend on genetic differences between individuals and species. This genetic control of photosynthetic traits, however, is generally not understood despite the thoroughness with which variation in photosynthetic traits is sometimes understood at the physiological level.

Maximum CO_2 assimilation rate (P_{max}) is a good example of a leaf-level photosynthetic trait. In vascular plants, P_{max} depends on a range of other sub-traits, such as high rates of electron transport and photosynthetic metabolism, and high stomatal and mesophyll conductances for CO_2 diffusion. These broad sub-traits can themselves be further refined into yet more specific sub-traits and the scaling from the cellular level of proteins and membranes to the whole leaf is reasonably well understood. However, the genetic basis for P_{max}—its genetic algorithm—is not known despite this trait being relatively well understood from a photosynthetic physiology perspective. Increasing P_{max} through marker-assisted selection breeding, therefore, is not currently feasible (although breeding based on phenotype selection alone is possible). Identifying the genes responsible for variation in P_{max} would make marker-assisted breeding possible, and would also make available targets for GM or gene editing approaches.

Apart from its value in breeding, understanding the genetic basis of variation in photosynthetic traits is also a way to understand the developmental process that builds the photosynthetic machinery and delivers the diverse range of photosynthetic phenotypes encountered in the natural world. The evolution of photosynthetic traits is recorded in these—as yet largely unknown—genetic underpinnings of photosynthetic variation. Understanding the genetic basis of natural variation in plant photosynthesis, therefore would also give us access to the evolutionary history of photosynthetic variation in plants, and to the adaptive mechanisms that allow plants to successfully occupy specific environmental niches (Flood, 2019; Flood *et al.*, 2011). This knowledge is of particular significance for agriculture given the extent and rate of environmental change and the rapidly growing human population that demands crop plant phenotypes to be creatively adapted.

Technological advances in photosynthetic measurement equipment and techniques have assisted in our understanding of photosynthetic physiology and its variation by allowing for highly integrated and detailed measurement of photosynthesis, including deep-phenotyping and high-throughput phenotyping. Deep phenotyping makes possible the analysis and identification of photosynthetic bottlenecks under diverse environmental/experimental conditions through the combined application of a range of destructive and non-destructive measurement techniques. The same chlorophyll fluorescence measurements that are often used in deep-phenotyping strategies are typically used in high-throughput phenotyping for photosynthetic traits (Flood *et al.*, 2016, 2020; Furbank and Tester, 2011). A high-throughput phenotyping approach can therefore provide the extensive data needed for the identification of the QTLs contributing to variation in trait properties. These phenotyping approaches, together with advances in genomics, make it possible to genetically compare related species with differing photosynthetic traits, and to understand the changes leading to different photosynthetic phenotypes. What remains is the identification of model species on which to effectively analyze complex physiological traits like photosynthesis based on trait function and genetics.

2.5 Photosynthetic limitation

The photosynthesis-irradiance response provides a useful handle to understand and quantify multiple limitations acting on a leaf. When crops and other plant species are exposed to increasing irradiances, leaf-level photosynthetic rates tend to become light-saturated well below the typical, full-sunlight levels of 2000 µmol m^{-2} s^{-1} or more encountered under natural conditions (Figure 2.1) (Gitelson *et al.*, 2015; Gu *et al.*, 2017; Monneveux *et al.*, 2003; Murchie *et al.*, 1999; Turner *et al.*, 2003). At strictly light-limiting irradiances, photosynthetic LUE is maximal; LUE is defined here as the ratio between gross photosynthesis (A_{gross}, µmol m^{-2} s^{-1}) and absorbed or incident photosynthetic photon flux density (µmol m^{-2} s^{-1}). Maximum, light-limited, photosynthetic efficiency is a complex trait and varies, inter alia, with the wavelengths of light used, photosynthetic metabolism (especially the activity of photorespiration Ehleringer and Björkman (1977)), state-transitions the presence of non-photosynthetic pigments (Hogewoning *et al.*, 2012), and photodamage to photosystems I and II (e.g., Kao and Forseth (1992); Sonoike (2011)).

At irradiances above the strictly light-limited range (for a typical C_3 crop plant, up to about 100 - 200 µmol m^{-2} s^{-1}) other physiological and metabolic limitations become significant. These limitations result in a decline in overall LUE from the light-limited maximum, and ultimately in the light saturation of photosynthesis. An implication of this is that in the absence of these constraints the photosynthesis–irradiance relationship would be linear with a gradient equal to the light-limited slope of the photosynthesis–irradiance relationship (Figure 2.1). Although the underlying biochemical components and mechanisms of achieving photosynthesis are highly conserved among C_3 species, the light-saturated assimilation rates arising from the loss of photosynthetic LUE with increasing irradiance can vary considerably among genotypes

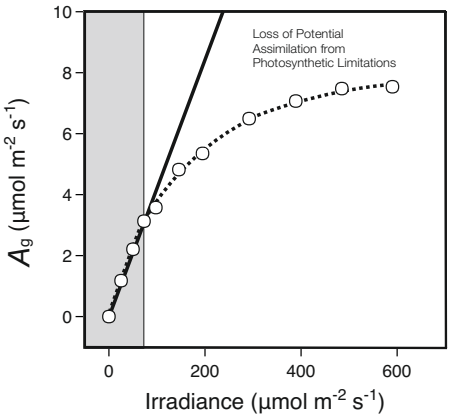

Figure 2.1: A non-rectangular hyperbola light response curve of gross photosynthetic rate (A_g) with increasing irradiance for a typical, low-light adapted C_3 plant leaf. "Low light" here refers to an irradiance of 200 µmol m^{-2} s^{-1}. The range of irradiances shaded in grey is where photosynthesis is significantly light-limited. The region between the extrapolation of the light-limited phase (solid line) and the assimilation response (dotted dash response curve fit estimate) is where photosynthetic efficiency declines. Understanding this natural variation in efficiency loss at saturating irradiances is key to providing research handles for future genetic research.

and species. While a higher P_{max} arises mainly from the ability to sustain a higher LUE at very high irradiances, this high LUE at high, near light saturating irradiances will be associated with higher light-use efficiencies and therefore assimilation rates at lower irradiances above the strictly light-limited region of the photosynthetic light-response curve (Harbinson and Yin, 2017). If the high P_{max}, high LUE of high light-adapted leaves is paralleled by a relatively high P_{max}, and thus a still high LUE, following adaptation to lower irradiances found deeper in the canopy, the high P_{max} phenotype will benefit not only the assimilation in those leaves exposed to a full irradiance but also those in more shaded parts of the canopy.

A sustained high rate of photosynthesis, which implies a high LUE for photosynthesis at high irradiances, requires that all of the supply and demand side processes can support high fluxes. A high light-saturated C_3 photosynthetic rate, in terms of a CO_2 fixation rate, is usually around 30 µmol m^{-2} s^{-1}, while C_4 leaves typically achieve rates of 30 - 60 µmol m^{-2} s^{-1}. Some C_3 leaves are capable of assimilation rates of up to 65 µmol m^{-2} s^{-1}, which is better than most C_4 crops, but how these leaves

achieve such high assimilation rates in terms of either their physiology or genetics is not clear. The potential flux limitations to steady-state C_3 leaf photosynthesis are numerous but can be summarized into four main groups of processes. Assuming that the product of photosynthesis is mesophyll cell carbohydrate, these four groups can be further split into three supply-side processes and one demand-side process. The main supply-side processes are the diffusive transport of CO_2 via the stomata and mesophyll conductive paths to the site of fixation; the light reactions with electron transport providing the reductant (NADPH) and ATP for the metabolic processes of assimilation; and photosynthetic metabolism which fixes CO_2 and produces carbohydrate. The demand-side process is the transport of carbohydrates to the sink tissues and their use by the sinks. Leaves with a high photosynthetic capacity must be able to develop high conductances for the gaseous diffusion of CO_2, a high metabolic capacity, and high rates of electron transport on the supply side, along with high carbohydrate transport and sink activity to allow for sustained high photosynthetic LUE at high irradiances. High photosynthetic capacity exemplar leaves offer models within which to explore the physiological and genetic basis of high P_{max}.

2.6 Naturally occurring high-photosynthetic capacity species

A. thaliana became the model C_3 plant species for plant research over the last 40 years mainly because it has many desirable characteristics for experimentation. These include a short regeneration time, a small plant size, a limited need for growth facilities, prolific self-fertilization, and an increasingly multidisciplinary research environment (Koornneef and Meinke, 2010). Despite the well-documented merits and benefits of research focused on *A. thaliana*, there have been concerns that the focus on *A. thaliana* may lose impetus as plant science research funding and community interest in the species decline (Koornneef and Meinke, 2010). Wild-type *A. thaliana* is of limited applicability in research directed toward enhancing C_3 photosynthesis by increased photosynthetic capacity because of its unexceptional photosynthetic capacity; light-saturated assimilation rates of 6.5 - 12 µmol m^{-2} s^{-1} have been reported for *A. thaliana* (Kaiser *et al.*, 2016; Tanaka *et al.*, 2013) grown at the low growth irradiances (100 - 200 µmol m^{-2} s^{-1}) typically used for this species under controlled environment conditions. When grown at high irradiances (1800 µmol m^{-2} s^{-1}) *A. thaliana* can achieve assimilation rates of 30 µmol m^{-2} s^{-1} (Garassino *et al.*, 2022), which while high when compared to results obtained from this species under commonly used irradiances, are not exceptional compared to some other C_3 species. Evidently, although *A. thaliana* is an excellent and well-researched general C_3 plant model, it lacks the photosynthetic properties that would make it a candidate model for exceptional P_{max}.

Among C_3 plants, sustained high LUE at high irradiances, resulting in a P_{max} of over 40 µmol m^{-2} s^{-1}, have only been reported for a few species. The majority of these are annuals endemic to desert environments of the South-Western USA,

where a high P_{max} is attributed to an adaptation to very short growing seasons in a high-light environment (Seemann *et al.*, 1980). (Nobel, 1991) calculated a theoretical maximum net photosynthetic rate for C_3 plants of about 55 µmol m^{-2} s^{-1} at an irradiance of 2000 µmol m^{-2} s^{-1}. (Werk *et al.*, 1983), however, reported the highest photosynthetic capacity for a C_3 plant of 65 µmol m^{-2} s^{-1} by *Palafoxia linearis*, a winter desert annual, grown in a controlled environment. The C_3 winter desert annuals *Camissonia claviformis* and *Camissonia brevipes* (currently part of the *Chylismia* genus) achieve an assimilation rate of about 60 µmol m^{-2} s^{-1} at an irradiance of 2000 µmol m^{-2} s^{-1} (Ehleringer *et al.*, 1979; Longstreth *et al.*, 1980; Seemann *et al.*, 1980). Other C_3 species have been reported to have high assimilation values of 42 – 55 µmol m^{-2} s^{-1} (Nobel, 1991) at an irradiance of 2000 µmol m^{-2} s^{-1}; even at this high irradiance, these assimilation rates were not fully light-saturated. Light-saturated leaf photosynthetic assimilation rates for a typical cereal crop, for example, wheat, are only 20 – 30 µmol m^{-2} s^{-1} under physiologically favorable conditions. The high photosynthetic capacities of the C_3 desert annual species referred to are comparable to, and even exceed, those of many C_4 crop species, which typically produce biomass more efficiently than C_3 crop plants on an intercepted light basis (von Caemmerer and Evans, 2010). The C_4 species *Amaranthus palmeri* (considered a problem weed in some regions), however, can achieve an assimilation rate of nearly 80 µmol m^{-2} s^{-1} when measured at an irradiance of 80 µmol m^{-2} s^{-1}, an optimal temperature of 42 ℃, and at natural ambient CO_2 concentrations (Ehleringer, 1983) - the highest known photosynthetic assimilation rate for any plant species (although this claim must be tempered by the lack of any extensive, focused effort to find better performing species). Despite the high assimilation rate potential revealed by these C_3 and C_4 species, many of the high photosynthetic capacity species mentioned here would not qualify as good experimental plant models. Good model plants demand simple genetics and a close genetic distance from major crop plant species or another genetically well-researched species; for example, neither *A. palmeri* (Amaranthaceae) nor *Camissionia*/*Chylismia* species (Onagraceae), share close phylogenetic relations with major commercial crops. Furthermore, they have scarce curated germplasm and are poorly studied genetically. To understand the genetic and physiological basis of high LUE at high irradiances, a model species that has high assimilation rates at high irradiance and is genetically close to crop or model plant species is therefore needed. *Hirschfeldia incana* appears to offer that special combination of physiological and genetic properties that support its candidacy as a model species. We need to recognize however that while *H. incana* may be the most convenient, suitable model species for the very high photosynthetic rate phenomenon, we should continue comparing this with other species that also have very high photosynthetic rates; there may be different routes to the same end.

2.7 *Hirschfeldia incana*

H. incana (L.) Lagr.-Foss. (Figure 2.2) is a nitrophilous and thermophilous winter annual or biennial C_3 species. It is the sole member of a monophyletic genus and

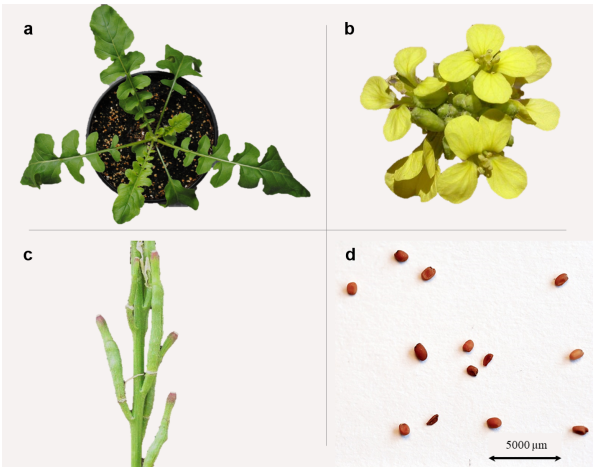

Figure 2.2: The morphology of *Hirschfeldia incana*. (a) Top view of a four weeks-old plant grown under artificial lighting in a controlled indoor environment, (b) typical inflorescence, (c) maturing siliques, (d) micrograph of ripe seeds.

is native to the Mediterranean basin and the Irano-Turanian floristic region (Manafzadeh *et al.*, 2017). Due to introductions elsewhere, the distribution of *H. incana* extends to cool and warm-temperate climate areas globally (Siemens, 2011), with wild populations concentrated in the sub-tropical to temperate regions of the world (Figure 2.3).

In these regions, *H. incana* occurs as a pioneer species in colonies often of only a few individuals or as solitary plants on disturbed sites such as roadsides and ditches (DiTomaso and Healy, 2007). It shows rapid, sustained growth over a short period, and is commonly seen as a weedy, invasive species (Lee *et al.*, 2004). *H. incana* is a prolific bearer of seeds, which are contained in numerous siliques. About 3–5 seeds per locule develop along many long racemes (DiTomaso *et al.*, 2013). Seeds obtained from a wild population at Oued El Himer in Eastern Morocco were 1.16 ±0.05 mm in length and 0.86 ±0.03 mm in width (n = 10, unpublished observation), and seeds are generally highly viable at maturation (Gresta *et al.*, 2010), germinating rapidly within a few days without requiring any seed pre-treatment (unpublished observation). Rapid germination was reported at 25–30 ℃ in light conditions and 20–25 ℃ in darkness by (Gresta *et al.*, 2010), while we have found high germination rates in fresh seed material exceeding 90% at 20 ℃ in darkness and 23 ℃ in light. Seeds of *H. incana* store well, consistent with its classification as a weedy species; when stored under non-ideal seed storage conditions of 23 ℃ and at a relative humidity of 50%, we found most *H. incana* seeds remained viable for more than 5 years. Under favorable growth conditions, *H. incana* has a relatively short reproductive cycle, completing germination to flowering in about 5–8 weeks. The maturation of the first seed set occurs approximately 3–4 weeks after flowering. Although *H. incana*

Figure 2.3: The global distribution of wild populations reported for *Hirschfeldia incana* (various sources, see Quiroz (2015)). Orange markers indicate approximate locations where at least one individual has been reported in the wild.

is considered self-incompatible due to protogyny (Al-Shehbaz, 1977), self-fertilization has been reported (Lee *et al.*, 2004) and self-fertilized progeny can be produced with hand pollination. Seedlings of *H. incana* establish and grow well in pots under neutral white LEDs or fluorescent lighting commonly used in growth chambers, and they thrive under warm conditions (Chronopoulos *et al.*, 2005).

As a member of the Brassicaceae, *H. incana* is a close relative of economically important vegetable and oil-seed crop species (FAO, 2022), including *Brassica rapa*, *Brassica oleracea*, *Brassica napus*, *Brassica nigra*, and *Camelina sativa* (Arias and Pires, 2012; Huang *et al.*, 2016; Warwick and Black, 1991). *H. incana*, therefore, belongs to one of the most well-studied plant families with regard to phylogeny, physiology, genetics, and genomics (Al-Shehbaz, 2012; Arias and Pires, 2012; Franzke *et al.*, 2011; Nikolov and Tsiantis, 2017), and which has been subject to in-depth comparative genomics studies (Koenig and Weigel, 2015; Schranz *et al.*, 2006; van den Bergh *et al.*, 2016). The fairly close phylogenetic relationship shared between *H. incana* and *A. thaliana* makes it feasible to swap tools and knowledge developed for either species. *H. incana* has a somatic chromosome number of 2n = 14 (Anderson and Warwick, 1999; Rollins, 1981), and an estimated genome size of about 450 Mb (Garassino *et al.*, 2022), which modern genomics technologies make accessible. The evolutionary history that *H. incana* shares with its relatives involve a series of ancient whole genome duplication and triplication events interspersed with diploidization events (Franzke *et al.*, 2011). This complex evolutionary history is believed to have resulted in a great potential for the emergence of new genetic functions or the preferential retention of favorable genes in Brassicaceae species exposed to different environmental constraints, thus resulting in their adaptation to these environments (Cheng *et al.*, 2014; Qi *et al.*, 2021; Zhang *et al.*, 2021). The ge-

netic blueprint for the evolution of high assimilation rate characteristics of *H. incana* may have arisen from the preferential retention or the evolution of specific photosynthetic efficiency-related genes because of its evolution in dry, hot, and sunny climates with ephemeral, low competition habitats.

The widespread distribution of *H. incana*, combined with its large seed production and ease of germination and growth, allows for the easy building of diverse panels, as demonstrated by the fact that without any targeted collection campaign our laboratory was able to obtain more than 30 accessions from three continents. Intercrossing between these accessions proved to be easy, and the resulting progeny can easily be propagated, implying that the construction of mapping populations in *H. incana* is a feasible task. Hybridization of *H. incana* has been shown to be possible with commercial Brassica crops such as *B. napus* (Darmency and Fleury, 2000; Lefol *et al.*, 1995; Siemens, 2002) and to some extent *B. oleracea* (Quiros *et al.*, 1988), on other Brassicaceae species such as *B. nigra* (Quiros *et al.*, 1988; Salisbury, 1991) and *B. carinata* (Mohanty *et al.*, 2007), as well as with very close relatives such as *Erucastrum elatum* and *Erucastrum virgatum*, based on our own experience. Thus, the introgression of genes underlying the photosynthetic efficiency of *H. incana* appears possible, with the best candidate for introgression being *B. napus* (Devos *et al.*, 2009).

2.8 Photosynthetic carachteristics of *H. incana*

H. incana exhibits exceptional C_3 photosynthetic characteristics (Canvin *et al.*, 1980). A photosynthetic capacity of 45 to 55 µmol m^{-2} s^{-1} was attained by young, fully expanded leaves when measured under 21% O_2, 400 ppm CO_2 and 70% relative humidity (Figure 2.4a). Although these rates are not as high as the highest assimilation rates recorded for *Camissionia*/*Chylismia* species or *Palafoxia linearis* (Ehleringer *et al.*, 1979; Longstreth *et al.*, 1980; Seemann *et al.*, 1980; Werk *et al.*, 1983), *H. incana* has the advantage of comparatively simple genetics and being a member of the Brassicaceae family. Although high photosynthetic capacity was recently reported for another species from the Brassicaceae family, *B. rapa* (Taylor *et al.*, 2020), *H. incana* currently represents the upper limit for what might be the highest achievable photosynthetic capacity in the family. While in an ideal scenario one would like to have a model species with large intra-species variation for P_{max} under similar growth conditions, this has not been reported for any plant species to date. Nevertheless, a degree of natural variation in photosynthetic capacity was already observed for *H. incana* (Garassino *et al.*, 2022), and the aforementioned widespread distribution of the species across geographical locations and climates suggests that more variation can be easily accessed, potentially allowing for the reporting of even higher photosynthetic capacity.

The rate constant for linear electron transport (measured using the dark relaxation of the 820 nm absorbance change as a measure of kinetic limitation imposed by the photosynthetic electron transport chain (Baker *et al.*, 2007)) generally had a $t_{1/2}$ of 1.7–2.5 ms (equivalent to a rate constant of 272–230 s^{-1}) — significantly higher than

values obtained, for example, pea and spinach leaves (Harbinson and Hedley, 1989; Schreiber *et al.*, 1989). The proton efflux rate constant via the ATPase (measured via the 520 nm light-induced absorbance change (Baker *et al.*, 2007)) in *H. incana* leaves is nearly constant at 80 s^{-1} in the irradiance range 1250–2400 µmol m^{-2} s^{-1}.

The decline of ΦPSII and ΦPSI with increasing irradiance was found to be modest for *H. incana*, decreasing to only 0.43 (ΦPSII) and 0.61 (ΦPSI) at an irradiance of 2400 µmol m^{-2} s^{-1} (Figure 2.4b). Similarly, the decline of F_v/F_m and *Fv'/Fm'* with irradiance was modest up to a saturating irradiance of 2400 µmol m^{-2} s^{-1} (Figure 2.4c). The relationship between ΦPSII and ΦPSI with increasing irradiance was also relatively linear, a finding consistent with results obtained from other species (Figure 2.4d). Therefore, the photosynthetic apparatus in *H. incana* leaves is remarkably robust to high irradiances, exhibiting little to no physiological stress or photoinhibition at measurement irradiances of up to 2400 µmol m^{-2} s^{-1} - levels higher than could be expected on a cloudless day at noon in the Equator during an equinox (˜2200 µmol m^{-2} s^{-1}) or for horizontal locations at 37°N during a summer solstice solar noon (˜2130 µmol m^{-2} s^{-1}) (Ritchie, 2010). In our hands, juvenile *H. incana* plants could be exposed for up to 12 h a day to 2400 µmol m^{-2} s^{-1} or for 16 h to 2000 µmol m^{-2} s^{-1} without detectable signs of physiological stress or a decrease in the dark-adapted F_v/F_m values. These exceptional high light performance and tolerance characteristics of *H. incana* support its candidacy as a model C_3 species in research focused on improving C_3 crop photosynthesis.

2.9 Concluding thoughts

A. thaliana did and continues to play a pivotal role as a model plant species in wide-ranging ecophysiological, photosynthesis, and genetic research. However, for research directed at bettering our understanding of the genetic, anatomical, and physiological adaptations required for sustained high rates of photosynthesis in a C_3 plant, *H. incana* can become a valuable complementary plant model. It is capable of withstanding, growing under, and photosynthetically making efficient use of very high irradiances. The diploid genetics and position within the Brassicaceae make *H. incana* genetically simple and allow the wealth of genomic information that exists for this family to be leveraged. In particular, the broad depth of genetic and physiological knowledge, and the techniques, available for *A. thaliana*, another Brassicaceae species, further strengthen the case for using *H. incana* as a high-light tolerant, high assimilation rate model species. In our hands the plant has also proven easy to grow in both greenhouses and growth cabinets — the greatest challenge was increasing fertilization and watering regimes to meet the high demands of the plant. At least in Western Europe, the seed is available from naturalized plant populations, although currently no diversity panels, or similar, exist for the species. We expect that more species with exceptional photosynthesis traits like *H. incana* will be established as models, which can collectively aid our understanding of how these traits function, and how they emerge from genetic specialization.

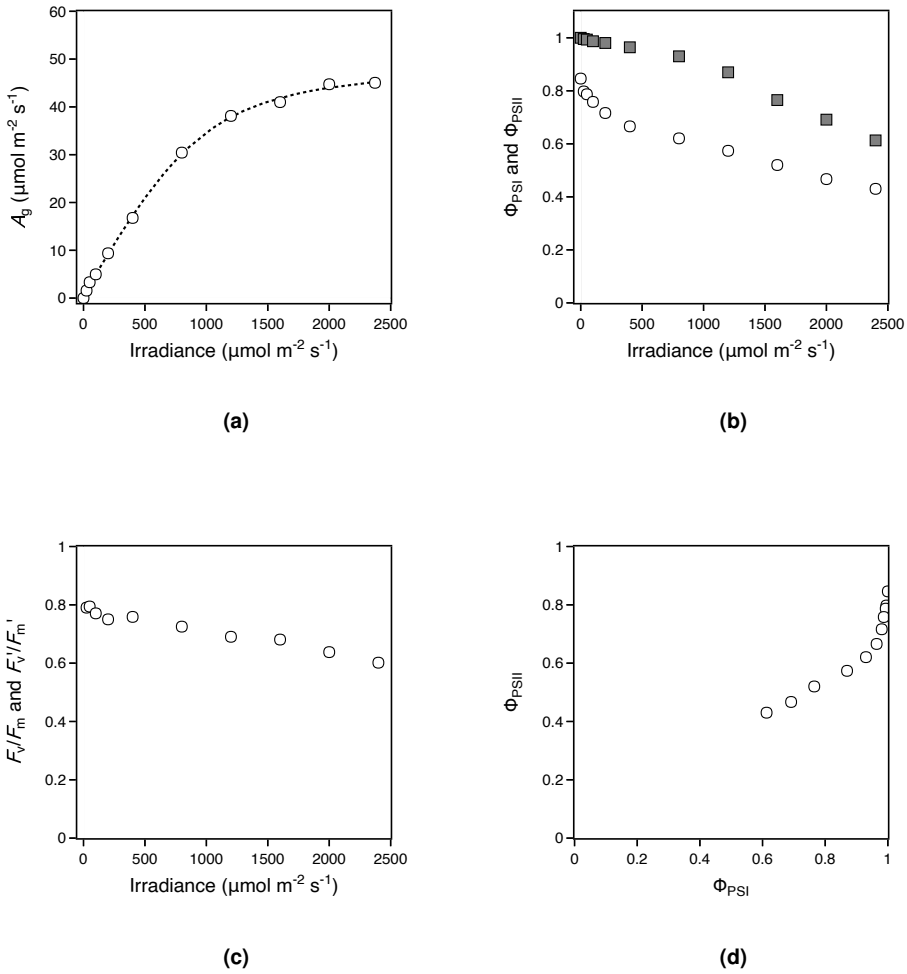

Figure 2.4: Photosynthetic characteristics of *Hirschfeldia incana*. All measurements conducted at 21% O_2, 400 ppm CO_2 and 70% relative humidity, with a 660 nm spectral peak actinic and chlorophyll fluorescence excitation light source, for a representative, young, fully expanded *H. incana* leaf grown under 1800 µmol m^{-2} s^{-1} PAR irradiance. **(a)** Gross photosynthetic assimilation rate (A_g) (○) and a fit estimate of the response to irradiance using a non-rectangular hyperbolic equation (Acock *et al.*, 1976; Thornley, 1976). **(b)** Φ_{PSI} (■) and ΦPSII (○) with increasing irradiance. **(c)** F_v/F_m (single dark-adapted measurement), and *Fv'/Fm'* (light-adapted measurements) with increasing irradiance. **(d)** Operating efficiencies of photosystems PSI (Φ_{PSI}) and PSII (ΦPSII) with increasing irradiance.

Author contributions

Graham Taylor, Francesco Garassino, and Jeremy Harbinson conceptualized the manuscript. Graham Taylor and Francesco Garassino drafted the manuscript using data derived from Graham Taylor; Graham Taylor designed the figures. Jeremy Harbinson and Mark G. M. Aarts revised and finalized the manuscript content and outline.

Acknowledgments

We thank Keir Morse for allowing the use (under license CC BY-NC-SA 3.0) of the picture used in Figure 2b, and Jim Boone for providing us with the picture used in Figure 2c.

References

Acevedo-Siaca, L. G., Coe, R., Quick, W. P., and Long, S. P. (2020a). Variation between rice accessions in photosynthetic induction in flag leaves and underlying mechanisms. *Journal of Experimental Botany*, **72**(4), 1282–1294.

Acevedo-Siaca, L. G., Coe, R., Wang, Y., Kromdijk, J., Quick, W. P., and Long, S. P. (2020b). Variation in photosynthetic induction between rice accessions and its potential for improving productivity. *New Phytologist*, **227**(4), 1097–1108.

Acock, B., Hand, D., Thornley, J., and Wilson, J. W. (1976). Photosynthesis in stands of green peppers. An application of empirical and mechanistic models to controlled-environment data. *Annals of Botany*, **40**(6), 1293–1307.

Al-Shehbaz, I. A. (1977). Protogyny in the Cruciferae. *Systematic Botany*, pages 327–333.

Al-Shehbaz, I. A. (2012). A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon*, **61**(5), 931–954.

Anderson, J. K. and Warwick, S. I. (1999). Chromosome number evolution in the tribe Brassiceae (Brassicaceae): evidence from isozyme number. *Plant Systematics and Evolution*, **215**(1), 255–285.

Aranda, I., Cano, F. J., Gascó, A., Cochard, H., Nardini, A., Mancha, J. A., López, R., and Sánchez-Gómez, D. (2014). Variation in photosynthetic performance and hydraulic architecture across European beech (Fagus sylvatica L.) populations supports the case for local adaptation to water stress. *Tree Physiology*, **35**(1), 34–46.

Arias, T. and Pires, J. C. (2012). A fully resolved chloroplast phylogeny of the brassica crops and wild relatives (Brassicaceae: Brassiceae): Novel clades and potential taxonomic implications. *Taxon*, **61**(5), 980–988.

Austin, R. B. (1989). Genetic variation in photosynthesis. *The Journal of Agricultural Science*, **112**(3), 287–294.

Baker, N. R., Harbinson, J., and Kramer, D. M. (2007). Determining the limitations and regulation of photosynthetic energy transduction in leaves. *Plant, Cell & Environment*, **30**(9), 1107–1125.

Beadle, C. and Long, S. (1985). Photosynthesis—is it limiting to biomass production? *Biomass*, **8**(2), 119–168.

Burgess, S. J., de Becker, E., Cullum, S., Causon, I., Floristeanu, I., Chan, K. X., Moore, C. E., Diers, B. W., and Long, S. P. (2020). Variation in relaxation of non-photochemical quenching in a soybean nested association mapping panel as a potential source for breeding improved photosynthesis. *bioRxiv*, page 2020.07.29.201210.

Canvin, D. T., Berry, J. A., Badger, M. R., Fock, H., and Osmond, C. B. (1980). Oxygen exchange in leaves in the light. *Plant Physiology*, **66**(2), 302–307.

Cardona, T., Shao, S., and Nixon, P. J. (2018). Enhancing photosynthesis in plants: the light reactions. *Essays in biochemistry*, **62**(1), 85–94.

Cheng, F., Wu, J., and Wang, X. (2014). Genome triplication drove the diversification of Brassica plants. *Horticulture Research*, **1**(1), 14024.

Chronopoulos, G., Theocharopoulos, M., and Christodoulakis, D. (2005). Phytosociological study of Hirschfeldia incana (L.) Lagraze-Fossat (Cruciferae) communities in mainland Greece. *Acta Botanica Croatica*, **64**(1), 75–114.

Dann, M. and Leister, D. (2017). Enhancing (crop) plant photosynthesis by introducing novel genetic diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**(1730), 20160380.

Darmency, H. and Fleury, A. (2000). Mating system in Hirschfeldia incana and hybridization to oilseed rape. *Weed Research (Oxford)*, **40**(2), 231–238.

Devos, Y., De Schrijver, A., and Reheul, D. (2009). Quantifying the introgressive hybridisation propensity between transgenic oilseed rape and its wild/weedy relatives. *Environmental monitoring and assessment*, **149**(1), 303–322.

Dimkpa, C. O. and Bindraban, P. S. (2016). Fortification of micronutrients for efficient agronomic production: a review. *Agronomy for Sustainable Development*, **36**(1), 7.

DiTomaso, J. M. and Healy, E. A. (2007). *Weeds of California and other western states*, volume 3488. UCANR Publications. Type: Book.

DiTomaso, J. M., Kyser, G. B., Oneto, S. R., Wilson, R. G., Orloff, S. B., Anderson, L. W., Wright, S. D., Roncoroni, J., Miller, T., and Prather, T. (2013). Weed control in natural areas in the western United States. *Weed Research and Information Center, University of California*, **544**.

Driever, S. M., Lawson, T., Andralojc, P. J., Raines, C. A., and Parry, M. A. J. (2014). Natural variation in photosynthetic capacity, growth, and yield in 64 field-grown wheat genotypes. *Journal of Experimental Botany*, **65**(17), 4959–4973.

Driever, S. M., Simkin, A. J., Alotaibi, S., Fisk, S. J., Madgwick, P. J., Sparks, C. A., Jones, H. D., Lawson, T., Parry, M. A. J., and Raines, C. A. (2017). Increased SBPase activity improves photosynthesis and grain yield in wheat grown in greenhouse conditions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**(1730), 20160384.

Ehleringer, J. (1983). Ecophysiology of Amaranthus palmeri, a Sonoran Desert summer annual. *Oecologia*, **57**(1), 107–112.

Ehleringer, J. and Björkman, O. (1977). Quantum yields for CO2 uptake in C3 and C4 plants: dependence on temperature, CO2, and O2 concentration. *Plant Physiology*, **59**(1), 86–90.

Ehleringer, J., Mooney, H. A., and Berry, J. A. (1979). Photosynthesis and microclimate of Camissonia claviformis, a desert winter annual. *Ecology*, **60**(2), 280–286.

Evans, J. R. (2013). Improving photosynthesis. *Plant physiology*, **162**(4), 1780–1793.

Fageria, N., Baligar, V., and Li, Y. (2008). The role of nutrient efficient plants in improving crop yields in the twenty first century. *Journal of plant nutrition*, **31**(6), 1121–1157.

FAO (2022). FAOSTAT Statistical Database. Type: Web Page.

Flood, P. J. (2019). Using natural variation to understand the evolutionary pressures on plant photosynthesis. *Current opinion in plant biology*, **49**, 68–73.

Flood, P. J., Harbinson, J., and Aarts, M. G. (2011). Natural genetic variation in plant photosynthesis. *Trends in plant science*, **16**(6), 327–335.

Flood, P. J., Kruijer, W., Schnabel, S. K., van der Schoor, R., Jalink, H., Snel, J. F., Harbinson, J., and Aarts, M. G. (2016). Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. *Plant methods*, **12**(1), 1–14.

Flood, P. J., Theeuwen, T. P., Schneeberger, K., Keizer, P., Kruijer, W., Severing, E., Kouklas, E., Hageman, J. A., Wijfjes, R., and Calvo-Baltanas, V. (2020). Reciprocal cybrids reveal how organellar genomes affect plant phenotypes. *Nature plants*, **6**(1), 13–21.

Franks, P. J., W. Doheny-Adams, T., Britton-Harper, Z. J., and Gray, J. E. (2015). Increasing water-use efficiency directly through genetic manipulation of stomatal density. *New Phytologist*, **207**(1), 188–195.

Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A., and Mummenhoff, K. (2011). Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in plant science*, **16**(2), 108–116.

Furbank, R. T. and Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends in plant science*, **16**(12), 635–644.

Garassino, F., Wijfjes, R. Y., Boesten, R., Becker, F. F. M., Clapero, V., van den Hatert, I., Holmer, R., Schranz, M. E., Harbinson, J., de Ridder, D., Smit, S., and Aarts, M. G. M. (2022). The genome sequence of Hirschfeldia incana, a species with high photosynthetic light-use efficiency. *bioRxiv*, page 2022.01.29.478283.

Gifford, R. M. and Evans, L. T. (1981). Photosynthesis, Carbon Partitioning, and Yield. *Annual Review of Plant Physiology*, **32**(1), 485–509.

Gilbert, M. E., Zwieniecki, M. A., and Holbrook, N. M. (2011). Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. *Journal of Experimental Botany*, **62**(8), 2875–2887.

Gitelson, A. A., Peng, Y., Arkebauer, T. J., and Suyker, A. E. (2015). Productivity, absorbed photosynthetically active radiation, and light use efficiency in crops: Implications for remote sensing of crop primary production. *Journal of plant physiology*, **177**, 100–109.

Gresta, F., Cristaudo, A., Onofri, A., Restuccia, A., and Avola, G. (2010). Germination response of four pasture species to temperature, light, and post-harvest period. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, **144**(4), 849–856.

Gu, J., Yin, X., Stomph, T., and Struik, P. C. (2014). Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. *Plant, cell & environment*, **37**(1), 22–34.

Gu, J., Zhou, Z., Li, Z., Chen, Y., Wang, Z., Zhang, H., and Yang, J. (2017). Photosynthetic properties and potentials for improvement of photosynthesis in pale green leaf rice under high light conditions. *Frontiers in Plant Science*, **8**, 1082.

Harbinson, J. and Hedley, C. (1989). The kinetics of P-700+ reduction in leaves: a novel in situ probe of thylakoid functioning. *Plant, Cell & Environment*, **12**(4), 357–369.

Harbinson, J. and Woodward, F. (1984). Field measurements of the gas exchange of woody plant species in simulated sunflecks. *Annals of Botany*, **53**(6), 841–851.

Harbinson, J. and Yin, X. (2017). A model for the irradiance responses of photosynthesis. *Physiologia Plantarum*, **161**(1), 109–123.

Heuer, S., Gaxiola, R., Schilling, R., Herrera-Estrella, L., López-Arredondo, D., Wissuwa, M., Delhaize, E., and Rouached, H. (2017). Improving phosphorus use efficiency: a complex trait with emerging opportunities. *The Plant Journal*, **90**(5), 868–885.

Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of experimental botany*, **58**(9), 2369–2387.

Hogewoning, S. W., Wientjes, E., Douwstra, P., Trouwborst, G., Van Ieperen, W., Croce, R., and Harbinson, J. (2012). Photosynthetic quantum yield dynamics: from photosystems to leaves. *The plant cell*, **24**(5), 1921–1935.

Huang, C.-H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., Zhang, Q., Koch, M. A., Al-Shehbaz, I., and Edger, P. P. (2016). Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular biology and evolution*, **33**(2), 394–412.

Kaiser, E., Morales, A., Harbinson, J., Heuvelink, E., Prinzenberg, A. E., and Marcelis, L. F. (2016). Metabolic and diffusional limitations of photosynthesis in fluctuating irradiance in Arabidopsis thaliana. *Scientific Reports*, **6**(1), 1–13.

Kao, W. and Forseth, I. (1992). Dirunal leaf movement, chlorophyll fluorescence and carbon assimilation in soybean grown under different nitrogen and water availabilities. *Plant, Cell & Environment*, **15**(6), 703–710.

Koenig, D. and Weigel, D. (2015). Beyond the thale: comparative genomics and genetics of Arabidopsis relatives. *Nature Reviews Genetics*, **16**(5), 285–298.

Koornneef, M. and Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal*, **61**(6), 909–921.

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**(6314), 857–861.

Lawson, T., Kramer, D. M., and Raines, C. A. (2012). Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. *Current Opinion in Biotechnology*, **23**(2), 215–220.

Leakey, A. D., Ferguson, J. N., Pignon, C. P., Wu, A., Jin, Z., Hammer, G. L., and Lobell, D. B. (2019). Water Use Efficiency as a Constraint and Target for Improving the Resilience and Productivity of C3 and C4 Crops. *Annual Review of Plant Biology*, **70**(1), 781–808.

Lee, P. L., Patel, R., Conlan, R., Wainwright, S., and Hipkin, C. (2004). Comparison of genetic diversities in native and alien populations of hoary mustard (Hirschfeldia incana [L.] Lagreze-Fossat). *International Journal of Plant Sciences*, **165**(5), 833–843.

Lefol, E., Danielou, V., Darmency, H., Boucher, F., Maillet, J., and Renard, M. (1995). Gene dispersal from transgenic crops. I. Growth of interspecific hybrids between oilseed rape and the wild hoary mustard. *Journal of Applied Ecology*, pages 803–808.

Lin, P., Tsai, Y., Hsu, S., Ou, J., Liao, C., and Tung, C. (2018). Identification of natural variants affecting chlorophyll content dynamics during rice seedling development. *Plant Breeding*, **137**(3), 355–363.

Long, S. P., Zhu, X.-G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant, Cell and Environment*, **29**(3), 315–330.

Longstreth, D. J., Hartsock, T. L., and Nobel, P. S. (1980). Mesophyll cell properties for some C3 and C4 species with high photosynthetic rates. *Physiologia plantarum*, **48**(4), 494–498.

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). *Frontiers in Plant Science*, **10**.

López-Calcagno, P. E., Brown, K. L., Simkin, A. J., Fisk, S. J., Vialet-Chabrand, S., Lawson, T., and Raines, C. A. (2020). Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants*, **6**(8), 1054–1063.

Manafzadeh, S., Staedler, Y. M., and Conti, E. (2017). Visions of the past and dreams of the future in the Orient: the Irano-Turanian region from classical botany to evolutionary studies. *Biological Reviews*, **92**(3), 1365–1388.

Mitchell, P. and Sheehy, J. E. (2006). Supercharging rice photosynthesis to increase yield. *New Phytologist*, pages 688–693.

Mohanty, A., Chrungu, B., Verma, N., and Shivanna, K. (2007). Wide Hybridization between Hirschfeldia incana L.(Lagreze Fossat) and Brassica carinata Braun: Towards Development of New Genetic Resource. *Indian Journal of Plant Genetic Resources*, **20**(3), 208–210.

Molero, G. and Reynolds, M. P. (2020). Spike photosynthesis measured at high throughput indicates genetic variation independent of flag leaf photosynthesis. *Field Crops Research*, **255**, 107866.

Monneveux, P., Pastenes, C., and Reynolds, M. P. (2003). Limitations to photosynthesis under light and heat stress in three high-yielding wheat genotypes. *Journal of Plant Physiology*, **160**(6), 657–666.

Monreal, C., DeRosa, M., Mallubhotla, S., Bindraban, P., and Dimkpa, C. (2016). Nanotechnologies for increasing the crop use efficiency of fertilizer-micronutrients. *Biology and fertility of soils*, **52**(3), 423–437.

Monteith, J. L. (1977). Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **281**(980), 277–294.

Murchie, E. H., Chen, Y.-z., Hubbart, S., Peng, S., and Horton, P. (1999). Interactions between senescence and leaf orientation determine in situ patterns of photosynthesis and photoinhibition in field-grown rice. *Plant physiology*, **119**(2), 553–564.

Murchie, E. H., Pinto, M., and Horton, P. (2009). Agriculture and the new challenges for photosynthesis research. *New Phytologist*, **181**(3), 532–552.

Nikolov, L. A. and Tsiantis, M. (2017). Using mustard genomes to explore the genetic basis of evolutionary change. *Current opinion in plant biology*, **36**, 119–128.

Nobel, P. S. (1991). Achievable productivities of certain CAM plants: basis for high values compared with C3 and C4 plants. *New phytologist*, **119**(2), 183–205.

Ortiz, D., Hu, J., and Salas Fernandez, M. G. (2017). Genetic architecture of photosynthesis in Sorghum bicolor under non-stress and cold stress conditions. *Journal of experimental botany*, **68**(16), 4545–4557.

Parry, M. A., Reynolds, M., Salvucci, M. E., Raines, C., Andralojc, P. J., Zhu, X.-G., Price, G. D., Condon, A. G., and Furbank, R. T. (2011). Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of experimental botany*, **62**(2), 453–467.

Pearcy, R. W. (1990). Sunflecks and photosynthesis in plant canopies. *Annual review of plant biology*, **41**(1), 421–453.

Piedade, M. T. F., Junk, W. J., and Long, S. P. (1991). The productivity of the C_4 Grass Echinochloa polystachya on the Amazon floodplain. *Ecology*, **72**(4), 1456–1463.

Poolman, M. G., Fell, D. A., and Thomas, S. (2000). Modelling photosynthesis and its control. *Journal of Experimental Botany*, **51**(suppl_1), 319–328.

Prinzenberg, A. E., Víquez-Zamora, M., Harbinson, J., Lindhout, P., and van Heusden, S. (2018). Chlorophyll fluorescence imaging reveals genetic variation and loci for a photosynthetic trait in diploid potato. *Physiologia Plantarum*, **164**(2), 163–175.

Prinzenberg, A. E., Campos-Dominguez, L., Kruijer, W., Harbinson, J., and Aarts, M. G. (2020). Natural variation of photosynthetic efficiency in Arabidopsis thaliana accessions under low temperature conditions. *Plant, Cell & Environment*, **43**(8), 2000–2013.

Qi, X., An, H., Hall, T. E., Di, C., Blischak, P. D., McKibben, M. T. W., Hao, Y., Conant, G. C., Pires, J. C., and Barker, M. S. (2021). Genes derived from ancient polyploidy have higher genetic diversity and are associated with domestication in Brassica rapa. *New Phytol*, **230**(1), 372–386.

Quiros, C., Ochoa, O., and Douches, D. (1988). Exploring the role of x= 7 species in Brassica evolution: hybridization with B. nigra and B. oleracea. *Journal of Heredity*, **79**(5), 351–358.

Quiroz, D. (2015). Hirschfeldia incana.

Raun, W. R. and Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy journal*, **91**(3), 357–363.

Richards, R. A., López-Castañeda, C., Gomez-Macpherson, H., and Condon, A. G. (1993). Improving the efficiency of water use by plant breeding and molecular biology. *Irrigation Science*, **14**(2), 93–104.

Ritchie, R. (2010). Modelling photosynthetic photon flux density and maximum potential gross photosynthesis. *Photosynthetica*, **48**(4), 596–609.

Rollins, R. C. (1981). Weeds of the cruciferae (Brassicaceae) in North America. *Journal of the Arnold Arboretum*, **62**(4), 517–540.

Salisbury, P. A. (1991). *Genetic variability in Australian wild crucifers and its potential utilisation in oilseed Brassica species*. Thesis.

Schranz, M. E., Lysak, M. A., and Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in plant science*, **11**(11), 535–542.

Schreiber, U., Neubauer, C., and Klughammer, C. (1989). Devices and methods for room-temperature fluorescence analysis. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **323**(1216), 241–251.

Seemann, J. R., Field, C., and Berry, J. A. (1980). Photosynthetic capacity of desert winter annuals measured in situ. Report.

Sheehy, J. E., Ferrer, A. B., and Mitchell, P. (2008). Harnessing photosynthesis in tomorrow's world: Humans, crop production and poverty alleviation. In *Photosynthesis. Energy from the sun*, pages 1237–1242. Springer. Type: Book Section.

Siemens, J. (2002). Interspecific hybridisation between wild relatives and Brassica napus to introduce new resistance traits into the oilseed rape gene pool. *Czech Journal of Genetics and Plant Breeding*, **38**(3/4), 155–157.

Siemens, J. (2011). Hirschfeldia. In *Wild Crop Relatives: Genomic and Breeding Resources*, pages 171–176. Springer. Type: Book Section.

Simkin, A. J., McAusland, L., Headland, L. R., Lawson, T., and Raines, C. A. (2015). Multigene manipulation of photosynthetic carbon assimilation increases CO2 fixation and biomass yield in tobacco. *Journal of Experimental Botany*, **66**(13), 4075–4090.

Sinclair, T. R., Purcell, L. C., and Sneller, C. H. (2004). Crop transformation and the challenge to increase yield potential. *Trends in plant science*, **9**(2), 70–75.

Sinclair, T. R., Rufty, T. W., and Lewis, R. S. (2019). Increasing photosynthesis: unlikely solution for world food problem. *Trends in Plant Science*, **24**(11), 1032–1039.

Sonoike, K. (2011). Photoinhibition of photosystem I. *Physiologia plantarum*, **142**(1), 56–64.

South, P. F., Cavanagh, A. P., Liu, H. W., and Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, **363**(6422).

Strigens, A., Freitag, N. M., Gilbert, X., Grieder, C., Riedelsheimer, C., Schrag, T. A., Messmer, R., and Melchinger, A. E. (2013). Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments. *Plant, cell & environment*, **36**(10), 1871–1887.

Swarbreck, S. M., Wang, M., Wang, Y., Kindred, D., Sylvester-Bradley, R., Shi, W., Bentley, A. R., and Griffiths, H. (2019). A roadmap for lowering crop nitrogen requirement. *Trends in plant science*, **24**(10), 892–904.

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**(3), 757–764.

Taylor, S. H., Orr, D. J., Carmo-Silva, E., and Long, S. P. (2020). During photosynthetic induction, biochemical and stomatal limitations differ between Brassica crops. *Plant, Cell & Environment*, **43**(11), 2623–2636.

Theeuwen, T. P. J. M., Logie, L. L., Harbinson, J., and Aarts, M. G. M. (2022). Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany*, **73**(10), 3122–3137.

Thornley, J. H. (1976). *Mathematical models in plant physiology*. Academic Press (Inc.) London, Ltd. Type: Book.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, **418**(6898), 671–677.

Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., and Gregory, M. (2003). A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biology*, **9**(3), 383–395.

van Bezouw, R. F. H. M., Keurentjes, J. J. B., Harbinson, J., and Aarts, M. G. M. (2019). Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. *The Plant Journal*, **97**(1), 112–133.

van den Bergh, E., Hofberger, J. A., and Schranz, M. E. (2016). Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany*, **103**(7), 1212–1222.

van Rooijen, R., Aarts, M. G., and Harbinson, J. (2015). Natural Genetic Variation for Acclimation of Photosynthetic Light Use Efficiency to Growth Irradiance in Arabidopsis. *Plant Physiology*, **167**(4), 1412–1429.

van Rooijen, R., Kruijer, W., Boesten, R., Van Eeuwijk, F. A., Harbinson, J., and Aarts, M. G. (2017). Natural variation of YELLOW SEEDLING1 affects photosynthetic acclimation of Arabidopsis thaliana. *Nature communications*, **8**(1), 1–9.

Veneklaas, E. J., Lambers, H., Bragg, J., Finnegan, P. M., Lovelock, C. E., Plaxton, W. C., Price, C. A., Scheible, W., Shane, M. W., and White, P. J. (2012). Opportunities for improving phosphorus-use efficiency in crop plants. *New phytologist*, **195**(2), 306–320.

von Caemmerer, S. and Evans, J. R. (2010). Enhancing C3 photosynthesis. *Plant Physiology*, **154**(2), 589–592.

Wang, Q., Xie, W., Xing, H., Yan, J., Meng, X., Li, X., Fu, X., Xu, J., Lian, X., Yu, S., Xing, Y., and Wang, G. (2015). Genetic Architecture of Natural Variation in Rice Chlorophyll Content Revealed by a Genome-Wide Association Study. *Mol Plant*, **8**(6), 946–57.

Warwick, S. I. and Black, L. D. (1991). Molecular systematics of Brassica and allied genera (subtribe Brassicinae, Brassiceae)—chloroplast genome and cytodeme congruence. *Theoretical and Applied Genetics*, **82**(1), 81–92.

Welch, R. M. and Graham, R. D. (2002). Breeding crops for enhanced micronutrient content. In *Food Security in Nutrient-Stressed Environments: Exploiting Plants' Genetic Capabilities*, pages 267–276. Springer. Type: Book Section.

Werk, K., Ehleringer, J., Forseth, I., and Cook, C. (1983). Photosynthetic characteristics of Sonoran Desert winter annuals. *Oecologia*, **59**(1), 101–105.

Xu, G., Tang, H., Zhang, H., Li, H., Ji, R., Qiu, S., Luo, W., and Xu, Y. (2021). Screening for Genetic Variation in Photosynthesis and High Photosynthetic Efficiency Germplasm in Potato. *Acta Horticulturae Sinica*, **48**(11), 2239–2250.

Zelitch, I. (1975). Improving the Efficiency of Photosynthesis. *Science*, **188**(4188), 626–633.

Zhang, H., Xie, J., Wang, W., and Wang, J. (2021). Comparison of Brassica Genomes reveals asymmetrical gene retention between functional groups of genes in recurrent polyploidizations. *Plant Molecular Biology*, **106**(1), 193–206.

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? *Current opinion in biotechnology*, **19**(2), 153–159.

Zhu, X.-G., Long, S. P., and Ort, D. R. (2010). Improving Photosynthetic Efficiency for Greater Yield. *Annual Review of Plant Biology*, **61**(1), 235–261.

CHAPTER 3

The genome sequence of *Hirschfeldia incana*, a new Brassicaceae model to improve photosynthetic light-use efficiency

The Plant Journal, 2022, 112: 1298-1315.

Francesco Garassino[1,*], Raúl Y. Wijfjes[2,5,*], René Boesten[1,*],Francisca Reyes Marquez[1], Frank F. M. Becker[1], Vittoria Clapero[1,6], Iris van den Hatert[2], Rens Holmer[2], M. Eric Schranz[3], Jeremy Harbinson[4], Dick de Ridder[2], Sandra Smit[1], Mark G. M. Aarts[1]

(1) Laboratory of Genetics, Wageningen University and Research, Wageningen (NL)

(2) Bioinformatics Group, Wageningen University and Research, Wageningen (NL)

(3) Biosystematics Group, Wageningen University and Research, Wageningen (NL)

(4) Laboratory of Biophysics, Wageningen University and Research, Wageningen (NL)

(5) Current affiliation: Faculty of Biology, Ludwig Maximilian University of Munich, Munich (DE)

(6) Current affiliation: Max Planck Institute for Molecular Plant Physiology, Golm (DE)

(*) These authors contributed equally to this work

Supplementary material available at:
https://onlinelibrary.wiley.com/doi/full/10.1111/tpj.16005

Abstract

Photosynthesis is a key process in sustaining plant and human life. Improving the photosynthetic capacity of agricultural crops is an attractive means to increase their yields. While the core mechanisms of photosynthesis are highly conserved in C_3 plants, these mechanisms are very flexible, allowing considerable diversity in photosynthetic properties. Among this diversity is the maintenance of high photosynthetic light-use efficiency at high irradiance as identified in a small number of exceptional C_3 species. *Hirschfeldia incana*, a member of the Brassicaceae family, is such an exceptional species, and because it is easy to grow, it is an excellent model for studying the genetic and physiological basis of this trait. Here, we present a reference genome of *H. incana* and confirm its high photosynthetic light-use efficiency. While *H. incana* has the highest photosynthetic rates found so far in the Brassicaceae, the light-saturated assimilation rates of closely related *Brassica rapa* and *Brassica nigra* are also high. The *H. incana* genome has extensively diversified from that of *B. rapa* and *B. nigra* through large chromosomal rearrangements, species-specific transposon activity, and differential retention of duplicated genes. Duplicated genes in *H. incana*, *B. rapa*, and *B. nigra* that are involved in photosynthesis and/or photoprotection show a positive correlation between copy number and gene expression, providing leads into the mechanisms underlying the high photosynthetic efficiency of these species. Our work demonstrates that the *H. incana* genome serves as a valuable resource for studying the evolution of high photosynthetic light-use efficiency and enhancing photosynthetic rates in crop species.

3.1 Introduction

Photosynthesis is the biophysical and biochemical process that sustains most life on planet Earth. The most common form of photosynthesis, oxygenic photosynthesis, uses solar energy to convert the inorganic carbon dioxide (CO_2) to organic carbon, typically represented as a carbohydrate, releasing molecular oxygen (O_2) from water in the process. Terrestrial plants provide by far the most conspicuous example of oxygenic photosynthesis (referred to as photosynthesis from now on for brevity) and are responsible for about 50% of the primary production of oxygen in the biosphere, with marine production by eukaryotic algae and cyanobacteria comprising the other 50%. Agriculture depends on primary production by plants, so expanding our knowledge of photosynthesis is crucial if we are to meet many of the pressing global challenges faced by mankind.

One of these challenges is the need to substantially increase the yield of agricultural crops to meet the increasing demand not only for food and fodder, but also for fibers and similar plant products, and organic precursors for the chemical industry as it transitions away from fossil carbon sources. A major yield-related trait is the conversion efficiency of absorbed solar irradiance to biomass (ϵ_c; Long *et al.* (2006)), a parameter which is strongly influenced by the light-use efficiency of photosynthesis. As light intensity, or irradiance, increases, the photosynthetic light-use efficiency of leaves and other photosynthetic organs decreases, which leads ultimately to the light-saturation of photosynthesis (Genty and Harbinson, 1996; Murchie *et al.*, 1999; Monneveux *et al.*, 2003; Turner *et al.*, 2003; Gitelson *et al.*, 2015; Gu *et al.*, 2017). Once light-saturation is reached, any additional light will not lead to a further increase in the photosynthetic rate and may even be detrimental to photosynthesis. The threshold for light saturation generally lies far below the maximum level of irradiance experienced in the field or greenhouse (Zhu *et al.*, 2010) and for most C_3 crops this light saturation phenomenon is an aspect of their photosynthesis which remains to be increased in order to increase yield. Improving the photosynthetic light-use efficiency of crop plants thus paves the way towards increasing their ϵ_c and ultimately their yield (Flood *et al.*, 2011; Furbank *et al.*, 2019; Lawson *et al.*, 2012; von Caemmerer and Evans, 2010; Zhu *et al.*, 2010), as recently shown in soybean (*Glycine max*; De Souza *et al.* (2022)).

The means with which to reduce the loss of photosynthetic light-use efficiency in crop plants might already exist in nature. Most temperate-zone crop species, alongside tropical crops species like rice (*Oryza sativa*), make use of the C_3 photosynthetic pathway, which is the original and ancestral photosynthetic pathway in vascular plants, with the alternative CAM and C_4 pathways having evolved as an adaptation to heat and drought, and low CO_2 levels. Due to several issues associated with the C_3 pathway compared to the C_4 pathway, the maximum photosynthesis rates commonly observed among C_3 species are generally lower than those of C_4 ones. Although the core mechanisms of photosynthesis are highly conserved (Leister, 2019; Shi *et al.*, 2005), natural variation in photosynthesis rates has been observed for major crops such as wheat (*Triticum aestivum*, Driever *et al.* (2014)), rice (Gu *et al.*, 2014, 2012), maize (*Zea mays*, Strigens *et al.* (2013)), soybean (Gilbert *et al.*, 2011), sorghum

(*Sorghum bicolor*, Ortiz *et al.* (2017)), as well as for the model species *Arabidopsis thaliana* (van Rooijen *et al.*, 2015, 2017). Much higher photosynthesis rates can be expected in species that are more ecologically specialised (van Bezouw *et al.*, 2019). Exceptionally high light-use efficiencies (and high assimilation rates) at high irradiance have been found previously in species growing in the Sonoran Desert, such as *Amaranthus palmeri*, *Chylismia claviformis*, *Eremalche rotundifolia*, and *Palafoxia linearis*. Although data collected on these species provided clues about the anatomical and physiological basis of their high photosynthesis rates (Werk *et al.*, 1983; Gibson, 1998), a comprehensive ecophysiological explanation of their phenotypes is still missing.

To understand the physiological and genetic basis of this more efficient photosynthesis at high irradiance, a suitable model species is needed. To date, of the handful of species showing high light-use efficiency that have been described (Ehleringer, 1985; Werk *et al.*, 1983), none would qualify as a model species due to a combination of complex genetics and difficulties in growing in laboratory conditions (e.g. difficult seed germination). Taking inspiration from *A. thaliana*, an attractive model species for high light-use efficiency would need to be easily grown in regular irradiance (typically up to 600 µmol m^{-2} s^{-1}) and high-light laboratory conditions; have a high-quality reference genome; be a diploid species capable of producing a large number of progeny (hundreds of seeds from a single mother plant) with a short generation time; germinate easily and have easily stored seed; and allow for both inbreeding and outcrossing (Somerville and Koornneef, 2002; Koornneef and Meinke, 2010).

Hirschfeldia incana (L.) Lagr.-Foss. is an excellent candidate that fulfils these requirements. *H. incana* is a thermophilous and nitrophilous annual species native to the Mediterranean basin and the Middle East, but currently widespread in most warm-temperate regions of the world (Siemens, 2011). It is generally self-incompatible and thus allogamous, but a degree of self-compatibility has been observed in natural populations (Lee *et al.*, 2004). Although it makes use of the C_3 pathway, *H. incana* has a very high photosynthesis rate at high irradiance (Canvin *et al.*, 1980), much higher than that of the C_3 crop species wheat (Driever *et al.*, 2014) and rice (Gu *et al.*, 2012), more in the range of C_4 species (Crafts-Brandner and Salvucci, 2002; Leakey *et al.*, 2006). Besides its exceptional physiological properties, *H. incana* is also an attractive model species for practical and genetic reasons. It shows fast and sustained growth in laboratory conditions and is a member of the Brassiceae tribe within the well-studied Brassicaceae family, allowing the use of many genetic and genomic resources developed for the model species *A. thaliana* and its close relatives *Brassica rapa* (Choi *et al.*, 2007; Kim *et al.*, 2009; The Brassica rapa Genome Sequencing Project Consortium, 2011; Belser *et al.*, 2018; Zhang *et al.*, 2018), *Brassica nigra* (Perumal *et al.*, 2020; Paritosh *et al.*, 2020), *Brassica oleracea* (Wang *et al.*, 2011; Liu *et al.*, 2014; Belser *et al.*, 2018), and *Brassica napus* (Bancroft *et al.*, 2011; Chalhoub *et al.*, 2014). Yet, *H. incana* has received little attention from the research community so far, being recognised mainly as a possible lead (Pb) hyperaccumulator (Auguy *et al.*, 2013; Fahr *et al.*, 2015; Auguy *et al.*, 2016) and for the ecological implications of its occurrence as a weed (Darmency and Fleury, 2000; Lee *et al.*, 2004; Sánchez-Yélamo, 2009; Liu *et al.*, 2013; Mira *et al.*, 2019).

Here we present a high-quality genomic assembly and gene set of *H. incana*. We expect these data to lay the foundation for studying photosynthetic light-use efficiency and improving this trait in crop species, through a process of candidate gene identification followed by phenotypic validation using genetic modification and/or gene editing. First, we directly compare the photosynthetic rate of *H. incana* at high irradiance to that of the Brassicaceae species *B. rapa*, *B. nigra*, and *A. thaliana* to affirm its high light-use efficiency. Second, we characterize how the *H. incana* genome differs from that of other members of the Brassicaceae family, specifically focusing on differences in numbers of gene copies. Finally, we report on whether such differences translate to differential expression of genes expected to mediate high light-use efficiency. Our work demonstrates how the assembly of *H. incana* serves as a valuable resource to elucidate the genetic basis of high photosynthetic performance and for studying the evolution of this trait in the Brassicaceae family.

3.2 Results

3.2.1 *Hirschfeldia incana* has an exceptionally high rate of photosynthesis

High photosynthesis rates have been reported for *H. incana* in 1980 (Canvin *et al.*, 1980). We performed new measurements in order to compare the performance of *H. incana* with that of close relatives and the well-established model species *Arabidopsis thaliana* (Figure 3.1, Table S1). Gross CO_2 assimilation rates, independent of CO_2 release by mitochondrial respiration and therefore a better indication of photosynthetic capacity than net photosynthesis rates, differed significantly between species (Table S2). The two *H. incana* accessions had the highest average gross CO_2 assimilation rates from 800 µmol m^{-2} s^{-1} to 2200 µmol m^{-2} s^{-1} irradiance, although only 'Burgos' had a statistically significant higher rate than the other species (Table S3). Net photosynthesis rates showed a similar trend, but larger differences in rates between the two *H. incana* genotypes (Figure S1, Table S3). Moreover, the two *H. incana* genotypes showed differences in rates of daytime dark respiration between them (R_d, Table S4).

3.2.2 A reference genome of *H. incana*

We assembled a scaffold-level reference genome of *H. incana* based on one genotype of the 'Nijmegen' accession, that was inbred for six generations and therefore expected to be substantially more homozygous than the natural accession 'Burgos', easing genomic assembly. Its haploid genome size was estimated to be 487 Mb, based on flow cytometry (Table S5). This estimate is smaller than the previously reported genome size estimates of *B. rapa* (529 Mb) and *B. nigra* (632 Mb) (Johnston *et al.*, 2005). Chromosome counts from root tip squashes showed seven pairs of chromosomes (2n=14) (Figure S2), consistent with previous reports (Anderson and Warwick, 1999; Siemens, 2011).

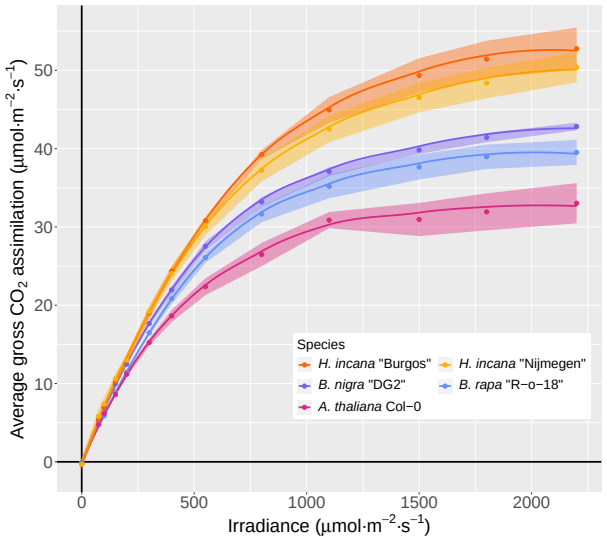

Figure 3.1: Two *H. incana* genotypes have a higher net CO_2 assimilation at high irradiance than genotypes of close relatives. Light-response curves for *H. incana*, *B. rapa*, *B. nigra*, and *A. thaliana* accessions adapted to high levels of irradiance. Each point represents the average net CO_2 assimilation value of three (*B. rapa*) or four leaves coming from independent plants. Error bars represent the standard error of means. Ribbons represent the standard error of the means. The lines indicate trends in gross assimilation for the various species and were obtained via LOESS smoothing.

We generated DNA sequencing data consisting of 56 Gb of PacBio long reads (115-fold genome coverage, based on the genome size estimate), 46 Gb of 10X Genomics synthetic long reads (94-fold coverage, referred to as “10X” from now on for brevity), and 33 Gb of Illumina paired-end short reads (68-fold coverage). In addition, we generated 7.5 Gb of RNA sequencing (RNA-seq) data from leaf tissue for annotation purposes. Summary statistics and accession numbers can be found in Table S6. A k-mer analysis of Illumina data resulted in a haploid genome size estimate of 325 Mb, with a low level of heterozygosity (1.2%).

Using a hybrid assembly strategy, we produced a nuclear genome assembly of 399 Mb of sequence in 384 scaffolds with an N50 of 5.1 Mb (Table 3.1, see Table S7 for the full report generated by QUAST (Gurevich *et al.*, 2013)). The assembly size is slightly larger than the genome size estimated from Illumina read k-mers (325 Mb), but smaller than the typical overestimate (Sun *et al.*, 2018) based on flow cytometry (487 Mb). Besides the nuclear genome, we assembled the mitochondrial and chloroplast genomes of *H. incana* into single sequences of 253 and 153 kb, and annotated the latter. The chloroplast assembly is typical for a Brassicaceae species, as it is nearly identical to chloroplast assemblies of *A. thaliana*, *B. rapa*, and *B. nigra* in terms of length and number of annotated genes (Table S8).

The assembly is near-complete and structurally consistent with the underlying

Table 3.1: Genomic properties of assemblies generated of *H. incana* 'Nijmegen' (this study), ***B. rapa* Chiifu 401-42** (Zhang *et al.*, 2018)**, and *B. nigra* Ni100** (Perumal *et al.*, 2020).

	H. incana	*B. rapa*	*B. nigra*
Technologies	PacBio, 10X, Illumina paired-end	PacBio, BioNano, Hi-C, Illumina mate-pair	Nanopore, Hi-C, genetic mapping
Size (Mb)	398.5	353.1	506
# scaffolds	384	1301	58
N50 (Mb)	5.1	4.4	60.8
Gaps (kb)	0.54	0.40	13
GC-content (%)	36.2	36.8	37.0
BUSCOs assembly*	96.2	97.7	97.0
BUSCOs annotation**	95.1 (80.2/14.9)	97.2 (84.2/13.0)	97.2 (81.9/15.3)
# protein-coding genes	32,313	46,250	59,852
# protein-coding transcripts	38,706	46,250	59,852
Repeat content (%)	49.4	37.5	54.0
Full-length LTR-RTs (%)	25.3	29.2	41.8

$*$ Complete BUSCOs assembly (%)

$**$ Complete BUSCOs annotation (single/duplicated) (%)

read data of *H. incana* 'Nijmegen' (Table S9). The high mapping rate of Illumina and 10X reads ($>$ 93%) suggest completeness, while the lower mapping rate of PacBio reads (81.5%) suggests some misassemblies or missing regions, likely repeats. The high mapping rate of RNA-seq reads (93.6%) again shows the gene space is near complete. We estimated the base-level error rate of the assembly to be 1 per 50 kb at most, based on variant calling using the mapped reads, resulting in 8,374 and 4,166 homozygous variants from the Illumina and 10X read alignments respectively.

We have annotated 32,313 gene models and 38,706 transcripts in the *H. incana* assembly (Table 3.1). This is a conservative annotation, based on filtering 64,546 initial gene models resulting from *ab initio*, protein alignment, and RNA-seq based predictions. Our filtering approach is more stringent than those used to generate the *B. rapa* and *B. nigra* annotations, which explains why we report a lower number of genes and transcripts for *H. incana* (Table 3.1) than for both *Brassica* species.

The annotation is expected to cover the large majority of the *H. incana* gene space. It contains 95.1% of 1,440 single-copy orthologs (BUSCOs) conserved in the Embryophyta plant clade, comparable to the percentages found for *B. rapa* and *B. nigra* (both 97.2%) (Table 3.1). The ratio of single to multiple copies is similar to that of *B. rapa* and *B. nigra* (Table 3.1), suggesting that the 14.9% of the BUSCOs present in multiple copies are true gene duplications shared by several species of the Brassiceae tribe. We additionally evaluated the completeness of the annotation by aligning protein sequences of *B. rapa* to the assembly and determining overlap between protein alignments and annotated genes. 30,552 out of the 37,387 protein alignments (81.7%) corroborate the annotation, as they completely or partially overlap with an annotated protein-coding gene. 2,570 (6.9%) of the protein alignments

completely or partially overlap with an annotated repeat, suggesting that the aligned *B. rapa* proteins correspond to transposable elements. The remainder of the *B. rapa* proteins completely or partially overlap with gene models that were filtered (3,945 or 10.6%) or do not overlap with any annotated element at all (320 or 0.9%), indicating a small number of genes that are potentially missing from the annotation. Based on these observations, we conclude that the *H. incana* assembly is mostly contiguous, correct, and complete, making it a solid foundation for comparative analyses with other Brassicaceae.

3.2.3 The genome of *H. incana* extensively diversified from that of *B. rapa* and *B. nigra*

We utilized our assembly to explore the genomic divergence between *H. incana*, *B. rapa*, and *B. nigra*, all members of the same Brassiceae tribe. A substantial degree of divergence is expected between the three species due to different processes of post-polyploid diploidization, i.e. the process in which polyploid genomes get extensively rearranged as they return to a diploid state (Mandáková and Lysak, 2018), following the ancient two-step genome triplication event shared by all Brassiceae (Lysak *et al.*, 2005; The Brassica rapa Genome Sequencing Project Consortium, 2011; He *et al.*, 2021). Part of this divergence may have facilitated the evolution of the exceptional rate of photosynthesis at high irradiance in *H. incana*.

We first assessed the phylogenetic relationship between *H. incana*, *B. rapa*, and *B. nigra* by constructing phylogenetic trees based on homologous nuclear and chloroplast genes, using *A. thaliana* as the outgroup. Both trees are congruent with each other and suggest that *H. incana* is more closely related to *B. nigra* than *B. rapa* (Figures 3.2a and 3.2b). This is corroborated by the median rate of synonymous substitutions between the syntenic orthologs (K_s) of the three species, which correspond to speciation events with an estimated time of 10.35 (*H. incana-B. nigra*) and 11.55 (*H. incana-B. rapa*) million years ago (mya) (Figure 3.2c), which were obtained by dividing the median K_s of each curve by the rate of 8.22×10^{-9} synonymous substitutions per year established for Brassicaceae species (Beilstein *et al.*, 2010). Our results are consistent with a previous phylogenetic analysis based on four intergenic chloroplast regions (Arias and Pires, 2012), but contradict a more recently constructed phylogeny of the Brassicaceae based on 113 nuclear genes (Huang *et al.*, 2016), possibly because the latter looks at single-copy genes only, while we also take into account multi-copy ones.

We determined rearrangements between the genomes of *H. incana-B. rapa* and *H. incana-B. nigra* by comparing the order of syntenic orthologs between their assemblies. On a small scale, most genomic regions of *H. incana* are syntenic (not rearranged) with *B. rapa* and *B. nigra*, as 77.7% and 81.0% of the genes of *H. incana* could be clustered in collinear blocks containing a minimum of four orthologous pairs of *H. incana-B. rapa* and *H. incana-B. nigra*, respectively. Gene order is less conserved when comparing larger blocks, indicating several rearrangements between the twenty largest scaffolds of *H. incana* (covering 43.6% of the assembly) and the

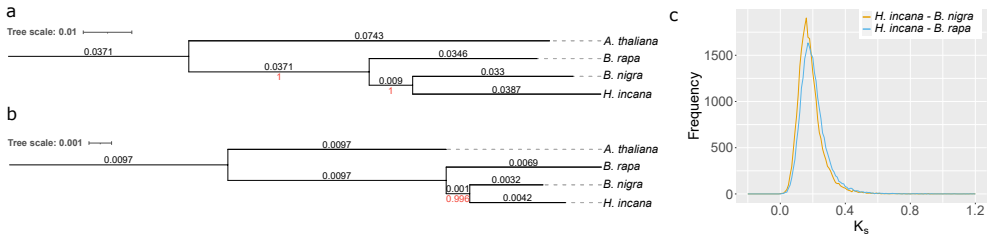

Figure 3.2: *Hirschfeldia incana* is more closely related to *Brassica nigra* than *Brassica rapa*. (a-b) Phylogenetic trees of *H. incana*, *B. rapa*, *B. nigra*, and *Arabidopsis thaliana* (outgroup), based on nuclear (a) and chloroplast (b) genes. Branch lengths (black) and bootstrap values (red) are displayed above and below each branch, respectively. **(c)** Distributions of the rates of synonymous substitutions between 25,127 and 26,137 orthologous gene pairs of *H. incana-B. rapa* and *H. incana-B. nigra*, respectively. Both distributions show a single peak corresponding to speciation events with an estimated time of 11.6 (*H. incana-B. rapa*) and 10.4 (*H. incana-B. nigra*) million years ago (mya).

chromosomes of the other two species (Figure 3.3a). For example, the two largest scaffolds of the *H. incana* assembly both contain inversions and/or translocations relative to their homologous chromosomes in *B. rapa* and *B. nigra* (Figures 3.3a and S3). For example, the two largest scaffolds of the *H. incana* assembly both contain inversions and/or translocations relative to their homologous chromosomes in *B. rapa* and *B. nigra*. A similar pattern of rearrangements of small collinear blocks was observed between the genomes of *B. rapa* and *B. nigra* in previous work (He *et al.*, 2021).

We further examined genomic differentiation between the three species by comparing their transposable element (TE) content. The assembly of *H. incana* consists of 49.4% repetitive elements (Table 3.1), of which most are long terminal repeat retrotransposons (LTR-RTs) (25.3% of the genome). These numbers are consistent with previous work that investigated the repeat content of the *H. incana* genome using genome skimming, and which reported a repeat content of 46.5% and LTR-RT content of 31.6% (Beric *et al.*, 2021). We specifically focused our analyses on LTR-RTs, as LTR-RT expansion and contraction has been previously identified as a major driver of genomic differentiation between Brassiceae (Xu *et al.*, 2018), even between different ecotypes of the same species (Cai *et al.*, 2020). The composition of LTR-RTs in the *H. incana* assembly differs from that of the *B. rapa* and *B. nigra* assembly, as the majority of LTR-RTs consist of Gypsy elements in *H. incana*, consistent with earlier work (Beric *et al.*, 2021), while Copia retrotransposons form the majority of LTR-RTs in the others (Figure 3.3b). Furthermore, the estimated insertion times of LTR-RTs vary between the three assemblies, as Gypsy and Copia elements in *H. incana* and *B. rapa* are predicted to have proliferated recently ($<$ 1 mya) (Figure 3.3c-d), while Gypsy elements in *B. nigra* show a more varied distribution of insertion times (Figure 3.3c). A possible explanation of this shift could be that the *B. nigra* assembly was generated using longer reads than those used for the assemblies of *H. incana* and

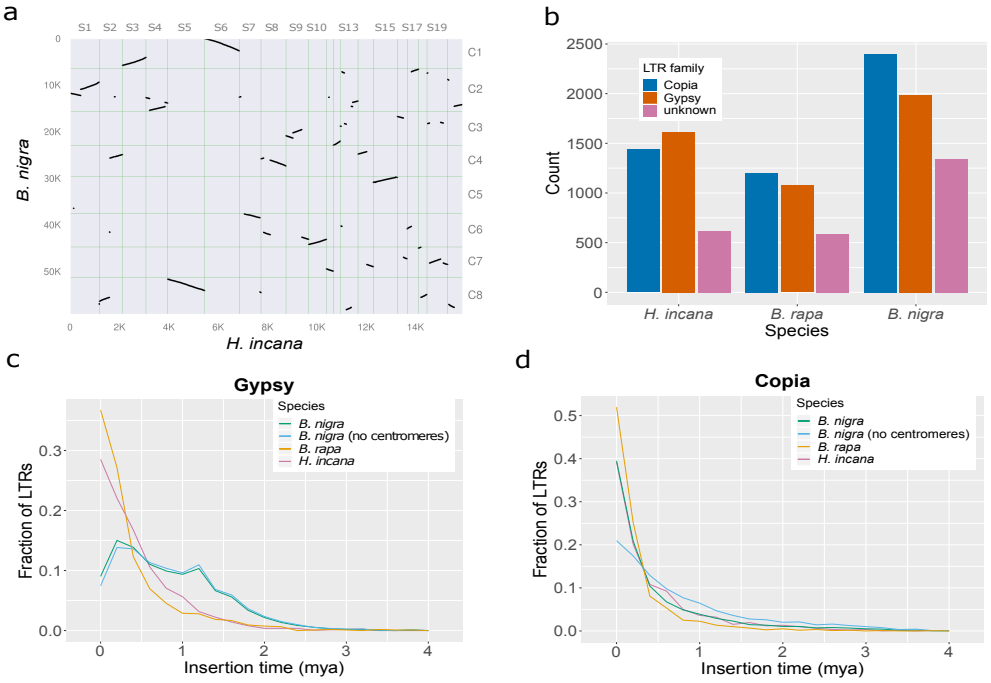

Figure 3.3: The genome of *Hirschfeldia incana* extensively diversified from that of *Brassica rapa* and *Brassica nigra*. **(a)** Orthologous syntenic blocks between the genomes of *H. incana* and *B. nigra*. Dots indicate pairs of syntenic orthologs that are found in the same order in both genomes according to sequence positions. Only the twenty largest scaffolds of *H. incana* (43.6% of the assembly) are shown for clarity. Axes labels correspond to the total number of genes annotated on the sequences (left and bottom) and identifiers of the scaffolds (top) or chromosomes (right). A dot plot visualizing orthologous syntenic blocks between *H. incana* and *B. rapa*, showing similar patterns, is found in Figure S3. **(b)** Frequency distribution of Long Terminal Repeat Retrotransposon (LTR-RT) families. LTR-RTs are classified as unknown if they contained elements of both Gypsy and Copia sequences and could thus not be reliably assigned to either of these families. **(c)** Frequency polygon (bin width = 0.2 mya) of the insertion times of Gypsy elements. **(d)** Frequency polygon (bin width = 0.2 mya) of insertion times of Copia elements.

B. rapa, enabling it to capture a larger proportion of the centromeric regions, but we found no evidence that this introduced a bias towards longer insertion times of Gypsy elements (Figure 3.3c).

Taken together, the breakdown of genomic synteny and divergence of LTR-RT content indicate that the genome of *H. incana* extensively diversified from that of *B. rapa* and *B. nigra* following their shared genome triplication event.

3.2.4 Gene copy number variation may contribute to high photosynthetic rates

Genomic differentiation can result in species-specific gains and losses of genes, which may explain the differences in photosynthetic light-use efficiency between *H. incana*, *B. rapa*, and *B. nigra*. Given that the three species all share the same ancient genome triplication event (Schranz *et al.*, 2006; He *et al.*, 2021), it is reasonable to assume that most differences originated through differential retention of duplicated genes, particularly those located in genomic blocks showing evidence of extensive fractionation since that event (He *et al.*, 2021). We investigated gene copy number variation between the three species, by clustering their annotated protein-coding genes with those of five other Brassicaceae species within (*Raphanus raphanistrum* and *Raphanus sativus*) and outside (*Aethionema arabicum*, *A. thaliana*, and *Sisymbrium irio*) the Brassiceae tribe into homology groups. The inclusion of *A. thaliana* allowed us to use its extensive genomic resources to functionally annotate the genes of other species. The other four species were included to put the analysis in a broader phylogenetic context. *A. arabicum* is part of the Aethionema tribe which diverged from the core group of the Brassicaceae family, thus allowing us to identify highly conserved genes. *S. irio* is part of a different tribe than *A. thaliana* (Sisymbrieae), that is more closely related to the Brassiceae tribe (Huang *et al.*, 2016), but did not undergo the ancient genome triplication. *R. raphanistrum* and *R. sativus* are part of the *Raphanistrum* genus within the Brassiceae tribe and thus represent another set of species that underwent the genome triplication shared by the whole tribe.

Our analysis resulted in 20,331 groups containing at least one *H. incana* gene (Table S10). The composition of the homology groups agrees with the currently established phylogeny of the Brassicaceae (Huang *et al.*, 2016), as groups containing *H. incana* genes share the fewest genes with *A. arabicum* (58.2%) and most genes with species part of the Brassiceae tribe (86.3-95.6%). *H. incana* has a low fraction of species-specific homology groups (3.4%) compared to the seven other species, which can be attributed to the stringent filtering of the predicted gene models.

We focused on a subset of 15,097 groups containing at least one gene of *A. thaliana* and one of *H. incana*, as these could be extensively annotated through the transfer of Gene Ontology (GO) terms from *A. thaliana* genes to their respective groups. According to the expectation that most genes quickly return to single-copy status following a whole genome duplication event (Li *et al.*, 2016), 70.2% of these groups contain a single gene of both *A. thaliana* and *H. incana*. Focusing on groups containing *A. thaliana* genes involved in photosynthesis (260 in total, Table S11), most contain a higher number of genes of *H. incana*, *B. rapa*, and *B. nigra*, compared to *A. thaliana* (Figure 3.4), consistent with the relatively higher photosynthetic light-use efficiency of the latter three (Figure 3.1). The higher efficiency of *H. incana* of all four species is not apparent from the gene copy number, as for most groups, *H. incana* contains the same or a lower number of copies, relative to *B. rapa* and *B. nigra*. This is not a result of our conservative filtering approach, as we explicitly retained putative photosynthesis-related genes during our filtering procedure (Methods S1). Besides photosynthesis-related genes, we also analysed copy numbers of a more general

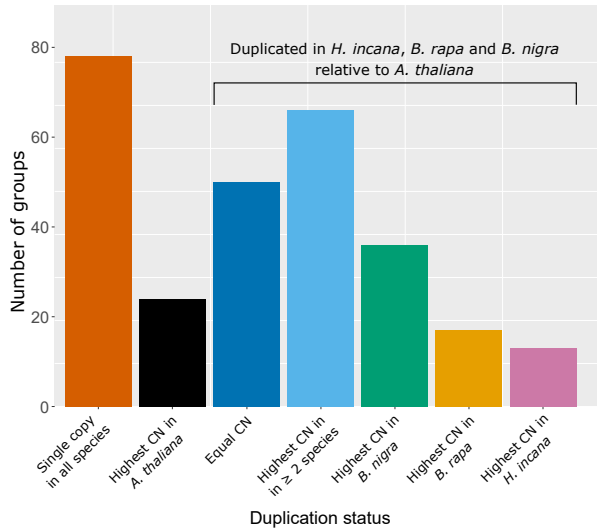

Figure 3.4: *Hirschfeldia incana* retained fewer duplicated copies of photosynthesis-associated genes than *Brassica rapa* and *Brassica nigra*. Bars show counts of homology groups containing genes associated with photosynthesis with different distributions of copy numbers (CN) in the four species (260 groups in total). For groups that contain a higher number of copies in *H. incana*, *B. rapa*, and *B. nigra* than in *Arabidopsis thaliana*, it has been indicated whether the same number of copies is found in all three species (equal CN), whether there are two or more species that contain a higher number of copies than the other(s) (highest CN in at least 2 species), or whether there is a single species containing the highest number of copies.

set. 4,901 homology groups contain genes of which the copy number in *H. incana* is higher than the one in *A. thaliana*, and equal to or higher than the ones in *B. rapa* and *B. nigra* (Table S12). We estimate that 74.4% of duplicated gene pairs in *H. incana* (16,788 of the 22,535 analysed pairs) were duplicated through whole-genome duplication, 1.8% through tandem duplication, and the remaining 23.6% through another mode of duplication. Given that the increased photosynthetic light-use efficiency of *H. incana*, relative to *A. thaliana*, *B. rapa*, and *B. nigra* is particularly pronounced at high levels of irradiance (Figure 3.1), genes annotated with GO terms associated with photosynthesis and/or photoprotection are of particular interest. The 4,901 homology groups contain ample examples of such genes (Table S12), although the groups were not significantly enriched for any GO term specifically linked to photosynthesis and/or photoprotection (Table S13).

As gene copy number variation can considerably affect expression levels (Żmieńko *et al.*, 2014), we hypothesized that retained copy number expansions of photosynthesis and photoprotection-associated genes in *H. incana*, *B. rapa* and *B. nigra* may aid the high photosynthetic capacities of these species (Figure 3.1). We therefore measured gene expression levels of nine genes for which there is inter-species

Table 3.2: Genes with inter-specific copy number variation of which expression was measured. All genes are annotated to a function in photosynthesis and/or photoprotection.

		Copy number			
Gene name	Reference	*A. thaliana*	*B. rapa*	*B. nigra*	*H. incana*
LHCA6	Peng *et al.* (2009)	1	1	1	3
ELIP1	Hutin *et al.* (2003)	1	3	3	3
SIGE/SIG5	Tsunoyama *et al.* (2004)	1	2	2	2
SIGD/SIG4	Favory *et al.* (2005)	1	2	2	3
BBX21	Crocco *et al.* (2018)	1	3	3	3
PETC	Maiwald *et al.* (2003)	1	2	2	2
ABC1K3	Martinis *et al.* (2013)	1	1	1	2
OHP2	Li *et al.* (2019)	1	2	2	3
CYFBP	Lee *et al.* (2008)	1	2	2	2

copy number variation in two contrasting light conditions (200 µmol m^{-2} s^{-1} and 1800 µmol m^{-2} s^{-1}), selecting genes with a function related to photosynthesis and/or photoprotection (Table 3.2). *A. thaliana*, the species with the lowest photosynthesis rates measured in this study, contains a single copy of each of the tested genes. For six genes, we observed a statistically significant positive correlation between gene expression level and gene copy number (Figure 3.5), with species showing higher or equal expression with an increasing number of copies (Figure S4). No such correlation was observed for the remaining three genes.

To test if the observed differences in gene expression are due to photosynthesis-related genes being more frequently upregulated in general in *B. rapa*, *B. nigra* and *H. incana*, compared to *A. thaliana*, we included nine additional genes in our experiment that are present in a single copy in all four species and involved in similar processes as the multi-copy genes. Although we find species-specific differences in expression for this set of genes, no consistently higher gene expression levels are found in *B. rapa*, *B. nigra* and *H. incana* compared to *A. thaliana* (Figure S5). Overall, our analyses suggest that the increased copy numbers of photosynthesis and photoprotection-associated genes in *H. incana*, *B. rapa* and *B. nigra*, relative to the *A. thaliana*, may contribute to their high photosynthetic efficiency, although this effect appears to not be specific to a particular species or level of irradiance.

3.3 Discussion

In this study, we generated a reference genome of *H. incana* to establish this species as a model for exceptional photosynthetic light-use efficiency at high irradiance. We find substantial differences in light-use efficiency, genomic structure, and gene content between *H. incana* and its close relatives. We discuss these results in terms of how they contributed to the evolution of the remarkable phenotype of *H. incana*.

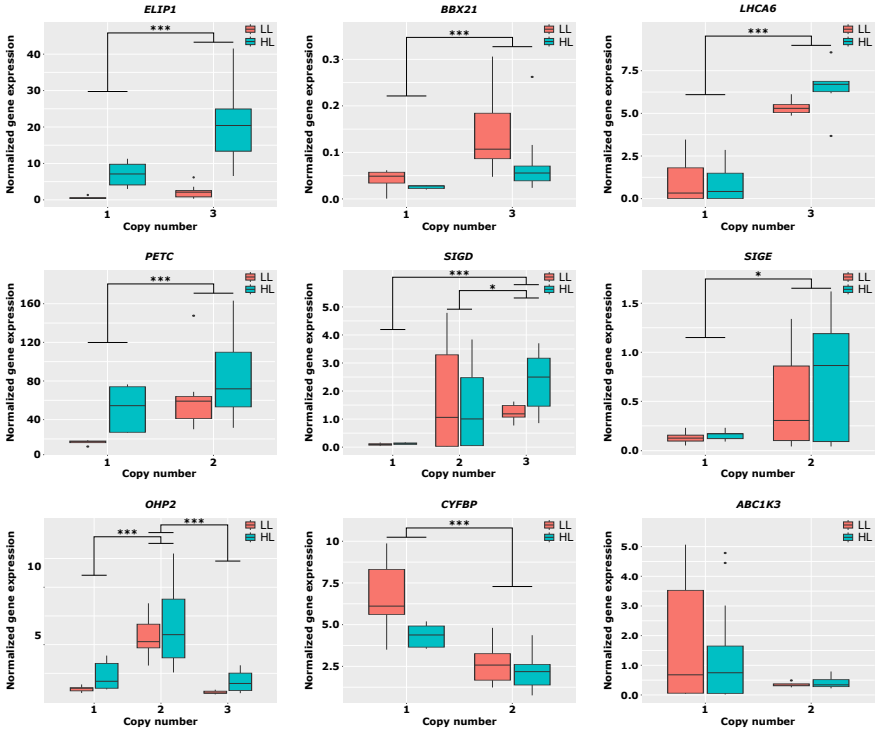

Figure 3.5: Copy numbers of photosynthesis- and photoprotection-associated genes correlate with expression level. Boxplots depict gene expression levels of *A. thaliana*, *B. rapa*, *B. nigra* and *H. incana* grown in 200 µmol m^{-2} s^{-1} and 1500 µmol m^{-2} s^{-1}. Gene expression levels were normalized against *H. incana* grown at 200 µmol m^{-2} s^{-1} and subsequently grouped per gene copy number. Titles of graphs indicate gene names based on the *A. thaliana* gene nomenclature. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Our results show an even higher photosynthetic light-use efficiency at high irradiance than previously reported for *H. incana* (Canvin *et al.*, 1980), with photosynthesis rates varying marginally between both accessions. Examination of a wider set of *H. incana* accessions may identify genotypes with larger differences in photosynthesis rates, that would allow a quantitative genetic approach to identify alleles conferring high photosynthesis rates. Our measurements imply that the photosynthetic rates of this C_3 are higher than the rates of photosynthesis of the C_4 crop maize (Crafts-Brandner and Salvucci, 2002; Leakey *et al.*, 2006) and and almost two times higher than those typically reported from key cereal crop species with a C_3 photosynthetic metabolism, such as wheat (Driever *et al.*, 2014) and rice (Gu *et al.*, 2012), respectively. Furthermore, these rates are higher than those of closely related Brassicaceae species *B. rapa*, *B. nigra*, and the more distantly related *A. thaliana*. The photosynthesis rates we measured in *B. rapa* are also higher than previously reported (Pleban

et al., 2018; Taylor *et al.*, 2020). Although the rates presented in this study were obtained from plants grown in controlled, favourable conditions and thus could be an overestimation of rates in natural environments, the magnitude of the differences suggests that the *H. incana* genome holds essential information for the improvement of photosynthetic light-use efficiency in crops.

The reference genome of *H. incana* generated in this study provides the means to elucidate the genetic basis of this plant's exceptional rate of photosynthesis and how it evolved in this species. We estimate that *H. incana* diverged 11.6 and 10.4 mya from *B. rapa* and *B. nigra*, respectively, consistent with an earlier study that used a smaller set of nuclear genes (Huang *et al.*, 2016). These time points are close to the reported time (11.5 mya) at which *B. rapa* and *B. nigra* diverged from each other (Perumal *et al.*, 2020) and the time at which the whole Brassicaceae family underwent a rapid radiation event (Franzke *et al.*, 2009). This event may have been mediated by the expansion of grass-dominated ecosystems in the region inhabited by Brassicaceae family members at that time, which created new open habitats that favoured rapid diversification (Franzke *et al.*, 2009). This expansion of grasslands is thought to have been driven by decreasing atmospheric CO_2 levels, increasing temperatures, and increasing aridity, which favoured the displacement of the then dominant C_3 plants by C_4 grasses (Edwards *et al.*, 2010). We argue that climatic changes also drove the evolution of the high photosynthetic rates observed in *H. incana*; grassland, i.e. non-forested ecosystems may have provided the ephemeral niches with high irradiances that favoured the evolution of high photosynthetic rates. Species with high photosynthetic rates are currently found in Mediterranean and desert ecosystems (Ehleringer, 1985; Werk *et al.*, 1983). The evolution of high rates of C_3 photosynthesis could therefore have paralleled the expansion of the C_4 photosynthesis pathway as an adaptation to low CO_2 levels and drought.

Our analyses suggest that the genome of *H. incana* extensively differentiated from that of *B. rapa* and *B. nigra* since their time of divergence through large genomic arrangements and differences in LTR-RT content. Previous analyses of natural *A. thaliana* accessions indicated that specific LTR-RT families show increased rates of proliferation in response to particular types of environmental stress (Baduel *et al.*, 2021), which may explain the species-specific amplification of Gypsy elements that we observed in *H. incana*. Such elements may have been retained because this particular LTR-RT family generally inserts outside of exons (Baduel *et al.*, 2021). We hypothesize that the differences in LTR-RT content between *H. incana*, *B. rapa*, and *B. nigra* were caused in part due to Gypsy elements being less efficiently purged from the genome of *B. nigra* than from those of the others. An increased rate of LTR-RT removal, based on the ratio of solo LTRs to intact LTR-RTs, has also been observed in *B. rapa* relative to *B. oleracea* and it was speculated that this is caused by the increased rate of genetic recombination in the former (Zhao *et al.*, 2013). Given that a similar negative correlation between local recombination rate and LTR-RT content was found in rice (Tian *et al.*, 2009), soybean (Du *et al.*, 2012), and eukaryotes in general (Kent *et al.*, 2017), the differences in predicted insertion times of Gypsy elements in *H. incana*, *B. rapa*, and *B. nigra* observed in this study may thus reflect different rates of genetic recombination in the three species. While it

has been suggested that changes in recombination rate can be adaptive, there is little empirical evidence that supports this (Ritz *et al.*, 2017). It would therefore be interesting to directly measure genome-wide rates of recombination of *H. incana*, *B. rapa*, and *B. nigra* and explore whether these are correlated with their rates of photosynthesis.

Further comparative analyses between the genomes of *H. incana*, *B. rapa*, *B. nigra*, and *A. thaliana* revealed numerous species-specific gains and losses of genes. For dosage-sensitive genes, such as those involved in transcriptional regulation, differences may not necessarily reflect adaptive selection. This category of genes was found to be consistently retained in multiple copies following polyploidy events across the Brassicaceae (Mandáková *et al.*, 2017) and a wide group of angiosperms (Li *et al.*, 2016), which is hypothesized to be due to dosage constraints (Edger and Pires, 2009). Differences in copy number of such genes may thus reflect different rates of relaxation of dosage balance constraints and subsequent loss of duplicates through time, which is a neutral process.

On the other hand, there is reason to believe that gene duplications contributed to the evolution of the high light-use efficiency of *H. incana*. Gene duplications have been identified as important drivers of plant evolution and differences in gene copy number between species are often enriched for adaptive evolutionary traits (Rizzon *et al.*, 2006; Suryawanshi *et al.*, 2016; Oh *et al.*, 2013; Dassanayake *et al.*, 2011). Moreover, RT-qPCR analysis of nine duplicated genes associated with photosynthesis and/or photoprotection showed that the expression levels of six of them correlate with gene copy number. In contrast, nine photosynthetic genes present in a single copy in all species did not show significantly increased expression of these genes in *H. incana*, *B. nigra*, and *B. rapa* compared to *A. thaliana*, indicating that photosynthetic genes are not overexpressed in the former three species in general. This supports a putative role for gene duplications in mediating the high light-use efficiency achieved by *H. incana*, *B. nigra* and *B. rapa*.

The most striking genes of which copy number correlated with gene expression are *LHCA6* and *ELIP1*, involved in response to high light and having the highest expression in *H. incana* growing under high light (Figure S4). *LHCA6* encodes a light-harvesting complex I (LHCI) protein of photosystem I (PSI), that together with LHCA5 is required to form a full-size NAD(P)H dehydrogenase (NDH)-PSI supercomplex (Peng *et al.*, 2009). Higher expression of *LHCA6* might help the formation of the NDH-PSI complex, thought to help stabilising NDH under high irradiance conditions. In turn, NDH has proposed roles in both supporting the Calvin-Benson cycle's activity (Harbinson *et al.*, 2022)) and photoprotection by preventing overreduction at high light intensities (Munekage *et al.*, 2004). *ELIP1* encodes for proteins with a proposed role in photoprotection, which is associated with high light stress (Norén *et al.*, 2003; Heddad *et al.*, 2006; Youssef *et al.*, 2010). Increased expression of this genes is expected to make the photosynthetic apparatus of *H. incana* more resistant to photoinhibition at high levels of irradiance. While the *H. incana* genome harboured the highest number of copies of *LHCA6* when compared to the genomes of *A. thaliana*, *B. rapa*, and *B. nigra*, this is not the case for *ELIP1*, for which *H. incana*, *B. nigra* and *B. rapa* all have three copies as opposed to the single copy of *A. thaliana*. Therefore,

although we can propose a role for gene duplications in achieving higher light-use efficiency, the exact nature of this role still remains unclear as it appears to not be completely dependent on species or light treatment.

While our gene expression analysis provides several promising leads, it only offers a glimpse of what may contribute to the high photosynthetic light-use efficiency of *H. incana*. Besides the nine genes included in this analysis, we identified many more genes with a high copy number in *H. incana* that warrant further investigations. Such investigations should not limit themselves to core photosynthetic genes, as *H. incana* can only attain high photosynthetic light-use efficiency through changes in many other traits that are outside the chloroplast, such as leaf architecture affecting mesophyll conductance to CO_2, the synthesis of carbohydrates in the cytosol, the transport of carbohydrates from the leaf, the uptake from the soil and the supply of nitrogen and other minerals to the leaf, the abundance and distribution of different leaf pigments, and (photo)respiration. Nor should they include duplicated genes only, as it is striking that *H. incana* shows a better high light-use efficiency than *B. rapa* and *B. nigra*, though it contains fewer photosynthesis-related genes than the latter two species. This points towards alternative scenarios in which adaptation of *H. incana* photosynthesis to high levels of irradiance occurred through regulation of expression of one copy of the photosynthesis-related genes, which relaxed selection on duplicate retention or even encouraged loss of duplicate copies, or through other traits, as described above.

To elucidate the exact genetic mechanisms underlying the high light-use efficiency of *H. incana*, a natural follow-up to this study is to perform comparative transcriptomic analyses of leaves of *H. incana*, *B. rapa*, and *B. nigra* under a range of different levels of irradiance and at different developmental stages. Genes that show copy number variation and are differentially expressed between *H. incana* and the latter two species, such as *LHCA6*, would then be prime candidates to further test for potential causality through e.g. knock-out mutant analysis. As previous work has shown that it is possible to cross distantly related Brassicaceae species (Katche *et al.*, 2019), a useful approach to further pinpoint the causal genes is to establish a genetic mapping population between *H. incana* and a Brassicaceae species with regular light-use efficiency and perform quantitative trait locus analyses of photosynthetic traits segregating within the population. It would also be useful to expand comparative genome and transcriptome analyses to plant species outside of the Brassicaceae clade that show high photosynthetic light-use efficiency, such as the aformentioned *A. palmeri*, *C. claviformis*, *E. rotundifolia*, and *P. linearis*. Such expanded analyses could be informative for instance, to investigate amino acid substitutions or lateral gene transfer specific to species with high photosynthetic light-use efficiency. Furthermore, transcriptomic data may indicate genes showing differences in expression between such species and those that are less efficient, providing further insight into which genes contribute to the evolution of this trait.

3.4 Conclusions

H. incana has an exceptional rate of photosynthesis at high irradiance. We generated a near-complete reference genome of this species and found evidence suggesting that its exceptional rate evolved through differential retention of duplicated genes. Taken together, our work provides several promising leads that may explain the high photosynthetic light-use efficiency of *H. incana* and we expect the reference genome generated in this study to be a valuable resource for improving this efficiency in crop cultivars.

3.5 Materials and Methods

Plant material

Hirschfeldia incana accessions 'Nijmegen' and 'Burgos' were used. 'Nijmegen' is an inbred line ($>$ six rounds of inbreeding) originally collected in Nijmegen, The Netherlands. Seeds of 'Burgos' were originally collected near Burgos, Spain. Furthermore, *Brassica nigra* accession 'DG2', sampled from a natural population near Wageningen, The Netherlands, the *Brassica rapa* inbred line 'R-o-18' (Stephenson *et al.*, 2010; Bagheri *et al.*, 2012), and the *Arabidopsis thaliana* Col-0 accession were used.

Measurements of photosynthesis rates

Seeds of *H. incana* 'Nijmegen', *H. incana* 'Burgos', *B. rapa* 'R-o-18', *B. nigra* 'DG2', and *A. thaliana* Col-0 were sown in 3 L pots filled with a peat-based potting mixture. Plants were grown in a climate chamber with a photoperiod of 12 hours and day and night temperatures of 23 and 20 °C, respectively. Humidity and CO_2 levels were set at 70% and 400 ppm. The chamber was equipped with high-output LED light modules (VYPR2p, Fluence by OSRAM). Plants were watered daily with a custom nutrient solution (0.6 mM NH_4^+, 3.6 mM K^+, 2 mM Ca^{2+}, 0.91 mM Mg^{2+}, 6.2 mM NO_3^-, 1.66 mM SO_4^{2-}, 0.5 mM P, 35 µM Fe^{3+}, 8 µM Mn^{2+}, 5 µM Zn^{2+}, 20 µM B, 0.5 µM Cu^{2+}, 0.5 µM Mo^{4+}). The seeds were germinated at an irradiance of 300 µmol m^{-2} s^{-1}, and the same irradiance was maintained to let seedlings establish. On day 14, 21, and 25 after sowing, the irradiance was raised to 600, 1200, and 1800 µmol m^{-2} s^{-1}, respectively.

The photosynthetic metabolism of young, fully expanded leaves developed under 1800 µmol m^{-2} s^{-1} of light was measured with a LI-COR 6400xt portable photosynthesis system (LI-COR Biosciences) equipped with a 2 cm^2 fluorescence chamber head. "Rapid" descending light-response curves were measured between 30 and 35 days after sowing to accommodate differences in growth rates of the different species on one leaf from four *H. incana* 'Nijmegen', *H. incana* 'Burgos', *B. nigra* 'DG2', and *A. thaliana* Col-0 plants, and three *B. rapa* 'R-o-18' plants. The net assimilation rates of the plants were measured at thirteen different levels of irradiance ranging from 0 to 2200 µmol m^{-2} s^{-1}. During measurements, leaf temperature was kept constant at

25 ℃ and reference CO_2 concentration was kept at 400 µmol mol^{-1}. Water in the reference air flux was regulated in order to achieve vapour-pressure deficit values comprised between 0.8 and 1.2 kPa.

Light response curve parameters (A_{max}: net CO_2 assimilation at saturating irradiance, φ: apparent quantum yield of CO_2 assimilation, R_d: daytime dark respiration rate, and θ: curve convexity) were estimated for each species through non-linear least squares regression of a non-rectangular hyperbola (Marshall and Biscoe, 1980) with the R package "photosynthesis" (version 2.0.0) (Stinziano *et al.*, 2020). An indication of gross assimilation rates for each species was subsequently generated by adding the daytime dark respiration rate (R_d) estimated for each species to the species' net assimilation rates.

Differences in net and gross assimilation rates were tested at each light level of the light-response curve with a one-way ANOVA on the "genotype" experimental factor. Pairwise comparisons between the assimilation rates of the different genotypes at each light level were subsequently performed and tested with the Tukey-Kramer extension of Tukey's range test. The p-value threshold for statistical significance was set at α = 0.05.

Flow cytometry

Leaf samples of the *H. incana* genotypes 'Burgos' and 'Nijmegen' and *A. thaliana* Col-0 were analysed for nuclear DNA content by flow cytometry (Plant Cytometry Services B.V., Didam, the Netherlands). Seven, three and five biological replicates were measured over separate rounds of analysis for *H. incana* 'Nijmegen' *H. incana* 'Burgos', and *A. thaliana* Col-0, respectively. Nuclei were extracted from leaf samples following the method by Arumuganathan and Earle (1991), and stained with 4',6-diamidino-2-phenylindole (DAPI). The DNA content of nuclei relative to that of the reference species *Monstera deliciosa* was determined on a CyFlow Ploidy Analyser machine (Sysmex Corporation, Kobe, Japan). A haploid flow cytometry estimate of 157 Mb was used for *A. thaliana*, resulting from comparisons of nuclear DNA content of this species and other model organisms (Bennett *et al.*, 2003). Haploid genome size estimates for the *H. incana* genotypes were obtained by multiplying the *H. incana*-to-*M. deliciosa* ratio by the haploid *A. thaliana* estimate and dividing this product by the average *A. thaliana*-to-*M. deliciosa* ratio.

Chromosome counting

Root tips (approximately 1 cm long) were collected from young, fast-growing rootlets of multiple *H. incana* 'Nijmegen' plants and pre-treated for 3 h at room temperature with a 0.2 mM 8-hydroxyquinoline solution. After pre-treatment, the 8-hydroxyquinoline solution was replaced with freshly prepared Carnoy fixative (1:3 (v/v) acetic acid - ethanol solution) and maintained at room temperature for half a day. Root tips were then rinsed with 70% ethanol for three times to remove remaining fixative and stored in 70% ethanol at 4 ℃ until further use. Prior to slide preparation, root tips were rinsed

twice in Milli-Q (MQ) water before adding 1:1 solution of a pectolytic enzymatic digestion solution (1% Cellulase from *Trichoderma*, 1% Cytohelicase from *Helix Promatia*, 1% Pecolyase from *Aspergillus japonicus*) and 10 mM citric buffer. After one hour incubation at 37 °C, the enzymatic digestion solution was replaced by MQ water. The digested root tips were spread in 45% acetic acid over microscopy slides on a hot plate set at 45 °C, cells were fixed with freshly prepared Carnoy fixative, dried, and stained with DAPI dissolved in Vectashield mounting medium (Vector Laboratories Inc., Burlingame, U.S.). Slides were imaged with an Axio Imager.Z2 fluorescence optical microscope coupled with an Axiocam 506 microscope camera (Carl Zeiss AG, Oberkochen, Germany) at 63x magnification. Chromosome numbers were counted in metaphase mitotic cells and averaged to obtain the reported number.

DNA and RNA isolation

Genomic DNA was extracted from *H. incana* 'Nijmegen' samples using a protocol modified from Chang *et al.* (1993). The modifications consisted of adding 300 µL β-mercaptoethanol to the extraction buffer just before use. We added 0.7% isopropanol to the supernatant instead of 10 M LiCl and then divided the total volume into 1 mL aliquots for subsequent extractions. The pellet was dissolved in 500 µL of SSTE which was preheated to 50 °C before use. The final pellets were dissolved in 50 µL MQ water and then pooled at the end of the extraction process. DNA used for Illumina and 10X Genomics sequencing was extracted from flower material, while leaf material was used for the PacBio sequencing, all originating from the same plant.

Total RNA was extracted from leaf material of *H. incana* 'Nijmegen' from a different plant than the one used for the DNA isolations with the Direct-zol RNA mini-prep kit (Zymo Research, Irvine, U.S.A.) according to the company's instructions and then subjected to a DNAse (Promega Corporation, Madison, U.S) treatment at 37 °C for one hour.

Generation of sequencing data

Sequencing of total-cellular DNA of *H. incana* 'Nijmegen' was performed by GenomeScan B.V. (Leiden, The Netherlands). A total of seven SMRT cells were used for sequencing on the Pacific Biosciences Sequel platform. Short read Illumina and 10X Genomics libraries with an insert size of approximately 500-700 bp were prepared with the NEBNext Ultra DNA Library Prep kit for Illumina and 10X Genomics Chromium™ Genome v1 kit, respectively. These libraries were sequenced using the Illumina X10 platform (2 x 151 bp). RNA paired-end sequencing libraries with an average insert size of 254 bp were prepared using the Illumina TruSeq RNA sample prep kit with polyA mRNA selection and sequenced using the Illumina HiSeq 2500 platform (2 x 125 bp).

k-mer analysis

A histogram of k-mer frequencies of Illumina reads predicted to be of nuclear origin (see Methods S1) was generated using Jellyfish (v2.2.6) (Marçais and Kingsford, 2011), using a k-mer length of 21. The resulting histogram was provided as input to Genomescope (v1.0.0) (Vurture *et al.*, 2017) to estimate genome size and heterozygosity.

Genomic assembly and annotation

The genome assembly and annotation process is more extensively described in Methods S1. In short, we generated an initial assembly based on the PacBio data only with Canu (Koren *et al.*, 2017) and used it to bin the PacBio, 10X, and Illumina reads according to whether they originated from nuclear, organellar, or contaminant DNA. The bins were used to separately assemble the nuclear and organellar genomes, yielding a nuclear assembly consisting of hundreds of contigs and mitochondrial and chloroplast assemblies that were both represented by a single sequence. Nuclear contigs representing alternative haplotypes were removed using purge_dups (Guan *et al.*, 2020), after which ARKS (Coombe *et al.*, 2018) was used to scaffold the remaining contigs using the 10X data. Scaffolds were polished using Arrow (https://github.com/PacificBiosciences/gcpp) and Freebayes (Garrison and Marth, 2012), followed by a manual filtering step to obtain the final nuclear assembly.

Repeats in the assembly were masked using RepeatMasker (Smit *et al.*, 2015) in combination with RepeatModeler2 (Flynn *et al.*, 2020) before starting the annotation procedure. Nuclear genes were annotated by using EvidenceModeler (Haas *et al.*, 2008) to generate consensus models of ab initio gene predictions, alignments of proteins from closely and distantly related plant species, and transcripts assembled from RNA-seq data. These models were manually filtered to obtain a final set of protein-coding genes.

Used datasets for comparative genome analyses

We mainly focused the comparative genome analyses on *H. incana*, *B. nigra*, and *B. rapa*, three species of the Brassiceae tribe of which all members underwent an ancient genome triplication (Lysak *et al.*, 2005; The Brassica rapa Genome Sequencing Project Consortium, 2011). For comparative gene analyses, we extended this group with the Brassicaceae species *Arabidopsis thaliana*, *Aethionema arabicum*, *Sisymbrium irio*, *Raphanus raphanistrum*, and *Raphanus sativus*. The latter two *Raphanus* species are also part of the Brassiceae tribe. Version numbers and locations of all genomes are listed in Table S14.

Analysis of pairwise gene synteny and long terminal repeat retrotransposons (LTR-RTs) in *H. incana*, *B. rapa*, and *B. nigra*

Analyses of pairwise gene synteny between scaffolds of *H. incana* and chromosomes of *B. rapa* and *B. nigra* were performed using the JCVI library (https://github.com/tanghaibao/jcvi) (v1.0.5) in Python. Orthologs were identified through all-vs-all alignment of genes with LAST (Kiełbasa *et al.*, 2011), retaining reciprocal best hits only (C-score of at least 0.99). Hits were filtered for tandem duplicates (hits located within 10 genes from each other) and chained using the Python implementation of MCScan (Tang *et al.*, 2008) to obtain collinear blocks containing at least four pairs of syntenic genes. Visualizations of collinearity between genomic assemblies were generated using custom scripts of JCVI.

K_s values of syntenic gene pairs were computed using the ks module of JCVI. Protein sequences of pairs were aligned against each other using MUSCLE (v3.8.1) (Edgar, 2004), after which PAL2NAL (v14) (Suyama *et al.*, 2006) was used to convert protein alignments to nucleotide ones. K_s values for each pair were computed from the nucleotide alignments using the method of Yang and Nielsen (2000) implemented in PAML (Yang, 2007) (v4.9). Times of divergence between species were estimated by dividing the median of the distributions of their K_s values by the rate of 8.22×10^{-9} synonymous substitutions per year that was established for Brassicaceae species based on extrapolation from the ancient triplication event in the Brassica clade (Beilstein *et al.*, 2010).

Putative LTR-RTs were identified using LTRharvest (v1.6.1) (Ellinghaus *et al.*, 2008) and LTR_finder (v1.1) (Xu and Wang, 2007), after which LTR_retriever (v2.9.0) (Ou and Jiang, 2018) was run with default parameters to filter and combine the output of both tools into a high confidence set. LTR_retriever was also used to provide estimates of the insertion time of each LTR-RT. Parameters of LTRharvest and LTR_finder were set as recommended in the LTR_retriever documentation. Centromeric regions of the *B. nigra* assembly were obtained from Table S21 of the manuscript describing the assembly (Perumal *et al.*, 2020).

Phylogenetic analysis of *H. incana*, *B. rapa*, and *B. nigra*

The longest isoforms of the nuclear genes of *H. incana*, *B. rapa*, *B. nigra*, and *A. thaliana* (outgroup) were provided to Orthofinder (version 2.3.11) (Emms and Kelly, 2019) to generate phylogenetic species trees. Orthofinder was run using the multiple sequence alignment (MSA) workflow with default parameters. The same analysis was performed using chloroplast genes. Trees were visualized using iTOL (version 6.3) (Letunic and Bork, 2021).

Comparative gene ontology analysis of eight Brassicaceae species

The longest isoforms of the genes of all eight Brassicaceae species described in the section “Used datasets for comparative genome analyses” were extracted using

AGAT (version 0.2.3) (https://github.com/NBISweden/AGAT) and clustered into homology groups using the "group" function of Pantools version 2 (Sheikhizadeh Anari *et al.*, 2018) with a relaxation parameter of 4. Groups were assigned GO slim terms of their associated *A. thaliana* genes (obtained from arabidopsis.org/download_files /GO_and_PO_Annotations/Gene_Ontology_Annotations/TAIR_GO_slim_categories.txt (last updated on 2020-07-01)) and GO terms assigned to protein domains of associated *H. incana*, *B. rapa*, and *B. nigra* genes using InterProScan (version 5.45-72.0) (Jones *et al.*, 2014) (ran using the Pfam and Panther databases only). GO term enrichment tests were performed using the Fisher Exact test and the Benjamini-Hochberg method for multiple testing correction (Benjamini and Hochberg, 1995). *A. thaliana* genes were considered to be involved in photosynthesis, if they fulfilled one of the following conditions:

- Annotated with one of the following GO terms: "photosynthesis", "electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity", or "electron transporter, transferring electrons within the noncyclic electron transport pathway of photosynthesis activity";
- Included in the KEGG pathways ath00195 (Photosynthesis), ath00710 (Carbon fixation in photosynthetic organisms), and ath00196 (Photosynthesis - Antenna Proteins);
- Protein products have been assigned the keyword "Photosynthesis" in the Swiss-Prot database.

The same criteria were used to retain photosynthesis-related genes of *H. incana* while filtering the gene annotation of the assembly (see Methods S1).

Investigating the mode of duplicated genes in *H. incana*

Dupgen_finder (Github commit hash 8001838) (Qiao *et al.*, 2019) was run with default parameters to determine the mode of duplication for duplicated gene pairs in *H. incana*, using the genome of *A. thaliana* as an outgroup to detect pairs duplicated through whole-genome duplication. Pairs were allowed to be assigned to a single category only. Input files containing alignments of the protein sequences of *H. incana* aligned to themselves and those of *A. thaliana* were prepared using DIAMOND (version 0.9.14) (Buchfink *et al.*, 2015).

Analysis of gene expression under high and low irradiance

Seeds of *H. incana* 'Nijmegen', *B. nigra* 'DG2', *B. rapa* 'R-o-18', and *A. thaliana* Col-0 were germinated on a peat-based potting mixture for nine days under an irradiance of 200 µmol m^{-2} s^{-1}. Twelve seedlings per species were then transferred to 2 L pots filled with a peat-based potting mixture enriched with perlite and 2.5 g/L Osmocote® Exact Standard 5-6M slow-release fertiliser (ICL Specialty Fertilizers, Geldermalsen, The Netherlands).

Plants were germinated and grown in a climate chamber with a photoperiod of 12 hours and day and night temperatures of 23 and 20 ℃, respectively. Humidity and CO_2 levels were set at 70% and 400 ppm. The chamber was equipped with high-output LED light modules (VYPR2p, Fluence by OSRAM, Austin, U.S.). Six plants per species were assigned to a high light (HL) treatment of 1800 µmol m^{-2} s^{-1} and the remaining six to a low light (LL) treatment of 200 µmol m^{-2} s^{-1}. Irradiance uniformity was very high for both HL and LL treatments, with an u2 value of 0.93. Plant positions were randomised across growing areas. Plants were watered with the same custom nutrient solution as the one used in the measurements of photosynthesis rates, daily for the LL treatment and twice a day for the HL treatment.

Twenty-eight days after sowing, one young fully adapted leaf from each plant was selected, excised, and snap-frozen in liquid nitrogen. Leaf samples were crushed with a mortar and pestle cooled with liquid nitrogen and further homogenised with glass beads for 2 min at 30 Hz in a MM300 Mixer Mill (Retsch GmbH, Haan, Germany). Total RNA was extracted with the RNeasy Plant Mini Kit (QIAGEN N.V., Venlo, The Netherlands) according to manufacturer's instructions and then subjected to a RQ1 DNAse treatment (Promega Corporation, Madison, U.S.) at 37 ℃ for 30 minutes. We validated the total removal of DNA by means of a no-reverse transcriptase PCR reaction on all RNA samples. The RNA quality was assessed for purity (A260/A280) with a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, U.S.) and for possible RNA degradation by means of a visual inspection of the RNA on a 1% agarose gel. cDNA was then synthesized from 2 µg total RNA (measured by spectrophotometer) with the SensiFAST™ cDNA Synthesis Kit (Meridian Bioscience, Cincinnati, U.S.) according to manufacturer's instructions.

To examine the expression of both single-copy and multi-copy photosynthesis and/or photoprotection-related genes (Table S15), species-specific RT-qPCR primers were designed with the following criteria: the PCR fragment size had to range between 80 and 120 bp, the maximum difference in melting temperature between primers of the same pair had to be 0.5 ℃, and overall melting temperatures had to be comprised between 58 and 62 ℃. Primers were designed to target a region of the gene as similar as possible in all species. Additionally, for multi-copy genes, the primer pair had to bind to all copies of a particular gene in one species. RT-qPCR reactions were performed with SYBR green on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories Inc., Hercules, U.S.). The efficiency of each designed primer set was assessed by means of a standard curve, and only primer sets with efficiencies ranging between 90% and 110% were used. All primer sequences can be found in Table S16.

Gene expression was normalized to the reference genes *ACT2*, *PGK*, *UBQ7* and *APR* (Joseph *et al.*, 2018; Løvdal and Lillo, 2009) using the delta-Ct (dCt) method (Livak and Schmittgen, 2001). Normalized gene expression values were calculated as 2^{-dCt}. For the statistical analysis, we performed two-way ANOVA on the dCt values with the copy number and light treatment as grouping variables for the multi-copy genes and species, and light treatment as grouping variables for the single copy genes. A Kenward-Roger approximation for the degrees of freedom was used and a post-hoc test was subsequently performed with Tukey's range test, with the

significance threshold set at ($\alpha = 0.05$).

Author contributions

JH and MGMA initiated the research on the genomic basis of high photosynthesis of *H. incana*. MGMA generated the inbred line of *H. incana* 'Nijmegen'. FG grew plants for the gene expression experiment and performed measurements of photosynthesis rates, flow cytometry experiments, and subsequent analyses of experimental data. FRM performed RNA extractions, cDNA synthesis, and qPCR reactions for the gene expression experiment. FG and VC selected genes and designed primers. RB analysed the results of the gene expression experiment. RB and FB performed DNA and RNA extractions for genome sequencing and assembly. RYW and SS generated strategies for the genome assembly and annotation, which were applied by RYW. RYW and IvdH performed comparative analyses of Brassicaceae genomes. RH was involved in genome annotation. MES helped interpreting the results of the comparative analyses and was involved in drafting the manuscript. JH was involved in the design and interpretation of experiments that were performed to measure photosynthesis rates and in drafting the manuscript. FG, RYW, RB, DDR, MGMA, and SS were majorly involved in overall experimental design and preparing the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We gratefully acknowledge Pádraic Flood for collecting the *H. incana* 'Nijmegen' accession and Graham Taylor for propagating it to perform the first gas exchange measurements. We thank Carlos Alonso Blanco for collecting the *H. incana* 'Burgos' accession, and we thank Niccolò Bassetti and Nina Fatouros for collecting the *B. nigra* 'DG2' accession. We thank the Unifarm staff at the Nergena greenhouse and Radix Klima climate chambers unit for taking excellent care of our plants. We thank Steven Driever for providing valuable insights on the modelling of light–response curves, José van de Belt for her help on chromosome counting, and the Applied Bioinformatics group of Wageningen Plant Research for performing RNA sequencing.

References

Anderson, J. K. and Warwick, S. I. (1999). Chromosome number evolution in the tribe *Brassiceae* (*Brassicaceae*): evidence from isozyme number. *Plant Systematics and Evolution*, **215**(1-4), 255–285.

Arias, T. and Pires, J. C. (2012). A fully resolved chloroplast phylogeny of the brassica crops and wild relatives (Brassicaceae: Brassiceae): Novel clades and potential taxonomic implications. *Taxon*, **61**(5), 980–988.

Arumuganathan, K. and Earle, E. (1991). Estimation of nuclear DNA content of plants by flow cytometry. *Plant Molecular Biology Reporter*, **9**(3), 229–241.

Auguy, F., Fahr, M., Moulin, P., Brugel, A., Laplaze, L., El Mzibri, M., Filali-Maltouf, A., Doumas, P., and Smouni, A. (2013). Lead tolerance and accumulation in *Hirschfeldia incana*, a mediterranean brassicaceae from metalliferous mine spoils. *PLoS One*, **8**(5), e61932.

Auguy, F., Fahr, M., Moulin, P., El Mzibri, M., Smouni, A., Filali-Maltouf, A., Béna, G., and Doumas, P. (2016). Transcriptome changes in *Hirschfeldia incana* in response to lead exposure. *Frontiers in Plant Science*, **6**, 1231.

Baduel, P., Leduque, B., Ignace, A., Gy, I., Gil, J., Loudet, O., Colot, V., and Quadrana, L. (2021). Genetic and environmental modulation of transposition shapes the evolutionary potential of *Arabidopsis thaliana*. *Genome Biology*, **22**(1), 138.

Bagheri, H., El-Soda, M., van Oorschot, I., Hanhart, C., Bonnema, G., Jansen-van den Bosch, T., Mank, R., Keurentjes, J., Meng, L., Wu, J., *et al.* (2012). Genetic analysis of morphological traits in a new, versatile, rapid-cycling *Brassica rapa* recombinant inbred line population. *Frontiers in Plant Science*, **3**, 183.

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., Baker, D., Long, Y., Meng, J., Wang, X., *et al.* (2011). Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature Biotechnology*, **29**(8), 762.

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R., and Mathews, S. (2010). Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, **107**(43), 18724–18728.

Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F.-C., Falentin, C., Genete, M., Berrabah, W., Chèvre, A.-M., Delourme, R., *et al.* (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nature Plants*, **4**(11), 879–887.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**(1), 289–300.

Bennett, M. D., Leitch, I. J., Price, H. J., and Johnston, J. S. (2003). Comparisons with *Caenorhabditis* (∼100 Mb) and *Drosophila* (∼175 Mb) using flow cytometry show genome size in *Arabidopsis* to be ∼157 Mb and thus ∼25% larger than the Arabidopsis Genome Initiative estimate of ∼125 Mb. *Annals of Botany*, **91**(5), 547–557.

Beric, A., Mabry, M. E., Harkess, A. E., Brose, J., Schranz, M. E., Conant, G. C., Edger, P. P., Meyers, B. C., and Pires, J. C. (2021). Comparative phylogenetics of repetitive elements in a diverse order of flowering plants (brassicales). *G3: Genes — Genomes — Genetics*.

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, **12**(1), 59–60.

Cai, X., Wu, J., Liang, J., Lin, R., Zhang, K., Cheng, F., and Wang, X. (2020). Improved *Brassica oleracea* JZS assembly reveals significant changing of LTR-RT dynamics in different morphotypes. *Theoretical and Applied Genetics*, **133**(11), 3187–3199.

Canvin, D. T., Berry, J. A., Badger, M. R., Fock, H., and Osmond, C. B. (1980). Oxygen exchange in leaves in the light. *Plant Physiology*, **66**(2), 302–307.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., *et al.* (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, **345**(6199), 950–953.

Chang, S., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, **11**(2), 113–116.

Choi, S. R., Teakle, G. R., Plaha, P., Kim, J. H., Allender, C. J., Beynon, E., Piao, Z. Y., Soengas, P., Han, T. H., King, G. J., *et al.* (2007). The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *Theoretical and Applied Genetics*, **115**(6), 777–792.

Coombe, L., Zhang, J., Vandervalk, B. P., Chu, J., Jackman, S. D., Birol, I., and Warren, R. L. (2018). ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers. *BMC Bioinformatics*, **19**(1), 1–10.

Crafts-Brandner, S. J. and Salvucci, M. E. (2002). Sensitivity of photosynthesis in a C_4 plant, maize, to heat stress. *Plant Physiology*, **129**(4), 1773–1780.

Crocco, C. D., Ocampo, G. G., Ploschuk, E. L., Mantese, A., and Botto, J. F. (2018). Heterologous expression of *AtBBX21* enhances the rate of photosynthesis and alleviates photoinhibition in *Solanum tuberosum*. *Plant Physiology*, **177**(1), 369–380.

Darmency, H. and Fleury, A. (2000). Mating system in *Hirschfeldia incana* and hybridization to oilseed rape. *Weed Research (Oxford)*, **40**(2), 231–238.

Dassanayake, M., Oh, D.-H., Haas, J. S., Hernandez, A., Hong, H., Ali, S., Yun, D.-J., Bressan, R. A., Zhu, J.-K., Bohnert, H. J., *et al.* (2011). The genome of the extremophile crucifer *Thellungiella parvula*. *Nature Genetics*, **43**(9), 913–918.

De Souza, A. P., Burgess, S. J., Doran, L., Hansen, J., Manukyan, L., Maryn, N., Gotarkar, D., Leonelli, L., Niyogi, K. K., and Long, S. P. (2022). Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science*, **377**(6608), 851–854.

Driever, S., Lawson, T., Andralojc, P., Raines, C., and Parry, M. (2014). Natural variation in photosynthetic capacity, growth, and yield in 64 field-grown wheat genotypes. *Journal of Experimental Botany*, **65**(17), 4959–4973.

Du, J., Tian, Z., Sui, Y., Zhao, M., Song, Q., Cannon, S. B., Cregan, P., and Ma, J. (2012). Pericentromeric effects shape the patterns of divergence, retention, and expression of duplicated genes in the paleopolyploid soybean. *The Plant Cell*, **24**(1), 21–32.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5), 1792–1797.

Edger, P. P. and Pires, J. C. (2009). Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Research*, **17**(5), 699.

Edwards, E. J., Osborne, C. P., Strömberg, C. A., Smith, S. A., and C_4 Grasses Consortium (2010). The origins of C_4 grasslands: integrating evolutionary and ecosystem science. *Science*, **328**(5978), 587–591.

Ehleringer, J. (1985). Annuals and perennials of warm deserts. In B. H. Chabot and H. A. Mooney, editors, *Physiological Ecology of North American Plant Communities*, pages 162–180. Springer.

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics*, **9**(1), 18.

Emms, D. M. and Kelly, S. (2019). Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, **20**(1), 1–14.

Fahr, M., Laplaze, L., El Mzibri, M., Doumas, P., Bendaou, N., Hocher, V., Bogusz, D., and Smouni, A. (2015). Assessment of lead tolerance and accumulation in metallicolous and non-metallicolous populations of *Hirschfeldia incana*. *Environmental and Experimental Botany*, **109**, 186–192.

Favory, J.-J., Kobayshi, M., Tanaka, K., Peltier, G., Kreis, M., Valay, J.-G., and Lerbs-Mache, S. (2005). Specific function of a plastid sigma factor for *ndh* f gene transcription. *Nucleic Acids Research*, **33**(18), 5991–5999.

Flood, P. J., Harbinson, J., and Aarts, M. G. (2011). Natural genetic variation in plant photosynthesis. *Trends in Plant Science*, **16**(6), 327–335.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., and Smit, A. F. (2020). Repeatmodeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, **117**(17), 9451–9457.

Franzke, A., German, D., Al-Shehbaz, I. A., and Mummenhoff, K. (2009). Arabidopsis family ties: molecular phylogeny and age estimates in brassicaceae. *Taxon*, **58**(2), 425–437.

Furbank, R. T., Jimenez-Berni, J. A., George-Jaeggli, B., Potgieter, A. B., and Deery, D. M. (2019). Field crop phenomics: enabling breeding for radiation use efficiency and biomass in cereal crops. *New Phytologist*, **223**(4), 1714–1727.

Garrison, E. and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*.

Genty, B. and Harbinson, J. (1996). Regulation of light utilization for photosynthetic electron transport. In N. R. Baker, editor, *Photosynthesis and the Environment*, pages 67–99. Springer.

Gibson, A. C. (1998). Photosynthetic organs of desert plants. *Bioscience*, **48**(11), 911–920.

Gilbert, M. E., Zwieniecki, M. A., and Holbrook, N. M. (2011). Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. *Journal of Experimental Botany*, **62**(8), 2875–2887.

Gitelson, A. A., Peng, Y., Arkebauer, T. J., and Suyker, A. E. (2015). Productivity, absorbed photosynthetically active radiation, and light use efficiency in crops: Implications for remote sensing of crop primary production. *Journal of Plant Physiology*, **177**, 100–109.

Gu, J., Yin, X., Struik, P. C., Stomph, T. J., and Wang, H. (2012). Using chromosome introgression lines to map quantitative trait loci for photosynthesis parameters in rice (*Oryza sativa* L.) leaves under drought and well-watered field conditions. *Journal of Experimental Botany*, **63**(1), 455–469.

Gu, J., Yin, X., Stomph, T.-J., and Struik, P. C. (2014). Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. *Plant, Cell & Environment*, **37**(1), 22–34.

Gu, J., Zhou, Z., Li, Z., Chen, Y., Wang, Z., Zhang, H., and Yang, J. (2017). Photosynthetic properties and potentials for improvement of photosynthesis in pale green leaf rice under high light conditions. *Frontiers in Plant Science*, **8**, 1082.

Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., and Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, **36**(9), 2896–2898.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, **29**(8), 1072–1075.

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., and Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, **9**(1), R7.

Harbinson, J., Kaiser, E., and Morales, A. S. (2022). Integrating the stages of photosynthesis. In A. Ruban, C. H. Foyer, and E. H. Murchie, editors, *Photosynthesis in Action*, pages 195–242. Elsevier.

He, Z., Ji, R., Havlickova, L., Wang, L., Li, Y., Lee, H. T., Song, J., Koh, C., Yang, J., Zhang, M., Parkin, I. A., Wang, X., Edwards, D., King, G. J., Zou, J., Liu, K., Snowdon, R. J., Banga, S. S., Machackova, I., and Bancroft, I. (2021). Genome structural evolution in brassica crops. *Nature Plants*, **7**(6), 757–765.

Heddad, M., Norén, H., Reiser, V., Dunaeva, M., Andersson, B., and Adamska, I. (2006). Differential expression and localization of early light-induced proteins in arabidopsis. *Plant Physiology*, **142**(1), 75–87.

Huang, C.-H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., Zhang, Q., Koch, M. A., Al-Shehbaz, I., Edger, P. P., *et al.* (2016). Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular Biology and Evolution*, **33**(2), 394–412.

Hutin, C., Nussaume, L., Moise, N., Moya, I., Kloppstech, K., and Havaux, M. (2003). Early light-induced proteins protect *Arabidopsis* from photooxidative stress. *Proceedings of the National Academy of Sciences*, **100**(8), 4921–4926.

Johnston, J. S., Pepper, A. E., Hall, A. E., Chen, Z. J., Hodnett, G., Drabek, J., Lopez, R., and Price, H. J. (2005). Evolution of genome size in Brassicaceae. *Annals of Botany*, **95**(1), 229–235.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., *et al.* (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**(9), 1236–1240.

Joseph, J. T., Poolakkalody, N. J., and Shah, J. M. (2018). Plant reference genes for development and stress response studies. *Journal of Biosciences*, **43**(1), 173–187.

Katche, E., Quezada-Martinez, D., Katche, E. I., Vasquez-Teuber, P., and Mason, A. S. (2019). Interspecific hybridization for *Brassica* crop improvement. *Crop Breeding, Genetics and Genomics*, **1**(1).

Kent, T. V., Uzunović, J., and Wright, S. I. (2017). Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**(1736), 20160458.

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., and Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Research*, **21**(3), 487–493.

Kim, H., Choi, S. R., Bae, J., Hong, C. P., Lee, S. Y., Hossain, M. J., Van Nguyen, D., Jin, M., Park, B.-S., Bang, J.-W., *et al.* (2009). Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. *BMC Genomics*, **10**(1), 432.

Koornneef, M. and Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal*, **61**(6), 909–921.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, **27**(5), 722–736.

Lawson, T., Kramer, D. M., and Raines, C. A. (2012). Improving yield by exploiting mechanisms underlying natural variation of photosynthesis. *Current Opinion in Biotechnology*, **23**(2), 215–220.

Leakey, A. D., Uribelarrea, M., Ainsworth, E. A., Naidu, S. L., Rogers, A., Ort, D. R., and Long, S. P. (2006). Photosynthesis, productivity, and yield of maize are not affected by open-air elevation of CO_2 concentration in the absence of drought. *Plant Physiology*, **140**(2), 779–790.

Lee, P. L., Patel, R., Conlan, R., Wainwright, S., and Hipkin, C. (2004). Comparison of genetic diversities in native and alien populations of hoary mustard (*Hirschfeldia incana* [L.] Lagreze-Fossat). *International Journal of Plant Sciences*, **165**(5), 833–843.

Lee, S.-K., Jeon, J.-S., Boernke, F., Voll, L., Cho, J.-I., Goh, C.-H., Jeong, S.-W., Park, Y.-I., Kim, S. J., Choi, S.-B., Miyoa, A., Hirochika, H., An, G., Cho, M.-H., Bhoo, S. H., Sonnewalk, U., and Hahn, T.-R. (2008). Loss of cytosolic fructose-1, 6-bisphosphatase limits photosynthetic sucrose synthesis and causes severe growth retardations in rice (*Oryza sativa*). *Plant, Cell & Environment*, **31**(12), 1851–1863.

Leister, D. (2019). Thawing out frozen metabolic accidents. *BMC Biology*, **17**(1), 1–11.

Letunic, I. and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, **49**(W1), W293–W296.

Li, Y., Liu, B., Zhang, J., Kong, F., Zhang, L., Meng, H., Li, W., Rochaix, J.-D., Li, D., and Peng, L. (2019). Ohp1, ohp2, and hcf244 form a transient functional complex with the photosystem ii reaction center. *Plant Physiology*, **179**(1), 195–208.

Li, Z., Defoort, J., Tasdighian, S., Maere, S., Van de Peer, Y., and De Smet, R. (2016). Gene duplicability of core genes is highly consistent across all angiosperms. *The Plant Cell*, **28**(2), 326–344.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., Zhao, M., Ma, J., Yu, J., Huang, S., *et al.* (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, **5**(1), 1–11.

Liu, Y., Wei, W., Ma, K., Li, J., Liang, Y., and Darmency, H. (2013). Consequences of gene flow between oilseed rape (*Brassica napus*) and its relatives. *Plant Science*, **211**, 42–51.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\delta\delta$ct method. *METHODS*, **25**(4), 402–408.

Long, S. P., Zhu, X.-G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment*, **29**(3), 315–330.

Løvdal, T. and Lillo, C. (2009). Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. *Analytical Biochemistry*, **387**(2), 238–242.

Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe *Brassiceae*. *Genome Research*, **15**(4), 516–525.

Maiwald, D., Dietzmann, A., Jahns, P., Pesaresi, P., Joliot, P., Joliot, A., Levin, J. Z., Salamini, F., and Leister, D. (2003). Knock-out of the genes coding for the rieske protein and the atp-synthase δ-subunit of arabidopsis. effects on photosynthesis, thylakoid protein composition, and nuclear chloroplast gene expression. *Plant Physiology*, **133**(1), 191–202.

Mandáková, T. and Lysak, M. A. (2018). Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology*, **42**, 55–65.

Mandáková, T., Li, Z., Barker, M. S., and Lysak, M. A. (2017). Diverse genome organization following 13 independent mesopolyploid events in Brassicaceae contrasts with convergent patterns of gene retention. *The Plant Journal*, **91**(1), 3–21.

Marçais, G. and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, **27**(6), 764–770.

Marshall, B. and Biscoe, P. (1980). A model for C_3 leaves describing the dependence of net photosynthesis on irradiance. *Journal of Experimental Botany*, **31**(1), 29–39.

Martinis, J., Glauser, G., Valimareanu, S., and Kessler, F. (2013). A chloroplast abc1-like kinase regulates vitamin e metabolism in arabidopsis. *Plant Physiology*, **162**(2), 652–662.

Mira, S., Veiga-Barbosa, L., and Pérez-García, F. (2019). Seed dormancy and longevity variability of *Hirschfeldia incana* L. during storage. *Seed Science Research*, **29**(2), 97–103.

Monneveux, P., Pastenes, C., and Reynolds, M. P. (2003). Limitations to photosynthesis under light and heat stress in three high-yielding wheat genotypes. *Journal of Plant Physiology*, **160**(6), 657–666.

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K.-I., Endo, T., Tasaka, M., and Shikanai, T. (2004). Cyclic electron flow around photosystem i is essential for photosynthesis. *Nature*, **429**(6991), 579–582.

Murchie, E. H., Chen, Y.-Z., Hubbart, S., Peng, S., and Horton, P. (1999). Interactions between senescence and leaf orientation determine in situ patterns of photosynthesis and photoinhibition in field-grown rice. *Plant Physiology*, **119**(2), 553–564.

Norén, H., Svensson, P., Stegmark, R., Funk, C., Adamska, I., and Andersson, B. (2003). Expression of the early light-induced protein but not the psbs protein is influenced by low temperature and depends on the developmental stage of the plant in field-grown pea cultivars. *Plant, Cell & Environment*, **26**(2), 245–253.

Oh, D.-H., Dassanayake, M., Bohnert, H. J., and Cheeseman, J. M. (2013). Life at the extreme: lessons from the genome. *Genome Biology*, **13**(3), 1–9.

Ortiz, D., Hu, J., and Salas Fernandez, M. G. (2017). Genetic architecture of photosynthesis in *Sorghum bicolor* under non-stress and cold stress conditions. *Journal of Experimental Botany*, **68**(16), 4545–4557.

Ou, S. and Jiang, N. (2018). LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiology*, **176**(2), 1410–1422.

Paritosh, K., Pradhan, A. K., and Pental, D. (2020). A highly contiguous genome assembly of *Brassica nigra* (BB) and revised nomenclature for the pseudochromosomes. *BMC Genomics*, **21**(1), 1–12.

Peng, L., Fukao, Y., Fujiwara, M., Takami, T., and Shikanai, T. (2009). Efficient operation of nad(p)h dehydrogenase requires supercomplex formation with photosystem i via minor lhci in *Arabidopsis*. *The Plant Cell*, **21**(11), 3623–3640.

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., Sankoff, D., Robinson, S. J., Kagale, S., Navabi, Z.-K., Tang, L., Horner, K. N., He, Z., Bancroft, I., Chalhoub, B., Sharpe, A. G., and Parkin, I. A. (2020). A high-contiguity *Brassica nigra* genome localizes active centromeres and defines the ancestral *Brassica* genome. *Nature Plants*, **6**(8), 929–941.

Pleban, J. R., Mackay, D. S., Aston, T. L., Ewers, B. E., and Weinig, C. (2018). Phenotypic trait identification using a multimodel Bayesian method: a case study using photosynthesis in *Brassica rapa* genotypes. *Frontiers in Plant Science*, **9**, 448.

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., Zhang, S., and Paterson, A. H. (2019). Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biology*, **20**(1), 1–23.

Ritz, K. R., Noor, M. A., and Singh, N. D. (2017). Variation in recombination rate: adaptive or not? *Trends in Genetics*, **33**(5), 364–374.

Rizzon, C., Ponger, L., and Gaut, B. S. (2006). Striking similarities in the genomic distribution of tandemly arrayed genes in arabidopsis and rice. *PLoS computational biology*, **2**(9), e115.

Sánchez-Yélamo, M. (2009). Relationships in the *Diplotaxis–Erucastrum–Brassica* complex (Brassicaceae) evaluated from isoenzymatic profiles of the accessions as a whole. Applications for characterisation of phytogenetic resources preserved *ex situ*. *Genetic Resources and Crop Evolution*, **56**(7), 1023.

Schranz, M. E., Lysak, M. A., and Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science*, **11**(11), 535–542.

Sheikhizadeh Anari, S., de Ridder, D., Schranz, M. E., and Smit, S. (2018). Efficient inference of homologs in large eukaryotic pan-proteomes. *BMC Bioinformatics*, **19**(1), 340.

Shi, T., Bibby, T. S., Jiang, L., Irwin, A. J., and Falkowski, P. G. (2005). Protein interactions limit the rate of evolution of photosynthetic genes in cyanobacteria. *Molecular Biology and Evolution*, **22**(11), 2179–2189.

Siemens, J. (2011). Hirschfeldia. In C. Kole, editor, *Wild Crop Relatives: Genomic and Breeding Resources*, pages 171–176. Springer.

Smit, A. F., Hubley, R., and Green, P. (2015). Repeatmasker open-4.0.

Somerville, C. and Koornneef, M. (2002). A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics*, **3**(11), 883–889.

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G. J., and Østergaard, L. (2010). A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biology*, **10**(1), 1–10.

Stinziano, J. R., Roback, C., Gamble, D., Murphy, B., Hudson, P., and Muir, C. D. (2020). photosynthesis: tools for plant ecophysiology & modeling. R package version 2.0.1.

Strigens, A., Freitag, N. M., Gilbert, X., Grieder, C., Riedelsheimer, C., Schrag, T. A., Messmer, R., and Melchinger, A. E. (2013). Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments. *Plant, Cell & Environment*, **36**(10), 1871–1887.

Sun, H., Ding, J., Piednoël, M., and Schneeberger, K. (2018). findGSE: estimating genome size variation within human and *Arabidopsis* using *k*-mer frequencies. *Bioinformatics*, **34**(4), 550–557.

Suryawanshi, V., Talke, I. N., Weber, M., Eils, R., Brors, B., Clemens, S., and Krämer, U. (2016). Between-species differences in gene copy number are enriched among functions critical for adaptive evolution in *Arabidopsis halleri*. *BMC Genomics*, **17**(13), 43–64.

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**(suppl_2), W609–W612.

Tang, H., Wang, X., Bowers, J. E., Ming, R., Alam, M., and Paterson, A. H. (2008). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Research*, **18**(12), 1944–1954.

Taylor, S. H., Orr, D. J., Carmo-Silva, E., and Long, S. P. (2020). During photosynthetic induction, biochemical and stomatal limitations differ between *Brassica* crops. *Plant, Cell & Environment*.

The Brassica rapa Genome Sequencing Project Consortium (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, **43**(10), 1035–1039.

Tian, Z., Rizzon, C., Du, J., Zhu, L., Bennetzen, J. L., Jackson, S. A., Gaut, B. S., and Ma, J. (2009). Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons? *Genome Research*, **19**(12), 2221–2230.

Tsunoyama, Y., Ishizaki, Y., Morikawa, K., Kobori, M., Nakahira, Y., Takeba, G., Toyoshima, Y., and Shiina, T. (2004). Blue light-induced transcription of plastid-encoded *psbD* gene is mediated by a nuclear-encoded transcription initiation factor, atsig5. *Proceedings of the National Academy of Sciences*, **101**(9), 3304–3309.

Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., and Gregory, M. (2003). A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biology*, **9**(3), 383–395.

van Bezouw, R. F., Keurentjes, J. J., Harbinson, J., and Aarts, M. G. (2019). Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. *The Plant Journal*, **97**(1), 112–133.

van Rooijen, R., Aarts, M. G., and Harbinson, J. (2015). Natural genetic variation for acclimation of photosynthetic light use efficiency to growth irradiance in Arabidopsis. *Plant Physiology*, **167**(4), 1412–1429.

van Rooijen, R., Kruijer, W., Boesten, R., Van Eeuwijk, F. A., Harbinson, J., and Aarts, M. G. (2017). Natural variation of yellow seedling1 affects photosynthetic acclimation of arabidopsis thaliana. *Nature Communications*, **8**(1), 1–9.

von Caemmerer, S. and Evans, J. R. (2010). Enhancing C_3 photosynthesis. *Plant Physiology*, **154**(2), 589–592.

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., and Schatz, M. C. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, **33**(14), 2202–2204.

Wang, X., Torres, M. J., Pierce, G., Lemke, C., Nelson, L. K., Yuksel, B., Bowers, J. E., Marler, B., Xiao, Y., Lin, L., *et al.* (2011). A physical map of *Brassica oleracea* shows complexity of chromosomal changes following recursive paleopolyploidizations. *BMC Genomics*, **12**(1), 1–15.

Werk, K., Ehleringer, J., Forseth, I., and Cook, C. (1983). Photosynthetic characteristics of Sonoran Desert winter annuals. *Oecologia*, **59**(1), 101–105.

Xu, C., Yinan, C., Zhang, L., Jian, W., Liang, J., Cheng, L., Xiaowu, W., and Cheng, F. (2018). Hotspots of independent and multiple rounds of LTR-retrotransposon bursts in Brassica species. *Horticultural Plant Journal*, **4**(4), 165–174.

Xu, Z. and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Research*, **35**(suppl_2), W265–W268.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8), 1586–1591.

Yang, Z. and Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution*, **17**(1), 32–43.

Youssef, A., Laizet, Y., Block, M. A., Maréchal, E., Alcaraz, J.-P., Larson, T. R., Pontier, D., Gaffé, J., and Kuntz, M. (2010). Plant lipid-associated fibrillin proteins condition jasmonate production under photosynthetic stress. *The Plant Journal*, **61**(3), 436–445.

Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., Liang, J., Cai, C., Liu, Z., Liu, B., *et al.* (2018). Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Horticulture Research*, **5**(1), 1–11.

Zhao, M., Du, J., Lin, F., Tong, C., Yu, J., Huang, S., Wang, X., Liu, S., and Ma, J. (2013). Shifts in the evolutionary rate and intensity of purifying selection between two brassica genomes revealed by analyses of orthologous transposons and relics of a whole genome triplication. *The Plant Journal*, **76**(2), 211–222.

Zhu, X.-G., Long, S. P., and Ort, D. R. (2010). Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biology*, **61**, 235–261.

Żmieńko, A., Samelak, A., Kozłowski, P., and Figlerowicz, M. (2014). Copy number polymorphism in plant genomes. *Theoretical and Applied genetics*, **127**(1), 1–18.

CHAPTER 4

Comparative transcriptomics of *Hirschfeldia incana* and relatives highlights differences in photosynthetic pathways

BioRxiv, 2023, 10.18.562717.

Francesco Garassino[1], Sofia Bengoa Luoni[1], Tommaso Cumerlato[1], Francisca Reyes Marquez[1], Jeremy Harbinson[2], Mark G. M. Aarts[1], Harm Nijveen[3], and Sandra Smit[3]

(1) Laboratory of Genetics, Wageningen University and Research, Wageningen (NL)

(2) Laboratory of Biophysics, Wageningen University and Research, Wageningen (NL)

(3) Bioinformatics Group, Wageningen University and Research, Wageningen (NL)

Supplementary material available at:
https://doi.org/10.1101/2023.10.18.562717

Abstract

Photosynthesis is the only yield-related trait that has not yet been substantially improved by plant breeding. The limited results of previous attempts to increase yield via improvement of photosynthetic pathways suggest that more knowledge is still needed to achieve this goal. To learn more about the genetic and physiological basis of photosynthetic light-use efficiency (LUE) at high irradiance, we study *Hirschfeldia incana*. Here, we compare the transcriptomic response to high light of *H. incana* with that of three other members of the Brassicaceae, *Arabidopsis thaliana*, *Brassica rapa*, and *Brassica nigra*, which have a lower photosynthetic LUE.

First, we built a high-light, high-uniformity growing environment in a climate-controlled room. Plants grown in this system developed normally and showed no signs of stress during the whole growth period. Then we compared gene expression in low and high-light conditions across the four species, utilizing a panproteome to group homologous proteins efficiently. As expected, all species actively regulate genes related to the photosynthetic process. An in-depth analysis on the expression of genes involved in three key photosynthetic pathways revealed a general trend of lower gene expression in high light conditions. However, *H. incana* distinguishes itself from the other species through higher expression of certain genes in these pathways, either through constitutive higher expression, as for *LHCB8*, ordinary differential expression, as for *PSBE*, or cumulative higher expression obtained by simultaneous expression of multiple gene copies, as seen for *LHCA6*.

These differentially expressed genes in photosynthetic pathways are interesting leads to further investigate the exact relationship between gene expression, protein abundance and turnover, and ultimately the LUE phenotype. In addition, we can also exclude thousands of genes from "explaining" the phenotype, because they do not show differential expression between both light conditions. Finally, we deliver a transcriptomic resource of plant species fully grown under, rather than briefly exposed to, a very high irradiance, supporting efforts to develop highly efficient photosynthesis in crop plants.

4.1 Background

Considering the projected global population growth, the increasing effects of global warming, and the need for a more sustainable means of production, it is evident that the agricultural sector is under substantial pressure to increase crop yields while reducing land use and inputs such as fertilisers and pesticides. Over the past decade photosynthesis has taken a central role in plant research aimed at increasing crop yields because it plays a major role in the crop energy conversion efficiency, the only yield-related trait of food and feed crops that has not yet been maximised or even substantially improved by plant breeding (Zhu *et al.*, 2010).

While increasing crop productivity via improved photosynthetic efficiency was proposed over forty years ago (Austin, 1989), limited results have been achieved so far due to the physiological and genetic complexity of the photosynthetic process. Studies based on modelling of the photosynthetic process, bottleneck identification, and genetic modification aimed at overcoming identified bottlenecks have proven successful in some field crops, with yield increases ranging between 15 and 28% (Kromdijk *et al.*, 2016; De Souza *et al.*, 2022; López-Calcagno *et al.*, 2020, 2019; Yoon *et al.*, 2020). However, the inconsistency of results over multiple seasons (De Souza *et al.*, 2022) and across species (Garcia-Molina and Leister, 2020) or growing conditions (Ruiz-Vera *et al.*, 2022) indicates that more knowledge is needed on photosynthetic processes and how they are influenced by the environment across a range of timescales if we are to systematically increase the photosynthetic efficiency in crops.

Plant photosynthesis is defined as the process in which energy from light radiation is converted into chemical energy via a complex series of reactions resulting in the production of carbohydrates and oxygen (Salisbury and Ross, 1992). A first set of photosynthetic reactions, catalyzed by photosystem complexes and an associated electron transport chain, is responsible for converting light radiation energy into chemical energy. This chemical energy is then stored in metabolically useful reducing agents (e.g. NADPH) and the energy-rich phosphate donor ATP. The processes linking light absorption to the formation of ATP and NADPH and other reducing agents collectively form the light reactions of photosynthesis (Schulze and Caldwell, 1995). The energy-rich reducing agents and ATP then drive a second set of photosynthetic reactions, the so-called dark reactions. These, generally referred as to the Calvin-Benson cycle, result in the conversion of the inorganic carbon substrate CO_2 into organic carbohydrate molecules (Schulze and Caldwell, 1995). Bottlenecks or constraints - sites whose modification could result in improved photosynthesis - have been identified in both the light and dark reactions (Zhu *et al.*, 2010). Furthermore, bottlenecks affecting photosynthesis have been identified in processes that would not be defined as strictly photosynthetic, such as the diffusion of CO_2 into and through leaves to the site of CO_2fixation in chloroplasts, and the transport of carbohydrates from photosynthetically active cells to carbon sinks elsewhere in the plant (Singh *et al.*, 2014).

Our current knowledge of the key mechanisms and components of photosynthesis is the result of decades of studies in plants and other photosynthetic organisms (Johnson, 2016). This amounts to a vast body of knowledge, but on its own

it is insufficient to improve crops' photosynthesis and their yield. Studies conducted so far have highlighted how the link between crop photosynthesis and productivity is much more complex than originally thought, as a result of interactions of this process with plant development and environmental factors (Araus *et al.*, 2021). One characteristic of photosynthesis that can have a major impact on crop productivity is the decreasing light-use efficiency that occurs with increasing irradiance, giving rise to the light-saturation of photosynthesis and limitation of assimilation rate. This limitation has a substantial impact on productivity at irradiance levels normally recorded during summer in temperate areas of our planet. We define photosynthetic light-use efficiency (LUE) as the ratio between photosynthetically assimilated CO_2 and incident light radiation, or irradiance. The decrease in LUE due to increasing irradiance is well-known and its causes are linked to both limitations in the photosynthetic process and other associated processes (Taylor *et al.*, 2022).

Evidence has been reported for large natural variation in photosynthesis rates, and therefore photosynthetic LUE, among crop and other plant species (Yin *et al.*, 2022; Theeuwen *et al.*, 2022). This suggests that a degree of plasticity exists for photosynthesis that could be leveraged to increase the photosynthesis of crop species. However, it is nowadays clear that increases might only be achieved if knowledge is accumulated on the regulation of the photosynthetic process as well as specific strategies some plant species might have evolved that result in photosynthesis optimised to meet unusual goals (Taylor *et al.*, 2022). One powerful way to map the genetic basis of complex biological processes is via the analysis of the associated transcriptional activity. Over the past years, several studies on transcriptional activity in a number of species have increased our knowledge on the response of photosynthesis to irradiances of different intensities or changes in irradiance. It was shown that *Arabidopsis thaliana* acclimates to high light by increasing expression of heat shock response genes and lipid remodelling genes (van Rooijen *et al.*, 2018), that rice (*Oryza sativa*) exposed to variations in irradiance associated to field conditions activates a large number of biotic and abiotic stress genes (Hashida *et al.*, 2022), and that barley (*Hordeum vulgare*) expresses genes involved in phenolic compounds accumulation at a higher level with increasing irradiance (Pech *et al.*, 2022). However, none of these studies applied a long-term, very high irradiance treatment, similar to what is experienced by plants growing in natural temperate environments during summer months, at high altitude conditions, or in the equatorial region. Neither did they include species with a particularly high photosynthetic LUE. We consider these two factors essential for unraveling the physiological and genetic basis of photosynthetic light-use efficiency, and for ultimately building more light-use efficient photosynthesis in our crops (Taylor *et al.*, 2022).

Here, we present the analysis of gene expression in *Hirschfeldia incana* (L.) Lagr.-Foss., the species we previously proposed as preferred model for studies on high photosynthetic LUE (Garassino *et al.*, 2022; Taylor *et al.*, 2022). The gene expression under contrasting high- and low-light irradiance conditions is compared to that of Brassicaceae family relatives *A. thaliana*, *Brassica rapa*, and *Brassica nigra*. While *A. thaliana* does not share the whole genome triplication that the other three species underwent and is therefore more distantly related, *B. rapa* and *B. nigra* represent the

different evolutionary history of two major lineages emerging after this event (Arias and Pires, 2012). Using transcriptomics we aim to elucidate which genes and thus pathways are involved in the maintenance of a high photosynthetic LUE at high irradiance in *H. incana*. First, we describe the experiment we conducted under high irradiance, and present the results of differential gene expression (DGE) analysis performed on each of the four species. Then, we report on the use of a panproteome to compare gene expression changes across the four species, and on the exploration of common and divergent trends in the gene expression response to high-light by means of untargeted enrichment analyses. We then present the results of targeted analysis of expression patterns across the four species for genes involved in key photosynthesis-related pathways. Lastly, we discuss the findings in the light of their implications for *H. incana's* higher photosynthetic light-use efficiency at high irradiance. Our work thus describes the transcriptional differences associated with plant growth under highly contrasting irradiance conditions, and serves as a resource for the elucidation of the genetic determinants of the striking photosynthetic capacity of *Hirschfeldia incana*.

4.2 Results

4.2.1 Plant growth under a reliable high-light environment

In this study, we aimed to identify genes and pathways responsible for the higher photosynthetic light-use efficiency of *H. incana* under high irradiance conditions. To create strongly contrasting growth conditions, we set our low light (LL) irradiance to 200 µmol m^{-2} s^{-1} and our high light (HL) irradiance to 1800 µmol m^{-2} s^{-1} for 12 h per day. We calculated U_2 irradiance uniformity values (Hu *et al.*, 2015; Sun *et al.*, 2014) over all growing positions designated for both treatments, and we selected positions on each growing table that resulted in the best irradiance uniformity. For the LL table, this resulted in an average irradiance of 227.5 µmol m^{-2} s^{-1} associated with an U_2 of 0.93, while for the whole HL table we measured an average irradiance of 1843.6 µmol m^{-2} s^{-1}, also associated with an U_2 of 0.93.

To compare the light treatments to conditions that plants would experience in natural environments, we calculated a Daily Light Integral (DLI) (Faust and Logan, 2018), a measure of the total irradiance delivered over the course of a day per unit of area, for each treatment. This resulted in DLIs of 9.82 mol m^{-2} d^{-1} and 79.64 mol m^{-2} d^{-1} for the LL and HL treatments, respectively.

Besides *H. incana*, this study featured three other Brassicaceae species: *A. thaliana*, *B. rapa*, and *B. nigra*. These are the same we used for previous work in which they showed to have lower photosynthetic LUE than *H. incana* (Garassino *et al.*, 2022). Plants of the four species established and grew well under both light treatments, albeit with differences in growth and architecture (Figures 4.1, S1). No stress symptoms were visible on the plants throughout the growing period. 20% of the *B. nigra* plants from the LL treatment appeared to grow more slowly and had paler leaf color than the other *B. nigra* plants (Figure S1).

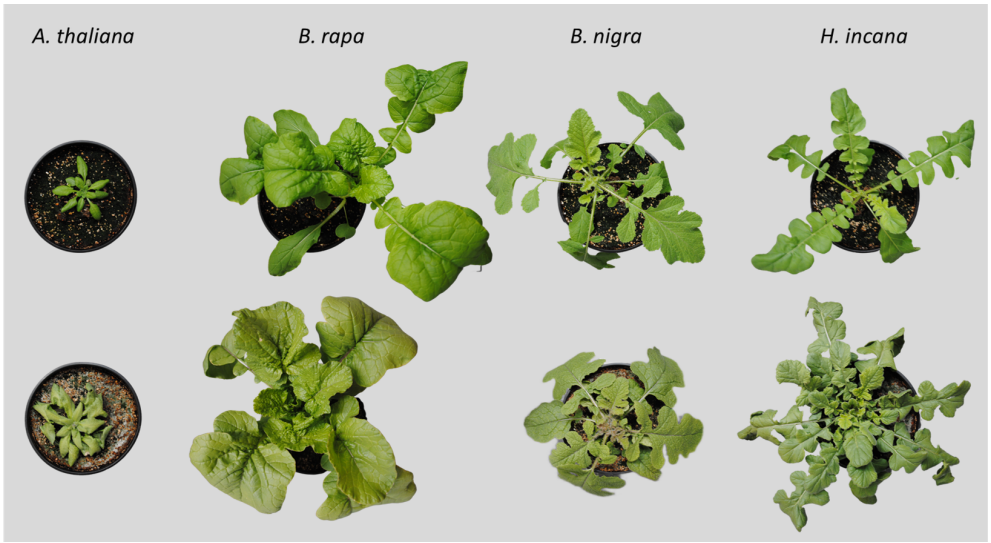

Figure 4.1: Pictures of representative plants at the end of the experiment for each of the four species grown under the two irradiance conditions. Left to right: *A. thaliana*, *B. rapa*, *B. nigra*, *H. incana*. Top row: low light (LL) irradiance. Bottom row: high light (HL) irradiance.

4.2.2 Per-species differential gene expression analysis

To study the gene expression under contrasting light conditions in the four species we sequenced forty mRNA libraries (4 species $\times$ 5 replicates $\times$ 2 conditions) generated from RNA extracted from whole plant canopies, with an average of 22.2 ± 2.4 million reads per library. The MultiQC inspection of all sequencing reads did not show any quality issues in our dataset. Percentages of reads mapped to reference genomes were overall high, ranging between 93.4 ± 1.5% and 96.3 ± 0.7% (Table S1).

We performed differential expression (DE) analysis on data from each species individually with DESeq2 and selected all differentially expressed genes (Tables 4.1, S2-S5). Per-species principal component analysis performed on count data transformed by regularized logarithm showed that the general patterns of gene expression are consistent across biological replicates belonging to the same species and originating from the same treatment (Figure S2). The percentages of genes significantly differentially expressed were similar for *A. thaliana*, *B. rapa*, and *H. incana*, while they were lower for *B. nigra* due to the high number of genes in the annotation. Since we are interested in differences and similarities between *A. thaliana* and the other species, and in particular *H. incana*, we compared the gene expression across the species.

Table 4.1: Numbers of differentially expressed and non-deferentially expressed genes for the various species. Percentages of the total number of genes are placed between brackets. Significant differences for $p<0.05$.

	Number of genes		
	Higher expression	Lower expression	Unchanged expression
A. thaliana	3,346 (12.1%)	3,027 (11.0%)	21,226 (76.9%)
B. rapa	7,292 (15.7%)	7,138 (15.4%)	32,050 (68.9%)
B. nigra	4,723 (7.9%)	4,052 (6.8%)	50,934 (85.3%)
H. incana	4,284 (13.2%)	4,334 (13.3%)	23,900 (73.5%)

4.2.3 Cross-species comparison using a panproteome

To enable the comparison of gene expression across species, we built a panproteome to group homologous genes (orthologs and paralogs) (Sheikhizadeh Anari *et al.*, 2018). A panproteome of eight Brassicaceae species (*Aethionema arabicum*, *A. thaliana*, *B. nigra*, *B. rapa*, *H. incana*, *Raphanus raphanistrum*, *Raphanus sativus*, *Sisymbrium irio*) yielded 106,511 homology groups (HGs, Figure S3, Table S6). We then selected HGs containing at least one gene from one of the four species for which RNA-Seq was performed, leaving 63,675 HGs for downstream analysis (Figure 4.2).

We also distinguished "differentially expressed" (DE) HGs, which contain at least one gene that was differentially expressed between both light conditions, and "non-DE" HGs (Table 4.2), which do not. Among the 19,012 DE HGs, the 10,770 which have an ortholog from each of the four species (i.e. the core proteome) form the main target of our research. Of particular interest are the differences and similarities between *A. thaliana* and the other species, all of which are members of tribe Brassiceae, and in particular *H. incana*. Similarly, we identified 3,688 HGs which contain genes that were not differentially expressed in any of the species and can therefore not explain the phenotypic differences (Table S7). Based on the *A. thaliana* genes contained in these groups, we found 267 Gene Ontology (GO) Biological Process (BP) terms enriched (Table S8). As expected, no terms related to photosynthesis or high-light adaptation were identified in this set of groups.

To compare transcriptional activity across species within a single light condition, we compared transcripts-per-million (TPM)-normalised transcript counts. To assess bias due to differences in sequencing libraries, which are not corrected for during TPM normalisation (Zhao *et al.*, 2020), we tested whether average transcript abundances were similar across species. We selected the non-DE homology groups containing a single expressed ortholog for each of the four species. We averaged TPM counts (regardless of treatment) for each species and calculated per-HG $\log_2$-ratios between average counts for the single orthologs of the various species. The distribution of these ratios showed that on average the *A. thaliana* transcript abundances are higher than those of *B. rapa* and *H. incana* (% of area under the curve (AUC) for

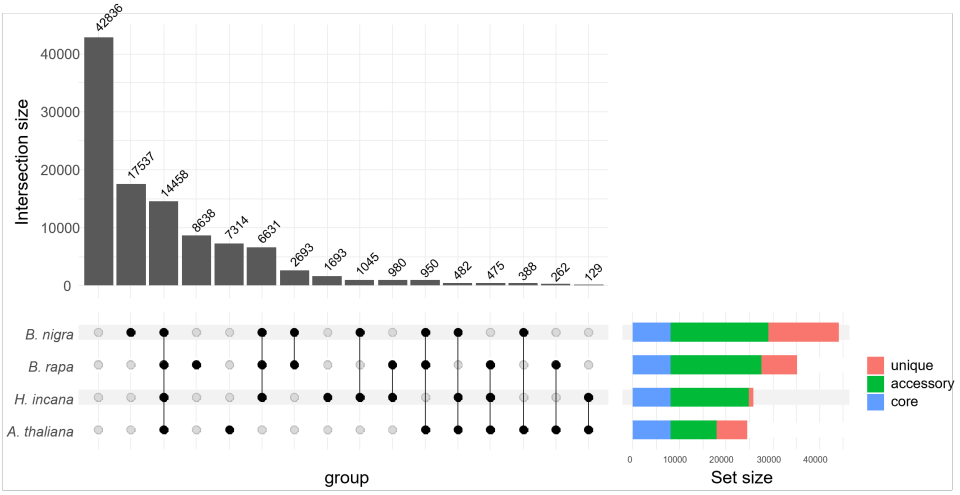

Figure 4.2: An UpSet plot of the panproteome homology groups (HGs), based on the four species for which RNA-Seq was done. Vertical bars represent the number of HGs, classified by presence/absence of genes from the various species as illustrated at the bottom of the figure. The first bar represents all the HGs of the panproteome constructed with proteomes from eight Brassicaceae species that do not contain any genes from the four species we sequenced. The horizontal colored bars show how many of the HGs containing at least one gene from a species fall within one of three categories: core HGs, i.e. those containing at least one gene from all four species; accessory HGs, i.e. those containing genes from more than one, but not all, species; and unique HGs, i.e. those containing only genes from a single species.

$\log_2$-ratios > 1: 59.9 and 62.8%, respectively), which are in turn higher than those of *B. nigra* (% of area under the curve for $\log_2$-ratios > 1: 67.4 % for *B. rapa* and 67.2% for *H. incana*) (Figure S4, Table S9). Given the fact that expression in *H. incana* is generally lower than in *A. thaliana* and similar to that in *B. rapa*, we conclude that detection of a significantly higher expression in *H. incana* is the effect of biological processes rather than an artifact.

4.2.4 Comparative analysis of core DE homology groups highlights photosynthetic pathways

Of the 10,770 core HGs (CHGs) containing at least one gene differentially expressed under HL, 10,352 showed non-ambiguous differential expression within each species and were selected for downstream analysis. We defined non-ambiguous DE as the situation in which the expression of all genes is exclusively increased or decreased. Clustering the CHGs with non-ambiguous responses allowed us to identify expression profiles for the four species (Figure 4.3, Table S10). Some CHGs show consist-

Table 4.2: Numbers of homology groups (HGs) in the constructed panproteome, classified based on their "differential expression" (DE) status. An HG is classified as "DE" if it contains at least one gene differentially expressed between both light conditions. If this condition is not met, the HG is classified as "non-DE".

	Homology groups		
	DE	Non-DE	Total
Core	10,770	3,688	14,458
Accessory	6,132	7,903	14,035
Unique	2,838	32,344	35,182
Total	19,012	44,663	63,675

ent higher or lower expression in all species (245 and 382, respectively). More often, an higher or lower expression is shared by some species (2,165 and 1,890 CHGs) or is unique to a species (2,163 and 1,946). Lastly, there are CHGs showing contrasting expression (higher expression in some species, lower expression in others) across species (1,561).

In order to get an overview of the role of the genes belonging to clusters of CHGs, we performed Gene Ontology (GO) (Ashburner *et al.*, 2000; Gene Ontology Consortium, 2021) and KEGG (Kanehisa and Goto, 2000; Kanehisa *et al.*, 2023) enrichment analyses separately for clusters of CHGs containing at least one gene model with significantly higher or lower expression in each species, in the three Brassiceae species (*B. rapa*, *B. nigra*, *H. incana*), and in *H. incana* alone. GO enrichment analysis for the sets of *A. thaliana* genes in the clusters of CHGs containing genes with higher expression under HL resulted in terms related to response to water deprivation and salt stress, heat, low cellular oxygen, and flavonoid biosynthesis (Table S11). A similar selection and analysis for genes with lower expression under HL resulted in terms involved in high-light response, chlorophyll metabolism, and growth regulation.

KEGG enrichment analysis highlighted seventeen over-represented pathways (Table S12). Twelve pathways were enriched in CHGs containing at least one gene with higher expression under HL in all species, Brassiceae species, or *H. incana* alone. The most notable of these pathways was "Carbon metabolism" (ath01200). The remaining five patwhays were enriched in CHGs containing at least one gene with lower expression under HL in all species, Brassiceae species, or *H. incana* alone. Notably, these five pathways comprised the two currently annotated in KEGG for photosynthesis: "Photosynthesis"(ath00195), and "Photosynthesis - antenna proteins" (ath00196). Since all three photosynthesis-related KEGG pathways were highlighted by our enrichment analysis, we decided to further explore the expression of the genes associated with these pathways in search of clues on the higher photosynthetic LUE of *H. incana*.

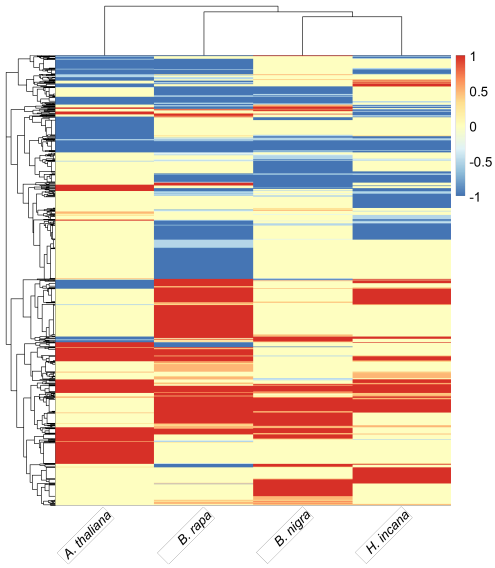

Figure 4.3: Heatmap of 10,352 groups showing only non-ambiguous responses per species. The color scale represents the ratio between the number of higher (positive numbers) or significantly lower expressed (negative numbers) genes and the total number of gene models present in each group per each species. Both rows and columns were clustered with hierarchical clustering based on Euclidean distances.

4.2.5 Targeted analysis of light-harvesting complex genes

We first analyzed the expression patterns of the *A. thaliana* genes annotated with the KEGG pathway "Photosynthesis - antenna proteins" (ath00196) and their orthologs in *B. rapa*, *B. nigra*, and *H. incana*. This allowed us to investigate transcriptional differences associated with light-harvesting complexes (LHCs), which are amongst the first complexes involved in the photosynthetic process. The KEGG pathway is made up of 22 *A. thaliana* genes assigned to 14 homology groups. These groups contain 34 genes for *B. rapa*, 33 genes for *B. nigra*, and 35 genes for *H. incana*. Inspection on these genes revealed $\log_2$ fold change ($\log_2$FC) values ranging between -3.15 and 1.01, with almost all genes showing significant lower expression under HL, except for *LHCB8*, *LHCB7*, and *LHCA5* (Figure S5). No differences across species were observed except for two genes coding for photosystem II (PSII) antenna proteins, *LHCB8* and *LHCB7*, and two coding for photosystem I (PSI) antenna proteins, *LHCA6* and *LHCA5* (Figure 4.4a). Considering the particular features explained below, we selected the *LHCB8* and *LHCA6* genes for further investigation.

LHCB8 was first investigated as member of a subset of rarely expressed light-harvesting complex (LHC) protein encoding genes (Klimmek *et al.*, 2006). The LHCB8 protein resembles the CP29.1 and CP29.2 proteins, encoded in *A. thali-*

ana by the *LHCB4.1* and *LHCB4.2* genes, and is therefore also known as CP29.3 (*LHCB4.3*). *AtLHCB8* shows a different expression pattern than *AtLHCB4.1* and *AtLHCB4.2*, suggesting a different role for the protein in the LHC. It seems to be present only in species of the eurosids, a subclade of the rosids (Klimmek *et al*., 2006). The LHCB8 protein is present as a monomer within the PSII supercomplex, forming the so-called "minor antenna" of photosystem II with a number of other LHCB proteins (LHCB4, LHCB5, LHCB6) (de Bianchi *et al*., 2011)). The expression of *A. thaliana LHCB8* is induced by high irradiance conditions (Flannery *et al*., 2021b)). *LHCB8* is a single-copy gene in the four species used for this study and under HL had moderately higher expression in *A. thaliana* and *B. nigra*, while it had moderarely lower expression in *B. rapa* and showed no significant changes in *H. incana*. Based on transcripts-per-million (TPM)-normalised read counts, *LHCB8* transcripts represent roughly 26.5% of the transcript pool for the minor antenna in *H. incana* plants grown under HL, while they represent only 5.6%, in *A. thaliana* (Figure 4.4b). This high representation is also found in *B. rapa* and *B. nigra*, with *LHCB8* making up 22.5% and 31.1% of minor antenna transcripts (Figure S6).

LHCA6 is a poorly-expressed gene coding for a protein associated with PSI as an antenna monomer. *LHCA6* is present as a single-copy gene in *A. thaliana*, *B. rapa*, and *B. nigra*, while it has three tandem copies in *H. incana* (Garassino *et al*., 2022)). *LHCA6* did not show statistically significant changes in expression in *B. nigra*, *A. thaliana* and in any of the three copies of *H. incana*. However, it had lower expression in *B. rapa* under HL. To determine expression across species, we summed TPM-normalized counts for the three *LHCA6* copies in *H. incana* and calculated all pairwise ratios between counts in the four species under the two treatments. Inspection of ratios between counts in *H. incana* and other species revealed the *LHCA6* paralogs to have higher expression in *H. incana* under both the irradiance conditions after correction for the transcriptional baseline differences (Figures 4.4c, S7).

4.2.6 Targeted analysis of light reactions genes

We then analyzed the expression patterns for the CHGs containing the 77 *A. thaliana* nuclear and chloroplast genes that are annotated with the KEGG pathway "Photosynthesis" (ath00195), which are involved in the light reactions of photosynthesis. Log_2FC values ranged between 1.98 and -1.78, with 142 of the total 368 genes significantly DE under HL. When considering only the significant expression changes, the trend across "Photosynthesis" pathway genes is lower expression under HL in all species: 29 genes out of 33 in *A. thaliana*, 48 out of 52 in *B. rapa*, 25 out of 26 in *B. nigra*, and 27 out of 31 in *H. incana* (Figure S8). A small number of CHGs contained at least one gene having higher expression under HL in one of the species (Figure 4.4a). For further analysis, we focused on those showing higher expression only in *H. incana*, thus selecting genes *PSBD* and *PSBE*, part of the PSII complex, and *PSAA* and *PSAJ*, part of the PSI complex.

The D2 protein, encoded by *PSBD*, forms the core of PSII along with the D1 protein, encoded by *PSBA*. These two subunits together bind three macromolecules that

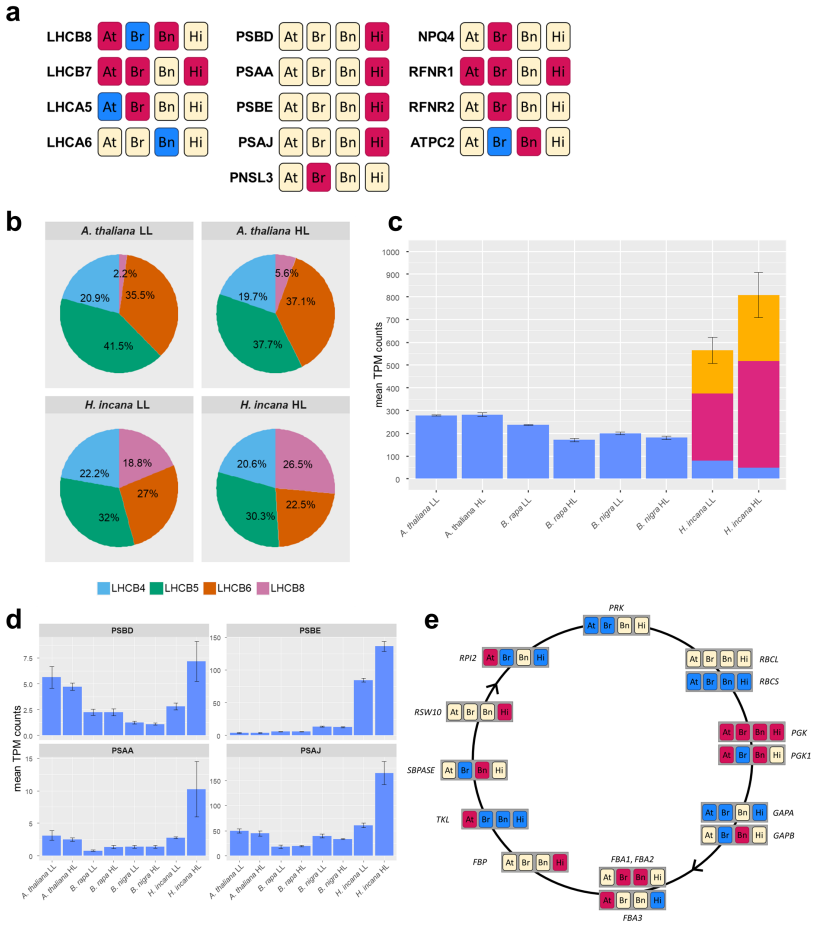

Figure 4.4: Results of the targeted analysis of photosynthetic pathways. (a) Schematic view of the restricted set of genes belonging to KEGG pathways ath00196 ("Photosynthesis - antenna proteins", left column) and ath00195 ("Photosynthesis", center and right columns) showing higher expression under HL in at least one of the species. Magenta box indicate significantly higher expression under HL, while blue boxes indicate significantly lower expression under HL. At: *A. thaliana*; Br: *B. rapa*; Bn: *B. nigra*; Hi: *H. incana*. **(b)** Pie charts representing the relative abundance of transcripts originating from genes encoding components of the PSII minor antenna. **(c)** Mean normalized abundance of *LHCA6* transcripts in plants of the four species grown under the two irradiance treatments. The different colors represent different paralogs. Error bars represent the standard errors of the mean. The full comparison of transcript abundances for *LHCA6* can be found in Figure S7. **(d)** Mean normalized abundances of *PSBE*, *PSBD*, *PSAA*, and *PSAJ* transcripts in plants of the four species grown under the two irradiance treatments. Error bars represent the standard errors of the mean. Full across-species comparisons can be found in Figures S9, S10, S11, and S12. **(e)** Schematic representation of the Calvin-Benson cycle, or the "dark reactions" of photosynthesis, and differential expression status of genes involved in the four species. Yellow: no significant differential expression under HL; red: significantly higher expression under HL; blue: significantly lower expression under HL. At: *A. thaliana*; Br: *B. rapa*; Bn: *B. nigra*; Hi: *H. incana*.

are fundamental for photosynthetic light reactions: the P680 reaction center, which transfers energy to water molecules, the Mn_4CaO_5 cluster responsible for the splitting of water molecules and retrieval of electrons, and components of the primary electron transfer chain, such as plastoquinones Q_A and Q_B (Leegood, 2013). The PSII reaction center is completed by the subunit encoded by the *PSBI* gene and cytochrome b559, composed of subunits encoded by the *PSBE* and *PSBF* genes and a heme cofactor (Johnson and Pakrasi, 2022). The *PSBD* gene is highly expressed in *A. thaliana* plants grown under both treatments, and *H. incana* plants grown under HL, while the *PSBE* gene is highly expressed only in *H. incana* plants grown under HL (Figures 4.4d, S9, S10).

The photosystem I (PSI) core is composed of proteins encoded by the *PSAA* and *PSAB* genes. The PSI complex is composed of several additional subunits, including one stabilized by the protein encoded by gene *PSAJ* (Scheller *et al.*, 2001). The expression of *PSAA* and *PSAJ* orthologs appears to be significantly higher in *H. incana* plants grown under HL, with plants growing under LL having similar transcript levels to those measured in the other species irrespective of the treatment (Figures 4.4d, S11, S12).

4.2.7 Targeted analysis of carbon metabolism genes

Continuing our analysis based on photosynthesis KEGG-related pathways, we studied the expression of the 273 *A. thaliana* nuclear and chloroplast genes associated to the KEGG pathway "Carbon metabolism" (ath01200) and their orthologs (Figure S13). Pathway ath01200 comprises genes involved in both catabolism and anabolism of carbon-based molecules, organized in a number of modules. Inspection of these modules revealed that the genes related to the Calvin-Benson cycle, and thus to assimilation of inorganic carbon into the end product of photosynthetic reactions, carbohydrates, were grouped into module "Reductive pentose phosphate cycle (Calvin cycle)" (ath_M00165). The expression of genes included in this module did not show an obvious profile (Figure S14). However, two of the 23 CHGs associated with this module contained genes that had higher expression uniquely in *H. incana* under HL. These are orthologs of the *A. thaliana* genes *FBP* and *RSW10* (Figure 4.4e). *A. thaliana* mutants for the *RSW10* gene has been linked to ribose-5-phosphate metabolism and cellulose biosynthesis, but no direct involvement with photosynthetic activity has been described to date (Howles *et al.*, 2006; Xiong *et al.*, 2009). Gene *FBP*, instead, has been associated with photosynthetic activity, and *FBP* overexpression has been proven to increase soluble sugar and starch contents, as well as photosynthetic CO_2 assimilation (Cho *et al.*, 2012).

4.3 Discussion

In this study, we explored the transcriptomes of plants of four Brassicaceae species (*A. thaliana*, *B. rapa*, *B. nigra*, *H. incana*) grown under contrasting irradiances to unravel the genetic determinants of *H. incana*'s high photosynthetic light-use efficiency under high irradiance. Considering the complexity of our dataset and based on the results of our untargeted enrichment analysis, we decided to restrict our exploration by focusing on genes related to photosynthesis in the KEGG ontology.

4.3.1 Increasing power of elimination through super-natural irradiance

To make sure we would observe any transcriptional differences associated with growth under high irradiance conditions, we designed and built a high-output, high-uniformity lighting system. The per-treatment daily light integral (DLI) of 9.82 mol m^{-2} d^{-1} that we measured for the low light (LL - 227.5 µmol m^{-2} s^{-1}) treatment is consistent with what has been reported for winter months in warm-temperate climate areas, while the DLI of 79.64 mol m^{-2} d^{-1} we measured for the high light treament (HL - 1843.6 µmol m^{-2} s^{-1}) is substantially higher than the values of 60-65 mol m^{-2} d^{-1} reported for summer months in the same climate areas (ENEA TER-SOLTERM, 2006; Faust and Logan, 2018; Korczynski *et al.*, 2002; Australian Government, 2022).

Our study differs from previous studies on high light responses not only because of our use of the "super-natural" magnitude of our high light treatment, but also for the way the treatment was applied. While previous high-light studies involving *A. thaliana* have employed irradiances ranging from between 150 and 2000 µmol m^{-2} s^{-1}, all of these studies applied the high light treatment to low light-adapted plants and focused on the response, or acclimation, to the high light (van Rooijen *et al.*, 2018; Pech *et al.*, 2022; Tiwari *et al.*, 2021; Bode *et al.*, 2016; Caldana *et al.*, 2011; Huang *et al.*, 2019; Page *et al.*, 2012; Alvarez-Fernandez *et al.*, 2021; Bobrovskikh *et al.*, 2022). We, on the other hand, focused on the steady-state transcriptional activity in the four species we examined grown from the seedling stage to maturity under either low or high light.

We have shown that between 68.9 and 85.3% of genes from the four species were not differentially expressed between the light treatments (Table 4.1). Furthermore, after performing homology grouping and integrating its results with gene differential expression analysis we identified a total of 44,663 HGs containing genes that did not respond to the treatment (Table 4.2), as well as 631 out of the total 10,352 CHGs containing genes that have the same response to the treatment in all species (Figure 5.8). None of these genes can, therefore, cause the higher photosynthetic LUE under high light of *H. incana*, and were therefore not considered in our further analysis. We thus believe that the combination of magnitude and application of treatment in our study gives us a sizeable "power of elimination" when dealing with complex transcriptomic datasets.

4.3.2 Dealing with the complexity of across-species transcriptomic comparisons

The limited set of studies comparing the transcriptomes of different plant species (Julca *et al.*, 2021; Aubry *et al.*, 2014; Wu *et al.*, 2022; Curci *et al.*, 2022; García de la Torre *et al.*, 2021) is proof of the novelty of between-species comparative transcriptomics. In contrast to previous studies, we used a panproteome built with PanTools (Sheikhizadeh *et al.*, 2016) to infer gene homology relationships. We made use of optimised homology grouping, based on the organization of universal single-copy orthologs (BUSCO gene sets, (Manni *et al.*, 2021b,a)), that is unique to PanTools (Sheikhizadeh Anari *et al.*, 2018; Jonkheer *et al.*, 2022). This method determines the optimal strictness of protein-clustering settings, given the phylogenetic distance between the proteomes in the data set.

Integrating HGs with per-species transcript abundance and differential expression data presented us with the challenge of comparing transcript abundances across species. Canonical normalization methods, such as the transcripts-per-million (TPM) normalization we used in our study, do not yield abundance measures that can be compared between species (Zhao *et al.*, 2020). In the absence of a widely accepted approach to compare normalised transcript abundances across species, we decided to estimate the transcriptional "baseline" of the four species. We extracted expression data for all the non-differentially expressed (non-DE) genes belonging to single-copy core HGs and calculating gene-by-gene $\log_2$-ratios between transcript abundances. Inspection of the distributions of these ratios revealed that *A. thaliana* has on average a slightly higher transcriptional baseline than *B. rapa* and *H. incana*, which in turn have a slightly higher baseline than *B. nigra*. We decided to control for these differences when comparing transcript abundance across species by calculating pairwise $\log_2$-ratios between TPM-normalized transcript counts and relating them to the ratios calculated for non-DE genes. As we have shown for the genes highlighted in the pathway analyses, the differences between TPM counts are much larger than what could be explained by differences in "baseline" transcription (Figures S7, S9, S10, S11, S12), and therefore have biological meaning.

4.3.3 Across-species comparison of differential gene expression highlights differences in photosynthetic pathways

The analysis of differential gene expression we performed individually on all four species in this study revealed similar percentages of differentially expressed (DE) genes for *A. thaliana*, *B. rapa*, and *H. incana*. Indeed, the cumulative percentage of DE genes in response to HL ranged between 23% and 31% (Table 4.1). This is in line with previous studies reporting that roughly 20% of the *A. thaliana* transcriptome is responsive to light (Ruckle *et al.*, 2012; Bode *et al.*, 2016). For *B. nigra*, on the other hand, only about 15% of the genes were DE under high light. We do not believe that this difference has a biological explanation, but that it is the result of the very large number of gene models included in the *B. nigra* annotation. Many of these

gene models are likely to be annotation artifacts rather than actual genes, as shown by the large number of *B. nigra* genes clustering separately from genes of the other species in the panproteome (Figures 4.2, S3).

After quantifying gene expression for our four species and inferring homology relationships between genes, we performed a number of untargeted analyses aimed at giving us a non-biased overview of the biological processes and pathways most affected when comparing the transcriptomes from the LL and HL treatments. By making use of the panproteome, we were able to perform these analyses on specific groups of genes, namely the core homology groups containing DE genes in all four species combined, the Brassiceae species (*B. rapa*, *B. nigra*) and *H. incana*) as a group, and *H. incana* on its own. Considering the higher photosynthesis rates we previously reported for the Brassiceae species (Garassino *et al.*, 2022), one might expect results linked to photosnynthetic LUE to come from the HGs showing differential expression for the Brassiceae species, or from the HGs with genes showing DE in *H. incana* alone. Nevertheless, the most promising results came from enrichment analyses on the HGs containing genes deferentially expressed across all four species. Indeed, out of a total of nine KEGG pathways enriched in this kind of HGs, three pathways mentioned photosynthesis in their name.

One striking finding of our targeted analysis of the three photosynthesis-related pathways was that most of the associated genes appeared to either have lower or unchanged expression under the HL treatment. This was expected for the "Photosynthesis - antenna proteins" pathway (ath00196), including all photosystem antenna genes, based on experimental evidence that plants growing under high light will reduce the size of their antennas (Ballottari *et al.*, 2007). However, this trend of lower or unchanged gene expression was unexpected for genes related to photosynthetic light reactions (included in the "Photosynthesis" pathway, ath00195) and carbon metabolism (included in the homonymous pathway, ath01200). Recent studies of changes in the *A. thaliana* proteome in response to irradiance increase or switch from controlled to field conditions have highlighted increases in abundance for most proteins involved in light reactions (Flannery *et al.*, 2021a,b). Furthermore, ample experimental evidence has been collected in the past showing that plants acclimating to high light develop a higher carbon fixation metabolism via increased protein levels (Schöttler and Tóth, 2014). A few considerations arise from the discrepancy between this evidence and the results of our transcriptome analysis. The first is that, as already discussed above, previous studies focused on acclimation responses to higher light, while ours was conducted on plants that grew under constant high or low light, and therefore the transcriptome snapshot obtained in our study might represents a much different gene and protein regulation situation than what was previously studied. Furthermore, it is important to point out that while transcriptome analysis highlights genes that are potentially involved in high photosynthetic LUE, it cannot inform us on downstream proteome dynamics. Thus, we currently cannot say whether higher gene expression is a consequence of higher protein turnover due to e.g. photodamage, or if it enables for higher protein abundance, thus potentially enabling for higher biochemical capacity in the photosynthetic reactions. The opposite is naturally true for lower gene expression, and therefore this analysis does not allow us to con-

clude whether that is the result of higher protein stability or lower protein abundance requirements.

Finally, our in-depth analysis of gene expression for three KEGG pathways revealed that differential gene expression is only one of the ways *H. incana* achieves higher transcript abundances, potentially enabling its higher photosynthetic light-use efficiency. While we identified four genes encoding photosystem subunits (*PSBD*, *PSBE*, *PSAA*, *PSAJ*) whose transcript levels were significantly higher in *H. incana* plants grown under HL, we identified other genes such as *LHCB8* and *LHCA6* having a striking transcript abundance in *H. incana* plants grown under both irradiances. This appears to be achieved in two additional ways: while the *LHCB8* gene is present in a single copy in *H. incana* and all other species, and the abundance of its transcript in *H. incana* can be explained with a constitutive overexpression of the gene, the *LHCA6* gene is present in three copies in *H. incana* as opposed as the single copy of the other three species. Each *LHCA6* copy is expressed in *H. incana* at levels that appear to be slightly higher than those of other species, but the cumulative expression of the three copies results in a substantially higher transcript abundance for the gene. These strategies to achieve higher gene expression form an interesting lead to further investigate the precise relationship between expression levels, protein abundance and turnover, and ultimately the photosynthetic light-use efficiency of *H. incana*.

4.3.4 Possibilities to explore other processes related to photosynthesis

While we decided to limit our research to KEGG photosynthesis pathways, we acknowledge that photosynthesis is a highly complex process involving other key pathways. We hypothesize that genes involved in transpiration, heat dissipation, stress response, and nutrient uptake and cycling will play a role in supporting higher photosynthetic efficiency. While previous studies identified transcriptional responses to high irradiance connected to heat-shock response (van Rooijen *et al.*, 2018; Bobrovskikh *et al.*, 2022), ribosome biogenesis and transcriptional activity (Bobrovskikh *et al.*, 2022), lipid remodeling (van Rooijen *et al.*, 2018), flavonoid biosynthesis (Page *et al.*, 2012; Pech *et al.*, 2022), a comprehensive picture of these responses is still far from being available (Burgess *et al.*, 2023). Based on what emerged from our targeted analysis on photosynthetic pathways, approaching our dataset in a different way than via enrichment analysis will likely reveal how the these processes are playing a role in high-light photosynthesis. Our resource will therefore provide means to further explore the genetic basis of high photosynthetic efficiency under high light.

4.3.5 Prospects for future research

In this study, we have highlighted three different strategies that *H. incana* can employ to achieve higher transcript abundances for genes that potentially play a key role in its

photosynthetic efficiency. Given that our analysis pipeline allows the retrieval of TPM-normalized counts for all expressed genes in each of the employed species, and that we established a method to estimate baseline differences in transcript abundances for the various species, an additional study of transcript abundances irrespective of differential expression might provide further clues on the mechanisms allowing *H. incana* to achieve higher photosynthetic light-use efficiency. Despite showing that most genes involved in photosynthetic reactions have lower expression as a response to high light in all analysed species, we have identified a number of genes that are either highly expressed in response to high light or have a constitutive higher expression in *H. incana*. Of these genes, *LHCB8* and *LHCA6* appear as very promising targets for further analysis, as the function of the first is still unclear and the higher expression in *H. incana* of the second cannot be explained with current literature.

It is important to stress once more how this experiment aimed at obtaining a snapshot of the operation of high photosynthesis rates, rather than at their establishment during leaf development. While our experiment uncovered some genes that might be playing a role in supporting high photosynthetic activity under high irradiance, future transcriptomics investigations on time series collected throughout leaf development will be crucial to understand which genes and processes enable the establishment of high photosynthetic light-use efficiency.

4.4 Conclusions

This study provides an analysis of the transcriptomes of *A. thaliana*, *B. rapa*, *B. nigra*, and *H. incana* plants grown under constant low and high irradiance, rather than the acclimation response to high irradiance. By combining gene expression quantification and differential expression analysis with a panproteome-based homology grouping, we quickly and efficiently identified expression patterns shared by the various species, or unique to one of them. Following an untargeted approach, we observed an enrichment for genes involved in photosynthetic pathways. A closer look at the expression of all genes belonging to these pathways allowed us to reveal that in comparison to other Brassicaceae species, *H. incana* growing under a high light treatment achieves higher expression of genes related to photosynthesis via three different modes: “canonical” differential expression between low and high light, constitutive higher expression of single-copy genes, or cumulative higher expression obtained by simultaneous expression of multiple gene copies. Besides identifying genes such as *LHCB8* and *LHCA6*, whose higher expression in *H. incana* growing under high light prompts for a detailed investigation of their role in photosynthetic LUE under high irradiance, we believe that analysing the genes undergoing differential expression specifically in *H. incana* will further clarify the role of non-strictly photosynthetic genes in supporting the species’ striking photosynthetic performance. Therefore, we expect the resource we established with this study to provide further, extensive knowledge on the genetic strategy employed by *H. incana* to support its high photosynthetic light-use efficiency.

4.5 Materials and methods

Construction of high-uniformity growth system

Two custom light ceilings were built for this study. Each ceiling measured approximately 4.3 m^2 (*l* 175 cm, *w* 245 cm), was equipped with six dimmable VYPR2p LED fixtures (Fluence, Austin, USA) arranged in three equally spaced rows (between-rows distance of 60 cm, and was hung so that fixtures would be one meter high over plants. We then centered two custom-made growing tables measuring approximately 1.6 m^2 (*l* 118 cm, *w* 137 cm) under the custom light ceilings, divided each table into thirty growing areas, each measuring approximately 0.05 m^2, and performed irradiance measurements at the centre of each growing area. By calculating averages over the thirty areas under each light ceiling, we optimised the output of the LED fixtures to have average irradiances as close as the reference values we chose for our treatments.

Plant material and growing conditions

Hirschfeldia incana accession 'Nijmegen', *Brassica nigra* accession 'DG1', *Brassica rapa* R-o-18, and *Arabidopsis thaliana* Col-0 were used for this experiment. 'Nijmegen' is an inbred line (over six rounds of inbreeding) from an *H. incana* plant originally collected in Nijmegen, The Netherlands; 'DG1' is a second-generation inbreeding line of *B. nigra* sampled from a natural population near Wageningen, The Netherlands; and 'R-o-18' is a *B. rapa* inbred line (Stephenson *et al.*, 2010; Bagheri *et al.*, 2012).

Seeds of all species were germinated on a peat-based potting mix for nine days under an irradiance of 200 µmol m^{-2} s^{-1}. Twelve healthy seedlings per species were then transferred to 2 L pots (⌀ 13.9 cm, h 17.4 cm, Soparco, Condé-sur-Huisne, France) filled with a peat-based potting mixture enriched with perlite and 2.5 g/L Osmocote® Exact Standard 5-6M slow-release fertiliser (ICL Specialty Fertilizers, Geldermalsen, The Netherlands).

Plants were germinated and grown in a climate-controlled room equipped with the custom arrays of high-output LED light modules described above, with a photoperiod of 12 h day and 12 h night, and air temperature set at 23 °C and 20 °C, respectively. Humidity and CO_2 levels were set at 70% and 400 ppm. Six plants per species were assigned to the high light (HL) treatment of 1800 µmol m^{-2} s^{-1} (measured irradiance average: 1843.6 µmol m^{-2} s^{-1}) and the remaining six to the low light (LL) treatment of 200 µmol m^{-2} s^{-1} (measured irradiance average: 227.5 µmol m^{-2} s^{-1}). Irradiance uniformity was very high for both HL and LL treatments, with a U_2 value (defined as *minimum irradiance/maximum irradiance*, (Hu *et al.*, 2015; Sun *et al.*, 2014)) of 0.93. Plant positions were randomised across growing areas. Plants assigned to the LL treatment were fertigated daily, while plants assigned to the HL treatment were fertigated twice a day, with a custom nutrient solution (0.6 mM NH_4^+, 3.6 mM K^+, 2 mM Ca^{2+}, 0.91 mM Mg^{2+}, 6.2 mM NO_3^-, 1.66 mM SO_4^{2-}, 0.5 mM P, 35 µM Fe^{3+}, 8 µM Mn^{2+}, 5 µM Zn^{2+}, 20 µM B, 0.5 µM Cu^{2+}, 0.5 µM Mo^{4+}).

Sampling and RNA extraction

Twenty-eight days after sowing, samples representative of the whole canopy were collected from all plants. All leaves (for smaller plants such as *A. thaliana* and *H. incana*, especially when grown under low light) or half the total number of leaves were excised from plants, transferred to 50 mL tubes, and flash-frozen in liquid nitrogen. All leaf samples were subsequently crushed with a mortar and pestle in excess liquid nitrogen, and further homogenised with glass beads for 2 min at 30 Hz in a MM300 Mixer Mill (Retsch GmbH, Haan, Germany). Total RNA was extracted with the RNeasy Plant Mini Kit (QIAGEN N.V., Venlo, The Netherlands) according to manufacturer's instructions, and eluted using 50 µL of DNAse/RNAse-free water. The following DNAse treatment and RNA recovery were performed as described in (Oñate-Sánchez and Vicente-Carbajosa, 2008). 6 µL of 10X DNAse buffer and 4 µL of RQ1 DNAse (Promega, Leiden, The Netherlands) were added to 50 µL of RNA, and incubated for 30 minutes at 37 ℃. The RNA was then precipitated overnight using ammonium acetate and ethanol, and resuspended in 25 µL of DNAse/RNAse-free water. To check RNA quality and integrity, 1 µL of RNA was used to (1) load a 1% agarose-Ethidium bromide gel and after electrophoresis observe the bands using standard imaging and (2) to determine spectrophotometric parameters with a Nanodrop 2000 (Thermo Fisher Scientific Inc., Waltham, U.S.A.). The RNA was further quantified using the Qubit RNA BR Assay kit and a Qubit 4 fluorometer (Thermo Fisher Scientific Inc., Waltham, U.S.A.).

Sequencing

RNA from five of the six plants of each species grown under each light treatment was sequenced by Novogene (UK) Company Ltd., Cambridge, U.K.. Poly-A enriched RNA was employed to prepare sequencing libraries with the NEBNext® Ultra™ RNA Library Prep Kit (New England Biolabs, Ipswich, U.S.A.). Paired-end, 150-bp-long reads (PE150) were generated with a NovaSeq 6000 system (Illumina Inc., San Diego, U.S.A.) aiming at obtaining 6 Gb of data per sample.

Selection and preparation of genome assemblies and annotations

For mapping of sequencing reads and quantification of gene expression, the TAIR10 (Swarbreck *et al.*, 2008) genome assembly and the Araport11 annotation (Cheng *et al.*, 2017) were used for *A. thaliana*, the "Chiifu" v3.0 assembly and annotation (Zhang *et al.*, 2018) were used for *B. rapa*, the "Ni100" v2.0 assembly and annotation (Perumal *et al.*, 2020) were used for *B. nigra*, and the "NIJ6" v1.0 assembly and annotation (Garassino *et al.*, 2022) were used for *H. incana*. For panproteome building, the v3.0/3.1 *A. arabicum* (Fernandez-Pozo *et al.*, 2021), the v1.0 *R. raphanistrum* (Moghe *et al.*, 2014), the v1.0 *R. sativus* (Kitashiba *et al.*, 2014), and the v1.0 *S. irio* (Haudry *et al.*, 2013) genome assemblies and annotations were employed together with the aforementioned ones. The exact locations where the various files were downloaded from can be found in Table S13.

Statistics were collected for all genome assemblies and annotations with custom Python (v3.11.0) scripts and are reported in Table S14. Given that not all genome annotations contained multiple transcript isoforms, all GFF files were processed with the agat_sp_keep_longest_isoform.pl script from the AGAT toolkit v1.0.0 (Dainat, 2021) to generate annotations containing only the longest transcript isoforms of all gene models. Subsequently, these GFF files were filtered with the agat_sp_filter_by_ORF_size.pl script to remove all gene models that would have yielded protein sequences shorter than 30 amino acids. Finally, a number of gene models identified in the *R. raphanistrum* and *R. sativus* annotation that would still not result in protein sequences (due to stop codons embedded in their sequence) were removed from the corresponding annotations with the agat_sp_filter_feature_from_kill_list.pl script. The resulting filtered annotation files are provided with the data package linked to this article.

Identification and Analysis of Differentially Expressed Genes

The quality of sequencing libraries was assessed with MultiQC (Ewels *et al.*, 2016) v1.11. A snakemake (v7.19.1) (Mölder *et al.*, 2021) pipeline was employed to automate subsequent read mapping and transcript quantification steps. Reads were aligned to reference genome assemblies with two passes of the STAR (Dobin *et al.*, 2013) v2.7.10a aligner (STAR indexing running with parameters -sjdbOverhang 139 and –genomeSAindexNbases 13, STAR aligner running with parameter –clip5pNbases 10 10). Assembly and quantification of full-length transcripts were then achieved with StringTie (Pertea *et al.*, 2015) v2.2.1 (running with option -e). Per-sample gene and transcripts counts were then grouped by species with the prepDE Python script included in the StringTie suite (running with parameter -l 140). Transcripts per million (TPM) counts (Wagner *et al.*, 2012) were extracted for visualisation purposes from the StringTie output with a custom Python script.

Relationships between samples of the same species were explored with PCA plots of transcript counts transformed by means of regularized logarithm (Love *et al.*, 2014). Differentially Expressed Genes (DEGs) were subsequently identified with DESeq2 (Love *et al.*, 2014) v1.34.0 running in R (R Core Team, 2021) v4.1.1.

Panproteome construction

Proteomes were created from the filtered annotations of all eight species with the AGAT toolkit agat_sp_extract_sequences.pl script, running with options -p, –cis, and –cfs. A panproteome was subsequently constructed by running PanTools v4.1.0 (Sheikhizadeh Anari *et al.*, 2018; Jonkheer *et al.*, 2022) commands build_panproteome, busco_protein, (with options -if brassicales_odb10 –version busco4), optimal_grouping, and change_grouping (with option –version 4, and thus running with a relaxation parameter of 4). A separate panproteome was constructed featuring chloroplast proteomes for *A. thaliana*, *B. rapa*, *B. nigra*, and *H. incana* with PanTools commands build_panproteome and group (with the same relaxation parameter of 4). The

panproteome was visualized by making UpSet plots (Lex *et al.*, 2014) with the ComplexUpset package (v1.3.3) running in R v4.4.2.

Integration of panproteome and DE results

The homology table resulting from panproteome construction was integrated with Differential Expression analysis results by means of a custom script running in Python v3.10.9 and leveraging NumPY v1.24.1 (Harris *et al.*, 2020), and Pandas v1.5.3 (McKinney, 2010). The resulting homology/DE status table was further processed and visualised with a custom script running in R v4.2.2. A heatmap of non-ambiguously responding core groups was generated with the Pheatmap v1.0.12 package. After specific categories of homology groups were selected, a Gene Ontology (GO) Biological Process (BP) enrichment analysis was performed for the *A. thaliana* gene identifiers present in said groups with TopGO v2.50.0 (Alexa *et al.*, 2006), relying on the org.At.tair.db v3.16.0 Bioconductor annotation data package, running the "Classic" algorithm, and performing Fisher tests. Enrichment results for each set of groups were filtered by keeping only terms which were associated to at least five genes. A KEGG pathway enrichment analysis was subsequently performed on the *A. thaliana* genes present in the same categories of homology groups with the enrichKEGG function of the ClusterProfiler v4.6.2 package (Yu *et al.*, 2012; Wu *et al.*, 2021). For both enrichment analyses, the set of background genes (i.e., the analysis "universe") was composed by all *A. thaliana* genes surviving the DE analysis (i.e., genes for which an adjusted *p*-value could be calculated by DESeq2).

Processing and visualization targeted analysis results

Expression profiles, TPM-normalized counts and homology relationships were processed and visualized with custom R scripts making use of packages dplyr (v1.1.0), ggplot2 (v3.4.1), janitor (v2.2.0), pheatmap (v1.0.12), scales(v1.2.1), stringr (v1.5.0), tidyr (v1.3.0). All scripts are published with DOI 10.4121/5b88cdf2-eb5f-4033-8ece-1f3f488a1f83.

Availability of data and materials

Raw sequencing data has been deposited to NCBI and can be found under BioProject PRJNA1001172. All scripts used for data analysis are available on the 4TU.ResearchData platform with DOI 10.4121/5b88cdf2-eb5f-4033-8ece-1f3f488a1f83. Supplementary information is available on the 4TU.ResearchData platform as well, with DOI 10.4121/d3455b3c-54d8-4ef8-8501-a70936a51dad.v1.

Author contributions

F.G. and S.S. designed the study in consultation with M.A. and J.H.. F.G. grew and sampled plants. F.R.M. extracted and evaluated RNA from all samples. T.C. performed preliminary transcriptome analyses, which F.G. extended and finalized with help from S.B.L.. H.N. and S.S supervised transcriptome analysis. F.G., S.B.L., H.N., and S.S. were involved in writing the manuscript. M.A. and J.H. provided guidance on the biological interpretation of results and revised the manuscript. All authors have read and approved the contents of the manuscript.

Acknowledgements

We thank Ludovico Caracciolo for his help during the building of the growth setup and for helping us with plant growth and sampling. We thank Nam Hoang for the critical discussions on the methodology and results of this study, as well as for helping us with plant sampling. Furthermore, we thank Vittoria Clapero for her involvement in the early stages of this study, Robin van Esch and Dirk-Jan van Workum for their support for parts of the analysis, and Eric Schranz for his useful contributions in discussions on this study.

References

Alexa, A., Rahnenführer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, **22**(13), 1600–1607.

Alvarez-Fernandez, R., Penfold, C. A., Galvez-Valdivieso, G., Exposito-Rodriguez, M., Stallard, E. J., Bowden, L., Moore, J. D., Mead, A., Davey, P. A., Matthews, J. S. A., Beynon, J., Buchanan-Wollaston, V., Wild, D. L., Lawson, T., Bechtold, U., Denby, K. J., and Mullineaux, P. M. (2021). Time-series transcriptomics reveals a BBX32-directed control of acclimation to high light in mature Arabidopsis leaves. *The Plant Journal*, **107**(5), 1363–1386.

Araus, J. L., Sanchez-Bragado, R., and Vicente, R. (2021). Improving crop yield and resilience through optimization of photosynthesis: Panacea or pipe dream? *Journal of Experimental Botany*, **72**(11), 3936–3955.

Arias, T. and Pires, J. C. (2012). A fully resolved chloroplast phylogeny of the brassica crops and wild relatives (Brassicaceae: Brassiceae): Novel clades and potential taxonomic implications. *TAXON*, **61**(5), 980–988.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: Tool for the unification of biology. *Nature genetics*, **25**(1), 25–29.

Aubry, S., Kelly, S., Kümpers, B. M. C., Smith-Unna, R. D., and Hibberd, J. M. (2014). Deep Evolutionary Comparison of Gene Expression Identifies Parallel Recruitment of Trans-Factors in Two Independent Origins of C4 Photosynthesis. *PLOS Genetics*, **10**(6), e1004365.

Austin, R. B. (1989). Genetic variation in photosynthesis. *The Journal of Agricultural Science*, **112**(3), 287–294.

Australian Government, B. o. M. (2022). Gridded Monthly Solar Exposure.

Bagheri, H., El-Soda, M., van Oorschot, I., Hanhart, C., Bonnema, G., Jansen-van den Bosch, T., Mank, R., Keurentjes, J., Meng, L., Wu, J., *et al.* (2012). Genetic analysis of morphological traits in a new, versatile, rapid-cycling *Brassica rapa* recombinant inbred line population. *Frontiers in Plant Science*, **3**, 183.

Ballottari, M., Dall'Osto, L., Morosinotto, T., and Bassi, R. (2007). Contrasting Behavior of Higher Plant Photosystem I and II Antenna Systems during Acclimation *. *Journal of Biological Chemistry*, **282**(12), 8947–8958.

Bobrovskikh, A. V., Zubairova, U. S., Bondar, E. I., Lavrekha, V. V., and Doroshkov, A. V. (2022). Transcriptomic Data Meta-Analysis Sheds Light on High Light Response in Arabidopsis thaliana L. *International Journal of Molecular Sciences*, **23**(8), 4455.

Bode, R., Ivanov, A. G., and Hüner, N. P. A. (2016). Global transcriptome analyses provide evidence that chloroplast redox state contributes to intracellular as well as long-distance signalling in response to stress and acclimation in Arabidopsis. *Photosynthesis Research*, **128**(3), 287–312.

Burgess, A. J., Masclaux-Daubresse, C., Strittmatter, G., Weber, A. P. M., Taylor, S. H., Harbinson, J., Yin, X., Long, S., Paul, M. J., Westhoff, P., Loreto, F., Ceriotti, A., Saltenis, V. L. R., Pribil, M., Nacry, P., Scharff, L. B., Jensen, P. E., Muller, B., Cohan, J.-P., Foulkes, J., Rogowsky, P., Debaeke, P., Meyer, C., Nelissen, H., Inzé, D., Klein Lankhorst, R., Parry, M. A. J., Murchie, E. H., and Baekelandt, A. (2023). Improving crop yield potential: Underlying biological processes and future prospects. *Food and Energy Security*, **12**(1), e435.

Caldana, C., Degenkolbe, T., Cuadros-Inostroza, A., Klie, S., Sulpice, R., Leisse, A., Steinhauser, D., Fernie, A. R., Willmitzer, L., and Hannah, M. A. (2011). High-density kinetic analysis of the metabolomic and transcriptomic response of Arabidopsis to eight environmental conditions. *The Plant Journal*, **67**(5), 869–884.

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: A complete reannotation of the Arabidopsis thaliana reference genome. *The Plant Journal*, **89**(4), 789–804.

Cho, M.-H., Jang, A., Bhoo, S. H., Jeon, J.-S., and Hahn, T.-R. (2012). Manipulation of triose phosphate/phosphate translocator and cytosolic fructose-1,6-bisphosphatase, the key components in photosynthetic sucrose synthesis, enhances the source capacity of transgenic Arabidopsis plants. *Photosynthesis Research*, **111**(3), 261–268.

Curci, P. L., Zhang, J., Mähler, N., Seyfferth, C., Mannapperuma, C., Diels, T., Van Hautegem, T., Jonsen, D., Street, N., Hvidsten, T. R., Hertzberg, M., Nilsson, O., Inzé, D., Nelissen, H., and Vandepoele, K. (2022). Identification of growth regulators using cross-species network analysis in plants. *Plant Physiology*, **190**(4), 2350–2365.

Dainat, J. (2021). AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format.

de Bianchi, S., Betterle, N., Kouril, R., Cazzaniga, S., Boekema, E., Bassi, R., and Dall'Osto, L. (2011). Arabidopsis mutants deleted in the light-harvesting protein Lhcb4 have a disrupted photosystem II macrostructure and are defective in photoprotection. *The Plant Cell*, **23**(7), 2659–2679.

De Souza, A. P., Burgess, S. J., Doran, L., Hansen, J., Manukyan, L., Maryn, N., Gotarkar, D., Leonelli, L., Niyogi, K. K., and Long, S. P. (2022). Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science*, **377**(6608), 851–854.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**(1), 15–21.

ENEA TER-SOLTERM (2006). ENEA - Solaritaly — DLI of Italy. http://www.solaritaly.enea.it/DLI/DLIMappeEn.php.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**(19), 3047–3048.

Faust, J. E. and Logan, J. (2018). Daily Light Integral: A Research Review and High-resolution Maps of the United States. *HortScience horts*, **53**(9), 1250–1257.

Fernandez-Pozo, N., Metz, T., Chandler, J. O., Gramzow, L., Mérai, Z., Maumus, F., Mittelsten Scheid, O., Theißen, G., Schranz, M. E., Leubner-Metzger, G., and Rensing, S. A. (2021). Aethionema arabicum genome annotation using PacBio full-length transcripts provides a valuable resource for seed dormancy and Brassicaceae evolution research. *The Plant Journal*, **106**(1), 275–293.

Flannery, S. E., Pastorelli, F., Wood, W. H. J., Hunter, C. N., Dickman, M. J., Jackson, P. J., and Johnson, M. P. (2021a). Comparative proteomics of thylakoids from Arabidopsis grown in laboratory and field conditions. *Plant Direct*, **5**(10), e355.

Flannery, S. E., Hepworth, C., Wood, W. H. J., Pastorelli, F., Hunter, C. N., Dickman, M. J., Jackson, P. J., and Johnson, M. P. (2021b). Developmental acclimation of the thylakoid proteome to light intensity in Arabidopsis. *The Plant Journal*, **105**(1), 223–244.

Garassino, F., Wijfjes, R. Y., Boesten, R., Reyes Marquez, F., Becker, F. F. M., Clapero, V., van den Hatert, I., Holmer, R., Schranz, M. E., Harbinson, J., de Ridder, D., Smit, S., and Aarts, M. G. M. (2022). The genome sequence of Hirschfeldia incana, a new Brassicaceae model to improve photosynthetic light-use efficiency. *The Plant Journal*, **112**(1), 1298–1315.

García de la Torre, V. S., Majorel-Loulergue, C., Rigaill, G. J., Alfonso-González, D., Soubigou-Taconnat, L., Pillon, Y., Barreau, L., Thomine, S., Fogliani, B., Burtet-Sarramegna, V., and Merlot, S. (2021). Wide cross-species RNA-Seq comparison reveals convergent molecular mechanisms involved in nickel hyperaccumulation across dicotyledons. *New Phytologist*, **229**(2), 994–1006.

Garcia-Molina, A. and Leister, D. (2020). Accelerated relaxation of photoprotection impairs biomass accumulation in Arabidopsis. *Nature Plants*, **6**(1), 9–12.

Gene Ontology Consortium (2021). The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Research*, **49**(D1), D325–D334.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, **585**(7825), 357–362.

Hashida, Y., Tezuka, A., Nomura, Y., Kamitani, M., Kashima, M., Kurita, Y., and Nagano, A. J. (2022). Fillable and unfillable gaps in plant transcriptome under field and controlled environments. *Plant, Cell & Environment*, **45**(8), 2410–2427.

Haudry, A., Platts, A. E., Vello, E., Hoen, D. R., Leclercq, M., Williamson, R. J., Forczek, E., Joly-Lopez, Z., Steffen, J. G., Hazzouri, K. M., Dewar, K., Stinchcombe, J. R., Schoen, D. J., Wang, X., Schmutz, J., Town, C. D., Edger, P. P., Pires, J. C., Schumaker, K. S., Jarvis, D. E., Mandáková, T., Lysak, M. A., van den Bergh, E., Schranz, M. E., Harrison, P. M., Moses, A. M., Bureau, T. E., Wright, S. I., and Blanchette, M. (2013). An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nature Genetics*, **45**(8), 891–898.

Howles, P. A., Birch, R. J., Collings, D. A., Gebbie, L. K., Hurley, U. A., Hocart, C. H., Arioli, T., and Williamson, R. E. (2006). A mutation in an Arabidopsis ribose 5-phosphate isomerase reduces cellulose synthesis and is rescued by exogenous uridine. *The Plant Journal*, **48**(4), 606–618.

Hu, Y., Luo, M. R., and Yang, Y. (2015). A study on lighting uniformity for LED smart lighting system. *2015 12th China International Forum on Solid State Lighting, SSLCHINA 2015*, pages 127–130.

Huang, J., Zhao, X., and Chory, J. (2019). The Arabidopsis Transcriptome Responds Specifically and Dynamically to High Light Stress. *Cell Reports*, **29**(12), 4186–4199.e3.

Johnson, M. P. (2016). Photosynthesis. *Essays in Biochemistry*, **60**(3), 255–273.

Johnson, V. M. and Pakrasi, H. B. (2022). Advances in the Understanding of the Lifecycle of Photosystem II. *Microorganisms*, **10**(5), 836.

Jonkheer, E. M., van Workum, D.-J. M., Sheikhizadeh Anari, S., Brankovics, B., de Haan, J. R., Berke, L., van der Lee, T. A. J., de Ridder, D., and Smit, S. (2022). PanTools v3: Functional annotation, classification and phylogenomics. *Bioinformatics*, **38**(18), 4403–4405.

Julca, I., Ferrari, C., Flores-Tornero, M., Proost, S., Lindner, A.-C., Hackenberg, D., Steinbachová, L., Michaelidis, C., Gomes Pereira, S., Misra, C. S., Kawashima, T., Borg, M., Berger, F., Goldberg, J., Johnson, M., Honys, D., Twell, D., Sprunck, S., Dresselhaus, T., Becker, J. D., and Mutwil, M. (2021). Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. *Nature Plants*, **7**(8), 1143–1159.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, **28**(1), 27–30.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, **51**(D1), D587–D592.

Kitashiba, H., Li, F., Hirakawa, H., Kawanabe, T., Zou, Z., Hasegawa, Y., Tonosaki, K., Shirasawa, S., Fukushima, A., Yokoi, S., Takahata, Y., Kakizaki, T., Ishida, M., Okamoto, S., Sakamoto, K., Shirasawa, K., Tabata, S., and Nishio, T. (2014). Draft Sequences of the Radish (Raphanus sativus L.) Genome. *DNA Research*, **21**(5), 481–490.

Klimmek, F., Sjödin, A., Noutsos, C., Leister, D., and Jansson, S. (2006). Abundantly and Rarely Expressed Lhc Protein Genes Exhibit Distinct Regulation Patterns in Plants. *Plant Physiology*, **140**(3), 793–804.

Korczynski, P. C., Logan, J., and Faust, J. E. (2002). Mapping Monthly Distribution of Daily Light Integrals across the Contiguous United States. *HortTechnology*, **12**(1), 12–16.

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**(6314), 857–861.

Leegood, R. C. (2013). Photosynthesis. In W. J. Lennarz and M. D. Lane, editors, *Encyclopedia of Biological Chemistry (Second Edition)*, pages 492–496. Academic Press, Waltham.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. *IEEE transactions on visualization and computer graphics*, **20**(12), 1983–1992.

López-Calcagno, P. E., Fisk, S., Brown, K. L., Bull, S. E., South, P. F., and Raines, C. A. (2019). Overexpressing the H-protein of the glycine cleavage system increases biomass yield in glasshouse and field-grown transgenic tobacco plants. *Plant Biotechnology Journal*, **17**(1), 141–151.

López-Calcagno, P. E., Brown, K. L., Simkin, A. J., Fisk, S. J., Vialet-Chabrand, S., Lawson, T., and Raines, C. A. (2020). Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants*, **6**(8), 1054–1063.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12), 550.

Manni, M., Berkeley, M. R., Seppey, M., and Zdobnov, E. M. (2021a). BUSCO: Assessing Genomic Data Quality and Beyond. *Current Protocols*, **1**(12), e323.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. (2021b). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution*, **38**(10), 4647–4654.

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In *Python in Science Conference*, pages 56–61, Austin, Texas.

Moghe, G. D., Hufnagel, D. E., Tang, H., Xiao, Y., Dworkin, I., Town, C. D., Conner, J. K., and Shiu, S.-H. (2014). Consequences of Whole-Genome Triplication as Revealed by Comparative Genomic Analyses of the Wild Radish Raphanus raphanistrum and Three Other Brassicaceae Species. *The Plant Cell*, **26**(5), 1925–1937.

Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., Forster, J., Lee, S., Twardziok, S. O., Kanitz, A., Wilm, A., Holtgrewe, M., Rahmann, S., Nahnsen, S., and Köster, J. (2021). Sustainable data analysis with Snakemake.

Oñate-Sánchez, L. and Vicente-Carbajosa, J. (2008). DNA-free RNA isolation protocols for Arabidopsis thaliana, including seeds and siliques. *BMC research notes*, **1**(1), 1–7.

Page, M., Sultana, N., Paszkiewicz, K., Florance, H., and Smirnoff, N. (2012). The influence of ascorbate on anthocyanin accumulation during high light acclimation in Arabidopsis thaliana: Further evidence for redox control of anthocyanin synthesis. *Plant, Cell & Environment*, **35**(2), 388–404.

Pech, R., Volná, A., Hunt, L., Bartas, M., Červeň, J., Pečinka, P., Špunda, V., and Nezval, J. (2022). Regulation of Phenolic Compound Production by Light Varying in Spectral Quality and Total Irradiance. *International Journal of Molecular Sciences*, **23**(12), 6533.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, **33**(3), 290–295.

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., Sankoff, D., Robinson, S. J., Kagale, S., Navabi, Z.-K., Tang, L., Horner, K. N., He, Z., Bancroft, I., Chalhoub, B., Sharpe, A. G., and Parkin, I. A. P. (2020). A high-contiguity Brassica nigra genome localizes active centromeres and defines the ancestral Brassica genome. *Nature Plants*, **6**(8), 929–941.

R Core Team (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Ruckle, M. E., Burgoon, L. D., Lawrence, L. A., Sinkler, C. A., and Larkin, R. M. (2012). Plastids Are Major Regulators of Light Signaling in Arabidopsis. *Plant Physiology*, **159**(1), 366–390.

Ruiz-Vera, U. M., Acevedo-Siaca, L. G., Brown, K. L., Afamefule, C., Gherli, H., Simkin, A. J., Long, S. P., Lawson, T., and Raines, C. A. (2022). Field-grown ictB tobacco transformants show no difference in photosynthetic efficiency for biomass relative to the wild type. *Journal of Experimental Botany*, **73**(14), 4897–4907.

Salisbury, F. B. and Ross, C. W. (1992). *Plant Physiology*. Wadsworth Publishing Company.

Scheller, H. V., Jensen, P. E., Haldrup, A., Lunde, C., and Knoetzel, J. (2001). Role of subunits in eukaryotic Photosystem I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1507**(1), 41–60.

Schöttler, M. A. and Tóth, S. Z. (2014). Photosynthetic complex stoichiometry dynamics in higher plants: Environmental acclimation and photosynthetic flux control. *Frontiers in Plant Science*, **5**.

Schulze, E. D. and Caldwell, M. M. (1995). *Ecophysiology of Photosynthesis*. Springer Berlin Heidelberg.

Sheikhizadeh, S., Schranz, M. E., Akdel, M., de Ridder, D., and Smit, S. (2016). PanTools: Representation, storage and exploration of pan-genomic data. *Bioinformatics*, **32**(17), i487–i493.

Sheikhizadeh Anari, S., de Ridder, D., Schranz, M. E., and Smit, S. (2018). Efficient inference of homologs in large eukaryotic pan-proteomes. *BMC Bioinformatics*, **19**(1), 340.

Singh, J., Pandey, P., James, D., Chandrasekhar, K., Achary, V. M. M., Kaul, T., Tripathy, B. C., and Reddy, M. K. (2014). Enhancing C3 photosynthesis: An outlook on feasible interventions for crop improvement. *Plant Biotechnology Journal*, **12**(9), 1217–1230.

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G. J., and Østergaard, L. (2010). A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biology*, **10**(1), 1–10.

Sun, W.-S., Tien, C.-L., Tsuei, C.-H., and Pan, J.-W. (2014). Simulation and comparison of the illuminance, uniformity, and efficiency of different forms of lighting used in basketball court illumination. *Applied optics*, **53**(29), H186.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., and Huala, E. (2008). The Arabidopsis Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Research*, **36**(suppl_1), D1009–D1014.

Taylor, G., Garassino, F., Aarts, M. G. M., and Harbinson, J. (2022). Improving C3 photosynthesis by exploiting natural genetic variation: Hirschfeldia incana as a model species. *Food and Energy Security*, **n/a**(e420).

Theeuwen, T. P. J. M., Logie, L. L., Harbinson, J., and Aarts, M. G. M. (2022). Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany*, **73**(10), 3122–3137.

Tiwari, B., Habermann, K., Arif, M. A., Top, O., and Frank, W. (2021). Identification of Small RNAs During High Light Acclimation in Arabidopsis thaliana. *Frontiers in Plant Science*, **12**.

van Rooijen, R., Harbinson, J., and Aarts, M. G. M. (2018). Photosynthetic response to increased irradiance correlates to variation in transcriptional response of lipid-remodeling and heat-shock genes. *Plant Direct*, **2**(7), e00069.

Wagner, G. P., Kin, K., and Lynch, V. J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, **131**(4), 281–285.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., and Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, **2**(3), 100141.

Wu, T.-Y., Hoh, K. L., Boonyaves, K., Krishnamoorthi, S., and Urano, D. (2022). Diversification of heat shock transcription factors expanded thermal stress responses during early plant evolution. *The Plant Cell*, **34**(10), 3557–3576.

Xiong, Y., DeFraia, C., Williams, D., Zhang, X., and Mou, Z. (2009). Deficiency in a cytosolic ribose-5-phosphate isomerase causes chloroplast dysfunction, late flowering and premature cell death in Arabidopsis. *Physiologia Plantarum*, **137**(3), 249–263.

Yin, X., Gu, J., Dingkuhn, M., and Struik, P. C. (2022). A model-guided holistic review of exploiting natural variation of photosynthesis traits in crop improvement. *Journal of Experimental Botany*, **73**(10), 3173–3188.

Yoon, D.-K., Ishiyama, K., Suganami, M., Tazoe, Y., Watanabe, M., Imaruoka, S., Ogura, M., Ishida, H., Suzuki, Y., Obara, M., Mae, T., and Makino, A. (2020). Transgenic rice overproducing Rubisco exhibits increased yields with improved nitrogen-use efficiency in an experimental paddy field. *Nature Food*, **1**(2), 134–139.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, **16**(5), 284–287.

Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., Liang, J., Cai, C., Liu, Z., Liu, B., Wang, F., Li, S., Liu, F., Li, X., Cheng, L., Yang, W., Li, M.-h., Grossniklaus, U., Zheng, H., and Wang, X. (2018). Improved Brassica rapa reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Horticulture Research*, **5**, 50.

Zhao, S., Ye, Z., and Stanton, R. (2020). Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *RNA*, **26**(8), 903–909.

Zhu, X.-G., Long, S. P., and Ort, D. R. (2010). Improving Photosynthetic Efficiency for Greater Yield. *Annual Review of Plant Biology*, **61**(1), 235–261.

CHAPTER 5

Analysis of natural variation in photosynthesis in a panel of Brassicaceae species

Francesco Garassino[1], Ludovico Caracciolo[2], José van de Belt[1], Marco Albiero[1], Rachel Pool[1], Femke Schreurs[1], Dick de Ridder[3], Jeremy Harbinson[2], and Mark G. M. Aarts[1]

(1) Laboratory of Genetics, Wageningen University & Research, Wageningen (NL)

(2) Laboratory of Biophysics, Wageningen University & Research, Wageningen (NL)

(3) Bioinformatics Group, Wageningen University & Research, Wageningen (NL)

Supplementary material available at:
https://doi.org/10.4121/79a62b5f-2881-4520-b031-e03334c02aad

Abstract

Photosynthesis research holds great potential to facilitate the transition to more sustainable agriculture, given its central role in the solar energy conversion efficiency of crops. While recent research aiming to increase the yield potential of crops has primarily focused on addressing specific bottlenecks in the photosynthetic process by means of genetic engineering, there is an increasing recognition of the importance of studying natural variation in photosynthetic efficiency. The Brassicaceae family contains one species, *Hirschfeldia incana*, which was shown to have a striking photosynthetic light-use efficiency. Elucidating the mechanisms behind such light-use efficiency and determining whetehr other members of the family possess similar photosynthetic characteristics to *H. incana* would contribute to our understanding of interspecific variation in photosynthesis, and its implications for plant productivity and adaptation. This, in turn, can contribute to efforts aimed at increasing the productivity of crops via more efficient photosynthesis.
Here, we present the results of our investigation into natural variation in photosynthesis within the Brassicaceae family. By combining high-throughput time-series measurements with destructive end-of-growth measurements, we examined several potential factors influencing high photosynthetic efficiency. Our study encompasses ten Brassicaceae species and provides insights into photosynthetic efficiency (dark-adapted F_v/F_m, and ΦPSII), Excess Green Index (ExGI), Normalized Difference Vegetation Index (NDVI), leaf anatomy and stomatal parameters, as well as chlorophyll content. We analyze the relationships between these species, their alignment with the phylogenetic relationships within the Brassicaceae family, and the correlations among the parameters.

In this study, we addressed a number of challenges to construct a dataset that serves as a valuable resource, describing natural variation in photosynthesis within the Brassicaceae family and providing a basis for further genomic and transcriptomic investigations. Our findings not only validate the exceptional photosynthetic efficiency of *H. incana*, but also identify *B. nigra*, *B. tournefortii*, and *Z. ait-atta* as species with high photosynthetic efficiency. Through a comprehensive analysis of twenty-one parameters, encompassing physiological and anatomical characteristics, we uncovered a correlation between ΦPSII and stomatal counts, showcasing the efficacy of combining high-throughput and end-of-growth destructive measurements for studying complex physiological traits like photosynthetic efficiency.

5.1 Background

Given the projected global population growth, the impacts of global warming, and the urgent need for sustainable food production, the agricultural sector is under pressure to increase crop yields while minimizing land use and reducing inputs such as fertilizers and pesticides. Photosynthesis can play a central role in increasing agricultural production as it determines the energy conversion efficiency of crops, a yield-related trait that has yet to see substantial improvement through conventional plant breeding methods (Zhu *et al.*, 2010). The limited progress in this area can be attributed to the physiological and genetic complexity of photosynthesis, as well as to complex interactions with environmental factors , as evidenced by inconsistent results from recent studies aimed at enhancing crop yields by manipulating photosynthetic processes across various species and cropping seasons (Kromdijk *et al.*, 2016; De Souza *et al.*, 2022; Ruiz-Vera *et al.*, 2022; López-Calcagno *et al.*, 2020, 2019; Yoon *et al.*, 2020; Garcia-Molina and Leister, 2020).

While recent research aiming to improve crop yields via photosynthesis has predominantly focused on addressing specific bottlenecks in the photosynthetic process, there is a growing recognition of the need to shift perspectives towards studying natural variation in photosynthetic light-use efficiency (LUE). Natural variation in photosynthesis has been identified as a potential avenue for improving crop productivity (Faralli and Lawson, 2020; Theeuwen *et al.*, 2022), although previous studies have primarily focused on crop species and the model plant *Arabidopsis thaliana* (Taylor *et al.*, 2022). Investigations on these species have significantly contributed to our understanding of physiological and genetic variation in photosynthesis. However, they have not elucidated the mechanisms responsible for high photosynthetic LUE at high irradiances because these species do not possess this trait. A plant is characterized by high photosynthetic LUE at high irradiance if the light saturation point for its photosynthetic reactions is high (Formighieri, 2015), meaning that its photosynthetic CO_2 assimilation rate will continue increasing until irradiance reaches very high values. While the photosynthetic apparatus of most major crops is saturated at irradiances corresponding to about one-quarter of maximum full sunlight (Long *et al.*, 2006), a much better performance has been reported for wild plant species growing in challenging environments (Werk *et al.*, 1983). Therefore, the collective effort to enhance the photosynthetic efficiency of crops would greatly benefit from a deeper understanding of natural variation in photosynthesis within non-crop and non-model species.

In our previous work we have confirmed that *Hirschfeldia incana* L. Lagr.-Foss. exhibits exceptional photosynthetic CO_2 assimilation rates, and therefore LUE, at high irradiances, and we have begun unraveling the genetic basis of this property (Garassino *et al.*, 2022). We also argued that *H. incana*, being a member of the well-studied and economically significant Brassicaceae family, provides a much more accessible way of exploring high photosynthetic LUE under high irradiance than other high-LUE species (Taylor *et al.*, 2022). The complex evolutionary history of the Brassicaceae family might have resulted in other member species evolving high photosynthetic LUE, and therefore exploring natural variation in photosynthesis across this family holds promise for understanding how high photosynthetic efficiency has evolved and

identifying species that may surpass *H. incana* in this regard.

The study of natural physiological and genetic variation in photosynthesis can be traced back over sixty years (Talling, 1961; Hodges, 1967; Avratovščuková and Fousová, 1975), and its use for the increase of crop yields was first proposed over forty years ago (Austin, 1989). However, up until recently the most common methods for measuring photosynthesis relied on gas analysis, which was inherently slow and hindered the scalability of research aimed at understanding the physiological and genetic basis of natural variation in photosynthesis (Siebers *et al.*, 2021; Du *et al.*, 2020). Nevertheless, recent advances in optical measurements of chlorophyll fluorescence have revolutionized photosynthesis phenotyping, liberating it from the constraints of low throughput. These modern methods enable the (quasi-)simultaneous measurement of photosynthetic parameters in hundreds or even thousands of plants across multiple time points, such as different times of day, and lasting over several days or weeks, or key growing season moments. The development of high-throughput phenotyping platforms creates an unprecedented opportunity to integrate chlorophyll fluorescence- and additional image-based parameters, allowing to gain detailed insights into photosynthetic performance. Multivariate analysis of high-throughput data holds significant promise to further our understanding of natural variation in photosynthesis and its implications for plant productivity and adaptation.

In this manuscript, we present a comprehensive analysis of photosynthetic efficiency at high irradiance in ten species that reflect key evolutionary events within the Brassicaceae family (Franzke *et al.*, 2011): *Arabidopsis thaliana*, *Brassica oleracea*, *Brassica nigra*, *Brassica rapa*, *Brassica tournefortii*, *Erucastrum littoreum*, *H. incana*, *Sinapis alba*, *Sisymbrium irio*, and *Zahora ait-atta*. Our study makes use of high-throughput phenotyping techniques to measure photosynthetic efficiency and integrates these measurements with other image-based parameters, such as the Excess Green Index (ExGI) and the Normalized Difference Vegetation Index (NDVI), as well as a range of anatomical and biochemical characteristics that potentially influence photosynthetic efficiency. We explore the resulting complex and multivariate dataset using various statistical methods to identify trends across species and investigate if more species within the family show high-photosynthetic LUE at high irradiance. Furthermore, we assess the alignment of these trends with the evolutionary history of the Brassicaceae family. Our study delivers a detailed description of inter-specific variation in photosynthetic parameters for the Brassicaceae family, completed by a selection of anatomical and biochemical characteristics that may play a role in supporting high photosynthetic LUE under high irradiance. The insights gained from this research will be important in developing strategies to enhance the photosynthetic LUE at high irradiance of crop species.

5.2 Results

Ten Brassicaceae species were selected for this study: *A. thaliana*, *B. oleracea*, *B. nigra*, *B. rapa*, *B. tournefortii*, *E. littoreum*, *H. incana*, *S. alba*, *S. irio*, and *Z. aitatta* (see Figure S1 for a summary of the phylogenetic relationships). To account for differences in growth speed, three faster-growing species (*B. oleracea*, *B. rapa*, *S. alba*) were sown and transferred to the phenotyping platform one week after the other seven species. Here we present the results of the various measurements we conducted during or at the end of their growth. Throughout their growth, we performed daily measurements on all plants for the maximum efficiency of photosystem II (dark-adapted F_v/F_m), Excess Green Index (ExGI), and Normalised Difference Vegetation Index (NDVI), as well as bi-daily measurements for the operating efficiency of photosystem II (ΦPSII). The parameters obtained from these measurements are derived from images, which, due to the design of our phenotyping platform, required segmentation for object identification and masking to enable appropriate quantification of the parameters. At the end of the growth period, we measured anatomical and biochemical parameters for all plants using more classical methods, such as pigment extraction in solvent, quantification by spectrophotometry, and microscope imaging. After presenting the key findings for each of the measured parameters, we show the results of a number of dimensionality-reduction and statistical analyses. These analyses were performed with the aim of highlighting relevant similarities and differences between the ten studied species.

5.2.1 Validating the robustness of custom masking approach and enhancing fluorescence data quality with dark-adapted F_v/F_m measurements

Due to the species' characteristics in our experiment, we observed that the standard masking and quantification workflow provided by the PSI FluorCam software, which relies on a fixed-area mask for measuring fluorescence parameters, did not perform well when plants exceeded the typical size of *A. thaliana*. To overcome this limitation, we developed a customized masking and quantification workflow using the PlantCV suite (Fahlgren *et al.*, 2015). Briefly, we first decoded the raw images in *.fimg* format generated by the PSI PlantScreen™ platform and imported them into Python as NumPy arrays (Harris *et al.*, 2020). Using PlantCv functions, we then processed individual images by thresholding them and identifying the plant at the center of the image, thus masking individual plants irrespective of their size. Finally, we made use of NumPy functions to calculate descriptive statistics for each processed image.

To evaluate the added value of our method over the manufacturer's standard approach, we fitted a linear regression model between the readings of photosystem II maximum quantum yield (F_v/F_m) obtained using the standard PSI software and our masking and quantification pipeline. The regression model demonstrated a strong linear trend between the two approaches (R^2 = 0.87), with the regression coefficients showing that values obtained from our custom pipeline are coherent with those ob-

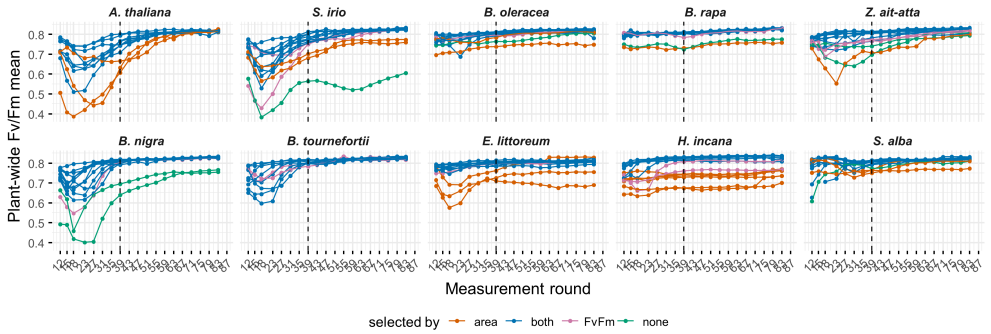

Figure 5.1: Overview of selection of plants based on measurements taken at 26 DAS_s. Each set of dots, and the line connecting them, represents a single plant's mean area measured during various days (i.e, measurement rounds). The round conducted at 26 DAS_s (round 39) is represented by the solid black line. Blue points and lines identify plants that were selected to be part of downstream analysis, orange ones identify plants selected based on area but not on F_v/F_m, pink ones those selected based on F_v/F_m but not on area, and green ones identify the plants that are excluded based on both methods.

tained from the FluorCam software (slope = 1.05, intercept = -0.04, Figure S2, Table S1). Using our own pipeline, we could compute the distribution of per-pixel F_v/F_m values over 200 equally-sized bins for each image in our dataset. This allowed us to extract the modal values of the per-image F_v/F_m distribution and compose time-series images of F_v/F_m distributions corresponding to each measured plant (File S1).

During the course of our experiment, we observed that the growth of a number of plants was stressed due to the experimental conditions, which subsequently influenced their chlorophyll fluorescence-derived parameters. To address this issue, we devised a filtering strategy to select non-stressed plants for downstream analysis of fluorescence parameters. We extracted dark-adapted F_v/F_m values measured 26 days after sowing of slow-growing species (DAS_s) (i.e., after one week of growth under high irradiance for all plants) and calculated the per-species means and standard deviations. Plants with F_v/F_m values lower than the corresponding species' mean minus 0.5 times the species-wide standard deviation were excluded from downstream analysis, except for *B. tournefortii* plants (Figure S3). Given the low standard deviation and high uniformity observed for this species, we decided not to apply the filtering criteria to this species, and to include all its plants in the subsequent analysis. In order to test whether the stress resulting in low F_v/F_m values impacted plants' growth, we related F_v/F_m values to plants' projected leaf areas. However, we could not confirm that plants filtered out based on F_v/F_m values had a consistently smaller size than plants selected for downstream analysis (Figure S4). Therefore, we complemented selection of plants based on F_v/F_m values by extracting corresponding plant areas and excluding plants with area values lower than the corresponding species' mean area minus the species' standard deviation (Figure S5).

Table 5.1: Number of plants per species retained after the F_v/F_m-based filtering and corresponding per-species F_v/F_m means (± standard deviations) at 26 DAS_s.

Species	*n*	F_v/F_m	Species	*n*	F_v/F_m
A. thaliana	8	0.758 (±0.030)	*B. nigra*	11	(0.807 ±0.008)
B. oleracea	11	0.804 (±0.006)	*B. rapa*	12	(0.807 ±0.004)
B. tournefortii	8	0.806 (±0.009)	*E. littoreum*	9	(0.799 ±0.011)
H. incana	7	0.823 (±0.007)	*S. alba*	10	(0.806 ±0.008)
S. irio	9	0.783 (±0.019)	*Z. ait-atta*	10	(0.803 ±0.017)

By applying both selection criteria, we selected 95 out of the 152 plants (62.5%) that grew during the experiment (Figure 5.1). The per-species-level mean F_v/F_m values at 26 DAS_s ranged from 0.758 to 0.823 (Table 5.1), thus falling within the optimal range of 0.75-0.83 (Krause and Weis, 1991; Hogewoning *et al.*, 2012; Maxwell and Johnson, 2000).

5.2.2 ΦPSII measurements confirm the photosynthetic efficiency of *H. incana*

Similarly to what was done for F_v/F_m, we compared the measurements of quantum yield of photosystem II (ΦPSII) obtained with the FluorCam software to ΦPSII measurements obtained with our custom masking and quantification workflow. ΦPSII measurements were conducted twice a day, in the morning and in the afternoon while under a growth irradiance of 960 µmol m^{-2} s^{-1}. The linear regression model fitted between FluorCam- and custom pipeline-derived plant-wide mean ΦPSII values showed good collinearity (R^2 = 0.87, Figure S6, Table S1). We confirmed the within-plant homogeneity of ΦPSII measures by inspecting distributions of per-pixel ΦPSII values for each image in our dataset organised in per-plant time series (File S2).

We then analysed the trends in ΦPSII over time for the 95 plants selected based on F_v/F_m. We decided to focus on measurements conducted while both the slow- and fast-growing species were in the phenotyping platform, and therefore We restricted the set of measurement rounds to those performed between 19 DAS_s and the end of the experiment (37 DAS_s). We examined trends in ΦPSII mean and modal values throughout the entire set of measurements with generalized linear mixed-effect models (GLMMs) that accounted for the block design and the repeated measurements of individual plants. We included information on measurement rounds as well in the model by calculating the hours elapsed between midnight on the day the ΦPSII measurement were started and the hour in which of each measurement round started. The calculated hours elapsed were included in the GLMMs as a second-order polynomial term. We evaluated the goodness of fit and the proportion of variance

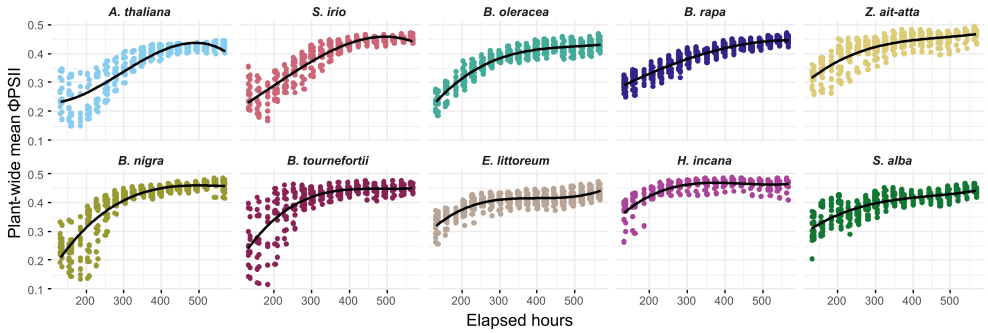

Figure 5.2: Per-species trends of plant-wide mean ΦPSII values over the course of the experiment. The black lines represent a linear model including the hours elapsed as a third-order polynomial fitted on the data to visualize trends in ΦPSII. The hours elapsed were calculated between midnight on the day the ΦPSII measurements were started and the hour in which of each measurement round started. The first points, at 130 hours elapsed, correspond to the measurement conducted in the morning at 19 DAS_s. The last points, at 567 hours elapsed, correspond to the measurement conducted in the afternoon at 37 DAS_s.

explained in the two GLMMs on means and modal values using marginal and conditional pseudo-R-squared (R_{GLMM}) measures.

The marginal R_{GLMM} measures were 0.96 for both GLMMs, conducted on plant-wide means and modal values of ΦPSII. These values ignore random effects and indicate the proportion of variance explained by the plant species and by elapsed hours, i.e. by the measurement rounds. The conditional R_{GLMM} measures, which consider both fixed and random effects, were 0.97 for both GLMMs. This demonstrates an excellent fit of the models and their ability to explain most of the variance in the ΦPSII datasets. We performed an analysis of variance (ANOVA) on the two separate models and observed a significant interaction effect between plant species and hours elapsed, confirming the presence of differences between species across the entire series of measurements (Table S5). Upon inspecting the coefficients for both models, we identified per-model significant differences in intercepts and most partial regression coefficients (Tables S6, S7). Therefore, we conclude that notable variations in trends over time could be established for both the plant-wide mean and modal values of ΦPSII.

We further employed linear mixed-effect models (LMMs) to test for differences in ΦPSII values across species during individual measurement rounds. Heritability (Visscher *et al.*, 2008), defined as the proportion of phenotypic variation in a trait that can be attributed to genetic variation among species, was used to identify the most promising rounds for downstream analysis. The heritability estimates, based on plant-wide mean ΦPSII values for the 24 selected measurement rounds, ranged between 0.27 and 0.61 (Table S2). Among the ten measurement rounds with the highest heritability, *H. incana* exhibited the highest mean ΦPSII values, with stat-

istically significant differences observed in comparison to most species, except for various combinations of *B. nigra*, *B. tournefortii*, and *Z. ait-atta* depending on the round (Figure S7, Table S3). An analogous analysis of ΦPSII modal values (i.e., the values corresponding to the most represented of the 200 bins in which the per-plant and per-round distribution of per-pixel ΦPSII values was divided) revealed a similar trend, with *H. incana* displaying high values, although the differences between *H. incana*, *B. nigra*, *B. tournefortii*, and *Z. ait-atta* were less pronounced (Figure S8, Table S4).

5.2.3 Analysis of Excess Green Index (ExGI) reveals significant variation in leaf greenness

RGB images taken between 21 and 37 DAS_s were processed by applying a mask and summarizing them using the Excess Green Index (ExGI) (Woebbecke *et al.*, 1995; Richardson *et al.*, 2007). The ExGI serves as a measure of leaf "greenness" (Richardson *et al.*, 2009) and has been linked both to carbon assimilation at a large scale in remote sensing applications (Wang *et al.*, 2020), as well as to leaf-level photosynthetic dynamics, such as the electron transport rate of photosystem II (Junker and Ensminger, 2016). Thus, the ExGI seems a useful parameter to study in the ten species included in this study.

After masking RGB images generated by the PSI PlantScreen™ platform and visualizing per-plant mean ExGI values, we observed that a technical camera or software issue caused images taken after 32 DAS_s to have slight, but consistent, variations in their color profile, resulting in a reduction of ExGI values (Figure S9). Having established that this reduction was not related to any biological effects, we adjusted the ExGI values. Specifically, we added the difference in ExGI calculated from images taken on each plant at 32 and 33 DAS_s to all subsequent ExGI values calculated from images captured from 33 DAS_s onwards. After applying the correction, data points from measurements conducted from 33 DAS_s did fit well the ExGI trends outlined by data collected up and until 32 DAS_s (Figure 5.3). Having observed some outliers that had a substantial impact on our analysis, we decided to exclude plants whose ExGI values deviated from the species' mean values by a factor greater than 0.8 times the species' standard deviations at 26 DAS_s, which corresponded to seven days after all plants were subjected to high light conditions.

To explore between-species differences in ExGI over plant growth (Figure 5.3), we fitted a GLMM identical to the one used on ΦPSII data to the ExGI values calculated for all selected plants at time points between 21 and 37 DAS_s (marginal R_{GLMM} = 0.77, conditional R_{GLMM} = 0.93). The model-estimated coefficients, as well as an ANOVA, confirmed the statistical significance of differences in intercept and partial regression coefficients (Tables S7, S8). We further investigated per-round differences in ExGI following a similar approach as employed for ΦPSII. Specifically, we utilized

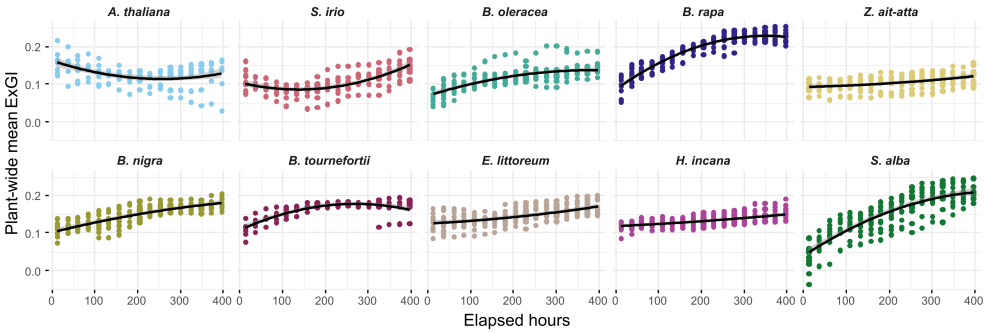

Figure 5.3: Per-species trends of plant-wide mean Excess Green Index (ExGI) values over all imaging rounds. The black lines represent a linear model including the hours elapsed as a third-order polynomial fitted on the data to visualize trends in ExGI. The hours elapsed were calculated between midnight on the day the RGB imaging started and the hour in which of each measurement round started. The first points, at 12 hours elapsed, correspond to the imaging round conducted at 21 DAS_s. The last points, at 396 hours elapsed, correspond to the imaging round conducted at 37 DAS_s.

per-round mixed linear models that accounted for the block design and examined differences in the ten rounds exhibiting the highest heritability with respect to plant species.

In these rounds, *B. rapa* and *S. alba* displayed the highest ExGI values, which were significantly different from those of most other species (Figure S10, Table S10). Based on the ExGI analysis, we categorized the species into three distinct groups. The first group consisted of “high green” species, including *B. rapa* and *S. alba*. The second group comprised “low green” species, namely *Z. ait-atta*, *S. irio*, *A. thaliana*, and *B. oleracea*. The third group represented “mid green” species, encompassing *H. incana*, *E. littoreum*, *B. nigra* and *B. tournefortii*.

5.2.4 Significant differences in potential vegetation health and photosynthetic activity highlighted by NDVI

Simultaneously with RGB imaging, we collected reflectance data in the red and near-infrared parts of the light spectrum to calculate the Normalized Difference Vegetation Index (NDVI). NDVI is used in high-throughput phenotyping to assess vegetation health and productivity potential, and it has been suggested to be linked to photosynthetic activity at a larger scale (Sellers, 1985). Similar to what observed for ExGI values, raw NDVI values exhibited substantial changes between 32 and 33 DAS_s (Figure S11). To address this, we applied the same correction method as described for ExGI, and we also removed outliers using the approach of excluding plants with NDVI values deviating from the species’ mean by more than 0.8 times the species’ standard deviation.

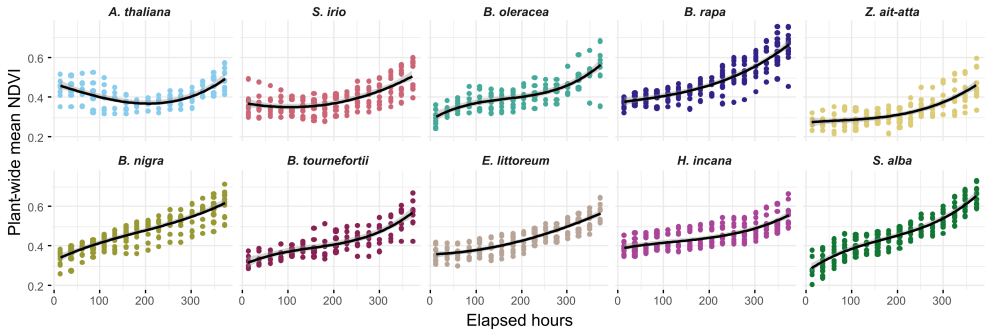

Figure 5.4: Per-species trends of plant-wide mean Normalized Difference Vegetation Index (NDVI) values. The hours elapsed were calculated between midnight on the day the NDVI imaging started and the hour in which of each measurement round started. The first points, at 12 hours elapsed, correspond to the imaging round conducted at 21 DAS_s. The last points, at 396 hours elapsed, correspond to the imaging round conducted at 37 DAS_s.

To investigate between-species differences in NDVI trends throughout plant growth (Figure 5.4), we followed the same methodology as employed for ΦPSII and ExGI analyses. The GLMM yielded a marginal R_{GLMM} of 0.82 and a conditional R_{GLMM} of 0.88. An ANOVA further confirmed the statistical significance of differences in intercept and partial regression coefficients (Tables S11, S12). Additionally, we analyzed the NDVI trends for individual measurement rounds, similarly to the approach described for ExGI. Among the ten rounds with the highest heritability, *B. nigra*, *B. rapa*, *S. alba*, *H. incana*, and *E. littoreum* stood out with the highest NDVI values, significantly differing from most other species (Figure S12). In rounds closer to the end of the experiment, the first three species exhibited even higher NDVI values, significantly differentiating them from the other two species (Table S13).

5.2.5 Variability of leaf anatomical properties

Leaf anatomy plays a key role in photosynthetic performance because it plays a major role in limiting CO_2 diffusion into and within leaves to the sites of CO_2 fixation (Terashima *et al.*, 2011; Adams III and Terashima, 2018). We examined the number and size of stomata on both the adaxial and abaxial surfaces of leaves from the ten plant species as a proxy for the conductance of CO_2 diffusion from the outside into the leaves. Utilizing a machine-learning-based approach to analyse micrographs of stomatal imprints, we estimated the number of stomata within a given leaf area and measured the lengths of the stomata major axes. For each micrograph, we calculated both the mean and median lengths of stomatal major axes. These parameters were analyzed with LMMs. The models revealed a better fit for the measurements conducted on the abaxial side of leaves compared to those on the adaxial side (conditional R_{LMM} of 0.64 and 0.45, respectively) when assessing the number of stomata.

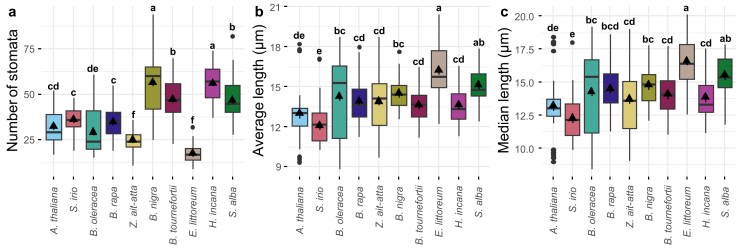

Figure 5.5: Per-species boxplots of parameters measured for stomata on the abaxial side of leaves. (a) Average number of stomata counted per-species, based on counts obtained from all images generated from samples for that species, each measuring 0.09 mm^2; (b) the average length of the major axis of stomata measured per-species, based on counts obtained from all images generated from samples for that species; (c) and the median length of the major axis of stomata measured per-species, based on counts obtained from all images generated from samples for that species. The "▲" symbols represent Estimated Marginal Means (EMMs) derived from LMMs accounting for block design, and their associated error bars represent the standard error of EMMs. Within each plot, letters present a Compact Letter Display (Gramm *et al.*, 2007) of significant differences between EMMs. If two species share a letter, differences between the estimates cannot be proven to be statistically significant.

H. incana and *B. nigra* exhibited the highest number of stomata on the abaxial side (Figure 5.5a) and, together with *B. tournefortii*, on the adaxial side (Figure S13a).

In terms of stomatal size, models for both the mean and median lengths did not provide a good fit for the measurements (Table S14). No differences in the trends between mean or median lengths could be identified. *S. alba* and *E. littoreum* displayed the highest major axis lengths on both the abaxial (Figure 5.5b,c) and adaxial surfaces of leaves (Figure S13b,c). The combination of larger stomatal dimensions observed in *E. littoreum* alongside its relatively low stomatal count on both leaf sides suggests that this species has fewer but larger stomata. Conversely, *S. alba* appears to have a relatively high number of stomata with larger dimensions, similar to the observations for *H. incana* and *B. nigra*.

Having assessed the potential for gas diffusion into the leaves of the ten species, we then investigated the limitations to CO_2 diffusion within leaves by performing measurements on cross-sectional micrographs of leaves. First, we focused on the square roots of areas of the palisade and spongy mesophyll as a proxy for cell volume. We transformed the area values by calculating their square roots to meet the assumptions of LLMs we employed in subsequent analysis. Our LMM analysis revealed significant differences in both palisade and spongy mesophyll cell areas (conditional R_{LMM} of 0.52 and 0.56, respectively). Among the ten species measured, *E. littoreum* exhibited the largest areas for both palisade and spongy mesophyll, while *B. nigra* and *S. irio* had the smallest areas (Table S15). Notably, *H. incana* displayed large areas for palisade mesophyll but small areas for spongy mesophyll (Figure

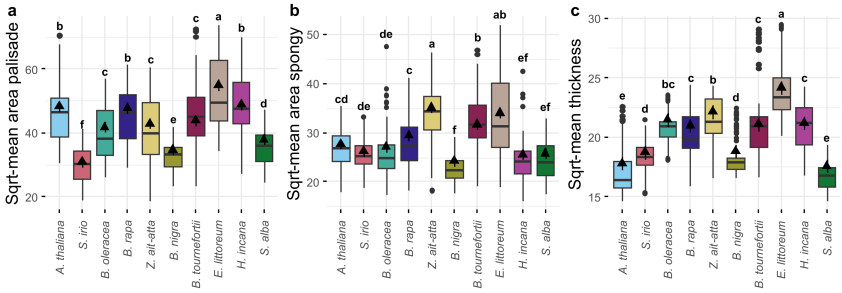

Figure 5.6: Per-species boxplots of measured anatomical parameters. (a) The square root of per-sample average area of palisade mesophyll cells (in pixels, px); (b) the square root of per-sample average area of spongy mesophyll cells (px); (c) and the square root of per-sample average leaf thickness (µm). The "▲" symbols represent Estimated Marginal Means (EMMs) derived from LMMs accounting for block design, and their associated error bars represent the standard error of EMMs. Letters present a Compact Letter Display of significant differences between EMMs.

5.6a,b).

To gain further insights into the impact of different cell volumes on overall leaf structure, we measured total leaf thickness from the micrographs. The LMM fitted on square roots of thickness measures resulted in a conditional R_{LMM} of 0.77. Based on this model, we found significant differences between species (Figure 5.6c). Once again, *E. littoreum* stood out with the highest values for leaf thickness, while *A. thaliana* and *S. alba* had the thinnest leaves (Table S15). Despite having a large area of palisade mesophyll cells, the leaf thickness of *H. incana* fell within the central range of the thickness distribution.

5.2.6 Moderate differences in leaf chlorophyll content

At the end of our experiment, we sampled the oldest leaf of each plant, which due to the short growing period was the larger and most developed as well, extracted the leaf chlorophylls using N,N-dimethylformamide (DMF), and quantified the concentration of chlorophylls *a* and *b* in the resulting extracts (Figure S14). By measuring the leaf thickness through microscopy on samples taken from the same leaves, we were able to estimate the volume of leaf tissue from which the extracted chlorophylls originated, which enabled us to express chlorophyll concentration in relation to leaf volume. The measurement of chlorophyll *a* and chlorophyll *b* content (Figures 5.7a,b) provides information on the abundance of these pigments, which are vital for capturing light energy during photosynthesis. Additionally, the chlorophyll *a*/*b* ratio (Figure 5.7c) offers insights into the relative distribution of these two chlorophyll types, highlighting potential variations in light-harvesting strategies among species. Further-

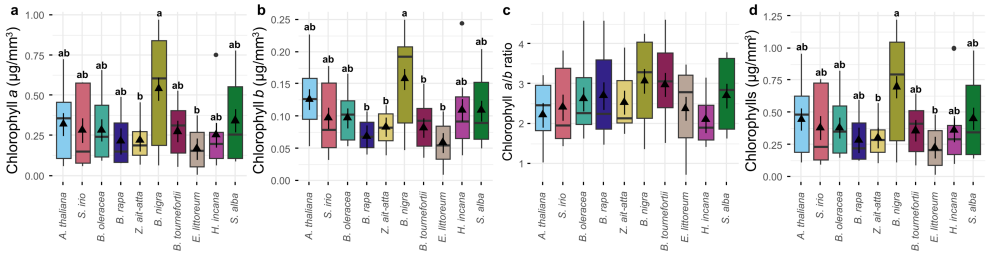

Figure 5.7: Per-species boxplots of parameters relating to chloropylls. (a) The per-sample chlorophyll *a* concentration, in µg per mm^3 of leaf volume; (b) the per-sample chlorophyll *b* concentration (µg /mm^3 leaf volume); (c) the per-sample chlorophyll *a*/chlorophyll *b* ratio, (d) and the per-sample total chlorophyll concentration (µg chlorophyll/mm^3 leaf volume). The "▲" symbols represent Estimated Marginal Means (EMMs) derived from LMMs accounting for block design, and their associated error bars represent the standard error of EMMs. Letters present a Compact Letter Display of significant differences between EMMs where any differences could be proven significant.

more, by examining the total chlorophyll content in the samples (Figure 5.7d), we gain an overall measure of the photosynthetic pigment concentration, which relates, albeit non-linearly, to the potential capacity for light absorption within the leaf.

Among the species investigated, *B. rapa*, *Z. ait-atta*, and *E. littoreum* exhibited the lowest levels of both chlorophyll *a* and *b*. The chlorophyll composition of *H. incana* did not differ significantly from these species, although *H. incana* displayed a low chlorophyll *a*/*b* ratio, indicating improved tolerance to high irradiance via reduced size of light-harvesting complexes. In contrast, *B. nigra* showed the highest chlorophyll contents and *a*/*b* ratio, which significantly differed from those of other species. These findings shed light on the inter-species variability in chlorophyll composition, reflecting differences in light-harvesting and high irradiance tolerance within the Brassicaceae family.

5.2.7 Integrative analysis of all measurements reveals little correlation between parameters

The extensive range of measurements conducted in this study resulted in a multivariate dataset. To simplify the dataset, we applied a filtering approach to focus on specific time-series parameters (ΦPSII, NDVI, ExGI) that corresponded to measurement rounds with highest heritability or were near the end of the experiment, while avoiding excessive noise due to overlapping of leaves from different plants. This allowed us to include two parameters (mean from round with highest heritability, mean from "late" round) for ExGI and NDVI each, and four parameters for ΦPSII (mean parameters as already described, modal values for the round with highest heritability

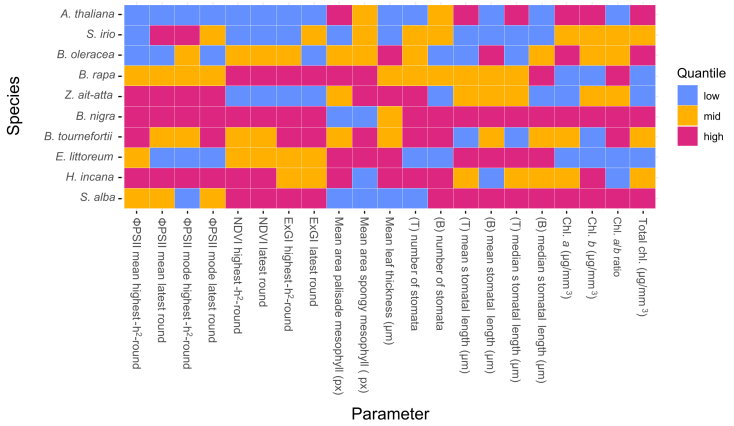

Figure 5.8: Heatmap summarising values for all parameters measured in this study on the ten Brassicaceae species. Each cell represents where the value of a given parameter for a given species falls in the distribution of all values for that parameter: "low" if the value is lower than 33rd quantile, "mid" if it is comprised between the 33rd and 66th quantile, "high" if it is higher than the 66th quantile. Parameters that were measured on both top and bottom sides of leaves are labelled with (T) and (B), respectively.

and for the "late" round) in our downstream analysis. By integrating these selected parameters with other measured variables, we obtained a total of 21 parameters for further analysis (Figure 5.8).

To reduce the dimensionality of the dataset and identify the key parameters driving the main differences among the ten studied species, we employed Principal Component Analysis (PCA). Since some measurements were missing for certain experimental units (i.e. plants, see Figure S15), we employed a regularized iterative Factor Analysis of Mixed Data (FAMD) algorithm to impute the missing values. Imputed values fell within the distribution of measured values and were therefore considered acceptable for downstream analysis (Figure S16). Examining the PCA eigenvalues, we found that the first three principal components accounted for a substantial portion of the data variation, nearly 70% (PC1=26.2%, PC2=25.0%, PC3=17.3%, Figure S17, S18). Consequently, we arranged the parameters based on their contributions to these components. Notably, four out of the five parameters with the highest contributions were related to ΦPSII, while the parameters with the lowest contributions were the mean areas of spongy and palisade mesophyll, median and average stomatal lengths on the adaxial side of leaves, and the chlorophyll *a*/*b* ratio (Table S16, Figure S19). Species-level differences in grouping of individuals emerged as expected from the PCA, albeit that visualisation of PCA results suffered from overlap of points and groups (Figures S20, S21).

We further investigated the distribution of individual plants based on the measured parameters using t-distributed stochastic neighbor embedding (t-SNE), a nonlinear dimensionality-reduction method (Van der Maaten and Hinton, 2008). The t-SNE

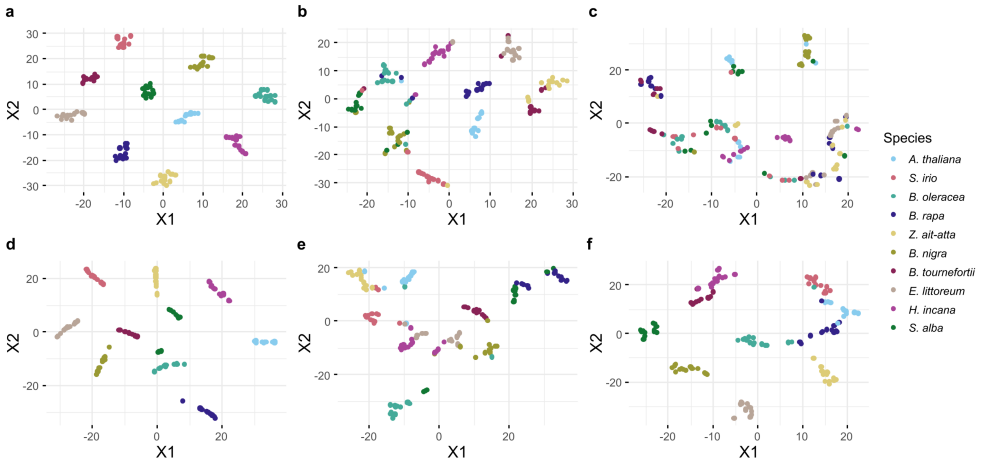

Figure 5.9: Visualization of the results of t-distributed stochastic neighbor embedding (t-SNE). t-SNE was run on **(a)** the complete dataset of 137 individuals and 21 parameters, **(b)** anatomical parameters, **(c)** chlorophyll parameters, **(d)** ΦPSII parameters, **(e)** NDVI and ExGI parameters, and **(f)** stomatal parameters.

plot generated over all measured parameters revealed the presence of 10 distinct groups, each predominantly consisting of individuals from a single species (Figure 5.9a). To determine which sets of parameters primarily contributed to the differences between species, we performed additional t-SNE analyses on specific parameter categories, including anatomical parameters, chlorophyll content, fluorescence parameters, RGB and reflectance parameters, and stomatal parameters (Figure 5.9b-f). Consistent with the contributions observed in the PCA results, we found that the fluorescence parameters determined the most pronounced separation between species. Stomatal parameters emerged as another determinant of good separation of species, while the remaining parameters resulted in weaker separation.

To statistically test the dissimilarities in sample groupings, we conducted a permutational analysis of variance (PERMANOVA) (Van der Maaten and Hinton, 2008) on the same distance matrix employed for PCA and t-SNE. The results demonstrated statistically significant differences between the centroids of groups corresponding to the various species, with a R^2 for the species factor of 0.91 (Table S17). The analysis of multivariate homogeneity of group variances (PERMDISP) we conducted as well, led us to conclude that group variances were homogeneous. We then proceeded to perform pairwise-PERMANOVA tests to highlight significant differences between individual species. Pairwise-PERMANOVA R^2 measures varied between 0.34 and 0.90, with median and mean R^2 of 0.71 and 0.67, respectively (Table S18). All pairwise differences between species resulted significant ($p<0.001$).

We further explored the relationships and similarities among the ten studied spe-

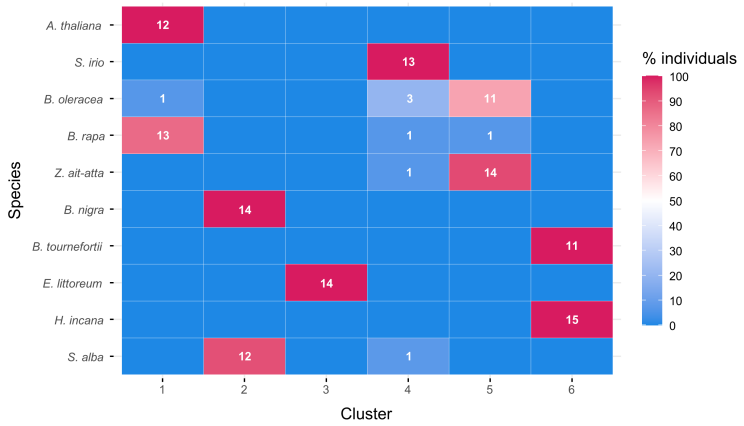

Figure 5.10: Visualization of the results of k-means clustering performed on all measured parameters for 137 individual plants. The color scale of tiles represents the percentage of the total number of individual plants for a given species that was grouped in a single cluster. The numbers in the tiles represent the number of individuals for a given species grouped in a single cluster.

cies based on the comprehensive dataset of twenty-one parameters with k-means clustering. After establishing that six clusters would be optimal (Figure S22), we grouped the species based on their similarities across all the measured parameters (Figure 5.10). Individuals from each species were mostly grouped in one of the six clusters, and three of the six clusters contained a high number of individuals from couples of phylogenetically close species (Arias and Pires, 2012; Koch and Lemmel, 2019), such as *H. incana* and *B. tournefortii*, *B. nigra* and *S. alba*, and *B. oleracea* and *Z. ait-atta*. *S. irio* and *E. littoreum* individuals clustered independently from those of any other species, and despite not being phylogenetically close, individuals of *A. thaliana* and *B. rapa* mostly clustered together.

Having observed this, and considering that the PCA reported ΦPSII-related parameters as the major contributors to the between-species separation, we analyzed the correlation between measurements for all twenty-one parameters to understand whether any other parameters could be related to ΦPSII. First, we computed and visualized the Pearson correlation coefficients (r) between all parameters along with their statistical significance (Figure S23). We then calculated a distance metric based on the calculated coefficients ($1 - r$) and used it to perform hierarchical clustering of all parameters (Figure S24). This analysis revealed a strong positive correlation between the four ΦPSII parameters and the stomatal counts on the adaxial leaf side ($0.51 < r < 0.82$, $p < 4e^{-4}$), as well as a positive correlation with the stomatal counts on the abaxial leaf side ($0.30 < r < 0.60$, $p < 2e^{-10}$). Furthermore, weaker yet significant negative correlations were observed between some of the ΦPSII parameters and the lengths of stomatal major axes on the abaxial leaf sides ($-0.22 < r < -0.16$, p

$<$0.03), and a weak correlation pattern emerged between ΦPSII parameters and the mean area of palisade mesophyll cells (-0.22 $<r<$0.19, $p<$0.03). Thus, ΦPSII, the proxy for photosynthetic LUE employed in this study, was observed to be correlated mostly to stomatal parameters, while no significant correlations emerged between ΦPSII and the other measured parameters.

5.3 Discussion

In this study we assessed the photosynthetic efficiency of ten Brassicaceae species using high-throughput image-based phenotyping approaches. Our analysis encompassed inter-specific variation in photosynthetic parameters and a selection of potential determinants, measured with both high- and low-throughput methods. The resulting multivariate dataset allowed us to compare species, identify similarities and differences, and evaluate the contribution of the various measured parameters to inter-species distinctions.

5.3.1 Challenges associated with testing the panel of Brassicaceae species

The experiments presented in this manuscript exemplify the challenges associated with high-throughput phenotyping (HTP) of diverse sets of plant species. We believe these challenges have greatly limited studies on inter-specific variation in controlled environments. The phenotyping platform used in this study allowed us to grow the selected species for only about four weeks (or three weeks for faster-growing species) due to leaf overlap and potential shade-avoidance responses. Consequently, our analysis captures the early growth phases of all species except *A. thaliana*, with plants still in the exponential growth phase (Figure S25) at the end of the experiment. We assessed the distribution of HTP parameters (such as ΦPSII, NDVI, and ExGI) and confirmed their uniformity, thus confirming the absence of large differences in these parameters between leaves of our plants. Additionally, we sampled the most developed leaf at the end of the experiment to minimize developmental differences while measuring low-throughput parameters (anatomical parameters and chlorophyll content). Therefore, while our conclusions may not be applicable to full-grown plants, we are confident in their robustness for the young plants we were able to grow and measure.

Our study highlights the need for caution when relying on “turnkey” image segmentation and analysis software associated with phenotyping platforms. During our experiment, we noticed rapid invasion of the fixed imaging areas employed by the standard masking method by leaves from faster-growing plants. This prompted us to develop and evaluate a custom masking strategy based on object identification. Beyond allowing us to explore the HTP datasets in more depth (e.g., by extracting per-pixel parameter values for all images and exploring their distribution), our approach yielded small but important improvements in the estimation of HTP parameters.

In line with our focus on photosynthetic light-use efficiency under high-irradiance conditions, we conducted the experiment at the highest irradiance levels allowed by our phenotyping platform, which was an average irradiance of 960 µmol m^{-2} s^{-1} of light at plant level for twelve hours a day, including dawn and dusk periods. This corresponded to a daily light integral (DLI) (Korczynski *et al.*, 2002) of 34.6 mol m^{-2} d^{-1} (Table S19), comparable to DLIs reported during late spring and early fall in warm-temperate global regions (Korczynski *et al.*, 2002; ENEA TER-SOLTERM, 2006). Although past studies considered irradiances well below 1000 µmol m^{-2} s^{-1} as "high light" in controlled-environment plant research (Balfagón *et al.*, 2019; Pech *et al.*, 2022; van Rooijen *et al.*, 2018), irradiances higher than 1000 µmol m^{-2} s^{-1} are widely experienced by plants growing outdoor. Considering that the average irradiance employed in this study is very close to 1000 µmol m^{-2} s^{-1}, we classify it as high irradiance as well.

Not anticipating growth issues, based on prior experience with significantly higher irradiances, we unexpectedly observed stress in a subset of plants during the experiment, as evidenced by the removal of several plants from the analysis of fluorescence parameters based on maximum quantum yield of Photosystem II (PSII), or dark-adapted F_v/F_m. We attribute this stress to difficulties in establishing seedlings within the phenotyping platform as a result of the heat they might have experienced due to a combination of the high irradiance and the design of the phenotyping platform. Seedlings had indeed to be transplanted directly into large rockwool blocks covered by a black plastic screen which heated up when exposed to the high irradiance generated by the LED modules. We believe that due to potentially uneven air circulation in the growth room, some seedlings could have then been exposed to high temperatures which could have compromised their establisment, and the following plant growth.

5.3.2 Pushing the boundaries of high-throughput measurements

High-throughput measurements enabled us to investigate photosynthesis, growth, plant pigmentation, and reflectance over a substantial part of the early stages of growth of these plants. The parametrization of time-series high-throughput measurements is a promising approach (Flood *et al.*, 2016) that has been employed in only a limited number of studies to date (van Eeuwijk *et al.*, 2019; Tardieu *et al.*, 2023; Pérez-Valencia *et al.*, 2022). Regarding ΦPSII, the prevailing analysis approach considers individual measurement rounds as separate entities (van Rooijen *et al.*, 2015, 2017; Flood *et al.*, 2020). In our study, we complemented the analysis at the single measurement-round level for ΦPSII, NDVI, and ExGI with an examination of the per-species trends over time for these parameters. We parameterized these trends by fitting generalised mixed-effect linear models on the high-throughput parameters' datasets (ΦPSII, ExGI, NDVI), and evaluating the statistical significance of differences between per-species model coefficients. With this approach, we revealed significant differences between the coefficients, and thus in the trends over time, for all parameters among the ten studied species, thus gaining insights into the

distribution and differences of the parameters across the entire time series.

In our study, we utilized our image analysis workflow to extract various statistics summarizing parameter distributions across single plant images. To evaluate the performance of mean and mode in describing plant performance, we conducted a case study using ΦPSII. The mean, as a summary statistic, is susceptible to outliers in the dataset. In the context of ΦPSII phenotyping, low values at the borders of leaves or other organs, such as stems, can act as outliers and significantly impact the plant-wide mean ΦPSII, despite their limited contribution to the overall plant area (Figure S26a). Additionally, the overlapping of leaves from different plants can lead to artificially low mean ΦPSII values for target plants (Figure S26c). To address these issues, we analyzed time-series images of ΦPSII and compared the plant-wide means and modes with the corresponding distribution of per-pixel ΦPSII values. Our findings indicated that the mode of ΦPSII is robust against the aforementioned disturbances (Figure S26b,d). Consequently, we examined the calculated heritabilities for plant-wide ΦPSII means and modes (Tables S2, S3), observing that the rounds with the highest heritability occurred around half-way the experiment for ΦPSII means, while for ΦPSII modes, the highest heritabilities were associated with the latest measuring rounds. The results of this case study imply that the modal values for image-wide parameter distributions can improve the reliability of parameter summary statistics calculated for individual plants.

In this study, we expanded the set of measured parameters by collecting RGB and near-infrared (NIR) images. In controlled-environment high-throughput (HTP) platforms, RGB imagery is commonly used to measure plant area parameters and morphological parameters through side-view cameras, if available (Yang *et al.*, 2020). However, in platforms which only offer a top-view perspective of plants like the one utilized in our study, RGB images have primarily been employed for calculating plant areas and growth (Ge *et al.*, 2016; Mazis *et al.*, 2020; Jiang *et al.*, 2022). The extensive use of RGB imagery in open-field HTP and remote sensing has led to the development of numerous vegetation indices that could be applied to images generated by controlled-environment HTP platforms (Adak *et al.*, 2021; Rufo *et al.*, 2021; De Swaef *et al.*, 2021). In our study, we conducted an analysis of RGB images to calculate the Excess Green Index (ExGI), which is one of the various measures of plant "greenness" (Richardson *et al.*, 2009). Notably, this index has been associated with carbon assimilation on a large scale in remote sensing applications (Wang *et al.*, 2020) and leaf-level photosynthetic dynamics, such as the electron transport rate of photosystem II (Junker and Ensminger, 2016). However, experimental evidence supporting these associations remains limited to date. In the context of controlled-environment HTP, our aim was to utilize ExGI for a different purpose, namely characterizing plant morphology based on top-view RGB images. The underlying assumption is that leaves, as the photosynthetic organs of plants, will accumulate more chlorophyll than other organs such as stems or leaf petioles. Consequently, in two plants covering the same area but displaying different morphologies, the plant with a greater amount of leaf tissue would exhibit a higher ExGI. However, this assumption holds true only if the accumulation of chlorophylls and other leaf pigments (e.g., anthocyanins and carotenoids) is consistent across different species. Upon inspecting the RGB images

from our experiment (Figure s27), we discovered that differences in greenness were significantly more pronounced than variations in morphology, rendering the latter interpretation incorrect.

5.3.3 Integration and correlation of results

Our study was designed with the ambitious goal of combining time-series high-throughput measurements with destructive end-of-growth ones to explore relationships and potential differences in photosynthetic efficiency. We confirmed the higher photosynthetic performance of *H. incana* compared to most species in our panel of Brassicaceae, which expands upon our previous findings based on low-throughput measurements of photosynthetic carbon assimilation (Garassino *et al.*, 2022). Moreover, the emergence of ΦPSII parameters as major determinants of species differentiation indicates a promising level of variation for photosynthesis within the Brassicaceae family. We found that *B. nigra*, *B. tournefortii*, and *Z. ait-atta* achieved similar levels of photosynthetic performance as *H. incana*, positioning them as prime candidates for further studies on their photosynthetic metabolism. However, we also observed striking differences in the parameters describing the time-series of ΦPSII measurements for these species, with *H. incana* reaching values close to its maximum ΦPSII during the early stages of growth. Therefore, our study reveals a significant degree of variation in photosynthetic efficiency within the Brassicaceae family, warranting further exploration and characterization.

We report strong positive correlations between ΦPSII parameters and stomatal counts, particularly those derived from the adaxial side of leaves. Recent studies have demonstrated the significant contribution of adaxial stomata to overall photosynthetic carbon assimilation in a wide range of species (Xiong and Flexas, 2020; Wall *et al.*, 2022). Therefore, the higher stomatal density on the abaxial side of leaves, and consequently the potential for increased abaxial stomatal conductance CO_2, can be linked to enhanced photosynthetic LUE at high irradiance, as reflected by ΦPSII. We did not observe strong or significant correlations between ΦPSII and our selected leaf anatomical parameters, indicating that these parameters may not serve as reliable proxies for mesophyll conductance to CO_2, which is another important determinant of photosynthetic efficiency alongside stomatal conductance (Mizokami *et al.*, 2022).

Despite some suggestions that NDVI would correlate to photosynthetic activity (Gamon *et al.*, 1995; Sellers, 1985), we were unable to confirm this correlation in our experiment. The only substantial correlation we identified for NDVI was with ExGI, with the opposite being true as well. Weaker correlations could be identified between both NDVI and ExGI and stomatal counts on the abaxial side of leaves, sizes of stomatal major axes on both the adaxial and abaxial side of leaves, and the chlorophyll *a*/*b* ratios. Explaining the relationships between NDVI or ExGI and all the correlated parameters is not trivial. The impact of the chlorophyll *a*/*b* ratio on leaf reflectance has not been clarified yet (Nyongesah *et al.*, 2015), and while a positive relationship has been established between NDVI and plant water content

(Zhang and Zhou, 2019; Zhou *et al.*, 2022), no obvious correlation between plant water content and physical characteristics of stomata has been found to date. The positive correlation between NDVI - and thus leaf water content - and stomatal characteristics might appear counterintuitive, as one would expect e.g. a lower stomatal count to be associated with higher leaf water content as a result of lower transpiration potential. However, this interpretation is an oversimplification, as it disregards stomatal regulation, which could not be measured in this study. Similarly, correlation between ExGI and stomatal physical characteristics has not been reported to date. Thus, a combined approach focusing on NDVI, ExGI, and stomatal density and conductance measurements would be an interesting target for future studies aiming at understanding the physiological basis of the correlation between vegetation indices and stomatal parameters.

5.3.4 Future directions

Naturally, our dataset could be expanded by incorporating additional measurements, with the most promising ones likely to come from hyperspectral imaging. In recent years, a growing body of evidence has accumulated regarding the possibility of directly modeling key photosynthetic parameters related to carbon assimilation and photosynthesis biochemistry, as well as parameters associated with plant constituents such as nitrogen or chlorophylls, based on hyperspectral readings (Meacham-Hensold *et al.*, 2020; Jin *et al.*, 2022; Buchaillot *et al.*, 2022). However, some caution is warranted regarding the applicability of methods developed thus far across different plant species and conditions (Khan *et al.*, 2021; Song *et al.*, 2022). Nonetheless, integrating hyperspectral imaging-based methods into future explorations, following a similar approach to our study, would undoubtedly contribute to a deeper understanding of physiological variation for photosynthetic and associated traits across the Brassicaceae family.

Furthermore, the purely physiological approach we employed in this study to explore variation in photosynthetic efficiency and potential associated parameters could be complemented by a genomic and transcriptomic dissection of such variation. In previous work we have shown how comparative genomics and transcriptomics, even when uncoupled from physiological investigations, can provide relevant insights into the genetic determinants of the high photosynthetic efficiency of *H. incana* (Garassino *et al.*, 2022, 2023). Moreover, recent comparative genomics approaches have elucidated the evolution of C_4 photosynthesis in the sister family of Brassicaceae, Cleomaceae (Hoang *et al.*, 2023), while comparative transcriptomics has been providing clues on the mechanisms responsible for photosynthetic efficiency for over ten years (Bräutigam *et al.*, 2011; Serba *et al.*, 2016). Given the availability of genome sequences for most of the species included in our study (reviewed for most species in (Garassino *et al.*, 2023), complemented by (Liu *et al.*, 2014)), the current accessibility of whole-genome and whole-transcriptome sequencing technologies, and the development of powerful tools enabling the construction of plant pangenomes (Petereit *et al.*, 2022; Li *et al.*, 2022; Shi *et al.*, 2023), comparat-

ive genomics and transcriptomics offer promising avenues for better elucidating the basis of high photosynthetic efficiency in the Brassicaceae family.

5.4 Conclusions

In this manuscript, we present an investigation of inter-species variation in photosynthetic efficiency and potentially associated anatomical and physiological parameters within the Brassicaceae family. Working with a diverse set of species posed several technical challenges that we addressed by implementing custom image-analysis pipelines and expanding the scope of our exploration to include time series of parameters instead of single measurements. The constructed dataset can serve as an initial resource describing natural variation in photosynthetic LUE across the Brassicaceae family and providing a foundation for genomic and transcriptomic dissection. Our findings confirm the remarkable photosynthetic efficiency of *H. incana* and identify *B. nigra*, *B. tournefortii*, and *Z. ait-atta* as additional species with high photosynthetic efficiency. Through our comprehensive analysis of twenty-one parameters encompassing six classes of physiological and anatomical characteristics, we have uncovered an unexpected correlation between ΦPSII parameters and stomatal counts, highlighting the power of combining high-throughput and end-of-growth destructive measurements to explore complex physiological traits, such as photosynthetic efficiency.

5.5 Materials and Methods

Plant material, experimental design, and growth conditions

A set of ten Brassicaceae species was used in this research: *Arabidopsis thaliana*, *Brassica nigra*, *Brassica oleracea*, *Brassica rapa*, *Brassica tournefortii*, *Erucastrum littoreum*, *Hirschfeldia incana*, *Sinapis alba*, *Sisymbrium irio*, and *Zahora ait-atta*. Information on individual accessions can be found in Table S20. Prior to germination, all seeds were surface-sterilized with chlorine gas for four hours. All seeds were sown on a custom sterile medium (8 g/L Daishin Agar in a 10 mM KNO_3 solution) and stratified at 4 °C for 96 h. Seeds were germinated for seven days at 24 °C with a photoperiod of 16 h under an irradiance of 150 µmol m^{-2} s^{-1} provided by fluorescent tubes. Given differences in growth rates between the various species, the seeds of seven of the species that were found in a pilot experiment to grow relatively slowly (*A. thaliana*, *B. nigra*, *B. tournefortii*, *E. littoreum*, *H. incana*, *S. irio*, *Z. ait-atta*) were sown, stratified, germinated and transferred to the phenotyping platform one week earlier than the seeds of the three other (*B. rapa*, *B. oleracea*, *S. alba*) (Table 5.2).

At the end of the germination period, seedlings were transferred to wet rockwool blocks (10x10x6.5 cm, Grodan by Rockwool B.V., Roermond, The Netherlands) placed in a flooding table part of the phenotyping platform. Fifteen seedlings per species were transplanted, giving a total of 150 plants, following a randomized complete

Table 5.2: Times at which the actions preceding growth in the final conditions were executed. All times are expressed as DASs, i.e. days after sowing of seeds of the slower growing species.

	Slower growing species	Faster growing species
Sowing and stratification	0 DASs	7 DASs
Germination	4 DASs	11 DASs
Transplant	11 DASs	18 DASs

block design (RCBD) with 15 blocks spanning a 7 m^2 growth area. The growth chamber housing the plants was equipped with high-output LED light modules (VYPR2x, Fluence, Austin, USA) delivering on average photosynthetically active irradiance of 960 µmol m^{-2} s^{-1} at plant level. Plants were grown with a photoperiod of 12 h including a two-hours dawn and a two-hours dusk period, day and night air temperature controlled to 20 °C (temperature at plant level during the day would be increased by the light modules), relative humidity controlled to 75%, and CO_2 levels were ambient. Plants were fertigated by flooding the growth table every 48 h with a custom nitrogen- and iron-rich solution (1.2 mM NH_4^+, 7.2 mM K^+, 4 mM Ca^{2+}, 1.82 mM Mg^{2+}, 12.4 mM NO_3^-, 3.32 mM SO_4^{2-}, 1.0 mM P, 35 µM Fe^{3+}, 8 µM Mn^{2+}, 5 µM Zn^{2+}, 20 µM B, 0.5 µM Cu^{2+}, 0.5 µM Mo^{4+}).

High-throughput phenotyping

High-throughput phenotyping of plants was conducted with a PlantScreen™ XY System (Photon Systems Instruments, Drásov, Czech Republic). Each day, every plant was subject to four imaging rounds: one dark-adapted F_v/F_m round (measurement starting at 04:00, running for 240 min), two ΦPSII rounds (measurements starting at 10:30 and 15:00, running for 80 min), and one combined RGB and NIR round (measurement starting at 12:00, running for 135 min). Imaging started three days after transplanting of seedlings of the slow-growing species, therefore 14 days after their sowing (DAS_s), and ran until 38 DAS_s. The order in which plants were measured was designed to maximize the time between measurements of neighboring plants, avoiding interference from the imaging processes (e.g., saturating light pulses shading by the camera and LED head).

Analysis of fluorescence images

Raw images corresponding to individual measurements needed to calculate fluorescence parameters F_v/F_m and ΦPSII (baseline and maximal fluorescence, defined as Fo and Fm, Fp and Fmp, respectively) were processed by means of custom Python v3.10.10 scripts making use of the modules NumPy v1.22.3, OpenCV v4.7.0,

PlantCV v3.14.3. Briefly, raw images in .fimg format generated by and exported from the PlantScreen™ platform were imported, decoded, and converted into NumPy arrays; images corresponding to maximal fluorescence readings (Fm and Fmp) were normalized and thresholded to convert them into binary images. These binary images were then subjected to object identification and selection of the object at the image center, resulting in a binary mask for such object; the resulting binary mask was applied to both maximal and baseline fluorescence images, and pixel-by-pixel fluorescence parameters were calculated as follows:

$$F_v/F_m = (F_m - F_o)/F_m$$
$$\Phi PSII = (F_{mp} - F_p)/F_{mp}$$

The mean, standard deviation, median, maximum and minimum values for the calculated parameters were then computed per plant with standard functions included in the NumPy module. The mode of the parameters was determined as the most-represented value in a histogram of per-plant parameter values with 200 equally sized bins. Calculated per-pixel values were then employed to reconstruct and output images of whole plants, as well as histograms of fluorescence values and summary statistics. Calculated per-plant F_v/F_m and ΦPSII values, as well as areas derived from reconstructed images, were correlated with measurements provided by the PSI FluorCam software. To reduce variation in the dataset, F_v/F_m values recorded at 39 DAS_s (one week of growth under high irradiance for all plants) were measured and plants having an F_v/F_m value lower than 0.5 times the standard deviation calculated over all replicates of their species were removed. Of the 152 plants that grew during the experiment, 107 were selected for downstream analysis.

Analysis of RGB images

RGB images in .png format, corrected for fisheye distortion by the PlantScreen software, were thresholded to binary images and subjected to object detection in the same way as done for fluorescence images. The resulting masks were applied to obtain masked images and calculate plant areas. Masked images were subsequently employed to calculate the Excess Green Index (ExGI) (Woebbecke *et al.*, 1995; Richardson *et al.*, 2007). The following calculation was performed for each pixel of each image:

$$r = R/(R + G + B)$$
$$g = G/(R + G + B)$$
$$b = B/(R + G + B)$$
$$ExGI = (2 * g) - r - b$$

Where R, G, and B are the intensities for the red, green and blue channels of a given pixel, respectively. Mean, standard deviation and mode of ExGI values corresponding to each plant were calculated with built-in functions, while the mode of ExGI values was calculated with the same approach as described above.

Since values measured and therefore calculated for imaging rounds conducted from 33 DAS_s appeared to be consistently reduced due to an unidentified technical issue, the measured values were corrected.The per-plant difference in ExGI calculated from images taken at 32 and 33 DAS_s was added back to all ExGI values calculated from pictures taken from 33 DAS_s.

Analysis of reflectance images

Images corresponding to reflectances at 660 and 740 nm in .fimg format were decoded as done for chlorohpyll fluorescence images. The Normalized Difference Vegetation Index (NDVI) was calculated per-pixel with the formula (Gamon *et al.*, 1995):

$$NDVI = (R_{740} - R_{660})/(R_{740} + R_{660})$$

Reconstructed NDVI images were then thresholded and masked as described for chlorophyll fluorescence parameters. Since a similar issue to what described for RGB images was detected for NDVI images as well, a similar correction of images generated during later imaging rounds was performed. Descriptive statistics were calculated as described for RGB statistics, and differences between per-plant mean NDVI values were tested following the same approach employed for fluorescence parameters.

Extraction and quantification of leaf chlorophylls

At the end of the growth period (38 DAS_s), two leaf discs (diameter 2 mm) were collected from the most developed leaf of each plant and immediately transferred to 1 mL N,N-dimethylformamide (DMF). Following incubation overnight at room temperature in the dark, absorbances at 647 and 664 nm of each extract were measured with a Cary 4000 UV-Vis spectrophotometer (Agilent Technologies, Santa Clara, U.S.A.). The concentrations of chlorophyll *a* and *b* in measured solutions, total chlorophyll content, and chlorophyll *a*/*b* ratios were calculated according to Minocha *et al.* (2009).

Measurement of leaf anatomical parameters by optical microscopy

Similar to what was done for leaf chlorophylls extraction, at 38 DAS_s two leaf discs (diameter 6 mm) were collected from the most developed leaf of each plant and transferred to 1 mL Accustain™ fixative (Sigma-Aldrich, St. Louis, USA). All samples underwent three cycles of vacuum infiltration at -800 mbar for 15 min followed by 30 min at ambient pressure and overnight incubation in the fixative. Samples were then dehydrated with an ethanol series and embedded in Technovit 7100 resin (Kulzer Technik, Wehrheim, Germany) with a modified version of the protocol by Yeung and Chan (Yeung and Chan, 2015). Embedded leaves were sectioned with a Reichert-Jung 2040 Autocut Rotary Microtome (Leica Biosystems, Nussloch, Germany) set at 5 µm section thickness. Sections mounted on microscopy slides were stained with a 0.05% toluidine blue-O, 1% sodium tetraborate solution and imaged in bright

field at 10x magnification with an AXIO Imager.A1 microscope equipped with an AxioCam MRc5 camera (Carl Zeiss, Köln, Germany). At least twelve captured images per plant leaf sample were analysed manually with Fiji v2.12.0 (Schindelin *et al.*, 2012) to determine the area of palisade and spongy mesophyll cells, and overall leaf thicnkess.

Measurement of abundance and dimensions of stomata

At 38 DAS_s, clear nail polish was applied to 2 cm^2 on both the ad- and abaxial sides of the most developed leaf of each plant and removed when dry with clear adhesive tape. Leaf impressions were then transferred to microscopy slides and imaged with an Axiophot microscope (Carl Zeiss, Köln, Germany) equipped with a DFC340 FX monochrome camera (Leica Microsystems, Wetzlar, Germany). At least twelve captured images per plant leaf sample were analysed. Counting and measuring of the stomatal length were automated with a Python v3.10.10 script running StarDist v0.8.3 (Schmidt *et al.*, 2018; Weigert *et al.*, 2020). A training set of annotated images was produced using Qupath v0.3.2 (Bankhead *et al.*, 2017).

Statistical analysis of measured parameters

Between-species differences for the various parameters were tested in different ways according to the nature of the respective datasets. For time-series data (ΦPSII, ExGI, NDVI), a generalized linear mixed-effect model (GLMM) accounting for block design, time of measurement, and repeated measurements of individual plants were fitted on the whole dataset with package glmmTMB v1.1.7 (Brooks *et al.*, 2017). Conditional and marginal coefficients of determination (pseudo-R^2) were calculated for the models with the R package MuMIn v1.47.5 (Barton, 2009).

In order to obtain a more detailed description of between-species differences, linear mixed-effect models (LMMs) accounting for block design were fitted on data corresponding to each measurement round with R package lme4 v1.1-33 (Bates *et al.*, 2015). The assumption of normal distribution of residuals was tested by scaling the residuals and testing the difference of their distribution to the normal distribution with the Kolmogorov-Smirnov test. Estimated marginal means (EMMs, also known as least-squares means) were then computed with package emmeans v1.8.6 (Lenth, 2023) using the Kenward-Roger approximation for degrees of freedom, and were subsequently employed to perform pairwise comparisons between species with package multcomp v1.4-23 (Hothorn *et al.*, 2008) using the Benjamini-Hochberg method for reduction of false discovery rate. Heritability of differences was calculated with the same approach and a modified linear model including the species experimental factor as a random-effects term. The methodology described for single measurement rounds was also employed to test between-species differences for single-time point data (anatomical parameters and chlorophyll content).

Integration and analysis of results

Block-effect corrected data were predicted from the (G)LMMs fitted to the various measurement datasets. Time-series parameter datasets were filtered to select only observations belonging to the round resulting in the highest heritability and the round closest to the end of the experiment in which overlap between leaves from different plants was judged to be minimal.

Missing values due to parameter-specific filtering or missing samples were imputed with package missMDA v1.18 (Josse and Husson, 2016), based on a Factorial Analysis for Mixed Data (FAMD) model. Principal Component Analysis (PCA) was performed on the generated dataset with package FactoMineR v2.8 (Lê *et al.*, 2008) and results were visualized with package FactoExtra v1.0.7 (Kassambara and Mundt, 2020). The dataset was further explored by means of nonlinear dimensionality reduction with t-distributed stochastic neighbor embedding (t-SNE) performed with package M3C v1.20.0 (John *et al.*, 2020). Individuals (i.e., measured plants) were clustered by means of k-means clustering (package stats v4.2.2), based on scaled non-centered values for all parameters. Clustering results were visualized with packages FactoExtra v1.0.7 and ggplot2 v3.4.2 (Wickham, 2016). Differences between species-based groups of individuals were tested with Permutational Analysis of Variance (PERMANOVA) and pairwise-PERMANOVA models from packages vegan v2.6-4 (Oksanen *et al.*, 2022) and RVAideMemoire v0.9-83 (Hervé, 2023). Prior to these analysis, an analysis of multivariate homogeneity of group dispersions was performed with package vegan v2.6-4. Pearson correlation coefficients (r) between parameters and associated p-values were calculated with package Hmisc v5.1-0 (Harrell Jr, 2023) and visualized with package corrplot v0.92 (Wei and Simko, 2021). Hierarchical clustering based on distances calculated as (1-r) was performed and visualized with packages stats v4.2.2 and graphics v4.2.2.

Availability of data and materials

All scripts employed to analyze data are available on the 4TU.ResearchData platform with DOI 10.4121/13f4d7fa-63d2-4892-ab7e-36ff08cb410e. Supplementary information accompanying this chapter and the data files employed for analysis are available on the 4TU.ResearchData platform with DOI 10.4121/79a62b5f-2881-4520-b031-e03334c02aad.

Author contributions

F.G. and L.C. designed the study in consultation with M.Aa. and J.H.. M.Al. optimized the microscopy protocol for imaging of leaf sections. L.C. developed the protocol for imaging and automated counting of stomata. J.vdB., R.P. and F.S. performed sample preparation, imaging and analysis for leaf anatomy parameters' measurements and stomatal counting. L.C. extracted and quantified leaf chlorophylls. F.G. supervised all microscopy work, analysed all data, and drafted the manuscript. D.dR. consulted on the integration of results and revized the manuscript. L.C., J.H., and M.Aa. contributed to the manuscript. All authors have read and approved the contents of the manuscript.

Acknowledgements

We thank Marcus Koch for providing us with the *Zahora ait-atta* seeds, Magdalena Smigaj for the help and discussions on vegetation indices, Sabine Schnabel for her guidance on statistical analysis, and Sofia Bengoa Luoni for her oversight of image analysis. Furthermore, we thank Nam Hoang for contributing to the selection of species, Eric Schranz for his contributions to the interpretation of results, and Louise Logie and Sandra Smit for providing feedback on the manuscript.

References

Adak, A., Murray, S. C., Božinović, S., Lindsey, R., Nakasagga, S., Chatterjee, S., Anderson, S. L., and Wilde, S. (2021). Temporal Vegetation Indices and Plant Height from Remotely Sensed Imagery Can Predict Grain Yield and Flowering Time Breeding Value in Maize via Machine Learning Regression. *Remote Sensing*, **13**(11), 2141.

Adams III, W. W. and Terashima, I., editors (2018). *The Leaf: A Platform for Performing Photosynthesis*, volume 44 of *Advances in Photosynthesis and Respiration*. Springer International Publishing, Cham.

Arias, T. and Pires, J. C. (2012). A fully resolved chloroplast phylogeny of the brassica crops and wild relatives (Brassicaceae: Brassiceae): Novel clades and potential taxonomic implications. *TAXON*, **61**(5), 980–988.

Austin, R. B. (1989). Genetic variation in photosynthesis. *The Journal of Agricultural Science*, **112**(3), 287–294.

Avratovščuková, N. and Fousová, S. (1975). Genetic Variation of Photosynthetic Rate in Leaf Discs of Zea Mays L. In Yu. S. Nasyrov and Z. Šesták, editors, *Genetic Aspects of Photosynthesis*, pages 343–347. Springer Netherlands, Dordrecht.

Balfagón, D., Sengupta, S., Gómez-Cadenas, A., Fritschi, F. B., Azad, R. K., Mittler, R., and Zandalinas, S. I. (2019). Jasmonic Acid Is Required for Plant Acclimation to a Combination of High Light and Heat Stress. *Plant Physiology*, **181**(4), 1668–1682.

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., and Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, **7**(1), 16878.

Barton, K. (2009). MuMIn: Multi-model inference. *http://r-forge. r-project. org/projects/mumin/*.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bräutigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K. L., Carr, K. M., Gowik, U., Maß, J., Lercher, M. J., Westhoff, P., Hibberd, J. M., and Weber, A. P. (2011). An mRNA Blueprint for C4 Photosynthesis Derived from Comparative Transcriptomics of Closely Related C3 and C4 Species. *Plant Physiology*, **155**(1), 142–156.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., and Bolker, B. M. (2017). Modeling zero-inflated count data with glmmTMB. *BioRxiv*, page 132753.

Buchaillot, M. L., Soba, D., Shu, T., Liu, J., Aranjuelo, I., Araus, J. L., Runion, G. B., Prior, S. A., Kefauver, S. C., and Sanz-Saez, A. (2022). Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models. *Planta*, **255**(4), 93.

De Souza, A. P., Burgess, S. J., Doran, L., Hansen, J., Manukyan, L., Maryn, N., Gotarkar, D., Leonelli, L., Niyogi, K. K., and Long, S. P. (2022). Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science*, **377**(6608), 851–854.

De Swaef, T., Maes, W. H., Aper, J., Baert, J., Cougnon, M., Reheul, D., Steppe, K., Roldán-Ruiz, I., and Lootens, P. (2021). Applying RGB- and Thermal-Based Vegetation Indices from UAVs for High-Throughput Field Phenotyping of Drought Tolerance in Forage Grasses. *Remote Sensing*, **13**(1), 147.

Du, T., Meng, P., Huang, J., Peng, S., and Xiong, D. (2020). Fast photosynthesis measurements for phenotyping photosynthetic capacity of rice. *Plant Methods*, **16**(1), 6.

ENEA TER-SOLTERM (2006). ENEA - Solaritaly — DLI of Italy. http://www.solaritaly.enea.it/DLI/DLIMappeEn.php.

Fahlgren, N., Feldman, M., Gehan, M. A., Wilson, M. S., Shyu, C., Bryant, D. W., Hill, S. T., McEntee, C. J., Warnasooriya, S. N., Kumar, I., Ficor, T., Turnipseed, S., Gilbert, K. B., Brutnell, T. P., Carrington, J. C., Mockler, T. C., and Baxter, I. (2015). A Versatile Phenotyping System and Analytics Platform Reveals Diverse Temporal Responses to Water Availability in Setaria. *Molecular Plant*, **8**(10), 1520–1535.

Faralli, M. and Lawson, T. (2020). Natural genetic variation in photosynthesis: An untapped resource to increase crop yield potential? *The Plant Journal*, **101**(3), 518–528.

Flood, P. J., Kruijer, W., Schnabel, S. K., van der Schoor, R., Jalink, H., Snel, J. F. H., Harbinson, J., and Aarts, M. G. M. (2016). Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. *Plant Methods*, **12**(1), 14.

Flood, P. J., Theeuwen, T. P. J. M., Schneeberger, K., Keizer, P., Kruijer, W., Severing, E., Kouklas, E., Hageman, J. A., Wijfjes, R., Calvo-Baltanas, V., Becker, F. F. M., Schnabel, S. K., Willems, L. A. J., Ligterink, W., van Arkel, J., Mumm, R., Gualberto, J. M., Savage, L., Kramer, D. M., Keurentjes, J. J. B., van Eeuwijk, F., Koornneef, M., Harbinson, J., Aarts, M. G. M., and Wijnker, E. (2020). Reciprocal cybrids reveal how organellar genomes affect plant phenotypes. *Nature Plants*, **6**(1), 13–21.

Formighieri, C. (2015). Light Saturation of Photosynthesis. In C. Formighieri, editor, *Solar-to-Fuel Conversion in Algae and Cyanobacteria*, SpringerBriefs in Environmental Science, pages 55–58. Springer International Publishing, Cham.

Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A., and Mummenhoff, K. (2011). Cabbage family affairs: The evolutionary history of Brassicaceae. *Trends in Plant Science*, **16**(2), 108–116.

Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., Penuelas, J., and Valentini, R. (1995). Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecological Applications*, **5**(1), 28–41.

Garassino, F., Wijfjes, R. Y., Boesten, R., Reyes Marquez, F., Becker, F. F. M., Clapero, V., van den Hatert, I., Holmer, R., Schranz, M. E., Harbinson, J., de Ridder, D., Smit, S., and Aarts, M. G. M. (2022). The genome sequence of Hirschfeldia incana, a new Brassicaceae model to improve photosynthetic light-use efficiency. *The Plant Journal*, **112**(1), 1298–1315.

Garassino, F., Harbinson, J., Aarts, M. G. M., and Smit, S. (2023). Transcriptional analysis of Hirschfeldia incana and relatives highlights differences in photosynthetic pathways.

Garcia-Molina, A. and Leister, D. (2020). Accelerated relaxation of photoprotection impairs biomass accumulation in Arabidopsis. *Nature Plants*, **6**(1), 9–12.

Ge, Y., Bai, G., Stoerger, V., and Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*, **127**, 625–632.

Gramm, J., Guo, J., Hüffner, F., Niedermeier, R., Piepho, H.-P., and Schmid, R. (2007). Algorithms for compact letter displays: Comparison and evaluation. *Computational Statistics & Data Analysis*, **52**(2), 725–736.

Harrell Jr, F. E. (2023). *Hmisc: Harrell Miscellaneous*.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, **585**(7825), 357–362.

Hervé, M. (2023). *RVAideMemoire: Testing and Plotting Procedures for Biostatistics*.

Hoang, N. V., Sogbohossou, E. O. D., Xiong, W., Simpson, C. J. C., Singh, P., Walden, N., van den Bergh, E., Becker, F. F. M., Li, Z., Zhu, X.-G., Brautigam, A., Weber, A. P. M., van Haarst, J. C., Schijlen, E. G. W. M., Hendre, P. S., Van Deynze, A., Achigan-Dako, E. G., Hibberd, J. M., and Schranz, M. E. (2023). The Gynandropsis gynandra genome provides insights into whole-genome duplications and the evolution of C4 photosynthesis in Cleomaceae. *The Plant Cell*, **35**(5), 1334–1359.

Hodges, J. D. (1967). Patterns of Photosynthesis under Natural Environmental Conditions. *Ecology*, **48**(2), 234–242.

Hogewoning, S. W., Wientjes, E., Douwstra, P., Trouwborst, G., van Ieperen, W., Croce, R., and Harbinson, J. (2012). Photosynthetic Quantum Yield Dynamics: From Photosystems to Leaves. *The Plant Cell*, **24**(5), 1921–1935.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, **50**(3), 346–363.

Jiang, Y., Yang, Y., Jiang, Y., and Yang, Y. (2022). High-throughput phenotyping for plant growth and biomass yield of switchgrass under a controlled environment. *Grass Research*, **2**(GR-2022-0004), 1–7.

Jin, J., Wang, Q., and Song, G. (2022). Selecting informative bands for partial least squares regressions improves their goodness-of-fits to estimate leaf photosynthetic parameters from hyperspectral data. *Photosynthesis Research*, **151**(1), 71–82.

John, C. R., Watson, D., Russ, D., Goldmann, K., Ehrenstein, M., Pitzalis, C., Lewis, M., and Barnes, M. (2020). M3C: Monte Carlo reference-based consensus clustering. *Scientific Reports*, **10**(1), 1–14.

Josse, J. and Husson, F. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, **70**(1), 1–31.

Junker, L. V. and Ensminger, I. (2016). Relationship between leaf optical properties, chlorophyll fluorescence and pigment changes in senescing Acer saccharum leaves. *Tree Physiology*, **36**(6), 694–711.

Kassambara, A. and Mundt, F. (2020). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*.

Khan, H. A., Nakamura, Y., Furbank, R. T., and Evans, J. R. (2021). Effect of leaf temperature on the estimation of photosynthetic and other traits of wheat leaves from hyperspectral reflectance. *Journal of Experimental Botany*, **72**(4), 1271–1281.

Koch, M. A. and Lemmel, C. (2019). Zahora, a new monotypic genus from tribe Brassiceae (Brassicaceae) endemic to the Moroccan Sahara. *PhytoKeys*, **135**, 119–131.

Korczynski, P. C., Logan, J., and Faust, J. E. (2002). Mapping Monthly Distribution of Daily Light Integrals across the Contiguous United States. *HortTechnology*, **12**(1), 12–16.

Krause, G. H. and Weis, E. (1991). Chlorophyll Fluorescence and Photosynthesis: The Basics. *Annual Review of Plant Physiology and Plant Molecular Biology*, **42**(1), 313–349.

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**(6314), 857–861.

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: A package for multivariate analysis. *Journal of Statistical Software*, **25**(1), 1–18.

Lenth, R. V. (2023). *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*.

Li, W., Liu, J., Zhang, H., Liu, Z., Wang, Y., Xing, L., He, Q., and Du, H. (2022). Plant pan-genomics: Recent advances, new challenges, and roads ahead. *Journal of Genetics and Genomics*, **49**(9), 833–846.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A. P., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, X., Wang, J., Lu, K., Fang, Z., Bancroft, I., Yang, T.-J., Hu, Q., Wang, X., Yue, Z., Li, H., Yang, L., Wu, J., Zhou, Q., Wang, W., King, G. J., Pires, J. C., Lu, C., Wu, Z., Sampath, P., Wang, Z., Guo, H., Pan, S., Yang, L., Min, J., Zhang, D., Jin, D., Li, W., Belcram, H., Tu, J., Guan, M., Qi, C., Du, D., Li, J., Jiang, L., Batley, J., Sharpe, A. G., Park, B.-S., Ruperao, P., Cheng, F., Waminal, N. E., Huang, Y., Dong, C., Wang, L., Li, J., Hu, Z., Zhuang, M., Huang, Y., Huang, J., Shi, J., Mei, D., Liu, J., Lee, T.-H., Wang, J., Jin, H., Li, Z., Li, X., Zhang, J., Xiao, L., Zhou, Y., Liu, Z., Liu, X., Qin, R., Tang, X., Liu, W., Wang, Y., Zhang, Y., Lee, J., Kim, H. H., Denoeud, F., Xu, X., Liang, X., Hua, W., Wang, X., Wang, J., Chalhoub, B., and Paterson, A. H. (2014). The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, **5**(1), 3930.

Long, S. P., Zhu, X.-G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment*, **29**(3), 315–330.

López-Calcagno, P. E., Fisk, S., Brown, K. L., Bull, S. E., South, P. F., and Raines, C. A. (2019). Overexpressing the H-protein of the glycine cleavage system increases biomass yield in glasshouse and field-grown transgenic tobacco plants. *Plant Biotechnology Journal*, **17**(1), 141–151.

López-Calcagno, P. E., Brown, K. L., Simkin, A. J., Fisk, S. J., Vialet-Chabrand, S., Lawson, T., and Raines, C. A. (2020). Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants*, **6**(8), 1054–1063.

Maxwell, K. and Johnson, G. N. (2000). Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany*, **51**(345), 659–668.

Mazis, A., Choudhury, S. D., Morgan, P. B., Stoerger, V., Hiller, J., Ge, Y., and Awada, T. (2020). Application of high-throughput plant phenotyping for assessing biophysical traits and drought response in two oak species under controlled environment. *Forest Ecology and Management*, **465**, 118101.

Meacham-Hensold, K., Fu, P., Wu, J., Serbin, S., Montes, C. M., Ainsworth, E., Guan, K., Dracup, E., Pederson, T., Driever, S., and Bernacchi, C. (2020). Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging. *Journal of Experimental Botany*, **71**(7), 2312–2328.

Minocha, R., Martinez, G., Lyons, B., and Long, S. (2009). Development of a standardized methodology for quantifying total chlorophyll and carotenoids from foliage of hardwood and conifer tree species. *Canadian Journal of Forest Research*, **39**(4), 849–861.

Mizokami, Y., Oguchi, R., Sugiura, D., Yamori, W., Noguchi, K., and Terashima, I. (2022). Cost–benefit analysis of mesophyll conductance: Diversities of anatomical, biochemical and environmental determinants. *Annals of Botany*, **130**(3), 265–283.

Nyongesah, M. J., Wang, Q., and Li, P. (2015). Effectiveness of photochemical reflectance index to trace vertical and seasonal chlorophyll a/b ratio in Haloxylon ammodendron. *Acta Physiologiae Plantarum*, **37**(2), 2.

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M. O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C. J., and Weedon, J. (2022). *Vegan: Community Ecology Package*.

Pech, R., Volná, A., Hunt, L., Bartas, M., Červeň, J., Pečinka, P., Špunda, V., and Nezval, J. (2022). Regulation of Phenolic Compound Production by Light Varying in Spectral Quality and Total Irradiance. *International Journal of Molecular Sciences*, **23**(12), 6533.

Pérez-Valencia, D. M., Rodríguez-Álvarez, M. X., Boer, M. P., Kronenberg, L., Hund, A., Cabrera-Bosquet, L., Millet, E. J., and van Eeuwijk, F. A. (2022). A two-stage approach for the spatio-temporal analysis of high-throughput phenotyping data. *Scientific Reports*, **12**(1), 3177.

Petereit, J., Bayer, P. E., Thomas, W. J. W., Tay Fernandez, C. G., Amas, J., Zhang, Y., Batley, J., and Edwards, D. (2022). Pangenomics and Crop Genome Adaptation in a Changing Climate. *Plants*, **11**(15), 1949.

Richardson, A. D., Jenkins, J. P., Braswell, B. H., Hollinger, D. Y., Ollinger, S. V., and Smith, M.-L. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, **152**(2), 323–334.

Richardson, A. D., Braswell, B. H., Hollinger, D. Y., Jenkins, J. P., and Ollinger, S. V. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications*, **19**(6), 1417–1428.

Rufo, R., Soriano, J. M., Villegas, D., Royo, C., and Bellvert, J. (2021). Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment. *Remote Sensing*, **13**(6), 1187.

Ruiz-Vera, U. M., Acevedo-Siaca, L. G., Brown, K. L., Afamefule, C., Gherli, H., Simkin, A. J., Long, S. P., Lawson, T., and Raines, C. A. (2022). Field-grown ictB tobacco transformants show no difference in photosynthetic efficiency for biomass relative to the wild type. *Journal of Experimental Botany*, **73**(14), 4897–4907.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, **9**(7), 676–682.

Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11*, pages 265–273. Springer.

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, **6**(8), 1335–1372.

Serba, D. D., Uppalapati, S. R., Krom, N., Mukherjee, S., Tang, Y., Mysore, K. S., and Saha, M. C. (2016). Transcriptome analysis in switchgrass discloses ecotype difference in photosynthetic efficiency. *BMC Genomics*, **17**(1), 1040.

Shi, J., Tian, Z., Lai, J., and Huang, X. (2023). Plant pan-genomics and its applications. *Molecular Plant*.

Siebers, M. H., Gomez-Casanovas, N., Fu, P., Meacham-Hensold, K., Moore, C. E., and Bernacchi, C. J. (2021). Emerging approaches to measure photosynthesis from the leaf to the ecosystem. *Emerging Topics in Life Sciences*, **5**(2), 261–274.

Song, G., Wang, Q., and Jin, J. (2022). Temporal instability of partial least squares regressions for estimating leaf photosynthetic traits from hyperspectral information. *Journal of Plant Physiology*, **279**, 153831.

Talling, J. F. (1961). Photosynthesis Under Natural Conditions. *Annual Review of Plant Physiology*, **12**(1), 133–154.

Tardieu, F., Bouidghaghen, J., Moreau, L., Beauchene, K., Chapuis, R., Mangel, N., Bosquet, L. C., Welcker, C., and Bogard, M. (2023). Robotized indoor phenotyping allows genomic prediction of adaptive traits in the field. *Research Square Preprint*.

Taylor, G., Garassino, F., Aarts, M. G. M., and Harbinson, J. (2022). Improving C3 photosynthesis by exploiting natural genetic variation: Hirschfeldia incana as a model species. *Food and Energy Security*, **n/a**(e420).

Terashima, I., Hanba, Y. T., Tholen, D., and Niinemets, Ü. (2011). Leaf Functional Anatomy in Relation to Photosynthesis. *Plant Physiology*, **155**(1), 108–116.

Theeuwen, T. P. J. M., Logie, L. L., Harbinson, J., and Aarts, M. G. M. (2022). Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany*, **73**(10), 3122–3137.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, **9**(11).

van Eeuwijk, F. A., Bustos-Korts, D., Millet, E. J., Boer, M. P., Kruijer, W., Thompson, A., Malosetti, M., Iwata, H., Quiroz, R., Kuppe, C., Muller, O., Blazakis, K. N., Yu, K., Tardieu, F., and Chapman, S. C. (2019). Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. *Plant Science*, **282**, 23–39.

van Rooijen, R., Aarts, M., and Harbinson, J. (2015). Natural Genetic Variation for Acclimation of Photosynthetic Light Use Efficiency to Growth Irradiance in Arabidopsis. *Plant Physiology*, **167**(4), 1412–1429.

van Rooijen, R., Kruijer, W., Boesten, R., van Eeuwijk, F. A., Harbinson, J., and Aarts, M. G. M. (2017). Natural variation of YELLOW SEEDLING1 affects photosynthetic acclimation of Arabidopsis thaliana. *Nature Communications*, **8**(1), 1421.

van Rooijen, R., Harbinson, J., and Aarts, M. G. M. (2018). Photosynthetic response to increased irradiance correlates to variation in transcriptional response of lipid-remodeling and heat-shock genes. *Plant Direct*, **2**(7), e00069.

Visscher, P. M., Hill, W. G., and Wray, N. R. (2008). Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics*, **9**(4), 255–266.

Wall, S., Vialet-Chabrand, S., Davey, P., Van Rie, J., Galle, A., Cockram, J., and Lawson, T. (2022). Stomata on the abaxial and adaxial leaf surfaces contribute differently to leaf gas exchange and photosynthesis in wheat. *New Phytologist*, **235**(5), 1743–1756.

Wang, H., Jia, G., Epstein, H. E., Zhao, H., and Zhang, A. (2020). Integrating a PhenoCam-derived vegetation index into a light use efficiency model to estimate daily gross primary production in a semi-arid grassland. *Agricultural and Forest Meteorology*, **288–289**, 107983.

Wei, T. and Simko, V. (2021). *R Package 'Corrplot': Visualization of a Correlation Matrix*.

Weigert, M., Schmidt, U., Haase, R., Sugawara, K., and Myers, G. (2020). Star-convex polyhedra for 3D object detection and segmentation in microscopy. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3666–3673.

Werk, K. S., Ehleringer, J., Forseth, I. N., and Cook, C. S. (1983). Photosynthetic characteristics of Sonoran Desert winter annuals. *Oecologia*, **59**(1), 101–105.

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, **38**(1), 259–269.

Xiong, D. and Flexas, J. (2020). From one side to two sides: The effects of stomatal distribution on photosynthesis. *New Phytologist*, **228**(6), 1754–1766.

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., and Yan, J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*, **13**(2), 187–214.

Yeung, E. C. and Chan, C. K. W. (2015). The Glycol Methacrylate Embedding Resins—Technovit 7100 and 8100. In E. C. T. Yeung, C. Stasolla, M. J. Sumner, and B. Q. Huang, editors, *Plant Microtechniques and Protocols*, pages 67–82. Springer International Publishing, Cham.

Yoon, D.-K., Ishiyama, K., Suganami, M., Tazoe, Y., Watanabe, M., Imaruoka, S., Ogura, M., Ishida, H., Suzuki, Y., Obara, M., Mae, T., and Makino, A. (2020). Transgenic rice overproducing Rubisco exhibits increased yields with improved nitrogen-use efficiency in an experimental paddy field. *Nature Food*, **1**(2), 134–139.

Zhang, F. and Zhou, G. (2019). Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize. *BMC Ecology*, **19**(1), 18.

Zhou, H., Zhou, G., Song, X., and He, Q. (2022). Dynamic Characteristics of Canopy and Vegetation Water Content during an Entire Maize Growing Season in Relation to Spectral-Based Indices. *Remote Sensing*, **14**(3), 584.

Zhu, X.-G., Long, S. P., and Ort, D. R. (2010). Improving Photosynthetic Efficiency for Greater Yield. *Annual Review of Plant Biology*, **61**(1), 235–261.

CHAPTER 6

General discussion

"Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."

Winston Churchill, Their Finest Hour

In the previous five chapters, I first introduced current trends in photosynthesis research and the rationale of the work I conducted. I then illustrated the most relevant results obtained from *H. incana* and a number of other Brassicaceae species. While detailed discussions of all results are included in each of the experimental chapters, this final chapter presents several more general considerations. The research presented in this thesis was started with the objective of identifying some of the genetic determinants, or part of the "genetic strategy", of *Hirschfeldia incana*'s striking photosynthetic light-use efficiency (LUE) and P_{max}. In this chapter, I first summarize all my findings about the photosynthetic LUE of *H. incana* and discuss how exceptional they are and how they relate to the LUE of other Brassicaceae species. Following a similar comparative approach, I then provide a unifying view of all findings based on comparative genomics and transcriptomics investigations coworkers and I conducted. I then interpret all the most relevant experimental results employing concepts borrowed from ecophysiology. Subsequently, I discuss how photosynthesis research will be more likely to achieve the goal of increasing crops' yields if we surpass the current leaf- and crop-centric perspectives. These perspectives were initially adopted to mitigate the high complexity associated with the photosynthetic process and its interactions with the environment. I argue that given the limited success obtained so far in improving crop yields via increased photosynthesis, the time has come to adopt broader, more holistic perspectives. Finally, I revisit my most important findings and suggest directions for future research.

6.1 Describing the photosynthetic performance of *Hirschfeldia incana*

Since the photosynthesis of *H. incana* had only been reported in a single work published over forty years ago (Canvin *et al.*, 1980), a more detailed study of *H. incana*'s photosynthesis was required. The results of this study are not organised within a dedicated chapter of this thesis but rather spread across three of the thesis' chapters.

In **Chapter 2**, I presented measurements conducted in preliminary laboratory studies on *H. incana* alone, highlighting that the photosynthetic apparatus in its leaves is remarkably robust to high irradiances, exhibiting little to no physiological stress or photoinhibition at measurement irradiances of up to 2400 $\mu mol\, m^{-2}\, s^{-1}$, levels higher than could be expected on a cloudless day at noon at the Equator during an equinox (approximately 2200 $\mu mol\, m^{-2}\, s^{-1}$) (Ritchie, 2010).

In **Chapter 3**, I reported on our first comparative exploration of photosynthetic rates between *H. incana*, *Brassica nigra*, *Brassica rapa*, and *Arabidopsis thaliana*, the species that have served as comparisons for most of this research. By growing plants under high, super-natural irradiances and measuring the response of CO_2 assimilation rate to irradiance, I demonstrated the higher photosynthetic rates of two different genotypes of *H. incana* compared to the other species. Growing plants under constant high irradiance in experimental settings was not trivial at the beginning of this research, and despite the introduction of increasingly powerful LED fixtures

over years still is not, as demonstrated by the custom-built lighting systems employed for research on the controlled-environment agriculture crop for which irradiance matters most, cannabis (*Cannabis spp.*) (Rodriguez-Morrison *et al.*, 2021; Moher *et al.*, 2022; Llewellyn *et al.*, 2022). Therefore, co-workers and I developed a custom solution to deliver super-natural, highly uniform irradiance at the plant level and integrated it into a growth room as described in **Chapter 4**.

The low throughput of the techniques employed in the studies reported so far prevented us from performing a more extended study of photosynthetic LUE across the Brassicaceae family, to determine if *H. incana* is an outlier within the family or one of many representatives of a clear evolutionary trend favoring higher photosynthetic LUE. Therefore, in **Chapter 5**, I present the results of an exploration conducted with a different, high-throughput method: chlorophyll fluorescence-based estimation of the operational quantum yield of Photosystem II (ΦPSII). Despite being grown under lower irradiances than used in previous studies, amounting to Daily Light Integrals (DLIs) comparable to those recorded in warm-temperate areas of the globe in spring and autumn, *H. incana* plants showed very high ΦPSII values, only matched by those of *B. nigra*, *Zahora ait-atta*, and occasionally *Brassica tournefortii*. Based on these results, one may be tempted to conclude that *H. incana* is not the sole example of high photosynthetic LUE within the Brassicaceae family, a conclusion that would be supported by additional gas-exchange-based data collected on *H. incana*, *B. nigra*, *B. rapa*, and *A. thaliana* in another project (Retta *et al.*, in preparation). This conclusion appears to be in contrast to what is described in **Chapter 3**, but an alternative conclusion can be proposed, which leads to an interesting hypothesis for follow-up research. In **Chapter 5** I also performed an analysis of ΦPSII trends over the plants' growth and demonstrated how *H. incana* achieved high values from very early stages of growth under high irradiance, contrary to what observed for most other species. This could suggest that *H. incana* is better equipped to maintain a higher photosynthetic LUE under challenging conditions than other Brassicaceae species. This hypothesis is supported by the higher leaf thickness reported for *H. incana* compared to *B. nigra* and *A. thaliana* in **Chapter 5** and that observed for *H. incana* compared to *B. nigra*, *B. rapa*, and *A. thaliana* in other research (Retta *et al.*, in preparation).

Considering that *H. incana* originated in warm-temperate climates, and possibly in the Mediterranean basin, its ability to achieve higher photosynthetic LUE under resource-limited conditions would be in line with what was observed for other Mediterranean species (Flexas *et al.*, 2014). Striking similarities in anatomical strategies can be found between *H. incana* and the more studied Mediterranean species *Capparis spinosa*, the caper plant. A link between leaf thickness, drought resistance, and higher photosynthetic LUE has been clarified for this species (Rhizopolou and Psaras, 2003; Levizou *et al.*, 2004; Gan *et al.*, 2013), and could be postulated for *H. incana* as well.

Furthermore, a leaf thickness increase in response to water scarcity has been reported in tomato (*Solanum lycopersicum*) and tobacco (*Nicothiana tabacum*) plants, further implicating leaf thickness adjustment as a mechanism to alleviate water stress (Galmés *et al.*, 2013; Khan *et al.*, 2023). Another observation supporting the hypothesis of a higher photosynthetic LUE of *H. incana* under stress conditions is the

higher leaf nitrogen content for *H. incana* than other species measured in a project related to this (Retta *et al.*, in preparation). In addition to being the fundamental component of proteins, nitrogen can play a key role in maintaining plants' photosynthetic performance in water or salt stress conditions, due to the accumulation of nitrogenous osmolytes (Tipirdamaz *et al.*, 2006; Ahanger *et al.*, 2019, 2021), and could be therefore another indication of a strategy evolved by *H. incana* to maintain its photosynthetic performance in water-scarce environments. Higher photosynthetic LUE means higher CO_2 assimilation and thus more energy for plant establishment and growth, so *H. incana*'s strategy could grant the species a competitive advantage over other species when establishing and growing in water-limited conditions.

6.2 Probing the genetic strategy responsible for the photosynthetic LUE of *H. incana* with genomics and transcriptomics

Photosynthesis is a complex biochemical process that can be subdivided into a number of traits, as I argued in **Chapter 2**. Although mostly interdependent, these traits are usually analyzed individually or in restricted sets during physiological, genetic, and ecological studies, due to their intricate genetic determinants (van Bezouw *et al.*, 2019). One could say that the molecular pathway from genotype to the photosynthetic LUE phenotype is therefore highly complex, similar to what is observed for many other traits determining plants' and crops' fitness and productivity (Paterson, 1998; Glazier *et al.*, 2002; Holland, 2007; Saini *et al.*, 2021). Since the field of genomics was revolutionized by the sequencing of the human genome (Lander *et al.*, 2001; Venter *et al.*, 2001), the scientific community has been probing such complexity and developing technologies making genetic information increasingly accessible. In this research, co-authors and I leveraged modern genomics and transcriptomics technologies and tools to explore the genomes of *H. incana* and relatives and understand the transcriptional activity in relation to irradiance, aiming to highlight some genetic determinants of photosynthetic light-use efficiency under high irradiance.

Given the little attention received by *H. incana* within the scientific community, it is not surprising that its genome had not been sequenced before the start of this project. In **Chapter 3**, co-authors and I reported the assembly and annotation of the *H. incana* genome. We then proceeded to test the hypothesis that differential retention of copies of genes duplicated following the genome triplication event the common ancestor of the three species underwent could contribute to *H. incana*'s higher photosynthetic LUE. By studying the gene expression of eighteen genes involved in photosynthesis and/or photoprotection, nine of which showing inter-species copy number variation (CNV) and nine present in single copies in all genomes, we found an indication of a positive correlation between photosynthetic/photoprotective genes' copy number and their expression. Such a relationship between gene copy number and expression had already been highlighted as a driver for traits' evolution (Kondrashov, 2012; Żmieńko *et al.*, 2014; Patterson *et al.*, 2018) and thus supports

our hypothesis that CNV of photosynthetic and photoprotective genes was an important component of *H. incana*'s genetic makeup for high photosynthetic LUE. However, having observed similar copy number retention and gene expression patterns in species that achieved lower photosynthetic light-use efficiencies, namely *B. nigra* and *B. rapa*, we concluded that more research was required to further clarify the transcriptional basis of *H. incana*'s performance.

Going beyond the few target genes studies in Chapter 3, in **Chapter 4** I presented the results of a genome-wide transcriptome analysis conducted on the four species grown under contrasting high and low irradiances. I analyzed gene expression differences in *H. incana*, *B. nigra*, *B. rapa*, and *A. thaliana* focusing on differences in expression of photosynthesis-related genes *per se*, rather than differences that could directly be linked to CNV. While RNA-Seq methodologies are now well-established, cross-species analyses are complicated by challenges such as the lack of high-quality reference genome assemblies for all species and difficulties in quantitatively comparing transcriptomes from distant species (Roux *et al.*, 2015; Parekh *et al.*, 2018; Rivarola Sena *et al.*, 2022).

I addressed these challenges by constructing a panproteome, clustering the proteins predicted from genes annotated for eight Brassicaceae genomes and using inferred gene homology relationships to group expression data for the various species. Enrichment analyses on *A. thaliana* orthologs from responsive core homology groups, i.e., those containing at least a gene from each the four studied species and at least one gene showing differential expression, highlighted photosynthetic pathways. Additionally to an expected generalized transcriptional response for photosynthetic genes in all species due to the contrasting irradiance conditions plants were exposed to, notable *H. incana*-specific transcriptional patterns could be identified. These patterns can be further classified based on the underlying genes' copy number status. Indeed, I identified cases in which single-copy *H. incana* genes underwent constitutive higher expression or differential expression as a result of irradiance, as well as cases in which genes having a higher copy number in *H. incana* achieved higher transcript abundances constitutively or in one of the irradiance conditions.

The combined findings of chapters 3 and 4, therefore, support the original hypothesis on the involvement of gene CNV in the genetic background of *H. incana*'s high photosynthetic LUE. However, the extent of the contribution of CNVs to this trait remains to be addressed, as the analysis of full transcriptomes revealed interesting *H. incana*-specific responses in single-copy genes as well.

6.3 An ecological interpretation of findings: many roads lead to high light-use efficiency?

So far, this Discussion has been focused on connecting selected findings from the various experiments reported in this thesis. As mentioned above, many photosynthetic traits appear to be interdependent or correlated. Thus, adopting a perspective encompassing this interdependence can help in formulating a collective summary of

the findings described in this thesis. Such a perspective can be derived from ecological studies and could be that of the global Leaf Economics Spectrum (LES) (Wright *et al.*, 2004; Osnas *et al.*, 2013). Essentially, the global LES describes a spectrum of coordinated biochemical, anatomical, and physiological properties that characterize the ecological leaf developmental and metabolic traits of different plant species. One extreme of the LES is represented by short-lived leaves of herbaceous annual plants, characterized by comparatively high photosynthetic rates but low resistance to stress factors over time, while the other is represented by the long-lived leaves of perennial plants, which are in turn resistant but show lower photosynthetic rates. Adoption of the LES concept has allowed for the description of the interdependence and trade-offs between leaf traits, including photosynthesis, between and within plant species (Marino *et al.*, 2010; Osnas *et al.*, 2013). As an example, greater amounts of cell wall have been postulated to be required to build long-lived leaves, which would reduce the efficiency of photosynthesis by lowering the fraction of leaf nitrogen invested in photosynthetic proteins and CO_2 diffusion rates through thicker and denser mesophyll cell walls (Onoda and Wright, 2018). However, inconsistencies in global LES trade-offs and relationships were reported when the principle was applied at smaller scales, such as restricted groups of plant species, individual plant species, or geographically-limited areas (Osnas *et al.*, 2018). These inconsistencies suggest that while the general idea of the short-lived, fast return leaf strategy opposed to the long-lived, slow return one holds true, different relationships between leaf traits can still result in these strategies (Niinemets, 2015; Xiong and Flexas, 2018), or as postulated by Anderegg *et al.* (2018), "there are still many ways to be a plant in most environments".

This is in line with what I described in **Chapter 5**, where I showed limited and weak correlations between photosynthetic LUE measured as ΦPSII and a set of presumably coordinated leaf anatomical and biochemical traits. The only significant and strong correlation to photosynthetic LUE was that of stomatal densities on adaxial and abaxial leaf sides, traits whose direct link to photosynthetic LUE is debated (Franks *et al.*, 2015; Schuler *et al.*, 2018; Harrison *et al.*, 2020). Beyond the lack of explanatory correlations, species showing the highest photosynthetic LUE did not consistently have thinner leaves than species with lower photosynthetic LUE. Thus, the lack of correlation between leaf thickness and photosynthetic LUE violates one of the key postulates of the global LES, stating that higher photosynthetic capacity is linked with lower leaf thickness. Another case of decoupling between leaf thickness and photosynthetic LUE in the Brassicaceae comes from the observations conducted in a related project (Retta *et al.*, in preparation), which showed that two species, *B. nigra* and *B. rapa*, can achieve similar photosynthetic capacity and thus LUE as *H. incana*, albeit via very different anatomical and biochemical strategies.

Taking a step away from physiological parameters, the work I presented in chapters 3 and 4 can be interpreted with the "many ways to be a plant" approach (Anderegg *et al.*, 2018). One finding co-authors and I reported in **Chapter 3** is that contrary to our original hypothesis, *H. incana* has, on average, a lower number of copies for genes involved in photosynthesis or photoprotection, despite having a higher photosynthetic LUE. This might be due to the strict filtering criteria we applied to pre-

dicted gene models in *H. incana* while annotating its genome, and will need to be confirmed by expanding the *H. incana* annotation with the use of larger RNA-Seq datasets. Interestingly however, *H. incana*, *B. nigra*, and *B. rapa* all had on average a higher number of copies of photosynthetic or photoprotective genes compared to *A. thaliana*, in line with their higher photosynthetic LUE. This hints at different photosynthesis-related genomic strategies being adopted by the three species. In **Chapter 4**, I focused on the expression profiles for genes directly involved in the photosynthetic reactions and highlighted a number of *H. incana*-specific gene expression patterns arising from irradiance-dependent expression regulation or constitutive gene expression. This, combined with the different transcriptional profiles observed for *B. nigra* and *B. rapa* but left to be studied in upcoming investigations, again hinted at the deployment of different transcriptional strategies, even at the level of the highly conserved photosynthetic genes, aimed at achieving high photosynthetic LUE by the three species.

The combined findings of Chapters 3 and 4, along with the ecological perspective of the global LES, offer valuable insights into the diverse strategies employed by *H. incana*, *B. nigra*, and *B. rapa* to achieve high photosynthetic LUE. These results challenge the notion of a direct correlation between leaf thickness and photosynthetic LUE and emphasize the complexity of the genetic and transcriptional mechanisms underlying photosynthesis in these species. While the hypothesis on the involvement of gene CNV in *H. incana*'s high photosynthetic LUE is partially supported, it is evident that other regulatory mechanisms and transcriptional strategies play crucial roles. Future research should delve deeper into the transcriptional strategies adopted by *H. incana*, and possibly other Brassicaceae species, to achieve and maintain high photosynthetic LUE under both optimal and challenging conditions. A logical next experimental step would therefore be a large-scale transcriptomics investigation, featuring species previously selected for their higher and lower photosynthetic LUE at high irradiance and utilizing a number of treatments that would expose plants to both optimal high-irradiance conditions and challenging ones (e.g., determined by the combination of high irradiance and water scarcity).

6.4 Photosynthesis beyond the leaf

This thesis, as the vast majority of photosynthesis research, was centered on the major photosynthetic organs of plants, i.e. leaves. However, future research aimed at improving plants' overall photosynthetic LUE should abandon this leaf-centric perspective for two main reasons.

First, a growing body of evidence is accumulating for the importance and contribution of photosynthesis carried out in non-foliar organs, such as stems and fruits, to plant productivity (Aschan and Pfanz, 2003; Simkin *et al.*, 2020; Lawson and Milliken, 2023). While the biochemical pathways of non-foliar photosynthesis are still to be fully understood (Henry *et al.*, 2020; Rangan *et al.*, 2022), the contribution of wheat ears' photosynthesis to overall assimilated carbon was shown to be relevant in modern cultivars and in challenging conditions (Tambussi *et al.*, 2021), similarly to

what is being described for rice panicles (Zhang *et al.*, 2022).

Second, ecophysiological reasons suggest that an approach focusing on more plant organs than just leaves will benefit future photosynthesis research. An extended formulation of the LES named the Plant Economics Spectrum was proposed to apply to whole plants (Reich, 2014; Shen *et al.*, 2019). "Fast" or "slow" strategies were therefore proposed to be applicable to whole plants, allowing establishing ecophysiological connections between their organ-specific traits (de la Riva *et al.*, 2016; Shen *et al.*, 2019). Somewhat similarly to what described above for the LES, the idea that plants could be classified along a single spectrum based on coordination between their traits has been shown to be less universal than originally proposed. Especially in the case of winter annual C_3 species, such as *H. incana* and the other Brassicaceae species studied in this thesis, a plant-wide decoupling between carbon acquisition, i.e. photosynthesis-related, traits and structural traits has been highlighted, albeit in controlled conditions (Kurze *et al.*, 2021). Interestingly, a similar decoupling has been reported as well for C_4 grasses grown in controlled conditions (Simpson *et al.*, 2020). Keeping this potential decoupling in mind, the Plant Economics Spectrum allows me to make hypotheses on traits of *H. incana* that were not studied within this thesis but could prove essential for future crop improvement strategies.

The traits I hypothesize playing an important role in enabling *H. incana* to have high photosynthetic LUE under challenging conditions relate to root architecture, uptake of nutrients, whole-plant water and photosynthate dynamics. In addition to presenting the decoupling between carbon acquisition and structural traits, Kurze *et al.* (2021) and Simpson *et al.* (2020) reported the coordination between root and leaf traits in non-woody species. Additionally, a substantial impact of root traits on plants' photosynthetic LUE is being reported (Reich *et al.*, 1998; Duan *et al.*, 2021; Kodama *et al.*, 2021; Miao *et al.*, 2023). Similarly, functional relatedness and coordination have been highlighted between photosynthesis and whole-plant water transport (Bucci *et al.*, 2019; Xu *et al.*, 2021). Based on these coordination patterns and a number of unpublished observations on the higher water-use efficiency of *H. incana*, I hypothesize that *H. incana*'s water dynamics are optimized to support its photosynthetic LUE, especially under water-limited conditions. An example of this would be a higher root:shoot ratio for *H. incana* than other Brassicaceae species, either constitutive or achievable via a large degree of plasticity for the ratio in relation to different environmental conditions. The latter has been reported for *Amaranthus palmeri*, an extremely successful weed characterized by the highest photosynthetic capacity reported to date (Ehleringer, 1983; Ward *et al.*, 2013). *A. palmeri*'s success was linked to its ability to outcompete crops for water (Berger *et al.*, 2015) and withstand drought (Chahal *et al.*, 2018). While a full description of the water dynamics of *A. palmeri* and their relationship with its high photosynthetic rates is still lacking (Cominelli and Patrignani, 2022), what was observed so far on this species strengthens my hypothesis on *H. incana*. Multi-organ water-related traits must play a role in the species' photosynthetic LUE, and especially in sustaining it in face of challenging environmental conditions, and their role should be further investigated. Similarly, I hypothesize an involvement in *H. incana*'s photosynthetic

LUE of whole-plant traits related to the uptake and utilization of key nutrients for photosynthesis, such as iron (Martín-Barranco *et al.*, 2021), or to "sink" traits, i.e. those concerning the transport, storage, and utilization of photosynthate, which have already been shown to play a role in photosynthetic LUE (McCormick *et al.*, 2006; Fabre *et al.*, 2020; Tejera-Nieves *et al.*, 2023). While clarifying the role of all these traits can appear as an overwhelming challenge, phenotyping platforms such as the NPEC Greenhouse module (https://www.npec.nl/phenotyping-modules/module-5-greenhouse-phenotyping) or the IPK PhenoSphere platform (Heuermann *et al.*, 2023) would allow to test a large number of plants of *H. incana* and relatives growing under high irradiance with differential water, nutrients, etc. treatments and quantify their photosynthetic LUE. This, in turn, would allow for a modelling effort considering the interdependence between these traits and photosynthetic LUE. The analysis of sink traits will likely be more difficult to perform, given the absence of high-throughput platforms allowing to measure them. X-ray Computed Tomography is emerging as the approach of choice for 3D imaging of the above- and below-ground organs of plants (Teramoto *et al.*, 2020; Piovesan *et al.*, 2021), and would allow for a first investigation aimed at identifying where the major sinks for photosynthate lie in *H. incana* and its relatives. While the integration of this technology is still limited in high-throughput phenotyping platforms, recent developments (Gerth *et al.*, 2021; Alle *et al.*, 2023) suggest that an experiment similar as what described above could be performed in the near future with focus on sink traits as well.

6.5 The value of a less crop-centric perspective

Besides including non-leaf traits in future plant photosynthesis research, our efforts towards improved crop productivity will benefit from a less crop-centric perspective as well. In **Chapter 2**, I compiled a list of studies that demonstrated the existence of natural genetic and physiological variation for a number of crop species. However, I argue that the relatively limited magnitude of differences reported for within-species variation does not align with the goal of substantially increasing crop productivity via photosynthesis. An alternative approach gaining traction in recent years involves screening crop wild relatives (CWRs) for interesting photosynthetic traits that could potentially be bred into elite crop varieties (McAusland *et al.*, 2020). CWRs have been defined as wild plant taxa that can find breeding use derived from their relatively close genetic relationship to a crop (Maxted *et al.*, 2006). They have been shown to harbor promising variation for photosynthesis traits such as enzymatic activity, stomatal dynamics, and leaf anatomical characteristics (Giuliani *et al.*, 2013; Prins *et al.*, 2016; McAusland *et al.*, 2020; Acevedo-Siaca *et al.*, 2021; Mathan *et al.*, 2021). Additionally, CWRs are recognized as valuable sources of traits directly linked with plant resilience and adaptation to biotic and abiotic stresses such as drought or salinity (Bohra *et al.*, 2022; Brozynska *et al.*, 2016; Renzi *et al.*, 2022; Satori *et al.*, 2022). Although these traits were often lost during breeding programs (Khoury *et al.*, 2022) because they could result in reduced yields under optimal conditions, they could contribute to increased photosynthetic LUE under challenging conditions, a much-

needed adaptation in light of the current rate of climate change (Tkemaladze and Makhashvili, 2016; Hussain *et al.*, 2021).

CWRs are genetically close to domesticated crops and, having their origin in the same areas as domesticated crops (Maxted and Vincent, 2021), were subjected to similar selection pressure during evolution. Therefore, CWRs are not likely to harbor the most radical photosynthetic solutions and adaptations that could benefit crop productivity. These solutions will likely be found in wild, non-model species that evolved in challenging environments. One example is the striking photosynthetic performance of some species growing as winter annuals in the Sonoran desert (Werk *et al.*, 1983), which appears to be coupled with a high degree of temperature tolerance (Downton *et al.*, 1984). Another example is *H. incana*, which, albeit native to warm-temperate areas coinciding with those of some key crops, has adapted to thrive on marginal lands, presenting it with similar challenges such as water and nutrient scarcity (Hart, 2001; Csikós and Tóth, 2023). Thus, a perspective shift in plant photosynthesis research that includes more wild species is likely to not only expand our understanding of the ecophysiological implications of the process, but also elucidate traits and strategies that will allow for breeding crops with higher, more resilient photosynthetic LUE and capacity.

6.6 Photosynthesis and yield research: where to now?

The assumption underlying the research presented in this thesis, and all research efforts on *H. incana*, is that understanding how the species achieves high photosynthetic LUE will contribute to the effort to increase the photosynthetic LUE of crops and consequently their yield potential. The idea that increasing photosynthetic LUE of crops can aid in increasing their yield potential is accepted across the scientific community, and supported by a recent life cycle assessment (LCA) analyzing both solar-derived energy and human-supplemented energy involved in the production of potato crops. This analysis highlighted how energy losses associated with photosynthetic inefficiencies are roughly eighty times greater than the energy supplemented with agricultural practices (Roney and Walker, 2023), suggesting that photosynthesis research has the potential to increase crop productivity beyond what can be achieved by any effort based on increasing supplements such as fertilizers. An improvement in photosynthetic LUE during plant evolution and domestication has been recently reported as the result of an ecophysiological investigation, raising hopes for further potential increases (Huang *et al.*, 2022). However, a number of studies reported that efforts carried out to increase the yield of crops by increasing their photosynthetic LUE have so far been essentially unsuccessful (Flexas, 2016; Sinclair *et al.*, 2019; Araus *et al.*, 2021; Sinclair *et al.*, 2023). The current lack of generalized success does not represent a reason to abandon photosynthesis research aimed at increasing crop yields, also considering the substantial measurement errors that can be associated with biomass and yield measurements which could hide small, yet significant, improvements (Monteith, 1994; Arslan and Colvin, 2002; Lyle *et al.*, 2014; Gollin and Udry, 2021). However, I will argue that this lack of success prompts in-

creased caution in designing and communicating future photosynthesis research and photosynthesis-oriented crop breeding efforts.

Crop yield has been proposed to be a super-complex trait (Chang *et al.*, 2019), and could even be defined as an omnigenic trait, i.e. one driven by very large numbers of (genetic) variants all having very small effects (Boyle *et al.*, 2017). Despite being responsible for a major part of energy losses in a crop, the complex "photosynthesis" trait is thus only one contributor to the intricate network of traits and environmental factors determining crop yield. It is therefore hard to imagine that single modifications to specific bottlenecks in the photosynthetic process will automatically result in increased yield for different crops, in different agricultural settings, or even in different cropping seasons (Araus *et al.*, 2021). It is much more reasonable to think of photosynthesis as one important tool that can help in achieving higher crop productivity, rather than the single process that will enable it.

Quite some efforts, some of which were presented in **Chapters 3** and **4** of this thesis, have been directed towards revealing the "genetic strategy" underlying photosynthetic LUE, and ways to improve it (Zhu *et al.*, 2007; van Bezouw *et al.*, 2019; Theeuwen *et al.*, 2022). In my opinion, however, two caveats apply to the idea of a genetic strategy for photosynthetic LUE. The first is that it tempts researchers to describe variation for photosynthetic LUE in reductionist terms, thus prompting "one trait" approaches in which a single plant trait contributing to photosynthetic LUE and its genetic determinants is elevated to the rank of the main determinant of photosynthetic LUE. The second caveat is that the idea of a genetic strategy for photosynthetic LUE can make researchers disregard the impact that environmental factors have on it. Indeed, the strategy for photosynthetic LUE should include a vast set of environmental determinants next to the genetic ones. As discussed in detail in the previous sections of this chapter, this implies the existence of many different strategies for high photosynthetic LUE, something which could make the use of this trait for crop yield increase seem like a hopeless task.

However, an approach that has higher chances of allowing the development of crops with high photosynthetic LUE and yield has been proposed and is being increasingly adopted. This approach is based on the idea of employing systems modeling to describe the relationship between photosynthesis, environmental factors, and yield (Lammerts van Bueren *et al.*, 2018; Verma *et al.*, 2021). Systems models of plants aim to describe and simulate quantitatively the behavior of a number of plant traits at scales ranging from single cells to whole ecosystems (Salvatori *et al.*, 2022). Such an approach appears to be promising for developing a coherent view of the impact of the many photosynthetic LUE strategies on plant productivity and highlighting specific combinations of environmental and genetic factors that will increase crop yields (Chang *et al.*, 2019). To create comprehensive models of yield as a function of photosynthetic LUE, a more complex and ecophysiological view of the processes underlying the latter will be required. To ensure this, the largely leaf-centric perspective adopted by many photosynthesis researchers should be surpassed. This will be supported by the increasing availability and accessibility of high-throughput phenotyping technologies, which are expanding the range of whole-plant physiological parameters that can be measured and modeled (van Bezouw *et al.*, 2019; Sharwood

et al., 2022). Furthermore, the responses of physiological traits to different, fluctuating environmental conditions closer to those experienced by crops growing in field environments should be better understood. The majority of current photosynthesis studies on a plant scale (as opposed to ecosystem-level studies, which are beyond the scope of this Discussion) have been conducted on plants grown in controlled, lowly variable conditions that do not mirror real-world conditions encountered by crops (Poorter *et al.*, 2016). The so-called lab-to-field transition of studies and results on photosynthetic LUE and other yield-determining traits is not trivial (Kuijken *et al.*, 2015; Rouphael *et al.*, 2018; Li *et al.*, 2021; Fu *et al.*, 2022). However, different methodologies are being developed based on the precise simulation of field-like conditions in large-scale controlled-environment facilities (Arend *et al.*, 2022; Langstroff *et al.*, 2022) or the mathematical prediction of traits' responses based on genetic and environmental information coupled with data collected in controlled conditions (Tardieu *et al.*, 2023).

Despite this proposed paradigm shift, objections to the feasibility of improving yield via photosynthesis could still be raised. Weiner (2019) argued that attempts to increase crop yields by "improving" photosynthesis may have failed so far because researchers and breeders have been trying to improve a template for photosynthesis that has already been optimized by evolution, possessing a force that is orders of magnitude greater than that of current scientific research. According to Weiner, "the idea that plant researchers can develop new mechanisms that nature has not found seems unlikely". Sinclair *et al.* (2019) argued that nitrogen availability poses a much more prominent constraint on crop yield than photosynthetic LUE, suggesting a much larger focus on the first. In my view, both opinions do not support a case for reducing the effort on linking photosynthetic LUE to crop yield, but rather support the idea of doing so by adopting more complex perspectives encompassing the interdependence of photosynthetic and other key plant traits and the plethora of solutions evolved in the natural world to achieve higher photosynthetic LUE and capacity, if not in absolute terms, at least in specific and often challenging conditions.

6.7 Conclusions

The research presented in this thesis initially aligned with the idea that a genetic strategy could be described for *H. incana* in a relatively straightforward manner, providing a template for increasing the photosynthetic LUE of crops. However, the findings presented in Chapters 3, 4, and 5 paint a more complex picture than originally hypothesized. In **Chapter 3**, I reported that while CNV correlated with higher photosynthetic LUE, other genetic factors might be involved in *H. incana*'s superior performance, as this species did not show a substantial increase in gene copies when compared to *B. nigra* and *B. rapa*. In **Chapter 4**, I focused on *H. incana* and revealed the species' adoption of unique transcriptional strategies under high irradiance, while also reporting diverging transcriptional profiles for *B. nigra* and *B. rapa*, which may contribute to different photosynthetic strategies. Additionally, **Chapter 5** highlighted different strategies, showing a lack of correlation between high photosyn-

thetic LUE and a series of anatomical and biochemical leaf parameters.

This thesis contributes essential genomic, transcriptomic, and phenomic resources, along with a set of novel approaches and methods that will support efforts to understand, and potentially enhance, high photosynthetic LUE. As argued in the previous sections, future research on photosynthetic LUE should embrace the complexity of this trait and expand experimental horizons beyond the current leaf- and optimal conditions-centric approaches. The results presented for *H. incana* are highly relevant in the context of a more holistic view of photosynthetic LUE. There is, therefore, still ample scope to work with *Hirschfeldia incana* and its relatives and delve deeper into the different mechanisms of high photosynthetic LUE in the Brassicaceae family.

Therefore, additional research on *H. incana* holds the potential to provide the first comprehensive description of the strategy adopted by a species to sustain high photosynthetic LUE. This description will, in turn, prompt for further investigations in other high-LUE species, and will provide an important template to inspire breeding projects aiming to increase the photosynthetic LUE of crops and realise its potential to increase their yields, Future studies hinged on *H. incana* should focus on delivering a robust and comprehensive description of photosynthetic LUE for this species and other selected Brassicaceae species grown under more challenging conditions than those employed in this thesis (e.g., drought, nutrient limitations, and higher temperatures). The differences observed will then need to be explained in terms of whole-plant ecophysiological trait trade-offs, where modern high-throughput phenotyping technologies will prove invaluable. Subsequently, the genetic background of the most explanatory traits should be investigated by determining patterns in gene expression and transcription translation, thus integrating genomics, transcriptomics, and proteomics. All this data could then be integrated through systems modelling, providing an algorithm for high photosynthetic LUE in the Brassicaceae family under various environmental conditions. Such an integrated set of data will undoubtedly be a fundamental resource in advancing our understanding of photosynthetic LUE and capacity, and could potentially direct future yield-oriented plant breeding approaches.

References

Acevedo-Siaca, L. G., Dionora, J., Laza, R., Paul Quick, W., and Long, S. P. (2021). Dynamics of photosynthetic induction and relaxation within the canopy of rice and two wild relatives. *Food and Energy Security*, **10**(3), e286.

Ahanger, M. A., Qin, C., Begum, N., Maodong, Q., Dong, X. X., El-Esawi, M., El-Sheikh, M. A., Alatar, A. A., and Zhang, L. (2019). Nitrogen availability prevents oxidative effects of salinity on wheat growth and photosynthesis by up-regulating the antioxidants and osmolytes metabolism, and secondary metabolite accumulation. *BMC Plant Biology*, **19**(1), 479.

Ahanger, M. A., Qi, M., Huang, Z., Xu, X., Begum, N., Qin, C., Zhang, C., Ahmad, N., Mustafa, N. S., Ashraf, M., and Zhang, L. (2021). Improving growth and photosynthetic performance of drought stressed tomato by application of nano-organic fertilizer involves up-regulation of nitrogen, antioxidant and osmolyte metabolism. *Ecotoxicology and Environmental Safety*, **216**, 112195.

Alle, J., Gruber, R., Wörlein, N., Uhlmann, N., Claußen, J., Wittenberg, T., and Gerth, S. (2023). 3D segmentation of plant root systems using spatial pyramid pooling and locally adaptive field-of-view inference. *Frontiers in Plant Science*, **14**.

Anderegg, L. D. L., Berner, L. T., Badgley, G., Sethi, M. L., Law, B. E., and HilleRisLambers, J. (2018). Within-species patterns challenge our understanding of the leaf economics spectrum. *Ecology Letters*, **21**(5), 734–744.

Araus, J. L., Sanchez-Bragado, R., and Vicente, R. (2021). Improving crop yield and resilience through optimization of photosynthesis: Panacea or pipe dream? *Journal of Experimental Botany*, **72**(11), 3936–3955.

Arend, D., Psaroudakis, D., Memon, J. A., Rey-Mazón, E., Schüler, D., Szymanski, J. J., Scholz, U., Junker, A., and Lange, M. (2022). From data to knowledge – big data needs stewardship, a plant phenomics perspective. *The Plant Journal*, **111**(2), 335–347.

Arslan, S. and Colvin, T. S. (2002). Grain Yield Mapping: Yield Sensing, Yield Reconstruction, and Errors. *Precision Agriculture*, **3**(2), 135–154.

Aschan, G. and Pfanz, H. (2003). Non-foliar photosynthesis – a strategy of additional carbon acquisition. *Flora - Morphology, Distribution, Functional Ecology of Plants*, **198**(2), 81–97.

Berger, S. T., Ferrell, J. A., Rowland, D. L., and Webster, T. M. (2015). Palmer Amaranth (Amaranthus palmeri) Competition for Water in Cotton. *Weed Science*, **63**(4), 928–935.

Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., and Varshney, R. K. (2022). Reap the crop wild relatives for breeding future crops. *Trends in Biotechnology*, **40**(4), 412–431.

Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, **169**(7), 1177–1186.

Brozynska, M., Furtado, A., and Henry, R. J. (2016). Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnology Journal*, **14**(4), 1070–1085.

Bucci, S. J., Carbonell Silletta, L. M., Garré, A., Cavallaro, A., Efron, S. T., Arias, N. S., Goldstein, G., and Scholz, F. G. (2019). Functional relationships between hydraulic traits and the timing of diurnal depression of photosynthesis. *Plant, Cell & Environment*, **42**(5), 1603–1614.

Canvin, D. T., Berry, J. A., Badger, M. R., Fock, H., and Osmond, C. B. (1980). Oxygen exchange in leaves in the light. *Plant Physiology*, **66**(2), 302–307.

Chahal, P. S., Irmak, S., Jugulam, M., and Jhala, A. J. (2018). Evaluating Effect of Degree of Water Stress on Growth and Fecundity of Palmer amaranth (Amaranthus palmeri) Using Soil Moisture Sensors. *Weed Science*, **66**(6), 738–745.

Chang, T.-G., Chang, S., Song, Q.-F., Perveen, S., and Zhu, X.-G. (2019). Systems models, phenomics and genomics: Three pillars for developing high-yielding photosynthetically efficient crops. *in silico Plants*, **1**(1), diy003.

Cominelli, S. and Patrignani, A. (2022). Transpiration response of palmer amaranth (Amaranthus palmeri) to drying soil in greenhouse conditions. *Frontiers in Agronomy*, **4**.

Csikós, N. and Tóth, G. (2023). Concepts of agricultural marginal lands and their utilisation: A review. *Agricultural Systems*, **204**, 103560.

de la Riva, E. G., Tosto, A., Pérez-Ramos, I. M., Navarro-Fernández, C. M., Olmo, M., Anten, N. P., Marañón, T., and Villar, R. (2016). A plant economics spectrum in Mediterranean forests along environmental gradients: Is there coordination among leaf, stem and root traits? *Journal of Vegetation Science*, **27**(1), 187–199.

Downton, W. J. S., Berry, J. A., and Seemann, J. R. (1984). Tolerance of Photosynthesis to High Temperature in Desert Plants 1. *Plant Physiology*, **74**(4), 786–790.

Duan, L., Pérez-Ruiz, J. M., Cejudo, F. J., and Dinneny, J. R. (2021). Characterization of CYCLOPHILLIN38 shows that a photosynthesis-derived systemic signal controls lateral root emergence. *Plant Physiology*, **185**(2), 503–518.

Ehleringer, J. (1983). Ecophysiology of Amaranthus palmeri, a sonoran desert summer annual. *Oecologia*, **57**(1), 107–112.

Fabre, D., Dingkuhn, M., Yin, X., Clément-Vidal, A., Roques, S., Soutiras, A., and Luquet, D. (2020). Genotypic variation in source and sink traits affects the response of photosynthesis and growth to elevated atmospheric CO2. *Plant, Cell & Environment*, **43**(3), 579–593.

Flexas, J. (2016). Genetic improvement of leaf photosynthesis and intrinsic water use efficiency in C3 plants: Why so much little success? *Plant Science*, **251**, 155–161.

Flexas, J., Diaz-Espejo, A., Gago, J., Gallé, A., Galmés, J., Gulías, J., and Medrano, H. (2014). Photosynthetic limitations in Mediterranean plants: A review. *Environmental and Experimental Botany*, **103**, 12–23.

Franks, P. J., W. Doheny-Adams, T., Britton-Harper, Z. J., and Gray, J. E. (2015). Increasing water-use efficiency directly through genetic manipulation of stomatal density. *New Phytologist*, **207**(1), 188–195.

Fu, P., Montes, C. M., Siebers, M. H., Gomez-Casanovas, N., McGrath, J. M., Ainsworth, E. A., and Bernacchi, C. J. (2022). Advances in field-based high-throughput photosynthetic phenotyping. *Journal of Experimental Botany*, **73**(10), 3157–3172.

Galmés, J., Ochogavía, J. M., Gago, J., Roldán, E. J., Cifre, J., and Conesa, M. À. (2013). Leaf responses to drought stress in Mediterranean accessions of Solanum lycopersicum: Anatomical adaptations in relation to gas exchange parameters. *Plant, Cell & Environment*, **36**(5), 920–935.

Gan, L., Zhang, C., Yin, Y., Lin, Z., Huang, Y., Xiang, J., Fu, C., and Li, M. (2013). Anatomical adaptations of the xerophilous medicinal plant, Capparis spinosa, to drought conditions. *Horticulture, Environment, and Biotechnology*, **54**(2), 156–161.

Gerth, S., Claußen, J., Eggert, A., Wörlein, N., Wininger, M., Wittenberg, T., and Uhlmann, N. (2021). Semiautomated 3D Root Segmentation and Evaluation Based on X-Ray CT Imagery. *Plant Phenomics*, **2021**.

Giuliani, R., Koteyeva, N., Voznesenskaya, E., Evans, M. A., Cousins, A. B., and Edwards, G. E. (2013). Coordination of Leaf Photosynthesis, Transpiration, and Structural Traits in Rice and Wild Relatives (Genus Oryza). *Plant Physiology*, **162**(3), 1632–1651.

Glazier, A. M., Nadeau, J. H., and Aitman, T. J. (2002). Finding Genes That Underlie Complex Traits. *Science*, **298**(5602), 2345–2349.

Gollin, D. and Udry, C. (2021). Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture. *Journal of Political Economy*, **129**(1), 1–80.

Harrison, E. L., Arce Cubas, L., Gray, J. E., and Hepworth, C. (2020). The influence of stomatal morphology and distribution on photosynthetic gas exchange. *The Plant Journal*, **101**(4), 768–779.

Hart, J. F. (2001). Half a Century of Cropland Change*. *Geographical Review*, **91**(3), 525–543.

Henry, R. J., Furtado, A., and Rangan, P. (2020). Pathways of Photosynthesis in Non-Leaf Tissues. *Biology*, **9**(12), 438.

Heuermann, M. C., Knoch, D., Junker, A., and Altmann, T. (2023). Natural plant growth and development achieved in the IPK PhenoSphere by dynamic environment simulation. *Nature Communications*, **14**(1), 5783.

Holland, J. B. (2007). Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, **10**(2), 156–161.

Huang, G., Peng, S., and Li, Y. (2022). Variation of photosynthesis during plant evolution and domestication: Implications for improving crop photosynthesis. *Journal of Experimental Botany*, **73**(14), 4886–4896.

Hussain, S., Ulhassan, Z., Brestic, M., Zivcak, M., Weijun Zhou, Allakhverdiev, S. I., Yang, X., Safdar, M. E., Yang, W., and Liu, W. (2021). Photosynthesis research under climate change. *Photosynthesis Research*, **150**(1), 5–19.

Khan, R., Ma, X., Hussain, Q., Chen, K., Farooq, S., Asim, M., Ren, X., Shah, S., and Shi, Y. (2023). Transcriptome and anatomical studies reveal alterations in leaf thickness under long-term drought stress in tobacco. *Journal of Plant Physiology*, **281**, 153920.

Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., de Haan, S., Engels, J. M. M., Guarino, L., Hoban, S., Mercer, K. L., Miller, A. J., Nabhan, G. P., Perales, H. R., Richards, C., Riggins, C., and Thormann, I. (2022). Crop genetic erosion: Understanding and responding to loss of crop diversity. *New Phytologist*, **233**(1), 84–118.

Kodama, A., Watanabe, T., Yamaguchi, M., Narita, R., Katsuhara, M., Sato, K., Ookawa, T., and Hirasawa, T. (2021). Accession difference in leaf photosynthesis, root hydraulic conductance and gene expression of root aquaporins under salt stress in barley seedlings. *Plant Production Science*, **24**(1), 73–82.

Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences*, **279**(1749), 5048–5057.

Kuijken, R. C., van Eeuwijk, Fred. A., Marcelis, L. F., and Bouwmeester, H. J. (2015). Root phenotyping: From component trait in the lab to breeding. *Journal of Experimental Botany*, **66**(18), 5389–5401.

Kurze, S., Engelbrecht, B. M. J., Bilton, M. C., Tielbörger, K., and Álvarez-Cansino, L. (2021). Rethinking the Plant Economics Spectrum for Annuals: A Multi-Species Study. *Frontiers in Plant Science*, **12**.

Lammerts van Bueren, E. T., Struik, P. C., van Eekeren, N., and Nuijten, E. (2018). Towards resilience through systems-based plant breeding. A review. *Agronomy for Sustainable Development*, **38**(5), 42.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.-F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H.-C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G. R., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F. A., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S.-P., Yeh, R.-F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A.,

Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Patrinos, A., Morgan, M. J., International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, C. f. G. R., The Sanger Centre:, Washington University Genome Sequencing Center, US DOE Joint Genome Institute:, Baylor College of Medicine Human Genome Sequencing Center:, RIKEN Genomic Sciences Center:, Genoscope and CNRS UMR-8030:, Department of Genome Analysis, I. o. M. B., GTC Sequencing Center:, Beijing Genomics Institute/Human Genome Center:, Multimegabase Sequencing Center, T. I. f. S. B., Stanford Genome Technology Center:, University of Oklahoma's Advanced Center for Genome Technology:, Max Planck Institute for Molecular Genetics:, Cold Spring Harbor Laboratory, L. A. H. G. C., GBF—German Research Centre for Biotechnology:, includes individuals listed under other headings):*Genome Analysis Group (listed in alphabetical order, a., Scientific management: National Human Genome Research Institute, U. N. I. o. H., Stanford Human Genome Center:, University of Washington Genome Center:, Department of Molecular Biology, K. U. S. o. M., University of Texas Southwestern Medical Center at Dallas:, Office of Science, U. D. o. E., and The Wellcome Trust: (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**(6822), 860–921.

Langstroff, A., Heuermann, M. C., Stahl, A., and Junker, A. (2022). Opportunities and limits of controlled-environment plant phenotyping for climate response traits. *Theoretical and Applied Genetics*, **135**(1), 1–16.

Lawson, T. and Milliken, A. L. (2023). Photosynthesis – beyond the leaf. *New Phytologist*, **238**(1), 55–61.

Levizou, E., Drilias, P., and Kyparissis, A. (2004). Exceptional Photosynthetic Performance of Capparis spinosa L. Under Adverse Conditions of Mediterranean summer. *Photosynthetica*, **42**(2), 229–235.

Li, D., Quan, C., Song, Z., Li, X., Yu, G., Li, C., and Muhammad, A. (2021). High-Throughput Plant Phenotyping Platform (HT3P) as a Novel Tool for Estimating Agronomic Traits From the Lab to the Field. *Frontiers in Bioengineering and Biotechnology*, **8**.

Llewellyn, D., Golem, S., Foley, E., Dinka, S., Jones, A. M. P., and Zheng, Y. (2022). Indoor grown cannabis yield increased proportionally with light intensity, but ultraviolet radiation did not affect yield or cannabinoid content. *Frontiers in Plant Science*, **13**.

Lyle, G., Bryan, B. A., and Ostendorf, B. (2014). Post-processing methods to eliminate erroneous grain yield measurements: Review and directions for future development. *Precision Agriculture*, **15**(4), 377–402.

Marino, G., Aqil, M., and Shipley, B. (2010). The leaf economics spectrum and the prediction of photosynthetic light–response curves. *Functional Ecology*, **24**(2), 263–272.

Martín-Barranco, A., Thomine, S., Vert, G., and Zelazny, E. (2021). A quick journey into the diversity of iron uptake strategies in photosynthetic organisms. *Plant Signaling & Behavior*, **16**(11), 1975088.

Mathan, J., Singh, A., Jathar, V., and Ranjan, A. (2021). High photosynthesis rate in two wild rice species is driven by leaf anatomy mediating high Rubisco activity and electron transport rate. *Journal of Experimental Botany*, **72**(20), 7119–7135.

Maxted, N. and Vincent, H. (2021). Review of congruence between global crop wild relative hotspots and centres of crop origin/diversity. *Genetic Resources and Crop Evolution*, **68**(4), 1283–1297.

Maxted, N., Ford-Lloyd, B. V., Jury, S., Kell, S., and Scholten, M. (2006). Towards a definition of a crop wild relative. *Biodiversity & Conservation*, **15**(8), 2673–2685.

McAusland, L., Vialet-Chabrand, S., Jauregui, I., Burridge, A., Hubbart-Edwards, S., Fryer, M. J., King, I. P., King, J., Pyke, K., Edwards, K. J., Carmo-Silva, E., Lawson, T., and Murchie, E. H. (2020). Variation in key leaf photosynthetic traits across wheat wild relatives is accession dependent not species dependent. *New Phytologist*, **228**(6), 1767–1780.

McCormick, A. J., Cramer, M. D., and Watt, D. A. (2006). Sink strength regulates photosynthesis in sugarcane. *New Phytologist*, **171**(4), 759–770.

Miao, Y., Ren, J., Zhang, Y., Chen, X., Qi, M., Li, T., Zhang, G., and Liu, Y. (2023). Effect of low root-zone temperature on photosynthesis, root structure and mineral element absorption of tomato seedlings. *Scientia Horticulturae*, **315**, 111956.

Moher, M., Llewellyn, D., Jones, M., and Zheng, Y. (2022). Light intensity can be used to modify the growth and morphological characteristics of cannabis during the vegetative stage of indoor production. *Industrial Crops and Products*, **183**, 114909.

Monteith, J. L. (1994). Validity of the correlation between intercepted radiation and biomass. *Agricultural and Forest Meteorology*, **68**(3), 213–220.

Niinemets, Ü. (2015). Is there a species spectrum within the world-wide leaf economics spectrum? Major variations in leaf functional traits in the Mediterranean sclerophyll Quercus ilex. *New Phytologist*, **205**(1), 79–96.

Onoda, Y. and Wright, I. J. (2018). The Leaf Economics Spectrum and its Underlying Physiological and Anatomical Principles. In W. W. Adams III and I. Terashima, editors, *The Leaf: A Platform for Performing Photosynthesis*, Advances in Photosynthesis and Respiration, pages 451–471. Springer International Publishing, Cham.

Osnas, J. L. D., Lichstein, J. W., Reich, P. B., and Pacala, S. W. (2013). Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum. *Science*, **340**(6133), 741–744.

Osnas, J. L. D., Katabuchi, M., Kitajima, K., Wright, S. J., Reich, P. B., Van Bael, S. A., Kraft, N. J. B., Samaniego, M. J., Pacala, S. W., and Lichstein, J. W. (2018). Divergent drivers of leaf trait variation within species, among species, and among functional groups. *Proceedings of the National Academy of Sciences*, **115**(21), 5480–5485.

Parekh, S., Vieth, B., Ziegenhain, C., Enard, W., and Hellmann, I. (2018). Strategies for quantitative RNA-seq analyses among closely related species.

Paterson, A. H. (1998). *Molecular Dissection of Complex Traits*. CRC press, Boca Raton.

Patterson, E. L., Pettinga, D. J., Ravet, K., Neve, P., and Gaines, T. A. (2018). Glyphosate Resistance and EPSPS Gene Duplication: Convergent Evolution in Multiple Plant Species. *Journal of Heredity*, **109**(2), 117–125.

Piovesan, A., Vancauwenberghe, V., Looverbosch, T. V. D., Verboven, P., and Nicolaï, B. (2021). X-ray computed tomography for 3D plant imaging. *Trends in Plant Science*, **26**(11), 1171–1185.

Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W. H., Kleyer, M., Schurr, U., and Postma, J. (2016). Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist*, **212**(4), 838–855.

Prins, A., Orr, D. J., Andralojc, P. J., Reynolds, M. P., Carmo-Silva, E., and Parry, M. A. J. (2016). Rubisco catalytic properties of wild and domesticated relatives provide scope for improving wheat photosynthesis. *Journal of Experimental Botany*, **67**(6), 1827–1838.

Rangan, P., Wankhede, D. P., Subramani, R., Chinnusamy, V., Malik, S. K., Baig, M. J., Singh, K., and Henry, R. (2022). Evolution of an intermediate C4 photosynthesis in the non-foliar tissues of the Poaceae. *Photosynthesis Research*, **153**(3), 125–134.

Reich, P. B. (2014). The world-wide 'fast–slow' plant economics spectrum: A traits manifesto. *Journal of Ecology*, **102**(2), 275–301.

Reich, P. B., Walters, M. B., Tjoelker, M. G., Vanderklein, D., and Buschena, C. (1998). Photosynthesis and respiration rates depend on leaf and root morphology and nitrogen concentration in nine boreal tree species differing in relative growth rate. *Functional Ecology*, **12**(3), 395–405.

Renzi, J. P., Coyne, C. J., Berger, J., von Wettberg, E., Nelson, M., Ureta, S., Hernández, F., Smýkal, P., and Brus, J. (2022). How Could the Use of Crop Wild Relatives in Breeding Increase the Adaptation of Crops to Marginal Environments? *Frontiers in Plant Science*, **13**.

Rhizopolou, S. and Psaras, G. K. (2003). Development and Structure of Drought-tolerant Leaves of the Mediterranean Shrub Capparis spinosa L. *Annals of Botany*, **92**(3), 377–383.

Ritchie, R. J. (2010). Modelling photosynthetic photon flux density and maximum potential gross photosynthesis. *Photosynthetica*, **48**(4), 596–609.

Rivarola Sena, A. C., Andres-Robin, A., Vialette, A. C., Just, J., Launay-Avon, A., Borrega, N., Dubreucq, B., and Scutt, C. P. (2022). Custom methods to identify conserved genetic modules applied to novel transcriptomic data from Amborella trichopoda. *Journal of Experimental Botany*, **73**(8), 2487–2498.

Rodriguez-Morrison, V., Llewellyn, D., and Zheng, Y. (2021). Cannabis Yield, Potency, and Leaf Photosynthesis Respond Differently to Increasing Light Levels in an Indoor Environment. *Frontiers in Plant Science*, **12**.

Roney, H. E. and Walker, B. J. (2023). Photon to plate: A holistic view of photosynthetic and anthropogenic energy fluxes. *Food and Energy Security*, **12**(1), e348.

Rouphael, Y., Spíchal, L., Panzarová, K., Casa, R., and Colla, G. (2018). High-Throughput Plant Phenotyping for Developing Novel Biostimulants: From Lab to Field or From Field to Lab? *Frontiers in Plant Science*, **9**, 408288.

Roux, J., Rosikiewicz, M., and Robinson-Rechavi, M. (2015). What to compare and how: Comparative transcriptomics for Evo-Devo. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **324**(4), 372–382.

Saini, D. K., Chopra, Y., Singh, J., Sandhu, K. S., Kumar, A., Bazzer, S., and Srivastava, P. (2021). Comprehensive evaluation of mapping complex traits in wheat using genome-wide association studies. *Molecular Breeding*, **42**(1), 1.

Salvatori, N., Carteni, F., Giannino, F., Alberti, G., Mazzoleni, S., and Peressotti, A. (2022). A System Dynamics Approach to Model Photosynthesis at Leaf Level Under Fluctuating Light. *Frontiers in Plant Science*, **12**.

Satori, D., Tovar, C., Faruk, A., Hammond Hunt, E., Muller, G., Cockel, C., Kühn, N., Leitch, I. J., Lulekal, E., Pereira, L., Ryan, P., Willis, K. J., and Pironon, S. (2022). Prioritising crop wild relatives to enhance agricultural resilience in sub-Saharan Africa under climate change. *PLANTS, PEOPLE, PLANET*, **4**(3), 269–282.

Schuler, M. L., Sedelnikova, O. V., Walker, B. J., Westhoff, P., and Langdale, J. A. (2018). SHORTROOT-Mediated Increase in Stomatal Density Has No Impact on Photosynthetic Efficiency. *Plant Physiology*, **176**(1), 757–772.

Sharwood, R. E., Quick, W. P., Sargent, D., Estavillo, G. M., Silva-Perez, V., and Furbank, R. T. (2022). Mining for allelic gold: Finding genetic variation in photosynthetic traits in crops and wild relatives. *Journal of Experimental Botany*, **73**(10), 3085–3108.

Shen, Y., Gilbert, G. S., Li, W., Fang, M., Lu, H., and Yu, S. (2019). Linking Aboveground Traits to Root Traits and Local Environment: Implications of the Plant Economics Spectrum. *Frontiers in Plant Science*, **10**.

Simkin, A. J., Faralli, M., Ramamoorthy, S., and Lawson, T. (2020). Photosynthesis in non-foliar tissues: Implications for yield. *The Plant Journal*, **101**(4), 1001–1015.

Simpson, K. J., Bennett, C., Atkinson, R. R. L., Mockford, E. J., McKenzie, S., Freckleton, R. P., Thompson, K., Rees, M., and Osborne, C. P. (2020). C4 photosynthesis and the economic spectra of leaf and root traits independently influence growth rates in grasses. *Journal of Ecology*, **108**(5), 1899–1909.

Sinclair, T., Specht, J., Cassman, K., Purcell, L., and Rufty, T. (2023). Comment on "Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection". *Science*, **379**(6634), eade8506.

Sinclair, T. R., Rufty, T. W., and Lewis, R. S. (2019). Increasing Photosynthesis: Unlikely Solution For World Food Problem. *Trends in Plant Science*, **24**(11), 1032–1039.

Tambussi, E. A., Maydup, M. L., Carrión, C. A., Guiamet, J. J., and Araus, J. L. (2021). Ear photosynthesis in C3 cereals and its contribution to grain yield: Methodologies, controversies, and perspectives. *Journal of Experimental Botany*, **72**(11), 3956–3970.

Tardieu, F., Bouidghaghen, J., Moreau, L., Beauchene, K., Chapuis, R., Mangel, N., Bosquet, L. C., Welcker, C., and Bogard, M. (2023). Robotized indoor phenotyping allows genomic prediction of adaptive traits in the field. *Research Square Preprint*.

Tejera-Nieves, M., Abraha, M., Chen, J., Hamilton, S. K., Robertson, G. P., and Walker, B. J. (2023). Seasonal decline in leaf photosynthesis in perennial switchgrass explained by sink limitations and water deficit. *Frontiers in Plant Science*, **13**.

Teramoto, S., Takayasu, S., Kitomi, Y., Arai-Sanoh, Y., Tanabata, T., and Uga, Y. (2020). High-throughput three-dimensional visualization of root system architecture of rice using X-ray computed tomography. *Plant Methods*, **16**(1), 66.

Theeuwen, T. P. J. M., Logie, L. L., Harbinson, J., and Aarts, M. G. M. (2022). Genetics as a key to improving crop photosynthesis. *Journal of Experimental Botany*, **73**(10), 3122–3137.

Tipirdamaz, R., Gagneul, D., Duhazé, C., Aïnouche, A., Monnier, C., Özkum, D., and Larher, F. (2006). Clustering of halophytes from an inland salt marsh in Turkey according to their ability to accumulate sodium and nitrogenous osmolytes. *Environmental and Experimental Botany*, **57**(1), 139–153.

Tkemaladze, G. S. and Makhashvili, K. A. (2016). Climate changes and photosynthesis. *Annals of Agrarian Science*, **14**(2), 119–126.

van Bezouw, R. F. H. M., Keurentjes, J. J. B., Harbinson, J., and Aarts, M. G. M. (2019). Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. *The Plant Journal*, **97**(1), 112–133.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V. D., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R.-R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z. Y., Wang, A., Wang, X., Wang, J., Wei, M.-H., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S. C., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.-H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigó, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.-H., Coyne, M., Dahlke, C., Mays, A. D., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001). The Sequence of the Human Genome. *Science*, **291**(5507), 1304–1351.

Verma, V., Vishal, B., Kohli, A., and Kumar, P. P. (2021). Systems-based rice improvement approaches for sustainable food and nutritional security. *Plant Cell Reports*, **40**(11), 2021–2036.

Ward, S. M., Webster, T. M., and Steckel, L. E. (2013). Palmer Amaranth (Amaranthus palmeri): A Review. *Weed Technology*, **27**(1), 12–27.

Weiner, J. (2019). Looking in the Wrong Direction for Higher-Yielding Crop Genotypes. *Trends in Plant Science*, **24**(10), 927–933.

Werk, K. S., Ehleringer, J., Forseth, I. N., and Cook, C. S. (1983). Photosynthetic characteristics of Sonoran Desert winter annuals. *Oecologia*, **59**(1), 101–105.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., and Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, **428**(6985), 821–827.

Xiong, D. and Flexas, J. (2018). Leaf economics spectrum in rice: Leaf anatomical, biochemical, and physiological trait trade-offs. *Journal of Experimental Botany*, **69**(22), 5599–5609.

Xu, H., Wang, H., Prentice, I. C., Harrison, S. P., and Wright, I. J. (2021). Coordination of plant hydraulic and photosynthetic traits: Confronting optimality theory with field measurements. *New Phytologist*, **232**(3), 1286–1296.

Zhang, Q., Tang, W., Peng, S., and Li, Y. (2022). Limiting factors for panicle photosynthesis at the anthesis and grain filling stages in rice (Oryza sativa L.). *The Plant Journal*, **109**(1), 77–91.

Zhu, X.-G., de Sturler, E., and Long, S. P. (2007). Optimizing the Distribution of Resources between Enzymes of Carbon Metabolism Can Dramatically Increase Photosynthetic Rate: A Numerical Simulation Using an Evolutionary Algorithm. *Plant Physiology*, **145**(2), 513–526.

Żmieńko, A., Samelak, A., Kozłowski, P., and Figlerowicz, M. (2014). Copy number polymorphism in plant genomes. *Theoretical and Applied Genetics*, **127**(1), 1–18.

Summary

Plant photosynthesis is the biochemical process directly or indirectly sustaining most life forms on planet Earth, by simultaneously providing them with chemical energy and the oxygen required to utilise it. Although the core mechanisms of photosynthesis are quite conserved across plants, not all of them are equal when it comes to their photosynthesis. This is generally referred to as natural variation for photosynthesis. While this variation is most apparent as physiological variation, hypotheses have been formulated on the possibility of tying this variation to genetic variation between plants, and tapping into this variation to uncover genetic traits that could be employed to increase the yield of agricultural crops. This thesis focuses on one plant species, the little-known *Hirschfeldia incana* (L.) Lagr.-Foss., and presents a number of findings arising from investigations on the genetic background of its photosynthetic performance, quantified in terms of photosynthetic light-use efficiency (LUE) or photosynthetic capacity.

Chapter 1 contains information that helps contextualising the research described in this thesis. In this introductory chapter, after compiling an account of the key moments in the history of photosynthesis research, I describe why photosynthesis is nowadays a key subject in plant sciences, and what implications it can have for breeding programs aimed at increasing crops' yield. I then define three concepts this thesis relies on, i.e. those of high irradiance, photosynthetic light-use efficiency (LUE) and photosynthetic capacity (P_{max}). Finally, I briefly review the most recent advances in technologies that can be applied to photosynthesis research, and summarise the rationale and the main findings of my thesis.

In **Chapter 2** co-authors and I build a case for greater focus on exploring and exploiting natural genetic variation in photosynthesis. We then argue for the use of high photosynthetic capacity species as models for exploring the physiological and genetic basis of high photosynthetic efficiency. Having described some of these species with the most striking efficiencies and discussed how they would not be suitable for research purposes, we identify Brassicaceae species *H. incana* as a promising candidate, and describe its basic biology, evolutionary history, and ecology. We then present some of the species' photosynthetic characteristics, highlighting its high photosynthetic capacity, which associates with high photosynthetic LUE, and the tolerance of its photosynthetic machinery to extreme irradiance conditions. Our conclusion is that *H. incana* is an excellent model species for studies aiming at understanding natural genetic variation in photosynthetic light-use efficiency.

Taking the first steps towards developing *H. incana* as a model species for photosynthesis research, in **Chapter 3** co-authors and I present a reference genome of *H. incana* and confirm the high photosynthetic light-use efficiency of the species. We then show how the *H. incana* genome has extensively diversified from those of *B. rapa* and *B. nigra* through large chromosomal rearrangements, species-specific transposon activity, and differential retention of duplicated genes. Based on this extensive diversification, we hypothesize that copy number variation for genes involved in photosynthesis or photoprotection, resulting in their higher expression, could be a major genetic determinant of *H. incana*'s higher photosynthetic LUE. We investigate this by performing a gene expression assay for selected photosynthetic and photoprotective targets, and while not finding evidence for *H. incana*-specific patterns, we report a positive correlation between the expression of genes duplicated in the *H. incana*, *B. rapa*, and *B. nigra* genomes and their copy number. Our conclusion is that while gene copy number variation appears to play a role in higher photosynthetic LUE, other genetic determinants are likely involved in it.

In **Chapter 4**, I analyse the transcriptional signature associated with growth under high light of *H. incana* and compare it with that of *A. thaliana*, *B. rapa*, and *B. nigra*, with the underlying idea of exploring the genetic determinants of photosynthetic LUE from a functional perspective. By analysing gene expression patterns using a panproteome, I confirm that all four species actively regulate genes associated with the photosynthetic process in response to high light. I then describe some unique photosynthetic gene expression patterns specific to *H. incana*. In a number of cases, *H. incana* is shown to exhibit differential expression for specific genes not observed in the other species, while in others significantly higher transcript abundance is reported for certain genes in *H. incana* compared to the other species. Since this chapter presents the analysis of only one part of the generated transcriptomic dataset, I conclude that the transcriptional strategies adopted by *H. incana* on photosynthetic genes are likely to be applied to other processes as well that are crucial to the species' performance.

Chapter 5 presents the results for a broader analysis of the photosynthetic efficiency and a number of biochemical and anatomical characteristics across a panel of ten Brassicaceae species. I report that *H. incana* outperforms most relatives in terms of photosynthetic efficiency, measured as ΦPSII. I then investigate the relationship between this and leaf chlorophyll content, vegetation indices, leaf thickness and composition in terms of palisade and spongy mesophyll, and stomatal densities and major axis lengths measured on both adaxial and abaxial sides of leaves. Interestingly, I report few correlations between these parameters and ΦPSII, with the only strong ones being for stomatal counts on both leaves' sides. While the investigated leaf traits are not coordinated at the species level, clustering of the species based on all parameters revealed similarities that are partially consistent with the phylogenetics of the Brassicaceae family. Considering these findings and the high photosynthetic LUE I report for *H. incana*, *B. nigra*, and *Z. ait-atta*, I conclude that this dataset supports the idea that multiple different strategies in the Brassicaceae family enable some of its members to achieve high photosynthetic LUE.

In **Chapter 6**, I fist draw some parallels between the results of the previous

chapters. I group all the data I collected on the photosynthetic LUE of *H. incana* and its relatives, and interpret them in terms of photosynthetic resilience to challenging environmental conditions. I then bridge between the results the genomic and transcriptomic investigations I performed to speculate on how much I explained of the genetic basis of *H. incana*'s photosynthetic LUE. I subsequently integrate all of my findings and contextualise them with ecophysiological concepts to conclude that different strategies can lead a plant to achieve high photosynthetic LUE. I further rely on ecophysiology to argue that future photosynthesis research should abandon the current leaf-centric and crop-centric perspectives, as only this will allow for a better understanding of photosynthetic adaptation and resilience to the various environmental conditions encountered in agroecosystems. Finally, I take a step back to argue that only by relying on more holistic approaches photosynthesis research will be able to contribute to crop yield increase.

About the author

Francesco Garassino was born in Alba, Italy, in 1994. Fascinated with biology by virtue of a brilliant high school teacher, he enrolled in the Biotechnology BSc programme at the University of Turin, which he graduated from *cum laude*. Having realised he had a greater interest for plants than other organisms, he moved to the Netherlands to pursue a Plant Biotechnology MSc at Wageningen University.

After graduating *cum laude*, he embarked in a PhD project stemming from his MSc thesis research. In the span of four and a half years, as part of a multidisciplinary research consortium, he studied *Hirschfeldia incana*, trying to understand how the plant can achieve high photosynthetic light-use efficiency.

Interested in (too) many things ranging from high-throughput image analysis to science communication, he was involved in a number of projects that together with his dog Pepa kept him busy outside of working hours. The most relevant can be accessed via linktr.ee/francescogarassino.

List of publications

The Plant Journal 2022; 112:1298-1315.
Francesco C. L. Garassino*, Raúl Y. Wijfjes*, René Boesten*, Francisca Reyes Marquez, Frank F. M. Becker, Vittoria Clapero, Iris van den Hatert, Rens Holmer, M. Eric Schranz, Jeremy Harbinson, Dick de Ridder, Sandra Smit, and Mark G. M. Aarts.
"The genome sequence of *Hirschfeldia incana*, a new Brassicaceae model to improve photosynthetic light-use efficiency"
*authors contributed equally to this publication

Food and Energy Security 2023; 12, e420.
Graham Taylor*, **Francesco C. L. Garassino***, Mark G. M. Aarts, and Jeremy Harbinson.
"Improving C_3 photosynthesis by exploiting natural genetic variation: *Hirschfeldia incana* as a model species"
*authors contributed equally to this publication

BioRxiv 2023; 10.18.562717.
Francesco C. L. Garassino, Sofia Bengoa Luoni, Tommaso Cumerlato, Francisca Reyes Marquez, Mark G. M. Aarts, Jeremy Harbinson, Harm Nijveen, and Sandra Smit.
"Comparative transcriptomics of *Hirschfeldia incana* and relatives highlights differences in photosynthetic pathways"
Submitted to NAR Genomics & Bioinformatics.

Acknowledgments

Doing a PhD is like reading all the sentences about doing a PhD: an exercise in patience and grit, made more manageable by the camaraderie of those around you. *Many* people got me to the end of this, so these will be some long acknowledgments. You've been warned!

First, I'd like to recognize the three supervisors who spent these four and a half years with me. **Mark**, thank you for giving me the opportunity to undertake this fascinating project and for believing in the project even when there was little to believe in. **Sandra**, thank you for guiding me through six challenging yet amazing months of RNA-Seq analysis, for being there for me anytime I needed it, and for the kind words at the right time. Your presence has been inspiring for me, and I would never have reached this goal without you on my team. **Jeremy**, thank you for all the good talks, the occasional jokes, and especially for sharing your unique and precious views on the world, science, and, of course, photosynthesis. I wish we had more time to spend just chatting about whatever came to mind.

All of this was even more enjoyable thanks to my fellow *Hirschfeldia* trailblazers, whom I want to acknowledge first for their determination to work together, regardless of the circumstances. **Ludovico**, you've been the best fellow PhDer I could have hoped for. You not only came to my rescue on numerous occasions when I struggled with everything from gardening to circuitry but also demonstrated a calm and composed approach to the challenges of a PhD, which I will always remember. **Sofia**, though we only worked together for about a year, what a year it was! I truly appreciated your kindness and your positive attitude. I can never thank you enough for coming aboard the RNA-Seq journey and helping me complete that chapter. Having you as a paranymph is the perfect way to celebrate our time together. **Nam** and **Moges**, even though we may have worked together a little less, I highly valued your company and contributions, and I look forward to seeing the results of your hard work in outstanding publications!

The trailblazing would not have been possible without the team supervising our progress and contributing to our work. Besides my direct supervisors, I'd like to express my gratitude to **Eric** for his support and the good laughs, and to **Steven** for introducing me to the intricate world of gas analysis, teaching me the significance of data

quality, and indulging in sports talk. My thanks also extend to **Dick** for demonstrating a no-nonsense approach to problem-solving, from my doubts about doing a PhD all the way down to some final clustering analysis, and to **Herbert**, **Paul**, and **Xinyou** for the insightful discussions about our research and maintaining their always-positive attitude.

Let me now take a step back and thank the people who made the whole project a reality. I will always be grateful to **Maarten**, **Menno**, and **Egbert**, not only for providing me and my fellow trailblazers with this incredible opportunity but also for their unconditional support and genuine interest in our work. I am very happy to see your dream of the Jan Ingenhousz Institute materialize, and wish that many other researchers will enjoy your support! Additionally, I want to acknowledge the members of our scientific advisory board. **Susanne**, **David**, and **Herbert**, thank you for keeping an eye on that big group of photosynthesis-ish people, and for truly putting effort in your supervision role.

As I've mentioned many times during these four and a half years, I consider the smiles and successes of "my" students to be my greatest achievement. Besides their invaluable contributions to my work, I am very grateful to each of them for offering me challenging perspectives and the opportunity to build wonderful personal relationships that I am proud to say are still intact, even years after the completion of their projects. My deepest respect and appreciation go to **Jeffrey**, **Loes**, **Stijn**, **Matthijs**, **Jacopo**, **Vittoria**, **Marco**, **Tommaso**, **Femke**, and **Rachel**. It has been an absolute pleasure to share parts of this journey with all of you! To Femke and Rachel, it is an honor to feature your artwork as my thesis cover - thank you for embarking on this final project with me!

Now, it's time to express my gratitude to all the wonderful people who collaborated with me. First and foremost, I want to thank the two individuals who have been by my side since my MSc days. **René**, I'm glad that we transitioned from a teacher-student relationship to being colleagues, and that I kept learning from you. Your passion for sharing knowledge, your dedication to fundamental science, and your love for reptiles are unparalleled. Many times, I found inspiration in your example when things got tough. **Raúl**, I still remember how well we "clicked" at the start of my MSc thesis, and I'm proud to say that our connection remains strong after all these years and the numerous rounds of revisions on the genome paper. Thank you for introducing me to the world of bioinformatics and, above all, for your endless patience and availability. Whenever I needed your assistance, no matter how trivial my (code) problem may have seemed, you were always there with a smile!

Onto the Genetics group we go. **Louise**, I wish we had more time to challenge the status quo. Thank you for the laughter, the mind-relieving breaks, the memes, the nonsense, and for standing by me on the defense stage! **Tom**, thank you for being my mentor in chlorophyll fluorescence, an occasional software engineer, a shining example of dedication, and a wonderful neighbor. **Ben**, thank you for... well, for be-

ing yourself. You've helped me in so many ways that I can't even remember them all, and I'm grateful to have shared this journey with you. **Roel**, you've moved on, but I haven't forgotten your camaraderie and your willingness to help me. **Jitpanu**, you've been an example of resilience for me. I treasure your kindness and all the times you sated our hunger! **Phuong** and **Laavanya**, your strength has been an inspiration. I will cherish the conversations we've had and your support when times were tough. **Alanna**, thank you for the great vibes, the occasional walks, and our discussions on France, Italy, and blue cheese. **Lennart**, you've shown me calm and perseverance. Thank you for that and for the nice bike rides we shared with Tom and Ben! **Maarten**, your stories and wisdom have been incredibly inspirational. If I ever change my mind and end up in academia, I'll consider myself accomplished if I can achieve a quarter of what you've done. **Bas**, you gave me hope and purpose during the most challenging part of my PhD, and I'll always be grateful for that. Thanks as well for the inspiring conversations we've had over the years! To **Anna**, **Alan**, **Alex**, **Bart**, **Duur**, **Eric B.**, **Eric W.**, **Eveline B.**, **Eveline S.**, **Fons**, **Gabriella**, **Hedayat**, **Helena**, **Hilje**, **Joost**, **Jordi**, **Lisa**, **Mariana**, **Mariska**, **Nicky**, and **Olga**, thank you for making the Genetics group the wonderful and unique place it is and for sharing a part of this journey with me.

In the previous paragraph, I purposefully did not mention a very specific group of Genetics colleagues: our support staff! They should not go unsung; quite the opposite. Let me begin with **Corrie**: thank you so much for always being available for pretty much anything happening in and around the lab, and even more so for all the laughter and our conversations about mountains. **Frank**, I cherish the times we've worked together and discussed bees, food, Nanopore sequencing, and wines. I hope we'll have the opportunity to open a truly exceptional bottle together someday. **José**, companion of microscopy days, gossip and banter, even nurse at times, I am lucky to have worked with you and learned from you. **Francisca**, while it might seem like your contribution to my project was minor, your indestructible good mood and kindness brought me joy on countless occasions! **Gema**, even though you moved on with your career around the halfway point of my PhD, I still want to express my gratitude for all the tedious work you did for me without hesitation, and for fuelling my desire to visit the Spanish mountains. Let me also extend my thanks to our secretaries **Marjan** and **Wytske**, who not only handled many less-than-exciting tasks but were often the glue that held our group together!

I'd now like to take a moment to express my gratitude to members of our "neighbor" groups. **Harm**, we accomplished something that not many others have with RNA-Seq, and it was an incredibly fun journey. Thank you so much for always being available to answer my questions and address my doubts, and for your endless kindness. **Justin**, I appreciate how, despite the randomness of it, you welcomed me as your office mate. I believe that the pleasant environment we created together played a significant role in pushing me toward this finish line. **Dirk-Jan** and **Robin**, you've been my guardian angels whenever pangenomics came up, and I am truly grateful for that. To **Cloé**, **Freek**, **Guusje**, **Marieke**, **Klaas**, **Nina**, **Patrick**, **Qiong**, **Siva**,

and **Wouter**, thank you for the good vibes, or participating in that crazy rap, or both! **Elias**, it's wonderful how photosynthesis brought us closer as friends. Thank you for involving me in some out-of-the-box thinking! **Martina**, your smile was like a ray of sunshine in the dark Klima corridors. Thank you for the good times! To **Maarten** and **Sarah**, your friendliness and competence with photosynthesis equipment were just what I needed during the many times I got hopelessly stuck. I'm glad you were there to help with a smile. **Nastassia**, thank you for pushing me to ask more challenging questions and for making me want to pick fights again.

We're more than three pages in, but we're not nearly done, so please hang tight. Now, I want to express my gratitude to the people who made the more technical aspects of my work possible. **Taede** and **Gerrit**, I've learned so much from both of you, and now I am no longer intimidated by working with plants. Thank you for supporting every wild idea we came up with for our climate room and for ultimately letting me play carpenter, electrician, and plumber. **Bertus**, **Chris**, **David**, **Felix**, **Jannick**, **Jeroen**, **Pinelopi**, **Rinie**, **Rohan**, and **Wilfred**, thank you for allowing me to sleep soundly when my plants were in the climate room, greenhouse, or even the field. **Rick**, I am delighted to have met you and to have engaged in exciting phenotyping work together. Thank you for making me an ambassador of NPEC and giving me the opportunity to meet so many interesting people!

During my project, I had the opportunity to collaborate with some exceptional individuals outside of Wageningen University, and I would like to acknowledge their valuable contributions. Let me begin by thanking **Graham** for being one of the initiators of the entire *Hirschfeldia* business and for agreeing to co-author a nice opinion piece with me. I appreciated the lively discussions and your dedication to pursuing your PhD despite working full-time in another job, and I look forward to reading your thesis soon! **Federico**, thank you for introducing me to the wild world of image analysis, for figuring out so many things I depended on, and for the nerdy giggles. **Carlos**, even though we never met in person, your support for my project through your sampling expeditions in the Spanish countryside has been invaluable. Thank you for providing us with a beautiful collection of *Hirschfeldia* accessions! **Bernard**, you welcomed me into your laboratory and your home, dedicating an amount of time that I could have never asked for, and you have been a true mentor to me. I am so happy that, despite all the circumstances, I had the opportunity to visit Cadarache and play with Li-Cors under your guidance.

During the second half of my PhD, I became increasingly interested in science communication. Here are my high-fives to those who helped me with it. **Annika**, thank you for creating the Science Communication Interest Group and then co-running it with me for a year. It's been a fantastic experience, and I've learned so much from it! **Jeroen**, thank you for being there when I started with my communication projects. **Sophie**, thank you for the enjoyable projects we worked on together and for the mood-lifting smiles. **Annet** and **Jan-Willem**, thank you for believing in me and giving me the opportunity to step onto the Dies stage. **Barend** and **Huib**, thank you

for making my first time in front of a camera so comfortable and for introducing me to the world of communication. **Simon**, thank you for involving me in a high-end project like your PhD stories and for the great time we had during filming. **Paola**, thank you for saying yes to the idea of creating a podcast and for sharing the responsibilities with me!

Now, it may seem that all my life in the past years was about work. Well, that couldn't be more wrong. Being in Wageningen allowed me to get to know the wildest, most amazing group of people. Here comes a lengthy list of friends, starting with those I've known for the longest time. **Cristina**, I'm not sure if you've been more of a friend or a grandmother, but at the end of the day, you've excelled at being both here. Thank you so much for always being there for me. **Michele**, **Renzo**, **Anastasia**, **Anna**, **Liam**, **Lucas**, **Eirini**, I've made countless memories with all of you that I will cherish for the rest of my life. I hope that, regardless of the geographical distance life may put between us, we'll remain one big family! **Lisa** and **Ivano**, you were the first Dutch friends I made here in the Netherlands, and I'm grateful for all the wonderful experiences we've had (and the delicious food we've enjoyed) together over these years. **Alex**, **Arne**, **Auke**, **Bas**, **Daniela**, **Donna**, **Eline**, **Giuditta**, **Jana**, **Joliese**, **Jorge**, **Max**, **Niccolò**, and **William**, we've certainly had a lot of fun during these years. Thank you for making this small town so much more interesting!

Upon moving into Lombardi, we were fortunate to encounter some truly wonderful people. **Aïsha**, **Alice**, **Anne**, **Evelien**, **Francesca**, **Jelle**, **Jenny**, **John**, **Kaylee**, **Lena**, **Lisette**, **Sander**, and **Saskia**, the memories we created are again countless. I want to thank you for keeping me sane during Covid, for all the times we talked (dogs), for the shared dinners, for the uplifting smiles and cups of coffee, and for so much more. **Stijn**, thank you for all of that, and also for joining me in all the recent nonsensical running adventures.

There are some people who are currently living far away that I would like to thank for their ongoing support through the years. **Alice** and **Federica**, thank you for repeatedly bringing some much-needed Italianness into our lives (yes, that apparently is a word). **Adi** and **Lucia**, thank you for all the good food and the fresh ideas you brought to the table! **Silvia**, even though you moved away, you always kept us in your thoughts, and your visits have been true highlights. **Alberto**, **Andi**, **Marco**, and **Matteo**, getting to see you when I'd return to Italy and witnessing how your paths were taking shape kept me connected to the homeland. While we might not have spoken for months or even years at times, every time we'd meet, it felt as if I had never left, something I consider pure luck.

I've already talked about the "Wageningen family", but please indulge me as I say a few words about my actual family as well. **Mamma** and **Papà**, you've always supported me in everything I wanted to do. You've shown me the beauty of simple things and have endured the distance that has separated us for over seven years. Thank you for every single thing you've done for me. **Laura**, I am immensely proud of your

life achievements and so happy to have you as my big sister, always by my side. I look forward to making many more memories with you and **Riccardo**. **Claudio**, I am honored to have you as my brother-in-law. Thank you for sharing all your passions with me over these years; hearing about them has always felt like breathing fresh air. **Ale**, **Alice**, **Arianna**, **Carlo**, **Diego**, **Gianpiero**, **Nanda**, **Olivia**, **Paolo**, and **Silvia**, I want to express my gratitude to all of you as well for your unconditional support over the years and for always making me feel at home.

If you've made it through all these pages, well, my respects! This long list of acknowledgements ends with the one person who has had the biggest impact on me over the past seven-and-a-bit years. My dear **Sara**, you've been with me through many of the best and worst times of my life so far. You've put up with the hard times in a manner that I wouldn't describe as anything less than heroic. You've been the true trailblazer, showing me the path out of some scarily dark places. If it wasn't for you and your unconditional love, I would never have finished this thesis. I cannot wait to see where life will take us and Pepa next!

I would maintain that thanks are the highest form of thought;
and that gratitude is happiness doubled by wonder.

Gilbert Keith Chesterton, A short history of England

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to: Francesco Garassino
Date: 11 December 2023
Group: Laboratory of Genetics
University: Wageningen University & Research

1) Start-Up Phase	date	cp
► **First presentation of your project**		
"Resolving the extreme photosynthetic efficiency at high light of grey mustard (*Hirschfeldia incana*)"	23 Apr 2019	1.5
► **Writing or rewriting a project proposal**		
► **MSc courses**		
WUR MSc course Advanced Bioinformatics (BIF-30806)	Nov-Dec 2019	6.0
Subtotal Start-Up Phase		7.5

2) Scientific Exposure	date	cp
► **EPS PhD days**		
EPS Get2Gether 2020, Soest (NL)	10-11 Feb 2020	0.6
EPS Get2Gether 2022, Soest (NL)	3-4 May 2022	0.6
► **EPS theme symposia**		
EPS theme 3 'Metabolism and Adaptation', Nijmegen (NL)	21 Oct 2019	0.3
EPS theme 3 'Metabolism and Adaptation', Wageningen (NL)	5 Nov 2021	0.3
► **National platform meetings**		
Annual Meeting 'Experimental Plant Sciences', Lunteren (NL)	8-9 Apr 2019	0.6
Annual Meeting 'Experimental Plant Sciences', online	12-13 Apr 2021	0.5
Annual Meeting 'Experimental Plant Sciences', Lunteren (NL)	11-12 Apr 2022	0.6
Annual Meeting 'Experimental Plant Sciences', Lunteren (NL)	17-18 Apr 2023	0.6
Netherlands Society for Evolutionary Biology (NLSEB) 2021 meeting & NLSEB2021 PhD and Postdoc meeting, online	19-20 Apr 2021	0.6
Dutch Bioinformatics & Systems Biology (BioSB) Conference 2021, online	15-16 Jun 2021	0.3
► **Seminars (series), workshops and symposia**		
Seminar "Metabolic engineering to enhance photosynthesis and increase crop yield" - Andrew Simkin	21 Mar 2019	0.1
Symposium "Land Plant Evolution & Improving Photosynthesis and Crops", Wageningen (NL)	20 Jun 2019	0.3
Workshop "The use of the Multispeq/Photosynq combination" - David Kramer	9 Oct 2019	0.1
Seminar "Harnessing natural variation, genetic diversity and detailed, high throughput phenotyping to understand the regulation of photosynthesis" - David Kramer	10 Oct 2019	0.1
SKiPR seminar "CO2 diffusion inside leaves during photosynthesis - Tory Clarke	28 Nov 2019	0.1
Symposium "Optimising photosynthesis for societal needs", Wageningen (NL)	10 March 2020	0.3
Seminar "Caring about code" - Antoine Languillaume	9 Jun 2020	0.1
Seminar "PSG regulations concerning the Nagoya Protocol" - Mirjam Lemmens-Pott	29 Sep 2020	0.1
Capitalise Photosynthesis 2030+ Webinar Series	5 Oct 2021	0.1
WUR SciCommIG: "What is Data Journalism?" - Marie-Louise Timcke	18 Jan 2022	0.1
WUR SciCommIG: "From the lab to SciComm" - Elodie Chabrol	5 Apr 2022	0.1
Seminar "The use of historical barley chlorophyll mutants to learn about chlorophyll biosynthesis" - Mats Hansson	7 Apr 2022	0.1
1st FREECLIMB workshop "Genome-based characterization of mediterranean fruit crop genetic resources for better resilience to climate extremes"	28 Apr 2022	0.1
WUR SciCommIG: "Science Communication, YouTube and me" - Simon Clark	14 Jun 2022	0.1
EPS Flying Seminar - Simon Clark	15 Jun 2022	0.1
CEPPG workshop "Recent advances in controlled environment phenotyping and lab-to-field translation", Wageningen (NL)	26 Sep 2022	0.2
SKiPR seminar, Peter Jahns and Tom van den Berg	2 Nov 2022	0.1
EMBL-EBI Webinar "Pangenomics and machine learning for crop improvement"- Dave Edwards	12 Apr 2023	0.1
Fraunhofer PhenoTalks "Phenotyping technologies: Bridging the gap between concept development and industrial deployment" - Bas van Eerdt	19 Apr 2023	0.1
Seminar "Quantitative links between gas exchange of leaves and photosynthetic biochemistry; considering species diversity" - Susanne von Caemmerer	31 May 2023	0.1
Seminar "The kinetic control of rapid excitation quenching and implications for photosynthetic energy storage in the real world" - David Kramer	31 May 2023	0.1
► **Seminar plus**		
Meeting with David Kramer	10 Oct 2019	0.1
Meeting with Tory Clarke	27 Nov 2019	0.1
► **International symposia and congresses**		
LXIII Annual Congress of the Italian Society of Agricultural Genetics, Naples (IT)	10-13 Sep 2019	1.0

CONTINUED ON NEXT PAGE

	Congress "Biophysics of photosynthesis - From molecules to the field", Rome (IT)	2-4 Oct 2019	0.9
	Plant Biology Europe 2021, online	28 Jun - 1 Jul 2021	1.2
	International Plant Phenotyping Symposium (IPPS) 2022, Wageningen (NL)	27-30 Sep 2022	1.1
►	**Presentations**		
	Poster: "*Hirschfeldia incana* as the model species for high light-use efficiency in photosynthesis", Rome (IT)	2-4 Oct 2019	1.0
	Poster: "*Hirschfeldia incana* , an attractive model to explore the physiology and genetics of high photosynthesis", online	28 Jun - 1 Jul 2021	1.0
	Seminar talk: "*Hirschfeldia incana* : a new Brassicaceae model to improve photosynthetic light-use efficiency", BIAM Cadarache (FR)	29 Mar 2022	1.0
	Poster and poster flash talk: "Exploring photosynthesis rates across the Brassicaceae family", Wageningen (NL)	27-30 Sep 2022	1.0
	Presentation: "Understanding high photosynthetic efficiency with *Hirschfeldia incana* ", Agricola Moderna scientific meeting, online	19 Jan 2023	1.0
	Presentation: "Comparative genomics/transcriptomics of *Hirschfeldia incana* and relatives highlight differences in photosynthetic pathways", Lunteren (NL)	18 Apr 2023	1.0
►	**Interviews**		
►	**Excursions**		
	Online EPS Company Visit, Rijk Zwaan	16 Jun 2021	0.2
	Subtotal Scientific Exposure		18.2

3) In-Depth Studies		*date*	*cp*
►	**Advanced scientific courses & workshops**		
	EPPN Course 'Experimental design and statistical analysis of phenotyping platform experiments', Wageningen (NL)	7 May 2019	0.2
	PE&RC/WIMEK Course 'Linear Models', Wageningen (NL)	12-14 Jun 2019	0.9
	PE&RC Course 'Tidy data transformation and visualization with R', online	8-18 Jun 2021	1.2
	PE&RC/WIMEK Course 'R and big data', Wageningen (NL)	14-15 Oct 2021	0.6
►	**Journal club**		
	PE&RC Frontier Literature in Plant Physiology (FLOP) - Photosynthesis PhD Discussion Group	2022-2023	0.6
►	**Individual research training**		
	Training on mesophyll conductance estimation - BIAM-CEA, Cadarache (FR)	28 Mar - 8 Apr 2022	3.0
	Subtotal In-Depth Studies		6.5

4) Personal Development		*date*	*cp*
►	**General skill training courses**		
	WGS Course 'Critical thinking and argumentation', Wageningen (NL)	21 May 2019	0.3
	WGS Course 'Presenting with Impact', Wageningen (NL)	6-27 May 2019	1.0
	WGS Course 'Ethics in Plant and Environmental Sciences', Wageningen (NL)	16 Oct 2019	0.3
	EPS Introduction Course, Wageningen (NL)	29 Oct 2019	0.3
	WGS Course 'Project and Time Management', Wageningen (NL)	30 Oct - 11 Dec 2020	1.5
	RBZS workshop series 'PhD Workshops - a free online skill-building event', online	1&8 Dec 2020	0.2
	EMBL Careers Webinar 'Core facility management', online	16 Apr 2021	0.1
	EPS Flying Seminar 'Modern Science Communication', Elizabeth Haswell	11 May 2021	0.1
	WGS PhD Carousel workshop 'Storytelling for academics', online	25 May 2021	0.1
	WGS PhD Carousel workshop 'Entrepreneurship as a skillset for career development', online	28 May 2021	0.1
	WGS Course 'Rmarkdown', Wageningen (NL)	6-13 Dec 2021	0.6
	Seminar 'Student Well-being by student psychologist'	1 Feb 2022	0.1
	WGS Workshop 'The Choice: Un-box your PhD process & take charge of your performance', Wageningen (NL)	24 Mar - 12 May 2022	0.5
	WGS Course 'Career Orientation', Wageningen (NL)	15 Nov - 6 Dec 2022	1.5
►	**Organisation of scientific meetings, PhD courses or outreach activities**		
	Organisation of Science Communication Interest Group (SciCommIG) sessions	Sep 2022 - Jun 2023	1.0
►	**Membership of EPS PhD Council**		
	Subtotal Personal Development		7.7

5) Teaching & Supervision Duties		*date*	*cp*
►	**Courses**		3.0
	GEN-30306 Genetic Analysis Trends and Concepts	Period 6 2018-2019	
	GEN-30306 Genetic Analysis Trends and Concepts	Period 1 2019-2020	
	GEN-20806 Plant Biotechnology	Period 2 2019-2020	

	GEN-20806 Plant Biotechnology	Period 2 2020-2021	
►	**Supervision of BSc/MSc projects**		3.0
	Research internship 'The effect of high photosynthesis efficiency genes in *Hirschfeldia incana* '	Aug 2019 - Jan 2020	
	M.Sc. thesis 'A comparative study of mechanisms underlying high photosynthetic efficiency through gas exchange and chlorophyll fluorescence measurements of *Hirschfeldia incana* and *Brassica nigra* '	Oct 2019 - Feb 2020	
	B.Sc. thesis 'Clarifying the (cyto)genetics of *Hirschfeldia incana* and allied species'	Oct 2019 - Jan 2020	
	M.Sc. Thesis 'The photosynthetic response of Brassicaceae species to variation in ammonium availability'	Jan 2020 - Oct 2020	
	B.Sc. Thesis 'Investigating essential biochemical aspects of the photosynthesis of *Hirschfeldia incana* and other Brassiceae'	Feb 2020 - Aug 2020	
	Erasmus+ internship M. Al.	Sep 2020 - Jan 2021	
	M.Sc. thesis 'Analysis of gene expression impact of differential irradiance on leaf development of *H. incana* and relatives'	Aug 2020 - Apr 2021	
	M.Sc. thesis 'RNA-sequencing analysis of *H. incana* and related species under contrasting light conditions'	Oct 2021 - Jun 2022	
	B.Sc. thesis 'Investigation of photosynthesis-related anatomical properties in a panel of Brassicaceae species'	May 2022 - Jul 2022	
	B.Sc. thesis 'Exploring natural variation for photosynthesis within the Brassicaceae family: a study of stomatal ratios and characteristics'	May 2022 - Jul 2022	
	Subtotal Teaching & Supervision Duties		6.0

TOTAL NUMBER OF CREDIT POINTS*	**45.9**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

** A credit represents a normative study load of 28 hours of study.*

The research described in this thesis was financially supported by the “Stichting Photosynthesis 2.0” Research Fund.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Rachel Pool (@rashelhanout) and Femke Schreurs (@femke_art_2000).

Printed by Ridderprint, The Netherlands.