Gene discovery using high-throughput phenotyping under fluctuating growth conditions and rapid identification of mutated genes

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität

Halle-Wittenberg,

vorgelegt

von Herrn Marc Christian Heuermann

geboren am 16.02.1987 in Herdecke

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

PREFACE

The rapid mapping part of this work was published as a technical advance article in The Plant Journal 2019:

Heuermann, M.C., Rosso, M.G., Mascher, M., Brandt, R., Tschiersch, H., Altschmied, L. and Altmann, T. (2019) Combining next-generation sequencing and progeny testing for rapid identification of induced recessive and dominant mutations in maize M_2 individuals. *The Plant Journal*, tpj.14431.

doi.org/10.1111/tpj.14431

I, Marc Christian Heuermann, am the first and corresponding author and wrote more than 90% of the manuscript. The only section not written by myself is "2.17 SNP calling" in the material and methods part, which was written by Dr. Martin Mascher. All sections are cited appropriately at the beginning and end of a relevant section and numbering of figures and tables was adapted accordingly to fit the design of this thesis.

ABBREVIATIONS

Abbrev.	Description	Abbrev.	Description
BLUE	best linear unbiased estimator	PCR	polymerase chain reaction
BLUP	best linear unbiased predictor	PKH	Plant Cultivation Hall
BMBF	Federal Ministry of Education and Research	Pos	position
BSA	bulked segregant analysis	PS I / II	photosystem I / II
Chr.	chromosome	PVE	phenotypic variance explained
cM	centimorgan	Q-Q	quantile-quantile
Ct	cycle threshold	QTL	quantitative trait locus
DAG	days after germination	Rfd	chlorophyll fluorescence decrease ratio
DAS	days after sowing	RGB	red, green, and blue
DNA	deoxyribonucleic acid	R_n	repeatability
DW	dry weight	RNA	ribonucleic acid
EMS	ethyl methanesulfonate	ROS	reactive oxygen species
Eq.	equation	SNP	single nucleotide polymorphism
F_m	maximum fluorescence	TKW	thousand kernel weight
F_q	photochemical quenching of fluorescence	tt	thermal time
F_v/F_m	maximum PS II efficiency	V	vegetative
G x E	genotype x environment interaction	VAZ	violaxanthin, antheraxanthin and zeaxanthin
GDD	growing degree days	vis	visible light
GO	gene ontology	VT	vegetative tassel
GOI	gene of interest	WGS	whole-genome-shotgun
GWAS	genome-wide association study	WT	wild-type
GWH	greenhouse	$\Phi_{(PSII)}$	PS II operating efficiency (F_q'/F_m')
HSV	hue, saturation, and value		
IAP	Integrated Analysis Pipeline		
Lab	L*,a*,and b* color space		
LD	linkage disequilibrium		
LED	light-emitting diode		
lm	lumen		
lx	Lux		
maf	minor allele frequency		
MTA	marker-trait association		
n.s.	not significant		
NAM	nested association mapping		
NGS	next-generation sequencing		
NPQ	non-photochemical quenching		
PC	principal component		
PCA	principal component analysis		

1. INTRODUCTION

During domestication (< 12,000 years ago), crop species were exposed to an ever-fluctuating, changing environment and diverse geographical locations due to human migration, to which crops had to adapt (Meyer and Purugganan, 2013). Seasonal changes and year to year variation led breeders to explore the genotypic performance of their breeding lines in multi-environment trials over several years (Ceccarelli, 2015). Climate change is altering environments with an increasing pace, which requires a large degree of phenotypic plasticity from a plant to keep or gain fitness benefits (Nicotra *et al.*, 2010). The adaptive value of phenotypic plasticity of fitness-related traits under dynamic environments can be evaluated by natural genetic variation in accession panels and by genetic variants introduced in mutant populations (Van Kleunen and Fischer, 2005; Laitinen and Nikoloski, 2019). The dynamics of physiological acclimation to fluctuating environments is an important feature for individual plants when their environment suddenly changes, like the sudden exposure to high irradiance for an obligatory understory crop like *Coffea arabica* (Gratani, 2014). To mimic the ever-fluctuating field and environmental conditions in greenhouses is of intrinsic difficulty and the translational step from a controlled environment to the field is of utmost importance to estimate the persistence of a greenhouse-identified trait under field conditions (Nelissen *et al.*, 2014). Yet, in the last decades, plant research has focused on controlling the growth condition of plants to peel off environmental variation and expose the differences that are the sole response to treatment or genotypic variation to get a better understanding about the function of specific genes.

1.1 Genome annotation in plants

The best analyzed and understood plant species to date is *Arabidopsis thaliana*. Its genome was sequenced in 2000 (Arabidopsis Genome Initiative, 2000) and in the same year the 2010 project was founded (Chory *et al.*, 2000). The aim of this project was to "understand the function of every gene by the year 2010" (Chory *et al.*, 2000). However, ten years later, only 37% of the genes were annotated to a biological process, molecular function or cellular component validated by experimental data (Lamesch *et al.*, 2011). Taking all evidence into account, such as sequence similarity, the number increases to an annotation of 77% of the genes (Lamesch *et al.*, 2011). Reciprocally that means that gene ontology of 40% of the annotated genes is not validated by

experimental data and that the gene ontology of 23% of all genes is totally unknown. With the release of Araport11 the functional annotation of protein-coding genes with meaningful product names was increased to 91.8%, which reduced the number of fully unknown genes to 8% (Cheng *et al.*, 2017). In crop plants, annotation by gene ontology is achieved to an even lesser extent. For example in rice, whose genome sequence was published four times (Sasaki and Burr, 2000; Barry, 2001; Goff *et al.*, 2002; Yu *et al.*, 2002), genes were annotated due to sequence similarity to several different plant species i.e. *Arabidopsis thaliana* (Ouyang *et al.*, 2007). The genome of maize was published by Schnable *et al.* (2009) and genes were annotated using evidence-based gene predictions based on *Arabidopsis thaliana* and *Oryza sativa* (Liang *et al.*, 2009) and the FGENESH pipeline developed for *Drosophila melanogaster* (Salamov and Solovyev, 2000). However, it is obvious that annotations by sequence similarity have only the potential to reach the precision of their templates. Thus gene annotation is still an important task to either validate genes, which were annotated *in silico*, experimentally or to unravel the function of genes with still unknown function. Such genes could have proven to be elusive because their mutants neither exhibit a distinct phenotype under standard growth conditions nor have been found in mutant populations screens.

1.2 High-throughput phenotyping under constant growth conditions

The German plant phenotyping network was started to establish high-throughput and field phenotyping methods and installments in Germany. At the IPK Gatersleben three different automated non-invasive plant phenotyping systems for small, medium and large plants, following the plant-to-sensor concept, have been successfully established and scientifically challenged by experiments on drought tolerance and different water regimes (Chen *et al.*, 2014; Harshavardhan *et al.*, 2014; Muscolo *et al.*, 2015; Neumann *et al.*, 2015). The phenotyping facilities at the IPK were designed to record precise phenotypes of plants under controlled growth conditions in a high-throughput manner (Chen *et al.*, 2014; Neumann *et al.*, 2015; Junker *et al.*, 2015; Muraya *et al.*, 2017; Tschiersch *et al.*, 2017). The expert knowledge gained from experiments on the IPK automated non-invasive plant phenotyping system for small and large plants was evaluated and incorporated into optimized cultivation protocols for performing high-throughput phenotyping experiments (Junker *et al.*, 2015). Tschiersch *et al.* (2017) have upgraded all three phenotyping platforms with a light adaptation tunnel and a subsequent FluorCam device, which is able to

evaluate maximum Photosystem II (PSII) efficiency (F_v/F_m) or PSII operating efficiency ($\Phi_{(PSII)}$) in a high-throughput manner. For the IPK automated non-invasive plant phenotyping system for small plants the throughput is 154 carriers per hour. Consequently, it would require two hours and 30 minutes to measure all 384 carriers. To measure the full range of kinetic chlorophyll fluorescence parameters like non-photochemical quenching (NPQ), the throughput would be significantly lower. The parameters extracted from kinetic chlorophyll fluorescence measurements, like $\Phi_{(PSII)}$, estimating the light use efficiency (Tschiersch *et al.*, 2017), or NPQ, a plant's ability to dissipate excess absorbed light into heat (Ruban, 2016), can act as important proxies for plant fitness under abiotic stress levels (Müller *et al.*, 2001; Horton, 2004; Harbinson *et al.*, 2012).

The high-throughput of the IPK automated non-invasive plant phenotyping systems allows to simultaneously grow and phenotype a large number of plants (up to 4,608, 520, and 396 plants for small, medium, and large plants, respectively; (Junker *et al.*, 2015)), which for example enabled the screen of a genetically diverse canola population on the system for large plants used in a subsequent genome-wide association study (GWAS) (Knoch *et al.*, 2019). GWAS can be used to assess phenotypic plasticity in crop species like maize (Kusmec *et al.*, 2017) and to identify genetic variants in natural accession panels underlying Arabidopsis' response to shade avoidance treatment (Filiault and Maloof, 2012). GWAS-based gene discovery of *Arabidopsis thaliana* candidate genes that respond to abiotic stress, like high light, has proven to be effective to identify the genetic basis of complex traits (Rungrat *et al.*, 2016). Nonetheless, all studies have hitherto relied on continuous and controlled growth conditions.

1.3 Influence of fluctuating light conditions on plant growth

In nature, plants are fully exposed to nature's feral wrath – high energetic solar radiation, strong wind and rainfall alternating with prolonged dry seasons and cloudy skies resulting in suboptimal conditions to perform photosynthesis. Phenotypic plasticity, the capacity of one genotype to produce different phenotypes, is vital for sessile organisms like plants to adjust to different environments or to rapid changes to which a plant needs to respond immediately to ensure its survival (Sultan, 2000). Hitherto, most greenhouses are built to provide optimal growth conditions and in the scientific field they were improved to deliver near perfect constant conditions to reduce environmental factors. Those artificial conditions will hardly unmask the effects of unknown,

potential climate-responsive genes. Thus, to discover such genes, there is a need for rapid methods to screen novel mutant populations as well as a need for new ways to mimic natural growth conditions. Additional research has to be done in Arabidopsis but even more importantly in crop plants, as those are fully exposed to the environment in the field.

Under natural conditions in the field or forest, plants are exposed to diurnally fluctuating light conditions. On the one hand, shading by clouds or surrounding plants due to wind can temporarily shade a leaf and drastically reduce the photon flux density. On the other hand, on cloudy days or in crop canopies, breaks in the cloud cover or wind movement of the canopy can drastically increase the light exposure of a leaf. The former are known as shadeflecks and the latter as sunflecks (Chazdon and Pearcy, 1991). Salisbury *et al.* (1968) were one of the first to analyze the frequency of short (seconds to hours) fluctuating light cycles due to cloud cover. They were monitoring the number of short cycles per day during the growing seasons of 1962, 1963 and 1964 and found that light/shade transitions were ranging from ten to over 60 per day (Salisbury *et al.*, 1968). Cloud cover was exemplarily analyzed for the understory growing species *Arnica latifolia* in a forest clearing at 2,800 m elevation for three consecutive years (Young and Smith, 1983). Up to 50% of the days during the growth period were cloudy for 40% or more of the total daylight hours (Young and Smith, 1983). Interestingly, due to the diffuse light during times of cloud cover, light in the understory exceeded values of 300 µmol m^{-2} s^{-1} photons in contrast to a maximum of only 200 µmol m^{-2} s^{-1} photons on clear days (Young and Smith, 1983). The frequency of light fluctuation in crop canopies depends on several factors like planting density, plant architecture, height and wind speed (Norman and Tanner, 1969; Desjardins *et al.*, 1973). Maize genotypes with a more rigid plant structure, for example with a dwarf phenotype or upright leaves, were found to have less light fluctuation in comparison to genotypes with purple or yellow leaves, which were more free to move in response to wind drag (Desjardins *et al.*, 1973). Taken together, the authors could directly link wind speed with increasing light fluctuations (Desjardins *et al.*, 1973). In 1990, Pearcy *et al.* measured between 500 and 1800 sunflecks on a cloudless day in soybean (*Glycine max*). Sunflecks contributed 20-93% of the photon flux density at locations receiving sunflecks. The rapid and reoccurring shift between direct sunlight and cloud cover requires plants to invest a lot of resources into compensating processes, which limits the potential yield of crops by up to 20% (Kromdijk *et al.*, 2016). The impact of fluctuating light on biomass and general plant fitness was

documented for *Arabidopsis thaliana* by knock-out of two important light-responsive proteins, violaxanthin de-epoxidase (*npq1*) and the PsbS protein (*npq4*) (Kulheim *et al.*, 2002). Plants having either of those genes knocked out produced about 35% fewer fruits and seeds per plant under fluctuating light in comparison to no significant differences under continuous light due to a lack of feedback de-excitation. A follow-up study could show that the double knock-out of *stn7* x *stn8*, depleting the capability of state-transition and phosphorylation of core photosystem II, caused the same extensive decrease in seeds per plant under natural, fluctuating light conditions (Frenkel *et al.*, 2006). Furthermore they found no additive effect for double knock-out of *stn7* x *npq4*, so there was no further reduction of plant fitness if both feedback de-excitation and state-transition were disturbed (Frenkel *et al.*, 2006). Wagner *et al.* (2008) could show that the *stn7* mutant is also lacking the beneficial potential of long-term response acclimation (shift every two- to three days between PSI- and PSII light source) if grown under fluctuating light. A more recent study shows that upon light intensity changes TAP38/PPH1 phosphatase is crucial in preventing state transition to maintain the excitation balance between PSII and PSI, or in other words, plants actively avoid state transitions upon changes in light intensity (Mekala *et al.*, 2015). The challenge of understanding the acclimation processes of plants to fluctuating and dynamic light conditions has been approached recently by several researchers from a physiological side in different plant species like Arabidosis, rice, and wheat (Mishra *et al.*, 2012; Allahverdiyeva *et al.*, 2015; Retkute *et al.*, 2015; Yamori, 2016; Yamori *et al.*, 2016; Vialet-Chabrand *et al.*, 2017; Annunziata *et al.*, 2018; Huang *et al.*, 2018; Kaiser *et al.*, 2018; Matsubara, 2018; Matthews *et al.*, 2018; Slattery *et al.*, 2018; Townsend *et al.*, 2018; Schneider *et al.*, 2019).

Variation in light intensities required Arabidopsis to invest into regulatory processes like NPQ to retain their competitiveness and to avoid yield loss (Kulheim *et al.*, 2002). Field grown *Arabidopsis thaliana* genotypes expressed substantially different phenotypes in comparison with the same genotypes grown in indoor experiments under low, normal and high light (Mishra *et al.*, 2012). On the one hand, leaf area, length and width and total chlorophyll amount of field-grown plants were substantially lower than of any indoor grown plants despite different continuous light regimes (Mishra *et al.*, 2012). A similar reduction was found for Arabidopsis plants grown under artificial fluctuating light (Vialet-Chabrand *et al.*, 2017; Schneider *et al.*, 2019). On the other hand, the chlorophyll a/b ratio was significantly increased as well as VAZ pigments (violaxanthin,

antheraxanthin and zeaxanthin) of the xanthophyll cycle pool, which reached 50% higher levels than indoor grown plants (Mishra *et al.*, 2012). Although F_v/F_m ratio, a proxy for the photosynthetic efficiency of photosystem II, did not significantly differ from field-grown plants to indoor plants under low light, non-photochemical quenching potential of field-grown plants exceeded indoor plants under high light by 50% (Mishra *et al.*, 2012). In comparison with normal light condition the F_v/F_m ratio was significantly reduced under fluctuating light conditions (Mishra *et al.*, 2012; Schneider *et al.*, 2019). In contrast to C4 crops and *Arabidopsis thaliana*, in C3 crop species like rice and wheat, F_v/F_m values were not significantly reduced when grown under fluctuating in comparison with constant light conditions (Kubásek *et al.*, 2013; Yamori *et al.*, 2016).

Under natural conditions, Arabidopsis extensively invests in protective measures for its photosystem, which results in overall less biomass production. Different light regimes also result in a different composition of the metabolome. Johansson Jänkänpää *et al.* (2012) grew *Arabidopsis thaliana* Col-0 in low-light, normal-light and high-light conditions and the metabolome of each growth regime was fully discriminable from the other two in a principal component analysis (PCA). Furthermore, the metabolome is greatly responsive to the transfer of plants grown under normal-light conditions to field conditions (Johansson Jänkänpää *et al.*, 2012). In an untargeted lipidomic approach a strong change of the lipidome of *Arabidopsis thaliana* was discovered in response to different light regimes (Burgos *et al.*, 2011). Metabolic snapshots monitored during a three day time course, had each a unique fingerprint which was clearly separable by PCA (Johansson Jänkänpää *et al.*, 2012). In other species like rice, wheat, *Setaria macrostachya*, *Amaranthus caudatus*, and *Celosia argentea* dry matter accumulation was also significantly reduced when grown under fluctuating light conditions (Kubásek *et al.*, 2013; Yamori *et al.*, 2016).

Fluctuating light conditions, or better natural light conditions, demand a great deal of plasticity from the plant to maintain balanced growth. Under continuous, artificial illumination, this adaptive machinery is barely required and masks differences in natural fitness of plants. Modern phenotyping facilities need to be adapted to measure the phenotypic plasticity of plants under changing environmental influences to get a better understanding how plants will react to short term and finally long term changes with climate change in mind.

As the high-throughput phenotyping systems at the IPK are now fully upgraded to discriminate subtle phenotypic trait differences under controlled conditions, the next improvement will have to approximate natural growth conditions. This would require the Illumination to be optimized to enable rapid changes and higher output. Currently, the Illumination can switch between on/off in about 10 - 15 minutes. Fluctuating light conditions would require a shift of seconds to mimic a moving canopy or cloud cover. Following the average of 50 light/dark transitions per day (Salisbury *et al.*, 1968) and a typical illumination time of 12 hours per day, the illumination of the system would need to adjust around every 14.4 minutes, which could already be provided by the growth chamber, ignoring the fact that cloud cover would alter light intensities immediately, which cannot be realized by the lamps. However, the majority of sunflecks, due to a moving canopy e.g. of soybean, have lengths of < 1 s and a frequency of 0.16 Hz (Pearcy *et al.*, 1990). In maize, the most rapid fluctuations did reach ten cycles per second (Desjardins *et al.*, 1973). The number of perceived sunflecks is furthermore strongly dependent on the architecture of the canopy of a crop species. During the same meteorological conditions in a six hour period, durum wheat experienced 2,606 and white mustard only 213 sunflecks, while the characteristic duration of less than five seconds and the average irradiance increase < 350 µmol m^{-2} s^{-1} was similar between the two species (Kaiser *et al.*, 2018).

To mimic such natural conditions, an upgrade of illumination to light-emitting diodes would be a feasible solution. An advantage of LEDs over other forms of illumination is the immediate controllability of the lamps. Furthermore, LEDs are on a par with gas discharge lamps regarding the low waste heat and efficacy (lm/W) and are expected to outperform any other form of illumination in the near future (DOE, 2013). Disadvantages are the limited spectral range, especially in the lower energy range per photon. However, to fully upgrade the illumination of the IPK automated non-invasive plant phenotyping system with modern LED technology would be costly. A cost reduced option could be provided by white LED light bars which could homogeneously illuminate the growth area given a high enough density. They have a continuous spectrum between 400 nm and 750 nm (between 3.09 eV and 1.65 eV per photon), which would be sufficient to emit photosynthetic active radiation. Those LED light bars can be coupled to time relays and turned on or off following a random or a predefined pattern supplementing the main illumination. Natural conditions could be simulated, mimicking shadeflecks or sunflecks. To mimic

shadeflecks in a green house, an automated, probably very complicated apparatus would be required to simulate cloud cover. In a more simple approach the illumination would be set to low light conditions, mimicking continuous cloud cover, and then temporarily light intensity would be increased to simulate sunflecks as described in Tikkanen *et al.* (2010). An upgraded IPK automated non-invasive plant phenotyping systems would image in visible light and measure the fluorescence of the plants (Junker *et al.*, 2015). The Integrated Analysis Pipeline (IAP) could subsequently be used for high-throughput analysis of the different recordings to extract parameters like shape and height, which are highly correlative to fresh and dry weight, and colorimetric parameters, which correspond to levels of fluorophores and data related to the water content (Klukas *et al.*, 2014; Junker *et al.*, 2015).

1.4 Performance of crops in dynamic environments

In the field crop species are exposed to different weather conditions, seasonal changes and biotic stress, which can negatively impact yield and plant performance with increasing temperatures due to climate change (Lobell and Field, 2007). Dynamic environments of field conditions are not only characterized by fluctuating light, but also by dynamic shifts in air and soil temperatures, air humidity, wind speed, water and nutrient supply, and soil compaction (Poorter *et al.*, 2016). Any of these parameters can have a substantial impact on yield performance and the phenotype of a crop species.

Maximum and minimum temperatures during the day or the growing season are strong determining factors for crop cultivation as every species responds to a different optimal temperature (Hatfield and Prueger, 2015). Plants adapted to species-specific temperature optima in regard to developmental speed and leaf elongation during evolution (Parent and Tardieu, 2012). Parent and Tardieu, (2012) found that crop breeding history was not successful so far in changing a species-specific temperature response. Yield performance in rice, maize, soybean, and cotton was reported to be highly susceptible to higher than optimal temperature regimes and yield performances were found to respond negatively to current and even more to predicted temperature increases due to climate change (Peng *et al.*, 2004; Schlenker and Roberts, 2009). Global warming will put pressure on breeders and researchers to immediately start to work on

methods to study and breed better adapted genotypes, which can better handle higher temperature regimes (Challinor *et al.*, 2016).

Outdoor-grown crop plants experience wind exposure, which is mostly absent in greenhouses or growth chambers. Wind drag and movement causes plants to respond with decreased maximum photosynthetic rate, increased root to shoot ratio, and plants are less predisposed to bending or lodging (Cordero, 1999). Exposure to wind causes maize seedlings to respond with growth retardation (Jaffe *et al.*, 1985), and can lead to increased lignification of cells and reinforcement of stems in non-woody plants to cope with wind drag (Gardiner *et al.*, 2016). Stronger wind can eventually cause lodging and significantly reduce grain yields dependent on the developmental stage of the plant at the time receiving the stress (Carter and Hudelson, 1988; Mi *et al.*, 2011).

Limiting factors for crops in fields, water availability and prolonged periods of drought, can significantly reduce the potential of plants to reach maximum height or productivity (Deikman *et al.*, 2012), which lead to efforts to dissect drought response in crops in greenhouses (Chen *et al.*, 2014; Neumann *et al.*, 2015) and in field trials (Chapuis *et al.*, 2012). While in one year water can be a limiting factor, in other years heavy rainfalls can cause waterlogging, which can also reduce plant height and ultimately yield performance (Ren *et al.*, 2014).

The observation of negative yield performances is not affecting all continents and regions equally. Yield of the same crop species, e.g. maize, can decrease in southern parts of Europe and Africa, increase in temperate Europe and northern USA, and stagnate in east Asia, all in a relatively short period from 1961 to 2008 (Ray *et al.*, 2012). This will require screening of germplasm for alleles, which can help crops to better adapt locally to future climate conditions (Varshney *et al.*, 2018). Furthermore, a better management of crop varieties in general is needed to use the best adapted varieties under a given climate scenario (van Etten *et al.*, 2019).

In contrast to field conditions, most plants in growth chambers were cultivated under constant temperatures near to optimal growth and low or intermediate light levels without regarding the substantial natural diurnal variation of either parameter found in nature (Poorter *et al.*, 2016). The strongest effect on biomass differences was caused by two parameters, daily integral of light and temperature, which can contribute between 60% and 600% in biomass increase, respectively, between different experiments and growth scenarios (Poorter *et al.*, 2016). There is a strong

discrepancy between field conditions and controlled indoor conditions, under which a lot of gene discovery took place.

Genotype x environment (G x E) interaction causes plants to respond with a great variety of phenotypes to different environments, which has led plant breeders to evaluate the breeding value of their crop varieties in multi-environment trials over several years (Cooper and DeLacy, 1994; Ceccarelli, 2015). In early to mid-phases of breeding cycles and to screen plant germplasm for beneficial agricultural traits, high-throughput phenotyping platforms are used to evaluate the performances of plant and crop varieties under distinct environmental scenarios (Fiorani and Schurr, 2013; Ceccarelli, 2015). In recent years many high-throughput phenotyping facilities have been established and used to study a diverse set of crop species like rice (Yang *et al.*, 2014; Schilling *et al.*, 2015), maize (Muraya *et al.*, 2017; Zhang *et al.*, 2017), barley (Chen *et al.*, 2014; Neumann *et al.*, 2015), canola (Pommerrenig *et al.*, 2018; Knoch *et al.*, 2019), lentils (Muscolo *et al.*, 2015), and of course the model plant Arabidopsis (Zhang *et al.*, 2012). The systems were highly productive to dissect genetic variation and characterize allelic variants. But all of the aforementioned systems were integrated into conventional growth chambers or glasshouses. Phenotyping of field stands has been approached differently and relies heavily on smaller mobile devices, which cannot perform phenotyping in comparable precision and throughput with the indoor systems but were still able to contribute significant knowledge (Araus and Cairns, 2014). Manual phenotyping of roots of field grown maize plants has been sped up to a respectable high-throughput level by creating the field of shovelomics (Trachsel *et al.*, 2011). Easily available tractors were fitted with sophisticated camera systems and together with an appropriate software environment were generating high-throughput phenotyping images of 1.5 ha field area in a respectable time frame of only three hours (Salas Fernandez *et al.*, 2017). The throughput is even higher with unmanned aerial vehicles, which can cover greater areas and phenotype plant height and color related traits with high heritability for plots in a field stand (Madec *et al.*, 2017; Yang *et al.*, 2017). Water use efficiency and canopy conductance can be measured outdoors for a whole field by wireless sensor networks providing real time information for crop management (Jones *et al.*, 2018). All of these field phenotyping technologies are using smart approaches to respond to challenges like throughput and precision, but they are all limited to dissect G x E interaction under the prevailing environment present at the point in time and space of the respective

experiment/trial (Araus and Cairns, 2014). Year to year environments in a field site can be very diverse and result in lower year to year than lab to field correlation making it nearly impossible to repeat a field experiment under similar conditions (Poorter *et al.*, 2016).

The IPK Plant Cultivation Hall (PKH) was designed and constructed to dissect genetic variants for agricultural traits under field-relevant environments in a controlled and repeatable way. The PKH provides an infrastructure for the plant scientific community to study most crop or model species in an environment defined by the researcher. According to the measurements of Kulheim *et al.*, (2002) and Slattery *et al.*, (2018), 500 – 2,000 µmol m^{-2} s^{-1} are enough to mimic field conditions. Temperature regimes can be programmed to range from 5 up to 40°C, which is covering the growth response curve of major crop species (Hatfield and Prueger, 2015). To simulate predicted CO_2 levels of > 800 ppm, which according to Pan *et al.*, (2015) will occur on our planet in the 2090s as a consequence of industrial use of fossil fuels, the PKH can elevate atmospheric CO_2 levels in compartment one from ambient 400 ppm up to 1,200 ppm.

1.5 Rapid identification of induced mutations

High-throughput phenotyping of natural accessions or inbreds under fluctuating light and dynamic environments combined with quantitative trait locus (QTL) mapping or genome-wide association (GWA) studies leads to the discovery of associated alleles. Next-generation sequencing technologies can be used for genomics-assisted breeding of quantitative trait loci or marker-trait associations gained from QTL or GWA studies (Varshney *et al.*, 2018). There is a great need to search for alleles from germplasm that have a beneficial effect on plant resilience under a climate change scenario (Varshney *et al.*, 2018). However, to precisely determine the underlying cause of a certain phenotype requires researchers to use reverse genetic, e.g. screening single gene knock-out mutants for phenotypes, and forward genetic, e.g. identifying the genetic cause for a phenotype found in a mutant screen, approaches. Mutant plants for an associated candidate gene can be further studied to support the results of the high-throughput phenotyping approach. Furthermore, mutant populations can be screened under dynamic environments to detect beneficial phenotypes, which consequently require researchers to find the responsible mutated allele. At the IPK, an ethyl methanesulfonate (EMS)-treated maize population was created.

THE FOLLOWING TEXT IS CITED FROM MY PUBLICATION (Heuermann *et al.*, 2019)

"Maize (*Zea mays* L.) is among the top yielding crops (> 1 gigaton) worldwide (FAO, 2017). Despite its agronomic importance and the availability of a high quality reference genome sequence, only a very restricted fraction of maize genes have been assigned to a biological role through direct experimental evidence, either through classical forward genetics like chemical mutagenesis (Gallavotti *et al.*, 2010), radiation mutagenesis (Kynast and Riera-Lizarazu, 2011), and transposon mutagenesis (Williams-Carrier *et al.*, 2010) or through reverse genetics approaches like TILLING (Till *et al.*, 2004; Weil and Monde, 2009). More recently, the function of several genes has been studied with the genome editing tool CRISPR/Cas9 (Svitashev *et al.*, 2015; Qi *et al.*, 2016; Char *et al.*, 2017). Conventional techniques like quantitative trait locus (QTL) mapping, genome-wide association studies (GWAS) and nested association mapping (NAM) are routinely used to dissect complex traits like (seed) yield, plant size or architecture, pathogen resistance and control of metabolic pathways (Peiffer *et al.*, 2014; Ding *et al.*, 2015; Tang *et al.*, 2015). However, maize barely profited from next-generation mapping techniques (Nannas and Dawe, 2015), although forward genetics have regained momentum over the last years in many species. With the introduction of next-generation sequencing (NGS) technologies, ethyl methanesulfonate (EMS)-induced mutations in model organisms like *Caenorhabditis elegans* (Sarin *et al.*, 2008), fission yeast (Irvine *et al.*, 2009) and *Drosophila melanogaster* (Blumenstiel *et al.*, 2009) were identified without the need to create classical mapping populations. Emerging technologies like bulked segregant analysis enabled combined mapping and mutant identification by sequencing pooled DNA of up to 500 F_2 plants (Schneeberger *et al.*, 2009). Zuryn *et al.* (2010) demonstrated that it is possible to identify EMS-induced mutations by sequencing DNA from three *C. elegans* EMS mutant populations after 4-6 rounds of backcrossing. Hitherto, the success of these methods depends on a large number of individuals or several rounds of backcrossing. Austin *et al.* (2011) omitted backcrossing and detected mutations in EMS treated *Arabidopsis thaliana* plants by sequencing a small pooled F_2 population and subsequent filtering for homozygous regions supporting the detection of recessive mutations only. The idea to directly compare corresponding plants with mutant and wild-type phenotypes to reveal the causal mutation has been approached in several ways in Arabidopsis, rice, maize and barley (Abe *et al.*, 2012; Hartwig *et al.*, 2012; Lindner *et al.*, 2012; Liu *et al.*, 2012; Fekih *et al.*, 2013; Nordström *et al.*, 2013; Tabata *et al.*, 2013; Mascher *et al.*, 2014). However, each of these approaches involved preceding backcrossing programs or analysis of large numbers of F_2 individuals for the assembly of bulks. Addo-Quaye *et al.* (2017) omitted

backcrossing and sequenced pools of *Sorghum bicolor* EMS mutant and wild-type individuals to filter for homozygous SNPs causing non-synonymous amino acid exchanges in the mutant population and could identify a recessive mutation. Although they were able to find a homozygous SNP in a protein coding region, their approach potentially missed mutations not caused by SNPs like insertions or deletions and would furthermore not have enough discriminative power to detect mutations in a heterozygous state, which is relevant for homozygous lethal alleles.

In crops like maize, it is very time consuming to carry out several rounds of backcrossing and the production of large bulks of phenotypically well-defined individuals may be limited by several constraints including appropriate cultivation area. Mutagenesis with EMS causes untargeted point mutations in the genome of the treated organism. Phenotypically selected individuals of an appropriately mutagenized population may suffer from low fecundity caused by a heavy genetic burden (high mutational load). Low numbers of mutant offspring, however, prevent immediate application of the aforementioned approaches, which rely on the assembly of large phenotypically defined pools of plants.

To overcome these limitations, we mapped EMS-induced mutations in populations of small sets of individually sequenced M_2 maize mutants and their corresponding wild-type siblings. Knowledge about the individual zygotic state of the mutation, which we gained by analyzing the segregation of the mutant phenotype in their M_3 offspring, enabled us to directly filter for the causative SNPs without the need of backcrossing. Thereby, we successfully isolated a recessive mutation causing a dwarf structure and a semi-dominant mutation responsible for a pale green leaf color using only sixteen individual plants for each mutation."[1]

END OF CITATION (Heuermann *et al.*, 2019)

1.6 The aims of this study

Despite the high quality annotation of the *Arabidopsis thaliana* genome, the annotation of genes with a functional response to diverse environmental stimuli is necessary and still an ongoing process (Cheng *et al.*, 2017), which is even more important in crops as gene annotation in crop species relies heavily on the annotation accuracy in model plants. The recent advancements in high-throughput phenotyping systems in greenhouses and growth chambers at the IPK

---

[1] pp.1-2, Heuermann *et al.*, (2019)

Gatersleben allowed to successfully dissect genetic variance under stable but artificial climate conditions and to gain valuable insights in regulation of genes and temporal dynamics of QTL action (Chen *et al.*, 2014; Neumann *et al.*, 2015; Junker *et al.*, 2015; Muraya *et al.*, 2017; Knoch *et al.*, 2019). To understand the regulatory response of plants to fluctuating light and dynamic environmental conditions in general, phenotyping needs to occur under natural-like environmental conditions and new tools need to be developed to speed up the identification of gene function. This work tackles these challenges by three different approaches.

The main aim is to study Arabidopsis' response to fluctuating light conditions by high-throughput phenotyping 384 natural accessions and to study the genome-wide association of phenotypic variation with fluctuating light specific gene regulation. Fluctuating light conditions will be achieved by upgrading the high-throughput phenotyping facility for small plants at the IPK with ancillary light sources and shading meshes of different opacities.

A supporting methodical advancement will explore the performance of IPK's Plant Cultivation Hall to simulate dynamic, field-relevant environments in a first benchmark. In three consecutive years a phenotypically diverse set of maize inbreds will be manually phenotyped twice a week and the hourly median weather data of these years will be replicated in the PKH. The same set of plants will then be phenotyped in parallel in the Plant Cultivation hall under field-relevant climatization and in a large glass house under standard greenhouse conditions for the same traits to compare plant performance with the three field experiments.

In a second methodical advancement a rapid mapping procedure will be developed to speed up the identification of the underlying mutated gene function in EMS-induced mutations in *Zea mays* directly in the M_2 generation without the need for prior backcrossing or for pools of large numbers of individuals.

2. MATERIAL AND METHODS

2.1 Plant material of Arabidopsis for the fluctuating light experiment

The plant material used in this study consisted of prior synchronized seed stocks from plants grown together in one cultivation, from an accession panel of 384 *Arabidopsis thaliana* accessions (Table S1), which partially originated from the RegMap panel genotyped with the Affymetrix 250k SNP array (Horton *et al.*, 2012). Analysis of the accessions using 140 SNP markers (Platt *et al.*, 2010) revealed that 58 of the accessions used in the experiments did not have public 250k SNP data available. DNA of the missing accessions was hybridized to the Affymetrix 250K SNP Array (performed by DNAVision, Charleroi, Belgium), and raw data subjected to the analysis pipeline established by Nordborg and colleagues (Atwell *et al.*, 2010) to ensure compatibility between the data sets. Version 75 was used for analyses (Horton *et al.*, 2012). The SNP data and synchronized seeds of the accession panel were kindly provided by Dr. Rhonda Meyer.

Finally, genotype data of 382 accessions were available for genome-wide association analysis (for more details see 3.1.2). Principal component analysis was performed on zero centered SNP data with the *prcomp* function in R (R Core Team, 2018) to check for population structure (see 3.1.2, Figure 1_2) and to correct for population structure in the GWAS model (see 2.7). The geographical location of the accessions was plotted using the R package *maps* (Becker *et al.*, 2018).

2.2 Lighting installation

Both sides of the IPK automated non-invasive plant phenotyping system for small plants (LemnaTec Scanalyzer system, (Junker *et al.*, 2015)) were upgraded with a rectangular aluminum frame at a height of 59 cm over the plant growth area. The fluctuating light side was upgraded with twenty 48" LED 5000K Cool White Light Bars (BML Horticulture Austin TX 78704) with a 60° beam angle. Furthermore, shading screens (Reimann, D-48282 Emsdetten) of different opacities, on the fluctuating light side a PyroSilver Ultra grey shading screen and on the constant side a Clear Day Screen, were attached to reduce the main illumination by 74% (to 65 µmol m^{-2} s^{-1}) and 24% (to 190 µmol m^{-2} s^{-1}), respectively (Figure 1_1). The differential was filled by turning the LED light bars (300 µmol m^{-2} s^{-1}) on and off with intervals of different length controlled by DESIGO software from Siemens. The homogeneity of illumination was checked with three LI-COR LI-190R sensors (LI-COR

Environmental – GmbH, D-61352 Bad Homburg), which were each positioned in a carrier of the IPK automated non-invasive plant phenotyping system for small plants and run through every carrier position to deliver a heat map of light intensity with three replicates (Figure S 1_1). Four modules (module 1: 1 min on/ 1.4 min off, module 2: 2.5 min on / 3.5 min off, module 3: 5 min on/ 7 min off, and module 4: 10 min on/ 14 min off) were each run for 120 minutes and repeated two times each day in a randomized order, together 16 hours of light. Three exemplary days, each with an individually randomized order of modules, were created and then randomized over the total growth period. Storage limitations in the lighting control program did not allow for further randomization. Spectral range of both illumination installments were checked at an exposure time of three milliseconds and show only minor differences between the metal-halide lamps and the fluctuating side supplied with additional LED light bars (Figure S 1_1; USB2000+XR1-ES, Ocean Optics, D-73760 Ostfildern). For the control experiment the LED light bars and the main illumination were dimmed to emit continuous light. The left (with LEDs) and right side (without LEDs) of the IPK automated non-invasive plant phenotyping system for small plants now received mean values of 176 and 175.5 µmol m^{-2} s^{-1}, respectively.

2.3 High-throughput phenotyping

The IPK automated non-invasive plant phenotyping system for small plants was used for growing the plants in 12 well trays under standard cultivation conditions for 21 days as described in Junker *et al.*, 2015. For each light condition the full set of 384 accessions was replicated a total of six times, two times three replicates in a randomized block design, together 4,608 plants. Illumination schemes were controlled as aforementioned, two experiments with parallel fluctuating and constant light conditions (1646MH and 1703MH) and one control experiment with constant light in both sides (1806MH). Single seeds were sown out into the trays and transferred into the climatized growth chamber of the IPK automated non-invasive plant phenotyping system for small plants. After two days (one day in 1703MH) of continuous darkness at 5°C and 95% relative humidity, temperature was increased to 16°C during the day and 14°C at night under constant 75% relative humidity for another two days. From the 5th day on, temperature and relative humidity were elevated to their final day/night cycle of 20°C/18°C and 60%/75% for the rest of the experiment, respectively. The light programs were run since the third (second day after sowing in 1703MH) day after sowing (see 3.1.1). Top view images were generated in three experiments

(1646MH = 119,808 images; 1703MH = 110,592 images; 1806MH = 147,456 images) by the red, green, and blue (RGB) and the fluorescence cameras of the IPK automated non-invasive plant phenotyping system for small plants. Time point of germination after sowing, which was used as a random effect in the linear mixed model, was determined individually for each plant and defined as the day of first occurrence of the cotyledons, visible as red fluorescence in the images.

Kinetic chlorophyll fluorescence analysis through imaging was performed with the FluorCam device (Photon Systems Instruments, PSI, Brno, Czech Republic, http://www.psi.cz/) on four consecutive days of the 1703MH fluctuating light experiment. Protocols were modified from Tschiersch *et al.* (2017) to record all possible parameters and their derivatives. After one hour of lights turned off, dark-adapted plants were subjected to eight saturated light pulses (800 ms; 4,100 µmol m^{-2} s^{-1} white light) over the course of 145.8 seconds. After the first saturating light pulse, light was turned off (five seconds) to measure F_0, followed by the second saturating light pulse and actinic light to measure maximum fluorescence (F_m) followed by ten seconds dark relaxation to measure the Kautsky kinetics. For the F_m' quenching analysis white light and actinic light were turned on for 120 seconds and supplemented by six saturating light pulses every 20 s followed by another ten seconds dark relaxation measurements. All chlorophyll fluorescence parameters (see Table 1 in Tschiersch *et al.*, 2017) were extracted by the FluorCam software package using a custom mask, to calculate averages of all pixel values contributing to a single plant out of the twelve slot trays.

2.4 Image analysis with the Integrated Analysis Pipeline software

The top view images from the IPK automated non-invasive plant phenotyping system for small plants were analyzed and image-related traits were extracted with the Java-based Integrated Analysis Pipeline (IAP) following the protocol of Klukas *et al.* (2014). For each of the three experiments, the RGB images had to be slightly rotated, zoomed and shifted to align with the fluorescence images. For the 1646MH/1703MH/1806MH experiments, the RGB images were rotated by -1°/0.5°/-0.15° and x- and y-axis were zoomed to 109.5%/110%/109.2% and 99.2%/99.5%/99.1%, respectively. Furthermore the images were shifted by 13.5/13/-3.5 pixels on the x-axis and by 40.5/38/27 pixels on the y-axis. The analysis was performed on a desktop

machine with an Intel i7-3700 and 16 GB of Ram, of which 12 GB of Ram attributed to IAP running on 64 bit version of Java.

2.5 Data processing and best linear unbiased estimators

After the IAP analysis, raw data of phenotypic traits were outlier corrected by eliminating data points that surpassed or undercut the median ± two-fold standard deviation. Statistical groups for the outlier test were defined as data from each experiment (1646MH, 1703MH and 1806MH) for each individual trait at one time point over all genotypes and over both light conditions, fluctuating and constant. A Bayesian principal component analysis (PCA) on the whole dataset was performed using the *pcaMethods* package (Stacklies *et al.*, 2007), as it tolerates missing data, to visualize global differences between the two light conditions for all phenotypic traits at each time point. Data were Pareto-scaled and centered and Bayes imputation was run for 1,000 steps with an alpha threshold of 1e-04.

For all subsequent tests statistical groups were further split for their specific light condition. Best linear unbiased estimators (BLUEs) were calculated between seven overlapping time points of 1646MH and 1703MH assuming fixed effects of the genotypes of each statistical group fitting linear mixed models with the *lme4* package (Bates *et al.*, 2015) according to recent work by He *et al.* (2016). The model was defined in Eq.1, where y_{ijtmk} is the vector of phenotypic data of the ith genotype in the mth experiment, at the tth time point of germination after sowing and in the kth carrier of the jth replication, μ is the intercept term, α_i the genetic effect of the ith genotype, R_j is the effect of the jth replication, G_t the effect of the tth time point of germination after sowing, E_m the effect of the mth experiment, C_k the effect of the kth carrier, δ_{jt} the effect of the interaction of the jth replicate with the tth time point of germination after sowing, Z_{jm} the effect of the interaction of the jth replicate with the mth experiment, γ_{jk} the effect of the interaction of the jth replicate with the kth carrier and ε_{ijtmk} is the remaining residual. BLUEs were estimated treating α_i as fixed effects and all other effects as random effects. The model was kept maximal as suggested by Barr *et al.* (2013) to estimate effects of the position and the time point of germination.

Eq.1: $y_{ijtmk} = \mu + \alpha_i + R_j + G_t + E_m + C_k + \delta_{jt} + Z_{jm} + \gamma_{jk} + \varepsilon_{ijtmk}$

Eq.2: $R_n = \dfrac{\sigma_\alpha^2}{\sigma_\alpha^2 + \frac{1}{n}\sigma_\varepsilon^2}$

Eq.3: $y_{ijtnkm} = \mu + \alpha_i + R_j + G_t + D_n + L_k + B_m + Z_{jn} + X_{jm} + \delta_{jt} + \gamma_{jk} + \varepsilon_{ijtnkm}$

Repeatability of the phenotypic traits was calculated with Eq.2, where n is the number of replicates, σ_α^2 is estimated from α_i treated as random effect and σ_ε^2 is estimated from ε_{ijtmk} from Eq.1 according to suggestions for obtaining repeatability directly from variances and number of measurements (McGraw and Wong, 1996; Nakagawa and Schielzeth, 2010).

The data from the FluorCam measurement were analyzed following the same work-flow as for the IAP raw data. Data was outlier corrected and the BLUEs were estimated treating α_i as fixed effect and all other effects as random. The linear mixed model was fitted in Eq.3, where y$_{ijtnkm}$ is the vector of phenotypic data of the ith genotype at the tth time point of germination after sowing at the nth day of measurement and in the kth lane and in the mth block of the jth replication, μ is the intercept term, α_i the genetic effect of the ith genotype, R_j was the effect of the jth replication, G_t the effect of the tth time point of germination after sowing, D_n the effect of the nth day of measurement, L_k the effect of the kth lane, B_m the effect of the mth block, Z_{jn} the effect of the interaction of the jth replicate with the nth day of measurement, X_{jm} the effect of the interaction of the jth replicate with the mth block, δ_{jt} the effect of the interaction of the tth time point of germination after sowing, γ_{jk} the effect of the interaction of the jth replicate with the kth lane and ε_{ijtnkm} is the remaining residual. Repeatability of the phenotypic traits was calculated with Eq.2, where n is the number of replicates, σ_α^2 is estimated from α_i treated as random effect and σ_ε^2 is estimated from ε_{ijtnkm} from Eq.3. Data processing and the estimation of BLUEs were implemented in a self-made R script on a desktop machine with an Intel i7-3700 and 16 GB of Ram (R Core Team, 2018).

2.6 Linkage disequilibrium

Linkage disequilibrium (LD) in Arabidopsis was calculated for the whole genome and for each chromosome individually, taking the SNPs as markers. The 214,054 heterozygous and homozygous SNP markers (see details 3.1.2) were filtered to keep only the 109,178 homozygous SNPs using the read.snpData function from the *QTCAT* package in R (Klasen *et al.*, 2016). Pairwise r^2 between SNPs were computed for each chromosome individually with the *LDheatmap* package (Shin *et al.*, 2006). The Euclidean distance between markers on a chromosome was calculated using R (R Core

Team, 2018). The proposed model of Hill and Weir (1988) was used to fit the decay of r^2 with the pairwise marker distance in nucleotides.

The simplified equation Eq.4: $E(r^2) = \left[\frac{10+C}{(2+C)(11+C)}\right]\left[1 + \frac{(3+C)(12+12C+C^2)}{n(2+C)(11+C)}\right]$ was used as suggested according to more recent work (Remington *et al.*, 2001; Marroni *et al.*, 2011), where n is the effective population size (764 gametes of 382 individuals) and c is the recombination fraction between sites and $C = 4nc$. The arbitrary C is estimated fitting a nonlinear model using the *nls* function in R and starting with $C = 0.1$. The estimated C is than refitted into the equation to model adjusted LD values aligned for their Euclidian distance along the chromosome. LD values from the adjusted data of the Arabidopsis accession population were plotted against the pairwise marker distance in nucleotides. The intercept of the half maximum LD value with the distance is the half LD decay value of the population – here 3365 bp. Calculations were run using 60 cores in parallel on a server architecture running CentOS 7.4 with four Intel Xeon E7-8880 v3, together 144 threads, and one TB Ram on solid state drives. The LD analysis was self-implemented in R on the server architecture with the help of R Studio accessed via a web browser (RStudio Team, 2015).

2.7 Genome-wide association studies

Genome-wide association studies (GWAS) were performed with the *FarmCPU* package for R with a minor allele frequency (maf) cut-off of 0.01, which removed 276 SNPs, a Bonferroni corrected *p*-value threshold (0.05/214,051 SNPs), a maximum of 100 loops, and a kinship value determined by *FarmCPU* (Liu *et al.*, 2016). The first four principal components from the marker PCA, calculated with the *prcomp* function package within *GAPIT* (Lipka *et al.*, 2012), were used as covariates. GWAS was individually run for the BLUEs of fluctuating and constant light between 1646MH and 1703MH of every trait at each of the six overlapping time points, 5, 6, 7, 12, 13, and 15 days after germination. The association tests (a combination of 286 IAP traits, six time points, and two treatments; *c.* 3,500 tests) were performed on the same computational resources as the linkage disequilibrium analysis. The association table generated by *FarmCPU*, containing the *p*-value, the minor allele frequency, and the effect for each of the 213,775 homo- and heterozygous SNPs, with a minor allele frequency greater than 0.01, was used to generate a Manhattan plot and a Q-Q-plot using *rMVP* in R https://github.com/XiaoleiLiuBio/rMVP.

Phenotypic variance was estimated for each association test individually. A linear regression model was fitted for each tested trait, being the dependent variable, and the Bonferroni corrected significant markers of their respective association test as the explanatory variables. The variance components of the model were determined by performing an ANOVA on the fitted linear model. The variance component of each significant marker, fitted as an explanatory variable, was evaluated against the sum of the variance of the significant markers tested and the residual error in Eq.5: $PVE\% = \frac{\sigma_m^2}{(\sum_{m=1}^{i} m = \sigma_m^2 + \sigma_{mi}^2) + \sigma_\varepsilon^2} * 100$, where i is the number of significant markers, σ_m^2 is the variance of a significant marker fitted as fixed factor and σ_ε^2 is the residual error.

Shared traits between the fluctuating and constant conditions of both experiments, which contributed significant associations after a Bonferroni correction in either of the two light conditions, were considered for comparisons. Significant marker-trait associations were further filtered to be either fluctuating or constant light condition specific. SNPs from significant marker-trait associations were annotated against the Araport11 annotation, containing genes and transposons (Cheng *et al.*, 2017) with an Arabidopsis Genome Initiative identifier, or gene identifier, if the SNPs were intersecting with an annotated gene identifier 3,365 bp up- and downstream (one full LD block) of the SNP position. BLUEs from the FluorCam measurements of fluctuating and constant light treatment were analyzed by GWAS for each trait following the same procedure as described above. The analysis pipeline of the GWAS, the phenotypic variance explained (PVE) calculation, the filter, and the annotation were implemented using a self-made R script on the aforementioned server architecture (R Core Team, 2018).

2.8 Gene ontology filter for photosynthesis and light related marker-trait associations

Significant marker-trait associations were annotated with gene ontology terms using the *topGO* R package (Alexa and Rahnenfuhrer, 2016) against the TAIR genome locus (TAIR10_2017) annotation from the agriGO v2.0 database (Tian *et al.*, 2017). The *topGO* package exports two tables, one documenting the statistical analysis with the gene ontology (GO) ID and its corresponding GO term, and a second table containing the gene identifiers attributed to the GO IDs used to perform the statistical analysis. The output limit was increased to allow the software to cover all gene identifiers which were present in each of the light specific annotated marker-trait association

(MTA) lists. The tool was only used for GO term annotation, not for enrichment analysis. With these tables it was possible to apply a reverse search algorithm to search for specific GO terms, here any ontology containing the terms "photo" or "light", and extract the corresponding GO IDs. These GO IDs were then used to extract the gene identifiers from the second table, which were annotated with the given GO IDs filtered for the search words "photo" or "light". A search-string combination of the time point and the gene identifier were used to find the same combination in the MTA table. Finally, only MTAs with a repeatability value higher than 0.5, which matched the search string, were isolated.

2.9 Phenotyping in the field 2016, 2017, and 2018

Twenty *Zea mays* inbred lines from the dent and flint breeding pool (B106, B73, DK4676A, DKFBHJ, EC136, EC334, EP2, F922, N22, P074, P089, P135, P148, PB116, PHG50, PHG84, PHT77, S052, W117, and W33) and ten accessions (amplified after single seed descent, SSD, passage) from the Genebank of the IPK Gatersleben (ZEA 132, ZEA 324, ZEA 332, ZEA 3425, ZEA 3660, ZEA 3682, ZEA 391, ZEA 392, ZEA 399, and ZEA 851) were grown in double row plots in two replicates in a randomized block design, in total 60 double row plots, in the field site at the IPK Gatersleben from April to September 2016/17/18 (for more information about the *Zea mays* lines see Table S 2_1). The stand was protected by a border planting of a commercially available maize hybrid to mimic commercial field conditions. Each double row plot consisted of 18 plants, nine in each row, with dimensions of 80 cm length and 60 cm width (Figure 2_1). The in-between distance of plants in one row was 10 cm and the distance between rows was 60 cm. The outer plants per row were self-pollinated for line maintenance. Crosses were allowed between the same genotypes if self-pollination was not possible. To avoid cross pollination by other genotypes, the growing ears were covered in plastic bags before the emergence of silks. From the inner six individuals of each double row plot, five representative individuals were selected and marked with an identifier, a label on the ground. Those 300 plants were visually rated twice a week (Mondays and Thursdays) for plant height, growth and vegetative stage (Abendroth *et al.*, 2011), time point of tassel emergence and the final ear height and seed yield. The first batch of seed stocks for the experiment in 2016 was kindly provided by Dr. Astrid Junker and the German plant phenotyping network project. The seed stocks of the later experiments were self-produced by line maintenance from the previous cultivations.

Seeds were sown out in Jiffy strips (5 x 5 x 5 cm bio-degradable pots, Hermann Meyer KG, D-01683 Nossen) filled with substrate 2 (Easy Growing, Klasmann-Deilmann GmbH, D-49744 Geeste) and cultivated under 16 hour days at 24°C/18°C day/night in a greenhouse with auxiliary illumination for seven days. Plants were then transferred to a protected outside soil bed, which was protected from animal pests, to acclimate to outdoor climate conditions. After four days, plants were brought to the field site in the early morning hours (7 am) and transplanted together with the well-watered Jiffy into the field at a depth of 5 – 7 cm to fully cover the Jiffy. The field was sufficiently watered to ensure that plants are able to root through the bio-degradable Jiffy pots. Supplementary irrigation would have been applied in the following two weeks if not enough rain had fallen. Fertilization was applied around one month after transplanting each year. In 2016, an ammonium sulfate fertilizer with 21% nitrogen and 24% water soluble sulfur was given and, in 2017 and 2018 the fertilizer Nitrophoska 15+15+15 (+2S) 15% Nitrogen, 15% P_2O_5, 15% K_2O, 2% S (EuroChem Agro GmbH, D-68165 Mannheim).

2.10 Parallel cultivation and phenotyping in the Plant Cultivation Hall and a large greenhouse

In a parallel experiment (2018/19) in the Plant Cultivation Hall (German abbreviation PKH) and in a large greenhouse (German abbreviation GWH), a selection of the aforementioned lines were grown in double row plots for 115 days in a container based system and in single pots, respectively. Five of the 20 inbred lines (B73, N22, P148, S052, and W33) and six of the ten Genebank lines (ZEA 132, ZEA 324, ZEA 332, ZEA 3660, ZEA 399, and ZEA 851) were selected, in total 11 genotypes. Plants for both experiments were sown out together in the same small greenhouse used for the field experiments under the same conditions for seven days and then transferred into the PKH and the GWH for a four day acclimatization after which they were transplanted.

In the PKH, in one container the double row plot layout per genotype from the field was replicated, which led to five containers per genotype and to a total of 55 containers per 11 genotypes. Each container had a surface area of one square meter and a soil height of 100 cm, which equals a volume of one cubic meter or 1,000 L of soil volume. Containers were filled with a mixture of material to mimic field-like conditions. The lowest layer (2.5 cm) was a filled with

expanded polystyrene granules covered by a mat of fibers to enable drainage and prevent water logging. The next layer (8 cm) consisted of coarse gravel, followed by a third layer (40 cm) of coarse sand. A fourth layer (25 cm) was filled with loess from the field. The final layer (25 cm) was filled with soil (mixture: two parts compost, one part white peat substrate, one part sand). Five liters of water per container (three liters in the early morning, two liters at night) were supplied by a droplet irrigation system thrice a week, in total 15 L water per week. Fertilization was applied one month after transplanting by Nitrophoska 15+15+15 (+2S) 15% Nitrogen, 15% P_2O_5, 15% K_2O, 2% S (EuroChem Agro GmbH, D-68165 Mannheim).

In the GWH, plants were grown in single pots (5.5 L volume), which were arranged in ten rows of 11 plants, each with a fully randomized set. Pots were filled with the same soil as used in the final layer of the container based system in the PKH. Irrigation was applied daily. Fertilization was applied once a week with the irrigation in a concentration of 2‰ resolved in the irrigation water. Before tassel emergence, Hakaphos blue (15% N (4% Nitrate and 11% Ammonium), 10% P_2O_5, 15% K_2O, 2% MgO) and after tassel emergence, Hakaphos red (8% N (3% Nitrate and 5% Ammonium), 12% P_2O_5, 24% K_2O, 4% MgO) was used (COMPO EXPERT GmbH, D-48155 Münster).

After a four day acclimation phase in the PKH and the large GWH, plants were transplanted. Climate conditions in the PKH were set to mimic the average weather of the same time of the years 2016/17/18 starting with the beginning of May and ending with the end of August. Local weather was recorded hourly on the IPK grounds by a Lambrecht weather station with a 180° pyranometer 16130 0-2,000 w/m² and global range of 285 – 3,000 nm (Lambrecht Meteo GmbH, D-37085 Göttingen). The yearly average of all days in a given week in the years 2016/17/18 was used as a normal day. The yearly average of a specific date was calculated for each week (Figure 2_2). A representative sunny and cloudy day with high and low temperature was defined for a given week. Those three characteristic days for a week on a field in a given week were distributed over the course of one week (normal: Friday, Tuesday, Wednesday; cloudy: Sunday, Thursday; sunny: Saturday, Monday). Humidity in the field was stable over the growth periods (Figure S 2_1) in all years. Therefore air humidity in the PKH was set to 90% relative humidity during the night and to 40% and 50% at daytime on the corresponding normal/sunny days and cloudy days, respectively. Sudden changes in humidity values were avoided by programming ramps between the different levels. The illumination was controlled for normal, sunny, and cloudy days

individually, and light intensities and fraction of day covered under clouds were predefined (Table S 2_2, S 2_3, and S 2_4). Assimilation light sources were General Electric Lighting CMH400 light bulbs (General Electric Deutschland Holding GmbH, D-60313 Frankfurt am Main) complemented by six different LUMILEDS LED light bars (cool white 5700K L1C1-5770, deep red L1C1-DRD1, far red L1C1-FRD1, blue L1C1-BLU1, royal blue L1C1-RYL1, and cyan L1C1-CYN1; LUMILEDS, D-52068 Aachen). Ultra violet radiation (UVA 315 – 400 nm) was provided by NARVA LT 36W T8/ 010 UV (NARVA, D-09618 Brand-Erbisdorf).

At normal days, five hours of cloud cover were simulated (light intensity shifting on a sinus curve). Sunny days included only two hours of cloud cover, and cloudy days eight hours of cloud cover. Light intensities were highest on sunny days, only marginally lower on normal days (around 1,450 and 1,500 µmol m^{-2} s^{-1} on normal and sunny days, respectively, Figure S 2_2) and low on cloudy days (400 µmol m^{-2} s^{-1}, Figure S 2_2). Length of the day was changed every week by ten minutes earlier sunrise and ten minutes later sunset with a reciprocal pattern after the 21st of June. Moderate wind movement was simulated by five big fans (PBT/4-630/32, Soler & Palau Deutschland GmbH, D-64293 Darmstadt) on the left and right side of the PKH chamber. Wind direction was changed every two hours and wind intensity oscillated between 10% and 20% of maximal rotation speed during night. Wind intensity during the days was split into ramping between 30% and 50% in the morning and in the late afternoon, and into constant 50% during midday. At 100% rotations, fans reached a maximum volumetric flow rate of 16,450 m^3 h^{-1}, which translated into wind speeds of six m s^{-1} at six m distance and gradually decreased to two m s^{-1} at 28 m distance.

In the GWH conditions were set to 25°C/18°C temperature and 80%/60% relative humidity at day and night, respectively, with 16h/8h day night cycle. Assimilation light was switched on when the sensor on the roof dropped below 60 klx between 7 am and 11 pm. The light intensity in the GWH was around 250 µmol m^{-2} s^{-1}.

Plants were harvested in the PKH and GWH after their cobs fully developed and matured (dry yellow phenotype of the full plants). The cobs (and kernels) were separated from the vegetative matter and individually weighed and counted. Shoot material was dried and weighed. Total dry matter was calculated by adding the weight of the dried cobs to the weight of the dried shoot.

2.11 Best linear unbiased estimators of the phenotypic traits

Best linear unbiased estimators (BLUEs) were obtained for the parameters plant height, growth stage, and vegetative stage. The model for the PKH is defined in Eq.6, where y_{ijpkr} is the vector of phenotypic data of the ith genotype in the pth position in a container and in the kth column and the rth row of a container in the PKH of the jth replication, μ is the intercept term, α_i the genetic effect of the ith genotype, R_j is the effect of the jth replication, P_p the effect of the pth position in a container, C_r the effect of the rth row of containers in the PKH, S_k the effect of the kth column of containers in the PKH, δ_{jp} the effect of the interaction of the jth replicate with the pth position in a container, Z_{jr} the effect of the interaction of the jth replicate with the rth row of containers in the PKH, γ_{jk} the effect of the interaction of the jth replicate with the kth column of containers in the PKH and ε_{ijpkr} is the remaining residual. BLUEs were estimated treating α_i as fixed effects and all other effects as random effects.

Eq.6: $y_{ijpkr} = \mu + \alpha_i + R_j + P_p + C_r + S_k + \delta_{jp} + Z_{jr} + \gamma_{jk} + \varepsilon_{ijpkr}$

The model for the GWH was defined in Eq.7, where y_{ijr} is the vector of phenotypic data of the ith genotype in the rth row of pots in the GWH of the jth replication, μ is the intercept term, α_i the genetic effect of the ith genotype, R_j is the effect of the jth replication, C_r the effect of the rth row of pots in the GWH and ε_{ijr} is the remaining residual.

Eq.7: $y_{ijr} = \mu + \alpha_i + R_j + C_r + \varepsilon_{ijr}$

The model for the field experiments in 2016/17/18 is defined in Eq.8, where y_{ijr} is the vector of phenotypic data of the ith genotype in the pth position in the field plot of the jth replication, μ is the intercept term, α_i the genetic effect of the ith genotype, R_j is the effect of the jth replication, P_p the effect of the position in the field plot, δ_{jp} the effect of the interaction of the jth replicate with the pth position in the field plot and ε_{ijp} is the remaining residual.

Eq.8: $y_{ijp} = \mu + \alpha_i + R_j + P_p + \delta_{jp} + \varepsilon_{ijp}$

The repeatability of the phenotypic traits was estimated as described for Eq.2.

To evaluate the performance of each cultivation scenario, the BLUEs of the plant height of the three field experiments, the GWH, and the PKH were used to determine the best linear unbiased

predictors (BLUPs) of each individual cultivation scenario. The model of the BLUPs of all cultivation scenarios is defined in Eq.9, where y_{icd} is the vector of the phenotypic data of the ith genotype in the cth cultivation scenario in the dth measurement date in one cultivation scenario, μ is the intercept term, G_i is the effect of the ith genotype, C_c the effect of the cth cultivation scenario, D_d the effect of the dth measurement date in one cultivation scenario, δ_{ic} the effect of the interaction of the ith genotype with the cth cultivation scenario, y_{id} the effect of the interaction of the ith genotype with the dth measurement date in one cultivation scenario, and ε_{icd} is the remaining residual. BLUPs were estimated treating all effects as random effects and extracted using the *ranef* function in R.

Eq.9: $y_{icd} = \mu + G_i + C_c + D_d + \delta_{ic} + \gamma_{id} + \varepsilon_{icd}$

2.12 Thermal time normalization of the phenotypic data

The temperature in the PKH and the GWH was monitored every 15 minutes and aggregated to hourly means. The temperature at the IPK field sites was measured and logged hourly. In rare occasions of missing data, values from the nearest public weather station were accepted to fill the gaps.

Eq.10: $tt_n = (tt_{n-1}) + \dfrac{1}{24} * \dfrac{T*e^{\left(\frac{-\Delta H_A}{R*T}\right)}}{1 + \left[e^{\left(\frac{-\Delta H_A}{R*T_{20}}\right)}\right]^{\alpha\left(1-\frac{T}{T_o}\right)}} \Bigg/ \dfrac{T_{20}*e^{\left(\frac{-\Delta H_A}{R*T_{20}}\right)}}{1 + \left[e^{\left(\frac{-\Delta H_A}{R*T_{20}}\right)}\right]^{\alpha\left(1-\frac{T_{20}}{T_o}\right)}}$

Cumulative hourly thermal time (tt) was computed with Eq.10 for every hour of a day and for every day of a growing season on the field, in the PKH, and in the GWH following the suggestions of Parent and Tardieu, 2012. In Eq.10 n is the hour beginning with 1, T is the temperature in Kelvin, $\Delta H_A = 73900\,J*mol^{-1}$ is the enthalpy of activation specific for maize, T_{20} = 293K is the temperature at 20°C in Kelvin, T_o = 306.4K is the maize specific maximum temperature in Kelvin, $R = 8.314\,J*mol^{-1}*K^{-1}$ is the gas constant, and α = 3.5 is a unitless parameter from Parent and Tardieu, 2012. The cumulative thermal time after 24 h was used to aggregate the thermal time over the growing period. All data analysis was implemented in a self-made R script (R Core Team, 2018).

THE FOLLOWING TEXT IS CITED FROM MY PUBLICATION (Heuermann *et al.*, 2019)

2.13 Plant material for rapid mapping

"A *Zea mays* pollen mutagenized EMS population of the PH207 inbred line was generated with EMS concentrations (10 – 20 mM) according to (Neuffer and Coe, 1978). Among others, a *dwarf* (Figure 3_1 a) and a *pale green* (Figure 3_1 b) mutant were found during phenotypic screening using IPK's automated phenotyping platform for large plants (described in Junker *et al.*, 2015). The mutant lines (internally numbered: 1744 and 1754) were named according to their phenotype, *dwarf* and *pale green*. *Zea mays* plants were grown in the Scanalyzer system for large plants under standard conditions as described in (Junker *et al.*, 2015). After EMS mutagenesis, M_1 mutants were self-pollinated and transferred into a segregating M_2 population. M_2 individuals of those populations, which were subjected to sequencing, were self-pollinated and transferred into the M_3 generation. For allelism tests, published mutants of the *an1* locus *116G an1* (Emerson and Emerson, 1922) and *116GA an1-93W1189,* and the D8 locus *121C-D8-1* (Phinney, 1956) and *131D-D8-N1452*, *131E-D8-N1591* and *131F-D8-N2023* (Neuffer, 1990) were ordered from the Maize Genetics Cooperation Stock Center (http://maizecoop.cropsci.uiuc.edu/).

2.14 Chlorophyll fluorescence

Photosystem II (PSII) operating efficiency ($\Phi_{PSII}=F_q'/F_m'$) was measured via imaging with a custom FluorCam device (Photon Systems Instruments, http://www.psi.cz/) integrated in the IPK phenotyping platform for small plants (Tschiersch *et al.*, 2017). Photochemical quenching of fluorescence by open PSII centers (F_q') was computed by subtracting the maximal fluorescence (F_m') in the light adapted leaf from the fluorescence emission (F') from the light adapted leaf (F_q' = F_m'-F'; (Baker, 2008). Young *dwarf*, *DWARF*, *pale green* and *PALE GREEN* plants (V1 stage), five replicates each, were light adapted to 500 µmol m^{-2} s^{-1} PAR for 15-20 minutes prior to the measurement. Imaging was performed after additional 20 seconds of illumination with adaptation light and then photosystem II was saturated with an 800 msec light pulse (4100 µmol m^{-2} s^{-1}). Mean values of leaf area of top view images were taken as one measurement.

2.15 Isolation of DNA

Shoot material from seven *dwarf*, nine *DWARF*, eight *pale green*, and eight *PALE GREEN* M_2 plants, each exhibiting an identified phenotype, was quenched in liquid nitrogen and homogenized. DNA

was extracted using the Qiagen DNeasy Plant mini Kit according to the manufacturer's protocol. In total 32 DNA samples were used for Illumina Sequencing. For qPCR, whole DNA (nuclear and plastidic) was extracted from *pale green* and *PALE GREEN* M_3 mutants (6 days after sowing (DAS); whole plant) via a cetyltrimethyl ammonium bromide (CTAB) protocol developed in the GABI-Kat project (Rosso *et al.*, 2003).

2.16 Whole-genome-shotgun sequencing

Thirty two Illumina uniquely barcoded paired-end (2 x 100 bp) libraries were generated from sheared DNA (fragment size 400 – 500 bp) as described by (Meyer and Kircher, 2010). Prior to WGS, each sample was partially sequenced to determine its precise molarity. This information was used to optimize the loading of the flow cell for WGS. The 32 libraries were paired-end sequenced in 32 lanes on 4 flow cells on the Illumina HiSeq 2000. All DNAs were individually barcoded and paired end sequenced together in one lane of the Illumina Flow Cells, 32 lanes altogether. The data was made publicly available under the accession number PRJEB31849 at the European Nucleotide Archive http://www.ebi.ac.uk/ena/data/view/PRJEB31849.

THE SECTION 2.17 WAS WRITTEN BY DR. MARTIN MASCHER

2.17 SNP calling

Reads were mapped to the maize reference genome AGPv3 (Schnable *et al.*, 2009) with BWA mem version 0.7.10 (Li, 2013). The resultant SAM files were converted to BAM format with SAMtools (Li *et al.*, 2009) and sorted by position with Novosort (http://www.novocraft.com/products/novosort/). Multiple BAM files per individual were merged with Picard (https://broadinstitute.github.io/picard/). SNP calling was performed with SAMtools version 0.1.19 using the parameters "-D -Q13 -q 10" for the command "mpileup". The SNPs was filtered using the AWK script "gen_call.awk" (Mascher *et al.*, 2013). Genotype calls with coverage (DP) below ten or genotype quality (GQ) below ten were set to missing. Only SNP with a minor allele frequency of at least 10%, with up to 10% missing data and no more than 60% heterozygous calls were retained and imported into the R statistical environment (R Core Team, 2018) for further analysis. Allele frequencies were computed for plants with mutant and wild-type

phenotypes separately and plotted along the genome using standard R plot functions, excluding SNPs with a mutant allele frequency > 0.7 in both sets.

2.18 Segregation analysis

The individual sequenced M_2 plants were selfed and up to 40 kernels (M_3) from the offspring were planted (Table 3_2). Two weeks after sowing, plants were screened for *dwarf* and *pale green* phenotype, respectively. Thereby, the precise state of zygosity of the parents could be determined, based on the specific segregation pattern of the offspring.

2.19 Allelism test

Allelism tests tested whether the mapped *dwarf* allele is allelic to either the *an1* (Emerson and Emerson, 1922) or the D8 (Neuffer, 1990) locus on the long arm of chromosome 1. Therefore M_3 *dwarf* plants were crossed with *116G-an1*, *116GA-an1-93W1189*, 121C-D8-1, *131D-D8-N1452*, *131E-D8-N1591*, and *131F-D8-N2023* mutants with and without dwarf phenotype. Offspring was sown and segregation frequencies were analyzed.

2.20 qPCR of chloroplast genes

Genomic DNAs of six *PALE GREEN* and six *pale green* M_3 plants were diluted to a concentration of 50ng/µL. qPCR was run on a 7900HT Fast Real-Time PCR System (Applied Biosystems, 64293 Darmstadt, Germany) with a standard protocol. Reaction volume was 10µL (1 X *Power* SYBR Green PCR Master Mix (ThermoFisher Scientific, 64293 Darmstadt, Germany); 500 nM Primers; 50ng DNA Template). Primer sequences of the chloroplast encoded genes (*atpA*, *atpF*, *ndhA*, *ndhC*, *psbA* and *psbC*) and the nuclear encoded gene (*HMG-1*) were designed according to the publication by (Udy *et al.*, 2012). qPCR was performed in six biological replicates, each consisting of three technical replicates for each primer tested. Delta Ct values (ΔCt) were computed by subtracting the arithmetic mean of the technical replicates of the nuclear encoded HMG-1 gene from the arithmetic mean of the gene of interest (GOI) *pale green* and GOI *PALE GREEN*, respectively (ΔCt = Ct GOI *pale green/PALE GREEN* – Ct HMG-1). Subsequently, delta delta Ct values ($\Delta\Delta$Ct) were computed (mean ΔCt *pale green* – mean ΔCt *PALE GREEN*) and transformed to a linear scale ($2^{-\Delta\Delta Ct}$). An unpaired *t*-test decided whether differences between the ΔCt *PALE GREEN* and *pale green* levels for each tested gene were significant.

2.21 Deep candidate resequencing

De novo assembled NGS data of a *DWARF* and a *PALE GREEN* plants was used as a template for primer design. Nested primers were designed to amplify the *an1* and *w2* locus (for primer sequences see Table S1). PCR fragment matching the expected size were purified using the QIAquick Gel Extraction Kit and sequenced by LGC Genomics GmbH. Sequences were aligned and analyzed with R statistical environment (R Core Team, 2018) and SeqMan Pro from DNASTAR Inc.

2.22 Analysis of structural variation

DNA sequences 5 Mb and 0.25 Mb up- and downstream of each candidate gene from B73_AGPv3 and PH207 were downloaded from the MaizeGDB (Andorf *et al.*, 2016) and structural rearrangements between the two ecotypes were visualized by creating dot plots using the Gepard software version 1.40 (Krumsiek *et al.*, 2007). "[1]

END OF CITATION (Heuermann *et al.*, 2019)

---

3. RESULTS

3.1 Arabidopsis' response to fluctuating light treatment

The main aim of this study was to record the phenotypic response of a panel of natural *Arabidopsis thaliana* accessions grown under fluctuating light conditions, to quantify phenes, and to associate those phenotypic differences with genetic loci. Plants were cultivated under two different illumination schemes (constant and fluctuating light treatment) in two independent experiments and were high-throughput phenotyped on a daily basis using IPK's automated non-invasive plant phenotyping system for small plants. The Integrated Analysis Platform (IAP) was used to extract 286 image derived traits covering architectural, color-related, and fluorescence-related phenotypic features like leaf area or number of pixels in color classes (Klukas *et al.*, 2014). Best linear unbiased estimators (BLUEs) of the phenotypic traits of the two independent repetitions of the experiments were subjected to genome-wide association studies (GWAS). Comparisons between the marker-trait associations (MTAs) for plants grown under constant and fluctuating light conditions revealed associations specific for Arabidopsis' responses to fluctuating light.

3.1.1 Illumination set-up for fluctuating light conditions

The IPK automated non-invasive plant phenotyping system for small plants was built into a growth chamber that was designed to tightly control light quality and intensity and to deliver repeatable climate conditions on both sides of a growth chamber (Junker *et al.*, 2015). Thus, to study the phenotypes of an Arabidopsis accession panel under fluctuating and continuous light condition in parallel, a modification to the illumination system was necessary (Figure 1_1). The intensity of the metal-halide lamps of the illumination system of the IPK automated non-invasive plant phenotyping system for small plants cannot be individually controlled for the left and right half, instead every shift in light intensity level is applied to the whole growth area. Therefore a shading system with two meshes of different opacities was installed to allow only a fraction of the light to pass the shades (Figure 1_1 a, b). The system was separated into one area, left half, subjected to fluctuating light conditions (Figure 1_1 c) and another area, right half, with continuous light conditions (Figure 1_1 d). The shading meshes on the left side allowed 26% and on the right side 76% of the intercepted light of the main illumination to pass the mesh, creating a differential in

light intensity, which was necessary to apply fluctuating light treatment. LED light bars were installed under the mesh and above the growth area on the left side, which could cycle between light intensities and on/off durations in a latency free control scheme.
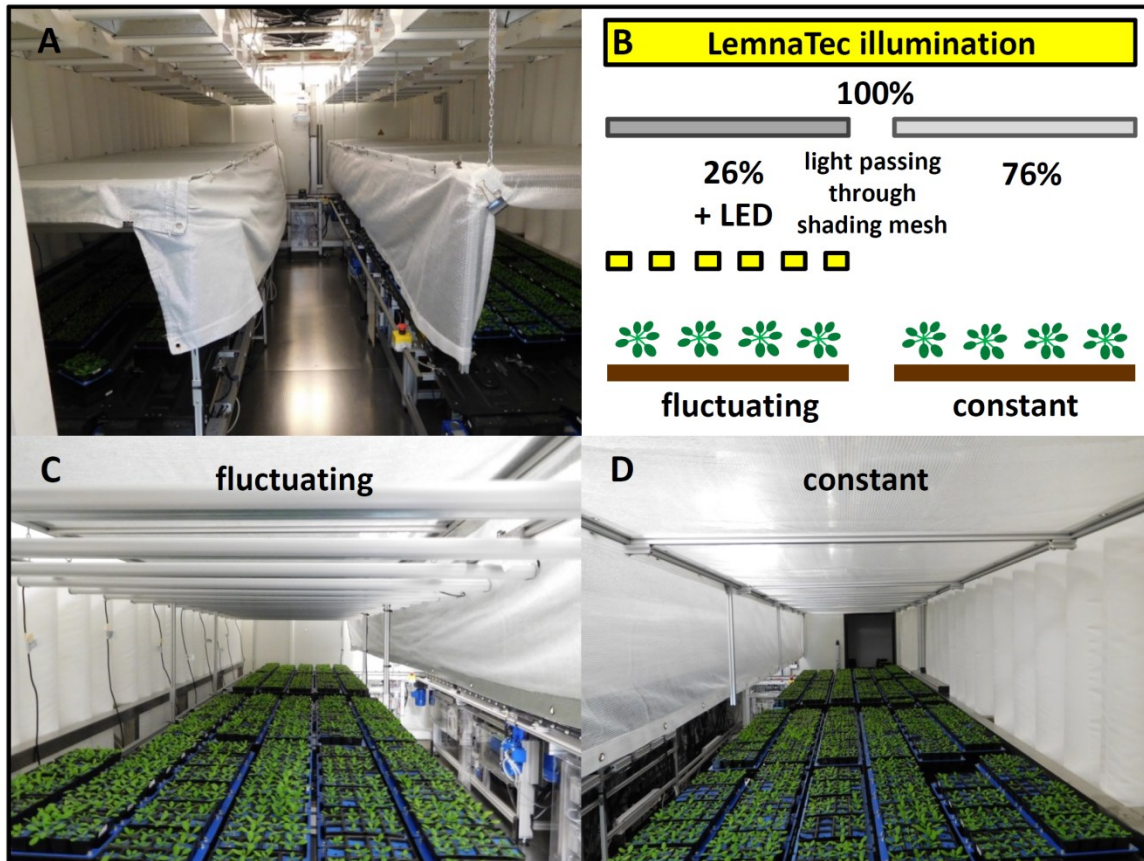
Figure 1_1: Illumination scheme of fluctuating light condition; (**A**) view of the shading installation in the growth chamber – left fluctuating and right constant light condition, (**B**) schematics of the installation; on the fluctuating side 26%, and on the constant side 76% of the incoming light passes the shading filter; the differential is filled by pulsing the LED light bars, (**C**) view under the shading mesh of the fluctuating side; LED light bars emit light in a 60° angle, (**D**) view under the shading mesh of the constant side.

Four modules were defined to control the LED light bars (Table 1_1) with the scope to expose the plants to different lengths of high light alternated by their respective low light period so that each module was an even fraction of 120 minutes. The four fluctuating light modules (module 1-4, Table 1_1) were repeated twice and randomly distributed over the 16h light period; thereby each module was running a total of 240 minutes per day. Trays were transported on conveyor belts and needed to traverse the whole growing area to reach the phenotyping chambers. Therefore, phenotyping was carried out during the dark periods to avoid trays from one lighting condition

being exposed to another lighting condition. Three model days (day A, B, and C, Table 1_1) were defined to randomize the order of the modules. The three model days themselves were randomly distributed over the experiments, so that the plants would not get acclimated to a pattern. Illumination in the growth chamber started in late afternoon, so that the eight hour dark period was synchronized with working hours to allow supervision of the phenotyping (Table 1_1).

Table 1_1: Fluctuating light modules 1 to 4 were randomized over the three model days A to C.

module	on [min]	off [min]	sum [min]
1	1	1.4	2.4
2	2.5	3.5	6
3	5	7	12
4	10	14	24
time (hh:mm)	day A (module)	day B (module)	day C (module)
17:00 - 19:00	3	1	4
19:00 - 21:00	1	3	1
21:00 - 23:00	3	4	4
23:00 - 01:00	4	4	1
01:00 - 03:00	2	2	3
03:00 - 05:00	2	1	2
05:00 - 07:00	1	3	3
07:00 - 09:00	4	2	2

The ratio between the on/off cycles was defined by the differential in light intensity, achieved by two shading meshes, to equalize the accumulated photosynthetically active radiation (PAR) for both sides over the course of a day. It is noteworthy that plants in both treatments were receiving the same amount of total PAR per day, with the difference that the left side was exposed to ever fluctuating light.

3.1.2 *Arabidopsis thaliana* population of natural accessions

To study Arabidopsis' response to fluctuating light, a synchronized Arabidopsis accession panel was used, which consisted of 384 natural accessions (Table S 1_1) of which 382 were characterized by 109,178 homozygous and 104,873 heterozygous SNPs, in total by 214,051 SNPs (Horton *et al.*, 2012) of which 192,498 SNPs were biallelic (Alonso-Blanco *et al.*, 2016). The composition of the panel was chosen to minimize population structure as represented by the low percentage of variance explained by individual principal components ranging from 4.1% of PC1 to 1.7% of PC4 (Figure 1_2). A mild separation between the accessions from the collection sites in North America versus the rest of the world was observed by PC2 (Figure 1_2). Ninety-two principal components were necessary to explain an accumulated portion of variance of more than 50%. This accession

panel covered natural phenotypic variation present in Arabidopsis as the collection sites from the accessions ranged from North America, Europe, Africa, Asia, and New Zealand (Figure 1_3).
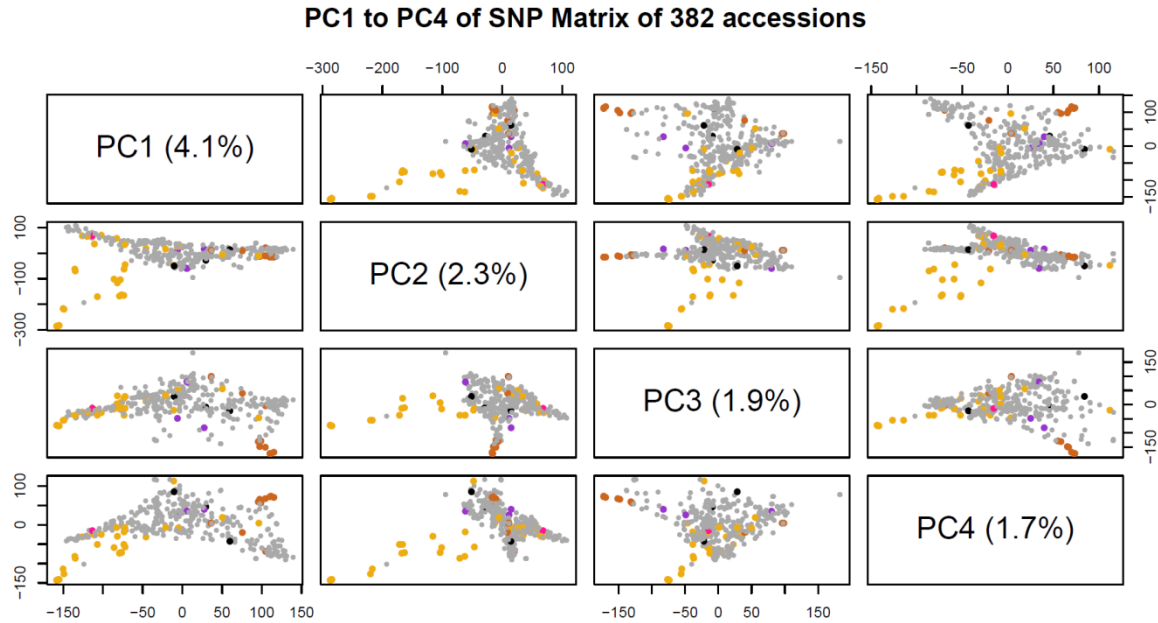
Figure 1_2: Principal component analysis of the 214,051 SNPs of the 382 natural accessions of the used Arabidopsis population. PCs are plotted as a scatterplot matrix. PC1 explained 4.1%, PC2 explained 2.3%, PC3 explained 1.9%, and PC4 explained 1.7% of the genotypic variance. European lines are colored in grey, North American lines in gold, Asian lines in brown, African lines in lilac, and the line from New Zealand in pink.
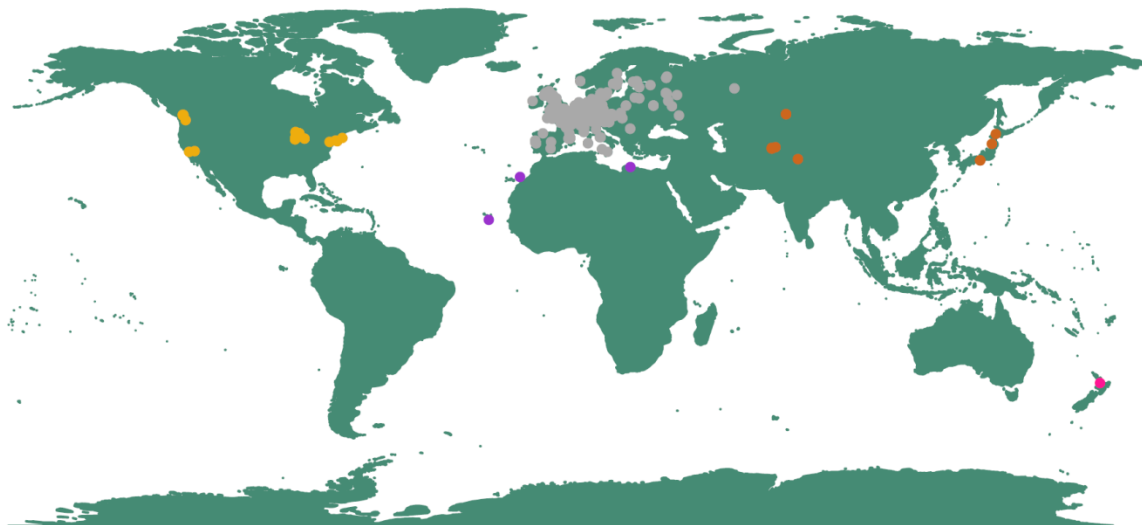
Figure 1_3: Collection sites of 380 Arabidopsis accessions, of which geographical data was available, plotted on a Mercator projection of the world map; European lines are colored in grey, North American lines in gold, Asian lines in brown, African lines in lilac, and the line from New Zealand in pink

The seed stocks of the accessions used for the phenotyping experiments originated from a prior synchronized cultivation of the natural accessions panel. The 214,051 SNP markers were, apart from centromeric regions, homogeneously distributed over the five Arabidopsis chromosomes with an average distance between two markers of 557 bp based on the golden path (Cheng *et al.*, 2017) of the Arabidopsis genome (Figure 1_4).
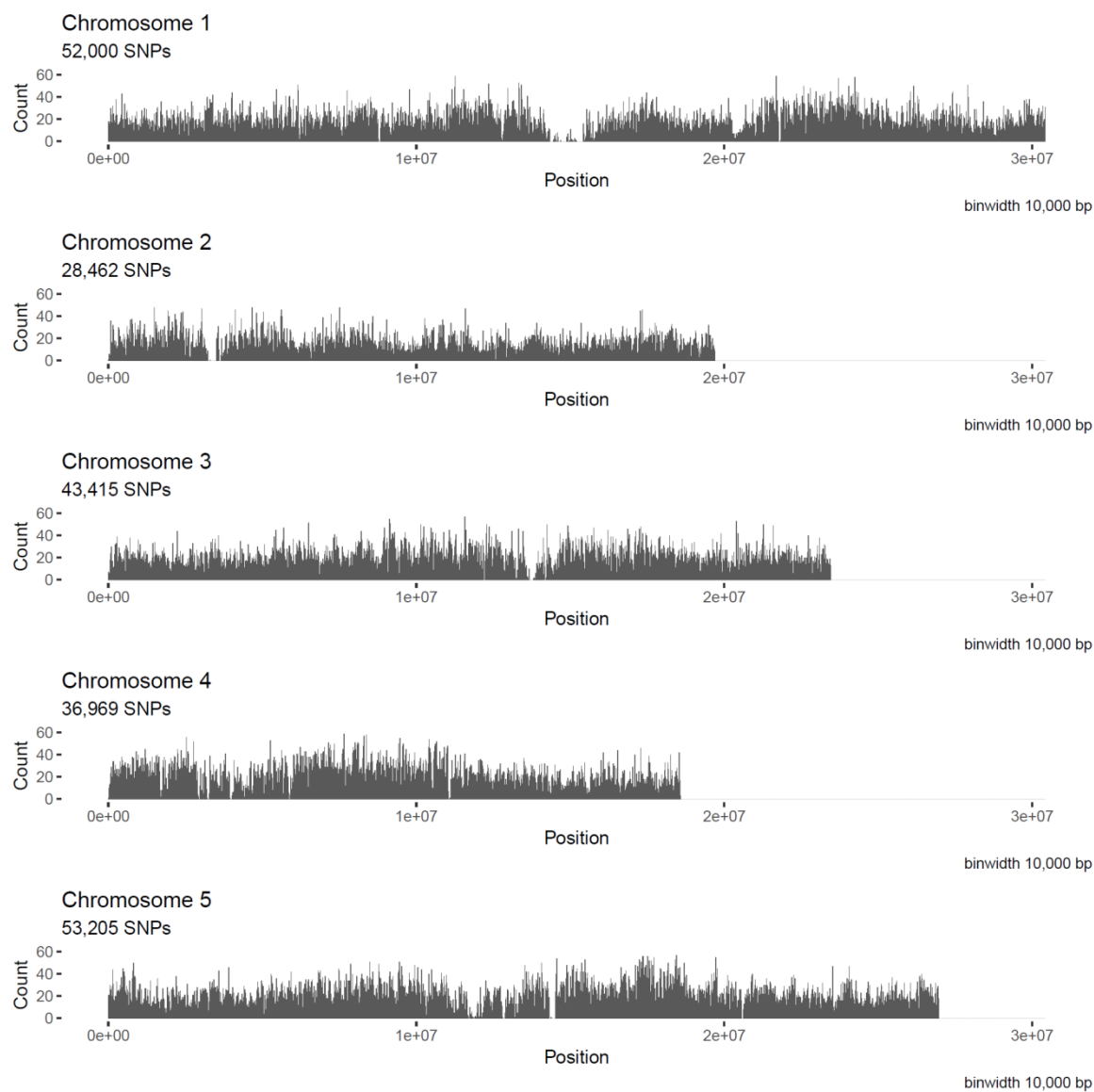
Figure 1_4: Marker density plot in bins of 10kb size, x-axis shows the position of a 10kb bin on the individual chromosomes and y-axis counts the SNPs in a single 10kb bin; average distance between markers was 557 bp; Chr.1 (3.04e+07 bp, 52,000 SNPs), Chr.2 (1.97e+07 bp, 28,462 SNPs), Chr.3 (2.35e+07 bp, 43,415 SNPs), Chr.4 (1.86e+07 bp, 36,969 SNPs), Chr.5 (2.7e+07 bp, 53,205 SNPs)

3.1.3 Linkage disequilibrium in the Arabidopsis accession panel

The resolution of a genome-wide association study is dependent on the linkage disequilibrium (LD) decay in the analyzed population, which is also contingent upon the given population structure (Korte and Farlow, 2013). A marker-trait association (MTA) found in a GWAS does not only indicate association with a given SNP but also with other genetic variants in LD surrounding the SNP. Thus, the LD decay pattern in the Arabidopsis accession panel was *de novo* calculated using the 109,178 homozygous SNP markers. The information about the position of each marker was used to calculate chromosome wide pairwise distance matrices by allelic correlation (r^2) using the *LDheatmap* package in R (Shin *et al.*, 2006).

Subsequently, the distances of a matrix of each chromosome were extracted and combined in a vector covering the whole genome of Arabidopsis. A second vector was created by combining the Euclidian distance matrices for each chromosome. Both vectors were combined into one data frame containing the Euclidian distances of the pairwise marker positions and the pairwise allelic correlation (r^2) between the markers. The combined r^2 values and their corresponding pairwise distances were fit into Eq.4 (section 2.6) calculating the LD decay following the proposed model by Hill and Weir, (1988). The half maximum LD of Eq.4 (section 2.6) was the half decay value of the Arabidopsis accession panel giving a genome-wide LD decay value of 3,365 base pairs (bp) (Figure 1_5).

The LD decay of the individual chromosomes ranged from a half decay of 2,909 bp on chr.1 to 4,260 bp on chr.2 (Figure S 1_3). A full LD block around a given marker was 6,730 bp, the length of the LD decay up- and downstream of the MTA position. The genome-wide LD decay was used to annotate the MTAs from the GWAS with Arabidopsis gene identifier in LD range. For each LD block, the Arabidopsis gene identifications were extracted from Araport11 (Cheng *et al.*, 2017).
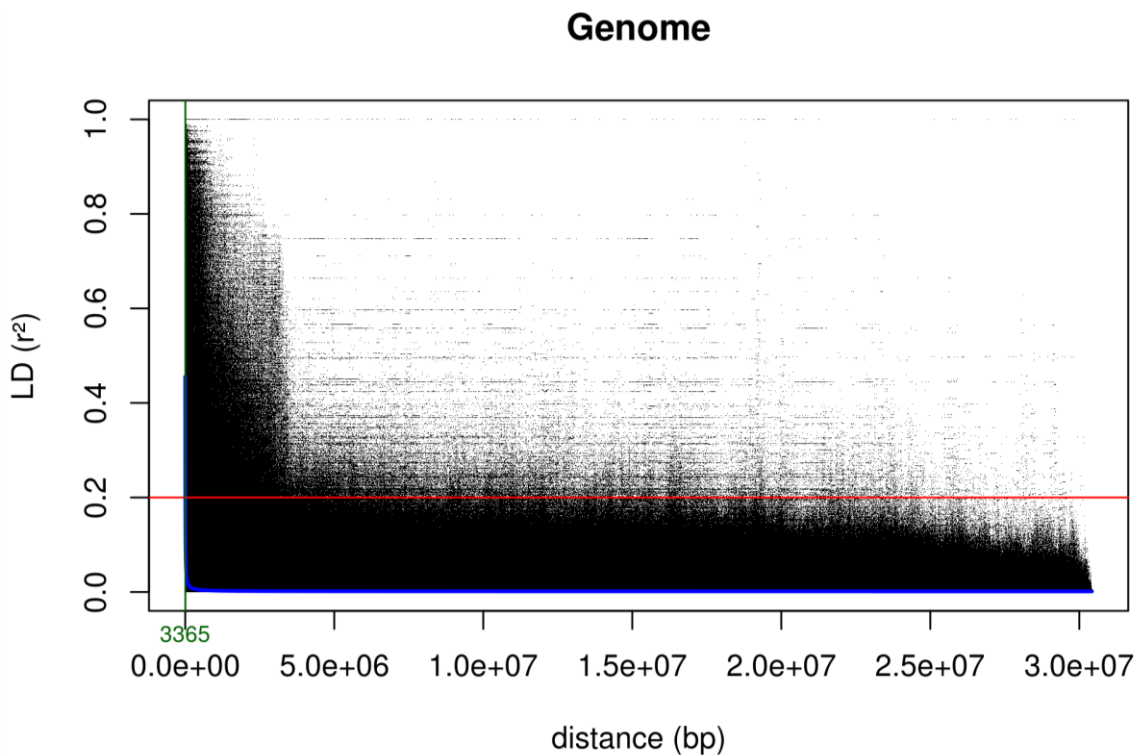
Figure 1_5: Genome-wide linkage disequilibrium (LD) plotted as *r²* against the pairwise marker distance in bp of 109,178 homozygous markers of the Arabidopsis accession panel. The blue line shows the LD decay as proposed by Remington *et al.* (2001). The green letters give the half maximum LD decay as a number – 3,365 bp.

3.1.4 The first fluctuating light experiment

The first fluctuating light experiment (1646MH) was run for 21 days, from Friday the 18th November 2016 to Friday the 9th December 2016. A full set of 384 accessions was grown under both light treatments in six replicates each, split into two times three replicates in a randomized block design. Each side consisted of 192 12-slot trays, which carried four genotypes with a replication of three, together 2,304 plants per light treatment. Each side was further split into two theoretical halves, 96 trays, which contained one full set of the accession panel with three replicates. Thus, a balanced randomized block design was realized, which later allowed to correct for positioning effects via linear mixed models. The illumination program started on the 3rd day after sowing, fluctuating light in the left side and constant light in the right side of the growth chamber. Thus, the Arabidopsis accessions directly germinated under their respective light treatments. The IPK automated non-invasive plant phenotyping system for small plants was used to record phenotypes of each individual plant nearly every day by taking top view images with a

RGB and a fluorescence camera. The images were subsequently analyzed with IAP and 286 plant architectural traits and color-related traits were extracted for each day of phenotyping (Klukas *et al.*, 2014). Color-related traits can be the normalized number of pixels in a certain color bin or a color intensity value. Architectural traits extract features like rosette diameter, leaf width, or top view leaf area, which itself is highly correlated with the fresh and dry weight of a plant (Klukas *et al.*, 2014). After an outlier correction the raw data from the IAP analysis were plotted to visualize differences in Arabidopsis' architecture and color specific for the two light treatments. Plants grown under fluctuating light had a significantly smaller plant area in comparison to their counterparts grown under constant light conditions (Figure 1_6 a). The total sum of the leaf areas, the aggregated leaf area incorporating data of all time points, was significantly smaller in plants exposed to fluctuating light conditions (Figure 1_6 b, *t*-test *p*-value < 0.001), which was in agreement with less dry matter accumulation under fluctuating light condition (fluctuating light: mean dry weight (DW) 4.1 g, constant light: mean DW 10.9 g; *t*-test *p*-value < 0.001). Principal components over all IAP traits per time point were calculated to visualize portions of variances according to the treatment applied. Five days after germination PC1 and PC2 mostly failed to separate variances of phenotypic traits expressed under different light treatments (Figure 1_6 c). Only PC3, which explains 0.8% of the variance, discriminates between fluctuating and constant light treatment (Figure 1_6 c). Fifteen days after germination PC1 and PC2, which together accounted for 98.6% of the phenotypic variance, separated the phenotypic traits according to the applied light condition (Figure 1_6 d). However, among the 286 IAP traits there were still traits whose components of variances overlapped, as no strict separation between the treatments was observed. Three days of this experiment were reserved for test measurements with the FluorCam panel (Tschiersch *et al.*, 2017) and consequently no phenotyping data was available for eight, ten, and 14 days after germination. The test measurements showed that the accessions differ in chlorophyll fluorescence parameters and a full measurement was conducted in the second iteration (1703MH) of fluctuating light treatment in the IPK automated non-invasive plant phenotyping system for small plants.
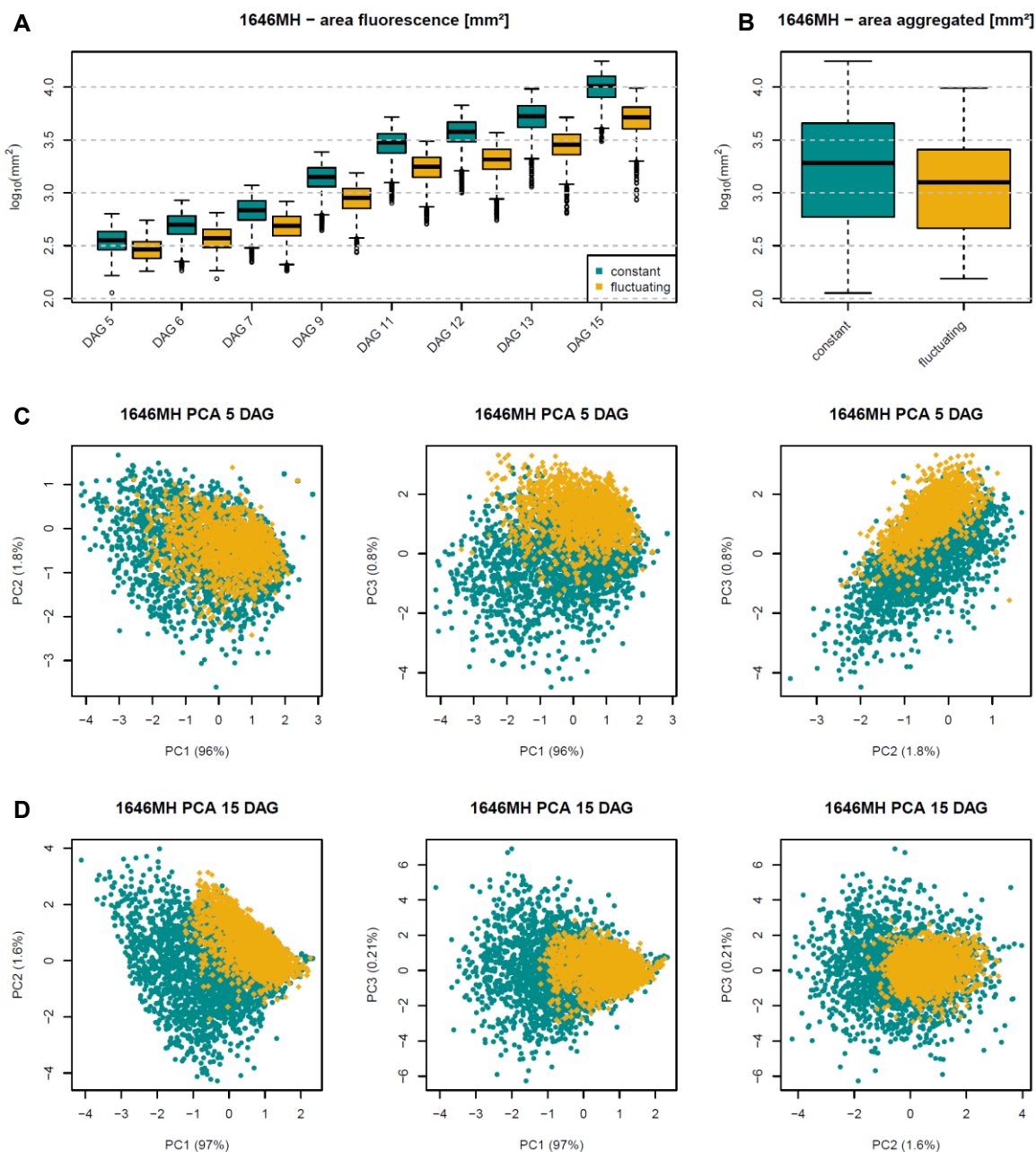
Figure 1_6: Effect of fluctuating light treatment on the Arabidopsis accession panel in experiment 1646MH on IAP raw data, gold color represents fluctuating light and blue color constant light condition, (**A**) $\log_{10}$ transformed leaf area extracted from fluorescence images at days after germination (DAG) 5,6,7,9,11,12,13, and 15, (**B**) aggregated leaf area of all time points; *p*-value < 0.001 *t*-test, (**C**) PCA of 286 raw IAP traits 5 DAG, (**D**) PCA of 286 raw IAP traits 15 DAG.

3.1.5 The second fluctuating light experiment

The second fluctuating light experiment (1703MH) was run from Friday the 20th January 2017 to Friday the 10th February 2017, for 21 days like the first experiment (1646MH). The second experiment was an iteration of the first experiment to provide enough data points to later allow the calculation of the best linear unbiased estimators (BLUEs) corrected for positioning, germination time and for technical variance in a linear mixed model. This experiment was run with the same set-up as the first one, except reducing the time of continuous darkness at 5°C and 95% relative humidity from two to one day. Hence, germination occurred one day earlier which subsequently led to larger plants at the end of the experiment as they were harvested 17 days after germination (DAG) in contrast to 16 DAG in the first experiment. Plants reached an average dry weight of 9.1 g under fluctuating light and 20.4 g under constant light, a difference which was found to be significant (t-test p-value < 0.001). Phenotyping of the leaf area revealed that plants under fluctuating light condition were smaller regardless of the measurement day (Figure 1_7 a). The difference in the aggregated leaf areas between the two treatments was significant (Figure 1_7 b, t-test p-value < 0.001). A PCA on the 286 IAP traits per time point explained a very similar portion of variance compared to the first experiment (Figure 1_7 c, d). At the earliest time point, five DAG, PC1, PC2 and PC3 accounted for 98.56% of variance in the phenotyping data, of which a fraction could be accounted for by the fluctuating light treatment (Figure 1_7 c).

At a later time point, 15 DAG, the separation by PC1 and PC2 (together 99% of the variance) was more pronounced (Figure 1_7 d). Principal component 3 did not contribute to a separation of phenotypic traits between the treatments (Figure 1_7 d). Top view images were recorded on a daily basis except for four consecutive days, eight to 11 DAG (Figure 1_7 a). At those days the phenotyping was replaced by kinetic chlorophyll fluorescence measurements with the FluorCam panel. Trays had to be phenotyped in the dark, like for the IAP traits, to avoid crossing the fluctuating or constant light side of the growth chamber while illumination was turned on. The throughput of the measurement of the fluorescence kinetic parameters allowed plants from 96 trays to be analyzed in one dark period, corresponding to one randomized block and containing three replicates per accession. Due to this constraint, each of the four randomized blocks had to be analyzed on four consecutive days, which made up the entire experiment.
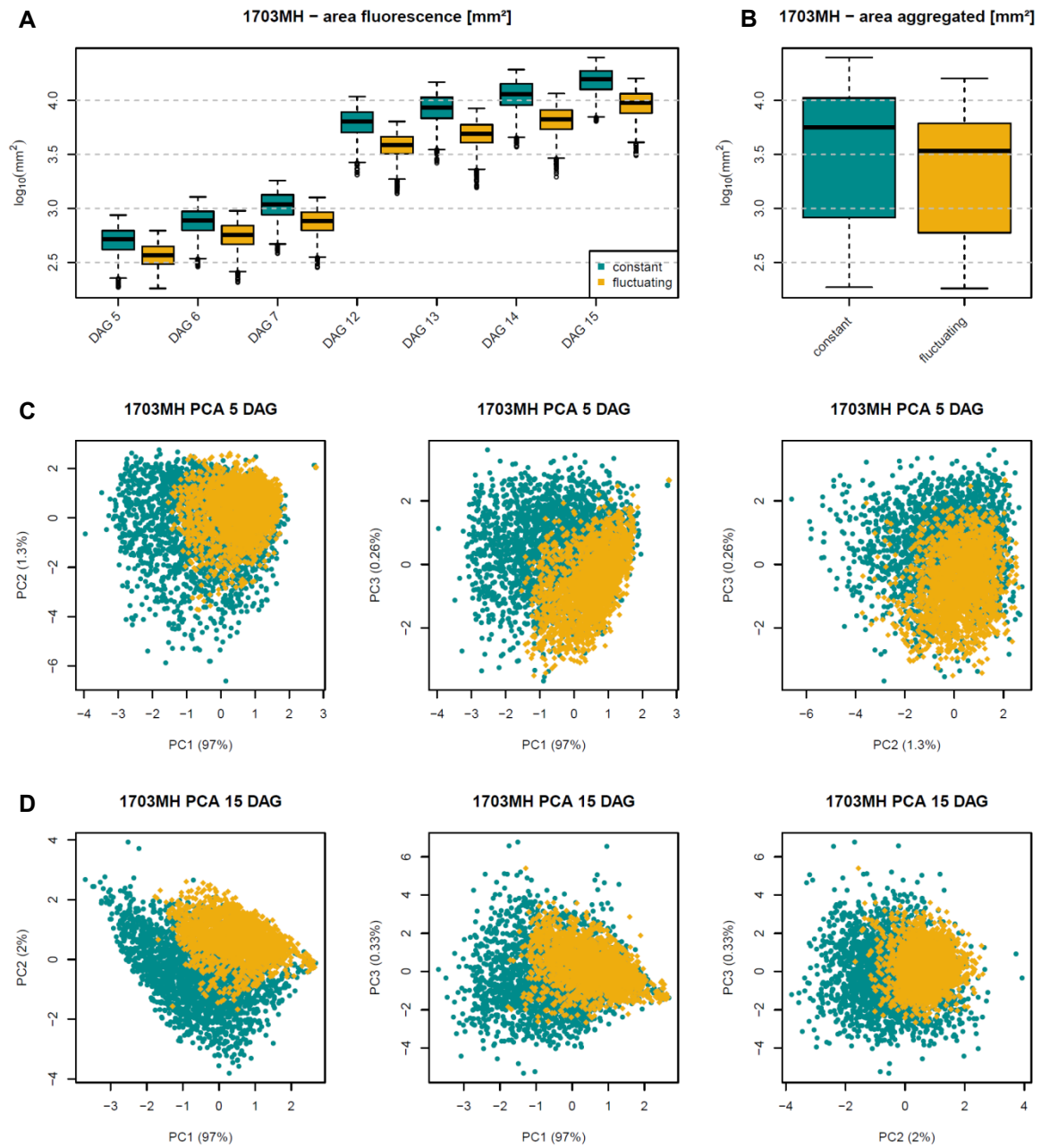
Figure 1_7: Effect of fluctuating light treatment on the Arabidopsis accession panel in experiment 1703MH on IAP raw data, gold color represents fluctuating light and blue color constant light condition, (**A**) $\log_{10}$ transformed leaf area extracted from fluorescence images at days after germination (DAG) 5,6,7,12,13,14, and 15, (**B**) aggregated leaf area of all time points; *p*-value < 0.001 *t*-test, (**C**) PCA of 286 raw IAP traits 5 DAG, (**D**) PCA of 286 raw IAP traits 15 DAG.

Hence, the trays were assigned into four groups, two per treatment. The first randomized block exposed to constant light condition was measured with the FluorCam at eight DAG, the first randomized block under fluctuating light condition was measured at nine DAG and the remaining randomized blocks per treatment were measured in the same order at ten and 11 days after germination. To account for the different measurement days, raw data from the FluorCam measurement was fitted into a linear mixed model (Eq.3, section 2.5) assuming fixed genotype effects to calculate BLUEs correcting for measurement day, day of germination and positioning effects. The model was effective in adjusting the means of the accessions, and PSII operating efficiency of Arabidopsis was found to be significantly reduced under fluctuating light (Figure 1_8).
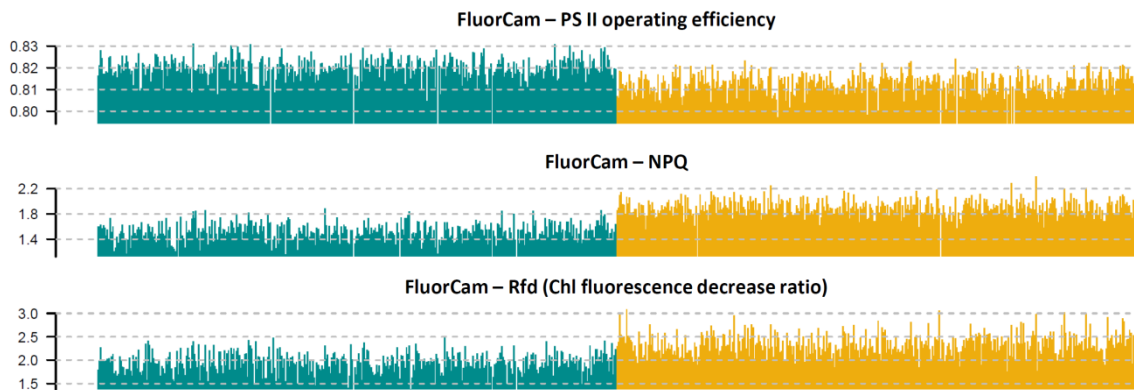
Figure 1_8: Best linear unbiased estimators of parameters (the PSII operating efficiency, non-photochemical quenching (NPQ), and Rfd Chlorophyll fluorescence decrease ratio) of the FluorCam measurement, under constant light (blue colored bars) and under fluctuating light condition (gold colored bars).

The repeatability of the PSII operating efficiency parameter (constant light: R_n 0.92, fluctuating R_n 0.96) was high under both light conditions, which indicated that most random variance was accounted for by the model whereas the remaining variation in the trait was under high genetic control. Non-photochemical quenching potential (NPQ; constant light: R_n 0.9, fluctuating light: R_n 0.87) and the chlorophyll fluorescence decrease ratio (Rfd; constant light: R_n 0.88, fluctuating light: R_n 0.85) were both significantly increased under fluctuating light conditions (Figure 1_8). BLUEs and the corresponding repeatabilities were calculated for every FluorCam trait (for list of traits see Tschiersch *et al.*, 2017) and their derivatives.

3.1.6 Control experiment under constant light

The third iteration of phenotyping of the accession panel in the IPK automated non-invasive plant phenotyping system for small plants (1806MH) was run for 21 days from the 9th February 2018 to the 2nd March 2018 under continuous light at both sides of the growth chamber in parallel. The aim of this experiment was to evaluate whether the different lighting sources, metal halide lamps versus metal halide lamps plus LEDs, had an influence on the phenotypes found in the two previous experiments with fluctuating light treatment. The experimental set-up was kept strictly the same compared to the previous experiments with the exception that the LEDs together with the main illumination were dimmed and not cycled in their intensity to reach the same amount of constant PAR in both sides of the growth chamber. Illumination was reduced to an average of 175.5 µmol m^{-2} s^{-1} on the right side (previously the side with constant light treatment) and to an average of 176 µmol m^{-2} s^{-1} on the left side (previously the side with fluctuating light treatment). Top view images were recorded daily from five DAG to 16 DAG (Figure 1_9 a). Image analysis with IAP extracted 123 architectural and color-related traits. The number was reduced in comparison to the two preceding experiments, as no binned color-related traits were extracted. Neither the aggregated leaf area (Figure 1_9 b) nor the accumulated dry matter at the end of the experiment significantly differed between the two light sources (left side: mean DW 15.3 g, right side: mean DW 15.7 g; *t*-test n.s.). The PCA supported this finding because neither PC1 nor PC2 nor PC3 separated the variances of the IAP data according to the side of the growth chamber 15 DAG (Figure 1_9 d). At the earliest time point, five DAG, PC1 indicated a minor difference for plants grown in the two sides of the chamber though much less than comparing plants grown under fluctuating light condition to plants grown under constant light condition (Figure 1_9 c). From that finding it was concluded that the differences in the spectral composition of the light in the two sides, as well as any other (potentially occurring but unknown) environmental difference of the two sides had negligible effects on the growth and the biomass accumulation of the plants.
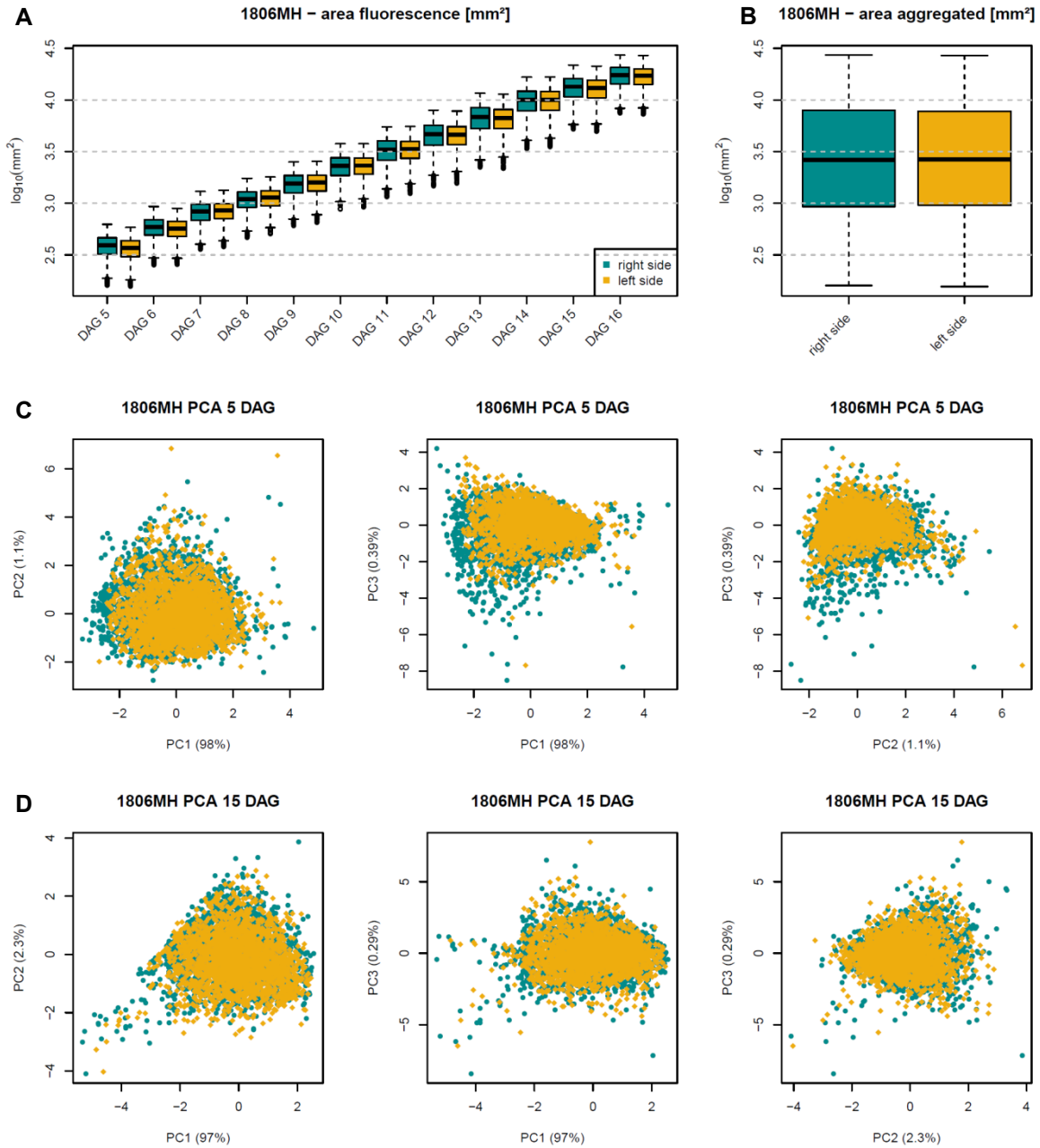
Figure 1_9: Effect of different light sources providing constant light for the Arabidopsis accession panel in experiment 1806MH on IAP raw data, gold color represents constant light at the left side with LEDs and blue color constant light without LEDs, (**A**) log10 transformed leaf area extracted from fluorescence images at days after germination (DAG) 5,6,7,8,9,10,11,12,13,14,15, and 16 (**B**) aggregated leaf area of all time points; n.s. *t*-test, (**C**) PCA of 123 raw IAP traits 5 DAG, (**D**) PCA of 123 raw IAP traits 15 DAG.

3.1.7 Best linear unbiased estimators of the first and second fluctuating light experiments

The best linear unbiased estimator (BLUE, Henderson, 1975) estimates the fixed effects of a linear mixed model under minimal variances of random effects and residual errors (see Eq.1, section 2.5). BLUEs were computed for each IAP trait at the equal day after germination of both fluctuating light experiments (1646MH and 1703MH) and adjusted for positioning effects, time point of germination, and the two independent iterations of the experiment. The accuracy of measurement, or repeatability (Nakagawa and Schielzeth, 2010), was calculated (Eq.2, section 2.5) and in this case explained the portion of phenotypic variation that was attributed to the genotypes (accessions). Higher values indicated that more variance could be explained by the genotypes of the underlying phenotypic trait. Repeatability values were highest when estimating fixed effects over both experiments in comparison to fixed effects of each individual experiment. In the example of rosette area seven DAG (Figure 1_10 a) repeatability of the first experiment (1646MH: constant light: R_n 0.69, fluctuating light: R_n 0.69) and the second experiment (1703MH: constant light: R_n 0.69, fluctuating light: R_n 0.73) were both lower than estimating fixed effects over both experiments (1646MH and 1703MH: constant light: R_n 0.81, fluctuating light: R_n 0.83).
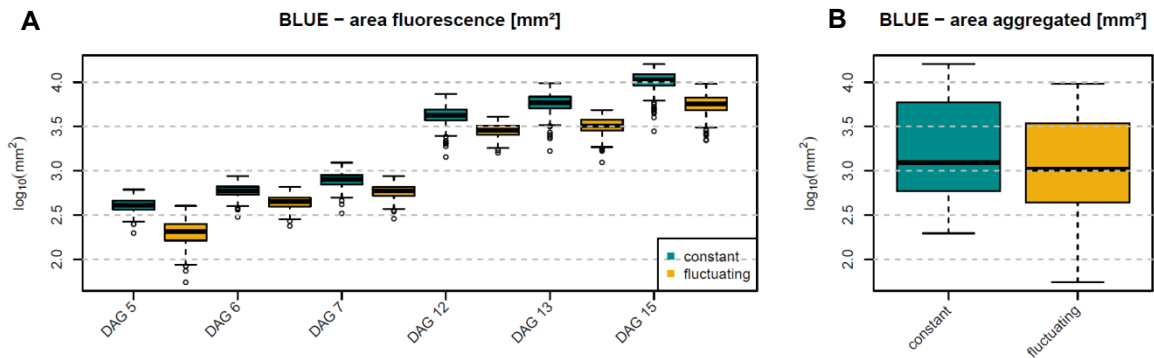
Figure 1_10: Best linear unbiased estimators (BLUEs) of the fluctuating and constant light treatment combining the two fluctuating light experiments, (**A**) BLUEs of the $\log_{10}$ transformed leaf area extracted from fluorescence images at 5,6,7,12,13, and 15 days after germination (DAG), blue color indicates constant light conditions and gold color indicates fluctuating light condition, (**B**) aggregated BLUEs of the $\log_{10}$ transformed leaf areas over all time points; p-value < 0.001 t-test.

Just as in each individual experiment, BLUEs of rosette area were smaller under fluctuating light condition in comparison to constant light condition (Figure 1_10 a). The aggregated area trait over

all DAGs was significantly different between the treatments (p < 0.001, Figure 1_10 b). Thus, BLUEs were calculated for 286 IAP traits over both experiments for 5, 6, 7, 12, 13, and 15 DAG.

3.1.8 Genome-wide association study with *FarmCPU* and post-GWAS analysis

Fluctuating light treatment-specific phenotypes of the same plants throughout the experiment were analyzed for potential genetic association, taking advantage of the dynamic non-invasive phenotyping. The BLUEs from the 286 IAP traits were used to perform a genome-wide association study with the 214k SNPs of the 382 accessions at each time point individually for both light treatments. In total, approximately 3,500 association tests had to be computed each for 214k SNPs. Hence, the *FarmCPU* package in R was used, which is computationally efficient and has improved statistical power in comparison to conventional methods (Liu *et al.*, 2016). The first four principal components from the PCA on the SNPs of the accessions (Figure 1_2) were fed into the *FarmCPU* GWAS model as covariates to reduce false positives attributed to potential population structure. The *FarmCPU* algorithm was allowed to run for a maximum number of 100 iterations, which was higher than default but still suitable due to the available computational resources, to ensure that the algorithm stops itself once reaching highest power and not being stopped prematurely. The *p*-value threshold was manually chosen to follow a stringent Bonferroni correction (0.05/214,051 SNPs) to avoid false positive detection of pseudo quantitative-trait nuclei. The minor allele frequency cutoff was set to 1%, which in this study required a SNP to be present in at least four accessions. The settings were chosen following the authors' advice from the user manual for *FarmCPU* (Liu *et al.*, 2016). Every association test yielded a Manhattan plot, which in *FarmCPU* is also called Helicopter plot as only the SNP with the highest $-\log_{10}(p)$ value represents a marker-trait association.

Furthermore, an association table was generated for each test containing a *p*-value, the minor allele frequency and the effect of every SNP tested in the algorithm, in total 214k rows for each individual association test. The quantile-quantile plot, Q-Q-plot, was taken as an indicator of how efficiently the model was able to correct for population structure. Phenotypic variance explained (PVE) by the significant SNPs was calculated by fitting a linear model for each trait taking the significant markers as fixed factor. The PVE was the variance of a significant marker as a fraction of the sum of variance of the significant markers and the residual errors (see Eq.5, section 2.7).

Significant associations were found dependent on the degree of repeatability of the measurement of the phenotypic trait and dependent on the phenotypic variation of the accessions' response to the two light treatments. The aim of this study was to find genome-wide associations with phenotypic traits that were specific for fluctuating light treatments. Therefore, every GWAS was performed in parallel for both light treatments for each time point and a pipeline was developed in R to extract the MTAs specific for fluctuating light and constant light treatment.

For each IAP trait the significant MTAs passing a Bonferroni correction were extracted from the association tables generated by the GWA at each time point and combined. This step was done for each light condition individually and resulted in 249 and 269 trait tables with every time point combined for fluctuating and constant light condition, respectively. The traits in those tables were filtered for traits present in both light treatments, 249 traits in total. Data distribution was checked manually for artefacts from image analysis of each of the remaining traits and the numbers were further reduced by ten, 239 traits were kept. The next step combined the remaining tables of both light treatments into one list, registering every MTA for each phenotypic trait and time point and treatment. This list contained 2,856 MTAs. The MTAs were further filtered to be specific to fluctuating light and vice versa (fluctuating light: 1,085 MTAs; constant light: 1,588 MTAs). The 183 MTAs that were detected under both treatments were omitted. That translated to 130 phenotypic traits yielding MTAs under fluctuating light and 171 phenotypic traits under constant light treatment, of which 102 phenotypic traits were in common for both treatments. Solely phenotypic traits with a repeatability value of higher than 0.5 were considered for further analysis. Thereby only IAP traits with variance under sufficient genetic control were selected. Of the 130 phenotypic traits with fluctuating light-specific MTAs, 103 were found to have a repeatability higher than 0.5 at least once during the measurements. Most traits were found to have a high repeatability over all time points and the average repeatability of those 103 traits in both treatments was 0.82 (Figure 1_11). Few traits were lacking measurement accuracy at early time points (five, six, and seven DAG) like leaf width or length, which were under tighter genetic control at later time points (12, 13, and 15 DAG; Figure 1_11).
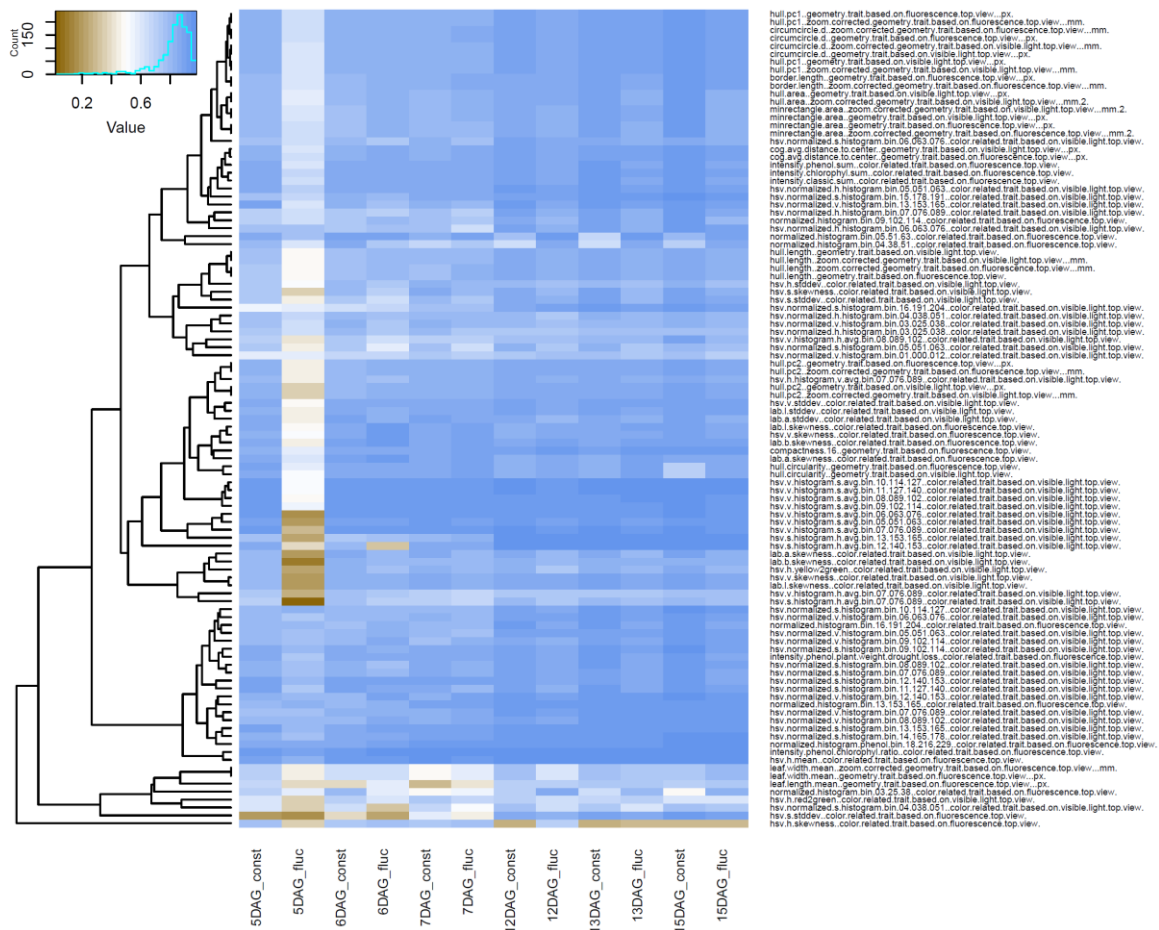
Figure 1_11: Repeatability values of 103 significantly associated traits under fluctuating light and their respective repeatability values under constant light treatment plotted in a heat map. Complete linkage clustering of Euclidian distance of the traits. Columns represent days after germination (DAG) and fluctuating and constant light treatment, and rows represent IAP traits. Brown to blue color gradient encodes repeatability values from 0 to 1, respectively, according to the key shown in the upper left part of the panel. Density function in the legend shows distribution of repeatability values in the heat map.

Every MTA was now annotated up- and downstream in range of the LD decay with at least one gene identifier, mostly several, and in rare cases with up to 16. In regions with up to 16 annotated gene identifiers most annotations could be attributed to transposable elements. The annotation was performed using the annotation file from Araport11 with the time stamp June 2016 (Cheng *et al.*, 2017). The 1,085 fluctuating light specific MTAs were within LD range of 5,354 gene identifiers and the 1,588 constant light specific MTAs with 7,610 gene identifiers. Manual interrogation of more than 10,000 gene identifiers is not feasible and therefore a GO term annotation was performed using software designed for GO term enrichment analysis. The *topGO* package in R (Alexa and Rahnenfuhrer, 2016) was set-up with the TAIR genome locus (TAIR10_2017) annotation

from the agriGO v2.0 database (Tian *et al.*, 2017) as a background reference. The GO term analysis was performed individually for each light treatment and each time point, which was necessary to later retain the time point information as all other identifiers except the gene identifier are dropped by the package. The gene identifiers of significant MTAs were filtered for GO annotation with the keywords "photo" or "light" specific for fluctuating light and constant light treatment, respectively.

Table 1_2: Marker IAP trait association fluctuating light specific GO filter for "photo" or "light"; DAG means days after germination, minor allele frequency (maf), effect was normalized by dividing by the mean of the best linear unbiased estimators, phenotypic variance explained (PVE) by the SNP, short description of the annotated gene identifier, and R_n was the repeatability

Trait	DAG	Chr.	Pos.	*p*-value	maf	effect norm.	PVE %	gene identifier	Short description	R_n
hsv.norm.h.his.bin. 06.063.076.vis.top	5	5	25,454,980	8.3e-11	0.06	-0.13	4.4	AT5G63590	flavonol synthase 3	0.75
hull.pc1.fluor.top.px	5	4	12,569,624	4.6e-11	0.07	0.04	2.8	AT4G24230	acyl-CoA-binding domain 3	0.64
hsv.norm.h.his.bin. 06.063.076.vis.top	6	5	25,454,980	3.1e-10	0.06	-0.12	4.3	AT5G63590	flavonol synthase 3	0.77
hull.pc1.fluor.top.px	6	3	21,827,353	4.4e-08	0.37	0.02	1.2	AT3G59060	phytochrome interacting factor 3-like 6	0.85
hsv.norm.v.hist.bin. 03.025.038.vis.top	7	4	461,470	8.2e-08	0.45	0.046	2.7	AT4G01050	thylakoid rhodanese-like	0.79
hull.pc2.vis.top.px	12	1	17,715,329	1.9e-08	0.2	0.023	3.1	AT1G48030	mitochondrial lipoamide dehydrogenase 1	0.88
hull.pc2.vis.top.px	12	3	21,232,451	9.6e-10	0.43	-0.02	4.8	AT3G57390	AGAMOUS-like 18	0.88
lab.a.skewness.fluor.top	12	5	14,012,328	2.5e-08	0.2	0.03	1.4	AT5G35840	phytochrome C	0.88
hsv.v.hist.h.avg.bin. 08.089.102.vis.top	13	2	19,034,511	2.4e-08	0.06	0.007	3	AT2G46370	Auxin-responsive GH3 family protein	0.82

The assumption was that genes that were annotated to respond to light stimuli or to be involved in photosynthetic processes, might also be involved in response to fluctuating light treatment. The filtered list contained 66 MTAs associated with the GO terms "photo" or "light", which were represented by 39 unique traits, of which 13 were architectural traits and 26 color-related traits. Traits originating from the same features extracted from both, fluorescence and RGB images, with and without scaling applied by IAP were causing a certain level of redundancy and yielding the

same MTAs. Because of this, a manual selection was applied to get the most promising candidates (Table 1_2). MTAs from redundant traits were preferably chosen according to higher repeatability and lower *p*-values. The full unselected list can be found in the ANNEX (Table S 1_2).

The selected MTAs included four color-related traits and two architectural traits. The architectural trait hull.pc1.fluor.top.px was a measurement of the maximum distance of any two pixels of a rosette taken from a fluorescence image, so basically the rosette diameter at the longest stretch. The other architectural trait hull.pc2.vis.top.px describes the summed maximal distances of other plant pixels left and right to the longest diameter (hull.pc1) traits, representing the width of the rosette orthogonal to the longest diameter axis. The four color-related traits were extracted out of RGB images from the HSV (hue, saturation, and value) color space. The color space was binned into 20 hue bins ranging from 0 to 255, each bin 12.5 arbitrary units. Bin zero to one were lying in the red spectrum, bin two were brown pixels, bin three were yellow pixels, and bins four to seven were carrying color information in the green spectrum. For example the trait hsv.norm.h.his.bin.06.063.076.vis.top, the sixth bin, contained information about green pixels with a hue value ranging from 63 to 76. The data dimension was a simple pixel count normalized to the plant size. The trait could be interpreted as the number of pixels in the given color ranges, with plants under fluctuating light having less pixel in the color range than plants under constant light conditions (Figure S 1_2). Two of the color-related traits represented pixel count and one at 13 DAG, hsv.v.hist.h.avg.bin.08.089.102.vis.top, encoded the intensity of the value parameter in the HSV color space in the eighth bin with green-bluish colored pixels. The trait was interpreted to exhibit differences in color intensity (see Figure S 1_2). The lab.a.skewness.fluor.top gave information about the skewness of the red color channel in the L*a*b color space. The trait showed more negative skewed values for plants grown under fluctuating light than for the plants grown under constant light (see Figure S 1_2).
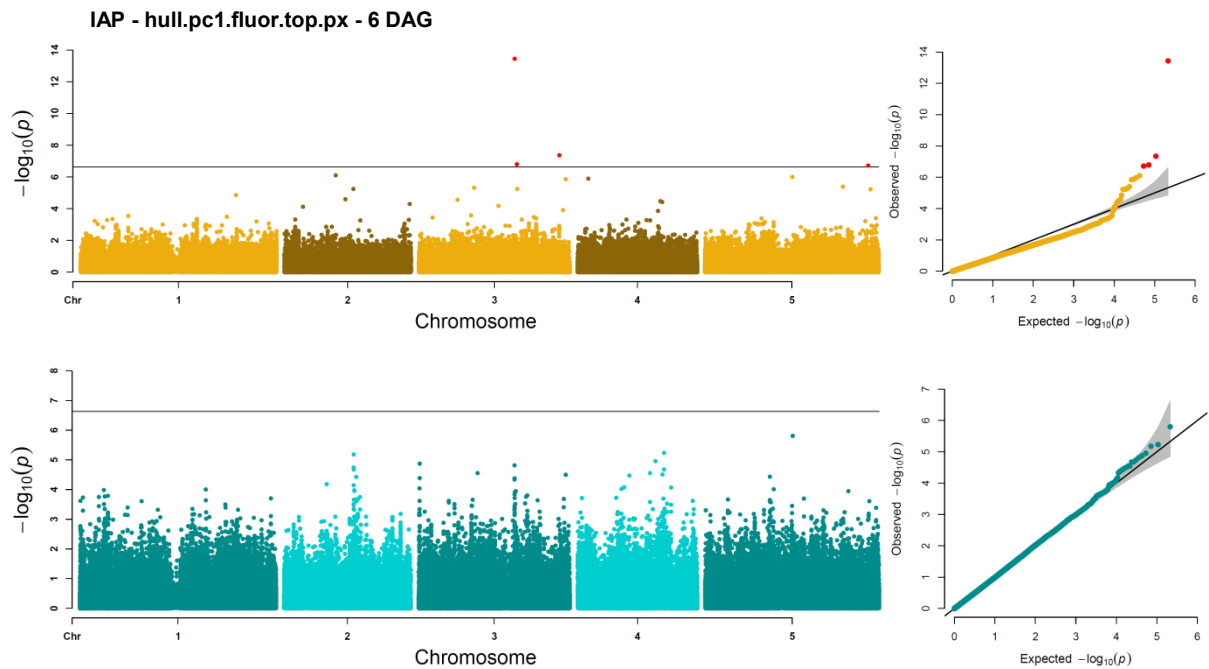
Figure 1_12: Genome-wide association study of the IAP trait hull.pc1.fluor.top.px 6 DAG describing the largest distance of any two pixels of the plant. The upper, golden Manhattan plot shows the GWA test performed on the best linear unbiased estimators of this trait of plants grown under fluctuating light and the lower, blue plot shows the test on data from plants grown under constant light. The corresponding Q-Q-plots indicate the efficiency of the model. Four significant marker-trait associations (red colored) were found under fluctuating light condition, while no MTA were found under constant light condition.

The quality of each significant association test was checked by investigation of the underlying repeatability values and the corresponding Q-Q-plots. The hull.pc1.fluor.top.px trait six DAG significantly associated with four markers after Bonferroni correction (Figure 1_12). The associations were only present under fluctuating light treatment and missing under constant light. The Q-Q-plot of the fluctuating light treatment supported the significant association by the observation of four significant markers deviating from the expected p-value distribution, which was missing under constant light (Figure 1_12). The repeatability of this trait at 6 DAG was at R_n 0.85 which indicated a strong genetic control (Table 1_2). The data distribution of the BLUEs of the accessions under both light treatments together with the Manhattan plots and the Q-Q-plots under fluctuating light of the selected MTAs from Table 1_2 can be found in the Figure S 1_2. The candidate genes of such selected MTAs are shown in Table 1_2.

The GWAS for the FluorCam traits was conducted with a similar analysis pipeline as applied for the IAP traits. Genome-wide association of 78 FluorCam traits was tested for both light treatments individually. For 34 of the 78 traits 144 Bonferroni corrected significant MTAs were found, 61 under fluctuating light and 83 under constant light. Traits with a repeatability value lower than 0.5 were discarded. Subsequently, the remaining MTAs were annotated with gene identifiers according to the calculated LD in the accession panel. Under fluctuating light condition 29 MTAs were kept for eight traits, which were inside a full LD block with 129 gene identifiers. The full unselected list can be found in the ANNEX (Table S 1_4). Of the constant light condition 73 MTAs were kept for 17 traits, which in total were annotated with 448 gene identifiers. The only fluctuating light trait containing GO terms with either "light" or "photo" was Rfd, which was also called the Chl fluorescence decrease ratio, a trait directly correlated with the CO_2 fixation rate in leaves (Lichtenthaler *et al.*, 2007). Two MTAs were found with Rfd (Table 1_3).

Table 1_3: Marker FluorCam trait association fluctuating light specific gene ontology filter for "photo" or "light"; minor allele frequency (maf), effect was normalized by dividing by the mean of the best linear unbiased estimators, phenotypic variance explained (PVE) by the SNP, short description of the annotated gene identifier, and R_n was the repeatability

Trait	Chr.	Pos.	*p*-value	maf	effect norm.	PVE%	gene identifier	Short description	R_n
Rfd_L3	1	28,418,423	1.6e-07	0.44	-0.02	2.7	AT1G75690	DnaJ/Hsp40 cysteine-rich domain superfamily protein	0.85
Rfd_L3	3	1,552,275	1.3e-09	0.48	0.02	1.4	AT3G05410	Photosystem II reaction center PsbP family protein	0.85

The Manhattan plot of the Rfd trait showed three significant associations under fluctuating light of which two were selected because of their GO term annotation (Figure 1_13). The same trait had no significant association under constant light. The Q-Q-plots showed slightly lower observed *p*-values than expected indicating an underestimation by *FarmCPU* (Figure 1_13).
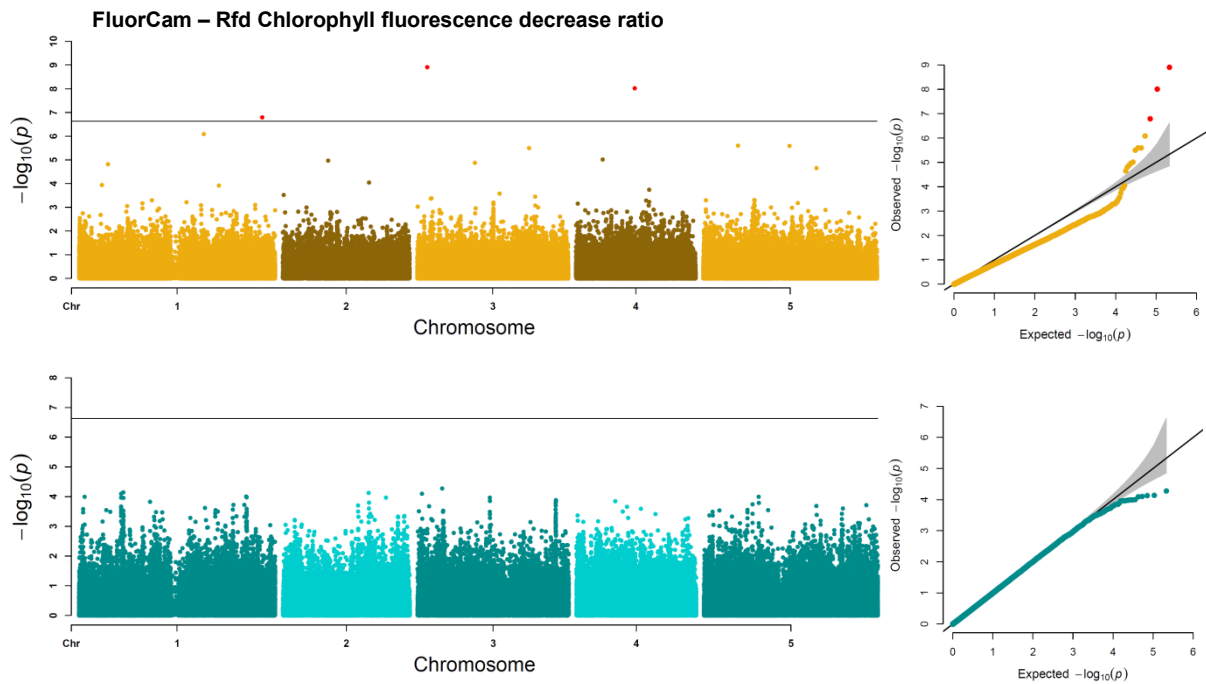
Figure 1_13: Genome-wide association study of the FluorCam trait Rfd Chlorophyll fluorescence decrease ratio. The upper, golden Manhattan plot shows the GWA test performed on the best linear unbiased estimators of this trait of plants grown under fluctuating light and the lower, blue plot shows the test on data from plants grown under constant light. The corresponding Q-Q-plots indicate the efficiency of the model. Three significant marker-trait associations (red colored) were found under fluctuating light condition, while no MTA were found under constant light condition.

3.1.9 Summary of growing *Arabidopsis thaliana* under fluctuating light condition

When grown under fluctuating light conditions, despite the same amount of intercepted photons, Arabidopsis responded with significantly reduced leaf area. Furthermore, high-throughput image analysis revealed broad changes in architectural and color-related features. In a GWAS, many MTAs were found specific for plants grown under fluctuating light conditions. After annotation of MTAs with genes in LD range, candidate genes, which passed a GO ontology filter for "photo" or "light", were reported.

3.2 Plant Cultivation Hall – indoor field simulation

The Plant Cultivation Hall (German abbreviation PKH) is a novel research infrastructure built to expose plants to field-like, dynamically changing environmental conditions in a precisely controlled manner. The PKH itself is split into four compartments of equal size, of which two are designed to provide field-relevant environments and the other two are currently equipped with an upgraded GROWSCREEN-Rhizo system following a modified design from Nagel *et al.*, (2012) with a sensor to plant concept to realize bigger rhizotrons to study root systems of larger plants or smaller plants for a longer period. In compartment one of the PKH a field experiment can be replicated with a size of up to 116 m², and more than six meters of isothermal airspace between the cultivation level and the light sources enable non-limiting growth conditions for crop species to explore their maximum plant height potential. A first benchmark experiment challenged the PKH by mimicking an average year on the IPK field site in the years 2016, 2017, and 2018. Reference phenotypes of 30 maize inbred lines were collected in three consecutive field phenotyping experiments in the maize growing seasons April – September in the years 2016, 2017, and 2018 on the IPK field sites. From October 2018 to February 2019 a smaller, but phenotypically diverse subset of 11 lines was cultivated in parallel in the PKH and in a large greenhouse (German abbreviation GWH) under standard conditions. The aim of this experiment was to test whether the PKH was able to produce field-like phenotypes. Therefore phenotypes of plants grown in the PKH were compared to those of plants grown in the GWH and in the three field experiments.

3.2.1 Experimental set-up in the Plant Cultivation Hall, a large greenhouse, and the field

The layout of the experiments in the field and the PKH was chosen to mimic field-relevant growth conditions. On the field sites, maize inbred lines were planted in double row plots, which consisted of 18 plants each, surrounded by a protective border planting of a commercially available maize hybrid to mimic a field stand and to shelter the inbreds from strong wind (Figure 2_1 a). Plants were fertilized once after one month in the field and manually watered only during the first two weeks. For further details see material and methods section. Thirty maize genotypes (for full list see Table S 2_1) were grown in two replicates, in total 60 double row plots. The two replicates were randomized in two complete blocks, indicated by the light and dark green color in Figure 2_1 a. In each row of a double row plot nine individual plants were planted ten centimeters apart from

each other. Out of the inner three plants on either side, six plants in total, five representative plants were chosen for the phenotyping experiment. Those plants were marked with labels and their phenotypes were manually measured twice a week in the years 2016 and 2017 and once a week in the year 2018. Phenotypes were recorded for plant height, growth stage, and vegetative stage. Plant height was defined as the tallest point of a plant measured above the ground level. Leaves were not stretched or elongated, which means that measured plant height was influenced by strong winds or by losing tassels due to breakages. The scoring of the vegetative stage was defined by counting only the mature leaves of a maize plant. Leaves were considered to be mature when they developed a leaf collar according to Abendroth *et al.*, (2011). In another scoring method, called the growth stage, every leaf on a maize plant was counted. The percentage of the total leaf length of the last leaf relative to the length of its preceding leaf was encoded by a fractional-part after the comma. The leaf length of the final leaf was scored in categories of increments of 0.2, 0.5, and 0.8 of the length of its preceding leaf. The two rating procedures together were sufficient to track the developmental stages of the plants until they reached vegetative tassel (VT) stage, when no more leaves developed and both counting methods converged when the final leaf matured. The two scoring methods will be referred to in the following text as vegetative stage and growth stage, respectively.

The design of the field experiments (Figure 2_1 a) was transferred to the PKH by growing the plants in a container based cultivation system with one cubic meter of soil each (Figure 2_1 b). A container was filled with four different layers of soil material starting with coarse gravel, followed by coarse sand, loess from the field and a top layer of soil. A selected subset of 11 phenotypically diverse genotypes from the field experiments was chosen and grown in 55 containers in five replicates each. The five replicates were grown in a randomized complete block design indicated by the red and green coloring in Figure 2_1 b. The inner five maize plants of each container were manually phenotyped twice a week following the same procedure as described for the field experiments.

In a parallel experiment the same selection of 11 *Zea mays* inbreds was grown under standard conditions in the GWH. Plants were grown in single pots with ten replicates per inbred. In the greenhouse the pots were arranged as ten rows of 11 pots each. Every row consisted of one

replicate of each genotype in a randomized order. Plants in the PKH and the GWH were watered regularly and fertilized as described in material and methods.
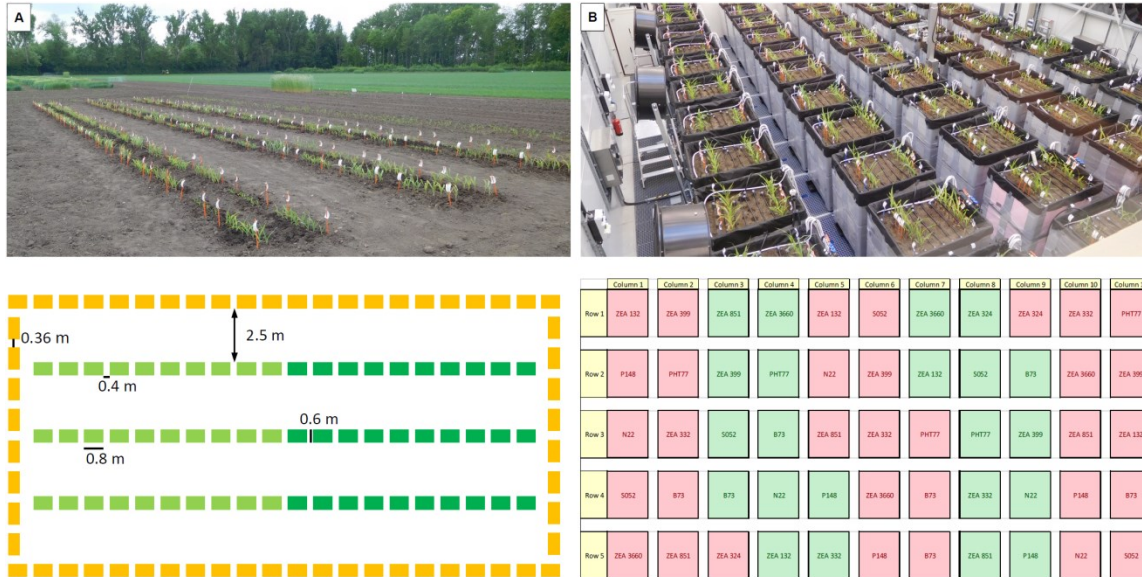
Figure 2_1: Layouts of the experiments of fields (2016/17/18) and the Plant Cultivation Hall (2018/19), (**A**) 30 maize inbred lines were grown on IPK field sites in two replicated (replication indicated by light and dark green color) in 60 double row plots (green rectangles) with a protective border planting (yellow rectangles), (**B**) 11 selected maize inbred lines were grown in five replicates (replication indicated by alternating red and green colors) in double row plots in a container based system in the Plant Cultivation Hall.

3.2.2 Climate conditions and settings

The phenotypes of the 30 maize lines were a product of the interaction of the genotypes and their accumulated exposure to the specific environmental conditions in the years 2016, 2017, and 2018 (Figure S 2_1). The years were highly diverse regarding accumulated temperature, global solar radiation and the amount of precipitation during the growing period of maize in Germany from April to September. The year 2017 exposed plants on the IPK field site to the overall lowest temperature and lowest levels of solar radiation but to the highest amount of precipitation relative to the other two years (Table 2_1). The year 2018 was the warmest and driest year of the three cultivation years and plants were exposed to the highest light levels. The year 2016 was a moderate year relative to the other two years (Table 2_1).

Weather at the IPK field site was recorded hourly by the IPK weather station (for a detailed overview see Figure S 2_1). The analysis of the weather data unraveled that wind directions and

humidity were stable throughout the time course of a field experiment. Wind at the IPK is blowing either from south-east (around 100°) or north-west (around 300°), two opposing directions (Figure S 2_1 wind direction). The humidity values were mostly reaching 100% relative humidity during the night (between 12 a.m. and 6 a.m.) and dropped to values between 30% and 50% relative humidity during the days (Figure S 2_1 humidity). Temperature and light conditions were following more diverse patterns.

Table 2_1: Sum of hourly climate values on the IPK field site from 1st April to 30th September 2016, 2017, and 2018, plotted are cumulative thermal time days (tt, explanation in section 3.2.7), precipitation in liters per m², and global solar radiation in W m⁻²

year	cumulative tt	precipitation [L m⁻²]	global solar radiation [W m⁻²]
2016	147.14	185.5	874,458.4
2017	133.25	367.7	791,397.6
2018	166.88	147.8	939,313.9

To test the PKH and to replicate certain climate conditions a careful observation and close meshed measuring of the weather was necessary. The PKH controls parameters regarding light conditions, temperature, humidity, and wind speed and direction. Each of these parameters was dynamically set and changed every hour and every day. A complex climatization program was developed for each week of the experiment in the PKH, in total 15 weeks and three days. The average climate (2016-2018) was simulated for the period from the 2nd May to the 28th August. The procedure is explained for the example of week 11 of the experiment, which translates to the week from the 11th July to the 17th July in the years 2016, 2017, and 2018 (Figure 2_2).

For each of the 21 days of this week in the three years (2016, 2017, and 2018; Figure 2_2 a), the hourly mean temperature values were computed and overlaid with a Tukey median smoothing function over the whole day to balance extreme deviations from an expected temperature curve (Figure 2_2 b). The mean profile of all 21 days of this week in the three years was defined as the "normal" day. Of the seven average days one "sunny" day (11th July 2016-2018) with the highest temperature and light levels and one "cloudy" day (15th July 2016-2018) with the lowest temperature and light levels per week were selected. For the week 11 of the climate simulation

three normal days, two cloudy days, and two sunny days were randomly distributed over the course of the week (Figure 2_2 c). This procedure of determining the average weather of a given week was applied to all 16 weeks of the experiment, with changing temperature values but the same order of normal, sunny, and cloudy days as shown in (Figure 2_2 c).
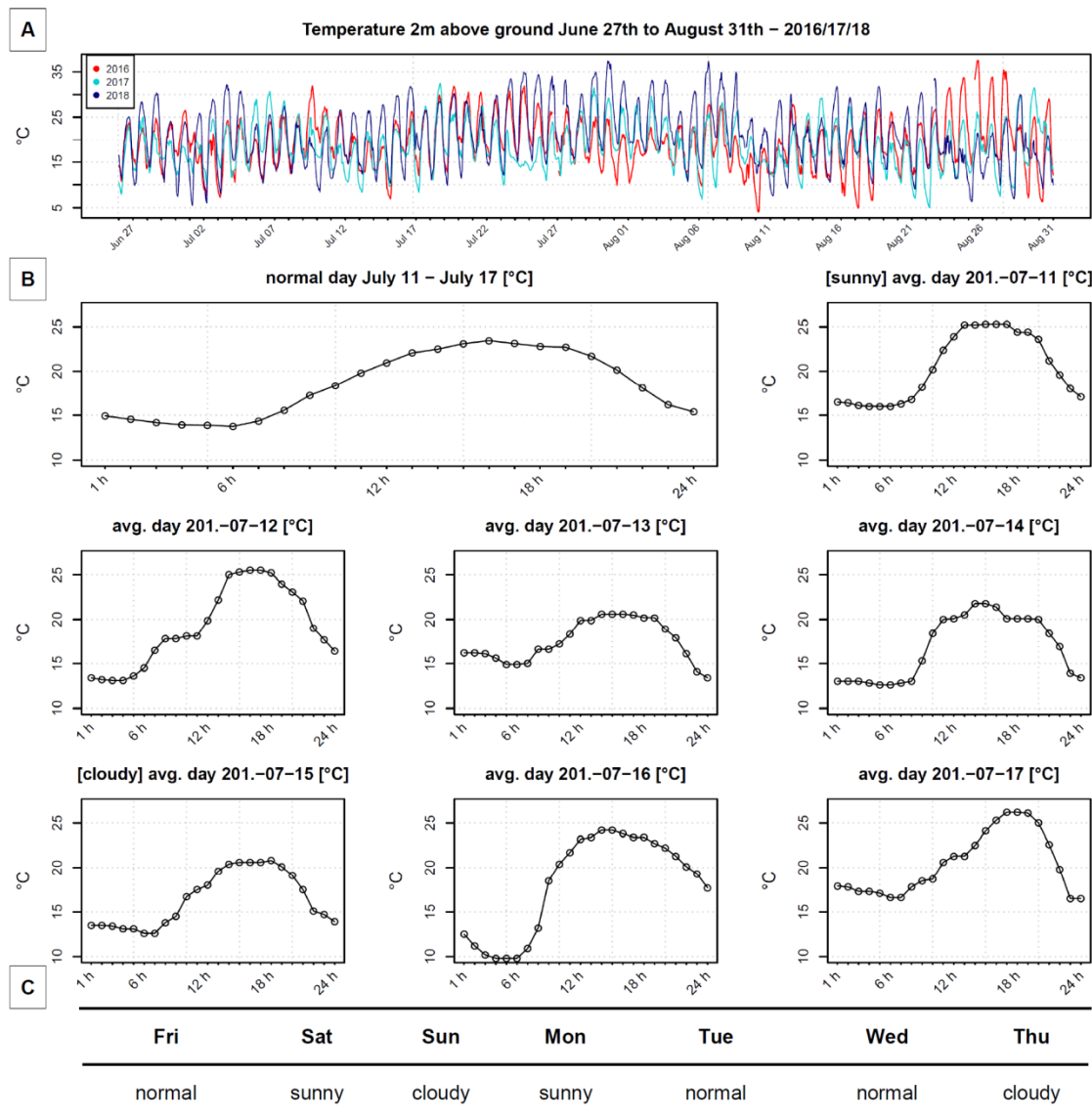
Figure 2_2: Plant Cultivation Hall temperature selection of week 11, (**A**) Above ground (2m) temperature of the years 2016, 2017, and 2018 from June 27th to August 31st, (**B**) Three year average of the selected week 11 (July 11th to July 17th), with the "normal" day as the average over all 21 days, and each three year average of one given day plotted with a Tukey median smoothing function. A "sunny" (July 11th) and a "cloudy" (July 15th) day were selected as they represented the hottest and coldest day per week, respectively, (**C**) The selected days (three normal days, two sunny days, and two cloudy days) were distributed over the course of a week.

Light was the second parameter which needed to be accounted for. Following the concept of three exemplary days per week ("normal", "cloudy", and "sunny" day), three independent light programs were created, in which each of the ten light sources was addressed independently (Tables S 2_2, S 2_3, and S 2_4). The strongest differences between the three light programs were light intensity and simulated cloud cover. The program of a normal day was set to reach intensities of around 1,450 µmol m^{-2} s^{-1} and five hours of cloud movement. Cloud movement was simulated by an oscillation following a sinus curve of light intensity, which can be seen in Figure S 2_2. The light intensity of the sunny day was set to be higher, reaching around 1,500 µmol m^{-2} s^{-1}, and cloud simulation was reduced to two hours. The cloudy day was characterized by drastically lower light intensities only reaching up to 400 µmol m^{-2} s^{-1} and eight hours of sinus curve oscillation reaching only one third of the levels of a sunny or normal day (Figure S 2_2). The same light programs exemplified for week 11, were applied to every week of the experiment with differing day length. Length of day was increased weekly by 10 minutes earlier sunrise and 10 minutes later sun set, in total 20 minutes, with reciprocal patterns after the 21st of June. In the first simulated week (2nd to 8th May) day length was 15 hours and 10 minutes, in the 7th week (20th to 26th June) day length was 16 hours and 45 minutes, and in the last week (22nd to 28th August) day length was set to 14 hours.

Humidity values were set according to the specific model days. Normal and sunny days were programmed to reach humidity values of 90% during night time and 40% during day time. At cloudy days, day time humidity was increased to 50%. In the field humidity often reached 100% from midnight until dawn, which was not possible to achieve with the climatization system of the PKH.

The last parameter which was accounted for was wind. Wind simulation was programmed to produce a steady current as the data from the IPK weather station showed, that the wind speed at the IPK never dropped to zero. Values of the fans were set to 10–20% of maximal rpm during the night, which converted to a wind speed of around one meter per second. During the light period wind speed was gradually increased and oscillated between 30–50% maximal rpm during the morning, decreased during the afternoon, and reached a steady 50% of maximal rpm during midday. Wind speed of 1.6–3 m s^{-1} were reached, which was in the range of typical wind speed at the IPK field site. The wind direction at the IPK was predominantly changing between South-East and North-West, two opposing directions, with sharp increments (Figure S 2_1). The rigid design of

the PKH fans also allowed for two opposing wind directions and directions were changed every two hours. The fans on both sides were alternating between push and pull configuration, meaning that when fans of one side pushed the air, fans from the other side pulled the air. Thereby a wind like air movement was realized. The climate data of the experiment in the GWH was following standard conditions as described in material and methods section, with constant day length, temperature and humidity values. Here, no wind simulation was applied.

3.2.3 *Zea mays* diversity panel

A diverse set of 30 *Zea mays* inbred lines was selected for the field experiments. The criteria for selection were variation in plant height, architecture and flowering time. Twenty lines from the dent and flint breeding pools were selected from Muraya *et al.*, 2017 and additional ten lines were provided by the IPK Genebank and selected by Dr. Astrid Junker from the Genebank collection following the above mentioned criteria (Table S 2_1). The selection of the 30 lines was based on phenotypes in experiments in the IPK automated non-invasive plant phenotyping system for large plants (Junker *et al.*, 2015). After three years of field phenotyping (Figure S 2_3) a set of the most diverse lines was selected for the phenotyping benchmark experiment in the PKH. The lines originated from the United States of America, Europe and Asia covering a broad spectrum of genotypic and phenotypic variation (Table 2_2).

Table 2_2: Selection of the 11 *Zea mays* lines grown in the field experiments and the Plant Cultivation Hall and a large greenhouse. Lines abbreviated with "ZEA" were provided by the IPK Genebank and the other lines were described in Muraya *et al.*, (2017).

Line	Taxa	Year	Donor/Breeder	Origin	Background
B73	*Zea mays* ssp. *mays*	1972	Iowa Agric & Home Econ Exp Stn	United States, Iowa	Stiff Stalk Synthetic
N22	*Zea mays* ssp. *mays*	unknown	unknown	United States, Nebraska	Krug Yellow Dent
P148	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk
PHT77	*Zea mays* ssp. *mays*	1988	Pioneer Hi-Bred International, Inc.	United States, Iowa	Non-Stiff Stalk Synthetic
S052	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk
ZEA 132	*Zea mays* L. subsp. *everta* (Sturtev.) Zhuk. var. *glaucornis* Alef.	1953	Genetikai Osztaly Budapest	unknown	Breeding line
ZEA 324	*Zea mays* L. subsp. *everta* (Sturtev.) Zhuk. var. *haematornis* Alef.	1968	Agricultural Botanical Garden Bucharest	unknown	Breeding line
ZEA 332	*Zea mays* L. subsp. *indurata* (Sturtev.) Zhuk. var. *vulgata* Körn.	1967	VIR Leningrad	Soviet Union	Breeding line

ZEA 3660	*Zea mays* L. subsp. *indentata* (Sturtev.) Zhuk. var. *flavorubra* Körn.	2003	BAZ, Braunschweig Genetic Resources Centre	China, Jilin	Research-, Breeding line
ZEA 399	*Zea mays* L.	1991	National Institute of Agrobiological Resources Tsukuba	DPR Korea	Research-, Breeding line
ZEA 851	*Zea mays* L.	1991	National Institute of Agrobiological Resources Tsukuba	DPR Korea	Research-, Breeding line

3.2.4 Phenotypes in the field

The field experiments were performed from April to September in the years 2016, 2017, and 2018. Plants were sown out at the 29th, 28th, and 20th April in the aforementioned years, respectively. After one week in the greenhouse plants were transferred into an outside shelter where they acclimated to the weather and were transplanted to the field after another approximately five days depending on the weather. First phenotyping started at the 17th, 18th, and 23rd May in the three years and last day of phenotyping was the 13th and 12th September in 2016 and 2017, respectively, and the 13th August in 2018. The final time point of phenotyping was identified by the developmental speed of the maize lines. Two consecutive phenotyping dates of no change in the plant height were the determining factor of the final plant height (Figure 2_3). The end of vegetative growth was cross-referenced by scoring the vegetative stage and the growth stage, which also had to converge at those time points for two consecutive phenotyping dates. After that date, individually for each genotype, the relevant genotypes were omitted from phenotyping and their reported final measurements were fixed in the lists.

The earliest genotype maturing in all three years was ZEA 332, which reached its final plant height with a vegetative stage of 12 mature leaves (Figure S 2_4). In the tasseling phase, plants were susceptible to high wind speeds around end of June and middle of July (Figure S 2_1), which reduced the plant height of ZEA 332 due to bending of the shoot in 2017 and 2018 (Figure 2_3). In both years plants fully recovered a week later. The latest maturing genotype was ZEA 324 with up to 25 leaves in 2016 (Figure S 2_4 and S 2_5). The average final plant height of all genotypes was affected by being exposed to the different weather of the three years. In 2016, the moderate year, the average of the BLUEs (Eq.8, see section 2.11) of the plant height was highest reaching an average of 184.1 cm (R_n 0.96) and an average vegetative stage of 18.1 mature leaves (R_n 0.99). In the wettest and coldest year, 2017, the average plant height dropped to 160.6 cm (R_n 0.98). The average number of mature leaves was 17.7 (R_n 0.99) and not significantly different from the

preceding year. In 2018, the hottest and driest year, the average plant height was affected even more and dropped to 158.7 cm (R_n 0.97). The vegetative stage proved to be a stable parameter and also reached 17.9 mature leaves (R_n 0.99) in 2018. The differences in plant height between 2016 and either of the two other years were significantly different with p < 0.001. In contrast, the differences in vegetative stage were not significant.
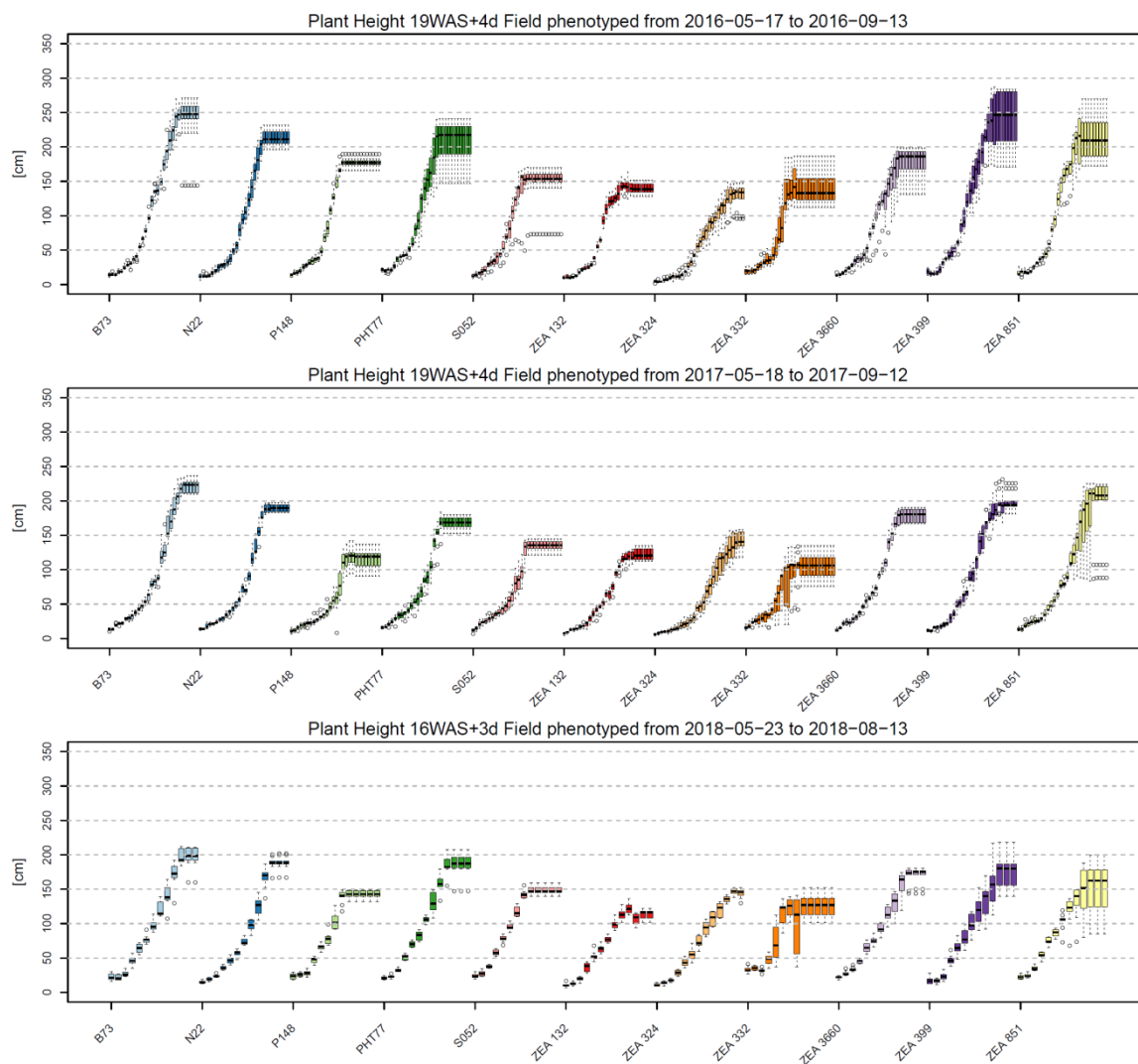
Figure 2_3: Plant height of the selected 11 maize lines in the field (n=10). Plant height was measured from May until September in the years 2016 and 2017 and until August in the year 2018. At the last time point of phenotyping plants were grown for 137 days (19WAS+4d) and 115 days (16WAS+3d) in the years 2016 and 2017, and 2018 respectively.

3.2.5 Phenotypes in the Plant Cultivation Hall and a large greenhouse

The phenotyping data from the field were the reference data for benchmarking the PKH at the IPK. The same set of 11 selected *Zea mays* lines was phenotyped twice a week in a parallel experiment in the PKH and a large GWH. The climatization in the PKH was programmed to simulate the average weather at the IPK field site beginning in week one simulating the 2th May to the 8th May until the 17th week simulating the 22nd August to the 28th August as described in section 3.2.2.
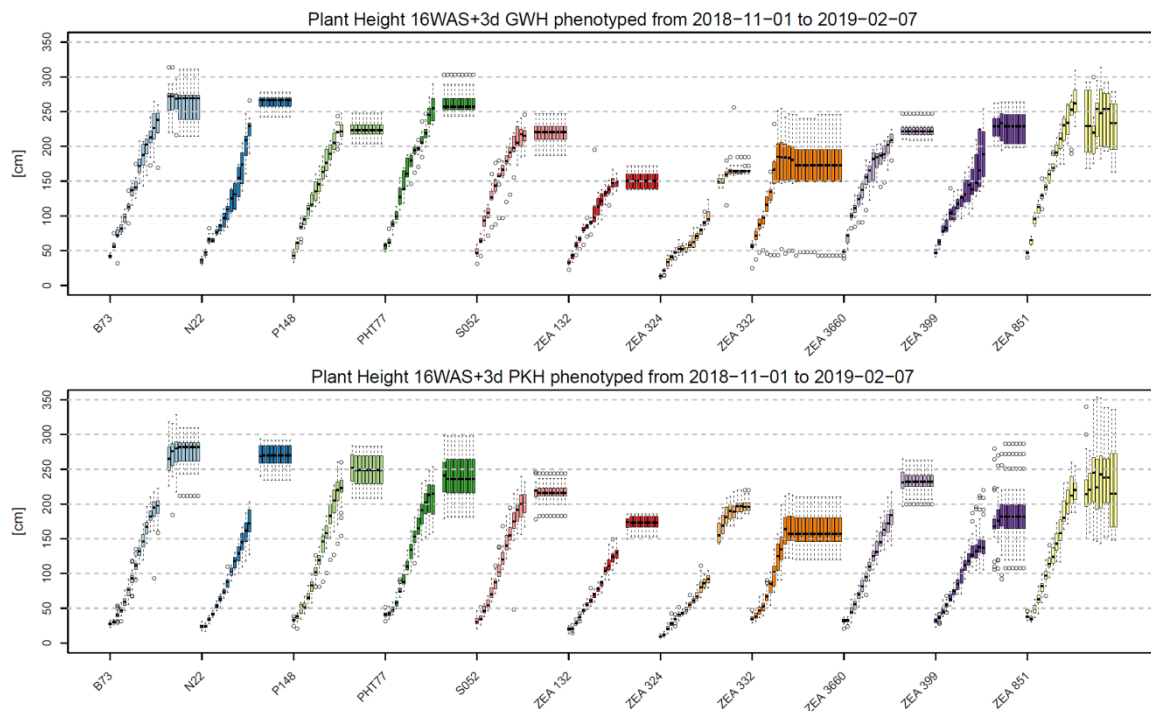
Figure 2_4: Plant height of the selected 11 maize lines in the large greenhouse (GWH) (n=10) and the Plant Cultivation Hall (PKH) (n=25). Plant height was measured from November 2018 until February 2019. At the last time point of phenotyping plants were grown for 115 days. In the gap no phenotyping occurred.

ZEA 332 matured earliest (Figure 2_4) with 12 leaves and ZEA 324 latest with 24 and 23 leaves in the PKH and GWH, respectively (Figure S 2_6 and S 2_7). The mean of the BLUEs (Eq.6 and Eq.7, see section 2.11) of the final plant height over all genotypes was 220.9 cm (R_n 0.97) in the PKH and 218.6 cm (R_n 0.96) in the GWH, which was not significantly different. In both cultivation scenarios plants matured on average with 18.7 and 18.8 leaves in the PKH and GWH, respectively. In the PKH and in the GWH plants did not develop significantly more leaves, but reached a significantly higher

final plant height compared to the three field cultivations (except the comparison between GWH and 2016, *p*-value = 0.059).

In the early growth period, the phenotypes in the PKH were much more similar to the field phenotypes than to those in the GWH, which is represented by a short arrest in plant growth after transplanting the seedlings from the greenhouse into their respective cultivation scenarios (Figure 2_3 and Figure 2_4). Early plant height and developmental speed in the PKH are closer to the field than to the plants grown in the GWH shown on the example of B73 (Figure 2_5). Plants in the field are much smaller and leaves mature at a slower pace than plants in the GWH. Phenotypes in the PKH were intermediate between the field and GWH phenotypes, an observation that was supported when modeling the best linear unbiased predictors (BLUPs) of the plant height (Eq.9, see section 2.11) in all cultivation scenarios (Figure 2_5). The BLUPs describe the effect of the cultivation scenarios on the plants height taking into account all measurement dates. While BLUPs in the GWH and the PKH both have a positive effect on plants height relative to the three field cultivations, the effect of the PKH is smaller and closer to the field trials (Figure 2_5).

Cultivation scenario	BLUP plant height
field 2016	-8.7
field 2017	-25.5
field 2018	-16.2
GWH	33.8
PKH	16.5

B73
2016 - 24 DAS
14.7 cm; V3

2017 - 24 DAS
13.6 cm; V3

2018 - 25 DAS
V3

B73
2018 - PKH 24 DAS
29.8 cm; V3.6

2018 - GWH 24 DAS
57.8 cm; V5
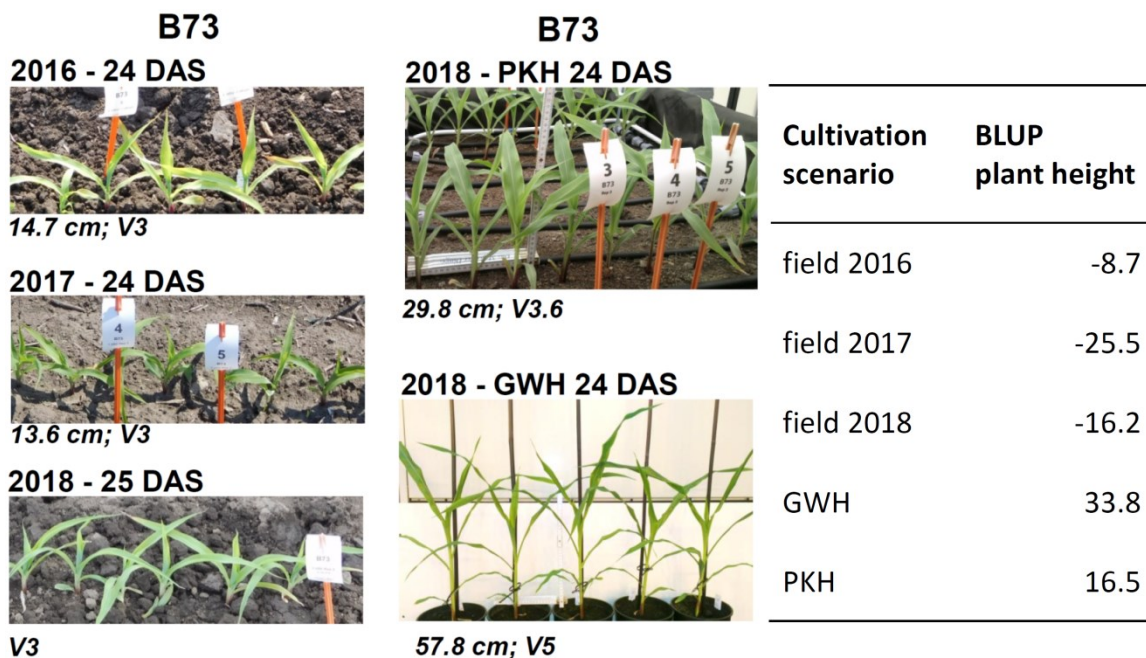
Figure 2_5: Phenotypes of B73 24/25 days after sowing (DAS) in the three field experiments 2016, 2017, and 2018, and in the Plant Cultivation Hall (PKH) and the large greenhouse (GWH). Plant height is presented in cm together with the vegetative stage describing the average number of mature leaves at the given time point. The table plots the best linear unbiased predictors (BLUP) of plant height for each of the five cultivation scenarios over all 11 genotypes.

3.2.6 Seed yield and dry weight

Phenotyping of plant height and leaf counting was supplemented with information about the seed yield and the accumulated dry matter of the genotypes grown under different cultivation scenarios to get a better impression concerning the performance of a genotype. After genotypes were exempted from phenotyping, when they had reached VT stage, ears were allowed to mature and irrigation was stopped. The last climate scenario (17th week) was repeatedly run for the rest of the desiccation period after the last genotype was exempted from phenotyping. Seed yield was determined for every inbred under all five growth scenarios. Ears were harvested from mature plants and total kernel weight, kernel count were measured for every individual. The sum of the BLUEs of the kernel count and kernel weight together with the average of the BLUEs of the thousand kernel weight (TKW) was calculated for each of the cultivations. The total seed yields of the same maize genotypes were affected by the exposure to the different climate scenarios. In the year 2016, total seed weight over all genotypes was highest, while 2017 and 2018 showed no significant differences (Table 2_3).

Table 2_3: Total seed yield of the five cultivation scenarios Plant Cultivation Hall (PKH), large greenhouse (GWH), field 2016, field 2017, and field 2018. Total sum of the best linear unbiased estimators (BLUEs) of the seed weight of all genotypes with the corresponding (R_n), total sum of the BLUEs of the kernel count of all genotypes with the corresponding repeatability (R_n), and the mean thousand kernel weights (TKW) from the summed BLUEs were recorded over all genotypes on either cultivation experiment.

experiment	sum kernel weight [g]	R_n kernel weight	sum kernel count	R_n kernel count	TKW [g]
PKH	435.30	0.92	3116.85	0.96	139.66
GWH	431.17	0.85	3054.60	0.94	141.15
field 2016	561.21	0.90	3492.71	0.96	160.68
field 2017	458.06	0.94	3927.56	0.98	116.63
field 2018	459.66	0.82	3417.97	0.93	134.48

The kernel weight produced in the PKH and the GWH was lower than either year on the field and the differences between the two indoor systems for total kernel weight and count were not significant. The repeatability (R_n) of the kernel weight was lower in the GWH. Likewise, ranking of genotypes according to kernel weight showed PHT77 and ZEA 399 yielding much less in the indoor

systems than in the fields (Figure S 2_9). The total number of kernels produced by these 11 genotypes in either system was ranking the PKH and the GWH lower than in any of the field experiments. Highest number of kernels was produced in the year 2017, while no significant difference was observed for the other two years. The ranking of kernel count between the genotypes was consistent with ZEA 132 and ZEA 324 showing the highest kernel count in either cultivation scenario (Figure S 2_8). The TKW was highest in the year 2016, but lowest in 2017. PKH, GWH, and the year 2018 were comparable (Table 2_3). Shoot dry weight was only measured for the plants grown in the PKH and the GWH. After the ears had matured, irrigation was stopped and plants were harvested one week later. Total shoot dry matter in the PKH was significantly heavier than in the GWH (Figure 2_6). Although ten of the 11 genotypes followed this pattern, PHT77 showed an opposing effect in which the genotype achieved a higher total biomass in the GWH than in the PKH (Figure 2_6).
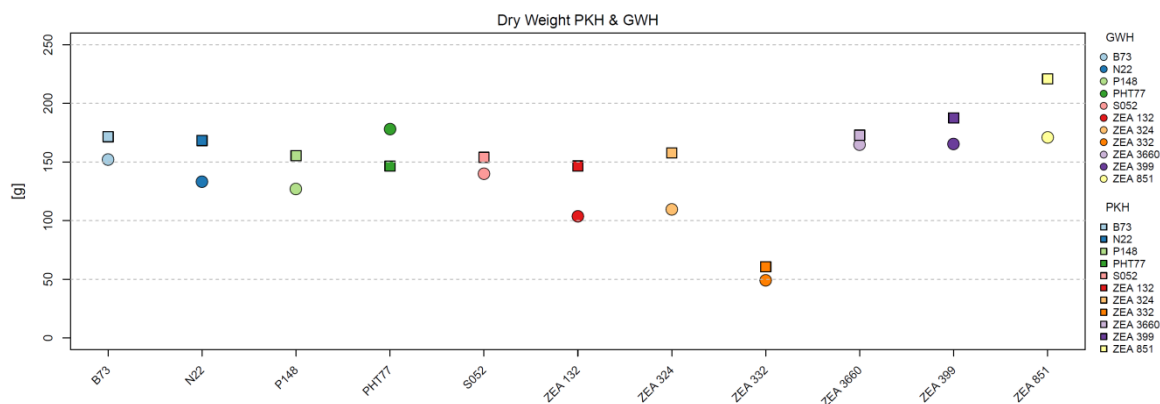
Figure 2_6: Best linear unbiased estimators of the dry weight of the shoot material with cobs of the Plant Cultivation Hall (PKH: n=25, R_n=0.86) and the large greenhouse (GWH: n=10, R_n=0.94), points symbolize the dry weight of the plants grown in the GWH and rectangles the plants grown in the PKH, differences between PKH and GWH are significant in the Welch's *t*-test with a *p*-value < 0.01.

3.2.7 Thermal time normalized development

Every cultivation scenario exposed the plants to different levels of temperature, light intensity and irrigation. The influence of temperature on many crops is calculated and expressed as growing degree days (McMaster and Smika, 1988) and was earlier described as thermal units (Cross and Zuber, 1972). The influence of cumulative temperature values has been extensively studied and many plant species have an optimum temperature at which the developmental speed reaches a

maximum (Parent and Tardieu, 2012). The concept of thermal time (tt) allows to normalize the time during which plants were exposed to a specific temperature, to an equivalent of days spend under constant 20°C (Parent and Tardieu, 2012). The hourly recorded weather data allowed for the calculation of a cumulative tt (Eq.10, see section 2.12) value for each day of the different experiments. Every phenotyping day was then assigned to a number of tt days, which can be interpreted as an equivalent amount of days the plant would have spent under constant 20°C to reach the same developmental stage. When the influence of temperature is modeled out, the remaining differences are then explained by other parameters like light intensity and quality or field-like fluctuations in the underlying climate scenarios. Plant height was plotted against the tt days in each individual experiment (Figure 2_7). Final plant height was defined as no detectable change at two consecutive measuring days. Around 100 tt days were required by the 11 selected *Zea mays* inbreds to reach final plant height in the three field experiments. While the plants in the PKH also required around 100 tt days to reach final plant height, the plants in the GWH were lagging behind and required more thermal time days to finalize plant height (Figure 2_7). The same pattern is reflected by the ratio of the growth stage to the vegetative stage, which approaches one when all leaves of the plants reach maturity (Figure S 2_10). The GWH requires more tt days to reach maturity than any other cultivation scenario. The rate of plants reaching maturity in the PKH is comparable to the three field experiments (Figure S 2_10).

3.2.8 Summary of benchmarking the Plant Cultivation Hall

A diverse set of *Zea mays* inbreds was grown in three consecutive years on the IPK field site and once in parallel in a large GWH and the PKH. Plant height and developmental changes were regularly phenotyped and developmental speed was normalized to cumulative thermal time days, an equivalent of days of growth at potential continuous 20°C temperature. Final plant height in the PKH and the GWH was significantly higher than in the field experiments. Final developmental stage was unaltered. During early growth phases the phenotypes of plants grown in the PKH were intermediate between the phenotypes of plants grown in the GWH and the three field cultivations. This was supported by smaller BLUP effects on plant height of the PKH in comparison with the GWH. The plants in the PKH and the field experiments required a similar amount of thermal time normalized days to reach maturity. The plants in the GWH needed between one and two weeks of thermal time days longer to reach final plant height and maturity.
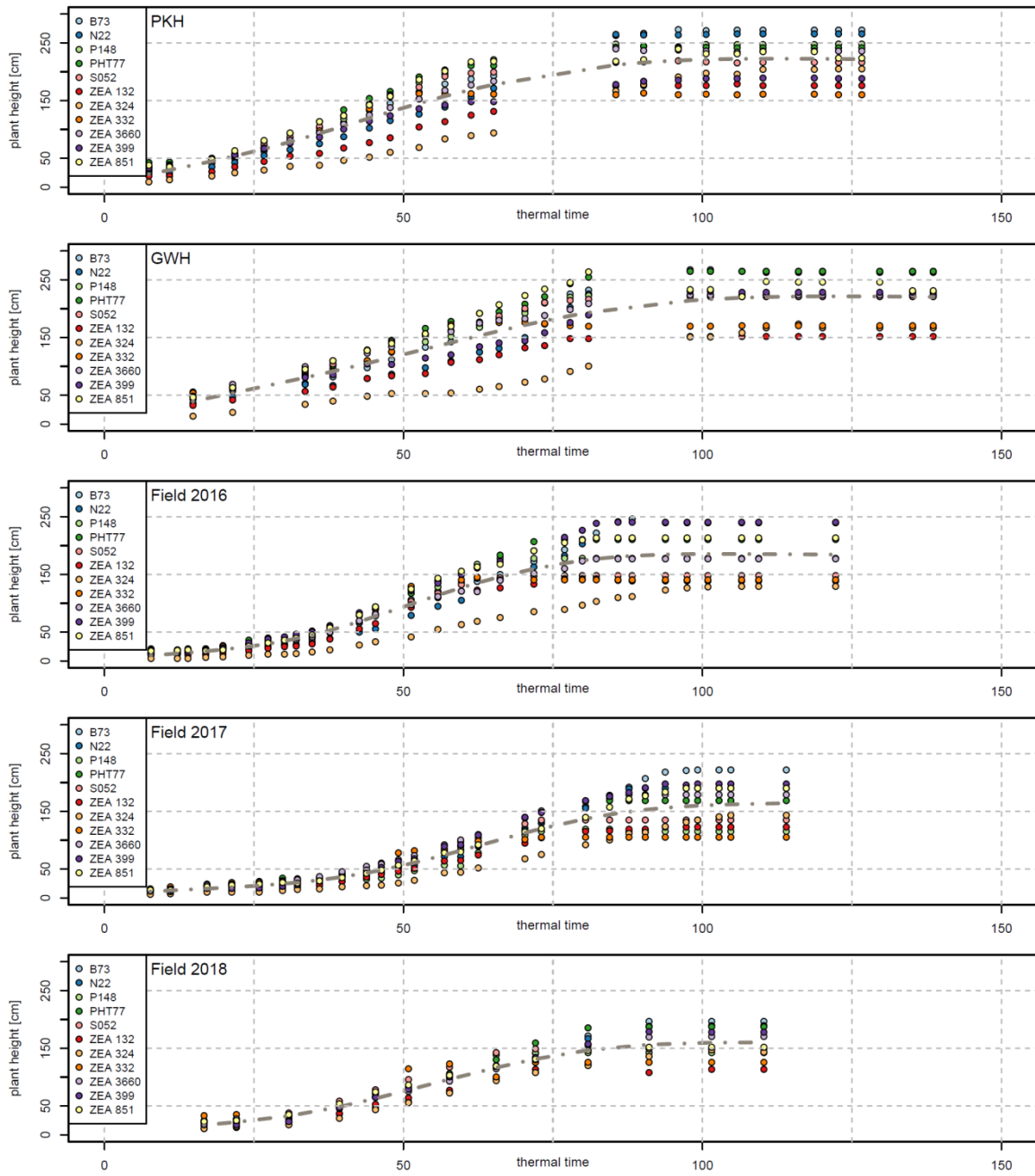
Figure 2_7: Best linear unbiased estimators (BLUEs) of plant height of 11 *Zea mays* genotypes in 5 experiments (Plant Cultivation Hall (PKH: n=25), large greenhouse (GWH: n=10), Field 2016 (n=10), Field 2017 (n=10), and Field 2018 (n=10)) plotted against thermal time (tt). Thermal time is the number of equivalent days at 20°C (Parent and Tardieu, 2012). PKH (115 days, 126.6 tt days), GWH (115 days, 138.6 tt days), Field 2016 (137 day, 122.6 tt days), Field 2017 (137 days, 114 tt days), Field 2018 (115 days, 110.9 tt days).

3.3 Rapid mapping of induced recessive and dominant mutations in maize

Next-generation sequencing boosted forward genetics approaches to map mutations to a genetic position. However, in most cases several rounds of prior backcrossing or large numbers of F_2 individuals are necessary to successfully perform such a mapping experiment. Here, a rapid mapping procedure is introduced and demonstrated for a *Zea mays* EMS population on mutants directly in the M_2 generation. A small number of M_2 siblings with and without a mutant phenotype were sequenced individually. The sequenced plants were selfed and the state of zygosity of the causative mutation was analyzed in a segregation analysis of the offspring based on Mendelian inheritance. The precise knowledge of the zygosity of the mutation was then used to apply a filter in which a SNP had to follow the zygosity pattern in each individually sequenced M_2 plant. Thereby a recessive dwarf mutation and a semi-dominant pale green mutation were successfully mapped to unique positions on the genome requiring only 16 maize plants for each mutation.

THE FOLLOWING TEXT IS CITED FROM MY PUBLICATION (Heuermann *et al.*, 2019)

3.3.1 The *dwarf* and *pale green* mutants

"We established a *Zea mays* mutant population based of the maize Iodent inbred line PH207 created by pollen EMS mutagenesis (modified protocol of Neuffer *et al.*, 2009) as a resource for forward genetics with mutant phenotypes. In total 7,000 M_1 seeds (unique mutation events) were planted, plants were self-pollinated, and transferred into the segregating M_2 population. For a proof of concept we chose two mutant families – a *dwarf* (M_1/M_2 family #1744) and a *pale green* (M_1/M_2 family #1754) – to demonstrate direct sequencing-based identification of EMS-induced mutations in M_2 maize mutant populations (Figure 3_1).

The *dwarf* phenotype was immediately expressed upon germination. During the early growth period (V1, V2; leaf collar method (Abendroth *et al.*, 2011)), internodes were drastically shortened and leaves were reduced in length and increased in width. In later stages (V5, Figure 3_1 a) symptoms were less severe but the plants did not catch up with siblings expressing a wild-type (*DWARF*) phenotype. Furthermore, the *dwarf* plants appeared to have a darker green coloration and their photosystem II (PSII) operating efficiency ($\Phi_{(PSII)}$) was increased in comparison to *DWARF*

plants (Figure 3_1 c). The *dwarf* mutation also caused the ears to develop anthers next to the kernels and to abort the upper part of the ear (Figure S 3_1).
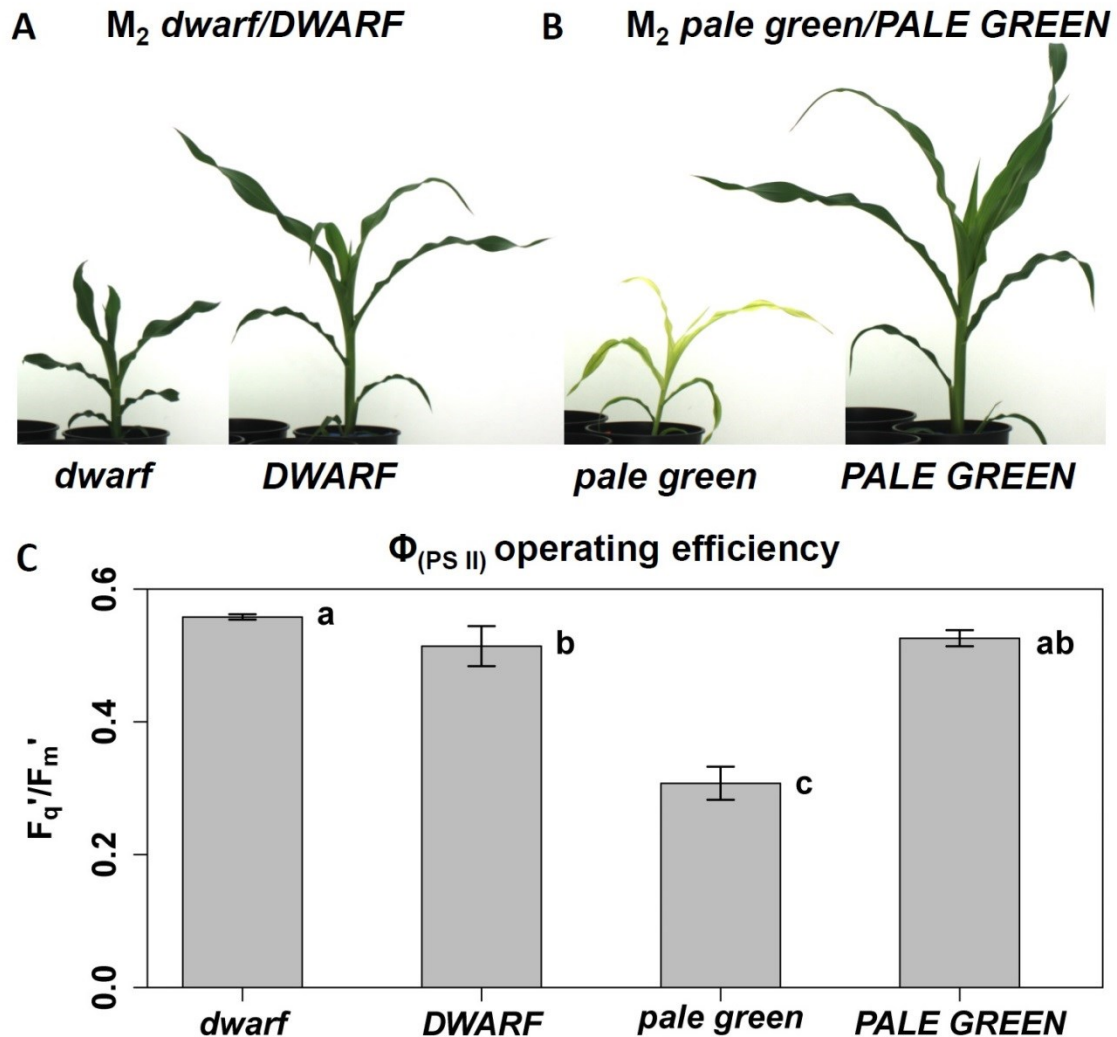
Figure 3_1: Phenotypes of the M_2 (**A**) *dwarf* mutant plant and its corresponding sister plant without a mutant phenotype (*DWARF*) and of the (**B**) *pale green* mutant plant and its corresponding sister plant without a mutant phenotype (*PALE GREEN*). (**C**) Operating efficiency of PSII ($\Phi_{(PSII)}$) in M3 progeny in the V1 stage of *dwarf*, *pale green*, *DWARF* and *PALE GREEN*; Barplot of F_q'/F_m' steady state data with standard error of the mean plotted; letters indicate significant differences (ANOVA followed by a post-hoc Tukey's test, n=5)

The *pale green* phenotype was characterized by its pale green coloration which coincided with a decreased plant height. The overall architecture of the plant did not differ from siblings with a wild-type (*PALE GREEN*) phenotype (V5, Figure 3_1 b). The pale green color was only observed in emerging and developing leaves. During maturation, leaves turned into darker green, which

caused the plant to express a pale green (emerging tissue) to dark green color (mature tissue) gradient (Figure 3_1 b). This coincided with a strong decrease in $\Phi_{(PSII)}$ in *pale green* plants relative to *PALE GREEN* plants, suggesting impaired function of the photosynthesis apparatus (Figure 3_1 c).

3.3.2 Whole-Genome-Shotgun sequencing of 32 individuals of segregating M_2 families

Genomic DNA of 32 individuals, composed of seven *dwarf*, nine *DWARF*, eight *pale green* and eight *PALE GREEN* plants was sequenced to an average of 19-fold coverage via WGS sequencing on an Illumina HiSeq 2000 platform (Table 3_1). About 99% of the reads were each successfully mapped to the B73 reference sequence and properly aligned, half of which were uniquely mapped (Table 3_1). SNPs, detected in regions with a minimal coverage of 10x, with a minor allele frequency (maf) of at least 10% and which did not surpass a maximal threshold of 60% heterozygosity (population *dw*/*DW* against *pg*/*PG*) and 10% missing data in both sequenced populations combined, were called between the sequences of the EMS mutagenized PH207-derived individuals and the B73 reference sequence resulting in 1,073,108 SNPs. Subsequently, SNPs segregating either among the 16 *dwarf*/*DWARF* (1,357 SNPs) or among the *pale green*/*PALE GREEN* plants (3,536 SNPs) were selected as induced mutations.

Table 3_1: Data alignment of paired end sequencing data from individual sequenced gDNAs for each of the mutant plants and their corresponding siblings without a mutant phenotype

Mutant plants *n* = 32	Total reads/plant	Mapped reads/plant	%	Properly paired/plant	%	Uniquely mapped/plant	%	Estimated coverage
dwarf n = 7	~ 432.9 M	~ 429.9 M	99.3	~ 428.6 M	99	~ 215.3 M	49.7	19 X
DWARF n = 9	~ 437.6 M	~ 434.5 M	99.3	~ 433.2 M	98.9	~ 218.5 M	49.9	19 X
pale green n = 8	~ 457.1 M	~ 453.9 M	99.3	~ 452.7 M	99	~ 229.7 M	50.3	20 X
PALE GREEN n = 8	~ 448.3 M	~ 445.3 M	99.3	~ 444.1 M	99.1	~ 222.6 M	49.7	19 X

3.3.3 Dominance/recessiveness of *dwarf* and *pale green* alleles and state of zygosity in sequenced M₂ individuals

The individually sequenced mutant plants (M₂) were, if possible, self-pollinated or else fertilized with B73 wild-type pollen and the segregation of the mutant phenotype was monitored in the offspring (M₃). This allowed the precise determination of the dominance/recessiveness of the mutations, *dwarf* and *pale green*, and the state of zygosity in each sequenced M₂ individual (Figure 3_2).
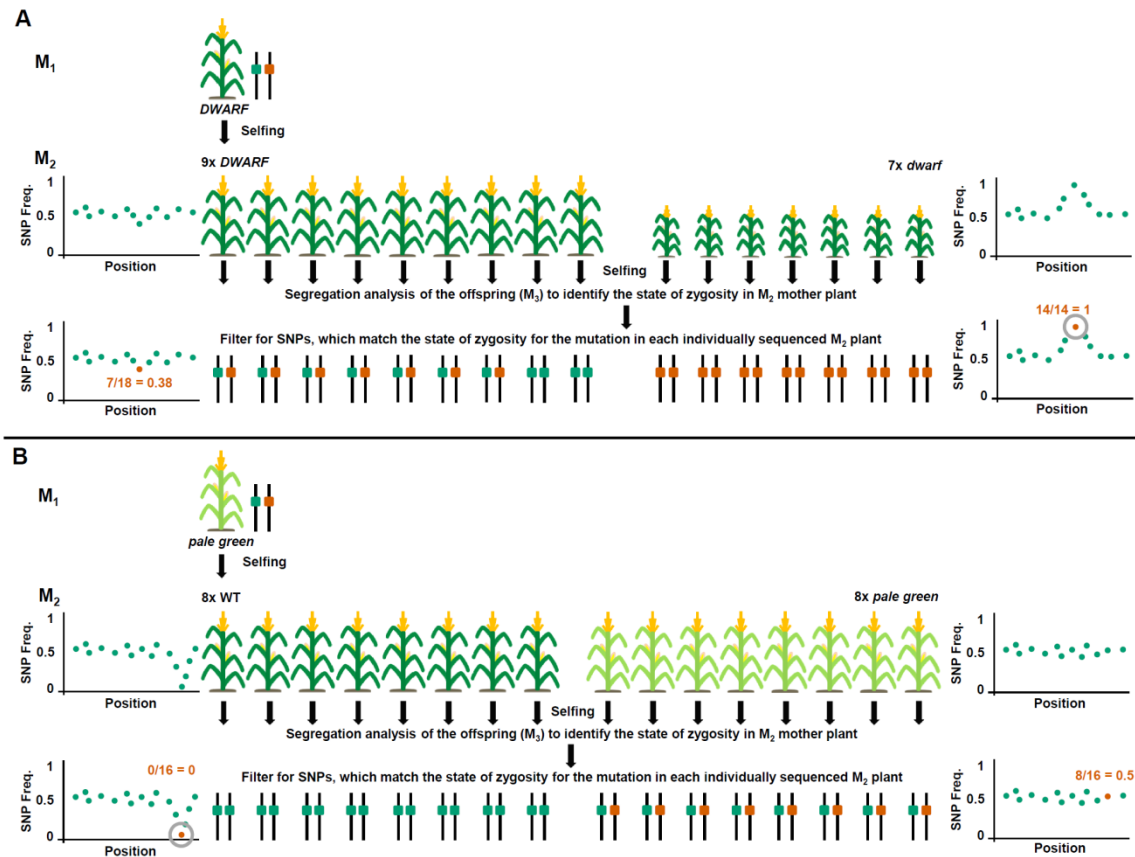
Figure 3_2: Schematic workflow of the SNP zygosity filter process for the *dwarf* and *pale* green populations. Individuals with and without a mutant phenotype of the M₂ population were individually sequenced and self-pollinated. The state of zygosity in M₂ mother plants was determined by a segregation analysis of the offspring (M₃). This information was used to filter for SNPs which match the state of zygosity in each individually sequenced M₂ plant. (**A**) In the nine *DWARF* plants (seven heterozygous and two homozygous wild-type) with 18 sequenced alleles, the frequency of the mutant SNP was 0.38 (7/18); and in the seven *dwarf* plants (all homozygous mutant), the expected frequency was 1 (14/14). (**B**) In the eight *PALE GREEN* plants (all homozygous wild-type) with 16 sequenced alleles, the expected mutant frequency was 0 (0/16); and in the eight *pale green* mutant plants (all heterozygous, as the allele is homozygous lethal), the expected mutant frequency was 0.5 (8/16).

The individuals expressing a *dwarf* phenotype were hard to self-pollinate. However, when cross-pollinated with B73, the resulting offspring expressed only wild-type (*DWARF*) phenotypes thereby proving recessiveness of the *dwarf* mutation (#15, 17; Table 3_2). One *DWARF* M_2 plant was pollinated with B73 and produced offspring with wild-type phenotype only (#21; Table 3_2). Progeny of five of the self-pollinated *DWARF* M_2 plants segregated for *dwarf* and *DWARF* phenotypes in a 1:3 relation, demonstrating a heterozygous state of the parental M_2 plant (#18, 19, 20, 23 and 24; Table 3_2). The offspring of one of the selfed *DWARF* M_2 plants exhibited only the wild-type phenotype, thus originating from a homozygous *DWARF* M_2 plant (# 22; Table 3_2). From this we concluded that the *dwarf* mutation is recessive and the sequenced *dwarf* M_2 plants thus were homozygous for the mutant (*dwarf*) allele (Figure 3_2 a). *DWARF* M_2 plants were either homozygous for the wild-type (*DWARF*) allele or heterozygous (*dwarf/DWARF*) (Figure 3_2 a, Figure S 3_4).

PALE GREEN M_2 plants were easily self-pollinated and their offspring produced only wild-type phenotypes (#8–14; Table 3_2), except one case where we did not get any seeds. The offspring of the *pale green* M_2 plants suffered from an overall low germination rate. Nonetheless, for six out of seven self-pollinated *pale green* M_2 plants, the offspring segregated for mutant (*pale green)* and wild-type (*PALE GREEN*) phenotypes (#1–4, 6 and 7; Table 3_2). One *pale green* M_2 plant produced only a single *pale green* offspring plant. Considering these data and the low germination rate, we concluded that the *pale green* mutation is semi-dominant causing the *pale green* phenotype in heterozygous individuals, while homozygous *pale green* individuals are embryo lethal. From this we derived that all sequenced *pale green* M_2 plants were heterozygous, and wild-type M_2 siblings were homozygous for the wild-type (*PALE GREEN*) allele (Figure 3_2 b).

Table 3_2: Segregation analysis of the offspring (M_3) of the individual sequenced M_2 *dwarf/ pale green* (*dw/pg*) and *DWARF/ PALE GREEN* (*DW/PG*) plants. M_3 offspring segregates into mutant (*mt*) and wild-type (*WT*) phenotypes.

#	M₂ Plant (*mt/WT*)	Pollinator	M_3 phenotypes		Germinated	Total
			mutant	WT		
1	*1754_5_pg*	selfed	21	11	32	40
2	*1754_40_pg*	selfed	3	3	6	35
3	*1754_46_pg*	selfed	11	9	20	40
4	*1754_49_pg*	selfed	4	1	5	21
5	*1754_50_pg*	selfed	1	0	1	2
6	*1754_53_pg*	selfed	2	2	4	16
7	*1754_63_pg*	selfed	8	8	16	31

8	*1754_48_PG*	selfed	0	33	33	40
9	*1754_52_PG*	selfed	0	40	40	40
10	*1754_54_PG*	selfed	0	40	40	40
11	*1754_55_PG*	selfed	0	39	39	40
12	*1754_56_PG*	selfed	0	38	38	40
13	*1754_58_PG*	selfed	0	39	39	40
14	*1754_60_PG*	selfed	0	39	39	40
15	*1744_10_dw*	B73	0	22	22	22
16	*1744_23_dw*	selfed	1	0	1	1
17	*1744_28_dw*	B73	0	7	7	7
18	*1744_5_DW*	selfed	12	28	42	42
19	*1744_11_DW*	selfed	10	29	40	40
20	*1744_13_DW*	selfed	9	31	40	40
21	*1744_16_DW*	B73	0	38	40	40
22	*1744_21_DW*	selfed	0	39	40	40
23	*1744_25_DW*	selfed	10	30	40	40
24	*1744_27_DW*	selfed	12	28	42	42

The allele frequency of SNPs expected to be causal for or linked with the mutation was calculated in both individually sequenced populations. For the seven homozygous *dwarf* individuals we expected all 14/14 represented alleles to carry the mutation resulting in a corresponding SNP allele frequency of 1 (Figure 3_2 a). Of the nine sequenced *DWARF* plants, seven were heterozygous and two were homozygous wild-type. Thus, seven out of the 18 sequenced alleles have to carry the mutation resulting in an expected SNP allele frequency of 0.38 (Figure 3_2 a). For the eight heterozygous *pale green* individuals, eight out of the 16 sequenced alleles have to carry the mutation resulting in an expected SNP frequency of 0.5 (Figure 3_2 b). In the eight homozygous wild-type *PALE GREEN* plants, all 16 represented alleles are wild-type and the expected mutant SNP allele frequency thus was 0 (Figure 3_2 b). The measured allele frequencies of all segregating SNPs are plotted for each population together with the state of the zygosity of every SNP in each individually sequenced M_2 plant in Figure S 3_4.

3.3.4 Filtering single-nucleotide polymorphisms for matching allelic states

The precise knowledge of the allelic state of the investigated mutations in each sequenced M_2 individual allowed filtering for SNPs whose state of zygosity exactly matched that of the mutation across all tested individuals (Figure S 3_4).

In the *dwarf/DWARF* M$_2$ family only three SNPs (1:241661229–1:243646893; AGPv3) matched the expected zygosity pattern: They were homozygous for the mutant allele in all seven *dwarf* plants (Table 3_3). In seven *DWARF* plants, the three SNPs were heterozygous and in two *DWARF* plants homozygous for the reference (wt) allele (Table 3_3). Although we were only able to filter for the precise state of zygosity of six *DWARF* plants (Table 3_2), calling the allelic state of these SNPs in every sequenced individual of the *dwarf/DWARF* population unveiled that at least two of the remaining *DWARF* individuals had to be heterozygous and one was homozygous wild-type (Table 3_3).

Table 3_3: SNP Position of the *dwarf/DWARF (dw/DW)* and *pale green/PALE GREEN (pg/PG)* population; Allelic State (*mt/WT*) counts the fraction of *mutant* or wild-type alleles in each population (*pg/PG and dw/DW*); Base exchange from *WT* (identical to B73) at a certain SNP position; Total allele count in both populations (*pg/PG and dw/DW*); Reference allele count of *pg/PG* and *dw/DW* in the whole population, respectively; Mutant allele count in the population with *mutant* and *WT* *p*henotype.

	Population (*mt/WT*)	SNP Position (Chr:Position [nt])	Allelic State (*mt/WT*)		Base exchange (WT/*mt*)	Total allele Count (*pg/PG/dw/DW*)	Ref allele count (WT)	Mutant allele count (*mt/WT*)
			mt	*WT*				
	pg/PG	10:144,036,710	0.5[a]	0[b]	C/T	64	56	8/0
Chr. 10 144 – 144.9 Mb	*pg/PG*	10:144,345,985	0.5[a]	0[b]	C/T	64	56	8/0
	pg/PG	10:144,991,851	0.5[a]	0[b]	C/T	64	56	8/0
	pg/PG	10:144,996,416	0.5[a]	0[b]	C/T	64	56	8/0
	dw/DW	1:241,661,229	1[c]	0.388[d]	G/A	64	43	14/7
Chr. 1 241 – 243 Mb	*dw/DW*	1:243,560,172	1[c]	0.388[d]	C/A	64	43	14/7
	dw/DW	1:243,646,893	1[c]	0.388[d]	G/A	64	43	14/7

a: *pale green* population is heterozygous (50% of alleles were mutant); **b**: *PALE GREEN* population is homozygous WT (0% of alleles were mutant); **c**: *dwarf* population is homozygous mutant (100% of alleles were mutant); **d**: *DWARF* population is either homozygous WT or heterozygous mutant (7/18 or 38.8% of alleles were mutant)

In the *pale green/PALE GREEN* M$_2$ family only four SNPs (10:144036710–10:144996416; AGPv3) matched the expected zygosity pattern and were found to be heterozygous for the mutant allele in the *pale green* plants and homozygous for the reference allele in the *PALE GREEN* plants (Table 3_3). We had no information about the M$_3$ phenotype for one sequenced *pale green* individual and one *PALE GREEN* plant as they failed to produce seeds. However, likewise to the situation in

the dwarf population, the mutant allele count in the sequenced individuals allowed the deduction that the *pale green* plant was heterozygous and the *PALE GREEN* was homozygous wild-type (Table 3_3).

3.3.5 Identification of mutant genes

All three SNPs, which exactly matched the zygosity pattern of the *dwarf* mutation, are located on the long arm of chromosome 1 at 241.6, 243.6, and 243.6 Mbp. One hundred seventy three gene loci of any kind are annotated in this interval ± 400 kb. Only five genes have previously been characterized experimentally, two transcription factors and three protein coding genes, including the *anther ear 1* gene according to the MaizeGDB (Andorf *et al.*, 2016). The mutated recessive *anther ear 1* gene (Zm00001d032961) has been described to cause a similar semi-dwarf phenotype as shown by the *dwarf* mutant in this study (Emerson and Emerson, 1922; Bensen *et al.*, 1995). To test whether the *dwarf* mutation was allelic to the *an1* locus or to the neighboring *D8* locus also responsible for a *dwarf* phenotype, 20 Mbp downstream of *an1*, we performed allelism tests between our *dwarf* mutant and two independent *an1* mutant lines (*116G* and *116GA*) and between three *D8* mutant lines (121C, 131E, and 131F) retrieved from the Maize Genetics Cooperation Stock Center (Scholl *et al.*, 2003; Andorf *et al.*, 2016). We crossed our *dwarf/DWARF* mutants with and without a phenotype to either of the publicly available mutants. Every cross between *dwarf* and either of the *116G-* or *116GA-dwarf* plants produced only offspring (F$_1$) with a dwarf phenotype (#2, 6, 7; Table 3_4), demonstrating that the tested mutations did not complement each other and thereby causing the recessive *dwarf* mutant phenotype of the F$_1$. Crosses between heterozygotes of our *dwarf/DWARF* phenotypes with heterozygotes of both, *116G* and *116GA*, produced *dwarf/DWARF* offspring in a 1:3 ratio (#3, 4; Table 3_4). Homozygous *dwarf* crossed with heterozygous *DWARF* segregated into half *dwarf* and half *DWARF* (#1, 5, 8; Table 3_4). If the *dwarf* allele was absent in at least one of the crossing partner, their offspring only expressed wild-type phenotypes (#9, 10; Table 3_4). From that we concluded that the *dwarf* mutation is fully allelic to both publicly available *an1* mutants.

As a negative control, we crossed *dwarf/DWARF* mutants with plants carrying a mutant dominant *D8* allele (#11–16; Table 3_4). *D8-dwarfs* segregated into 1:1 mutant and wild-type offspring regardless of the phenotype of their crossing partner. *D8-DWARF* plants crossed to *dwarf/DWARF*

only produced wild-type phenotypes. Therefore, the *dwarf* mutation is not allelic to the *D8* locus also present on chromosome 1.

Table 3_4: Allelism test between *dwarf* and public *an1* and *D8* mutants. Female and male parents exhibited either *dwarf* or *DWARF* phenotypes and the offspring of the crosses segregated into *dwarf* and *DWARF* phenotypes.

#	Female plant (dwarf/DWARF)	Male plant (dwarf/DWARF)	dwarf	DWARF	Total
1	*dwarf*	*116G_DWARF*	25	23	48
2	*dwarf*	*116G_dwarf*	41	0	41
3	*DWARF*	*116G_DWARF*	15	33	48
4	*DWARF*	*116GA_DWARF*	11	39	50
5	*dwarf*	*116GA_DWARF*	24	26	50
6	*116GA_dwarf*	*dwarf*	35	0	35
7	*116G_dwarf*	*dwarf*	48	0	48
8	*116G_DWARF*	*dwarf*	20	24	44
9	*dwarf*	*116G_DWARF*	0	48	48
10	*DWARF*	*116G_DWARF*	0	50	50
11	*dwarf*	*121C_D8_DWARF*	0	40	40
12	*DWARF*	*131E_D8_dwarf*	18	22	40
13	*DWARF*	*131E_D8_DWARF*	0	40	40
14	*dwarf*	*131E_D8_dwarf*	17	23	40
15	*dwarf*	*131F_D8_dwarf*	18	12	30
16	*DWARF*	*131F_D8_DWARF*	0	30	30

All four SNPs, which exactly matched the zygosity pattern of the *pale green* mutation, are located on the long arm of chromosome 10 at 144, 144.3, 145, and 145 Mbp. In this interval ± 400 kb, 145 gene loci of any kind are predicted of which seven are transcription factors and three are protein coding genes previously characterized experimentally, including the *white seedling 2* gene, which expressed a similar mutant phenotype according to the MaizeGDB (Andorf *et al.*, 2016). One of the *pg/PG* SNPs (10:144036710) causes a nonsynonymous amino acid exchange in the derived amino acid sequence of the Zm00001d026402 gene corresponding to the *white seedling 2* locus. This locus was published to condition shades of pale green mutant phenotypes, dependent on the homo- or heteroallelic state of the mutation, caused by a defective nuclear encoded chloroplast DNA polymerase, necessary for the replication of the chloroplast DNA (Udy *et al.*, 2012). Thus, the relative amount of chloroplast DNA in mutant (*pale green*) vs. wild-type (*PALE GREEN*) plants was analyzed by qPCR using primers for specific chloroplast encoded genes.
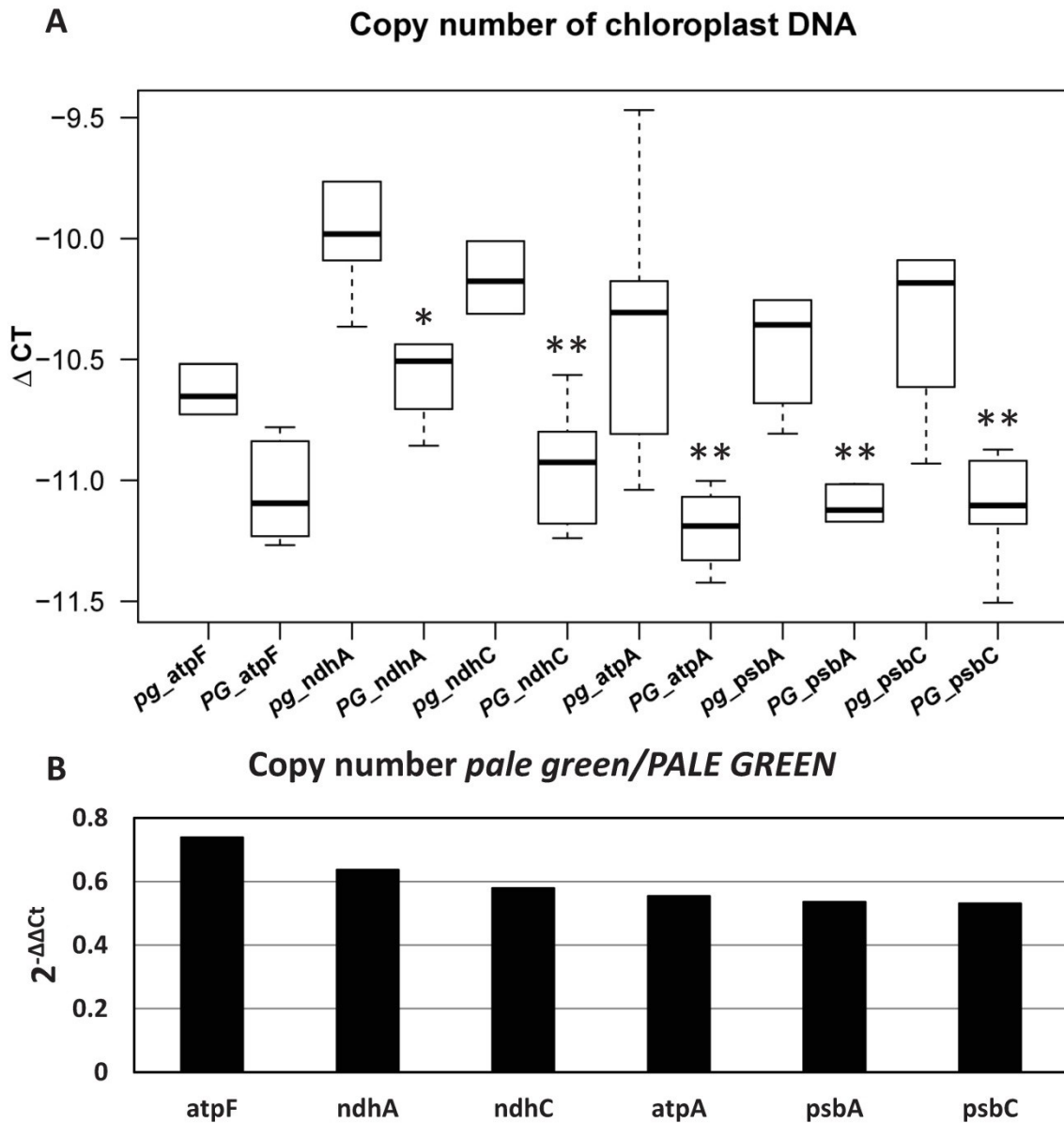
Figure 3_3: Chloroplast DNA copy number variation between *pale green* and *PALE GREEN* mutants. (**A**) ΔCt values (cpGOI – nuclear reference gene) of six chloroplast encoded genes; asterisks indicate significance (n=6; unpaired *t*-test * ≤ 0.05 and ** ≤ 0.01 between ΔCt(cpGOI) *pale green* and ΔCt(cpGOI) *PALE GREEN*. (**B**) cpDNA levels of *pale green* relative to *PALE GREEN* plotted as $2^{-\Delta\Delta Ct}$ (mean ΔCt *pale green* – mean ΔCt *PALE GREEN*)

The cycle threshold values (Ct) for each chloroplast gene were normalized to Ct values of a single copy nuclear gene, in other words, the chloroplast DNA levels were normalized to the amount of nuclear DNA to correct for differences in concentration, resulting in delta Ct (ΔCt) values (Figure 3_3 a). Copy number of chloroplast DNA in *pale green* plants reached between 73.9% (atpF) and

53.2% (psbC) relative to the cpDNA levels of the *PALE GREEN* mutants (Figure 3_3 a, b). The amino acid exchange of the 810th amino acid (serine to phenylalanine) was predicted to be of deleterious nature (PROVEAN score -5.8), which can be interpreted as a negative effect on the function or even a loss of function for the mutated protein (Choi *et al.*, 2012; Choi and Chan, 2015). Furthermore, the nonsynonymous amino acid exchange at the given position is located in the DNA polymerase A domain part of the protein (The UniProt Consortium, 2017).

3.3.6 Full length sequence analysis of the affected genes

The SNP in the *white seedling 2* gene found in the *pale green* mutant caused an altered sequence of the encoded protein and causes a non-synonymous amino acid exchange. However, it could not be excluded that additional EMS-induced sequence alterations were present in this gene. Furthermore, the detected SNPs associated with the recessive *dwarf* mutation rather acted as markers than as the cause of the mutant phenotype, as the *an1* gene is found 300 kb upstream of the SNP at position 1:241,661,229. Therefore, we re-sequenced the *dwarf* allele of the *anther ear 1* gene and the *pale green* allele of the *white seedling 2* gene and the corresponding wild-type PH207 alleles.

An insertion was found to be present at position 259 of the large exon (Zm00001d032961_T003, AGPv4) (Andorf *et al.*, 2016) of the *anther ear 1* gene in the *dwarf* mutant, which was absent in the *DWARF* and PH207 wild-type sequence (Figure S 3_2; Primer 612f, Alignment position 3111 ff). According to BLAST and the P-MITE database the detected additional 256-bp sequence had a 95% sequence identity with a DTH_Zem60 DNA transposon belonging to the PIF/Harbinger superfamily with a published length of 133 kb (Altschul *et al.*, 1997; Chen *et al.*, 2014).

Despite some sequencing gaps especially in the intron regions, the genomic sequence of the *pale green* gene was consistent with the PH207 sequence from the Genome Browser (Figure S 3_3)."[1]

END OF CITATION (Heuermann *et al.*, 2019)

3.3.7 Summary of the rapid mapping

Genomic DNA from 32 EMS-induced M_2 mutant plants in the *Zea mays* PH207 genotype, of one population of plants with seven *dwarf*/nine *DWARF* and one population of plants with eight *pale green*/eight *PALE GREEN* phenotypes, were individually whole-genome sequenced. SNPs were

called against the B73 reference sequence (AGPv3) and subsequently filtered for segregating SNPs in both populations, resulting in 1,357 and 3,536 segregating SNPs, respectively. The state of zygosity of the *dwarf* and *pale green* mutation in the individually sequenced M_2 plants was determined by analyzing the segregation of the mutant phenotype in their self-pollinated offspring (M_3). The SNPs causal for or linked with the mutation were determined by only keeping SNPs if they followed the exact state of zygosity of each individually sequenced M_2 plant. The zygosity filter identified three and four SNPs in linkage with the homozygous *dwarf* and heterozygous *pale green* mutation, respectively. An insertion in the *AN1* gene was found to be responsible for the *dwarf* and a non-synonymous amino acid exchange in the *W2* gene for the *pale green* mutation.

4. DISCUSSION

The aim of this work was to monitor the growth performance of plants exposed to dynamically changing environmental conditions and to detect genetic factors that specifically affect growth under fluctuating illumination. The main part of this work was thus devoted to a comparative genome-wide association study in *Arabidopsis thaliana* using plants exposed to either constant or fluctuating light to uncover determinants of the fast acclimatization reactions underlying variation in growth performance.

Two supporting methodical advancements were explored and verified. First, the Plant Cultivation Hall was tested to simulate dynamic and field-relevant environments, which will later help to study the performance of crops in field-like experiments under expected future climate change scenarios. Second, a next-generation sequencing based mapping method was developed to rapidly identify mutant genes/alleles requiring only a small number of individuals to provide an efficient tool for forward genetic studies in crop species.

In contrast to the field, where a plethora of simultaneously acting dynamic environmental factors affect the phenotype of a plant, the controlled modulation of only one environmental factor, here fluctuating versus constant light, enables the attribution of conditional phenotypic variation to allelic variation found in a natural population. The effects of other environmental factors like temperature, CO_2 concentration, and humidity were identical in both treatment and thus did not affect the phenotypes of plants. In this study the IPK automated non-invasive plant phenotyping system for small plants was upgraded with LED light bars to enable high-throughput phenotyping of an *Arabidopsis thaliana* accession panel under fluctuating light condition and to associate the resulting phenotypic variation with the natural genetic variation present in the population. In two independent experiments, 1646MH and 1703MH, 384 accessions were grown in parallel under fluctuating and constant light in the same growth chamber for 21 days and daily high-throughput phenotyping recorded the phenes of the plants.

4.1 The design of the fluctuating light conditions

The constant light conditions were comparable to standard indoor cultivation scenarios of *Arabidopsis thaliana* (Meyer *et al.*, 2007; Junker *et al.*, 2015; Tschiersch *et al.*, 2017). The

fluctuating light treatment was chosen to deliver alternating high and low light conditions with distinct and immediate switches between the two settings. The idea was to mimic the effects of wind movements of understory or canopy movement (Salisbury *et al.*, 1968; Young and Smith, 1983; Pearcy, 1990; Pearcy *et al.*, 1990; Slattery *et al.*, 2018; Townsend *et al.*, 2018) and of cloud cover, which both can lead to frequent and rapid immediate changes in light intensity over the course of a day (Kulheim *et al.*, 2002; Kaiser *et al.*, 2018; Matsubara, 2018). Plants do develop a light memory effect which is determined by the levels of de-epoxidation of violaxanthin to zeaxanthin, when being exposed to different light intensities for a specific time (Matuszyńska *et al.*, 2016). Taking this into account, low and high light intervals of alternating length were randomly distributed over the course of a day and randomized between days to prevent plants from potentially adapting to reoccurring patterns.

Automatic movement of plant carriers in the IPK automated non-invasive plant phenotyping system for small plants is based on a conveyor belt transportation system (Junker *et al.*, 2015), which means that every carrier has to occupy every position once in the chamber while automatic phenotyping occurs. Consequently carriers would have to pass through the opposing light treatments, which were set-up in the two halves of the chamber, while doing so. Moreover, the possibility of the system to position the plants differently each day to dilute potential positioning effects in the growth chamber was lost. Over the course of the experiment a specific carrier had to stay at the same position in the growth chamber. Phenotypes therefore had to be recorded during the night so that plants never experienced a contrasting light treatment. Thus potential position effects for example caused by minor inhomogeneity in light intensities (see Figure S 1_1) were not diluted but accumulated throughout the experiment. To account for positioning effects, time point of germination, and other systematic effects, which cannot be accounted for by prior knowledge, the best linear unbiased estimators (BLUEs) were calculated to extract the genotypic means for further genetic association studies (Möhring and Piepho, 2009; He *et al.*, 2016). The linear mixed models were fitted to dissect the variance components in the dataset and extract only the genetic contribution as fixed factors to the phenotypes. The dimension of the genetic contribution itself to a phenotype was expressed by the repeatability value, which was calculated from the dissected variance components from the linear mixed models treating the genotypes also as random factors. Hence, traits with higher repeatability values could attribute more variance to genetic differences

and less variance to environmental effects. Best repeatability was achieved for BLUEs calculated in a one-stage analysis, a unified linear mixed model treating the two experiments as random factors, which yielded more robust BLUEs than calculating the BLUEs for each experiment individually. This effect can be explained by a larger number of replicates tested and by more phenotypic plasticity being explored by the accessions due to slightly different environmental effects experienced in the two independent experiments, even though the IPK automated non-invasive plant phenotyping system for small plants is designed to deliver stable repeatable conditions. The finding that phenotyping in two or more different experiments leads to a more robust representation of the phenotype of a genotype is in concordance with Zhao *et al.* (2015).

The BLUEs were able to account for variance components, which affected the final phenotypes. However, only variance components inside a statistical group, here the two light treatments, were accounted for. Differences that overlapped with the treatments, like differences in spectra of the light sources, cannot be distinguished from the biological variance. Plants do respond differently to diverse light sources and light spectra, which can significantly alter total biomass accumulation (Hogewoning *et al.*, 2010). Therefore, the spectra between the two lighting set-ups were checked (Figure S 1_1) and the combination of metal halide lamps shaded with a mesh and supplementary LEDs for the fluctuating light treatment, showed similar characteristics between the two sides. Despite the LED light bars supplementing the shaded main illumination, no differences in plant area and dry weight were detectable in experiment 1806MH under continuous equal levels of light intensities between the sides of the growth chamber. Thus, the differences detected in experiments 1646Mh and 1703MH can be solely contributed to the treatment with fluctuating light.

4.2 Arabidopsis' response to fluctuating light treatment

When exposed to fluctuating light, the Arabidopsis accessions responded with significantly lower leaf area and dry matter accumulation. Plants germinated under their corresponding light regimes and slower growth was already observable at five days after germination, the earliest time point recorded. The initial difference in leaf area was becoming more pronounced over the course of the experiment, which was true for the two independent experiments. The observation that Arabidopsis grows slower under fluctuating light is consistent with results reported previously.

Smaller plant size for plants grown under natural light was reported by Kulheim *et al.* (2002) and Mishra *et al.* (2012), and reduced plant size was also observed under indoor grown plants subjected to a fluctuating light pattern (Vialet-Chabrand *et al.*, 2017; Schneider *et al.*, 2019). The image data were recorded on a daily basis and analyzed for overlapping time points between both experiments (1646MH and 1703MH). Two hundred and eighty architectural and color related traits could be extracted with the IAP software for both experiments. Differences between the fluctuating and constant light treatment were analyzed with principal component analysis individually for each time point. Principal components one, two, and three were able to discriminate between the effects of the light treatments on image related traits from the earliest time point. In agreement with the observed increase in the differential in dry matter accumulation at later time points, the discriminating power between the two light treatments by principal component one increased in later time points. Again, this effect was reproducible in both experiments and supports the effectiveness of the applied fluctuating light treatment in this study. The discriminative power between light regimes by principal component analysis observed in this study is in line with observations of Annunziata *et al.* (2018) who were able to distinguish between different light regimes using metabolic phenotypes.

Kinetic chlorophyll fluorescence analysis was performed using the FluorCam device (Tschiersch *et al.*, 2017). The throughput of the FluorCam device would allow measuring the full set of carrier in one night, if only PSII-operating efficiency was measured as a single parameter, but is not high enough to process 384 plant carriers in the eight hour dark period measuring multiple parameters. As parameters like the non-photochemical quenching potential (NPQ) were previously reported to be an important parameter for describing the biological response of plants under fluctuating light or high light treatment (Kulheim *et al.*, 2002; Allahverdiyeva *et al.*, 2015; Yamori, 2016; Yamori *et al.*, 2016; Vialet-Chabrand *et al.*, 2017; Slattery *et al.*, 2018; Townsend *et al.*, 2018; Schneider *et al.*, 2019), a FluorCam measuring protocol suitable to measure NPQ was used. The full set of plants was measured at four consecutive nights and data were analyzed by calculating the BLUEs accounting for the different measuring days. The genetic variation contributing to the phenotype was successfully extracted as indicated by the high repeatability values and the differential data distribution between the two treatments (Figure 1_8). PSII-operating efficiency was found to be significantly reduced while NPQ potential was significantly increased in plants grown under

fluctuating light conditions, a finding in agreement with other studies on Arabidopsis grown under fluctuating light (Mishra *et al.*, 2012; Schneider *et al.*, 2019).

In summary, the phenotypic responses to fluctuating light, measured by the RGB and fluorescence cameras and the FluorCam device, were consistent with results reported in previous studies on this subject. Thus, the BLUEs from the phenotypic traits over both independent experiments were then subjected to a genome-wide study to associate the phenotypic variation between the genotypes with the genotypic variation.

4.3 The Arabidopsis accession panel was appropriate for genome-wide association study

A powerful and suitable experimental design is the basis to study Arabidopsis' response to fluctuating light condition in a GWAS. But equally or even more important for a successful GWAS, is a suitable natural population. In this study a natural panel of 384 *Arabidopsis thaliana* accessions, with geographic origin from North America, Africa, Europe, Eurasia, Japan, and New Zealand was used. The panel was selected by the Heterosis research group at the IPK to cover the natural phenotypic variation and genetic variation while reaching a suitable population size for the IPK automated non-invasive plant phenotyping system for small plants. While the first principal component described the variation in the whole population without any separation, the second and third principal components could distinguish between the origins of the accessions from North America and from Asia versus the rest of the population, respectively, without separating them. No distinct patterning between the accessions was observed in the first four principal components and only 10% of the genotypic variation was explained, indicating a low population structure in the panel. The subset of the accession panel was designed to express low population structure, which was supported by the PCA. Principal component analysis can be misinterpreted if long range LD regions, the physical linkage between markers in a population, are present or SNP data is of low quality (Price *et al.*, 2010), which was both not relevant in this population. Beside low population structure and phenotypic variation, marker quality and density are the determining factors of GWA resolution. The accession panel was genotyped by 214,051 SNPs, which were homogeneously distributed over all five Arabidopsis chromosomes (Figure 1_4). The average distance between markers (557 bp) would be low enough to associate quantitative phenotypic

traits with single Arabidopsis genome initiative gene identifiers. Nonetheless, LD limits the possible resolution of a GWA to distinguish a causal marker from markers linked with the causal marker (Bush and Moore, 2012). In Arabidopsis, linkage is published to decay on average within ten kb, which was calculated for a population consisting of 19 accessions (Kim *et al.*, 2007). The population used to study the response of fluctuating light was substantially larger and thus half decay of the LD was calculated based on pairwise distances by allelic correlation (r^2). Due to the availability of a modern server cluster infrastructure it was not necessary to approximate the LD decay by defining bins of SNPs, but to directly calculate the pairwise distance matrices for each chromosome considering every homozygous SNP present. Thereby, the precise genome-wide LD decay could be reconstructed and the half maximum of the modeled decay in this Arabidopsis population was found to occur after 3,365 bases on average. A full LD block around a marker adds up to 6,730 bp for the given population and marker coverage, which allows for a threefold higher resolution in the GWA study than in the previously reported 20 kb LD block in the smaller population (Kim *et al.*, 2007).

4.4 Genome-wide association of the image related-traits

The GWAS was performed with the *FarmCPU* R package (Liu *et al.*, 2016) which has been proven to be a suitable tool to study marker traits associations in plants (Dai *et al.*, 2017; Morosini *et al.*, 2017; Gage *et al.*, 2018; Xu *et al.*, 2018; Carlson *et al.*, 2019). In the interest of its computational power and speed, *FarmCPU* was chosen in this study. The temporally resolved (six overlapping time points) high-throughput phenotyping (286 traits IAP; 78 traits FluorCam) of the Arabidopsis accessions under two light conditions, which were genotyped for 214,051 SNPs, was resulting in about one billion marker-trait association tests, which had to be analyzed. *FarmCPU* is not only characterized by its improved statistical power but by its ability to operate with huge data sets as its computing time scales linear with the number of both, traits and markers, in comparison to exponential scaling of conventional methods (Liu *et al.*, 2016). Thus by applying this package, the full potential of the IPK automated non-invasive plant phenotyping system for small plants to record the phenotypic variation of the Arabidopsis accession panel could be exploited for GWAS.

The combination of eight time points, nearly 300 traits and two light treatments resulted in the detection of thousands of significant MTAs. To extract the most relevant data out of thousands of

MTAs the data were filtered. The filters were designed to only keep phenotypic traits which can be attributed to biological explained variation regarding the traits of interest (traits yielding MTAs under both light conditions) and were under reasonable genetic control ($R_n > 0.5$). Thereby, two third of the traits were discarded. The chosen threshold is debatable as also traits with lower repeatability values were successfully associated with genetic variation (Filiault and Maloof, 2012) and potential interesting associations could have been missed. However, due to the large number of associations found with the strict filter, the threshold was deemed appropriate.

For remaining MTAs the constant light condition was treated as a control and only MTAs specific for the fluctuating light treatment were accepted. Linkage disequilibrium was detected in the population and it was therefore necessary that every MTA was annotated with several Arabidopsis gene identifiers of any kind in range of the average genome-wide half maximum modelled LD decay up- and downstream of the associated SNPs.

Taking advantage of the annotated gene identifiers allowed filtering for their published biological function. It was expected that only one of the annotated gene identifiers in a certain stretch of DNA was actually causative for the observed phenotype, and the other gene identifiers to be just in LD. To find those causal gene identifiers, the annotated genes were filtered to have gene ontology with the terms "photo" or "light". This restricted the identification to genes that had been already published to be involved with photosynthesis or light responses in other experiments. Candidates with unknown function would be missed by such a filter. Genes, who were not directly annotated with light responses but other regulatory functions, could be potentially involved in response to fluctuating light but these would also be missed. Hence, the full table of unfiltered gene identifiers for every fluctuating light-specific MTA was recorded and can be found in Table S 1_3. Although potential candidates were missed by the GO filter, 42 candidate genes, associated with fluctuating light treatment-specific traits, were found from RGB and fluorescence imaging and measuring kinetic chlorophyll fluorescence using this strategy (Table 1_2 IAP; Table 1_3 FluorCam).

4.5 Fluctuating light specific candidate genes

Two geometrical traits, hull.pc1 and hull.pc2, which describe the maximal length and width of a rosette, respectively, were significantly associated with genes passing the GO filter. Both traits

may be affected by petiole length, which is responsive to light intensity (Tsukaya *et al.*, 2002), and were significantly correlated with the trait leaf area (hull.pc1: DAG5, r = 0.71, p < 0.001 and DAG6, r = 0.9, p < 0.001; hull.pc2: DAG12, r = 0.78, p < 0.001). Five days after germination the acyl-CoA-binding domain 3 gene (*ACBP3*, *AT4G24230*) was found to be significantly associated with hull.pc1 (maximal rosette diameter) under fluctuating light treatment. ACBP3 belongs to a family of six acyl-CoA-binding proteins (Xiao and Chye, 2009), and responds to abiotic stress like absence of light and is under circadian control (Xiao *et al.*, 2010; Zheng *et al.*, 2012). The gene is ubiquitously expressed in Arabidopsis and functions in the regulation of leaf senescence by modulating membrane phospholipid metabolism and stability of the autophagy (ATG)- related protein ATG8 (Xiao *et al.*, 2010; Zheng *et al.*, 2012).

The same trait, hull.pc1, was significantly associated one day later (DAG6) with the *PIL6* gene, also called *PIF5* (*AT3G59060*), the phytochrome interacting factor 3-like 6, which represents a basic helix-loop helix (bHLH) transcription factor. Extensive research has been published discussing the function of *PIL6*, suggesting that the gene is under circadian control, its protein stability is light controlled and involved in differentially-phased leaf growth (Takase *et al.*, 2013; Dornbusch *et al.*, 2014; Seaton *et al.*, 2015). Furthermore *PIL6* regulates dark-induced leaf senescence (Sakuraba *et al.*, 2014; Song *et al.*, 2014; Zhang *et al.*, 2015) and mediates growth by controlling auxin (Nozue *et al.*, 2011; Hornitschek *et al.*, 2012) and gibberellic acid signaling (Filo *et al.*, 2015). In a GWAS with 180 Arabidopsis accessions, 115 of which overlapped with this study, *PIL6* was found to be associated with hypocotyl growth under shade avoidance (Filiault and Maloof, 2012). The identified natural variant is common, minor allele frequency 0.37, and the allele positively (norm. effect 0.02) contributes to rosette diameter under fluctuating light conditions. The overlap of described function between the two candidate genes, mediating plant size by controlling leaf senescence and differential growth expansion as a response to shade-avoidance treatments, suggest a similar response to fluctuating light treatment.

The second architectural trait, hull.pc2, was found to be significantly associated with two genes, *mtLPD1* (*AT1G48030*) and *AGL18* (*AT3G57390*), in later stages of the experiment (DAG12). The *mtLPD1* gene, a nuclear encoded mitochondrial lipoamide dehydrogenase 1, responds to light stimulus and affects steady state contents of the tricarboxylic acid cycle, photorespiration, CO_2 assimilation and plant growth and its regulation is dependent on light-capture and light use

efficiency (Timm *et al.*, 2015). AGL18 is a MADS-box containing protein, likely a transcription factor, involved in regulating and arresting flowering initiation in vegetative organs and it also affects leaf morphology (Adamczyk *et al.*, 2007; Fernandez *et al.*, 2014; Serivichyaswat *et al.*, 2015). Moreover, AGL18 is published to be a substrate of MPK3 (Popescu *et al.*, 2009), a mitogen-activated protein kinase which together with MPK6 regulates stomatal development and patterning as response to environmental stimuli (Wang *et al.*, 2007). Furthermore, it regulates reactive oxygen species-responsive gene transcription under a fluctuating light environment (Wang *et al.*, 2013). The two candidate genes found for the MTA with hull.pc2 suggest that Arabidopsis responds to fluctuating light in the vegetative state by affecting photorespiration and CO_2 assimilation in the mitochondria and plant growth in general. Minor allele frequencies (0.2 and 0.43) suggested that the associated alleles are common in the natural accessions panel.

Next to the two architectural traits, four color-related traits were found to be significantly associated with genes that passed the GO filter. The earliest color-related trait found was hsv.norm.h.his.bin.06.063.076.vis.top, which gives information about the number of pixels normalized for the leaf area, which fall into the sixth color bin representing a green color in the hsv color space. Five and six days after germination, the same trait was found to be associated with the flavonol synthase 3 gene *FLS3* (*AT5G63590*). Flavonol glycosides, products from FLS, were found to function in protecting plants against UV light (Jin *et al.*, 2000). Subsequently, Hartmann *et al.*, (2005) discovered light-responsive promoter sequences of importance for the coregulation of *FLS* pathway genes, which are targets of transcription factors for example belonging to the basic helix-loop helix family (bHLH). Transcript levels of the *FLS3* gene were found to be higher in light than in dark grown plants and in plants subjected to high light treatment (Nakamura *et al.*, 2007; Lillo *et al.*, 2008). Under standard greenhouse conditions no phenotype could be attributed to reduced FLS3 protein levels (Stracke *et al.*, 2009). *FLS3* was found to be the closest homolog to *FLS1*, and to encode a functional protein (Stracke *et al.*, 2009; Preuss *et al.*, 2009), whereas the *FLS2* gene in perfect LD to *FLS3* encodes a pseudogene in Col-0 (Stracke *et al.*, 2009). *FLS1* is located more than 20 Mbp upstream on the same chromosome, making *FLS3* the most likely candidate for the association. Flavonoids are published to be the major pigments in plants for the colors red, blue, and purple (Winkel-Shirley, 2001), which if altered could explain a lower share of green colored pixels found for the associated allele.

The second color-related MTA was found between the trait hsv.norm.v.hist.bin.03.025.038.vis.top, number of pixels in the yellow color space, and the gene thylakoid rhodanese-like (*TROL*, *AT4G01050*) at seven days after germination. TROL is required for tethering the flavoenzyme ferredoxin:NADP$^+$ oxidoreductase (FNR) to the thylakoid membrane to sustain efficient linear electron flow (LEF) from H_2O to NADP$^+$ (Jurić *et al.*, 2009). Knock out of *trol* lowers electron transport rates under high-light and leads to higher NPQ levels (Jurić *et al.*, 2009). Furthermore, *trol* mutant phenotypes were smaller and had yellowish siliques and inflorescences (Jurić *et al.*, 2009), which is in concordance with the phenotypic trait measuring yellow pixels. The minor allele frequency of this allele was high (0.45), suggesting that in natural accessions *TROL* is a common target for responsiveness to fluctuating light and might affect the maintenance of linear electron flow under dynamic light conditions.

The trait lab.a.skewness.fluor.top was the third color-related trait and was associated with phytochrome C (*PHYC*, *AT5G35840*). The trait measures the skewness of the red color channel in the L*a*b color space. PHYC was found to interact with PIL6/PIF5 (Khanna *et al.*, 2007), which was also found to be associated with fluctuating light treatment by hull.pc1 in this study. Like *PIL6*, *PHYC* was found to be significantly associated with shade avoidance in a natural accession panel (Filiault and Maloof, 2012). Méndez-Vigo *et al.*, (2011) found environment-genotype association in PHYC to be associated with winter temperatures and spring and winter precipitation, suggesting that allelic variation in *PHYC* is involved in climatic adaptation. PHYC among other phytochromes regulates the Arabidopsis circadian clock in a light-dependent manner (Hu *et al.*, 2013). In response to the prevailing light environment, phytochromes arrest or promote progression of plant development (Hu *et al.*, 2013). The interaction of PIL6 and PHYC, which is likely mediated by PIL6 controlling the degradation of PHYC (Khanna *et al.*, 2007), and the fact that both genes were found in this study in a natural accession panel grown under fluctuating light conditions, suggest an important role for both genes to regulate and control Arabidopsis' response to fluctuating light treatment.

The last MTA (DAG13) was found between the trait hsv.v.hist.h.avg.bin.08.089.102.vis.top and the auxin-responsive GH3 family protein (*JAR1*, *AT2G46370*). The trait encodes color intensity values for green-bluish colored pixels. *Fin219*, later found to be a mutated allele of *JAR1* (Staswick *et al.*, 2002), attributed a critical role to *JAR1*, it regulates the phyA-mediated far-red inactivation of

COP1 and embodies the possible cross-talk juncture between auxin regulation and phytochrome signaling (Hsieh *et al.*, 2000). COP1 itself was found to be involved in controlling sensitivity to shade in fluctuating light conditions of plant canopies (Pacín *et al.*, 2013). Reduced sensitivity to far-red light was observed in *jar1* mutants (Chen *et al.*, 2007). In rice, longer coleoptile phenotypes for *osjar1* mutants were only observed in specific far-red and blue light conditions (Svyatyna *et al.*, 2014). In response to far-red light JAR1 regulates a number of transcription factors including 94 basic helix-loop helix (bHLH) TFs (Chen *et al.*, 2015), notably one of which is the *PIL6*/*PIF5* gene found to be associated with the hull.pc1 trait six days after germination in this study.

The role of phytochrome A signaling in response to fluctuating light treatment can be further supported by looking back into the full GO filtered list (Table S 1_2), where the traits hsv.norm.v.his.bin.05.051.063.vis.top., number of green colored pixels, and lab.l.skewness.vis.top., skewness of lightness distribution, were found to be associated with *FRS1* (*AT4G19990*) and *FRS10* (*AT5G28530*), respectively. The gene family of FAR1-related sequence (FRS) genes is published to play a phyA-mediated role in light control in Arabidopsis development, likely by regulating nuclear gene expression as most FRS genes contain a nuclear localization sequence (Wang and Deng, 2002; Lin and Wang, 2004).

The kinetic fluorescence parameters yielded two MTAs between the chlorophyll fluorescence decrease ratio (Rfd) and two candidate genes that passed the GO filter. The Rfd trait, which is directly correlated with the CO_2 fixation rate in leaves, was published to act as an indicator of photosynthetic rates and to be higher in sun than in shade exposed leaves of trees (Lichtenthaler *et al.*, 2007). A similar trend was observed in Arabidopsis, despite being a different species, that leaves of plants grown under fluctuating light treatment had significantly increased Rfd values compared to leaves of plants grown under constant light (Figure 1_8). It would be interesting to measure Rfd in sun and shade exposed Arabidopsis leaves to compare fluctuating light conditions to sunlight exposure. One of the two genes Rfd was associated with, was the thylakoid localized DnaJ/Hsp40 cysteine-rich domain superfamily protein *LQY1* (*AT1G75690*). LQY1 is a small zinc finger protein, which interacts with the PSII core complex and responds to high light treatment (Lu *et al.*, 2011). LQY1 was found to comigrate with PSII core monomers and the CP43-less PSII monomer, which is a marker for PSII repair and reassembly (Lu *et al.*, 2011). The role of LQY1 is predicted to maintain PSII activity under high light (Lu *et al.*, 2011). Homologs of *LQY1* are found in

many land plants like rice (*Oryza sativa*) or great millet (*Sorghum bicolor*), but not in algae or waterborne plants, which may reflect an evolutionary adaptation to excess light stress of plants during transition to the land (Lu, 2011). Mutants of *lqy1* were found to have higher NPQ values under high light exposure (Lu, 2011). The role of LQY1 in the repair of PSII was supported by Jin *et al.*, (2014), who found that it together with hypersensitive to high-light 1 (HHL1) forms a complex and participates in the repair of photo damage.

The second gene found to associate with Rfd was the Photosystem II reaction center PsbP family protein (*PsbP*, *AT3G05410*). This gene was found to belong to the PsbP superfamily but to be distant to the other PsbP proteins (Sato, 2010). In an analysis of the function of the members of this family *PsbP* was labeled *PPD7*, but no protein levels were detectable and no function could be attributed (Bricker *et al.*, 2013). Other PsbP proteins like PsbP-1 and PsbP-2 are required for the assembly of the PSII complex and its stability (Yi *et al.*, 2007). The marker-trait association with this exotic *PsbP*/*PPD7* gene may suggest a function present under fluctuating light conditions.

The validity of detecting the two above mentioned genes by associating the chlorophyll fluorescence parameter Rfd with the natural accession panel, was supported by the fact that *LQY1* and *PsbP* family genes were found to be measurable by chlorophyll fluorescence measurements in other studies (Lu *et al.*, 2011; Yi *et al.*, 2007). Minor allele frequencies were high (0.44 and 0.48) for both alleles suggesting an important role in potentially maintaining PSII core complexes as a response to fluctuating light conditions. Despite the high allele frequency, the phenotypic variance explained by these MTAs (2.7 and 1.4%, respectively) was low, which indicated that the effect on the phenotype, and thus the selective pressure, was likely not large enough to remove the negative allele from the population.

The GWAS found thousands of MTAs, which resulted in the 11 above discussed most promising candidates after application of the gene ontology filter for "photo" and "light". For all of the ten unique genes associated with the 11 MTAs, polymorphic SNPs and for eight genes indels were found, indicating that all are targets of natural variation (Alonso-Blanco *et al.*, 2016). Phenotypic variation associated with fluctuating light could be attributed to genes like *FLS3* and *PsbP*/*PPD7*, for which until now no phenotypes were reported. Three candidate genes, the *PIL6* transcription factor together with the *PHYC* and the *JAR1* gene were all published to act together in

phytochrome signaling in response to far-red light or shade avoidance suggesting that Arabidopsis' response to fluctuating light may be mediated by the same pathways and signaling cascades. The signaling role of phytochromes in response to changing light environments in Arabidopsis is further supported by Ushijima *et al.*, (2017) who could link phytochromes to genome-wide changes in alternative promoter selection to modulate protein localization, thereby allowing plants to adapt to light environments.

The *PIL6* transcription factor belongs to the basic helix-loop helix transcription factor family (bHLH), which are also published to control *FLS* genes (Hartmann *et al.*, 2005), among which the *FLS3* was found to be significantly associated with fluctuating light. In the unfiltered MTA table (Table S 1_3) nine MTAs, over all time points, were found to be in LD with other annotated *bHLH* transcription factors, of which seven unique members of this family were found to be significantly associated with fluctuating light treatment (*AT1G10610*, *AT1G12860*, *AT1G25330*, *AT1G27660*, *AT3G07340*, *AT5G50915*, *AT5G54680*). Those MTA were missed by the GO filter as so far no ontology which contained the filter words "photo" or "light" are annotated, but six out of seven (except *AT1G12860*) were found to be targets of JAR1 in response to far-red light treatment (Chen *et al.*, 2015). This further supports the essential role of the *bHLH* transcription factor family in mediating Arabidopsis' response to light in general and in particular to fluctuating light. The candidate genes extracted from the Rfd trait, *LQY1* and *PsbP*/*PPD7*, suggest that fluctuating light treatment affects the PSII complex, as both genes are likely involved in maintenance and repair. Another MTA was found for the trait hull.length.geom.trait.fluor.top seven days after germination (Table S 1_2) with the gene *LPA19* (*AT1G05385*), which was also published to be involved in PSII biogenesis (Wei *et al.*, 2010).

The GWAS was rigorously filtered to deem only MTAs to be significant, when they passed a conservative Bonferroni correction, which reduced the return of false positives but may result in false negatives. This may explain that members of the same signaling cascade like *PIL6*, *PHYC*, and *JAR1* were not found to be significant at the same time point.

A major advantage of non-invasive high-throughput phenotyping performed at multiple time points is the detection of temporal QTL actions, which were published to be a common occurrence in plant species like maize and canola (Muraya *et al.*, 2017; Knoch *et al.*, 2019). Most associations

would have been missed if phenotyping had been performed only at one developmental time point. Moreover, not all traits were precisely measurable at all time points which resulted in reduced repeatability for example in early traits regarding leaf morphology (Figure 1_11). In such instances, more variation is attributed to residual than to genetic origin, which makes it hard to significantly associate such traits with allelic variation.

The vast number of significant MTAs, which were not further explored in this study, suggests that fluctuating light treatment is causing a systemic remodeling of pathways and signal cascades. The phenotypic variance explained (PVE) by single MTAs was ranging from < 1% to 14.12% and the average PVE of the fluctuating light specific MTAs and the subsequently GO filtered MTAs was 2.35% and 2.39%, respectively. This fits the expectation that complex traits like adaptation processes, here to fluctuating light, are not caused by natural variants of few genes which would be responsible for the majority of PVE, but by many genes all contributing a small portion of PVE to the global changes a plant has to undergo while adapting to a certain environment (Ingvarsson and Street, 2011).

The here applied rigorous filter of annotated genes by factors like gene ontology was leading to significant knowledge because of the high quality annotation of the *Arabidopsis thaliana* genome and the low LD range in the underlying population, which reduced the search intervals to very few genes of interest. In other plant species, like in many crop plants, where LD blocks are large, genes were annotated *in silico* and the marker quality is often poor, the usefulness of such a filter would be greatly reduced.

In future experiments, results from this GWAS can be used to select a smaller set of contrasting accessions for treatment-associated traits and perform untargeted metabolite profiling in combination with labeling experiments on such selected lines under both light conditions, which would enable metabolic flux analysis and help to visualize differences in the distribution of metabolic resources into pathways like flavonoid biosynthesis. Knock-out and overexpression mutants of candidate genes, like the eight members of the *bHLH* gene family, could be screened for novel phenotypes under fluctuating light conditions, for example the reduced yield found in the *npq* mutants under variable light (Kulheim *et al.*, 2002). Alleles which were found to be beneficial for the adaptation to dynamic light conditions in Arabidopsis could be ultimately

transferred to crop systems, after the presence of homologs in such crop species was checked, and help to breed better adapted crop varieties.

4.6 Environments in the field are dynamic

Crop plants are naturally exposed to dynamic environments and fluctuating light when grown in the field. Modern crop cultivars have an extensive breeding history in which breeders explored genetic variation using multi-environment trials studying genotype-environment interaction over many years to increase response to selection and breed new and better adapted cultivars (Ceccarelli, 2015). Crop production and yield are threatened by the ongoing climate change, which increases global temperatures, CO_2 levels and the frequency of droughts (Ceccarelli *et al.*, 2010; Henry and Nevo, 2014). The understanding of how wild populations evolved to those threats in nature can help breeding efforts to generate climate-resilient crop varieties (Henry and Nevo, 2014). The necessity for science to study crop species and plants in general under more natural field-like fluctuating and dynamic environments, to understand the regulation of genes contributing to outdoor phenotypes, has been described extensively (Fiorani and Schurr, 2013; Poorter *et al.*, 2016; Matsubara, 2018; Kaiser *et al.*, 2018; Slattery *et al.*, 2018; Townsend *et al.*, 2018). High-throughput phenotyping technologies have been used to dissect genetic variation in response to treatments changing one factor at a time or to phenotype natural population panels under highly controlled and stable environments (Chen *et al.*, 2014; Neumann *et al.*, 2015; Junker *et al.*, 2015; Muraya *et al.*, 2017; Knoch *et al.*, 2019). Nowadays, to study crops under natural conditions and under expected altered climate conditions due to climate change, high-throughput phenotyping platforms need to be able to simulate changing environments and create field-relevant climate scenarios to enable fast breeding programs to ensure global food security (Araus and Cairns, 2014; Fahlgren *et al.*, 2015). To tackle those challenges the IPK Plant Cultivation Hall (PKH) was developed to study crops under field-relevant conditions in replicable dynamic environments.

4.7 Plant Cultivation Hall – indoor field simulation

In this study the PKH was challenged in a benchmark experiment to assess how growth conditions in the PKH and in the GWH would compare to reference phenotypic data from three consecutive field cultivation experiments on 11 *Zea mays* inbreds. This small population originated from three

continents and covered a diverse set of genotypic variation, which resulted in a wide range of phenotypic response to the three field cultivation scenarios. The three cultivation years in the field were different in their exposure to local weather during the growth period of maize in Germany from April to September. The year 2018 was the hottest and driest year of the three, while 2017 was the wettest year with more than doubled rainfall in comparison to 2018. The year 2016 was a moderate year relative to the other two. While the temperature between April and September was different, the average temperatures of the full three years were similar between the years. Over the whole year of 2016 the IPK received 366 L m^{-2}, 506 L m^{-2} in the year 2017, and only 263 L m^{-2} in the by far driest year 2018. Global solar radiation was equal between 2016 and 2017 but about 10% higher in the 2018. Thus, the three diverse years were covering a wide range of the possible environmental variation at the IPK field sites. This was reflected by the average of the BLUEs of the final plant height. In 2016, plant height was highest, while in the other two years it was significantly reduced, since the lower temperature and the rainy weather in 2017 and the prolonged drought and high temperature in 2018 had a negative effect on final plant height.

The benchmark experiment in the PKH was the first large-scale experiment in the system, which was run with a complex climate regime. The PKH was designed to deliver light intensities and light qualities comparable to natural sunlight. To map the climate conditions from the three field experiments to the phenotyping experiment in the PKH, the median hourly weather data from 2016 to 2018 were calculated. This captured the most common weather characteristics of a certain time point at the IPK field sites but also buffered extreme values from any individual year. The control software of the PKH allows for complex programming, but it is not yet capable of directly translating the environmental information from weather data into a usable template to precisely reproduce the input weather in an automatic fashion. In an approximation, characteristic weather for each week was summarized in three exemplary days, based on their maximal daily temperature, and randomly distributed over the course of each week. Light intensities of the experiment (Figure S 2_2) reached field-relevant levels of up to 1,500 µmol m^{-2} s^{-1} which was in the range of other studies measuring light intensities in the field (Kulheim *et al.*, 2002; Slattery *et al.*, 2018; Townsend *et al.*, 2018). Thus, the illumination of the PKH was not run at maximum for the first benchmark experiment. On "cloudy" days, light intensity was lowered by two thirds to

achieve a differential which was also found in natural environments (Kulheim *et al.*, 2002; Slattery *et al.*, 2018).

The best linear unbiased predictors (BLUPs) estimated the random effects of the linear mixed model, which was applied on the BLUEs of the plant height at each measurement day in every cultivation scenario. Relative to the PKH and the GWH the effect of the field sites was negative in regard to development of plant height over the course of the experiments (Figure 2_5). The BLUPs of each year coincided with the weather observation, and the wettest year 2017 had the worst effect on plant height, followed by the hot and dry year 2018. The moderate year had the least negative estimated effect on development of plant height relative to the PKH and GWH. The BLUPs of the PKH and the GWH were both positive, however, the effect of the PKH was smaller than that of the GWH. Instead of only comparing the final plant heights, the modeling of BLUPs estimates the effect of each cultivation scenario in the context of the whole experiment. Thus, the smaller effect of the PKH can be explained by the early phenotypes. Plants in the PKH experienced a lag phase of (slow) growth similar to the plants grown in the field where the effect was even more pronounced due to the exposure to harsher weather conditions after transplanting (Figure 2_5). This effect was absent in the GWH and plants continuously grew until they reached maturity. The linear mixed model included those differences in the estimation of the random effects, which thereby explains the smaller effect of the PKH in comparison with the GWH. Plant height in general was higher in the PKH than in the field cultivation, moreover ranking of genotypes was different. A possible factor to explain this phenomenon could be the applied wind simulation, the light spectra of the different artificial light sources, and the daily temperature variation in the PKH.

While wind movement can help plants to adjust their stature, and in maize increases accumulation in strengthening sclerenchyma (Gardiner *et al.*, 2016), wind drag can alter phenotypes in a process called thigmomorphogenesis (Börnke and Rocksch, 2018). Especially early maize seedlings respond with reduced growth speeds when exposed to wind stimulus (Jaffe *et al.*, 1985), a process which was also observed in the PKH and in the fields, but not in the GWH where no wind was applied (Figure 2_7). The PKH wind simulation alternated between opposing directions to ensure a homogeneous mechanical stimulus. In contrast, at the IPK field sites the wind directions were also alternating between two opposing directions (Figure S 2_1), but often remained stable for several days. The plants in the PKH experienced changing wind drag every two hours, while the plants in

the fields experienced wind from one direction for a much longer period, enabling the plant to respond with adaptation processes. This effect could explain the observed genotypic differences regarding final plant height and the phenotypes in general. Plants in the PKH were more similar to the plants grown in the field than to the plants in the GWH in the early stages, when plants are most sensitive to wind drag (Figure 2_5).

Final plant heights were higher in both the PKH and the GWH than in the three field cultivations, which can partially be explained by higher temperatures in both systems (Figure 2_7). The average plant height between the PKH and GWH was not significantly different. The stable environments in the GWH exposed the plants to the highest temperature regime of all experiments. The climate simulation in the PKH was programmed to deliver the hourly median of the three field cultivation years, which effectively buffered the daily extreme values. Maize in particular is susceptible to low temperatures and developmental rates approach zero at temperatures lower than 10°C (Parent and Tardieu, 2012). The downside of this buffer was the overall increased temperature, and the avoidance of lower temperature extremes in the PKH leads to continuous growth and thereby resulted in taller plants than in the field. Another factor was the irrigation system. Plants in the PKH were watered weekly with 15 L m^{-2} water which was as much as in the wettest year 2017, normalized to the number of days. In the GWH irrigation was applied daily. Taken together, higher overall temperatures and greater availability of water could result in more vigorous growing, thus taller plants. Nonetheless, those two factors cannot solely account for the biomass increase, as final dry weight in the PKH was higher than dry weight in the GWH, while plant height was the same. The wind simulation in the PKH was run until the end of the experiment, and plants were developing bushy phenotypes, which could explain the greater accumulation of biomass while maintaining a similar plant height.

A very important parameter for a phototroph organism is the availability of light, and light quality is known to strongly influence biomass accumulation (Hogewoning *et al.*, 2010). It was shown that temperature responses in crops like maize are highly conserved and have not been affected by breeding processes nor by intraspecific evolution (Parent and Tardieu, 2012). To distinguish the effect of light quality and light intensity from the influence of temperature, the hourly temperature exposure was converted into species specific thermal time days (Parent and Tardieu, 2012). Thus the effect of the temperature was normalized to thermal time days and made

comparable between the experiments. Differences after the same number of thermal time days were then only explained by other factors like light quality and intensity, water availability, and wind stimulus. In the PKH and in the field cultivations maize plants required fewer temperature units to reach final plant height than in the GWH, suggesting that the above-mentioned or other environmental factors are positively influencing developmental speed. The differences in light intensities and/or light qualities could also explain the greater biomass accumulation in the PKH than in GWH. Furthermore, the shape of the thermal time normalized curves indicate lower growth rates in the early phases in the field and also in the PKH than in comparison with the GWH (Figure 2_7), which suggests that the climate parameters apart from temperature demanded a longer acclimation time from the plants.

An often neglected factor is the available soil composition when cultivating crops. In the PKH substrate layers were stacked in a large container system to produce field-relevant soil volume and soil type composition, which can have a substantial effect on predicted yield (Poorter *et al.*, 2012; Folberth *et al.*, 2016; Hohmann *et al.*, 2016). The small pots used in the GWH had only one type of soil and volume of 5.5 L was much lower than in the PKH. The different soil compositions and volumes could also explain higher biomass accumulation in the PKH than in the GWH, but were partially compensated by daily watering and weekly fertilization in the GWH.

Maize plants in all five experiments were grown until maturity and final seed yield was determined. The BLUEs of the seed traits displayed lower kernel weight and count for both cultivation scenarios in the indoor systems (PKH, and GWH). The yield of all three field experiments was higher for both aforementioned traits. The calculated thousand kernel weight of the indoor system was lower than of the moderate year 2016 and higher than of the wet year 2017 but similar to the hottest year 2018 (Table 2_3). The repeatability was high for every yield-related trait in every cultivation scenario, suggesting tight genotypic control and little environmental influence within a particular scenario. A potential explanation for the overall lower yield in the indoor systems could be the number of genotypes grown in parallel during the experiments. The full set of 30 maize genotypes was cultivated during the field experiments, while only 11 inbreds were grown in parallel in the indoor systems due to space limitations. The phenotyped plants, which were later used for yield determination, were open pollinated by themselves or by neighboring genotypes. Moreover the field plants were protected by a

commercially available maize hybrid, whose pollen was also contributing to the available pollen pool. Thereby more plants were shedding pollen simultaneously and at overlapping time points in the field than indoors. The earliest flowering genotype, ZEA 332, was flowering before any other maize genotype and could only be self-pollinated. ZEA 332 yielded comparable kernel numbers for all cultivations (Figure S 2_8). Other genotypes like ZEA 399 or ZEA 132, which were flowering later in the growth period, were yielding more kernels in field scenarios than in either of the indoor systems (Figure S 2_8). The increase in yield, likely due to more kernels, could be partially explained by a higher availability of fertile pollen during fertilization. A genotypic effect of different paternal genotypes to the pollen as contributing factors to yield in this generation can be neglected, except reduced pollen fertility, as kernel sink capacity and mature kernel mass is under strong maternal influence (Jones *et al.*, 1996).

In the current experiment, measurements were performed manually, as the imaging system is not available yet, and thus detailed high-throughput phenotyping was not possible. Traits like chlorophyll fluorescence, architectural traits, and color-related traits, which were recorded in the fluctuating light experiments for the Arabidopsis panel, and used to characterize the response of dynamic fluctuations, were not measured. Nonetheless, plant height, developmental stage, and final biomass accumulation and yield were suitable tools to characterize the *Zea mays* population and distinguish the performances from the different cultivation systems.

The first benchmark experiment was successful in its purpose to evaluate the performance of the PKH in relation to conventional greenhouses and field cultivations. But before the PKH can be fully used to mimic field scenarios, further questions need to be addressed, like whether the PKH can precisely replicate identical experiments and whether this results in reproduced phenotypes. An exact climate scenario from a specific year should be simulated in the PKH to evaluate the ability of the system to mimic certain field-like environmental scenarios. The PKH can account for most of the important environmental stimuli to which a plant is exposed to in the field, maybe with the exception of rain or biotic interaction, both positive and negative. However, so much possible variation and combinations requires careful and detailed preplanning of climate regimes and a close meshed supervision and maintenance of the systems to achieve the targeted performance. The planned upgrades of the system like the automatic phenotyping platform, the PhenoCrane, which mounts cameras of different spectral range, mounted to a freely movable crane system on

the x, y, and z axis developed by Photon Systems Instruments, will enable high-throughput phenotyping under field-relevant but controlled environments. The dissection of genetic traits, which are responsible for adaptation processes of crops to dynamic environments, will be possible. The PKH will be a valuable tool to elucidate mechanisms underlying responses to dynamic environments, especially by simulating future climate scenarios, and to characterize natural and mutant populations. Its rating to operate with genetically modified systems will even enable the performance estimation of transgenic plant varieties under field-relevant conditions, which enables the evaluation of the performance of genetically modified populations under favorable and unfavorable field-like conditions.

4.8 Rapid mapping of induced recessive and dominant mutations in maize

High-throughput phenotyping of mutated crop populations under dynamic environments can help to detect climate-responsive alleles with a positive effect on plant vigor, whose potential was untapped in the breeding pools due to a lack of genetic variation in many elite breeding programs (Mba, 2013). More than 3,200 elite crop varieties, as of the year 2012, were successfully bred by induced mutation programs (Mba *et al.*, 2012). Until now, gene discovery in EMS-induced mutant populations relied on mapping procedures, which required prior backcrossing or the analysis of large bulks of F_2 individuals (Heuermann *et al.*, 2019).

THE FOLLOWING TEXT IS CITED FROM MY PUBLICATION (Heuermann *et al.*, 2019)

"Here we devised and validated a strategy suitable for rapid detection of induced mutations closely linked to, and even causal to, phenotypically defined mutant loci. This strategy relies on the enormous capacity of the next-generation sequence technology and the power of computational sequence analyses making high-coverage WGS sequencing of multiple individuals affordable and comprehensive mutation detection and analysis feasible even for species with large genomes such as maize. The second decisive feature of the strategy is the outstanding discriminative power of Mendelian genetic segregation, which provides the means of filtering thousands of sequence variants for the very few (induced) mutation that are in the immediate vicinity of the phenotypically defined mutant locus and even within the responsible gene.

The success of this approach is dependent on covering a high fraction of the genome by reaching a sufficient sequencing depth for reliable SNP calling and genotyping of the mutants and their corresponding wild-type siblings. EMS mutagenesis, a process which is strongly influenced by the genotype and the climate conditions of the day of mutagenesis (Neuffer *et al.*, 2009), mutates base pairs in a random fashion, so that one can expect the mutations to occur evenly distributed over the chromosome. In the segregating *dwarf* and *pale green* M_2 populations the mutational load (1,357 and 3,536 induced mutations) was slightly lower compared to previous EMS mutagenesis in maize (Till *et al.*, 2004), which can be explained by lower concentrated EMS used in our mutagenesis. These few thousand sequence variants per mutated genome/per M_2 family are distributed among the 10 maize chromosomes covering approximately 2.3 Gbp of DNA sequence, which corresponds to *c.* 2,200 cM of genetic distance (Su *et al.*, 2017). The maximal achievable genetic resolution for each mutant population – 16 individuals derived from 32 meiotic events – is *c*. (1/32) 3 cM, which corresponds to unresolved regions of *c*. 6 cM around the causal mutation assuming identical segregation for all sequence variants. In relation to the 2,200 cM of the genome one would expect four and 10 SNPs for the *dwarf* and *pale green* population to be in the same linkage block, respectively. The numbers of SNPs which passed our zygosity filter were slightly lower than expected which can be explained by the fact that both mutant loci are located on the long arms of their chromosomes, in regions of relatively high recombination frequency (Liu *et al.*, 2015).

The low number/density of EMS-induced mutations in comparison to the occurrence of sequence polymorphism in classical linkage mapping of phenotypically defined F_2 populations, derived from a cross of a mutant with a genetically divergent wild-type, has a strong advantage as in the latter approach one is generally confronted with a lot more polymorphisms (by at least two to three orders of magnitude dependent on the species), which will not be linked to the causal mutation. Furthermore, mapping carried out within the same genotypic background eliminates influences of QTL potentially present in mapping populations derived from genetically divergent lines, which may be particularly relevant for the analysis of complex traits.

A further advantage, which should not be overlooked, is saving time. The rapid mapping strategy does neither require crosses nor the creation of mapping populations, which can save years in species with longer generation times as is the case with most crop plants. Any segregating family

derived from a selfed heterozygous individual, like an M_2 family, is sufficient, and only the sequenced individuals need to be subjected to progeny analysis.

The time-limiting step in this analysis is the phenotyping of the M_3 generation. Early phenotypes, like those caused by the *dwarf* or *pale green* mutation, which were unambiguously discriminable one week after germination, have no additional impact on the time of the experiment. Late phenotypes like flowering time or yield would scale the time of the experiment with the generation time of a species.

Hitherto, we expected that the majority of the mutant phenotypes in our *dwarf* and *pale green* populations were caused by transition mutations as a direct result of the mutagen EMS. The *pale green* mutation was identified as a transition SNP in the *white seedling 2* (Zm00001d026402) gene leading to a non-synonymous amino acid exchange. This gene encodes a chloroplast DNA polymerase, and, if mutated, the plant expresses a pale green phenotype (Udy *et al.*, 2012). Although other alleles of the *w2* locus are described to segregate recessively (Andorf *et al.*, 2016), the *pale green* mutation is inherited semi-dominantly. The mutant allele is semi-dominant over the wild-type allele and thus makes the wild-type *w2* gene haploinsufficient to produce a wild-type phenotype in a heterozygous state. The mutation is also affected by age-related penetrance of the phenotype as in later developmental stage the pale green color diminishes, probably because the chloroplast biogenesis catches up, which could explain a slight discrepancy in the expected Mendelian segregation.

The mutation was observed in the M_1 generation, where only dominant mutations would produce a phenotype. Despite the offspring from *pale green* mutants suffering from a low germination rate, every *pale green* M_2 individual produced segregating mutant offspring with pale green mutant phenotypes, while *PALE GREEN* M_2 individuals did not. In a diploid organism the zygosity of an allele can only be homozygous or heterozygous, therefore the *pale green* mutation had to be inherited dominantly as supported by the four SNPs in Chromosome 10 passing the stringent zygosity filter. The non-synonymous amino acid exchange was predicted to have a deleterious effect on the protein function as it changes the DNA polymerase A domain of the protein, which we could demonstrate by showing that the chloroplast DNA levels were strongly reduced in the *pale green* mutants (Figure 3_3). Furthermore, the pale green phenotype was visible with the

naked eye and reduced photosystem II operating efficiency was measured (Figure 3_1 b, c), indicating a lower amount of chlorophylls likely because of an impaired chloroplast biogenesis. The other three SNPs which also passed the required zygosity filter were probably in linkage. The *pale green* mutant population supported our expectation and allowed the direct identification of an EMS-induced mutation.

The *dwarf* mutation, conversely, was not caused by a point mutation. The mutation likely occurred spontaneously during, but not necessarily because of, the mutagenesis or during seed amplification immediately before. The dwarf phenotype caused by the *116GA-an1-93W1189* allele was also found to derive from a spontaneous mutation which arose in the Maize Genetics Cooperation Stock Center stock (Andorf *et al.*, 2016). We identified publicly available mutants of the *anther ear 1* gene (Zm00001d032961) to be fully allelic to our *dwarf* mutation and during re-sequencing identified an insertion of a DNA transposon in the largest exon which likely disrupts the function of the gene and leads to the *dwarf* phenotype, which is in concordance with other studies (Bensen *et al.*, 1995; Landoni *et al.*, 2007). *Anther ear 1* encodes an *ent*-kaurenoic acid synthase that acts as an entry point enzyme in the pathway of bioactive gibberellins and if mutated leads to a semi-dwarf structure (Katsumi *et al.*, 1964; Bensen *et al.*, 1995; Landoni *et al.*, 2007; Nelissen *et al.*, 2012). Moreover, the anther ear phenotype has already been known since the early 20th century (Emerson and Emerson, 1922) and was also expressed in our *dwarf* mutants (Figure S1). We concluded that the SNPs in the dwarf population, which passed the zygosity filter, were in linkage with the insertion in the *an1* gene rather than being causal. Although this was not expected prior to the experiment, it demonstrates the power of the rapid mapping approach, which not only identifies causative SNPs, but also discovers linked mutations using the SNPs solely as markers. This is a strong advantage of sequencing individual plants, as knowledge of the exact allelic state in each sequenced individual allows for unambiguous identification SNPs in linkage with the causal mutation.

A classical approach to identify SNPs in linkage with the causal mutation is bulked segregant analysis (BSA). The power of discrimination of closely linked/causal SNPs from unlinked SNPs and the achieved mapping resolution in BSA depends on the size of the analyzed bulks and the sequencing depth. In a bulk of 16 homozygous mutant plants (equivalent to the number of plants analyzed in this study) an unlinked SNP will show an allele frequency of 1 (occurrence of only one

allele) at a theoretical likelihood of (0.25^{16}) 2.33e-10 provided non-limiting read coverage. Erroneous detection of unlinked SNPs will therefore be rare. The achievable mapping resolution, that is the range covered by linked SNPs, which cannot be resolved with a bulk size of 16 plants (32 meiotic events) will be around 9 cM at an (alpha) error threshold of 5% ($1 - \sqrt[32]{0.05} = 0.09$). In experimental studies the observed mapping resolution of sequenced bulks was close to the theoretical possible resolution. Abe *et al.* (2012) found mapping intervals of 2 Mb in pools of only 20 mutant rice plants, a resolution of 6.66 cM, which is very close to the expected 7 cM ($1 - \sqrt[40]{0.05} = 0.07$). Similar high precision was found in a bulked maize RNA-Seq analysis by Liu *et al.* (2012), who achieved a mapping resolution of 2 Mb (*c*. 2 cM) matching the expected 2 cM ($1 - \sqrt[128]{0.05} = 0.02$) resolution. In Arabidopsis, Lindner *et al.* (2012) realized a mapping resolution of 4.7 cM with pools of 53 plants compared to expected 3 cM ($1 - \sqrt[106]{0.05} = 0.03$). Even though increasing pool sizes tend to lead to higher mapping resolution this parameter is strongly dependent on coverage of the NGS data. If the coverage is considerably lower than the pool size, allele frequencies cannot be resolved accurately enough and resolution drops, as seen by Hartwig *et al.* (2012) who achieved around 7 cM with pools of 318 plants despite an theoretically achievable resolution of 0.5 cM.

For zygosity mapping, the likelihood of unlinked SNP to be considered as causative is also (0.25^{16}) 2.33e-10. However, in contrast to the bulk segregant analysis, here the zygosity for each individual SNP position can be assessed. Therefore, only SNPs in perfect linkage exactly follow the same zygosity pattern as the causal SNP whereas any SNP outside the strong linkage range will deviate from the exact pattern and therefore will not pass the zygosity filter (Table 3_3). The SNPs in full linkage are indistinguishable from each other and therefore the likelihood is the same for all SNPs sharing the same linkage block. Provided sufficient sequencing depth, the achievable resolution of the rapid mapping approach is therefore only determined by the number of meiotic events and the SNP density.

To reach a similar resolution by BSA a larger number of plants would be required and subsequently, a sufficient sequencing depth needs to be achieved, which depends on the number of individuals in the pool, to unequivocally identify SNPs linked with the mutation. To achieve a mapping resolution of 3 cM with BSA in a diploid species and an alpha threshold of 5%, a pool of

50 individuals, 100 meiotic events ($1 - \sqrt[100]{0.05} = 0.03$), and a minimal read coverage of 100x would be required, to sequence every allele in the pool once, assuming an ideal distribution of alleles in the pools. As an ideal distribution is experimentally challenging, read coverage of 2-3x per allele would be necessary to assess the real allele frequency in such pools. However, in most BSA experiments the coverage is lower than the pool of individual alleles and therefore the assessment of the real allele frequency on whole-genome sequencing data alone was rarely possible (Hartwig *et al.*, 2012). In contrast, 16 individuals with 20x sequencing coverage each (16* 20x = 320x) are on average needed for a similar resolution in the zygosity filter approach. In absolute numbers the zygosity filter procedure is two- to three-fold more efficient than a classical BSA.

One of the strongest arguments for zygosity mapping is the fact that mutations can be mapped in a heterozygous state without the need of prior backcrossing, which is not possible with BSA. The *pale green* mutation mapped in this study would not be detectable in sequenced pools of M_2 mutants as its SNP frequency of 0.5 would be indistinguishable from that of unlinked SNPs segregating in the population. Although it has been shown that heterozygous mutations can be generally mapped in pools of plants (Lindner *et al.*, 2012), they require several rounds of prior backcrossing, which can be skipped performing a zygosity filter on sequenced individuals.

The only drawback of individually sequencing plants instead of performing bulk segregant analysis, is the increase in cost due to the necessity to reach the coverage 16 times per population instead of only once following the SHORE map approach (Schneeberger *et al.*, 2009). However, working with maize rather than with Arabidopsis the coincidence of a drastically longer generation time, limited greenhouse space, and most importantly the gain in time, justifies our rapid mapping technique. Furthermore, the steady decrease in sequencing cost according to (Wetterstrand, 2017) will likely continue in the coming years, making our rapid mapping approach much more feasible for a wide range of plant species and even with higher level of ploidy like wheat or oilseed rape. The only requirement for a successful adaptation for the rapid mapping approach is an available high quality reference sequence, which does not necessarily have to be from the same genotype. We compared sequence variation in 10 Mb regions around the two candidate mutations between B73_AGPv3 and PH207 and found no major structural rearrangements (Figure S 3_5). Even though we performed SNPs calling of the individually sequenced plants in the PH207

background against the reference sequence of B73, as the genome of PH207 was not published when the work in this study was performed, which lead to a much larger number of SNPs, mainly not EMS specific, filtering for only segregating SNPs was enough to enable the zygosity filter to precisely identify the causative or linked SNPs."[1]

END OF CITATION (Heuermann *et al.*, 2019)

SUMMARY

Plants evolved under fluctuating light and dynamic environmental conditions, yet plant research was so far mainly focused on gene discovery under continuous, controlled growth conditions. The main part of this work explored Arabidopsis' response to fluctuating light and, in two supporting methodical advancements, plant performance of maize under simulated field-like environments was evaluated in the Plant Cultivation Hall and a rapid mapping procedure was developed to identify EMS-induced mutations in maize.

A panel of 384 natural *Arabidopsis thaliana* accessions was analyzed in two independent experiments under fluctuating and constant light conditions in parallel in a high-throughput phenotyping system for small plants, upgraded with supplemental LEDs. Automated image analyses at six time points and kinetic chlorophyll fluorescence measurements at one time point were used to extract 286 architectural- and/or color-related, and 78 photosynthesis-related traits. Genome-wide association studies were conducted for all phenotypic traits and all time points with 214,051 single nucleotide polymorphisms (SNPs) using the R package *FarmCPU*. Genes in full linkage disequilibrium with fluctuating light specific marker-trait associations with a repeatability value > 0.5 were filtered for gene ontology terms containing "photo" or "light", which finally resulted in ten candidate genes from eleven marker-trait associations. Three of those candidate genes, the *PIL6* transcription factor together with the *PHYC* and the *JAR1* gene, act together in phytochrome signaling in response to far red light or shade avoidance suggesting that Arabidopsis's response to fluctuating light may be mediated by the same pathways and signaling cascades.

The Plant Cultivation Hall at the IPK was designed and built to mimic field-relevant environments in a controlled and repeatable way. As a first benchmark experiment plant performance of eleven diverse *Zea mays* inbreds in the Plant Cultivation Hall was compared to those of lines grown in the greenhouse and for three consecutive seasons in the field. Hourly recorded temperature data for the three growing seasons served as template for the climate conditions in the Plant Cultivation Hall and were combined with a complex weekly regime regarding light intensity, humidity, wind speed and direction, and adjusted day length. Overall, the chosen climate conditions in the Plant Cultivation Hall proved to be superior to those in the greenhouse to mimic performance of field grown plants, especially in the early growth phase, however, individual parameters, like wind simulation need further refinement.

In a rapid mapping approach, whole genome shotgun sequencing was performed on 32 individual ethyl methansulfonate mutagenized *Zea mays* plants from two M_2 populations with 16 plants each, which both segregated for plants with and without a mutant phenotype. The state of zygosity of the mutation was determined for each sequenced M_2 individual by analyzing the segregation pattern of the mutant phenotype in up to 40 of their respective M_3 offspring plants. The knowledge about the precise state of zygosity allowed filtering for SNPs in either population, which strictly followed the determined state of zygosity of a mutation in each individual member. Three SNPs were found in linkage with the homozygous *dwarf* mutation, which was likely caused by an insertion in the large exon of the *an1* gene. Four SNPs were in linkage with the heterozygous *pale green* mutation, of which one was non-synonymously changing an amino acid in the *w2* gene.

ZUSAMMENFASSUNG

Pflanzen evolvierten unter schwankenden Licht- und dynamischen Umweltbedingungen, während sich die Pflanzenforschung bislang hauptsächlich auf die Entdeckung von Genen unter kontinuierlichen, kontrollierten Wachstumsbedingungen konzentrierte. Der Hauptteil dieser Arbeit untersuchte die Reaktion von Arabidopsis auf schwankendes Licht und, in zwei unterstützenden methodischen Weiterentwicklungen, bewertete die pflanzliche Leistung von Mais unter simulierten feldähnlichen Bedingungen in der Pflanzenkulturhalle und entwickelte ein schnelles Kartierungsverfahren zur Identifizierung von EMS-induzierten Mutationen in Mais.

Ein Panel von 384 natürlichen *Arabidopsis thaliana* Akzessionen wurde in zwei unabhängigen Experimenten parallel unter wechselnden und konstanten Lichtverhältnissen in einem mit zusätzlichen LEDs aufgerüsteten Hochdurchsatzphänotypisierungssystem für kleine Pflanzen analysiert. Automatisierte Bildanalyse zu sechs Zeitpunkten und kinetische Chlorophyll-Fluoreszenzmessungen zu einem Zeitpunkt wurden verwendet, um 286 architektonische und/oder farbbezogene bzw. 78 Photosynthese-bezogene Merkmale zu extrahieren. Genomweite Assoziationsstudien wurden für alle phänotypischen Merkmale und alle Zeitpunkte mit 214.051 Einzelnukleotid-Polymorphismen (SNPs) mit dem R-Paket *FarmCPU* durchgeführt. Gene im Kopplungsungleichgewicht mit Marker-Merkmalassoziationen spezifisch für schwankendes Licht mit einem Wiederholungswert > 0,5 wurden hinsichtlich der genontologischen Begriffe "Foto" oder "Licht" gefiltert, was schließlich zu zehn Kandidatengenen aus elf Marker-Merkmalassoziationen führte. Drei dieser Kandidatengene, der *PIL6*-Transkriptionsfaktor zusammen mit dem *PHYC*- und dem *JAR1*-Gen, wirken zusammen im Phytochrom Signalweg als Antwort auf dunkelrotes Licht oder Schattenvermeidung, was darauf hindeutet, dass Arabidopsis als Reaktion auf schwankendes Licht die gleichen Wege und Signalkaskaden nutzt.

Die Pflanzenkulturhalle am IPK wurde so konzipiert und gebaut, dass sie feldrelevante Umgebungen kontrolliert und wiederholbar nachbildet. Als erster Benchmark-Versuch wurde die Pflanzenleistung von elf verschiedenen *Zea mays* Inzuchtlinien in der Pflanzenkulturhalle mit der von Linien im Gewächshaus und drei aufeinander folgenden Jahren auf dem Feld verglichen. Stündlich aufgezeichnete Temperaturdaten für die drei Vegetationsperioden dienten als Vorlage für die Klimabedingungen in der Pflanzenkulturhalle und wurden hinsichtlich Lichtintensität, Luftfeuchtigkeit, Windgeschwindigkeit und -richtung sowie angepasster Tageslänge in einem komplexen wöchentlichen Ablauf kombiniert. Insgesamt erwiesen sich die gewählten Klimabedingungen in der Pflanzenkulturhalle für den Pflanzenanbau als besser als im Gewächshaus, um die Leistung von Feldpflanzen abzubilden, besonders in frühen Wachstumsphasen, obwohl einzelne Parameter, wie die Windsimulation, weiter verfeinert werden müssen.

In einem schnellen Kartierungsansatz wurde die Sequenzierung des gesamten Genoms von 32 einzelnen ethylmethansulfonatmutagenisierten *Zea mays* Pflanzen aus zwei M_2-Populationen mit je 16 Pflanzen durchgeführt, die beide in Pflanzen mit und ohne mutierten Phänotyp aufspalteten. Der Zustand der Zygotie der Mutation wurde für jedes sequenzierte M_2 Individuum bestimmt, indem das Segregationsmuster des mutierten Phänotyps in bis zu 40 ihrer jeweiligen M_3-Nachkommenpflanzen analysiert wurde. Das Wissen über den genauen Zustand der Zygotie erlaubte es, SNPs in beiden Populationen so zu filtern, dass sie dem erwarteten Zygotiemuster in jeder einzelnen Pflanze folgten. Drei SNPs wurden in Verknüpfung mit der homozygoten *dwarf* Mutation gefunden, welche wahrscheinlich durch eine Insertion im größten Exon des *an1*-Gens verursacht wurde. Vier SNPs waren mit der heterozygoten *pale green* Mutation verknüpft, von denen ein SNP eine Aminosäure im *w2*-Gen nicht-synonym veränderte.

REFERENCES

Abe, A., Kosugi, S., Yoshida, Kentaro, et al. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nature Biotechnology*, **30**, 174–178.

Abendroth, L.J., Elmore, R., Boyer, M. and Marlay, S. (2011) *Corn growth and development*, Iowa State University, University Extension. Available at: https://store.extension.iastate.edu/Product/6065.

Adamczyk, B.J., Lehti-Shiu, M.D. and Fernandez, D.E. (2007) The MADS domain factors AGL15 and AGL18 act redundantly as repressors of the floral transition in Arabidopsis. *The Plant Journal*, **50**, 1007–1019.

Addo-Quaye, C., Buescher, E., Best, N., Chaikam, V., Baxter, I. and Dilkes, B.P. (2017) Forward Genetics by Sequencing EMS Variation-Induced Inbred Lines. *G3*, **7**, 413–425.

Alexa, A. and Rahnenfuhrer, J. (2016) *topGO: Enrichment Analysis for Gene Ontology.*, R package version 2.26.0.

Allahverdiyeva, Y., Suorsa, M., Tikkanen, M. and Aro, E.-M. (2015) Photoprotection of photosystems in fluctuating light intensities. *Journal of Experimental Botany*, **66**, 2427–2436.

Alonso-Blanco, C., Andrade, J., Becker, C., et al. (2016) 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell*, **166**, 481–491.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Andorf, C.M., Cannon, E.K., Portwood, J.L., et al. (2016) MaizeGDB update: new tools, data and interface for the maize model organism database. *Nucleic Acids Research*, **44**, D1195–D1201.

Annunziata, M.G., Apelt, F., Carillo, P., Krause, U., Feil, R., Koehl, K., Lunn, J.E. and Stitt, M. (2018) Response of Arabidopsis primary metabolism and circadian clock to low night temperature in a natural light environment. *Journal of Experimental Botany*, **69**, 4881–4895.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.

Araus, J.L. and Cairns, J.E. (2014) Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.*, **19**, 52–61.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., et al. (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, **465**, 627–631.

Austin, R.S., Vidaurre, D., Stamatiou, G., et al. (2011) Next-generation mapping of Arabidopsis genes. *The Plant Journal*, **67**, 715–725.

Baker, N.R. (2008) Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.*, **59**, 89–113.

Barr, D.J., Levy, R., Scheepers, C. and Tily, H.J. (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**.

Barry, G.F. (2001) The use of the Monsanto draft rice genome sequence in research. *Plant Physiol.*, **125**, 1164–1165.

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015) Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, **67**.

Becker, R.A., Wilks, A.R., Brownrigg, R., Minka, T.P. and Deckmyn, A. (2018) *maps: Draw Geographical Maps*, Available at: https://CRAN.R-project.org/package=maps.

Bensen, R.J., Johal, G.S., Crane, V.C., Tossberg, J.T., Schnable, P.S., Meeley, R.B. and Briggs, S.P. (1995) Cloning and characterization of the maize *An1* gene. *Plant Cell*, **7**, 75–84.

Blumenstiel, J.P., Noll, A.C., Griffiths, J.A., Perera, A.G., Walton, K.N., Gilliland, W.D., Hawley, R.S. and Staehling-Hampton, K. (2009) Identification of EMS-induced mutations in *Drosophila melanogaster* by whole-genome sequencing. *Genetics*, **182**, 25–32.

Börnke, F. and Rocksch, T. (2018) Thigmomorphogenesis – Control of plant growth by mechanical stimulation. *Scientia Horticulturae*, **234**, 344–353.

Bricker, T.M., Roose, J.L., Zhang, P. and Frankel, L.K. (2013) The PsbP family of proteins. *Photosynthesis Research*, **116**, 235–250.

Burgos, A., Szymanski, J., Seiwert, B., Degenkolbe, T., Hannah, M.A., Giavalisco, P. and Willmitzer, L. (2011) Analysis of short-term changes in the *Arabidopsis thaliana* glycerolipidome in response to temperature and light. *The Plant Journal*, **66**, 656–668.

Bush, W.S. and Moore, J.H. (2012) Chapter 11: Genome-wide association studies. *PLoS Comput. Biol.*, **8**, e1002822.

Carlson, C.H., Gouker, F.E., Crowell, C.R., Evans, L., DiFazio, S.P., Smart, C.D. and Smart, L.B. (2019) Joint linkage and association mapping of complex traits in shrub willow (*Salix purpurea* L.). *Annals of Botany*, mcz047.

Carter, P.R. and Hudelson, K.D. (1988) Influence of Simulated Wind Lodging on Corn Growth and Grain Yield. *Journal of Production Agriculture*, **1**, 295.

Ceccarelli, S. (2015) Efficiency of Plant Breeding. *Crop Science*, **55**, 87.

Ceccarelli, S., Grando, S., Maatougui, M., et al. (2010) Plant breeding and climate changes. *The Journal of Agricultural Science*, **148**, 627–637.

Challinor, A.J., Koehler, A.-K., Ramirez-Villegas, J., Whitfield, S. and Das, B. (2016) Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nature Climate Change*, **6**, 954–958.

Chapuis, R., Delluc, C., Debeuf, R., Tardieu, F. and Welcker, C. (2012) Resiliences to water deficit in a phenotyping platform and in the field: How related are they in maize? *European Journal of Agronomy*, **42**, 59–67.

Char, S.N., Neelakandan, A.K., Nahampun, H., et al. (2017) An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnology Journal*, **15**, 257–268.

Chazdon, R.L. and Pearcy, R.W. (1991) The Importance of Sunflecks for Forest Understory Plants. *BioScience*, **41**, 760–766.

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T. and Klukas, C. (2014) Dissecting the Phenotypic Components of Crop Plant Growth and Drought Responses Based on High-Throughput Image Analysis. *Plant Cell*, **26**, 4636–4655.

Chen, H.-J., Chen, C.-L. and Hsieh, H.-L. (2015) Far-Red Light-Mediated Seedling Development in Arabidopsis Involves FAR-RED INSENSITIVE 219/JASMONATE RESISTANT 1-Dependent and -Independent Pathways. *PLoS ONE*, **10**, e0132723.

Chen, I.-C., Huang, I.-C., Liu, M.-J., Wang, Z.-G., Chung, S.-S. and Hsieh, H.-L. (2007) Glutathione S-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in Arabidopsis. *Plant Physiol.*, **143**, 1189–1202.

Chen, J., Hu, Q., Zhang, Y., Lu, C. and Kuang, H. (2014) P-MITE: a database for plant miniature inverted-repeat transposable elements. *Nucleic Acids Research*, **42**, D1176–D1181.

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**, 789–804.

Choi, Y. and Chan, A.P. (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, **31**, 2745–2747.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R. and Chan, A.P. (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE*, **7**, e46688.

Chory, J., Ecker, J.R., Briggs, S., et al. (2000) National Science Foundation-Sponsored Workshop Report: "The 2010 Project" functional genomics and the virtual plant. A blueprint for

understanding how plants are built and how to improve them. *Plant Physiol.*, **123**, 423–426.

Cooper, M. and DeLacy, I.H. (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theoret. Appl. Genetics*, **88**, 561–572.

Cordero, R.A. (1999) Ecophysiology of *Cecropia schreberiana* saplings in two wind regimes in an elfin cloud forest: growth, gas exchange, architecture and stem biomechanics. *Tree Physiol.*, **19**, 153–163.

Cross, H.Z. and Zuber, M.S. (1972) Prediction of Flowering Dates in Maize Based on Different Methods of Estimating Thermal Units. *Agronomy Journal*, **64**, 351.

Dai, L., Wu, L., Dong, Q., Zhang, Z., Wu, N., Song, Y., Lu, S. and Wang, P. (2017) Genome-wide association study of field grain drying rate after physiological maturity based on a resequencing approach in elite maize germplasm. *Euphytica*, **213**, 182.

Deikman, J., Petracek, M. and Heard, J.E. (2012) Drought tolerance through biotechnology: improving translation from the laboratory to farmers' fields. *Current Opinion in Biotechnology*, **23**, 243–250.

Desjardins, R.L., Sinclair, T.R. and Lemon, E.R. (1973) Light Fluctuations in Corn. *Agronomy Journal*, **65**, 904.

Ding, J., Ali, F., Chen, G., et al. (2015) Genome-wide association mapping reveals novel sources of resistance to northern corn leaf blight in maize. *BMC Plant Biol.*, **15**, 206.

DOE (2013) Solid-state lighting technology fact sheet: energy efficiency of LEDs. *U.S. Department of Engineering - Energy Efficiency and Renewable Energy, Building Technologies Office. Fact Sheet PNNL-SA-94206*. Available at: http://www.hi-led.eu/wp-content/themes/hiled/pdf/led_energy_efficiency.pdf.

Dornbusch, T., Michaud, O., Xenarios, I. and Fankhauser, C. (2014) Differentially phased leaf growth and movements in Arabidopsis depend on coordinated circadian and light regulation. *Plant Cell*, **26**, 3911–3921.

Emerson, R.A. and Emerson, S.H. (1922) Genetic Interrelations of Two Andromonoecious Types of Maize, Dwarf and Anther Ear. *Genetics*, **7**, 203–236.

Etten, J. van, Sousa, K. de, Aguilar, A., et al. (2019) Crop variety management for climate adaptation supported by citizen science. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 4194–4199.

Fahlgren, N., Gehan, M.A. and Baxter, I. (2015) Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*, **24**, 93–99.

FAO (2017) Food and Agricultural Organization of the United Nations. FAOSTAT Statistics Database, Food and Agricultural commodities production [cited 2 Nov 2017]. Available from: http://www.fao.org/faostat/en/#data/QC.

Fekih, R., Takagi, H., Tamiru, M., et al. (2013) MutMap+: genetic mapping and mutant identification without crossing in rice. *PLoS ONE*, **8**, e68529.

Fernandez, D.E., Wang, C.-T., Zheng, Y., Adamczyk, B.J., Singhal, R., Hall, P.K. and Perry, S.E. (2014) The MADS-Domain Factors AGAMOUS-LIKE15 and AGAMOUS-LIKE18, along with SHORT VEGETATIVE PHASE and AGAMOUS-LIKE24, Are Necessary to Block Floral Gene Expression during the Vegetative Phase. *Plant Physiol.*, **165**, 1591–1603.

Filiault, D.L. and Maloof, J.N. (2012) A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genet.*, **8**, e1002589.

Filo, J., Wu, A., Eliason, E., Richardson, T., Thines, B.C. and Harmon, F.G. (2015) Gibberellin driven growth in *elf3* mutants requires PIF4 and PIF5. *Plant Signal Behav*, **10**, e992707.

Fiorani, F. and Schurr, U. (2013) Future Scenarios for Plant Phenotyping. *Annu. Rev. Plant Biol.*, **64**, 267–291.

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L.B., Obersteiner, M. and Velde, M. van der (2016) Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nature Communications*, **7**, 11872.

Frenkel, M., Bellafiore, S., Rochaix, J.-D. and Jansson, S. (2006) Hierarchy amongst photosynthetic acclimation responses for plant fitness. *Physiologia Plantarum*, **129**, 455–459.

Friedman, W.E. (1999) Expression of the cell cycle in sperm of Arabidopsis: implications for understanding patterns of gametogenesis and fertilization in plants and other eukaryotes. *Development*, **126**, 1065–1075.

Gage, J.L., White, M.R., Edwards, J.W., Kaeppler, S. and Leon, N. de (2018) Selection Signatures Underlying Dramatic Male Inflorescence Transformation During Modern Hybrid Maize Breeding. *Genetics*, **210**, 1125–1138.

Gallavotti, A., Long, J.A., Stanfield, S., Yang, X., Jackson, D., Vollbrecht, E. and Schmidt, R.J. (2010) The control of axillary meristem fate in the maize ramosa pathway. *Development*, **137**, 2849–2856.

Gardiner, B., Berry, P. and Moulia, B. (2016) Review: Wind impacts on plant growth, mechanics and damage. *Plant Science*, **245**, 94–118.

Gerdes, J.T., Behr, C.F., Coors, J.G. and Tracy, W.F. (1993) *Compilation of North American Maize Breeding Germplasm*, Crop Science Society of America.

Goff, S.A., Ricke, D., Lan, T.-H., et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*, **296**, 92–100.

Gratani, L. (2014) Plant Phenotypic Plasticity in Response to Environmental Factors. *Advances in Botany*, **2014**, 1–17.

Harbinson, J., Prinzenberg, A.E., Kruijer, W. and Aarts, M.G. (2012) High throughput screening with chlorophyll fluorescence imaging and its use in crop improvement. *Current Opinion in Biotechnology*, **23**, 221–226.

Harshavardhan, V.T., Van Son, L., Seiler, C., et al. (2014) AtRD22 and AtUSPL1, Members of the Plant-Specific BURP Domain Family Involved in *Arabidopsis thaliana* Drought Tolerance. *PLoS ONE*, **9**, e110065.

Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R. and Weisshaar, B. (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Mol. Biol.*, **57**, 155–171.

Hartwig, B., James, G.V., Konrad, K., Schneeberger, K. and Turck, F. (2012) Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks. *Plant Physiol.*, **160**, 591–600.

Hatfield, J.L. and Prueger, J.H. (2015) Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes*, **10**, 4–10.

He, S., Schulthess, A.W., Mirdita, V., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., Reif, J.C. and Jiang, Y. (2016) Genomic selection in a commercial winter wheat population. *Theor. Appl. Genet.*, **129**, 641–651.

Henderson, C.R. (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics*, **31**, 423–447.

Henry, R.J. and Nevo, E. (2014) Exploring natural selection to guide breeding for agriculture. *Plant Biotechnology Journal*, **12**, 655–662.

Heuermann, M.C., Rosso, M.G., Mascher, M., Brandt, R., Tschiersch, H., Altschmied, L. and Altmann, T. (2019) Combining next-generation sequencing and progeny testing for rapid identification of induced recessive and dominant mutations in maize M_2 individuals. *The Plant Journal*, tpj.14431.

Hill, W.G. and Weir, B.S. (1988) Variances and covariances of squared linkage disequilibria in finite populations. *Theor Popul Biol*, **33**, 54–78.

Hogewoning, S.W., Douwstra, P., Trouwborst, G., Ieperen, W. van and Harbinson, J. (2010) An artificial solar spectrum substantially alters plant development compared with usual climate room irradiance spectra. *Journal of Experimental Botany*, **61**, 1267–1276.

Hohmann, M., Stahl, A., Rudloff, J., Wittkop, B. and Snowdon, R.J. (2016) Not a load of rubbish: simulated field trials in large-scale containers: Large-container yield phenotyping. *Plant, Cell & Environment*, **39**, 2064–2073.

Hornitschek, P., Kohnen, M.V., Lorrain, S., et al. (2012) Phytochrome interacting factors 4 and 5 control seedling growth in changing light conditions by directly controlling auxin signaling. *The Plant Journal*, **71**, 699–711.

Horton, M.W., Hancock, A.M., Huang, Y.S., et al. (2012) Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics*, **44**, 212–216.

Horton, P. (2004) Molecular design of the photosystem II light-harvesting antenna: photosynthesis and photoprotection. *Journal of Experimental Botany*, **56**, 365–373.

Hsieh, H.L., Okamoto, H., Wang, M., Ang, L.H., Matsui, M., Goodman, H. and Deng, X.W. (2000) *FIN219*, an auxin-regulated gene, defines a link between phytochrome A and the downstream regulator COP1 in light control of Arabidopsis development. *Genes Dev.*, **14**, 1958–1970.

Hu, W., Franklin, K.A., Sharrock, R.A., Jones, M.A., Harmer, S.L. and Lagarias, J.C. (2013) Unanticipated regulatory roles for Arabidopsis phytochromes revealed by null mutant analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 1542–1547.

Huang, W., Cai, Y.-F., Wang, J.-H. and Zhang, S.-B. (2018) Chloroplastic ATP synthase plays an important role in the regulation of proton motive force in fluctuating light. *Journal of Plant Physiology*, **226**, 40–47.

Ingvarsson, P.K. and Street, N.R. (2011) Association genetics of complex traits in plants: Tansley review. *New Phytologist*, **189**, 909–922.

Irvine, D.V., Goto, D.B., Vaughn, M.W., Nakaseko, Y., McCombie, W.R., Yanagida, M. and Martienssen, R. (2009) Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing. *Genome Res.*, **19**, 1077–1083.

Jaffe, M.J., Wakefield, A.H., Telewski, F., Gulley, E. and Biro, R. (1985) Computer-Assisted Image Analysis of Plant Growth, Thigmomorphogenesis, and Gravitropism. *Plant Physiol.*, **77**, 722–730.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. and Martin, C. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J.*, **19**, 6150–6161.

Jin, H., Liu, B., Luo, L., et al. (2014) HYPERSENSITIVE TO HIGH LIGHT1 interacts with LOW QUANTUM YIELD OF PHOTOSYSTEM II1 and functions in protection of photosystem II from photodamage in Arabidopsis. *Plant Cell*, **26**, 1213–1229.

Johansson Jänkänpää, H., Mishra, Y., Schröder, W.P. and Jansson, S. (2012) Metabolic profiling reveals metabolic shifts in Arabidopsis plants grown under different light conditions. *Plant, Cell & Environment*, **35**, 1824–1836.

Jones, H.G., Hutchinson, P.A., May, T., Jamali, H. and Deery, D.M. (2018) A practical method using a network of fixed infrared sensors for estimating crop canopy conductance and evaporation rate. *Biosystems Engineering*, **165**, 59–69.

Jones, R.J., Schreiber, B.M.N. and Roessler, J.A. (1996) Kernel Sink Capacity in Maize: Genotypic and Maternal Regulation. *Crop Science*, **36**, 301.

Junker, A., Muraya, M.M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A.E., Meyer, R.C., Riewe, D. and Altmann, T. (2015) Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Front Plant Sci*, **5**.

Jurić, S., Hazler-Pilepić, K., Tomasić, A., et al. (2009) Tethering of ferredoxin:NADP+ oxidoreductase to thylakoid membranes is mediated by novel chloroplast protein TROL. *The Plant Journal*, **60**, 783–794.

Kaiser, E., Morales, A. and Harbinson, J. (2018) Fluctuating Light Takes Crop Photosynthesis on a Rollercoaster Ride. *Plant Physiol.*, **176**, 977–989.

Katsumi, M., Phinney, B.O., Jefferies, P.R. and Henrick, C.A. (1964) Growth Response of the d-5 and an-1 Mutants of Maize to Some Kaurene Derivatives. *Science*, **144**, 849–850.

Khanna, R., Shen, Y., Marion, C.M., Tsuchisaka, A., Theologis, A., Schäfer, E. and Quail, P.H. (2007) The basic helix-loop-helix transcription factor PIF5 acts on ethylene biosynthesis and phytochrome signaling by distinct mechanisms. *Plant Cell*, **19**, 3915–3929.

Kim, S., Plagnol, V., Hu, T.T., Toomajian, C., Clark, R.M., Ossowski, S., Ecker, J.R., Weigel, D. and Nordborg, M. (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **39**, 1151–1155.

Klasen, J.R., Barbez, E., Meier, L., Meinshausen, N., Bühlmann, P., Koornneef, M., Busch, W. and Schneeberger, K. (2016) A multi-marker association method for genome-wide association studies without the need for population structure correction. *Nature Communications*, **7**, 13299.

Klukas, C., Chen, D. and Pape, J.-M. (2014) Integrated Analysis Platform: An Open-Source Information System for High-Throughput Plant Phenotyping. *Plant Physiol.*, **165**, 506–518.

Knoch, D., Abbadi, A., Grandke, F., Meyer, R.C., Samans, B., Werner, C.R., Snowdon, R.J. and Altmann, T. (2019) Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. *Plant Biotechnology Journal*, pbi.13171.

Korte, A. and Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, **9**, 29.

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

Krumsiek, J., Arnold, R. and Rattei, T. (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics*, **23**, 1026–1028.

Kubásek, J., Urban, O. and Šantrůček, J. (2013) C_4 plants use fluctuating light less efficiently than do C_3 plants: a study of growth, photosynthesis and carbon isotope discrimination. *Physiol Plantarum*, **149**, 528–539.

Kulheim, C., Agren, J. and Jansson, S. (2002) Rapid Regulation of Light Harvesting and Plant Fitness in the Field. *Science*, **297**, 91–93.

Kusmec, A., Srinivasan, S., Nettleton, D. and Schnable, P.S. (2017) Distinct genetic architectures for phenotype means and plasticities in *Zea mays*. *Nature Plants*, **3**, 715–723.

Kynast, R.G. and Riera-Lizarazu, O. (2011) Development and use of oat-maize chromosome additions and radiation hybrids. *Methods Mol. Biol.*, **701**, 259–284.

Laitinen, R.A.E. and Nikoloski, Z. (2019) Genetic basis of plasticity in plants. *Journal of Experimental Botany*, **70**, 739–745.

Lamesch, P., Berardini, T.Z., Li, D., et al. (2011) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, D1202–D1210.

Landoni, M., Vecchia, F.D., Gavazzi, G., Giulini, A., La Rocca, N., Rascio, N., Colombo, M., Bononi, M. and Consonni, G. (2007) The an1-4736 mutation of anther ear1 in maize alters scotomorphogenesis and the light response. *Plant Science*, **172**, 172–180.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. , **arXiv:1303.3997v2 [q-bio.GN]**. Available at: https://arxiv.org/abs/1303.3997.

Li, H., Handsaker, B., Wysoker, A., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Liang, C., Mao, L., Ware, D. and Stein, L. (2009) Evidence-based gene predictions in plant genomes. *Genome Res.*, **19**, 1912–1923.

Lichtenthaler, H.K., Babani, F. and Langsdorf, G. (2007) Chlorophyll fluorescence imaging of photosynthetic activity in sun and shade leaves of trees. *Photosynthesis Research*, **93**, 235–244.

Lillo, C., Lea, U.S. and Ruoff, P. (2008) Nutrient depletion as a key factor for manipulating gene expression and product formation in different branches of the flavonoid pathway. *Plant Cell Environ.*, **31**, 587–601.

Lin, R. and Wang, H. (2004) Arabidopsis FHY3/FAR1 gene family and distinct roles of its members in light control of Arabidopsis development. *Plant Physiol.*, **136**, 4010–4022.

Lindner, H., Raissig, M.T., Sailer, C., Shimosato-Asano, H., Bruggmann, R. and Grossniklaus, U. (2012) SNP-Ratio Mapping (SRM): identifying lethal alleles and mutations in complex genetic backgrounds by next-generation sequencing. *Genetics*, **191**, 1381–1386.

Lipka, A.E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P.J., Gore, M.A., Buckler, E.S. and Zhang, Z. (2012) GAPIT: genome association and prediction integrated tool. *Bioinformatics*, **28**, 2397–2399.

Liu, H., Niu, Y., Gonzalez-Portilla, P.J., et al. (2015) An ultra-high-density map as a community resource for discerning the genetic basis of quantitative traits in maize. *BMC Genomics*, **16**, 1078.

Liu, S., Yeh, C.-T., Tang, H.M., Nettleton, D. and Schnable, P.S. (2012) Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLoS ONE*, **7**, e36406.

Liu, X., Huang, M., Fan, B., Buckler, E.S. and Zhang, Z. (2016) Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet.*, **12**, e1005767.

Lobell, D.B. and Field, C.B. (2007) Global scale climate–crop yield relationships and the impacts of recent warming. *Environ. Res. Lett.*, **2**, 014002.

Lu, Y. (2011) The occurrence of a thylakoid-localized small zinc finger protein in land plants. *Plant Signal Behav*, **6**, 1881–1885.

Lu, Y., Hall, D.A. and Last, R.L. (2011) A small zinc finger thylakoid protein plays a role in maintenance of photosystem II in *Arabidopsis thaliana*. *Plant Cell*, **23**, 1861–1875.

Madec, S., Baret, F., Solan, B. de, Thomas, S., Dutartre, D., Jezequel, S., Hemmerlé, M., Colombeau, G. and Comar, A. (2017) High-Throughput Phenotyping of Plant Height: Comparing Unmanned Aerial Vehicles and Ground LiDAR Estimates. *Front Plant Sci*, **8**, 2002.

Marroni, F., Pinosio, S., Zaina, G., Fogolari, F., Felice, N., Cattonaro, F. and Morgante, M. (2011) Nucleotide diversity and linkage disequilibrium in *Populus nigra* cinnamyl alcohol dehydrogenase (CAD4) gene. *Tree Genetics & Genomes*, **7**, 1011–1023.

Mascher, M., Jost, M., Kuon, J.-E., Himmelbach, A., Aßfalg, A., Beier, S., Scholz, U., Graner, A. and Stein, N. (2014) Mapping-by-sequencing accelerates forward genetics in barley. *Genome Biol.*, **15**, R78.

Mascher, M., Wu, S., Amand, P.S., Stein, N. and Poland, J. (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS ONE*, **8**, e76925.

Matsubara, S. (2018) Growing plants in fluctuating environments: why bother? *Journal of Experimental Botany*, **69**, 4651–4654.

Matthews, J.S.A., Vialet-Chabrand, S. and Lawson, T. (2018) Acclimation to Fluctuating Light Impacts the Rapidity of Response and Diurnal Rhythm of Stomatal Conductance. *Plant Physiol.*, **176**, 1939–1951.

Matuszyńska, A., Heidari, S., Jahns, P. and Ebenhöh, O. (2016) A mathematical model of non-photochemical quenching to study short-term light memory in plants. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1857**, 1860–1869.

Mba, C. (2013) Induced Mutations Unleash the Potentials of Plant Genetic Resources for Food and Agriculture. *Agronomy*, **3**, 200–231.

Mba, C., Guimaraes, E.P. and Ghosh, K. (2012) Re-orienting crop improvement for the changing climatic conditions of the 21st century. *Agric Food Secur*, **1**, 7.

McGraw, K.O. and Wong, S.P. (1996) Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, **1**, 30–46.

McMaster, G.S. and Smika, D.E. (1988) Estimation and evaluation of winter wheat phenology in the central Great Plains. *Agricultural and Forest Meteorology*, **43**, 1–18.

Mekala, N.R., Suorsa, M., Rantala, M., Aro, E.-M. and Tikkanen, M. (2015) Plants Actively Avoid State Transitions upon Changes in Light Intensity: Role of Light-Harvesting Complex II Protein Dephosphorylation in High Light. *Plant Physiol.*, **168**, 721–734.

Méndez-Vigo, B., Picó, F.X., Ramiro, M., Martínez-Zapater, J.M. and Alonso-Blanco, C. (2011) Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in Arabidopsis. *Plant Physiol.*, **157**, 1942–1955.

Meyer, M. and Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*, **2010**.

Meyer, R.C., Steinfath, M., Lisec, J., et al. (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 4759–4764.

Meyer, R.S. and Purugganan, M.D. (2013) Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics*, **14**, 840–852.

Mi, C., Zhang, X., Li, S., Yang, J., Zhu, D. and Yang, Y. (2011) Assessment of environment lodging stress for maize using fuzzy synthetic evaluation. *Mathematical and Computer Modelling*, **54**, 1053–1060.

Mishra, Y., Johansson Jänkänpää, H., Kiss, A.Z., Funk, C., Schröder, W.P. and Jansson, S. (2012) Arabidopsis plants grown in the field and climate chambers significantly differ in leaf morphology and photosystem components. *BMC Plant Biol.*, **12**, 6.

Mogensen, H.L., Leduc, N., Matthys-Rochon, E. and Dumas, C. (1995) Nuclear DNA amounts in the egg and zygote of maize (*Zea mays* L.). *Planta*, **197**.

Möhring, J. and Piepho, H.-P. (2009) Comparison of Weighting in Two-Stage Analysis of Plant Breeding Trials. *Crop Science*, **49**, 1977.

Morosini, J.S., Freitas Mendonça, L. de, Lyra, D.H., Galli, G., Vidotti, M.S. and Fritsche-Neto, R. (2017) Association mapping for traits related to nitrogen use efficiency in tropical maize lines under field conditions. *Plant Soil*, **421**, 453–463.

Müller, P., Li, X.-P. and Niyogi, K.K. (2001) Non-Photochemical Quenching. A Response to Excess Light Energy. *Plant Physiol.*, **125**, 1558–1566.

Muraya, M.M., Chu, J., Zhao, Y., Junker, A., Klukas, C., Reif, J.C. and Altmann, T. (2017) Genetic variation of growth dynamics in maize (*Zea mays* L.) revealed through automated non-invasive phenotyping. *The Plant Journal*, **89**, 366–380.

Muscolo, A., Junker, A., Klukas, C., Weigelt-Fischer, K., Riewe, D. and Altmann, T. (2015) Phenotypic and metabolic responses to drought and salinity of four contrasting lentil accessions. *Journal of Experimental Botany*, **66**, 5467–5480.

Nagel, K.A., Putz, A., Gilmer, F., et al. (2012) GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biol.*, **39**, 891.

Nakagawa, S. and Schielzeth, H. (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev Camb Philos Soc*, **85**, 935–956.

Nakamura, Y., Kimura, A., Saga, H., et al. (2007) Differential metabolomics unraveling light/dark regulation of metabolic activities in Arabidopsis cell culture. *Planta*, **227**, 57–66.

Nannas, N.J. and Dawe, R.K. (2015) Genetic and genomic toolbox of *Zea mays*. *Genetics*, **199**, 655–669.

Nelissen, H., Moloney, M. and Inzé, D. (2014) Translational research: from pot to plot. *Plant Biotechnology Journal*, **12**, 277–285.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Van Lijsebettens, M., Kamiya, Y., Inzé, D. and Beemster, G.T.S. (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. *Curr. Biol.*, **22**, 1183–1187.

Neuffer, M.G. (1990) Location and description of dominant dwarf mutants. *Maize Genetics Cooperation News Letter*, **64**, 51.

Neuffer, M.G. and Coe, E.H.J. (1978) Paraffin oil technique for treating mature corn pollen with chemical mutagens. *Maydica*, **23**, 21–28.

Neuffer, M.G., Johal, G., Chang, M.T. and Hake, S. (2009) Mutagenesis – the Key to Genetic Analysis. In J. L. Bennetzen and S. Hake, eds. *Handbook of Maize*. New York, NY: Springer New York, pp. 63–84.

Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., Stein, N., Graner, A. and Kilian, B. (2015) Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis: Biomass accumulation in barley. *Plant, Cell & Environment*, **38**, 1980–1996.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., et al. (2010) Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, **15**, 684–692.

Nordström, K.J.V., Albani, M.C., James, G.V., Gutjahr, C., Hartwig, B., Turck, F., Paszkowski, U., Coupland, G. and Schneeberger, K. (2013) Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *Nature Biotechnology*, **31**, 325–330.

Norman, J.M. and Tanner, C.B. (1969) Transient Light Measurements in Plant Canopies. *Agronomy Journal*, **61**, 847.

Nozue, K., Harmer, S.L. and Maloof, J.N. (2011) Genomic analysis of circadian clock-, light-, and growth-correlated genes reveals PHYTOCHROME-INTERACTING FACTOR5 as a modulator of auxin signaling in Arabidopsis. *Plant Physiol.*, **156**, 357–372.

Ouyang, S., Zhu, W., Hamilton, J., et al. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Research*, **35**, D883-887.

Pacín, M., Legris, M. and Casal, J.J. (2013) COP1 re-accumulates in the nucleus under shade. *The Plant Journal*, **75**, 631–641.

Pan, S., Tian, H., Dangal, S.R.S., Yang, Q., Yang, J., Lu, C., Tao, B., Ren, W. and Ouyang, Z. (2015) Responses of global terrestrial evapotranspiration to climate change and increasing atmospheric CO_2 in the 21st century. *Earth's Future*, **3**, 15–35.

Parent, B. and Tardieu, F. (2012) Temperature responses of developmental processes have not been affected by breeding in different ecological areas for 17 crop species. *New Phytologist*, **194**, 760–774.

Pearcy, R.W. (1990) Sunflecks and Photosynthesis in Plant Canopies. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **41**, 421–453.

Pearcy, R.W., Roden, J.S. and Gamon, J.A. (1990) Sunfleck dynamics in relation to canopy structure in a soybean (*Glycine max* (L.) Merr.) canopy. *Agricultural and Forest Meteorology*, **52**, 359–372.

Peiffer, J.A., Romay, M.C., Gore, M.A., et al. (2014) The genetic architecture of maize height. *Genetics*, **196**, 1337–1356.

Peng, S., Huang, J., Sheehy, J.E., Laza, R.C., Visperas, R.M., Zhong, X., Centeno, G.S., Khush, G.S. and Cassman, K.G. (2004) Rice yields decline with higher night temperature from global warming. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 9971–9975.

Phinney, B.O. (1956) GROWTH RESPONSE OF SINGLE-GENE DWARF MUTANTS IN MAIZE TO GIBBERELLIC ACID. *Proc. Natl. Acad. Sci. U.S.A.*, **42**, 185–189.

Platt, A., Horton, M., Huang, Y.S., et al. (2010) The scale of population structure in *Arabidopsis thaliana*. *PLoS Genet.*, **6**, e1000843.

Pommerrenig, B., Junker, A., Abreu, I., Bieber, A., Fuge, J., Willner, E., Bienert, M.D., Altmann, T. and Bienert, G.P. (2018) Identification of Rapeseed (*Brassica napus*) Cultivars With a High Tolerance to Boron-Deficient Conditions. *Front Plant Sci*, **9**, 1142.

Poorter, H., Bühler, J., Dusschoten, D. van, Climent, J. and Postma, J.A. (2012) Pot size matters: a meta-analysis of the effects of rooting volume on plant growth. *Functional Plant Biol.*, **39**, 839.

Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., Putten, W.H. van der, 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**, 838–855.

Popescu, S.C., Popescu, G.V., Bachan, S., Zhang, Z., Gerstein, M., Snyder, M. and Dinesh-Kumar, S.P. (2009) MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev.*, **23**, 80–92.

Preuss, A., Stracke, R., Weisshaar, B., Hillebrecht, A., Matern, U. and Martens, S. (2009) *Arabidopsis thaliana* expresses a second functional flavonol synthase. *FEBS Lett.*, **583**, 1981–1986.

Price, A.L., Zaitlen, N.A., Reich, D. and Patterson, N. (2010) New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, **11**, 459–463.

Qi, W., Zhu, T., Tian, Z., Li, C., Zhang, W. and Song, R. (2016) High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. *BMC Biotechnology*, **16**.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*, Vienna, Austria. Available at: http://www.R-project.org/.

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**, 1293.

Remington, D.L., Thornsberry, J.M., Matsuoka, Y., Wilson, L.M., Whitt, S.R., Doebley, J., Kresovich, S., Goodman, M.M. and Buckler, E.S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 11479–11484.

Ren, B., Zhang, J., Li, X., Fan, X., Dong, S., Liu, P. and Zhao, B. (2014) Effects of waterlogging on the yield and growth of summer maize under field conditions. *Canadian Journal of Plant Science*, **94**, 23–31.

Retkute, R., Smith-Unna, S.E., Smith, R.W., Burgess, A.J., Jensen, O.E., Johnson, G.N., Preston, S.P. and Murchie, E.H. (2015) Exploiting heterogeneous environments: does photosynthetic acclimation optimize carbon gain in fluctuating light? *Journal of Experimental Botany*, **66**, 2437–2447.

Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., et al. (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics*, **44**, 217–220.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol. Biol.*, **53**, 247–259.

RStudio Team (2015) *RStudio: Integrated Development for R.*, RStudio, Inc., Boston, MA. Available at: http://www.rstudio.com/.

Ruban, A.V. (2016) Nonphotochemical Chlorophyll Fluorescence Quenching: Mechanism and Effectiveness in Protecting Plants from Photodamage. *Plant Physiol.*, **170**, 1903–1916.

Rungrat, T., Awlia, M., Brown, T., et al. (2016) Using Phenomic Analysis of Photosynthetic Function for Abiotic Stress Response Gene Discovery. *Arabidopsis Book*, **14**, e0185.

Sakuraba, Y., Jeong, J., Kang, M.-Y., Kim, J., Paek, N.-C. and Choi, G. (2014) Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in Arabidopsis. *Nature Communications*, **5**, 4636.

Salamov, A.A. and Solovyev, V.V. (2000) Ab initio gene finding in Drosophila genomic DNA. *Genome Res.*, **10**, 516–522.

Salas Fernandez, M.G., Bao, Y., Tang, L. and Schnable, P.S. (2017) A High-Throughput, Field-Based Phenotyping Technology for Tall Biomass Crops. *Plant Physiol.*, **174**, 2008–2022.

Salisbury, F.B., Spomer, G.G., Sobral, M. and Ward, R.T. (1968) Analysis of an alpine environment. *Botanical Gazette*, **129**, 16–32.

Sarin, S., Prabhu, S., O'Meara, M.M., Pe'er, I. and Hobert, O. (2008) *Caenorhabditis elegans* mutant allele identification by whole-genome sequencing. *Nature Methods*, **5**, 865–867.

Sasaki, T. and Burr, B. (2000) International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.*, **3**, 138–141.

Sato, N. (2010) Phylogenomic and structural modeling analyses of the PsbP superfamily reveal multiple small segment additions in the evolution of photosystem II-associated PsbP protein in green plants. *Mol. Phylogenet. Evol.*, **56**, 176–186.

Schilling, S., Gramzow, L., Lobbes, D., et al. (2015) Non-canonical structure, function and phylogeny of the B$_{sister}$ MADS-box gene *OsMADS30* of rice (*Oryza sativa*). *The Plant Journal*, **84**, 1059–1072.

Schlenker, W. and Roberts, M.J. (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 15594–15598.

Schnable, P.S., Ware, D., Fulton, R.S., et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, **326**, 1112–1115.

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A.H., Nielsen, K.L., Jørgensen, J.-E., Weigel, D. and Andersen, S.U. (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nature Methods*, **6**, 550–551.

Schneider, T., Bolger, A., Zeier, J., Preiskowski, S., Benes, V., Trenkamp, S., Usadel, B., Farré, E.M. and Matsubara, S. (2019) Fluctuating Light Interacts with Time of Day and Leaf Development Stage to Reprogram Gene Expression. *Plant Physiol.*, **179**, 1632–1657.

Scholl, R., Sachs, M.M. and Ware, D. (2003) Maintaining Collections of Mutants for Plant Functional Genomics. In *Plant Functional Genomics*. New Jersey: Humana Press, pp. 311–326.

Seaton, D.D., Smith, R.W., Song, Y.H., et al. (2015) Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Mol. Syst. Biol.*, **11**, 776.

Serivichyaswat, P., Ryu, H.-S., Kim, W., Kim, S., Chung, K.S., Kim, J.J. and Ahn, J.H. (2015) Expression of the floral repressor miRNA156 is positively regulated by the AGAMOUS-like proteins AGL15 and AGL18. *Mol. Cells*, **38**, 259–266.

Shin, J.-H., Blay, S., Graham, J. and McNeney, B. (2006) LDheatmap: An *R* Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. *Journal of Statistical Software*, **16**.

Slattery, R.A., Walker, B.J., Weber, A.P.M. and Ort, D.R. (2018) The Impacts of Fluctuating Light on Crop Performance. *Plant Physiol.*, **176**, 990–1003.

Song, Y., Yang, C., Gao, S., Zhang, W., Li, L. and Kuai, B. (2014) Age-triggered and dark-induced leaf senescence require the bHLH transcription factors PIF3, 4, and 5. *Mol Plant*, **7**, 1776–1787.

Stacklies, W., Redestig, H., Scholz, M., Walther, D. and Selbig, J. (2007) pcaMethods a bioconductor package providing PCA methods for incomplete data. *Bioinformatics*, **23**, 1164–1167.

Staswick, P.E., Tiryaki, I. and Rowe, M.L. (2002) Jasmonate response locus JAR1 and several related Arabidopsis genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. *Plant Cell*, **14**, 1405–1415.

Stracke, R., De Vos, R.C.H., Bartelniewoehner, L., Ishihara, H., Sagasser, M., Martens, S. and Weisshaar, B. (2009) Metabolomic and genetic analyses of flavonol synthesis in Arabidopsis thaliana support the in vivo involvement of leucoanthocyanidin dioxygenase. *Planta*, **229**, 427–445.

Su, C., Wang, W., Gong, S., Zuo, J., Li, S. and Xu, S. (2017) High Density Linkage Map Construction and Mapping of Yield Trait QTLs in Maize (*Zea mays*) Using the Genotyping-by-Sequencing (GBS) Technology. *Front Plant Sci*, **8**.

Sultan, S.E. (2000) Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.*, **5**, 537–542.

Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol.*, **169**, 931–945.

Svyatyna, K., Jikumaru, Y., Brendel, R., Reichelt, M., Mithöfer, A., Takano, M., Kamiya, Y., Nick, P. and Riemann, M. (2014) Light induces jasmonate-isoleucine conjugation via OsJAR1-dependent and -independent pathways in rice. *Plant Cell Environ.*, **37**, 827–839.

Tabata, R., Kamiya, T., Shigenobu, S., Yamaguchi, K., Yamada, M., Hasebe, M., Fujiwara, T. and Sawa, S. (2013) Identification of an EMS-induced causal mutation in a gene required for boron-mediated root development by low-coverage genome re-sequencing in Arabidopsis. *Plant Signal Behav*, **8**, e22534.

Takase, M., Mizoguchi, T., Kozuka, T. and Tsukaya, H. (2013) The unique function of the Arabidopsis circadian clock gene *PRR5* in the regulation of shade avoidance response. *Plant Signal Behav*, **8**, e23534.

Tang, J.D., Perkins, A., Williams, W.P. and Warburton, M.L. (2015) Using genome-wide associations to identify metabolic pathways involved in maize aflatoxin accumulation resistance. *BMC Genomics*, **16**, 673.

The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, **45**, D158–D169.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W. and Su, Z. (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research*, **45**, W122–W129.

Tikkanen, M., Grieco, M., Kangasjarvi, S. and Aro, E.-M. (2010) Thylakoid Protein Phosphorylation in Higher Plant Chloroplasts Optimizes Electron Transfer under Fluctuating Light. *Plant Physiol.*, **152**, 723–735.

Till, B.J., Reynolds, S.H., Weil, C., et al. (2004) Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.*, **4**, 12.

Timm, S., Wittmiß, M., Gamlien, S., et al. (2015) Mitochondrial Dihydrolipoyl Dehydrogenase Activity Shapes Photosynthesis and Photorespiration of *Arabidopsis thaliana*. *Plant Cell*, **27**, 1968–1984.

Townsend, A.J., Retkute, R., Chinnathambi, K., Randall, J.W.P., Foulkes, J., Carmo-Silva, E. and Murchie, E.H. (2018) Suboptimal Acclimation of Photosynthesis to Light in Wheat Canopies. *Plant Physiol.*, **176**, 1233–1246.

Trachsel, S., Kaeppler, S.M., Brown, K.M. and Lynch, J.P. (2011) Shovelomics: high throughput phenotyping of maize (*Zea mays* L.) root architecture in the field. *Plant and Soil*, **341**, 75–87.

Tschiersch, H., Junker, A., Meyer, R.C. and Altmann, T. (2017) Establishment of integrated protocols for automated high throughput kinetic chlorophyll fluorescence analyses. *Plant Methods*, **13**, 54.

Tsukaya, H., Kozuka, T. and Kim, G.-T. (2002) Genetic control of petiole length in *Arabidopsis thaliana*. *Plant Cell Physiol.*, **43**, 1221–1228.

Udy, D.B., Belcher, S., Williams-Carrier, R., Gualberto, J.M. and Barkan, A. (2012) Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize. *Plant Physiol.*, **160**, 1420–1431.

Ushijima, T., Hanada, K., Gotoh, E., et al. (2017) Light Controls Protein Localization through Phytochrome-Mediated Alternative Promoter Selection. *Cell*, **171**, 1316-1325.e12.

Van Kleunen, M. and Fischer, M. (2005) Constraints on the evolution of adaptive phenotypic plasticity in plants: Research review. *New Phytologist*, **166**, 49–60.

Varshney, R.K., Singh, V.K., Kumar, A., Powell, W. and Sorrells, M.E. (2018) Can genomics deliver climate-change ready crops? *Curr. Opin. Plant Biol.*, **45**, 205–211.

Vialet-Chabrand, S., Matthews, J.S.A., Simkin, A.J., Raines, C.A. and Lawson, T. (2017) Importance of Fluctuations in Light on Plant Photosynthetic Acclimation. *Plant Physiol.*, **173**, 2163–2179.

Wagner, R., Dietzel, L., Bräutigam, K., Fischer, W. and Pfannschmidt, T. (2008) The long-term response to fluctuating light quality is an important and distinct light acclimation mechanism that supports survival of *Arabidopsis thaliana* under low light conditions. *Planta*, **228**, 573–587.

Wang, H. and Deng, X.W. (2002) Arabidopsis FHY3 defines a key phytochrome A signaling component directly interacting with its homologous partner FAR1. *EMBO J.*, **21**, 1339–1349.

Wang, H., Ngwenyama, N., Liu, Y., Walker, J.C. and Zhang, S. (2007) Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in Arabidopsis. *Plant Cell*, **19**, 63–73.

Wang, P., Du, Y., Zhao, X., Miao, Y. and Song, C.-P. (2013) The MPK6-ERF6-ROS-responsive cis-acting Element7/GCC box complex modulates oxidative gene transcription and the oxidative response in Arabidopsis. *Plant Physiol.*, **161**, 1392–1408.

Wei, L., Guo, J., Ouyang, M., Sun, X., Ma, J., Chi, W., Lu, C. and Zhang, L. (2010) LPA19, a Psb27 homolog in *Arabidopsis thaliana*, facilitates D1 protein precursor processing during PSII biogenesis. *J. Biol. Chem.*, **285**, 21391–21398.

Weil, C. and Monde, R. (2009) TILLING and Point Mutation Detection. In J. L. Bennetzen and S. Hake, eds. *Handbook of Maize*. New York, NY: Springer New York, pp. 585–596.

Wetterstrand, K. (2017) DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata. Accessed [11/21/2017].

Williams-Carrier, R., Stiffler, N., Belcher, S., Kroeger, T., Stern, D.B., Monde, R.-A., Coalter, R. and Barkan, A. (2010) Use of Illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy Mutator lines of maize. *The Plant Journal*, **63**, 167–177.

Winkel-Shirley, B. (2001) Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. *Plant Physiol.*, **126**, 485–493.

Xiao, S. and Chye, M.-L. (2009) An Arabidopsis family of six acyl-CoA-binding proteins has three cytosolic members. *Plant Physiol. Biochem.*, **47**, 479–484.

Xiao, S., Gao, W., Chen, Q.-F., Chan, S.-W., Zheng, S.-X., Ma, J., Wang, M., Welti, R. and Chye, M.-L. (2010) Overexpression of Arabidopsis acyl-CoA binding protein ACBP3 promotes starvation-induced and age-dependent leaf senescence. *Plant Cell*, **22**, 1463–1482.

Xu, C., Zhang, H., Sun, J., et al. (2018) Genome-wide association study dissects yield components associated with low-phosphorus stress tolerance in maize. *Theor. Appl. Genet.*, **131**, 1699–1714.

Yamori, W. (2016) Photosynthetic response to fluctuating environments and photoprotective strategies under abiotic stress. *J Plant Res*, **129**, 379–395.

Yamori, W., Makino, A. and Shikanai, T. (2016) A physiological role of cyclic electron transport around photosystem I in sustaining photosynthesis under fluctuating light in rice. *Sci Rep*, **6**, 20147.

Yang, G., Liu, J., Zhao, C., et al. (2017) Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front Plant Sci*, **8**, 1111.

Yang, W., Guo, Z., Huang, C., et al. (2014) Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*, **5**, 5087.

Yi, X., Hargett, S.R., Liu, H., Frankel, L.K. and Bricker, T.M. (2007) The PsbP Protein Is Required for Photosystem II Complex Assembly/Stability and Photoautotrophy in *Arabidopsis thaliana*. *J. Biol. Chem.*, **282**, 24833–24841.

Young, D.R. and Smith, W.K. (1983) Effect of Cloudcover on Photosynthesis and Transpiration in the Subalpine Understory Species *Arnica latifolia*. *Ecology*, **64**, 681.

Yu, J., Hu, S., Wang, J., et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science*, **296**, 79–92.

Zhang, X., Hause, R.J. and Borevitz, J.O. (2012) Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in *Arabidopsis thaliana*. *G3*, **2**, 29–34.

Zhang, X., Huang, C., Wu, D., et al. (2017) High-Throughput Phenotyping and QTL Mapping Reveals the Genetic Architecture of Maize Plant Growth. *Plant Physiol.*, **173**, 1554–1564.

Zhang, Y., Liu, Z., Chen, Y., He, J.-X. and Bi, Y. (2015) PHYTOCHROME-INTERACTING FACTOR 5 (PIF5) positively regulates dark-induced senescence and chlorophyll degradation in Arabidopsis. *Plant Sci.*, **237**, 57–68.

Zhao, Y., Mette, M.F. and Reif, J.C. (2015) Genomic selection in hybrid breeding. *Plant Breed*, **134**, 1–10.

Zheng, S.-X., Xiao, S. and Chye, M.-L. (2012) The gene encoding Arabidopsis acyl-CoA-binding protein 3 is pathogen inducible and subject to circadian regulation. *Journal of Experimental Botany*, **63**, 2985–3000.

Zuryn, S., Le Gras, S., Jamet, K. and Jarriault, S. (2010) A Strategy for Direct Mapping and Identification of Mutations by Whole-Genome Sequencing. *Genetics*, **186**, 427–430.

CURRICULUM VITAE

Marc Christian Heuermann, M.Sc.

Place & Date of birth:

Herdecke | 16 February 1987

School:

1993 – 1997 Grundschule Aplerbecker-Mark

1997 – 2004 Gymnasium an der Schweizer Allee

2004 – 2007 Phoenix Gymnasium

(**Graduation: Abitur**)

Alternative civilian service:

2007 – 2008

St. Josephs Hospital Dortmund Hörde

Education:

2008 – 2011 **Bachelor's degree** at the University Bayreuth

(**Graduation: Bachelor of Science**)

Bachelor's thesis: „Untersuchung zum Nutzen der nukleären Primer At2g06530a und At2g34620b

für phylogenetische Untersuchungen bei *Marsdenieae*"

2011 – 2014 **Master's degree** at the Martin-Luther-University Halle-Wittenberg

(**Graduation: Master of Science**)

Master's thesis: „Characterization of the tyrosine aminotransferase gene family in *Arabidopsis thaliana*"

since 2014 PhD student at IPK-Gatersleben

EIDESSTATTLICHE ERKLÄRUNG / DECLARATION UNDER OATH

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

_____   _____

Datum / Date   Unterschrift des Antragstellers / Signature of the applicant

Erklärung über bestehende Vorstrafen und anhängige Ermittlungsverfahren / Declaration concerning criminal record and pending investigations

Hiermit erkläre ich, dass ich weder vorbestraft bin noch dass gegen mich Ermittlungsverfahren anhängig sind.

I hereby declare that I have no criminal record and that no preliminary investigations are pending against me.

_____   _____

Datum / Date   Unterschrift des Antragstellers / Signature of the applicant

ANNEX

Table of ANNEX Figures/Tables

Arabidopsis' response to fluctuating light treatment

Table S 1_1: The accession panel with 384 natural accessions. AKZ_ID is the internal accession id number, name the trivial name of the accession, longitude and latitude give information about the coordinates of sample collections, and the region categorizes the coordinates into broad regions of the planet.

AKZ_ID	name	longitude	latitude	region	genotyped
A001	11PNA4_IPK	-86.3253	42.0945	North America	1
A002	328PNA_IPK	-86.3253	42.0945	North America	1
A004	Aa-0_IPK	9.57073	50.9167	Europe	1
A006	Ak-1_MPI	7.62551	48.0683	Europe	1
A007	Akita_MPI	140.1	39.77	Asia	1
A009	Alc-0_IPK	-3.36	40.49	Europe	1
A010	ALL1-2_IPK	1.48333	45.2667	Europe	1
A011	Alst-1_IPK	-2.4333	54.8	Europe	1
A013	Amel-1_IPK	5.73	53.448	Europe	1
A015	An-1_IPK	4.4	51.2167	Europe	1
A017	An-2_IPK	4.4	51.2167	Europe	1
A018	Ang-0_IPK	5.3	50.3	Europe	1
A020	Ang-0_UHOH	5.3	50.3	Europe	1
A021	Ang-1_UHOH	5.3	50.3	Europe	1
A022	Ann-1_IPK	6.13028	45.9	Europe	1
A024	Arby-1_IPK	16.7999	59.4308	Europe	1
A026	Ba-1_IPK	-4.79821	56.5459	Europe	1
A027	Ba1-2_IPK	12.9	56.4	Europe	1
A028	Baa-1_IPK	6.1	51.3333	Europe	1
A030	Bay-0_IPK	11	49	Europe	1
A032	Bch-1_MPI	9.3166	49.5166	Europe	1
A033	Bd-0_MPI	13.287	52.4584	Europe	1
A035	Be-0_MPI	8.6161	49.6803	Europe	1
A036	Be-1_IPK	8.6161	49.6803	Europe	1
A037	Belm_IPK	12.4833	42.1167	Europe	1
A038	Benk-1_IPK	5.675	52	Europe	1
A039	Bg-2_IPK	-122.305	47.6479	North America	1
A040	Bl-1_MPI	11.3396	44.5041	Europe	1
A041	Bla-1_IPK	2.79	41.6833	Europe	1
A042	Bla-11_MPI	2.79	41.6833	Europe	1
A044	Blh-1_IPK	19.85	48.3	Europe	1
A045	Blh-1_MPI	19.85	48.3	Europe	1
A046	Blh-2_IPK	19.85	48.3	Europe	1
A047	Boot-1_IPK	-3.2667	54.4	Europe	1
A050	Bor-1_IPK	16.2326	49.4013	Europe	1
A054	Bor-4_IPK	16.2326	49.4013	Europe	1
A056	Br-0_MPI	16.6166	49.2	Europe	1
A057	Bs-1_UHOH	7.5	47.5	Europe	1
A058	Bs-2_IPK	7.5	47.5	Europe	1
A059	Bsch-0_IPK	8.6667	40.0167	Europe	1
A060	Bsch-2_MPI	8.6667	40.0167	Europe	1
A062	Bu-0_IPK	9.5	50.5	Europe	1
A064	Bu-2_MPI	9.5	50.5	Europe	1
A068	Bur-0_IPK	-9.39	53.87	Europe	1
A070	C24_IPK	-8.42639	40.2077	Europe	1
A072	Ca-0_IPK	8.26607	50.2981	Europe	1
A074	Cal-0_UHOH	-1.64293	53.2699	Europe	1
A075	CAM-16_IPK	-4.58333	48.2667	Europe	1
A076	CAM-61_IPK	-4.58333	48.2667	Europe	1
A078	Can-0_MPI	-13.4811	29.2144	Africa	1
A080	Cen-0_IPK	0.5	49	Europe	1
A081	Cha-0_IPK	7.1167	46.0333	Europe	1
A084	Chat-1_IPK	1.33867	48.0717	Europe	1
A085	Chi-0_UHOH	34.7361	53.7502	Europe	1
A086	CIBC-17_IPK	-0.6383	51.4083	Europe	1
A087	CIBC-2_IPK	-0.6383	51.4083	Europe	1
A088	CIBC-4_IPK	-0.6383	51.4083	Europe	1

A089	CIBC-5_IPK	-0.6383	51.4083	Europe	1
A091	Cit-0_IPK	2.54038	43.3779	Europe	1
A092	Cl-0_MPI	NA	NA	unknown	1
A093	CLE-6_IPK	-0.483333	48.9167	Europe	0
A094	Cnt-1_IPK	1.1	51.3	Europe	1
A095	Co_IPK	-8.42639	40.2077	Europe	1
A096	Co-2_IPK	-8.25	40.12	Europe	1
A097	Co-3_MPI	-8.25	40.12	Europe	1
A098	Co-4_IPK	-8.25	40.12	Europe	1
A099	Col-0_IPK	8.008504	50.392033	Europe	1
A101	Com-1_IPK	2.823	49.416	Europe	1
A102	CSHL-5_IPK	-73.4675	40.8585	North America	1
A103	Ct-1_IPK	15	37.3	Europe	1
A105	CUR-3_IPK	1.75	45	Europe	1
A106	Cvi-0_IPK	-23.6167	15.1111	Africa	1
A109	Da(1)-12_IPK	NA	NA	Europe	1
A111	Da-0_IPK	8.65081	49.8724	Europe	1
A113	Db-0_IPK	8.324	50.3055	Europe	1
A114	Db-1_UHOH	8.324	50.3055	Europe	1
A115	Di-1_IPK	5	47	Europe	1
A116	Dijon-M_MPI	5	47	Europe	1
A117	Do-0_IPK	8.2372	50.7224	Europe	1
A118	Dr-0_MPI	13.7336	51.051	Europe	1
A119	Dra-0_MPI	16.2667	49.4167	Europe	1
A120	Dra-2_IPK	16.2667	49.4167	Europe	1
A121	DraIV1-14_IPK	16.2815	49.4112	Europe	1
A122	DraIV1-5_IPK	16.2815	49.4112	Europe	1
A123	DraIV1-7_IPK	16.2815	49.4112	Europe	0
A124	DraIV6-16_IPK	16.2815	49.4112	Europe	1
A125	DraIV6-35_IPK	16.2815	49.4112	Europe	1
A126	Duk_IPK	16.2	49.1	Europe	1
A127	Durh-1_UHOH	-1.5733	54.7761	Europe	1
A128	Ede-1_IPK	5.66667	52.0333	Europe	1
A129	Eden-1_IPK	18.177	62.877	Europe	1
A131	Edi-0_IPK	-3	56	Europe	1
A133	Ei-2_MPI	6.3	50.3	Europe	1
A134	Ei-4_IPK	6.3	50.3	Europe	1
A136	El-0_MPI	9.68253	51.5105	Europe	1
A137	En-1_UHOH	8.5	50	Europe	1
A138	Enkheim-D_MPI	8.5	50	Europe	1
A139	Ep-0_IPK	8.38912	50.1721	Europe	1
A141	Er-0_MPI	11.0087	49.5955	Europe	1
A142	Es-0_IPK	24.5682	60.1997	Europe	1
A143	Est-0_IPK	25.3	58.3	Europe	1
A144	Est-1_IPK	25.3	58.3	Europe	1
A146	Fei-0_IPK	-8.54	40.92	Europe	1
A150	Fi-0_UHOH	8.0167	50.5	Europe	1
A151	Fi-1_IPK	8.0167	50.5	Europe	1
A153	Fr-2_UHOH	8.6822	50.1102	Europe	1
A154	Fr-4_IPK	8.6822	50.1102	Europe	1
A156	Ga-0_IPK	8	50.3	Europe	1
A158	Ga-2_IPK	8	50.3	Europe	1
A159	Gd-1_IPK	10.5	53.5	Europe	1
A163	Ge-1_IPK	6.08	46.5	Europe	1
A164	Ge-2_MPI	6.08	46.5	Europe	1
A166	Gel-1_IPK	5.86667	51.0167	Europe	1
A167	Gie-0_IPK	8.67825	50.584	Europe	1
A169	Gö-0_IPK	9.9355	51.5338	Europe	1
A170	Gö-2_MPI	9.9355	51.5338	Europe	1
A171	Golm-1_MPI	12.969162	52.416213	Europe	1
A173	GOT-7_MPI	9.9355	51.5338	Europe	1
A174	Gr_MPI	15.5	47	Europe	1
A175	Gr-1_IPK	15.5	47	Europe	1
A178	Gr-5_IPK	15.5	47	Europe	1
A179	Gre-0_MPI	-85.2532	43.178	North America	1
A180	Gu-0_UHOH	8	50.3	Europe	1
A181	Gu-1_IPK	8	50.3	Europe	1

A183	Gy-0_IPK	2	49	Europe	1
A185	Ha-0_IPK	9.73569	52.3721	Europe	1
A187	Hau-0_IPK	12.5686	55.675	Europe	1
A188	Hey-1_IPK	5.9	51.25	Europe	1
A189	Hh-0_IPK	9.88682	54.4175	Europe	1
A190	Hi-0_IPK	5	52	Europe	1
A191	Hl-3_MPI	9.37827	52.1444	Europe	1
A192	Hn-0_IPK	8.28844	51.3472	Europe	1
A194	Hod_IPK	17.1	48.8	Europe	1
A195	HOG_MPI	69.712	38.717	Asia	1
A196	Hoh-1_UHOH	9.2116	48.7129	Europe	1
A197	Hovdala-2_IPK	13.74	56.1	Europe	1
A198	HR-10_UHOH	-0.6383	51.4083	Europe	1
A199	HR-5_IPK	-0.6383	51.4083	Europe	1
A202	Hs-0_IPK	9.44	52.24	Europe	1
A205	HSm_IPK	15.76	49.33	Europe	1
A206	In-0_IPK	11.5	47.5	Europe	1
A208	Is-1_MPI	7.5	50.5	Europe	1
A209	Je-0_IPK	11.587	50.927	Europe	1
A211	Je54_MPI	15	49	Europe	1
A212	JEA_IPK	7.33333	43.6833	Europe	1
A213	Jea_MPI	7.33	43.69	Europe	1
A215	Jl-3_IPK	16.61	49.31	Europe	1
A216	Jm-1_IPK	15	49	Europe	1
A218	Kä-0_IPK	14	47	Europe	1
A221	Kas-2_IPK (Kas-1)	77	35	Asia	1
A223	Kb-0_UHOH	8.5	50.183	Europe	1
A224	KBS-Mac-8_IPK	-85.398	42.405	North America	1
A225	Kelsterbach-2_IPK	8.5333	50.0667	Europe	1
A226	Kelsterbach-4_IPK	8.5333	50.0667	Europe	1
A228	Kil-0_UHOH	-5.66364	55.6395	Europe	1
A229	Kin-0_IPK	-85.25472	43.362776	North America	1
A231	Kl-0_MPI	6.9666	50.95	Europe	1
A232	Kl-5_IPK	6.9666	50.95	Europe	1
A233	Kn-0_IPK	23.8924	54.8969	Europe	1
A235	KNO-11_IPK	-86.621	41.2816	North America	1
A237	Kno-18_IPK	-86.621	41.2816	North America	1
A240	Koln_IPK	7	51	Europe	1
A241	Kondara_MPI	68.49	38.48	Asia	1
A242	Kr-0_IPK	6.55934	51.3317	Europe	1
A244	Kro-0_IPK	8.96617	50.0742	Europe	1
A245	Krot-2_IPK	11.5722	49.631	Europe	1
A247	Kz-1_UHOH	73.1	49.5	Asia	1
A248	Kz-9_UHOH	73.1	49.5	Asia	1
A249	LAC-3_IPK	6.81667	47.7	Europe	1
A250	LAC-5_IPK	6.81667	47.7	Europe	1
A252	Lan-0_MPI	-3.78181	55.6739	Europe	1
A253	Laud-1_UHOH	-2.75	55.7	Europe	1
A254	Lc-0_IPK	-4	57	Europe	1
A255	LDV-14_IPK	-4.06667	48.5167	Europe	1
A256	LDV-25_IPK	-4.06667	48.5167	Europe	1
A257	LDV-34_IPK	-4.06667	48.5167	Europe	1
A258	LDV-58_IPK	-4.06667	48.5167	Europe	1
A260	Ler-1_IPK	10.8719	47.984	Europe	1
A262	Li-3_IPK	8.0666	50.3833	Europe	1
A263	Li-5:2_IPK	8.0666	50.3833	Europe	1
A264	Li-6_IPK	8.0666	50.3833	Europe	1
A266	Li-7_IPK	8.0666	50.3833	Europe	1
A267	Liarum_IPK	13.85	55.95	Europe	1
A268	Limeport_UHOH	-75.4472	40.5088	North America	1
A270	Lip-0_IPK	19.3	50	Europe	1
A273	Lis-1_IPK	14.7	56	Europe	1
A275	LL-0_IPK	2.49	41.59	Europe	1
A277	Ll-OF-095_IPK	-72.9069	40.7777	North America	1
A278	Lm_MPI	0.5	48	Europe	1
A280	Lm-2_IPK	0.5	48	Europe	1
A281	Lom1-1_IPK	13.9	56.09	Europe	1

A282	Löv-5_MPI	18.079	62.801	Europe	1
A283	Lp2-2_IPK	16.81	49.38	Europe	1
A285	Lp2-6_IPK	16.81	49.38	Europe	1
A286	Lu_MPI	13.2	55.71	Europe	1
A287	Lz-0_IPK	3.3	46	Europe	1
A289	Map-42_IPK	-86.412	42.166	North America	1
A290	Mc-0_IPK	-2.3	54.6167	Europe	1
A291	Me-0_UHOH	10.1138	51.9183	Europe	1
A292	Mh-0_IPK	7.5	50.95	Europe	1
A293	Mh-1_MPI	7.5	50.95	Europe	1
A294	MIB-15_IPK	5.31667	47.3833	Europe	1
A295	MIB-22_IPK	5.31667	47.3833	Europe	1
A296	MIB-28_IPK	5.31667	47.3833	Europe	1
A297	MIB-84_IPK	5.31667	47.3833	Europe	1
A298	MNF-Pot-48_IPK	-86.2657	43.595	North America	1
A299	Mnz-0_IPK	8.26664	50.001	Europe	1
A300	MOG-37_IPK	-4.06667	48.6667	Europe	1
A302	Mrk-0_IPK	9.3	49	Europe	1
A304	Ms-0_MPI	37.6322	55.7522	Europe	1
A305	Mt-0_IPK	22.46	32.34	Africa	1
A307	Mz-0_IPK	8.3	50.3	Europe	1
A310	N13_IPK	34.15	61.36	Europe	1
A313	N4_IPK	34.36777	61.84423	Europe	1
A314	N7_IPK	34.15	61.36	Europe	1
A315	Na-1_IPK	1.5	47.5	Europe	1
A317	Nc-1_IPK	6.25	48.6167	Europe	1
A318	Nd_MPI	10	50	Europe	1
A320	Nd-1_IPK	10	50	Europe	1
A323	NFA-10_IPK	-0.6383	51.4083	Europe	1
A325	NFA-8_IPK	-0.6383	51.4083	Europe	1
A327	NFC-20_IPK	-0.6383	51.4083	Europe	1
A329	No-0_IPK	13.2995	51.0581	Europe	1
A332	Nok-1_IPK	4.45	52.24	Europe	1
A333	Nok-2_MPI	4.45	52.24	Europe	1
A334	Nw-0_IPK	8.5	50.5	Europe	1
A335	Nw-2_IPK	8.5	50.5	Europe	1
A336	Nw-3_MPI	8.5	50.5	Europe	1
A337	Nz1_IPK	175.283	-37.7871	New Zealand	1
A338	Ob-1_IPK	8.5833	50.2	Europe	1
A339	Old-1_IPK	8.2	53.1667	Europe	1
A341	Or-0_IPK	8.01161	50.3827	Europe	1
A342	Ors-1_IPK	22.3955	44.7203	Europe	1
A343	Ors-2_IPK	22.3955	44.7203	Europe	1
A344	Ost-0_IPK	18.37	60.25	Europe	1
A346	Ove-0_MPI	8.42255	53.3422	Europe	1
A347	Ove-0_UHOH	8.42255	53.3422	Europe	1
A348	Oy-0_IPK	6.13	60.23	Europe	1
A351	Oy-1_UHOH	6.13	60.23	Europe	1
A352	Pa-1_UHOH	13.22	38.07	Europe	1
A353	Pa-2_IPK	13.22	38.07	Europe	1
A354	PAR-3_IPK	-0.25	46.65	Europe	1
A355	PAR-4_IPK	-0.25	46.65	Europe	1
A356	Per-1_IPK	56.3167	58	Europe	1
A358	Petergof_IPK	29	59	Europe	1
A361	PHW-10_IPK	0.40907	51.29273	Europe	1
A362	PHW-13_IPK	0.40907	51.29273	Europe	1
A363	PHW-14_IPK	0.40907	51.29273	Europe	1
A364	PHW-20_IPK	0.40907	51.29273	Europe	1
A365	PHW-22_IPK	-1.7167	51.4167	Europe	1
A366	PHW-26_IPK	-3.8404	50.6728	Europe	1
A367	PHW-28_IPK	-3.5833	50.35	Europe	1
A369	PHW-31_IPK	-3.2	51.4666	Europe	1
A370	PHW-33_IPK	4.5667	52.25	Europe	1
A372	PHW-35_IPK	2.3086	48.6103	Europe	1
A373	PHW-36_IPK	2.3086	48.6103	Europe	1
A374	PHW-37_IPK	2.3086	48.6103	Europe	1
A375	Pi-0_UHOH	10.51	47.04	Europe	1

A376	Pla-0_IPK	2.25	41.5	Europe	1
A377	Pn-0_IPK	-2.96591	48.0653	Europe	1
A379	Pna-10 _UHOH	-86.3253	42.0945	North America	1
A380	Pna-17_IPK	-86.3253	42.0945	North America	1
A382	Pna-17_UHOH	-86.3253	42.0945	North America	1
A383	Po-0_MPI	7.1	50.7167	Europe	1
A384	Pog-0_IPK	-123.206	49.2655	North America	1
A386	Pr-0_IPK	8.60706	50.1448	Europe	1
A390	Pro-0_IPK	-6	43.25	Europe	1
A392	Pt-0_MPI	10.6065	53.476	Europe	1
A394	Pu2-23_IPK	16.36	49.42	Europe	1
A396	Pu2-24_IPK	16.36	49.42	Europe	1
A397	Pu2-7_UHOH	16.36	49.42	Europe	1
A398	Pyl-1_MPI	4.04	47.19	Europe	1
A400	Ra-0_IPK	3.3	46	Europe	1
A402	Rak-2_IPK	16	49	Europe	1
A405	Ren-1_IPK	-1.41	48.5	Europe	1
A407	Ren-11 _UHOH	-1.41	48.5	Europe	1
A409	Rhen-1_IPK	5.56667	51.9667	Europe	1
A410	Ri-0_IPK	-123.137	49.1632	North America	1
A412	RLD-1_MPI	NA	NA	unknown	1
A413	RLD-2_IPK	34.3167	56.25	Europe	1
A414	Rmx-A02_UHOH	-86.511	42.036	North America	1
A417	Rmx-A-180_IPK	-86.511	42.036	North America	1
A418	Rou-0_IPK	1.09849	49.4424	Europe	1
A419	RRS-10_IPK	-86.4251	41.5609	North America	1
A421	RRS-7_IPK	-86.4251	41.5609	North America	1
A423	Rsch-0_MPI	34	56.3	Europe	1
A424	Rsch-4_IPK	34	56.3	Europe	1
A425	Rubeszhnoe_MPI	38.28	49	Europe	1
A426	S96_IPK	NA	NA	unknown	1
A427	Santa Clara_UHOH	-121.16	37.21	North America	1
A428	Sap-0_IPK	14.24	49.49	Europe	1
A431	Sapporo-0_IPK	141.346	43.0553	Asia	1
A433	Sav-0_IPK	15.8833	49.1833	Europe	1
A434	Sav-0_IPK	15.8833	49.1833	Europe	1
A436	Se-0_IPK	-3.53333	38.3333	Europe	1
A438	Sei-0_IPK	11.5614	46.5438	Europe	1
A439	Sg-1_IPK	9.5	47.6667	Europe	1
A440	Sh-0_IPK	10.2144	51.6832	Europe	1
A442	Shahdara_IPK	68.48	38.35	Asia	1
A444	Si-0_IPK	8.02341	50.8738	Europe	1
A445	SLSP-30_IPK	-86.496	43.665	North America	1
A446	Sorbo_MPI	68.48	38.35	Asia	1
A447	Sp-0_IPK	13.181	52.5339	Europe	1
A448	Sq-1_UHOH	-0.6383	51.4083	Europe	1
A449	Sq-8_IPK	-0.6383	51.4083	Europe	1
A451	St-0_IPK	18	59	Europe	1
A453	Ste-0_IPK	11.8558	52.6058	Europe	1
A454	Ste-3_IPK	-86.514	42.03	North America	1
A455	Stw-0_MPI	36	52	Europe	1
A456	Ta-0_IPK	14.5	49.5	Europe	1
A458	TAMM-2_MPI	23.5	60	Europe	1
A459	Tamm-27_UHOH	23.5	60	Europe	1
A460	TDr-1_IPK	14.1386	55.7683	Europe	1
A461	TDr-3_IPK	14.1381	55.7686	Europe	1
A462	Te-0_MPI	23.2982	60.0585	Europe	1
A463	Tha-1_IPK	4.3	52.08	Europe	1
A464	Ting-1_IPK	14.9	56.5	Europe	1
A465	Tiv-1_IPK	12.8	41.96	Europe	1
A467	Tol-0_UHOH	-83.5553	41.6639	North America	1
A468	Tottarp-2_IPK	13.85	55.95	Europe	1
A469	TOU-A1-115_IPK	4.11667	46.6667	Europe	1
A470	TOU-A1-116_IPK	4.11667	46.6667	Europe	1
A472	TOU-A1-43_IPK	4.11667	46.6667	Europe	1
A473	TOU-A1-62_IPK	4.11667	46.6667	Europe	1
A475	TOU-A1-96_IPK	4.11667	46.6667	Europe	1

A476	TOU-C-3_IPK	4.11667	46.6667	Europe	1
A477	TOU-E-11_IPK	4.11667	46.6667	Europe	1
A478	TOU-H-13_IPK	4.11667	46.6667	Europe	1
A479	TOU-I-17_IPK	4.11667	46.6667	Europe	1
A480	TOU-I-2_IPK	4.11667	46.6667	Europe	1
A481	TOU-I-6_IPK	4.11667	46.6667	Europe	1
A482	Ts-1_IPK	2.93056	41.7194	Europe	1
A484	Ts-5_UHOH	2.93056	41.7194	Europe	1
A485	Tscha-1_IPK	9.9042	47.0748	Europe	1
A486	Tsu-0_IPK	136.31	34.43	Asia	1
A487	Tsu-1_MPI	136.31	34.43	Asia	1
A488	Tu-0_UHOH	7.5	45	Europe	1
A489	Tul-0_UHOH	-85.2563	43.2708	North America	1
A490	Ty-0_IPK	-5.23439	56.4278	Europe	1
A492	Udul1-34_IPK	16.6314	49.2771	Europe	1
A493	Uk-1_IPK	7.7667	48.0333	Europe	1
A494	Uk-2_IPK	7.7667	48.0333	Europe	1
A495	Uk-4_UHOH	7.7667	48.0333	Europe	1
A496	UKID48_IPK	-2.7	54.7	Europe	1
A497	UKNW06-059_IPK	-3	54.4	Europe	1
A498	UKNW06-060_IPK	-3	54.4	Europe	1
A499	UKNW06-386_IPK	-3.1	54.6	Europe	1
A500	UKSE06-429_IPK	0.4	51.3	Europe	1
A501	UKSE06-466_IPK	0.4	51.2	Europe	1
A502	UKSE06-482_IPK	0.6	51.2	Europe	1
A503	UKSE06-520_IPK	1.1	51.3	Europe	1
A504	UKSE06-628_IPK	0.4	51.1	Europe	1
A506	Ull2-3_IPK	13.9707	56.0648	Europe	1
A508	Ull2-5_IPK	13.9707	56.0648	Europe	1
A509	Uod-1_UHOH	14.45	48.3	Europe	1
A511	Uod-7_IPK	14.45	48.3	Europe	1
A513	Utrecht_IPK	5.1145	52.0918	Europe	1
A515	Van-0_IPK	-123	49.3	North America	1
A517	Var-2-1_IPK	14.334	55.58	Europe	1
A518	Ven-1_IPK	5.55	52.0333	Europe	1
A519	Wa-1_IPK	21	52.3	Europe	1
A520	Wa-1_UHOH	21	52.3	Europe	1
A522	Wag-3_IPK	5.6666	51.9666	Europe	1
A523	Wag-4_IPK	5.6666	51.9666	Europe	1
A524	Wag-5_IPK	5.6666	51.9666	Europe	1
A525	WAR_IPK	-71.2825	41.7302	North America	1
A526	Wc-2_IPK	10.0667	52.6	Europe	1
A529	Wei-(1)_MPI	8.26	47.25	Europe	1
A531	Wei-0_IPK	8.26	47.25	Europe	1
A533	Wil_MPI	25.3167	54.6833	Europe	1
A536	Wil-2_UHOH	25.3167	54.6833	Europe	1
A537	Wl-0_IPK	10.8134	47.9299	Europe	1
A538	Ws_IPK	30	52.3	Europe	1
A540	Ws-0_UHOH	30	52.3	Europe	1
A541	Ws-2_UHOH	30	52.3	Europe	1
A542	Ws-3_MPI	30	52.3	Europe	1
A543	Wt-3_IPK	9.3	52.3	Europe	1
A544	Wt-5_IPK	9.3	52.3	Europe	1
A546	Yo-0_IPK	-119.35	37.45	North America	1
A549	Zdr-1_UHOH	16.2544	49.3853	Europe	1
A550	Zdr-6_IPK	16.2544	49.3853	Europe	1
A553	Zdrl2-24_IPK	16.2544	49.3853	Europe	1
A554	Zdrl2-25_IPK	16.2544	49.3853	Europe	1
A555	Zü-1_IPK	8.55	47.3667	Europe	1

Lane_12	Lane_11	Lane_10	Lane_9	Lane_8	Lane_7
301.00	331.33	334.33	346.67	349.00	323.67
321.67	343.00	354.00	357.00	354.67	326.00
341.67	359.00	364.00	363.00	358.33	328.33
348.67	366.67	366.67	363.00	355.33	323.33
357.33	380.67	384.67	369.67	362.00	327.00
373.00	389.67	383.00	371.67	361.67	325.00
372.67	386.00	374.67	369.67	360.00	325.67
369.33	379.33	369.00	368.67	362.67	333.00
349.67	364.00	359.33	364.67	372.67	348.00
328.67	353.00	349.00	358.00	357.67	341.00
328.67	353.67	356.33	358.67	360.67	336.67
338.67	358.67	364.33	364.00	366.67	336.33
342.00	363.33	366.00	369.67	363.33	331.67
336.33	364.33	367.67	365.00	359.33	325.00
350.33	371.33	373.67	377.67	369.67	338.33
353.33	374.00	367.67	373.67	369.67	343.00
348.00	368.67	366.33	358.67	356.67	343.67
353.67	367.33	367.33	360.00	353.00	344.67
353.33	370.33	364.33	358.33	349.00	338.67
362.00	373.33	371.33	357.67	354.67	340.67
361.33	370.00	366.00	358.33	354.00	337.67
348.33	365.67	360.33	358.33	356.33	346.00
351.33	370.33	374.67	369.00	373.00	355.67
355.67	378.67	385.00	381.67	380.00	354.33
352.33	377.00	384.33	387.00	380.67	345.00
358.00	381.67	391.33	389.00	388.67	354.00
350.00	380.67	380.33	394.67	380.33	350.33
345.00	374.33	378.67	388.33	381.67	354.00
336.67	367.33	370.00	381.67	374.67	349.67
335.67	364.00	372.00	375.67	370.67	343.33
331.33	361.33	363.67	370.33	364.33	337.67
326.33	353.33	355.67	362.67	355.67	335.67

Lane_6	Lane_5	Lane_4	Lane_3	Lane_2	Lane_1
221.33	226.00	221.33	212.33	204.67	182.67
215.33	216.00	210.67	201.33	194.00	188.67
216.67	219.67	215.67	205.67	199.00	192.33
214.67	221.33	217.33	207.33	201.00	185.33
210.67	217.67	214.00	204.67	197.67	187.00
208.00	212.00	208.33	202.00	193.67	188.33
205.67	210.00	207.67	200.00	192.33	181.67
201.67	208.67	207.00	199.67	192.33	180.00
197.33	205.33	204.67	196.33	189.67	116.00
195.00	204.67	202.67	194.00	186.00	172.00
194.33	204.00	202.00	194.67	187.33	173.33
191.00	199.33	201.00	192.67	188.00	169.33
185.67	190.33	195.67	185.33	182.33	161.67
183.33	189.33	193.33	181.67	177.67	159.67
184.00	195.67	193.00	183.67	179.00	164.67
177.67	198.67	192.33	187.67	181.33	166.33
177.00	193.00	192.33	188.00	179.33	167.33
183.67	192.67	192.67	186.00	177.33	167.00
186.67	196.00	192.33	187.33	182.33	170.33
190.00	198.33	196.33	190.33	184.33	171.67
194.67	197.67	196.67	191.00	183.67	172.33
196.00	194.00	195.67	188.33	181.33	169.33
202.33	200.00	197.33	190.00	184.33	174.33
207.33	201.00	198.33	193.00	187.67	175.67
215.00	201.33	201.00	194.33	189.00	177.67
217.33	200.33	202.33	194.67	190.67	177.67
218.00	201.33	204.67	194.00	192.00	178.33
220.33	204.67	205.67	196.33	194.67	181.00
223.00	206.67	207.00	199.00	196.00	183.67
221.67	204.33	202.00	200.33	192.00	180.67
231.00	205.67	208.67	199.00	200.00	184.33
234.33	212.33	211.67	202.00	201.67	191.00

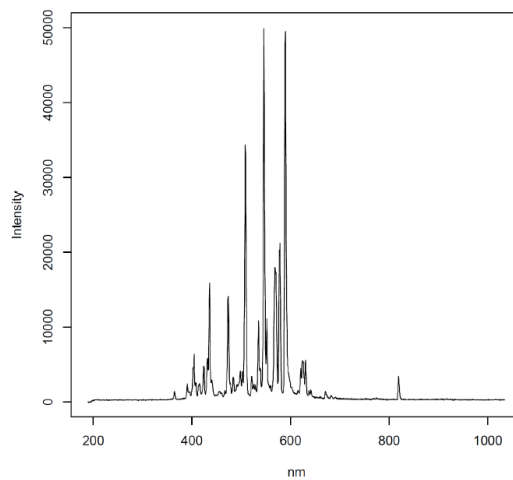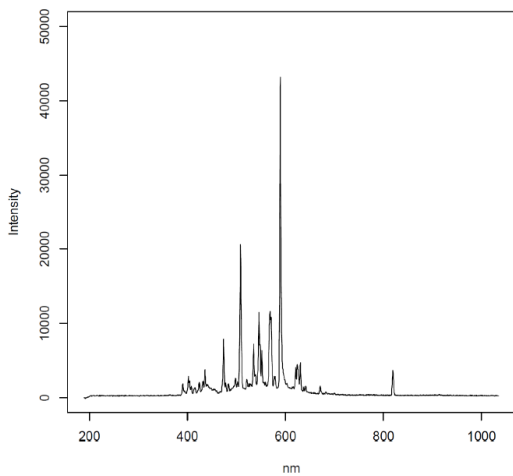**Metal−halide + LED**

Metal−halide

Figure S 1_1: Heatmap of the maximal light intensity in the left (with LED) and right side (w/o LED) of the IPK automated non-invasive plant phenotyping system for small plants. Every carrier position was measured by three independent sensors and mean values are plotted. Light spectra were recorded with 3 ms exposure time.

Table S 1_2: Full list of marker IAP trait association fluctuating light specific GO filter for "photo" or "light"; DAG means days after germination, minor allele frequency (maf), phenotypic variance explained (PVE) by the SNP, short description of the annotated gene identifiers, and R_n was the repeatability

Trait	DAG	Chr	Pos.	*p*-value	maf	PVE %	gene identifier	Short description	R_n
hsv.norm.v.his.bin.01.000.012.vis.top.	5	1	22062493	3.51E-08	0.28	3.97	AT1G59940	response regulator 3	0.59
norm.his.bin.03.25.38.fluor.top.	5	1	27937725	1.23E-10	0.10	2.36	AT1G74310	heat shock protein 101	0.59
hsv.norm.h.his.bin.03.025.038.vis.top.	5	2	11218900	5.95E-09	0.37	2.18	AT2G26350	peroxin 10	0.65
hsv.norm.s.his.bin.12.140.153.vis.top.	5	3	5271626	1.38E-07	0.26	0.16	AT3G15570	Phototropic-responsive NPH3 family protein	0.81
hsv.norm.s.his.bin.08.089.102.vis.top.	5	4	6099811	3.38E-09	0.16	2.54	AT4G09650	ATP synthase delta-subunit gene	0.77
circumcircle.d.geom.trait.fluor.top.px.	5	4	12569624	3.20E-10	0.07	2.83	AT4G24230	acyl-CoA-binding domain 3	0.63
circumcircle.d.z.correct.geom.trait.fluor.top.mm.	5	4	12569624	3.20E-10	0.07	2.83	AT4G24230	acyl-CoA-binding domain 3	0.63
hull.pc1.geom.trait.fluor.top.px.	5	4	12569624	4.57E-11	0.07	2.81	AT4G24230	acyl-CoA-binding domain 3	0.64
hull.pc1.geom.trait.vis.top.px.	5	4	12569624	3.49E-09	0.07	2.71	AT4G24230	acyl-CoA-binding domain 3	0.62
hull.pc1.z.correct.geom.trait.fluor.top.mm.	5	4	12569624	4.57E-11	0.07	2.81	AT4G24230	acyl-CoA-binding domain 3	0.64
hull.pc1.z.correct.geom.trait.vis.top.mm.	5	4	12569624	3.49E-09	0.07	2.71	AT4G24230	acyl-CoA-binding domain 3	0.62
cog.avg.distance.to.center.geom.trait.vis.top.px.	5	5	21416265	1.05E-07	0.17	0.60	AT5G52840	NADH-ubiquinone oxidoreductase-related	0.62
hsv.norm.h.his.bin.06.063.076.vis.top.	5	5	25454980	8.27E-11	0.05	4.38	AT5G63590	flavonol synthase 3	0.75
hsv.norm.h.his.bin.06.063.076.vis.top.	5	5	25454980	8.27E-11	0.05	4.38	AT5G63570	glutamate-1-semialdehyde-2	0.75
norm.his.phenol.bin.18.216.229.fluor.top.	5	5	25454980	2.33E-07	0.05	3.13	AT5G63590	flavonol synthase 3	0.86
norm.his.phenol.bin.18.216.229.fluor.top.	5	5	25454980	2.33E-07	0.05	3.13	AT5G63570	glutamate-1-semialdehyde-2	0.86
lab.a.skewness.vis.top.	6	1	4890205	1.50E-07	0.45	2.16	AT1G14320	Ribosomal protein L16p/L10e family protein	0.82
hsv.norm.s.his.bin.11.127.140.vis.top.	6	1	27936667	1.27E-07	0.09	2.26	AT1G74310	heat shock protein 101	0.86
circumcircle.d.geom.trait.fluor.top.px.	6	3	21827353	1.13E-07	0.37	1.17	AT3G59060	phytochrome interacting factor 3-like 6	0.85
circumcircle.d.z.correct.geom.trait.fluor.top.mm.	6	3	21827353	1.13E-07	0.37	1.17	AT3G59060	phytochrome interacting factor 3-like 6	0.85
hull.pc1.geom.trait.fluor.top.px.	6	3	21827353	4.36E-08	0.37	1.19	AT3G59060	phytochrome interacting factor 3-like 6	0.85
hull.pc1.geom.trait.vis.top.px.	6	3	21827353	8.20E-09	0.37	1.09	AT3G59060	phytochrome interacting factor 3-like 6	0.85
hull.pc1.z.correct.geom.trait.fluor.top.mm.	6	3	21827353	4.36E-08	0.37	1.19	AT3G59060	phytochrome interacting factor 3-like 6	0.85

hull.pc1.z.correct.geom.trait.vis.top.mm.	6	3	21827353	8.20E-09	0.37	1.09	AT3G59060	phytochrome interacting factor 3-like 6	0.85
hsv.h.yellow2green.vis.top.	6	4	18219021	1.55E-07	0.10	0.72	AT4G39100	PHD finger family protein / bromo-adjacent homology (BAH) domain-containing protein	0.77
hsv.norm.h.his.bin.06.063.076.vis.top.	6	5	25454980	3.14E-10	0.05	4.25	AT5G63590	flavonol synthase 3	0.77
hsv.norm.h.his.bin.06.063.076.vis.top.	6	5	25454980	3.14E-10	0.05	4.25	AT5G63570	glutamate-1-semialdehyde-2	0.77
hull.length.geom.trait.fluor.top.	7	1	1586415	4.80E-08	0.26	0.75	AT1G05385	photosystem II 11 kDa protein-related	0.82
hull.length.z.correct.geom.trait.fluor.top.mm.	7	1	1586415	4.80E-08	0.26	0.75	AT1G05385	photosystem II 11 kDa protein-related	0.82
hsv.norm.v.his.bin.13.153.165.vis.top.	7	1	26744556	2.84E-10	0.27	2.44	AT1G70940	Auxin efflux carrier family protein	0.79
hsv.s.skewness.vis.top.	7	4	248308	1.36E-07	0.46	2.02	AT4G00585		0.79
hsv.norm.v.his.bin.03.025.038.vis.top.	7	4	461470	8.24E-08	0.45	2.66	AT4G01050	thylakoid rhodanese-like	0.79
hsv.h.his.v.avg.bin.07.076.089.vis.top.	7	4	6103010	6.69E-11	0.26	2.33	AT4G09650	ATP synthase delta-subunit gene	0.83
hsv.norm.v.his.bin.05.051.063.vis.top.	7	4	10834608	9.06E-11	0.19	1.85	AT4G19990	FAR1-related sequence 1	0.77
hsv.norm.s.his.bin.09.102.114.vis.top.	7	4	15874961	1.32E-07	0.35	1.39	AT4G32890	GATA transcription factor 9	0.86
hsv.s.skewness.vis.top.	7	5	6004834	2.15E-07	0.28	0.73	AT5G18170	glutamate dehydrogenase 1	0.79
lab.l.skewness.vis.top.	7	5	10525671	6.84E-09	0.34	1.80	AT5G28530	FAR1-related sequence 10	0.82
hsv.norm.s.his.bin.12.140.153.vis.top.	12	1	4947128	2.11E-09	0.17	2.80	AT1G14450	NADH dehydrogenase (ubiquinone)s	0.86
hull.pc2.geom.trait.fluor.top.px.	12	1	17715329	8.23E-08	0.20	2.96	AT1G48030	mitochondrial lipoamide dehydrogenase 1	0.88
hull.pc2.geom.trait.vis.top.px.	12	1	17715329	1.91E-08	0.20	3.09	AT1G48030	mitochondrial lipoamide dehydrogenase 1	0.88
hull.pc2.z.correct.geom.trait.fluor.top.mm.	12	1	17715329	8.23E-08	0.20	2.96	AT1G48030	mitochondrial lipoamide dehydrogenase 1	0.88
hull.pc2.z.correct.geom.trait.vis.top.mm.	12	1	17715329	1.91E-08	0.20	3.09	AT1G48030	mitochondrial lipoamide dehydrogenase 1	0.88
hsv.h.skewness.fluor.top.	12	1	26143221	7.31E-09	0.08	2.90	AT1G69530	expansin A1	0.71
hsv.v.skewness.vis.top.	12	1	26228048	2.64E-09	0.37	1.05	AT1G69720	heme oxygenase 3	0.88
lab.l.skewness.vis.top.	12	1	26228048	7.13E-08	0.37	1.51	AT1G69720	heme oxygenase 3	0.88
hsv.v.skewness.vis.top.	12	2	7713722	1.08E-09	0.29	3.88	AT2G17750	NEP-interacting protein 1	0.88
lab.l.skewness.vis.top.	12	2	7713722	7.77E-10	0.29	3.21	AT2G17750	NEP-interacting protein 1	0.88
hsv.v.his.h.avg.bin.08.089.102.vis.top.	12	2	17270449	1.51E-07	0.40	6.15	AT2G41430	dehydration-induced protein (ERD15)	0.76
hsv.s.stddev.vis.top.	12	3	17129797	5.52E-08	0.12	1.91	AT3G46510	plant U-box 13	0.83
hsv.s.stddev.fluor.top.	12	3	18200160	9.20E-08	0.38	1.40	AT3G49110	peroxidase CA	0.83
hull.pc2.geom.trait.fluor.top.px.	12	3	21232451	5.18E-09	0.43	5.13	AT3G57390	AGAMOUS-like 18	0.88

|								
hull.pc2.geom.trait.vis.top.px.	12	3	21232451	9.57E-10	0.43	4.81	AT3G57390	AGAMOUS-like 18	0.88
hull.pc2.z.correct.geom.trait.fluor.top.mm.	12	3	21232451	5.18E-09	0.43	5.13	AT3G57390	AGAMOUS-like 18	0.88
hull.pc2.z.correct.geom.trait.vis.top.mm.	12	3	21232451	9.57E-10	0.43	4.81	AT3G57390	AGAMOUS-like 18	0.88
norm.his.bin.04.38.51.fluor.top.	12	4	10168474	3.29E-08	0.15	1.81	AT4G18390	TEOSINTE BRANCHED 1	0.86
lab.a.skewness.fluor.top.	12	5	4602080	2.22E-07	0.28	1.19	AT5G14250	Proteasome component (PCI) domain protein	0.88
norm.his.bin.04.38.51.fluor.top.	12	5	4951797	2.98E-09	0.23	2.79	AT5G15250	FTSH protease 6	0.86
hsv.s.stddev.fluor.top.	12	5	8442557	1.79E-07	0.38	2.02	AT5G24655	response to low sulfur 4	0.83
lab.a.skewness.fluor.top.	12	5	14012328	2.45E-08	0.20	1.36	AT5G35840	phytochrome C	0.88
hsv.norm.s.his.bin.05.051.063.vis.top.	13	1	22455428	1.26E-07	0.08	1.33	AT1G60980	gibberellin 20-oxidase 4	0.72
norm.his.bin.03.25.38.fluor.top.	13	2	18427267	1.72E-09	0.23	1.54	AT2G44680	casein kinase II beta subunit 4	0.81
hsv.v.his.h.avg.bin.08.089.102.vis.top.	13	2	19034511	2.42E-08	0.06	2.99	AT2G46370	Auxin-responsive GH3 family protein	0.82
hsv.v.his.s.avg.bin.05.051.063.vis.top.	13	4	11011073	4.59E-09	0.25	2.33	AT4G20400	JUMONJI 14	0.90
hsv.norm.v.his.bin.05.051.063.vis.top.	15	1	22062493	1.83E-07	0.28	1.51	AT1G59940	response regulator 3	0.89
hsv.s.stddev.vis.top.	15	5	5906720	1.97E-07	0.20	0.66	AT5G17880	disease resistance protein (TIR-NBS-LRR class)	0.85
hsv.norm.v.his.bin.13.153.165.vis.top.	15	5	16901625	1.01E-07	0.40	0.36	AT5G42270	FtsH extracellular protease family	0.90
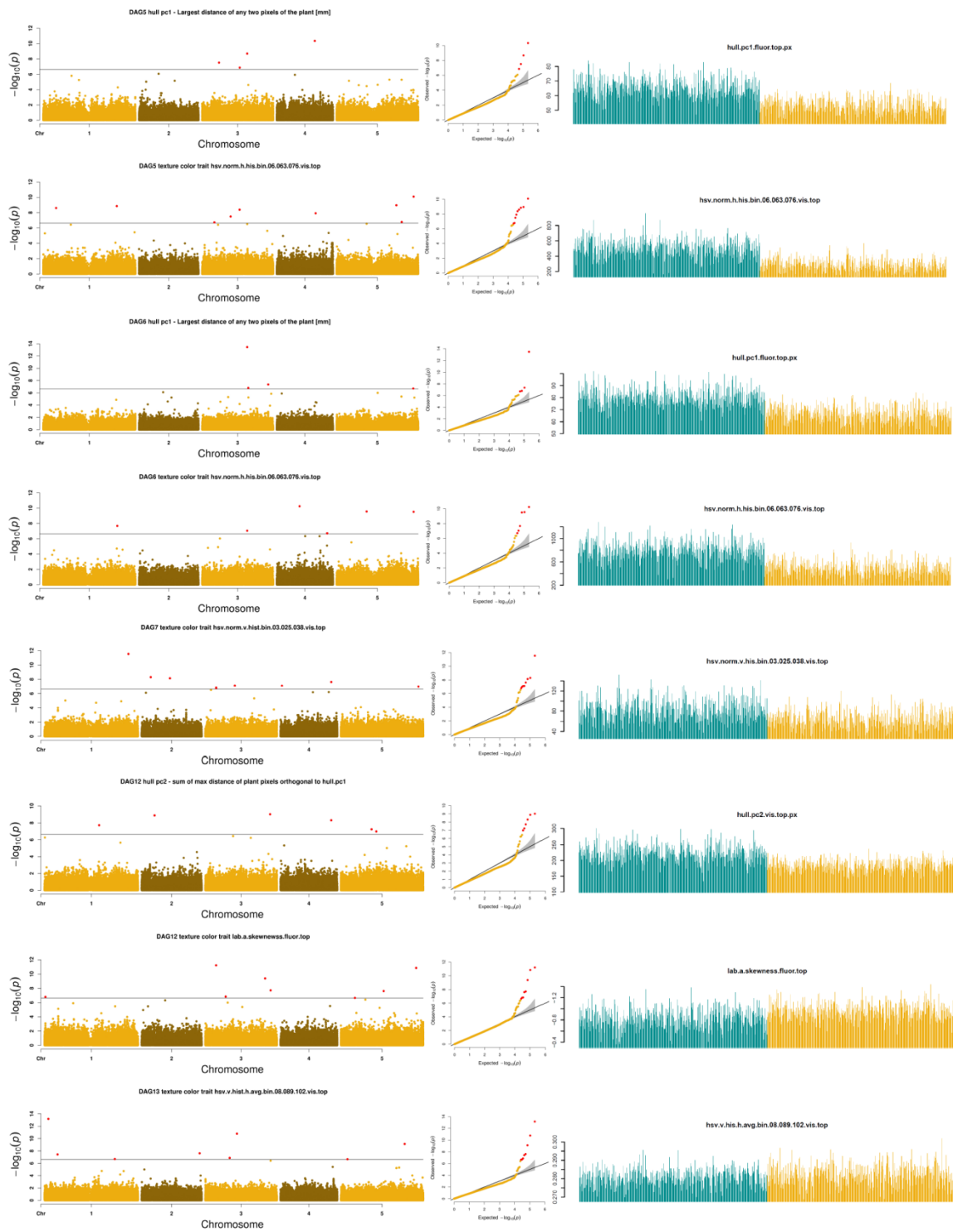
Figure S 1_2: Manhattan plots, Q-Q-plots, and data distribution of the BLUEs of the selected candidate MTAs filtered for GOs with "photo" or "light" from Table 1_2.
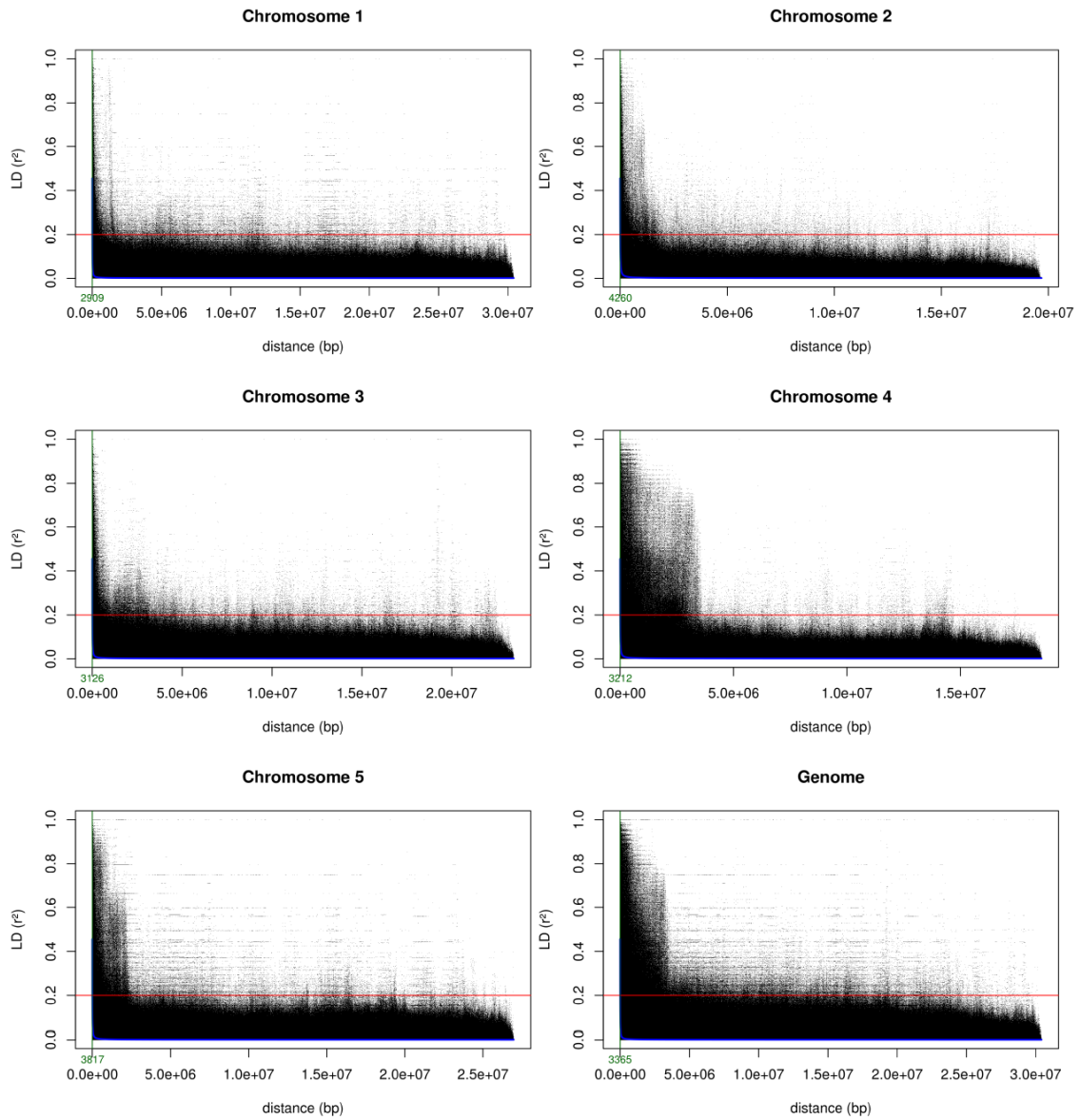
Figure S 1_3: LD decay of each Arabidopsis chromosome: Chr.1, 2909; Chr.2, 4260; Chr.3, 3126; Chr.4, 3212; Chr.5, 3817; Genome, 3365.

Table S 1_3: Full list of gene identifiers annotated in a full LD block with the fluctuating light specific MTAs from the IAP analysis.

Table can be found on the additional supplementary disk.

Table S 1_4: Full list of gene identifiers annotated in full LD block with the fluctuating light specific MTAs from The FluorCam measurement

Table can be found on the additional supplementary disk.

Plant Cultivation Hall – indoor field simulation

Table S 2_1: Full list of the 30 *Zea mays* line phenotyped on the field 2016/17/18. Lines abbreviated with ZEA were provided by the IPK Genebank and the other lines by Muraya *et al.*, 2017. Additional information about origins of lines was taken from publication and public databases (Gerdes *et al.*, 1993; Riedelsheimer *et al.*, 2012; Andorf *et al.*, 2016).

Line	Taxa	Year	Donor/Breeder	Origin	Background	Gene pool
B106	*Zea mays* ssp. *mays*	1996	Iowa Agric & Home Econ Exp Stn	USA, Iowa	Lancaster Sure Crop heterotic group	Dent
B73	*Zea mays* ssp. *mays*	1972	Iowa Agric & Home Econ Exp Stn	USA, Iowa	Stiff Stalk Synthetic	Dent
DK4676A	*Zea mays* ssp. *mays*	unknown	DeKalb-Pfizer Genetics	USA, Illinois	unknown	Dent
DKFBHJ	*Zea mays* ssp. *mays*	unknown	DeKalb-Pfizer Genetics	USA, Illinois	Stiff Stalk Synthetic	Dent
EC136	*Zea mays* ssp. *mays*	unknown	unknown	Spain	unknown	Dent
EC334	*Zea mays* ssp. *mays*	unknown	unknown	Spain	unknown	Dent
EP2	*Zea mays* ssp. *mays*	unknown	unknown	Spain	unknown	Dent
F922	*Zea mays* ssp. *mays*	unknown	unknown	France, Montpellier	unknown	Dent
N22	*Zea mays* ssp. *mays*	unknown	unknown	USA, Nebraska	Krug Yellow Dent	Dent
P074	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk	Dent
P089	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk	Dent
P135	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk	Dent
P148	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk	Dent
PB116	*Zea mays* ssp. *mays*	unknown	unknown	USA	unknown	Dent
PHG50	*Zea mays* ssp. *mays*	1983	Pioneer Hi-Bred International, Inc.	USA, Iowa	Iodent	Dent
PHG84	*Zea mays* ssp. *mays*	1986	Pioneer Hi-Bred International, Inc.	USA, Iowa	Oh07-Midland/Broadbase	Dent
PHT77	*Zea mays* ssp. *mays*	1988	Pioneer Hi-Bred International, Inc.	USA, Iowa	Non-Stiff Stalk Synthetic	Dent
S052	*Zea mays* ssp. *mays*	unknown	unknown	Germany	Non-Stiff Stalk	Dent
W117	*Zea mays* ssp. *mays*	1976	Plant Introduction Stn	USA, Wisconsin	643 x Minn. #13	Dent
W33	*Zea mays* ssp. *mays*	1948	Wisconsin Agric Exp Stn	USA, Wisconsin	W9 x WH	Dent
ZEA 132	*Zea mays* L. subsp. *everta* (Sturtev.) Zhuk. var. *glaucornis* Alef.	1953	Genetikai Osztaly Budapest	unknown	Breeding line	unknown
ZEA 324	*Zea mays* L. subsp. *everta* (Sturtev.) Zhuk. var. *haematornis* Alef.	1968	Agricultural Botanical Garden Bucharest	unknown	Breeding line	unknown
ZEA 332	*Zea mays* L. subsp. *indurata* (Sturtev.) Zhuk. var. *vulgata* Körn.	1967	VIR Leningrad	Soviet Union	Breeding line	unknown
ZEA 3425	*Zea mays* L. subsp. *indurata* (Sturtev.) Zhuk. var. *vulgata* Körn.	2001	Anhaltische Pflanzenzucht GmbH	Germany	Breeding line	unknown
ZEA 3660	*Zea mays* L. subsp. *indentata* (Sturtev.) Zhuk. var. *flavorubra* Körn.	2003	BAZ, Braunschweig Genetic Resources Centre	China, Jilin	Research-, Breeding line	unknown

ZEA 3682	*Zea mays* L. subsp. *indurata* (Sturtev.) Zhuk. var. *vulgata* Körn.	2002	KWS, Bernburg, Dr. Giso Zieger	GDR	Breeding line	unknown
ZEA 391	*Zea mays* L. subsp. *everta* (Sturtev.) Zhuk. var. *oryzoides* Körn.	1989	Redwood City Seed Co., Calif. USA	unknown	Breeding line	unknown
ZEA 392	*Zea mays* L. subsp. *saccharata* (Körn.) Zhuk. var. *flavodulcis* Körn.	1989	Redwood City Seed Co., Calif. USA	unknown	Breeding line	unknown
ZEA 399	*Zea mays* L.	1991	National Institute of Agrobiological Resources Tsukuba	Democratic People's Republic of Korea	Research-, Breeding line	unknown
ZEA 851	*Zea mays* L.	1991	National Institute of Agrobiological Resources Tsukuba	Democratic People's Republic of Korea	Research-, Breeding line	unknown

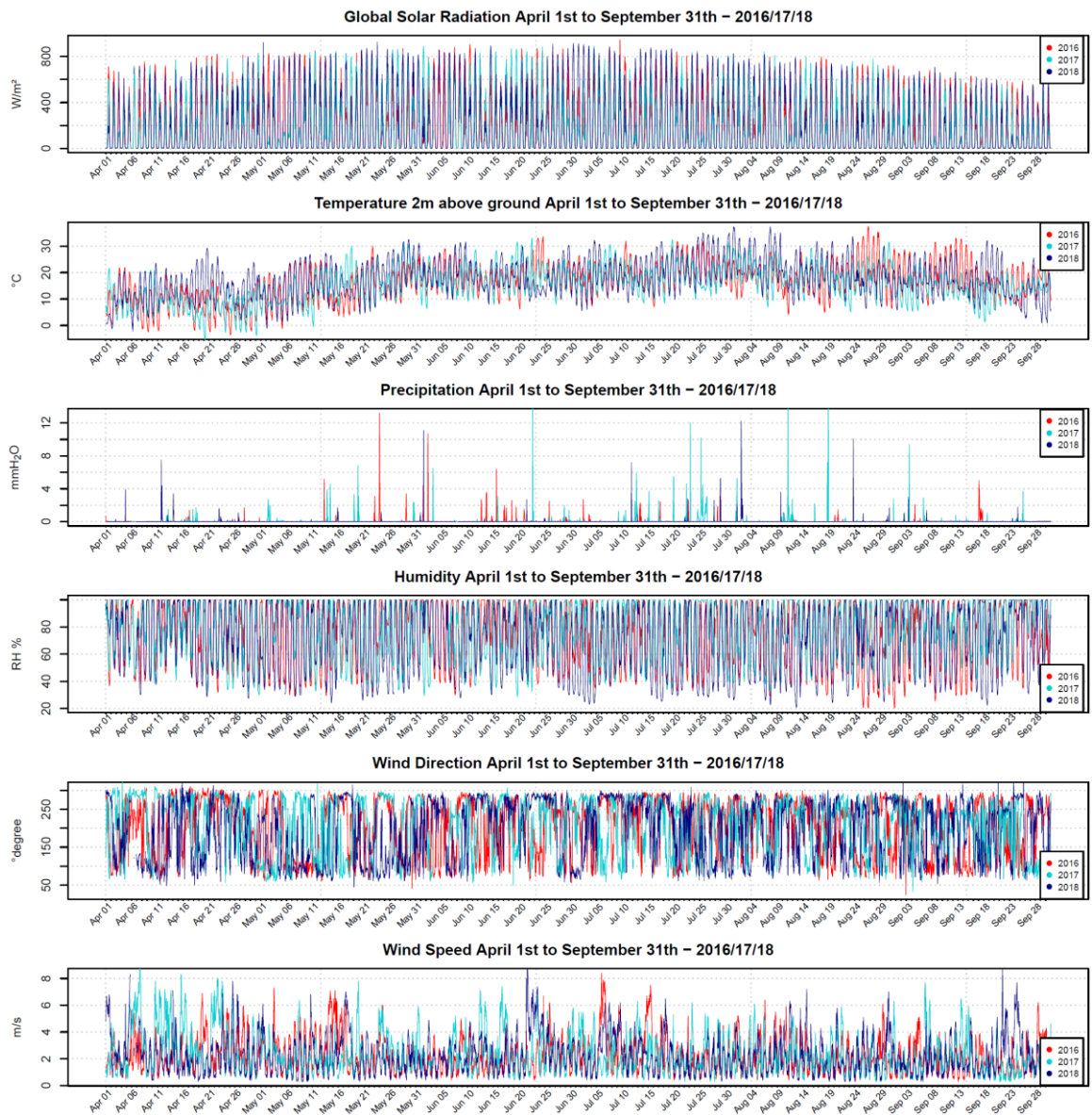**Figure S 2_1**: Local weather data from the IPK weather station of the years 2016/17/18. Parameters are temperature in °C, global solar radiation in W m⁻², precipitation in mmH$_2$O, relative humidity in %, wind direction in °degree with 0° being north, and wind speed in m s⁻¹.

Table S 2_2: Illumination program of the PKH at a normal day. Values for each lighting channel are given in percent of maximum. K stands for constant and R for a linear ramping. LED light channels (W, white; DB, dark blue; R, red; FR, far red), UV, and 4 channels of equal intensity of metal halide lamps. Sin symbolizes an oscillation between the two values following a sinus curve.

Time	LED-W	LED-DB	LED-CB	LED-R	LED-FR	UV(A)	CMH-Gr.1	CMH-Gr.2	CMH-Gr.3	CMH-Gr.4
0:00	0	0	0	0	0	0	0	0	0	0
	K		K	K	K	K				
6:00	0		0	0	0	0				
	R		R	R	R	R	K			
7:00	50		50	50	50	15	100			
	R						K			
7:03	10		10	10	10		100			
	R		R	R	R	R	K			
8:00	50		50	50	50	30	100			
	Sin_50-100		Sin_50-100	Sin_50-100	Sin_50-100	R	K	K		
9:00	50		50	50	50	50	100	100		
	R		R	R	R		K	K		
9:03	30		30	30	30					
	R		R	R	R	R	K	K	K	
10:00	70		70	70	70	60	100	100	100	
	R		R	R	R					
10:03	20		20	20	20					
	Sin_10-50	K	Sin_10-50	Sin_10-50	Sin_10-50	R	K	K	K	
11:00	50	0	50	50	50	70	100	100	100	
	R	R	R	R	R	R				
12:00	75	15	75	8	75	75				
	K	K	K	K	K	K	K	K	K	K
13:00	75	15	75	8	75	75	0	100	100	100
	Sin_10-75	Sin_2-15	Sin_10-75	Sin_1-8	Sin_10-75	K	K	K	K	K
14:00	75	15	75	8	75	75	0	100	100	100
	K	K	K	K	K	K	K	K	K	K
15:00	75	15	75	8	75	75	0	100	100	100
	R	R	R	R	R	R	K	K	K	K
16:00	20	0	20	20	20	70	0	100	100	100
	Sin_10-50	K	Sin_10-50	Sin_10-50	Sin_10-50	R		K	K	K
17:00	70		70	70	70	60		0	100	100
	R		R	R	R	R			K	K
18:00	30		30	30	30	50			0	100
	Sin_50-100		Sin_50-100	Sin_50-100	Sin_50-100	R				K
19:00	50		50	50	50	30				100
	R		R	R	R	R				
20:00	10		10	10	10	15				K
20:00	50		50	50	50	15				0
	R		R	R	R	R				
21:00	0		0	0	0	0				
	K	K	K	K	K	K	K	K	K	K
0:00	0	0	0	0	0	0	0	0	0	0

Table S 2_3: Illumination of the PKH at a sunny day. Values for each lighting channel are given in percent of maximum. K stands for constant and R for a linear ramping. LED light channels (W, white; DB, dark blue; R, red; FR, far red), UV, and 4 channels of equal intensity of metal halide lamps. Sin symbolizes an oscillation between the two values following a sinus curve.

Time	LED-W	LED-DB	LED-CB	LED-R	LED-FR	UV(A)	CMH-Gr.1	CMH-Gr.2	CMH-Gr.3	CMH-Gr.4
0:00	0	0	0	0	0	0	0	0	0	0
	K		K	K	K	K				
6:00	0		0	0	0	0				
	R		R	R	R	R	K			
7:00	50		50	50	50	15	100			
	R						K			
7:03	10		10	10	10		100			
	R		R	R	R	R	K			
8:00	50		50	50	50	30	100			
	Sin_50-90		Sin_50-90	Sin_50-90	Sin_50-90	R	Sin_80-100	K		
9:00	100		100	100	100	50	100	100		
	R		R	R	R		K	K		
9:03	30		30	30	30					
	R		R	R	R	R	K	K	K	
10:00	70		70	70	70	60	100	100	100	
	R		R	R	R					
10:03	20		20	20	20					
	R	K	R	R	R	R	K	K	K	
11:00	50	0	50	50	50	70	100	100	100	
	R	R	R	R	R	R				
12:00	80	20	80	8	80	80				
	K	K	K	K	K	K	K	K	K	K
13:00	100	20	100	8	100	80	0	100	100	100
13:03	80		80		80	80				
	K	K	K	K	K	K	K	K	K	K
14:00	80	20	80	8	80	80	0	100	100	100
	K	R	R	R	R	R	K	K	K	K
15:00	80	15	75	8	75	75	0	100	100	100
	R	R	R	R	R	R	K	K	K	K
16:00	30	0	30	30	30	70	0	100	100	100
	Sin_20-50	K	Sin_20-50	Sin_20-50	Sin_20-50	R		Sin_80-100	Sin_80-100	Sin_80-100
17:00	70		70	70	70	60		0	100	100
	R		R	R	R	R			K	K
18:00	30		30	30	30	50			0	100
18:00	100		100	100	100	R				K
	R		R	R	R					
19:00	50		50	50	50	30				100
	R		R	R	R	R				
20:00	10		10	10	10	15				K
20:00	50		50	50	50	15				0
	R		R	R	R	R				
21:00	0		0	0	0	0				
	K	K	K	K	K	K	K	K	K	K
0:00	0	0	0	0	0	0	0	0	0	0

Table S 2_4: Illumination of the PKH at a cloudy day. Values for each lighting channel are given in percent of maximum. K stands for constant and R for a linear ramping. LED light channels (W, white; DB, dark blue; R, red; FR, far red), UV, and 4 channels of equal intensity of metal halide lamps. Sin symbolizes an oscillation between the two values following a sinus curve.

| Time	LED-W	LED-DB	LED-CB	LED-R	LED-FR	UV(A)	CMH-Gr.1	CMH-Gr.2	CMH-Gr.3	CMH-Gr.4
0:00	0	0	0	0	0	0	0	0	0	0
	K		K	K	K	K				
6:00	0		0	0	0	0				
	R		R	R	R	R				
7:00	15		15	15	15	6				
	Sin_10_30		Sin_10_30	Sin_10_30	Sin_10_30					
8:00										
	Sin_20_50		Sin_20_50	Sin_20_50	Sin_20_50	R				
9:00	45		45	45	45	15				
	R		R	R	R	R				
10:00	55		55	55	55	20				
	Sin_45_70		Sin_45_70	Sin_45_70	Sin_45_70	R				
11:00						25				
	Sin_55_75		Sin_55_75	Sin_55_75	Sin_55_75	R				
12:00						25				
13:00	65		65	65	65	25				
	R		R	R	R	R				
14:00	55		55	55	55	20				
	Sin_60_40		Sin_60_40	Sin_60_40	Sin_60_40	R				
15:00	45		45	45	45	15				
	R		R	R	R	R				
16:00	35		35	35	35	10				
	Sin_40_25		Sin_40_25	Sin_40_25	Sin_40_25	R				
17:00	30		30	30	30	8				
	R		R	R	R	R				
18:00	25		25	25	25	7				
	Sin_30_20		Sin_30_20	Sin_30_20	Sin_30_20	R				
19:00						6				
	Sin_25_10		Sin_25_10	Sin_25_10	Sin_25_10	K				
20:00	15		15	15	15	6				
	R		R	R	R	R				
21:00	0		0	0	0	0	0	0	0	0
	K	K	K	K	K	K	K	K	K	K
0:00	0	0	0	0	0	0	0	0	0	0
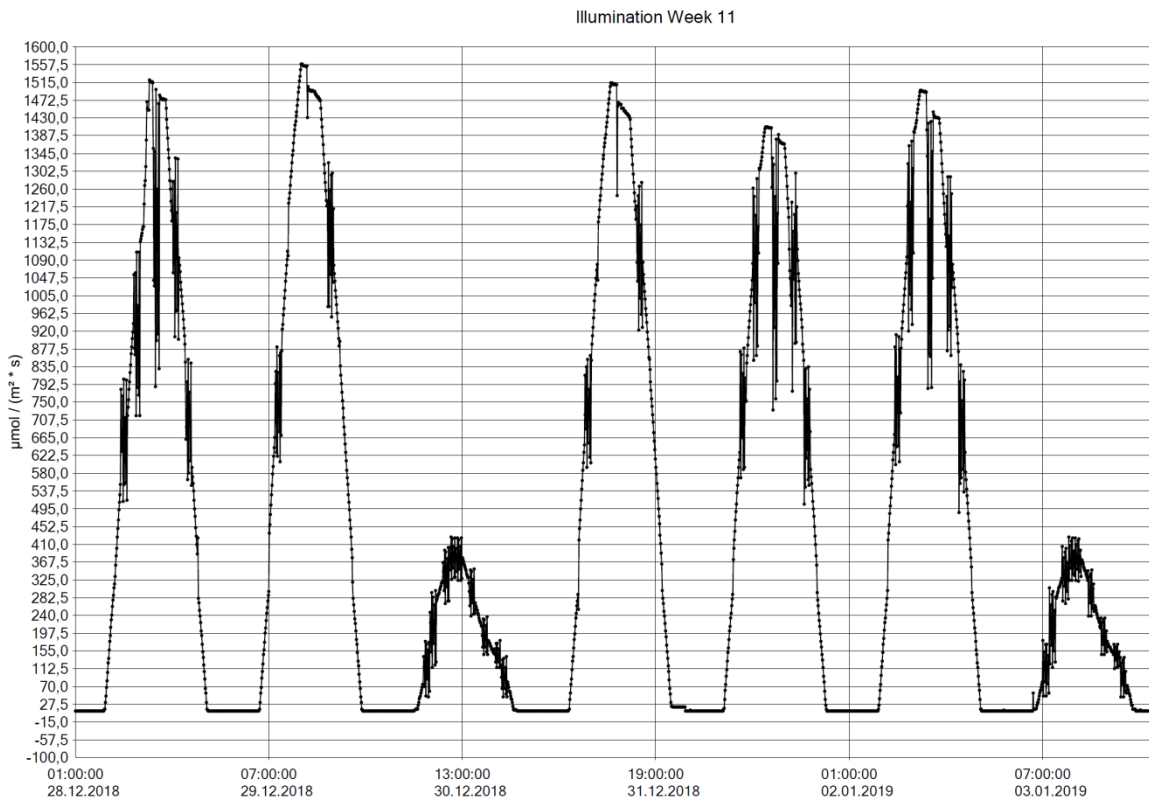
Figure S 2_2: Illumination scheme of week 11 in the PKH. The order of days is normal, sunny, cloudy, sunny, normal, normal, and cloudy.
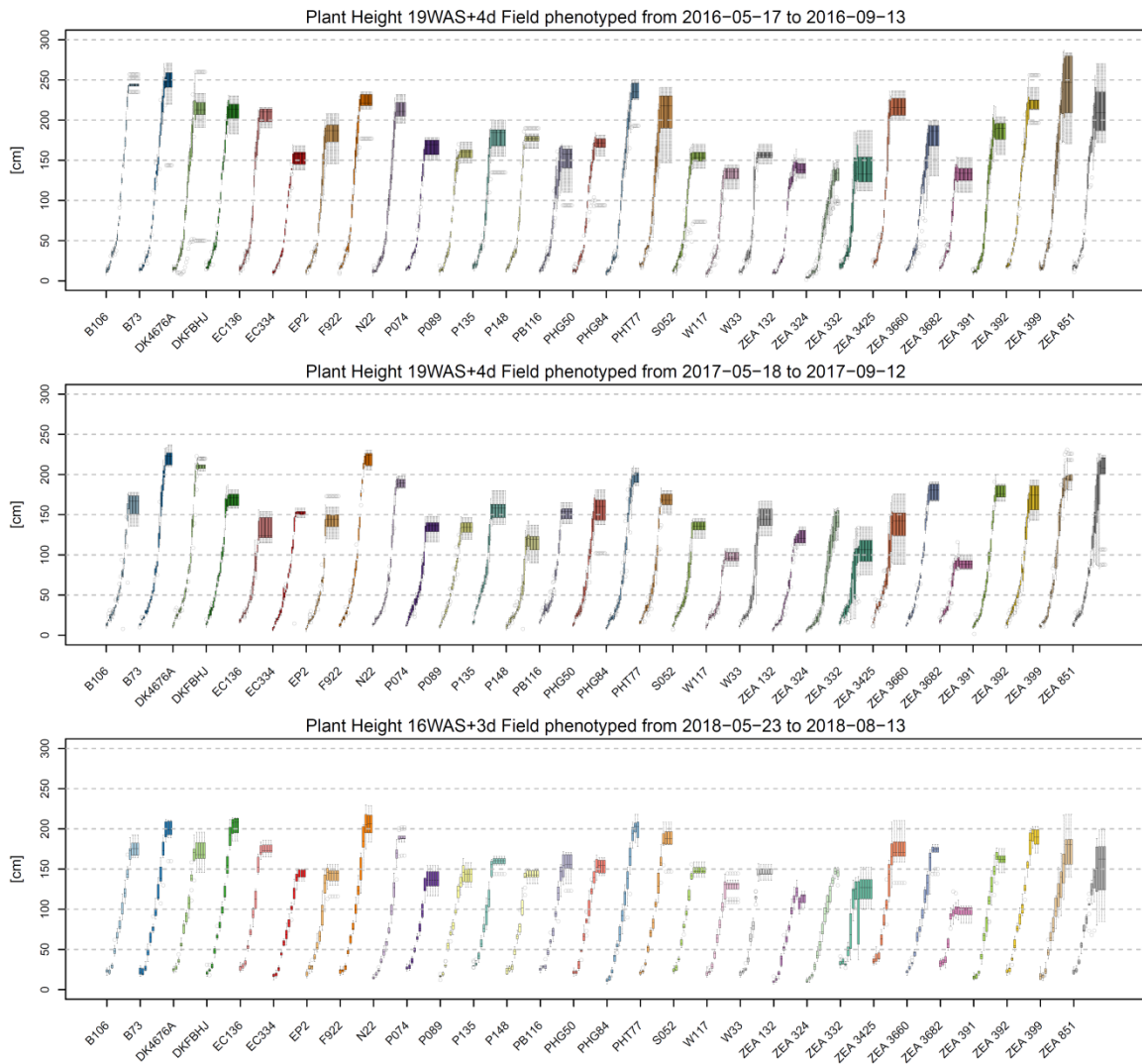
Figure S 2_3: Plant height of the full set of the 30 *Zea mays* inbred phenotyped in the field in 2016, 2017, and 2018. Plant height was measured as the highest point of plant matter above ground.
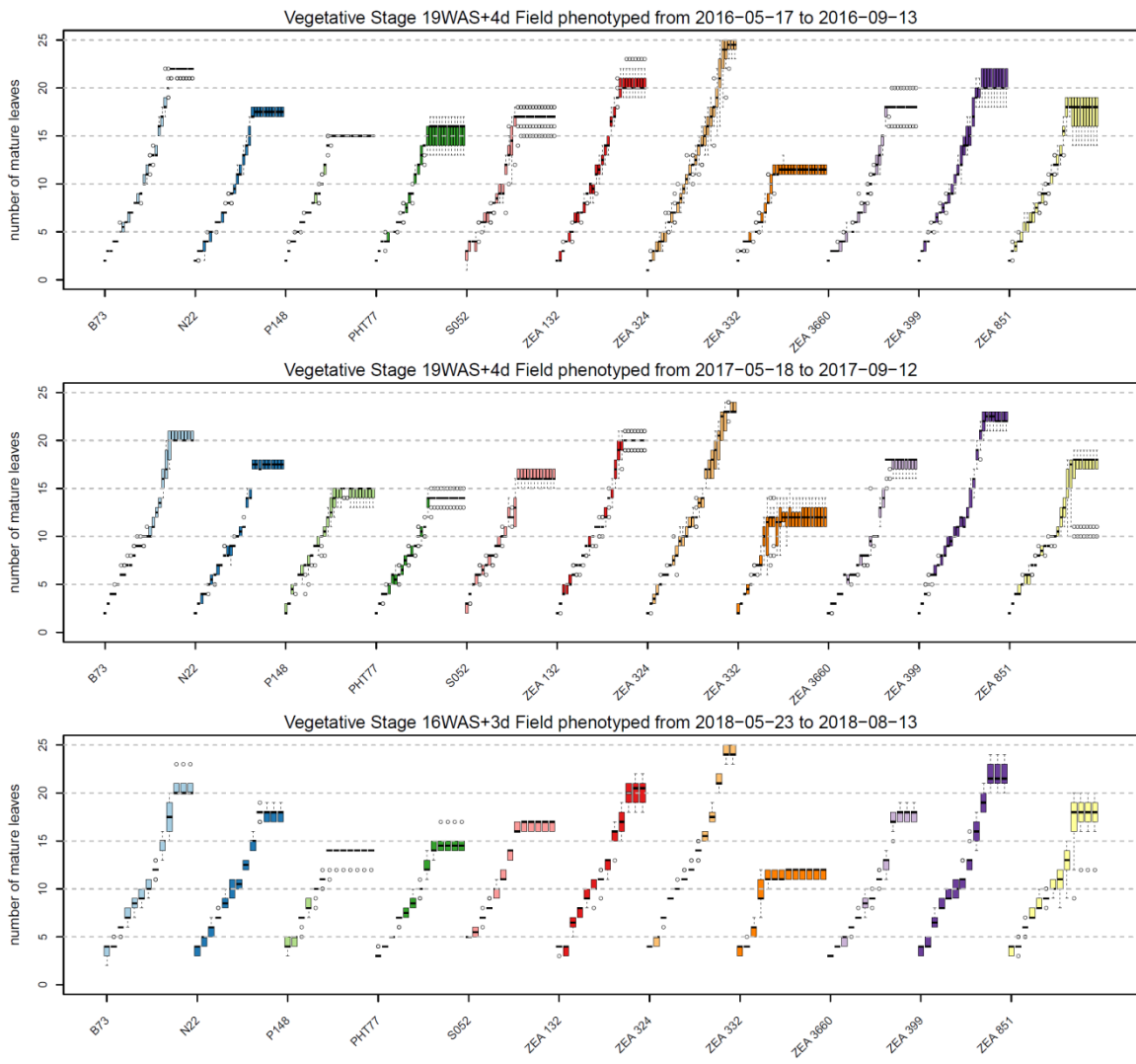
Figure S 2_4: Vegetative Stage of the selected 11 *Zea mays* inbred phenotyped in the field in 2016, 2017, and 2018. Vegetative stage counts mature leaves with a leaf collar.
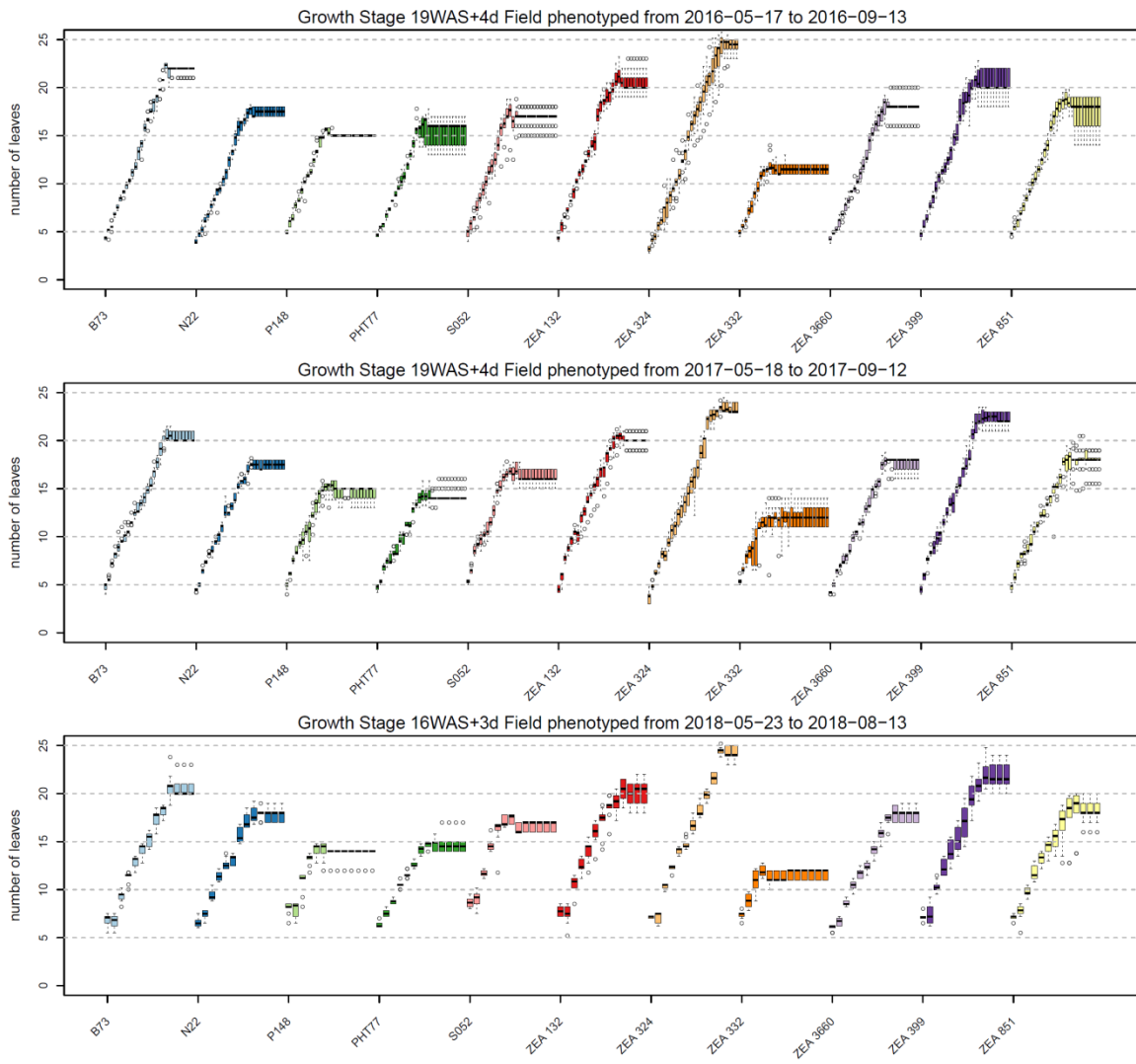
Figure S 2_5: Growth Stage of the selected 11 *Zea mays* inbred phenotyped in the field in 2016, 2017, and 2018. Growth stage counts all leaves of a plant with a fractional-part denoting the leaf length of the last leaf relative to the length of the preceding leaf.
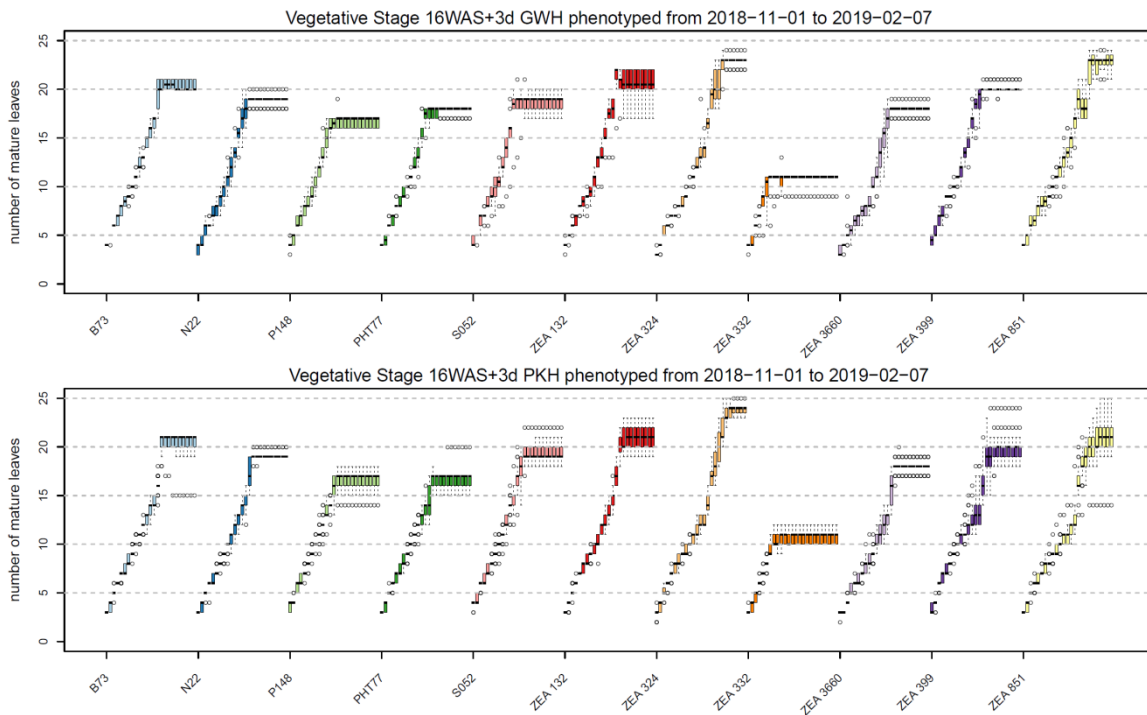
Figure S 2_6: Vegetative Stage of the selected 11 *Zea mays* inbred phenotyped in the PKH and GWH. Vegetative stage counts mature leaves with a leaf collar.
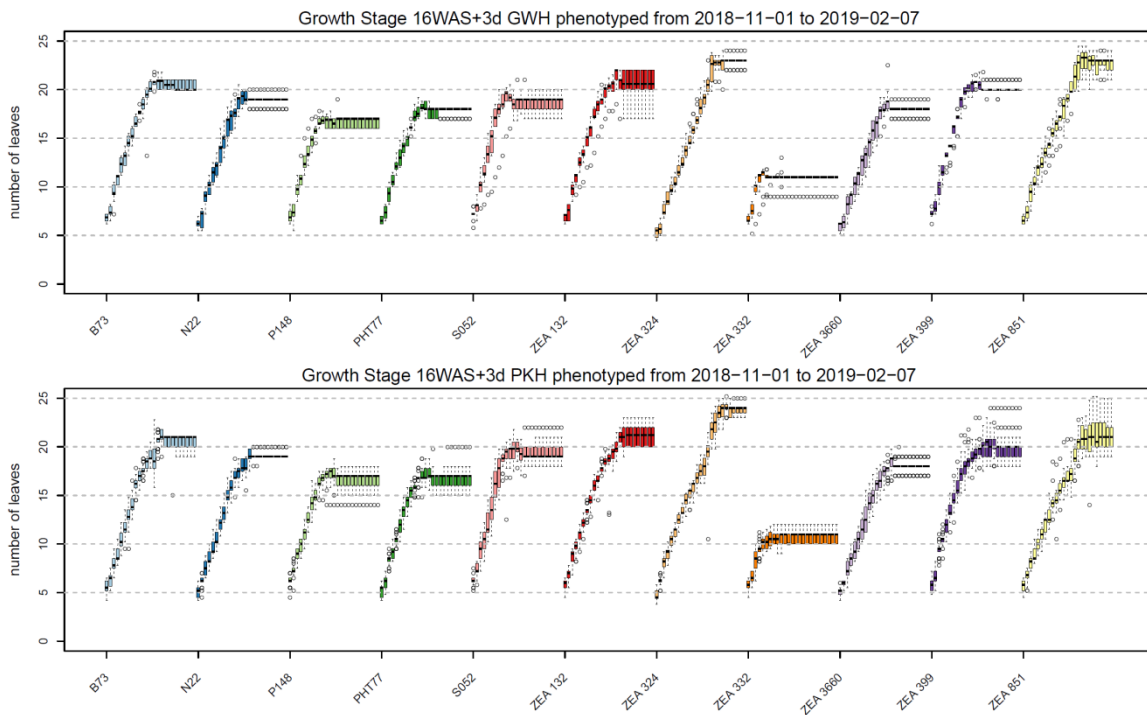
Figure S 2_7: Growth Stage of the selected 11 *Zea mays* inbred phenotyped in the PKH and GWH. Growth stage counts all leaves of a plant with a fractional-part denoting the leaf length of the last leaf relative to the length of the preceding leaf.
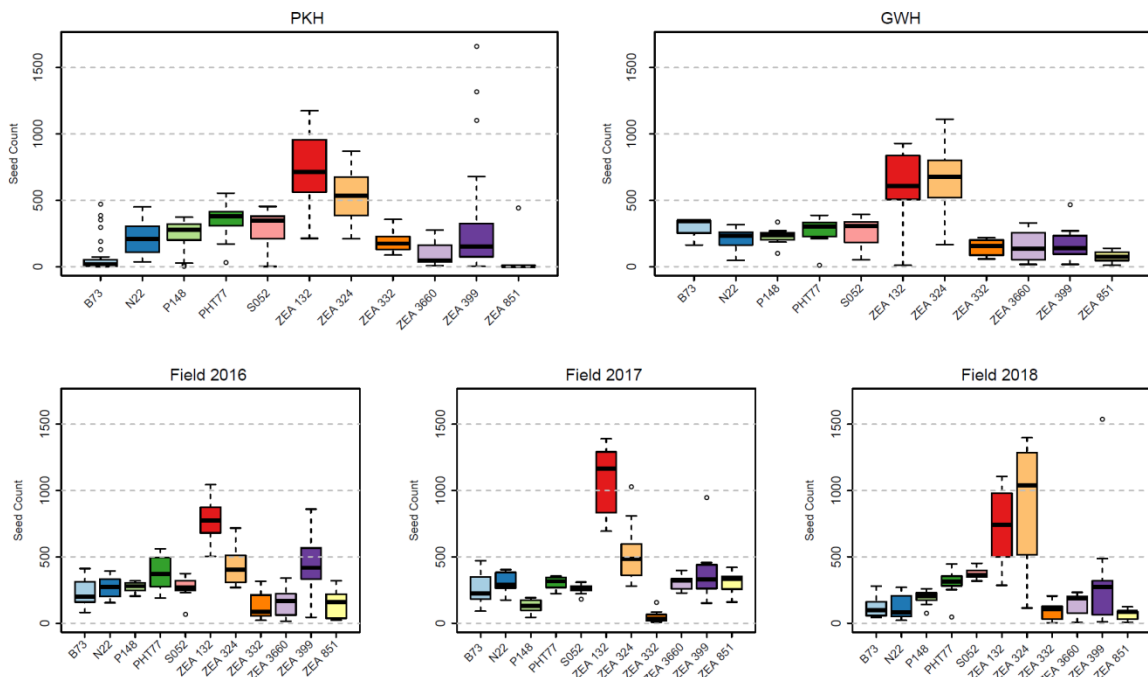
Figure S 2_8: Kernel count of the 11 selected *Zea mays* inbred lines used in the phenotyping experiments PKH (n=25), GWH (n=10), field 2016 (n=10), field 2017 (n=10), and field 2018 (n=10).
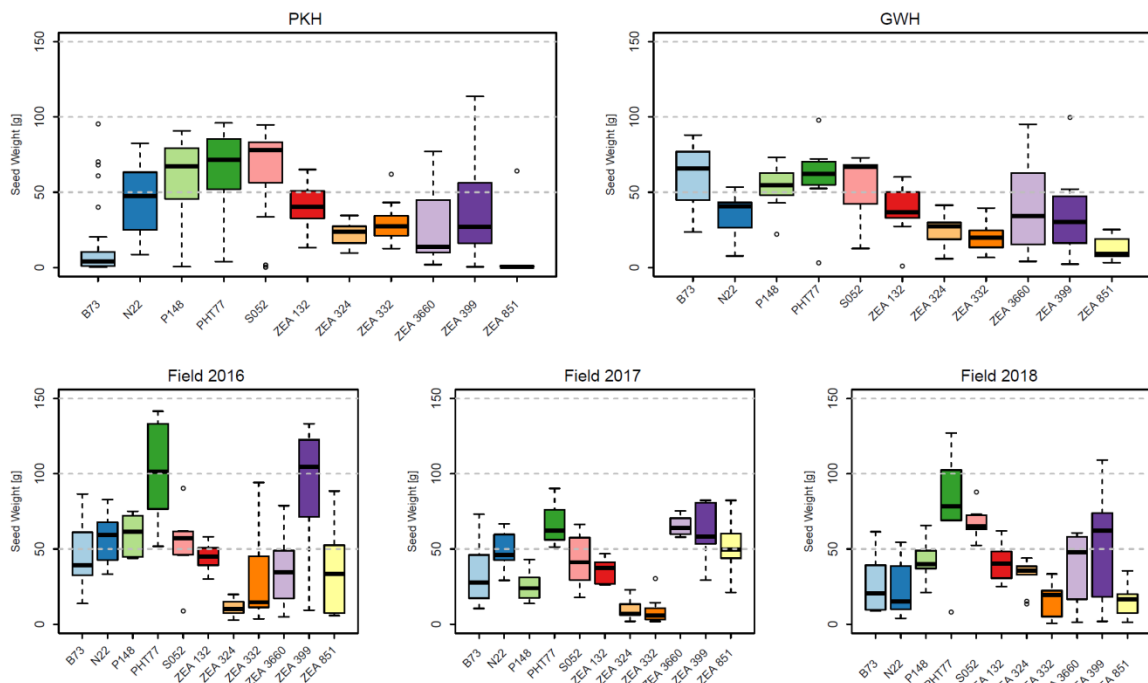
Figure S 2_9: Kernel weights of the 11 selected *Zea mays* inbred lines used in the phenotyping experiments PKH (n=25), GWH (n=10), field 2016 (n=10), field 2017 (n=10), and field 2018 (n=10).
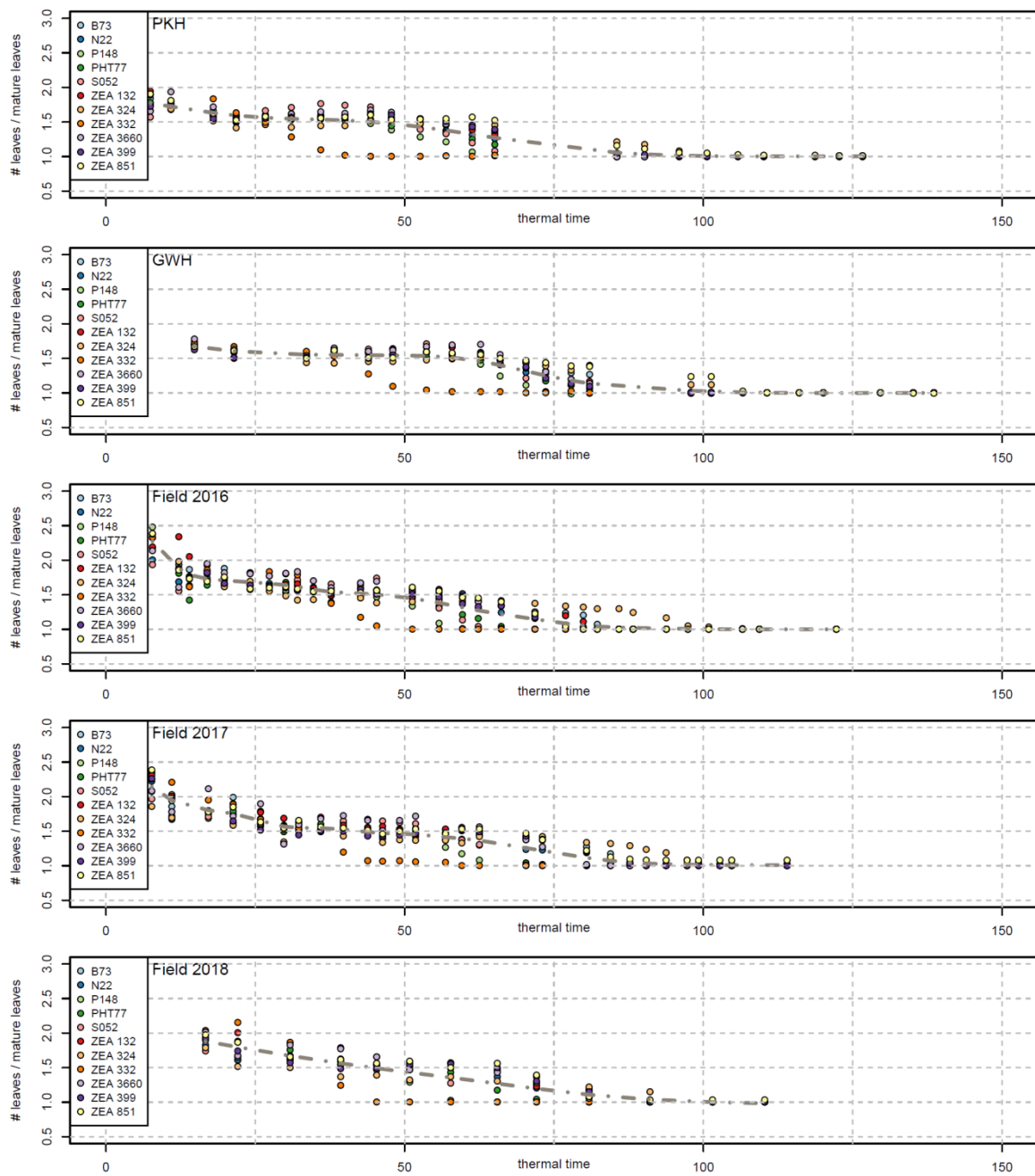
Figure S 2_10: Ratio of the BLUEs of the growth stage to the vegetative stage of the 11 selected *Zea mays* lines plotted against thermal time days. Values of 1 represent VT stage.

Rapid mapping of induced recessive and dominant mutations in maize
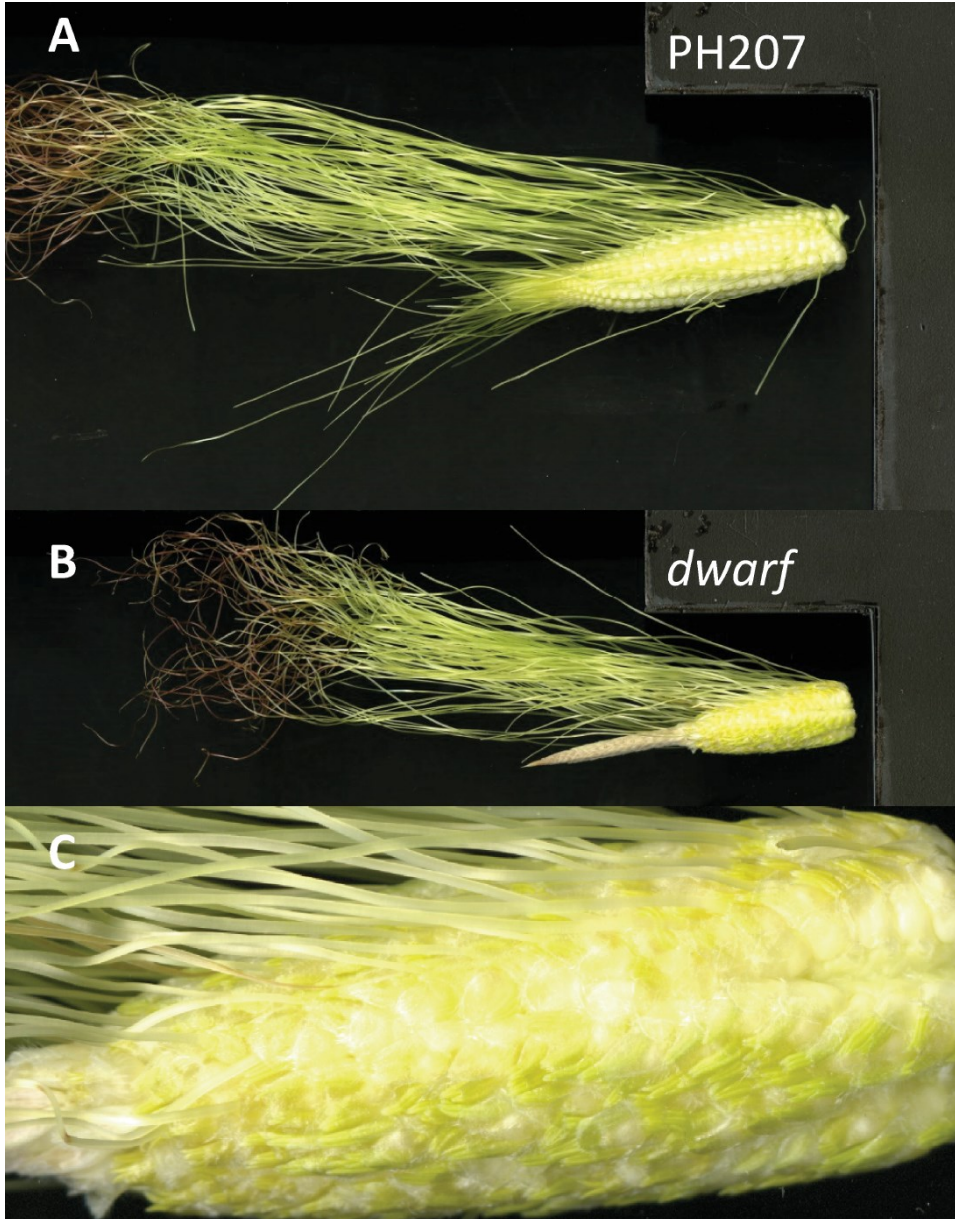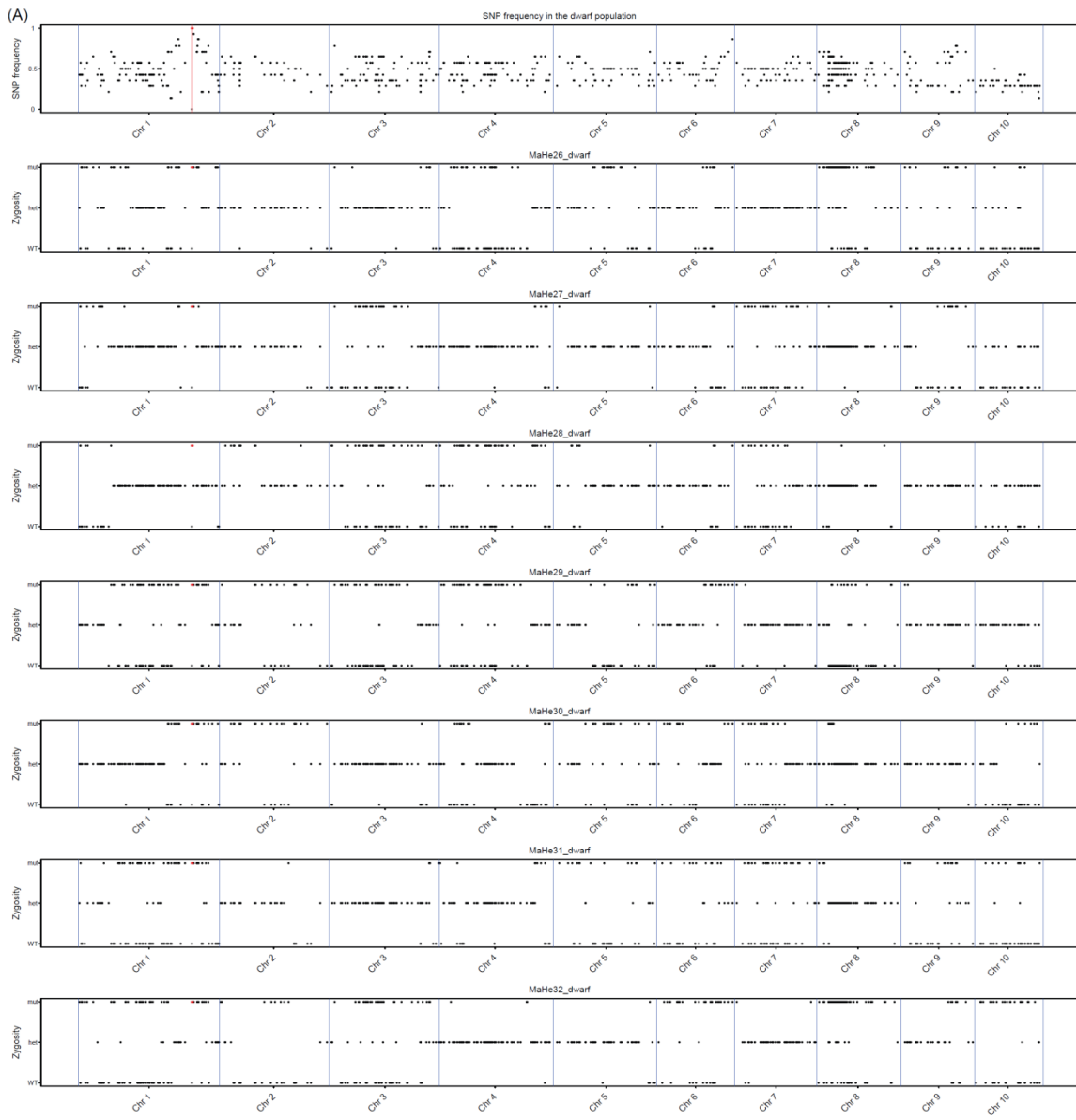
Figure S 3_1: Phenotype of young dwarf and PH207 ears. (**A**) PH207 ear (**B**) dwarf ear (**C**) Magnification of dwarf ear to show anthers next to the kernels
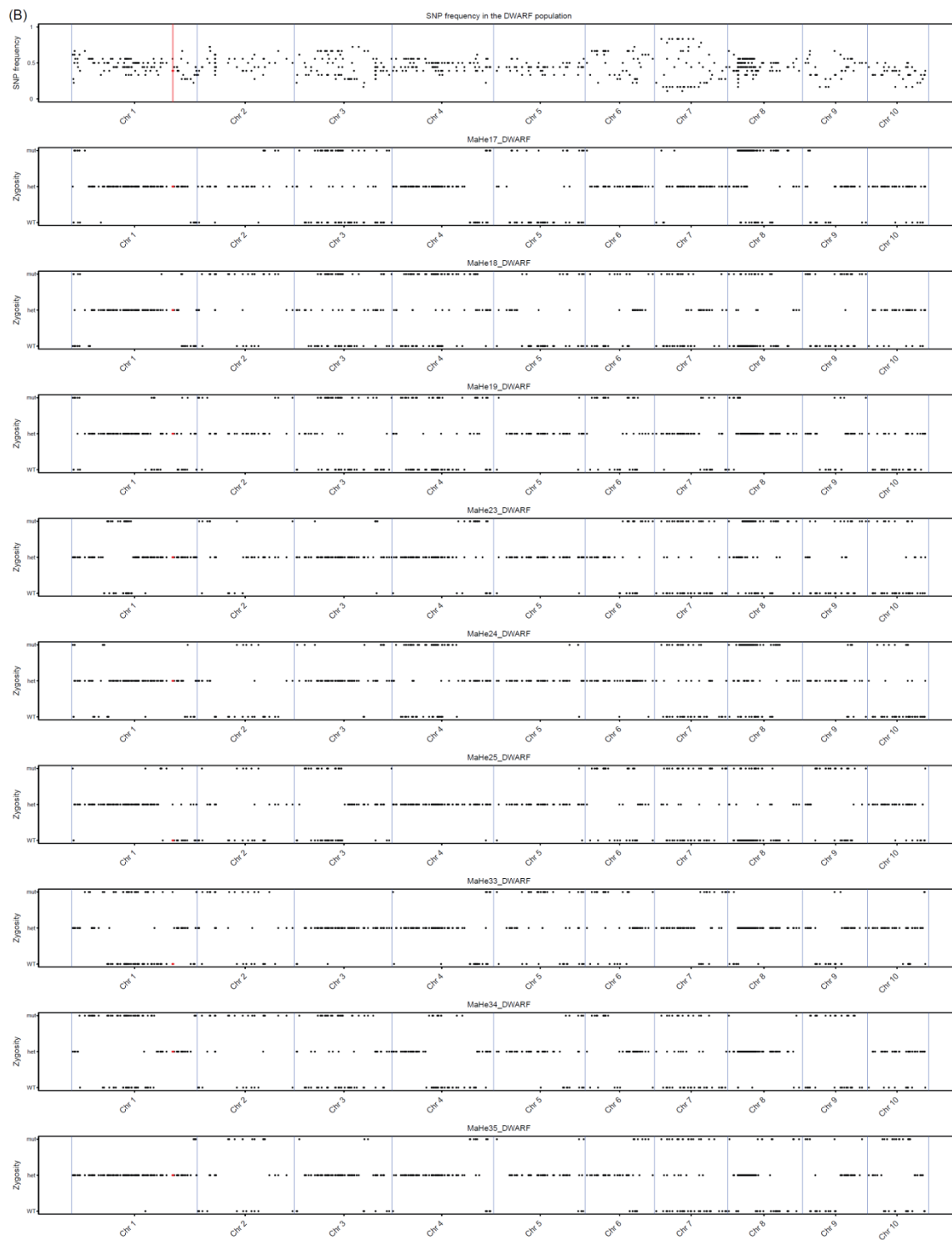
Figure S 3_2: Alignment and strategy view of the re-sequenced *an1* gene in the *dwarf* (1744) and *DWARF* (WT) mutants, B73 (AGPv4; Chr1:244856531-244869541) and PH207 (Chr1:244828512-244841534)
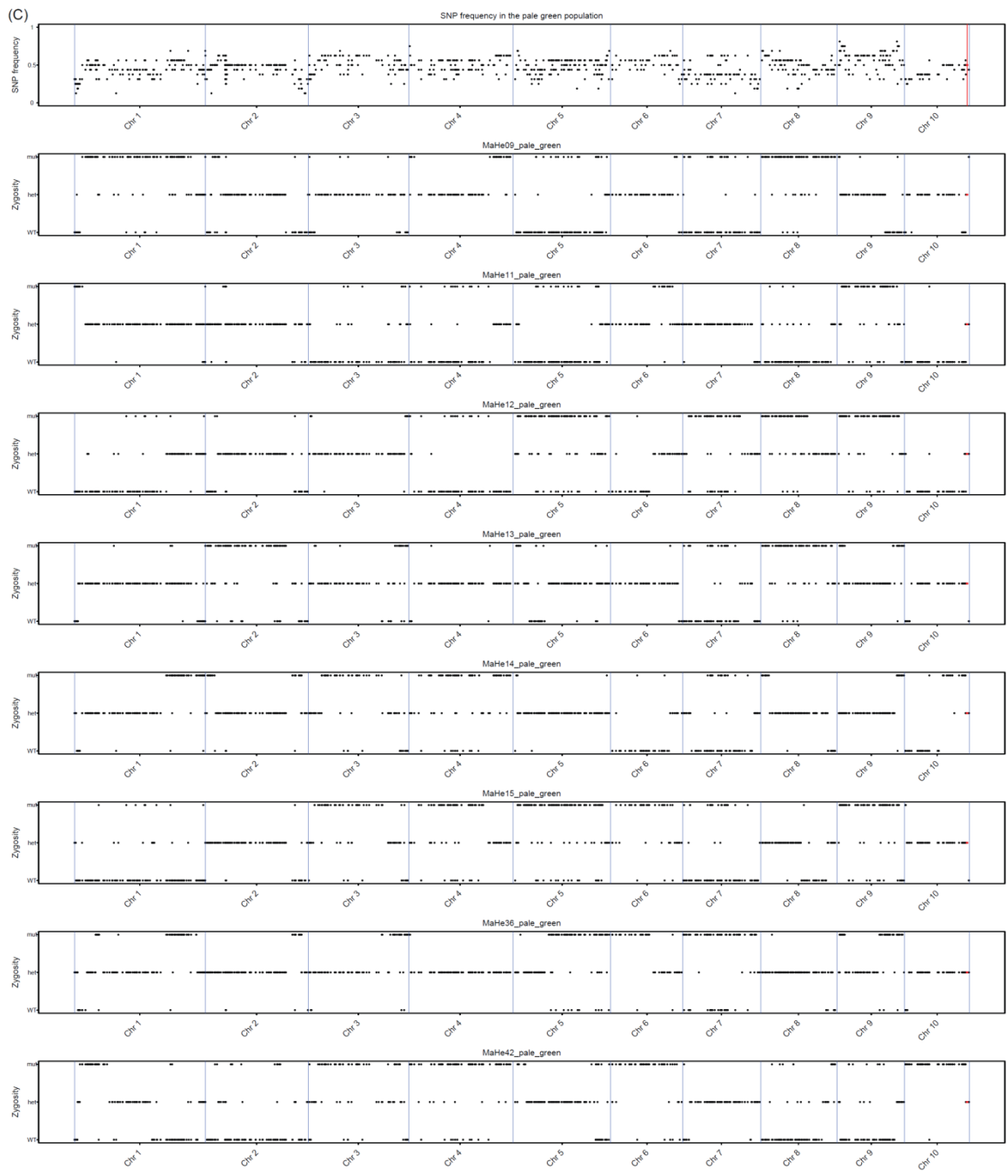
Figure S 3_3: Alignment and strategy view of the re-sequenced *w2* gene in the *pale green* (1754) and *PALE GREEN* (WT) mutants, PH207_w2_flanking and PH207_ws (Chr10:140702976-140735202 and 140707976-140724202)

(A)

SNP frequency in the dwarf population

MaHe26_dwarf

MaHe27_dwarf

MaHe28_dwarf

MaHe29_dwarf

MaHe30_dwarf

MaHe31_dwarf

MaHe32_dwarf

(B)

SNP frequency in the DWARF population

MaHe17_DWARF

MaHe18_DWARF

MaHe19_DWARF

MaHe23_DWARF

MaHe24_DWARF

MaHe25_DWARF

MaHe33_DWARF

MaHe34_DWARF

MaHe35_DWARF

(C)

SNP frequency in the pale green population

MaHe09_pale_green

MaHe11_pale_green

MaHe12_pale_green

MaHe13_pale_green

MaHe14_pale_green

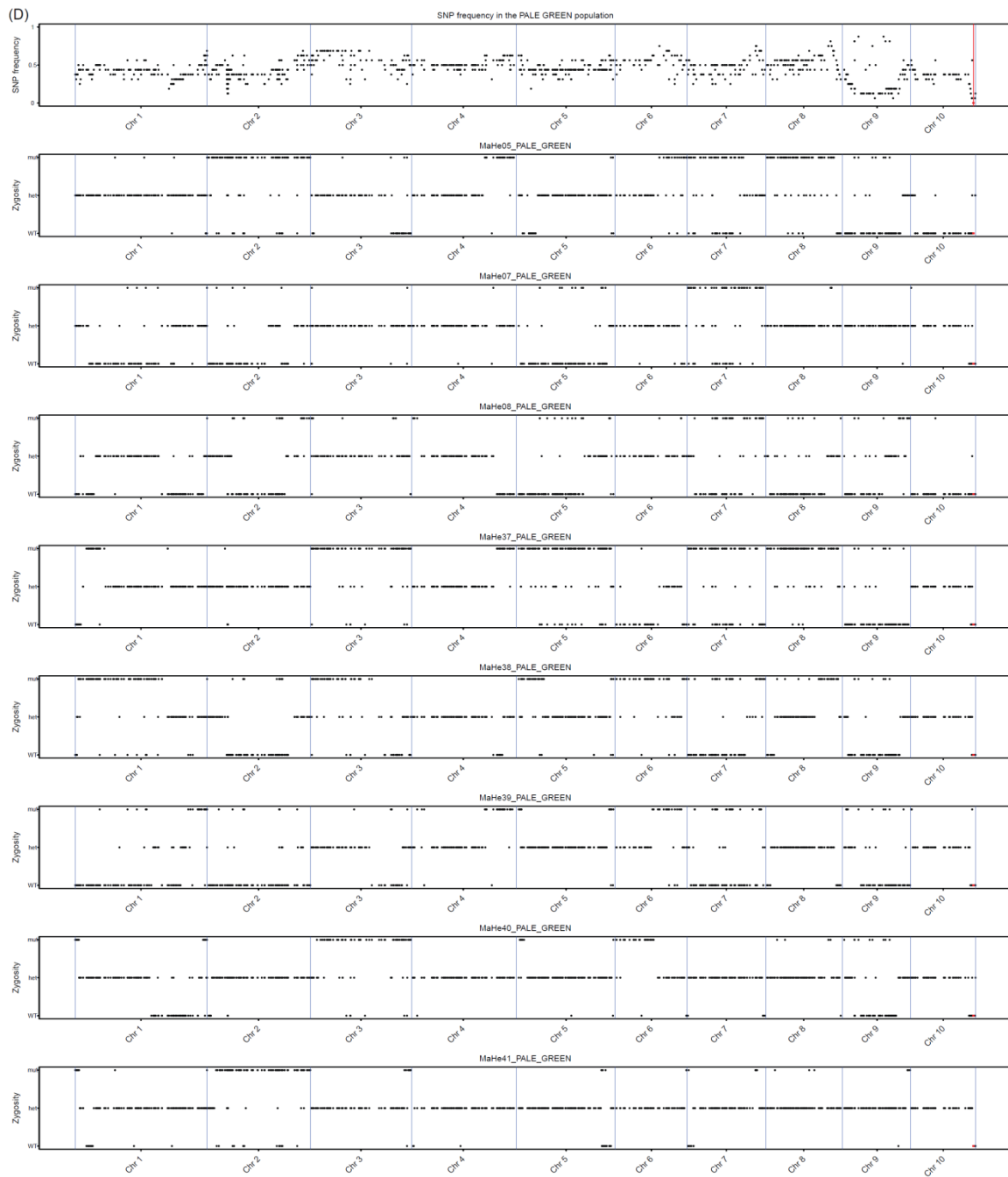MaHe15_pale_green

MaHe36_pale_green

MaHe42_pale_green

Figure S 3_4: Allele frequencies of the segregating SNPs each in aggregated plots and individual plots depicting the zygosity of each SNP position for each sequenced individual, annotated as mut (homozygous mutant), as het (heterozygous mutant/WT), and as WT (homozygous WT). Blue lines indicate start and end point of each chromosome (B73_AGPv3) and the filtered SNPs from Table 3_3 are plotted in red and their position is marked with a red line in the frequency plots. SNP frequencies in each population and the state of zygosity of every SNP in the individuals for (**A**) the seven homozygous *dwarf* individuals, (**B**) the nine *DWARF* individuals (seven heterozygotes, two homozygotes: MaHe25_DWARF and MaHe33_DWARF), (**C**) the eight heterozygote *pale green* individuals, (**D**) and in the eight homozygous *PALE GREEN*

individuals. The occurrence of two alternative haplotypes which are each composed of strings of syntenic mutant and wt alleles of neighboring loci in each of the two mutant families is most probably the result of the performed pollen EMS mutagenesis: Mature maize pollen is in the tricellular G1 stage (Mogensen *et al.*, 1995; Friedman, 1999) and the EMS-induced base (Guanine) modifications result in manifested mutations during the following replication, which takes place after fertilization and leads to two alternative double stranded DNA molecules in the chromatids of the paternal chromosomes. Their mitotic separation lead to chimeric M_1 plants and their propagation to the selfed progeny result in segregation of the aforementioned alternative haplotypes among the individuals of the corresponding M_2 families. Thereby patterns of neighboring homozygous mut and WT SNPs alternating in close vicinity are explained. The rare occurrence of heterozygous SNPs within (otherwise) homozygous chromosomal regions can be attributed to zygosity miscalling by SAMtools erroneously identifying SNPs as heterozygous instead of homozygous which is potentially enhanced in positions of low read coverage.
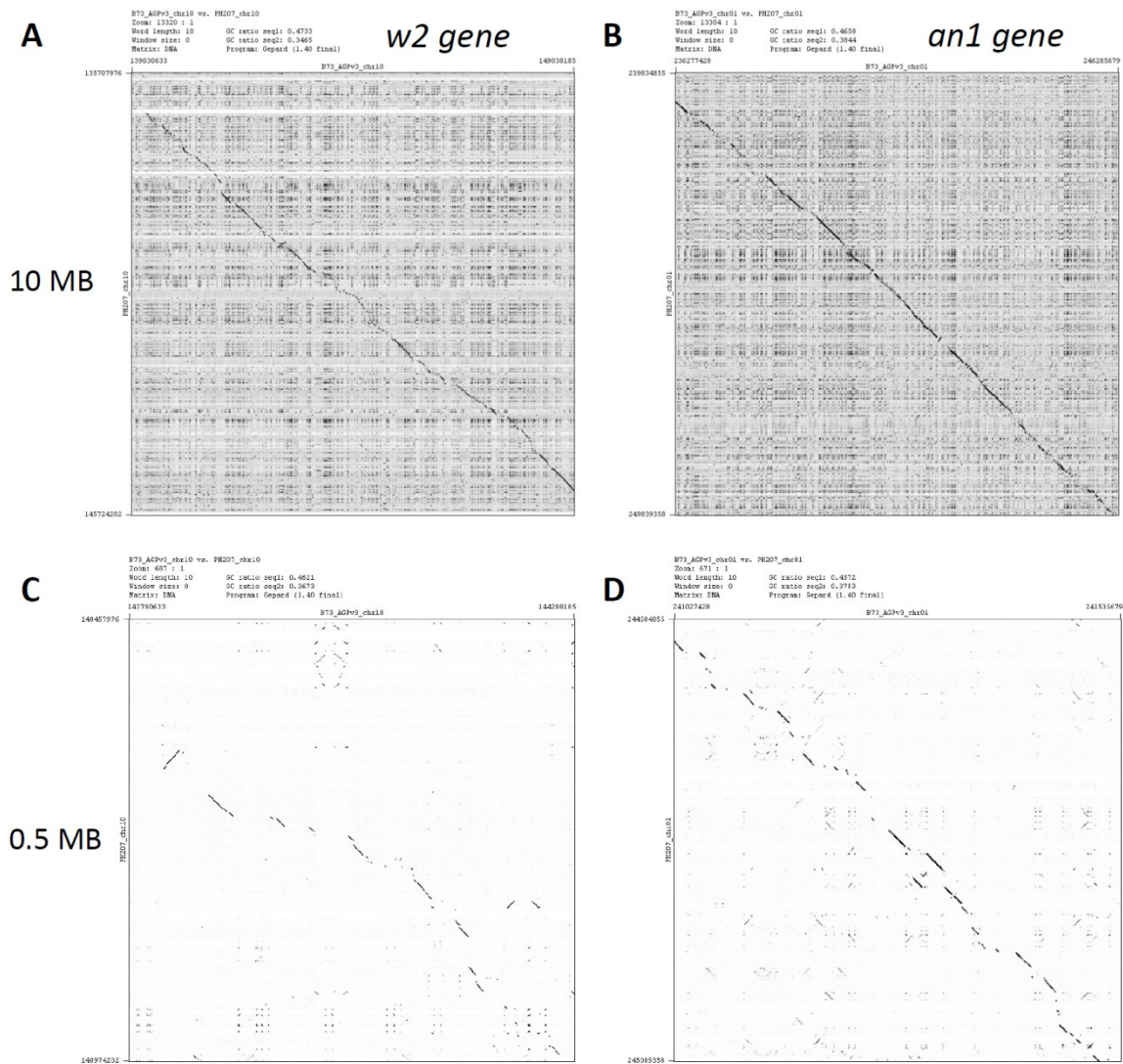
Figure S 3_5: DotPlots with the Gepard software (Krumsiek *et al.*, 2007) of the 10 MB and 500 kB regions surrounding the *an1* (Chr.1) and *w2* (Chr.10) gene loci between the PH207 sequence and the B73_AGPv3 sequence. (**A**) B73_AGPv3 sequence (Chr.10:139030633-149038185; *w2* ± 5 MB) against PH207 (Chr.10:135707976-145724202; *w2* ± 5 MB), (**B**) B73_AGPv3 sequence (Chr.1:236277428-246285679; *an1* ± 5 MB) against PH207 (Chr.1:239834855-249839358; *an1* ± 5 MB), (**C**) B73_AGPv3 sequence (Chr.10:143780633-144288185; *w2* ± 250 kB) against PH207 (Chr.10:140457976-140974202; *w2* ± 250 kB), (**D**) B73_AGPv3 sequence (Chr.1:241027428-241535679; *an1* ± 250 kB) against PH207 (Chr.1:244584855-245089358; *an1* ± 250 kB)

END OF CITATION (Heuermann *et al.*, 2019)