Molecular responses to UV-B in *Arabidopsis thaliana*

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Thomas Przemyslaw Piofczyk

aus Beuthen, Polen

Köln, 2014

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung für Pflanzenzüchtung und Genetik (Prof. Dr. Maarten Koornneef) in der Arbeitsgruppe von Dr. Ales Pecinka angefertigt.

Berichterstatter/in: Prof. Dr. Maarten Koornneef

                            Prof. Dr. Ute Höcker

Tag der mündlichen Prüfung: 20. Oktober 2014

"Wir erkennen die Wahrheit nicht nur
durch die Vernunft,
sondern auch durch das Herz."

Blaise Pascal (1623-1662)

Abstract

As an unavoidable consequence of the requirement for light, higher plants are exposed to UV-B radiation. While several key factors of UV-B signalling, damage prevention and repair are known, their interplay, regulation and variation across natural populations remain to a large extent undiscovered. Although it has been proposed that UV-B radiation can cause mutations, realistic estimations of the mutagenic potential of natural UV-B on a genome-wide level in plants are missing.

In order to explore natural variation in response to UV-B, growth related phenotypes of 345 *Arabidopsis thaliana* accessions in response to UV-B irradiation were analysed by genome-wide association studies and complemented by QTL mapping. A complex genetic basis underlying these traits was revealed and identified the *RECEPTOR-LIKE PROTEIN 7* as a putative candidate gene partially responsible for this variation. In addition, RNA sequencing of UV-B sensitive and resistant accessions followed by mutant analysis revealed several novel candidate genes involved in the response to UV-B irradiation in Arabidopsis.

The mutagenic effects of natural-like UV-B radiation were analysed by whole genome sequencing of over 120 *A. thaliana* genomes from plants grown for up to three full generations under UV-B conditions corresponding to different sites within the natural distribution range. This revealed unaffected mutation rates in Col-0 wild-type plants, irrespective of the UV-B dosage applied. Similarly, the lack of flavonoid biosynthesis, UV-B photoreception and photolyase UVR3 had no influence on the accumulation of UV-B induced mutations. In contrast, UV-B induced many mutations in plants deficient in the UVR2 photolyase, suggesting the importance of this factor in maintaining genome stability in Arabidopsis. The majority of mutations were transitions, which were specifically enriched by UV-B irradiation. Furthermore, UV-B induced mutations were non-randomly distributed. Mutations occurred preferentially in methylated cytosines of transposable elements in pericentromeric regions.

In summary these findings revealed natural variation in *A. thaliana* in response to UV-B and identified several novel genes associated with the response to UV-B that will be further analysed. Moreover, the mutagenic potential of natural-like UV-B irradiation was estimated and gave insight into the mutational spectrum of UV-B induced mutations and the importance of certain genetic factors to prevent the accumulation of mutations.

Zusammenfassung

Als Teil natürlicher Sonnenstrahlung sind Pflanzen zwangsläufig auch UV-B-Strahlung ausgesetzt. Unter anderem durch die Identifizierung des UV-B-Photorezeptors sind bereits einige Komponenten der UV-B-Rezeption, Signaltransduktion und Reparatur UV-B-induzierter DNA-Schäden bekannt. Allerdings ist weitgehend unklar wie das Zusammenspiel dieser Faktoren im Detail abläuft und reguliert ist, ob weitere, bislang unbekannte Faktoren daran beteiligt sind, sowie die genetischen Variationen dieser Komponenten in natürlichen Populationen. Ferner gilt UV-B-Strahlung als ein natürlicher Auslöser von Mutationen. Allerdings sind bislang realistische Berechnungen der Mutationsraten auf genomweiter Ebene für Pflanzen nicht verfügbar.

Zu diesem Zweck wurden einerseits 345 *Arabidopsis thaliana* Ökotypen auf ihre Reaktion hinsichtlich UV-B Bestrahlung phänotypisch untersucht, und genomweite Assoziationsstudien (GWAS) und QTL Kartierungen durchgeführt. Dabei wurde neben einer komplexen genetischen Grundlage der beobachteten Phänotypen das Rezeptor-ähnliche Protein 7 ermittelt, welches zumindest teilweise für die natürliche Variation verantwortlich zu sein scheint. Zu dem wurden RNA-Sequenzierungen einiger sensitiver und resistenter Ökotypen durchgeführt, gefolgt von der Untersuchung ausgewählter Mutanten. Zusammen konnten damit einige Gene ermittelt werden die an der Reaktion auf UV-B-Strahlung beteiligt scheinen.

Andererseits wurden zur Bestimmung der mutagenen Wirkung naturähnlicher UV-B-Strahlung über 120 *A. thaliana* Genome von Pflanzen, die bis zu drei Generationen in UV-B-Bedingungen angezogen wurden die natürlichen Standorten entsprechen, vollständig sequenziert. Dabei stellte sich heraus, dass, unabhängig von der Strahlungsdosis, UV-B keinen Einfluss auf die Mutationsrate in Col-0 Wildtyp-Pflanzen hatte. Im Gegensatz dazu wurden zahlreiche Mutationen in der *uvr2*-Mutante, die keine funktionale UVR2-Photolyase aufwies, ausgelöst. Die UVR2-Photolyase stellte sich dabei als essentiell für die Reparatur von durch UV-B-Strahlung ausgelösten Mutationen heraus. Flavonoide, der UVR8-Photorezeptor sowie die UVR3-Photolyase schienen nur eine untergeordnete Rolle bei der Prävention von durch UV-B-Strahlung ausgelösten Mutationen zu haben. Darüber hinaus zeigte die Analyse der ausgelösten Mutationen eine ungleichmäßige Verteilung. Mutationen traten überwiegend in Transposons (TEs) und in perizentromeren Genomregionen auf. Außerdem lagen methylierte Cytosine häufiger mutiert.

Zusammenfassend konnten durch Untersuchungen der natürlichen Variation von

A. thaliana Populationen mehrere Gene identifiziert werden die bislang nicht mit der Reaktion auf UV-B-Strahlung assoziiert waren und als Grundlage für weitere Untersuchungen dienen können. Daneben konnte die mutagene Wirkung von UV-B-Strahlung durch Bestimmung der Mutationsraten nach UV-B-Bestrahlung ermittelt werden. Dadurch konnten Erkenntnisse hinsichtlich des Mutationsspektrums und des Einflusses bestimmter genetischer Faktoren auf das Auftreten von Mutationen gewonnen werden.

Table of contents

List of abbreviations

6-4PP	6,4-pyrimidine pyrimidone dimer
A	Adenine
ATM	Ataxia Telangiectasia mutated
ATR	Ataxia Telangiectasia and Rad3-related
bp	Base pair
C	Cytosine
CAPS	cleaved amplified polymorphic sequences
CCR2	Cold, circadian rhythm, and RNA binding 2
cDNA	complementary DNA
Col-0	*Arabidopsis thaliana* accession Columbia
COP1	Constitutively photomorphogenic 1
CPD	cyclobutane pyrimidine dimer
CpG	C-phosphate-G
Cry	Cryptochrome
DET1	De-etiolated 1
DGE	Differential gene expression
DNA	Desoxyribonucleic acid
DSB	Double strand break
E. coli	*Escherichia coli*
EMS	Ethylmethansulfonat
et al	et alii / et aliae
FAD	Flavin adenine dinucleotide
FMN	flavin mononucleotide
G	Guanine
G1/2/3	Generation 1/2/3
Gbp	Giga base pair
gDNA	Genomic DNA
GGR	Global genome repair
GWAS	Genome wide association studies
HR	Homologous recombination
HY5	Elongated hypocotyl 5
HYH	HY5 homolog
JA	Jasmonic acid
LUC	Luciferase
MAP	mitogen-activated protein
Mbp	Mega base pair
MKP1	MAP kinase phosphatase

MMR	Mismatch repair
MQM	Multiple QTL model
MRE11	Meiotic recombination 11
MRN	Complex consisting of MRE11, RAD50 and NBS1
MTHF	methenyltetrahydrofolate
NASC	European Arabidopsis Stock Center
NBS1	Nijmegen breakage syndrom 1
NER	Nucleotide excision repair
NHEJ	non-homologous end joining
NIL	Near isogenic line
PCR	Polymerase chain reaction
Phot	Phototropin
Phy	Phytochrome
qRT-PCR	Quantitative reverse Transcriptase PCR
QTL	Quantitative trait locus / loci
RAD50	Ras associated with diabetes 50 protein
RIL	Recombinant inbred line
RNA	Ribonucleic acid
ROS	reactive oxygen species
RUP	Repressor of UV-B photomorphogenesis
SA	Salicylic acid
SD	Standard deviation
SHORE	Short read mapping
SNP	Single nucleotide polymorphism
SSB	Single strand break
SSLP	simple sequence length polymorphism
T	Thymine
Taq	*Thermus aquaticus*
TCR	transcription-coupled repair
TE	Transposable element
TT4	Transparent testa 4
UV	Ultraviolet
UVH	Ultraviolet hypersensitive 1
UVR 2/3/8	Ultraviolet resistance 2/3/8
WEE1	Protein kinase, regulates cell cycle checkpoint in response to DNA damage
WGD	Whole genome duplication
ZTL	Zeitlupe

1 Introduction

Due to their sessile and auto-phototrophic lifestyle, plants are influenced by and dependent on the light environment they are growing in throughout their whole life cycle. As an important environmental cue, light affects many physiological and developmental processes like photosynthesis rate, germination or flowering initiation (Neff et al, 2000). Altogether the development mediated by light is termed photomorphogenesis. In the absence of light, plants develop a drastically different phenotype. In dark-grown (etiolated) Arabidopsis seedlings this includes elongation of the hypocotyl, maintenance of an apical hook, reduced growth of cotyledons, lack of chlorophyll production and few other, less obvious, developmental differences (Fig. 1-1).
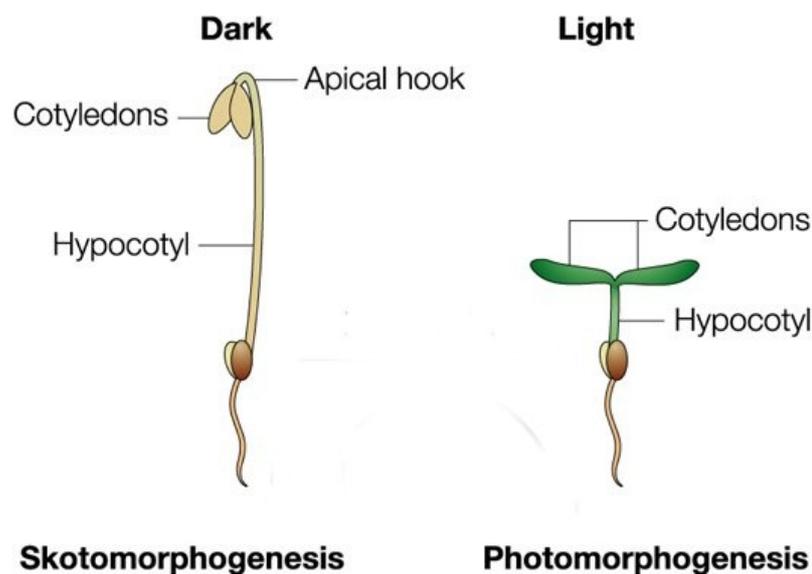
Fig. 1-1. Effects of light during early plant development. A schematic comparison of Arabidopsis seedlings grown in dark and light conditions (modified after Sullivan et al, 2003).

In order to respond to the ever-changing light conditions adequately, plants have evolved mechanism to perceive different light spectra and react accordingly.
Plants can detect light spectra ranging from far-red light to ultraviolet-B (UV-B) through a variety of photoreceptors (Heijde and Ulm, 2012). In the model organism in plant biology, *Arabidopsis thaliana*, 13 photoreceptors are currently known. Many of these are evolutionary conserved among plants (Lariguet and Dunand, 2005).

Plant photoreceptors

The plant photoreceptors so far identified can be subdivided into four families, including phytochromes, cryptochromes, phototropins and Zeitlupe (Fig. 1-2; Sullivan and Deng, 2003; Kami et al, 2010; Heijde and Ulm, 2012). In addition, the UV-B photoreceptor UVR8 has recently been identified (Favory et al., 2009, Rizzini et al, 2011).

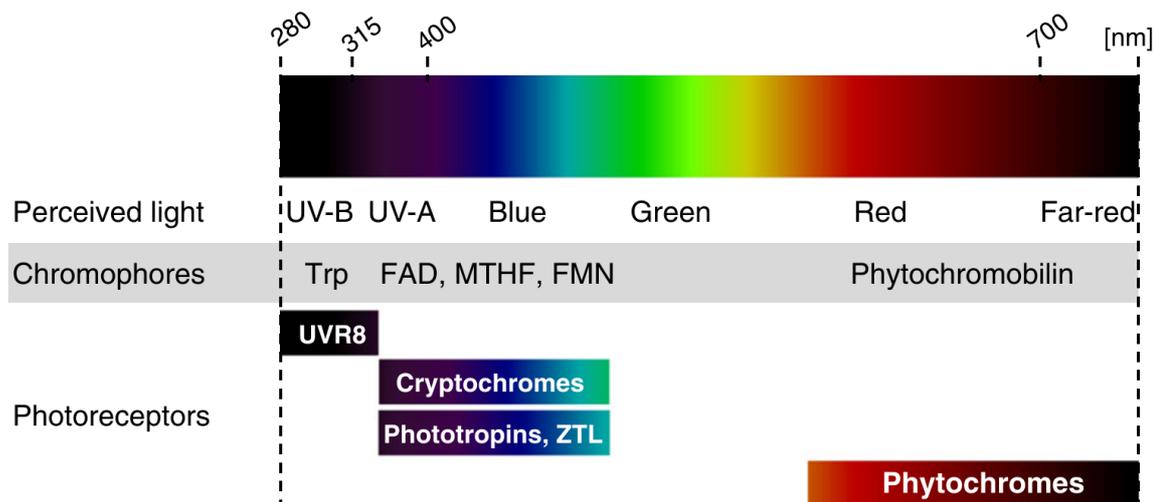**Fig. 1-2**. Photoreceptors in plants. The four classes of photoreceptors in plants and the UV-B photoreceptor UVR8 are depicted together with their absorption spectra and chromophores. Trp=Tryptophans, FAD=flavin adenine dinucleotide, MTHF=methenyltetrahydrofolate, FMN=flavin mononucleotide, ZTL=Zeitlupe family (modified after Heijde and Ulm, 2012).

The phytochrome family in Arabidopsis consists of five phytochromes (PhyA to PhyE) perceiving light in the red to far-red spectrum with their chromophore phytochromobilin, a linear tetrapyrrol. Among other functions, phytochromes are involved in the initiation of germination and the de-etiolation processes (Kami et al., 2010).

The other three classes of photoreceptors - cryptochromes, phototropins and Zeitlupe (ZTL) - absorb light in the UV-A to blue light spectrum (Kami et al., 2010).

Arabidopsis possesses two cryptochromes (Cry1 and Cry2), which are important for example for stomata development and opening. Cryptochromes use flavin adenine dinucleotide (FAD) and methenyltetrahydrofolate (MTHF) as chromophores to capture light (Kami et al., 2010).

Phototropins (Phot1 and Phot2) detect light via the chromophore flavin mononucleotide (FMN) and are associated with phototropic processes like stomata opening, but are also involved in vegetative growth processes (Kami et al., 2010).

The fourth class of photoreceptors in Arabidopsis consists of three Zeitlupe photo-

receptors (ZTL, FKF1 and LKP2). They also use flavin mononucleotide (FMN) as chromophore and are related to the circadian clock and the control of flowering (Kami et al., 2010).

UV-B perception and the UV-B photoreceptor UVR8

UV-B is an intrinsic part of natural solar radiation and as an unavoidable consequence of the requirement for light, higher plants are exposed to UV-B radiation. In general UV radiation can be subdivided into three types of radiation. UV-A, ranging from 400 nm to 315 nm, UV-B from 315 nm to 280 nm and UV-C from 280 nm to 100 nm. UV-A is the least energetic type of UV radiation, whereas UV-C is the most energetic and most harmful.

The ozone layer in the stratosphere filters out the UV-C part of solar radiation and it is thus not naturally present on earth (Caldwell et al, 2007; McKenzie et al., 2007). The same layer largely blocks UV-B, but a small percentage of solar UV-B extends to earth's surface, whereas UV-A penetrates the stratosphere almost unaffected. Up to 99% of the total UV radiation that reaches the earth's surface is UV-A (Caldwell et al, 2007; McKenzie et al., 2007).

In contrast, UV-B levels are highly variable. Biologically effective UV-B radiation on earth is influenced by many factors, like latitude, altitude and cloud coverage, but also anthropogenic factors that influence the ozone content in the stratosphere (McKenzie et al., 2011).

Depending on the intensity of UV-B radiation on the level of the plant, the physiological response is twofold (Fig. 1-3; Jenkins, 2009). Lower ambient UV-B doses stimulate photomorphogenic signalling (Brown and Jenkins, 2008). The UV-B signal is directly perceived by the UV-B photoreceptor UVR8 and transduced via several key factors like COP1, HY5 and HYH to induce specific target genes and downstream responses (Ulm et al., 2004; Brown et al, 2005; Oravecz et al, 2006; Favory et al, 2009). Besides, photomorphogenic signalling can also be accomplished via UVR8-independent pathways, for example via a cascade of kinases, as recently shown by MAP kinase phosphatase 1, MAP kinase 3 and MAP kinase 6 (Gonzalez-Besteiro et al., 2011).

Higher doses of UV-B stimulate nonspecific signalling, which leads to higher levels of DNA damage, accumulation of reactive oxygen species (ROS) and elevated levels of stress-related phytohormones, such as jasmonic acid (JA) and salicylic acid (SA) (A-H-Mackerness et al., 1999; Apel and Hirt, 2004; Ballare, 2014). Together this induces, partially overlapping, the expression of target genes and further responses. As indicated,

the responses following nonspecific and photomorphogenic signalling show a more smooth transition rather than strictly separated pathways.
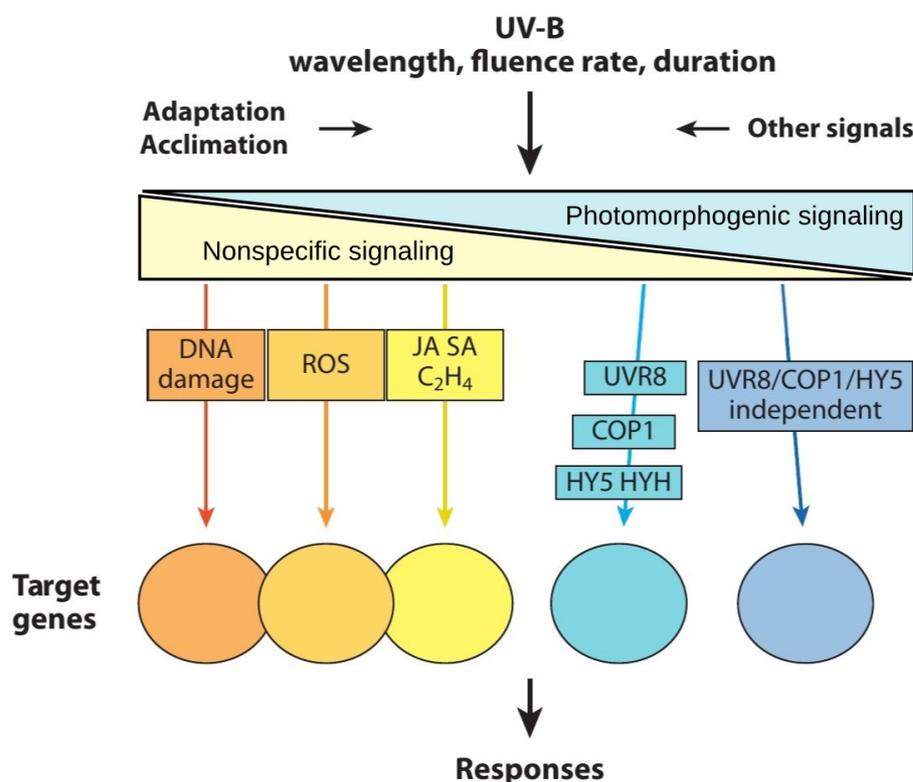
Fig. 1-3. UV-B signal transduction pathways. UV-B stimulates nonspecific and photomorphogenic signalling, depending on the UV-B intensity as well as other factors, leading to the induction of target genes and downstream responses. ROS = reactive oxygen species, JA = jasmonic acid, SA = salicylic acid, UVR8 = UV photoreceptor UV RESISTANCE LOCUS8, COP1 = CONSTITUTIVELY PHOTOMORPHOGENIC1, HY5 = ELONGATED HYPOCOTYL5, HYH = HY5 HOMOLOG (modified after Jenkins, 2009).

UVR8 was first identified in a genetic screen for mutants hypersensitive to UV-B and after further characterization recognized as the UV-B photoreceptor (Kliebenstein et al., 2002; Favory et al., 2009, Rizzini et al., 2011). Unlike other photoreceptors, which detect different light qualities using a bound chromophore, UVR8 detects UV-B via intrinsic tryptophan residues (Christie et al., 2012, Wu et al., 2012). In its inactive form UVR8 forms a homodimer, which is mainly located in the cytoplasm. UVR8 can be found almost ubiquitously in every plant cell (Kaiserli and Jenkins, 2007; Kilian et al., 2007; Rizzini et al., 2011). After exposure to UV-B, the UVR8 dimer breaks down into monomers, which accumulate in the nucleus and initiate UV-B responses upon interaction with the E3 ubiquitin ligase COP1 (Kaiserli and Jenkins, 2007; Rizzini et al.,

2011; Tilbrook et al., 2013). Monomerization after UV-B exposure is a process that is directly reversible and independent of other factors in the absence of UV-B. However, *in vivo* re-dimerization is accelerated by the UVR8-interacting proteins RUP1 and RUP2 (Gruber et al., 2010; Heijde and Ulm, 2013; Heilmann and Jenkins, 2013).

UV-B as a DNA damaging agent and mechanisms to prevent damage

DNA is the main storage molecule for genetic information for all living organisms and therefore its sequence and integrity need to be protected. However, due to its relatively high energy content and its natural occurrence on earth's surface, UV-B radiation can cause damages to macromolecules like DNA, which makes it one of the most potent naturally occurring DNA damaging agents (Kimura et al., 2004; Kimura and Sakaguchi, 2006).

At the level of DNA, UV-B radiation can induce non-native bonds between adjacent pyrimidine bases (pyrimidine dimers) that hinder transcription and replication (Clark et al., 2011). Pyrimidine dimers are thought to be the most common type of DNA damage in plant tissues (Kimura et al., 2004; Kimura and Sakaguchi, 2006). Two types of pyrimidine dimers can be distinguished, 6,4 pyrimidine pyrimidone dimer (6-4PP) and cyclobutane pyrimidine dimer (CPD) (Britt, 2004). CPDs make up approximately 75% of pyrimidine dimers induced by UV-B, with 6-4PP making up the rest (Britt et al., 1996; Lo et al., 2005). The main repair pathway for these lesions is photoreactivation (Britt et al., 2004). In photoreactivation pyrimidine dimers are repaired by a direct reversal mechanism, which involves photolyase enzymes. Photolyases are able to use the energy conferred by UV-A and blue light to resolve pyrimidine dimers and restore the original DNA sequence (Sancar et al., 2003). Arabidopsis has two photolyase genes, *UVR2* (At1g12370) and *UVR3* (At3g15620) (Jiang et al., 1997). As the structure of the two pyrimidine dimers is different, each photolyase can only resolve one type of lesion, with UVR2 repairing CPDs and UVR3 being specific to 6-4PPs (Britt et al., 2004). Apart from photoreactivation plants have also light-independent repair pathways.

In the nucleotide excision repair mechanism (NER) after recognition of the lesion, DNA is unwound, the lesion containing sequence is excised and then synthesized by using the intact opposite strand as template (Shuck et al., 2008). As NER recognizes and repairs various kinds of DNA damage ubiquitously, the NER pathway is also referred to as global genome repair (GGR). In addition to that, a NER pathway acting specifically on transcribed DNA exists, the transcription-coupled repair (TCR), which emphasizes the importance of the NER mechanism (Kimura and Sakaguchi, 2006; Tuteja et al., 2009).

Mutants impaired in the NER pathway, like *uvh1*, a component of an endonuclease that excises dimers, are hypersensitive to UV-B radiation (Liu et al., 2000).

A third mechanisms is that of mismatch repair (MMR). In MMR a mismatched base pair formed by the incorporation of an incorrect base during DNA replication and recombination is repaired by excision of a part of the erroneous base-containing strand (Iyer et al., 2006; Kimura and Sakaguchi, 2006). As the newly synthesized strand will commonly contain the mismatching base, it can be easily distinguished from the template strand by missing DNA methylation. After recognition and excision of the lesion, the DNA strand will be newly synthesized again using the parental strand as a template. Although the importance of MMR for UV-B induced mutations outside of active DNA replication is unclear, there is evidence that non-functional MMR leads to an increased number of DNA lesions, which can potentially lead to mutations (Lario et al., 2011).

A fourth mechanism that plants may employ for repair of DNA damage is recombination repair. If breaks occur on both strands of DNA (double strand break, DSB), the lesioned DNA sequences can be repaired by two processes. Either the DNA sequence is exchanged between similar or identical molecules, e.g. sister chromatids, in a process called homologous recombination (HR), in which the original sequence integrity is maintained, or by non-homologous end joining (NHEJ). In NHEJ the break ends of the DNA strands are directly ligated together without the use of a template (Bray and West, 2005; Puchta, 2005; Schuermann et al., 2005). But it remains unclear if natural doses of UV-B can cause DSBs that will lead to mutations in considerable numbers, or even at all. DSBs are more likely to occur as a result of strong, ionizing radiation like UV-C or gamma radiation. In contrast, the UV-B radiation reaching the earth's surface is solely a non-ionizing type of radiation. However, natural-like UV-B doses have been shown to have a negative effect on plant genome stability by increasing somatic homologous recombination frequencies of transgenic reporter constructs in plant leaves (Ries et al., 2000).

If the described repair mechanisms fail and/or a mutation occurs during DNA replication and cannot be repaired in time, the photoproducts can be bypassed by error-prone DNA polymerases that may incorporate incorrect nucleotides opposite to the pyrimidine dimer (Vaisman et al., 2003; Britt et al., 2004; Choi et al., 2006). Thus, unrepaired lesions can lead to fixed mutations. However, it is unknown how many UV-B induced mutations are generated and transmitted to their offspring under natural conditions.

As the maintenance of DNA integrity is crucial for the survival and proliferation of a plant, plants have also evolved "passive" ways to shield themselves against the deleterious

effects of UV-B radiation, in addition to the "active" repair systems described.

Flavonoids, a class of plant secondary metabolites, have been shown to accumulate in response to UV-B radiation and to have a protective value (Mazza et al., 2000; Bieza and Lois, 2001; Emiliani et al., 2013). Mutants impaired in flavonoid production, like *transparent testa 4* (*tt4*), encoding a chalcone synthase enzyme, which catalyzes one of the first steps in the flavonoid biosynthesis pathway, are more sensitive to UV-B radiation than wild-type plants (Li et al., 1993). However, it is not clear, whether flavonoids primarily act in directly shielding plants against UV-B radiation by absorbing or filtering UV-B, or rather indirectly as scavengers of UV-B caused by-products like reactive oxygen species (Agati and Tattini, 2010; Fini et al., 2011). Indeed, it has been demonstrated that reactive oxygen species (ROS) production increases under UV-B conditions in plants (Hideg and Vass, 1996; Gerhardt et al., 2005). Potentially ROS accumulate due to photosynthesis apparatus protein damage caused by UV-B light, leading to a reduced ability to quench excitation energy (Barta et al., 2004). Although the mechanism of protection is unknown, in either case flavonoids are undoubtedly involved in the photoprotection of plants.

Beside the activity of the repair machinery and plant secondary metabolites, the organization of DNA can also play a role in the accumulation of DNA lesions. The genome of plants, as that of all eukaryotic cells, is organized in chromatin, a complex and very dynamic combination of DNA and many different protein factors (He et al., 2011). The primary function of chromatin is to achieve a compaction of DNA to physically store large DNA molecules in a relatively small volume of the nucleus, to package DNA for the migration during mitosis and to control gene expression and DNA replication (He et al., 2011). At the same time, chromatin compaction can also influence the amount of DNA damage caused and/or the efficiency of the repair of DNA lesions (Suter et al., 1997; Campi et al., 2012).

DNA damage recognition and transduction

In order to activate the described repair mechanisms, DNA damage has to be recognized as such. In plants, the protein kinases ATM (Ataxia Telangiectasia mutated) and ATR (Ataxia Telangiectasia and Rad3-related) play a central role in the signalling of DNA damage (Garcia et al., 2003; Culligan et al., 2004; Kurz and Less-Miller, 2004; Shechter et al., 2004; Culligan et al., 2006).

DSBs are recognized by a complex called MRN, consisting of the three factors MRE11, RAD50 and NBS1, which identify and bind to the DSB sites and activate mainly ATM, but

also ATR (Lee and Paull, 2005; Bleuyard et al., 2004; Waterworth et al., 2007; Amiard et al., 2010). ATM then phosphorylates many proteins involved in the DNA repair, the initiation of translation or checkpoint control during DNA replication and the mitotic cell cycle (Shiloh, 2006; Shiloh and Ziv; 2013). ATR mutants in *Arabidopsis thaliana* have been found to be hypersensitive to UV-B light (Culligan et al., 2004), whereas mutants deficient of ATM are not (Garcia et al., 2003).

On the other hand, single strand breaks or lesions (SSB) are believed to mainly be transduced by ATR (Culligan et al., 2004; Sancar et al., 2004; Culligan et al., 2006). Thereby, ATR is primarily activated by replication fork stalling during DNA replication, which can be a consequence of a DNA lesion caused by UV-B, as described before (Culligan et al., 2004; Sancar et al., 2004). Thereafter, ATR would regulate the cell cycle via cell cycle dependent kinases and regulators, in which the kinase WEE1 has been shown to be one of the key factors (Abraham, 2001; De Schutter et al., 2007). ATR, as well as the phosphatase MKP1, have been shown to play a distinct role in the response to UV-B stress in Arabidopsis (Gonzalez-Besterio and Ulm, 2013). It is proposed that both work in parallel in different tissues.

The recognition of SSBs and UV-induced pyrimidine dimers is less well understood. Depending on the repair mechanism employed, it can be a multi component complex, including damaged-DNA binding (DDB) factors, chromatin-remodelling factors and histone-modifying enzymes, for example for the NER pathway (Sancar et al., 2004; Palomera-Sanchez and Zurita, 2011). The initial trigger in this, is believed to be a change in the chromatin or nucleosome structure caused by the pyrimidine dimer.

For repair via photoreactivation, photolyase enzymes both directly recognize and repair the DNA lesion (Sancar et al., 2004; Essen and Klar, 2006; Yang, 2011).

Arabidopsis natural variation and evolution

Plant genomes show a remarkable variability in terms of their size and organization, ranging from small genomes like in *Genlisea margaretae* with only about 63 Mbp to very large genomes like in *Paris japonica* with about 150 Gbp (Greilhuber et al., 2006; Pellicer et al., 2010; Proost et al., 2011), an almost 2400-fold range. Chromosome numbers can range from n=2 for *Haplopappus gracilis* up to n=630 in the adder's-tongue family of ferns (e.g. *Ophioglossum reticulatum*) (Jackson, 1959; van den Burg, 2004; Roberto, 2005).

Many different mechanisms and phenomena can explain this enormous plasticity, including whole genome duplications (WGD), activity and amplification of transposable

elements, sequence insertions and deletions and point mutations (Proost et al., 2011). If a point mutation is stably fixed within a population with higher allele frequencies it will be referred to as a single nucleotide polymorphism (SNP). In the context of evolution, WGDs exhibit a rare event with a huge impact on genome size and organization, whereas SNPs are much more frequent with usually low impact, depending on the type of the mutation (Adams and Wendel, 2005; Proost et al., 2011). This can vary from no effect (neutral) e.g. in case of a silent mutation, to a non-sense mutation that alters the function of a gene, causing decreased or increased fitness, respectively. Indeed, some polymorphisms have been found to be associated with an altered gene function, hence having an influence on the fitness of the plant (Alonso-Blanco et al., 2009; Weigel, 2011). With respect to UV radiation, several studies indicated variability in response to UV-A and/or UV-B among natural accessions of *A. thaliana*, although the underlying genetic basis could not be identified (Torabinejad and Caldwell 2000; Cooley et al., 2001; Kalbina and Strid, 2006; Jansen et al., 2010).

Only due to recent technical advances, especially in the field of next generation sequencing, SNPs and short insertions/deletions can be identified accurately and on a genome-wide scale (Ossowski et al., 2008; Schneeberger et al., 2009; Grimm et al., 2013). This has not only revealed an enormous natural variation among Arabidopsis accessions, but together with phenotypic data it has been used for genetic mapping of different traits (Ossowski et al., 2008; Atwell et al., 2010; Cao et al., 2011; Weigel, 2011). Natural variation within a species like *A. thaliana* is not only limited to SNPs, but also to variation on a larger scale. Whole protein coding genes can be naturally absent in one accession compared to another, and large structural re-arrangements, e.g. deletions of several kb-size, can occur (Clark et al., 2007; Cao et al., 2011; Gan et al., 2011).

There is an increasing amount of data available for *A. thaliana*, with several hundreds of natural accessions already whole-genome-sequenced within the 1001 genomes project and over 1000 accessions densely genotyped for 250,000 markers using a microarray SNP chip (Weigel and Mott, 2009; Horton et al., 2012).

Genome-wide analysis of spontaneous mutations of *A. thaliana* plants revealed a basic mutation rate of around 7×10^{-9} per site and generation (Ossowski et al., 2010). It could be shown that the mutation rate can vary substantially between species, from 1.2×10^{-8} in humans, to 2.8×10^{-9} in the fruitfly *Drosophila melanogaster* to 1.7×10^{-10} in the budding yeast *Saccharomyces cerevisiae* (Kong et al., 2012; Keightley et al., 2014; Zhu et al., 2014). The mutation rate was analysed for only one other plant species so far, *Chlamydomonas reinhardtii*, and identified a much lower mutation rate with 3.2×10^{-10}

compared to *A. thaliana* (Ness et al., 2012). However, the number of lines sequenced and the total number of identified mutations was very low and the unicellular organism *C. reinhardtii* might be evolutionary already too distant to draw further conclusions.

For *A. thaliana,* the mutational spectrum was strongly skewed towards CG to TA nucleotide transitions (Ossowski et al., 2010). As many of the mutations identified in this study overlapped with di-pyrimidine sites and/or methylated cytosines, two phenomena could explain this bias in the mutation spectrum. On the one hand, deamination of methylated cytosines can lead to thymine substitutions (Lindthal and Nyberg, 1974; Coulondre et al, 1978; Duncan and Miller, 1980; Xia et al., 2012). On the other hand, UV-B induced pyrimidine dimer and the DNA replication of those by error-prone DNA polymerases can lead to the incorporation of an incorrect nucleotide, as described before. However, in the course of this study plants were grown in standard greenhouses, which are devoid of UV-B radiation. Yet, the presence of low doses of UV-A radiation can not be excluded. This suggests that under more natural conditions with physiological doses of UV-B radiation higher mutation rates may be present, and that UV radiation could play a role in generating DNA sequence variation in the plant genome. Both mechanisms, methylation and UV-B, are suggested to work together, meaning that methylation can influence the probability of damage formation induced by UV-B radiation, as shown in a reporter gene construct and *in-vitro* experiments (Ikehata et al., 2003; Rochette et al., 2008). However, for unmethylated cytosines a bias in favour of transitions may not always be the case, as seen in the mutation spectrum of pseudogene sequences in a grasshopper species (Keller et al., 2007).

Recently a study addressed these questions more thoroughly. This study identified nearly 1,000 spontaneous mutations by whole-genome sequencing which accumulated in over ~311,000 generations in 145 diploid lines of the budding yeast *S. cerevisiae* (Zhu et al., 2014). It could be confirmed that both previously described phenomena exist also in yeast. A strong bias in favour of CG to AT transition was identified, as well as a higher rate of mutation at CpG dinucleotides in two specific contexts consistent with cytosine methylation (Zhu et al., 2014). However, it is questionable how applicable this results are to other species like *A. thaliana*, as the propagation of the yeast strains within this study was carried out under rather unnatural conditions and with yeast being only an unicellular organism.

Aim of this study

Recent progress in the field of UV-B research identified with UVR8, RUP1 and RUP2, and their interaction with COP1 and HY5, the key players in perception of UV-B and photomorphogenic signalling (Ulm et al., 2004; Brown et al, 2005; Oravecz et al, 2006; Favory et al, 2009; Gonzalez-Besteiro and Ulm, 2011). On the other hand little is known about the signalling and the responses following UV-B stress. As pointed out, ATR and MKP1, as well as DET1, are important factors in the UV-B stress signalling pathway (Castells et al., 2011; Gonzalez-Besteiro and Ulm, 2013), but more detailed information of the complex interaction within the stress signalling pathway is lacking.

Natural variation has been shown to be a useful resource to understand adaptation to the environment in which plants are growing in, including different UV-B doses, and to identify the underlying genes of this variation (Alonso-Blanco et al., 2009; Bergelson and Roux, 2010; Weigel, 2011). Previous studies have already indicated natural variation in plants in response to UV-B radiation, although the numbers of accessions were low and the genetic basis underlying the observed differences was not identified (Torabinejad and Caldwell 2000; Cooley et al., 2001; Kalbina and Strid, 2006; Jansen et al., 2010).

This study was conducted with two major aims, focussing on understanding the effects of UV-B radiation on plant growth and genome stability.

The first aim was to perform a natural variation screen in response to UV-B stress with a large number of accessions of *A. thaliana* and to carry out genetic mapping to get more insight into the UV-B stress response pathway.

It has been proposed that UV-B is one of or even the most important environmental factor causing mutations and hence be a driving force in evolution (Henri and Henri, 1914; McLennan, 1987; Kimura et al., 2004; Kimura and Sakaguchi, 2006). Interestingly, the majority of studies supporting this hypothesis used unnaturally high doses of UV-B, or even UV-C radiation, which does not occur naturally. Even if more ambient doses were applied, often, due to the experimental design, it is questionable if or how these results can be translated e.g. from *in-vitro* grown unicellular organisms to higher, more complex, organisms grown under realistic conditions.

Therefore, the second aim of this study was to identify the mutagenic effects of UV-B radiation on a genome-wide scale in wild-type plants and UV-B hypersensitive mutants. To this end, a mutation accumulation project was performed in which *A. thaliana* plants were grown under natural-like UV-B radiation conditions.

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

All chemicals were purchased from one of the following companies: Applichem (Darmstadt, Germany), Bio-Budget Technologies (Krefeld, Germany), Bio-Rad (Hercules, USA), Carl-Roth (Karlsruhe, Germany), Duchefa Biochemie (Haarlem, Netherlands), Fermentas / ThermoScientific (St. Leon-Rot, Germany), Life Technologies (Karlsruhe, Germany), Merck (Darmstadt, Germany), New England Biolabs (Frankfurt am Main, Germany), peqLab (Erlangen, Germany), Promega (Mannheim, Germany), Roche (Basel, Switzerland), Sigma-Aldrich (Steinheim, Germany), VWR International (Darmstadt, Germany).

2.1.2 Enzymes

All restriction enzymes were purchased from New England Biolabs or Fermentas / ThemoScientific.

Taq DNA polymerase for standard PCR reactions was obtained from Ampliqon (Odense, Denmark) or Fermentas / ThemoScientific, Phusion proof-reading polymerase from New England Biolabs. DNase I was obtained from Roche. Reverse Transcriptase Superscript II was purchased from Life Technologies.

2.1.3 Molecular biology kits

All kits were used according to the manufacturer's instructions.

Agencourt AMPure XP Kit	Beckman Coulter Genomics	Bernried, Germany
Agilent RNA 6000 Nano Kit	Agilent Technologies	Böblingen, Germany
Agilent DNA 1000 Kit	Agilent Technologies	Böblingen, Germany
BioSprint 96 DNA Plant Kit	Qiagen	Hilden, Germany
First Strand cDNA Synthesis Kit	Fermentas / ThermoScientific	St. Leon-Rot, Germany
Gateway BP clonase II enzyme	Life Technologies	Karlsruhe, Germany
Gateway LR clonase II enzyme	Life Technologies	Karlsruhe, Germany
Gateway TOPO TA Cloning Kit	Life Technologies	Karlsruhe, Germany
MinElute Gel Extraction Kit	Qiagen	Hilden, Germany
Nucleospin Gel and PCR Kit	Macherey-Nagel	Düren, Germany
NucleoSpin Plasmid Kit	Macherey-Nagel	Düren, Germany
Nucleon PhytoPure gDNA Kit	GE healthcare	München, Germany
RNeasy Plant Mini Kit	Qiagen	Hilden, Germany

TruSeq DNA Sample Prep Kit	Illumina	San Diego, USA
TruSeq RNA Sample Prep Kit	Illumina	San Diego, USA
SensiMix SYBR & Fluorescein Kit	peqLab	Erlangen, Germany
Qubit dsDNA HS assay Kit	Life Technologies	Karlsruhe, Germany
Qubit RNA BR assay Kit	Life Technologies	Karlsruhe, Germany

2.1.4 Oligonucleotides

All oligonucleotides were synthesized by Metabion (Martinsried, Germany) or Life Technologies and dissolved in sterile water to a concentration of 100 µM.

2.1.5 Plant material and growth

All experiments were carried out using the model plant species *Arabidopsis thaliana*. The different ecotypes or accessions, respectively, are enlisted in the appendix and were provided by Maarten Koornneef's lab. Mutant lines were obtained from the European Arabidopsis Stock Center (NASC), except *uvr8-6*, *uvr2*, *uvr3* and *uvr2/uvr3*. The mutant line *uvr8-6* was kindly provided by Dr. Roman Ulm (University of Geneva), *uvr2*, *uvr3* and *uvr2/uvr3* by Dr. Chris Bowler (Ecole Normale Supérieure, Paris). A recombinant inbred line (RIL) population, derived from a cross between Col-0 and C24 accessions, and its related near isogenic lines (NILs) were kindly provided by Dr. Thomas Altmann (IPK Gatersleben). Detailed information can be found in the appendix.

2.1.6 Bacterial strains

For standard cloning procedures the *Escherichia coli* strain DH5α was used. For transformation of *A. thaliana* the *Agrobacterium tumefaciens* strain GV3101 was used.

2.1.7 Plasmids

Apart from plasmids included in the GATEWAY TOPO TA Cloning kit, the plasmid pGWB435 was used. The plasmid is suitable for the Gateway cloning environment and has a C-terminal Luciferase gene (Nakagawa et al., 2007). It was kindly provided by the lab of Dr. Jane Parker at the MPIPZ.

2.1.8 Software and databases

For *in-silico* analyses of DNA sequences the software DNASTAR Lasergene v8.0, SnapGene Viewer v2.2.2 and MEGA v6.0 were used. Statistical analyses were performed with R (http://www.r-project.org) and the appropriate packages.

Genome and transcriptome sequencing analysis was performed using the following software:

Bowtie2 v2.0.5, FASTX-toolkit v0.0.13.2, fastQC v0.10.1, QualiMap v0.6, SAM tools v0.1.18, SHORE v0.9 and TopHat v2.0.10, DEseq package for R (Langmead and Salzberg, 2012; Garcia-Alcalde et al., 2012; Li et al., 2009; Schneeberger et al., 2009; Kim et al., 2011; Anders and Huber, 2010)

Image processing was performed with the ImageJ software v1.4.7a.

For different purposes the following databases/tools have been used in the course of this study:

- Various information about *A. thaliana* - http://www.arabidopsis.org
- Information about *A. thaliana* T-DNA mutants - http://signal.salk.edu/cgi-bin/tdnaexpress (Alonso et al., 2003)
- Database for scientific publications - http://www.ncbi.nlm.nih.gov/
- Primer desgining tool - http://frodo.wi.mit.edu/primer3/ (Koressaar and Remm, 2003; Untergasser et al., 2012)
- Database for microarray expression data for *A. thaliana* - http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Kilian et al., 2007; Winter et al., 2007)
- Database for transcriptome data for *A. thaliana* - https://www.genevestigator.com/gv/ (Hruz et al., 2008)
- 1001 genomes project browser - http://signal.salk.edu/atg1001/index.php (Weigel and Mott, 2009)
- Co-expression database for *A. thaliana* - http://atted.jp/ (Obayashi et al., 2011)

2.2 Methods

All standard and not further described methods were conducted as described in Sambrook and Russell, 2001 or Weigel and Glazebrook, 2002.

2.2.1 DNA extraction

Extraction of genomic DNA for standard PCR reactions, e.g. for genotyping, was carried out with the Biosprint96 DNA Plant kit. Samples for whole genome sequencing were obtained by using the Nucleon PhytoPure gDNA kit. Plasmid-DNA was extracted with the NucleoSpin Plasmid kit. All three kits were used according to the manufacturer's instructions.

2.2.2 RNA extraction

RNA extraction for gene expression analysis and whole transcriptome sequencing was performed using the RNeasy Plant Mini Kit with an on-column DNA digestion using DNase I enzyme.

2.2.3 Polymerase chain reaction (PCR)

Standard PCR reactions were carried out using Taq DNA polymerase according to the manufacturer's instructions. For applications where DNA sequence integrity was essential, Phusion proof-reading polymerase was used with no alteration of the enclosed protocol.

2.2.4 Purification of PCR products and gel extraction of DNA fragments

PCR products and gel separated DNA fragments were purified using the Nucleospin Gel and PCR kit according to the manufacturer's instructions.

2.2.5 Gateway-Cloning

For cloning of DNA sequences the Gateway system from Invitrogen was employed, using the Gateway BP clonase II enzyme kit, the Gateway LR clonase II enzyme and the Gateway TOPO TA Cloning kit according to the manufacturer's protocol. The destination vector used, was the vector pGWB435 with a C-terminal coding sequence for firefly Luciferase (Nakagawa et al., 2007).

2.2.6 Gene expression analysis

For gene expression analysis RNA was extracted as described and 1 µg of total RNA was converted into cDNA using the First Strand cDNA Synthesis kit. RNA concentration was quantified with the NanoDrop ND-1000 spectral photometer from peqLab. Quantitative reverse Transcriptase PCR (qRT-PCR) was carried out by using the SensiMix SYBR & Fluorescein kit on an Eppendorf EP Realplex. Gene expression was calculated with the standard curve method and normalized to an internal control, represented by the reference gene *Actin7* (AT5G09810).

2.2.7 Single read sequencing

Single read sequencing was carried out in the Max Planck Genome Center Cologne on an Applied Biosystems 3730XL Genetic Analyser using BigDye-Terminator Chemistry.

2.2.8 Whole genome and whole transcriptome sequencing

For whole genome and whole transcriptome sequencing DNA and RNA were extracted and purified as described. The exact concentration was measured by Qubit fluorometer quantification and the RNA integrity analysed with the Bioanalyzer assay using the Agilent RNA 6000 Nano kit prior to library preparation. Both DNA and RNA library preparations were carried out according to the Illumina TruSeq RNA or DNA Preparation kit, respectively. Library insert size and integrity was then analysed on the Bioanalyzer with the Agilent DNA 1000 kit and the concentration determined by Qubit fluorometer quantification. Sequencing was performed on an Illumina HiSeq2500 Sequencer. Sequencing run conditions were 100 bp paired end for whole genome sequencing and 100 bp single end for whole transcriptome sequencing. On average each sample for whole genome analysis was sequenced at a sequencing depth of around 30 million reads, whereas samples for whole transcriptome analysis were sequenced at a sequencing depth of around 15 million reads.

2.2.9 Plant growth

Plants were sown on moist standard soil and kept for 4 days at 4° C in darkness for synchronization of germination. Then, plants were transferred to greenhouse or growth chamber, depending on the following experiment. In both greenhouse and growth chamber conditions were similar: Long day conditions with 16 h daylight from 6 am to 10 pm and night from 10 pm to 6 am; day temperature 22° C and night temperature 18° C; relative humidity 50%; light intensity ~150 µmol m^{-2} s^{-1}.

2.2.10 UV-B treatment

UV-B treatment was carried out in a growth chamber, where the UV-B containing area consisted of standard white light bulbs (Philips TLD 50 W / 840) supplemented with four Philips TL40 W / 12RS SLV UV fluorescent tubes (fluence rate at 312 nm: 4.35 W / cm²). After 30 min of irradiation supplementary UV-B radiation was switched off. The treatment was repeated for 10 consecutive days, where UV-B was applied at 12.15 pm to 12.45 pm. Control plants were treated in parallel but shielded from UV-B radiation by a lightproof plastic foil.

2.2.11 Crossing of plants

For crossing of Arabidopsis plants open flowers and developed siliques were removed from the inflorescence of the mother plant, which served as a pollen acceptor. Three to four closed and immature buds remained on this inflorescence and were cleared from sepals, petals and stamens. After two days the stigmata of these buds were pollinated by rubbing the anthers of the father plant as the pollen donor.

2.2.12 Plant transformation

Plants were transformed using the floral dip method (Clough and Bent, 1998), following an adjusted protocol (Zhang et al, 2006). Selection of positive transformants was carried out according to the same protocol.

2.2.13 Luciferase assay

For detection of Luciferase enzyme activity, plants potentially expressing the Luficerase enzyme were sprayed with Luciferin solution (100 mM D-Luciferin sodium salt in 0.01% Triton-X-100) and incubated for 5 min at room temperature. Photon counting was performed using a set-up consisting of a Hammatsu Image Intensifier Controller M4314 and a Hammatsu ICCD camera C2400-40H with a Nikon lens Nikkor 50 mm. Analysis was performed with the Hammatsu software Hokawo Image 2.1.

2.2.14 Quantitative trait locus mapping

Quantitative trait locus (QTL) mapping was performed with the R package "qtl", using the multiple QTL model (MQM) approach (Broman et al., 2003; Arends et al., 2010). Significance threshold was calculated by a permutation test analysis with 1000 permutations.

2.2.15 Genome wide associations studies

Genome wide associations studies were performed by using the EMMA-X tool kit with default parameters (Kang et al., 2010). Further analysis and graphical output of the results was done with R.

2.2.16 Leaf area measurements

For leaf area measurements normal colour photographs taken with a Sony Alpha DLSR camera were transformed into binary pictures with ImageJ. Pictures were normalized to a standard scale. The calculation of pixels of the plant leafs in this binary picture was a measure for the leaf area.

## 2.2.17 Whole transcriptome analysis

After sequencing, basic quality control of the raw reads was performed. This included adapter trimming and clipping with the FASTX-toolkit using standard parameters. Prior to mapping, quality was reviewed with fastQC. Samples with sufficient quality were mapped to a reference genome, which was specific for individual accessions, using TopHat2 and Bowtie2, allowing for 2 mismatches by default. After mapping, read counts were estimated with QualiMap. Differential gene expression analysis was performed in R, using the DEseq package.

2.2.18 Plant irradiation in sun simulators

For simulation of natural-like solar radiation, exposure chambers that are technically equipped for realistic simulation of natural climate and radiation, have been used (Thiel et al., 1996; Müller et al., 2013). The so called sun simulators are located at the Research Unit Environmental Simulation at the Helmholtz Zentrum München. By a combination of metal halide lamps (HQI/D, 400 W, Osram, München, Germany), quartz halogen lamps (Halostar, 300 W and 500 W, Osram, München, Germany), blue fluorescent (TLD 18, 36 W, Philips, Amsterdam, The Netherlands) and UV-B fluorescent tubes (TL12, 40 W, Philips, Amsterdam, The Netherlands) a simulation of natural sun-like spectra, ranging from 280 nm to 850 nm, can be achieved (Thiel et al., 1996; Müller et al., 2013). Additionally, various filters, including different borosilicate and lime glass filters, acrylic glass and a filter consisting of a water layer, are used to exclude different wavelengths, remove excessive heat generated by the lamps, and to generate different UV-B scenarios (Thiel et al., 1996; Müller et al., 2013).

The standard growth conditions were set to resemble the main growing season of

Arabidopsis in April/May. The conditions were: 14 h daylight; 21 °C at day, 10 °C at night; relative humidity at day 60%, at night 80%; PAR at day 340 µmol m^{-2} s^{-1} , at night 0. Natural diurnal variations were also taken into account by switching on/off different groups of lamps. Three different UV-B conditions were applied, as well as a control environment devoid of UV-B. UV-1 conditions correspond to 150 mW / m^2 biologically effective UV-B, UV-2 conditions correspond to 230 mW / m^2 biologically effective UV-B and UV-3 conditions correspond to 300 mW / m^2 biologically effective UV-B (Fig. 2-1). UV-B conditions mimic realistic conditions, corresponding to UV-B conditions beginning of May in Berlin in Germany (UV1), Rome in Italy (UV2) and Athens in Greece (UV3).

Directly after sowing on standard soil, plants were kept in the described UV-B conditions until siliques were fully developed and ready to be harvested, hence each plant spent the whole life cycle in the given conditions.

A

B
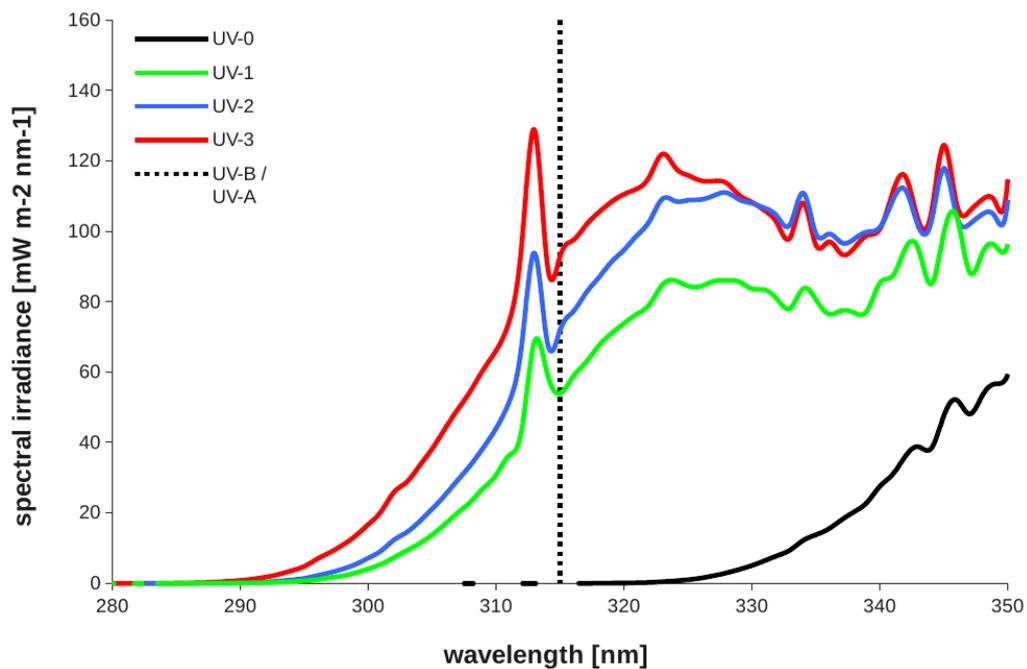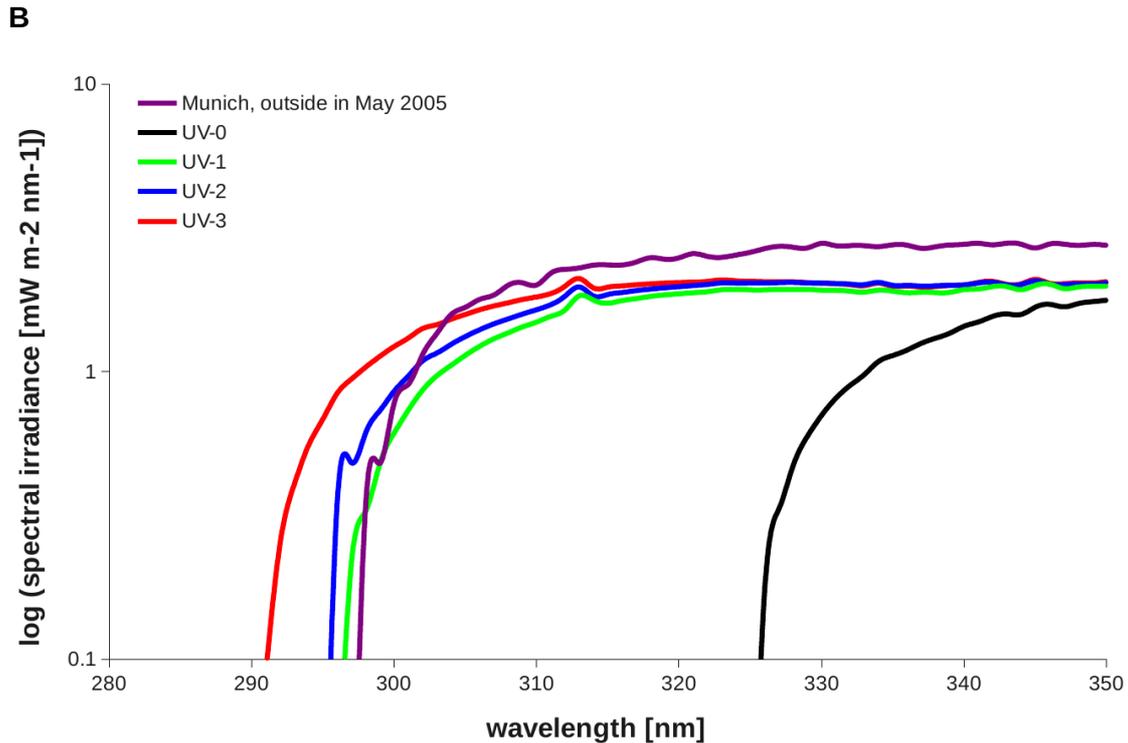
Fig. 2-1. Spectrum of UV-B irradiance in different compartments of the sun simulators. The spectral irradiance in the UV-B/UV-A range of three different UV-B conditions and one control condition without UV-B that were applied in the sun simulators are depicted (measurements were taken and provided by Dr. Andreas Albert and Dr. Barbro Winkler [Helmholtz Zentrum München]).

2.2.19 Sampling of plants grown in the sun simulators

Plants were kept for their whole life cycle in the described conditions, and after harvesting of seeds of self pollinated plants, the seeds were sown out in UV-free conditions and tissue propagated from these plants was used for sequencing. Hence, mutations were always estimated in the offspring of treated plants (Fig. 2-2). In total six different genotypes – Col-0, *uvr2*, *uvr3*, *uvr8*, *tt4* and *uvr2/uvr3* - and three UV-B and one mock condition were used. For Col-0 and *uvr2/uvr3* mutations were followed for three generations (Fig. 2-2; Tab. 2-1). For every genotype and condition 15 plants were grown of which 5 were sequenced (Tab. 2-1). Sequencing was carried out using Illumina technology as described (see 2.2.8). Five libraries were pooled per Illumina sequencing lane.

Sequencing
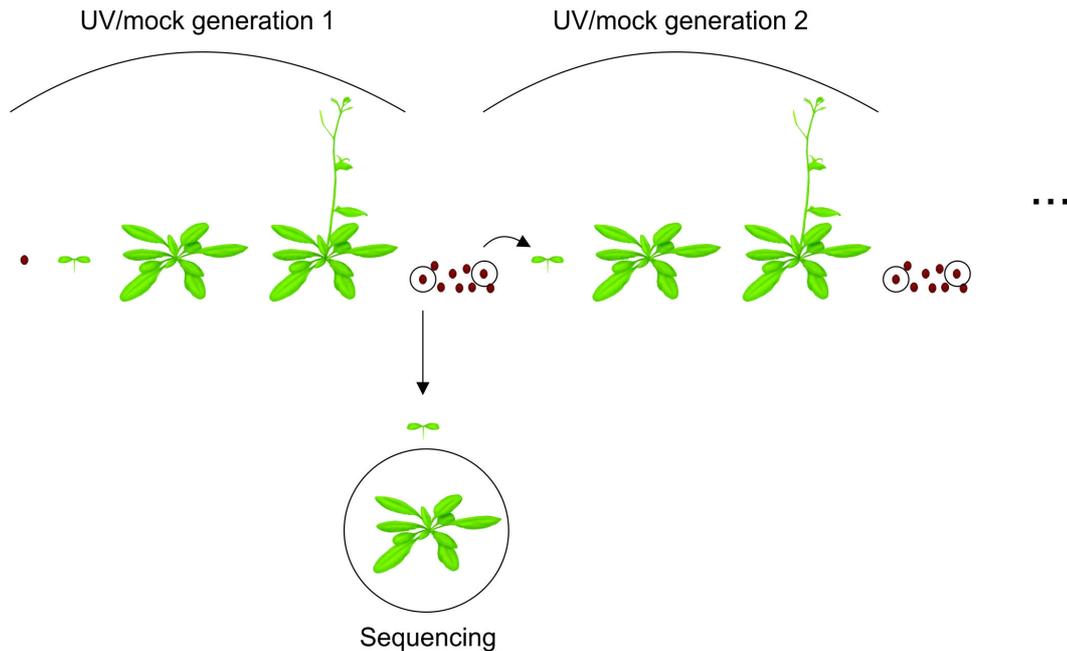
Fig. 2-2. Experimental procedure for identification of UV-B caused mutations. Plants were grown for a whole life cycle in the same conditions. At seed harvest one seed was randomly selected to remain under the same conditions for an additional life cycle or generation, respectively, and another seed was selected to generate plant tissue in UV-free conditions for sequencing

2.2.20 SNP calling

For identification of SNPs the SHOrt REad (SHORE) mapping and analysis program for read data produced on the Illumina platform was used (Ossowski et al., 2008). The 100 bp paired end reads were mapped to the Col-0 reference genome with TAIR10 annotation allowing for 10% mismatches. The genome consensus matrix was generated for each genome and every position in each genome was divided into different classes, depending on coverage and allele frequency:

- homozygous wt allele: coverage > 20x and wt allele frequency > 0.9
- homozygous mutant allele: coverage > 20x and mutant allele frequency > 0.9
- heterozygous: coverage > 20x and mutant allele frequency between 0.3 and 0.9
- position not considered, coverage too low: coverage < 20x
- other: coverage > 20x but 20% of reads have "N"
- undefined: coverage > 20x but allele frequency between 0.1 and 0.3

Each individual list of genome positions was filtered against 9 other genome position lists from the same genotype from the first generation as background filtering in order to identify new and unique mutations. A mutation was discarded when:

- 1 of 9 background genomes heterozygous or homozygous for the same variant
- 2 of 9 background genomes undefined at that position
- 3 of 9 background genomes defined as "other" (reads with too many "Ns")
- 4 of 9 background genomes with low coverage
- 5 of 9 background genomes fulfil one of the three above mentioned criteria

For background filtering of plants from generation 2 and 3 as well as from UV1 and UV2 treatments, the genome position lists were filtered against 9 randomly chosen genome position lists from the first generation from UV0 and UV3 treatment.

Tab. 2-1. Overview of samples for analysis of the mutagenic potential of natural-like UV-B. In total 120 genomes were sequenced on a genome-wide scale. Three UV-B conditions and one UV-B-free condition were used (UV0 - UV3), as well as several mutant genotypes. G1=first irradiated generation, G2=second irradiated generation, G3=third irradiated generation.

		G1	G2	G3
Col-0	UV0	5	5	5
	UV1	5		
	UV2	5		
	UV3	5	5	5
uvr2/uvr3	UV0	5	5	5
	UV1	5		
	UV2	5		
	UV3	5	5	5
uvr2	UV0	5		
	UV3	5		
uvr3	UV0	5		
	UV3	5		
tt4	UV0	5		
	UV3	5		
uvr8	UV0	5		
	UV3	5		

3 Results

3.1 Natural variation of *Arabidopsis thaliana* in response to UV irradiation

3.1.1 Phenotypic screen for natural variation in response to UV irradiation

In order to identify and quantify natural variation of *A. thaliana* in response to UV irradiation, 14-days-old plants were grown for 10 consecutive days either under UV free conditions or supplementary UV radiation. Afterwards, three global growth related traits - leaf length, leaf area and fresh aerial biomass - were measured in a total of 345 accessions. All three traits revealed a broad diversity of responses among accessions, ranging from very sensitive accessions to relatively resistant accessions with almost no change relative to mock (UV free) conditions (Fig. 3-1).

The range of responses was greatest for the biomass trait, with the most sensitive accession exhibiting a reduction to 22% after UV treatment compared to mock treated plants, whereas the most insensitive accession showed an increase to 152%. The range for the leaf length trait was smallest, spanning from 40% to 97%. Leaf area values were in the range of 19% to 123%. All three traits showed a strong and significant correlation (Tab. 3-1.).

The values of three reproducibly sensitive accessions, T1080, C24 and Ba-1, and three insensitive accessions, Lip-0, Got-7 and Mh-0, as well as Col-0 as the reference accession are depicted in more detail (Fig. 3-2.).

Enhanced sensitivity or insensitivity did not correlate with latitude or longitude, suggesting that the response towards UV under the given experimental conditions was not connected to a geographical parameter.

A

B

C
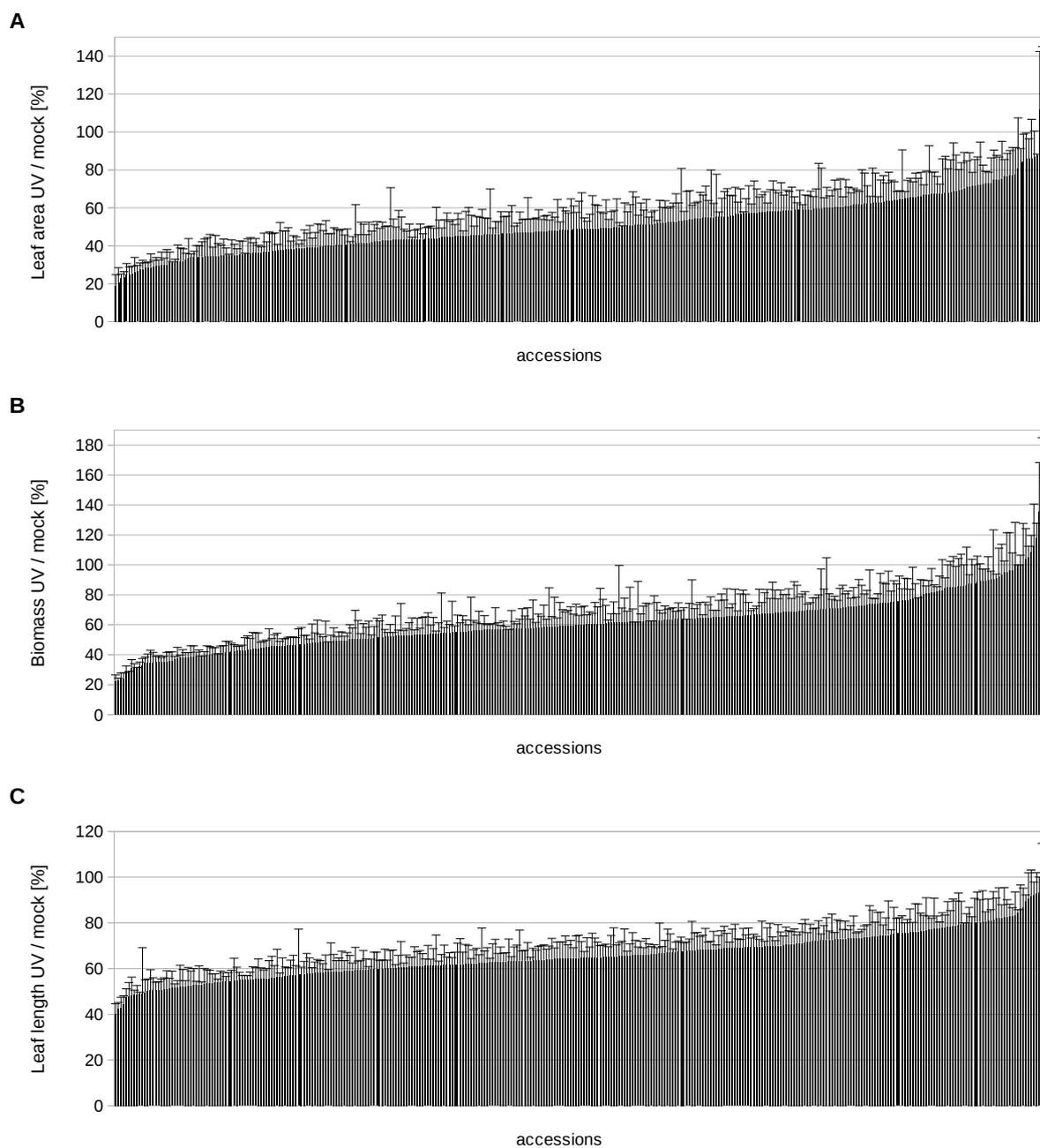
Fig. 3-1. Natural variation screen of *A. thaliana* in response to UV irradiation. Three growth related traits were measured: (A) - leaf area, (B) - biomass, (C) - leaf length. N=2-4 plants. Each bar represents the mean of one accession, error bars=SD.

Tab. 3-1. Pearson-Correlation between UV related traits. All correlations are significant, tested at α=0.05.

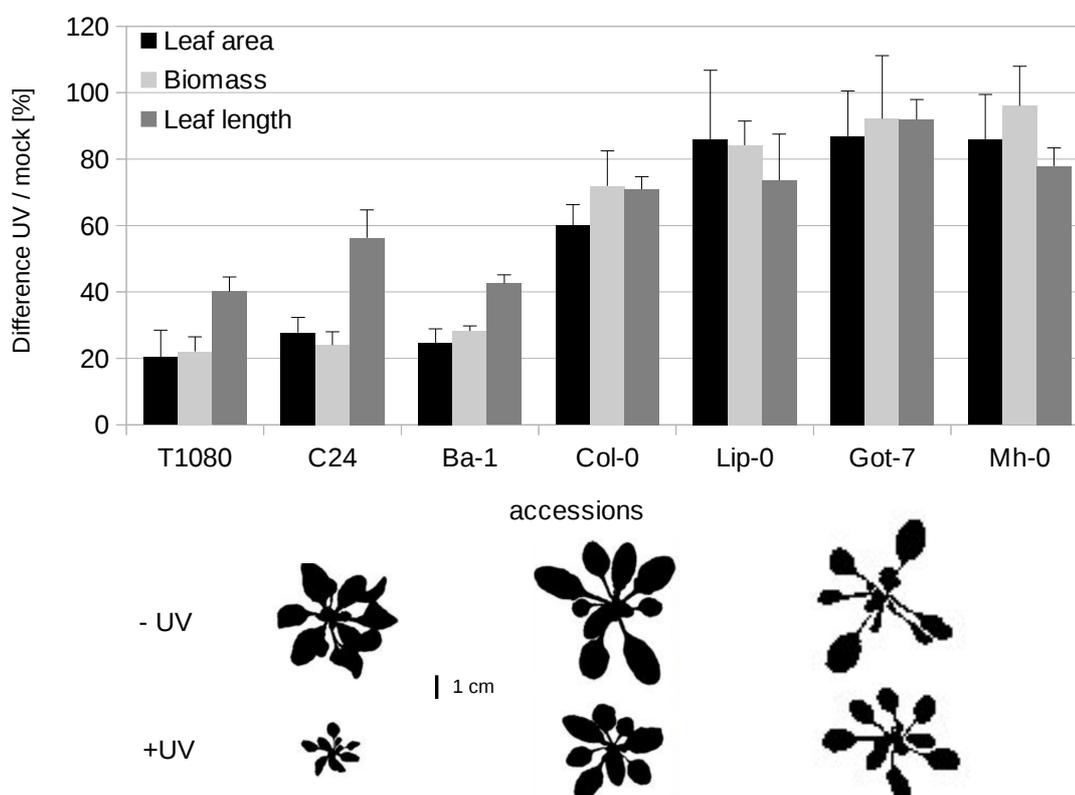| Leaf length	Biomass	Leaf area	
1	0.76[a]	0.86[b]	Leaf length
	1	0.83[c]	Biomass
		1	Leaf area

a,b,c: p-value < $2.2e^{-16}$

Fig. 3-2. Natural variation of selected *A. thaliana* accessions in response to UV irradiation. Three growth related traits were measured. N=2-4 plants. Each bar represents the mean of one accession, error bars=SD. Images below represent typical rosettes of UV free and UV treated plants, from left to right C24, Col-0 and Got-7.
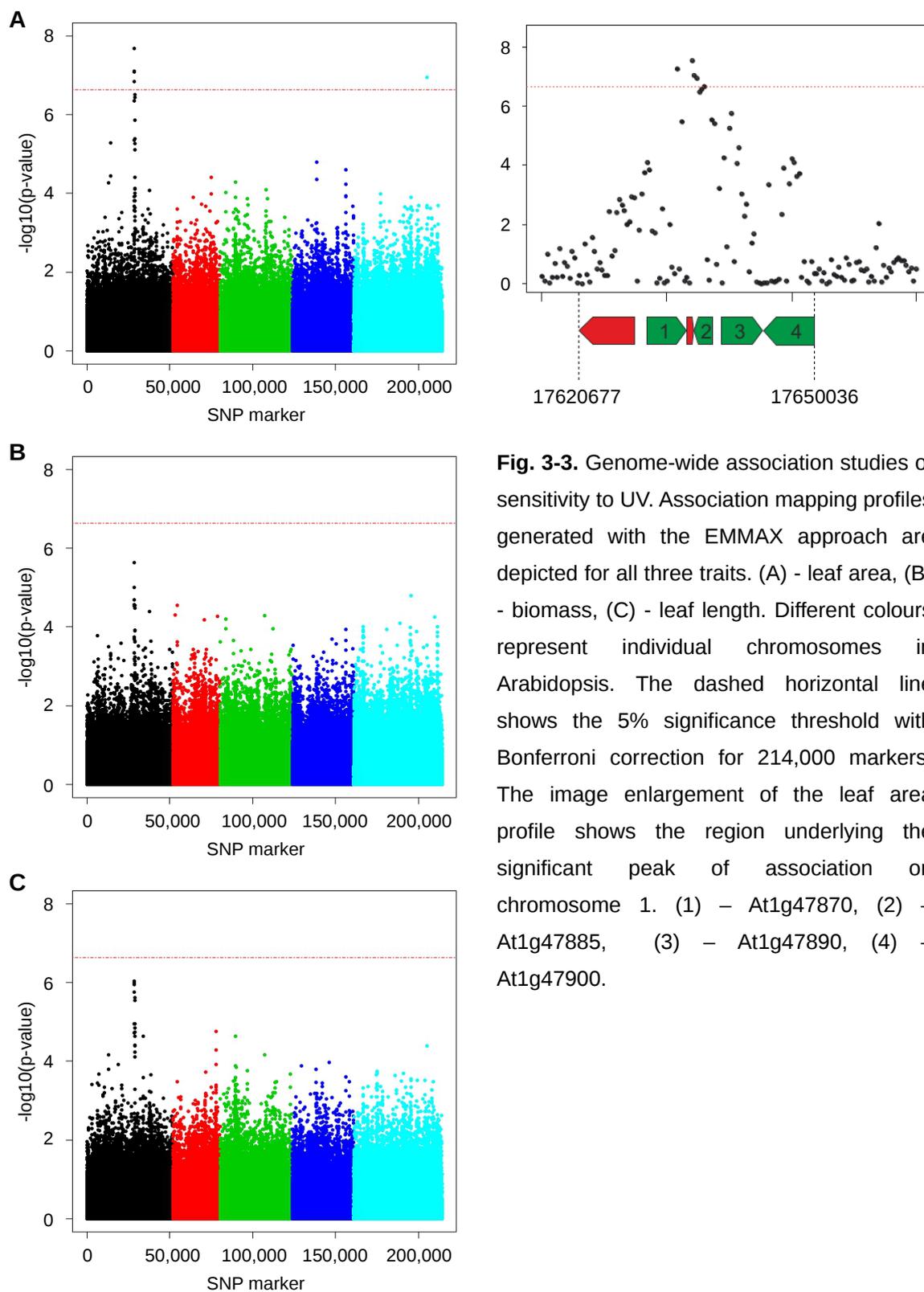
3.1.2 Genome-wide association studies

With the previously described data genome-wide association studies (GWAS) have been employed to identify genes causing or affecting the diverse sensitivity to the same UV treatment among the accessions tested (Fig. 3-3). All accessions used in this study were genotyped with the 250k SNP chip array by Affymetrix with the data publicly available (Kim et al., 2007; Atwell et al., 2010; Li et al., 2010). Every accession contained 214,051 markers, which led to a marker density of about one marker per 500 bp. This high resolution allowed association mapping on the level of individual genes.

Although no geographical effect or correlation could be identified in the sensitivity among accessions, the EMMAX approach for association mapping was applied to take population structure into account (Kang et al, 2010).

In all three traits a strong peak of association on chromosome 1 was detected. Yet only for the leaf area trait some SNP markers underlying this peak were significant (Bonferroni corrected significance threshold, P<0.05). In the region underneath of or surrounding these significant markers 6 genes are located (Fig. 3-3). The first red box represents a transposable element gene, the green box marked as #1 represents a gene encoding an E2F transcription factor (At1g47870), #2 encodes a putative ribonuclease inhibitor (At1g47885), #3 encodes a receptor like protein (At1g47890) and #4 encodes a protein with unknown function (At1g47900). The red box in between #1 and #2 represents a pseudogene (At1g47880). All significant markers within this peak showed a strong linkage disequilibrium, that is they are most likely inherited together to their offspring and form a haplotype group. In general this region is highly polymorphic with hundreds of SNPs and smaller insertions and deletions relative to Col-0.

Another single marker passed the significance threshold for the leaf area trait on chromosome 5. Underneath this marker a single gene with an unknown function (At5g53905) is located.

As the association mapping approach with a linear model and a calculation of probabilities after Wilcoxon showed almost exactly the same results, e.g. only one significant peak of association in only the leaf area trait, it can be safely assumed that population structure does not play an important role in the given dataset.

Fig. 3-3. Genome-wide association studies of sensitivity to UV. Association mapping profiles generated with the EMMAX approach are depicted for all three traits. (A) - leaf area, (B) - biomass, (C) - leaf length. Different colours represent individual chromosomes in Arabidopsis. The dashed horizontal line shows the 5% significance threshold with Bonferroni correction for 214,000 markers. The image enlargement of the leaf area profile shows the region underlying the significant peak of association on chromosome 1. (1) – At1g47870, (2) – At1g47885, (3) – At1g47890, (4) – At1g47900.

27

3.1.3 Single read sequencing of candidate region revealed by GWAS

Although almost all accessions that have been used throughout this study have been sequenced on a whole genome wide scale and the data is available via databases or data repositories, few small selected regions have been re-sequenced by single read (Sanger) sequencing. The pseudogene At1g47880 was sequenced in the accessions C24, Ba-1, T1080, Lip-0 and Mh-0. In all these accessions, as well as several others analysed in the Salk 1001 genome browser, At1g47880 has multiple stop codons, which do not lead to a reasonable gene product. Thus, At1g47880 is indeed a pseudogene in the reference genome of Col-0 and other tested accessions.

The gene At1g47890 was sequenced in few accessions, including the sensitive C24, T1080 and Ba-1, and the resistant Got-7, Lip-0 and Mh-0. Interestingly, four major differences have been detected, that were not included in the mentioned genome browser. One deletion of 12 bp was found in the promoter region ad two deletions and one insertion in the coding sequence. The two deletions in the coding region, one 12 bp long and one 9 bp long, maintain the reading frame, whereas the four bp insertion in the coding region leads to a frame shift and a premature stop codon (Fig. 3-4). As all three sensitive accessions that were initially re-sequenced were carrying all four aberrations, several more accessions were analysed to see a potential correlation between sensitivity to UV-B and these sequence aberrations. In total, 105 accessions were analysed by sequencing, cleaved amplified polymorphic sequences (CAPS) and simple sequence length polymorphism (SSLP) markers (Fig. 3-5) for the 12 bp deletion in the promoter region and the four bp insertion in the coding sequence. About half of the accessions carried both alterations. In only very few accessions these alterations were separated from each other. Taken together a weak but significant linear correlation between both the presence of the 12 bp deletion in the promoter region and the 4 bp insertion in the coding sequence with sensitivity existed for all three traits. For the leaf area trait and the 12 bp promoter deletion, for example, it was r=0.33 ($p_{\alpha=0.05}$=3.7 x 10^{-4}).

The gene At1g47900 was sequenced in C24 and revealed a 21 bp deletion compared to Col-0 reference, not annotated in the 1001 genomes dataset. This deletion allowed designing a SSLP marker, that has been used to detect the presence of this deletion in other accessions.

This deletion was detected in all accessions, but Col-0, irrespective of their UV sensitivity. Therefore it is unlikely that this deletion is causal for the phenotypes observed after UV treatment.

Fig. 3-4. DNA sequence variation in the gene At1g47890. Few selected accessions were sequenced by single read sequencing revealing a 12 bp deletion in the promoter region (upper panel) and a 4 bp insertion in the coding region (lower panel). Red = sensitive accession, green = resistant accessions.
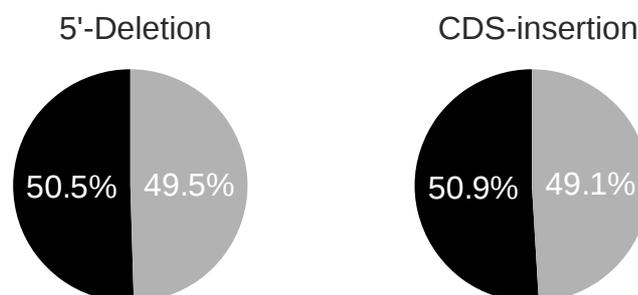
Fig 3-5. Frequencies of DNA sequence variations in the gene At1g47890. Presence of a 12 bp deletion in the promoter region and a 4 bp insertion in the coding region are shown. In total 110 accessions were analysed. Grey insertion/deletion present, black insertion/deletion absent.

3.1.4 Expression analysis of candidate genes revealed by GWAS

To elucidate whether differential expression of the candidate genes revealed by GWAS could be responsible for the diverse phenotypes observed, reverse quantitative transcriptase PCR (RT-qPCR) was employed. Samples were taken from 14 days-old seedlings irradiated with the same dose as described for the screen and harvested at the indicated time points (Fig. 3-6).

Expression was monitored in three sensitive (Ba-1, C24 and T1080) and three insensitive (Amel-1, Mh-0 and Lip-0) accessions and Col-0 as control. All four protein coding candidate genes, as well as the pseudogene At1g47880, were analysed (Fig. 3-6).
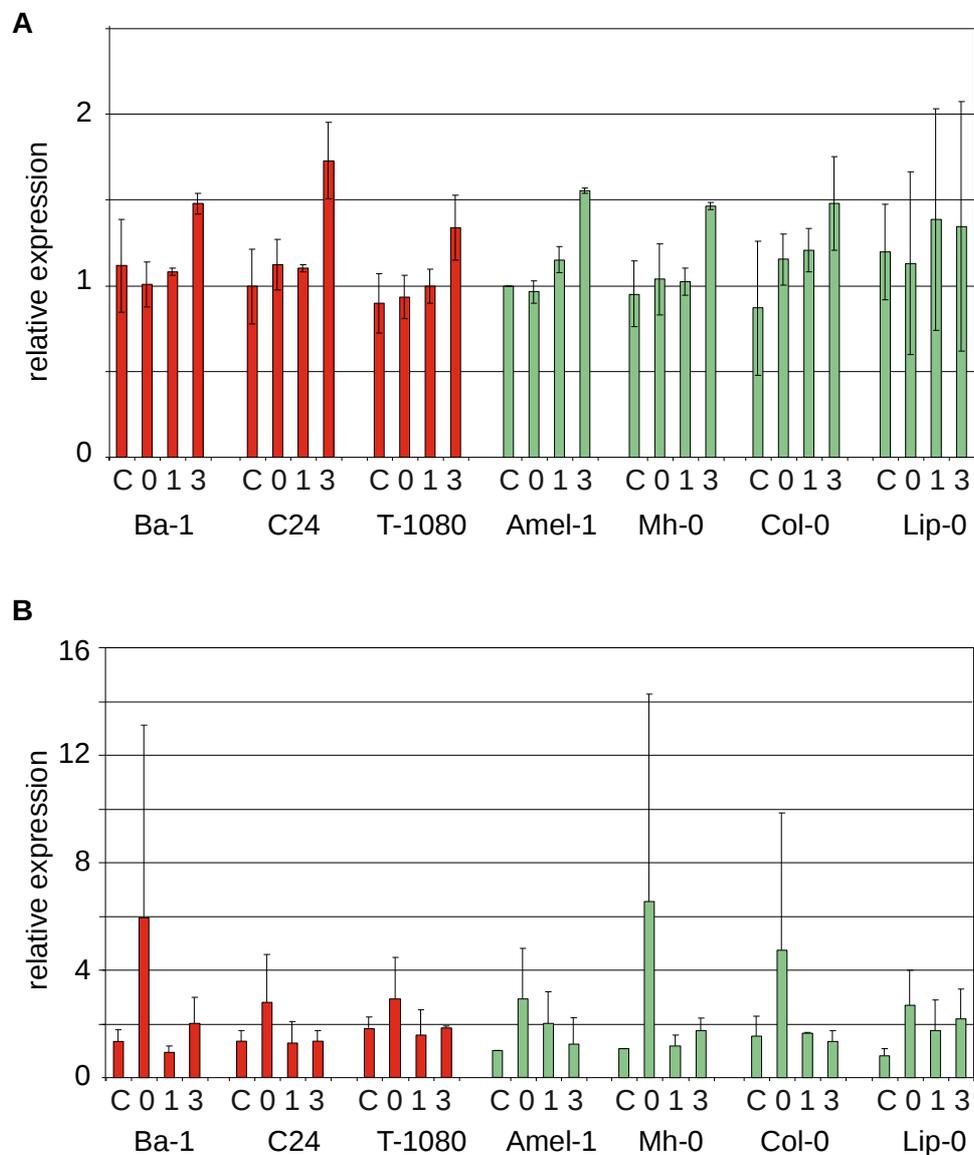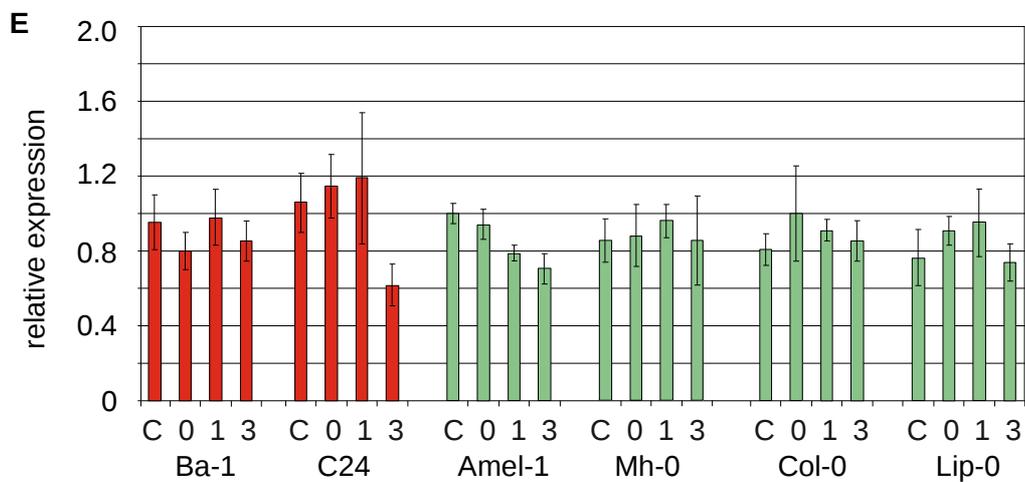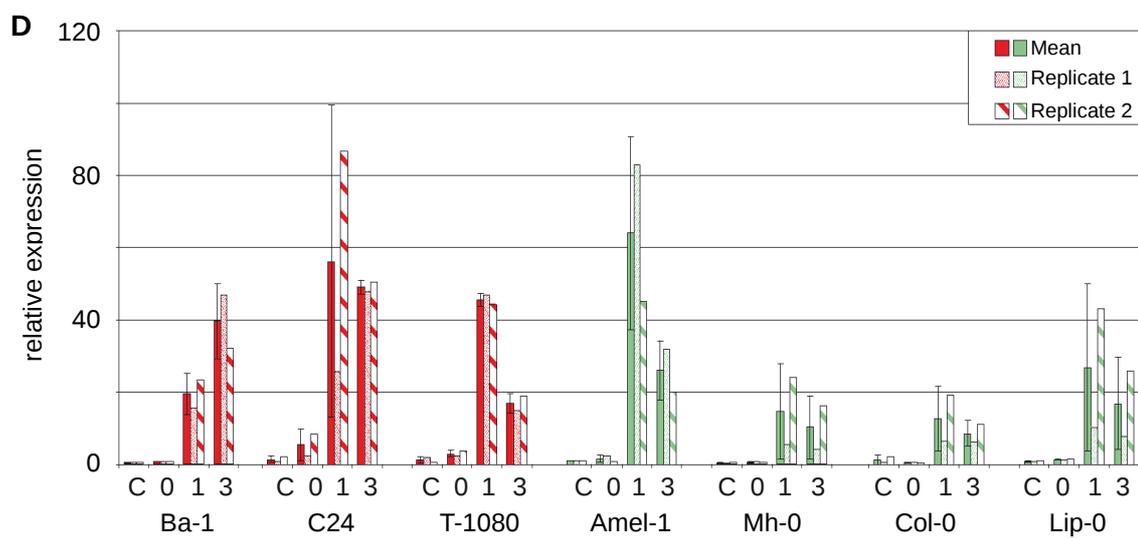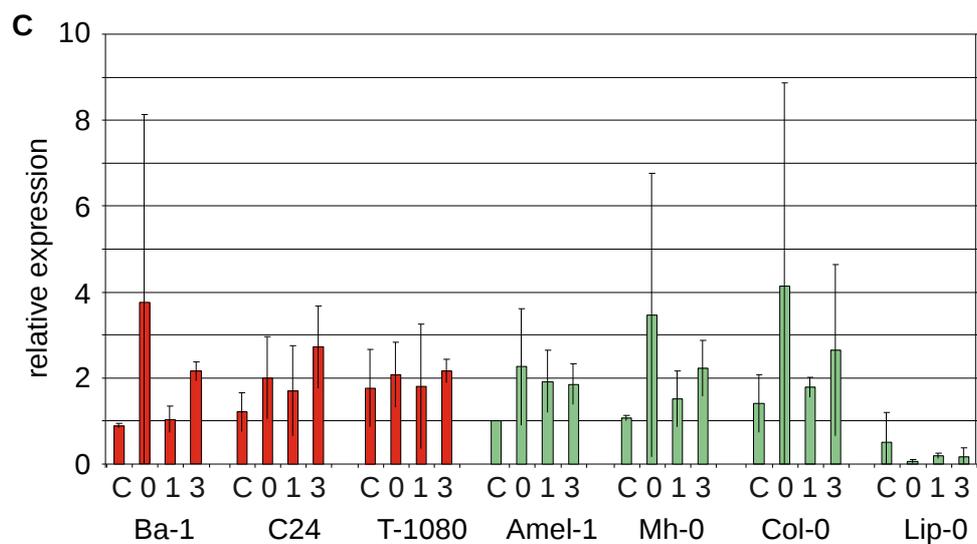
Fig. 3-6. Expression analysis of candidate genes revealed by GWAS. Each block of four columns represents one accession. Red = sensitive accession, green = resistant accession, C = mock control, 0 = no recovery after UV treatment, 1 = 1h recovery, 3 = 3 h recovery. Expression was normalized to Actin 7 (At5g09810). Values are means of two biological replicates with each three technical replicates. Error bars = SD.

For candidate gene #1 (At1g47870), no expression differences between sensitive and insensitive accessions could be observed (Fig. 3-6 A). Furthermore, for almost all accessions, but Lip-0, a very weak induction 3 hours after UV treatment was seen.

For the pseudogene (At1g47880) no obvious differences between sensitive and insensitive accessions could be identified (Fig. 3-6 B). In all accessions the expression seemed to be induced directly after the UV treatment, but as this differed strongly between individual replicates of the expression analysis no conclusion could be drawn at this point.

The same pattern of expression could be observed for the putative ribonuclease inhibitor (At1g47885) (Fig. 3-6 C). Again, no differences between sensitive and insensitive accessions were found, and for some accessions, like Ba-1, Mh-0 and Col-0, expression seemed to be induced upon UV irradiation, but differed vastly between replicates. The only truly deviant behaviour was seen in Lip-0, were expression seemed to be completely shut down.

The expression of the receptor like protein (At1g47890) showed the biggest differences of all candidate genes (Fig. 3-6 D). In all accessions it was strongly and significantly induced 1 and 3 hours after UV irradiation, with up to about 60-fold induction in C24 and Amel-1 one hour after treatment. However, no clear pattern between expression of this gene in sensitive and insensitive accessions could be observed. To further investigate this, several more accessions have been analysed for their expression after UV treatment (Fig. 3-7). Yet, also with more accession no clear correlation between the expression of At1g47890 and sensitivity or insensitivity could be identified.

The fourth candidate gene At1g47900 showed neither differences among accession, nor responsiveness to UV, except 3 hours post treatment in C24, where expression was slightly reduced (Fig. 3-6 E).
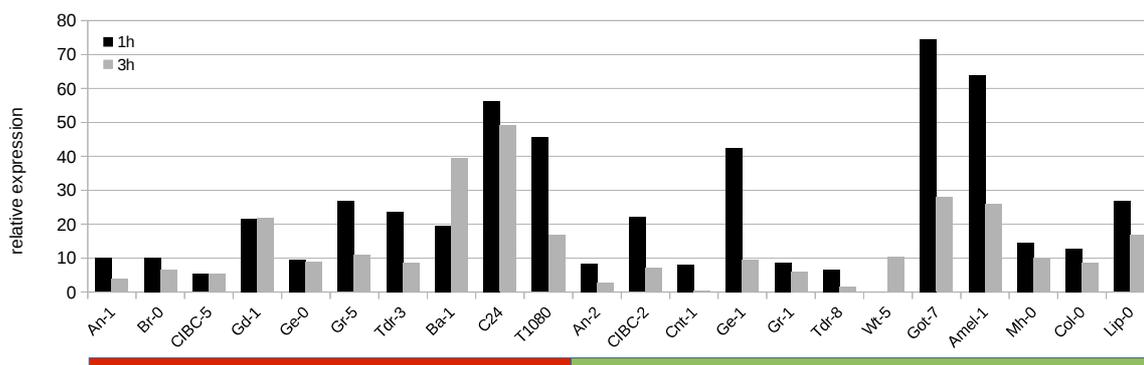
Fig. 3-7. Expression analysis of candidate gene At1g47890 in various accessions. Each block of two columns represents one accession. Red = sensitive accession, green = resistant accession, 1 = 1h recovery, 3 = 3 h recovery. Expression was normalized to Actin 7 (At5g09810).

3.1.5 Analysis of T-DNA mutants revealed by GWAS

For the four protein coding candidate genes underlying the peak of association in the leaf area trait in association mapping, homozygous T-DNA insertion mutants have been obtained. None of the mutants showed significant difference compared to Col-0 wild-type plants. Nonetheless, all mutants were significantly different compared to the C24 accession, which represented a UV sensitive genotype.
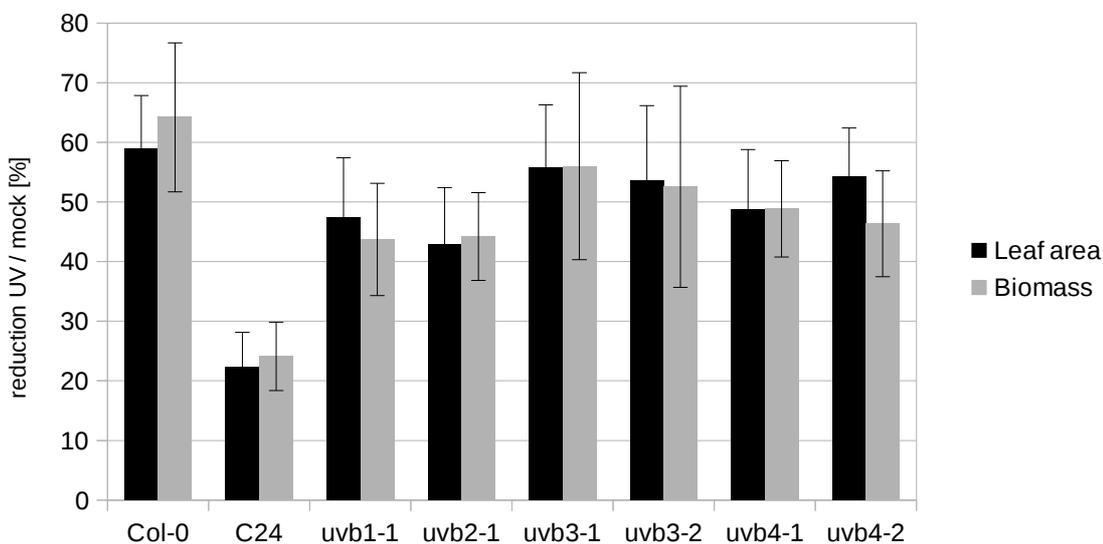
Fig. 3-8.. T-DNA insertion mutants of candidate genes identified by GWAS. Leaf area and biomass traits were measured. N≥15 plants. Each bar represents the mean of one genotype, error bars=SD.

3.1.6 Quantitative trait locus (QTL) mapping of UV sensitivity

Previous data suggested a more complex genetic interaction, instead of a single gene causing the observed phenotypes after UV treatment. In order to dissect the genetic interaction controlling these effects, QTL mapping in addition to GWAS was employed. A recombinant inbred line population (RIL), derived from a cross of the UV sensitive C24 and the UV resistant Col-0 accessions was used. 122 RILs were selected to cover the maximum genotype diversity and phenotyped for the leaf area and biomass traits (Fig. 3-9). QTL mapping was performed using the multiple QTL mapping (MQM) approach (Broman et al., 2003; Arends et al., 2010).
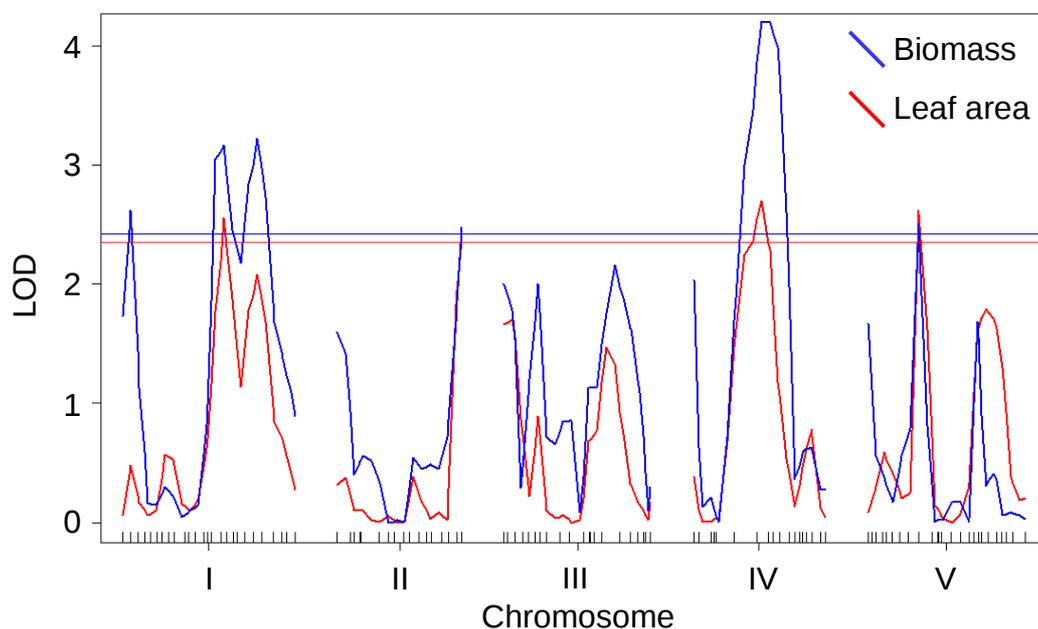
Fig. 3-9. QTL mapping of biomass (blue) and leaf area (red) changes in response to UV irradiation using a C24 x Col-0 RIL population. Markers along the five chromosomes are plotted against the logarithm of the odds (LOD) scores. Horizontal lines show the significance threshold by running 1000 permutations.

QTL mapping identified several loci throughout the genome for both traits that passed the significance threshold as well as few minor effect loci, although overall the LOD scores of these loci were low. Both traits had three significant QTL in common. One each on chromosomes 1, 4 and 5. The QTL on chromosome 1, that reached the significance threshold in both traits also overlapped with the region of association identified by the GWAS approach.

In order to analyse the influence of a single QTL on the UV sensitivity several near isogenic lines (NILs) were phenotyped (Fig. 3-10).
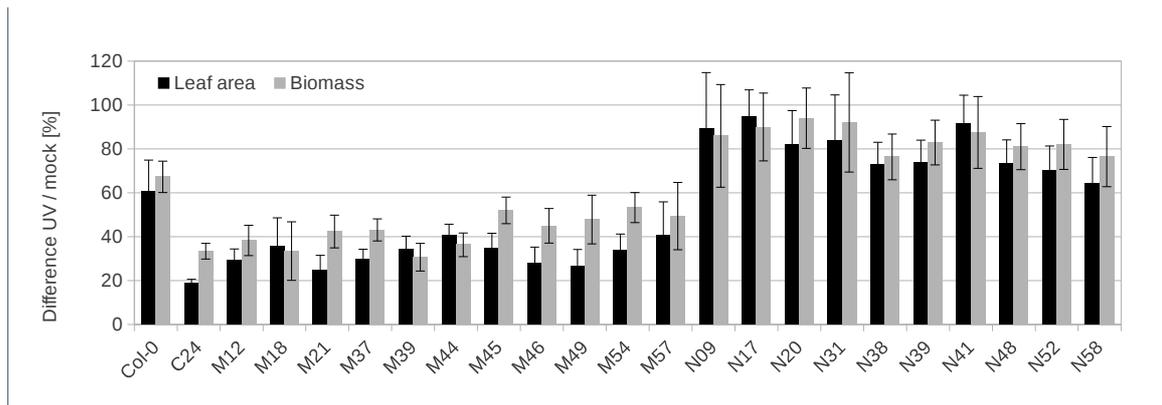
Fig. 3-10. Leaf area (black) and biomass (grey) changes of NILs in response to UV irradiation. Each bar represents the mean of one genotype. N≥15 plants , error bars=SD.

NILs starting with "M" have a C24 background throughout their genome and small introgressions of the Col-0 genotype covering one of the loci identified by QTL mapping. NILs starting with "N" have opposite genotypes, hence having Col-0 genome background and C24 introgressions.

Taken together, all lines with Col-0 background were more resistant to UV than lines with C24 background. They showed Col-0 wild-type phenotypes or often even outperformed their parental lines and no C24 introgression in Col-0 background provoked a C24-like phenotype, nor a partially sensitive phenotype.

On the other hand all lines with C24 background and a Col-0 introgression showed sensitivity. However, all sensitive NILs with C24 background were both significantly different compared to NILs with Col-0 background, as well as to their parental lines.

3.1.7 Whole transcriptome sequencing

As the expression analysis indicated that sensitivity to UV might be connected with expression of certain genes, like candidate gene At1g47890, but many other, yet unknown genes, a whole transcriptome sequencing approach (Illumina RNA sequencing) was conducted, which in contrast to previous microarray studies might identify novel transcripts, e.g. with low abundance (Ulm et al., 2004; Killian et al., 2007; Zhao et al., 2014).

The three sensitive accessions T1080, C24 and Ba-1, as well as the three insensitive accessions Lip-0, Got-7 and Mh-0 were selected (Fig. 3-2). Plants were irradiated with the same dose as for the phenotypic screening at a 14-days-old seedling stage. Samples were harvested after 3 hours of recovery after the treatment. In order to identify

common patterns of expression among sensitive versus insensitive accessions, analysis was carried out iteratively. First, the UV-B treated samples were compared to mock treated plants for each accession. Second, the identified differentially expressed genes in each accession were intersected with the other two sensitive or insensitive accessions, respectively. Third, the list of differentially regulated genes in all sensitive accessions was compared to genes differentially regulated in all insensitive accessions. Of the 24,185 genes that were accessible for analysis, around one quarter, in total 5421 genes (22.4%), showed differential regulation in all three sensitive accessions (Fig. 3-11). But only 315 of these genes were significantly differently regulated from all resistant accessions. According to GO term analysis these genes were enriched for kinase activity (p-value = 9.25 x 10^{-5}), transferase activity (p-value = 2.14 x 10^{-4}), DNA binding (p-value = 6.59 x 10^{-4}) and transmembrane transporter activity (p-value = 6.92 x 10^{-4}) indicating differentially expressed genes mainly in the signalling / regulatory part.
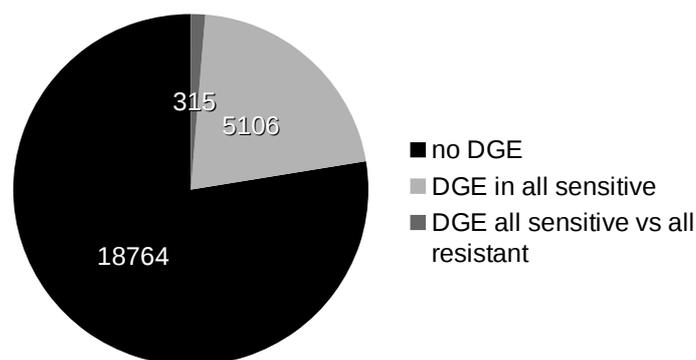
Fig. 3-11. Whole transcriptome analysis of UV treated accessions. Differential gene expression (DGE) was analysed in three UV sensitive and three UV resistant accessions 3 h after UV treatment with two replicates each.

3.1.8 Analyses of UV sensitivity of T-DNA insertion mutants derived from whole transcriptome analysis

Of the 315 genes that were identified by whole transcriptome analysis with different UV responses, as a first step some were selected to be analysed as T-DNA knock-out mutants by testing their UV sensitivity.

In order to further reduce the candidate gene list, several selection criteria were applied. These included fold change expression difference, genomic position with considering the QTL mapping and GWAS results, as well as functional connection to known UV-B genes and pathways, e.g. expression dependency on UVR8 and/or HY5 (Brown et al., 2005;

Oravecz et al., 2006; Kilian et al, 2007; Lee et al., 2007; Winter et al, 2007; Gonzalez-Besteiro et al., 2011; Obayashi et al., 2011). In a lot of cases selection was limited simply due to lack of T-DNA insertion mutants. Based on this, in total 50 lines were selected of which 45 could be propagated and genotyped or confirmed for their homozygosity.

Most lines showed wild-type like behaviour grown under the described UV treatment conditions, yet five lines differed markedly (Tab. 3-2; Fig. 3-12). Lines 32 and 40 had reduced leaf area to 68% and 64%, respectively, and biomass reduced to 79% and 63%, after UV treatment. On the other hand, lines 5, 22 and 25 showed an increased leaf area to 170%, 171% and 195%, and an increased biomass to 148%, 171% and 152%, respectively. Some additional information of these candidate genes are depicted in more detail (Tab. 3-2).

Tab. 3-2. Candidate genes suggested by whole transcriptome sequencing which showed diverse phenotypes in response to UV irradiation.

#	5	22	25	32	40
Gene ID	AT4G22530	AT1G48240	AT1G56510	AT1G76360	AT2G37040
Function	Methyl-transferase	SNARE protein	TIR-NB-LLR protein	Protein kinase	Phenylalanin-ammonium-lyase 1
Fold change in sensitive acc.	8.10	0.58	4.07	2.15	6.54
Fold change in resistant acc.	14.30	0.47	0.47	3.00	3.48
T-DNA mutant [NASC ID, insertion]	N655074; exon	N664298; intron	N660251; exon	N65580; exon	N661248; exon
Difference in leaf area UV / mock [%]	170.3	170.7	194.8	68.1	64.2
Difference in biomass UV / mock [%]	148.2	171.5	151.6	79.3	63.1
Additional features*	COP1 and HY5 dependent	Co-localization with QTL and GWAS	Co-expression with UV-B dependent MAPK	Co-expression with UV-B dependent MAPK and At1g47890 (GWAS candidate gene)	COP1 and HY5 dependent

*(Brown et al., 2005; Oravecz et al., 2006; Kilian et al, 2007; Lee et al., 2007; Winter et al, 2007; Gonzalez-Besteiro et al., 2011; Obayashi et al., 2011)
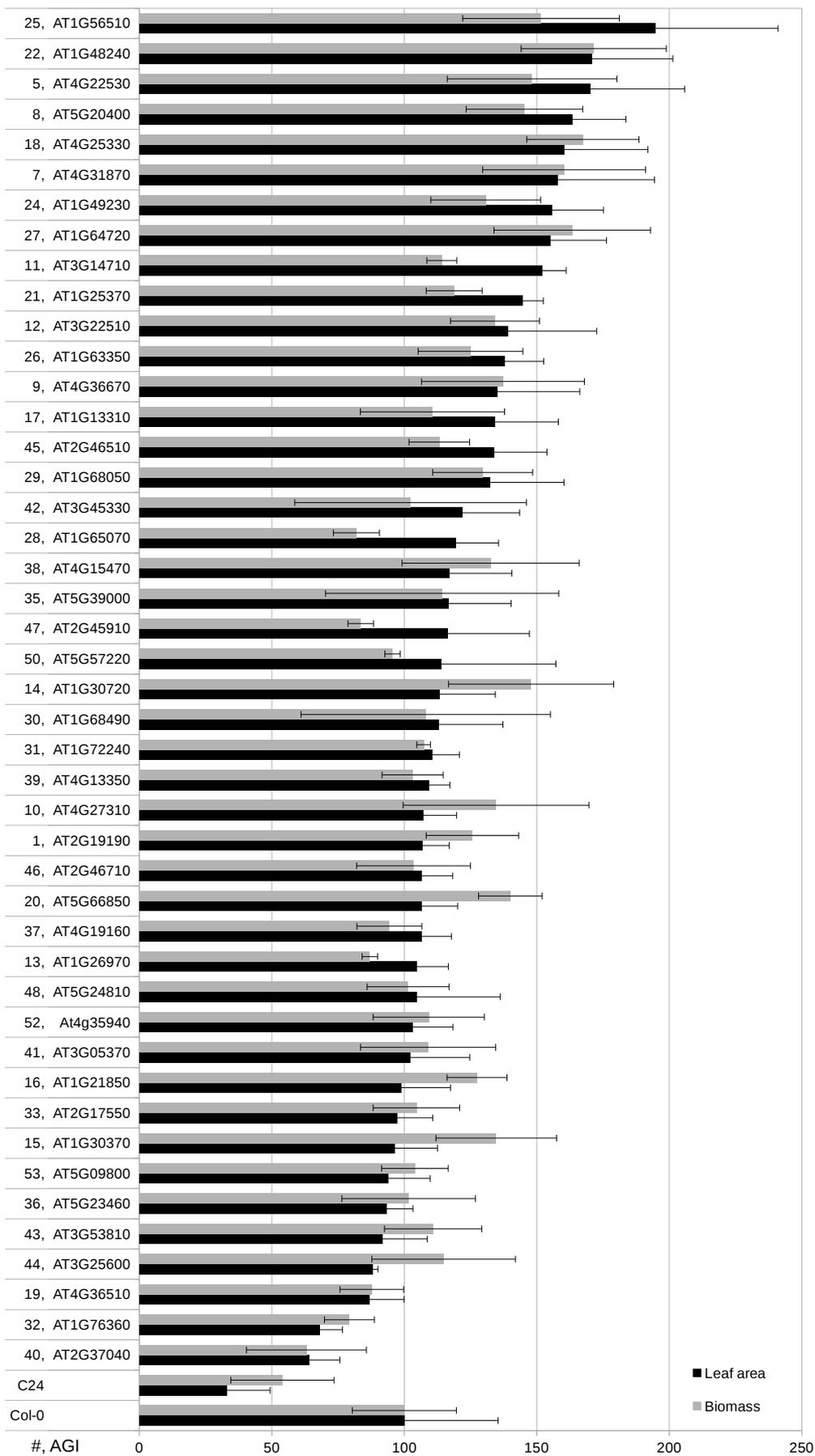
Difference UV / mock (normalized to Col-0l [%])

Fig. 3-12. Phenotype analysis of UV sensitivity of selected candidate genes suggested by whole transcriptome sequencing. The growth related traits leaf area and biomass were measured. N≥4 plants. Each bar represents the mean of one accession, error bars=SD. Data sorted by increasing leaf area values.

3.2 The mutagenic potential of natural-like UV-B

3.2.1 Sampling and identification of mutations

In order to analyse the mutagenic potential of natural-like UV-B radiation, a mutation accumulation project was initiated. To address different aspects, three UV-B conditions with increasing UV-B dosages (UV1, UV2 and UV3), as well as one control condition devoid of UV-B radiation (UV0) were applied. UV-B dosages mimic realistic UV-B conditions, corresponding to UV-B conditions beginning of May in Berlin in Germany (UV1), Rome in Italy (UV2) and Athens in Greece (UV3) with biologically effective UV-B dosages of 150 mW / m^2 UV-B$_{be}$, 230 mW / m^2 UV-B$_{be}$ and 300 mW / m^2 UV-B$_{be}$, respectively.

In addition to Col-0 wild-type plants, several mutant genotypes were analysed. These included the UV-B photoreceptor mutant *uvr8*, the flavonoid impaired mutant *tt4*, the photolyase mutants *uvr2* and *uvr3*, and a photolyase double mutant *uvr2/uvr3*. To monitor mutations on a whole genome wide scale, mutations were identified by a sequencing approach.

In total, 120 genotypes were sequenced. On average, samples had a sequencing depth of over 36 million reads. This yielded an average genome coverage of over 30x. Based on the results of mapping and SNP calling 20 low quality genomes were discarded from the analysis due to, e.g. low number of genomic positions accessible for SNP calling and analysis. Taken together, 100 genomes were analysed with on average over 70% of each genome accessible for identification of SNPs (Tab. 3-3).

Tab. 3-3. Overview of samples for analysis of the mutagenic potential of natural-like UV-B. In total 100 genomes were analysed. Three UV-B conditions and one UV-B-free condition were applied (UV0 – UV3). Beside Col-0 wild-type plants with *uvr2/uvr3, uvr2, uvr3, tt4* and *uvr8* also several mutant genotypes were used. G1=first irradiated generation, G2=second irradiated generation, G3=third irradiated generation.

		G1	G2	G3
Col-0	UV0	4	4	4
	UV1	5		
	UV2	4		
	UV3	5	4	4
uvr2/uvr3	UV0	4	3	2
	UV1	4		
	UV2	4		
	UV3	4	4	4
uvr2	UV0	5		
	UV3	4		
uvr3	UV0	4		
	UV3	5		
tt4	UV0	5		
	UV3	5		
uvr8	UV0	4		
	UV3	5		

3.2.2 Mutational spectrum

In the 100 sequenced samples that were analysed for their mutational spectrum in total 1980 mutations were identified, of which 1467 were heterozygous and 513 were homozygous. This resulted in a ratio of heterozygous to homozygous mutations of about 3 to 1. Since mutations were identified in the progeny of irradiated plants, 2/3 heterozygous mutations and 1/3 homozygous mutations would be expected, which yields a ratio of 2:1, assuming a normal mendelian inheritance of mutations. The experimentally found ratio differed from the expected one and was slightly skewed towards heterozygous mutations. To confirm that the identified mutations were indeed true mutations, 25 were randomly selected and validated by single read sequencing. All 25 mutations could be verified.

41

Depending on the type of change of the nucleotide base caused by the point mutation, transversions and transitions can be distinguished (Futuyma, 2005). Transversions refer to the substitution of a purine (A or G) for a pyrimidine (T or C) or vice versa. Transitions are point mutations that change a purine nucleotide to another purine or a pyrimidine to another pyrimidine. There are twice as many possibilities for transversions than for transitions.

Of the total number of mutations identified, 89.2% were transitions, 9.6% transversions and 1.2% single base pair deletions (Fig. 3-14). The ratio of transitions over transversions for all SNPs identified was estimated as 9.3 to 1, whereas it was previously identified as a spontaneous mutation rate in UV-B free conditions as 2.4 to 1 (Ossowski et al., 2010; Fig. 3-14).
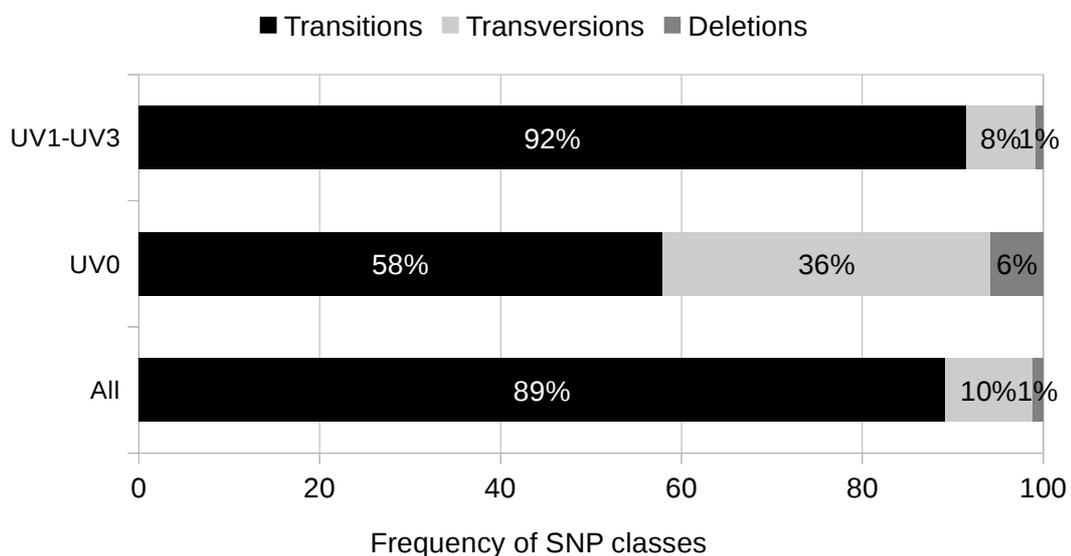
Fig. 3-14. Mutation spectrum of all identified mutations. Depending on the change of the nucleotide base, mutations were divided into the three groups: transitions, transversions and deletions.

If SNPs that occurred under mock conditions (still including natural-like UV-A radiation) were excluded from analysis, the ratio increased to 12 to 1 for transitions over transversions, whereas in mock plants only, it dropped down to 1.6 to 1.

Compared to UV-B free conditions, GC to AT transitions were specifically enriched by UV-B radiation (Fig. 3-15; Ossowski et al., 2010). The frequencies of different SNP classes that occurred under UV-B free conditions were in the same range as previous estimations of spontaneous mutations of greenhouse grown UV-B free plants.
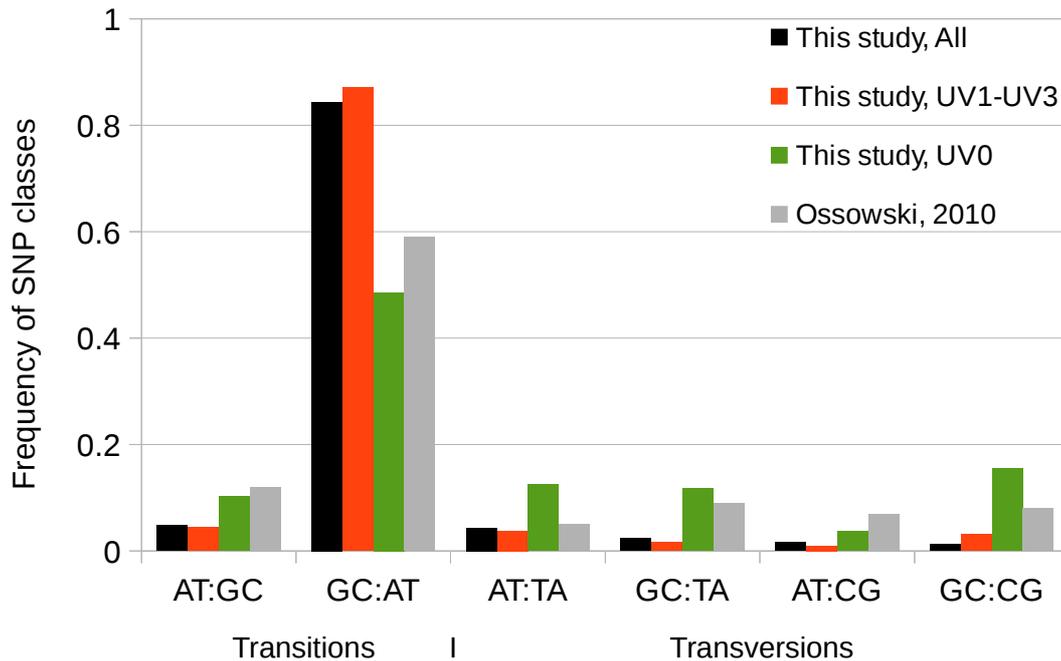
Fig. 3-15. Frequency of different mutation classes. All possible transitions and transversions classes are shown and compared to previously identified rates estimated in UV-B free conditions (Ossowski et al., 2010).

3.2.3 Genomic distribution of mutations

In order to determine whether mutations were evenly and randomly distributed over the genome or certain genomic regions were favourably mutated, the genomic distribution of identified mutations was compared to the *A. thaliana* reference genome.

The *A. thaliana* genome in the TAIR10 annotation consists of 50.83% genes, 29.93% intergenic regions and 21.37% transposable elements (TEs). Of all identified SNPs 33.7% were found in genes, 25.4% in intergenic regions and 40.0% in TEs (Fig. 3-16). Compared to the genome composition, the occurrence of SNPs differs strongly and is skewed from genes towards TEs. The occurrence of SNPs in intergenic regions as well as pseudogenes resembles the expected ratio.

Across the five chromosomes SNPs were evenly distributed, with no chromosome being more prone to mutations than another (Fig. 3-17).
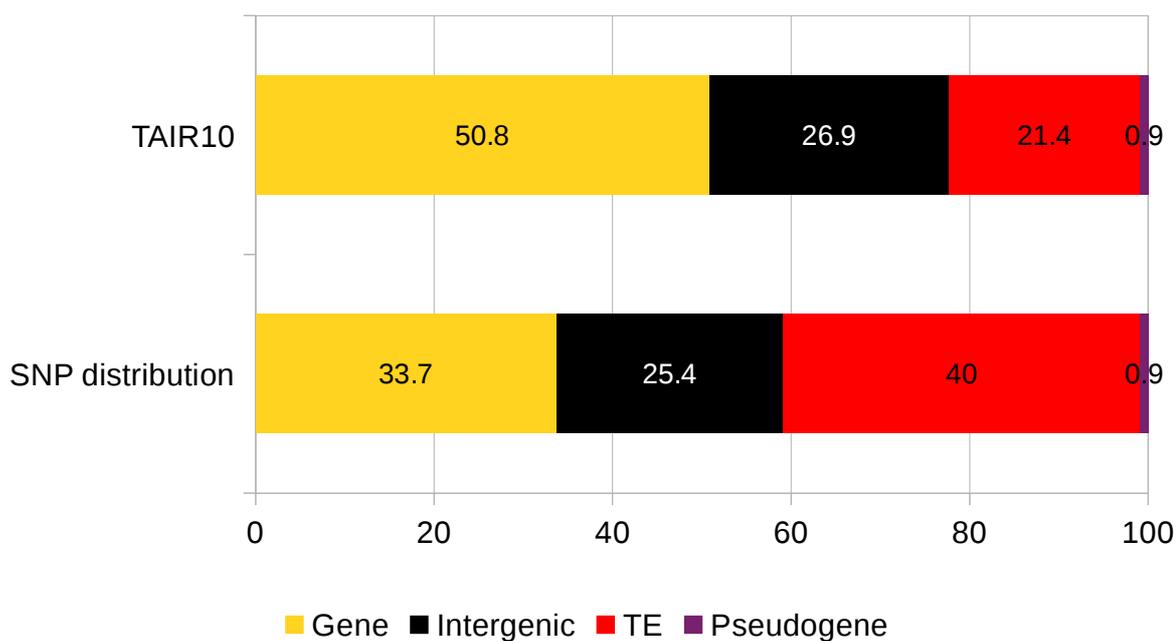
Fig. 3-16. Genomic distribution of mutations. Comparison between the distribution of identified mutations and the genome composition of *A. thaliana* TAIR10. The golden path length (~ 119 Mbp) of the TAIR10 annotation was used as reference.
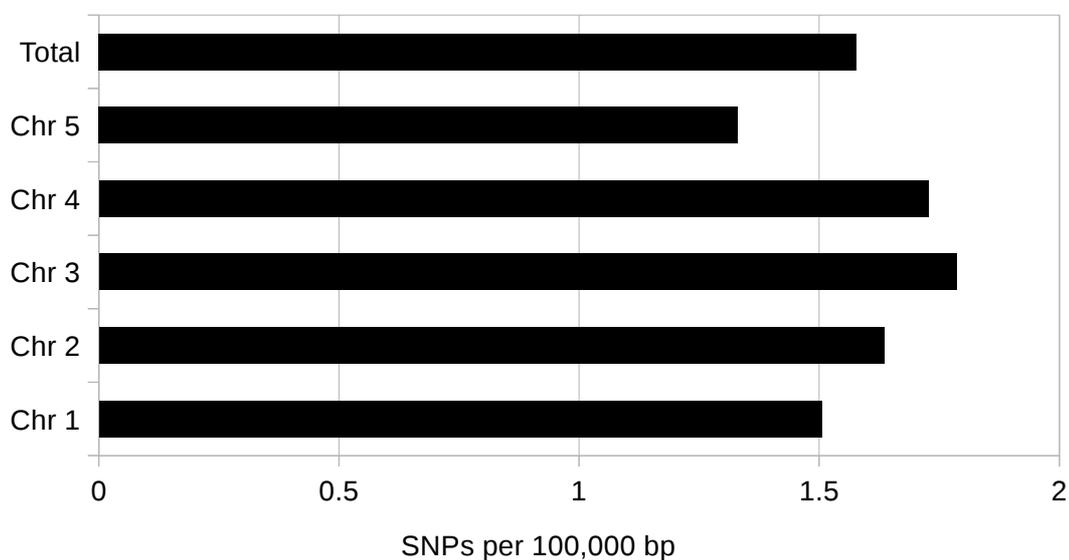
Fig. 3-17. Mutation rates for individual chromosomes. The mutation rates per chromosome per 100,000 bp were calculated, considering all identified mutations and the golden path length of the TAIR10 annotation as reference.

Although the mutation rates for the five chromosomes did not differ significantly, as the minor differences represented normal variation, the frequency of SNPs along each chromosome was calculated and suggested a non-random distribution (Fig. 3-18). The frequency of SNPs was highest in pericentromeric regions and lower along chromosomal arms. In addition, the frequency of SNPs seemed to run almost parallel to the frequency of TEs. This could be observed for all five chromosomes and also the heterochromatic knob on chromosome 4, in which the gene frequency decreased and both the TE frequency and the SNP frequency increased (McCombie et al., 2000). The TE frequency and the SNP frequency were positively correlated with a Pearson value r = 0.78 (p-value < $1e^{-22}$), whereas the occurrence of SNPs and genes were negatively correlated, and rather mutually exclusive with a Pearson value r = -0.71 (p-value < $1e^{-22}$).
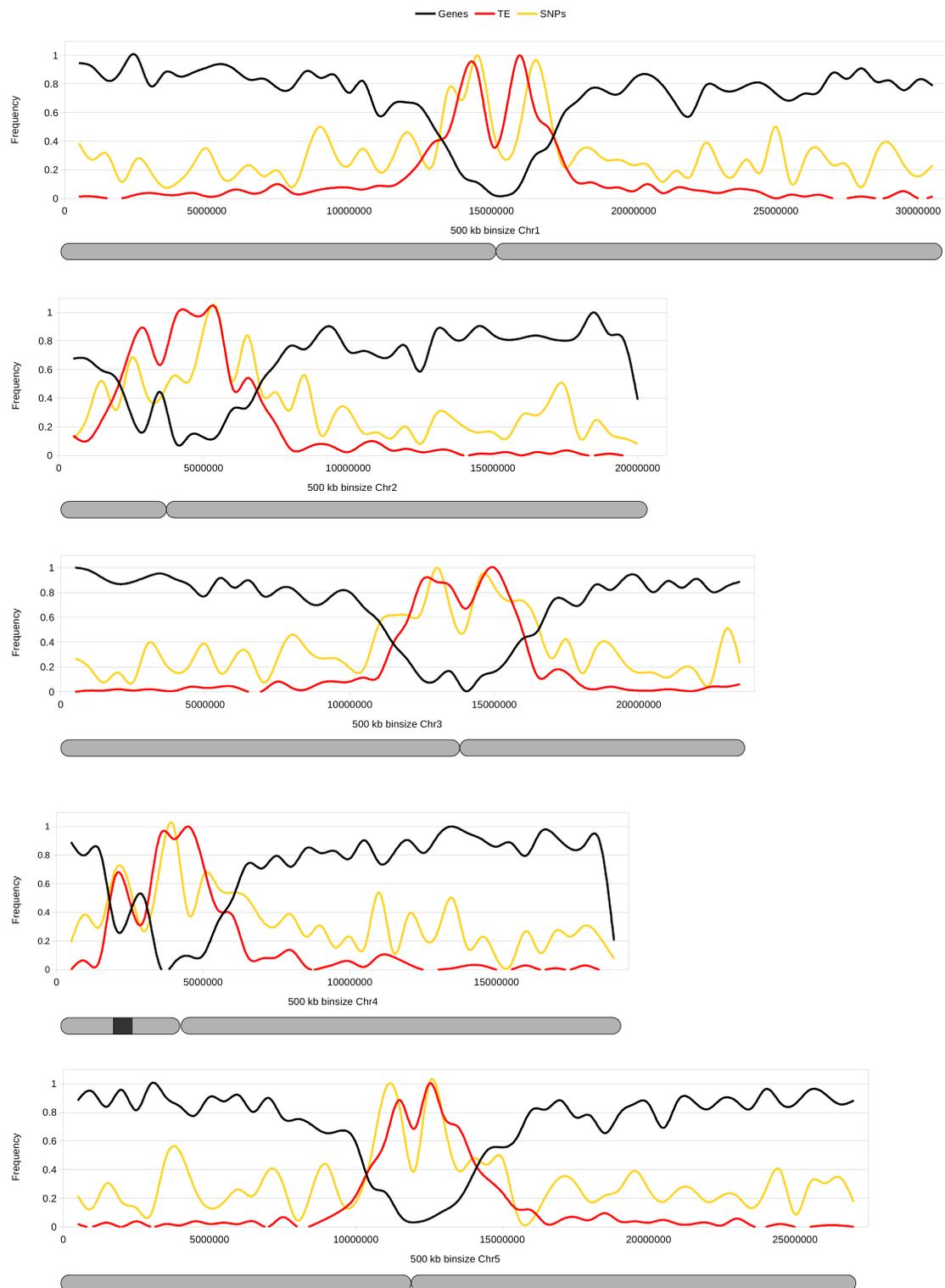
Fig. 3-18. Frequency of SNPs in relationship to genomic regions. Along the five chromosomes of *A. thaliana* the frequency of genes, TEs and all identified mutations summed in 500 kb bins is depicted. Grey bars represent schematicly drawn chromosomes with constrictions indicating position of centromeres. Dark grey box on chromosome 4 represents a heterochromatic island (McCombie et al., 2000).

3.2.4 The relationship between SNPs and DNA methylation

As the identified mutations were enriched in TEs and pericentromeric regions, regions that are frequently methylated, the relationship between SNPs and cytosine methylation was analysed. For every SNP position that contained a cytosine on the leading or lagging strand, the methylation was determined using publicly available DNA methylation data (Zhang et al., 2006; Cokus et al., 2008; Stroud et al. 2013). The DNA methylation was analysed in all three sequence contexts – CG, CHG and CHH (where H is any base but G) – as well as regardless of the the genomic context (C) (Fig. 3-19).
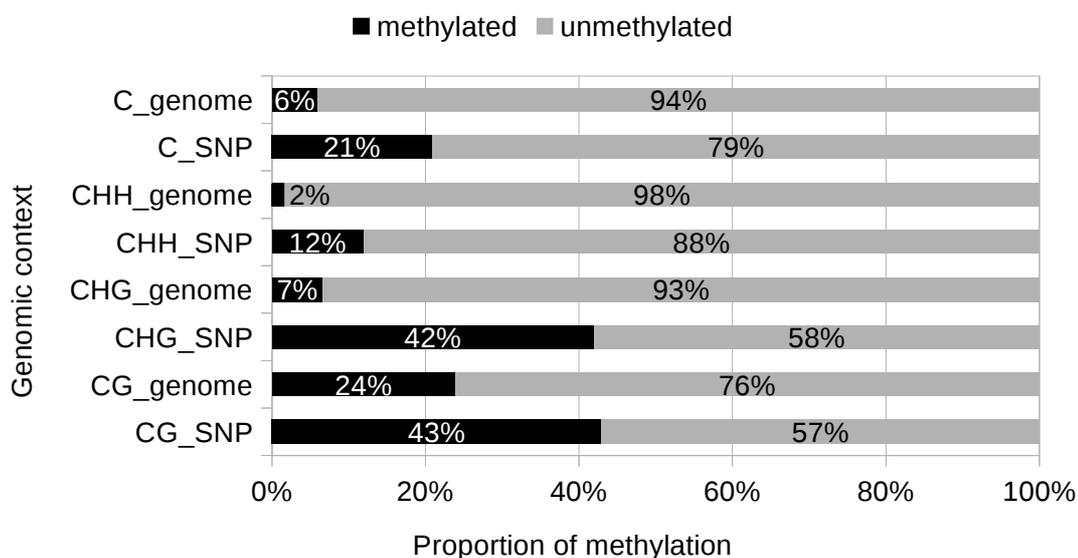
Fig. 3-19. The relationship between SNPs and DNA methylation. Methylation status of cytosine containing SNPs was analysed in all three sequence contexts – CG, CHG and CHH (where H is any base but G) – and independent of the the genomic context (C). Reference values for cytosine methylation on a genome-wide scale (XXX_genome) are taken from Cokus et al., 2008, and Lister et al., 2008.

In total, 1107 of the identified SNPs overlapped with cytosines on one of the DNA strands. Of these 1107 cytosines, 21% (n = 228) were found to be DNA methylated. In contrast, on average only around 6% of cytosines in the Arabidopsis genome are DNA methylated (Lister et al., 2008). Hence, UV-B induced SNPs were found to overlap with DNA methylated cytosines 3-fold more frequently than expected at random.

Subsequently, the DNA methylation was analysed in all three possible sequence contexts in order to identify associations with specific DNA methylation contexts. In the CHH context out of 799 cytosines 12% (n = 97) were DNA methylated. Compared to published data, where CHH methylation occurs in around 1.7% of cytosines in the

Arabidopsis genome, this was an increase of more than 7-fold (Cokus et al., 2008). In total 170 SNPs were found to overlap with cytosines in CHG context, of which 42% (n = 72) were DNA methylated. This represented an increase of around 6-fold compared to the 6.7% genome-wide DNA methylation rate in CHG context (Cokus et al., 2008). In the CG context 138 SNPs were identified to overlap with cytosines in that sequence background. Of these, 43% (n = 59) were methylated. Compared to published data, where CG methylation occurs in around 24% of cytosines in the Arabidopsis genome, this is an increase of around 1.8-fold (Cokus et al., 2008). In contrast to the methylation in CHG and CHH contexts, that almost exclusively occur in heterochromatic regions, CG methylation occurs both in silent DNA regions, as well as in actively transcribed genes in form of so called gene body methylation (Saze and Kakutani, 2011). Therefore SNPs in CG context were also analysed separately for TEs and genes. In total, 53 SNPs were found in CG context within TEs, of which 43, which equals 81%, were DNA methylated. In contrast, only 7 DNA methylated cytosines of 52 SNPs found in CG context within genes were methylated (=13%).

Taken together, DNA methylated cytosines in all possible contexts in the Arabidopsis genome were found to be mutated in response to UV-B irradiation more often than expected by chance.

3.2.5 Dependency of UV-B irradiation dosage and accumulation of mutations

To elucidate whether with increasing UV-B irradiation dosages plants accumulate more mutations, the occurrence of SNPs in relationship to the UV-B irradiation dosage was monitored for Col-0 wild-type and *uvr2/uvr3* mutant plants (Fig. 3-20).

In Col-0 mock plants (UV0) on average 3.5 SNPs per haploid genome were found. A similar number of SNPs was found in the three different UV-B treatments with around 4.2 SNPs per haploid genome for UV1, 2.1 SNPs for UV2 and 3.1 SNPs for UV3. Irrespective of the UV-B irradiation dosage applied, the number of identified SNPs per genome was not significantly different from mock grown plants. For *uvr2/uvr3* mutant plants, UV-B significantly increased the accumulation of mutations for all three UV-B irradiation dosages applied, compared to mock grown mutant plants. However, there were no significant differences in the accumulation of mutations between the three UV-B dosages. Hence, in the given dataset UV-B irradiation seems to have no impact on the accumulation of mutation for Col-0 wild-type plants. In contrast, in plants lacking UVR2 photolyase repair the mutation rate was drastically increased by UV-B irradiation. Apart
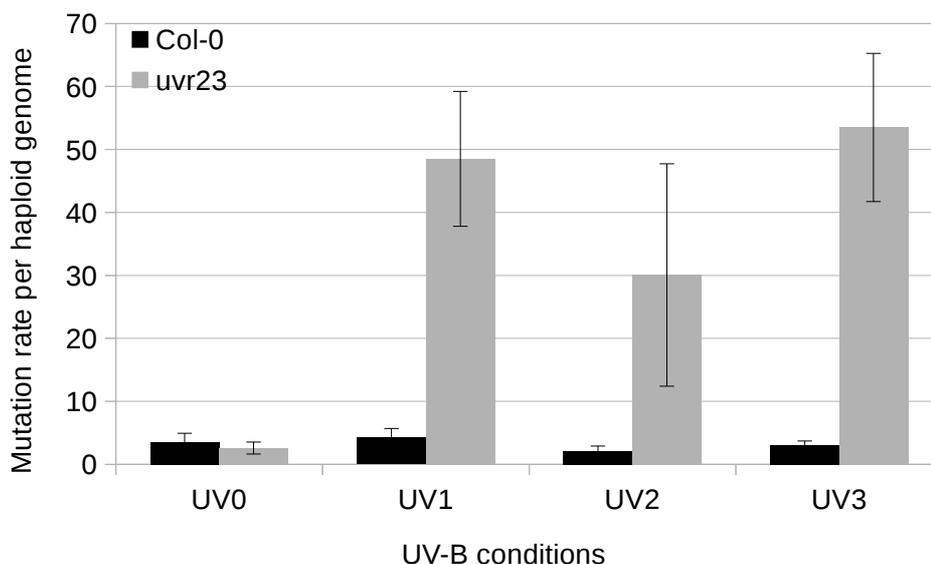
Fig. 3-20. Relationship between UV-B irradiation dosage and accumulation of mutations. Col-0 and *uvr2/uvr3* mutant plants were grown in three different UV-B conditions (UV1 to UV3) and mock conditions (UV0). The mutation rate per haploid genome was calculated. Each bar represents the mean of one genotype/condition, error bars = SD. N = 4-5 genomes.

from that, no correlation between the dosages applied and the accumulation of mutations for both genotypes could be observed.

3.2.6 Importance of repair pathways in the accumulation of UV-B induced mutations

The occurrence of SNPs in different mutant backgrounds was analysed to elucidate the involvement of different repair pathways in the prevention of UV-B induced mutations (Fig. 3-21).

Under control conditions without UV-B (UV0) all mutant genotypes had a similar basic mutation rate to Col-0 wild-type plants. This indicated that none of the mutants suffered from genomic instabilities that increased the mutation rate already under mock conditions. As described before, UV-B did not increase the mutation rate in Col-0 wild-type plants. This was true also for the flavonoid impaired mutant *tt4.* Furthermore, the *uvr8* mutant, lacking the UV-B photoreceptor, did not have an increased mutation rate caused by UV-B treatment. Also the photolyase mutant *uvr3* did not show more mutations upon UV-B irradiation. The only mutant genotypes that had strong and

significant increases in the number of mutations were the photolyase single mutant *uvr2* and the photolyase double mutant *uvr2/uvr3*. Of all genotypes analysed, the *uvr2* single mutant had the highest mutation rate observed with an average of 66 SNPs per haploid genome. As *uvr3* did not have an altered mutation rate, the elevated number of mutations in *uvr2/uvr3* double mutant probably were caused only by the absence of functional photolyase UVR2.
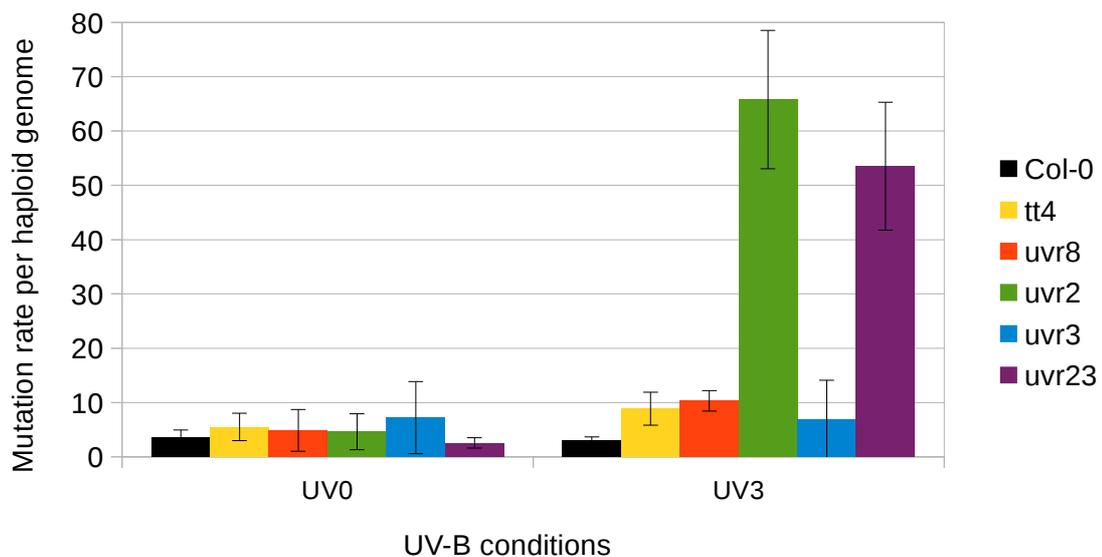
Fig. 3-21. Mutation rates in different mutant backgrounds. The mutation rates per genotype and condition were calculated, and normalized to the golden path length of the TAIR10 annotation as reference genome. Plants were grown in conditions with strong UV-B irradiation (UV3) and mock conditions (UV0). Each bar represents the mean of one genotype/condition, error bars = SD. N = 4-5 genomes.

3.2.7 Parental origin of mutations

As described before, the ratio between heterozygous mutations to homozygous mutations of all mutations identified differed from the expected ratio and was skewed towards heterozygous mutations. A possible explanation for this observation would be an uneven contribution of mutations to the offspring between maternal and paternal germ cells. Consequently, this would imply that mutations occur late in the development of a plant rather than uniformly over the whole life cycle of the plant. In order to explore whether mutations were equally contributed by both male and female germ cells or there was a difference in accumulation of mutations, reciprocal crosses between UV-B treated

(UV3) and untreated (UV0) *uvr2/uvr3* plants were performed and the mutations in their offspring monitored. To this end, anthers from four-week-old mock plants were removed to avoid self-pollination. Then the stigma was manually pollinated with pollen taken from plants grown for four weeks in UV-B conditions and *vice versa.*

When the the pollen acceptor plant (mother) was untreated and the pollen carrying plant (father) was UV-B irradiated 6.7 ± 1.2 mutations per genome were identified (n = 3 genomes). When the mother plant was UV-B irradiated and the father plant was untreated 12.7 ± 11.0 mutations per genome were found (n = 3 genomes). Although one value in the latter crossing direction deviated from the other two values, leading to a rather high standard deviation, with the given dataset no differences between the number of mutations in these plants could be observed. Hence, there was no indication that mutations were preferably originating from one parent, but rather an even contribution of both parents to the number of mutations in their offspring. To verify this result further experiments will be conducted.

3.2.8 Accumulation of mutations over generations

In order to explore to which extent mutations caused by UV-B irradiation are passed on to their offspring and how UV-B mutagenesis can aid in shaping natural diversity the accumulation of mutations was followed for three generations in Col-0 wild-type and *uvr2/uvr3* mutant plants grown in mock and UV-B conditions. Although UV-B irradiation had no effect on the mutation rate in Col-0 wild-type plants after one generation, mutations were identified in the following generations to exclude e.g. a additive or synergistic effect visible only after several generations.

For both genotypes no differences were observed in the number of identified mutations (Fig. 3-22; Fig. 3-23). Hence, mutations occurred in the same amount and speed over several generations, suggesting a constant mutation rate. When analysed separately for homozygous and heterozygous mutations, more homozygous mutations were found in the third generation compared to the second generation (Fig. 3-24). The ratio of heterozygous to homozygous mutations in UV-B treated plants was around 3.4 in generation 2 and 2.0 in generation 3. Although the differences were not significant, the increase in the number of homozygous mutations in generation 3 might indicate that some heterozygous mutations that were segregating in the previous generations, are manifesting themselves as homozygous mutations in the later generations.
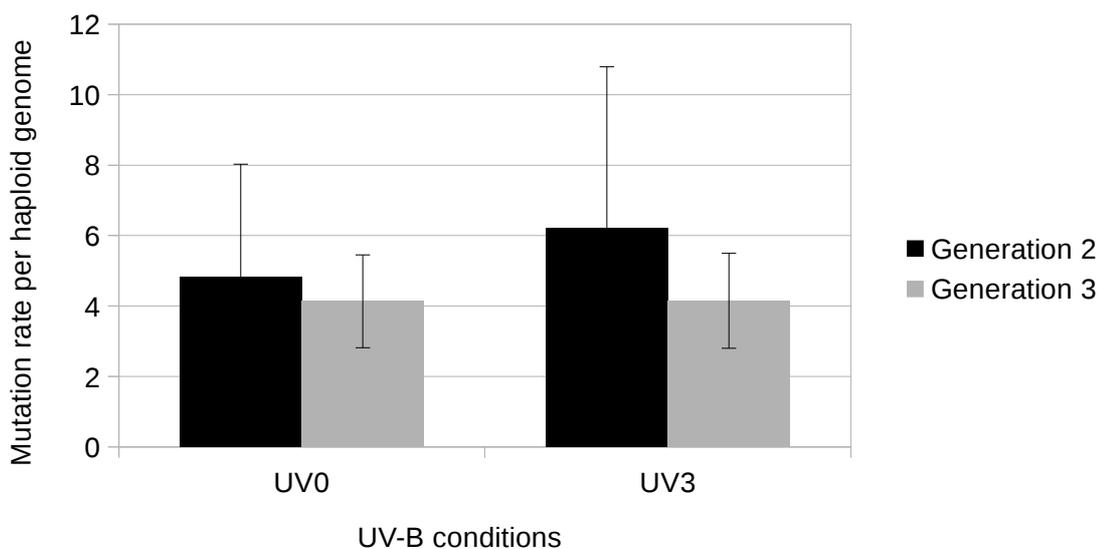
Fig. 3-22. Accumulation of mutations over generations in Col-0. Occurrence of SNPs in Col-0 wild-type plants grown for 3 generations in mock (UV0) or UV-B conditions (UV3). The mutation rates per genotype/condition/generation were calculated, filtered against mutations in generation 1 and normalized to the golden path length of the TAIR10 annotation as reference genome. Each bar represents the mean of one genotype/condition, error bars = SD. N = 4 genomes.
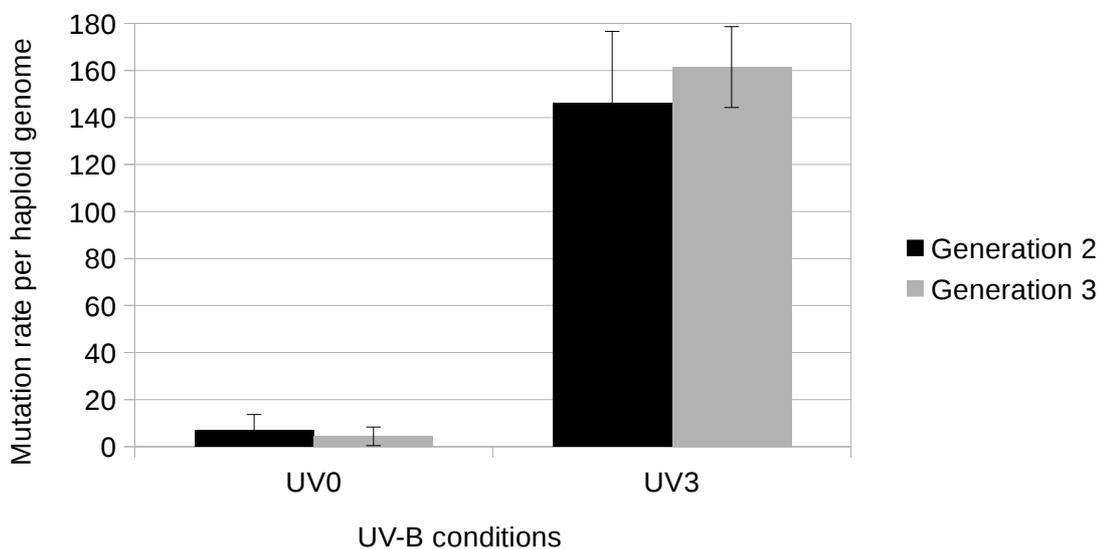
Fig. 3-23. Accumulation of mutations over generations in *uvr2/uvr3* mutant plants. Occurrence of SNPs in *uvr2/uvr3* mutant plants grown for 3 generations in mock (UV0) or UV-B conditions (UV3). The mutation rates per genotype/condition/generation were calculated, filtered against mutations in generation 1 and normalized to the golden path length of the TAIR10 annotation as reference genome. Each bar represents the mean of one genotype/condition, error bars = SD. N = 2-4 genomes.
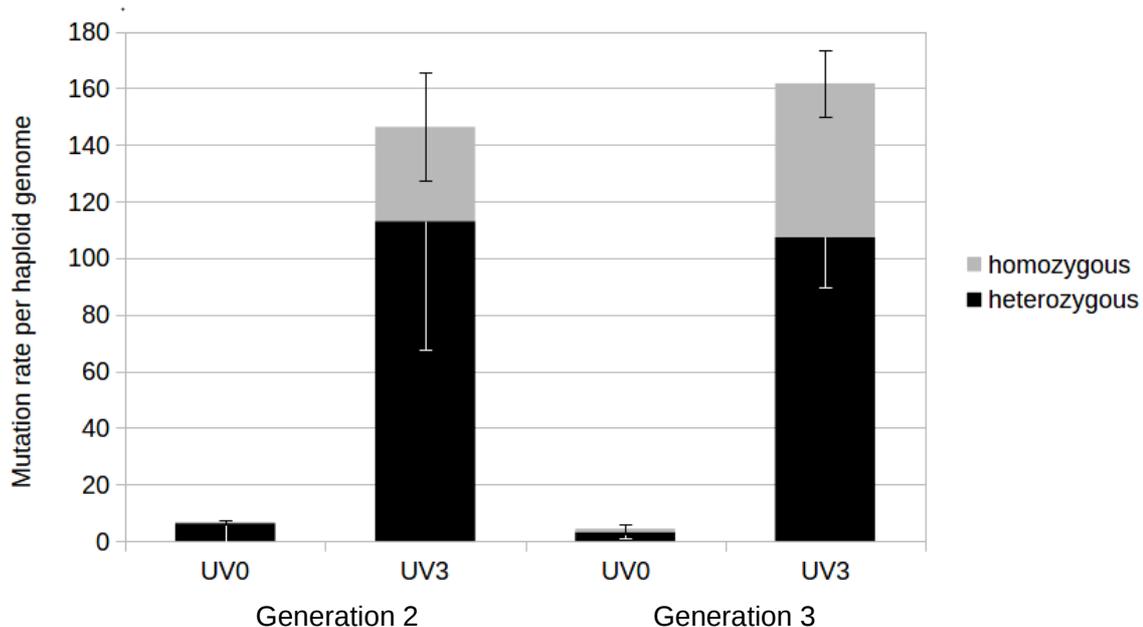
Fig. 3-24. Accumulation of mutations over generations in *uvr2/uvr3* mutant plants. Occurrence of SNPs in *uvr2/uvr3* mutant plants grown for 3 generations in mock (UV0) or UV-B conditions (UV3) separated for homozygous and heterozygous mutations. The mutation rates per genotype / condition / generation were calculated, filtered against mutations in generation 1 and normalized to the golden path length of the TAIR10 annotation as reference genome. Each bar represents the mean of one genotype/condition, error bars = SD. N = 2-4 genomes.

3.3 Photolyase reporter lines for monitoring UV-B induced DNA damage repair

As previous results pointed out the importance of the photolyase UVR2 for the repair of UV-B induced DNA damages, UVR2 reporter lines have been created in order to monitor potential tissue specific UVR2 accumulation in response to irradiation (Fig. 3-25). Two constructs have been created to monitor both transcription and translation of photolyase UVR2 by fusion of the *UVR2* promotor to the luciferase gene (*pUVR2::LUC*) and the whole genomic sequence of *UVR2* to the luciferase gene (*pUVR2::UVR2::LUC*).

Luciferase expressing plants indicated a basal and ubiquitous expression during vegetative growth (Fig. 3-26 A). According to own RT-qPCR and publicly available microarray data the expression of *UVR2* is induced upon UV-B irradiation 2- to 4-fold compared to mock conditions. This induction could not be observed on the protein level in the reporter lines (Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007).

In general the LUC-signal was stronger in lines where the luciferase gene was fused to the whole genomic sequence of *UVR2*, than to the promotor fusion.

Strong luciferase signals could be identified in flowering plants in flowers and siliques (Fig. 3-26 B). Again the signal was stronger for the whole genomic constructs. Microarray data suggested a strong expression of *UVR2* exclusively in pollen and a mild increase of 2- to 4-fold in floral tissue (Fig. 3-26; Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007). Although the resolution was limited, both reporter lines indicated a rather strong expression throughout whole floral tissue. To resolve the localization of UVR2 an approach with high resolution, e.g. confocal microscopy, will be conducted.

In order to identify potential modulators of *UVR2* expression and gain more insight into the regulation of *UVR2*, especially in floral tissue, a forward genetic EMS screen was initiated.

A

B
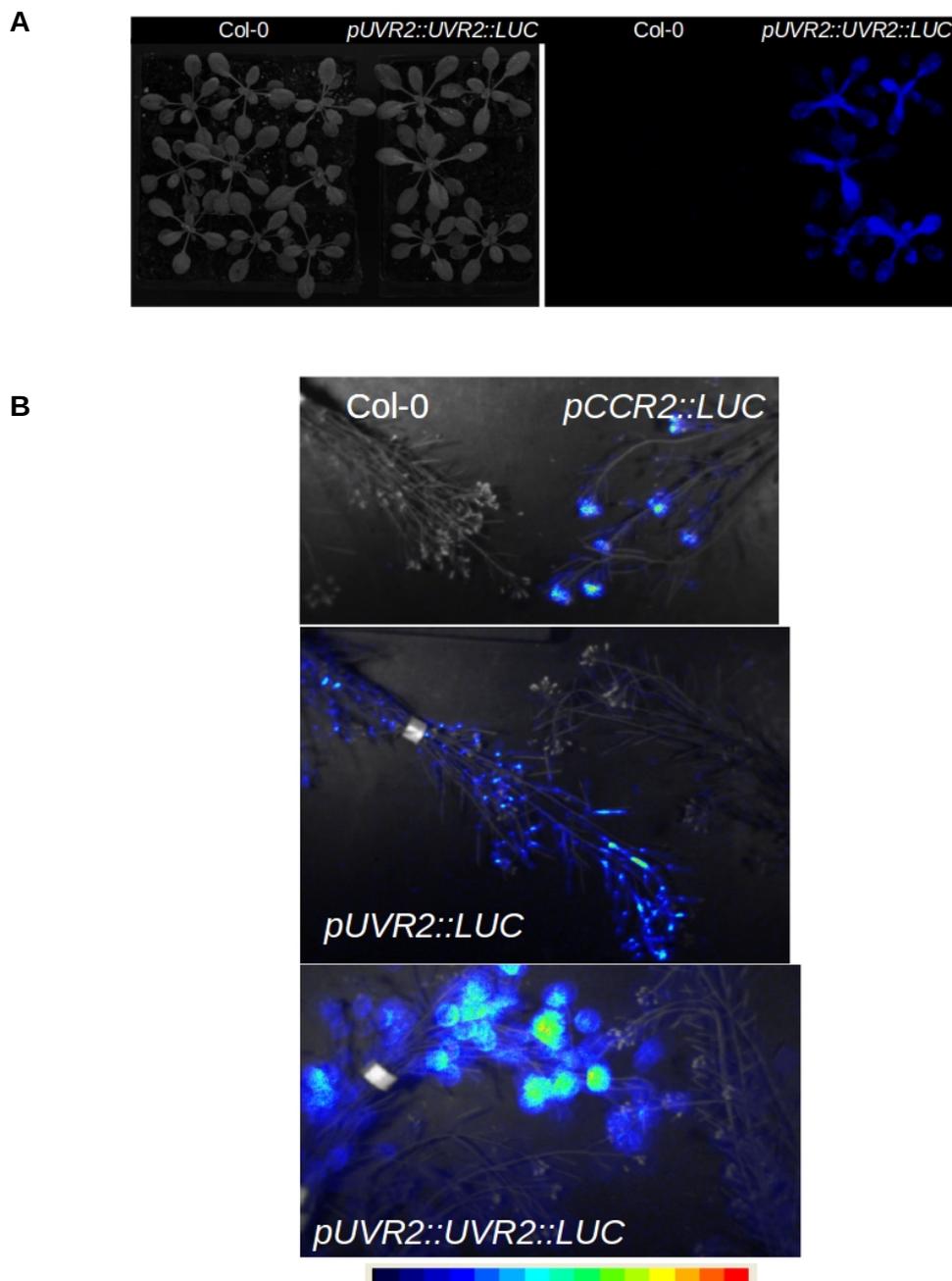
Fig. 3-25. Luciferase expressing photolyase reporter lines. Firefly's luciferase gene was fused to the promotor sequence of the photolyase *UVR2* gene and to the whole genomic sequence of *UVR2*. Luciferase signal could be detected during vegetative growth (A), and in flowering plants (B). Col-0 wild-type plants and a *CCR2* (At2g21660) promotor fusion have been used as controls (*pCCR2::LUC* line was a courtesy of Niels Müller).
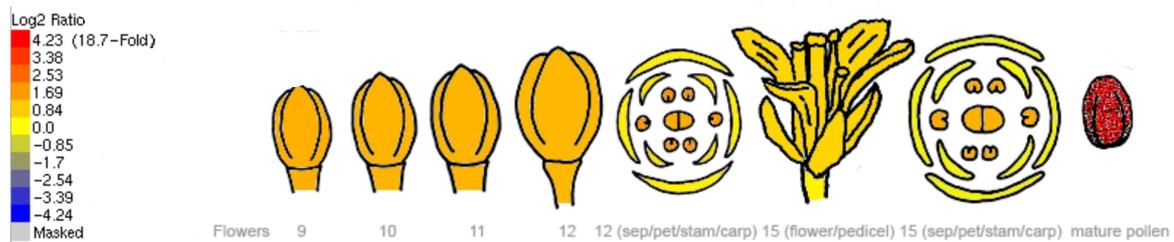
Fig. 3-26. Expression of *UVR2* photolyase during flower development. Image taken from efp-browser (Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007)

Discussion

UV-B radiation, as a natural part of solar radiation, is a ubiquitous environmental factor that plants are inevitably exposed to. As such, UV-B is an important factor influencing plant morphogenesis and development. It is perceived by the UV-B photoreceptor UVR8 which induces activation of specific target genes and downstream responses mainly in the photomorphogenic pathway (Favory et al., 2009; Jenkins, 2009; Rizzini et al., 2011; Tilbrook et al., 2013). Beside a few key components in the UV-B perception, signalling and UV-B induced damage response, the complex interplay of UV-B responses and their regulation is poorly understood. Many components may also be still not identified. In addition, little is known about UV-B responses in the context of natural variation.

Therefore a natural variation screen in response to UV stress has been performed. In total 345 *A. thaliana* accessions were tested and several sensitive (e.g. C24, Ba-1 and T1080) and resistant accessions (e.g. Mh-0, Lip-0 and Got-7) have been identified. This extends previous data where C24 was identified as a UV sensitive accessions. C24 was shown to develop curled leaves with necrotic spots after exposure to UV-B (Kalbrina and Strid, 2006). This was associated with increased expression of the *PR5* gene, a defence-related marker gene (Kalbrina and Strid, 2006). However, the genetic cause of these observations was not determined.

With leaf length, leaf area and fresh aerial biomass three growth related traits were monitored in response to UV. For none of them enhanced sensitivity or insensitivity was found to be correlated with the geographic origin of the accessions, suggesting that the response towards UV under the given experimental conditions was not connected to a geographical parameter. A similar observations was made for the analysis of the susceptibility of photosystem II in response to UV radiation where only a very weak association between the constitutive UV tolerance and the geographic origin of accessions was found (Jansen et al., 2010). Instead, most of the variance in constitutive UV-B protection of photosynthesis was observed at the level of local Arabidopsis populations originating in the same geographic and climatic area, which rather indicated an adaptation on a local scale (Jansen et al., 2010). In contrast, a correlation between UV tolerance and geographic parameters was seen for buckwheat and spruce species (Pukacki and Modrzynski, 1998; Yao et al. 2007). However, these studies cannot be compared directly because of different growth and UV-B conditions that were applied and different traits measured. Hence, a correlation between the variation of a specific trait in response to UV-B and a geographic component could not be detected in

Arabidopsis but might not be generally excluded for other plant species. In this regard, natural UV-B irradiation might represent only weak selection pressure and have only a mild effect on natural variation and adaptation. Detection of correlations between traits and geographical parameters might also be hindered by uneven sampling of Arabidopsis accessions, unknown environmental conditions of the sites the accessions were originated and fast dispersal of Arabidopsis (Trontin et al., 2011; Weigel, 2012).

A GWAS approach of the generated phenotype data led to the identification of one region of association that contained four candidate genes encoding for an E2F transcription factor, a putative ribonuclease inhibitor, a receptor-like protein and a protein with unknown function.

In general, E2F transcription factors (E2F TFs) are a family of transcription factors that have been found to be important for the mitotic and endoreduplication cell cycle, embryogenesis and DNA stress response (Ramirez-Parra et al., 2004; Vlieghe et al., 2005; Li et al., 2008). E2F TFs are functionally conserved among plants and animals, and based on their structural differences they can be subdivided into typical and atypical E2F (Lammens et al., 2009). In Arabidopsis six E2F TFs have been identified - E2Fa, E2Fb and E2Fc as typical E2F members, and E2Fd, E2Fe and E2Ff as atypical E2F. The identified candidate gene is the E2Fc TF and hence a typical E2F (Lammens et al., 2009). Typical E2F possess a conserved DNA binding domain and a dimerization domain, enabling the binding to a dimerization partner to form heterodimeric proteins (van de Heuvel and Dyson, 2008; Lammens et al., 2009). Typically the dimerization partner itself carries also a DNA binding domain that together with the DNA binding domain of the E2F facilitate a highly specific binding to promotors of target genes, including cell cycle regulation and DNA stress response (Vanderpoele et al., 2005; DeGregori and Johnson, 2006). The E2Fc TF has been reported to be involved in controlling the balance between cell proliferation and onset of the endocycle program (del Pozo et al., 2006). E2Fc itself is regulated through the cell cycle, expressed in both dividing and differentiated cells, and degraded in a ubiquitin-dependent manner in response to light (del Pozo et al., 2002). In contrast to E2Fa and E2Fb, that are proposed to work as transcriptional activators, E2Fc is repressing the expression of target genes. However, E2Fc is not only working antagonistically of the other two TFs, but functions also outside of cell cycle regulation (Jager et al., 2009). E2Fc may indirectly be involved in the responses following UV-B irradiation / stress. It has been shown that E2Fc is engaged in the control of the atypical E2Fe TF (also referred to as DEL1), which itself was identified as a transcriptional repressor of the *UVR2* photolyase

(Berckmans et al., 2011; Radziejwoski et al., 2011). Upon UV-B treatment, plants knocked out for *E2Fe* had improved DNA repair abilities compared to control plants, whereas those overexpressing it were more sensitive (Radziejwoski et al., 2011). However, no expression differences of *E2Fc* upon UV-B treatment among the sensitive and resistant accessions could be found. Also no relevant sequence differences between the protein coding sequences of sensitive and resistant accessions in public databases (1001 genomes) and single-read sequencing of few selected accessions carried out within this study could be identified. However, the promotor sequence of the *E2Fc* gene shows many polymorphisms, yet without any clear correlation between sensitivity and resistance.

Except their annotation as putative ribonuclease inhibitor and protein with unknown function no further information were available for the other two candidate genes within the region of association. In general, ribonuclease inhibitors are proposed to bind to certain ribonucleases and therefore regulate the degradation of RNAs (Dickson et al., 2005).

The fourth candidate gene was a receptor-like protein. Receptor-like proteins and kinases have often been shown to be involved in abiotic stress responses, disease resistance and plant development (Morris and Walker, 2003; Wang et al., 2008). For example the Arabidopsis protein TMM was shown to be involved in stomatal patterning and the *RPP27* gene conferring resistance to downy mildew (Nadeau and Sack, 2002; Tör et al., 2007). In total, 57 RLP genes have been identified in the Arabidopsis genome (Wang et al., 2008). In addition, many RLPs are functionally conserved among plants (Fritz-Laylin et al., 2005). Beside strong relative expression upon UV-B treatment with up to 60-fold induction in C24 and Amel-1 accessions one hour after UV-B treatment, several sequence aberrations were found. With few exceptions, these sequence aberrations formed two distinct haplotype groups. RLPs commonly exhibit great variation at the sequence level, as well as in the number of extracellular leucine-rich-repeats, which are thought to mediate the ligand perception (Kobe and Kajava, 2001; Fritz-Laylin et al., 2005; Kinoshita et al., 2005). A weak but significant linear correlation between the occurrence of a 12 bp deletion in the promoter region and sensitivity for all three traits was observed. Hence, the deletion of the 12 bp promotor region might represent, e.g. a lack of negative control, leading to higher expression of this gene under stress conditions. T-DNA knock-out mutants of the RLP candidate gene showed no altered sensitivity nor phenotypic differences. This might be due to functional redundancy with other RLP genes in Arabidopsis. This phenomenon typically masks the exploration of the

biological function of genes in families with a large number of members (Krakauer and Plotkin, 2002).

A GWAS approach offers the possibility for fast and high resolution mapping, however it also bears certain disadvantages, e.g. impaired detection of genetic cause due to genetic and allelic heterogeneity (Bergelson and Roux, 2010; Korte and Farlow, 2013). Therefore, a classical QTL mapping approach with a Col-0 x C24 RIL mapping population was performed to combine advantages of both approaches and complement the GWAS approach. QTL mapping revealed several genomic regions associated with the response to UV-B. One QTL co-localized with the peak of association identified by GWAS, hence supporting the GWAS results. Yet together with the other QTL identified, a more complex genetic interaction was indicated. Several NILs derived from the same mapping population were tested under the same conditions in order to identify the importance and relative contribution to the phenotype of each individual QTL. However, no differences between NILs with the same parental background were detected, so that sensitivity could be observed only in lines with the genomic background of the sensitive parent and *vice versa*.

As single knock-out mutants and individual NILs showed no altered phenotypes aiding in understanding of this complex genetic scenario, and the lack of just one genetic component seemed to not increase or decrease sensitivity in response to the UV-B treatment, sensitivity might have been conferred by the coordinated expression, or the lack of this coordinated expression, respectively, of different genes. In addition, UV-B has been shown to greatly modulate the transcription profile in Arabidopsis (Ulm et al., 2004; Killian et al., 2007; Zhao et al., 2014). In order to explore natural variation in this response and potentially link it with specific genetic components, a RNA sequencing experiment was conducted. Three sensitive (Ba-1, C24 and T1080) and three resistant accessions (Got-7, Lip-0, Mh-0) were selected and the transcriptome analysed 3 h after the UV-B treatment. In contrast to previous microarray studies this could identify novel transcripts, e.g. with low abundance or highly polymorphic genes that did not hybridize to probes on microarrays (Ulm et al., 2004; Killian et al., 2007; Zhao et al., 2014).

In total 315 genes have been identified that were differentially regulated in all three sensitive accessions and at the same time showed significantly different expression compared to all three resistant accessions. Genes were enriched for kinase activity, DNA binding, transferase activity and transmembrane transporter activity, indicating differentially expressed genes mainly in the signalling / regulatory part. Based on their expression, but also considering the previous QTL mapping and GWAS results, as well

as functional connections to known UV-B genes and pathways, e.g. expression dependency on *UVR8* and/or *HY5*, several candidate genes have been selected and T-DNA knock-out mutant tested for their response to UV-B irradiation (Brown et al., 2005; Oravecz et al., 2006; Kilian et al, 2007; Lee et al., 2007; Winter et al, 2007; Gonzalez-Besteiro et al., 2011; Obayashi et al., 2011). Mutants of five genes, encoding a methyl-transferase, a snare protein, a TIR-NB-LRR protein, a protein kinase and the phenylalanin-ammonium-lyase 1 (*PAL1*), showed remarkably different phenotypes. *PAL1* is known to be induced by UV-B irradiation, yet its connection in the context of natural variation has not been established (Kubasek et al., 1992). PAL1 catalyses the first step in the phenylpropanoid pathway and is involved in the biosynthesis of polyphenolic compounds such as flavonoids, phenylpropanoids, and lignin (Taiz and Zeigler, 2002). Therefore *pal1* mutants might have lower amounts of UV-protective polyphenolic compounds which leads to increased sensitivity under UV-B. The other candidate genes have not been associated with UV-B responses. More experiments need to be carried out to understand the function of these candidate genes in the response to UV-B.

Since the first demonstration of mutagenic effects of UV irradiation in 1914, in which sub-lethal doses of UV radiation modified the metabolism of an bacterial species so that it could grow on deficient media, but also by later research carried out with other species, including plants, it has been proposed that UV and in particular UV-B is one of or even the most important environmental factor causing mutations and hence be a driving force in evolution (Henri and Henri, 1914; McLennan, 1987; Kimura et al., 2004; Kimura and Sakaguchi, 2006). However, the majority of studies that proposed this hypothesis used unnaturally high doses of UV-B, or even UV-C radiation, which does not naturally occur on earth (Hockberger et al., 2002; Caldwell et al, 2007; McKenzie et al., 2007). Even if more ambient doses were applied, often, due to the experimental design, it is questionable if or how these results can be translated e.g. from *in-vitro* grown unicellular organisms to higher, more complex, organisms grown under natural conditions. In order to analyse the mutagenic potential of natural-like UV-B radiation in more realistic conditions for a higher organism, a mutation accumulation project was initiated. To address different aspects, three UV-B conditions with increasing UV-B dosages (UV1, UV2 and UV3) and one control condition devoid of UV-B radiation (UV0) were applied, which were mimicking realistic UV-B conditions and corresponding to UV-B conditions beginning of May in Berlin in Germany (UV1), Rome in Italy (UV2) and Athens in Greece (UV3). In addition to Col-0 wild-type plants, several mutant genotypes were analysed. These included the UV-B photoreceptor mutant *uvr8*, the flavonoid impaired

mutant *tt4*, the photolyase mutants *uvr2* and *uvr3*, and a photolyase double mutant *uvr2/uvr3*. To monitor mutations on a genome-wide scale, mutations were identified by a whole genome sequencing approach. In total, over 120 genotypes were sequenced and mutations identified by a short read mapping approach (Ossowski et al., 2008). As mutations were analysed in the progeny of treated plants, only mutations that occurred in the germ line were inspected.

In total, 1980 mutations were identified together in all genomes, conditions and generations. The mutation rate of Col-0 under mock conditions was similar to the mutation rates of all other mutants under mock conditions. This indicated that none of the mutants suffered from genomic instabilities that increased the mutation rate already under mock conditions. The average mutation rate for mock grown plants was estimated as around 2.5 x 10^{-8} per haploid genome and generation. This is around 3 times higher than the previous estimation of the spontaneous mutation rate in *A. thaliana* which revealed a basic mutation rate of around 7 x 10^{-9} per site and generation (Ossowski et al., 2010). However, in the cited study plants were grown under unknown conditions, presumably under greenhouse conditions with rather unclear and unstable UV-A and PAR conditions that might vary drastically, e.g. depending on the season and cloud coverage, whereas plants grown within this study were grown under stable UV-A and PAR conditions, which might be higher than in greenhouses and hence might explain the higher basic mutation rate. Compared to spontaneous mutation rates in other species, the estimated mutation rate is in the range of mutation rates described for humans (1.2 x 10^{-8}) and the roundworm *Caenorhabditis elegans* (2.1 x 10^{-8}) and higher than mutation rates in the fruitfly *Drosophila melanogaster* (2.8 x 10^{-9}) or in the budding yeast *Saccharomyces cerevisiae* (1.7 x 10^{-10}) (Denver et al., 2009; Kong et al., 2012; Keightley et al., 2014; Zhu et al., 2014). The mutation rate for the only other plant species so far analysed, the unicellular green algae *Chlamydomonas reinhardtii*, identified a lower mutation rate with 3.2 x 10^{-10} compared to the one presented here (Ness et al., 2012). However, in addition to a species specific component influencing the mutation rates, the observed variations in mutation rates might be also due to strongly varying growth conditions, sample sizes and estimations.

As described, under control conditions without UV-B (UV0) all mutant genotypes had a similar basic mutation rate to Col-0 wild-type. Furthermore, UV-B did not increase the mutation rate in Col-0 wild-type plants, the flavonoid impaired mutant *tt4,* the *uvr8* mutant, lacking the UV-B photoreceptor, and the photolyase mutant *uvr3*. In contrast, the photolyase single mutant *uvr2* and the photolyase double mutant *uvr2/uvr3* had strong

and significant increase in the number of mutations. Of all genotypes analysed, the *uvr2* single mutant had the highest mutation rate observed with an average of 66 mutations per haploid genome. As *uvr3* did not have an altered mutation rate, the elevated number of mutations in *uvr2/uvr3* double mutant was most likely caused only by the absence of functional photolyase UVR2. As no increased mutation rate was detected in the photoreceptor-deficient mutant *uvr8,* the repair of mutations induced upon UV-B irradiation seemed to be triggered independently of the photoreceptor, as also shown for photomorphogenic signalling that was accomplished via different MAP kinases independently of UVR8 (Gonzalez-Besteiro et al., 2011). Although mutants impaired in flavonoid production, like *tt4*, were shown to be more sensitive to UV-B irradiation than wild-type plants, it is not clear which precise function flavonoids fulfil in preventing UV-B stress-related responses (Li et al., 1993; Agati and Tattini, 2010; Fini et al., 2011). As *tt4* plants showed no increase in mutations upon UV-B irradiation, flavonoids are presumably more important as scavengers of UV-B induced ROS, for example to protect the photosynthesis apparatus, than directly involved in prevention of DNA damage (Hideg and Vass, 1996; Barta et al., 2004; Gerhardt et al., 2005; Agati and Tattini, 2010). Furthermore, no increased mutation rate was observed in the *uvr3* photolyase mutant. The majority of UV-B induced mutations are CPDs, which are repaired by the photolyase UVR2 (Britt, 2004). Only up to 25% of pyrimidine dimer are 6-4PPs which are repaired specifically by UVR3 (Britt et al., 1996; Lo et al., 2005). Hence, already due to the lower occurrence of 6-4PPs a lower mutation rate in *uvr3* would have been expected compared to *uvr2*. However, as no increased amount of mutations in *uvr3* was found at all, this indicated another repair pathway working redundantly to the UVR3 repair. This might be NER or MMR. For example, mutants impaired in the NER pathway, like *uvh1*, were shown to be hypersensitive to UV-B radiation (Liu et al., 2000). Taken together, the UVR2 photolyase seemed to be essential for the repair of UV-B induced mutations and worked independently of UVR8. Flavonoids, the UV-B photoreceptor and the photolyase UVR3 were not required for the prevention of UV-B induced mutations.

To elucidate whether with increasing UV-B irradiation plants accumulate more mutations, the occurrence of mutations in relationship to the UV-B irradiation dosage that was applied was monitored for Col-0 wild-type and *uvr2/uvr3* mutant plants. Only in the *uvr2/uvr3* mutant UV-B significantly increased the number of mutations. However, no correlation between the dosage of UV-B and the number of mutations could be observed. In Col-0 wild-type UV-B did not elevate mutation rates. Either UV-B was indeed not capable of inducing mutations in wild-type plants, as the applied conditions

simulated natural-like conditions to which plants may have adapted to over a long period of time and hence could avoid the accumulation of mutations, or a single threat to the genomic stability like UV-B was alone not capable to increase mutation rates. In true natural conditions, plants may face multiple stresses simultaneously, e.g. increased UV-B levels, lack of a certain nutrient, drought stress, and biotic stress by pathogens, and possibly only in combination these could cause mutations in germ cells that are inherited to the progeny. For flavonols a crosstalk between UV-B stress and biotic stress signalling was shown to prevent flavonol accumulation in favour of pathogen defence compound production (Schenke et al., 2011). Hence, under limiting conditions plants may have to set priorities to deal with the most urgent threat and at the same time become more susceptible to other threats. In heat stressed plants epigenetic control of certain repetitive elements was hindered, allowing activation and transposition of these elements, which could lead to genomic instabilities (Pecinka et al., 2010; Ito et al., 2011). The majority of the identified mutations was heterozygous, with a ratio of around 3:1 of heterozygous to homozygous mutations. Since mutations were identified in the progeny of irradiated plants, 2/3 heterozygous mutations and 1/3 homozygous mutations would be expected, suggesting a ratio of 2:1, assuming a normal mendelian inheritance of mutations. One possible explanation for this skewed ratio would be the occurrence of mutations at later developmental stages, supposedly after the split of male and female lineages, which could lead to an uneven contribution of mutations to the offspring between maternal and paternal germ cells. In order to explore whether mutations were equally contributed by both male and female germ cells or there was a difference in accumulation of mutations, reciprocal crosses between UV-B treated (UV3) and untreated (UV0) *uvr2/uvr3* plants were performed and the mutations in their offspring monitored. Preliminary analysis revealed similar mutation rates between both crosses, which suggested a rather even contribution of both parents to the number of mutations in their offspring. Further experiments need to be conducted to elucidate, e.g. if a particular developmental stage is more prone to accumulate UV-B induced mutations.

Mutations were also distinguished as transitions or transversions. Most mutations were transitions, and this mutation type was specifically enhanced by UV-B irradiation, as the amount of transitions increased from 58% in mock conditions to 92% in UV-B conditions. Especially GC to AT transitions increased in UV-B treated plants. The amount of transitions under mock conditions was in agreement with previous estimates for Arabidopsis under UV-B free conditions (Ossowski et al., 2010). The bias towards transitions could also be observed for (spontaneous) mutations in different bacterial

species, the budding yeast *S. cerevisiae*, the fruitfly *D. melanogaster* and humans, ranging from around 30% to 70% (Keightly et al., 2009; Hershberg and Petrov, 2010; Lynch, 2010; Kong et al., 2012; Zhu et al., 2014). However, this might not be a universal phenomenon but vary between species, as seen in the mutation spectrum of pseudogene sequences in a grasshopper species (Keller et al., 2007). In general, transitions are more likely to manifest in a genome because this type of mutation changes the chemical structure less dramatically than transversions. This is also because the wobble position of the DNA, which to a large extent is responsible for the degeneracy of the genetic code, is more tolerant to a transition than to a transversion, such that a transition mutation is more likely to lead to the same amino acid being encoded for, whereas a transversion is more likely going to change the encoded amino acid (Alberts et al., 2002).

If mutations occurred randomly, an even distribution over the whole genome should be expected. In order to analyse this, the distribution of mutations was compared to the genome composition of *A. thaliana*. The majority of mutations (40%) was found in TEs and only 34% of mutations were found in genes. Compared to the genome composition with 51% genic regions and 21% TEs, the occurrence of UV-B induced mutations seemed to be skewed from genic regions towards TEs and suggested a non-random distribution. Also along the individual chromosomes mutations showed a non-uniform distribution. Mutation frequencies were higher in pericentromeric regions than on chromosomal arms and were correlated with the frequencies of TEs, whereas the occurrence of mutations and genes was negatively correlated. Remarkably, this held true also at a relatively fine genomic scale as a similar pattern could also be observed for the heterochromatic knob on chromosome four. Taken together mutations were predominantly found in TEs and in pericentromeric regions than expected at random. This might be caused by a certain genomic feature that favours these regions to be mutated or by impaired repair of induced mutations in these regions. Both phenomena might also contribute together to the increased number of mutations. It could be shown that the higher condensation in heterochromatic and (peri-)centromeric regions, represented by a higher nucleosome density, impaired the repair efficiency of UV-B induced mutations (Suter and Thoma, 2002; Thoma, 2005). In yeast cells it was demonstrated that nucleosomal DNA was repaired by photolyases within seconds under conditions where photolyases were not limited, whereas repair of UV-caused damages in heterochromatin was more slowly repaired within several minutes and centromeres could not be repaired at all (Bucceri et al., 2006). In addition, chromatin dynamics and

chromatin remodelling proteins have been shown to be involved in the DNA damage response upon UV-B (Yu et al., 2005; Palomera-Sanchez and Zurita, 2010; Campi et al., 2012).

On the other hand, a common feature between TEs, pericentromeric and hetero-chromatic regions in Arabidopsis is a high degree of DNA methylation. Therefore, for every position that was found to be mutated upon UV-B irradiation within this study and contained a cytosine on the leading or lagging strand, the DNA methylation was determined using publicly available DNA methylation data (Zhang et al., 2006; Cokus et al., 2008; Stroud et al. 2013). The DNA methylation was analysed in all three sequence contexts – CG, CHG and CHH (where H is any base but G) – and regardless of the sequence context. In all sequence contexts mutations overlapped with DNA methylated cytosines more frequently than expected at random with the biggest differences in CHH context. Approximately 1.7% of cytosines in CHH context in the Arabidopsis genome are DNA methylated, whereas 12% of the mutated cytosines were DNA methylated in the CHH context, which was an increase of more than 7-fold. In contrast to the methylation in CHG and CHH contexts, that almost exclusively occurs in transcriptionally inactive regions, CG methylation occurs also in actively transcribed genes in form of so called gene body methylation (Saze and Kakutani, 2011). Therefore SNPs in CG context were also analysed separately for TEs and genes. Mutated cytosines in CG context in TEs were in 81% of cases DNA methylated, whereas only 13% of cytosines in CG context in genes were DNA methylated. On average genes have around 14% CG methylation, hence the estimated ratio reflects the expected outcome suggested by whole genome methylation data (Cokus et al., 2008). In contrast, mutated cytosines in TEs were found more frequently to be DNA methylated compared to the average CG methylation in TEs of around 66% on a genome-wide scale (Cokus et al., 2008). Taken together, DNA methylation influenced the probability of a cytosine to be mutated. At the same time, these mutations seemed to be efficiently repaired in genes, as a lower number of mutations was found in genes than expected at random. In contrast to genes, these mutated cytosines were enriched in TEs, especially in heterochromatic and (peri-)centromeric regions, as more mutations were observed in these genomic regions. This might indicate a putative mechanism employed by plants, in which the energy provided by UV-B radiation is used to accelerate the occurrence of mutations at specific regions, like TEs, in order to enhance inactivation of them. A similar observation was made in cell cultures, where constitutive methylation of cytosines increased the frequency of UV-B induced CPD formation by 1.7-fold, suggesting that methylation *per*

se was influencing the probability of damage formation, due to different absorption properties of methylated cytosines compared to non-methylated cytosines (Rochette et al., 2009). Methylated cytosines showed higher absorption in the UV-B spectrum (Rochette et al., 2009). Also in the human genome methylated cytosines have been shown to have elevated mutation rates (Fryxell and Moon, 2005; Xia et al., 2012; Hernando-Herraez et al., 2013). However, although the majority of mutations identified within this study contained cytosine residues of which a large percentage was DNA methylated, around 1/3 of the identified mutations was found in non-cytosine context, suggesting that DNA methylation is not exclusively responsible for enhanced mutation rates.

Furthermore, it was analysed if mutation rates were stable over generations and how UV-B mutagenesis could aid in shaping natural diversity by accumulation of mutations. Therefore, mutations were followed for three generations in Col-0 wild-type and *uvr2/3* mutant plants grown in mock and UV-B conditions. For both genotypes mutations occurred in the same amount and speed over the three followed generations, suggesting a constant mutation rate.

As the presented results pointed out the importance of the photolyase UVR2 for the repair of UV-B induced DNA damages, two UVR2 reporter lines have been created in order to monitor UVR2 accumulation in response to irradiation throughout development both at the transcriptional and translational level.

A basal and ubiquitous expression of the UVR2 reporter during vegetative growth could be observed. However, in contrast to expression data that indicated an induction of UVR2 upon UV-B irradiation 2- to 4-fold compared to mock conditions, no differences could be observed on the protein level in these reporter lines (Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007). This could be due to lacking sensitivity of the detection system of only minor differences during vegetative growth. However, strong signals could be identified in flowers and siliques. Microarray data suggested a strong expression of *UVR2* exclusively in pollen and a mild 2- to 4-fold increase in floral tissue (Fig. 3-26; Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007). Both reporter lines indicated a rather strong expression throughout whole floral tissue. This would be also more in agreement with the preliminary estimations of mutation rates identified in reciprocal crosses of *uvr2/uvr3*, that suggested an even contribution of maternal and paternal germ cells to the mutation spectrum in the offspring, which would stand in contrast to a very strong expression of UVR2 only in pollen and not in the other flower parts. To resolve the localization of UVR2 more clearly an approach with higher

resolution, e.g. confocal microscopy, will be conducted. With the translational UVR2 reporter line a forward genetic EMS screen was initiated in order to identify potential modulators of *UVR2* expression and gain more insight into the regulation of *UVR2*, especially in floral tissue.

Concluding remarks and outlook

Through the combined use of GWAS, QTL mapping and transcriptome sequencing following a natural variation screen in response to UV-B irradiation, several novel candidate genes involved in response to UV-B irradiation were identified. At the same time a complex genetic basis for sensitivity to UV-B irradiation was indicated. Hence, validation of candidate genes might be hindered, so as to sensitivity or resistance conferred by the lack or presence of one candidate gene might be only visible in a certain genetic background, presumably also due to redundant gene functions. Certainly, complementation tests of candidate genes in the appropriate genetic background would provide utmost proof of the relevance of these genes. To this end, combinations of different T-DNA insertion mutants or NILs might be necessary. Further analysis could also include expression studies of the identified genes on a finer scale, such as in different tissues or developmental stages as well as at more time points to gain more insight into the precise function of these candidate genes.

Furthermore, the mutagenic potential of natural-like UV-B radiation was explored. To this end, *A. thaliana* wild-type plants and selected mutants in UV-B signalling, DNA damage repair and in the flavonoid production pathway were analysed for their accumulation of mutations on a genome-wide scale. Mutations were shown to occur non-randomly and were enriched in TEs in (peri-)centromeric regions. In addition, methylated cytosines were favourably mutated. However, many questions remain open. UV-B induced many mutations in mutants deficient in the UVR2 photolyase, which was shown to be essential for the repair of UV-B induced mutations. However, no elevated mutation rates were found in *uvr3* photolyase mutant. It would be interesting to analyse mutants impaired in other repair pathways to see the importance or contribution of these pathways in the repair or prevention of UV-B induced mutations. Especially mutants in the NER and MMR pathways would be insightful. As no UV-B induced mutations were found in *uvr8* mutants it would be also of interest how UVR2 repair is triggered in these conditions. This could be achieved by UVR8-independent pathways via different MAP kinases or an other, yet unknown pathway. To this end, mutants knocked-out for several genes might be used. As UV-B alone was not capable of inducing mutations in wild-type plants it

would be interesting to see whether combined stresses could cause elevated mutation rates also in wild-type plants. Furthermore, the sensitivity of different developmental stages to induction of UV-B caused mutations could be tested by irradiation of plants for short periods at defined developmental stages. As DNA methylated cytosines were shown to be more prone to be mutated upon UV-B irradiation, it would be interesting to analyse the mutational spectrum in DNA methylation deficient mutants.

Table A.1: *Arabidopsis thaliana* accessions

Accession	Latitude	Longitude	Country
11ME1.32	42.093	-86.359	USA
11PNA4.101	42.0945	-86.3253	USA
328PNA054	42.0945	-86.3253	USA
627ME-4Y1	42.093	-86.359	USA
Aa-0	50.9167	9.57073	GER
Ag-0	45	1.3	FRA
ALL1-2	45.2667	1.48333	FRA
ALL1-3	45.2667	1.48333	FRA
Alst-1	54.8	-2.4333	UK
Amel-1	53.448	5.73	NED
An-1	51.2167	4.4	BEL
An-2	51.2167	4.4	BEL
Ang-0	50.3	5.3	BEL
Ann-1	45.9	6.13028	FRA
App1-16	56.3333	15.9667	SWE
Ã–r-1	56.45	16.11	SWE
Arby-1	59.4308	16.7999	SWE
BÃ¥1-2	56.4	12.9	SWE
Ba-1	56.5459	-4.79821	UK
Baa-1	51.3333	6.1	NED
Bay-0	49	11	GER
Be-1	49.6803	8.6161	GER
Belmonte-4-94	42.1167	12.4833	ITA
Benk-1	52	5.675	NED
Bg-2	47.6479	-122.305	USA
Bla-1	41.6833	2.8	ESP
Blh-1	48	19	CZE
Blh-2	48	19	CZE
Boot-1	54.4	-3.2667	UK
Bor-1	49.4013	16.2326	CZE
Bor-4	49.4013	16.2326	CZE
Br-0	49.2	16.6166	CZE
Bro1-6	56.3	16	SWE
Bs-2	47.5	7.5	SUI
Bsch-0	40.0167	8.6667	GER
Bu-0	50.5	9.5	GER
Bu-8	50.5	9.5	GER
BUI	48.3667	0.933333	FRA
Bur-0	54.1	-6.2	IRL
C24	41.25	-8.45	POR
Ca-0	50.2981	8.26607	GER
CAM-16	48.2667	-4.58333	FRA
CAM-61	48.2667	-4.58333	FRA
Can-0	29.2144	-13.4811	ESP
Cen-0	49	0.5	FRA
Cha-0	46.0333	7.1167	SUI
Chat-1	48.0717	1.33867	FRA
CIBC-17	51.4083	-0.6383	UK
CIBC-2	51.4083	-0.6383	UK
CIBC-4	51.4083	-0.6383	UK
CIBC-5	51.4083	-0.6383	UK
Cit-0	43.3779	2.54038	FRA
CLE-6	48.9167	-0.48333	FRA

Cnt-1	51.3	1.1	UK
Co-2	40.12	-8.25	POR
Co-4	40.12	-8.25	POR
Col-0	38.3	-92.3	USA
Com-1	49.416	2.823	FRA
CSHL-5	40.8585	-73.4675	USA
Ct-1	37.3	15	ITA
CUR-3	45	1.75	FRA
Cvi-0	15.1111	-23.6167	CPV
Da(1)-12	NA	NA	CZE
Da-0	49.8724	8.65081	GER
Db-0	50.3055	8.324	GER
Di-1	47	5	FRA
Do-0	50.7224	8.2372	GER
Dra-2	49.4167	16.2667	CZE
DraIV 1-14	49.4112	16.2815	CZE
DraIV 1-5	49.4112	16.2815	CZE
DraIV 1-7	49.4112	16.2815	CZE
DraIV 6-16	49.4112	16.2815	CZE
DraIV 6-35	49.4112	16.2815	CZE
Duk	49.1	16.2	CZE
Ede-1	52.0333	5.66667	NED
Eden-2	62.877	18.177	SWE
Edi-0	56	-3	UK
Ep-0	50.1721	8.38912	GER
Es-0	60.1997	24.5682	FIN
Est-0	58.3	25.3	RUS
Est-1	58.3	25.3	RUS
Fab-4	63.0165	18.3174	SWE
Fei-0	40.5	-8.32	POR
Fi-1	50.5	8.0167	GER
FjÃ¤1-2	56.06	14.29	SWE
FjÃ¤1-5	56.06	14.29	SWE
Fr-4	50.1102	8.6822	GER
Ga-0	50.3	8	GER
Ga-2	50.3	8	GER
Gd-1	53.5	10.5	GER
Ge-0	46.5	6.08	SUI
Ge-1	46.5	6.08	SUI
Gel-1	51.0167	5.86667	NED
Gie-0	50.584	8.67825	GER
Go-0	51.5338	9.9355	GER
Got-7	51.5338	9.9355	GER
Gr-1	47	15.5	AUT
Gr-5	47	15.5	AUT
Gu-1	50.3	8	GER
Gul1-2	56.3	16	SWE
Gy-0	49	2	FRA
Ha-0	52.3721	9.73569	GER
Hau-0	55.675	12.5686	DEN
Hey-1	51.25	5.9	NED
Hh-0	54.4175	9.88682	GER
Hi-0	52	5	NED
Hn-0	51.3472	8.28844	GER
Hod	48.8	17.1	CZE
Hov4-1	56.1	13.74	SWE
Hovdala-2	56.1	13.74	SWE
HR-5	51.4083	-0.6383	UK

Hs-0	52.24	9.44	GER
HSm	49.33	15.76	CZE
In-0	47.5	11.5	AUT
Je-0	50.927	11.587	GER
JEA	43.6833	7.33333	FRA
Jl-3	49.2	16.6166	CZE
Jm-1	49	15	CZE
Ka-0	47	14	AUT
Kas-1	35	77	IND
KBS-Mac-8	42.405	-85.398	USA
Kelsterbach-2	50.0667	8.5333	GER
Kelsterbach-4	50.0667	8.5333	GER
Kin-0	44.46	-85.37	USA
Kl-5	50.95	6.9666	GER
Kn-0	54.8969	23.8924	LTU
KNO-11	41.2816	-86.621	USA
Kno-18	41.2816	-86.621	USA
Köln	51	7	GER
Kr-0	51.3317	6.55934	GER
Kro-0	50.0742	8.96617	GER
Krot-2	49.631	11.5722	GER
Kulturen-1	55.705	13.196	SWE
LAC-3	47.7	6.81667	FRA
LAC-5	47.7	6.81667	FRA
Lc-0	57	-4	UK
LDV-14	48.5167	-4.06667	FRA
LDV-25	48.5167	-4.06667	FRA
LDV-34	48.5167	-4.06667	FRA
LDV-58	48.5167	-4.06667	FRA
Ler-1	47.984	10.8719	GER
Li-3	50.3833	8.0666	GER
Li-5:2	50.3833	8.0666	GER
Li-6	50.3833	8.0666	GER
Li-7	50.3833	8.0666	GER
Liarum	55.95	13.85	SWE
Lillo-1	56.1512	15.7844	SWE
LI-OF-095	40.7777	-72.9069	USA
Lip-0	50	19.3	POL
Lis-1	56	14.7	SWE
Lis-2	56	14.7	SWE
Lisse	52.25	4.5667	NED
LL-0	41.59	2.49	ESP
Lm-2	48	0.5	FRA
Lom1-1	56.09	13.9	SWE
Lov-5	62.801	18.079	SWE
Lp2-2	49.38	16.81	CZE
Lp2-6	49.38	16.81	CZE
Lz-0	46	3.3	FRA
Map-42	42.166	-86.412	USA
Mc-0	54.6167	-2.3	UK
Mh-0	50.95	7.5	POL
MIB-15	47.3833	5.31667	FRA
MIB-22	47.3833	5.31667	FRA
MIB-28	47.3833	5.31667	FRA
MIB-84	47.3833	5.31667	FRA
MNF-Che-2	43.5251	-86.1843	USA
MNF-Jac-32	43.5187	-86.1739	USA
MNF-Pot-48	43.595	-86.2657	USA

MNF-Pot-68	43.595	-86.2657	USA
Mnz-0	50.001	8.26664	GER
MOG-37	48.6667	-4.06667	FRA
Mrk-0	49	9.3	GER
Mt-0	32.34	22.46	LIB
Mz-0	50.3	8.3	GER
N13	61.36	34.15	RUS
N4	61.36	34.15	RUS
N7	61.36	34.15	RUS
Na-1	47.5	1.5	FRA
Nc-1	48.6167	6.25	FRA
NC-6	35	-79.18	USA
Nd-1	50	10	SUI
NFA-10	51.4083	-0.6383	UK
NFA-8	51.4083	-0.6383	UK
NFC-20	51.4083	-0.6383	UK
No-0	51.0581	13.2995	GER
Nok-1	52.24	4.45	NED
Nw-0	50.5	8.5	GER
Nw-2	50.5	8.5	GER
Nz1	-37.7871	175.283	NZL
Ob-1	50.2	8.5833	GER
Old-1	53.1667	8.2	GER
Omo2-1	56.14	15.78	SWE
Or-0	50.3827	8.01161	GER
Ors-1	44.7203	22.3955	ROU
Ost-0	60.25	18.37	SWE
Oy-0	60.23	6.13	NOR
Pa-2	38.07	13.22	ITA
PAR-3	46.65	-0.25	FRA
PAR-4	46.65	-0.25	FRA
PAR-5	46.65	-0.25	FRA
Paw-3	42.148	-86.431	USA
Pent-1	43.7623	-86.3929	USA
Per-1	58	56.3167	RUS
Petergof	59	29	RUS
PHW-13	51.2878	0.0565	UK
PHW-14	51.2878	0.0565	UK
PHW-20	51.2878	0.0565	UK
PHW-22	51.4167	-1.7167	UK
PHW-26	50.6728	-3.8404	UK
PHW-28	50.35	-3.5833	UK
PHW-31	51.4666	-3.2	UK
PHW-33	52.25	4.5667	NED
PHW-34	48.6103	2.3086	FRA
PHW-35	48.6103	2.3086	FRA
PHW-36	48.6103	2.3086	FRA
PHW-37	48.6103	2.3086	FRA
Pla-0	41.5	2.25	ESP
Pn-0	48.0653	-2.96591	FRA
Pna-17	42.0945	-86.3253	USA
Pog-0	49.2655	-123.206	CAN
Pr-0	50.1448	8.60706	GER
Pro-0	43.25	-6	ESP
Pu2-23	49.42	16.36	CZE
Pu2-24	49.42	16.36	CZE
Ra-0	46	3.3	FRA
Rak-2	49	16	CZE

Ren-1	48.5	-1.41	FRA
Rev-2	55.7	13.4	SWE
Rhen-1	51.9667	5.56667	NED
Rmx-A180	42.036	-86.511	USA
ROM-1	45.5333	4.85	FRA
Rou-0	49.4424	1.09849	FRA
RRS-10	41.5609	-86.4251	USA
RRS-7	41.5609	-86.4251	USA
Rsch-4	56.3	34	RUS
S96	NA	NA	UNK
Sanna-2	62.69	18	SWE
Sap-0	49.49	14.24	CZE
Sapporo-0	43.0553	141.346	JPN
Sav-0	49.1833	15.8833	CZE
Se-0	38.3333	-3.53333	ESP
Sei-0	46.5438	11.5614	ITA
Sg-1	47.6667	9.5	GER
Sh-0	51.6832	10.2144	GER
Shahdara	38.35	68.48	TJK
Si-0	50.8738	8.02341	GER
SLSP-30	43.665	-86.496	USA
Sp-0	52.5339	13.181	GER
Sparta-1	55.7097	13.0489	SWE
Sq-8	51.4083	-0.6383	UK
St-0	59	18	SWE
Ste-0	52.6058	11.8558	GER
Ste-3	42.03	-86.514	USA
T1040	55.6494	13.2147	SWE
T1060	55.6472	13.2225	SWE
T1080	55.6561	13.2178	SWE
T1110	55.6	13.2	SWE
T510	55.7936	13.1233	SWE
T540	55.7967	13.1044	SWE
T620	55.7	13.2	SWE
TÃ…D 01	62.8714	18.3447	SWE
Ta-0	49.5	14.5	CZE
Tamm-2	60	23.5	FIN
TDr-1	55.7683	14.1386	SWE
TDr-18	55.7714	14.1208	SWE
TDr-3	55.7686	14.1381	SWE
TDr-8	55.7706	14.1342	SWE
Tha-1	52.08	4.3	NED
Ting-1	56.5	14.9	SWE
Tiv-1	41.96	12.8	ITA
Tomegap-2	55.7	13.2	SWE
Tottarp-2	55.95	13.85	SWE
TOU-A1-115	46.6667	4.11667	FRA
TOU-A1-116	46.6667	4.11667	FRA
TOU-A1-12	46.6667	4.11667	FRA
TOU-A1-43	46.6667	4.11667	FRA
TOU-A1-62	46.6667	4.11667	FRA
TOU-A1-67	46.6667	4.11667	FRA
TOU-A1-96	46.6667	4.11667	FRA
TOU-C-3	46.6667	4.11667	FRA
TOU-E-11	46.6667	4.11667	FRA
TOU-H-12	46.6667	4.11667	FRA
TOU-H-13	46.6667	4.11667	FRA
TOU-I-17	46.6667	4.11667	FRA

TOU-I-2	46.6667	4.11667	FRA
TOU-I-6	46.6667	4.11667	FRA
TOU-J-3	46.6667	4.11667	FRA
TOU-K-3	46.6667	4.11667	FRA
Ts-1	41.7194	2.93056	ESP
Tscha-1	47.0748	9.9042	AUT
Tsu-0	34.43	136.31	JPN
Ty-0	56.4278	-5.23439	UK
Udul 1-34	49.2771	16.6314	CZE
Uk-1	48.0333	7.7667	GER
Uk-2	48.0333	7.7667	GER
UKID101	53.2	-1.4	UK
UKID37	51.3	1.1	UK
UKID48	54.7	-2.7	UK
UKID80	54.7	-2.9	UK
UKNW06-059	54.4	-3	UK
UKNW06-060	54.4	-3	UK
UKNW06-386	54.6	-3.1	UK
UKNW06-436	54.7	-3.4	UK
UKNW06-460	54.7	-3.4	UK
UKSE06-062	51.3	0.5	UK
UKSE06-192	51.3	0.5	UK
UKSE06-272	51.3	0.4	UK
UKSE06-278	51.3	0.4	UK
UKSE06-349	51.3	0.4	UK
UKSE06-351	51.3	0.4	UK
UKSE06-414	51.3	0.4	UK
UKSE06-429	51.3	0.4	UK
UKSE06-466	51.2	0.4	UK
UKSE06-482	51.2	0.6	UK
UKSE06-520	51.3	1.1	UK
UKSE06-628	51.1	0.4	UK
UKSW06-202	50.4	-4.9	UK
Ull2-3	56.0648	13.9707	SWE
Ull3-4	56.06	13.97	SWE
Uod-7	48.3	14.45	AUT
Utrecht	52.0918	5.1145	NED
VÃ¥r2-1	55.58	14.334	SWE
Van-0	49.3	-123	CAN
Ven-1	52.0333	5.55	NED
VOU-1	46.65	0.166667	FRA
VOU-2	46.65	0.166667	FRA
Wa-1	52.3	21	POL
Wag-3	51.9666	5.6666	NED
Wag-4	51.9666	5.6666	NED
Wag-5	51.9666	5.6666	NED
WAR	41.7302	-71.2825	USA
Wc-2	52.6	10.0667	GER
Wei-0	47.25	8.26	SUI
Wil-1	54.6833	25.3167	LTU
Wl-0	47.9299	10.8134	GER
Ws	52.3	30	RUS
Ws-0	52.3	30	RUS
Wt-3	52.3	9.3	GER
Wt-5	52.3	9.3	GER
Yo-0	37.45	-119.35	USA
Zdr-6	49.3853	16.2544	CZE
ZdrI 2-24	49.3853	16.2544	CZE

| ZdrI 2-25 | 49.3853 | 16.2544 | CZE |
| Zu-1 | 47.3667 | 8.55 | SUI |

Table A.2: *Arabidopsis thaliana* mutant lines used in this study

Mutant line	Gene ID	Back-ground	NASC ID	Source / Reference
uvr8-6	AT5G63860	Col	N533468	Roman Ulm; Alonso et al. 2003
uvr2	AT1G12370	Col	N857140	Chris Bowler; Castells et al. 2010
uvr3	AT3G15620		N864134	
uvr2 /	AT1G12370 /			
uvr3	AT3G15620			
tt4	AT5G13930	Col	N661740	European Arabidopsis Stock Center
uvb1-1	AT1G47870	All Col	N370059	European Arabidopsis Stock Center
uvb2-1	AT1G47885		N453858	
uvb3-1	AT1G47890		N855468	
uvb3-2	AT1G47890		N110784	
uvb3-3	AT1G47890		N65440	
uvb3-4	AT1G47890		N858010	
uvb3-5	AT1G47890		N643685	
uvb4-1	AT1G22080		N859250	
uvb5-1	AT1G56290		N684094	
uvb7-1	AT2G43000		N536473	
uvb8-1	AT2G42960		N656656	
uvb10-1	AT3G09720		N643440	
uvb11-1	AT3G25030		N24992	
uvb12-1	AT3G47660		N571329	
uvb13-1	AT3G61340		N500440	
uvb14-1	AT4G12120		N457681	
uvb16-1	AT5G46330		N670812	
uvb17-1	AT5G47040		N362989	
uvb18-1	AT5G53910		N117225	
uvb19-1	AT5G58130		N678222	
uvb20-1	AT5G63630		N590068	
uvb21-1	AT5G63920		N369999	
uvb26-1	AT1G47900		N657114	
uvb26-2	AT1G47900		N860238	
uvb27-1	AT4G31990		N653279	
uvb28-1	AT4G32000		N560167	
uvb29-1	AT5G53905		N655844	
uvb30-1	AT5G53902		N569428	
uvb31-1	AT1G47920		N639332	
uvb31-2	AT1G47920		N639334	
uvb31-3	AT1G47920		N820502	
uvb32-1	AT1G47880		N513123	
uvb32-2	AT1G47880		N597080	
uvb33-1	AT1G47840		N530722	
RNA1	AT2G19190	All Col	N818232	European Arabidopsis Stock Center
RNA5	AT4G22530		N655074	
RNA7	AT4G31870		N658985	
RNA8	AT5G20400		N678544	

RNA9	AT4G36670		N664159	
RNA10	AT4G27310		N545412	
RNA11	AT3G14710		N665756	
RNA12	AT3G22510		N682700	
RNA13	AT1G26970		N426351	
RNA14	AT1G30720		N55769	
RNA15	AT1G30370		N858390	
RNA16	AT1G21850		N538666	
RNA17	AT1G13310		N668333	
RNA18	AT4G25330		N655092	
RNA19	AT4G36510		N662005	
RNA20	AT5G66850		N121879	
RNA21	AT1G25370		N661684	
RNA22	AT1G48240		N664298	
RNA24	AT1G49230		N666146	
RNA25	AT1G56510		N660251	
RNA26	AT1G63350		N655376	
RNA27	AT1G64720		N682211	
RNA28	AT1G65070		N65670	
RNA29	AT1G68050		N677753	
RNA30	AT1G68490		N657365	
RNA31	AT1G72240		N666454	
RNA32	AT1G76360		N65580	
RNA33	AT2G17550		N667853	
RNA35	AT5G39000		N661219	
RNA36	AT5G23460		N665470	
RNA37	AT4G19160		N653742	
RNA38	AT4G15470		N807336	
RNA39	AT4G13350		N667833	
RNA40	AT2G37040		N661248	
RNA41	AT3G05370		N65469	
RNA42	AT3G45330		N685929	
RNA43	AT3G53810		N836974	
RNA44	AT3G25600		N524662	
RNA45	AT2G46510		N686454	
RNA46	AT2G46710		N658254	
RNA47	AT2G45910		N664874	
RNA48	AT5G24810		N661440	
RNA50	AT5G57220		N664866	
RNA52	AT4G35940		N654513	
RNA53	AT5G09800		N676366	

Table A.3: *Arabidopsis thaliana* RILs and NILs from ColxC24 mapping population

Recombinant Inbred Lines (RILs)	Near Isogenic Lines (NILs)
Q002, Q004, Q006, Q007, Q008, Q020, Q021, Q025, Q027, Q030, Q034, Q038, Q041, Q050, Q055, Q056, Q061, Q062, Q063, Q066, Q067, Q084, Q087, Q092, Q095, Q104, Q109, Q112, Q113, Q115, Q116, Q119, Q120, Q123, Q129, Q136, Q152, Q154, Q156, Q159, Q165, Q166, Q168, Q173, Q174, Q178, Q179, Q180, Q182, Q183, Q185, Q186, Q190, Q191, Q192, Q194, Q195, Q196, Q199, Q201, Q204, Q205, Q206, Q207, Q208, Q210, Q213, Q216, Q218, Q223, Q234, Q239, R002, R003, R005, R010, R015, R016, R017, R018, R034, R047, R053, R061, R062, R067, R070, R072, R098, R104, R108, R120, R123, R129, R130, R133, R134, R136, R143, R144, R148, R149, R156, R158, R164, R168, R170, R171, R174, R175, R179, R180, R182, R183, R188, R189, R194, R195, R196, R202, R216, R222	M12 , M18 , M21 , M37 , M39 , M44 , M45 , M46 , M49 , M54 , M57 , N09 , N17 , N20 , N31 , N38 , N39 , N41 , N48 , N52 , N58,

Table A.4 Oligonucleotides used in this study

Name	Locus		Purpose
E2F_CDS_seq_1	AT1G47870	TCCCGCTGTTTCCTCTCTTC	Sequencing
E2F_CDS_seq_2	AT1G47870	CACAACTGTCTCCTCTAATTTCAC	Sequencing
E2F_CDS_seq_3	AT1G47870	CGAATTCAGCTTGGATCTATTGC	Sequencing
E2F_CDS_seq_4	AT1G47870	CTGCCGTTATGACAGTTCTTTAG	Sequencing
E2F_CDS_seq_5	AT1G47870	CCTGCGATCCCTCTTAACAG	Sequencing
E2F_CDS_seq_6	AT1G47870	CCCCATTACTGACCGGTTACAT	Sequencing
E2F_CDS_seq_7	AT1G47870	CGAGTCTCCCACGGTTTCAG	Sequencing
E2F_CDS_seq_8	AT1G47870	CTCTCTGTGTCTGGTCTTATCTA	Sequencing
885_CDS_seq_1	AT1G47885	CTAAAGTGTTGATATCCGCAGTAT	Sequencing
885_CDS_seq_2	AT1G47885	GCTTCATACAAGCATGAGTGGT	Sequencing
UVR2_seq1	AT1G12370	GACAGCTTTTGCTTCTTCGTCT	Sequencing
UVR2_seq2	AT1G12370	GTGAGATTGTAAATGTGAAGAACC	Sequencing
UVR2_seq3	AT1G12370	GATAAGTCAAGTACCTCTTCTTCT	Sequencing
UVR2_seq4	AT1G12370	CCCACCTCCTCGAATAGTTGAT	Sequencing
UVR2_seq5	AT1G12370	CTACTACACAATAGTCAGTAAACAAA	Sequencing
UVR2_seq6	AT1G12370	ATTTAGGGCAGTTGATTGCGCT	Sequencing
UVR2_seq7	AT1G12370	GTTTGGAATTAAGACTAAGCTTTGG	Sequencing
UVR2_seq8	AT1G12370	CCAAATTACCTACCAAATTTTCCATTT	Sequencing
UVR2_seq9	AT1G12370	GATATCAGCTTTTGTCTCCCTATA	Sequencing
UVR2_seq10	AT1G12370	CTTGTATTTGTCACTTATTCATAGCA	Sequencing
UVR2_seq11	AT1G12370	GAGTAAACCCTAGATAAGCACGAT	Sequencing
UVR2_seq12	AT1G12370	CGACCCGACCAAATCAACTCA	Sequencing
UVR2_seq13	AT1G12370	CACAGTCTGACACGTGTTGATAA	Sequencing
UVR2_seq14	AT1G12370	GGGTTTGCGTCAGCTTCACCA	Sequencing
UVR2_seq15	AT1G12370	GGGCTGCTTCGAGTAAATTGG	Sequencing
UVR2_seq16	AT1G12370	GGCAGGTATCAGCTCAGCGTT	Sequencing
UVR2_seq17	AT1G12370	GAGAAGGGACTGACAGCAGAT	Sequencing
Chr1-Seq-1	GWAS peak Chr1	AACGCGTTTGCGAACGGGTTT	Sequencing
Chr1-Seq-2	GWAS peak Chr1	CCCTAAAGTACTCCAAATACCC	Sequencing
Chr1-Seq-3	GWAS peak Chr1	GCGAACGACTTTAAAGTTACTTTCT	Sequencing
Chr1-Seq-4	GWAS peak Chr1	GGCCTCTTCTCTCACATGTCTT	Sequencing

Chr1-Seq-5	GWAS peak Chr1	GAGTCAAGAGTAGTTATAAATCAATGA	Sequencing
Chr1-Seq-6	GWAS peak Chr1	CAACGTTAGATAACTTGTAAGATCACT	Sequencing
Chr1-Seq-7	GWAS peak Chr1	TGTCCAGTTTTACGAAAACAAGTTAA	Sequencing
Chr1-Seq-8	GWAS peak Chr1	GCTACAAAGTATGAAGCTCTAGAA	Sequencing
Chr1-Seq-10	GWAS peak Chr1	CTGTTAAAGGCTGGGAAACGAAT	Sequencing
Chr1-Seq-11	GWAS peak Chr1	CAAATGTCCTTGAAGGAATTGGAT	Sequencing
Chr1-Seq-12	GWAS peak Chr1	CGAGAGTGTTGCTAAGATCAATGA	Sequencing
Chr1-Seq-14	GWAS peak Chr1	CGAACCACCTGATTTGGACAAT	Sequencing
Chr1-Seq-15	GWAS peak Chr1	CCTGTCAAATAAGAGAAGGTATAGA	Sequencing
Chr1-Seq-16	GWAS peak Chr1	GGGCCGAGCTGATATATCACT	Sequencing
Chr1-Seq-17	GWAS peak Chr1	CTCACTGAGAACGAGTTTTGTTCTT	Sequencing
Chr1-Seq-18	GWAS peak Chr1	GTCAAAGTGTGTGTTTGGTTAAAGA	Sequencing
Chr1-Seq-20	GWAS peak Chr1	GGAGACCTAGTGTTTGTTTGGAA	Sequencing
Chr1-Seq-21	GWAS peak Chr1	GCCACAACAACAGTTGATATTAAG	Sequencing
Chr1-Seq-22	GWAS peak Chr1	GGTGAAGTTATTATTGAATCCCAAAA	Sequencing
Chr1-Seq-23	GWAS peak Chr1	CGGTTTTCCATTTTCCACATTGTT	Sequencing
Chr1-Seq-24	GWAS peak Chr1	TCTACATATATTTAATCTTGGCTCATA	Sequencing
Chr1-Seq-26	GWAS peak Chr1	CCAATCAATCTTCTTCAGCTAACAA	Sequencing
Chr1-Seq-27	GWAS peak Chr1	TATGTTGGTGGTAACAAGCTTAGT	Sequencing
Chr1-Seq-28	GWAS peak Chr1	AGCGGTTTCCATGTATCCGTAA	Sequencing
Chr1-Seq-29	GWAS peak Chr1	GCCAGAGTACATTCAGAATCCTA	Sequencing
Chr1-Seq-30	GWAS peak Chr1	GGTTCATGAAGACTTTCGGACG	Sequencing
Chr1-Seq-31	GWAS peak Chr1	AGTGATCTTACAAGTTATCTAACGTTG	Sequencing
Chr1-Seq-32	GWAS peak Chr1	TTCTAGAGCTTCATACTTTGTAGC	Sequencing
Chr1-Seq-33	GWAS peak Chr1	ATTCGTTTCCCAGCCTTTAACAG	Sequencing
Chr1-Seq-34	GWAS peak Chr1	TCTATACCTTCTCTTATTTGACAGG	Sequencing
Chr1-Seq-35	GWAS peak Chr1	TCTTTAACCAAACACACACTTTGAC	Sequencing
Chr1-Seq-36	GWAS peak Chr1	TAGGATTCTGAATGTACTCTGGC	Sequencing
Chr1-Seq-37	GWAS peak Chr1	ATTTATGGAGCAACTTCAACAGCTGA	Sequencing
Chr1-Seq-38	GWAS peak Chr1	CTCACCCACACACGAGAAGAT	Sequencing
Chr1-Seq-39	GWAS peak Chr1	CGGAGAACCGAACGTACAGAA	Sequencing
Chr1-Seq-40	GWAS peak Chr1	GTGAAGTTGAATGACTAGTTTAGAG	Sequencing
Chr1-Seq-41	GWAS peak Chr1	GGGATGGCGTGTGATGCTAA	Sequencing
Chr1-Seq-42	GWAS peak Chr1	GCTTCTTCAGCTAACCAAGTTGT	Sequencing
Chr1-Seq-43	GWAS peak Chr1	AGAAAACCATAGACCATGGGCTA	Sequencing
Chr1-Seq-44	GWAS peak Chr1	CGTAAATAGTAGAAACTGGGATAATT	Sequencing
Chr1-Seq-45	GWAS peak Chr1	GGCGTATACACTCGTCGTAATAT	Sequencing
Chr1-Seq-46	GWAS peak Chr1	AGAAAGTAACTTTAAAGTCGTTCGC	Sequencing
Chr1-Seq-47	GWAS peak Chr1	GGGAAATGTGACATGGAAATTAAAAG	Sequencing
Chr1-Seq-48	GWAS peak Chr1	GTTGTCACATGCCATAGGGACG	Sequencing
Chr1-Seq-49	GWAS peak Chr1	AATGAGCTTCCTTATTCGTTCCAT	Sequencing
Chr1-Seq-50	GWAS peak Chr1	GAACAATCCTAATCTGAGAGGCAA	Sequencing
Chr1-Seq-51	GWAS peak Chr1	AACCAGTTCACAGGTTCCCTTCCA	Sequencing
Chr1-Seq-52	GWAS peak Chr1	GGTCAAGTACCAGACTGGTTGT	Sequencing
Chr1-Seq-53	GWAS peak Chr1	AATGGAGGGAAAGCTTCCTGGTT	Sequencing
Chr1-Seq-LipTS49a	GWAS peak Chr1	CCATTGATAAATCTTTTCTTCCTCTAC	Sequencing
Chr1-Seq-C24TS24a	GWAS peak Chr1	GTCAATATTGACGTTGGTTTTAGGTT	Sequencing
47860-1_f	GWAS peak Chr1	CTGAAGTGCATAGCTTTGGAAGT	Sequencing
47860-2_f	GWAS peak Chr1	CGGGCATTACCGTAAATAACATCA	Sequencing
47870_r	GWAS peak Chr1	AGTAATTCGAAGCGAAAAGGAGAT	Sequencing
E2FcDNA_f	GWAS peak Chr1	ATGGCCGCGACATCAAACTCAG	Sequencing
E2FcDNA_r	GWAS peak Chr1	TCAGCTGTTGAAGTTGCTCCATAAAT	Sequencing
47870_f	GWAS peak Chr1	CAAAAAGCTCCAGTTGGAGTAGA	Sequencing
47880_r	GWAS peak Chr1	GCGTAGTCACTATATTAATAGCCT	Sequencing
47880_f	GWAS peak Chr1	ATTTCCAAGCAGTGTTCATCTGATA	Sequencing
47890_r	GWAS peak Chr1	ATGGAACGAATAAGGAAGCTCATT	Sequencing
47890_f	GWAS peak Chr1	ATATATGTAGTCGGGAAGACACTT	Sequencing
47900_r	GWAS peak Chr1	GCTCTCAAGGTCAGGTTCATCTA	Sequencing
47890_CDS-Insert_CAPS_F	AT1G47890	CCAGTTCACAGGTTCCCTTCCA	CAPS
47890_CDS-Insert_CAPS_R1	AT1G47890	GCTGCTTGAGAGACGAGAAGACAT	CAPS
47890_CDS-Insert_CAPS_R2	AT1G47890	GTGAGATATATAATGTGCCTAGCTG	CAPS
47890_5-Deletion_SSLP	AT1G47890	TTACTTTTCCTCACGAAACTCCAA	SSLP
47890_5-Deletion_SSLP	AT1G47890	AGAAAACATAAAGTTCCATAGACTTTGA	SSLP
47900DEL_F	AT1G47900	CGGCATAGTTCTTTGGGACTGA	SSLP
47900DEL_R		TTGGGGAAGCAGCTGAAGTCAT	SSLP
SSG1_SS52_F	1_3328157	CGGTTCCGCCTTTGAGTTTGT	Sequencing
SSG1_SS52_R		ACGAGGCAGAATCCAACATGG	Sequencing
SSG1_SS35_F	1_9308108	ACTCACACACCTTCTAGGTCCA	Sequencing
SSG1_SS35_R		AGTTTGATGGTCCAGCGTCAC	Sequencing
SSG1_SS3_F	1_14300109	TTACTCATGGGCAAGCTTTTATATTC	Sequencing
SSG1_SS3_R		AGCTCAAATAAAAGCCCCAAACG	Sequencing
SSG1_SS42_F	1_16579994	GTGCCAAAGGGACATGGCTAA	Sequencing

SSG1_SS42_R		TGCAGTTCCCTTCATTCGAGC	Sequencing
SSG1_SS41_F	1_24939108	AGCAAGGTCCACCAGGATTGA	Sequencing
SSG1_SS41_R		TCGGATAGAGGACATGGCTGT	Sequencing
SSG1_SS4_F	1_28779044	TCTAAGTGGTTGGATGCGTCA	Sequencing
SSG1_SS4_R		TGCCTCCTTGGCTGAAAACC	Sequencing
SSG1_SS9_F	2_9486871	CTCCATCCAGGAGCGGATCT	Sequencing
SSG1_SS9_R		TGGAACGAGTTTCACATCTCCA	Sequencing
SSG1_SS10_F	2_9892771	AGCGCCACACACATGTAACG	Sequencing
SSG1_SS10_R		TTCACCGGCCATAGTGGTCT	Sequencing
SSG1_SS50_F	2_12924867	GGGTGTTTTGTATTGGACTACGC	Sequencing
SSG1_SS50_R		CGACGTCAATCATGCGTCAGA	Sequencing
SSG1_SS44_F	2_18064955	CAGCATGCTCGAAAGTACGGT	Sequencing
SSG1_SS44_R		CCAAGGCTTTTTGCAGCATCC	Sequencing
SSG1_SS47_F	3_6300738	TTCAGGCGAGCCAGAAGTTTG	Sequencing
SSG1_SS47_R		TTCTGGCTTCGATGGCTTGAC	Sequencing
SSG1_SS46_F	3_8260080	TCTCTCTCTTGTTGGGTTGGC	Sequencing
SSG1_SS46_R		CTACGACGCTCCTCTAGCTCA	Sequencing
SSG1_SS48_F	3_11834921	CCCTCACCCCATCAGTGACTT	Sequencing
SSG1_SS48_R		GGTCGACACCTCACACGTTTA	Sequencing
SSG1_SS15_F	3_13059542	CTGCACACCGCGAAAGAGAT	Sequencing
SSG1_SS15_R		TCACTTCTCACCTTTCTTCCATGA	Sequencing
SSG1_SS53_F	3_13542758	AGCCACTATCTTCAGGGGACC	Sequencing
SSG1_SS53_R		GTGGGCGGACTGAGTTCTTTC	Sequencing
SSG1_SS54_F	3_21296433	AGACAAAGTCGACCACCAGGA	Sequencing
SSG1_SS54_R		TTGGCCTTTGGCTGTGTGTAG	Sequencing
SSG1_SS19_F	4_559207	ACAGGTATAGCCAACACAGGTA	Sequencing
SSG1_SS19_R		GGATGTCACGATTTAATATATGACT	Sequencing
SSG1_SS20_F	4_2908151	GATGCCTCGTATGGCTGTCA	Sequencing
SSG1_SS20_R		AGCCACACTTGGTCTCAGGA	Sequencing
SSG1_SS51_F	4_6695937	CCCAAATCGCCGCATACTTCT	Sequencing
SSG1_SS51_R		ATAATATGACCCGTGGCCCGT	Sequencing
SSG1_SS22_F	4_6990674	CCAGGTCCCGGGAAAGAAGA	Sequencing
SSG1_SS22_R		GCGAAGAGGCAAGTGGCAAA	Sequencing
SSG1_SS25_F	5_5672375	TCTGCCCATGGTCACCCATT	Sequencing
SSG1_SS25_R		CGGTGCTCCTGCTATTGTTGA	Sequencing
SSG1_SS43_F	5_11068534	TCGATGAGGGGGTTCTAGCAG	Sequencing
SSG1_SS43_R		GGAGGCAATGGAAGTTGGAGG	Sequencing
SSG1_SS27_F	5_17593720	ATGTGTGGTGAGAGAGACCGA	Sequencing
SSG1_SS27_R		TGGAGCTGAGAAAAGGTGGGA	Sequencing
SSG1_SS55_F	5_22936341	CACGAGCCAAAGATCCGTGAA	Sequencing
SSG1_SS55_R		AAAGTTTCCAACGAAGCCGGT	Sequencing
SSG1_SS28_F	5_24258109	TTTCAGCAAGAGACGCCGTT	Sequencing
SSG1_SS28_R		CGGTTGAGCCAGAAAACCCA	Sequencing
SAIL-LB3-Primer	/	TAGCATCTGAATTTCATAACCAATCTCGATACAC	Genotyping
SM_Primer	/	TACGAATAAGAGCGTCCATTTTAGAGTGA	Genotyping
LB-TP	/	GTCCGCAATGTGTTATTAAGTTGTC	Genotyping
LB-b1.3-TP	/	GTGGTTCCGAAATCGGCAAAATC	Genotyping
UVB1-1_LP	AT1G47870	GCTCACCTCTGCATCTGACTC	Genotyping
UVB1-1_RP		GACAAAAGGACCTGGGAGATC	Genotyping
UVB2-1_LP	AT1G47885	GTCTCGTGGCACACTTGCCAA	Genotyping
UVB2-1_RP		AAACCTGCTGCTACCCAACTC	Genotyping
UVB3-1_LP	AT1G47890	ATTCAACGTTGGCATTCTCCACAA	Genotyping
UVB3-1_RP		TCAACAATTCGCCAATACCAGCT	Genotyping
UVB3-2_LP2	AT1G47890	ATAAGAAACGGAAAGGGTCAAAG	Genotyping
UVB3-2_RP2		AGGTTTGAAATTGGCTAGAGGAC	Genotyping
UVB4-1_LP	AT1G47900	CTGCAGCTACCAGAGTCGTCATT	Genotyping
UVB4-1_RP		AGTGAAGCAGCACTCTAAGGTTG	Genotyping
UVB4-2_LP	AT1G47900	TCCATACAAGCATTCACATCTTGC	Genotyping
UVB4-2_RP		GGAGTTTGAAAAGAGGATGTGTG	Genotyping
tt4-2_LP	AT5G13930	TGAGTTTGAGACCTTGTGGATGTA	Genotyping
tt4-2_RP		AAACCCACACATGTGTGCTTACA	Genotyping
UVR2ko_f	AT1G12370	GAACTCTCCGACAATTTTTGCTA	Genotyping
UVR2ko_r		CCAACACAGGGCAAAGTAGTC	Genotyping
UVR3ko_f2	AT3G15620	GTTCCGGAAAGGTCTCCGGGTT	Genotyping
UVR3ko_r2		GGAGTCCAATCAGCCTACAGAC	Genotyping
RNA_1_LP	AT2G19190	TTCCACCGATCATTAACGCCCT	Genotyping
RNA_1_RP		CATGAGAAGGTCAACCTGCGA	Genotyping
RNA_10_LP2	AT4G27310	ATCGAAAGAACTATGGTGAACAAAG	Genotyping
RNA_10_RP2		AGGAGGAGGTTAATAGTTTCGAGTC	Genotyping
RNA_12_2_LP	AT3G22510	AACGACCTAAAGCTTCACCATTAC	Genotyping
RNA_12_2_RP		TCTTTGAAGAACAGAGTCGTCATC	Genotyping
RNA_13_LP	AT1G26970	TGTTGATTTAACAGGCGGAGGT	Genotyping

RNA_13_RP		GAAATAGAACCCGACTTCAACG	Genotyping
RNA_14_LP	AT1G30720	AGATCGGTTCTCGAAATCAGAC	Genotyping
RNA_14_RP		CGGACTAATAGTTAATCCCTTGACA	Genotyping
RNA_15_LP	AT1G30370	ATTCAATCCTTAAGCTGCCAAG	Genotyping
RNA_15_RP		CTCGTCCTAAGATCCATGAACC	Genotyping
RNA_16_LP	AT1G21850	TACACCGTTCTCATCGGTGACT	Genotyping
RNA_16_RP		AGGCTTTAGAGGACTCTCACCTG	Genotyping
RNA_20_LP2	AT5G66850	TCAGCTCTTTGTCCAGTAAGCTGAAGA	Genotyping
RNA_20_RP		TTCTGCCGTTCCCTTACCCTTA	Genotyping
RNA_38_LP	AT4G15470	TGCATGATTGCATGTGATGTGTG	Genotyping
RNA_38_RP		AAGATTTTGGTCGGTTTGGGCT	Genotyping
RNA_43_LP	AT3G53810	AAAGAGGTCTGGTCTACCTCCC	Genotyping
RNA_43_RP		GAAGTTGTTTCCGGTAGACGAC	Genotyping
RNA_44_LP	AT3G25600	ACCAAGTGCATGCACTTAAATG	Genotyping
RNA_44_RP		CTGTGTGATGTGCTTCCTACCC	Genotyping
UVR2_TOPO_F	AT1G12370	CATTTGAATCTCTTTTTTGATCAATTTGAT	Gateway clon.
UVR2_TOPO_R	AT1G12370	GTAAACAATAGTTATCTTGGGATCAACA	Gateway clon.
UVR2pro_TOPO_R	AT1G12370	TACGAAGGAGGAGAGACGGATTTT	Gateway clon.
TAIL_AD1	/	NGTCGASWGANAWGAA	TAIL-PCR
TAIL_AD2	/	TGWGNAGSANCASAGA	TAIL-PCR
TAIL_AD3	/	AGWGNAGWANCAWAGG	TAIL-PCR
TAIL_AD4	/	STTGNTASTNCTNTGC	TAIL-PCR
TAIL_AD5	/	NTCGASTWTSGWGTT	TAIL-PCR
TAIL_AD6	/	WGTGNAGWANCANAGA	TAIL-PCR
LB1	/	GGCCGGAGAACCTGCGTGCAAT	TAIL-PCR
LB2	/	CTAATTGGATACCGAGGGGAATTT	TAIL-PCR
LB3	/	GGCGACTTTTGAACGCGCAATAA	TAIL-PCR
RBnew1	/	ATCAGGCGGGCAAGAATGTGAAT	TAIL-PCR
RBnew2	/	GCTCCTGAAAATCTCGCCGGAT	TAIL-PCR
RBnew3	/	GTGACTGGGAAAACCCTGGCGT	TAIL-PCR
UVR2q_f	AT1G12370	CAGTTAGAGAAGGGACTGACAGCAGAT	RT-qPCR
UVR2q_r	AT1G12370	TTCAGGTCCCTTGGTCCATTCTAGAAT	RT-qPCR
Act7qF	AT5G09810	TGGATCGGAGGATCAATCCTTG	RT-qPCR
Act7qR	AT5G09810	GACTCATCGTACTCACTCTTTGAA	RT-qPCR
UVR2/TAG_F	AT1G12370	TGAAGAACAGCTCACCAGAGACTCT	RT-qPCR
UVR2/LUC_R	AT1G12370	GTTGCTCTCCAGCGGTTCCATCTT	RT-qPCR
LUC_F	/	AAACGGATTACCAGGGATTTCAGT	RT-qPCR
LUC_R	/	GTCTTGTCCCTATCGAAGGACTCT	RT-qPCR
E2Fc_q_F	AT1G47870	TGAGTTTTCCGCAACAGTACCGGAT	RT-qPCR
E2Fc_q_R		AGTGTCTACTCCAACTGGAGCTTTTT	RT-qPCR
47880_qF	AT1G47880	GCAGTGTTCATCTGATACCCAGCT	RT-qPCR
47880_qR		GATGCGTAGTCACTATATTAATAGCCT	RT-qPCR
47885_qF	AT1G47885	CTAAAGTCATGGGACGATACAAAAGA	RT-qPCR
47885_qR		ATACCTCAGTCAATATGTGGACTCAA	RT-qPCR
47890_qF	AT1G47890	CACATAACCAGCTTGTAGGTTCCA	RT-qPCR
47890_qR2		GAAGGACCATTAAGTCCGGGGTTT	RT-qPCR
47890_qF2	AT1G47890	CAAGCTTAGTGGAAACTTGCCAGCTA	RT-qPCR
47890_qR2		GTGGAAGGGAACCTGTGAACTGGTT	RT-qPCR
47890_qF3	AT1G47890	CCTACGAAGGAAACCCCGGACTTA	RT-qPCR
47890_qR3		CTTCTTTTGTCTCCAACGGTTCG	RT-qPCR
47900_qF	AT1G47900	CGATGTCGTCGTATCGCTAGTCTCT	RT-qPCR
47900_qR		GAACTAGGCACCACCACTACCTCA	RT-qPCR
47890_qF4	AT1G47890	CACAGGTTCCCTTCCACCTAGCAT	RT-qPCR
47890_qR4		TCAATCCCAACAAGGTCGTTGAGT	RT-qPCR
UVR8qF	AT5G63860	TCCATTTGTACCTCTACCCCAA	RT-qPCR
UVR8qR		TTCTCCTGTGCAAGTGCGATTT	RT-qPCR

7 References

Abraham RT (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev **15**: 2177–2196

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol **8**: 135–141

Agati G, Tattini M (2010) Multiple functional roles of flavonoids in photoprotection. New Phytol **186**: 786–793

A-H-Mackerness S, Surplus SL, Blake P, John CF, Buchanan-Wollaston V (1999). Ultraviolet-B-induced stress and changes in gene expression in Arabidopsis thaliana: role of signalling pathways controlled by jasmonic acid, ethylene and reactive oxygen species. Plant Cell Environ. **22**: 1413–23

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) Molecular Biology of the Cell, 4th ed. Garland Science

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, Reymond M, Vreugdenhil D, Koornneef M (2009) What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell **21**: 1877–1896

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science **301**: 653–657

Amiard S, Charbonnel C, Allain E, Depeiges A, White CI, Gallego ME (2010) Distinct roles of the ATR kinase and the Mre11-Rad50-Nbs1 complex in the maintenance of chromosomal stability in Arabidopsis. Plant Cell **22**: 3020–3033

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol **11**: R106

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol **55**: 373–399

Arends D, Prins P, Jansen RC, Broman KW (2010) R/qtl: high-throughput multiple QTL mapping. Bioinforma Oxf Engl **26**: 2990–2992

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, et al (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature **465**: 627–631

Ballaré CL (2014) Light regulation of plant defense. Annu Rev Plant Biol **65**: 335–363

Barta C, Kalai T, Hideg K, Vass I, Hideg E (2004). Differences in the ROS- generating efficacy of various ultraviolet wavelengths in detached spinach leaves. Functional Plant Biology **31**: 23-28.

Berckmans B, Lammens T, Van Den Daele H, Magyar Z, Bögre L, De Veylder L (2011) Light-dependent regulation of DEL1 is determined by the antagonistic action of E2Fb and E2Fc. Plant Physiol **157**: 1440–1451

Bergelson J, Roux F (2010) Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. Nat Rev Genet **11**: 867–879

Bieza K, Lois R (2001) An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol **126**: 1105–1115

Bleuyard J-Y, Gallego ME, White CI (2004) Meiotic defects in the Arabidopsis rad50 mutant point to conservation of the MRX complex function in early stages of meiotic recombination. Chromosoma **113**: 197–203

Bray CM, West CE (2005) DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytol **168**: 511–528

Britt, AB (1996). DNA damage and repair in plants. Annu Rev Plant Physiol Plant Mol Biol **47**: 75–100

Britt, AB (2004). Repair of DNA damage induced by solar UV. Photosynthesis Research **81**: 105–112

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinforma Oxf Engl **19**: 889–890

Brown BA, Cloix C, Jiang GH, Kaiserli E, Herzyk P, Kliebenstein DJ, Jenkins GI (2005) A UV-B-specific signaling component orchestrates plant UV protection. Proc Natl Acad Sci U S A **102**: 18225–18230

Brown BA, Jenkins GI (2008) UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature Arabidopsis leaf tissue by requirement for UVR8, HY5, and HYH. Plant Physiol **146**: 576–588

Bucceri A, Kapitza K, Thoma F (2006) Rapid accessibility of nucleosomal DNA in yeast on a second time scale. EMBO J **25**: 3123–3132

Caldwell MM, Bornman JF, Ballaré CL, Flint SD, Kulandaivelu G (2007) Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with other climate change factors. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol **6**: 252–266

Campi M, D'Andrea L, Emiliani J, Casati P (2012) Participation of chromatin-remodeling proteins in the repair of ultraviolet-B-damaged DNA. Plant Physiol **158**: 981–995

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, et al (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet **43**: 956–963

Castells E, Molinier J, Benvenuto G, Bourbousse C, Zabulon G, Zalc A, Cazzaniga S, Genschik P, Barneche F, Bowler C (2011) The conserved factor DE-ETIOLATED 1 cooperates with CUL4-DDB1DDB2 to maintain genome integrity upon UV stress. EMBO J **30**: 1162–1172

Choi J-H, Besaratinia A, Lee D-H, Lee C-S, Pfeifer GP (2006) The role of DNA polymerase iota in UV mutational spectra. Mutat Res **599**: 58–65

Christie JM, Arvai AS, Baxter KJ, Heilmann M, Pratt AJ, O'Hara A, Kelly SM, Hothorn M, Smith BO, Hitomi K, et al (2012) Plant UVR8 photoreceptor senses UV-B by tryptophan-mediated disruption of cross-dimer salt bridges. Science **335**: 1492–1496

Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA, et al (2007) Common sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana. Science **317**: 338–342

Clark TA, Spittle KE, Turner SW, Korlach J (2011) Direct detection and sequencing of damaged DNA bases. Genome Integr **2**: 10

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J Cell Mol Biol **16**: 735–743

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature **452**: 215–219

Cooley NM, Higgins JT, Holmes MG, Attridge TH (2001) Ecotypic differences in responses of Arabidopsis thaliana L. to elevated polychromatic UV-A and UV-B+A radiation in the natural environment: a positive correlation between UV-B+A inhibition and growth rate. J Photochem Photobiol B **60**: 143–150

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in Escherichia coli. Nature **274**: 775–780

Culligan KM, Robertson CE, Foreman J, Doerner P, Britt AB (2006) ATR and ATM play both distinct and additive roles in response to ionizing radiation. Plant J Cell Mol Biol **48**: 947–961

Culligan K, Tissier A, Britt A (2004) ATR regulates a G2-phase cell-cycle checkpoint in Arabidopsis thaliana. Plant Cell **16**: 1091–1104

DeGregori J, Johnson DG (2006) Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. Curr Mol Med **6**: 739–748

Denver DR, Dolan PC, Wilhelm LJ, Sung W, Lucas-Lledó JI, Howe DK, Lewis SC, Okamoto K, Thomas WK, Lynch M, et al (2009) A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci U S A **106**: 16310–16314

Duncan BK, Miller JH (1980) Mutagenic deamination of cytosine residues in DNA. Nature **287**: 560–561

De Jager SM, Scofield S, Huntley RP, Robinson AS, den Boer BGW, Murray JAH (2009) Dissecting regulatory pathways of G1/S control in Arabidopsis: common and distinct targets of CYCD3;1, E2Fa and E2Fc. Plant Mol Biol **71**: 345–365

Del Pozo JC, Boniotti MB, Gutierrez C (2002) Arabidopsis E2Fc functions in cell division and is degraded by the ubiquitin-SCF(AtSKP2) pathway in response to light. Plant Cell **14**: 3057–3071

Del Pozo JC, Diaz-Trivino S, Cisneros N, Gutierrez C (2006) The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCFSKP2A pathway in Arabidopsis. Plant Cell **18**: 2224–2235

De Schutter K, Joubès J, Cools T, Verkest A, Corellou F, Babiychuk E, Van Der Schueren E, Beeckman T, Kushnir S, Inzé D, et al (2007) Arabidopsis WEE1 kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint. Plant Cell **19**: 211–225

Dickson KA, Haigis MC, Raines RT (2005) Ribonuclease inhibitor: structure and function. Prog Nucleic Acid Res Mol Biol **80**: 349–374

Emiliani J, Grotewold E, Falcone Ferreyra ML, Casati P (2013) Flavonols protect Arabidopsis plants against UV-B deleterious effects. Mol Plant **6**: 1376–1379

Essen LO, Klar T (2006) Light-driven DNA repair by photolyases. Cell Mol Life Sci CMLS **63**: 1266–1277

Favory J-J, Stec A, Gruber H, Rizzini L, Oravecz A, Funk M, Albert A, Cloix C, Jenkins GI, Oakeley EJ, et al (2009) Interaction of COP1 and UVR8 regulates UV-B-induced photomorphogenesis and stress acclimation in Arabidopsis. EMBO J **28**: 591–601

Fini A, Brunetti C, Di Ferdinando M, Ferrini F, Tattini M (2011) Stress-induced flavonoid biosynthesis and the antioxidant machinery of plants. Plant Signal Behav **6**: 709–711

Fransz PF, Armstrong S, de Jong JH, Parnell LD, van Drunen C, Dean C, Zabel P, Bisseling T, Jones GH (2000) Integrated cytogenetic map of chromosome arm 4S of A. thaliana: structural organization of heterochromatic knob and centromere region. Cell **100**: 367–376

Fritz-Laylin LK, Krishnamurthy N, Tör M, Sjölander KV, Jones JDG (2005) Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. Plant Physiol **138**: 611–623

Fryxell KJ, Moon W-J (2005) CpG mutation rates in the human genome are highly dependent on local GC content. Mol Biol Evol **22**: 650–658

Futuyma D (2005) Evolution. 2nd edition. Sinauer Associates, Inc.

Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, et al (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. Nature **477**: 419–423

García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A (2012) Qualimap: evaluating next-generation sequencing alignment data. Bioinforma Oxf Engl **28**: 2678–2679

Garcia V, Bruchet H, Camescasse D, Granier F, Bouchez D, Tissier A (2003) AtATM is essential for meiosis and the somatic response to DNA damage in plants. Plant Cell **15**: 119–132

Gerhardt KE, Wilson MI, Greenberg BM (2005) Ultraviolet wavelength dependence of photomorphological and photosynthetic responses in Brassica napus and Arabidopsis thaliana. Photochem Photobiol **81**: 1061–1068

González Besteiro MA, Bartels S, Albert A, Ulm R (2011) Arabidopsis MAP kinase phosphatase 1 and its target MAP kinases 3 and 6 antagonistically determine UV-B stress tolerance, independent of the UVR8 photoreceptor pathway. Plant J Cell Mol Biol **68**: 727–737

González Besteiro MA, Ulm R (2013a) Phosphorylation and stabilization of Arabidopsis MAP kinase phosphatase 1 in response to UV-B stress. J Biol Chem **288**: 480–486

González Besteiro MA, Ulm R (2013b) ATR and MKP1 play distinct roles in response to UV-B stress in Arabidopsis. Plant J Cell Mol Biol **73**: 1034–1043

Greilhuber J, Borsch T, Müller K, Worberg A, Porembski S, Barthlott W (2006) Smallest angiosperm genomes found in lentibulariaceae, with chromosomes of bacterial size. Plant Biol Stuttg Ger **8**: 770–777

Grimm D, Hagmann J, Koenig D, Weigel D, Borgwardt K (2013) Accurate indel prediction using paired-end short reads. BMC Genomics **14**: 132

Gruber H, Heijde M, Heller W, Albert A, Seidlitz HK, Ulm R (2010) Negative feedback regulation of UV-B-induced photomorphogenesis and stress acclimation in Arabidopsis. Proc Natl Acad Sci U S A **107**: 20132–20137

He G, Elling AA, Deng XW (2011) The epigenome and plant development. Annu Rev Plant Biol **62**: 411–435

Heijde M, Ulm R (2012) UV-B photoreceptor-mediated signalling in plants. Trends Plant Sci **17**: 230–237

Heilmann M, Jenkins GI (2013) Rapid reversion from monomer to dimer regenerates the ultraviolet-B photoreceptor UV RESISTANCE LOCUS8 in intact Arabidopsis plants. Plant Physiol **161**: 547–555

Henri Mme V, V Henri (1914). Variation du pouvoir abiotique des rayons ultraviolets avec leur longueur d'onde. C R Seances Soc Biol Fil **73**: 321–322

Hernando-Herraez I, Prado-Martinez J, Garg P, Fernandez-Callejo M, Heyn H, Hvilsom C, Navarro A, Esteller M, Sharp AJ, Marques-Bonet T (2013) Dynamics of DNA methylation in recent human and great ape evolution. PLoS Genet **9**: e1003763

Hershberg R, Petrov DA (2010) Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet **6**: e1001115

Hideg E, Vass I (1996). UV-B induced free radical production in plant leaves and isolated thylakoid membranes. Plant Science **115**: 251-260

Hockberger PE (2002) A history of ultraviolet photobiology for humans, animals and microorganisms. Photochem Photobiol **76**: 561–579

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, et al (2012) Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. Nat Genet **44**: 212–216

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Adv Bioinforma **2008**: 420747

Ikehata H, Masuda T, Sakata H, Ono T (2003) Analysis of mutation spectra in UVB-exposed mouse skin epidermis and dermis: frequent occurrence of C-->T transition at methylated CpG-associated dipyrimidine sites. Environ Mol Mutagen **41**: 280–292

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature **472**: 115–119

Iyer RR, Pluciennik A, Burdett V, Modrich PL (2006) DNA mismatch repair: functions and mechanisms. Chem Rev **106**: 302–323

Jackson RC (1959). A study of meiosis in Haplopappus gracilis. American Journal of Botany **46**: 550-554

Jansen MAK, Martret BL, Koornneef M (2010) Variations in constitutive and inducible UV-B tolerance; dissecting photosystem II protection in Arabidopsis thaliana accessions. Physiol Plant **138**: 22–34

Jenkins GI (2009) Signal transduction in responses to UV-B radiation. Annu Rev Plant Biol **60**: 407–431

Jiang CZ, Yee J, Mitchell DL, Britt AB (1997) Photorepair mutants of Arabidopsis. Proc Natl Acad Sci U S A **94**: 7441–7445

Kaiserli E, Jenkins GI (2007) UV-B promotes rapid nuclear translocation of the Arabidopsis UV-B specific signaling component UVR8 and activates its function in the nucleus. Plant Cell **19**: 2662–2673

Kalbina I, Strid A (2006) Supplementary ultraviolet-B irradiation reveals differences in stress responses between Arabidopsis thaliana ecotypes. Plant Cell Environ **29**: 754–763

Kami C, Lorrain S, Hornitschek P, Fankhauser C (2010) Light-regulated plant growth and development. Curr Top Dev Biol **91**: 29–66

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet **42**: 348–354

Keightley PD, Ness RW, Halligan DL, Haddrill PR (2014) Estimation of the spontaneous mutation rate per nucleotide site in a Drosophila melanogaster full-sib family. Genetics **196**: 313–320

Keller I, Bensasson D, Nichols RA (2007) Transition-transversion bias is not universal: a counter example from grasshopper pseudogenes. PLoS Genet **3**: e22

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J Cell Mol Biol **50**: 347–363

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol **14**: R36

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nat Genet **39**: 1151–1155

Kimura S, Sakaguchi K (2006) DNA repair in plants. Chem Rev **106**: 753–766

Kimura S, Tahira Y, Ishibashi T, Mori Y, Mori T, Hashimoto J, Sakaguchi K (2004) DNA repair in higher plants; photoreactivation is the major DNA repair pathway in non-proliferating cells while excision repair (nucleotide excision repair and base excision repair) is active in proliferating cells. Nucleic Acids Res **32**: 2760–2767

Kinoshita T, Caño-Delgado A, Seto H, Hiranuma S, Fujioka S, Yoshida S, Chory J (2005) Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1. Nature **433**: 167–171

Kliebenstein DJ, Lim JE, Landry LG, Last RL (2002) Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. Plant Physiol **130**: 234–243

Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol **11**: 725–732

Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Jonasdottir A (2012) Rate of de novo mutations and the importance of father's age to disease risk. Nature **488**: 471–475

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinforma Oxf Engl **23**: 1289–1291

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

Krakauer DC, Plotkin JB (2002) Redundancy, antiredundancy, and the robustness of genomes. Proc Natl Acad Sci U S A **99**: 1405–1409

Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, Ausubel FM (1992) Regulation of Flavonoid Biosynthetic Genes in Germinating Arabidopsis Seedlings. Plant Cell **4**: 1229–1236

Kurz EU, Lees-Miller SP (2004) DNA damage-induced activation of ATM and ATM-dependent signaling pathways. DNA Repair **3**: 889–900

Lammens T, Li J, Leone G, De Veylder L (2009) Atypical E2Fs: new players in the E2F transcription factor family. Trends Cell Biol **19**: 111–118

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods **9**: 357–359

Lariguet P, Dunand C (2005) Plant photoreceptors: phylogenetic overview. J Mol Evol **61**: 559–569

Lario LD, Ramirez-Parra E, Gutierrez C, Casati P, Spampinato CP (2011) Regulation of plant MSH2 and MSH6 genes in the UV-B-induced DNA damage response. J Exp Bot **62**: 2925–2937

Lee J, He K, Stolc V, Lee H, Figueroa P, Gao Y, Tongprasit W, Zhao H, Lee I, Deng XW (2007) Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. Plant Cell **19**: 731–749

Lee J-H, Paull TT (2005) ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. Science **308**: 551–554

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. Bioinforma Oxf Engl **25**: 2078–2079

Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL (1993) Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. Plant Cell **5**: 171–179

Li J, Ran C, Li E, Gordon F, Comstock G, Siddiqui H, Cleghorn W, Chen H-Z, Kornacker K, Liu C-G, et al (2008) Synergistic function of E2F7 and E2F8 is essential for cell survival and embryonic development. Dev Cell **14**: 62–75

Lindahl T, Nyberg B (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. Biochemistry (Mosc) **13**: 3405–3410

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell **133**: 523–536

Liu Z, Hossain GS, Islas-Osuna MA, Mitchell DL, Mount DW (2000) Repair of UV damage in plants by nucleotide excision repair: Arabidopsis UVH1 DNA repair gene is a homolog of Saccharomyces cerevisiae Rad1. Plant J Cell Mol Biol **21**: 519–528

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO (2010) Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. Proc Natl Acad Sci U S A **107**: 21199–21204

Lo H-L, Nakajima S, Ma L, Walter B, Yasui A, Ethell DW, Owen LB (2005) Differential biologic effects of CPD and 6-4PP UV-induced DNA damage on the induction of apoptosis and cell-cycle arrest. BMC Cancer **5**: 135

Lynch M (2010) Rate, molecular spectrum, and consequences of human mutation. Proc Natl Acad Sci U S A **107**: 961–968

Mazza CA, Boccalandro HE, Giordano CV, Battista D, Scopel AL, Ballaré CL (2000) Functional significance and induction by solar radiation of ultraviolet-absorbing sunscreens in field-grown soybean crops. Plant Physiol **122**: 117–126

McCombie W R, De la Bastide M, Habermann K, Parnell L, Dedhia N, Gnoj L, Schutz K, Huang E, Spiegel L, Yordan C, Sehkon M, Murray J, Sheet P, Cordes M, Threideh J, Stoneking T, Kalicki J, Graves T, Harmon G, Edwards J, Latreille P, Courtney L, Cloud J, Abbott A, Scott K, Johnson D, Minx P, Bentley D, Fulton B, Miller N, Greco T, Kemp K, Kramer J, Fulton L, Mardis E, Dante M, Pepin K, Hillier L, Nelson J, Spieth J, Simorowski J, May B, Ma P, Preston R, Vil D, Lei HS, Shekher M, Matero A, Shah R, Swaby I, O'Shaughnessy A, Rodriguez M, Hoffman J, Till S, Granat S, Shohdy N, Hasegawa A, Hameed A, Lodhi M, Johnson A, Chen E, Marra M, Wilson RK, Martienssen R (2000). The complete sequence of a heterochromatic island from a higher eukaryote. Cell 100: 377-386

McKenzie RL, Bjorn LO, Bais A, Ilyasd M (2003). Changes in biologically active ultraviolet radiation reaching the Earth's surface. Photochem Photobiol Sci **2**: 5-15

McKenzie RL, Aucamp PJ, Bais AF, Björn LO, Ilyas M (2007) Changes in biologically-active ultraviolet radiation reaching the Earth's surface. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol **6**: 218–231

McKenzie RL, Aucamp PJ, Bais AF, Björn LO, Ilyas M, Madronich S (2011) Ozone depletion and climate change: impacts on UV radiation. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol **10**: 182–198

McLennan, AG (1987). The repair of ultraviolet light-induced DNA damage in plant cells. Mutation Research **181**: 1-7

Morris ER, Walker JC (2003) Receptor-like protein kinases: the keys to response. Curr Opin Plant Biol **6**: 339–342

Müller V, Albert A, Barbro Winkler J, Lankes C, Noga G, Hunsche M (2013) Ecologically relevant UV-B dose combined with high PAR intensity distinctly affect plant growth and accumulation of secondary metabolites in leaves of Centella asiatica L. Urban. J Photochem Photobiol B **127**: 161–169

Nadeau JA, Sack FD (2002) Control of stomatal distribution on the Arabidopsis leaf surface. Science **296**: 1697–1700

Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J Biosci Bioeng **104**: 34–41

Neff MM, Fankhauser C, Chory J (2000) Light: an indicator of time and place. Genes Dev **14**: 257–271

Ness RW, Morgan AD, Colegrave N, Keightley PD (2012) Estimate of the spontaneous mutation rate in Chlamydomonas reinhardtii. Genetics **192**: 1447–1454

Obayashi T, Nishida K, Kasahara K, Kinoshita K (2011) ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. Plant Cell Physiol **52**: 213–219

Oravecz A, Baumann A, Máté Z, Brzezinska A, Molinier J, Oakeley EJ, Adám E, Schäfer E, Nagy F, Ulm R (2006) CONSTITUTIVELY PHOTOMORPHOGENIC1 is required for the UV-B response in Arabidopsis. Plant Cell **18**: 1975–1990

Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D (2008) Sequencing of natural strains of Arabidopsis thaliana with short reads. Genome Res **18**: 2024–2033

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M (2010) The rate and molecular spectrum of spontaneous mutations in Arabidopsis thaliana. Science **327**: 92–94

Palomera-Sanchez Z, Zurita M (2011) Open, repair and close again: chromatin dynamics and the response to UV-induced DNA damage. DNA Repair **10**: 119–125

Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N, Mittelsten Scheid O (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. Plant Cell **22**: 3118–3129

Pellicer J, Fay MF, Leitch IJ (2010). The largest eukaryotic genome of them all? Botanical Journal of the Linnean Society **164**: 10-15

Proost S, Pattyn P, Gerats T, Van de Peer Y (2011) Journey through the past: 150 million years of plant genome evolution. Plant J Cell Mol Biol **66**: 58–65

Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot **56**: 1–14

Pukacki PM, Modrzyński J (1998) The influence of ultraviolet-B radiation on the growth, pigment production and chlorophyll fluorescence of Norway spruce seedlings. Acta Physiol Plant **20**: 245–250

Radziejwoski A, Vlieghe K, Lammens T, Berckmans B, Maes S, Jansen MAK, Knappe C, Albert A, Seidlitz HK, Bahnweg G, et al (2011) Atypical E2F activity coordinates PHR1 photolyase gene transcription with endoreduplication onset. EMBO J **30**: 355–363

Ramirez-Parra E, López-Matas MA, Fründt C, Gutierrez C (2004) Role of an atypical E2F transcription factor in the control of Arabidopsis cell growth and differentiation. Plant Cell **16**: 2350–2363

Ries G, Heller W, Puchta H, Sandermann H, Seidlitz HK, Hohn B (2000) Elevated UV-B radiation reduces genome stability in plants. Nature **406**: 98–101

Rizzini L, Favory J-J, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI, et al (2011) Perception of UV-B by the Arabidopsis UVR8 protein. Science **332**: 103–106

Rochette PJ, Lacoste S, Therrien J-P, Bastien N, Brash DE, Drouin R (2009) Influence of cytosine methylation on ultraviolet-induced cyclobutane pyrimidine dimer formation in genomic DNA. Mutat Res **665**: 7–13

Roberto C (2005). Low chromosome number angiosperms. Caryologia **58**: 403-409

Sambrook J, Russell DW (2001). Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sancar A (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. Chem Rev **103**: 2203–2237

Sancar A, Lindsey-Boltz LA, Unsal-Kaçmaz K, Linn S (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem **73**: 39–85

Saze H, Kakutani T (2011) Differentiation of epigenetic modifications between transposons and genes. Curr Opin Plant Biol **14**: 81–87

Schenke D, Böttcher C, Scheel D (2011) Crosstalk between abiotic ultraviolet-B stress and biotic (flg22) stress signalling in Arabidopsis prevents flavonol accumulation in favor of pathogen defence compound production. Plant Cell Environ **34**: 1849–1864

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet **37**: 501–506

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

Schuermann D, Molinier J, Fritsch O, Hohn B (2005) The dual nature of homologous recombination in plants. Trends Genet TIG **21**: 172–181

Shechter D, Costanzo V, Gautier J (2004) Regulation of DNA replication by ATR: signaling in response to DNA intermediates. DNA Repair **3**: 901–908

Shiloh Y (2006) The ATM-mediated DNA-damage response: taking shape. Trends Biochem Sci **31**: 402–410

Shiloh Y, Ziv Y (2013) The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. Nat Rev Mol Cell Biol **14**: 197–210

Shuck SC, Short EA, Turchi JJ (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. Cell Res **18**: 64–72

Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell **152**: 352–364

Sullivan JA, Deng XW (2003) From seed to seed: the role of photoreceptors in Arabidopsis development. Dev Biol **260**: 289–297

Sullivan JA, Shirasu K, Deng XW (2003) The diverse roles of ubiquitin and the 26S proteasome in the life of plants. Nat Rev Genet **4**: 948–958

Suter B, Livingstone-Zatchej M, Thoma F (1997) Chromatin structure modulates DNA repair by photolyase in vivo. EMBO J **16**: 2150–2160

Suter B, Thoma F (2002) DNA-repair by photolyase reveals dynamic properties of nucleosome positioning in vivo. J Mol Biol **319**: 395–406

Taiz L, Zeigler E (2002). Plant Physiology. Palgrave Macmillan. 3rd edition.

Thiel S, Döhring T, Köfferlein M, Kosak A, Martin P, Seidlitz HK (1996). A Phytotron for plant stress research: How far can artificial lighting compare to natural sunlight? J Plant Physiol **148**: 456-463

Thoma F (2005) Repair of UV lesions in nucleosomes--intrinsic properties and remodeling. DNA Repair **4**: 855–869

Tilbrook K, Arongaus AB, Binkert M, Heijde M, Yin R, Ulm R (2013) The UVR8 UV-B Photoreceptor: Perception, Signaling and Response. Arab Book Am Soc Plant Biol **11**: e0164

Torabinejad J, Caldwell MM (2000) Inheritance of UV-B tolerance in seven ecotypes of Arabidopsis thaliana L. Heynh. and their F1 hybrids. J Hered **91**: 228–233

Tör M, Brown D, Cooper A null, Woods-Tör A, Sjölander K, Jones J D G null, Holub E B null (2007) Arabidopsis downy mildew resistance gene RPP27 encodes a receptor-like protein similar to CLAVATA2 and tomato Cf-9. Plant Physiol **143**: 1079

Trontin C, Tisné S, Bach L, Loudet O (2011) What does Arabidopsis natural variation teach us (and does not teach us) about adaptation in plants? Curr Opin Plant Biol **14**: 225–231

Tuteja N, Ahmad P, Panda BB, Tuteja R (2009) Genotoxic stress in plants: shedding light on DNA damage, repair and DNA repair helicases. Mutat Res **681**: 134–149

Van den Heuvel S, Dyson NJ (2008) Conserved functions of the pRB and E2F families. Nat Rev Mol Cell Biol **9**: 713–724

Ulm R, Baumann A, Oravecz A, Máté Z, Adám E, Oakeley EJ, Schäfer E, Nagy F (2004) Genome-wide analysis of gene expression reveals function of the bZIP transcription factor HY5 in the UV-B response of Arabidopsis. Proc Natl Acad Sci U S A **101**: 1397–1402

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res **40**: e115

Vaisman A, Frank EG, Iwai S, Ohashi E, Ohmori H, Hanaoka F, Woodgate R (2003) Sequence context-dependent replication of DNA templates containing UV-induced lesions by human DNA polymerase iota. DNA Repair **2**: 991–1006

van der Burg WJ (2004). Ophioglossum reticulatum L. In: Grubben, G.J.H. & Denton, O.A. (Editors). PROTA 2: Vegetables/Légumes. [CD-Rom]. PROTA, Wageningen, Netherlands

Vandepoele K, Vlieghe K, Florquin K, Hennig L, Beemster GTS, Gruissem W, Van de Peer Y, Inzé D, De Veylder L (2005) Genome-wide identification of potential plant E2F target genes. Plant Physiol **139**: 316–328

Vlieghe K, Boudolf V, Beemster GTS, Maes S, Magyar Z, Atanassova A, de Almeida Engler J, De Groodt R, Inzé D, De Veylder L (2005) The DP-E2F-like gene DEL1 controls the endocycle in Arabidopsis thaliana. Curr Biol CB **15**: 59–63

Wang G, Ellendorff U, Kemp B, Mansfield JW, Forsyth A, Mitchell K, Bastas K, Liu C-M, Woods-Tör A, Zipfel C, et al (2008) A genome-wide functional investigation into the roles of receptor-like proteins in Arabidopsis. Plant Physiol **147**: 503–517

Waterworth WM, Altun C, Armstrong SJ, Roberts N, Dean PJ, Young K, Weil CF, Bray CM, West CE (2007) NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants. Plant J Cell Mol Biol **52**: 41–52

Weigel D (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiol **158**: 2–22

Weigel D, Glazebrook J (2002). Arabidopsis: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Weigel D, Mott R (2009) The 1001 genomes project for Arabidopsis thaliana. Genome Biol **10**: 107

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PloS One **2**: e718

Wu D, Hu Q, Yan Z, Chen W, Yan C, Huang X, Zhang J, Yang P, Deng H, Wang J, et al (2012) Structural basis of ultraviolet-B perception by UVR8. Nature **484**: 214–219

Xia J, Han L, Zhao Z (2012) Investigating the relationship of DNA methylation with mutation rate and allele frequency in the human genome. BMC Genomics **13 Suppl 8**: S7

Yang W (2011) Surviving the sun: repair and bypass of DNA UV lesions. Protein Sci Publ Protein Soc **20**: 1781–1789

Yao Y, Xuan Z, He Y, Lutts S, Korpelainen H, Li C (2007) Principal component analysis of intraspecific responses of tartary buckwheat to UV-B radiation under field conditions. Environ Exp Bot **61**: 237–245

Yu Y, Teng Y, Liu H, Reed SH, Waters R (2005) UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. Proc Natl Acad Sci U S A **102**: 8650–8655

Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H (2006a) Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protoc **1**: 641–646

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW-L, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al (2006b) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell **126**: 1189–1201

Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X (2014) Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PloS One **9**: e78644

Zhu YO, Siegal ML, Hall DW, Petrov DA (2014) Precise estimates of mutation rate and spectrum in yeast. Proc Natl Acad Sci U S A.

Acknowledgements

First of all I want to thank Dr. Ales Pecinka for the opportunity to work on this exciting topic in his group, his helpful discussions and his support during these years.

I also want to thank Prof. Dr. Maarten Koornneef for the opportunity to do a doctorate in his department and his interest in this work.

I am grateful to Prof. Dr. Ute Höcker at the University of Cologne for her evaluation of this doctoral thesis and for being part of my thesis committee. Furthermore, I would like to thank the other members of my thesis committee, Prof. Dr. Wolfgang Werr for holding the chair and Dr. Wim Soppe as assessor.

This thesis would not have come to paper without the help and support of many people. I want to thank all collaborators that participated in this work, for sharing their knowledge with me, for their patience and their guidance – Dr. Eva-Maria Willing, Dr. Korbinian Schneeberger, Dr. Andreas Albert and Dr. Barbro Winkler.

I would also like to thank Niels Müller, Ganga Jeena, Emma Miatton, Navratan Bagwan and Arunkumar Srinivasan for their help solving different problems that arose during this work.

Besides, I want to thank all members of the group Pecinka for their help and discussions in everyday lab life, especially Barbara Eilts and Regina Gentges for their technical support in countless experiments and in the greenhouse.

For the critical review of this thesis I want to thank Jeremy David Woodward.

Furthermore, I want to thank Marcel von Reth, Arunkumar Srinivasan and Navratan Bagwan for supporting me in this endeavour to do a PhD and becoming more than just colleagues, but friends. I wish you all a lifetime of happiness and strength to pursue the goals in your life.

I am deeply and forever grateful to my parents and my brother for their faith in me and their unconditional support through all these years.

Words cannot express my gratitude to M and S. I will be forever indebted to you – for your encouragement, your understanding and your kindness. You truly helped to make me the person I have become. In my heart there will always be a place for you...

Lebenslauf

Thomas Piofczyk
Widdersdorfer Straße 221
50825 Köln, Deutschland
Email: piofczyk@mpipz.mpg.de

■ Persönliche Daten

Geburtsdatum:	10 Januar 1985
Geburtsort:	Beuthen (Polen)
Staatsangehörigkeit:	deutsch

■ Akademische Ausbildung

Nov 2010 – Okt 2014	Promotion in der Abteilung für Pflanzenzüchtung und Genetik am Max-Planck-Institut für Pflanzenzüchtungs-forschung, Köln Titel der Doktorarbeit: „Molecular responses to UV-B in *Arabidopsis thaliana*"
Okt 2005 – Sept 2010	Studium der Biologie an der Ruhr-Universität Bochum am Lehrstuhl für Pflanzenphysiologie Abschluss und Note: Diplom (equiv MSc), 1.0 Titel der Diplomarbeit: „The Myrosinase homologous Glycosidases of *Arabidopsis thaliana* – Heterologous expression and activity analyses"

■ Praktische Lehrerfahrungen

Jan 2013	Organisation und Durchführung des Workshop „Toolbox for Molecular Biology" im Rahmen von UV4growth / COST-action FA0906
Jan 2010 – Apr 2010	studentische Hilfskraft am Lehrstuhl für Pflanzenphysiologie
Sept 2008 – Nov 2008	studentische Hilfskraft am Lehrstuhl für Pflanzenphysiologie
Mai 2007 – Sept 2007	studentische Hilfskraft am Lehrstuhl für Allgemeine und Molekulare Botanik

■ Auszeichnungen und Stipendien

Nov 2010 – Okt 2014	Promotionsstipendium der Max-Planck-Gesellschaft
Sept 2010	Studienabschlussstipendium der Ruth und Gert-Massenberg Stiftung

■ Schulische Ausbildung

1995 – 2004	Max-Planck-Gymnasium, Gelsenkirchen Abschluss: Allgemeine Hochschulreife

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit − einschließlich Tabellen, Karten und Abbildungen −, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie abgesehen von unten angegebenen Teilpublikationen noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maarten Koornneef betreut worden.

Köln, August 2014

   Thomas Piofczyk