Regina Kölzsch
Institut für die Sicherheit biotechnologischer Verfahren bei Pflanzen

Approaches to generate herbicide resistant *Taraxacum koksaghyz* by directed and undirected mutagenesis of the acetohydroxyacid synthase

Dissertationen aus dem Julius Kühn-Institut

Julius Kühn-Institut
Bundesforschungsinstitut für Kulturpflanzen

Kontakt | Contact

Regina Kölzsch

regina.koelzsch@gmail.com

Die Schriftenreihe „Dissertationen aus dem Julius Kühn-Institut" veröffentlicht Doktorarbeiten, die in enger Zusammenarbeit mit Universitäten an Instituten des Julius Kühn-Instituts entstanden sind.

The publication series „Dissertationen aus dem Julius Kühn-Institut" publishes doctoral dissertations originating from research doctorates and completed at the Julius Kühn-Institut (JKI) either in close collaboration with universities or as an outstanding independent work in the JKI research fields.

Der Vertrieb dieser Schriftenreihe erfolgt über den Buchhandel (Nachweis im Verzeichnis lieferbarer Bücher - VLB). Einige der Dissertationen erscheinen außerdem online open access und werden unter einer Creative Commons Namensnennung 4.0 Lizenz (CC-BY 4.0) zur Verfügung gestellt (https://creativecommons.org/licenses/by/4.0/deed.de). Die Schriftenreihe ist nachgewiesen in unserem Repositorium OpenAgrar: https://www.openagrar.de/receive/openagrar_mods_00005667.

The publication series is distributed through the book trade (listed in German Books in Print - VLB). Some of the dissertations are published online open access under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/deed.en). The publication series is documented within our repository OpenAgrar: https://www.openagrar.de/receive/openagrar_mods_00005667.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie: detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
(German National Library)
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

ISBN 978-3-95547-121-7
ISSN (elektronisch) 2510-0602
ISSN (Druck) 2510-0599
DOI 10.5073/20220913-131243

Herausgeber l Editor
Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Quedlinburg, Deutschland
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Quedlinburg, Germany

Regina Kölzsch

Approaches to generate herbicide resistant *Taraxacum koksaghyz* by directed and undirected mutagenesis of the acetohydroxyacid synthase

2022

Biologie

Approaches to generate herbicide resistant *Taraxacum koksaghyz* by directed and undirected mutagenesis of the acetohydroxyacid synthase

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Bologie
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
Regina Kölzsch
aus Rostock

-2022-

Dekan:	Prof. Dr. Jürgen Gadau
Erster Gutachter:	Prof. Dr. Dirk Prüfer
Zweiter Gutachter:	Dr. Frank Hartung
Tag der mündlichen Prüfung:	29.08.2022
Tag der Promotion:	14.10.2022

Abstract

Taraxacum koksaghyz is an upcoming alternative bioresource for natural rubber. A major obstacle for extensive cultivation is the lacking weed management. Here, two different strategies were conducted to obtain herbicide resistant *T. koksaghyz* by targeting the essential enzyme acetohydroxyacid synthase (AHAS) as a well-studied locus for resistance-conferring mutations.

For the diploid *T. koksaghyz* a single AHAS gene was identified with a range of highly homologous alleles in the considered selection of plants.

Undirected EMS mutagenesis led to one putative imidazolinone tolerant and two resistant individuals. The resistant plants featured the same, well-described mutation (C572T causing Ala191Val) which was successfully inherited to vital and fertile progeny. Transcriptional studies gave little evidence of the involvement of a metabolic tolerance.

Several gene regions of *TkoAHAS1* were targeted successfully by CRISPR/Cas9 *in vitro*. *In vivo* targeting of two sites led to detection of InDels at the 3' end. However, their influence on herbicide resistance remains to be conclusively assessed.

Table of contents

1 Introduction

1.1 *Taraxacum koksaghyz* is a natural rubber producing plant

The Russian Dandelion *Taraxacum koksaghyz* L. Rodin (*T. koksaghyz*) is a perennial plant of the family of Cichorioideae, clade of Asteraceae (*1, 2*). It is diploid (n = 16), self-incompatible and sexually reproducing (*3, 4*). Due to its natural habitat – the high valleys of the Tian Shan Mountains of Kazakhstan – the plant is accepting moderate climate conditions and marginal soils (*2, 5–7*). Flowering can or cannot be inducible by vernalisation (*8, 9*). Pollination is naturally occurring by thrips and beetles (Phalacridae) (*2*) and mature seeds are presented as seed heads (*8*). *T. koksaghyz* contains a milky sap, also known as latex (*7*).

Fig. 1-1 Appearance (A-D) (*10*) and origin (E) of *T. koksaghyz* L. Rodin. A) *T. koksaghyz* on the field, B) Latex stream after cutting the root, C) Inflorescence, D) Capitula with mature seeds, E) A modified map of Kazakhstan (based on (*11*)): the red circle indicates the region, where *T. koksaghyz* L. Rodin has been discovered (*2, 5–7*).

Approximately 10 % (20,000 species) of flowering plants produce latex (*12*), which is the cytoplasm of specifically differentiated cells, named laticifers (*13*). Laticifers can occur in any plant part and are divided into two forms. A non-articulated laticifer originates from one coenocytic cell and grows tube-like between the parenchyma cells. Branching can occur and

elongation is proceeding consistently during plant existence. Instead, articulated laticifers are forming a continuous system by a longitudinal arrangement of originally meristematic cells with perforated or degraded cell walls (*14–16*). Those structures enable the presence of latex and its ingredients throughout the plant: bioactive substances like alkaloids, terpenes, terpenoids or proteases depict a potent defence mechanism against herbivores and infections (*13*). In the latex of around 2,500 species *cis*-1,4-polyisoprene (natural rubber) particles are contained (*17, 18*). They are assumed to be responsible for sealing wounds, thereby inhibiting microorganisms to spread and also for generating a stickiness which is preventing the herbivore from further feed (*13*). Three species are known, which are able to produce long chain poly-isoprenes, that are suitable for rubber downstream processing industries: the Pará rubber tree (*Hevea brasiliensis*), Guayule (*Parthenium argentatum* Gray) and the Russian Dandelion (*T. koksaghyz* L. Rodin) *(18).*

Natural rubber is a highly valuable bio-product with a wide field of applications. Over 40,000 products are made of it, including e.g. tires, gloves and medical equipment (*18*). The world production was 13.804 million tons of natural rubber in 2019 (a plus of 11 % compared to 2016) and consumption was 13.661 million tonnes (*19, 20*). For 2021, prices were reported from US$ 1.68 to US$ 2.17 per kg (*21*), with the market generally being described as volatile (*22*). The tire industry is accounting for 70 % of the worlds' rubber consumption (*23*). Synthetic analogues have been developed but the natural rubber provides unique properties, which cannot be met by those and therefore they are only partially used. The elasticity at low temperatures, its resilience, abrasion and shock absorbance behaviours as well as the tolerance towards high temperatures make natural rubber inalienable (*18, 24, 25*).

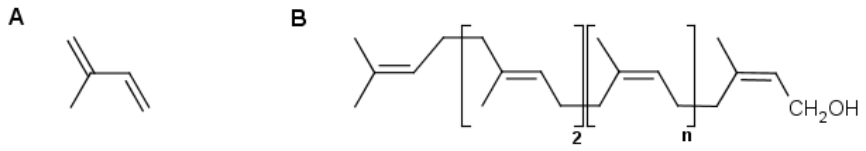

Fig. 1-2 Chemical structure of A) isoprene and B) *cis*-1,4-polyisoprene, with "n" depending on the source (Tab. 1-1) (*18*).

The three species mentioned above present different challenges in terms of cultivation, harvesting and processing due to the biological conditions (see below). They all produce long-chain poly-isoprenes with varying lengths, which is influencing the chemical and physical properties of the natural rubber. An overview of the characteristics and the rubber yields that have been reported so far is provided in Tab. 1-1.

Tab. 1-1 Comparison of the characteristics of natural rubber from three different species. References are listed; *cis*-1,4-polyisoprene chain length was calculated (chain length = (rubber MW)/(isoprene MW)). Isoprene molecular weight (MW) = 68.119 Da (*26*).

Species	**Maximum rubber MW [kDa]**	***cis*-1,4-polyiso-prene chain length**	**Rubber particle diameter [µm]**	**Rubber content in dry matter [%]**	**Rubber yield [kg/ha]**
H. brasiliensis	1,310 (*4*)	19,230	0.96-1.23 (*27*)	30-50 of latex; 2 of tree dry weight (*4*)	500-3000 (*4*)
***P. argentatum* Gray**	1,280 (*4*)	18,790	1.27-1.41 (*27*)	3-12 (*4*)	300-1000 (*4*)
***T. koksaghyz* L. Rodin**	2,180 (*4*)	32,000	0.2-1.00 (*28*)	0-20 of dry root (*7*)	30-60 (*29*); 389 (potentially) (*30*)

The well-established Pará rubber tree *H. brasiliensis* is currently the almost sole commercial source for natural rubber and the reference point in research for new alternatives (*18*). The rubber tree originates from Brazil and only tolerates the tropical climate conditions near the equator (*31*). *H. brasiliensis* consists of articulated laticifers in concentric rows in the primary and secondary phloem (*32*). The latex sap is harvested by tapping of the secondary laticifers in the bark (*18, 32*) (Fig. 1-3 A). First tapping is possible after initial seven years of growth and is profitable for about 25 years; afterwards the trees are deforested. Processing of the latex is established: it can be directly used as a concentrated mixture by removing water or dried after acid mediated coagulation of the rubber particles (*33*).

To date, *H. brasiliensis* is mainly cultivated in South East Asia (particularly Malaysia, Taiwan, Indonesia) and cropping in Africa and South America is accounting for a small proportion (*18, 34, 35*). Latter one is threatened by the fungus *Pseudocercospora ulei*, causing the South American Leaf Blight (*36*). The disease is spreading over South America, resulting in massive dieback of trees and whole plantations. Currently, this issue is not affecting other continents but the low genetic variability of *H. brasiliensis* worldwide could ease spreading of the disease (*18, 37*). Another, already global problem is the white root rot disease caused by *Rigidoporus microporus*, first described in 1904 (*38, 39*). Furthermore, cutting down rainforest areas for new *H. brasiliensis* plantations which are strictly grown as monocultures is a severe ecological problem (*40*). A medical issue is the growing percentage of latex allergies. Residual latex proteins in the rubber are recognized by immunoglobulin E antibodies of the human immune system, causing a type I allergic reaction which is triggered by ascending exposure (*18, 41–43*). However, to meet the increasing demands for natural rubber, establishment of new plantations is going on (*40*).

Fig. 1-3 Plantations of A) *H. brasiliensis* (*44*) and B) *P. argentatum* Gray (*45*).

Due to the above mentioned problems with cultivation of *H. brasiliensis*, studies on the Guayule shrub *P. argentatum* Gray as a rubber source have been conducted. During World War II, and in 1979/1982, during the oil crisis, the U.S.A. were not able to assess sufficient amounts of *H. brasiliensis* rubber and therefore focused research on Guayule (*45, 46*). The plant is producing natural rubber particles in the parenchyma cells of the bark (*18*) as it is lacking laticifers (*47, 48*). The rubber particles are found in specific epithelial cells which are laying around the resin canals (*47*). Therefore, tapping of Guayule is not possible. Perennial cultivation with frequent harvest of branches to enable regrowth of the plant is exerted (*42*) (Fig. 1-3 B). Harvesting is less laborious but processing of the natural rubber – dividing it from residual plant material – is more extensive than for *H. brasiliensis* rubber (*49*). The purified product contains less and other residual proteins compared to *H. brasiliensis* rubber, hence harbouring a lower allergenic potential; type I allergic reactions are not occurring (*42, 50, 51*). Thus, Guayule is already in commercial use, but only to a small extent (*52*). The shrub accepts only semi-arid climate with temperatures between 18 and 49 °C as it originates in the Chihuahuan desert of North Mexico and Texas (*35*), diminishing the potential cultivation area. Present cultivation and development occur in Arizona (*42, 52*).

T. koksaghyz was discovered in 1931 by L. E. Rodin (*7*). It has an articulated laticifer system like *H. brasiliensis* (*18*) and is able to store the main proportion of rubber in the roots (Fig. 1-1 B) (*14*). Therefore, cropping is annual or biennial and the whole plant is harvested (*8*). As with *P. argentatum* Gray, rubber needs to be separated from the harvested plant material by complex processing (*14, 46*). Besides natural rubber, *T. koksaghyz* contains noteworthy amounts of inulin useable in food industries or for production of bioethanol (*4*).

Discovery of the Russian Dandelion happened in the context of a series of expeditions in the Soviet Union. Extensive search for domestic plants, which were able to produce natural rubber was conducted to become independent from *H. brasiliensis* rubber imports. From the city of Ketmen (south-east Kazakhstan) the region of the Tian Shan mountains was explored (Fig. 1-1 E). Besides *T. koksaghyz* L. Rodin ("*Koksaghyz*"), other species have been identified, e.g. *Taraxacum hybernum* Stev. ("*Krim-Saghyz*") and *Scorzonera tausaghyz* Lipsch. et Bosse („*Tau-Saghyz*"). In 1933, the Soviet Union started extensive research on selected species including the above-mentioned ones and also Guayule. By this, 350 t raw natural rubber could be gained already in 1934. In 1941 it was concluded that *T. koksaghyz* was the best candidate for production of rubber and in the following years, *T. koksaghyz* was cropped in the regions of Kazakhstan, Ukraine and Moldova with preferential use of peat soils (*7*). Besides the Soviet Union, also e.g. Germany, Italy and the U.S.A. were investigating into alternative rubber sources (*7*), but after World War II most countries stopped or abandoned further research due

to reaccessibility of cheap *H. brasiliensis* rubber (*53, 54*). Half a century later, research has restarted: projects in the U.S.A. (e.g. PENRA (*55*)) and in Europe (e.g. EU-PEARLS (*56*), DRIVE4EU (*57*), RUBIN 2 (*58*)) concentrated/ are concentrating on establishing the Russian Dandelion as a crop. Several issues are under examination, e.g. the flowering habit, the unfavourable branched root morphology (*59*), selection and breeding of high-quality and high-yield rubber producing plants (*60*), revealing of the isoprene synthesis (*61, 62*), agronomy (cultivation, maintenance and harvest) (*8, 63–65*) and processing for use in downstream industries (*14*). A draft genome was published in 2017 by Lin *et al.* (*66*) that offers new insights and opportunities. *T. koksaghyz* being a promising source for natural rubber is also reflected by the interest of tire companies like Bridgestone (*67*) or Continental (*68*), that are also involved in research and were able to present prototype tires and mounts (*68*).

1.2 The acetohydroxyacid synthase

1.2.1 The physiological role of the enzyme

The acetohydroxyacid synthase (EC 2.2.1.6) (AHAS) is an essential enzyme in archaea, bacteria, fungi and plants whereas it is not occurring in animals/ mammals (*69*). In the biosynthesis of the branched-chain amino acids (BCAA) valine (Val), leucine (Leu) and isoleucine (Ile) it catalyses the crucial reaction of either two pyruvate (PYR) molecules or one PYR and one 2-ketobutyrate (KB) yielding 2-acetolactate (AL) or 2-aceto-2-hydroxybutyrate (AHB), respectively. The products are precursors for further three to six enzyme mediated reactions in the anabolism of stated amino acids (Fig. 1-4) and cannot be delivered by any other biosynthetic pathway (*70*).

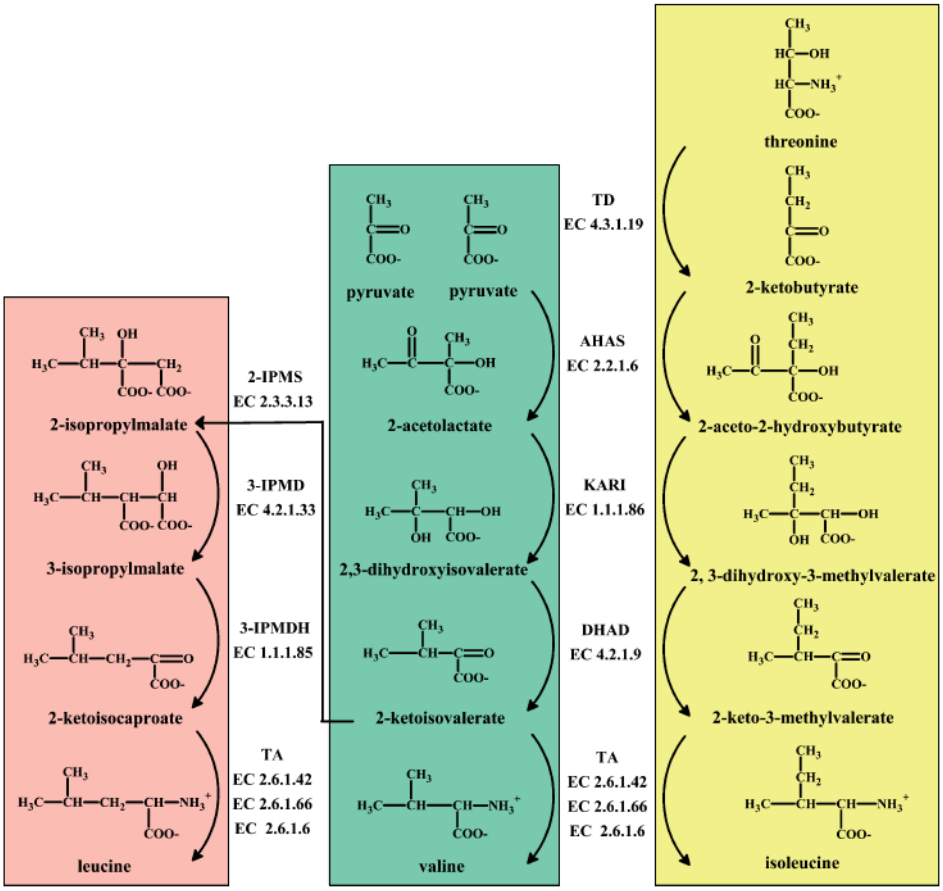

Fig. 1-4 Biosynthetic pathways of the anabolism of leucine, valine and isoleucine.
TD: threonine deaminase; KARI: ketol-acid reductoisomerase; DHAD: dihydroxyacid dehydratase; TA: transaminase; 2-PIMS: 2-isopropylmalate synthase; 3-IPMD: 3-isopropylmalate dehydratase; 3-IPMH: 3-isopropylmalate dehydrogenase. Adopted from McCourt *et al.* (*69*).

This feature makes AHAS interesting as a target for herbicide control in crops but also for the development of new antimicrobial drugs against pathogenic bacteria and fungi (*71, 72*). In literature the designation “acetolactate synthase” (ALS) instead of AHAS is frequently used, although Duggleby *et al*. proposed that it “shall be reserved for a different enzyme that produces AL only” (*70*). In contrast to AHAS, ALS is not able to yield AHB from KB and further does not consist of a regulatory subunit, nor is it flavin adenine dinucleotide dependent. Both enzymes are sometimes referred as to catabolic (ALS) and anabolic ALS (AHAS) (*71*).

The functional AHAS enzyme is an oligomer of catalytic subunits (CSU) and regulatory subunits (RSU), whereas the CSU implement the synthase reaction and the RSU convey feedback inhibition and promote synthase activity. A CSU is active without the RSU *in vitro* but then shows no sensitivity towards end-product inhibition (*73*). Main efforts on elucidating the structure have been put on AHAS from *Escherichia coli* (EcoAHAS), *Arabidopsis thaliana* (AthAHAS) and *Saccharomyces cerevisiae* (SceAHAS) (*69, 70*). To date (June 2022) 58 crystal structures for the CSU, RSU or the whole enzyme are available at the National Center for Biotechnology Information (NCBI) (*74*). In 2020, crystallisation of the AthAHAS holo-enzyme (PDB 6U9H) was successful, revealing a hetero-16-mer with a total weight of approx. 954 kDA (*74, 75*). Before that, the functional quaternary structure of AthAHAS was reported to consist of four CSU and four RSU leading to a 500 kDa complex (*73*).

Protein sequences and sizes of RSU across species differ widely: less than 25 % identity consists between bacterial RSU (*76*). RSU from bacteria show molecular weights from 9.5 to 17 kDa, whereas those of plants and fungi range between 34-50 kDa, due to a sequence repeat (*77*). The regulatory subunit is sensitive to feedback inhibition (*73*). Plant AHAS are inhibited by all BCAA (*78*) and additionally, synergistic effects for Leu and Val have been reported (*79*). This is due to the internal duplication by which two different binding sites for Leu and Val/ Ile are formed (*78*). Besides sensing feedback inhibition, the RSU is indispensable for activating the CSU; only 5 to 15 % of the full activity is reached by the CSU *in vitro* alone (*71*).
AHAS CSU of bacteria, fungi and plants range in molecular weight between 59 and 66 kDa (*70*). Because AHAS is nuclear-encoded and the synthesis of BCAA is mainly occurring in plastids of meristematic plant tissue (*80*–*82*) a signal peptide is afforded (*69*). The *N*-terminal 70-85 amino acids (AA) (*69*) have this targeting function with a typical preponderance of serine (Ser) (*83*). The mature plant protein often starts with the motif TFCS(K/R)(F/Y)AP and is then colinear with the CSU from bacteria (*79, 84*). Instead, AHAS from algae do not need a transit peptide due to the gene being encoded in the plastid genome (*79*). Organisms can have more than one AHAS gene and in Fig. 1-5 a protein alignment of CSU from *A. thaliana*, *E. coli* and *S. cerevisiae* AHAS reflects the similarities and depicts conserved domains which are described later.

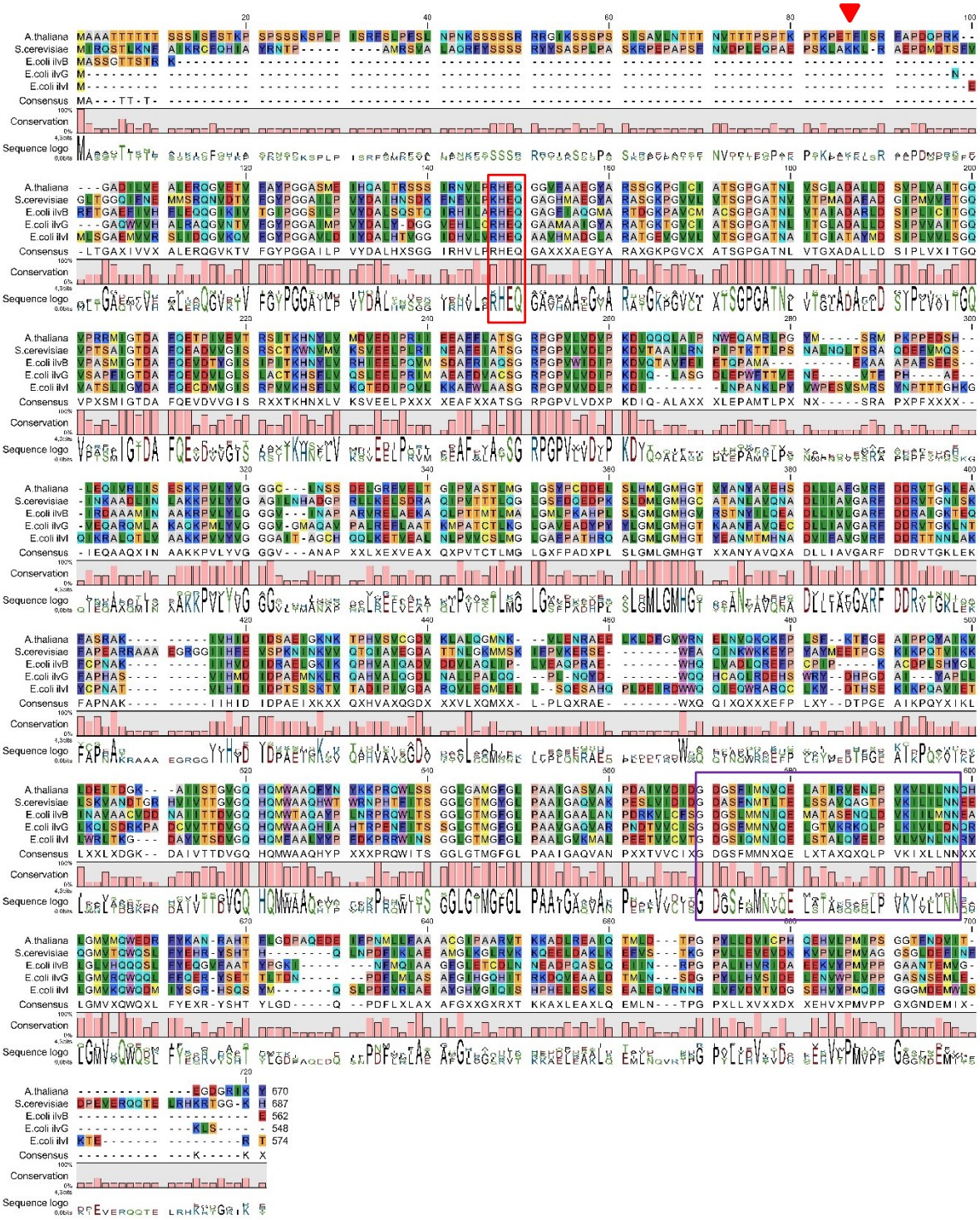

Fig. 1-5 Protein alignment of AHAS CSU from *A. thaliana* ALS (P17597.1), *S. cerevisiae* ilv2 (NP_013826.1) and *E. coli* ilvB/G/I (NP_418127.1, P0DP90.1, YP_025294.2) (*85*).
Red triangle: the start of the mature protein after the signal peptide in *A. thaliana* and *S. cerevisiae*. Red box: the RHEQ-motif with the catalytic Glu. Purple box: the ThDP-binding motif.

AthAHAS and SceAHAS CSU dimers and tetramers are well described in literature (*70, 73, 86*). Each CSU binds one molecule of thiamine diphosphate (ThDP), flavin adenine dinucleotide (FAD) and a divalent metal ion, e.g. Mg^{2+}. The active site is located at the interface between two catalytic monomers. This means a minimal setup to achieve activity is a dimer – as it was described for some bacteria and yeasts (*87*). Two separate binding sites for the first substrate (PYR) and the second substrate (PYR/ KB) have been elucidated – a simultaneous binding and a “waiting” of the second substrate in a pocket for conversion is assumed (*71*).

A catalytic monomer consists of the three domains α, β and γ. Each domain is folded as a parallel β-sheet, surrounded by several α-helices; whereas α and γ domain share a similar tertiary structure (*70*). The fold is common for proteins of the pyruvate oxidase (POX)-like family to which AHAS belongs (*71*). In AthAHAS, 23 C-terminal residual AA are forming a structured tail which is looping over the active site (*79*). 16 residues of the γ domain create a "mobile loop" that enables a variable tunnel for the two identical substrate channels composed by two monomers. The C-terminal tail and the mobile loop are forming a "capping region" whose conformation is changing by binding of an inhibitor (herbicide) (*88*). The crystal structure without inhibitor reveals an exposition of the substrate channel to the solvent and a barely organized conformation – this is not the case if an inhibitor is bound. Therefore, these are designated as "open" and "closed" forms, assuming the latter one represents the catalytically active one (*69*).
The crystal structure of the CSU tetramer of AthAHAS provides a suitable illustration of the above-described behaviour of two monomers forming a functional dimer (Fig. 1-6). Therefore, the quaternary structure is often termed as a dimer of dimers (*88*).

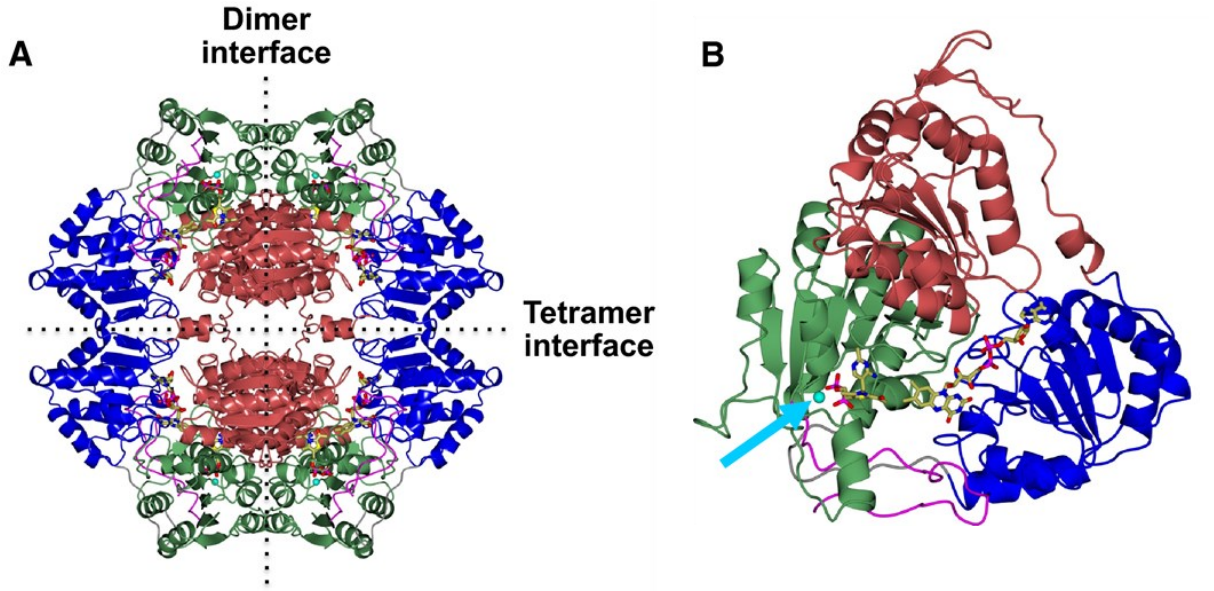

Fig. 1-6 The overall fold of AthAHAS. A) The tetrameric structure. B) A single subunit. The individual domains α, β and γ and the C-terminal tail are colored red, blue, green and pink, respectively. ThDP and FAD are shown as ball-and-stick models (gold: carbon; blue: nitrogen; red: oxygen; yellow: sulfur; magenta: phosphorus). Mg^{2+} is shown as cyan sphere; the cyan arrow pointing towards it. Dashed lines depict dimer and tetramer interfaces. Adopted and modified after Garcia *et al.* (*88*).

Different strategies are followed to decide between PYR and 2-KB as a second substrate. Bacterial AHAS isozymes vary in kinetics and substrate specificities to meet varying environmental conditions (*69*). Despite, plant and fungal AHAS generally favour 2-KB over PYR via different rate constants in the catalytic cycle (*89*) as an adaption towards the intracellular prevailing conditions of rather low 2-KB and rather high PYR concentrations (*90*). Additionally, seven residues (Phe206, Met351, Arg377, Gln207, Gly121, Trp574, Met570; AthAHAS) are described to contribute to substrate recognition and binding (*69*). On the contrary, the conserved Glu (Glu144 in AthAHAS) of the RHEQ motif (Fig. 1-5) is important for initiation of the catalytic cycle (*70*). Anyhow, catalysation of the synthase reaction still can only be achieved by interaction of all cofactors. The divalent metal ion is of great importance for anchoring ThDP in the enzyme. Several metal ions are accepted, but all the reported crystal structures have bound a Mg^{2+}, probably due to an over excess in the crystallization buffer (*70*).

ThDP is known to be required for several enzymes that catalyze decarboxylation reactions. The reaction cycle of ThDP is common and in the end a new C-C-bond between two substrates is build (*69*). The characteristic ThDP binding motif is retraceable in Fig. 1-5 (*91, 92*). FAD is held in position in the enzyme by six residues. These and the motives in which they are embedded are well conserved among various organisms (Tab. 6-1) (*90*). The role of FAD, a cofactor for redox reactions, has been diversely discussed and in 2004 Tittmann *et al.* ascertained that AHAS indeed has an oxygenase and a redox side activity (comparable to that of POX) (*89*). In 2016, Lonhienne *et al.* could reveal additional insights into the inhibition mechanism of AHAS via redox-reactions of FAD and herbicide binding (*93*). In fact, AHAS initially needs to be activated for catalysis by reduction of the FAD cofactor into the required $FADH_2$ via the POX side reactivity (*93, 94*). Under oxidative stress $FADH_2$ is oxidized, which leads to an inactivation of synthase activity. By the POX reaction, FAD can be re-reduced (*89*). Furthermore, an oxygenase side reaction is leading to formation of reactive oxygen species (ROS) (peracetate, singlet oxygen), which can oxidize $FADH_2$, too. This reaction proceeds only with 1 % of the AHAS synthase activity but is promoted by most herbicides (Fig. 6-1) (*72, 89, 93*).

The biosynthetical pathway of BCAA anabolism is regulated by various mechanisms, ensuring adaption of AHAS activity dependent on the energy level. A signal for this level is adenosine triphosphate (ATP). Accessibility of ATP implies a capacity for anabolic reactions and therefore ATP partially repeals the feedback inhibition of AHAS by BCAA to promote further BCAA synthesis. This is described for AHAS from fungi, algae and bacteria (*73, 79, 94*) (Fig. 1-7).
Quinones are important electron carrier in e.g. photosynthesis and respiration as well as antioxidants for free radicals. Lonhienne *et al.* could show that binding of oxidized ubiquinone (and quinone derivatives) leads to inactivation of AHAS (Fig. 1-7) (*94*). The ubiquinone pool in general reflects the redox and therefore the energy level of the cell. A low redox level in mitochondria/ cells is correlating to a low energy level in general. An accumulation of oxidized quinone is occurring if NADH (nicotinamide adenine dinucleotide) is lacking, resulting in drag down of ATP production. A stop of protein biosynthesis has to be mediated, too. By a redox reaction, the ubiquinone is reduced and the $FADH_2$ cofactor of AHAS is oxidized to FAD – the catalytically inactive form. Re-activation can be achieved via the POX side reactivity of AHAS. As mentioned, pyruvate is needed for this reaction. A low level of NADH however is the result of a low pyruvate level. Importantly, the binding affinity for quinones is low to enable binding of different quinones and ensure an active AHAS at standard quinone concentrations. Whereas the inhibition by quinones is even more effective if the PYR level is low, because the re-reduction of FAD is slower (*94*).

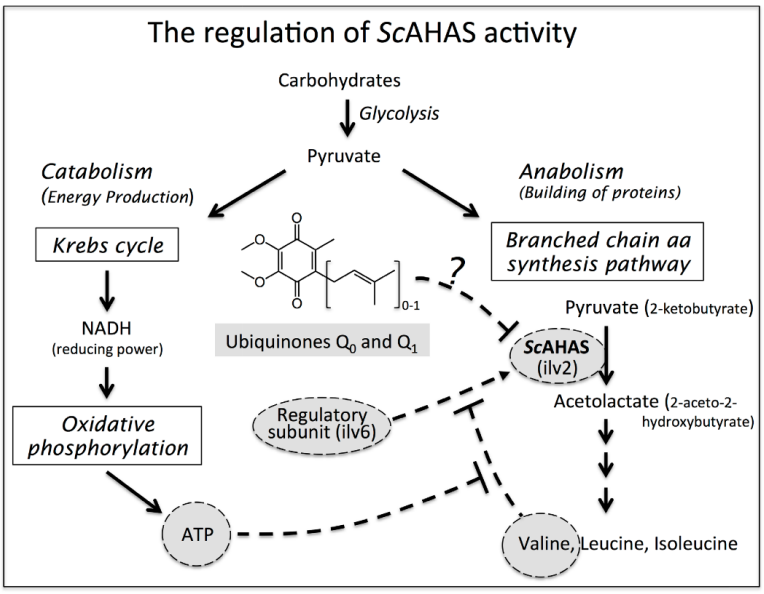

Fig. 1-7 Regulation of SceAHAS (here named *Sc*AHAS) activity in the metabolism. Adopted from Lonhienne *et al*. (*94*).

Since 1988 it is known that AHAS targeting herbicides are using the same binding pocket as quinones – but the mode of action is differing (*94, 95*). Whilst quinones directly undergo a redox reaction with $FADH_2$, herbicides rather stabilize an intermediate of the oxygenase side reaction leading to formation of ROS, which then oxidize FAD (*93, 94*). Quinones are altered by that mechanism of AHAS inhibition and are not accessible for further inhibition. In contrast, herbicides are not modified, enabling inhibition of further AHAS proteins (*94*).

1.2.2 Occurrence in plants

Besides mentioned investigations on structure, activity and mechanism of AthAHAS, several reports about AHAS from other plants exist. In those reports, mainly the herbicide resistance endowing mutations (chapter 1.2.3) and the occurrence of multiple genes are discussed. The AHAS CSU is encoded intron-free and the number of AHAS genes in planta is varying strongly (*79*). Whilst *A. thaliana* (*96*) and *Xanthium sp*. (*79*) possess one constitutively expressed gene, *Zea mays* (*97*), *Nicotiana tabacum* (*98*) and *Oryza sativa* (*99*) have two genes each. *Triticum spp.* (*100*) and *Helianthus annuus* (*82*) both have three AHAS genes. But even more complex gene families are known from *Brassica napus* with five genes (*84*) and *Gossypium hirsutum* with six genes (*101*).

The first crystal structure of AthAHAS CSU with a bound sulfonylurea was published in 2004 by Pang *et al*. (*102*) followed by further crystal structures of the protein bound with different herbicides (*69*). To date *A. thaliana* is the only plant from which AHAS crystal structures are published (*74*) – but numerous reports exist about activity, selectivity and inhibition of others (*74*).

In the late 1980ies and early 1990ies, investigations on AHAS have been challenging due to its instability and low abundance (0.1 % of the total protein) in the plant (*69, 79*). Only few reports on AHAS from cellular plant extracts or of purified protein were published (*103–108*). By using *E. coli* expression vectors, characterization of AHAS from e.g. *B. napus* (*84*), *Z. mays* (*109*), *Xanthium sp.* (*110*), *G. hirsutum* (*101*), *N. tabacum* (*86*), *N. pumbaginifolia* (*111*) and *A. thaliana* (*73, 86, 112*) was possible.

Plants can consist of one or more AHAS genes, but at least one gene is constitutively expressed while differing regulation of the others is possible (*79*). In such organisms, that originate from genome combination (e.g. *N. tabacum*, *B. napus*, *G. hirsutum*) even two

constitutively expressed genes exist (*79, 84, 101, 113*). Activity and transcript accumulation is highest in roots and leaves from young plantlets and in general in developing tissue (*82, 98, 114–116*). Thereby it is possible that a coordinated expression of several AHAS genes occurs – like in *N. tabacum* (*98*) – or even differential expression like in *Echinochloa phyllopogon* (*117*). But still tissue specific expression patterns were found in *G. hirsutum* and *B. napus* (*101, 115*). In addition, *H. annuus*, *Z. mays*, *Sorghum bicolor* and *Hordeum vulgare* have higher expression levels in leaves than in roots (*118–121*) whereas for *B. napus* and *Chicorium intybus* similar levels were detected (*114, 121*).
In Fig. 1-8 a phylogenetic tree of several AHAS (CSU) reveals the distant cognation between plant and non-plant protein and reflects the high resemblance within the planta branch. The grouping into monocotyledons, dicotyledons and asterids is well comprehensible.

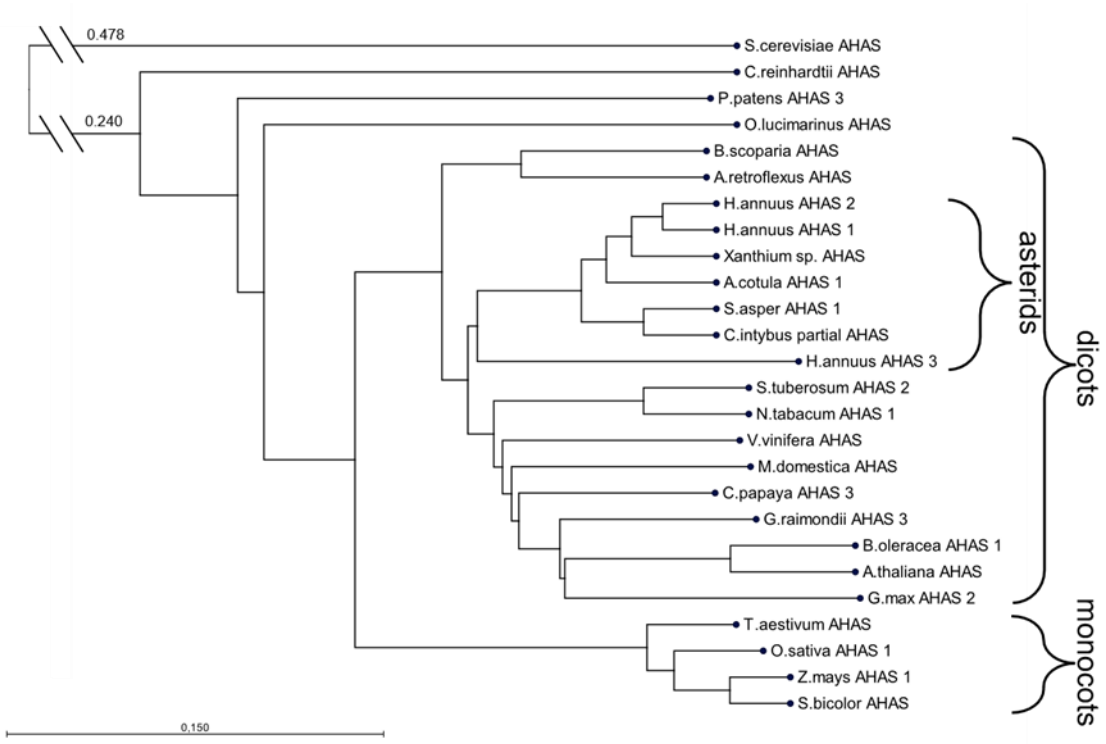

Fig. 1-8 Phylogentic tree for AHAS of exemplary organisms from monocots, dicots, as well as chlorophytes, bryopsidas and fungi (references are listed in Tab. 6-2).

A gene or protein sequence for AHAS of *T. koksaghyz* has not been published, so far. The sole draft genome available, was presented by Lin *et al.* in 2017 (*66*). It has a size of 1.29 Gb (*66*), whereas Kirschner *et al.* determined a genome size of 1.45 pg in 2013 (*53*) which is equivalent to a size of ~1.42 Gb (*122*). Annotations are not provided for the draft genome (*66*). Additionally to this draft genome, thirteen data collections from *T. koksaghyz* or Asteraceae in general are listed at NCBI (Tab. 6-3); either raw sequence reads, transcriptome data or assemblies (*74*). Only eight of these ten sets are publicly available, whereas annotations, at least for AHAS, do not exist.

1.2.3 The enzyme as a target for herbicides

Control of plant growth by herbicides is possible by different modes of action: photosynthesis related processes (e.g. photosystem I/ II, carotenoid biosynthesis), growth/ cell division or the cell metabolism can be targeted. Several classes of herbicides for specific applications have

been developed and are categorized by the Herbicide Resistance Action Committee (HRAC). One of the major groups is the HRAC class B of AHAS (ALS) inhibitors (*123*).
In the late 1970s, chemists of DuPont developed the first sulfonylurea Chlorsulfuron that was commercially available in 1982. The imidazolinone Imazaquin was marketed 1986 by American Cyanamid. Two years after the launch of Chlorsulfuron the site of action – the enzyme AHAS – was identified (*69*). The AHAS inhibitors revolutionized the market: despite other herbicides used at that time, these compounds could be applied at rather low rates (g instead of kg/ha) but with a wide crop selectivity in combination with zero toxicity for mammals (*69, 124*).
Five classes of herbicides targeting AHAS are nowadays marketed: imidazolinones (IMI), sulfonylureas (SU), triazolopyrimidines (TP), pyrimidinyl-thiobenzoates (PYB) and sulfonyl-aminocarbonyl-triazolinones (SCT) (*124, 125*). The first two classes represent the mostly applied inhibitors for broad spectrum weed control (*69*); to date six different IMI and 34 SU are commercially available, worldwide (*123*).

IMI consist of a carboxylated pyridine/benzene/quinoline ring with substituents (*69*). Furthermore, it is crucial for activity that the acid group is *ortho*-positioned towards the imidazolinone ring (*126*) (Fig. 1-9). The lipophilic property of IMI and the uncharged state due to the low pH in the plant apoplast enable passive transport through cell membranes. The higher pH inside a cell results in the charged IMI form and by this, the herbicide is trapped inside the cell (*127*). Varying IMI absorption properties and behaviour of translocation to meristematic tissues, where the activity of AHAS is the highest, are the reasons for different herbicidal activities. Also, trapping in the phloem is related to this (*128–130*): transport through the plant is enabled by the phloem, but the degree of lipophilicity determines the ability to diffuse away from it. This means highly active IMI are those, which are more lipophilic and therefore distributed fast in the plant. Less lipophilic IMI are trapped in the phloem, slowly diffusing and by this having a higher post-emergent activity (*126*).

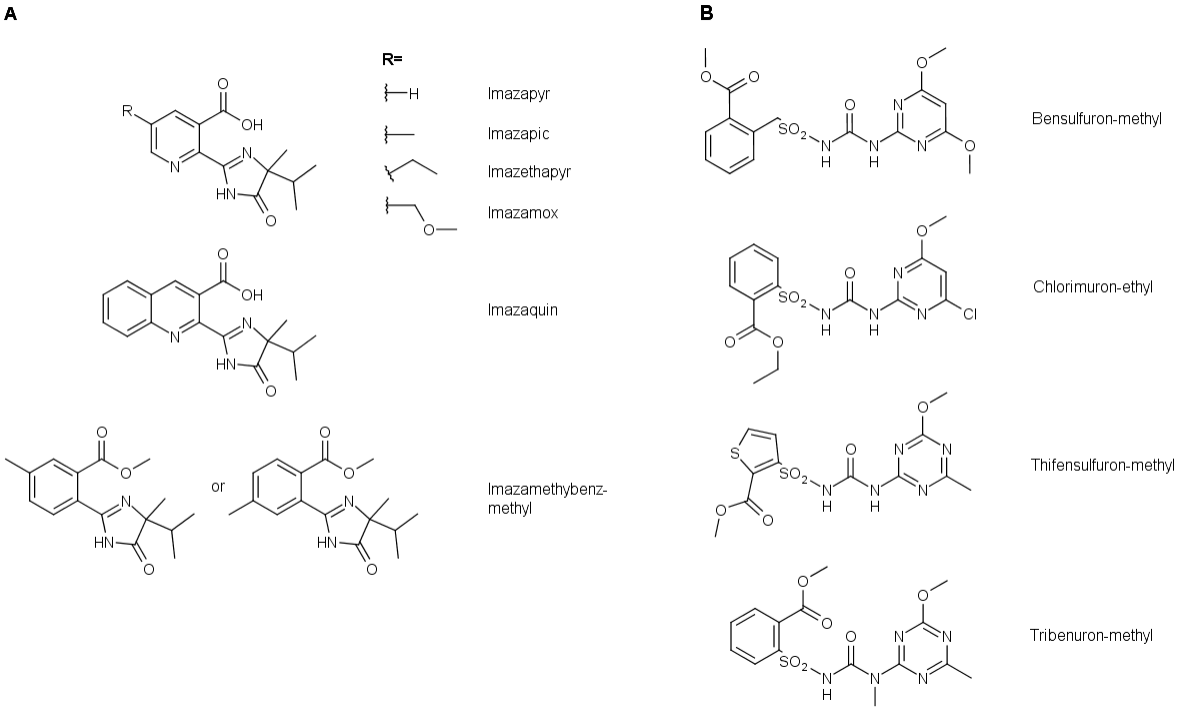

Fig. 1-9 Chemical structures and names of (A) the six commercially available IMI and (B) four exemplary SU (*126, 131*); drawn with MarvinSketch (*132*).

Structures of SU are varying more than those of IMI but always contain a sulfonylurea bridge and often two methoxy/methyl substituents at the heterocycle, conferring high herbicidal activity (Fig. 1-9) (*69*). The water solubility of SU increases with increasing pH due to the characteristic of being weak acids. At neutral or basic pH the anionic form is present, which is less susceptible to hydrolysis (*133*). The herbicide class is easily distributed in the plant via phloem and xylem (*134*).

Application of IMI or SU provoke a similar plant response: BCAA synthesis is inhibited, therefore protein biosynthesis is impaired, leading to cessation of cell division and finally to cell death. These reactions are phenotypically traceable: shortly after herbicide application, growth arrest occurs and chlorosis as well as anthocyanin production, resulting in a reddish colour of the leaves, appear. Plant death is occurring between seven and 20 days after application (*126, 131*). Although IMI and SU do not share structure similarities, they attempt to bind at overlapping sites in AHAS and cross-tolerance for these two pesticides exists in some cases. Generally, AHAS inhibiting herbicides do not compete with the substrates for binding, as already mentioned in chapter 1.2.1. Instead, they bind next to the active site, in a conserved quinone-binding pocket, blocking the substrate access channel and causing inactivation of AHAS by oxidation of the FAD cofactor and modification of the ThDP. This inhibition is not altering the herbicide, thus enabling deactivation of further AHAS, which can recover by the POX side reaction only with slow rate (*72*). Several crystal structures of AthAHAS (CSU tetramer) free as well as bound with different IMI and SU revealed 18 positions important for herbicide binding (Tab. 1-2) (*70*).

Tab. 1-2 AHAS AA residues that contribute to binding of SU and IMI ("x" = direct contact; "(x)" = no direct contact but contribution to binding; "-" = no contact). Based on Duggleby *et al.* (*70*) and McCourt *et al.* (*69*).

AA residue AthAHAS numbering	**Binding of SU**	**Binding of IMI**
Gly121	x	-
Ala122	x	x
Val196	x	x
Pro197	x	-
Arg199	(x)	x
Met200	x	x
Ala205	x	-
Phe206	x	x
Gln207	x	x
Lys256	x	x
Met351	x	x
Asp376	x	x
Arg377	x	x
Met570	x	-
Val571	x	-
Trp574	x	x
Ser653	x	x
Gly654	-	x

Generally, SU are better AHAS inhibitors than IMI – nano molarities of SU are needed instead of micro molarities of IMI to inhibit the enzyme (*135*). This is assumed to be due to the better

and tighter fit of SU in the protein (*70*). In 1992 first IMI resistant maize was marketed – nowadays known as Clearfield®corn. But also for oilseed rape, rice and wheat and sunflower resistant lines are commercially available (Tab. 6-4) (*136*). The above listed residues, which contribute to the herbicide binding (Tab. 1-2), are potential candidates for mutations to impede binding and inhibition of AHAS. Mutations of six AA positions are most common in commercial crops: Ala122, Pro197, Ala205, Asp376, Trp574 and Ser653 (AthAHAS numbering). At least one AA substitution can confer herbicide resistance. Mutations of Ala122 and Ser653 confer resistance towards IMI and changing Pro197 provides resistance to SU and can result in IMI resistance (*70, 137*). Resistance to both IMI and SU is covered by mutations at Trp574, whereas those at position Ala205 confer resistance against all AHAS inhibitors (*110, 137, 138*). Pro197 and Ala205 are not contributing to IMI binding directly, still substitution of e.g. Pro197 by a bulky AA impedes IMI to enter the access channel (*70, 139, 140*).
Since the introduction of AHAS inhibiting herbicides in the 1980ies, field-evolved herbicide resistant or tolerant weeds appeared. Rather specific point mutations in the AHAS gene than metabolic tolerance effects are the reason for weed persistence (*69*). To date, biotypes of 66 weed species with evolved AHAS herbicide resistance have been listed by the International Survey of Herbicide Resistant Weeds (*141*). The Survey also provides information about reported AHAS AA substitutions and whether those are conferring resistance towards AHAS inhibitors. Such substitutions have been found at eight distinct positions concerning IMI and SU resistance: Ala122, Pro197, Ala205, Asp376, Arg377, Trp574, Ser653 and Gly654 (AthAHAS numbering). These positions have already been mentioned in conjunction with herbicide binding in Tab. 1-2. Up to eleven different AA substitutions for one position are described and resistance towards both herbicides were observed, partially (Tab. 6-5). Most often described weeds with resistance towards any AHAS inhibitor are *Apera spica-venti*, *Descurainia Sophia*, *Kochia scoparia* and *Schoenoplectus juncoides* (*141*).

The HRAC developed criteria for definition and confirmation of herbicide resistant weeds based on the “resistance” definition by the Weed Science Society of America (WSSA) (*142*). That society already defined in 1998 “resistance” as “the inherited ability of a plant to survive and reproduce following exposure to a dose of herbicide normally lethal to the wild type. In a plant, resistance may be naturally occurring or induced by such techniques as genetic engineering or selection of variants produced by tissue culture or mutagenesis” (*143*). Despite this, tolerance is “the inherent ability of a species to survive and reproduce after herbicide treatment. This implies that there was no selection or genetic manipulation to make the plant tolerant; it is naturally tolerant” (*143*). The HRAC further stated that considering a plant (or weed) as resistant should be based upon the scientific and the agricultural definition, both. Scientifically, resistance is a “genetically inherited statistical difference in herbicide response” of two plants of the same species (*144*). The survival of recommended application rates of an herbicide under field conditions is taking the agricultural definition into account. Only if these two prerequisites are complied, a plant can be considered as resistant (*144*).
Baucom (*142*) gave a résumé about the varying use of the terms “herbicide resistance” and “herbicide tolerance” and summarized that tolerance shall be understood as the “fitness response after herbicide application”. This would lead to survival of the plant but in contrary to ”resistance”, damage of the plant is not prevented; at least, the tolerant plant is able to compensate for these damages (*142, 145*).
By this description, “resistance” of a plant is only to be understood due to genetic modification of the herbicide target protein, whilst this trait is independent from the plant having contact with the herbicide. In contrast, “tolerance” is acquired due to (several) exposures to an herbicide.

The plant has developed mechanisms to cope with a chemical by either increased metabolism, sequestration, compartmentalization or over-expression of the target-protein (*142*). Further, strategies of reduced uptake, translocation or reduced intracellular activation of the herbicide are known (*146–148*).

The metabolism of herbicides is often also referred to as non-target-site resistance but shall herein be termed as tolerance (as stated above). Despite the fact that a resistant plant is able to survive without damages, it is still an obstacle for the organism to cope with the organic compound itself. The herbicide's characteristics of moderate sizes (< 50 kDa) and a good lipophilicity enable symplastic passaging and transport via the phloem and xylem (*149*). Mostly, herbicides diffuse into the symplast; uptake by carriers is unlikely, but was detected for e.g. glyphosate (*150, 151*). The herbicidal compound can e.g. influence the pH, interfere with the electron transport of photosynthesis or with the biosynthesis of several metabolites like cellulose, porphyrins, fatty acids or AA (*152, 153*). Generally, plants consist of mechanisms and pathways to detoxify and metabolize xenobiotics. Plants, which have enhanced this strategy to stand herbicidal treatment are considered as (metabolic) tolerant (*154*). Metabolism is achieved by changing the lipophilic to an hydrophilic character in order to reduce the ability for membrane interaction and distribution in the cell (*155*). Three phases of detoxification of xenobiotics, in general, are described: transformation (phase I), conjugation (phase II) and compartmentation (phase III) (*156*).
In phase I, oxidation, reduction or hydrolysis occurs to increase the solubility of the compound. In addition, a reactive group, e.g. AA, hydroxyl or carboxylic acid group, can be introduced for activating the compound for phase II of detoxification. The enzyme class of cytochrome P450 monooxygenases is playing a major role by catalyzing a broad range of reactions like *N*- and *O*-dealkylations, *N*-demethylations, aromatic hydroxylations, and beta-oxidations (*149*). Nitroreductases, deaminases, peroxidases and esterases are further enzymes important for phase I (*147, 148*). It is possible that after this modification the xenobiotic is even more toxic. Phase II ensures formation of less or non-toxic compounds by conjugation reactions with glucose, malonate or glutathione (*155*). Thereby, the water solubility is increased and/ or the mobility in the plant as well as the biological activity is reduced. Glucosyl-, malonyl- and glutathione-transferases to form glucosides and glucose esters are of importance for this phase (*155–157*). In phase III, the conjugates are sequestered to the extracellular space or the vacuole, mostly by ABC transporters. In the compartments, further degradation of the conjugates occurs, also sometimes designated as phase IV (*149, 158*).
These described pathways take place in plants whether or not target-site resistance against the xenobiotic exists, to reduce the above-mentioned influences of xenobiotics on pH etc. Though, if a target-site resistance is lacking, the plant is either able to efficiently detoxify and deposit the xenobiotic (tolerance) or will die by trying so (susceptibility) (*159*).

Besides the mentioned metabolism, Shaner and Singh described the physiological plant response towards AHAS inhibitors (*160*). First, cessation of mitosis occurs (between G2 and M phase) and within hours after application, the thymidine incorporation into DNA is inhibited. These reactions are reversible by supplementation with BCAA. Second, Shaner and Singh assumed an inhibited photosynthate transport; no source to sink transport is occurring anymore, thereby neutral sugars accumulate in affected leaves (*160*).
Royuela *et al.* did not find evidence that AHAS inhibition causes sink/source deficiency but rather assumed that phytotoxic effects result from deregulation of AA biosynthesis or accumulation of AA (*161*). Zhou *et al.* (*162*) described phytotoxic effects by accumulation of α-

ketobutyrate and its transaminated derivative 2-aminobutyrate after herbicidal inhibition of AHAS. Orcaray *et al.* found AHAS inhibitors to promote toxic quinate accumulation (*163*).
Metabolism of IMI and SU depends on the substituents: 5'-substituted IMI (e.g. Imazamox, Imazethapyr, Imazamethabenz-methyl) are detoxified by hydroxylation and subsequent glucose-conjugation to the hydroxyl group (*164*). In contrast, for Imazapyr and Imazaquin internal condensation of the carboxylic acid from the aromatic ring and the nitrogen in the imidazolinone ring with subsequent cleavage of the imidazolinone ring is described (*126, 164*). The metabolic mechanisms and metabolites of exemplary IMI and SU are listed in Tab. 1-3.

Tab. 1-3 Mechanisms and metabolites of the metabolism of IMI (Imazamox, Imazethapyr, Imazamethabenz methyl) and SU (Chlorsulfuron, Bensulfuron). Adopted and modified after Shaner and Singh (*160*) and for Imazamox according to Rojano-Delgado et al. (*165*), respectively.

AHAS inhibitor	**Primary metabolic inactivation route**	**Metabolites**
Imazamox	Hydroxylation, glucose conjugation	5'-hydroxymethyl metabolite, glucose-conjugate metabolite
Imazethapyr	Hydroxylation, glucose conjugation	5'-hydroxyethyl metabolite, glucose-conjugate metabolite
Imazamethabenz-methyl	Hydroxylation, glucose conjugation	Imazamethabenz, glucose -conjugate
Chlorsulfuron	Hydroxylation, glucose conjugation	Glucose-conjugate
Bensulfuron	O-dealkylation	*O*-dealkylated metabolite

Due to the specific metabolic pathways of weeds for detoxification of herbicides, IMI and SU herbicides control different weeds. Imazamox, for instance, controls grasses better than Imazethapyr, whilst Imazapyr has the broadest range of weed control (*126, 166*).

1.3 Introduction of mutations into the genome of plants

10,000 to 12,000 years ago, the history of plant breeding started by first selections of natural varieties for beneficial traits like increased seed count. Selection and propagation resulted in about 2,500 domesticated plant species (*167–171*).

Induced mutagenesis has led to various optimized crops and new varieties in the 20th century (*172*). Since the 1990s insertional mutagenesis on plants was conducted for basic research (*173*) and transgenic approaches led to genetically modified organisms (GMO) which could not have been gained through classical breeding (*174, 175*), e.g. “Bt maize” (*176*), “Golden Rice” (*169*). Since the end of the 20th century, site-directed mutagenesis by using site-directed nucleases (SDN) enables targeted, specific editing of genomic sequences, broadening the spectrum of possible new varieties (*175*).

1.3.1 Induced and insertional mutagenesis

Undirected or untargeted mutations in plants can be induced by various chemical and physical mutagens. Over 70 % of the existing varieties from induced mutagenesis have been gained by latter ones (*177, 178*). Gamma and X-rays rays are the predominantly used physical tools for induction of mutations by generating DNA double strand breaks (*172, 178*). Instead, high energy ionizing particles or ion beams are known to cause damages ranging from large DNA deletions to chromosome changes and aberrations (*172, 178, 179*).

By use of chemical agents, rather single point mutations are occurring, which can lead to loss-of-function but also to gain-of-function phenotypes (*172*). The mechanisms by which those mutations are induced are varying widely. Base analogues (e.g. 5-bromouracil) are incorporated into the DNA instead of the correct bases and are provoking transitions during replication. Transitions are exchanges of a purine by another purine (adenine (A)/guanine (G)) or a pyrimidine by another pyrimidine (cytosine (C)/thymine (T)). This is also inducible by nitrous acid via deamination of cytosine. Intercalators (e.g. acridines orange) bind between the DNA bases, thereby stretching the DNA strand and causing insertion of additional bases during replication, finally leading to frameshifts (*172*). Mutagens like psoralen can induce intra-and inter-strand crosslinks, which are preventing DNA replication and transcription. The mostly used chemical mutagens are alkylating agents, e.g. ethyl methanesulfonate (EMS), 2-chlorethyl-dimethyl, *N*-ethyl-*N*-nitrosourea or ethylene oxide (*178*). More than 80 % of chemically mutagenized plants were generated by using those (*172, 180*). Alkylating chemicals react with DNA bases by methylating or ethylating them, leading to complete degradation of the base and frame shift or to mispairing and mutations during DNA replication (*178*). EMS in particular can ethylate the DNA bases guanine and thymine resulting in mismatches leading to the introduction of wrong bases and thereby introducing point mutations (*181*). With more than 99 % (in *A. thaliana*) transition of G/C to A/T pairing is occurring (Fig. 1-10) (*182*).

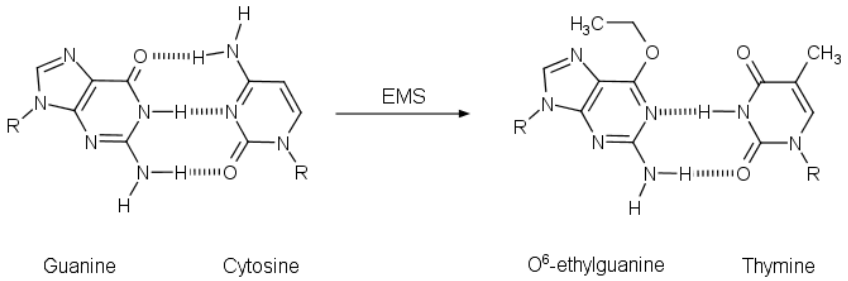

Fig. 1-10 Ethylation of guanine to O^6-ethylguanine causes pairing with thymine instead of cytosine, thereby provoking a transition from G/C to A/T (*183*).

The mutation rate is depending on the organism as well as the conditions during mutation (see next paragraph). Anyhow, EMS mutation rates of 1/170 kb (*138*) and 1/300 kb (*182*) for *A. thaliana* and 1/210 kb for *Drosophila melanogaster* (*184*) have been described, revealing values of the same range. Greene *et al.* further described that EMS mutagenesis in *A. thaliana* leads to ~45 % silent and ~55 % non-silent mutations, whereas ~50 % missense mutations and ~5 % truncation occur due to the introduction of nonsense codons or splice junction losses (*182*). The International Atomic Energy Agency (IAEA) provides a “Mutant Variety Database” where plant mutant varieties resulting from induced mutagenesis are listed. Currently, 388 varieties are registered, that have been gained by chemical mutagenesis. Of those, more than 100 have been developed by using EMS, e.g. from rice, wheat, soybean, faba bean, lettuce, barley, maize or rapeseed (*180*).
When using physical or chemical mutagens, extensive preliminary tests have to be conducted to reach an optimal dose-mutation frequency-rate. Many factors have to be taken into account, e.g. the target tissue, treatment duration and temperature, pH or post-treatment handling (*185*). Above mentioned techniques are not exclusively used in plants, but also in model organisms like *E. coli* (*186–188*), *S. cerevisiae* (*186, 189, 190*) and *Caenorhabditis elegans* (*191*).
Insertional mutagenesis is a technique to identify genes, their functional role and possibly their position in the genome (*192*). For this purpose, transposons – DNA sequences that can change their positions in the genome and influencing the genome function (*193–195*) – or transfer-DNA (T-DNA) – a DNA sequence inserted via Agrobacteria (chapter 1.4) – are used. By introduction of such external DNA sequences, which insert randomly into the genome of the targeted organism genes are disrupted. The effect for the phenotype can be directly associated to the genotype, as the gene is tagged by the insertion. This kind of tagging is also useful to identify the gene position in the genome if a reference genome is available (*196*).

1.3.2 Site-directed mutagenesis

Methods of site-directed mutagenesis have been developed during the last two decades and belong to the group of new plant breeding techniques (*174*). The use of SDN enables targeted genome modification, opening up new possibilities for precise breeding (*175*). SDN are introducing DNA double-strand breaks (DSBs), which have to be repaired by the targeted organism (*197*). They are categorized in three types (SDN-1/ -2/ -3) depending on whether rather point mutations (SDN1), insertions/ editing of few or more base-pairs (SDN2) or insertion of longer strands (SDN3) takes place (*198*). By non-homologous end joining (NHEJ), which is not always precise, insertions and deletions (InDels) occur, often leading to frame shifts and premature stop codons that can result in loss of function for the translated protein. Alternatively, homologous recombination (HR) takes place: a repair template (sister chromatid, homologous chromosome, designed homologous DNA) is directing the exact and correct repair (*199*). NHEJ is dominating over HR (*200, 201*), but these two repair mechanisms inherit the potential for new mutations to disrupt, add, adapt or replace genes (Fig. 1-11) (*197*). Thereby site specificity due to usage of SDN for introduction of the initial DNA DSB enables less off-target-effects than by using classical breeding technologies (*175, 202, 203*).

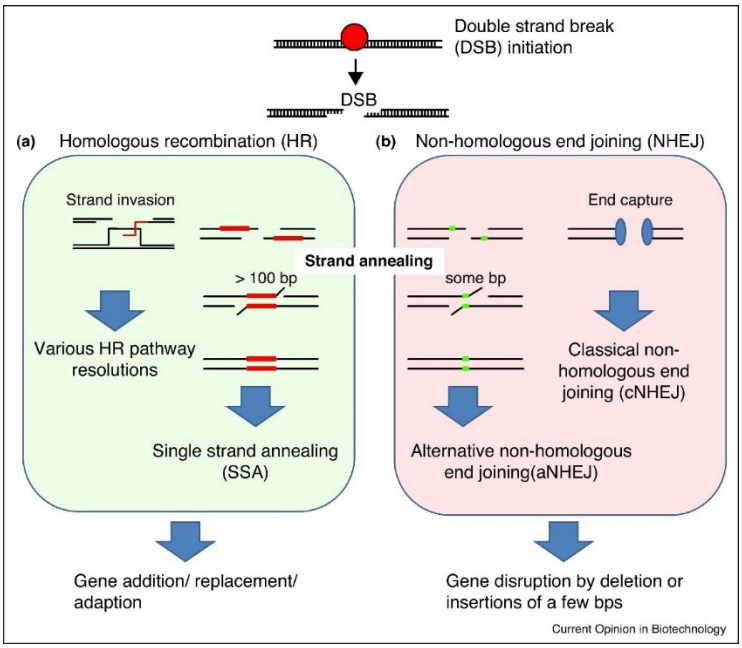

Fig. 1-11 The HR and NHEJ repair mechanisms after introduction of a DNA DSB. a) HR can occur either after strand invasion leading to double holliday junctions (several resolution pathways) or after strand annealing of homologous tandem sequences in the broken DNA-fragment or by introduction of a foreign DNA template. HR leads to addition, replacement or adaption of genes. b) Via NHEJ the DNA ends are simply fused together by classical/canonical NHEJ or microhomologies of the broken DNA strands lead to alternative NHEJ. Both mechanisms sometimes evoke InDels, causing gene disruption. Adopted from Sprink *et al.* (*197*).

In 1991, Pavletich and Pabo first described zinc finger nucleases (ZFN) and how those SDN are targeting DNA to induce DSBs (*204*). ZFN are fusion proteins of a DSB inducing nuclease (nowadays used: FokI (*197*)) and a DNA binding domain. Latter one consists of alpha-helices, known as zinc finger domains, with each recognizing three DNA base pairs (*205*). This recognition mechanism is complicating the design of new ZFN and involving extensive screening (*197*). Furthermore, meganucleases like I-SceI were described to introduce DSB by naturally preferences, too (*206*). Yet, these preferences are restricting possible targets (*197*). The transcription activator-like effector nucleases (TALEN) enabled new possibilities in the field of genome editing in 2010 (*207*). TALEN are fusions proteins like ZFN, too, but the DNA binding domain consists of several short circa 34 AA long sequences, which bind only one nucleotide instead of three as in ZFN (*205, 208*). Therefore, TALEN are somewhat easier to design towards the target (*197*). Besides the SDN, oligonucleotide directed mutagenesis (ODM) aroused a huge scientific interest between 1999 and 2004 (*174*). Chemically synthesized RNA, DNA or chimeric oligonucleotides are used, which are homologous to the targeted DNA sequence – except for a few nucleotides which are intended to be changed (*174, 175*). Introduction of point, insertion and even short deletion mutations is possible with ODM (*167, 174, 209*).

The bottleneck of design and applicability of ZFN and TALEN has been overcome by a new type of SDN. Clustered regularly interspaced short palindromic repeats (CRISPRs) fused to a CRISPR associated (Cas) nuclease represent an easier and faster system (*197*). The origin of this system lies in the natural adaptive immune system of microbes against phages. After phage infection, the foreign DNA is cleaved by specific Cas protein(s) and is integrated as spacers, separated by repeats, into specific CRISPR array(s) of the host. If the cell is attacked again, the spacer regions are processed to CRISPR RNAs (crRNAs), bind to other kinds of Cas nuclease(s) and together detect the foreign DNA, leading to its degradation and disposal (*210*).

Three CRISPR systems (type I, II, III) have been described (*211*). Certainly, the type II CRISPR system is of most interest for application in site-specific mutagenesis (*212*). It is in principle consisting of three components: a trans-activating RNA (tracrRNA), mature crRNA and the nuclease Cas9. The tracrRNA binds to the repeats of the pre-crRNA transcript. RNase III processing yields hybrids of tracrRNAs and crRNAs, called dual guide RNAs. Together with only one Cas protein (Cas9), targeting and degradation of foreign DNA is achieved (*210*). Like other Cas proteins, Cas9 consists of a RuvC (cleavage of DNA-strand not complementary to RNA) and an HNH (cleavage of DNA-strand complementary to RNA) nuclease domain that both provide nickase activity, together enabling a DNA DSB (*213*). Further, specific protospacer adjacent motives (PAM), 3' downstream of the target DNA, are of importance for recognition and binding by the guide RNA-Cas9 complex (*210*).
Like ZFN and TALEN, CRISPR/Cas9 is built of a DNA targeting domain and a nuclease, but the markedly difference is the RNA-guided mechanism. This is facilitating targeting of gene sequences for site-directed mutagenesis/genome editing. Because the target DNA is detected by the crRNA via complementary base-pairing, design of customized crRNAs is simple and straightforward. Targets are only restricted by the availability of a PAM (*210*). In 2012, investigations by working groups of Charpentier, Doudna and Siksnys revealed that Cas9 from *Streptococcus thermophiles* and *S. pyogenes* can be used for crRNA-guided cleavage of DNA *in vitro* (*214, 215*). The former two also showed, that it is possible to construct a fusion of tracrRNA and crRNA, which further facilitates application (Fig. 1-12) (*214*). This so-called single guide RNA (sgRNA) has a DNA-binding part (from crRNA) and a Cas9 interacting part (from tracrRNA) (*208*). Fauser *et al.* published *A. thaliana* genome editing using such a sgRNA together with *S. pyogenes* Cas9 (*213*). This Cas is requesting an NGG PAM and the target DNA length that is bound by crRNA is defined as 20 nt. This means that any DNA sequence with N_{21}GG can be targeted by this CRISPR system. The DNA DSB is then introduced three bp upstream of the PAM (*208, 213*). Several applications of CRISPR/Cas9 for genome editing in plants and model organisms have been published – focussing on both repair mechanisms NHEJ and HR (*216, 217*). Moreover, extensive basic research is conducted, considering e.g. site-specificity and off-target effects (*218–224*).

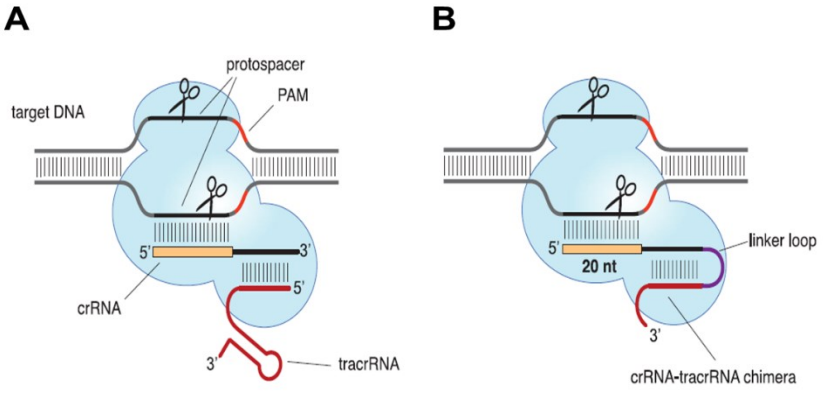

Fig. 1-12 CRISPR/Cas9 way of function. A) The naturally occurring mechanism by interaction of crRNA, tracrRNA and Cas9 (blue bubble) to bind to target DNA and to introduce a DSB upstream of the PAM. B) The engineered CRISPR/Cas9 system using a sgRNA (crRNA-tracrRNA chimera) for targeting. Adopted and modified from Jinek *et al.* (*214*).

1.4 Introduction of foreign DNA into plants

Besides using *A. tumefaciens* for insertional mutagenesis e.g. to detect phenotype-associated genes (chapter 1.3.1), the infective T-DNA system is used to introduce specific DNA sequences into the plant genome (*225*). Agrobacteria are plant pathogens, naturally infecting the host by transferring a tumour-inducing plasmid (Ti-plasmid) into the plant cell. A part of the plasmid, the transfer DNA (T-DNA), inserts into the host genome. This is causing the production of plant hormones (stimulating tumour growth) and of opines (AA-sugar conjugates that serve as an energy source for the bacterium) (*226*). This system is frequently used to transfer foreign DNA into plants. Therefore a two-plasmid system has been developed: the DNA of interest is cloned into a T-DNA binary vector, which is then introduced into a "disarmed" Agrobacterium consisting of an engineered *virulence* (*vir*) helper plasmid, lacking inducers for opine and hormone production (*227–229*). The T-DNA binary vector consists of the DNA of interest and a selection marker, flanked by short sequences mediating integration into the plant genome (left and right border), and origins of replication for both *A. tumefaciens* and *E. coli* to facilitate cloning (Fig. 1-13) (*227*). Introduction of DNA is achieved by floral dip (*230*), wounding the plant or incubation of the bacteria with plant explants, followed by regeneration (*231*). *A. rhizogenes* is another possible organism for the introduction of foreign DNA into plants, leading to formation of so called hairy roots which are indicators for successful transformation, simplifying screening (*232, 233*).

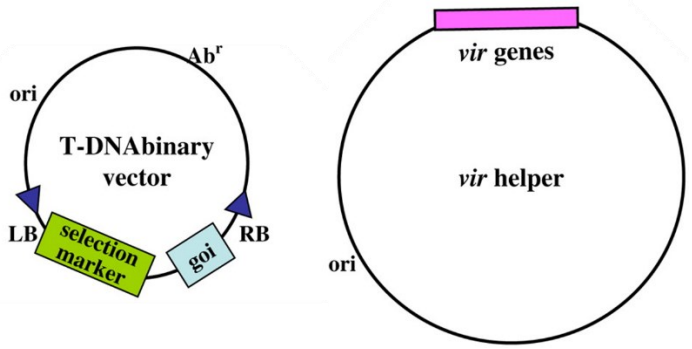

Fig. 1-13 Two-plasmid system for Agrobacterium-mediated delivery of DNA into plants. The T-DNA binary vector consists of: specific origins of replication (ori), an antibiotic resistance gene for selection (Ab^r) in bacteria and left and right borders (LB, RB) wherein the gene of interest (goi) and a selection marker for selection in plants are arranged. The disarmed Agrobacterium strain carries the *vir* helper plasmid, with an ori and virulence (*vir*) genes for enabling infection/transformation. Adopted and modified from Lee and Gelvin (*227*).

Several systems have been developed to directly deliver DNA into a plant cell: particle bombardment of plant parts or tissue/ suspension cultures, electroporation of plant cells/ tissues using silicon carbide whiskers with cell cultures or micro-injection of DNA into plant cells/ tissue. (*232*). Protoplasts instead of intact cells are frequently used in plant research; since protoplasts are lacking the cell wall, DNA can pass the membrane if it becomes porous by using electroporation (*234*), microinjection (*235*) or poly-ethylene glycol (PEG) (*236*).
These methods can be used to transfer the DNA of interest into a plant cell either stably (*237*) or transiently (*238*).

Currently, the CRISPR/Cas9 system is frequently applied, using both Agrobacteria transformation and protoplast transfection (*239*). For the former one, a plasmid consisting of the Cas9 as well as the sgRNA (and a repair template) sequence with appropriate promoters is transformed into plant cells/ tissues. Due to selection during regeneration, only stably transformed cells are maintained; transformed plants, anyhow, are rather chimeras. Therefore, those progeny plants are of interest that are lacking the T-DNA and consist of CRISPR/Cas9

induced stable and plant-uniform sequence changes (*240*). Transient transformation by Agrobacterium is described, too, evading the T-DNA problem (*241*). Protoplast transfection by delivering sgRNA, purified Cas9 and possibly a repair template instead of an expression plasmid represents a faster (and DNA-free) tool. Anyhow, protoplast isolation, transfection and regeneration are complicated and have to be established for each species. Often regeneration is a bottleneck, nevertheless the method is suitable for investigating into e.g. mutagenesis efficiency (*239, 242, 243*). In Tab. 1-4 exemplary plant species and methods for CRISPR/Cas9 approaches are listed.

Tab. 1-4 Exemplary plant species edited with CRISPR/Cas9 using PEG-induced, plasmid-free transfection of protoplasts or Agrobacterium-mediated transformation.

Plant species	**Method for introduction of CRISPR/Cas9**	**Reference**
A. thaliana	protoplasts	Li *et al.* (*244*), Lin et al. (*245*)
	Agrobacterium	Li *et al.* (*244*), Mao *et al.* (*246*), Xing *et al.* (*247*)
Nicotiana benthamiana	protoplasts	Gao and Zhao (*248*)
	Agrobacterium	Nekrasov et al. (*249*), Gao et al. (*250*)
O. sativa	protoplasts	Lin et al. (*245*)
	Agrobacterium	Ma et al. (*251*), Li et al. (*252*),

1.5 Thesis objectives

T. koksaghyz is a perennial Asteraceae, which is able to produce natural rubber, qualitatively comparable to that of the rubber tree. The lacking results of former breeding efforts from the early 20th century and therefore lacking homogeneous material for research nowadays are impeding the success of the plant on the field as a serious bioresource. In contrast to other established crops, an efficient weed management is lacking. This is a crucial obstacle that could be overcome if an herbicide resistant *T. koksaghyz* would be at hand.
For generation of herbicide resistant *T. koksaghyz* following issues are to be investigated:

- Research on the gene number of the well-known herbicide target enzyme AHAS in *T. koksaghyz*. In the context of this work, the catalytic subunit of AHAS is discussed.
- Undirected mutagenesis of *T. koksaghyz* seed material and selection for AHAS conferred herbicide resistant plants.
- Site-directed genome editing of the AHAS gene in *T. koksaghyz* and screening for herbicide resistance conferring mutations.
- Transcriptomic analysis of herbicide resistant and putative tolerant *T. koksaghyz* plants derived from EMS mutagenesis.

2 Materials and methods

2.1 Materials

2.1.1 Bacterial strains

A. tumefaciens EHA105	from the working group of Prof. Dr. D. Prüfer, WWU Münster (*253, 254*)
E. coli DH5alpha	*fhuA2* (*argF-lacZ*)*U169 phoA glnV44 80* (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17* (*255*)

2.1.2 Chemicals and consumables

Chemicals and consumables including kits, polymerases, primer, DNA ladder and restriction enzymes were purchased from the following companies: Agilent Technologies, Inc. (Santa Clara, California, USA), Becton, Dickinson and Company (Le Pont de Chaix, France), Carl Roth GmbH + Co. KG (Karlsruhe, Germany), Duchefa Biochemie (Haarlem, Netherlands), Invitek Molecular GmbH (Berlin, Germany), Linde AG (Pullach, Germany), Merck KGaA (Darmstadt, Germany), MP Biomedicals (Eschwege, Germany), metabion international AG/metabion GmbH 2018 (Planegg/Steinkirchen, Germany), New England Biolabs GmbH, Takara Bio USA, Inc. (Kusatsu, Japan), Pro-Pac Ostendorf Plastic (Vechte, Germany), Thermo Fisher Scientific (Waltham, Massachusetts, USA).
Pulsar® 40 was obtained as a sample from BASF Österreich GmbH (Vienna, Austria).
Insects were purchased from SAUTTER & STEPPER GmbH (Ammerbuch, Germany).

2.1.3 Equipment

Agilent 2100 Bioanalyzer	Agilent Technologies, Inc. (Santa Clara, United States)
Autoclave Certoclav Classic	CertoClav Sterilizer GmbH (Traun/ Austria)
Autoclave Laboklav 80MV-FA	SHP Steriltechnik AG (Satuelle/ Germany)
Centrifuge Eppendorf 5415D	Eppendorf AG (Hamburg/ Germany)
Centrifuge Sigma 4K15C	Sigma Laborzentrifugen GmbH (Osterode am Harz/ Germany)
Eppendorf Thermomixer Comfort	Eppendorf AG (Hamburg/ Germany)
Gel caster	bsb11 biotech service blu (Kassel, Germany)
Gel caster system Biostep	biostep GmbH (Burkhardtsdorf/ Germany)
Gel electrophoresis chamber Biostep	biostep GmbH (Burkhardtsdorf/ Germany)
Gel electrophoresis chamber PeqLab Class II	PeqLab/ VWR International, LLC. (Darmstadt/ Germany)
Gel electrophoresis documentation: Gel iX20 Imager & Mitsubishi P93D	Intas Science Imaging Instruments GmbH (Göttingen/ Germany) Mitsubishi Electric Corporation (Chiyoda, Japan)
Gel electrophoresis Power Supply GE Healthcare EPS301	General Electric (Chicago, U.S.A.)
HANNA edge HI12300 pH meter	Hanna Instruments Deutschland GmbH (Vöhringen/ Germany)
Harris Uni-Core Puncher (0.5 mm)	Merck KGaA (Darmstadt, Germany)
Heraeus Biofuge Stratos Centrifuge	Fisher Scientific GmbH (Schwerte/ Germany)
Milli-Q® Reference Water Purification System	Merck KGaA (Darmstadt/ Germany)
Mill Retsch MM301	Retsch GmbH (Haan/ Germany)

Nanodrop 8000	Fisher Scientific GmbH (Schwerte/ Germany)
PCR cycler GeneTouch Thermal Cylcer	Biozym Scientific GmbH (Hessisch Oldendorf/ Germany)
RUMED 1200/ 1301	Rubarth Apparate GmbH (Laatzen/ Germany)
Shaker Infors HT Minitron	Infors AG (Bottmingen, Switzerland)
Shaker innova44	New Brunswick Scientific/ Eppendorf AG (Hamburg/ Germany)
Sonifier W-250-D	Branson, Emerson Electric Co (Langenfeld Rheinland/ Germany)
Spectrophotometer Genesys 20	Thermo Fisher Scientific Inc. (Waltham, U.S.A)
Tissue Grinder Mixy Professional	NIPPON Genetics Europe (Düren, Germany)
TissueLyser Adapter Set 2 x 24	QIAGEN (Germantown, Maryland, USA)
Vortexer VortexGenie 2	Bender& Habein AG (Zürich, Switzerland)

2.1.4 Media for bacteria and plants; solutions

Media for the work with E. coli and A. tumefaciens

- Lysogenic broth: 25 g/ L LB broth (Luria/Miller); solid: plus 16 g/ L Agar-Agar, Kobe I
- Yeast extract broth: 5 g/ L beef extract, 5 g/ L saccharose, 5 g/ L peptone, 1 g/ L yeast extract, 0.2 % $MgSO_4x7H_2O$; solid: plus 16 g/ L Agar-Agar, Kobe I
- Additives: 100 mg/ L Rifampicin, 100 mg/ L Spectinomycin
- 10x Super Optimal Broth (SOC) buffer: 2 g/L KCl, 20 g/ L $MgCl_2$, 20 g/ L $MgSO_4$, 40 g/ L glucose

Media for T. koksaghyz transformation

- MS medium: 4.4 g/ L Murashige & Skoog Medium, 20 g/ L Saccharose, 1 mg/ L BAP, 0.2 mg/ L IAA; pH 5.8; solid: plus 6 g/ L Plant agar
- Infiltration medium: MS medium plus 200 µM Acetosyringone
- Callus inducing medium (CIM): MS medium plus 400 mg/ L Amoxycillin sodium / Clavulanate potassium, 2.4 mg/ L PPT, 6 g/ L Plant agar
- Shoot inducing medium (SIM): MS medium plus 0.5 mg/ L Kinetin, 0.1 mg/ L IAA, 400 mg/ L Amoxycillin sodium /Clavulanate potassium, 2.4 mg/ L PPT, 6.3 g/ L Plant agar
- Root inducing medium (RIM): MS medium plus 0.2 mg/ L IAA, 400 mg/ L Amoxycillin sodium /Clavulanate potassium, 2.4 mg/ L PPT, 6.5 g/ L Plant agar

Media for T. koksaghyz clonal propagation and regeneration

- MS medium: 4.4 g/ L Murashige & Skoog Medium, 20 g/ L Saccharose, 1 mg/ L BAP, 0.2 mg/ L IAA; pH 5.8; solid: plus 6 g/ L Plant agar
- Callus inducing medium (CIM): MS medium plus 400 mg/ L Amoxycillin sodium / Clavulanate potassium, 6 g/ L Plant agar
- Shoot inducing medium (SIM): MS medium plus 0.5 mg/ L Kinetin, 0.1 mg/ L IAA, 400 mg/ L Amoxycillin sodium /Clavulanate potassium, 6.3 g/ L Plant agar
- Root inducing medium (RIM): MS medium plus 0.2 mg/ L IAA, 400 mg/ L Amoxycillin sodium /Clavulanate potassium, 6.5 g/ L Plant agar

Other solution

- TAE buffer for gel electrophoresis:40 mM Tris-Acetate, pH 8.0, 0.1 mM Na_2EDTA

2.1.5 Plant material and growth conditions

T. koksaghyz field mix seed material was received from the working group of K. Thiele (JKI Quedlinburg). It originates from initially received material from ESKUSA GmbH (cooperation partner of the project EVITA (031A285B)) and is a mixed population of five USDA lines (W6 35156, W6 35170, W6 35172, W6 35173, W6 35182).
Plant material of the *T. koksaghyz in vitro* line Tks203 was received from the Prüfer group.
"Reference plants" (no. 4A, 68A, 69A, 78A, 81A, 83A, 131A, 139A) were received from ESKUSA, selected material by the breeder derived from a mixed population of five USDA lines (W6 35156, W6 35170, W6 35172, W6 35173, W6 35182).
T. koksaghyz in vitro plants were sterile cultivated in plastic containers (Pro-Pac Ostendorf Plastic) on ½ MS medium (2.2 g/ L Murashige and Skoog salt solution, 10 g/ L glucose, and 6.5 g/ L plant agar, pH 5.8) at 16 °C with a 16 h photoperiod, according to Wahler *et al.* (*14*). Long-term storage of *T. koksaghyz in vitro* plants was carried out on medium containing additionally activated carbon (2 g/ L).
T. koksaghyz plants in the green house were cultivated on sieved sand-humus mixture (2:1 v/v) (in-house preparation) at 18 °C (day) and 16 °C (night), respectively with a 16 h photoperiod of 25,000 lx (if necessary complemented by artificial light). Flowering induction, if necessary, was achieved by vernalisation for two weeks at 4 °C with an 8 h photoperiod of 5,000 lx.

2.1.6 Primer

Tab. 2-1 List of the primer used.

Name	Sequence 5'-3'	Melting temperature [°C] or further description
For amplification and sequencing of *TkoAHAS1* (see also Fig. 6-2)		
ALS-5'UTRkok	GTTCAACATGTATAAACAACTG	58
AHAS1_-20_fwd	CAAAGAAACACTCCAACCCT	58
AHAS1_60_fwd	CCCATTTCAGCCTCGCAC	58
ALS-R2kok	TCCACGAGGACATCGGAAC	59
AHAS_438_fw	ACCCGGCGTATGTATTGCTA	60
AHAS_499_fw	GATGCGCTGCTTGACAGT	58
AHAS1_692_rev	CGGAAGATGCGAGATAGAAAG	62
ALS-3A	GTGTTTGGATTCGAGCGTTG	60
AHAS1_1022_rev	CAGTTCCATGCATTCCGAG	58
I-ALS-R2	AGCAGAATCGATGTCGATGTG	59
AHAS1_1185_rev	ACATGGGGTTGTTTGTTCT	57
I-ALS-R1	AGACCCTGTAATGCGATCTTG	59
AHAS1_1295_rev	GTCTAATTCCTTCCTCCATG	57
I-ALS-1	ATCTAGGTATGGTCGTTCAATG	58
I-ALS-2	ACGTTATTGTCCCACATCAAG	57
ALS-STOPkok	TCAATATTTGATTCGGCCATC	55
ALS_R00	TTCATAAACTAGTTATAACAAAACG	56
ALS-R0 kok	GAAGCATCATGTGTAACATAAG	58
For CRISPR/Cas9 cleavage assay		
AHAS_CRISRNA_363_FW	GAGAGGAGATTCGACGATCGTGTAAC	Oligos for intention to target AA 363
AHAS_CRISRNA_363_RV	AAACGTTACACGATCGTCGAATCTCC	

Name	Sequence 5'-3'	Melting temperature [°C] or further description
		Cleavage site: 1091 nt
AHAS_CRISRNA_639_FW	GAGAGGTTGCCTATGATCCCCGCCGG	Oligos for intention to target AA 639 Cleavage site: 1916 nt
AHAS_CRISRNA_639_RV	AAACCCGGCGGGGATCATAGGCAACC	
AHAS_CRISRNA_324_fw	GAGAGGCTACCCTGGCGGCGCATCCA	Oligos for intention to target AA 108 Cleavage site: 325 nt
AHAS_CRISRNA_324_rv	AAACTGGATGCGCCGCCAGGGTAGCC	
AHAS_CRISRNA_549_fw	GAGAGGCCATCACCGGCCAAGTTCCC	Oligos for intention to target AA 183 Cleavage site: 546 nt
AHAS_CRISRNA_549_rv	AAACGGGAACTTGGCCGGTGATGGCC	
AHAS_CRISRNA_572_fw	GAGAGGGATCGGAACTGACGCATTTC	Oligos for intention to target AA 191 Cleavage site: 574 nt
AHAS_CRISRNA_572_rv	AAACGAAATGCGTCAGTTCCGATCCC	
AHAS_CRISRNA_1680_fw	GAGAGGATCTAGGTATGGTCGTTCAA	Oligos for intention to target AA 560 Cleavage site: 1674 nt
AHAS_CRISRNA_1680_rv	AAACTTGAACGACCATACCTAGATCC	
M13_fwd(+20)	GTAAAACGACGGCCAGTG	56
M13_rev(-20)	CACAGGAAACAGCTATGAC	54
For CRISPR/Cas9 cloning		
AHAS1_CRISPR_363_FWD	ATTGAGATTCGACGATCGTGTAAC	Oligos for targeting AA 363 Target site: 1091 nt
AHAS1_CRISPR_363_REV	AAACGTTACACGATCGTCGAATCT	
AHAS1_CRISPR_639_FWD	ATTGTTGCCTATGATCCCCGCCGG	Oligos for targeting AA 639 Target site: 1916 nt
AHAS1_CRISPR_639_REV	AAACCCGGCGGGGATCATAGGCAA	
For verification of efficient cloning/transformation of CRISPR/Cas9 constructs		
M13_rev(-20)	CACAGGAAACAGCTATGAC	54
pSBE2_rev	CACTATCTTCACAATAAAGTG	56
SS42	TCCCAGGATTAGAATGATTAGG	60
SS43	CGACTAAGGGTTTCTTATATGC	60
For testing of mcs from cPCR of pJET/ pTZ57R/T		
pJET1.2-FWD	CGACTCACTATAGGGAGAGCGGC	60
pJET1.2-REV	AAGAACATCGATTTTCCATGGCAG	60
M13_fwd(+20)	GTAAAACGACGGCCAGTG	56
M13_rev(-20)	CACAGGAAACAGCTATGAC	54
For detection of T-DNA		
SS42	TCCCAGGATTAGAATGATTAGG	60
SS43	CGACTAAGGGTTTCTTATATGC	60
For searching for further AHAS sequences based on WWU transcriptome data		
2736434_fwd1	CTTCGATCTCCTTCTCCAC	58
2636062(rev)_fwd1	AGTACGTTGATGGGGCTG	56
2724878(rev)_rev1	GACCTTAATCGCATACTGTG	58
21826(rev)_fwd1	TCGGCTGAGATTGGGAAG	56
93774(rev)_fwd1	ACAATACGCGATTCAGGTC	56
2552443_rev1	GATTCTCTACCCTGATTGTG	58
199828_fwd	AGAGTGACGCACAAGGAC	58

Name	Sequence 5'-3'	Melting temperature [°C] or further description
199828_rev	CAACAACCGGATAGTCATAG	58
587450(rev)_fwd_1	GCAGATTGTTAGGTTGATTTC	58
587450(rev)_rev_1	TCGCAACAGGAATACCCGT	58
526463(rev)_fwd_1	CACAATCAGGGTAGAGAATC	58
526463(rev)_rev_1	GCTGGAATCCCACAAGCTG	60
3068(rev)_fwd_1	TGATCCCGAGTGGTGGCAC	62
3068(rev)_rev_1	CACCGGTTTTCACATCTCAG	60
For RACE		
c170_FWD	GAAATTTGCAGAGGCTTGTG	58
c170_REV	CACTGGGAATCATAGGTAGA	58
29774full_3'RACE_GSP_out	GTAGGGCAACATCAAATGTGGGCCGCAC	73
29774full_3'RACE_GSP_in	GACAATGGCTGACCTCCGTTGGGACAG	73
29774full_5'RACE_GSP_out	CTGTCCCAACGGAGGTCAGCCATTGTC	73
29774full_5'RACE_GSP_in	CAACGGAGGTCAGCCATTGTCGAGGTCTATTG	75

2.1.7 Vectors

The applied vector constructs used in this work a subsequently listed.

pEn-Chimera	pEn-Chimera was a gift from Holger Puchta (Addgene plasmid # 61432 ; http://n2t.net/addgene:61432 ; RRID:Addgene_61432) (*213*)
pEn-Chimera_T7	Derivative of pEn-Chimera, was a gift from Dr. Thorben Sprink, Quedlinburg Region of nt 1-851 of pEn-Chimera is replaced by 394 nt consisting of T7-promoter, mcs for sgRNA-protospacer, sgRNA backbone, T7 terminator (vector map: Fig. 6-3)
pDe-CAS9	pDe-CAS9 was a gift from Holger Puchta (Addgene plasmid # 61433 ; http://n2t.net/addgene:61433 ; RRID:Addgene_61433) (*213*)
pDe-Cas9_PcUbi	Derivative of pDe-CAS9, was a gift from Dr. Thorben Sprink, Quedlinburg Region of nt 14,092-14,142 of pDe-CAS9 is replaced by 1029 nt consisting of a PcUbi (promoter) (vector map: Fig. 6-4)
pJET1.2	Thermo Fisher Scientific (Waltham, Massachusetts, USA)
pTZ57R/T	Thermo Fisher Scientific (Waltham, Massachusetts, USA)

2.1.8 Polymerases

Tab. 2-2 List of DNA and RNA polymerases used.

Polymerase	Company	Product number
Standard DNA polymerase		
DreamTaq DNA Polymerase (5 U/µL)	Thermo Fisher Scientific	EP0701
High-fidelity DNA polymerases		
Phusion Hot Start II DNA Polymerase, 20 U	Thermo Fisher Scientific	F549
Q5® High-Fidelity DNA Polymerase	New England Biolabs GmbH	M0491L
RNA polymerase		
T7 RNA polymerase	Thermo Fisher Scientific	EP0111

2.1.9 RNA data

The unpublished database “Tks root transcriptome” from 09.10.2014 from the working group of Dr. Prüfer was available for targeted search inquiries (chapter 2.2.14).
RNA data obtained as part of this work: chapter 2.2.17.

2.2 Methods

2.2.1 Standard Methods

Tab. 2-3 List of standard methods polymerases used.

Method	Kit/ protocol
Plant tissue grinding	A) Using the Mill Retsch MM301 (2x1 min at 30 Hz), followed by isolation of genomic DNA method A) B) Using the Tissue Grinder Mixy Professional (15 sec), followed by isolation of genomic DNA method B)
Isolation of genomic DNA	A) Isolation method by Doyle and Doyle (*256*) B) Isolation method adapted from Edwards *et al.* (*257, 258*)
Purification of isolated genomic DNA	GENECLEAN SPIN KIT UN1760-III (MP Biomedicals)
Total RNA extraction	Bio&SELL RNA-Minikit (Bio&SELL GmbH)
Purification of extracted total RNA	GeneJET RNA Purification Kit (Thermo Fisher Scientific)
Purification of sgRNA	GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific)
Determination on RNA purity	Agilent RNA 6000 Nano Kit (Agilent Technologies, Inc.) Manufacturer´s instructions
Isolation of plasmid DNA	Invisorb® Spin Plasmid Mini Two
Isolation of PCR products, restriction enzyme digest products	GeneJET PCR Purification Kit (Thermo Fisher Scientific)
Restriction enzyme digest	Manufacturer´s instructions (New England Biolabs) (Life Technologies GmbH)
Ligation	Manufacturer´s instructions (Thermo Fisher Scientific)
Direct PCR	Phire Plant DirectPCR MM (Thermo Fisher Scientific)
Cloning of fragments (Subcloning)	CloneJET PCR Cloning Kit, InsTAclone PCR Cloning Kit (Thermo Fisher Scientific)
RACE	SMARTer® RACE 5’/3’ Kit User Manual (Takara Inc.)

2.2.2 Cloning, transformation and analysis of vectors into *E. coli*

Cloning experiments were planned with the software CLC Main Workbench 8.1.3 (QIAGEN Aarhus, Sweden). Vector maps for this work were created with SnapGene software (from Insightful Science; available at snapgene.com) Version 6.0.2. Ligations of amplicons into the vectors pJET1.2 and pTZ57R/T, respectively were done according to the manufacturer's manuals (Thermo Fisher Scientific). Transformations (*259*) into *E. coli* DH5alpha was followed by analysis of colonies for the mcs by using the appropriate primer (Tab. 2-1) and the standard polymerase. For this purpose, a part of the colony was picked with a sterile tip and dispensed directly in the PCR approach.

2.2.3 Crossing of *T. koksaghyz* plants

Regenerates and progeny of plants from the EMS mutagenesis were pollinated manually or by using *Lucilia sericata* or *Bombus terrestris* according to Fig. 3-3. Insects were purchased from SAUTTER & STEPPER GmbH. Regenerates and progeny of plants from the CRISPR/Cas9 approaches were only pollinated manually.

2.2.4 Design, cloning and transformation of CRISPR/Cas9 constructs

For targeting the gene regions according to positions Arg363 and Ala639 of TkoAHAS1 binary vectors were generated according to Fauser *et al.* (*213*). The same protospacer regions as described in Tab. 2-4 were used; oligos (Tab. 2-1) and the vectors pEn-Chimera and pDe-Cas9_PcUbi were applied for the approach. Binary vectors were transformed into *A. tumefaciens* EHA105 with the freeze-thaw transformation (*260*) or by electroporation (*261*). For the analysis of colonies, a part of a colony was picked and resuspended in 100 µL ddH_2O, pelleted and resuspended twice in 200 µL of fresh water. Ultrasonic disruption of cells took place for 30 sec with 10 % amplitude. After vortexing, 2 µL of the solution was used in a 50 µL PCR approach with the standard DNA polymerase with the primer pSBE2_rev and AHAS1_CRISPR_363_FWD or AHAS1_CRISPR_639_FWD, respectively to verify containment of the correct construct.
These constructs (Fig. 6-5 and Fig. 6-6) were used for *A. tumefaciens* transformation likewise to Wahler *et al.* (*14*) (chapter 2.2.16).

2.2.5 DNA Sequencing

Sequencing of DNA samples was performed by GATC Biotech AG (Konstanz, Germany), nowadays Eurofins Genomics Germany GmbH (Ebersberg, Germany). Analysis of the results was achieved with the software CLC Main Workbench 8.1.3 (QIAGEN Aarhus, Sweden).

2.2.6 EMS mutagenesis, selection and analysis for resistant plants

Three EMS mutagenesis approaches with 30,000 seeds (approx. 15 g), each, were conducted with *T. koksaghyz* field mix seed material, harvested in 2013 (Thiele group). *T. koksaghyz* seeds were EMS-mutagenized at the WWU Münster (*262*). Afterwards, seeds were grown on sieved sand-humus mixture (2:1 v/v) (in-house preparation) at 18 to 25 °C with a 16 h photoperiod (if necessary complemented by artificial light). Seedlings at the third leaf stage were treated with the Imazamox herbicide Pulsar® 40. Survived plants were pricked separately. After symptoms subside and plants developed further leaves (6-8 weeks), Pulsar® 40 was applied, again. After plants survived several Pulsar® 40 treatments, they have been cut at ground level, waiting for regrowth and were Pulsar® 40 treated again.

Survived plants were sampled at the leaf tips, genomic DNA was isolated and PCR with a high-fidelity polymerase (Tab. 2-2) and appropriate primer (Tab. 2-1) was conducted according to the manufacturers manual to amplify *TkoAHAS1*.
PCR amplicons were cloned into pJET1.2 or pTZ57R/T according to manufactures guidance lines. Several positive clones were picked and amplification of the MCS was conducted with DreamTaq polymerase according to the manufacturer's manual using the appropriate primer (Tab. 2-1). PCR products were cleaned up as described under standard methods and subsequently sequenced.

2.2.7 *In vitro* cleavage assay

Six positions in *TkoAHAS1* were chosen to potentially target with the CRISPR/Cas9 system (Tab. 3-15). Protospacer regions were picked according to guidelines described by *Fauser et al.* (*213*), choosing the closest "NGG" to the target site in *TkoAHAS1*. Chosen protospacer regions/ designed sgRNAs were checked for reliable performance with an *in vitro* cleavage assay according to Jinek *et al.* (*263*). For this approach the BbsI (New England Biolabs) cut vector pEn-Chimera_T7 and the listed oligos (Tab. 2-1) were used. Cas9 (*Streptococcus pyogenes*) recombinant protein was in-house provided by Dr. Metje-Sprink. The cleavage reaction was performed for four hours and directly afterwards the analysis via gel electrophoresis (2 % agarose in TAE) was done. In Tab. 2-4 the investigated target sites are listed together with the used primer for generation of target DNA for the assay, the referring amplicon length and band sizes to be expected after a successful cleavage assay.

Tab. 2-4 List of investigated positions of *TkoAHAS1* by the *in vitro* cleavage assay and expected fragment sizes after successful cleavage of amplicons by the Cas9. DNA targets were derived by amplification of different regions from *TkoAHAS1* cloned into pJET1.2.

Target AA TkoAHAS1	Protospacer sequence 5' -3' in *TkoAHAS1*	Used primer for amplify-cation of target DNA	Amplicon length [bp]	DSB by Cas9 after nt …	Fragment sizes [bp] after successful *in vitro* cleavage
Ala108	CTACCCTGGCGGCGCATCCA	5'UTRkok ALS-R00	2173	325	444 1729
Pro183	CCATCACCGGCCAAGTTCCC			546	665 1508
Ala191	AGATTCGACGATCGTGTAAC			574	693 1480
Arg363	AGATTCGACGATCGTGTAAC			1091	1209 964
Trp560	ATCTAGGTATGGTCGTTCAA			1674	1793 380
Ala639	TTGCCTATGATCCCCGCCGG	ALS-3A ALS-R2kok	4535	1916	1033 3502

2.2.8 Phylogenetic Analysis

Phylogenetic analysis was performed with CLC Main Workbench 8.1.3. The analysis method clustalW (*264*) with default parameters was used and phylogenetic trees were constructed with maximum likelihood phylogeny (construction method: UPGMA, Protein substitution model: Bishop-Friday (*265*)).

2.2.9 Plant tissue grinding and genomic DNA isolation

For receiving high quality genomic DNA, the method A) (chapter 2.2.1) for plant tissue grinding followed by method A) for isolation of genomic DNA was used and further purification of isolated genomic DNA was pursued if needed. For plants that do not contain of high biomass or small plants from sterile cultivation or if a high number of plants had to be analysed, the method B) for plant tissue grinding followed by method B) for isolation of genomic DNA was used. Purification of isolated genomic DNA only took place if PCR approaches did not succeed due to disturbing contents (e.g. salt, tissue residues).

2.2.10 Pulsar®40 treatment and evaluation

Pulsar®40 is an herbicidal formula approved for commercial use in Austria. It contains 40 g/ L Imazamox. Its spectrum of action is not described against Taraxacum species. Further its usage is not approved in cultivation of Taraxacum species. The guidelines for usage in Imazamox tolerant sunflower cultivation (1.25 L/ha in 200-400 L water) were taken into account. Due to the late response (four weeks) of the *T. koksaghyz* plants towards this concentration, it was decided to use a 1.5-fold higher concentration to receive an earlier response (approx. two weeks). Immediately before a treatment with Pulsar®40, the required Pulsar®40/water mixture was freshly made. Application of the mixture was done using a 50 mL spray bottle. For evaluation of successful application, Tks203 clones were used as a control. Two weeks after treatment Pulsar®40 susceptible plants show reddish stains of the leaf edge and/ or browning of the heart of the rosette and/ or small and bar leaves (compare to Fig. 3-4) followed by dying of the whole plant. It occurred that re-growth out of the root was observed for older plants. Another Pulsar®40 treatment had to be performed then.

2.2.11 Rapid Amplification of cDNA ends (RACE)

In order to detect potential further AHAS sequences upstream or downstream the known *TkoAHAS1* as well as to identify 5’ and 3’ untranslated regions (UTRs) that are not congruent to the already known sequences, indicating further AHAS sequences, RACE was conducted. Experiments were run according to the manufacturer’s manual with the primer listed in Tab. 2-1.

2.2.12 Regeneration of plants from explants

If an interesting *T. koksaghyz* individual from the greenhouse was intended for clonal propagation under sterile conditions, an explant was taken, surface sterilised (according to chapter 2.2.15) and put on CIM, SIM and RIM consecutively (chapter 2.1.4). If the regenerated plants were approx. 10 cm high, clonal propagation by vertical division of one plant and recovery of the two plant parts proceeded on MS media.

2.2.13 Screening of plants via restriction enzyme assays

Plant material was sampled using a Harris Uni-Core Puncher (0.5 mm) according to the manufacturer's manual of Phire Plant DirectPCR MM (Thermo Fisher Scientific). Two leaf punches were collected in 10 µL of Dilution Buffer in a sterile, nuclease free PCR tube. Grinding of the plant material was conducted with a sterile 200 µL pipette tip. After short centrifugation, the supernatant was used for PCR. Samples were stored short-term at 4 °C and long-term (> 8 weeks) at -20 °C.

2.2.13.1 Screening of progeny from the EMS mutagenized material

The assay was performed to investigate the EMS induced sequence changes at the *TkoAHAS1* position 572. Therefore, a PCR fragment of 747 bp (*TkoAHAS1* nt 438 to 1185) was amplified. TaiI (recognition sequence ACGT) is not able to cut the fragment of the WT sequence (nt 572 is a Cytosine) but that of the modified sequence (nt 572 is a Thymine; cut after nt 572). Fragments are 135 bp and 612 bp in size.

PCR setup for one sample was as follows:

5 µL	2x PhirePlant Direct PCR Master Mix
0.1 µL	Primer forward (AHAS_438_fw)
0.1 µL	Primer reverse (AHAS_1185_rev)
4.55 µL	ddH_2O
0.25 µL	Supernatant of the sample

PCR programme was as follows:

98 °C	5:00	
98	0:05	40 x
59	0:05	
72	0:20	
72	1:00	
4	∞	

The PCR product was TaiI digested for 10 min at 65 °C in the following mixture:

0.5 µL	FD Green Buffer
0.25 µL	FASTDIGEST TaiI
4.25 µL	ddH_2O
2.5 µL	PCR product

The result of the digest was immediately analysed via gel electrophoresis: 2 % agarose gel, 20 min, 250 V, 250 mA. Visible DNA bands were evaluated as follows:

1 band at	747 bp	corresponds to	Homozygous genotype C/C (WT)
2 bands at	612 bp, 135 bp	corresponds to	Homozygous genotype T/T
3 bands at	747 bp, 612 bp, 135 bp	corresponds to	Heterozygous genotype C/T

2.2.13.2 Screening of progeny of the CRISPR/Cas9 transformation approach

The assay was performed to screen the progeny of the plants from the CRISPR/Cas9 transformation approach targeting *TkoAHAS1* nt position 1916. Screening reduced the number of individuals, which had to be investigated in detail (*TkoAHAS1* amplification and sequencing). Therefore, a PCR fragment of 1257 bp was amplified (*TkoAHAS1* nt 885 to +170 (3'UTR region)). PdiI (recognition sequence GCCGGC) is able to cut the fragment of the WT sequence (cut after nt 1917) but not if any modification occurred in the recognition site (*TkoAHAS1* 1915-1920). Fragments are 1033 bp and 224 bp in size. *TkoAHAS1* from individuals with a single/ uncut fragment was further analysed by Sanger sequencing.

PCR setup for one sample was as follows:

5 µL	2x PhirePlant Direct PCR Master Mix
0.1 µL	Primer forward (ALS-3A)
0.1 µL	Primer reverse (ALS-R0kok)
4.55 µL	ddH_2O
0.25 µL	Supernatant of the sample

PCR programme was as follows:

98 °C	5:00	
98	0:05	40 x
58	0:05	
72	0:26	
72	1:00	
4	∞	

The PCR product was TaiI digested for 5 min at 37 °C followed by incubation for 15 min at 65 °C in the following mixture:

0.5 µL	FD Green Buffer
0.25 µL	FASTDIGEST PdiI
4.25 µL	ddH_2O
2.5 µL	PCR product

The result of the digest was immediately analysed via gel electrophoresis: 2 % agarose gel, 20 min, 250 V, 250 mA. Visible DNA bands were evaluated as follows:

2 bands at	1033 bp, 244 bp	corresponds to	Unmodified DNA sequence (WT)
1 band at	1257 bp	corresponds to	Modified DNA sequence

2.2.14 Searching for further AHAS sequences based on WWU transcriptome data

The transcriptome data (chapter 2.1.9) from the Prüfer group (WWU Münster) was available for targeted search inquiries to investigate for further AHAS sequences. BLAST searches with the known *TkoAHAS1* against the transcriptome data were performed. *TkoAHAS1* was retrieved. Contigs were identified possibly belonging to another AHAS sequence(s) (Tab. 2-5). Primer were designed (Tab. 2-1) and genome walking to identify genomic sequences adjacent to the identified contigs or amplification of fragments out of *T. koksaghyz* genomic DNA (Tks203) was attempted.

Tab. 2-5 List of identified contigs from the transcriptome data of the Prüfer group. Either genome walking or amplification of the discovered fragments from genomic DNA was performed.

Contig name	Aim
MUE_TK_C2736434 8.0 Length=918	Genome walking
MUE_TK_C2636062 15.0 Length=356	
MUE_TK_C2724878 19.0 Length=751	
noHitAssembly_c21826_g1_i2 len=187	
noHitAssembly_c93774_g1_i1 len=155	
MUE_TK_C2552443 5.0 Length=254	
JKI_TK_C199828 9.0 Length=411 noHitAssembly_c587450_g1_i1 len=126	Amplification of fragments out of genomic DNA
noHitAssembly_c526463_g1_i1 len=241 noHitAssembly_c3068_g1_i1 len=233	

2.2.15 Surface sterilisation of seed material

Whenever small proportions of progeny were generated, germination of seeds was not directly exercised on sand-humus mixture. Instead, the seed material was surface sterilised for 13 min in 3 % sodium hypochlorite solution including 50 µL Tween, followed by 8 times of washing with sterile distilled water. Sterilized seeds were applied on MS plates (chapter 2.1.4) for germination at 16 °C with a 16 h photoperiod and after seedlings germinated they were planted into the greenhouse (parameter according to chapter 2.1.5).

2.2.16 *T. koksaghyz* transformation experiments

Transformations with binary plasmids of *T. koksaghyz* Tks203 *in vitro* material was conducted likewise to Wahler *et al.* (*14*) using the *A. tumefaciens* strain EHA 105. However, plant explants were cut out with a scalpel and freshly transformed material was incubated on MS medium for 48 h in the dark at 20 °C. Media used are listed in chapter 2.1.4. Regenerates and progeny of the regenerates were checked for inheritance of *A. tumefaciens* tDNA. Therefore, a leaf sample was collected, whole DNA was isolated and PCR was performed using the standard polymerase and the appropriate primer (Tab. 2-1).

2.2.17 Transcriptome approaches

Transcriptome data sets are available at the Hartung group and at the Thiele group at the JKI Quedlinburg. Additionally, Dataset mentioned in chapter 2.2.17.1 is available at the Prüfer group at WWU Münster.

2.2.17.1 Transcriptome Set "INVIEW Transcriptome Discover" from 29.12.2015

Samples:

T. koksaghyz Tks203 plants were planted from *in vitro* to the greenhouse and further to open land at the JKI Quedlinburg. After acclimatisation, samples were taken from flowering plants (age unknown but not of importance for the aim of the approach) and were instantly frozen in liquid nitrogen. The sampled inflorescence (sample "flower") was at the stage of beginning to open. The collected leaf sample (sample "leaf") was the 3rd youngest leaf of the plant. Before milling the samples with mortal and pistil, the midvein was removed from the leaf sample. RNA isolation, cleaning and quality check was done according to the listed standard methods (2.2.1).

Key data of the approach:

Company: GATC Biotech AG (Konstanz, Germany)

Requirement for the samples:

- Total RNA amount: 1 µg per sample
- Concentration: > 20 ng/ µL
- Purity: OD 260/280: 1.8 - 2.2 or RIN value ≥ 8
- Sample: solved in clean, DNase, RNase, protease free water

Technology: Genome Sequencer Illumina HiSeq2500

Library: Strand-specific cDNA library

Method: Isolation of poly(A)+ RNA, mRNA fragmentation, random-primed cDNA synthesis (strand-specific), adapter ligation and PCR amplification

Parameter: Paired End Run

Read length: 2 x 125 bp

Output: FastQ files (Tab. 6-6)

Processing:

Processing of raw data was done by using the in-house Galaxy (*266*) server.

- Trimming was done using Trim Galore! adaptive quality and adapter trimmer (Galaxy Version 0.4.0), using default parameter, besides: Trim low-quality ends from reads in addition to adapter removal (30), Discard reads that became shorter than length INT (60), Unpaired single-end read length cutoff needed for read 1 and 2 to be written (61)
- Trimmed data of the paired reads was assembled separately via Trinity (Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.0.6.1), default parameter) (*267, 268*). These assemblies are serving as references for the further work.
- A tblastx (NCBI BLAST+ tblastx Search translated nucleotide database with translated nucleotide query sequence(s) (Galaxy Version 0.3.0), default parameter) was conducted with the known TkoAHAS1 sequence against the two transcriptome assemblies in order to search for further AHAS genes.
- A tblastn (NCBI BLAST+ tblastn Search translated nucleotide database with protein query sequence(s) (Galaxy Version 0.3.0), default parameter) was conducted with following AHAS protein sequences in order to search for further AHAS genes in *T. koksaghyz*.

Tab. 2-6 List of AHAS protein IDs used for the tblastn approach.

Organism	**ID (Source)** **Source:** **NCBI Reference Sequence/ GenBank (N) or UniProtKB/Swiss-Prot (U)**
Xanthium sp.	CBV23123.1 (N)
T. koksaghyz	"TkokAHAS_allSNPs" (Annex 6-4)
Sonchus asper	ACF47583.1 (N)
Phaseolus_vulgaris	AGZ15379.1 (N)
Helianthus annuus	AAT07329.1 (N)
Helianthus annuus	AAT07328.1 (N)
Helianthus annuus	AAT07326.1 (N)
Helianthus annuus	AAT07324.1 (N)
Glycine max	XP_003528106.1 (N)
Galium aparine	ADI23953.1 (N)
Eupatorium cannabinum	CAX48741.1 (N)
Bassia scoparia	AAC69629.1 (N)
Arabidopsis thaliana	P17597.1 (U)
Amaranthus retroflexus	AAK50820.1 (N)
Anthemis cotula	AEL89168.1 (N)

- SNP calling (MPileup DC call variants (Galaxy Version 2.0), default parameter, besides: additional output of DP (Number of high-quality bases) and DV (Number of high-quality non-reference bases)) was done for TkoAHAS1.

2.2.17.2 Transcriptome Set "TruSeq Stranded mRNA" from 26.03.2018

Samples:

The three plants M_2-1/ 2/ 3 were transferred from the greenhouse to the *in vitro* cultivation for clonal propagation: leaf explants were sterilized and cultivated on CIM, SIM and RIM consecutively to generate clonal plants. After several weeks of growth, until the plants reached approx. 10 cm heights, clones were planted back into the greenhouse in trays (33.5 x 51.5 cm, 24 pots à 7.5 cm diameter). Field mix seeds (from 2016) were grown in the greenhouse. Six

weeks later, all plants were treated with Imazamox or water as control. Imazamox or water (= control) treatment was conducted simultaneously and samples were collected 24 hours later (according to Balabanova *et al.* (*269*)). 2 qcm of a leaf point from an approx. three-week-old leaf was collected each. Samples were immediately frozen in liquid nitrogen and milled using mortar and pistil. RNA isolation, cleaning and quality check was done according to the listed standard methods.

Key data of the approach:
Company: Macrogen Inc.
Requirement for the samples:
Total RNA amount: 1 µg per sample
Concentration: > 20 ng/ µL
Purity: RIN value > 7
Sample: solved in clean, DNase, RNase, protease free water
Technology: Genome Sequencer Illumina NovaSeq 6000
Library: Strand-specific cDNA library (TruSeq Stranded mRNA)
Method: Isolation of poly(A)+ RNA, random fragmentation of cDNA, 5' and 3' adapter ligation and PCR amplification
Parameter: Paired End Run
Read length: 2 x 101 bp
Output: FastQ files (Tab. 6-7)

Processing:
Processing of raw data was done by using the in-house Galaxy (*266*) server.

- Trimming of raw data was done using Trim Galore! adaptive quality and adapter trimmer (Galaxy Version 0.4.0), using default parameter, besides: Trim low-quality ends from reads in addition to adapter removal (30), Discard reads that became shorter than length INT (50), Unpaired single-end read length cut off needed for read 1 and 2 to be written (51).
- Trimmed data sets were mapped (Map with BWA-MEM (Galaxy Version 0.8.0), default parameter) against each other to receive paired and unpaired reads as flagstats-data.
- All processed RNAseq data are assembled via Trinity (Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.0.6.1), default parameter) (*267, 268*), which serves as a reference (common transcriptome assembly) for the further work (approx. 224,000 partial transcripts).
- Mapping (Map with BWA-MEM (Galaxy Version 0.8.0), default parameter) of every processed data set against the common transcriptome.
- Executing Transdecoder (TransDecoder Find coding regions within transcripts (Galaxy Version 3.0.1), default parameter) (*267*) on the common transcriptome to specify functional regions, in order to annotate them. With the annotation, a comparison of different data sets is possible.
- Executing Feature Count (featureCounts Measure gene expression in RNA-Seq experiments from SAM or BAM files. (Galaxy Version 1.4.6.p5), default parameter) on the mapping data to get information, which and what amount of fragments did map against the common transcriptome.
- Finally, the DeSeq (DESeq2 Determines differentially expressed features from count tables (Galaxy Version 2.11.40.1), default parameter) was done to compare

differentially expressed (DE) transcripts from different samples against each other (using the feature count data) (Imazamox treated versus untreated samples).

- DE transcripts were (partially) annotated by using Blast2Go (Blast2GO Maps BLAST results to GO annotation terms (Galaxy Version 0.0.9), default parameter).
- Only the DeSeq data from the comparison of M_2-1/ 2/ 3 was used: 755 DE transcripts were found (180 without annotation). Among the 575 remaining, annotated DE transcripts, it was selected for those with annotations according to Tab. 3-21.
- For the 25 transcripts that were found solely in M_2-3 (Tab. 3-21) SNP calling (MPileup DC call variants (Galaxy Version 2.0), default parameter, besides: additional output of DP (Number of high-quality bases) and DRP (Number of high-quality bases for each observed allele)) was done for all samples.
- SNPs were filtered by quality (cutoff: genotype quality ≥ 40; error probability ≤ 0.0001).
- SNPs only occurring in M_2-3 but not in M_2-1/ 2 were detected by searching for genotype (GT) information on the MPileup data.
- In order to evaluate the potential impact of a SNP on the protein level, a blastx for each transcript was done. Always, the best fitting result (identity) was used as a reference sequence to identify the reading frame of the transcript fragment. Afterwards it was possible to identify silent and non-silent mutations of the fragments.
- It was checked whether the detected SNPs of M_2-3 also exist in the field mix data. The evaluation was conducted using the continuous genotype (cg) information from the MPileup data. The cg can be calculated as follows:

  $$cg\ [\%] = n_0 / (n_0 + n_1) * 100$$

  n_0 = absolute read count of reads matching the reference at position xy

  n_1 = absolute read count of reads not matching the reference at position xy
- Apart from literature search, the protein reference sequences were investigated for protein domain and structure information by using the database UniProt Knowledgebase (*270*) and the Simple Modular Architecture Research Tool (*271*).

3 Results

3.1 The acetohydroxyacid synthase of *T. koksaghyz*

Initially, in 2014, neither a genome sequence, nor an *AHAS* sequence of the diploid *T. koksaghyz* L. Rodin was publicly available. Internal investigations by the working group of F. Hartung (JKI Quedlinburg) on samples from *T. koksaghyz* field trials from the working group of K. Thiele (JKI Quedlinburg) resulted in two first preliminary *AHAS* allele sequences with eight identified allelic positions (Tab. 3-1). However, further investigations into SNP compositions have not been accomplished and it was unclear if *T. koksaghyz* consists of more *AHAS* genes.

Tab. 3-1 Preliminary alleles of *TkoAHAS1* (preliminary allele 1 sequence see Annex 6-1) and the corresponding positions and character on AA level.

nt level			**AA level**		
Position	**Character Preliminary Allele 1**	**Preliminary Allele 2**	**Position**	**Character Preliminary Allele 1**	**Preliminary Allele 2**
104	T	C	**35**	Leu	Pro
387	C	T	**129**	His	
522	G	A	**174**	Pro	
762	C	T	**254**	Asp	
936	C	T	**312**	Ile	
1020	T	A	**340**	Thr	
1095	C	T	**365**	Thr	
1642	T	A	**548**	Leu	Ile

The identified AHAS gene of *T. koksaghyz* (further named *TkoAHAS1* or *AHAS1*) has a length of 1971 bp, is intron free and encodes for a 656 AA long protein. Determination of allelic positions of *AHAS1* and search for further genes were supposed to be achieved by using eight selected, specific *T. koksaghyz* plants from the EVITA cooperation partner and breeding company ESKUSA GmbH (chapter 2.1.5). Those results were ought to serve as a reference for evaluating the results of the EMS mutagenesis approach, too.
Thirteen different *AHAS1* alleles were identified among the reference plants (Tab. 3-2).

Tab. 3-2 Thirteen different *TkoAHAS1* allele sequences detected in the reference plants from ESKUSA. For a better overview, the former listed allelic character is represented by a blank space. The complete CDS is exemplarily given for allele sequence 1 in Annex 6-2 and the CDS with all SNP positions together in one sequence are given in Annex 6-3.

Allele sequence no.	**nt position and character**											
	13	**104**	**110**	**154**	**240**	**417**	**928**	**936**	**1020**	**1095**	**1385**	**1642**
	C/A	**T/C**	**C/A**	**C/G**	**C/T**	**T/C**	**A/G**	**T/C**	**A/T**	**T/C**	**G/C**	**T/A**
	AA position and character											
	5	**35**	**37**	**52**	**80**	**139**	**310**	**312**	**340**	**365**	**462**	**548**
	Pro/ Thr	**Leu/ Pro**	**Ala/ Asp**	**Leu/ Val**	**Asp**	**Gly**	**Thr/ Ala**	**Ile**	**Thr**	**Thr**	**Gly/ Ala**	**Leu/ Ile**
1		C			T			C	T			
2	A											
3		C		G		C		C	T	C		
4												

Allele sequence no.	nt position and character											
	13	104	110	154	240	417	928	936	1020	1095	1385	1642
	C/A	T/C	C/A	C/G	C/T	T/C	A/G	T/C	A/T	T/C	G/C	T/A
	AA position and character											
	5	35	37	52	80	139	310	312	340	365	462	548
	Pro/ Thr	Leu/ Pro	Ala/ Asp	Leu/ Val	Asp	Gly	Thr/ Ala	Ile	Thr	Thr	Gly/ Ala	Leu/ Ile
5												A
6											C	
7		C			T						C	
8		C			T			C	T	C		
9	A							C	T	C		
10		C						C	T	C		
11			A									
12			A								C	
13		C					G		T	C		

Besides the plants from ESKUSA, two other *T. koksaghyz* sources were investigated: an *in vitro* line ("Tks203") from the working group of D. Prüfer (WWU Münster) and seed material from field trials ("field mix") from the working group of K. Thiele (JKI Quedlinburg). For the *in vitro* line the *AHAS1* allele sequences have been determined, too. In Tab. 3-3 it is summarized which alleles occurred in the Tks203 line and the reference plants, referring to the numeration of Tab. 3-2. The diploid status of Tks203 has been verified by the Prüfer group in the past and is reflected by the occurrence of two alleles. Despite, 50 % of the reference plants consist of three instead of two alleles.

Tab. 3-3 The occurrence of *AHAS1* alleles in the reference plants from ESKUSA and the *in vitro* line Tks203; numeration of allele sequences are referring to Tab. 3-2.

Plant designation	Occurring allele sequences (no.)		
4A	1	2	
68A	3	4	5
69A	6	7	8
78A	4	6	
81A	2	9	10
83A	6	11	12
131A	2	4	
139A	2	13	
Tks203	4	10	

The thirteen different allele sequences show variation on the nt level between the different sources. From the twelve identified allelic positions only seven are provoking AA changes while five are silent. By translation of the thirteen allele sequences (Tab. 3-2), ten different AA sequences are resulting (Tab. 3-4), revealing a lower variation range on AA than on nt level.

Tab. 3-4 The thirteen identified *AHAS1* allele sequences encode for ten different AA sequences. For a better overview, the former listed AA character is represented by a blank space.

Allele sequence no.	**AA position and character**												**AA sequence no.**
	5	**35**	**37**	**52**	**80**	**139**	**310**	**312**	**340**	**365**	**462**	**548**	
	Pro/ Thr	**Leu/ Pro**	**Ala/ Asp**	**Leu/ Val**	**Asp**	**Gly**	**Thr/ Ala**	**Ile**	**Thr**	**Thr**	**Gly/ Ala**	**Leu/ Ile**	
1		Pro											**I**
2	Thr												**II**
3		Pro		Val									**III**
4													**IV**
5												Ile	**V**
6											Ala		**VI**
7		Pro									Ala		**VII**
8		Pro											**I**
9	Thr												**II**
10		Pro											**I**
11			Asp										**VIII**
12			Asp								Ala		**IX**
13		Pro					Ala						**X**

The *AHAS1* sequence of the field mix material was also examined to accurately evaluate the results of the EMS mutagenesis approach. A certain heterogeneity of this material was expected due to lack of selection, and hence it was decided to determine the SNP profile instead of sequencing the target gene of individual plants. Transcriptome data (chapter 3.2.1) collected from four mixed samples of 20 individual plants of the field mix material have been evaluated. Already presented allelic positions (Tab. 3-2) could be retraced, but further, presence of four other silent SNPs could be detected (Tab. 3-5). Three of those were already known from the preliminary sequence studies on samples from the field mix (Tab. 3-1).

Tab. 3-5 Allelic positions and characters of *AHAS1* found in the field mix material from JKI Quedlinburg by analysis of transcriptomic data from 4 x 20 individuals. Additionally, to Tab. 3-2, four further allelic positions could be detected (marked in bold). Despite to Tab. 3-2, at nt positions 110 and 240 occurences of the second allelic character were negligible (2 %), hence only one is listed (grey). The CDS with all SNP positions together in one sequence are given in Annex 6-5.

nt position and character

13	**48**	104	110	154	240	**387**	417	**522**	**762**	928	936	1020	1095	1385	1642
C/A	**T/G**	T/C	C	C/G	C	**T/C**	T/C	**A/G**	**T/C**	A/G	T/C	A/T	T/C	G/C	T/A

AA position and character

5	**16**	35	37	52	80	**129**	139	**174**	**254**	310	312	340	365	462	548
Pro/ Thr	**Pro**	Leu / Pro	Ala	Leu / Val	Asp	**His**	Gly	**Pro**	**Asp**	Thr/ Ala	Ile	Thr	Thr	Gly/ Ala	Leu/ Ile

According to Tab. 3-5 14 allelic positions were detected in the field mix material compared to 12 positions in the other material mentioned (Tab. 3-2). Due to the silent character of the four additional SNPs at nt positions 48, 387, 522 and 762 and the unbiased characters of nt

positions 110 and 240, six different altering AA positions in the field mix material compared to seven in the other were present.

Phylogenetic analysis revealed that among the Asteraceae family up to three AHAS genes (*H. annuus*) occur. In Fig. 3-1 the phylogenetic tree from chapter 1.2.2 is complemented by TkoAHAS1, which is grouping in the clade of *C. intybus* and *S. asper*. Due to ten different identified TkoAHAS1 AA sequences, AA sequence no. I (Tab. 3-4) was chosen representatively for the comparison as it was found the most. Furthermore, the described conserved motifs of AHAS, for example for cofactor binding, could be recovered in the AA sequence of TkoAHAS1 (indicated in Annex 6-4).

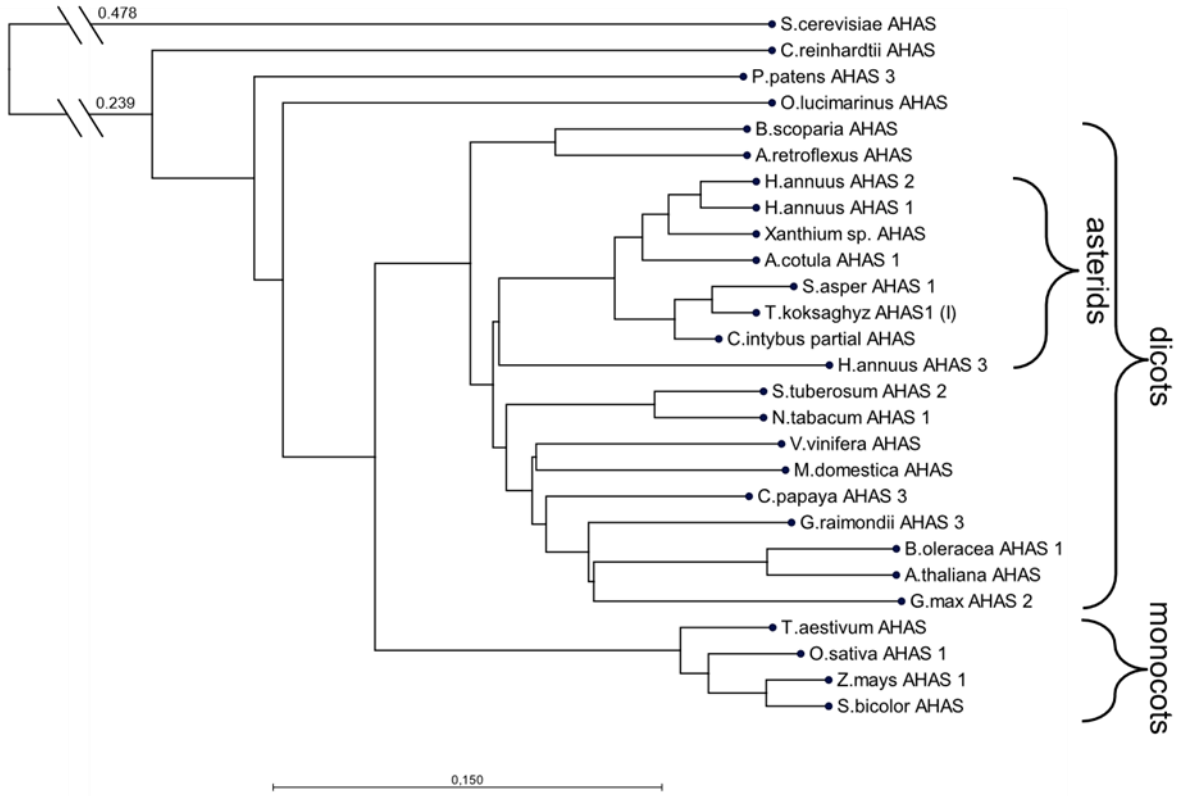

Fig. 3-1 Phylogenetic tree for AHAS of exemplary organisms from monocots, dicots, as well as chlorophytes, bryopsidas, fungi and TkoAHAS1 (AA sequence no. I) (references are listed in Tab. 6-2).

A relatively closely related Asteraceae towards *T. koksaghyz* with adequate literature references concerning the AHAS genes is *H. annuus*. It has three AHAS isozymes which share high (AHAS 1 and AHAS 2) and relatively low homology (AHAS 1 and 2 towards AHAS 3) on gene and protein level (Tab. 3-6). Hence, it was assumed that *T. koksaghyz* may consist of further AHAS genes, too. Due to the identity ratios for *H. annuus* AHAS it was searched for highly homologous genes but also for those with less homologies in *T. koksaghyz*.

Tab. 3-6 Pairwise comparison of the three isozymes AHAS 1, 2, 3 of *H. annuus* on gene and protein level (references: Tab. 6-8). Upper comparison: percent identity; lower comparison: differences (absolute).

Gene level	***AHAS 1* 1965 bp**	***AHAS 2* 1941 bp**	***AHAS 3* 1941 bp**	**Protein level**	**AHAS 1 654 AA**	**AHAS 2 646 AA**	**AHAS 3 646 AA**
AHAS 1	-	92.9 %	73.5 %	**AHAS 1**	-	94.4 %	78.6 %
AHAS 2	140	-	74.0 %	**AHAS 2**	37	-	78.1 %
AHAS 3	523	514	-	**AHAS 3**	141	144	-

A lacking *T. koksaghyz* genome impeded a simple data search at the beginning of this work. Instead, a RACE approach was conducted but with this, only DNA fragments of the known *TkoAHAS1* could be amplified. Unpublished transcriptome data (chapter 2.1.9) from the Prüfer group enabled a more selective and faster search for further AHAS of *T. koksaghyz* (chapter 2.2.14).
Using the transcriptome data, BLAST searches with the known *TkoAHAS1* resulted in retrieving the mentioned sequence on the one hand. On the other hand, fitting contigs with less homologies were found that could also assemble to build a larger fragment. Anyhow, amplification of that fragment out of *T. koksaghyz* genomic DNA failed. A BLAST search using the NCBI database revealed that the assembled fragment belongs to *AthAHAS* because of a contaminated RNA sample.
Due to the *A. thaliana* contamination of that transcriptome data and to verify the existence of only a single AHAS gene, new transcriptome data have been collected in 2016. A leaf and a flower sample of Tks203 were sequenced. These data were free of contaminants and 96.5 % and 97.1 % of the collected data were usable, respectively (Tab. 3-7).

Tab. 3-7 Parameters concerning the collected transcriptome data from two different sources of Tks203.

Parameter	**Transcriptome data set**	
	Leaf	**Flower**
Technology	Genome Sequencer Illumina HiSeq2500	
Read type	Paired end	
Read length	2 x 125 bp	
Pre-processing		
Sequenced bases	8,029,050,000	11,593,900,000
Sequenced reads	32,116,200	46,375,600
Post-processing		
Paired reads	28,834,132	41,405,625
Unpaired reads	2,166,002	3,605,613
Total reads	31,000,134	45,011,238
Percentage of usable reads Post-processing/ Pre-processing	96.5 %	97.1 %

By using the in-house platform Galaxy (*266*) with implemented open source tools (Trinity (*267, 268*) and BLAST (*272, 273*)) analysis of the transcriptome data was performed (2.2.17.1). The processed and assembled transcriptome data served as a reference database for BLAST searches for possible further AHAS transcripts. Those searches were conducted by using several AHAS sequences from other Asteraceae. However, the matching contigs always belonged to the known *TkoAHAS1* sequence; no transcript indicated another *AHAS* in *T. koksaghyz*.
As mentioned above, Tks203 inherits the alleles number 4 and 10 (Tab. 3-2). The alleles could be retraced in both transcriptome data sets (Tab. 3-8) with values between 37 % and 63 %;

the occurrence of SNPs lower than 1 % is negligible and can be explained by sequencing errors.

Tab. 3-8 Occurrence of SNPs in the transcriptome data sets of Tks203 leaf and flower. The percentages of the detected nucleotides at all known twelve *TkoAHAS1* allelic positions are listed. The differing characters between the allele sequences no. 4 and 10 are highlighted in bold.

nt position		**13**	**104**	**110**	**154**	**240**	**417**	**928**	**936**	**1020**	**1095**	**1385**	**1642**
nt character	**Allele sequ. no. 4**	C	**T**	C	C	C	T	A	**T**	**A**	**T**	G	T
	Allele sequ. no. 10	C	**C**	C	C	C	T	A	**C**	**T**	**C**	G	T
nt appearance [%]	**Leaf dataset**	>99	**37**	100	100	100	100	>99	**47**	**42**	**38**	100	100
		<1	**63**	0	0	0	0	<1	**53**	**58**	**62**	0	0
	Flower dataset	100	**55**	100	100	100	>99	100	**52**	**52**	**55**	100	100
		0	**45**	0	0	0	<1	0	**48**	**48**	**45**	0	0

In 2017 a 1.29 Gb draft genome of *T. ko*ksaghyz was published (*66*). Although annotations were not provided, it was intended to verify the identified *TkoAHAS1* sequence within this genome assembly. Furthermore, this dataset was also examined for additional AHAS sequences. A tblastn of the *TkoAHAS1* (SNPs translated as X in the AA sequence) against the genome assembly was conducted. The *TkoAHAS1* could not be recovered completely, but only in parts (Tab. 3-9). Due to this output, no further investigations concerning this data have been conducted.

Tab. 3-9 Overview of the recovered parts of *Tko*AHAS1 in the draft genome from Lin *et al.* (*66*).

Recovered parts of *Tko*AHAS1	**Identities of the fragments**	**Number of fitting fragments**	**Gaps in fitting fragments**	**Missing parts of *Tko*AHAS1**
379-476 513-524 565-591	91.7-97.3 %	6	Without gaps	1-378 477-512 525-564 592-656
352-591	70.8-87.5 %	7	With gaps	1-351 592-656
21-466 519-594	51.2-68.1 %	8	With gaps	1-20 595-656

3.2 Introducing an AHAS-dependent herbicide resistance into *T. koksaghyz*

3.2.1 Undirected mutagenesis by using EMS

The EMS mutagenesis approach in this framework was conducted for three times with 30,000 seeds, each, of *T. koksaghyz* field mix material. Investigations by the Prüfer group estimated a mutation rate by EMS of 1/ 67 kb (*274*) in *T. koksaghyz*. In order to evaluate the mutagenesis results concerning the target gene *TkoAHAS1*, the SNP pattern of *TkoAHAS1* from the non mutagenized *T. koksaghyz* field mix material was determined (chapter 3.1, Tab. 3-5). By this it was feasible to differentiate between already present mutations and probably by EMS induced mutations.

The EMS mutagenesis experimental setup differentiates from the classical TILLING (Targeting Induced Local Lesions IN Genomes) approach (*275, 276*) (Fig. 3-2). Seeds have been EMS treated and the emerging plants (M_1) – instead of the M_2 plants – were immediately treated with imidazolinone (Imazamox) to select for resistant plants. This procedure was chosen since *T. koksaghyz* is an obligate cross-pollinator and therefore no true M_2 could be generated for selection. Surviving plants were screened for mutations in *AHAS1*. Furthermore, they were intercrossed as well as crossed with “wild type” (WT) *T. koksaghyz* (field mix) plants to increase the fitness after mutagenesis. Moreover, this was conducted to stabilize putative EMS induced mutations from the chimeric plants in the subsequent M_2 generation. This approach was considered appropriate since the AHAS conferred herbicide resistance is dominant, thus a heterozygous mutation is sufficient for the plant to survive herbicide treatment (*124*). In contrast to classical TILLING (*275, 276*), identification of mutations was done for each plant individually by *AHAS1* amplification and subsequent Sanger sequencing. Despite the named differences in handling, the approach is further referred to as TILLING in this work since the general procedure is comparable.

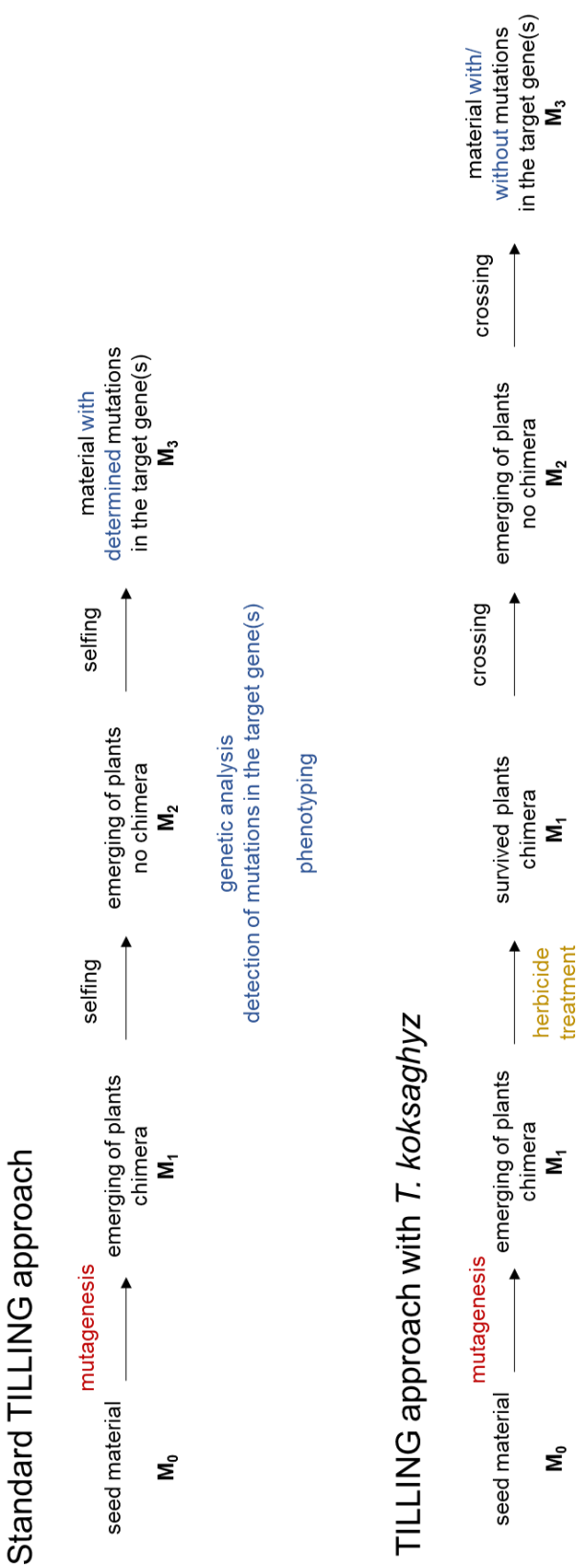

Fig. 3-2 The standard TILLING approach (according to Colbert *et al.* (*275*) and Kurowska *et al.* (*276*)) and the modified TILLING approach used for the work with *T. koksaghyz*.

The outcome of the EMS mutagenesis approach is presented in the following passages. To simplify comprehension, an overview is already given in Fig. 3-3 afore the detailed description of the results.

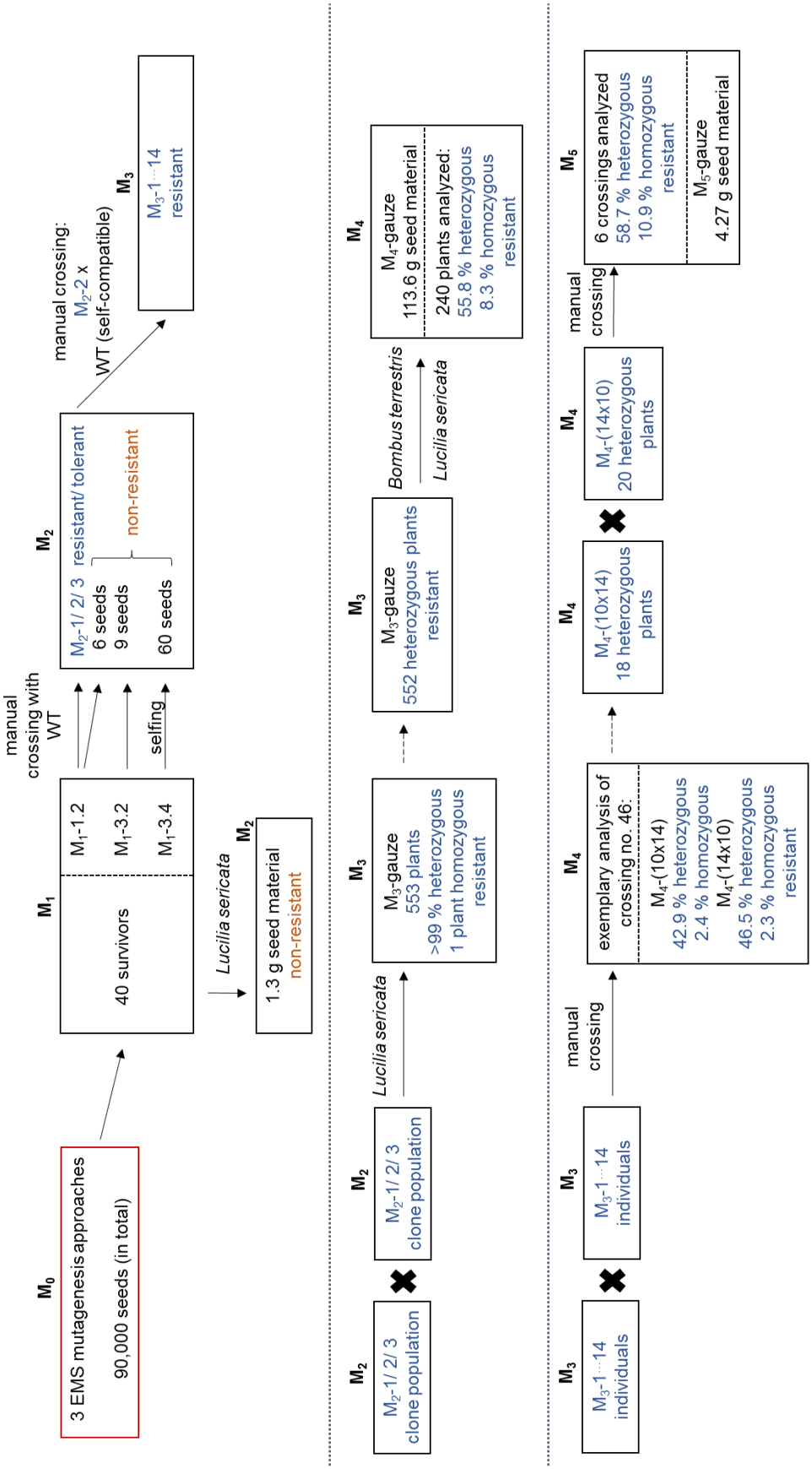

Fig. 3-3 Overview about the EMS mutagenesis approach, the associated crossings and the outcome concerning hetero- and homozygous individuals (referring to nt-position 572 in *AHAS1*). First row: from initial EMS mutagenesis M_0 to M_3 generation. Second row: crossing attempts with the M_2-1/2/3 plants (clone population). Third row: crossing attempts with the M_3-1…14 population.

From the 90,000 treated seeds, 40 M_1 plants were selected that endured (sequential) herbicide treatment(s). Among the M_1 plants, EMS typical adverse effects like leaf colour chimera were observed. It was tried to provoke regrowth yielding a plant with homogeneous, non-chimeric genetic material by cutting these M_1 plants at ground level. However, no mutations could be detected which are known to confer herbicide resistance. Additionally, propagation of those plants was highly impaired: reduced growth and either no flowering or no production of seeds after crossing were mostly observed. Several manual crossing approaches with WT *T. koksaghyz* (selected individuals from field mix material 2014, Thiele group) led to 15 seeds, from which no plant was Imazamox resistant. Crossing of the M_1 plants with each other and with WT by insects (*Lucilia sericata*) lead to 1.3 g of seed material, from which also no plant showed Imazamox resistance. A M_1 individual was able to overcome self-incompatibility and produced 60 seeds by selfing, but none of the progenies was Imazamox resistant.
Eventually, crossing efforts yielded in three plants (M_2-1/ 2/ 3) which survived several Imazamox treatments. Those M_2 plants originate from a crossing of the individual "M_1-1.2" with WT *T. koksaghyz*. Sequencing of *AHAS1* revealed that two M_2 plants were heterozygous for a mutation at nt-position 572 (C to T), which leads to an AA substitution of Ala191 to Val191 (Tab. 3-10). This position is equivalent to Ala205 in *A. thaliana*, from which it is known, that specific AA substitutions (Phe, Val) can confer herbicide resistance (*136, 141*).

Tab. 3-10 Overview of the plants M_2-1/ 2/ 3 with/ without a mutation at nt-position 572 in *AHAS1* and the effect on AA level and assumed Imazamox resistance based on reports (*136, 141*) (see also Tab. 6-4, Tab. 6-5)

Plant	**nt 572**	**AA 191**	**Imazamox resistance**
M_2-1	C/**T**	Ala/**Val**	**Yes**
M_2-2	C/**T**	Ala/**Val**	**Yes**
M_2-3	C/C	Ala/Ala	No

In contrast to the observations with the M_1 plants, these M_2 individuals showed no restraints in growth and fertility. The vitality after treatment with Imazamox was not impaired for two of the three plants. The exception was plant M_2-3: it showed a retardation of growth after herbicide application (two to three weeks post application) but it was able to recover and continue growing after approximately six weeks. This was different from the behaviour of WT *T. koksaghyz*: it displayed a growth stop and a necrosis of the youngest leaves in the inner of the rosette and the plant was not able to recover from the treatment. This behaviour is traceable in Fig. 3-4.

Fig. 3-4 M_2-1/ 2/ 3 and WT *T. koksaghyz* 2.5 weeks after Imazamox treatment. A: three plants of M_2-1; B: one plant of Tks203 (as WT control); C: three plants of M_2-2; D: three plants of M_2-3. Plants were propagated *in vitro* and planted into the green house; after six weeks of acclimatisation, the plants were Imazamox treated.

Crossing of M_2-2 with a selected *T. koksaghyz* field mix individual (seed material from 2015, Thiele group) that was identified as self-compatible yielded in 51 seeds. 14 (M_3-1···14) germinated plants were Imazamox resistant, too. Sanger sequencing of *AHAS1* revealed, that they consist of the same mutation at nt-position 572 as the parent M_2-2.
To produce seed material, M_2-1/ 2/ 3 were taken to *in vitro* cultivation, clones were propagated and subsequently planted into the green house and after acclimatisation planted into a gauze tent. Pollination by insects (*Lucilia sericata*) was successful and progenies ("M_3-gauze") were Imazamox treated after germination. 651 progenies survived the treatment and were screened for the mutation at nt-position 572 by using a restriction enzyme assay (chapter 2.2.13.1). The short procedure of tissue collection, followed by Direct PCR, digestion of the amplicon with the restriction enzyme TaiI and evaluation of the fragment pattern via gel electrophoresis enabled a fast analysis of the progenies. Due to the mutation C572T, a new cleavage site for TaiI was generated in *AHAS1*. Therefore, determination whether individuals are hetero- or homozygous was possible (Fig. 3-5).

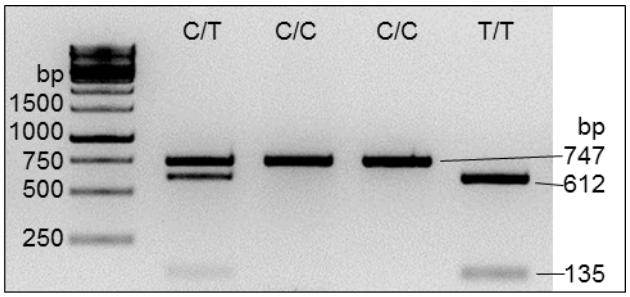

Fig. 3-5 Result of a TaiI screening of EMS progeny plants to determine which nucleotide exists at nt-position 572 in *AHAS1*. Amplicon: 747 bp (from nt 438 to 1185 of *TkoAHAS1*). For WT (C/C) no amplicon is cut by TaiI; for a heterozygous mutant (C/T), 50 % of the amplicons is cut, leading to three fragments of 747, 612 and 135 bp. For a homozygous mutant (T/T), all amplicons are cut by TaiI, leading to two fragments of 612 and 135 bp. DNA ladder: GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific).

552 of the 651 M_3-gauze plants were heterozygous (C/T) at nt-position 572 in *AHAS1*, whilst still 65 plants without a mutation (C/C) were found; despite prior Imazamox treatment. Only one homozygous (T/T) plant ("M_3-gauze-54") was identified that showed normal phenotypical appearance and behaviour (Fig. 3-6).

Fig. 3-6 The solely identified homozygous (T/T) plant M_3-gauze-54 (blue label) from the 651 M_3-gauze plants. On the right, arranged behind the plant is a heterozygous M_3-gauze plant. Both plants show comparable growth.

Due to the selection by Imazamox no conclusion concerning Mendelian splitting in the M_3-gauze population could be drawn. Therefore, the 552 heterozygous M_3-gauze plants were crossed with each other by using insects (*Bombus terrestris*, *Lucilia sericata*). A total of 113.6 g seed material was harvested. A trial with 240 of these M_4 plants, showed roughly Mendelian segregation concerning WT and heterozygous mutants – but not towards the ratio of homozygous mutants (Tab. 3-11).

Tab. 3-11 Segregation of WT (C/C), heterozygous (C/T) and homozygous (T/T) mutants in 240 M_4-gauze plants, progenies of heterozygous M3-gauze plants, concerning nt-position 572 in *AHAS1*.

	Number of plants	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])			
		WT (C/C)	**Heterozygous mutant (C/T)**	**Homozygous mutant (T/T)**	**Dead; n. d.**
M_4 gauze	240	76/ 31.7	134/ 55.8	20/ 8.3	10/ 4.2

Besides the crossings with M_2-1/ 2/ 3, investigations into crossing and segregation behaviour of the plants M_3-1···14 were conducted, too. This was of interest because of the above-mentioned parent that had overcome self-incompatibility. Manual crossings between the 14 siblings were performed. For evaluation of segregation one crossing (no. 46) was selected from which equal seed material was achieved regarding both parents. Like the other M_4 plants (from M_3-gauze population) (Tab. 3-11), this M_4 plants also show roughly Mendelian segregation, whereas the ratios of WT and homozygous mutant do not match, again (Tab. 3-12). The heterozygous mutants and the WT plants were phenotypically similar (besides a few exceptions; see further paragraphs) and showed normal growth and flowering behaviour (Fig. 3-7).

Tab. 3-12 Segregation of WT (C/C), heterozygous (C/T) and homozygous (T/T) mutants in the M_4 progenies of crossing no. 46 (M_3-10 x M_3-14) concerning nt-position 572 in *AHAS1*.

	Amount of M_4 seeds	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])			
		WT (C/C)	**Heterozygous mutant (C/T)**	**Homozygous mutant (T/T)**	**Dead; n. d.**
M_4 [♀ M_3-10 x ♂ M_3-14]	42 seeds	13/ 31.0	18/ 42.9	1/ 2.4	10/ 23.8
M_4 [♀ M_3-14 x ♂ M_3-10]	43 seeds	15/ 34.9	20/ 46.5	1/ 2.3	7/ 16.3

Fig. 3-7 Both, WT and heterozygous progenies of crossing no. 46 in the greenhouse.
A: M_4 of [♀ M_3-14 x ♂ M_3-10], 35 plants; B: M_4 of [♀ M_3-10 x ♂ M_3-14], 29 plants (two heterozygous plants not shown, see Fig. 3-8).

To check if a further propagation within this established M_4 population is possible and to compare the segregation pattern with the ancestral generations, manual crossings were conducted. Propagation was possible but with varying success: analysis of selected crossings

revealed differences concerning recovered seed amount and the distribution of hetero- and homozygous progenies (Tab. 3-13).

Tab. 3-13 Absolute amounts and percentages of WT (C/C), heterozygous (C/T) and homozygous (T/T) mutants concerning nt-position 572 in *AHAS1* of selected M_5 progenies from six crossings within the M_4 population (Tab. 3-12) (details listed in Tab. 6-9).

	Number of M_5 plants	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])		
		WT (C/C)	**Heterozygous mutant (C/T)**	**Homozygous mutant (T/T)**
M_5 [M_4 (10 x 14) x M_4 (14 x 10)]	46	14/ 30.4	27/ 58.7	5/ 10.9

After conducting manual crossings, isolated propagation of population M_4 [♀ M_3-10 x ♂ M_3-14] (18 heterozygous plants) and population M_4 [♀ M_3-14 x ♂ M_3-10] (20 heterozygous plants) was continued in gauze tents by insects (*Lucilia sericata*). From the former, 0.5 g of seeds could be gained, whereas the latter yielded 3.77 g, reflecting different success in reproduction, too. This M_5-gauze population was not further investigated.

In Tab. 3-14 an overview is given about the determined distributions concerning the zygosity in the above presented generations. Whilst the percentages of heterozygous plants were roughly around 50 % according to Mendelian inheritance, the proportions of WT plants were 30 - 35 % (average 32 %) and significantly above the expected value of 25 %. The proportion of homozygous mutant individuals ranges between 2 - 11 % (average 7 %) and is far lower than the expected 25 %. The number of plants which died shortly after germination or did not germinate at all fluctuates widely, too; an average rate of 7 % was determined.

Tab. 3-14 Overview for four populations of the M_4 and M_5 generation (371 individuals) from the EMS mutagenesis approach concerning the segregation of the mutation at nt-position 572 in *AHAS1*. Data relying on Tab. 3-11, Tab. 3-12, Tab. 3-13.

Population	**Number of tested plants**	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])			
		WT (C/C) [%]	Heterozygous mutant (C/T) [%]	Homozygous mutant (T/T) [%]	Dead; n. d. [%]
M_4 gauze	240	32	56	8	4
M_4 [♀ M_3-10 x ♂ M_3-14]	42	31	43	2	24
M_4 [♀ M_3-14 x ♂ M_3-10]	43	35	47	2	16
M_5 [M_4 (10 x 14) x M_4 (14 x 10)]	46	30	59	11	0
TOTAL	371	32	54	7	7

Concerning the identified homozygous mutants from M_3-gauze plants, related M_4 plants as well as from M_4 [♀ M_3-10 x ♂ M_3-14], M_4 [♀ M_3-14 x ♂ M_3-10] and the related M_5-gauze population, different phenotypes were identified.
The above-mentioned homozygous mutant M_3-gauze-54 not solely showed a healthy phenotype with unimpeded flowering habit. From the 240 tested M_4-gauze-plants, 20 homozygous mutants were identified; of these four developed and grew normally (Fig. 3-8). The other showed a kind of dwarf phenotype (comparable to individuals in Fig. 3-9): they developed eight to ten very small leaves, did not flower and died eventually.

Similar behaviour was observed for the progenies of crossing no. 46: plants no. M_4-(10x14)-31 and M_4-(14x10)-36 were the only identified homozygous mutants and exhibited the described dwarf phenotype (Fig. 3-9). Certainly, from this population also two heterozygous mutants with dwarf phenotype could be identified: plants no. M_4-(10x14)-30 and -32 (Fig. 3-9).

Fig. 3-8 Comparison of heterozygous and homozygous mutants from the M_4-gauze plants. A-D: plant on the left side is no. M_4-gauze-14 (heterozygous), plant on the right side is a homozygous mutant: A: no. M_4-gauze-19; B: no. M_4-gauze-20; C: no. M_4-gauze-24; D: no. M_4-gauze-40. Pot sizes: small: 8 x 8 cm; big: 10 x 10 cm.

Fig. 3-9 Plants from crossing no. 46 (M_3-10 x M_3-14) showing the dwarf phenotype. A-B: heterozygous mutants; C-D: homozygous mutants. A: M_4-(10x14)-32; B: M_4-(10x14)-30; C: M_4-(10x14)-31; D: M_4-(10x14)-36. Pot size: 8 x 8 cm.

Despite the above-mentioned vital homozygous mutants from M_3-gauze and M_4-gauze populations, also in the M_5 progeny from [M_4-(10x14) x M_4-(14x10)] a single homozygous mutant was identified that showed no impairment in growth and flowering habit (Fig. 3-10).

Fig. 3-10 Comparison of a heterozygous and a homozygous mutant. Plant on the left side: no. M_4-gauze-14 (heterozygous). Plant on the right side: M_5 progeny from [M_4-(10x14) x M_4-(14x10)] (plant no. 3 of crossing no. 59 B) (Tab. 3-13, Tab. 6-9). Pot sizes: small: 8 x 8 cm; big: 10 x 10 cm.

In summary, based on the TILLING approach described here, crossing experiments and analyses of the progeny from the M_1 to the M_4 and M_5 populations were successfully performed.

3.2.2 Site-directed genome editing by using the CRISPR/Cas9 system

In chapter 1.2.3 different AA are described, whose substitutions can confer resistance towards imidazolinones and account to binding of the herbicide in AHAS. Accordingly, six prominent positions in *TkoAHAS1* were chosen to potentially target with the CRISPR/Cas9 system. Ahead of *A. tumefaciens* mediated stable transformation of the system into *T. koksaghyz*, an *in vitro* cleavage assay according to Jinek *et al.* (*263*) was conducted to check the designed sgRNAs for reliable performance (chapter 2.2.7). Cleavage of the DNA target after efficient binding of the sgRNA and Cas9 nuclease is simply retraceable by electrophoresis (Fig. 3-11).

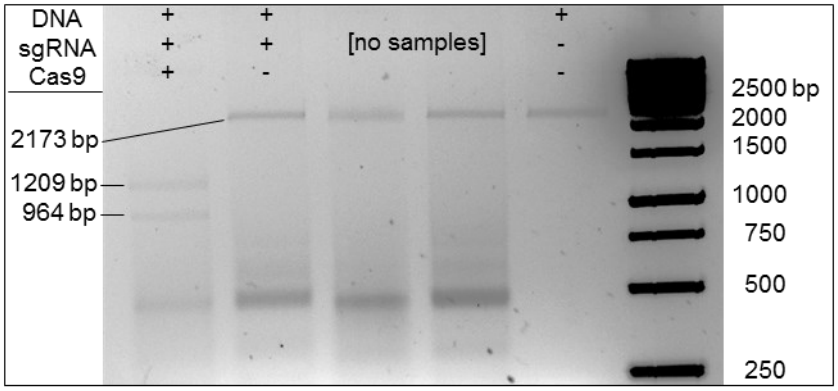

Fig. 3-11 Exemplary results of an *in vitro* cleavage assay (gel electrophoresis). Target DNA: 2173 nt (from nt -119 (5'UTR region) to +83 (3'UTR region) of *TkoAHAS1*); sgRNA designed, so Cas9 is cutting after nt position 1091; fragments after cleavage: 1209 bp and 964 bp. DNA ladder: GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific).

The *in vitro* cleavage assay was conducted with all six sgRNAs for the different target sites in *TkoAHAS1*. Efficient cleavage of target DNA could be verified in four cases (Tab. 3-15).

Tab. 3-15 *In vitro* cleavage assay results giving information about reliable functioning of designed sgRNAs for six different prominent target sites.

Target AA TkoAHAS1	**Referring AA AthAHAS1**	**Referring nt position *TkoAHAS1***	**DSB by Cas9 after nt …**	**Result *in vitro* assay/ functionality of sgRNA**
Ala108	Ala122	322-324	325	Negative
Pro183	Pro197	547-549	546	Negative
Ala191	Ala205	571-573	574	Positive
Arg363	Arg377	1087-1089	1091	Positive
Trp560	Trp574	1678-1680	1674	Positive
Ala639	Ser653	1915-1917	1916	Positive

Two positions were selected to target *in vivo* by stable transformation of the CRISPR/Cas9 system into *T. koksaghyz*: Arg363 and Ala639. Thus, a position in the middle and one at the 3' end of the gene and C-terminal end of the protein, respectively, were to be investigated. Binary vectors were cloned that encoded for a sgRNA, the Cas9 nuclease and two resistance cassettes for selection in bacteria ($Spec^R$) and plants (PPT^R). No template sequence for HR was donated, choosing the way of NHEJ for DSB repair. *A. tumefaciens* transformation was conducted with leaf explants likewise to Wahler *et al.* (*14*) (Fig. 3-12).

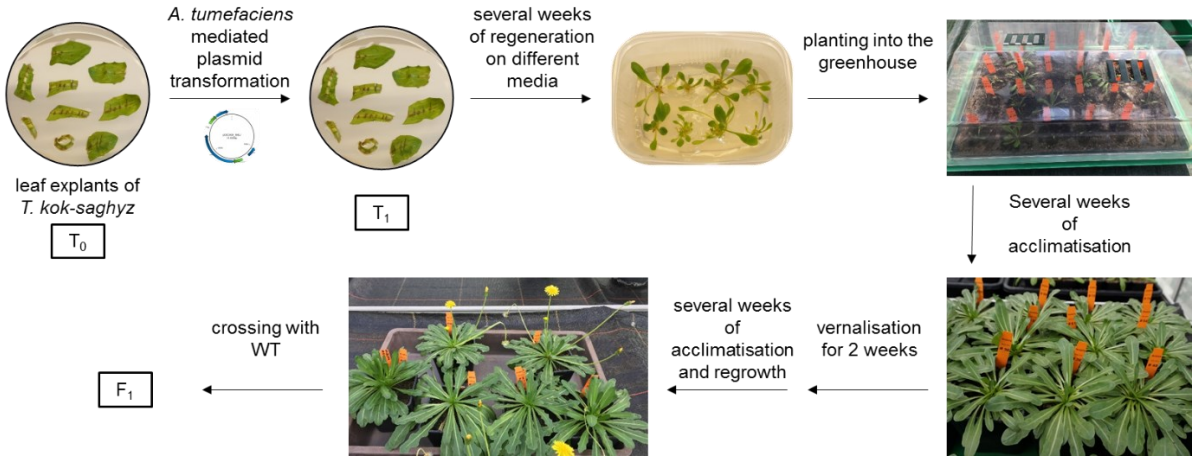

Fig. 3-12 Methodical approach of *A. tumefaciens* mediated stable transformation of the CRISPR/Cas9 system into *T. koksaghyz* and the subsequently efforts for generation of progenies.

Several transformation approaches were conducted. Problems with low survival rates of the explants and overgrowth by the Agrobacteria occurred. Finally, 30 regenerated plants were received by transformation with the vector for targeting nt-position 1091 (corresponding to Arg363). The presence of T-DNA was verified for all regenerates, but no sequence changes in *TkoAHAS1* could be detected in the T_1 plants. In contrast, 88 plants were recovered after transformation with the vector for targeting nt-position 1916 (corresponding to Ala639). T-DNA was present in all regenerated individuals, too. Sequence changes in *TkoAHAS1* occurred to 18.2 % (16 plants); 81.8 % of the regenerates showed no changes. Due to the chimeric characteristic of the T_1 regenerates, several different sequence changes, even in one regenerate could be identified (Tab. 3-16).

Tab. 3-16 Identified sequence changes in 16 T_1 plants, stably transformed with the CRISPR/Cas9 system targeting nt-position 1916 in *AHAS1*. Trace data is given in Fig. 6-7.

Detected changes in *AHAS1*	**Frequency in 16 plants**
In 1916 **t**	14
In 1916 **a**	6
In 1916 **g**	4
DEL 1910-1916 (7 nt)	6
In 1916 **a** & Subst 1946 **g**	1
In 1916 **t** & Subst. 1914 **t**	1

The CRISPR/Cas9 system is working in a stably transformed plant as long as the protospacer and the PAM sites are not modified. It is of interest to achieve progenies of the transformed plants, which do not contain the T-DNA and/ or consist of *AHAS1* changes that impede further effects by CRISPR/Cas9. As long as the T-DNA is inherent and at least one allele of *AHAS1* is vulnerable by CRISPR/Cas9, a plant is still chimeric concerning the targeted gene regions. The chimera status is not allowing to draw conclusions concerning effects on gene/ protein level towards herbicidal resistance. For this reason, all 88 T_1 plants were ought to be used for propagation. Finally, after adaption and vernalisation in the greenhouse, 27 T_1 regenerates were usable, therefore. Crossings were conducted with WT *T. koksaghyz* (field mix) or between the T_1 plants. 54 crossings were performed manually, from which 19 were successful, leading to a total of 24 harvested seed heads. Noteworthy, a seed head was often developing only on one crossing partner. A selection of 13 crossings (17 seed heads, 315 seeds) was chosen for planting in the greenhouse, resulting in 271 F_1 plants. A restriction enzyme screening method (chapter 2.2.13.2) was used for fast and efficient search for individuals with

AHAS1 mutations at/ around nt-position 1916; similar to the restriction assay presented in chapter 3.2.1. A tissue sample was collected, followed by direct PCR and the amplicon was digested with the restriction enzyme PdiI. As long as the WT *AHAS1* sequence was present, the enzyme was able to cut and the fragments were retraceable via gel electrophoresis. The digest by PdiI was not complete, but the ratio of the fragments indicated, whether WT and/ or a changed *TkoAHAS1* sequences were existent (Fig. 3-13).

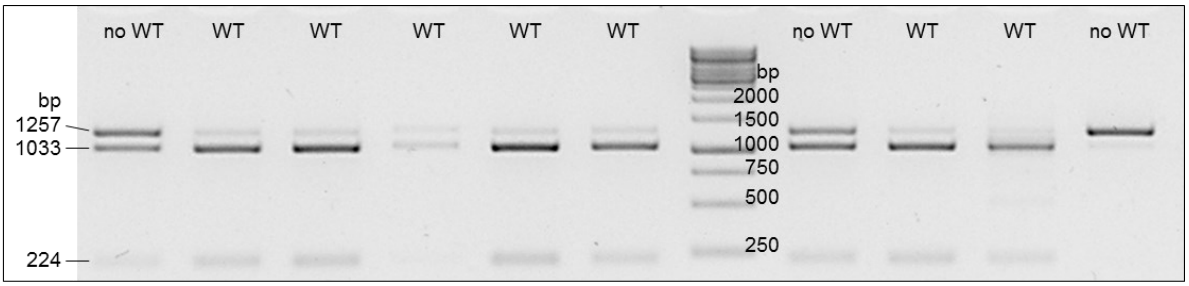

Fig. 3-13 Exemplary results of a PdiI screening of F_1 plants to determine whether mutations at nt-position 1916 in *AHAS1* exist or not. Amplicon: 1257 bp (from nt 885 to +170 (3'UTR region) of *TkoAHAS1*). The amplicons of unchanged *AHAS1* are cut by PdiI into two fragments of 1033 bp and 224 bp length. Amplicons of changed *AHAS1* are not cut. The ratio of cut and uncut amplicons indicates if WT sequence exists or not.

From the 271 tested F_1 plants, 228 (84.1 %) comprised only unchanged *AHAS1* sequences and 35 (12.9 %) revealed to have mutations (Tab. 3-17). Among these 35 plants, six individuals were detected which did not contain the T-DNA, anymore.

Tab. 3-17 Results of the PdiI screening (chapter 2.2.13.2) of 271 F_1 plants for *AHAS1* mutations at nt-position 1916 (corresponding to Ala639).

Number of crossings T_1 x WT	**F_1 progeny**			
	Number of seeds	**Number of vital plants**	**Number of vital plants with mutations in *AHAS1***	**Dead; n.d.**
13	315	271	35	44

Of the 13 crosses examined (Tab. 3-17), only seven were causative of the 35 progeny with mutations in *AHAS1*. To determine the mutations, Sanger sequencing of *AHAS1* was conducted for a selection of 31 of these 35 F_1 plants, recovering a variety of sequence changes (Tab. 3-18). The numbering and designations from the table are explained by Fig. 3-14 where the referring sequence changes on nt and AA level are shown in detail. It occurred that F_1 plants had *AHAS1* sequence changes although the ancestor T_1 did not and even differing sequence mutations between T_1 regenerates and F_1 progeny were detected. Besides sequence mutations, WT sequence was also recovered in most plants, indicating the chimeric status is also occurring in F_1 plants. As explained above, no conclusions about herbicide resistance related to the changes in *TkoAHAS1* can be drawn for these chimeras.

Among the 31 analysed individuals, four plants did not contain T-DNA, anymore. These four progeny originated from three different crosses of the same T_1 parent (as father) with WT (as mother) and all exhibited the same type of mutation: an insertion of A at position 1916 (designation number (1)) (Tab. 3-18).

Tab. 3-18 Results of the Sanger sequencing of selected 31 F_1-plants from five crossing events concerning *AHAS1* mutations at nt-position 1916 (Ala639). In = insertion; Del = deletion; InDel = insertion plus deletion; Subst&In = substitution plus insertion. The asterisk marks where the four individuals without T-DNA belong to.

T_1 (♂)	**F_1**		
Detected mutations in *AHAS1*	**Number of analysed crossings**	**Amount of analysed plants**	**(Number) Designation for detected mutation in *AHAS1***
None	1	4	(1) In 1916 **a** (2) In 1916 **t** (5) Subst&In 1914 **a**C**a**
In 1916 **a** In 1916 **t** In 1916 **g**	3	24	(1) In 1916 **a** * (2) In 1916 **t** (3) In 1916 **c** (4) In 1916 **g** (6) InDel 1917 **t** 1919 (7) Del 1914-1917 (8) Del1913-1917
In 1916 **a**	1	3	(1) In 1916 **a** (4) In 1916 **g**

The *TkoAHAS1* mutations that have been detected among the 31 F_1 plants (Tab. 3-18) describe in total eight different kinds of insertions, deletions and InDels (Fig. 3-14).

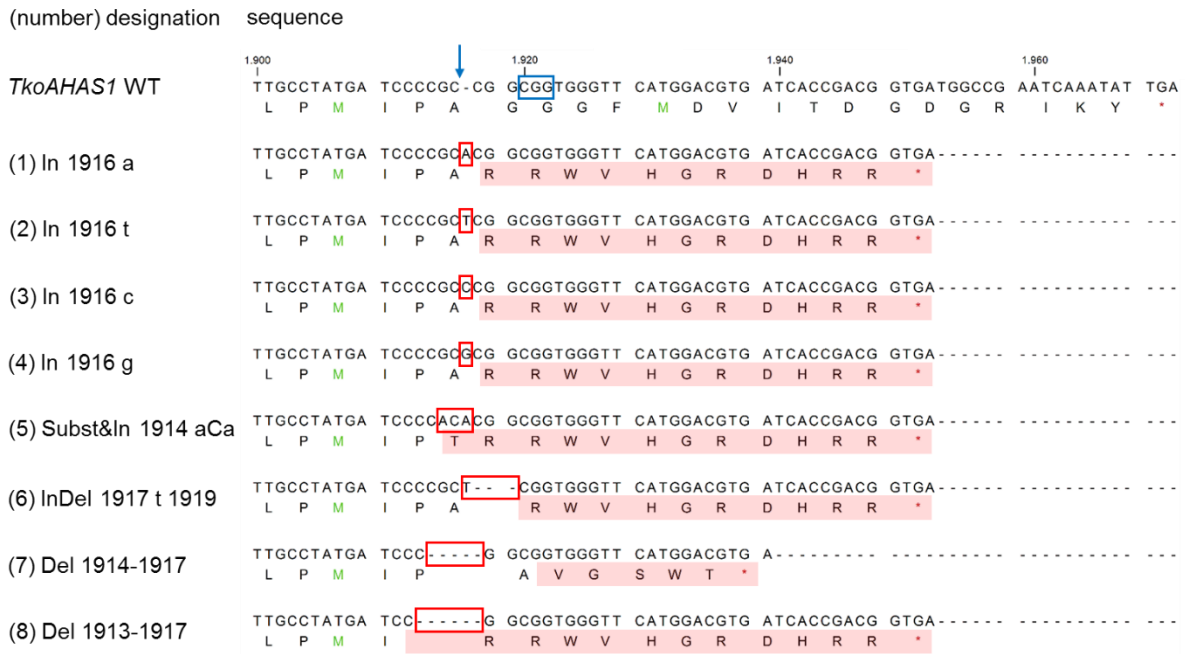

Fig. 3-14 Sequence changes on nt-level and corresponding changes on AA level of the eight detected mutation types in the F_1 progenies in comparison to the WT *AHAS1* sequence. The blue box marks the PAM, the blue arrow depicts the site of DSB induction by Cas9, red boxes indicate the changes on nt-level and red stripes highlight the changes on AA level. Trace data is given in Fig. 6-8.

In all cases the mutation(s) number (1) to (8) lead to a change of the AA sequence and due to a generated premature STOP-codon also to a truncation of the protein. The insertion types (1) to (4) all result in the same AA output. The mutations (5), (6) and (8) are resulting in nearly the same AA sequence, despite one or two residues, respectively. Deletion number (7) is the exception that encodes for an even more truncated protein than the other mutations. By the

discovered eight different mutations, the targeted AA Ala639 was either not changed (six variants, (1) – (4), (6), (7)) or substituted by Thr (5) or Arg (8). The adjacent Gly640 is described in literature, too, as a resistance conferring position. Here, it is mostly substituted by Arg (six variants, number (1)-(6)) and once each by Val (7) and Trp (8), respectively.

Further crossings with these F_1 plants were successful but have not been further investigated in this framework.

3.3 Transcriptome analysis of EMS mutagenized material

The plants M_2-1/ 2/ 3 originated from one crossing and were able to survive Imazamox treatment (chapter 3.2.1). Two behaved similarly (M_2-1/ 2) and showed strong resistance most likely due to the mutation at nt position 572. M_2-3 was coping with a treatment differently but was still not behaving like the WT *T. koksaghyz* (Fig. 3-3). To determine if this behaviour is explainable by a metabolic tolerance, transcriptome analysis was conducted to detect differences between these relatives.

The three plants were transferred from the greenhouse to the *in vitro* cultivation for clonal propagation. Clones were planted back into the greenhouse and six weeks later treated with Imazamox or water as control. Besides the plants of interest M_2-1/ 2/ 3, another plant source as a control was needed. The parents of the M_2 plants were not available anymore, but since they originate from the field mix material, this was used as a negative control instead of Tks203. Imazamox treatment was conducted simultaneously and samples were collected 24 hours later (according to Balabanova *et al.* (*269*)). An overview about the experimental setup is given in Tab. 3-19.

Tab. 3-19 Experimental setup for investigating metabolic differences between M_2-1/ 2/ 3. Plants were planted from the *in vitro* cultivation into the greenhouse and treated with Imazamox (Pulsar®40) or water (as control) six weeks afterwards; sample collection 24 hours after treatment.

Plant	Treatment	Samples	Collected plant(s)/ sample
M_2-1	Imazamox	2	1
M_2-2		2	1
M_2-3		2	1
Field mix		2	20
M_2-1	Water (=control)	2	1
M_2-2		2	1
M_2-3		2	1
Field mix		2	20

Transcriptome data was generated for the described 16 samples with the parameters listed in Tab. 3-20. The RNAseq raw datasets were processed (chapter 2.2.17.2) comparable to the data processing presented in chapter 3.1; but instead of building an assembly for every dataset, data was collected and a common transcriptome was assembled. With the TransDecoder programme (*267*), coding regions were identified and functionally annotated by using Blast2Go. Furthermore, a comparison of differentially expressed transcripts between the samples was conducted using the R package DeSeq (*277*). With these processing steps, it was possible to compare the samples between each other and to name differentially expressed transcripts.

Tab. 3-20 Parameters of the acquired transcriptome data. Detailed results: Tab. 6-7.

Parameter	Transcriptome data
Technology	Genome Sequencer Illumina NovaSeq 6000
Read type	Paired end
Read length	2 x 101 bp
Read quantity	30 – 47 million per sample

The result of a principal component analysis (PCA) for the sixteen datasets of Imazamox treated versus untreated samples is represented in the plot Fig. 3-15. The first principal

component (PC) explains 64 % of the variance and separates M_2-1/ 2/ 3 from the field mix samples. The second PC can only explain 14 % of the variance, separating M_2-2 from M_2-1 and M_2-3. Overall, all samples for each plant, whether herbicide treated or not, are lying very close together but are still apparently different.

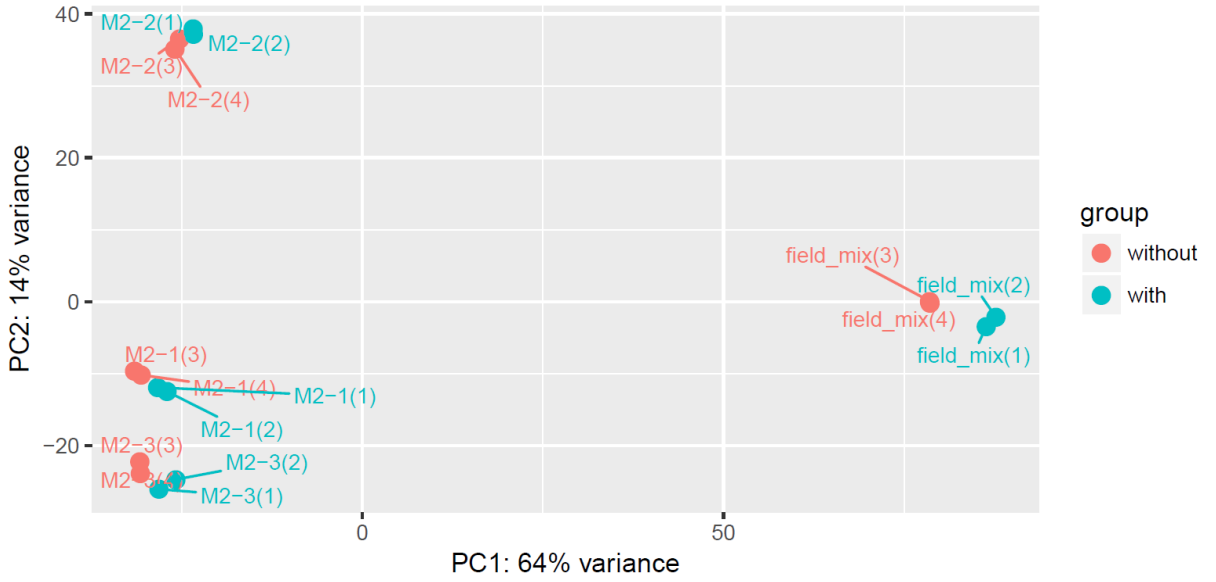

Fig. 3-15 PCA plot for the 16 datasets (Tab. 3-19) to compare Imazamox treated ("with") and untreated ("without") samples. The numbers in parenthesis refer to the serial numbers used for analysis (e.g. M2-1(1) refers to one of the four datasets of M_2-1, in this case the treated sample with Imazamox).

By selecting only the data of M_2-1/ 2/ 3 for PCA plot analysis, a more detailed discrimination is possible (Fig. 3-16). The first and the second PCs explain 44 % and 41 % of the variance, respectively.

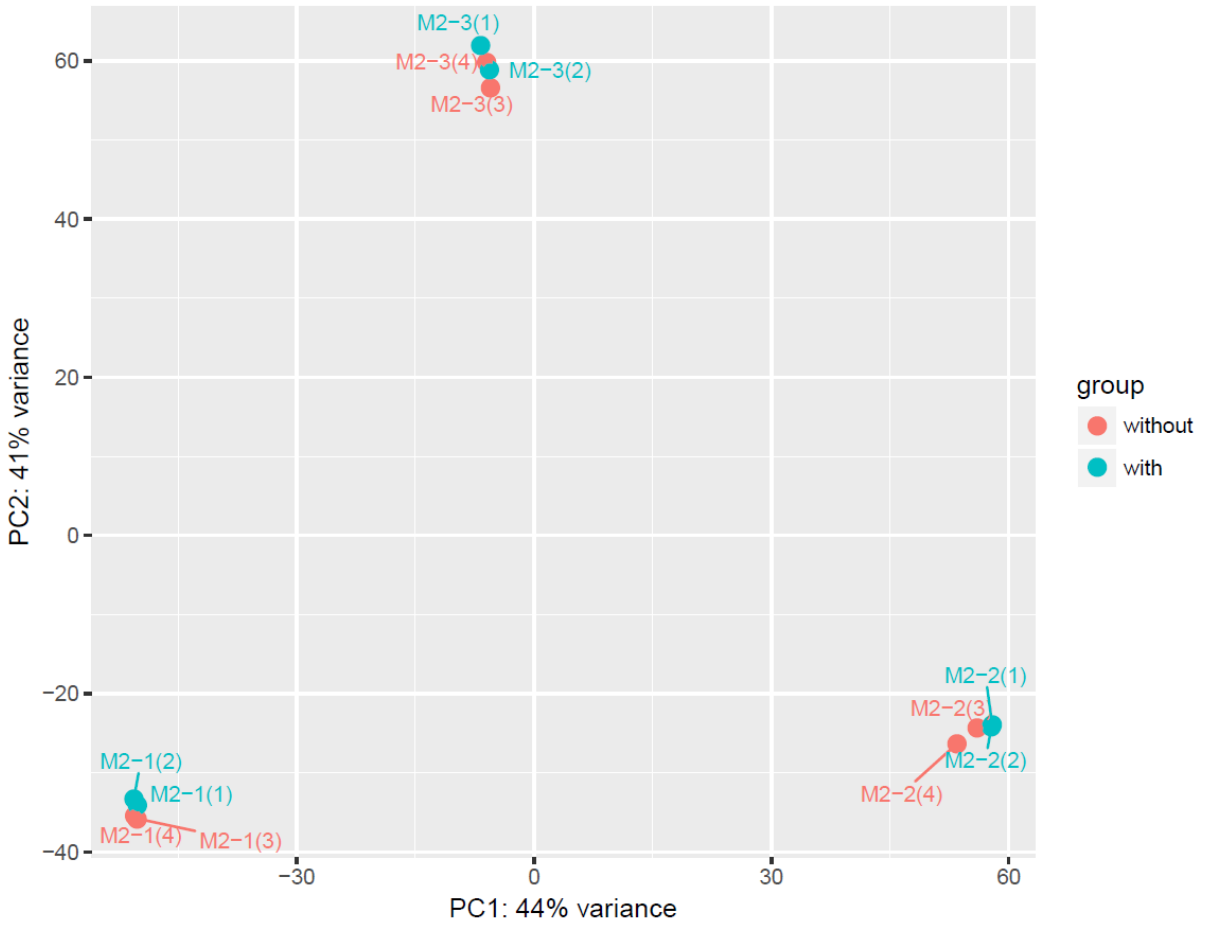

Fig. 3-16 PCA plot for twelve of the transcriptome datasets (Tab. 3-19) to compare Imazamox treated ("with") and untreated ("without") samples of M_2-1/ 2/ 3. The number in parenthesis refer to the numbering used for analysis (e.g. M2-1(1) refers to one of the four datasets of M_2-1, in this case the treated sample with Imazamox).

Due to the mentioned high variance of the data from M_2-1/ 2/ 3 compared to the field mix (Fig. 3-15), the comparison of the RNASeq data concerning differential expressions was conducted between M_2-1/ 2/ 3 alone. The expression values received by running DeSeq were compared between each other: treated samples in relation to the untreated samples for each plant. The resulting quantity of transcripts reflects only those, which are up- or down-regulated in the treated plants. In Fig. 3-17 the associated Venn diagram of M_2-1/ 2/ 3 describes these amounts of differentially expressed (DE) transcripts. 755 transcripts have been identified by these constraints. Intersections between the plants describe shared transcripts, regardless of up- or down-regulation. 214 transcripts in total are shared somehow, which is 28.3 % – the remaining 71.7 % represent those transcripts that are differentially expressed only individually. All three plants share 47 differentially expressed transcripts (6.2 %). The Imazamox resistant plants M_2-1 and M_2-2 share additional 77 differentially expressed transcripts.

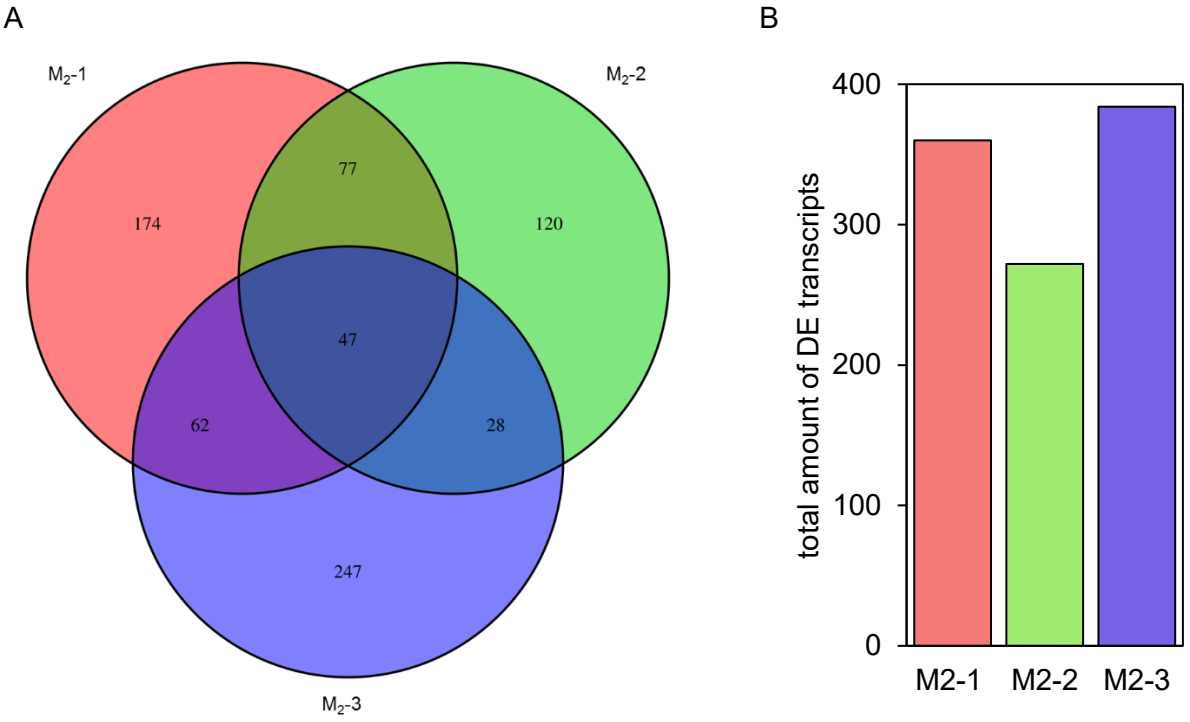

Fig. 3-17 Overview for the DE transcripts of M_2-1/ 2/ 3. A) Venn diagram; each circle represents the amount of DE transcripts of the Imazamox treated samples in relation to the untreated samples. B) Bar chart: the total amount of DE transcripts for M_2-1/ 2/ 3.

For 180 of the 755 DE transcripts, no functional annotation was assigned by using Blast2Go. In the remaining 575 out of the 755 DE transcripts, no transcript was found that matched an AHAS annotation. However, an AHAS transcript could be recovered in the assembled transcriptome, but it is not differentially expressed in any variant and hence not included/ detected in the named DE transcripts.

Investigating potential detoxification genes, the focus was on those which contribute to phase I, II and III of xenobiotic metabolism as those are well described in contrast to those of phase IV. The 575 functional annotated transcripts were selected for the following annotations: P450 monooxygenases (according to Breccia *et al.* (*278*)), glutathione-*S*-transferases and -reductases (according to Balabanova *et al.* (*269*)) as well as nitroreductases, deaminases, peroxidases, esterases (according to Hoagland et al. (*148*)) and ABC transporter (according to Yuan *et al.* (*158*)). Rojano-Delgado *et al.* described the occurrence of two metabolites by phytodegradation of Imazamox in wheat cultivars: a hydroxylated and glucose-conjugated metabolite (*165*). Former one was found in susceptible plants and latter one additionally in resistant plants. Oxidative hydroxylation and carbohydrate conjugation is appearing (*165*), therefore the DE transcript annotations were also screened for glycosyl- and glucosyltransferases. The amounts of matching transcripts/ gene IDs are listed in Tab. 3-21.

Tab. 3-21 Amounts of matching gene IDs from the common transcriptome assembly relying on the Venn diagram in Fig. 3-17 towards different detoxification enzyme types (annotation).

Annotation	**Amounts of matching gene IDs**			
	In the whole Venn diagram	Shared by M_2-1 and M_2-2	Only occurring in M_2-3	Shared by all
P450 monooxygenase	**35**	7	6	5
Glutathione-*S*-transferase	**17**	3	4	3

Annotation	Amounts of matching gene IDs			
	In the whole Venn diagram	Shared by M_2-1 and M_2-2	Only occurring in M_2-3	Shared by all
Glutathione-*S*-reductase	**-**	-	-	-
Nitroreductase	**-**	-	-	-
Deaminase	**-**	-	-	-
Peroxidase	**9**	-	7	2
Esterase	**-**	-	-	-
ABC transporter	**12**	-	2	-
Glycosyltransferase	**7**	-	2	-
Glucosyltransferase	**24**	2	4	4
In total	**104**	**12**	**25**	**14**

Gene IDs from the DESeq matched for six of the ten search terms. P450 monooxygenases and glucosyltransferases are prevailing. Less hits were detected for glutathione-S-transferases, ABC transporter, peroxidases and glycosyltransferases.

The shared DE transcripts from the samples with herbicide resistance (M_2-1 and M_2-2) only include those of three different detoxification enzyme types. The hits shared by all samples only include four different enzyme types. Despite, for the putative metabolic tolerant M_2-3 more hits and a wider range of DE transcripts matching detoxification enzymes were identified, which are uniquely differentially expressed in that sample. Comparing the results overall, there is no clear tendency for DE transcripts of one detoxification enzyme type to occur mainly in the putative tolerant, resistant or all plants; however, peroxidases were mainly present in M_2-3.

For 21 of the identified 25 DE transcripts in M_2-3 an upregulation was detected (Tab. 6-10).

SNP calling is a general approach to find SNPs that are potentially involved in an altered phenotype. Therefore, SNP calling for the above listed 25 DE transcripts found in M_2-3 was conducted. It was of interest to determine SNPs that did not occur in M_2-1/ 2 but in M_2-3, implicating potential impact in detoxification and herbicide tolerance mechanisms.

For the SNP call all datasets – also of the field mix – were considered. After filtering for SNP quality, a total of 553 SNPs in the 25 transcripts could be identified. SNPs were further evaluated using the genotype (GT) character which is a common bioinformatic procedure for evaluation of SNPs in diploid plants. The 553 identified SNPs were checked for alleles, only occurring in M_2-3 but not in M_2-1/ 2 leading to a number of 32 SNPs in ten of the 25 transcripts. For identifying the non-silent SNPs, every transcript was blasted (blastx against land plants) and the match with the highest identity was used as a reference for evaluating the SNP effects on AA level. With the reference protein sequences it was possible to identify the probable reading frames of the transcripts. Almost half of the SNPs turned out as silent mutations, while 17 SNPs in eight transcripts lead to altered AA compositions (Tab. 6-11).

The question arose whether these SNPs are to be justified by EMS mutagenesis or whether they were already present in the starting material. For this purpose, the SNP data of the field mix had to be consulted. However, the GT produced by the programme MPileup (used for the SNP call) is mathematically based on a single plant sample of a diploid organism and, hence, it is not a meaningful quantity for the data of the field mix. The intention of the increased number of plants in the field mix (bulk sample) was to obtain a comprehensive spectrum of naturally existing SNPs. Therefore, the SNPs from the field mix samples were not evaluated by GT, but by continuous genotype (cg), which describes the probability to find the corresponding interesting SNP from M_2-3 in the field mix material. Thus, in Tab. 3-22 an overview is provided, listing the 17 non-silent SNPs detected in M_2-3 in relation to the cg of the field mix data and the GTs of M_2-1/ 2/ 3. Although the cg is an appropriate parameter for SNP evaluation, it should

be noted that the cg values only suggest trends with regard to the potential relevance of SNPs. This is due on the one hand to the differences in the number of plants per sample and on the other hand it cannot be assumed that there are balanced transcription rates between alleles. The determined cg values vary strongly (between 0.51 and 80 %). SNPs which were introduced during mutagenesis and possibly contributing to metabolic tolerance should have lower cg values than SNPs naturally occurring in the population. Four SNPs with cg values below 1 % stand out.

Tab. 3-22 List of non-silent SNPs that only exist in M_2-3 and not in M_2-1 and M_2-2 (based on GT). For each SNP the transcript (Gene ID) and SNP position are listed with the referring cg value of the field mix and the GTs of M_2-3/ 2/ 1. For each transcript the SNPs are sorted after cg value. Cg values lower than 1 % are highlighted in bold. (GT: allele 1/allele 2, with 0 = reference nucleotide, 1 = alternative nucleotide). The reference proteins according to Tab. 6-11 are listed.

Transcript (Gene ID)	Reference protein	SNP position	cg Field mix	GT M_2-3	GT M_2-2	GT M_2-1
Gene.111970:: 145060::g.111970 ::m.111970.exon1	cytochrome P450 CYP72A219-like [Lactuca sativa]	28	2.75%	0/1	0/0	0/0
Gene.57283:: 74240::g.57283 ::m.57283.exon1	cytochrome P450 CYP72A219-like isoform X1 [Lactuca sativa]	349	**0.51%**	0/1	0/0	0/0
		343	**0.53%**	0/1	0/0	0/0
		164	**0.79%**	0/1	0/0	0/0
		266	4.44%	0/1	0/0	0/0
		186	20.76%	0/1	0/0	0/0
Gene.57295:: 74262::g.57295 ::m.57295.exon1	cytochrome P450 CYP72A219-like isoform X1 [Lactuca sativa]	64	**0.88%**	0/1	0/0	0/0
		353	1.17%	0/1	0/0	0/0
Gene.5553:: 6790::g.5553 ::m. 5553.exon1	peroxidase N1-like [Lactuca sativa]	827	31.10%	0/1	1/1	1/1
		830	31.21%	0/1	1/1	1/1
		831	31.40%	0/1	1/1	1/1
Gene.80226:: 103690::g.80226 ::m. 80226.exon1	peroxidase 42 [Lactuca sativa]	1008	70.59%	0/1	1/1	1/1
Gene.135775:: 175077::g.135775 ::m. 135775.exon1	UDP-glycosyltransferase 71E1-like [Lactuca sativa]	1611	16.83%	0/1	0/0	0/0
		1191	21.21%	0/1	0/0	0/0
		641	80.00%	0/0	1/1	1/1
Gene.159476:: 206896::g.159476 ::m. 159476.exon1	UDP-glycosyltransferase 83A1-like [Lactuca sativa]	621	38.73%	0/1	1/1	1/1
Gene.59442:: 76796::g.59442 ::m.59442.exon1	UDP-glycosyltransferase 73C5-like [Lactuca sativa]	99	11.02%	0/1	0/0	0/0

The four SNPs with cg values below 1 % belong to two different DE transcripts which share the same reference protein: cytochrome P450 CYP72A219-like isoform X1 from *Lactuca sativa* (XP_023736509.1). CYP72A219 is already described being upregulated in tolerant versus susceptible individuals of *Myosoton aquaticum* towards tribenuron-methyl (ALS inhibition) (*279*) but also in *Amaranthus palmeri* towards glufosinate (glutamine synthetase inhibition)

(*280*) and is therefore assumed to contribute to plant herbicide tolerance mechanisms. By aligning the two transcripts (Fig. 6-9), an identity of 87 % for the congruous region (Fig. 6-10) was determined. The SNP at position 164 from transcript Gene.57283[…] and the SNP at position 64 from transcript Gene.57295[…] describe the same position also regarding the reference protein (Tab. 6-12, Fig. 6-11)). This reduces the count of interesting SNPs to three (Tab. 6-13). Assuming that these three SNPs in *T. koksaghyz* actually belong to one CYP gene, the combination of SNPs would describe a specific haplotype. However, this haplotype is also retraceable in the field mix, verifying the existence in the material already before the mutagenesis – yet to a very small proportion.

4 Discussion

4.1 Acetohydroxyacid synthase gene(s) in *T. koksaghyz*

Several methods were employed to reveal the AHAS gene(s) of *T. koksaghyz*. In addition to DNA-based applications, RNA data from external and internal approaches were considered. Still, all results led to the conclusion, of *T. koksaghyz* only coding for one single intron-free AHAS gene of 1971 bp, named here *TkoAHAS1*. Although the plant is diploid and therefore two alleles of *TkoAHAS1* exist per individual, a wide range of different allele sequences has been elucidated. DNA data from eight different reference plants from a breeder and the *in vitro* line Tks203 from the Prüfer group was evaluated. Twelve allelic positions in the gene leading to thirteen different allele variants were identified. Only seven of the twelve positions are causing AA changes.

Although *T. koksaghyz* is diploid, three instead of only two *TkoAHAS1* alleles could be detected in 50 % of the reference plants. This could indicate triploidy. Since these plants were selected material from the cooperation partner (chapter 2.1.5), the exact genetic background was not known. Possibly, the cooperation partner had succeeded in identifying hybrids, which accidentally received a higher ploidy (*281, 282*). However, this was not investigated further in this work.

By examining the field mix material from the Thiele group, additional four allelic positions could be discovered in the gene. This leads to a total of 16 allelic positions in *TkoAHAS1* identified in this work. The high number of identified allelic variants in a rather small focus group is comprehensible; the heterogeneity of *T. koksaghyz* material was (*5*) and still is discussed (*53, 64*). Research is substantially affected by the heterogeneity issue. Compared to other crops like wheat and barley or model plants like *A. thaliana*, no elite lines with specified characteristics but a range of accessions is available (*8, 64*). This may be on the one hand due to the commotions after the abrupt stop of cultivation and research after WWII (*53, 54*). On the other hand, the self-incompatibility of the plant is strongly impeding breeding efforts. Additionally, the necessity of vernalisation – yet, for a small percentage of the material (*283*) – may be another limiting factor.

The publication of the draft genome from Lin *et. al* (*66*) in 2017 can thus definitely be considered a milestone. Nevertheless, it was not possible to fully recover the TkoAHAS1 sequence information described here in the draft genome (Tab. 3-9). High conservation of this essential enzyme was expected. In contrast, the results showed that only about 20% of the gene sequence could be recovered. Consideration of results with a low identity value is not reasonable for such a comparison of the same gene in individuals of the same species. One possible reason for these divergent results could be the highly error-prone procedure of data processing to generate a draft genome (*284*). However, these results did not negatively affect this work because the own studies to identify the *AHAS* sequence had already been completed.

Further, the conserved domains for e.g. binding of the cofactors FAD and ThDP or the RHEQ-motif with the catalytic active Glutamate were recovered in the AA sequence of TkoAHAS1 (Annex 6-4) – confirming the identified sequence and the highly conserved characteristic of the AHAS protein itself. By phylogenetic analysis (Fig. 3-1) TkoAHAS1 appears to group with AHAS of *C. intybus* and *S. asper*. These two species, like *T. koksaghyz*, belong to the tribe of Cichorideae in the family of Asteraceae (*1*), so the assignment in the phylogenetic tree can be considered reasonable. The members of the adjacent Asteraceae group including AHAS of *H. annuus* (AHAS1 and AHAS2), *A. cotula* and *Xanthium sp.* belong to the Asteroideae tribe (*1*). Information about *H. annuus* AHAS proteins and herbicide resistance was considered, as it is

a member of the Asteraceae family with a respectable amount of public information. Contrary to the expectations no second AHAS gene was identified in *T. koksaghyz*. Certainly, that is plausible: *H. annuus* is an exception concerning genome evolution (e.g. the species did undergo a genome triplication followed by genome duplication) (*285*) and hence was not the ideal but as explained the most appropriate reference for investigations.

As mentioned in the first paragraph, RNA data was exploited to search for further *AHAS* sequences – but without success. However, the data were used to verify the identified allelic positions of *TkoAHAS1*, revealing that there was a) not a balanced ratio of the two *TkoAHAS1* alleles in a dataset and b) the allele ratios between the data sets were different (Tab. 3-8). On the one hand this may be due to errors during the sequencing process as well as the processing of the data itself: by quality trimming, reads are sorted out which otherwise would have accounted to the mapping (chapter 2.2.17.1). On the other hand, even more consistent, a differential transcription of the two alleles can be assumed. A transcription disequilibrium can be due to several factors e.g. the genomic environment or the tissue (*286*). In fact, both datasets were generated from different plant parts – leaf and inflorescence. It must be emphasised that the data were collected for the purpose of sequence search - but not for deeper transcript analysis. Therefore, transcriptional equilibrium was not addressed in detail.

4.2 Mutagenesis approaches

4.2.1 Herbicide resistant EMS mutagenized plants

EMS mutagenesis was conducted with *T. koksaghyz* seeds according to the strategy named in chapter 3.2.1. Fig. 3-2 illustrates the differences between the standard and the applied TILLING procedure. Usually, plant seed material (M_0) is mutated and emerging plants (M_1) are obliged to self-fertilize. These individual plants of the M_1 generation are considered as chimeric. This means, that genetic information is not identical throughout the organism due to individual EMS induced mutation events in individual cells of the embryo in the seed. After meiosis in M_1, however, the progeny (M_2) are no longer chimeric. Thus, the M_2 material is feasible for genetic characterization and for screening of plants with desired phenotypic characteristics. M_2 individuals of interest to the particular research are propagated to characterize and study the progeny (M_3) in more detail with respect to genotypic and phenotypic traits (*275, 276*).
The herein applied TILLING procedure with *T. koksaghyz* however, was different due to two main reasons. First, self-incompatibility is a huge bottleneck and impedes the selfing step, hence no true M_2 population could be generated. Second, the Prüfer group experienced a reduced fertility after EMS mutagenesis, which is among other adverse effects a common observation in TILLING populations e.g. in *H. annuus* (*287*) as well as *A. thaliana* (*182*).
Hence, instead of conducting extensive and time-consuming crossing approaches amongst the M_1 plants to generate a non-chimeric M_2 population, it was decided to use the M_1 plants directly. After herbicide application, surviving plants were selected, further investigated on genomic level (*TkoAHAS1*) and used for crossing studies. The dominance of the AHAS conferred herbicide resistance is enabling this procedure, since a heterozygous mutation is sufficient for the plant to survive herbicide treatment (*124*). This approach reduced the number of plants drastically, which also facilitated handling in view of the limited space in the greenhouse.

It can be concluded that the procedure of herbicidal selection in already the M_1 generation was successful: the herbicide was lethal to most M_1 plants without resistance conferring mutations in the gene, minimizing the count of plants for investigations to a manageable number of 40. Within these 40 M_1 plants it was not possible to identify any herbicide resistance conferring mutations in *TkoAHAS1* (Fig. 3-3). However, *TkoAHAS1* mutations were identified in the M_2 generation (after stabilization via the germline). This may be due the chimeric character of the M_1 plants. Trials to provoke non-chimeric regrowth after cutting the plants at ground level did not overcome the chimera problem. Since genotyping of *TkoAHAS1* was done by leaf sampling, it can be assumed that inherited *TkoAHAS1* mutations in M_2 were due to chimeric cell tissue in M_1, which was not detected by this sampling.
Furthermore, especially in combination with this assumption, one could consider a possible contribution of metabolic resistance (see also chapter 4.3). Moreover, in the mature M_1 plants, the robustness and ability of regrowth of chimeric tissue from the root after recovery from herbicide treatments should also be regarded.
Generally, Imazamox is incorporated via the leaves and distributed over xylem and phloem. IMI application to seeds results in seedlings not developing the first leaves. Being applied to older plants, the lethal impact of IMI becomes evident delayed by e.g. cessation of growth, chlorosis and anthocyanin expression. In contrast, exposure to sublethal doses can lead to temporary stunting of the plant and shortening of internodes (*288*).
In this work, Imazamox was always applied on seedlings with at least three leaves and lethality of the herbicide was observed. Usage of sub-lethal doses as well as a wrong application mode can therefore be ruled out. However, M_1 plants were able to survive repeated treatments with Imazamox. Stunting of plants was observed for some individuals, which then have been sorted out. 40 M_1 individuals were retained that stalled development but recovered several weeks after treatment – but none of these plants contained an AHAS conferred resistance. It can be assumed that these plants could tolerate several treatments as they probably metabolized the herbicide. Another possible reason for their survival may lie in the architecture of *T. koksaghyz*: the root is a huge sink for nutrients and therefore enables recovery of the plant from biotic and abiotic stresses. A selection by herbicide application may thus be rather inefficient with mature plants having already evolved a robust root system – as it was the case for the 40 M_1 plants that were cultivated over several month and repetitively treated with Imazamox. Such a recovering behaviour after herbicidal treatment has also been observed by Eggert in field trials (*289*).

Finally, crossing efforts on the low flowering and low fertile M_1 population with wild type led to three interesting, related progenies (M_2 generation): a putative IMI tolerant individual (M_2-3; discussed in chapter 4.3) and two IMI resistant individuals (M_2-1/ 2). The resistance is most likely due to a mutation which is located at nt position 572 (C/T) and causing the mutation Ala191Val in TkoAHAS1. It is described that EMS mutagenesis is favouring G/C to A/T transition (*290*), which is perfectly reflected by the detected mutation. Ala191 in TkoAHAS1 corresponds to Ala205 in AthAHAS. Referring to chapter 1.2.3 it is one of the six most prominent positions for herbicide resistance mutations. In fact, the mutation Ala205Val is already described for IMI resistant *H. annuus* (Tab. 6-4) and to occur in IMI and SU resistant weed species (Tab. 6-5). Interestingly, Ala205 contributes to binding of SU but not of IMI (Tab. 1-2). However, the reports of evolved resistance towards IMI in crops and weed species are covering the findings here.

To investigate on the hereditary behaviour of this resistance conferring mutation C572T, further crossings were conducted with the two individuals M_2-1 and M_2-2. Advantageously, the initial

EMS mutagenesis did not affect the fertility or growth behaviour of these plants and the majority of the progeny; the phenotype did not differ from the WT.
Fig. 3-3 gives a comprehensive overview about all investigations concerning the untargeted mutagenesis approach to receive herbicide resistant *T. koksaghyz* plants.
Manual crossing of M_2-2 with a self-compatible WT yielded in 14 heterozygous IMI resistant progeny (M_3 generation). With these individuals of the M_2 and M_3 generation further crossings up to the M_5 generation were performed and evaluated.
In these, herbicide selection was omitted and the *AHAS1* mutation C572T was verified using a restriction enzyme assay for each individual plant. The reliability and benefits of this procedure were demonstrated by reviewing the resistance results of the M_3-gauze population obtained by IMI selection by identifying some individuals with homozygous WT *AHAS* sequences that were incorrectly designated as resistant.
A minority of plants with abnormal growth behaviour (dwarfing) was detected in the M_4 generations (Fig. 3-8, Fig. 3-9) and identified as either heterozygous or homozygous for the resistance conferring mutation. These plants did neither grow properly – instead, growth paused at an early stage – nor did they develop any flower although favourable growth conditions were present. This phenotype, resembling IMI treated susceptible plants, would have been discarded in an IMI selection and would have led to an erroneous ratio concerning the zygosity studies in the M_4 and M_5 generations. It cannot be ruled out that such dwarf mutants have already occurred in and before the M_3 generations, where selection was done via IMI. However, no connection between dwarfing and the mutation in *TkoAHAS1* can be drawn because plants heterozygous as well as homozygous for the mutation occur to show the dwarf phenotype. Dwarfism is an often-described phenotype caused by mutagenesis – as well as e.g. leaf colour chimera, deformations and chlorosis (*291*), which have been observed in the M_1 generation of EMS mutated *T. koksaghyz*, too.

Besides verifying the zygosity of over 650 plants of the M_3-gauze population, the appearance of the mutation was determined for over 370 individuals of the M_4 and M_5 generations, generated by crossing of heterozygous plants. Several crossings were analysed to receive reliable results and to preserve seed material for potential further investigations. Interestingly, Mendelian inheritance was not found thoroughly (Tab. 3-14). Whilst the proportions of homozygous WT and heterozygous plants were roughly following Mendelian inheritance, the percentage of homozygous mutant plants was, with 7 %, lower than expected. Furthermore, the values of individuals that did not germinate or died shortly thereafter were about 7 %, too. Adding this proportion to the homozygous mutants, the inheritance followed approximately Mendelian distribution. The zygosity ratios of the related M_4 and M_5 generations (Tab. 3-14) differed by reason that all fertile seeds of M_5 germinated. This may be due to a putative loss of the disadvantageous dwarf effect by crossing only between healthy, fertile plants.
Irrespective of this, it must be taken into account that these progenies are the result of EMS mutagenesis. Crosses of these may have resulted in recombination of unfavourable mutations which haven't been further investigated in this work.
Remarkably, the origin of all discussed progeny lies in three plants from the M_2 generation. The 14 M_3 plants, in turn, were the progeny of only one M_2 plant crossed with a self-fertile *T. koksaghyz* WT plant. Hence, all crossings (Fig. 3-3) were conducted between highly related individuals, although *T. koksaghyz* is a known self-incompatible species. The initial M_2 parent plants may have overcome their self-incompatibility due to a side effect of EMS treatment. For the subsequent crossings of the 14 plants this effect was either stabilised or enhanced by the self-fertile WT parent. The circumstance of overcoming self-incompatibility enabled these

highly successful crossings resulting in a high number of hetero- and homozygous Imazamox resistant progeny.

EMS treatment to obtain mutations in AHAS is well-described and has been successfully applied in several plant species, e.g. maize, rice and wheat (*136*) and in *A. thaliana* (*138*). In the latter, five different AHAS mutations were found in 12 individuals in a pool of 250,000 M_2 plants (*138*). It is not possible to directly compare the results and efficiency of this EMS approach on AHAS in *A. thaliana* with the results presented here on *T. koksaghyz*. Differences exist not only in species, but also in the ability of *A. thaliana* to self-fertilize, the different protocols for EMS mutagenesis, the different population sizes, and the application of the standard/ modified TILLING strategy. Nevertheless, it can be summarized that even in the versatile and widely studied model organism *A. thaliana*, no disproportionately higher number of resistance-mediated mutations could be identified.

In conclusion, the applied TILLING strategy did bypass on the one hand A) the selfing/fertility problem and by early herbicide selection B) a time- and space-consuming crossing of many M_1 plants but on the other hand, it did not solve C) the chimera issue.
Anyhow, the approach resulted in the identification of an herbicide resistance conferring mutation in *T. koksaghyz* that was successfully passed on to several subsequent generations. Since there are no published results for such an approach targeting AHAS in *T. koksaghyz* so far, a comparison was not possible. Whether these results could have been gained by a standard TILLING procedure with comparable effort, cannot be determined, as this comparison was not part of this work.

4.2.2 CRISPR/Cas9 edited plants

The use of SDN and in particular of CRISPR/Cas9 simplifies site-directed modification of genes. Still, the chosen system is setting restrictions towards its application. The used *S. pyogenes* Cas9 is strictly accepting “NGG” PAM sequences although it was described that “NAG” PAM sequences are recognized with reduced efficiency, too (*292*). Hence, it was not possible to target all the chosen six positions in *TkoAHAS1* exactly (Tab. 3-15). Proper target cleavage was verified for four of the six designed sgRNAs *in vitro*. Possible reasons for misfunctioning are discussed in literature e.g. the thermodynamic stability of the sgRNA-DNA complex (*220*). SNPs are not existent in the protospacer regions of the chosen target sites (Tab. 3-2, Tab. 3-3, Tab. 3-15), ruling out binding instabilities due to SNP based mismatches; although distinct mismatches are accepted (*292*). Several tools are available to test the characteristics and suitability of sgRNAs also regarding to off-target activity, e.g. Mfold (*293*), CCTop (*294*) or CRISPOR (*295*). Since no reference genome for *T. koksaghyz* was implemented in such online tools and the targets in *TkoAHAS1* were set, these additional applications were not used. Two AA in TkoAHAS1 have been selected to target by site-directed mutagenesis on gene level: one in the middle, corresponding to Arg363, and the other at the 3’ end, corresponding to Ala639. Search for off-targeting was not in the scope of this work.

The results of the approaches towards both positions differed significantly in terms of regeneration efficiency as well as identification of target modifications.
For the position Arg363, only transformation regenerates containing *TkoAHAS1* WT alleles were detected. Furthermore, the regeneration percentage for this position was distinctively lower, possibly due to the following reasons.

First, Arg363 is described as a crucial AA for catalysis of the enzymatic reaction; this is reflected by the high level of conservation throughout various organisms (Tab. 4-1) (*69*). In the catalytic cycle of AHAS, the reaction occurs through ThDP mediated transition steps. Arg363 recognizes the second 2-ketoacid substrate (PYR or 2-KB) by stabilizing the negatively charged carboxylate group via its positively charged amino group. Substitution experiments with EcoAHAS showed that other AA are not able to fit sterically or to provide carboxylate coordination due to a lack of positive charge (*296*). Hence, it is plausible that only a single functional mutation (Arg377His, AthAHAS numbering) found in the weed *Apera spica-venti* is known to date, giving resistance to SU, SCT, TP (*297*) as well as IMI (*123, 298*) (Tab. 6-5). Further, the participation in binding of herbicides was described (*125*) and discussed to be a relic of the quinone binding domain of the protein ancestor POX (*140*). AHAS is also regulated by quinones as a mediator of the cell's redox state and herbicides inhibiting AHAS use the quinone binding pocket (*94*). Besides Arg363, Asp362 could have been targeted by the applied sgRNA, too. Asp362 corresponds to Asp376 in AthAHAS and the position is described to be important for interaction with the catalytical relevant Arg377 (*296*). Moreover, it is one of the six most prominent positions for herbicide resistance conferring mutations and is highly conserved throughout species (Tab. 4-1) (*69*). Only the substitution Asp376Glu (AthAHAS numbering) is described to confer IMI resistance (*298, 299*) and was also detected in resistant weeds (*123*) (Tab. 6-5). The triplets for both AA Arg363 and Asp362 are next to the DSB induced by Cas9 cleavage (two and five bp, respectively). Whilst for the mutation Asp362Glu (GAT to GAG/ GAA) a single nt substitution would have been sufficient, for Arg363His (CGT to CAG/ CAC) already at least two nt would have been needed to be exchanged.
Anyhow, none of these or any other mutations were detected. Since both alleles are transcribed (chapter 3.1) one would have expected at least one modified allele and compensation of enzymatic functionality by the other allele. Otherwise, it is also conceivable that for this position in the middle of the gene and the protein, respectively, no huge sequence changes can be tolerated to maintain functionality. In 2016, Iaffaldano *et al.* (*300*) used CRISPR/Cas9 for targeting successfully the inulin biosynthesis associated enzyme fructan:fructan 1-fructosyltransferase (1-FFT) in *T. koksaghyz*. Instead of *A. tumefaciens, A. rhizogenes* was used for transformation. Anyhow, the CRISPR/Cas9 approach yielded a range of insertions and substitutions of single nucleotides, as well as deletions of up to 22 nucleotides. These caused frame shifts and premature stop codons, yielding significantly truncated and unfunctional proteins (*300*). Since AHAS is an essential enzyme, such a truncation around the middle of the gene would have led to loss of function. Considering the transformation system/callus regeneration, one functional allele may not be sufficient for a regenerating callus. It is therefore questionable whether a different result could have been achieved with supplementation of the amino acids.

Second, it is possible that the sgRNA is working *in vivo* not as proper as *in vitro*. Several factors are known influencing the efficiency: the DNA accessibility (*301*), the thermodynamic stability of the sgRNA-DNA complex (*220*), the affinity between sgRNA and Cas9, the GC content of the sgRNA, the targeted DNA strand (transcribed or non-transcribed) as well as the purine content of the protospacer region (*302*). The latter three factors are unaffected concerning the *in vivo* or *in vitro* setup. The general functionality of the system *in vivo* was confirmed by the positive results of the second targeting approach. Concerning the DNA accessibility: for the *in vitro* assay a linear PCR amplicon of *TkoAHAS1* was provided. This was lacking the features of genomic DNA e.g. methylation or chromatin structures. Whereas methylation is not affecting the CRISPR efficiency (*303*), chromatin blockage can make the PAM or protospacer region inaccessible for the detection by Cas9 or binding by the sgRNA (*301*). This discrepancy

between *in vivo* and *in vitro* results was e.g. described for rice protoplasts by Liu *et al.* (*301*) but also for CRISPR approaches in zebrafish (*304*). Several approaches have been described for identification of accessible DNA regions (*305*). Such methods can deliver more information and in combination with a genome sequence they can be helpful for CRISPR target selection in *T. koksaghyz*.

Tab. 4-1 List of TkoAHAS1 positions and the referring positions in AthAHAS, SceAHAS and EcoAHAS isozyme II (adopted from McCourt and Duggleby (*69*)) as well as the known IMI and SU resistance conferring mutations (adopted from Yu and Powles (*124*)).

TkoAHAS1	AthAHAS	SceAHAS	EcoAHAS isozyme II	Resistance conferring mutations
Asp362	Asp376	Asp379	Asp275	Glu, Ala, Asn, Cys, Gly, Pro, Ser, Trp, Val
Arg363	Arg377	Arg380	Arg276	His
Ala639	Ser653	Gly657	Pro536	Asn, Thr, Ile, Phe
Gly640	Gly654	Gly537	Gly658	Glu, Asp

In contrast to Arg363 and Asp362, Ala639 is not described as a conserved AA throughout species, reflected by the variations listed in Tab. 4-1 (*69*). In AthAHAS Ser653 belongs to the C-terminal arm which is looping over the active site (*79*). Residues Pro649 to Ser653 influence the arrangement of the catalytically relevant AA that bind the FAD cofactor, e.g. Arg377 – when not bound to an herbicide (*88*). Despite, in the presence of SU or IMI, Ser653 is stabilizing the binding of the herbicide to the protein, e.g. via hydrogen bonds (*69*). Several resistance conferring substitutions are known e.g. Asn, Thr, Ile, Phe – either found in resistant weed species (*123*) or artificially generated (*124*) – indicating proper catalytic functioning for a range of AA although stabilization of FAD might be affected. Next to Ala639 residue Gly640 is situated, which occurs to be more conserved than its neighbour and is described to stabilize binding of IMI and SU, too. It is therefore also a known position to confer herbicide resistance (substitutions Glu and Asp) (*69*).

For the second targeting approach in this study on position Ala639, a higher number of regenerates with a range of different substitutions was identified (Tab. 3-16). Hence, compared to the other position (Arg363) this site was probably more accessible for the CRISPR/Cas9 system *in vivo*. Additionally, the targeted position is very close to the 3' end of the gene and hence, the protein; sequence changes may be more tolerable for proper functionality and activity. Transformation regenerates (T_1 plants) contained insertions, deletions or substitutions at the target site in *AHAS* but also T-DNA, causing continuous expression of the CRISPR/Cas9-System and ongoing cleavage activity. This resembles a chimeric character of the regenerates. Hence, testing for herbicide resistance in this early generation was not suitable.
Surviving T_1 individuals were available for crossbreeding and with WT *T. koksaghyz*. Several crossings with WT were selected for analysis. Altered *TkoAHAS1* sequences were detected in 12.9 % (from 271 F_1 plants). Detailed Sanger sequencing of F_1 individuals revealed either a) mutations that were already found in the parent T_1, b) new kinds of InDels or c) no changes at all. A range of eight different mutations was determined (Tab. 3-18), mostly inducing AA sequence changes and leading to premature STOP codons. Since T-DNA was detected in most individuals, as with the T1 plants, a chimeric character impeded IMI resistance tests.

Four T_1 individuals without T-DNA were identified, carrying the same “In1916a” insertion. In the altered AA sequence Ala639 remains, but the substitution Gly640Arg occurred. Glu and Asp are described to confer herbicide resistance – both consist of a negatively charged carboxylic function and are bulkier than Gly (*69*). Despite, Arg is positively charged and even bulkier. It is possible that the bulky character of Arg is sufficient to inhibit binding of IMI.
In addition to the Gly640Arg substitution, the insertion “In1916a” effects a 6 AA shorter and altered sequence. The truncated sequence is unlikely to affect the functionality of the above-mentioned structural arm or the AHAS itself. Hence, it is reasonable to expect an IMI resistant and functional AHAS caused by the insertion “In1916a”.
Verification of this assumption was not conducted due to the small sample number not allowing reliable results. Instead, further crossings to preserve the interesting mutation were successfully conducted but functional review on the F2 was not continued further.

The other mutations (according to Fig. 3-14), although identified in individuals carrying the tDNA, shall also be discussed towards their potential of conferring herbicide resistance.
AHAS mutation number (6) and (8) are comparable to the above-named insertion on AA level by consisting of the same Gly640Arg substitution but an even more truncated sequence. This may effect an herbicide resistant and functional AHAS. Mutation number (5) consists of a Ala639Thr substitution, which is also described to confer IMI resistance (Tab. 4-1). Mutation number (7) results in a 12 AA shorter sequence and a Gly640Val substitution.
Due to the chimeric character (regarding the tDNA) of the plants carrying these mutations, the effects on herbicide resistance could not be verified.

The low number of regenerated plants discussed here is due, on the one hand, to problems in establishing the transformation process at the JKI Quedlinburg and, on the other hand, to limited human resources. Whilst an *A. tumefaciens* transformation is done quickly, maintenance of regenerates is laborious, time- and space-consuming. Furthermore, it would have been of interest to investigate a HR approach and compare the results of NHEJ and HR experiments. HR allows exact nucleotide substitutions leading to defined AA change(s) of interest (*197*). Thus, herbicide resistance, could be implemented in a more targeted manner.

4.2.3 Comparison of the EMS and CRISPR/Cas9 method

In view of the initial aim of this work to receive an herbicide resistant *T. koksaghyz*, the untargeted EMS mutagenesis approach was successful: individuals with the IMI resistance conferring mutation Ala191Val were obtained. Several crossing series proved propagation of the mutation to the next generations without negatively affecting growth behaviour or fertility. In contrast, the individuals obtained from the CRISPR/Cas9 approach could not be tested for herbicide resistance. However, it was of further interest to compare the methods in general which is discussed in the following.
For the *A. tumefaciens* mediated transformation of the CRISPR/Cas9 system into leaf explants less initial plant material than for the EMS mutagenesis is required: instead of thousands of seeds at least one individual is sufficient. Thereby an identical genetic background for the regenerates is provided, allowing a direct comparison of physiological effects caused only by mutations. This is particularly advantageous for those species of which no accessions or breeding lines with homogeneous genetic composition are available – as is it the case for *T. koksaghyz*. The EMS approach was conducted with genetically heterogenic material. Contrary, all regenerates of the CRISPR/Cas9 approach have the same genetic origin – neglecting somaclonal variation, which occurs during longer *in vitro* cultivation (*306*).

The recovered number of individuals with mutations in *AHAS1* after EMS mutagenesis (M_2 generation) appears to be low regarding the initial high amount of mutated seed material. Genome editing via the CRISPR/Cas9 method resulted in a range of different mutations for the targeted position at the 3' end of the gene. Crossing efforts led to four individuals without T-DNA which were not used for herbicide resistance tests.
In both experiments the problem of chimera occurred. In case of the EMS approach this is overcome after the first crossing step. In contrast, *A. tumefaciens* mediated transformation of the CRISPR/Cas9 system can lead to the persistence of the T-DNA over several generations. As explained, preservation of the T-DNA is in most cases equivalent to a chimeric character. Hence, the status of *AHAS1* in the individual plant is in a constant change and is not enabling herbicide application trials, since the output cannot be related to one certain mutation.
EMS mutagenesis is a tool which has been successfully used for decades; still, it has to be considered the targeted genome is mutated randomly (*172, 180*). Yielding a healthy and fertile individual with the desired genotypic/ phenotypic characteristics is a result of several crossbreeding steps to lose negative side effects (e.g. dwarfism, reduced fertility) of unwanted mutations. In contrast, off-target activity of CRISPR/Cas9 is possible but can be reduced to a minimum by choosing sgRNAs with less off-target probability. Tools to evaluate this are available online as explained. CRISPR/Cas9 has revolutionized the field of SDN; the design of experiments is faster and simpler in comparison to e.g. TALENs. Ideally, the resulting plants genetics (and vigour) is equivalent to the original plant used for transformation.
Comparing the time and cost effectiveness, CRISPR/Cas9 outmatches other SDN as well as undirected mutagenesis (*307*). *Chen et al*. described a time advantage of four years for CRISPR/Cas9 over mutation breeding – from the initial experiment to the desired new variety (*308*). Anyhow, regarding the regulatory aspect, at least in the European Union, the chances of using CRISPR/Cas9 for breeding of new varieties are limited since the decision of Court of Justice of the European Union in 2018 (*198, 309*).
The decisions towards articles 2(2) and 3(1) of Directive 2001/18 declared that a) "organisms obtained by means of techniques/ methods of mutagenesis constitute GMO" and b) "techniques/ methods which have conventionally been used in a number of applications and have a long safety record are excluded from the scope of that directive" (*309*). Thereby the strict process-based legislation was further manifested, and the alternative of product-based decision was ruled out. This means, although EMS mutagenesis is introducing a lot of unintended and unidentified mutations in an organism, this technique is regarded as safe due to a long history of safe use and is further allowed for application breeding in the EU. Despite, CRISPR/Cas9 generated plants will be considered as GMO (*198*). This is cutting down the chances for plants from CRISPR/Cas9 approaches to be used in agricultural aspects now and in the near future in the EU. Still, research is in charge to show foresight and hence approaches whose output maybe of importance in the distant future have to be pursued. Nevertheless, regulation on new plant breeding techniques is varying over the globe: in contrast to the EU e.g. U.S.A. and Canada are pursuing a case-to-case decision for every request (*198, 310*).
Plant breeding, at least in the EU, has been deprived of a powerful, simple and fast tool, so that it must continue to rely on classical methods that work but lag behind new breeding techniques in many aspects as described.

4.3 Transcriptome studies on resistant and putative metabolic tolerant *T. koksaghyz*

As addressed in chapter 4.2.1 a putative Imazamox metabolic tolerant plant (M_2-3) was identified from the EMS mutagenesis approach. To get an insight into the possibly changed or

intensified pathways for phytodegradation of the herbicide, transcriptome data was evaluated. Data of clonal plants of the tolerant individual were compared to that of the two related resistant ones (M_2-1/ 2) in background of herbicide application.

Besides investigating into detoxification enzymes, expression of the herbicide target AHAS was examined. No changes of the *AHAS* transcription rates could be detected – neither for the resistant or the putative tolerant individuals, nor for the WT control. There are two possible reasons for this. Firstly, (at least one gene of) AHAS is constitutively expressed throughout the plant, whereas it is highly expressed in the meristematic, metabolically active tissues (*79*). Since three-week-old leaves of *T. koksaghyz* plants were sampled for the transcriptome approach, clearly no differential expression could be detected. Secondly, Salas-Perez *et al.* (*280*) described no increased amplification of the target gene glutamine synthetase for glufosinate-tolerant and -resistant plants after herbicide application. Similarly, Yu *et al.* (*311*) also demonstrated comparable mRNA levels for AHAS in susceptible and resistant *Raphanus raphanistrum* populations, but reported altered AHAS activities. This is consistent with the present results: AHAS expression in *T. koksaghyz* was also not differentially regulated after imazamox exposure - regardless of whether resistant, tolerant or susceptible plants were considered. However, an involvement of AHAS in plant response mechanisms to herbicide exposure cannot be ruled out.

Therefore, the focus was set on screening the transcriptome data for DE transcripts of potentially involved detoxification enzymes. These were found to be present in both, tolerant and resistant plants (Tab. 3-21). This fits the statement of Salas-Perez *et al.* (*280*) and is comprehensible with respect to the explanations from chapter 1.2.3: the plant is generally forced to deal with the herbicide as a phytotoxic substance and to develop a strategy to catabolize or modify it. Anyhow, it was of interest to determine whether the putative tolerant plant possesses further individual mechanisms. Since the herbicide is A) an inhibitor of the essential AHAS enzyme and is B) a xenobiotic substance, that the plant must cope with, it represents a double burden for the organism – but not for a resistant plant whose AHAS is not inhibited.

As already mentioned, Rojano-Delgado *et al.* (*165*) were able to demonstrate that Imazamox is converted to an hydroxylated metabolite in susceptible/ resistant plants and further to a glucose-conjugated metabolite in resistant plants. This corresponds to the detoxification phases I and II. The hydroxylated conjugate can be generated by the conversion of Imazamox via a P450 monooxygenase (CYP). Glycosyltransferases, in particular glucosyltransferases, are further able to generate the glucose-conjugated form of it. The results of Domínguez-Mendez *et al.* (*312*) on Imazamox resistant durum and wheat varieties support the thesis of CYP dependent metabolism. They could show that a lower amount of conjugates is detected when malathion – an inhibitor of CYP – was used. Anyhow, it is less known which CYPs play major roles in xenobiotic degradation – only some have been investigated from different species, e.g. soybean or tobacco (*313*). For the enzyme classes CYP and glucosyltransferases DE transcripts were found in high numbers and together account for more than 50% of the total transcripts listed. They are found in similar proportions not only in the intersection of the two resistant variants, but also in the putative tolerant plant alone and in the intersection of all three variants. Thus, it can be assumed that both phase I and II reactions occur in both, the resistant and the putative tolerant variants.

However, detoxification does not end with the glucose conjugate – phase III follows. The sequestration of the glucose-conjugated metabolite into the vacuole is carried out by ATP-dependent ABC transporters (*149, 155, 158*). In fact, DE transcripts could also be detected for this enzyme class.

Besides these three classes of enzymes, DE transcripts of Glutathione-*S*-transferases were also detected. It is generally possible that in phase II, instead of glycosylation, glutathionylation also takes place (*155*). However, only the detection of the two mentioned metabolites of Imazamox have been described so far (*165, 312*). Studies on glutathione pathways by Balabanova *et al.* (*269*) on Imazamox resistant *H. annuus* indicate that Glutathione-*S*-transferases are involved in detoxification – but not in susceptible plants. Possibly the herein invested variants of *T. koksaghyz* use this pathway – although this hypothesis needs to be confirmed by an appropriate metabolite analysis.
In addition to DE transcripts of glucosyltransferases, transcripts of glycosyltransferases were also found. These generally enable UDP-dependent *N*-/ *O*-glycosylation of substrates. Plants possess different enzymes and isoenzymes that catalyse similar reactions, but with different substrate preferences. This variance ensures the organism an efficient degradation of different xenobiotics (*314*). Thus, the differential expression of glycosyltransferases appears to be comprehensible here as well.
DE transcripts of peroxidases were also found at low levels, particularly in the putative tolerant plant. Peroxidases are important in rendering ROS harmless as part of the overall protection against oxidative stress (*147*). Such oxidative stress can be caused by e.g. xenobiotics; however, herbicides targeting AHAS do not induce oxidative stress (*315*). Although these herbicides trigger the oxygenase side reaction of AHAS, resulting in the production of ROS, there is no accumulation of ROS. This is due to the fact that the ROS formed cause oxidation of the cofactor FAD, which means inactivation of AHAS (*72*).
The non-involvement of oxidative stress is confirmed by the absence of a) glutathione peroxidases as well as b) Glutathione-*S*-reductases among the DE transcripts. These two enzymes are able to build up a system for efficient detoxification of ROS via glutathione (*316*). The lack of DE transcripts indicates that glutathione does not play a role in the control of oxidative stress. This relationship is already described for Imazamox resistant *H. annuus* (*269*). Transcript profiling of wild non-target-site resistant *H. annuus* does not reveal an expression of peroxidases, either (*317*). Among the peroxidase annotations, terms of “peroxidase 42”, “peroxidase n1(-like)” or “heme peroxidase” were found. Anyhow, no detailed information about the roles of those peroxidases could be received, hence no exact explanation can be given why slightly more DE transcripts of peroxidases were found in the putative tolerant variant.
In view of the points mentioned and regarding the chemical structure of Imazamox, it is understandable that nitroreductases, deaminases and esterases do not experience differential expression after Imazamox treatment in any case studied.
For none of the discussed enzymes an over- or under-distribution in M_2-3 was observed in comparison to the other variants (Tab. 3-21). Only peroxidases were found to be more abundant in the putative tolerant plant.

The high variance in the field mix (Fig. 3-15) did not allow a comparison with these data. Although this material was applied for EMS mutagenesis, it was not genetically homogeneous. The pooled samples used for the analysis were therefore not suitable for a one-to-one comparison with the samples of a single plant. Likewise, the parent plant of M_2-1/2/3 would not have been suitable due to its chimeric character. A comparison of samples – with and without point mutation(s) – is only reasonable if the transcription patterns show high similarities. This is usually achieved by using genetically homogeneous material or material that is sufficiently closely related. Nevertheless, data collected from the control was used. The mentioned DE transcripts, which were exclusively found in the putative tolerant variant, were examined for non-silent mutations. It was assumed that the mutations may have been caused by EMS

mutagenesis and may have an influence on metabolic tolerance. Hence, the mutations found in the putative tolerant variant were compared to those of the unmutated field mix samples. The results showed that the identified SNPs/ mutations were most likely already present in the field mix material before the EMS mutagenesis.
However, three SNPs, most likely belonging to the transcripts of a CYP-like gene (CYP72A219), attracted further interest. CYP72A219 is known for its role in tolerance to AHAS inhibitors (*279*) or glutamine synthetase inhibitors (*280*); to date, no information concerning tolerance relevant SNPs is known. Since the identified SNPs were also present in the control, they are unlikely to provide an explanation for metabolic tolerance.

The metabolic tolerance of plants against phytotoxic substances is based on a complex interaction of different enzymes and pathways. Transcriptomic profiles of detoxification enzymes were highly similar in both, IMI resistant and putative tolerant individuals due to their high degree of relationship.
Transcriptomic data provide information on the first step of protein biosynthesis, but do not provide information on translation or enzyme activity. The experiments performed could not explain the observed tolerance. This would have required further studies on the metabolome and proteome and, if necessary, the degradation products of the herbicide, but these would have been beyond the scope of this work.

5 Summary

T. koksaghyz is an upcoming alternative source for natural rubber. The lacking results of breeding efforts from the early 20th century and lacking homogenous material for breeding nowadays are impeding the success of the plant as a serious bioresource. A major bottleneck is the agronomic aspect; until now cropping is highly laborious since a weed management is not available. Hence, in this work two different strategies were followed to obtain an herbicide resistant *T. koksaghyz* plant. The metabolically essential enzyme acetohydroxyacid synthase (AHAS) is a well-studied target for resistance-conferring mutations. Such were ought to be introduced on the one hand by undirected EMS mutagenesis and on the other hand by the site-directed genome editing method CRISPR/Cas9.

The findings of the herein discussed approaches are:

- The diploid *T. koksaghyz* consists of a single, intron-free, 1971 bp long AHAS gene (*TkoAHAS1*). The identified variety of allelic positions/SNPs and allele sequences with nevertheless high homology among each other reflect the genetic heterogeneity of the *T. koksaghyz* material studied.
- Undirected EMS mutagenesis of *T. koksaghyz* led to three related individuals, of which two were imidazolinone resistant and one putative tolerant. The resistant plants featured the same mutation (C572T causing Ala191Val; corresponding to Ala205 in AthAHAS), which is a well-known resistance conferring mutation in crop as well as weed species. Vital and fertile progeny were derived following nearly Mendelian distribution concerning the mutation. Successful crossings between highly related individuals imply the overcoming of self-incompatibility as a side effect of EMS mutagenesis and crossing with a self-fertile wild type, respectively.
- Several gene regions of *TkoAHAS1* were targeted successfully via the CRISPR/Cas9 system *in vitro*. *In vivo* targeting of two sites led to detection of InDels only for the position at the 3’ end of *TkoAHAS1*. The effects of these mutations towards herbicide resistance are ought to be determined.
- The transcriptomic profiles after imidazolinone treatment of the putative tolerant plant derived from the EMS approach were compared to those of its two resistant relatives. However, the transcriptional studies gave little evidence for the involvement of metabolic tolerance and could not explain the observed phenotype.

Seed material of both approaches is available at the Thiele Group for future investigations. Fertile and viable material from the EMS approach is available for further crossing trials, for example to obtain herbicide-resistant *T. koksaghyz* breeding lines.
In addition, further studies could be conducted to address the assumed herbicide tolerance. Material from the CRISPR approach is also available to investigate the influences of the identified mutations with regard to herbicide resistance.

6 Appendix

Tab. 6-1 Amino acid residues important for FAD bending/ coordination in the closed AHAS conformation. Data from SceAHAS structure with bound sulfonylurea; corresponding AthAHAS positions are listed. Table modified after McCourt *et al.* (*90*).

Interactions are H-bonds or van der Waals (vdW) interactions which can be repelling (-) or attracting (+). Conservation indicates the number of sequences in which the residues are conserved based on results from aligning 21 sequences published by Duggleby and Pang, 2000. The sequence motives that are surrounding the conserved residues are listed with respect to degree of conservation; bold: conservation in 21 sequences; normal font: conservation in at least 18 sequences; italic: conservation in 16 or 17 sequences; underlined: contact residues.

Residue SceAHAS	**Residue AthAHAS**	**FAD atom**	**Inter-action**	**Conser-vation**	**Sequence motif**
Phe201	Phe206	C8/C8M	vdW-	21	AspAlaPheGlnGlu
Leu335	Leu332	O2	H-bond	20	[**Ser**/**Thr**]Leu*Met*Gly
His355	His352	O4	H-bond	21	Met*Leu*GlyMetHisGly
Val381	Val378	O2	vdW-	17	**ArgPhe**Asp**AspArg***Val*Thr*Gly*
Val497	Val485	C8M	vdW+	21	ValGlyGlnHisGlnMetTrp
Met582	Met570	C7M	vdW+	20	GlyMetValXxxGlnTrp[Glu/Gln]

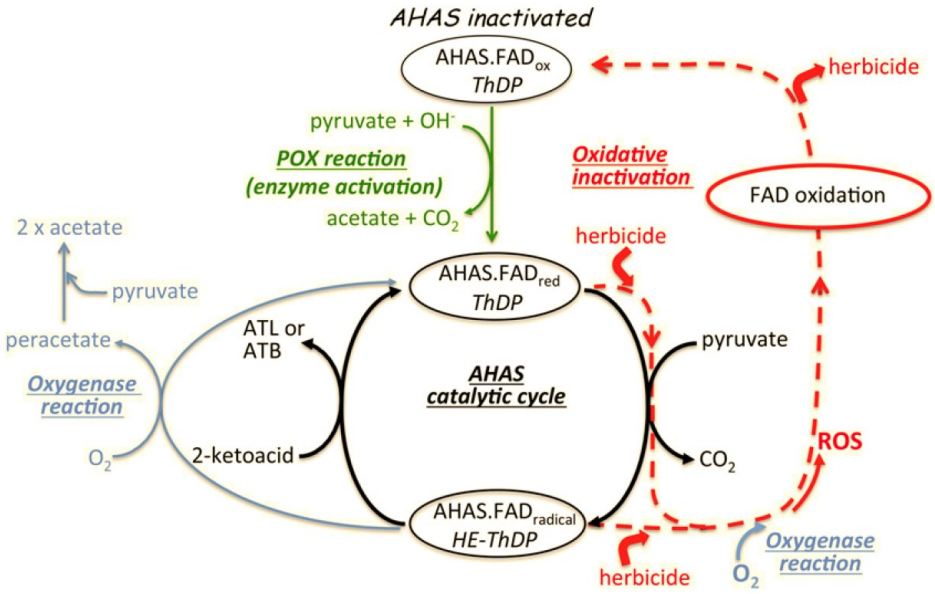

Fig. 6-1 Scheme for oxidative inhibition of SceAHAS by herbicides. ATL: 2-acetolactate; ATB: 2-aceto-2-hydroxybutyrate; ROS: reactive oxygen species. Adopted from Lonhienne *et al.* (*72*).

Tab. 6-2 List of AHAS references included in the phylogenetic analysis.

Organism	AHAS Reference in the phylogenetic tree	ID (Source) Source: NCBI Reference Sequence/ GenBank (N) or UniProtKB/Swiss-Prot (U)	Description
Saccharomyces cereviseae	AHAS	NP_013826.1 (N)	acetolactate synthase catalytic subunit [Saccharomyces cerevisiae S288C]
Chlamydomonas reinhardtii	AHAS	XP_042920710.1 (N)	uncharacterized protein CHLRE_09g386758v5 [Chlamydomonas reinhardtii]
Physcomitrella patens	AHAS 3	XP_024386701.1 (N)	acetolactate synthase 3, chloroplastic-like [Physcomitrella patens]
Ostreococcus lucimarinus	AHAS	XP_001421626.1 (N)	predicted protein [Ostreococcus lucimarinus CCE9901]
Bassia scoparia	AHAS	ACD62485.1 (N)	acetolactate synthase [Bassia scoparia]
Amaranthus retroflexus	AHAS	AAK50820.1 (N)	acetolactate synthase [Amaranthus retroflexus]
Helianthus annuus	AHAS 1	AAT07323.1(N)	acetohydroxyacid synthase 1 [Helianthus annuus]
Helianthus annuus	AHAS 2	AAT07327.1 (N)	acetohydroxyacid synthase 2 [Helianthus annuus]
Helianthus annuus	AHAS 3	AAT07329.1 (N)	acetohydroxyacid synthase 3 [Helianthus annuus]
Xanthium Sp.	AHAS	AAA74913.1 (N)	acetolactate synthase precursor [Xanthium sp.]
Anthemis cotula	AHAS 1	AEL89170.1 (N)	acetolactate synthase 1, partial [Anthemis cotula]
Sonchus asper	AHAS 1	ACF47583.1 (N)	acetolactate synthase 1 [Sonchus asper]
Cichorium intybus	partial AHAS	CAE01110.1 (N)	acetolactase synthase (ALS), partial [Cichorium intybus]
Solanum tuberosum	AHAS 2	XP_006361740.1 (N)	PREDICTED: acetolactate synthase 2, chloroplastic [Solanum tuberosum]
Nicotiana tabacum	AHAS 1	P09342.1 (U)	Acetolactate synthase 1, chloroplastic
Vitis vinifera	AHAS	CBI21345.3 (N)	Unnamed protein product, partial [Vitis vinifera]
Malus domestica	AHAS	RXH85640.1 (N)	hypothetical protein DVH24_009461 [Malus domestica]
Carica papaya	AHAS 3	XP_021905626.1 (N)	Acetolactate synthase 3, chloroplastic [Carica papaya]
Gossypium raimondii	AHAS 3	XP_012455043.1 (N)	PREDICTED: acetolactate synthase 3, chloroplastic-like [Gossypium raimondii]
Brassica oleracea	AHAS 1	XP_013603602.1 (N)	PREDICTED: acetolactate synthase 1, chloroplastic [Brassica oleracea var. Oleracea]
Arabidopsis thaliana	AHAS	P17597 (U)	Acetolactate synthase, chloroplastic (ILVB_ARATH)
Glycine max	AHAS 2	XP_003545907.1 (N)	acetolactate synthase 2, chloroplastic [Glycine max]

Organism	AHAS Reference in the phylogenetic tree	ID (Source) Source: NCBI Reference Sequence/ GenBank (N) or UniProtKB/Swiss-Prot (U)	Description
Triticum aestivum	AHAS	AAO53549.1 (N)	acetohydroxyacid synthase, partial [Triticum aestivum] GenBank: AAO53549.1
Oryza sativa	AHAS 1	XP_015626459.1 (N)	acetolactate synthase 1, chloroplastic [Oryza sativa Japonica Group]
Zea mays	AHAS 1	NP_001151761.2 (N)	acetolactate synthase 1, chloroplastic [Zea mays]
Sorghum bicolor	AHAS	CUS31209.1 (N)	Acetolactate synthase, chloroplastic [Sorghum bicolor]

Tab. 6-3 List of publicly available sequence data for *T. koksaghyz* at NCBI (status 23.03.2022) (*74*).

No	Description	Project data type	Scope	Institute	Accession	Submission date	Additional information	Data details
1	Taraxacum koksaghyz Raw sequence reads	Raw sequence reads	Mono-isolate	Shihezi University	PRJNA745038	09.07.2021	none	66 Gbases; 27907 Mbytes
2	Taraxacum koksaghyz Raw sequence reads	Raw sequence reads	Multi-isolate	Xinjiang Academy of Agricultural Sciences	PRJNA686722	20.12.2020	none	54 Gbases; 20074 Mbytes
3	Investigation of the PAR gene conservation in Taraxacum germplasms	Other	Mono-isolate	Keygene N.V.	PRJEB40739	08.12.2020	high-throughput re-sequencing	473 Gbases; 0.18 Tbytes
4	A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae	Other	Multi-specis	Smithsonian Institution	PRJNA540287	29.04.2019	target capture	166 Gbases; 91148 Mbytes
5	Taraxacum koksaghyz Rodin roots	Raw sequence reads	Multi-species	Shihezi University	PRJNA539838	26.04.2019	none	99 Gbases; 61698 Mbytes
6	Population genetics of rubber dandelion (Taraxacum koksaghyz)	Raw sequence reads	Multi-species	University of Tennessee, Institute of Agriculture	PRJNA505305	13.11.2018	de novo MiSeq	11 Gbases; 6110 Mbytes
7	Taraxacum koksaghyz R.	Transcriptome or Gene expression	Multi-species	ENEA	PRJNA413689	09.10.2017	none	491 Mbases; 1344 Mbytes
8	Taraxacum koksaghyz Assembly	Assembly	Multi-isolate	ENEA	PRJNA413688	09.10.2017	cDNA assembly	no public data
9	Taraxacum koksaghyz Transcriptome or Gene expression	Transcriptome or Gene expression	Mono-isolate	OSU	PRJNA378120	05.03.2017	RNAseq transcriptome analysis	GFJE00000000.1 shotgun assembly
10	Taraxacum koksaghyz Raw sequence reads	Raw sequence reads	Multi-isolate	OSU	PRJNA361575	17.01.2017	RNAseq transcriptome analysis: 6 SRA	36 Gbases; 26247 Mbytes

No	Description	Project data type	Scope	Institute	Accession	Submission date	Additional information	Data details
							experiment data	
11	Taraxacum koksaghyz RefSeq Genome	RefSeq Genome	Mono-isolate	NCBI	PRJNA356593	07.12.2016	genome reference project	NC_032057.1; chloroplast genome; 151338 bp
12	Taraxacum koksaghyz Raw sequence reads	Raw sequence reads	Mono-isolate	Korea Research Institute of Bioscience and Biotechnology	PRJNA310379	01.02.2016	RNASeq	no public data
13	Asteraceae Raw sequence reads	Raw sequence reads	Multi-species	University of British Columbia	PRJNA288472	29.06.2015	survey sequencing of Asteraceae to assess genome composition	29 Gbases; 19012 Mbytes

Tab. 6-4 List of exemplary crop lines with the corresponding IMI resistance conferring mutation. Based on Tan *et al.* (*136*).

Crop/line	**Mutation AthAHAS no.**
Z. mays/XI12	Ser653Asn
Z. mays/XA17	Trp574Leu
O. sativa/93AS3510	Gly654Glu
T. aestivum/TEAIIMI 11A	Ser653Asn
B. napus/PM1	Ser653Asn
H. annuus/Two	Ala205Val

Tab. 6-5 Overview of AA substitutions in AHAS that confer herbicide resistance to at least one of the five classes of AHAS inhibitors. Only responses to SU and IMI are shown with referring numbers of resistant species. Based on Tranel et al. (*141*). S = susceptible; r = moderate resistance (< 10-fold relative to sensitive biotype); R = high resistance (> 10-fold); ND = not determined.

AA position AthAHAS no.	**AA substitution**	**Response towards SU**	**Number of resistant species**	**Response towards IMI**	**Number of resistant species**
Ala122	Asn	R	1	R	1
	Thr	ND/S	0	ND/R	6
	Tyr	R	1	R	1
	Val	ND/R	1	ND/R	1
Pro197	Ala	R	11	ND/S/r	1
	Arg	R	5	ND/S/r	1
	Asn	R	1	ND	0
	Gln	R	7	ND/S/r	1
	Glu	R	1	R	1
	His	R	9	ND/S/R/r	4
	Ile	R	1	r	1
	Leu	R	13	ND/S/R/r	7
	Ser	ND/R/r	26	ND/S/r	4
	Thr	R/r	14	ND/S/r	6
	Tyr	R	1	ND	0
Ala205	Phe	R	1	R	1
	Val	S/R/r	4	R/r	5
Asp376	Glu	ND/R/r	11	ND/R/r	8
Arg377	His	R	1	ND	0
Trp574	Arg	R	1	R	1
	Gly	R	1	ND	0
	Leu	ND/R/r	36	ND/R	29
	Met	R	1	ND/R	0
Ser653	Asn	ND/S/r	2	ND/R	6
	Ile	r	1	R	1
	Thr	ND/S/r	1	R	6
Gly654	Asp	r	1	R	1
	Glu	ND	0	R	1

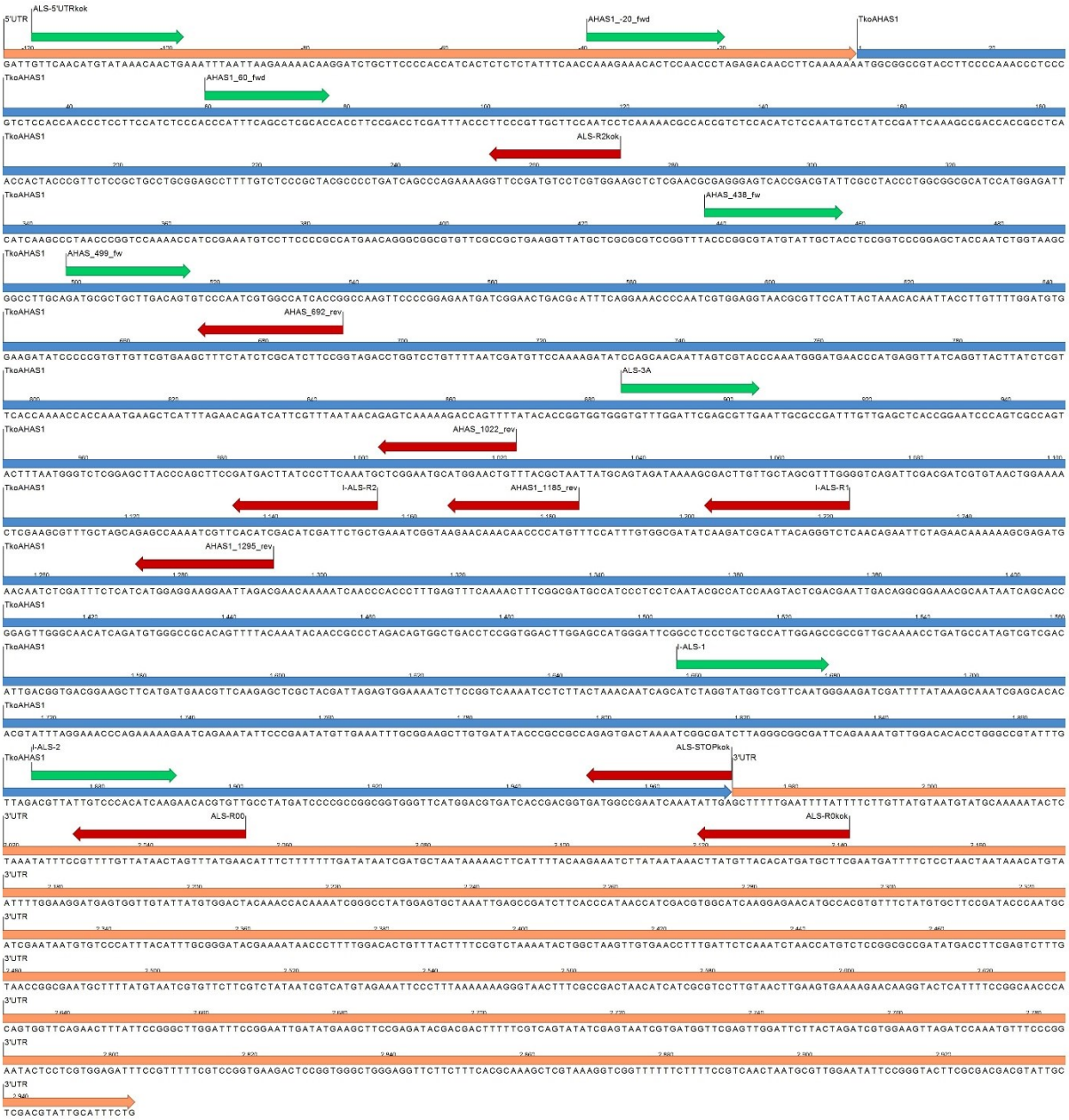

Fig. 6-2 The primer (Tab. 2-1) used in this work with reference to *TkoAHAS1* for better understanding. It is shown the CDS of *TkoAHAS1* allele sequence I (according to Tab. 3-2; Annex 6-2) (blue) with 5' and 3' regions (orange). Forward (green) and reverse (red) primer are visualized by arrows.

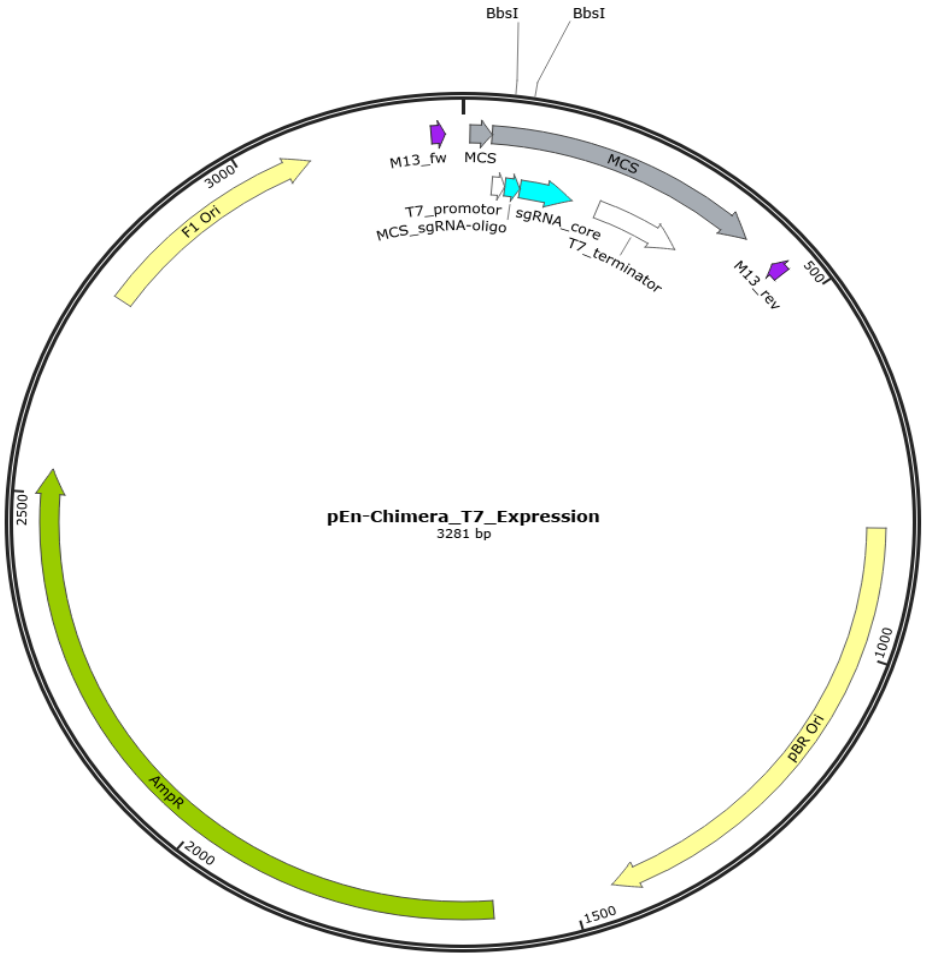

Fig. 6-3 Map of pEn-Chimera-T7_Expression used for the method mentioned in chapter 2.2.7. Figure created with SnapGene. Abbreviations: MCS: multiple cloning site; MCS_sgRNA-oligo: MCS for the designed oligos for the SgRNA of interest; sgRNAcore: sgRNA backbone; Ori: origin of replication; AmpR: Ampicillin Resistance Cassette.

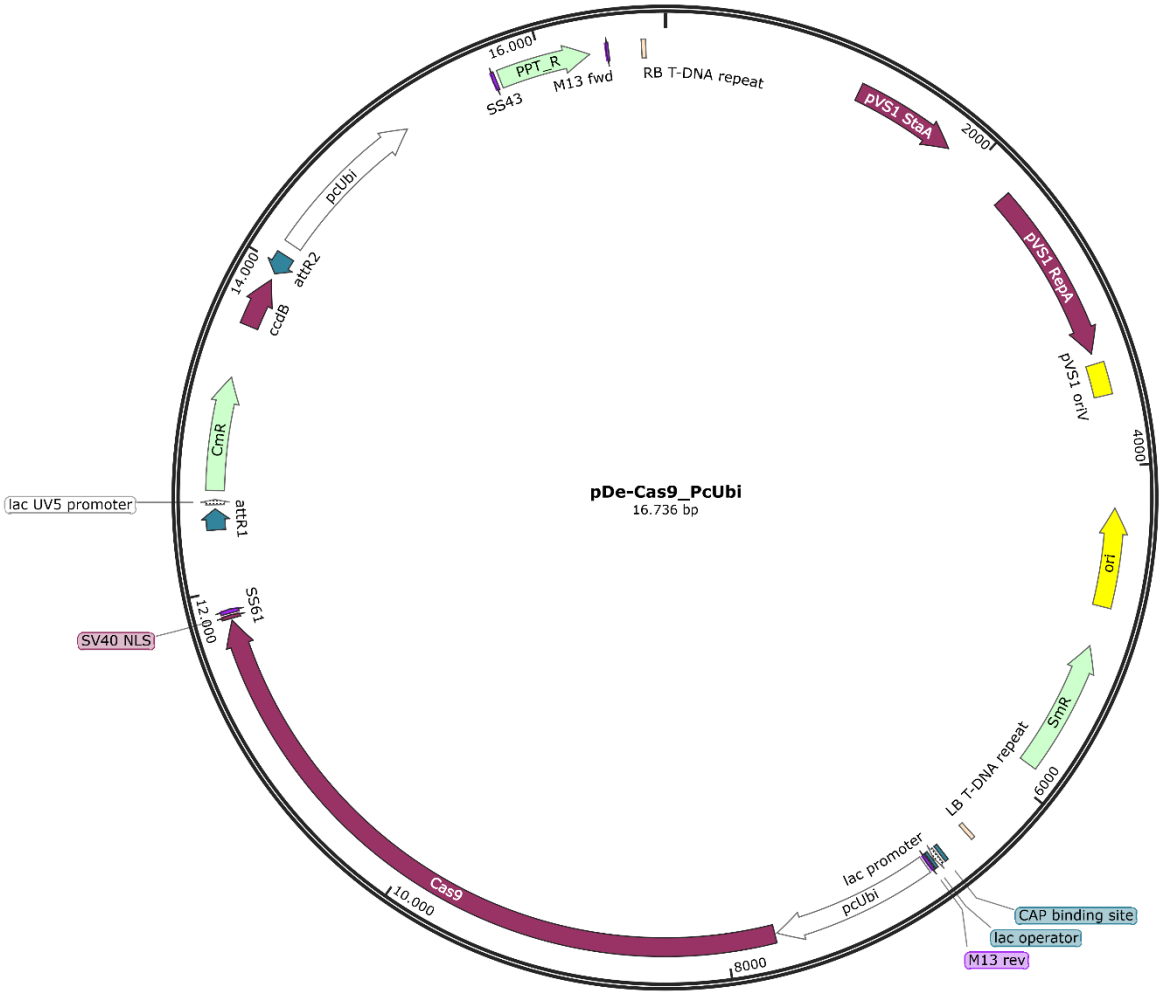

Fig. 6-4 Map of pDe-Cas9_PcUbi used for the method mentioned in chapter 2.2.4. Figure created with SnapGene. Abbreviations: pVS1 StaA, pVS1 RepA, pVS1 oriV: features for replication of vector in *A. tumefaciens*; Ori: origin of replication; SmR: streptomycin resistance cassette; LB/ RB T-DNA repeat; left/ right border of T-DNA; lac promoter: Promoter from lac operon; pcUbi: Plant promoter from maize ubiquitin gene; Cas9: sequence encoding Cas9 nuclease; CmR: chloramphenicol-resistance cassette; attR1, attR2: borders for Gateway® cloning; ccdB: codes for the toxic protein (CcdB), is a tool for Gateway® cloning; PPT_R: phosphinothricin acetyltransferase (resistance cassette).

Fig. 6-5 Map of the vector obtained by the cloning method described in chapter 2.2.4 for targeting the nt position 1091 in *TkoAHAS1* corresponding to AA position 363. Figure created with SnapGene. Abbreviations: pVS1 StaA, pVS1 RepA, pVS1 oriV: features for replication of vector in *A. tumefaciens*; Ori: origin of replication; SmR: streptomycin resistance cassette; LB/ RB T-DNA repeat; left/ right border of T-DNA; lac promoter: Promoter from lac operon; pcUbi: Plant promoter from maize ubiquitin gene; Cas9: sequence encoding Cas9 nuclease; protospacer_sg363: region for the protospacer sequence used to target the gene region encoding for TkoAHAS1 AA position 363; gRNA scaffold: backbone of the sgRNA construct (without protospacer region); PPT_R: phosphinothricin acetyltransferase (resistance cassette).

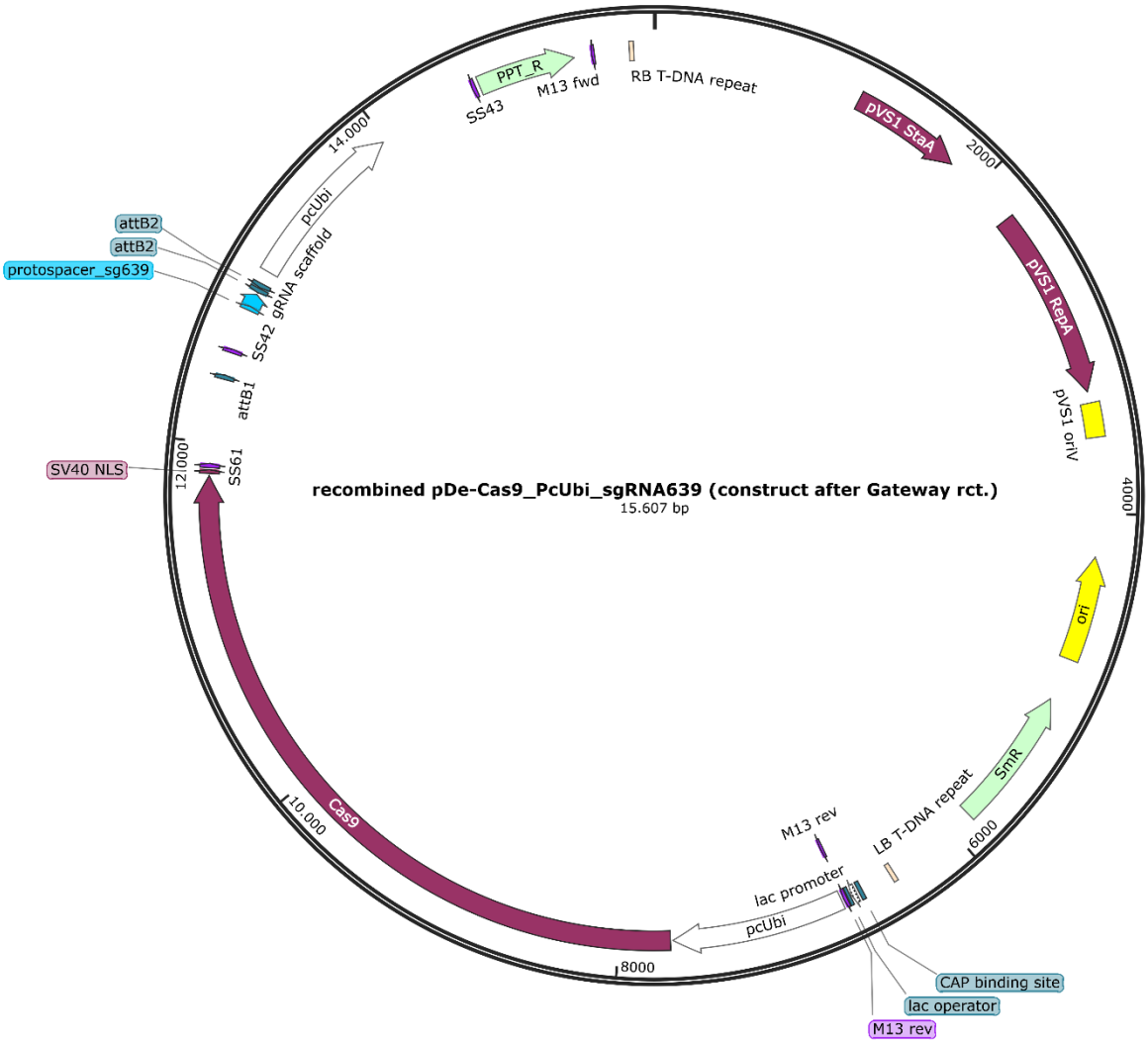

Fig. 6-6 Map of the vector obtained by the cloning method described in chapter 2.2.4 for targeting nt position 1916 in *TkoAHAS1* corresponding to AA position 639. Figure created with SnapGene. Abbreviations: pVS1 StaA, pVS1 RepA, pVS1 oriV: features for replication of vector in *A. tumefaciens*; Ori: origin of replication; SmR: streptomycin resistance cassette; LB/ RB T-DNA repeat; left/ right border of T-DNA; lac promoter: Promoter from lac operon; pcUbi: Plant promoter from maize ubiquitin gene; Cas9: sequence encoding Cas9 nuclease; protospacer_sg639: region for the protospacer sequence used to target the gene region encoding for TkoAHAS1 AA position 639; gRNA scaffold: backbone of the sgRNA construct (without protospacer region); PPT_R: phosphinothricin acetyltransferase (resistance cassette).

Tab. 6-6 Transcriptome set “INVIEW Transcriptome Discover” from 29.12.2015 (chapter 2.2.17.1): list of samples (datasets) and parameters of the raw data.

Dataset	Sequenced bases	Sequenced reads
***T. koksaghyz* leaf**	8,029,050,000	32,116,200
***T. koksaghyz* flower**	11,593,900,000	46,375,600

Tab. 6-7 Transcriptome Set “TruSeq Stranded mRNA” from 26.03.2018 (chapter 2.2.17.2): list of samples (datasets) and parameters of the raw data.

Dataset number	Sample information			Sequenced bases	Sequenced reads
	Plant	Individuals/ sample	Treatment		
1	M_2-1	1	Imazamox	3,711,131,072	36,743,872
2	M_2-1	1		3,574,152,448	35,387,648
3	M_2-2	1		3,522,371,768	34,874,968
4	M_2-2	1		4,556,669,742	45,115,542
5	M_2-3	1		4,461,659,850	44,174,850
6	M_2-3	1		3,166,558,464	31,352,064
7	Field mix	20		4,776,742,682	47,294,482
8	Field mix	20		3,108,539,014	30,777,614
9	M_2-1	1	Water	3,429,205,732	33,952,532
10	M_2-1	1		3,493,865,528	34,592,728
11	M_2-2	1		3,310,461,850	32,776,850
12	M_2-2	1		3,523,446,408	34,885,608
13	M_2-3	1		4,051,758,420	40,116,420
14	M_2-3	1		4,008,774,032	39,690,832
15	Field mix	20		3,924,788,694	38,859,294
16	Field mix	20		3,618,665,572	35,828,372

Tab. 6-8 References for analysis of *H. annuus* AHAS I, II and III relationships.

H. annuus	**GenBank Reference CDS**	**GenBank Reference Protein**
AHAS I	AY541452.1	AAT07323.1
AHAS II	AY541456.1	AAT07327.1
AHAS III	AY541458.1	AAT07329.1

Tab. 6-9 Details on selected M_5 progenies from 6 different crossings within the M_4 population (from Tab. 3-12).

crossing no.	♀	♂	**number of M_5 plants**	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])		
				WT (C/C)	**Hetero-zygous mutant (C/T)**	**Homo-zygous mutant (T/T)**
19	M_4-(14x10)-14	M_4-(10x14)-12	4	3/ 75.0	1/ 25.0	0/ 0
59A	M_4-(14x10)-19	M_4-(14x10)-26	7	4/ 57.1	3/ 42.9	0/ 0
59B	M_4-(14x10)-26	M_4-(14x10)-19	3	1/ 33.3	1/ 33.3	1/ 33.3
122	M_4-(10x14)-5	M_4-(10x14)-4	16	1/ 6.3	12/ 75.0	3/ 18.8

crossing no.	♀	♂	number of M_5 plants	***AHAS1* genotype at nt position 572** Amount of plants (absolute/ percentage [%])		
				WT (C/C)	**Hetero-zygous mutant (C/T)**	**Homo-zygous mutant (T/T)**
281A	M_4-(10x14)-4	M_4-(14x10)-26	5	0/ 0	4/ 80.0	1/ 20.0
281B	M_4-(14x10)-26	M_4-(10x14)-4	11	5/ 45.5	6/ 54.5	0/ 0

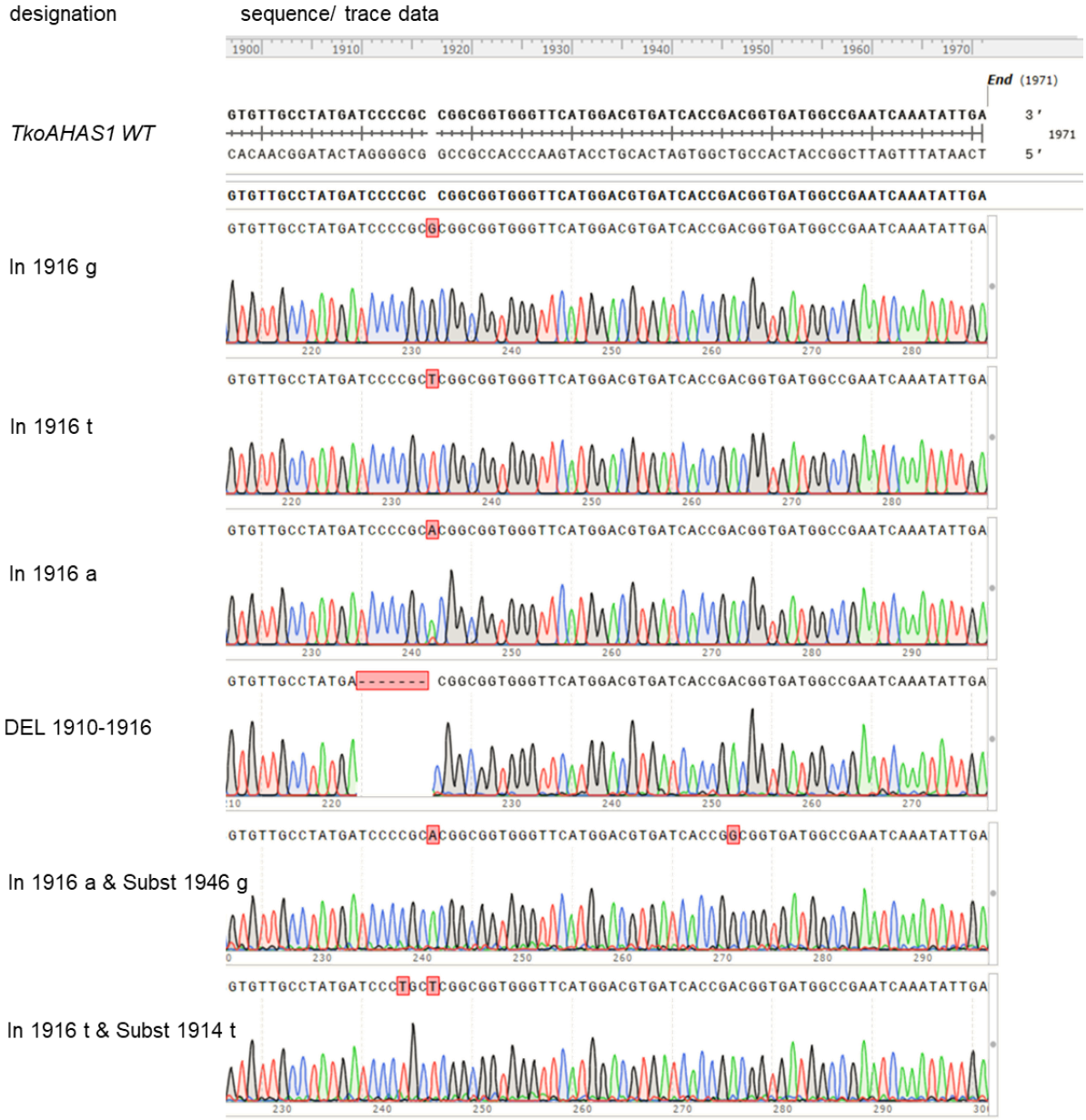

Fig. 6-7 Exemplary DNA sequences including the trace data of the detected mutations in T_1 plants from the CRISPR/Cas9 approach targeting nt-position 1916 (corresponds to TkoAHAS1 Ala639) aligned to the *TkoAHAS1 WT* sequence. Figure created with SnapGene.

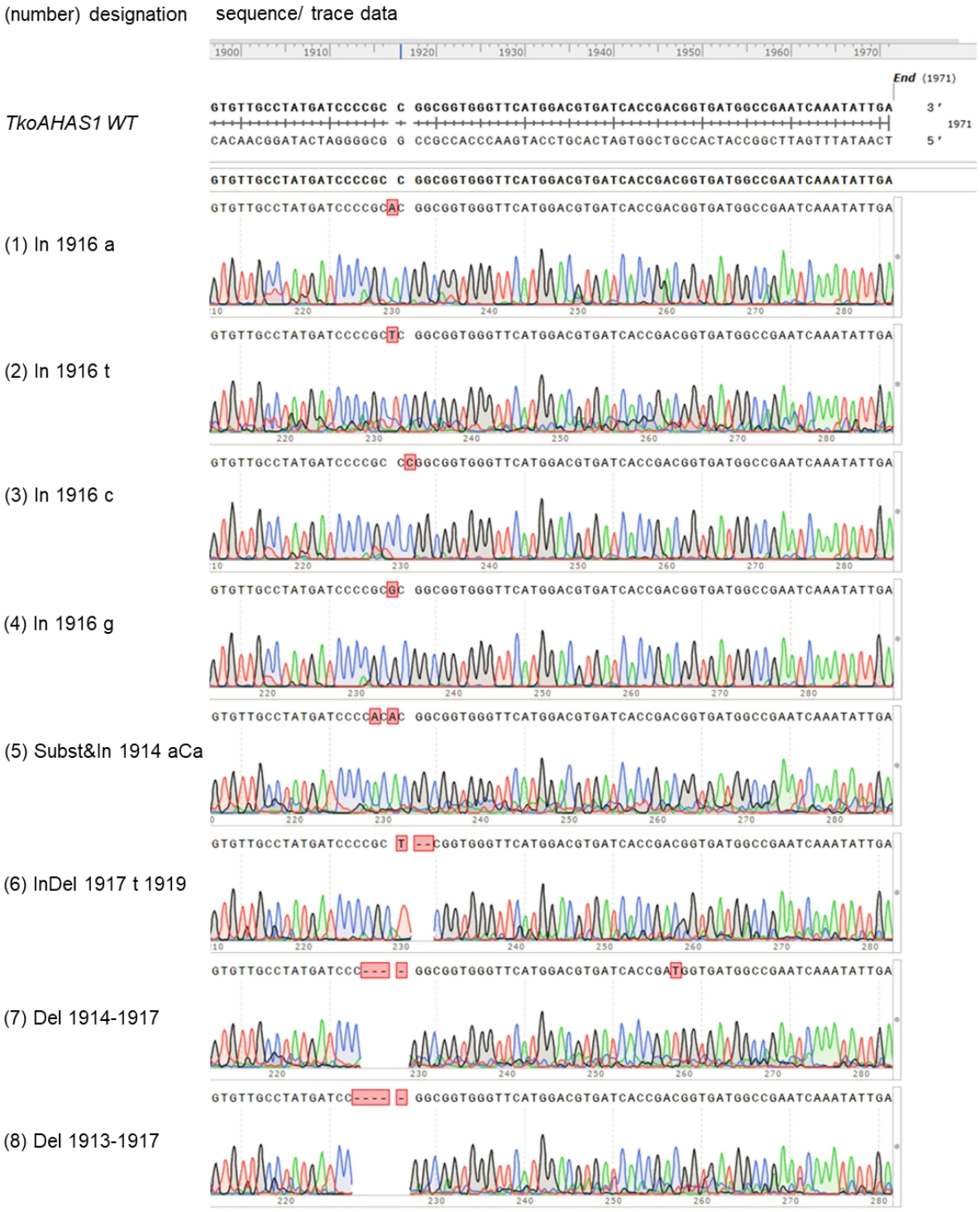

Fig. 6-8 Exemplary sequences including the trace data of the detected mutations in F_1 plants from the CRISPR/Cas9 approach targeting nt-position 1916 (corresponds to TkoAHAS1 Ala639) aligned to the *TkoAHAS1 WT* sequence. Figure created with SnapGene.

Tab. 6-10 Gene IDs of the 25 transcripts exclusively occurring and differentially expressed in M_2-3, listed with the respective log2(FC) value (Imazamox treated in relation to untreated; negative = down-regulation; positive = up-regulation), the Blast2GO annotation and the matching GO-Terms.

Gene ID	log2(FC)			Blast2GO annotation	GO-TERM
	M_2-1	M_2-2	M_2-3		
Gene.111970::145060::g.111970::m.111970.exon1	-	-	4.117	cytochrome p450	GO:0004497, GO:0005506, GO:0016021, GO:0016705, GO:0020037, GO:0055114
Gene.115073::148708::g.115073::m.115073.exon1	-	-	-2.042	cytochrome p450 78a5-like	GO:0005506, GO:0016021, GO:0020037, GO:0033772, GO:0044550, GO:0055114
Gene.12171::15044::g.12171::m.12171.exon1	-	-	-2.119	cytochrome p450	GO:0004497, GO:0005506, GO:0016021, GO:0016705, GO:0020037, GO:0055114
Gene.142645::184474::g.142645::m.142645.exon1	-	-	3.111	cytochrome p450 cyp72a219-like	GO:0004497, GO:0005506, GO:0016021, GO:0016705, GO:0020037, GO:0055114
Gene.57283::74240::g.57283::m.57283.exon1	-	-	2.219	cytochrome p450	GO:0004497, GO:0005506, GO:0016021, GO:0016705, GO:0020037, GO:0055114
Gene.57295::74262::g.57295::m.57295.exon1	-	-	2.280	cytochrome p450	GO:0004497, GO:0005506, GO:0016021, GO:0016705, GO:0020037, GO:0055114
Gene.124191::160752::g.124191::m.124191.exon1	-	-	3.073	glutathione s-transferase f9-like	GO:0004364
Gene.152629::198642::g.152629::m.152629.exon1	-	-	5.010	probable glutathione s-transferase	GO:0004364, GO:0004462, GO:0005737, GO:0006749, GO:0009407
Gene.152641::198653::g.152641::m.152641.exon1	-	-	5.846	probable glutathione s-transferase parc	GO:0004364, GO:0005737, GO:0006749, GO:0009407
Gene.74068::95831::g.74068::m.74068.exon1	-	-	5.602	glutathione s-transferase	GO:0016740
Gene.115223::148914::g.115223::m.115223.exon1	-	-	2.715	peroxidase n1	GO:0004601, GO:0005576, GO:0006979, GO:0009505, GO:0009664, GO:0016021, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.115225::148917::g.115225::m.115225.exon1	-	-	3.269	peroxidase n1	GO:0004601, GO:0005576, GO:0006979, GO:0009505, GO:0009664, GO:0016021, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.5553::6790::g.5553::m.5553.exon1	-	-	4.846	peroxidase n1-like	GO:0004601, GO:0005576, GO:0006979, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.5560::6795::g.5560::m.5560.exon1	-	-	3.012	heme peroxidase	GO:0004601, GO:0005576, GO:0006979, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869

Gene ID	log2(FC)			Blast2GO annotation	GO-TERM
	M_2-1	M_2-2	M_2-3		
Gene.5564::6799:: g.5564::m.5564.exon1	-	-	2.993	heme peroxidase	GO:0004601, GO:0005576, GO:0006979, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.80226::103690:: g.80226::m.80226.exon1	-	-	-2.273	peroxidase 42	GO:0004601, GO:0005576, GO:0006979, GO:0009505, GO:0009664, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.80228::103692:: g.80228::m.80228.exon1	-	-	-2.302	peroxidase 42	GO:0004601, GO:0005576, GO:0006979, GO:0020037, GO:0042744, GO:0046872, GO:0055114, GO:0098869
Gene.135775::175077:: g.135775::m.135775.exon1	-	-	2.565	udp-glycosyltransferase 71e1	GO:0009813, GO:0016020, GO:0043231, GO:0052696, GO:0080043, GO:0080044
Gene.155828::202616:: g.155828::m.155828.exon1	-	-	2.714	probable glycosyltransferase at5g03795	GO:0016021, GO:0050508
Gene.159460::206885:: g.159460::m.159460.exon1	-	-	2.761	udp-glucuronosyl udp-glucosyltransferase	GO:0009813, GO:0043231, GO:0052696, GO:0080043, GO:0080044
Gene.159476::206896:: g.159476::m.159476.exon1	-	-	2.641	udp-glucuronosyl udp-glucosyltransferase	GO:0008152, GO:0016758
Gene.171482::223123:: g.171482::m.171482.exon1	-	-	2.063	7-deoxyloganetic acid glucosyltransferase-like	GO:0008152, GO:0016758
Gene.59442::76796:: g.59442::m.59442.exon1	-	-	3.046	udp-glucuronosyl udp-glucosyltransferase	GO:0008152, GO:0016758
Gene.166973::216701:: g.166973::m.166973.exon1	-	-	2.255	abc transporter b family member 25	GO:0005524, GO:0016021, GO:0042626, GO:0055085
Gene.30796::38551:: g.30796::m.30796.exon1	-	-	2.829	abc transporter c family member 3-like	GO:0005524, GO:0005737, GO:0006535, GO:0009001, GO:0016021, GO:0042626, GO:0055085

Tab. 6-11 Results of the SNP call for SNPs that did not occur in M_2-1/ 2 but in M_2-3 for the 25 DE transcripts of M_2-3 referring to Tab. 3-21: 32 SNPs in ten of the 25 transcripts were identified. Referring Gene ID, log2(FC), Blast2GO annotation, number of SNPs and of non-silent SNPs thereof are listed. The results of SNP analysis are with regard to a respective reference sequence, identified via blastx against land plants (blastx results are listed).

Gene ID	log2(FC) M_2-3	Blast2GO annotation	SNPs (total)	Non-silent SNPs	BLASTX result				
					Reference for SNP analysis	Acc.no.	Query cover	E value	Per. Ident
Gene.111970::145060::g.111970::m.111970.exon1	4.12	cytochrome p450	4	1	cytochrome P450 CYP72A219-like [Lactuca sativa]	XP_023767467.1	93.0 %	4E-100	85.9 %
Gene.57283::74240::g.57283::m.57283.exon1	2.22	cytochrome p450	7	5	cytochrome P450 CYP72A219-like isoform X1 [Lactuca sativa]	XP_023736509.1	60.0 %	6E-79	84.4 %
Gene.57295::74262::g.57295::m.57295.exon1	2.28	cytochrome p450	4	2	cytochrome P450 CYP72A219-like isoform X1 [Lactuca sativa]	XP_023736509.1	99.0 %	6E-89	84.2 %
Gene.5553::6790::g.5553::m.5553.exon1	4.85	peroxidase n1-like	6	3	peroxidase N1-like [Lactuca sativa]	XP_023748165.1	72.0 %	0	84.2 %
Gene.5564::6799::g.5564::m.5564.exon1	2.99	heme peroxidase	1	0	hypothetical protein LSAT_9X85400 [Lactuca sativa]	PLY62876.1	61.0 %	4E-40	81.5 %
Gene.80226::103690::g.80226::m.80226.exon1	-2.27	peroxidase 42	1	1	peroxidase 42 [Lactuca sativa]	XP_023742891.1	71.0 %	2E-174	98.8 %
Gene.80228::103692::g.80228::m.80228.exon1	-2.30	peroxidase 42	1	0	peroxidase 42 [Lactuca sativa]	XP_023742891.1	73.0 %	4E-38	92.9 %
Gene.135775::175077::g.135775::m.135775.exon1	2.56	udp-glycosyl-transferase 71e1	4	3	UDP-glycosyltransferase 71E1-like [Lactuca sativa]	XP_023770763.1	75.0 %	0	86.7 %
Gene.159476::206896::g.159476::m.159476.exon1	2.64	udp-glucuronosyl udp-glucosyl-transferase	3	1	UDP-glycosyltransferase 83A1-like [Lactuca sativa]	XP_023732388.1	73.0 %	2E-92	68.6 %
Gene.59442::76796::g.59442::m.59442.exon1	3.05	udp-glucuronosyl udp-glucosyl-transferase	1	1	UDP-glycosyltransferase 73C5-like [Lactuca sativa]	XP_023745919.1	77.0 %	5E-76	88.0 %

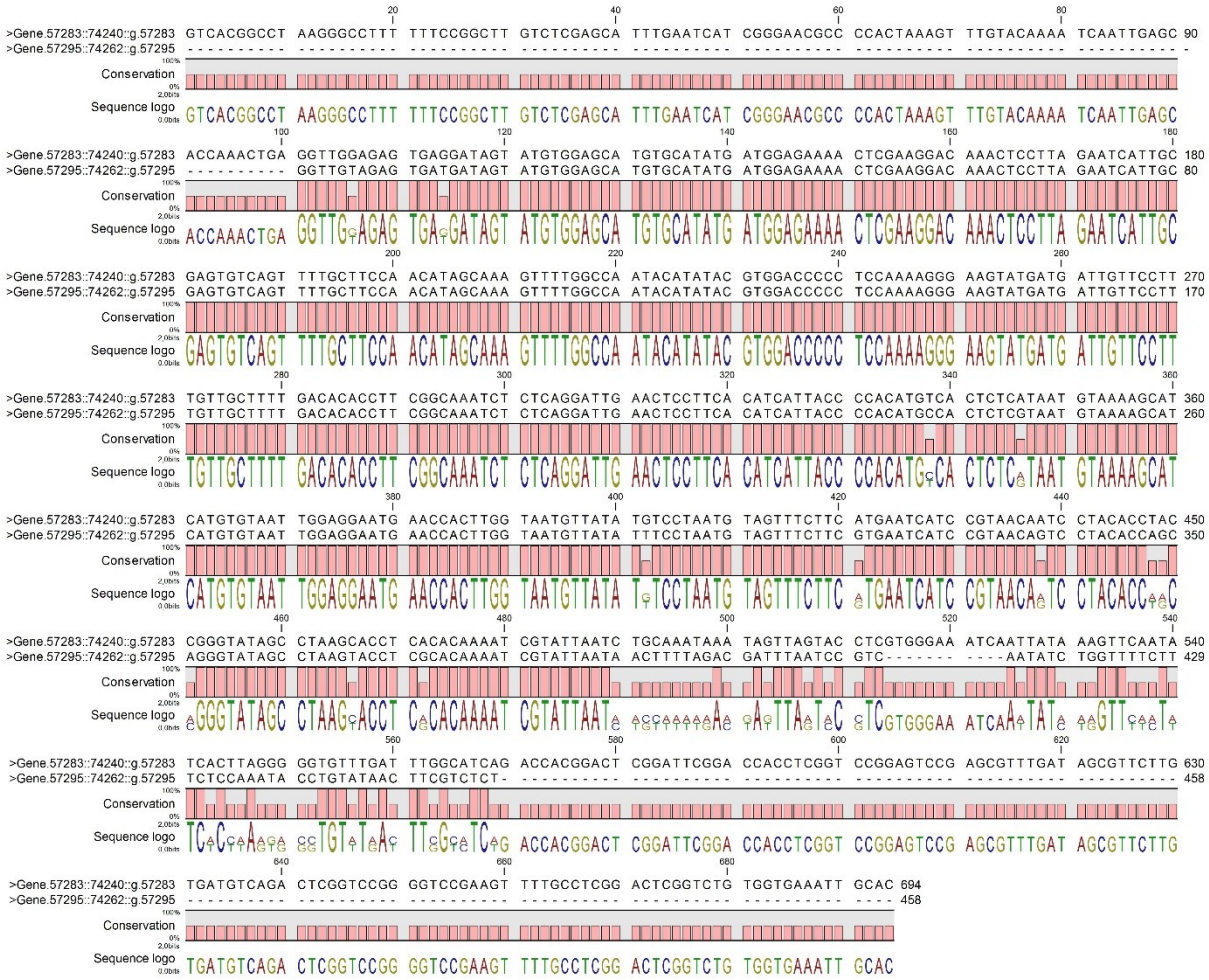

Fig. 6-9 Alignment of the two transcripts Gene.57283::74240::g.57283(::m.57283.exon1) and Gene.57295::74262: :g.57295(::m.57295.exon1).

Fig. 6-10 Selection of the alignment shown in Fig. 6-9. SNP positions mentioned in Tab. 3-22 are marked with a blue triangle.

Tab. 6-12 List of the four interesting SNPs (Tab. 3-22) with the referring nt characters found in the data of M_2-3 and *T. koksaghyz* field mix and resulting AA characters together with the corresponding AA positions and characters of the reference protein (Blast2GO annotation) (Tab. 6-11).

Transcript (Gene ID)	**SNP position**	**Detected SNPs in M_2-3 and field mix**		**Reference protein XP_023736509.1 cytochrome P450 CYP72A219-like isoform X1 [*Lactuca sativa*]**	
		nt 0/1	**AA 0/1**	**AA position**	**AA character**
Gene.57283:: 74240::g.57283 ::m.57283.exon1	164	C/ A	Ser/ Ile	488	Gly
	343	C/ A	Glu/ Asp	428	Asp
	349	A/ T	His/ Gln	426	His
Gene.57295:: 74262::g.57295 ::m.57295.exon1	64	C/ A	Ser/ Ile	488	Gly

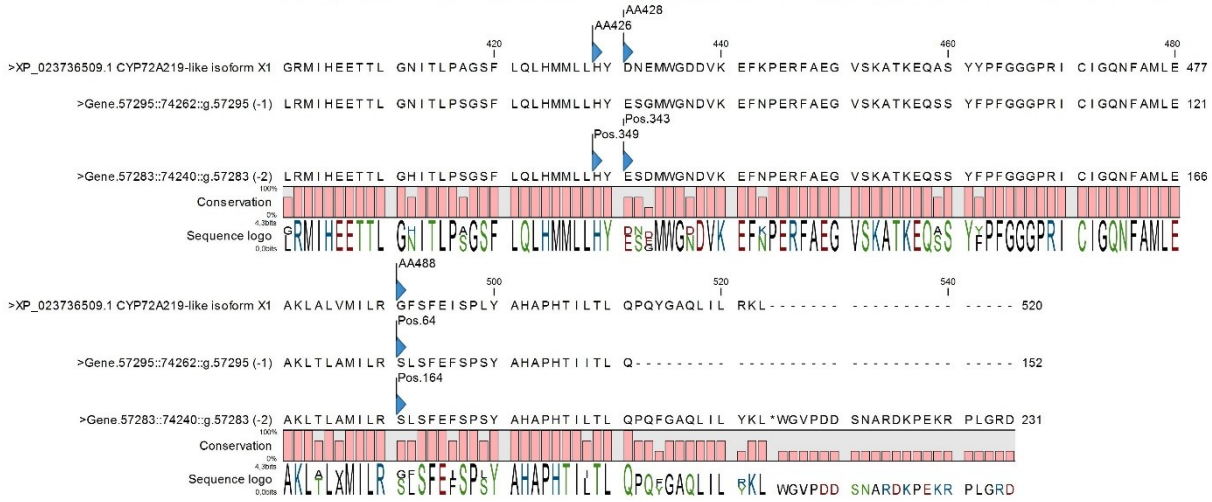

Fig. 6-11 Selection of the protein alignment of the reference protein cytochrome P450 CYP72A219-like isoform X1 from *Lactuca sativa* (Tab. 6-11) and the translated transcripts (Fig. 6-9).

Tab. 6-13 List of the four interesting SNPs (Tab. 3-22) with the referring nt characters found in the data of M_2-3 and field mix together with the detected absolute counts for the characters in the data sets. The SNP at position 164 from transcript Gene.57283::74240::g.57283::m.57283.exon1 and the SNP at position 64 from transcript Gene.57295::74262::g.57295 m.57295.exon1 describe the same position (Tab. 6-12) – therefore latter one is shown in grey. This yields two haplotypes (CCA (reference) and AAT (alternative)) for M_2-3, which are both recovered in the field mix, too.

Transcript (Gene ID)	**SNP position**	**Detected SNPs in M_2-3 and field mix**		**Absolute counts for 0/1**	
		nt 0/1	**AA 0/1**	**M_2-3**	**Field mix**
Gene.57283:: 74240::g.57283 ::m.57283.exon1	164	C/ A	Ser/ Ile	712/ 206	2272/ 18
	343	C/ A	Glu/ Asp	210/ 180	2614/ 14
	349	A/ T	His/ Gln	213/ 183	2526/ 13
Gene.57295:: 74262::g.57295 m.57295.exon1	64	C/ A	Ser/ Ile	331/ 147	1577/ 14

Annex 6-1 The CDS (5' - 3') of the preliminary *TkoAHAS1* allele 1 from the working group of F. Hartung (JKI Quedlinburg).

ATGGCGGCCGTACCTTCCCCAAACCCTCCCGTCTCCACCAACCCTCCTTCCATCTCCCA
CCCATTTCAGCCTCGCACCACCTTCCGACCTCGATTTACCCTTCTCGTTGCTTCCAATCC
TCAAAAACGCCACCGTCTCCACATCTCCAATGTCCTATCCGATTCAAAGCCGACCACCG
CCTCAACCACTACCCGTTCTCCGCTGCCTGCGGAGCCTTTTGTCTCCCGCTACGCCCCT
GACCAGCCCAGAAAAGGTTCCGATGTCCTCGTGGAAGCTCTCGAACGCGAGGGAGTCA
CCGACGTATTCGCCTACCCTGGCGGCGCATCCATGGAGATTCATCAAGCCCTAACCCG
GTCCAAAACCATCCGAAATGTCCTTCCCCGCCACGAACAGGGCGGCGTGTTCGCCGCT
GAAGGTTATGCTCGCGCGTCCGGTTTACCCGGCGTATGTATTGCTACCTCCGGTCCCG
GAGCTACCAATCTGGTAAGCGGCCTTGCAGATGCGCTGCTTGACAGTGTCCCGATCGT
GGCCATCACCGGCCAAGTTCCCCGGAGAATGATCGGAACTGACGCATTTCAGGAAACC
CCAATCGTGGAGGTAACGCGTTCCATTACTAAACACAATTACCTTGTTTTGGATGTGGAA
GATATCCCCCGTGTTGTTCGTGAAGCTTTCTATCTCGCATCTTCCGGTAGACCTGGTCC

TGTTTTAATCGATGTTCCAAAAGATATCCAGCAACAATTAGTCGTACCCAAATGGGACGA
ACCCATGAGGTTATCAGGTTACTTATCTCGTTCACCAAAACCACCAAATGAAGCTCATTT
AGAACAGATCATTCGTTTAATAACAGAGTCAAAAAGACCAGTTTTATACACCGGTGGTGG
GTGTTTGGATTCGAGCGTTGAATTGCGCCGATTTGTTGAGCTCACCGGAATCCCAGTCG
CCAGTACTTTAATGGGTCTCGGAGCTTACCCAGCTTCCGATGACTTATCCCTTCAAATGC
TCGGAATGCATGGAACTGTTTACGCTAATTATGCAGTAGATAAAAGCGACTTGTTGCTAG
CGTTTGGGGTCAGATTCGACGATCGTGTAACCGGAAAACTCGAAGCGTTTGCTAGCAGA
GCCAAAATCGTTCACATCGACATCGATTCTGCTGAAATCGGTAAGAACAAACAACCCCA
TGTTTCCATTTGTGGCGATATCAAGATCGCATTACAGGGTCTCAACAGAATTCTAGAACA
AAAAAGCGAGATGAACAATCTCGATTTCTCATCATGGAGGAAGGAATTAGACGAACAAA
AATCAACCCACCCTTTGAGTTTCAAAACTTTCGGCGATGCCATCCCTCCTCAATACGCCA
TCCAAGTACTCGACGAATTGACAGGCGGAAACGCAATAATCAGCACCGGAGTTGGGCA
ACATCAGATGTGGGCCGCACAGTTTTACAAATACAACCGCCCTAGACAGTGGCTGACCT
CCGGTGGACTTGGAGCCATGGGATTCGGCCTCCCTGCTGCCATTGGAGCCGCCGTTGC
AAAACCTGATGCCATAGTCGTCGACATTGACGGTGACGGAAGCTTCATGATGAACGTTC
AAGAGCTCGCTACGATTAGAGTGGAAAATCTTCCGGTCAAAATCCTCTTACTAAACAATC
AGCATCTAGGTATGGTCGTTCAATGGGAAGATCGATTTTATAAAGCAAATCGAGCACAC
ACGTATTTAGGAAACCCAGAAAAAGAATCAGAAATATTCCCGAATATGTTGAAATTTGCG
GAAGCTTGTGATATACCCGCCGCCAGAGTGACTAAAATCGGCGATCTTAGGGCGGCGA
TTCAGAAAATGTTGGACACACCTGGGCCGTATTTGTTAGACGTTATTGTCCCACATCAAG
AACACGTGTTGCCTATGATCCCCGCCGGCGGTGGGTTCATGGACGTGATCACCGACGG
TGATGGCCGAATCAAATATTGA

Annex 6-2 The CDS (5’ - 3’) of *TkoAHAS1* allele sequence 1 (according to numbering given in Tab. 3-2).

ATGGCGGCCGTACCTTCCCCAAACCCTCCCGTCTCCACCAACCCTCCTTCCATCTCCCA
CCCATTTCAGCCTCGCACCACCTTCCGACCTCGATTTACCCTTCCCGTTGCTTCCAATC
CTCAAAAACGCCACCGTCTCCACATCTCCAATGTCCTATCCGATTCAAAGCCGACCACC
GCCTCAACCACTACCCGTTCTCCGCTGCCTGCGGAGCCTTTTGTCTCCCGCTACGCCC
CTGATCAGCCCAGAAAAGGTTCCGATGTCCTCGTGGAAGCTCTCGAACGCGAGGGAGT
CACCGACGTATTCGCCTACCCTGGCGGCGCATCCATGGAGATTCATCAAGCCCTAACC
CGGTCCAAAACCATCCGAAATGTCCTTCCCCGCCATGAACAGGGCGGCGTGTTCGCCG
CTGAAGGTTATGCTCGCGCGTCCGGTTTACCCGGCGTATGTATTGCTACCTCCGGTCCC
GGAGCTACCAATCTGGTAAGCGGCCTTGCAGATGCGCTGCTTGACAGTGTCCCAATCG
TGGCCATCACCGGCCAAGTTCCCCGGAGAATGATCGGAACTGACGCATTTCAGGAAAC
CCCAATCGTGGAGGTAACGCGTTCCATTACTAAACACAATTACCTTGTTTTGGATGTGGA
AGATATCCCCCGTGTTGTTCGTGAAGCTTTCTATCTCGCATCTTCCGGTAGACCTGGTC
CTGTTTTAATCGATGTTCCAAAAGATATCCAGCAACAATTAGTCGTACCCAAATGGGATG
AACCCATGAGGTTATCAGGTTACTTATCTCGTTCACCAAAACCACCAAATGAAGCTCATT
TAGAACAGATCATTCGTTTAATAACAGAGTCAAAAAGACCAGTTTTATACACCGGTGGTG
GGTGTTTGGATTCGAGCGTTGAATTGCGCCGATTTGTTGAGCTCACCGGAATCCCAGTC
GCCAGTACTTTAATGGGTCTCGGAGCTTACCCAGCTTCCGATGACTTATCCCTTCAAAT
GCTCGGAATGCATGGAACTGTTTACGCTAATTATGCAGTAGATAAAAGCGACTTGTTGCT
AGCGTTTGGGGTCAGATTCGACGATCGTGTAACTGGAAAACTCGAAGCGTTTGCTAGCA
GAGCCAAAATCGTTCACATCGACATCGATTCTGCTGAAATCGGTAAGAACAAACAACCC
CATGTTTCCATTTGTGGCGATATCAAGATCGCATTACAGGGTCTCAACAGAATTCTAGAA
CAAAAAAGCGAGATGAACAATCTCGATTTCTCATCATGGAGGAAGGAATTAGACGAACA
AAAATCAACCCACCCTTTGAGTTTCAAAACTTTCGGCGATGCCATCCCTCCTCAATACGC
CATCCAAGTACTCGACGAATTGACAGGCGGAAACGCAATAATCAGCACCGGAGTTGGG

CAACATCAGATGTGGGCCGCACAGTTTTACAAATACAACCGCCCTAGACAGTGGCTGAC
CTCCGGTGGACTTGGAGCCATGGGATTCGGCCTCCCTGCTGCCATTGGAGCCGCCGTT
GCAAAACCTGATGCCATAGTCGTCGACATTGACGGTGACGGAAGCTTCATGATGAACGT
TCAAGAGCTCGCTACGATTAGAGTGGAAAATCTTCCGGTCAAAATCCTCTTACTAAACAA
TCAGCATCTAGGTATGGTCGTTCAATGGGAAGATCGATTTTATAAAGCAAATCGAGCACA
CACGTATTTAGGAAACCCAGAAAAAGAATCAGAAATATTCCCGAATATGTTGAAATTTGC
GGAAGCTTGTGATATACCCGCCGCCAGAGTGACTAAAATCGGCGATCTTAGGGCGGCG
ATTCAGAAAATGTTGGACACACCTGGGCCGTATTTGTTAGACGTTATTGTCCCACATCAA
GAACACGTGTTGCCTATGATCCCCGCCGGCGGTGGGTTCATGGACGTGATCACCGACG
GTGATGGCCGAATCAAATATTGA

Annex 6-3 The CDS (5' - 3') of "TkoAHAS1_allSNPs": The sequence includes all identified SNP positions referring to Tab. 3-2. IUPAC ambigious codes: M = C/A, Y = C/T, S = G/C, R = A/G, W = A/T.

ATGGCGGCCGTAMCTTCCCCAAACCCTCCCGTCTCCACCAACCCTCCTTCCATCTCCCA
CCCATTTCAGCCTCGCACCACCTTCCGACCTCGATTTACCCTTCYCGTTGMTTCCAATC
CTCAAAAACGCCACCGTCTCCACATCTCCAATGTCSTATCCGATTCAAAGCCGACCACC
GCCTCAACCACTACCCGTTCTCCGCTGCCTGCGGAGCCTTTTGTCTCCCGCTACGCCC
CTGAYCAGCCCAGAAAAGGTTCCGATGTCCTCGTGGAAGCTCTCGAACGCGAGGGAGT
CACCGACGTATTCGCCTACCCTGGCGGCGCATCCATGGAGATTCATCAAGCCCTAACC
CGGTCCAAAACCATCCGAAATGTCCTTCCCCGCCATGAACAGGGCGGCGTGTTCGCCG
CTGAAGGYTATGCTCGCGCGTCCGGTTTACCCGGCGTATGTATTGCTACCTCCGGTCC
CGGAGCTACCAATCTGGTAAGCGGCCTTGCAGATGCGCTGCTTGACAGTGTCCCAATC
GTGGCCATCACCGGCCAAGTTCCCCGGAGAATGATCGGAACTGACGCATTTCAGGAAA
CCCCAATCGTGGAGGTAACGCGTTCCATTACTAAACACAATTACCTTGTTTTGGATGTGG
AAGATATCCCCCGTGTTGTTCGTGAAGCTTTCTATCTCGCATCTTCCGGTAGACCTGGT
CCTGTTTTAATCGATGTTCCAAAAGATATCCAGCAACAATTAGTCGTACCCAAATGGGAT
GAACCCATGAGGTTATCAGGTTACTTATCTCGTTCACCAAAACCACCAAATGAAGCTCAT
TTAGAACAGATCATTCGTTTAATAACAGAGTCAAAAAGACCAGTTTTATACACCGGTGGT
GGGTGTTTGGATTCGAGCGTTGAATTGCGCCGATTTGTTGAGCTCRCCGGAATYCCAGT
CGCCAGTACTTTAATGGGTCTCGGAGCTTACCCAGCTTCCGATGACTTATCCCTTCAAAT
GCTCGGAATGCATGGAACWGTTTACGCTAATTATGCAGTAGATAAAAGCGACTTGTTGC
TAGCGTTTGGGGTCAGATTCGACGATCGTGTAACYGGAAAACTCGAAGCGTTTGCTAGC
AGAGCCAAAATCGTTCACATCGACATCGATTCTGCTGAAATCGGTAAGAACAAACAACC
CCATGTTTCCATTTGTGGCGATATCAAGATCGCATTACAGGGTCTCAACAGAATTCTAGA
ACAAAAAAGCGAGATGAACAATCTCGATTTCTCATCATGGAGGAAGGAATTAGACGAAC
AAAAATCAACCCACCCTTTGAGTTTCAAAACTTTCGGCGATGCCATCCCTCCTCAATACG
CCATCCAAGTACTCGACGAATTGACAGSCGGAAACGCAATAATCAGCACCGGAGTTGG
GCAACATCAGATGTGGGCCGCACAGTTTTACAAATACAACCGCCCTAGACAGTGGCTGA
CCTCCGGTGGACTTGGAGCCATGGGATTCGGCCTCCCTGCTGCCATTGGAGCCGCCGT
TGCAAAACCTGATGCCATAGTCGTCGACATTGACGGTGACGGAAGCTTCATGATGAACG
TTCAAGAGCTCGCTACGATTAGAGTGGAAAATCTTCCGGTCAAAATCCTCWTACTAAAC
AATCAGCATCTAGGTATGGTCGTTCAATGGGAAGATCGATTTTATAAAGCAAATCGAGCA
CACACGTATTTAGGAAACCCAGAAAAAGAATCAGAAATATTCCCGAATATGTTGAAATTT
GCGGAAGCTTGTGATATACCCGCCGCCAGAGTGACTAAAATCGGCGATCTTAGGGCGG
CGATTCAGAAAATGTTGGACACACCTGGGCCGTATTTGTTAGACGTTATTGTCCCACATC
AAGAACACGTGTTGCCTATGATCCCCGCCGGCGGTGGGTTCATGGACGTGATCACCGA
CGGTGATGGCCGAATCAAATATTGA

Annex 6-4 The amino acid sequence according to “TkoAHAS1_allSNPS” (Annex 6-3). Ambigious code: X = different amino acids present at this position. AHAS conserved/ typical domains according to chapter 1.2.1 and Tab. 6-1: start of the mature plant protein (light grey), RHEQ-motif with catalytic Glu (grey), FAD-binding sites (underlined), ThDP binding site (framed).

MAAVXSPNPPVSTNPPSISHPFQPRTTFRPRFTLXVXSNPQKRHRLHISNVXSDSKPTTAST
TTRSPLPAEPFVSRYAPDQPRKGSDVLVEALEREGVTDVFAYPGGASMEIHQALTRSKTIRN
VLPRHEQGGVFAAEGYARASGLPGVCIATSGPGATNLVSGLADALLDSVPIVAITGQVPRRM
IGTDAFQETPIVEVTRSITKHNYLVLDVEDIPRVVREAFYLASSGRPGPVLIDVPKDIQQQLVV
PKWDEPMRLSGYLSRSPKPPNEAHLEQIIRLITESKRPVLYTGGGCLDSSVELRRFVELXGIP
VASTLMGLGAYPASDDLSLQMLGMHGTVYANYAVDKSDLLLAFGVRFDDRVTGKLEAFASR
AKIVHIDIDSAEIGKNKQPHVSICGDIKIALQGLNRILEQKSEMNNLDFSSWRKELDEQKSTHP
LSFKTFGDAIPPQYAIQVLDELTXGNAIISTGVGQHQMWAAQFYKYNRPRQWLTSGGLGAM
GFGLPAAIGAAVAKPDAIVVDIDGDGSFMMNVQELATIRVENLPVKILJLNNQHLGMVVQWE
DRFYKANRAHTYLGNPEKESEIFPNMLKFAEACDIPAARVTKIGDLRAAIQKMLDTPGPYLLD
VIVPHQEHVLPMIPAGGGFMDVITDGDGRIKY*

Annex 6-5 The CDS (5’- 3’) of “TkoAHAS1_fieldmix_allSNPs”: The sequence includes all identified SNP positions referring to Tab. 3-5. IUPAC ambigious codes: M = C/A, K = G/T, Y = C/T, S = G/C, R = A/G, W = A/T.

ATGGCGGCCGTAMCTTCCCCAAACCCTCCCGTCTCCACCAACCCTCCKTCCATCTCCCA
CCCATTTCAGCCTCGCACCACCTTCCGACCTCGATTTACCCTTCYCGTTGCTTCCAATCC
TCAAAAACGCCACCGTCTCCACATCTCCAATGTCSTATCCGATTCAAAGCCGACCACCG
CCTCAACCACTACCCGTTCTCCGCTGCCTGCGGAGCCTTTTGTCTCCCGCTACGCCCCT
GACCAGCCCAGAAAAGGTTCCGATGTCCTCGTGGAAGCTCTCGAACGCGAGGGAGTCA
CCGACGTATTCGCCTACCCTGGCGGCGCATCCATGGAGATTCATCAAGCCCTAACCCG
GTCCAAAACCATCCGAAATGTCCTTCCCCGCCAYGAACAGGGCGGCGTGTTCGCCGCT
GAAGGYTATGCTCGCGCGTCCGGTTTACCCGGCGTATGTATTGCTACCTCCGGTCCCG
GAGCTACCAATCTGGTAAGCGGCCTTGCAGATGCGCTGCTTGACAGTGTCCCRATCGT
GGCCATCACCGGCCAAGTTCCCCGGAGAATGATCGGAACTGACGCATTTCAGGAAACC
CCAATCGTGGAGGTAACGCGTTCCATTACTAAACACAATTACCTTGTTTTGGATGTGGAA
GATATCCCCCGTGTTGTTCGTGAAGCTTTCTATCTCGCATCTTCCGGTAGACCTGGTCC
TGTTTTAATCGATGTTCCAAAAGATATCCAGCAACAATTAGTCGTACCCAAATGGGAYGA
ACCCATGAGGTTATCAGGTTACTTATCTCGTTCACCAAAACCACCAAATGAAGCTCATTT
AGAACAGATCATTCGTTTAATAACAGAGTCAAAAAGACCAGTTTTATACACCGGTGGTGG
GTGTTTGGATTCGAGCGTTGAATTGCGCCGATTTGTTGAGCTCRCCGGAATYCCAGTCG
CCAGTACTTTAATGGGTCTCGGAGCTTACCCAGCTTCCGATGACTTATCCCTTCAAATGC
TCGGAATGCATGGAACWGTTTACGCTAATTATGCAGTAGATAAAAGCGACTTGTTGCTA
GCGTTTGGGGTCAGATTCGACGATCGTGTAACYGGAAAACTCGAAGCGTTTGCTAGCA
GAGCCAAAATCGTTCACATCGACATCGATTCTGCTGAAATCGGTAAGAACAAACAACCC
CATGTTTCCATTTGTGGCGATATCAAGATCGCATTACAGGGTCTCAACAGAATTCTAGAA
CAAAAAAGCGAGATGAACAATCTCGATTTCTCATCATGGAGGAAGGAATTAGACGAACA
AAAATCAACCCACCCTTTGAGTTTCAAAACTTTCGGCGATGCCATCCCTCCTCAATACGC
CATCCAAGTACTCGACGAATTGACAGSCGGAAACGCAATAATCAGCACCGGAGTTGGG
CAACATCAGATGTGGGCCGCACAGTTTTACAAATACAACCGCCCTAGACAGTGGCTGAC
CTCCGGTGGACTTGGAGCCATGGGATTCGGCCTCCCTGCTGCCATTGGAGCCGCCGTT
GCAAAACCTGATGCCATAGTCGTCGACATTGACGGTGACGGAAGCTTCATGATGAACGT
TCAAGAGCTCGCTACGATTAGAGTGGAAAATCTTCCGGTCAAAATCCTCWTACTAAACA
ATCAGCATCTAGGTATGGTCGTTCAATGGGAAGATCGATTTTATAAAGCAAATCGAGCAC

ACACGTATTTAGGAAACCCAGAAAAAGAATCAGAAATATTCCCGAATATGTTGAAATTTG
CGGAAGCTTGTGATATACCCGCCGCCAGAGTGACTAAAATCGGCGATCTTAGGGCGGC
GATTCAGAAAATGTTGGACACACCTGGGCCGTATTTGTTAGACGTTATTGTCCCACATCA
AGAACACGTGTTGCCTATGATCCCCGCCGGCGGTGGGTTCATGGACGTGATCACCGAC
GGTGATGGCCGAATCAAATATTGA

7 References

1. P. F. Stevens, *Angiosperm Phylogeny Website.* (2001 onwards) (available at http://www.mobot.org/MOBOT/research/APweb/).
2. S. Volis, K. Uteulin, D. Mills, Russian dandelion (Taraxacum kok-saghyz): one more example of overcollecting in the past? *Journal of Applied Botany and Food Quality*. **83**, 60–63 (2009).
3. G. Krotkov, A review of literature onTaraxacum koksaghyz Rod. *The Botanical Review*. **11**, 417–461 (1945), doi:10.1007/BF02861139.
4. J. B. van Beilen, Y. Poirier, Guayule and Russian Dandelion as Alternative Sources of Natural Rubber. *Critical reviews in biotechnology*. **27**, 217–231 (2007), doi:10.1080/07388550701775927.
5. W. G. Whaley, J. S. Bowen, Russian dandelion (kok-saghyz): an emergency source of natural rubber (US Department of Agriculture) (1947).
6. P. van Dijk, J. Kirschner, J. Štěpánek, I. O. Baitulin, T. Černý, Taraxacum koksaghyz Rodin definitely is not an example of overcollecting in the past. A reply to S. Volis et al. (2009). *2012*. **83**, 3 (2012).
7. M. Ulmann, Wertvolle Kautschukpflanzen des gemaessigten Klimas, 1-6, 17–45, 136 (1951).
8. M. Kreuzberger, T. Hahn, S. Zibek, J. Schiemann, K. Thiele, Seasonal pattern of biomass and rubber and inulin of wild Russian dandelion (Taraxacum koksaghyz L. Rodin) under experimental field conditions. *European Journal of Agronomy*. **80**, 66–77 (2016), doi:10.1016/j.eja.2016.06.011.
9. K. J. Hodgson-Kratky, D. J. Wolyn, Inheritance of Flowering Habit in Russian Dandelion. *J. Amer. Soc. Hort. Sci.* **140**, 614–619 (2015).
10. M. Eggert, personal communication, 15 October 2018.
11. NordNordWest/Wikipedia, *Political and regional map of Kazakhstan*, *https://creativecommons.org/licenses/by-sa/3.0/de/legalcode*, CC-BY-SA-3.0-DE (2009).
12. T. M. Lewinsohn, The geographical distribution of plant latex. *CHEMOECOLOGY*. **2**, 64–68 (1991), doi:10.1007/BF01240668.
13. A. A. Agrawal, K. Konno, Latex: A Model for Understanding Mechanisms, Ecology, and Evolution of Plant Defense Against Herbivory. *Annual Review of Ecology, Evolution, and Systematics*. **40**, 311–331 (2009), doi:10.1146/annurev.ecolsys.110308.120307.
14. D. Wahler *et al.,* Polyphenoloxidase Silencing Affects Latex Coagulation in Taraxacum Species. *Plant Physiology*. **151**, 334–346 (2009), doi:10.1104/pp.109.138743.
15. M. D. Serpe, A. J. Muir, A. Driouich, Immunolocalization of beta-D-glucans, pectins, and arabinogalactan-proteins during intrusive growth and elongation of nonarticulated laticifers in Asclepias speciosa Torr. *Planta*. **215**, 357–370 (2002), doi:10.1007/s00425-002-0756-y.
16. X.-Q. Zhao, The significance of the structure of laticifer with relation to the exudation of latex in Hevea brasiliensis. *Journal of Natural Rubber Research* (1987).
17. A. Hillebrand *et al.,* Down-Regulation of Small Rubber Particle Protein Expression Affects Integrity of Rubber Particles and Rubber Content in *Taraxacum brevicorniculatum*. *PLoS ONE*. **7**, e41874 (2012), doi:10.1371/journal.pone.0041874.
18. J. B. van Beilen, Y. Poirier, Establishment of new crops for the production of natural rubber. *Trends in Biotechnology*. **25**, 522–529 (2007), doi:10.1016/j.tibtech.2007.08.009.
19. ANRPC, *ANRPC Releases Natural Rubber Trends & Statistics, December 2017* (available at http://www.anrpc.org/html/market-news-details.aspx?ID=26&PID=28&NID=1703).
20. ANRPC, *ANRPC releases Natural Rubber Trends & Statistics Jan- Feb 2020* (09.03.2020) (available at http://www.anrpc.org/html/news-secretariat-details.aspx?ID=9&PID=39&NID=4668).

21. ANRPC, *ANRPC Releases Natural Rubber Trends December 2021* (09.02.2022) (available at http://www.anrpc.org/html/news-secretariat-details.aspx?ID=9&PID=39&NID=8535).
22. H. H. Goh, K. L. Tan, C. Y. Khor, S. L. Ng, Volatility and Market Risk of Rubber Price in Malaysia: Pre-and Post-Global Financial Crisis. *Journal of Quantitative Economics*. **14**, 323–344 (2016).
23. European Tyre & Rubber manufacturers' association, *European Tyre & Rubber Industry Statistics Edition 2021* (2021) (available at https://www.etrma.org/wp-content/uploads/2021/12/20211030-Statistics-booklet-2021VF.pdf).
24. K. Cornish, Similarities and differences in rubber biochemistry among plant species. *Phytochemistry*. **57**, 1123–1134 (2001).
25. F. Cataldo, Guayule rubber: A new possible world scenario for the production of natural rubber? *Progress in rubber and plastics technology*. **16**, 31–59 (2000).
26. *isoprene unit MW* (available at https://pubchem.ncbi.nlm.nih.gov/compound/isoprene#section=Top).
27. K. Cornish, D. J. Siler, O.-K. Grosjen, N. Goodman, Fundamental similarities in rubber particle architecture and function in three evolutionarily divergent plant species. *Journal of Natural Rubber Research*. **8**, 275 (1993).
28. T. Schmidt *et al.,* Characterization of rubber particles and rubber chain elongation in Taraxacum koksaghyz. *BMC Biochemistry*. **11**, 11 (2010).
29. I. Uhlemann, M. Eggert, J. SCHIEMANN, K. THIELE, Zum Wiederanbau von Taraxacum koksaghyz (Asteraceae) als Kautschuklieferant in Deutschland. *Kochia*. **12**, 19–35 (2019).
30. D. A. Ramirez-Cadavid, K. Cornish, F. C. Michel Jr, Taraxacum kok-saghyz (TK): compositional analysis of a feedstock for natural rubber and other bioproducts. *Industrial Crops and Products*. **107**, 624–640 (2017).
31. S. C. Espy, J. D. Keasling, J. Castillon, K. Cornish, Initiator-independent and initiator-dependent rubber biosynthesis in Ficus elastica. *Archives of biochemistry and biophysics*. **448**, 13–22 (2006).
32. A. M. Patten, D. G. Vassão, M. P. Wolcott, L. B. Davin, N. G. Lewis, in *Comprehensive Natural Products II*, H.-W. (Liu, L. Mander, Eds. (Elsevier, Oxford, 2010), pp. 1173–1296.
33. C. C. Ho, in *Natural rubber materials*, S. Thomas, R. K. R., H. J. Maria, C. H. Chan, L. A. Pothen, Eds. (RSC Publ, Cambridge, 2014), vol. **1,** pp. 73–106.
34. C. H. AJG Simoes, The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development. Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence. (2011).
35. P. Venkatachalam, N. Geetha, P. Sangeetha, A. Thulaseedharan, Natural rubber producing plants. *African Journal of Biotechnology*. **12**, 1297–1310 (2013).
36. B. T. da Hora Júnior *et al.,* Erasing the past: a new identity for the Damoclean pathogen causing South American leaf blight of rubber. *PLoS ONE*. **9**, e104750 (2014), doi:10.1371/journal.pone.0104750.
37. W. Davies, The rubber industry's biological nightmare. *Fortune*. **136**, 86–95 (1997).
38. A. O. Oghenekaro *et al.,* Molecular phylogeny of Rigidoporus microporus isolates associated with white rot disease of rubber trees (Hevea brasiliensis). *Fungal Biology*. **118**, 495–506 (2014), doi:10.1016/j.funbio.2014.04.001.
39. P. Holliday, in *Fungus diseases of tropical crops*, P. Holliday, Ed. (Cambridge University Press, 1980), 495 pp.
40. H. Chen *et al.,* Pushing the Limits: The Pattern and Dynamics of Rubber Monoculture Expansion in Xishuangbanna, SW China. *PLoS ONE*. **11**, e0150062 (2016), doi:10.1371/journal.pone.0150062.

41. J. Bousquet *et al.,* Natural rubber latex allergy among health care workers: a systematic review of the evidence. *Journal of Allergy and Clinical Immunology*. **118**, 447–454 (2006).
42. *Guayule* (available at https://www.wur.nl/en/Research-Results/Projects-and-programmes/eu-pearls-projects/Projects/Guayule.htm).
43. S. Wagner, H. Breiteneder, Hevea brasiliensis latex allergens: current panel and clinical relevance. *International archives of allergy and immunology*. **136**, 90–97 (2005).
44. POWO (2019), *Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet* (available at http://www.plantsoftheworldonline.org/).
45. D. Rasutis, K. Soratana, C. McMahan, A. E. Landis, A sustainability review of domestic rubber from the guayule plant. *Industrial Crops and Products*. **70**, 383–394 (2015), doi:10.1016/j.indcrop.2015.03.042.
46. F. S. Nakayama, Guayule future development. *Industrial Crops and Products*. **22**, 3–13 (2005), doi:10.1016/j.indcrop.2004.05.006.
47. H. Kajiura, N. Suzuki, H. Mouri, N. Watanabe, Y. Nakazawa, Elucidation of rubber biosynthesis and accumulation in the rubber producing shrub, guayule (Parthenium argentatum Gray). *Planta*. **247**, 513–526 (2018), doi:10.1007/s00425-017-2804-7.
48. M. G. Gilliland, J. van Staden, Detection of rubber in guayule (Parthenium argentatum Gray) at the ultrastructural level. *Zeitschrift für Pflanzenphysiologie*. **110**, 285–291 (1983).
49. W. W. Schloman, Processing guayule for latex and bulk rubber. *Industrial Crops and Products*. **22**, 41–47 (2005), doi:10.1016/j.indcrop.2004.04.031.
50. D. J. Siler, K. Cornish, R. G. Hamilton, Absence of cross-reactivity of IgE antibodies from subjects allergic to Hevea brasiliensislatex with a new source of natural rubber latex from guayule(Parthenium argentatum). *Journal of Allergy and Clinical Immunology*. **98**, 895–902 (1996), doi:10.1016/S0091-6749(96)80005-4.
51. K. Cornish *et al.,* Guayule latex provides a solution for the critical demands of the non-allergenic medical products market. *AGRO FOOD INDUSTRY HI TECH*. **12**, 27–32 (2001).
52. *Yulex Guayule Info* (available at http://yulex.com/products/guayule/).
53. J. Kirschner, J. Štěpánek, T. Černý, P. de Heer, P. van Dijk, Available ex situ germplasm of the potential rubber crop Taraxacum koksaghyz belongs to a poor rubber producer, T. brevicorniculatum (Compositae–Crepidinae). *Genet Resour Crop Evol*. **60**, 455–471 (2013), doi:10.1007/s10722-012-9848-0.
54. H. Mooibroek, K. Cornish, Alternative sources of natural rubber. *Appl Microbiol Biotechnol*. **53**, 355–365 (2000), doi:10.1007/s002530051627.
55. *PENRA* (http://u.osu.edu/penra/penra-development-tracks/).
56. *EU PEARLS* (available at https://www.wur.nl/en/Research-Results/Projects-and-programmes/eu-pearls-projects.htm).
57. *DRIVE4EU* (available at https://cordis.europa.eu/project/id/613697/de).
58. *RUBIN 2* (available at https://www.fisaonline.de).
59. O. Munt *et al.,* Fertilizer and planting strategies to increase biomass and improve root morphology in the natural rubber producer Taraxacum brevicorniculatum. *Industrial Crops and Products*. **36**, 289–293 (2012), doi:10.1016/j.indcrop.2011.10.014.
60. *EVITA project* (available at https://fisaonline.de).
61. J. Post *et al.,* Establishment of an ex vivo laticifer cell suspension culture from Taraxacum brevicorniculatum as a production system for cis-isoprene. *Journal of Molecular Catalysis B: Enzymatic*. **103**, 85–93 (2014), doi:10.1016/j.molcatb.2013.07.013.
62. T. Schmidt *et al.,* Molecular Cloning and Characterization of Rubber Biosynthetic Genes from Taraxacum koksaghyz. *Plant Mol Biol Rep*. **28**, 277–284 (2010), doi:10.1007/s11105-009-0145-9.

63. J. Luo, J. Cardina, Germination patterns and implications for invasiveness in three Taraxacum (Asteraceae) species. *Weed Research*. **52**, 112–121 (2012), doi:10.1111/j.1365-3180.2011.00898.x.
64. M. Eggert, J. Schiemann, K. Thiele, Yield performance of Russian dandelion transplants (Taraxacum koksaghyz L. Rodin) in flat bed and ridge cultivation with different planting densities. *European Journal of Agronomy*. **93**, 126–134 (2018), doi:10.1016/j.eja.2017.12.003.
65. M. Eggert, K. Thiele, Selektivität von Herbiziden im Russischen Löwenzahn (Taraxacum kok-saghyz L. Rodin).
66. T. Lin *et al.,* Genome analysis of Taraxacum kok-saghyz Rodin provides new insights into rubber biosynthesis. *National Science Review*, nwx101-nwx101 (2017), doi:10.1093/nsr/nwx101.
67. *Bridgestone Finds Russian Dandelion May Be a Sustainable Source of Natural Rubber* (available at https://www.bridgestone.com/corporate/news/2012051701.html).
68. *Continental Dandelion* (available at https://www.continental-reifen.de/autoreifen/media-services/newsroom/taraxagum).
69. J. A. McCourt, R. G. Duggleby, Acetohydroxyacid synthase and its role in the biosynthetic pathway for branched-chain amino acids. *Amino acids*. **31**, 173–210 (2006), doi:10.1007/s00726-005-0297-3.
70. R. G. Duggleby, J. A. McCourt, L. W. Guddat, Structure and mechanism of inhibition of plant acetohydroxyacid synthase. *Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale*. **46**, 309–324 (2008), doi:10.1016/j.plaphy.2007.12.004.
71. Y. Liu, Y. Li, X. Wang, Acetohydroxyacid synthases: evolution, structure, and function. *Appl Microbiol Biotechnol*. **100**, 8633–8649 (2016), doi:10.1007/s00253-016-7809-9.
72. T. Lonhienne *et al.,* Structural insights into the mechanism of inhibition of AHAS by herbicides. *Proceedings of the National Academy of Sciences*. **115**, E1945 (2018), doi:10.1073/pnas.1714392115.
73. Y.-T. Lee, R. G. Duggleby, Identification of the Regulatory Subunit of Arabidopsis thaliana Acetohydroxyacid Synthase and Reconstitution with Its Catalytic Subunit. *Biochemistry*. **40**, 6836–6844 (2001), doi:10.1021/bi002775q.
74. L. Y. Geer *et al.,* The NCBI BioSystems database. *Nucleic Acids Research*. **38**, D492-6 (2010), doi:10.1093/nar/gkp858.
75. T. Lonhienne *et al.,* Structures of fungal and plant acetohydroxyacid synthases. *Nature*. **586**, 317–321 (2020), doi:10.1038/s41586-020-2514-3.
76. A. Kaplun *et al.,* Structure of the Regulatory Subunit of Acetohydroxyacid Synthase Isozyme III from Escherichia coli. *Journal of Molecular Biology*. **357**, 951–963 (2006), doi:10.1016/j.jmb.2005.12.077.
77. N. M. Karanth, S. P. Sarma, The coil-to-helix transition in IlvN regulates the allosteric control of Escherichia coli acetohydroxyacid synthase I. *Biochemistry*. **52**, 70–83 (2012).
78. Y. T. Lee, R. G. Duggleby, Regulatory interactions in Arabidopsis thaliana acetohydroxyacid synthase. *FEBS letters*. **512**, 180–184 (2002).
79. R. G. Duggleby, S. S. Pang, Acetohydroxyacid synthase. *J Biochem Mol Biol*. **33**, 1–36 (2000).
80. B. J. Miflin, The Location of Nitrite Reductase and Other Enzymes Related to Amino Acid Biosynthesis in the Plastids of Root and Leaves. *Plant Physiology*. **54**, 550–555 (1974).
81. B. K. Singh, D. L. Shaner, Biosynthesis of Branched Chain Amino Acids: From Test Tube to Field. *The Plant Cell*. **7**, 935 (1995), doi:10.1105/tpc.7.7.935.

82. G. Breccia, T. Vega, S. A. Felitti, L. Picardi, G. Nestares, Differential expression of acetohydroxyacid synthase genes in sunflower plantlets and its response to imazapyr herbicide. *Plant Science*. **208**, 28–33 (2013), doi:10.1016/j.plantsci.2013.03.008.
83. G. von Heijne, J. Steppuhn, R. G. Herrmann, Domain structure of mitochondrial and chloroplast targeting peptides. *European journal of biochemistry*. **180**, 535–545 (1989).
84. R. G. Rutledge, T. Quellet, J. Hattori, B. L. Miki, Molecular characterization and genetic origin of the Brassica napus acetohydroxyacid synthase multigene family. *Molecular & general genetics : MGG*. **229**, 31–40 (1991).
85. The UniProt Consortium, UniProt: the universal protein knowledgebase. *Nucleic Acids Research*. **45**, D158-D169 (2017), doi:10.1093/nar/gkw1099.
86. A. K. CHANG, R. G. Duggleby, Expression, purification and characterization of Arabidopsis thaliana acetohydroxyacid synthase. *Biochemical Journal*. **327**, 161–169 (1997).
87. S. S. Pang, R. G. Duggleby, L. W. Guddat, Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors1. *Journal of Molecular Biology*. **317**, 249–262 (2002).
88. M. D. Garcia, J.-G. Wang, T. Lonhienne, L. W. Guddat, Crystal structure of plant acetohydroxyacid synthase, the target for several commercial herbicides. *The FEBS journal*. **284**, 2037–2051 (2017), doi:10.1111/febs.14102.
89. K. Tittmann *et al.,* Electron Transfer in Acetohydroxy Acid Synthase as a Side Reaction of Catalysis. Implications for the Reactivity and Partitioning of the Carbanion/Enamine Form of (α-Hydroxyethyl)thiamin Diphosphate in a “Nonredox” Flavoenzyme. *Biochemistry*. **43**, 8652–8661 (2004), doi:10.1021/bi049897t.
90. J. A. McCourt, S. S. Pang, L. W. Guddat, R. G. Duggleby, Elucidating the Specificity of Binding of Sulfonylurea Herbicides to Acetohydroxyacid Synthase. *Biochemistry*. **44**, 2330–2338 (2005), doi:10.1021/bi047980a.
91. C. F. Hawkins, A. Borges, R. N. Perham, A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS letters*. **255**, 77–82 (1989).
92. C. Vogel, J. Pleiss, The modular structure of ThDP-dependent enzymes. *Proteins: Structure, Function, and Bioinformatics*. **82**, 2523–2537 (2014).
93. Lonhienne Thierry *et al.,* Commercial Herbicides Can Trigger the Oxidative Inactivation of Acetohydroxyacid Synthase. *Angew. Chem.* **128**, 4319–4323 (2016), doi:10.1002/ange.201511985.
94. T. Lonhienne, M. D. Garcia, L. W. Guddat, The Role of a FAD Cofactor in the Regulation of Acetohydroxyacid Synthase by Redox Signaling Molecules. *Journal of Biological Chemistry*. **292**, 5101–5109 (2017), doi:10.1074/jbc.M116.773242.
95. J. V. Schloss, D. E. van Dyk, Acetolactate synthase isozyme II from Salmonella typhimurium. *Methods in enzymology*. **166**, 445 (1988).
96. B. J. Mazur, C. F. Chui, J. K. Smith, Isolation and characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicides. *Plant physiology*. **85**, 1110–1117 (1987).
97. K. Newhouse, T. Wang, P. Anderson, in *The Imidazolinone Herbicides*, D. L. Shaner, S. L. O'Connor, Ed. (CRC Press, Boca Raton, Florida, 1991), pp. 139–150.
98. S. J. Keeler, P. Sanders, J. K. Smith, B. J. Mazur, Regulation of Tobacco Acetolactate Synthase Gene Expression. *Plant Physiology*. **102**, 1009 (1993), doi:10.1104/pp.102.3.1009.
99. Q. Feng *et al.,* Sequence and analysis of rice chromosome 4. *Nature*. **420**, 316–320 (2002), doi:10.1038/nature01183.

100. C. Pozniak *et al.,* Physiological and Molecular Characterization of Mutation-Derived Imidazolinone Resistance in Spring Wheat. *Crop Science - CROP SCI*. **44** (2004), doi:10.2135/cropsci2004.1434.
101. J. W. Grula, R. L. Hudspeth, S. L. Hobbs, D. M. Anderson, Organization, inheritance and expression of acetohydroxyacid synthase genes in the cotton allotetraploid Gossypium hirsutum. *Plant molecular biology*. **28**, 837–846 (1995), doi:10.1007/BF00042069.
102. S. S. Pang, L. W. Guddat, R. G. Duggleby, Crystallization of Arabidopsis thaliana acetohydroxyacid synthase in complex with the sulfonylurea herbicide chlorimuron ethyl. *Acta Crystallographica Section D: Biological Crystallography*. **60**, 153–155 (2004).
103. F. Bekkaoui, P. Schorr, W. L. Crosby, Acetolactate synthase from Brassica napus: Immunological characterization and quaternary structure of the native enzyme. *Physiologia Plantarum*. **88**, 475–484 (1993).
104. E. Delfourne, J. Bastide, R. Badon, A. Rachon, P. Genix, Specificity of plant acetohydroxyacid synthase: formation of products and inhibition by herbicides. *Plant Physiology and Biochemistry (France)* (1994).
105. M. D. Southan, L. Copeland, Physical and kinetic properties of acetohydroxyacid synthase from wheat leaves. *Physiologia Plantarum*. **98**, 824–832 (1996).
106. J. Durner, P. Böger, Acetolactate synthase from barley (Hordeum vulgare L.): purification and partial characterization. *Zeitschrift für Naturforschung C*. **43**, 850–856 (1988).
107. B. K. Singh, M. A. Stidham, D. L. Shaner, Separation and characterization of two forms of acetohydroxy acid synthase from black mexican sweet corn cells. *Journal of Chromatography A*. **444**, 251–261 (1988), doi:10.1016/S0021-9673(01)94028-2.
108. B. J. Mazur, C.-F. Chui, J. K. Smith, Isolation, Characterization of plant genes coding for acetolactate synthase, the target enzyme for two classes of herbicides. *Plant physiology*, 1110–1117 (1987).
109. L. Y. Fang, P. R. Gross, C. H. Chen, M. Lillis, Sequence of two acetohydroxyacid synthase genes from Zea mays. *Plant molecular biology*. **18**, 1185–1187 (1992), doi:10.1007/BF00047723.
110. P. Bernasconi, A. R. Woodworth, B. A. Rosen, M. V. Subramanian, D. L. Siehl, A naturally occurring point mutation confers broad range tolerance to herbicides that target acetolactate synthase. *The Journal of biological chemistry*. **270**, 17381–17385 (1995).
111. H. P. Hershey, L. J. Schwartz, J. P. Gale, L. M. Abell, Cloning and functional expression of the small subunit of acetolactate synthase from Nicotiana plumbaginifolia. *Plant molecular biology*. **40**, 795–806 (1999).
112. B. Singh, G. Schmitt, M. Lillis, J. M. Hand, R. Misra, Overexpression of acetohydroxyacid synthase from arabidopsis as an inducible fusion protein in Escherichia coli: production of polyclonal antibodies, and immunological characterization of the enzyme. *Plant Physiology*. **97**, 657–662 (1991).
113. K. Y. Lee *et al.,* The molecular basis of sulfonylurea herbicide resistance in tobacco. *The EMBO journal*. **7**, 1241–1248 (1988).
114. D. Degrande, E. Dewaele, S. Rambour, The AHAS gene of Cichorium intybus is expressed in fast growing and inflorescential organs. *Physiologia Plantarum*. **110**, 224–231 (2000).
115. T. Ouellet, R. G. Rutledge, B. L. Miki, Members of the acetohydroxyacid synthase multigene family of Brassica napus have divergent patterns of expression. *The Plant Journal*. **2**, 321–330 (1992).
116. Schmitt, G. K., Singh, B. K., Tissue distyibution of acetohydroxyacid synthase activity at various development stages of lima bean. *Pest. Sci.* **30**, 418–419 (1990).

117. S. Iwakami, A. Uchino, H. Watanabe, Y. Yamasue, T. Inamura, Isolation and expression of genes for acetolactate synthase and acetyl-CoA carboxylase in Echinochloa phyllopogon, a polyploid weed species. *Pest. Manag. Sci*. **68**, 1098–1106 (2012).
118. G. Forlani, E. Nielsen, P. Landi, R. Tuberosa, Chlorsulfuron tolerance and acetolactate synthase activity in corn (Zea mays L.) inbred lines. *Weed Science*. **39**, 553–557 (1991).
119. M. Barrett, Reduction of Imazaquin Injury to Corn (Zea mays) and Sorghum (Sorghum bicolor) by Antidotes. *Weed Science*. **37**, 34–41 (1989).
120. F. Dastgheib, R. J. Field, Acetolactate synthase activity and chlorsulfuron sensitivity of wheat cultivars. *Weed Research, Oxford (United Kingdom)* (1998).
121. F. Zhang *et al.,* Spatial and temporal changes in acetolactate synthase activity as affected by new herbicide ZJ0273 in rapeseed, barley and water chickweed. *Pesticide Biochemistry and Physiology*. **95**, 63–71 (2009), doi:10.1016/j.pestbp.2009.06.005.
122. J. Dolezel, J. Bartos, H. Voglmayr, J. Greilhuber, Nuclear DNA content and genome size of trout and human. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*. **51**, 127-8; author reply 129 (2003), doi:10.1002/cyto.a.10013.
123. I. Heap, *The International Survey of Herbicide Resistant Weeds. Online.* (2020) (available at www.weedscience.org).
124. Q. Yu, S. B. Powles, Resistance to AHAS inhibitor herbicides. *Pest Management Science*. **70**, 1340–1350 (2014), doi:10.1002/ps.3710.
125. M. D. Garcia, A. Nouwens, T. G. Lonhienne, L. W. Guddat, Comprehensive understanding of acetohydroxyacid synthase inhibition by different herbicide families. *Proceedings of the National Academy of Sciences of the United States of America*. **114**, E1091-E1100 (2017), doi:10.1073/pnas.1616142114.
126. D. L. Shaner, M. Stidham, B. Singh, S. Tan, in *Modern crop protection compounds, 3 volume set*, W. Krämer, U. Schirmer, P. Jeschke, M. Witschel, Eds. (John Wiley & Sons, 2012), pp. 88–98.
127. M. R. van Ellis, D. L. Shaner, Mechanism of cellular absorption of imidazolinones in soybean (Glycine max) leaf discs. *Pesticide science*. **23**, 25–34 (1988).
128. M. A. Stidham, B. K. Singh, D. L. Shaner, S. L. O'Connor, The Imidazolinone Herbicides (1991).
129. P. J. Wepplo, D. L. Shaner, S. L. O'Connor, Imidazolinone Herbicides. *Princeton: Academic Press*, 15–29 (1991).
130. D. L. Little, D. W. Ladner, D. L. Shaner, R. D. Ilnicki, Modeling root absorption and translocation of 5-substituted analogs of the imidazolinone herbicide, imazapyr. *Pesticide science*. **41**, 171–185 (1994).
131. O. Ort, in *Modern crop protection compounds, 3 volume set*, W. Krämer, U. Schirmer, P. Jeschke, M. Witschel, Eds. (John Wiley & Sons, 2012), p. 50.
132. *MarvinSketch version 18.20* (ChemAxon).
133. J. V. Hay, Chemistry of Sulfonylurea Herbicides. *Pesticide science*. **29** (1990), doi:10.1002/ps.2780290303.
134. D. M. G. Anderson, V. A. Carolan, S. Crosland, K. R. Sharples, M. R. Clench, Examination of the translocation of sulfonylurea herbicides in sunflower plants by matrix-assisted laser desorption/ionisation mass spectrometry imaging. *Rapid Communications in Mass Spectrometry*. **24**, 3309–3319 (2010).
135. A. K. CHANG, R. G. Duggleby, Herbicide-resistant forms of Arabidopsis thaliana acetohydroxyacid synthase: characterization of the catalytic properties and sensitivity to inhibitors of four defined mutants. *Biochemical Journal*. **333**, 765–777 (1998).
136. S. Tan, R. R. Evans, M. L. Dahmer, B. K. Singh, D. L. Shaner, Imidazolinone-tolerant crops. *Pest Management Science*. **61**, 246–257 (2005), doi:10.1002/ps.993.

137. H. Lee *et al.,* Single nucleotide mutation in the barley acetohydroxy acid synthase (AHAS) gene confers resistance to imidazolinone herbicides. *Proceedings of the National Academy of Sciences*. **108**, 8909–8913 (2011).
138. G. Jander *et al.,* Ethylmethanesulfonate Saturation Mutagenesis in Arabidopsis to Determine Frequency of Herbicide Resistance. *Plant Physiology*. **131**, 139–146 (2003), doi:10.1104/pp.102.010397.
139. M. Sibony, A. Michel, H. U. Haas, B. Rubin, K. Hurle, Sulfometuron-resistant Amaranthus retroflexus: cross-resistance and molecular basis for resistance to acetolactate synthase-inhibiting herbicides. *Weed Research*. **41**, 509–522 (2001).
140. J. A. McCourt, S. S. Pang, J. King-Scott, L. W. Guddat, R. G. Duggleby, Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase. *Proceedings of the National Academy of Sciences of the United States of America*. **103**, 569–573 (2006), doi:10.1073/pnas.0508701103.
141. Tranel, P.J., Wright, T.R, and Heap, I.M., *Mutations in herbicide-resistant weeds to ALS inhibitors.* (available at http://www.weedscience.com.).
142. R. S. Baucom, in *Weedy and Invasive Plant Genomics*, C. N. Stewart, Ed. (Wiley, 2009), pp. 163–175.
143. Weed Technology Volume 12, Issue 4 (October-December), p. 789 (1998).
144. I. Heap, Criteria for confirmation of herbicide-resistant weeds. *International Survey of Herbicide-Resistant Weeds* (2005).
145. R. Mauricio, Natural selection and the joint evolution of toleranceand resistance as plant defenses. *Evolutionary Ecology*. **14**, 491–507 (2000).
146. J. Menendez, M. A. Rojano-Delgado, R. de Prado, in *Retention, uptake, and translocation of agrochemicals in plants*, K. Myung, N. M. Satchivi, C. K. Kingston, Eds. (American Chemical Society, Washington, DC, 2014), vol. **1171,** pp. 141–157.
147. C. Délye, Unravelling the genetic bases of non-target-site-based resistance (NTSR) to herbicides: a major challenge for weed science in the forthcoming decade. *Pest Management Science*. **69**, 176–187 (2013), doi:10.1002/ps.3318.
148. R. E. Hoagland, R. M. Zablotowicz, J. C. Hall, in *Pesticide Biotransformation in Plants and Microorganisms* (American Chemical Society, 2000), vol. **777,** pp. 2–27.
149. K. Kreuz, R. Tommasini, E. Martinoia, Old Enzymes for a New Job (Herbicide Detoxification in Plants). *Plant Physiology*. **111**, 349–353 (1996).
150. D. Coupland, Ed., *Predicting and optimising the translocation of foliage-applied herbicide--a plant physiologist's perspective* (1991).
151. M.-H. Denis, S. Delrot, Carrier-mediated uptake of glyphosate in broad bean (Vicia faba) via a phosphate transporter. **87**, 569–575 (1993), doi:10.1034/j.1399-3054.1993.870418.x.
152. P. Boger, G. Sandmann, *Target sites of herbicide action* (CRC Press, 1989).
153. M. D. Devine, S. O. Duke, C. Fedtke, Physiology of herbicide action. PTR Prentice-Hall. *Inc., New Jersey*, 441 (1993).
154. Q. Yu, S. Powles, Metabolism-Based Herbicide Resistance and Cross-Resistance in Crop Weeds: A Threat to Herbicide Sustainability and Global Crop Production. *Plant Physiology*. **166**, 1106–1118 (2014), doi:10.1104/pp.114.242750.
155. J. Coleman, M. Blake-Kalff, E. Davies, Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. *Trends in Plant Science*. **2**, 144–151 (1997), doi:10.1016/S1360-1385(97)01019-4.
156. H. Sandermann, Plant metabolism of xenobiotics. *Trends in Biochemical Sciences*. **17**, 82–84 (1992), doi:10.1016/0968-0004(92)90507-6.

157. D. Bowles, E.-K. Lim, B. Poppenberger, F. E. Vaistij, Glycosyltransferases of lipophilic small molecules. *Annual Review of Plant Biology*. **57**, 567–597 (2006), doi:10.1146/annurev.arplant.57.032905.105429.
158. J. S. Yuan, P. J. Tranel, C. N. Stewart, Non-target-site herbicide resistance: a family business. *Trends in Plant Science*. **12**, 6–13 (2007), doi:10.1016/j.tplants.2006.11.001.
159. K. K. Hatzios, *Regulation of enzymatic systems detoxifying xenobiotics in plants* (Springer Science & Business Media, 1997).
160. D. L. Shaner, B. K. Singh, in *Herbicide Activity: Toxicology, Biochemistry and Molecular Biology*, R. M. Roe, J. D. Burton, R. J. Kuhr, Eds. (IOS Press, 1997), pp. 91–95.
161. M. Royuela *et al.,* Physiological consequences of continuous, sublethal imazethapyr supply to pea plants. *Journal of plant physiology*. **157**, 345–354 (2000), doi:10.1016/S0176-1617(00)80057-7.
162. Q. Zhou, W. Liu, Y. Zhang, K. K. Liu, Action mechanisms of acetolactate synthase-inhibiting herbicides. *Pesticide Biochemistry and Physiology*. **89**, 89–96 (2007), doi:10.1016/j.pestbp.2007.04.004.
163. L. Orcaray, M. Igal, D. Marino, A. Zabalza, M. Royuela, The possible role of quinate in the mode of action of glyphosate and acetolactate synthase inhibitors. *Pest Management Science*. **66**, 262–269 (2010), doi:10.1002/ps.1868.
164. D. Shaner, Herbicides, Imidazolinones. *Encyclopedia of Agrochemicals* (2003).
165. A. M. Rojano-Delgado, F. Priego-Capote, Luque de Castro, María Dolores, R. de Prado, Mechanism of imazamox resistance of the Clearfield® wheat cultivar for better weed control. *Agronomy for Sustainable Development*. **35**, 639–648 (2015), doi:10.1007/s13593-014-0232-7.
166. D. L. Shaner, B. Tecle, J. C. Hall, R. E. Hoagland, R. M. Zablotowicz, Eds., *Designing herbicide tolerance based on metabolic alteration: the challenges and the future* (Washington, DC; American Chemical Society; 1999, 2001).
167. D. D. Songstad, J. F. Petolino, D. F. Voytas, N. A. Reichert, Genome Editing of Plants. *Critical Reviews in Plant Sciences*. **36**, 1–23 (2017), doi:10.1080/07352689.2017.1281663.
168. R. S. Meyer, A. E. DuVal, H. R. Jensen, Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytologist*. **196**, 29–48 (2012).
169. P. Gepts, A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Science*. **42**, 1780–1790 (2002).
170. R. S. Meyer, M. D. Purugganan, Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics*. **14**, 840 (2013).
171. R. Dirzo, P. H. Raven, Global State of Biodiversity and Loss. *Annu. Rev. Environ. Resour.* **28**, 137–167 (2003), doi:10.1146/annurev.energy.28.050302.105532.
172. Y. Oladosu *et al.,* Principle and application of plant mutagenesis in crop improvement: a review. *Biotechnology & Biotechnological Equipment*. **30**, 1–16 (2016), doi:10.1080/13102818.2015.1087333.
173. R. Azpiroz-Leehan, K. A. Feldmann, T-DNA insertion mutagenesis in Arabidopsis: going back and forth. *Trends in Genetics*. **13**, 152–156 (1997).
174. M. Lusser, C. Parisi, D. Plan, E. Rodriguez-Cerezo, Deployment of new biotechnologies in plant breeding. *Nat Biotech*. **30**, 231–239 (2012).
175. F. Hartung, J. Schiemann, Precise plant breeding using new genome editing techniques. *The Plant Journal*. **78**, 742–752 (2014), doi:10.1111/tpj.12413.
176. E. Pellegrino, S. Bedini, M. Nuti, L. Ercoli, Impact of genetically engineered maize on agronomic, environmental and toxicological traits: a meta-analysis of 21 years of field data. *Scientific Reports*. **8**, 3113 (2018).

177. C. Mba, R. Afza, Q. Y. Shu, B. P. Forster, H. Nakagawa, Mutagenic radiations: X-rays, ionizing particles and ultraviolet. *Plant mutation breeding and biotechnology. CABI, Oxfordshire*, 83–90 (2012).
178. C. Mba, Induced mutations unleash the potentials of plant genetic resources for food and agriculture. *Agronomy*. **3**, 200–231 (2013).
179. in *Encyclopedia of Genetics, Genomics, Proteomics and Informatics*, G. P. Rédei, Ed. (Springer Netherlands, Dordrecht, 2008), pp. 642–923.
180. *FAO/IAEA Mutant Variety Database* (available at https://mvd.iaea.org/).
181. E. Strasburger, P. Sitte, *Lehrbuch der Botanik für Hochschulen* (Spektrum Akad. Verl., Heidelberg, ed. 35, 2002).
182. E. A. Greene *et al.,* Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics*. **164**, 731–740 (2003).
183. A. J. F. Griffiths, *An introduction to genetic analysis* (Freeman, New York, NY, ed. 8, 2004).
184. A. Bentley, B. MacLennan, J. Calvo, C. R. Dearolf, Targeted recovery of mutations in Drosophila. *Genetics*. **156**, 1169–1173 (2000).
185. C. Mba, R. Afza, S. Bado, S. Jain, Induced Mutagenesis in Plants Using Physical and Chemical Agents. *Plant Cell Culture: Essential Methods* (2010), doi:10.1002/9780470686522.ch7.
186. W. Siede, F. Eckardt-Schupp, A mismatch repair-based model can explain some features of UV mutagenesis in yeast. *Mutagenesis*. **1**, 471–474 (1986).
187. D. Botstein, E. W. Jones, Nonrandom mutagenesis of the Escherichia coli genome by nitrosoguanidine. *Journal of bacteriology*. **98**, 847 (1969).
188. B. A. Bridges, J. Law, R. J. Munson, Mutagenesis in Escherichia coli. *Molecular and General Genetics MGG*. **103**, 266–273 (1968), doi:10.1007/BF00273698.
189. F. Winston, EMS and UV mutagenesis in yeast. *Current protocols in molecular biology*. **Chapter 13**, Unit 13.3B (2008), doi:10.1002/0471142727.mb1303bs82.
190. J. F. T. Spencer, D. M. Spencer, in *Yeast Protocols, Methods in Cell and Molecular Biology*, I. H. Evans, Ed. (Humana Press, Totowa, NJ, 1996), pp. 17–38.
191. J. Chaudhuri, M. Parihar, A. Pires-daSilva, An Introduction to Worm Lab: from Culturing Worms to Mutagenesis. *Journal of Visualized Experiments: JoVE* (2011), doi:10.3791/2293.
192. S. Duangpan, W. Zhang, Y. Wu, S. H. Jansky, J. Jiang, Insertional Mutagenesis Using Tnt1 Retrotransposon in Potato1OPEN. *Plant Physiology*. **163**, 21–29 (2013), doi:10.1104/pp.113.221903.
193. E. Bucher, J. Reinders, M. Mirouze, Epigenetic control of transposon transcription and mobility in Arabidopsis. *Current opinion in plant biology*. **15**, 503–510 (2012), doi:10.1016/j.pbi.2012.08.006.
194. B. McClintock, *The significance of responses of the genome to challenge* (1993).
195. N. V. Fedoroff, The Suppressor-mutator element and the evolutionary riddle of transposons. *Genes to Cells*. **4**, 11–19 (1999).
196. R. Nath Radhamony, A. Mohan Prasad, R. Srinivasan, T-DNA insertional mutagenesis in Arabidopsis: a tool for functional genomics. *Electro Journal of Biotech*. **8** (2005), doi:10.2225/vol8-issue1-fulltext-4.
197. T. Sprink, J. Metje, F. Hartung, Plant genome editing by novel tools. *Current Opinion in Biotechnology*. **32**, 47–53 (2015), doi:10.1016/j.copbio.2014.11.010.
198. J. Menz, D. Modrzejewski, F. Hartung, R. Wilhelm, T. Sprink, Genome Edited Crops Touch the Market: A View on the Global Development and Regulatory Environment. *Front. Plant Sci.* **11**, 319 (2020), doi:10.3389/fpls.2020.586027.

199. A. Knoll, F. Fauser, H. Puchta, DNA recombination in somatic plant cells: mechanisms and evolutionary consequences. *Chromosome research*. **22**, 191–201 (2014).
200. A. Ray, M. Langer, Homologous recombination: ends as the means. *Trends in Plant Science*. **7**, 435–440 (2002), doi:10.1016/S1360-1385(02)02327-0.
201. A. Britt, Re-engineering plant gene targeting. *Trends in Plant Science*. **8**, 90–95 (2003), doi:10.1016/S1360-1385(03)00002-5.
202. J. D. Sander *et al.,* In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Research*. **41**, e181 (2013), doi:10.1093/nar/gkt716.
203. S. W. Cho *et al.,* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Research*. **24**, 132–141 (2014), doi:10.1101/gr.162339.113.
204. N. P. Pavletich, C. O. Pabo, Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science*. **252**, 809–817 (1991).
205. Y. G. Kim, J. Cha, S. Chandrasegaran, Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences*. **93**, 1156–1160 (1996), doi:10.1073/pnas.93.3.1156.
206. H. Puchta, B. Dujon, B. Hohn, Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Research*. **21**, 5034–5040 (1993).
207. M. Christian *et al.,* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. **186**, 757–761 (2010).
208. T. O. Auer, F. Del Bene, CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods*. **69**, 142–150 (2014), doi:10.1016/j.ymeth.2014.03.027.
209. C. Dong, P. Beetham, K. Vincent, P. Sharp, Oligonucleotide-directed gene repair in wheat using a transient plasmid gene repair assay system. *Plant Cell Rep*. **25**, 457–465 (2006), doi:10.1007/s00299-005-0098-x.
210. P. D. Hsu, E. S. Lander, F. Zhang, Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*. **157**, 1262–1278 (2014), doi:10.1016/j.cell.2014.05.010.
211. S. J. J. Brouns *et al.,* Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science (New York, N.Y.)*. **321**, 960–964 (2008), doi:10.1126/science.1159689.
212. K. S. Makarova *et al.,* Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*. **9**, 467 (2011).
213. F. Fauser, S. Schiml, H. Puchta, Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *The Plant Journal*. **79**, 348–359 (2014), doi:10.1111/tpj.12554.
214. M. Jinek *et al.,* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. **337**, 816–821 (2012), doi:10.1126/science.1225829.
215. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*. **109**, E2579-E2586 (2012).
216. J. D. Sander, J. K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*. **32**, 347 (2014).
217. K. Belhaj, A. Chaparro-Garcia, S. Kamoun, V. Nekrasov, Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant methods*. **9**, 39 (2013).
218. B. Farboud, B. J. Meyer, Dramatic Enhancement of Genome Editing by CRISPR/Cas9 Through Improved Guide RNA Design. *Genetics*. **199**, 959–971 (2015), doi:10.1534/genetics.115.175166.

219. B. P. Kleinstiver *et al.,* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. **529**, 490 (2016).
220. X. Wu, A. J. Kriz, P. A. Sharp, Target specificity of the CRISPR-Cas9 system. *Quantitative biology*. **2**, 59–70 (2014), doi:10.1007/s40484-014-0030-x.
221. J. Park *et al.,* Digenome-seq web tool for profiling CRISPR specificity. *Nature Methods*. **14**, 548 EP - (2017), doi:10.1038/nmeth.4262.
222. J. Tycko, V. E. Myer, P. D. Hsu, Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Molecular Cell*. **63**, 355–370 (2016), doi:10.1016/j.molcel.2016.07.004.
223. J. G. Doench *et al.,* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature biotechnology*. **34**, 184 EP - (2016), doi:10.1038/nbt.3437.
224. I. M. Slaymaker *et al.,* Rationally engineered Cas9 nucleases with improved specificity. *Science (New York, N.Y.)*. **351**, 84–88 (2016), doi:10.1126/science.aad5227.
225. P. J. Hooykaas, R. A. Schilperoort, Agrobacterium and plant genetic engineering. *Plant molecular biology*. **19**, 15–38 (1992).
226. J. Tempé, A. Petit, in *Molecular biology of plant tumors* (Elsevier, 1982), pp. 451–459.
227. L.-Y. Lee, S. B. Gelvin, T-DNA Binary Vectors and Systems. *Plant Physiology*. **146**, 325–332 (2008), doi:10.1104/pp.107.113001.
228. A. Hoekema, P. R. Hirsch, P. J. J. Hooykaas, R. A. Schilperoort, A binary plant vector strategy based on separation of vir-and T-region of the Agrobacterium tumefaciens Ti-plasmid. *Nature*. **303**, 179 (1983).
229. A. J. de Framond, K. A. Barton, M.-D. Chilton, Mini–Ti: a new vector strategy for plant genetic engineering. *Nature biotechnology*. **1**, 262 (1983).
230. S. J. Clough, A. F. Bent, Floral dip. *The Plant Journal*. **16**, 735–743 (1998), doi:10.1046/j.1365-313x.1998.00343.x.
231. M. Chabaud *et al.,* Agrobacterium tumefaciens-Mediated Transformation and in vitro Plant Regeneration of M. truncatula. *Medicago truncatula handbook. http://www. noble. org/MedicagoHandbook/pdf/AgrobacteriumTumefaciens* (2007).
232. C. A. Newell, Plant Transformation Technology: Developments and Applications. *MB*. **16**, 53–66 (2000), doi:10.1385/MB:16:1:53.
233. M.-D. Chilton *et al.,* Agrobacterium rhizogenes inserts T-DNA into the genomes of the host plant root cells. *Nature*. **295**, 432 (1982).
234. M. Fromm, L. P. Taylor, V. Walbot, Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proceedings of the National Academy of Sciences of the United States of America*. **82**, 5824–5828 (1985).
235. A. Crossway *et al.,* Integration of foreign DNA following microinjection of tobacco mesophyll protoplasts. *Mol Gen Genet*. **202**, 179–185 (1986), doi:10.1007/BF00331634.
236. F. A. Krens, L. Molendijk, G. J. Wullems, R. A. Schilperoort, In vitro transformation of plant protoplasts with Ti-plasmid DNA. *Nature*. **296**, 72 EP - (1982), doi:10.1038/296072a0.
237. H. D. Jones, C. A. Sparks, Stable transformation of plants. *Methods in molecular biology (Clifton, N.J.)*. **513**, 111–130 (2009), doi:10.1007/978-1-59745-427-8_7.
238. H. D. Jones, A. Doherty, C. A. Sparks, Transient transformation of plants. *Methods in molecular biology (Clifton, N.J.)*. **513**, 131–152 (2009), doi:10.1007/978-1-59745-427-8_8.
239. V. M. G. Borrelli, V. Brambilla, P. Rogowsky, A. Marocco, A. Lanubile, The Enhancement of Plant Disease Resistance Using CRISPR/Cas9 Technology. *Frontiers in Plant Science*. **9** (2018), doi:10.3389/fpls.2018.01245.
240. Q. Yu *et al.,* CRISPR/Cas9-induced Targeted Mutagenesis and Gene Replacement to Generate Long-shelf Life Tomato Lines. *Scientific Reports*. **7**, 11874 (2017), doi:10.1038/s41598-017-12262-1.

241. L. Chen *et al.,* A method for the production and expedient screening of CRISPR/Cas9-mediated non-transgenic mutant plants. *Horticulture Research*. **5**, 13 (2018), doi:10.1038/s41438-018-0023-4.
242. C.-S. Lin *et al.,* Application of protoplast technology to CRISPR/Cas9 mutagenesis. *Plant biotechnology journal* (2017), doi:10.1111/pbi.12870.
243. L. Arora, A. Narula, Gene Editing and Crop Improvement Using CRISPR-Cas9 System. *Frontiers in Plant Science*. **8** (2017), doi:10.3389/fpls.2017.01932.
244. J.-F. Li *et al.,* Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature biotechnology*. **31**, 688–691 (2013), doi:10.1038/nbt.2654.
245. C.-S. Lin *et al.,* Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant biotechnology journal*. **16**, 1295–1310 (2018), doi:10.1111/pbi.12870.
246. Y. Mao *et al.,* Application of the CRISPR-Cas system for efficient genome engineering in plants. *Molecular plant*. **6**, 2008–2011 (2013), doi:10.1093/mp/sst121.
247. H.-L. Xing *et al.,* A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology*. **14**, 327 (2014), doi:10.1186/s12870-014-0327-y.
248. Y. Gao, Y. Zhao, Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *Journal of integrative plant biology*. **56**, 343–349 (2014), doi:10.1111/jipb.12152.
249. V. Nekrasov, B. Staskawicz, D. Weigel, J. D. G. Jones, S. Kamoun, Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature biotechnology*. **31**, 691–693 (2013), doi:10.1038/nbt.2655.
250. J. Gao *et al.,* CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. *Plant molecular biology*. **87**, 99–110 (2015), doi:10.1007/s11103-014-0263-0.
251. X. Ma *et al.,* A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular plant*. **8**, 1274–1284 (2015), doi:10.1016/j.molp.2015.04.007.
252. M. Li *et al.,* Reassessment of the Four Yield-related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System. *Frontiers in Plant Science*. **7**, 377 (2016), doi:10.3389/fpls.2016.00377.
253. J. Post *et al.,* Laticifer-specific cis-prenyltransferase silencing affects the rubber, triterpene, and inulin content of Taraxacum brevicorniculatum. *Plant physiology*. **158**, 1406–1417 (2012), doi:10.1104/pp.111.187880.
254. E. E. Hood, S. B. Gelvin, L. S. Melchers, A. Hoekema, NewAgrobacterium helper plasmids for gene transfer to plants. *Transgenic Research*. **2**, 208–218 (1993), doi:10.1007/BF01977351.
255. *Source DH5alpha* (available at https://www.neb-online.de/).
256. J. J. Doyle, J. L. Doyle, Isolation ofplant DNA from fresh tissue. *Focus*. **12**, 39–40 (1990).
257. K. Edwards, C. Johnstone, C. Thompson, A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*. **19**, 1349 (1991), doi:10.1093/nar/19.6.1349.
258. Shorty buffer stock solution (5X). *Cold Spring Harbor Protocols*. **2009**, pdb.rec11660-pdb.rec11660 (2009), doi:10.1101/pdb.rec11660.
259. A. Froger, J. E. Hall, Transformation of plasmid DNA into E. coli using the heat shock method. *Journal of Visualized Experiments : JoVE*, 253 (2007), doi:10.3791/253.

260. H. Chen, R. S. Nelson, J. L. Sherwood, Enhanced recovery of transformants of Agrobacterium tumefaciens after freeze-thaw transformation and drug selection. *BioTechniques*. **16**, 664—8, 670 (1994).
261. M. Mersereau, G. J. Pazour, A. Das, Efficient transformation of Agrobacterium tumefaciens by electroporation. *Gene*. **90**, 149–151 (1990), doi:10.1016/0378-1119(90)90452-w.
262. K. Unland, *Functional characterisation of squalene synthases, squalene epoxidases and oxidosqualene cyclases involved in the biosynthesis of triterpenoids identified in separated natural rubber of Taraxacum koksaghyz, PhD thesis, University of Münster* (2018).
263. C. Anders, M. Jinek, In vitro Enzymology of Cas9. *Methods in enzymology*. **546**, 1–20 (2014), doi:10.1016/B978-0-12-801185-0.00001-5.
264. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**, 4673–4680 (1994), doi:10.1093/nar/22.22.4673.
265. M. J. Bishop, A. E. Friday, Evolutionary trees from nucleic acid and protein sequences. *Proceedings of the Royal society of London. Series B. Biological sciences*. **226**, 271–302 (1985).
266. E. Afgan *et al.,* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*. **46**, W537-W544 (2018), doi:10.1093/nar/gky379.
267. B. J. Haas *et al.,* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*. **8**, 1494–1512 (2013), doi:10.1038/nprot.2013.084.
268. M. G. Grabherr *et al.,* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech*. **29**, 644–652 (2011).
269. D. Balabanova, T. Remans, A. Vassilev, A. Cuypers, J. Vangronsveld, Possible involvement of glutathione S-transferases in imazamox detoxification in an imidazolinone-resistant sunflower hybrid. *Journal of plant physiology*. **221**, 62–65 (2018), doi:10.1016/j.jplph.2017.12.008.
270. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*. **47**, D506-D515 (2019), doi:10.1093/nar/gky1049.
271. I. Letunic, P. Bork, 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*. **46**, D493-D496 (2018), doi:10.1093/nar/gkx922.
272. A. Morgulis *et al.,* Database indexing for production MegaBLAST searches. *Bioinformatics (Oxford, England)*. **24**, 1757–1764 (2008), doi:10.1093/bioinformatics/btn322.
273. Z. Zhang, S. Schwartz, L. Wagner, W. Miller, A greedy algorithm for aligning DNA sequences. *Journal of computational biology : a journal of computational molecular cell biology*. **7**, 203–214 (2000), doi:10.1089/10665270050081478.
274. C. Schulze Gronover, personal communication, 9 May 2017.
275. T. Colbert *et al.,* High-Throughput Screening for Induced Point Mutations. *Plant Physiology*. **126**, 480–484 (2001), doi:10.1104/pp.126.2.480.
276. M. Kurowska *et al.,* TILLING: a shortcut in functional genomics. *Journal of applied genetics*. **52**, 371–390 (2011), doi:10.1007/s13353-011-0061-1.
277. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biology*. **11**, R106 (2010), doi:10.1186/gb-2010-11-10-r106.

278. G. Breccia *et al.,* Contribution of non-target-site resistance in imidazolinone-resistant Imisun sunflower. *Bragantia*. **76**, 536–542 (2017), doi:10.1590/1678-4499.2016.336.
279. W. Liu *et al.,* Non-target site-based resistance to tribenuron-methyl and essential involved genes in Myosoton aquaticum (L.). *BMC Plant Biology*. **18**, 225 (2018), doi:10.1186/s12870-018-1451-x.
280. R. A. Salas-Perez *et al.,* RNA-Seq transcriptome analysis of Amaranthus palmeri with differential tolerance to glufosinate herbicide. *PLoS ONE*. **13**, e0195488 (2018), doi:10.1371/journal.pone.0195488.
281. I. C. Q. Tas, P. J. van Dijk, Crosses between sexual and apomictic dandelions (Taraxacum). I. The inheritance of apomixis. *Heredity*. **83**, 707–714 (1999).
282. P. J. van Dijk, I. C. Q. Tas, M. Falque, T. Bakx-Schotman, Crosses between sexual and apomictic dandelions (Taraxacum). II. The breakdown of apomixis. *Heredity*. **83**, 715–721 (1999).
283. K. J. M. Hodgson-Kratky, M. N. K. Demers, O. M. Stoffyn, D. J. Wolyn, Harvest date, post-harvest vernalization and regrowth temperature affect flower bud induction in Russian dandelion (Taraxacum kok-saghyz). *Canadian Journal of Plant Science*. **95**, 1221–1228 (2015).
284. M. G. Claros *et al.,* Why assembling plant genome sequences is so challenging. *Biology (Basel)*. **1**, 439–459 (2012), doi:10.3390/biology1020439.
285. H. Badouin *et al.,* The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*. **546**, 148–152 (2017), doi:10.1038/nature22380.
286. Campbell, N. A., Reece, J, Markl, J. (editor), *Biologie* (Pearson Studium, München, ed. 6, 2006).
287. W. Sabetta, V. Alba, A. Blanco, C. Montemurro, sunTILL: a TILLING resource for gene function analysis in sunflower. *Plant methods*. **7**, 20 (2011), doi:10.1186/1746-4811-7-20.
288. D. L. Shaner, in *Plant Nitrogen Metabolism*, J. E. Poulton, J. T. Romeo, E. E. Conn, Eds. (Springer US, 2012), pp. 240–243.
289. M. Eggert, personal communication, 18 October 2018.
290. D. R. Krieg, Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants. *Genetics*. **48**, 561 (1963).
291. T. P. Alcantara, P. W. Bosland, D. W. Smith, Ethyl methanesulfonate-induced seed mutagenesis of Capsicum annuum. *Journal of heredity*. **87**, 239–241 (1996).
292. W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. **31**, 233–239 (2013), doi:10.1038/nbt.2508.
293. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. **31**, 3406–3415 (2003), doi:10.1093/nar/gkg595.
294. M. Stemmer, T. Thumberger, M. del Sol Keyer, J. Wittbrodt, J. L. Mateo, CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS ONE*. **10**, e0124633 (2015), doi:10.1371/journal.pone.0124633.
295. M. Haeussler *et al.,* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol*. **17**, 148 (2016), doi:10.1186/s13059-016-1012-2.
296. S. Engel *et al.,* Role of a conserved arginine in the mechanism of acetohydroxyacid synthase catalysis of condensation with a specific ketoacid substrate. *The Journal of biological chemistry*. **279**, 24803–24812 (2004).
297. D. MASSA, B. KRENZ, R. GERHARDS, Target-site resistance to ALS-inhibiting herbicides in Apera spica-venti populations is conferred by documented and previously

unknown mutations. *Weed Research*. **51**, 294–303 (2011), doi:10.1111/j.1365-3180.2011.00843.x.

298. S. B. Powles, Q. Yu, Evolution in action: plants resistant to herbicides. *Annual Review of Plant Biology*. **61**, 317–347 (2010), doi:10.1146/annurev-arplant-042809-112119.
299. C. M. Whaley, H. P. Wilson, J. H. Westwood, A New Mutation in Plant Als Confers Resistance to Five Classes of Als-inhibiting Herbicides. *Weed Science*. **55**, 83–90 (2007), doi:10.1614/WS-06-082.1.
300. B. Iaffaldano, Y. Zhang, K. Cornish, CRISPR/Cas9 genome editing of rubber producing dandelion Taraxacum kok-saghyz using Agrobacterium rhizogenes without selection. *Industrial Crops and Products*. **89**, 356–362 (2016), doi:10.1016/j.indcrop.2016.05.029.
301. G. Liu, K. Yin, Q. Zhang, C. Gao, J.-L. Qiu, Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency in vivo. *Genome Biology*. **20**, 145 (2019), doi:10.1186/s13059-019-1762-8.
302. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. **343**, 80–84 (2014).
303. P. D. Hsu *et al.,* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*. **31**, 827–832 (2013), doi:10.1038/nbt.2647.
304. M. I. E. Uusi-Mäkelä *et al.,* Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). *PLOS ONE*. **13**, e0196238-e0196238 (2018), doi:10.1371/journal.pone.0196238.
305. D. C. Klein, S. J. Hainer, Genomic methods in profiling DNA accessibility and factor localization. *Chromosome research* (2019), doi:10.1007/s10577-019-09619-9.
306. C. Miguel, L. Marum, An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. *Journal of experimental botany*. **62**, 3713–3725 (2011), doi:10.1093/jxb/err155.
307. A. Scheben, F. Wolter, J. Batley, H. Puchta, D. Edwards, Towards CRISPR/Cas crops – bringing together genomics and genome editing. *New Phytol*. **216**, 682–698 (2017), doi:10.1111/nph.14702.
308. K. Chen, Y. Wang, R. Zhang, H. Zhang, C. Gao, CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annual Review of Plant Biology*. **70**, 667–697 (2019), doi:10.1146/annurev-arplant-050718-100049.
309. *Preliminary Ruling on Directive 2001/18/EC, Articles 2 and 3* (25.07.2018).
310. E. Waltz, Gene-edited CRISPR mushroom escapes US regulation. *Nature News*. **532**, 293 (2016).
311. Q. Yu, X. Q. Zhang, A. Hashem, M. J. Walsh, S. B. Powles, ALS gene proline (197) mutations confer ALS herbicide resistance in eight separated wild radish (Raphanus raphanistrum) populations. *Weed sci.* **51**, 831–838 (2003), doi:10.1614/02-166.
312. R. Domínguez-Mendez *et al.,* Multiple mechanisms are involved in new imazamox-resistant varieties of durum and soft wheat. *Scientific Reports*. **7**, 14839 (2017), doi:10.1038/s41598-017-13874-3.
313. S. Kumar, Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opin Drug Metab Toxicol*. **6**, 115–131 (2010), doi:10.1517/17425250903431040.
314. D. E. Riechers, K. Kreuz, Q. Zhang, Detoxification without Intoxication. *Plant Physiology*. **153**, 3 (2010), doi:10.1104/pp.110.153601.
315. A. Zabalza, S. Gaston, L. Sandalio, L. Del Río, M. Royuela, Oxidative stress is not related to the mode of action of herbicides that inhibit acetolactate synthase. *Environmental and Experimental Botany - ENVIRON EXP BOT*. **59**, 150–159 (2007), doi:10.1016/j.envexpbot.2005.11.003.

316. M.-S. Rahantaniaina, A. Tuzet, A. Mhamdi, G. Noctor, Missing links in understanding redox signaling via thiol/disulfide modulation: how is glutathione oxidized in plants? *Frontiers in Plant Science*. **4**, 477 (2013), doi:10.3389/fpls.2013.00477.
317. M. Gil, A. C. Ochogavía, T. Vega, S. A. Felitti, G. Nestares, Transcript Profiling of Non-Target-Site Imidazolinone Resistance in Imisun Sunflower. *Crop Science*. **58**, 1991 (2018), doi:10.2135/cropsci2018.01.0074.

8 Abbreviations

AA	amino acid
AHAS	acetohydroxyacid synthase
AHB	2-aceto-2-hydroxybutyrate
AL	2-acetolactate
ALS	acetolactate synthase
ATP	adenosine triphosphate
BAP	6-benzylaminopurine
BCAA	branched chain amino acid
BLAST	basic local alignment search tool
bp	base pairs
° C	degree Celsius
Cas	CRISPR associated
CDS	coding region of a gene
cg	continuous genotype
CRISPR	clustered regularly interspaced short palindromic repeats
CSU	catalytic subunit
Da	Dalton
DE	differentially expressed
Del	deletion
DNA	deoxyribonucleic acid
DSB	double-strand break
EDTA	ethylenediaminetetraacetic acid
e.g.	*exempli gratia*, for example
EMS	ethyl methanesulfonate
etc.	et cetera, and so forth
EU	European Union
FAD / $FADH_2$	flavin adenine dinucleotide
fw /fwd	forward
G	giga
g	gram
GMO	genetically modified organism
GT	genotype
h	hour
ha	hectar
HR	homologous recombination
IAA	indole-3-acetic acid
IMI	imidazolinone
In	insertion
InDel	insertion(s) and deletion(s)
IUPAC	International Union of Pure and Applied Chemistry
k	kilo
KB	2-ketobuyrate
L	liter
lx	lux
m	milli
mcs	multiple cloning site
µ	micro
n	nano
NAD / NADH	nicotinamide adenine dinucleotide
n.d.	not determined
NHEJ	non-homologous end joining
no.	number
%	percent
p	pico
PAM	protospacer-adjacent motif
PCR	polymerase chain reaction
POX	pyruvate oxidase
PPT^R	phosphinothricin resistance
PYR	pyruvate
RACE	Rapid Amplification of cDNA ends
rev	reverse
RNA	ribonucleic acid
ROS	reactive oxygen species
RSU	regulatory subunit
SDN	site-directed nuclease
sgRNA	single guide RNA
SNP	single-nucleotide polymorphism
$Spec^R$	spectomycine resistance
SU	sulfonylurea
Subst.	substitution
t-DNA	transfer DNA
ThDP	thiamine diphosphate
TILLING	Targeting Induced Local Lesions IN Genomes
U.S.A.	United States of America
UTR	untranslated region
v	volume
WT	wild type

9 Lebenslauf / Curriculum vitae

10 Thanks

This work was part of a bilateral cooperation in the framework of the project EVITA (031A285B) funded by the Federal Ministry of Education and Research.

Many people have played their part in the completion of this work and I want to express my sincere thanks to them:

Prof. Dr. Dirk Prüfer for the supervision of the PhD at WWU, the support and for the patience.

Dr. Joachim Schiemann and Dr. Ralf Wilhelm for enabling the research activities and completion of the work at the Institute for Biosafety in Plant Biotechnology at the Julius Kuehn-Institute in Quedlinburg.

Prof. Dr. Eva Liebau for her support as a member of the committee.

Katja Thiele and Dr. Frank Hartung for the discussions, ideas and encouragement from the very beginning to the very end. I enjoyed spending the years of doctoral studies with you and will have cherished memories of the time!

Dr. Thorben Sprink for providing the vectors and introduction to CRISPR and Dr. Janina Metje-Sprink for providing in-house prepared Cas9.

The bioinformatics group of the Julius Kuehn-Institute headed by Dr. Jens Keilwagen for the support in analyzing the RNA data; for helping to search for and find correlations in data.

Tina, Bärbel, Kathrin, Astrid, Angelika, Antje and Antje, Vivien, Sandra, Yvonne for their helping hands in sampling, crossing, planting, watering and advice and simply for the memorable time spent together in the lab or outside among the dandelions.

Christian, Nicole, Gianina and Daniela and the "Prüfer"-PhD students for the support and for making my time in Münster a good memory because of you.

Marie – for the conversations around Taraxacum and everything else important in life. I very much enjoyed sharing the office with you and continue to appreciate your advice and food for thought.

Enikö – for the many evenings together, which have always given me strength for new things.

My friends – for listening and sharing good words.

My parents and sister – for their continued support and encouragement.

Jochen – for your encouragement, criticism, motivation and existence.


JKI
Julius Kühn-Institut
Bundesforschungsinstitut für Kulturpflanzen