Studying salicylic acid function in natural accessions of *Arabidopsis thaliana* using CRISPR/Cas technology

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Efthymia Symeonidi

aus Larissa, Griechenland

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	29.10.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Detlef Weigel
2. Berichterstatter:	Prof. Dr. Thomas Lahaye

To my dad,

for always supporting and believing in me.

Acknowledgments

I would like to thank,

Detlef Weigel, for this great opportunity of working in his department, for his constant support and guidance through the hard times, for always being available to discuss with me and of course for always believing in me.

Thomas Lahaye, for agreeing to be my second supervisor, for evaluating my thesis, being member of my PhD and Thesis committee and for his support.

Hernán Burbano, Martin Bayer and Klaus Harter for being part of my PhD and/or Thesis committee and for their support and inputs over the years.

Rebecca Schwab, for her continuous support, help, encouragement, advices and kindness for always sharing gummy bears with me.

Anette Habring-Müller, Anna-Lena Van de Weyer, Christa Lanz, Derek Lundberg, Felix Bemm, Fernando Rabanal, Gautam Shirsekar, Ilja Bezrukov, Julia Hildebrandt, Julian Regalado, Lei Li, Alba Gonzalez, Kavita Venkataramani, Manuela Neumann, Moises Exposito-Alonso, Rui Wu, Sonja Kersten, Talia Karasov, Ulrich Lutz, Wangsheng Zhu, Wei Yuan, Beth Rowan, Diep Tran, Dino Jolic, Eunyoung Chae, Maricris Zaidem, Subhashini Muralidharan, Chang Liu, Wanyan Xi, Clemens Weiss, Kelly Swarts, Laura Morales, Patricia Lang, Sergio Latorre, Giovanna Capovilla and Ignacio Rubio Somoza for all the questions that they answered and the discussions.

Patricia Lang for the translation of my abstract to German and for doing what Google translate could not.

Julian Regalado for writing PlexSeq and for being an amazing flat mate.

Giovanna Capovilla, Talia Karasov, Kavita Venkataramani, Patricia Lang, Bogdan Sieriebriennikov, Michael Werner, Gautam Shirsekar, Madhavi Patil, Noemi Skorzinski, Cristina Barragan, Alessandro Mongera, Moises Exposito-Alonso, Diep Tran, Dino Jolic, Maricris Zaidem, Subhashini Muralidharan, Derek Lundberg, Julian Regalado, Anna Yuan, Sergio Latorre, Clemens Weiss, Leily Rabbani, Hernán Burbano, Rebecca Schwab, Thanvi Srikant, Bridgit Waithaka, Fiona Paul, Max Collenberg, Daphne Welter, Zach Henseler, Jessica Sutter, Anna-Lena Van de Weyer, Fernando Rabanal, Christian Kubica, Rui Wu, Sonja Kersten, Wei Yuan, Beth Rowan, Kelly Swarts, Laura Morales, Ezgi Dogan, Sebastian Petersen, Silvio Collani, Travis Wrightsman, Rafal Gutaker, Jorge Kageyama, Anna Yuan and Emanuele Scacchi for their friendship and making life in Tübingen so amazing.

Natasa Heretidou, Ellie Hatzokou, Bill Papakostas, George Papadopoulos, Karolina Zaborska, Niki Kourtoglou and Eirini-Maria Lampraki for all their support and love from far away.

Gio, Talia, Mike, Bogdan and Laika for always being there for me.

Laika for his unconditional love and excitement every time he “sees” me and for being the best godson one can ask.

My parents, my sister and my family for their constant support and motivation to follow my dreams and being there for every step.

Table of Contents

Abstract

The plant hormone salicylic acid (SA) has been used for millennia to relieve pain. In modern times, its acetylate form – known as aspirin – has become one of the most popular painkillers. SA production in plants has been implicated in multiple aspects of plant life including thermogenesis, response to stresses, seed viability, leaf senescence, flowering time regulation and immune responses. It has been reported that regulation of SA accumulation exhibits natural variation, for example due to the positive regulator *ACCELERATED CELL DEATH 6* (*ACD6*). But still little is known regarding variation of the main biosynthetic pathway of SA, its involvement in flowering time regulation or its effect regarding the virulence of microbes.

To elucidate natural variation in the synthesis of SA and its effect on other processes, I utilised a newly introduced genome editing approach called CRISPR/Cas system. This tool allows the user to edit a targeted region using an endonuclease (Cas9) and an artificial RNA (sgRNA). With this method I selectively knocked out the *ICS1* gene and disrupted the main SA biosynthetic pathway in seven natural accessions of *A. thaliana*. To efficiently screen the large number of individuals that required genotyping for this workflow (>900) I developed a preparation and analysis pipeline using deep amplicon sequencing (CRISPR-finder). Using this pipeline, individuals carrying variants at the targeted region can be identified within a few days in a cost-effective and precise manner.

Using the *ics1* lines I generated, in spite of the non-functional *ICS1* allele, I was able to detect residual SA in most of the lines. These results suggested that the alternative PAL biosynthetic pathway or the *ICS2* gene may be responsible for appreciable levels of SA production in some genotypes. Additionally, I concluded that flowering time was not significantly affected by decreased levels of SA in these seven natural accessions, even though SA had been linked to flowering before. Significant reduction of SA accumulation was observed for all the mutant lines, but to different degrees, when compared to the corresponding wild types, during *Pseudomonas syringae* (*Psm*4326) infection. When the oomycete *Hyaloperonospora arabidopsidis* (14OHMLP04) was used for infecting individuals, no significant induction of SA accumulation was detected when plants exhibited resistance (flecking necrosis or trailining necrosis) or susceptibility. For the genotypes characterised with complete resistance (flecking necrosis) no effect of SA

was observed. Interestingly, increased severity of trailing necrosis was observed for the *ics1* mutant lines derived from partially resistant genotypes (trailing necrosis symptom). There were also accessions and their corresponding *ics1* mutant lines that showed susceptibility and pathogen growth. These findings suggest that resistance to the pathogen isolate is SA-independent and that decreased levels of SA allow an increased manifestation of the potential virulence of 14OHMLP04.

Finally, the commonly used Col-0 accession was used for investigating a controversial aspect of the CRISPR/Cas system – off-target cleavage. This can occur due to non specific Cas9 activity. For assessing off-target variants, I used the same sgRNAs as the ones used for generating the *ics1* mutant lines. During the investigation, I was able to detect a very small number of incidents (0, 2 and 5 depending on the line in question) that could potentially be attributed to Cas9's cleavage activity. These numbers do not differ dramatically from the expected *de novo* mutation rate in *A. thaliana* which is approximately one mutation per generation. This makes it hard to confidently assign detected variants to Cas9 cleavage or *de novo* generation. These findings do not exclude that off-target cleavage events can occur when the CRISPR/Cas system is used, but they are rare and may be difficult to detect.

Zusammenfassung

Das Pflanzenhormon Salicylsäure (SA) wird seit Jahrtausenden zur Schmerzlinderung eingesetzt. In der heutigen Zeit ist seine Acetylatform - bekannt als Aspirin - eines der beliebtesten Schmerzmittel. SA-Produktion in Pflanzen wird mit vielfältigen Aspekten der Pflanzenwelt in Verbindung gebracht, darunter Thermogenese, die Reaktion auf Stress, Samenlebensfähigkeit, Blattseneszenz, Blühzeitregulierung und Immunantworten. Es ist bekannt dass die Regulation von SA-Akkumulation natürliche Schwankungen aufweist, beispielsweise aufgrund des positiven Reglers ACCELERATED CELL DEATH 6 (ACD6). Über die Variation des Hauptbiosynthesewegs von SA, die Beteiligung von SA an der Blühzeitregulierung oder seine Wirkung auf die Virulenz von Mikroben ist jedoch noch wenig bekannt.

Um die natürlichen Unterschiede in der Synthese von SA und seine Auswirkungen auf andere Prozesse aufzuklären, verwendete ich eine neue Methode zur Genomeditierung, das sogenannte CRISPR / Cas-System. Dieses Werkzeug erlaubt dem Benutzer das Editieren einer Zielregion im Genom mittels einer Endonuklease (Cas9) und einer künstlichen RNA (sgRNA). Mit dieser Methode generierte ich selektiv knock-out Varianten des ICS1-Genes und unterbrach den Haupt-Biosyntheseweg von SA in sieben natürlichen Akzessionen von *Arabidopsis thaliana*. Um die große Anzahl von Individuen, die in diesem Arbeitsablauf genotypisiert werden mussten (>900), auf effiziente Art und Weise zu sichten, entwickelte ich einen Verarbeitungs- und Analyseablauf, der Deep Amplicon Sequencing (CRISPR-Finder) integriert. Dieser Arbeitsfluss ermöglicht es, Individuen, die genetische Varianten in der Zielregion aufweisen, innerhalb weniger Tag kosteneffizient und präzise zu identifizieren.

In den generierten *ics1*-Linien konnte ich, trotz des nicht-funktionalen *ICS1*-Allels, in den meisten Fällen Restbestände von SA nachweisen. Diese Ergebnisse legen nahe, dass der alternative PAL-Biosyntheseweg oder das *ICS2*-Gen in manchen Genotypen für nachweisbare Produktion von SA verantwortlich sein könnten. Zusätzlich kam ich zu dem Schluss, dass die Blütezeit in diesen sieben natürlichen Akzessionen nicht signifikant von den verringerten SA-Konzentrationen beeinflusst wurde, obwohl SA zuvor mit Pflanzenblüte in Verbindung gebracht worden war. Bei Infektion mit *Pseudomonas syringae* (*Psm*4326) wurde in allen Mutantenlinien im Vergleich mit dem entsprechenden Wildtypen eine signifikante,

wenngleich in ihrem Ausmaß variierende, Verringerung der SA-Akkumulation beobachtet. Wenn der Oomycet *Hyaloperonospora arabidopsidis* (14OHMLP04) zur Infektion von Individuen verwendet wurde, konnte keine signifikante Induktion von SA-Akkumulation nachgewiesen werden, wenn Pflanzen Resistenz (gesprenkelte (flecking) oder Folge-Nekrose (trailing necrosis)) oder Anfälligkeit zeigten. Für die Genotypen, die als vollständig resistent (gesprenkelte Nekrose) charakterisiert waren, wurde keine Wirkung von SA festgestellt. Interessanterweise wurde für die *ics1*-Mutantenlinien, die von teilweise resistenten Genotypen abstammten, ein erhöhtes Ausmaß der Folge-Nekrose beobachtet (Folge-Nekrose-Symptom). Es gab auch Akzessionen und ihre entsprechenden *ics1*-Mutantenlinien, die Anfälligkeit und Pathogenwachstum zeigten. Diese Befunde legen nahe, dass die Resistenz gegenüber dem Pathogen-Isolat SA-unabhängig ist, und dass verminderte SA Niveaus eine verstärkte Manifestierung der potentiellen Virulenz von 14OHMLP04 ermöglichen.

Schließlich wurde die häufig verwendete Col-0 Akzession zur Untersuchung eines umstrittenen Aspekts des CRISPR / Cas Systems verwendet: unspezifische Gen-Editierung, sogenannte off-target Effekte. Diese können aufgrund unspezifischer Aktivität von Cas9 auftreten. Um off-target Varianten auszuwerten, verwendete ich die gleichen sgRNAs, die auch für die Erzeugung der *ics1* Mutantenlinien benutzt wurden. Im Lauf der Nachforschungen konnte ich eine sehr kleine Anzahl an Vorfällen (0, 2 und 5, abhängig von der jeweiligen Linie) nachweisen, die möglicherweise auf die “Genscheren-Aktivität” von Cas9 zurückzuführen sind. Diese Zahlen unterscheiden sich nicht dramatisch von der erwarteten *de novo* Mutationsrate in *A. thaliana*, die etwa einer Mutation pro Generation entspricht. Dies macht es schwierig, die Varianten sicher entweder Cas9-Aktivität oder natürlicher Mutation zuzuweisen. Diese Befunde schließen nicht aus, dass off-target “Genscheren-Aktivität” bei der Verwendung des CRISPR-Cas Systems auftreten kann, jedoch ist solche Aktivität selten und kann schwierig zu detektieren sein.

Introduction

Natural variation

Variation between different species (inter-specific) and within a species (intra-specific) is a daily observation. This variation can result from genetic, epigenetic and/or environmental differences. There are two distinct types of variation, the qualitative and the quantitative variation. The qualitative variation is used for describing traits in a population that are found in two or more distinct classes. These phenotypes usually derive from a single gene that has multiple alleles, like the flower color in peas during Mendel's experiments, although they can also involve complex interactions in which multiple genes are involved. The combination of the alleles describe the genotype of an individual. In quantitative variation the phenotypes exhibit a ged distribution without a distinct separation. The genetic basis of quantitative variation is largely thought to be the result of the activity of a large number of genes. Beside genetic variation, the environmental influence is very prominent in quantitative variation regarding the phenotype of an individual. For this reason the interaction between genotype and phenotype are often complex. This complexity obscures the identification of one-to-one correspondences between genotype and phenotype. Quantitative variation is more common in nature and can be observed in the plant and animal kingdoms (fruit size, crop yield etc.) including in human populations (height, weight etc.). The intermediate phenotypes of quantitative variation are more abundant than the extreme ones and the distribution of the phenotypic frequencies often follow a bell shape (Griffiths et al. 1999) (Figure 1).

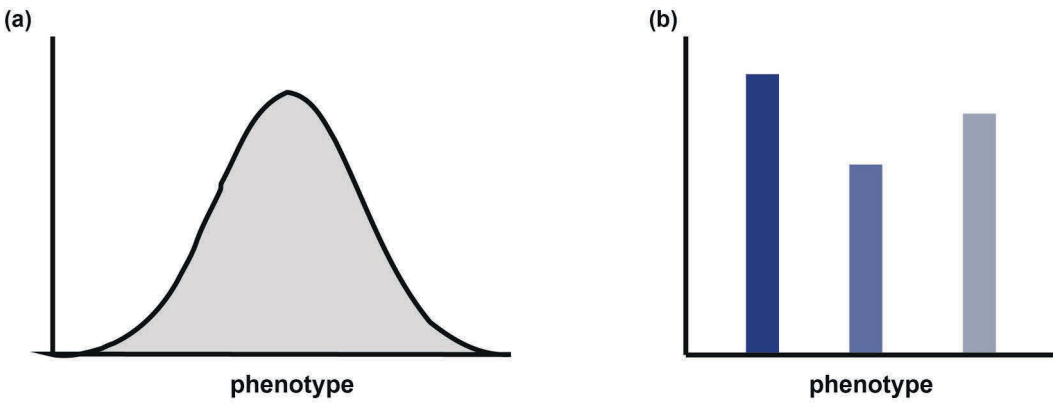


Figure 1: Schematic representation of phenotypic distribution of traits exhibiting (a) quantitative (black line) or (b) qualitative (blue bars) variation. (a) The trait with quantitative variation shows a continuous distribution of phenotypes (black line) **(b)** while the one with qualitative variation shows discrete phenotypes (blue bars).

Natural phenotypic variation is observed in part as a result of adaptation to different environments or niches within an environment and is, in the absence of environmental variation, largely the result of genetic variation. An as-of-yet underappreciated type of genetic variation is the cryptic variation that does not have an immediate effect on the phenotype but can secure survival in the case of extreme changes (environmental or genetic nature).

Animal kingdom

A great example of natural variation within species is the example of differences in wing patterning between butterflies. These differences function as differential defence mechanisms against predators, camouflage and/or indication of toxicity (Merrill et al. 2012; Olofsson et al. 2013). More specifically two genera, *Heliconius* and *Papilio* show intra-specific variation in wing pigmentation due to different genetic loci interactions and this variation leads to mimicry with other species for avoiding predators (Nadeau 2016). Putatively adaptive variation in exoskeleton color is observed also in beetle populations. Two dung beetles species *Geotrupes auratus* Motsch and *G. laevistriatus* Motsch exhibit different coloration depending on their habitat with individuals of the different species looking more similar when they share habitats (Watanabe et al. 2002). North American ants also exhibit inter-specific variation in thermotolerance, which is vital for climate adaptation (Verble-Pearson et al. 2015). Two sibling species of cricket frogs that occupy different habitats exhibit variation in body size and foot length that is correlated with different environmental cues (Nevo 1973).

Plant kingdom

The plant kingdom offers great examples of natural variation and putative adaptation maintaining this variation. There is great variation in flower shape, scent, color all of which seem to have a strong genetic basis (Fenster & Ritland 1994;

Galen 1999; Schemske & Bierzychudek 2007; Delle-Vedove et al. 2017). Variation found in plant species and especially in crops is very important due to their economical value when bred for desired traits. For example, the kernel size in maize shows differences in different populations and it is affected by at least 10 different loci in the genome (Chen et al. 2016). Additionally, defense against insect herbivores in maize shows variation that is attributed to numerous loci as well (Meihls et al. 2012). Rice biomass was linked to genetic variation that can help improving biofuels production (Jahn et al. 2011). Many great example come from *Arabidopsis thaliana* that exhibits variation in various traits like flowering time, cell death pattern, immune responses, seed dormancy, germination and mineral accumulation (Alonso-Blanco et al. 2009). For example it has been shown that two main regulators of flowering time, *FRIGIDA* and *FLOWERING LOCUS C* (*FLC*), are responsible for most of the differences in flowering time that are observed (Lempe et al. 2005). Another important trait that shows variation is the activity of the immune system against pathogens. More specifically, it is well documented that different accessions of *A. thaliana* show divergent defense responses to the same pathogen strain (for example against *Fusarium graminearum*) (Chen et al. 2006). But usually the response to a pathogen is complex and hard to unravel due to involvement of many loci with different functions, such as the response to *B. cinerea* (Rowe & Kliebenstein 2008). Finally, when the transcriptome of natural accessions of *A. thaliana* was evaluated in respect to exogenous salicylic acid treatment a considerable level of genetic variation was revealed (one third of the differences was due to accession x treatment interactions) (van Leeuwen et al. 2007)

I am particularly interested in the contribution of salicylic acid (SA) to differences in flowering time and the immune responses in natural accessions of *A. thaliana*. SA has been implicated in flowering time regulation and has been shown to to be a vital component during pathogen infection (Delaney et al. 1994; Nawrath & Métraux 1999).

Salicylic acid (SA)

Plants that exhibit exceptional levels of salicylic acid (SA), and its derivatives (salicylates), have been used in medicine for millennia. Already in the fourth century BC, the father of modern medicine, Hippocrates, observed that willow leaves or bark

extract can relieve from fever and childbirth pain. Other civilizations similarly used plants rich in salicylates such as the Babylonians, Chinese, Assyrians. In 1763 reverend Edward Stone was the first to study clinically the effects of the willow tree (Stone 1763) and in 1828, Johann Buchner purified the active ingredient in the form of yellow, bitter taste crystals that he named salicin. Ten years later, Raffaele Piria managed to separate salicin in a sugar and an aromatic compound that he converted through hydrolysis and oxidation to colourless crystals, which he named SA (as a tribute to *Salix alba*, the taxonomic name of white willow) (Jourdier, Sophie. "Miracle Drug." *Royal Society of Chemistry*, 1999.) The salicylates received from different plants are converted to SA naturally in humans and animals after digestion (Klessig et al. 2016). In 1874 production of synthetic SA started due to high demand in the market but soon its prolonged use was linked with some negative effects like stomach irritation. Felix Hofmann was the first to acetylate SA (ASA) converting it this way to a more compatible form for the human body with all the medicinal properties intact. In 1897 scientists at Bayer and Company initiated the production of ASA and in 1899 it was named aspirin from "a" acetyl and "spirin" in honour to the plant that was used for extracting SA (*Spiraea ulmaria* or meadowsweet) (Witthauer 1899).

SA involvement in plant life

But why do plants need SA? First of all, SA has been shown to be involved in a diversity of processes including regulation of response to biotic and abiotic stresses, thermotolerance, seedling viability, leaf senescence, thermogenesis, seed germination, legumes' nodulation (Rate et al. 1999; Morris et al. 2000; Metwally et al. 2003; Clarke et al. 2004; Norman et al. 2004; Stacey et al. 2006; Rajjou et al. 2006; Vlot et al. 2009). It has also been suggested that it is involved in the flowering time regulation. The first observation to document this relationship was reported in (Cleland & Ajami 1974) in which SA was shown to induce flowering of *Lemna gibba* G3 in short day conditions. This is supported by studies where the flowering phenotype of *Arabidopsis* mutants correlated with the levels of SA (Jin et al. 2007; Li et al. 2012; Villajuana-Bonequi et al. 2014a). A more direct example was presented by (Martínez et al. 2004b) where a SA-deficient mutant (NahG) showed delayed flowering. However, other studies were unable to observe such a connection. In the

case of *win3* mutants, the accelerated transition to the reproductive phase was associated with decreased levels of SA and in the case of *MYB30,* overexpression leads to accumulation of SA without any effect on the flowering time of the mutants (Wang et al. 2011; Liu et al. 2014).

PTI and ETI

SA is mainly known for its involvement during plant immune responses against hemi-/biotrophic pathogens. The mechanism against infections consists of different layers depending on the pathogen and the ability of the plant to respond (Jones & Dangl 2006). There are two levels of recognition during immune responses. The first is based on receptors (pattern recognition receptors (PRRs)) on the cell surface that can recognise molecules with specific patterns that are unique and conserved among microbes. These patterns are pathogen-/microbe-associated molecular patterns (PAMPs/MAMPs) and their recognition leads to the pattern-triggered immunity (PTI). The plant has also the ability to recognise damage-associated molecular patterns (DAMPs) that derive from plant products that have been degraded during pathogen invasion (Boller & Felix 2009). Initialisation of PTI leads to accumulation of reactive oxygen species, intracellular calcium influx, transient activation of mitogen-activated protein kinases (MAPKs), and defense hormone production (Tsuda et al. 2008; Tsuda & Katagiri 2010). Some pathogens have managed to overcome PTI through its suppression. They accomplish this through secretion of molecules called effector proteins (Dangl & Jones 2001; Vleeshouwers & Oliver 2014). These molecules have taken the name due to their property of affecting the genotype of the host (Vleeshouwers & Oliver 2014). Secretion of effectors could lead to two different outcomes regarding the response of the host, either to the susceptibility of the host during which the pathogen can successfully establish infection (incompatible interaction) or to the elicitation of the effector-triggered immunity (ETI) response against the pathogen (compatible interaction) with the host regaining resistance (Dangl & Jones 2001; Katagiri et al. 2002). ETI resistance is based on the direct or indirect recognition of the effectors by endogenous products of resistance (R) genes and often follows the “gene-for-gene” model in which resistance is conferred only when an R gene product can recognise the effector expressed by a pathogen avirulence gene (Flor 1956; Dangl & Jones

2001). The response to ETI resembles the one that is observed during PTI with the difference of being more rapid, robust and prolonged. Additionally, localised cell death, called hypersensitive response (HR), is often observed (Thomma et al. 2011; Hamdoun et al. 2013). ETI can induce systemic acquired resistance (SAR) during which the uninfected tissues show broad-spectrum resistance and SA is involved during the establishment of the response (Ryals et al. 1996). Some evidence suggests that SA could have antimicrobial properties when intercellular washing fluids (IWFs) from young or NahG plants were not able to prevent *Pseudomonas syringae* growth comparing to the one incubated with intercellular washing fluids (IWFs) from old plants (Kus et al. 2002). Production of SA is induced during both PTI and ETI in response to biotrophic or hemibiotrophic pathogen infection (Glazebrook 2005).

SA and immune responses

The first evidence that SA can confer resistance against pathogens was published in 1979 when SA applied to viral infected tobacco plants triggered defense gene expression and repressed the infection (White 1979). Later on, in 1990, two studies showed that SA is produced endogenously and in response to pathogen infection, where after treating tobacco and cucumber plants with pathogens high levels of SA were detected as response to local and/or systemic resistance (Malamy et al. 1990; Métraux et al. 1990). Stronger evidence supporting SA involvement at PTI and ETI surfaced when tobacco and *Arabidopsis* SA deficient mutant lines were studied. Transgenic lines of tobacco or *Arabidopsis* expressing *NahG* gene (salicylate hydroxylase that converts SA to catechol (Yamamato et al. 1965)) were used in different studies showing that after infection plants were unable to accumulate SA, express defense related genes or develop SAR and exhibit enhanced susceptibility (Gaffney et al. 1993; Delaney et al. 1994; Vernooij et al. 1994; Lawton et al. 1995). (Delaney et al. 1994) showed that providing the plants with the SA synthetic analog (2,6-dichloro-isonicotinic acid - INA) the resistance and defense genes' expression were restored. Similar findings were supported by studies where tobacco or *Arabidopsis* plants with defective SA production pathways were used (Pallas et al. 1996; Nawrath & Métraux 1999; Wildermuth et al. 2001; Nawrath et al. 2002). The role of SA in monocots is still not very well understood. SA

or its synthetic analogs INA or benzo(1,2,3) thiadiazole-7-carbothioic acid S-methyl ester (BTH) initiate the expression of PR genes and/or pathogen resistance in wheat (Görlach et al. 1996), rice (Schweizer et al. 1999; Hwang et al. 2008), maize (Morris et al. 1998) and barley (Muradov et al. 1993; Kogel et al. 1994).

Biosynthetic pathways

There are two biosynthetic pathways in plants that can produce SA and both of them require chorismate. The first pathway utilises the enzyme PHENYLALANINE AMMONIA LYASE (PAL) for catalysing the initial reaction (Phenylalanine to trans-Cinnamic acid (tCA)) (Figure 2). Radio-labelling studies using different treatments of different species (pathogen-inoculated tobacco and cucumber, elicitor treated potato) or untreated rice seedling support the production of SA from Phenylalanine (Phe) via Benzoic acid (BA) (Yalpani et al. 1993; Meuwly et al. 1995; Silverman et al. 1995; Coquoz et al. 1998). Furthermore, increased levels of PAL expression and SA were detected in tobacco and *Arabidopsis* plants that exhibit resistance during infection, support its involvement in SA biosynthesis and activation upon infection (Pellegrini et al. 1994; Mauch-Mani & Slusarenko 1996; Dempsey et al. 1999).

The second pathway is based on the conversion of isochorismate to SA on a two-step process using the ISOCHORISMATE SYNTHASE (ICS) and ISOCHORISMATE PYRUVATE LYASE (IPL) enzymes (Figure 2) (Wildermuth et al. 2001; Strawn et al. 2007). In *Arabidopsis thaliana* there are two *ICS* genes encoding proteins that share 88% similarity at the amino acid level (Garcion et al. 2008). According to (Wildermuth et al. 2001) when plants are infected with *Pseudomonas maculicola* only transcripts of *ICS1* are detected and no *ICS2*. *Ics1* mutant lines accumulate approximately 10% of SA comparing to the wild type while *ics2* mutants present similar levels of SA as the wild type (Garcion et al. 2008). Additionally, when *ics2* mutant lines are treated with UV exposure the levels of SA before and after the treatment do not present a significant difference from the wild type (Garcion et al. 2008). All these findings suggest that *ICS2* involvement is minimal in comparison to *ICS1*. Furthermore, mutations in the *ICS1* gene increased the susceptibility to pathogens, decreased PR expression and fail to initiate SAR (Nawrath & Métraux 1999; Dewdney et al. 2000; Wildermuth et al. 2001). Finally, the remaining SA that can be detected in an *ics1/ics2* double mutant indicates that the

ICS pathway is the dominant but not the only one for the production of SA in *Arabidopsis* (Garcion et al. 2008).

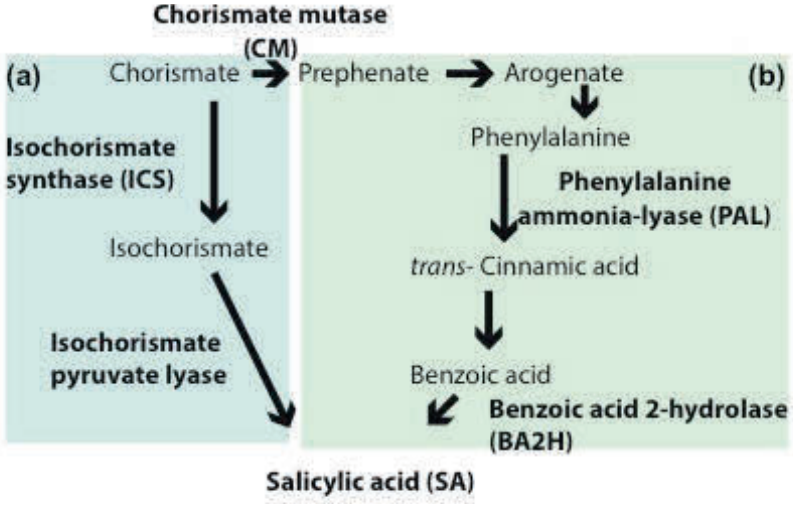


Figure 2: Schematic representation of the two salicylic acid (SA) biosynthetic pathways in *A. thaliana*. (a) First, there is the Isochorismate Synthase (ICS) pathway during which chorismate is converted to Isochorismate and then to SA. With purple writing the targeted gene of this study is noted. (b) Secondly, there is the Phenylalanine ammonia-lyase (PAL) during which Phenylalanine is converted to SA via sequential conversion to trans-Cinnamic acid and Benzoic acid. The main biosynthetic pathway for *A. thaliana* is the ICS (Wildermuth et al. 2001).

SA modifications

Post-production modifications of SA can generate different conjugated forms. One of the most common modifications generates the SA 2-O-β-D-glucoside (SAG) after the glucosylation of SA at its hydroxyl group. *Arabidopsis* has two genes that can catalyse this conversion that are called UDP-glucosyltransferases, *UGT74F1* and *UGT74F2* (Lim et al. 2002; Song 2006; Dean & Delaney 2008). The first preferentially catalyses the conversion of SA to SAG where the second catalyses the conjugation of SA at the carboxyl group generating salicylate glucose ester (SGE) (Dean & Delaney 2008). Similar to tobacco (Dean et al. 2005) and soybean (Dean et al. 2003), SAG in Arabidopsis is believed to be produced in the cytosol and then transported to the vacuole, where it is stored for later use during immune responses after conversion back to SA (Hennig et al. 1993).

Another modification is the methylation of SA that gives rise to the methyl SA (MeSA). This process is catalysed by the BA/SA carboxyl methyltransferase 1 (BSMT1) (Chen et al. 2003). MeSA has been implicated in resistance against

insects by attracting their predators and pathogens by acting as the main signal for SAR (Van Poecke et al. 2001; Park et al. 2007). In the beginning, it was believed that SA is responsible for SAR signaling due to its increased levels and PR expression with or just before SAR and its detection in the phloem of infected tobacco and cucumber plants (Métraux et al. 1990; Rasmussen et al. 1991; Yalpani et al. 1991). Later on, leaf detachment assays, in cucumber, indicated that the molecule responsible for SAR left the infected leaf before increased levels of SA were detected in the corresponding petiole (Rasmussen et al. 1991; Smith et al. 1991). Additionally, chimeric NahG-expressing rootstock tobacco plants were able to induce SAR and PR expression after TMV infection in wild type scion leaves (Vernooij et al. 1994). It has been shown, in tobacco, that SA-binding protein 2 (SABP2) is important for establishing SAR in the systemic tissues but not in the primarily infected (Park et al. 2007). The same group has shown that SABP2-deficient tobacco plants are unable to establish SAR, that MeSA levels are increased in primary infected leaves, in the phloem and in systemic tissues and that MeSA treatment in lower leaves can induce SAR in higher tissues (Park et al. 2007). SABP2 exhibits SA-inhibitable methyl salicylate esterase activity which means that can convert MeSA to SA and SA inhibits SABP2 activity which leads to a negative feedback loop (Forouhar et al. 2005). All these findings suggest that MeSA is the mobile signal for SAR in tobacco (Vlot et al. 2008). A similar model had been proposed for *Arabidopsis* but is still unclear if MeSA solely delivers SA to distal tissues or has also other roles (Gao et al. 2015). Interestingly, BSMT1 (converts SA to MeSA) involvement in SAR establishment can be bypassed by longer exposure to light after infection, suggesting that SAR establishment is a complex procedure with many players (Attaran et al. 2009; Liu et al. 2011).

SA can also be found conjugated to amino acids. The best-characterized example of this is Salicyloyl-L-aspartate (SA-Asp). The process is catalysed by acyl-adenylate/thioester-forming enzyme (GH3.5) which can also conjugate the auxin indole acetic acid (Staswick et al. 2005). (Mackelprang et al. 2017) have shown that GH3.5 acts more in the conversion of IAA during growth and more in the conversion of SA during (hemi-) biotrophic pathogen infection. SA-Asp cannot be converted back into active SA and is able to induce only very weak PR-1 expression which indicates that SA-Asp could serve as a limiting factor for extended HR or is mobile

and responsible for promoting defense priming (Y. Chen et al. 2013; Mackelprang et al. 2017).

It has been shown that AtSOT12 acts as a sulfotransferase and conjugates sulphuryl group to SA in vitro (Baek et al. 2010). The authors suggest that the sulfonation of SA occurs mainly upon infection and is used for detoxification of the excess SA but at the same time it plays a role in the resistance since the *sot12* mutant shows compromised responses (Baek et al. 2010). Another modification is the hydroxylation of SA that occurs through a non-enzymatic reaction and generates 2,3- and 2,5-DHBA (dihydrobenzoates) (Maskos et al. 1990). (Bartsch et al. 2010) have shown that the levels of 2,3-DHBA are increased upon infection or due to age and that external application can induce weak PR-1 expression that lead to a hypothesis of being an inactive form of SA. The same authors have suggested that it can also serve a protective role against oxidative stress, like the same way that acts in pathogens (Bartsch et al. 2010). When 2,5-DHBA is applied externally in tomato, cucumber and *Gynura* plants it initiates expression of a subset of *PR* genes that is also present after SA application (Bellés et al. 1999; Bellés et al. 2006).

SA targets

It is believed that animal and plant hormones bind to one or a small number of receptors. It is still not entirely clear exactly how many receptors SA has in *A. thaliana*. It has been proposed that SA binds to non-expressor of pathogenesis-related genes (NPR) protein family members (Fu et al. 2012; Wu et al. 2012). SA can bind to NPR1 and by changing the conformation of the protein abolish its auto-inhibition and promotes activation of defense gene *PR-1* expression (Cao et al. 1997; Wu et al. 2012). SA is involved in other process in plant life which are NPR1-independent. Indeed, almost 30 SA-binding proteins (SABPs) have been identified with different affinities for SA and change in their activity is observed after binding to SA (Klessig et al. 2016). It has been hypothesized that SA interacts with its receptors depending on the location and its affinity based on the changes of SA levels that are affected by the developmental stage and/or (a)biotic stress (Klessig et al. 2016).

Approach

Even though there are naturally occurring null alleles of genes involved in numerous genes there are other loci for which null alleles are not naturally available. Nowadays, with constant improvements in genome editing approaches one can introduce modifications into targeted loci. Being interested in SA involvement in flowering time regulation and during infections I utilised the CRISPR/Cas system in order to knock out *ICS1* gene in seven different natural accessions in *A. thaliana*.

Genome editing

Genome editing is the deliberate alteration (insertion, deletion, replacement or modification) of the genetic code of an organism. It is one of the biotechnological approaches that is most useful for forward or reverse genetics. The ability to alter the expression of genes gives the unique opportunity to investigate gene function and its relationship to the observed phenotype. With the tools that we have at our disposal and the ability to generate mutants in different genetic backgrounds of the same species we have the unique opportunity to tackle questions regarding natural variation. How much does a single gene contribute to a specific phenotype? Is the effect size the same for all the populations? Are some alleles more active in specific genetic backgrounds?

The main tools for genome editing that have been used over the last three decades are based on engineered nucleases, specifically meganucleases (Rudin & Haber 1988), Zinc-Finger nucleases (ZFNs) (Bibikova et al. 2002), Transcription Activator-like Effector nucleases (TALENs) (Huang et al. 2011) and Clustered Regularly Interspaced Short Palindromic Repeats/Cas (CRISPR/Cas) (Cong et al. 2013). The mechanism of action of these nucleases is based on cleavage activity that can be directed (with varying degrees of ease) to desired regions of the genome. These tools will be described in greater detail in the next paragraphs.

The outcome of the cleavage is the generation of a double stranded break (DSB) that can be repaired with one of two different pathways. The first pathway that can be utilised is that of nonhomologous end joining (NHEJ) where the two broken DNA ends are modified in order to become compatible and then are ligated back together (Lieber 1999). In this pathway no template is used for repair, the

modifications are permanent and genetic information can be lost completely. In the other pathway, homology directed repair (HDR), the break is repaired using the homologue information of the intact chromosome. Using this method for genome editing the user can provide a template that will be used for repair (Liang et al. 1998). In this case directed mutagenesis can occur where the outcome can be predicted. In either case the genome alteration can be achieved and utilised for further studies.

Meganucleases

One of the first attempts for genome editing was based on meganucleases. Meganucleases are endodeoxyribonucleases that have the ability to recognise a long stretch of nucleotides (12 to 40) and create double stranded breaks (DBS) (Silva et al. 2011). Due to the length of the recognition site, they cleave double stranded DNA with high specificity. The first application in mammalian cells was performed in mice in 1994 by (Rouet et al. 1994) using the I-SceI meganuclease (endogenous to *Saccharomyces cerevisiae*) to introduce DSBs in the genome. Since then the technology has been used for homologous recombination in yeast and mammalian cells (Epinat et al. 2003) and to modify and correct genes in mammalian cells (Arnould et al. 2007).

With great specificity come great limitations! The natural occurring recognition sites for a meganuclease in the genome are extremely rare. This problem creates obstacles when attempting to edit a specific region of the genome. In order to overcome these difficulties, several researchers attempted to alter the specificity by introducing mutations in the recognition domain of the endonucleases, swapping domains among different endonucleases or using yeast surface display for identifying endonucleases with desired sequence specificity (Ashworth et al. 2010; Grizot et al. 2010; Jacoby et al. 2017). These attempts of improvement have not always been successful and the efficiency of the system is frequently very low. Meganucleases are still in use but newer technologies have been developed for genome editing that are more efficient.

Zinc-Finger nucleases (ZFNs)

This genome editing tool is based on the properties of the natural type IIS restriction enzyme *FokI* and Zinc-Finger (ZF) proteins. (Li et al. 1992) described that the FokI recognition and cleavage domains are different and can be manipulated separately in a way where by altering the recognition domain anything can be cleaved by the enzyme (Kim & Chandrasegaran 1994; Kim et al. 1996). The main innovation of the properties of the enzyme were elucidated by (Kim et al. 1996) in which the authors used Zinc-Finger protein fused with the FokI cleavage domain for cutting λ DNA.

The first ZF protein that was described was identified in *Xenopus laevis* and it was the transcription factor IIIA that has nine repeated domains, each with a zinc ligand (Miller et al. 1985). These proteins bind to Zn^{+2} in order to form their structural domains, called zinc-fingers. When the crystal structure of a Zif268-DNA complex was resolved, it was revealed that each ZF domain interacts with three nucleotides (Pavletich & Pabo 1991). Numerous ZF proteins have been identified and described since the first one (Jayakanthan et al. 2009). Each ZF domain differs from the next one in specific amino acid positions that are flexible allowing for different domains to recognise and bind to different triplet. Since each ZF domain is binding in an independent manner from the others, one can create a customised ZF protein by linking different ZF domains that can recognise and bind in a desired location in the genome (Desjarlais & Berg 1992; Desjarlais & Berg 1993; Desjarlais & Berg 1994).

The first results in which ZF were fused with the FokI cleavage domains were not promising (Kim et al. 1996) but it was later shown by (Bitinaite et al. 1998) and (Smith et al. 2000) that the FokI cleavage domain needs to be dimerised to cut DNA. For this reason, two different ZF proteins, fused with one FokI cleavage domain each, were used for targeting neighbouring sequences in order to create DSB. Each ZF domain was fused to the next one through a linker and the spacer between the target sites was 5 or 6 bp (Bibikova et al. 2001; Händel et al. 2009; Shimizu et al. 2009). Since both ZFNs have to bind precisely to the targeted regions for the DSB to occur, the system has high specificity (Figure 3a).

The first time that ZFN technology was used for generating target mutations was in the fruit fly *Drosophila melanogaster* (Bibikova et al. 2002). A year later also gene replacement was carried out successfully in *Drosophila* (Bibikova et al. 2003).

In both cases the *yellow* locus was targeted in somatic cells but also most importantly in the germline. Since then, the ZFN technology has been used in numerous studies in different organisms. For example, ZFNs have successfully been used in nematodes (*C. elegans*) (Morton et al. 2006), zebrafish (*D. rerio)* (Meng et al. 2008; Doyon et al. 2008), mouse (*M. musculus*) (Meyer et al. 2010; Carbery et al. 2010), rat (*R. norvegicus*) (Geurts et al. 2009), thale cress (*A. thaliana*) (Lloyd et al. 2005; de Pater et al. 2009; Zhang et al. 2010), maize (*Z. mays*) (Shukla et al. 2009) and in mammalian cell lines like human cell lines (Porteus & Baltimore 2003; Urnov et al. 2005; DeKelver et al. 2010) and pig (*S. domestica*) (Watanabe et al. 2010).

Transcription Activator-like effector nucleases (TALENs)

Another genome editing tool that has been used over the last two decades is based on the transcription activator-like effector (TALE) proteins. TALEs are proteins that are naturally found in the plant pathogenic bacteria *Xanthomonas*. After injection of TALEs through the type III secretion system, the TALEs bind to host DNA loci and alter the transcription patterns (Boch & Bonas 2010). The first discovery of a TALE was by (Bonas et al. 1989). The unique feature of this protein family is that they can recognise and bind to DNA. The code of this specificity has been described by (Schornack et al. 2006).

It has been described that the AvrBS3 TALE protein (one of the most well studied) consists of a nuclear localisation signal and a region of 17.5 repeats in which each repeat has 34 amino acids. All repeats are identical to each other except for the 12th and 13th amino acids which are called hypervariable region (HVR) (Schornack et al. 2006). The AvrBS3 protein interacts with genes containing the UPA box (Römer et al. 2007; Kay & Bonas 2009). The authors determined that each repeat can recognise one base pair in the target sequence (Boch et al. 2009). This characteristic can be used for generating user specific recognition sites. On the same issue another paper by (Moscou & Bogdanove 2009) presented their results supporting the same hypothesis.

Beside directing the TALE protein to specific regions of interest in the genome the user has also the ability to fuse the TALE protein to different modification domains (activator, repressor etc). Different groups have used these

features in order to repress (Cong et al. 2012) or activate gene expression (Zhang et al. 2011), to drive histone modifications (Mendenhall et al. 2013), DNA demethylation (Maeder et al. 2013) or recombination (Mercer et al. 2012) to specific regions. It has also extensively been used for genome editing following the same principle as the ZFN. Two TALE proteins fused with one FokI cleavage domain each targeting neighbouring regions of the genome, this combination is termed TALE nucleases (TALENs) (Figure 3b) (Christian et al. 2010; Miller et al. 2011; Mahfouz et al. 2011).

Since then TALEN technology has been used in zebrafish (*D. rerio*) (Gupta et al. 2013; Huang et al. 2016), in fruit fly (*D. melanogaster*) (Liu et al. 2012; Katsuyama et al. 2013), in mouse (*M. musculus*) (Wefers et al. 2013; Jones & Meisler 2014), in thale cress (*A. thaliana*) (Christian et al. 2013; Forner et al. 2015) and in human cell lines (Piganeau et al. 2013; Ochiai et al. 2014).

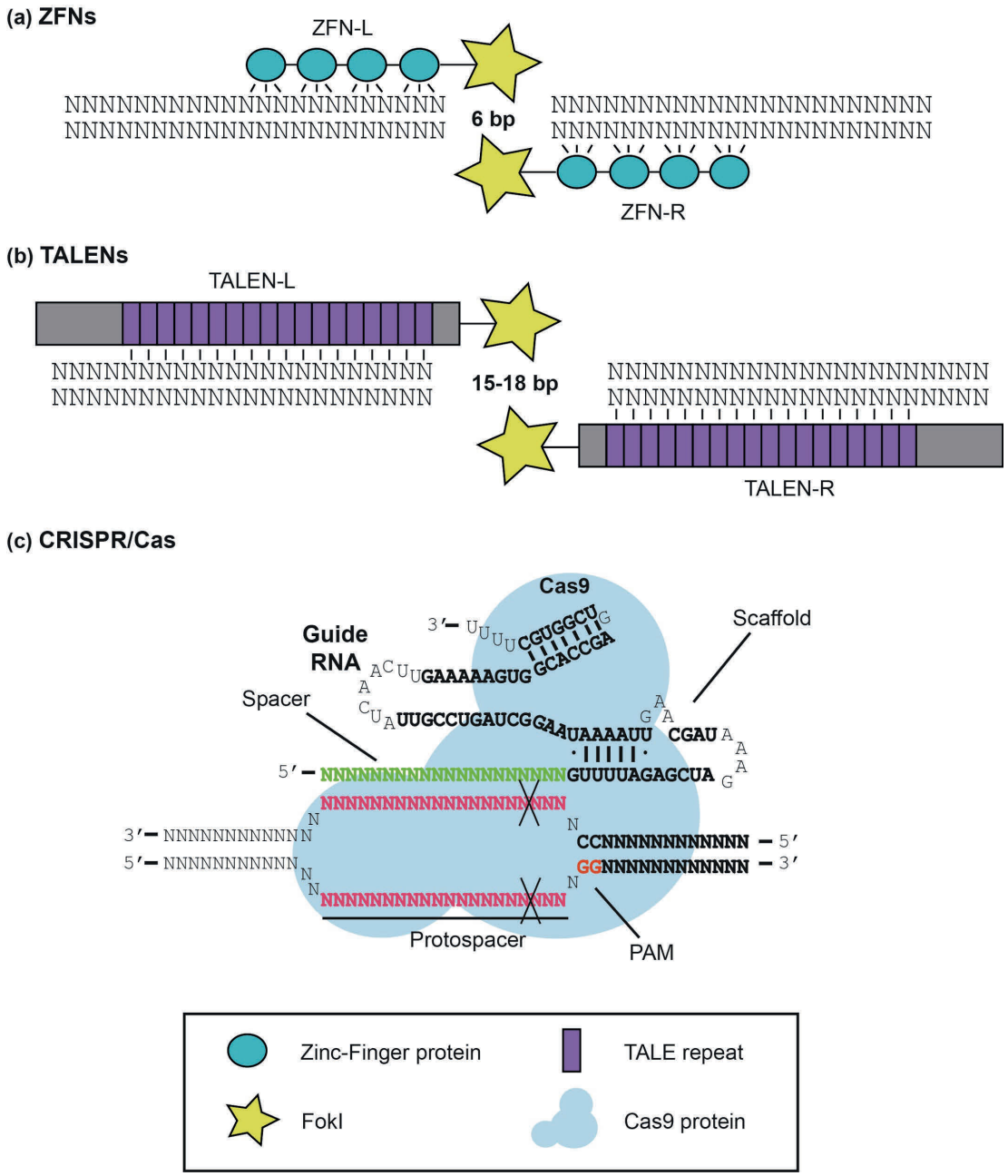


Figure 3: Graphical representation of three genome editing tools, ZFNs, TALENs and CRISPR/Cas. (a) A zinc-finger nuclease dimer with each recognising 12 nucleotides and carrying one FokI cleavage domain (yellow star). An individual zinc-finger (blue oval) recognises three nucleotides. When both ZFNs bind to the recognised regions, dimerisation of FokI domains takes place and a double strand break can occur. The distance between two targeted regions has to be 6 basepairs. (b) A TALE nuclease dimer with each recognising 18 nucleotides and carrying one FokI cleavage domain. Each TALE protein carries 17.5 repeats and each repeat recognises one nucleotide. The principal of TALENs is the same as ZFNs with the difference that the distance between the two targeted regions has to be 15 to 18 basepairs. (c) The structure of the active guide RNA-Cas9 complex. The sgRNA consists of the spacer that is complementary to the strand on which the protospacer (target site) is located and a scaffold that creates an RNA structure that is vital for the stabilisation of the complex. Cas9 is recruited to the target site, recognises the PAM site and is

stabilised in order to create a double strand break three nucleotides upstream of the PAM site (it is indicated with the X) (adapted from (Mali et al. 2013).

Cluster Regularly Short Palindromic Repeat/Cas (CRISPR/Cas)

The newest genome editing tool that has quickly come to dominate the field of genome editing over the last decade is based on a prokaryotic adaptive immune mechanism called Clustered Regularly Short Palindromic Repeat (CRISPR) type.

The story of CRISPR starts in 1987 where (Ishino et al. 1987) discovered a region in the *E. coli* genome that consists of tandem repeats. In 1989 (Nakata et al. 1989)described this region but did not yet know the function of the repeats. Four years later, (Mojica et al. 1993) described a very similar region of tandem repeats found in the archaeon *Haloferax mediterranei*. The authors were also able to find a similar region in the closely related species *H. volcanii* and described both regions in 1995 (Mojica et al. 1995). More groups described similar regions in other archaea or bacterial genomes (Hermans et al. 1991; Flamand et al. 1992; Masepohl et al. 1996; Bult et al. 1996; Sensen et al. 1998; Nelson et al. 1999). In 2002 after agreement between Mojica and Jansen the region was named CRISPR. The same year two new features were described, first the non-repetitive regions of similar size (spacers) that interspaced the repeats and second genes that show homology among different genomes that are associated with the CRISPR regions, that take the name of CRISPR associated (Cas) genes (Jansen et al. 2002). In 2005 three different papers described findings that the CRISPR region and the associated genes are playing a role in resistance against viruses and other extrachromosomal element invasions through complementarity of the spacer region with the invading DNA (Mojica et al. 2005; Pourcel et al. 2005; Bolotin et al. 2005). (Bolotin et al. 2005) also noticed that the foreign regions that correspond to the spacers of the CRISPR locus of *S. thermophilus* are lying next to a short signature of nucleotides (protospacer adjacent motif PAM). Others also made the same observation and the discovery was later proven be a very important feature of the system (Deveau et al. 2008; Horvath et al. 2008; Mojica et al. 2009; Lillestøl et al. 2009; Semenova et al. 2009).

The first proposed mechanism of the CRISPR system as defense against exogenous DNA was based on the predicted functionality of the different domains of

the Cas proteins and it led to its comparison to the siRNA mechanism that takes place in eukaryotes (Makarova et al. 2006). This hypothesis was proven to be incorrect in the next years. When bacterial cultures of *S. thermophilus* were incubated with viruses, new resistant strains survived from the culture, suggesting that the CRISPR system confers acquired resistance against viruses (Barrangou et al. 2007). This hypothesis was proven correct and the resistance was shown to be achieved through addition of new spacers in the CRISPR locus (Barrangou et al. 2007). Over the years following the 2007 publication, the function of different CRISPR associated (Cas) proteins was described along with the possibility to transfer CRISPR systems to distant species (Deveau et al. 2008; Brouns et al. 2008; Sapranauskas et al. 2011). It was further discovered that Cas9 is the only protein needed to provide resistance in the *S. thermophilus* system (Deveau et al. 2008; Brouns et al. 2008; Sapranauskas et al. 2011).

In 2008 (Marraffini & Sontheimer 2008) were the first to prove that the interference occurs through DNA targeting after they introduced a self-splicing intron in a gene. They also showed that the system can be programmable by introducing desired spacer sequences and they also recognized the biotechnological value of the method. Only in 2011 the final component of the system was identified. Another RNA molecule (tracrRNA) that is important for the maturation of the CRISPR array (tracrRNA:crRNA) that is transcribed by RNAseIII polymerase was identified (Deltcheva et al. 2011). In 2012 an artificial fusion RNA (sgRNA) was introduced in order to replace the tracrRNA:crRNA hybrid and skip the maturation of the complex with promising results (Figure 3c) (Jinek et al. 2012). Using this sgRNA the Cas9 complex can be easily programmed to target any sequence that is followed by a PAM site (Figure 3c) (Jinek et al. 2012). They also showed that Cas9 has two active endonuclease domains which means that it can create DSBs (Figure 3c) (Jinek et al. 2012). A couple of months later another group was able to co-purify Cas9 and the rest of the complex (tracrRNA:crRNA) and use it *in vitro* for cleavage successfully; they also showed that 20 nt spacer is enough for efficient targeting (Gasiunas et al. 2012).

CRISPR types

In nature different types of CRISPR systems have been found and classified in two classes. In Class I there are the types that consist of multisubunit (multiple proteins) complex for recognition and cleavage of the target site (CRISPR type I,III and IV) where in class II there are the types that require single subunit (one protein) for activation (CRISPR types II and V) (Makarova et al. 2015). The first type I system was identified by (Brouns et al. 2008) in *E. coli* and this type is characterised by 8 Cas cluster genes produce a series of proteins that are involved as a cascade for the maturation of the pre-crRNA and the identification of the target site so the Cas3 protein will come and cut the dsDNA (Brouns et al. 2008; Pougach et al. 2010; Jore et al. 2011; Westra & Brouns 2012). This type is more prevalent in bacteria when compared to the type III system, which is more abundant in archaea (Makarova et al. 2006). The type III system also consists of multiple Cas proteins but what makes it different is the presence of Cas10 (endonuclease activity) and the fact that it can also target RNA (Makarova et al. 2011). The first observations of type III were made by (Marraffini & Sontheimer 2008) in *S. epidermidis* and by (Hale et al. 2009) in *P. furiosus*. Finally, a related class, the putative type IV, is usually encoded on plasmids and consists of a smaller subunit (4 Cas proteins, Csf1 endonuclease activity) (Makarova et al. 2015). The hypothesis was formulated that this type could serve as a mobile unit (Makarova et al. 2015).

A member of the second class is the type II CRISPR which is the one that was first adapted as a genome editing tool (Jinek et al. 2012; Cong et al. 2013). It is the best studied type mainly due to its presence in the bacterium *S. thermophilus* which has economic value (lactic-acid bacterium) (Bolotin et al. 2005; Barrangou et al. 2007; Sapranauskas et al. 2011; Gasiunas et al. 2012). This type only needs the Cas9 protein for cleavage of the targeted region (Jinek et al. 2012; Gasiunas et al. 2012). The same CRISPR type is found in the *S. pyogenes* which was used for the identification of the tracrRNA and also its Cas9 sequence is the most commonly used nowadays (Deltcheva et al. 2011; Jinek et al. 2012; Cong et al. 2013; Mali et al. 2013). Finally, the last type in this class is the type V CRISPR. This system was described recently by (Zetsche et al. 2015) in studies on *F. novicida*. It also utilises a single protein Cpf1 for cleavage of the targeted region but in contrast to the other types, the maturation process occur without the use of a tracrRNA (Zetsche et al.

2015). Another difference is that the Cpf1 protein creates staggered DNA breaks (Zetsche et al. 2015).

CRISPR/Cas genome editing tool

The type of CRISPR that has been adapted as a genome editing tool is the type II. In this system only Cas9 is needed for cleavage (Sapranauskas et al. 2011) and 20nt of complementarity between DNA and RNA is enough for efficient targeting (Gasiunas et al. 2012). The main advancement arrived in 2012 when (Jinek et al. 2012) created an artificial small RNA (sgRNA), replacing the mature tracrRNA:crRNA of the native system. Also when using the *S. pyogenes* Cas9 (SpCas9), only a very simple PAM site (5'-NGG-3') is required to be located downstream of the targeted site (Jinek et al. 2012). In 2013 (Cong et al. 2013) described the first time that the CRISPR/Cas system was used to edit mammalian cell lines. After adding a nuclear localisation signal in SpCas9's sequence the protein was used for targeting and creating DSBs in *EMX1* gene (Cong et al. 2013). An sgRNA was designed and evaluated for efficiency. Their results showed that the system was able to generate mutations in the targeted positions and create deletions when two regions are targeted (Cong et al. 2013).

Applications and advancements of CRISPR/Cas

Since the first application of the CRISPR/Cas system in 2013 for editing mammalian cells, the number of its applications has increased dramatically. It has already been used for editing zebrafish (*D. rerio*) (Hwang et al. 2013; Auer et al. 2014; Irion et al. 2014), mouse (*M. musculus*) (Wang et al. 2013; Hirose et al. 2017), rat (*R. norvegicus*) (D. Li et al. 2013; Chapman et al. 2015), fruit fly (*D. melanogaster*) (Gratz et al. 2013; Ren et al. 2013), monkey (*Macaca fascicularis*) (Niu et al. 2014), pig (*S. scrofa domestica*) (Yan et al. 2018), nematode (*C. elegans*) (Friedland et al. 2013) rice (*O. sativa*) (Endo et al. 2014; Lowder et al. 2015), tomato (*S. lycopersicum*) (Nekrasov et al. 2017), thale cress (*A. thaliana*) (J.-F. Li et al. 2013; Hyun et al. 2015; Peterson et al. 2016), and mammalian cells (Canver et al. 2014; Bauer et al. 2015; Bressan et al. 2017).

Creating a deactivated Cas9 (dCas9) in which both endonuclease domains are mutated one can recruit activators or repressors to the targeted region without

directly altering the genome, and only affecting gene expression (Qi et al. 2013; Gilbert et al. 2013; Larson et al. 2013). (Hilton et al. 2015) described a dCas9 (both *Sp* dCas9 and *Nisseria meningitidis* dCas9) fused with the human p300 Core, that has acetyltransferase activity, and was able to regulate the gene expression when the corresponding promoters or enhancers were targeted. On the other hand, (Kearns et al. 2015) investigated a *Nm* dCas9 fused with the histone demethylase LSD1 targeting enhancers of genes in mouse embryonic stem cells. The authors showed that successful suppression of gene expression took place through reducing the H3K4me2 and H3K27ac markers (Kearns et al. 2015). Finally, a study described that the dCas9-KRAB fusion can induce H3K9me3 and by targeting HS2 enhancers suppresses all genes that are regulated by HS2 (Thakore et al. 2015). A new approach arose based on the dCas9 properties but in this method the sgRNA scaffold is fused with RNA aptamers (e.g MS2,com). These aptamers can recruit RNA binding proteins (e.g MCP, Com) that are fused to activators (e.g VP64) or repressors (e.g KRAB), and gene activation or repression can occur in the same cell for different genes (Zalatan et al. 2015). dCas9 has been proven to be very useful for labeling and live imaging in cell lines. It has been used for visualising repetitive regions like telomeres or genes that are targeted simultaneously with multiple sgRNAs by fusing flurescent protins to dCas9 or the sgRNA scaffold (B. Chen et al. 2013; Ma et al. 2015; Ye et al. 2017; Qin et al. 2017). Increasing the recognition specificity of Cas9 have also been achieved by mutating different domains without affecting the cleavage efficiency (Slaymaker et al. 2016; Kleinstiver et al. 2016; Chen et al. 2017).

Comparison of genome editing tools

Within the last five years that CRISPR/Cas system become applicable as a genome editing tool, its recorded usage has exceeded that of its predecessors (Figure 4). When it comes to ZFNs or TALENs two proteins are needed to target neighbouring regions in order for a DSB. Since both tools are based on protein-DNA recognition and binding, the site-specific targeting is more difficult to be manipulated because every new target region requires a new protein to be generated (Bibikova et al. 2001; Boch et al. 2009). When it comes to meganucleases the recognised region is pre-defined but attempts to improve the recognition site were carried out

(fusing different domains or by mutating various amino acids) (Ashworth et al. 2010; Grizot et al. 2010). A drawback of this method is that the recognition and endonuclease domains are encoded as single proteins and it is therefore difficult to alter them independently. CRISPR/Cas targeting is simpler in comparison. The system needs only one protein (Cas9) that can create a DSB and the targeting, that is based on RNA-DNA interaction, can be easily manipulated by changing the 20nt of the spacer region of the sgRNA (Figure 3c) (Jinek et al. 2012). The elaborated targeting process of the first three tools causes to be more costly than the CRISPR/Cas system. Another great advantage of the CRISPR/Cas system is the ability of the user to achieve multiplex targeting at the same time easier. The sgRNA, that directs the recognition, is a small molecule allowing this way the multiplexing by introducing multiple sgRNAs at once (Peterson et al. 2016; J.-F. Li et al. 2013; Tothova et al. 2017).

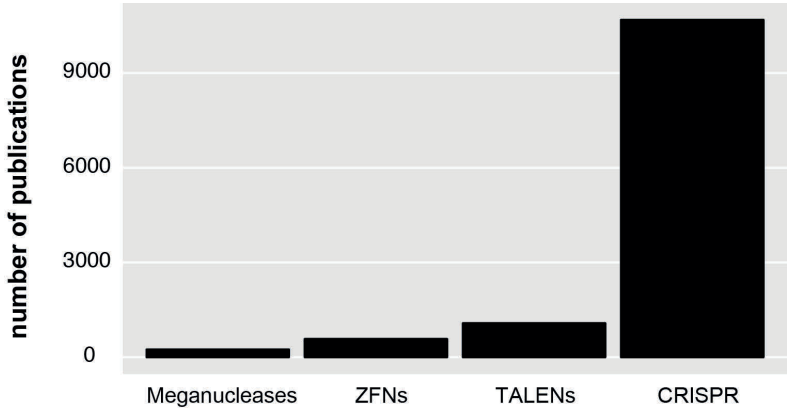


Figure 4: Number of publications for each genome editing tool within the last 30 years (1988-2018). Web of Science was used for searching publications where each one of the genome editing tools was mentioned. The data were acquired in August 2018. (http://apps.webofknowledge.com).

It is important to note, however, that the CRISPR/Cas system does have some limitations in its activity. One of them is the requirement of a PAM site for recognition and cleavage (Sapranauskas et al. 2011). The most commonly used system is based on the *S. pyogenes* Cas9 in which a 5'-NGG-3' acts as PAM (Jinek et al. 2012). This pattern can be found in abundance in a genome but still can be limiting in specific regions. As mentioned previously, the meganucleases also have targeting limitations based on the domains, while the ZFNs and TALENs can target any region (Desjarlais & Berg 1992; Schornack et al. 2006). Finally, another limitation of the system is the potential generation of off-target events. It has been

shown that Cas9 could tolerate some mismatches between sgRNA and target site (Sapranauskas et al. 2011; Jinek et al. 2012). Details on the rates of off-target mutations are not settled. Indeed, there are studies that support the presence of off-target but also studies that describe no off-target activity (Hsu et al. 2013; Pattanayak et al. 2013; Veres et al. 2014; Peterson et al. 2016). I investigated the presence of off-target events regarding my approach and this is discussed in Chapter 3. The off-target problem was also prominent with the other tools (ZFNs and TALENs) but similarly improved versions of the nucleases have been generated with increased specificity and decreased toxicity (Miller et al. 2007; Doyon et al. 2011).

Aims

Several studies investigate the involvement of SA in different aspects of plant such as those analyzing the effect of SA on flowering time, its effect under stress conditions and during immune responses. The majority of these studies have been conducted using the common Col-0 accession. The nearly exclusive focus on Col-0 (or a few other canonical accessions) has led to a gap in our knowledge with respect to gene interactions and networks that could be specific to other natural accessions. It has already been shown that there is substantial genetic variation in response to external application of SA in natural accessions of *A. thaliana* (van Leeuwen et al. 2007). Additionally, the *ACCELERATED CELL DEATH 6* (*ACD6*) gene that acts in a positive feedback loop with SA and promotes cell death, has been shown to be active in the natural accession Est-1 but not in others (like Col-0) indicating variation regarding *ACD6* and SA regulation (Todesco et al. 2010). Using a genome editing approach one can create null alleles in different genetic backgrounds and investigate their performance. The CRISPR/Cas technology has been utilised extensively over the last five years. Nonetheless, there continue to be aspects of the technology that are poorly understood, or for which there are contradictory results and hypotheses. One example is the controversy surrounding the rate of off-target cleavage due to non specific Cas9 activity.

By knocking out *ICS1* in seven natural accessions of *A. thaliana* using the CRISPR/Cas system I was able to eliminate the main biosynthetic pathway of SA. I then proceeded to:

1. Generate an economic and high-throughput screening pipeline (CRISPR-con), starting from DNA extraction to identification of edited individual using amplicon sequencing;
2. Investigate SA involvement in flowering time regulation, the SA levels in absence of the *ICS1* gene during *P. syringae* infection and the effect of SA on *H. arabidopsidis* pathogenicity in natural accessions of *A. thaliana*;
3. Identify off-target cleavage sites when the *ICS1* gene is targeted with two sgRNAs at the 3rd exon using CRISPR/Cas system.

Chapter 1

CRISPR-finder: A high throughput, cost efficient and precise method for identification of successfully edited individuals in *A. thaliana*

Introduction

"Genome editing", or making precise changes to an organism's genome, has become a routine approach to investigate gene function *in vivo*. While there are other systems available for editing genomes, such as Zinc Finger nucleases (Bibikova et al. 2002; Lloyd et al. 2005; Porteus & Baltimore 2003) and Transcription activator-like effector nucleases (TALENs) (Christian et al. 2010; Miller et al. 2011; Forner et al. 2015), both of these systems are not as straightforward to use because targeting and recognition is carried out by a fine-tuned protein/DNA interaction (Pavletich & Pabo 1991; Schornack et al. 2006). The difficulty of designing the required proteins leads to lower success rates (Gaj et al. 2013). Over the last few years, a new system was discovered and described called Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). In nature, this system is found in bacteria and archaea and acts as an adaptive immune mechanism which recognizes and cleaves specific phage DNA sequences. In the lab, it has now been successfully adapted as a genome engineering tool which can recognize and cleave user-defined sequences, thereby editing the genome (Jinek et al. 2012; Cong et al. 2013).

The discovery of the CRISPR/CAS system opened new doors for genome editing by simplifying the requirements for genome targeting, particularly in comparison to Zinc finger nucleases and TALENs (Gaj et al. 2013). The system requires a nuclease (Cas9), an artificial single guide RNA (sgRNA), and a short sequence upstream of the sgRNA binding site called a Protospacer Adjacent Motif (PAM) site, which has the sequence 5'-NGG-3' (Jinek et al. 2012; Gasiunas et al. 2012). Part of the sgRNA is complementary to 20 nucleotides in the targeted region of the genome, and the rest is responsible for the stabilisation of the Cas9/sgRNA complex.

Interaction of the Cas9/sgRNA complex with the target site enables Cas9's endonuclease domain to generate a double stranded break (DSB). Such breaks can be repaired through either the non-homologous end joining (NHEJ) or homology directed repair (HDR) pathways. NHEJ is error-prone, and can introduce small insertions or deletions that can lead to the disruption of the reading frame (Phillips & Morgan 1994; Ma et al. 2004). In the case of HDR, a donor template complementary to the target needs to be present to introduce a specific region to the genome of interest (Liang et al. 1998; Gratz et al. 2014). The CRISPR/CAS system has been successfully used for generating knock out lines (D. Li et al. 2013; Chang et al. 2013), knock in lines (Auer et al. 2014; Platt et al. 2014) and for deleting loci (Canver et al. 2014) in several organisms including *Arabidopsis thaliana* (Feng et al. 2013; Feng et al. 2014; Hyun et al. 2015; Peterson et al. 2016).

Although this system is now routinely used, identifying successful mutational events is often challenging and tedious, since the number of transformants and mutants that must be screened is usually large. For example, when applied to *Arabidopsis thaliana,* one might have to screen hundreds of individuals in order to find the desired event. The most common approaches for screening for mutations are either to use Sanger sequencing (Feng et al. 2014; Fauser et al. 2014) or to use the T7 Endonuclease 1 (T7E1) assay (Xie & Yang 2013; Ablain et al. 2015) applied to PCR products. Neither of these methods provides immediately a precise identification of mutations in the desired region. For example, in the case of screening using Sanger sequencing, the final readout merges the most abundant products in the template into one chromatogram (Strauss et al. 1986; Sanger & Coulson 1975). This can lead to secondary peaks and sometimes a mixed signal due to other amplified molecules in the mixture, and can make it very hard to detect desired but rare events that might have occurred during editing. Confirmation of successful editing through subsequent cloning of such a mixed PCR product followed by retrieval of bacterial colonies that carry the rare variant is additionally time-consuming and expensive. Use of T7E1 for screening can also yield inconclusive results. This can happen because the assay is based on identifying heterozygous individuals after denaturing/annealing steps, in which the T7 Endonuclease 1 digests the mismatched fragments (Mashal et al. 1995). This approach lacks the ability to identify homozygous individuals, as there are no

mismatched fragments available for digestion. Finally, both techniques can be expensive if one wants to screen a large number of samples (>100).

These imperfections concerning the routinely-used approaches for screening CRISPR'ed individuals led me to develop a more robust way of efficiently screening large numbers of samples. Here I introduce a high throughput screening approach for identifying mutations using Illumina sequencing, called CRISPR-finder. I describe both the library preparation of the samples and the analysis pipeline for identifying editing events. My approach is based on amplicon sequencing using MiSeq as a platform of choice, but my design yields amplicons that are compatible with the HiSeq3000 platform as well (by altering the adapter sequence one could use the pipeline for other platforms too). High throughput amplicon sequencing has been successfully used in the past for various purposes like for the description of marine microbial communities using the 18S rDNA (Huber et al. 2007), the characterization of the human gut microbiome using 16S rDNA (Wu et al. 2010) and the identification of the bacterial community in soil (Nacke et al. 2011). Amplicon sequencing has also been used in the case of cancer research; (Takeda et al. 2015) used the Ion Torrent^{TM} platform for amplicon sequencing panels of 22 genes that are linked to lung and colon cancer using formalin-fixed, paraffin-embedded (FFPE) samples, and the MiSeq platform was used for sequencing 212 amplicons in 48 gene related to colorectal cancer (Betge et al. 2015). Both studies successfully investigated the presence of variants at the genes of interest using hard-to-process source material.

Aims

Using the CRISPR/Cas genome editing tool has allowed the alteration of genomes almost effortlessly. The extensive use leads to the generation of large numbers of individuals to be screened for identifying the desired mutations. The aim of my study was to overcome this issue. The great number of samples in my work (>900) led me to adapt a previously described preparation method by (Lundberg et al. 2013). This method gave me the opportunity to generate a pipeline (CRISPR-finder) for successfully identifying individuals carrying mutations at the targeted regions (Figure 1.1).

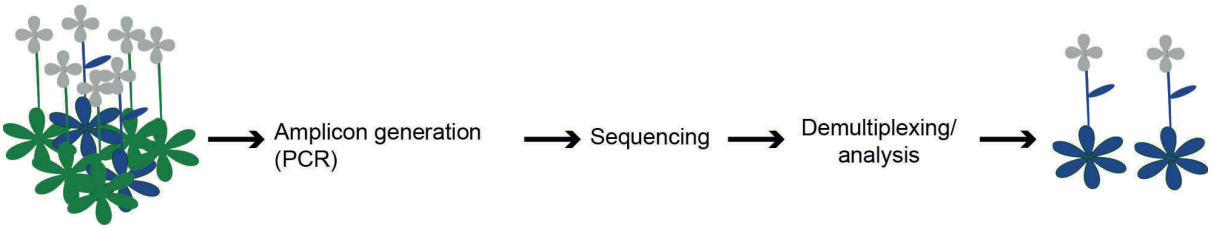


Figure 1.1: Schematic representation of the screening pipeline, CRISPR-con. Starting from >900 samples, the amplicon generation takes place by preparing the individuals for sequencing by DNA extraction and library generation. Following the sequencing run, the demultiplexing and analysis lead to the identification of desired edited individuals. Green represents unedited plants and blue represents plants that carry editing events.

Results

Motivation of work

My aim was to improve the CRISPR/Cas system application and efficiency. Specifically, I sought to simplify the screening of individuals to identify those that have been edited. To demonstrate the efficacy of my new approach, I targeted the *FLOWERING LOCUS C* (*FLC*) or *ISOCHORISMATE SYNTHASE 1* (*ICS1*) genes in different *A. thaliana* accessions (Table 1.1) using the CRISPR/Cas system, and then screened these accessions for successful events.

Table 1.1: List of all the accessions used throughout optimisation and verification of the amplicon sequencing for different runs.

Agu-1	ICE107	ICE153	ICE228	ICE75	Lerik1-3	Tü-Scha-9
Bak-2	ICE111	ICE163	ICE29	ICE79	Nermut-1	Tü-V-12
Bak-7	ICE112	ICE169	ICE33	ICE91	Nie1-2	TüWa1-2
Col-0	ICE119	ICE173	ICE36	ICE92	Pre-6	Vash-1
Don-0	ICE120	ICE181	ICE60	ICE93	Qui-0	Wal-HäsB-4
Ey1.5-2	ICE127	ICE21	ICE61	ICE97	Rü3.1-27	Xan-1
ICE1	ICE130	ICE212	ICE7	ICE98	Sha	Yeg-1
ICE102	ICE134	ICE213	ICE70	Kastel-1	Star-8	
ICE106	ICE152	ICE216	ICE73	Koch-1	TüSB30-2	

FLC encodes a flowering repressor with rich natural allelic variation in *A. thaliana* (Michaels & Amasino 1999; Li et al. 2014). *ICS1* encodes an enzyme involved in salicylic acid biosynthesis (Wildermuth et al. 2001).

The accessions of *A. thaliana* used in this study are from the first phase of the 1001 Genomes Project (Cao et al. 2011). I used the polymorph tool

(http://polymorph.weigelworld.org) to align sequences of *FLC* and *ICS1* from the different accessions. I identified target sites without sequence variation among the accessions to develop the guide RNAs (Figure 1.2).

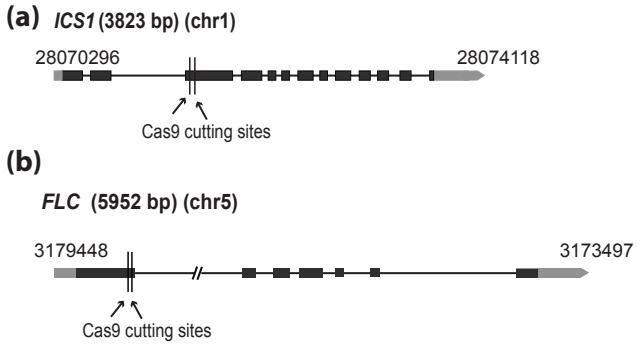


Figure 1.2: Diagrams of the two targeted genes, *ISOCHORISMATE SYNTHASE 1* (*ICS1*) and *FLOWERING LOCUS C* (*FLC*). Black boxes indicate coding regions in exons, and grey boxes untranslated regions. The arrow shows the directionality of the transcription.

Plants were transformed separately with either the *FLC* or *ICS1* targeting constructs (Appendix Table 1) and the primary transformants were found to have somatic editing events. Two versions of Cas9 were used, which were either plant-codon-optimised (pcoCas9) (J.-F. Li et al. 2013) or Arabidopsis-codon-optimized (AthCas9) (Fauser et al. 2014). The selection of the transgene was based on glufosinate or the seed specific expression of mCherry (Kroj et al. 2003; Gao et al. 2016). Details concerning the cloning system and the different components are in Chapter 2.

Generation and sequencing of amplicons spanning CRISPR target sites

Sanger sequencing and the T7E1 assay are time consuming, imprecise, and expensive (Feng et al. 2014; Ablain et al. 2015). Moreover, the T7E1 must be combined with further screening like Sanger sequencing. To overcome these problems, I developed CRISPR-finder which is based on an amplicon sequencing approach (modified from (Lundberg et al. 2013)) and a novel analysis pipeline (Figure 1.1).

In order to quickly and unambiguously identify CRISPR/Cas9-induced mutations in a large number of plants, I amplified the target regions by PCR using

barcodes to identify amplicons from different individual plants, and then sequenced pools of barcoded PCR products on an Illumina MiSeq (or HiSeq) instrument. I modified the *FLC* and *ICS1* loci to determine the efficacy of my method in (i) different genetic backgrounds and (ii) at different genetic loci. The distance between the expected cutting sites for the two target sites was 32 bp for *FLC* and 72 bp for *ICS1,* and the insert sizes following PCR were 197 bp and 211 bp respectively.

The amplicons were prepared based on a two-step PCR amplification (Figure 1.3a-c). During the first amplification the specific region of interest was amplified, and the frameshifting nucleotides and part of the TruSeq adapters were added (Figure 1.3b) (Appendix Table 2). The cleaned PCR product was used as a template for the second amplification where the remainder of the TruSeq adapters and one of 96 barcodes were added (Lundberg et al. 2013) (Figure 1.3c) (Appendix Table 2). Each PCR amplification step was carried out for 15 cycles.

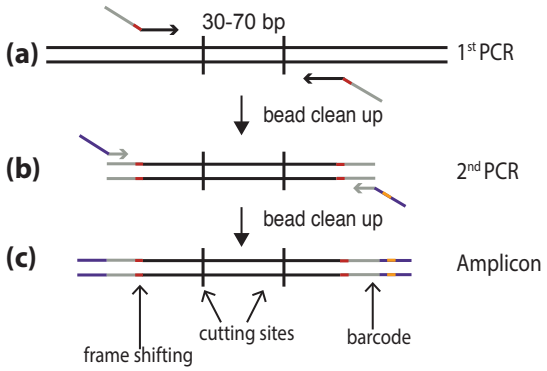


Figure 1.3: Amplicon preparation. (a) The 1st PCR amplifies a specific region. The oligonucleotide primers in this step include the first part of the TruSeq adapters (grey) and the frameshifting nucleotides (red). **(b)** The 2nd PCR amplification adds the last part of the TruSeq adapters (purple) and one of the 96 barcodes (orange). **(c)** The final amplicon with the frame shifting base pairs (red), the TruSeq adapters (grey and purple) and the barcode (orange).

The PCR products were quantified using the Quant-iT™ PicoGreen® dsDNA assay, normalized (described below), and pooled. For the sequencing on the MiSeq platform the MiSeq reagent kit v2 (300-cycles) (MS-102-2002) was used. I designed the adapters in order to be compatible both with MiSeq and HiSeq3000 platforms (Illumina, USA); on both I had successful runs.

Demultiplexing process

After sequencing, the pooled reads were demultiplexed in a two-step process. 96 batches of combined samples were first identified via the indices that were located at the TruSeq adapters incorporated in the 2nd PCR amplification. This process was carried out by bcl2fastq (1.8.4) software, provided by Illumina, which also trims the sequence of the barcodes (https://my.illumina.com) (Figure 1.4a).

Subsequently, sequencing reads from different samples were mixed under the same barcode. In order to assign each read to the individual from which it came, we took advantage of the frameshifting nucleotides incorporated during the first step of the two-step PCR amplification. The first nine nucleotides from each read were used as "secondary" barcodes to determine from which sample each read in the sequencing run originated; 9 bases are sufficient to capture the unique frameshifting nucleotides used during the amplicon generation (Figure 1.4b).

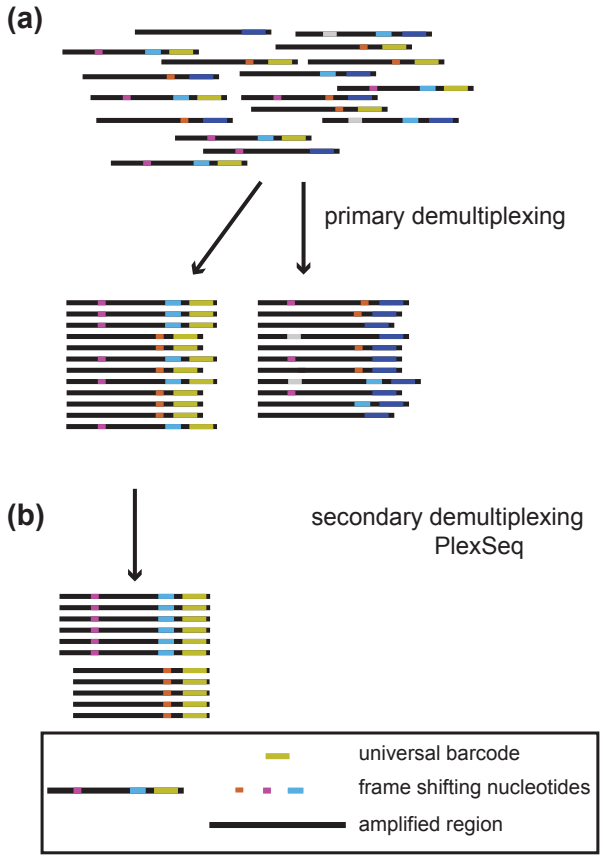


Figure 1.4: Diagrams of the demultiplexing procedure. (a) The primary demultiplexing step is carried out by the illumina software and separates the samples based on the indices that are located

within the adapter region into 96 pools. **(b)** The secondary demultiplexing script (PlexSeq) then assigns the reads to individuals based on the frame shifting nucleotides.

To achieve this, an in-house python script (PlexSeq) (available at https://gitlab.localnet/jregalado/plexSeq) was developed by fellow PhD student Julian Regalado that successfully demultiplexes >98% of reads in my dataset (Figure 1.5a-d). Since PlexSeq was run without allowing any mismatches of the "secondary" barcodes, around 2% of the data could not be separated because of errors in primers or errors introduced during the sequencing process, which makes it impossible to unambiguously determine the sample origin of these reads. However, a loss of 2% of data was deemed acceptable.

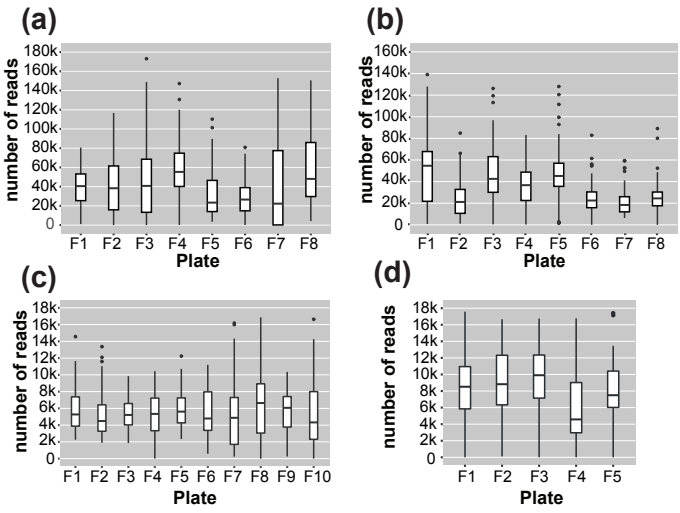


Figure 1.5: Graphs representing the average number of reads per plate (≤96 samples) per run. (a) MiSeq run 010, **(b)** MiSeq run 024, **(c)** MiSeq run 083 and **(d)** HiSeq 3000 run 058. For the HiSeq 3000 run, the library preparation was the same as for the MiSeq runs, but only 1% of a lane was used.

Unassigned reads are not used for downstream analysis. A file with the expected "secondary" barcodes needs to be provided in order for the script to successfully proceed with the demultiplexing (Figure 1.6).

#ID	Seq	R2
A1F1	CAGTCGAAT	ACTCACAGA
A1F2	TCAGTCGAA	CTCACAGAT
A1F3	CTCAGTCGA	TCACAGATG
A1F4	ACTCAGTCG	TGCAGATGA
A1F5	CAGTCGAAT	TCACAGATG
A1F6	TCAGTCGAA	ACTCACAGA
A1F7	CTCAGTCGA	ACTCACAGA
A1F8	ACTCAGTCG	CTCACAGAT
A1F9	CAGTCGAAT	CTCACAGAT
A1F10	CTCAGTCGA	TGCAGATGA

Figure 1.6: Screenshot of the text file that is needed for the PlexSeq to demultiplex the raw reads based on the frame shifting nucleotides. The file consists of three columns. The first column has the name that the file should take after demultiplexing. The second and third columns have the first 9 expected nucleotides for each read based on the oligonucleotides that were used for amplification during the library preparation.

Analysis pipeline

After the demultiplexing process, each sample was genotyped in order to detect single nucleotide polymorphisms (SNPs), as well as small insertions and deletions in the region of interest.

For each sample, reads were mapped back to the reference sequence for the genes of interest (Gene ID: 830878 and 843810) using the MEM algorithm of the BWA read mapping tool (Li & Durbin 2009) with standard parameters (Figure 1.7a). Afterwards, alignments were filtered for suitable quality with a mapping quality cutoff of 30 or higher (Li et al. 2009). The resulting alignment files were genotyped with freebayes using standard parameters (Figure 1.7a) (Garrison & Marth 2012a). The resulting VCF file was then filtered with vcftools (Danecek et al. 2011) to only keep samples in which high quality variants were detected at regions of interest.

At the end, IGV (Robinson et al. 2011; Thorvaldsdóttir et al. 2013) was used for visual inspection of read mapping and variant calls (Figure 1.7b,c). All software was used with standard parameters unless otherwise noted. The required memory for the analysis can be 5-20 Gb depending on the output of the run.

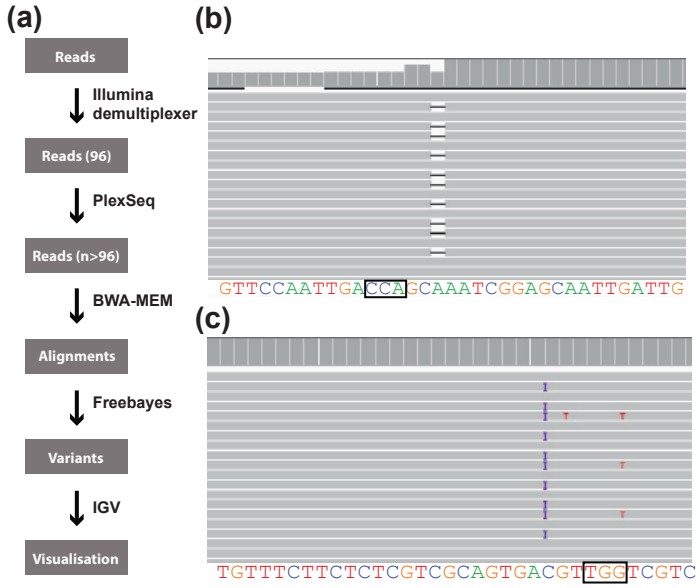


Figure 1.7: The analysis pipeline and visualised alignments. (a) Diagram of the analysis pipeline. BWA-MEM is used for the alignment and the Freebayes algorithm for variant calling. Finally, IGV is used for visualizing the alignment or the vcf files. **(b)** Alignment that shows a deletion visualised in IGV. On the top track one can notice the coverage panel and how it is decreased at the location of the deletion. The black open box indicates the location of the PAM site. **(c)** Alignment that shows an insertion (purple '**I**') visualised in IGV. The identified insertion was one base pair. The black open box indicates the location of the PAM site.

Identifying Mutations

Using CRISPR-finder, I was able to identify plants that were either hetero- or homozygous for the desired mutations. In order to evaluate the accuracy of the screening approach and also the impact of the mutation that was caused using the CRISPR/CAS system, I focused on an *ics1* mutant in the TüWa1-2 background (TüWa1-2 *c-ics1-1*) that I identified after screening more than one hundred individuals. The parental genotype was originally collected in Germany and its phenotype shows extensive necrotic lesions on the leaves (Figure 1.8c) which can be attributed to local cell death. I hypothesized that this is caused by elevated levels of SA. Using a biosensor assay I was able to successfully quantify the SA content in the plants (Huang et al. 2006; Defraia et al. 2008). As expected, the levels of free SA in the *ics1* mutant were significantly lower than the wild type accession (Figure 1.8a,b). These results demonstrate firstly, that using our approach of screening one can easily and precisely identify individuals that were edited through the

CRISPR/CAS system. Secondly, our system of targeting can efficiently generate knockout lines.

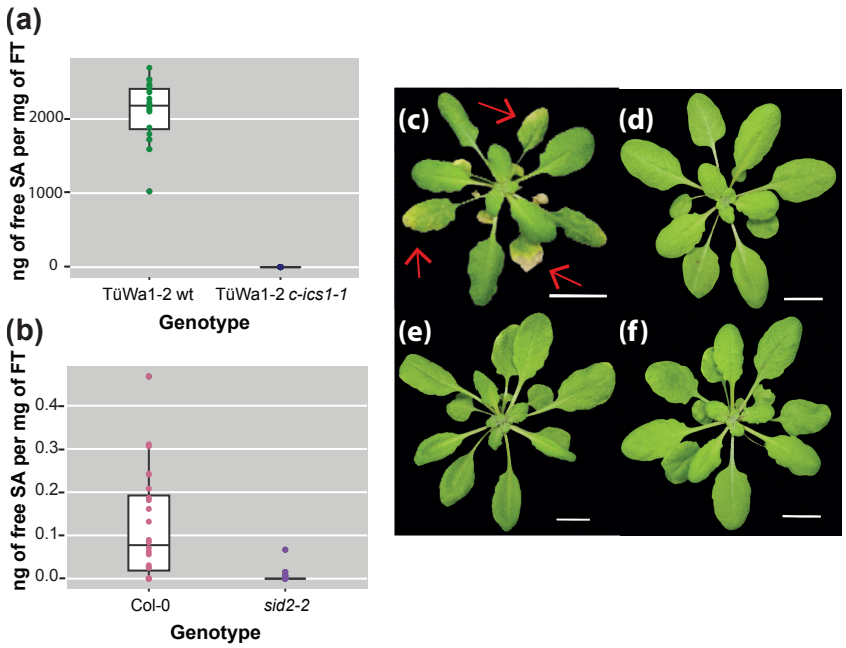


Figure 1.8: Salicylic acid levels and morphology of *ics1/sid2* mutants. (a) SA content of TüWa1-2 wild-type and the derivative TüWa1-2 *c-ics1-1* mutant. **(b)** SA content of Col-0 reference wild type and derivative *sid2-2* mutant for comparison. (*SID2* is a synonym for *ICS1)*. Note the very different scales of the two graphs. The measurements were obtained using plants that were growing in 23°C short-day conditions (8 h light/ 16 h dark) for 43 days. **(c)-(f)** Pictures representing the morphology of the wild types and the corresponding *ics1/sid2* mutants. **(c)** TüWa1-2 wild type, **(d)** TüWa1-2 *c-ics1-1*, **(e)** Col-0 reference, and **(f)** *sid2-2*. The plants were growing in 23 °C short-day conditions and the pictures were taken 40 days after sowing. The chlorotic and necrotic regions on the leaves are indicated by red arrows. The white line indicates 1 cm, note the different size of the scale bar in each image.

Discussion

The CRISPR/Cas9 system is a widely used approach for genome editing, with increasingly diverse and broad applications (Belhaj et al. 2013; Auer et al. 2014; Canver et al. 2014; Hyun et al. 2015). Many of these applications require genotyping of large numbers of individuals for the identification of induced mutations. The available screening approaches lack efficiency, precision and remain expensive. Here I describe a high throughput screening approach, called CRISPR-finder that increases the accuracy and reduces the time and money required for screening a great number of individuals.

I implemented an amplicon-based high throughput screening approach for identifying individuals with desired CRISPR generated events (Figure 1.1). Amplicon sequencing has been used in the past in various studies for addressing questions such as bacterial communities in different environments (Huber et al. 2007; Wu et al. 2010; Nacke et al. 2011) and variant identification in cancer related genes (Takeda et al. 2015; Betge et al. 2015). I exploited this method to use the full genomic DNA from each plant to generate amplicons of the targeted region through a two-step PCR amplification (Figure 1.3). Based on the precision of amplicon sequencing, 1000 reads per individual can generate enough coverage (1000x) where every molecule in the mixture will be sequenced and even the low frequency variants will be detected. Based on this coverage per individual and the expected output of a MiSeq run (~15 to 20 million reads) one could ambitiously attempt to multiplex up to 20000 samples in a single run. Of course the coverage can be adjusted to the needs of different experimental set ups. I also describe the demultiplexing approach and an analysis pipeline (Figure 1.4 and 1.7). For each individual a unique combination of frameshifting nucleotides and index sequence is used; in this way the same region from many different individuals can be sequenced.

The major cost of CRISPR-finder is the amplicon sequencing kit which has the capacity for producing millions of sequences. With one reaction one can comfortably sequence thousands of samples, so the preparation cost per sample decreases as more samples are added to the pool. Additionally, “spiking in” to another flow cell to use only part of a flowcell’s capacity is possible and allows high

flexibility and the lowest costs. In contrast, the T7E1 assay and the Sanger sequencing costs remain the same, regardless the number of samples.

There have been attempts over the last few years to address the difficulties posed by the available screening methods. There are available packages for the downstream analysis of demultiplexed reads, or packages that demultiplex reads that originated from different regions of the genome (Boel et al. 2016; Pinello et al. 2016). There is also an available R package that one can use to summarize variants' features like their type, their location and their frequency (Lindsay et al. 2016). The input for this package can be either Sanger or NGS data, but a method to sequence numerous individuals is not described in the pipeline (Lindsay et al. 2016). The advantage of CRISPR-finder is that I introduce not only the preparation method for the amplicons but also methods to multiplex numerous amplicons from hundreds of samples with high resolution. Our method uses frameshifting nucleotides for higher sequence quality without the necessity of phiX and has higher multiplexing capability in comparison with others (Brocal et al. 2016).

Finally, the first oligonucleotide set does not need HPLC purification, keeping the cost of the multiplexing procedure low. This is because oligonucleotides are synthesized 3'->5', therefore any minor truncations or errors will be concentrated towards the ends of the amplicon during the first amplification. These ends serve as the binding sites for the oligonucleotides that are used at the second amplification, which will anneal despite minor errors and synthesize products with the correct adapter sequence.

I was also able to successfully sequence CRISPR-finder libraries using the HiSeq3000 platform. The adapters and the size of the amplicons were designed in a way that the libraries can be compatible with both the MiSeq and HiSeq platforms.

Finally, in order to assess CRISPR-finder as a screening approach, I identified and characterised an *ics1* mutant. The levels of SA in this mutant were significantly lower when compared to the wild-type (Figure 1.8). This finding supports our hypothesis and expectations since the *ICS1* gene is involved in the production of SA in *Arabidopsis thaliana (Wildermuth et al. 2001)*. While my method was developed for screening *Arabidopsis thaliana* edited individuals, it can be easily adopted for any organism that has been genome edited using the CRISPR/Cas9 system. A fact that one has to keep in mind is that recently has been shown that

large deletions can occur during CRISPR/Cas editing (Kosicki et al. 2018). These deletions would escape from our proposed pipeline.

In conclusion, during this study a full pipeline form DNA extraction to identification of individuals carrying mutations generated with CRISPR/Cas system is described-- CRISPR-finder. When compared to the more traditional methods (Sanger and T7E1 assay), amplicon sequencing is more robust and less expensive.

Chapter 2

Investigation of the role of salicylic acid (SA) in flowering time regulation and the immune response against the oomycete *Hyaloperonospora arabidopsidis*

Introduction

Different *Arabidopsis thaliana* genotypes exhibit a broad spectrum of phenotypic variation in several aspects of development, such as flowering time, and defensive mechanisms (Lempe et al. 2005; Todesco et al. 2010). One such phenotype that varies between individuals is the regulation of salicylic acid production. The hormone salicylic acid (SA) is a regulator of both growth and defense (Santner et al. 2009; Vlot et al. 2009). Several studies have demonstrated that natural accessions differing in their levels of SA production show an autoimmune phenotype, identifiable by necrotic lesions that are visible on the leaves and resemble local cell death (HR; hypersensitive response) (Todesco et al. 2010; Zhu et al. 2018). The mechanism linking this autoimmune phenotype to the production of salicylic acid is somewhat elucidated (Todesco et al. 2010; Zhu et al. 2018). Salicylic acid is a well-established component of the HR response and numerous genes have been identified that influence the production of SA to ultimately mount a defense response (Delaney et al. 1994). One case that has been shown to potentiate a positive defense response in an SA-dependent manner is the *ACCELERATED CELL DEATH 6* (*ACD6*) gene. Natural genetic variation in this gene results in clear phenotypic differences between accessions (Lu et al. 2003; Todesco et al. 2010). Finally, another interesting aspect of natural variation in the plant immune system is the hybrid incompatibility between natural accessions that has been observed (Chae et al. 2014). The severity in the phenotype varies from small lesions to extensive cell death or even lethality (Chae et al. 2014).

Salicylic acid is a phenolic compound involved in various plant processes like response to abiotic stresses, allelopathy, flowering time regulation, leaf senescence and disease resistance (Raskin 1992; Martínez et al. 2004a; Vlot et al. 2009; Kumar 2014). SA is synthesised under two different pathways in *A. thaliana*, the

PHENYLALANINE AMMONIA LYASE (PAL) and ISOCHORISMATE SYNTHASE (ICS) pathways, with the latter being the dominant one (Vlot et al. 2009).

It is still unclear how SA may influence flowering time. The connection between flowering time and SA has been reported since 1974 when phloem sap or products of aphid infestation were able to accelerate flowering in SD conditions of the long-day plant *L. gibba* (Cleland & Ajami 1974). SA was identified as the main compound responsible for the phenotype and external application showed similar results (Cleland & Ajami 1974). A study in *A. thaliana* has also shown that plants exposed to UV-C radiation, which acts as an abiotic stress, exhibit elevated levels of SA and flower earlier while the *NahG* mutant (decreased SA levels) is characterised by late flowering in SD conditions (Martínez et al. 2004b).

There is evidence in the literature that there is cross-talk between SAR and flowering time and between sumoylation (post-translation modification that involves the SMALL UBIQUITIN-LIKE MODIFIER (SUMO) peptides that are conjugated to proteins and alter their function (Miura et al. 2007)) and flowering time (Lee et al. 2007; Singh et al. 2013; Villajuana-Bonequi et al. 2014a). *SIZ1* encodes a plant SUMO E3 ligase (Miura et al. 2005) and its mutants show accumulation of SA and SAG, resistance to pathogens and constitutive SAR marker *PR-1* expression (Lee et al. 2007). Additionally, *SIZ1* has been reported to promote and stabilise *FLC* expression (Jin et al. 2008; Son et al. 2014), which encodes a MADS box transcription factor that represses flowering time (Michaels & Amasino 1999). This indicates an indirect interplay between flowering time and SAR. The *early in short days 4* (*esd4*) mutant defective in a gene encodes a SUMO protease exhibits early flowering phenotype and it might be due to the accumulation of SUMO conjugates (Villajuana-Bonequi et al. 2014a). Even though the mechanism is still unknown a suppressor was identified as a mutant allele at the *ICS1* locus (Villajuana-Bonequi et al. 2014a). Finally, it has been shown that *FLOWERING LOCUS D* (*FLD*) is involved in flowering time regulation and in SAR (He et al. 2003; Singh et al. 2013). *FLD* is a known negative regulator of *FLC* expression and thereby can promote flowering (He et al. 2003; Michaels & Amasino 1999). The causal mutation in the *reduced in systemic immunity 1* (*rsi1*) mutant was identified in the *FLD* locus (Singh et al. 2013). In addition to the delay these mutants exhibit in flowering time, they can produce SAR signals but are unable to establish the SAR response. This result

suggests that *FLD* is necessary for receiving the SAR signal (Singh et al. 2013). Finally, the *HOPW-1-INTERACTING3* (*WIN3*) gene regulates SAR and the *win3* mutant shows decreased levels of SA but flowers earlier than the wild type (Wang et al. 2011) while MYB30 overexpression lines with elevated levels of SA show no difference in flowering time (Liu et al. 2014). While the aforementioned studies establish a link between different processes (SAR and (a)biotic stress response) and flowering time through molecules involved in both, a direct link between SA and flowering time has not been reported.

SA is mainly known for its involvement during the defense responses of *A. thaliana* against pathogens (Kumar 2014). Production of SA is induced during PTI and ETI in response to biotrophic or hemibiotrophic pathogen infection and is vital for establishment of systemic acquired resistance (SAR) (Glazebrook 2005; Loake & Grant 2007; Dempsey et al. 2011). Increased levels of SA have been detected upon infection indicating resistance (Zheng et al. 2007; Tsuda et al. 2008). In support of this observation, plants that are unable to accumulate SA due to *NahG* expression (convert SA to catechol) show enhanced susceptibility during infections and inability to mount SAR (Gaffney et al. 1993; Delaney et al. 1994; Vernooij et al. 1994; Lawton et al. 1995). Additionally, external application of SA or its synthetic analogues, 2,6,-dichloroisonicotinic (INA) or benzo(1,2,3) thiadiazole-7-carbothioic acid S-methyl ester (BTH), can aid in the retainment or regain of resistance (Pallas et al. 1996; Görlach et al. 1996; Nawrath & Métraux 1999). A key player for HR induction is *EDS1* (*ENHANCED DISEASE SUSCEPTIBILITY 1*) which controls SA production that is necessary for defense against virulent pathogens (Vlot et al. 2008). Mutants with a non functional *EDS1* are highly susceptible to several pathogens and unable to accumulate SA (Parker et al. 1996; Falk et al. 1999).

A common approach for evaluating SA contribution and induction in *A. thaliana* is pathogen infection. *Pseudomonas syringae* and *Hyaloperonospora arabidopsidis* have been extensively used in the literature for investigating and identifying genes that are involved during immune response. *P. syringae* is a hemibiotrophic bacterium and during its infection common symptoms are chlorotic lesions and “water-soaked” patches that turn necrotic (Dong et al. 1991; Whalen et al. 1991; Dangl et al. 1992; Katagiri et al. 2002). It has been shown that during *P. syringae* infection SA accumulation occurs (Gupta et al. 2000; Zheng et al. 2007). *H.*

arabidopsidis is an obligate biotrophic oomycete that naturally infects *A. thaliana* (Holub 2008). SA contribution towards the observed phenotypes during *H. arabidopsidis* infection have been reported for various isolates of the pathogen (Delaney et al. 1994; Nawrath & Métraux 1999; McDowell et al. 2005). The phenotypes that have been reported after infection with *H. arabidopsis* are quantitative regarding the response against the pathogen (Holub 2008).

First, there is the flecking necrosis in which small, almost invisible to the naked eye, necrotic lesions occur at the location where the spores landed and attempted to enter the leaf (Holub 2008). The next two types are characterised by more extensive HR in the surrounding cells (Holub 2008). The fourth type of response exhibits a characteristic pattern of necrosis called trailing necrosis (TN) during which indicates that the plant responds with HR around the positions that spores landed but the elicitation is not rapid enough which leads to pathogen growth (Holub 2008). As I will show with my results, which of these phenotypes presents is an indication of the quantitative differences in the physiological response of the plant.

The response to either pathogen infection follows often the “gene-for-gene” model according to which resistance is achieved when the product of a plant R gene counteracts the effector encoded by an avirulence gene (Katagiri et al. 2002; Coates & Beynon 2010). By utilising these two pathogens I can examine the levels of SA accumulation during infection (*P. syringae*) and its effect on the virulence (*H. arabidopsidis*). I use the term virulence to describe the degree of pathogenicity which is the ability of a microbe to cause disease (Agrios, 1988).

Aims

While there are reports of natural variation affecting the regulation of SA (e.g. *ACD6*) little is known regarding differences in its biosynthesis in different accessions (Todesco et al. 2010; Zhu et al. 2018). While the available evidence points to ICS as being a central component of the main SA biosynthetic pathway in *A. thaliana*, it is still unclear how much the PAL pathway contributes in different accessions (Wildermuth et al. 2001). These results may be due to the fact that the majority of available SA deficient mutants are generated in the commonly used Col-0 accession that does not exhibit high SA accumulation when unchallenged.

By eliminating the expression of *ICS1*, which is the gene for the first enzyme involved in the conversion of isochorismate to SA in the ICS pathway, I aim to investigate the effect of SA in flowering time regulation and assess the performance of SA deficient mutant lines during infection using different pathogens in different accessions.

Results

Generation of the mutant lines

The aim was to investigate the extent to which and the mechanism by which SA contributes to the flowering time phenotype and to pathogen responses respectively in different natural accessions. For this reason, I generated targeted mutations at the *ICS1* locus in different natural accessions using the CRISPR/Cas system. I employed the CaMV35S or the *A. thaliana EGG CELL1.1* promoter (pEC1.1) to express a plant codon optimised Cas9 (pcoCas9) (J.-F. Li et al. 2013) and two sgRNAs driven by two *U6* promoters targeting the 3rd exon of *ICS1* locus (Figure 2.1). The *EC1.1* promoter specifies an egg-cell specific expression pattern (Sprunck et al. 2012). The egg cell is fertilised by the sperm and along with the fertilised central cell and the ovule integuments give rise to the seed's embryo, endosperm and seed coat (Maheshwari 1950; Yadegari & Drews 2004). I chose to include the promoter in order to increase the possibility of the generation of inherited targeted mutations.

The two regions targeted for genome editing were identical across the accessions that I used as described in Chapter 1 (Figure 2.1a). I selected the primary transformants using glufosinate resistance or fluorescent mCherry expression (Bouchez, D. (Institut National de la Recherche Agronomique, Versailles (France). Centre de Versailles, Biologie Cellulaire) et al. oct1993; Gao et al. 2016). The generation of the final binary constructs was based on the Greengate cloning system (Lampropoulos et al. 2013) and each was carrying the promoter (CaMV35S/p*EC1.1*), the pcoCas9, the terminator (trbcs), the two sgRNA cassettes and the selection marker (BASTA/mCherry) (Figure 2.1c,d) (Appendix Table 1). I examined the expression of pcoCas9 by transient expression of pEF008 in *N. benthamiana* plants prior to the transformation of the natural accessions (Figure 2.1b). Finally, I recovered Cas9-free homozygous *ics1* mutant lines in the T_3 generation using either the glufosinate brushing approach as described in the Material and Methods or the fluorescent mCherry as described by (Gao et al. 2016), based on uniform segregation in T_3 families.

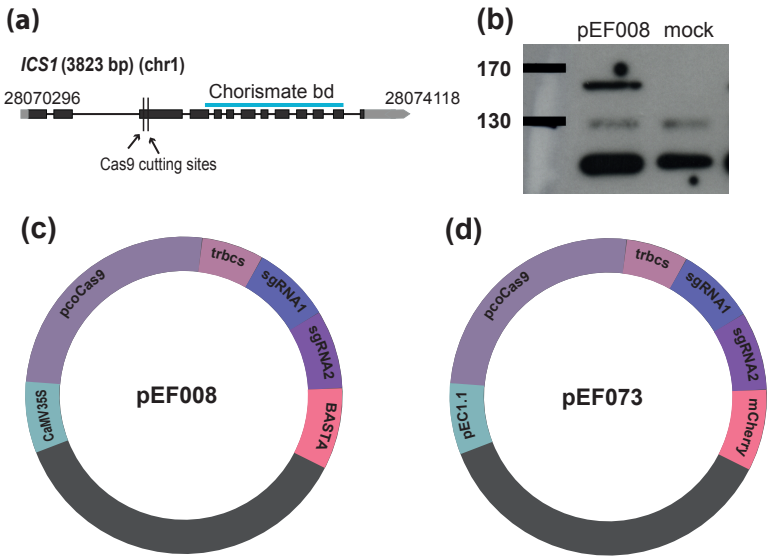


Figure 2.1:Graphical representation of the two binary constructs and expression of pcoCas9. (a) Diagram of *ICS1* gene. Black boxes indicate exons, and grey boxes untranslated regions. The arrow shows the directionality of the transcription. The blue line indicates the important for its activity Chorismate binding-domain (van Tegelen et al. 1999; Wildermuth et al. 2001). (b) pEF008, (c) pEF073 and (d) evidence for expression of pcoCas9 using protein blot. Anti-FLAG antibody (SIGMA-ALDRICH®, #F1804, 2012) was used for detection of Cas9 protein tagged with a FLAG epitope. *N. benthamiana* plants were grown in 23°C SD conditions for 5 weeks and were infiltrated either with the pEF008 construct or with infiltration medium only (mock).

I successfully generated *ics1* mutant lines using the pEF008 construct for the Yeg-1 and TüWa1-2 accessions and with pEF073 for the Col-0, Koch-1, Fei-0, Ey1.5-2 and ICE50 accessions. Not all the *ics1* lines were carrying the same mutations at the targeted region but for all of them the generated mutations were giving rise to a premature stop codon within the 3rd exon (Table 2.1) (Appendix Figure 1). Due to the generated mutations the chorismate binding-domain, which is essential for ICS1 function, is eliminated (van Tegelen et al. 1999; Wildermuth et al. 2001).

Table 2.1: List of all *ics1* mutant lines that were generated with the corresponding mutations and the effect in the final gene product. The TüWa1-2 *c-ics1-1* and TüWa1-2 *c-ics1-2* are two independent lines of the same accession carrying different mutations at the targeted region.

***ics1* line**	**Reference**	**Alteration**	**Effect**
Col-0 *c-ics1*	GAC	GAAC	Premature stop codon
Koch-1 *c-ics1*	CAA	CA	Premature stop codon
ICE50 *c-ics1*	CAGTGAC	CT	Premature stop codon
Ey1.5-2 *c-ics1*	GAC	GATC	Premature stop codon
Fei-0 *c-ics1*	GAC	GAAC	Premature stop codon
Yeg-1 *c-ics1*	GAC	GATC	Premature stop codon
TüWa1-2 *c-ics1-1*	TCGCTGTTTCTTCTC TCGTCGCAGTGAC	TC	Premature stop codon
TüWa1-2 *c-ics1-2*	GAC	GATC	Premature stop codon

Flowering time and measurement of SA of unchallenged mutant lines

In this study, I directly tested the effect of SA production on flowering time regulation by generating *ics1* mutant lines (SA deficient) in different genetic backgrounds. All plants used in this study (for measuring flowering time phenotype and for pathogen assays) were grown in 23^{o}C short days (8h light/12h dark) based on an incomplete randomised block design. The *ENHANCED DISEASE SUSCEPTIBILITY 1* (*eds1-1*) mutant in Ws-0 background and *eds1-2* in Col-0 background were added in the experiment because they are known immune deficient mutant lines with defects in bacteria SA signaling (Parker et al. 1996; Falk et al. 1999; Feys et al. 2001; Bartsch et al. 2006). Finally, the fast-neutron-generated mutant, *sid2-2*, which carries a null allele of *ICS1* in Col-0 genetic background, was added, serving as control because it is a well-studied SA deficient line (Dewdney et al. 2000; Wildermuth et al. 2001). Representative pictures of plants are shown in Figure 2.2.

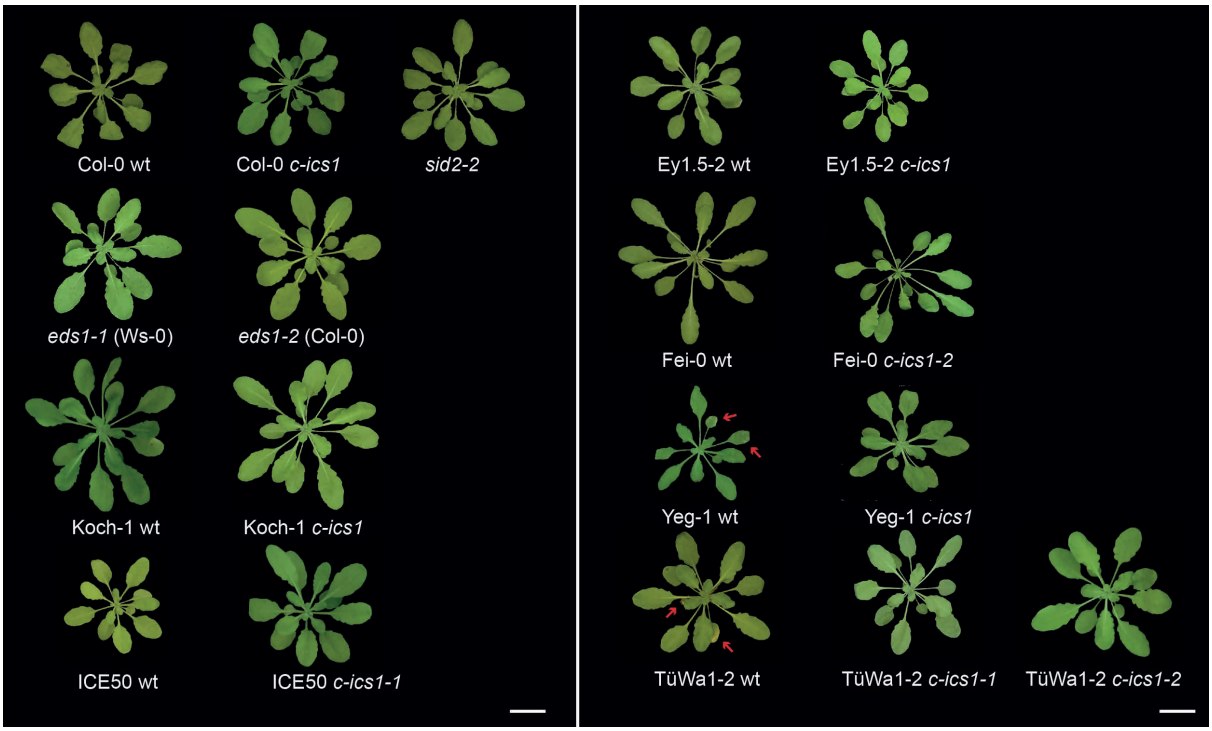


Figure 2.2: Representative pictures of all the lines used during this study. Plants were grown in 23°C SD for 35 days. The white bar indicates 2 cm. The red arrows indicate necrotic spots on the leaves. Pictures were taken using RAPA system (Vasseur et al. 2018).

The Col-0 *c-ics1* mutant showed slightly delayed flowering (defined as main shoot reaching 1 cm, p=0.013) when compared to the wild type but no such difference was observed for *sid2-2* , which is a null *ics1* mutant in Col-0 (Figure 2.3a) (Nawrath & Métraux 1999; Wildermuth et al. 2001). Another similar case was the Ey1.5-2 *c-ics1* mutant for which bolting was slower than the wild type (p=0.005) suggesting delayed flowering time. For both lines no other obvious trait was significantly affected compared to the wild type (rosette and cauline leaf number) (Figure 2.3). One mutant line of TüWa1-2, TüWa1-2 *c-ics1-1*, showed a significant increase in the number of rosette and cauline leaves (p=0.011 and p=0.025 respectively), indicating accelerated flowering time (Figure 2.3b,c). The second line TüWa1-2 *c-ics1-2*, that was carrying an independently generated mutation for *ICS1*, did not show the same pattern (Figure 2.3b,c). Again no other trait was significantly affected but all of them exhibited the same trend of delayed flowering time (Figure 2.3). Koch-1 *c-ics1* plants showed an increase of small significance (p=0.047) for the number of cauline leaves in comparison to wild type with no other trait significantly affected but following the same pattern (Figure 2.3c).

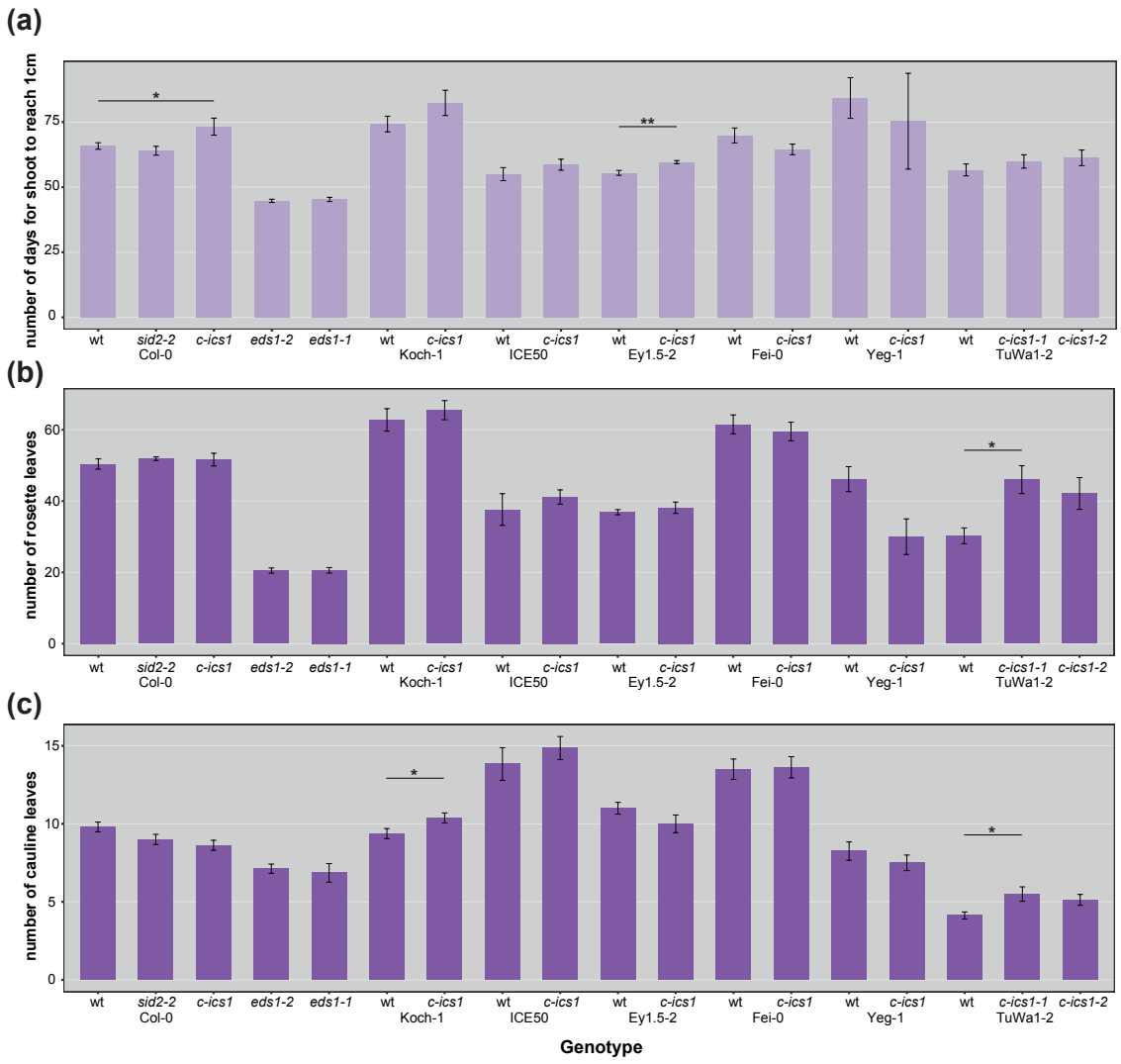


Figure 2.3: Flowering time of *ics1* mutant lines and their wild types. (a) Flowering time is given as days to flowering (shoot elongation to reach 1 cm) and number of (b) rosette and (c) cauline leaves. Plants were grown in 23°C SD conditions. Statistical analysis was performed using linear mixed effect model and Dunnett's test (Bates et al. 2014). ***p<0.001, **p<0.01 and *p<0.05.

I also quantified the free and total SA content using a biosensor assay based on the *Acinetobacter* sp. ADPWH_lux system (Huang et al. 2006; Defraia et al. 2008). I used plants that had been grown in 23°C SD conditions for 38 days. In measurements of free SA, Koch-1 *c-ics1*, Ey1.5-2 *c-ics1*, Fei-0 *c-ics1*, Yeg-1 *c-ics1* and TüWa1-2 *c-ics1-1*,*2*, showed significantly decreased accumulation of SA (p=0.014, p=0.015, p=0.007, p=0.016, p=0.002 and p=0.002 respectively) (Figure 2.4). Col-0 *c-ics1* and *sid2-2* had similar levels of SA with Col-0 wild type (p=0.1) and the ICE50 *c-ics1* despite the fact that low content of SA was not significantly

different from its wild type (p=0.7) (Figure 2.4). When I measured the total SA content (SA and SAG) I was able to verify that Col-0 *c-ics1* plants are unable to accumulate SAG, exhibiting the same pattern as the well-studied *sid2-2* mutant line (Defraia et al. 2008). All samples showed increased amounts of total SA despite the mutation at the *ICS1* locus which is not unexpected since all the other *ics1* mutant lines manage to accumulate free SA (Figure 2.4).

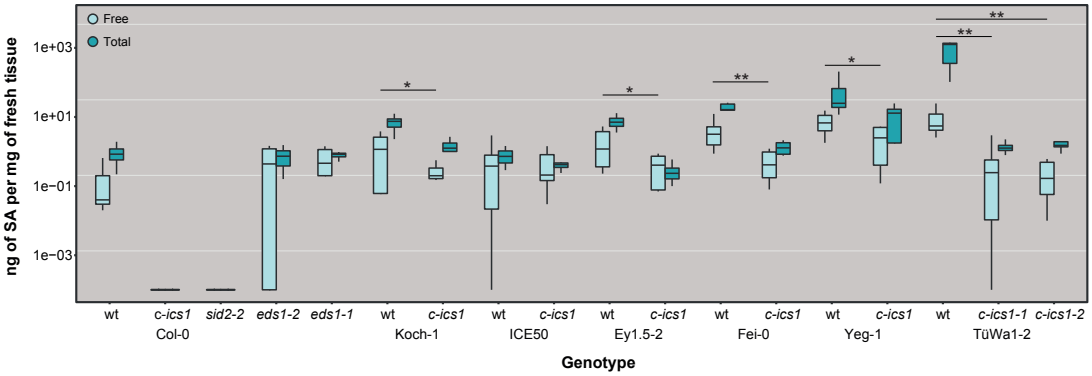


Figure 2.4: SA content of the *ics1* mutant lines and the corresponding wild type parents. Plants were grown in 23°C SD for 38 days. The quantification of free and total SA (SA and SAG) was carried out using the biosensor assay based on the *Acinetobacter* sp. ADPWH_lux (Huang et al. 2006; Defraia et al. 2008). Light blue indicates free SA and dark blue total SA. The statistical analysis for the free SA measurements was performed using a linear mixed effect model and Dunnett's test (Bates et al. 2014). ***p<0.001, **p<0.01 and *p<0.05.

Role of SA in response to *Pseudomonas syringae* pv. *maculicola* ES4326 (*Psm*4326) infections

Multiple studies have used Psm4326 for evaluating SA induction after *Pseudomonas* infection, both in the Col-0 accession and different immune deficient mutants (Gupta et al. 2000; Glazebrook et al. 2003; Zheng et al. 2007; Defraia et al. 2008; Wang et al. 2008). Specifically, (Wang et al. 2008) have shown that SA induction by *Psm*4326 is stronger than after infection with *P. syringae* pv tomato, the most common bacterial model pathogen for *A. thaliana*.

For addressing the response of SA accumulation after pathogen infection I infiltrated the *Psm*4326 strain into 35-day old plants. The plants used were part of an incomplete block design randomisation and from each replicate of the design I used two plants, one for *Psm*4326 infiltration and one for mock infiltration (buffer) (Figure

2.5). I infiltrated leaves using *Psm*4326 suspension of $O.D_{600}$ equal to 0.001 which corresponds to about 10^6 cfu and I collected the infected tissue for SA quantification two days later.

TüWa1-2 *c-ics1-2*	*sid2-2*	Ey1.5-2 wt	Koch-1 *c-ics1*	Fei-0 wt
Ey1.5-2 *c-ics1*	ICE50 *c-ics1*	Ey1.5-2 wt	Col-0 wt	*eds1-1*
Fei-0 wt	Yeg-1 wt	*eds1-1*	*sid2-2*	X
ICE50 wt	Fei-0 *c-ics1*	Col-0 *c-ics1*	Yeg-1 wt	Ey1.5-2 *c-ics1*
ICE50 wt	Col-0 *c-ics1*	TüWa1-2 wt	X	X
TüWa1-2 *c-ics1-2*	Koch-1 wt	Yeg-1 *c-ics1*	Yeg-1 *c-ics1*	*eds1-2*
ICE50 *c-ics1*	Col-0 wt	Koch-1 wt	X	TüWa1-2 wt
Koch-1 *c-ics1*	Fei-0 *c-ics1*	TüWa1-2 *c-ics1-1*	TüWa1-2 *c-ics1-1*	*eds1-2*

Figure 2.5: Representative picture of *Psm*4326 infiltrated and mock treated lines two days after infiltration. The red arrows indicate "water-soaked" symptoms in *Psm*4326 infiltrated plants. The red circle marks a *Psm*4326 infiltrated plant while the blue circle marks a mock treated plant of the same genotype. The table at the bottom of the picture shows the block design. Cells filled with red represent *Psm*4326 infiltrated ones while blue represent mock infiltrated plants.

All accessions and their *ics1* mutant lines were susceptible to the *Psm*4326 strain, as evidenced by the presence of "water-soaked" leaf symptoms (Figure 2.5)

(Dong et al. 1991; Whalen et al. 1991). All the lines showed induction of SA accumulation with the wild type accessions typically having more dramatic changes (Figure 2.6a). I used R and *lme4* (Bates et al. 2015) to create linear mixed effect models for evaluating the relationship between SA levels, the genetic background and the treatment (for the wild-type measurements). Both the genetic background and the treatment effects significantly increased the model fit with the addition of the genotype showing a more significant improvement based on likelihood ratio tests (genetic background: χ^2=16.87, p=3.99e-5, treatment: χ^2=9.85, p=0.0016). Additionally, the Fei-0, TüWa1-2, Yeg-1 and ICE50 *ics1* mutant lines did not show any significant induction of SA during infection (Figure 2.6b). On the contrary, the Koch-1, Ey1.5-2 and Col-0 *ics1* mutant lines showed significant induction of SA with values ranging between 0.1 to 0.5 ng of free SA per mg of fresh tissue, when infected (Figure 2.6b). Interestingly, the *sid2-2* did not show significant SA induction in comparison to Col-0 *c-ics1* despite the fact that they derive from the same genetic background (Figure 2.6b). Finally, all *ics1* mutant lines showed significantly decreased SA levels when compared to the corresponding wild type (Figure 2.6c).

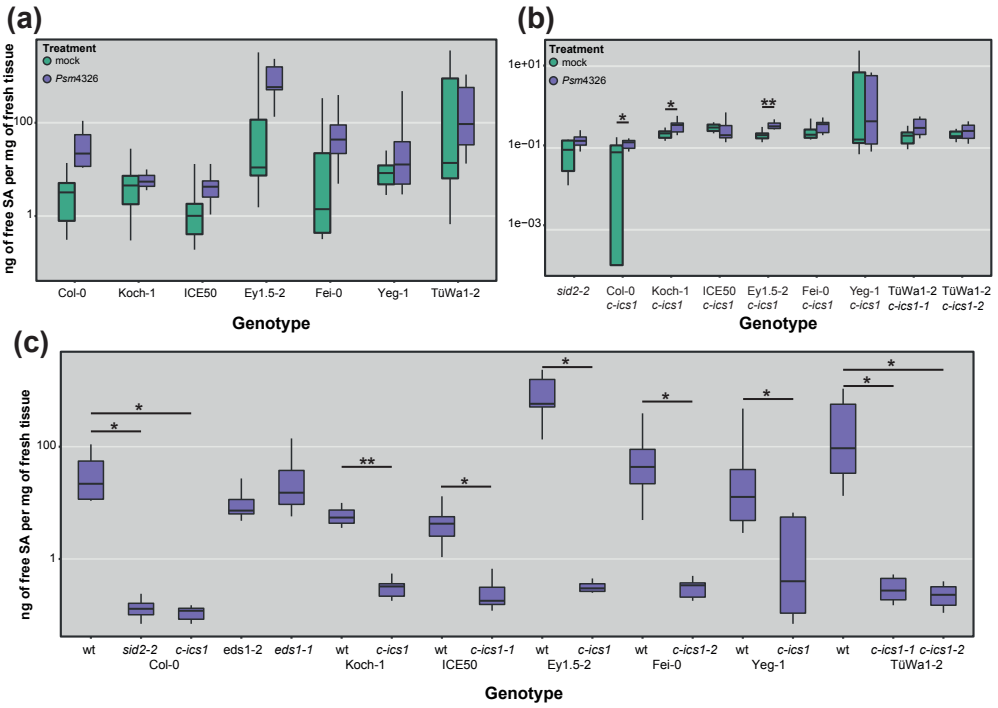


Figure 2.6: Free SA accumulation during *Psm*4326 infection. (a) Free SA quantification of the wild type plants for each accession during *Psm*4326 or mock treatment. (b) Free SA quantification of the *ics1* mutant lines during *Psm*4326 or mock treatment. Note the different scales for graphs (a) and (b). (c) Free SA quantification during *Psm*4326 infection of the wild type plants from each accession

along with their corresponding *ics1* mutant lines. The statistical analysis for panels b and c was performed using Wilcoxon test (McDonald 2014).

Role of SA in response to *Hyaloperonospora arabidopsidis* (Hpa) infections

To examine the contribution of SA towards the pathogenicity of the Hpa, I used the 14OHMLN4 isolated. I collected inoculated plants (Hpa or mock) 11 days post inoculation for Trypan blue staining, SA quantification and phenotyping. Based on susceptibility and the symptoms I categorised the plants into three groups. The first group included the resistant genotypes, Col-0, Koch-1 and ICE50. These accessions showed no sporulation but only the characteristic flecking necrosis, where the cells died due to HR preventing spread of the pathogen (Figure 2.7 and 2.8). In the second group, Ey1.5-2 and Fei-0 accessions were classified as partially resistant with trailing necrosis (Holub 2006), indicative of partial resistance and indicating that HR was elicited but not fast enough to prevent pathogen growth (Figure 2.7 and 2.8). Finally, the last group consisted of the susceptible accessions TüWa1-2 and Yeg-1, which showed prolific pathogen growth (Figure 2.7 and 2.8). In all cases, the *ics1* mutant lines were classified into the same groups as the accessions that they were derived from. The *eds1* mutants, both in Col-0 and Ws-0 background, were susceptible as expected, since the *EDS1* gene is required for SA-dependent ETI signaling and resistance against *Hpa* (Parker et al. 1996; Falk et al. 1999).

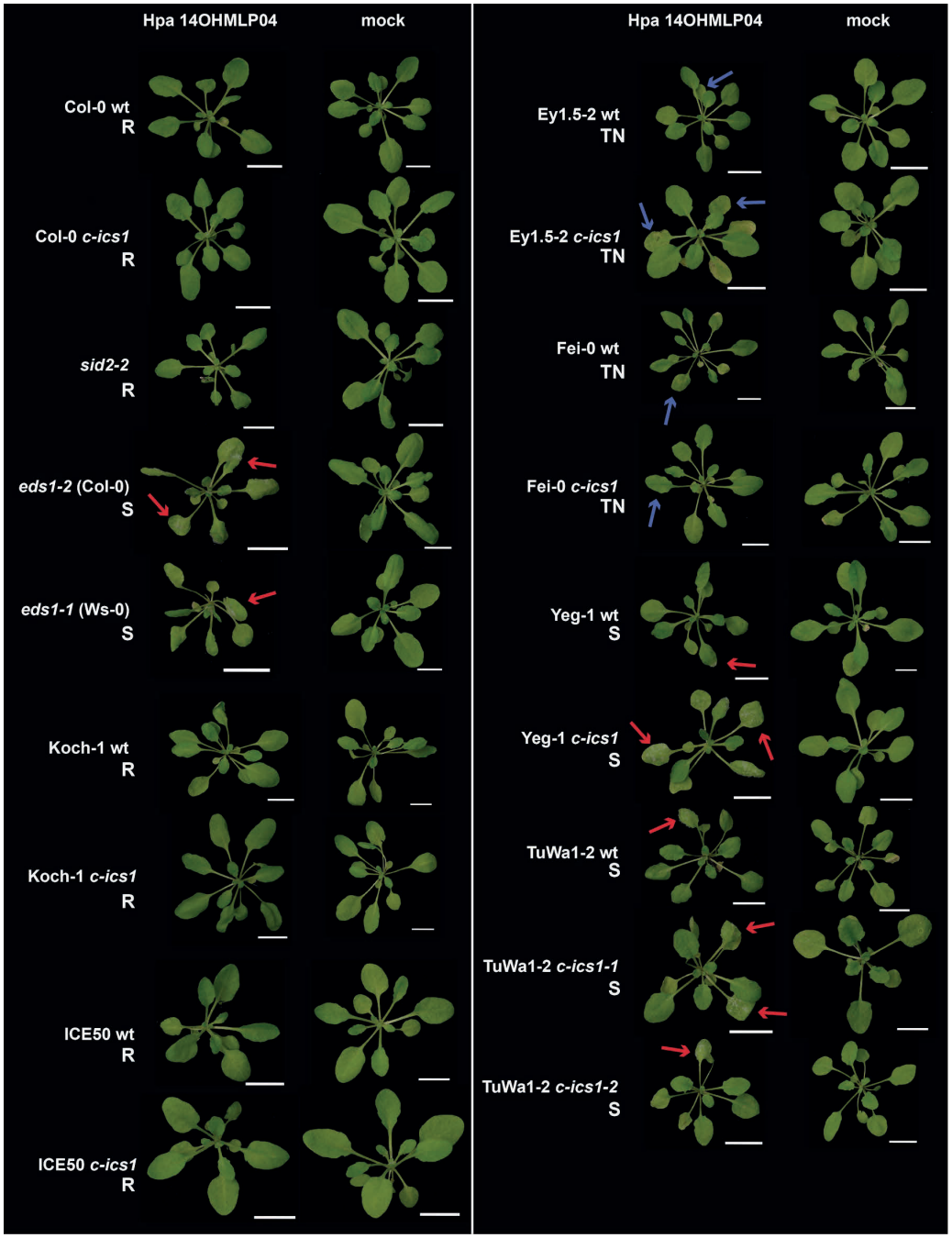


Figure 2.7: Representative images of plants infected with 14OHMLP04 (Hpa) or mock (H_2O). Red arrows point leaves that exhibit sporulation of the pathogen and blue arrows indicate the trailing necrosis symptoms. The white bar indicates 1 cm. R: resistant, flecking necrosis; S: susceptible and TN: partial resistance with trailing necrosis. Pictures were taken 12 dpi.

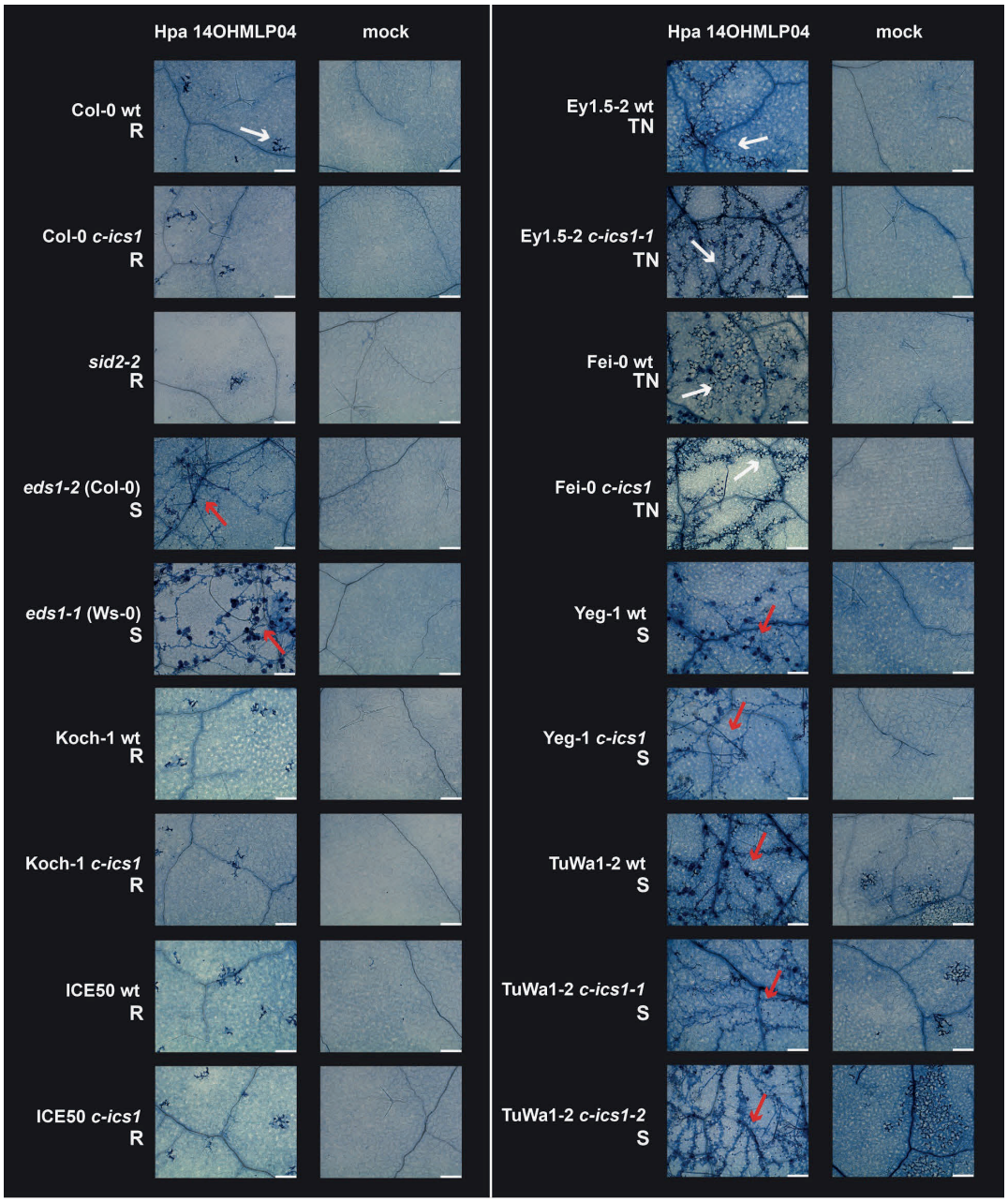


Figure 2.8: Lactophenol Trypan blue–stained leaves of 14OHMLP04 (Hpa) or mock (H_2O) treated plants. Red arrows indicate pathogen growth (hyphae, spores and oospores), the white arrows indicate HR symptoms. R: resistant, flecking necrosis; S: susceptible and TN: partial resistance, trailing necrosis. Leaves were collected 13 dpi. The white bar indicates 200 µm.

In a similar way as before (with *Psm*4326) using likelihood ratio tests I could show that the addition of the genetic background increases significantly the model fit of SA accumulation in response to 14OHMLP4 (χ^2=7.25, p=0.007). The addition of the treatment did not have a significant improvement to the model fit (χ^2=2.7, p=0.1) (Figure 2.9). In both cases only the SA content of the wild type plants, that was

reported, was used. I did not observe any significant induction of SA in the wild type nor the *ics1* mutant lines (Figure 2.9). Additionally, the decrease of SA in the *ics1* mutant lines was not significant when compared to their corresponding wild types (Figure 2.9).

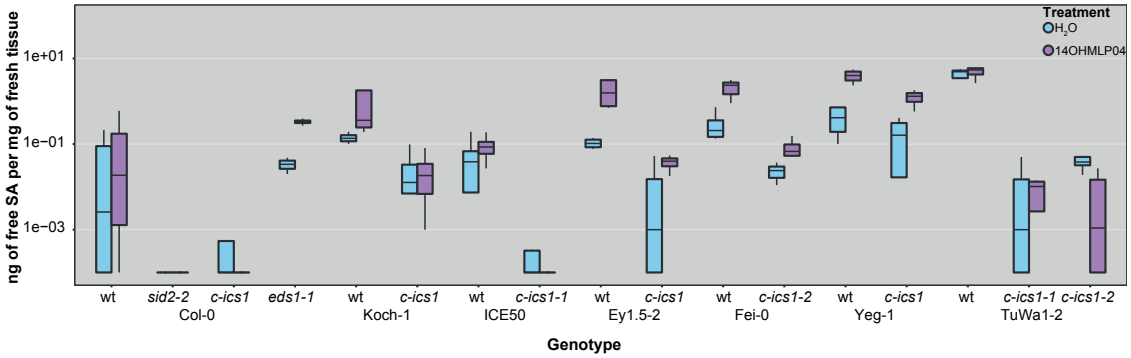


Figure 2.9: Free SA levels during 14OHMLP04 or H_2O treatment. Light blue boxes indicate free SA of plants that were treated with H_2O and the light purple boxes free SA of plants treated with 14OHMLP04. Whole rosettes were collected 12 dpi. The statistical analysis of differences between wild type and corresponding *ics1* mutant line was performed using Wilcoxon test, with no significant difference being reported (McDonald 2014).

Discussion

Salicylic acid (SA) has been implicated in different aspects of the *Arabidopsis thaliana* life cycle such as flowering, leaf senescence, and it is a key signaling molecule for immune responses (Morris et al. 2000; Martínez et al. 2004a; Malamy et al. 1990; Métraux et al. 1990). Most of the studies investigating the role of SA have been conducted using the commonly studied Col-0 genotype (Villajuana-Bonequi et al. 2014b; Kus et al. 2002; Wildermuth et al. 2001) but much less is known about its importance in other natural accessions.

The aim of this Chapter was to evaluate the contribution of SA to differences in flowering time and immune responses in natural accessions of *A. thaliana*. To achieve this, I generated SA compromised mutants lines by introducing premature stop codons in the *ICS1* gene of different accessions using CRISPR/Cas technology. The *ICS1* gene product is part of the main SA biosynthetic pathway in *A. thaliana* (Wildermuth et al. 2001). The mutant lines that I generated carried slightly different mutations at the *ICS1* locus but all of them were predicted to encode truncated proteins missing the important chorismate binding domain (Appendix Figure 1).

Flowering time is not affected by decreased levels of SA

The first trait that I examined was flowering time. The Col-0 *c-ics1* and Ey1.5-2 *c-ics1* mutant lines showed some difference in onset of bolting and TüWa1-2 *c-ics1-1* and Koch-1 *c-ics1* showed some differences in the number of rosette and cauline leaves, but there were no generally consistent effects of SA on multiple aspects of flowering onset (Figure 2.3). Moreover, in the case of Col-0 *c-ics1* and TüWa1-2 *c-ics1-1* the phenotype was not supported by the second mutant line analyzed (Col-0 *sid2-2* and TüWa1-2 *c-ics1-2*), consistent with the marginally significant flowering time results having been spurious (Figure 2.3). My findings are not surprising, given the contradictory flowering time effects attributed to SA in the literature. While some studies suggested that elevated SA levels lead to late flowering (and vice versa) others reported the opposite or indicated no effect (Martínez et al. 2004a; Jin et al. 2007; Wang et al. 2011; Li et al. 2012; Villajuana-Bonequi et al. 2014b; Liu et al. 2014). The SA contribution to flowering in these

studies was examined with lines that carried mutations in genes that affected SA accumulation and at the same time genetically interacted with other genes involved in flowering time regulation (such as *FLC* and *FLD*) or caused abiotic stress, in turn also known to affect flowering (Lee et al. 2007; Singh et al. 2013; Villajuana-Bonequi et al. 2014b). In addition, transgenic NahG plants, in which SA is metabolized, were reported once to be early flowering (Martínez et al. 2004a) but not in another study (Liu et al. 2014). MYB30 overexpressers, which have high levels of SA, also flower early (Liu et al. 2014).

In conclusion, SA had small effects on flowering in some genetic backgrounds, but not in others. Because in some cases only the chronological time to flowering or the number of leaves were affected, it is difficult to come to definitive conclusions. However, my data so far suggest that there are SA effects on flowering time, but that these might depend on the genetic background.

ICS1 is important but not essential for SA accumulation in natural accessions

Significantly decreased SA levels were observed for most *ics1* mutant lines except for Col-0 and ICE50. In the case of Col-0, this is likely due to free SA being very low in wild-type plants ((Defraia et al. 2008), and a further decrease thus very difficult to measure. Inability of *sid2-2* mutant plants (and Col-0 *c-ics1* line) to accumulate SA becomes apparent when total SA (SA and SAG) is measured instead (Wildermuth et al. 2001; Defraia et al. 2008), which I could confirm. The situation is similar in ICE50, which also has very low accumulation of SA (Figure 2.4). In addition, my evidence suggests that the mutations generated in all lines, during editing with the CRISPR/Cas, successfully eliminated *ICS1* function (Table 2.1, Appendix Figure 1).

Not all the *ics1* mutants were devoid of free SA (Figure 2.4). The residual SA can potentially be attributed to differential expression of *ICS2*, which has no significant contribution towards SA accumulation under abiotic stress or pathogen infection in Col-0 (Garcion et al. 2008), or differential activity of the PAL biosynthetic pathway in these accessions. These possibilities can be examined in future studies by expression and genetic studies of *ICS2* and *PAL* in these other accessions.

No significant accumulation of SA during *Psm*4326 infection

To examine the SA response when plants are infected with different pathogens I used a *P. syringae* strain (*Psm*4326) and an oomycete (*H. arabidopsidis*, 14OHMLP04). During *Psm*4326 infection, SA production was induced and all infected plants developed "water-soaked" leaf symptoms (Figure 2.5) (Dong et al. 1991; Whalen et al. 1991). Mutants accumulated significantly less SA (Figure 2.6c). I could show that differences in SA induction was dependent on both infection and genetic background, with the latter being more important.

No significant induction of SA was detected in most *ics1* lines (Figure 2.6b). Ey1.5-2 *c-ics1*, Koch-1 *c-ics1* and Col-0 *c-ics1* lines accumulated only between 0.1 to 0.4 ng of free SA per mg of fresh tissue during the infection (Figure 2.6c). A trend towards induction of SA, albeit non-significant, during infection of *ics1* mutants has been observed previously by (Defraia et al. 2008) with *sid2-2* plants, which have reduced background levels of SA. My observations with natural accessions are in agreement by previous studies when different immune response deficient lines, which also accumulated less SA after infection (Gupta et al. 2000; Glazebrook et al. 2003; Zheng et al. 2007).

SA deficiency affects the severity of symptoms during *Hpa* infection but not *Hpa* resistance

Plants were infected with a specific strain of the oomycete, *H. arabidopsidis* (*Hpa*), fell into three groups: resistant with flecking necrosis (Col-0, ICE50, Koch-1 and their mutants), partially resistant with trailing necrosis (Fei-0, Ey1.5-2 and their mutant) and fully susceptible (Yeg-1, TüWa1-2 and their mutants). Regardless of the grouping, all *ics1* mutant lines accumulated less SA after induction, although the difference was never significant. Considering all the data it seems that the resistance against the *Hpa* strain tested is SA-independent but decreased levels of SA affect the severity of HR during trailing necrosis (partial resistance). It has been shown that infected Col-0 *NahG* plants with the Noco isolate of *Hpa* support increased pathogen growth and even more when infected with the Wela isolate, while Col-0 wild type is resistant for these isolates (Delaney et al. 1994). It has been

shown that resistance to Wela and Noco isolates is linked with SA accumulation (Delaney et al. 1994; Nawrath & Métraux 1999). In the case of the Landsberg-*erecta* (Ler) accession, wild type plants are resistant to Noco2, while Ler *NahG* develop trailing necrosis, indicating that accumulation of SA is important for *RPP5*-dependent resistance to Noco2 (Feys et al. 2001). The situation is similar for Col-0 and the Emco5 isolate (McDowell et al. 2005). The aforementioned studies suggested involvement of SA-dependent signaling in Noco/Col-0, Wela/Col-0 and Noco2/Ler interactions. This is in contrast to my observations, since inactivation of *ICS1* had only minor effects on *Hpa* resistance in the first group of resistant accessions. SA-independent signaling in *EDS1*-dependent responses has been suggested to involve the *FLAVIN-DEPENDENT MONOOXYGENASE 1* (*FMO1*) gene (Bartsch et al. 2006), since *fmo1* mutant plants are able to accumulate SA during *Pst* infection but lose resistance to Cala2 (Bartsch et al. 2006).

Another gene that has been associated with pathogen responses through EDS1 signaling is the *Nudix Hydrolase NUDT7* (Bartsch et al. 2006). The phenotype of *nudt-7* mutant plants, which are characterised by extended cell death and reduced growth is abolished in the *nudt7-1/eds1-2* double mutant but not in a *nudt7-1/sid2-1* mutant line (Bartsch et al. 2006). This observation indicates that *NUDT7* regulation relies entirely in EDS1 signaling (Bartsch et al. 2006). Additionally, both double mutant lines exhibit similar levels of free SA suggesting that the cell death is not caused by accumulation of SA (Bartsch et al. 2006). These findings show that priming of cell death in *nudt7* mutant lines requires *EDS1* but is antagonised by SA that derives from the ICS1 biosynthetic pathway (Bartsch et al. 2006). A *nudt7-1/eds1-2/sid2-1* line abolishes the *nudt7-1/sid2-1* phenotype but shows the same increased susceptibility to Noco2 indicating balance between promoting and inhibiting SA production is vital for resistance (Straus et al. 2010). The increased severity of trailing necrosis that is observed in Ey1.5-2 *c-ics1* and Fei-0 *c-ics1* can be attributed to the absence of SA being the reason of the “uncontrolled” cell death due to *EDS1* signaling.

To conclude, SA accumulation does not appear to contribute to the resistant phenotype of flecking necrosis but it does affect the severity of the trailing necrosis symptoms during 14OHMLP04 infection of partially resistant accessions (Figure 2.7 and 2.8). Further investigation is needed to elucidate the mechanism of SA signaling

regarding the pathogenicity of the 14OHMLP04. Quantification of *FMO1* expression levels during infection in the wild type and *ics1* mutants plants could provide information as to whether the *FMO1* pathway can partially bypass the *ICS1* pathway in these accessions.

Conclusions

This study led to two major conclusions. First, our findings showed an effect of SA on flowering time regulation that was much smaller than what was reported in several other studies.

During *Psm*4326 infection I was not able to observe significant induction of SA in the *ics1* mutant lines. The accumulation of SA in the wild-type plants during the infection was guided strongly by the genetic background along with the treatment.

Finally, during the H. arabidopsisdis (14OHMLP04) infection no significant accumulation of SA was observed suggesting an SA-independent response regarding the resistance. A potential response during the resistance could be attributed to *FMO1* signalling. During partial resistance (trailing necrosis), increased cell death was observed suggesting that SA can antagonise the *EDS1* primed cell death. Further investigation is required to unravel the exact mechanism of the response to the 14OHMLP04 isolate.

Chapter 3

Investigation of off-target cleavage sites in the *Arabidopsis thaliana* genome when *ICS1* gene is targeted using the CRISPR/Cas system

Introduction

Genome editing describes the targeted alteration of the genetic code of an organism (insertion, deletion, replacement or modification). It can be achieved by using nucleases (ZFNs, TALENs or CRISPR/Cas) that have been engineered for targeting and cleaving DNA by creating double stranded breaks (DSBs). Unfortunately, all current nucleases have one common limitation: the unintentional editing of additional regions in the genome that have not been targeted, called "off-targets" (Figure 3.1) (Gaj et al. 2013; Osborn et al. 2016; Gutierrez-Guerrero et al. 2018). Many studies have been focused on predicting potential off-target regions based on their similarity with the targeted one (Sapranauskas et al. 2011; Jinek et al. 2012; Gaj et al. 2013). Since their discovery, off-target editing events have been studied with type II and III CRISPR systems, and single mismatches at the 5' end of the crRNA to the respective PAM sites have been shown to be insufficient to prevent these (Sapranauskas et al. 2011; Jinek et al. 2012).

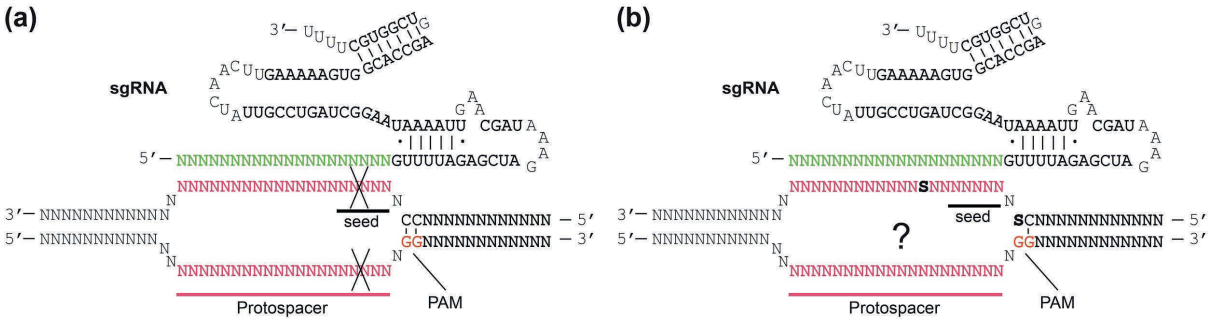


Figure 3.1: Graphical representation of "on" and "off" targeting. The red underline indicates the target site in the genome (protospacer) that is followed by the PAM site and the black underline (5 nt) indicates the seed region of the protospacer. The green region of the sgRNA is the spacer (crRNA) that recognizes the target site and the black region is the scaffold (tracrRNA). The "X"s point to the position of cleavage by Cas9. (a) On-targeting: cleavage occurs with perfect complementarity between sgRNA and one strand of target site. (b) Off-targeting: potential cleavage occurs when there is recognition between sgRNA and a target sequence (eg. a mismatch ("S") in the target) because of partial complementarity.

The same year as the first CRISPR/Cas application was published, studies focusing on investigating the specificity of the sgRNA/DNA complex in human cell lines and bacteria reported that cleavage may happen with imperfect complementarity to the guide RNA (Cong et al. 2013; Fu et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). Using a fluorescent density (eGFP) approach, off-target cleavage was evaluated using mutagenised single guide RNAs (sgRNAs) (Fu et al. 2013). Differences in the cleavage activity of each sgRNA were attributed to the number of mismatches and their position in the sgRNA (Fu et al. 2013). Two mismatches in the 10 bp region upstream of the PAM site decreased cleavage activity to ~5% (Fu et al. 2013). When endogenous genomic regions were targeted, the number of potential off-target sites based on the sequence similarity was between 43 to 64 sites, and the detected off-target cleavage fluctuated between 0%-26% for the different sites (Fu et al. 2013). Additionally, when endogenous loci in the genome were targeted in different regions, a list of potential off-target sites was generated based on similarity and at 3.4% of them, considerable alteration activity could be detected (Hsu et al. 2013). The scaffold of the sgRNA also affects off-target cleavage, with longer scaffolds being more problematic (Pattanayak et al. 2013). The differences in off-target cleavage activity was dependent on the specific sgRNA sequence, with sites showing ~40-90% of difference in off-target cleavage activity (Pattanayak et al. 2013).

For better selection of sgRNAs, *in silico* tools were developed to predict off-target sites based on features that were previously reported by (Hsu et al. 2013) and (Fu et al. 2013) regarding number of mismatches and their position. These tools are programmed either for designing target sites and reporting predicted off-target sites, or exclusively for the prediction of the off-targets for a given target site. There are numerous online tools that have been designed to improve the selection of a targeted region like CRISPR-P 2.0, CHOPCHOP v2.0, CCTop, Cas-OFFinder and CRISPR design (Bae et al. 2014; Stemmer et al. 2015; Labun et al. 2016; Liu et al. 2017); http://crispr.mit.edu/). The developers used different pipelines to predict and score off-targets sites but all of them follow the same principle. Off-target sites are ranked based on their similarity to the targeted site. They are scored based on the number and position of mismatches, the presence of a PAM site, and location of the

off-target region. Finally, the sgRNA is also scored regarding its specificity based on the number of predicted off-target sites and their scores. These tools' prediction ability is limited by the assumption of similarity between regions, the number of allowed mismatches, and the PAM requirement (Bae et al. 2014; Stemmer et al. 2015; Labun et al. 2016; Liu et al. 2017); http://crispr.mit.edu/).

While it is understandable that initial efforts focused on predicted off-targets, it is important to conduct unbiased evaluation of off-target cleavage without assumptions regarding sequence similarity ot the sgRNA (Crosetto et al. 2013; Frock et al. 2015; Kim et al. 2015; Tsai et al. 2015; Wang et al. 2015). IDLV (integration defective lentiviral vector), GUIDE-seq and BLESS (direct in situ breaks labeling, enrichment on streptavidin and next-generation sequencing) are based on identification of DSBs that occur during application of nucleases, either by integrating known molecules to the break or by labelling it (Gabriel et al. 2011; Crosetto et al. 2013; Tsai et al. 2015). Others are based on the identification of chromosomal translocations that occur due to on-/off- target cleavage (HTGS; high-throughput, genome-wide, translocation sequencing) (Frock et al. 2015), or the *in vitro* identification of indels generated during *in vivo* application (digenome-seq; *in vitro* Cas9-digested whole-genome sequencing) (Kim et al. 2015).

All aforementioned studies were conducted using human cell lines. Studies for plants are much more limited (Scheben et al. 2017). *In silico* predicted off-target sites based on similarity exhibited no off-target cleavage when they were tested during CRISPR/Cas application in *Brassica napus,* bread wheat (*Triticum aestivum*), cotton (*Gossypium hirsutum*), cacao (*Theobroma cacao*), tomato, rice or *A. thaliana* (Feng et al. 2014; Zhang et al. 2016; Nekrasov et al. 2017; Ueta et al. 2017; Yang et al. 2017; Wang et al. 2017; Fister et al. 2018; Macovei et al. 2018; Sánchez-León et al. 2018). However, evidence of off-target cleavage has been reported in cases where the off-target sites are nearly identical with the targeted region (1-6 nucleotides difference) in rice, soybean and maize (Shan et al. 2013; Jacobs et al. 2015; Svitashev et al. 2015; Li et al. 2016). According to (Jacobs et al. 2015), the mutation frequency at the off-target site was 2-13% comparing to 95% at the targeted region, while in another study, by (Li et al. 2016), the off-target modification frequency was 47.5% and 67% for regions differing by one or two nucleotides from

the targeted site, respectively. Based on all the aforementioned reports, the specificity of sgRNAs is complex with many unknown variables yet to be determined.

Aims

Exploration of off-target cleavage in plant species has been conducted mainly by focusing on *in silico* predicted sites that rely on sequence similarity between the regions and presence of canonical PAM site (Feng et al. 2014; Zhang et al. 2014; Peterson et al. 2016; Nekrasov et al. 2017; Macovei et al. 2018; Yang et al. 2017). These approaches create biases since noncanonical PAM sites have been observed when studies assess off-target cleavage (Tsai et al. 2015). Additionally, in human cell lines, regions without similarity to the targeted one have been identified as putative off-target cleavage sites (Veres et al. 2014). In this study, I investigated the presence of off-target cleavage sites of Cas9 when the *ICS1* gene is targeted with two sgRNAs in the commonly used Col-0 accession (Figure 2.1, Chapter 2). Using the pEC1.1 promoter, which is active at the egg cell of the female gametophyte, for Cas9 expression allowed me to focus on germline mutations (Sprunck et al. 2012). I generated a bioinformatic pipeline to analyze sequencing data for potential off-target events in a genome-wide, unbiased manner and without *a priori* assumptions on putative hotspots of off-target cleavage sites regarding sequence similarity.

Results

An experimental set-up to investigate CRISPR-Cas9 off-target editing events

My aim was to investigate the prevalence of off-target editing events following the application of CRISPR/Cas to introduce loss-of-function mutations in the *ICS1* locus in *A. thaliana*. To this end, I engineered a transgenic construct with Cas9 driven by the EGG CELL 1.1 (EC1.1) promoter, which is specifically expressed in the egg cell tissue, and two sgRNAs driven by two U6 promoters targeting the *ICS1* gene in the 3rd exon (refer to Figure 2.1 of Chapter 2). The commonly used Col-0 TAIR10 reference accession was used as genetic background for transformation. Because the EC1.1 promoter is active in the egg cell (Sprunck et al. 2012), mutations should be primarily induced in the germline and therefore be heritable, as confirmed in the previous chapter.

To select the primary transformants, I used the seed coat expressed mCherry marker (Gao et al. 2016). I genotyped selected plants at the targeted positions using amplicon deep sequencing as described in Chapter 1. I detected no mutations in any of the primary transformants (0/15), as expected, and propagated two randomly chosen T_1 plants (line 3 and 4) carrying the Cas9 transgene to the next (T_2) generation (Figure 3.2). The T_2 plants were classified as wild type or heterozygous at the respective targeted position; no homozygous or trans-heterozygous events were detected. Based on the experimental evidence regarding on-target cleavage and the expression pattern of the EC1.1 promoter (Sprunck et al. 2012), I hypothesised that inheritance of off-target events will occur in the same manner as for on-target mutations.

To this end, I followed the progeny of three T_2 plants for which the Cas9 transgene was carried as a single copy and would therefore be lost in ¼ of the progeny. Two were heterozygous mutant at the target position (plant lines 3 7; het 3_7 and line 3 32; het 3_32) and derived from the same parent, and one exhibited wild type reads for the target position (plant line wt 4 7; wt 4_7) and derived from a different parent (Figure 3.2). Focusing on T_3 plants that lost the Cas9 transgene due to segregation, I randomly selected twenty of them from each of the three families

and prepared three equimolar DNA pools. By using a large number of randomly selected plants from each line, the mutant sequencing reads at the target locus should reflect segregation of mutant plants in each family, and serve as positive control during data processing. From now on, I will term each pool/sample based on their genotype in the T_2 generation (het 3_7, het 3_32 and wt 4_7). If my hypothesis that off-target variants occur in an inherited manner is correct, and considering the Mendelian law of segregation, the mutated allele should have a frequency of 50% in the final pool. Finally, I used a fourth pool of twenty untransformed T_1 plants as a control for the experiments (Figure 3.2).

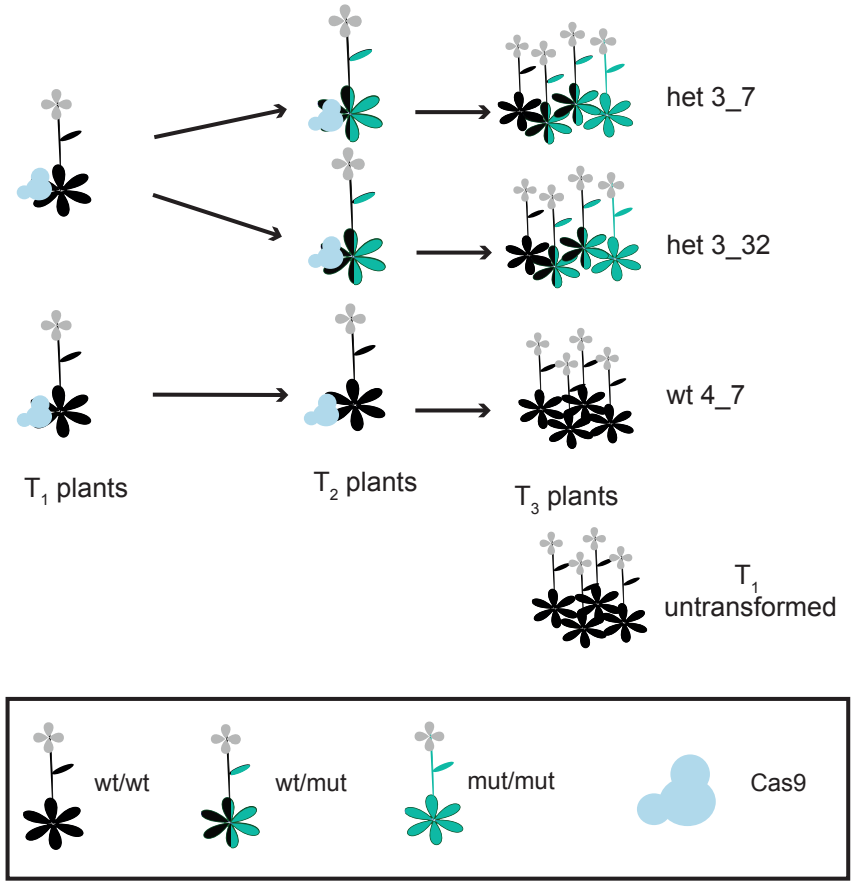


Figure 3.2: Representation of selection process and selected plants. Two T_1 plants that were carrying Cas9 and showed no mutation in the targeted region were followed to the subsequent generation with Cas9. Three T_2 plants were selected that were carrying Cas9. Two of the plants were heterozygous offspring from the same parent, while the third one was offspring of a different parent and had no mutations in the targeted region. At the next generation, only plants without Cas9 were sown out from each parent; 20 of them were randomly selected without genotyping the targeted region and pooled in order to generate the samples for sequencing. As a control I used a pool of 20 randomly picked T_1 plants that underwent the transformation procedure but the transgene was not

integrated (untransformed). The different graphical representation of each genotype and component of the figure are explained at the bottom.

Sample sequencing and BAM file processing

To investigate off-target cleavage occurrence, the four pools were sequenced on a HiSeq3000 Illumina (Illumina, USA) instrument aiming for average genome coverage of 50x. Reads were mapped to the TAIR10 reference using BOWTIE2 (v2.2.3) and downstream processing was performed using samtools (v1.3.1), GATK (v3.5) and PICARD (v2.2.1) softwares (Li et al. 2009; McKenna et al. 2010; Langmead & Salzberg 2012); (Arabidopsis Genome Initiative 2000)http://broadinstitute.github.io/picard(Arabidopsis Genome Initiative 2000). Using samtools (v1.3.1), an estimation of the average number of reads covering a base pair in the genome was calculated (Li et al. 2009) (Figure 3.3).

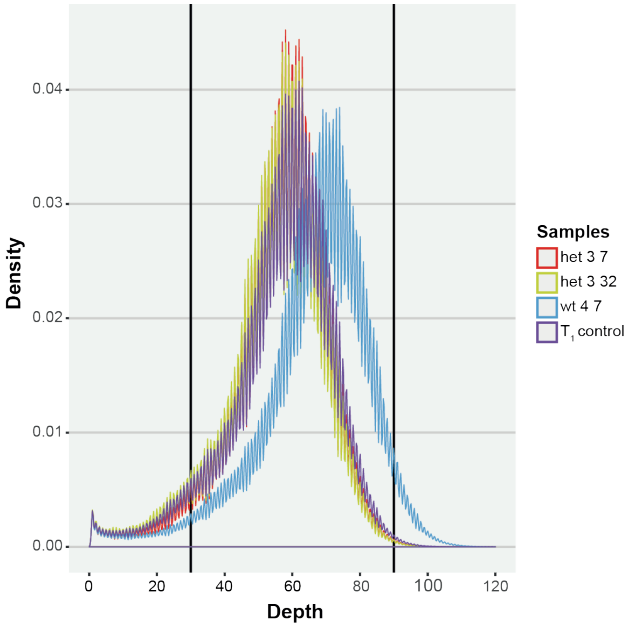


Figure 3.3: Density plot of read coverage (sequencing depth) in four samples. The black lines indicate the depth cut off (30 to 90) that I used during the filtering.

Variants were called using FREEBAYES (v1.1.0) with standard parameters, and the datasets of variants (VCF files) for each sample were filtered for high quality calls (Garrison & Marth 2012b) (Figure 3.4a). No dramatic difference in the initial number of variants was observed among the CRISPR/Cas9 samples as well as the control (pool of untransformed T_1 plants) (Figure 3.4a). Finally, applying each

filter either one at a time after intersection, or one after the other, I was able to eliminate the possibility of bias towards a specific filtering parameter or a sample (Figure 3.4b,c). The filters used were controlling for the reads covering a region (DP), the mean mapping quality of observed variant (MQM), the probability of observing a variant in the specific region (QUAL), the number of reads supporting a variant (AO) and the number of reads covering the variant further away from each end of the read (RPR and RPL). After all filters, between 20 and 48 variants were left in each set. At this point the genotype of the remaining variants has not been controlled.

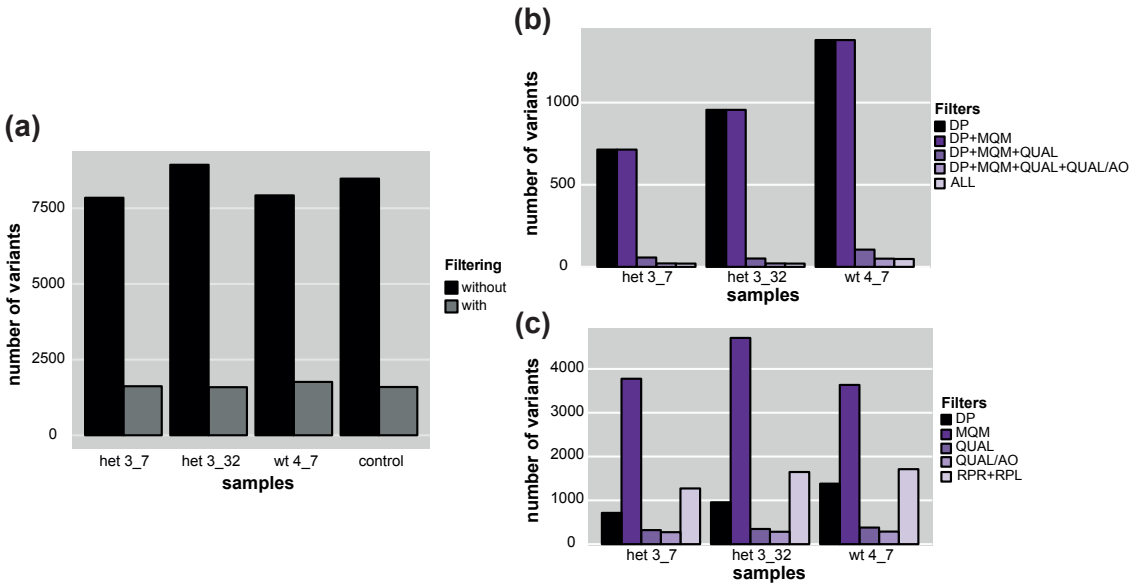


Figure 3.4: Initial variant calling and filtering. (a) Variants were called with FREEBAYES v1.1.0 with default settings (Garrison & Marth 2012b), and the filtering was performed with the vcffilter option from the vcflib package. (b) Effects of sequential filtering. (c) Results of separate filtering. Parameters used were DP=30-90, MQM >30, QUAL >30, QUAL/AO >10, RPR >1 and RPL >1. For (b) and (c) the filters were applied in the intersected vcf files of each sample and for this reason the control sample is missing.

Identification and validation of unique variants

The analyzed samples came from the reference accession Col-0, but it is expected that the Col-0 line used in the lab has a number of shared variants that distinguish it from the TAIR reference genome, since new mutations accumulate at a rate of about 1 per generation (Ossowski et al. 2010). Mutations induced by

CRISPR/Cas9 are expected to be unique to each sample. In order to identify such unique variants in the samples, the variants called in the control sample (pool of T_1 untransformed plants) were intersected with the variants called in each of the experimental samples. After filtering based on the aforementioned parameters (read depth, mapping quality, supporting reads) (details in material and methods) further removal of variants located in repeats regions was performed (removed variants 8-16). Shared variants between samples het 3_7 and het 3_32 (10 in total) were also removed in order to eliminate mutations that were unique to the shared T_1 founder. Moreover, variants with strong strand bias that was calculated based on (Guo et al. 2012) were removed along with variants that were supported even by one read at the control sample (T_1 untransformed variants). Finally, homozygous variants were removed based on the hypothesis that the CRISPR/Cas mutations were occurred in a heterozygote manner in the previous generation.

het 3_7 and het 3_32 samples both had different indels in the targeted region, and these were identified with my filtering approach (Figure 3.5). The number of unique variants after removing the on-target variants were two, zero and five for the het 3_7, het 3_32 and wt 4_7 samples, respectively (Figure 3.5).

I evaluated the unique variants with Sanger sequencing of six individual plants from each line (i.e., before pooling) (Table 3.1). All variants were confirmed.

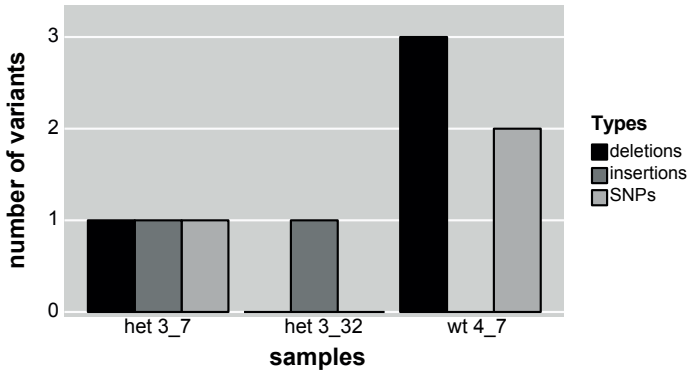


Figure 3.5: Unique variants in each sample. For the samples het 3_7 and het 3_32, the only insertion is in the targeted region. After intersection and filtering of the vcf file, variants in repeat regions, with strong strand bias and variants that were supported with at least one read in the control sample were removed.

Spontaneous mutations occur primarily in the form of single nucleotide substitutions (Ossowski et al. 2010). I recovered more indels than SNPs. All the

variants were located within 5 nt distance from a canonical or non-canonical PAM site (PAM sites that do not follow the expected NGG motif), (Table 3.1) (Hsu et al. 2013; Jiang et al. 2013; Tsai et al. 2015).

The rate of *de novo* generated mutations in *A. thaliana* is approximately one per haploid genome per generation, which means that there is an approximately 14% chance that no new mutation occurs in a single generation, 27% that a single new mutation occurs, and 59% that more than one new mutation occurs, which does not differ from the unique variants that I identified in this study (Ossowski et al. 2010). Thus, the number of observed variants is close to that expected from the background mutation rate.

Table 3.1: Sanger validation of unique variants. 6 individual plants were sequenced for each variant. On-target mutations highlighted in purple. For het 3_7 in position chr1 117403331, one sample failed sequencing. The parenthesis next to the proposed PAM site indicates the distance of the variants from the PAM site. The "-" indicates that the proposed PAM was part of the identified deletion. The alignments of the sgRNA sequences to the positions of the off-target cleavage sites can be found at the Appendix Figure 2.

sample	chr	position	variant	location	gene	ref	alt	effect	PAM	individual genetoyping		
										hom	het	wt
het 3_7	chr1	11740331	deletion$_1$	5'UTR	AT1G32470	TTTTATTGACG	TG	no predicted effect	NTG (3)	1	2	2
	chr1	28071633	insertion$_1$	exon	AT1G74710	GAC	GAAC	frameshifting	NGG (4)	1	4	1
	chr5	16997961	SNP$_1$	integenic	-	G	C	no predicted effect	NAC (4)	NA	1	NA
het 3_32	chr1	28071634	insertion$_2$	exon	AT1G74710	GAC	GATC	frameshifting	NGG (3)	2	2	2
wt 4_7	chr1	16902771	SNP$_2$	integenic	-	G	A	no predicted effect	NGC (5)	2	3	1
	chr1	22057548	deletion$_2$	integenic	-	TGTTAAGAGTAAG	TG	no predicted effect	NGT (-)	2	3	1
	chr1	26836216	deletion$_3$	intron	AT1G71200	ATAGATCT	AT	no predicted effect	NGG (1)	2	2	2
	chr3	23201179	deletion$_4$	5'UTR	AT3G62720	CTTCCTTTCCAT	CT	no predicted effect	NAG (5)	3	1	1
	chr5	25073734	SNP$_3$	exon	AT5G62440	C	T	missense	NGT (3)	4	12	4

Prediction off-target tools

Using several *in silico* methods (CRISPR-P 2.0, CCTop, CRISPR design, CHOPCHOP v2 and Cas-OFFinder), I generated a list with predicted off-target cleavage sites (Bae et al. 2014; Stemmer et al. 2015; Labun et al. 2016; Liu et al. 2017); http://crispr.mit.edu/). The CHOPCHOP v2 prediction tool did not predict off-target cleavage sites for either of my sgRNAs, while Cas-OFFinder generated a list of >31,000 potential off-target cleavage sites. This is not surprising because it allows up to 9 mismatched nucleotides and DNA/sgRNA bulges (Bae et al. 2014; Lin et al. 2014). A more reasonable number of potential off-target sites, of 7 to 19, was generated by the CCTop, CRISPR design and CRISPR-P 2.0 tools (Stemmer et al. 2015; Liu et al. 2017); http://crispr.mit.edu/) (Appendix Table 4). I investigated my

sequencing data for the presence of variants in all predicted off-target cleavage sites either by visualization of reads in IGV or bioinformatically. No variants were detected.

Discussion

A controversial issue in the use of genome editing is the question of unintended off-target effects (Yee 2016; Scheben et al. 2017). This question has only been partially addressed in plants, and I used a line in which Cas9 activity is restricted to the egg cell to investigate heritable off-target effects. Using deep sequencing, I found a mutation rate that was not measurably different from the rate of spontaneous *de novo* mutations in *A. thaliana* (Ossowski et al. 2010).

There have been multiple studies conducted in human cell lines, investigating off-target cleavage during CRISPR/Cas system application by evaluating the specificity of sgRNAs through its mutagenesis or by creating libraries of fragments carrying potential off-target cleavage sites (Fu et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). Their conclusion was that mutations can be tolerated by Cas9, especially 8 nucleotides distal to the PAM site, and that off-target cleavage can take place, but that more mutations decrease off-target cleavage likelihood (Fu et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). Among the mutations that I found in my lines, none affected a sequence with fewer than 10 mismatches to the target sequence (10-13 mismatches) (Appendix Figure 2). Based on the aforementioned studies and the number of mismatches between the sgRNAs and the regions with the identified mutations it seems that it is unlikely to be occurred due to Cas9 activity.

Off-target mutations induced by Cas9 have been investigated in several plants, including rice, *Brassica napus*, bread wheat, cotton, cacao, tomato and *A. thaliana.* None reported mutations at predicted off-target sites (Feng et al. 2014; Zhang et al. 2016; Nekrasov et al. 2017; Ueta et al. 2017; Yang et al. 2017; Wang et al. 2017; Fister et al. 2018; Macovei et al. 2018; Sánchez-León et al. 2018). This is a concern because in other systems, off-target cleavage sites were identified in regions of the genome adjacent to non-canonical PAM sites (NAG, NGA, NAA, NGT, NGC and NCG) (Tsai et al. 2015) (Table 3.1). In contrast to these prior studies, I have not restricted my analyses to predicted off-target cleavage sites, but searched for off-target mutations in an entirely unbiased manner. It is unclear how well available tools can predict potential off-target cleavage sites. Several tools (CRISPR-P 2.0, CHOPCHOP v2, Cas-OFFinder, CRISPR design and CCTop)

predicted various number of different off-target sites or none at all, generating inconsistent results (Bae et al. 2014; Stemmer et al. 2015; Labun et al. 2016; Liu et al. 2017); http://crispr.mit.edu/).

Surprisingly, when *in silico* tools were used for predicting off-target cleavage sites in bread wheat, no Cas9 cleavage activity was detected at these sites; when very close homologues were investigated, however, it was found that one mismatch can be tolerated and cleavage activity was observed at similar rates as the targeted region (Zhang et al. 2016). Findings have been reported in rice, maize and soybean studies, with the mutation rate at off-target cleavage sites fluctuating between 2% and 67% indicating great variability (Shan et al. 2013; Jacobs et al. 2015; Svitashev et al. 2015; Li et al. 2016). Another study in rice exploited the off-target cleavage activity of Cas9 for targeting three highly homologous regions using a single sgRNA (Endo et al. 2014). These findings suggest that off-target cleavage can occur in plant genomes in regions that share high homology (Shan et al. 2013; Endo et al. 2014; Jacobs et al. 2015; Svitashev et al. 2015; Li et al. 2016). Based on my findings I was not able to detect any mutation in regions with high homology to the targeted (10-13 mismatches between sgRNAs and region carrying identified mutation).

Approaches for investigation of potential off-target cleavage sites have been developed mainly based on the identification of DSBs that are created due to Cas9 cleavage activity, generated translocations, or *in vitro* identification of InDels (Crosetto et al. 2013; Frock et al. 2015; Tsai et al. 2015; Kim et al. 2015; Wang et al. 2016). These methods can lead to overestimation of off-target cleavage incidents due to naturally occuring DSBs because of cellular physiology, genomic regions that are more fragile, and the likelihood of mutation generation given a particular cell line (Lieber 2010; Hussein et al. 2011; Gore et al. 2011; Ma et al. 2012). The approach used in the present study was based on generation of stable, heritable mutations (Table 3.1).

To conclude, off-target cleavage has been described in the literature as the unintended binding and cleavage of Cas9 to regions beside the targeted one (Yee 2016). During this study, I investigated the presence of off-target cleavage when the *ICS1* gene was targeted with two sgRNAs in *A. thaliana*. Based on my findings off-target cleavage can potentially occur with a low frequency (0, 2 or 5 variants depending on the line in question) similar to *de novo* generated mutations (~ one per

generation) (Ossowski et al. 2010). This similarity in numbers makes it hard to unambiguously assign the variants to Cas9 cleavage activity.

General Conclusions

The studies described here further our understanding of the involvement of salicylic acid in different natural accessions of *A. thaliana* in flowering time regulation and immune responses. By eliminating the *ICS1* activity using CRISPR/Cas system, the main SA biosynthetic pathway was disrupted. This allowed us to observe individuals with decreased levels of SA. The study followed three main directions.

To streamline the identification of CRISPR/Cas9 induced mutations, I introduced a high throughput (up to 900 samples/run), precise and cost-efficient screening method for reliably identifying individuals that carry events at targeted positions (CRISPR-finder). Using this method I was able to successfully identify individuals from seven natural accessions carrying mutations that generate a premature stop codon at the *ICS1* gene.

With the *ics1* mutant lines in hand, I investigated the effects of reduced SA accumulation on flowering and immune responses. There is considerable controversy in the literature regarding a role of SA in flowering time control (Martínez et al. 2004a; Wang et al. 2011; Villajuana-Bonequi et al. 2014b; Liu et al. 2014). My *ics1* mutants did not have consistent flowering time changes, arguing against an important role of SA in the control of this life history trait. .

In all genetic backgrounds tests, induction of SA in *ics1* mutants after infection with *Psm*4326 was either eliminated or greatly reduced. The residual induction of SA accumulation in some backgrounds points to differential activity of *ICS2* and/or PAL pathways in different accessions, such that they are partially redundant with the ICS1 pathway in some, but not in other accessions.

In contrast to *Psm*4326, infection with *H. arabidopsidis* (14OHMLP04) does not significantly induce SA. In agreement, accessions that were completely resistant to this pathogen did not develop disease when *ICS1* was mutated, suggesting an SA-independent response for conferring resistance. However, partially resistant accessions showed increased cell death during the trailing necrosis symptom when *ICS1* was mutated. This observation could be attributed to the lack of sufficient SA accumulation to antagonise the cell death that is primed due to *EDS1* signalling (Bartsch et al. 2006; Straus et al. 2010). A third group of accessions was completely

susceptible to *H. arabidopsidis* (14OHMLP04) already when *ICS1* was functional, and no conclusion can be drawn as to the importance of SA in these systems.

Finally, the third part of this study was focused on the investigation of mutations generated due to off-target cleavage of Cas9 (Yee 2016). It has been reported in various experiments in human cell lines and in few cases in plant genomes--only in regions with high homology (Hsu et al. 2013; Pattanayak et al. 2013; Jacobs et al. 2015; Li et al. 2016). Using the Col-0 accession and its high quality genome sequence I was able to examine the occurrence of such events when *ICS1* gene was targeted with the two sgRNAs. I did not find evidence for an increased mutation rate in three lines examined in detail, indicating that off-target effects are not an issue of great concern in *A. thaliana*, and potentially neither in other plants.

Material and Methods

Chemicals were mainly obtained from Sigma-Aldrich (Taufkirchen, Germany), Roth (Karlsruhe, Germany), and VWR/Merck (Darmstadt, Germany) unless otherwise stated. Buffers, media and protocols were mostly prepared according to Sambrook and Russell (2001).

Plant growth conditions

Arabidopsis thaliana accessions were used throughout this study. The seeds were kept at -80°C prior to sterilisation for an overnight and then surface-sterilised with 70% ethanol and 0.05% (v/v) Triton X-100 for 5 minutes, followed by 100% ethanol for 5 minutes. Seeds were air-dried in a sterile hood until all residual ethanol had evaporated. Seeds were stratified in 0.1% (w/v) agar-agar for 7 days in the dark at 4°C prior to sowing. Late flowering accessions were vernalized for seven weeks in 4°C short-day (8 h light/16 h dark) and then were moved to 23°C long-day (16 h light/8 h dark). The other accessions were grown from the beginning in 23°C long-day (16 h light/8 h dark).

Plants used for the Chapter 2 phenotyping experiment were grown in 23°C short-day (8 h light/16 h dark).

Unchallenged plants used in Chapter 2 for measuring salicylic acid (SA) were grown in 23^{o}C short-day (8 h light/16 h dark) for 38 days.

Plants used for *Psm*4326 infections in Chapter 2 were grown in 23°C short-day (8 h light/16 h dark) for 35 days, and after infiltration, were returned to the same conditions for another 2 days.

Plants used for 14OHMLP04 infection in Chapter 2 were grown in 23°C short-day (8 h light/16 h dark) for 20 days and then moved to a Percival Arabidopsis growth cabinet in 16°C for 3 days. After spray inoculation with spores, plants were returned to the percival for another 12 days (more detailed description in another paragraph).

The light bulbs for the growth rooms were in a 2:1 ratio of TLED 18W (deep red/white, ca. 15% blue) and 20W (white, ca. 25% blue) Philips GreenPower with a photon stream of 70%.

pcoCas9 mutagenesis

In order to introduce the pcoCas9 gene to the cloning system (Greengate) I had to remove Eco31I sites that were present in the gene. Eco31I is the main enzyme being utilised for the final cloning of the system (Lampropoulos et al. 2013) and by removing present sites I aimed to increase the efficiency of the reaction. For mutating Eco31I sites in the pcoCas9 gene, overlapping PCR amplification was carried out to alter one base pair of each recognition site. The choice of the new nucleotide was based on the frequency of the codon in *Arabidopsis thaliana*.

pcoCas9 was separated into six parts based on the location of Eco31I recognition sites (Figure 5). For amplifying five individual fragments, a plasmid containing the pcoCas9 sequence (courtesy of Prof. Dr. Jen Sheen; (J.-F. Li et al. 2013)) was used as template (2.5 ng) with 0.5 µM forward oligonucleotide, 0.5 µM reverse oligonucleotide, 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (Thermo Fisher Scientific, #R0182) and 0.02 U/µL Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific Inc.) to a final volume of 25 µL. The oligonucleotide set for the generation of each fragment and the cycling conditions for each amplification can be found in Table 1.

Figure 5: Schematic representation of pcoCas9 gene. The purple lines indicate positions of Eco31I recognition sites that were mutated. The arrows represent the oligonucleotide position for altering the recognition sites. The blue segment indicates the location of the potato IV2 intron (J.-F. Li et al. 2013).

Table 1: Oligonucleotide sets and cycling conditions used for generating the initial six fragments from the pcoCas9 gene.

fragment	Oligonucleotide set	Conditions	
		Annealing temp (^{o}C)	Extension time
1	G-38382/G-38369	55	30''
2	G-38368/G-38371	55	30''
3	G-38370/G-38373	55	1'
4	G-38372/G-38375	55	1' 40''
5	G-38374/G-38377	55	30''
6	G-38376/G-38383	55	1'

The fragments 1-2-3 and 4-5 were fused. Equimolar pools of fragments 1-2-3 and 5-6 were used as template for each reaction. The 1st amplification of the overlap PCRs consisted of 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (Thermo Fisher scientific, #R0182) and 0.02 U/μL Phusion High-Fidelity DNA polymerase (Thermo Fisher scientific, #F530) to a final volume of 25 μL. The cycling program was designed with 60^{o}C annealing temperature for 30 seconds and extension of 1 minute at 72^{o}C for 15 cycles. In the same tube, another reaction of 0.5 μM forward oligonucleotide, 0.5 μM reverse oligonucleotide, 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (#R0182) and 0.02 U/μL Phusion High-Fidelity DNA polymerase (#F530) to a final volume of 50 μL. The cycling program was designed with annealing temperature at 55^{o}C for 30 seconds and extension time of 2 minutes and 30 seconds at 72^{o}C for 35 cycles following with 10 minutes of final extension at 72^{o}C.

Finally, the same approach was followed for fusing the final three fragments (1/2/3 + 4 + 5/6). The reaction and the first cycling program remained the same,except that the cycling program was carried out for 20 cycles. The added reaction and the final cycling program remained the same as before with extension of 3 minutes and 30 seconds for 35 cycles. Oligos and detailed cycling conditions for fusing the pcoCas9 fragments together can be found in Table 2.

Table 2: Overview of oligonucleotide sets and amplification conditions used for fusing the different fragments of pcoCas9 together. T_m: annealing temperature

Fused fragment	Oligonucleotide set for 2nd PCR	Conditions			
		1st PCR		2nd PCR	
		T_m (°C)	Ext. time	T_m (°C)	Ext. time
1+2+3	G-38382/G-38373	60	1’	55	2’ 30’’
5+6	G-38368/G-38371	60	1’	55	2’ 30’’
1/2/3+4+5/6	G-38382/G-38383	60	1’ 30’’	55	3’ 30’’

Plasmid generation

The transgene constructs were generated using the GreenGate cloning system (Lampropoulos et al. 2013). The five different constructs used are described in Supplementary File 1. Two versions of Cas9 were used: the plant codon optimised (pcoCas9) (J.-F. Li et al. 2013) and the Arabidopsis codon optimised (AthCas9) (Fauser et al. 2014). The promoters used were CaMV35S, *ICU2* and EC1.1 (courtesy of Dr. Martin Bayer) (Sprunck et al. 2012; Hyun et al. 2013; Lampropoulos et al. 2013).

Plant transformation

Plants were transformed using the flora dip method as described by (Clough & Bent 1998).

Selection of Cas9-free plants

Two selection markers were used, resistance to glufosinate ammonium (BASTA SL, Bayer Crop Science, Leverkusen, Germany) and AT2S3::mCherry (Gao et al. 2016). To select transgenic free plants that no longer carried the BASTA resistance, leaves were brushed a solution with diluted from the original stock (200 g/l) BASTA (1:1000

or 1:2000). The treatment caused leaves from plants without the transgene to become wrinkled and yellowish.

Seeds from plants that were carrying the AT2S3::mCherry (Kroj et al. 2003) cassette were screened for fluorescence or absence thereof under a fluorescent microscope (LEICA MZFLIII fluorescence stereoscope with SOLA SM Light Engine© lamp).

DNA isolation

Genomic DNA was extracted following a protocol by (Edwards et al. 1991), with an additional ethanol wash. DNA was resuspended in 100 µL of ddH_2O.

Plasmid DNA was isolated using GeneJET Plasmid Miniprep kit (K0503) from Thermo Fisher Scientific according to the protocols provided by the manufacturers.

GeneJET Gel Extraction kit (K0692) from Thermo Fisher Scientific was used for the extraction of DNA fragments from agarose gels according to the protocols provided by the manufacturers.

Polymerase chain reaction (PCR)

For all PCR amplification, the same protocol was used unless otherwise stated. For each reaction, 1-6 ng of plasmid or 5-50 ng of genomic DNA were used as template with 0.5 µM forward oligonucleotide, 0.5 µM reverse oligonucleotide, 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (#R0182) and 0.02 U/µL Phusion High-Fidelity DNA polymerase (#F530) to a final volume of 25 µL. Oligonucleotide sets can be found summarised in Appendix Table 2.

Agarose gel electrophoresis

PCR products and restriction enzyme fragments were separated in agarose gel prepared with 1x TAE [2.0 M Tris Acetate, 0.05 M EDTA buffered by glacial acetic acid (~57.1 mL per liter) to pH 8.2 – 8.4]. The concentration of the gel varied (1-1.5 % w/v, with the expected fragment size of products loaded. Loading buffer (50% (v/v) glycerol with orange G) was mixed with PCR products and restriction fragments, for visualization and tracking of the electrophoresis progression. A molecular-weight size marker (GeneRuler DNA ladder mix, #SM0331) was loaded

on to wells adjacent to the samples being resolved. The gels were run at 130 V/cm until the desired separation was observed. The fragments were visualised with UV light (302 nm) and documented using Alphamager (Alpha Innotech, Genetic Technologies, Inc., Florida, USA).

Salicylic acid quantification

The protocol was adapted from (Marek et al. 2010). Fresh tissue was collected and frozen at -80°C overnight. For every 175 mg of fresh tissue, 250 µL of 0.1 M pH 5.5 sodium acetate was added post grinding for further vortexing. *Acinetobacter* sp. ADPWH_lux strain was used (Huang et al. 2006) for the quantification of salicylic acid. Overnight culture of *Acinetobacter* sp. ADPWH_lux at 37°C was diluted (1:20) and grown at 37°C shaking at 200 rpm until it reached O.D.600 of 0.4. For measuring free and 2-O-β-D-glucoside (SAG) SA, plant crude from the samples was incubated at 37^{o}C for 1.5 hours with 0.4 U/µL of β-glucosidase prior to measurement.

Black optiplates (96 wells, greiner bio-one, ref:655906) were used for the measurements. They were prepared with 50 µL of LB, 60 µL of the cell culture and 30 µL of the plant extract. The standards that were prepared with 50 µL of LB, 60 µL of the cell culture, 10 µL of known SA concentrations and 20 µL of plant extract from NahG plant (Col-0 background) (prepared the same way as the samples). The plates were incubated at 37°C for 2 hours without shaking and the luminescence was measured using the TECAN infinite F200 instrument and the i-control 1.12 software.

Chapter 1

Amplicon Libraries

The amplicon libraries were generated by a two-step PCR. The first reaction consisted of 1 µL of genomic DNA as template, 0.5 µM forward oligonucleotide (G-40598/ G-40599/ G-40600/ G-40604/ G-40605/ G-40606/ G-40606/ G-42015), 0.5 µM reverse oligonucleotide (G-40601/ G-40602/ G-40603/ G-40607/ G-40608/ G-40609/ G-42016), 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (#R0182)

and 0.02 U/µL Phusion High-Fidelity DNA polymerase (#F530) to a final volume of 25 µL.

The second PCR amplification consisted of 2.5 µL of the cleaned PCR product of the previous reaction, 0.5 µM forward oligonucleotide (G-40610), 0.25 µM reverse oligonucleotide that had one of the 96 indices(Lundberg et al. 2013), 1x Phusion HF buffer (1.5 mM $MgCl_2$), 0.2 mM dNTPs (ThermoFisher scientific, #R0182) and 0.02 U/µL Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific Inc.) to a final volume of 25 µL.

Sequencing libraries were also prepared using the Q5® High-Fidelity DNA polymerase (New England BioLabs® Inc., #M0491) in a final concentration of 0.02 U/µL along with 1x Q5 reaction buffer (2 mM $MgCl_2$). The rest of the reaction components (DNA template, dNTPs) remained the same.

The MJ Research PTC225 peltier or the BIO-RAD C1000 Touch™ thermal cyclers were used. The PCR programs had 15 cycles in which the denaturing temperature was 94°C for 30 s, followed by annealing at 60°C for 30 s, and extension at 72°C for 10 s for program 1, and 15 s for program 2. A final extension step was at 72°C for 2 minutes.

Bead clean up

For the generation of the amplicon libraries, two bead-based clean-up steps were carried out using SPRI beads (Sera-Mag™ Magnetic SpeedBeads™ (GE Healthcare No.:65152105050250)). The first PCR product was cleaned using a ratio of 1:0.9 and resuspended in 17 µL of ddH2O. The second PCR product was cleaned using a ratio of 1:0.9 for *ICS1* and 1:0.8 for *FLC*. The ratios of clean ups were chosen after optimisation.

Quant-iT™ PicoGreen® dsDNA assay

Amplicons were quantified using the Quant-iT™ PicoGreen® dsDNA assay. One µL of each amplicon was used according to the manufacturer's instructions for the quantification. The samples were prepared in black microplates with 96 well, F-bottom, non-binding (Item No.: 655906), and the TECAN infinite M200 PRO plate reader was used for all the measurements using the Magellan 7.2.

Pooling procedure

To roughly normalize samples when pooling, the DNA concentration of all samples in each 96 well plate was first measured fluorometrically (PicoGreen assay). First, all the 96 samples from each plate were pooled, creating subpools. From samples with concentrations less than half of the mean, 6 µL were taken. From samples with concentrations more than twice the mean, 1.5 µL was taken. For all other samples falling between these extremes, 3 µL was taken. After each plate was pooled in this way, the subpools representing entire plates were again measured fluorometrically (Qubit® assay) and pooled in an equimolar manner to create a final pool containing all samples.

Qubit® fluorometer analysis

The concentration of the subpools and the final pool were evaluated using the Qubit dsDNA-HS assay kit where the samples were prepared according to the manufacturer's instructions.

Bioanalyzer

Each pool was analyzed on the Agilent 2100 Bioanalyzer (Agilent Technologies) according to the manufacturer's instructions. DNA1000 chips were used for the amplicon libraries, and the High Sensitivity chip was used for the off-target cleavage investigation libraries.

Chapter 2

N. benthamiana infiltration

Agrobacterium tumefaciens (ASE) carrying pEF008 (Cas9 expression cassette) were grown overnight at 28°C in an orbital shaker (HTC Infors Multitron, Einsbach, Germany) at 200 rpm in 5 mL of LB containing antibiotics (kanamycin 50 µg/mL, spectinomycin 100 µg/mL, chloramphenicol 25 µg/mL and tetracyclin 5 µg/mL). The next day, 2 mL of the cultured cells were diluted into 50 mL of LB with appropriate antibiotics for another overnight incubation at the same conditions. After incubation, the cells were recovered by centrifugation at 2000 x g for 15 minutes. The cells were

resuspended in infiltration medium (10 mM $MgCl_2$, 10 mM MES, 150 uM acetosyringone) to an OD_{600} of 0.5, and further incubated for 3 hours at room temperature with gentle shaking. The suspension was used for manual infiltration using 1 mL needleless syringe into the abaxial side of 4- to 5-week-old *N. benthamiana* leaves of plants growing at 23^{o}C SD. As a control, leaves were infiltrated with uninoculated infiltration medium. The infiltrated plants were kept in the growth room for 2 more days, then the leaves were collected for protein extraction.

Protein extraction

Infiltrated *N. benthamiana* leaf tissue equivalent to 100 µL was collected and homogenised in 100 µL of CelLytic^{TM} P cell lysis buffer. The supernatant containing the protein extract was used for downstream analysis.

Western blot

Proteins were resuspended in 4 x Laemmli sample buffer (240 mM Tris-HCl pH 6.8, 8% SDS, 40% Glycerol, 5% B- mercaptoethanol, 0.04% Bromophenol blue), boiled for 10 min at 95°C, separated on a 7% SDS-PAGE (in a volume of 8 mL gel: 1.4 mL of 40% Acrylamide, 2 mL of 1.5 M Tris pH 8.8, 80 µL of 10% SDS, 80 µL of 10% APS, 8 µL of TEMED) and immunoblotted on PVDF membrane (Bio-Rad, Foster City, CA, USA). The membrane was incubated with anti-FLAG primary antibody (Sigma-Aldrich, #F1804, 2012) at the recommended dilution. As a secondary, the anti-mouse antibody was used (Millipore, #AP124p, 2012).

Incomplete block design

The lines used for this part of the study were randomised based on an incomplete block design. In total there were 8 replicates divided into 2 sets. Each line was represented once in each replicate with five plants. Two of these plants were used for the *Psm*4326 infections, one for measuring SA of unchallenged plants, one for evaluating flowering time phenotype and one as a back-up (Figure 6). For the generation of each set the agricolae R package was used (De Mendiburu & Simon 2015).

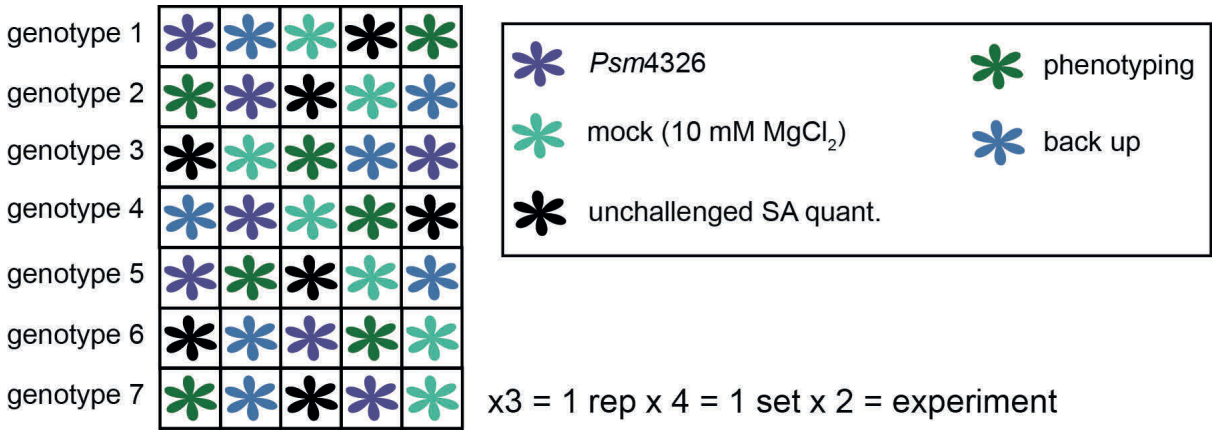


Figure 6: Schematic representation of the randomisation plan and the tray organisation of the experiment. Each rosette represents one plant and each row one genotype. One replicate consisted of three trays and each set of four replicates. There were two sets in total. From each genotype one plant was used for different experiments or measurements as noted at the box on the top right.

P. syringae pv. *maculicola* ES4326 infections

The *P. syringae* pv. *maculicola* ES4326 strain was streaked on KB medium agar plates with 25 mg/L streptomycin and and incubated at 30°C for 2 days. A single colony was selected and transferred to a liquid culture of 4 mL KB medium with 25 mg/L streptomycin for ~18 hours at 28°C. The next day, the cells were spun down at 4,000 x g for 8 minutes and the pellet was washed twice with 2 mL of 10 mM $MgCl_2$. Finally the pellet was resuspended in 500 µL of 10 mM $MgCl_2$ and the OD_{600} was adjusted at 0.001 which is equivalent to 10^6 cfu/mL.

The suspension was used for manual infiltration using 1 mL needle-less syringe into the abaxial side of 35 days old *A. thaliana* plants until leaves were saturated. As control, 10 mM $MgCl_2$ was used for infiltration. In total, 8 biological replicates were used for each treatment (*Psm*4326 or buffer) and there was a pair at each tray. Two days after inoculation, infiltrated leaf tissue was collected and processed for SA quantification.

H. arabidopsidis 14OHMLP04 infections

The isolate was provided by Dr Gautam Shirsekar and it was collected in Indiana, USA. Active spores were already available in the laboratory and they were propagated by infection of *eds1-1* mutant plants. Leaves of infected plants were collected and washed in ddH_2O, and spore counting was performed using a improved Bauer haemocytometer chamber. Spores were adjusted to 50,000 spores/mL and the

suspension was used for spray inoculation of 21 day old *A. thaliana* plants. In parallel, the same set of *A. thaliana* lines was spray inoculated with ddH_2O as control. Each set was kept in separate trays in order to avoid contamination.

Each line was represented by 4 biological replicates per treatment and each biological replicate consisted of 4-5 plants. After inoculation trays were kept in a Percival Arabidopsis growth cabinet (at 15^oC with 50% humidity) in the dark overnight. The next day, the trays were uncovered for the inoculum to dry. Seven days later, the humidity was increased by spraying the lids of the trays to promote sporulation. Twelve days post inoculation, plants were imaged for phenotyping and leaf tissue was collected for SA measurements. Thirteen days post inoculation, leaf tissue was collected for cell death and pathogen growth assess (Trypan blue staining).

Trypan blue staining

Leaf tissue harvested 13 days post inoculation was fully immersed in lacto-phenol/Trypan blue solution (10 mL lactic acid, 10 mL glycerol, 10 mL phenol, 10 mg Trypan blue and 10 mL water) in 24-well plate for staining. Plates were placed at a 70^oC incubator for 1 hour and then left at room temperature for 1 hour more. Next, the staining solution was aspirated out and replaced with chloral hydrate solution (2.5g/mL) for destaining the leaf tissues. Samples were kept at room temperature for ~16 hours, then the chloral hydrate was replaced with fresh solution and incubated for another day. Samples were kept in 50% glycerol for long-term storage or prior to fixation.

Statistical analysis

The packages used during the statistical analysis were:

- *lme4* (Bates et al. 2015),
- *MASS* (Anon n.d.),
- multcomp (Anon n.d.)

The flowering time phenotype data were fitted in a linear mixed effect model calculating the variance originate from the genetic background and the incomplete block design. Dunnett tests was carried out for calculating the significance between

wild type and corresponding *ics1* mutant line using the multcomp package (Anon n.d.).

flowering_time <- lmer(trait~ genotype+(1 | background)+(1 | set)+(1 | rep:set)+(1 | block:rep:set), data=$data)

For evaluating the effect of the genotype and treatment during pathogen infection experiment regarding SA accumulation only the measurements of the wild type plants were used. The quantified weight of SA was box-cox normalized and three different linear models were generated. Likelihood ratio tests were carried out between the full model and each one of the two reduced ones.

modelFull <-lmer(weightbc~treatment+(1|genotype), data = $data) (full model)
modelGenotype <-lmer(weightbc~(1|genotype), data = $data) (reduced 1)
modelTreatment<-lm(weightbc~treatment, data = $data) (reduced 2)

lrtest(modelFull,modelGenotype)
lrtest(modelFull,modelTreatment).

Chapter 3

Whole genome library preparation

Genomic DNA was extracted as described by (Karasov et al. 2018) and was quantified using Qubit. From each line (het 3_7, het 3_32, wt 4_7 and ctrl_t1) 20 plants were used for library preparation. 25 ng of DNA was pooled from each plant so each pool consisted of 500 ng of DNA. Covaris™ S2 (Covaris, Inc., USA) was used to shear the DNA using standard parameters (10% duty factor, 5 intensity, 200 cycles/burst, 45 seconds, sweeping frequency at 5-6°C). Each pool was cleaned using SPRI beads in 1:1.8 ratio prior to library preparation. The NEBNext® Ultra™ II DNA library prep kit for Illumina® (New England BioLabs® Inc., E7645) was used for preparing the libraries according to manufacturer's instruction. Each library was amplified for 6 cycles of 98°C for 10 seconds, 65°C for 1 minute and 15 seconds followed by a 5-minute extension at 65°C. Final libraries were evaluated using bioanalyzer and quantified using qubit. Final, libraries were loaded in two lanes of a HiSeq3000 Illumina platform (Illumina, USA).

Off-target identification pipeline

The reads were trimmed and filtered using the SKEWER (v0.2.2) software (-q 20, -l 30 -n) (Jiang et al. 2014). The processed data were aligned to the reference genome (TAIR10 (Arabidopsis Genome Initiative 2000) with the used construct added as an extra chromosome) using BOWTIE2 (v2.2.3) with default parameters (Arabidopsis Genome Initiative 2000; Langmead & Salzberg 2012). Only mapped reads with mapping quality above 30 were used for the downstream analysis using Samtools (v1.3.1) (Li et al. 2009). Additionally, the PCR duplicates were removed using the PICARD (v2.2.1) algorithm (http://broadinstitute.github.io/picard). Furthermore, I used the GATK (v3.5) software (DePristo et al. 2011; McKenna et al. 2010) and specifically the RealignerTargetCreator and IndelRealigner functions for improving the alignments around positions with insertions and deletions (InDels). Samtools depth was used for estimating the average number of reads covering a base pair in the genome.

FREEBAYES (v1.1.0) was used for variant calling using default parameters (Garrison & Marth 2012b). The variant calling format (VCF) files of the samples were intersected using the vcflib package, written by the same provider, with VCF file of the control sample for identification of the unique variants (https://github.com/vcflib/vcflib). The remaining unique variants were filtered using vcflib filter option (vcffilter) for high quality. The filtering parameters concerned the number of reads covering a region (DP; 30 < DP < 90), the mean mapping quality of observed variant (MQM; MQM > 30), the probability of observing a variant in the specific region (QUAL; QUAL > 30), the number of reads supporting a variant (AO; QUAL / AO > 10) and the number of reads covering the variant further away from each end of the read (RPR and RPL; RPR > 1 and RPL >1).

A Browser Extensible Data (BED) file, generated using the RepeatMasker software with default parameters, was used for removing variants in repeat regions (Smit, AFA et al. 2013-2015). Finally, the strand bias was calculated based on (Guo et al. 2012) to remove affected variants.

Unique variant evaluation

The unique variants were verified by PCR amplification of the surrounding region and Sanger sequencing of 6 or 20 samples. The PCR programs had 42 cycles in

which the denaturing temperature was 98°C for 30 seconds, followed by annealing at 55°C for 30 seconds (for deletion$_{1,2,3,4}$ and SNP$_{2,3}$) or 60°C for 30 seconds (for SNP$_1$ and the on-target variants), and extension at 72°C for 45 seconds. A final extension step was at 72°C for 3 minutes. Next, PCR fragments were Sanger sequenced and the genotype of the respective variant was evaluated by examining the peaks of the chromatograph.

In silico prediction tools

CCTop (Stemmer et al. 2015)

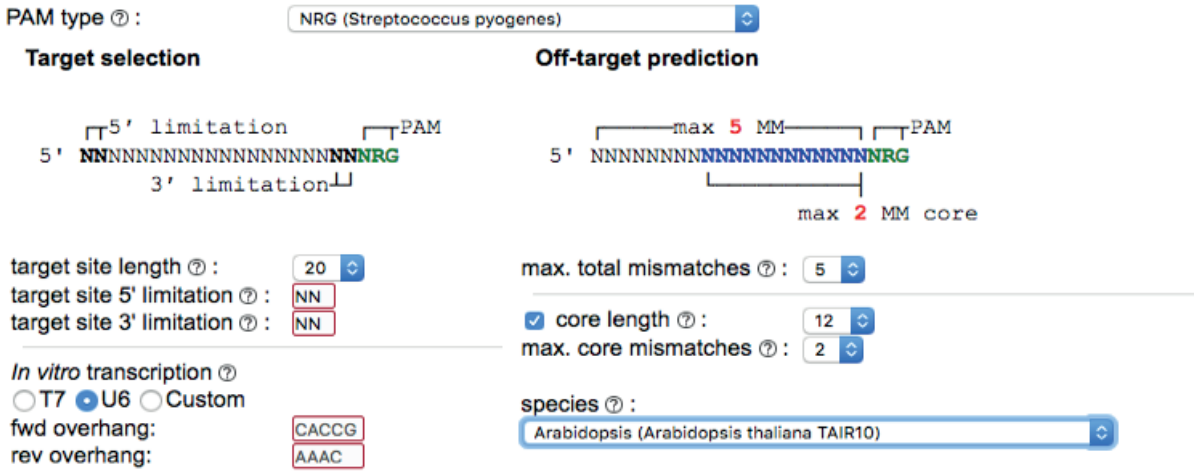


Figure 7: Screenshot of the specifications used for predicting of off-target cleavage sites with CCTop (Stemmer et al. 2015)**.** The software was accessed January 2018.

CRISPR-P 2.0 and CRISPR design (Liu et al. 2017); http://crispr.mit.edu/(Liu et al. 2017)

After selecting the appropriate organism as input, default parameters were used for predicting off-target cleavage sites for the two sgRNAs.

CHOPCHOP v2 (Labun et al. 2016)

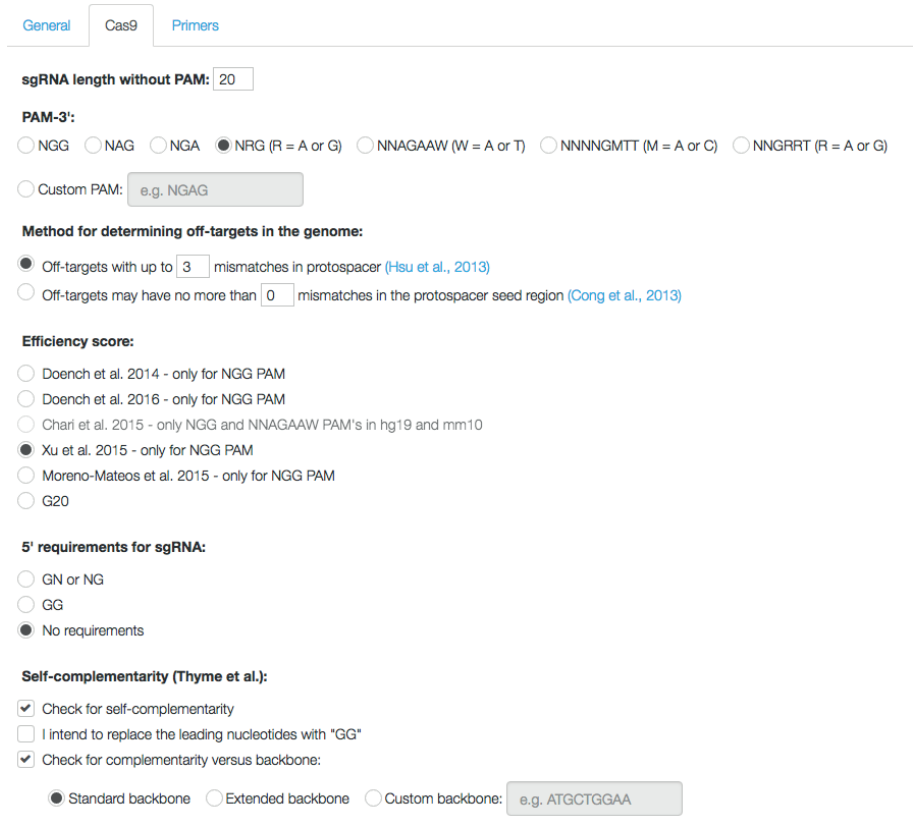


Figure 8: Screenshot of the specifications used for predicting of off-target cleavage sites with CHOPCHOP v2 (Labun et al. 2016)**.** The software was accessed January 2018.

Cas-OFFinder (Bae et al. 2014)

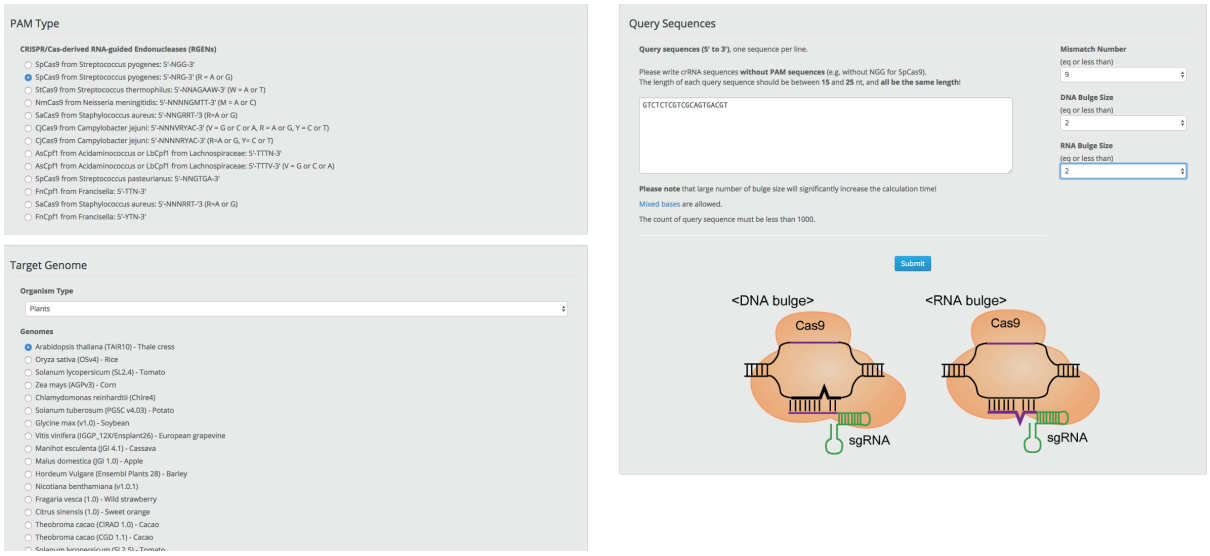


Figure 9: Screenshot of the specifications used for predicting of off-target cleavage sites with Cas-OFFinder (Bae et al. 2014)**.** The software was accessed January 2018.

References

Ablain, Julien, Ellen M. Durand, Song Yang, Yi Zhou, and Leonard I. Zon. 2015. “A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish.” *Developmental Cell* 32 (6): 756–64.

Agrios, G.N., 1988. Plant Pathology. Academic Press, San Diego, CA.

Alonso-Blanco, Carlos, Mark G. M. Aarts, Leonie Bentsink, Joost J. B. Keurentjes, Matthieu Reymond, Dick Vreugdenhil, and Maarten Koornneef. 2009. “What Has Natural Variation Taught Us about Plant Development, Physiology, and Adaptation?” *The Plant Cell* 21 (7): 1877–96.

Arabidopsis Genome Initiative. 2000. “Analysis of the Genome Sequence of the Flowering Plant Arabidopsis Thaliana.” *Nature* 408 (6814): 796–815.

Arnould, Sylvain, Christophe Perez, Jean-Pierre Cabaniols, Julianne Smith, Agnès Gouble, Sylvestre Grizot, Jean-Charles Epinat, Aymeric Duclert, Philippe Duchateau, and Frédéric Pâques. 2007. “Engineered I-CreI Derivatives Cleaving Sequences from the Human XPC Gene Can Induce Highly Efficient Gene Correction in Mammalian Cells.” *Journal of Molecular Biology* 371 (1): 49–65.

Ashworth, Justin, Gregory K. Taylor, James J. Havranek, S. Arshiya Quadri, Barry L. Stoddard, and David Baker. 2010. “Computational Reprogramming of Homing Endonuclease Specificity at Multiple Adjacent Base Pairs.” *Nucleic Acids Research* 38 (16): 5601–8.

Attaran, Elham, Tatiana E. Zeier, Thomas Griebel, and Jürgen Zeier. 2009. “Methyl Salicylate Production and Jasmonate Signaling Are Not Essential for Systemic Acquired Resistance in Arabidopsis.” *The Plant Cell* 21 (3): 954–71.

Auer, Thomas O., Karine Duroure, Anne De Cian, Jean-Paul Concordet, and Filippo Del Bene. 2014. “Highly Efficient CRISPR/Cas9-Mediated Knock-in in Zebrafish by Homology-Independent DNA Repair.” *Genome Research* 24 (1): 142–53.

Baek, Dongwon, Prasanth Pathange, Jung-Sung Chung, Jiafu Jiang, Liqiong Gao, Akira Oikawa, Masami Yokota Hirai, Kazuki Saito, Paul W. Pare, and Huazhong Shi. 2010. “A Stress-Inducible Sulphotransferase Sulphonates Salicylic Acid and Confers Pathogen Resistance in Arabidopsis.” *Plant, Cell & Environment* 33 (8): 1383–92.

Bae, Sangsu, Jeongbin Park, and Jin-Soo Kim. 2014. “Cas-OFFinder: A Fast and Versatile Algorithm That Searches for Potential off-Target Sites of Cas9 RNA-Guided Endonucleases.” *Bioinformatics* 30 (10): 1473–75.

Barrangou, Rodolphe, Christophe Fremaux, Hélène Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. 2007. “CRISPR Provides Acquired Resistance against Viruses in Prokaryotes.” *Science* 315 (5819): 1709–12.

Bartsch, Michael, Paweł Bednarek, Pedro D. Vivancos, Bernd Schneider, Edda von Roepenack-Lahaye, Christine H. Foyer, Erich Kombrink, Dierk Scheel, and Jane E. Parker. 2010. “Accumulation of Isochorismate-Derived 2, 3-Dihydroxybenzoic 3-O-β-D-Xyloside in Arabidopsis Resistance to Pathogens and Ageing of Leaves.” *The Journal of Biological Chemistry* 285 (33): 25654–65.

Bartsch, Michael, Enrico Gobbato, Pawel Bednarek, Svenja Debey, Joachim L. Schultze, Jaqueline Bautor, and Jane E. Parker. 2006. “Salicylic Acid-Independent ENHANCED DISEASE SUSCEPTIBILITY1 Signaling in Arabidopsis Immunity and Cell Death Is Regulated by the Monooxygenase FMO1 and the Nudix Hydrolase NUDT7.” *The Plant Cell* 18 (4): 1038–51.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2014. “Fitting Linear Mixed-Effects Models Using lme4.” *arXiv [stat.CO]*. arXiv. http://arxiv.org/abs/1406.5823.

———. 2015. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software, Articles* 67 (1): 1–48.

Bauer, Daniel E., Matthew C. Canver, and Stuart H. Orkin. 2015. “Generation of Genomic Deletions in Mammalian Cell Lines via CRISPR/Cas9.” *Journal of Visualized Experiments: JoVE*, no. 95 (January): e52118.

Belhaj, Khaoula, Angela Chaparro-Garcia, Sophien Kamoun, and Vladimir Nekrasov. 2013. “Plant Genome Editing Made Easy: Targeted Mutagenesis in Model and Crop Plants Using the CRISPR/Cas System.” *Plant Methods* 9 (1): 39.

Bellés, José María, Rafael Garro, Joaquín Fayos, Pilar Navarro, Jaime Primo, and Vicente Conejero. 1999. “Gentisic Acid As a Pathogen-Inducible Signal, Additional to Salicylic Acid for Activation of Plant Defenses in Tomato.” *Molecular Plant-Microbe Interactions: MPMI* 12 (3): 227–35.

Bellés, José M., Rafael Garro, Vicente Pallás, Joaquín Fayos, Ismael Rodrigo, and Vicente Conejero. 2006. “Accumulation of Gentisic Acid as Associated with Systemic Infections but Not with the Hypersensitive Response in Plant-Pathogen Interactions.” *Planta* 223 (3): 500–511.

Betge, Johannes, Grainne Kerr, Thilo Miersch, Svenja Leible, Gerrit Erdmann, Christian L. Galata, Tianzuo Zhan, et al. 2015. “Amplicon Sequencing of Colorectal Cancer: Variant Calling in Frozen and Formalin-Fixed Samples.” *PloS One* 10 (5): e0127146.

Bibikova, Marina, Kelly Beumer, Jonathan K. Trautman, and Dana Carroll. 2003. “Enhancing Gene Targeting with Designed Zinc Finger Nucleases.” *Science* 300 (5620): 764.

Bibikova, Marina, Mary Golic, Kent G. Golic, and Dana Carroll. 2002. “Targeted Chromosomal Cleavage and Mutagenesis in Drosophila Using Zinc-Finger Nucleases.” *Genetics* 161 (3): 1169–75.

Bibikova, M., D. Carroll, D. J. Segal, J. K. Trautman, J. Smith, Y. G. Kim, and S. Chandrasegaran. 2001. “Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.” *Molecular and Cellular Biology* 21 (1): 289–97.

Bitinaite, J., D. A. Wah, A. K. Aggarwal, and I. Schildkraut. 1998. “FokI Dimerization Is Required for DNA Cleavage.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (18): 10570–75.

Boch, Jens, and Ulla Bonas. 2010. “Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function.” *Annual Review of Phytopathology* 48: 419–36.

Boch, Jens, Heidi Scholze, Sebastian Schornack, Angelika Landgraf, Simone Hahn, Sabine Kay, Thomas Lahaye, Anja Nickstadt, and Ulla Bonas. 2009. “Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors.” *Science* 326 (5959): 1509–12.

Boel, Annekatrien, Woutert Steyaert, Nina De Rocker, Björn Menten, Bert Callewaert, Anne De Paepe, Paul Coucke, and Andy Willaert. 2016. “BATCH-GE: Batch Analysis of Next-Generation Sequencing Data for Genome Editing Assessment.” *Scientific Reports* 6 (July): 30330.

Boller, Thomas, and Georg Felix. 2009. “A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors.” *Annual Review of Plant Biology* 60: 379–406.

Bolotin, Alexander, Benoit Quinquis, Alexei Sorokin, and S. Dusko Ehrlich. 2005. “Clustered Regularly Interspaced Short Palindrome Repeats (CRISPRs) Have Spacers of Extrachromosomal Origin.” *Microbiology* 151 (Pt 8): 2551–61.

Bonas, Ulla, Robert E. Stall, and Brian Staskawicz. 1989. “Genetic and Structural Characterization of the Avirulence Gene avrBs3 from Xanthomonas Campestris Pv. Vesicatoria.” *Molecular & General Genetics: MGG* 218 (1): 127–36.

Bouchez, D. (Institut National de la Recherche Agronomique, Versailles (France). Centre de Versailles, Biologie Cellulaire), C. Camilleri, and M. Caboche. oct1993. “A Binary Vector Based on Basta Resistance for in Planta Transformation of Arabidopsis Thaliana [T-DNA, Gene-Tagging].” *Comptes Rendus de l’Academie Des Sciences Serie 3 Sciences de La Vie (France)* 316. http://agris.fao.org/agris-search/search.do?recordID=FR9307336.

Bressan, Raul Bardini, Pooran Singh Dewari, Maria Kalantzaki, Ester Gangoso, Mantas Matjusaitis, Claudia Garcia-Diaz, Carla Blin, et al. 2017. “Efficient CRISPR/Cas9-Assisted Gene Targeting Enables Rapid and Precise Genetic Manipulation of Mammalian Neural Stem Cells.” *Development* 144 (4): 635–48.

Brocal, Isabel, Richard J. White, Christopher M. Dooley, Samantha N. Carruthers, Richard Clark, Amanda Hall, Elisabeth M. Busch-Nentwich, Derek L. Stemple, and Ross N. W. Kettleborough. 2016. “Efficient Identification of CRISPR/Cas9-Induced Insertions/deletions by Direct Germline Screening in Zebrafish.” *BMC Genomics* 17 (March): 259.

Brouns, Stan J. J., Matthijs M. Jore, Magnus Lundgren, Edze R. Westra, Rik J. H. Slijkhuis, Ambrosius P. L. Snijders, Mark J. Dickman, Kira S. Makarova, Eugene V. Koonin, and John van der Oost. 2008. “Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes.” *Science* 321 (5891): 960–64.

Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, et al. 1996. “Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus Jannaschii.” *Science* 273 (5278): 1058–73.

Canver, Matthew C., Daniel E. Bauer, Abhishek Dass, Yvette Y. Yien, Jacky Chung, Takeshi Masuda, Takahiro Maeda, Barry H. Paw, and Stuart H. Orkin. 2014. “Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells.” *The Journal of Biological Chemistry* 289 (31): 21312–24.

Cao, H., J. Glazebrook, J. D. Clarke, S. Volko, and X. Dong. 1997. “The Arabidopsis NPR1 Gene That Controls Systemic Acquired Resistance Encodes a Novel Protein Containing Ankyrin Repeats.” *Cell* 88 (1): 57–63.

Cao, Jun, Korbinian Schneeberger, Stephan Ossowski, Torsten Günther, Sebastian Bender, Joffrey Fitz, Daniel Koenig, et al. 2011. “Whole-Genome Sequencing of Multiple Arabidopsis Thaliana Populations.” *Nature Genetics* 43 (10): 956–63.

Carbery, Iara D., Diana Ji, Anne Harrington, Victoria Brown, Edward J. Weinstein, Lucy Liaw, and Xiaoxia Cui. 2010. “Targeted Genome Modification in Mice Using Zinc-Finger Nucleases.” *Genetics* 186 (2): 451–59.

Chae, Eunyoung, Kirsten Bomblies, Sang-Tae Kim, Darya Karelina, Maricris Zaidem, Stephan Ossowski, Carmen Martín-Pizarro, et al. 2014. “Species-Wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis.” *Cell* 159 (6): 1341–51.

Chang, Nannan, Changhong Sun, Lu Gao, Dan Zhu, Xiufei Xu, Xiaojun Zhu, Jing-Wei Xiong, and Jianzhong Jeff Xi. 2013. “Genome Editing with RNA-Guided Cas9 Nuclease in Zebrafish Embryos.” *Cell Research* 23 (4): 465–72.

Chapman, Karen M., Gerardo A. Medrano, Priscilla Jaichander, Jaideep Chaudhary, Alexandra E. Waits, Marcelo A. Nobrega, James M. Hotaling, Carole Ober, and F. Kent Hamra. 2015. “Targeted Germline Modifications in Rats Using CRISPR/Cas9 and Spermatogonial Stem Cells.” *Cell Reports* 10 (11): 1828–35.

Chen, Baohui, Luke A. Gilbert, Beth A. Cimini, Joerg Schnitzbauer, Wei Zhang, Gene-Wei Li, Jason Park, et al. 2013. “Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System.” *Cell* 155 (7): 1479–91.

Chen, Feng, John C. D’Auria, Dorothea Tholl, Jeannine R. Ross, Jonathan Gershenzon, Joseph P. Noel, and Eran Pichersky. 2003. “An Arabidopsis Thaliana Gene for Methylsalicylate Biosynthesis, Identified by a Biochemical Genomics Approach, Has a Role in Defense.” *The Plant Journal: For Cell and Molecular Biology* 36 (5): 577–88.

Chen, Janice S., Yavuz S. Dagdas, Benjamin P. Kleinstiver, Moira M. Welch, Alexander A. Sousa, Lucas B. Harrington, Samuel H. Sternberg, J. Keith Joung, Ahmet Yildiz, and Jennifer A. Doudna. 2017. “Enhanced Proofreading Governs CRISPR-Cas9 Targeting Accuracy.” *Nature*, September. https://doi.org/10.1038/nature24268.

Chen, Jiafa, Luyan Zhang, Songtao Liu, Zhimin Li, Rongrong Huang, Yongming Li, Hongliang Cheng, et al. 2016. “The Genetic Basis of Natural Variation in Kernel Size

and Related Traits Using a Four-Way Cross Population in Maize." *PloS One* 11 (4): e0153428.
Chen, Xinwei, Andrew Steed, Clare Harden, and Paul Nicholson. 2006. "Characterization of Arabidopsis Thaliana-Fusarium Graminearum Interactions and Identification of Variation in Resistance among Ecotypes." *Molecular Plant Pathology* 7 (5): 391–403.
Chen, Ying, Hui Shen, Muyang Wang, Qun Li, and Zuhua He. 2013. "Salicyloyl-Aspartate Synthesized by the Acetyl-Amido Synthetase GH3.5 Is a Potential Activator of Plant Immunity in Arabidopsis." *Acta Biochimica et Biophysica Sinica* 45 (10): 827–36.
Christian, Michelle, Tomas Cermak, Erin L. Doyle, Clarice Schmidt, Feng Zhang, Aaron Hummel, Adam J. Bogdanove, and Daniel F. Voytas. 2010. "Targeting DNA Double-Strand Breaks with TAL Effector Nucleases." *Genetics* 186 (2): 757–61.
Christian, Michelle, Yiping Qi, Yong Zhang, and Daniel F. Voytas. 2013. "Targeted Mutagenesis of Arabidopsis Thaliana Using Engineered TAL Effector Nucleases." *G3* 3 (10): 1697–1705.
Clarke, Shannon M., Luis A. J. Mur, Jacqueline E. Wood, and Ian M. Scott. 2004. "Salicylic Acid Dependent Signaling Promotes Basal Thermotolerance but Is Not Essential for Acquired Thermotolerance in Arabidopsis Thaliana." *The Plant Journal: For Cell and Molecular Biology* 38 (3): 432–47.
Cleland, C. F., and A. Ajami. 1974. "Identification of the Flower-Inducing Factor Isolated from Aphid Honeydew as Being Salicylic Acid." *Plant Physiology* 54 (6): 904–6.
Clough, Steven J., and Andrew F. Bent. 1998. "Floral Dip: A Simplified Method forAgrobacterium-Mediated Transformation ofArabidopsis Thaliana." *The Plant Journal: For Cell and Molecular Biology* 16 (6): 735–43.
Coates, Mary E., and Jim L. Beynon. 2010. "Hyaloperonospora Arabidopsidis as a Pathogen Model." *Annual Review of Phytopathology* 48: 329–45.
Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *Science* 339 (6121): 819–23.
Cong, Le, Ruhong Zhou, Yu-Chi Kuo, Margaret Cunniff, and Feng Zhang. 2012. "Comprehensive Interrogation of Natural TALE DNA-Binding Modules and Transcriptional Repressor Domains." *Nature Communications* 3 (July): 968.
Coquoz, J. L., A. Buchala, and J. P. Metraux. 1998. "The Biosynthesis of Salicylic Acid in Potato Plants." *Plant Physiology* 117 (3): 1095–1101.
Crosetto, Nicola, Abhishek Mitra, Maria Joao Silva, Magda Bienko, Norbert Dojer, Qi Wang, Elif Karaca, et al. 2013. "Nucleotide-Resolution DNA Double-Strand Break Mapping by next-Generation Sequencing." *Nature Methods* 10 (4): 361–65.
Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27 (15): 2156–58.
Dangl, J. L., and J. D. Jones. 2001. "Plant Pathogens and Integrated Defence Responses to Infection." *Nature* 411 (6839): 826–33.
Dangl, J. L., C. Ritter, M. J. Gibbon, L. A. Mur, J. R. Wood, S. Goss, J. Mansfield, J. D. Taylor, and A. Vivian. 1992. "Functional Homologs of the Arabidopsis RPM1 Disease Resistance Gene in Bean and Pea." *The Plant Cell* 4 (11): 1359–69.
Dean, John V., and Sean P. Delaney. 2008. "Metabolism of Salicylic Acid in Wild-Type, ugt74f1 and ugt74f2 Glucosyltransferase Mutants of Arabidopsis Thaliana." *Physiologia Plantarum* 132 (4): 417–25.
Dean, John V., Leila A. Mohammed, and Terry Fitzpatrick. 2005. "The Formation, Vacuolar Localization, and Tonoplast Transport of Salicylic Acid Glucose Conjugates in Tobacco Cell Suspension Cultures." *Planta* 221 (2): 287–96.
Dean, John V., Reena P. Shah, and Leila A. Mohammed. 2003. "Formation and Vacuolar Localization of Salicylic Acid Glucose Conjugates in Soybean Cell Suspension Cultures." *Physiologia Plantarum* 118 (3): 328–36.

Defraia, Christopher T., Eric A. Schmelz, and Zhonglin Mou. 2008. “A Rapid Biosensor-Based Method for Quantification of Free and Glucose-Conjugated Salicylic Acid.” *Plant Methods* 4 (December): 28.

DeKelver, Russell C., Vivian M. Choi, Erica A. Moehle, David E. Paschon, Dirk Hockemeyer, Sebastiaan H. Meijsing, Yasemin Sancak, et al. 2010. “Functional Genomics, Proteomics, and Regulatory DNA Analysis in Isogenic Settings Using Zinc Finger Nuclease-Driven Transgenesis into a Safe Harbor Locus in the Human Genome.” *Genome Research* 20 (8): 1133–42.

Delaney, T. P., S. Uknes, B. Vernooij, L. Friedrich, K. Weymann, D. Negrotto, T. Gaffney, et al. 1994. “A Central Role of Salicylic Acid in Plant Disease Resistance.” *Science* 266 (5188): 1247–50.

Delle-Vedove, Roxane, Bertrand Schatz, and Mathilde Dufay. 2017. “Understanding Intraspecific Variation of Floral Scent in Light of Evolutionary Ecology.” *Annals of Botany* 120 (1): 1–20.

Deltcheva, Elitza, Krzysztof Chylinski, Cynthia M. Sharma, Karine Gonzales, Yanjie Chao, Zaid A. Pirzada, Maria R. Eckert, Jörg Vogel, and Emmanuelle Charpentier. 2011. “CRISPR RNA Maturation by Trans-Encoded Small RNA and Host Factor RNase III.” *Nature* 471 (7340): 602–7.

De Mendiburu, Felipe, and Reinhard Simon. 2015. “Agricolae - Ten Years of an Open Source Statistical Tool for Experiments in Breeding, Agriculture and Biology.” e1748. PeerJ PrePrints. https://doi.org/10.7287/peerj.preprints.1404v1.

Dempsey, D’maris Amick, Jyoti Shah, and Daniel F. Klessig. 1999. “Salicylic Acid and Disease Resistance in Plants.” *Critical Reviews in Plant Sciences* 18 (4): 547–75.

Dempsey, D’maris Amick, A. Corina Vlot, Mary C. Wildermuth, and Daniel F. Klessig. 2011. “Salicylic Acid Biosynthesis and Metabolism.” *The Arabidopsis Book / American Society of Plant Biologists* 9 (December): e0156.

DePristo, Mark A., Eric Banks, Ryan Poplin, Kiran V. Garimella, Jared R. Maguire, Christopher Hartl, Anthony A. Philippakis, et al. 2011. “A Framework for Variation Discovery and Genotyping Using next-Generation DNA Sequencing Data.” *Nature Genetics* 43 (5): 491–98.

Desjarlais, J. R., and J. M. Berg. 1992. “Toward Rules Relating Zinc Finger Protein Sequences and DNA Binding Site Preferences.” *Proceedings of the National Academy of Sciences of the United States of America* 89 (16): 7345–49.

———. 1993. “Use of a Zinc-Finger Consensus Sequence Framework and Specificity Rules to Design Specific DNA Binding Proteins.” *Proceedings of the National Academy of Sciences of the United States of America* 90 (6): 2256–60.

———. 1994. “Length-Encoded Multiplex Binding Site Determination: Application to Zinc Finger Proteins.” *Proceedings of the National Academy of Sciences of the United States of America* 91 (23): 11099–103.

Deveau, Hélène, Rodolphe Barrangou, Josiane E. Garneau, Jessica Labonté, Christophe Fremaux, Patrick Boyaval, Dennis A. Romero, Philippe Horvath, and Sylvain Moineau. 2008. “Phage Response to CRISPR-Encoded Resistance in Streptococcus Thermophilus.” *Journal of Bacteriology* 190 (4): 1390–1400.

Dewdney, J., T. L. Reuber, M. C. Wildermuth, A. Devoto, J. Cui, L. M. Stutius, E. P. Drummond, and F. M. Ausubel. 2000. “Three Unique Mutants of Arabidopsis Identify Eds Loci Required for Limiting Growth of a Biotrophic Fungal Pathogen.” *The Plant Journal: For Cell and Molecular Biology* 24 (2): 205–18.

Dong, X., M. Mindrinos, K. R. Davis, and F. M. Ausubel. 1991. “Induction of Arabidopsis Defense Genes by Virulent and Avirulent Pseudomonas Syringae Strains and by a Cloned Avirulence Gene.” *The Plant Cell* 3 (1): 61–72.

Doyon, Yannick, Jasmine M. McCammon, Jeffrey C. Miller, Farhoud Faraji, Catherine Ngo, George E. Katibah, Rainier Amora, et al. 2008. “Heritable Targeted Gene Disruption in Zebrafish Using Designed Zinc-Finger Nucleases.” *Nature Biotechnology* 26 (6): 702–8.

Doyon, Yannick, Thuy D. Vo, Matthew C. Mendel, Shon G. Greenberg, Jianbin Wang, Danny F. Xia, Jeffrey C. Miller, Fyodor D. Urnov, Philip D. Gregory, and Michael C. Holmes. 2011. “Enhancing Zinc-Finger-Nuclease Activity with Improved Obligate Heterodimeric Architectures.” *Nature Methods* 8 (1): 74–79.

Edwards, K., C. Johnstone, and C. Thompson. 1991. “A Simple and Rapid Method for the Preparation of Plant Genomic DNA for PCR Analysis.” *Nucleic Acids Research* 19 (6): 1349.

Endo, M., M. Mikami, and S. Toki. 2014. “Multi-Gene Knockout Utilizing off-Target Mutations of the CRISPR/Cas9 System in Rice.” *Plant & Cell Physiology*. http://pcp.oxfordjournals.org/content/early/2014/11/11/pcp.pcu154.short.

Epinat, Jean-Charles, Sylvain Arnould, Patrick Chames, Pascal Rochaix, Dominique Desfontaines, Clémence Puzin, Amélie Patin, Alexandre Zanghellini, Frédéric Pâques, and Emmanuel Lacroix. 2003. “A Novel Engineered Meganuclease Induces Homologous Recombination in Yeast and Mammalian Cells.” *Nucleic Acids Research* 31 (11): 2952–62.

Falk, A., B. J. Feys, L. N. Frost, J. D. Jones, M. J. Daniels, and J. E. Parker. 1999. “EDS1, an Essential Component of R Gene-Mediated Disease Resistance in Arabidopsis Has Homology to Eukaryotic Lipases.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (6): 3292–97.

Fauser, Friedrich, Simon Schiml, and Holger Puchta. 2014. “Both CRISPR/Cas-Based Nucleases and Nickases Can Be Used Efficiently for Genome Engineering in Arabidopsis Thaliana.” *The Plant Journal: For Cell and Molecular Biology* 79 (2): 348–59.

Feng, Zhengyan, Yanfei Mao, Nanfei Xu, Botao Zhang, Pengliang Wei, Dong-Lei Yang, Zhen Wang, et al. 2014. “Multigeneration Analysis Reveals the Inheritance, Specificity, and Patterns of CRISPR/Cas-Induced Gene Modifications in Arabidopsis.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (12): 4632–37.

Feng, Zhengyan, Botao Zhang, Wona Ding, Xiaodong Liu, Dong-Lei Yang, Pengliang Wei, Fengqiu Cao, et al. 2013. “Efficient Genome Editing in Plants Using a CRISPR/Cas System.” *Cell Research* 23 (10): 1229–32.

Fenster, Charles B., and Kermit Ritland. 1994. “Evidence for Natural Selection on Mating System in Mimulus (Scrophulariaceae).” *International Journal of Plant Sciences* 155 (5): 588–96.

Feys, B. J., L. J. Moisan, M. A. Newman, and J. E. Parker. 2001. “Direct Interaction between the Arabidopsis Disease Resistance Signaling Proteins, EDS1 and PAD4.” *The EMBO Journal* 20 (19): 5400–5411.

Fister, Andrew S., Lena Landherr, Siela N. Maximova, and Mark J. Guiltinan. 2018. “Transient Expression of CRISPR/Cas9 Machinery Targeting TcNPR3 Enhances Defense Response in Theobroma Cacao.” *Frontiers in Plant Science* 9 (March): 268.

Flamand, Marie Christine, Jean Pierre Goblet, Gérard Duc, Michel Briquet, and Marc Boutry. 1992. “Sequence and Transcription Analysis of Mitochondrial Plasmids Isolated from Cytoplasmic Male-Sterile Lines of Vicia Faba.” *Plant Molecular Biology* 19 (6): 913–23.

Flor, H. H. 1956. “The Complementary Genic Systems in Flax and Flax Rust**Joint Contribution from the Field Crops Research Branch, Agricultural Research Service, United States Department of Agriculture and the North Dakota Agricultural Experiment Station.” In *Advances in Genetics*, edited by M. Demerec, 8:29–54. Academic Press.

Forner, Joachim, Anne Pfeiffer, Tobias Langenecker, Pablo A. Manavella, Pablo Manavella, and Jan U. Lohmann. 2015. “Germline-Transmitted Genome Editing in Arabidopsis Thaliana Using TAL-Effector-Nucleases.” *PloS One* 10 (3): e0121056.

Forouhar, Farhad, Yue Yang, Dhirendra Kumar, Yang Chen, Eyal Fridman, Sang Wook Park, Yiwen Chiang, et al. 2005. “Structural and Biochemical Studies Identify Tobacco SABP2 as a Methyl Salicylate Esterase and Implicate It in Plant Innate Immunity.”

Proceedings of the National Academy of Sciences of the United States of America 102 (5): 1773–78.
Friedland, Ari E., Yonatan B. Tzur, Kevin M. Esvelt, Monica P. Colaiácovo, George M. Church, and John A. Calarco. 2013. "Heritable Genome Editing in C. Elegans via a CRISPR-Cas9 System." *Nature Methods* 10 (8): 741–43.
Frock, Richard L., Jiazhi Hu, Robin M. Meyers, Yu-Jui Ho, Erina Kii, and Frederick W. Alt. 2015. "Genome-Wide Detection of DNA Double-Stranded Breaks Induced by Engineered Nucleases." *Nature Biotechnology* 33 (2): 179–86.
Fu, Yanfang, Jennifer A. Foden, Cyd Khayter, Morgan L. Maeder, Deepak Reyon, J. Keith Joung, and Jeffry D. Sander. 2013. "High-Frequency off-Target Mutagenesis Induced by CRISPR-Cas Nucleases in Human Cells." *Nature Biotechnology* 31 (9): 822–26.
Fu, Zheng Qing, Shunping Yan, Abdelaty Saleh, Wei Wang, James Ruble, Nodoka Oka, Rajinikanth Mohan, et al. 2012. "NPR3 and NPR4 Are Receptors for the Immune Signal Salicylic Acid in Plants." *Nature* 486 (7402): 228–32.
Gabriel, Richard, Angelo Lombardo, Anne Arens, Jeffrey C. Miller, Pietro Genovese, Christine Kaeppel, Ali Nowrouzi, et al. 2011. "An Unbiased Genome-Wide Analysis of Zinc-Finger Nuclease Specificity." *Nature Biotechnology* 29 (9): 816–23.
Gaffney, T., L. Friedrich, B. Vernooij, D. Negrotto, G. Nye, S. Uknes, E. Ward, H. Kessmann, and J. Ryals. 1993. "Requirement of Salicylic Acid for the Induction of Systemic Acquired Resistance." *Science* 261 (5122): 754–56.
Gaj, Thomas, Charles A. Gersbach, and Carlos F. Barbas 3rd. 2013. "ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering." *Trends in Biotechnology* 31 (7): 397–405.
Galen, Candace. 1999. "Why Do Flowers Vary?The Functional Ecology of Variation in Flower Size and Form within Natural Plant Populations." *Bioscience* 49 (8): 631–40.
Gao, Qing-Ming, Shifeng Zhu, Pradeep Kachroo, and Aardra Kachroo. 2015. "Signal Regulators of Systemic Acquired Resistance." *Frontiers in Plant Science* 6 (April): 228.
Gao, Xiuhua, Jilin Chen, Xinhua Dai, Da Zhang, and Yunde Zhao. 2016. "An Effective Strategy for Reliably Isolating Heritable and Cas9-Free Arabidopsis Mutants Generated by CRISPR/Cas9-Mediated Genome Editing." *Plant Physiology* 171 (3): 1794–1800.
Garcion, Christophe, Antje Lohmann, Elisabeth Lamodière, Jérémy Catinot, Antony Buchala, Peter Doermann, and Jean-Pierre Métraux. 2008. "Characterization and Biological Function of the ISOCHORISMATE SYNTHASE2 Gene of Arabidopsis." *Plant Physiology* 147 (3): 1279–87.
Garrison, Erik, and Gabor Marth. 2012a. "Haplotype-Based Variant Detection from Short-Read Sequencing." *arXiv [q-bio.GN]*. arXiv. http://arxiv.org/abs/1207.3907.
———. 2012b. "Haplotype-Based Variant Detection from Short-Read Sequencing." *arXiv [q-bio.GN]*. arXiv. http://arxiv.org/abs/1207.3907.
Gasiunas, Giedrius, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. 2012. "Cas9-crRNA Ribonucleoprotein Complex Mediates Specific DNA Cleavage for Adaptive Immunity in Bacteria." *Proceedings of the National Academy of Sciences of the United States of America* 109 (39): E2579–86.
Geurts, Aron M., Gregory J. Cost, Yevgeniy Freyvert, Bryan Zeitler, Jeffrey C. Miller, Vivian M. Choi, Shirin S. Jenkins, et al. 2009. "Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases." *Science* 325 (5939): 433.
Gilbert, Luke A., Matthew H. Larson, Leonardo Morsut, Zairan Liu, Gloria A. Brar, Sandra E. Torres, Noam Stern-Ginossar, et al. 2013. "CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes." *Cell* 154 (2): 442–51.
Glazebrook, Jane. 2005. "Contrasting Mechanisms of Defense against Biotrophic and Necrotrophic Pathogens." *Annual Review of Phytopathology* 43: 205–27.
Glazebrook, Jane, Wenqiong Chen, Bram Estes, Hur-Song Chang, Christiane Nawrath, Jean-Pierre Métraux, Tong Zhu, and Fumiaki Katagiri. 2003. "Topology of the Network

Integrating Salicylate and Jasmonate Signal Transduction Derived from Global Expression Phenotyping." *The Plant Journal: For Cell and Molecular Biology* 34 (2): 217–28.

Gore, Athurva, Zhe Li, Ho-Lim Fung, Jessica E. Young, Suneet Agarwal, Jessica Antosiewicz-Bourget, Isabel Canto, et al. 2011. "Somatic Coding Mutations in Human Induced Pluripotent Stem Cells." *Nature* 471 (7336): 63–67.

Görlach, J., S. Volrath, G. Knauf-Beiter, G. Hengy, U. Beckhove, K. H. Kogel, M. Oostendorp, et al. 1996. "Benzothiadiazole, a Novel Class of Inducers of Systemic Acquired Resistance, Activates Gene Expression and Disease Resistance in Wheat." *The Plant Cell* 8 (4): 629–43.

Gratz, Scott J., Alexander M. Cummings, Jennifer N. Nguyen, Danielle C. Hamm, Laura K. Donohue, Melissa M. Harrison, Jill Wildonger, and Kate M. O'Connor-Giles. 2013. "Genome Engineering of Drosophila with the CRISPR RNA-Guided Cas9 Nuclease." *Genetics* 194 (4): 1029–35.

Gratz, Scott J., Fiona P. Ukken, C. Dustin Rubinstein, Gene Thiede, Laura K. Donohue, Alexander M. Cummings, and Kate M. O'Connor-Giles. 2014. "Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila." *Genetics* 196 (4): 961–71.

Griffiths, Anthony J. F., William M. Gelbart, Jeffrey H. Miller, and Richard C. Lewontin. 1999. *Genetics Begins with Variation*. W. H. Freeman.

Grizot, Sylvestre, Jean-Charles Epinat, Séverine Thomas, Aymeric Duclert, Sandra Rolland, Frédéric Pâques, and Philippe Duchateau. 2010. "Generation of Redesigned Homing Endonucleases Comprising DNA-Binding Domains Derived from Two Different Scaffolds." *Nucleic Acids Research* 38 (6): 2006–18.

Guo, Yan, Jiang Li, Chung-I Li, Jirong Long, David C. Samuels, and Yu Shyr. 2012. "The Effect of Strand Bias in Illumina Short-Read Sequencing Data." *BMC Genomics* 13 (November): 666.

Gupta, Ankit, Victoria L. Hall, Fatma O. Kok, Masahiro Shin, Joseph C. McNulty, Nathan D. Lawson, and Scot A. Wolfe. 2013. "Targeted Chromosomal Deletions and Inversions in Zebrafish." *Genome Research* 23 (6): 1008–17.

Gupta, Vaijayanti, Michael G. Willits, and Jane Glazebrook. 2000. "Arabidopsis Thaliana EDS4 Contributes to Salicylic Acid (SA)-Dependent Expression of Defense Responses: Evidence for Inhibition of Jasmonic Acid Signaling by SA." *Molecular Plant-Microbe Interactions: MPMI* 13 (5): 503–11.

Gutierrez-Guerrero, Alejandra, Sabina Sanchez-Hernandez, Giuseppe Galvani, Javier Pinedo-Gomez, Rocio Martin-Guerra, Almudena Sanchez-Gilabert, Araceli Aguilar-González, et al. 2018. "Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus." *Human Gene Therapy* 29 (3): 366–80.

Hale, Caryn R., Peng Zhao, Sara Olson, Michael O. Duff, Brenton R. Graveley, Lance Wells, Rebecca M. Terns, and Michael P. Terns. 2009. "RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex." *Cell* 139 (5): 945–56.

Hamdoun, Safae, Zhe Liu, Manroop Gill, Nan Yao, and Hua Lu. 2013. "Dynamics of Defense Responses and Cell Fate Change during Arabidopsis-Pseudomonas Syringae Interactions." *PloS One* 8 (12): e83219.

Händel, Eva-Maria, Stephen Alwin, and Toni Cathomen. 2009. "Expanding or Restricting the Target Site Repertoire of Zinc-Finger Nucleases: The Inter-Domain Linker as a Major Determinant of Target Site Selectivity." *Molecular Therapy: The Journal of the American Society of Gene Therapy* 17 (1): 104–11.

Hennig, J., J. Malamy, G. Grynkiewicz, J. Indulski, and D. F. Klessig. 1993. "Interconversion of the Salicylic Acid Signal and Its Glucoside in Tobacco." *The Plant Journal: For Cell and Molecular Biology* 4 (4): 593–600.

Hermans, P. W., D. van Soolingen, E. M. Bik, P. E. de Haas, J. W. Dale, and J. D. van Embden. 1991. "Insertion Element IS987 from Mycobacterium Bovis BCG Is Located

in a Hot-Spot Integration Region for Insertion Elements in Mycobacterium Tuberculosis Complex Strains." *Infection and Immunity* 59 (8): 2695–2705.
He, Yuehui, Scott D. Michaels, and Richard M. Amasino. 2003. "Regulation of Flowering Time by Histone Acetylation in Arabidopsis." *Science* 302 (5651): 1751–54.
Hilton, Isaac B., Anthony M. D'Ippolito, Christopher M. Vockley, Pratiksha I. Thakore, Gregory E. Crawford, Timothy E. Reddy, and Charles A. Gersbach. 2015. "Epigenome Editing by a CRISPR-Cas9-Based Acetyltransferase Activates Genes from Promoters and Enhancers." *Nature Biotechnology* 33 (5): 510–17.
Hirose, Michiko, Ayumi Hasegawa, Keiji Mochida, Shogo Matoba, Yuki Hatanaka, Kimiko Inoue, Tatsuhiko Goto, et al. 2017. "CRISPR/Cas9-Mediated Genome Editing in Wild-Derived Mice: Generation of Tamed Wild-Derived Strains by Mutation of the a (nonagouti) Gene." *Scientific Reports* 7 (February): 42476.
Holub, Eric B. 2006. "Evolution of Parasitic Symbioses between Plants and Filamentous Microorganisms." *Current Opinion in Plant Biology* 9 (4): 397–405.
———. 2008. "Natural History of Arabidopsis Thaliana and Oomycete Symbioses." *European Journal of Plant Pathology / European Foundation for Plant Pathology* 122 (1): 91–109.
Horvath, Philippe, Dennis A. Romero, Anne-Claire Coûté-Monvoisin, Melissa Richards, Hélène Deveau, Sylvain Moineau, Patrick Boyaval, Christophe Fremaux, and Rodolphe Barrangou. 2008. "Diversity, Activity, and Evolution of CRISPR Loci in Streptococcus Thermophilus." *Journal of Bacteriology* 190 (4): 1401–12.
Hsu, Patrick D., David A. Scott, Joshua A. Weinstein, F. Ann Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li, et al. 2013. "DNA Targeting Specificity of RNA-Guided Cas9 Nucleases." *Nature Biotechnology* 31 (9): 827–32.
Huang, Peng, An Xiao, Xiangjun Tong, Shuo Lin, and Bo Zhang. 2016. "Targeted Mutagenesis in Zebrafish by TALENs." In *TALENs: Methods and Protocols*, edited by Ralf Kühn, Wolfgang Wurst, and Benedikt Wefers, 191–206. New York, NY: Springer New York.
Huang, Peng, An Xiao, Mingguo Zhou, Zuoyan Zhu, Shuo Lin, and Bo Zhang. 2011. "Heritable Gene Targeting in Zebrafish Using Customized TALENs." *Nature Biotechnology* 29 (8): 699–700.
Huang, Wei E., Linfeng Huang, Gail M. Preston, Martin Naylor, John P. Carr, Yanhong Li, Andrew C. Singer, Andrew S. Whiteley, and Hui Wang. 2006. "Quantitative in Situ Assay of Salicylic Acid in Tobacco Leaves Using a Genetically Modified Biosensor Strain of Acinetobacter Sp. ADP1." *The Plant Journal: For Cell and Molecular Biology* 46 (6): 1073–83.
Huber, Julie A., David B. Mark Welch, Hilary G. Morrison, Susan M. Huse, Phillip R. Neal, David A. Butterfield, and Mitchell L. Sogin. 2007. "Microbial Population Structures in the Deep Marine Biosphere." *Science* 318 (5847): 97–100.
Hussein, Samer M., Nizar N. Batada, Sanna Vuoristo, Reagan W. Ching, Reija Autio, Elisa Närvä, Siemon Ng, et al. 2011. "Copy Number Variation and Selection during Reprogramming to Pluripotency." *Nature* 471 (7336): 58–62.
Hwang, Seon-Hee, In Ah Lee, Se Won Yie, and Duk-Ju Hwang. 2008. "Identification of an OsPR10a Promoter Region Responsive to Salicylic Acid." *Planta* 227 (5): 1141–50.
Hwang, Woong Y., Yanfang Fu, Deepak Reyon, Morgan L. Maeder, Prakriti Kaini, Jeffry D. Sander, J. Keith Joung, Randall T. Peterson, and Jing-Ruey Joanna Yeh. 2013. "Heritable and Precise Zebrafish Genome Editing Using a CRISPR-Cas System." *PloS One* 8 (7): e68708.
Hyun, Youbong, Jungeun Kim, Seung Woo Cho, Yeonhee Choi, Jin-Soo Kim, and George Coupland. 2015. "Site-Directed Mutagenesis in Arabidopsis Thaliana Using Dividing Tissue-Targeted RGEN of the CRISPR/Cas System to Generate Heritable Null Alleles." *Planta* 241 (1): 271–84.
Hyun, Youbong, Hyein Yun, Kyunghyuk Park, Hyonhwa Ohr, Okchan Lee, Dong-Hwan Kim, Sibum Sung, and Yeonhee Choi. 2013. "The Catalytic Subunit of Arabidopsis DNA

Polymerase α Ensures Stable Maintenance of Histone Modification." *Development* 140 (1): 156–66.
Irion, Uwe, Jana Krauss, and Christiane Nüsslein-Volhard. 2014. "Precise and Efficient Genome Editing in Zebrafish Using the CRISPR/Cas9 System." *Development* 141 (24): 4827–30.
Ishino, Y., H. Shinagawa, K. Makino, M. Amemura, and A. Nakata. 1987. "Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia Coli, and Identification of the Gene Product." *Journal of Bacteriology* 169 (12): 5429–33.
Jacobs, Thomas B., Peter R. LaFayette, Robert J. Schmitz, and Wayne A. Parrott. 2015. "Targeted Genome Modifications in Soybean with CRISPR/Cas9." *BMC Biotechnology* 15 (1): 16.
Jacoby, Kyle, Abigail R. Lambert, and Andrew M. Scharenberg. 2017. "Characterization of Homing Endonuclease Binding and Cleavage Specificities Using Yeast Surface Display SELEX (YSD-SELEX)." *Nucleic Acids Research* 45 (3): e11.
Jahn, Courtney E., John K. Mckay, Ramil Mauleon, Janice Stephens, Kenneth L. McNally, Daniel R. Bush, Hei Leung, and Jan E. Leach. 2011. "Genetic Variation in Biomass Traits among 20 Diverse Rice Varieties." *Plant Physiology* 155 (1): 157–68.
Jansen, Ruud, Jan D. A. van Embden, Wim Gaastra, and Leo M. Schouls. 2002. "Identification of Genes That Are Associated with DNA Repeats in Prokaryotes." *Molecular Microbiology* 43 (6): 1565–75.
Jayakanthan, Mannu, Jayaraman Muthukumaran, Sanniyasi Chandrasekar, Konika Chawla, Ankita Punetha, and Durai Sundar. 2009. "ZifBASE: A Database of Zinc Finger Proteins and Associated Resources." *BMC Genomics* 10 (September): 421.
Jiang, Hongshan, Rong Lei, Shou-Wei Ding, and Shuifang Zhu. 2014. "Skewer: A Fast and Accurate Adapter Trimmer for next-Generation Sequencing Paired-End Reads." *BMC Bioinformatics* 15 (June): 182.
Jiang, Wenyan, David Bikard, David Cox, Feng Zhang, and Luciano A. Marraffini. 2013. "RNA-Guided Editing of Bacterial Genomes Using CRISPR-Cas Systems." *Nature Biotechnology* 31 (3): 233–39.
Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. "A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science* 337 (6096): 816–21.
Jin, Jing Bo, Yin Hua Jin, Jiyoung Lee, Kenji Miura, Chan Yul Yoo, Woe-Yeon Kim, Michael Van Oosten, et al. 2007. "The SUMO E3 Ligase, AtSIZ1, Regulates Flowering by Controlling a Salicylic Acid-Mediated Floral Promotion Pathway and through Affects on FLC Chromatin Structure: Floral Repression by the SUMO E3 Ligase SIZ1." *The Plant Journal: For Cell and Molecular Biology* 53 (3): 530–40.
———. 2008. "The SUMO E3 Ligase, AtSIZ1, Regulates Flowering by Controlling a Salicylic Acid-Mediated Floral Promotion Pathway and through Affects on FLC Chromatin Structure." *The Plant Journal: For Cell and Molecular Biology* 53 (3): 530–40.
Jones, Jonathan D. G., and Jeffery L. Dangl. 2006. "The Plant Immune System." *Nature* 444 (7117): 323–29.
Jones, Julie M., and Miriam H. Meisler. 2014. "Modeling Human Epilepsy by TALEN Targeting of Mouse Sodium Channel Scn8a." *Genesis* 52 (2): 141–48.
Jore, Matthijs M., Magnus Lundgren, Esther van Duijn, Jelle B. Bultema, Edze R. Westra, Sakharam P. Waghmare, Blake Wiedenheft, et al. 2011. "Structural Basis for CRISPR RNA-Guided DNA Recognition by Cascade." *Nature Structural & Molecular Biology* 18 (5): 529–36.
Karasov, Talia L., Juliana Almario, Claudia Friedemann, Wei Ding, Michael Giolai, Darren Heavens, Sonja Kersten, et al. 2018. "Arabidopsis Thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales." *Cell Host & Microbe* 24 (1): 168–79.e4.

Katagiri, Fumiaki, Roger Thilmony, and Sheng Yang He. 2002. “The Arabidopsis Thaliana-Pseudomonas Syringae Interaction.” *The Arabidopsis Book / American Society of Plant Biologists* 1 (March): e0039.

Katsuyama, Tomonori, Arslan Akmammedov, Makiko Seimiya, Samuel C. Hess, Cem Sievers, and Renato Paro. 2013. “An Efficient Strategy for TALEN-Mediated Genome Engineering in Drosophila.” *Nucleic Acids Research* 41 (17): e163.

Kay, Sabine, and Ulla Bonas. 2009. “How Xanthomonas Type III Effectors Manipulate the Host Plant.” *Current Opinion in Microbiology* 12 (1): 37–43.

Kearns, Nicola A., Hannah Pham, Barbara Tabak, Ryan M. Genga, Noah J. Silverstein, Manuel Garber, and René Maehr. 2015. “Functional Annotation of Native Enhancers with a Cas9-Histone Demethylase Fusion.” *Nature Methods* 12 (5): 401–3.

Kim, Daesik, Sangsu Bae, Jeongbin Park, Eunji Kim, Seokjoong Kim, Hye Ryeong Yu, Jinha Hwang, Jong-Il Kim, and Jin-Soo Kim. 2015. “Digenome-Seq: Genome-Wide Profiling of CRISPR-Cas9 off-Target Effects in Human Cells.” *Nature Methods* 12 (3): 237–43, 1 p following 243.

Kim, Y. G., J. Cha, and S. Chandrasegaran. 1996. “Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (3): 1156–60.

Kim, Y. G., and S. Chandrasegaran. 1994. “Chimeric Restriction Endonuclease.” *Proceedings of the National Academy of Sciences of the United States of America* 91 (3): 883–87.

Kleinstiver, Benjamin P., Vikram Pattanayak, Michelle S. Prew, Shengdar Q. Tsai, Nhu T. Nguyen, Zongli Zheng, and J. Keith Joung. 2016. “High-Fidelity CRISPR–Cas9 Nucleases with No Detectable Genome-Wide off-Target Effects.” *Nature* 529 (7587): 490–95.

Klessig, Daniel F., Miaoying Tian, and Hyong Woo Choi. 2016. “Multiple Targets of Salicylic Acid and Its Derivatives in Plants and Animals.” *Frontiers in Immunology* 7 (May): 206.

Kogel, K. H., U. Beckhove, J. Dreschers, and S. Munch. 1994. “Acquired Resistance in Barley (the Resistance Mechanism Induced by 2, 6-Dichloroisonicotinic Acid Is a Phenocopy of a Genetically Based Mechanism Governing Race ….” *Plant*. http://www.plantphysiol.org/content/106/4/1269.short.

Kosicki, Michael, Kärt Tomberg, and Allan Bradley. 2018. “Repair of Double-Strand Breaks Induced by CRISPR-Cas9 Leads to Large Deletions and Complex Rearrangements.” *Nature Biotechnology*, July. https://doi.org/10.1038/nbt.4192.

Kroj, Thomas, Gil Savino, Christiane Valon, Jérôme Giraudat, and François Parcy. 2003. “Regulation of Storage Protein Gene Expression in Arabidopsis.” *Development* 130 (24): 6065–73.

Kumar, Dhirendra. 2014. “Salicylic Acid Signaling in Disease Resistance.” *Plant Science: An International Journal of Experimental Plant Biology* 228 (November): 127–34.

Kus, Julianne V., Kasia Zaton, Raani Sarkar, and Robin K. Cameron. 2002. “Age-Related Resistance in Arabidopsis Is a Developmentally Regulated Defense Response to Pseudomonas Syringae.” *The Plant Cell* 14 (2): 479–90.

Labun, Kornel, Tessa G. Montague, James A. Gagnon, Summer B. Thyme, and Eivind Valen. 2016. “CHOPCHOP v2: A Web Tool for the next Generation of CRISPR Genome Engineering.” *Nucleic Acids Research* 44 (W1): W272–76.

Lampropoulos, Athanasios, Zoran Sutikovic, Christian Wenzl, Ira Maegele, Jan U. Lohmann, and Joachim Forner. 2013. “GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis.” *PloS One* 8 (12): e83043.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): nmeth.1923.

Larson, Matthew H., Luke A. Gilbert, Xiaowo Wang, Wendell A. Lim, Jonathan S. Weissman, and Lei S. Qi. 2013. “CRISPR Interference (CRISPRi) for Sequence-Specific Control of Gene Expression.” *Nature Protocols* 8 (11): 2180–96.

Lawton, K., K. Weymann, L. Friedrich, B. Vernooij, S. Uknes, and J. Ryals. 1995. “Systemic Acquired Resistance in Arabidopsis Requires Salicylic Acid but Not Ethylene.” *Molecular Plant-Microbe Interactions: MPMI* 8 (6): 863–70.

Lee, J., J. Nam, H. C. Park, G. Na, K. Miura, and J. B. Jin. 2007. “Salicylic Acid-mediated Innate Immunity in Arabidopsis Is Regulated by SIZ1 SUMO E3 Ligase.” *The Plant*. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-313X.2006.02947.x/full.

Leeuwen, Hans van, Daniel J. Kliebenstein, Marilyn A. L. West, Kyunga Kim, Remco van Poecke, Fumiaki Katagiri, Richard W. Michelmore, Rebecca W. Doerge, and Dina A. St.Clair. 2007. “Natural Variation among Arabidopsis Thaliana Accessions for Transcriptome Response to Exogenous Salicylic Acid.” *The Plant Cell* 19 (7): 2099–2110.

Lempe, Janne, Sureshkumar Balasubramanian, Sridevi Sureshkumar, Anandita Singh, Markus Schmid, and Detlef Weigel. 2005. “Diversity of Flowering Responses in Wild Arabidopsis Thaliana Strains.” *PLoS Genetics* 1 (1): 109–18.

Liang, F., M. Han, P. J. Romanienko, and M. Jasin. 1998. “Homology-Directed Repair Is a Major Double-Strand Break Repair Pathway in Mammalian Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (9): 5172–77.

Li, Dali, Zhongwei Qiu, Yanjiao Shao, Yuting Chen, Yuting Guan, Meizhen Liu, Yongmei Li, et al. 2013. “Heritable Gene Targeting in the Mouse and Rat Using a CRISPR-Cas System.” *Nature Biotechnology* 31 (8): 681–83.

Lieber, Michael R. 2010. “The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway.” *Annual Review of Biochemistry* 79: 181–211.

Lieber, M. R. 1999. “The Biochemistry and Biological Significance of Nonhomologous DNA End Joining: An Essential Repair Process in Multicellular Eukaryotes.” *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 4 (2): 77–85.

Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform.” *Bioinformatics* 25 (14): 1754–60.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79.

Li, Jian-Feng, Julie E. Norville, John Aach, Matthew McCormack, Dandan Zhang, Jenifer Bush, George M. Church, and Jen Sheen. 2013. “Multiplex and Homologous Recombination-Mediated Genome Editing in Arabidopsis and Nicotiana Benthamiana Using Guide RNA and Cas9.” *Nature Biotechnology* 31 (8): 688–91.

Lillestøl, Reidun K., Shiraz A. Shah, Kim Brügger, Peter Redder, Hien Phan, Jan Christiansen, and Roger A. Garrett. 2009. “CRISPR Families of the Crenarchaeal Genus Sulfolobus: Bidirectional Transcription and Dynamic Properties.” *Molecular Microbiology* 72 (1): 259–72.

Li, L., L. P. Wu, and S. Chandrasegaran. 1992. “Functional Domains in Fok I Restriction Endonuclease.” *Proceedings of the National Academy of Sciences of the United States of America* 89 (10): 4275–79.

Li, Meiru, Xiaoxia Li, Zejiao Zhou, Pingzhi Wu, Maichun Fang, Xiaoping Pan, Qiupeng Lin, Wanbin Luo, Guojiang Wu, and Hongqing Li. 2016. “Reassessment of the Four Yield-Related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System.” *Frontiers in Plant Science* 7 (March): 377.

Lim, E. K., C. J. Doucet, Y. Li, L. Elias, and D. Worrall. 2002. “The Activity of Arabidopsisglycosyltransferases toward Salicylic Acid, 4-Hydroxybenzoic Acid, and Other Benzoates.” *Journal of Biological*. http://www.jbc.org/content/277/1/586.short.

Lindsay, Helen, Alexa Burger, Berthin Biyong, Anastasia Felker, Christopher Hess, Jonas Zaugg, Elena Chiavacci, et al. 2016. “CrispRVariants Charts the Mutation Spectrum of Genome Engineering Experiments.” *Nature Biotechnology* 34 (7): 701–2.

Lin, Yanni, Thomas J. Cradick, Matthew T. Brown, Harshavardhan Deshmukh, Piyush Ranjan, Neha Sarode, Brian M. Wile, Paula M. Vertino, Frank J. Stewart, and Gang Bao. 2014. “CRISPR/Cas9 Systems Have off-Target Activity with Insertions or Deletions between Target DNA and Guide RNA Sequences.” *Nucleic Acids Research* 42 (11): 7473–85.

Li, Peijin, Daniele Filiault, Mathew S. Box, Envel Kerdaffrec, Cock van Oosterhout, Amity M. Wilczek, Johanna Schmitt, et al. 2014. “Multiple FLC Haplotypes Defined by Independent Cis-Regulatory Variation Underpin Life History Diversity in Arabidopsis Thaliana.” *Genes & Development* 28 (15): 1635–40.

Liu, Hao, Yuduan Ding, Yanqing Zhou, Wenqi Jin, Kabin Xie, and Ling-Ling Chen. 2017. “CRISPR-P 2.0: An Improved CRISPR-Cas9 Tool for Genome Editing in Plants.” *Molecular Plant* 10 (3): 530–32.

Liu, Jiyong, Changqing Li, Zhongsheng Yu, Peng Huang, Honggang Wu, Chuanxian Wei, Nannan Zhu, et al. 2012. “Efficient and Specific Modifications of the Drosophila Genome by Means of an Easy TALEN Strategy.” *Journal of Genetics and Genomics = Yi Chuan Xue Bao* 39 (5): 209–15.

Liu, Liangyu, Jian Zhang, Jessika Adrian, Lionel Gissot, George Coupland, Diqiu Yu, and Franziska Turck. 2014. “Elevated Levels of MYB30 in the Phloem Accelerate Flowering in Arabidopsis through the Regulation of FLOWERING LOCUS T.” *PloS One* 9 (2): e89799.

Liu, Po-Pu, Caroline C. von Dahl, and Daniel F. Klessig. 2011. “The Extent to Which Methyl Salicylate Is Required for Signaling Systemic Acquired Resistance Is Dependent on Exposure to Light after Infection.” *Plant Physiology* 157 (4): 2216–26.

Li, Wei, Il-Pyung Ahn, Yuese Ning, Chan-Ho Park, Lirong Zeng, Justin G. A. Whitehill, Haibin Lu, et al. 2012. “The U-Box/ARM E3 Ligase PUB13 Regulates Cell Death, Defense, and Flowering Time in Arabidopsis.” *Plant Physiology* 159 (1): 239–50.

Lloyd, Alan, Christopher L. Plaisier, Dana Carroll, and Gary N. Drews. 2005. “Targeted Mutagenesis Using Zinc-Finger Nucleases in Arabidopsis.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (6): 2232–37.

Loake, Gary, and Murray Grant. 2007. “Salicylic Acid in Plant Defence--the Players and Protagonists.” *Current Opinion in Plant Biology* 10 (5): 466–72.

Lowder, Levi G., Dengwei Zhang, Nicholas J. Baltes, Joseph W. Paul 3rd, Xu Tang, Xuelian Zheng, Daniel F. Voytas, Tzung-Fu Hsieh, Yong Zhang, and Yiping Qi. 2015. “A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation.” *Plant Physiology* 169 (2): 971–85.

Lu, Hua, Debra N. Rate, Jong Tae Song, and Jean T. Greenberg. 2003. “ACD6, a Novel Ankyrin Protein, Is a Regulator and an Effector of Salicylic Acid Signaling in the Arabidopsis Defense Response.” *The Plant Cell* 15 (10): 2408–20.

Lundberg, Derek S., Scott Yourstone, Piotr Mieczkowski, Corbin D. Jones, and Jeffery L. Dangl. 2013. “Practical Innovations for High-Throughput Amplicon Sequencing.” *Nature Methods* 10 (10): 999–1002.

Mackelprang, Rebecca, Rachel A. Okrent, and Mary C. Wildermuth. 2017. “Preference of Arabidopsis Thaliana GH3.5 Acyl Amido Synthetase for Growth versus Defense Hormone Acyl Substrates Is Dictated by Concentration of Amino Acid Substrate Aspartate.” *Phytochemistry* 143 (November): 19–28.

Macovei, Anca, Neah R. Sevilla, Christian Cantos, Gilda B. Jonson, Inez Slamet-Loedin, Tomáš Čermák, Daniel F. Voytas, Il-Ryong Choi, and Prabhjit Chadha-Mohanty. 2018. “Novel Alleles of Rice eIF4G Generated by CRISPR/Cas9-Targeted Mutagenesis Confer Resistance to Rice Tungro Spherical Virus.” *Plant Biotechnology Journal*, March. https://doi.org/10.1111/pbi.12927.

Maeder, Morgan L., James F. Angstman, Marcy E. Richardson, Samantha J. Linder, Vincent M. Cascio, Shengdar Q. Tsai, Quan H. Ho, et al. 2013. “Targeted DNA Demethylation and Activation of Endogenous Genes Using Programmable TALE-TET1 Fusion Proteins.” *Nature Biotechnology* 31 (12): 1137–42.

Ma, Hanhui, Ardalan Naseri, Pablo Reyes-Gutierrez, Scot A. Wolfe, Shaojie Zhang, and Thoru Pederson. 2015. “Multicolor CRISPR Labeling of Chromosomal Loci in Human Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (10): 3002–7.

Maheshwari, Panchanan. 1950. “An Introduction to the Embryology of Angiosperms.” McGraw-Hill,. http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=agrono.xis&method=post&formato=2&cantidad=1&expresion=mfn=000706.

Mahfouz, Magdy M., Lixin Li, Md Shamimuzzaman, Anjar Wibowo, Xiaoyun Fang, and Jian-Kang Zhu. 2011. “De Novo-Engineered Transcription Activator-like Effector (TALE) Hybrid Nuclease with Novel DNA Binding Specificity Creates Double-Strand Breaks.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (6): 2623–28.

Makarova, Kira S., Nick V. Grishin, Svetlana A. Shabalina, Yuri I. Wolf, and Eugene V. Koonin. 2006. “A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action.” *Biology Direct* 1 (March): 7.

Makarova, Kira S., Daniel H. Haft, Rodolphe Barrangou, Stan J. J. Brouns, Emmanuelle Charpentier, Philippe Horvath, Sylvain Moineau, et al. 2011. “Evolution and Classification of the CRISPR–Cas Systems.” *Nature Reviews. Microbiology* 9 (6): 467–77.

Makarova, Kira S., Yuri I. Wolf, Omer S. Alkhnbashi, Fabrizio Costa, Shiraz A. Shah, Sita J. Saunders, Rodolphe Barrangou, et al. 2015. “An Updated Evolutionary Classification of CRISPR-Cas Systems.” *Nature Reviews. Microbiology* 13 (11): 722–36.

Ma, Ke, Li Qiu, Kristin Mrasek, Jun Zhang, Thomas Liehr, Luciana Gonçalves Quintana, and Zheng Li. 2012. “Common Fragile Sites: Genomic Hotspots of DNA Damage and Carcinogenesis.” *International Journal of Molecular Sciences* 13 (9): 11974–99.

Malamy, J., J. P. Carr, D. F. Klessig, and I. Raskin. 1990. “Salicylic Acid: A Likely Endogenous Signal in the Resistance Response of Tobacco to Viral Infection.” *Science* 250 (4983): 1002–4.

Mali, Prashant, Kevin M. Esvelt, and George M. Church. 2013. “Cas9 as a Versatile Tool for Engineering Biology.” *Nature Methods* 10 (10): 957–63.

Marek, George, Ryan Carver, Yezhang Ding, Deepak Sathyanarayan, Xudong Zhang, and Zhonglin Mou. 2010. “A High-Throughput Method for Isolation of Salicylic Acid Metabolic Mutants.” *Plant Methods* 6 (September): 21.

Marraffini, Luciano A., and Erik J. Sontheimer. 2008. “CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA.” *Science* 322 (5909): 1843–45.

Martínez, Cristina, Elsa Pons, Guillermo Prats, and José León. 2004a. “Salicylic Acid Regulates Flowering Time and Links Defence Responses and Reproductive Development.” *The Plant Journal: For Cell and Molecular Biology* 37 (2): 209–17.

———. 2004b. “Salicylic Acid Regulates Flowering Time and Links Defence Responses and Reproductive Development: Salicylic Acid and Flowering Time in Arabidopsis.” *The Plant Journal: For Cell and Molecular Biology* 37 (2): 209–17.

Masepohl, B., K. Görlitz, and H. Böhme. 1996. “Long Tandemly Repeated Repetitive (LTRR) Sequences in the Filamentous Cyanobacterium Anabaena Sp. PCC 7120.” *Biochimica et Biophysica Acta* 1307 (1): 26–30.

Mashal, R. D., J. Koontz, and J. Sklar. 1995. “Detection of Mutations by Cleavage of DNA Heteroduplexes with Bacteriophage Resolvases.” *Nature Genetics* 9 (2): 177–83.

Maskos, Z., J. D. Rush, and W. H. Koppenol. 1990. “The Hydroxylation of the Salicylate Anion by a Fenton Reaction and γ-Radiolysis: A Consideration of the Respective Mechanisms.” *Free Radical Biology and Medicine* 8 (2): 153–62.

Mauch-Mani, B., and A. J. Slusarenko. 1996. “Production of Salicylic Acid Precursors Is a Major Function of Phenylalanine Ammonia-Lyase in the Resistance of Arabidopsis to Peronospora Parasitica.” *The Plant Cell* 8 (2): 203–12.

Ma, Yunmei, Haihui Lu, Brigette Tippin, Myron F. Goodman, Noriko Shimazaki, Osamu Koiwai, Chih-Lin Hsieh, Klaus Schwarz, and Michael R. Lieber. 2004. “A Biochemically Defined System for Mammalian Nonhomologous DNA End Joining.” *Molecular Cell* 16 (5): 701–13.

McDonald, J. H. 2014. “Wilcoxon Signed-Rank Test. Handbook of Biological Statistics.” *Www.biostathandbook.com/wilcoxonsignedrank.html*. Sparky House Publishing, Baltimore, Maryland. Available at: http://www. biostathandbook. com/wilcoxonsignedrank. html [accessed 14 Dec 2014]. http://www.biostathandbook.com/wilcoxonsignedrank.html.

McDowell, John M., Scott G. Williams, Nicholas T. Funderburg, Thomas Eulgem, and Jeffery L. Dangl. 2005. “Genetic Analysis of Developmentally Regulated Resistance to Downy Mildew (Hyaloperonospora Parasitica) in Arabidopsis Thaliana.” *Molecular Plant-Microbe Interactions: MPMI* 18 (11): 1226–34.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, et al. 2010. “The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data.” *Genome Research* 20 (9): 1297–1303.

Meihls, L. N., H. Kaur, and G. Jander. 2012. “Natural Variation in Maize Defense against Insect Herbivores.” *Cold Spring Harbor Symposia on Quantitative Biology* 77 (December): 269–83.

Mendenhall, Eric M., Kaylyn E. Williamson, Deepak Reyon, James Y. Zou, Oren Ram, J. Keith Joung, and Bradley E. Bernstein. 2013. “Locus-Specific Editing of Histone Modifications at Endogenous Enhancers.” *Nature Biotechnology* 31 (12): 1133–36.

Meng, Xiangdong, Marcus B. Noyes, Lihua J. Zhu, Nathan D. Lawson, and Scot A. Wolfe. 2008. “Targeted Gene Inactivation in Zebrafish Using Engineered Zinc-Finger Nucleases.” *Nature Biotechnology* 26 (6): 695–701.

Mercer, Andrew C., Thomas Gaj, Roberta P. Fuller, and Carlos F. Barbas 3rd. 2012. “Chimeric TALE Recombinases with Programmable DNA Sequence Specificity.” *Nucleic Acids Research* 40 (21): 11163–72.

Merrill, Richard M., Richard W. R. Wallbank, Vanessa Bull, Patricio C. A. Salazar, James Mallet, Martin Stevens, and Chris D. Jiggins. 2012. “Disruptive Ecological Selection on a Mating Cue.” *Proceedings. Biological Sciences / The Royal Society* 279 (1749): 4907–13.

Métraux, J. P., H. Signer, J. Ryals, E. Ward, M. Wyss-Benz, J. Gaudin, K. Raschdorf, E. Schmid, W. Blum, and B. Inverardi. 1990. “Increase in Salicylic Acid at the Onset of Systemic Acquired Resistance in Cucumber.” *Science* 250 (4983): 1004–6.

Metwally, Ashraf, Iris Finkemeier, Manfred Georgi, and Karl-Josef Dietz. 2003. “Salicylic Acid Alleviates the Cadmium Toxicity in Barley Seedlings.” *Plant Physiology* 132 (1): 272–81.

Meuwly, P., W. Molders, A. Buchala, and J. P. Metraux. 1995. “Local and Systemic Biosynthesis of Salicylic Acid in Infected Cucumber Plants.” *Plant Physiology* 109 (3): 1107–14.

Meyer, Melanie, Martin Hrabé de Angelis, Wolfgang Wurst, and Ralf Kühn. 2010. “Gene Targeting by Homologous Recombination in Mouse Zygotes Mediated by Zinc-Finger Nucleases.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (34): 15022–26.

Michaels, S. D., and R. M. Amasino. 1999. “FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering.” *The Plant Cell* 11 (5): 949–56.

Miller, Jeffrey C., Michael C. Holmes, Jianbin Wang, Dmitry Y. Guschin, Ya-Li Lee, Igor Rupniewski, Christian M. Beausejour, et al. 2007. “An Improved Zinc-Finger Nuclease

Architecture for Highly Specific Genome Editing." *Nature Biotechnology* 25 (7): 778–85.
Miller, Jeffrey C., Siyuan Tan, Guijuan Qiao, Kyle A. Barlow, Jianbin Wang, Danny F. Xia, Xiangdong Meng, et al. 2011. "A TALE Nuclease Architecture for Efficient Genome Editing." *Nature Biotechnology* 29 (2): 143–48.
Miller, J., A. D. McLachlan, and A. Klug. 1985. "Repetitive Zinc-Binding Domains in the Protein Transcription Factor IIIA from Xenopus Oocytes." *The EMBO Journal* 4 (6): 1609–14.
Miura, Kenji, Jing Bo Jin, and Paul M. Hasegawa. 2007. "Sumoylation, a Post-Translational Regulatory Process in Plants." *Current Opinion in Plant Biology* 10 (5): 495–502.
Miura, Kenji, Ana Rus, Altanbadralt Sharkhuu, Shuji Yokoi, Athikkattuvalasu S. Karthikeyan, Kashchandra G. Raghothama, Dongwon Baek, et al. 2005. "The Arabidopsis SUMO E3 Ligase SIZ1 Controls Phosphate Deficiency Responses." *Proceedings of the National Academy of Sciences of the United States of America* 102 (21): 7760–65.
Mojica, F. J., C. Ferrer, G. Juez, and F. Rodríguez-Valera. 1995. "Long Stretches of Short Tandem Repeats Are Present in the Largest Replicons of the Archaea Haloferax Mediterranei and Haloferax Volcanii and Could Be Involved in Replicon Partitioning." *Molecular Microbiology* 17 (1): 85–93.
Mojica, F. J., G. Juez, and F. Rodríguez-Valera. 1993. "Transcription at Different Salinities of Haloferax Mediterranei Sequences Adjacent to Partially Modified PstI Sites." *Molecular Microbiology* 9 (3): 613–21.
Mojica, F. J. M., C. Díez-Villaseñor, J. García-Martínez, and C. Almendros. 2009. "Short Motif Sequences Determine the Targets of the Prokaryotic CRISPR Defence System." *Microbiology* 155 (Pt 3): 733–40.
Mojica, Francisco J. M., César Díez-Villaseñor, Jesús García-Martínez, and Elena Soria. 2005. "Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements." *Journal of Molecular Evolution* 60 (2): 174–82.
Morris, K., S. A. MacKerness, T. Page, C. F. John, A. M. Murphy, J. P. Carr, and V. Buchanan-Wollaston. 2000. "Salicylic Acid Has a Role in Regulating Gene Expression during Leaf Senescence." *The Plant Journal: For Cell and Molecular Biology* 23 (5): 677–85.
Morris, S. W., B. Vernooij, S. Titatarn, M. Starrett, S. Thomas, C. C. Wiltse, R. A. Frederiksen, A. Bhandhufalck, S. Hulbert, and S. Uknes. 1998. "Induced Resistance Responses in Maize." *Molecular Plant-Microbe Interactions: MPMI* 11 (7): 643–58.
Morton, Jason, M. Wayne Davis, Erik M. Jorgensen, and Dana Carroll. 2006. "Induction and Repair of Zinc-Finger Nuclease-Targeted Double-Strand Breaks in Caenorhabditis Elegans Somatic Cells." *Proceedings of the National Academy of Sciences of the United States of America* 103 (44): 16370–75.
Moscou, Matthew J., and Adam J. Bogdanove. 2009. "A Simple Cipher Governs DNA Recognition by TAL Effectors." *Science* 326 (5959): 1501.
Muradov, A., L. Petrasovits, A. Davidson, and K. J. Scott. 1993. "A cDNA Clone for a Pathogenesis-Related Protein 1 from Barley." *Plant Molecular Biology* 23 (2): 439–42.
Nacke, Heiko, Andrea Thürmer, Antje Wollherr, Christiane Will, Ladislav Hodac, Nadine Herold, Ingo Schöning, Marion Schrumpf, and Rolf Daniel. 2011. "Pyrosequencing-Based Assessment of Bacterial Community Structure along Different Management Types in German Forest and Grassland Soils." *PloS One* 6 (2): e17000.
Nadeau, Nicola J. 2016. "Genes Controlling Mimetic Colour Pattern Variation in Butterflies." *Current Opinion in Insect Science* 17 (October): 24–31.
Nakata, A., M. Amemura, and K. Makino. 1989. "Unusual Nucleotide Arrangement with Repeated Sequences in the Escherichia Coli K-12 Chromosome." *Journal of Bacteriology* 171 (6): 3553–56.
Nawrath, Christiane, Silvia Heck, Nonglak Parinthawong, and Jean-Pierre Métraux. 2002. "EDS5, an Essential Component of Salicylic Acid-Dependent Signaling for Disease

Resistance in Arabidopsis, Is a Member of the MATE Transporter Family." *The Plant Cell* 14 (1): 275–86.
Nawrath, Christiane, and Jean-Pierre Métraux. 1999. "Salicylic Acid Induction–Deficient Mutants of Arabidopsis Express PR-2 and PR-5 and Accumulate High Levels of Camalexin after Pathogen Inoculation." *The Plant Cell* 11 (8): 1393–1404.
Nekrasov, Vladimir, Congmao Wang, Joe Win, Christa Lanz, Detlef Weigel, and Sophien Kamoun. 2017. "Rapid Generation of a Transgene-Free Powdery Mildew Resistant Tomato by Genome Deletion." *Scientific Reports* 7 (1): 482.
Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, et al. 1999. "Evidence for Lateral Gene Transfer between Archaea and Bacteria from Genome Sequence of Thermotoga Maritima." *Nature* 399 (6734): 323–29.
Nevo, Eviatar. 1973. "Adaptive Variation in Size of Cricket Frogs." *Ecology* 54 (6): 1271–81.
Niu, Yuyu, Bin Shen, Yiqiang Cui, Yongchang Chen, Jianying Wang, Lei Wang, Yu Kang, et al. 2014. "Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos." *Cell* 156 (4): 836–43.
Norman, Christel, Katharine A. Howell, A. Harvey Millar, James M. Whelan, and David A. Day. 2004. "Salicylic Acid Is an Uncoupler and Inhibitor of Mitochondrial Electron Transport." *Plant Physiology* 134 (1): 492–501.
Ochiai, Hiroshi, Tatsuo Miyamoto, Akinori Kanai, Kosuke Hosoba, Tetsushi Sakuma, Yoshiki Kudo, Keiko Asami, et al. 2014. "TALEN-Mediated Single-Base-Pair Editing Identification of an Intergenic Mutation Upstream of BUB1B as Causative of PCS (MVA) Syndrome." *Proceedings of the National Academy of Sciences of the United States of America* 111 (4): 1461–66.
Olofsson, Martin, Hanne Løvlie, Jessika Tibblin, Sven Jakobsson, and Christer Wiklund. 2013. "Eyespot Display in the Peacock Butterfly Triggers Antipredator Behaviors in Naïve Adult Fowl." *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology* 24 (1): 305–10.
Osborn, Mark J., Beau R. Webber, Friederike Knipping, Cara-Lin Lonetree, Nicole Tennis, Anthony P. DeFeo, Amber N. McElroy, et al. 2016. "Evaluation of TCR Gene Editing Achieved by TALENs, CRISPR/Cas9, and megaTAL Nucleases." *Molecular Therapy: The Journal of the American Society of Gene Therapy* 24 (3): 570–81.
Ossowski, Stephan, Korbinian Schneeberger, José Ignacio Lucas-Lledó, Norman Warthmann, Richard M. Clark, Ruth G. Shaw, Detlef Weigel, and Michael Lynch. 2010. "The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis Thaliana." *Science* 327 (5961): 92–94.
Pallas, J. A., N. L. Paiva, C. Lamb, and R. A. Dixon. 1996. "Tobacco Plants Epigenetically Suppressed in Phenylalanine Ammonia-lyase Expression Do Not Develop Systemic Acquired Resistance in Response to Infection by …." *The Plant Journal: For Cell and Molecular Biology*. http://onlinelibrary.wiley.com/doi/10.1046/j.1365-313X.1996.10020281.x/full.
Parker, J. E., E. B. Holub, L. N. Frost, A. Falk, N. D. Gunn, and M. J. Daniels. 1996. "Characterization of eds1, a Mutation in Arabidopsis Suppressing Resistance to Peronospora Parasitica Specified by Several Different RPP Genes." *The Plant Cell* 8 (11): 2033–46.
Park, Sang-Wook, Evans Kaimoyo, Dhirendra Kumar, Stephen Mosher, and Daniel F. Klessig. 2007. "Methyl Salicylate Is a Critical Mobile Signal for Plant Systemic Acquired Resistance." *Science* 318 (5847): 113–16.
Pater, Sylvia de, Leon W. Neuteboom, Johan E. Pinas, Paul J. J. Hooykaas, and Bert J. van der Zaal. 2009. "ZFN-Induced Mutagenesis and Gene-Targeting in Arabidopsis through Agrobacterium-Mediated Floral Dip Transformation." *Plant Biotechnology Journal* 7 (8): 821–35.
Pattanayak, Vikram, Steven Lin, John P. Guilinger, Enbo Ma, Jennifer A. Doudna, and David R. Liu. 2013. "High-Throughput Profiling of off-Target DNA Cleavage Reveals

RNA-Programmed Cas9 Nuclease Specificity." *Nature Biotechnology* 31 (August): 839.
Pavletich, N. P., and C. O. Pabo. 1991. "Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 A." *Science* 252 (5007): 809–17.
"[PDF]Multiple Comparisons Using R." n.d. http://www.ievbras.ru/ecostat/Kiril/R/Biblio/R_eng/Bretz%20Multiple%20Comparisons.pdf.
"[PDF]Package 'MASS' - CRAN.R-Project.org." n.d. https://cran.r-project.org/web/packages/MASS/MASS.pdf.
"[PDF]Package 'Multcomp' - CRAN.R-Project.org." n.d. https://cran.r-project.org/web/packages/multcomp/multcomp.pdf.
Pellegrini, L., O. Rohfritsch, B. Fritig, and M. Legrand. 1994. "Phenylalanine Ammonia-Lyase in Tobacco. Molecular Cloning and Gene Expression during the Hypersensitive Reaction to Tobacco Mosaic Virus and the Response to a Fungal Elicitor." *Plant Physiology* 106 (3): 877–86.
Peterson, Brenda A., David C. Haak, Marc T. Nishimura, Paulo Jpl Teixeira, Sean R. James, Jeffery L. Dangl, and Zachary L. Nimchuk. 2016. "Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis." *PloS One* 11 (9): e0162169.
Phillips, J. W., and W. F. Morgan. 1994. "Illegitimate Recombination Induced by DNA Double-Strand Breaks in a Mammalian Chromosome." *Molecular and Cellular Biology* 14 (9): 5794–5803.
Piganeau, Marion, Hind Ghezraoui, Anne De Cian, Lionel Guittat, Mark Tomishima, Loic Perrouault, Oliver René, et al. 2013. "Cancer Translocations in Human Cells Induced by Zinc Finger and TALE Nucleases." *Genome Research* 23 (7): 1182–93.
Pinello, Luca, Matthew C. Canver, Megan D. Hoban, Stuart H. Orkin, Donald B. Kohn, Daniel E. Bauer, and Guo-Cheng Yuan. 2016. "Analyzing CRISPR Genome-Editing Experiments with CRISPResso." *Nature Biotechnology* 34 (7): 695–97.
Platt, Randall J., Sidi Chen, Yang Zhou, Michael J. Yim, Lukasz Swiech, Hannah R. Kempton, James E. Dahlman, et al. 2014. "CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling." *Cell* 159 (2): 440–55.
Porteus, Matthew H., and David Baltimore. 2003. "Chimeric Nucleases Stimulate Gene Targeting in Human Cells." *Science* 300 (5620): 763.
Pougach, Ksenia, Ekaterina Semenova, Ekaterina Bogdanova, Kirill A. Datsenko, Marko Djordjevic, Barry L. Wanner, and Konstantin Severinov. 2010. "Transcription, Processing and Function of CRISPR Cassettes in Escherichia Coli." *Molecular Microbiology* 77 (6): 1367–79.
Pourcel, C., G. Salvignol, and G. Vergnaud. 2005. "CRISPR Elements in Yersinia Pestis Acquire New Repeats by Preferential Uptake of Bacteriophage DNA, and Provide Additional Tools for Evolutionary Studies." *Microbiology* 151 (Pt 3): 653–63.
Qi, Lei S., Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin, and Wendell A. Lim. 2013. "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression." *Cell* 152 (5): 1173–83.
Qin, Peiwu, Mahmut Parlak, Cem Kuscu, Jigar Bandaria, Mustafa Mir, Karol Szlachta, Ritambhara Singh, Xavier Darzacq, Ahmet Yildiz, and Mazhar Adli. 2017. "Live Cell Imaging of Low- and Non-Repetitive Chromosome Loci Using CRISPR-Cas9." *Nature Communications* 8 (March): 14725.
Rajjou, Loïc, Maya Belghazi, Romain Huguet, Caroline Robin, Adrien Moreau, Claudette Job, and Dominique Job. 2006. "Proteomic Investigation of the Effect of Salicylic Acid on Arabidopsis Seed Germination and Establishment of Early Defense Mechanisms." *Plant Physiology* 141 (3): 910–23.
Raskin, I. 1992. "Role of Salicylic Acid in Plants." *Annual Review of Plant Physiology and Plant Molecular Biology* 43 (1): 439–63.

Rasmussen, J. B., R. Hammerschmidt, and M. N. Zook. 1991. “Systemic Induction of Salicylic Acid Accumulation in Cucumber after Inoculation with Pseudomonas Syringae Pv Syringae.” *Plant Physiology* 97 (4): 1342–47.

Rate, D. N., J. V. Cuenca, G. R. Bowman, D. S. Guttman, and J. T. Greenberg. 1999. “The Gain-of-Function Arabidopsis acd6 Mutant Reveals Novel Regulation and Function of the Salicylic Acid Signaling Pathway in Controlling Cell Death, Defenses, and Cell Growth.” *The Plant Cell* 11 (9): 1695–1708.

Ren, Xingjie, Jin Sun, Benjamin E. Housden, Yanhui Hu, Charles Roesel, Shuailiang Lin, Lu-Ping Liu, et al. 2013. “Optimized Gene Editing Technology for Drosophila Melanogaster Using Germ Line-Specific Cas9.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (47): 19012–17.

Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. 2011. “Integrative Genomics Viewer.” *Nature Biotechnology* 29 (1): 24–26.

Römer, Patrick, Simone Hahn, Tina Jordan, Tina Strauss, Ulla Bonas, and Thomas Lahaye. 2007. “Plant Pathogen Recognition Mediated by Promoter Activation of the Pepper Bs3 Resistance Gene.” *Science* 318 (5850): 645–48.

Rouet, P., F. Smih, and M. Jasin. 1994. “Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease.” *Molecular and Cellular Biology* 14 (12): 8096–8106.

Rowe, Heather C., and Daniel J. Kliebenstein. 2008. “Complex Genetics Control Natural Variation in Arabidopsis Thaliana Resistance to Botrytis Cinerea.” *Genetics* 180 (4): 2237–50.

Rudin, N., and J. E. Haber. 1988. “Efficient Repair of HO-Induced Chromosomal Breaks in Saccharomyces Cerevisiae by Recombination between Flanking Homologous Sequences.” *Molecular and Cellular Biology* 8 (9): 3918–28.

Ryals, J. A., U. H. Neuenschwander, M. G. Willits, A. Molina, H. Y. Steiner, and M. D. Hunt. 1996. “Systemic Acquired Resistance.” *The Plant Cell* 8 (10): 1809–19.

Sambrook, J. and D. W. Russell (2001). Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Sánchez-León, Susana, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, and Francisco Barro. 2018. “Low-Gluten, Nontransgenic Wheat Engineered with CRISPR/Cas9.” *Plant Biotechnology Journal* 16 (4): 902–10.

Sanger, F., and A. R. Coulson. 1975. “A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase.” *Journal of Molecular Biology* 94 (3): 441–48.

Santner, Aaron, Luz Irina A. Calderon-Villalobos, and Mark Estelle. 2009. “Plant Hormones Are Versatile Chemical Regulators of Plant Growth.” *Nature Chemical Biology* 5 (5): 301–7.

Sapranauskas, Rimantas, Giedrius Gasiunas, Christophe Fremaux, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. 2011. “The Streptococcus Thermophilus CRISPR/Cas System Provides Immunity in Escherichia Coli.” *Nucleic Acids Research* 39 (21): 9275–82.

Scheben, Armin, Felix Wolter, Jacqueline Batley, Holger Puchta, and David Edwards. 2017. “Towards CRISPR/Cas Crops - Bringing Together Genomics and Genome Editing.” *The New Phytologist* 216 (3): 682–98.

Schemske, Douglas W., and Paulette Bierzychudek. 2007. “Spatial Differentiation for Flower Color in the Desert Annual Linanthus Parryae: Was Wright Right?” *Evolution; International Journal of Organic Evolution* 61 (11): 2528–43.

Schornack, Sebastian, Annett Meyer, Patrick Römer, Tina Jordan, and Thomas Lahaye. 2006. “Gene-for-Gene-Mediated Recognition of Nuclear-Targeted AvrBs3-like Bacterial Effector Proteins.” *Journal of Plant Physiology* 163 (3): 256–72.

Schweizer, Patrick, Edith Schlagenhauf, Ulrich Schaffrath, and Robert Dudler. 1999. “Different Patterns of Host Genes Are Induced in Rice by Pseudomonas Syringae, a Biological Inducer of Resistance, and the Chemical Inducer Benzothiadiazole (BTH).” *European Journal of Plant Pathology / European Foundation for Plant Pathology* 105 (7): 659–65.

Semenova, Ekaterina, Maxim Nagornykh, Mikhail Pyatnitskiy, Irena I. Artamonova, and Konstantin Severinov. 2009. “Analysis of CRISPR System Function in Plant Pathogen Xanthomonas Oryzae.” *FEMS Microbiology Letters* 296 (1): 110–16.

Sensen, C. W., Robert L. Charlebois, Cynthia Chow, Ib Groth Clausen, Bruce Curtis, W. Ford Doolittle, Michel Duguet, et al. 1998. “Completing the Sequence of the Sulfolobus Solfataricus P2 Genome.” *Extremophiles: Life under Extreme Conditions* 2 (3): 305–12.

Shan, Qiwei, Yanpeng Wang, Jun Li, Yi Zhang, Kunling Chen, Zhen Liang, Kang Zhang, et al. 2013. “Targeted Genome Modification of Crop Plants Using a CRISPR-Cas System.” *Nature Biotechnology* 31 (8): 686–88.

Shimizu, Yuka, Mital S. Bhakta, and David J. Segal. 2009. “Restricted Spacer Tolerance of a Zinc Finger Nuclease with a Six Amino Acid Linker.” *Bioorganic & Medicinal Chemistry Letters* 19 (14): 3970–72.

Shukla, Vipula K., Yannick Doyon, Jeffrey C. Miller, Russell C. DeKelver, Erica A. Moehle, Sarah E. Worden, Jon C. Mitchell, et al. 2009. “Precise Genome Modification in the Crop Species Zea Mays Using Zinc-Finger Nucleases.” *Nature* 459 (7245): 437–41.

Silva, George, Laurent Poirot, Roman Galetto, Julianne Smith, Guillermo Montoya, Philippe Duchateau, and Frédéric Pâques. 2011. “Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy.” *Current Gene Therapy* 11 (1): 11–27.

Silverman, P., M. Seskar, D. Kanter, P. Schweizer, J. P. Metraux, and I. Raskin. 1995. “Salicylic Acid in Rice (Biosynthesis, Conjugation, and Possible Role).” *Plant Physiology* 108 (2): 633–39.

Singh, Vijayata, Shweta Roy, Mrunmay Kumar Giri, Ratnesh Chaturvedi, Zulkarnain Chowdhury, Jyoti Shah, and Ashis Kumar Nandi. 2013. “Arabidopsis Thaliana FLOWERING LOCUS D Is Required for Systemic Acquired Resistance.” *Molecular Plant-Microbe Interactions: MPMI* 26 (9): 1079–88.

Slaymaker, Ian M., Linyi Gao, Bernd Zetsche, David A. Scott, Winston X. Yan, and Feng Zhang. 2016. “Rationally Engineered Cas9 Nucleases with Improved Specificity.” *Science* 351 (6268): 84–88.

Smith, J., M. Bibikova, F. G. Whitby, A. R. Reddy, S. Chandrasegaran, and D. Carroll. 2000. “Requirements for Double-Strand Cleavage by Chimeric Restriction Enzymes with Zinc Finger DNA-Recognition Domains.” *Nucleic Acids Research* 28 (17): 3361–69.

Smith, Jennifer A., Raymond Hammerschmidt, and Dennis W. Fulbright. 1991. “Rapid Induction of Systemic Resistance in Cucumber by Pseudomonas Syringae Pv. Syringae.” *Physiological and Molecular Plant Pathology* 38 (3): 223–35.

Son, Ga Hyun, Bong Soo Park, Jong Tae Song, and Hak Soo Seo. 2014. “FLC-Mediated Flowering Repression Is Positively Regulated by Sumoylation.” *Journal of Experimental Botany* 65 (1): 339–51.

Song, Jong Tae. 2006. “Induction of a Salicylic Acid Glucosyltransferase, AtSGT1, Is an Early Disease Response in Arabidopsis Thaliana.” *Molecules and Cells* 22 (2): 233–38.

Sprunck, Stefanie, Svenja Rademacher, Frank Vogler, Jacqueline Gheyselinck, Ueli Grossniklaus, and Thomas Dresselhaus. 2012. “Egg Cell–Secreted EC1 Triggers Sperm Cell Activation During Double Fertilization.” *Science* 338 (6110): 1093–97.

Stacey, Gary, Crystal Bickley McAlvin, Sung-Yong Kim, José Olivares, and María José Soto. 2006. “Effects of Endogenous Salicylic Acid on Nodulation in the Model Legumes Lotus Japonicus and Medicago Truncatula.” *Plant Physiology* 141 (4): 1473–81.

Staswick, Paul E., Bogdan Serban, Martha Rowe, Iskender Tiryaki, Marién T. Maldonado, Mitsa C. Maldonado, and Walter Suza. 2005. “Characterization of an Arabidopsis Enzyme Family That Conjugates Amino Acids to Indole-3-Acetic Acid.” *The Plant Cell* 17 (2): 616–27.

Stemmer, Manuel, Thomas Thumberger, Maria del Sol Keyer, Joachim Wittbrodt, and Juan L. Mateo. 2015. “CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool.” *PloS One* 10 (4): e0124633.

Stone, Edward. 1763. “XXXII. An Account of the Success of the Bark of the Willow in the Cure of Agues. In a Letter to the Right Honourable George Earl of Macclesfield, President of R. S. from the Rev. Mr. Edward Stone, of Chipping-Norton in Oxfordshire.” *Philosophical Transactions* 53 (January): 195–200.

Straus, Marco R., Steffen Rietz, Emiel Ver Loren van Themaat, Michael Bartsch, and Jane E. Parker. 2010. “Salicylic Acid Antagonism of EDS1-Driven Cell Death Is Important for Immune and Oxidative Stress Responses in Arabidopsis.” *The Plant Journal: For Cell and Molecular Biology* 62 (4): 628–40.

Strauss, E. C., J. A. Kobori, G. Siu, and L. E. Hood. 1986. “Specific-Primer-Directed DNA Sequencing.” *Analytical Biochemistry* 154 (1): 353–60.

Strawn, Marcus A., Sharon K. Marr, Kentaro Inoue, Noriko Inada, Chloe Zubieta, and Mary C. Wildermuth. 2007. “Arabidopsis Isochorismate Synthase Functional in Pathogen-Induced Salicylate Biosynthesis Exhibits Properties Consistent with a Role in Diverse Stress Responses.” *The Journal of Biological Chemistry* 282 (8): 5919–33.

Svitashev, Sergei, Joshua K. Young, Christine Schwartz, Huirong Gao, S. Carl Falco, and A. Mark Cigan. 2015. “Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA.” *Plant Physiology* 169 (2): 931–45.

Takeda, M., K. Sakai, M. Terashima, H. Kaneda, H. Hayashi, K. Tanaka, K. Okamoto, et al. 2015. “Clinical Application of Amplicon-Based next-Generation Sequencing to Therapeutic Decision Making in Lung Cancer.” *Annals of Oncology: Official Journal of the European Society for Medical Oncology / ESMO* 26 (12): 2477–82.

Tegelen, L. J. van, P. R. Moreno, A. F. Croes, R. Verpoorte, and G. J. Wullems. 1999. “Purification and cDNA Cloning of Isochorismate Synthase from Elicited Cell Cultures of Catharanthus Roseus.” *Plant Physiology* 119 (2): 705–12.

Thakore, Pratiksha I., Anthony M. D’Ippolito, Lingyun Song, Alexias Safi, Nishkala K. Shivakumar, Ami M. Kabadi, Timothy E. Reddy, Gregory E. Crawford, and Charles A. Gersbach. 2015. “Highly Specific Epigenome Editing by CRISPR-Cas9 Repressors for Silencing of Distal Regulatory Elements.” *Nature Methods* 12 (12): 1143–49.

Thomma, Bart P. H. J., Thorsten Nürnberger, and Matthieu H. A. J. Joosten. 2011. “Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy.” *The Plant Cell* 23 (1): 4–15.

Thorvaldsdóttir, Helga, James T. Robinson, and Jill P. Mesirov. 2013. “Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration.” *Briefings in Bioinformatics* 14 (2): 178–92.

Todesco, Marco, Sureshkumar Balasubramanian, Tina T. Hu, M. Brian Traw, Matthew Horton, Petra Epple, Christine Kuhns, et al. 2010. “Natural Allelic Variation Underlying a Major Fitness Trade-off in Arabidopsis Thaliana.” *Nature* 465 (7298): 632–36.

Tothova, Zuzana, John M. Krill-Burger, Katerina D. Popova, Catherine C. Landers, Quinlan L. Sievers, David Yudovich, Roger Belizaire, et al. 2017. “Multiplex CRISPR/Cas9-Based Genome Editing in Human Hematopoietic Stem Cells Models Clonal Hematopoiesis and Myeloid Neoplasia.” *Cell Stem Cell* 21 (4): 547–55.e8.

Tsai, Shengdar Q., Zongli Zheng, Nhu T. Nguyen, Matthew Liebers, Ved V. Topkar, Vishal Thapar, Nicolas Wyvekens, et al. 2015. “GUIDE-Seq Enables Genome-Wide Profiling of off-Target Cleavage by CRISPR-Cas Nucleases.” *Nature Biotechnology* 33 (2): 187–97.

Tsuda, Kenichi, and Fumiaki Katagiri. 2010. “Comparing Signaling Mechanisms Engaged in Pattern-Triggered and Effector-Triggered Immunity.” *Current Opinion in Plant Biology* 13 (4): 459–65.

Tsuda, Kenichi, Masanao Sato, Jane Glazebrook, Jerry D. Cohen, and Fumiaki Katagiri. 2008. “Interplay between MAMP-Triggered and SA-Mediated Defense Responses.” *The Plant Journal: For Cell and Molecular Biology* 53 (5): 763–75.

Ueta, Risa, Chihiro Abe, Takahito Watanabe, Shigeo S. Sugano, Ryosuke Ishihara, Hiroshi Ezura, Yuriko Osakabe, and Keishi Osakabe. 2017. “Rapid Breeding of Parthenocarpic Tomato Plants Using CRISPR/Cas9.” *Scientific Reports* 7 (1): 507.

Urnov, Fyodor D., Jeffrey C. Miller, Ya-Li Lee, Christian M. Beausejour, Jeremy M. Rock, Sheldon Augustus, Andrew C. Jamieson, Matthew H. Porteus, Philip D. Gregory, and Michael C. Holmes. 2005. “Highly Efficient Endogenous Human Gene Correction Using Designed Zinc-Finger Nucleases.” *Nature* 435 (7042): 646–51.

Van Poecke, Rmp, M. A. Posthumus, and M. Dicke. 2001. “Herbivore-Induced Volatile Production by Arabidopsis Thaliana Leads to Attraction of the Parasitoid Cotesia Rubecula: Chemical, Behavioral, and Gene-Expression ….” *Journal of Chemical Ecology*. https://link.springer.com/article/10.1023/A:1012213116515.

Vasseur, Francois, George Wang, Justine Bresson, Rebecca Schwab, and Detlef Weigel. 2018. “Image-Based Methods for Phenotyping Growth Dynamics and Fitness in Arabidopsis Thaliana.” *bioRxiv*. https://doi.org/10.1101/208512.

Verble-Pearson, Robin M., Matthew E. Gifford, and Stephen P. Yanoviak. 2015. “Variation in Thermal Tolerance of North American Ants.” *Journal of Thermal Biology* 48 (February): 65–68.

Veres, Adrian, Bridget S. Gosis, Qiurong Ding, Ryan Collins, Ashok Ragavendran, Harrison Brand, Serkan Erdin, Chad A. Cowan, Michael E. Talkowski, and Kiran Musunuru. 2014. “Low Incidence of off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing.” *Cell Stem Cell* 15 (1): 27–30.

Vernooij, B., L. Friedrich, A. Morse, R. Reist, R. Kolditz-Jawhar, E. Ward, S. Uknes, H. Kessmann, and J. Ryals. 1994. “Salicylic Acid Is Not the Translocated Signal Responsible for Inducing Systemic Acquired Resistance but Is Required in Signal Transduction.” *The Plant Cell* 6 (7): 959–65.

Villajuana-Bonequi, Mitzi, Nabil Elrouby, Karl Nordström, Thomas Griebel, Andreas Bachmair, and George Coupland. 2014a. “Elevated Salicylic Acid Levels Conferred by Increased Expression of ISOCHORISMATE SYNTHASE 1 Contribute to Hyperaccumulation of SUMO1 Conjugates in the Arabidopsis Mutant Early in Short Days 4.” *The Plant Journal: For Cell and Molecular Biology* 79 (2): 206–19.

———. 2014b. “Elevated Salicylic Acid Levels Conferred by Increased Expression of ISOCHORISMATE SYNTHASE 1 Contribute to Hyperaccumulation of SUMO1 Conjugates in the Arabidopsis Mutant Early in Short Days 4.” *The Plant Journal: For Cell and Molecular Biology* 79 (2): 206–19.

Vleeshouwers, Vivianne G. A. A., and Richard P. Oliver. 2014. “Effectors as Tools in Disease Resistance Breeding against Biotrophic, Hemibiotrophic, and Necrotrophic Plant Pathogens.” *Molecular Plant-Microbe Interactions: MPMI* 27 (3): 196–206.

Vlot, A. Corina, D’maris Amick Dempsey, and Daniel F. Klessig. 2009. “Salicylic Acid, a Multifaceted Hormone to Combat Disease.” *Annual Review of Phytopathology* 47: 177–206.

Vlot, Anna Corina, Po-Pu Liu, Robin K. Cameron, Sang-Wook Park, Yue Yang, Dhirendra Kumar, Fasong Zhou, et al. 2008. “Identification of Likely Orthologs of Tobacco Salicylic Acid-Binding Protein 2 and Their Role in Systemic Acquired Resistance in Arabidopsis Thaliana.” *The Plant Journal: For Cell and Molecular Biology* 56 (3): 445–56.

Wang, Fujun, Chunlian Wang, Piqing Liu, Cailin Lei, Wei Hao, Ying Gao, Yao-Guang Liu, and Kaijun Zhao. 2016. “Enhanced Rice Blast Resistance by CRISPR/Cas9-Targeted

Mutagenesis of the ERF Transcription Factor Gene OsERF922." *PloS One* 11 (4): e0154027.
Wang, Guan-Feng, Savanna Seabolt, Safae Hamdoun, Gina Ng, Jin Park, and Hua Lu. 2011. "Multiple Roles of WIN3 in Regulating Disease Resistance, Cell Death, and Flowering Time in Arabidopsis." *Plant Physiology* 156 (3): 1508–19.
Wang, Haoyi, Hui Yang, Chikdu S. Shivalila, Meelad M. Dawlaty, Albert W. Cheng, Feng Zhang, and Rudolf Jaenisch. 2013. "One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering." *Cell* 153 (4): 910–18.
Wang, Lin, Raka M. Mitra, Keegan D. Hasselmann, Masanao Sato, Lisa Lenarz-Wyatt, Jerry D. Cohen, Fumiaki Katagiri, and Jane Glazebrook. 2008. "The Genetic Network Controlling the Arabidopsis Transcriptional Response to Pseudomonas Syringae Pv. Maculicola: Roles of Major Regulators and the Phytotoxin Coronatine." *Molecular Plant-Microbe Interactions: MPMI* 21 (11): 1408–20.
Wang, Pengcheng, Jun Zhang, Lin Sun, Yizan Ma, Jiao Xu, Sijia Liang, Jinwu Deng, et al. 2017. "High Efficient Multi-Sites Genome Editing in Allotetraploid Cotton (Gossypium Hirsutum) Using CRISPR/Cas9 System." *Plant Biotechnology Journal*, May. https://doi.org/10.1111/pbi.12755.
Wang, Xiaoling, Yebo Wang, Xiwei Wu, Jinhui Wang, Yingjia Wang, Zhaojun Qiu, Tammy Chang, He Huang, Ren-Jang Lin, and Jiing-Kuan Yee. 2015. "Unbiased Detection of off-Target Cleavage by CRISPR-Cas9 and TALENs Using Integrase-Defective Lentiviral Vectors." *Nature Biotechnology* 33 (2): 175–78.
Watanabe, Masahito, Kazuhiro Umeyama, Hitomi Matsunari, Shuko Takayanagi, Erika Haruyama, Kazuaki Nakano, Tsukasa Fujiwara, Yuka Ikezawa, Hiromitsu Nakauchi, and Hiroshi Nagashima. 2010. "Knockout of Exogenous EGFP Gene in Porcine Somatic Cells Using Zinc-Finger Nucleases." *Biochemical and Biophysical Research Communications* 402 (1): 14–18.
Watanabe, Takashi, Taketo Tanigaki, Hirotaka Nishi, Atushi Ushimaru, and Tsuyoshi Takeuchi. 2002. "A Quantitative Analysis of Geographic Color Variation in Two Geotrupes Dung Beetles." *Zoological Science* 19 (3): 351–58.
Wefers, Benedikt, Melanie Meyer, Oskar Ortiz, Martin Hrabé de Angelis, Jens Hansen, Wolfgang Wurst, and Ralf Kühn. 2013. "Direct Production of Mouse Disease Models by Embryo Microinjection of TALENs and Oligodeoxynucleotides." *Proceedings of the National Academy of Sciences of the United States of America* 110 (10): 3782–87.
Westra, Edze R., and Stan J. J. Brouns. 2012. "The Rise and Fall of CRISPRs--Dynamics of Spacer Acquisition and Loss." *Molecular Microbiology* 85 (6): 1021–25.
Whalen, M. C., R. W. Innes, A. F. Bent, and B. J. Staskawicz. 1991. "Identification of Pseudomonas Syringae Pathogens of Arabidopsis and a Bacterial Locus Determining Avirulence on Both Arabidopsis and Soybean." *The Plant Cell* 3 (1): 49–59.
White, R. F. 1979. "Acetylsalicylic Acid (aspirin) Induces Resistance to Tobacco Mosaic Virus in Tobacco." *Virology* 99 (2): 410–12.
Wildermuth, M. C., J. Dewdney, G. Wu, and F. M. Ausubel. 2001. "Isochorismate Synthase Is Required to Synthesize Salicylic Acid for Plant Defence." *Nature* 414 (6863): 562–65.
Wu, Gary D., James D. Lewis, Christian Hoffmann, Ying-Yu Chen, Rob Knight, Kyle Bittinger, Jennifer Hwang, et al. 2010. "Sampling and Pyrosequencing Methods for Characterizing Bacterial Communities in the Human Gut Using 16S Sequence Tags." *BMC Microbiology* 10 (July): 206.
Wu, Y., D. Zhang, J. Y. Chu, P. Boyle, Y. Wang, and I. D. Brindle. 2012. "The Arabidopsis NPR1 Protein Is a Receptor for the Plant Defense Hormone Salicylic Acid." *Cell Reports*. https://www.sciencedirect.com/science/article/pii/S2211124712001337.
Xie, Kabin, and Yinong Yang. 2013. "RNA-Guided Genome Editing in Plants Using a CRISPR–Cas System." *Molecular Plant* 6 (6): 1975–83.

Yadegari, Ramin, and Gary N. Drews. 2004. “Female Gametophyte Development.” *The Plant Cell* 16 Suppl (April): S133–41.

Yalpani, N., J. Leon, M. A. Lawton, and I. Raskin. 1993. “Pathway of Salicylic Acid Biosynthesis in Healthy and Virus-Inoculated Tobacco.” *Plant Physiology* 103 (2): 315–21.

Yalpani, N., P. Silverman, T. M. Wilson, D. A. Kleier, and I. Raskin. 1991. “Salicylic Acid Is a Systemic Signal and an Inducer of Pathogenesis-Related Proteins in Virus-Infected Tobacco.” *The Plant Cell* 3 (8): 809–18.

Yamamato, S., M. Katagiri, H. Maeno, and O. Hayaishi. 1965. “Salicylate Hydroxylase, a Monooxygenase* Requiring Flavin Adenine Dinucleotide.” *The Journal of Biological Chemistry* 240: 3408–14.

Yang, Hong, Jia-Jing Wu, Ting Tang, Ke-De Liu, and Cheng Dai. 2017. “CRISPR/Cas9-Mediated Genome Editing Efficiently Creates Specific Mutations at Multiple Loci Using One sgRNA in Brassica Napus.” *Scientific Reports* 7 (1): 7489.

Yan, Sen, Zhuchi Tu, Zhaoming Liu, Nana Fan, Huiming Yang, Su Yang, Weili Yang, et al. 2018. “A Huntingtin Knockin Pig Model Recapitulates Features of Selective Neurodegeneration in Huntington’s Disease.” *Cell*, March. https://doi.org/10.1016/j.cell.2018.03.005.

Yee, Jiing-Kuan. 2016. “Off-Target Effects of Engineered Nucleases.” *The FEBS Journal* 283 (17): 3239–48.

Ye, Huiying, Zhili Rong, and Ying Lin. 2017. “Live Cell Imaging of Genomic Loci Using dCas9-SunTag System and a Bright Fluorescent Protein.” *Protein & Cell* 8 (11): 853–55.

Zalatan, Jesse G., Michael E. Lee, Ricardo Almeida, Luke A. Gilbert, Evan H. Whitehead, Marie La Russa, Jordan C. Tsai, et al. 2015. “Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds.” *Cell* 160 (1-2): 339–50.

Zetsche, Bernd, Jonathan S. Gootenberg, Omar O. Abudayyeh, Ian M. Slaymaker, Kira S. Makarova, Patrick Essletzbichler, Sara E. Volz, et al. 2015. “Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System.” *Cell* 163 (3): 759–71.

Zhang, Feng, Le Cong, Simona Lodato, Sriram Kosuri, George M. Church, and Paola Arlotta. 2011. “Efficient Construction of Sequence-Specific TAL Effectors for Modulating Mammalian Transcription.” *Nature Biotechnology* 29 (2): 149–53.

Zhang, Feng, Morgan L. Maeder, Erica Unger-Wallace, Justin P. Hoshaw, Deepak Reyon, Michelle Christian, Xiaohong Li, et al. 2010. “High Frequency Targeted Mutagenesis in Arabidopsis Thaliana Using Zinc Finger Nucleases.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (26): 12028–33.

Zhang, Hui, Jinshan Zhang, Pengliang Wei, Botao Zhang, Feng Gou, Zhengyan Feng, Yanfei Mao, et al. 2014. “The CRISPR/Cas9 System Produces Specific and Homozygous Targeted Gene Editing in Rice in One Generation.” *Plant Biotechnology Journal* 12 (6): 797–807.

Zhang, Yi, Zhen Liang, Yuan Zong, Yanpeng Wang, Jinxing Liu, Kunling Chen, Jin-Long Qiu, and Caixia Gao. 2016. “Efficient and Transgene-Free Genome Editing in Wheat through Transient Expression of CRISPR/Cas9 DNA or RNA.” *Nature Communications* 7 (August): 12617.

Zheng, Zuyu, Stephen L. Mosher, Baofang Fan, Daniel F. Klessig, and Zhixiang Chen. 2007. “Functional Analysis of Arabidopsis WRKY25 Transcription Factor in Plant Defense against Pseudomonas Syringae.” *BMC Plant Biology* 7 (January): 2.

Zhu, Wangsheng, Maricris Zaidem, Anna-Lena Van de Weyer, Rafal M. Gutaker, Eunyoung Chae, Sang-Tae Kim, Felix Bemm, et al. 2018. “Modulation of ACD6 Dependent Hyperimmunity by Natural Alleles of an Arabidopsis Thaliana NLR Resistance Gene.” *bioRxiv*. https://doi.org/10.1101/300798.

Appendix

Appendix Table 1: A list of all the binary constructs that were used during the amplicon screening development.

Name	Promoter	Cas9	Terminator	SgRNA1	sgRNA2	Selection marker		
						promoter	CDS	terminator
pEF008	35S	pcoCas9	rbcs	*ICS1*	*ICS1*	MAS	BASTA	MAS
pEF009	35S	pcoCas9	rbcs	*FLC*	*FLC*	MAS	BASTA	MAS
pEF071	35S	pcoCas9	rbcs	*ICS1*	*ICS1*	AT2S3	mCherry	MAS
pEF073	EC1.1	pcoCas9	rbcs	*ICS1*	*ICS1*	AT2S3	mCherry	MAS
pEF070	ICU2	AthCas9	rbcs	*ICS1*	*ICS1*	AT2S3	mCherry	MAS

Appendix Table 2: List of all oligonucleotides used during this study. The blue bold nucleotides of the oligonucleotides G-40599, G-40600, G-40602, G-40603, G-40605, G-40606, G-42015, G-40608, G-40609 and G-42016 indicate the frameshifting nucleotides used during the amplicon generation.

name	sequence	target region	orientation	purpose
G-39418	GTGGATCCCAACGCTGCCAGATTCGAT	pICU2	F	for generating pEF044
G-39419	GGTACCTTTTACAAATCCGGTCAAT		R	
G-39590	GGCGGAGACCTCTCCTTACA		F	Sanger sequencing
G-39672	GTGGATCCATGGATAAGAAGTACTCTAT	AthCas9	F	for generating pEF048
G-39673	ATCTGCAGTCAAACCTTCCTCTTCTTCT		R	
G-39860	CTTCAAGTCTAACTTCGATC		F	Sanger sequencing
G-39861	GATCTCTTGTTCAAGACCAA		F	
G-39862	CATTCTTGAAGGATGATTCT		F	
G-39863	CGGTGGTTTCGATTCTCCTA		F	
G-40025	TTGGTCTTGAACAAGAGATC		R	
G-40026	TAGGAGAATCGAAACCACCG		R	
G-40094	GGATAAGGGTGCTAGTGCTC		F	
G 39211	GGTCTCAACCTCCCGCCTTATGATTTCTTCGGT	pEC1.1	F	for generating pEF019
G 39247	GGTCTCATGTTCCTTCTCAACAGATTGATAA		R	
G-38860	GATTGGAAAAGGCGTACTTATCGCGTTTTAGAGCTATGCTG	DsgRNA/FLC	F	for generating pEF014
G-38861	CTAAAACGCGATAAGTACGCCTTTTCCAATCACTACTTCGAC		R	
G-38862	GATTGTTCTCGTCGTCGCCGGCTCGTTTTAGAGCTATGCTG	EsgRNA/FLC	F	for generating pEF015
G-38863	CTAAAACGAGGCGGAGACGACGAGAACAATCACTACTTCGAC		R	
G-38958	GATTGAATCAATTGCTCCGATTTGCGTTTTAGAGCTATGCTG	EsgRNA/ICS1	F	for generating pEF016
G-38959	CTAAAACGCAAATCGGAGCAATTGATTCAATCACTACTTCGAC		R	
G-38956	GATTGTTCTCTCGTCGCAGTGACGTGTTTTAGAGCTATGCTG	DsgRNA/ICS1	F	for generating pEF017
G-38957	CTAAAACACGTCACTGCGACGAGAGAACAATCACTACTTCGAC		R	
G-38719	GGTCTCAGGCTAAAGCTTTCGTTCGTATCATCG	trbcs	F	for generating pEF003
G-38720	GGTCTCACTGACGATTGATGCATGTTGTCAATC		R	
G-38717	GGATCCAAAGAACAATAGTATTTC	sgRNA	F	for generating pEF004, pEF005
G-38718	GGTACCAAAAAAAAGCACCGAC		R	
G-38382	GTGGATCCATGGATTACAAGGATGATG	pcoCas9	F	for generating pcoCas9 without 7 ECO31I sites
G-38369	CTCAGCGGTTTCTCCAGAA		R	
G-38368	TTCTGGAGAAACCGCTGAG		F	
G-38371	GGTTAAGGTCACCCTCGATA		R	
G-38370	TATCGAGGGTGACCTTAACC		F	
G-38373	GGGTGATGGTTTCCTCAGAC		R	
G-38372	GTCTGAGGAAACCATCACCC		F	
G-38375	CTGTCTGGTTTCAACAAGC		R	
G-38374	GCTTGTTGAAACCAGACAG		F	
G-38377	CTCCGGTTTCTCCGTTGGTTTCGATAAGTGGCCTCTTTC		R	
G-38376	GAAAGAGGCCACTTATCGAAACCAACGGAGAAACCGGAG		F	
G-38383	ATCTGCAGTCACTTCTTCTTCTTAGCC		R	
G-38923	GAGGAGCTTCTTGTTAAGTT	pcoCas9	F	sanger sequencing
G-38924	GTACGTTACCGAGGGAATGA		F	
G-38925	GGAATCCTTCAGACCGTTAA		F	
G-44245	GAGGAAGCCCACATTCTCAGT	Chr1:11740331	F	generating fragment for Sanger sequencing deletion1
G-44246	TGGGCCGAAATAACCCCGTT		R	
G-44256	GGACGCAGTTTCATCTTCTC	Chr5:16997961	F	generating fragment for Sanger sequencing SNP1
G-44257	TCGCGGTGTATGGACGGAAT		R	

Appendix Table 2: List of all oligonucleotides used during this study. The blue bold nucleotides of the oligonucleotides G-40599, G-40600, G-40602, G-40603, G-40605, G-40606, G-42015, G-40608, G-40609 and G-42016 indicate the frameshifting nucleotides used during the amplicon generation (continue).

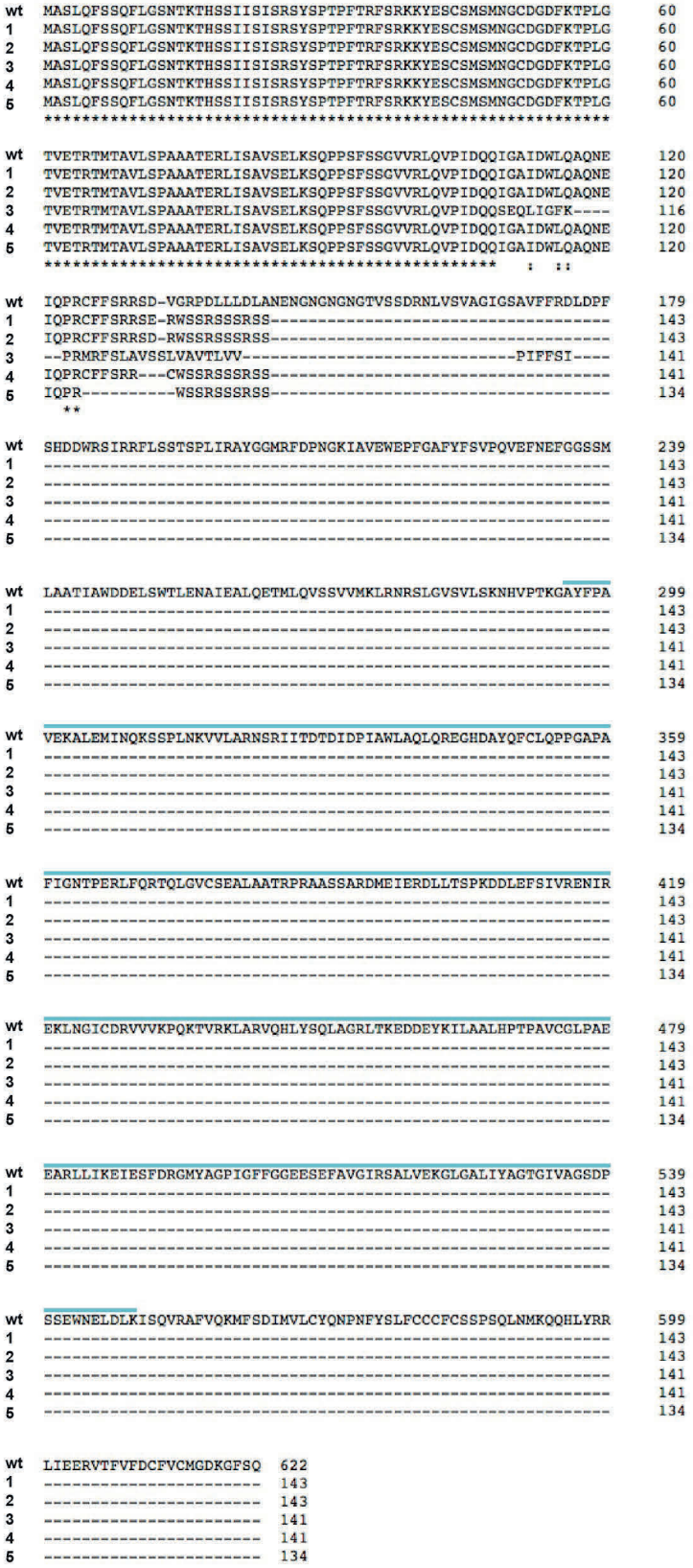
name	sequence	target region	orientation	purpose
G-44319	TGTAACTTGCAAGGTCCTCG	Chr1:16902771	F	generating fragment for Sanger sequencing SNP2
G-44320	AAACCTATACCGATCAACCG		R	
G-44321	TGACACGTGAAACGACAAGG	Chr1:22057548	F	generating fragment for Sanger sequencing deletion2
G-44322	GTCGTCCATTAAATACACCC		R	
G-44252	GTGTTGTCACGAAGCGAAGA	Chr1:26836216	F	generating fragment for Sanger sequencing deletion3
G-44253	CAACTTCCATGTCAAGGTAC		R	
G-44254	TTAACGCGTCCCAGACAGAC	Chr3:23201179	F	generating fragment for Sanger sequencing deletion4
G-44255	TTCGGACGACGTCGTTTTGG		R	
G-44343	ATGAATGGCTCCGTCGAAAC	Chr5:20396387	F	generating fragment for Sanger sequencing deletion5
G-44344	AAGATTGTTTCTGCTCTGGG		R	
G-44345	CAAAGACGAAGAGTAAACCC	Chr5:25073734	F	generating fragment for Sanger sequencing SNP3
G-44346	ACACCCTCAAGCTTAAACCC		R	
G-41950	GTTGTAAAGGCAAGGCGGCG	*ICS1*	F	generating fragment for Sanger sequencing of targeted ICS1 region
G-41951	ATCAAAACGCATACCACCAT		R	
G-40598	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGAAAAGTAA AAGAGCACAAA	FLC/TruSeq adapters	F	enerating the 1st PCR product for amplicon screening for identiying mutations at the targeted location
G-40599	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**T**CAGAAAAGTA AAAGAGCACAAA		F	
G-40600	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CT**CAGAAAAGT AAAAGAGCACAAA		F	
G-40601	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTTCTCCAA ACGTCGCAACGG		R	
G-40602	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**TT**GTTCTCCA AACGTCGCAACGG		R	
G-40603	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CT**TGTTCTCCA AACGTCGCAACGG		R	
G-40604	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTCGAATTT TCCCGCGCAATA	ICS1/TruSeq adapters	F	enerating the 1st PCR product for amplicon screening for identiying mutations at the targeted location
G-40605	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**T**CAGTCGAATT TTCCCGCGCAATA		F	
G-40606	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CT**CAGTCGAAT TTTCCCGCGCAATA		F	
G-42015	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**ACT**CAGTCGAA TTTTCCCGCGCAATA		F	
G-40607	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCAGATGAC ACTGTTCCGTTTC		R	
G-40608	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**T**CACAGATGA CACTGTTCCGTTTC		R	
G-40609	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CT**CACAGATGA CACTGTTCCGTTTC		R	
G-42016	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**ACT**CACAGAT GACACTGTTCCGTTTC		R	
G-40610	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTC	TruSeq adapter	F	adding the last part of the forward TruSeq adapter at the final amplicons

Appendix Table 3: List of constructs used during this study.

name	Alias	purpose of the construct	backbone
EF004	D000-sgRNA	template for generating new sgRNAs	pUC19
EF005	E000-sgRNA	template for generating new sgRNAs	pUC19
EF007	B000-pcoCas9	carrying the pcoCas9 gene without 7 Eco31I sites	pUC19
EF003	C000-rbcs	carrying the trbcs sequence	pUC19
EF008	CaMV35:pcoCas9:sgRNA(*ICS1*):BASTA	targeting the *ICS1* locus	pGreenIIS
EF009	CaMV35:pcoCas9:sgRNA(*FLC*):BASTA	targeting the *FLC* locus	pGreenIIS
EF014	D000-1st target FLC	carrying one target for *FLC* locus	pUC19
EF015	E000-2nd target FLC	carrying one target for *FLC* locus	pUC19
EF016	E000-5th target ICS1	carrying one target for *ICS1* locus	pUC19
EF017	D000-6th target ICS1	carrying one target for *ICS1* locus	pUC19
EF019	A000-EC1	carrying the pEC1.1	pUC19
EF044	A000-ICU2 pr	carrying the pICU2	pUC19
EF048	B000-AthCas9	carrying the AthCas9	pUC19
EF071	CaMV35:pcoCas9:sgRNA(*ICS1*):mCherry	targeting the *ICS1* locus	pGreenIIS
EF073	pEC1.1:pcoCas9:sgRNA(*ICS1*):mCherry	targeting the *ICS1* locus	pGreenIIS
EF070	pICU2:AthCas9:sgRNA(*ICS1*):mCherry	targeting the *ICS1* locus	pGreenIIS
KF264451		carrying original pcoCas9 sequence(courtesy of Prof. Sheen)	
pDe-Cas9		carrying the AthCas9 (courtesy of Prof. Puchta)	

Appendix Table 4: List of predicted off-target cleavage sites using the *in silico* tools, CRISPR-P 2.0, CRISPR design and CCTop (Stemmer et al. 2015; Liu et al. 2017).

in silico tool	chromosome	position	location	gene
CCTop	chr1	7164736-7164758	exon	AT1G20670
	chr1	18954593-18954615	exon	AT1G51170
	chr1	25648355-25648377	exon	AT1G68400
	chr1	25951172-25951194		intergenic
	chr1	28375255-28375277		intergenic
	chr2	6976847-6976869	EXON	AT2G16040
	chr2	15523508-15523530	UTR	at2g36960
	chr2	17188546-17188568	UTR	AT2G41220
	chr2	18184237-18184259		intergenic
	chr3	826177-826199	exon	AT3G03470
	chr3	9662119-9662141	exon	AT3G26380
	chr3	9756922-9756944	exon	ATG26570
	chr3	11679384-11679406	exon	AT3G29783
	chr3	13286207-13286229	exon	AT3G32360
	chr3	14430302-14430324	exon	AT3G42256
	chr3	1982506-1982528		intergenic
	chr3	21210468-21210490	exon	AT3G57320
	chr4	2269951-2269973		intergenic
	chr4	7613446-7613468	UTR	AT4G13040
	chr4	7958724-7958746	exon	AT4G13700
	chr4	9359779-9359801	exon	AT4G16620
	chr4	9708820-9708842	intron	AT4G17380
	chr4	17507387-17507409	exon	AT4G37190
	chr5	3181406-3181428		intergenic
	chr5	10044681-10044703	exon	AT5G28050
	chr5	14560649-14560671		intergenic
CRISPR-P 2.0	chr1	1217285	UTR	AT1G04480
	chr1	7164742	exon	AT1G20670
	chr2	16184505	UTR	AT2G28710
	chr3	12372309	exon	AT3G30742
	chr4	2269968		intergenic
	chr4	17507404	exon	AT4G37190
	chr5	3441804	UTR	AT5G10920
CRISPR design	chr1	1217266	UTR	AT1G04480
	chr1	7164731	exon	AT1G20670
	chr2	16184486	UTR	AT2G28710
	chr2	17399772	exon	AT2G41705
	chr2	19448487	UTR	AT2G47410
	chr3	826172	exon	AT3G03470
	chr3	12372298	exon	AT3G30742
	chr3	13286205	exon	AT3G32360
	chr4	2269949		intergenic
	chr4	11127906	exon	AT4G20740
	chr4	17507385	exon	AT4G37190
	chr5	3441793	UTR	AT5G10920



Appendix Figure 1: Alignment of the amino acid sequences that were generated after the CRISPR mutations at the *ICS1* targeted region against the wild type sequence. The blue line indicates the chorismate binding domain of the protein. 1: Col-0 *c-ics1* and Fei-0 *c-ics1*, 2: Ey1.5-2 *c-ics1*, Yeg-1 *c-ics1* and TüWa1-2 *c-ics1-2*, 3: Koch-1 *c-ics1*, 4: ICE50 *c-ics1* and 5: TüWa1-2 *c-ics1-1*.

Appendix Figure 2: Alignments of the regions where the off-target variants were located against one of the sgRNA sequence. For each region that potential target was chosen based on the best alignments (lower number of mismatches). The underlined part of each region indicates the location of the variant (deletion or SNP).