



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Hochschule für Angewandte Wissenschaften Hamburg

Fakultät Life Sciences

Etablierung eines *in planta* Funktionalitätstests für Zielsequenz- spezifisch programmierbare Endonukleasen

Bachelorarbeit

im Studiengang Biotechnologie

vorgelegt von

Hannes Trautwein
1765844

Hamburg Bergedorf

am 08. November 2016

Erstgutachter: Prof. Dr. Jörg Andrä
Hochschule für Angewandte Wissenschaften Hamburg

Zweitgutachter: Dr. Jochen Kumlehn
Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung
(IPK) Gatersleben

Danksagung

Ich möchte mich rechtherzlich beim Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben für die Möglichkeit zur Bearbeitung dieser interessanten Thematik im Zuge meiner Bachelorarbeit bedanken.

Mein besonderer Dank gilt meinem Betreuer und Arbeitsgruppenleiter am IPK Herrn Dr. Jochen Kumlehn, meinem Kollegen Stefan Hiekel und meiner Kollegin Dr. Nagaveni Budhagatapalli, sowie der ganzen Arbeitsgruppe Pflanzliche Reproduktionsbiologie, die mich bei allen Fragen und Problemen während meiner Arbeit stets unterstützten.

Weiterhin möchte ich mich rechtherzlich bei Herrn Prof. Dr. Jörg Andrä von der HAW Hamburg für seine Bereitschaft sich als Erstgutachter zur Verfügung zu stellen bedanken.

Für die immerwährende Unterstützung und Geduld gilt meiner Familie, insbesondere meinen Eltern, der größte Dank.

Inhaltsverzeichnis

Danksagung	2
Inhaltsverzeichnis	3
Tabellenverzeichnis	6
Abbildungsverzeichnis	7
Abkürzungsverzeichnis	8
1 Theoretische Grundlagen	11
1.1 Einleitung	11
1.2 Genomveränderung mit programmierbaren Endonukleasen.....	12
1.2.1 Grundlegende Mechanismen.....	12
1.2.1.1 Nichthomologe End-zu-End-Verknüpfung.....	12
1.2.1.2 Homologie-basierte DNA-Reparatur	14
1.2.1.3 DSB-induzierte, ortsspezifische Mutagenese	17
1.2.2 Sequenz-spezififizierbare, synthetische Endonukleasen	18
1.2.2.1 Meganukleasen	18
1.2.2.2 Zinkfinger-Nukleasen	19
1.2.2.3 TALE-Nukleasen	22
1.2.2.4 RNA-vermittelte Endonukleasen	29
1.3 Biolistische Transformation von Pflanzenzellen.....	35
1.4 Zielstellung der Arbeit.....	37
2 Material und Methoden	38
2.1 Materialien.....	38
2.1.1 Chemikalien und Enzyme	38
2.1.2 Kits und Verbrauchsmaterialien.....	39
2.1.3 Geräte und Software	39
2.1.4 Pflanzenmaterial und Inkubationsmedium.....	40
2.1.5 Bakterienstamm und Kulturmedien.....	41
2.1.6 Gelelektrophoresepuffer	42
2.1.7 <i>TALEN</i> -Proteine	42

2.1.8	Plasmide	47
2.1.9	Oligonukleotide	48
2.2	Methoden.....	50
2.2.1	Allgemeine Klonierungsarbeiten.....	50
2.2.1.1	Agarose-Gelelektrophorese.....	50
2.2.1.2	Polymerase- Kettenreaktion.....	51
2.2.1.3	Enzymatische Spaltungen von DNA mit Restriktionsendonukleasen	52
2.2.1.4	Ligation von DNA-Fragmenten	53
2.2.1.5	Transformation von <i>E. coli</i> mittels Hitzeschock-Verfahren	54
2.2.1.6	Kultivierung von <i>E. coli</i>	54
2.2.1.7	Isolierung und Reinigung von Plasmid-DNA.....	54
2.2.1.8	Konzentrationsbestimmungen von Plasmid-DNA.....	55
2.2.1.9	Identifikationen rekombinanter Plasmid-DNA.....	55
2.2.2	Klonierung der Testvektoren.....	55
2.2.2.1	Modifikation des Vektorrückgrats <i>pUbi-AB</i>	55
2.2.2.2	Klonierung des generischen Testvektor <i>pHT1</i>	56
2.2.2.3	Testvektor <i>pHT2</i> für das <i>gfp</i> -spezifische <i>TALEN</i> -Paar	56
2.2.2.4	Testvektor <i>pHT3</i> für das <i>cenH3α</i> -spezifische <i>TALEN</i> -Paar	57
2.2.2.5	Testvektor <i>pHT4</i> für das <i>mlo</i> -spezifische <i>TALEN</i> -Paar	58
2.2.3	Transformation und Analyse von Gerstenzellen	58
2.2.3.1	Biolistischer Gentransfer	58
2.2.3.2	Expressionsanalysen mittels Licht- und konfokaler laser scanning Mikroskopie	59
3	Ergebnisse	61
3.1	Klonierung der Testvektoren.....	62
3.1.1	Modifikation des Vektorrückgrats <i>pUbi-AB</i>	62
3.1.2	Klonierung des generischen Testvektors <i>pHT1</i>	63
3.1.3	Klonierung des Testvektors <i>pHT2</i> für das <i>gfp</i> -spezifische <i>TALEN</i> -Paar.....	64
3.1.4	Klonierung des Testvektors <i>pHT3</i> für das <i>cenH3α</i> -spezifische <i>TALEN</i> -Paar ...	66

3.1.5	Klonierung des Testvektors <i>pHT4</i> für das <i>mlo</i> -spezifische <i>TALEN</i> -Paar	67
3.2	Expressionsversuche an Gerstenzellen.....	67
3.2.1	Benetzungsreaktionen und Partikelbombardment.....	67
3.2.2	Expressionsanalysen am Licht- und konfokalen <i>Laser Scanning</i> -Mikroskop ...	68
3.2.2.1	Testvektor pHT1	68
3.2.2.2	Testvektor pHT2 für das <i>gfp</i> -spezifische <i>TALEN</i> -Paar	68
3.2.2.3	Testvektor pHT3 für das <i>cenH3α</i> -spezifische <i>TALEN</i> -Paar	69
3.2.2.4	Testvektor pHT4 für das <i>mlo</i> -spezifische <i>TALEN</i> -Paar	73
4	Diskussion.....	76
	Literaturverzeichnis.....	83
	Eidesstattliche Erklärung.....	94

Tabellenverzeichnis

Tabelle 1: <i>Vektoren für Gensequenzamplifikationen</i>	48
Tabelle 2: <i>TALEN-kodierende Vektoren</i>	48
Tabelle 3: <i>PCR- und Oligonukleotidprimer</i>	49
Tabelle 4: <i>Zusammensetzungen der PCR-Reaktionsansätze</i>	52
Tabelle 5: <i>PCR-Prozessverlauf von Genamplifikationen</i>	52
Tabelle 6: <i>Zusammensetzungen der Restriktionsreaktionsansätze</i>	53
Tabelle 7: <i>Zusammensetzungen der Ligationsansätze mit T4-DNA-Ligase</i>	53
Tabelle 8: <i>Zusammensetzungen der Ligationsansätze mit Quick-Ligase</i>	54
Tabelle 9: <i>Ergebnisse der Expressionsanalysen mit pHT2 und den gfp-spezifischen TALEN</i>	69
Tabelle 10: <i>Ergebnisse der Kontrollversuche mit pHT2</i>	69
Tabelle 11: <i>Ergebnisse der lichtmikroskopischen Expressionsanalysen mit pHT3 und den ScCenH3α-spezifizierten TALEN</i>	71
Tabelle 12: <i>Ergebnisse der am Lichtmikroskop ausgewerteten Kontrollversuche mit pHT3</i> ..	71
Tabelle 13: <i>Ergebnisse der Expressionsanalysen am CLSM mit pHT3 und den ScCenH3α- spezifizierten TALEN</i>	72
Tabelle 14: <i>Ergebnisse der am CLSM ausgewerteten Kontrollversuche mit pHT3</i>	73
Tabelle 15: <i>Ergebnisse der Expressionsanalysen am Licht- und CLS-Mikroskop mit pHT4 und den mlo-spezifisierten TALEN</i>	74
Tabelle 16: <i>Ergebnisse der Kontrollversuche mit pHT4</i>	75

Abbildungsverzeichnis

Abb. 1 Übersicht der wesentlichen Modelle von DSB-Reparaturmechanismen in somatischen Pflanzenzellen	15
Abb. 2 Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz durch ein Zinkfinger-Nuklease-Paar zur Induktion eines Doppelstrangbruchs	21
Abb. 3 Schematische Struktur eines <i>TAL</i> -Effektorproteins am Beispiel <i>PthXo1</i> aus <i>X. oryzae</i>	24
Abb. 4 3D Strukturmodell der DNA-Bindung durch ein <i>TALE</i> -Protein.	25
Abb. 5 Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz zur Induktion eines Doppelstrangbruchs durch ein <i>TALEN</i> -Paar.....	28
Abb. 6 Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz zur Induktion eines Doppelstrangbruchs durch eine RNA-geleitete Endonuklease.....	32
Abb. 7 Abfolgen der RVD-Motive und Gesamtzielsequenz des <i>gfp</i> -spezifischen TALEN-Paars.....	43
Abb. 8 Abfolgen der RVD-Motive und Gesamtzielsequenz des <i>cenH3a</i> -spezifischen TALEN-Paars.	45
Abb. 9 Abfolgen der RVD-Motive und Gesamtzielsequenz des <i>mlo</i> -spezifischen TALEN-Paars	46
Abb. 10 Schema des Expressionsvektors <i>pUbi-AB</i>	47
Abb. 11 Komplettes Ablaufschema eines einzelnen Klonierungsschrittes.	50
Abb. 12 Konzept der Testvektoren zur Ermittlung der <i>in planta</i> Restriktionsaktivitäten von Sequenz-spezifisierten Endonukleasen.....	62
Abb. 13 Ausschnitt des Elektrophoresegels zur Identifikation von rekombinanten <i>pUbi-AB</i> -Vektoren.....	63
Abb. 14 Ausschnitt des Elektrophoresegels zur Identifikation von <i>pUbi-AB_M</i> Vektoren mit inserierter <i>Egfp</i> -Sequenz.....	63
Abb. 15 Ausschnitt des Elektrophoresegels zur Identifikation von <i>pHT1</i>	63
Abb. 16 Schematische Darstellung des generischen Vektors <i>pHT1</i>	64
Abb. 17 Ausschnitt des Elektrophoresegels zur Identifikation von <i>pUbi-AB_M</i> Vektoren mit inserierter <i>mCherry</i> -Sequenz.....	65
Abb. 18 Ausschnitt des Elektrophoresegels zur Identifikation von <i>pHT2</i>	65
Abb. 19 Schematische Darstellung des Aufbaus und der Funktionsweise des Testvektors <i>pHT2</i>	65
Abb. 20 Schematische Darstellung des Testvektors <i>pHT3</i>	66
Abb. 21 Ausschnitt des Elektrophoresegels zur Identifikation von <i>pHT4</i>	67
Abb. 22 Schematische Darstellung des Testvektors <i>pHT4</i>	67
Abb. 23 <i>CLSM</i> -Aufnahme der <i>Egfp</i> -Emission von einer mit <i>pHT2</i> und den <i>gfp</i> -spezifischen <i>TALEN</i> -kodierenden Vektoren Co-transformierte Gerstenzelle.....	69
Abb. 24 <i>CLSM</i> -Aufnahmen von Gerstenzellen, die mit <i>pHT3</i> und den <i>cenH3a</i> -spezifischen <i>TALEN</i> -kodierenden Vektoren Co-transformiert wurden	72
Abb. 25 <i>CLSM</i> -Aufnahmen von Gerstenzellen, die mit <i>pHT4</i> und den <i>mlo</i> -spezifischen <i>TALEN</i> -kodierenden Vektoren Co-transformierten wurden.	74

Abkürzungsverzeichnis

Abkürzung	Beschreibung
A	Adenin
Abb.	Abbildung
Amp/ amp _R	Ampicillin/ -resistenz
AS	Aminosäure
Au	Gold
bp	Basenpaar
bspw.	beispielsweise
bzgl.	bezüglich
bzw.	beziehungsweise
C	Cytosin
CaCl ₂	Calciumdichlorid
Cas	<i>CRISPR-associated protein</i>
<i>cenH3α</i>	<i>centromeric histone 3α</i>
<i>CLSM</i>	<i>confocal laser scanning microscopy</i>
<i>CRISPR</i>	<i>clustered regularly interspaced short palindromic repeat</i>
cv.	cultivar
DNA	Desoxiribonukleinsäure
dNTPs	Desoxynukleosidtriphosphate
ds	Doppelstrang
DSB	Doppelstrangbruch
<i>E. coli</i>	<i>Escherichia coli</i>
EDV	elektronische Datenverarbeitung
<i>Egfp/ gfp</i>	<i>Enhanced green fluorescence protein</i>
engl.	englisch
<i>et al.</i>	und andere
EtOH	Ethanol
FokI/ FokIR	<i>Flavobacterium okeanokoites</i> Restriktionsendonuklease I
G	Guanin
g, mg, µg	Gramm, Milligramm, Mikrogramm
GE	<i>genome engineering</i>
ggf.	gegebenenfalls
gRNA	<i>guide RNA</i>
H	Wasserstoff
h	Stunde
H ₃ BO ₃	Borsäure
HDR	<i>homology directed repair</i>
HE	<i>homing endonuclease</i>
His	Histidin
<i>Hv</i>	<i>Hordeum vulgare</i>
i.d.R.	in der Regel
<i>in Hg</i>	<i>inches of mercury</i>

<i>InDel</i>	<i>Insert-Deletion</i>
kbp	Kilobasenpaare
KCl	Kaliumchlorid
l, ml, μ l	Liter, Milliliter, Mikroliter
<i>LB</i>	<i>lysogeny broth</i>
m, cm, mm, μ m, nm	Meter, Centimeter, Millimeter, Mikrometer, Nanometer
M, mM, μ M	Molar, Millimolar, Mikromolar
mA	Milliamper
Mbp	Megabasenpaare
MgCl ₂	Magnesiumdichlorid
MgSO ₄	Magnesiumsulfat
min.	Minuten
<i>mlo</i>	<i>mildew resistance locus</i>
N	beliebige Nukleotidbase
n	ganze Zahl
NaCl	Natriumchlorid
nfW	nukleasefreies Wasser
<i>NHEJ</i>	<i>nonhomologous end-joining</i>
<i>NLS</i>	<i>nuclear localization signal</i>
<i>PAM</i>	<i>protospacer adjacent motif</i>
OD	Optische Dichte
PCR	Polymerase-Kettenreaktion
<i>psi</i>	<i>pound per square inch</i>
<i>pv.</i>	<i>pathovar</i>
RGEN	<i>RNA-guided endonuclease</i>
RNA	Ribonukleinsäure
<i>rpm</i>	<i>revolutions per minute</i>
RT	Raumtemperatur
<i>RVD</i>	<i>repeat variable diresidue</i>
s	Sekunden
S.	Seite
<i>Sc</i>	<i>Secale cereale</i>
<i>SDSA</i>	<i>synthesis-dependend strand annealing</i>
sog.	sogenannten
<i>SP</i>	<i>Signalpeptid</i>
Spec/ spec _R	Spectinomycin/ -resistenz
ss	Einzelstrang
<i>SSA</i>	<i>single strand annealing</i>
T	Thymin
t NOS	Nopalinsynthase-Terminator
<i>Ta</i>	<i>Triticum aestivum</i>
Tab.	Tabelle
<i>TALE/ TALEN</i>	<i>Transcription activator-like effector/ nuclease</i>
TBE	Tris-Borat-EDTA
T-DNA	Transfer-DNA

U	<i>Units</i>
u.a.	unter anderem
u.U.	unter Umständen
<i>Ubi-P/ Ubi-P-Int</i>	<i>UBIQUITIN</i> -Promotor/ mit Intron
UV	Ultraviolett
v	<i>volume</i>
V	Volt
VE	vollentsalzt
Vgl.	Vergleich
w	<i>weight</i>
X.	<i>Xanthomonas</i>
z.B.	zum Beispiel
z.T.	zum Teil
ZF/ ZFN	Zinkfinger/ -Nuklease
<i>Zm</i>	<i>Zea maize</i>
Zn	Zink
μmol	Mikromol
λ	Wellenlänge Lambda
°C	Grad Celsius

1 Theoretische Grundlagen

1.1 Einleitung

Das natürliche Bestreben des Menschen seine Umwelt zu erkunden, sie zu beobachten und zu verstehen, aber auch sie gezielt nach seinen Vorstellungen zu gestalten, zu organisieren und zum eigenen Vorteil zu nutzen, ermöglichten es ihm sich zu einer der erfolgreichsten Spezies des Planeten zu entwickeln. Mit dem Aufkommen der einfachen Landwirtschaft während des Neolithikums vor etwa 11.000 Jahren gingen erste Domestizierungen einher, die die einfachste Form der Züchtung zur Nutzbarmachung pflanzlicher und tierischer Wildarten durch Selektion und Vermehrung und damit erste Eingriffe in die natürliche Entwicklung von Lebewesen darstellen. Neben verschiedenen Nutztieren spielten dabei vor allem Süßgräser (*Poaceae*) wie z.B. Teosinte eine zentrale Rolle zur Sicherung der Nahrungsversorgung. Durch anthropogene Zuchtwahl konnten über die Jahrhunderte hinweg einige der heute bedeutendsten und verbreitetsten Formen von Kulturpflanzenarten, wie z.B. Weizen, Gerste, Reis und Mais, aus oftmals unscheinbaren Wildgräsern erzeugt werden (1-3). Auch die traditionelle Kreuzungszüchtung nutzte frühzeitig und ohne grundlegende Kenntnisse Vererbungsprinzipien die zufälligen genomischen Neuverteilungen zwischen verwandten und sexuell kompatiblen Pflanzen, um deren nützliche Eigenschaften in Nachkommen neu zu kombinieren (1). Die zunehmenden Erfordernisse an effizientere, ertragreichere Getreidesorten führten seit Beginn des 20. Jahrhunderts zur Entwicklung neuartiger Zuchtmethoden wie der Hybridzüchtung und der chemisch oder radioaktiv induzierten Mutagenese zur Generierung vorteilhafter, in der Natur nicht auftretender Merkmalsausprägungen (1, 4). Diese Verfahren wirken grundsätzlich unspezifisch und mehr oder weniger zufällig. Die resultierenden Veränderungen des Genoms sind dabei nicht vorhersehbar. Konventionelle Anwendungen erfordern daher in der Regel große Mutantenpopulationen, langwierige Analysen und Selektionen, sowie Rückkreuzungen zur Eliminierung ungewollter Hintergrundmutationen. Die Entwicklung moderner Zellkulturmethoden und molekularbiologischer Verfahren führte seit den 1980iger Jahren zu bedeutenden Fortschritten in der Pflanzenzüchtung. Molekulare Marker- und Transformationstechnologien ermöglichen die Identifizierung kausaler Mutationen und verkürzen den Selektionsprozess erheblich (2-4). Seit einigen Jahren stehen Wissenschaftlern und Züchtern nun Techniken zur Verfügung, mit denen sich ortsspezifische und nahezu beliebige Modifikationen im Genom eines Organismus vornehmen lassen. Die Möglichkeiten reichen dabei vom gezielten Austausch einzelner Basenpaare (bp), über die Integration von

gut charakterisierten, kodierenden und regulatorischen DNA-Sequenzen in frei wählbaren genomischen Loci bis hin zur Entfernung ganzer Chromosomenabschnitte. In Kombination mit geno- und phänotypischen Analysen wird diese Technologieplattform allgemein als *genomisches Ingenieurwesen* (engl.: *genome engineering*, GE) oder auch *Genomeditierung* (engl.: *genome editing*) bezeichnet (5). Die Optimierung und Erweiterung der Verfahren des GE können als Schlüssel zur schnellen, problemorientierten Züchtung neuer Kultursorten angesehen werden. Sie stellen eine neue Qualität in einer Jahrtausende andauernden Entwicklung dar und ermöglichen Lösungen für komplexe Aufgaben, Probleme und Herausforderungen bezüglich der Versorgung der Weltbevölkerung mit Nahrungsmitteln und im Besonderen einen effektiven Beitrag zur Bekämpfung von Hunger und Mangelernährung. Neben dem enormen Potential für die moderne Landwirtschaft eröffnen die Verfahren des GE zudem neuartige molekularbiologische Ansätze und Methoden in vielen Bereichen der Grundlagenforschung, die bereits heute zur Gewinnung von wissenschaftlichen Erkenntnissen genutzt werden.

1.2 Genomveränderung mit programmierbaren Endonukleasen

1.2.1 Grundlegende Mechanismen

Zwei bedeutende Erkenntnisse rückten zu Beginn der 1990er Jahre, die Sequenz-spezifisch DNA-spaltenden Reagenzien in den Interessensfokus des GE. Zum einen erkennen Zellen Doppelstrangbrüche (DSB) in ihrer DNA als potentiell letalen Schaden und nutzen vielfältige Mechanismen ihres endogenen Reparatursystems zu deren Behebung (6, 7). Zum anderen kann die Häufigkeit von homologen Rekombinationsereignissen in eukaryotischen Zellen durch die Generierung von DSB um mehrere Größenordnungen gesteigert werden (8, 9). Die gezielte Spaltung chromosomaler Doppelstrang- (ds) DNA an präzise definierten Stellen in einem Genom bildet damit eine interessante und neuartige Grundlage zur Erzeugung lokaler Mutationen bzw. Modifikationen unter Ausnutzung der zellulären DSB-Reparaturverfahren, insbesondere der nichthomologen End-zu-End-Verknüpfung und der Homologie-basierten DNA-Reparatur (engl.: *homology directed repair*, HDR) (7, 9, 10).

1.2.1.1 Nichthomologe End-zu-End-Verknüpfung

Die erste bedeutende, auch als *illegitimate Rekombination* bezeichnete, Form der DNA-Reparatur stellt den primären Mechanismus höherer eukaryotischer Zellen zur Behebung von DSB dar. Der Begriff der *nichthomologen End-zu-End-Verknüpfung* (engl.: *nonhomologous end joining*, NHEJ) umfasst alle daran beteiligten Prozesse (11, 12). NHEJ tritt in allen

Stadien des Zellzyklus auf und erfolgt im Allgemeinen ohne Beteiligung homologer Sequenzen (13). Es handelt sich um vergleichsweise schnelle, jedoch inakkurate Prozesse, bei denen der Erhalt ursprünglicher Sequenzinformationen keine hinreichende Berücksichtigung findet und welche daher zum Teil zu verschiedenartigen Mutationen an der Bruchstelle führen (12, 14). Molekularbiologische Studien belegen deutliche Parallelen zwischen pflanzlicher DSB-Reparatur mit den gut untersuchten Mechanismen des NHEJ und der Homologiebasierten Reparatur in Hefe- und Säugetierzellen (15, 16). Anhand beteiligter Reaktionspartner und der resultierenden, charakteristisch ausgeprägten Sequenzmuster an den Verknüpfungsstellen konnten in verschiedenen Pflanzenarten bislang zwei konkurrierende Formen des NHEJ nachgewiesen werden (15, 17, 18). Die Existenz mindestens eines weiteren, der *syntheseabhängigen Stranghybridisierung* (engl.: *synthesis-dependent strand annealing, SDSA*; siehe S. 16) -ähnlichen Mechanismus' gilt als nahezu gesichert (13, 19).

Zentrales Element des *klassischen NHEJ* (c-NHEJ) bilden die Heterodimer-Untereinheiten *Ku70* und *Ku80*, welche in der Lage sind die ds-Enden eines DSB zu erkennen, gezielt an diese zu binden und sie vor Degradierung zu schützen (20). Durch die Dimerisierung der Ku-Protein-DNA-Komplexe werden die Stränge räumlich zusammengeführt und unter Beteiligung von DNA-Ligase IV verknüpft (S. 15, Abb. 1) (14, 21). In der überwiegenden Mehrheit der Fälle erfolgen dabei geringfügige Prozessierungen der Bruchenden durch Exonukleasen oder Polymerasen vor der Religation, die z.T. in kleineren Insertionen und/oder Deletionen von Nukleotiden (*InDels*) innerhalb der ursprünglichen DNA-Sequenz an der Bindungsstelle resultieren (11, 13). Die Integration von Sequenzkopien anderer genomischer Abschnitte mittels c-NHEJ ist vergleichsweise selten, aber ebenso belegt (17, 22).

Der zweite mögliche Verlauf wird als *alternative NHEJ* (a-NHEJ) oder auch *back-up NHEJ* (b-NHEJ) bezeichnet (13, 23). Aufgrund der großen Analogie zum Modell der *Einzelstranghybridisierung* (engl.: *single strand annealing, SSA*) der HDR wird a-NHEJ von vielen Autoren auch als SSA-ähnlicher Mechanismus behandelt (Vgl. S. 16; S. 15, Abb. 1) (24). Im Falle exonukleolytisch aktiver Enzymkomplexe kann es zur Bildung erweiterter 3'-Überhänge an den Enden eines DSB kommen, in deren Folge eine Hybridisierung einzelner, komplementärer Nukleotide (≥ 2 bp bis 25 bp) innerhalb dieser Einzelstrang- (engl.: *single strand, ss*) Abschnitte eintreten kann. Nach enzymatischer Abtrennung ggf. überstehender, nichtkomplementärer 3'-Enden und DNA-Reparatursynthese zur Komplementierung partieller ss-Regionen erfolgt die Ligation der zu verknüpfenden Enden, wobei sämtlicher

Informationsgehalt zwischen Bruchstelle und Hybridisierung verloren geht. In Folge dessen zeichnen sich resultierende Verknüpfungen typischerweise durch Mikrohomologien in Begleitung mit Sequenzdeletionen in variierenden Umfang aus (13, 15, 16, 18).

Während der DSB-Prozessierung durch NHEJ auftretende Insertionen und Deletionen können bis zu 1,2 kbp umfassen (22) und zu einer Verschiebung im Leseraster der Basentriplettabfolge der DNA (engl.: *frame-shift*) führen. Diese Mutationen bewirken eine umfassende Änderung der in Leserichtung folgenden Aminosäuresequenz kodierter Proteine und resultieren daher zumeist in einer funktionellen Stilllegung, dem sog. *Knockout*, eines Gens. Zudem enthält die resultierende *non-sense* Transkriptionssequenz zumeist zufällig entstehende transkriptionsbeendende Nukleotidtripletts (*Stopp-Codons*). Auch sog. *In-Frame*-Mutationen, bei denen der Gesamtumfang der Veränderung drei bp oder ein Vielfaches davon beträgt, können die natürliche Funktion eines Proteins durch Auswirkungen auf dessen räumliche Struktur oder spezifische funktionelle Domänen beeinträchtigen oder auch zur Stilllegung des Gens, u.a. ebenfalls durch ein zufällig entstehendes Stopp-Codon, führen (7, 10). Sollten mehrere DSB zeitgleich auftreten kann es neben großen Sequenzdeletionen auch zu komplexeren, genetischen Variationen in Form von DNA-Inversionen oder -Translokationen innerhalb eines oder zwischen verschiedenen Chromosomen kommen (7, 25). Ebenso kann eine NHEJ-vermittelte Integration artfremder DNA-Sequenzen, wie z.B. T-DNA von *Agrobacterium tumefaciens*, in ein Pflanzengenom erfolgen (26).

1.2.1.2 Homologie-basierte DNA-Reparatur

Ein zweites bedeutendes Prinzip der DSB-Reparatur ist die *homologe Rekombination*. Sie erfordert die Einbeziehung einer intakten DNA-Sequenzen mit Homologie zum defekten genomischen Locus und ist ein zentraler Bestandteil der sexuellen Rekombination höherer eukaryotischer Zellen während der Meiose (27). Davon abgesehen sind Homologie-basierte DNA-Reparaturen in somatischen Zellen von geringerer Bedeutung und kommen dort seltener vor als NHEJ-Ereignisse (etwa $1:10^3$ bis $1:10^7$) (12, 28). Die Häufigkeit und Effizienz variiert in Abhängigkeit von der Pflanzenspezies, dem Status des Zellzyklus, sowie dem Vorkommen von DSB und der Verfügbarkeit einer homologen Reparaturvorlage. Die Rekombinationsfrequenz wird zudem durch die Art des DNA-Homologs, sowie dessen räumliche Nähe zum DSB beeinflusst (12, 22). Als natürliche Reparaturvorlage können (*ektopische*) Sequenzkopien von anderen Orten im Genom, (*allelische*) Sequenzen des homologen Chromosoms, Sequenzen des gleichen Chromosoms (*intrachromosomale Rekombination*) oder in der späten S- und G₂-Phase des Zellzyklus auch des

Schwesterchromatids verwendet werden (12, 18). Zudem können artfremde DNA-Sequenzen mit partieller Homologie als Matrize dienen (9, 17). Es existieren mehrere Modelle anhand derer sich die vielfältigen genomischen Neuordnungen, welche durch die Prozesse der HDR vermittelt werden, erklären lassen (15). Zu den in somatischen Pflanzenzellen am umfassendsten beschriebenen Mechanismen der homologen Rekombination zählen die *Einzelstranghybridisierung* (29, 30) und die *syntheseabhängige Stranghybridisierung* (31, 32). Für meiotisch aktive Zellen ist zudem das klassische *Doppelstrangbruchreparatur-Modell* (DSBR) charakteristisch (13, 33). Allen gemeinsam ist die Erkennung und Bindung der DSB-Enden durch enzymatisch aktive Proteinkomplexe wie MRE11/RAD50/NBS1, in dessen Folge es zur ATM-Kinase-Signalkaskade-vermittelten und exonukleolytisch katalysierten Generierung von 3'-ss-Überhängen an den Strangenden kommt (Abb. 1)(16, 27, 34).

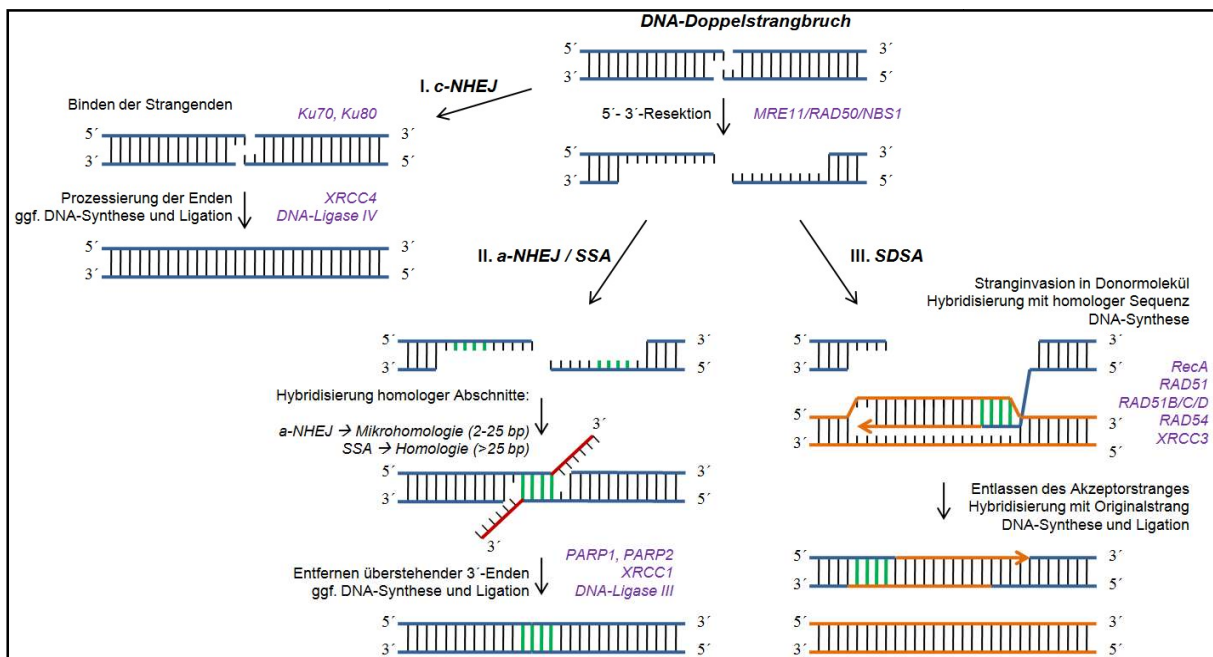


Abb. 1 | Übersicht der wesentlichen Modelle von DSB-Reparaturmechanismen in somatischen Pflanzenzellen. I.) **Klassische nichthomologe End-zu-End-Verknüpfung (c-NHEJ):** Detektion und Bindung der ds-Enden eines DNA-Doppelstrangbruchs (DSB) durch Ku70 und Ku80. Mit der Dimerisierung der Ku-Protein-DNA-Komplexe werden die Bruchenden zusammengeführt. Vor der Ligation kommt es z.T. zu enzymatischen Prozessierungen der Strangenden, welche u.U. in *InDels* resultieren. In Konkurrenz zur c-NHEJ erfolgt eine exonukleolytisch-katalysierte Generierung von 3'-ss-Überhängen an den DSB-Enden. II.) **Alternatives NHEJ (a-NHEJ) und Einzelstrang-Hybridisierung (SSA):** Nach 5'-Resektion der Strangenden kommt es zur Hybridisierung komplementärer Nucleotide bzw. Sequenzen (grün) innerhalb der gebildeten ss-Abschnitte. a-NHEJ zeichnet sich dabei durch Verwendung gegebener Mikrohomologien von ≥ 2 bis 25 bp, SSA durch Sequenzhomologien über 25 bp aus. Ggf. erfolgt eine katalysierte Trimmung nicht-komplementärer 3'-Überhänge (rot) und die Komplementierung von ss-Lücken, bevor die Strangenden enzymatisch verknüpft werden. III.) **Syntheseabhängige Stranghybridisierung (SDSA):** Eindringen eines 3'-Überhangs des DSB-Akzeptormoleküls in ein DNA-Donormolekül (orange) unter Ausbildung einer D-Schleufe und Hybridisierung homologer Sequenzabschnitte (grün, oben). Nach komplementärer Erweiterung des Akzeptorstranges durch DNA-Synthese entlang des Donorstrangs (Pfeil, oben) wird der z.T. neu synthetisierte 3'-Einzelstrang aus dem Donormolekül entlassen. Es kommt zur Hybridisierung mit dem Ausgangsstrang (grün, unten), erneuter DNA-Synthese zur Komplementierung (Pfeil, unten) und der Ligation des Strangs. Violett: Auswahl von in Pflanzen identifizierten und an den Reaktionen essentiell beteiligten Proteinen. (Darstellung abgewandelt nach Vu et al., 2014; Gorbunova & Levy, 1999)

Die *Einzelstranghybridisierung* stellt die häufigste Form der HDR in somatischen Zellen dar (12). Die sehr effizienten enzymatischen Prozesse verlaufen dabei prinzipiell analog zum a-NHEJ (Vgl. S. 13), wobei der wesentliche Unterschied zwischen beiden Modellen einzig auf der Größenordnung der homologen Region beruht (> 25 bp)(22). SSA tritt typischerweise an DSB zwischen DNA-Sequenzwiederholungen auf, welche sich in einer räumlich nahen, hintereinander liegenden Orientierung zueinander befinden. Nach 5'-Resektion der Strangenden kann eine Hybridisierung zwischen den gebildeten, partiell komplementären ss-Abschnitten zweier Wiederholungen erfolgen (S. 15, Abb. 1). Aufgrund des resultierenden Informationsverlusts zwischen den homologen Sequenzen gilt die SSA als nicht konservative HDR-Reaktion (12, 13).

Die *syntheseabhängige Stranghybridisierung* zeichnet sich im Gegensatz dazu als Hauptmechanismus für konservative HDR in somatischen Pflanzenzellen aus. Beim klassischen Verlauf erfolgt eine genetische Konversion durch die Übertragung von Sequenzinformationen durch eine Strangaustauschreaktion von einem intakten, homologen DNA-Donormolekül zum DSB des Akzeptors ohne Informationsverluste in der Donorsequenz. Durch die Generierung von ss-Überhängen an den Bruchenden kann es zum Eindringen eines freien 3'-Endes in die homologe ds-Kopiervorlage unter Ausbildung einer D-Schleife kommen (S. 15, Abb. 1). Nach Anlagerung der komplementären Sequenzabschnitte erfolgt die DNA-Synthese entlang des neu gebildeten Strangs. Im Anschluss an die Verlängerung des Akzeptors kommt es zur Freisetzung des Einzelstrangs aus der Schleife und der anschließenden Hybridisierung mit komplementären Basen des zweiten, freien 3'-ss-Überhangs des DSBs. Ist die Sequenz der Kopiervorlage mit der zu reparierenden Sequenz identisch, kann auf diese Weise die ursprüngliche genetische Information des gebrochenen Doppelstrangs wieder hergestellt werden. Im Gegensatz zum DSBR-Modell erfolgt während der SDSA keine Ausbildung einer *Doppel-Holliday*-Struktur, welche zu unvorhersehbaren genomischen Änderungen durch *DNA-Crossover* führen kann (12, 13, 35).

Die Abläufe pflanzlicher DSB-Reparaturen und deren beteiligter Reaktionspartner sind nicht vollkommen aufgeklärt und weiterhin Gegenstand der Forschung. Obwohl die Mechanismen als konserviert gelten, kommt es z.T. zu großen, artspezifischen Unterschieden bezüglich Vorkommen und Effizienz verschiedener, in unterschiedlichem Maße bevorzugt zum Tragen kommender Reparaturwege. Auch eine strikte Trennung der einzelnen Prozesse ist nicht immer möglich, da diese nicht nur miteinander um die DSB-Reparatur konkurrieren,

sondern sich ggf. auch ergänzen und kooperieren. Zum Beispiel kann es im Fall nicht kompatibler 3'-ss-Enden nach der Akzeptorstrangsynthese der SDSA zu einer NHEJ-vermittelten Ligation der DSB-Enden kommen. Die resultierenden genetischen Neuordnungen der pflanzlichen DNA-Reparaturen sind wesentlich vielfältiger und komplexer als die anderer Gruppen von Lebewesen und tragen daher in weitaus stärkerem Maße zu deren Genomentwicklung bei (12, 36).

1.2.1.3 DSB-induzierte, ortsspezifische Mutagenese

Die einfachste Anwendung der Genommodifikation mittels Sequenz-spezifisierbarer Endonukleasen ist die *zielgerichtete Mutagenese* (engl.: *site-directed mutagenesis*). Durch Einführung eines DSB an einer zur Mutation vordefinierten Position im Genom werden die zellulären DNA-Reparaturprozesse initiiert. Bei Pflanzen erfolgt dabei zumeist eine NHEJ-vermittelte Ligation der Bruchenden. Aufgrund der potentiellen Fehleranfälligkeit dieses Reparatursystems resultieren solche DSB-Induktionen mit gewisser Häufigkeit in *InDel*-Mutationen. Treten diese in proteinkodierenden Abschnitten auf kann sich das Leseraster der betroffenen Gensequenz nach der Bruchstelle verschieben (*frame-shift*), was in der Regel zu einem funktionellen Ausfall führt (Vgl. S. 14)(5). Durch individuelle Anpassung der Nuklease kann dies für einzelne Gene, aber auch zum simultanen Ausschalten ganzer Genfamilien genutzt werden (37). Zudem können durch *In-Frame*-Mutationen einzelne, funktionell bedeutende Aminosäuren eines Proteins oder einer Proteindomäne gezielt manipuliert werden (38). Neben der Abänderung kodierender Sequenzen kann eine gesteuerte Mutation regulierender DNA-Regionen, wie z.B. Bindestellen für Pathogenitätsfaktoren in Promotoren, erfolgen (39). Auch das Entfernen gesamter Gensegmente mit flankierenden Erkennungssequenzen für entsprechend entworfene Endonukleasen ist möglich, wodurch sich bspw. transgene Selektionsmarker aus einem Genom eliminieren lassen (40, 41). Der Erfolg dieser Anwendungen hängt maßgeblich von der Aktivität der eingesetzten Nukleasen ab und kann zudem durch Manipulation der Enzymmaschinerie der einzelnen DSB-Reparaturwege beeinflusst werden (42).

Wesentlich weitreichendere Genommodifikationen eröffnen die Anwendungen der *Homologie-basierten Genomeditierung* (engl.: *homology-directed genome editing*). Statt zufälliger Sequenzänderungen erfolgt dabei eine gezielte Manipulation unter Nutzung der natürlichen DSB-Reparaturmechanismen der HDR. Durch das Einbringen künstlicher DNA-Reparaturvorlagen mit gewünschter Nukleotid- bzw. Gensequenz und flankierenden Homologiearmen, deren Sequenzen passend zu den Enden eines induzierten DSB gestaltet

werden, können nahtlose genetische Modifikationen an theoretisch jeder Position im Genom vorgenommen werden. Dies ermöglicht ein gezieltes Einfügen, Entfernen, Korrigieren oder Ersetzen einzelner Basenpaare, funktioneller Domänen und Gene oder auch komplexer Expressionskassetten an einem definierten genomischen Locus. Aufgrund des Wettbewerbs mit dem ebenfalls homologen Chromosom sind dabei die Menge an bereitgestellter DNA-Reparaturvorlagen und der Umfang der enthaltenen Sequenzhomologien von entscheidender Bedeutung für den Erfolg der gewünschten Rekombination. Neben der enzymatischen Aktivität der Nuklease hat auch die Konkurrenz zur bevorzugten DSB-Reparatur mittels NHEJ einen großen Einfluss auf die Frequenz der HDR. Durch Manipulation der Enzymmaschinerie einzelner Prozesse kann diese Effizienz gezielt beeinflusst und gesteigert werden (5, 38, 42, 43).

Die auf Zellebene resultierenden genetischen Änderungen der zielgerichteten Mutagenese und der Reparaturvorlage-vermittelten GE können ggf. auf die Keimbahn bzw. Gameten übertragen und damit weiter vererbt werden (44, 45).

1.2.2 Sequenz-spezifisierbare, synthetische Endonukleasen

Für die gezielte Induktion von DNA-Doppelstrangbrüchen innerhalb gewünschter Nukleotidsequenzen stehen verschiedene Plattformen von Sequenz-spezifisierbaren Endonukleasen zur Verfügung, deren Bindestrukturen sich mehr oder weniger gut umgestalten lassen um sie an neue, vom Nutzer definierte DNA-Zielsequenzen anzupassen. Anhand der Struktur und den Mechanismen zur Erkennung und Spaltung von DNA werden derzeit vier Klassen von synthetisierbaren Endonukleasen unterschieden, mit denen ortsspezifische, genetische Modifikationen in Pflanzengenomen initiiert und katalysiert werden können.

1.2.2.1 Meganukleasen

Die erste Art von Sequenz-spezifisch DNA-schneidenden Reagenzien, welche als Vermittler für gezielte Rekombinationsprozesse in Pflanzen zur Anwendung kamen, bilden die sogenannten *Meganukleasen* (engl.: *meganucleases* oder *homing endonucleases*, HE)(6, 46). Dabei handelt es sich um natürlich vorkommende Restriktionsendonukleasen, die meist als selbstspleißendes Element auf mobilen, intervenierenden Klasse I und II Introns oder Inteinen kodiert sind (47). Sie agieren typischerweise als Homodimere oder auch Monomer mit zwei separaten Enzymdomänen und erkennen verhältnismäßig große DNA-Motive von ≥ 12 bis 45 bp mit hoher Spezifität (48). Innerhalb dieser katalysieren sie einen DSB um den gezielten

Integrationsprozess des mobilen genetischen Elements, auch *homing* genannt, oder andere mögliche Rekombinationsereignisse zu initiieren (46, 47). Bislang konnten über 250 HE mit einer Vielzahl verschiedener Erkennungssequenzen in Eukaryoten, Bakterien und Archaeen identifiziert werden (47, 49, 50). Sie zählen zu den kleinsten Reagenzien des GE, eine Monomereinheit umfasst durchschnittlich nur etwa 165 AS. Dies ermöglicht ein einfaches Einbringen in Zellen mit fast allen derzeit verfügbaren Transformationsmethoden (siehe S. 35), was speziell im Fall begrenzter Vektorkapazitäten einen entscheidenden Vorteil darstellt (38). Ein wesentlicher Unterschied zu anderen Sequenz-spezifisierbaren Endonukleasen ist die strukturelle Überlagerung der DNA-Binde- und der Restriktionsdomäne. Eine Änderung der AS-Sequenz der Bindedomäne kann sich direkt auf die katalytische Enzymaktivität auswirken, was eine künstliche Umgestaltung zur Generierung neuer Sequenz-Spezifitäten wesentlich komplizierter und anspruchsvoller macht (48, 50). Die Erzeugung artifizieller Meganukleasen zur Erkennung neuer DNA-Zielsequenzen wurde zwar mehrfach erfolgreich umgesetzt, ist wegen der Komplexität der kaum modularen Proteinstrukturen jedoch noch immer eine Herausforderung und in Folge dessen für routinemäßige GE nicht geeignet (5, 50-52). Aufgrund des Umfangs der spezifischen Erkennungssequenz ist die Anzahl möglicher Bindungsstellen in einem Pflanzengenom äußerst gering. Häufig existieren keine oder lediglich eine bis ein paar wenige entsprechende DNA-Motive, wodurch auch das Einsatzgebiet natürlicher HE limitiert wird (48). Bisherige Anwendungen erfolgten daher vorwiegend mit selektierten Derivaten von natürlich vorkommenden HE, wie z.B. *I-SceI*, *I-CreI* und *HO*, oder durch vorherige Integration einer transgenen Erkennungssequenz in das Wirtsgenom (52, 53). Trotz der erwähnten Probleme konnten mit Meganukleasen diverse ortsspezifisch DSB-induzierte Rekombinationsereignisse von endo- und transgenen DNA-Sequenzen in Acker-Schmalwand (*Arabidopsis thaliana*) (41, 54), Tabak (*Nicotiana tabacum*) (9, 17, 55), Baumwolle (*Gossypium*) (56) und Mais (*Zea mays*) (44, 51, 57) stimuliert und katalysiert werden, darunter NHEJ- und HDR-vermittelte Fremd-DNA-Integrationen, umfangreiche Sequenzdeletionen und gezielte Genstilllegungen.

1.2.2.2 Zinkfinger-Nukleasen

Zinkfinger-Nukleasen (ZFN) stellen die ersten tatsächlich künstlich generierten Sequenz-spezifisierbaren Endonukleasen zur Erkennung und Spaltung von beliebigen nutzerdefinierten Abschnitten in Chromosomen dar (28, 38). Es handelt sich um chimärische Restriktionsenzyme aus zwei separaten und funktionell verschiedenen Domärentypen. Sie setzen sich aus einer N-terminalen, modularen DNA-Bindedomäne basierend auf Zinkfinger-

(ZF) Proteinkomplexen und einer daran fusionierten C-terminalen, katalytischen FokIR-Domäne zusammen (58, 59).

FokI-Enzyme sind natürlich vorkommende Typ-II-Restriktionsendonukleasen aus *Flavobacterium okeanoikoites*, welche ds-DNA im präzisen Abstand von neun bzw. 13 Nukleotiden strangabwärts, außerhalb ihrer N-terminalen fünf bp-Erkennungssequenz spaltet und dabei adhäsive 5'-Überhänge generiert (60). Die dafür zuständige 196 AS umfassende und ebenfalls C-Terminal gelegene FokIR-Domäne besitzt generell keine Sequenz-Spezifität, jedoch ist eine Dimerisierung zur Induktion der katalytischen Aktivität erforderlich (61). Wildtyp-FokI bildet natürlicherweise homodimere Strukturen und sind u.U. in der Lage, bereits bei Bindung nur eines Monomers an einen DNA-Abschnitt zu dimerisieren und einen DSB zu prozessieren (62, 63). Die erfolgreiche Isolation der Restriktionsdomäne unter Aufrechterhaltung der enzymatischen Aktivität eröffnete letztlich eine Anwendung als Fusionsprotein in Kombination mit verschiedenen alternativen DNA-Bindedomänen (58, 61).

Natürliche Zinkfinger-Domänen sind multimere Sequenzbindungsaggregate aus einer Aneinanderreihung einzelner ZF-Proteinkomplexe (64, 65). Als ein zentraler Bestandteil vieler eukaryotischer Transkriptionsfaktoren binden sie zumeist spezifische *cis*-Elemente in Promotorregionen ihrer Wirtsgenome und regulieren darüber die Expression verschiedenster Gene (5). Einige Vertreter gehen zudem auch RNA- oder Protein-Protein-Interaktionen ein (66-68). Aufgrund der Vielfalt an natürlich vorkommenden Derivaten werden sie in großen Proteinfamilien zusammengefasst. Eine der umfangreichsten und am besten charakterisierten Klassen sind dabei die *Cys₂His₂*-Motive. Sie zählen zu den häufigsten und weitverbreitetsten Sequenz-spezifischen Transkriptionsfaktoren in höheren Eukaryoten (59, 69). Ähnlich wie Meganukleasen sind auch Zinkfinger-Nukleasen mit durchschnittlich etwa 300 AS pro Monomer relativ klein und eignen sich dadurch gut für die meisten gängigen Transformationsmethoden (siehe S. 35)(53). Die DNA-Bindedomänen bislang synthetisch erzeugter ZFN bestehen zumeist aus drei bis sechs miteinander verknüpften ZF-Proteinen vom Typ *Cys₂His₂*, die Zielsequenz eines einzelnen Poly-Finger-Monomers beträgt damit neun bis 18 bp (7). Aufgrund der erforderlichen Dimerisierung der FokI-Nukleasedomänen sind für die effiziente Spaltung von ds-DNA zwei ZFN-Einheiten mit entsprechend spezifisch gestalteten Erkennungssequenzen notwendig, welche in entgegengesetzter Orientierung und geeigneter Distanz zueinander an jeweils einen Strang der DNA binden (S. 21, Abb. 2)(70). Dadurch verdoppelt sich das effektive Bindungsmotiv auf 18 bis zu 36 bp, was zu einer signifikanten Steigerung der Zielspezifität und einer deutlichen Verringerung unerwünschter

Aktivität im Hintergrund des Wirtsgenoms führt. Der Umfang der Erkennungssequenz eines ZFN-Paars ist für gewöhnlich ausreichend um nur einmal im Genom der meisten höheren Eukaryoten aufzutreten (43, 71). Dennoch ist der Einsatz von ZFN häufiger mit zytotoxischen Effekten durch kaum vermeidbare, nicht zielgerichtete DSB-Induktionen verbunden (52, 59). Zur Verringerung dieses Problems basieren neuere Strategien auf dem Einsatz von ZFN-Paaren mit speziell entwickelten, strikt heterodimeren FokIR-Derivaten oder nutzen proteinregulierende Co-Faktoren, wodurch das spaltende Reagenz nur assembliert und aktiviert wird wenn beide ZFN-Einheiten tatsächlich an ihre Zielsequenzen binden (72-76). Die erforderliche Distanz zwischen zwei gebundenen ZFN variiert dabei geringfügig in Abhängigkeit von ggf. zur Verbindung der funktionellen Domänen zusätzlich in die Sequenzen der einzelnen Monomere eingefügten AS. Der Abstand für eine effektive Dimerisierung der *FokIR*-Domänen beträgt für direkt fusionierte ZFN ohne eine solche AS-Verknüpfung fünf bis sieben bp, optimal sind genau sechs bp (10, 77, 78).

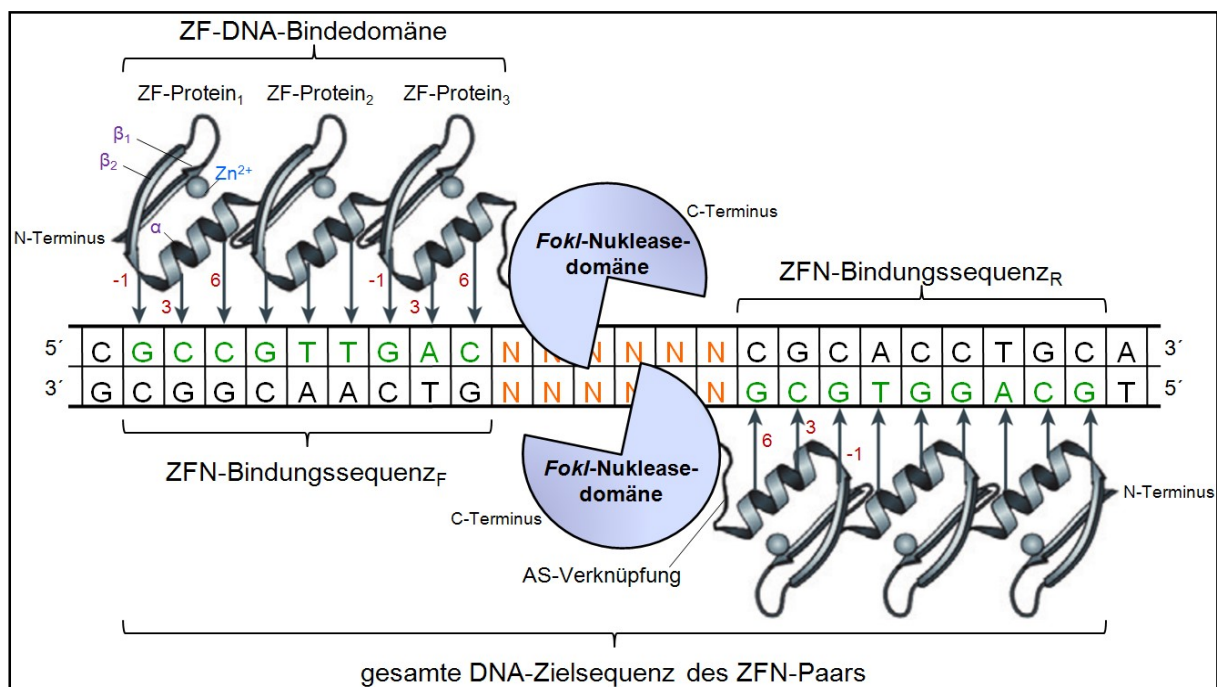


Abb. 2 | Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz durch ein Zinkfinger-Nuklease-(ZFN) Paar zur Induktion eines Doppelstrangbruchs (DSB). Jedes ZFN-Monomer besteht aus N-terminaler DNA-Bindedomäne und C-terminaler FokI-Restriktionsdomäne, welche ggf. durch zusätzliche AS verknüpft sind. Die Bindedomänen setzen sich aus mehreren (hier je drei) variablen und selektiv DNA-interagierenden Zinkfinger- (ZF) Proteinen zusammen. Durch die Art der AS an den Positionen -1, 3, 6 (rot) und 2 (nicht dargestellt) der α -Helix werden Spezifität und Affinität einzelner ZF-Proteine für verschiedene Nukleotidtripletts definiert. Die zu bindende DNA-Sequenz (grün) eines ZFN-Monomers sollte mind. neun bp betragen. Der primäre DNA-Kontakt erfolgt durch hydrophobe Wechselwirkungen und Wasserstoffbrückenbindungen zwischen den genannten AS der einzelnen Helices und aufeinander folgenden Nukleotiden im DNA-Rückgrat (graue Pfeile). Zur DSB-Induktion durch FokIR müssen zwei entsprechend spezifisch gestaltete ZFN-Monomere in entgegengesetzter Orientierung links- und rechtsseitig einer Überbrückungssequenz (orange, N = beliebige Base), zeitgleich an jeweils einen DNA-Strang binden. Dies ermöglicht die Dimerisierung der Nukleasedomänen, in Folge dessen es zur Katalyse eines DSB innerhalb der Überbrückungssequenz kommt. Die Gesamtzielsequenz eines ZFN-Paars (hier 24 bp) variiert in Abhängigkeit von der Anzahl an verwendeten ZF-Proteinen und der Nukleotidsequenz zwischen den ZFN-Domänen (Darstellung abgewandelt nach *Porteus & Carroll, 2005*; Bildquelle ZF-Proteine: *Gaj et al., 2013*)

Im Vergleich zu Meganukleasen (Vgl. S. 19) ist die Erzeugung artifizierter ZFN nicht ganz so anspruchsvoll, dennoch ist die Generierung neuer Zielspezifitäten eine außerordentliche Herausforderung und grundsätzlich mit Unsicherheiten bzgl. der erzielten Funktionalität verbunden (53). *Sigma-Aldrich* (St. Louis, MO, USA) bietet die Generierung von ZFN nach Kundenwunsch an, die auf einem internen, patentierten Protokoll von *Sangamo Biosciences* (Richmond, CA, USA) basieren.

Durch ZFN ortsspezifisch induzierte Genommodifikationen umfassen ein breites Spektrum an z.T. vererblichen Änderungen von endo- und transgenen DNA-Sequenzen in somatischen Pflanzenzellen (28, 53). Dazu zählen u.a. funktionelle Genstilllegungen mit Mutationsraten von bis zu 16 % in Acker-Schmalwand (*A. thaliana*) (79, 80), Tabak (*N. tabacum*), Petunie (*Petunia hybrida*) (81) und Soja (*Glycine max*) (37), umfangreiche Sequenzdeletionen von bis zu neun Mbp in *A. thaliana* (82) und *N. tabacum* (40), sowie DNA-Integrationen oder Sequenzänderungen mit Effizienzen von bis zu zehn Prozent in *A. thaliana* (83, 84), *N. tabacum* (85, 86) und Mais (*Z. mays*) (87, 88). Trotz der zahlreichen, erfolgreichen Anwendungen von ZFN in Modell- und Kulturpflanzenarten beschränkten sich die Veröffentlichungen bislang zumeist auf vielversprechende Erkenntnisse in der Grundlagenforschung und wenige anwendungsorientierte Beispiele zur Überprüfung und Veranschaulichung neuer Konzepte und geltender Prinzipien. Aufgrund der genannten Limitierungen und Nebenwirkungen sind ZFN in den letzten Jahren, auch zugunsten neu entwickelter und effektiverer Plattformen von Sequenz-spezifisierbaren Endonukleasen zunehmend aus dem Fokus der anwendungsorientierten Wissenschaft verschwunden (5).

1.2.2.3 TALE-Nukleasen

Die Entschlüsselung der relativ simplen Kodierung der spezifischen Interaktion zwischen einer Gruppe von hochspezialisierten pflanzenpathogenen Wirkstoffen und der Erbinformation infizierter Wirtszellen führte seit Ende des Jahres 2009 zur Entwicklung und Etablierung einer neuartigen Plattform der Sequenz-spezifisierbaren Endonukleasen (89, 90). Bei *transcription activator-like effector nucleases (TALEN)* handelt es sich ebenfalls um artifizelle Fusionsproteine aus zwei funktionell verschiedenen Proteindomänen, deren Prinzip und generelle strukturelle Organisation mit den zuvor beschriebenen ZFN vergleichbar ist (Vgl. S. 20)(91). Auch *TALEN* nutzen eine C-Terminal lokalisierte katalytische FokI-Restriktionsdomäne (Vgl. S. 20) zur Induktion eines DSB in einem gewünschten DNA-Zielabschnitt. Der wesentliche Unterschied im Vergleich zu ZFN besteht in der N-Terminal gelegenen, modularen DNA-Erkennungs- und Bindungsstruktur, die sich durch eine

charakteristische und hochrepetitive Abfolge von nahezu identischen AS-Sequenzkopien auszeichnet. Jeweils eine AS innerhalb eines jeden dieser Sequenzwiederholungsmodule (engl.: *repeat*) einer *TALEN*-Bindedomäne tritt mit einer einzelnen, individuellen Base eines fortlaufenden DNA-Strangs in Wechselwirkung und vermittelt auf diese Weise den spezifischen Kontakt zur nutzerdefinierten Zielsequenz (89, 90).

Diese neuartige und in der Natur einmalige Klasse variabel gestaltbarer Bindungsmotive ist von Vertretern umfangreicher Familien natürlicher *Avirulenz* (*Avr*) - Proteine abgeleitet, welche erstmals aus dem phytopathogenen Bakterium *Xanthomonas campestris pv. vesicatoria* isoliert wurden (92, 93). Zur Infektion einer befallenen Pflanzenzelle bilden *Xanthomonaden* einen als *Typ-III-Translokation* bezeichneten, transmembranen und kanalähnlichen Porenkomplex, über den sie einen Cocktail aus etwa 20 bis 40 verschiedenen bakteriellen Effektorproteinen in das eukaryotische Zytoplasma sekretieren (94, 95). Die Virulenzfaktoren sind in der Lage, eine Wirtszelle auf sehr vielfältige Art und Weise zu manipulieren. Einige greifen bspw. in verschiedene Stoffwechselwege oder Signalkaskaden ein und sabotieren diese, um mögliche Immunantworten zu unterdrücken oder den Metabolismus der Zelle zum Vorteil des Bakteriums zu beeinflussen (96, 97). Andere Effektorgruppen weisen strukturelle Übereinstimmungen und typische Charakteristiken von eukaryotischen Transkriptionsfaktoren auf. Sie zielen auf eine direkte Manipulation des Wirtstranskriptom ab und werden daher auch als *transcription activator-like effector* (*TALE*) bezeichnet (98). Diese werden mit Hilfe einer N-terminalen Kernlokalisationssequenz (engl.: *nuclear localization signal*; NLS) in den Nukleus der Pflanzenzelle eingeschleust, wo sie hochspezifisch an Promotorregionen von Zielgenen binden und deren Expression über die C-terminale Transkriptionsaktivierungsdomäne regulieren (99, 100). Ein besonders interessanter Aspekt der *TALE*-Proteine ist das zentral gelegene repetitive DNA-Bindungsmotiv. Die mehr als 112 bekannten, natürlich vorkommenden *TAL*-Effektoren verfügen alle über eine hochkonservierte Erkennungsdomäne bestehend aus 1,5 bis 33,5 Kopien von nahezu identischen und direkten Sequenzwiederholungen aus jeweils 30 bis z.T. 42 Aminosäuren (AS)(92, 98). Ein Großteil der nativen *TALE*-Varianten setzen sich aus 17,5 Wiederholungsmodulen mit je 34 AS pro Abfolge zusammen, wobei das typische halbe *Repeat* am C-terminalen Ende der Domänen lediglich die ersten 20 AS aufweist (89). Die meisten der Proteinderivate besitzen strukturelle Übereinstimmungen von mehr als 80 bis zu 97 Prozent ihrer Sequenzen. Die Unterschiede beschränken sich neben der Anzahl und ggf. auch Länge der Kopienabfolgen im Wesentlichen auf einige, weniger konservierte AS

innerhalb der einzelnen Wiederholungsmodule (92, 101). Variationen treten insbesondere an den Positionen zwölf und dreizehn einer jeden Sequenzwiederholung auf. Diese hypervariablen di-AS-Motive werden als *repeat variable diresidues (RVD)* bezeichnet (90). Die Entschlüsselung des spezifischen Interaktionsmechanismus zwischen den Zentraldomänen von *TALE*-Proteinen und DNA-Zielsequenzen offenbarte, dass jedes der *RVD*-Motive innerhalb der Wiederholungsabfolgen eins zu eins mit einer Base im kodierenden Strang einer fortlaufenden Zielsequenz korreliert und die direkte Bindung mit dieser vermittelt (89, 90). Die Kombination aus zwölfter und dreizehnter AS spezifiziert dabei die Art von Nukleotidbase, die gebunden wird. Einige der *RVD*-Motive besitzen eine starke bis exklusive Präferenz für eine der vier kodierenden Basen und favorisieren deren Bindung, während andere flexibler oder sogar nichtselektiv interagieren und Bindungen mit verschiedenen Basentypen tolerieren. Nahezu alle bekannten DNA-Zielsequenzen von natürlichen *TALE*-Proteinen beginnen typischerweise mit einer konservierten und als T₀ deklarierten Thyminbase am 5'-Ende, welche einen direkten Einfluss auf die spezifische Wahrnehmung von Zielsequenzen hat (102).

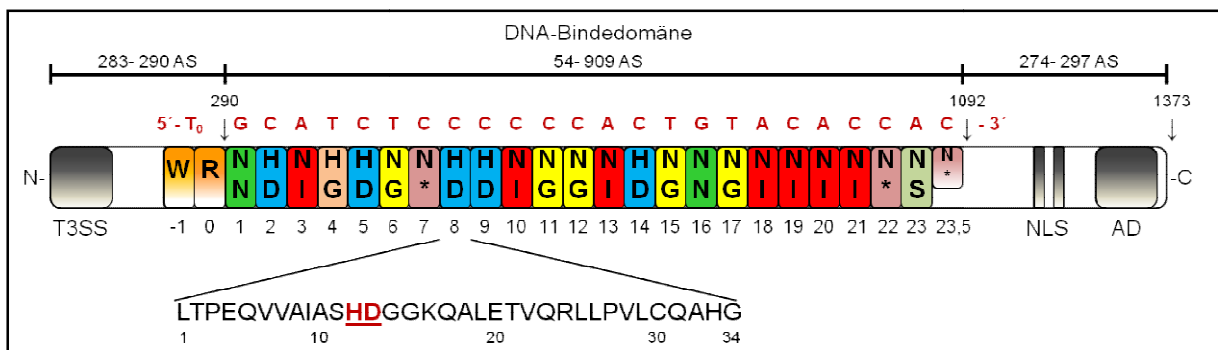


Abb. 3 | Schematische Struktur eines *TAL*-Effektorproteins am Beispiel *PthXo1* aus *X. oryzae*. Charakteristisches Element aller *TALE*-Proteine ist die hochrepetitive, zentrale DNA-Bindedomäne, die sich aus einzelnen Modulen (sog. *Repeats*; hier 23,5) mit nahezu identischen Sequenzen von zumeist je 34 Aminosäuren zusammensetzt (untere Sequenz). Variationen treten vorwiegend an den AS-Positionen zwölf und dreizehn (rot unterstrichen) eines jedes *Repeats* auf, die als *repeat variable diresidues (RVD)* bezeichnet werden. Jedes der *RVD*-Motive korreliert eins zu eins mit einer Base der DNA-Zielsequenz (rote Sequenz oberhalb des Proteins), wobei die AS-Kombination der *RVD* die Art der gebundenen Base spezifiziert. Der Kontakt mit der DNA wird über den -1. und 0. *Repeat* (W, R) in Wechselwirkung mit der T₀-Base am Beginn nahezu jeder Zielsequenz vermittelt. Zu den am häufigsten auftretenden *RVD*-Motiven zählen HD für Cytosin (blau), NI für Adenin (rot), NG für Thymin (gelb) und NN für Guanin oder Adenin (grün). Daneben existieren zahlreiche Varianten mit stark variierenden Spezifitäten und Affinitäten für einzelne Basen. *RVD*'s vom Typ NS (olive) agieren bspw. unspezifisch und können mit allen Arten von Basen Bindungen eingehen (98). T3SS: Typ-III-Segregationssignal; NLS: Kernlokalisationssignal; AD: Aktivierungsdomäne; *= fehlende AS (Darstellung abgewandelt nach Mak *et al.*, 2012)

Die initiale Interaktion mit dem DNA-Strang erfolgt durch zwei degenerierte Wiederholungen am erweiterten N-terminalen Beginn der repetitiven Domäne, welche als 0. und -1. Wiederholung bezeichnet werden. Beide bilden eine α -Helix-Schleufe-Struktur, die

bei räumlicher Annäherung an eine Thyminbase konvergieren und stabilisierende Wechselwirkungen eingehen (102, 103). Durch die lokale Fixierung wird die Anlagerung und Bindung des folgenden, ersten Sequenzmoduls an die nächste DNA-Base koordiniert und gefördert (103). Von der zweiten Hälfte dieses ersten *Repeats* ausgehend bilden die Moleküle der Bindedomäne eine kontinuierlich linksgängige und superhelikale Packung entlang der DNA-Achse (Abb. 4). Die einzelnen Wiederholungen besitzen eine charakteristische Helix-Schleife-Helix-Struktur, wobei die dritte bis elfte AS am N-terminalen Beginn eine kurze, nach innen gerichtete α -Helix und die 15. bis 33. AS eine längere und von der DNA nach außen abgeneigte α -Helix bilden. Zwischen den beiden ist die sog. *RVD-Schleife* lokalisiert, bestehend aus zwölfter, dreizehnter und nichtvariabler vierzehnter AS. Die größeren C-terminalen Helices formen gemeinsam eine Art äußere Hülle um den Bindungskomplex, während die kleineren Helices einen internen Verbund entlang der DNA-Achse konstruieren. Auf diese Weise nehmen die *RVD*-Schleifen der einzelnen Wiederholungsmodule eine nach innen gerichtete und direkt zur DNA exponierte Lage ein. Die zwölfte AS dient der Stabilisierung der lokalen Schlaufenstruktur und unterstützt die Kontaktvermittlung lediglich indirekt. Die Bindung mit einer Base erfolgt exklusiv durch die dreizehnte AS. In Abhängigkeit von deren Beschaffenheit kommt es z.T. zur Ausbildung von stabilen, polaren Wasserstoffbrückenbindungen, wie auch schwächeren und unspezifischeren *Van-der-Waals*-Wechselwirkungen zwischen dem Protein und der DNA (10, 103, 104).

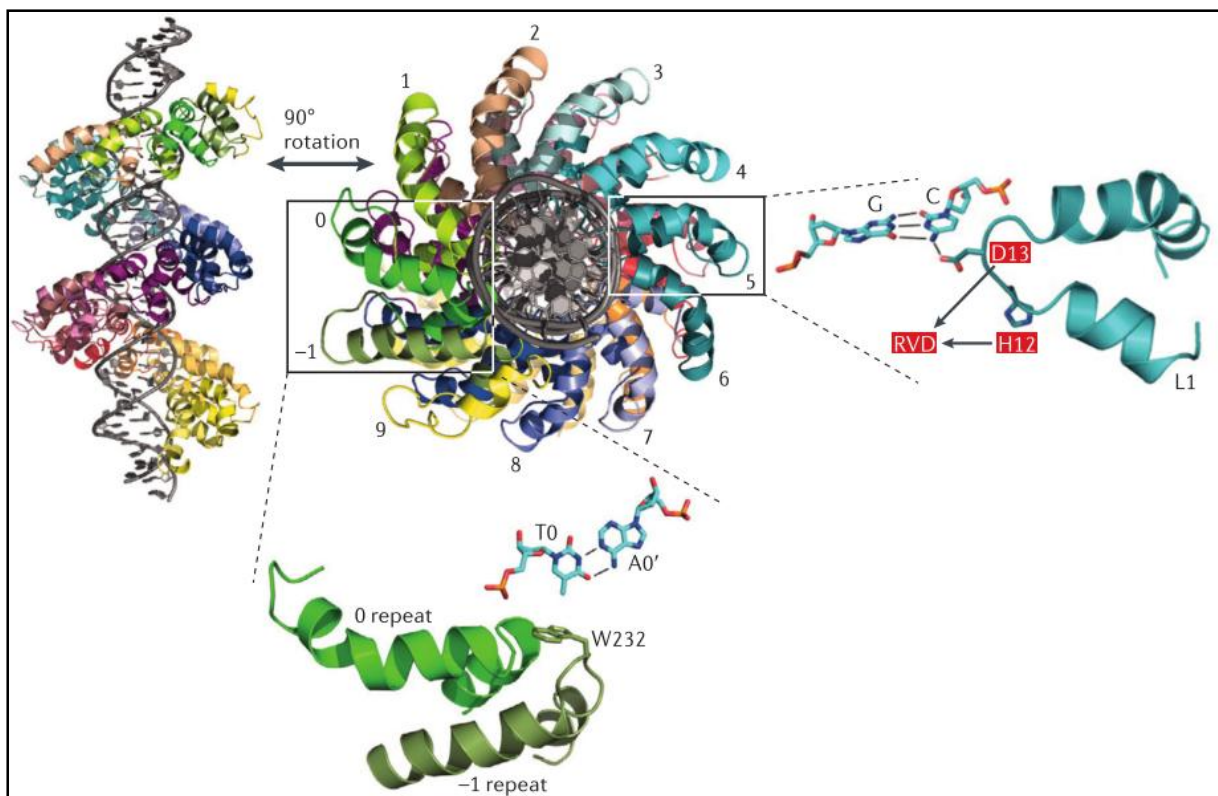


Abb. 4 | 3D Strukturmodell der DNA-Bindung durch ein *TALE*-Protein.

Aufgrund der modulartigen Architektur des Bindungsmotivs und der Eins-zu-eins-Korrespondenzen mit den Basen eines Zielstrangs, können durch systematische Verknüpfungen von einzelnen, individuellen Wiederholungseinheiten in einer entsprechend angepassten Reihenfolge, künstliche *TALE*-Domänen mit neuen Erkennungsspezifitäten generiert werden. Diese sind im Stande, nahezu jede vorgegebene bzw. gewünschte DNA-Sequenz zu binden (91, 105). Die einzige Limitierung ergibt sich aus der erforderlichen T₀-Base am 5'-Ende einer Zielsequenz (5, 7). Das System gewährt wesentlich mehr Flexibilität bei der Gestaltung von Bindungsmotiven als die auf Triplets beschränkten ZF-Proteinkomplexe (43). Die Anzahl der Sequenzmodule kann der vorgegebenen Zielsequenz sehr leicht angepasst werden. Ein Minimum von 6,5 Wiederholungen ist für die Funktionalität einer Bindedomäne erforderlich, wobei ab 10,5 *Repeats* eine deutlich gesteigerte Aktivität erkennbar ist (89). Die bislang bevorzugt genutzten *TALE*-Architekturen bestehen typischerweise aus etwa 14 bis 20 Sequenzabfolgen pro Protein bzw. Bindedomäne (5, 53). Für die Konstruktion von artifiziellen *TALE*-Motiven stehen mittlerweile zahlreiche natürliche und künstliche Modulvarianten für jede der vier Nukleotidbasen zur Verfügung, deren verschiedene *RVD*-Kombinationen unterschiedlich gut zur Aktivität einer Bindedomäne beitragen (106). Die Spezifität und Affinität einer *TALE*-Struktur für ein DNA-Zielmotiv wird im Wesentlichen von der Art, Anzahl und Abfolge der einzelnen Sequenzwiederholungen mit deren individuell abgestimmten und variierend effizienten *RVD* bestimmt. Die vier häufigsten und für *TALEN* bis heute mehrheitlich genutzten *RVD*-Motive sind Histidin-Asparaginsäure (HD) für Cytosin, Asparagin-Glycin (NG) für Thymin, Asparagin-Isoleucin (NI) für Adenin und Asparagin-Asparagin (NN) für Guanin oder Adenin (5, 38).

Es existieren zahlreiche Methoden und Schemata zur Erstellung und Assemblierung nutzerdefinierter *TALE*-Abfolgen, wie bspw. die sequentielle Verknüpfung einzelner Module (38, 91), auf einer Polymerase-Kettenreaktion (engl.: *polymerase chain reaction*; *PCR*) basierten oder kompletten *in vitro* Gensynthese mittels Hochdurchsatz-Festphasenassemblierung (107), durch eine Mutagenese von natürlichen Sequenzwiederholungen (108), sowie mittels verschiedener modularer oder ligationsunabhängiger Klonierungstechniken (109, 110). Aufgrund der repetitiven Beschaffenheit der langen und hochgradig homologen Wiederholungsstrukturen stellen einfache Klonierungen und *PCR*-basierte Assemblierungsmethoden bislang jedoch häufig eine technisch herausfordernde und zeitintensive Prozedur zur Konstruktion Zielsequenzspezifizierter, *TALE*-kodierender Expressionseinheiten dar (43, 111). Ein praktikables, effizientes und erprobt zuverlässiges Verfahren für die Synthese neuer Bindestrukturen bildet

bspw. die sogenannte *Golden-Gate*-Klonierungsplattform (112, 113), welche eine spezifisch geordnete Assemblierung von bis zu zehn Modulen in nur einem Reaktionsschritt erlaubt (114). Die meisten der Sequenzmodulvarianten und weitere für diverse Konstruktionsverfahren erforderlichen Reagenzien sind in öffentlich zugänglichen Bibliotheken katalogisiert und werden bspw. durch die Internetplattform *Addgene.org* zur freien Verfügung gestellt. Für die Planung und Konstruktion von *TALE*-Domänen stehen verschiedene, webbasierte Programme zum Auffinden und zur Auswahl von Zielsequenzen in einem gegebenen DNA-Abschnitt zur Verfügung, wie auch zur Suche nach weiteren möglichen Bindungssequenzen im chromosomalen Hintergrund eines bekannten Wirtsgenoms (115-117). Diese berücksichtigen bislang jedoch weder potentielle epigenetische Modifikationen der Zielsequenzen wie bspw. methylierte DNA-Basen, noch die deutlich unterschiedlichen Affinitäten der vier bevorzugten *RVD*-Motive (5, 118).

Die Fusion von natürlichen und künstlichen *TALE*-Bindestrukturen mit verschiedenen Arten von Effektorgruppen führte sehr schnell auch zu einer Kopplung an eine katalytische FokI-Nukleasedomäne und damit zur Etablierung eines nahezu universell einsetzbaren, chimären Restriktionsenzym für Sequenz-spezifizierte DSB-Induktionen (91). Wie bei ZFN sind dafür zwei *TALEN*-Einheiten mit unterschiedlich gestalteten Bindedomänen erforderlich, die in entgegengesetzter Orientierung und mit passendem Abstand zueinander an je einen der DNA-Stränge binden, um den beiden Nukleasedomänen die Dimerisierung zu ermöglichen (119). Die DNA-Zielsequenz eines *TALEN*-Paares beträgt damit im Durchschnitt etwa 28 bis 40 bp, was in den meisten Fällen deutlich länger im Vergleich zu denen anderen Klassen von SSE ist und einen wesentlichen Vorteil mit sich bringt (5). *TALEN* gelten als die spezifischste aller Nukleaseplattformen und besitzen das geringste Risiko für ungewollte, potentiell toxische Genomveränderungen durch unspezifische Bindungen (5, 53). Im Vergleich zu ZFN (Vgl. S. 20) agieren die einzelnen Wiederholungsmodule größtenteils unabhängig voneinander und werden weniger vom molekularen Kontext beeinflusst (38). Dennoch können sich vielfältige und z.T. nur geringfügige strukturelle Eigenschaften problematisch auf die Bindungseffizienz und DSB-Induktionsrate von *TALEN* auswirken (106). Die flankierenden AS-Sequenzen der Bindemotive bilden das sog. *TALEN-Rückgrat*, welches eine wichtige Rolle bei der Wahrnehmung von DNA spielt und die Aktivität einer Domäne in Organismen wesentlich mitbestimmt (38). Stark verkürzte oder lediglich aus Wiederholungsmodulen bestehende Proteine funktionieren nicht effizient (91). Jedoch zeigen *TALEN* mit teilweise verkürzten N- oder C-Termini häufig eine z.T. über 25-fach gesteigerte Mutageneserate gegenüber längeren Rückgratstrukturen (38). In Vergleichsstudien in Hefen

konnten die besten *TALEN*-Aktivitäten mit einer C-terminalen 63 bp-Sequenz und einem Abstand von 14 bis 18 bp zwischen den Enzymeinheiten erzielt werden (91, 120). Zwecks struktureller Faltung und Stabilität des Proteins empfiehlt es sich Sequenzen von mind. 15 bis 21 bp als Verknüpfung vor der FokIR-Domäne zu nutzen (10). Der optimale Abstand zweier *TALEN*-Monomeren zueinander variiert leicht in Abhängigkeit der zusätzlichen AS-Sequenzen zwischen der Binde- und Restriktionsdomänen einer jeden *TALEN*-Einheit. Die Distanz sollte mindestens zehn bp betragen, zwischen zwölf bis 20 bp sind zumeist üblich (5, 10). Die noch immer offenen Fragen an eine robuste Rückgratarchitektur bezüglich der Aktivität und Effizienz von *TALEN*-DNA-Bindungen und -Spaltungen, stellt bislang eines der größten Mankos der Plattform dar. Der oftmals unberücksichtigte Einfluss möglicher epigenetischer Modifikationen der Zielsequenz und die signifikant variierende Affinität der zahlreichen, von natürlichen Proteinen abgeleiteten AS-Module, die häufig behandelt werden als besäßen sie essentiell dieselben Eigenschaften, sind vermutlich mit die Hauptgründe für die unzureichende Funktionalität einiger *TALEN*-Konstrukte (5, 10). Zur Analyse der *in vivo* Aktivität neu gestalteter *TALEN* wurden daher verschiedene Funktionalitätstests entwickelt, welche bislang vorwiegend auf *Hefe-Hybridisierungssystemen* basieren (52). Eine weitere potentielle Einschränkung ergibt sich aus der Größe der je Monomereinheit durchschnittlich rund 950 AS umfassenden Proteine. Im Vergleich zu ZFN sind *TALEN* damit pro zu bindender Base etwa drei- bis viermal größer, was ein Einbringen in Pflanzenzellen unter Umständen verkompliziert (53). Ein kommerzieller Vertrieb von Reaktionssets oder auch kompletten *TALEN*-kodierenden Expressionskonstrukten nach Kundenwunsch erfolgt durch

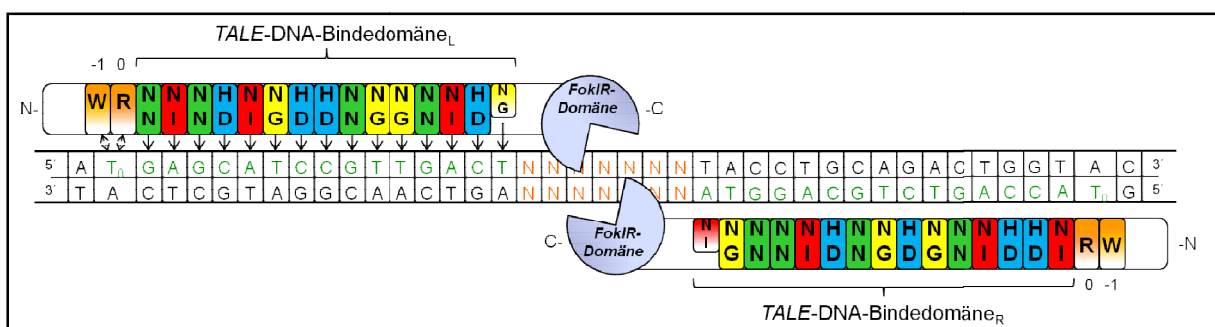


Abb. 5 | Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz zur Induktion eines Doppelstrangbruchs (DBS) durch ein *TALEN*-Paar. Jedes der *TALEN*-Proteine setzt sich aus einer N-terminal gelegenen, spezifisch gestalteten DNA-Bindedomäne und einer C-terminalen FokI-Restriktionsdomäne zusammen. Die Kontaktaufnahme mit der DNA wird vom -1. und 0. Repeat (W, R) vermittelt. Die Bindung der Basen der Zielsequenz (grün) erfolgt ausschließlich durch die 13. AS der RVD-Motive der einzelnen Wiederholungsmodule, die zwölfte AS dient zur Stabilisierung der Struktur. Zur Induktion eines DSB innerhalb der Zielsequenz müssen zwei entsprechend gestaltete *TALEN*-Monomere in entgegengesetzter Orientierung links- und rechtsseitig einer Überbrückungssequenz (orange, N= beliebige Base) an jeweils einen DNA-Strang binden, um die Dimerisierung der Nukleasedomänen zu ermöglichen. (Darstellung abgewandelt nach Kim& Kim, 2014)

die Firmen *DNA Cloning Service* (Hamburg, Deutschland), *Collectis BioResearch* (Paris, Frankreich), *Life Technologies* (Grand Island, New York, USA) und *Transposagen Biopharmaceuticals* (Lexington, Kentucky, USA).

Trotz der Limitierungen eignet sich die *TALEN*-Plattform zur ortsspezifischen Induktion von genetischen Modifikationen in Pflanzenzellen (39, 121). *TALEN*-stimulierte Genomänderungen umfassen NHEJ- und HDR-vermittelte Modifikationen trans- und endogener DNA-Sequenzen von mono- wie auch dikotylen Modell- und Kulturpflanzen (53, 122). Dazu zählen bspw. Acker-Schmalwand (*A. thaliana*) (114, 123), die Zwenke (*Brachypodium distachyon*) (124), Tabak (*N. benthamiana*, *N. tabacum*) (121, 125), Reis (*Oryza sativa*) (39, 124), Mais (*Z. mays*) (126, 127), Weizen (*Triticum aestivum*) (128), Gerste (*Hordeum vulgare*) (129-131), Soja (*G. max*) (132), Kartoffel (*Solanum tuberosum*) (133, 134) und Tomate (*Solanum lycopersicum*) (135). Die erzielten genetischen Variationen beinhalten die Stilllegung von Genfunktionen, die gezielte Integration und das Ersetzen von DNA-Sequenzen sowie das Entfernen von Gen- oder Chromosomenabschnitten (53, 122). Die Änderungen wurden z.T. erfolgreich an die Nachkommenschaften vererbt. Aufgrund der häufig eher geringen DSB-Induktionsfrequenzen, einiger bestehender Unklarheiten bezüglich einer funktionell verlässlichen *TALEN*-Architektur und der vergleichsweise komplizierten Assemblierung von kodierenden Expressionseinheiten gegenüber den, im nächsten Kapitel beschriebenen *RNA-vermittelte Endonukleasen* (engl.: *RNA-guided endonucleases*, RGEN), bevorzugen viele Forschungsgruppen im Bereich der GE heute zumeist das zwar weniger spezifische, dafür aber leichter zu handhabende und effizientere RGEN-Format (5). Dennoch stellt die *TALEN*-Plattform eine bedeutende und nützliche Erweiterung in der Palette von präzise DNA-interagierenden Werkzeugen zur GE dar und wurde 2012 gemeinsam mit ZFN vom Fachmagazin *Nature* zur 'Methode des Jahres' gekürt (136).

1.2.2.4 *RNA-vermittelte Endonukleasen*

RNA-vermittelte Endonukleasen bilden die derzeit effizienteste Plattform zur gezielten Induktion von Mutationen in Genomen eukaryotischer Organismen. Im Gegensatz zu den zuvor beschriebenen Nukleaseplattformen, welche allesamt auf Protein-DNA-Interaktionen zur Erkennung und Bindung von Zielsequenzen beruhen, basieren RGEN-Systeme auf dem einfachen Prinzip der komplementären Basenpaarung zur spezifischen Erkennung und Wechselwirkung mit einer DNA-Zielsequenz und einem RNA-Leitmolekül (engl.: *guide RNA*, gRNA) mit variabler, individuell gestaltbarer Erkennungssequenz. Die gRNA dirigiert

eine zuvor gebundene Endonuklease zur Hybridisierungssequenz, welche die enzymatische Spaltung der ds-DNA katalysiert und einen DSB generiert (10, 53, 137).

Der Mechanismus der RGEN ist von nativen, in Bakterien und Archaeen weit verbreiteten adaptiven Immunsystemen zur Abwehr eindringender Phagen- oder Plasmid-DNA abgeleitet, welche als *CRISPR/Cas*- (engl. Kurzform für: *clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein*) Systeme bezeichnet werden (138, 139). *CRISPR* sind spezielle DNA-Abschnitte aus einer Serien von kurzen, durch regelmäßig integrierte, variierende Zwischensequenzen (engl.: *spacer*) getrennte und zumeist palindrome Sequenzwiederholungen von jeweils etwa 20 bis 50 bp. Die *Repeat*-Strukturen innerhalb eines *CRISPR*-Lokus sind stets identisch und hochkonserviert. Verschiedene *CRISPR*-Regionen unterscheiden sich jedoch z.T. deutlich bezüglich der Anzahl, der Länge und der Nukleotidsequenz ihrer Wiederholungen (140, 141). Die prokaryotischen Motive sind mehrheitlich auf Chromosomen, seltener auch auf Plasmiden lokalisiert und werden für gewöhnlich von einer variierenden Anzahl sehr verschiedener *Cas*-Gene flankiert, die eine umfangreiche Familie von *CRISPR*-assoziierten Proteinen kodieren (138). Über 40 Subklassen natürlicher *Cas*-Proteine mit diversen funktionellen Domärentypen konnten bislang identifiziert werden, darunter bspw. Nukleasen, Helikasen, Hydrolasen, Polymerasen und nukleotidbindende Proteine (142). Zur spezifischen Immunisierung einer befallenen Zelle mittels *CRISPR/Cas* übernehmen spezialisierte *Cas*-Proteine kurze, als *Protospacer* (Ursprung der *Spacer*) bezeichnete Fragmente der exogenen DNA und integrieren diese jeweils einzeln zwischen zwei Wiederholungen der *CRISPR*-Region im eigenen Genom. Auf diese Weise werden verschiedene Signaturen von eingedrungenen, pathogenen DNA-Molekülen als *Spacer*-Sequenzen in der *CRISPR*-Region gesammelt und konserviert (143-146). Bei einer erneuten Invasion durch DNA des gleichen oder eines verwandten Virustyp bzw. derselben Plasmidfamilie erfolgt die Transkription der *CRISPR*-Kassette und eine Prozessierung der gebildeten RNA-Produkte. Dabei werden die zuvor akquirierten DNA-Abschnitte zusammen mit Teilen der angrenzenden Wiederholungssequenzen als Mustervorlage für die Biogenese von einzelnen kurzen *CRISPR*-RNA (crRNA) Molekülen genutzt, welche ggf. partiell komplementär und damit spezifisch passend zur *Protospacer*-Sequenz des Exogens sind (147-149). Mit diesem Zielmotiv entsprechend korrelierende crRNA leitet ein *Cas*-Protein zum *Protospacer*, welches die enzymatische Spaltung der fremden DNA katalysiert und diese dadurch inaktiviert (150, 151). Auf Grundlage der allgemeinen Organisationsstruktur eines Lokus und den Mechanismen der

Genkonservierung werden native *CRISPR/Cas*-Systeme in drei Haupttypen (Typ I bis Typ III *CRISPR*-assoziierte Systeme) untergliedert (152).

Die in den letzten Jahren konstruierten und für das GE in Eukaryoten bevorzugt genutzten RGEN basieren fast ausschließlich auf modifizierten und laboroptimierten Versionen eines Typ II *CRISPR*-assoziierten Systems von *Streptococcus pyogenes* (137). Für eine gezielte Spaltung beliebiger gewünschter DNA-Sequenzen sind bei diesen Varianten lediglich nur noch zwei wesentliche Kernkomponenten erforderlich, das multifunktionelle Protein Cas9 und ein Zielsequenz-spezifisch gestaltetes RNA-Leitmolekül (5, 10). Zur vereinfachten Handhabung basieren die gRNA-Konstrukte moderner RGEN-Systeme auf einer Fusion von crRNA und sog. *transaktivierender CRISPR-RNA* (tracrRNA), einem essentiellen Element, das für die Verbindung der Zielsequenz-spezifisierenden crRNA mit der nukleolytischen Cas9 erforderlich ist. Die synthetischen gRNA von RGEN vereinen crRNA und tracrRNA in einem Molekül, das daher auch als *single guide RNA* bezeichnet wird. Zielsequenz-spezifizierte gRNA-Konstrukte unterscheiden sich daher zumeist lediglich in der vom Anwender definierten und zum Ziel komplementären Leitsequenz von gewöhnlich 20 Basen am 5'-Ende des Transkripts (10). Die Bindung der gRNA an Cas9 induziert eine funktionelle Konformationsänderung in der Proteinstruktur und resultiert in der Aktivierung der Endonuklease (153). Die spezifische Erkennung der DNA-Zielsequenz durch einen gRNA-Cas9-Komplex erfolgt über die entsprechend homolog korrelierende Leitsequenz nach dem Prinzip der komplementären Basenpaarung. Bei einer Übereinstimmung mit der Zielsequenz (dem *Protospacer* des zugrunde liegenden bakteriellen Immunsystems) kommt es zu einer partiellen Trennung der DNA-Stränge und der Ausbildung eines DNA-RNA-Hybrids, wobei Fehlpaarungen u.U. toleriert werden (154-156). Die Spezifität einer RGEN für eine Zielsequenz wird primär von den Basen nahe dem 3'-Ende des Leitmotivs bestimmt. Diese bilden eine sog. *Seed* Sequenz aus etwa sieben bis zwölf Nukleotiden, von denen die Kontaktaufnahme zur DNA ausgeht. Im Gegensatz zum 5'-Ende werden innerhalb dieses Bereiches keine Basenfehlpaarungen toleriert (157, 158). Cas9 besitzt zwei verschiedene nukleolytische Domänen der Typen RuvC und HNH (137). Zur Stimulation der enzymatischen Aktivität von Cas9 ist ein spezielles, direkt strangabwärts an den *Protospacer* angrenzendes Triplettmotiv (engl.: *protospacer adjacent motif*, *PAM*) erforderlich, für welches das Sequenzmuster 5'-NGG-3' charakteristisch ist (144). Dieses Motiv hat zudem weitreichende Funktionen bei der Wahrnehmung von Zielsequenzen durch Cas9 selbst, sowie der Stabilisierung der Bindung und beeinflusst die Spezifität und Aktivität einer RGEN maßgeblich (159, 160). Wird ein entsprechendes Triplet vom Cas9-Molekül erkannt und

gebunden, erfolgt eine weitere Konformationsänderung in der Enzymarchitektur, die zu einer Spaltung der beiden Ziel-DNA-Stränge jeweils genau drei bp strangaufwärts vom *PAM* führt und von je einer der beiden Exonukleasen katalysieren wird (153). Im Gegensatz zur FokI-Nuklease, welche adhäsive und überstehende DSB-Enden generiert (Vgl. S. 20), induziert Cas9 einen glatten Bruch ohne Überhänge und ist zudem in der Lage, auch methylierte DNA zu spalten (7, 155).

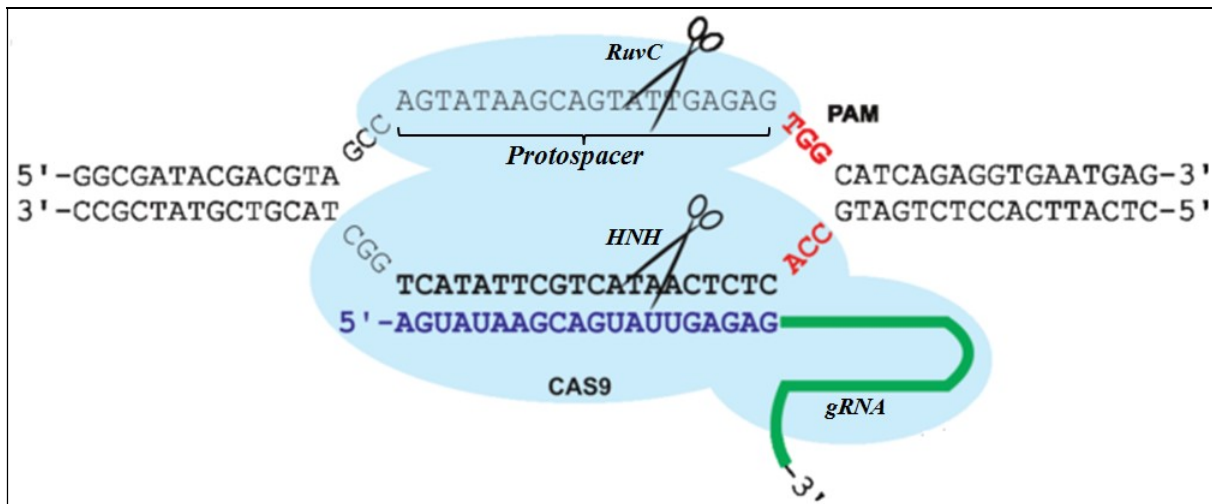


Abb. 6 | Schematische Darstellung der Erkennung und Bindung einer spezifischen DNA-Zielsequenz zur Induktion eines Doppelstrangbruchs durch eine RNA-geleitete Endonuklease (RGEN). RGEN bestehen typischerweise aus einem RNA-Leitmolekül (gRNA; blau und grün), das eine Bindung mit einer Cas9-Endonuklease eingeht. Die gRNA beinhaltet eine zum Zielabschnitt komplementäre Leitsequenz (blau) aus für gewöhnlich 20 Basen, mit deren Hilfe das Cas9-Protein zur Zielsequenz dirigiert und die Bindung an die DNA nach dem Prinzip der komplementären Basenpaarung vermittelt wird. Zur Stimulation der enzymatischen Aktivität von Cas9 ist ein strangabwärts an die *Protospacer*-Region angrenzendes Triplettmotiv (*PAM*; rot, oben) erforderlich, welches sich durch ein charakteristisches Sequenzmuster vom Typ 5'-NGG-3' (hier 5'-TGG-3') auszeichnet. Wird ein entsprechendes Triplet vom Cas9-Molekül erkannt und gebunden, kommt es infolge einer Konformationsänderung innerhalb der Cas9-Struktur zur Spaltung der beiden Ziel-DNA-Stränge durch je eine der beiden Exonukleasedomänen (RuvC und HNH). (Bildquelle: *Puchta & Fauser, 2013*)(161)

Ein wesentlicher Vorteil von RGEN gegenüber anderen Sequenz-spezifisierbaren Endonukleasen ist das vergleichsweise einfache Reprogrammieren zur Generierung neuer Zielspezifitäten für die Bindung von gewünschten Sequenzen. Da der RGEN-Mechanismus nur ein einziges und stets identisches Protein zum Spalten von DNA nutzt, ist ein aufwendiges und anspruchsvolles Umgestalten komplexer Proteinstrukturen nicht notwendig (7, 122). Die Konstruktion von neuen zielkomplementären Leitsequenzen erfordert lediglich grundlegende Kenntnisse in den Regeln der spezifischen Basenpaarung nach Watson und Crick und ist durch das Klonieren von entsprechenden gRNA-kodierenden DNA-Vektoren mit verschiedensten Techniken, wie bspw. *Gibson*[®]- oder *Golden-Gate*-Klonierungen vergleichsweise einfach umsetzbar (112, 113, 162). Alternativ kann eine Assemblierung durch die Hybridisierung von zwei komplementären Oligonukleotiden und anschließender *in-*

in vitro Replikation erfolgen (163). Zudem bieten zahlreiche Firmen und Zulieferer eine kommerzielle Fertigung synthetischer RGEN oder auch komplette Reaktionssätze mit Protokollen zur eigenen Herstellung an. Die geringe Größe der zielführenden Elemente erlaubt ein Unterbringen mehrerer, verschiedener gRNA-Varianten in einer Expressionskassette, womit ein simultanes Abzielen auf multiple DNA-Sequenzen in einer Zelle möglich ist (164-166). Ein generiertes Leitmotiv sollte mindestens die Basen der *Seed* Sequenz umfassen. Kurze Bindungsabfolgen zeichnen sich i.d.R. durch eine höhere Spezifität aus, besitzen dafür aber zumeist eine geringere Affinität für den Zielabschnitt (157, 167). Zur Steigerung der Rekombinationsfrequenz empfiehlt es sich daher, Sequenzen von 20 Basen und drei komplementären Basen der *PAM* zu nutzen (5, 168). Eine entsprechende Gesamtzielsequenz besitzt damit das charakteristische Muster 5'-X₂₀NGG-3', wobei X je ein spezifisches Nukleotid der Zielsequenz darstellt. Cas9 von *Streptococcus pyogenes* erkennt und toleriert neben 5'-NGG-3' auch Triplets vom Typ 5'-NGA-3', wobei die Effizienz jedoch stark vermindert ist (155). Proteinhomologe von *CRISPR/Cas*-Systemen anderer Arten besitzen zudem verschiedene individuell bevorzugte *PAM* (144, 169, 170). Die Präferenz von RGEN für solche Triplets führt zu einer geringfügigen Reduktion der absoluten Anzahl an potentiellen Zielsequenzen in einem Genom (5). Ein Nachteil von RGEN ergibt sich aus der verhältnismäßig großen Toleranz gegenüber Basenfehlpaarungen, die in einer vergleichsweise hohen Anfälligkeit für mögliche unerwünschte Bindungen und potentiell toxische Spaltungen von anderen strukturell ähnlichen DNA-Abschnitten im Wirtsgenom resultiert (155, 171). Der einfachste Weg zur Vermeidung von ungerichteten Interaktionen ist eine strikte Auswahl von gRNA mit minimalen Homologien bzw. möglichst einzigartig abweichenden Signaturen zu anderen Sequenzen im chromosomalen Hintergrund (172). Speziell bei Arten mit einem kleinen Genom ist diese Vorgehensweise häufig ausreichend, um zielgerichtete Modifikationen ohne Nebenwirkungen zu induzieren. Mit steigender Größe eines Genoms erhöht sich jedoch die Anzahl von wahrscheinlichen sekundären Bindungsmöglichkeiten, was jedoch durch den entsprechend höheren Anteil an nicht-kodierenden Bereichen weitgehend ausgeglichen wird, der Mutationen ohne Weiteres toleriert (5). Verschiedene Organisationen bieten Programme zur Ermittlung von einmalig bzw. selten auftretenden Bindungsmotiven eines gegebenen DNA-Abschnitts oder auch weiterer tendenziell tolerierter Nukleotidabfolgen einer RGEN in bekannten Genomen an (7, 173). Alternative Ansätze zur Steigerung der Spezifität nutzen bspw. ausschließlich ss-spaltende Cas9-Derivate (174-176) oder auch eine Fusion von katalytisch inaktiven Cas9 mit funktionellen FokI-Domänen (177, 178), wodurch das Zielmotiv verdoppelt wird und es nur bei Bindung eines gRNA-Paares zur

Induktion eines DSB kommt. Auch eine Reduktion der Leitsequenz kann die Spezifität einer RGEN für DNA-Ziele erhöhen, führt aber im Gegenzug häufig zu einer verringerten Aktivität bei der Wahrnehmung (167). Eine andere potentielle Einschränkung ergibt sich aus der Größe des Cas9-Moleküls. Mit etwa 1.350 Aminosäuren ist das Protein eines der größten Reagenzien des GE und für Anwendungen mit geringer Vektorkapazität daher u.U. nicht geeignet. Eine Alternative bieten Cas-Nukleasen von *CRISPR*-Systemen anderer Arten, wie bspw. die etwa ein kbp kleinere und dennoch effizient DNA-spaltende Cas9 von *Staphylococcus aureus* (153, 179). Ein neuartiger Ansatz zur Umgehung der Problematik nutzt zwei separat kodierte und unabhängig voneinander exprimierte Cas9-Untereinheiten, die eine Selbstassemblierung *in vivo* vollziehen und in Anwesenheit der gRNA eine aktive Endonuklease bilden (180).

Trotz der geringfügigen Einschränkungen wurde die RGEN-Plattform seit der Dekodierung des *CRISPR/Cas9*-Mechanismus (137) sehr schnell für effektive Sequenz-spezifische Modifikationen in einer Vielzahl von Eukaryoten etabliert und angewendet, darunter auch verschiedene mono- und dikotyle Modell- und Kulturpflanzen wie bspw. Acker-Schmalwand (*A. thaliana*), Zwenke (*B. distachyon*), Schneckenklee (*Medicago truncatula*), Tabak (*N. benthamiana*, *N. tabacum*), Reis (*O. sativa*), Mais (*Z. mays*), Weizen (*T. aestivum*), Hirse (*Sorghum bicolor*), Soja (*G. max*), Kartoffel (*S. tuberosum*), Tomate (*S. lycopersicum*) und Orange (*Citrus sinensis*) (122, 181, 182). Das breite Spektrum an NHEJ- oder HDR-vermittelten Editierungen von trans- und endogenen Nukleotidsequenzen umfasst alle Arten von genomischen Änderungen, darunter zahlreiche funktionelle Inaktivierungen einzelner oder multipler Gene durch *InDel*-Mutationen, die Entfernung gesamter Gengruppen und Chromosomenabschnitte von bis zu 245 kbp mit Hilfe mehrerer verschiedener gRNA-Konstrukte, die ortsspezifische Insertion von Transgenen sowie das Abändern bzw. Ersetzen von Sequenzen oder Basen durch artifizielle DNA-Reparaturvorlagen (164, 183). Die generierten Mutationen konnten z.T. erfolgreich durch eine normale Vererbung nach Mendel an die T₁- und weitere Pflanzengenerationen übertragen werden (184, 185). RGEN-Systeme zeichnen sich durch vergleichsweise hohe Mutageneseraten aus. Abhängig von der Art und dem Zelltyp des Wirts, dem Transformationssystem, sowie weiteren Faktoren konnten in Studien an Acker-Schmalwand (*A. thaliana*) und Reis (*O. sativa*) mitunter Rekombinationsfrequenzen von über 90 % erzielt werden (185, 186). Aufgrund der vielfältigen Vorteile von RGEN gegenüber anderen Plattformen von Sequenz-spezifisierbaren Endonukleasen, wurde die Technologie vergleichsweise schnell weiterentwickelt und

avancierte schließlich zur bevorzugten Methode für die Forschung und Anwendungen im Bereich des pflanzlichen *Genome Engineering*.

1.3 Biolistische Transformation von Pflanzenzellen

Der Begriff der *genetischen Transformation* beschreibt allgemein ein Einschleusen von DNA in lebende Zellen. Wird die DNA in Folge dieses Prozess dauerhaft in das Wirtsgenom integriert, spricht man von stabiler Transformation bzw. stabil transformierten Zellen. Findet hingegen keine Integration statt, wird das Exogen dennoch vorübergehend exprimiert, geht im Verlauf der Zellentwicklung jedoch verloren. Dieser Vorgang wird als *transiente Expression* bezeichnet (111). Der Erfolg vieler praktischer Anwendung des GE hängt direkt von effizienten Methoden zur Transformation von potentiellen Zielzellen und -geweben für die Bereitstellung von Sequenz-spezifischen Endonukleasen sowie ggf. DNA-Reparaturvorlagen und anderen Substanzen ab (10). Es existieren vielfältige biologische, physikalische und chemische Methoden, mit denen die natürlichen Barrieren von Pflanzenzellen in Form von Zellwand und Plasmamembranen überwunden werden können (187). Zu den bedeutendsten und am besten etablierten Verfahren zählen die *biolistische Transformation* (188) und die *Agrobacterium tumefaciens-vermittelte Transformation* (189). Einen besonderen Status bezogen auf die Art der Akzeptorzellen hat die Transformation von *Protoplasten*, da diese keine Zellwand besitzen und Fremd-DNA mit diversen Techniken eingeschleust werden kann. Ein großes Potential besitzen zudem auf DNA- oder RNA-Viren basierende transiente Expressionssysteme, die in den letzten Jahren immer häufiger zur Applikation ohne stabile Transformation genutzt wurden (111, 190, 191).

Die biolistische Transformation verwendet Hochgeschwindigkeits-Mikroprojekte um DNA und andere Substanzen in intakte Zellen zu transportieren. Der Prozess wird daher häufig auch als *Mikroprojektil-Bombardement* bezeichnet. Das genetische Material wird in einer Vernetzungsreaktion zunächst an der Oberfläche kleiner Metallpartikel – gewöhnlich aus Gold oder Wolfram – immobilisiert. Die benetzten Partikel werden anschließend mit hoher Geschwindigkeit auf ein Zielgewebe geschossen, wobei sie die Zellbarrieren durchbrechen und die exogene DNA in die Zellen einschleusen. Die verwendeten Partikel sollten einen Durchmesser von etwa 0,1 bis 0,2 μm besitzen, um keine schwerwiegenden oder gar letalen Schäden an den Wirtszellen zu verursachen. Zur Beschleunigung der Partikel werden mit Schwarzpulver, Gasdruck oder elektrischen Entladungen betriebene *Genkanonen* genutzt. Die Biolistik stellt eine recht effiziente, schnelle und einfache Methode zur direkten Transformation von mono- und dikotylen Pflanzenzellen dar. Das Verfahren unterliegt kaum

Beschränkungen und gilt als nahezu unabhängig bezüglich der Art und dem Gewebetyp des Wirts. Mikroprojektil-Bombardement wurde vielfach zur Erzeugung von stabil transgenen Pflanzen genutzt. Durch die üblicherweise hohe Anzahl an einbrachten Fremd-DNA-Kopien und der Möglichkeit verschiedene Moleküle in einem Schritt zu transferieren, einem sog. *Co-Bombardement*, eignet sich die Methode jedoch insbesondere für eine starke transiente Expression (192-194).

1.4 Zielstellung der Arbeit

Ziel dieser Arbeit war es zunächst detaillierte Hintergrundinformationen zu den unterschiedlichen Plattformen des *Genome Engineerings* und der Mechanismen der DNA-Reparatur zusammenzustellen und in Form dieser Dokumentation der Arbeitsgruppe zur Verfügung zu stellen. Der experimentelle Kern der Arbeit sollte darüber hinaus in der Klonierung und Erprobung eines flexibel anpassbaren DNA-Testvektors zur Funktionalitätsprüfung und quantitativen Ermittlung der nukleolytischen Aktivität von Sequenz-spezifisierten Endonukleasen in pflanzlichen Geweben bestehen.

Das zu entwickelnde Testsystem sollte auf der Reaktivierung eines Reportergens beruhen, das im inaktiven Leserahmen des Testvektors vorliegt und N-terminal mit einer beliebigen zu mutierenden Testsequenz bestückt werden kann. Ein signifikanter Anteil der in dieser Testsequenz mittels Sequenz-spezifizierter Endonuklease induzierten *InDel*-Mutationen würde dann zur Korrektur des Reportergen-Leserahmens führen und aufgrund von fluoreszierenden Zellen detektierbar sein. Die Erprobung des Versuchsprinzips sollte mit Hilfe von transienten Expressionsanalysen in biolistisch transformierten Zellen von jungen Gerstenblättern (*Hordeum vulgare*) erfolgen. Dazu sollten die Testvektoren mit der jeweiligen DNA-Zielsequenz gemeinsam mit entsprechenden *TALEN*-kodierenden Plasmiden in das Pflanzengewebe eingebracht werden. Die Aufgaben im Rahmen der vorliegenden Arbeit sollten dementsprechend die detaillierte Konzeption und Erstellung von Testvektoren in verschiedener Ausführung sowie die Durchführung, Auswertung und Interpretation von Experimenten zur transienten Expression mit dem Ziel der Validierung des entwickelten Testsystems umfassen.

2 Material und Methoden

2.1 Materialien

2.1.1 Chemikalien und Enzyme

Agarose	<i>Serva</i>
Ampicillin	<i>Duchefa Biochemie</i>
Borsäure	<i>Roth</i>
Benzimidazol	<i>Sigma-Aldrich</i>
Calciumdichlorid	<i>Sigma-Aldrich</i>
Chloramphenicol	<i>Sigma-Aldrich</i>
Desoxynukleosidtriphosphate	<i>Invitrogen</i>
DNA Loading Dye	<i>Thermo Scientific</i>
DNA Stain Clear G	<i>Serva</i>
Ethylendiamintetraessigsäure	<i>Sigma-Aldrich</i>
Ethanol	<i>Roth</i>
Ethidiumbromid	<i>Serva</i>
<i>Fast Digest Green Puffer</i>	<i>Thermo Scientific</i>
<i>Gene Ruler DNA Ladder Mix</i>	<i>Thermo Scientific</i>
Glucose	<i>Sigma-Aldrich</i>
Goldkugeln	<i>BIO-RAD</i>
<i>Green GoTaq-Reaktionspuffer</i>	<i>Promega</i>
Hefeextrakt	<i>Duchefa Biochemie</i>
Helium	<i>Alphagaz</i>
Isopropanol	<i>Roth</i>
Kaliumchlorid	<i>Merck</i>
<i>LB</i>	<i>Duchefa Biochemie</i>
Magnesiumdichlorid	<i>Thermo Scientific</i>
Magnesiumsulfat	<i>Duchefa Biochemie</i>
Natriumchlorid	<i>Roth</i>
Nukleasefreies Wasser	<i>Ambion</i>
Phyto Agar	<i>Duchefa Biochemie</i>
<i>Phytigel</i>	<i>Sigma-Aldrich</i>
<i>Quick Ligase-Puffer</i>	<i>NewEngland Biolabs</i>

<i>Substrat 2</i>	<i>Klasmann-Deilmann</i>
Spectinomycin	<i>Duchefa Biochemie</i>
Spermidin	<i>Sigma-Aldrich</i>
<i>T4 DNA-Ligase-Puffer</i>	<i>Promega</i>
Tris-HCl	<i>Roth</i>
Trypton	<i>Duchefa Biochemie</i>
Vollentsalztes Wasser	<i>Millipore Milli-Q Plus</i>
<i>Fast Digest Restriktionsendonukleasen</i>	<i>Thermo Scientific</i>
<i>GoTaq DNA-Polymerase</i>	<i>Promega</i>
<i>T4 DNA-Ligase</i>	<i>Promega</i>
<i>Quick-Ligase</i>	<i>NewEngland Biolabs</i>

2.1.2 Kits und Verbrauchsmaterialien

<i>QIAprep Spin Miniprep Kit</i>	<i>QIAGEN</i>
<i>QIAquick Gel Extraction Kit</i>	<i>QIAGEN</i>
<i>QIAquick PCR Purification Kit</i>	<i>QIAGEN</i>
Deckgläser	<i>Menzel-Gläser</i>
Erlenmeyer Schikanenkolben	<i>Fisherbrand</i>
Kulturflaschen	<i>Fisherbrand</i>
Kulturröhrchen	<i>Sigma</i>
Kunststoffröhrchen (50 ml) <i>Cellstar</i>	<i>Greiner Bio-One</i>
Latex-/ Nitrilhandschuhe	<i>Top Glove/ Kimtech</i>
Objektträger <i>Micro Slides</i>	<i>Assistent</i>
<i>PCR-Reaktionsgefäße</i>	<i>Brand</i>
Partikelbombardmentscheiben	<i>BIO-RAD</i>
Petrischalen	<i>Greiner Bio-One</i>
Pipettenspitzen <i>Safe Seal-Tips</i>	<i>Biozym</i>
Reaktionsgefäße (0,5 - 2,0 ml)	<i>Eppendorf</i>

2.1.3 Geräte und Software

Autoklave <i>ST 6200</i>	<i>Heraeus Instruments</i>
Bakterienbrutschrank <i>Typ B6</i>	<i>Heraeus Instruments</i>

Bunsenbrenner <i>Schuett Phoenix</i>	<i>Schütt Labortechnik</i>
Elektrophoresekammern <i>Sub-Cell Modell 192</i>	<i>BIO-RAD</i>
Elektrophorese Energieversorger <i>EPS 301</i>	<i>Amersham Pharmacia</i>
Feinwaage <i>Quintix 124-18</i>	<i>Sartorius</i>
Genkanone <i>PDS-1000/He</i>	<i>BIO-RAD</i>
Heizblock <i>ThermoMixer C</i>	<i>Eppendorf</i>
Heizplatte <i>SLK 12</i>	<i>SI Analytics</i>
Geldokumentation <i>ImageMaster VDS</i>	<i>INTAS</i>
Konfokal <i>laser scanning</i> Mikroskop <i>LSM 780</i>	<i>Carl Zeiss</i>
Lichtmikroskop <i>Axio Imager M2</i>	<i>Carl Zeiss</i>
Magnetrührer <i>MR Hei-Standard</i>	<i>Heidolph Instruments</i>
Mikrowelle <i>7017</i>	<i>Privileg</i>
PCR-Thermocycler <i>Mastercycler EP gradient S</i>	<i>Eppendorf</i>
pH-Meter <i>pH 538</i>	<i>WTW</i>
Pipetten <i>Pipetman</i>	<i>Gilson</i>
Schnellkochtopf <i>Perfekt Plus</i>	<i>WMF</i>
Schüttelinkubatoren <i>Minitron</i>	<i>Infors HAT</i>
Spektrophotometer <i>Nanodrop 2000c</i>	<i>Thermo Scientific</i>
Sterilwerkbank <i>HeraSafe HS18</i>	<i>Heraeus Instruments</i>
Tischzentrifugen <i>Centrifuge 5415 R</i>	<i>Eppendorf</i>
Ultraschallbad <i>Transsonic T 310</i>	<i>Elma</i>
Vakuumpumpe <i>RZ 8</i>	<i>Vacuubrand</i>
Vortexer <i>Vortex-Genie 2 G-560E</i>	<i>Scientific Industries</i>
<i>Clone Manager 9 Professional Edition</i>	<i>Scientific & Educational Software</i>
DNA-Schmelztemperaturkalkulatoren	<i>Thermofisher.com/NEB.com</i>
<i>ZEN 2 Black Edition</i>	<i>Carl Zeiss</i>

2.1.4 Pflanzenmaterial und Inkubationsmedium

Anzucht der Gerste

Die Kultivierung des Pflanzenmaterials der Sommergerstenlinie *Golden Promise* erfolgte mit Saatgut aus den Beständen des *IPK Gatersleben*. Zur Keimung wurden bis zu 30 Samen in 16,0 cm Töpfe mit einer Mischung aus Komposterde, *Substrat 2* und Sand (2:2:1) eingebracht. Die Pflanzen wurden anschließend für maximal zwei Wochen im Gewächshaus

bei 20,0 bis 22,0 °C und einem Tag-Nacht-Rhythmus von 16 h bei 50 bis 100 $\mu\text{mol s}^{-1} \text{m}^{-2}$ zu acht h Dunkelheit kultiviert.

Inkubationsmedium für Gerstenblätter

<u>Phyto Agar-Platten:</u>		<u>Endkonzentration</u>
Phyto Agar	10,00 g/l _{VE-Wasser}	1,0 % (w/v)
pH 7,2 einstellen; autoklavieren		
Benzimidazol (40,0 g/l; sterilfiltriert)	0,50 ml/l _{Medium}	20,0 mg/l _{Medium}
Chloramphenicol (20,0 g/l; sterilfiltriert)	0,50 ml/l _{Medium}	10,0 mg/l _{Medium}
Lösung in Petrischalen gießen		

2.1.5 Bakterienstamm und Kulturmedien

Escherichia Coli XLI-Blue

Genotyp:

supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1 lac [F' *proAB lacI^fZAM15 Tn10* (Tet^r)]

Kompetente Zellen des zur Klonierung und Vermehrung von Plasmiden verwendeten *E. coli* Stamms *XLI-Blue* (*Stratagene*, Heidelberg) wurden von *DNA Cloning Service*, Hamburg, bezogen und bei -80 °C gelagert. Der tetracyclinresistente Stamm ist sowohl Endonukleasedefizient (*endA*⁻) als auch Rekombinationsdefizient (*recA*⁻), was zur Erhöhung der Insertionsstabilität und einer Qualitätssteigerung von isolierter DNA beiträgt. Eine Blau-Weiß-Selektion rekombinanter Plasmide auf IPTG-/X-Gal-haltigem Medium wird durch die *lacZAM15*-Mutation zur α -Komplementation des β -Galaktosidase-Gens ermöglicht. Das Spalten klonierter DNA durch das *EcoK*-Endonukleasesystem ist durch eine Mutation des *hsdR*-Gens unterbunden (195, 196).

Medien zur Kultivierung von E. coli

<u>LB-Flüssigmedium:</u>		<u>Endkonzentration</u>
LB	25,00 g/l _{VE-Wasser}	2,5 % (w/v)
pH 7,2 einstellen; autoklavieren		
Amp oder Spec (100,0 g/l; sterilfiltriert)	1,00 ml/l _{Medium}	0,10 g/l _{Medium}

LB-Agar-Platten:

LB	25,00 g/l _{VE-Wasser}	2,5 % (w/v)
Agar	10,00 g/l _{VE-Wasser}	1,0 % (w/v)
pH 7,2 einstellen; autoklavieren		
Amp oder Spec (100,0 g/l; sterilfiltriert)	1,00 ml/l _{Medium}	0,10 g/l _{Medium}
Lösung in Petrischalen gießen		

SOC-Regenerationsmedium:

Trypton	20,00 g/l _{VE-Wasser}	2,0 % (w/v)
Hefeextrakt	5,00 g/l _{VE-Wasser}	0,5 % (w/v)
KCl (1,0 M)	2,50 ml/l _{VE-Wasser}	2,50 mM
MgCl ₂ (1,0 M)	10,00 ml/l _{VE-Wasser}	10,00 mM
MgSO ₄ (1,0 M)	10,00 ml/l _{VE-Wasser}	10,00 mM
pH 7,2 einstellen; autoklavieren		
Glucose (1,0 M; sterilfiltriert)	20,00 ml/l _{Medium}	20,00 mM

2.1.6 Gelelektrophoresepuffer0,5x TBE-Puffer:

		<u>Endkonzentration</u>
Tris-HCl	54,50 g/l _{VE-Wasser}	89,00 mM
H ₃ BO ₃	27,50 g/l _{VE-Wasser}	89,00 mM
Na EDTA	4,65 g/l _{VE-Wasser}	2,00 mM

2.1.7 TALEN-Proteine*gfp-spezifische TALEN*

Die kodierenden Sequenzen für das auf einen spezifischen Sequenzabschnitt innerhalb von *Egfp* abzielende *TALEN*-Paar wurden von *Collectis BioResearch* (Paris, Frankreich) bezogen. Jede der 3.132 bp großen DNA-Kassetten wurde in einen Expressionsvektor unter Kontrolle eines *UBIQUITIN-1 (Ubi1)* -Promotors mit erstem Intron und *NOPALINE SYNTHASE (NOS)* -Terminator überführt (Dr. G. Hensel, *IPK Gatersleben*). Am Transkriptionsstart des offenen Leserahmens wurde zusätzlich eine *S40*-Kernlokalisationssequenz eingefügt. Für die hier beschriebenen Arbeiten wurde das daraus hervorgegangene Vektorpaar *pGH297* und *pGH400* (siehe S. 48, Tabelle 1) genutzt. *pGH297* kodiert die linksseitig der *Spacer*-Region bindenden und *pGH400* die rechtsseitig bindende *TALEN*-Einheit.

Die strukturelle Basis der jeweils 1.061 AS großen Proteine bildet die AS-Sequenz von AvrBs3 aus *X. campestris pv. vesicatoria* (92). Insbesondere der 289 AS große N-

Terminus und die 530 AS umfassende zentrale *Repeat*-Domäne weisen eine Übereinstimmung von über 92% mit der natürlichen AS-Abfolge des Avirulenzproteins auf. Die DNA-Bindedomänen bestehen aus 15 kompletten Wiederholungseinheiten von je 34 AS und ein halbes *Repeat* aus 20 AS. Die Module besitzen die geläufigen *RVD*-Motive HD für Cytosin, NG für Thymin, NI für Adenin und NN für Guanin. Die *RVD*-Abfolgen der beiden *TALE*-Proteine und die entsprechende DNA-Zielsequenz innerhalb von *Egfp* sind in der nachfolgenden Abbildung 7 dargestellt. Neben den zu erwartenden Unterschieden an den *RVD*-Positionen zwölf und dreizehn kommt es bei einigen der *Repeats* an den Positionen vier, 16 und 17 oder auch 24 zu weiteren Abweichungen von der *AvrBs3*-Sequenz. Auf Adenin abzielenden Bindemodule mit einer *RVD* vom Typ NI besitzen bspw. ein Alanin statt Arginin an Position 24. Die Unterschiede an Position vier scheinen dagegen alle Arten von *RVD*'s zu betreffen, in den meisten Fällen handelt es sich um eine Substitution von Glutaminsäure zu Glutamin bei NG- und NN-Motiven oder Alanin zu Glutaminsäure bei HD- und NI-*Repeats*. Die C-terminale Sequenz der *TALE*-Proteine besteht aus nur 11 AS und ist damit im Vergleich zu *AvrBs3* stark verkürzt. Der Rest des C-Terminus wird von einer 13 AS Verknüpfungsregion und der 196 AS großen FokI-Restriktionsdomäne gebildet. Die DNA-Zielsequenz des *TALEN*-Paar befindet sich an Position 362 bis 407 des 720 bp umfassenden *Egfp*-Gens. Sie setzt sich aus zwei je 16 bp großen *TALE*-Bindungsabschnitten, die links- und rechtsseitig einer 12 bp *Spacer*-Region lokalisiert sind zusammen. Inklusive der beiden T₀-Basen besitzt die Zielsequenz damit einen Umfang von 46 bp.



Abb. 7 | a) Abfolgen der *RVD*-Motive der je 16 Wiederholungsmodule umfassenden *gfp*-spezifischen *TALEN*. b) Gesamtzielsequenz des *TALEN*-Paares inkl. T₀-Basen (grün). Links- und rechtsseitig der 12 bp umfassenden *Spacer*-Region sind die beiden je 16 bp umfassenden *TALE*-Bindungsabschnitte (blau unterstrichen) lokalisiert. c) Position der *TALEN*-Zielsequenz innerhalb der 720 bp großen *Egfp*-Sequenz.

Die Restriktionsaktivität der Proteine wurde vom Hersteller mittels Einzelstrang-Hybridisierungstests in Hefezellen geprüft und für „*sehr gut*“ befunden (197). Die Funktionalität der Vektoren und Aktivität der *gfp*-spezifischen *TALEN* konnte auch *in planta* bereits mehrfach demonstriert werden (129, 131, 198).

cenH3a-spezifische TALEN

Die Expressionsvektoren des *TALEN*-Paars mit einer spezifischen DNA-Zielsequenz innerhalb des *centromeric histone 3 α* -Gens (*cenH3a*) von Roggen (*Secale cereale*; *Sc*) wurden mittels *Golden Gate* Assemblierung nach *Cermak et al.* erstellt (S. Hiekel, *PhD*, *IPK* Gatersleben)(114). Die *Repeat*-Sequenzen wurden in einen von *Sangamo BioSciences* (Richmond, CA, USA) bezogenen DNA-Vektor kloniert, welcher die C- und N-terminale AS-Sequenz der *TALEN*-Proteine nach *Cermak et al.* kodiert. Das Rückgrat der beiden Genkassetten bilden Expressionsvektoren vom Typ *pUbi-AB* (siehe S. 47). Am Transkriptionsstart der offenen Leserahmen wurden eine Markierungs- und eine *S40*-Kernlokalisationssequenz eingefügt. Das Plasmid *pSH36* kodiert das 982 AS große, linksseitig der FokI-Restriktionsschnittstelle bindende *TALEN*-Protein. Der Vektor *pSH37* enthält die Sequenz für die 911 AS umfassende, rechtsseitig der *Spacer*-Region bindende Nukleaseeinheit.

Die strukturelle Grundlage der *TALEN*-Proteine bildet eine Gensequenz des von *Cermak et al.* veröffentlichten *TALEN*-kodierenden Vektors *pTAL4* (114, *Supplementary Data*). Die einzelnen Wiederholungsmodule basieren auf dem ersten *Repeat* des *Tal1c*-Effektors von *X. oryzae pv. oryzicola*. Dieser besteht aus allgemein geläufigen *Codons* und eignet sich gut als „*Konsens-Repeat*“ (114). Die *RVD*-Motive der je 34 AS großen Module bestehen aus den AS-Kombinationen HD für Cytosin, NG für Thymin, NI für Adenin und NN für Guanin. Die DNA-Bindedomäne des linksseitig bindenden *TALE*-Proteins ist aus 16,5 *Repeats*, die des rechtsseitig bindenden Proteins aus 14,5 *Repeats* zusammengesetzt. Das letzte halbe Modul besteht aus 20 AS. Die *RVD*-Zusammensetzungen und die DNA-Zielsequenz des *TALEN*-Paars sind in Abbildung 8 dargestellt (siehe S. 45). Die N- und C-terminalen Proteineabschnitte sind Vergleich zum Vorbild deutlich verkürzt. Der N-Terminus umfasst 132 AS und der C-Terminus 46 AS. Jeweils drei AS dienen als Verknüpfung zwischen den *TALE*-Proteinen und der Restriktionsdomäne. Die FokI-kodierende bp-Sequenz wurde stark modifiziert, die genutzten *Codons* bilden auf AS-Ebene aber die exakt gleiche Proteinstruktur. Die DNA-Zielsequenz des *cenH3a*-spezifischen *TALEN*-Paar umfasst

inklusive der beiden T₀-Basen insgesamt 48 bp. Die *Spacer*-Region zwischen den 17 bp und 15 bp großen Bindungsabschnitten beträgt 14 bp.

Die Tauglichkeit der *TALEN*-Architektur nach *Cermak et al.* konnte bereits in Protoplasten von *A. thaliana* gezeigt werden (114). Die Funktionalität der beiden genutzten Expressionsvektoren *pSH36* und *pSH37*, wie auch die Aktivität des kodierten *TALEN*-Paares wurden bislang nicht geprüft.

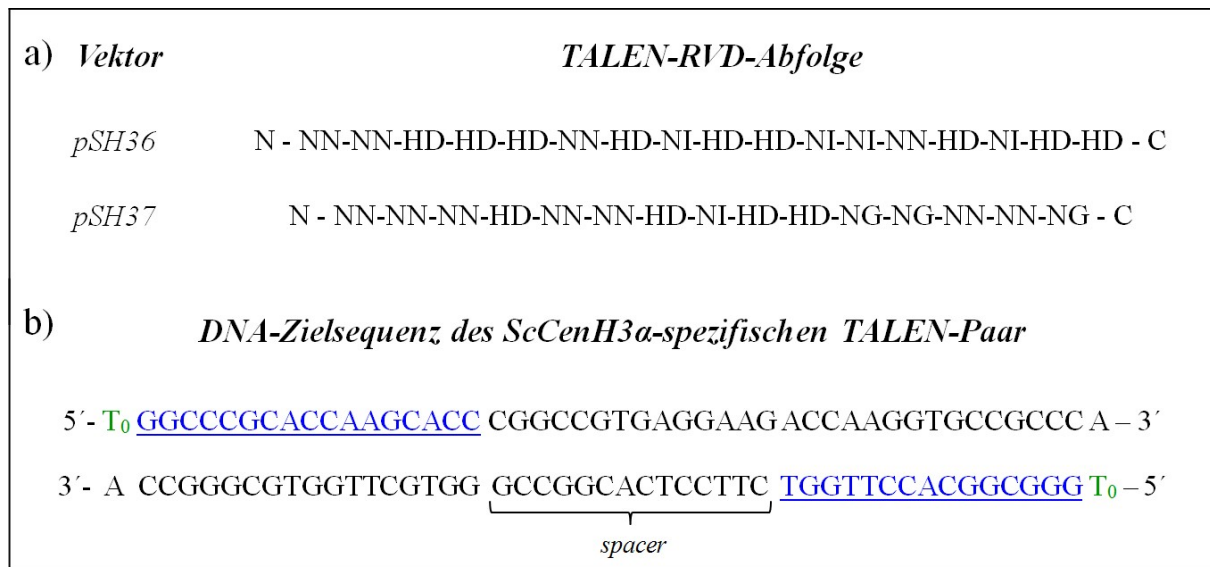


Abb. 8 | a) Abfolgen der *RVD*-Motive des *ScCenH3α*-spezifischen *TALEN*-Paares. Die linksseitig der *Spacer*-Region bindende *TALE*-Domäne umfasst 17, die der rechtsseitigen Bindedomäne 15 Wiederholungsmodule. b) Gesamtzielsequenz des *TALEN*-Paares inkl. T₀-Basen (grün). Die beiden *TALE*-Bindungsabschnitte (blau unterstrichen) sind durch einen 14 bp *Spacer* voneinander getrennt.

mlo-spezifische *TALEN*

Die DNA-Sequenzen für ein auf den Mehltaresistenzloкус (engl.: *mildew resistance locus*, *mlo*) aus Gerste (*Hordeum vulgare*, *Hv*) und Weizen (*Triticum aestivum*, *Ta*) (199, 200) abzielendes *TALEN*-Paar stammt ebenfalls von *Collectis BioResearch* (Paris, Frankreich). *Mlo* steht im Zusammenhang mit einer Infektionsanfälligkeit der Pflanzen für den Mehltauerreger *Blumeria graminis*. Wirte ohne funktionelle *Mlo*-Proteine zeichnen sich durch Resistenz gegenüber dem Pathogen aus (198). Die bezogenen *TALEN*-Genkassetten wurden in *pUbi-AB*-Vektoren (siehe S. 47) kloniert und mit einer Kernlokalisationssequenz und *HA*- bzw. *S*-Markierung ausgestattet (S. Hiekel, *PhD*, *IPK* Gatersleben). Das Plasmid *pSH36* kodiert ein 938 AS großes Peptid, dessen DNA-Bindungsabschnitt strangaufwärts der FokI-Restriktionsschnittstelle liegt. Die aus 944 AS bestehende und strangabwärts der *Spacer*-Region bindende Nukleaseeinheit wird vom Vektor *pSH37* kodiert.

Auch bei den *mlo*-spezifischen *TALEN* handelt es sich um eine modifizierte Version des Avirulenzproteins AvrBs3 aus *X. campestris* pv. *vesicatoria* (92). Der Hersteller macht keine konkreten Angaben bezüglich der Herkunft der *Repeat*-Strukturen oder des *TALEN*-Rückgrats, jedoch zeigt ein Vergleich der bp-Sequenzen von *pSH42* mit *avrBs3* abschnittsweise eine Übereinstimmung von 94 %. Wie bei den *gfp*-spezifischen *TALEN* ist der N-Terminus der Proteine mit 137 AS stark verkürzt. Die 530 AS großen Zentraldomänen setzen sich aus je 15,5 Wiederholungsmodulen zusammen. Jeder *Repeat* besteht aus 34 AS, das letzte halbe Modul aus 20 AS. Die *TALEN* nutzen die *RVD*-Kombinationen HD für Cytosin, NG für Thymin, NI für Adenin und NN für Guanin. Neben den *RVD*-Motiven kommt es zum Teil zu weiteren Abweichungen von der AvrBs3-Struktur innerhalb der einzelnen Sequenzwiederholungen, dies betrifft insbesondere die AS-Positionen vier und vierundzwanzig. Die AS-Substitutionen an diesen Stellen entsprechen denen der zuvor beschriebenen *gfp*-spezifischen Bindedomänen (Vgl. S. 43). Der C-terminale Abschnitt der *TALE*-Proteine besteht aus je 40 AS, gefolgt von einer sechs AS langen Verknüpfung zur 196 AS FokI-Restriktionsdomäne. Die DNA-Zielsequenz innerhalb des *mlo*-Gens umfasst je 16 bp *TALE*-Bindungssequenz und eine dazwischenliegende *Spacer*-Region von 15 bp. Inklusive der beiden T₀-Basen umfasst die Zielsequenz damit 49 bp.

Im Hefe-Hybridisierungstest des Herstellers zeigte das *TALEN*-Paar eine „sehr gute“ Restriktionsaktivität (201). Die Funktionalität der Expressionsvektoren *pSH42* und *pSH43* wurde bislang nicht geprüft.

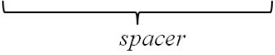
a) <i>Vektor</i>	<i>TALEN-RVD-Abfolge</i>
<i>pSH42</i>	N - HD-NG-NG-HD-HD-NI-NG-NN-NG-HD-NI-HD-HD-NG-NI-NG - C
<i>pSH43</i>	N - NG-NN-NI-NN-NI-HD-NN-NN-HD-NG-NG-NI-NN-NI-NN-NG - C
b)	<i>DNA-Zielsequenz des mlo-spezifischen TALEN-Paar</i>
	5' - T ₀ <u>CTTCCATGTCACCTAC</u> AGCGTCATCATCATG GCTCTAAGCCGTCTCA A - 3'
	3' - A GAAGGTACAGTGGATG TCGCAGTAGTAGTAC <u>CGAGATTCGGCAGAGT</u> T ₀ - 5'
	 <i>spacer</i>

Abb. 9 | **a)** Abfolgen der *RVD*-Motive der je 16 Wiederholungsmodulen umfassenden *mlo*-spezifischen *TALEN*. **b)** DNA-Zielsequenz des *TALEN*-Paares inkl. T₀-Basen (grün). Die beiden *TALE*-Bindungsabschnitte (blau unterstrichen) sind durch eine 15 bp *Spacer*-Region voneinander getrennt.

2.1.8 Plasmide

Expressionsvektor *pUbi-AB*

Die 4.745 bp großen Plasmide vom Typ *pUbi-AB* stammen von *DNA Cloning Service* (Hamburg). Der Expressionsvektor verfügt über einen *Zea maise UBIQUITIN-1* Promotor (*ZmUbi1-P*) samt untransliertem Exon und erstem Intron (*Int*). Die Familie der *Ubi*-Promotoren zeichnet sich durch hohe Aktivitäten in monokotyledonen Pflanzen aus und ist insbesondere für die transiente Expression von Selektionsmarkern und Reportergenen in diesen geeignet (202). Ein Polylinkerabschnitt (*multiple cloning site, MCS*) mit einmalig auftretenden Erkennungssequenzen für verschiedene Restriktionsendonukleasen ermöglicht das Einfügen von DNA-Abschnitten in Folge des starken Promotors. Das *NOPALINE SYNTHASE (NOS)* Fragment von *A. tumefaciens* (203) bildet den Transkriptionsterminator. Das Plasmid verfügt über ein Ampicillinresistenzgen als prokaryotischen Selektionsmarker sowie einen *ColE1*-Replikationsursprung für die Vermehrung in *E. coli*-Zellen.

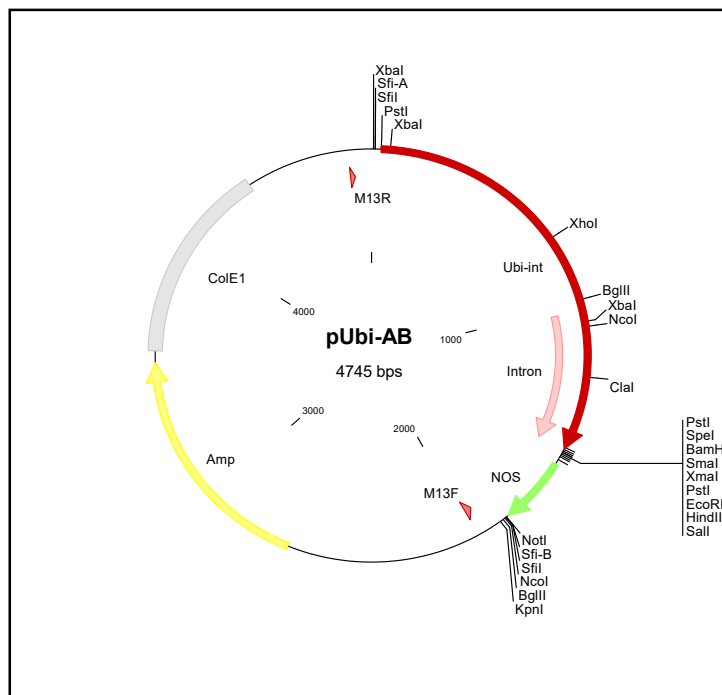


Abb. 10 | Schema des Expressionsvektors *pUbi-AB*. *Ubi-Int*: *Zea maise UBIQUITIN-1* Promotor mit erstem Intron (rote Pfeile); ***PstI* bis *SalI*:** Polylinkerregion mit Erkennungssequenzen für dargestellte Restriktionsendonukleasen; ***NOS*:** *NOPALINE SYNTHASE* Terminator (grüner Pfeil); ***Amp*:** Ampicillinresistenzgen (gelber Pfeil); ***ColE1*:** *E. coli* Replikationsursprung (grau) (Abbildung erstellt mit *Clone Manager*)

Weitere Vektoren

Tabelle 1: Vektoren für Gensequenzamplifikationen

Bezeichnung	Amplifikation von	Herkunft
<i>pGH119</i>	<i>Egfp</i> (720 bp)	Dr. G. Hensel, IPK Gatersleben
<i>pSH04</i>	<i>mCherry</i> (704 bp)	S. Hiekel, PhD, IPK Gatersleben

Tabelle 2: TALEN-kodierende Vektoren

Bezeichnung	bp	Relevante Merkmale	Herkunft
<i>pGH297</i>	14.334	<i>Ubi-P-Int; S40-NLS; Egfp-TALEN_{links}; t NOS; spec_R</i>	Collectis BioResearch, Paris Dr. G. Hensel, IPK Gatersleben
<i>pGH400</i>	14.361	<i>Ubi-P-Int; S40-NLS; HA-Tag; Egfp-TALEN_{rechts}; t NOS; spec_R</i>	Collectis BioResearch, Paris Dr. G. Hensel, IPK Gatersleben
<i>pSH36</i>	7.716	<i>pUbi-AB; S40-NLS; 6xHis-Tag; ScCenH3α-TALEN_{links}; amp_R</i>	Sangamo BioSciences, Richmond S. Hiekel, PhD, IPK Gatersleben
<i>pSH37</i>	7.494	<i>pUbi-AB; S40-NLS; HA-Tag; ScCenH3α-TALEN_{rechts}; amp_R</i>	Sangamo BioSciences, Richmond S. Hiekel, PhD, IPK Gatersleben
<i>pSH42</i>	7.612	<i>pUbi-AB; S40-NLS; HA-Tag; Hv/TaMlo-TALEN_{links}; amp_R</i>	Collectis BioResearch, Paris S. Hiekel, PhD, IPK Gatersleben
<i>pSH43</i>	7.630	<i>pUbi-AB; S40-NLS; S-Tag; Hv/TaMlo-TALEN_{rechts}; amp_R</i>	Collectis BioResearch, Paris S. Hiekel, PhD, IPK Gatersleben

2.1.9 Oligonukleotide

Die für die Amplifikation von Genen und zur Generierung von Oligonukleotidsequenzen verwendeten PCR-Primer wurden von *metabion International AG* (Planegg, Deutschland) bezogen. Die Basensequenzen der einzelnen Primer sind in Tabelle 3 dargestellt (siehe S. 49). Die Synthese von sequenzierungsspezifischen Primern erfolgte ggf. durch *LGC Genomics* (Berlin, Deutschland).

Tabelle 3: PCR- und Oligonukleotidprimer

Primer zur Genamplifikation	
Bezeichnung	Basensequenz (5' → 3')
<i>gfp-1-fw</i>	GCACTGACGCGTTAGTAACGTGAGCAAGGGCGAGGAGCTG
<i>gfp-1-rv</i>	CATAGCGTCGACCTATCATTCCCTTGTACAGCTCGTCCATGCCGAGAGTG
<i>mCh-1-fw</i>	CTGTGCGGATCCCAGTCAATGGTGAGCAAGGGCGAGGAGGATAACATGGC
<i>mCh-1-rv</i>	CCAGCTGAATTCTATAGCCTTGTACAGCTCGTCCATGCCGCCGGTG
<i>gfp-2-fw</i>	GTACGCGGATCCAATATAATGGTGAGCAAGGGCGAGGAGCTGTTC
<i>gfp-2-rv</i>	GTCCGTGAATTCTCACTATTCCTTGTACAGCTCGTCCATGCC
<i>mCh-2-fw</i>	CTGGATGAATTCCACTCAGCGTGAGCAAGGGCGAGGAGGATAAC
<i>mCh-2-rv</i>	GTCACGAAGCTTTCACTACTTGTACAGCTCGTCCATGC
Primer zur Erzeugung von Oligonukleotidsequenzen	
Bezeichnung	Basensequenz (5' → 3')
<i>SP-MCS-fw</i>	CGTCGGACTAGTACGATGGCTTCCAAACCTTTTCTATCTTTGCTTTCACCTTCCTTGC TTCTCTTTACAAGCACATGTTTAGCAGG
<i>SP-MCS-rv</i>	CCGATCGTCGACCCTCCTACGCGTGGACCTAAGCTTCCACCAGAATTCCTTGTGGA TCCCGTTGAGCCTGCTAAACATGTGCTTG
<i>tg-CH3-fw</i>	ATTGCTGAATTCGTACGAGGCCCGCACCAAGCACCCGGCCGTGAGGAAGAC
<i>tg-CH3-rv</i>	TAACGTAAGCTTAAGCCTGTGGGCGGCACCTTGGTCTTCCTCACGGCCGGGTG
<i>tg-Mlo-fw</i>	TGGCGAGAATTCGCGTCTTCCATGTCACCTACAGCGTCATCATCATGGCTCTAAG
<i>tg-Mlo-rv</i>	CGATCTACGCGTGCGGCTTGAGACGGCTTAGAGCCATGATGATGACGCTGTAGGTG
<i>tg-gfp-fw</i>	TCCACTGAATTCGACCCTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG
<i>tg-gfp-rv</i>	ACAGTGGAATTCTCGCGATTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATG
Primer für Testamplifikationen oder Sequenzierungen	
Bezeichnung	Basensequenz (5' → 3')
<i>Bie475</i>	TTTAGCCCTGCCTTCATACG
<i>GH-NOS-R1</i>	GAAAACAAAATATAGCGCGC
<i>HT-Mlo</i>	TCCCACAACGAGGACTACAC

2.2 Methoden

2.2.1 Allgemeine Klonierungsarbeiten

Die Herstellung der in dieser Arbeit verwendeten Plasmide erfolgte mittels *PCR*- und Oligonukleotidklonierungen. In insgesamt sieben Klonierungsschritten wurden dabei vier Testvektoren erstellt. Der generelle Ablauf einer gesamten Klonierung ist in der nachfolgenden Abbildung 11 dargestellt. Die einzelnen Arbeitstechniken werden im folgenden Abschnitt erläutert. Der Aufbau und die Klonierung der Expressionsvektoren *pHT1* bis *pHT4* sind im Kapitel 2.2.2 ab S. 55 beschrieben.

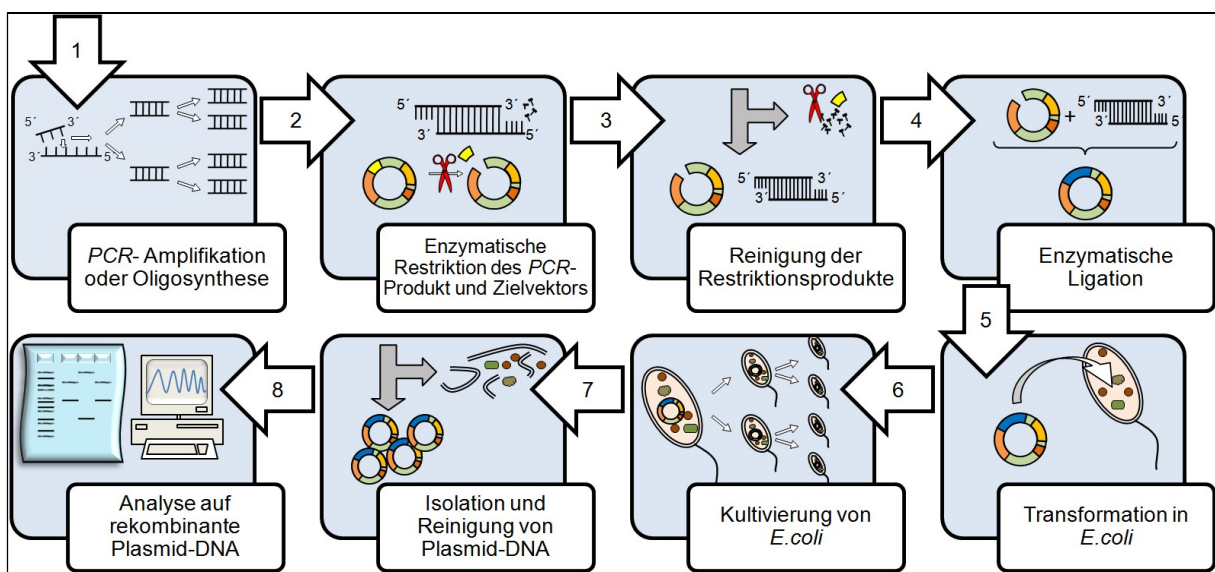


Abb. 11 | Komplettes Ablaufschema eines einzelnen Klonierungsschrittes. 1) *PCR*-basierte Erzeugung des zu integrierenden DNA-Fragments durch Genamplifikation aus einem vorhandenen DNA-Template oder durch die Synthese von Oligonukleotiden mit Hilfe von spezifisch gestalteten Primern. 2) Enzymatische Spaltung des erzeugten Insertionssequenzen und des DNA-Zielvektors mit identischen Restriktionsenzymen zur Erzeugung von kompatiblen Basenüberhängen. 3) Reinigung der Reaktionsprodukte der Restriktion mittels *PCR*-Reinigungs- oder Gelelektrophorese und anschließender DNA-Extraktion mittels Gelelektrophorese. 4) Enzymatische Ligation der Vektor-DNA mit dem zu inserierenden DNA-Fragment 5) Einbringen der Ligationsprodukte in *E. coli*-Zellen mittels *Hitzeschock-Transformation* 6) Kultivierung von transformierten *E. coli*-Zellen unter Selektionsdruck zur Vervielfältigung der eingeschleusten Vektor-DNA 7) Isolation und Reinigung der Vektor-DNA aus den *E. coli*-Suspensionskulturen mittels Plasmidpräparationskit 8) Identifikation von rekombinanten Vektoren mit Hilfe von Testamplifikationen oder -restriktionen und anschließender geleelektrophoretischer Analyse, sowie Sequenzanalysen mittels *Next Generation Sequencing* durch *LGC Genomics* (Berlin).

2.2.1.1 Agarose-Gelelektrophorese

Zur Herstellung von 400 ml Gelen mit Agarosekonzentrationen von 0,8 bis 2,0 % (w/v) wurden pro Gel 200 ml 0,5x TBE-Puffer vorgelegt und die entsprechende Menge Agarose unter Rühren zugesetzt. Die Lösung wurde für 5,0 min. in der Mikrowelle aufgeköcht. Im Anschluss erfolgte der Zusatz von 8,0 µl Ethidiumbromid oder *DNA Stain Clear G* und die

Zugabe weiterer 200 ml 0,5x TBE-Puffer unter rühren. Die noch warme Lösung wurde in eine Gelkammer mit Kämmen gegossen und zur Polymerisation für 30 min. bei RT aufbewahrt. Die Proben wurden ggf. mit 10 % (v/v) *Loading Dye* Probenpuffer versetzt. Die Auftrennung von *PCR*- und Restriktionsprodukten erfolgte über 45 bis 60 min. bei gewöhnlich 150 V und 200 mA Gleichstrom in horizontalen Elektrophoresekammern mit 0,5x TBE als Laufpuffer. Die erhaltenen Gele wurden mittels *ImageMaster VDS Gel Documentation System* dokumentiert. Als Größenstandard diente *Gene Ruler™ DNA Ladder Mix*.

2.2.1.2 Polymerase- Kettenreaktion

Die *PCR* wurde zur Amplifikation der Reportergensequenzen *Egfp* und *mCherry* aus vorhandenen Plasmidtemplates (siehe S. 48, Tab. 1) genutzt. Mit Hilfe der erweiterten 5'-Enden der verwendeten Primer (siehe S. 49, Tab. 3) wurden die Amplikons mit verschiedenen flankierenden Restriktionsschnittstellen versehen. Desweiteren erfolgten *PCR*-basierte Anwendungen zur Identifikation von rekombinanter Vektor-DNA. Testamplifikationen wurden mit isolierten Plasmidlösungen oder auch kleinen Mengen vereinzelter Bakterienkolonien, einer sog. *Kolonie-PCR*, durchgeführt. Die Präparation der Ansätze erfolgte auf Eis, die DNA-Polymerase wurde stets als letztes Reagenz zugesetzt. Nachfolgend sind die Zusammensetzung (siehe S. 52, Tab. 4) und der Verlauf (Tab. 5) eines 20,0 µl Standard-*PCR*-Ansatzes aufgeführt. Für Kolonieamplifikationen wurden 50,0 µl Ansätze vorbereitet und mit der Hilfe von Bakterien-benetzten Zahnstocherspitzen "beimpft". Zur Synthese von ds-Oligonukleotidsequenzen wurden 20,0 µl *PCR*-Ansätze mit je 3,0 µl Primerlösung, jedoch ohne weiteres DNA-Template eingesetzt. Die zueinander komplementären 3'-Enden der Primerpaare wurden nach 3,0 min. Denaturierung einmalig für 45 s bei entsprechender Temperatur miteinander hybridisiert und anschließend für 10,0 min. direkt final elongiert. Zur Ermittlung geeigneter Hybridisierungstemperaturen wurden auf *Thermofisher.com* und *NEB.com* zur Verfügung gestellte Schmelztemperaturkalkulationssoftware genutzt. Die Extensionsreaktionen erfolgten mit *Go-Taq®* DNA-Polymerase.

Tabelle 4: Zusammensetzungen der *PCR*-Reaktionsansätze

Komponente	Genamplifikation Vol. [μ l]	Kolonie- <i>PCR</i> Vol. [μ l]	Oligosynthese Vol. [μ l]
<i>green Go-Taq</i> Reaktionspuffer (5x)	4,0	10,0	4,0
MgCl ₂ (25 mM)	1,5	2,0	1,5
dNTP-Mix (je 10 mM)	1,0	1,5	2,0
Primer (je 10 pM)	je 1,0	je 1,5	je 3,0
DNA-Template	0,5 - 2,0	/	/
nukleasefreies Wasser	9,0 - 10,5	32,5	6,0
<i>Go-Taq</i> [®] DNA-Polymerase (5U/ μ l)	0,5	1,0	0,5
	20,0	50,0	20,0

Tabelle 5: *PCR*-Prozessverlauf von Genamplifikationen

Phase	Dauer	Temperatur [°C]	
Initiale Denaturierung	5 min.	95,0	
Denaturierung	30 s	95,0	← 35x
Hybridisierung	30 - 45 s	58,0 - 68,0	
Extension	45 s	72,0	
Finale Extension	10 min.	72,0	

Zur Analyse vorgesehene *PCR*-Produkte wurden gelelektrophoretisch aufgetrennt und unter UV-Anregung dokumentiert. Für die Klonierung bestimmte Reaktionsprodukte wurden direkt für die enzymatische Restriktion genutzt.

2.2.1.3 Enzymatische Spaltungen von DNA mit Restriktionsendonukleasen

Die für die Restriktionsreaktionen verwendeten *fast digest* Enzyme Acc65I, BamHI, EcoRI, HindIII, MluI, Sall und SpeI erkennen jeweils eine Sequenz von sechs bp und erzeugen vier bp lange, adhäsive 3'-Überhänge (engl.: *sticky ends*) an diesen. Die Zusammensetzung der 20,0 μ l Doppelverdau-Reaktionsansätze für Plasmid- und Testrestriktionen ist in der nachfolgenden Tabelle 6 (siehe S. 53) dargestellt. Die enzymatische Spaltung von *PCR*-Fragmenten erfolgte mit 50,0 μ l Ansätzen. Die Inkubation erfolgte jeweils für eine Stunde bei 37°C.

Tabelle 6: Zusammensetzungen der Restriktionsreaktionsansätze

Komponente	Plasmidrestriktion Vol. [μ l]	PCR-Produktrestriktion Vol. [μ l]
<i>fast digest green</i> Reaktionspuffer (10x)	2,0	5,0
Restriktionsenzyme (je 20 - 100 U/ μ l)	je 1,0	je 1,0
Vektor-DNA (0,2 - 1,0 μ g)	0,5 - 1,0	/
PCR-Produktlösung	/	10,0
nukleasefreies Wasser	15,0 - 15,5	33,0
	20,0	50,0

Restriktionsprodukte von DNA-Fragmenten kleiner als 100 bp wurden mit einem PCR-Reinigungskit nach den Angaben des Herstellers behandelt (*QIAGEN*, Hilden). Verdaute Plasmid-DNA und PCR-Produkte größer als 500 bp wurden gelelektrophoretisch aufgetrennt. Die Extraktion der gewünschten DNA aus den Gelbanden erfolgte mit Hilfe eines Gelelutionskits nach dem Protokoll des Herstellers (*QIAGEN*, Hilden). Proben von Testverdaureaktionen wurden ebenfalls mittels Gelelektrophorese aufgetrennt und anschließend unter UV-Anregung dokumentiert.

2.2.1.4 Ligation von DNA-Fragmenten

Für die *in vitro* Rekombinationen wurde mit den beiden entsprechend gleichen Restriktionsenzymen vorbehandelte Vektor- und Insert-DNA verwendet. Pro Ligation erfolgten zwei Reaktionsansätze im Verhältnis 1:3 und 1:7 (Vektor: Insertionsfragment; v/v). Als Enzym diente T4-DNA- oder *Quick*-Ligase. In den nachfolgenden Tabelle 7 und 8 sind die Zusammensetzungen der Ligationsansätze und die Inkubationsbedingungen aufgeführt.

Tabelle 7: Zusammensetzungen der Ligationsansätze mit T4-DNA-Ligase

Komponente	Vol. [μ l]	Vol. [μ l]
Ligase-Puffer (10x)	2,0	3,0
Vektor-DNA-Lösung	3,0	3,0
Insert-DNA-Lösung	9,0	21,0
Nuklease-freies Wasser	5,0	2,0
T4-DNA-Ligase (3U/ μ l)	1,0	1,0
	20,0	30,0
Inkubation: 12 bis 16 h bei RT		

Tabelle 8: Zusammensetzungen der Ligationsansätze mit *Quick-Ligase*

Komponente	Vol. [μ l]	Vol. [μ l]
<i>Quick-Ligase-Puffer</i> (2x)	15,0	25,0
Vektor-DNA-Lösung	3,0	3,0
Insert-DNA-Lösung	9,0	21,0
nukleasefreies Wasser	2,0	-
<i>Quick-Ligase</i>	1,0	1,0
	30,0	50,0
Inkubation: 15 bis 20 min. bei RT		

2.2.1.5 Transformation von *E. coli* mittels Hitzeschock-Verfahren

Die Bakterientransformationen erfolgten mit jeweils 100 μ l superkompetenten *E. coli XL-1 blue*-Aliquotes. Nach fünf minütigem auftauen auf Eis wurden diesen jeweils 3,0 μ l eines auf 4,0 °C gekühlten Ligationsansatzes oder Plasmidlösung zugesetzt und für 35,0 min. auf Eis inkubiert. Im Anschluss erfolgte ein zwei minütiger Hitzeschock bei 42 °C im Thermoblock und das sofortige Abkühlen auf Eis für weitere zwei min. Nach der Zugabe von 450 ml *SOC*-Regenerationsmedium wurden die Bakterien für eine Stunde bei 37 °C und 550 *rpm* inkubiert.

2.2.1.6 Kultivierung von *E. coli*

Zur Selektion und Anzucht von Einzelkolonien wurden jeweils 50,0 μ l Transformationsansatz auf *LB_{Amp}*- oder *LB_{Spec}*-Platten ausgestrichen und über Nacht bei 37 °C im Brutschrank inkubiert. Zum Beimpfen von Suspensionskulturen wurden vereinzelte Kolonien entnommen und in sterile Kulturröhrchen mit 2,0 ml *LB*-Flüssigmedium und Antibiotikazusatz überführt. Die Zellernte erfolgte nach einer 24 h Inkubation bei 37 °C und 180 *rpm*. Für die ausschließliche Vermehrung von Plasmiden gepickte Bakterienkolonien wurden in 100 ml Kulturkolben mit 20,0 ml *LB*-Medium samt Antibiotika überführt und für 24 bis 36 h bei sonst gleichen Bedingungen inkubiert. In vereinzelten Fällen wurden auch 20,0 ml *LB*-Medium mit 100 μ l von vorhandenen Glycerinstocks beimpft.

2.2.1.7 Isolierung und Reinigung von Plasmid-DNA

Die präparative Isolierung von Plasmid-DNA aus den *E. coli*-Suspensionskulturen erfolgte unter Zuhilfenahme eines Plasmidpräparationskits (*QIAGEN*, Hilden) und wurde nach dem Protokoll des Herstellers durchgeführt. Pro Reaktionsansatz wurden dabei 2,0 bis 6,0 ml einer Übernachtkultur in Reaktionsgefäße überführt und für 3,5 min. mit 13.000 *rpm* zentrifugiert.

Nach der Resuspension der Pallets wurden die Zellen alkalisch lysiert und über eine Silikatsäulenmatrix von den restlichen Zellbestandteilen getrennt. Die Elution der Plasmide erfolgte mit 10,0 oder 20,0 µl Nuklease-freien Wassers.

2.2.1.8 Konzentrationsbestimmungen von Plasmid-DNA

Die Bestimmung von DNA-Konzentrationen erfolgte über die absorptionspektrometrische Messung der optischen Dichte bei einer Wellenlänge von 260 nm mittels *NanoDrop 2000c*-Spektrophotometer (*Thermo Scientific*, Wilmington, DE, USA). Pro Messung wurden dabei 2,0 µl Plasmidpräparationsprodukt aufgetragen, als Referenzwert dienten 2,0 µl Nuklease-freies Wasser. Aus den Verhältnissen der ermittelten OD_{260nm}- und OD_{280nm}-Werte kann eine Aussage über mögliche Proteinkontaminationen in den Lösungen getroffen werden.

2.2.1.9 Identifikationen rekombinanter Plasmid-DNA

Zur Bestimmung von rekombinanter Vektor-DNA wurden Restriktionsreaktionen und *PCR*-Amplifikationen mit isolierten Plasmidlösungen oder entnommenen Bakterienkolonien und den jeweils zugehörigen Kontrollen durchgeführt. Die Auftrennung der Reaktionsprodukte erfolgte durch Gelelektrophorese. Die Gele wurden unter UV-Anregung dokumentiert und rekombinante Plasmide anhand des Restriktionsbandenmusters identifiziert. Nach jedem Klonierungsschritt wurde eine Auswahl potentieller Kandidaten von *LGC Genomics* (Berlin, Deutschland) mittels *Next Generation Sequencing* sequenziert.

2.2.2 Klonierung der Testvektoren

2.2.2.1 Modifikation des Vektorrückgrats *pUbi-AB*

Als Grundlage für die Erstellung der Testvektoren erfolgte im ersten Klonierungsschritt die Substitution der 41 bp großen Polylinkerregion strangabwärts der Promotorsequenz des Basisplasmid *pUbi-AB* mit einer 134 bp umfassenden DNA-Sequenz, die neben den erforderlichen Restriktionsschnittstellen *SpeI*, *BamHI*, *EcoRI*, *HindIII*, *MluI* und *SallI* ein Transkriptionsstart-*Codon* (*atg*) und eine darauf folgende *Legumin B4*-Signalpeptidsequenz der Ackerbohne (*Vicia faba*) enthält. Zur Erzeugung der zu inserierenden Sequenz wurden die Oligonukleotidprimer *SP-MCS-fw* und *SP-MCS-rv* (siehe S. 49, Tab. 3) im *PCR-Thermocycler* bei 46 °C einmalig hybridisiert und anschließend für 10 min. bei 72 °C elongiert. Die Restriktionsreaktionen der *PCR*-Produkte und des Basisvektors *pUbi-AB* erfolgten mit *SpeI* und *SallI*. Die Restriktionsprodukte wurden gereinigt und im Anschluss zwei Ligationsansätze mit T4-DNA-Ligase (siehe S. 53, Tab. 7) über Nacht inkubiert. Von

den im weiteren Verlauf gewonnen *E. coli*-Kulturen wurden zwölf Einzelkolonien gepickt und in je 2,0 ml LB_{Amp} -Medium vermehrt. Aus den Kulturen isolierte Plasmid-DNA wurde durch Testrestriktionen mit SpeI und Acc65I überprüft und einzelne Vektoren anschließend von *LGC Genomics* (Berlin) sequenziert. Das erzeugte Plasmid wird im weiteren Textverlauf unter der Bezeichnung *pUbi-AB_M* geführt.

2.2.2.2 Klonierung des generischen Testvektor *pHT1*

In zwei Klonierungsschritten wurde aus dem zuvor beschriebenen Konstrukt *pUbi-AB_M* der Vektor *pHT1* erzeugt (siehe S. 64, Abb. 16). Im ersten Schritt erfolgte die Insertion der 720 bp großen *Egfp*-Sequenz. Das Gen wurde mit Hilfe der PCR-Primer *gfp-1-fw* und *gfp-1-rv* (siehe S. 49, Tab. 3) vom Plasmid *pGH119* amplifiziert, die Hybridisierungen erfolgten bei 60 °C. Über die erweiterten 5'-Enden der verwendeten Primer wurden vor und hinter der *Egfp*-Sequenz je ein Paar Stopp-Codons und die Erkennungssequenzen für MluI und Sall angefügt. Die Restriktionen des Vektorrückgrats und der PCR-Produkte erfolgten mit selbigen Enzymen. Die Produkte wurden gereinigt, über Nacht ligiert und in *E. coli XLI-blue* transformiert. Die Kultivierung der Bakterien erfolgte mit LB_{Amp} -Medium. Isolierte Plasmidproben wurden einem Testverdau mit BamHI und Sall unterzogen und sechs der Vektoren anschließend sequenziert. Im zweiten Klonierungsschritt wurde die 704 bp umfassende Sequenz von *mCherry* in das Plasmid integriert. Zur PCR-Amplifikation des Gens wurde das Primerpaar *mCh-1-fw* und *mCh-1-rv* (siehe S. 49, Tab. 3) genutzt, als DNA-Template diente der Vektor *pSH04*. Die Hybridisierungsreaktionen erfolgten bei 62 °C. Über die erweiterten 5'-Enden der Primer wurden der *mCherry*-Sequenz die benötigten flankierenden Erkennungssequenzen für BamHI und EcoRI angefügt. Nach dem Restriktionsverdau mit den genannten Enzymen und einer Reinigung der Produkte erfolgten zwei Ligationsansätze mit *Quick-Ligase* (siehe S. 54, Tab. 8). Aus *E. coli*-Kulturen gewonnene Plasmidproben wurden durch Testrestriktionen mit BamHI und EcoRI überprüft und fünf der Vektoren zur Sequenzierung übergeben.

2.2.2.3 Testvektor *pHT2* für das *gfp*-spezifische *TALEN*-Paar

Die Generierung des 6.277 bp großen Testvektors *pHT2* zur Ermittlung der Nukleaseaktivität des *gfp*-spezifische *TALEN*-Paares verlief im Prinzip wie die direkt zuvor beschriebenen Klonierung von *pHT1*. Da sich die Zielsequenz der *TALEN* in diesem Fall innerhalb des *Egfp*-Reportergens befindet, musste eine entsprechende strukturelle Anpassung der Expressionskassette vorgenommen werden (siehe S. 65, Abb. 19). Im ersten

Klonierungsschritt erfolgte die Insertion der um zwei bp des exprimierten Leserasters versetzten *mCherry*-Sequenz in den Zwischenvektor *pUbi-AB_M* (S. 55). Für die Amplifikation des Reportergens vom Plasmid *pSH04* wurden die Primer *mCh-2-fw* und *mCh-2-rv* (siehe S. 49, Tab. 3) genutzt, die Hybridisierungstemperatur betrug dabei 62 °C. Die erweiterten 5'-Enden der Primer enthalten zudem die *mCherry*-flankierenden Stopp-Codons, sowie die zur Integration in das Vektorrückgrat genutzten Restriktionsschnittstellen EcoRI und HindIII. Die Ligationsreaktionen erfolgten mit *Quick-Ligase* (siehe S. 54, Tab. 8). Mit den im weiteren Verlauf gewonnenen Plasmidproben wurden Testrestriktionen mit EcoRI und HindIII durchgeführt. Anschließend erfolgte die Sequenzierung von fünf der Vektoren. Der DNA-Abschnitt der im zweiten Klonierungsschritt inserierten *Egfp*-Sequenz wurde mit Hilfe des Primerpaars *gfp-2-fw* und *gfp-2-rv* (siehe S. 49, Tab. 3) erzeugt. Als DNA-Template diente der Vektor *pGH119*, die Hybridisierungen erfolgten bei 60 °C. Die Erkennungssequenzen der zur Restriktion genutzten Enzyme BamHI und EcoRI wurden über die erweiterten 5'-Enden der Primer an das Genfragment angefügt. Die Ligation erfolgte mit T4-DNA-Ligase über Nacht bei RT (siehe S. 53, Tab. 7). Zur Identifikation von rekombinanten Plasmiden wurden PCR-Testamplifikationen mit den Primern *gfp-2-fw* und *gfp-2-rv* durchgeführt. Vier der Vektoren wurden im Anschluss zur Sequenzierung übergeben.

2.2.2.4 Testvektor *pHT3* für das *cenH3 α* -spezifische TALEN-Paar

Vom Plasmid *pHT1* ausgehend erfolgte die Generierung des 6.276 bp großen Testvektors *pHT3*, der in einem Klonierungsschritt mit einer DNA-Zielsequenz für das *cenH3 α* -spezifische TALEN-Paar ausgestattet wurde (siehe S. 66, Abb. 20). Das 46 bp umfassende Fragment setzt sich aus 17 bp links- und 15 bp rechtsseitiger TALE-Bindungssequenz und einem dazwischenliegenden *Spacer*-Abschnitt von 14 bp zusammen. Zur Erzeugung der DNA-Sequenz wurden die Oligonukleotidprimer *tg-CH3-fw* und *tg-CH3-rv* (siehe S. 49, Tab. 3) im PCR-Thermocycler bei 60 °C einmalig hybridisiert und über die angefügten Restriktionsschnittstellen EcoRI und HindIII zwischen den Fluoreszenzreportergensequenzen des Vektors *pHT1* inseriert. Die Ligation erfolgte mit *Quick-Ligase* (siehe S. 54, Tab. 8). Isolierte Plasmidproben wurden mittels Testrestriktionen mit EcoRI und HindIII überprüft und fünf der Vektoren anschließend sequenziert. Bei der Klonierung ist auf die korrekte Integration des TALEN-Zielabschnitts zu achten. Der bp-Versatz der *Egfp*-Sequenz muss erhalten bleiben um die Funktionalität des *Frame-shift*-Mechanismus zu gewährleisten.

2.2.2.5 Testvektor *pHT4* für das *mlo*-spezifische TALEN-Paar

Beim 6.315 bp umfassenden Testvektor *pHT4* wurde ebenfalls ausschließlich die DNA-Zielsequenz für das spezifisch innerhalb der *mlo*-Region bindende TALEN-Paar in den generischen Vektor *pHT1* integriert (siehe S. 67, Abb. 22). Das 63 bp große Fragment wurde mit Hilfe der Oligonukleotidprimer *tg-Mlo-fw* und *tg-Mlo-rv* (siehe S. 49, Tab. 3) im *Thermocycler* erzeugt und über die Restriktionsschnittstellen EcoRI und MluI in den Basistestvektor *pHT1* ligiert. Aus transformierten *E.coli*-Kulturen isolierte Plasmid-DNA wurde mittels PCR-Amplifikation überprüft, als Primer dienten dabei *tg-Mlo-fw* und *GH-NOS-RI*. Sieben der Proben mit einer entsprechenden Gelbande im Bereich von 1.000 bp wurden im Anschluss sequenziert. Auch bei dieser Klonierung musste die korrekte Insertion des Fragments beachtet werden, sodass *Egfp* im inaktiven Leseraster des Testplasmids vorliegt.

2.2.3 Transformation und Analyse von Gerstenzellen

2.2.3.1 Biolistischer Gentransfer

Benetzen von Goldkugeln mit Plasmid-DNA

Die im Durchmesser 1,0 μm großen Goldpartikel wurden in 1,5 ml Reaktionsgefäße eingewogen, mit 100 % EtOH versetzt ($c = 30,0 \mu\text{g/ml}$) und bis zur Verwendung bei $-20 \text{ }^\circ\text{C}$ gelagert. Die Präparation der Ansätze und das anschließenden Partikelbombardement erfolgten in der Sterilwerkbank, alle Reagenzien wurden dabei auf Eis gekühlt. Für die Transformation von Zellen mit verschiedenen Plasmiden (*Co-Transformation*) wurde die vorgesehene Vektor-DNA zunächst separat bis zu einer Gesamtmenge von 5,0 μg gemischt. Versuche mit identischen Plasmiden erfolgten mit 3,0 bis 5,0 μg DNA. Pro Ansatz wurden 30,0 μl Goldsuspension unter vortexen entnommen und in 0,5 ml Reaktionsgefäße überführt. Das Gold wurde mittels zwei minütiger Zentrifugation bei 13.000 *rpm* sedimentiert und der Überstand vorsichtig mit einer Pipette entfernt. Anschließend wurden zwei identische Waschschrte mit je 100,0 μl Nuklease-freien Wasser durchgeführt. In Abhängigkeit von den Konzentrationen der Plasmidlösungen erfolgte die Zugabe einer entsprechenden Menge Nuklease-freien Wassers für ein finales Reaktionsvolumen von 65,0 μl . Die resuspendierten Goldpartikel wurden zum Entklumpen für 45 bis 60 s im Ultraschallbad geschwenkt. Direkt im Anschluss erfolgte die Benetzungsreaktion unter permanentem vortexen. Als erstes wurde die jeweilige Plasmidlösung zugesetzt, unmittelbar danach je 10,0 μl Spermidin (0,1 M) und 25,0 μl CaCl_2 (2,5 M). Nach weiteren zwei min. vortexen wurde der Reaktionsmix kurz

zentrifugiert und der Überstand verworfen. Die DNA-Gold-Ansätze wurden zweimal in 100 % EtOH gewaschen und für den Beschuss in 15,0 μ l reinem EtOH resuspendiert.

Partikelbombardement

Für die Transformationsexperimente an Gerste diente jeweils nur das erste Keimblatt von acht bis vierzehn Tage alten Pflanzen. Pro Versuch erfolgten zwei Schüsse auf die untere Seite von zwei Blättern, welche mit Rührmagneten in einer Petrischale mit Inkubationsmedium (siehe S. 41) fixiert und mittig in der Schussskammer positioniert wurden. Als Projektile dienten je 5,0 μ l der Gold-DNA-Suspensionen, welche auf Trägerscheiben pipettiert und zur Verflüchtigung des EtOH kurz luftgetrocknet wurden. Nach dem Einsetzen aller Komponenten wurde in der Kammer ein Unterdruck von 27 *in Hg* (entspricht 0,91 bar) erzeugt und anschließend der Beschussdruck mit Helium aufgebaut. Die zum Einsatz gekommenen Reiss-Scheiben halten einem Druck von 1.100 *psi* (entspricht 75,84 bar) stand. Bei Erreichen des Werts zerplatzen sie, wodurch der aufgebaute Druck schlagartig entweicht und die Trägerscheibe samt Goldkugeln in Richtung Zielgewebe beschleunigt wird. Ein eingesetztes Stahlsieb stoppt den Träger, während die Partikel sich lösen und mit hoher Geschwindigkeit in die Zellen eindringen. Beschossene Blattproben wurden auf denselben Medienplatten für 32 bis 40 h in Dunkelheit und bei RT inkubiert.

2.2.3.2 Expressionsanalysen mittels Licht- und konfokaler laser scanning Mikroskopie

Als Nachweis der *Egfp*- und *mCherry*-Expression in den Gerstenzellen wurden Messungen mittels Licht- oder *confocal laser scanning*-Mikroskop (*CLSM*) durchgeführt. Zur lichtmikroskopischen Detektion der *Egfp*-Expression wurden die Proben mit UV-Licht angeregt, für *mCherry* wurde eine Lichtquelle im grünen Spektralbereich genutzt. Als quantitativer Nachweis der Reportergenexpressionen erfolgten Analysen am *CLSM*. Dazu wurden die Blattproben mit Hilfe einer Vakuumpumpe zunächst entgast und auf Objektträgern präpariert. Zur Messung der Intensität der *Egfp*-Emissionen wurden die Proben mit einem Argon-Laser im Wellenlängenbereich $\lambda = 488$ nm angeregt, was dem Anregungsmaximum des Signalpeptids entspricht. Die Emission wurde durch einen Bandpassfilter im Bereich von 505 bis 530 nm detektiert. Das Emissionsmaximum von *Egfp* liegt bei $\lambda = 509$ nm. Die Anregung der Proben zur Ermittlung der Intensität der *mCherry*-Emissionen erfolgte mit einem Helium-Neon-Laser der Wellenlänge $\lambda = 561$ nm. Die Detektion erfolgte durch einen Bandpassfilter im Bereich von 570 bis 620 nm. Das Emissionsmaximum des Proteins liegt bei 610 nm. Die Messungen wurden mit Hilfe des

Programms *ZEN 2* (*Carl Zeiss*, Jena) dokumentiert und ausgewertet. Jede der erfolgreich transformierten Zellen wurde auf die Bildung der beiden Signalproteine überprüft und deren Anzahl dokumentiert. Die untersuchten Proben wurden nach der Analyse und Auszählung verworfen.

3 Ergebnisse

Das für die Testvektoren entwickelte Konzept zur Funktionalitätsprüfung und quantitativen Ermittlung der nukleolytischen Aktivität von Sequenz-spezifisierten Endonukleasen basiert auf der Fehleranfälligkeit der natürlichen DNA-Reparaturmechanismen infolge von gezielt induzierten Doppelstrangbrüchen innerhalb der Testplasmidsequenzen. Das entwickelte Testsystem nutzt einen *Frame-shift*-Mechanismus durch *InDel*-Mutationen zur Reaktivierung eines Reportergens, das zunächst im inaktiven Leseraster des Testvektors vorliegt (siehe S. 62, Abb. 12). Ein zweites, N-terminal lokalisiertes und innerhalb des exprimierten Leserahmens befindliches Reportergergen wird als Marker zur Identifikation von erfolgreich transformierten Zellen genutzt. Das zentrale Element des Testsystems bildet die Polylinkerregion mit einmalig auftretenden Restriktionsschnittstellen zwischen den beiden Reportergergen. Sie dienen zum Einklonieren von gewünschten DNA-Testsequenzen für Sequenz-spezifizierte Endonukleasen. Sind diese in der Lage einen DSB innerhalb der inserierten Zielsequenz zu induzieren, kommt es infolge von fehlerhaften NHEJ-vermittelten DNA-Reparaturen zu *InDel*-Mutationen. Resultieren diese Mutationen in einem *Frame-shift* der gesamten bp-Abfolge hinter der Restriktionsschnittstelle, welcher zu einer Korrektur des Leserahmens der C-terminalen Reportergergensequenz führt, kommt es zur Bildung eines Fusionsproteins aus zwei Fluoreszenzreportern. Als genereller Expressionsindikator kommt dabei das rotfluoreszierende Protein mCherry (204) und als Restriktionsindikator das grünfluoreszierende Egfp (engl.: *Enhanced green fluorescent protein*) (205) zum Einsatz. Aus dem Verhältnis der Gesamtanzahl an transformierten, mCherry-bildenden Zellen zur Anzahl der ebenfalls *Egfp*-exprimierenden Zellen, in denen ein *Frame-shift* infolge einer erfolgreichen DSB-Induktion durch die Sequenz-spezifisierten Endonukleasen stattgefunden hat, kann eine Aussage über die Restriktionsaktivität der Nukleasen im Wirtsorganismus getroffen werden.

Von der nachfolgenden Darstellung abweichend (S. 62, Abb. 12) wurde zudem die Klonierung eines Testvektors für ein spezifisch auf *gfp*-abzielendes *TALEN*-Paar umgesetzt (siehe S. 65, Abb. 19). Im Gegensatz zu den anderen zur Verfügung stehenden *TALEN*-Paaren konnten die sehr guten Restriktionseigenschaften der *gfp*-spezifischen *TALEN* in Gerste bereits mehrfach demonstriert werden (129, 131). Aus diesem Grund sollte der Testvektor zur Überprüfung des Versuchskonzepts und der Funktionalität des Testmechanismus verwendet werden. Zudem könnten sich die ermittelten Aktivitätswerte als Referenzgröße für die

Einschätzung und Beurteilung der erzielten Nukleaseaktivitäten der übrigen *TALEN*-Paare eignen.

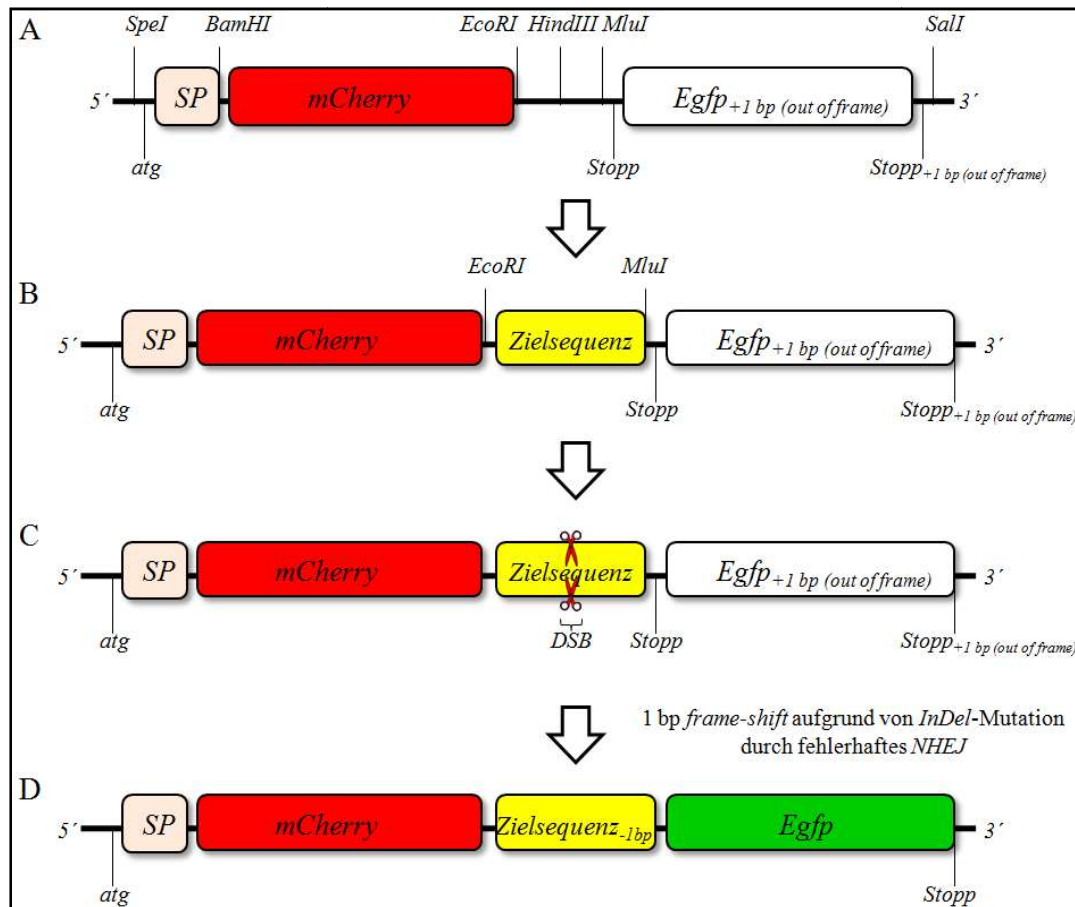


Abb. 12 | Konzept der Testvektoren zur Ermittlung der *in planta* Restriktionsaktivitäten von Sequenz-spezifischen Endonukleasen. A) Der generische Vektor (*pHT1*) dient als Ausgangspunkt für die Konstruktion von Testvektoren mit individuell gestaltbaren DNA-Zielsequenzen. B) Über die Restriktionsschnittstellen *EcoRI*, *HindIII* und *MluI* können in verschiedene DNA-Zielsequenzen (gelb) für diverse Sequenz-spezifizierte Endonukleasen in den Testvektor inseriert werden. C) Kommt es innerhalb der Zielsequenz zu einer DSB-Induktion kann ein *Frame-shift* als Folge von *InDel*-Mutationen stattfinden. D) Ein signifikanter Anteil der *InDel*-Mutationen führt zur Korrektur des Reporter-gen-Leserahmens, infolge dessen es zur Bildung eines Fusionsproteins aus mCherry (rot) und dem zuvor nicht exprimierten Egfp (weiß, dann grün) kommt.

3.1 Klonierung der Testvektoren

Die Klonierungsarbeiten zur Generierung des Basistestvektors *pHT1* und der genspezifischen Testplasmide *pHT2* für *gfp*-, *pHT3* für *cenH3α*- und *pHT4* für *mlo*-spezifische *TALEN*-Paare wurden größtenteils entsprechend der Planungen umgesetzt.

3.1.1 Modifikation des Vektorrückgrats *pUbi-AB*

Die Substitution des 41 bp *SpeI*-*SalI*-Fragments von *pUbi-AB* mit der 134 bp großen Oligonukleotidsequenz, deren Polylinkerregion als Ausgangspunkt für die weiteren

Vektorklonierungen diente, verlief ohne Komplikationen. Bei der Testrestriktion mit SpeI und Acc65I zeigten vier der zwölf Plasmidproben eine Bande im erwarteten Bereich von 440 bp (Abb. 13). Die Sequenzanalyse bestätigte dem Vektor mit der Nummer zehn eine vollständige Übereinstimmung mit der geplanten DNA-Sequenz.

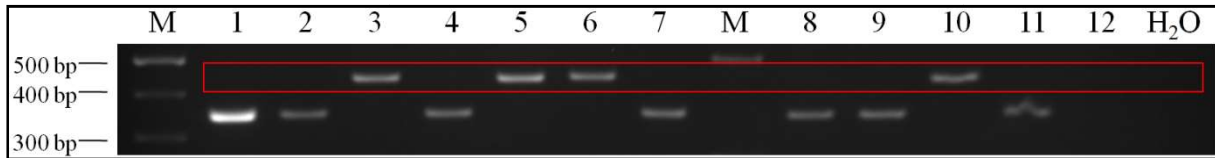


Abb. 13 | Ausschnitt des Elektrophoresegels zur Identifikation von rekombinanten *pUbi-AB*-Vektoren.

3.1.2 Klonierung des generischen Testvektors *pHT1*

In zwei weiteren erfolgreichen Klonierungsschritten konnte das 6.276 bp große Plasmid *pHT1* aus dem modifizierten Vektor *pUbi-AB_M* erzeugt werden. Nach der Insertion der *gfp*-Reportergensequenz zeigten neun von zehn isolierten Plasmidproben in einer Testrestriktion mit BamHI und Sall eine entsprechende Gelbande im Bereich von 770 bp (Abb. 14). Die Sequenzanalyse von sechs der Vektoren ergab eine Übereinstimmung des achten Plasmids mit der gewünschten DNA-Sequenz. Im zweiten Klonierungsschritt kam es zu einer einzelnen Punktmutation an der bp-Position 2.355. Es handelt sich um eine Substitution von Adenin zu Thymin. Die direkt strangaufwärts der EcoRI-Erkennungssequenz liegende Mutation hat jedoch keinen Einfluss auf die Reportergenexpression oder die Funktionalität des *Frame-shift*-Mechanismus und wurde daher übernommen. Im Testverdau mit EcoRI und BamHI zeigten fünf von zehn isolierten Proben eine Restriktionsbande im erwarteten Bereich von 725 bp (Abb. 15). Nach der Sequenzierung der Vektoren wurde das Plasmid Nummer vier für die weiteren Arbeiten ausgewählt und unter der Bezeichnung *pHT1* weitergeführt.



Abb. 14 | Ausschnitt des Elektrophoresegels zur Identifikation von *pUbi-AB_M* Vektoren mit inserierter *Egfp*-Sequenz.

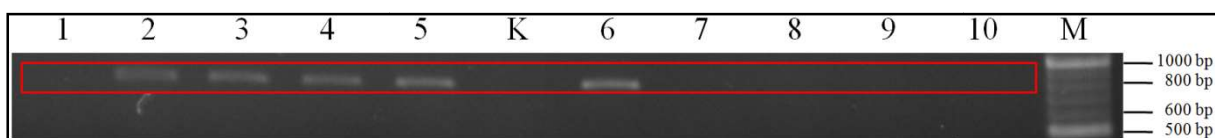


Abb. 15 | Ausschnitt des Elektrophoresegels zur Identifikation von *pHT1*.

Der Vektor *pHT1* ist so konzipiert, dass die erzeugte Expressionskassette zunächst nur die *Legumin B4*-Signalpeptid- und die *mCherry*-Sequenz beinhaltet. *mCherry* befindet sich generell im exprimierten Leserahmen der Testplasmid-DNA und fungiert als genereller

Reporter. Aufgrund einer einzelnen, zusätzlich eingefügten Base zwischen der ersten, strangaufwärts befindlichen Stopp- und der darauffolgenden *Egfp*-Sequenz befindet sich *Egfp* nicht im exprimierten Leseraster der DNA, sondern ist bereits *out of frame*. Die vorangestellten Stopp-Codons liegen hingegen im offenen Leserahmen der Expressionskassette und verhindern eine fortlaufende Transkription der folgenden *non-sense*-Sequenz wenn kein *Frame-shift* erfolgt ist.

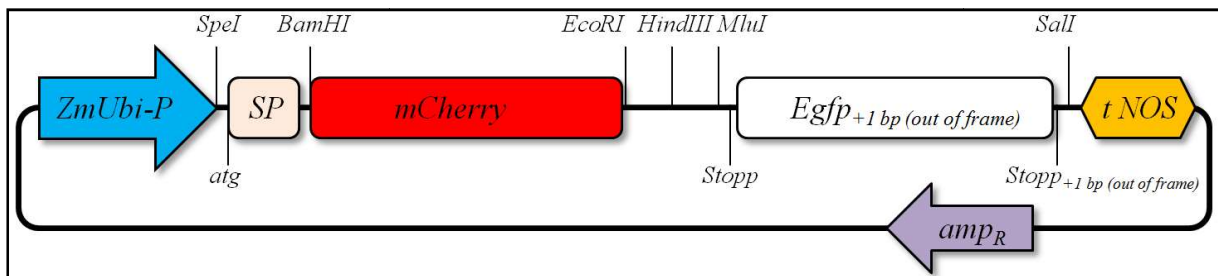


Abb. 16 | Schematische Darstellung des generischen Vektors *pHT1*. Der unter Kontrolle eines *Ubiquitin-1*-Promotors aus *Zea mays* (*ZmUbi-P*) stehenden offene Leserahmen der Expressionskassette kodiert ein *Legumin B4*-Signalpeptidsequenz (SP) und die als Transformationsindikator vorgesehene Reportergensequenz des rotfluoreszierenden *mCherry*-Proteins. Die strangabwärts folgenden Restriktionsschnittstellen *EcoRI*, *HindIII* und *MluI* ermöglichen das Einklonieren von DNA-Zielabschnitte für verschiedene Sequenz-spezifizierte Endonukleasen. Eine zweite Reportergensequenz (*Egfp*) befindet sich in einem inaktiven Leserahmen des Vektors (+1 bp *out of frame*). Die Sequenz soll im weiteren Verlauf als Indikator für einen *Frame-shift* infolge von *InDel*-Mutationen durch fehlerhafte NHEJ-vermittelte DSB-Reparaturen dienen. *t NOS*: *NOPALINE SYNTHASE* Terminator; *amp_R*: Ampicillinresistenzgen

3.1.3 Klonierung des Testvektors *pHT2* für das *gfp*-spezifische *TALEN*-Paar

Die Generierung des ebenfalls in zwei Klonierungsschritten aus dem modifizierten Plasmid *pUbi-AB_M* erzeugten Testvektor *pHT2* für das *gfp*-spezifische *TALEN*-Paar konnte erfolgreich umgesetzt werden. Die nach der *mCherry*-Insertion gewonnenen Plasmidproben wurden durch Testrestriktion mit *EcoRI* und *HindIII* oder mittels Kolonie-PCR auf *mCherry* überprüft. Sechs der insgesamt dreizehn Kandidaten zeigten eine deutliche Gelbande im Bereich von 724 bp (S. 65, Abb. 17). Die mit fünf der Vektoren durchgeführten Sequenzanalysen ergaben für die Proben eins und drei eine Übereinstimmung mit dem geplanten DNA-Konstrukt. Im zweiten Klonierungsschritt konnte die 738 bp große *Egfp*-Sequenz in den Vektor integriert werden. Der Nachweis von rekombinanter Plasmid-DNA erfolgte durch PCR-Testamplifikationen mit den zuvor genutzten Primern *gfp-2-fw* und *gfp-2-rv* (siehe S. 49, Tab. 3). Sieben der zehn untersuchten Proben zeigten dabei eine Bande im Bereich von 756 bp (S. 65, Abb. 18), von denen vier daraufhin sequenziert wurden. Der 6.277 bp große Vektor mit der Nummer fünf besitzt die gewünschte Sequenz und wurde als *pHT2* weitergeführt.



Abb. 17 | Ausschnitt des Elektrophoresegels zur Identifikation von *pUbi-AB_M* Vektoren mit inserierter *mCherry*-Sequenz.

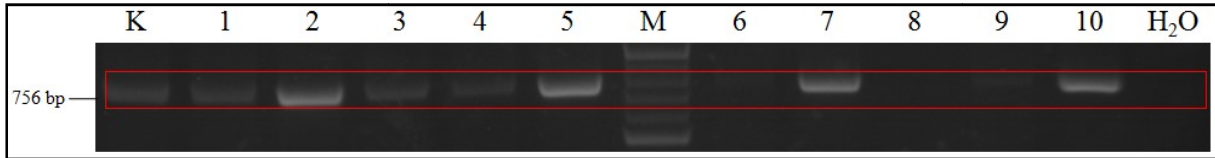


Abb. 18 | Ausschnitt des Elektrophoresegels zur Identifikation von *pHT2*.

Die Funktionsweise von *pHT2* beruht auf dem gleichen *Frame-shift*-Mechanismus wie die der anderen Testvektoren. Allerdings bildet bei *pHT2* das *Egfp* den generellen Reporter zur Identifikation von erfolgreich transformierten Gerstenzellen (Abb. 19). Die Sequenz befindet sich am Transkriptionsstart des offenen Leserahmens der DNA und enthält zugleich die *TALEN*-Zielsequenz. Die zum exprimierten Leseraster um zwei bp versetzte *mCherry*-Sequenz dient als Restriktionsindikator zur Bestimmung der Aktivität der *gfp*-spezifischen *TALEN*. Sie wird lediglich im Falle eines *Frame-shifts* um $n-2$ oder $n+1$ bp, wobei n einem kompletten Basentriplett entspricht, gebildet. Ein *Frame-shift* führt zu einer Veränderung der gesamten Sequenz strangabwärts der Restriktionsschnittstelle, wodurch auch der C-terminale Abschnitt von *Egfp* partiell *out of frame* ist.

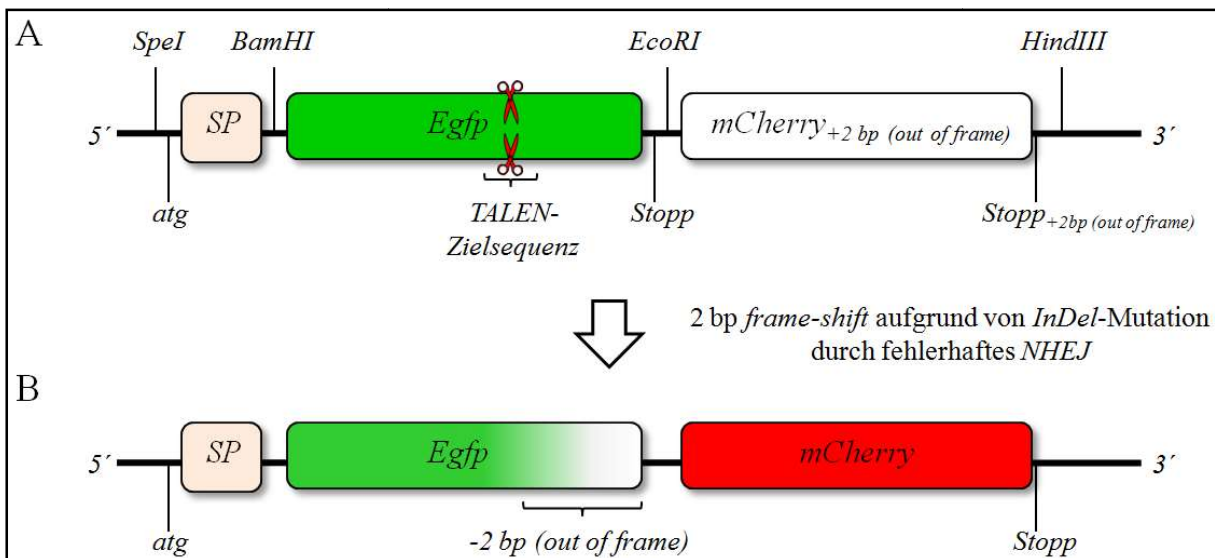


Abb. 19 | Schematische Darstellung des Aufbaus und der Funktionsweise des Testvektors *pHT2*. **A)** Die Funktionsweise des Testvektors *pHT2* zur Ermittlung der Restriktionsaktivität der *gfp*-spezifizierten *TALEN* beruht ebenfalls auf einem *Frame-shift*-Mechanismus. Da sich die Zielsequenz der *TALEN* in diesem Fall jedoch innerhalb der *Egfp*-Sequenz (grün) befindet, wurden die Reportergensequenzen im Vergleich zu den anderen Testvektoren vertauscht. *Egfp* bildet den generellen Indikator zur Identifikation von erfolgreich transformierten Zellen, während die *mCherry*-Sequenz (weiß) um +2 bp gegenüber dem exprimierten Leseraster versetzt ist und das Protein als Restriktionsindikator dient. **B)** Kommt es zu einer DSB-Induktion innerhalb der *TALEN*-Zielsequenz, in deren Folge ein durch *InDel*-Mutationen vermittelter *Frame-shift* stattfindet, wird die Funktionalität von *mCherry* in einen signifikanten Anteil der Zellen wieder hergestellt (rot). Die *Egfp*-

Sequenz befindet sich dann partiell außerhalb des aktiven Leserasters der DNA (grün-weiß), was zu einer beeinträchtigten Emission durch die resultierende *non-sense*-Sequenz führen kann. Der Testvektor *pHT2* nutzt das Plasmidrückgrat *pUbi-AB* unter Kontrolle eines *Zea maize Ubiquitin-1*-Promotors und *NOPALINE SYNTHASE*-Terminators. Der Vektor verfügt zudem über eine Ampicillinresistenz als Selektionsmarker (nicht dargestellt). *SP*: *Legumin B4*-Signalpeptidsequenz

3.1.4 Klonierung des Testvektors *pHT3* für das *cenH3α*-spezifische *TALEN*-Paar

pHT3 ist der erste vom generischen Plasmid *pHT1* abstammende Testvektor, der lediglich mit einer 46 bp DNA-Zielsequenz für ein spezifisch darauf abzielendes *TALEN*-Paar ausgestattet wurde. Nach der Insertion der Oligonukleotidsequenz zeigten drei der fünf aus transformierten *E.coli*-Kolonien isolierten Plasmidlösungen bei Testrestriktionen mit *Bam*HI und *Hind*III eine Gelbande im erwarteten Bereich von 792 bp (nicht dargestellt). Die Sequenzierung der Vektoren ergab für die Proben zwei und fünf eine Übereinstimmung mit der geplanten Struktur. Das 6.330 bp große Plasmid mit der Nummer fünf wurde ausgewählt und unter der Bezeichnung *pHT3* für die Versuche an Gerstenzellen genutzt. Die T₀-Base am 5'-Ende des linksseitigen *TALEN*-Bindungsabschnitts der *cenH3α*-Zielsequenz war aufgrund eines Planungsfehlers bei den Oligonukleotidprimern allerdings nicht vorhanden. Die übrige DNA-Sequenz entspricht jedoch dem Entwurf und eignet sich für die Erprobung der *TALEN* und des *Frame-shift*-Mechanismus.

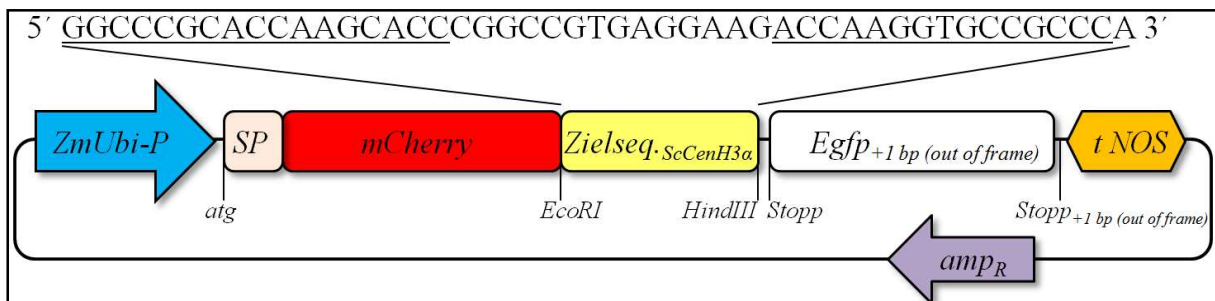


Abb. 20 | Schematische Darstellung des Testvektors *pHT3*. Der mit einer 48 bp großen DNA-Zielsequenz (obere Basensequenz, gelb) für die *cenH3α*-spezifizierte *TALEN* ausgestattete Vektor wurde vom Plasmid *pHT1* abgeleitet. Die *Legumin B4*-Signalpeptid- (*SP*) und *mCherry*- (rot) Sequenz befinden sich im aktiven Leseraster der Expressionskassette. Die strandabwärts der Zielsequenz lokalisierte Sequenz von *Egfp* (weiß) befindet sich hingegen zunächst nicht im exprimierten Leseraster des Testvektors, sondern ist um +1 bp zu diesem versetzt. Kommt es innerhalb der Zielsequenz zu einer DSB-Induktion durch das *TALEN*-Paar, die letztlich in einem entsprechenden *Frame-shift* resultiert, wird die *Egfp*-Reportersequenz reaktiviert und es kommt zur Bildung eines *mCherry-Egfp*-Fusionsproteins. **ZmUbi-P**: *Zea maize Ubiquitin-1*-Promotor; **tNOS**: *NOPALINE SYNTHASE* Terminator; **amp_R**: Ampicillinresistenzgen

In der dargestellten Form wird die Transkription direkt hinter der Zielsequenz abgebrochen. Die Expression des *Egfp*-Restriktionsindikators wird lediglich nach einer erfolgreichen DSB-Induktion innerhalb der Zielsequenz und einem resultierenden *Frame-shift* der nachfolgenden Sequenz um n-1 oder n+2 bp ermöglicht, wobei n einem kompletten Basentriplett entspricht.

3.1.5 Klonierung des Testvektors *pHT4* für das *mlo*-spezifische *TALEN*-Paar

In einer weiteren Klonierung konnte die 49 bp große Zielsequenz für das *mlo*-spezifische *TALEN*-Paar in den Vektor *pHT1* inseriert werden (Abb. 22). Von den 15 gewonnen und mittels Testamplifikation überprüften Plasmidproben zeigen acht ein *PCR*-Produkt mit einer Gelbande im erwarteten Bereich von 1.013 bp auf (Abb. 21). Sieben der Proben wurden im Anschluss sequenziert, wobei der Vektor Nummer eins die korrekte Basensequenz aufwies. Das 6.315 bp umfassende Plasmid *pHT4* bildet den dritten erfolgreich generierten Testvektor zur Ermittlung der *in planta* Endonukleaseaktivität eines spezifizierten *TALEN*-Paares mittels *Frame-shift*-Mechanismus.

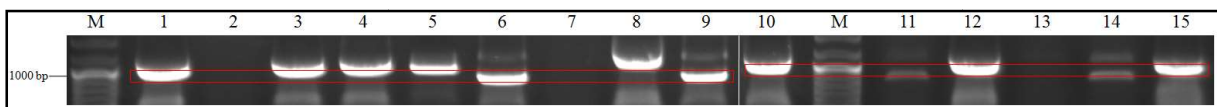


Abb. 21 | Ausschnitt des Elektrophoresegels zur Identifikation von *pHT4*.

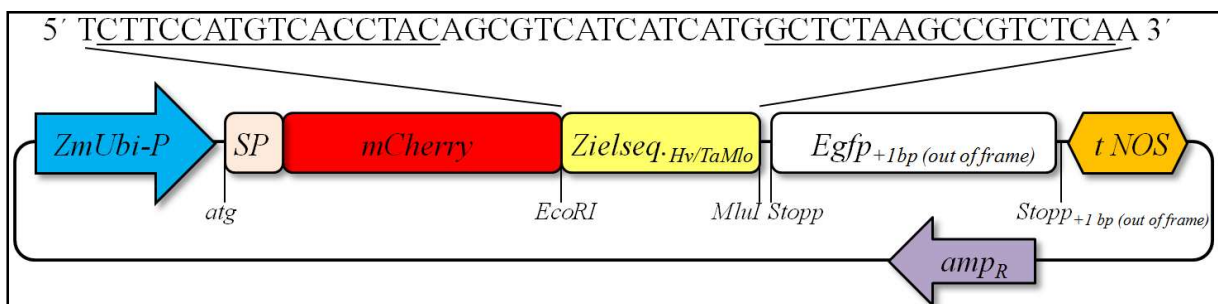


Abb. 22 | Schematische Darstellung des Testvektors *pHT4*. Der mit einer 49 bp großen DNA-Zielsequenz (obere Basensequenz, gelb) für das *mlo*-spezifische *TALEN*-Paar ausgestattete Vektor wurde ebenfalls vom Plasmid *pHT1* abgeleitet verfügt über genau den gleichen strukturellen Aufbau (abgesehen von der Zielsequenz) wie der zuvor beschriebene Testvektor *pHT3* (siehe S. 66, Abb. 20). **ZmUbi-P**: *Zea mays* Ubiquitin-1-Promotor; **SP**: *Legumin B4*-Signalpeptidsequenz; **tNOS**: *NOPALINE SYNTHASE* Terminator; **amp_R**: Ampicillinresistenzgen

Die Vervielfältigung der erzeugten Testvektoren in transformierten *E.coli*-Flüssigkulturen ergab je nach Menge an eingesetztem Medium und der Kultivierungsdauer gereinigte Plasmidlösungen mit Konzentrationen von 0,75 bis 1,23 µg/µl.

3.2 Expressionsversuche an Gerstenzellen

3.2.1 Benetzungsreaktionen und Partikelbombardment

Die Benetzung der Test- und *TALEN*-Plasmide an die Goldpartikel verliefen überwiegend erfolgreich. Während der Experimente kam es zum Teil zu stark variierenden Goldkonzentrationen in den erhaltenen Reaktionslösungen. Die Menge an verbliebenen Au-DNA-Partikeln innerhalb der Ansätze fiel dabei mitunter zu hoch oder sehr gering aus, was direkte Auswirkungen auf die Gesamtzahl transformierter Zellen hatte. Nach einer Anpassung

des Transformationsprotokolls und der Optimierung der Handhabung konnte die Qualität der Benetzungsreaktionen hinreichend standardisiert werden. Insgesamt wurden 24 Co-Bombardmentexperimente mit den verschiedenen Test- und dazugehörigen *TALEN*-kodierenden Vektoren an Gerstenblättern durchgeführt und in die Auswertungen einbezogen. Daneben fanden mehrere Vor- und Kontrollversuche statt.

3.2.2 Expressionsanalysen am Licht- und konfokalen *Laser Scanning*-Mikroskop

3.2.2.1 Testvektor *pHT1*

Eine erste Kontrolle und Überprüfung der Funktionalität der Reporterexpression erfolgte mit dem generischen Testvektor *pHT1*. Die Auswertung am Licht- und konfokalen *Laser Scanning*-Mikroskop ergab eine deutliche Expression von *mCherry* in einzelnen Epidermal- und Stomatazellen der Gerstenblätter. *Egfp* konnte dagegen nicht nachgewiesen werden. Dies entspricht den Erwartungen an die geplante und durch Sequenzanalyse bestätigte Expressionskassette. *mCherry* befindet sich im offenen Leseraster der DNA und wird als Indikator in jeder der durch das Partikelbombardment transformierten Zelle gebildet. Die Expression von *Egfp* wird hingegen durch die strangaufwärts der Reportersequenz befindlichen Transkriptionsstopp-*Codons* unterbunden.

3.2.2.2 Testvektor *pHT2* für das *gfp*-spezifische *TALEN*-Paar

Von den mit *pHT2* und den *gfp*-*TALEN*-kodierenden Vektoren *pGH297* und *pGH400* durchgeführten Co-Bombardmentversuchen wurden lediglich die ersten Vier exemplarisch ausgezählt und in die Auswertung einbezogen (siehe S. 69, Tab. 9). Als Kontrolle wurde *pHT2* ohne die *TALEN*-Plasmide mitgeführt (Tab. 10). Die Analysen der Proben am *CLSM* zeigten in allen Fällen deutlich *Egfp*-exprimierende Zellen (S. 69, Abb. 23). Entgegen den Erwartungen konnte jedoch in keiner der transformierten Zellen eine Fluoreszenz des Restriktionsindikators *mCherry* nachgewiesen werden. Auch die Variation der Konditionen der Benetzungsreaktion, des DNA-Mengenverhältnisses von Testvektor zu *TALEN*-Plasmiden und Versuche mit Zwischenvektoren, die während der Klonierung von *pGH297* und *pGH400* entstanden sind, blieben erfolglos. Damit konnte in keiner der durchschnittlich 74 transformierten Zellen die Bildung des Doppelreporterfusionsproteins infolge eines *Frame-shifts* um $n-2$ oder $n+1$ bp nach einer gezielten DSB-Induktion durch die *Egfp*-spezifischen *TALEN* nachgewiesen werden.

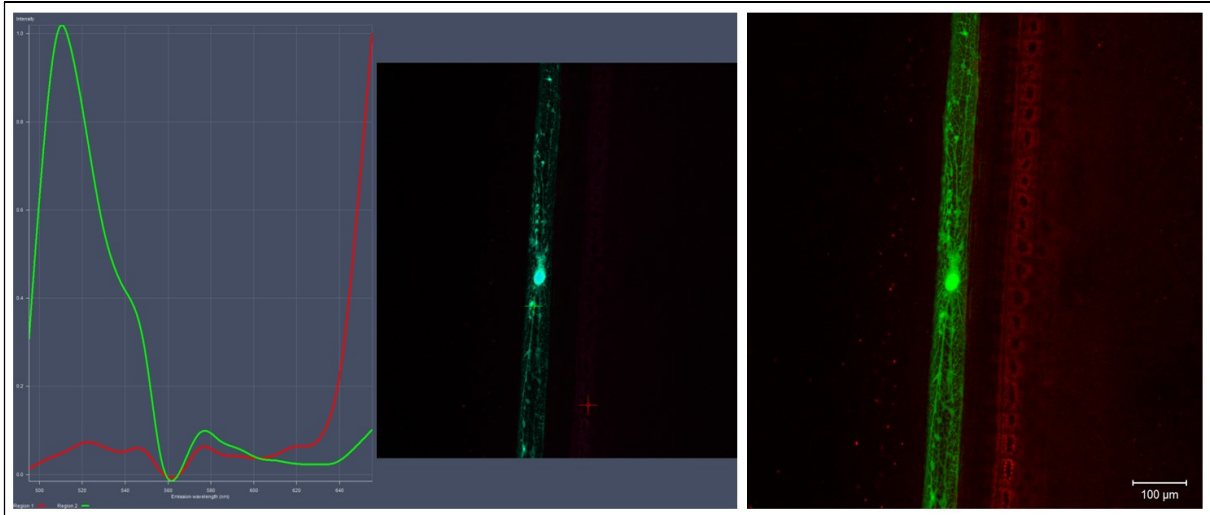


Abb. 23 | CLSM-Aufnahme der *Egfp*-Emission von einer mit *pHT2* und den *gfp*-spezifischen *TALEN*-kodierenden Vektoren Co-transformierte Gerstenzelle. Der grüne Peak stellt den auf die fluoreszierende Zelle gerichteten Messpunkt dar. Zu sehen ist eine deutliche Emission im Wellenlängenbereich von $\lambda = 509$ nm, was dem Emissionsmaximum von *Egfp* entspricht. Eine Emission im Bereich von $\lambda = 610$ nm, dem Emissionsmaximum von mCherry, ist nicht zu erkennen. Der rote Messpunkt dient der Kontrolle.

Tabelle 9: Ergebnisse der Expressionsanalysen mit *pHT2* und dem *gfp*-spezifischen *TALEN*-Paar

Versuch	Anzahl <i>Egfp</i> Zellen	Anzahl <i>mCherry</i> Zellen	Verhältnis <i>mCherry/Egfp</i> [%]
1	26	0	0
2	35	0	0
3	123	0	0
4	110	0	0
Durchschnitt	74 ± 43	0	0

Tabelle 10: Ergebnisse der Kontrollversuche mit *pHT2*

Kontrolle	Anzahl <i>Egfp</i> Zellen	Anzahl <i>mCherry</i> Zellen
1	44	0
2	120	0
Durchschnitt	82 ± 38	0

3.2.2.3 Testvektor *pHT3* für das *cenH3 α* -spezifische *TALEN*-Paar

Mit dem Testvektor *pHT3* und den *cenH3 α* -*TALEN*-kodierenden Plasmiden *pSH36* und *pSH37* wurden insgesamt 13 Co-Bombardmentexperimente durchgeführt. Die mitgeführten

Kontrollversuche erfolgten dabei ohne die *TALEN*-Plasmide. Die Expression der beiden Reporterproteine wurde zunächst am *CLSM* überprüft. Dabei konnte in einer Vielzahl von transformierten Epidermal- und Stomatazellen die zu erwartende, intensiv rote Fluoreszenz von mCherry im Wellenlängenbereich von $\lambda = 610$ nm detektiert werden (siehe S. 72, Abb. 24). Desweiteren zeigte ein gewisser Anteil dieser Zellen eine schwache bis z.T. sehr deutliche Expression des Restriktionsindikators *Egfp*. Die Kontrollen wiesen ebenfalls eine große Anzahl an mCherry-bildenden Zellen auf, *Egfp* konnte dagegen nicht nachgewiesen werden. Dies war ein erster Beleg für die Funktionalität des Testvektorsystems mittels *Frame-shift*-Mechanismus. Die Bildung des Doppelreporterfusionsproteins wird durch fehlerhafte DNA-Reparaturen infolge der *cenH3 α -TALEN* induzierten DSB innerhalb der inserierten Zielsequenz von *pHT3* ermöglicht. Zur Verifikation der Messungen und Ermittlung der *TALEN*-Restriktionsaktivität wurden im Anschluss umfangreichere Bombardmentexperimente durchgeführt. Die Auswertung einer ersten Versuchsreihe erfolgte am Lichtmikroskop. Die Ergebnisse sind in Tabelle 11 (S. 71) dargestellt, die der Kontrolle in Tabelle 12 direkt darunter aufgeführt. Aufgrund der starken Autofluoreszenzen des Gewebehintergrunds und der im Vergleich dazu oftmals schwachen Intensität der *Egfp*-Fluoreszenz erwies sich die lichtmikroskopische Analyse im Fall der *cenH3 α -spezifischen TALEN* jedoch als nicht praktikabel. Eine eindeutige, rein visuelle Erkennung von *Egfp*-exprimierenden Zellen war oftmals nicht möglich. Daher wurde zunächst sowohl die Mindestanzahl an zweifelsfrei *Egfp*-bildenden Zellen, als auch die maximale Anzahl an tendenziell *Egfp*-bildenden Zellen erfasst und beide Werte zur Ermittlung der *in planta* Restriktionsaktivität der *TALEN* mit der Anzahl an *mCherry*-exprimierenden Zellen ins Verhältnis gesetzt. Von den im Durchschnitt 140 transformierten Zellen zeigten dabei mindestens 14,6 % und maximal 32,2 % die Bildung des Reporterfusionsproteins infolge eines durch DSB-Induktion herbeigeführten *Frame-shifts* um n-1 oder n+2 bp auf. Die Varianz zwischen Mindest- und Maximalanzahl beträgt damit mehr als das Doppelte. Eine eindeutige *Egfp*-Fluoreszenz konnte in keiner der Kontrollen festgestellt werden.

Tabelle 11: Ergebnisse der lichtmikroskopischen Expressionsanalysen von Versuchen mit *pHT3* und den *cenH3 α* -spezifischen *TALEN*

Versuch	mCherry	Egfp (min.)	Verhältnis [%]	Egfp (max.)	Verhältnis [%]
1	166	24	14,46	61	36,75
2	74	6	8,11	26	35,14
3	102	13	12,75	30	29,41
4	81	16	19,75	28	34,57
5	278	43	15,47	81	29,14
Durchschnitt	140 \pm 76	20 \pm 13	14,55	45 \pm 22	32,24

Tabelle 12: Ergebnisse der am Lichtmikroskop ausgewerteten Kontrollversuche mit *pHT3*

Kontrolle	Anzahl mCherry Zellen	Anzahl Egfp Zellen
1	117	0
2	51	0
3	138	0
Durchschnitt	102 \pm 37	0

Aufgrund der unzureichenden Genauigkeit der lichtmikroskopischen Untersuchungen und Auswertungen wurden die Analysen zur Ermittlung einer realistischen *TALEN*-Restriktionsaktivität in einer zweiten Versuchsreihe komplett mittels quantitativer Emissionsmessungen am *CLSM* durchgeführt. Die Ergebnisse der Co-Bombardements sind in der nachfolgenden Tabelle 13 (S. 72), die der mitgeführten Kontrollen in Tabelle 14 dargestellt. Im Durchschnitt zeigten dabei 23 von 64 transformierten Zellen die Bildung beider Reporterfluoreszenzen, was einem Anteil von 35,6 % entspricht. Der erzielte Wert weicht mit 3,3 % nur geringfügig von der lichtmikroskopisch ermittelten, maximalen Anzahl an Egfp-bildenden Zellen ab, was für die geringe Intensität von Egfp in einer Vielzahl von Zellen spricht. Die dokumentierten Messungen am *CLSM* belegen die oftmals nur schwache Expression des Fluoreszenzgens im Vergleich zu normal *Egfp*-exprimierenden Zellen (Vgl. S. 69, Abb. 23). Dennoch sind die hier erzielten Ergebnisse ein erster Beleg für die Funktionalität und Aktivität der *ScCenH3 α* -spezifischen *TALEN* in Gerste.

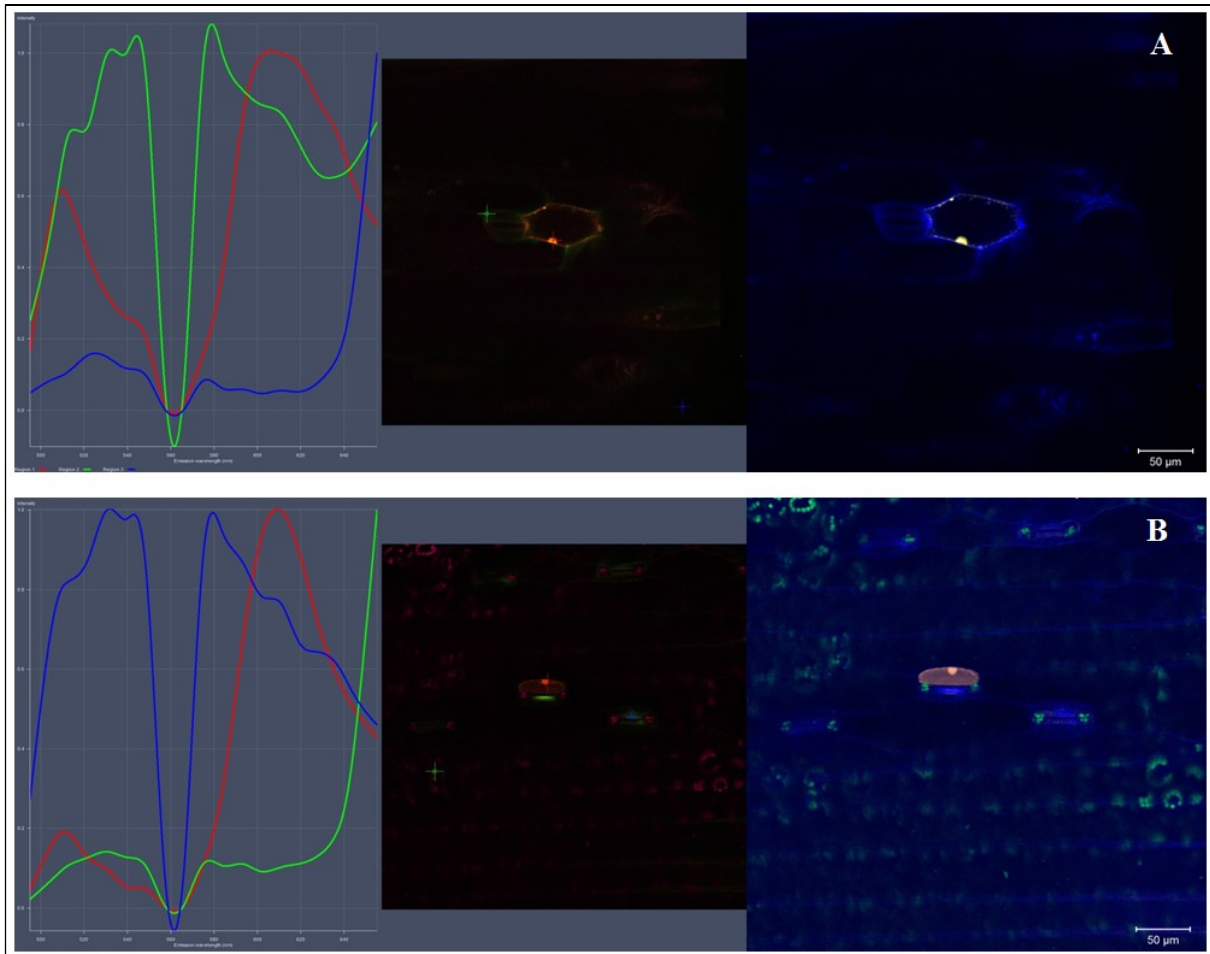


Abb. 24 | CLSM-Aufnahmen von Gerstenzellen, die mit *pHT3* und den *cenH3 α* -spezifischen TALEN-kodierenden Vektoren Co-transformiert wurden. Die roten Graphen besitzen jeweils deutliche Peaks im Wellenlängenbereich von $\lambda = 610$ nm und belegen zweifelsfrei die Expression des *mCherry*-Reportergens. Die Intensität der Egfp-Emission variiert hingegen sehr stark, deutlich erkennbar an den roten Peaks im Bereich von $\lambda = 509$ nm. Die Zelle A zeigt eine etwa 40 % stärkere Emission gegenüber der Zelle B.

Tabelle 13: Ergebnisse der am CLSM ausgewerteten Expressionsversuche mit *pHT3* und dem *cenH3 α* -spezifischen TALEN-Paar

Versuch	Anzahl mCherry Zellen	Anzahl Egfp Zellen	Verhältnis Egfp/mCherry [%]
1	51	17	33,33
2	59	21	35,59
3	73	16	21,92
4	77	29	37,66
5	157	62	39,49
6	13	3	23,08
7	38	17	44,74
8	44	17	38,64
Durchschnitt	64 ± 40	23 ± 16	35,55

Tabelle 14: Ergebnisse der am *CLSM* ausgewerteten Kontrollversuche mit *pHT3*

Kontrolle	Anzahl <i>mCherry</i> Zellen	Anzahl <i>Egfp</i> Zellen
1	27	0
2	143	0
3	83	0
4	43	0
5	48	0
Durchschnitt	69 ± 41	0

3.2.2.4 Testvektor *pHT4* für das *mlo*-spezifische *TALEN*-Paar

Von den sieben mit *pHT4* und den Vektoren *pSH42* und *pSH43* für das *HvMlo*-/ *TaMlo*-spezifische *TALEN*-Paar durchgeführten Co-Bombardementversuchen wurden zwei am *CLSM* und fünf am Lichtmikroskop ausgewertet. Die Ergebnisse der Expressionsanalysen sind in der nachfolgenden Tabelle 15 (S. 75) aufgeführt, die der Kontrollen in Tabelle 16 (S. 75). Neben der erwarteten Bildung von *mCherry* zeigten die Beobachtungen und Messungen bei allen Proben zudem eine große Anzahl an ebenfalls deutlich bis intensiv grün fluoreszierenden Zellen (siehe S. 74, Abb. 25), während in den mitgeführten Kontrollen keinerlei *Egfp* nachgewiesen werden konnte. Von den pro Experiment durchschnittlich 57 transformierten Zellen besaßen 40 Zellen neben der *mCherry*- auch eine *Egfp*-Fluoreszenz, was einem Anteil von 70,2 % entspricht. Eine Übersicht der in den einzelnen Versuchen erzielten Nukleaseaktivitäten ist in Tabelle 15 (S. 74) zu finden. Die im Vergleich zu den Versuchen mit *pHT3* (Vgl. S. 72) wesentlich stärkere Intensität von *Egfp* ermöglichte eine einfache und i.d.R. zweifelsfreie Identifikation des Reporters am Lichtmikroskop. Die Wiederherstellung des Fusionsproteins infolge von gezielt induzierten DSB innerhalb der Testvektorsequenz ist ein weiterer Beleg für die Funktionalität und Eignung des *Frame-shift*-Mechanismus für das Endonukleasentestsystem. Auch im Fall von *pHT4* konnte die Restriktionsaktivität des *HvMlo*-/ *TaMlo*-spezifischen *TALEN*-Paars in Gerste erstmals ermittelt und belegt werden.

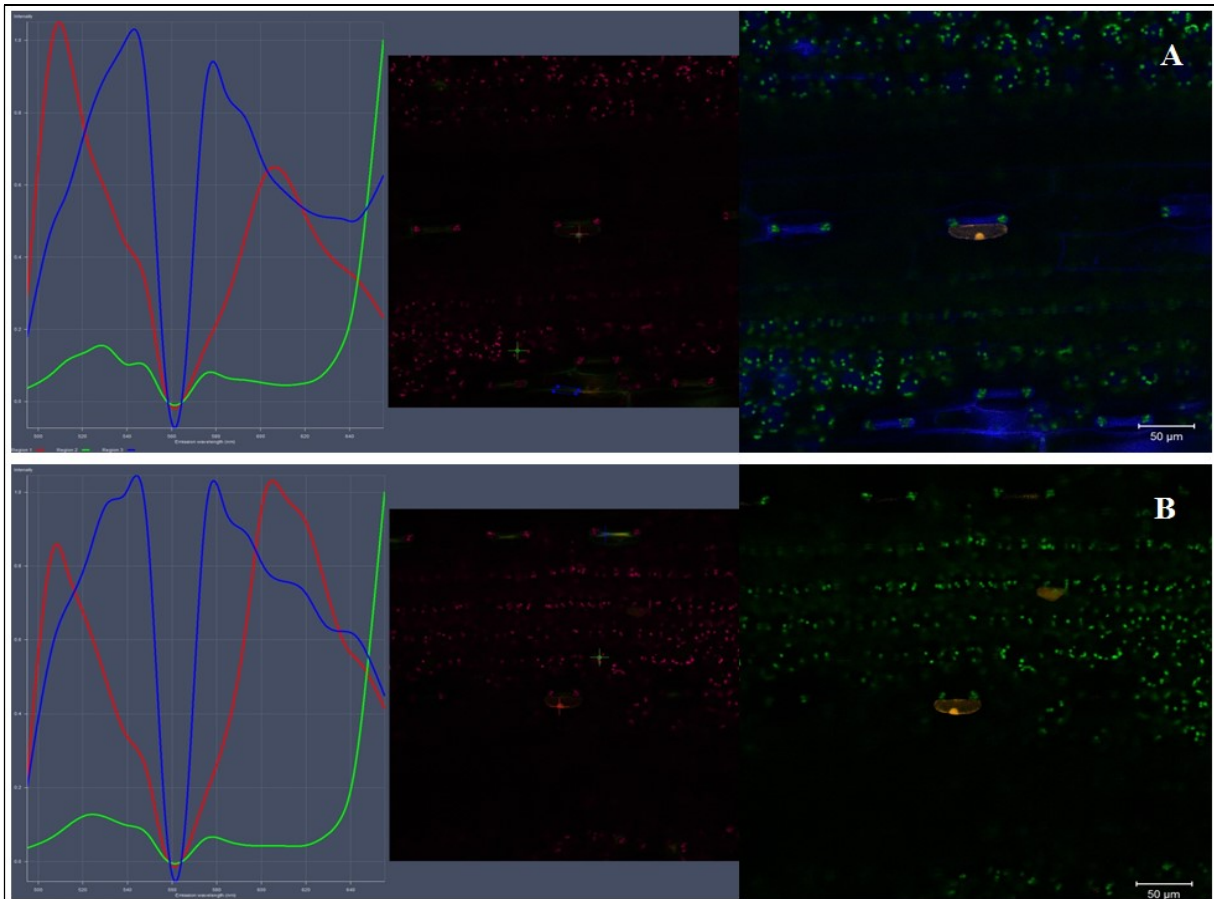


Abb. 25| CLSM-Aufnahmen von Gerstenzellen, die mit *pHT4* und den *mlo*-spezifischen TALEN-kodierenden Vektoren Co-transformiert wurden. Ausschlaggebend sind die Graphen der roten Messpunkte. In beiden Fällen sind zwei deutlich ausgeprägte Peaks zu sehen, deren Maxima in den Wellenlängenbereichen $\lambda = 509$ nm und $\lambda = 610$ nm liegt. Sie zeigen die deutlich bis sehr stark ausgeprägte Emissionsintensität von beiden Fluoreszenzreportern.

Tabelle 15: Ergebnisse der Expressionsanalysen am Licht- und CLS-Mikroskop mit *pHT4* und den *mlo*-spezifischen TALEN-Paar

Versuch	Anzahl <i>mCherry</i> Zellen	Anzahl <i>Egfp</i> Zellen	Verhältnis <i>Egfp/mCherry</i> [%]	Analyse
1	84	63	75,00	CLSM
2	45	30	66,67	CLSM
Durchschnitt	65 ± 20	47 ± 17	72,09	CLSM
3	58	37	63,79	LM
4	14	6	42,86	LM
5	72	50	69,44	LM
6	84	70	83,33	LM
7	40	23	57,50	LM
Durchschnitt	54 ± 25	37 ± 22	69,40	LM
Gesamt-Durchschnitt	57 ± 24	40 ± 21	70,28	CLSM + LM

Tabelle 16: Ergebnisse der Kontrollversuche mit *pHT4*

Kontrolle	Anzahl <i>mCherry</i> Zellen	Anzahl <i>Egfp</i> Zellen
1	98	0
2	87	0
3	165	0
Durchschnitt	117 ± 34	0

4 Diskussion

Die eingangs an diese Arbeit gestellte Zielsetzung zur Klonierung und Erprobung eines modularen, flexibel anpassbaren DNA-Testvektors für Funktionalitätsprüfungen und zur quantitativen Ermittlung der nukleolytischen Aktivitäten von verschiedenen Sequenz-spezifizierten Endonukleasen in pflanzlichen Geweben konnte erfolgreich umgesetzt werden. Neuartig ist dabei die Verwendung eines Fusionsproteins zur Ermittlung der Restriktionsaktivitäten, anstatt eines Co-Bombardments mit zwei auf separaten Plasmiden lokalisierten Reporter genen. Dadurch sollte sich ein geringerer statistischer Fehler bei der Auswertung ergeben, wodurch eine präzisere Beurteilung und Einschätzung der Funktionalität der SSE ermöglichen werden würde. In der vorliegenden Form kann *pHT1* als generischer Vektor für die Insertion von nahezu jeder gewünschten Ziel-DNA zum Testen von Sequenz-spezifizierten Endonukleasen mittels *Frame-shift*-Mechanismus genutzt werden. Bei der Insertion von entsprechenden DNA-Fragmenten zwischen den Fluoreszenzreporter genen ist lediglich darauf zu achten, dass die *Egfp*-Sequenz bereits um ein bp versetzt ist und daher kein funktionales Leserasters aufweist. Durch die Insertion einer weiteren zusätzlichen Base infolge der DNA-Zielsequenz kann das System leicht auf einen +2 bp-Versatz umgestellt werden. Die Funktionalität des genutzten *Frame-shift*-Mechanismus konnte anhand zweier Beispiele in Gerstenzellen demonstriert werden.

Die im Fall von *pHT3* ermittelte, vergleichsweise geringe Enzymaktivitätsrate von 35,55 % und die oftmals nur schwache Intensität der *Egfp*-Emissionen können als Indikatoren für eine stattfindende, aber anscheinend eingeschränkte Restriktion der Zielsequenz durch die *TALEN* angesehen werden. Als Grund dafür kommen verschiedene mögliche Ursachen in Frage, deren individuelle Einflussnahmen zum vorliegenden Gesamtergebnis beitragen können. Die Architektur der *TALEN*-Bindedomänen entspricht mit mehr als 14 Wiederholungsmodulen aus je 34 AS und einem halben *Repeat* von 20 AS pro Nukleaseeinheit dem damaligen Stand der Literatur und orientiert sich sehr stark an natürlich vorkommenden *TALE*-Proteinen (121, 36, 54, 138). Die guten Bindungseigenschaften der verwendeten *RVD*-Motive ist an einer Vielzahl von Arten belegt (36, 38). Allerdings ist die zugrunde liegende *Codon*-Nutzung der von *Cermak et al.* (2011) als „*Konsens-Repeat*“ eingesetzten *Tal1c*-Sequenz wahrscheinlich nicht optimal für eine Expression in Gerstenzellen. Ein Großteil der Resultate der Studie von *Cermak et al.* stammt von Versuchen mit Hefezellen. Die Autoren machen leider keine Angaben zur Mutagenesefrequenz der *TALEN* in *A. thaliana*, die als Vergleich für die hier erzielten Ergebnisse dienen könnten. Es

werden auch keine weiteren Aussagen bezüglich der genutzten *Codons* getroffen, eine Optimierung für Gerste liegt in jedem Fall nicht vor. Gleiches gilt für die *Codon*-Nutzung bei der FokIR-Domäne. Daher können artspezifische oder auch kontextabhängige Einflüsse, die in einer verminderten Bildung der *TALEN*-Proteine resultieren nicht vollständig ausgeschlossen werden. Auch die vor dem *atg* befindlichen Klonierungsnarben können sich auf die Expression auswirken, da der Transkriptionsstart stets möglichst nahe am Promotor lokalisiert sein und einen bestimmten Kontext aufweisen sollte (14). Die Architektur des *TALE*-Rückgrats entspricht ebenfalls zum Großteil dem Stand des Wissens. Die N- und C-Termini der Proteine sind nur leicht verkürzt, wodurch eine gesteigerte Aktivität der *TALEN* erreicht werden soll (38). Zu stark verkürzte Abschnitte können deren Effizienz im Gegenzug jedoch auch erheblich beeinträchtigen, was im vorliegenden Fall allerdings eher unwahrscheinlich ist (123). Die Verknüpfung zur FokIR-Domäne fällt mit nur drei AS deutlich zu kurz aus. Für gewöhnlich wird ein Minimum von fünf bis sieben AS empfohlen um die korrekte Faltung und strukturelle Integrität des Proteins zu gewährleisten (5). Unter Umständen wird dies durch den mit 46 AS vergleichsweise großen C-Terminus der *TALE*-Proteine kompensiert, ein Einfluss auf die Restriktion oder Bindung kann jedoch auch hier nicht vollkommen ausgeschlossen werden. Die vermutlich wesentliche Ursache für die geringe Restriktionsrate ist jedoch mit hoher Wahrscheinlichkeit die nicht vorhandene T₀-Base am 5'-Beginn der linksseitigen Bindungssequenz, über welche für gewöhnlich die initiale Kontaktaufnahme der *TALE*-Domäne dirigiert und die Bindung vermittelt wird (145, 2, 36). Es existieren *TALEN*-Varianten die andere Basen am 5'-Ende der Zielsequenz registrieren (206) und Veröffentlichungen die das Erfordernis einer T₀-Base auch grundsätzlich in Frage stellen (7). Die meisten Autoren empfehlen jedoch die Nutzung entsprechender Zielsequenzen und weisen auf die dringende Notwendigkeit der Base für eine effiziente Bindung der Proteine an die DNA hin (145, 2, 36). Infolge der fehlenden T₀-Base könnte das Interaktionsvermögen der *TALEN* also durchaus auch sehr stark eingeschränkt sein, sodass es nur zu sehr wenigen vereinzelt Bindungen oder zufälligen Wechselwirkungen mit der DNA kommt. Dies spiegelt sich direkt in der geringen Anzahl an *Egfp*-exprimierenden Zellen und deren vergleichsweise schwachen Emissionen bei Verwendung des Vektors *pHT3* wieder. Ein interessanter Versuch zur Überprüfung dieser Theorie wäre ein Vergleich der Restriktionsrate der *ScCenH3α*-spezifischen *TALEN* in Kombination mit einem Testvektor dessen Zielsequenz über beide T₀-Basen verfügt. Die Wahrscheinlichkeit das sterische Probleme innerhalb des Fusionsproteins als Ursache für die geringe Intensität der *Egfp*-Fluoreszenz in Frage kommen ist gering, aber nicht vollkommen

auszuschließen. *pHT4* nutzt das gleiche Reporterkonstrukt und zeigt dabei in einer Vielzahl von Zellen eine deutliche bis intensive Expression des Fusionsproteins. Den einzigen, möglicherweise einflussreichen Unterschied zu *pHT3* stellt die bp-Abfolge zwischen dem Beginn der *TALEN*-Zielsequenz und dem *Egfp*-Gen dar. Die jeweils resultierenden AS können das Potential für ungewollte strukturelle Wechselwirkungen besitzen, welche die Intensität der Emission beeinträchtigen (205). Zu guter letzt bleibt die Frage bezüglich einer möglichen Kontextabhängigkeit der *TALEN* offen. Es gibt zahlreiche Berichte und Belege über eine z.T. unzureichende Funktionalität von synthetischen *TALEN*-Konstrukten in Wirtsorganismen (5, 36). Zudem kann davon ausgegangen werden, dass negative experimentelle Ergebnisse in der Literatur auch stark unterrepräsentiert sind. Oftmals verhalten sich Enzyme zudem anders als vorhergesagt und zeigen nicht die gewünschten Aktivitäten. Eine robuste und universell einsetzbare *TALEN*-Architektur existiert bislang nicht. Daher sind wirts- oder auch zellspezifische, kontextbasierende Einflüsse stets unvorhersagbare Faktoren bei allen Erstanwendungen.

Mit den *mlo*-spezifischen *TALEN* konnten die besten Resultate innerhalb der Versuchsreihen erzielt werden. Bei den auf dem natürlichen Avirulenzfaktor Bs3 basierenden Endonukleasen handelt es sich um professionell gestaltete, kommerzielle *TALEN* der ersten Generation, die in ihrem Aufbau und ihrer Struktur dem damaligen Stand der Literatur entsprechen. Die Restriktionsrate von 70,3 % reflektiert die durchaus guten Bindungseigenschaften der je 15,5 Wiederholungsmodule bzw. genutzten *RVD*-Motive und belegt erstmals die Funktionalität der *TALEN* in Gerste. Die auch in diesem Fall verkürzten N- und C-Termini könnten sich dabei Vorteilhaft auf die Aktivität der *TALE*-Proteine auswirken (38). Die *Spacer*-Sequenz zwischen den beiden Bindungsabschnitten besitzt mit 15 bp einen optimalen Umfang für die Dimerisierung der FokIR-Homodimere (5,36). Im Vergleich zu *pHT3* verfügen die Verknüpfungsabschnitte zwischen den *TALE*- und Restriktionsdomänen mit sechs AS über eine ausreichende Größe (5). Dies steigert die Wahrscheinlichkeit für die korrekte Faltung der Proteine um den DNA-Strang, was die räumlichen Ausrichtung der Nukleasedomänen ggf. begünstigen und zu einer erleichterten Dimerisierung beitragen kann. Selbst solche eher geringfügige Faktoren können sich Vorteilhaft auf die Funktionalität der Endonukleasen auswirken, da neben der zentralen Bindedomäne vor allem die Rückgratarchitektur von *TALEN* im Allgemeinen eine wichtige Rolle bei der Wahrnehmung von Zielsequenzen spielt und deren Aktivität wesentlich mitbestimmt (38). Da dieses Paar, ebenso wie die *cenH3a*-spezifischen *TALEN*, bisher für keinerlei andere Experimente genutzt wurden, existieren keine direkt vergleichbaren

Ergebnisse. Trotz des durchaus guten Resultats von 70,3 % müsste man dennoch zunächst von einer zu erwartenden Egfp-Detektionsrate von 100 % ausgehen. Aufgrund der hohen Kopienanzahl an Plasmid-DNA-Molekülen pro Goldpartikel sollte entsprechend der Wahrscheinlichkeit in mehr oder weniger jeder transformierten Zelle auch eine Expression des Restriktionsindikators zu beobachten sein. Vorteilhafterweise ist dies nicht der Fall, da ein Vergleich von Rekombinationsfrequenzen mit dem vorliegendem Testsystem dann in den meisten Fällen nicht möglich wäre. Eine Ursache für die geringere Rate könnte bei der DNA-Reparatur zu finden sein. Trotz der Fehleranfälligkeit sind die Mechanismen des NHEJ recht effizient und beheben Strangbrüche in den meisten Fällen fehlerfrei (12). Durch die hohe Kopienzahl an Testvektoren in den transformierten Zellen liegen zudem genügend homologe Sequenzen vor, die als Matrix für eine fehlerfreie, HD-vermittelte DNA-Reparatur genutzt werden können. Diese treten zwar generell sehr seltenen auf, sind aber nicht auszuschließen (8). Eine hohe Kopienzahl fördert zudem die schnellere Zersetzung der Fremd-DNA (14). Sollte es zu einer Mutation infolge einer fehlerhaften DNA-Reparatur kommen, muss diese weiterhin in einem *Frame-shift* mit dem entsprechenden bp-Versatz zur Wiederherstellung des Restriktionsreportergens resultieren. Eine erste Analyse von Zielsequenzmustern aus Gerstenzellen deutet darauf hin, dass bei etwa einem Drittel aller auftretenden Mutation tatsächlich ein *Frame-shift* erfolgt (*Gurushidze et al.*, Manuskript in Vorbereitung). Ein von *Johnson et al.* entwickeltes Testsystem zur quantitativen Ermittlung von *in planta* Spaltungseffizienzen von Sequenz-spezifisierten Endonukleasen, welches ebenfalls auf der Korrektur einer Reportergensequenz beruht, konnte eine um das 6,3-fach gesteigerte Aktivität von *TALEN* mit heterodimeren FokIR-Domänen gegenüber *TALEN* mit homodimeren FokIR-Domänen ermittelt werden (207). Die Autoren vermuten das Zielsequenzen von *TALEN* mit einem hohen Vorkommen an invertierten Sequenzwiederholungen nicht effizient von als Homodimer wirkenden *TALE*-Nukleasedomänen gespalten werden. Ein Vergleich dieser Art wäre auch mit dem hier vorliegenden Testsystem möglich. Die Nutzung von als Heterodimerisierenden Nukleasedomänen könnte zu einer Steigerung der Rekombinationsfrequenzen beitragen. In einem in der Arbeitsgruppe parallel entwickelten und zu diesem hier ähnlichen Testverfahren wurde ein weiteres *TALEN*-Paar mit einer anderen Zielsequenz innerhalb von *mlo* untersucht (198). Dabei konnte eine durchschnittliche Restriktionsfrequenz von 42,0 % ermittelt werden. Auch hier zeigte sich, dass eine Restriktionsrate von 100 % nicht zwangsläufig erwartet werden kann. Dennoch ist es denkbar und wahrscheinlich, dass eine sehr effiziente Endonuklease eine 100 %-Rate erreichen wird. In diesem Zusammenhang

empfiehlt sich bspw. die Erprobung von RGEN-Konstrukten, welche sich zumeist durch hohe Restriktionsaktivitäten auszeichnen (36).

Komplett anders verhält es sich im Fall des Testvektors *pHT2* und den *gfp*-spezifischen *TALEN*. Die Funktionalität und gute Aktivität der Endonukleasen in Gerstenzellen wurde mehrfach belegt, auch in der hier verwendeten Sommergerstensorte *Golden Promise* (129, 198). Im transienten Testsystem von *Budhagatapalli et al.* konnte eine Restriktionsaktivität von durchschnittlich 30,7 % ermittelt werden. *Gurushidze et al.* erzielte mit den *TALEN* eine Mutationsfrequenz von rund 22 %. Das Paar basiert auf der gleichen Rückgrat- und Bindedomänenarchitektur wie die *mlo*-spezifischen *TALEN*, denen eine effiziente Restriktion *in planta* bescheinigt werden konnten. Im Hefe-Hybridisierungstest des Herstellers erzielten die *gfp*-spezifischen *TALEN* einen Restriktionswert von 0,9 und wurden als „*sehr gute Spalter*“ ausgewiesen (197). Die ebenfalls für „*sehr gut*“ befundenen *mlo*-spezifischen *TALEN* erhielten im Vergleich dazu einen Wert von 0,86 (201). Da eine Beeinträchtigung der *TALEN*-Expression, Einflüsse durch strukturelle Eigenheiten und kontextabhängige Wechselwirkungen als Ursachen für die nicht vorhandene Bildung des Restriktionsindikators ausgeschlossen werden können, ist der Fehler vorrangig beim Testvektor selbst zu suchen. Das Plasmid *pHT2* nutzt prinzipiell den gleichen Mechanismus wie die anderen Testvektoren, jedoch bestehen mehrere bedeutsame Unterschiede die sich auf die Bildung von mCherry auswirken können. Zwei wesentliche Aspekte sind die vertauschte Abfolge der beiden Reportergene und die DNA-Zielsequenz der *gfp*-spezifischen *TALEN*, die sich in diesem Fall nicht infolge sondern innerhalb der Sequenz des ersten Reporters befindet. Kommt es infolge einer DSB-Induktion im knapp hinter der Hälfte der Gensequenz liegende Zielabschnitt zu dem gewünschten *Frame-shift* befindet sich ein großer Teil der hinteren *Egfp*-Sequenz partiell außerhalb des exprimierten Leserasters. Die stattdessen gebildete, etwa 115 AS umfassende *non-sense*-Sequenz zwischen den Reportern könnte zu schwerwiegenden und weitreichenden sterischen Problemen führen, welche die strukturelle Integrität des Fusionsproteins kompromittieren. Der genutzte -2 bp *Frame-shift* sollte keinen großen Unterschied zum -1 bp *Frame-shift* der anderen Testvektoren ausmachen. Genaue Analysen zur Häufigkeit bestimmter Mutationen liegen bislang nicht vor, bestätigt ist jedoch das die auftretenden Mutationen alle Arten von Variationen umfassen (130)(*Gurushidze et al.*, Manuskript in Vorbereitung). Daher sollte zumindest in einigen Zellen eine geringe Expression detektierbar sein. Die Eignung einer mit dem Testvektor *pHT2* ermittelten Aktivitätsrate als Maß- und Richtgröße zur Beurteilung der anderen Endonukleasepaare wäre fraglich, da die strukturellen Unterschiede im Vergleich mit den anderen Testvektoren zu

Umfangreich sind. Eine Möglichkeit zur Überwindung der Probleme wäre eine Umgestaltung des Testvektors nach dem Vorbild von *pHT3* und *pHT4*. Dazu bedarf es eines weiteren Reportergens und lediglich der Zielsequenz der *gfp*-spezifischen *TALEN*.

Ein angestrebter direkter Vergleich der Intensitäten der Reporteremissionen zur quantitativen Ermittlung von Restriktionsaktivitäten anhand nur einer oder weniger Zellen war leider nicht möglich. Die Idee bestand darin das stets gleiche Mengenverhältnis der Reporter-DNA in einer Zelle zu nutzen und die Intensitäten der Fluoreszenz ins Verhältnis zu einander zu setzen. Damit hätte eine vermutlich recht präzise Aussage über den prozentualen Anteil an Plasmidmolekülen die das Fusionsprotein bilden getroffen werden können. Da die Anregung der Proben zur Messung der Egfp- und mCherry-Emissionen mit zwei verschiedenen Lasern in unterschiedlichen Wellenlängenbereichen und die Detektion ebenfalls durch verschiedene Passfilter erfolgt ist ein solcher direkter Vergleich mit dem zur Verfügung stehenden Mikroskopiesystem nicht möglich. Die Berechnung des Intensitätswerts erfolgt erst im Anschluss an die Messung anhand des erfassten Bildpunktes. Für einen Vergleich müsste in jeder Zelle ein interner Standard der beiden Reporter mitgeführt werden, was nicht möglich ist. Dies wurde auf Rückfrage vom Hersteller *Zeiss* (Jena) bestätigt.

Gemeinsam mit dem in der Arbeitsgruppe parallel entwickelten Ansatz (198) sind die durchgeführten Arbeiten ein erster Versuch zur Etablierung eines vielseitigen, modularen und leicht zu handhabenden *in vivo* Testsystems für Kulturpflanzen, welches sich für alle Plattformen von Sequenz-spezifisierbaren Endonukleasen eignen sollte. Mit dem generischen Testvektor *pHT1* konnte ein neuartiges und nützliches Werkzeug mit vielseitigen Anwendungsmöglichkeiten in der Grundlagen- und angewandten Forschung im Bereich des pflanzlichen *Genome Engineering* geschaffen werden, welches zum besseren Verständnis von Kontextabhängigkeiten oder Fragen bezüglich Endonukleasearchitekturen beitragen kann. Der Vektor kann als schneller und reproduzierbarer Indikator vor langfristigen Studien genutzt werden und damit evtl. helfen den Umfang von Mutationszüchtungen besser einschätzen zu können oder sogar zu reduzieren. Neben der Feststellung von *in planta* Restriktionsaktivitäten können bspw. Plasmide zur Ermittlung eines Interaktionspotentials von synthetischen Endonukleasen mit nichtziel-Sequenzen generiert werden, die dabei helfen könnten molekulare Zusammenhänge aufzuklären. Die Klonierung einer neuen Zielsequenz in den Basistestvektor ist in jedem Fall schneller, günstiger und einfacher als die Generierung eines neuen *TALEN*-Paars. Im nächsten Schritt sollte eine Überprüfung der Tauglichkeit im Zusammenhang mit RGEN erfolgen. Durch die Möglichkeit zum einfachen Austausch der

Promotorsequenz kann der Vektor zudem mit nur einem Klonierungsschritt für Versuche in weiteren Kulturpflanzen optimiert werden. Theoretisch wäre sogar ein Einsatz in Hochdurchsatzverfahren denkbar, bspw. mittels Durchfluss-zytometrisch analysierter Pollenkulturen. Die Autoren *Kim et al.* nutzen ein zu diesem hier nahezu identisches Testsystem für Versuche in humanen Leberzellen, die mittels Durchfluss-Zytometrie analysiert werden (208).

Literaturverzeichnis

1. Miedaner T. Fundamentals of plant breeding. Frankfurt am Main: DLG-Verlag; 2010. 261 pp. p.
2. Moose SP, Mumm RH. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant physiology*. 2008;147(3):969-77.
3. Schmidt R. Taschenatlas der Biotechnologie und Gentechnik: Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim; 2002.
4. Mba C. Induced mutations unleash the potentials of plant genetic resources for food and agriculture. *Agronomy*. 2013;3(1):200-31.
5. Hiekel S, Schedel S, Hensel G, Gurushidze M, Budhagatapalli N, Kumlehn J, editors. Synthetic Endonucleases: Novel Tools for the Site-Directed Genetic Modification of Plants. XXV International EUCARPIA Symposium Section Ornamentals: Crossing Borders 1087; 2015.
6. Puchta H, Dujon B, Hohn B. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Research*. 1993;21(22):5034-40.
7. Kim H, Kim J-S. A guide to genome engineering with programmable nucleases. *Nat Rev Genet*. 2014;15(5):321-34.
8. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*. 1994;14(12):8096-106.
9. Puchta H, Dujon B, Hohn B. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proceedings of the National Academy of Sciences*. 1996;93(10):5055-60.
10. Carroll D. Genome Engineering with Targetable Nucleases. *Annual Review of Biochemistry*. 2014;83(1):409-39.
11. Gorbunova V, Levy AA. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Research*. 1997;25(22):4650-7.
12. Puchta H. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*. 2005;56(409):1-14.
13. Knoll A, Fauser F, Puchta H. DNA recombination in somatic plant cells: mechanisms and evolutionary consequences. *Chromosome Res*. 2014;22(2):191-201.
14. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molekularbiologie der Zelle*: Wiley-VCH Verlag & Co. KGaA, Weinheim, Germany; 2011.
15. Britt AB, May GD. Re-engineering plant gene targeting. *Trends in Plant Science*. 2003;8(2):90-5.
16. Amiard S, Gallego ME, White CI. Signaling of double strand breaks and deprotected telomeres in *Arabidopsis*. *Frontiers in Plant Science*. 2013;4:405.
17. Salomon S, Puchta H. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *The EMBO Journal*. 1998;17(20):6086-95.
18. Vu GTH, Cao HX, Watanabe K, Hensel G, Blattner FR, Kumlehn J, et al. Repair of Site-Specific DNA Double-Strand Breaks in Barley Occurs via Diverse Pathways Primarily Involving the Sister Chromatid. *The Plant Cell*. 2014;26(5):2156-67.
19. Charbonnel C, Gallego ME, White CI. Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in *Arabidopsis* plants. *The Plant Journal*. 2010;64(2):280-90.

20. Tamura K, Adachi Y, Chiba K, Oguchi K, Takahashi H. Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: evidence for a role in the repair of DNA double-strand breaks. *The Plant Journal*. 2002;29(6):771-81.
21. West CE, Waterworth WM, Jiang Q, Bray CM. *Arabidopsis* DNA ligase IV is induced by γ -irradiation and interacts with an *Arabidopsis* homologue of the double strand break repair protein XRCC4. *The Plant Journal*. 2000;24(1):67-78.
22. Gorbunova V, Levy AA. How plants make ends meet: DNA double-strand break repair. *Trends in Plant Science*. 1999;4(7):263-9.
23. Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 2011;711(1-2):61-72.
24. Nicolás AL, Munz PL, Young CSH. A modified single-strand annealing model best explains the joining of DNA double-strand breaks in mammalian cells and cell extracts. *Nucleic Acids Research*. 1995;23(6):1036-43.
25. Pacher M, Schmidt-Puchta W, Puchta H. Two Unlinked Double-Strand Breaks Can Induce Reciprocal Exchanges in Plant Genomes via Homologous Recombination and Nonhomologous End Joining. *Genetics*. 2007;175(1):21-9.
26. Tzfira T, Frankman LR, Vaidya M, Citovsky V. Site-Specific Integration of *Agrobacterium tumefaciens* T-DNA via Double-Stranded Intermediates. *Plant Physiology*. 2003;133(3):1011-23.
27. Schuermann D, Molinier J, Fritsch O, Hohn B. The dual nature of homologous recombination in plants. *Trends in Genetics*. 2005;21(3):172-81.
28. Rinaldo A, Ayliffe M. Gene targeting and editing in crop plants: a new era of precision opportunities. *Mol Breeding*. 2015;35(1):1-15.
29. Lin FL, Sperle K, Sternberg N. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Molecular and Cellular Biology*. 1984;4(6):1020-34.
30. Puchta H, Hohn B. The mechanism of extrachromosomal homologous DNA recombination in plant cells. *Molec Gen Genet*. 1991;230(1-2):1-7.
31. Nassif N, Penney J, Pal S, Engels WR, Gloor GB. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Molecular and Cellular Biology*. 1994;14(3):1613-25.
32. Puchta H. Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. *The Plant Journal*. 1998;13(3):331-9.
33. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. *Cell*. 1983;33(1):25-35.
34. Daoudal-Cotterell S, Gallego ME, White CI. The plant Rad50-Mre11 protein complex. *FEBS Letters*. 2002;516(1-3):164-6.
35. Roth N, Klimesch J, Dukowic-Schulze S, Pacher M, Mannuss A, Puchta H. The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. *The Plant Journal*. 2012;72(5):781-90.
36. Kirik A, Salomon S, Puchta H. Species-specific double-strand break repair and genome evolution in plants. *The EMBO Journal*. 2000;19(20):5562-6.
37. Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, et al. Targeted Mutagenesis of Duplicated Genes in Soybean with Zinc-Finger Nucleases. *Plant Physiology*. 2011;156(2):466-73.

38. Voytas DF. Plant Genome Engineering with Sequence-Specific Nucleases. *Annual Review of Plant Biology*. 2013;64(1):327-50.
39. Li T, Liu B, Spalding MH, Weeks DP, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotech*. 2012;30(5):390-2.
40. Petolino J, Worden A, Curlee K, Connell J, Strange Moynahan T, Larsen C, et al. Zinc finger nuclease-mediated transgene deletion. *Plant Mol Biol*. 2010;73(6):617-28.
41. Antunes MS, Smith JJ, Jantz D, Medford JI. Targeted DNA excision in Arabidopsis by a re-engineered homing endonuclease. *BMC Biotechnology*. 2012;12:86-.
42. Qi Y, Zhang Y, Zhang F, Baller JA, Cleland SC, Ryu Y, et al. Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Research*. 2013;23(3):547-54.
43. Gaj T, Gersbach CA, Barbas Iii CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*. 2013;31(7):397-405.
44. Yang M, Djukanovic V, Stagg J, Lenderts B, Bidney D, Carl Falco S, et al. Targeted mutagenesis in the progeny of maize transgenic plants. *Plant Mol Biol*. 2009;70(6):669-79.
45. Puchta H, Hohn B. Breaking news: Plants mutate right on target. *Proceedings of the National Academy of Sciences*. 2010;107(26):11657-8.
46. Paques F, Duchateau P. Meganucleases and DNA Double-Strand Break-Induced Recombination: Perspectives for Gene Therapy. *Current Gene Therapy*. 2007;7(1):49-66.
47. Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research*. 2001;29(18):3757-74.
48. Prieto J, Redondo P, Padró D, Arnould S, Epinat J-C, Pâques F, et al. The C-terminal loop of the homing endonuclease I-CreI is essential for site recognition, DNA binding and cleavage. *Nucleic Acids Research*. 2007;35(10):3262-71.
49. Sprink T, Metje J, Hartung F. Plant genome editing by novel tools: TALEN and other sequence specific nucleases. *Current Opinion in Biotechnology*. 2015;32(0):47-53.
50. Smith J, Grizot S, Arnould S, Duclert A, Epinat J-C, Chames P, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Research*. 2006;34(22):e149.
51. Gao H, Smith J, Yang M, Jones S, Djukanovic V, Nicholson MG, et al. Heritable targeted mutagenesis in maize using a designed endonuclease. *The Plant Journal*. 2010;61(1):176-87.
52. Tzfira T, Weinthal D, Marton I, Zeevi V, Zuker A, Vainstein A. Genome modifications in plant cells by custom-made restriction enzymes. *Plant Biotechnology Journal*. 2012;10(4):373-89.
53. Baltes NJ, Voytas DF. Enabling plant synthetic biology through genome engineering. *Trends in Biotechnology*. 2015;33(2):120-31.
54. Fauser F, Roth N, Pacher M, Ilg G, Sánchez-Fernández R, Biesgen C, et al. In planta gene targeting. *Proceedings of the National Academy of Sciences*. 2012;109(19):7535-40.
55. Siebert R, Puchta H. Efficient Repair of Genomic Double-Strand Breaks by Homologous Recombination between Directly Repeated Sequences in the Plant Genome. *The Plant Cell*. 2002;14(5):1121-31.
56. D'Halluin K, Vanderstraeten C, Van Hulle J, Rosolowska J, Van Den Brande I, Pennewaert A, et al. Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnology Journal*. 2013;11(8):933-41.

57. D'Halluin K, Vanderstraeten C, Stals E, Cornelissen M, Ruiters R. Homologous recombination: a basis for targeted genome optimization in crop species such as maize. *Plant Biotechnology Journal*. 2008;6(1):93-102.
58. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences*. 1996;93(3):1156-60.
59. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotech*. 2005;23(8):967-73.
60. Hiroyuki S, Susumu K. New restriction endonucleases from *Flavobacterium okeanoikoites* (FokI) and *Micrococcus luteus* (MluI). *Gene*. 1981;16(1-3):73-8.
61. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. *Proceedings of the National Academy of Sciences*. 1992;89(10):4275-9.
62. Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. FokI dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences*. 1998;95(18):10570-5.
63. Wah DA, Bitinaite J, Schildkraut I, Aggarwal AK. Structure of FokI has implications for DNA cleavage. *Proceedings of the National Academy of Sciences*. 1998;95(18):10564-9.
64. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *The EMBO Journal*. 1985;4(6):1609-14.
65. Klug A, Rhodes D. 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends in biochemical sciences*. 1987;12:464-9.
66. Searles MA, Lu D, Klug A. The role of the central zinc fingers of transcription factor IIIA in binding to 5 S RNA1. *Journal of Molecular Biology*. 2000;301(1):47-60.
67. Lu D, Alexandra Searles M, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature*. 2003;426(6962):96-100.
68. Wolfe SA, Nekludova L, Pabo CO. DNA Recognition by Cys2His2 Zinc Finger Proteins. *Annual Review of Biophysics and Biomolecular Structure*. 2000;29(1):183-212.
69. Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2-His2 Zinc Finger Proteins. *Annual Review of Biochemistry*. 2001;70(1):313-40.
70. Mani M, Smith J, Kandavelou K, Berg JM, Chandrasegaran S. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochemical and Biophysical Research Communications*. 2005;334(4):1191-7.
71. Beerli RR, Barbas CF. Engineering polydactyl zinc-finger transcription factors. *Nat Biotech*. 2002;20(2):135-41.
72. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotech*. 2007;25(7):786-93.
73. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotech*. 2007;25(7):778-85.
74. Pruetz-Miller SM, Reading DW, Porter SN, Porteus MH. Attenuation of Zinc Finger Nuclease Toxicity by Small-Molecule Regulation of Protein Levels. *PLoS Genet*. 2009;5(2):e1000376.
75. Ramalingam S, Kandavelou K, Rajenderan R, Chandrasegaran S. Creating Designed Zinc-Finger Nucleases with Minimal Cytotoxicity. *Journal of Molecular Biology*. 2011;405(3):630-41.
76. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Meth*. 2011;8(1):74-9.

77. Wilson KA, McEwen AE, Pruett-Miller SM, Zhang J, Kildebeck EJ, Porteus MH. Expanding the Repertoire of Target Sites for Zinc Finger Nuclease-mediated Genome Modification. *Mol Ther Nucleic Acids*. 2013;2:e88.
78. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim Y-G, et al. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology*. 2001;21(1):289-97.
79. Osakabe K, Osakabe Y, Toki S. Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. *Proceedings of the National Academy of Sciences*. 2010;107(26):12034-9.
80. Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, et al. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. *Proceedings of the National Academy of Sciences*. 2010;107(26):12028-33.
81. Marton I, Zuker A, Shklarman E, Zeevi V, Tovkach A, Roffe S, et al. Nontransgenic Genome Modification in Plant Cells. *Plant Physiology*. 2010;154(3):1079-87.
82. Qi Y, Li X, Zhang Y, Starker CG, Baltus NJ, Zhang F, et al. Targeted Deletion and Inversion of Tandemly Arrayed Genes in Arabidopsis thaliana Using Zinc Finger Nucleases. *G3: Genes|Genomes|Genetics*. 2013.
83. De Pater S, Pinas JE, Hooykaas PJJ, van der Zaal BJ. ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium-mediated floral dip transformation. *Plant Biotechnology Journal*. 2013;11(4):510-5.
84. De Pater S, Neuteboom LW, Pinas JE, Hooykaas PJ, Van Der Zaal BJ. ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. *Plant biotechnology journal*. 2009;7(8):821-35.
85. Wright DA, Townsend JA, Winfrey RJ, Irwin PA, Rajagopal J, Lonosky PM, et al. High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *The Plant Journal*. 2005;44(4):693-705.
86. Cai C, Doyon Y, Ainley WM, Miller J, DeKolver R, Moehle E, et al. Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol Biol*. 2009;69(6):699-709.
87. Shukla VK, Doyon Y, Miller JC, DeKolver RC, Moehle EA, Worden SE, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*. 2009;459(7245):437-41.
88. Ainley WM, Sastry-Dent L, Welter ME, Murray MG, Zeitler B, Amora R, et al. Trait stacking via targeted genome editing. *Plant Biotechnology Journal*. 2013;11(9):1126-34.
89. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*. 2009;326(5959):1509-12.
90. Moscou MJ, Bogdanove AJ. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science*. 2009;326(5959):1501.
91. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*. 2010;186(2):757-61.
92. Bonas U, Stall R, Staskawicz B. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molec Gen Genet*. 1989;218(1):127-36.
93. White FF, Potnis N, Jones JB, Koebnik R. The type III effectors of *Xanthomonas*. *Molecular Plant Pathology*. 2009;10(6):749-66.
94. Alfano JR, Collmer A. The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *Journal of Bacteriology*. 1997;179(18):5655-62.

-
95. Büttner D, Bonas U. Getting across—bacterial type III effector proteins on their way to the plant cell. *2002;21(20):5313-22.*
 96. Saijo Y, Schulze-Lefert P. Manipulation of the Eukaryotic Transcriptional Machinery by Bacterial Pathogens. *Cell Host & Microbe.* 2008;4(2):96-9.
 97. Nino-Liu DO, Ronald PC, Bogdanove AJ. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Molecular Plant Pathology.* 2006;7(5):303-24.
 98. Boch J, Bonas U. *Xanthomonas AvrBs3* Family-Type III Effectors: Discovery and Function. *Annual Review of Phytopathology.* 2010;48(1):419-36.
 99. Kay S, Hahn S, Marois E, Hause G, Bonas U. A Bacterial Effector Acts as a Plant Transcription Factor and Induces a Cell Size Regulator. *Science.* 2007;318(5850):648-51.
 100. Römer P, Recht S, Strauß T, Elsaesser J, Schornack S, Boch J, et al. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist.* 2010;187(4):1048-57.
 101. Hopkins CM, White F, Choi S, Guo A, Leach J. Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *Mol Plant-Microbe Interact.* 1992;5(6):451-9.
 102. Schreiber T, Bonas U. Repeat 1 of TAL effectors affects target specificity for the base at position zero. *Nucleic Acids Research.* 2014.
 103. Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. *Science.* 2012;335(6069):716-9.
 104. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J-K, et al. Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science.* 2012;335(6069):720-3.
 105. Boch J. TALEs of genome targeting. *Nat Biotech.* 2011;29(2):135-6.
 106. Streubel J, Blucher C, Landgraf A, Boch J. TAL effector RVD specificities and efficiencies. *Nat Biotech.* 2012;30(7):593-5.
 107. Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Research.* 2012.
 108. Morbitzer R, Römer P, Boch J, Lahaye T. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences.* 2010;107(50):21617-22.
 109. Morbitzer R, Elsaesser J, Hausner J, Lahaye T. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic acids research.* 2011;39(13):5790-9.
 110. Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotech.* 2013;31(1):76-81.
 111. Kumlehn J, Gurushidze M, Hensel G, Stein N. *Genetic Engineering. Biotechnological Approaches to Barley Improvement: Springer; 2014. p. 393-407.*
 112. Engler C, Kandzia R, Marillonnet S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE.* 2008;3(11):e3647.
 113. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE.* 2009;4(5):e5553.

-
114. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research*. 2011.
115. Doyle EL, Boohar NJ, Standage DS, Voytas DF, Brendel VP, VanDyk JK, et al. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Research*. 2012;40(W1):W117-W22.
116. Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J. Computational Predictions Provide Insights into the Biology of TAL Effector Target Sites. *PLoS Comput Biol*. 2013;9(3):e1002962.
117. Grau J, Boch J, Posch S. TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics*. 2013;29(22):2931-2.
118. Valton J, Dupuy A, Daboussi F, Thomas S, Maréchal A, Macmaster R, et al. Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation. *Journal of Biological Chemistry*. 2012;287(46):38427-32.
119. Kumlehn J. Neue Werkzeuge der Pflanzenzüchtung im Fokus: Programmierbare Endonukleasen. *Neue Verfahren in der Pflanzenzüchtung - Nutzen und Herausforderungen*; 27.01.2015; Bern, Swiss: Forum for Genetic Research, Swiss Academy of Sciences (SCNAT); 2015. p. 10 - 2.
120. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Research*. 2011;39(1):359-72.
121. Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu J-K. 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*. 2011;108(6):2623-8.
122. Weeks DP, Spalding MH, Yang B. Use of designer nucleases for targeted gene and genome editing in plants. *Plant Biotechnology Journal*. 2015:n/a-n/a.
123. Christian M, Qi Y, Zhang Y, Voytas DF. Targeted Mutagenesis of *Arabidopsis thaliana* Using Engineered TAL Effector Nucleases. *G3: Genes|Genomes|Genetics*. 2013;3(10):1697-705.
124. Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, et al. Rapid and Efficient Gene Modification in Rice and *Brachypodium* Using TALENs. *Molecular Plant*. 2013;6(4):1365-8.
125. Zhang Y, Zhang F, Li X, Baller JA, Qi Y, Starker CG, et al. Transcription Activator-Like Effector Nucleases Enable Efficient Plant Genome Engineering. *Plant Physiology*. 2013;161(1):20-7.
126. Liang Z, Zhang K, Chen K, Gao C. Targeted Mutagenesis in *Zea mays* Using TALENs and the CRISPR/Cas System. *Journal of Genetics and Genomics*. 2014;41(2):63-8.
127. Char SN, Unger-Wallace E, Frame B, Briggs SA, Main M, Spalding MH, et al. Heritable site-specific mutagenesis using TALENs in maize. *Plant Biotechnology Journal*. 2015:n/a-n/a.
128. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotech*. 2014;32(9):947-51.
129. Gurushidze M, Hensel G, Hiekel S, Schedel S, Valkov V, Kumlehn J. True-Breeding Targeted Gene Knock-Out in Barley Using Designer TALE-Nuclease in Haploid Cells. *PLoS ONE*. 2014;9(3):e92046.
130. Wendt T, Holm P, Starker C, Christian M, Voytas D, Brinch-Pedersen H, et al. TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. *Plant Mol Biol*. 2013;83(3):279-85.
131. Budhagatapalli N, Rutten T, Gurushidze M, Kumlehn J, Hensel G. Targeted modification of gene function exploiting homology-directed repair of TALEN-mediated double strand breaks in barley. *bioRxiv*. 2015.

132. Haun W, Coffman A, Clasen BM, Demorest ZL, Lowy A, Ray E, et al. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnology Journal*. 2014;12(7):934-40.
133. Clasen BM, Stoddard TJ, Luo S, Demorest ZL, Li J, Cedrone F, et al. Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnology Journal*. 2015:n/a-n/a.
134. Nicolia A, Proux-Wéra E, Åhman I, Onkokesung N, Andersson M, Andreasson E, et al. Targeted gene mutation in tetraploid potato through transient TALEN expression in protoplasts. *Journal of Biotechnology*. 2015;204(0):17-24.
135. Lor VS, Starker CG, Voytas DF, Weiss D, Olszewski NE. Targeted Mutagenesis of the Tomato PROCERA Gene Using Transcription Activator-Like Effector Nucleases. *Plant Physiology*. 2014;166(3):1288-91.
136. Baker M. Gene-editing nucleases. *Nat Meth*. 2012;9(1):23-6.
137. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012;337(6096):816-21.
138. Jansen R, Embden JDAV, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular microbiology*. 2002;43(6):1565-75.
139. Makarova K, Grishin N, Shabalina S, Wolf Y, Koonin E. 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. *Biol Direct*. 2006;1(1):1-26.
140. Horvath P, Barrangou R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*. 2010;327(5962):167-70.
141. Sorek R, Lawrence CM, Wiedenheft B. CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea. *Annual Review of Biochemistry*. 2013;82(1):237-66.
142. Haft DH, Selengut J, Mongodin EF, Nelson KE. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. *PLoS Comput Biol*. 2005;1(6):e60.
143. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*. 2005;151(8):2551-61.
144. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009;155(3):733-40.
145. Mojica FM, Díez-Villaseñor Cs, García-Martínez J, Soria E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol*. 2005;60(2):174-82.
146. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*. 2007;315(5819):1709-12.
147. Gesner EM, Schellenberg MJ, Garside EL, George MM, MacMillan AM. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat Struct Mol Biol*. 2011;18(6):688-92.
148. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 2011;471(7340):602-7.
149. Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJJ, van der Oost J, et al. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature*. 2011;477(7365):486-9.
150. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuys RJH, Snijders APL, et al. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*. 2008;321(5891):960-4.
151. Garneau JE, Dupuis M-E, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010;468(7320):67-71.

-
152. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR–Cas systems. *Nat Rev Micro*. 2011;9(6):467-77.
153. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science*. 2014;343(6176).
154. Nishimasu H, Ran FA, Hsu Patrick D, Konermann S, Shehata Soraya I, Dohmae N, et al. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*. 2014;156(5):935-49.
155. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotech*. 2013;31(9):827-32.
156. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotech*. 2013;31(9):833-8.
157. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(25):10098-103.
158. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences*. 2011;108(25):10092-7.
159. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. 2014;513(7519):569-73.
160. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014;507(7490):62-7.
161. Puchta H, Fauser F. Gene targeting in plants: 25 years later. *Int J Dev Biol*. 2013;57:629-37.
162. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. 2009;6(5):343-5.
163. Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotech*. 2013;31(3):230-2.
164. Mao Y, Zhang H, Xu N, Zhang B, Gou F, Zhu J-K. Application of the CRISPR-Cas system for efficient genome engineering in plants. *Molecular plant*. 2013;6(6):2008-11.
165. Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, et al. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotech*. 2013;31(8):688-91.
166. Xing H-L, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, et al. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology*. 2014;14(1):327.
167. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotech*. 2014;32(3):279-84.
168. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotech*. 2013;31(9):839-43.
169. Shah SA, Erdmann S, Mojica FJM, Garrett RA. Protospacer recognition motifs. *RNA Biology*. 2013;10(5):891-9.
170. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécroivain A-L, Bzdrenga J, et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR–Cas systems. *Nucleic Acids Research*. 2014;42(4):2577-90.

-
171. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotech.* 2013;31(9):822-6.
172. Xie K, Zhang J, Yang Y. Genome-Wide Prediction of Highly Specific Guide RNA Spacers for CRISPR-Cas9-Mediated Genome Editing in Model Plants and Major Crops. *Molecular Plant.* 2014;7(5):923-6.
173. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Research.* 2014;42(W1):W401-W7.
174. Ran FA, Hsu Patrick D, Lin C-Y, Gootenberg Jonathan S, Konermann S, Trevino AE, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell.* 2013;154(6):1380-9.
175. Schiml S, Fauser F, Puchta H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* 2014;80(6):1139-50.
176. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *The Plant Journal.* 2014;79(2):348-59.
177. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotech.* 2014;32(6):577-82.
178. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotech.* 2014;32(6):569-76.
179. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature.* 2015;520(7546):186-91.
180. Wright AV, Sternberg SH, Taylor DW, Staahl BT, Bardales JA, Kornfeld JE, et al. Rational design of a split-Cas9 enzyme complex. *Proceedings of the National Academy of Sciences.* 2015;112(10):2984-9.
181. Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V. Editing plant genomes with CRISPR/Cas9. *Current Opinion in Biotechnology.* 2015;32:76-84.
182. Bortesi L, Fischer R. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances.* 2015;33(1):41-52.
183. Gao J, Wang G, Ma S, Xie X, Wu X, Zhang X, et al. CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. *Plant Mol Biol.* 2015;87(1-2):99-110.
184. Jiang W, Yang B, Weeks DP. Efficient CRISPR/Cas9-Mediated Gene Editing in Arabidopsis thaliana and Inheritance of Modified Genes in the T2 and T3 Generations. *PLoS ONE.* 2014;9(6):e99225.
185. Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang D-L, et al. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proceedings of the National Academy of Sciences.* 2014;111(12):4632-7.
186. Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, et al. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research.* 2013;23(10):1233-6.
187. Mülhardt C. *Der Experimentator: Molekularbiologie, Genomics: Spektrum Akademischer Verlag;* 2009.
188. Sanford JC, Klein TM, Wolf ED, Allen N. Delivery of substances into cells and tissues using a particle bombardment process. *Particulate Science and Technology.* 1987;5(1):27-37.
189. Hansen G, Wright MS. Recent advances in the transformation of plants. *Trends in Plant Science.* 1999;4(6):226-31.
190. Vainstein A, Marton I, Zuker A, Danziger M, Tzfira T. Permanent genome modifications in plant cells by transient viral vectors. *Trends in Biotechnology.* 2011;29(8):363-9.

-
191. Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF. DNA Replicons for Plant Genome Engineering. *The Plant Cell*. 2014;26(1):151-63.
 192. Sanford JC. Biolistic plant transformation. *Physiologia Plantarum*. 1990;79(1):206-9.
 193. Birch R, Franks T. Development and Optimisation of Microprojectile Systems for Plant Genetic Transformation. *Functional Plant Biology*. 1991;18(5):453-69.
 194. Christou P. Genetic transformation of crop plants using microprojectile bombardment. *The Plant Journal*. 1992;2(3):275-81.
 195. Bullock WO. XL-1 Blue : a high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *BioTechniques*. 1987;5:376-9.
 196. Stratagene. XL1-Blue Competent Cells.
 197. BioResearch C. gfp-Talen Delivery Report.
 198. Budhagatapalli N. A simple test for the cleavage activity of customized endonucleases in plants. *Plant Methods*. 2016.
 199. Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, et al. The Barley Mlo Gene: A Novel Control Element of Plant Pathogen Resistance. *Cell*. 1997;88(5):695-705.
 200. Elliott C, Zhou F, Spielmeier W, Panstruga R, Schulze-Lefert P. Functional Conservation of Wheat and Rice Mlo Orthologs in Defense Modulation to the Powdery Mildew Fungus. *Molecular Plant-Microbe Interactions*. 2002;15(10):1069-77.
 201. BioResearch C. mlo-Talen Delivery Report.
 202. Christensen A, Quail P. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research*. 1996;5(3):213-8.
 203. Bevan M, Barnes WM, Chilton M-D. Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucleic Acids Research*. 1983;11(2):369-85.
 204. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotech*. 2004;22(12):1567-72.
 205. Stotz M. Green Fluorescent Protein (GFP) and it's relatives. 2005.
 206. Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ. TAL effectors: highly adaptable phyto-bacterial virulence factors and readily engineered DNA-targeting proteins. *Trends in Cell Biology*. 2013;23(8):390-8.
 207. Johnson R, Gurevich V, Levy A. A rapid assay to quantify the cleavage efficiency of custom-designed nucleases in planta. *Plant Mol Biol*. 2013;82(3):207-21.
 208. Kim H, Um E, Cho S-R, Jung C, Kim H, Kim J-S. Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nature methods*. 2011;8(11):941-3.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind in allen Fällen unter Angabe der Quelle kenntlich gemacht.

Hannes Trautwein

Hamburg, den 08.11.2016