

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
03. Oktober 2024 (03.10.2024)



(10) Internationale Veröffentlichungsnummer
WO 2024/199650 A1

(51) Internationale Patentklassifikation:
C12N 15/10 (2006.01)

(21) Internationales Aktenzeichen: PCT/EP2023/058156

(22) Internationales Anmeldedatum:
29. März 2023 (29.03.2023)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(71) Anmelder: WACKER CHEMIE AG [DE/DE]; Gise-
la-Stein-Str. 1, 81671 München (DE).

(72) Erfinder: RICHTER, Hagen; c/o Wacker Chemie AG,
Hanns-Seidel-Platz 4, 81737 München (DE). BORN-
HOEVD, Carsten; c/o Wacker Chemie AG, Hanns-Sei-
del-Platz 4, 81737 München (DE). CRONET, Philip-
pe; c/o Wacker Chemie AG, Hanns-Seidel-Platz 4, 81737
München (DE). KUJAU, Marian; c/o Wacker Chemie
AG, Hanns-Seidel-Platz 4, 81737 München (DE). WICH,
Guenter; c/o Wacker Chemie AG, Hanns-Seidel-Platz 4,
81737 München (DE).

(74) Anwalt: GROSS, Christine et al.; Hanns-Seidel-Platz 4,
81737 München (DE).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW,
BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ,
DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH,
GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO,
JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR,
LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY,
MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST,
SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, WS, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,
CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD,
SL, ST, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY,
KG, KZ, RU, TJ, TM), europäisches (AL, AT, BE, BG, CH,
CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS,
IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO,
RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht (Artikel 21 Absatz 3)
- mit dem Sequenzprotokollteil der Beschreibung (Regel 5 Absatz 2 Buchstabe a)



WO 2024/199650 A1

(54) Title: CRISPR-CAS9-BASED LINEARIZATION OF PLASMID DNA TEMPLATES

(54) Bezeichnung: CRISPR-CAS9 BASIERTE LINEARISIERUNG VON PLASMID DNA-TEMPLATEN

(57) Abstract: The invention relates to a method for producing a linear double-stranded DNA molecule, wherein the linear DNA molecule has smooth ends, and wherein the linear DNA molecule has a poly T sequence at the 5' end of the matrix strand, using a leader DNA and a Cas nuclease.

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matri-zen-Strangs aufweist, unter Verwendung einer Führungs-DNA und einer Cas-Nuklease.

CRISPR-Cas9 basierte Linearisierung von Plasmid DNA-Templaten

Die Erfindung betrifft ein Verfahren zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweist, umfassend die Schritte:

- (1) Bereitstellen eines doppelsträngigen zirkulären DNA-Moleküls, umfassend auf dem Matrizen-Strang in der 3'- zu 5'-Richtung:
 - (i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,
 - (ii) eine Poly T-Sequenz,
 - (iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon darstellen, und
 - (iv) ein Protospacer Adjacent Motif (PAM),
- (2) Bereitstellen einer Führungs-RNA, umfassend eine Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii), und eine Region, die mit einer Cas-Nuklease interagiert,
- (3) Bereitstellen einer Cas-Nuklease vom Typ II,
- (4) In Kontakt bringen des doppelsträngigen zirkulären DNA-Moleküls mit der Führungs-RNA und der CAS-Nuklease.

25

Ribonukleinsäuren (RNAs) spielen in vielen zellulären Prozessen eine entscheidende Rolle. Im zentralen Dogma der Biologie vermitteln sie die Übersetzung der DNA in Proteine und spielen mit messenger RNAs (mRNAs) und transfer RNAs (tRNAs) die Schlüsselrolle bei der Transkription und Translation. Aber auch viele andere Prozesse, vor allem auf regulatorischer Ebene, also der Feinabstimmung von biochemischen Reaktionen innerhalb und außerhalb von Zellen, werden durch RNAs maßgeblich gesteuert. Daher ist es nicht verwunderlich, dass RNAs ein Ziel von vielen Forschungsarbeiten sind, um diese Prozesse besser zu verstehen.

35

Neben grundlegenden Arbeiten zum allgemeinen Verständnis von RNAs hat sich vor allem in den letzten Jahren gezeigt, wie nützlich RNAs auch für therapeutische Anwendungen sind, wie z.B. die Impfstoffe gegen SARS-Cov2. Für solche Anwendungsgebiete ist es wichtig, dass geeignete Verfahren zur Verfügung stehen, um ausreichende Mengen an mRNA herstellen zu können. Diese Verfahren bedienen sich natürlich vorkommender Moleküle, beispielsweise RNA-Polymerasen, welche RNAs produzieren. Die synthetische Herstellung von RNAs durch DNA abhängige RNA-Polymerasen ist ein Verfahren, welches für Forschungs- und Industrieanwendungen routinemäßig eingesetzt wird. Bei diesem Verfahren werden eine DNA-Matrize, eine RNA-Polymerase, z.B. eine T7 RNA-Polymerase, Nukleotide wie ATP, CTP, GTP, UTP sowie geeignete Pufferbedingungen benötigt. Dann synthetisiert die Polymerase in Abhängigkeit vom DNA-Templat eine entsprechende komplementäre RNA. Die hergestellten RNAs können unterschiedlicher Art sein: mRNAs, *self-amplifying* RNAs (saRNAs), lange, nicht kodierende RNAs (lncRNAs), *guide* RNAs (gRNAs), *short interfering* RNAs (siRNA), *micro* RNAs (miRNAs), zirkuläre RNAs (circRNAs) und andere. Vor allem bei kürzeren RNAs kann industriell auch eine chemische Synthese verwendet werden. Für längere RNAs von über 100 Nukleotiden, beispielsweise mRNAs, ist eine chemische Synthese allerdings nicht ökonomisch, weswegen längere RNAs meist mittels einer enzymatischen *in vitro*-Transkription hergestellt werden.

In der großtechnischen Produktion wird für eine *in vitro*-Transkription in der Regel Plasmid-DNA als DNA-Templat verwendet, da Plasmid-DNA, anders als durch PCR hergestellte DNA-Templates, mittels Fermentation in den notwendigen Mengen produziert werden kann. Die in sich geschlossene, zirkuläre Natur von Plasmiden führt bei einer RNA-Synthese zu einem theoretisch unendlichen langen RNA-Transkript, in der Praxis immer zumindest anteilig zu ungewollt heterogenen, verlängerten RNAs. Diese Verlängerungen wirken sich negativ auf die Effizienz

des Transkriptionsprozesses aus, wodurch die Herstellungskosten steigen. In der Literatur werden zwei mögliche Ansätze beschrieben, um die ungewollten verlängerten RNA-Transkripte zu vermeiden.

5

Eine Lösung ist die klassische Nutzung von Restriktionsendonukleasen vom Typ II. Diese Enzyme erkennen eine spezifische DNA-Sequenz und führen in der Regel innerhalb oder nahe dieser Sequenz zu einem gezielten Schnitt im Doppelstrang der DNA. Mit Hilfe eines solchen Schnittes in der Plasmid-DNA wird eine so genannte *in vitro Run-Off*-Transkription ermöglicht, bei welcher die RNA-Polymerase am Doppelstrangbruch „abfällt“ und erneut an einen Promoter binden und eine neue Syntheserunde starten kann. Für einen solchen Schnitt in der DNA-Matrize werden typischerweise klassische Restriktionsenzyme vom Typ II eingesetzt. Restriktionsenzyme der Untergruppe Typ IIP schneiden DNA an definierten Positionen innerhalb ihrer Erkennungssequenzen. Zu den Restriktionsenzymen vom Typ IIP gehören zum Beispiel *EcoRI* und *XbaI*. Alternativ können auch Enzyme der Untergruppe Typ IIS verwendet werden, die außerhalb ihrer Erkennungssequenzen schneiden. Mit beiden Restriktionsenzymklassen vom Typ II (Typ IIP und Typ IIS) lassen sich aus der zirkulären Plasmid-DNA lineare DNA-Fragmente erzeugen, die sich zur effizienteren RNA-Synthese einsetzen lassen. Zu beachten ist hierbei, dass das DNA-Templat keine weiteren Erkennungssequenzen für die jeweils verwendeten Restriktionsenzyme enthalten sollte. Somit ergibt sich eine Einschränkung bei der Templat-Sequenz bzw. ein aufwändigeres Design derselben. Das kann einzelne Basenpaarungen betreffen, aber auch ganze Codone des DNA-Templats oder Domänen der Zielproteine, beispielsweise Motive oder Strukturen. Zusätzlich prozessieren viele Typ II-Restriktionsenzyme, vor allem Enzyme des Typs IIS, die DNA mit einem Überhang, wobei so genannte klebrige Enden 5' oder 3' entstehen. Dabei ist besonders ein 3' Überhang ungünstig, da dieser durch die Promiskuität der RNA-

10
15
20
25
30
35

Polymerase zu längeren Transkripten führen kann als ursprünglich geplant. Um diese ungewollten Transkripte zu verhindern, müsste der Überhang im Anschluss an den Restriktionsschritt abgebaut werden. Damit ergeben sich in der Planung und der Erzeugung der DNA-Template deutlich mehr Schritte und es steigt der entsprechende Entwicklungsaufwand.

Die zweite Lösung zur Vermeidung von zu langen RNA-Transkripten ist die Verwendung von Rho-unabhängigen Terminationssignalen. Im Gegensatz zu einer Verwendung von Restriktionseendonukleasen wird das DNA-Templat hierbei nicht linearisiert. Rho-unabhängige Terminatoren sind üblicherweise sekundäre Strukturen innerhalb der synthetisierten Ziel-RNA und verhindern das Entstehen von zu langen RNA-Transkripten mittels Syntheseabbruch. Nach einem erfolgten Syntheseabbruch kann die Polymerase, ähnlich wie bei der *Run-Off*-Transkription, erneut einen Promoter binden und eine neue Syntheserunde starten. Dabei bleibt eine hohe Transkriptionsrate erhalten, es werden allerdings auch Sequenzbereiche angefügt, welche unerwünscht sind, z.B. die Terminator-Sequenz. Um diese unerwünschten Elemente zu entfernen, können neben RNA-schneidenden Enzymen, z.B. RNasen, auch katalytisch aktive Nukleinsäuren eingesetzt werden. Zu diesen zählen zum Beispiel DNAsen, aber auch sogenannte Ribozyme. Ribozyme sind strukturierte RNA-Moleküle, welche *in trans* (von außen kommend) oder *in cis* (von der Sequenz kommend) RNAs an definierten Stellen schneiden. Hierfür ist in der Regel eine bestimmte RNA-Struktur und ein zweiwertiges Metallion nötig, typischerweise Mg^{2+} . Dadurch erfolgt ein nucleophiler Angriff an der Zielsequenz, wodurch diese gespalten wird und im Falle einer RNA-Synthese überschüssige Elemente entfernt werden können. Da die T7 RNA-Polymerase ebenfalls auf Mg^{2+} Ionen angewiesen ist, wird die Ribozym-Aktivität innerhalb der Synthese-Reaktion einen Einfluss auf die Effizienz der *in vitro*-Transkription haben, da eine Konkurrenz um die Mg^{2+} Ionen besteht. Damit ist eine weitere

Optimierung der Reaktionsbedingungen erforderlich, was mit Effizienzeinbußen einhergehen kann. Beim Entwurf entsprechender Sequenzen mit einem *Cis*-Ribozym ist darauf zu achten, dass ein Ribozym gefolgt von mindestens einer Terminator-Sequenz in die Ziel-RNA eingebaut wird. Entsprechend wird der gesamte Prozess zur Herstellung von DNA-Templaten für die mRNA-Herstellung teurer und somit weniger wirtschaftlich.

Für die enzymatische Herstellung von funktionalen reifen mRNAs und saRNAs, welche in Säugetierzellen aktiv sein sollen, ist es zwingend notwendig, dass die RNA einen Poly A-Anhang aufweist. Dieser Poly A-Anhang erhöht die Stabilität der RNA und spielt auch während der Translation in der eukaryotischen Zielzelle eine entscheidende Rolle. Der Poly A-Anhang wird in der Zielzelle durch Poly A-bindende Proteine erkannt und gebunden. Durch die gebundenen Proteine kommt es zur Interaktion des 3' Endes der mRNA mit den Proteinen der Translationsmaschinerie auf dem 5' Ende der mRNA. Dieses komplexe Zusammenspiel erhöht neben der Stabilität der RNA auch die Translationseffizienz. Daraus ergibt sich eine Schlüsselrolle des Poly A-Anhangs innerhalb der reifen mRNA. Um eine möglichst natürliche mRNA bzw. saRNA darzustellen, sollte der Poly A-Anhang möglichst lang (> 100 A-Nukleotide) und idealerweise homogen sein (ohne andere Basen; C, G oder U). Man spricht dann von einem unverdeckten Poly A-Anhang am 3' Ende der RNA.

Zwei Arten zur Produktion von RNA-Molekülen mit einheitlichem Poly A-Anhang sind derzeit üblich. Zum einen kann nach der Synthese der Ziel-RNA-Sequenz der Poly A-Anhang über eine Poly A-Polymerase generiert werden. Im Falle dieses enzymatischen Schrittes wird die hergestellte RNA zunächst gereinigt, bevor durch Zugabe einer Poly A-Polymerase und ATP als Substrat ein Poly A-Anhang synthetisiert wird. Die Dauer der Reaktion bestimmt die durchschnittliche Länge der Poly A-Anhänge. Diese

enzymatische Synthese funktioniert in der Regel sehr gut, birgt aber auch einige Nachteile. So ist es mit dieser Technik nicht möglich, eine klar definierte Länge des Poly A-Anhangs zu erzeugen, und es entsteht ein Produkt mit inhomogenen RNA-Molekülen von normal-verteilter Länge. Für eine industrielle Nutzung der so hergestellten RNA-Moleküle sind daher weitere Prozessschritte zur Anreicherung bzw. zur Reinigung der RNAs nötig, wodurch die Ausbeuten sinken, und die Herstellungskosten steigen.

10

Demgegenüber steht eine Methode, bei der die Kodierung des Poly A-Anhangs direkt durch das DNA-Templat erfolgt, welches als Vorlage zur Herstellung der RNA dient. Für die Polyadenylierung wird hierfür beim Design der Templat-DNA darauf geachtet, dass die doppelsträngige DNA auf dem Matrizen-Strang das nötige Poly T-Motiv enthält. Der Vorteil dieses Verfahrens liegt darin, dass innerhalb eines Reaktionsschrittes ein wesentlich definierteres Produkt mit einem Poly A-Anhang mit einheitlicher Länge entsteht.

15

20

Wenn PCR-Produkte als Templat für die Synthese von RNA-Molekülen mit homogenen und unverdeckten Poly A-Anhängen genutzt werden, kann der Poly A-Anhang direkt in den für die PCR verwendeten Primern kodiert werden. Für die Herstellung eines DNA-Templats mittels PCR kann ein reverser Primer verwendet werden, der neben der komplementären Sequenz am 3' Ende eine Abfolge von bis zu 120 Adeninen oder auch mehr enthält. In der PCR wird dieser Anhang an die kodierende DNA angefügt und steht somit für die *in vitro*-Transkription zur Verfügung. Da sich PCR-Reaktionen technisch meist nur in kleinen Volumina bis zu wenigen hundert Mikrolitern durchführen lassen, würden für eine großtechnische Verarbeitung zur RNA mehrere hundert PCR-Reaktionen benötigt. Eine Hochskalierung des PCR-Ansatzes durch ein sogenanntes Scale-Out, bei dem mehrere Reaktionen parallel durchgeführt werden, ist aufgrund des hohen Arbeits- und Kostenaufwands im industriellen Maßstab aktuell nicht wirtschaftlich.

30

35

Wenn Plasmid-DNA als Templat für die *in vitro*-Transkription von RNA verwendet wird, so ist eine fermentative Produktion der Plasmide, eine Aufreinigung der pDNA und anschließend eine Linearisierung der pDNA notwendig. Nach der Linearisierung muss
5 eine erneute Aufreinigung der linearisierten pDNA erfolgen, um eine möglichst reine DNA-Matrize für die *in vitro*-Transkription bereitzustellen, wodurch die Komplexität und die Kosten des Verfahrens steigen. Wie bereits zuvor angemerkt liegt ein
10 Nachteil bei der Verwendung von Typ IIP Restriktionsenzymen darin, dass diese Enzyme keinen freien unverdeckten Poly A-Anhang ergeben, da die Erkennungssequenz von Typ IIP Restriktionsenzymen nahezu immer auch andere Nukleotide wie C, G, oder T enthält. Zwar könnte diesbezüglich die Nutzung von Typ IIS Enzymen Abhilfe schaffen, da diese außerhalb ihrer
15 Erkennungssequenz schneiden und so der gewünschte unverdeckte Poly A-Anhang geschaffen werden könnte. Allerdings führt die Verwendung von Typ IIS Enzymen üblicherweise zu 3' bzw. 5' Überhängen, die wie oben beschrieben, zunächst abgebaut werden müssten.

20

Neben der Verwendung von Restriktionsenzymen ist es auch möglich, Ribozyme zur Linearisierung der Templat-DNA zu verwenden. Auch diese können so entworfen werden, dass ein homogener und freier Poly A-Anhang entstehen kann. Ribozyme werden bevorzugt so
25 entworfen, dass sie direkt am Ende der Poly A-Sequenz schneiden. Der Nachteil der Ribozyme liegt darin, dass bei *trans* Ribozymen eine weitere RNA bzw. bei *cis* Ribozymen eine deutlich längere RNA hergestellt werden muss. Dieses Vorgehen erhöht den Produktionsaufwand und entsprechend die Kosten.

30

CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*) und CRISPR-assoziierte (Cas) Nukleasen wurden als Teile bakterieller Immunsysteme identifiziert. Da Cas-Nukleasen DNA-Elemente präzise an einer bestimmten Stelle
35 schneiden können, wurden CRISPR-Cas-Systeme als effiziente

Geneditierungswerkzeuge weiterentwickelt, hauptsächlich für Anwendungen im Bereich der Genomeditierung, und haben in diesem Bereich zu einer Revolution der Molekular- und Mikrobiologie geführt (Hille et al., Cell 172, Issue 6, 1239-1259, 2018; 5 Plagens et al. FEMS Microbiology Reviews, 39(3):442-63, 2015; Richter et al. Int J Mol Sci. 2013 Jul; 14(7): 14518-14531).

Wird eine bakterielle Zelle von einer fremden Nukleinsäure infiziert, beispielsweise Phagen-DNA, Transposons, pDNA oder 10 RNA, reagiert das Bakterium mit einer Immunantwort, um die fremde DNA oder RNA zu eliminieren. Das Besondere an der CRISPR-Cas-Immunität ist, dass es sich um eine adaptive Immunität handelt, was bedeutet, dass die Bakterienzelle diese Art von Immunantwort erlernt. Während einer Infektion wird ein kurzes 15 DNA-Fragment der fremden Nukleinsäure für das Immunsystem verwendet, wodurch später bei einer gleichen Art der Infektion eine schnellere Antwort induziert werden kann. Dieses DNA-Fragment wird innerhalb des genetischen Locus für das CRISPR-Cas System im CRISPR-array eingebaut und wird dort als *spacer* 20 bezeichnet. Diese *spacer*-Sequenz wird später Teil der Führungs-RNA (gRNA) bzw. CRISPR-RNA (crRNA) und trägt somit zur Sequenzspezifität bei. Im Rahmen dieser bakteriellen Immunabwehr wird die Cas-Nuklease durch die crRNA zu einer spezifischen Zielsequenz auf der fremden Nukleinsäure dirigiert. Für die 25 Erkennung der Sequenz ist zum einen die Komplementarität der crRNA mit der fremden Ziel-DNA nötig, und zum anderen eine kurze Signalsequenz in direkter Nähe der Ziel-DNA, die als *Protospacer Adjacent Motif* (PAM) bezeichnet wird. Die PAM-Sequenz besteht je nach CRISPR-Cas System aus einer unterschiedlichen Anzahl an 30 spezifischen Nukleotiden und ist die initiale Erkennungssequenz für die Cas-Nuklease auf der zu schneidenden DNA. Beispielsweise erkennt die Cas-Nuklease Cas9 von *Streptococcus pyogenes* (SpCas9) die PAM-Sequenz 5'-NGG-3'. Ohne die PAM-Sequenz in direkter Nähe zur Zielsequenz erfolgt kein Schnitt der fremden 35 DNA. Die fremde Zielsequenz wird nach der Erkennung durch den

Ribonukleoprotein-Komplex aus Cas Nuklease und crRNA spezifisch geschnitten und im Folgenden kann die fremde DNA durch weitere zelluläre Nukleasen abgebaut werden. Nur die passende Kombination aus PAM-Sequenz und Zielsequenz in entsprechender räumlicher Nähe führen zur Schneideaktivität der Cas-Nuklease. Da eine spezifische Kombination aus PAM und Zielsequenz in der Regel im Genom der Bakterien nicht vorhanden ist, wird eine Autoimmunreaktion, d.h. das Schneiden der eigenen DNA verhindert.

10

Es gibt unterschiedliche CRISPR-Cas-Systeme, wobei sich diese insbesondere in der Art und Anzahl der Proteine unterscheiden, die erforderlich ist, damit der CRISPR-Komplex an der richtigen Stelle der Zielsequenz binden und schneiden kann. Besonders gut verstanden sind die Systeme von Typ II-A, welche für die Interferenz mit der Ziel-DNA das Cas9-Protein als Effektor-Nuklease verwendet. Neben dem Namensgebenden CRISPR Array, welcher die verschiedenen crRNAs kodiert, umfassen die Typ II-Systeme eine trans-aktivierende RNA (tracrRNA) die Adaptionsproteine Cas1, Cas2 und Cas4, sowie die Effektornuklease Cas9. Die tracrRNA bildet ein Hybrid mit der vom CRISPR Array kodierten crRNA. Dieses Hybrid wiederum wird spezifisch von Cas9 gebunden. Der crRNA Anteil des Hybrid-Konstruktes leitet Cas9 zur Ziel-DNA. Das zweiteilige System aus crRNA und tracrRNA ist für eine Anwendung in der Molekularbiologie relativ umständlich, daher werden die crRNA und die tracrRNA oft mit einem Linker verbunden. Diese resultierende RNA nennt man sgRNA (single guide RNA).

30

Eine Aufgabe der Erfindung ist es, ein vereinfachtes bzw. verbessertes Verfahren zur Linearisierung zirkulärer DNA-Moleküle zur Verfügung zu stellen, welche als DNA-Template für die enzymatische RNA-Synthese verwendet werden können. Dazu soll das linearisierte DNA-Molekül über glatte Enden verfügen und eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweisen.

35

Insbesondere soll das Verfahren linearisierte DNA-Moleküle bereitstellen, welche die Synthese von RNAs mit einem homogenen und unverdeckten Poly A-Anhang ermöglichen. Eine andere Aufgabe der Erfindung besteht darin, ein Verfahren bereitzustellen, 5 welches weniger Verfahrensschritte benötigt als die im Stand der Technik offenbarten Verfahren. Eine andere Aufgabe besteht darin, ein Verfahren bereitzustellen, welches lineare DNA-Moleküle ergibt, die eine verbesserte Homogenität hinsichtlich der Länge des Poly-A-Anhangs aufweisen. Eine andere Aufgabe ist die 10 Bereitstellung eines Verfahrens, welches effizienter bzw. kostengünstiger durchgeführt werden kann als die im Stand der Technik beschriebenen Verfahren.

Die Aufgabe wird gelöst durch ein Verfahren zur Herstellung eines 15 linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweist, umfassend die Schritte:

- (1) Bereitstellen eines doppelsträngigen zirkulären DNA- 20 Moleküls, umfassend auf dem Matrizen-Strang in der 3'- zu 5-Richtung:
 - (i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,
 - (ii) eine Poly T-Sequenz,
 - 25 (iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon darstellen, und
 - (iv) ein Protospacer Adjacent Motif (PAM),
- (2) Bereitstellen einer Führungs-RNA, umfassend eine Region, die 30 komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii), und eine Region, die mit einer Cas-Nuklease interagiert,
- (3) Bereitstellen einer Cas-Nuklease vom Typ II,
- (4) In Kontakt bringen des doppelsträngigen zirkulären DNA- 35 Moleküls mit der Führungs-RNA und der CAS-Nuklease.

Eine Anwendung von CRISPR-Cas Systemen im Bereich der Linearisierung von Plasmid-DNA, insbesondere als Templat für eine *in vitro*-Transkription von RNA, ist bisher noch nicht bekannt. Durch eine Kombination einer Führungs-RNA, einer Cas-Nuklease, beispielsweise Cas9, und einem doppelsträngigen zirkulären DNA-Molekül kann ein CRISPR-Cas System für die effiziente und homogene Linearisierung von Plasmid-DNA eingesetzt werden.

10 Durch die flexible Gestaltungsmöglichkeit der Führungs-RNA kann das Verfahren zur Linearisierung flexibel für sämtliche doppelsträngige zirkuläre DNA-Moleküle verwendet werden, wobei ein spezifischer Schnitt an einer vorab ausgewählten Stelle der DNA möglich ist.

15

Durch die flexible Gestaltungsmöglichkeit des PAM kann das Verfahren zur Linearisierung flexibel mit unterschiedlichen Cas-Nukleasen durchgeführt werden.

20 Cas-Nukleasen des Typs II, beispielsweise Cas9, bilden beim Schnitt des Doppelstrangs im Zielmolekül sogenannte „glatte“ oder „stumpfe“ DNA-Enden, d.h. DNA-Enden ohne 3'- oder 5'-Überhang. Derartige DNA-Matrizen mit glatten Enden sind für die RNA-Synthese besonders gut geeignet. Bei der mRNA-Synthese ist es wünschenswert, dass die resultierende mRNA einen homogenen und unverdeckten Poly A-Anhang mit möglichst definierter Länge besitzt. Je homogener die Poly A-Anhänge der reifen mRNA sind, d.h. desto weniger andere Basen als Adenin enthalten sind, desto einfacher und effizienter ist die Bindung des Poly A-bindenden Proteins. Die Bindung dieses Proteins wiederum ist entscheidend für eine effektive Translation der reifen mRNA innerhalb der Zielzelle. Negative Effekte von mRNAs, die nicht direkt mit Adenin-Resten bzw. mit unterschiedlichen Basen enden, wie sie beispielsweise bei der Verwendung von vielen Typ II-Restriktionsenzymen entstehen, können so vermieden werden. Damit

25
30
35

ermöglicht das erfindungsgemäße Verfahren die Erzeugung von linearen DNA-Templaten, die zur Herstellung einer natürlichen, reifen RNA, beispielsweise einer mRNA, besonders gut geeignet sind, die bei therapeutischen Anwendungen Vorteile aufweist, z.B. eine verbesserte Stabilität, längere Halbwertszeit oder eine höhere Translationseffizienz.

Das erfindungsgemäße Verfahren bietet einen hohen Freiheitsgrad beim Sequenzdesign und eine absolute Sequenzunabhängigkeit, d.h. es gibt keine Einschränkungen bei Nukleotiden sowie Sequenz- und Strukturmotiven.

Das erfindungsgemäße, auf CRISPR-Cas basierende Verfahren ist unabhängig von Methylierungen und anderen Modifikationen der DNA, wie es bei vielen Restriktionsenzymen der Fall ist.

Erfindungsgemäß kann eine Linearisierung des doppelsträngigen zirkulären DNA-Moleküls, d.h. des DNA-Templats, schon während der Vervielfältigung der Plasmid-DNA, also beispielsweise innerhalb der bakteriellen Produktionszelle bzw. vor dem Zellaufschluss erfolgen. Dies bietet den Vorteil, dass die Anzahl der Prozessschritte reduziert werden kann. So kann am Ende der pDNA-Fermentation bereits in der Zelle die Linearisierung durch eine Cas-Nuklease, beispielsweise Cas9, erfolgen. Die lineare DNA kann nach Abschluss der Fermentation und intrazellulären Linearisierung mittels herkömmlicher Methoden aufgereinigt werden, eine separate Aufarbeitung des doppelsträngigen zirkulären DNA-Moleküls entfällt. Anschließend kann das lineare DNA-Templat direkt in einer *in vitro*-Transkription zur Produktion von RNA, beispielsweise mRNA mit einem unverdeckten Poly A-Anhang verwendet werden.

Das erfindungsgemäß hergestellte lineare doppelsträngige DNA-Molekül weist glatte Enden und eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs auf. Solche linearen DNA-Moleküle sind besonders

gut geeignet als DNA-Template für eine *in vitro*-Transkription. Bei der *in vitro*-Transkription handelt es sich um eine enzymatische Katalyse, bei welcher *in vitro* eine RNA hergestellt wird, dabei wird mit einer DNA-abhängigen RNA-Polymerase in
5 Abhängigkeit einer DNA-Matrize die dazu komplementäre RNA synthetisiert. Zur Synthese werden zudem Ribonukleotide (NTPs), ein Puffersystem, z.B. TRIS-HCl, Magnesium-Ionen, z.B. MgCl₂ und optional ein RNase-Inhibitor, eine anorganische Pyrophosphatase und Spermidin benötigt. Typischerweise werden monomere RNA-
10 Polymerasen verwendet, welche in der Regel Bakteriophagen-Enzyme sind. Die am häufigsten verwendeten Polymerasen sind zum Beispiel die T7- RNA-Polymerase (Uniprot: P00573) oder die T3-RNA-Polymerase (Uniprot: Q778M8).

15 Definitionen

Sofern es nicht anderweitig definiert ist, haben die hierin verwendeten technischen und wissenschaftlichen Fachbegriffe die gleiche Bedeutung, wie sie von einem Fachmann auf dem Gebiet der
20 vorliegenden Erfindung üblicherweise verstanden wird. Jedes technische Merkmal, das bei den folgenden Definitionen erwähnt wird, kann auf jede Ausführungsform der Erfindung angewendet werden.

25 Die Begriffe Nukleinsäuresequenz, Nukleotidsequenz, DNA-Sequenz und RNA-Sequenz sind einem Fachmann bekannt und beziehen sich auf eine individuelle und bestimmte Abfolge von Nukleotiden. Die Begriffe DNS und DNA sowie die Begriffe RNS und RNA werden im Rahmen dieser Erfindung synonym verwendet.

30

Die Begriffe Polynukleotid und Nukleinsäure werden hierin austauschbar verwendet und beziehen sich auf eine Abfolge von Nukleotiden beliebiger Länge, entweder Ribonukleotide oder Desoxyribonukleotide. Diese Begriffe umfassen einzelsträngige
35 und doppelsträngige DNA oder RNA, genomische DNA, cDNA, mRNA,

saRNA, gRNA, siRNA, miRNA oder circRNA, welche Purin- und Pyrimidin-Basen, Nukleotidanaloga oder auch andere natürlich, chemische oder biochemisch modifizierte, nicht-natürliche oder derivatisierte Nukleotid-Basen umfassen kann.

5

Ein Nukleotidanalogen bezieht sich auf eine chemische Verbindung, die strukturell und funktionell dem Nukleotid ähnlich ist, d.h. das Nukleotidanalogen kann von einer Polymerase als Substrat erkannt werden. Unter Nukleotidanaloga versteht man hierin insbesondere nativ und nicht-nativ vorkommende Varianten der natürlich vorkommenden Nukleotide Adenosin, Cytosin, Thymidin, Guanosin und Uridin, beispielsweise chemisch derivatisierte Nukleotide mit nicht-nativ vorkommenden funktionellen Gruppen, die an das natürlich vorkommende Nukleotid angefügt oder aus ihm entfernt werden, oder die die natürlich vorkommenden funktionellen Gruppen eines Nukleotids ersetzen. Beispiele für Nukleotidanaloga umfassen N1-Methylpseudouridin, 5-Methoxyuridin, 5-Methylcytidin, Pseudouridin, N4-Acetylcytidin und N6-Methyladenin.

20

Unter einer Poly T-Sequenz wie hierin verwendet ist eine Abfolge von Thymin-Nukleotiden zu verstehen, worin bevorzugt keine anderen Nukleotide enthalten sind, d.h. kein Adenin, Guanin oder Cytosin.

25

Unter einer Poly A-Sequenz wie hierin verwendet ist eine Abfolge von Adenin-Nukleotiden zu verstehen, worin bevorzugt keine anderen Nukleotide enthalten sind, d.h. kein Thymin, Guanin oder Cytosin.

30

Eine Poly T-Sequenz bzw. eine Poly A-Sequenz kann in zwei oder mehr unterschiedliche Segmente mit jeweils mindestens etwa 40 Thymin- bzw. Adenin-Nukleotiden segmentiert sein. Die einzelnen Segmente können jeweils von Sequenzen von etwa 10 - 20 Nukleotiden Länge unterbrochen sein, die keine Poly T bzw. Poly A-Sequenzen

35

sind, und auch G, C und U enthalten können. Eine solche Segmentierung kann eine verbesserte Stabilität der Poly T-Sequenz oder Poly A-Sequenz bewirken.

5 Unter einem unverdecktem Poly A-Anhang einer RNA ist hierin eine endständige Poly A-Sequenz zu verstehen, bei welcher auf die Abfolge von Adenin-Nukleotiden am 3' Ende der RNA keine anderen Nukleotide außer Adenin folgen, also kein Cytosin, Uracil oder Guanin.

10

Ein Poly A-Anhang einer RNA aus einer *in vitro*-Transkription wird hierin als homogen bezeichnet, wenn dieser neben dem unverdecktem 3' Ende eine enge Größenverteilung um eine definierte Ziellänge aufweist. Bevorzugt weist die durchschnittliche Abweichung zur Ziellänge nicht mehr als 10% bezogen auf die Anzahl der Nukleotide auf, mehr bevorzugt nicht mehr als 5% und besonders bevorzugt weniger als 5%. Ein segmentierter Poly A-Anhang gilt als homogen, wenn die Gesamtheit aller Segmente zusammen die oben genannte durchschnittliche Abweichung zur Ziellänge der Gesamtheit der Segmente aufweist. Eine Segmentierung kann eine Stabilisierung des gesamten Poly A-Anhangs bewirken. Auch bei einem segmentierten Poly A-Anhang einer RNA ist es von Vorteil, dass das 3' gelegene Segment des Poly A-Anhanges unverdeckt ist, um eine hohe Translationseffizienz zu gewährleisten.

25

Unter einem Matrizen-Strang einer doppelsträngigen DNA versteht man den DNA-Strang, welcher während einer Transkription von einer RNA-Polymerase abgelesen wird, um die RNA zu synthetisieren.

30 Unter einer Templat-Sequenz oder Templat-DNA bzw. einem DNA-Templat versteht man die DNA-Sequenz, die für eine *in vitro*-Transkription verwendet wird. Dabei dient diese Sequenz als Vorlage für eine DNA-abhängige RNA-Polymerase, um entsprechend dieser Sequenz eine RNA zu synthetisieren.

35

Eine Templat-Sequenz kann eine Nukleinsäure umfassen, welche für ein Antigen, beispielsweise eine Tumorantigen, ein virales oder bakterielles Antigen, ein therapeutisches Protein oder für andere Proteine wie z.B. Wachstumsfaktoren oder Transkriptionsfaktoren
5 kodiert.

Wie hierin verwendet, ist eine Promotor-Sequenz oder ein Promotor ein DNA-regulatorischer Bereich, der in der Lage ist, eine RNA-Polymerase zu binden und die Transkription einer
10 stromabwärts (in 3'-Richtung) gelegenen, kodierenden oder nicht-kodierenden Sequenz auszulösen. Geeignete Promotoren können von beliebigen Organismen abgeleitet werden, einschließlich prokaryotischen und eukaryotischen Organismen.

15 Wie hierin verwendet, bedeutet operativ mit einem Promotor verbunden, dass der Promotor die Transkription einer DNA bewirkt oder reguliert, die für ein Gen kodiert, wie beispielsweise das Gen, das durch die Templat-Sequenz kodiert wird, oder das Gen, das für die Führungs-RNA kodiert, oder das Gen, das für eine Cas-
20 Nuklease vom Typ II kodiert. Der Promotor kann ein nativer Promotor sein, d.h. ein Promoter, der in der Zelle, in die der Vektor eingeführt wird, vorliegt. In einer Ausführungsform ist der Promotor ein induzierbarer Promotor, d.h. der Promotor wird reguliert, um eine induzierbare Transkription eines Gens zu
25 ermöglichen, wie beispielsweise des Gens, das für die Führungs-RNA kodiert, oder des Gens, das für eine Cas-Nuklease vom Typ II kodiert.

Der Promotor kann ein konstitutiv aktiver Promotor sein, d.h. ein
30 Promotor, der sich konstitutiv in einem aktiven Zustand befindetet, oder es kann ein induzierbarer Promotor sein, d.h. ein Promotor, dessen Zustand durch einen externen Stimulus kontrolliert wird, bzw. vom inaktiven in einen aktiven Zustand gebracht wird. Durch einen Stimulus bzw. Induktor wird die Expression eines Zielgens
35 induziert, also freigeschaltet. Diese Methode wird bevorzugt

verwendet, wenn die Expression bzw. Überexpression eines Zielgens negative Auswirkungen auf den Produktionsorganismus oder auf ein Verfahren hat. Ein solcher externer Stimulus kann beispielsweise eine bestimmte Temperatur, eine Verbindung oder ein Protein sein.

5

Beispiele für induzierbare Promotoren umfassen den Lac-Promotor, der durch IPTG induzierbar ist (Jacob und Monod, Volume 3, Issue 3, Journal of Molecular Biology, June 1961, Pages 318-356, 1961), den Tetracyclin-regulierten Promotor (Gossen et al. 1995 *Science* 268 (5218): 1766-9.), den Rhamnose-induzierbaren Promotor (Kelly et al., ACS Synth. Biol.; 2016, 5 Pages 1136-1145), und den Arabinose-Promotor, der durch Arabinose induzierbar ist (Guzman et al., J Bacteriol. 1995, Jul;177(14):4121-30.).

15

Der Begriff *Protospacer Adjacent Motif* oder PAM beschreibt eine Signalsequenz aus etwa 2 bis 8 Basenpaaren, die die spezifische Erkennungsstelle für eine Cas-Nuklease oder für einen Multiprotein Cas-Effektor Komplex (z.B. CASCADE oder CMR) auf der zu schneidenden DNA darstellt. Ohne Vorliegen einer PAM-Signalsequenz erfolgt kein Schnitt der Ziel-DNA. Die Länge und die Basensequenz der PAM-Sequenz unterscheiden sich je nach CRISPR-Cas System und Cas-Nuklease sowie dem dazugehörigen Organismus, in dem die Cas-Nuklease natürlicherweise vorkommt. Beispielsweise ist die PAM-Sequenz, die mit der Cas9-Nuklease von *Streptococcus pyogenes* assoziiert ist (*SpCas9*), 5'-NGG-3', wobei „N“ ein beliebiges Nukleotid ist, also A, T, G oder C sein kann, gefolgt von zwei Guanin-Nukleotiden. Weitere Spezies und die damit assoziierten PAM-Sequenzen sind dem Fachmann bekannt und im Stand der Technik beschrieben (Fonfara et al. *Nucleic Acids Research*, 2014, Vol. 42, No. 4 2577-2590).

20

25

30

35

Eine Führungs-RNA, gRNA oder *guide RNA* ist eine spezifische RNA-Sequenz, an die eine Cas-Nuklease binden kann und die diese Cas-Nuklease zur Templat-DNA leitet. Die Führungs-RNA hybridisiert spezifisch an eine bestimmte Zielstelle der Templat-DNA, wo die

Cas-Nuklease die Templat-DNA schneidet. Die Führungs-RNA kann zweiteilig sein und ein Hybrid aus crRNA und tracrRNA darstellen. Die Führungs-RNA kann alternativ eine einzelne sgRNA sein (*single guide RNA*), worin die wesentlichen Bereiche der crRNA und tracrRNA miteinander verbunden sind, so dass eine durchgängige chimäre RNA entsteht. Eine sgRNA imitiert die Haarnadelstruktur des crRNA::tracrRNA-Duplex und ist ebenfalls in der Lage, eine Cas-Nuklease zu binden und zur Zielsequenz auf der Templat-DNA zu dirigieren. Die Führungs-RNA wird gestaltet anhand der Sequenzinformation der Templat-DNA, die geschnitten werden soll.

Komplementär oder Komplementarität, wie hierin verwendet, beschreibt die Beziehung zwischen zwei DNA- oder RNA-Nukleinsäureregionen, welche auf Grund ihrer Nukleobasen nach Watson-Crick zueinander passen und Wasserstoffbrücken untereinander eingehen, d.h. hybridisieren. Im Stand der Technik sind beispielsweise Standard-Watson-Crick-Basenpaarungen bekannt: Adenin (A) paart sich mit Thymin (T), Adenin (A) paart sich mit Uracil (U), und Guanin (G) paart sich mit Cytosin (C).

Unter homologen Genen bzw. homologen Sequenzen ist hierin zu verstehen, dass die DNA-Sequenzen dieser Gene bzw. DNA-Abschnitte zu mindestens 70%, bevorzugt zu mindestens 80%, bevorzugt zu mindestens 90% und besonders bevorzugt zu mindestens 95% identisch sind. Der Grad der DNA-Identität wird durch das Programm „nucleotide blast“, zu finden auf der Seite <http://blast.ncbi.nlm.nih.gov/>, bestimmt, welches auf dem blastn-Algorithmus basiert. Als Algorithmus-Parameter für ein Alignment zweier oder mehrerer Nukleotidsequenzen wurden die voreingestellten Parameter genutzt. Die voreingestellten generellen Parameter sind: Max target sequences = 100; Short queries = "Automatically adjust parameters for short input sequences"; Expect Threshold = 10; Word size = 28; Automatically adjust parameters for short input sequences = 0. Die entsprechenden voreingestellten Scoring Parameter sind: Match/Mismatch Scores = 1,-2; Gap Costs = Linear.

Unter einer homologen Aminosäuresequenz ist eine Sequenz zu verstehen, die zu mindestens 70%, bevorzugt mindestens 80%, bevorzugt mindestens 90% und besonders bevorzugt mindestens 95% identisch ist, wobei jede Änderung in der homologen Sequenz ausgewählt ist aus Insertion, Addition, Deletion und Substitution einer oder mehrerer Aminosäuren. Für den Vergleich von Proteinsequenzen wird das Programm „protein blast“, auf der Seite <http://blast.ncbi.nlm.nih.gov/>, genutzt. Dieses Programm greift auf den blastp-Algorithmus zurück. Als Algorithmus-Parameter für ein Alignment zweier oder mehrerer Proteinsequenzen wurden die voreingestellten Parameter genutzt. Die voreingestellten generellen Parameter sind: Max target sequences = 100; Short queries = "Automatically adjust parameters for short input sequences"; Expect Threshold = 10; Word size = 3; Automatically adjust parameters for short input sequences = 0. Die voreingestellten Scoring Parameter sind: Matrix = BLOSUM62; Gap Costs = Existence: 11 Extension: 1; Compositional adjustments = Conditional compositional score matrix adjustment.

20 Ausführliche Beschreibung der Erfindung

In einem Aspekt betrifft die vorliegende Erfindung ein Verfahren zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweist, umfassend die Schritte:

- (1) Bereitstellen eines doppelsträngigen zirkulären DNA-Moleküls, umfassend auf dem Matrizen-Strang in der 3'- zu 5'-Richtung:
 - 30 (i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,
 - (ii) eine Poly T-Sequenz,
 - (iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein
35 Nukleotidanalogen davon darstellen, und

(iv) ein Protospacer Adjacent Motif (PAM),

(2) Bereitstellen einer Führungs-RNA, umfassend eine Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii), und eine Region, die mit einer Cas-Nuklease interagiert,

(3) Bereitstellen einer Cas-Nuklease vom Typ II,

(4) In Kontakt bringen des doppelsträngigen zirkulären DNA-Moleküls mit der Führungs-RNA und der CAS-Nuklease.

10 Für das erfindungsgemäße Verfahren wird ein spezifisch gestaltetes, doppelsträngiges, zirkuläres DNA-Molekül bereitgestellt. Das zirkuläre DNA-Molekül umfasst auf dem Matrizen-Strang in der 3'- zu 5'-Richtung eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,
15 eine Poly T-Sequenz, gefolgt von einer Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon darstellen, und ein *Protospacer Adjacent Motif* (PAM).

20 Der RNA-Polymerase-Promotor ist bevorzugt eine Erkennungssequenz für eine DNA-abhängige RNA-Polymerase, beispielsweise der T7 RNA-Polymerase-Promotor, der SP6 RNA-Polymerase-Promotor oder der T3 RNA-Polymerase-Promotor.

25 Die Templat-Sequenz umfasst beispielsweise eine Nukleinsäure, welche für ein Antigen, beispielsweise ein Tumorantigen, ein virales oder ein bakterielles Antigen kodiert. In einem anderen Beispiel umfasst die Templat-Sequenz eine Nukleinsäure, welche für ein therapeutisches Protein oder für ein anderes Protein wie
30 z.B. einen Wachstumsfaktor oder Transkriptionsfaktor kodiert.

Die Poly T-Sequenz auf dem Matrizen-Strang des doppelsträngigen zirkulären DNA-Moleküls besteht beispielsweise aus zwischen etwa 40 bis etwa 250 T-Nukleotiden. Bevorzugt besteht die Poly T-

Sequenz aus zwischen etwa 100 und etwa 140 T-Nukleotiden, und mehr bevorzugt aus etwa 120 T-Nukleotiden.

5 In der Nukleotidsequenz, die dargestellt ist durch $N_1N_2N_3$, stellen N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon dar. Bevorzugt stellen N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C dar. In einer besonders bevorzugten Ausführungsform stellen N_1 bis N_3 jeweils unabhängig voneinander G, A, oder C dar.

10

Auf dem Matrizen-Strang des doppelsträngigen zirkulären DNA-Moleküls befindet sich zudem ein *Protospacer Adjacent Motif* (PAM). Bevorzugt folgt das PAM unmittelbar auf die Nukleotidsequenz $N_1N_2N_3$.

15

Das PAM kann ein beliebig ausgewähltes PAM sein, in Abhängigkeit von der spezifischen Cas-Nuklease, welche im erfindungsgemäßen Verfahren eingesetzt werden soll. Beispielsweise umfasst das PAM eine Sequenz, welche von einer Cas-Nuklease des Typs II erkannt wird. Beispielsweise ist das PAM ausgewählt aus NGG, NGGNG, NNAAAAW, NG, NNNNACA, NNNNGATT, GNNNCNNA und NNAGNN, worin N jeweils unabhängig ausgewählt ist aus G, C, A und T, und wobei die PAM-Sequenzen in 5'-3'-Richtung dargestellt sind. In einer bevorzugten Ausführungsform umfasst das PAM die Sequenz 5'-NGG-20 3', worin N unabhängig ausgewählt ist aus G, C, A und T. In einer anderen bevorzugten Ausführungsform weist das PAM die Sequenz 5'-NGG-3' auf, wobei N unabhängig ausgewählt ist aus G, C, A oder T. In einer weiteren bevorzugten Ausführungsform wird das PAM von einer Cas9-Nuklease aus *Streptococcus* erkannt.

30

Die Zielstelle, an der die zirkuläre doppelsträngige DNA geschnitten werden soll, liegt bevorzugt 3 Nukleotide 5' vom PAM.

Die Führungs-RNA für das erfindungsgemäße Verfahren ist so 35 gestaltet, dass sie am 5'-Ende eine Sequenz aufweist, welche

komplementär zum Matrizen-Strang des doppelsträngigen zirkulären DNA-Moleküls ist. Diese Komplementaritätsregion besteht beispielsweise aus etwa 20 Nukleotiden, die komplementär sind zu etwa 20 aufeinander folgenden Nukleotiden des Matrizen-Strangs.

5

Die Führungs-RNA umfasst eine Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii). Bevorzugt weist die Region der Führungs-RNA, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii) die Sequenz 3'-G(A)_n-N₄N₅N₆-5' auf. Dabei ist n eine Zahl zwischen 17 und 23, bevorzugt eine Zahl zwischen 17 und 20 und mehr bevorzugt ist n gleich 17.

10

N₄ bis N₆ stellen jeweils unabhängig G, U, A, C oder ein Nukleotidanalogen davon dar, bevorzugt stellen N₄ bis N₆ jeweils unabhängig G, U, C oder ein Nukleotidanalogen dar, besonders bevorzugt stellen N₄ bis N₆ jeweils unabhängig G, U oder C dar.

15

Die Komplementaritätssequenz innerhalb der Führungs-RNA weist beispielsweise die Sequenz 5'-GAAAAAAAAAAAAAAAAANN-3' auf, worin N jeweils unabhängig ein beliebiges Nukleotid ist. Bevorzugt ist das Verfahren dadurch gekennzeichnet, dass es sich bei der Führungs-RNA um SEQ ID NO:1 handelt.

20

Die Komplementarität der Region der Führungs-RNA zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii) beträgt mindestens 90%, beispielsweise 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, oder 100%. Bevorzugt beträgt die Komplementarität 98%, 99% oder 100%, besonders bevorzugt 100%.

25

30

Die Nukleotide N₄N₅N₆ der Führungs-RNA sind komplementär zu den Nukleotiden N₁N₂N₃ aus der Nukleotidsequenz (iii) aus dem Matrizen-Strang. Das bedeutet, dass N₄ komplementär ist zu N₁, N₅ komplementär ist zu N₂, und N₆ komplementär ist zu N₃. Bevorzugt beträgt die Komplementarität der Nukleotide N₄N₅N₆ und der

35

Nukleotide $N_1N_2N_3$ 100%. Bevorzugte komplementäre Basenpaarungen umfassen beispielsweise G und C, A und T bzw. A und U. Komplementäre Sequenzen sind beispielsweise CAG für $N_1N_2N_3$ und GUC für $N_4N_5N_6$.

5

Weiter umfasst die Führungs-RNA eine Region, die mit einer Cas-Nuklease interagiert. Bevorzugt umfasst die Führungs-RNA eine Region, die mit einer Cas-Nuklease vom Typ II interagiert.

10 Weiter wird in dem erfindungsgemäßen Verfahren eine Cas-Nuklease vom Typ II bereitgestellt. Es kann hierbei eine beliebige Cas-Nuklease vom Typ II eingesetzt werden. Beispielsweise wird eine Cas Nuklease vom Typ II aus *Streptococcus pyogenes* (SpCas9; Uniprot Q99ZW2), *Streptococcus mutans* (SmCas9; Uniprot Q8DTE3),
15 *Streptococcus thermophilus* (StCas9; Unitprot G3ECR1), *Staphylococcus aureus* (SaCas9; Uniprot J7RUA5), *Francisella novicida* (FnCas9; Uniprot A0Q5Y3), *Campylobacter jejuni* (CjCas9; Uniprot A0A698TVI8), *Neisseria meningitidis* (MnCas9; Uniprot C9X1G5), *Pasteurella multocida* (PmCas9; AddGene#68703;
20 <https://www.addgene.org/68703/>) bereitgestellt. Bevorzugt wird eine Cas-Nuklease vom Typ II aus *Streptococcus spec.* bereitgestellt. Mehr bevorzugt wird eine Cas-Nuklease vom Typ II aus *Streptococcus pyogenes* bereitgestellt. Am meisten bevorzugt wird als Cas-Nuklease vom Typ II Cas9 aus *Streptococcus pyogenes*
25 eingesetzt.

Bevorzugt wird als Cas-Nuklease vom Typ II Cas9 eingesetzt. Beispielsweise wird Cas9 ausgewählt aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen
30 Sequenz, bevorzugt mindestens 80% homologen Sequenz, mehr bevorzugt mindestens 90% homologen Sequenz, mehr bevorzugt mindestens 95% homologen Sequenz und besonders bevorzugt mindestens 98% homologen Sequenz; *Staphylococcus aureus* mit der SEQ ID NO:3 oder einer dazu zu mindestens 70% homologen Sequenz,
35 bevorzugt mindestens 80% homologen Sequenz, mehr bevorzugt

mindestens 90% homologen Sequenz, mehr bevorzugt mindestens 95% homologen Sequenz und besonders bevorzugt mindestens 98% homologen Sequenz, oder *Streptococcus thermophilus* mit der SEQ ID NO:4 oder einer dazu zu mindestens 70% homologen Sequenz, 5 bevorzugt mindestens 80% homologen Sequenz, mehr bevorzugt mindestens 90% homologen Sequenz, mehr bevorzugt mindestens 95% homologen Sequenz und besonders bevorzugt mindestens 98% homologen Sequenz.

10 In einer bevorzugten Ausführungsform handelt es sich bei der Cas-Nuklease vom Typ II um Cas9 aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen Sequenz, bevorzugt mindestens 80% homologen Sequenz, mehr bevorzugt mindestens 90% homologen Sequenz, mehr bevorzugt mindestens 95% 15 homologen Sequenz und besonders bevorzugt mindestens 98% homologen Sequenz.

Wenn im erfindungsgemäßen Verfahren als Cas-Nuklease Cas9 aus *S. pyogenes* eingesetzt wird, dann wird als PAM bevorzugt 5'-NGG-3' 20 verwendet. Wenn als Cas-Nuklease Cas9 aus *S. mutans* eingesetzt wird, dann wird als PAM bevorzugt 5'-NGG-3' verwendet. Wenn als Cas-Nuklease Cas9 eine Variante aus *S. thermophilus* eingesetzt wird, dann wird als PAM bevorzugt 5'-NGGNG-3' oder 5'-NNAAA-3' verwendet. Wenn als Cas-Nuklease Cas9 aus *Francisella novicida* 25 eingesetzt wird, dann wird als PAM bevorzugt 5'-NG-3' verwendet. Wenn als Cas-Nuklease Cas9 aus *Campylobacter jejuni* eingesetzt wird, dann wird als PAM bevorzugt 5'-NNNNACA-3' verwendet. Wenn als Cas-Nuklease Cas9 aus *Neisseria meningitidis* eingesetzt wird, dann wird als PAM bevorzugt 5'-NNNGATT-3' verwendet. Wenn als 30 Cas-Nuklease Cas9 aus *Pasteurella multocida* eingesetzt wird, dann wird als PAM bevorzugt 5'-GNNNCNNA-3', verwendet, wobei die vorliegende Erfindung nicht auf diese Beispiele beschränkt sein soll.

In einer weiteren bevorzugten Ausführungsform betrifft die vorliegende Erfindung ein Verfahren zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweist, umfassend die Schritte:

(1) Bereitstellen eines doppelsträngigen zirkulären DNA-Moleküls, umfassend auf dem Matrizen-Strang in der 3'- zu 5-Richtung:

(i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,

(ii) eine Poly T-Sequenz, worin die Poly T-Sequenz aus etwa 120 T-Nukleotiden besteht,

(iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon darstellen, und

(iv) ein PAM, welches die Sequenz 5'-NGG-3' aufweist, wobei N jeweils unabhängig ausgewählt ist aus G, C, A und T davon,

(2) Bereitstellen einer Führungs-RNA, umfassend eine Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii), und eine Region, die mit einer Cas-Nuklease interagiert, worin die Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii) die Sequenz 3'-G(A)_nN₄N₅N₆-5' aufweist, worin n gleich 17 ist, worin

N₄ bis N₆ jeweils unabhängig G, U oder C darstellt,

(3) Bereitstellen einer Cas-Nuklease vom Typ II, worin die Cas-Nuklease vom Typ II Cas9 aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen Sequenz ist,

(4) In Kontakt bringen des doppelsträngigen zirkulären DNA-Moleküls mit der Führungs-RNA und der CAS-Nuklease.

Die Führungs-RNA kann im erfindungsgemäßen Verfahren als RNA-Molekül bereitgestellt werden. Das Bereitstellen der Führungs-RNA kann erfindungsgemäß auch durch die Expression einer Nukleotidsequenz erfolgen, welche die Führungs-RNA kodiert. Bevorzugt wird die Führungs-RNA durch die Expression einer

Nukleotidsequenz bereitgestellt, welche die Führungs-RNA kodiert.

Die Cas-Nuklease vom Typ II kann im erfindungsgemäßen Verfahren als Protein bereitgestellt werden. Das Bereitstellen der Cas-Nuklease vom Typ II kann erfindungsgemäß auch durch die Expression einer Nukleotidsequenz erfolgen, welche für die Cas-Nuklease vom Typ II kodiert. Bevorzugt wird die Cas-Nuklease vom Typ II durch die Expression einer Nukleotidsequenz bereitgestellt, welche die Cas-Nuklease vom Typ II kodiert. Besonders bevorzugt werden sowohl die Führungs-RNA als auch die Cas-Nuklease durch die Expression einer Nukleotidsequenz bereitgestellt. Die Nukleinsäure, welche die Führungs-RNA und/oder die Cas-Nuklease kodiert, kann auf einem Expressionsplasmid lokalisiert sein. Beispielsweise kann die Nukleinsäure, welche die Führungs-RNA kodiert, auf einem Expressionsplasmid lokalisiert sein, und die Nukleinsäure, welche die Cas-Nuklease kodiert, kann auf einem separaten Expressionsplasmid lokalisiert sein. In einem anderen Beispiel sind die Nukleinsäure, welche die Führungs-RNA kodiert und die Nukleinsäure, welche die Cas-Nuklease kodiert, auf demselben Expressionsplasmid lokalisiert.

Beispielsweise kann ein Expressionsplasmid mit Cas9 und Führungs-RNA wie folgt generiert werden. Bei der Klonierung kann zunächst eine Cas9-Nuklease in ein geeignetes Expressionsplasmid kloniert werden, und anschließend kann beispielsweise in 5' - 3' Richtung ein tracrRNA Konstrukt und ein minimaler CRISPR Array kloniert werden. Dies ermöglicht beispielsweise eine natürliche Formierung einer tracrRNA::crRNA Duplex einer Führungs-RNA. Ein geeigneter Expressionsvektor ist beispielsweise pWAC003 (SEQ ID NO:5). Dabei handelt es sich um ein *low-copy* Plasmid mit ca. 10 Kopien pro Zelle auf Basis des pACYC-Replikationsursprungs.

Erfindungsgemäß erfolgt ein in Kontakt bringen des doppelsträngigen zirkulären DNA-Moleküls mit der Führungs-RNA und der Cas-Nuklease. Das in Kontakt bringen kann beispielsweise *in vitro* erfolgen. In einer Ausführungsform, worin das Bereitstellen
5 der Führungs-RNA und/oder der Cas-Nuklease vom Typ II durch die Expression einer Nukleotidsequenz erfolgt, welche die Führungs-RNA und/oder die Cas-Nuklease vom Typ II kodiert, erfolgt das in Kontakt bringen bevorzugt in einer Zelle, insbesondere in einer Bakterienzelle, welche geeignet ist, Plasmid-DNA zu vermehren.

10

Eine konstitutive Expression der Cas-Endonuklease, beispielsweise Cas9, kann unspezifische DNA-Schnittereignisse an verschiedenen Stellen im Plasmid-Produkt, d.h. im doppelsträngigen zirkulären DNA-Molekül, oder im Chromosom der
15 verwendeten Bakterienzelle verursachen. Ein frühes Prozessieren des doppelsträngigen zirkulären DNA-Moleküls durch eine Cas-Nuklease, beispielsweise Cas9, in den Zellen kann so zum Abbau des gewünschten linearen DNA-Produkts führen, da die freien Enden einer linearen DNA von entsprechenden DNasen erkannt werden.
20 Zudem kann auch die Zelle als Produzent geschädigt werden. Um diesen Effekt zu verringern ist es bevorzugt, für die Expression der Cas-Nuklease vom Typ II, beispielsweise Cas9, einen regulierbaren Promotor zu verwenden, der eine möglichst geringe Basalexpression aufweist.

25

Für die induzierbare Expression der Cas-Nuklease, beispielsweise Cas9, ist zusätzlich dazu eine mögliche Begrenzung der Promoter-Derepression, d.h. der induzierbaren Expressionshöhe bevorzugt. Eine zu hohe Expression der Cas-Nuklease, beispielsweise Cas9,
30 in einer kurzen Zeitspanne, kann ähnlich negative Effekte auf die pDNA-Qualität und die Produzentenzelle haben, wie eine hohe Basalexpression über eine längere Zeitspanne. Daher sind solche Promotoren besonders bevorzugt, die eine geringe Basalexpression und nach Induktion eine moderate Expressionsstärke aufweisen.

35

Eine induzierbare Expression der Führungs-RNA hat gegenüber einer konstitutiven Expression den Vorteil, dass die Expression keine dauerhafte Belastung für die Produzentenzelle darstellt. Durch den geringeren Stress kann die Zelle mehr vom eigentlichen Zielprodukt, der zirkulären DNA herstellen.

Durch die Verwendung eines regulierbaren bzw. induzierbaren Promotors kann der Zeitpunkt der Produktion der Cas-Nuklease und der Führungs-RNA gezielt durch die Induktion der Expression gewählt werden. Durch eine regulierbare bzw. induzierbare Expression des Prozessierungsapparates, d.h. der Cas9-Nuklease und der Führungs-RNA, innerhalb der Zelle, können im Folgeprozess einzelne Schritte eingespart werden, was zu wirtschaftlichen Vorteilen führt. Beispielsweise können zusätzliche Linearisierungsschritte oder eine separate Aufreinigung der Plasmid-DNA entfallen.

Bevorzugt ist das erfindungsgemäße Verfahren dadurch gekennzeichnet, dass die Expression der Cas-Nuklease vom Typ II und/oder der Führungs-RNA induzierbar ist. Das bedeutet, dass die Nukleotidsequenz kodierend für die Führungs-RNA und/oder die Nukleotidsequenz kodierend für die Cas-Nuklease operativ mit einem induzierbaren Promotor verbunden ist. Bevorzugt ist sowohl die Nukleotidsequenz kodierend für die Führungs-RNA als auch die Nukleotidsequenz kodierend für die Cas-Nuklease operativ mit einem induzierbaren Promotor verbunden.

In einer Ausführungsform ist die Nukleotidsequenz kodierend für die Führungs-RNA und die Nukleotidsequenz kodierend für die Cas-Nuklease operativ mit dem gleichen induzierbaren Promotor verbunden. Bei Verwendung eines gleichen oder ähnlichen Induktionssystems beginnt die Expression der Führungs-RNA und der Cas-Nuklease gleichzeitig. Zum einen erlaubt dieses Vorgehen die Senkung des zellulären Stresses, andererseits wird durch den zeitlich abgestimmten Beginn der Expression, ein passenderes

Verhältnis der Moleküle Cas-Endonuklease vs. Führungs-RNA ermöglicht.

In einer anderen Ausführungsform ist die Nukleotidsequenz
5 kodierend für die Führungs-RNA und die Nukleotidsequenz kodierend
für die Cas-Nuklease operativ mit unterschiedlichen induzierbaren
Promotoren, z.B. für sequentielle Expressionen, verbunden.

Induzierbare Promotoren sind einem Fachmann bekannt. Als
10 induzierbare Promotoren eignen sich beispielsweise der lac-,
tac-, trc-, lambda PL, ara-, cumat- oder tet-Promotor oder davon
abgeleitete Sequenzen.

Bevorzugt wird erfindungsgemäß als induzierbarer Promotor ein
15 durch Lactose-induzierbarer Promotor verwendet, insbesondere der
Promotor WAC003 mit der Sequenz SEQ ID NO:6 oder der Promotor
WAC003-LacIDel mit der Sequenz SEQ ID NO:7.

In einer Ausführungsform ist die Nukleotidsequenz kodierend für
20 die Führungs-RNA und/oder die Nukleotidsequenz kodierend für die
Cas-Nuklease in das Genom von *E. coli* integriert. In einer
bevorzugten Ausführungsform ist sowohl die Nukleotidsequenz
kodierend für die Führungs-RNA als auch die Nukleotidsequenz
kodierend für die Cas-Nuklease in das Genom von *E. coli*
25 integriert. In einer anderen bevorzugten Ausführungsform ist die
Nukleotidsequenz kodierend für die Cas-Nuklease in das Genom von
E. coli integriert, nicht aber die Nukleotidsequenz kodierend für
die Führungs-RNA. Verfahren zum Integrieren von
Nukleotidsequenzen in bakterielle Genome sind einem Fachmann
30 bekannt (z.B. Datsenko et al. 2000, Proc. Natl. Acad. Sci. U S
A. 97: 6640 - 5).

In einem anderen Aspekt betrifft die vorliegende Erfindung die
Verwendung einer Cas-Nuklease vom Typ II zur Linearisierung eines
35 doppelsträngigen, zirkulären DNA-Moleküls, worin das lineare DNA-

Molekül auf dem Matrizen-Strang in der 3'- zu 5'-Richtung umfasst:

- (i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,
- 5 (ii) eine Poly T-Sequenz,
- (iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C oder ein Nukleotidanalogen davon darstellen, und
- (iv) ein Protospacer Adjacent Motif (PAM).

10

In einem weiteren Aspekt betrifft die vorliegende Erfindung die Verwendung eines rekombinanten Mikroorganismus, worin die Nukleotidsequenz kodierend für eine Führungs-RNA und/oder die Nukleotidsequenz kodierend für eine Cas-Nuklease vom Typ II in
15 das Genom integriert ist, zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Stranges aufweist.

20 Eine Ausführungsform betrifft die Verwendung eines rekombinanten Mikroorganismus, worin die Nukleotidsequenz kodierend für eine Führungs-RNA und die Nukleotidsequenz kodierend für eine Cas-Nuklease vom Typ II in das Genom integriert ist. Eine andere Ausführungsform betrifft die Verwendung eines rekombinanten
25 Mikroorganismus, worin die Nukleotidsequenz kodierend für eine Cas-Nuklease vom Typ II in das Genom integriert ist, zur Herstellung eines linearen doppelsträngigen DNA-Moleküls.

Der rekombinante Mikroorganismus kann ein beliebiger, für die
30 Erzeugung von Plasmid-DNA geeigneter Mikroorganismus sein. Bevorzugt ist der rekombinante Mikroorganismus *E. coli*, mehr bevorzugt ist der rekombinante Mikroorganismus ein *E. coli* K12 Stamm, besonders bevorzugt ist der rekombinante Mikroorganismus *E. coli* NEB10Beta.

35

Die Erfindung wird weiterhin im Folgenden anhand von Ausführungsbeispielen und den beigefügten Abbildungen näher beschrieben, ohne dadurch beschränkt zu werden.

5 Abbildung 1 zeigt ein Agarose-Gel mit einer gelelektrophoretisch aufgetrennten Plasmid-DNA und verschiedenen Kontrollen: 1: Marker (GeneRuler™ 1 kb DNA Ladder) 2: pGFP1 ohne Inkubation (Negativkontrolle 1) 3: pGFP1 in Wasser mit Inkubation (Negativkontrolle 2) 4: pGFP1 in Schneidepuffer (Negativkontrolle
10 3) 5: pGFP1 in Schneidepuffer + 1 mM EDTA (Negativkontrolle 4) 6: pGFP1 inkubiert mit XbaI (Positivkontrolle 1) 7: pGFP1 inkubiert mit NcoI (Postivkontrolle 2) 8: pGFP1 inkubiert mit 25 nM Cas9 (Negativkontrolle 5) 9: pGFP1 inkubiert mit 25 nM sgRNA (Negativkontrolle 6) 10-15: Testproben 1-5 (1-facher, 2-facher,
15 3-facher, 4-facher, 5-facher und 10-facher Überschuss Cas9/Führungs-RNA zu pDNA).

Abbildung 2 zeigt ein Agarosegel mit gelelektrophoretisch aufgetrennten mRNAs, die in *in vitro*-Transkriptionen unter
20 Verwendung unterschiedlicher DNA-Template gebildet wurden. Cas9 bezeichnet die Probe, worin die mittels *Sp*Cas9/Führungs-RNA linearisierte DNA als Templat-DNA verwendet wurde. Typ IIS bezeichnet die Probe, worin die mittels Typ IIS Restriktionsenzym linearisierte DNA als Template-DNA verwendet wurde. PCR
25 bezeichnet die Probe, worin die mittels PCR erzeugte DNA als Templat verwendet wurde. pGFP und pGFP2 bezeichnen die Proben, worin die zirkulären Plasmide pGFP1 und pGFP2 als DNA-Template verwendet wurden.

30 Abbildung 3 zeigt die Analyse des Poly A-Anhangs mittels Acrylamid-Gelelektrophorese. 1: Marker, GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific). Cas9: RT-PCR-Produkt für die RNA, welche mit einer mittels *Sp*Cas9-linearisierten Templat-DNA hergestellt wurde. TypIIS: RT-PCR-Produkt für die RNA, welche mit
35 einer mittels Typ IIS Restriktionsenzym linearisierten Templat-

DNA hergestellt wurde. PCT: RT-PCR-Produkt für die RNA, welche mit einer mittels PCR hergestellten Templat-DNA hergestellt wurde.

5 Abbildung 4 zeigt die Basalexpression und Expressionsstärke der veränderten Promotoren WAC003 und WAC003-LacIDel im Vergleich zur pTac1 Ausgangsvariante am Beispiel der Expression des Reporterproteins GFP in *E.coli* K12.

10 Abbildung 5 zeigt eine Schematische Darstellung des Plasmids pCas9_tracrRNA_CRISPR.

Abbildung 6 ist eine schematische Darstellung eines Konstrukts für die genomische Integration mittels λ -Red Rekombination.

15

Beispiele

Beispiel 1: Plasmid-DNA-Schneideaktivität durch Cas9

20 Zunächst wurde für die Herstellung des Plasmids ein Konstrukt mit Templat-Sequenz, Poly A-Anhang sowie Zielsequenz für den Schnitt einer Cas-Nuklease und PAM entworfen. Dieses Konstrukt kodiert neben einer GFP-mRNA Templat-Sequenz am 3' Ende dieser Sequenz für einen Poly A-Anhang, welcher wiederum von drei weiteren
25 Nukleotiden (5'-GTC-3') sowie dem *Protospacer Adjacent Motif* (PAM) 5'-CGG-3' gefolgt wird. Die PAM-Sequenz stellt die Cas9-gRNA Erkennungssequenz zur Prozessierung dar. Für die Erstellung der Templat-DNA Sequenz wurde das *gfp*-Gen (Uniprot: C5MKY7) mit den beiden dafür entsprechenden Primern (SEQ ID NO:8 und 9) in
30 einer PCR nach Standardprotokoll mit der Phusion Polymerase amplifiziert (Thermo Fisher Scientific). Die beiden Primer umfassten zum einen den T7-RNA-Polymerase-Promotor (Forward Primer, SEQ ID NO:8) und zum anderen die Elemente: Poly T, 5' GAC 3' und PAM (5'-CGG-3') (Reverse Primer, SEQ ID NO:9). Das PCR-
35 Produkt (SEQ ID Nr:10) wurde anschließend über die *Xba*I- und

*Eco*RI-Schnittstellen in einen pUC18-Vektor (New England Biolabs) kloniert. Das resultierende Plasmid pGFP1 (SEQ ID NO:11) wurde in den folgenden Versuchen/Beispielen als Ausgangsplasmid für die Herstellung der linearisierten Templat-DNA eingesetzt.

5

Nach erfolgreicher Herstellung des Plasmids pGFP1 wurde die Plasmid-DNA mit Hilfe von *E. coli* NEB10Beta (New England Biolabs) produziert. Dazu wurden die NEB10Beta Zellen nach Herstellerangaben mittels Hitzeschock-Methode mit dem Plasmid pGFP1 transformiert. Hierfür wurden die Zellen und 100 ng pGFP1 für 30 min auf Eis inkubiert, bevor ein 30 Sekunden Hitzeschock bei 42 °C erfolgte. Nach einer Regenerationsphase (1 h bei 37 °C) folgte die Selektion der Transformanten auf Selektionsmedium (LB-Ampicillin (LB-Amp)). Entstandene Klone wurden in 3 mL-Kulturen unter selektiven Bedingungen angezogen und das Plasmid der Klone über Restriktionsmuster und Sequenzierung überprüft. Klone mit dem gewünschten Plasmid wurden in 50 mL LB-Amp angezogen (schüttelnd, 37 °C über Nacht), um ausreichende Mengen an Plasmid-DNA zu generieren. Je nach Klon wurde eine OD₆₀₀ (optische Dichte bei 600nm gegen Wasser) von 8-14 im Schüttelkolben erreicht. Die Plasmid-DNA wurde abschließend mit dem Gene Jet Plasmid Kit (Thermo Fisher Scientific) entsprechend den Angaben des Herstellers aus der Über-Nacht-Kultur (stationäre Phase wurde erreicht) isoliert und gereinigt.

25

In einem nächsten Schritt wurde eine Templat-DNA für die *in vitro*-Herstellung einer Führungs-RNA (SEQ ID NO:1) entworfen. Um die Synthese einer Führungs-RNA mit spezifischer Länge und spezifischer Sequenz zu ermöglichen, wurde die DNA-Sequenz komplementär zu der Führungs-RNA 3' hinter eine T7 RNA-Polymerase-Promotorsequenz angefügt (SEQ ID NO:12). Zusammen mit der komplementären Sequenz des T7 Promoters (SEQ ID NO:13) wurde dieses als lange Einzelstränge (Primer) synthetisiert (Fa. Metabion).

35

Zur Herstellung der Führungs-RNA (SEQ ID NO:1) wurden diese beiden DNA-Oligonukleotide (SEQ ID NO:12 und 13) miteinander hybridisiert, indem beide in äquimolaren Mengen (100 pmol) gemischt, für 5 min bei 95 °C aufgeköcht und anschließend langsam auf Raumtemperatur heruntergekühlt wurden. Das erzeugte Hybrid wurde anschließend als Matrize genutzt, um die Führungs-RNA mit einem *in vitro*-Transkriptionskit (HighYield T7 RNA Synthesis Kit von Jena Bioscience) nach Herstellerangaben zu synthetisieren. Anschließend wurde die Transkriptionsreaktion mit dem Monarch RNA Cleanup Kit von New England Biolabs (NEB) nach Herstellerangaben gereinigt und die Konzentration mittels UV-Absorption bestimmt.

Nach der Herstellung des Plasmids pGFP1 (SEQ ID NO:11) und der Führungs-RNA (SEQ ID NO:1), wurden beide zusammen mit der Cas-Nuklease *SpCas9* (New England Biolabs) in einem Schneidepuffer inkubiert (20 mM Hepes pH 7.5; 150 mM KCl, 10 mM MgCl₂, 0.5 mM DTT, 0.1 mM EDTA). Dabei wurden *SpCas9* und die Führungs-RNA äquimolar eingesetzt und jeweils in verschiedenen Überschüssen zur pDNA (5 nM) (1-fach, 2-fach, 3-fach, 4-fach, 5-fach und 10-fach) verwendet. Zur besseren Komplexbildung, wurden *SpCas9* und die Führungs-RNA zunächst allein inkubiert (15 min, 37 °C). Anschließend wurden die pDNA (5 nM) und der Reaktionspuffer hinzugefügt. Nach 60 min Inkubationszeit bei 37 °C wurde die Reaktion mit Zugabe von 20 µL 2x Formamid-Puffer (Thermo Fisher Scientific) abgestoppt und die pDNA mittels Agarose-Gelelektrophorese analysiert (Abbildung 1).

Für die Analyse wurde die geschnittene pDNA über ein 1%-iges Agarose-Gel (TAE gepuffert) getrennt und anschließend mittels Midori Green (Nippon Genetics Europe) an einer Geldokumentationsanlage (BioRad) visualisiert (siehe Abbildung 1). Als Kontrollen dienten ungeschnittene und nicht inkubierte Plasmid-DNA, ungeschnittene inkubierte (1h, 37 °C) pDNA, pDNA mit Schneidepuffer, pDNA mit Schneidepuffer und 1 mM EDTA, pDNA mit 25 nM Cas9 ohne Führungs-RNA, pDNA mit 25 nM Führungs-RNA ohne

Cas9, sowie als Positivkontrolle pDNA (SEQ ID NO:11), die mittels *Xba*I bzw. *Nco*I (verwendet nach Herstellerangaben, Thermo Fisher Scientific) linearisiert wurde.

5 Abbildung 1 zeigt, dass die Plasmid-DNA pGFP1 durch den Cas9/Führungs-RNA Komplex linearisiert wird. Aufgetragen wurde in Spur 1 der Marker, GeneRuler™ 1 kb DNA Ladder (Thermo
Fisher Scientific). Als Kontrollen dienten: pGFP1 ohne
10 Inkubation (Spur 2, Ausgangsmaterial), pGFP1 in Wasser mit
Inkubation (Spur 3, ohne Cas9 und Führungs-RNA), pGFP1 in
Schneidepuffer (Spur 4, ohne Cas9 und Führungs-RNA), pGFP1 in
Schneidepuffer + 1 mM EDTA (Spur 5, ohne Cas9 und Führungs-
RNA), pGFP1 inkubiert mit *Xba*I (Spur 6), pGFP1 inkubiert mit
15 *Nco*I (Spur 7), pGFP1 inkubiert mit 25 nM Cas9 (ohne Führungs-
RNA; Spur 8), pGFP1 inkubiert mit 25 nM Führungs-RNA (ohne
Cas9; Spur 9). Die Testproben wurden in den Spuren 10 – 15
aufgetragen (1-fach, 2-facher, 3-facher, 4-facher, 5-facher
und 10-facher Überschuss Cas9/Führungs-RNA zu pDNA).
Abbildung 1 zeigt zudem, dass bei zunehmender Menge des Cas9/
20 Führungs-RNA-Komplexes mehr pDNA linearisiert wird. Bei einem
10-fachen molaren Überschuss (Spur 15) ist nahezu der gleiche
Grad der Linearisierung erreicht, wie bei der Verwendung von
herkömmlichen Type II-Restriktionsenzymen. Dies ist
exemplarisch mit den Reaktionen unter Nutzung von *Xba*I und
25 *Nco*I (Spuren 6 & 7) dargestellt. Die Abhängigkeit von der
Cas9/Führungs-RNA-Komplexbildung zeigen die beiden Spuren 8
und 9, in denen bei einem 10-fachen molaren Überschuss jeweils
eine Komponente, entweder die Führungs-RNA (Spur 8) oder Cas9
(Spur 9), weggelassen wurde. Hier ist keine Linearisierung zu
30 beobachten, ebenso wie bei pGFP1 Proben, in denen keine Enzyme
verwendet wurden (Spur 3, Spur 4, Spur 5).

Beispiel 2: mRNA-Synthese aus pDNA, welche durch Cas9 geschnitten wurde

Die Plasmid-DNA pGFP1, (SEQ ID NO:11) wurde wie in Beispiel 1
5 beschrieben mittels SpCas9 und der Führungs-RNA (SEQ ID NO:1)
linearisiert (10-facher molarer Überschuss, siehe auch Abbildung
1, Spur 15). Das linearisierte Plasmid wurde anschließend über
das Gene Jet PCR Purification Kit (Thermo Fisher Scientific)
gereinigt. Die lineare, aufgereinigte pDNA wurde anschließend als
10 Templat-DNA in einer *in vitro*-Transkriptionsreaktion (HighYield
T7 RNA Synthesis Kit, Jena Bioscience) nach Herstellerangaben
eingesetzt, um eine mRNA zu synthetisieren. Die Reaktion wurde
für 2 h bei 37 °C durchgeführt, dabei wurden 200 ng Templat-DNA
eingesetzt. Um Verunreinigungen und Kitbestandteile, wie
15 beispielsweise Enzyme und Nukleotide, aus der Reaktion zu
entfernen, wurde die mRNA entsprechend den Herstellerangaben über
ein Monarch RNA Cleanup Kit (NEB) aufgereinigt.

Um mRNA-Herstellungsprozesse unter Verwendung des
20 erfindungsgemäßen Verfahrens mit den Methoden aus dem Stand der
Technik zu vergleichen, d.h. Linearisierung einer Plasmid-DNA
mittels Typ IIS Restriktionsenzymen bzw. Herstellung eines PCR-
Produktes als lineare Ausgangs-DNA, wurde pGFP2 (SEQ ID NO:14,
hergestellt in Analogie zu pGFP1, durch Amplifikation des gfp
25 Gens mittels Primern der SEQ ID NO:9 und 15.) mit dem Typ IIS
Restriktionsenzym *BpiI* nach Herstellerangaben (Thermo Fisher
Scientific) behandelt und linearisiert. Zusätzlich wurde eine
lineare Ausgangs-DNA über eine Polymerase-Ketten-Reaktion (PCR-
Produkt) (SEQ ID NO:16) hergestellt. Dabei wurde ein Primer mit
30 einem T7 RNA-Polymerase-Promoter (Forward Primer, SEQ ID NO:17)
zusammen mit einem Primer mit einem 60 nt Poly A-Anhang (Reverse
Primer, SEQ ID NO:18) sowie mit pGFP1 als PCR-Vorlage (10 ng) und
der Phusion DNA Polymerase verwendet. Die Reaktion wurde im
mitgelieferten Puffersystem nach Herstellerangaben durchgeführt
35 (Thermo Fisher Scientific). Das PCR-Produkt wurde nach der

Reaktion mit *DpnI* verdaut, um verbliebene methylierte Plasmid-DNA aus der PCR-Reaktion zu entfernen. Abschließend wurden sowohl das PCR-Produkt als auch das Typ IIS linearisierte Plasmid über ein PCR Purification Kit (Thermo Fisher Scientific) nach
5 Herstellerangaben gereinigt. Während die lineare DNA aus der PCR-Reaktion bzw. die mit *SpCas9* linearisierte DNA jeweils glatte („blunt“) Enden aufweisen, weist die Type IIS linearisierte DNA einen 3' Überhang auf.

10 Die linearen DNAs aus dem Restriktionsverdau und der PCR-Reaktion wurden in den gleichen Mengen wie die mittels *SpCas9* linearisierte pDNA (200 ng) nach Herstellerangaben in einer *in vitro*-Transkriptionsreaktion als Templat-DNA zur Synthese einer mRNA eingesetzt (HighYield T7 RNA Synthesis Kit, Jena Bioscience). Die
15 Synthese wurde bei 37°C für 2h durchgeführt. Abschließend wurden die mRNAs über das Monarch RNA Cleanup Kit (NEB) entsprechend der Herstellerangaben gereinigt.

Die Analyse der hergestellten mRNAs erfolgte zunächst über eine
20 Polyacrylamid-Gelelektrophorese. Dabei wurden 200 ng mRNA von jedem Ansatz mit RNA-Ladepuffer (Gel Loading Buffer 2, Thermo Fisher Scientific) versetzt und für 5 min bei 95 °C aufgeköcht. Dann erfolgte die Analyse der RNA-Spezies über die Auftrennung auf einem 15 % TBE-UREA Polyacrylamid Gel (Thermo Fisher
25 Scientific; 250 V, 50 min). Im Anschluss an die Elektrophorese wurde das Gel mittels SYBR Green II RNA Gel Stain (Thermo Fisher Scientific) gefärbt und in einer Geldokumentationsanlage (BioRad) visualisiert (Abbildung 2). Als Kontrollen dienten *in vitro*-Transkriptionsansätze mit 200 ng ungeschnittenem pGFP1 und pGFP2.
30 pGFP2 ist eine pGFP1 Variante, welche keine PAM-Sequenz enthält (SEQ ID NO:14). Die Konzentration der hergestellten mRNAs wurde photometrisch mit einem NanoDrop-Photospektrometer (Thermo Fisher Scientific) bestimmt.

Abbildung 2 ist eine Visualisierung der hergestellten mRNAs. In Spur 1 befindet sich der Century Marker™ (Thermo Fisher Scientific) und in den folgenden Spuren die RNA-Proben entsprechend der Ausgangs-DNAs: SpCas9/Führungs-RNA linearisierte DNA (Cas9), mit Typ IIS Restriktionsenzym linearisierte DNA (Type IIS), DNA aus der PCR-Reaktion (PCR) sowie die ungeschnittenen Plasmide pGFP1 und pGFP2. Abbildung 2 zeigt deutlich, dass die erhaltenen mRNA-Hauptprodukte aus den Reaktionen mit linearisierten Plasmiden (Cas9, Type IIS) und PCR-Produkt (PCR) identische Größen aufweisen. Nur bei Reaktionen mit nicht linearisierten Plasmiden (pGFP1 und pGFP2) zeigen sich größere Produkte. Diese Produkte resultieren aus der „ungestoppten“ *in vitro*-Transkription der T7 RNA-Polymerase. Durch die fehlende Linearisierung ist kein „Run-OFF“ der Polymerase möglich was zu einem „Durchlesen“ und damit zu heterogenen, größeren Produkten führt.

Für die Analyse des Poly A-Anhanges der mRNAs wurde mittels einer T4 RNA-Ligase 1 (NEB) ein Adapter-Oligonukleotid (SEQ ID NO:19) an das 3' Ende der mRNA ligiert. Die T4 RNA-Ligase 1 ist ein Enzym, welches einzelsträngige RNA ligieren kann. Dabei wird ein 5' Monophosphat-haltiges Substrat an ein 3' OH-Ende der Empfänger-mRNA ligiert. Das 3' OH-Ende enthält die hergestellte RNA natürlicherweise. Das Adapter-Oligonukleotid wurde so synthetisiert, dass es ein 5' Monophosphat enthält (Metabion). Um eine Selbstligation der Adapter-Oligonukleotide zu verhindern, wurde auch das 3' Ende des Adapter-Oligonukleotides mit einem synthetischen Phosphatrest versehen. Für die Ligation wurden 5 µg RNA und 40 pmol des Adapter-Oligonukleotids eingesetzt. Die Reaktion (50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 0.5 U RNase Inhibitor, 1 mM ATP, 10 U T4 RNA Ligase) inkubierte für 2h bei 25° C. Dann wurde ein Pufferwechsel zu Wasser mittels eines Amicons (Merck) (100 kDa cut-off) durchgeführt.

Die erhaltenen Adapter-Ligationsprodukte (Adapter Oligonukleotid + mRNA) wurden im Anschluss über eine reverse Transkriptase - Reaktion (RT-PCR) amplifiziert. Für die RT-PCR fand das SuperScript™ Kit entsprechend der Herstellerangaben Verwendung
5 (Thermo Fisher Scientific). Für die Reaktion wurden spezifische Primer verwendet, die direkt 5' vor dem Poly A (Forward, SEQ ID NO:20) und an der Adaptersequenz binden (Reverse, SEQ ID NO:21). Die aus der RT-PCR enthaltenen Produkte wurden über ein PCR Purification Kit von Thermo Fisher nach Herstellerangaben
10 gereinigt und die Proben anschließend mittels Elektrophorese in einem 15 % Polyacrylamid Gel analysiert. Dazu wurde die aufgetrennte DNA mittels Midori Green (Biozym) gefärbt und in einer Geldokumentationsanlage (BioRad) visualisiert (siehe Abbildung 3).

15

Abbildung 3 zeigt die Analyse des Poly A-Anhangs mittels Acrylamid-Gelelektrophorese. In Spur 1 wurde der Marker, GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific) aufgetragen. In den folgenden 3 Spuren wurden die RT-PCR-Produkte für die RNAs,
20 welche mit unterschiedlich linearisierten DNA-Templaten hergestellt wurden, aufgetrennt (*SpCas9*-linearisiert (Spur2), mit Typ IIS Restriktionsenzym linearisiert (Spur3) und PCR-Produkt (Spur4)).

25 In Abbildung 3 ist zu erkennen, dass in allen 3 RNA-Proben die gleichen Produkte mit der exakt gleichen Größenverteilung der RT-PCR-Produkte erhalten wurden. Da die Ursprungs-RNAs bereits sehr große Ähnlichkeiten aufweisen (vgl. Abbildung 2) ist davon auszugehen, dass die RT-PCR hier die gleichen bzw. sehr ähnliche
30 Template als Ausgangsmaterial hatte. Die Gelmuster zeigen also, dass mit der erfindungsgemäßen Methode nicht nur pDNA effektiv linearisiert werden kann, sondern dass die pDNA auch an der korrekten, vorher definierten Stelle linearisiert wird und der gewünschte unverdeckte freie Poly A-Anhang entsteht.

35

Beispiel 3: Promotor-Design für eine induzierbare Expression von Cas9

Da eine konstitutive Expression einer Cas-Endonuklease, 5
beispielsweise Cas9, in einer Bakterienzelle *in vivo*
unspezifische DNA-Schnittereignisse an verschiedenen Stellen
sowohl in Plasmiden als auch im Chromosom einer Bakterienzelle
verursachen kann, und so die Bakterienzelle als Plasmidproduzent
schädigen kann, ist eine induzierbare Expression der Cas-
10 Endonuklease, beispielsweise Cas9, unter einem regulierbaren
Promoter mit möglichst geringer Basalexpression bevorzugt.

Für die induzierbare Expression der Cas-Nuklease und damit der
zu einem gewünschten Kultivierungszeitpunkt via Induktor-Zugabe
15 „anschaltbaren“ *in vivo*-Linearisierung des Plasmidprodukts ist
zusätzlich eine Begrenzung der induzierbaren Expressionshöhe
bevorzugt. Eine zu hohe Expression einer Cas-Nuklease in einer
kurzen Zeitspanne kann ähnlich negative Effekte auf die Plasmid-
DNA und die Produzentenzelle haben wie eine hohe Basalexpression
20 über eine längere Zeitspanne. Daher sind Promotoren besonders
bevorzugt, die sowohl eine geringe Basalexpression als auch nach
einer Induktion eine moderate Expressionsstärke aufweisen.

Der pTac1 Promoter (de Boehr, 1983, Proceedings of the National
25 Academy of Sciences 80(1):21-5), ein Hybrid aus *trp*- und *lac*-
Promoter, besitzt vom ursprünglichen *Lac*-Operon nur noch eine
verbliebene Operator-Site *LacO1* an Position +1
(Transkriptionsstart) zum Binden des Repressorproteins *LacI*. Das
Repressorprotein *LacI* selbst, bestehend aus einer N-terminalen
30 „headpiece“ DNA-Bindeeinheit, einer *hinge*-Region, einem Core-
Bereich mit N-terminalen und C-terminalen Subdomänen inkl.
Lactose-Bindungsereich und Dimerisierungsmotiv, sowie einem C-
terminalen Mini-Zipper für die Tetramerisierung, assoziiert zu
einem Dimer von Dimeren. Jedes *LacI*-Dimer bindet eine 21 bp Duplex
35 DNA-Operator Sequenz *LacO* (Lewis 2011. J Mol Biol. 409, 14-27).
Vom assoziierten *LacI*-Tetramer werden im nativen *Lac*-Operon zwei

Duplex-DNA-Operator-Sites (LacO1 und LacO3) simultan gebunden, was im DNA-Strang einen sogenannten „*repression loop*“ für eine maximale Repression ausbildet (Oehler et al. 2006, *Nucleic Acids Res.* 34, 606-612).

5

In WAC003 (SEQ ID NO:6) und WAC003-LacIdel (SEQ ID NO:7) wurde eine synthetische LacO-Bindungsstelle über die -35 Region des pTac1 Promoters eingefügt, ohne die Nukleotidsequenzen der -35 und -10 Promoter-Regionen, den Abstand (Spacer) zwischen beiden
10 Regionen oder die Nukleotidsequenz des Spacers selbst zu verändern. WAC003 interagiert mit dem nativen, also zur Tetramerisierung fähigem LacI-Repressorprotein. Im Konstrukt WAC003-LacIdel wurde das Gen *lacI* durch eine Variante *LacIdel* ersetzt, die in Folge eines Nukleotid-Frameshifts am C-Terminus
15 des kodierten LacI Proteins eine andere Aminosäuresequenz ausbildet und so das LacI-Tetramer nicht mehr ausbilden kann, sondern nur noch das LacI-Dimer.

Die Basalexpression (Repression) und die nach Induktor-Zugabe erreichbare Expression (Derepression) wurde mit dem
20 Reporterprotein GFP geprüft. GFP als zytoplasmatisch lokalisiertes Protein stellt selbst keine Belastung für die Zellen dar und eignet sich somit gut als Referenzprotein zur Messung der Expressionsleistung. Das Gen für GFP wurde unter die
25 Kontrolle der WAC003 und WAC003-LacIdel-Promotorkonstrukte gestellt, auf ein Plasmid mit *ColE1 origin of replication* kloniert, die Plasmide in einen *E. coli* K12 Stamm transformiert und die Basalexpression sowie die induzierbare Expression in Schüttelkultur in LB-Medium wie folgt getestet.

30

Aus Vorkulturen wurden Hauptkulturen auf $OD_{600} = 0,1$ angeimpft. Nach 2 h, bei einer korrelierenden Zelldichten von $OD_{600} \sim 0,4$, erfolgte keine (= Basalexpression) bzw. die Zugabe von 0,1 mM IPTG (= Induktion). Nach weiteren 24 h Kultivierung wurde die
35 Menge an exprimierten GFP vermessen (Anregung 380 nm, Emission 500 nm) und auf die jeweils zugrunde liegende Zelldichte OD_{600} der

Kultur normiert. Als Kontrolle wurde eine Zelllinie mit dem pTac1 Promoter (hohe Basalexpression, hohe induzierbare Expressionsstärke) mitgeführt.

5 Abbildung 4 zeigt die Basalexpression und Expressionsstärke der WAC003 und WAC003-LacIdel-Promotorkonstrukte im Vergleich zur Tac1 Ausgangsvariante. Die WAC003 Variante zeigte eine gegenüber Tac1 deutlich verringerte Basalexpression. Nach IPTG-Zugabe war die Derepression des WAC003 Promoters ebenfalls deutlich
10 reduziert (nur ~16 % GFP/OD₆₀₀ verglichen mit Tac1). Im WAC003-LacIdel Konstrukt ist die Basalexpression ebenfalls niedrig, die induzierbare Expressionshöhe ist zwar etwas höher als bei WAC003, aber immer noch sehr niedrig im Vergleich zu Tac1. Die Promotoren WAC003 und WAC003-LacIdel eignen sich somit sehr gut für die
15 induzierbare Expression einer Cas-Nuklease, beispielsweise Cas9, welche *in vivo* zur spezifischen Linearisierung von Plasmid-DNA exprimiert werden soll.

20 **Beispiel 4: In vivo-Induktion von Cas9 und Aufreinigung von linearer DNA nach Inkubation**

Zur Analyse der *in vivo*-Funktionalität des in Beispiel 1 und Beispiel 2 beschriebenen Systems, wurden sowohl das Gen für SpCas9 kodiert (SEQ ID NO:2), als auch das Konstrukt für die Führungs-
25 RNA (SEQ ID NO:1) im selben Plasmid (pWAC003, SEQ ID NO:5) jeweils unter den induzierbaren Promoter WAC003-LacIdel (SEQ ID NO:7) kloniert. Dabei wurden zwei SpCas9-spezifische Primer verwendet, (SEQ ID NO:22 und 23) um das *Sp_cas9*-Gen für die Klonierung aus dem Plasmid pCas9-ts1 (SEQ ID NO:24) zu amplifizieren. Das
30 resultierende PCR-Produkt wurde über *EcoRI* und *XbaI* in das Plasmid pWAC003 (SEQ ID NO:5) hinter den Promotor WAC003-LacIdel (SEQ ID NO:7) kloniert. Die Cas9-Expressionskassette des resultierende Plasmids pCas9 (SEQ ID NO:25) wurde mit spezifischen Primern (SEQ ID NO:26 und SEQ ID NO:27) sequenziert (Sequise), um die
35 gewünschte Sequenz zu bestätigen.

Da in einem *in vivo*-System die Herstellung der Führungs-RNA nicht durch eine Run-Off Transkription, wie in Beispiel 2 beschrieben, durchgeführt werden kann, bedarf es einer Anpassung des RNA-Designs. Für die Erzeugung einer funktionalen Führungs-RNA muss auf die natürliche Form der RNA-Prozessierung von CRISPR-Cas9 Systemen zurückgegriffen werden. Das heißt, es wird ein minimaler CRISPR-Array und eine tracrRNA bereitgestellt.

In das Plasmid pCas9 (SEQ ID NO:25) wurden eine tracrRNA-Sequenz (SEQ ID NO:28) sowie eine CRISPR-Sequenz (SEQ ID NO:29) eingefügt. Die Sequenzen SEQ ID NO:28 und SEQ ID NO:29 umfassen neben der tracrRNA-Sequenz bzw. der CRISPR-Sequenz jeweils eine WAC003-LacIdel-Promotorsequenz und eine Rho-unabhängige Terminator Sequenz, in welche die tracrRNA bzw. CRISPR-Sequenz eingebettet sind (5'-3'). Dafür wurden zunächst die beiden DNA-Fragmente jeweils über eine PCR hergestellt, für SEQ ID NO:28 wurden die Primer mit den SEQ ID NO:30 - 35 verwendet und für die SEQ ID NO:29 wurden die Primer mit den SEQ ID NO:36 - 41 genutzt. Die tracrRNA-Sequenz wurde über die beiden Restriktionsenzyme *KpnI* und *XbaI* in pCas9 kloniert, so dass sie 3' nach dem Cas9 Konstrukt liegt. In das entstandene Plasmid pCas9_tracrRNA (SEQ ID NO:42) wurde die CRISPR-Sequenz über die Schnittstellen *XmaI* und *XbaI* in 3' Position hinter die tracrRNA Sequenz kloniert. Die korrekte Assemblierung des entstandenen Plasmids pCas9_tracrRNA_CRISPR (SEQ ID NO:43) wurde mittels Sequenzierung überprüft, dazu wurde der spezifische Primer (SEQ ID NO:27) verwendet.

Abbildung 4 zeigt eine Schematische Darstellung des Plasmids pCas9_tracrRNA_CRISPR. Im Plasmid pCas9_tracrRNA_CRISPR sind alle drei Elemente Cas9, tracrRNA und CRISPR jeweils unter Kontrolle eines WAC003-LacIdel Promoters und verfügen zusätzlich über eine Rho-unabhängige Terminator Sequenz.

Das Plasmid pCas9_tracrRNA_CRISPR (SEQ ID NO:43) wurde zusammen mit pGFP1 (SEQ ID NO:11) in *E. coli* NEB10Beta transformiert. Nach

Selektion auf LB-Ampicillin/Tetracyclin wurden Klone in 20 mL LB mit Antibiotika (100 mg/L Ampicillin oder 20 mg/L Tetracyclin) angeimpft und über Nacht bei 37 °C inkubiert (200 rpm). Mit den Übernachtskulturen wurde 100 mL frisches LB-Medium mit Ampicillin
5 bzw. Tetracyclin (100 mg/L bzw. 20 mg/L) angeimpft. Die resultierende Hauptkultur wurde bis zu einer OD_{600nm} von 0.8 unter Schütteln bei 37°C inkubiert. Bei Erreichen der Ziel-OD wurde IPTG auf eine finale Konzentration von 0,5 mM hinzugefügt, um die Expression der Cas9, der CRISPR-Kasette und der tracrRNA
10 gleichzeitig zu induzieren. Aufgrund der Co-Expression dieser Elemente wird das Plasmid pGFP1 spezifisch gespalten. 2 h nach Induktion wurden die Zellen mittels Zentrifugation sedimentiert. Die DNA im Überstand (linearisiertes pGFP1 und zirkuläres pCas9_tracrRNA_crRNA) wurde anschließend nach Herstellerangaben
15 mit dem Gene Jet Plasmid Kit (Thermo Fisher Scientific) isoliert und gereinigt.

Die Zusammensetzung der isolierten DNA und der Anteil des linearisierten pGFP1 wurde mittels Agarose Gelelektrophorese und
20 Kapillarelektrophorese analysiert. Für die Agarose Gelelektrophorese wurde die isolierte DNA über ein 1,5%-iges Agarose-Gel getrennt und anschließend mittels Midori Green (Nippon Genetics Europe) und einer Geldokumentationsanlage (BioRad) visualisiert.

25 Die Verwendbarkeit der linearisierten DNA als Templat für die Herstellung von mRNA wurde wie in Beispiel 2 beschrieben über eine *in vitro*-Transkription überprüft. Die erhaltenen RNAs wurden mittels Polyacrylamid Gelelektrophorese analysiert und die Poly
30 A-Anhänge der mRNAs wurden über die beschriebene Methode mit T4 RNA Ligase 1, RT-PCR analysiert.

Beispiel 5: Inkorporation des *cas9*-CRISPR-tracrRNA Konstruktes auf dem Chromosom eines ausgewählten *E. coli* Stamms

5 Das in Beispiel 4 beschriebene CRISPR-Cas Konstrukt pCas9_tracrRNA_CRISPR (SEQ ID NO:43) wurde in das Chromosom von *E. coli* NEB10Beta integriert. Für die Integration wurde ein intergenischer Bereich des Genlokus von *atpI* gewählt. Für die Integration des Cas9-gRNA Konstruktes wurde die Gen-Deletions-
10 Methode mittels λ -Red Rekombination nach Datsenko und Wanner (Datsenko et al. 2000, Proc. Natl. Acad. Sci. U S A. 97: 6640 - 5) genutzt.

Für die Integration des Cas9-CRISPR-tracrRNA Konstrukts in den
15 intergenischen Bereich des *atpI* Lokus in das Genom des verwendeten *E. coli* NEB10Beta Stammes wurde zunächst das zu integrierende Konstrukt entworfen. Dieses enthält neben dem in Beispiel 3 beschriebenen *cas9_tracrRNA_CRISPR* Konstrukt (SEQ ID NO:43) eine Chloramphenicol Resistenzkassette (*cat*), zwei FRT
20 sites, welche die *cat* Kassette flankieren, sowie jeweils einen flankierenden homologen Bereich zum Integrationsort. Das Konstrukt (SEQ ID NO:44) für die genomische Integration mittels λ -Red Rekombination ist schematisch in Abbildung 6 dargestellt. Zwei jeweils 150 bp lange homologe Bereiche zum Integrationsort
25 flankieren das Konstrukt am 3' und am 5' Ende. Dazwischen befinden sich in 5' - 3' Orientierung das *cas9_tracrRNA_CRISPR* Konstrukt, eine FRT site, die *cat* Kassette und eine weitere FRT site.

Die *cat* Kassette dient zur späteren Selektion, um eine
30 erfolgreiche Integration zu identifizieren. Die FRT sites werden im späteren Verlauf genutzt, um mit Hilfe einer Rekombinase die *cat* Kassette wieder aus dem Genom zu entfernen. Die homologen Regionen, welche das gesamte Konstrukt flankieren, sind 150 Basenpaare lang und dienen zur Integration am Zielort. Das

Einbauen des Konstruktes in das Genom findet durch homologe Rekombination statt.

Das vollständige Konstrukt wurde von GeneArt synthetisiert und
5 in den Vektor pMX Vektor (GeneArt) kloniert. Aus diesem Plasmid wurde das Konstrukt über Restriktionsenzyme (XhoI) herausgeschnitten und mittels GeneJet Gel-Extraktionskit (Thermo Fisher Scientific) nach Herstellerangaben gereinigt.

10 Um die homologe Rekombination für die Integration zu vermitteln, wurde das Lambda Red System verwendet, welches auf dem Plasmid pKD46 (CGSC: 7736) kodiert ist. Das Plasmid wurde in NEB10Beta Zellen transformiert und anschließend werden nach den Angaben von Datsenko und Wanner ((Datsenko et al. 2000, Proc. Natl. Acad.
15 Sci. U S A. 97: 6640 - 5) kompetente Zellen hergestellt. Diese kompetenten Zellen wurden mit dem gereinigten Integrationskonstrukt transformiert. Die Selektion auf Integration der Chloramphenicol-Resistenzkassette (*cat* = Chloramphenicol-Acetyltransferase) und damit auch des Cas9 bzw.
20 Führungs-RNA-Konstrukts in das Chromosom erfolgte auf LB-Agar-Platten, die 20 mg/l Chloramphenicol enthalten. Die Kontrolle bzw. Bestätigung der Integration an der gewünschten Position im Chromosom erfolgte mittels PCR unter Verwendung der Oligonukleotide Check Fwd (SEQ ID NO:54) und Check Rev (SEQ ID
25 NO:55) und chromosomaler DNA der Chloramphenicol-resistenten Zellen als Matrize. Als Negativkontrolle wurde die PCR auch mit chromosomaler DNA des Ausgangsstamms *E. coli* NEB10Beta durchgeführt. Auf diese Weise wurden *E. coli*-Zellen erhalten, bei denen das Cas9- bzw. Führungs-RNA-Konstrukt ins Genom integriert
30 ist.

Anschließend wurden die Zellen nach der von Datsenko und Wanner (s.o.) beschriebenen Prozedur von dem Plasmid pKD46 kuriert und der auf diese Weise erzeugte Stamm wurde als *E. coli* NEB10Beta
35 *atpI::cas9-CRISPR-tracrRNA-cat* bezeichnet. Die Entfernung der

Chloramphenicol-Resistenzkassette aus dem Chromosom erfolgte nach der Vorschrift von Datsenko und Wanner (s.o.) mit Hilfe des Plasmids pCP20 (CGSC: 7629), welches für das FLP-Rekombinase-Gen kodiert. Der schließlich mit dieser Methode erhaltene
5 Chloramphenicol sensitive Cas9-CRISPR-tracrRNA enthaltende Stamm wurde als *E. coli* NEB10Beta *atpI::cas9*-CRISPR-tracrRNA bezeichnet. Die erfolgreiche Integration des Konstruktes (SEQ ID NO:44) wurde abschließend über Sequenzierung, unter Verwendung verschiedener Primer (SEQ ID NO:45-54), überprüft.

10

In den neuen Stamm (*E. coli* NEB10Beta *atpI::cas9*-CRISPR-tracrRNA) wurde anschließend zur Prüfung des Systems das Plasmid pGFP1 (SEQ ID NO:11) transformiert. Nach der Regeneration und Selektion (über Nacht auf LB-Ampicillin Agar Platten) wurde eine Kolonie
15 ausgewählt, mit welcher zunächst eine Übernachtskultur in 20 mL LB-Amp angesetzt wurde (37 °C, 200 rpm). Diese Vorkultur diente zum Animpfen der 100 mL Hauptkultur. Die resultierende Hauptkultur wurde bis zu einer OD_{600nm} von 0.8 inkubiert. Bei Erreichen der Ziel-OD wurde IPTG hinzugefügt, um die Expression
20 sowohl von Cas9 als auch der Führungs-RNA zu starten. Nach 2h unter induzierenden Bedingungen wurde die Kultur mittels Zentrifugation geerntet. Die Plasmid-DNA in den Zellen wurde anschließend nach Herstellerangaben mit dem Gene Jet Plasmid Kit (Thermo Fisher Scientific) isoliert und gereinigt.

25

Die isolierte Plasmid-DNA wurde mittels Agarose-Gelelektrophorese und Kapillarelektrophorese analysiert. Für die Gelelektrophorese wurde die isolierte Plasmid-DNA mittels eines 1,5%-igen Agarose-Gels entsprechend der Größe und Eigenschaft (linarisiert, coiled)
30 getrennt und anschließend mittels Midori Green (Nippon Genetics Europe) an einer Geldokumentationsanlage (BioRad) visualisiert. Die Verwendbarkeit der DNA für die Herstellung von mRNA wurde wie in Beispiel 2 beschrieben über eine *in vitro*-Transkription geprüft.

35

Patentansprüche

1. Verfahren zur Herstellung eines linearen doppelsträngigen DNA-Moleküls,

5 worin das lineare DNA-Molekül glatte Enden aufweist, und
worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Matrizen-Strangs aufweist, umfassend die Schritte:

(1) Bereitstellen eines doppelsträngigen zirkulären DNA-Moleküls, umfassend auf dem Matrizen-Strang in der 3'- zu 5-
10 Richtung:

(i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,

(ii) eine Poly T-Sequenz,

(iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin N_1
15 bis N_3 jeweils unabhängig voneinander G, T, A oder C darstellen, und

(iv) ein Protospacer Adjacent Motif (PAM),

(2) Bereitstellen einer Führungs-RNA, umfassend eine Region, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und
20 zur Nukleotidsequenz (iii), und eine Region, die mit einer Cas-Nuklease interagiert,

(3) Bereitstellen einer Cas-Nuklease vom Typ II,

(4) In Kontakt bringen des doppelsträngigen zirkulären DNA-Moleküls mit der Führungs-RNA und der CAS-Nuklease.

25

2. Verfahren nach Anspruch 1, worin die Poly T-Sequenz (ii) aus zwischen etwa 40 bis etwa 250 T-Nukleotiden besteht, bevorzugt aus zwischen etwa 100 und etwa 140 T-Nukleotiden, und besonders bevorzugt aus etwa 120 T-Nukleotiden.

30

3. Verfahren nach Anspruch 1 oder 2, worin die Region der Führungs-RNA, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii) die Sequenz $3'-G(A)_nN_4N_5N_6-5'$ aufweist, worin n eine Zahl zwischen 17 und
35 23 ist, worin N_4 bis N_6 jeweils unabhängig G, U, A oder C

darstellt, und worin die Nukleotide $N_4N_5N_6$ komplementär sind zu den Nukleotiden $N_1N_2N_3$ aus (iii).

4. Verfahren nach einem der Ansprüche 1 bis 3, worin es sich bei der Cas-Nuklease vom Typ II um Cas9 handelt, bevorzugt ausgewählt aus Cas9 aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen Sequenz, Cas9 aus *Streptococcus aureus* mit der SEQ ID NO:3 oder einer dazu zu mindestens 70% homologen Sequenz und Cas9 aus *Streptococcus thermophilus* mit der SEQ ID NO:4 oder einer dazu zu mindestens 70% homologen Sequenz.
5. Verfahren nach einem der Ansprüche 1 bis 4, worin es sich bei der Cas-Nuklease vom Typ II um Cas9 aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen Sequenz handelt.
6. Verfahren nach einem der Ansprüche 1 bis 5, worin das PAM die Sequenz 5'-NGG-3' aufweist, wobei N jeweils unabhängig ausgewählt ist aus G, C, A und T.
7. Verfahren nach einem der Ansprüche 1 bis 6, worin die Poly T-Sequenz (ii) aus etwa 120 T-Nukleotiden besteht, worin die Region der Führungs-RNA, die komplementär ist zum 5'-Ende der Poly T-Sequenz (ii) und zur Nukleotidsequenz (iii) die Sequenz 3'-G(A)_n-N₄N₅N₆-5' aufweist, worin n gleich 17 ist, worin N₄ bis N₆ jeweils unabhängig G, U oder C darstellt, worin das PAM die Sequenz 5'-NGG-3' aufweist, wobei N jeweils unabhängig ausgewählt ist aus G, C, A und T, und worin es sich bei der Cas-Nuklease vom Typ II um Cas9 aus *Streptococcus pyogenes* mit der SEQ ID NO:2 oder einer dazu zu mindestens 70% homologen Sequenz handelt.

8. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass das Bereitstellen der Führungs-RNA durch die Expression einer Nukleotidsequenz, kodierend für die Führungs-RNA erfolgt.

5

9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass das Bereitstellen der Cas-Nuklease vom Typ II durch die Expression einer Nukleotidsequenz kodierend für die Cas-Nuklease erfolgt.

10

10. Verfahren nach einem der Ansprüche 8 oder 9, dadurch gekennzeichnet, dass die Nukleotidsequenz kodierend für die Führungs-RNA und/oder die Nukleotidsequenz kodierend für die Cas-Nuklease operativ mit einem induzierbaren Promotor verbunden ist.

15

11. Verfahren nach Anspruch 10, worin der induzierbare Promotor ein durch Lactose-induzierbarer Promotor ist, insbesondere der Promotor WAC003 mit der Sequenz SEQ ID NO:6 oder der Promotor WAC003-LacDel mit der Sequenz SEQ ID NO:7.

20

12. Verfahren nach einem der Ansprüche 8-11, dadurch gekennzeichnet, dass die Nukleotidsequenz kodierend für die Führungs-RNA und/oder die Nukleotidsequenz kodierend für die Cas-Nuklease in das Genom von *E. coli* integriert ist.

25

13. Verwendung einer Cas-Nuklease vom Typ II zur Linearisierung eines doppelsträngigen, zirkulären DNA-Moleküls, worin das lineare DNA-Molekül auf dem Matrizen-Strang in der 3'- zu 5'-Richtung umfasst:

30

(i) eine Templat-Sequenz, die operativ mit einem RNA-Polymerase-Promotor verbunden ist,

(ii) eine Poly T-Sequenz,

(iii) eine Nukleotidsequenz, dargestellt durch $N_1N_2N_3$, worin

35

N_1 bis N_3 jeweils unabhängig voneinander G, T, A oder C darstellen, und

(iv) ein Protospacer Adjacent Motif (PAM).

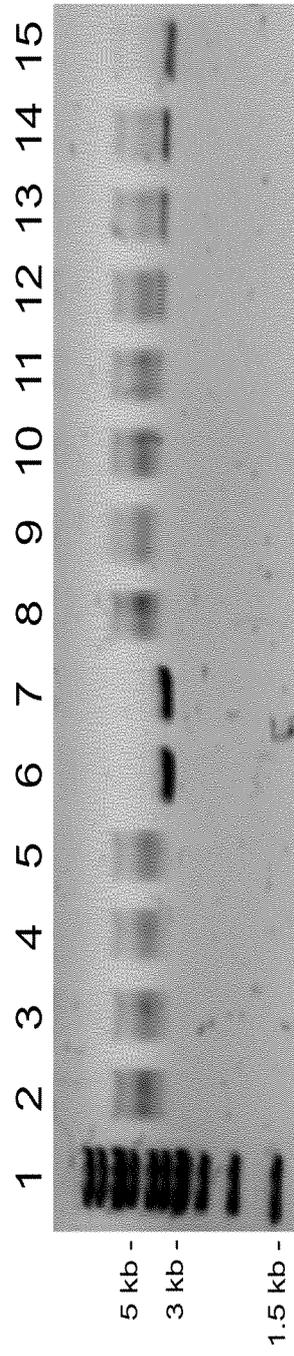
14. Verwendung eines rekombinanten Mikroorganismus, worin die Nukleotidsequenz kodierend für eine Führungs-RNA und/oder die Nukleotidsequenz kodierend für eine Cas-Nuklease vom Typ II in das Genom integriert ist, zur Herstellung eines linearen doppelsträngigen DNA-Moleküls, worin das lineare DNA-Molekül glatte Enden aufweist, und worin das lineare DNA-Molekül eine Poly T-Sequenz am 5' Ende des Templat-Stranges aufweist.

5
10

15. Verwendung nach Anspruch 14, worin der rekombinante Mikroorganismus *E. coli* ist, bevorzugt ein *E. coli* K12 Stamm, besonders bevorzugt *E. coli* NEB10Beta.

Abbildung 1

5



2/6

Abbildung 2

5

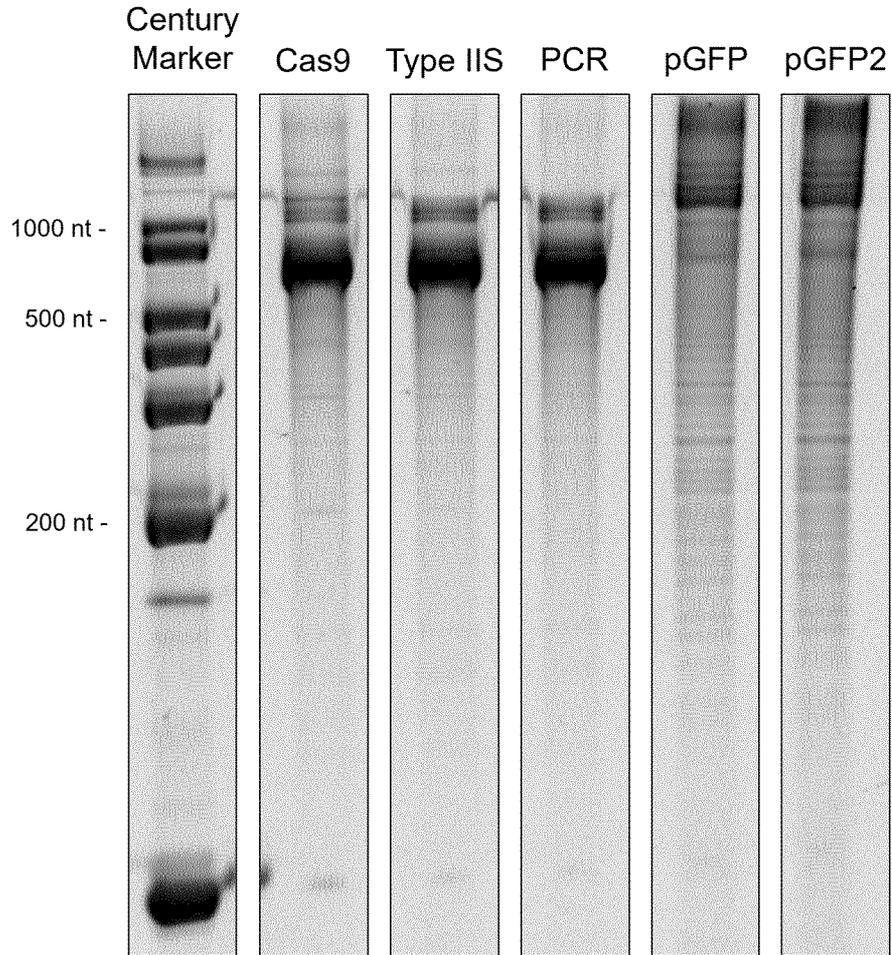
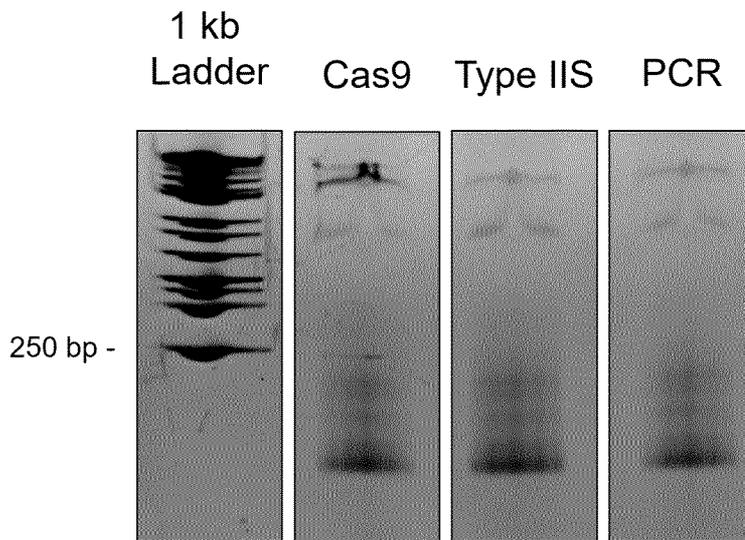


Abbildung 3

5



4/6

Abbildung 4

5

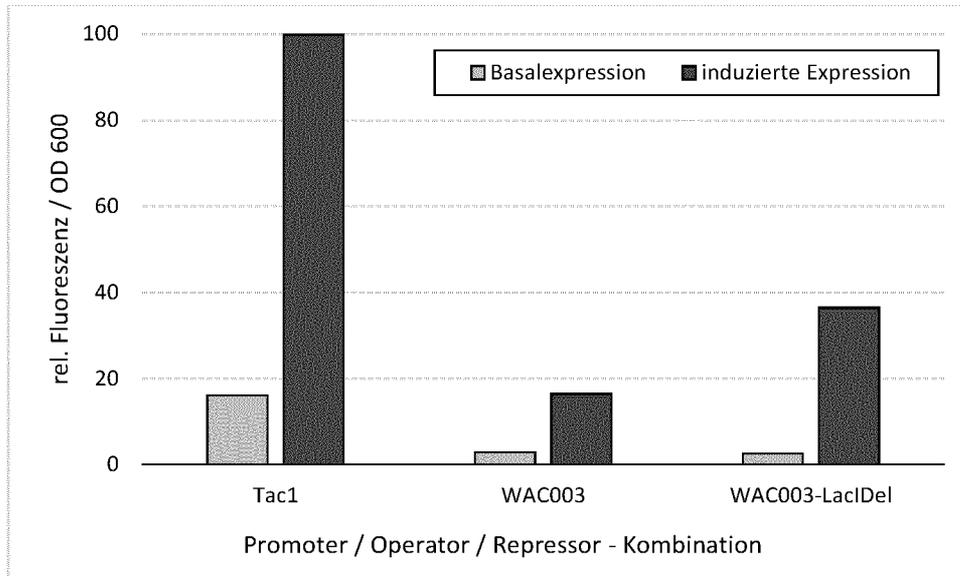


Abbildung 5

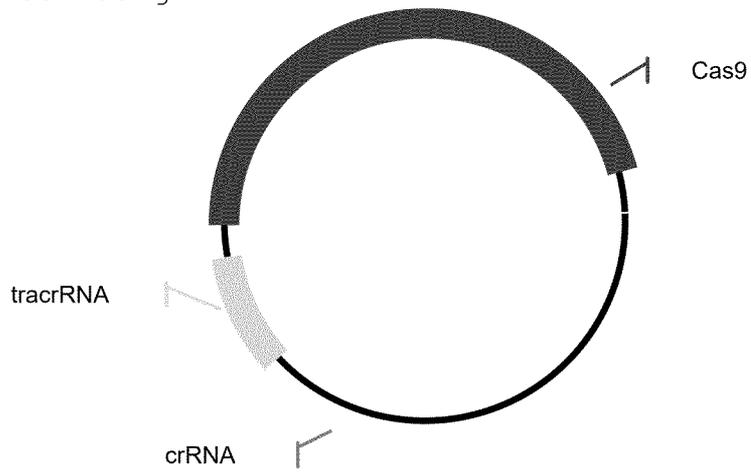
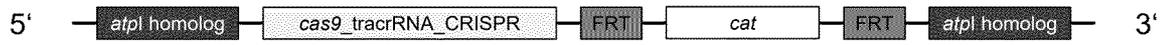


Abbildung 6:

5



INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2023/058156

A. CLASSIFICATION OF SUBJECT MATTER <i>C12N 15/10</i> (2006.01) According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, BIOSIS, EMBASE, FSTA		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 3289077 B1 (CUREVAC AG [DE]) 15 April 2020 (2020-04-15)	1-15
A	EP 3294885 B1 (CUREVAC REAL ESTATE GMBH [DE]) 01 July 2020 (2020-07-01)	1-15
A	CA 3171495 A1 (TRANSLATE BIO INC [US]) 26 August 2021 (2021-08-26)	1-15
A	US 10106800 B2 (BIONTECH AG [DE]) 23 October 2018 (2018-10-23)	1-15
A	US 10760070 B2 (CUREVAC REAL ESTATE GMBH [DE]) 01 September 2020 (2020-09-01)	1-15
A	US 2020392518 A1 (EBERLE FLORIAN [DE] ET AL) 17 December 2020 (2020-12-17)	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 11 October 2023		Date of mailing of the international search report 18 October 2023
Name and mailing address of the ISA/EP European Patent Office p.b. 5818, Patentlaan 2, 2280 HV Rijswijk Netherlands (Kingdom of the) Telephone No. (+31-70)340-2040 Facsimile No. (+31-70)340-3016		Authorized officer Bassias, Ioannis Telephone No.

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. forming part of the international application as filed:
 - in the form of an Annex C/ST.25 text file.
 - on paper or in the form of an image file.
 - b. furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. furnished subsequent to the international filing date for the purposes of international search only:
 - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/EP2023/058156

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
EP	3289077	B1	15 April 2020	CN	108064307	A	22 May 2018
				EP	3289077	A1	07 March 2018
				SG	11201708681S	A	29 November 2017
				US	2018208957	A1	26 July 2018
				US	2022389470	A1	08 December 2022
				WO	2016174227	A1	03 November 2016

EP	3294885	B1	01 July 2020	DK	3294885	T3	10 August 2020
				EP	3294885	A1	21 March 2018
				US	2016326575	A1	10 November 2016
				US	2018282790	A1	04 October 2018
				US	2020308634	A1	01 October 2020
				US	2022154253	A1	19 May 2022
				US	2023053382	A1	23 February 2023
				WO	2016180430	A1	17 November 2016

CA	3171495	A1	26 August 2021	AU	2021224892	A1	13 October 2022
				BR	112022016258	A2	18 October 2022
				CA	3171495	A1	26 August 2021
				CN	115398546	A	25 November 2022
				EP	4107733	A1	28 December 2022
				IL	295669	A	01 October 2022
				JP	2023513836	A	03 April 2023
				KR	20220162125	A	07 December 2022
				WO	2021168052	A1	26 August 2021

US	10106800	B2	23 October 2018	AT	550427	T	15 April 2012
				AU	2006296702	A1	05 April 2007
				CA	2621444	A1	05 April 2007
				CA	3023101	A1	05 April 2007
				CY	1113526	T1	22 June 2016
				CY	1115246	T1	04 January 2017
				DE	102005046490	A1	29 March 2007
				DK	1934345	T3	23 July 2012
				DK	2357230	T3	02 June 2014
				EP	1934345	A2	25 June 2008
				EP	2357230	A1	17 August 2011
				ES	2384113	T3	29 June 2012
				ES	2467678	T3	12 June 2014
				HK	1155477	A1	18 May 2012
				HR	P20120480	T1	31 July 2012
				HR	P20140561	T1	18 July 2014
				JP	5435949	B2	05 March 2014
				JP	6058928	B2	11 January 2017
				JP	2009509516	A	12 March 2009
				JP	2012235786	A	06 December 2012
				JP	2016027811	A	25 February 2016
				PL	1934345	T3	28 September 2012
				PL	2357230	T3	29 August 2014
				PT	1934345	E	21 June 2012
PT	2357230	E	15 July 2014				
RS	52353	B	31 December 2012				
RS	53351	B	31 October 2014				

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/EP2023/058156

Patent document cited in search report				Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)	
				SI	1934345	T1	31 July 2012
				SI	2357230	T1	31 July 2014
				US	2010129877	A1	27 May 2010
				US	2017009244	A1	12 January 2017
				US	2019062762	A1	28 February 2019
				US	2023193296	A1	22 June 2023
				WO	2007036366	A2	05 April 2007
US	10760070	B2	01 September 2020	CN	107873055	A	03 April 2018
				CN	113897350	A	07 January 2022
				DE	22190069	T1	02 March 2023
				DE	202016009003	U1	28 May 2021
				DK	3303583	T3	13 July 2020
				DK	4108769	T3	09 October 2023
				EP	3303583	A1	11 April 2018
				EP	3744843	A1	02 December 2020
				EP	4108769	A1	28 December 2022
				ES	2798271	T3	10 December 2020
				PL	3303583	T3	19 October 2020
				PT	3303583	T	07 July 2020
				US	2018298372	A1	18 October 2018
				US	2021040473	A1	11 February 2021
				US	2022325273	A1	13 October 2022
				US	2023151352	A1	18 May 2023
				US	2023287388	A1	14 September 2023
				WO	2016193206	A1	08 December 2016
US	2020392518	A1	17 December 2020	AU	2015286820	A1	23 February 2017
				AU	2021277776	A1	23 December 2021
				CA	2954706	A1	14 January 2016
				DE	202015010019	U1	02 June 2023
				DE	202015010020	U1	05 June 2023
				DK	3167059	T3	05 August 2019
				EP	3167059	A1	17 May 2017
				EP	3594337	A1	15 January 2020
				ES	2735728	T3	20 December 2019
				HU	E046120	T2	28 February 2020
				JP	6759196	B2	23 September 2020
				JP	7144481	B2	29 September 2022
				JP	2017522050	A	10 August 2017
				JP	2020188758	A	26 November 2020
				PL	3167059	T3	31 December 2019
				PT	3167059	T	29 July 2019
				SI	3167059	T1	30 October 2019
				TR	201910424	T4	21 August 2019
				US	2017166905	A1	15 June 2017
				US	2020392518	A1	17 December 2020
				WO	2016005004	A1	14 January 2016
				WO	2016005324	A1	14 January 2016

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2023/058156

Feld Nr. I Nucleotid- und/oder Aminosäuresequenz(en) (Fortsetzung von Punkt 1 c) auf Blatt 1)

1. Hinsichtlich der **Nucleotid- und/oder Aminosäuresequenz**, die in der internationalen Anmeldung offenbart wurde, ist die internationale Recherche auf der Grundlage eines Sequenzprotokolls durchgeführt worden, das
 - a) im Anmeldezeitpunkt Bestandteil der internationalen Anmeldung war.
 - b) nach dem internationalen Anmeldedatum nur für die Zwecke der internationalen Recherche eingereicht wurde (Regel 13^{ter}.1.a),
 zusammen mit einer Erklärung, wonach das Sequenzprotokoll nicht über den Offenbarungsgehalt der internationalen Anmeldung im Anmeldezeitpunkt hinausgeht.
2. Hinsichtlich der Nucleotid- und/oder Aminosäuresequenz, die in der internationalen Anmeldung offenbart wurde, ist dieser Bericht insoweit erstellt worden, als eine sinnvolle Recherche auch ohne ein dem WIPO-Standard ST.26 entsprechendes Sequenzprotokoll durchgeführt werden konnte.
3. Zusätzliche Bemerkungen:

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2023/058156

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES INV. C12N15/10 ADD.		
Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC		
B. RECHERCHIERTE GEBIETE		
Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) C12N		
Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen		
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) EPO-Internal, BIOSIS, EMBASE, FSTA		
C. ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	EP 3 289 077 B1 (CUREVAC AG [DE]) 15. April 2020 (2020-04-15) -----	1-15
A	EP 3 294 885 B1 (CUREVAC REAL ESTATE GMBH [DE]) 1. Juli 2020 (2020-07-01) -----	1-15
A	CA 3 171 495 A1 (TRANSLATE BIO INC [US]) 26. August 2021 (2021-08-26) -----	1-15
A	US 10 106 800 B2 (BIONTECH AG [DE]) 23. Oktober 2018 (2018-10-23) -----	1-15
A	US 10 760 070 B2 (CUREVAC REAL ESTATE GMBH [DE]) 1. September 2020 (2020-09-01) -----	1-15
A	US 2020/392518 A1 (EBERLE FLORIAN [DE] ET AL) 17. Dezember 2020 (2020-12-17) -----	1-15
<input type="checkbox"/> Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen <input checked="" type="checkbox"/> Siehe Anhang Patentfamilie		
* Besondere Kategorien von angegebenen Veröffentlichungen : "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist "E" frühere Anmeldung oder Patent, die bzw. das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt) "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist "T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist "&" Veröffentlichung, die Mitglied derselben Patentfamilie ist		
Datum des Abschlusses der internationalen Recherche 11. Oktober 2023		Absendedatum des internationalen Recherchenberichts 18/10/2023
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Bevollmächtigter Bediensteter Bassias, Ioannis

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2023/058156

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung				
EP 3289077	B1	15-04-2020	CN 108064307 A	22-05-2018			
			EP 3289077 A1	07-03-2018			
			SG 11201708681S A	29-11-2017			
			US 2018208957 A1	26-07-2018			
			US 2022389470 A1	08-12-2022			
			WO 2016174227 A1	03-11-2016			
			EP 3294885	B1	01-07-2020	DK 3294885 T3	10-08-2020
EP 3294885 A1	21-03-2018						
US 2016326575 A1	10-11-2016						
US 2018282790 A1	04-10-2018						
US 2020308634 A1	01-10-2020						
US 2022154253 A1	19-05-2022						
US 2023053382 A1	23-02-2023						
WO 2016180430 A1	17-11-2016						
CA 3171495	A1	26-08-2021				AU 2021224892 A1	13-10-2022
			BR 112022016258 A2	18-10-2022			
			CA 3171495 A1	26-08-2021			
			CN 115398546 A	25-11-2022			
			EP 4107733 A1	28-12-2022			
			IL 295669 A	01-10-2022			
			JP 2023513836 A	03-04-2023			
			KR 20220162125 A	07-12-2022			
			WO 2021168052 A1	26-08-2021			
			US 10106800	B2	23-10-2018	AT 550427 T	15-04-2012
						AU 2006296702 A1	05-04-2007
CA 2621444 A1	05-04-2007						
CA 3023101 A1	05-04-2007						
CY 1113526 T1	22-06-2016						
CY 1115246 T1	04-01-2017						
DE 102005046490 A1	29-03-2007						
DK 1934345 T3	23-07-2012						
DK 2357230 T3	02-06-2014						
EP 1934345 A2	25-06-2008						
EP 2357230 A1	17-08-2011						
ES 2384113 T3	29-06-2012						
ES 2467678 T3	12-06-2014						
HK 1155477 A1	18-05-2012						
HR P20120480 T1	31-07-2012						
HR P20140561 T1	18-07-2014						
JP 5435949 B2	05-03-2014						
JP 6058928 B2	11-01-2017						
JP 2009509516 A	12-03-2009						
JP 2012235786 A	06-12-2012						
JP 2016027811 A	25-02-2016						
PL 1934345 T3	28-09-2012						
PL 2357230 T3	29-08-2014						
PT 1934345 E	21-06-2012						
PT 2357230 E	15-07-2014						
RS 52353 B	31-12-2012						
RS 53351 B	31-10-2014						
SI 1934345 T1	31-07-2012						
SI 2357230 T1	31-07-2014						
US 2010129877 A1	27-05-2010						
US 2017009244 A1	12-01-2017						
US 2019062762 A1	28-02-2019						

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2023/058156

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
		US 2023193296 A1	22-06-2023
		WO 2007036366 A2	05-04-2007

US 10760070	B2	01-09-2020	
		CN 107873055 A	03-04-2018
		CN 113897350 A	07-01-2022
		DE 22190069 T1	02-03-2023
		DE 202016009003 U1	28-05-2021
		DK 3303583 T3	13-07-2020
		EP 3303583 A1	11-04-2018
		EP 3744843 A1	02-12-2020
		EP 4108769 A1	28-12-2022
		ES 2798271 T3	10-12-2020
		PL 3303583 T3	19-10-2020
		PT 3303583 T	07-07-2020
		US 2018298372 A1	18-10-2018
		US 2021040473 A1	11-02-2021
		US 2022325273 A1	13-10-2022
		US 2023151352 A1	18-05-2023
		US 2023287388 A1	14-09-2023
		WO 2016193206 A1	08-12-2016

US 2020392518	A1	17-12-2020	
		AU 2015286820 A1	23-02-2017
		AU 2021277776 A1	23-12-2021
		CA 2954706 A1	14-01-2016
		DE 202015010019 U1	02-06-2023
		DE 202015010020 U1	05-06-2023
		DK 3167059 T3	05-08-2019
		EP 3167059 A1	17-05-2017
		EP 3594337 A1	15-01-2020
		ES 2735728 T3	20-12-2019
		HU E046120 T2	28-02-2020
		JP 6759196 B2	23-09-2020
		JP 7144481 B2	29-09-2022
		JP 2017522050 A	10-08-2017
		JP 2020188758 A	26-11-2020
		PL 3167059 T3	31-12-2019
		PT 3167059 T	29-07-2019
		SI 3167059 T1	30-10-2019
		TR 201910424 T4	21-08-2019
		US 2017166905 A1	15-06-2017
		US 2020392518 A1	17-12-2020
		WO 2016005004 A1	14-01-2016
		WO 2016005324 A1	14-01-2016
