



(12) Wirtschaftspatent

Erteilt gemäß § 17 Absatz 1 Patentgesetz

(19) DD (11) 263 081 A1

4(51) C 12 N 15/00

AMT FÜR ERFINDUNGS- UND PATENTWESEN

In der vom Anmelder eingereichten Fassung veröffentlicht

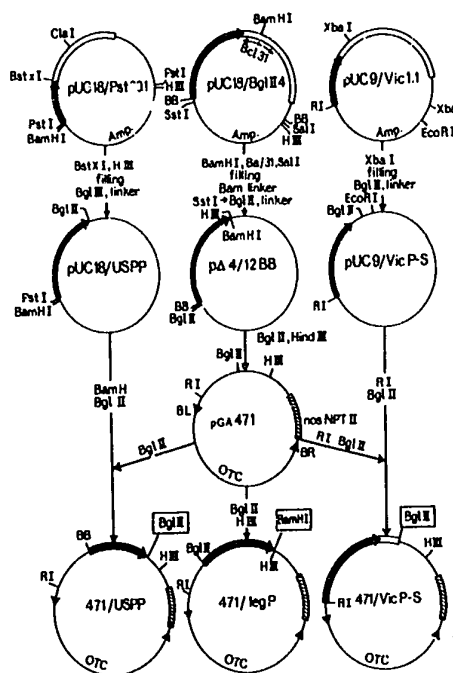
(21) WPC 12 N / 305 136 6 (22) 20.07.87 (44) 21.12.88

(71) Akademie der Wissenschaften der DDR, Otto-Nuschke-Straße 22/23, Berlin, 1080, DD  
 (72) Bäumlein, Helmut, Dr. rer. nat.; Bassüner, Ronald, Dr. rer. nat.; Jung, Rudolf, Dr. rer. nat.; Müntz, Klaus, Prof. Dr. rer. nat. habil.; Saalbach, Gerhard, Dr. rer. nat.; Weschke, Winfriede, Dr. rer. nat.; Wobus, Ulrich, Prof. Dr. sc. nat., DD

(54) Verfahren zum Einführen von DNS-Sequenzen in das Genom von höheren Pflanzen

(55) Transformation, Fremdgen, Pflanze, Expression, Protein, Pflanzensamen

(57) Die Erfindung betrifft das Gebiet der Transformation von Fremdgenen in Pflanzen, vor allem zur Expression bestimmter, ernährungsphysiologisch wichtiger Proteine in Pflanzensamen, speziell von Leguminosen oder Getreiden. Dazu wird die Verwendung einer Expressionskassette vorgesehen, die mehrteilig aus einem Vektorplasmid und einer Deletionsmutante besteht und einen unikalsten Restriktionsort zur Einligierung einer beliebigen, vor allem aber proteinogenen DNS-Sequenz aufweist, die nach der mit bekannten Mitteln durchgeführten Transformation exprimiert wird. Figur



**Patentansprüche:**

1. Verfahren zum Einführen von beliebigen DNS-Sequenzen in das Genom von höheren Pflanzen, vorzugsweise von proteinogenen DNS-Sequenzen für die Expression in Pflanzensamen, wobei
  - zunächst ein vollständiges, beispielsweise pflanzensamenprotein-spezifisches Gen isoliert wird und
  - danach die proteinkodierenden und regulativen DNS-Abschnitte durch Sequenzanalyse ermittelt werden,  
**dadurch gekennzeichnet, daß**
    - die proteinkodierenden DNS-Abschnitte vollständig oder fast vollständig aus dem in einem Vektorplasmid befindlichen Gen herausgeschnitten werden,
    - die regulativen DNS-Abschnitte vollständig oder fast vollständig erhalten bleiben und unter Herbeiführung eines unikalsten Restriktionsortes, der im Wirkungsbereich der regulativen DNS-Abschnitte liegt, zu einer Deletionsmutante vereinigt werden,
    - die Deletionsmutante aus dem Vektorplasmid isoliert und in einen Pflanzen-Transformationsvektor so einligiert wird, daß der unikale Restriktionsort erhalten bleibt und eine Expressionskassette entsteht, in welcher der Pflanzen-Transformationsvektor die Transformation der gesamten Expressionskassette in höhere Pflanzen übernimmt,
    - in den unikalsten Restriktionsort der Expressionskassette eine beliebige DNS-Sequenz einligiert wird und
    - die DNS-Sequenz in Pflanzensamen höherer Pflanzen exprimiert wird.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß die Expressionskassette auf der Basis eines Gens für ein Protein in Pflanzensamen, beispielsweise von Leguminosen oder Getreiden, konstruiert wird, welches aus einer Genbank selektiert wird.
3. Verfahren nach Anspruch 1 und 2, **dadurch gekennzeichnet**, daß das Gen, welches für das in Pflanzensamen gespeicherte Protein Legumin kodiert, aus einer Genbank der Ackerbohne selektiert wird.
4. Verfahren nach Anspruch 1 und 2, **dadurch gekennzeichnet**, daß das Gen, welches für das in Pflanzensamen gespeicherte Protein Vicilin kodiert, aus einer Genbank der Ackerbohne selektiert wird.
5. Verfahren nach Anspruch 1 und 2, **dadurch gekennzeichnet**, daß das Gen, welches für das in Pflanzensamen exprimierte Protein USP kodiert, aus einer Genbank der Ackerbohne selektiert wird.
6. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß aus einer cDNS-Bank der Ackerbohne Klone mit spezifischen Inserts von Legumin-, Vicilin- oder USP-DNS selektiert und durch Sequenzierung deren cDNS-Nukleotidsequenzen ermittelt werden.
7. Verfahren nach Anspruch 1 und 2, einem der Ansprüche 3 bis 5, **dadurch gekennzeichnet**, daß die genomische Sequenz eines Proteins mit der zugehörigen cDNS-Nukleotidsequenz verglichen und daraus bestimmt wird, wo sich die proteinkodierenden, die regulativen DNS-Abschnitte und wo sich Intronsequenzen befinden.
8. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß das in dem Vektorplasmid vorliegende Gen an einem im proteinkodierenden DNS-Abschnitt gelegenen unikalsten Restriktionsort geschnitten wird, daß die beidseitig vom Restriktionsort gelegenen beiden Teile des proteinkodierenden DNS-Abschnittes entfernt, die beiden so entstandenen Enden des Gens stumpf gemacht und über Linker unter Wiederherstellung eines unikalsten Restriktionsortes rezirkularisiert werden.
9. Verfahren nach Anspruch 1 und 8, **dadurch gekennzeichnet**, daß die beidseitig vom unikalsten Restriktionsort gesetzte Deletion den gesamten proteinkodierenden DNS-Abschnitt des Gens erfaßt.
10. Verfahren nach Anspruch 1 und 8, **dadurch gekennzeichnet**, daß die beidseitig vom unikalsten Restriktionsort gesetzte Deletion den proteinkodierenden DNS-Abschnitt ausschließlich der Signalpeptidsequenz und der durch die Signalpeptidase erkennbaren Proteinstruktur betrifft.
11. Verfahren nach Anspruch 1, 8 und 10, **dadurch gekennzeichnet**, daß die Deletion die Signale zum intra- und extrazellulären Transport der Proteine nicht berührt.
12. Verfahren nach Anspruch 1 und 8, **dadurch gekennzeichnet**, daß die im Vektorplasmid vorliegende Deletionsmutante mittels Restriktasen gewonnen und in den durch kompatible Restriktasen geöffneten Pflanzen-Transformationsvektor unter Erhalt der Expressionskassette so einligiert wird, daß der 3'seitig vom Promotor des Gens gelegene unikale Restriktionsort zur Einligierung der DNS-Sequenz erhalten bleibt.

13. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß eine DNS-Sequenz, die für ein Protein mit physiologischer Wirkung im menschlichen Organismus kodiert, aus einer cDNS-Bank selektiert wird.
14. Verfahren nach Anspruch 1 und 13, **dadurch gekennzeichnet**, daß aus einer cDNS-Bank ein Klon selektiert wird, der die vollständige doppelsträngige cDNS-Kopie der mRNS des menschlichen Gewebe-Plasminogen-Aktivators enthält.
15. Verfahren nach Anspruch 1, 13 und 14, **dadurch gekennzeichnet**, daß die DNS-Sequenz des gesamten proteinkodierenden DNS-Abschnittes mittels Restriktasen gewonnen wird.
16. Verfahren nach Anspruch 1, 8 und 12 bis 15, **dadurch gekennzeichnet**, daß die proteinkodierende DNS-Sequenz des menschlichen Gewebe-Plasminogen-Aktivators in den mittels Restriktasen geöffneten unikalen Restriktionsort der Expressionskassette einligiert wird und mit dieser Konstruktion höhere Pflanzen transformiert werden.

Hierzu 1 Seite Zeichnungen

#### Anwendungsgebiet der Erfindung

Die Erfindung betrifft ein Verfahren zum Einführen von beliebigen DNS-Sequenzen in das Genom von höheren Pflanzen nach dem Oberbegriff des Patentanspruches 1.

#### Charakteristik des bekannten Standes der Technik

Es ist seit langem bekannt, speziell für diesen Zweck isolierte Gene in höhere Pflanzen einzuschleusen, um auf diese Weise bestimmte, vererbare Veränderungen im Genom dieser Pflanzen herbeizuführen. Eine Voraussetzung für die Expression der übertragenen Gene ist dabei, daß sie über eine pflanzenspezifische Promotorsequenz verfügen. So ist es beispielsweise bereits gelungen, bestimmte Gene in höheren Pflanzen zu exprimieren, um deren Resistenz gegen Antibiotika und Herbizide zu verbessern. Man verwendet dazu sogenannte konstitutive Promotoren wie den Promotor des Nopalinsynthase-Gens<sup>1/</sup>, den TR-Doppelpromotor<sup>4/</sup> oder den Promotor des 35S-Transkriptes des Blumenkohl-Mosaikvirus<sup>5/</sup>.

Diese Promotoren sind in allen oder mindestens in vielen Geweben der manipulierten Pflanzen aktiv, und das ist nicht in jedem Fall erforderlich oder erwünscht. Es sind vielmehr häufig bessere Ergebnisse zu erwarten, wenn die Promotoren nur dort wirksam sind, wo das Pflanzengewebe tatsächlich verändert werden soll. Besonders vorteilhaft wäre es an sich auch, wenn der Vorgang in das günstigste Entwicklungsstadium der Pflanzen fiel. Die angegebenen bekannten Promotoren vermögen diesen Forderungen aber nicht gerecht zu werden. Aus diesem Grunde sind auch schon gewebe- und entwicklungspezifische Promotoren isoliert worden, wobei man sich entweder sogenannter Promotorsuchplasmide oder des screening in Gen-Banken mit entsprechenden mRNS-Populationen als Hybridisierungsproben bedient; es konnten antheren-, ovarien-, blüten-, blätter-, laubblätter-, stengel-, wurzel- und samenspezifische Gene mit ihren zugehörigen Promotoren isoliert werden<sup>2/</sup>. Davon sind vor allem die samenspezifischen Promotoren wegen der Bedeutung von Pflanzensamen für die Ernährung von großer Wichtigkeit. Als ergiebige Quelle für samenspezifische Regulationssequenzen stehen die Promotoren der von verschiedenen Pflanzenarten bereits klonierten und sequenzierten pflanzensamenproteinspezifischen Gene zur Verfügung<sup>3; 6; 7/</sup>. Es fehlt jedoch bisher ein universell einsetzbares Übertragungsvehikel für Fremdgene zu deren pflanzensamenspezifischer Expression.

#### Ziel der Erfindung

Die Erfindung hat das Ziel, Gene in Pflanzensamen zur Expression zu bringen, deren Expressionsprodukte Proteine sind, welche als solche aus den Pflanzensamen zu gewinnen sind oder ihre physiologische Wirkung bei der ernährungsmäßigen Nutzung der Pflanzensamen entfalten.

#### Darlegung des Wesens der Erfindung

Demzufolge hat sich die Erfindung die Aufgabe gestellt, für Gene der eingangs näher bezeichneten Art, deren proteinkodierende und regulative DNS-Abschnitte durch Sequenzanalyse ermittelt worden sind, ein geeignetes Transportmittel für die Einschleusung in das Genom der Empfängerpflanzen mit nachfolgender samenspezifischer Expression bereitzustellen, das einfach herstell- und universell anwendbar ist.

Erfindungsgemäß wird die Aufgabe dadurch gelöst, daß die proteinkodierenden DNS-Abschnitte vollständig oder fast vollständig aus dem in einem Vektorplasmid befindlichen Gen herausgeschnitten werden, daß die regulativen DNS-Abschnitte vollständig oder fast vollständig erhalten bleiben und unter Herbeiführung eines unikalen Restriktionsortes, der im Wirkungsbereich der regulativen DNS-Abschnitte liegt, zu einer Deletionsmutante vereinigt werden, daß die Deletionsmutante aus dem Vektorplasmid isoliert und in einen Pflanzen-Transformationsvektor so einligiert wird, daß der unikale Restriktionsort erhalten bleibt und eine Expressionskassette entsteht, in welcher der Pflanzen-Transformationsvektor die Transformation der gesamten Expressionskassette in höhere Pflanzen übernimmt, daß in den unikalen Restriktionsort der Expressionskassette eine beliebige DNS-Sequenz einligiert wird und daß die DNS-Sequenz in Pflanzensamen höherer Pflanzen exprimiert wird. Es ist zweckmäßig, daß die Expressionskassette auf der Basis eines Gens für ein Protein in Pflanzensamen, beispielsweise von Leguminosen oder Getreiden, konstruiert wird, welches aus einer Genbank selektiert wird, wobei das Gen, welches für das in Pflanzensamen gespeicherte Protein Legumin<sup>6/</sup> kodiert, aus einer Genbank der Ackerbohne selektiert wird. Es ist auch möglich, daß das Gen, welches für das in Pflanzensamen gespeicherte Protein Vicilin (Anlage 1) kodiert, aus einer

Genbank der Ackerbohne selektiert wird, genauso wie das Gen, welches für das in Pflanzensamen exprimierte Protein USP (Anlage 2) kodiert. Es ist erforderlich, daß aus einer cDNS-Bank der Ackerbohne Klone mit spezifischen Inserts von Legumin-, Vicilin- oder USP-DNS selektiert und durch Sequenzierung deren cDNS-Nukleotidsequenzen ermittelt werden, und daß die genomische Sequenz eines Proteins mit der zugehörigen cDNS-Nukleotidsequenz verglichen und daraus bestimmt wird, wo sich die proteinkodierenden, die regulativen DNS-Abschnitte und wo sich Intronsequenzen befinden. Im einzelnen ist es dann zweckmäßig, daß das in dem Vektorplasmid vorliegende Gen an einem im proteinkodierenden DNS-Abschnitt gelegenen, unikalen Restriktionsort geschnitten wird, daß die beidseitig vom Restriktionsort gelegenen beiden Teile des proteinkodierenden DNS-Abschnittes entfernt, die beiden so entstandenen Enden des Gens stumpf gemacht und über Linker unter Wiederherstellung eines unikalen Restriktionsortes rezirkularisiert werden. Es ist denkbar, daß die beidseitig vom unikalen Restriktionsort gesetzte Deletion den gesamten proteinkodierenden DNS-Abschnitt des Gens erfaßt, aber günstiger, wenn dies ausschließlich der Signalpeptidsequenz und der durch die Signalpeptidase erkennbaren Proteinstruktur erfolgt, und wenn die Deletion die Signale zum intra- und extrazellulären Transport der Proteine nicht berührt. Es ist auch vorteilhaft, wenn die im Vektorplasmid vorliegende Deletionsmutante mittels Restriktasen gewonnen und in den durch kompatible Restriktasen geöffneten Pflanzen-Transformationsvektor unter Erhalt der Expressionskassette so einligiert wird, daß der 3'seitig vom Promotor des Gens gelegene unikale Restriktionsort zur Einligierung der DNS-Sequenz erhalten bleibt. Es ist möglich, daß eine DNS-Sequenz, die für ein Protein mit physiologischer Wirkung im menschlichen Organismus kodiert, aus einer cDNS-Bank selektiert wird bzw. aus dieser cDNS-Bank ein Klon selektiert wird, der die vollständige doppelsträngige cDNS-Kopie der mRNA des menschlichen Gewebe-Plasminogen-Aktivators enthält. Es ist schließlich möglich, daß die DNS-Sequenz des gesamten proteinkodierenden DNS-Abschnittes mittels Restriktasen gewonnen wird und daß die proteinkodierende DNS-Sequenz des menschlichen Gewebe-Plasminogen-Aktivators in den mittels Restriktasen geöffneten unikalen Restriktionsort der Expressionskassette einligiert wird und mit dieser Konstruktion höhere Pflanzen transformiert werden. Entsprechend dem erfindungsgemäßen Verfahren werden beliebige, vor allem proteinkodierende DNS-Sequenzen durch Überführung in eine Expressionskassette in den Zustand der direkten Transformierbarkeit in das Genom höherer Pflanzen versetzt. Ihre Expression erfolgt bevorzugt im Zellgewebe von Pflanzensamen. Die Vielfalt der verfahrensgemäß herstellbaren Expressionskassetten erlaubt eine Wahl nach dem Kriterium des gewünschten intra- oder extrazellulären Ortes der Lokalisation der exprimierten Proteine.

#### Ausführungsbeispiel

Die Erfindung wird nachstehend — zum Teil an Hand einer Zeichnung — an einem Ausführungsbeispiel näher erläutert. Dabei werden drei Varianten der Herstellung einer erfindungsgemäßen Expressionskassette dargestellt und auf der einzigen Figur der Zeichnung verdeutlicht.

##### a) Konstruktion einer Expressionskassette für das Samenprotein USP

Das mit USP bezeichnete, unbekannte Samenprotein USP hat ein Molekulargewicht von 30 kD. Seine Funktion konnte noch nicht ermittelt werden. USP-kodierende mRNA hat einen Anteil von 80% am poly-A<sup>+</sup>-RNS-Gehalt in frühen Stadien sich entwickelnder Samen der Ackerbohne/8/. Durch Translation hybridselektierter mRNA werden USP-spezifische cDNS-Klone identifiziert. Einige dieser cDNS-Klone werden sequenziert und als spezifische Hybridisierungsprobe für die Isolierung entsprechender genomischer Klone aus einer Genbank benutzt/9/. Ein etwa 3,5 kb großes Pst-I-Fragment eines genomischen Klones  $\lambda$  Vf30.1 wird im Sequenzierungsvektor H 13 mp 18 und mittels Sanger-Technik/10/ sequenziert. Durch Vergleich der cDNS-Nukleotidsequenzen mit der genomischen Sequenz können nun Promotorbereich, proteinkodierende und regulative DNS-Abschnitte sowie Intronsequenzen bestimmt werden (Anlage 2). Die Klonierung des 3,5 kb großen Pst-I-Fragmentes in einem Vektorplasmid pUC 18 ergibt das für die Konstruktion einer Expressionskassette verwendbare Plasmid pUC 18/Pst30.1 (Fig.). Die in diesem Plasmid jeweils unikalen Restriktionsorte BstXI am Übergang vom proteinkodierenden zum regulativen DNS-Abschnitt und HindIII (in der Zeichnung teilweise verkürzend mit HIII bezeichnet) im Polylinker des Plasmids werden gespalten. Das restliche Plasmid wird nach Erzeugung glatter Enden mit Hilfe von T<sub>4</sub>-Polymerase und Einfügung von Bgl-II-Linkern zirkularisiert/11/. Es entsteht ein mit pUC 18/USPP bezeichnetes Plasmid (Fig.). Das USP-Promotorfragment liegt in diesem Plasmid als ein Bam-HI-Bgl-II-Fragment vor und wird in ein durch Bgl-II-Spaltung geöffnetes Vektorplasmid pGA471/12/ kloniert. Dabei entsteht auf der einen Seite ein weder durch BglII noch durch BAM HI spaltbarer, mit BB bezeichneter Ort und auf der anderen Seite ein unikalere Bgl-II-Ort. Das so konstruierte Plasmid wird mit 471/USPP bezeichnet und ist die Expressionskassette. Es handelt sich um einen für den Ti-Plasmid-vermittelten Gentransfer geeigneten Binärvektor, der in Agrobakterien-Stämme konjugiert und für die Transformation höherer Pflanzen benutzt wird. Der Einbau einer beliebigen DNS-Sequenz in den unikalen Bgl-II-Ort der Expressionskassette stellt diese unter die Kontrolle des USP-Promotors. Es ist vorteilhaft, die einzubauende DNS-Sequenz mit GATC-Enden zu versehen, d. h., sie sollte als Bam-HI-, Bgl-II-, Bcl-I-, Sma-3A- oder Mbo-I-Fragment vorliegen.

##### b) Konstruktion einer Expressionskassette für das Samenprotein Legumin

Ein Gen LEB4/13; 14/ des Samenproteins Legumin liegt in einem Plasmid pUC 18/BglII4 (Fig.) als ein 4,7 kb großes Bgl-II-Fragment vor. Dieses Plasmid wird am unikalen BAM-HI-Ort geöffnet und partiell mit der Exonuklease Bal31 und anschließend zur Entfernung der 3'seitig vom ursprünglichen BAM-HI-Ort gelegenen proteinkodierenden Sequenzen mit SalI verdaut. Nach der Erzeugung von glatten Enden wird das Plasmid unter Einfügung von BAM-HI-Linkern zirkularisiert. Aus der so entstandenen Plasmid-Schar mit unterschiedlicher Lage des Delektionspunktes wird ein Klon p $\Delta$ 4/12 ausgewählt. Der BAM-HI-Linker liegt bei diesem Klon unmittelbar vor dem AUG-Startkodon/13/, so daß die gesamte 5'seitig nichttranslatierte Sequenz erhalten bleibt. Im so erhalten Plasmid wird zusätzlich der unikalere Sst-I-Ort im Polylinker durch Einfügung eines Bgl-II-Linkers in einen Bgl-II-Ort umgewandelt, wodurch das Promotorfragment als Bgl-II-Hind-III-Fragment isoliert werden kann. Es entsteht ein als p $\Delta$ 4/12 BB bezeichnetes Plasmid (Fig.). Das Bgl-II-Hind-III-Promotorfragment wird nun in den Bgl-II-Hind-III-geöffneten Binärvektor pGA471 wie unter a) beschrieben kloniert. Es entsteht so eine samenspezifische Expressionskassette mit einem unikalen BAM-HI-Ort auf der Basis des Legumin-Promotors analog a).

Aus der Plasmid-Schar kann ferner ein Klon p $\Delta$ 4/4 ausgewählt werden, in dessen Bam-HI-Linker drei Nukleotide hinter dem die Signalsequenz kodierenden Bereich geschnitten wird. Durch Umwandlung des unikaligen Sst-I-Ortes in einen Bgl-II-Ort durch Einfügung entsprechender Linker wird der Abschnitt Promotor/Signalsequenz/kodierender Teil als Bgl-II-Hind-III-Fragment isolierbar und kann in den Bgl-II-Hind-III-geöffneten Binärvektor pGA471 inkloniert werden. Das so erhaltene Plasmid 471/legP ist die Expressionskassette. In ihr stehen die in den Bam-HI-Ort inklonierten DNS-Sequenzen nicht nur unter der Kontrolle des Legumin-Promotors, sondern darüber hinaus die Genprodukte unter der Kontrolle der Legumin-Signalsequenz, die eine entsprechende Kanalisierung in vakuoläre Kompartimente gewährleistet.

**c) Konstruktion einer Expressionskassette für das Samenprotein Vicilin**

Vicilin ist eines der beiden wichtigsten Reserveproteine in den Samen der Ackerbohne. Über isolierte polyA<sup>+</sup>-RNS aus frühen bis mittleren Reifestadien dieser Samen wird doppelsträngige DNS synthetisiert und eine cDNS-Bank eingerichtet, wobei vicilin-spezifische Klone mittels Hybridisierungsselektion und in-vitro-Translation der mRNA identifiziert werden/13/. Diese Klone dienen wie oben als Hybridisierungssprobe für die Isolierung entsprechender genomischer Klone. Ein etwa 3,2 kb großes Eco-RI-Fragment wird in einem Sequenzierungs-Vektor M13mp18 kloniert und sequenziert/11/. Durch Vergleich mit cDNS-Sequenzen/9/ werden proteinkodierende von regulativen DNS-Abschnitten unterscheiden. Das in ein Plasmid pUC9 umklonierte, 3,2 kb große genomische Eco-RI-Fragment ergibt das Plasmid pUC9/Vic1.1 (Fig.). Durch Verdauung mit XbaI wird der Teil des Genes entfernt, der 3'seitig von der Spaltungsstelle der Signalsequenz liegt. Durch Einligieren eines Bgl-II-Linkers wird das Plasmid zirkularisiert zu dem Plasmid pUC9/VicP-S (Fig.). Der Promotorteil des Vicilin-Genes einschließlich des für die Signalsequenz kodierenden Bereiches wird durch Eco-RI-Bgl-II-Verdauung isoliert und in den Eco-RI-Bgl-II-geöffneten Binärvektor pGA471 einligiert: es entsteht als Expressionskassette das Plasmid 471/VicP-S (Fig.). Beim Einbau einer beliebigen DNS-Sequenz in den unikaligen Bgl-II-Ort dieses Plasmids steht diese in transgenischen Pflanzen unter der Kontrolle des Vicilin-Promotors und das Genprodukt unter Kanalisierungskontrolle der Vicilin-Signalsequenz.

**Literatur**

- 1 Herrera-Estrella, L., A. Depicker, M. van Montagu, J. Schell, *Nature (London)* 303 (1983) 209-213.
- 2 Goldberg, R. B., *Phil. Trans. R. Soc. Lond. B* 314 (1986) 343-353.
- 3 Croy, R. R. D., J. A. Gatehouse (1985) *Genetic Engineering of Seed Proteins: current and potential applications*, in: *Plant Genetic Engineering*, ed. J. H. Dodds, Cambridge University Press.
- 4 Velten, J., L. Velten, R. Hani, J. Schull, *EMBO J.* 3 (1984) 2723-2730.
- 5 Koziel, M. G., T. L. Adams, M. A. Hazlet, D. Damm, J. Miller, D. Dahlbeck, S. Jayne, B. J. Slaskawicz, *J. Mol. Appl. Genet.* 2 (1984) 549-562.
- 6 Casey, R., C. Domoney, N. Ellis (1986) *Legume storage proteins and their genes*, *Oxford Surveys of Plant Molecular and Cell Biol.*, im Druck.
- 7 Müntz, K., C. Horstmann, B. Schlesier, *Biol. Zbl.* 105 (1986) 107-120.
- 8 Bassüner, R., R. Manteuffel, K. Müntz, P. Püchel, P. Schmidt, und E. Weber, *Biochem. Physiol. Pflanzen* 178 (1983) 665-684.
- 9 Patentschrift DD 221 756.
- 10 Sanger, F., S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci., USA*, 74 (1977) 5463.
- 11 Maniatis, T., E. F. Fritsch und J. Sambrook, *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory (1982).
- 12 An, G., B. D. Watson, S. Stachel, M. P. Gordon, E. W. Nester, *EMBO J.* 4 (1985) 277-284.
- 13 DD-Patentschrift 240911.
- 14 Wobus, U. H. Bäumlein, R. Bassüner, U. Heim, R. Jung, K. Müntz, G. Saalbach und W. Wescheke, *FEBS Lett.* 201 (1986) 74-80.

10 20 30 40 50 60  
\* \* \* \* \*  
TTAATCCAAAATGATTCTATTAAATTC AATGTTAAAATTTTGCTACAAAATCAGCTGCAA

70 80 90 100 110 120  
\* \* \* \* \*  
GCATGGAAGCTGCCAAGAATCCATTTTCATTTTAGTAATGCATTTTCTCAATTCTATATG

130 140 150 160 170 180  
\* \* \* \* \*  
ATGAGATAGAGGTAAC TATTTAATCTCATAATAGAAGAATTTCTCGTGAGACACAATTTG

190 200 210 220 230 240  
\* \* \* \* \*  
GTGAGTGATGAAGACAAATTTAAATCATCCTCCTATTCTTAAGAGTGTCCAGATTCATC

250 260 270 280 290 300  
\* \* \* \* \*  
AAATTCAAAACAAAAC TCCACCACCCCAATTTATGTTATTTTCATTTTGCCACCTCAATTTT

310 320 330 340 350 360  
\* \* \* \* \*  
GTACATTTCAACACACG TCCATATGCATGGACATCATGACCAATGTTGGTGCATGTC

370 380 390 400 410 420  
\* \* \* \* \*  
ATTACATAATGTTGTTCTATACCTATAAATATCCTTTCACTCAATGTACTTCATTCATC

430 440 450 460 470 480  
\* \* \* \* \*  
CACATCGAGTAAAGTTAGACTTTAATTC A A A T T L K D  
AATCAACATGGCAGCTACCACATTGAAAGAT

490 500 510 520 530 540  
\* \* \* \* \*  
TCATTTCCACTTTT GACGTTGTTGGGAATCGCTTTTCTAGCTTCAGTGTGTCTTTCTTCT  
S F P L L T L L G I A F L A S V C L S S

550 560 570 580 590 600  
\* \* \* \* \*  
AGATCCGATCAAGATAACCCCTTTTGTCTTTGAGTCTAACAGGTTTCAAAC TCTTTTGGAG  
R S D Q D N P F V F E S N R F Q T L F E

610 620 630 640 650 660  
\* \* \* \* \*  
AATGAAAACGGTCACAT TCGTCTTCTCCAAAATTTGATCAACATTCCAAACTACTTGAG  
N E N G H I R L L Q K F D Q H S K L L E

670 680 690 700 710 720  
\* \* \* \* \*  
AATCTCCAAAATTACCGTCTTTTGG AATATAAATCTAAACCTCACACCATATTTCTTCCA  
N L Q N Y R L L E Y K S K P H T I F L F

730 740 750 760 770 780  
\* \* \* \* \*  
CAACAAACTGACGCGGATTTTCATCCTTGT TGTCTTAGTGGTAATTAATTATTTATCAAG  
Q Q T D A D F I L V V L S G

790 800 810 820 830 840  
\* \* \* \* \*  
TTAAATATTAAAAACATATTATAAATAGATATTTAAAAAGAATGGTATTTCTCGCTTTGCA

850 860 870 880 890 900  
\* \* \* \* \*  
TGTTTGTTTTGTAAATGATAAATTAAGTTCAATTTTTAAATCTAATGTCATTTGTAATTT

910 920 930 940 950 960  
\* \* \* \* \*  
GCAGGAAAAGCTATACTCACAGTGTATTACCCCAACGACAGAAACTCCTTCAGTCTTGAG  
K A I L T V L L P N D R N S F S L E

970 980 990 1000 1010 1020  
\* \* \* \* \*  
CGAGGAGATACDATAAAAACCTCCTGCTGGTACAATTGGTTATTTAGTTAACCGAGATGAC  
R G D T I K L P A G T I G Y L V N R D D

1030 1040 1050 1060 1070 1080  
\* \* \* \* \*  
GAGGAGGATCTTAGAGTACTAGATCTTGTTCATCCCCGTAATAGACCAGGTGAACCTCAG  
E E D L R V L D L V I P V N R P G E P Q

1090 1100 1110 1120 1130 1140  
\* \* \* \* \*  
GTAATATATCCAATAATACTTTCTAATATCAACTAATGTATGCATGCAATTGCATAAAAT  
V

1150 1160 1170 1180 1190 1200  
\* \* \* \* \*  
AGATTTTGGAGATCTTTTCATAACATCACACAACAATTATTACTATGTTTATCCACATT

1210 1220 1230 1240 1250 1260  
\* \* \* \* \*  
TTTCTACAATTTTTTCTATTGGATTCTTATAATTTTTTTTCTTTCAACAGTCTTTCTTA  
F L

1270 1280 1290 1300 1310 1320  
\* \* \* \* \*  
TTGTCTGGTAATCAAACCAACCGTCCATCTTATCTGGATTCAAGCAAGAACATTTTAGAG  
L S G N Q N Q P S I L S G F S K N I L E

1330 1340 1350 1360 1370 1380  
\* \* \* \* \*  
GCTTCTTCAATGTAAGCATAAAACAACACTATTTCTTCATTTTTTCATGTATTAGTTTCAT  
A S F N V

1390 1400 1410 1420 1430 1440  
\* \* \* \* \*  
T TACTCTTTGTAATATTCACTTTGTCCTTGTATATATTGTAATATTGTAGACCGATTAT  
D Y

1450 1460 1470 1480 1490 1500  
\* \* \* \* \*  
AAAGAGATAGAGAAGGTTCTCTTAGAAGAGCATGGGAAAGAGAAATATCACAGGAGAGGC  
K E I E K V L L E E H G K E K Y H R R G





```
3010      3020      3030      3040      3050      3060
*         *         *         *         *         *
AAGAAGATGATTGTTACAATCGTCACCATTACTAGTGGAGGTTGTCTAGACCCGTTGCACT

3070      3080      3090      3100      3110      3120
*         *         *         *         *         *
ATATATTACTCGAGATTGGGGCCAAAGATTTTTCAAAGGTGGGTTAGAATAGGCTTCTGA

3130      3140      3150      3160      3170      3180
*         *         *         *         *         *
CACACAGATTTTGGCAATTGTTGTTCAAATTGTTGCAACCCCAATAATAGCGATAATTAGAA

3190      3200      3210
*         *         *
CCGCCACAACACTCATACAAAACCTCTCACCAAAA
```

Anlage 1

```

      10      20      30      40      50      60
      *      *      *      *      *      *
CTGCAGCAAATTTACACATTGCCACTAAACGCTAAACCCCTTGTAAATTTGTTTTGTTTTAC

      70      80      90      100     110     120
      *      *      *      *      *      *
TATGTGTGTTATGTATTTGATTTGCGATAAATTCTATATTTGGTACTAAATTTATAACA

      130     140     150     160     170     180
      *      *      *      *      *      *
CCTTTTATGCTAACGTTTGCCAACACTTAGCAATTTGCAAGTTGATTAATTGATTCTAAA

      190     200     210     220     230     240
      *      *      *      *      *      *
TTATTTTTGTCTTCTAAATACATATACCTAATCAACTGGAAATGTAAATATTTGCTAATA

      250     260     270     280     290     300
      *      *      *      *      *      *
TTTCTACTATAGGAGAATTAAGTGAGTGAATATGGTACCACAAGGTTTGGAGATTTAAT

      310     320     330     340     350     360
      *      *      *      *      *      *
TGTTCGAATGCTGCATGGATGGCATATACCCAAACATCAATAATCTGAGGATAATAATG

      370     380     390     400     410     420
      *      *      *      *      *      *
GTACCACACAAGATTGAGTGCATGAACGTCACGTGGACAAAAGGTTTAGTAATTTTCAAG

      430     440     450     460     470     480
      *      *      *      *      *      *
ACAACAATGTACCATATATAAGTTTAGTGCATGCATGGATGCCCTGTGGAAAGTTTAAA

      490     500     510     520     530     540
      *      *      *      *      *      *
AAATATTTTTGGAAATGATTTGCTAGGAAGGCCTATGTGTAAAACCATGACATCCACTTG

      550     560     570     580     590     600
      *      *      *      *      *      *
GAGGATGCAATAATGAAGAAAACCTACAAATTTACATGCAACTAGTTATGCATGTAGTCTA

      610     620     630     640     650     660
      *      *      *      *      *      *
TATHATGAGGATTTTGCAACTACTTTCATTCATACACACTCACTAAGTTTTACACGATTAT

      670     680     690     700     710     720
      *      *      *      *      *      *
AATTTCTTCATAGCCAGTCAAATGGAGTTTGCACATCTCACTGTTTTATCTCTCTTTTGC
      M E F A H L T V L S L F C

      730     740     750     760     770     780
      *      *      *      *      *      *
GTAGCTAAATATTTTCTAATCTTCAAGATTTAACGTATGCATGGTTTTTATTTCAAACG

      790     800     810     820     830     840
      *      *      *      *      *      *
AGTCATGAATTTATTTTGCAGCTGGCTTTTGGTGGGAATCACTGCGACATCATCCGGAGAA
      L A F V G I T A T S S G E

```



1570      1580      1590      1600      1610      1620  
\*           \*           \*           \*           \*           \*  
TTGCTTTATGAAGCACTCGAAGTTACTCCTGGAAGTGTCCCTGTATGTCATTTTCATTGGC  
L L Y E A L E V T P G T V P V C H F I G

1630      1640      1650      1660      1670      1680  
\*           \*           \*           \*           \*           \*  
AATAAAGCCCGCTGCTTGGGTACCAAATCATACCGCAGACAATCTTTGTGTCATGTAATG  
N K A A A W V P N H T A D N L C V M -

1690      1700      1710      1720      1730      1740  
\*           \*           \*           \*           \*           \*  
TTAAATGGTTGTGCACTCATCTCTCATTCAACAATTTTCATGTTTGTGATTTGAGGATTTT

1750      1760      1770      1780      1790      1800  
\*           \*           \*           \*           \*           \*  
CATCTTCATTTGACTCTCATGTTTTGCGGGAGGAATGTGATGTGCAAGAACCGGATCATA

1810      1820      1830      1840      1850      1860  
\*           \*           \*           \*           \*           \*  
TGTATATTAATAAAACTAGCTAGCACTATGGGTGTATCCCAGTATCTAATAGTATCCTA

1870      1880      1890      1900      1910      1920  
\*           \*           \*           \*           \*           \*  
GTATCGTATGCCAGTACATATGAATGTATGCTTGAATTTATAATTAAAAAAAATGTT

1930      1940      1950      1960      1970      1980  
\*           \*           \*           \*           \*           \*  
GTGTTTGAATTAATATATTGATGATATATGTGCCTTAAAAGATTACAACATGTTAGGTA

1990      2000      2010      2020      2030      2040  
\*           \*           \*           \*           \*           \*  
CATTTTCATACTTATTAAATTATAATTGGTTGGAAAAGTTTIGATAAGAGCTGTAACACTT

2050      2060      2070      2080      2090      2100  
\*           \*           \*           \*           \*           \*  
ATTAATATATATCATCTTTTCAGATAGGGGTGTGCAAAAAACCAATAATCCATAAACCAA

2110      2120      2130      2140      2150      2160  
\*           \*           \*           \*           \*           \*  
ATTAATATCCAAATTAATCCATTTTTTCAAACCCAAACCCATCATTACATTTTAAACCCAG

2170      2180      2190      2200      2210      2220  
\*           \*           \*           \*           \*           \*  
TGACCAATTTATCAAACCAAATAAAAACCAAAAACDATTAAATCAATTCTGGATCAACCT

2230      2240      2250      2260      2270      2280  
\*           \*           \*           \*           \*           \*  
ATAGGATACCACATTTGAAACTCTCTCTTCGGTGTGTCAACAATCACAATGTAATTGCTA

2290      2300      2310      2320      2330      2340  
\*           \*           \*           \*           \*           \*  
AACCCATAATTTCCCAAATCCTATCTTCGAAGATTTTGTGTCTCGCAATCCCAAATCTTG

2350      2360      2370      2380      2390      2400  
\*           \*           \*           \*           \*           \*  
AGTGAAGTTCAGGTAAGAAACATGAAAGAAATTTGTGTGAGTGGTGAAGAAAAGGAG

2410        2420        2430        2440        2450        2460  
\*           \*           \*           \*           \*           \*  
CAGAGATTGATAAAGATGAAGACAAACTCGCTTTGCAATTTATGGATTCACCTTCATAACT

2470        2480        2490        2500        2510        2520  
\*           \*           \*           \*           \*           \*  
ATCTATCCTTCACTAATTCCTCTGTCTTATACNCTTTGACAAGTAAAAAATTACAGTTTAA

2530        2540        2550        2560        2570        2580  
\*           \*           \*           \*           \*           \*  
CCTATTCAATTCAAATTTATATTATCATCTTCATCATAGAATAATACCCTTCAATAATTIC

2590        2600        2610        2620        2630        2640  
\*           \*           \*           \*           \*           \*  
TGTATTCACTCATTGCATCTACTTTGATTCCGGTATACTCTAGGATCATGATTTATTGTAC

2650        2660        2670        2680        2690        2700  
\*           \*           \*           \*           \*           \*  
TTTCTTTCAAATCTATACGTTGGGGTTAATAGGGATTATGGAAACCATTCTCTGATCATGT

2710        2720        2730        2740        2750        2760  
\*           \*           \*           \*           \*           \*  
TGTTATGTATTGGTCATTCAAATGGGATTTTCATGTAATATGAGATGAAAGGTTGTGAACCT

2770        2780        2790        2800        2810        2820  
\*           \*           \*           \*           \*           \*  
TGTAGTATATGTAAGGATCTTTTGGATAATAGCTCAAAGATAGAATGAAGTTAACTTTTA

2830        2840        2850        2860        2870        2880  
\*           \*           \*           \*           \*           \*  
TAAACTGCTCGAATGATTTTTAATTAAACTGGTTTAGAAAAATCAGTTTAAATTTGTT

2890        2900        2910        2920        2930        2940  
\*           \*           \*           \*           \*           \*  
TTTTTTAAACTGGTTTTATGAGAAACCAATTCAAAACCGAACCGATCCCCATGTAACCG

2950        2960        2970        2980        2990        3000  
\*           \*           \*           \*           \*           \*  
GTTTTAGAAATATACACTGGTTTTAATGAAGTAGTTTTAAGATCCAAACCAATCATATA

3010        3020        3030        3040        3050        3060  
\*           \*           \*           \*           \*           \*  
TATGGTTC AATTGGTTGGTTATTGATCCATGCACACACCTACTTCAGATACTCCACACAG

3070        3080        3090        3100        3110        3120  
\*           \*           \*           \*           \*           \*  
TGTTAATCAAATAACCGCTACGGGTATCCTCGTTATCTTGTTCCTTCACCTTTACCCAAAT

3130        3140        3150        3160        3170        3180  
\*           \*           \*           \*           \*           \*  
AAAGAATAAGCTTCCAACCTCTGGCAAAGATAACTCCAATCACCACCAATCCAAGTATAT

3190        3200        3210        3220        3230        3240  
\*           \*           \*           \*           \*           \*  
ACCTTCTTCCAAATACTATCCAAAATCGCGCATTATCGGAAAAGATGACGTAAACATTCC

3250        3260        3270        3280        3290        3300  
\*           \*           \*           \*           \*           \*  
TCATTTCGTACTACAAGAAGACCGCAAGATCTATCCGATAAAACCCCTCTNNNNNNNNNTTT

3310      3320      3330      3340      3350      3360  
\*           \*           \*           \*           \*           \*  
AATCTCCTAAGATGTTTTATTAATGCACTGCGTTACTGTCAAACGGGATAGCATAATAT

3370      3380      3390      3400      3410      3420  
\*           \*           \*           \*           \*           \*  
TCTCTGAAAAGCTCATAGAAAGATTTGACCCAAHAATACCTTTGTGGCCAAACCTCAAG

3430      3440      3450      3460  
\*           \*           \*           \*  
CTGAAACTGTTATTGGCACTACTTGCTAGCGACTTCCCTGCAG

Anlage 2

