Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2443865 C 2012/02/07

(11)(21) 2 443 865

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2002/05/03
(87) Date publication PCT/PCT Publication Date: 2002/11/14
(45) Date de délivrance/Issue Date: 2012/02/07
(85) Entrée phase nationale/National Entry: 2003/10/07
(86) N° demande PCT/PCT Application No.: US 2002/013898
(87) N° publication PCT/PCT Publication No.: 2002/089561
(30) Priorité/Priority: 2001/05/09 (US60/289,527)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01),
A01H 3/00 (2006.01), *A01H 5/00* (2006.01),
A23J 1/00 (2006.01), *C11B 1/00* (2006.01),
C12N 15/11 (2006.01), *C12N 15/31* (2006.01),
C12N 15/53 (2006.01), *C12N 15/54* (2006.01),
C12N 5/10 (2006.01), *C12Q 1/68* (2006.01)

(72) Inventeurs/Inventors:
VALENTIN, HENRY E., US;
MITSKY, TIMOTHY A., US;
HAO, MING, US;
KARUNANANDAA, BALASULOJINI, US;
QI, QUNGANG, US

(73) Propriétaire/Owner:

(54) Titre : GENES TyrA ET UTILISATIONS ASSOCIEES
(54) Title: TyrA GENES AND USES THEREOF

(57) Abrégé/Abstract:
The present invention is in the field of plant genetics and biochemistry. More specifically, the invention relates to genes associated with the tocopherol biosynthesis pathway. The present invention provides and includes nucleic acid molecules, proteins, and antibodies associated with the genes of the tocopherol biosynthesis pathway. The present invention also provides methods for utilizing such agents, for example in gene isolation, gene analysis and the production of transgenic plants. Moreover, the present invention includes transgenic plants modified to express proteins associated with the tocopherol pathway. In addition, the present invention includes methods for the production of products from the tocopherol biosynthesis pathway.

CA 2443865 C 2012/02/07

(11)(21) **2 443 865**

(13) **C**

(73) **Propriétaires(suite)/Owners(continued)**:MONSANTO TECHNOLOGY LLC, US

(74) **Agent**: MCFADDEN, FINCHAM

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 November 2002 (14.11.2002)

PCT

(10) International Publication Number
WO 02/089561 A1

(51) International Patent Classification[7]: A01H 13/00, 5/00, C12N 5/04, 5/10, 15/31, 15/53, 15/54, 15/74, 15/82

(21) International Application Number: PCT/US02/13898

(22) International Filing Date: 3 May 2002 (03.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/289,527 9 May 2001 (09.05.2001) US

(71) Applicant: **MONSANTO TECHNOLOGY LLC.** [US/US]; 800 N. Lindbergh Boulevard, St. Louis, MO 63167 (US).

(72) Inventors: **VALENTIN, Henry, E.**; 873 M Foxsprings Drive, Chesterfield, MO 63017 (US). **MITSKY, Timothy, A.**; 2262 A Rule Avenue, Maryland Heights, MO 63043 (US).

(74) Agents: **MARSH, David, R.** et al.; ARNOLD & PORTER, 555 Twelfth Street, N.W., Washington, DC 20004 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/089561 A1

(54) Title: TYRA GENES AND USES THEREOF

(57) Abstract: The present invention is in the field of plant genetics and biochemistry. More specifically, the invention relates to genes associated with the tocopherol biosynthesis pathway. The present invention provides and includes nucleic acid molecules, proteins, and antibodies associated with the genes of the tocopherol biosynthesis pathway. The present invention also provides methods for utilizing such agents, for example in gene isolation, gene analysis and the production of transgenic plants. Moreover, the present invention includes transgenic plants modified to express proteins associated with the tocopherol pathway. In addition, the present invention includes methods for the production of products from the tocopherol biosynthesis pathway.

TyrA GENES AND USES THEREOF

FIELD OF THE INVENTION

The present invention is in the field of plant genetics and biochemistry. More specifically, the invention relates to genes associated with the tocopherol biosynthesis pathway. The present invention provides and includes nucleic acid molecules, proteins, and antibodies associated with the genes of the tocopherol biosynthesis pathway. The present invention also provides methods for utilizing such agents, for example in gene isolation, gene analysis and the production of transgenic plants. Moreover, the present invention includes transgenic plants modified to express proteins associated with the tocopherol pathway. In addition, the present invention includes methods for the production of products from the tocopherol biosynthesis pathway.

BACKGROUND OF THE INVENTION

Tocopherols are an essential component of mammalian diets. Epidemiological evidence indicates that tocopherol supplementation can result in decreased risk for cardiovascular disease and cancer, can aid in immune function, and is associated with prevention or retardation of a number of degenerative disease processes in humans (Traber and Sies, *Annu. Rev. Nutr. 16*:321-347 (1996)). Tocopherol functions, in part, by stabilizing the lipid bilayer of biological membranes (Skrypin and Kagan, *Biochim.*

Biophys. Acta 815:209 (1995); Kagan, *N.Y. Acad. Sci.* p 121, (1989); Gomez-Fernandez *et al.*, *Ann. N.Y. Acad. Sci.* p 109 (1989)), reducing polyunsaturated fatty acid (PUFA) free radicals generated by lipid oxidation (Fukuzawa *et al.*, *Lipids 17*: 511-513 (1982)), and scavenging oxygen free radicals, lipid peroxy radicals and singlet oxygen species (Diplock *et al. Ann. N Y Acad. Sci.* 570: 72 (1989); Fryer, *Plant Cell Environ. 15(4)*:381-392 (1992)).

α-Tocopherol, often referred to as vitamin E, belongs to a class of lipid-soluble antioxidants that includes α, β, γ, and δ-tocopherols and α, β, γ, and δ-tocotrienols. Although α, β, γ, and δ-tocopherols and α, β, γ, and δ-tocotrienols are sometimes referred to collectively as "vitamin E", vitamin E is more appropriately defined chemically as α-tocopherol. α-Tocopherol is significant for human health, in part because it is readily absorbed and retained by the body, and has a higher degree of bioactivity than other tocopherol species (Traber and Sies, *Annu. Rev. Nutr. 16*:321-347 (1996)). However, other tocopherols such as β, γ, and δ-tocopherols, also have significant health and nutritional benefits.

Tocopherols are primarily synthesized only by plants and certain other photosynthetic organisms, including cyanobacteria. As a result, mammalian dietary tocopherols are obtained almost exclusively from these sources. Plant tissues vary considerably in total tocopherol content and tocopherol composition, with α-tocopherol the predominant tocopherol species found in green, photosynthetic plant tissues. Leaf tissue can contain from 10-50 μg of total tocopherols per gram fresh weight, but most of the world's major staple crops (*e.g.*, rice, maize, wheat, potato) produce low to extremely low levels of total tocopherols, of which only a small percentage is α-tocopherol (Hess,

Vitamin E, α-tocopherol, *In* Antioxidants in Higher Plants, R. Alscher and J. Hess, Eds., CRC Press, Boca Raton. pp. 111-134 (1993)). Oil seed crops generally contain much higher levels of total tocopherols, but α-tocopherol is present only as a minor component (Taylor and Barnes, *Chemy Ind., Oct.*:722-726 (1981)).

The recommended daily dietary intake of 15-30 mg of vitamin E is quite difficult to achieve from the average American diet. For example, it would take over 750 grams of spinach leaves in which α-tocopherol comprises 60% of total tocopherols, or 200-400 grams of soybean oil to satisfy this recommended daily vitamin E intake. While it is possible to augment the diet with supplements, most of these supplements contain primarily synthetic vitamin E, having six stereoisomers, whereas natural vitamin E is predominantly composed of only a single isomer. Furthermore, supplements tend to be relatively expensive, and the general population is disinclined to take vitamin supplements on a regular basis. Therefore, there is a need in the art for compositions and methods that either increase the total tocopherol production or increase the relative percentage of α-tocopherol produced by plants.

In addition to the health benefits of tocopherols, increased α-tocopherol levels in crops have been associated with enhanced stability and extended shelf life of plant products (Peterson, *Cereal-Chem. 72(1)*:21-24 (1995); Ball, *Fat-soluble vitamin assays in food analysis. A comprehensive review,* London, Elsevier Science Publishers Ltd. (1988)). Further, tocopherol supplementation of swine, beef, and poultry feeds has been shown to significantly increase meat quality and extend the shelf life of post-processed meat products by retarding post-processing lipid oxidation, which contributes to the

undesirable flavor components (Sante and Lacourt, *J. Sci. Food Agric. 65(4)*:503-507 (1994); Buckley *et al.*, *J. of Animal Science 73*:3122-3130 (1995)).

TOCOPHEROL BIOSYNTHESIS

The plastids of higher plants exhibit interconnected biochemical pathways leading to secondary metabolites including tocopherols. The tocopherol biosynthetic pathway in higher plants involves condensation of homogentisic acid and phytylpyrophosphate to form 2-methyl-6 phytylplastoquinol (Fiedler *et al.*, *Planta 155*: 511-515 (1982); Soll *et al.*, *Arch. Biochem. Biophys. 204*: 544-550 (1980); Marshall *et al.*, *Phytochem. 24*: 1705-1711 (1985)). This plant tocopherol pathway can be divided into four parts: 1) synthesis of homogentisic acid, which contributes to the aromatic ring of tocopherol; 2) synthesis of phytylpyrophosphate, which contributes to the side chain of tocopherol; 3) cyclization, which plays a role in chirality and chromanol substructure of the vitamin E family; 4) and S-adenosyl methionine dependent methylation of an aromatic ring, which affects the relative abundance of each of the tocopherol species.

SYNTHESIS OF HOMOGENTISIC ACID

Homogentisic acid is the common precursor to both tocopherols and plastoquinones. In at least some bacteria the synthesis of homogentesic acid is reported to occurs via the conversion of chorismate to prephenate and then to p-hydroxyphenyl-pyruvate via a bifunctional prephenate dehydrogenase. Examples of bifunctional bacterial prephenate dehydrogenase enzymes include the proteins encoded by the *tyrA* genes of *Erwinia herbicola* and *Escherichia coli*. The *tyrA* gene product catalyzes the production of prephenate from chorismate, as well as the subsequent dehydrogenation of prephenate to form p-hydroxyphenylpyruvate (p-HPP), the immediate precursor to

homogentisic acid. p-HPP is then converted to homogentisic acid by hydroxyphenyl-pyruvate dioxygenase (HPPD). In contrast, plants are believed to lack prephenate dehydrogenase activity, and it is generally believed that the synthesis of homogentesic acid from chorismate occurs via the synthesis and conversion of the intermediate arogenate. Since pathways involved in homogentesic acid synthesis are also responsible for tyrosine formation, any alterations in these pathways can also result in the alteration in tyrosine synthesis and the synthesis of other aromatic amino acids.

SYNTHESIS OF PHYTYLPYROPHOSPHATE

Tocopherols are a member of the class of compounds referred to as the isoprenoids. Other isoprenoids include carotenoids, gibberellins, terpenes, chlorophyll and abscisic acid. A central intermediate in the production of isoprenoids is isopentenyl diphosphate (IPP). Cytoplasmic and plastid-based pathways to generate IPP have been reported. The cytoplasmic based pathway involves the enzymes acetoacetyl CoA thiolase, HMGCoA synthase, HMGCoA reductase, mevalonate kinase, phosphomevalonate kinase, and mevalonate pyrophosphate decarboxylase.

Recently, evidence for the existence of an alternative, plastid based, isoprenoid biosynthetic pathway emerged from studies in the research groups of Rohmer and Arigoni (Eisenreich *et al.*, *Chem. Bio.*, *5*:R221-R233 (1998); Rohmer, *Prog. Drug. Res.*, *50*:135-154 (1998); Rohmer, *Comprehensive Natural Products Chemistry*, Vol. 2, pp. 45-68, Barton and Nakanishi (eds.), Pergamon Press, Oxford, England (1999)), who found that the isotope labeling patterns observed in studies on certain eubacterial and plant terpenoids could not be explained in terms of the mevalonate pathway. Arigoni and coworkers subsequently showed that 1-deoxyxylulose, or a derivative thereof, serves as

an intermediate of the novel pathway, now referred to as the MEP pathway (Rohmer *et al.*, *Biochem. J.*, *295*:517-524 (1993); Schwarz, Ph.D. thesis, Eidgenössiche Technische Hochschule, Zurich, Switzerland (1994)). Recent studies showed the formation of 1-deoxyxylulose 5-phosphate (Broers, Ph.D. thesis (Eidgenössiche Technische Hochschule, Zurich, Switzerland) (1994)) from one molecule each of glyceraldehyde 3-phosphate (Rohmer, *Comprehensive Natural Products Chemistry*, Vol. 2, pp. 45-68, Barton and Nakanishi, eds., Pergamon Press, Oxford, England (1999)) and pyruvate (Eisenreich *et al.*, *Chem. Biol.*, *5*:R223-R233 (1998); Schwarz *supra*; Rohmer *et al.*, *J. Am. Chem. Soc.*, *118*:2564-2566 (1996); and Sprenger *et al.*, *Proc. Natl. Acad. Sci. USA, 94*:12857-12862 (1997)) by an enzyme encoded by the *dxs* gene (Lois *et al.*, *Proc. Natl. Acad. Sci. USA, 95*:2105-2110 (1997); and Lange *et al.*, *Proc. Natl. Acad. Sci. USA, 95*:2100-2104 (1998)). 1-Deoxyxylulose 5-phosphate can be further converted into 2-*C*-methylerythritol 4-phosphate (Arigoni *et al.*, *Proc. Natl. Acad. Sci. USA, 94*:10600-10605 (1997)) by a reductoisomerase catalyzed by the *dxr* gene (Bouvier *et al.*, *Plant Physiol, 117*:1421-1431 (1998); and Rohdich *et al.*, *Proc. Natl. Acad. Sci. USA, 96*:11758-11763 (1999)).

Reported genes in the MEP pathway also include *ygbP,* which catalyzes the conversion of 2-C-methylerythritol 4-phosphate into its respective cytidyl pyrophosphate derivative and *ygbB,* which catalyzes the conversion of 4-phosphocytidyl-2C-methyl-D-erythritol into 2C-methyl-D-erythritol, 3, 4-cyclophosphate. These genes are tightly linked on the *E. coli* genome (Herz *et al.*, *Proc. Natl. Acad. Sci. U.S.A., 97(6)*:2485-2490 (2000)).

Once IPP is formed by the MEP pathway, it is converted to GGDP by GGDP synthase, and then to phytylpyrophosphate, which is the central constituent of the tocopherol side chain.

COMBINATION AND CYCLIZATION

Homogentisic acid is combined with either phytyl-pyrophosphate or solanyl-pyrophosphate by phytyl/prenyl transferase forming 2-methyl-6-phytyl plastoquinol or 2-methyl-6-solanyl plastoquinol respectively. 2-methyl-6-solanyl plastoquinol is a precursor to the biosynthesis of plastoquinones, while 2-methyl-6-phytyl plastoquinol is ultimately converted to tocopherol.

METHYLATION OF THE AROMATIC RING

The major structural difference between each of the tocopherol subtypes is the position of the methyl groups around the phenyl ring. Both 2-methyl-6-phytyl plastoquinol and 2-methyl-6-solanyl plastoquinol serve as substrates for 2-methyl-6-phytylplatoquinol/2-methyl-6-solanylplastoquinol-9 methyltransferase (Methyl Transferase 1 or MT1), which catalyzes the formation of plastoquinol-9 and γ-tocopherol respectively, by methylation of the 7 position. Subsequent methylation at the 5 position of γ-tocopherol by γ-methyl-transferase generates the biologically active α-tocopherol. Tocopherol methyl transferase 2 (TMT2) shows similar activity to MT1.

There is a need in the art for nucleic acid molecules encoding enzymes involved in tocopherol biosysnthesis, as well as related enzymes and antibodies for the enhancement or alteration of tocopherol production in plants. There is a further need for transgenic organisms expressing those nucleic acid molecules involved in tocopherol biosynthesis, which are capable of nutritionally enhancing food and feed sources.

Gene ID	Enzyme name
tyrA	Prephanate dehydrogenase
slr1736	Phytylprenyl transferase from Synechocystis
ATPT2	Phytylprenyl transferase from Arabidopsis thaliana
DXS	1-Deoxyxylulose-5- phosphate synthase
DXR	1-Deoxyxylulose-5- phosphate reductoisomerase
GGPPS	Geranylgeranyl pyrophosphate synthase
HPPD	p-Hydroxyphenylpyruvate dioxygenase
AANT1	Adenylate transporter
slr1737	Tocopherol cyclase
IDI	Isopentenyl diphosphate isomerase
GGH	Geranylgeranyl reductase
MT1	Methyl transferase 1
tMT2	Tocopherol methyl transferase 2
GMT	Gamma Methyl Transferase

As used herein, homogentisate phytyl transferase (HPT), phytylprenyl transferase (PPT), slr1736, and ATPT2, each refer to proteins or genes encoding proteins that have the same enzymatic activity.

SUMMARY OF THE INVENTION

The present invention includes and provides a substantially purified nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule that encodes an enzyme with chorismate mutase and prephenate dehydrogenase activities or a fragment thereof of at least 20 contiguous amino acids of said enzyme.

The present invention includes and provides a substantially purified nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule that encodes an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4 and fragments thereof of at least 20 contiguous amino acids.

The present invention includes and provides a nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein having chorismate mutase and prephenate dehydrogenase activities.

The present invention includes and provides a nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4.

The present invention includes and provides a transformed plant having a nucleic acid molecule which comprises as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) an exogenous nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4 or fragments thereof encoding at least 20 contiguous amino acids, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a transformed plant having a nucleic acid molecule comprising as operably linked components: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; which is linked to (B) a transcribed nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 4 and fragments thereof comprising at least 20 contiguous amino acids.

The present invention includes and provides a method of producing a plant having increased tocopherol levels comprising: (A) transforming said plant with a nucleic acid molecule, wherein said nucleic acid molecule comprises a promoter region, wherein said promoter region is linked to a nucleic acid sequence that encodes a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4; and, growing said plant.

The present invention includes and provides a method for reducing tocopherol levels in a plant comprising: (A) transforming said plant with a nucleic acid molecule, wherein said nucleic acid molecule comprises as operably linked components an exogenous promoter region which functions in plant cells to cause the production of an mRNA molecule, a transcribed nucleic acid molecule having a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3; and wherein said transcribed nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant cells to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA sequence; and (B) growing said transformed plant.

The present invention includes and provides a method for screening for increased tocopherol levels in a plant comprising interrogating genomic DNA for the presence or

absence of a marker molecule that specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3 and complements thereof; and detecting said presence or absence of said marker.

The present invention includes and provides a method for determining a genomic polymorphism in a plant that is predictive of an increased tocopherol level comprising the steps: (A) incubating a marker nucleic acid molecule and a complementary nucleic acid molecule obtained from said plant under conditions permitting nucleic acid hybridization, wherein said marker nucleic acid molecule specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3 and complements thereof; (B) permitting hybridization between said marker nucleic acid molecule and said complementary nucleic acid molecule obtained from said plant; and (C) detecting the presence of said polymorphism.

The present invention includes and provides a method for determining a level or pattern of expression of a protein in a plant cell or plant tissue comprising:

(A) incubating under conditions permitting nucleic acid hybridization: a marker nucleic acid molecule, said marker nucleic acid molecule having a nucleic acid sequence selected from SEQ ID NO: 1 and SEQ ID NO: 3, complements of either or fragments comprising at least about 20 nucleotides of said sequences, with a complementary nucleic acid molecule obtained from a plant cell or plant tissue, (B) permitting hybridization between said marker nucleic acid molecule and said complementary nucleic acid molecule obtained from said plant cell or plant tissue; and (C) detecting said level or pattern of said complementary nucleic acid, wherein detection of said complementary nucleic acid is predictive of said level or pattern of said expression of said protein.

The present invention includes and provides a method for determining a level or pattern of expression of a protein in a plant cell or plant tissue under evaluation, comprising: assaying a concentration of an indicator molecule in said plant cell or plant tissue under evaluation, wherein said concentration of said indicator molecule is

dependent upon expression of a gene, and wherein said gene specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3 and complements thereof; and, comparing said concentration of said indicator molecule with known concentrations of said indicator molecule that occur in plant cells or plant tissues with known levels or patterns of expression of said protein.

The present invention includes and provides a cell comprising a nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule, wherein said heterologous nucleic acid molecule encodes an enzyme with chorismate mutase and prephenate dehydrogenase activities or a fragment of said nucleic acid molecule comprising at least 20 contiguous amino acids.

The present invention includes and provides oil derived from a seed of a transformed plant having a nucleic acid molecule which comprises as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) an exogenous nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a method of preparing tocopherols which comprises: transforming a plant with a nucleic acid comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) an exogenous nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and, growing said plant.

The present invention includes and provides a method of preparing homogentesic acid which comprises transforming a plant with a nucleic acid comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) an exogenous nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides a method of preparing plastoquinones which comprises transforming a plant with a nucleic acid comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) an exogenous nucleic acid molecule encoding a protein or fragment thereof comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention includes and provides feedstock comprising a transformed plant or part thereof, wherein said transformed plant has an exogenous nucleic acid molecule comprising a sequence selected from the group of SEQ ID NOs: 1 and 3.

The present invention includes and provides a meal comprising plant material manufactured from a transformed plant, wherein said transformed plant contains an exogenous nucleic acid molecule comprising a sequence selected from the group of SEQ ID NOs: 1 and 3.

The present invention includes and provides a transformed plant having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4.

The present invention includes and provides a transformed plant having an exogenous nucleic acid molecule that comprises a nucleic acid sequence of SEQ ID NO: 1 or 3.

The present invention includes and provides a method of producing a plant having seeds with increased tocopherol level comprising: (A) transforming said plant with a nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities ; and (b) growing said transformed plant.

The present invention includes and provides a method of producing a plant having seeds with increased tocopherol level comprising: (A) transforming said plant with a nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4; and (b) growing said transformed plant.

The present invention includes and provides a seed derived from a transformed plant having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said seed has an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides a seed derived from a transformed plant having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides oil derived from a seed of a transformed plant having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said transformed

plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides oil derived from a seed of a transformed plant having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides feedstock comprising a transformed plant or part thereof, having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides feedstock comprising a transformed plant or part thereof, having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides feedstock comprising a transformed plant or part thereof having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides a meal comprising plant material manufactured from a transformed plant, having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides a meal comprising plant material manufactured from a transformed plant having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides a meal comprising plant material manufactured from a transformed plant, having an exogenous nucleic acid molecule that encodes a protein with chorismate mutase and prephenate dehydrogenase activities, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking said exogenous nucleic acid molecule.

The present invention includes and provides a meal comprising plant material manufactured from a transformed plant having an exogenous nucleic acid molecule which comprises a nucleic acid sequence that encodes for a protein having an amino acid sequence of SEQ ID NO: 2 or 4, wherein said transformed plant has a seed with an increased tocopherol level relative to seeds from a plant having a similar genetic background but lacking the exogenous nucleic acid molecule.

The present invention includes and provides nucleic acid constructs, as well as plants and organisms containing those constructs, having combinations of two or more genes involved in tocopherol and tocotrienol biosynthesis. Any combination of the

following genes with *tyrA* is prefered: *slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase. In a particularly preferred embodiment, *tyrA* is combined with *HPPD* and either slr1736 or *ATPT2*.

DESCRIPTION OF THE NUCLEIC AND AMINO ACID SEQUENCES

SEQ ID NO: 1 sets forth a nucleic acid sequence of DNA molecule that encodes an *Erwinia herbicola* bifunctional prephenate dehyrdrogenase.

SEQ ID NO: 2 sets forth a derived amino acid sequence of an *Erwinia herbicola* bifunctional prephenate dehyrdrogenase.

SEQ ID NO: 3 sets forth a nucleic acid sequence of DNA molecule that encodes an *Escherichia coli* bifunctional prephenate dehyrdrogenase.

SEQ ID NO: 4 sets forth a derived amino acid sequence of an *Escherichia coli* bifunctional prephenate dehyrdrogenase.

SEQ ID NO: 5 sets forth a 5' primer used for amplification of an *Erwinia herbicola* tyrA sequence.

SEQ ID NO: 6 sets forth a 3' primer used for amplification of an *Erwinia herbicola* tyrA sequence.

SEQ ID NO: 7 sets forth a 5' primer used for amplification of an *Escherichia coli* tyrA sequence.

SEQ ID NO: 8 sets forth a 3' primer used for amplification of an *Escherichia coli* tyrA sequence.

SEQ ID NO: 9 sets forth a primer sequence.

SEQ ID NO: 10 sets forth a primer sequence.

SEQ ID NO: 11 sets forth a primer sequence.

SEQ ID NO: 12 sets forth a primer sequence.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a graph comparing total seed tocopherol levels of *Arabidopsis thaliana* lines harboring an *Erwinia herbicola tyrA* expression construct with a plastidal target sequence, a wild type plant, and a plant with a vector control.

Figure 2 is a schematic of construct pMON26588.

Figure 3 is a schematic of construct pMON26589.

Figure 4 is a schematic of construct pMON26591.

Figure 5 is a schematic of construct pMON26590.

Figure 6 is a schematic of construct pMON36510.

Figure 7 is a schematic of construct pMON36512.

Figure 8 is a schematic of construct pMON36511.

Figure 9 is a schematic of construct pMON36520.

Figure 10 is a schematic of construct pCGN10822.

Figure 11 is a schematic of construct pMON36528.

Figure 12 is a schematic of construct pMON69907.

Figure 13 is a schematic of construct pMON69909.

Figure 14 depicts the total tocopherol and tocotrienol content of *Arabidopsis* seeds from wild type plants and several plant lines transformed with the plasmid vector pMON69907.

Figure 15 depicts the total tocopherol content of *Arabidopsis* seeds from wild type plants and several plant lines transformed with the plasmid vector pMON69907.

Figure 16 depicts the total tocopherol and tocotrienol content of *Arabidopsis thaliana* seeds from wild type plants and several plant lines separately transformed with

the plasmid vectors pCGN10822, pMON36528, pMON69907 and pMON69909.

Figure 17 depicts a detailed analysis of tocopherol and tocotrienol content of arabidopsis seeds from plant lines transformed with the vector pMON69909 relative to wild type plant seeds.

Figures 18a and 18b show exepmlary constructs for transformation into plants.

Figure19 is a schematic construct of pMON36582.

Figure 20 is an example of a shuttle vector (pMON36586) harboring an expression cassette of the *Arabidopsis* homogentisate phytyltransferase (HPT) as a Bsp120 I/Not I cassette. The napin promoter and napin terminator are fused to the 5' and 3' ends, respectively, to drive seed specific expression.

Figure 21 represents various gene expression cassettes shown in a shuttle vector (A) or in a binary vector (B).

Figure 22 is a schematic of construct pMON36596.

Figure 23 is a schematic of construct pMON36597.

Figure 24 is a schematic of construct pMON77601.

Figure 25 is a schematic of construct pMON77602.

Figure 26 is a schematic of construct pMON66657.

Figure 27 is a schematic of construct pMON66659.

Figure 28 is a schematic of construct pMON26541.

Figure 29 is a schematic of construct pMON26543.

Figure 30 is a schematic of construct pMON36176.

Figure 31 is an LC/MS standard graph.

Figure 32 is an LC/MS graph.

Figure 33 is an HPLC chromatograph.

Figure 34 is an HPLC chromatograph.

Figure 35 is a schematic of construct pMON69915.

Figure 36 is a schematic of construct pMON69919.

Figure 37 is a schematic of construct pMON10098.

Figure 38 is a schematic of construct pMON36520.

Figure 39 is a schematic of construct pMON43853.

Figure 40 is a schematic of construct pMON36525.

Figure 41 is a schematic of construct pMON43861.

Figure 42 is a schematic of construct pCGN7770.

Figure 43 is a schematic of construct pMON69911.

Figure 44 is a schematic of construct pCGN11301.

Figure 45 is a schematic of construct pCGN3223.

Figure 46 is a schematic of construct pMON36575.

Figure 47 is a schematic of construct pMON38207R.

Figure 48 is a schematic of construct pMON36571.

Figure 49 is a schematic of construct pMON36576.

Figure 50 is a schematic of construct pMON69924.

Figure 51 is a schematic of construct pMON69943.

Figure 52 is a schematic of construct pMON69929.

Figure 53 is a schematic of construct pMON69936.

Figure 54 is a schematic of construct pMON36592.

Figure 55 is a schematic of construct pMON69945.

Figure 56 is a schematic of construct pMON16602.

Figure 57 is a schematic of construct pMON36525.

Figure 58 is a schematic of construct pMON58171.

Figure 59 is a schematic of construct pMON58172.

Figure 60 is a schematic of construct pMON58170.

Figure 61 is a schematic of construct pMON36591.

Figure 62 is a schematic of construct pMON36588.

Figure 63 is a schematic of construct pMON36592.

Figure 64 is a schematic of construct pMON58182.

Figure 65 is a schematic of construct pMON58176.

Figure 66 is a schematic of construct pMON58183.

Figure 67 is a schematic of construct pMON58185.

Figure 68 is a schematic of construct pMON36593.

Figure 69 is a schematic of construct pMON36589.

Figure 70 is a schematic of construct pMON36590.

Figure 71 is a schematic of construct pMON67162.

Figure 72 is a schematic of construct pMON58178.

Figure 73 is a schematic of construct pMON58186.

Figure 74 is a schematic of construct pMON58188.

Figure 75 shows tocopherol and tocotrienol as well as HGA levels in selected lines.

Figure 76 shows tocotrienol and 2M6PPQ levels in selected lines.

DETAILED DESCRIPTION

Any of the nucleic acid molecules disclosed herein can be enhanced or over-expressed in a variety of organisms, such as plants, which can result in higher levels of tocopherol precursors such as homogentisic acid (HGA) and ultimately in enhanced levels of tocopherols in such organisms. In addition, the enhanced expression or over expression of proteins set forth herein can also result in the production of increased levels of plastoquinones. Moreover, the present invention provides a number of agents, for example, nucleic acid molecules and proteins associated with the production of tocopherols, and provides uses of such agents.

The present invention includes and provides for nucleic acid constructs for expression of bifunctional prephenate dehydrogenases in organisms in which it is desirable to produce an increased yield of homogentesic acid, plastoquinones, or tocopherols. Such nucleic acid constructs may be used in organisms for which an increased level of prephenate dehydrogenase activity is desirable. The invention also includes and provides for nucleic acids constructs for the expression of phytyl prenyltransferases in organisms in which it is desirable to produce an increased yield of plastoquinones, or tocopherols, and the use of constructs producing antisense nucleic acids against phytyl prenyltransferases in organisms in which it is desirable to produce an increased yield of homogentesic acid.

AGENTS:

The agents of the invention will preferably be "biologically active" with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid molecule, or the ability of a protein to be bound by an antibody (or to compete with another molecule for such binding). Alternatively, such an attribute may be catalytic and thus involve the capacity of the agent to mediate a chemical reaction or response. The agents will preferably be "substantially purified." The term "substantially

purified," as used herein, refers to a molecule separated from substantially all other molecules normally associated with it in its native state. More preferably a substantially purified molecule is the predominant species present in a preparation. A substantially purified molecule may be greater than 60% free, preferably 75% free, more preferably 90% free, and most preferably 95% free from the other molecules (exclusive of solvent) present in the natural mixture. The term "substantially purified" is not intended to encompass molecules present in their native state.

The agents of the invention may also be recombinant. As used herein, the term recombinant means any agent (*e.g.*, DNA, peptide *etc.*), that is, or results, however indirectly, from human manipulation of a nucleic acid molecule.

It is understood that the agents of the invention may be labeled with reagents that facilitate detection of the agent (*e.g.*, fluorescent labels, Prober *et al., Science 238*:336-340 (1987); Albarella *et al.*, EP 144914; chemical labels, Sheldon *et al.*, U.S. Patent 4,582,789; Albarella *et al.*, U.S. Patent 4,563,417; modified bases, Miyoshi *et al.*, EP 119448).

NUCLEIC ACID MOLECULES

Agents of the invention include nucleic acid molecules which encode a bifunctional prephenate dehydrogenase, having both chorismate mutase and prephenate dehydrogenase activities. In a preferred aspect of the present invention, the nucleic acid molecule comprises a nucleic acid sequence which encodes a bacterial homologue of a bifuntional prephenate dehydrogenase. In a preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence having SEQ ID NOs: 1 or 3. In another preferred embodiment, the nucleic acid molecule is a fragment of any nucleic acid sequence disclosed herein encoding an amino acid sequence having prephenate dehydrogenase activity.

In another preferred aspect of the present invention the nucleic acid molecule comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3, complements thereof, and fragments of either. In a further aspect of the present invention the nucleic acid molecule comprises a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4, and fragments thereof.

In another preferred aspect of the present invention nucleic acid molecules comprise both a nucleic acid sequence encoding a bifuntional prephenate dehydrogenase and an expression cassette for expressing phyty prenyltransferase. In a further aspect of the present invention, nucleic acid constructs separately encoding a bifunctional prephenate dehydrogenase and phyty prenyltransferase may be employed.

In another preferred aspect of the present invention a nucleic acid molecule comprises nucleotide sequences encoding a plastid transit peptide operably fused to a nucleic acid molecule that encodes a protein or fragment of the present invention.

It is understood that in a further aspect of nucleic acid sequences of the present invention, the nucleic acids can encode a protein that differs from any of the proteins in that one or more amino acids have been deleted, substituted or added without altering the function. For example, it is understood that codons capable of coding for such conservative amino acid substitutions are known in the art.

One subset of the nucleic acid molecules of the invention is fragment nucleic acids molecules. Fragment nucleic acid molecules may consist of significant portion(s) of, or indeed most of, the nucleic acid molecules of the invention, such as those specifically disclosed. Alternatively, the fragments may comprise smaller oligonucleotides (having from about 15 to about 400 nucleotide residues and more preferably, about 15 to about 30 nucleotide residues, or about 50 to about 100 nucleotide residues, or about 100 to about 200 nucleotide residues, or about 200 to about 400 nucleotide residues, or about 275 to about 350 nucleotide residues).

A fragment of one or more of the nucleic acid molecules of the invention may be a probe and specifically a PCR probe. A PCR probe is a nucleic acid molecule capable of initiating a polymerase activity while in a double-stranded structure with another nucleic acid. Various methods for determining the structure of PCR probes and PCR techniques exist in the art. Computer generated searches using programs such as Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www-genome.wi.mit.edu/cgi-bin/www-STS_Pipeline), or GeneUp (Pesole *et al.*, *BioTechniques 25*:112-123 (1998)), for example, can be used to identify potential PCR primers.

Another subset of the nucleic acid molecules of the invention include nucleic acid molecules that encode a protein or fragment thereof.

Nucleic acid molecules or fragments thereof of the present invention are capable of specifically hybridizing to other nucleic acid molecules under certain circumstances. Nucleic acid molecules of the present invention include those that specifically hybridize to nucleic acid molecules having a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3, and complements thereof. Nucleic acid molecules of the present invention also include those that specifically hybridize to nucleic acid molecules encoding an amino acid sequences selected from SEQ ID NOs: 2, 4, complements thereof, and fragments of either.

As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure.

A nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if they exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit

them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are described by Sambrook *et al., Molecular Cloning, A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press*, Cold Spring Harbor, New York (1989), and by Haymes *et al., Nucleic Acid Hybridization, A Practical Approach, IRL Press*, Washington, DC (1985). Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for a nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Appropriate stringency conditions which promote DNA hybridization are, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 20-25°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 X SSC at 50°C to a high stringency of about 0.2 X SSC at 65°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NOs: 1 and 3 and complements thereof under moderately stringent conditions, for example at about 2.0 X SSC and about 65°C.

In a particularly preferred embodiment, a nucleic acid of the present invention will include those nucleic acid molecules that specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NOs: 1 and 3 and complements thereof under high stringency conditions such as 0.2 X SSC and about 65°C.

In one aspect of the present invention, the nucleic acid molecules of the present invention have one or more of the nucleic acid sequences set forth in SEQ ID NOs: 1 and 3 and complements thereof. In another aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 90% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NOs: 1 and 3 and complements thereof and fragments of either. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 95% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NOs: 1 and 3, complements thereof, and fragments of either. In a more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 98% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NOs: 1 and 3, complements thereof and fragments of either. In an even more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 99% sequence identity with one or more of the sequences set forth in SEQ ID NOs: 1 and 3, complements thereof, and fragments of either.

In a preferred embodiment the percent identity calculations are performed using the Megalign program of the LASERGENE bioinformatics computing suite (default parameters, DNASTAR Inc., Madison, Wisconsin).

A nucleic acid molecule of the invention can also encode a homolog protein. As used herein, a homolog protein molecule or fragment thereof is a counterpart protein molecule or fragment thereof in a second species (*e.g.*, maize rubisco small subunit is a homolog of *Arabidopsis* rubisco small subunit). A homolog can also be generated by

molecular evolution or DNA shuffling techniques, so that the molecule retains at least one functional or structure characteristic of the original protein (*see*, for example, U.S. Patent 5,811,238).

In another embodiment, the homolog is selected from the group consisting of alfalfa, *Arabidopsis*, barley, *Brassica campestris, Brassica napus,* broccoli, cabbage, canola, citrus, cotton, garlic, oat, onion, flax, an ornamental plant, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, soybean, maize, *Phaseolus,* crambe, mustard, castor bean, sesame, cottonseed, linseed, safflower, and oil palm. More particularly, preferred homologs are selected from canola, maize, *Arabidopsis*, *Brassica campestris, Brassica napus*, soybean, crambe, mustard, castor bean, peanut, sesame, cottonseed, linseed, safflower, oil palm, flax, and sunflower. In an even more preferred embodiment, the homolog is selected from the group consisting of canola, maize, *Arabidopsis, Brassica campestris, Brassica napus*, soybean, sunflower, safflower, oil palms, and peanut. In a preferred embodiment, the homolog is soybean. In a preferred embodiment, the homolog is canola. In a preferred embodiment, the homolog is *Brassica napus*.

In a preferred embodiment, nucleic acid molecules having SEQ ID NOs: 1 and 3, complements thereof, and fragments of either; or more preferably SEQ ID NOs: 1 and 3 and complements thereof, can be utilized to obtain such homologs.

In another further aspect of the present invention, nucleic acid molecules of the present invention can comprise sequences that differ from those encoding a protein or fragment thereof in SEQ ID NOs: 2 and 4 due to fact that a protein can have one or more conservative amino acid changes, and nucleic acid sequences coding for the protein can therefore have sequence differences. It is understood that codons capable of coding for such conservative amino acid substitutions are known in the art.

It is well known in the art that one or more amino acids in a native sequence can be substituted with other amino acid(s), the charge and polarity of which are similar to that of the native amino acid, *i.e.*, a conservative amino acid substitution. Conservative substitutes for an amino acid within the native polypeptide sequence can be selected from other members of the class to which the amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids, (2) basic amino acids, (3) neutral polar amino acids, and (4) neutral, nonpolar amino acids. Representative amino acids within these various groups include, but are not limited to, (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cysteine, cystine, tyrosine, asparagine, and glutamine; and (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine.

Conservative amino acid substitution within the native polypeptide sequence can be made by replacing one amino acid from within one of these groups with another amino acid from within the same group. In a preferred aspect, biologically functional equivalents of the proteins or fragments thereof of the present invention can have ten or fewer conservative amino acid changes, more preferably seven or fewer conservative amino acid changes, and most preferably five or fewer conservative amino acid changes. The encoding nucleotide sequence will thus have corresponding base substitutions, permitting it to encode biologically functional equivalent forms of the proteins or fragments of the present invention.

It is understood that certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Because it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence

substitutions can be made in a protein sequence and, of course, its underlying DNA coding sequence and, nevertheless, a protein with like properties can still be obtained. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the proteins or fragments of the present invention, or corresponding DNA sequences that encode said peptides, without appreciable loss of their biological utility or activity. It is understood that codons capable of coding for such amino acid changes are known in the art.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte and Doolittle, *J. Mol. Biol. 157*, 105-132 (1982)). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like.

Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, *J. Mol. Biol. 157:*105-132 (1982)); these are isoleucine (+4.5), valine (+4.2), leucine (+3.8), phenylalanine (+2.8), cysteine/cystine (+2.5), methionine (+1.9), alanine (+1.8), glycine (-0.4), threonine (-0.7), serine (-0.8), tryptophan (-0.9), tyrosine (-1.3), proline (-1.6), histidine (-3.2), glutamate (-3.5), glutamine (-3.5), aspartate (-3.5), asparagine (-3.5), lysine (-3.9), and arginine (-4.5).

In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101 states that the

greatest local average hydrophilicity of a protein, as govern by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0), lysine (+3.0), aspartate (+3.0±1), glutamate (+3.0±1), serine (+0.3), asparagine (+0.2), glutamine (+0.2), glycine (0), threonine (-0.4), proline (-0.5±1), alanine (-0.5), histidine (-0.5), cysteine (-1.0), methionine (-1.3), valine (-1.5), leucine (-1.8), isoleucine (-1.8), tyrosine (-2.3), phenylalanine (-2.5), and tryptophan (-3.4).

In making such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those for which a specific sequence is provided herein because one or more codons has been replaced with a codon that encodes a conservative substitution of the amino acid originally encoded.

Agents of the invention include nucleic acid molecules that encode at least about a contiguous 10 amino acid region of a protein of the present invention, more preferably at least about a contiguous 25, 40, 50, 100, or 125 amino acid region of a protein of the present invention.

In a preferred embodiment, any of the nucleic acid molecules of the present invention can be operably linked to a promoter region which functions in a plant cell to cause the production of an mRNA molecule, where the nucleic acid molecule that is linked to the promoter is heterologous with respect to that promoter. As used herein, "heterologous" means not naturally occurring together.

PROTEIN AND PEPTIDE MOLECULES

A class of agents includes one or more of the proteins or fragments thereof or peptide molecules encoded by a nucleic acid agent of the invention. A particular preferred class of proteins is that having an amino acid sequence selected from the group consisting of SEQ ID NOs: 2 and 4 and fragments thereof. Protein or peptide agents may have C-terminal or N-terminal amino acid sequence extensions. One class of N-terminal extensions employed in a preferred embodiment are plastid transit peptides. When employed, plastid transit peptides can be operatively linked to the N-terminal sequence, thereby permitting the localization of the agent peptide or proteins to plastids. In a preferred embodiment the plastid targeting sequence is a CTP1 sequence. In another embodiment the sequence is a CTP2 sequence.

As used herein, the term "protein" or "peptide molecule" includes any molecule that comprises five or more amino acids. It is well known in the art that proteins may undergo modification, including post-translational modifications, such as, but not limited to, disulfide bond formation, glycosylation, phosphorylation, or oligomerization. Thus, as used herein, the term "protein" or "peptide molecule" includes any protein that is modified by any biological or non-biological process. The terms "amino acid" and "amino acids" refer to all naturally occurring L-amino acids. This definition is meant to include norleucine, norvaline, ornithine, homocysteine, and homoserine.

One or more of the protein or fragments thereof or peptide molecules may be produced via chemical synthesis, or more preferably, by expression in a suitable bacterial or eukaryotic host. Suitable methods for expression are described by Sambrook *et al.*, In: *Molecular Cloning, A Laboratory Manual, 2nd Edition*, *Cold Spring Harbor Press*, Cold Spring Harbor, New York (1989) or similar texts.

A "protein fragment" is a peptide or polypeptide molecule whose amino acid sequence comprises a subset of the amino acid sequence of that protein. A protein or fragment thereof that comprises one or more additional peptide regions not derived from that protein is a "fusion" protein. Such molecules may be derivatized to contain

carbohydrate or other moieties (such as keyhole limpet hemocyanin). Fusion protein or peptide molecules of the invention are preferably produced via recombinant means.

Another class of agents comprise protein or peptide molecules or fragments or fusions thereof comprising SEQ ID NOs: 2 and 4 and fragments thereof in which conservative, non-essential or non-relevant amino acid residues have been added, replaced or deleted. Computerized means for designing modifications in protein structure are known in the art (Dahiyat and Mayo, *Science* 278:82-87 (1997)).

A protein of the invention can also be a homolog protein. As used herein, a homolog protein or fragment thereof is a counterpart protein or fragment thereof in a second species. A homolog can also be generated by molecular evolution or DNA shuffling techniques, so that the molecule retains at least one functional or structure characteristic of the original (*see*, for example, U.S. Patent 5,811,238).

In another embodiment, the homolog is selected from the group consisting of alfalfa, *Arabidopsis*, barley, broccoli, cabbage, canola, citrus, cotton, garlic, oat, onion, flax, an ornamental plant, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, soybean, maize, and *Phaseolus*. More particularly, preferred homologs are selected from canola, maize, *Arabidopsis, Brassica campestris, Brassica napus*, soybean, crambe, mustard, castor bean, peanut, sesame, cottonseed, linseed, safflower, oil palm, flax, and sunflower. In an even more preferred embodiment, the homolog is selected from the group consisting of canola, maize, *Arabidopsis, Brassica campestris, Brassica napus*, soybean, sunflower, safflower, oil palms, and peanut. In a preferred embodiment, the homolog is soybean. In a preferred embodiment, the homolog is canola. In a preferred embodiment, the homolog is *Brassica napus*.

In a preferred embodiment, the nucleic acid molecules of the present invention or complements and fragments of either can be utilized to obtain such homologs.

Agents of the invention include proteins and fragments thereof comprising at least about a contiguous 10 amino acid region preferably comprising at least about a contiguous 20 amino acid region, even more preferably comprising at least a contiguous 25, 35, 50, 75 or 100 amino acid region of a protein of the present invention. In another preferred embodiment, the proteins of the present invention include between about 10 and about 25 contiguous amino acid region, more preferably between about 20 and about 50 contiguous amino acid region, and even more preferably between about 40 and about 80 contiguous amino acid region.

PLANT CONSTRUCTS AND PLANT TRANSFORMANTS

One or more of the nucleic acid molecules of the invention may be used in plant transformation or transfection. Exogenous genetic material may be transferred into a plant cell and the plant cell regenerated into a whole, fertile or sterile plant. Exogenous genetic material is any genetic material, whether naturally occurring or otherwise, from any source that is capable of being inserted into any organism. In a preferred embodiment, the exogenous genetic material encodes a bifunctional prephenate dehydrogenase or fragments thereof, more preferably a. bifunctional prephenate dehydrogenase from a prokaryotic organism, and even more preferably a bifunctional prephenate dehydrogenase from *Erwinia herbicola* or *Escherichia coli.* In a preferred embodiment, the exogenous genetic material includes a nucleic acid molecule of the present invention, and preferably a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NOs: 1 and 3, complements thereof, and fragments of either. In another embodiment, the exogenous genetic material includes a nucleic acid molecule of the present invention, preferably a nucleic acid encoding a protein or fragment thereof having phytyl prenyltransferase activity.

In an embodiment of the present invention, exogenous genetic material comprising a TyrA homolog or fragment thereof is introduced into a plant with one or more additional

genes. In one embodiment, preferred combinations of genes include two or more of the following genes: *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase (Krindl *et al., Seed Sci. Res.* 1:209:219 (1991); Keegstra, *Cell 56(2)*:247-53 (1989); Nawrath, *et al., Proc. Natl. Acad. Sci. U.S.A. 91*:12760-12764 (1994); Xia *et al., J. Gen. Microbiol. 138*:1309-1316 (1992); Cyanobase on the world wide web at www.kazusa.or.jp/cyanobase; Lois *et al., Proc. Natl. Acad. Sci. U.S.A. 95 (5)*:2105-2110 (1998); Takahashi et al. *Proc. Natl. Acad. Sci. U.S.A. 95 (17)*, 9879-9884 (1998); Norris *et al., Plant Physiol. 117*:1317-1323 (1998); Bartley and Scolnik, *Plant Physiol. 104*:1469-1470 (1994), Smith *et al., Plant J. 11: 8*3-92 (1997); WO 00/32757; WO 00/10380; Saint Guily, *et al., Plant Physiol., 100(2)*:1069-1071 (1992); Sato *et al., J. DNA Res. 7 (1)*:31-63 (2000)). In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA*, *HPPD* and either *slr1736* or *ATPT2*.

In such combinations, one or more of the gene products can be directed to the plastid by the use of a plastid targeting sequence. Alternatively, one or more of the gene products can be localized in the cytoplasm. Such genes can be introduced, for example, with the TyrA homolog or fragment thereof on a single construct, introduced on different constructs but the same transformation event or introduced into separate plants followed by one or more crosses to generate the desired combination of genes. In such combinations, a preferred promoter is a napin promoter and a preferred plastid targeting sequence is a CTP1 sequence.

Such genetic material may be transferred into either monocotyledons and dicotyledons including, but not limited to canola, maize, soybean, *Arabidopsis* phaseolus, peanut, alfalfa, wheat, rice, oat, sorghum, rye, tritordeum, millet, fescue, perennial ryegrass, sugarcane, cranberry, papaya, banana, safflower, oil palms, flax, muskmelon, apple, cucumber, dendrobium, gladiolus, chrysanthemum, liliacea, cotton, eucalyptus, sunflower, *Brassica campestris, Brassica napus*, turfgrass, sugarbeet, coffee

and dioscorea (Christou, In: *Particle Bombardment for Genetic Engineering of Plants,* Biotechnology Intelligence Unit. Academic Press, San Diego, California (1996)), with canola, maize, *Arabidopsis, Brassica campestris, Brassica napus*, soybean, crambe, mustard, castor bean, peanut, sesame, cottonseed, linseed, safflower, oil palm, flax, and sunflower preferred, and canola, maize, *Arabidopsis, Brassica campestris, Brassica napus*, soybean, sunflower, safflower, oil palms, and peanut preferred. In a more preferred embodiment, the genetic material is transferred into canola. In another more preferred embodiment, the genetic material is transferred into *Brassica napus*. In another more preferred embodiment, the genetic material is transferred into soybean.

Transfer of a nucleic acid that encodes a protein can result in expression or overexpression of that protein in a transformed cell or transgenic plant. One or more of the proteins or fragments thereof encoded by nucleic acid molecules of the invention may be overexpressed in a transformed cell or transformed plant. Such expression or overexpression may be the result of transient or stable transfer of the exogenous genetic material.

In a preferred embodiment, expression or overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of tocotrienols.

In a preferred embodiment, expression or overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of α-tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of γ-tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of homogentesic acid

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, relative to an untransformed plant with a similar genetic background, an increased level of plastoquinols or plastoquinones.

In some embodiments, the levels of one or more products of the tocopherol biosynthesis pathway, including any one or more of tocotrienols, tocopherols, α-tocopherols, γ-tocopherols plastoquinols, plastoquinones or homogentesic acid, are increased by 10%, or more preferably 25%, 50%, 100%, 200%, 250%, 1,000%, 2,000%, or 2,500%. The levels of products may be increased throughout an organism such as a plant or localized in one or more specific organs or tissues of the organism. For example the levels of products may be increased in one or more of the tissues and organs of a plant include without limitation: roots, tubers, stems, leaves, stalks, fruit, berries, nuts, bark, pods, seeds and flowers.

In another embodiment, overexpression of a protein or fragment thereof of the present invention in a plant provides in that plant, or a tissue of that plant, relative to an untransformed plant or plant tissue, with a similar genetic background, an increased level of prephenate dehydrogenase protein.

In another preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a transformed plant may provide tolerance to a variety of stress, *e.g.* oxidative stress tolerance such as to oxygen or ozone, UV tolerance, cold tolerance, or fungal/microbial pathogen tolerance.

As used herein in a preferred aspect, a tolerance or resistance to stress is determined by the ability of a plant, when challenged by a stress such as cold to produce a plant having a higher yield than one without such tolerance or resistance to stress. In a particularly preferred aspect of the present invention, the tolerance or resistance to stress

is measured relative to a plant with a similar genetic background to the tolerant or resistance plant except that the plant expresses or over expresses a protein or fragment thereof of the present invention.

Exogenous genetic material may be transferred into a host cell by the use of a DNA vector or construct designed for such a purpose. Design of such a vector is generally within the skill of the art (*See*, *Plant Molecular Biology: A Laboratory Manual*, Clark (ed.), Springer, New York (1997)). In some embodiments of this invention a single gene sequence selected from the group consisting of *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase can be transferred into the desired target plant. In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA*, *HPPD* and either *slr1736* or *ATPT2*. Target plants expressing the desired activity from the transferred gene sequence can be subject to one or more crosses with plants having been transformed with one or more other gene sequences selected from the group consisting of *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase in order to obtain plants expressing two or more of the desired activities from the transferred gene sequence. In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA*, *HPPD* and either *slr1736* or *ATPT2*. In another embodiment, DNA vector constructs may be multiple gene constructs that comprise two or more gene sequences selected from the group consisting of *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase, such that transformation with a single DNA vector construct will result in the expression of the two or more of the gene sequences. In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA*, *HPPD* and either *slr1736* or *ATPT2*.

A construct or vector may include a plant promoter to express the protein or protein fragment of choice. In a preferred embodiment, any nucleic acid molecules described herein can be operably linked to a promoter region which functions in a plant cell to cause the production of an mRNA molecule. For example, any promoter that functions in a plant cell to cause the production of an mRNA molecule, such as those promoters described herein, without limitation, can be used. In a preferred embodiment, the promoter is a plant promoter.

A number of promoters that are active in plant cells have been described in the literature. These include the nopaline synthase (NOS) promoter (Ebert *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 84*:5745-5749 (1987)), the octopine synthase (OCS) promoter (which is carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*), the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton *et al.*, *Plant Mol. Biol. 9*:315-324 (1987)) and the CaMV 35S promoter (Odell *et al.*, *Nature 313*:810-812 (1985)), the figwort mosaic virus 35S-promoter, the light-inducible promoter from the small subunit of ribulose-1,5-bis-phosphate carboxylase (ssRUBISCO), the Adh promoter (Walker *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 84*:6624-6628 (1987)), the sucrose synthase promoter (Yang *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 87*:4144-4148 (1990)), the R gene complex promoter (Chandler *et al.*, *The Plant Cell 1*:1175-1183 (1989)) and the chlorophyll a/b binding protein gene promoter, *etc.* These promoters have been used to create DNA constructs that have been expressed in plants; *see, e.g.*, PCT publication WO 84/02913. The CaMV 35S promoters are preferred for use in plants. Promoters known or found to cause transcription of DNA in plant cells can be used in the invention.

For the purpose of expression in source tissues of the plant, such as the leaf, seed, root or stem, it is preferred that the promoters utilized have relatively high expression in these specific tissues. Tissue-specific expression of a protein of the present invention is a particularly preferred embodiment. For this purpose, one may choose from a number of

promoters for genes with tissue- or cell-specific or enhanced expression. Examples of such promoters reported in the literature include the chloroplast glutamine synthetase GS2 promoter from pea (Edwards *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 87*:3459-3463 (1990)), the chloroplast fructose-1,6-biphosphatase (FBPase) promoter from wheat (Lloyd *et al.*, *Mol. Gen. Genet. 225*:209-216 (1991)), the nuclear photosynthetic ST-LS1 promoter from potato (Stockhaus *et al.*, *EMBO J. 8*:2445-2451 (1989)), the serine/threonine kinase (PAL) promoter and the glucoamylase (CHS) promoter from *Arabidopsis thaliana*. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase (RbcS) promoter from eastern larch (*Larix laricina*), the promoter for the *cab* gene, cab6, from pine (Yamamoto *et al.*, *Plant Cell Physiol. 35*:773-778 (1994)), the promoter for the Cab-1 gene from wheat (Fejes *et al.*, *Plant Mol. Biol. 15*:921-932 (1990)), the promoter for the CAB-1 gene from spinach (Lubberstedt *et al.*, *Plant Physiol. 104*:997-1006 (1994)), the promoter for the cab1R gene from rice (Luan *et al.*, *Plant Cell. 4*:971-981 (1992)), the pyruvate, orthophosphate dikinase (PPDK) promoter from maize (Matsuoka *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 90*: 9586-9590 (1993)), the promoter for the tobacco Lhcb1*2 gene (Cerdan *et al.*, *Plant Mol. Biol. 33*:245-255 (1997)), the *Arabidopsis thaliana* SUC2 sucrose-H+ symporter promoter (Truernit *et al.*, *Planta. 196*:564-570 (1995)) and the promoter for the thylakoid membrane proteins from spinach (*psa*D, *psa*F, *psa*E, PC, FNR, *atp*C, *atp*D, *cab*, *rbc*S). Other promoters for the chlorophyll a/b-binding proteins may also be utilized in the invention, such as the promoters for *Lhc*B gene and *Psb*P gene from white mustard (*Sinapis alba*; Kretsch *et al.*, *Plant Mol. Biol. 28*:219-229 (1995)).

For the purpose of expression in sink tissues of the plant, such as the tuber of the potato plant, the fruit of tomato, or the seed of maize, wheat, rice and barley, it is preferred that the promoters utilized in the invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or tuber-enhanced expression are known, including the class I patatin promoter (Bevan *et al.*,

EMBO J. 8:1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol. 14*:995-1006 (1990)), the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter (Salanoubat and Belliard, *Gene 60*:47-56 (1987), Salanoubat and Belliard, *Gene 84*:181-185 (1989)), the promoter for the major tuber proteins including the 22 kd protein complexes and protease inhibitors (Hannapel, *Plant Physiol. 101*:703-704 (1993)), the promoter for the granule-bound starch synthase gene (GBSS) (Visser *et al.*, *Plant Mol. Biol. 17*:691-699 (1991)) and other class I and II patatins promoters (Koster-Topfer *et al.*, *Mol. Gen. Genet. 219*:390-396 (1989); Mignery *et al.*, *Gene. 62*:27-44 (1988)).

Other promoters can also be used to express a protein or fragment thereof in specific tissues, such as seeds or fruits. Indeed, in a preferred embodiment, the promoter used is a seed specific promoter. Examples of such promoters include the 5' regulatory regions from such genes as napin (Krindl *et al.*, *Seed Sci. Res.* 1:209:219 (1991)), phaseolin (Bustos, *et al., Plant Cell*, 1(9):839-853 (1989)), soybean trypsin inhibitor (Riggs, *et al., Plant Cell* 1(6):609-621 (1989)), ACP (Baerson, *et al., Plant Mol. Biol.,* 22(2):255-267 (1993)), stearoyl-ACP desaturase (Slocombe, *et al., Plant Physiol.* 104(4):167-176 (1994)), soybean a' subunit of b-conglycinin (soy 7s, (Chen *et al., Proc. Natl. Acad. Sci.,* 83:8560-8564 (1986))), and oleosin (see, for example, Hong, *et al., Plant Mol. Biol.,* 34(3):549-555 (1997)). Further examples include the promoter for β-conglycinin (Chen *et al.*, *Dev. Genet. 10*: 112-122 (1989)). Also included are the zeins, which are a group of storage proteins found in maize endosperm. Genomic clones for zein genes have been isolated (Pedersen *et al.*, *Cell 29*:1015-1026 (1982), and Russell *et al.*, *Transgenic Res. 6*(2):157-168) and the promoters from these clones, including the 15 kD, 16 kD, 19 kD, 22 kD, 27 kD and genes, could also be used. Other promoters known to function, for example, in maize include the promoters for the following genes: *waxy, Brittle, Shrunken 2*, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins and sucrose synthases. A particularly preferred promoter for maize

endosperm expression is the promoter for the glutelin gene from rice, more particularly the Osgt-1 promoter (Zheng *et al.*, *Mol. Cell Biol. 13*:5829-5842 (1993)). Examples of promoters suitable for expression in wheat include those promoters for the ADPglucose pyrosynthase (ADPGPP) subunits, the granule bound and other starch synthase, the branching and debranching enzymes, the embryogenesis-abundant proteins, the gliadins and the glutenins. Examples of such promoters in rice include those promoters for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases and the glutelins. A particularly preferred promoter is the promoter for rice glutelin, Osgt-1. Examples of such promoters for barley include those for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases, the hordeins, the embryo globulins and the aleurone specific proteins. A preferred promoter for expression in the seed is a napin promoter.

Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene (Samac *et al.*, *Plant Mol. Biol. 25*:587-596 (1994)). Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV35S promoter that have been identified (Lam *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 86*:7890-7894 (1989)). Other root cell specific promoters include those reported by Conkling *et al.* (Conkling *et al.*, *Plant Physiol. 93*:1203-1211 (1990)).

Additional promoters that may be utilized are described, for example, in U.S. Patent Nos. 5,378,619; 5,391,725; 5,428,147; 5,447,858; 5,608,144; 5,608,144; 5,614,399; 5,633,441; 5,633,435; and 4,633,436. In addition, a tissue specific enhancer may be used (Fromm *et al.*, *The Plant Cell 1*:977-984 (1989)).

Constructs or vectors may also include, with the coding region of interest, a nucleic acid sequence that acts, in whole or in part, to terminate transcription of that region. A number of such sequences have been isolated, including the Tr7 3' sequence and the NOS 3' sequence (Ingelbrecht *et al.*, *The Plant Cell 1*:671-680 (1989); Bevan *et*

al., *Nucleic Acids Res. 11*:369-385 (1983)). Regulatory transcript termination regions can be provided in plant expression constructs of this invention as well. Transcript termination regions can be provided by the DNA sequence encoding the gene of interest or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. The skilled artisan will recognize that any convenient transcript termination region which is capable of terminating transcription in a plant cell can be employed in the constructs of the present invention.

A vector or construct may also include regulatory elements. Examples of such include the Adh intron 1 (Callis *et al.*, *Genes and Develop. 1*:1183-1200 (1987)), the sucrose synthase intron (Vasil *et al.*, *Plant Physiol. 91*:1575-1579 (1989)) and the TMV omega element (Gallie *et al.*, *The Plant Cell 1*:301-311 (1989)). These and other regulatory elements may be included when appropriate.

A vector or construct may also include a selectable marker. Selectable markers may also be used to select for plants or plant cells that contain the exogenous genetic material. Examples of such include, but are not limited to: a *neo* gene (Potrykus *et al.*, *Mol. Gen. Genet. 199*:183-188 (1985)), which codes for kanamycin resistance and can be selected for using kanamycin, RptII, G418, hpt *etc.*; a bar gene which codes for bialaphos resistance; a mutant EPSP synthase gene (Hinchee *et al.*, *Bio/Technology 6*:915-922 (1988); Reynaerts *et al.,* Selectable and Screenable Markers. In Gelvin and Schilperoort. Plant Molecular Biology Manual, Kluwer, Dordrecht (1988); Reynaerts *et al.,* Selectable and screenable markers. In Gelvin and Schilperoort. Plant Molecular Biology Manual, Kluwer, Dordrecht (1988)), *aad*A (Jones *et al.,* Mol. Gen. Genet. (1987)),) which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil (Stalker *et al.*, *J. Biol. Chem. 263*:6310-6314 (1988)); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204 (Sept. 11, 1985)), ALS (D'Halluin *et al.,* Bio/Technology 10: 309-

314 (1992)), and a methotrexate resistant DHFR gene (Thillet *et al.*, *J. Biol. Chem.* *263*:12500-12508 (1988)).A vector or construct may also include a transit peptide.

Incorporation of a suitable chloroplast transit peptide may also be employed (European Patent Application Publication Number 0218571). Translational enhancers may also be incorporated as part of the vector DNA. DNA constructs could contain one or more 5' non-translated leader sequences which may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the mRNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence. For a review of optimizing expression of transgenes, see Koziel *et al., Plant Mol. Biol. 32*:393-405 (1996). A preferred transit peptide is CTP1. In another embodiment the transit peptide is a CTP2 sequence.

A vector or construct may also include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include: a β-glucuronidase or *uid*A gene (GUS) which encodes an enzyme for which various chromogenic substrates are known (Jefferson, *Plant Mol. Biol, Rep. 5*:387-405 (1987); Jefferson *et al., EMBO J. 6*:3901-3907 (1987)); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, Stadler Symposium *11*:263-282 (1988)); a β-lactamase gene (Sutcliffe *et al., Proc. Natl. Acad. Sci. (U.S.A.) 75*:3737-3741 (1978)), a gene which encodes an enzyme for which various chromogenic substrates are known (*e.g.*, PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al., Science 234*:856-859 (1986)); a *xyl*E gene (Zukowsky *et al., Proc. Natl. Acad. Sci. (U.S.A.) 80*:1101-1105 (1983)) which encodes a catechol dioxygenase that can convert chromogenic catechols; an α-amylase gene (Ikatu *et al., Bio/Technol. 8*:241-242 (1990)); a tyrosinase gene (Katz *et al.*, *J. Gen. Microbiol. 129*:2703-2714 (1983)) which encodes an enzyme capable of

oxidizing tyrosine to DOPA and dopaquinone which in turn condenses to melanin; an α-galactosidase, which will turn a chromogenic α-galactose substrate.

Included within the terms "selectable or screenable marker genes" are also genes which encode a secretable marker whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected catalytically. Secretable proteins fall into a number of classes, including small, diffusible proteins which are detectable, (*e.g.*, by ELISA), small active enzymes which are detectable in extracellular solution (*e.g.*, α-amylase, β-lactamase, phosphinothricin transferase), or proteins which are inserted or trapped in the cell wall (such as proteins which include a leader sequence such as that found in the expression unit of extension or tobacco PR-S). Other possible selectable and/or screenable marker genes will be apparent to those of skill in the art.

There are many methods for introducing transforming nucleic acid molecules into plant cells. Suitable methods are believed to include virtually any method by which nucleic acid molecules may be introduced into a cell, such as by *Agrobacterium* infection or direct delivery of nucleic acid molecules such as, for example, by PEG-mediated transformation, by electroporation or by acceleration of DNA coated particles, etc (Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol. 42*:205-225 (1991); Vasil, *Plant Mol. Biol. 25*:925-937 (1994)). For example, electroporation has been used to transform maize protoplasts (Fromm *et al.*, *Nature 312*:791-793 (1986)).

Other vector systems suitable for introducing transforming DNA into a host plant cell include but are not limited to binary artificial chromosome (BIBAC) vectors (Hamilton *et al., Gene 200*:107-116 (1997)); and transfection with RNA viral vectors (Della-Cioppa *et al., Ann. N.Y. Acad. Sci.* (1996), 792 (Engineering Plants for Commercial Products and Applications), 57-61). Additional vector systems also include

plant selectable YAC vectors such as those described in Mullen *et al., Molecular Breeding 4*:449-457 (1988).

Technology for introduction of DNA into cells is well known to those of skill in the art. Four general methods for delivering a gene into cells have been described:

(1) chemical methods (Graham and van der Eb, *Virology* 54:536-539 (1973));

(2) physical methods such as microinjection (Capecchi, *Cell* 22:479-488 (1980)), electroporation (Wong and Neumann, *Biochem. Biophys. Res. Commun. 107*:584-587 (1982); Fromm *et al., Proc. Natl. Acad. Sci. (U.S.A.) 82*:5824-5828 (1985); U.S. Patent No. 5,384,253); the gene gun (Johnston and Tang, *Methods Cell Biol. 43*:353-365 (1994)); and vacuum infiltration (Bechtold *et al., C.R. Acad. Sci. Paris, Life Sci. 316*:1194-1199. (1993)); (3) viral vectors (Clapp, *Clin. Perinatol. 20*:155-168 (1993); Lu *et al., J. Exp. Med. 178*:2089-2096 (1993); Eglitïs and Anderson, *Biotechniques 6*:608-614 (1988)); and (4) receptor-mediated mechanisms (Curiel *et al., Hum. Gen. Ther. 3*:147-154 (1992), Wagner *et al., Proc. Natl. Acad. Sci. (USA) 89*:6099-6103 (1992)).

Acceleration methods that may be used include, for example, microprojectile bombardment and the like. One example of a method for delivering transforming nucleic acid molecules into plant cells is microprojectile bombardment. This method has been reviewed by Yang and Christou (eds.), *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994)). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum and the like.

A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly transforming monocots, is that neither the isolation of protoplasts (Cristou *et al., Plant Physiol. 87*:671-674 (1988)) nor the susceptibility to *Agrobacterium* infection is required. An illustrative embodiment of a method for delivering DNA into maize cells by acceleration is a biolistics α-particle delivery system,

which can be used to propel particles coated with DNA through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with maize cells cultured in suspension. Gordon-Kamm *et al.,* describes the basic procedure for coating tungsten particles with DNA (Gordon-Kamm *et al.*, *Plant Cell 2*:603-618 (1990)). The screen disperses the tungsten nucleic acid particles so that they are not delivered to the recipient cells in large aggregates. A particle delivery system suitable for use with the invention is the helium acceleration PDS-1000/He gun, which is available from Bio-Rad Laboratories (Bio-Rad, Hercules, California)(Sanford *et al.*, *Technique 3*:3-16 (1991)).

For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein one may obtain 1000 or more loci of cells transiently expressing a marker gene. The number of cells in a focus which express the exogenous gene product 48 hours post-bombardment often ranges from one to ten, and average one to three.

In bombardment transformation, one may optimize the pre-bombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for bombardment are important in this technology. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target

cells to help alleviate the trauma associated with bombardment and also the nature of the transforming DNA, such as linearized DNA or intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

In another alternative embodiment, plastids can be stably transformed. Methods disclosed for plastid transformation in higher plants include the particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (Svab *et al., Proc. Natl. Acad. Sci. (U.S.A.) 87*:8526-8530 (1990); Svab and Maliga, *Proc. Natl. Acad. Sci. (U.S.A.) 90:*913-917 (1993); Staub and Maliga, *EMBO J. 12:*601-606 (1993); U.S. Patents 5, 451,513 and 5,545,818).

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance and helium pressure. One may also minimize the trauma reduction factors by modifying conditions that influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

Agrobacterium-mediated transfer is a widely applicable system for introducing genes into plant cells because the DNA can be introduced into whole plant tissues, thereby bypassing the need for regeneration of an intact plant from a protoplast. The use of *Agrobacterium*-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art. See, for example the methods described by Fraley *et al., Bio/Technology 3*:629-635 (1985) and Rogers *et al., Methods Enzymol. 153*:253-277 (1987). Further, the integration of the Ti-DNA is a relatively precise process resulting in few rearrangements. The region of DNA to be transferred is defined by the border

sequences and intervening DNA is usually inserted into the plant genome as described (Spielmann *et al., Mol. Gen. Genet. 205*:34 (1986)).

Modern *Agrobacterium* transformation vectors are capable of replication in *E. coli* as well as *Agrobacterium*, allowing for convenient manipulations as described (Klee *et al.*, In: *Plant DNA Infectious Agents*, Hohn and Schell (eds.), Springer-Verlag, New York, pp. 179-203 (1985)). Moreover, technological advances in vectors for *Agrobacterium*-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes and are suitable for present purposes (Rogers *et al., Methods Enzymol. 153*:253-277 (1987)). In addition, *Agrobacterium* containing both armed and disarmed Ti genes can be used for the transformations. In those plant strains where *Agrobacterium*-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer.

A transgenic plant formed using *Agrobacterium* transformation methods typically contains a single gene on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. More preferred is a transgenic plant that is homozygous for the added structural gene; *i.e.*, a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an independent segregant, transgenic plant that contains a single added gene, germinating some of the seed produced and analyzing the resulting plants produced for the gene of interest.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating, exogenous genes. Selfing of appropriate progeny can produce plants that are homozygous for both added,

exogenous genes that encode a polypeptide of interest. Back-crossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation.

Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation and combinations of these treatments (*See, for example*, Potrykus *et al., Mol. Gen. Genet. 205*:193-200 (1986); Lorz *et al., Mol. Gen. Genet. 199*:178 (1985); Fromm *et al., Nature 319*:791 (1986); Uchimiya *et al., Mol. Gen. Genet. 204*:204 (1986); Marcotte *et al., Nature 335*:454-457 (1988)).

Application of these systems to different plant strains depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from protoplasts are described (Fujimura *et al., Plant Tissue Culture Letters 2*:74 (1985); Toriyama *et al., Theor. Appl. Genet. 205*:34 (1986); Yamada *et al., Plant Cell Rep. 4*:85 (1986); Abdullah *et al., Biotechnology 4*:1087 (1986)).

To transform plant strains that cannot be successfully regenerated from protoplasts, other ways to introduce DNA into intact cells or tissues can be utilized. For example, regeneration of cereals from immature embryos or explants can be effected as described (Vasil, *Biotechnology 6*:397 (1988)). In addition, "particle gun" or high-velocity microprojectile technology can be utilized (Vasil *et al., Bio/Technology 10*:667 (1992)).

Using the latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al., Nature 328:70* (1987); Klein *et al., Proc. Natl. Acad. Sci. (U.S.A.) 85*:8502-8505 (1988); McCabe *et al., Bio/Technology 6:923* (1988)). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

Other methods of cell transformation can also be used and include but are not limited to introduction of DNA into plants by direct DNA transfer into pollen (Hess *et*

al., Intern Rev. Cytol. 107:367 (1987); Luo *et al., Plant Mol Biol. Reporter 6*:165 (1988)), by direct injection of DNA into reproductive organs of a plant (Pena *et al., Nature 325*:274 (1987)), or by direct injection of DNA into the cells of immature embryos followed by the rehydration of desiccated embryos (Neuhaus *et al., Theor. Appl. Genet. 75*:30 (1987)).

The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: *Methods for Plant Molecular Biology*, Academic Press, San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous gene that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens* and obtaining transgenic plants have been published for cotton (U.S. Patent No. 5,004,863; U.S. Patent No. 5,159,135; U.S. Patent No. 5,518,908); soybean (U.S. Patent No. 5,569,834; U.S. Patent No. 5,416,011; McCabe *et. al., Biotechnology 6*:923 (1988);

Christou *et al., Plant Physiol. 87*:671-674 (1988)); *Brassica* (U.S. Patent No. 5,463,174); peanut (Cheng *et al., Plant Cell Rep. 15:*653-657 (1996), McKently *et al., Plant Cell Rep. 14:*699-703 (1995)); papaya; pea (Grant *et al., Plant Cell Rep. 15*:254-258 (1995)); and *Arabidopsis thaliana* (Bechtold *et al., C.R. Acad. Sci. Paris, Life Sci. 316*:1194-1199 (1993)). The latter method for transforming *Arabidopsis thaliana* is commonly called "dipping" or vacuum infiltration or germplasm transformation.

Transformation of monocotyledons using electroporation, particle bombardment and *Agrobacterium* have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier *et al., Proc. Natl. Acad. Sci. (USA) 84*:5354 (1987)); barley (Wan and Lemaux, *Plant Physiol 104:*37 (1994)); maize (Rhodes *et al., Science 240:*204 (1988); Gordon-Kamm *et al., Plant Cell* 2:603-618 (1990); Fromm *et al., Bio/Technology 8*:833 (1990); Koziel *et al., Bio/Technology 11*:194 (1993); Armstrong *et al., Crop Science 35:*550-557 (1995)); oat (Somers *et al., Bio/Technology 10*:1589 (1992)); orchard grass (Horn *et al., Plant Cell Rep. 7*:469 (1988)); rice (Toriyama *et al., Theor Appl. Genet. 205*:34 (1986); Part *et al., Plant Mol. Biol. 32*:1135-1148 (1996); Abedinia *et al., Aust. J. Plant Physiol. 24:*133-141 (1997); Zhang and Wu, *Theor. Appl. Genet. 76*:835 (1988); Zhang *et al., Plant Cell Rep. 7*:379 (1988); Battraw and Hall, *Plant Sci. 86*:191-202 (1992); Christou *et al., Bio/Technology 9*:957 (1991)); rye (De la Pena *et al., Nature 325*:274 (1987)); sugarcane (Bower and Birch, *Plant J. 2*:409 (1992)); tall fescue (Wang *et al., Bio/Technology 10*:691 (1992)) and wheat (Vasil *et al., Bio/Technology 10*:667 (1992); U.S. Patent No. 5,631,152).

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte *et al., Nature 335*:454-457 (1988); Marcotte *et al., Plant Cell 1:*523-532 (1989); McCarty *et al., Cell 66:*895-905 (1991); Hattori *et al., Genes Dev. 6:*609-618 (1992); Goff *et al., EMBO J. 9:*2517-2522 (1990)). Transient expression systems may be

used to functionally dissect gene constructs (*see generally*, Mailga *et al., Methods in Plant Molecular Biology,* Cold Spring Harbor Press (1995)).

Any of the nucleic acid molecules of the invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers, *etc.* Further, any of the nucleic acid molecules of the invention may be introduced into a plant cell in a manner that allows for expression or overexpression of the protein or fragment thereof encoded by the nucleic acid molecule.

Cosuppression is the reduction in expression levels, usually at the level of RNA, of a particular endogenous gene or gene family by the expression of a homologous sense construct that is capable of transcribing mRNA of the same strandedness as the transcript of the endogenous gene (Napoli *et al., Plant Cell 2*:279-289 (1990); van der Krol *et al., Plant Cell 2*:291-299 (1990)). Cosuppression may result from stable transformation with a single copy nucleic acid molecule that is homologous to a nucleic acid sequence found with the cell (Prolls and Meyer, *Plant J. 2:*465-475 (1992)) or with multiple copies of a nucleic acid molecule that is homologous to a nucleic acid sequence found with the cell (Mittlesten *et al., Mol. Gen. Genet. 244:*325-330 (1994)). Genes, even though different, linked to homologous promoters may result in the cosuppression of the linked genes (Vaucheret, *C.R. Acad. Sci. III 316:*1471-1483 (1993); Flavell, *Proc. Natl. Acad. Sci. (U.S.A.) 91*:3490-3496 (1994)); van Blokland *et al., Plant J. 6*:861-877 (1994); Jorgensen, *Trends Biotechnol. 8*:340-344 (1990); Meins and Kunz, In: *Gene Inactivation and Homologous Recombination in Plants*, Paszkowski (ed.), pp. 335-348, Kluwer Academic, Netherlands (1994)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the cosuppression of an endogenous protein.

Antisense approaches are a way of preventing or reducing gene function by targeting the genetic material (Mol *et al., FEBS Lett. 268*:427-430 (1990)). The objective

of the antisense approach is to use a sequence complementary to the target gene to block its expression and create a mutant cell line or organism in which the level of a single chosen protein is selectively reduced or abolished. Antisense techniques have several advantages over other 'reverse genetic' approaches. The site of inactivation and its developmental effect can be manipulated by the choice of promoter for antisense genes or by the timing of external application or microinjection. Antisense can manipulate its specificity by selecting either unique regions of the target gene or regions where it shares homology to other related genes (Hiatt *et al.,* In: *Genetic Engineering,* Setlow (ed.), Vol. 11, New York: Plenum 49-63 (1989)).

Antisense RNA techniques involve introduction of RNA that is complementary to the target mRNA into cells, which results in specific RNA:RNA duplexes being formed by base pairing between the antisense substrate and the target mRNA (Green *et al., Annu. Rev. Biochem. 55*:569-597 (1986)). Under one embodiment, the process involves the introduction and expression of an antisense gene sequence. Such a sequence is one in which part or all of the normal gene sequences are placed under a promoter in inverted orientation so that the 'wrong' or complementary strand is transcribed into a noncoding antisense RNA that hybridizes with the target mRNA and interferes with its expression (Takayama and Inouye, *Crit. Rev. Biochem. Mol. Biol. 25*:155-184 (1990)). An antisense vector is constructed by standard procedures and introduced into cells by transformation, transfection, electroporation, microinjection, infection, *etc.* The type of transformation and choice of vector will determine whether expression is transient or stable. The promoter used for the antisense gene may influence the level, timing, tissue, specificity, or inducibility of the antisense inhibition.

It is understood that the activity of a protein in a plant cell may be reduced or depressed by growing a transformed plant cell containing a nucleic acid molecule whose non-transcribed strand encodes a protein or fragment thereof.

Posttranscriptional gene silencing (PTGS) can result in virus immunity or gene silencing in plants. PTGS is induced by dsRNA and is mediated by an RNA-dependent RNA polymerase, present in the cytoplasm, that requires a dsRNA template. The dsRNA is formed by hybridization of complementary transgene mRNAs or complementary regions of the same transcript. Duplex formation can be accomplished by using transcripts from one sense gene and one antisense gene colocated in the plant genome, a single transcript that has self-complementarity, or sense and antisense transcripts from genes brought together by crossing. The dsRNA-dependent RNA polymerase makes a complementary strand from the transgene mRNA and RNAse molecules attach to this complementary strand (cRNA). These cRNA-RNase molecules hybridize to the endogene mRNA and cleave the single-stranded RNA adjacent to the hybrid. The cleaved single-stranded RNAs are further degraded by other host RNases because one will lack a capped 5' end and the other will lack a poly(A) tail (Waterhouse *et al., PNAS 95*: 13959-13964 (1998)).

It is understood that one or more of the nucleic acids of the invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the postranscriptional gene silencing of an endogenous transcript.

Antibodies have been expressed in plants (Hiatt *et al., Nature 342:*76-78 (1989); Conrad and Fielder, *Plant Mol. Biol. 26:*1023-1030 (1994)). Cytoplasmic expression of a scFv (single-chain Fv antibody) has been reported to delay infection by artichoke mottled crinkle virus. Transgenic plants that express antibodies directed against endogenous proteins may exhibit a physiological effect (Philips *et al., EMBO J. 16:*4489-4496 (1997); Marion-Poll, *Trends in Plant Science 2:*447-448 (1997)). For example, expressed anti-abscisic antibodies have been reported to result in a general perturbation of seed development (Philips *et al., EMBO J. 16:* 4489-4496 (1997)).

Antibodies that are catalytic may also be expressed in plants (abzymes). The principle behind abzymes is that since antibodies may be raised against many molecules, this recognition ability can be directed toward generating antibodies that bind transition states to force a chemical reaction forward (Persidas, *Nature Biotechnology 15:*1313-1315 (1997); Baca *et al., Ann. Rev. Biophys. Biomol. Struct. 26:*461-493 (1997)). The catalytic abilities of abzymes may be enhanced by site directed mutagenesis. Examples of abzymes are, for example, set forth in U.S. Patent No: 5,658,753; U.S. Patent No. 5,632,990; U.S. Patent No. 5,631,137; U.S. Patent 5,602,015; U.S. Patent No. 5,559,538; U.S. Patent No. 5,576,174; U.S. Patent No. 5,500,358; U.S. Patent 5,318,897; U.S. Patent No. 5,298,409; U.S. Patent No. 5,258,289 and U.S. Patent No. 5,194,585.

It is understood that any of the antibodies of the invention may be expressed in plants and that such expression can result in a physiological effect. It is also understood that any of the expressed antibodies may be catalytic.

Alterations of plant phenotypes of the present invention can be relative to a plant having a similar genetic background that lacks the introduced nucleic acid of interest. In a preferred aspect, a similar genetic background is a background where the organisms being compared share 50% or greater of their nuclear genetic material. In a more preferred aspect a similar genetic background is a background where the organisms being compared share 75% or greater, even more preferably 90% or greater of their nuclear genetic material. In another even more preferable aspect, a similar genetic background is a background where the organisms being compared are plants, and the plants are isogenic except for any genetic material originally introduced using plant transformation techniques.

The present invention also provides for parts of the plants, particularly reproductive or storage parts, of the present invention. Plant parts, without limitation, include seed, endosperm, ovule and pollen. In a particularly preferred embodiment of the

present invention, the plant part is a seed. In one embodiment the seed is a constituent of animal feed.

In another embodiment, the plant part is a fruit, more preferably a fruit with enhanced shelf life. In another preferred embodiment, the fruit has increased levels of a tocopherol.

The present invention also provides a container of over 10,000, more preferably 20,000, and even more preferably 40,000 seeds where over 10%, more preferably 25%, more preferably 50% and even more preferably 75% or 90% of the seeds are seeds derived from a plant of the present invention.

The present invention also provides a container of over 10 kg, more preferably 25 kg, and even more preferably 50 kg seeds where over 10%, more preferably 25%, more preferably 50% and even more preferably 75% or 90% of the seeds are seeds derived from a plant of the present invention.

Any of the plants or parts thereof of the present invention may be processed to produce a feed, meal, protein or oil preparation. A particularly preferred plant part for this purpose is a seed. In a preferred embodiment the feed, meal, protein or oil preparation is designed for ruminant animals. Methods to produce feed, meal, protein and oil preparations are known in the art. See, for example, U.S. Patents 4,957,748, 5,100,679, 5,219,596, 5,936,069, 6,005,076, 6,146,669, and 6,156,227. In a preferred embodiment, the protein preparation is a high protein preparation. Such a high protein preparation preferably has a protein content of greater than 5% w/v, more preferably 10% w/v, and even more preferably 15% w/v. In a preferred oil preparation, the oil preparation is a high oil preparation with an oil content derived from a plant or part thereof of the present invention of greater than 5% w/v, more preferably 10% w/v, and even more preferably 15% w/v. In a preferred embodiment the oil preparation is a liquid and of a volume greater than 1, 5, 10 or 50 liters. The present invention provides for oil produced from plants of the present invention or generated by a method of the present

invention. Such oil may be a minor or major component of any resultant product. Moreover, such oil may be blended with other oils. In a preferred embodiment, the oil produced from plants of the present invention or generated by a method of the present invention constitutes greater than 0.5%, 1%, 5%, 10%, 25%, 50%, 75% or 90% by volume or weight of the oil component of any product. In another embodiment, the oil preparation may be blended and can constitute greater than 10%, 25%, 35%, 50% or 75% of the blend by volume. Oil produced from a plant of the present invention can be admixed with one or more organic solvents or petroleum distillates.

Plants of the present invention can be part of or generated from a breeding program. The choice of breeding method depends on the mode of plant reproduction, the heritability of the trait(s) being improved, and the type of cultivar used commercially (*e.g.*, F_1 hybrid cultivar, pureline cultivar, etc). Selected, non-limiting approaches, for breeding the plants of the present invention are set forth below. A breeding program can be enhanced using marker assisted selection of the progeny of any cross. It is further understood that any commercial and non-commercial cultivars can be utilized in a breeding program. Factors such as, for example, emergence vigor, vegetative vigor, stress tolerance, disease resistance, branching, flowering, seed set, seed size, seed density, standability, and threshability etc. will generally dictate the choice.

For highly heritable traits, a choice of superior individual plants evaluated at a single location will be effective, whereas for traits with low heritability, selection should be based on mean values obtained from replicated evaluations of families of related plants. Popular selection methods commonly include pedigree selection, modified pedigree selection, mass selection, and recurrent selection. In a preferred embodiment a backcross or recurrent breeding program is undertaken.

The complexity of inheritance influences choice of the breeding method. Backcross breeding can be used to transfer one or a few favorable genes for a highly heritable trait into a desirable cultivar. This approach has been used extensively for

breeding disease-resistant cultivars. Various recurrent selection techniques are used to improve quantitatively inherited traits controlled by numerous genes. The use of recurrent selection in self-pollinating crops depends on the ease of pollination, the frequency of successful hybrids from each pollination, and the number of hybrid offspring from each successful cross.

Breeding lines can be tested and compared to appropriate standards in environments representative of the commercial target area(s) for two or more generations. The best lines are candidates for new commercial cultivars; those still deficient in traits may be used as parents to produce new populations for further selection.

One method of identifying a superior plant is to observe its performance relative to other experimental plants and to a widely grown standard cultivar. If a single observation is inconclusive, replicated observations can provide a better estimate of its genetic worth. A breeder can select and cross two or more parental lines, followed by repeated selfing and selection, producing many new genetic combinations.

The development of new cultivars requires the development and selection of varieties, the crossing of these varieties and the selection of superior hybrid crosses. The hybrid seed can be produced by manual crosses between selected male-fertile parents or by using male sterility systems. Hybrids are selected for certain single gene traits such as pod color, flower color, seed yield, pubescence color, or herbicide resistance, which indicate that the seed is truly a hybrid. Additional data on parental lines, as well as the phenotype of the hybrid, influence the breeder's decision whether to continue with the specific hybrid cross.

Pedigree breeding and recurrent selection breeding methods can be used to develop cultivars from breeding populations. Breeding programs combine desirable traits from two or more cultivars or various broad-based sources into breeding pools from which cultivars are developed by selfing and selection of desired phenotypes. New cultivars can be evaluated to determine which have commercial potential.

Pedigree breeding is used commonly for the improvement of self-pollinating crops. Two parents who possess favorable, complementary traits are crossed to produce an F_1. A F_2 population is produced by selfing one or several F_1's. Selection of the best individuals from the best families is carried out. Replicated testing of families can begin in the F_4 generation to improve the effectiveness of selection for traits with low heritability. At an advanced stage of inbreeding (*i.e.*, F_6 and F_7), the best lines or mixtures of phenotypically similar lines are tested for potential release as new cultivars.

Backcross breeding has been used to transfer genes for a simply inherited, highly heritable trait into a desirable homozygous cultivar or inbred line, which is the recurrent parent. The source of the trait to be transferred is called the donor parent. The resulting plant is expected to have the attributes of the recurrent parent (*e.g.*, cultivar) and the desirable trait transferred from the donor parent. After the initial cross, individuals possessing the phenotype of the donor parent are selected and repeatedly crossed (backcrossed) to the recurrent parent. The resulting parent is expected to have the attributes of the recurrent parent (*e.g.*, cultivar) and the desirable trait transferred from the donor parent.

The single-seed descent procedure in the strict sense refers to planting a segregating population, harvesting a sample of one seed per plant, and using the one-seed sample to plant the next generation. When the population has been advanced from the F_2 to the desired level of inbreeding, the plants from which lines are derived will each trace to different F_2 individuals. The number of plants in a population declines each generation due to failure of some seeds to germinate or some plants to produce at least one seed. As a result, not all of the F_2 plants originally sampled in the population will be represented by a progeny when generation advance is completed.

In a multiple-seed procedure, breeders commonly harvest one or more pods from each plant in a population and thresh them together to form a bulk. Part of the bulk is

used to plant the next generation and part is put in reserve. The procedure has been referred to as modified single-seed descent or the pod-bulk technique.

The multiple-seed procedure has been used to save labor at harvest. It is considerably faster to thresh pods with a machine than to remove one seed from each by hand for the single-seed procedure. The multiple-seed procedure also makes it possible to plant the same number of seed of a population each generation of inbreeding.

Descriptions of other breeding methods that are commonly used for different traits and crops can be found in one of several reference books (*e.g.* Fehr, *Principles of Cultivar Development* Vol. 1, pp. 2-3 (1987))).

A transgenic plant of the present invention may also be reproduced using apomixis. Apomixis is a genetically controlled method of reproduction in plants where the embryo is formed without union of an egg and a sperm. There are three basic types of apomictic reproduction: 1) apospory where the embryo develops from a chromosomally unreduced egg in an embryo sac derived from the nucellus, 2) diplospory where the embryo develops from an unreduced egg in an embryo sac derived from the megaspore mother cell, and 3) adventitious embryony where the embryo develops directly from a somatic cell. In most forms of apomixis, pseudogamy or fertilization of the polar nuclei to produce endosperm is necessary for seed viability. In apospory, a nurse cultivar can be used as a pollen source for endosperm formation in seeds. The nurse cultivar does not affect the genetics of the aposporous apomictic cultivar since the unreduced egg of the cultivar develops parthenogenetically, but makes possible endosperm production. Apomixis is economically important, especially in transgenic plants, because it causes any genotype, no matter how heterozygous, to breed true. Thus, with apomictic reproduction, heterozygous transgenic plants can maintain their genetic fidelity throughout repeated life cycles. Methods for the production of apomictic plants are known in the art. *See*, U.S. Patent No. 5,811,636.

OTHER ORGANISMS

A nucleic acid of the present invention may be introduced into any cell or organism such as a mammalian cell, mammal, fish cell, fish, bird cell, bird, algae cell, algae, fungal cell, fungi, or bacterial cell. A protein of the present invention may be produced in an appropriate cell or organism. Preferred host and transformants include: fungal cells such as *Aspergillus*, yeasts, mammals, particularly bovine and porcine, insects, bacteria, and algae. Methods to transform such cells or organisms are known in the art (EP 0 238 023; Yelton *et al., Proc. Natl. Acad. Sci. (U.S.A.), 81*:1470-1474 (1984); Malardier *et al., Gene, 78*:147-156 (1989); Becker and Guarente, In: Abelson and Simon (eds.), *Guide to Yeast Genetics and Molecular Biology, Method Enzymol.,* Vol. 194, pp. 182-187, Academic Press, Inc., New York; Ito et al., *J. Bacteriology, 153*:163 (1983) Hinnen *et al., Proc. Natl. Acad. Sci. (U.S.A.), 75*:1920 (1978); Bennett and LaSure (eds.), *More Gene Manipualtionins in fungi,* Academic Press, CA (1991)). Methods to produce proteins of the present invention are also known (Kudla *et al., EMBO, 9*:1355-1364 (1990); Jarai and Buxton, *Current Genetics, 26*:2238-2244 (1994); Verdier, *Yeast, 6*:271-297 (1990; MacKenzie *et al., Journal of Gen. Microbiol., 139*:2295-2307 (1993); Hartl *et al., TIBS, 19*:20-25 (1994); Bergenron *et al., TIBS, 19*:124-128 (1994); Demolder *et al., J. Biotechnology, 32*:179-189 (1994); Craig, *Science, 260:*1902-1903 (1993); Gething and Sambrook, *Nature, 355*:33-45 (1992); Puig and Gilbert, J. *Biol. Chem., 269*:7764-7771 (1994); Wang and Tsou, *FASEB Journal, 7*:1515-1517 (1993); Robinson *et al., Bio/Technology, 1*:381-384 (1994); Enderlin and Ogrydziak, *Yeast, 10*:67-79 (1994); Fuller *et al., Proc. Natl. Acad. Sci. (U.S.A.), 86*:1434-1438 (1989); Julius *et al., Cell, 37*:1075-1089 (1984); Julius *et al., Cell 32*:839-852 (1983).

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to a cell or organism with a similar genetic background, an increased level of tocotrienols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to an untransformed cell or organism with a similar genetic background, an increased level of tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to an untransformed cell or organism with a similar genetic background, an increased level of α-tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to an untransformed cell or organism with a similar genetic background, an increased level of γ-tocopherols.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to an untransformed cell or organism with a similar genetic background, an increased level of homogentisic acid.

In a preferred embodiment, overexpression of a protein or fragment thereof of the present invention in a cell or organism provides in that cell or organism, relative to an untransformed cell or organism with a similar genetic background, an increased level of plastoquinols or plastoquinones.

ANTIBODIES

One aspect of the invention concerns antibodies, single-chain antigen binding molecules, or other proteins that specifically bind to one or more of the protein or peptide molecules of the invention and their homologs, fusions or fragments. In a particularly preferred embodiment, the antibody specifically binds to a protein having the amino acid sequence set forth in SEQ ID NOs: 2 and 4 or a fragment thereof. In another

embodiment, the antibody specifically binds to a fusion protein comprising an amino acid sequence selected from the amino acid sequence set forth in SEQ ID NOs: 2 or 4 or a fragment thereof. In another embodiment the antibody specifically binds to a fusion protein comprising an amino acid sequence selected from the amino acid sequence set forth in SEQ ID NOs: 2 or 4 or a fragment thereof. Antibodies of the invention may be used to quantitatively or qualitatively detect the protein or peptide molecules of the invention, or to detect post translational modifications of the proteins. As used herein, an antibody or peptide is said to "specifically bind" to a protein or peptide molecule of the invention if such binding is not competitively inhibited by the presence of non-related molecules.

Nucleic acid molecules that encode all or part of the protein of the invention can be expressed, via recombinant means, to yield protein or peptides that can in turn be used to elicit antibodies that are capable of binding the expressed protein or peptide. Such antibodies may be used in immunoassays for that protein. Such protein-encoding molecules, or their fragments may be a "fusion" molecule (*i.e.*, a part of a larger nucleic acid molecule) such that, upon expression, a fusion protein is produced. It is understood that any of the nucleic acid molecules of the invention may be expressed, via recombinant means, to yield proteins or peptides encoded by these nucleic acid molecules.

The antibodies that specifically bind proteins and protein fragments of the invention may be polyclonal or monoclonal and may comprise intact immunoglobulins, or antigen binding portions of immunoglobulins fragments (such as (F(ab'), $F(ab')_2$), or single-chain immunoglobulins producible, for example, via recombinant means. It is understood that practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (*see*, for example, Harlow and Lane, In: *Antibodies: A Laboratory Manual,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1988)).

As discussed below, such antibody molecules or their fragments may be used for diagnostic purposes. Where the antibodies are intended for diagnostic purposes, it may be desirable to derivatize them, for example with a ligand group (such as biotin) or a detectable marker group (such as a fluorescent group, a radioisotope or an enzyme).

The ability to produce antibodies that bind the protein or peptide molecules of the invention permits the identification of mimetic compounds derived from those molecules. These mimetic compounds may contain a fragment of the protein or peptide or merely a structurally similar region and nonetheless exhibits an ability to specifically bind to antibodies directed against that compound.

EXEMPLARY USES

Nucleic acid molecules and fragments thereof of the invention may be employed to obtain other nucleic acid molecules from the same species (nucleic acid molecules from maize may be utilized to obtain other nucleic acid molecules from maize). Such nucleic acid molecules include the nucleic acid molecules that encode the complete coding sequence of a protein and promoters and flanking sequences of such molecules. In addition, such nucleic acid molecules include nucleic acid molecules that encode for other isozymes or gene family members. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries. Methods for forming such libraries are well known in the art.

Nucleic acid molecules and fragments thereof of the invention may also be employed to obtain nucleic acid homologs. Such homologs include the nucleic acid molecules of plants and other organisms, including bacteria and fungi, including the nucleic acid molecules that encode, in whole or in part, protein homologues of other plant species or other organisms, sequences of genetic elements, such as promoters and transcriptional regulatory elements. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries obtained from such plant species. Methods for forming such libraries are well

known in the art. Such homolog molecules may differ in their nucleotide sequences from those found in one or more of SEQ ID NOs: 1 and 3 and complements thereof because complete complementarity is not needed for stable hybridization. The nucleic acid molecules of the invention therefore also include molecules that, although capable of specifically hybridizing with the nucleic acid molecules may lack "complete complementarity."

Any of a variety of methods may be used to obtain one or more of the above-described nucleic acid molecules (Zamechik *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 83*:4143-4146 (1986); Goodchild *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:5507-5511 (1988); Wickstrom *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:1028-1032 (1988); Holt *et al.*, *Molec. Cell. Biol. 8*:963-973 (1988); Gerwirtz *et al.*, *Science 242*:1303-1306 (1988); Anfossi *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 86*:3379-3383 (1989); Becker *et al.*, *EMBO J. 8*:3685-3691 (1989)). Automated nucleic acid synthesizers may be employed for this purpose. In lieu of such synthesis, the disclosed nucleic acid molecules may be used to define a pair of primers that can be used with the polymerase chain reaction (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol. 51*:263-273 (1986); Erlich *et al.*, European Patent 50,424; European Patent 84,796; European Patent 258,017; European Patent 237,362; Mullis, European Patent 201,184; Mullis *et al.*, U.S. Patent 4,683,202; Erlich, U.S. Patent 4,582,788; and Saiki *et al.*, U.S. Patent 4,683,194) to amplify and obtain any desired nucleic acid molecule or fragment.

Promoter sequences and other genetic elements, including but not limited to transcriptional regulatory flanking sequences, associated with one or more of the disclosed nucleic acid sequences can also be obtained using the disclosed nucleic acid sequence provided herein. In one embodiment, such sequences are obtained by incubating nucleic acid molecules of the present invention with members of genomic libraries and recovering clones that hybridize to such nucleic acid molecules thereof. In a second embodiment, methods of "chromosome walking," or inverse PCR may be used to

obtain such sequences (Frohman *et al., Proc. Natl. Acad. Sci. (U.S.A.) 85*:8998-9002 (1988); Ohara *et al., Proc. Natl. Acad. Sci. (U.S.A.) 86*:5673-5677 (1989); Pang *et al., Biotechniques 22*:1046-1048 (1977); Huang *et al., Methods Mol. Biol. 69*:89-96 (1997); Huang *et al., Method Mol. Biol. 67*:287-294 (1997); Benkel *et al., Genet. Anal. 13*:123-127 (1996); Hartl *et al., Methods Mol. Biol. 58*:293-301 (1996)). The term "chromosome walking" means a process of extending a genetic map by successive hybridization steps.

The nucleic acid molecules of the invention may be used to isolate promoters of cell enhanced, cell specific, tissue enhanced, tissue specific, developmentally or environmentally regulated expression profiles. Isolation and functional analysis of the 5' flanking promoter sequences of these genes from genomic libraries, for example, using genomic screening methods and PCR techniques would result in the isolation of useful promoters and transcriptional regulatory elements. These methods are known to those of skill in the art and have been described (See, for example, Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, (1997), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). Promoters obtained utilizing the nucleic acid molecules of the invention could also be modified to affect their control characteristics. Examples of such modifications would include but are not limited to enhancer sequences. Such genetic elements could be used to enhance gene expression of new and existing traits for crop improvement.

Another subset of the nucleic acid molecules of the invention includes nucleic acid molecules that are markers. The markers can be used in a number of conventional ways in the field of molecular genetics. Such markers include nucleic acid molecules SEQ ID NOs: 1 and 3, complements thereof, and fragments of either that can act as markers and other nucleic acid molecules of the present invention that can act as markers.

Genetic markers of the invention include "dominant" or "codominant" markers. "Codominant markers" reveal the presence of two or more alleles (two per diploid individual) at a locus. "Dominant markers" reveal the presence of only a single allele per locus. The presence of the dominant marker phenotype (e.g., a band of DNA) is an

indication that one allele is in either the homozygous or heterozygous condition. The absence of the dominant marker phenotype (*e.g.*, absence of a DNA band) is merely evidence that "some other" undefined allele is present. In the case of populations where individuals are predominantly homozygous and loci are predominately dimorphic, dominant and codominant markers can be equally valuable. As populations become more heterozygous and multi-allelic, codominant markers often become more informative of the genotype than dominant markers. Marker molecules can be, for example, capable of detecting polymorphisms such as single nucleotide polymorphisms (SNPs).

The genomes of animals and plants naturally undergo spontaneous mutation in the course of their continuing evolution (Gusella, *Ann. Rev. Biochem. 55*:831-854 (1986)). A "polymorphism" is a variation or difference in the sequence of the gene or its flanking regions that arises in some of the members of a species. The variant sequence and the "original" sequence co-exist in the species' population. In some instances, such co-existence is in stable or quasi-stable equilibrium.

A polymorphism is thus said to be "allelic," in that, due to the existence of the polymorphism, some members of a population may have the original sequence (i.e., the original "allele") whereas other members may have the variant sequence (i.e., the variant "allele"). In the simplest case, only one variant sequence may exist and the polymorphism is thus said to be di-allelic. In other cases, the species' population may contain multiple alleles and the polymorphism is termed tri-allelic, etc. A single gene may have multiple different unrelated polymorphisms. For example, it may have a di-allelic polymorphism at one site and a multi-allelic polymorphism at another site.

The variation that defines the polymorphism may range from a single nucleotide variation to the insertion or deletion of extended regions within a gene. In some cases, the DNA sequence variations are in regions of the genome that are characterized by short tandem repeats (STRs) that include tandem di- or tri-nucleotide repeated motifs of

nucleotides. Polymorphisms characterized by such tandem repeats are referred to as "variable number tandem repeat" ("VNTR") polymorphisms. VNTRs have been used in identity analysis (Weber, U.S. Patent 5,075,217; Armour *et al., FEBS Lett. 307*:113-115 (1992); Jones *et al., Eur. J. Haematol. 39*:144-147 (1987); Horn *et al.*, PCT Patent Application WO91/14003; Jeffreys, European Patent Application 370,719; Jeffreys, U.S. Patent 5,175,082; Jeffreys *et al., Amer. J. Hum. Genet. 39*:11-24 (1986); Jeffreys *et al., Nature 316*:76-79 (1985); Gray *et al., Proc. R. Acad. Soc. Lond. 243*:241-253 (1991); Moore *et al., Genomics 10*:654-660 (1991); Jeffreys *et al., Anim. Genet. 18*:1-15 (1987); Hillel *et al., Anim. Genet. 20*:145-155 (1989); Hillel *et al., Genet. 124*:783-789 (1990)).

The detection of polymorphic sites in a sample of DNA may be facilitated through the use of nucleic acid amplification methods. Such methods specifically increase the concentration of polynucleotides that span the polymorphic site, or include that site and sequences located either distal or proximal to it. Such amplified molecules can be readily detected by gel electrophoresis or other means.

In an alternative embodiment, such polymorphisms can be detected through the use of a marker nucleic acid molecule that is physically linked to such polymorphism(s). For this purpose, marker nucleic acid molecules comprising a nucleotide sequence of a polynucleotide located within 1 mb of the polymorphism(s) and more preferably within 100kb of the polymorphism(s) and most preferably within 10kb of the polymorphism(s) can be employed.

The identification of a polymorphism can be determined in a variety of ways. By correlating the presence or absence of it in a plant with the presence or absence of a phenotype, it is possible to predict the phenotype of that plant. If a polymorphism creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (e.g., a VNTR polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, organisms that possess a variant sequence can be distinguished from those having the

original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed "restriction fragment length polymorphisms" ("RFLPs") (Glassberg, UK Patent Application 2135774; Skolnick *et al.*, *Cytogen. Cell Genet. 32*:58-67 (1982); Botstein *et al.*, *Ann. J. Hum. Genet. 32*:314-331 (1980); Fischer *et al.*, (PCT Application WO90/13668; Uhlen, PCT Application WO90/11369).

Polymorphisms can also be identified by Single Strand Conformation Polymorphism (SSCP) analysis (Elles, *Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases*, Humana Press (1996)); Orita *et al., Genomics 5:*874-879 (1989)). A number of protocols have been described for SSCP including, but not limited to, Lee *et al., Anal. Biochem. 205:*289-293 (1992); Suzuki *et al., Anal. Biochem. 192:*82-84 (1991); Lo *et al., Nucleic Acids Research 20:*1005-1009 (1992); Sarkar *et al., Genomics 13:*441-443 (1992). It is understood that one or more of the nucleic acids of the invention, may be utilized as markers or probes to detect polymorphisms by SSCP analysis.

Polymorphisms may also be found using a DNA fingerprinting technique called amplified fragment length polymorphism (AFLP), which is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA to profile that DNA (Vos *et al.*, *Nucleic Acids Res. 23*:4407-4414 (1995)). This method allows for the specific co-amplification of high numbers of restriction fragments, which can be visualized by PCR without knowledge of the nucleic acid sequence. It is understood that one or more of the nucleic acids of the invention may be utilized as markers or probes to detect polymorphisms by AFLP analysis or for fingerprinting RNA.

Polymorphisms may also be found using random amplified polymorphic DNA (RAPD) (Williams *et al., Nucl. Acids Res. 18*:6531-6535 (1990)) and cleaveable amplified polymorphic sequences (CAPS) (Lyamichev *et al., Science 260*:778-783 (1993)). It is understood that one or more of the nucleic acid molecules of the invention,

may be utilized as markers or probes to detect polymorphisms by RAPD or CAPS analysis.

Single Nucleotide Polymorphisms (SNPs) generally occur at greater frequency than other polymorphic markers and are spaced with a greater uniformity throughout a genome than other reported forms of polymorphism. The greater frequency and uniformity of SNPs means that there is greater probability that such a polymorphism will be found near or in a genetic locus of interest than would be the case for other polymorphisms. SNPs are located in protein-coding regions and noncoding regions of a genome. Some of these SNPs may result in defective or variant protein expression (*e.g.*, as a result of mutations or defective splicing). Analysis (genotyping) of characterized SNPs can require only a plus/minus assay rather than a lengthy measurement, permitting easier automation.

SNPs can be characterized using any of a variety of methods. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes (Botstein *et al.*, *Am. J. Hum. Genet. 32*:314-331 (1980); Konieczny and Ausubel, *Plant J. 4*:403-410 (1993)), enzymatic and chemical mismatch assays (Myers *et al.*, *Nature 313*:495-498 (1985)), allele-specific PCR (Newton *et al.*, *Nucl. Acids Res. 17*:2503-2516 (1989); Wu *et al.*, *Proc. Natl. Acad. Sci. USA 86*:2757-2760 (1989)), ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. USA 88*:189-193 (1991)), single-strand conformation polymorphism analysis (Labrune *et al.*, *Am. J. Hum. Genet. 48*: 1115-1120 (1991)), single base primer extension (Kuppuswamy *et al.*, *Proc. Natl. Acad. Sci. USA 88*:1143-1147 (1991), Goelet US 6,004,744; Goelet 5,888,819), solid-phase ELISA-based oligonucleotide ligation assays (Nikiforov *et al.*, *Nucl. Acids Res. 22*:4167-4175 (1994), dideoxy fingerprinting (Sarkar *et al.*, *Genomics 13*:441-443 (1992)), oligonucleotide fluorescence-quenching assays (Livak *et al.*, *PCR Methods Appl. 4*:357-362 (1995a)), 5'-nuclease allele-specific hybridization TaqMan™ assay (Livak *et al.*, *Nature Genet. 9*:341-342 (1995)), template-directed dye-terminator incorporation (TDI) assay (Chen

and Kwok, *Nucl. Acids Res. 25*:347-353 (1997)), allele-specific molecular beacon assay (Tyagi *et al.*, *Nature Biotech. 16*: 49-53 (1998)), PinPoint assay (Haff and Smirnov, *Genome Res. 7*: 378-388 (1997)), dCAPS analysis (Neff *et al.*, *Plant J. 14*:387-392 (1998)), pyrosequencing (Ronaghi *et al, Analytical Biochemistry 267*:65-71 (1999); Ronaghi *et al* PCT application WO 98/13523; Nyren *et al* PCT application WO 98/28440; http//www.pyrosequencing.com), using mass spectrometry, *e.g.* the Masscode ™ system (Howbert *et al* WO 99/05319; Howber *et al* WO 97/27331; http//www.rapigene.com; Becker *et al* PCT application WO 98/26095; Becker *et al* PCT application; WO 98/12355; Becker *et al* PCT application WO 97/33000; Monforte *et al* US 5,965,363), invasive cleavage of oligonucleotide probes (Lyamichev *et al Nature Biotechnology 17*:292-296; http//www.twt.com), and using high density oligonucleotide arrays (Hacia *et al Nature Genetics 22*:164-167; http//www.affymetrix.com).

Polymorphisms may also be detected using allele-specific oligonucleotides (ASO), which, can be for example, used in combination with hybridization based technology including southern, northern, and dot blot hybridizations, reverse dot blot hybridizations and hybridizations performed on microarray and related technology.

The stringency of hybridization for polymorphism detection is highly dependent upon a variety of factors, including length of the allele-specific oligonucleotide, sequence composition, degree of complementarity (*i.e.* presence or absence of base mismatches), concentration of salts and other factors such as formamide, and temperature. These factors are important both during the hybridization itself and during subsequent washes performed to remove target polynucleotide that is not specifically hybridized. In practice, the conditions of the final, most stringent wash are most critical. In addition, the amount of target polynucleotide that is able to hybridize to the allele-specific oligonucleotide is also governed by such factors as the concentration of both the ASO and the target polynucleotide, the presence and concentration of factors that act to "tie up" water molecules, so as to effectively concentrate the reagents (*e.g.*, PEG, dextran, dextran

sulfate, *etc.*), whether the nucleic acids are immobilized or in solution, and the duration of hybridization and washing steps.

Hybridizations are preferably performed below the melting temperature (T_m) of the ASO. The closer the hybridization and/or washing step is to the T_m, the higher the stringency. T_m for an oligonucleotide may be approximated, for example, according to the following formula: $T_m = 81.5 + 16.6 \times (\log10[Na+]) + 0.41 \times (\%G+C) - 675/n$; where [Na+] is the molar salt concentration of Na+ or any other suitable cation and n = number of bases in the oligonucleotide. Other formulas for approximating T_m are available and are known to those of ordinary skill in the art.

Stringency is preferably adjusted so as to allow a given ASO to differentially hybridize to a target polynucleotide of the correct allele and a target polynucleotide of the incorrect allele. Preferably, there will be at least a two-fold differential between the signal produced by the ASO hybridizing to a target polynucleotide of the correct allele and the level of the signal produced by the ASO cross-hybridizing to a target polynucleotide of the incorrect allele (*e.g.*, an ASO specific for a mutant allele cross-hybridizing to a wild-type allele). In more preferred embodiments of the present invention, there is at least a five-fold signal differential. In highly preferred embodiments of the present invention, there is at least an order of magnitude signal differential between the ASO hybridizing to a target polynucleotide of the correct allele and the level of the signal produced by the ASO cross-hybridizing to a target polynucleotide of the incorrect allele.

While certain methods for detecting polymorphisms are described herein, other detection methodologies may be utilized. For example, additional methodologies are known and set forth, in Birren *et al.*, *Genome Analysis,* 4:135-186, *A Laboratory Manual. Mapping Genomes*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1999); Maliga *et al.*, *Methods in Plant Molecular Biology. A Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1995);

Paterson, *Biotechnology Intelligence Unit: Genome Mapping in Plants*, R.G. Landes Co., Georgetown, TX, and Academic Press, San Diego, CA (1996); *The Maize Handbook*, Freeling and Walbot, eds., Springer-Verlag, New York, NY (1994); *Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases*, Elles, ed., Humana Press, Totowa, NJ (1996); Clark, ed., *Plant Molecular Biology: A Laboratory Manual*, Clark, ed., Springer-Verlag, Berlin, Germany (1997).

Requirements for marker-assisted selection in a plant breeding program are: (1) the marker(s) should co-segregate or be closely linked with the desired trait; (2) an efficient means of screening large populations for the molecular marker(s) should be available; and (3) the screening technique should have high reproducibility across laboratories and preferably be economical to use and be user-friendly.

The genetic linkage of marker molecules can be established by a gene mapping model such as, without limitation, the flanking marker model reported by Lander and Botstein, *Genetics 121*:185-199 (1989) and the interval mapping, based on maximum likelihood methods described by Lander and Botstein, *Genetics 121*:185-199 (1989) and implemented in the software package MAPMAKER/QTL (Lincoln and Lander, *Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL,* Whitehead Institute for Biomedical Research, Massachusetts, (1990). Additional software includes Qgene, Version 2.23 (1996), Department of Plant Breeding and Biometry, 266 Emerson Hall, Cornell University, Ithaca, NY). Use of Qgene software is a particularly preferred approach.

A maximum likelihood estimate (MLE) for the presence of a marker is calculated, together with an MLE assuming no QTL effect, to avoid false positives. A $\log_{10}$ of an odds ratio (LOD) is then calculated as: LOD = $\log_{10}$ (MLE for the presence of a QTL/MLE given no linked QTL).

The LOD score essentially indicates how much more likely the data are to have arisen assuming the presence of a QTL than in its absence. The LOD threshold value for

avoiding a false positive with a given confidence, say 95%, depends on the number of markers and the length of the genome. Graphs indicating LOD thresholds are set forth in Lander and Botstein, *Genetics 121*:185-199 (1989) and further described by Arús and Moreno-González , *Plant Breeding*, Hayward *et al.,* (eds.) Chapman & Hall, London, pp. 314-331 (1993).

In a preferred embodiment of the present invention the nucleic acid marker exhibits a LOD score of greater than 2.0, more preferably 2.5, even more preferably greater than 3.0 or 4.0 with the trait or phenotype of interest. In a preferred embodiment, the trait of interest is altered tocopherol levels or compositions.

Additional models can be used. Many modifications and alternative approaches to interval mapping have been reported, including the use non-parametric methods (Kruglyak and Lander, *Genetics 139:*1421-1428 (1995)). Multiple regression methods or models can be also be used, in which the trait is regressed on a large number of markers (Jansen, *Biometrics in Plant Breeding*, van Oijen and Jansen (eds.), Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp. 116-124 (1994); Weber and Wricke, *Advances in Plant Breeding,* Blackwell, Berlin, 16 (1994)). Procedures combining interval mapping with regression analysis, whereby the phenotype is regressed onto a single putative QTL at a given marker interval and at the same time onto a number of markers that serve as 'cofactors,' have been reported by Jansen and Stam, *Genetics 136*:1447-1455 (1994), and Zeng, *Genetics 136*:1457-1468 (1994). Generally, the use of cofactors reduces the bias and sampling error of the estimated QTL positions (Utz and Melchinger, *Biometrics in Plant Breeding,* van Oijen and Jansen (eds.) Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp.195-204 (1994), thereby improving the precision and efficiency of QTL mapping (Zeng, *Genetics 136*:1457-1468 (1994)). These models can be extended to multi-environment experiments to analyze genotype-environment interactions (Jansen *et al., Theo. Appl. Genet. 91:*33-37 (1995)).

It is understood that one or more of the nucleic acid molecules of the invention may be used as molecular markers. It is also understood that one or more of the protein molecules of the invention may be used as molecular markers.

In a preferred embodiment, the polymorphism is present and screened for in a mapping population, *e.g.* a collection of plants capable of being used with markers such as polymorphic markers to map genetic position of traits. The choice of appropriate mapping population often depends on the type of marker systems employed (Tanksley *et al., J.P. Gustafson and R. Appels* (eds.). Plenum Press, New York, pp. 157-173 (1988)). Consideration must be given to the source of parents (adapted vs. exotic) used in the mapping population. Chromosome pairing and recombination rates can be severely disturbed (suppressed) in wide crosses (adapted x exotic) and generally yield greatly reduced linkage distances. Wide crosses will usually provide segregating populations with a relatively large number of polymorphisms when compared to progeny in a narrow cross (adapted x adapted).

An F_2 population is the first generation of selfing (self-pollinating) after the hybrid seed is produced. Usually a single F_1 plant is selfed to generate a population segregating for all the genes in Mendelian (1:2:1) pattern. Maximum genetic information is obtained from a completely classified F_2 population using a codominant marker system (Mather, Measurement of Linkage in Heredity: Methuen and Co., (1938)). In the case of dominant markers, progeny tests (*e.g.*, F_3, BCF_2) are required to identify the heterozygotes, in order to classify the population. However, this procedure is often prohibitive because of the cost and time involved in progeny testing. Progeny testing of F_2 individuals is often used in map construction where phenotypes do not consistently reflect genotype (*e.g.* disease resistance) or where trait expression is controlled by a QTL. Segregation data from progeny test populations *e.g.* F_3 or BCF_2) can be used in map construction. Marker-assisted selection can then be applied to cross progeny based on

marker-trait map associations (F_2, F_3), where linkage groups have not been completely disassociated by recombination events (*i.e.*, maximum disequilibrium).

Recombinant inbred lines (RIL) (genetically related lines; usually >F_5, developed from continuously selfing F_2 lines towards homozygosity) can be used as a mapping population. Information obtained from dominant markers can be maximized by using RIL because all loci are homozygous or nearly so. Under conditions of tight linkage (*i.e.*, about <10% recombination), dominant and co-dominant markers evaluated in RIL populations provide more information per individual than either marker type in backcross populations (Reiter. *Proc. Natl. Acad. Sci. (U.S.A.) 89*:1477-1481 (1992)). However, as the distance between markers becomes larger (*i.e.*, loci become more independent), the information in RIL populations decreases dramatically when compared to codominant markers.

Backcross populations (*e.g.*, generated from a cross between a successful variety (recurrent parent) and another variety (donor parent) carrying a trait not present in the former) can be utilized as a mapping population. A series of backcrosses to the recurrent parent can be made to recover most of its desirable traits. Thus a population is created consisting of individuals nearly like the recurrent parent but each individual carries varying amounts or mosaic of genomic regions from the donor parent. Backcross populations can be useful for mapping dominant markers if all loci in the recurrent parent are homozygous and the donor and recurrent parent have contrasting polymorphic marker alleles (Reiter *et al., Proc. Natl. Acad. Sci. (U.S.A.) 89*:1477-1481 (1992)). Information obtained from backcross populations using either codominant or dominant markers is less than that obtained from F_2 populations because one, rather than two, recombinant gamete is sampled per plant. Backcross populations, however, are more informative (at low marker saturation) when compared to RILs as the distance between linked loci increases in RIL populations (*i.e.* about 0.15% recombination). Increased recombination can be

beneficial for resolution of tight linkages, but may be undesirable in the construction of maps with low marker saturation.

Near-isogenic lines (NIL) (created by many backcrosses to produce a collection of individuals that is nearly identical in genetic composition except for the trait or genomic region under interrogation) can be used as a mapping population. In mapping with NILs, only a portion of the polymorphic loci is expected to map to a selected region.

Bulk segregant analysis (BSA) is a method developed for the rapid identification of linkage between markers and traits of interest (Michelmore *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*:9828-9832 (1991)). In BSA, two bulked DNA samples are drawn from a segregating population originating from a single cross. These bulks contain individuals that are identical for a particular trait (resistant or susceptible to particular disease) or genomic region but arbitrary at unlinked regions (*i.e.* heterozygous). Regions unlinked to the target region will not differ between the bulked samples of many individuals in BSA.

In an aspect of the present invention, one or more of the nucleic molecules of the present invention are used to determine the level (*i.e.*, the concentration of mRNA in a sample, *etc.*) in a plant (preferably canola, maize, *Brassica campestris, Brassica napus*, soybean, crambe, mustard, castor bean, peanut, sesame, cottonseed, linseed, safflower, oil palm, flax or sunflower) or pattern (*i.e.*, the kinetics of expression, rate of decomposition, stability profile, *etc.*) of the expression of a protein encoded in part or whole by one or more of the nucleic acid molecule of the present invention (collectively, the "Expression Response" of a cell or tissue).

As used herein, the Expression Response manifested by a cell or tissue is said to be "altered" if it differs from the Expression Response of cells or tissues of plants not exhibiting the phenotype. To determine whether a Expression Response is altered, the Expression Response manifested by the cell or tissue of the plant exhibiting the phenotype is compared with that of a similar cell or tissue sample of a plant not exhibiting the phenotype. As will be appreciated, it is not necessary to re-determine the

Expression Response of the cell or tissue sample of plants not exhibiting the phenotype each time such a comparison is made; rather, the Expression Response of a particular plant may be compared with previously obtained values of normal plants. As used herein, the phenotype of the organism is any of one or more characteristics of an organism (*e.g.* disease resistance, pest tolerance, environmental tolerance such as tolerance to abiotic stress, male sterility, quality improvement or yield *etc.*). A change in genotype or phenotype may be transient or permanent. Also as used herein, a tissue sample is any sample that comprises more than one cell. In a preferred aspect, a tissue sample comprises cells that share a common characteristic (e.g. derived from root, seed, flower, leaf, stem or pollen etc.).

In one aspect of the present invention, an evaluation can be conducted to determine whether a particular mRNA molecule is present. One or more of the nucleic acid molecules of the present invention are utilized to detect the presence or quantity of the mRNA species. Such molecules are then incubated with cell or tissue extracts of a plant under conditions sufficient to permit nucleic acid hybridization. The detection of double-stranded probe-mRNA hybrid molecules is indicative of the presence of the mRNA; the amount of such hybrid formed is proportional to the amount of mRNA. Thus, such probes may be used to ascertain the level and extent of the mRNA production in a plant's cells or tissues. Such nucleic acid hybridization may be conducted under quantitative conditions (thereby providing a numerical value of the amount of the mRNA present). Alternatively, the assay may be conducted as a qualitative assay that indicates either that the mRNA is present, or that its level exceeds a user set, predefined value.

A number of methods can be used to compare the expression response between two or more samples of cells or tissue. These methods include hybridization assays, such as northerns, RNAse protection assays, and *in situ* hybridization. Alternatively, the methods include PCR-type assays. In a preferred method, the expression response is compared by hybridizing nucleic acids from the two or more samples to an array of

nucleic acids. The array contains a plurality of suspected sequences known or suspected of being present in the cells or tissue of the samples.

An advantage of *in situ* hybridization over more conventional techniques for the detection of nucleic acids is that it allows an investigator to determine the precise spatial population (Angerer *et al., Dev. Biol. 101*:477-484 (1984); Angerer *et al., Dev. Biol. 112*:157-166 (1985); Dixon *et al., EMBO J. 10*:1317-1324 (1991)). *In situ* hybridization may be used to measure the steady-state level of RNA accumulation (Hardin *et al., J. Mol. Biol. 202*:417-431 (1989)). A number of protocols have been devised for *in situ* hybridization, each with tissue preparation, hybridization and washing conditions (Meyerowitz, *Plant Mol. Biol. Rep. 5*:242-250 (1987); Cox and Goldberg, In: *Plant Molecular Biology: A Practical Approach,* Shaw (ed.), pp. 1-35, IRL Press, Oxford (1988); Raikhel *et al., In situ RNA hybridization in plant tissues*, In: *Plant Molecular Biology Manual*, vol. B9:1-32, Kluwer Academic Publisher, Dordrecht, Belgium (1989)).

In situ hybridization also allows for the localization of proteins within a tissue or cell (Wilkinson, *In Situ Hybridization,* Oxford University Press, Oxford (1992); Langdale, *In Situ Hybridization* In: *The Maize Handbook,* Freeling and Walbot (eds.), pp. 165-179, Springer-Verlag, New York (1994)). It is understood that one or more of the molecules of the invention, preferably one or more of the nucleic acid molecules or fragments thereof of the invention or one or more of the antibodies of the invention may be utilized to detect the level or pattern of a protein or mRNA thereof by *in situ* hybridization.

Fluorescent *in situ* hybridization allows the localization of a particular DNA sequence along a chromosome, which is useful, among other uses, for gene mapping, following chromosomes in hybrid lines, or detecting chromosomes with translocations, transversions or deletions. *In situ* hybridization has been used to identify chromosomes in several plant species (Griffor *et al., Plant Mol. Biol. 17*:101-109 (1991); Gustafson *et al., Proc. Natl. Acad. Sci. (U.S.A.) 87*:1899-1902 (1990); Mukai and Gill, *Genome*

34:448-452 (1991); Schwarzacher and Heslop-Harrison, *Genome 34*:317-323 (1991); Wang *et al., Jpn. J. Genet. 66*:313-316 (1991); Parra and Windle, *Nature Genetics 5*:17-21 (1993)). It is understood that the nucleic acid molecules of the invention may be used as probes or markers to localize sequences along a chromosome.

Another method to localize the expression of a molecule is tissue printing. Tissue printing provides a way to screen, at the same time on the same membrane many tissue sections from different plants or different developmental stages (Yomo and Taylor, *Planta 112*:35-43 (1973); Harris and Chrispeels, *Plant Physiol.* 56:292-299 (1975); Cassab and Varner, *J. Cell. Biol. 105*:2581-2588 (1987); Spruce *et al., Phytochemistry 26:*2901-2903 (1987); Barres *et al., Neuron 5:*527-544 (1990); Reid and Pont-Lezica, *Tissue Printing: Tools for the Study of Anatomy, Histochemistry and Gene Expression,* Academic Press, New York, New York (1992); Reid *et al., Plant Physiol. 93:*160-165 (1990); Ye *et al., Plant J. 1:*175-183 (1991)).

One skilled in the art can refer to general reference texts for detailed descriptions of known techniques discussed herein or equivalent techniques. These texts include *Current Protocols in Molecular Biology* Ausubel, *et al.,* eds., John Wiley & Sons, N. Y. (1989), and supplements through September (1998), *Molecular Cloning, A Laboratory Manual*, Sambrook *et al,* 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989), *Genome Analysis: A Laboratory Manual 1: Analyzing DNA*, Birren *et al.,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1997); *Genome Analysis: A Laboratory Manual 2: Detecting Genes*, Birren *et al.,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1998); *Genome Analysis: A Laboratory Manual 3: Cloning Systems*, Birren *et al.,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); *Genome Analysis: A Laboratory Manual 4: Mapping Genomes*, Birren *et al.,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1999); *Plant Molecular Biology: A Laboratory Manual*, Clark, Springer-Verlag, Berlin, (1997), *Methods in Plant Molecular Biology*, Maliga *et al.,* Cold Spring Harbor Press, Cold Spring Harbor, New York (1995).

These texts can, of course, also be referred to in making or using an aspect of the invention. It is understood that any of the agents of the invention can be substantially purified and/or be biologically active and/or recombinant.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

EXAMPLE 1

CLONING OF tyrA FROM *Erwinia herbicola*

Vectors pJX1, pJX181, and pJX184 are obtained (Zhao and Jensen of Molecular Evolution 36(2):107-20 (1993)). The tyrA gene is amplified by PCR using primers tyrA5' (ACT GCC ATG GTG GCT GAA CTG ACC G (SEQ ID NO: 5)) and tyrA3' (ACT GGA ATT CTT ATT ATG GGC GGC TGT CAT TG (SEQ ID NO: 6)) and plasmid DNA from vectors pJX1, pJX181, and pJX184 as template DNA. PCR reactions using the Expand ™ high fidelity PCR kit from Boehringer Mannheim are performed in a total volume of 50 µl according to the manufacturer's protocol. The tyrA gene is amplified by 30 PCR cycles under the following conditions: 10 min incubation at 95°C, followed by 30 cycles of 1 min at 95°C, 1 min annealing at 56°C and 1.5 min extension at 72°C. These reactions were followed by a 5 min incubation at 72°C. The PCR product form pJX184 is chosen for gene cloning and digested with NcoI and EcoRI. The gel purified restriction fragment is ligated into NcoI/EcoRI-digested and gel purified pSE280

(Invitrogen Co., Carlsbad, CA) resulting in the formation of pMON26588 (Figure 2). The tyrA insert in pMON26588 is verified by DNA sequencing.

EXAMPLE 2

CLONING OF tyrA FROM *Escherichia coli*

The *tyrA* gene from *E. coli* is amplified by PCR using primers tyrAecoli5' (ACT GCC ATG GTT GCT GAA TTG ACC G (SEQ ID NO: 7)) and tyrAecoli3' (ACT GGA ATT CTT ATT ACT GGC GAT TG (SEQ ID NO:8)) and *E. coli* DH5α total genomic DNA as template DNA. *E. coli* total genomic DNA is isolated using the Qiaamp Tissue Kit from Qiagen (Qiagen Inc. Valencia, CA). PCR reactions using the Expand ™ high fidelity PCR kit from Boehringer Mannheim are performed in a total volume of 50 µl according to the manufacturer's protocol. The *tyrA* gene is amplified by 30 PCR cycles under the following conditions: 10 min incubation at 95°C, followed by 30 cycles of 1 min at 95°C, 1 min annealing at 56°C and 1.5 min extension at 72°C. These reactions are followed by a 5 min incubation at 72°C. The PCR product is digested with NcoI and EcoRI. The gel purified restriction fragment is ligated into NcoI/EcoRI-digested and gel purified pSE280 (Invitrogen Co., Carlsbad, CA) resulting in the formation of pMON26589 (Figure 3). The tyrA insert in pMON26589 is verified by DNA sequencing.

EXAMPLE 3

EXPRESSION OF BIFUNCTIONAL PREPHENATE DEHYDROGENASE IN *E. COLI*

Vectors pMON26588 and pMON26589 are transformed into *E. coli* DH5α and cells are grown in a 15 ml LB culture to an optical density at 600 nm of about 0.6, and induced by adding IPTG to a final concentration of 0.66 μM. After incubation for 2 to 3 hours, cells are harvested. The cell pellet iss resuspended in 0.5 ml 25 mM Tris/HCl, pH8.2 and cells are disrupted by sonication. Membranes and cell debris are sedimented by centrifugation at 100,000 x g for three hours. The supernatant is used in enzyme assays as a crude cell extract. Prephenate dehydrogenase activity is measured in a final volume of 1.5 ml containing 1 mM EDTA, 1 mM DTE, 1 mM NAD, and 1 mM prephenate (Ba-salt) in 25 mM Tris/HCl pH 8.2. The specific activity of prephenate dehydrogenase is determined by monitoring the conversion of NAD^+ to NADH as described in Methods in Enzymology Vol. 17 (Part A) pages 564-574 (1970). Results are shown in table 1, below.

TABLE 1

Vector Designation	Gene	Specific activity
		μmole/mg x min
Wild type control		2×10^{-5}
pMON26588	*tyrA Erwinia herbicola*	5.75
pMON26589	*tyrA Escherichia coli*	3.44

EXAMPLE 4

PLACING THE tyrA GENE UNDER THE CONTROL OF THE T7 PROMOTER

The *E. coli* and *E. herbicola* tyrA genes are cleaved as NcoI / EcoRI fragments from pMON26589 and pMON26588, gel purified, and cloned into NcoI / EcoRI-digested and gel purified pMON26541 (Figure 28), resulting in the formation of pMON26591 and pMON26590, respectively (Figures 4 and 5). These vectors place the tyrA gene under the control of the T7 promoter.

EXAMPLE 5

PREPARATION OF A PLANT EXPRESSION VECTOR WITH tyrA

The *E. herbicola tyrA* gene is chosen for plant expression. The gene is cleaved from pMON26590 by NcoI / EcoRI restriction digest, gel purified, and ligated into NcoI / EcoRI-digested and gel purified pMON26541 resulting in the formation of the shuttle vector pMON36510 (Figure 6). These ligations fuse the bacterial tyrA gene to CTP1, which is the chloroplast target peptide of the small subunit of the ribulose bisphosphate carboxylase from *Arabidopsis*, and place it under e35S promoter control.

To place the *tyrA* gene under Napin promoter control, pMON36510 is digested with EcoRI, ends are filled in using the Klenow fragment (Maniatis), and the gel purified vector is digested with Bgl II. The smaller fragment encoding the tyrA gene fused to CTP1 is gel purified and ligated for ligation into digested and gel purified pCGN3223 (Figure 45). To perform this ligation, pCGN3224 is digested with PstI, ends are filled in with Klenow fragment (Maniatis) and subsequently the vector is digested with Bgl II and gel purified. Ligation of the purified vector and the purified CTP1::tyrA fusion results in the formation of pMON36512 (Figure 7).

To transfer the *E. herbicola tyrA* gene into an *Arabidopsis* binary vector pMON36510 is digested with HindIII and Sac I and the gel purified fragment carrying the e35S promoter is fused to CTP1 and tyrA is ligated into HindIII / SacI-digested and gel purified pMON26543 (Figure 29), which results in the formation of pMON36511 (Figure 8). This vector contains tyrA under e35S promoter control. The pNapin binary expression vector is obtained by ligating the gel purified NotI fragment harboring the

pNapin::CTP1::tyrA::napin 3' expression cassette into NotI digested pMON36176 (Figure 30), which results in the formation of pMON36520 (Figure 9).

EXAMPLE 6

Transformation of *Arabidopsis* with pMON36520 and pMON36511

Agrobacterium that have been transformed with pMON36520 and pMON36511 are prepared as follows. 100µl of an overnight culture are spread on an agar LB plate with antibiotics. The plate is placed upside down in a 30°C chamber overnight. The plates are removed after colonies have grown (24-48 hours).

A small scale culture is started by placing 10ml of liquid LB media in a 50ml tube. 10µl Kanamycin (50 µg/µL), 10µl Spectinomycin (75-100 µg/µL), and 10µl Chloramphenicol (25 µg/µL) are added. Agrobacterium is added from a plate, and the tube is shaken and placed in a 30°C shaker overnight.

Following overnight growth of the 10ml culture, the culture is removed to a 500ml flask. 200ml of liquid LB is placed in a flask, 200µl Kanamycin (50 µg/µL), 200µl Spectinomycin (75-100 µg/µL), and 200µl of Chloramphenicol (25 µg/µL) are added, and the entire 10ml overnight culture is then added. The 500ml flask is placed in a 30°C shaker and grown overnight.

The entire 200ml culture in placed in a centrifuge tube and centrifuged for 25 minutes at 3,750 rpm and 19°C. After centrifugation, the liquid is poured off and the pellet is resuspended in 25ml of 5% Sucrose (0.05% Silwet) solution.

900µl of the sucrose solution and 100µl of the 25ml bacterial culture are placed in a cuvette, and the cuvette is shaken with a covering of parafilm. A blank OD reading is

taken with a 1ml of sucrose solution, and then readings of all the bacterial solutions are taken. The OD (at a wavelength of 600) of each culture is recorded. The following calculations are then performed:

$C_1V_1 = C_2V_2$; $C_1V_1 = (0.8)(200ml)$; $C_1V_1 = 160$; $V_1 = 160 / C_1$; and $V_1 = X$ ml/10 to determine $OD_{600} = 0.8$ of an *Agrobacterium* culture.

Plants are soaked for at least 30 minutes in water prior to dipping. The bacterial solution is poured into a shallow plastic container, and above ground parts of the plant (bolts, rosettes) are dipped into the solution for 3-5 seconds with gentle agitation. Dipped plants are placed on their side in a diaper lined black tray, and covered by a dome overnight (16-24 hours) to maintain a high humidity. The cover is removed and normal plant growth conditions are resumed for 4 weeks.

Following the transformation and high humidity treatment, plants are maintained at 22°C, 60% RH, and a 16 hour photoperiod for 4 weeks. 5-7 days after transformation , plants are coned. Fertilization with a weak 20-20-20 fertilizer is done weekly. After 4 weeks of growth, plants are placed in the greenhouse and all watering is stopped to encourage plant dry down for seed harvest. Plants are ready for seed harvest after 1-1.5 weeks of dry down.

Seeds are harvested by cutting the base of the plant below the cones, holding the plant over a seed sieve and a white piece of paper, running bolts through the cone hole, and collecting clean seeds through sieving.

Seeds are sterilized by connecting a vacuum dessicator hose to a vacuum in a fume hood/flow bench. 100ml of bleach is placed in a 250ml beaker, and 3ml of concentrated HCl is added to the bleach. The beaker is placed in the dessicator, and

seeds in seed tubes in a tube holder are placed in the dessicator. A cover is placed on the dessicator, and the vacuum is operated. The dessicator is left overnight but no longer than 16 hours.

Once sterilized, seeds are plated on selection media (prepared by adding 10g (2g/L) Phyta-Gel, 10.75 g (2.15 g/L) MS Basal Salts (M-5524 from Sigma), 50 g (10g/L) Sucrose, and 6ml (1.2 ml/L) Kanamycin solution (950mg/ml), 5ml (1ml/L) Cefotaxime Solution (250 mg/ml), and 5 ml (1 ml/L) Carbenecillin Solution (250 mg/ml) to a total volume of 5 liters at a pH or 5.7). Seed tubes are tapped lightly over a plate in order to distribute the seeds sparsely. The plates are wrapped in parafilm and placed in a 4°C refrigerator for 1-2 days of cold treatment. After this cold treatment the plates are placed in a 28°C chamber for germination.

Selected plantlets are green and have secondary leaves developing. The selected plantlets are moved to soil after secondary leaves have developed.

The plantlets are potted in soil and covered with a dome for 5 days to maintain high humidity. The plantlets are moved to a greenhouse after the bottom siliques begin to turn yellow.

Seeds from the selected plantlets are grown in 2.5 inch pots with soil (½ Metro-200; ½ PGX Mix). The soil is mounded and the pot is covered with mesh screen. The screen is fastened to the pot with a rubber band. Seeds are sown and covered with a germination dome.

The seedlings are grown in a 12hr. photoperiod in 70% relative humidity at 22°C. Water is supplied every other day as needed and Peter's 20-20-20 fertilizer is applied from below, bi-weekly.

EXAMPLE 7

Transformed seed plants from Example 6 representing 20 independent transformation events are grown and seeds harvested to produce T_2 seeds. The T_2 seeds are grown and tested for tocopherol levels. In Figure 1, tocopherol levels are expressed as nanograms of total tocopherol per milligram of seed. Tocopherol levels are determined by adding 10 to 15 mg of *Arabidopsis* seed into a 2 mL microtube. A mass of 1 g of 0.5mm microbeads (Biospecifics Technologies Corp., Lynbrook, NY) and 500 µl 1% pyrogallol (Sigma Chem, St. Louis, MO) in ethanol containing 5 µg/mL tocol, are added to the tube. The sample is shaken twice for 45 seconds in a FastPrep (Bio101/Savant) at a speed of 6.5. The extract is filtered (Gelman PTFE acrodisc 0.2 µm, 13 mm syringe filters, Pall Gelman Laboratory Inc, Ann Arbor, MI) into an autosampler tube. HPLC is performed on a Zorbax silica HPLC column, 4.6 mm x 250 mm (5 µm) with a fluorescent detection using a Hewlett Packard HPLC (Agilent Technologies, Palo Alto CA). Sample excitation is performed at 290 nm, and emission is monitored at 336 nm. Tocopherols are separated with a hexane methyl-t-butyl ether gradient using an injection volume of 20 µl, a flow rate of 1.5 ml/min, and a run time of 12 min (40°C). Tocopherol concentration and composition is calculated based on standard curves for α, β, δ, and γ-tocopherol and α, β, δ, and γ-tocotrienols using Chemstation software (Agilent Technologies, Palo Alto CA).

EXAMPLE 8

TRANSFORMED PLANTS WITH TyrA AND OTHER TOCOPHEROL BIOSYNSTHESIS GENES

Canola, *Brassica napus, Arabidopsis* and soybean plants are transformed with a variety of DNA constructs using a particle bombardment approach essentially as set forth in Christou, In *Particle Bombardment for the Genetic Engineering of Plants,* Biotechnology Intelligence Unit, Academic Press, San Diego, California (1996) or using *Agrobacterium* mediated transformation. Two sets of DNA constructs are produced. The first set of constructs are "single gene constructs". Each of the following genes is inserted into a separate plant DNA construct under the control of a napin promoter (Krindl *et al., Seed Sci. Res.* 1:209:219 (1991)) and the products of the genes can be targeted to the plastid by an encoded plastid target peptide such as CTP1 (Keegstra, *Cell 56(2)*:247-53 (1989); Nawrath, *et al., Proc. Natl. Acad. Sci. U.S.A. 91*:12760-12764 (1994)) or CTP2): an *E. herbicola tyrA* gene (Xia *et al., J. Gen. Microbiol. 138*:1309-1316 (1992)), an slr1736 gene (in Cyanobase on the world wide web at: kazusa.or.jp/cyanobase), a plant ATPT2 gene (Smith *et al., Plant J. 11: 8*3-92 (1997)), a dxs gene (Lois *et al., Proc. Natl. Acad. Sci. U.S.A. 95 (5):*2105-2110 (1998)), a dxr gene (Takahashi et al. *Proc. Natl. Acad. Sci. U.S.A. 95 (17)*, 9879-9884 (1998)), an *Arabidopsis thaliana* HPPD gene (Norris *et al., Plant Physiol. 117*:1317-1323 (1998)), a GGH gene (Keller et al., *Eur. J. Biochem. 251*:413-417 (1998)), an *Arabidopsis thaliana* GGPPS gene (Bartley and Scolnik, *Plant Physiol. 104*:1469-1470 (1994)), a AANT1 gene (Saint Guily, *et al., Plant Physiol., 100(2)*:1069-1071 (1992)), an MT1 gene (The sequence of the Synechocystis MT1 (NCBI General Identifier Number1653572) was used in a blast search against ESTs of Anabaena sp. strain PCC 7120 (Kaneko 2001). A sequence with substantial homology to the Synechocystis MT1 was found in a blast

search against ESTs of Anabaena sp. strain PCC 7120 (Kaneko et al., DNA Research 8(5): 205-213 (2001)), a TMT2 gene (as disclosed in U.S. Patent No. 7,262,339), a GMT gene (as disclosed in U.S. Patent No. 7,244,877, WO 00/32757, WO 00/10380), and a slr1737 gene (in Cyanobase),

and an antisense construct for homogentisic acid dioxygenase (Sato *et al.*, *J. DNA Res. 7 (1)*:31-63 (2000))). Each construct is transformed into at least one canola, *Brassica napus, Arabidposis* and soybean plant. Plants expressing each of these genes are selected to participate in additional crosses. The tocopherol composition and level in each plant is also analyzed using the method set forth in example 7. Crosses are carried out for each species to generate transgenic plants having one or more of the following combination of introduced genes: *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase. In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA, HPPD* and either *slr1736* or *ATPT2*.

The tocopherol composition and level in each plant generated by the crosses (including all intermediate crosses) is also analyzed using the method set forth in example 7. Progeny of the transformants from these constructs will be crossed with each other to stack the additional genes to reach the desired level of tocopherol.

A second set of DNA constructs is generated and referred to as the "multiple gene constructs." The multiple gene constructs contain multiple genes each under the control of a napin promoter (Krindl *et al., Seed Sci. Res.* 1:209:219 (1991)) and the products of each of the genes are targeted to the plastid by an encoded plastid target peptide. The

multiple gene construct can have two or more of the following genes: *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase. In a preferred combination, the nucleic acid construct or constructs encode, in addition to *tyrA*, *HPPD* and either *slr1736* or *ATPT2*.

Each construct is then transformed into at least one canola, *Brassica napus*, *Arabidopsis* and soybean plant. The tocopherol composition and level in each plant is also analyzed using the method set forth in example 7. Progeny of the transformants from these constructs are crossed with each other to stack the additional genes to reach the desired level of tocopherol.

EXAMPLE 9

TRANSFORMED ARABIDOPSIS PLANTS WITH TyrA, ATPT2, AND OTHER TOCOPHEROL BIOSYNTHESIS GENES

Wild type *Arabidopsis* plants and *Arabidopsis* plant lines are transformed with the plasmid vector pMON69907 (Figure 12), are grown, and seed is collected as described in the above examples, and the seed is analyzed for tocopherol and tocotrienol content as described above. Plasmid pMON69907 encodes a bifunctional prephenate dehydrogenase (*tyrA*) and a phytyl prenyltransferase (ATPT2). Figure 14 depicts the total tocopherol and tocotrienol content of *Arabidopsis* seeds from wild type plants and several plant lines transformed with the plasmid vector pMON69907. Figure 15 depicts the total tocopherol content of *Arabidopsis* seeds from a wild type plant and several plant lines transformed with the plasmid vector pMON69907. Figure 31 shows LC/MS standards for tocopherol and tocotrienol. Figure 32 shows LC/MS results for selected

lines, showing presence of tocotrienols. Figure 33 shows an HPLC/FLD chromatogram of control seed extract showing no presence of tocotrienols. Figure 34 shows an HPLC/FLD chromatogram of control seed extract showing the presence of tocotrienols in selected lines.

EXAMPLE 10

TRANSFORMED ARABIDOPSIS PLANTS WITH TyrA AND OTHER TOCOPHEROL BIOSYNTHESIS GENES

Expression constructs pCGN10822, pMON36528, pMON69907 and pMON69909, shown in figures 10-13 respectively are prepared.

Arabidopsis plants are transformed with the indicated vectors using the transformation techniques described in Example 8. Transformants are isolated and grown into individual lines by self pollination and seed from each line collected. The total tocopherol and tocotrienol composition of the seeds from each line are analyzed using the method set forth in Example 7. Figure 16 shows total tocopherol and tocotrienol levels for plant lines harboring the described contructs or a control. An analysis of T2 seeds from plant lines derived by transformation with the vector pMON69909 relative to wild type is shown in figure 17. Plant lines transformed with pMON69909 demonstrate a substantial increase in total tocopherols and total tocotrienols, with the largest increases in delta tocopherol, alpha tocotrieneol, delta tocotrienol, and gamma tocotrienol. Some seed from plants harboring the vector pMON69909 show a dark coloration as the result of homogentisic acid accumulation, which is confirmed by LC/MS analysis (see Figures 31 and 32).

Heterologous expression of *tyrA* in seeds of transgenic *Arabidopsis* plants

produces a 1.6-fold increase in seed tocopherol levels as compared to control lines. Another key enzyme essential for tocopherol biosynthesis is HPT, which is involved in the condensation of phytyl pyrophosphate (PPP) and homogentisate (HGA) to produced 2-methyl-6-phytylplastoquinol (2M6PPQ), a precursor for synthesis of four different isoforms of tocopherols. Overexpression of $HPT_{Arabidopsis}$ (ATPT2) and the $HPT_{Synechocystis}$ (slr1736) independently in seeds of transgenic *A. thaliana* results in a 1.6-fold increase in seed tocopherols. A putative adenylate transporter from *A. thaliana* (AANT1) expressed as a single gene is shown to increase seed tocopherol levels to 1.4-fold in *A. thaliana*. To test whether a combination of these genes would result in synergistic effect on tocopherol biosynthesis, various combinations are tested in *A. thaliana*.

T2 *Arabidopsis* seeds harboring *ATPT2* and *tyrA* double gene constructs (pMON69907), *ATPT2, tyrA*, and *HPPD* triple gene constructs (pMON69909), ATPT2, *tyrA*, and *GGPPS* triple gene constructs (pMON69915 (Figure 35)), and *ATPT2*, *tyrA*, and *AANT1* triple gene constructs (pMON69919 (Figure 36)), are analyzed for seed tocopherol content and composition. Total seed tocopherol and tocotrienol content increases to approximately 2.4-fold in lines transformed with pMON69907 (double gene vector) and up to 5-fold in the lines carrying the triple gene vector (pMON69909) (See Figure 16 and 17). *HPPD* expressed as a single gene in *A. thaliana* result in a barely detectable increase of tocopherol levels. The combination of *HPPD* with *ATPT2* does not result in a further increase of tocopherol levels as compared to lines harboring *ATPT2* alone (data not shown). In contrast, when *HPPD* is combined with *tyrA* and *ATPT2*, tocopherol and tocotrienol levels double compared to the *tyrA*, ATPT2 combination. Seeds harboring the triple gene construct pMON69909 appear much darker in color than

control seeds.

Furthermore, it is known that wild-type dicotyledonous plants do not accumulate tocotrienols. However, the transgenic *A. thaliana* seeds harboring all four constructs accumulates substantial levels of tocotrienols (confirmed by HPLC, and for selected samples by LC-MS, (See Figures 31, 32, 33, and 34). The tocopherol and tocotrienol content of seeds harboring the triple gene expression construct, pMON69909, consist of 60% tocotrienols and 40% tocopherols. When the availability of endogenous HGA is elevated by overexpression of the HGA biosynthetic enzymes (*tyrA* & *HPPD*) along with *HPT*, the *HPT* would utilize geranylgeranyl pyrophosphate (*GGPP*) and *HGA* to produce tocotrienols instead of tocopherols under conditions limited by the availability of endogenous level of geranylgeranyl reductase (*GGH*). The *GGH* functions on hydrogenating the *GGPP* to PPP, a substrate for *HPT* in tocopherol synthesis. Hence, increased tocotrienols accumulation seen in the constructs tested can be overcome by overexpression of GGH in combination with *tyrA*, *HPPD*, and *HPT*.

EXAMPLE 11

TRANSFORMED PLANTS WITH TyrA AND OTHER TOCOPHEROL BIOSYNTHESIS GENES

Plants are transformed with the DNA constructs shown in tables 2 and 3 below, employing the techniques described in Example 8. The constructs contain one or more genes under the control of a napin promoter (Krindl *et al., Seed Sci. Res.* 1:209:219 (1991)), the 7Sα' promoter (Chen et al., *PNAS 83(22)*:8560-8564 (1998)) or the Arc5 promoter (Goossens et al., *Plant Physiol. 120*:1095-1104 (1999)). The products of the genes can be targeted to the plastid by an encoded plastid target peptide such as CTP1

(Keegstra, *Cell 56(2)*:247-53 (1989); Nawrath, *et al., Proc. Natl. Acad. Sci. U.S.A. 91*:12760-12764 (1994)) or CTP2. One or more of the following genes are used: an *E. herbicola tyrA* gene (Xia *et al., J. Gen. Microbiol. 138*:1309-1316 (1992)), an slr1736 gene (in Cyanobase on the world wide web at: kazusa.org.jp/cyanobase), an ATPT2 gene (Smith *et al., Plant J. 11: 83*-92 (1997)), an *E. coli* dxs gene (Lois *et al., Proc. Natl. Acad. Sci. U.S.A. 95 (5)*:2105-2110 (1998)), a dxr gene (Takahashi et al. *Proc. Natl. Acad. Sci. U.S.A. 95 (17)*, 9879-9884 (1998)), an HPPD gene (Norris *et al., Plant Physiol. 117*:1317-1323 (1998)), a GGH gene (Keller et al., *Eur. J. Biochem. 251*:413-417 (1998)), an *Arabidopsis thaliana* GGPPS gene (Bartley and Scolnik, *Plant Physiol. 104*:1469-1470 (1994)), an AANT1 gene (Saint Guily, *et al., Plant Physiol., 100(2)*:1069-1071 (1992)), an MT1 gene (as above for Example 8), a TMT2 gene (as above for example 8), a GMT gene (as above for example 8, and WO 00/32757, WO 00/10380), an slr1737 gene (in Cyanobase on the world wide web at kazusa.org.jp/cyanobase), and an antisense construct for homogentisic acid dioxygenase (denoted HGD_{AS})(Sato *et al., J. DNA Res. 7 (1)*:31-63 (2000)). Each construct is transformed into at least one canola, *Brassica napus, Arabidposis*, and soybean plant. The tocopherol composition and level in each plant is also analyzed using the method set forth in example 7. Examples of transformed plants with *tyrA* and other tocopherol biosynthesis genes include *Arabidopsis* plants transformed with the constructs set forth in Table 2 and soy plants transformed with the constructs in Table 3.

Plants with desired characteristics may be subject to further crosses to generate transgenic plants having one or more of the following combination of introduced genes: *tyrA, slr1736, ATPT2, dxs, dxr, GGH, GGPPS, HPPD, MT1, TMT2, GMT, AANT1, slr*

1737, and an antisense construct for homogentisic acid dioxygenase. Alternatively the plants may be crossed to stack multiple copies of one or more of the aforementioned genes in a transgenic plant.

TABLE 2

The following gene combinations are prepared and tested in *Arabidopsis thaliana*	
1	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::CTP2::HPPD$_{A.\ thaliana}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'
2	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::AANT1$_{A.\ thaliana}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'
3	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::GGPPS$_{A.\ thaliana}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'
4	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::CTP1::DXS$_{E.\ coli}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'/pNapin::GGH::Napin 3'
5	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::CTP1::DXR$_{E.\ coli}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'/pNapin::GGH::Napin 3'
6	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::HGD$_{AS}$::Napin3'
7	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::HGD$_{AS}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$
8	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::HGD$_{AS}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'/pNapin::GGH::Napin 3'
9	pNapin::CTP1::TyrA$_{E.\ herbicola}$::Napin 3'/pNapin::HGD$_{AS}$::Napin3'/pNapin::ATPT2$_{A.\ thaliana}$::Napin 3'/pNapin::CTP1::DXS$_{E.\ coli}$::Napin 3'/pNapin::GGH::Napin 3'

TABLE 3

The following gene combinations are prepared and tested in *Glycine max*	
1	p7S::CTP2::HPPD::E9 3'/p7Sα'::CTP1::TyrA::E9 3'
2	pArc5::ATPT2::Arc 3'/ p7Sα'::CTP1::TyrA::E9 3'/pNapin::GGH::Napin3'
3	pArc5::ATPT2::Arc 3'/p7S::CTP1::TyrA::E9 3'/ pNapin::GGH::Napin3'/pNapin::CTP1::DXS::Napin 3'

WO 02/089561 PCT/US02/13898 CA 02443865 2003-10-07

EXAMPLE 12
CONSTRUCTION OF PLANT BINARY VECTORS HARBORING COMBINATIONS OF *tyrA* WITH OTHER TOCOPHEROL SYNTHESIS GENES

Components of each gene expression cassette include a promoter, in this example the napin promoter, a terminator, a plastid target peptide (which can be the native plastid target peptide or an N-terminal fused chloroplast target peptide), and a gene of interest, as shown in Figures 18a and 18b. The expression cassettes can be oriented head to tail, head to head, or the orientation can vary.

Cloning is performed using expression cassettes flanked by Bsp120 I and Not I restriction sites. A shuttle vector (pMON36582 (Figure 19)) is constructed by annealing primers SV MCS 1A and SV MCS 1B:

```
                                Xma I    Bsp120 I Eag     Xba I EcoRI
SV MCS 1A          GATCT CCCGGG AA GGGCCC CGGCCG TCTAGA GAATTC

  Not I     Asc I  Age Iv
GCGGCCGC GGCGCGCC ACCGGT (SEQ ID NO:9)
```

```
                               Xma I    Bsp120 I Eag     Xba I EcoRI
SV MCS 1B          TCGA ACCGGT GGCGCGCC GCGGCCGC GAATTC TCTAGA

  Not I     Asc I  Age I
CGGCCG GGGCCC TT CCCGGG A (SEQ ID NO: 10)
```

and ligating them into Bgl II and Xho I digested and gel purified pSP72 (Promega, www.promega.com). The resulting vector is designated pMON36582 (Figure 19). The vector is confirmed by DNA sequencing.

All gene expression cassettes are set up to be flanked by Not I restriction sites. These cassettes are isolated by digesting the previous vectors with Not I, followed by gel purification of the expression cassettes. pMON36582 is digested with Eag I, which cuts twice in this vector, once within the Not I site, and once 19 bp upstream of the Not I site.

Both overhangs are compatible with Not I. The Not I expression cassettes are ligated into gel purified Eag I digested pMON36582, resulting in a vector with a single Not I site. The expression cassette is therefore available as a Bsp120 I / Not I cassette. An example of an expression cassette for the *Arabidopsis* homogentisate phytyltransferase available as a Bsp120 I / Not I cassette is shown as pMON36586 in Figure 20. This vector is obtained as described above.

The assembly of expression cassettes are performed in a shuttle vector, such as pMON36586. Gene expression cassettes are released from other shuttle vectors by Bsp120 I / Not I digests, and ligated into a shuttle vector such as pMON36586, which has been digested with Not I. The resulting vector harbors one additional gene expression cassette and a single Not I site. This procedure can be repeated as required. Upon completion of the gene assembly, the combined expression cassettes can be released by Bsp120 I / Not I digest (pMON10098 (Figure 37)). The resulting fragment carrying the expression cassettes is then purified and ligated into a single Not I site of a binary vector. Alternatively the assembly of gene expression cassettes can be performed directly in a binary vector (Figure 21). A binary vector is defined by the presence of the right and left border sequences, which are necessary for DNA transfer from *Agrobacterium* into plant cells. All chemical reagents and enzymes for this manipulation are molecular grades. These reagents and enzymes are utilized according to the supplier's instructions. Standard molecular cloning techniques are used.

Several examples of plant binary constructs, their components and plasmid maps are depicted. The examples depicted containing tyrA combinations with other genes of interest for tocopherol pathway engineering are listed in Figures 18a and 18b.

Components of these constructs are also provided in the Table 4. Vector maps shown as Figures 22-27 represent various constructs.

TABLE 4. List of binary vectors to be transformed into Arabidopsis thaliana to engineer tocopherol biosynthesis

pMON #	Gene Combination	Genetic Elements
36596	$HPPD_{At}$/tyrA	Napin 5' & Napin 3'; CTP1&2
36597	$HPPD_{At}$/tyrA/GGH_{Syn}	Napin 5' & Napin 3'; CTP1&2; native CTPs
77601	HPPDAt/tyrA/GGHAt/ATPT2	Napin 5' & Napin 3'; CTP1&2; native CTPs
77602	$HPPD_{At}$/tyrA/GGH_{Syn}/ ATPT2	Napin 5' & Napin 3'; CTP1&2; native CTPs
66657	$HPPD_{At}$/tyrA/GGH_{Syn}/ slr1736	Napin 5' & Napin 3'; CTP1&2; native CTPs
66659	$HPPD_{At}$/tyrA/GGH_{At}/ ATPT2/TMT2	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{syn}/ ATPT2/MT1	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/ ATPT2/TMT2	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{At}/ ATPT2/MT1/ DxS	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{At}/ ATPT2/$DxS_{E.coli}$	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{At}/ slr1736/MT1/$GGPPS_{At}$	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/ ATPT2/GGH_{At}/DxR_{At}	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{At}/ slr1736/$Cyclase_{syn}$/MT1	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{syn}/ slr1736/$Cyclase_{syn}$	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/slr1736/$Cyclase_{syn}$	Napin 5' & Napin 3'; CTP1&2; native CTPs
	$HPPD_{At}$/tyrA/GGH_{At}/ ATPT2/GMT_{At}	Napin 5' & Napin 3'; CTP1&2; native CTPs

WO 02/089561 PCT/US02/13898 CA 02443865 2003-10-07

EXAMPLE 13

CONSTRUCTION OF VECTORS ENCODING MULTIPLE ENZYMES

This example set forth the use of a prephenate dehydrogenase (tyrA) such as the *Erwinia herbicola* tyrA in combination with other key enzymes in the tocopherol biosynthetic pathway to enhance tocopherol production in transgenic plant seeds such as *Arabidopsis thaliana* seeds. The enzymes combined with tyrA include ATPT2, *p*-hydroxyphenylpyruvate dioxygenase ($HPPD_{Arabidopsis}$), and geranylgeranylpyrophosphate synthase ($GGPPS_{Arabidopsis}$) from *Arabidopsis thaliana*. In addition tyrA was also tested in combination with ATPT2 and a putative adenylate transporter ($AANT1_{Arabidopsis}$) from *Arabidopsis thaliana*.

Construction of a double gene vector harboring seed specific *tyrA* and *ATPT2* expression cassettes is performed as follows. Purified plasmid DNA of pMON36520 (Figure 38) is subjected to a partial KpnI digest and ligated with a 4.2 kbp gel purified Kpn I-fragment isolated from pMON43853 (Figure 39). The 4.2 kb insert from pMON43853 contains the PPT gene expression cassette (pNapin::ATPT2::Napin 3'). The resulting plant binary vector pMON69907 (Figure 12) is used for transformation of *Arabidopsis thaliana* to test the combinatorial effect of seed specific expression of tyrA and ATPT2.

To further increase tocopherol biosyntheses the $HPPD_{Arabidopsis}$ is expressed in addition to *tyrA*, and ATPT2 in *Arabidopsis thaliana* seed. This was achieved by adding a seed specific expression cassette for $HPPD_{Arabidopsis}$ to pMON69907 resulting in the formation of pMON69909. The binary vector pMON69909 is constructed by partially digesting pMON69907 with KpnI. The single KpnI-cut pMON69907 is gel purified and ligated with a 4.6 kb KpnI/KpnI insert isolated from pMON36525 (Figure 40). The 4.6 kb KpnI/KpnI insert from pMON36525 contains the HPPD gene expression cassette, pNapin::CTP2::$HPPD_{Arabidopsis}$::Napin 3' to direct seed specific plastid targeted expression of HPPD. The CTP2 is a chloroplast-target signal from the *Arabidopsis* 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene.

The binary vector pMON69915 (Figure 35) is constructed to test the effect of three gene combinations, tyrA, ATPT2, and $GGPPS_{Arabidopsis}$ on seed tocopherol production. Vector pMON69907 is digested partially with KpnI. Single KpnI-cut pMON69907 is gel purified and ligated with a gel purified 4.3 kb, KpnI/KpnI fragment from pMON43861 (Figure 41) to create pMON69915. The KpnI fragment from pMON43861 contains the gene expression cassette for the *Arabidopsis* geranylgeranyldiphosphate synthase from *Arabidopsis thaliana* (pNapin::$GGPPS_{Arabidopsis}$::Napin 3'). The GGPPS cDNA is identified as an EST clone, by searching an EST database with sequence information available in the literature (Okada et al., *Plant Physiol. 122*:1045-1056 (2000)). The EST clone is digested with NcoI and blunt-ended by filling the 5' overhang with the klenow-fragment. Subsequently the clone is digested with BamHI and to excise the cDNA fragment. The gel purified BamHI/blunt cDNA fragment is ligated with BglII/SalI digested and (SalI blunt-ended) vector pCGN7770 (Figure 42) to create pMON43861.

The plant binary vector pMON69919 (Figure 36) is constructed to test combined expression of tyrA, ATPT2, and $AANT1_{Arabidopsis}$ on seed tocopherol levels. To generate this vector, pMON69907 is partially digested with KpnI. Single KpnI-cut pMON69907 is gel purified, and ligated with a 4.2 kb gel-purified KpnI/KpnI fragment from pMON69911 (Figure 43). The 4.2 kb fragment contains a seed specific expression cassette for the *Arabidopsis* adenylate transporter AANT1 (pNapin::$AANT1_{Arabidopsis}$::napin 3'). pMON69911 is generated by excising the AANT1 fragment from pCGN11301 (Figure 44) with SalI and PstI (the PstI site is blunted by removing 3' overhang with Klenow) and then ligated to SalI/XhoI digested (XhoI blunt-ended) pCGN7770.

Using the published partial sequence of AANT1 (Saint-Guily et al., *Plant Physiol. 100(2)*:1069-1071 (1992)) several full-length clones are identified in EST databases. The AANT1 coding region is PCR-amplified using primers, AANT1F 5'-GGATCCGCGGCCGCACC***ATG***GTTGATCAAGTTCAGCA (SEQ ID NO: 11) and AANT1R 5'-GAGCTCCTGCAGGAAGCTT***TTA***GGCACCTCCTGATCCGT-3' (SEQ ID NO: 12). The *Not*I site (underlined) is placed upstream of the start codon (italics) in primer AANT1F while the *Sse*8387I site (underlined) was placed downstream of the stop codon (italics) in AANT1R. The PCR products are first cloned into pCR2.1 and the inserts are verified by

sequencing of both strands. Subsequently, the *Not*I/*Sse*8387I fragments are inserted into the *Not*I/*Sse*8387I sites of the napin expression cassette in pCGN9979 in sense orientation with respect to the napin promoter to generate pCGN11301. The plant expression constructs are used for transformation of *Arabidopsis thaliana* by *Agrobacterium* mediated transformation as described in above.

EXAMPLE 14

EXPRESSION OF VECTORS ENCODING MULTIPLE ENZYMES IN PLANTS

Using the transformation technique given in Example 8, *Arabidopsis thaliana* plants are transformed with the vectors from Example 13. Results for pMON69909 are given in Figures 14, 15, 16, 17 and 31-34. Further results are given in the table below and in Figures 75 and 76, which show tocopherol, tocotrienol, homogentisate, and 2-methylphytlyplastoquinol levels for transformed plants having pMON69909.

Construct	Genetic elements	Tocopherol/ Tocotrienol increase	% tocopherol	% tocotrienol
pMON69907	pNapin::HPT$_{Arabidopsis}$::napin 3'/ pNapin::CTP1::tyrA$_{Erwinia\ herbicola}$::napin 3'	2.4-fold	91	9
pMON69909	pNapin::HPT$_{Arabidopsis}$::napin 3' /pNapin::CTP1::tyrA$_{Erwinia\ herbicola}$::napin 3'/ pNapin::CTP2::HPPD$_{Arabidopsis}$:: napin 3'	5-fold	38	62
pMON69915	pNapin::HPT$_{Arabidopsis}$::napin 3'/ pNapin::CTP1::tyrA$_{E.\ herbicola}$ ::napin 3'/ pNapin::GGPPS$_{Arabidopsis}$::napin3'	2.9-fold	86	14
pMON69919	pNapin::HPT$_{Arabidopsis}$::napin 3'/pNapin::CTP1::tyrA$_{E.\ herbicola}$::napin 3' /pNapin::AANT1$_{Arabidopsis}$::napin3'	3-fold	89	11

EXAMPLE 15

EXPRESSION OF CONSTRUCTS IN SOYBEAN

This example describes the method involved in preparation of plant binary vectors to test tyrA alone and in combination with other key enzymes in the tocopherol biosynthetic pathway to enhance tocopherol production in transgenic *Glycine max* seeds.

The table below describes the plant binary vectors prepared for *G. max* transformation with their respective gene of interest expression cassettes for seed-specific expression of the transgenes.

List of constructs transformed to *G. max*.

Construct number	Genetic elements
pMON36575	p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3'
pMON69924	p7Sα'::CTP2::HPPD$_{Arabidopsis}$::E9 3'/ p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3'
pMON69943	p7Sα'::CTP2::HPPD$_{Arabidopsis}$::E9 3'/ p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3'/ pArcelin-5::CTP1::slr1736::Arcelin 3'
pMON69945	p7Sα'::CTP2::HPPD$_{Arabidopsis}$::E9 3'/ p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3'/ pArcelin-5::CTP1::slr1736::Arcelin 3'/ pNapin ::GGH$_{Arabidopsis}$::napin 3'

The pMON36575 (Figure 46) is prepared by ligating the 3 kb gel purified NotI fragment from pMON38207R (Figure 47) at the NotI site of pMON36571 (Figure 48) that contains the p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3' expression cassette. The CTP1 encodes the chloroplast-target signal sequence from the *Arabidopsis* RUBISCO small subunit. The 3 kb NotI fragment contains the selectable marker cassette, pFMV::CTP2::CP4syn::E9 3'. The CTP2 encodes the chloroplast-target signal sequence from the *Arabidopsis* 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). The CP4syn is an EPSPS synthetic gene. Vector pMON36575 is further digested with HindIII to release 3 kb fragment containing the p7Sα'::CTP1::tyrA$_{E.\ herbicola}$::E9 3' expression cassette. The fragment is blunt-ended by filling the 5' overhangs with the klenow-fragment, gel purified and ligated at the PmeI site of pMON36576 (Figure 49) which carries the expression cassette of p7Sα'::CTP2::HPPD$_{Arabidopsis}$::E9 3' to generate the pMON69924 (Figure 50).

The plant binary vector pMON69943 (Figure 51) is prepared by digesting pMON69929 (Figure 52), containing the p7Sα'::CTP2::HPPD$_{Arabidopsis}$::E9 3' expression cassette, with NotI and ligating with 7.3 kb gel purified fragment generated by digestion of pMON69936 (Figure 53) with Bsp120I and NotI. This fragment contains the expression cassettes of p7Sα'::CTP1::tyrA$_{E.}$

herbicola::E9 3' and pArcelin-5::CTP1::slr1736::Arcelin 3'. Vector pMON69943 is further digested with NotI and ligated with a 4.5 kb Bsp120I/NotI gel purified fragment from pMON36592 (Figure 54) to generate pMON69945 (Figure 55). The fragment from pMON36592 contains the expression cassette of pNapin ::$GGH_{Arabidopsis}$::napin 3'.

Gluycine max is transformed with the described vectors according to the procedure set forth in WO 00/61771 A3 on pages 99-100.

EXAMPLE 16

Multiple gene combinations expressed in canola.

All gene expression cassettes used for expression in canola are prepared as Not I cassettes containing the napin promoter, a gene of interest, and the napin terminator. Genes of interest are N-terminally fused to a chloroplast target peptide, unless a natural chloroplast target peptide is present. All gene combinations are assembled in a single multi gene vector.

In order to ease the construction of multi gene vectors, the Not I expression cassettes are isolated by Not I digestion from pMON16602 (Figure 56), pMON36525 (Figure 57), pMON36520 (Figure 38) and cloned into Eag I digested and gel purified pMON36582 (Figure 19), resulting in the formation of pMON58171 (Figure 58)(*slr1736* expression cassette), pMON58172 (Figure 59)($HPPD_{Arabidopsis}$ expression cassette), and pMON58170 (Figure 60) ($tyrA_{E.\ herbicola}$ expression cassette). All of the resulting expression cassettes are flanked by Bsp120 I and Not I.

A napin driven expression cassette for the *Arabidopsis* GGH is obtained by isolation and gel purification of a 3191 bp Not I / Hind III fragment from pMON36591 (Figure 61) and a 5612 bp Not I / Hind III-fragment from pMON36588 (Figure 62). These two purified fragments are

ligated, resulting in the formation of pMON36592 (Figure 63). Vector pMON36592 is digested with Bsp120I and Not I, the GGH expression cassette is gel purified, and ligated into Eag I digested and gel purified pMON36582 (Figure 19), resulting in the formation of pMON58182 (Figure 64).

Multi gene vectors combining these four genes are obtained by digesting vectors pMON58171, pMON58172, pMON58170, and pMON58182 with Bsp120I and Not I, followed by gel purification of the larger fragments from each construct. These fragments contain the slr1736, HPPD, tyrA, and GGH expression cassettes, respectively. The tyrA expression cassette from pMON58170 is ligated into Not I digested and alkaline phosphatase treated pMON58171, resulting in the formation of the double gene vector pMON58176 (Figure 65) containing gene expression cassettes for tyrA and slr1736, respectively. This vector is again digested with Not I, alkaline phosphatase treated, and ligated with the HPPD expression cassette from pMON58172. The resulting triple gene vector pMON58183 (Figure 66) contains the HPPD, tyrA, and slr1736 expression cassettes. Also pMON58183 is digested with Bsp120 I, alkaline phosphatase treated, and ligated with the gel purified GGH expression cassette (see purification above), resulting in the formation of pMON58185 (Figure 67).

The shuttle vector pMON36593 (Figure 68) (containing tyrA and HPPD expression cassettes) is prepared by ligating a Bsp120I / Not I digested gel purified tyrA expression cassette from pMON36589 (Figure 69) into NotI digested and alkaline phosphatase treated pMON36590 (Figure 70).

The combined gene expression cassettes are excised by Bsp120 I / Not I digest from pMON36593 (HPPD/tyrA), pMON58183 (HPPD/tyrA/slr1736), and pMON58185 (HPPD, tyrA,

slr1736, GGH). These combined gene expression cassettes are gel purified, and ligated into Not I digested, alkaline phosphatase treated pMON67162 (Figure 71), resulting in the formation of binary vectors pMON58178 (Figure 72), pMON58186 (Figure 73), and pMON58188 (Figure 74), respectively. The latter three binary vectors are used for canola transformation.

SEQUENCE LISTING

<110> MONSANTO TECHNOLOGY LLC

<120> TyrA Genes and Uses Thereof

<130> 6305-17

<140>
<141> May 3, 2002

<150> PCT/US02/13898
<151> May 3, 2002

<150> US 60/289,527
<151> May 9, 2001

<160> 12

<170> PatentIn version 3.0

<210> 1
<211> 1122
<212> DNA
<213> Erwinia herbicola

<400> 1
atggtggctg aactgaccgc gttacgcgat caaattgaca gtgtagataa agcgctgctg 60

gatctgctgg ctaagcgact ggaactggtg gccgaggtag gtgaggtgaa gagccgttac 120

ggcctgccta tctatgtgcc tgagcgtgag gcgtcgatgc tggcttcgcg tcgcaaagag 180

gccgaagcgc tcggcgtacc accggatctg attgaggatg tgctgcgtcg cgtgatgcgg 240

gaatcctata ccagcgagaa tgataaaggc tttaaaaccc tctgtcctga actgcgcccg 300

gtggtgattg tcggtggtaa gggccagatg ggccggctgt ttgaaaaaat gctcgggcta 360

tcaggctaca cggttaaaac gctggataaa gaggactggc ctcaggctga gactctgctc 420

agcgatgccg gaatggtgat cattagcgtg ccgattcacc tgaccgagca ggtgattgcc 480

caactgccac cactgccgga agattgtatt ctggtcgatc tggcgtcagt caaaaaccgg 540

cctctgcagg caatgctggc tgcccataac gggcctgtac tgggtctgca tccgatgttt 600

ggcccggaca gcggcagcct ggcaaaacag gtggtggtct ggtgtgatgg aagacaaccg 660

gaagcgtatc agtggttcct ggagcagatt caggtctggg gtgcgcgtct gcatcgtatc 720

agcgctgttg agcatgacca gaacatggca ttcattcagg cgctgcgtca ctttgctacc 780

ttcgcttatg gtctgcattt agccgaagag aacgtcaatc tggatcagct gctggcgctc 840

tcgtcgccca tttaccggct tgaactggcg atggtggggc ggttgttcgc tcaggatccg 900

caactctatg cggatatcat catgtcttca gagagtaatc tggcgctgat aaaacgctat 960

taccagcggt ttggtgaagc gattgcgctg ctggagcagg gcgacaagca ggcgtttatc 1020

gccagcttta accgggttga acagtggttt ggcgatcacg caaaacgctt cctggtcgaa 1080

agccgaagcc tgttgcgatc ggccaatgac agccgcccat aa 1122

<210> 2
<211> 373
<212> PRT
<213> Erwinia herbicola

<400> 2

Met Val Ala Glu Leu Thr Ala Leu Arg Asp Gln Ile Asp Ser Val Asp
1 5 10 15

Lys Ala Leu Leu Asp Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu
20 25 30

Val Gly Glu Val Lys Ser Arg Tyr Gly Leu Pro Ile Tyr Val Pro Glu
35 40 45

Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Lys Glu Ala Glu Ala Leu
50 55 60

Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg
65 70 75 80

Glu Ser Tyr Thr Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro
85 90 95

Glu Leu Arg Pro Val Val Ile Val Gly Gly Lys Gly Gln Met Gly Arg
100 105 110

Leu Phe Glu Lys Met Leu Gly Leu Ser Gly Tyr Thr Val Lys Thr Leu
115 120 125

Asp Lys Glu Asp Trp Pro Gln Ala Glu Thr Leu Leu Ser Asp Ala Gly
130 135 140

Met Val Ile Ile Ser Val Pro Ile His Leu Thr Glu Gln Val Ile Ala
145 150 155 160

Gln Leu Pro Pro Leu Pro Glu Asp Cys Ile Leu Val Asp Leu Ala Ser
165 170 175

Val Lys Asn Arg Pro Leu Gln Ala Met Leu Ala Ala His Asn Gly Pro
180 185 190

Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala
195 200 205

Lys Gln Val Val Val Trp Cys Asp Gly Arg Gln Pro Glu Ala Tyr Gln
210 215 220

Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile
225 230 235 240

Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg
245 250 255

His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val
260 265 270

Asn Leu Asp Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu
275 280 285

Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala
290 295 300

Asp Ile Ile Met Ser Ser Glu Ser Asn Leu Ala Leu Ile Lys Arg Tyr
305 310 315 320

Tyr Gln Arg Phe Gly Glu Ala Ile Ala Leu Leu Glu Gln Gly Asp Lys
325 330 335

Gln Ala Phe Ile Ala Ser Phe Asn Arg Val Glu Gln Trp Phe Gly Asp
340 345 350

His Ala Lys Arg Phe Leu Val Glu Ser Arg Ser Leu Leu Arg Ser Ala
355 360 365

Asn Asp Ser Arg Pro
370

<210> 3
<211> 1122
<212> DNA
<213> Escherichia coli

<400> 3
atggttgctg aattgaccgc attacgcgat caaattgatg aagtcgataa agcgctgctg 60

aatttattag cgaagcgtct ggaactggtt gctgaagtgg gcgaggtgaa aagccgcttt 120

ggactgccta tttatgttcc ggagcgagag gcatctatgt tggcctcgcg tcgtgcagag 180

gcggaagctc tgggtgtacc gccagatctg attgaggatg ttttgcgtcg ggtgatgcgt 240

gaatcttact ccagtgaaaa cgacaaagga tttaaaacgc tttgtcctgc gttacgcccg 300

gtagttatcg ttggcggcgg cggtcagatg ggacgtctgt tcgagaagat gctgacactc 360

tcgggttatc aggtgcggat tctggagcaa catgactggg atcgagcggc tgatattgtt 420

gccgatgccg gaatggtgat tgttagtgtg ccaatccacg ttactgagca agttattggc 480

aaattaccgc ctttaccgaa agattgtatt ctggttgatc tggcatcagt gaaaaatgga 540

ccattacagg ccatgctggc ggcgcacgat ggcccggtac tggggttaca cccaatgttc 600

ggtccggaca gcggtagcct ggcaaagcaa gttgtggtct ggtgtgatgg acgtaaaccg 660

gaagcatacc aatggtttct ggagcaaatt caggtctggg gcgctcggtt gcatcgtatt 720

agcgccgtcg agcacgatca gaatatggcg tttattcagg cactgcgcca ctttgctact 780

tttgcttacg ggctgcacct ggcagaagaa aatgttcagc ttgagcaact tctggcgctc 840

tcttcgccga tttaccgcct tgagctggcg atggtcgggc gactgttcgc tcaggatccg 900

cagctttatg ccgacattat tatgtcgtca gagcgtaatc tggcgttaat caaacgttac 960

tataagcgtt tcggcgaggc gattgagttg ctggagcagg gcgataagca ggcgtttatt 1020

gacagtttcc gcaaggtgga gcactggttc ggcgattacg cacagcgttt tcagagtgaa 1080

agccgcgtgt tattgcgtca ggcgaatgac aatcgccagt aa 1122

<210> 4
<211> 373
<212> PRT
<213> Escherichia coli

<400> 4

Met Val Ala Glu Leu Thr Ala Leu Arg Asp Gln Ile Asp Glu Val Asp
1 5 10 15

Lys Ala Leu Leu Asn Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu
20 25 30

Val Gly Glu Val Lys Ser Arg Phe Gly Leu Pro Ile Tyr Val Pro Glu
35 40 45

Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Ala Glu Ala Glu Ala Leu
50 55 60

Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg
65 70 75 80

Glu Ser Tyr Ser Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro
85 90 95

Ala Leu Arg Pro Val Val Ile Val Gly Gly Gly Gly Gln Met Gly Arg
100 105 110

Leu Phe Glu Lys Met Leu Thr Leu Ser Gly Tyr Gln Val Arg Ile Leu
115 120 125

Glu Gln His Asp Trp Asp Arg Ala Ala Asp Ile Val Ala Asp Ala Gly
130 135 140

Met Val Ile Val Ser Val Pro Ile His Val Thr Glu Gln Val Ile Gly
145 150 155 160

Lys Leu Pro Pro Leu Pro Lys Asp Cys Ile Leu Val Asp Leu Ala Ser
165 170 175

Val Lys Asn Gly Pro Leu Gln Ala Met Leu Ala Ala His Asp Gly Pro
180 185 190

Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala
195 200 205

Lys Gln Val Val Val Trp Cys Asp Gly Arg Lys Pro Glu Ala Tyr Gln
210 215 220

Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile
225 230 235 240

Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg
245 250 255

His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val
260 265 270

Gln Leu Glu Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu
275 280 285

Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala
290 295 300

Asp Ile Ile Met Ser Ser Glu Arg Asn Leu Ala Leu Ile Lys Arg Tyr
305 310 315 320

Tyr Lys Arg Phe Gly Glu Ala Ile Glu Leu Leu Glu Gln Gly Asp Lys
325 330 335

Gln Ala Phe Ile Asp Ser Phe Arg Lys Val Glu His Trp Phe Gly Asp
340 345 350

Tyr Ala Gln Arg Phe Gln Ser Glu Ser Arg Val Leu Leu Arg Gln Ala
355 360 365

Asn Asp Asn Arg Gln
370

<210> 5
<211> 25
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 5
actgccatgg tggctgaact gaccg 25

<210> 6
<211> 32
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 6
actggaattc ttattatggg cggctgtcat tg 32

<210> 7
<211> 25
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 7
actgccatgg ttgctgaatt gaccg 25

<210> 8
<211> 26
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 8
actggaattc ttattactgg cgattg 26

<210> 9
<211> 59
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 9
gatctcccgg gaagggcccc ggccgtctag agaattcgcg gccgcggcgc gccaccggt 59

<210> 10
<211> 59
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 10
tcgaaccggt ggcgcgccgc ggccgcgaat tctctagacg gccggggccc ttcccggga 59

<210> 11
<211> 37
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 11
ggatccgcgg ccgcaccatg gttgatcaag ttcagca 37

<210> 12
<211> 39
<212> DNA
<213> artificial sequence

<220>
<223> synthetic primer

<400> 12
gagctcctgc aggaagcttt taggcacctc ctgatccgt 39

CLAIMS:

1. A substantially purified nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule that encodes an enzyme with chorismate mutase and prephenate dehydrogenase activities or a fragment thereof of at least 20 contiguous amino acids of said enzyme, said heterologous nucleic acid molecule is selected from the group consisting of the nucleic acid sequence SEQ ID NO:1, the nucleic acid sequence SEQ ID NO:3, a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:2, and a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:4

2. A substantially purified nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule that encodes an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4 and fragments thereof of at least 20 contiguous amino acids.

3. The nucleic acid molecule of claim 1, further comprising a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

4. The nucleic acid molecule of claim 1, wherein said heterologous nucleic acid molecule encodes *Erwinia herbicola tyrA* or *Escherichia coli tryA*.

5. The nucleic acid molecule of claim 1, wherein said heterologous nucleic acid molecule comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1 and 3.

6. The nucleic acid molecule of claim 1, wherein said heterologous nucleic acid molecule further comprises an expression cassette which expresses phytyl prenyltransferase.

7. The nucleic acid molecule of claim 6, wherein said heterologous nucleic acid molecule further comprises an expression cassette which expresses hydroxyphenyl-pyruvate dehydrogenase.

8. The nucleic acid molecule of claim 1, wherein said fragment encodes a polypeptide having prephenate dehydrogenase activity.

9. The nucleic acid molecule of claim 1, wherein said heterologous nucleic acid molecule further comprises two or more expression cassettes, each of which express a member selected from the group consisting of *slr 1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, and an antisense construct for homogentistic acid dioxygenase.

10. The nucleic acid molecule of claim 1, wherein said heterologous nucleic acid molecule further comprises a nucleic acid sequence encoding *HPPD* and either *slr 1736* or *ATPT2*.

11. A nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein having chorismate mutase and prephenate dehydrogenase activities, said heterologous nucleic acid is selected from the group consisting of the nucleic acid sequence SEQ ID NO:1, the nucleic acid sequence SEQ ID NO:3, a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:2, and a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:4.

12. A nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) a heterologous nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of SEQ ID NOs:2 and 4.

13. A transformed plant cell having a nucleic acid molecule which comprises as operably linked components: (A) a promoter region which functions in a plant cell to cause the production of an mRNA molecule; (B) an exogenous nucleic acid molecule encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs:2 and 4 or fragments thereof of at least 20 contiguous amino acids, and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of said mRNA molecule.

14. The transformed plant cell of claim 13, wherein said exogenous nucleic acid molecule further comprises an expression cassette which expresses phytyl prenyltransferase.

15. The transformed plant cell of claim 14, wherein said heterologous nucleic acid molecule further comprises an expression cassette which expresses hydroxyphenyl-pyruvate dehydrogenase.

16. The transformed plant cell of claim 13, wherein said heterologous nucleic acid molecule further comprises two or more expression cassettes, each of which expresses a member of the group consisting of: *slr 1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, and an antisense construct for homogentistic acid dioxygenase.

17. The transformed plant cell of claim 13, wherein said heterologous nucleic acid molecule further comprises a nucleic acid sequence encoding *HPPD* and either *slr 1736* or *ATPT2*.

18. The transformed plant cell according to claim 13, wherein said plant cell is selected from the group consisting of a canola cell, a maize cell, an *Arabidiosus* cell, a *Brassica campestris* cell, a *Brassica napus* cell, a soybean cell, a crambe cell, a mustard cell, a castor bean cell, a peanut cell, a sesame cell, a cottonseed cell, a linseed cell, a safflower cell, an oil palm cell, a flax cell and a sunflower cell.

19. The transformed plant cell according to claim 13, wherein said plant cell is a soybean cell.

20. The transformed plant cell according to claim 13, wherein said plant cell is a canola cell.

21. The transformed plant cell according to claim 13, wherein said plant cell is a *Brassica nupus* cell.

22. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocopherol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

23. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocopherol levels of at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

24. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocopherol levels of at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

25. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocopherol levels of at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

26. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocotrienol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

27. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocotrienol levels of at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

28. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocotrienol levels of at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

29. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased tocotrienol levels of at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

30. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocopherol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

31. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocopherol levels of at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

32. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocopherol levels at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

33. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocopherol levels at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

34. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocotrienol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

35. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocotrienol levels at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

36. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocotrienol levels at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

37. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased α-tocotrienol levels at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

38. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocopherol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

39. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocopherol levels of at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

40. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocopherol levels of at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

41. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocopherol levels of at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

42. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocotrienol levels relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

43. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocotrienol levels of at least about 25% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

44. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocotrienol levels of at least about 250% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

45. The transformed plant cell according to claim 13, wherein said plant cell exhibits increased γ-tocotrienol levels of at least about 2500% relative to a plant cell with a similar genetic background but lacking said exogenous nucleic acid molecule.

46. The transformed plant cell of claim 13, wherein said nucleic acid molecule further comprises a plastid targeting sequence, wherein said plastid targeting sequence is operably linked to said exogenous nucleic acid molecule to cause a transcript of said exogenous nucleic acid molecule to further encode a plastid peptide target sequence operably linked to said amino acid sequence.

47. The transformed plant cell of claim 13, further comprising an expression cassette which expresses phytyl prenyltransferase.

48. The transformed plant cell of claim 47, wherein said nucleic acid molecule further comprises said expression cassette.

49. The transformed plant cell of claim 13, wherein said nucleic acid molecule encodes a fragment of SEQ ID NOs:2 or 4, wherein said fragment has prephenate dehydrogenase activity.

50. A transformed plant cell having a nucleic acid molecule comprising as operably linked components: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a heterologous nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule encoding a protein comprising an amino acid sequence selected from the group consisting of: SEQ ID NO:2, SEQ ID NO:4 and fragments thereof comprising at least 20 contiguous amino acids.

51. A method of producing a plant having increased tocopherol levels comprising:

(A) transforming said plant with a nucleic acid molecule, wherein said nucleic acid molecule comprises a promoter region, wherein said promoter region is linked to a nucleic acid sequence that encodes a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs:2 and 4; and

(B) growing said plant.

52. The method of claim 51, wherein said nucleic acid molecule further comprises an expression cassette which expresses phytyl prenyltransferase.

53. The method of claim 52, wherein said nucleic acid molecule further comprises an expression cassette which expresses hydroxyphenylpyruvate dehydrogenase.

54. The method of claim 51, wherein said nucleic acid molecule further comprises two or more expression cassettes, each of which expresses a member selected from the group consisting of *slr 1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, and an antisense construct for homogentistic acid dioxygenase.

55. The method of claim 51, wherein said nucleic acid molecule further comprises a nucleic acid sequence encoding *HPPD* and either *slr 1736* or *ATPT2*.

56. The method of claim 51, wherein said nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of an mRNA molecule, and wherein expression of said nucleic acid molecule results in overexpression of said protein.

57. The method of producing a plant according to claim 51, wherein said plant is selected from the group of canola, a maize, *Arabidiosus*, *Brassica campestris*, a *Brassica napus*, soybean, crambe, mustard, castor bean, peanut, sesame, cottonseed, linseed, safflower, oil palm, flax and sunflower.

58. The method of producing a plant according to claim 51, wherein said plant is canola.

59. The method of producing a plant according to claim 51, wherein said plant is soybean.

60. The method of producing a plant according to claim 51, wherein said plant is *Brassica napus*.

61. The method of producing a plant according to claim 51, wherein said plant exhibits increased α-tocopherol levels relative to a plant with a similar genetic background but lacking said nucleic acid molecule.

62. The method of producing a plant according to claim 51, wherein said plant exhibits increased γ-tocopherol levels relative to a plant with a similar genetic background but lacking said nucleic acid molecule.

63. The method of producing a plant according to claim 51, wherein said plant

exhibits increased tocopherol levels relative to a plant with a similar genetic background but lacking said nucleic acid molecule.

64. The method of producing a plant according to claim 51, wherein said plant exhibits increased tocotrienol levels relative to a plant with a similar genetic background but lacking said nucleic acid molecule.

65. A method for reducing tocopherol levels in a plant comprising:

(A) transforming said plant with a nucleic acid molecule, wherein said nucleic acid molecule comprises as operably linked components an exogenous promoter region which functions in plant cells to cause the production of an mRNA molecule, a heterologous nucleic acid molecule having a transcribed strand and a non-transcribed strand, wherein said transcribed strand is complementary to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1 and 3; and wherein said nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant cells to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA sequence; and (B) growing said transformed plant.

66. A method for screening for increased tocopherol levels in a plant comprising interrogating genomic DNA for the presence or absence of a marker molecule that specifically hybridizes, under conditions 2 X SSC at 65°C, to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1 and 3 and complements thereof; and detecting said presence or absence of said marker, wherein said marker molecule is selected from the group consisting of SEQ ID NOs:1,3, complements thereof and fragments thereof.

67. A cell comprising a nucleic acid molecule comprising as operably linked components: (A) a promoter region which functions in a plant cell to cause the

production of an mRNA molecule; (B) a heterologous nucleic acid molecule, wherein said heterologous nucleic acid molecule encodes an enzyme with chorismate mutase and prephenate dehydrogenase activities or a fragment of said nucleic acid molecule comprising at least 20 contiguous amino acids; said heterologous nucleic acid molecule is selected from the group consisting of the nucleic acid sequence SEQ ID NO:1, the nucleic acid sequence SEQ ID NO:3, a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:2, and a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO:4.

68. The cell of claim 67, wherein said nucleic acid molecule further comprises an expression cassette which expresses phytyl prenyltransferase.

69. The cell of claim 68, wherein said nucleic acid molecule further comprises an expression cassette which expresses hydroxyphenylpyruvate dehydrogenase.

70. The cell of claim 67, wherein said nucleic acid molecule further comprises two or more expression cassettes, each of which expresses a member from the group consisting of: *slr 1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, and an antisense construct for homogentistic acid dioxygenase.

71. The cell according to claim 67, wherein said cell further comprises a nucleic acid sequence encoding *HPPD* and either *slr 1736* or *ATPT2*.

72. A cell according to claim 67, wherein said cell is a bacterial cell.

73. A cell according to claim 67, wherein said cell is a blue green algae cell.

74. Oil derived from a seed of a transformed plant having the nucleic acid molecule according to claim 1 and (C) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule, said seed having the nucleic acid molecule according to claim 1 and having an increased tocopherol level relative to seeds from a plant having similar genetic background but lacking said nucleic acid molecule.

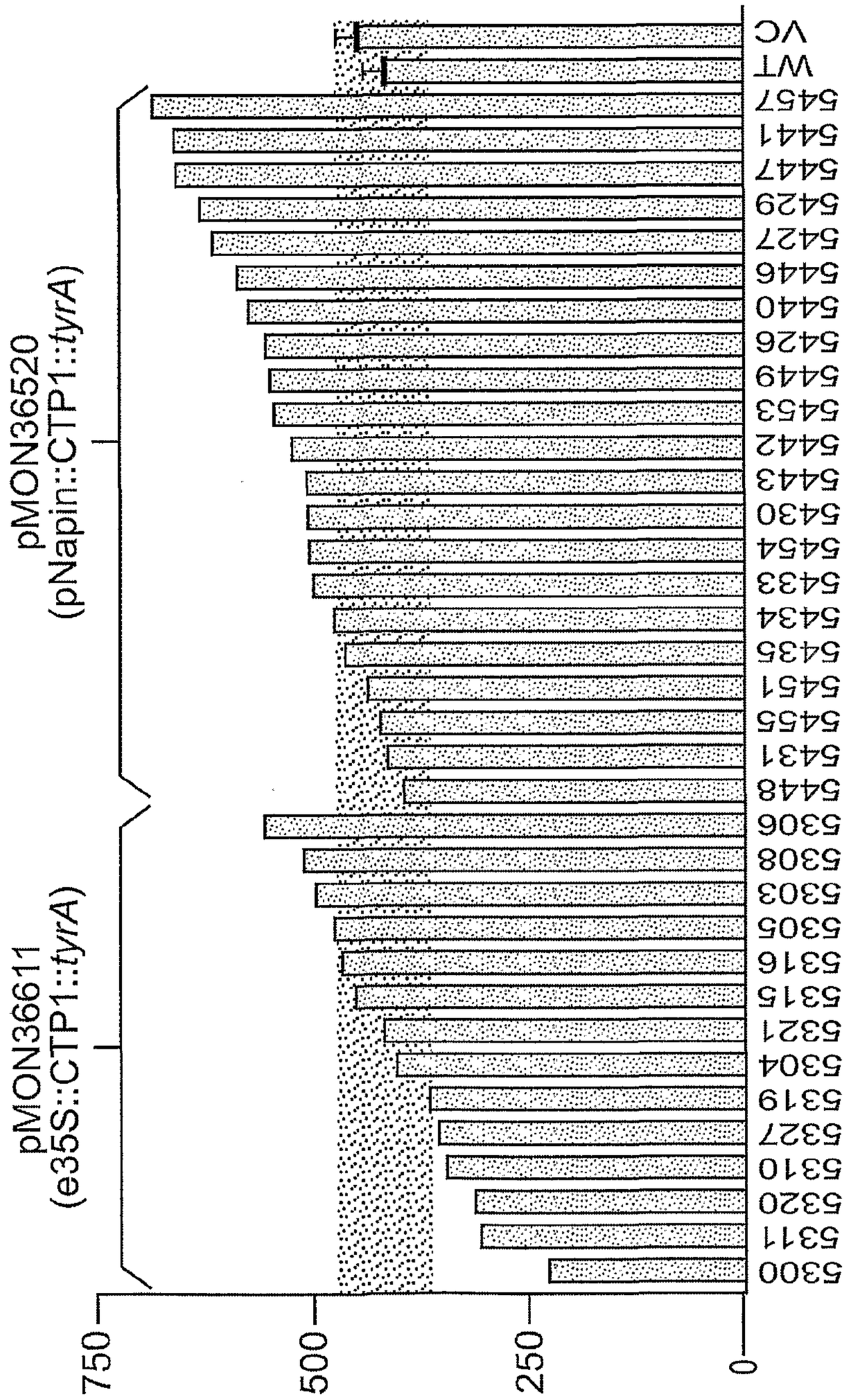

pMON36520
(pNapin::CTP1::tyrA)
pMON36611
(e35S::CTP1::tyrA)
LINE DESIGNATION
VC
WT
5457
5441
5447
5429
5427
5446
5440
5426
5449
5453
5442
5443
5430
5454
5433
5434
5435
5451
5455
5431
5448
5306
5308
5303
5305
5316
5315
5321
5304
5319
5327
5310
5320
5311
5300
750
500
250
0
TOTAL SEED TOCOPHEROL
[ng/mg]

FIG. 1

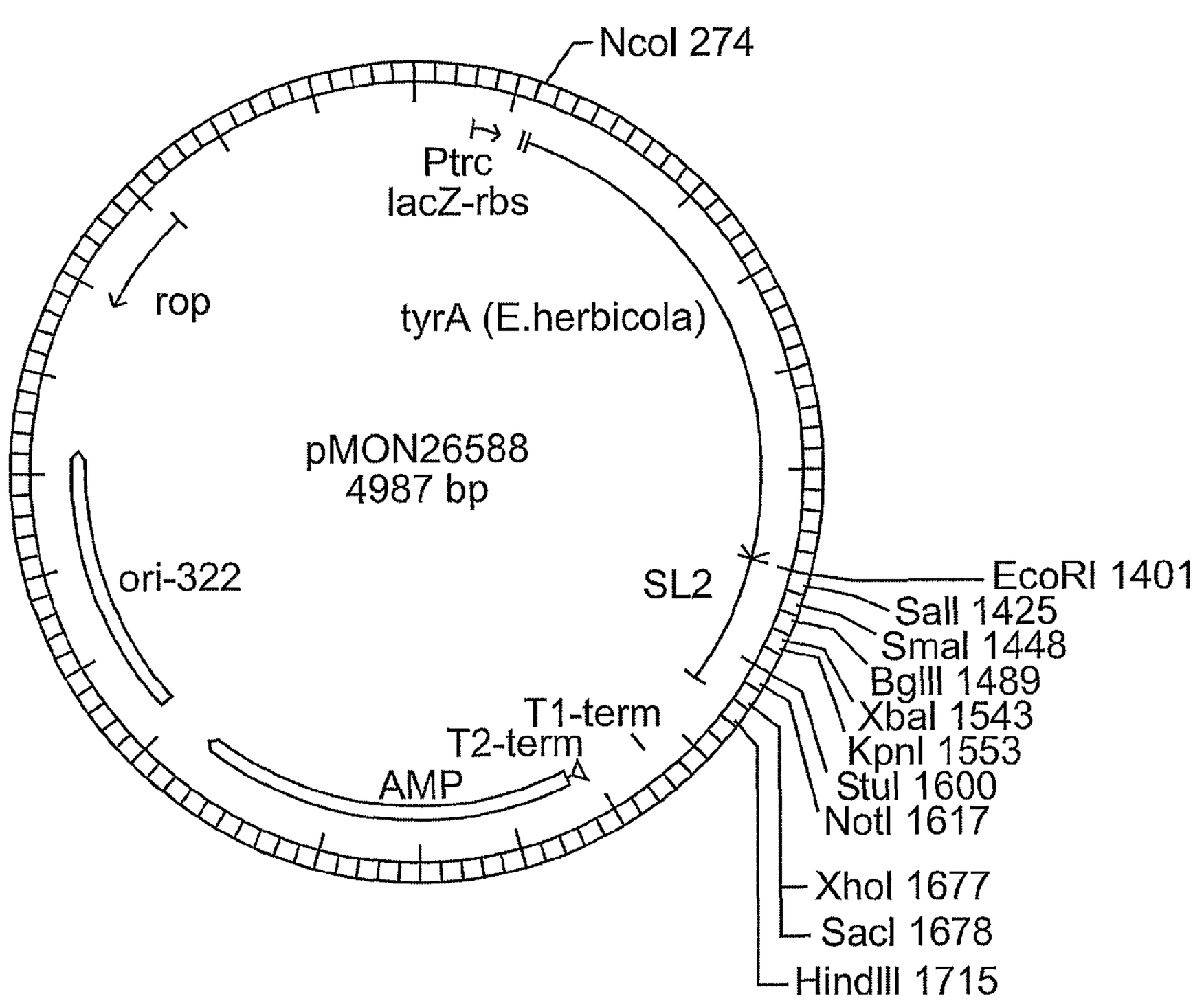

NcoI 274
Ptrc
lacZ-rbs
rop
tyrA (E.herbicola)
pMON26588
4987 bp
ori-322
SL2
EcoRI 1401
SalI 1425
SmaI 1448
BglII 1489
XbaI 1543
KpnI 1553
StuI 1600
NotI 1617
T1-term
T2-term
AMP
XhoI 1677
SacI 1678
HindIII 1715

FIG. 2

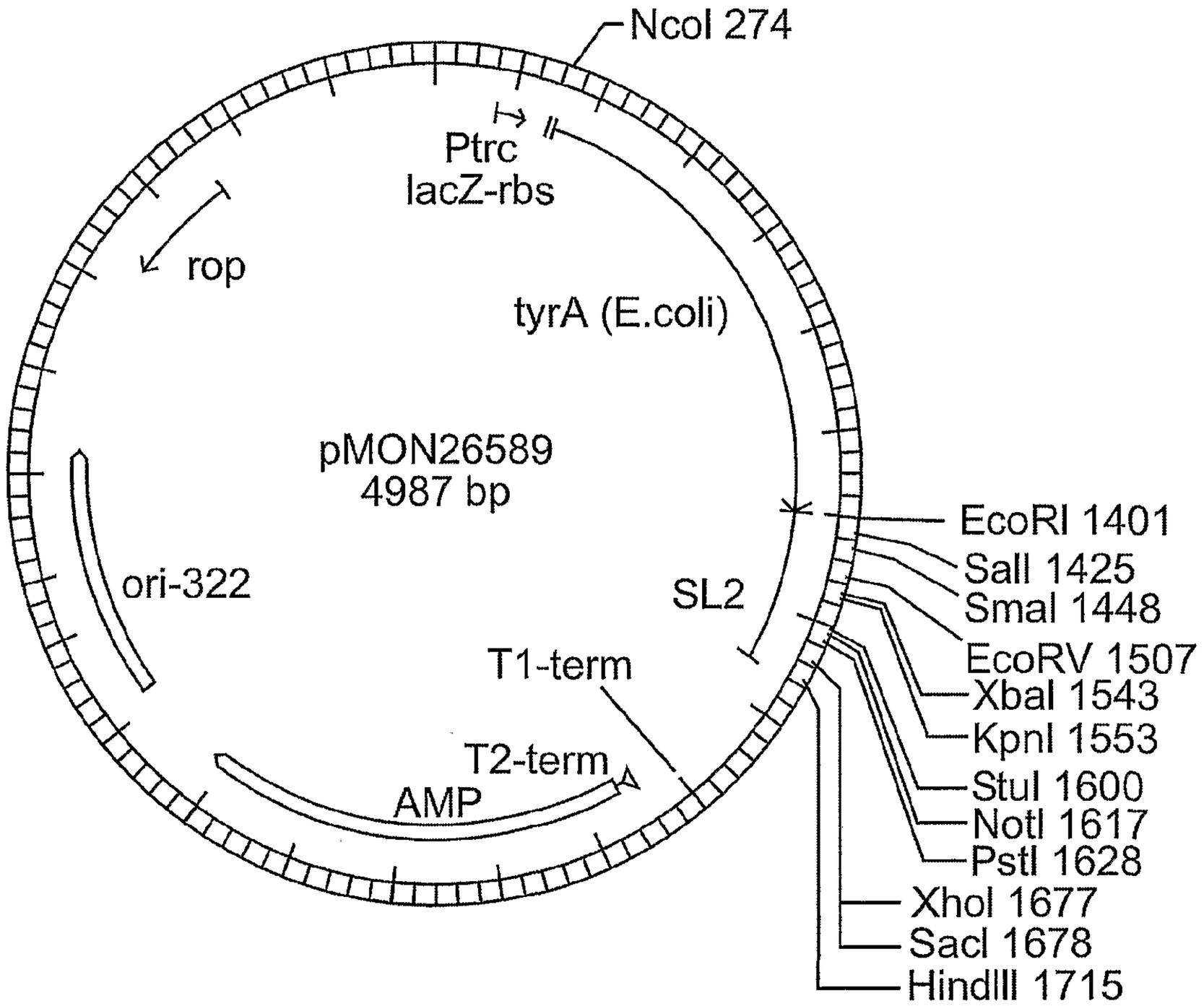

NcoI 274
Ptrc
lacZ-rbs
rop
tyrA (E.coli)
pMON26589
4987 bp
ori-322
SL2
T1-term
T2-term
AMP
EcoRI 1401
SalI 1425
SmaI 1448
EcoRV 1507
XbaI 1543
KpnI 1553
StuI 1600
NotI 1617
PstI 1628
XhoI 1677
SacI 1678
HindIII 1715

FIG. 3

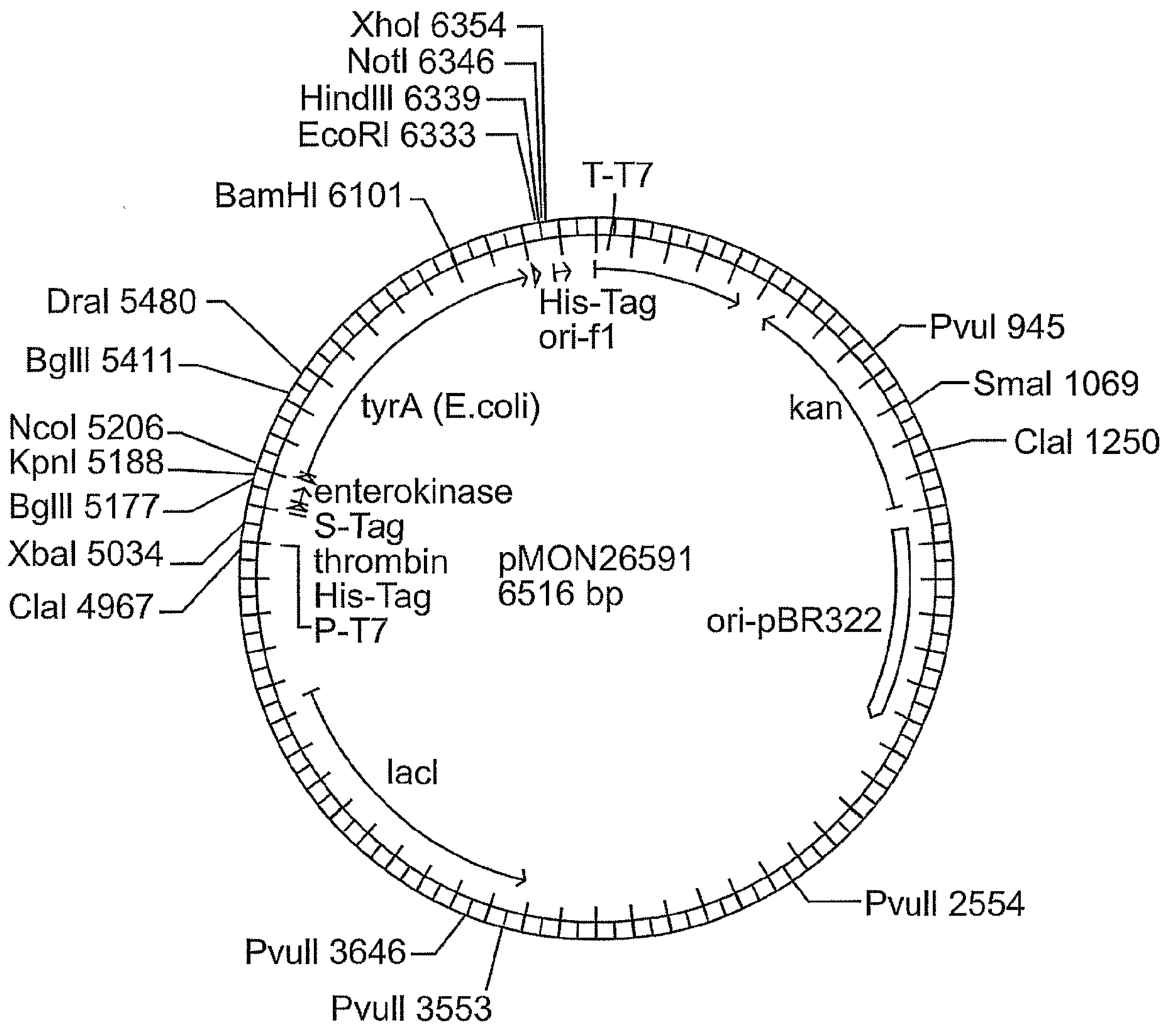

XhoI 6354
NotI 6346
HindIII 6339
EcoRI 6333
BamHI 6101
T-T7
His-Tag
ori-f1
DraI 5480
BglII 5411
NcoI 5206
KpnI 5188
BglII 5177
XbaI 5034
ClaI 4967
tyrA (E.coli)
enterokinase
S-Tag
thrombin
His-Tag
P-T7
pMON26591
6516 bp
kan
PvuI 945
SmaI 1069
ClaI 1250
ori-pBR322
lacI
PvuII 2554
PvuII 3646
PvuII 3553

FIG. 4

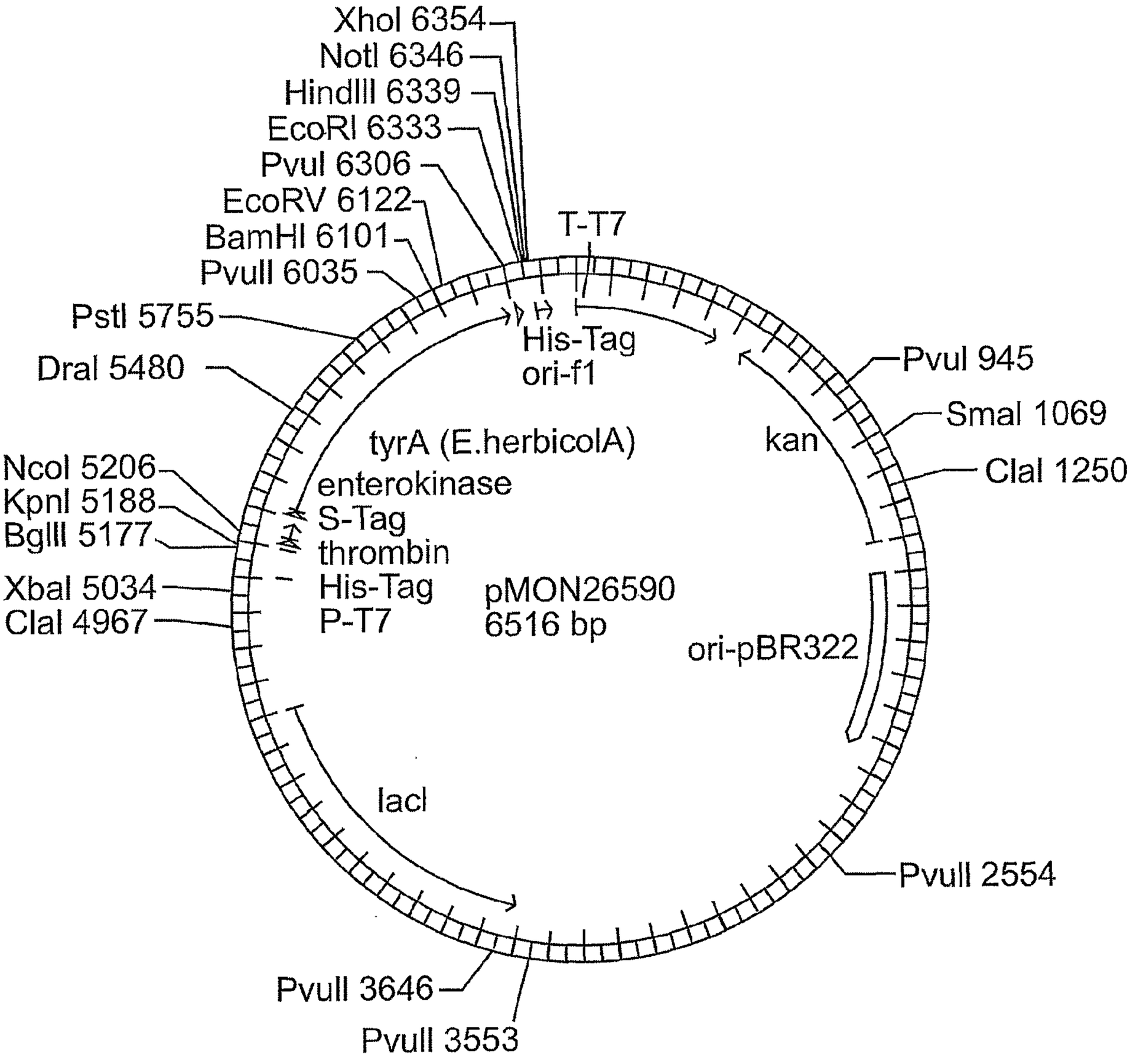

XhoI 6354
NotI 6346
HindIII 6339
EcoRI 6333
PvuI 6306
EcoRV 6122
BamHI 6101
PvuII 6035
PstI 5755
DraI 5480
NcoI 5206
KpnI 5188
BglII 5177
XbaI 5034
ClaI 4967
T-T7
His-Tag
ori-f1
PvuI 945
SmaI 1069
ClaI 1250
kan
tyrA (E.herbicolA)
enterokinase
S-Tag
thrombin
His-Tag
P-T7
pMON26590
6516 bp
ori-pBR322
lacI
PvuII 2554
PvuII 3646
PvuII 3553

FIG. 5

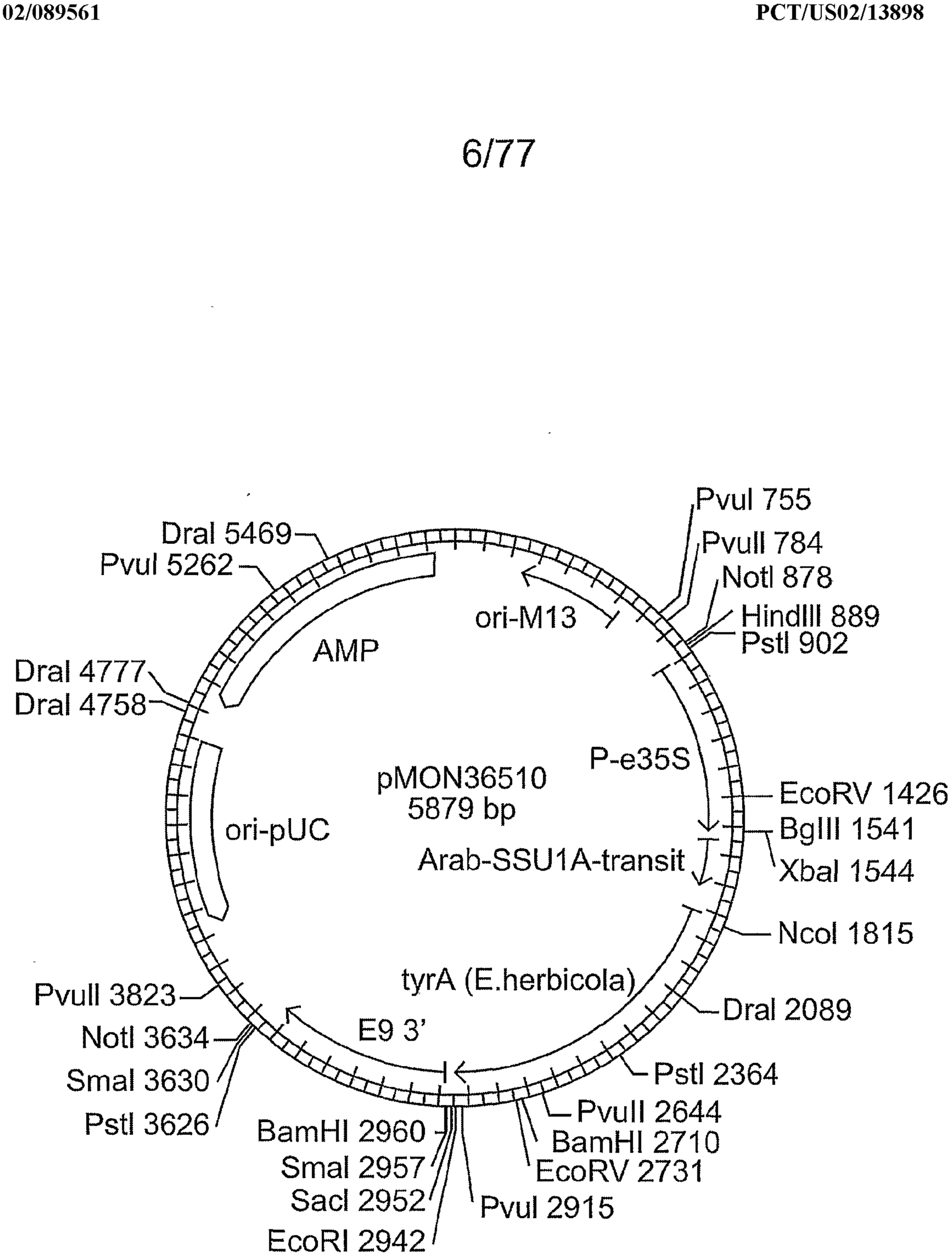

Pvul 755
PvuII 784
NotI 878
HindIII 889
PstI 902
ori-M13
AMP
DraI 5469
Pvul 5262
DraI 4777
DraI 4758
P-e35S
pMON36510
5879 bp
EcoRV 1426
BglII 1541
XbaI 1544
Arab-SSU1A-transit
ori-pUC
NcoI 1815
tyrA (E.herbicola)
DraI 2089
PvuII 3823
NotI 3634
SmaI 3630
PstI 3626
E9 3'
PstI 2364
PvuII 2644
BamHI 2710
EcoRV 2731
BamHI 2960
SmaI 2957
SacI 2952
EcoRI 2942
Pvul 2915

FIG. 6

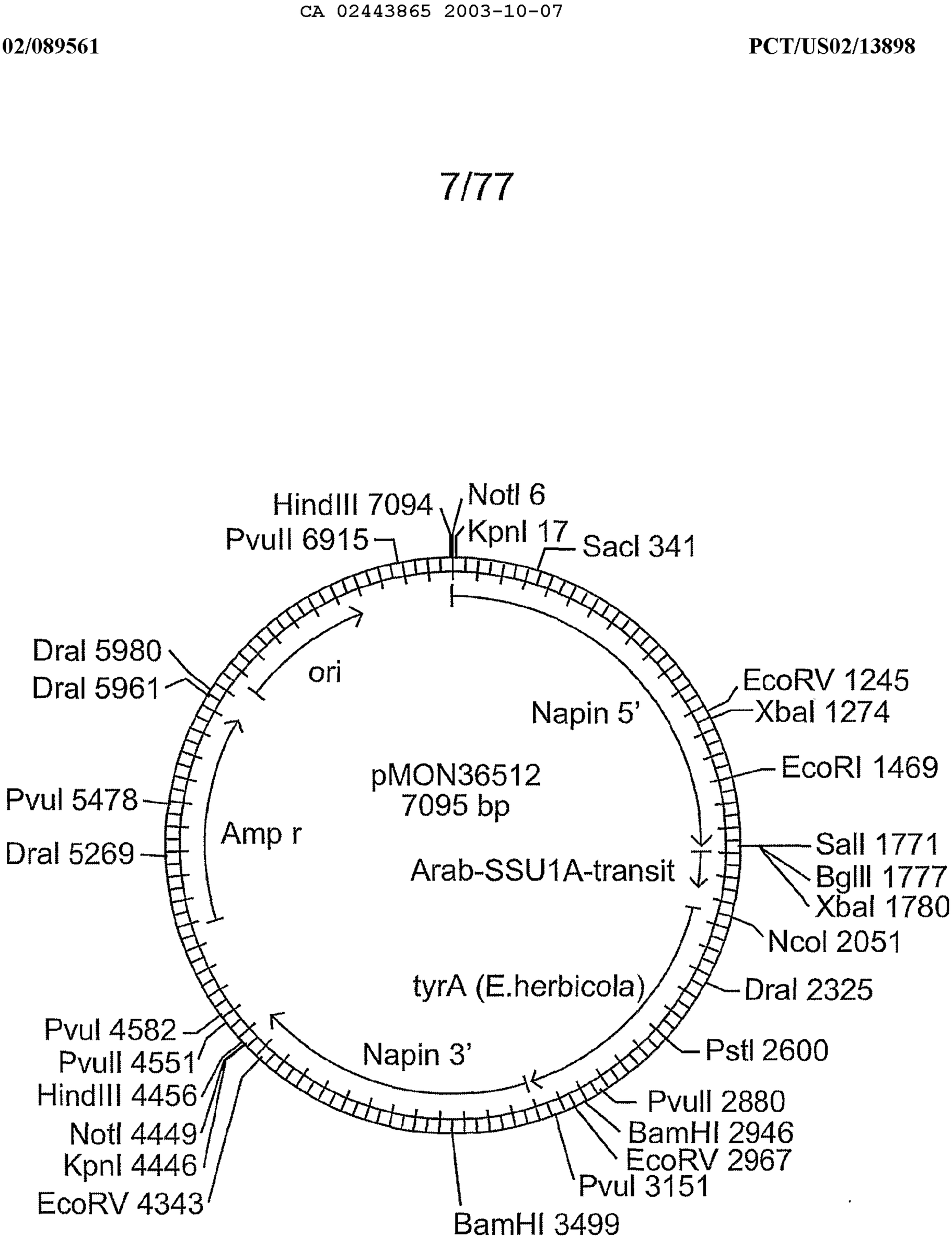

HindIII 7094
NotI 6
KpnI 17
PvuII 6915
SacI 341
DraI 5980
DraI 5961
ori
EcoRV 1245
XbaI 1274
Napin 5'
pMON36512
7095 bp
EcoRI 1469
PvuI 5478
Amp r
SalI 1771
BglII 1777
XbaI 1780
DraI 5269
Arab-SSU1A-transit
NcoI 2051
tyrA (E.herbicola)
DraI 2325
PvuI 4582
PvuII 4551
Napin 3'
PstI 2600
HindIII 4456
PvuII 2880
BamHI 2946
NotI 4449
EcoRV 2967
KpnI 4446
PvuI 3151
EcoRV 4343
BamHI 3499

FIG. 7

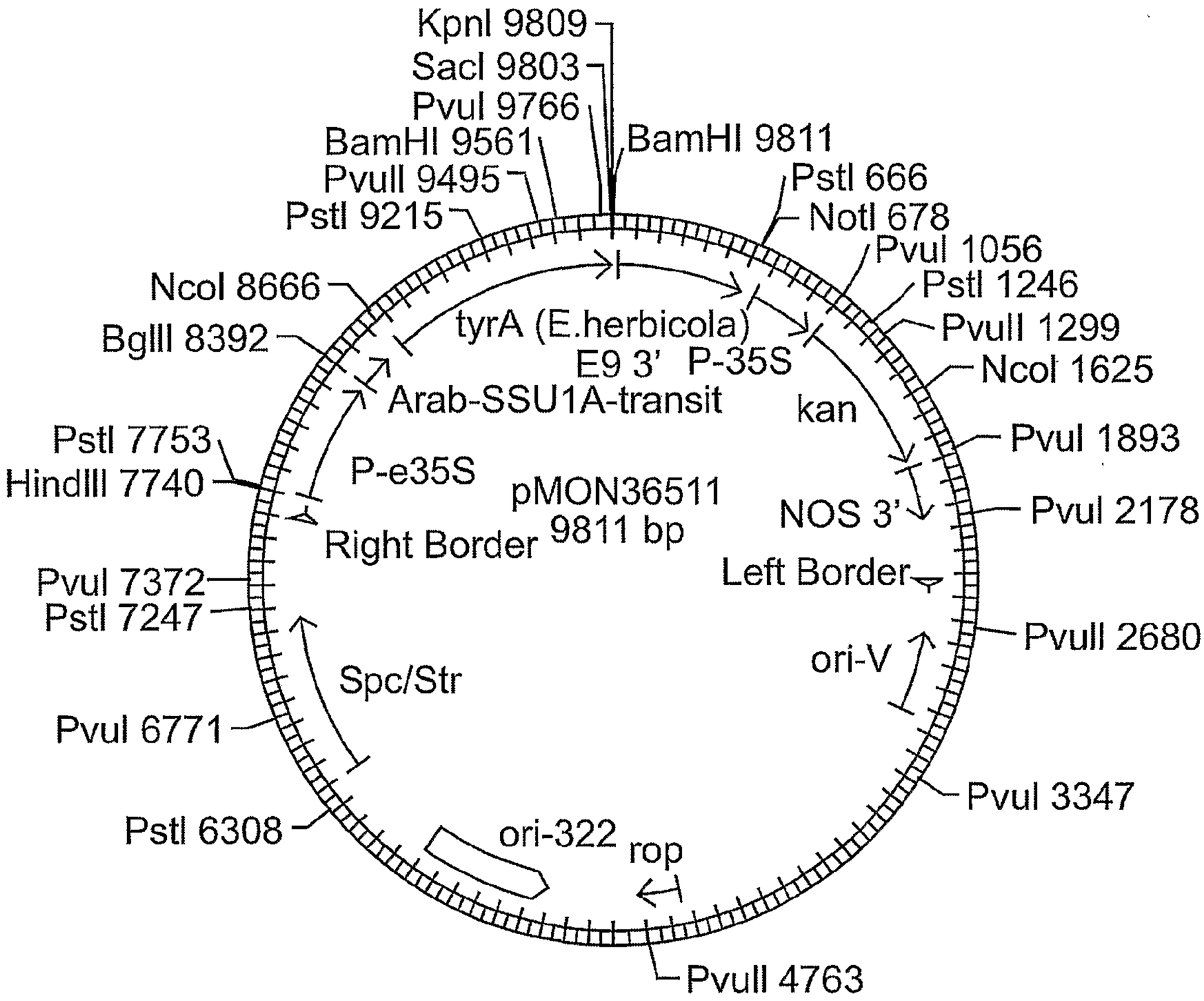

KpnI 9809
SacI 9803
PvuI 9766
BamHI 9561
BamHI 9811
PvuII 9495
PstI 666
PstI 9215
NotI 678
PvuI 1056
NcoI 8666
PstI 1246
tyrA (E.herbicola)
BglII 8392
PvuII 1299
E9 3' P-35S
NcoI 1625
Arab-SSU1A-transit
kan
PstI 7753
PvuI 1893
HindIII 7740
P-e35S
pMON36511
9811 bp
NOS 3'
PvuI 2178
Right Border
Left Border
PvuI 7372
PstI 7247
PvuII 2680
ori-V
Spc/Str
PvuI 6771
PvuII 3347
PstI 6308
ori-322
rop
PvuII 4763

FIG. 8

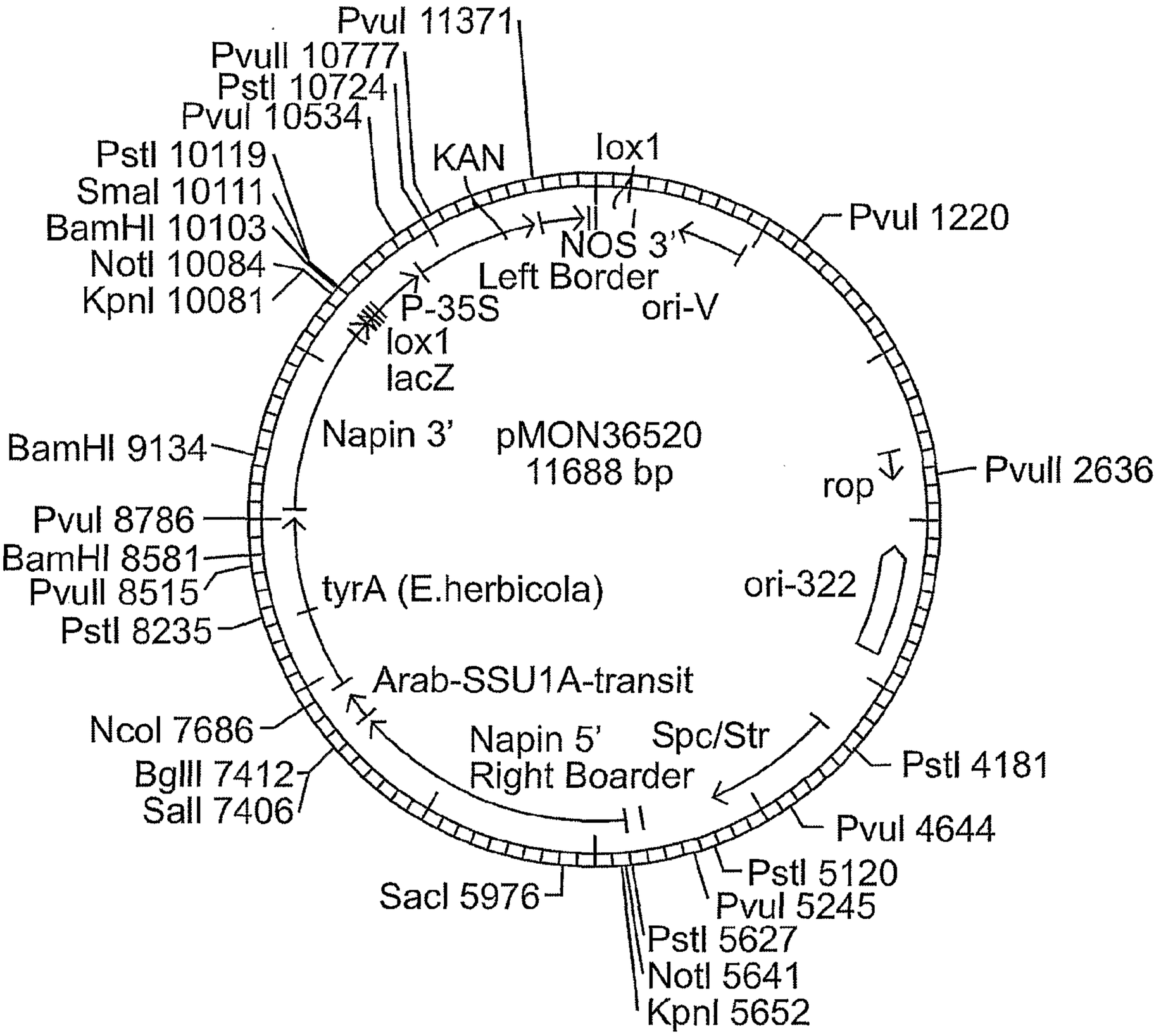

Pvul 11371
Pvull 10777
Pstl 10724
Pvul 10534
Pstl 10119
SmaI 10111
BamHI 10103
NotI 10084
KpnI 10081
KAN
lox1
Pvul 1220
NOS 3'
Left Border
P-35S
ori-V
lox1
lacZ
Napin 3'
pMON36520
11688 bp
BamHI 9134
rop
Pvull 2636
Pvul 8786
BamHI 8581
Pvull 8515
Pstl 8235
tyrA (E.herbicola)
ori-322
Arab-SSU1A-transit
NcoI 7686
Napin 5'
Spc/Str
Right Boarder
BglII 7412
SalI 7406
Pstl 4181
Pvul 4644
Pstl 5120
SacI 5976
Pvul 5245
Pstl 5627
NotI 5641
KpnI 5652

FIG. 9

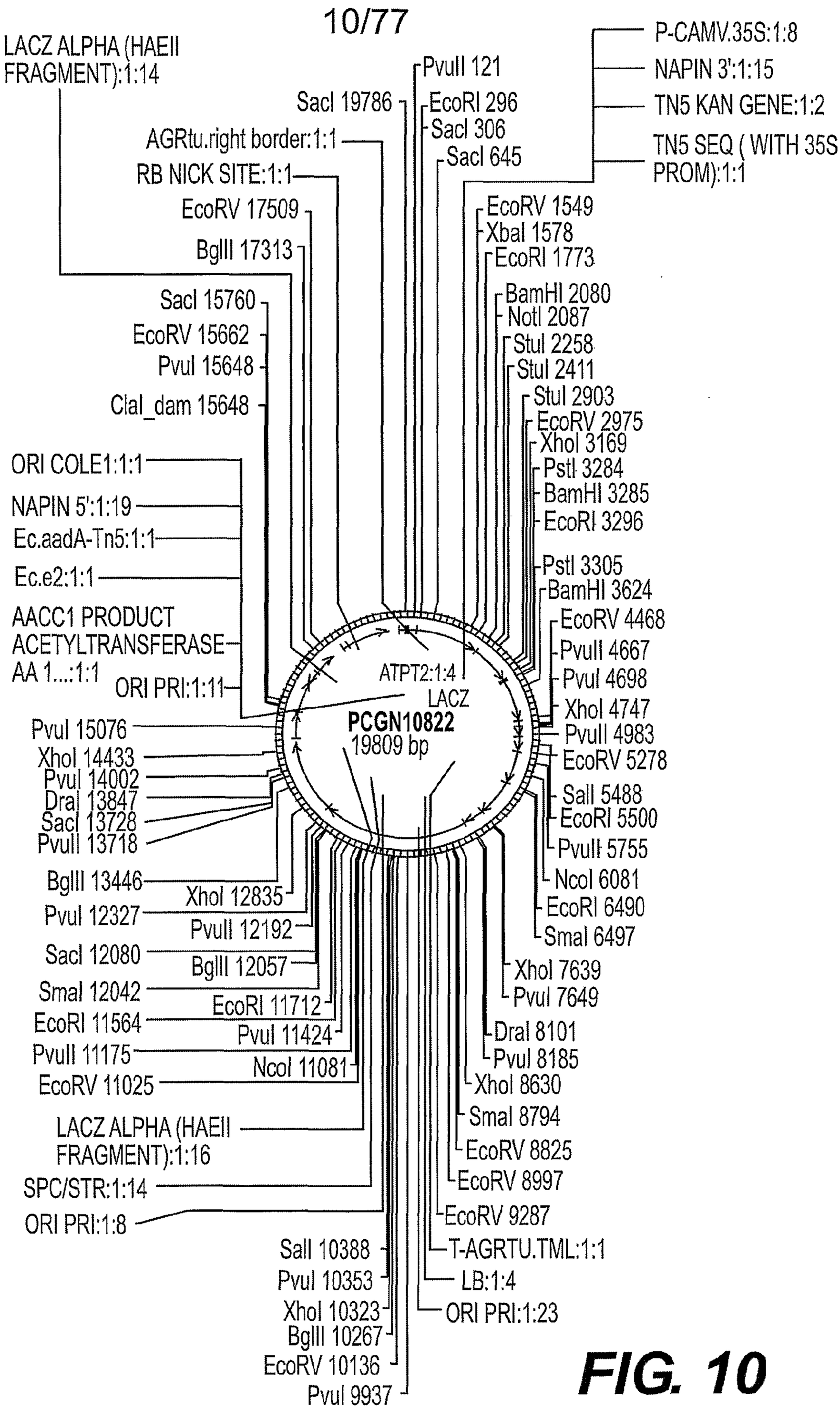

LACZ ALPHA (HAEII FRAGMENT):1:14
AGRtu.right border:1:1
RB NICK SITE:1:1
EcoRV 17509
BglII 17313
SacI 15760
EcoRV 15662
PvuI 15648
ClaI_dam 15648
ORI COLE1:1:1
NAPIN 5':1:19
Ec.aadA-Tn5:1:1
Ec.e2:1:1
AACC1 PRODUCT ACETYLTRANSFERASE AA 1...:1:1
ORI PRI:1:11
PvuI 15076
XhoI 14433
PvuI 14002
DraI 13847
SacI 13728
PvuII 13718
BglII 13446
PvuI 12327
SacI 12080
SmaI 12042
EcoRI 11564
PvuII 11175
EcoRV 11025
LACZ ALPHA (HAEII FRAGMENT):1:16
SPC/STR:1:14
ORI PRI:1:8
SalI 10388
PvuI 10353
XhoI 10323
BglII 10267
EcoRV 10136
PvuI 9937
SacI 19786
PvuII 121
EcoRI 296
SacI 306
SacI 645
P-CAMV.35S:1:8
NAPIN 3':1:15
TN5 KAN GENE:1:2
TN5 SEQ (WITH 35S PROM):1:1
EcoRV 1549
XbaI 1578
EcoRI 1773
BamHI 2080
NotI 2087
StuI 2258
StuI 2411
StuI 2903
EcoRV 2975
XhoI 3169
PstI 3284
BamHI 3285
EcoRI 3296
PstI 3305
BamHI 3624
EcoRV 4468
PvuII 4667
PvuI 4698
XhoI 4747
PvuII 4983
EcoRV 5278
SalI 5488
EcoRI 5500
PvuII 5755
NcoI 6081
EcoRI 6490
SmaI 6497
XhoI 7639
PvuI 7649
DraI 8101
PvuI 8185
XhoI 8630
SmaI 8794
EcoRV 8825
EcoRV 8997
EcoRV 9287
T-AGRTU.TML:1:1
LB:1:4
ORI PRI:1:23
XhoI 12835
PvuII 12192
BglII 12057
EcoRI 11712
PvuI 11424
NcoI 11081
ATPT2:1:4
LACZ
PCGN10822
19809 bp

FIG. 10

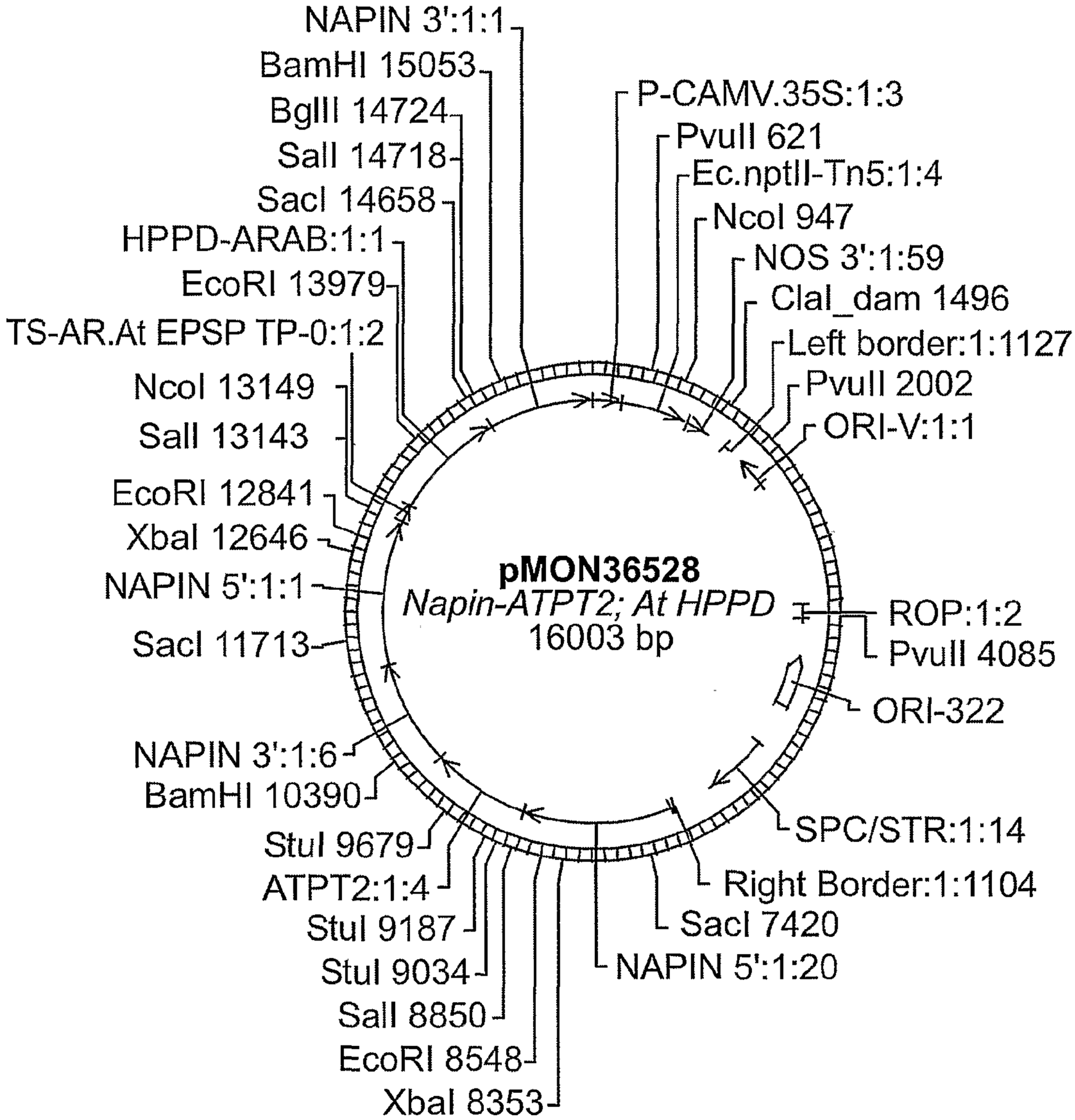

NAPIN 3':1:1
BamHI 15053
BglII 14724
SalI 14718
SacI 14658
HPPD-ARAB:1:1
EcoRI 13979
TS-AR.At EPSP TP-0:1:2
NcoI 13149
SalI 13143
EcoRI 12841
XbaI 12646
NAPIN 5':1:1
SacI 11713
NAPIN 3':1:6
BamHI 10390
StuI 9679
ATPT2:1:4
StuI 9187
StuI 9034
SalI 8850
EcoRI 8548
XbaI 8353
P-CAMV.35S:1:3
PvuII 621
Ec.nptII-Tn5:1:4
NcoI 947
NOS 3':1:59
ClaI_dam 1496
Left border:1:1127
PvuII 2002
ORI-V:1:1
pMON36528
Napin-ATPT2; At HPPD
16003 bp
ROP:1:2
PvuII 4085
ORI-322
SPC/STR:1:14
Right Border:1:1104
SacI 7420
NAPIN 5':1:20

FIG. 11

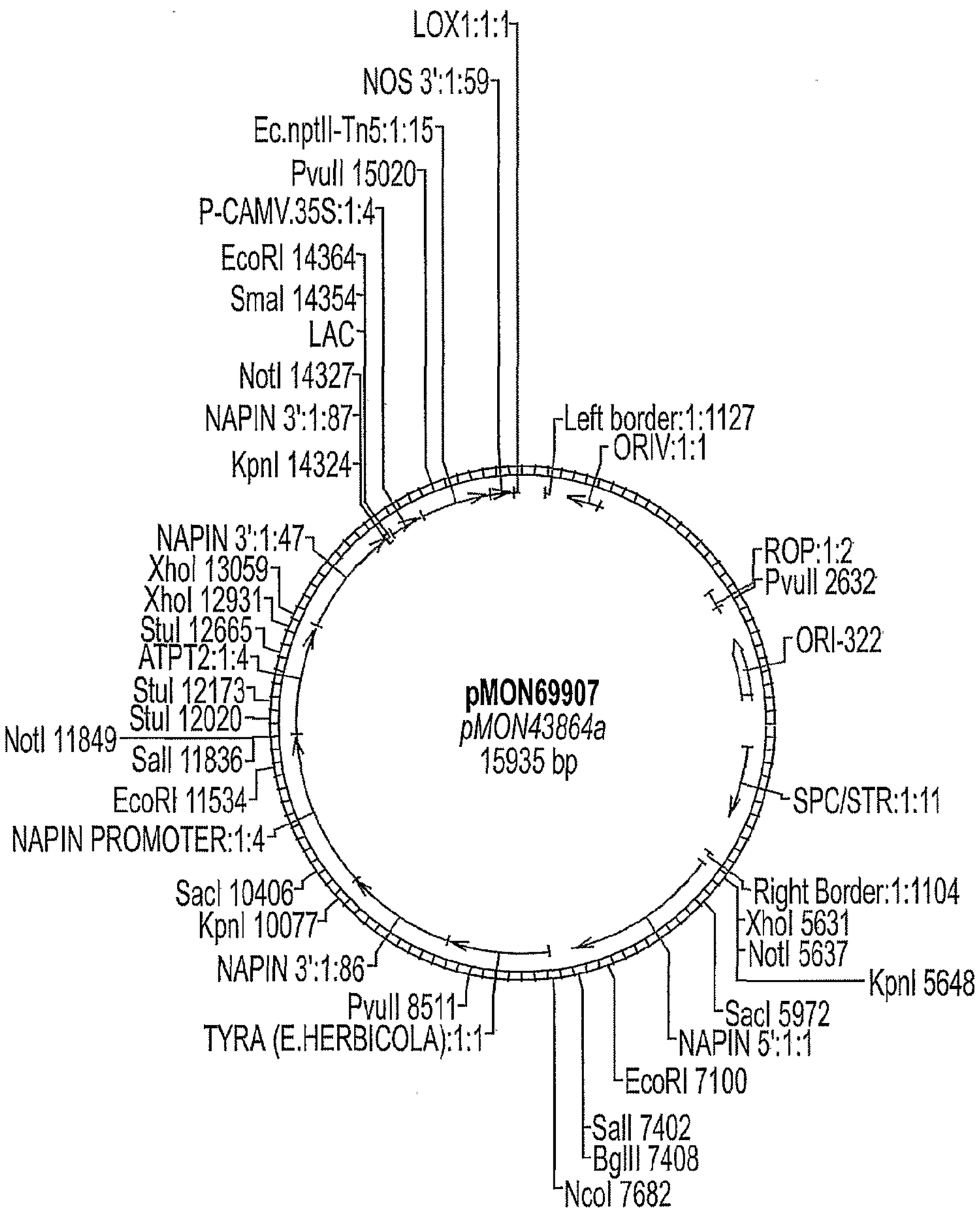

LOX1:1:1
NOS 3':1:59
Ec.nptII-Tn5:1:15
PvuII 15020
P-CAMV.35S:1:4
EcoRI 14364
SmaI 14354
LAC
NotI 14327
NAPIN 3':1:87
KpnI 14324
Left border:1:1127
ORIV:1:1
NAPIN 3':1:47
XhoI 13059
XhoI 12931
StuI 12665
ATPT2:1:4
StuI 12173
StuI 12020
NotI 11849
SalI 11836
EcoRI 11534
NAPIN PROMOTER:1:4
SacI 10406
KpnI 10077
NAPIN 3':1:86
PvuII 8511
TYRA (E.HERBICOLA):1:1
pMON69907
pMON43864a
15935 bp
ROP:1:2
PvuII 2632
ORI-322
SPC/STR:1:11
Right Border:1:1104
XhoI 5631
NotI 5637
KpnI 5648
SacI 5972
NAPIN 5':1:1
EcoRI 7100
SalI 7402
BglII 7408
NcoI 7682

FIG. 12

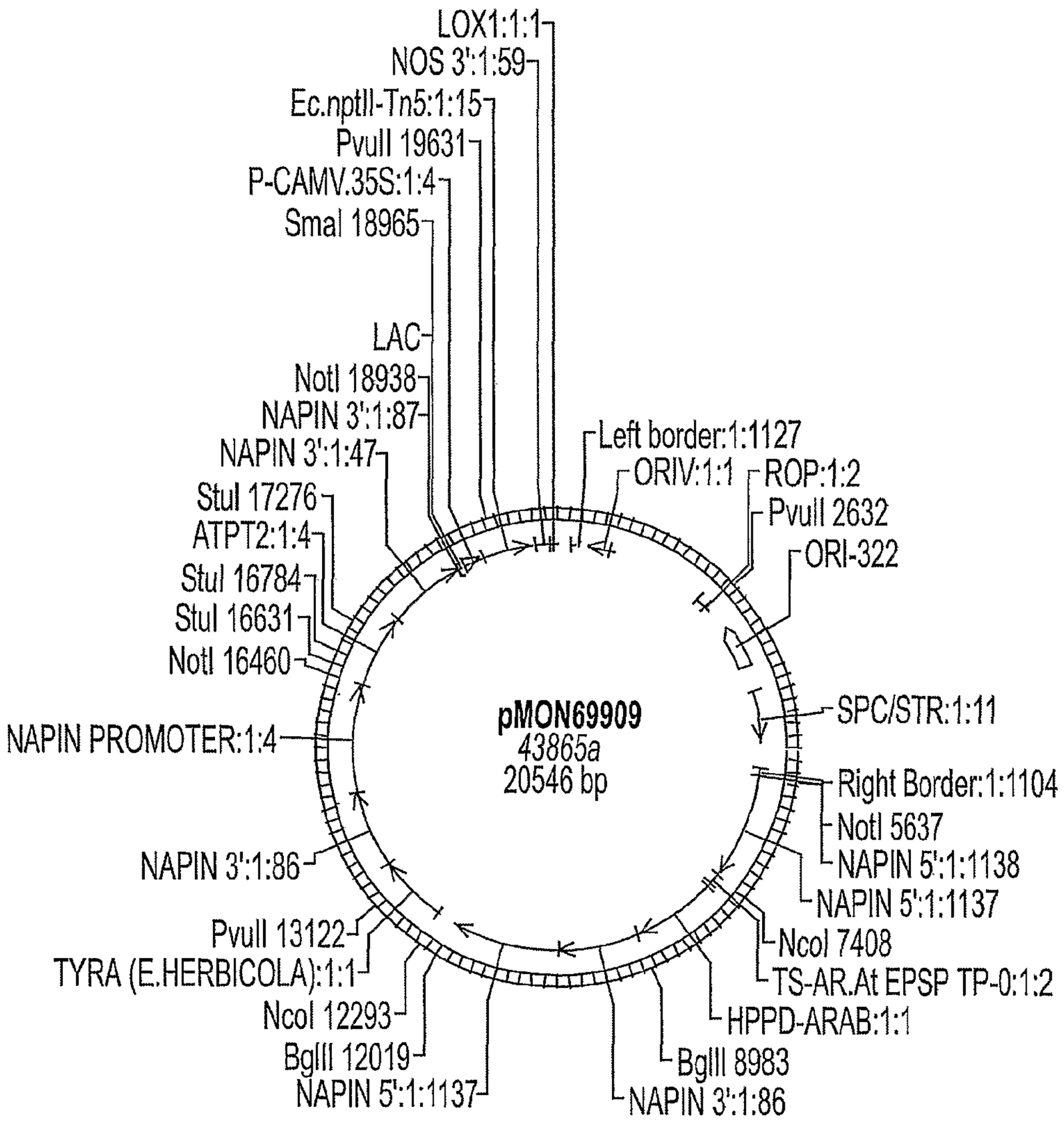

LOX1:1:1
NOS 3':1:59
Ec.nptII-Tn5:1:15
PvuII 19631
P-CAMV.35S:1:4
SmaI 18965
LAC
NotI 18938
NAPIN 3':1:87
NAPIN 3':1:47
StuI 17276
ATPT2:1:4
StuI 16784
StuI 16631
NotI 16460
NAPIN PROMOTER:1:4
NAPIN 3':1:86
PvuII 13122
TYRA (E.HERBICOLA):1:1
NcoI 12293
BglII 12019
NAPIN 5':1:1137
Left border:1:1127
ORIV:1:1
ROP:1:2
PvuII 2632
ORI-322
SPC/STR:1:11
Right Border:1:1104
NotI 5637
NAPIN 5':1:1138
NAPIN 5':1:1137
NcoI 7408
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
BglII 8983
NAPIN 3':1:86
pMON69909
43865a
20546 bp

FIG. 13

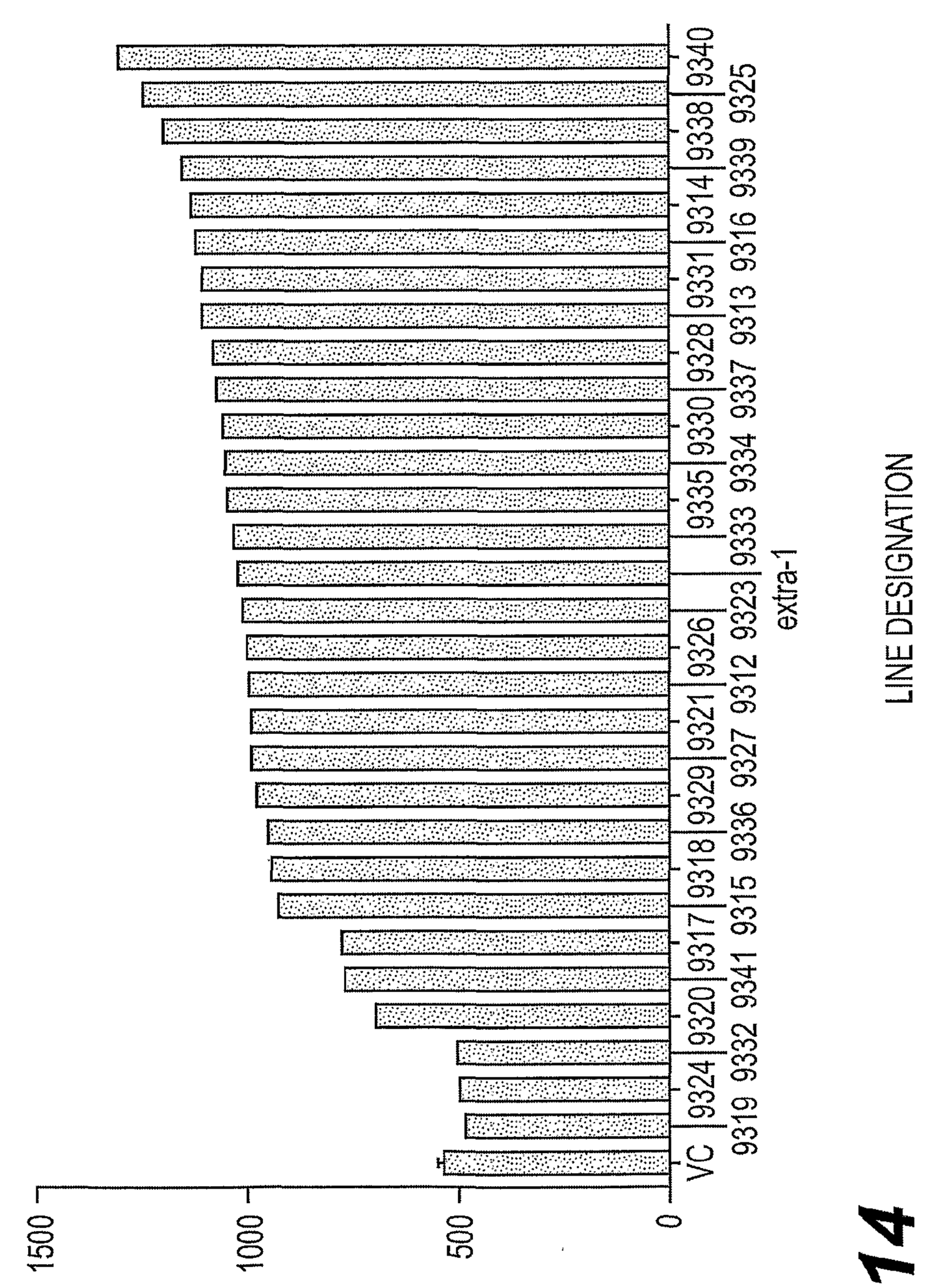

pMON 69907 [pnapin::ATPT2::napin 3'::pnapin::CTP1::tyrA::napin 3']
TOTAL SEED TOCOPHEROLS+TOCOTRIENOLS [ng/mg]
1500
1000
500
0
VC
9324
9320
9317
9318
9329
9321
9326
9335
9330
9328
9331
9314
9338
9340
9319 9332 9341 9315 9336 9327 9312 9323 9333 9334 9337 9313 9316 9339 9325
extra-1
LINE DESIGNATION

FIG. 14

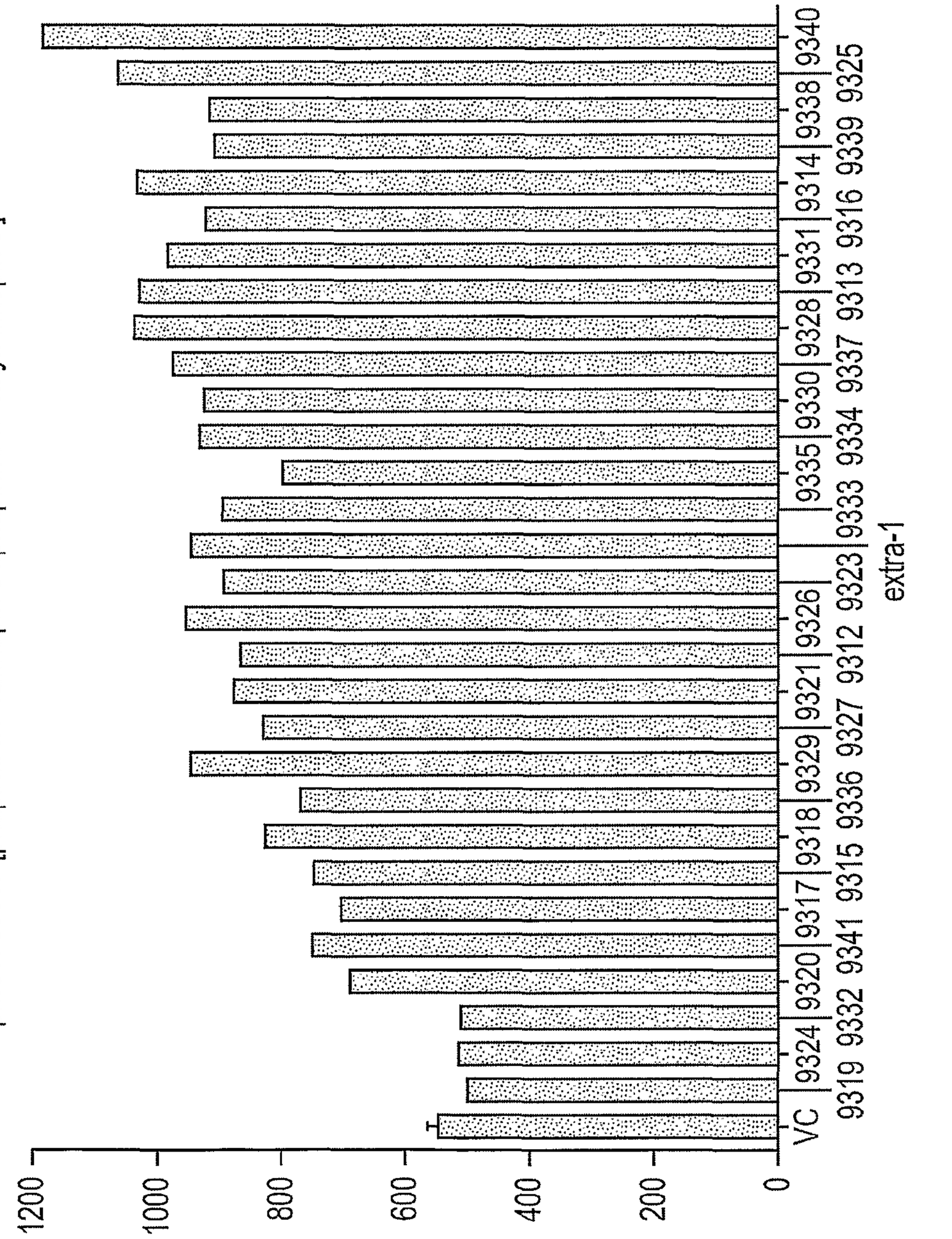

TOTAL SEED TOCOPHEROLS [ng/mg]
1200
1000
800
600
400
200
0
pMON 69907 [pnapin::ATPT2::napin 3'::pnapin::CTP1::tyrA::napin 3']
VC 9324 9320 9317 9318 9329 9321 9326 9335 9330 9328 9331 9314 9338 9340
9319 9332 9341 9315 9336 9327 9312 9323 9333 9334 9337 9313 9316 9339 9325
extra-1
LINE DESIGNATION

FIG. 15

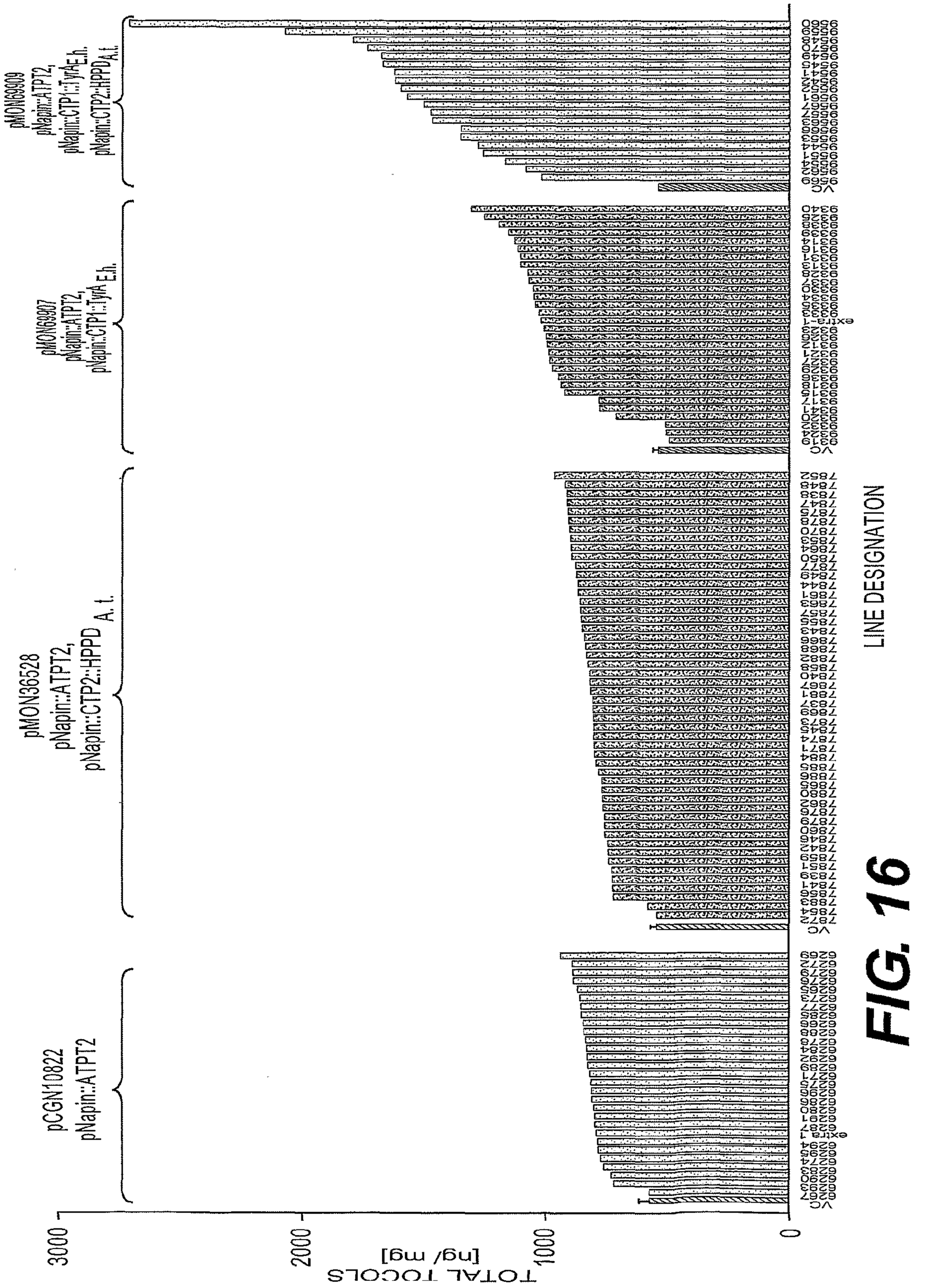

pMON69909
pNapin::ATPT2,
pNapin::CTP1::TyrA E.h.,
pNapin::CTP2::HPPD A.t.
pMON69907
pNapin::ATPT2,
pNapin::CTP1::TyrA E.h.
pMON36528
pNapin::ATPT2,
pNapin::CTP2::HPPD A.t.
pCGN10822
pNapin::ATPT2
3000
2000
1000
0
TOTAL TOCOLS
[ng/ mg]
LINE DESIGNATION

FIG. 16

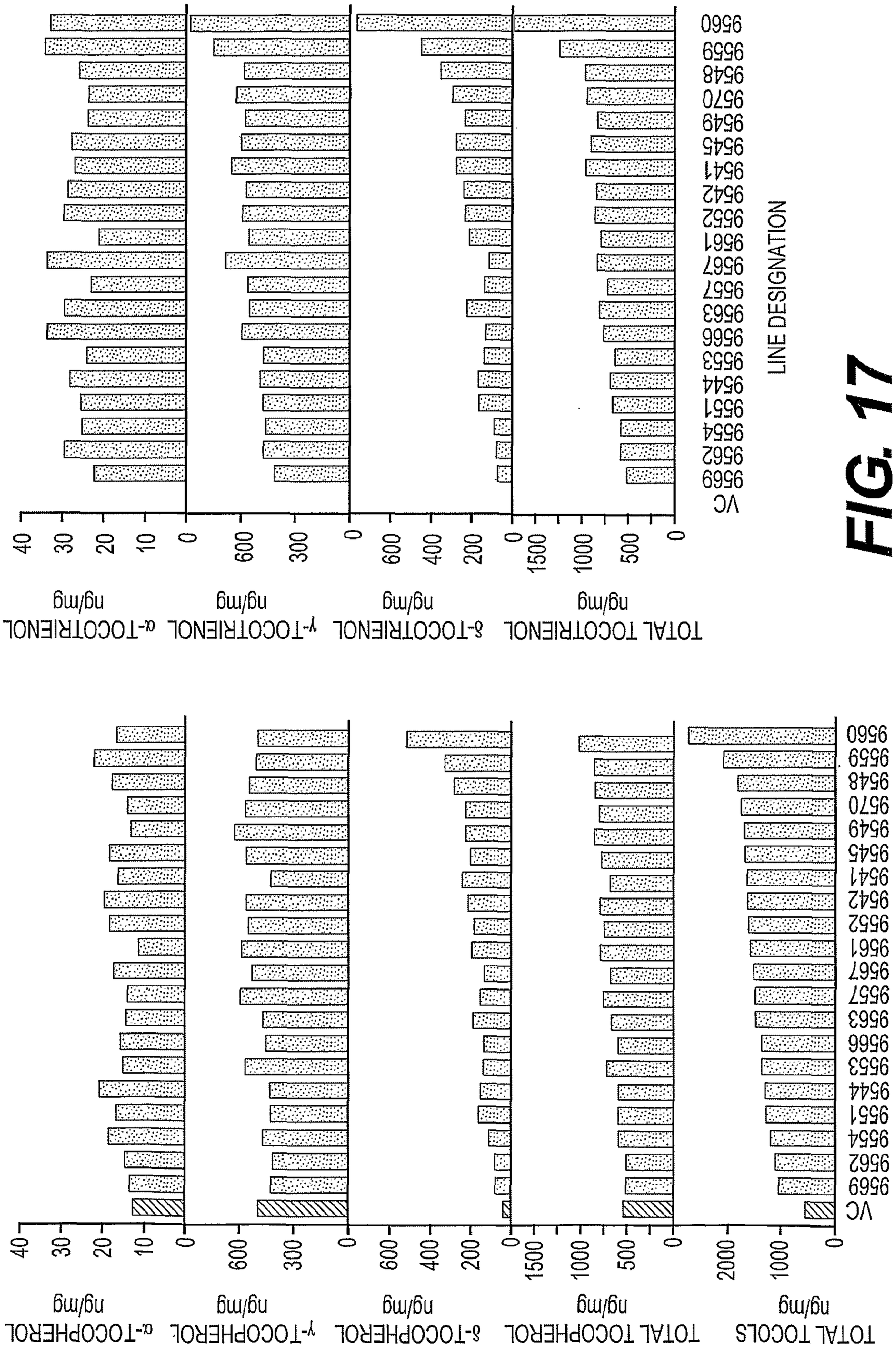

TOTAL TOCOTRIENOL ng/mg
δ-TOCOTRIENOL ng/mg
γ-TOCOTRIENOL ng/mg
α-TOCOTRIENOL ng/mg
LINE DESIGNATION
VC
9569
9562
9554
9551
9544
9553
9566
9563
9557
9567
9561
9552
9542
9541
9545
9549
9570
9548
9559
9560
TOTAL TOCOLS ng/mg
TOTAL TOCOPHEROL ng/mg
δ-TOCOPHEROL ng/mg
γ-TOCOPHEROL ng/mg
α-TOCOPHEROL ng/mg

FIG. 17

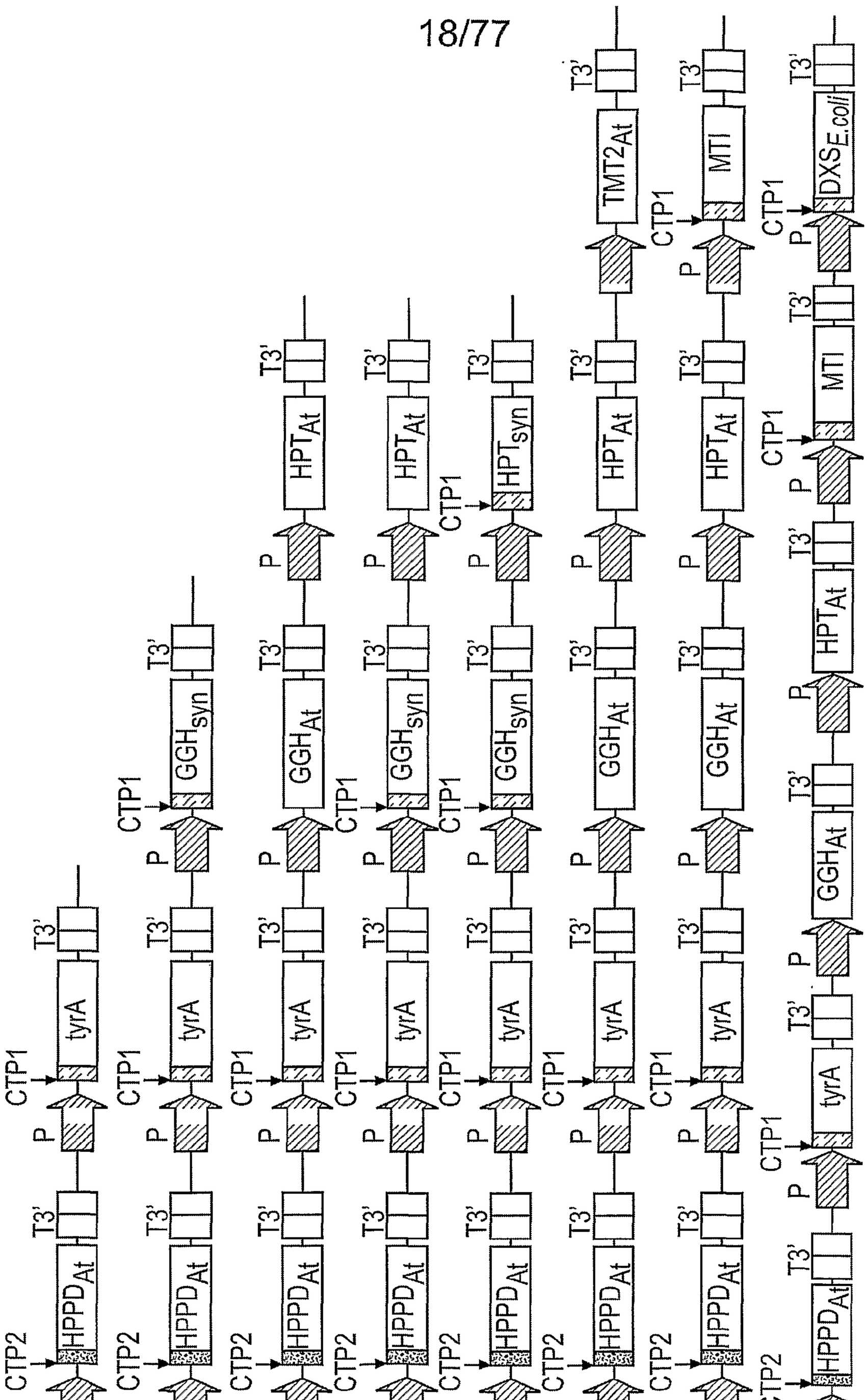

T3'
TMT2At
CTP1
P
MTI
DXSE.coli
HPTAt
HPTsyn
GGHsyn
GGHAt
tyrA
HPPDAt
CTP2

FIG. 18A

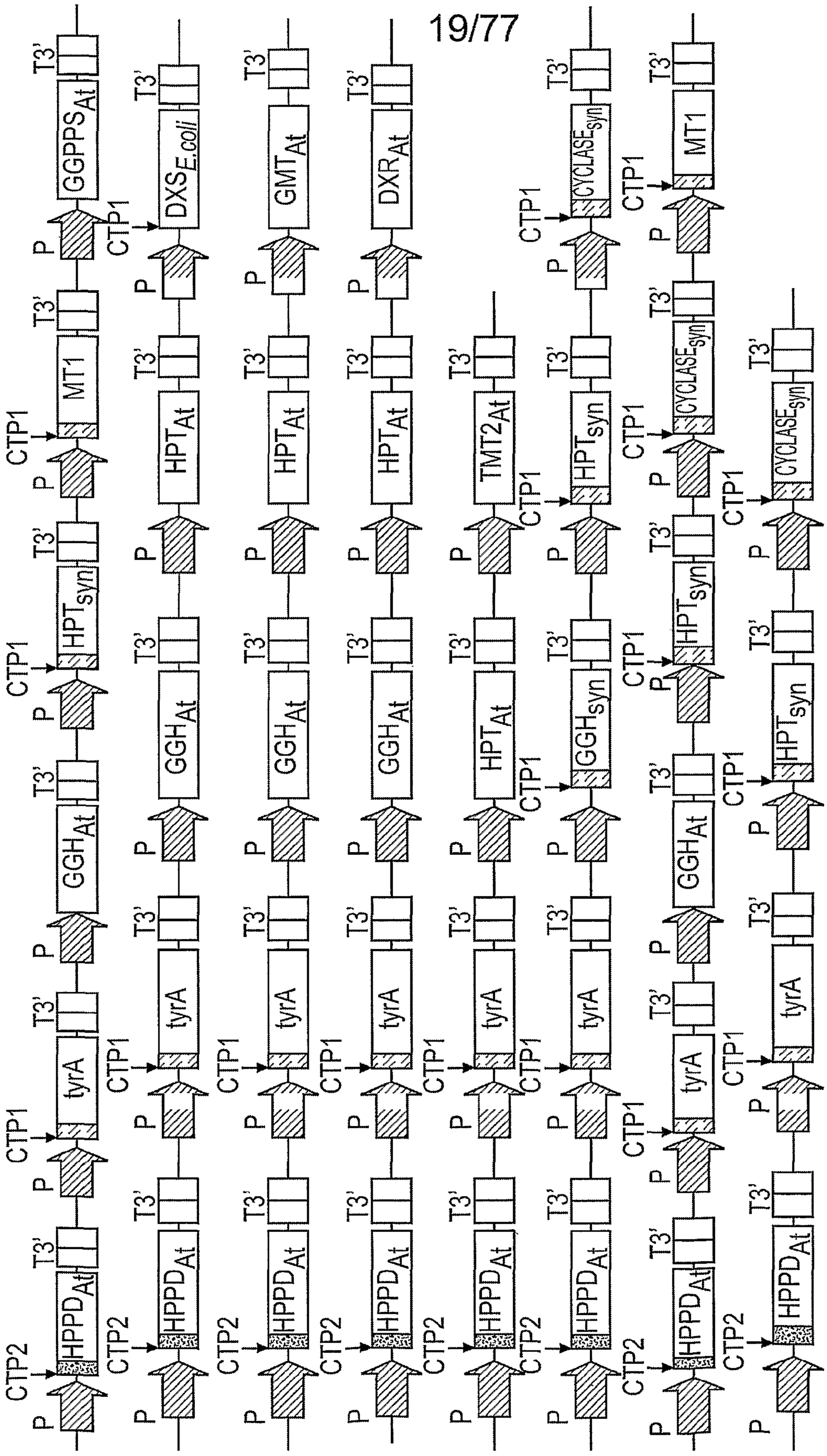

T3'
GGPPS$_{At}$
P
CTP1
DXS$_{E.coli}$
GMT$_{At}$
DXR$_{At}$
CYCLASE$_{syn}$
MT1
HPT$_{syn}$
HPT$_{At}$
TMT2$_{At}$
GGH$_{At}$
GGH$_{syn}$
tyrA
HPPD$_{At}$
CTP2

FIG. 18B

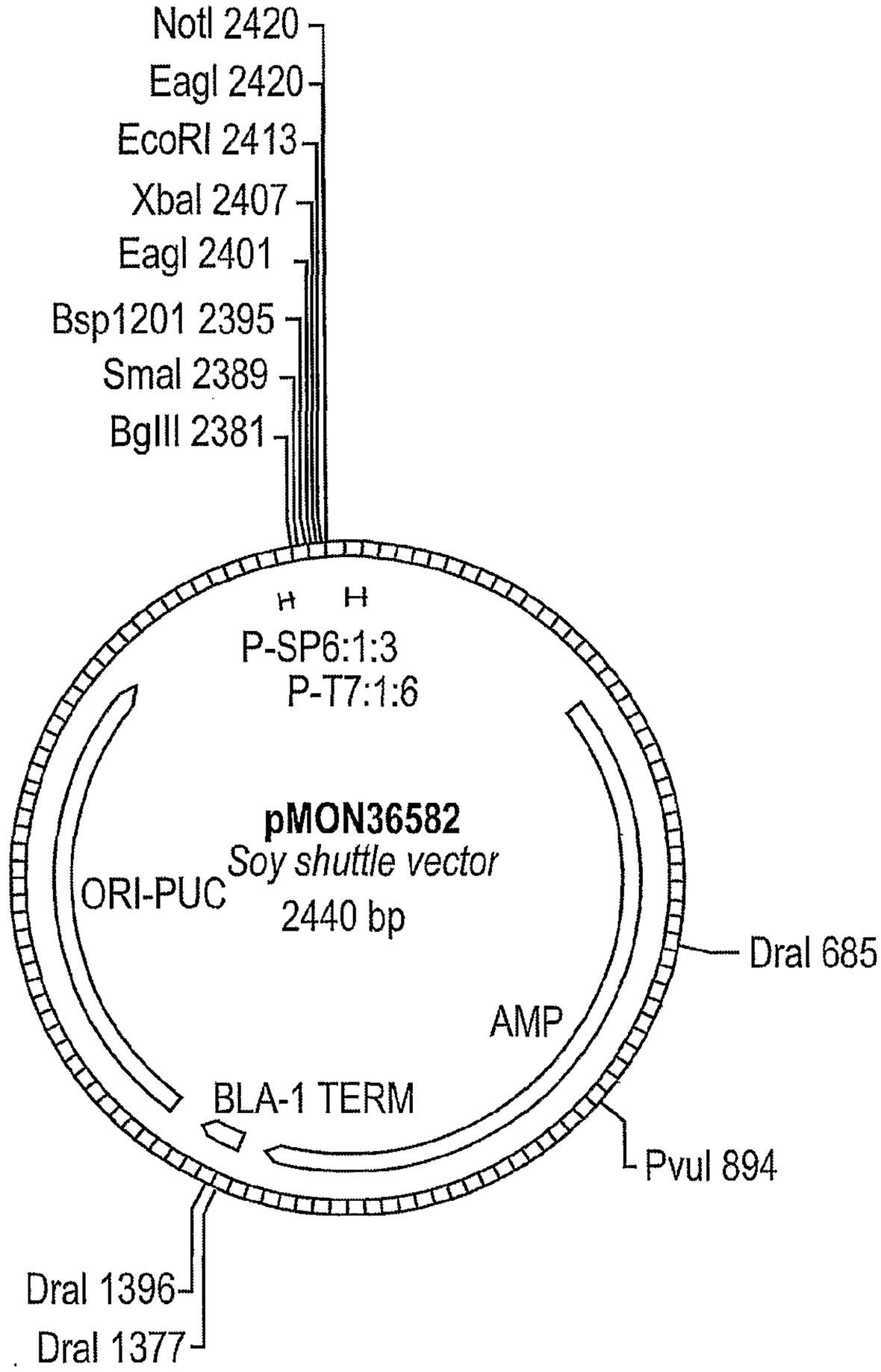

NotI 2420
EagI 2420
EcoRI 2413
XbaI 2407
EagI 2401
Bsp1201 2395
SmaI 2389
BglII 2381
H H
P-SP6:1:3
P-T7:1:6
pMON36582
Soy shuttle vector
2440 bp
ORI-PUC
AMP
BLA-1 TERM
DraI 685
PvuI 894
DraI 1396
DraI 1377

FIG. 19

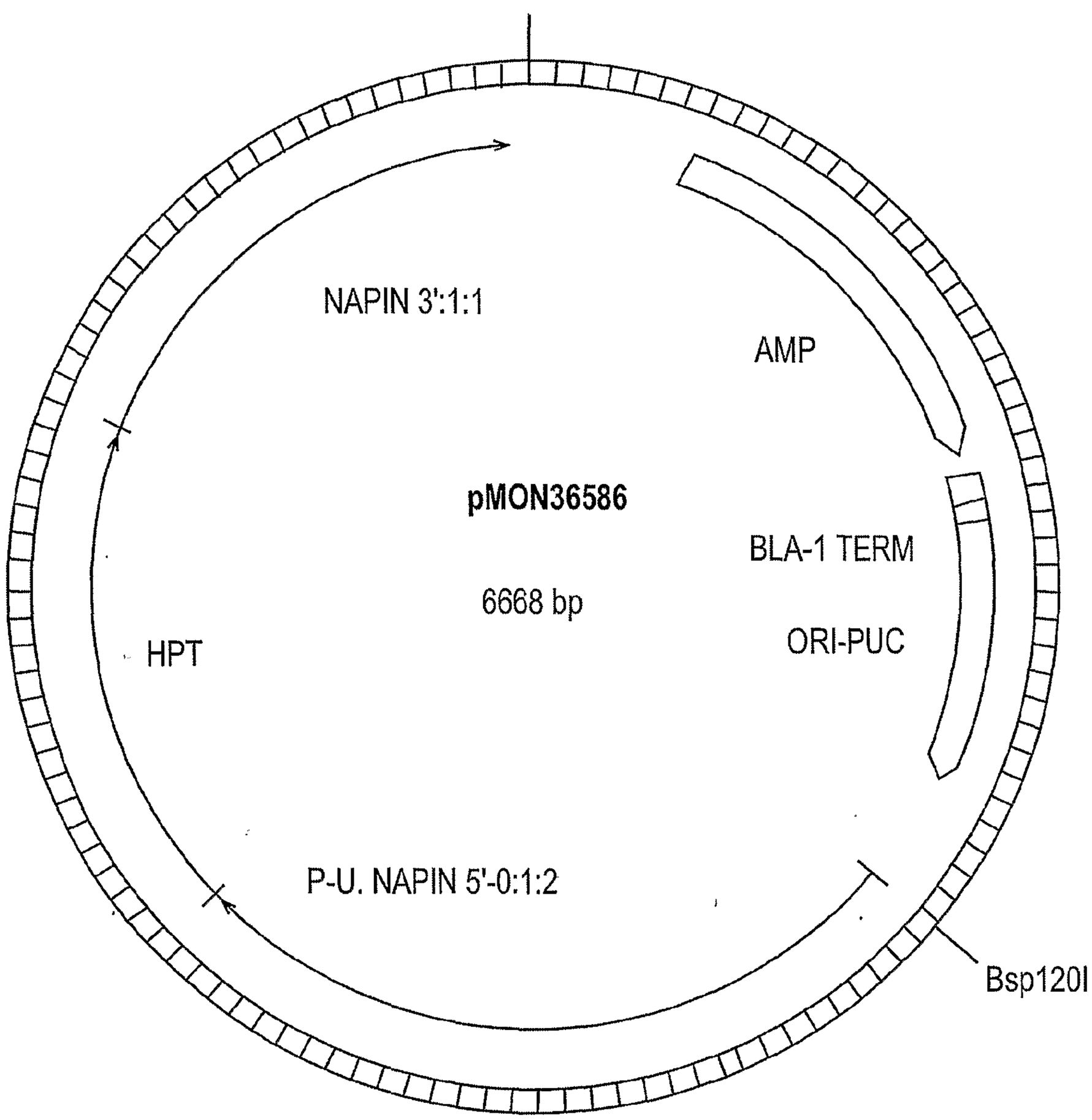

NotI
NAPIN 3':1:1
AMP
pMON36586
BLA-1 TERM
6668 bp
HPT
ORI-PUC
P-U. NAPIN 5'-0:1:2
Bsp120I

FIG. 20

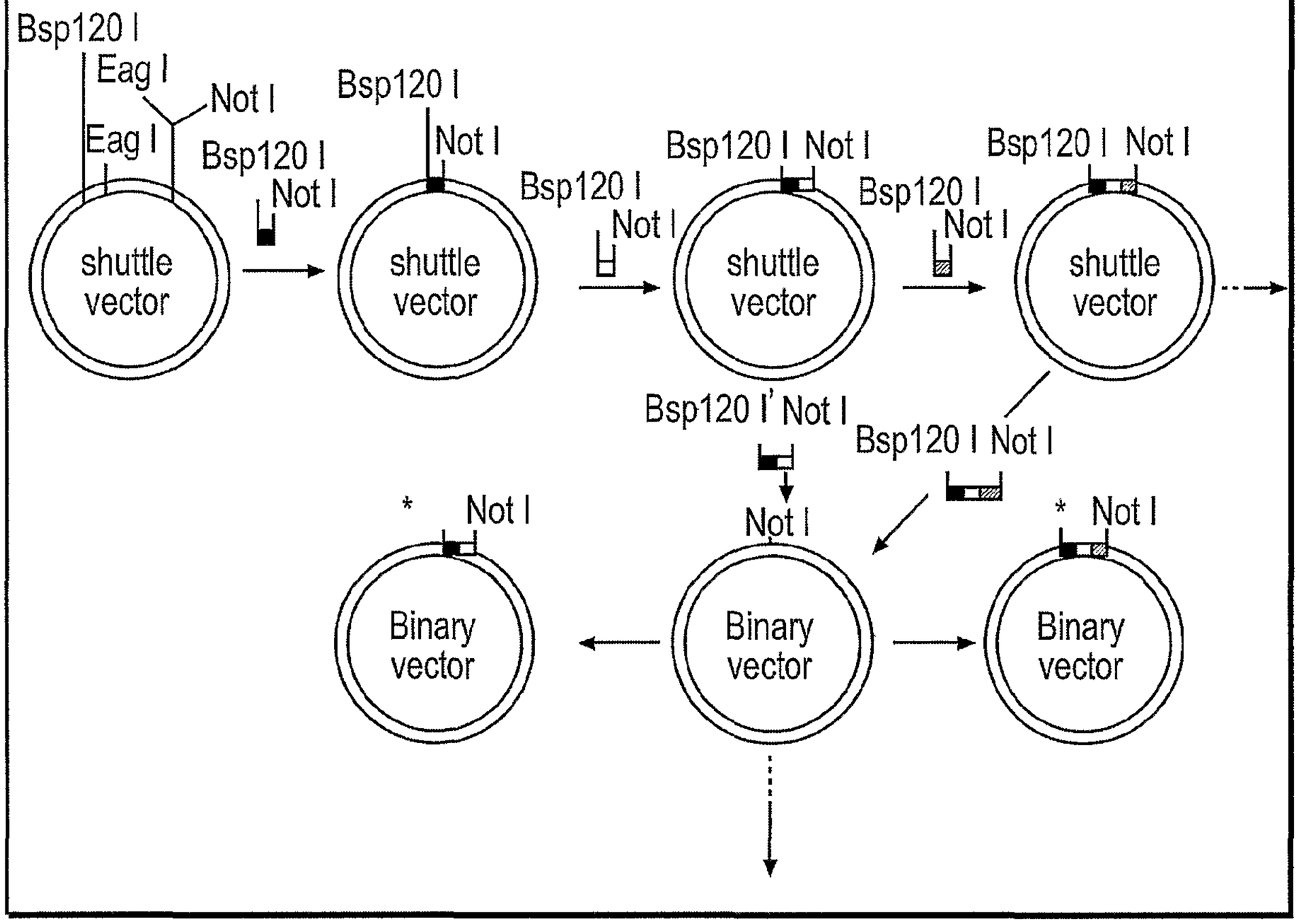


FIG. 21A

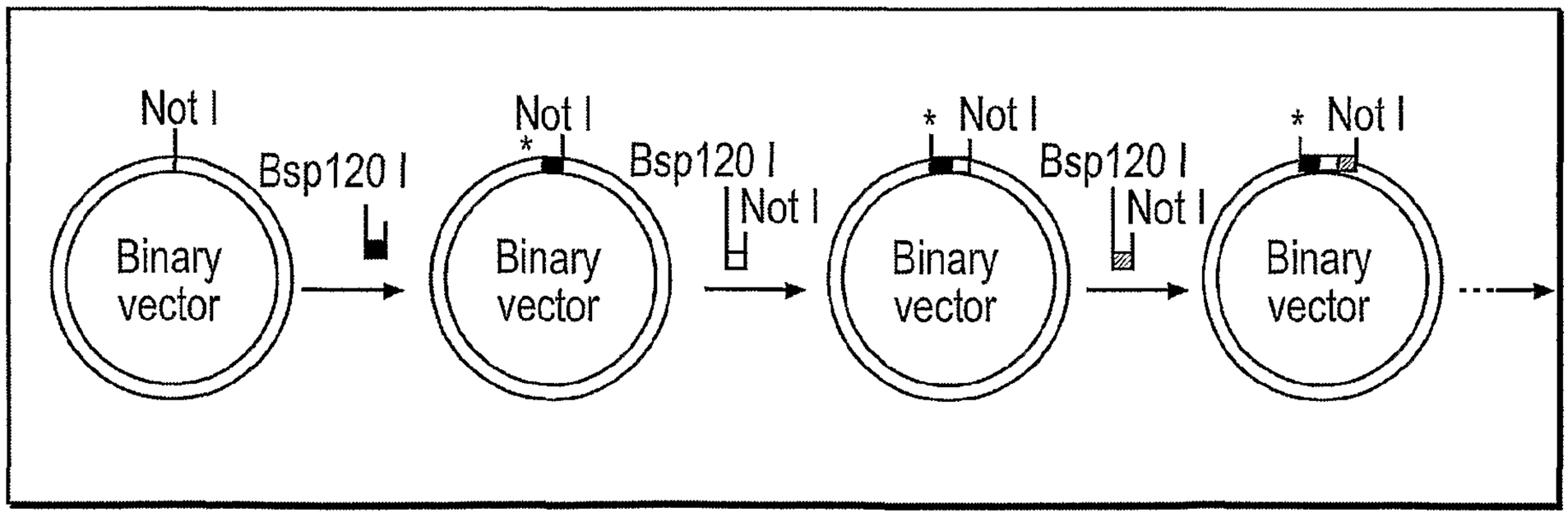


FIG. 21B

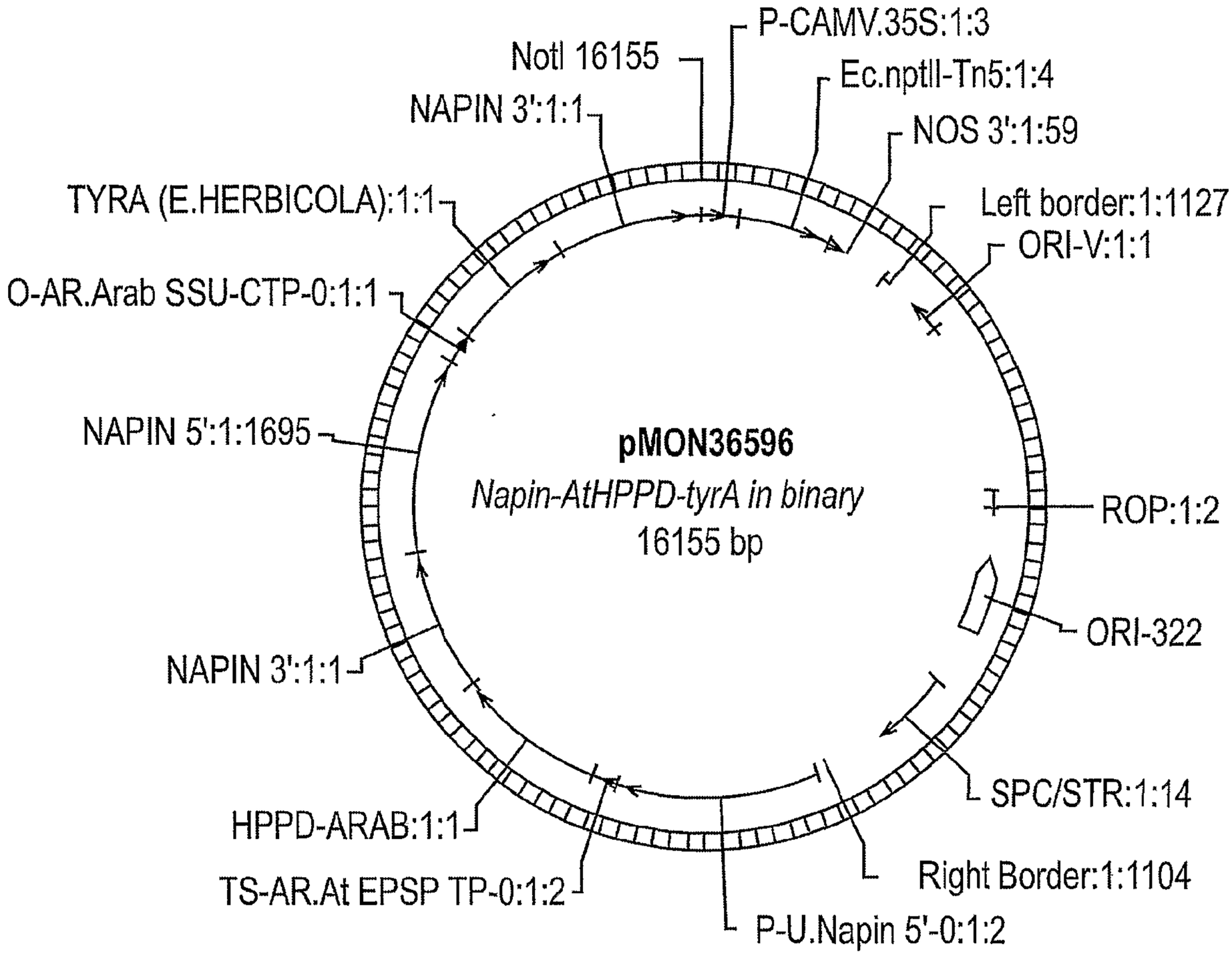

P-CAMV.35S:1:3
NotI 16155
Ec.nptII-Tn5:1:4
NAPIN 3':1:1
NOS 3':1:59
TYRA (E.HERBICOLA):1:1
Left border:1:1127
ORI-V:1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695
pMON36596
Napin-AtHPPD-tyrA in binary
16155 bp
ROP:1:2
ORI-322
NAPIN 3':1:1
SPC/STR:1:14
HPPD-ARAB:1:1
Right Border:1:1104
TS-AR.At EPSP TP-0:1:2
P-U.Napin 5'-0:1:2

FIG. 22

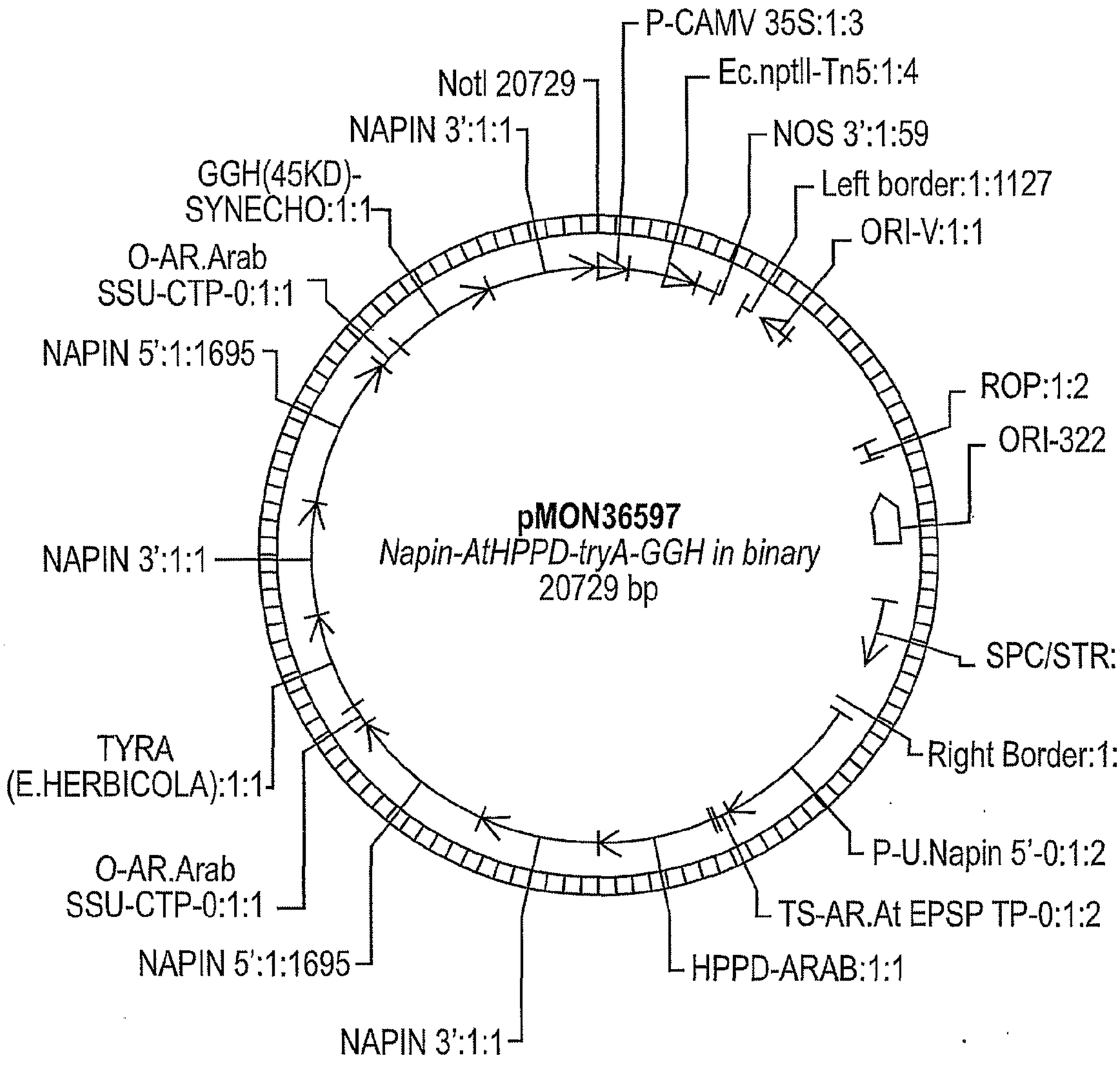

P-CAMV 35S:1:3
Ec.nptII-Tn5:1:4
NOS 3':1:59
Left border:1:1127
ORI-V:1:1
ROP:1:2
ORI-322
SPC/STR:1:14
Right Border:1:1104
P-U.Napin 5'-0:1:2
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 5':1:1695
O-AR.Arab SSU-CTP-0:1:1
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:1
NAPIN 5':1:1695
O-AR.Arab SSU-CTP-0:1:1
GGH(45KD)-SYNECHO:1:1
NAPIN 3':1:1
NotI 20729
pMON36597
Napin-AtHPPD-tryA-GGH in binary
20729 bp

FIG. 23

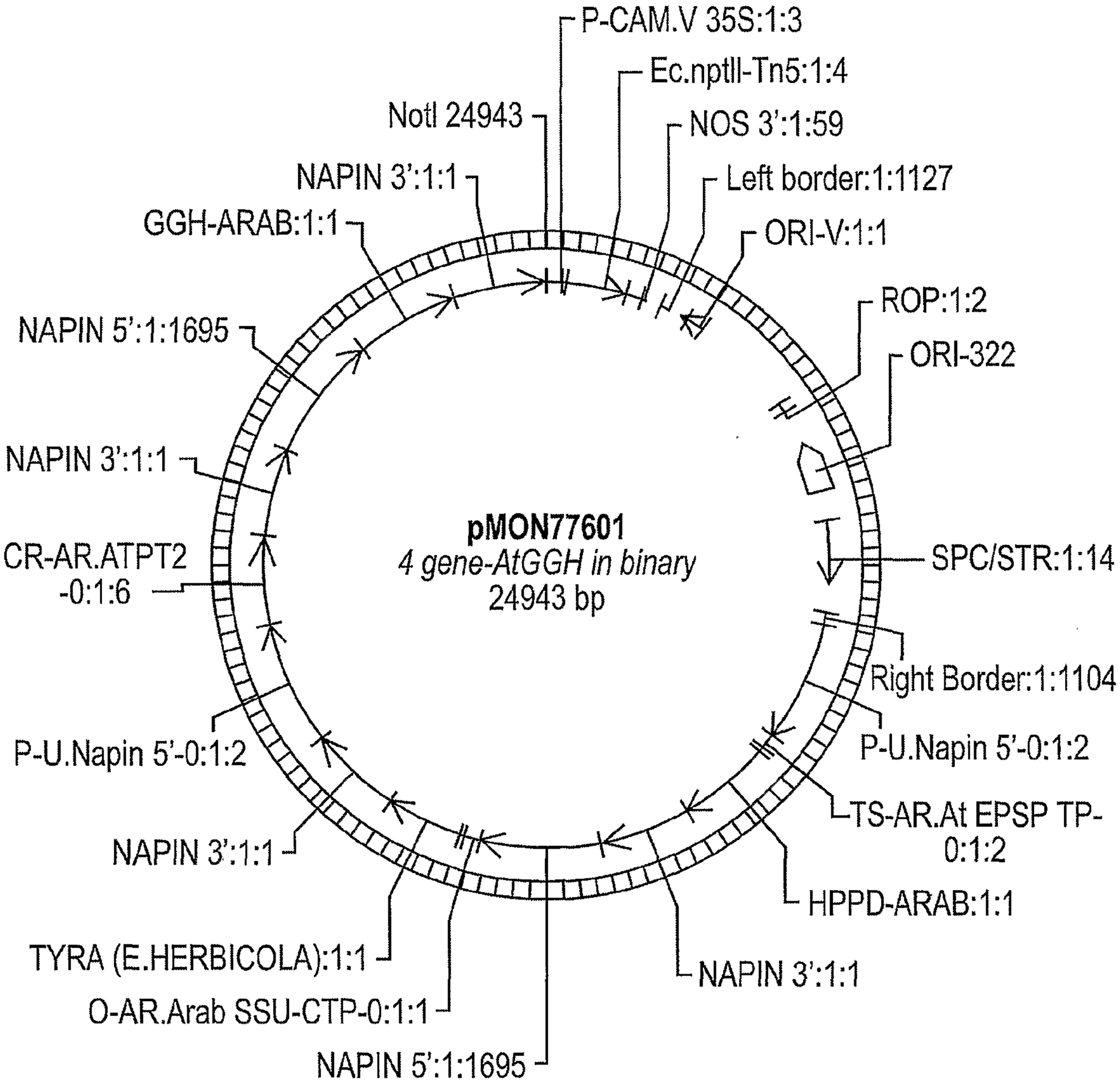

P-CAM.V 35S:1:3
Ec.nptII-Tn5:1:4
NotI 24943
NOS 3':1:59
NAPIN 3':1:1
Left border:1:1127
GGH-ARAB:1:1
ORI-V:1:1
ROP:1:2
NAPIN 5':1:1695
ORI-322
NAPIN 3':1:1
pMON77601
4 gene-AtGGH in binary
24943 bp
CR-AR.ATPT2
-0:1:6
SPC/STR:1:14
Right Border:1:1104
P-U.Napin 5'-0:1:2
P-U.Napin 5'-0:1:2
TS-AR.At EPSP TP-
0:1:2
NAPIN 3':1:1
HPPD-ARAB:1:1
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695

FIG. 24

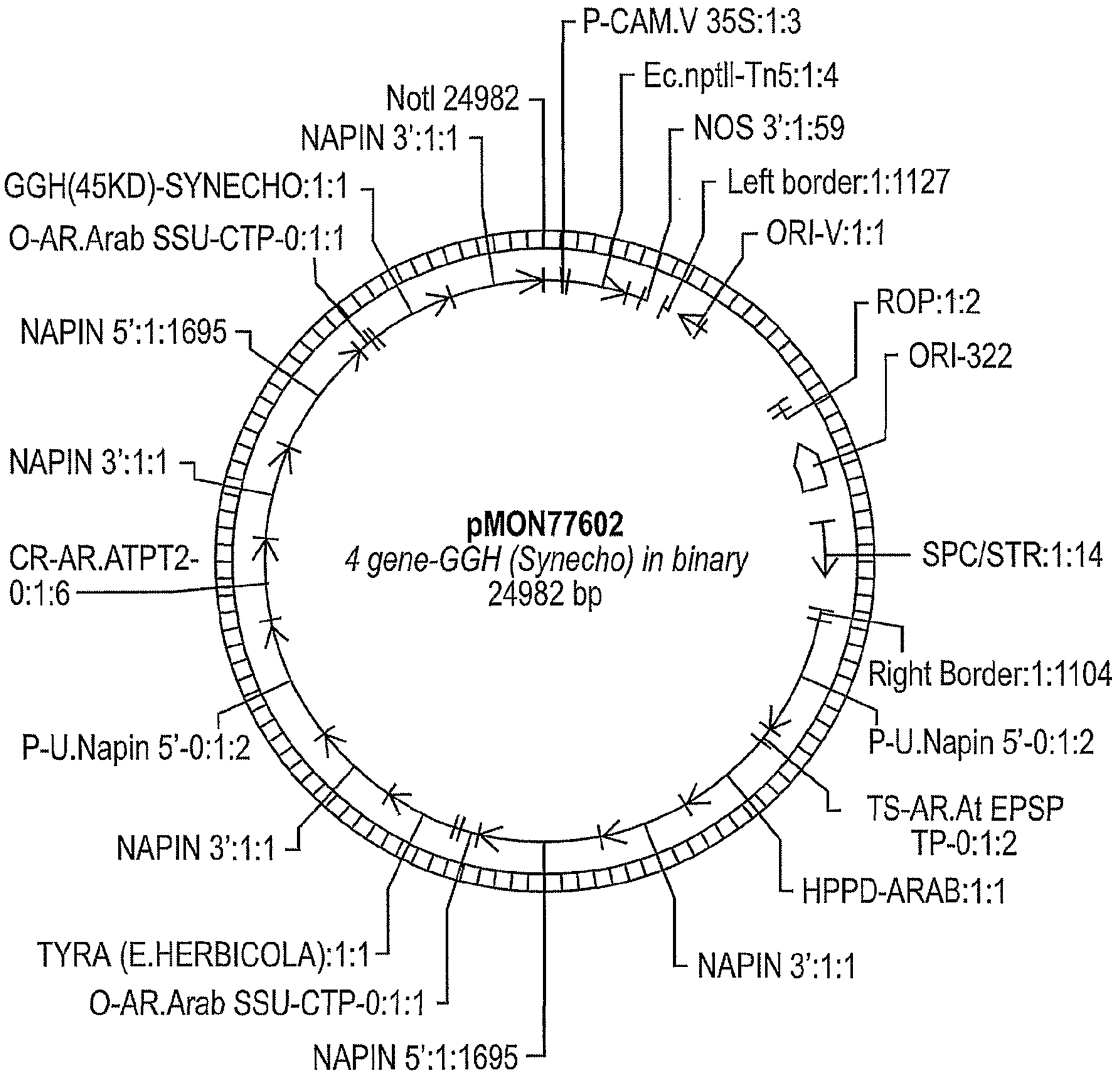

P-CAM.V 35S:1:3
Ec.nptII-Tn5:1:4
NotI 24982
NOS 3':1:59
NAPIN 3':1:1
Left border:1:1127
GGH(45KD)-SYNECHO:1:1
ORI-V:1:1
O-AR.Arab SSU-CTP-0:1:1
ROP:1:2
NAPIN 5':1:1695
ORI-322
NAPIN 3':1:1
pMON77602
4 gene-GGH (Synecho) in binary
24982 bp
CR-AR.ATPT2-
0:1:6
SPC/STR:1:14
Right Border:1:1104
P-U.Napin 5'-0:1:2
P-U.Napin 5'-0:1:2
TS-AR.At EPSP
TP-0:1:2
NAPIN 3':1:1
HPPD-ARAB:1:1
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695

FIG. 25

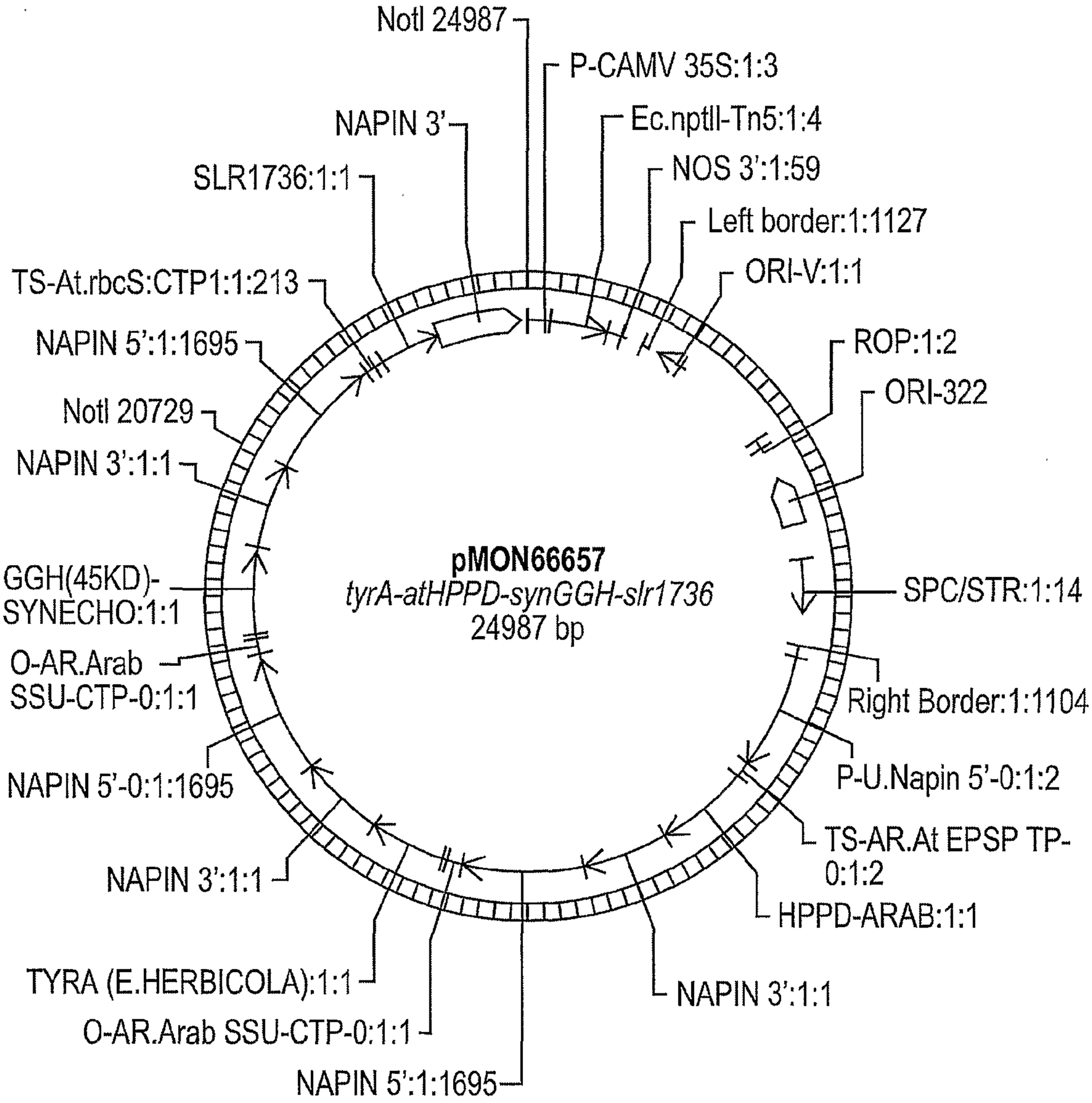

NotI 24987
P-CAMV 35S:1:3
Ec.nptII-Tn5:1:4
NOS 3':1:59
Left border:1:1127
ORI-V:1:1
ROP:1:2
ORI-322
SPC/STR:1:14
Right Border:1:1104
P-U.Napin 5'-0:1:2
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 5':1:1695
NAPIN 3'
SLR1736:1:1
TS-At.rbcS:CTP1:1:213
NAPIN 5':1:1695
NotI 20729
NAPIN 3':1:1
GGH(45KD)-SYNECHO:1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5'-0:1:1695
NAPIN 3':1:1
TYRA (E.HERBICOLA):1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695
pMON66657
tyrA-atHPPD-synGGH-slr1736
24987 bp

FIG. 26

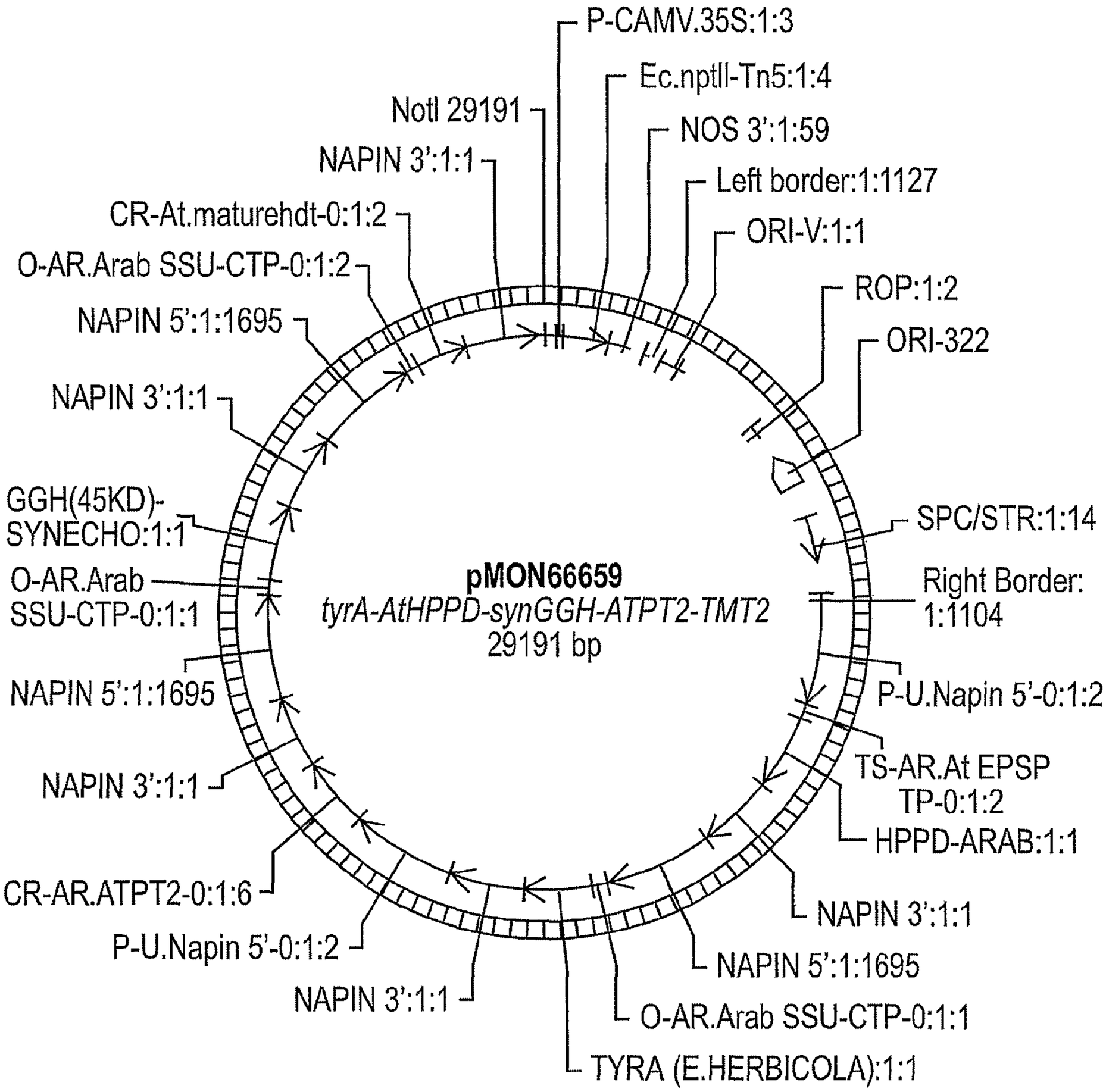

P-CAMV.35S:1:3
Ec.nptII-Tn5:1:4
NotI 29191
NOS 3':1:59
NAPIN 3':1:1
Left border:1:1127
CR-At.maturehdt-0:1:2
ORI-V:1:1
O-AR.Arab SSU-CTP-0:1:2
ROP:1:2
NAPIN 5':1:1695
ORI-322
NAPIN 3':1:1
GGH(45KD)-SYNECHO:1:1
SPC/STR:1:14
O-AR.Arab SSU-CTP-0:1:1
pMON66659
tyrA-AtHPPD-synGGH-ATPT2-TMT2
29191 bp
Right Border: 1:1104
NAPIN 5':1:1695
P-U.Napin 5'-0:1:2
NAPIN 3':1:1
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
CR-AR.ATPT2-0:1:6
NAPIN 3':1:1
P-U.Napin 5'-0:1:2
NAPIN 5':1:1695
NAPIN 3':1:1
O-AR.Arab SSU-CTP-0:1:1
TYRA (E.HERBICOLA):1:1

FIG. 27

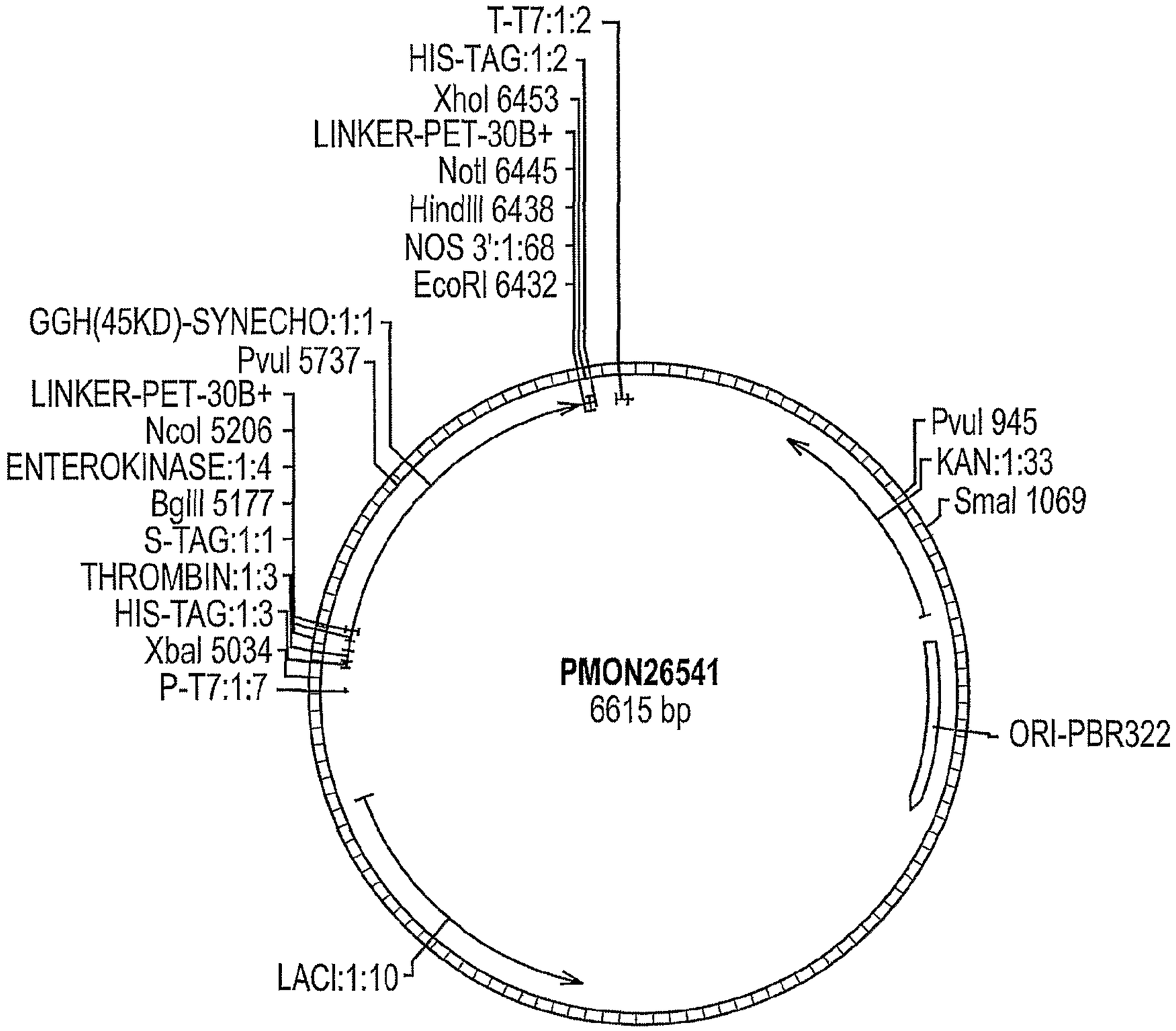

T-T7:1:2
HIS-TAG:1:2
XhoI 6453
LINKER-PET-30B+
NotI 6445
HindIII 6438
NOS 3':1:68
EcoRI 6432
GGH(45KD)-SYNECHO:1:1
PvuI 5737
LINKER-PET-30B+
NcoI 5206
ENTEROKINASE:1:4
BglII 5177
S-TAG:1:1
THROMBIN:1:3
HIS-TAG:1:3
XbaI 5034
P-T7:1:7
PMON26541
6615 bp
PvuI 945
KAN:1:33
SmaI 1069
ORI-PBR322
LACI:1:10

FIG. 28

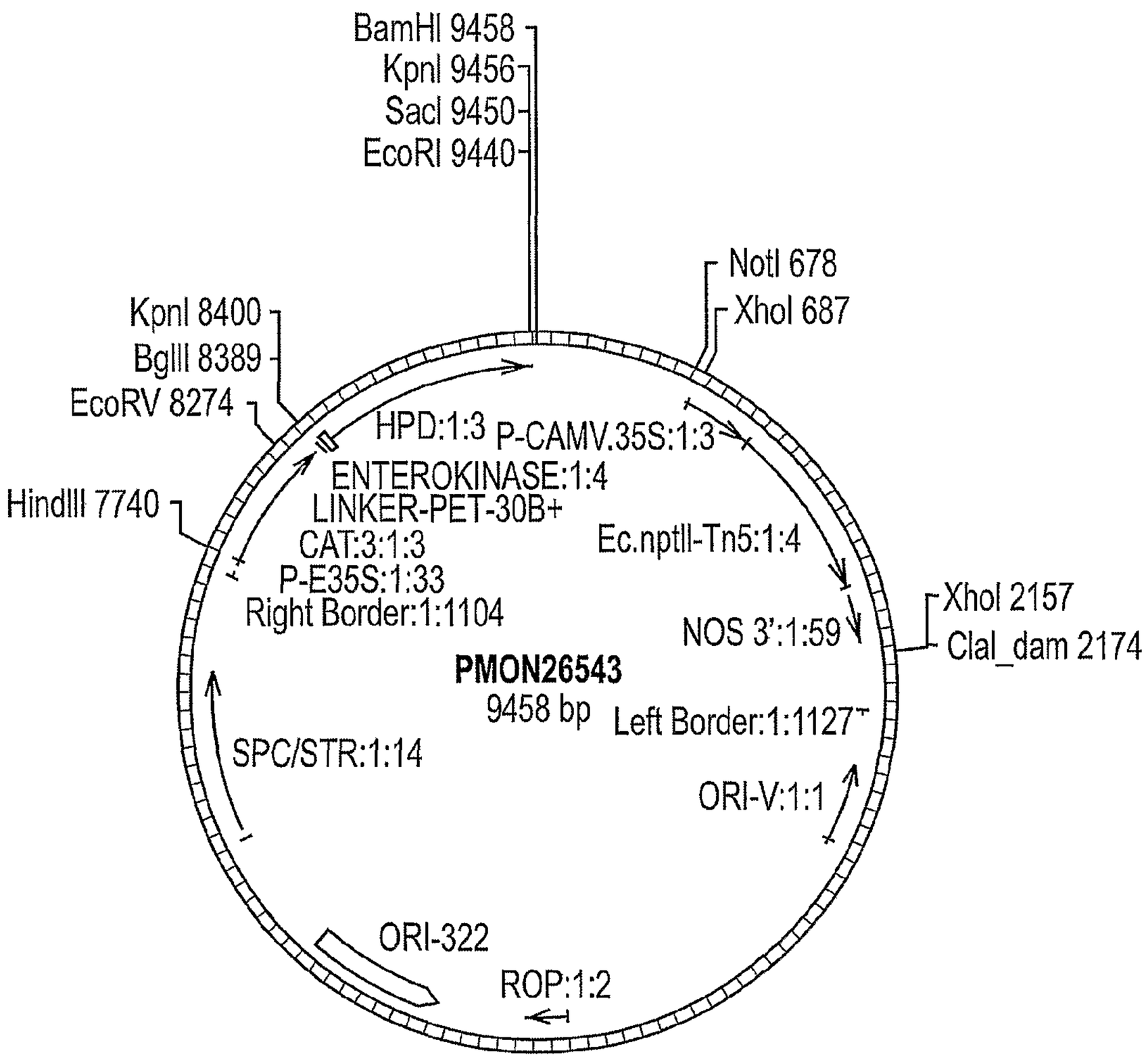

BamHI 9458
KpnI 9456
SacI 9450
EcoRI 9440
NotI 678
XhoI 687
KpnI 8400
BglII 8389
EcoRV 8274
HPD:1:3
P-CAMV.35S:1:3
ENTEROKINASE:1:4
LINKER-PET-30B+
HindIII 7740
CAT:3:1:3
Ec.nptII-Tn5:1:4
P-E35S:1:33
Right Border:1:1104
XhoI 2157
NOS 3':1:59
ClaI_dam 2174
PMON26543
9458 bp
Left Border:1:1127
SPC/STR:1:14
ORI-V:1:1
ORI-322
ROP:1:2

FIG. 29

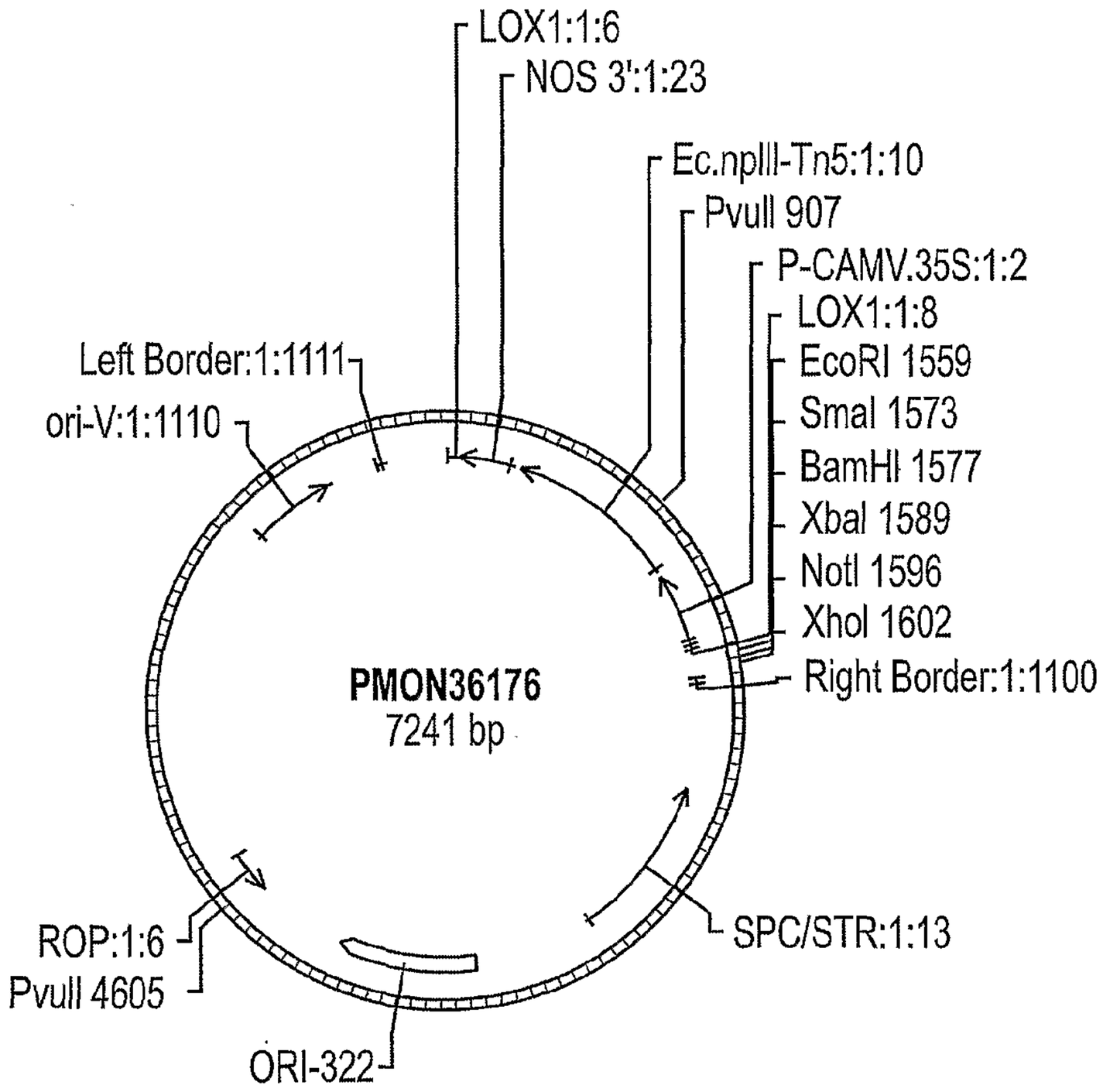

LOX1:1:6
NOS 3':1:23
Ec.nplll-Tn5:1:10
Pvull 907
P-CAMV.35S:1:2
LOX1:1:8
EcoRI 1559
SmaI 1573
BamHI 1577
XbaI 1589
NotI 1596
XhoI 1602
Right Border:1:1100
Left Border:1:1111
ori-V:1:1110
PMON36176
7241 bp
ROP:1:6
Pvull 4605
ORI-322
SPC/STR:1:13

FIG. 30

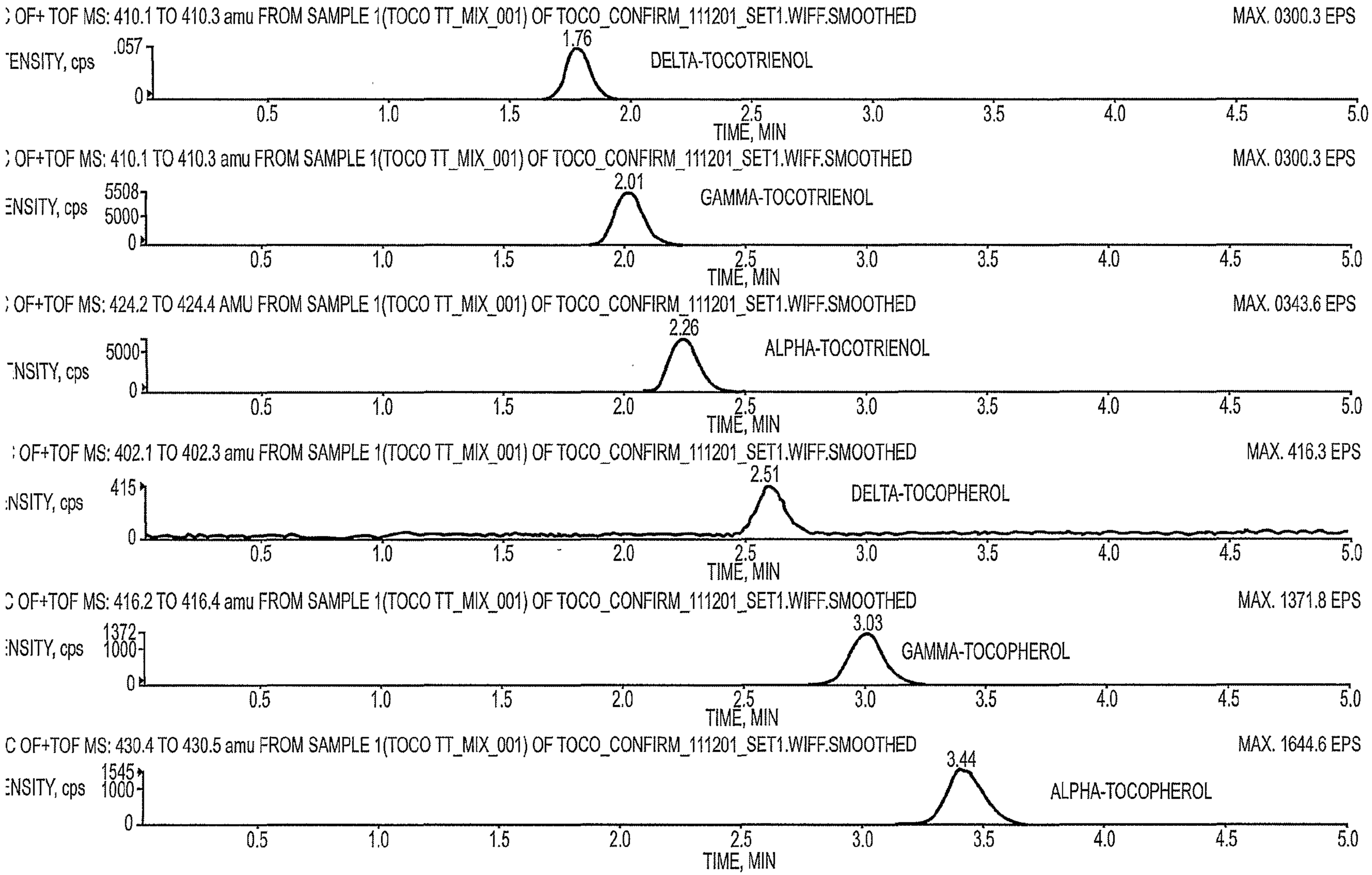


FIG. 31

WO 02/089561
32/77
CA 02443865 2003-10-07
PCT/US02/13898

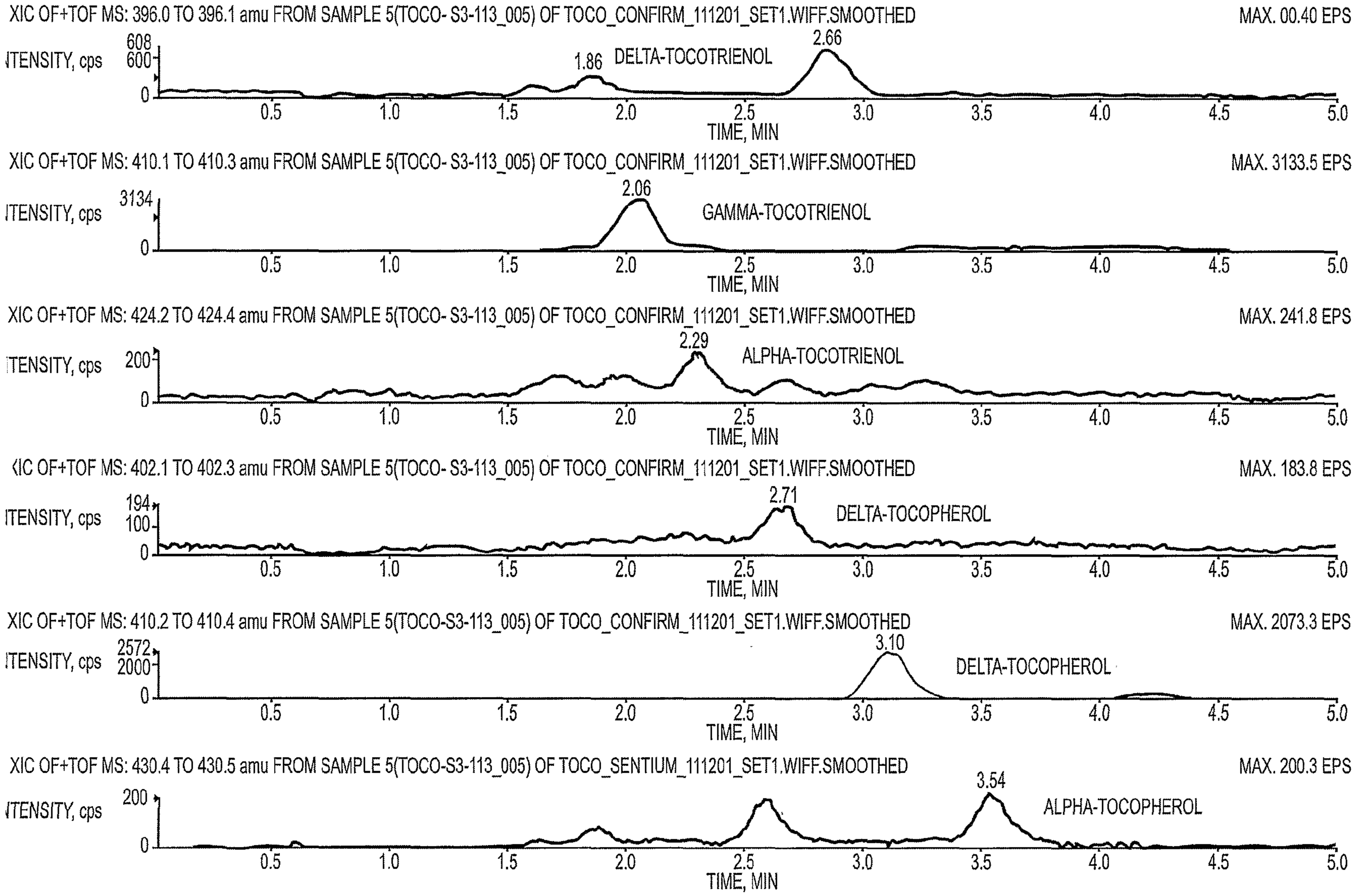

XIC OF+TOF MS: 396.0 TO 396.1 amu FROM SAMPLE 5(TOCO- S3-113_005) OF TOCO_CONFIRM_111201_SET1.WIFF.SMOOTHED
MAX. 00.40 EPS
ITENSITY, cps
608
600
0
1.86
DELTA-TOCOTRIENOL
2.66
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN
XIC OF+TOF MS: 410.1 TO 410.3 amu FROM SAMPLE 5(TOCO- S3-113_005) OF TOCO_CONFIRM_111201_SET1.WIFF.SMOOTHED
MAX. 3133.5 EPS
ITENSITY, cps
3134
0
2.06
GAMMA-TOCOTRIENOL
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN
XIC OF+TOF MS: 424.2 TO 424.4 amu FROM SAMPLE 5(TOCO- S3-113_005) OF TOCO_CONFIRM_111201_SET1.WIFF.SMOOTHED
MAX. 241.8 EPS
ITENSITY, cps
200
0
2.29
ALPHA-TOCOTRIENOL
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN
XIC OF+TOF MS: 402.1 TO 402.3 amu FROM SAMPLE 5(TOCO- S3-113_005) OF TOCO_CONFIRM_111201_SET1.WIFF.SMOOTHED
MAX. 183.8 EPS
ITENSITY, cps
194
100
0
2.71
DELTA-TOCOPHEROL
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN
XIC OF+TOF MS: 410.2 TO 410.4 amu FROM SAMPLE 5(TOCO-S3-113_005) OF TOCO_CONFIRM_111201_SET1.WIFF.SMOOTHED
MAX. 2073.3 EPS
ITENSITY, cps
2572
2000
0
3.10
DELTA-TOCOPHEROL
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN
XIC OF+TOF MS: 430.4 TO 430.5 amu FROM SAMPLE 5(TOCO-S3-113_005) OF TOCO_SENTIUM_111201_SET1.WIFF.SMOOTHED
MAX. 200.3 EPS
ITENSITY, cps
200
0
3.54
ALPHA-TOCOPHEROL
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
TIME, MIN

FIG. 32

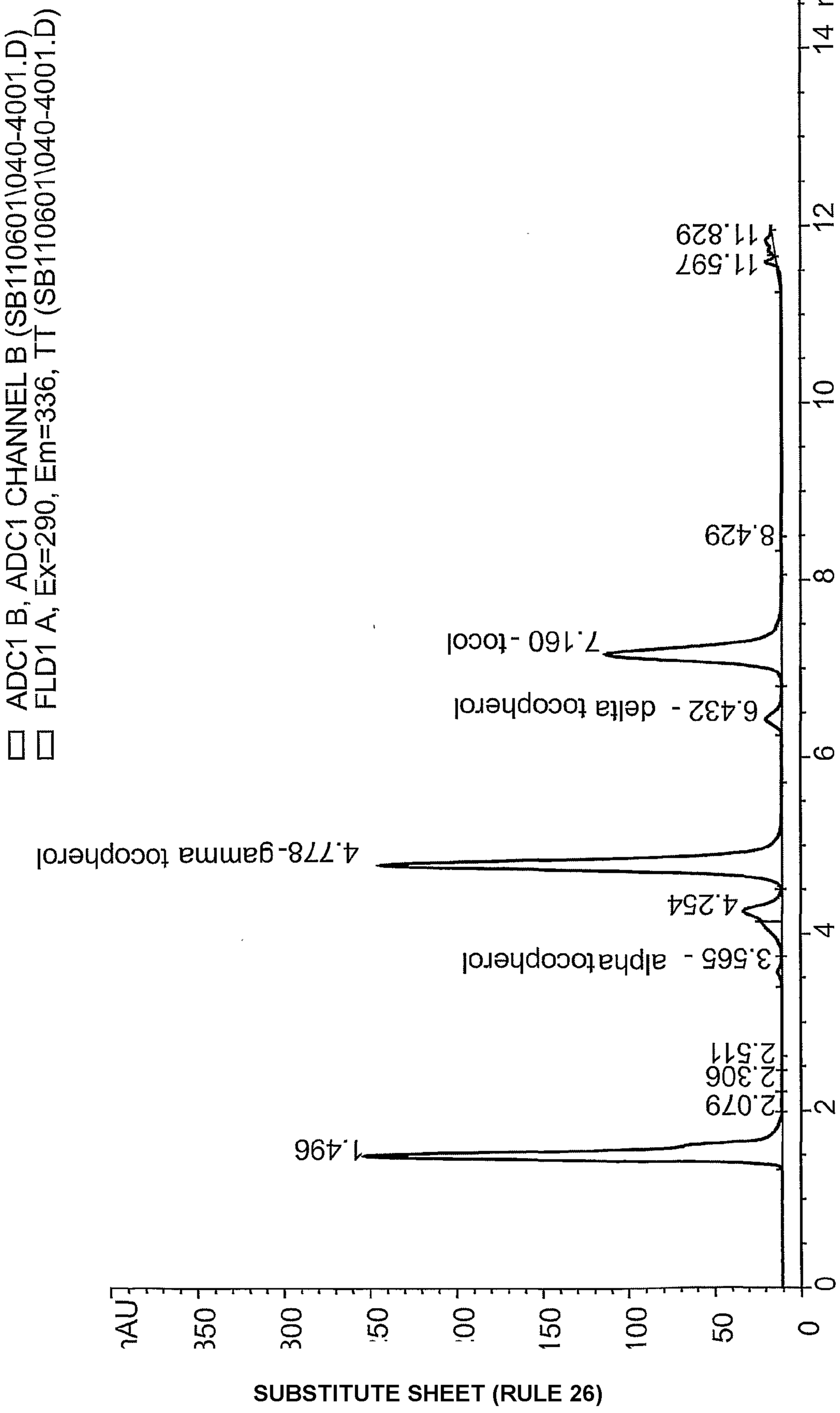

ADC1 B, ADC1 CHANNEL B (SB110601\040-4001.D)
FLD1 A, Ex=290, Em=336, TT (SB110601\040-4001.D)
14 min
12
10
8
6
4
2
0
11.829
11.597
8.429
7.160 - tocol
6.432 - delta tocopherol
4.778 - gamma tocopherol
4.254
3.565 - alpha tocopherol
2.511
2.306
2.079
1.496
nAU
350
300
150
100
50
0

FIG. 33

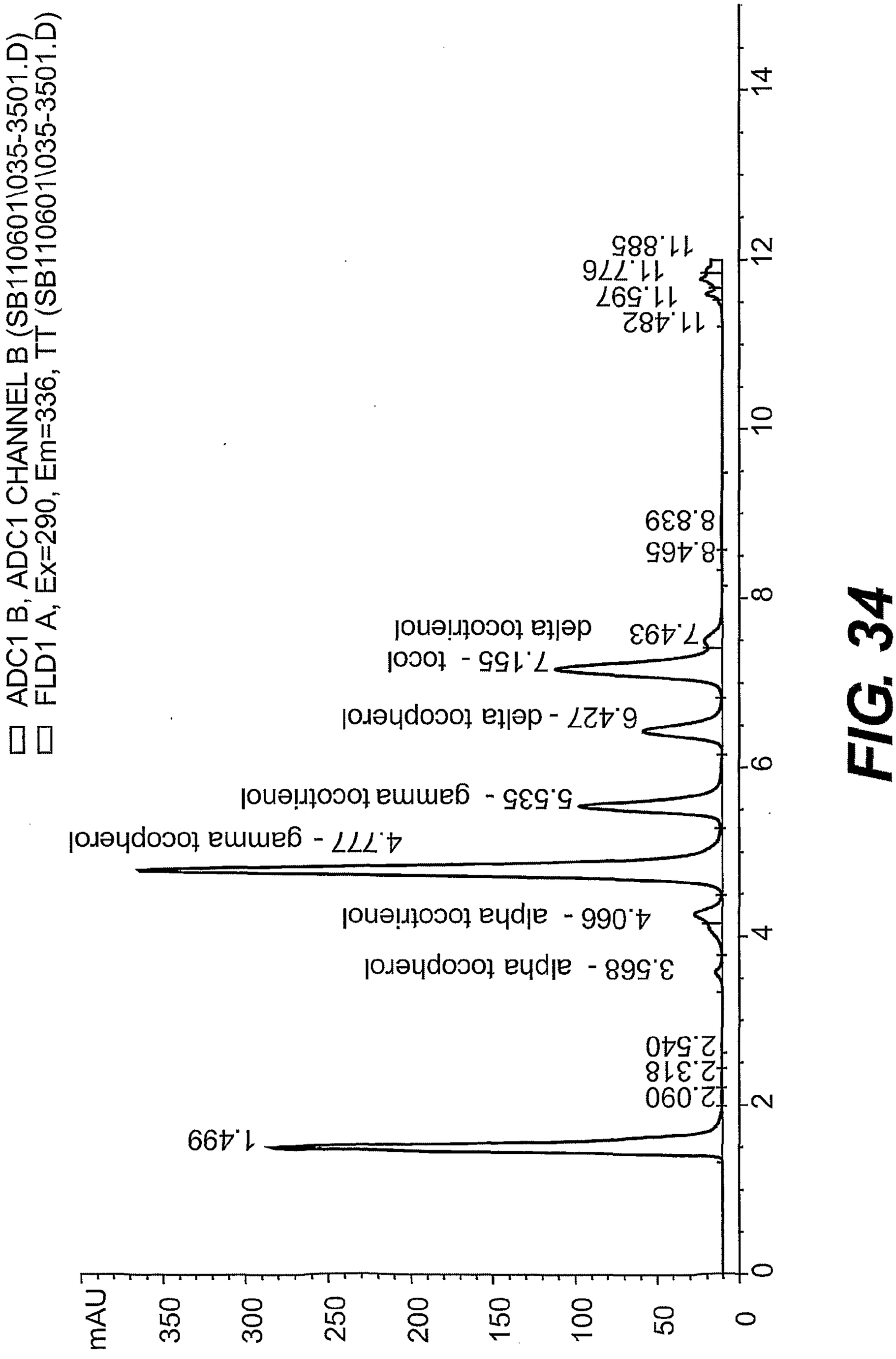

ADC1 B, ADC1 CHANNEL B (SB110601\035-3501.D)
FLD1 A, Ex=290, Em=336, TT (SB110601\035-3501.D)
mAU
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
14
1.499
2.090
2.318
2.540
3.568 - alpha tocopherol
4.066 - alpha tocotrienol
4.777 - gamma tocopherol
5.535 - gamma tocotrienol
6.427 - delta tocopherol
7.155 - tocol
7.493 delta tocotrienol
8.465
8.839
11.482
11.597
11.776
11.885

FIG. 34

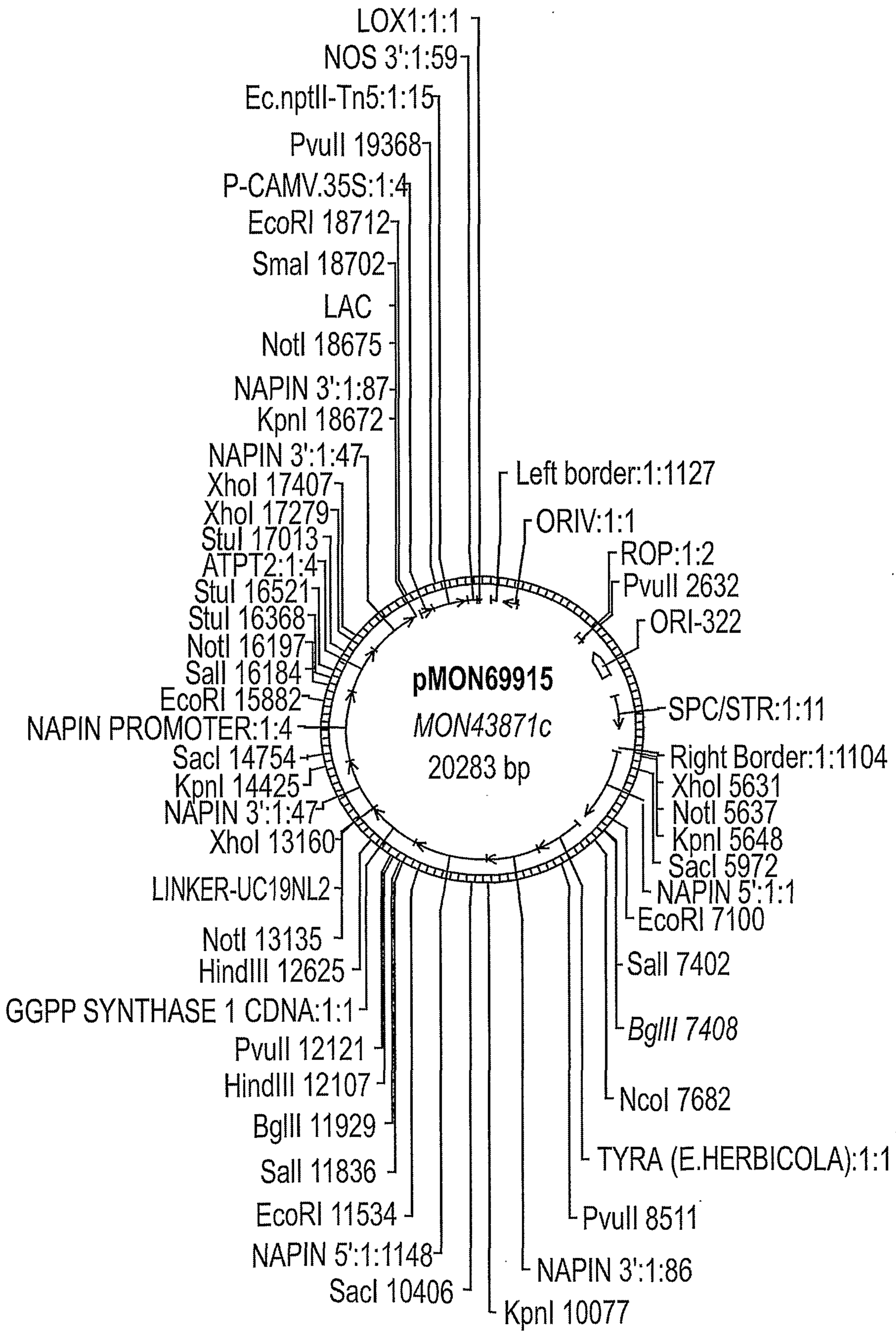

LOX1:1:1
NOS 3':1:59
Ec.nptII-Tn5:1:15
PvuII 19368
P-CAMV.35S:1:4
EcoRI 18712
SmaI 18702
LAC
NotI 18675
NAPIN 3':1:87
KpnI 18672
NAPIN 3':1:47
XhoI 17407
XhoI 17279
StuI 17013
ATPT2:1:4
StuI 16521
StuI 16368
NotI 16197
SalI 16184
EcoRI 15882
NAPIN PROMOTER:1:4
SacI 14754
KpnI 14425
NAPIN 3':1:47
XhoI 13160
LINKER-UC19NL2
NotI 13135
HindIII 12625
GGPP SYNTHASE 1 CDNA:1:1
PvuII 12121
HindIII 12107
BglII 11929
SalI 11836
EcoRI 11534
NAPIN 5':1:1148
SacI 10406
pMON69915
MON43871c
20283 bp
Left border:1:1127
ORIV:1:1
ROP:1:2
PvuII 2632
ORI-322
SPC/STR:1:11
Right Border:1:1104
XhoI 5631
NotI 5637
KpnI 5648
SacI 5972
NAPIN 5':1:1
EcoRI 7100
SalI 7402
BglII 7408
NcoI 7682
TYRA (E.HERBICOLA):1:1
PvuII 8511
NAPIN 3':1:86
KpnI 10077

FIG. 35

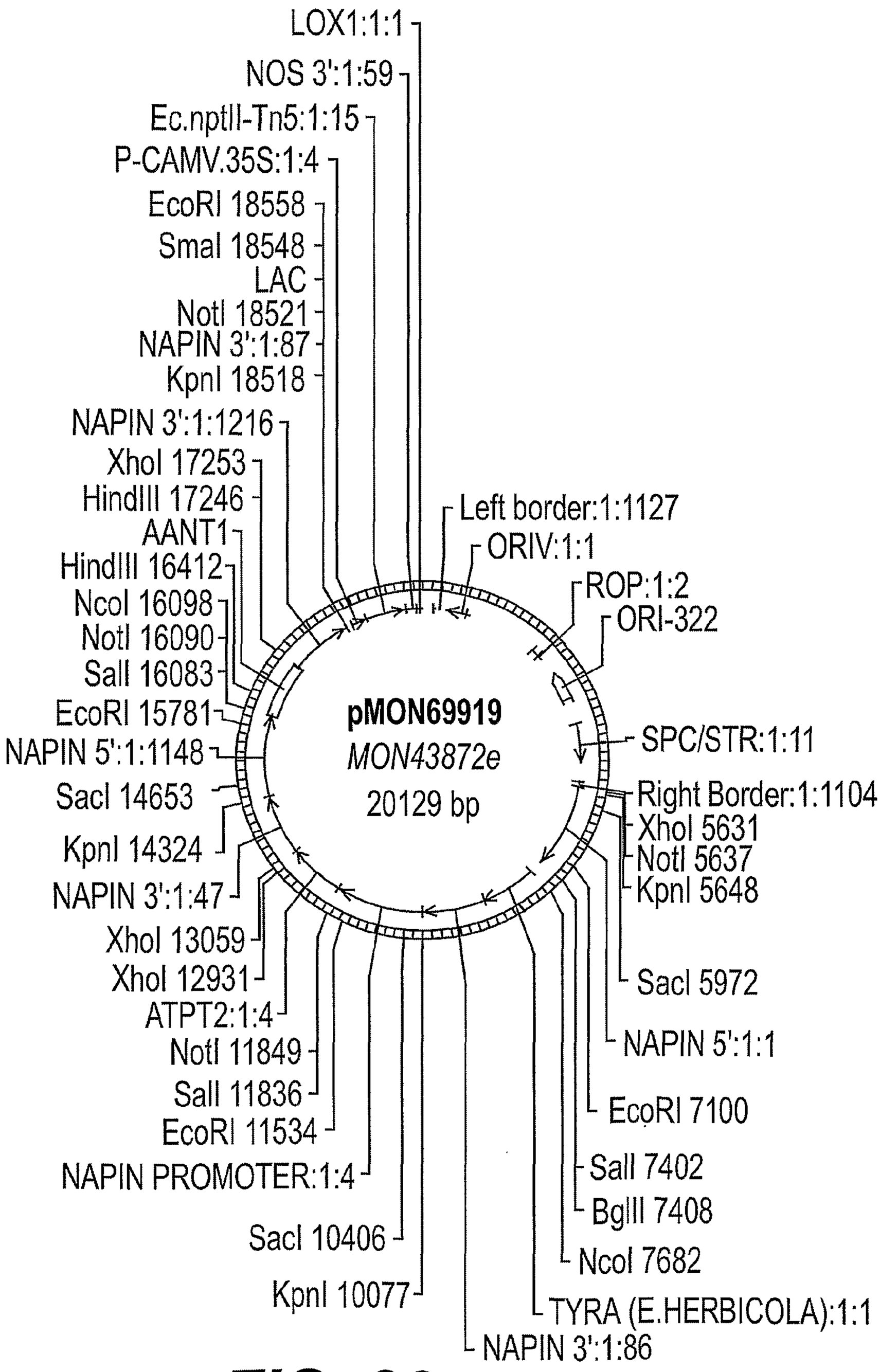

LOX1:1:1
NOS 3':1:59
Ec.nptII-Tn5:1:15
P-CAMV.35S:1:4
EcoRI 18558
SmaI 18548
LAC
NotI 18521
NAPIN 3':1:87
KpnI 18518
NAPIN 3':1:1216
XhoI 17253
HindIII 17246
AANT1
HindIII 16412
NcoI 16098
NotI 16090
SalI 16083
EcoRI 15781
NAPIN 5':1:1148
SacI 14653
KpnI 14324
NAPIN 3':1:47
XhoI 13059
XhoI 12931
ATPT2:1:4
NotI 11849
SalI 11836
EcoRI 11534
NAPIN PROMOTER:1:4
SacI 10406
KpnI 10077
Left border:1:1127
ORIV:1:1
ROP:1:2
ORI-322
SPC/STR:1:11
Right Border:1:1104
XhoI 5631
NotI 5637
KpnI 5648
SacI 5972
NAPIN 5':1:1
EcoRI 7100
SalI 7402
BglII 7408
NcoI 7682
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:86
pMON69919
MON43872e
20129 bp

FIG. 36

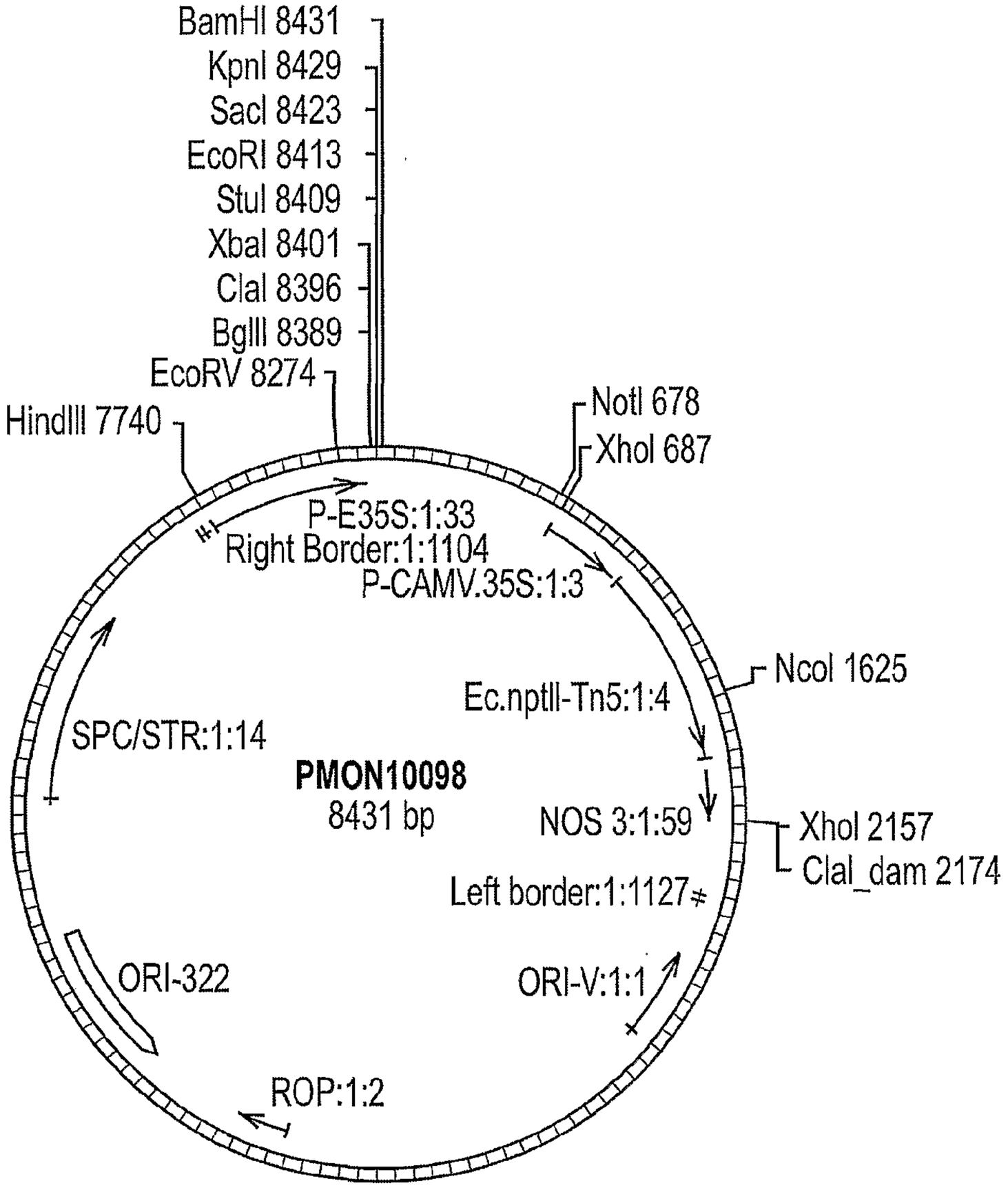

BamHI 8431
KpnI 8429
SacI 8423
EcoRI 8413
StuI 8409
XbaI 8401
ClaI 8396
BglII 8389
EcoRV 8274
HindIII 7740
NotI 678
XhoI 687
P-E35S:1:33
Right Border:1:1104
P-CAMV.35S:1:3
NcoI 1625
Ec.nptII-Tn5:1:4
SPC/STR:1:14
PMON10098
8431 bp
NOS 3:1:59
XhoI 2157
ClaI_dam 2174
Left border:1:1127
ORI-322
ORI-V:1:1
ROP:1:2

FIG. 37

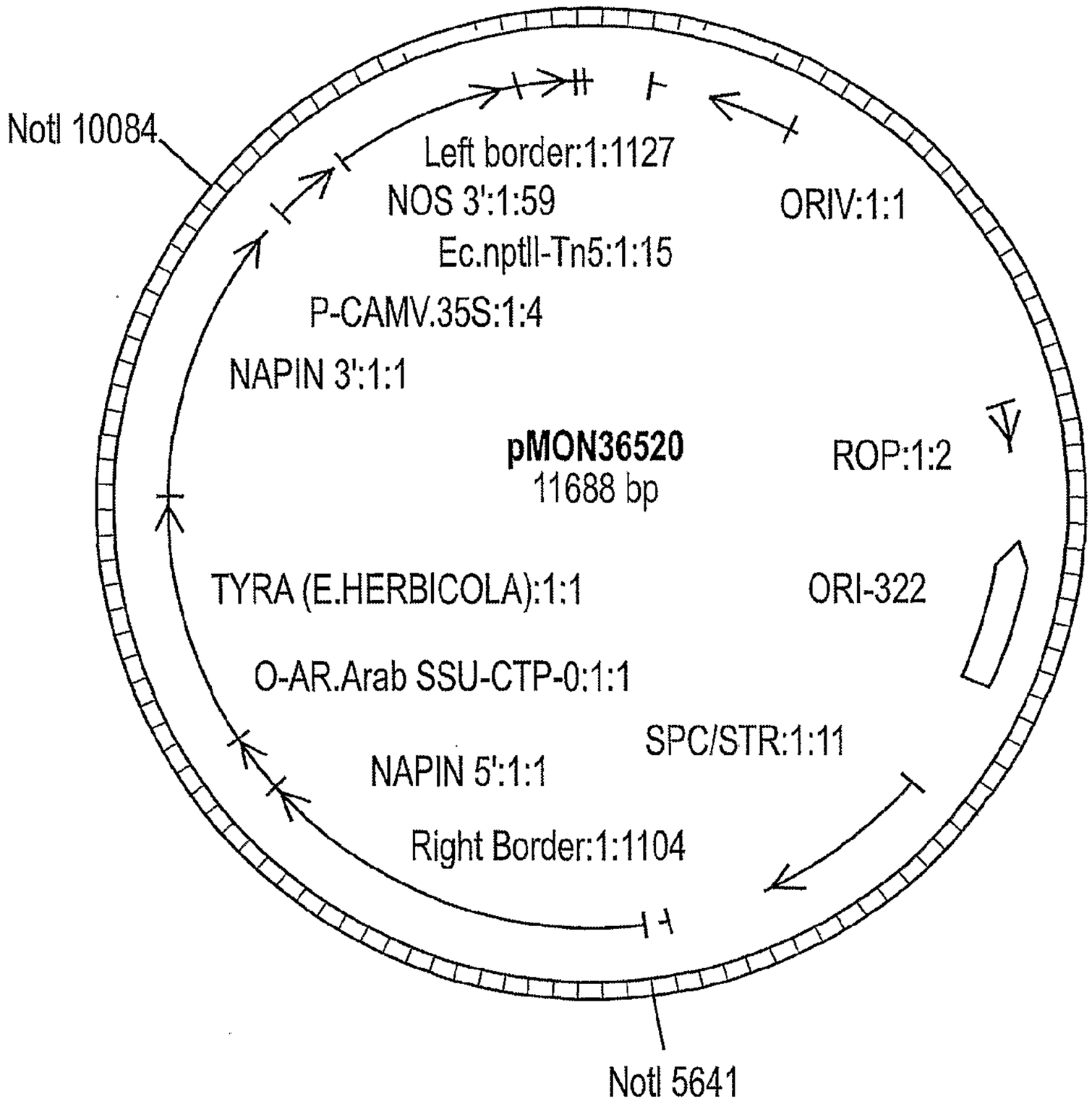

NotI 10084
Left border:1:1127
NOS 3':1:59
Ec.nptII-Tn5:1:15
P-CAMV.35S:1:4
NAPIN 3':1:1
ORIV:1:1
pMON36520
11688 bp
ROP:1:2
TYRA (E.HERBICOLA):1:1
ORI-322
O-AR.Arab SSU-CTP-0:1:1
SPC/STR:1:11
NAPIN 5':1:1
Right Border:1:1104
NotI 5641

FIG. 38

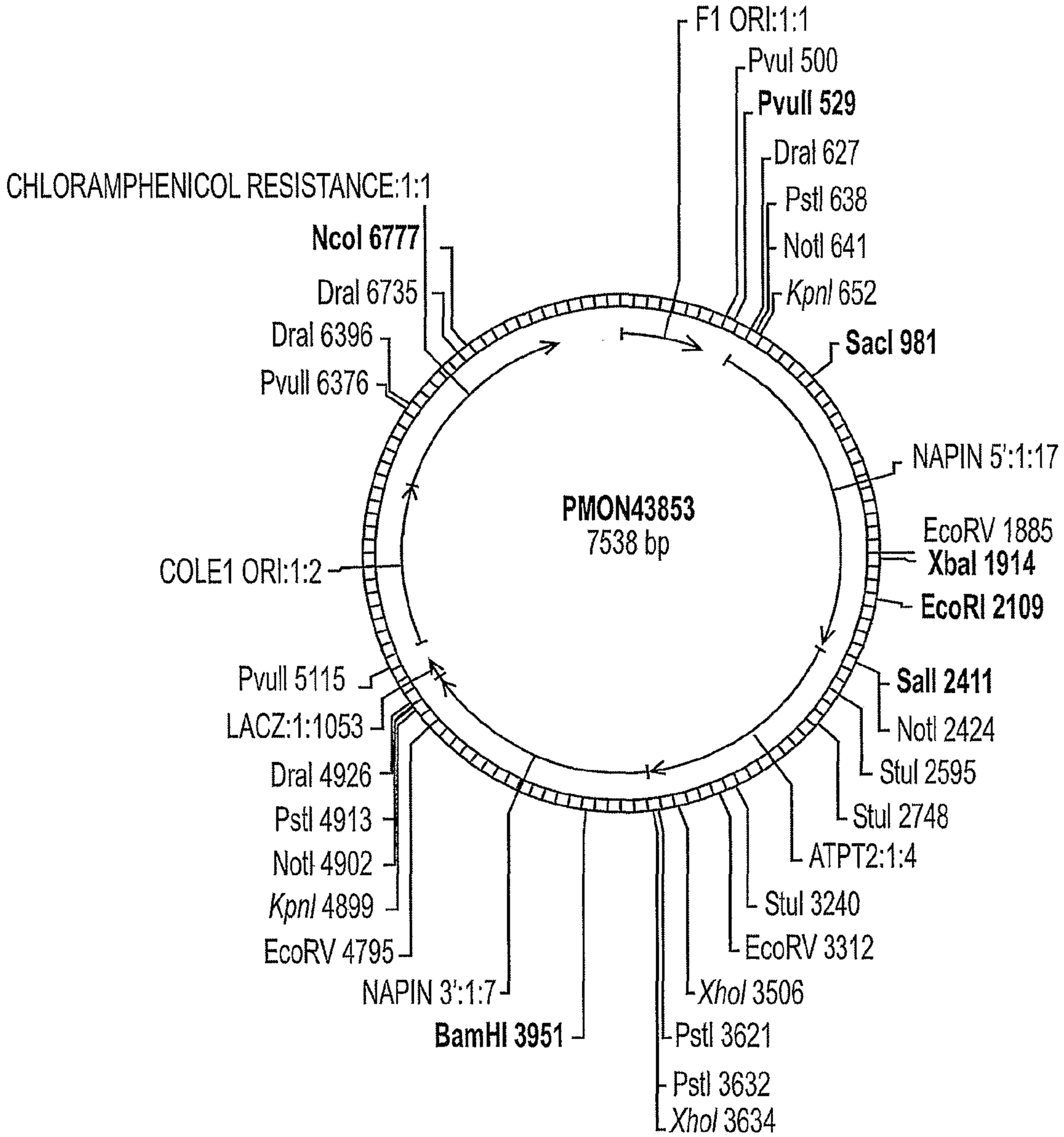

F1 ORI:1:1
Pvul 500
PvuII 529
DraI 627
PstI 638
NotI 641
KpnI 652
SacI 981
NAPIN 5':1:17
EcoRV 1885
XbaI 1914
EcoRI 2109
SalI 2411
NotI 2424
StuI 2595
StuI 2748
ATPT2:1:4
StuI 3240
EcoRV 3312
XhoI 3506
PstI 3621
PstI 3632
XhoI 3634
PMON43853
7538 bp
CHLORAMPHENICOL RESISTANCE:1:1
NcoI 6777
DraI 6735
DraI 6396
PvuII 6376
COLE1 ORI:1:2
PvuII 5115
LACZ:1:1053
DraI 4926
PstI 4913
NotI 4902
KpnI 4899
EcoRV 4795
NAPIN 3':1:7
BamHI 3951

FIG. 39

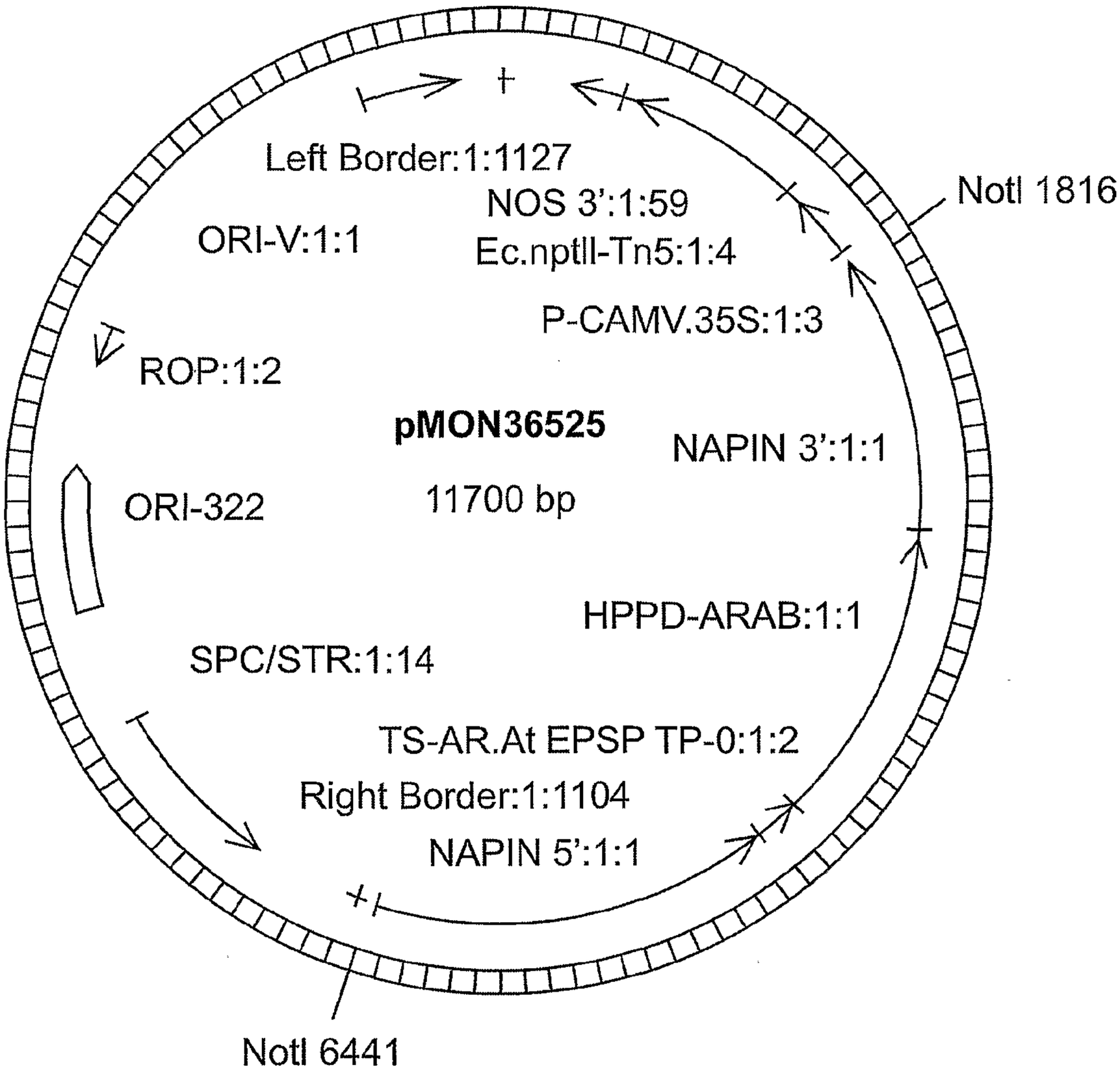

Left Border:1:1127
NOS 3':1:59
Ec.nptII-Tn5:1:4
P-CAMV.35S:1:3
NotI 1816
ORI-V:1:1
ROP:1:2
pMON36525
11700 bp
NAPIN 3':1:1
ORI-322
HPPD-ARAB:1:1
SPC/STR:1:14
TS-AR.At EPSP TP-0:1:2
Right Border:1:1104
NAPIN 5':1:1
NotI 6441

FIG. 40

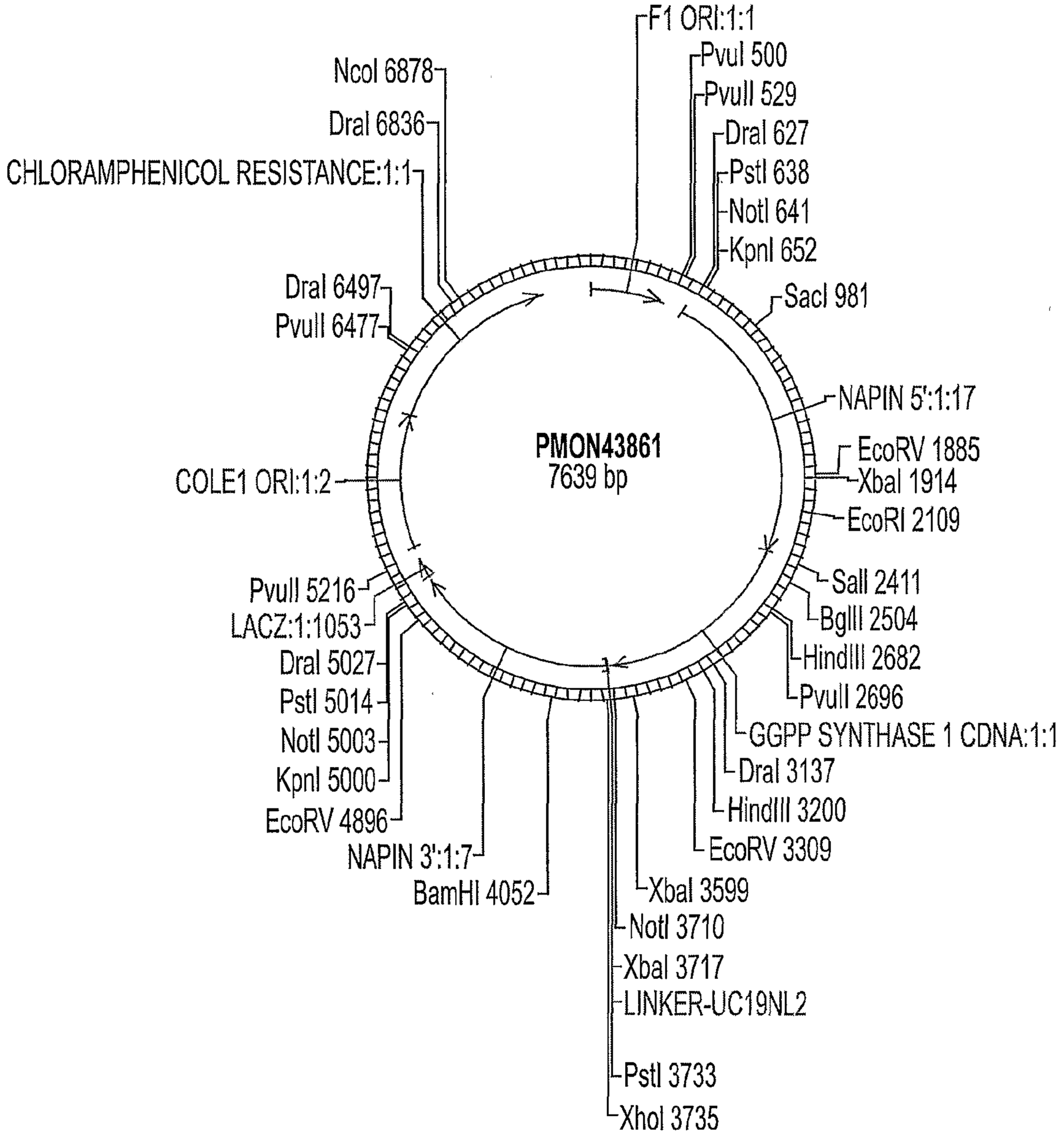

F1 ORI:1:1
Pvul 500
Pvull 529
Dral 627
Pstl 638
Notl 641
Kpnl 652
SacI 981
NAPIN 5':1:17
EcoRV 1885
Xbal 1914
EcoRI 2109
Sall 2411
BglII 2504
HindIII 2682
Pvull 2696
GGPP SYNTHASE 1 CDNA:1:1
Dral 3137
HindIII 3200
EcoRV 3309
Xbal 3599
Notl 3710
Xbal 3717
LINKER-UC19NL2
Pstl 3733
Xhol 3735
Ncol 6878
Dral 6836
CHLORAMPHENICOL RESISTANCE:1:1
Dral 6497
Pvull 6477
COLE1 ORI:1:2
Pvull 5216
LACZ:1:1053
Dral 5027
Pstl 5014
Notl 5003
Kpnl 5000
EcoRV 4896
NAPIN 3':1:7
BamHI 4052
PMON43861
7639 bp

FIG. 41

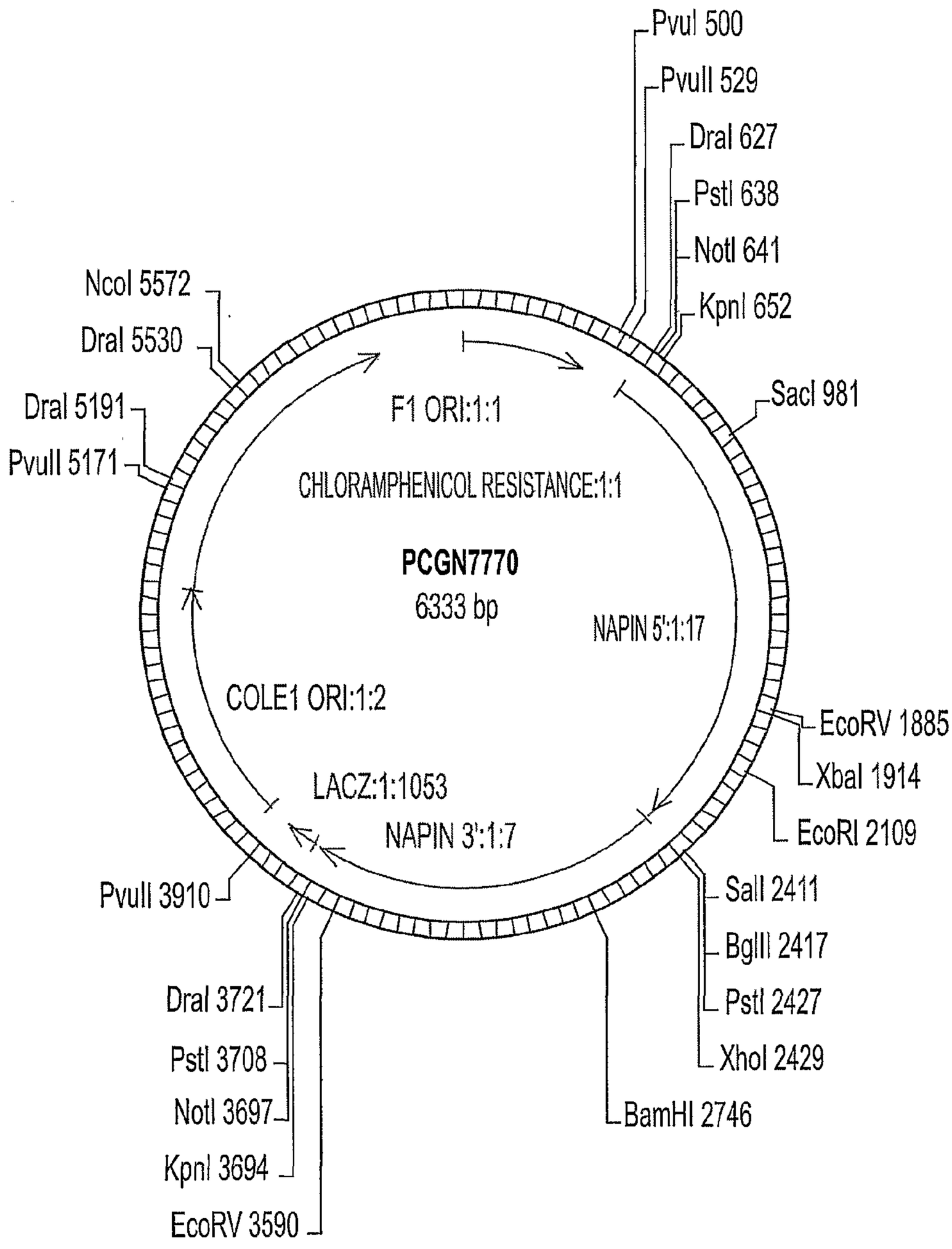

Pvul 500
PvuII 529
DraI 627
PstI 638
NotI 641
KpnI 652
SacI 981
EcoRV 1885
XbaI 1914
EcoRI 2109
SalI 2411
BglII 2417
PstI 2427
XhoI 2429
BamHI 2746
NcoI 5572
DraI 5530
DraI 5191
PvuII 5171
F1 ORI:1:1
CHLORAMPHENICOL RESISTANCE:1:1
PCGN7770
6333 bp
NAPIN 5':1:17
COLE1 ORI:1:2
LACZ:1:1053
NAPIN 3':1:7
PvuII 3910
DraI 3721
PstI 3708
NotI 3697
KpnI 3694
EcoRV 3590

FIG. 42

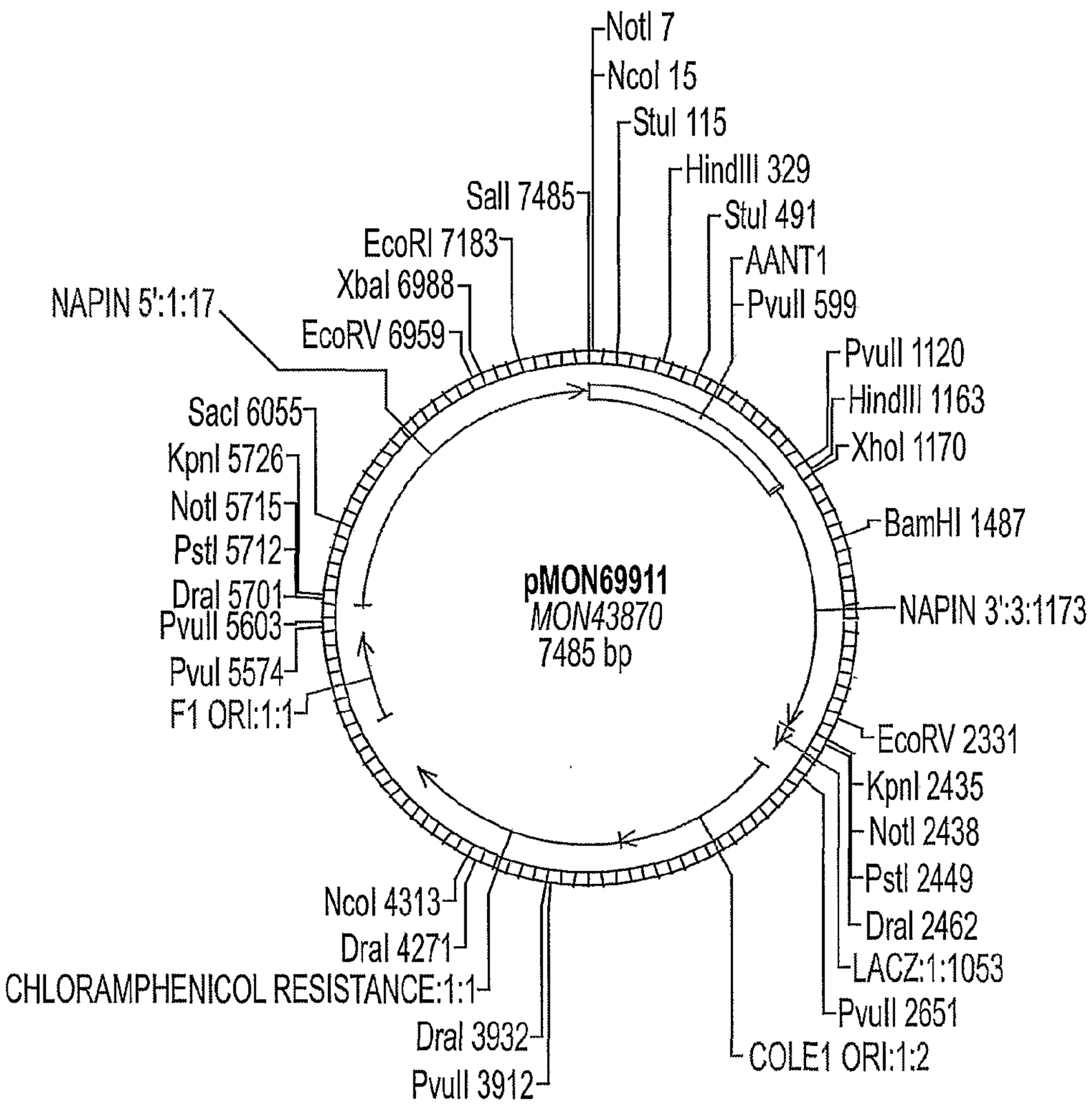

NotI 7
NcoI 15
StuI 115
HindIII 329
StuI 491
AANT1
PvuII 599
PvuII 1120
HindIII 1163
XhoI 1170
BamHI 1487
NAPIN 3':3:1173
EcoRV 2331
KpnI 2435
NotI 2438
PstI 2449
DraI 2462
LACZ:1:1053
PvuII 2651
COLE1 ORI:1:2
PvuII 3912
DraI 3932
CHLORAMPHENICOL RESISTANCE:1:1
DraI 4271
NcoI 4313
F1 ORI:1:1
PvuI 5574
PvuII 5603
DraI 5701
PstI 5712
NotI 5715
KpnI 5726
SacI 6055
NAPIN 5':1:17
EcoRV 6959
XbaI 6988
EcoRI 7183
SalI 7485
pMON69911
MON43870
7485 bp

FIG. 43

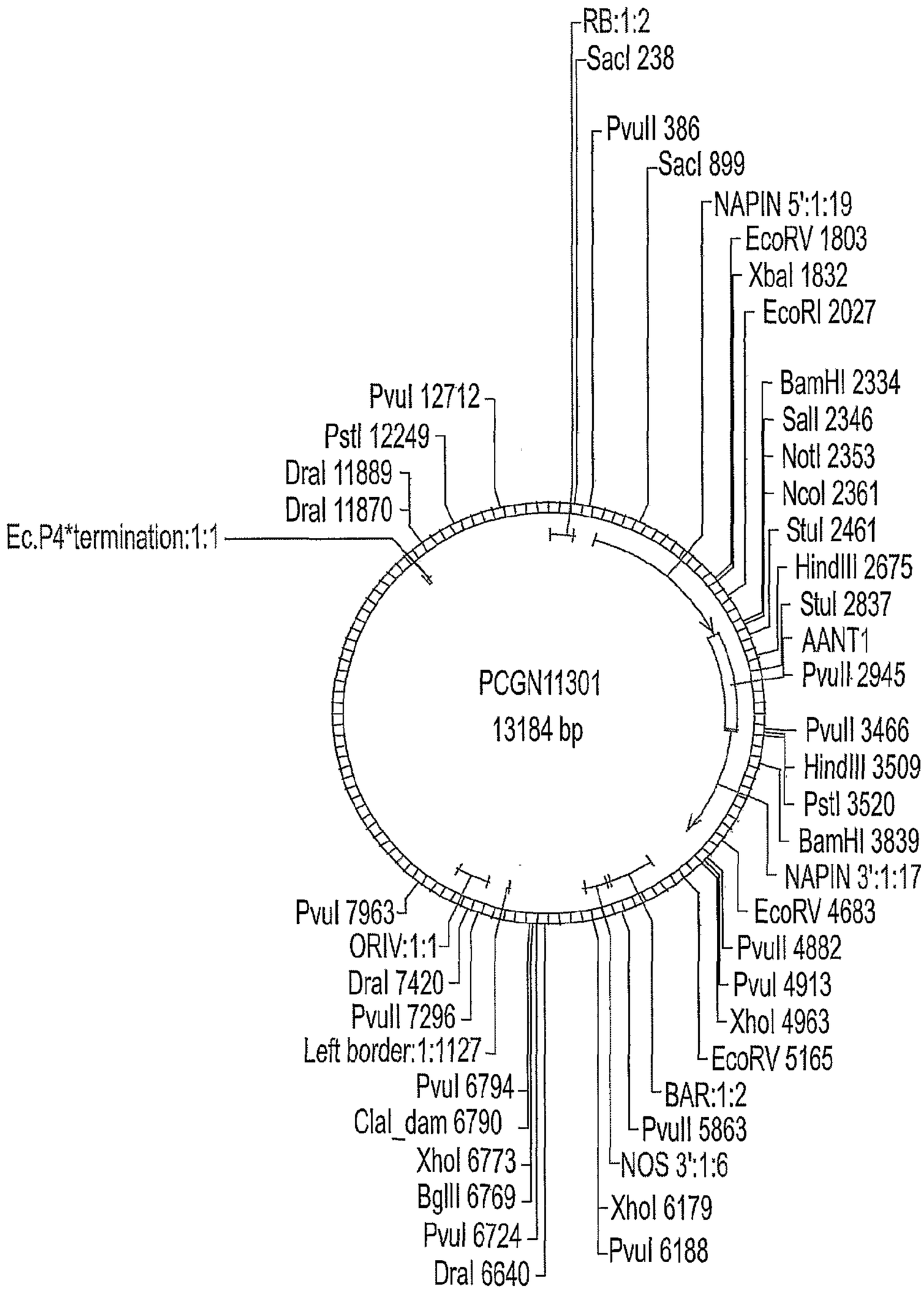

RB:1:2
SacI 238
PvuII 386
SacI 899
NAPIN 5':1:19
EcoRV 1803
XbaI 1832
EcoRI 2027
BamHI 2334
SalI 2346
NotI 2353
NcoI 2361
StuI 2461
HindIII 2675
StuI 2837
AANT1
PvuII 2945
PvuII 3466
HindIII 3509
PstI 3520
BamHI 3839
NAPIN 3':1:17
EcoRV 4683
PvuII 4882
PvuI 4913
XhoI 4963
EcoRV 5165
BAR:1:2
PvuII 5863
NOS 3':1:6
XhoI 6179
PvuI 6188
PvuI 12712
PstI 12249
DraI 11889
DraI 11870
Ec.P4*termination:1:1
PCGN11301
13184 bp
PvuI 7963
ORIV:1:1
DraI 7420
PvuII 7296
Left border:1:1127
PvuI 6794
ClaI_dam 6790
XhoI 6773
BglII 6769
PvuI 6724
DraI 6640

FIG. 44

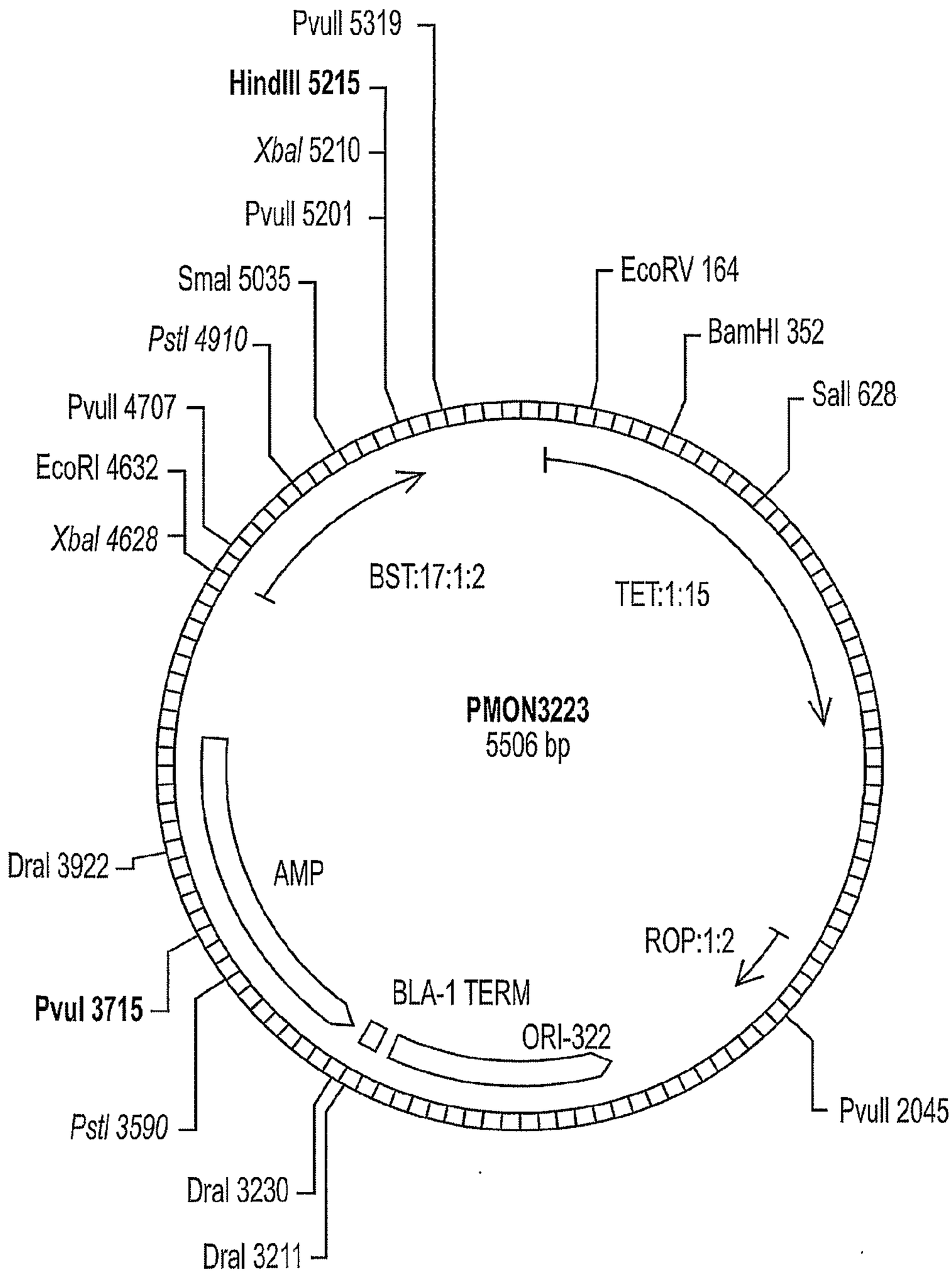

PvuII 5319
HindIII 5215
XbaI 5210
PvuII 5201
SmaI 5035
PstI 4910
PvuII 4707
EcoRI 4632
XbaI 4628
EcoRV 164
BamHI 352
SalI 628
BST:17:1:2
TET:1:15
PMON3223
5506 bp
DraI 3922
AMP
ROP:1:2
BLA-1 TERM
PvuI 3715
ORI-322
PstI 3590
PvuII 2045
DraI 3230
DraI 3211

FIG. 45

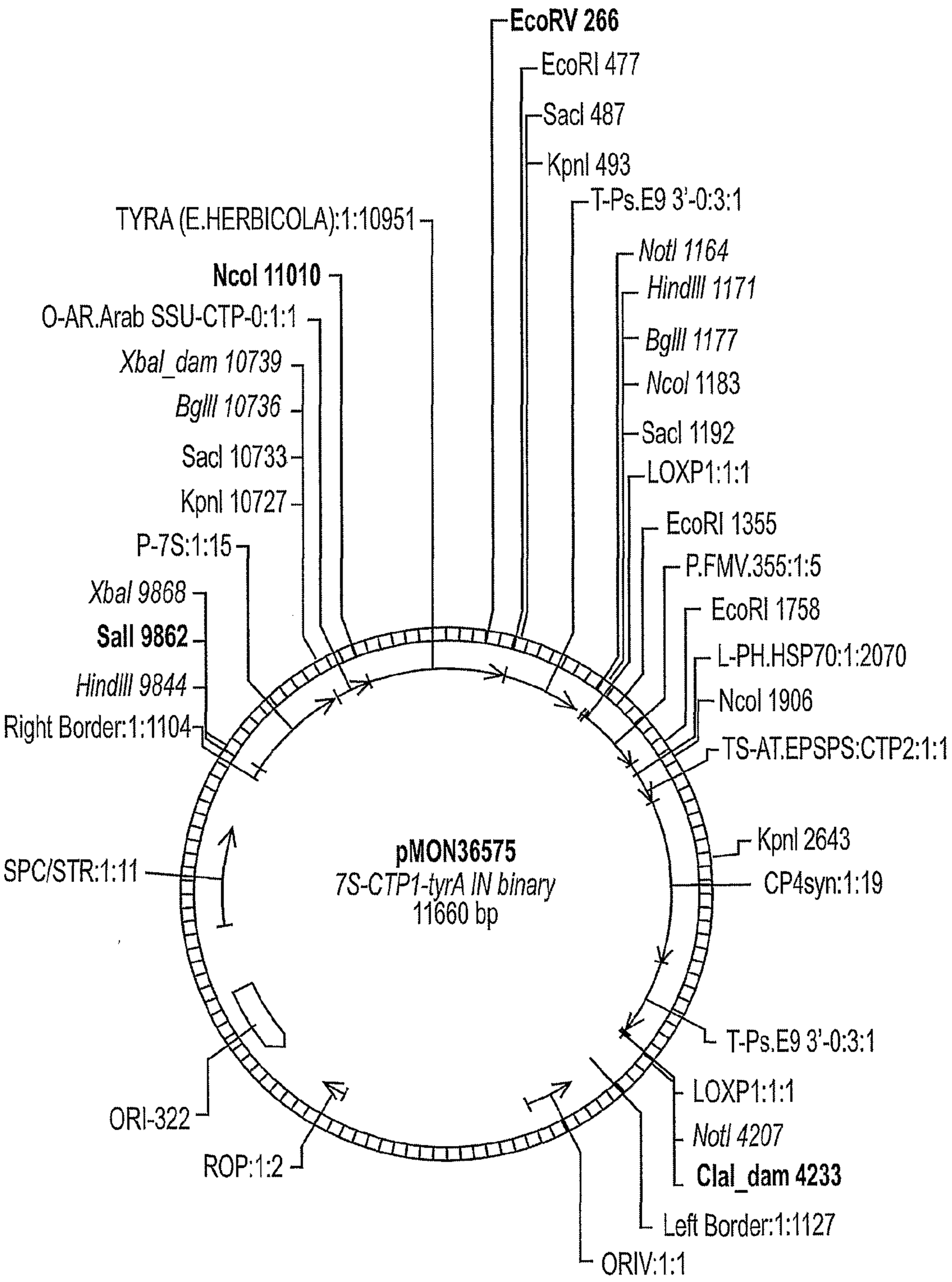

EcoRV 266
EcoRI 477
SacI 487
KpnI 493
T-Ps.E9 3'-0:3:1
NotI 1164
HindIII 1171
BglII 1177
NcoI 1183
SacI 1192
LOXP1:1:1
EcoRI 1355
P.FMV.355:1:5
EcoRI 1758
L-PH.HSP70:1:2070
NcoI 1906
TS-AT.EPSPS:CTP2:1:1
KpnI 2643
CP4syn:1:19
T-Ps.E9 3'-0:3:1
LOXP1:1:1
NotI 4207
ClaI_dam 4233
Left Border:1:1127
ORIV:1:1
TYRA (E.HERBICOLA):1:10951
NcoI 11010
O-AR.Arab SSU-CTP-0:1:1
XbaI_dam 10739
BglII 10736
SacI 10733
KpnI 10727
P-7S:1:15
XbaI 9868
SalI 9862
HindIII 9844
Right Border:1:1104
SPC/STR:1:11
ORI-322
ROP:1:2
pMON36575
7S-CTP1-tyrA IN binary
11660 bp

FIG. 46

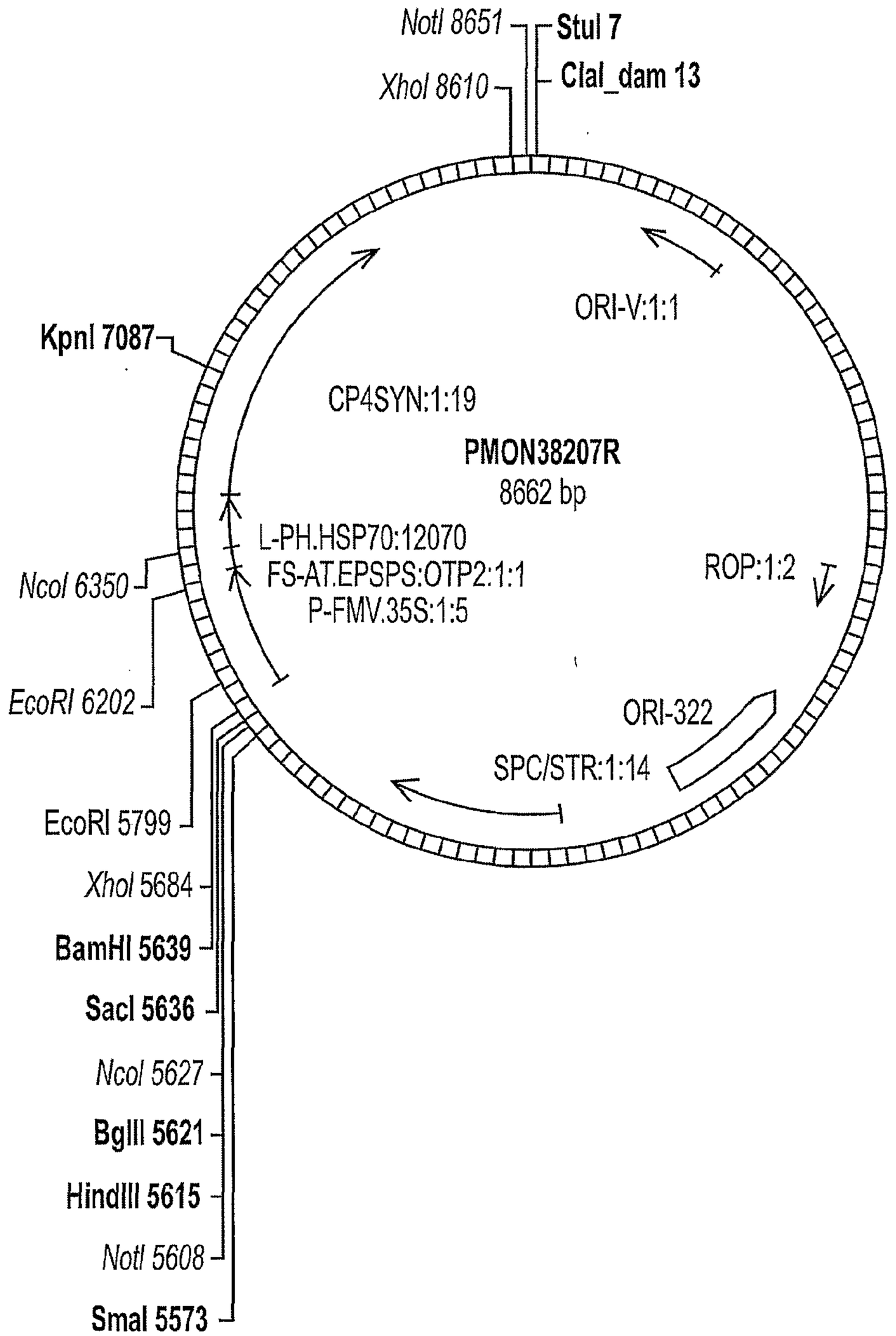

NotI 8651
Stul 7
ClaI_dam 13
XhoI 8610
ORI-V:1:1
KpnI 7087
CP4SYN:1:19
PMON38207R
8662 bp
L-PH.HSP70:12070
FS-AT.EPSPS:OTP2:1:1
P-FMV.35S:1:5
ROP:1:2
NcoI 6350
EcoRI 6202
ORI-322
SPC/STR:1:14
EcoRI 5799
XhoI 5684
BamHI 5639
SacI 5636
NcoI 5627
BglII 5621
HindIII 5615
NotI 5608
SmaI 5573

FIG. 47

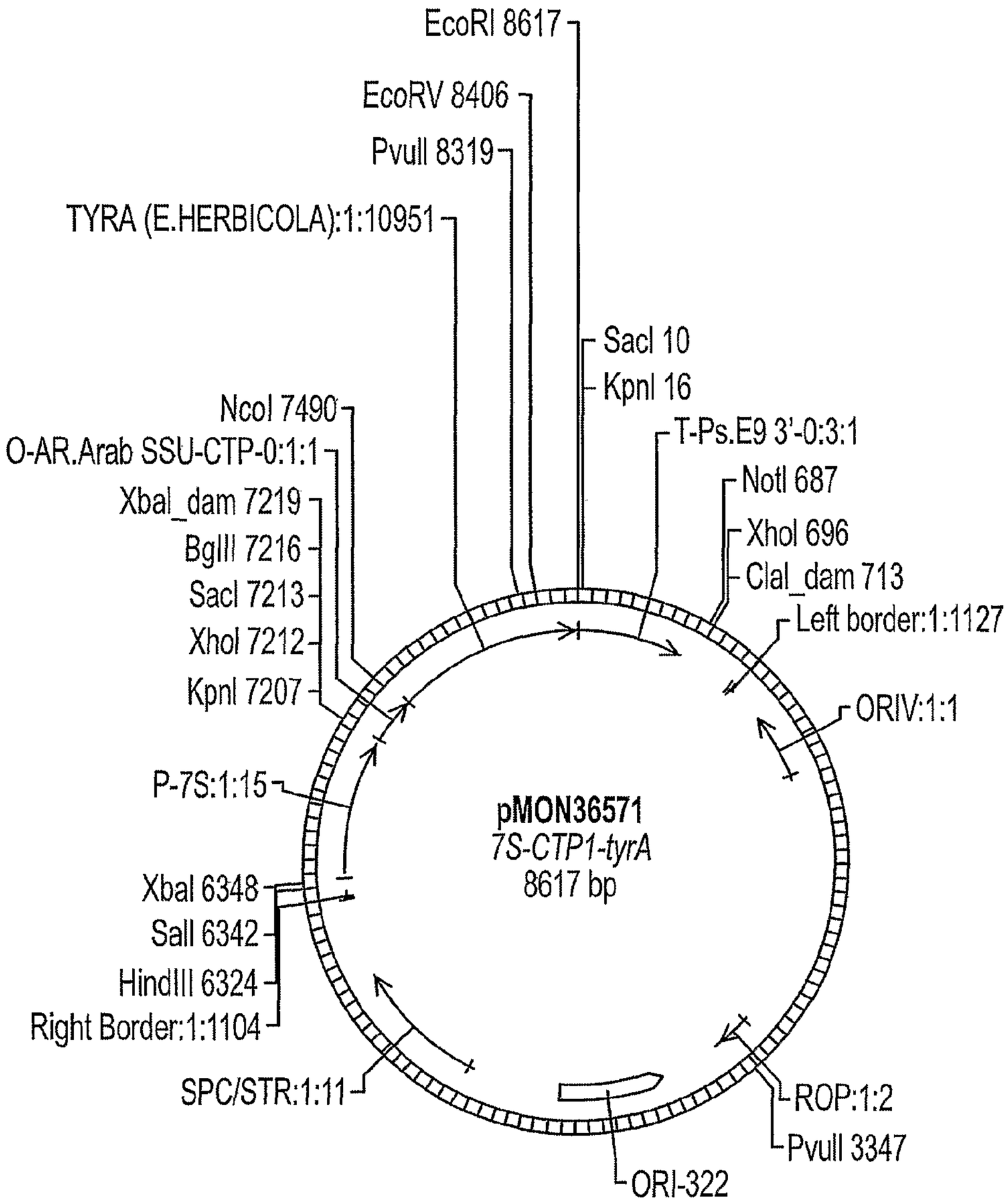

EcoRI 8617
EcoRV 8406
PvuII 8319
TYRA (E.HERBICOLA):1:10951
SacI 10
KpnI 16
NcoI 7490
O-AR.Arab SSU-CTP-0:1:1
T-Ps.E9 3'-0:3:1
NotI 687
XbaI_dam 7219
XhoI 696
BglII 7216
ClaI_dam 713
SacI 7213
Left border:1:1127
XhoI 7212
KpnI 7207
ORIV:1:1
P-7S:1:15
pMON36571
7S-CTP1-tyrA
8617 bp
XbaI 6348
SalI 6342
HindIII 6324
Right Border:1:1104
SPC/STR:1:11
ROP:1:2
PvuII 3347
ORI-322

FIG. 48

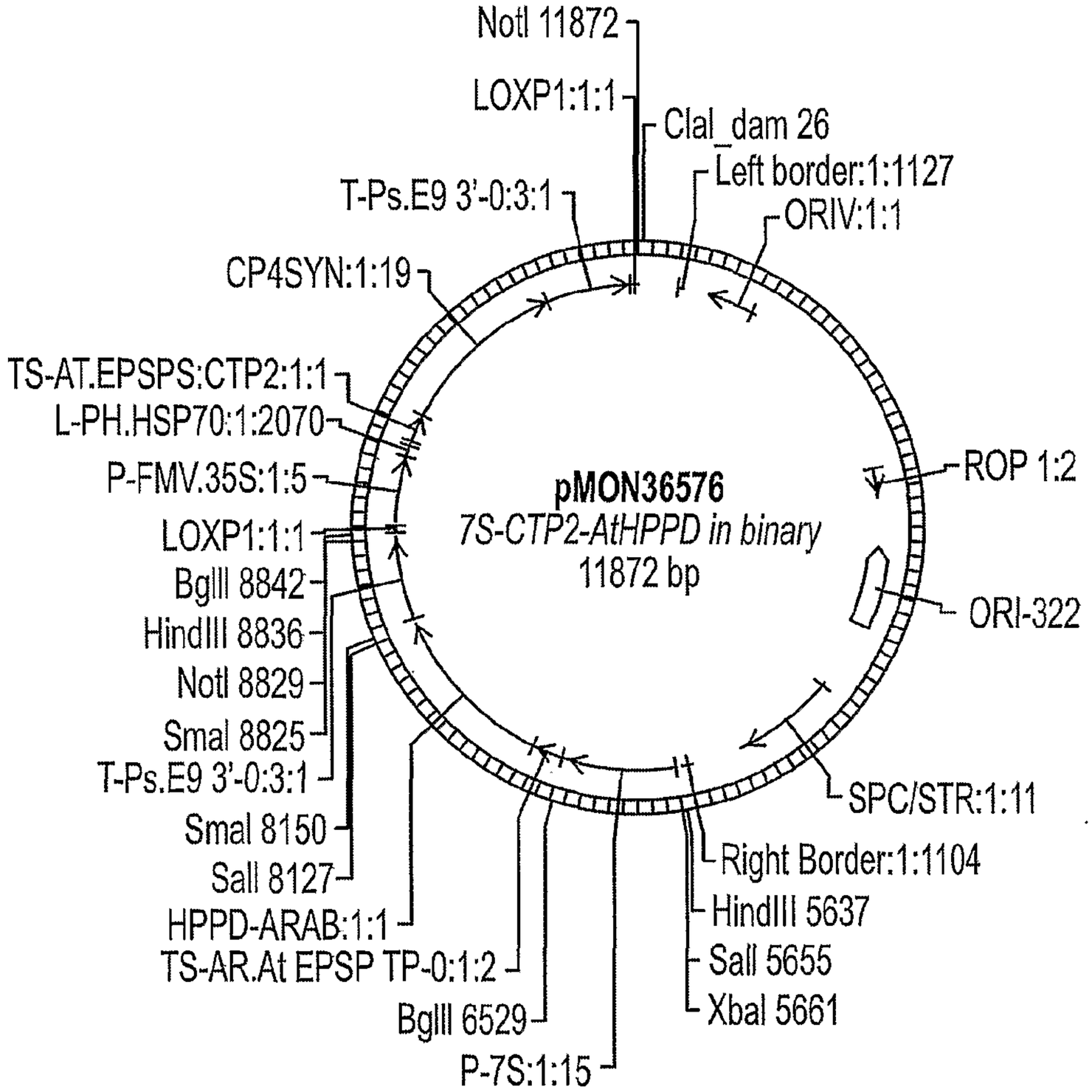

NotI 11872
LOXP1:1:1
ClaI_dam 26
Left border:1:1127
T-Ps.E9 3'-0:3:1
ORIV:1:1
CP4SYN:1:19
TS-AT.EPSPS:CTP2:1:1
L-PH.HSP70:1:2070
P-FMV.35S:1:5
ROP 1:2
pMON36576
7S-CTP2-AtHPPD in binary
11872 bp
LOXP1:1:1
BglII 8842
HindIII 8836
ORI-322
NotI 8829
SmaI 8825
T-Ps.E9 3'-0:3:1
SPC/STR:1:11
SmaI 8150
SalI 8127
Right Border:1:1104
HindIII 5637
HPPD-ARAB:1:1
TS-AR.At EPSP TP-0:1:2
SalI 5655
BglII 6529
XbaI 5661
P-7S:1:15

FIG. 49

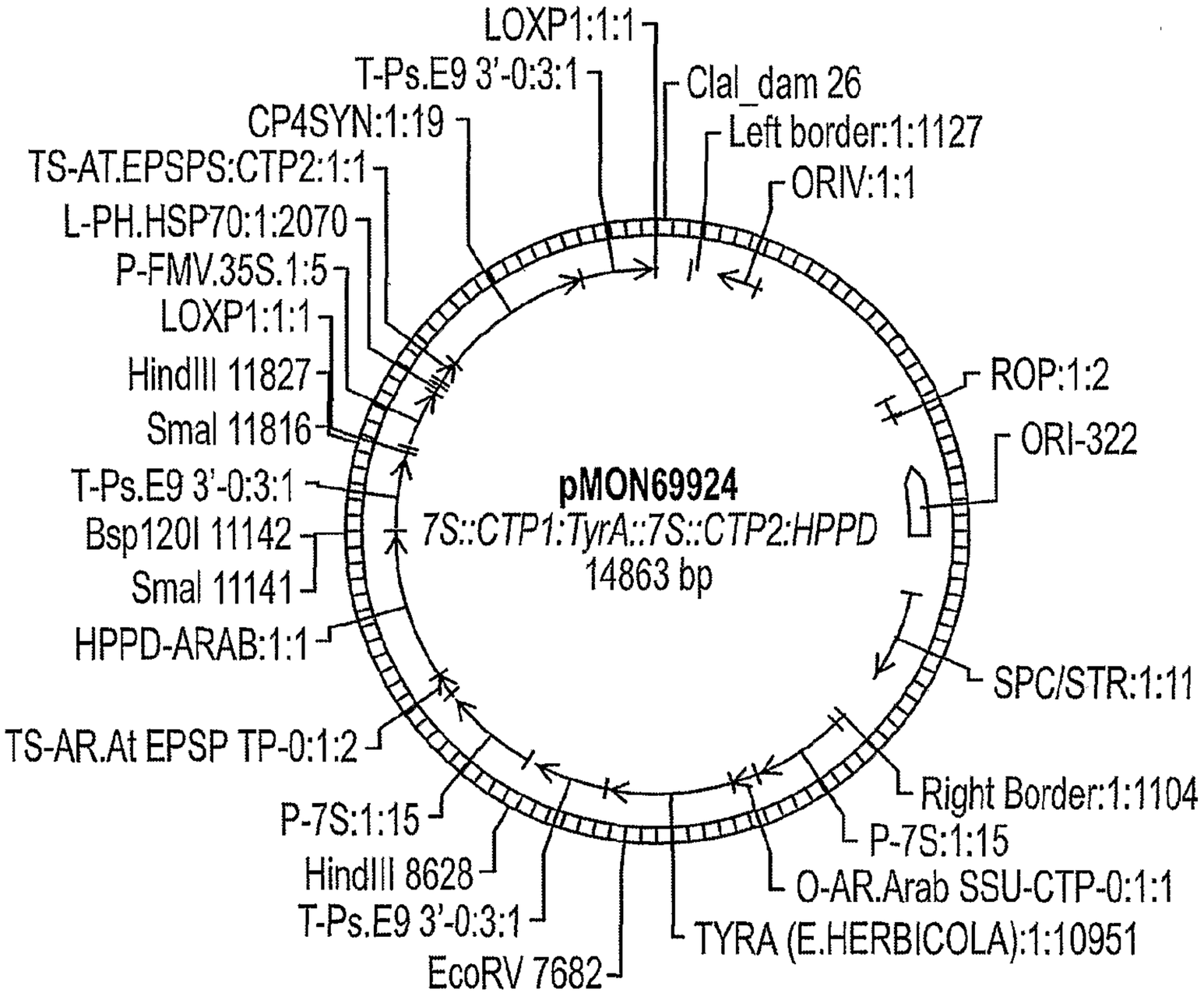

LOXP1:1:1
T-Ps.E9 3'-0:3:1
ClaI_dam 26
CP4SYN:1:19
Left border:1:1127
TS-AT.EPSPS:CTP2:1:1
ORIV:1:1
L-PH.HSP70:1:2070
P-FMV.35S.1:5
LOXP1:1:1
HindIII 11827
ROP:1:2
SmaI 11816
ORI-322
T-Ps.E9 3'-0:3:1
pMON69924
7S::CTP1:TyrA::7S::CTP2:HPPD
14863 bp
Bsp120I 11142
SmaI 11141
HPPD-ARAB:1:1
SPC/STR:1:11
TS-AR.At EPSP TP-0:1:2
Right Border:1:1104
P-7S:1:15
P-7S:1:15
HindIII 8628
O-AR.Arab SSU-CTP-0:1:1
T-Ps.E9 3'-0:3:1
TYRA (E.HERBICOLA):1:10951
EcoRV 7682

FIG. 50

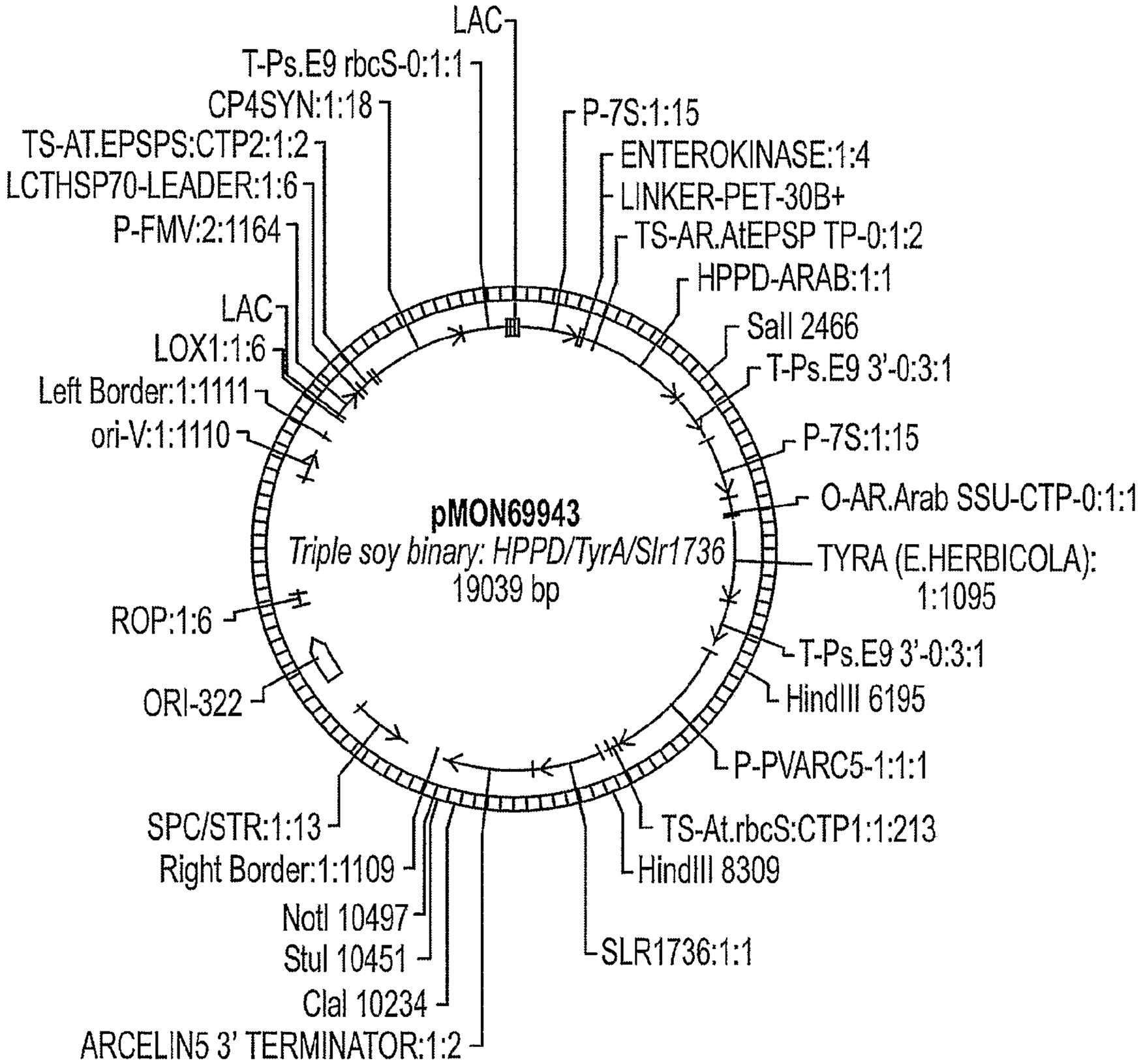

LAC
T-Ps.E9 rbcS-0:1:1
CP4SYN:1:18
TS-AT.EPSPS:CTP2:1:2
LCTHSP70-LEADER:1:6
P-FMV:2:1164
LAC
LOX1:1:6
Left Border:1:1111
ori-V:1:1110
ROP:1:6
ORI-322
SPC/STR:1:13
Right Border:1:1109
NotI 10497
StuI 10451
ClaI 10234
ARCELIN5 3' TERMINATOR:1:2
pMON69943
Triple soy binary: HPPD/TyrA/Slr1736
19039 bp
P-7S:1:15
ENTEROKINASE:1:4
LINKER-PET-30B+
TS-AR.AtEPSP TP-0:1:2
HPPD-ARAB:1:1
SalI 2466
T-Ps.E9 3'-0:3:1
P-7S:1:15
O-AR.Arab SSU-CTP-0:1:1
TYRA (E.HERBICOLA):
1:1095
T-Ps.E9 3'-0:3:1
HindIII 6195
P-PVARC5-1:1:1
TS-At.rbcS:CTP1:1:213
HindIII 8309
SLR1736:1:1

FIG. 51

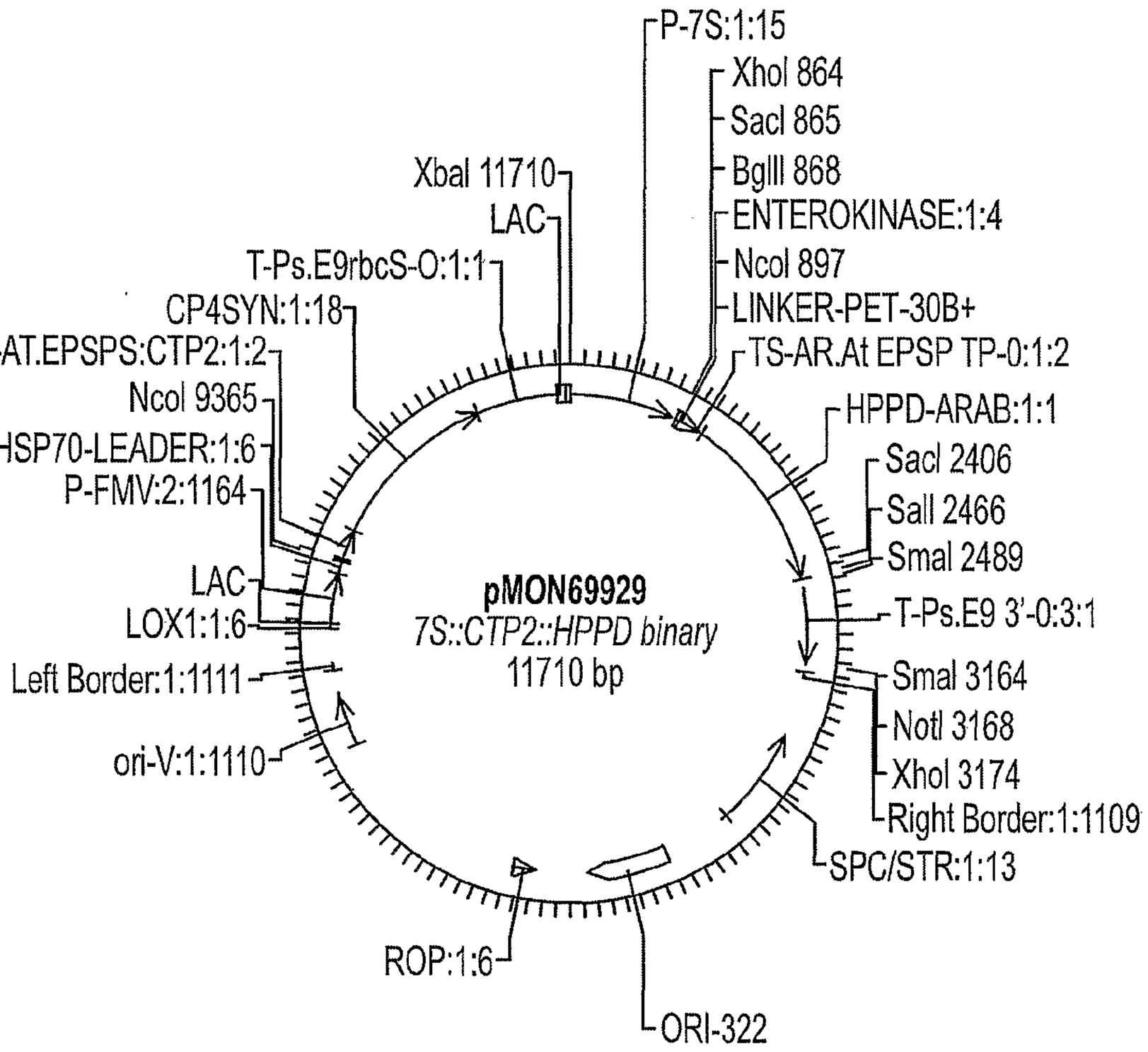

P-7S:1:15
XhoI 864
SacI 865
BglII 868
ENTEROKINASE:1:4
NcoI 897
LINKER-PET-30B+
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
SacI 2406
SalI 2466
SmaI 2489
T-Ps.E9 3'-0:3:1
SmaI 3164
NotI 3168
XhoI 3174
Right Border:1:1109
SPC/STR:1:13
ORI-322
ROP:1:6
ori-V:1:1110
Left Border:1:1111
LOX1:1:6
LAC
NcoI 9365
PCTHSP70-LEADER:1:6
P-FMV:2:1164
TS-AT.EPSPS:CTP2:1:2
CP4SYN:1:18
T-Ps.E9rbcS-O:1:1
XbaI 11710
LAC
pMON69929
7S::CTP2::HPPD binary
11710 bp

FIG. 52

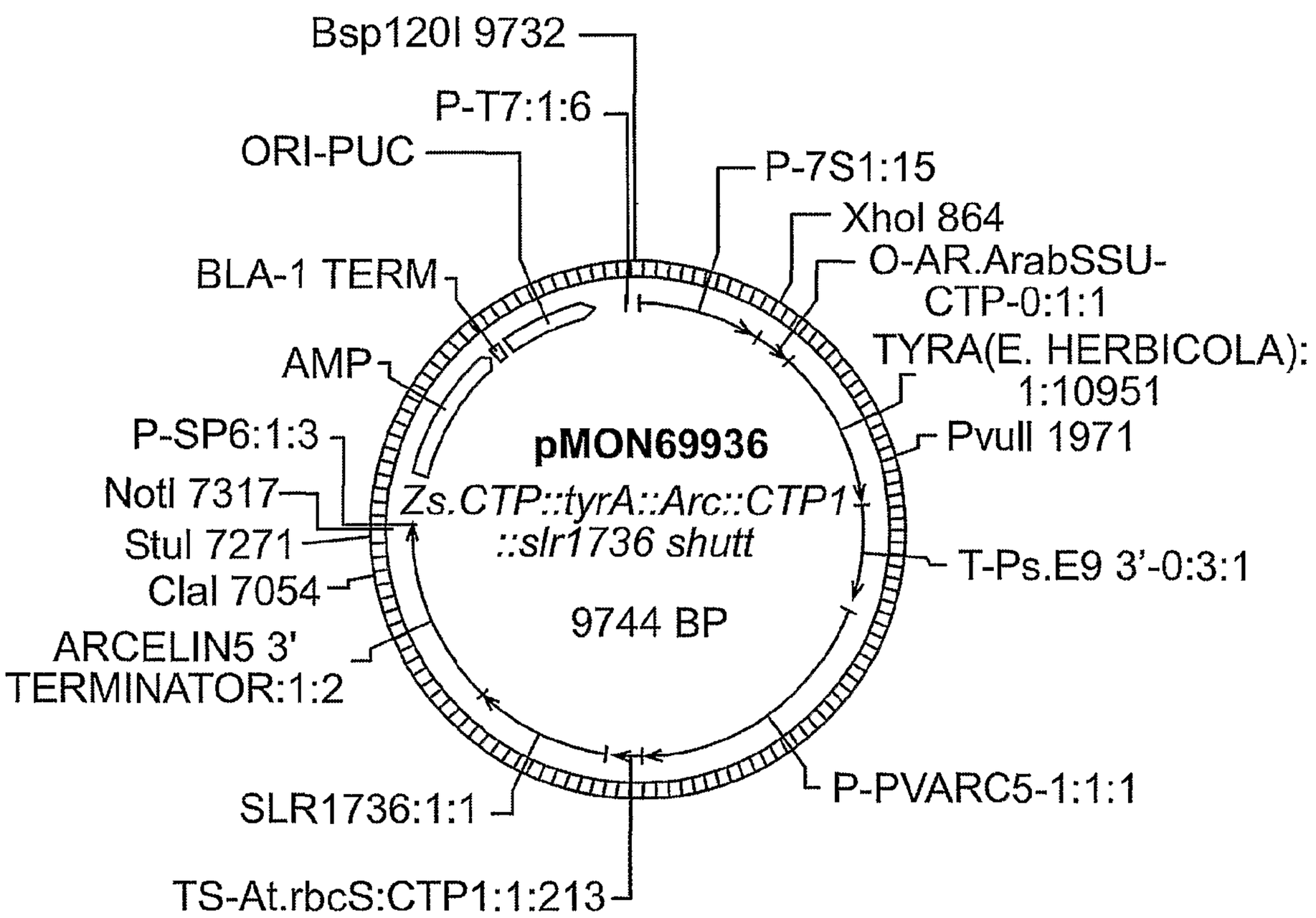

Bsp120I 9732
P-T7:1:6
ORI-PUC
P-7S1:15
XhoI 864
O-AR.ArabSSU-CTP-0:1:1
TYRA(E. HERBICOLA): 1:10951
PvuII 1971
T-Ps.E9 3'-0:3:1
P-PVARC5-1:1:1
BLA-1 TERM
AMP
P-SP6:1:3
NotI 7317
StuI 7271
ClaI 7054
ARCELIN5 3' TERMINATOR:1:2
SLR1736:1:1
TS-At.rbcS:CTP1:1:213
pMON69936
Zs.CTP::tyrA::Arc::CTP1 ::slr1736 shutt
9744 BP

FIG. 53

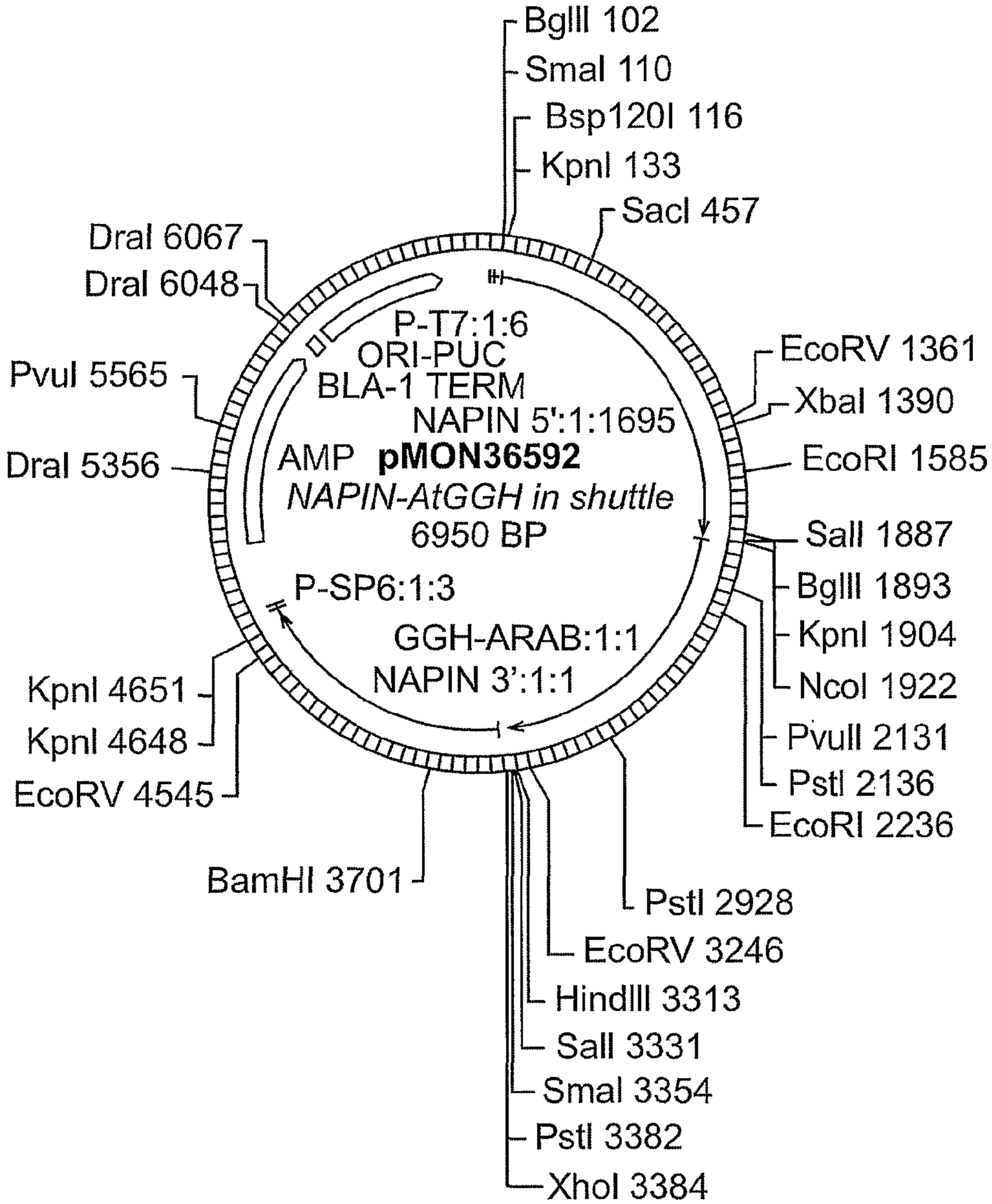

BglII 102
SmaI 110
Bsp120I 116
KpnI 133
SacI 457
DraI 6067
DraI 6048
P-T7:1:6
ORI-PUC
BLA-1 TERM
NAPIN 5':1:1695
AMP pMON36592
NAPIN-AtGGH in shuttle
6950 BP
P-SP6:1:3
GGH-ARAB:1:1
NAPIN 3':1:1
EcoRV 1361
XbaI 1390
EcoRI 1585
SalI 1887
BglII 1893
KpnI 1904
NcoI 1922
PvuII 2131
PstI 2136
EcoRI 2236
PvuI 5565
DraI 5356
KpnI 4651
KpnI 4648
EcoRV 4545
BamHI 3701
PstI 2928
EcoRV 3246
HindIII 3313
SalI 3331
SmaI 3354
PstI 3382
XhoI 3384

FIG. 54

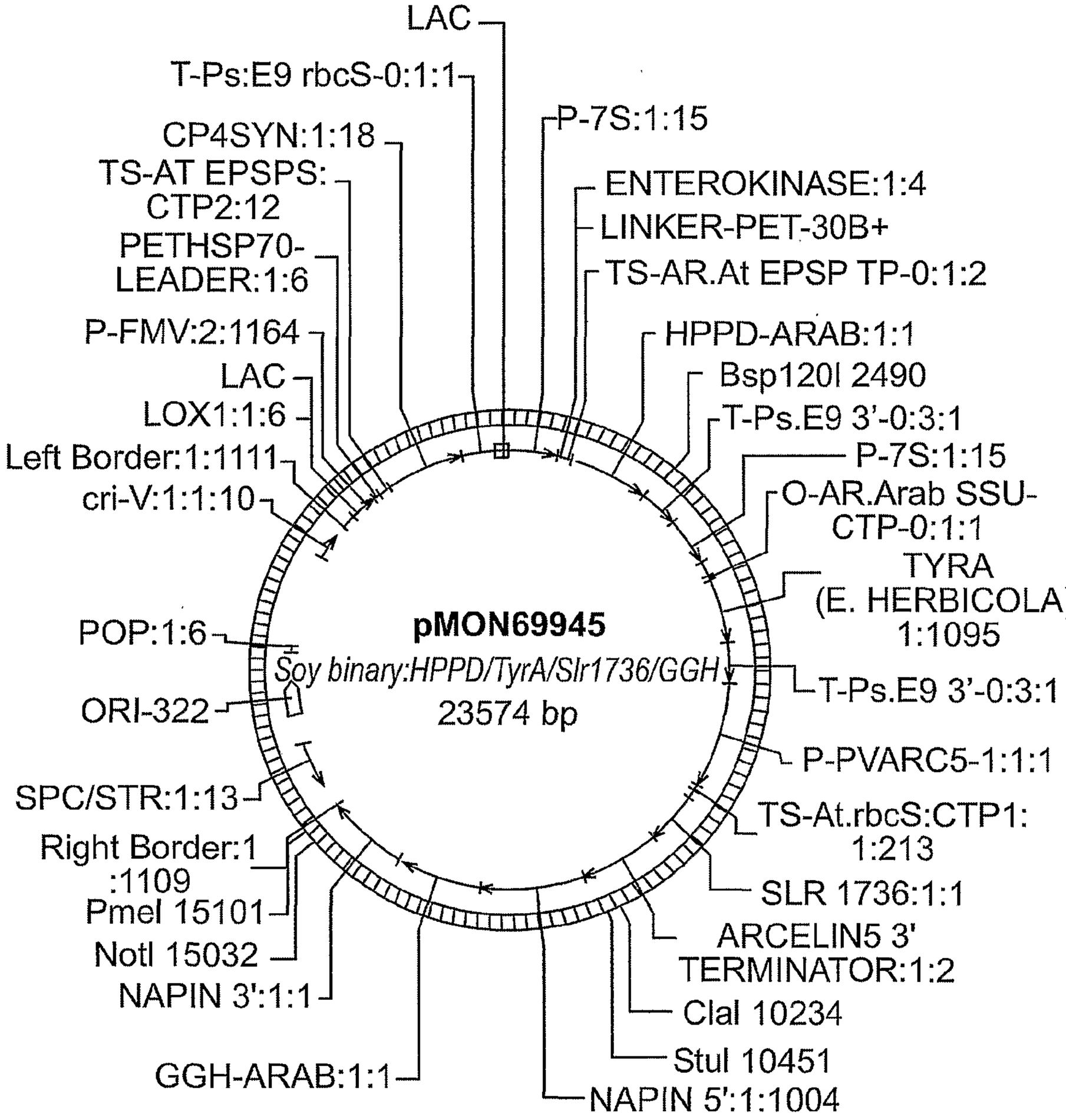

pMON69945
Soy binary:HPPD/TyrA/Slr1736/GGH
23574 bp
LAC
T-Ps:E9 rbcS-0:1:1
CP4SYN:1:18
TS-AT EPSPS: CTP2:12
PETHSP70-LEADER:1:6
P-FMV:2:1164
LAC
LOX1:1:6
Left Border:1:1111
cri-V:1:1:10
POP:1:6
ORI-322
SPC/STR:1:13
Right Border:1:1109
PmeI 15101
NotI 15032
NAPIN 3':1:1
GGH-ARAB:1:1
P-7S:1:15
ENTEROKINASE:1:4
LINKER-PET-30B+
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
Bsp120I 2490
T-Ps.E9 3'-0:3:1
P-7S:1:15
O-AR.Arab SSU-CTP-0:1:1
TYRA (E. HERBICOLA): 1:1095
T-Ps.E9 3'-0:3:1
P-PVARC5-1:1:1
TS-At.rbcS:CTP1: 1:213
SLR 1736:1:1
ARCELIN5 3' TERMINATOR:1:2
ClaI 10234
StuI 10451
NAPIN 5':1:1004

FIG. 55

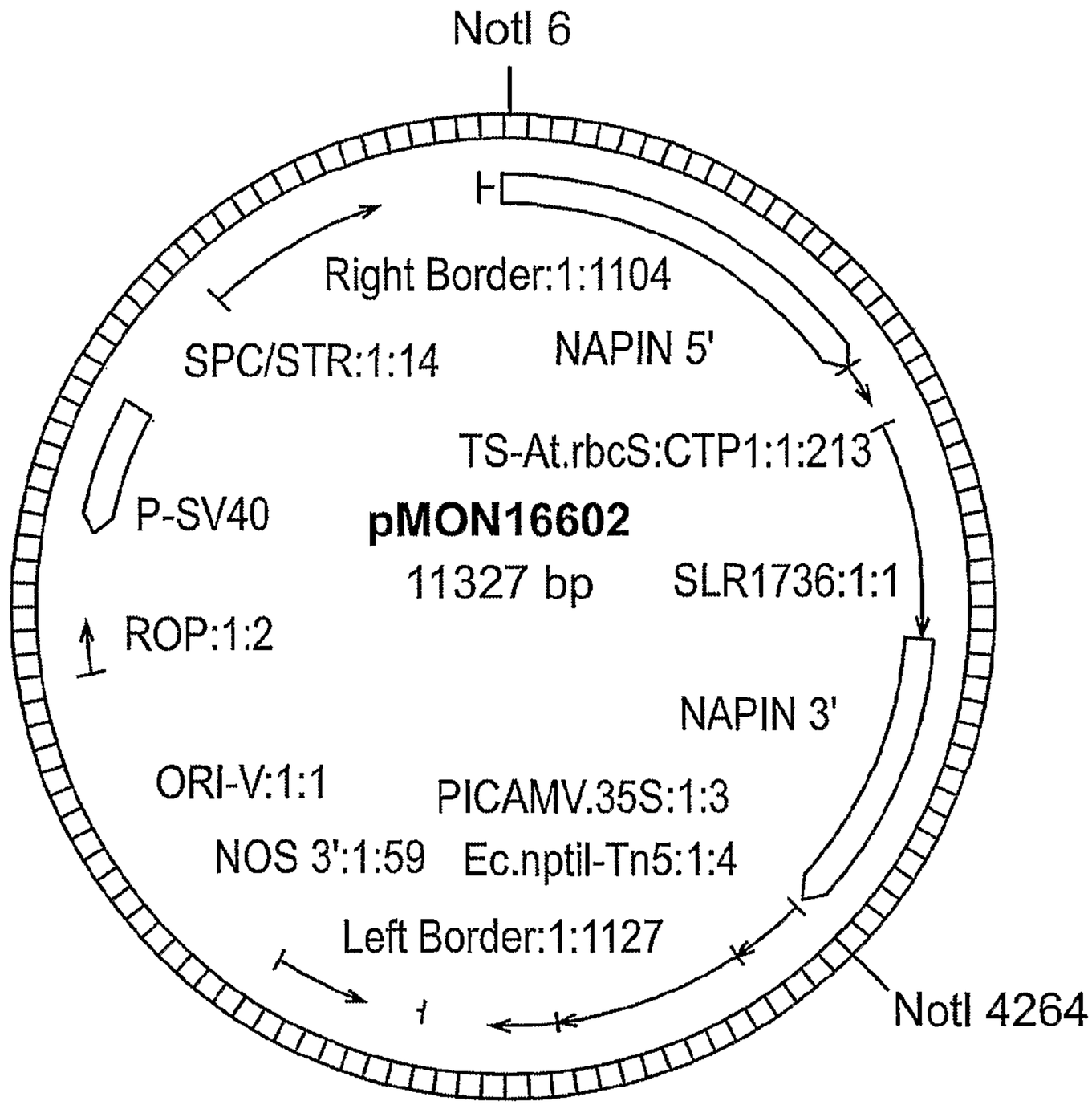

NotI 6
Right Border:1:1104
SPC/STR:1:14
NAPIN 5'
TS-At.rbcS:CTP1:1:213
P-SV40
pMON16602
11327 bp
SLR1736:1:1
ROP:1:2
NAPIN 3'
ORI-V:1:1
PICAMV.35S:1:3
NOS 3':1:59
Ec.nptII-Tn5:1:4
Left Border:1:1127
NotI 4264

FIG. 56

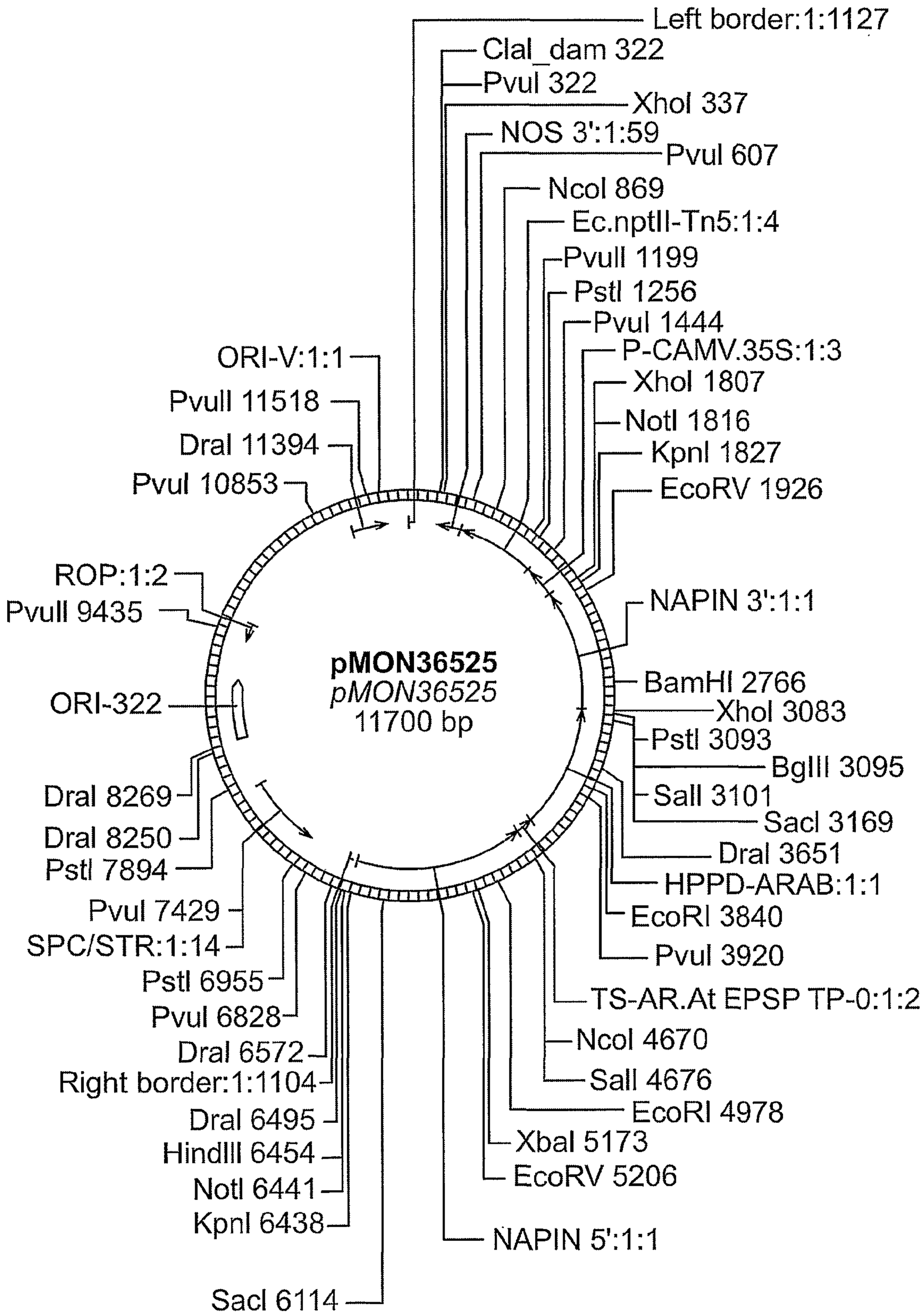

Left border:1:1127
ClaI_dam 322
Pvul 322
XhoI 337
NOS 3':1:59
Pvul 607
NcoI 869
Ec.nptII-Tn5:1:4
PvuII 1199
PstI 1256
Pvul 1444
P-CAMV.35S:1:3
XhoI 1807
NotI 1816
KpnI 1827
EcoRV 1926
NAPIN 3':1:1
BamHI 2766
XhoI 3083
PstI 3093
BglII 3095
SalI 3101
SacI 3169
DraI 3651
HPPD-ARAB:1:1
EcoRI 3840
Pvul 3920
TS-AR.At EPSP TP-0:1:2
NcoI 4670
SalI 4676
EcoRI 4978
XbaI 5173
EcoRV 5206
NAPIN 5':1:1
ORI-V:1:1
PvuII 11518
DraI 11394
Pvul 10853
ROP:1:2
PvuII 9435
pMON36525
pMON36525
11700 bp
ORI-322
DraI 8269
DraI 8250
PstI 7894
Pvul 7429
SPC/STR:1:14
PstI 6955
Pvul 6828
DraI 6572
Right border:1:1104
DraI 6495
HindIII 6454
NotI 6441
KpnI 6438
SacI 6114

FIG. 57

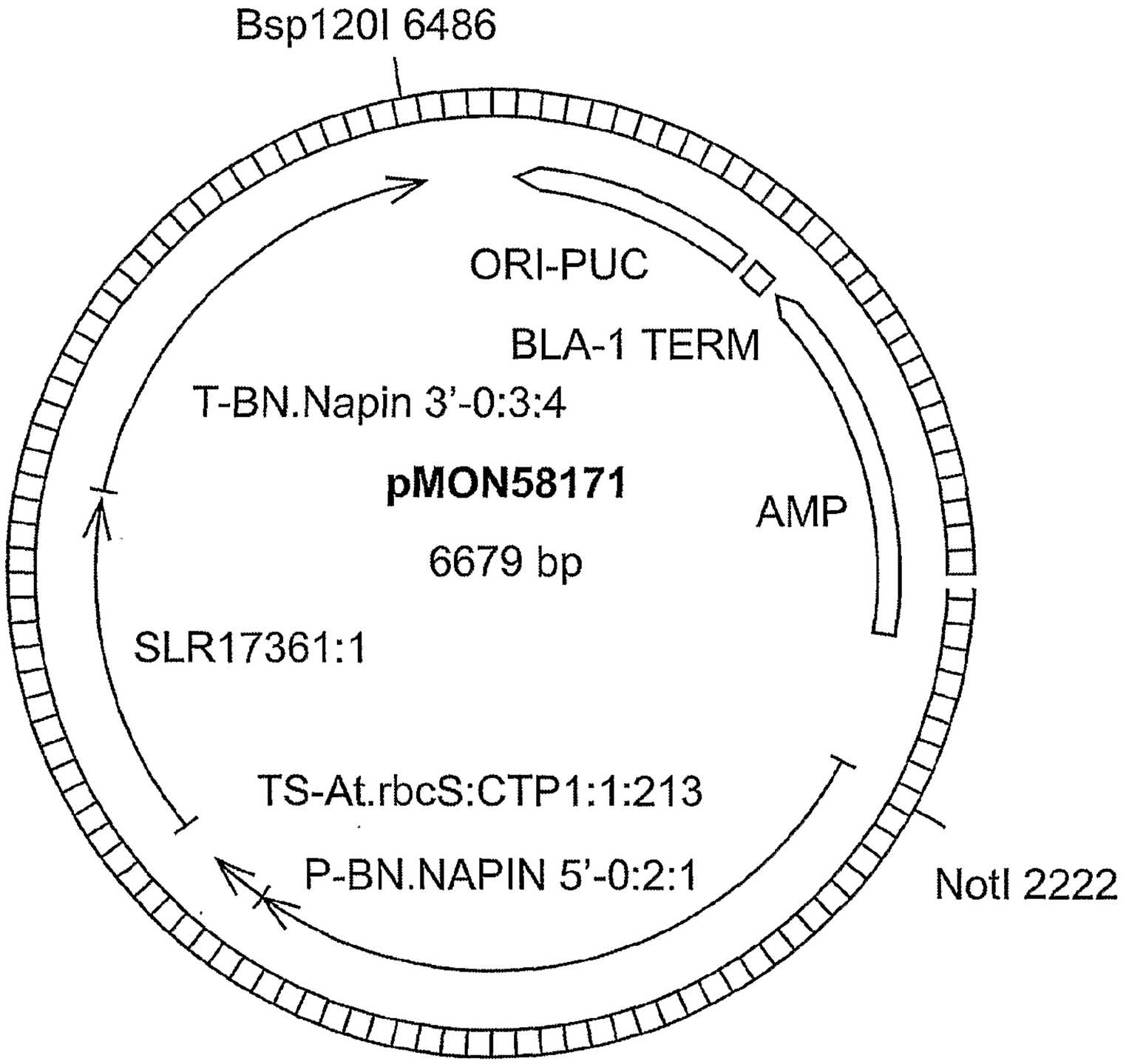

Bsp120I 6486
ORI-PUC
BLA-1 TERM
T-BN.Napin 3'-0:3:4
pMON58171
6679 bp
AMP
SLR17361:1
TS-At.rbcS:CTP1:1:213
P-BN.NAPIN 5'-0:2:1
NotI 2222

FIG. 58

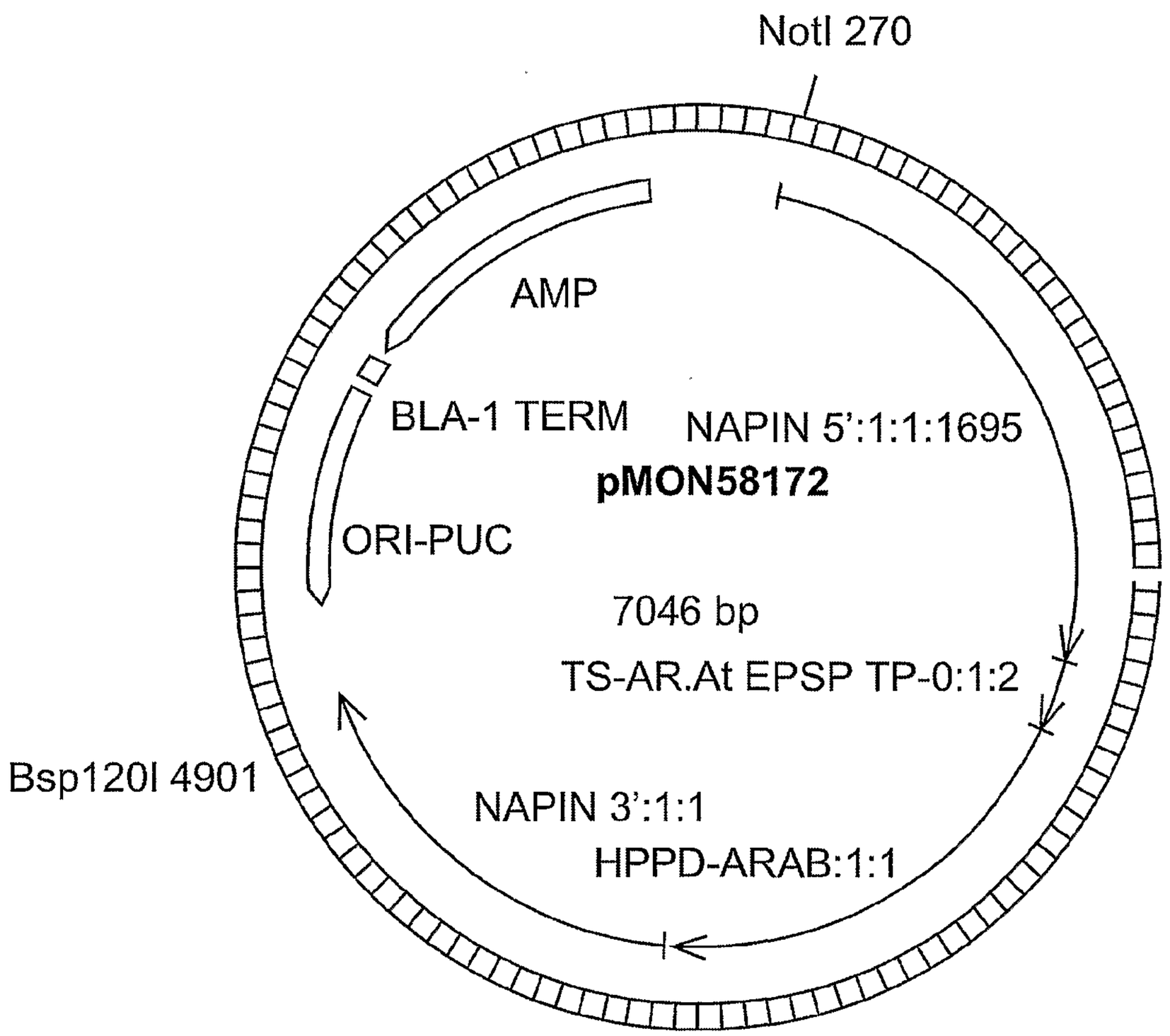

NotI 270
AMP
BLA-1 TERM
NAPIN 5':1:1:1695
pMON58172
ORI-PUC
7046 bp
TS-AR.At EPSP TP-0:1:2
Bsp120I 4901
NAPIN 3':1:1
HPPD-ARAB:1:1

FIG. 59

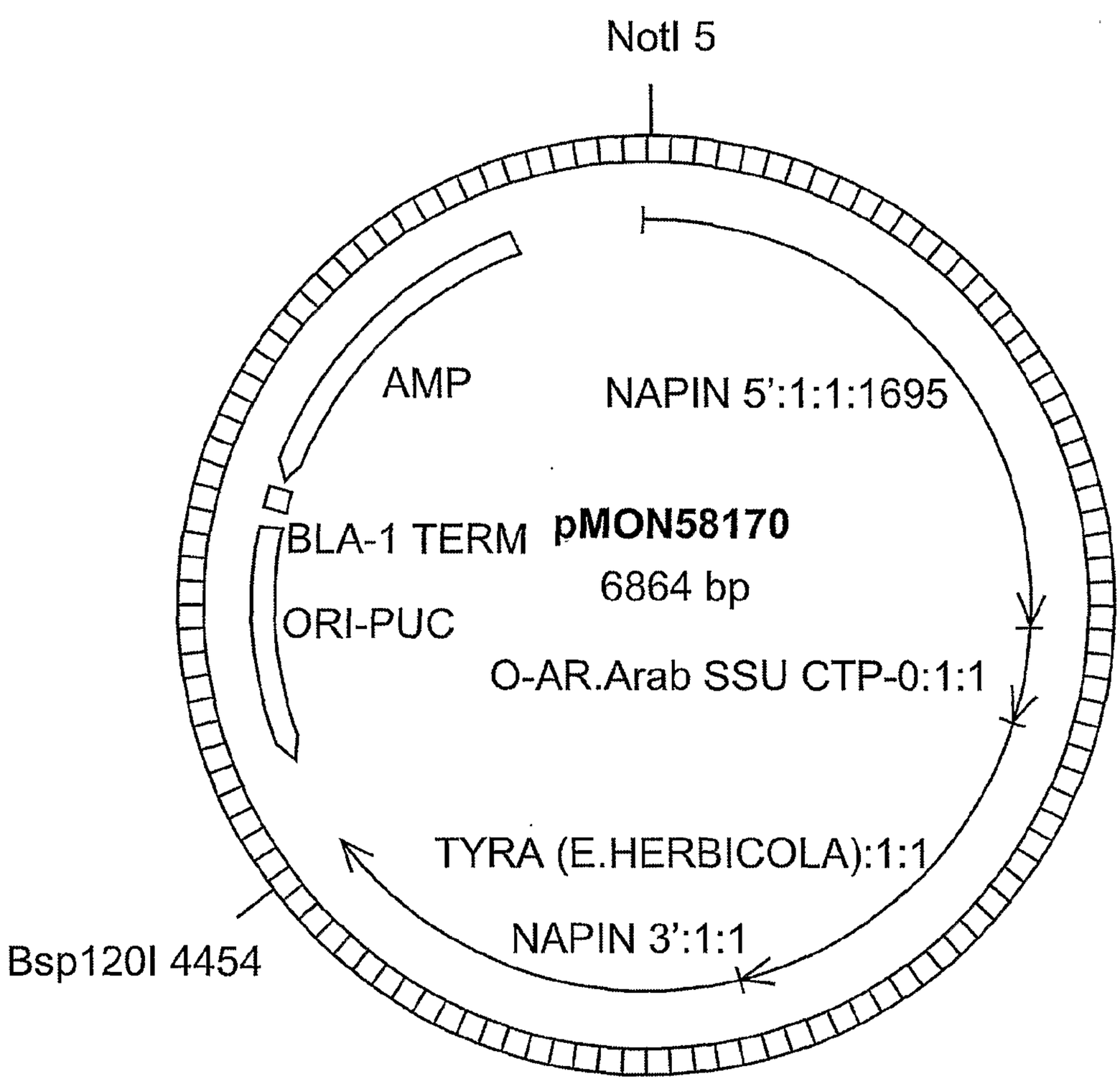

NotI 5
AMP
NAPIN 5':1:1:1695
BLA-1 TERM
pMON58170
6864 bp
ORI-PUC
O-AR.Arab SSU CTP-0:1:1
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:1
Bsp120I 4454

FIG. 60

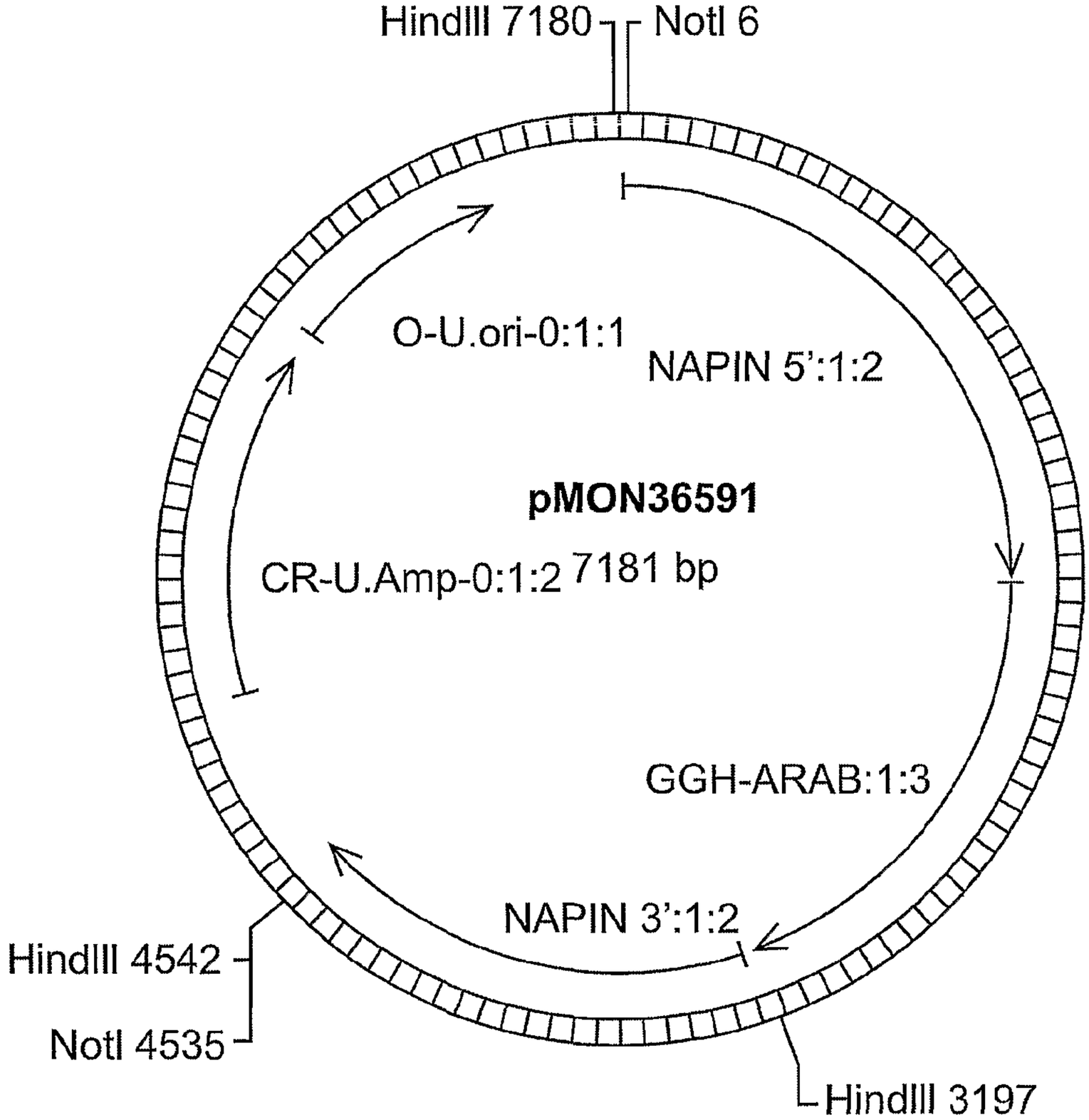
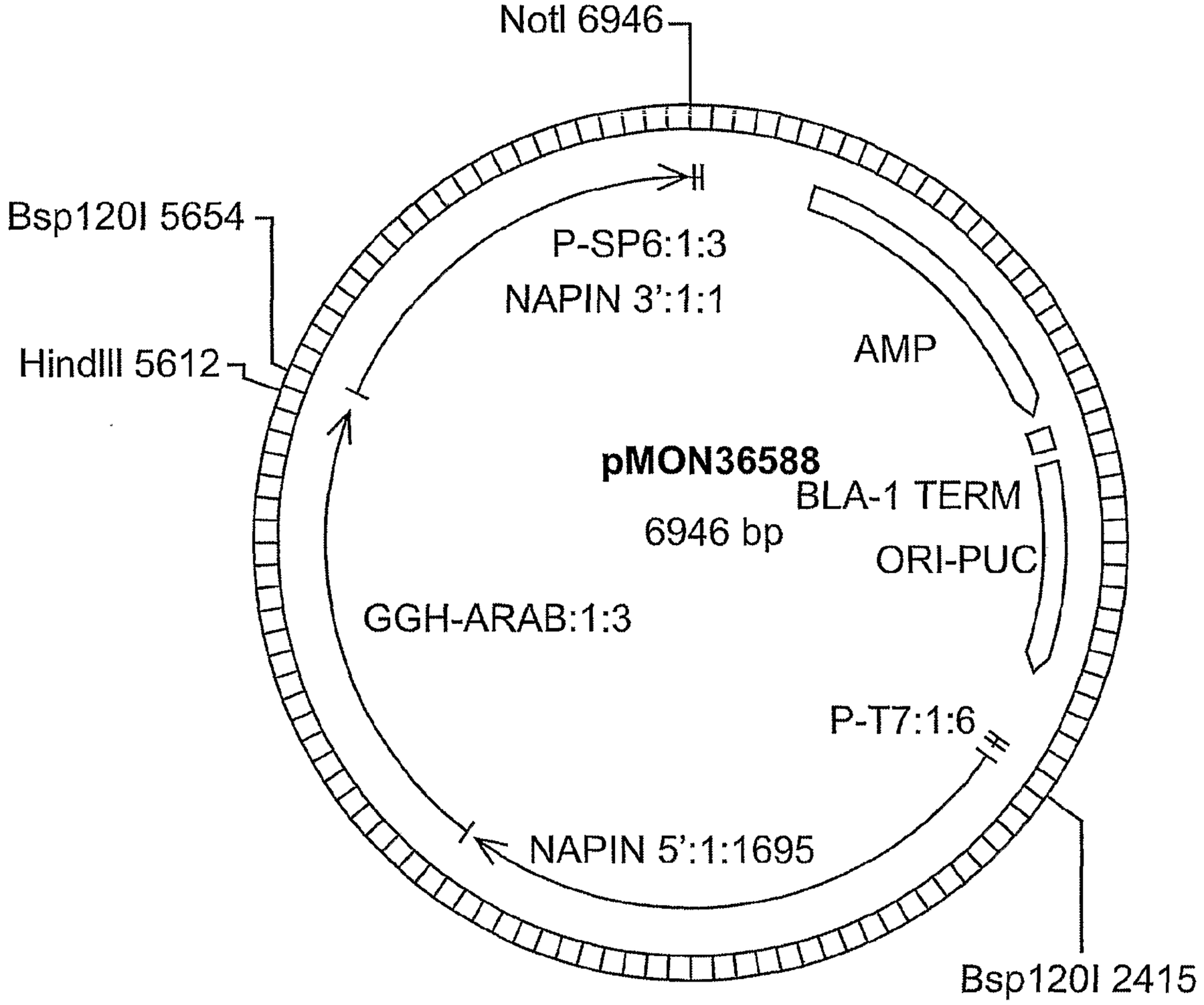
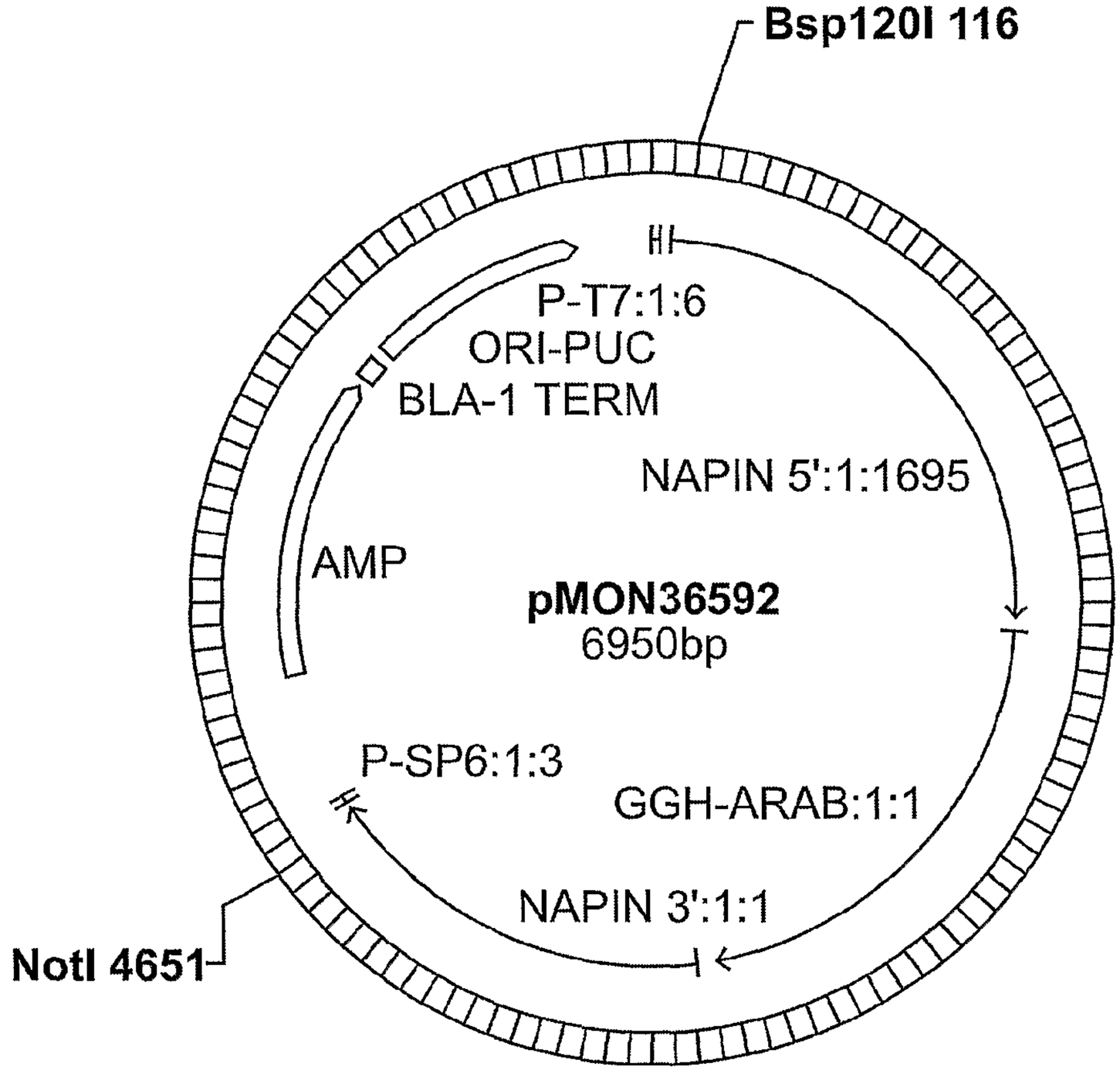

HindIII 7180
NotI 6
O-U.ori-0:1:1
NAPIN 5':1:2
pMON36591
CR-U.Amp-0:1:2
7181 bp
GGH-ARAB:1:3
NAPIN 3':1:2
HindIII 4542
NotI 4535
HindIII 3197

FIG. 61



pMON36588

FIG. 62



pMON36592

FIG. 63

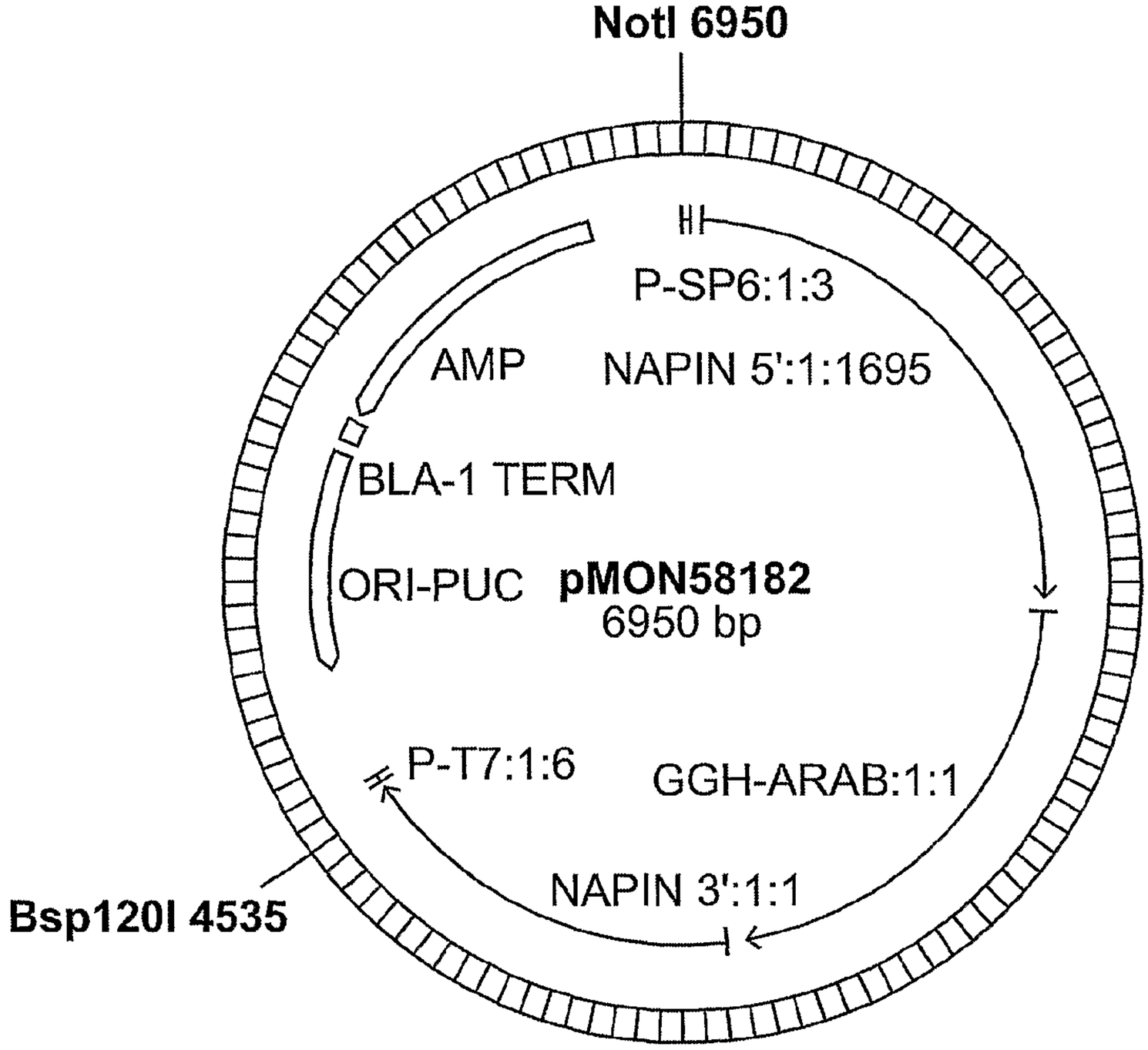

NotI 6950
P-SP6:1:3
AMP
NAPIN 5':1:1695
BLA-1 TERM
ORI-PUC
pMON58182
6950 bp
P-T7:1:6
GGH-ARAB:1:1
NAPIN 3':1:1
Bsp120I 4535

FIG. 64

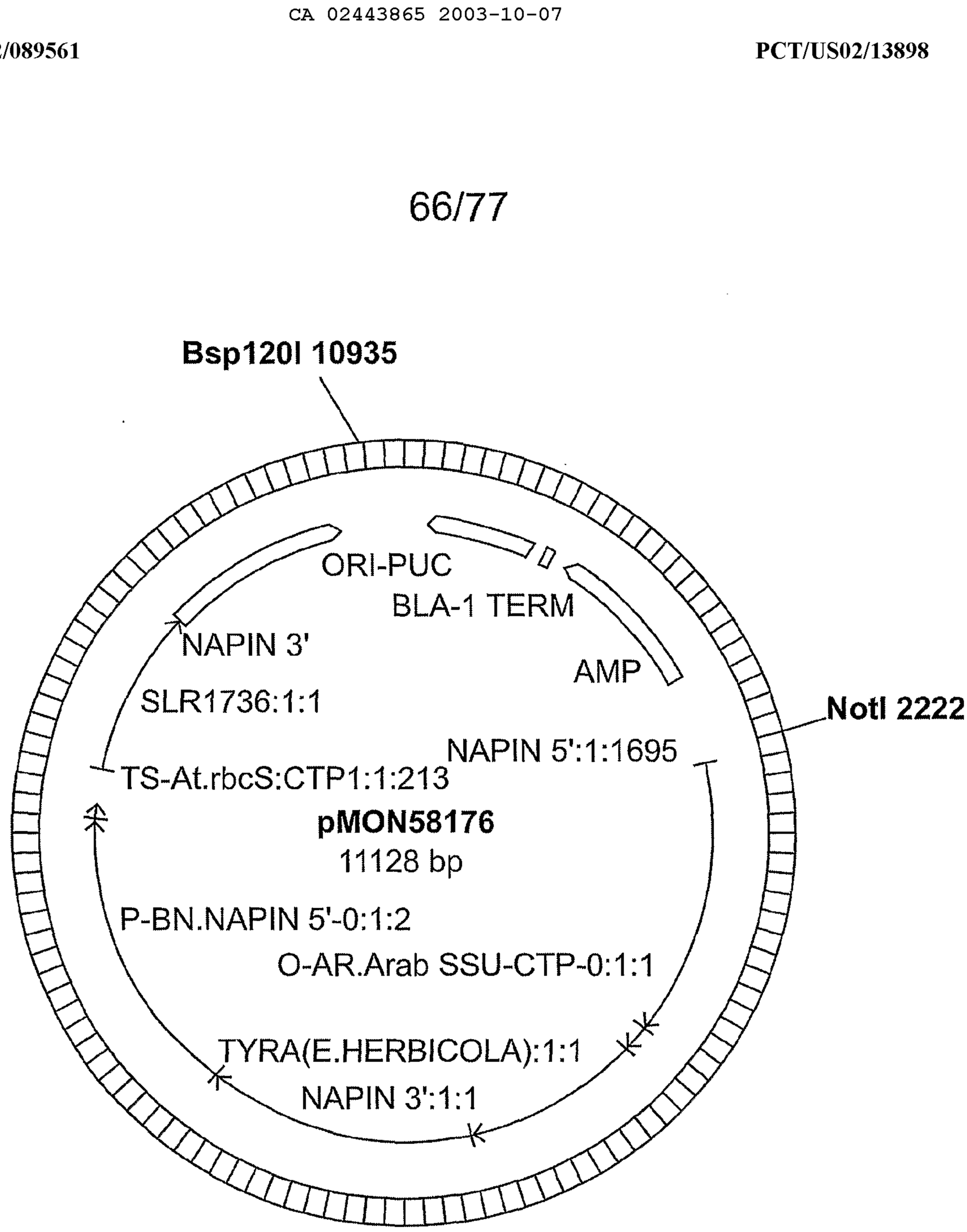

Bsp120I 10935
ORI-PUC
BLA-1 TERM
NAPIN 3'
AMP
SLR1736:1:1
NotI 2222
NAPIN 5':1:1695
TS-At.rbcS:CTP1:1:213
pMON58176
11128 bp
P-BN.NAPIN 5'-0:1:2
O-AR.Arab SSU-CTP-0:1:1
TYRA(E.HERBICOLA):1:1
NAPIN 3':1:1

FIG. 65

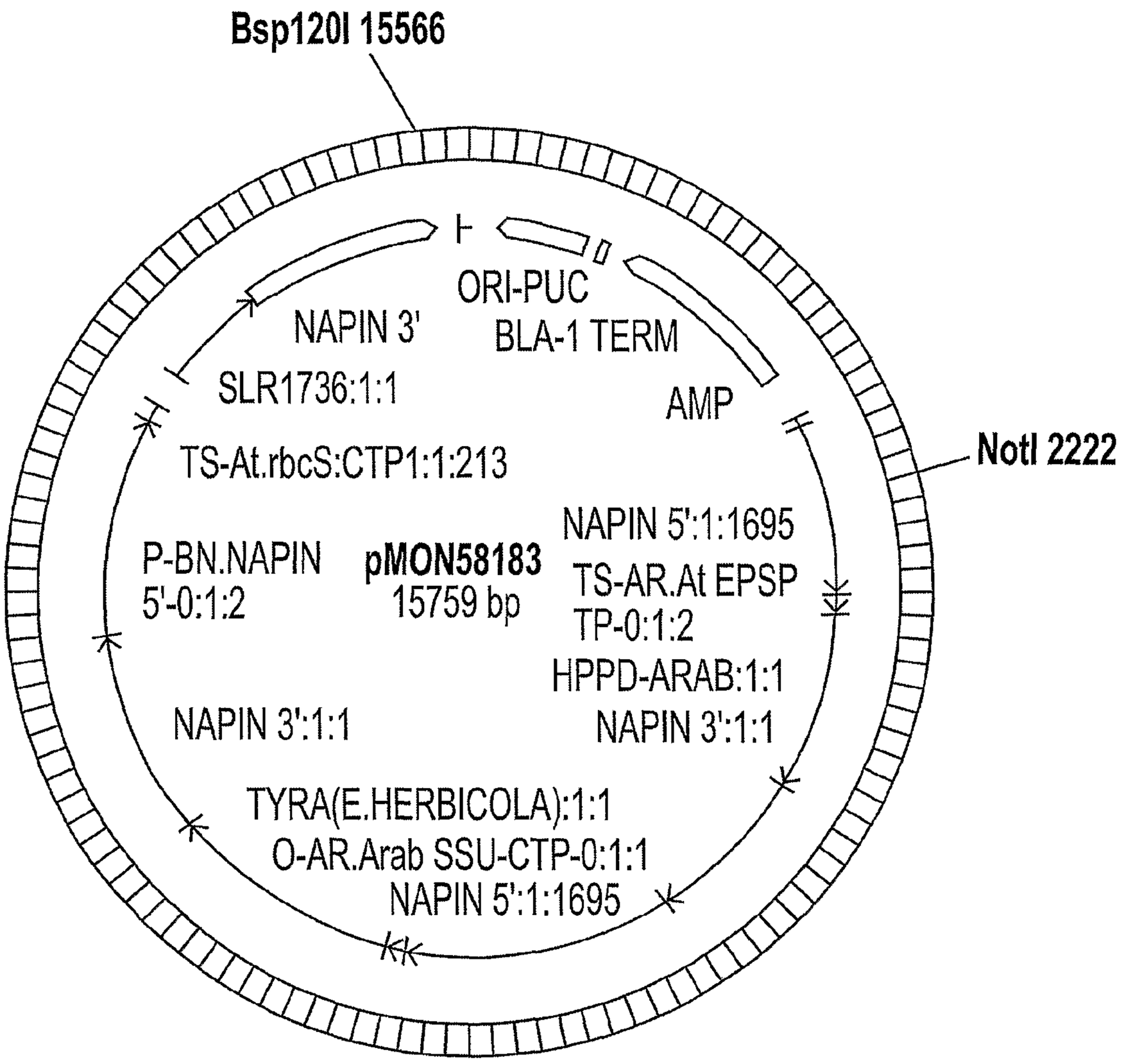

Bsp120I 15566
ORI-PUC
BLA-1 TERM
AMP
NotI 2222
NAPIN 3'
SLR1736:1:1
TS-At.rbcS:CTP1:1:213
NAPIN 5':1:1695
P-BN.NAPIN
5'-0:1:2
pMON58183
15759 bp
TS-AR.At EPSP
TP-0:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 3':1:1
TYRA(E.HERBICOLA):1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695

FIG. 66

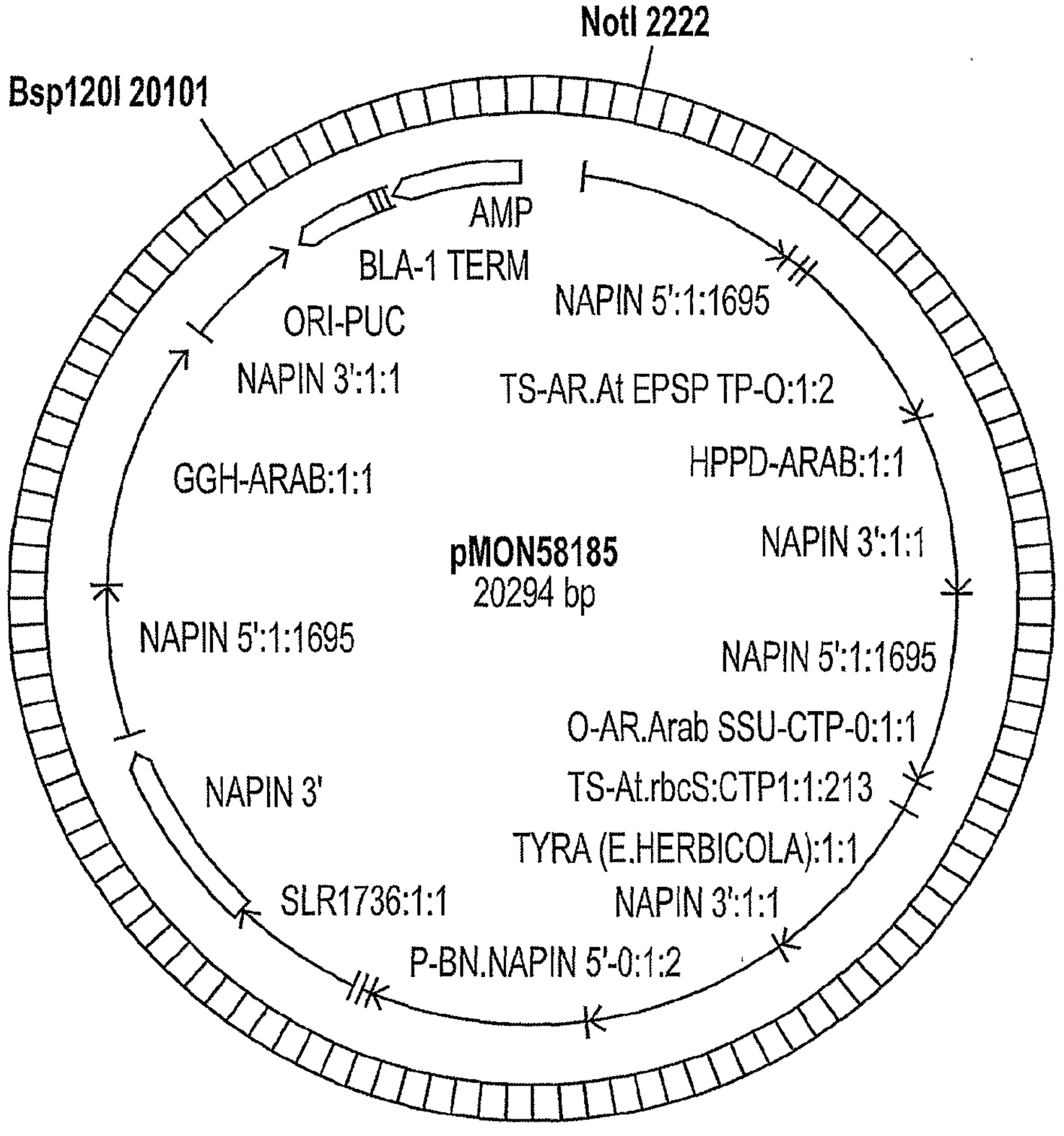

NotI 2222
Bsp120I 20101
AMP
BLA-1 TERM
ORI-PUC
NAPIN 3':1:1
GGH-ARAB:1:1
NAPIN 5':1:1695
NAPIN 3'
SLR1736:1:1
P-BN.NAPIN 5'-0:1:2
NAPIN 5':1:1695
TS-AR.At EPSP TP-O:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 5':1:1695
O-AR.Arab SSU-CTP-0:1:1
TS-At.rbcS:CTP1:1:213
TYRA (E.HERBICOLA):1:1
NAPIN 3':1:1
pMON58185
20294 bp

FIG. 67

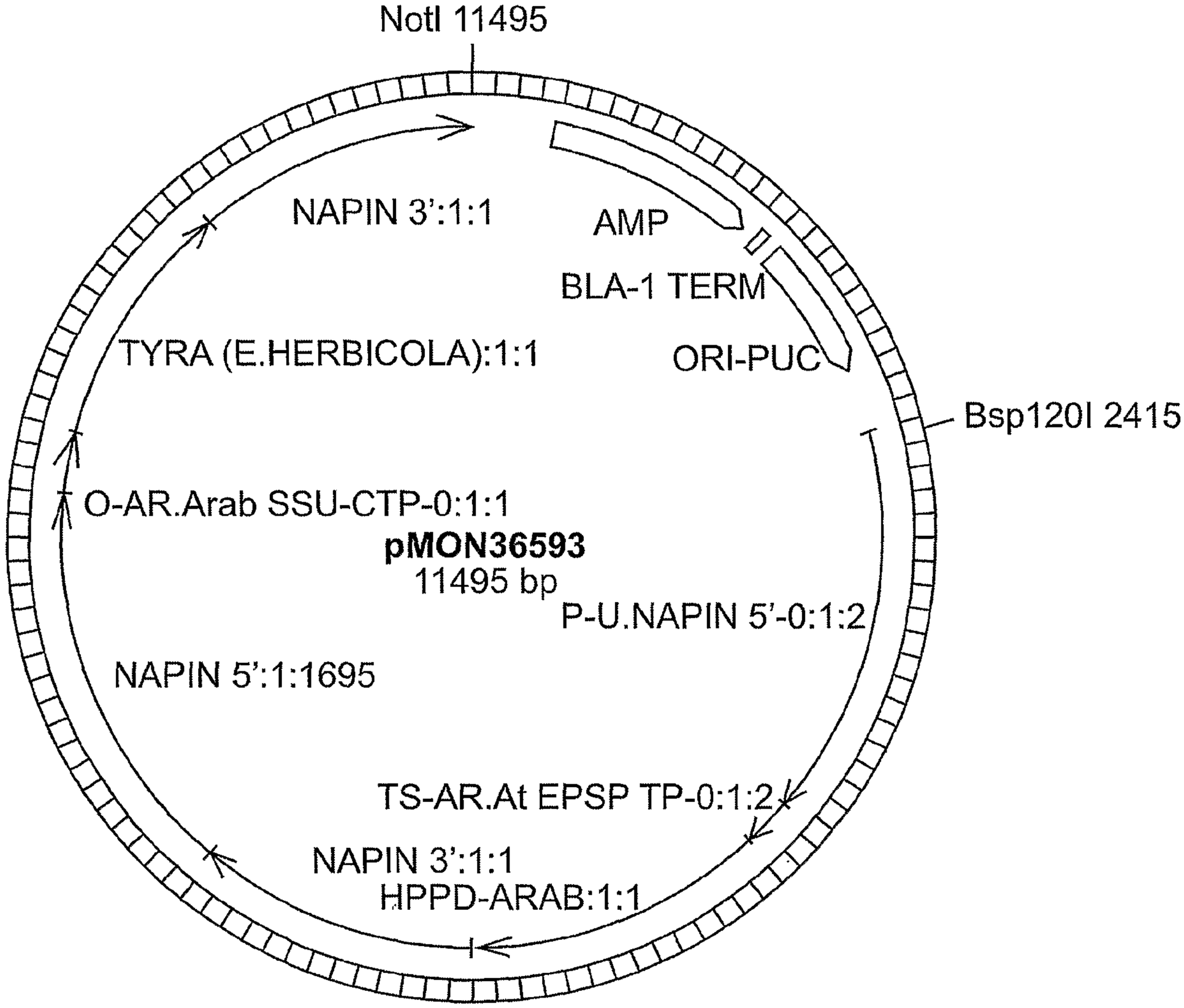

Notl 11495
NAPIN 3':1:1
AMP
BLA-1 TERM
TYRA (E.HERBICOLA):1:1
ORI-PUC
Bsp120I 2415
O-AR.Arab SSU-CTP-0:1:1
pMON36593
11495 bp
P-U.NAPIN 5'-0:1:2
NAPIN 5':1:1695
TS-AR.At EPSP TP-0:1:2
NAPIN 3':1:1
HPPD-ARAB:1:1

FIG. 68

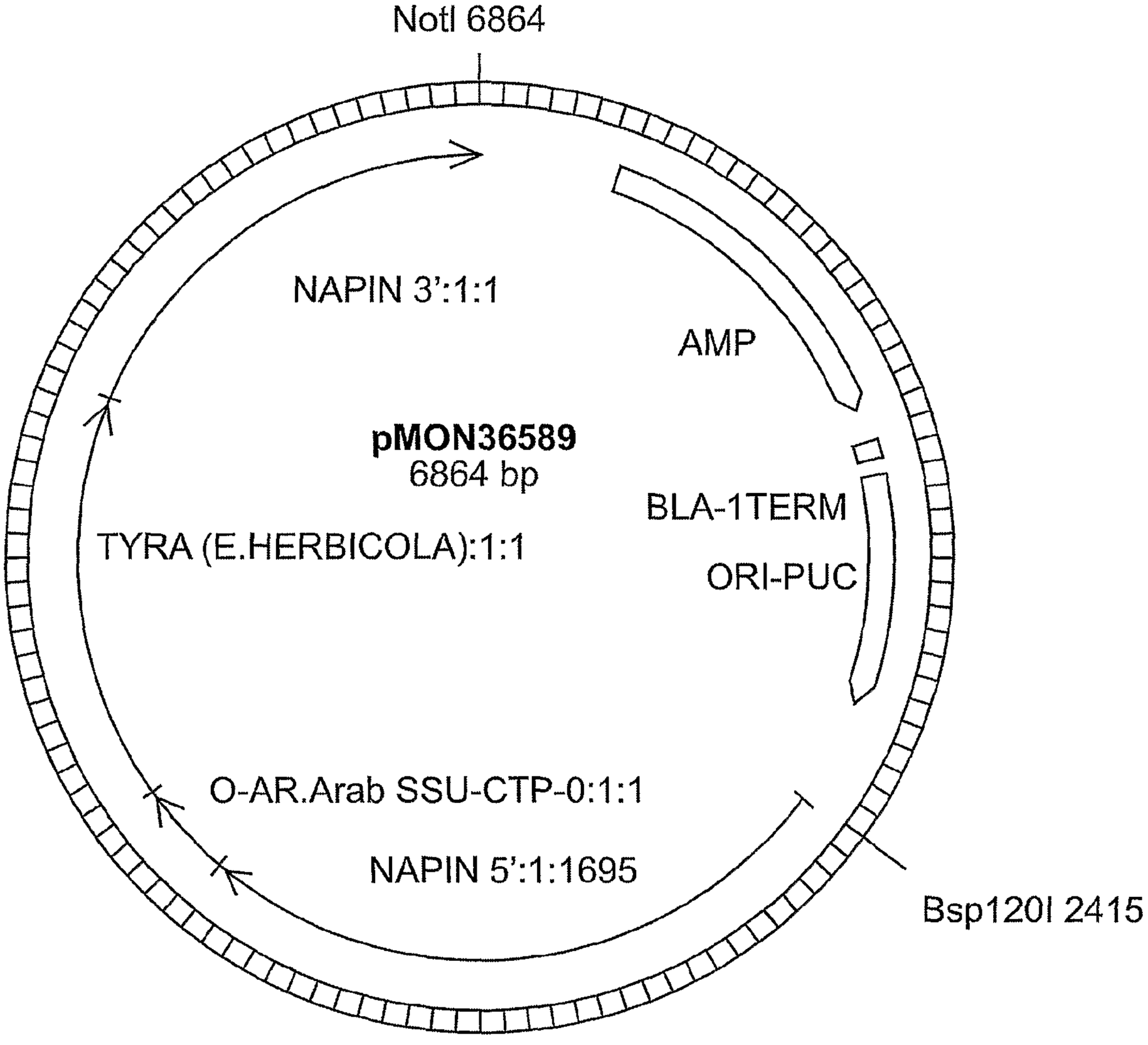

Notl 6864
NAPIN 3':1:1
AMP
pMON36589
6864 bp
BLA-1TERM
TYRA (E.HERBICOLA):1:1
ORI-PUC
O-AR.Arab SSU-CTP-0:1:1
NAPIN 5':1:1695
Bsp120I 2415

FIG. 69

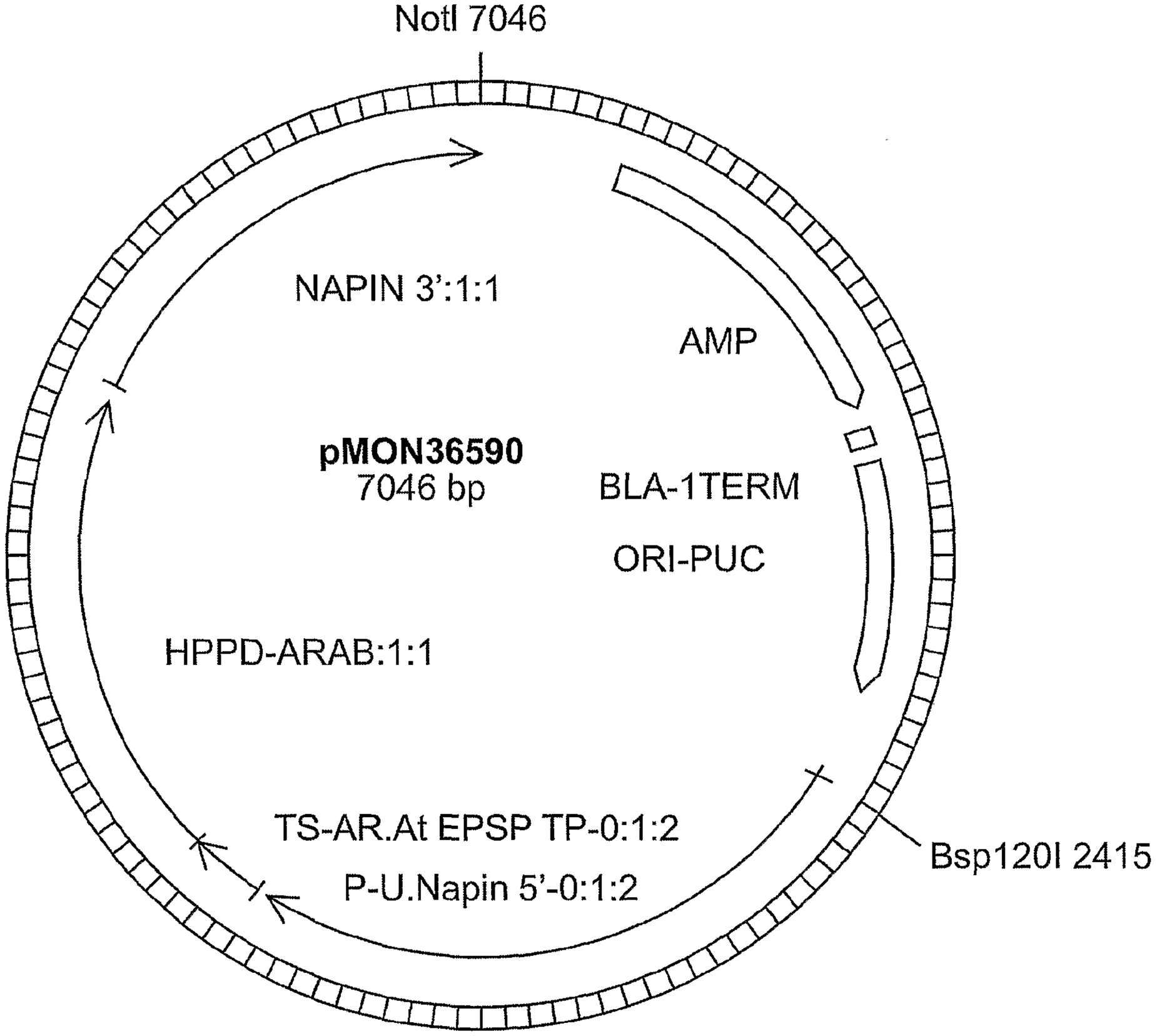

Notl 7046
NAPIN 3':1:1
AMP
pMON36590
7046 bp
BLA-1TERM
ORI-PUC
HPPD-ARAB:1:1
TS-AR.At EPSP TP-0:1:2
P-U.Napin 5'-0:1:2
Bsp120I 2415

FIG. 70

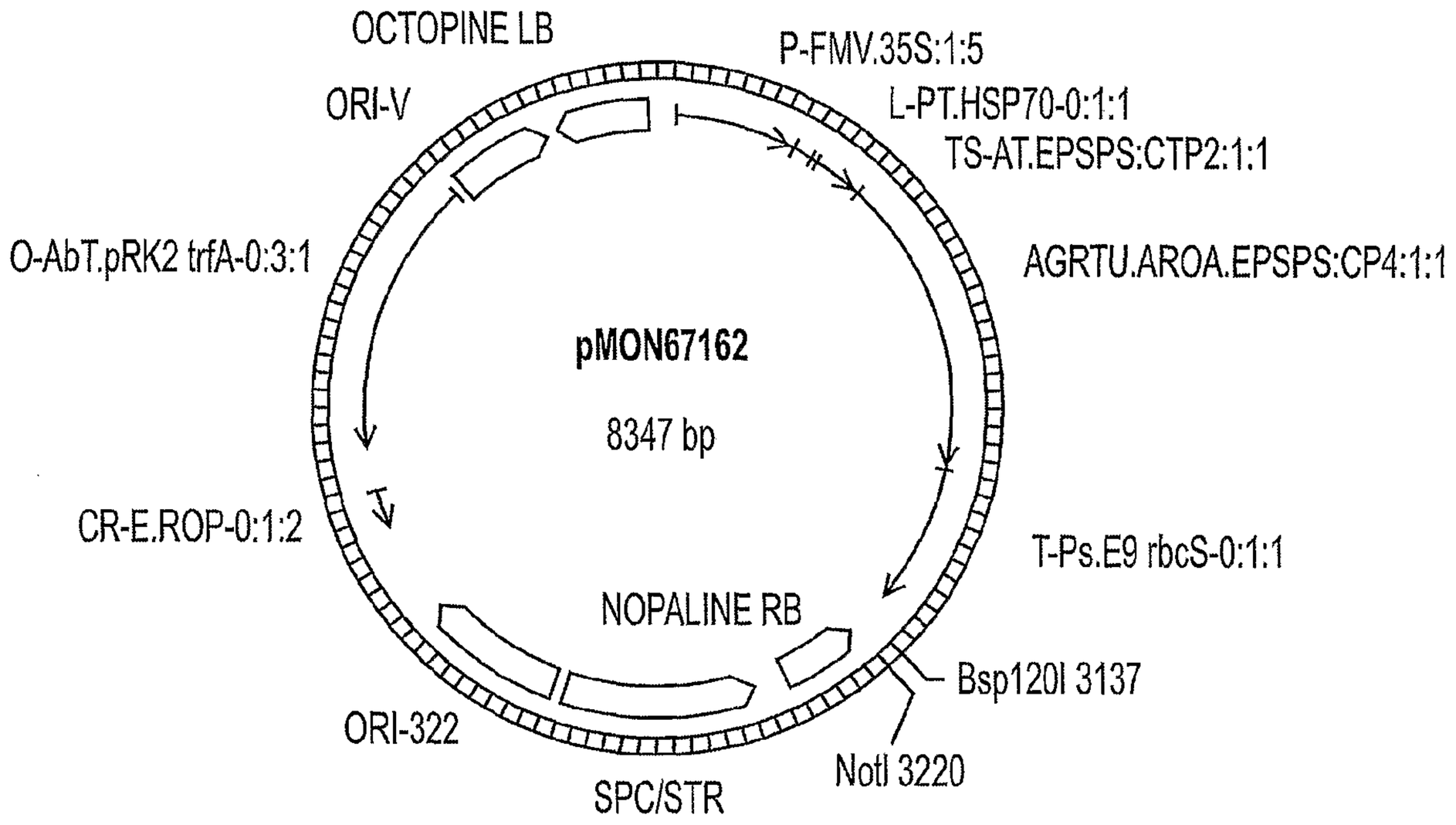

OCTOPINE LB
P-FMV.35S:1:5
ORI-V
L-PT.HSP70-0:1:1
TS-AT.EPSPS:CTP2:1:1
O-AbT.pRK2 trfA-0:3:1
AGRTU.AROA.EPSPS:CP4:1:1
pMON67162
8347 bp
CR-E.ROP-0:1:2
T-Ps.E9 rbcS-0:1:1
NOPALINE RB
Bsp120I 3137
ORI-322
NotI 3220
SPC/STR

FIG. 71

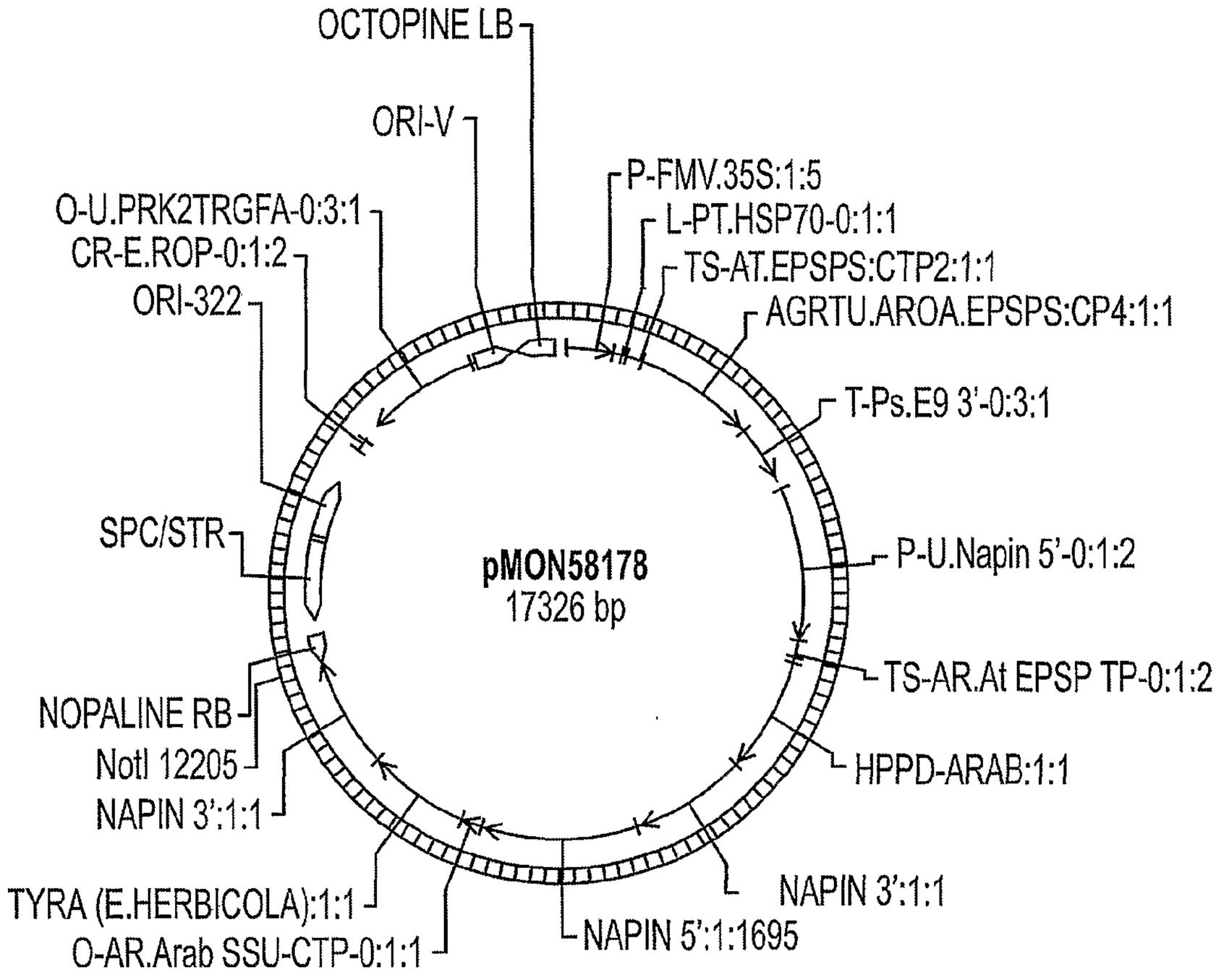

OCTOPINE LB
ORI-V
P-FMV.35S:1:5
L-PT.HSP70-0:1:1
TS-AT.EPSPS:CTP2:1:1
AGRTU.AROA.EPSPS:CP4:1:1
T-Ps.E9 3'-0:3:1
P-U.Napin 5'-0:1:2
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 5':1:1695
TYRA (E.HERBICOLA):1:1
O-AR.Arab SSU-CTP-0:1:1
NOPALINE RB
NotI 12205
NAPIN 3':1:1
SPC/STR
ORI-322
CR-E.ROP-0:1:2
O-U.PRK2TRGFA-0:3:1
pMON58178
17326 bp

FIG. 72

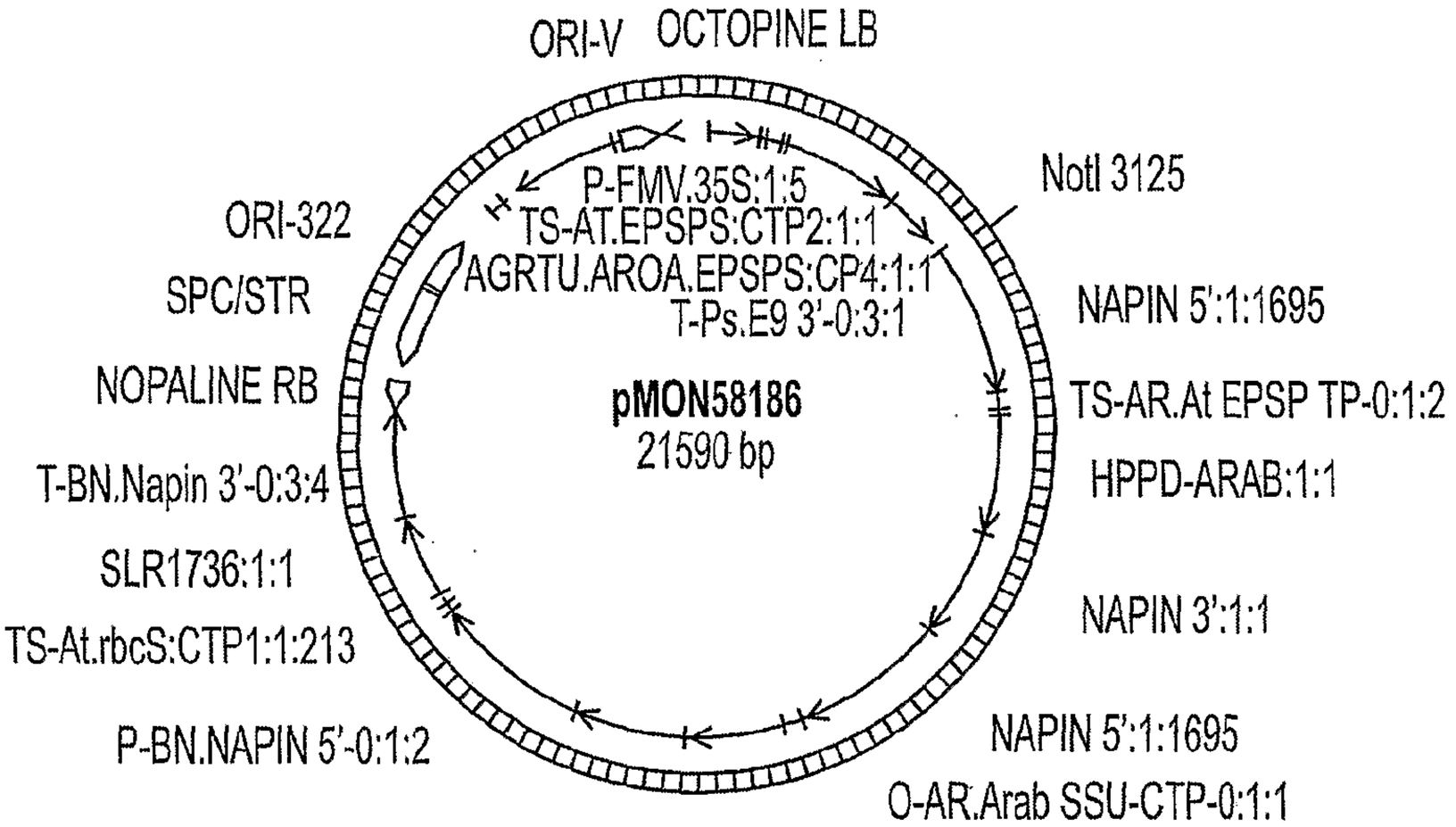

ORI-V
OCTOPINE LB
NotI 3125
P-FMV.35S:1:5
TS-AT.EPSPS:CTP2:1:1
AGRTU.AROA.EPSPS:CP4:1:1
T-Ps.E9 3'-0:3:1
ORI-322
SPC/STR
NOPALINE RB
T-BN.Napin 3'-0:3:4
SLR1736:1:1
TS-At.rbcS:CTP1:1:213
P-BN.NAPIN 5'-0:1:2
NAPIN 3':1:1
pMON58186
21590 bp
NAPIN 5':1:1695
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
NAPIN 3':1:1
NAPIN 5':1:1695
O-AR.Arab SSU-CTP-0:1:1
NAPIN 3':1:1
TYRA (E.HERBICOLA):1:1

FIG. 73

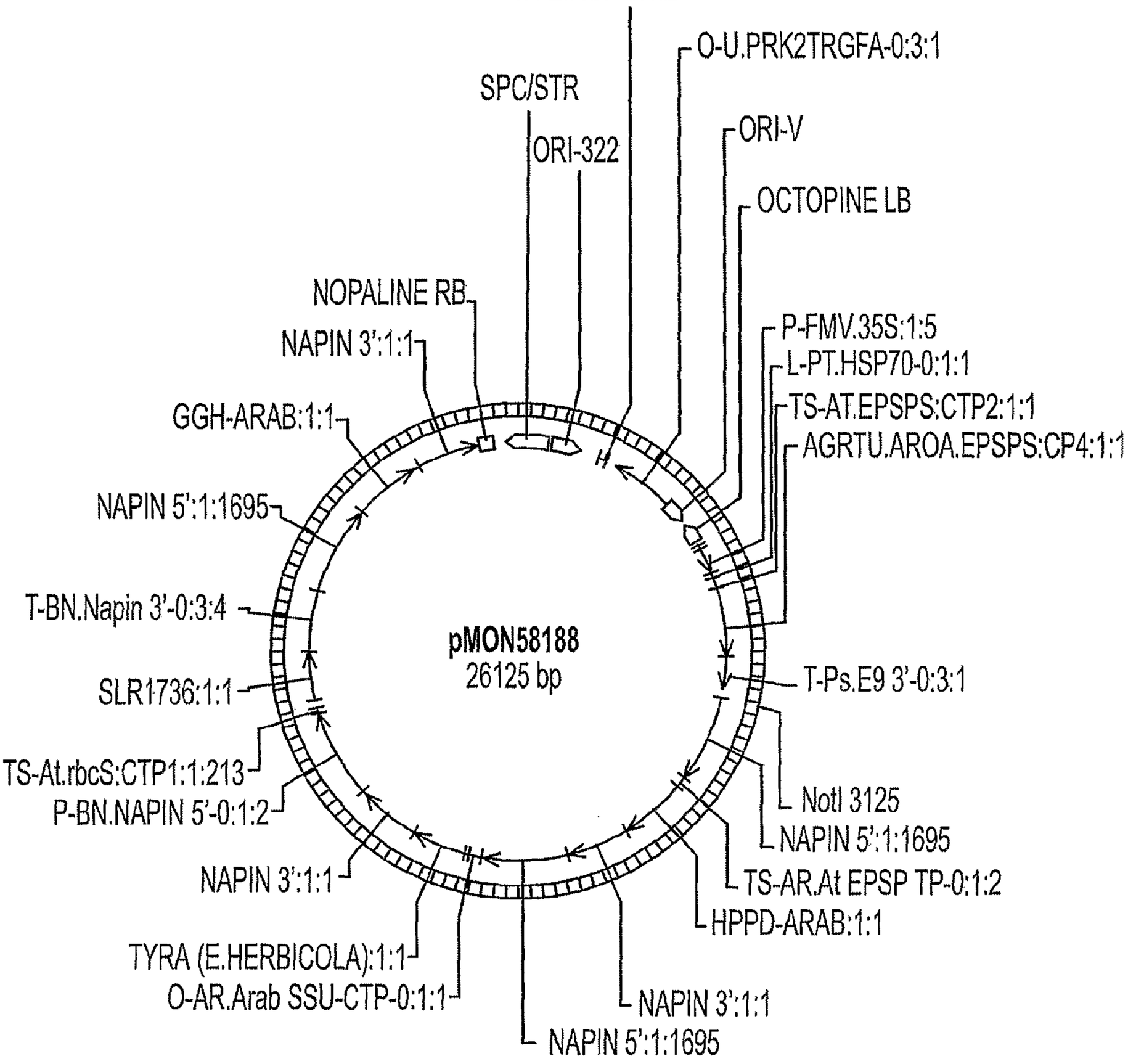

CR-E.ROP-0:1:2
O-U.PRK2TRGFA-0:3:1
SPC/STR
ORI-V
ORI-322
OCTOPINE LB
NOPALINE RB
P-FMV.35S:1:5
NAPIN 3':1:1
L-PT.HSP70-0:1:1
GGH-ARAB:1:1
TS-AT.EPSPS:CTP2:1:1
AGRTU.AROA.EPSPS:CP4:1:1
NAPIN 5':1:1695
T-BN.Napin 3'-0:3:4
pMON58188
26125 bp
T-Ps.E9 3'-0:3:1
SLR1736:1:1
NotI 3125
TS-At.rbcS:CTP1:1:213
NAPIN 5':1:1695
P-BN.NAPIN 5'-0:1:2
NAPIN 3':1:1
TS-AR.At EPSP TP-0:1:2
HPPD-ARAB:1:1
TYRA (E.HERBICOLA):1:1
O-AR.Arab SSU-CTP-0:1:1
NAPIN 3':1:1
NAPIN 5':1:1695

FIG. 74

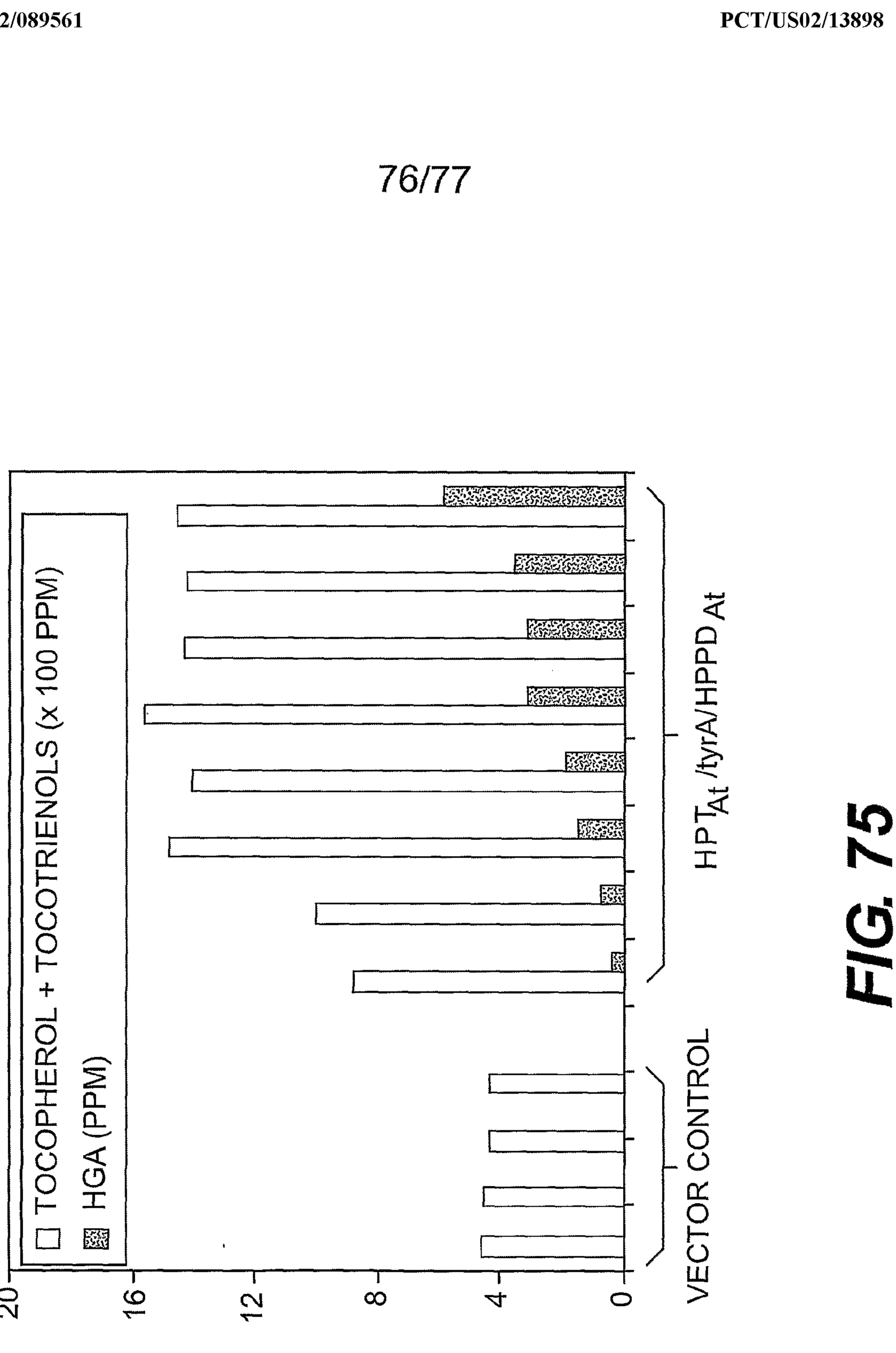

TOCOPHEROL + TOCOTRIENOLS (x 100 PPM)
HGA (PPM)
20
16
12
8
4
0
VECTOR CONTROL
HPT$_{At}$ /tyrA/HPPD At
PPM

FIG. 75

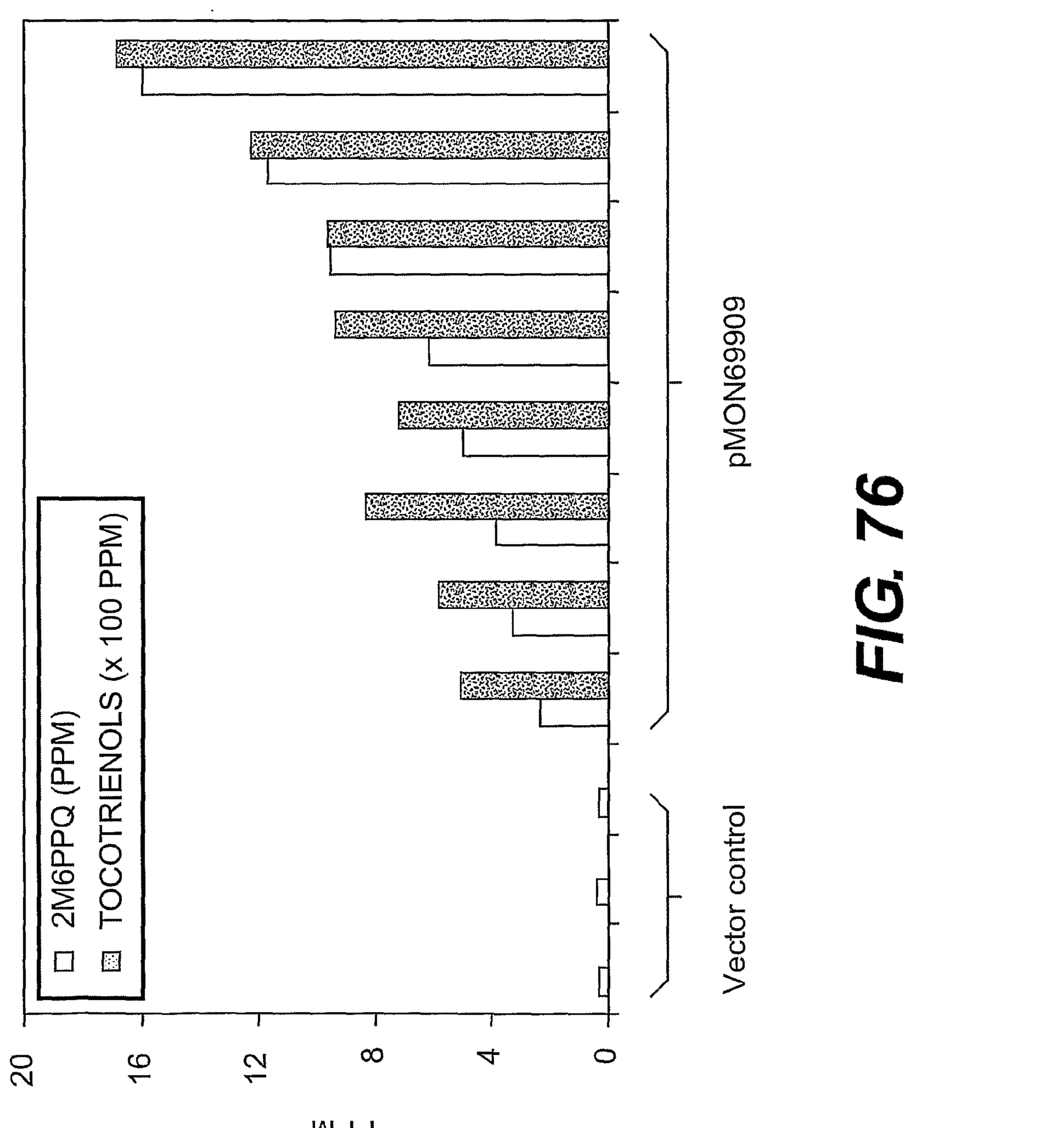

2M6PPQ (PPM)
TOCOTRIENOLS (x 100 PPM)
pMON69909
Vector control
20
16
12
8
4
0
PPM

FIG. 76