



- (72) BABYAK, Lyubov Ya., UA
- (72) BACHER, Adelbert, DE
- (72) BORETSKYY, Yuriy R., UA
- (72) DEMCHYSHYN, Vasyl V., UA
- (72) EBERHARDT, Sabine, DE
- (72) FEDOROVYCH, Dariya, UA
- (72) LÜTTGEN, Holger, DE
- (72) RICHTER, Gerald, DE
- (72) VAN LOON, Adolphus, CH
- (71) F.HOFFMANN-LA ROCHE AG, CH
- (51) Int.Cl.⁶ C12N 1/19, A23L 1/302, A23K 1/16, C12P 25/00
- (30) 1998/04/23 (98107380.2) EP
- (54) **SURPRODUCTION DE RIBOFLAVINE DANS LA LEVURE**
- (54) **OVERPRODUCTION OF RIBOFLAVIN IN YEAST**

(57) Recombinant strains of *Bacillus subtilis* for the production of riboflavin by fermentation have been described, e.g. in EP 405 370. These strains carry the riboflavin operon under the control of a strong promoter directing the production of the cognate enzymes in large amount. The gene constructs of the riboflavin operon under the control of a strong promoter can be present at one or several different locations on the *B. subtilis* chromosome. The incorporation of an additional gene of the riboflavin pathway under the control of a strong promoter at a separate locus on the *B. subtilis* chromosome has also been shown to increase the yield of riboflavin obtained by fermentation, see EP 821 063. Thus the present specification provides a yeast strain which has been transformed by a recombinant DNA sequence comprising a DNA sequence which upon expression in a suitable host cell encodes at least one polypeptide with riboflavin biosynthetic activity and which DNA sequence is transcriptionally linked to a promotor functional in such yeast strain. The yeast strain may belong to the group of flavinogenic yeasts which overproduce riboflavin under conditions of iron starvation. The specification discloses the following groups: *Schwanniomyces*, preferably *Schwanniomyces occidentalis*, *Debaryomyces*, preferably *Debaryomyces hansenii*, *Torulopsis*, preferably *Torulopsis candida*, or *Candida*, preferably *Candida guilliermondii* or *Candida famata*.

Abstract

Recombinant strains of *Bacillus subtilis* for the production of riboflavin by fermentation have been described, e.g. in EP 405 370. These strains carry the riboflavin operon under the control of a strong promoter directing the production of the cognate enzymes in large amount. The gene constructs of the riboflavin operon under the control of a strong promoter can be present at one or several different locations on the *B. subtilis* chromosome. The incorporation of an additional gene of the riboflavin pathway under the control of a strong promoter at a separate locus on the *B. subtilis* chromosome has also been shown to increase the yield of riboflavin obtained by fermentation, see EP 821 063. Thus the present specification provides a yeast strain which has been transformed by a recombinant DNA sequence comprising a DNA sequence which upon expression in a suitable host cell encodes at least one polypeptide with riboflavin biosynthetic activity and which DNA sequence is transcriptionally linked to a promoter functional in such yeast strain. The yeast strain may belong to the group of flavinogenic yeasts which overproduce riboflavin under conditions of iron starvation. The specification discloses the following groups: Schwanniomyces, preferably Schwanniomyces occidentalis, Debaryomyces, preferably Debaryomyces hansenii, Torulopsis, preferably Torulopsis candida, or Candida, preferably Candida guilliermondii or Candida famata.

Ref. 20'071

Derivatives of riboflavin (the flavocoenzymes FMN and FAD) are universally required for redox reactions in all cellular organisms. Riboflavin (vitamin B₂) is produced by all plants and by many microorganisms [Demain
5 A.L. Riboflavin oversynthesis. Ann. Rev. Microbiol. 1972, 26, 369]. The compound is not produced in vertebrates. Riboflavin is therefore an essential nutrient for man and animals.

Riboflavin can be produced by chemical synthesis and by various fermentation procedures using strains of *Bacillus* (e.g. *Bacillus subtilis*), the
10 ascomycetes *Ashbya gossypii* and *Eremothecium ashbyi* [Demain A.L. Riboflavin oversynthesis. Ann. Rev. Microbiol. 1972, 26, 369 and Mitsuda H, Nakajima K., Effects of 8-azaguanine on riboflavin production and on the nucleotide pools in non-growing cells of *Eremothecium ashbyi*. J. Nutr Sci Vitaminol (Tokyo) 1973; 19(3):215-227], various yeast strains such as *Candida*
15 *guilliermondii*, *Candida famata* [F.W. Tanner, Jr.C. Vojnovich, J.M. Van Lanen. Riboflavin production by *Candida* species. Nature, 1945, 101 (2616):180-181] and related strains, as well as other microorganisms.

The pathway of riboflavin biosynthesis in yeast is shown in Fig. 1. The precursors for the biosynthesis of the vitamin are guanosine triphosphate
20 (GTP) and ribulose 5-phosphate. One mol of GTP and two mol of ribulose 5-phosphate are required to biosynthetically generate one mol of riboflavin.

In the yeast, *Saccharomyces cerevisiae*, the biosynthesis of the vitamin requires at least six genes, specifically the genes *RIB1*, *RIB2*, *RIB3*, *RIB4*, *RIB5* and *RIB7* [Oltmanns O., Bacher A., Lingens F. and Zimmermann F.K.

Biochemical and genetic classification of riboflavine deficient mutants of *Saccharomyces cerevisiae*. Mol. Gen. Genet. 1969, 105, 306]. In *C. guilliermondii*, the biosynthesis of riboflavin has also been shown to require the products of at least six genes, specifically the genes *RIB1*, *RIB2*, *RIB3*, *RIB4*,
5 *RIB5* and *RIB6* (2). The enzymes specified by these *C. guilliermondii* genes and their roles in the biosynthetic pathway are summarized in Fig. 1. In contrast to the situation in *B. subtilis*, the riboflavin biosynthetic genes are not clustered in the eukaryots *S. cerevisiae* and *C. guilliermondii*.

The initial step in the biosynthetic pathway is the opening of the
10 imidazole ring of GTP catalyzed by the enzyme, GTP cyclohydrolase II. The product of this enzyme has been reported to be 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate. This intermediate is converted to 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione by a sequence of side chain reduction, ring deamination, and dephosphorylation. The hypothetical enzyme
15 involved in the dephosphorylation of 5-amino-6-ribitylamino 5'-phosphate is still unknown. The conversion of 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione to 6,7-dimethyl-8-ribityllumazine by the enzyme, 6,7-dimethyl-8-ribityllumazine synthase, requires a second substrate, 3,4-dihydroxy-2-butanone 4-phosphate, which is obtained from ribulose 5-
20 phosphate by the catalytic action of 3,4-dihydroxy-2-butanone-4-phosphate synthase. Finally, 6,7-dimethyl-8-ribityllumazine is converted to riboflavin by a dismutation reaction catalyzed by riboflavin synthase. The sequence of the *RIB1* gene directing the synthesis of GTP cyclohydrolase II, the initial enzyme of the riboflavin pathway, has been established in the yeast, *C. guilliermondii*
25 (4).

Recombinant strains of *Bacillus subtilis* for the production of riboflavin by fermentation have been described, e.g. in EP 405 370. These strains carry the riboflavin operon under the control of a strong promoter directing the production of the cognate enzymes in large amount. The gene constructs of the
30 riboflavin operon under the control of a strong promoter can be present at one or several different locations on the *B. subtilis* chromosome. The incorporation of an additional gene of the riboflavin pathway under the control of a strong promoter at a separate locus on the *B. subtilis* chromosome has also been shown to increase the yield of riboflavin obtained by fermentation, see EP 821
35 063.

Whereas the production of riboflavin by strains of yeasts such as *C. guilliermondii* has been reported, recombinant DNA technology has not been

- 3 -

applied for the overexpression of riboflavin biosynthetic genes in *C. guilliermondii* or in related flavinogenic yeasts so far.

It is therefore an object of the present invention to provide recombinant means which should allow the production of yeast strains which overproduce
5 riboflavin. More specifically it is an object of the present invention to provide a yeast strain which has been transformed by a recombinant DNA sequence comprising a DNA sequence which upon expression in a suitable host cell encodes at least one polypeptide with riboflavin biosynthetic activity and which DNA sequence is transcriptionally linked to a promotor functional in
10 such yeast strain and even more specifically such a yeast strain which belongs to the groups of flavinogenic yeasts which overproduce riboflavin under conditions of iron starvation like a yeast strain which is selected from the following group: Schwanniomyces, preferably Schwanniomyces occidentalis, Debaryomyces, preferably Debaryomyces kloeckeri, Torulopsis, preferably
15 Torulopsis candida, or, especially Candida, preferably Candida guilliermondii or Candida famata (Logvinenko et al., Ukrainskii Biokhimicheskii Zhurnal 61(1), 28-32, 1989; Logvinenko et al., Mikrobiologiya 57(2), 181-186, 1988 and Nakase and Suzuki, Journal of General and Appl. Mikrobiology 31(1), 49-70 (1985). It is furthermore an object of the present invention into provide such
20 yeast strains wherein the polypeptide encoding DNA sequence is from yeast, preferably flavinogenic yeasts which overproduce riboflavin under conditions of iron starvation, more preferably Candida, e.g. Candida guilliermondii or Candida famata.

It is also an object of the present invention to provide such yeast strains
25 wherein the polypeptide encoding DNA sequence encodes a protein with GTP cyclohydrolase II activity and is selected from the following DNA sequences:

- a) the DNA sequence as shown in Fig. 5 or its complementary strand;
- b) DNA sequences which hybridize under standard conditions to the protein coding regions of the DNA sequences defined in (a) or fragments
30 thereof; and
- c) DNA sequences which, but for the degeneracy of the genetic code, would hybridize to the DNA sequences defined in (a) and (b).

It is furthermore an object of the present invention to provide such yeast strains wherein the promotor is the TEF *S. cerevisiae* promotor.

- 4 -

It is also an object of the present invention to provide a process for the production of riboflavin characterized therein that a yeast strain as described above is cultured under suitable culture conditions and the riboflavin produced is isolated from the medium or the yeast strain by methods known to the man skilled in the art, and a process for the production of a food or feed composition characterized therein that riboflavin as obtained by such process is mixed with one or more suitable food or feed ingredients by a process known to the man skilled in the art.

Drawings

Figure 1 shows the pathway of riboflavin biosynthesis in yeast.

Figure 2 is a nucleotide sequence of the p19R1 insert.

Figure 3 is the nucleotide sequence of the pTC2 insert.

Figure 4 is a nucleotide sequence of the pTCdXS-2 insert.

Figure 5 is a nucleotide sequence of PCR product obtained with primers ShBle_V and PGgtpCY_nco using boiled cells of *P. guilliermondii* XS-3 recombinant strain.

Figure 6 is a scheme of the 2-step PCR amplification.

All *C. guilliermondii* strains used in the practice of the present invention are derivatives of the *C. guilliermondii* strain obtained from the American Type Culture Collection (ATCC) under accession No. ATCC 9058 (1). *Candida guilliermondii* (ATCC 9058) has been redeposited as a Budapest Treaty deposit on April 1, 1998 and has been assigned accession No. ATCC 74437. *Candida guilliermondii* is a representative of yeast species which overproduce riboflavin (vitamin B2) under conditions of iron starvation. The group includes also *Schwanniomyces occidentalis*, (or called *Debaryomyces occidentalis*) *Debaryomyces cloeckeri*, *Torulopsis candida* and *Candida famata*. The latter species is used for industrial production of riboflavin. Regarding the taxonomic assignments of yeast species a man skilled in the art knows that these assignments are handled variably by different authors, for example: *Candida famata* - *Debaryomyces hansenii* - *Torulaspora hansenii*, *Candida guilliermondii* - *Pichia guilliermondii* - *Yamadazyma guilliermondii* are used as synonyms. A man skilled in the art knows that microorganisms which can be used for the practice of the present invention, either as host cells or source for the isolation of DNA sequences, are available from depository authorities, e.g. the American Type Culture Collection (ATCC), the Centraalbureau voor Schimmelcultures (CBS) or the Deutsche Sammlung für Mikroorganismen und Zellkulturen GmbH (DSM) or any other depository authority as listed in the Journal "Industrial Property" [(1991) 1, pages 29-40].

4A

5

DNA sequences useful for the practice of the present invention and encoding a polypeptide with riboflavin biosynthetic activity can be obtained from any microorganism known to produce riboflavin (see above) in form of e.g., genomic or c-DNA sequences by methods known to the man skilled in the art or by using the wellknown PCR-Technology. The principles of the polymerase chain reaction (PCR) methode are outlined e.g. by White et al., Trends in Genetics, 5, 185-189 (1989), whereas improved methods are

described, e.g. in Innis et al. [PCR Protocols: A guide to Methods and Applications, Academic Press, Inc. (1990)].

The sequence information needed for the design of the PCR-primers can be obtained from any sequence data base, for example like Genbank
 5 (Intelligenetics, California, USA), European Bioinformatics Institute (Hinton Hall, Cambridge, GB), NBRF (Georgetown University, Medical Centre, Washington DC, USA) and Vecbase (University of Wisconsin, Biotechnology Centre, Madison, Wisconsin, USA).

Once such DNA sequences have been obtained they can be expressed in
 10 any desirable host and the riboflavin biosynthetic activity of the encoded polypeptide can be determined by any assay known to the man skilled in the art and described, e.g. in Bacher A., G. Richter, H. Ritz, S. Eberhardt, M. Fisher and C. Krieger, Biosynthesis of riboflavin: GTP cyclohydrolase II, deaminase, and reductase. Methods in enzymology 1997; 280: 382-389; K. Kis,
 15 R. Volk and A. Bacher, Biosynthesis of riboflavin. Studies on the reaction mechanism of 6,7-dimethyl-8-ribityllumazine synthase. Biochemistry 1995, 34, 2883-2892; Logvinenko EM, Shavlovskii GM, Zakal'skii AE, Kontorovskaia Niu. Properties of 2,5-diamino-4-oxy-6-ribosylaminopyrimidine-5'-phosphate reductase, a enzyme of the second stage of flavinogenesis in *Pichia*
 20 *guilliermondii* yeast Ukr Biokhim Zh 1989 Jul; 61(4): 47-54; G. Richter, M. Fischer, C. Krieger, S. Eberhardt, H. Lüttgen, I. Gerstenschläger and A. Bacher. Biosynthesis of riboflavin. Characterization of the bifunctional deaminase/reductase of *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. 1997, 179, 2022-2028; K. Ritsert, D. Turk, R. Huber, R. Ladenstein, K.
 25 Schmidt-Bäse and A. Bacher. Studies on the lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*. Crystal structure analysis of reconstituted icosahedral β subunit capsid at 2.4 Å resolution. J. Mol. Biol. 1995, 253, 151-167.

The DNA sequences used for the practice of the present invention
 30 comprise at least one DNA sequence which encodes a polypeptide with riboflavin biosynthetic activity. It is however, understood by the man skilled in the art that also more than one, e.g. all enzymes of the riboflavin biosynthetic pathway can be encoded by such DNA sequences and one or more of these enzymes can be encoded by DNA sequences of different species origin or
 35 can be of partial or total synthetic origin as long as they show at least one desired riboflavin biosynthetic activity. One example of such a DNA sequence is given in Fig. 5 coding for a GTP cyclohydrolase II. However, DNA sequences

- 6 -

which hybridize under standard conditions to this DNA sequence and encode such a GTP cyclohydrolase are also useful for the practice of the present invention.

"Standard conditions" for hybridization in this context are conditions
5 which are generally used by a man skilled in the art to detect specific hybridization signals and which are described, e.g. by Sambrook et al., "Molecular Cloning" second edition, Cold Spring Harbor Laboratory Press 1989, New York, or preferably so called stringent hybridization and non-stringent washing conditions or more preferably so called stringent
10 hybridization and stringent washing conditions a man skilled in the art is familiar with and which are described, e.g. in Sambrook et al. (s.a.).
"Fragment of the DNA sequences" means in this context a fragment which codes for a polypeptide still having the enzymatic activity as specified above.

For the overexpression of the proteins encoded by the DNA sequences of
15 the present invention these sequences can be linked to promoters which are functional in the desired yeast and are, e.g. the *S.cerevisiae* TEF-promotor (see Example 2) or the *pho5*-promotor [Vogel et al., Mol. Cell. Biol., 2050-2057 (1989); Rudolf and Hinnen, Proc, Natl. Acad. Sci. 84, 1340-1344 (1987)] or the gap-promotor or the *aox1*-promotor [Koutz et al., Yeast 5, 167-177 (1989);
20 Sreekrishna et al., J. Basic Microbiol. 28, 265-278 (1988)] or the FMD promoter [Hollenberg et al., EP 299108] or MOS-promotor [Ledeboer et al., Nucleic Acids Res. 13, 3063-3082 (1985)].

The DNA sequences useful for the practice of the present invention can also comprise so called "ARS" elements (autonomously replicating sequence)
25 as described, e.g. in Example 1.

Examples

If not specifically indicated or referred to by references standard procedures have been used as described, e.g. in Sambrook et al. "Molecular Cloning", (s.a.) and Cregg, J.M., K.J. Barriner, A.Y.Hessler, and K.R. Madden
30 (1985). *Pichia pastoris* as a host system for transformations. Mol. Cell. Biol. 5, 3376-3385.

Example 1Autonomous replication of plasmid p19R1 in *C. guilliermondii*

The plasmid pFR1 carrying the *RIB1* gene of *C. guilliermondii* has been described (Zakalsky et al. *Genetika* 26, 614-620, 1990). In order to subclone the *RIB1* gene, plasmid pFR1 was digested with the restriction nuclease *SalI*. The resulting fragments were cloned into the *SalI* site of the pUC19 vector. The ligation mixture was transformed into the *E. coli* mutant strain BSV821 carrying a mutation of the *ribA* gene conducting to riboflavin deficiency. Colonies growing in the absence of riboflavin were isolated and were shown to contain a plasmid p19R1.

The plasmid p19R1 was sequenced and was shown to contain a 2.18 kb fragment of *C. guilliermondii* DNA in the *SalI* site of the pUC19 vector. The sequence of this insert is shown in Fig. 2.

The DNA sequence shown in Fig. 2 carries the *RIB1* gene of *C. guilliermondii*. The plasmid transforms *C. guilliermondii* mutant defective in the *RIB1* gene to riboflavin prototrophy and can replicate autonomously in this yeast species. The replication was shown to be due to the presence of an autonomously replicating sequence (ARS) comprising approximately base pairs 1542 to 1755 in Fig. 2 and extending into the structural gene *RIB1*.

Example 2Construction of a plasmid for hyperexpression of the *RIB1* gene of *C. guilliermondii*

The *TEF* gene of the yeast, *S. cerevisiae*, specifies the translation elongation factor 1-alpha. This gene is known to be transcribed in *S. cerevisiae* with high efficiency.

A DNA fragment carrying the *S. cerevisiae* *TEF* promoter and the 5' part of the *RIB1* gene of *C. guilliermondii* was obtained by PCR amplification. Initially, a DNA sequence located upstream from the 5' end of the *S. cerevisiae* *TEF* gene was amplified by PCR with primers ShBle_V and TEF1_H using chromosomal DNA of *S. cerevisiae* as a template. The amplified DNA fragment (subsequently designated *TEF* promoter) comprises bp 15 984 to 16 344 of the sequence listed under EMBL accession number gb/U51033/YSCP9513.

Independently, a DNA fragment comprising the 5'-terminal part of *C. guilliermondii* GTP cyclohydrolase II structural gene was obtained by PCR with primers PGgtpCY_V and PGgtpCY_nco using p19R1 plasmid as a template. The amplified DNA fragment (subsequently designated 5'GTPcII)
5 comprises bp 460 to 1145 of the sequence shown in Fig. 2.

The DNA amplicates obtained in the two PCR reactions described above comprising parts of the *TEF* gene of *S. cerevisiae* and of the 5'GTPcII were mixed, and a third PCR amplification was performed using the primers ShBle_V and PGgtpCY_nco. This reaction yielded a DNA fragment which
10 contains the complete *TEF* promoter and the 5' part of the *C. guilliermondii* *RIB1* gene.

All primers used in these PCRs are shown in Table 2. The sequence of the final amplicate is shown in Fig. 3.

The amplicate contains a cutting site for the restriction nucleases *Sph*I
15 which had been introduced via the primer ShBle_V. The amplicate also contains an *Msc*II site which is a feature of the *RIB1* gene. The amplicate was digested with *Sph*I and *Msc*I.

The plasmid p19R1 (whose construction is described above) was also digested with the same enzymes, and the PCR amplicate was ligated into the
20 digested plasmid.

The ligation mixture was transformed into a mutant designated Rib7 of *Escherichia coli* carrying a mutation of the *ribA* gene which specifies GTP cyclohydrolase II. Transformation was performed by electroporation according to the protocol of Invitrogen (5).

25 The *E. coli* cells were plated on Luria-Broth plates supplemented with ampicillin (100 mg/ml) which did not contain riboflavin. Colonies growing on this medium were isolated and were shown to contain a plasmid designated pTC2.

The sequence of the insert of the plasmid pTC2 is shown in Fig. 3.

30 Plasmid pTC2 was digested with *Xho*I and *Sal*I restriction endonucleases yielding 4.4 kb and 0.5 kb fragments. The 4.4 kb fragment was circularized with T4 DNA ligase. The ligation mixture was transformed into the Rib7 mutant of *E. coli* carrying the *RibA* mutation. Transformation was achieved by

electroporation. The cells were plated on Luria-Broth plates containing ampicillin and no riboflavin (see above). Colonies growing on these plates were isolated and were shown to contain a plasmid pTCdXS2. The procedure had resulted in the removal of 0.5 kb base pairs from the plasmid pTC2.

- 5 The sequence of the insert of the plasmid pTCdXS2 is shown in Fig. 4. The open reading frame of the *RIB1* gene of *C. guilliermondii* is indicated. The DNA segment representing the promoter of the *TEF* gene of *S. cerevisiae* is also shown.

Example 3

10 Construction of a recombinant *C. guilliermondii* strains

The riboflavin deficient mutant rh-21 with an apparent defect of the *RIB1* gene specifying GTP cyclohydrolase II (2) has been obtained after chemical mutagenesis of the L2 strain (2) which was previously obtained from the ATCC 9058 *C. guilliermondii* strain.

- 15 The plasmid pTCdXS2 was transformed into the *RIB1* mutant strain rh-21 of *C. guilliermondii* by the LiCl procedure, respectively (6). The cells were plated on YPD medium without added riboflavin. Colonies growing without riboflavin were isolated. They were monitored for GTP cyclohydrolase II activity and for riboflavin production as described below.

- 20 The prototrophic strains were monitored for the presence of DNA segments introduced with the plasmid by PCR analysis. PCR was performed using the primers ShBle_V and PGgtpCY_nco and boiled *C. guilliermondii* recombinant strains cells as template. Primer ShBle_V is complementary to the *TEF* promoter and primer PGgtpCY_nco is complementary to the *RIB1*
25 structural gene. Amplificates of the expected length (1175 base pairs) were obtained from all transformants isolated. The amplificate obtained from strain XS-3 was isolated and was sequenced by the fluorescent dideoxy terminator method. The sequence is shown in Fig. 5. This sequence is identical with base pairs 1 to 1168 of the insert of plasmid pTCdXS2.

- 30 The recombinant transformants were genetically stable. Specifically, they did not segregate riboflavin deficient subclones.

Example 4

GTP cyclohydrolase activity in recombinant *C. guilliermondii* strains

- 10 -

The level of GTP cyclohydrolase II activity in the recombinant strains described above was determined as follows.

The recombinant *C. guilliermondii* cells were grown aerobically in synthetic Burkholder medium supplemented with trace elements (Science 101, 180, 1945) but without asparagine, during 2-3 days at 30°C. *C. guilliermondii* L2 strain (wild type) served as a control in these experiments.

Cells from exponential growth phase were harvested by centrifugation (5000 g, 15 min), washed twice with 20 mM Tris HCl, pH 8.2, containing 1 mM DTT and 1 mM MgCl₂. Cells were stored at -20°C. Frozen cells mass (1-3 g) was thawed in 3-9 ml of washing buffer. Cells were disrupted by agitation with glass beads (d = 0.8 mm). After centrifugation, cell extract was dialyzed overnight against 100 volumes of washing buffer. Protein concentration was measured by the Lowry method.

Reaction mixtures for GTP cyclohydrolase assays contained 20 mM Tris HCl, pH 8.2, 3 mM DTT, 2 mM MgCl₂, 1 mM GTP, and protein (protein concentration, 1-3 mg/ml, total volume, 4 ml). They were incubated at 37°C for 20 min in the dark.

After incubation, 2 ml aliquots were removed, and 2,3-butanedione was added to a final concentration of 0.5 mg/ml. The mixtures were incubated at 95°C for 30 min. Blank values were processed in the same way but without added diacetyl.

Differences in specific fluorescence of both types of aliquots were determined and were used to calculate the concentrations of 6,7-dimethylpteridin and the activity of GTP cyclohydrolase II. Results are shown in Table 3.

Strain L2 from which the mutant rh-21 had been derived was used as a control. The enzyme activity in strain L2 was 2,9 nmol mg⁻¹ h⁻¹. No enzyme activity was found in the riboflavin deficient recipient strain rh-21 carrying a mutation of the *RIB1* gene. The recombinant strains obtained by transformation with plasmid pTCdXS2 showed enzyme levels between 6,5 and 13,4 nmol mg⁻¹ h⁻¹. Thus, the enzyme level in recombinant strains was 2.3 - 4.6-fold higher as compared with the *C. guilliermondii* strain L2.

The activity of riboflavin synthase was also measured in the recombinant strains. The activity of riboflavin synthase was not affected by the

transformation with the plasmids p19RI, pTC2, pTCdXS2. All strains analyzed had riboflavin synthase activities in the range of 20 nmol mg⁻¹ h⁻¹ (Table 3).

Example 5

5 Production of riboflavin by recombinant *C. guilliermondii* strains

C. guilliermondii strains (wild type and recombinant strains) were grown aerobically in synthetic Burkholder medium supplemented with trace elements [Science 101, p. 180, (1945)] but without asparagine during 4 days at 30 °C [F.W. Tanner, Jr.C. Vojnovich, J.M. Van Lanen, Riboflavin production
10 by Candida species. Nature, 1945, 101 (2616): 180-181]. The suspension was centrifuged. Riboflavin concentration was determined fluorometrically. Results are shown in Table 4.

The wild strain L2 produced 1.2 mg riboflavin per liter under the conditions described. The recombinant strain XS-3 produced a 3-fold increased
15 level of riboflavin (3.6 mg/l).

Example 6

Isolation of riboflavin

Two Erlenmeyer flasks (2.5 l) each containing 0.5 l of synthetic Burkholder medium containing trace elements but no asparagine were
20 inoculated with the recombinant *C. guilliermondii* strain XS-3. The cultures were incubated with shaking at 30°C for 50 h. The solution was centrifuged. The supernatant was passed through a column of Florisil (4 ml bed volume) at a velocity of 500 ml/h. The column was washed with distilled water (7 ml). Riboflavin was eluted by a mixture of acetone/1M aqueous NH₄OH. The
25 effluent was evaporated to dryness. The yield of riboflavin was determined photometrically.

References

1. Sibirny, A. A., Zharova, V. P., Kshanovskaya, B. V., Shavlovsky, G. M. Selection of a genetic line of yeast *Pichia guilliermondii* capable to formation of significant amount of spores. *Tsytologia i genetika* 11, 330-333, 1977 (in Russian).
5
2. Shavlovskyy, G. M., Sibirnyy, A. A., Kshanovs'ka, B. V. Genetical classification of *C. guilliermondii* riboflavin auxotroph mutants. *Genetika* 15, 1561-1568, 1979 (in Russian).
3. Zakalsky, A. E., Zlochevsky, M. L., Stasiv, Y. Z., Logvynenko, E. M.,
10 Beburov, M. Y., Shavlovsky, G. M. Cloning of the *Pichia guilliermondii* *RIB1* gene coding for GTP cyclohydrolase II - first enzyme of flavinogenesis - in *E. coli* cells. *Genetika* 26, 614-620, 1990 (in Russian).
4. Liauta-Teglivets, O., Hasslacher, M., Boretskyy, Y., Kohlwein, S. D.,
15 Shavlovskii, G. M. Molecular cloning of the GTP cyclohydrolase. Structural gene *RIB1* of *Pichia guilliermondii* involved in riboflavin biosynthesis. *Yeast* 11, 945-952, 1995.
5. Zero Background/Kan Cloning Kit. Version A. 151204. Instruction manual. INVITROGENE.
6. Logvinenko, E. M., Stasiv, Yu. Z., Zlochevsky, M. L., Voronovsky, A. Ya.,
20 Beburov, M. Yu., Shavlovsky, G. M. Cloning of the *RIB7* gene encoding the riboflavin synthase of the yeast *Pichia guilliermondii*. *Genetika* 29, 922-927, 1993 (in Russian).

Table 1. Enzymes and genes of the riboflavin pathway

Enzyme	Gene		
	<i>S. cerevisiae</i>	<i>C. guilliermondii</i>	<i>E. coli</i>
A GTP cyclohydrolase	<i>RIB1</i>	<i>RIB1</i>	<i>ribA</i>
B bacterial deaminase			<i>ribD</i>
C yeast reductase	<i>RIB7</i>	<i>RIB2</i>	
D yeast deaminase	<i>RIB2</i>	<i>RIB3</i>	
E bacterial reductase			<i>ribD</i>
F unknown phosphatase			
G lumazine synthase	<i>RIB4</i>	<i>RIB5</i>	<i>ribE</i>
H riboflavin synthase	<i>RIB5</i>	<i>RIB7</i>	<i>ribC</i>
I 3,4-dihydroxy-2-butanon 4-phosphate synthase	<i>RIB3</i>	<i>RIB6</i>	<i>ribB</i>

Table 2. Nucleotide sequences of the primers used.

N	P r i m e r	S e q u e n c e (5' - 3')
1	ShB1e_V	GGGCATGCAATTCGAGCTCGGTACCCG
2	TEF1_H	CGACTCACTATAGGAGGAAGCTTGGCGC
3	PGgtpCY_V	AGGAGGAAGCTTGGCGCTATGGCATCGAAGG
4	PGgtpCY_nco	GCTGGTCGGTTAATGGGTGAAGCTGGG

- 14 -

Table 3. Activity of GTP cyclohydrolase II and riboflavin synthase in *C. guilliermondii* recombinant strains (time of growth: 40-48 h).

N	Strain	Riboflavin synthase activity nmol mg ⁻¹ h ⁻¹	GTP cyclohydrolase II activity nmol mg ⁻¹ h ⁻¹	Ratio*
1	L2 (wild type)	21.6	2.88	1.00
2	R1-1	n.d.	10.08	3.50
3	R1-2	n.d.	4.20	1.46
4	R1-3	20.4	8.76	3.04
5	R1-4	19.8	7.80	2.70
6	R1-5	21.6	7.80	2.70
7	TC-1	20.4	9.60	3.33
8	TC-2	n.d.	8.40	2.92
9	TC-3	n.d.	7.56	2.63
10	XS-1	22.8	13.38	4.60
11	XS-2	n.d.	12.60	4.37
12	XS-3	n.d.	6.60	2.29

n.d. not determined.

* GTP cyclohydrolase activity of recombinant strain divided by GTP
5 cyclohydrolase activity of strain L2

Table 4. Riboflavin production by recombinant *C. guilliermondii* strains (time of growth: 110 h, incubation temperature: 30 °C).

N	Strain	Riboflavin production [mg/l]	Relative riboflavin production
1	L2 (wild type)	1.2	1.0
4	R1-3	1.4	1.2
5	R1-4	3.6	3.0
6	R1-5	3.0	2.5
7	TC-1	1.4	1.2
9	TC-3	1.3	1.0
12	XS-3	3.6	3.0
13	XS-4	2.0	1.7
14	XS-5	2.3	1.9

-16-

Sequence Listing

GENERAL INFORMATION

- (i) APPLICANT: F. HOFFMANN-LA ROCHE AG
- (ii) TITLE OF INVENTION: OVERPRODUCTION OF RIBOFLAVIN IN YEAST
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADRESSEE: Gowling, Strathy & Henderson
 - (B) STREET: 160 Elgin Street, Suite 2600
 - (C) CITY: Ottawa
 - (D) PROVINCE: Ontario
 - (E) COUNTRY: Canada
 - (F) POSTAL CODE: K1P 1C3
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SOFTWARE: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 2,268,382
 - (B) FILING DATE: APRIL 21, 1999
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: EP 98107380.2
 - (B) FILING DATE: April 23, 1998
- (viii) AGENT INFORMATION:
 - (A) NAME: Eli J. McKhool
 - (B) REFERENCE NUMBER: 08-883093CA
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (613) 233-1781
 - (B) TELEFAX: (613) 563-9869

INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2181 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) ORIGINAL SOURCE:
 - ORGANISM: Candida guilliermondii
- (iv) FEATURE:
 - FEATURE KEY: misc feature
 - POSITION: (1450)..(1659)

-17-

DESCRIPTION: autonomously replicating sequence

(iv) FEATURE:
 FEATURE KEY: CDS
 POSITION: (474)..(1472)
 DESCRIPTION: translated portion of RIB1 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

gtcgactttc actccgaagg taggtgcggc tggaagacgt cgtcccaagt cgtatgcggt 60
agctgagagc gacggaaacg aaagtgatga agattacatg ctggaataat ccatagctag 120
tgtacttgct aatacaaccg gtaaagctag ccaattgcag cggtattcac caccgccgtg 180
gatcgggtta gtcacgtgaa ctggccggtg ggtcctgcac gtcgcttcat tattcatata 240
ttagtgagag tcttcctata tcagtcagca gacgtatcgg ttgatttcag gtcaaaaaga 300
gaaaagggtg tcttacaaaa gcgaaatagc tgatacattt ttactcacag cagcatcata 360
tttgtggaac ctttaaactt gacttttcat ttcaagcaag ttattttgaa attcaaata 420
tttgaaaatc aaaaaagaac atctaagttc tgaaaaattg tacgaacaac gct atg      476
                                     Met
                                     1

gca tcg aag gac ata gta cat ccg caa cca gag cgc cgg cac ggg tcg      524
Ala Ser Lys Asp Ile Val His Pro Gln Pro Glu Arg Arg His Gly Ser
          5                      10                      15

gaa act cac gaa ttt acc atg cct ctc tta tct cct aca ttg aca cca      572
Glu Thr His Glu Phe Thr Met Pro Leu Leu Ser Pro Thr Leu Thr Pro
          20                      25                      30

tcc cat att cca tcg caa acg cct caa att cct ccg gaa gtg cca gca      620
Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro Pro Glu Val Pro Ala
          35                      40                      45

gaa gtc agg gat cgc ttg ccc ctt cct gaa acg ttg cct gtg gtg aaa      668
Glu Val Arg Asp Arg Leu Pro Leu Pro Glu Thr Leu Pro Val Val Lys
          50                      55                      60                      65

tgc atg gcg aga gct cgt ata ccg acc act cag ggg ccg gag ata ttt      716
Cys Met Ala Arg Ala Arg Ile Pro Thr Thr Gln Gly Pro Glu Ile Phe
          70                      75                      80

ctc cat ttg tac gag aat aac gtt gac aat aaa gag cat ttg gct att      764
Leu His Leu Tyr Glu Asn Asn Val Asp Asn Lys Glu His Leu Ala Ile
          85                      90                      95

gtt ttt ggg gaa gat gtg cgg tcg aaa acg ctc tat cag aaa cgt ccc      812
Val Phe Gly Glu Asp Val Arg Ser Lys Thr Leu Tyr Gln Lys Arg Pro
          100                      105                      110

```

-18-

aat gag acc cag caa gat aga atg act cgt ggt gct tat gtg ggc aga	860
Asn Glu Thr Gln Gln Asp Arg Met Thr Arg Gly Ala Tyr Val Gly Arg	
115 120 125	
ttg ttt cct gga aga acc gag gca gac tat gac agt gag tct aat ttg	908
Leu Phe Pro Gly Arg Thr Glu Ala Asp Tyr Asp Ser Glu Ser Asn Leu	
130 135 140 145	
aga ttg aat ttc gat gaa aat ggc caa ctt atc aga gat ccg agt acc	956
Arg Leu Asn Phe Asp Glu Asn Gly Gln Leu Ile Arg Asp Pro Ser Thr	
150 155 160	
acc tgt agt ggt gag ccc att ttg gcc cgt att cat tcg gaa tgt tat	1004
Thr Cys Ser Gly Glu Pro Ile Leu Ala Arg Ile His Ser Glu Cys Tyr	
165 170 175	
acg ggg gaa acc gca tgg agt gct cgt tgc gat tgt gga gaa caa ttc	1052
Thr Gly Glu Thr Ala Trp Ser Ala Arg Cys Asp Cys Gly Glu Gln Phe	
180 185 190	
gat gaa gct ggt cgg tta atg ggt gaa gct ggg cac ggg tgt atc gtg	1100
Asp Glu Ala Gly Arg Leu Met Gly Glu Ala Gly His Gly Cys Ile Val	
195 200 205	
tac ctt cgt cag gaa ggt cgt gga att gga ctt ggg gaa aag ttg aag	1148
Tyr Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu Gly Glu Lys Leu Lys	
210 215 220 225	
gct tat aat ttg caa gac ttg gga gcg gat acc gtc cag gcc aat ttg	1196
Ala Tyr Asn Leu Gln Asp Leu Gly Ala Asp Thr Val Gln Ala Asn Leu	
230 235 240	
atg tta cga cat cct gct gat gcg aga tct ttt tcg ctc gct aca gcc	1244
Met Leu Arg His Pro Ala Asp Ala Arg Ser Phe Ser Leu Ala Thr Ala	
245 250 255	
ata ctc ttg gac ttg ggg ctc aac gag atc aag ttg ttg acc aac aat	1292
Ile Leu Leu Asp Leu Gly Leu Asn Glu Ile Lys Leu Leu Thr Asn Asn	
260 265 270	
ccc gat aaa att gct gca gtt gag gga aga aac aga gag gtc aag gta	1340
Pro Asp Lys Ile Ala Ala Val Glu Gly Arg Asn Arg Glu Val Lys Val	
275 280 285	
gtg gaa cgg gtg cct atg gtg ccg ttg gca tgg aga agt gag aat gga	1388
Val Glu Arg Val Pro Met Val Pro Leu Ala Trp Arg Ser Glu Asn Gly	
290 295 300 305	
atc aag tca aaa gag ata gag ggc tac ttg agt gct aag att gaa agg	1436
Ile Lys Ser Lys Glu Ile Glu Gly Tyr Leu Ser Ala Lys Ile Glu Arg	
310 315 320	
atg ggg cac ttg ctt gaa aag cca ctc aag ata tga tagaagagat	1482
Met Gly His Leu Leu Glu Lys Pro Leu Lys Ile	
325 330	

-19-

gaagttaagg acttaagaaa taaatgatga attaaatgac gcaaattgtca ctactcgatt 1542
 agagaaatag ctataatgaa gaatttttgca tttcgcaaaa tttaagataa atgcaaaaat 1602
 tgcaaattac gaaatatgca tatgatacaa gacaagaaaa gactactaaa agtctctcga 1662
 gaagaatact gggtaacctt catctcttga ttatgcactg gggctattca tatgcagatt 1722
 cgcacgccga ggtgcagcgt ttaggcgcgg ctcaacggaa gccaacggcc gccacaaatt 1782
 gtccggaaag tcgccgaaac tgatccactg gtaccacagc cccataagaa cccctttaa 1842
 tattaanaac cgttcttcag ccacttttga tcacattggt tgcagccgcc cgttgctgcc 1902
 atccaaacac cacgcgtccc ccgcacctt tacgggtgcc actgcattgg aatttgcata 1962
 aaacagcctc acgaagtgga ttaattttta gagcactcaa gtcacatgc tgcaatctct 2022
 gcatcatgaa atgactcccg ttgatacagg gaactcagac cgcaagcggc gaagagtcac 2082
 aagagcgtgt gatgtgtgtc gactctagag atccccgggt accgagctcg aattcactgg 2142
 ccgtcgtttt acaacgtcgt gactgggaaa accctggcg 2181

INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 332 residues
 - (B) TYPE: amino acid
 - (C) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) ORIGINAL SOURCE:
 - ORGANISM: Candida guilliermondii

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met	Ala	Ser	Lys	Asp	Ile	Val	His	Pro	Gln	Pro	Glu	Arg	Arg	His	Gly
1				5					10					15	
Ser	Glu	Thr	His	Glu	Phe	Thr	Met	Pro	Leu	Leu	Ser	Pro	Thr	Leu	Thr
			20					25					30		
Pro	Ser	His	Ile	Pro	Ser	Gln	Thr	Pro	Gln	Ile	Pro	Pro	Glu	Val	Pro
		35					40					45			
Ala	Glu	Val	Arg	Asp	Arg	Leu	Pro	Leu	Pro	Glu	Thr	Leu	Pro	Val	Val
	50					55				60					
Lys	Cys	Met	Ala	Arg	Ala	Arg	Ile	Pro	Thr	Thr	Gln	Gly	Pro	Glu	Ile
65					70					75				80	

-20-

Phe	Leu	His	Leu	Tyr	Glu	Asn	Asn	Val	Asp	Asn	Lys	Glu	His	Leu	Ala		
				85					90					95			
Ile	Val	Phe	Gly	Glu	Asp	Val	Arg	Ser	Lys	Thr	Leu	Tyr	Gln	Lys	Arg		
			100					105					110				
Pro	Asn	Glu	Thr	Gln	Gln	Asp	Arg	Met	Thr	Arg	Gly	Ala	Tyr	Val	Gly		
		115					120					125					
Arg	Leu	Phe	Pro	Gly	Arg	Thr	Glu	Ala	Asp	Tyr	Asp	Ser	Glu	Ser	Asn		
	130					135					140						
Leu	Arg	Leu	Asn	Phe	Asp	Glu	Asn	Gly	Gln	Leu	Ile	Arg	Asp	Pro	Ser		
145					150					155					160		
Thr	Thr	Cys	Ser	Gly	Glu	Pro	Ile	Leu	Ala	Arg	Ile	His	Ser	Glu	Cys		
				165					170					175			
Tyr	Thr	Gly	Glu	Thr	Ala	Trp	Ser	Ala	Arg	Cys	Asp	Cys	Gly	Glu	Gln		
			180					185					190				
Phe	Asp	Glu	Ala	Gly	Arg	Leu	Met	Gly	Glu	Ala	Gly	His	Gly	Cys	Ile		
	195						200					205					
Val	Tyr	Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Gly	Glu	Lys	Leu		
	210					215					220						
Lys	Ala	Tyr	Asn	Leu	Gln	Asp	Leu	Gly	Ala	Asp	Thr	Val	Gln	Ala	Asn		
225				230					235						240		
Leu	Met	Leu	Arg	His	Pro	Ala	Asp	Ala	Arg	Ser	Phe	Ser	Leu	Ala	Thr		
				245					250					255			
Ala	Ile	Leu	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Lys	Leu	Leu	Thr	Asn		
			260					265					270				
Asn	Pro	Asp	Lys	Ile	Ala	Ala	Val	Glu	Gly	Arg	Asn	Arg	Glu	Val	Lys		
		275					280					285					
Val	Val	Glu	Arg	Val	Pro	Met	Val	Pro	Leu	Ala	Trp	Arg	Ser	Glu	Asn		
	290					295					300						
Gly	Ile	Lys	Ser	Lys	Glu	Ile	Glu	Gly	Tyr	Leu	Ser	Ala	Lys	Ile	Glu		
305					310					315					320		
Arg	Met	Gly	His	Leu	Leu	Glu	Lys	Pro	Leu	Lys	Ile						
				325					330								

INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2264 residues
- (B) TYPE: nucleic acid

-21-

- (ii) MOLECULE TYPE: DNA (artificial sequence)
 DESCRIPTION: Amplified DNA containing parts of the complete TEF promoter of *S. cerevisiae* and the 5' part of the *C. guilliermondii* RIB1 gene
- (iv) FEATURE:
 FEATURE KEY: promoter
 POSITION: (1)..(552)
 DESCRIPTION: TEF promoter fragment
- (iv) FEATURE:
 FEATURE KEY: CDS
 POSITION: (557)..(1555)
 DESCRIPTION: translated portion of RIB1 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

caattcgagc tcggtacccg gggatccccc acacaccata gcttcaaaat gtttctactc 60
cttttttact cttccagatt ttctcggact ccgcgcacgc ccgtaccact tcaaaacacc 120
caagcacagc atactaaatt tccctctttc ttctcttagg gtgtcggttaa ttacccgtag 180
taaaggtttg gaaaagaaaa aagagaccgc ctcgtttctt tttcttcgtc gaaaaaggca 240
ataaaaatth ttatcacgtt tctttttctt gaaatttttt tttttgattt ttttctcttt 300
cgatgacctc ccattgatat ttaagtcaat aaacgggtctt caatttctca agtttcagtt 360
tcatttttct tgttctatta caactttttt tacttcttgc tcattagaaa gaaagcatag 420
caatctaadc taagggcgag ctccaattcg aactagtact gcagcacgtg accggcgcct 480
agtgttgaca attaatacgc ggcatagtat atcggcatag tataatacga ctactatag 540
gaggaagctt ggcgcct atg gca tcg aag tac ata gta cat ccg caa cca gag 592
          Met Ala Ser Lys Tyr Ile Val His Pro Gln Pro Glu
              1              5              10

cgc cgg cac ggg tcg gaa act cac gaa ttt acc atg cct ctc tta tct 640
Arg Arg His Gly Ser Glu Thr His Glu Phe Thr Met Pro Leu Leu Ser
          15              20              25

cct aca ttg aca cca tcc cat att cca tcg caa acg cct caa att cct 688
Pro Thr Leu Thr Pro Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro
          30              35              40

ccg gaa gtg cca gca gaa gtc agg gat cgc ttg ccc ctt cct gaa acg 736
Pro Glu Val Pro Ala Glu Val Arg Asp Arg Leu Pro Leu Pro Glu Thr
          45              50              55              60

ttg cct gtg gtg aaa tgc atg gcg aga gct cgt ata ccg acc act cag 784
Leu Pro Val Val Lys Cys Met Ala Arg Ala Arg Ile Pro Thr Thr Gln
          65              70              75

```

-22-

ggg	ccg	gag	ata	ttt	ctc	cat	ttg	tac	gag	aat	aac	gtt	gac	aat	aaa	832
Gly	Pro	Glu	Ile	Phe	Leu	His	Leu	Tyr	Glu	Asn	Asn	Val	Asp	Asn	Lys	
			80					85					90			
gag	cat	ttg	gct	att	gtt	ttt	ggg	gaa	gat	gtg	cgg	tcg	aaa	acg	ctc	880
Glu	His	Leu	Ala	Ile	Val	Phe	Gly	Glu	Asp	Val	Arg	Ser	Lys	Thr	Leu	
		95					100					105				
tat	cag	aaa	cgt	ccc	aat	gag	acc	cag	caa	gat	aga	atg	act	cgt	ggt	928
Tyr	Gln	Lys	Arg	Pro	Asn	Glu	Thr	Gln	Gln	Asp	Arg	Met	Thr	Arg	Gly	
	110					115				120						
gct	tat	gtg	ggc	aga	ttg	ttt	cct	gga	aga	acc	gag	gca	gac	tat	gac	976
Ala	Tyr	Val	Gly	Arg	Leu	Phe	Pro	Gly	Arg	Thr	Glu	Ala	Asp	Tyr	Asp	
125					130					135					140	
agt	gag	tct	aat	ttg	aga	ttg	aat	ttc	gat	gaa	aat	ggc	caa	ctt	atc	1024
Ser	Glu	Ser	Asn	Leu	Arg	Leu	Asn	Phe	Asp	Glu	Asn	Gly	Gln	Leu	Ile	
			145					150						155		
aga	gat	ccg	agt	acc	acc	tgt	agt	ggg	gag	ccc	att	ttg	gcc	cgt	att	1072
Arg	Asp	Pro	Ser	Thr	Thr	Cys	Ser	Gly	Glu	Pro	Ile	Leu	Ala	Arg	Ile	
		160						165					170			
cat	tcg	gaa	tgt	tat	acg	ggg	gaa	acc	gca	tgg	agt	gct	cgt	tgc	gat	1120
His	Ser	Glu	Cys	Tyr	Thr	Gly	Glu	Thr	Ala	Trp	Ser	Ala	Arg	Cys	Asp	
		175				180						185				
tgt	gga	gaa	caa	ttc	gat	gaa	gct	ggg	cgg	tta	atg	ggg	gaa	gct	ggg	1168
Cys	Gly	Glu	Gln	Phe	Asp	Glu	Ala	Gly	Arg	Leu	Met	Gly	Glu	Ala	Gly	
	190					195				200						
cac	ggg	tgt	atc	gtg	tac	ctt	cgt	cag	gaa	ggg	cgt	gga	att	gga	ctt	1216
His	Gly	Cys	Ile	Val	Tyr	Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	
205					210					215					220	
ggg	gaa	aag	ttg	aag	gct	tat	aat	ttg	caa	gac	ttg	gga	gcg	gat	acc	1264
Gly	Glu	Lys	Leu	Lys	Ala	Tyr	Asn	Leu	Gln	Asp	Leu	Gly	Ala	Asp	Thr	
			225					230					235			
gtc	cag	gcc	aat	ttg	atg	tta	cga	cat	cct	gct	gat	gcg	aga	tct	ttt	1312
Val	Gln	Ala	Asn	Leu	Met	Leu	Arg	His	Pro	Ala	Asp	Ala	Arg	Ser	Phe	
			240					245					250			
tcg	ctc	gct	aca	gcc	ata	ctc	ttg	gac	ttg	ggg	ctc	aac	gag	atc	aag	1360
Ser	Leu	Ala	Thr	Ala	Ile	Leu	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Lys	
		255					260					265				
ttg	ttg	acc	aac	aat	ccc	gat	aaa	att	gct	gca	gtt	gag	gga	aga	aac	1408
Leu	Leu	Thr	Asn	Asn	Pro	Asp	Lys	Ile	Ala	Ala	Val	Glu	Gly	Arg	Asn	
	270					275					280					
aga	gag	gtc	aag	gta	gtg	gaa	cgg	gtg	cct	atg	gtg	ccg	ttg	gca	tgg	1456
Arg	Glu	Val	Lys	Val	Val	Glu	Arg	Val	Pro	Met	Val	Pro	Leu	Ala	Trp	
285					290					295					300	

Pro Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro Pro Glu Val Pro
35 40 45

-24-

Ala	Glu	Val	Arg	Asp	Arg	Leu	Pro	Leu	Pro	Glu	Thr	Leu	Pro	Val	Val	50	55	60	
Lys	Cys	Met	Ala	Arg	Ala	Arg	Ile	Pro	Thr	Thr	Gln	Gly	Pro	Glu	Ile	65	70	75	80
Phe	Leu	His	Leu	Tyr	Glu	Asn	Asn	Val	Asp	Asn	Lys	Glu	His	Leu	Ala	85	90	95	
Ile	Val	Phe	Gly	Glu	Asp	Val	Arg	Ser	Lys	Thr	Leu	Tyr	Gln	Lys	Arg	100	105	110	
Pro	Asn	Glu	Thr	Gln	Gln	Asp	Arg	Met	Thr	Arg	Gly	Ala	Tyr	Val	Gly	115	120	125	
Arg	Leu	Phe	Pro	Gly	Arg	Thr	Glu	Ala	Asp	Tyr	Asp	Ser	Glu	Ser	Asn	130	135	140	
Leu	Arg	Leu	Asn	Phe	Asp	Glu	Asn	Gly	Gln	Leu	Ile	Arg	Asp	Pro	Ser	145	150	155	160
Thr	Thr	Cys	Ser	Gly	Glu	Pro	Ile	Leu	Ala	Arg	Ile	His	Ser	Glu	Cys	165	170	175	
Tyr	Thr	Gly	Glu	Thr	Ala	Trp	Ser	Ala	Arg	Cys	Asp	Cys	Gly	Glu	Gln	180	185	190	
Phe	Asp	Glu	Ala	Gly	Arg	Leu	Met	Gly	Glu	Ala	Gly	His	Gly	Cys	Ile	195	200	205	
Val	Tyr	Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Gly	Glu	Lys	Leu	210	215	220	
Lys	Ala	Tyr	Asn	Leu	Gln	Asp	Leu	Gly	Ala	Asp	Thr	Val	Gln	Ala	Asn	225	230	235	240
Leu	Met	Leu	Arg	His	Pro	Ala	Asp	Ala	Arg	Ser	Phe	Ser	Leu	Ala	Thr	245	250	255	
Ala	Ile	Leu	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Lys	Leu	Leu	Thr	Asn	260	265	270	
Asn	Pro	Asp	Lys	Ile	Ala	Ala	Val	Glu	Gly	Arg	Asn	Arg	Glu	Val	Lys	275	280	285	
Val	Val	Glu	Arg	Val	Pro	Met	Val	Pro	Leu	Ala	Trp	Arg	Ser	Glu	Asn	290	295	300	
Gly	Ile	Lys	Ser	Lys	Glu	Ile	Glu	Gly	Tyr	Leu	Ser	Ala	Lys	Ile	Glu	305	310	315	320
Arg	Met	Gly	His	Leu	Leu	Glu	Lys	Pro	Leu	Lys	Ile					325	330		

-25-

INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1742 residues
 (B) TYPE: nucleic acid
- (ii) MOLECULE TYPE: DNA (artificial sequence)
 DESCRIPTION: Nucleotide sequence of pTCdXS-2 insert
- (iv) FEATURE:
 FEATURE KEY: promoter
 POSITION: (1)..(552)
 DESCRIPTION: TEF promoter fragment
- (iv) FEATURE:
 FEATURE KEY: misc feature
 POSITION: (1533)..(1742)
 DESCRIPTION: autonomously replicating sequence
- (iv) FEATURE:
 FEATURE KEY: CDS
 POSITION: (557)..(1555)
 DESCRIPTION: translated portion of RIB1 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

caattcgagc tcggtaccgc gggatccccc acacaccata gcttcaaaat gtttctactc 60
cttttttact cttccagatt ttctcggact ccgcgcacgc ccgtaccact tcaaaacacc 120
caagcacagc atactaaatt tccctctttc ttctctagg gtgtcggttaa ttaccggtac 180
taaagggttg gaaaagaaaa aagagaccgc ctcgtttctt tttcttcgtc gaaaaaggca 240
ataaaaattt ttatcacggt tcttttttctt gaaatttttt tttttgattt ttttctcttt 300
cgatgacctc ccattgatat ttaagtcaat aaacgggtctt caatttctca agtttcagtt 360
tcatttttct tgttctatta caactttttt tacttcttgc tcattagaaa gaaagcatag 420
caatctaata taagggcgag ctggaattcg aactagtact gcagcacgtg accggcgctt 480
agtgttgaca attaatacat gccatagtat atcggcatag tataatacga ctactatag 540
gaggaagctt ggcgct atg gca tcg aag tac ata gta cat ccg caa cca gag 592
          Met Ala Ser Lys Tyr Ile Val His Pro Gln Pro Glu
              1              5              10

cgc cgg cac ggg tcg gaa act cac gaa ttt acc atg cct ctc tta tct 640
Arg Arg His Gly Ser Glu Thr His Glu Phe Thr Met Pro Leu Leu Ser
      15              20              25

cct aca ttg aca cca tcc cat att cca tcg caa acg cct caa att cct 688
Pro Thr Leu Thr Pro Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro
      30              35              40

```

-26-

ccg gaa gtg cca gca gaa gtc agg gat cgc ttg ccc ctt cct gaa acg	736
Pro Glu Val Pro Ala Glu Val Arg Asp Arg Leu Pro Leu Pro Glu Thr	
45 50 55 60	
ttg cct gtg gtg aaa tgc atg gcg aga gct cgt ata ccg acc act cag	784
Leu Pro Val Val Lys Cys Met Ala Arg Ala Arg Ile Pro Thr Thr Gln	
65 70 75	
ggg ccg gag ata ttt ctc cat ttg tac gag aat aac gtt gac aat aaa	832
Gly Pro Glu Ile Phe Leu His Leu Tyr Glu Asn Asn Val Asp Asn Lys	
80 85 90	
gag cat ttg gct att gtt ttt ggg gaa gat gtg cgg tcg aaa acg ctc	880
Glu His Leu Ala Ile Val Phe Gly Glu Asp Val Arg Ser Lys Thr Leu	
95 100 105	
tat cag aaa cgt ccc aat gag acc cag caa gat aga atg act cgt ggt	928
Tyr Gln Lys Arg Pro Asn Glu Thr Gln Gln Asp Arg Met Thr Arg Gly	
110 115 120	
gct tat gtg ggc aga ttg ttt cct gga aga acc gag gca gac tat gac	976
Ala Tyr Val Gly Arg Leu Phe Pro Gly Arg Thr Glu Ala Asp Tyr Asp	
125 130 135 140	
agt gag tct aat ttg aga ttg aat ttc gat gaa aat ggc caa ctt atc	1024
Ser Glu Ser Asn Leu Arg Leu Asn Phe Asp Glu Asn Gly Gln Leu Ile	
145 150 155	
aga gat ccg agt acc acc tgt agt ggt gag ccc att ttg gcc cgt att	1072
Arg Asp Pro Ser Thr Thr Cys Ser Gly Glu Pro Ile Leu Ala Arg Ile	
160 165 170	
cat tcg gaa tgt tat acg ggg gaa acc gca tgg agt gct cgt tgc gat	1120
His Ser Glu Cys Tyr Thr Gly Glu Thr Ala Trp Ser Ala Arg Cys Asp	
175 180 185	
tgt gga gaa caa ttc gat gaa gct ggt cgg tta atg ggt gaa gct ggg	1168
Cys Gly Glu Gln Phe Asp Glu Ala Gly Arg Leu Met Gly Glu Ala Gly	
190 195 200	
cac ggg tgt atc gtg tac ctt cgt cag gaa ggt cgt gga att gga ctt	1216
His Gly Cys Ile Val Tyr Leu Arg Gln Glu Gly Arg Gly Ile Gly Leu	
205 210 215 220	
ggg gaa aag ttg aag gct tat aat ttg caa gac ttg gga gcg gat acc	1264
Gly Glu Lys Leu Lys Ala Tyr Asn Leu Gln Asp Leu Gly Ala Asp Thr	
225 230 235	
gtc cag gcc aat ttg atg tta cga cat cct gct gat gcg aga tct ttt	1312
Val Gln Ala Asn Leu Met Leu Arg His Pro Ala Asp Ala Arg Ser Phe	
240 245 250	

-27-

```

tcg ctc gct aca gcc ata ctc ttg gac ttg ggg ctc aac gag atc aag 1360
Ser Leu Ala Thr Ala Ile Leu Leu Asp Leu Gly Leu Asn Glu Ile Lys
      255              260              265

ttg ttg acc aac aat ccc gat aaa att gct gca gtt gag gga aga aac 1408
Leu Leu Thr Asn Asn Pro Asp Lys Ile Ala Ala Val Glu Gly Arg Asn
      270              275              280

aga gag gtc aag gta gtg gaa cgg gtg cct atg gtg ccg ttg gca tgg 1456
Arg Glu Val Lys Val Val Glu Arg Val Pro Met Val Pro Leu Ala Trp
      285              290              295              300

aga agt aag aat gga atc aag tca aaa gag ata gag ggc tac ttg agt 1504
Arg Ser Lys Asn Gly Ile Lys Ser Lys Glu Ile Glu Gly Tyr Leu Ser
              305              310              315

gct aag att gaa agg atg ggg cac ttg ctt gaa aag cca ctc aag ata 1552
Ala Lys Ile Glu Arg Met Gly His Leu Leu Glu Lys Pro Leu Lys Ile
              320              325              330

tga tagaagagat gaagttaagg acttaagaaa taaatgatga attaaatgac 1605

gcaaattgtca ctactcgatt agagaaatag ctataatgaa gaatttttgca tttcgcaaaa 1665

tttaagataa atgcaaaaat tgcaaattac gaaatatgca tatgatacaa gacaagaaaa 1725

gactactaaa agtctct 1742

```

INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 332 residues
 (B) TYPE: amino acid
 (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein (artificial sequence)

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

Met Ala Ser Lys Tyr Ile Val His Pro Gln Pro Glu Arg Arg His Gly
  1              5              10              15

Ser Glu Thr His Glu Phe Thr Met Pro Leu Leu Ser Pro Thr Leu Thr
              20              25              30

Pro Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro Pro Glu Val Pro
              35              40              45

Ala Glu Val Arg Asp Arg Leu Pro Leu Pro Glu Thr Leu Pro Val Val
              50              55              60

```

-28-

Lys	Cys	Met	Ala	Arg	Ala	Arg	Ile	Pro	Thr	Thr	Gln	Gly	Pro	Glu	Ile	65	70	75	80
Phe	Leu	His	Leu	Tyr	Glu	Asn	Asn	Val	Asp	Asn	Lys	Glu	His	Leu	Ala	85	90	95	
Ile	Val	Phe	Gly	Glu	Asp	Val	Arg	Ser	Lys	Thr	Leu	Tyr	Gln	Lys	Arg	100	105	110	
Pro	Asn	Glu	Thr	Gln	Gln	Asp	Arg	Met	Thr	Arg	Gly	Ala	Tyr	Val	Gly	115	120	125	
Arg	Leu	Phe	Pro	Gly	Arg	Thr	Glu	Ala	Asp	Tyr	Asp	Ser	Glu	Ser	Asn	130	135	140	
Leu	Arg	Leu	Asn	Phe	Asp	Glu	Asn	Gly	Gln	Leu	Ile	Arg	Asp	Pro	Ser	145	150	155	160
Thr	Thr	Cys	Ser	Gly	Glu	Pro	Ile	Leu	Ala	Arg	Ile	His	Ser	Glu	Cys	165	170	175	
Tyr	Thr	Gly	Glu	Thr	Ala	Trp	Ser	Ala	Arg	Cys	Asp	Cys	Gly	Glu	Gln	180	185	190	
Phe	Asp	Glu	Ala	Gly	Arg	Leu	Met	Gly	Glu	Ala	Gly	His	Gly	Cys	Ile	195	200	205	
Val	Tyr	Leu	Arg	Gln	Glu	Gly	Arg	Gly	Ile	Gly	Leu	Gly	Glu	Lys	Leu	210	215	220	
Lys	Ala	Tyr	Asn	Leu	Gln	Asp	Leu	Gly	Ala	Asp	Thr	Val	Gln	Ala	Asn	225	230	235	240
Leu	Met	Leu	Arg	His	Pro	Ala	Asp	Ala	Arg	Ser	Phe	Ser	Leu	Ala	Thr	245	250	255	
Ala	Ile	Leu	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Lys	Leu	Leu	Thr	Asn	260	265	270	
Asn	Pro	Asp	Lys	Ile	Ala	Ala	Val	Glu	Gly	Arg	Asn	Arg	Glu	Val	Lys	275	280	285	
Val	Val	Glu	Arg	Val	Pro	Met	Val	Pro	Leu	Ala	Trp	Arg	Ser	Lys	Asn	290	295	300	
Gly	Ile	Lys	Ser	Lys	Glu	Ile	Glu	Gly	Tyr	Leu	Ser	Ala	Lys	Ile	Glu	305	310	315	320
Arg	Met	Gly	His	Leu	Leu	Glu	Lys	Pro	Leu	Lys	Ile					325	330		

INFORMATION FOR SEQ ID NO: 7:

-29-

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1175 residues
 (B) TYPE: nucleic acid
- (ii) MOLECULE TYPE: DNA (artificial sequence)
 DESCRIPTION: PCR fragment containing base pairs 1-1168 of the pTCDXS-2 insert
- (iv) FEATURE:
 FEATURE KEY: promoter
 POSITION: (8)..(559)
 DESCRIPTION: TEF promoter fragment
- (iv) FEATURE:
 FEATURE KEY: CDS
 POSITION: (564)..(1172)
 DESCRIPTION: partially translated portion of RIB1 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```

gggcatgcaa ttcgagctcg gtacccgggg atccccaca caccatagct tcaaaatggt 60
tctactcctt ttttactctt ccagattttc tcggactccg cgcacgcccg taccacttca 120
aaacacccaa gcacagcata ctaaattttc ctctttcttc ctctaggggtg tcgttaatta 180
cccgacttaa aggtttggaa aagaaaaaag agaccgcctc gtttcttttt cttcgtcgaa 240
aaaggcaata aaaattttta tcacgtttct ttttcttgaa attttttttt ttgatttttt 300
tctctttcga tgacctccca ttgatattta agtcaataaa cggctcttcaa tttctcaagt 360
ttcagtttca tttttcttgt tctattacaa ctttttttac ttcttgctca ttagaaagaa 420
agcatagcaa tctaattctaa gggcgagctc gaattcgaac tagtactgca gcacgtgacc 480
ggcgccctagt gttgacaatt aatcatcggc atagtatatc ggcatagtat aatacgactc 540
actataggag gaagcttggc gct atg gca tcg aag tac ata gta cat ccg caa 593
                Met Ala Ser Lys Tyr Ile Val His Pro Gln
                  1                   5                   10

cca gag cgc cgg cac ggg tcg gaa act cac gaa ttt acc atg cct ctc 641
Pro Glu Arg Arg His Gly Ser Glu Thr His Glu Phe Thr Met Pro Leu
                15                   20                   25

tta tct cct aca ttg aca cca tcc cat att cca tcg caa acg cct caa 689
Leu Ser Pro Thr Leu Thr Pro Ser His Ile Pro Ser Gln Thr Pro Gln
                30                   35                   40

att cct ccg gaa gtg cca gca gaa gtc agg gat cgc ttg ccc ctt cct 737
Ile Pro Pro Glu Val Pro Ala Glu Val Arg Asp Arg Leu Pro Leu Pro
                45                   50                   55

```

-30-

gaa acg ttg cct gtg gtg aaa tgc atg gcg aga gct cgt ata ccg acc	785
Glu Thr Leu Pro Val Val Lys Cys Met Ala Arg Ala Arg Ile Pro Thr	
60 65 70	
act cag ggg ccg gag ata ttt ctc cat ttg tac gag aat aac gtt gac	833
Thr Gln Gly Pro Glu Ile Phe Leu His Leu Tyr Glu Asn Asn Val Asp	
75 80 85 90	
aat aaa gag cat ttg gct att gtt ttt ggg gaa gat gtg cgg tcg aaa	881
Asn Lys Glu His Leu Ala Ile Val Phe Gly Glu Asp Val Arg Ser Lys	
95 100 105	
acg ctc tat cag aaa cgt ccc aat gag acc cag caa gat aga atg act	929
Thr Leu Tyr Gln Lys Arg Pro Asn Glu Thr Gln Gln Asp Arg Met Thr	
110 115 120	
cgt ggt gct tat gtg ggc aga ttg ttt cct gga aga acc gag gca gac	977
Arg Gly Ala Tyr Val Gly Arg Leu Phe Pro Gly Arg Thr Glu Ala Asp	
125 130 135	
tat gac agt gag tct aat ttg aga ttg aat ttc gat gaa aat ggc caa	1025
Tyr Asp Ser Glu Ser Asn Leu Arg Leu Asn Phe Asp Glu Asn Gly Gln	
140 145 150	
ctt atc aga gat ccg agt acc acc tgt agt ggt gag ccc att ttg gcc	1073
Leu Ile Arg Asp Pro Ser Thr Thr Cys Ser Gly Glu Pro Ile Leu Ala	
155 160 165 170	
cgt att cat tcg gaa tgt tat acg ggg gaa acc gca tgg agt gct cgt	1121
Arg Ile His Ser Glu Cys Tyr Thr Gly Glu Thr Ala Trp Ser Ala Arg	
175 180 185	
tgc gat tgt gga gaa caa ttc gat gaa gct ggt cgg tta atg ggt gaa	1169
Cys Asp Cys Gly Glu Gln Phe Asp Glu Ala Gly Arg Leu Met Gly Glu	
190 195 200	
gct ggg	1175
Ala	

INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 203 residues
 - (B) TYPE: amino acid
 - (C) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein (artificial sequence)

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met	Ala	Ser	Lys	Tyr	Ile	Val	His	Pro	Gln	Pro	Glu	Arg	Arg	His	Gly
1					5				10					15	

-31-

Ser Glu Thr His Glu Phe Thr Met Pro Leu Leu Ser Pro Thr Leu Thr
 20 25 30
 Pro Ser His Ile Pro Ser Gln Thr Pro Gln Ile Pro Pro Glu Val Pro
 35 40 45
 Ala Glu Val Arg Asp Arg Leu Pro Leu Pro Glu Thr Leu Pro Val Val
 50 55 60
 Lys Cys Met Ala Arg Ala Arg Ile Pro Thr Thr Gln Gly Pro Glu Ile
 65 70 75 80
 Phe Leu His Leu Tyr Glu Asn Asn Val Asp Asn Lys Glu His Leu Ala
 85 90 95
 Ile Val Phe Gly Glu Asp Val Arg Ser Lys Thr Leu Tyr Gln Lys Arg
 100 105 110
 Pro Asn Glu Thr Gln Gln Asp Arg Met Thr Arg Gly Ala Tyr Val Gly
 115 120 125
 Arg Leu Phe Pro Gly Arg Thr Glu Ala Asp Tyr Asp Ser Glu Ser Asn
 130 135 140
 Leu Arg Leu Asn Phe Asp Glu Asn Gly Gln Leu Ile Arg Asp Pro Ser
 145 150 155 160
 Thr Thr Cys Ser Gly Glu Pro Ile Leu Ala Arg Ile His Ser Glu Cys
 165 170 175
 Tyr Thr Gly Glu Thr Ala Trp Ser Ala Arg Cys Asp Cys Gly Glu Gln
 180 185 190
 Phe Asp Glu Ala Gly Arg Leu Met Gly Glu Ala
 195 200

INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 residues
 - (B) TYPE: nucleic acid
- (ii) MOLECULE TYPE: DNA (artificial sequence)
- DESCRIPTION: PGgtpCY_nco Primer

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

gctggtcggt taatgggtga agctggg

27

INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 residues

-32-

(B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA (artificial sequence)
DESCRIPTION: PGgtpCY_V Primer

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

aggaggaagc ttggcgctat ggcacgaag g

31

INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 residues
(B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA (artificial sequence)
DESCRIPTION: ShBle_V Primer

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

gggcatgcaa ttcgagctcg gtacccg

27

INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 residues
(B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA (artificial sequence)
DESCRIPTION: TEF1_H Primer

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

cgactcacta taggaggaag cttggcgc

28

- 33 -

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE
PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A yeast strain which has been transformed by a recombinant DNA
sequence comprising a DNA sequence which upon expression in a suitable
host cell encodes at least one polypeptide with riboflavin biosynthetic activity
5 and which DNA sequence is transcriptionally linked to a promotor functional
in such yeast strain.

2. A yeast strain of claim 1 which belongs to the group of flavinogenic
yeasts which overproduce riboflavin under conditions of iron starvation.

3. A yeast strain of claim 2 which is selected from the following group:
10 Schwanniomyces, preferably Schwanniomyces occidentalis, Debaryomyces,
preferably Debaryomyces hansenii, Torulopsis, preferably Torulopsis candida,
or Candida, preferably Candida guilliermondii or Candida famata.

4. The yeast strain of claim 3 which is Candida guilliermondii or Candida
famata.

15 5. A yeast strain as claimed any one of claims 1 to 4 wherein the
polypeptide encoding DNA sequence is from yeast, preferably flavinogenic
yeasts which overproduce riboflavin under conditions of iron starvation more
preferably Candida, e.g. Candida guilliermondii or Candida famata.

6. A yeast strain as claimed in any one of claims 1 to 4, wherein the
20 polypeptide encoding DNA sequence encodes a protein with GTP
cyclohydrolase II activity and is selected from the following DNA sequences:

a) the DNA sequence as shown in Fig. 4 or its complementary strand;

b) DNA sequences which hybridize under standard conditions to the
protein coding regions of the DNA sequences defined in (a) or fragments
25 thereof; and

c) DNA sequences which, but for the degeneracy of the genetic code,
would hybridize to the DNA sequences defined in (a) and (b).

7. A yeast strain as claimed in any one of claims 1 to 6 wherein the
promotor is the TEF *S. cerevisiae* promotor.

30 8. A process for the production of riboflavin characterized therein that a
yeast strain as claimed in any one of claims 1 to 7 is cultured under suitable

- 34 -

culture conditions and the riboflavin produced is isolated from the medium or the yeast strain by methods known to the man skilled in the art.

9. A process for the production of a food or feed composition characterized therein that riboflavin as obtained by the process of claim 8 is mixed with one
5 or more suitable food or feed ingredients.

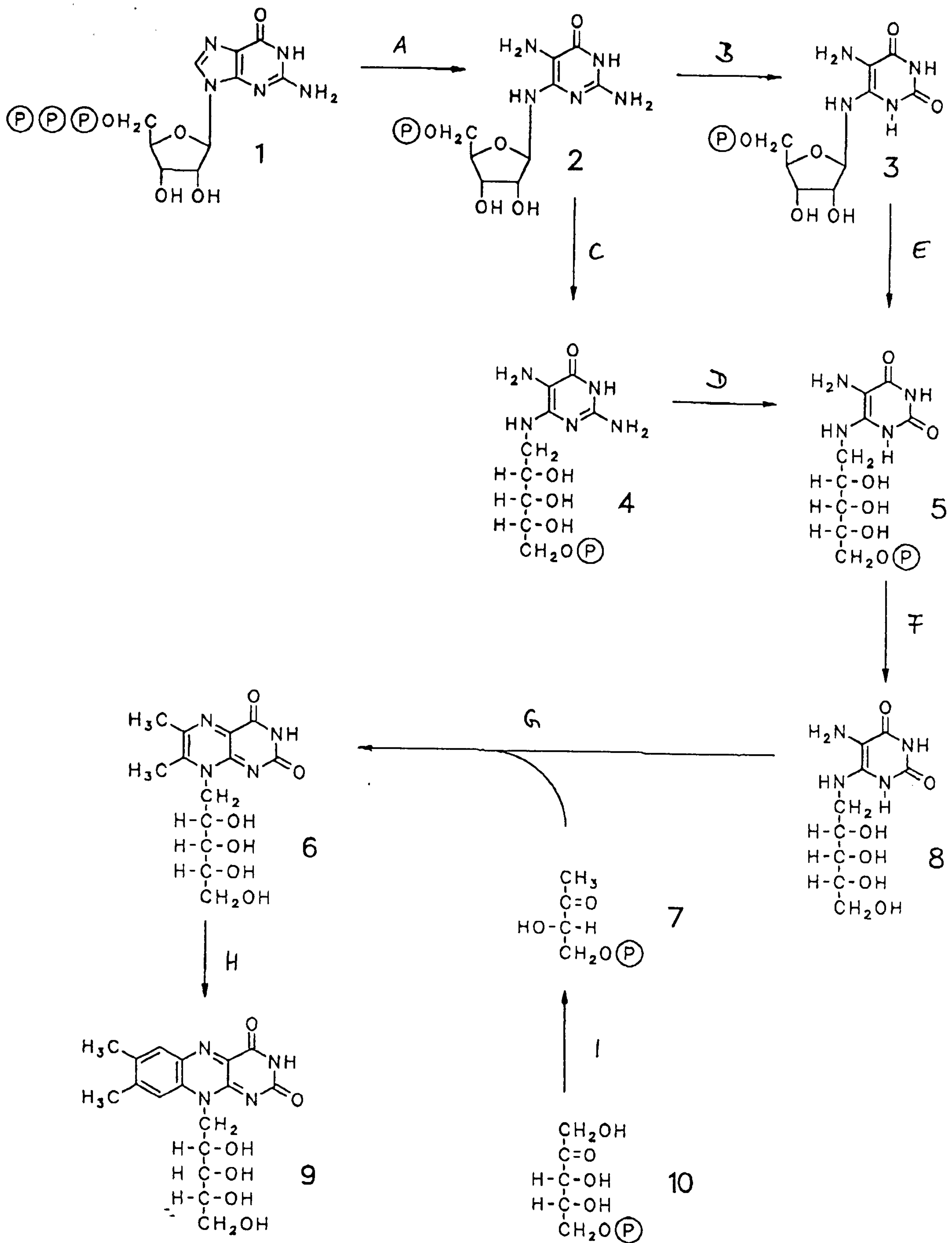


Fig. 1

Fig.2.
Nucleotide sequence of the p19R1 insert.

10	20	30	40	50	60
GTCGACTTTC	ACTCCGAAGG	TAGGTGCGGCT	GGAAGACGTC	GTCCCAAGTC	GTATGCGTT
70	80	90	100	110	120
AGCTGAGAGC	GACGGAACGA	AAGTGATGAAG	ATTACATGCT	GGAATAATCC	ATAGCTAG
130	140	150	160	170	180
TGTACTTGCT	AATAACAACCG	GTAAGCTAGCC	AATTGCAGCGT	TATTCACCACC	GCCGCTG
190	200	210	220	230	240
GATCGGGTTA	GTACGTGAACT	GGCCGTTGGGT	CCTGCACGTC	GCTTCATTATT	CATATA
250	260	270	280	290	300
TTAGTGAGAG	TCTTCCTATAT	CAGTCAGCAGAC	GTATCGGTTG	ATTTCAGGTCA	AAAAGA
310	320	330	340	350	360
GAAAAGGTGG	TCTTACAAAAG	CGAAATAGCTG	AACATTTTACT	CACAGCAGCAT	CATA
370	380	390	400	410	420
TTTGTGGAAC	CTTAAACTTG	ACTTTTCATTT	CAAGCAAGTT	ATTTTGAAATT	CAAATCA
430	440	450	460	470	480
TTTGGAATCA	AAAAAGAACAT	CTAAGTTCTG	AAAAATTGTAC	GAACAACGCT	ATGGCAT
					MetAlaSer
490	500	510	520	530	540
CGAAGGACAT	AGTACATCCG	CAACCAGAGC	CGCCGGCACG	GGTCGGAAACT	CACGAATTTA
LysAspIleValHisProGlnProGluArgArgHisGlySerGluThrHisGluPheThr					
550	560	570	580	590	600
CCATGCCTCT	CTTATCTCCT	AATTGACACCA	TCCCATATTCC	ATCGCAAACGC	CTCAAA
METProLeuLeuSerProThrLeuThrProSerHisIleProSerGlnThrProGlnIle					

- 2 -

610 620 630 640 650 660
| | | | | |
TTCCTCCGGAAGTGCCAGCAGAAGTCAGGGATCGCTTGCCCCTTCCTGAAACGTTGCCTG
ProProGluValProAlaGluValArgAspArgLeuProLeuProGluThrLeuProVal

670 680 690 700 710 720
| | | | | |
TGGTGAAATGCATGGCGAGAGCTCGTATACCGACCACTCAGGGGCCGGAGATATTTCTCC
ValLysCysMETAlaArgAlaArgIleProThrThrGlnGlyProGluIlePheLeuHis

730 740 750 760 770 780
| | | | | |
ATTTGTACGAGAATAACGTTGACAATAAAGAGCATTTGGCTATTGTTTTTGGGGAAGATG
LeuTyrGluAsnAsnValAspAsnLysGluHisLeuAlaIleValPheGlyGluAspVal

790 800 810 820 830 840
| | | | | |
TGCGGTCGAAAACGCTCTATCAGAAACGTCCCAATGAGACCCAGCAAGATAGAATGACTC
ArgSerLysThrLeuTyrGlnLysArgProAsnGluThrGlnGlnAspArgMETThrArg

850 860 870 880 890 900
| | | | | |
GTGGTGCTTATGTGGGCAGATTGTTTCCTGGAAGAACCGAGGCAGACTATGACAGTGAGT
GlyAlaTyrValGlyArgLeuPheProGlyArgThrGluAlaAspTyrAspSerGluSer

910 920 930 940 950 960
| | | | | |
CTAATTTGAGATTGAATTTTCGATGAAAATGGCCAACTTATCAGAGATCCGAGTACCACCT
AsnLeuArgLeuAsnPheAspGluAsnGlyGlnLeuIleArgAspProSerThrThrCys

970 980 990 1000 1010 1020
| | | | | |
GTAGTGGTGAGCCCATTTTGGCCCGTATTCATTCGGAATGTTATACGGGGGAAACCGCAT
SerGlyGluProIleLeuAlaArgIleHisSerGluCysTyrThrGlyGluThrAlaTrp

1030 1040 1050 1060 1070 1080
| | | | | |
GGAGTGCTCGTTGCGATTGTGGAGAACAATTCGATGAAGCTGGTCGGTTAATGGGTGAAG
SerAlaArgCysAspCysGlyGluGlnPheAspGluAlaGlyArgLeuMETGlyGluAla

- 3 -

1090 1100 1110 1120 1130 1140
| | | | |
CTGGGCACGGGTGTATCGTGTACCTTCGTCAGGAAGGTCGTGGAATTGGACTTGGGGAAA
GlyHisGlyCysIleValTyrLeuArgGlnGluGlyArgGlyIleGlyLeuGlyGluLys

1150 1160 1170 1180 1190 1200
| | | | |
AGTTGAAGGCTTATAATTTGCAAGACTTGGGAGCGGATACCGTCCAGGCCAATTTGATGT
LeuLysAlaTyrAsnLeuGlnAspLeuGlyAlaAspThrValGlnAlaAsnLeuMETLeu

1210 1220 1230 1240 1250 1260
| | | | |
TACGACATCCTGCTGATGCGAGATCTTTTTCGCTCGCTACAGCCATACTCTTGGACTIONTGG
ArgHisProAlaAspAlaArgSerPheSerLeuAlaThrAlaIleLeuLeuAspLeuGly

1270 1280 1290 1300 1310 1320
| | | | |
GGCTCAACGAGATCAAGTTGTTGACCAACAATCCCGATAAAATTGCTGCAGTTGAGGGAA
LeuAsnGluIleLysLeuLeuThrAsnAsnProAspLysIleAlaAlaValGluGlyArg

1330 1340 1350 1360 1370 1380
| | | | |
GAAACAGAGAGGTCAAGGTAGTGAACGGGTGCCTATGGTGCCGTTGGCATGGAGAAGTG
AsnArgGluValLysValValGluArgValProMETValProLeuAlaTrpArgSerGlu

1390 1400 1410 1420 1430 1440
| | | | |
AGAATGGAATCAAGTCAAAAGAGATAGAGGGCTACTTGAGTGCTAAGATTGAAAGGATGG
AsnGlyIleLysSerLysGluIleGluGlyTyrLeuSerAlaLysIleGluArgMETGly

1450 1460 1470 1480 1490 1500
| | | | |
GGCACTTGCTTGAAAAGCCACTCAAGATATGATAGAAGAGATGAAGTTAAGGACTTAAGA
HisLeuLeuGluLysProLeuLysIle-----

1510 1520 1530 1540 1550 1560
| | | | |
AATAAATGATGAATTAAATGACGCAAATGTCACTACTCGATTAGAGAAATAGCTATAATG

1570 1580 1590 1600 1610 1620
| | | | |
AAGAATTTTGCATTTTCGCAAAATTTAAGATAAATGCAAAAATTGCAAATTACGAAATATG

- 4 -

1630	1640	1650	1660	1670	1680
CATATGATACAAGACAAGAAAAGACTACTAAAAGTCTCTCGAGAAGAATACTGGGTAACC					

1690	1700	1710	1720	1730	1740
TTCATCTCTTGATTATGCACTGGGGCTATTCATATGCAGATTTCGCACGCCGAGGTGCAGC					

1750	1760	1770	1780	1790	1800
GTTTAGGCGCGGCTCAACGGAAGCCAACGGCCGCCACAAATTGTCCGGAAGTCGCCGAA					

1810	1820	1830	1840	1850	1860
ACTGATCCACTGGTACCACAGCCCCATAAGAACCCCTTTAATATTAA ~~AA~~ ACCGTTCTTC					

1870	1880	1890	1900	1910	1920
AGCCACTTTTGATCACATTGTTTGCAGCCGCCCGTTGCTGCCATCCAA ~~A~~ CACCACGCGTC					

1930	1940	1950	1960	1970	1980
CCCCGCACCTTTTACGGTGCCCACTGCATTGGAATTTGCATAAAACAGCCTCACGAAGTG					

1990	2000	2010	2020	2030	2040
GATTAATTTTTAGAGCACTCAAGTCATCATGCTGCAATCTCTGCATCATGAAATGACTCC					

2050	2060	2070	2080	2090	2100
CGTTGATACAGGGA ~~A~~ ACTCAGACCGCAAGCGGCGAAGAGTCACAAGAGCGTGTGATGTGTG					

2110	2120	2130	2140	2150	2160
TCGACTCTAGAGATCCCCGGGTACCGAGCTCGAATTCAGTGGCCGTCGTTTACAACGTC					

2170	2180
GTGACTGGGAAAACCCTGGCG	

Sequence of the *RIB1* gene is translated.

Sequence of ARS element is typed in bold letters.

Fig.3
Nucleotide sequence of the pTC2 insert.

```
      10      20      30      40      50      60
      |      |      |      |      |      |
CAATTCGAGCTCGGTACCCGGGGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTC

      70      80      90     100     110     120
      |      |      |      |      |      |
CTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACC

     130     140     150     160     170     180
     |      |      |      |      |      |
CAAGCACAGCATACTAAATTTCCCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTAC

     190     200     210     220     230     240
     |      |      |      |      |      |
TAAAGGTTTGGAAAAGAAAAAGAGACCGCCTCGTTTCTTTTCTTCGTCGAAAAGGCA

     250     260     270     280     290     300
     |      |      |      |      |      |
ATAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTGATTTTTTCTCTTT

     310     320     330     340     350     360
     |      |      |      |      |      |
CGATGACCTCCCATTGATATTTAAGTCAATAAACGGTCTTCAATTTCTCAAGTTTCAGTT

     370     380     390     400     410     420
     |      |      |      |      |      |
TCATTTTCTTGTTCTATTACAACCTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG

     430     440     450     460     470     480
     |      |      |      |      |      |
CAATCTAATCTAAGGGCGAGCTCGAATTCGAACTAGTACTGCAGCACGTGACCGGCGCCT

     490     500     510     520     530     540
     |      |      |      |      |      |
AGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAG

     550     560     570     580     590     600
     |      |      |      |      |      |
-GAGGAAGCTTGGCGCTATGGCATCGAAGTACATAGTACATCCGCAACCAGAGCGCCGGCA
```

METAlaSerLysTyrIleValHisProGlnProGluArgArgHis

- 2 -

610 620 630 640 650 660
CGGGTCGGAACTCACGAATTTACCATGCCTCTCTTATCTCCTACATTGACACCATCCCA
GlySerGluThrHisGluPheThrMETProLeuLeuSerProThrLeuThrProSerHis

670 680 690 700 710 720
TATTCCATCGCAAACGCCTCAAATTCCTCCCGAAGTCCAGCAGAAGTCAGGGATCGCTT
IleProSerGlnThrProGlnIleProProGluValProAlaGluValArgAspArgLeu

730 740 750 760 770 780
GCCCCTTCCTGAAACGTTGCCTGTGGTGAAATGCATGGCGAGAGCTCGTATACCGACCAC
ProLeuProGluThrLeuProValValLysCysMETAlaArgAlaArgIleProThrThr

790 800 810 820 830 840
TCAGGGGCCGAGATATTTCTCCATTTGTACGAGAATAACGTTGACAATAAAGAGCATT
GlnGlyProGluIlePheLeuHisLeuTyrGluAsnAsnValAspAsnLysGluHisLeu

850 860 870 880 890 900
GCCTATTGTTTTTGGCGAAGATGTGCGGTGCGAAACGCTCTATCAGAAACGTCCCAATGA
AlaIleValPheGlyGluAspValArgSerLysThrLeuTyrGlnLysArgProAsnGlu

910 920 930 940 950 960
GACCCAGCAAGATAGAATGACTCGTGGTGCTTATGTGGGCAGATTGTTTCCTGGAAGAAC
ThrGlnGlnAspArgMETThrArgGlyAlaTyrValGlyArgLeuPheProGlyArgThr

970 980 990 1000 1010 1020
CGAGGCAGACTATGACAGTGAGTCTAATTTGAGATTGAATTTGATGAAAATGGCCAACT
GluAlaAspTyrAspSerGluSerAsnLeuArgLeuAsnPheAspGluAsnGlyGlnLeu

1030 1040 1050 1060 1070 1080
TATCAGAGATCCGAGTACCACCTGTAGTGGTGAGCCCATTTTGGCCCGTATTCATTCGGA
IleArgAspProSerThrThrCysSerGlyGluProIleLeuAlaArgIleHisSerGlu

Fig.4**Nucleotide sequence of the pTCdXS-2 insert.**

```
      10      20      30      40      50      60
      |      |      |      |      |      |
CAATTCGAGCTCGGTACCCGGGATCCCCACACACCATAGCTTCAAAATGTTTCTACTC

      70      80      90     100     110     120
      |      |      |      |      |      |
CTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACC

     130     140     150     160     170     180
     |      |      |      |      |      |
CAAGCACAGCATACTAAATTTCCCTCTTTCTTCCTCTAGGGTGTCTGTTAATTACCCGTAC

     190     200     210     220     230     240
     |      |      |      |      |      |
TAAAGGTTTGGAAAAGAAAAAGAGACCGCCTCGTTTCTTTTTCTTCGTCGAAAAGGCA

     250     260     270     280     290     300
     |      |      |      |      |      |
ATAAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTTTGATTTTTTCTCTTT

     310     320     330     340     350     360
     |      |      |      |      |      |
CGATGACCTCCCATTGATATTTAAGTCAATAAACGGTCTTCAATTTCTCAAGTTTCAGTT

     370     380     390     400     410     420
     |      |      |      |      |      |
TCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG

     430     440     450     460     470     480
     |      |      |      |      |      |
CAATCTAATCTAAGGGCGAGCTCGAATTCGAAGTAGTACTGCAGCACGTGACCGGCGCCT

     490     500     510     520     530     540
     |      |      |      |      |      |
AGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAG

     550     560     570     580     590     600
     |      |      |      |      |      |
GAGGAAGCTTGGCGCTATGGCATCGAAGTACATAGTACATCCGCAACCAGAGCGCCGGCA
```

METAlaSerLysTyrIleValHisProGlnProGluArgArgHis

- 2 -

610 620 630 640 650 660
| | | | |
CGGGTCGGAACTCACGAATTTACCATGCCTCTCTTATCTCCTACATTGACACCATCCCA

GlySerGluThrHisGluPheThrMETProLeuLeuSerProThrLeuThrProSerHis

670 680 690 700 710 720
| | | | |
TATTCCATCGCAAACGCCTCAAATTCCTCCGGAAGTGCCAGCAGAAGTCAGGGATCGCTT

IleProSerGlnThrProGlnIleProProGluValProAlaGluValArgAspArgLeu

730 740 750 760 770 780
| | | | |
GCCCCTTCCTGAAACGTTGCCTGTGGTGAAATGCATGGCGAGAGCTCGTATACCGACCAC

ProLeuProGluThrLeuProValValLysCysMETAlaArgAlaArgIleProThrThr

790 800 810 820 830 840
| | | | |
TCAGGGGCCGAGATATTTCTCCATTTGTACGAGAATAACGTTGACAATAAAGAGCATTT

GlnGlyProGluIlePheLeuHisLeuTyrGluAsnAsnValAspAsnLysGluHisLeu

850 860 870 880 890 900
| | | | |
GGCTATTGTTTTTGGGGAAGATGTGCGGTGCGAAACGCTCTATCAGAAACGTCCCAATGA

AlaIleValPheGlyGluAspValArgSerLysThrLeuTyrGlnLysArgProAsnGlu

910 920 930 940 950 960
| | | | |
GACCCAGCAAGATAGAATGACTCGTGGTGCTTATGTGGGCAGATTGTTTCCTGGAAGAAC

ThrGlnGlnAspArgMETThrArgGlyAlaTyrValGlyArgLeuPheProGlyArgThr

970 980 990 1000 1010 1020
| | | | |
CGAGGCAGACTATGACAGTGAGTCTAATTTGAGATTGAATTTTCGATGAAAATGGCCAACT

GluAlaAspTyrAspSerGluSerAsnLeuArgLeuAsnPheAspGluAsnGlyGlnLeu

1030 1040 1050 1060 1070 1080
| | | | |
TATCAGAGATCCGAGTACCACCTGTAGTGGTGAGCCCATTTTGGCCCGTATTCATTTCGGA

IleArgAspProSerThrThrCysSerGlyGluProIleLeuAlaArgIleHisSerGlu

- 3 -

1090	1100	1110	1120	1130	1140
ATGTTATACGGGGGAAACCGCATGGAGTGCTCGTTGCGATTGTGGAGAACAAATTCGATGA					
CysTyrThrGlyGluThrAlaTrpSerAlaArgCysAspCysGlyGluGlnPheAspGlu					

1150	1160	1170	1180	1190	1200
AGCTGGTCGGTTAATGGGTGAAGCTGGGCACGGGTGTATCGTGTACCTTCGTCAGGAAGG					
AlaGlyArgLeuMETGlyGluAlaGlyHisGlyCysIleValTyrLeuArgGlnGluGly					

1210	1220	1230	1240	1250	1260
TCGTGGAATTGGACTTGGGGAAAAGTTGAAGGCTTATAATTTGCAAGACTTGGGAGCGGA					
ArgGlyIleGlyLeuGlyGluLysLeuLysAlaTyrAsnLeuGlnAspLeuGlyAlaAsp					

1270	1280	1290	1300	1310	1320
TACCGTCCAGGCCAATTTGATGTTACGACATCCTGCTGATGCGAGATCTTTTTCGCTCGC					
ThrValGlnAlaAsnLeuMETLeuArgHisProAlaAspAlaArgSerPheSerLeuAla					

1330	1340	1350	1360	1370	1380
TACAGCCATACTCTTGGACTTGGGGCTCAACGAGATCAAGTTGTTGACCAACAATCCCGA					
ThrAlaIleLeuLeuAspLeuGlyLeuAsnGluIleLysLeuLeuThrAsnAsnProAsp					

1390	1400	1410	1420	1430	1440
TAAAATTGCTGCAGTTGAGGGAAGAAACAGAGAGGTCAAGGTAGTGGAACGGGTGCCTAT					
LysIleAlaAlaValGluGlyArgAsnArgGluValLysValValGluArgValProMET					

1450	1460	1470	1480	1490	1500
GGTGCCGTTGGCATGGAGAAGTGAGAATGGAATCAAGTCAAAAGAGATAGAGGGCTACTT					
ValProLeuAlaTrpArgSerGluAsnGlyIleLysSerLysGluIleGluGlyTyrLeu					

1510	1520	1530	1540	1550	1560
GAGTGCTAAGATTGAAAGGATGGGGCACTTGCTTGAAAAGCCACTCAAGATATGATAGAA					
SerAlaLysIleGluArgMETGlyHisLeuLeuGluLysProLeuLysIle-----					

- 4 -

1570	1580	1590	1600	1610	1620
GAGATGAAGTTAAGGACTTAAGAAATAAATGATGAATTAAATGACGCAAATGTCACTACT					
1630	1640	1650	1660	1670	1680
CGATTAGAGAAATAGCTATAATGAAGAATTTTGCATTTTCGCAAATTTAAGATAAATGCA					
1690	1700	1710	1720	1730	1740
AAAATTGCAAATTACGAAATATGCATATGATACAAGACAAGAAAAGACTACTAAAAGTCTCT					

Sequence of the *TEF* promoter is underlined.
Sequence of the *RIB1* gene is translated.
Sequence of ARS element is typed in bold letters.

Fig.5

Nucleotide sequence of PCR product obtained with primers ShBle_V and PGgtpCY_nco using boiled cells of *P.guilliermondii* XS-3 recombinant strain.

```

      10      20      30      40      50      60
      |      |      |      |      |      |
GGGCATGCAATTCGAGCTCGGTACCCGGGGATCCCCCACACACCATAGCTTCAAAATGTT

      70      80      90     100     110     120
      |      |      |      |      |      |
TCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCA

      130     140     150     160     170     180
      |      |      |      |      |      |
AAACACCCAAGCACAGCATACTAAATTTCCCTCTTTCTTCCTCTAGGGTGTCGTTAATTA

      190     200     210     220     230     240
      |      |      |      |      |      |
CCCGTACTAAAGGTTTGGAAAAGAAAAAGAGACCGCCTCGTTTCTTTTCTTCGTCGAA

      250     260     270     280     290     300
      |      |      |      |      |      |
AAAGGCAATAAAAATTTTATCACGTTTCTTTTCTTGAAATTTTTTTTTTGATTTTTT

      310     320     330     340     350     360
      |      |      |      |      |      |
TCTCTTTCGATGACCTCCCATTGATATTTAAGTCAATAAACGGTCTTCAATTTCTCAAGT

      370     380     390     400     410     420
      |      |      |      |      |      |
TTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAA

      430     440     450     460     470     480
      |      |      |      |      |      |
AGCATAGCAATCTAATCTAAGGGCGAGCTCGAATTCGAACTAGTACTGCAGCACGTGACC

      490     500     510     520     530     540
      |      |      |      |      |      |
GGCGCCTAGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTC

      550     560     570     580     590     600
      |      |      |      |      |      |
ACTATAGGAGGAAGCTTGGCGCTATGGCATCGAAGTACATAGTACATCCGCAACCAGAGC

```

METAlaSerLysTyrIleValHisProGlnProGluArg

- 2 -

610 620 630 640 650 660
| | | | |
GCCGGCACGGGTCGGAACTCACGAATTTACCATGCCTCTCTTATCTCCTACATTGACAC
ArgHisGlySerGluThrHisGluPheThrMETProLeuLeuSerProThrLeuThrPro

670 680 690 700 710 720
| | | | |
CATCCCATATTCCATCGCAAACGCCTCAAATTCCTCCGGAAGTGCCAGCAGAAGTCAGGG
SerHisIleProSerGlnThrProGlnIleProProGluValProAlaGluValArgAsp

730 740 750 760 770 780
| | | | |
ATCGCTTGCCCCTTCCTGAAACGTTGCCTGTGGTGAAATGCATGGCGAGAGCTCGTATAC
ArgLeuProLeuProGluThrLeuProValValLysCysMETAlaArgAlaArgIlePro

790 800 810 820 830 840
| | | | |
CGACCACTCAGGGGCCGGAGATATTTCTCCATTTGTACGAGAATAACGTTGACAATAAAG
ThrThrGlnGlyProGluIlePheLeuHisLeuTyrGluAsnAsnValAspAsnLysGlu

850 860 870 880 890 900
| | | | |
AGCATTTGGCTATTGTTTTTGGGGAAGATGTGCGGTGCGAAAACGCTCTATCAGAAACGTC
HisLeuAlaIleValPheGlyGluAspValArgSerLysThrLeuTyrGlnLysArgPro

910 920 930 940 950 960
| | | | |
CCAATGAGACCCAGCAAGATAGAATGACTCGTGGTGCTTATGTGGGCAGATTGTTTCCTG
AsnGluThrGlnGlnAspArgMETThrArgGlyAlaTyrValGlyArgLeuPheProGly

970 980 990 1000 1010 1020
| | | | |
GAAGAACCGAGGCAGACTATGACAGTGAGTCTAATTTGAGATTGAATTTTCGATGAAAATG
ArgThrGluAlaAspTyrAspSerGluSerAsnLeuArgLeuAsnPheAspGluAsnGly

1030 1040 1050 1060 1070 1080
| | | | |
GCCAACTTATCAGAGATCCGAGTACCACCTGTAGTGGTGAGCCCATTTTGGCCCGTATTC
GlnLeuIleArgAspProSerThrThrCysSerGlyGluProIleLeuAlaArgIleHis

1090 1100 1110 1120 1130 1140
| | | | |
ATTCGGAATGTTATACGGGGGAAACCGCATGGAGTGCTCGTTGCGATTGTGGAGAACAAAT
SerGluCysTyrThrGlyGluThrAlaTrpSerAlaArgCysAspCysGlyGluGlnPhe

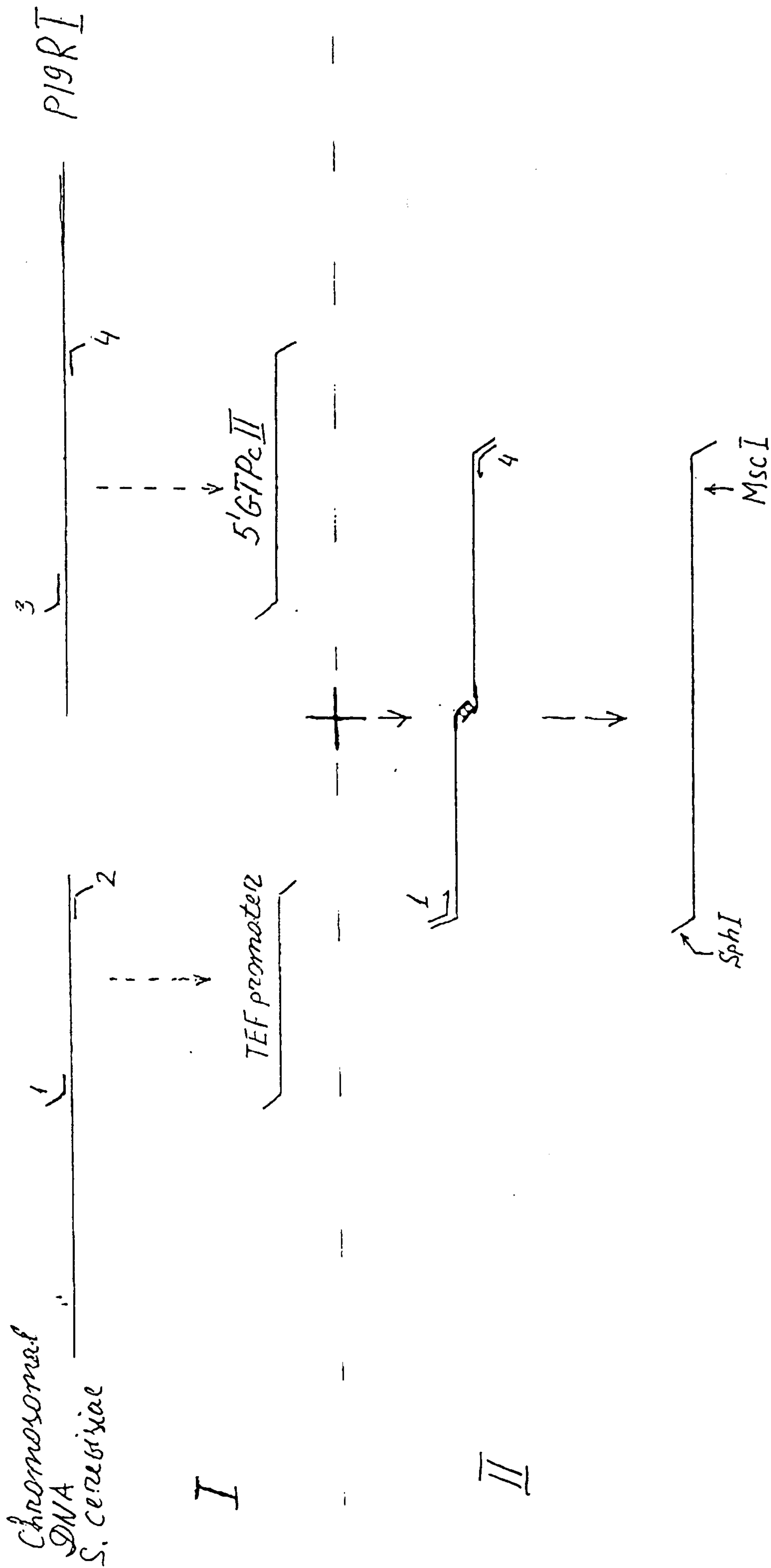
- 3 -

1150 1160 1170
| | |
TCGATGAAGCTGGTCGGTTAATGGGTGAAGCTGGG

AspGluAlaGlyArgLeuMETGlyGluAlaGly

Sequence of the *TEF* promoter is underlined.
Partial sequence of the *RIB1* gene is translated.

Scheme of the 2-Step PCR Amplification



1. ShBle-V
2. TEF1-H
3. PGgtPCY-V
4. PGgtPCY-ncO

FIG. 6