Title: YEAST/BACTERIAL TWO-HYBRID SYSTEM AND METHODS OF USE THEREOF

Abstract: A combined yeast/bacterial two-hybrid system is disclosed.
Yeast/Bacterial Two-hybrid System and Methods of Use Thereof

Ilya G. Serebriiskii
Erica A. Golemis
J. Keith Joung

This application claims priority to US Provisional Application 60/606,266 filed August 31, 2004, the entire disclosure of which is incorporated herein by reference.

Pursuant to 35 U.S.C. §202(c) it is acknowledged that the U.S. Government has certain rights in the invention described herein, which was made in part with funds from the National Institutes of Health, Grant Numbers RO1CA63366 and K08 DK02883.

FIELD OF THE INVENTION

This invention relates to the field of molecular biology. More specifically, the invention provides novel compositions and methods to facilitate the isolation and characterization of novel, protein-protein interactions involved in the regulation of cell growth and metabolism.

BACKGROUND OF THE INVENTION

Several publications and patent documents are cited throughout this application to better define the state of the art to which this invention pertains. Each of the foregoing citations is incorporated by reference herein.

Yeast two-hybrid systems (Chien et al. 1991; Fields and Song 1989; Gyuris et al. 1993; Vojtek et al. 1993) are standard tools used to identify novel protein-protein interactions and to perform structure-function analysis on previously defined protein-protein interactions. Such systems are effective with a substantial fraction of eukaryotic proteins and have played an important role in high throughput proteomic analyses aimed at establishing sets of interacting proteins (e.g. (Giot et al. 2003; Ito et al. 2000; Li et al. 2004; Uetz et al. 2000). In order to increase the power of a two-hybrid approach to identify and analyze protein interactions in high throughput applications, one approach has been to translate the basic components of the yeast
two-hybrid system to a bacterial host organism (Dove et al. 1997; Joung et al. 2000).
To date, the relative effectiveness of protein interaction detection in bacterial and
yeast backgrounds has not been directly compared. However, there are a number of
reasons to anticipate that differences might be observed. As yeast are eukaryotes,
eukaryotic proteins used as "baits" in two-hybrid screens may be more likely to be
appropriately folded and post-translationally modified in yeast than in bacteria,
thereby increasing their chances of identifying physiological partners. However,
certain proteins can be problematic as baits in the yeast two-hybrid system; for
example, proteins that are normally excluded from the nucleus in eukaryotes, that are
potentially sequestered via interaction with an abundant partner evolutionarily
conserved in yeast, or that stimulate transcription in yeast (i.e.—that "autoactivate").
All of these potential issues would be expected to be less problematic in the bacterial
two-hybrid system. To maximize chances of obtaining all relevant interactors for a
protein of interest, it would be desirable to have the capability to rapidly test a given
bait in both yeast and bacterial milieus.

SUMMARY OF THE INVENTION

In accordance with the present invention, plasmids and strains suitable for use
in both yeast and bacterial protein interaction systems are provided. A novel series of
vectors are disclosed in which a single plasmid containing a modified promoter drives
the efficient expression of a bait protein in either yeast or bacteria, thereby permitting
parallel studies in both organisms. In addition, optimized supporting yeast and
bacterial reporter strains are provided.

Thus, in one aspect of the invention, an isolated nucleic acid comprising a
promoter sequence of SEQ ID NO: 1 which drives expression of an operably linked
coding sequence in both yeast and bacteria is provided. Also provided are plasmids
comprising this promoter selected from the group consisting of SEQ ID NO: 1
(pGLS20), SEQ ID NO: 2 (pGLS22) and SEQ ID NO: 3 (pGLS23) and pBR-AMP-
aLPL (SEQ ID NO: 5). Host cells comprising these plasmids are also disclosed.

Such host cell are preferably E. coli and S. cerevisiae cells.

In yet another aspect of the invention, new strains of E. coli E.coli KJ1567 and
E. coli AG58A(RP28) are provided. New strains of yeast are also disclosed, e.g., S.
cerevisiae PRT50 and S. cerevisiae PRT475.
In a preferred embodiment of the invention, a method for comparing binding interactions between a first protein and a second protein in both bacterial and yeast organismal milieus using a constructs which function in both organisms, comprising

a) providing yeast and bacterial host cells, each comprising,

i) a reporter gene operably linked to a DNA sequence comprising a protein binding site;

ii) a first fusion gene which expresses a first fusion protein, said first fusion protein comprising said first protein covalently bonding to a binding moiety which is capable of specifically binding to said protein binding site which is driven by a the promoter element as claimed in claim 1 and

iii) a second fusion gene which expresses a second fusion protein, said second fusion protein comprising said second protein covalently bonded to gene activating moiety;

b) allowing said first and second proteins to interact; and

c) measuring expression of said reporter gene as a measure of said interaction between said first and second proteins in both E.coli and yeast.

Also provided in the present invention are kits for practicing the method described above. An exemplary kit comprises

a) a plasmid selected from the group consisting of (pGLS20), SEQ ID NO: 2 (pGLS22) and SEQ ID NO: 3 (pGLS23);

b) at least one of pAC-AMP-αLPL, and/or pBR-AMP- αLPL;

c) an E. coli strain selected from the group consisting of E.coli KJ1567 or E. coli AG58A(RP28); and

d) a yeast strain selected from the group consisting of PRT50, and

PRT475.
BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Schematics of the Yeast and Bacterial two hybrid systems. A. In the yeast two-hybrid system shown, a dimeric λ cl-bait hybrid protein interacts with an activation domain (AD)-prey hybrid protein thereby stimulating transcription from an adjacent promoter that directs expression of a quantitative GusA or selectable LYS2 reporter gene. B. In the bacterial two-hybrid system shown, a dimeric λ cl-bait hybrid protein interacts with an E. coli RNA polymerase (RNAP) α-subunit-prey hybrid protein, thereby recruiting RNAP to an adjacent promoter that directs expression of a quantitative lacZ or selectable HIS3 reporter gene. Note that both systems utilize a λ cl-bait hybrid protein from a single plasmid effective in either organism. C-E. Sequence in pGLS20 (SEQ ID NO: 2). F-H. Sequence in pGLS22 (SEQ ID NO: 3). I-J. Sequence in pBR_AMP_alphaLP-IS-B.gb (SEQ ID NO: 5).

Figure 2. Bait Expression from a combined bacterial/yeast expression plasmid. A. Plasmid pGLS20 and pGLS23 use a combined TEF1/uvLac promoter to express λ cl fused baits in yeast or bacteria. Plasmids are selected in yeast by selection for G418 resistance (pGLS20) or HIS5 complementation (pGLS23), and in bacteria by selection for Kanamycin resistance (pGLS20) or chloramphenicol resistance (pGLS23). Relative expression of cl baits from these plasmids, versus the previously described pGBS10 (yeast two-hybrid, (Serebriiskii et al. 2002)) or pBT (bacterial two-hybrid, Stratagene) vectors is shown in bacteria (center panel). B. To demonstrate relative bait levels, equal total protein concentration was confirmed by Coomassie staining of a PAGE gel loaded with equivalent amounts of cell lysate for bacteria expressing each plasmid (not shown). Then, equal volumes of 1:40 (for pGLS20) or 1:100 (for pBT) dilutions of extracts in sample buffer were loaded in parallel with the same volume of undiluted extract from pGBS10-bearing cells. Western blots using anti-cl antibodies are shown. C. pGBS10 and pGLS20 express comparable levels of λ cl baits in yeast, based on Western analysis with antibodies to λ cl. 1, 2 denotes two independent transformants in bacteria or yeast; -, denotes yeast containing no bait plasmid.
Figure 3. Activation of colorimetric and auxotrophic reporters by zipper interaction in yeast. Lane numbers below bar graph represent pairs of samples defined in Table 1. Bar graph reflects relative reporter activity measured by beta-glucuronidase assay using PNP-gluc as a substrate. Inset, indicated samples re-analyzed using MU-gluc as a substrate. For context, values obtained for combination 6 (with a Kd of $1 \times 10^{-8}$ M), were more than 10-fold higher than those with combination 5 with the MU-gluc substrate, indicating a significant discriminating function of the yeast two-hybrid system in this affinity range (not shown). Shown below bar graph is the growth of two representative spots of colonies 2 days after plating to selective medium.

Figure 4. Activation of colorimetric and auxotrophic reporters by zipper interaction in bacteria. Lane numbers below bar graph represent pairs of samples defined in Table 1. Bar graph reflects relative reporter activity measured by $\beta$-galactosidase assay using ONPG as a substrate. Shown below bar graph is the growth of two representative spots of colonies 24 days after plating to selective medium.

Figure 5. Bacterial prey plasmid map of pBR_AMP_alphaLP_IS_B.gb. This plasmid differs from pAC-AMP-αLPL in that it contains a different origin of replication, pBR as opposed to pAC.

Figure 6. The hybrid promoter sequence of SEQ ID NO: 1.

DETAILED DESCRIPTION OF THE INVENTION

Two-hybrid screening is a standard methodology to identify and characterize protein-protein interactions and has become an integral component of many proteomic investigations. The two-hybrid system was initially developed using yeast as a host organism. However, bacterial two-hybrid systems have also become common laboratory tools and are preferred in some circumstances, although yeast and bacterial two-hybrid systems have never been directly compared. In accordance with the present invention, a unified yeast and bacterial two-hybrid system is provided in which a single bait expression plasmid is used in both organismal milieus. Additionally, an extensive series of leucine zipper fusion proteins of known affinities were generated to compare the efficiency of interaction detection using both systems.
While both two-hybrid systems detected interactions occurring with a comparable range of interaction affinities, each demonstrated unique advantages. The yeast system produced quantitative readout over a greater dynamic range than that observed with bacteria. However, the phenomenon of “auto-activation” by baits was far less problematic in the bacterial system than in yeast. The ability to rapidly shift between yeast and bacterial systems provided by these new reagents provides a marked advantage for two-hybrid investigations. In addition, the modified expression vectors should be useful for any application requiring facile expression of a protein of interest in both yeast and bacteria. Conventional two hybrid systems have been disclosed in US Patents 5,580,736 and 6,326,150, the contents of which are incorporated herein by reference. Also provided in the present invention are kits useful for performing the methods disclosed herein.

The following definitions are provided to facilitate an understanding of the present invention.

As used herein, “reporter gene” refers to a gene whose expression may be assayed; such genes include, without limitation, LacZ, β-glucuronidase (GUS), amino acid biosynthetic genes, e.g., the yeast LEU2, HIS3, LYS2, or URA3 genes, nucleic acid biosynthetic genes, the mammalian chloramphenicol transacetylase (CAT) gene, the green fluorescent protein (GFP) or any surface antigen gene for which specific antibodies are available.

A "promoter" is a DNA sequence located proximal to the start of transcription at the 5' end of an operably linked transcribed sequence. The promoter may contain one or more regulatory elements or modules which interact in modulating transcription of the operably linked gene.

"Operably linked" describes two macromolecular elements arranged such that modulating the activity of the first element induces an effect on the second element. In this manner, modulation of the activity of a promoter element may be used to alter and/or regulate the expression of an operably-linked coding sequence. For example, the transcription of a coding sequence that is operably-linked to a promoter element is induced by factors that "activate" the promoter's activity; transcription of a coding sequence that is operably-linked to a promoter element is inhibited by factors that "repress" the promoter's activity. Thus, a promoter region is operably-linked to the coding sequence of a protein if transcription of such coding sequence activity is influenced by the activity of the promoter.
"Fusion construct" refers generally to recombinant genes which encode fusion proteins.

A "fusion protein" is a hybrid protein, i.e., a protein which has been constructed to contain domains from at least two different proteins. As used herein, a fusion protein is a hybrid protein which possesses (a) transcriptional regulatory domain from a transcriptional regulatory protein, or (b) a DNA binding domain from a DNA binding protein linked to a heterologous protein to be assayed for interaction. The structure of the fusion protein is such that the transcriptional regulatory domain and the DNA binding domain are arranged in a manner that allows both domains to be biologically active. The protein that is the source of the transcriptional regulatory domain is different from the protein that is the source of the DNA binding domain. In other words, the two domains are heterologous to each other.

The transcriptional regulatory domain of the fusion protein may either activate or repress transcription of target genes, depending on the native biological activity of the domain. The bait proteins of the invention are also fusion proteins encoded by a fusion gene which comprises a protein of interest operably linked to a DNA binding moiety.

The term "fusion protein gene" refers to a DNA sequence which encodes a fusion protein. A fusion protein gene may further provide transcriptional and translational regulatory elements for the transcriptional and translational control thereof.

"Expression" is the process by which the information encoded within a gene is revealed. If the gene encodes a protein, expression involves both transcription of the DNA into mRNA, the processing of mRNA (if necessary) into a mature mRNA product, and translation of the mature mRNA into protein.

A nucleic acid molecule, such as a DNA or gene is said to be "capable of expressing" a polypeptide if the molecule contains the coding sequences for the polypeptide and the expression control sequences which, in the appropriate host environment, provide the ability to transcribe, process and translate the genetic information contained in the DNA into a protein product, and if such expression control sequences are operably-linked to the nucleotide sequence that encodes the polypeptide.

As used herein, a "cloning vehicle" is any entity that is capable of delivering a nucleic acid sequence into a host cell for cloning purposes. Examples of cloning
vehicles include plasmids or phage genomes. A plasmid that can replicate autonomously in the host cell is especially desired. Alternatively, a nucleic acid molecule that can insert (integrate) into the host cell’s chromosomal DNA is useful, especially a molecule which inserts into the host cell’s chromosomal DNA in a stable manner, that is, a manner which allows such molecule to be inherited by daughter cells.

Cloning vehicles are often characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the vehicle, and into which DNA may be spliced in order to bring about its replication and cloning.

The cloning vehicle may further contain a marker suitable for use in the identification of cells transformed with the cloning vehicle. For example, "a marker gene" may be a gene which confers resistance to a specific antibiotic on a host cell.

The word "vector" is sometimes used interchangeably with "cloning vehicle".

As used herein, an “expression vehicle” is a vehicle or vector similar to the cloning vehicle but is especially designed to provide an environment which allows the expression of the cloned gene after transformation into the host. One manner of providing such an environment is to include transcriptional and translational regulatory sequences on such expression vehicle, such transcriptional and translational regulatory sequences being capable of being operably linked to the cloned gene. Another manner of providing such an environment is to provide a cloning site or sites on such vehicle, wherein a desired cloned gene and desired expression regulatory elements may be cloned.

In an expression vehicle, the gene to be cloned is usually operably-linked to certain control sequences such as promoter sequences. Expression control sequences will vary depending on whether the vector is designed to express the operably-linked gene in a prokaryotic or eukaryotic host and may additionally contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements, and/or translational initiation and termination sites.

A “host” refers to any organism that is the recipient of a cloning or expression vehicle. In preferred embodiments, the host of the invention is a yeast cell or a cultured animal cell such as a mammalian or insect cell. In an especially preferred embodiment, the yeast host is Saccharomyces cerevisiae.
A “binding moiety” is a stretch of amino acids which is capable of directing specific polypeptide binding to a particular DNA sequence (i.e., a “protein binding site”). Also referred to herein as a DNA binding domain, these proteins may be homodimers or monomers that bind DNA in a sequence specific manner. Exemplary DNA binding domains of the invention include LexA, cI, glucocorticoid receptor binding domains and the Ume6 domain.

A “gene activating moiety” is a stretch of amino acids which is capable of weakly inducing the expression of a gene to whose control region it is bound. As used herein, “weakly” is meant below the level of activation effected by GAL4 activation region II (Ma and Ptashne, Cell, 48: 347, 1987) and is preferably at or below the level of activation effected by the B42 activation domain of Ma and Ptashne (Cell, 51: 413, 1987). Levels of activation may be measured using any downstream reporter gene system and comparing, in parallel assays, the level of expression stimulated by the GAL4 region II-polypeptide with the level of expression stimulated by the polypeptide to be tested.

“Purified DNA” is DNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one of the 5' end and one of the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

“Substantially identical”, in reference to an amino acid sequence, means an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the function of the protein (assayed, e.g., as described herein). A “substantially identical” nucleic acid sequence codes for a substantially identical amino acid sequence as defined above.
A “transformed cell” is a yeast or bacterial cell into which (or into an ancestor of which) exogenous DNA has been introduced by means of recombinant DNA techniques.

The phrase “positioned for expression” refers to a DNA coding molecule which is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence.

A “purified antibody” is an antibody at least 60 weight percent of which is free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation comprises antibody in an amount of at least 75 weight percent, more preferably at least 90 weight percent, and most preferably at least 99 weight percent.

The following examples are provided to facilitate an understanding of the present invention. They are not intended to limit the invention in any way.

**EXAMPLE I**

The following materials and methods are provided to facilitate the practice of the present example.

*Molecular and Microbiological Manipulation.* Cloning of novel constructs was performed using conventional protocols. Details of the sequences and cloning sites encompassed in the plasmids described in the Results section, as well as other basic characterizations of expression properties of these plasmids, are provided hereinafter.

Briefly, plasmid pGLS20 was constructed by replacing the ADH1 promoter of pGKS9 with a combination of the *TEF1* promoter (from the pLexZeo plasmid, Invitrogen) and a lacUV5 promoter (from the pBT plasmid, Stratagene). To produce pGLS23, a HIS5CmR cassette was constructed in pCR2.1 vector by combining a *HIS5* cassette from pJFK (R. Hopkins, unpublished) and a CmR cassette from pMW108. This cassette was then used to replace the G418R cassette in pGLS20. Consequently, the difference between pGLS22 and pGLS23 is minor – a change in position 3267 of pGLS22 (gaattC --&gt; gaattA) makes it pGLS23 by destroying an EcoRI site. Fig.4 shows results using the pGLS23 (no EcoRI site) vector.
The bacterial two-hybrid prey plasmid pAC-AMP-αLPL was constructed by replacing the chloramphenicol resistance gene present in plasmid pKJ1267 (J.K.J., unpublished) with the ampicillin resistance gene from plasmid pACYC177. Leucine zipper sequences were chosen from among peptides described in (Krylov et al. 1998; Krylov et al. 1994; Moll et al. 2001): DNA was synthesized artificially to encode the described peptide sequences. To fuse the various leucine zippers to the amino-terminal domain and inter-domain linker of the E. coli RNA polymerase-α subunit, DNA fragments encoding the zipper variants were inserted into the plasmid using unique Not I and Xho I restriction sites.

Bait and prey expression. Expression of bait and prey proteins (except for bacterial RNA polymerase-α fusions, for which no antibody was available) was confirmed by Western analysis, with primary antibody to cl for baits, or hemagglutinin for preys expressed in yeast. While cl is the bait exemplified herein many different baits are known and available to the skilled artisan. To compare expression levels of cl protein in E. coli, corresponding plasmids were transformed into the DH5α strain and protein extracts prepared from exponentially growing cultures. Equal protein concentration was confirmed by Coomassie staining of a PAGE gel, then equal volumes of 1:40 (for pGLS20) or 1:100 (for pBT) dilutions of extracts in sample buffer were loaded in parallel with the same volume of undiluted extract from pGLS10-bearing cells. Proteins were resolved on a PAGE gel, and Western blot analysis was performed, using anti-cl antibodies. To compare expression levels of cl protein in yeast, corresponding plasmids were transformed in SKY191 strain and protein extracts prepared from the exponentially growing cultures. Equal protein concentration was confirmed by Coomassie staining of a PAGE gel (not shown). Then, equal volumes of extracts in sample buffer were loaded on the gel, and Western blot analysis was performed.

Reporter assays. For yeast, the activity of quantitative reporters was determined on a plate reader using a technique modified from Serebriiskii et al. 2000. Briefly, 50 μl of cultures exponentially growing in the wells of 96-well cells was added an equal volume 2 x Z-buffer containing 2mg/ml of the corresponding substrate and 50% Y-PER (Pierce), for yeast. Activity was calculated as (OD_{420f}-OD_{420i}) divided by OD_{600}, where the difference between OD_{420i} and OD_{420f} (initial and final readings) reflects the
conversion of the colorless substrate (PNPGluc) into yellow product over a period of time from ~10-30 minutes, and OD_{600} is a measure of cell density in a given sample. For each data point for each yeast experiment, activities of 5 to 8 clones were measured and averaged. All readings were taken in a plate reader; it was previously shown (Serebriiskii et al. 2000) that plate reader measurements and derivative units are proportionally correlated with the OD units taken on a spectrophotometer.

For bacterial β-galactosidase reporter gene measurements, assays were performed essentially as described (Thibodeau et al. 2004). Briefly, cultures inoculated from a fresh single colony were grown to mid-log phase and lysed by adding 1/10 volume PopCulture™ (Novagen). In a 96 well microtiter plate, 15 μl cell lysate was added to a mixture of 135 μl Z buffer and 30 μl 4 mg/ml ONPG to start the reaction. Kinetic assays were carried out by monitoring OD_{415} from 0-30 minutes using a plate reader. All bacterial β-galactosidase assays were performed in triplicate.

Auxotrophic reporters were assayed as described in (Serebriiskii and Joung 2002). Bait and prey plasmids were transformed into corresponding selection strain, S. cerevisiae SKY191 or E. coli KJ1567. Growth on selection plates was measured over 5 days (yeast; note, all colonies that grew were prominent at 2 days) or 1 day (bacteria).

RESULTS

We have developed plasmids which facilitate expression and parallel screening of a single bait protein in either a yeast or bacterial two-hybrid system using a single expression plasmid (Figure 1). As shown in Figure 1, bait proteins are expressed as fusions to the λ ci protein in both the yeast and bacterial two-hybrid systems. To enable this, we made several modifications to the plasmid pGBS9 (Serebriiskii et al. 2002), originally developed to express bait proteins as fusions to the λ ci repressor in a yeast two-hybrid system (Figure 2A). The ADHI promoter from this plasmid was replaced with a tandem promoter, in which the extremely powerful TEF1 promoter (Nagashima et al. 1986) from S. cerevisiae and the E.coli lacUV5 promoter both direct expression of a λ ci coding sequence and polylinker cloning site. The resulting plasmid, pGLS20, can be maintained in yeast or bacteria based on G418 or kanamycin resistance, respectively (Figure 2A). Other closely related plasmid
derivatives (pGLS22, pGLS23) harbor the HIS5 gene to confer selection in yeast, and chloramphenicol resistance for selection in bacteria (Figure 2A). As shown in Figure 2B, expression of λ cl repressor using plasmid pGLS20 in bacteria is comparable to that obtained with plasmid pBT (a vector optimized for the bacterial two-hybrid system, Stratagene) and is more than 40-fold higher than that provided by the standard yeast two-hybrid expression plasmid pGBS10 (Serebriiskii et al. 2002). In yeast, expression of cl repressor fusions from pGLS20 and its derivatives is comparable to or exceeds that from pGBS10 (Figure 2C).

We used these bi-functional pGLS plasmids to determine whether the yeast and bacterial two-hybrid systems exhibited any differences in their abilities to detect a series of interactions with differing affinities. To do this, we created a series of bait and prey fusion proteins using a set of previously characterized leucine zipper variants (Krylov et al. 1998; Krylov et al. 1994; Moll et al. 2001) with defined interaction affinities ranging from Kd >10^{-4} to 10^{-15} M as determined in vitro (Table 1). For analysis in the bacterial two-hybrid system, plasmid pAC-AMP-αLPL (Table 2) was used to express preys from the strong inducible lpp/lacUV5 tandem promoter as fusions to the amino-terminal domain of the RNA polymerase α subunit. For the yeast two-hybrid system, pJG4-5 (Gyuris et al. 1993) was used to express preys from the inducible GAL1 promoter as fusions to the synthetic transcriptional activation domain B42 (Figure 1). The ability of each zipper pair to activate transcription of a quantitative and an auxotrophic reporter was then assessed in bacteria and in yeast.

Table 1. Properties of leucine zippers used in this Example.

<table>
<thead>
<tr>
<th>Combination</th>
<th>Bait</th>
<th>pI</th>
<th>Prey</th>
<th>Kd for bait-prey (in M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EE_{12345}L</td>
<td>4.2</td>
<td>EE_{12345}L</td>
<td>Not detectable</td>
</tr>
<tr>
<td>2</td>
<td>RR_{12}EE_{345}L</td>
<td>6.5</td>
<td>RR_{12}EE_{345}L</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>EE_{34}</td>
<td>5.3</td>
<td>EE_{34}</td>
<td>8.1 x 10^{-4}</td>
</tr>
<tr>
<td>4</td>
<td>RR_{34}</td>
<td>10.5</td>
<td>RR_{34}</td>
<td>3.9 x 10^{-5}</td>
</tr>
<tr>
<td>5</td>
<td>RR_{1234} L</td>
<td>11.8</td>
<td>RR_{1234} L</td>
<td>2.5 x 10^{-7}</td>
</tr>
<tr>
<td>6</td>
<td>RR_{34}</td>
<td>10.5</td>
<td>EE_{34}</td>
<td>1.0 x 10^{-8}</td>
</tr>
<tr>
<td>7</td>
<td>EE_{34}</td>
<td>5.3</td>
<td>RR_{34}</td>
<td>1.0 x 10^{-8}</td>
</tr>
</tbody>
</table>

Leucine zippers for many of the preys were originally described, and interaction properties characterized in vitro in (Krylov et al. 1994).
Table 2. Strains and plasmids used in this study.
* reagent produced in this study. pRG61 (Serebriiskii et al. 2002), pDR8 (Serebriiskii et al. 2002), SKY191 (Serebriiskii et al. 1999) and pJG4-5 (Gyuris et al. 1993) have been described.

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Selection in yeast / in E.coli</th>
<th>Comment/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR12EE345L</td>
<td>6.5</td>
<td>RR12EE345L 1.3 x 10^{-11}</td>
</tr>
<tr>
<td>EE12RR345L</td>
<td>10.4</td>
<td>RR12EE345L 1.3 x 10^{-11}</td>
</tr>
<tr>
<td>RR12345L</td>
<td>12.2</td>
<td>EE12345L 1.1 x 10^{-11}</td>
</tr>
<tr>
<td>RR1234 L</td>
<td>11.8</td>
<td>EE1234 L 1.0 x 10^{-15}</td>
</tr>
<tr>
<td>EE1234 L</td>
<td>4.3</td>
<td>RR1234 L 1.0 x 10^{-15}</td>
</tr>
</tbody>
</table>

Baits

- pGLS20* G418\(^R\) Km\(^R\) TEF promoter ensures expression of cl in yeast, while lpp/lacUV5 promoter provides for expression in E. coli
- pGLS22/23* HIS5 Cm\(^R\) Similar to pGLS20, see text for details

Reporters

- pRG61 URA3* (Km\(^R\)) \(\lambda cl\) operators direct transcription of the gusA gene; pRG61 is less sensitive and lower background reporter than pDR8.

Activation Domain Fusions

- pJG4-5 TRP1 Ap\(^R\) \(GAL1\) promoter provides efficient expression in yeast of a gene fused to a cassette consisting of nuclear localization sequence, transcriptional activation domain, and HA epitope tag.
- pAC-AMP-\(\alpha\)LPL* N/A Ap\(^R\) In each of the two prey plasmids, tandem lpp/lacUV5 promoters provide efficient expression in E.coli of a gene fused to E. coli RNAP alpha subunit residues 1-248. pAC-AMP-\(\alpha\)LPL has pACYC origin of replication, while pBR-AMP-\(\alpha\)LPL has the pBR322 origin of replication. Hence, copy numbers of the bait/prey plasmids, and therefore bait/prey expression levels can be regulated.
- pBR-AMP-\(\alpha\)LPL* N/A Ap\(^R\) Strains

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Comment/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.cerevisiae SKY191</td>
<td>MAT(\alpha) trpl, his3, ura3, clon-LYS2</td>
<td>Reporter strains in which the expression of the LYS2 reporter gene is directed by a weak promoter bearing a (\lambda cl) DNA binding site.</td>
</tr>
<tr>
<td>S.cerevisiae PRT50*</td>
<td>MAT(\alpha) trpl, his3, his5, ura3, clon-LYS2</td>
<td></td>
</tr>
</tbody>
</table>
Our results in the yeast-based system demonstrate that zipper bait-prey combinations activate transcription of a quantifiable β-glucuronidase (GusA) reporter over a substantial range of affinities (Figure 3, bar graph). In this assay, zipper pairs with reported interaction dissociation constants of $1 \times 10^{-8}$ or lower (lanes 6-12) strongly activated reporter gene expression, as detected using a colorimetric substrate (PNP-gluc). Those with Kd values of $2.5 \times 10^{-7}$ M or higher (with one exception—see below) did not strongly activate the reporter gene (Figure 3A, lanes 1-5). β-glucuronidase activity was generally induced ~30-180 fold over baseline values with the higher affinity leucine zipper pairs. Additional testing of the lower affinity interacting pairs using a more sensitive fluorescent substrate for beta-glucuronidase, MU-gluc (Figure 3A, inset), indicated that it was also possible to convincingly detect interactions in the range of $10^{-7}$ M, although the stimulation of GusA gene expression seen in these samples is markedly less strong than those obtained with interactions in the $10^{-8}$ M range. With the auxotrophic reporter strain (Figure 3, panels below bar graph), cells grew under selective conditions only if the interacting zippers possessed dissociation constants of $1 \times 10^{-8}$ M or lower, paralleling the results obtained with the quantitative GusA reporter. The system did not have significant ability to discriminate interactions that interact with dissociations constants of $10^{-8}$ M or lower, suggesting the expression of the reporter gene was saturated. Importantly, for some of the baits examined, expression of the bait alone in the absence of the prey was sufficient to strongly activate transcription of the reporters, making it difficult to convincingly demonstrate protein interaction (see Figure 3, samples 1, 10, and 12).

We next examined the abilities of the same zipper bait-prey combinations to activate transcription in the bacterial two-hybrid system (Figure 1) using the quantifiable lacZ reporter (Figure 4). Consistent with our results in the yeast-based system, leucine zipper pairs with reported dissociation constants lower than $10^{-8}$ M clearly stimulated expression of the lacZ reporter gene (Figure 3, samples 6-12).
whereas interaction pairs with dissociation constants \(2.5 \times 10^{-7}\) M or higher failed to stimulate \(lacZ\) expression (Figure 3, samples 1-5). We also analyzed zipper-based activation of the auxotrophic reporter \(HIS3\) (Figure 4, panels below bar graph). Results obtained using the auxotrophic \(HIS3\) reporter gene closely paralleled those obtained with the \(lacZ\) reporter: only cells harboring zipper pairs with dissociation constants of \(1 \times 10^{-8}\) M or lower showed growth after 24 hours on selective plates. In contrast to the results obtained in the yeast-based system, none of the baits tested exhibited autoactivation in the absence of prey partners (compare samples 1, 10 and 12 in Figures 3 and 4).

Our results using a closely related set of small leucine zipper bait and prey fusions suggest differential advantages for detecting protein-protein interactions in the yeast and bacterial two-hybrid systems. First, our results using quantifiable reporters suggest that the yeast-based system possesses a broader dynamic range for detecting interactions (contrast Figures 3 and 4). In the yeast system, interactions characterized by dissociation constants as high as \(10^{-8}\) M could be detected as an increase in \(GusA\) reporter gene expression (or as high as \(10^{-7}\) M if a more sensitive substrate for \(GusA\) detection was used). In contrast, in the bacterial system, only interactions characterized by dissociation constants \(10^{-8}\) M or lower could be detected as an increase in \(lacZ\) expression. Second, we note that the experiments performed using bacterial two-hybrid system yield colonies on selective medium somewhat more quickly than those done in the yeast system (one day versus two). Third, our results also suggest that autoactivation by bait proteins is likely to be less problematic in bacteria than in yeast (compare lanes 1, 10, 12 in Figures 3 and 4). This finding is not entirely surprising given the fundamental differences in mechanisms of gene activation and the evolutionary distance between prokaryotes and eukaryotes. The lower frequency of auto-activation by baits in the bacterial two-hybrid system is a potentially significant advantage of this system compared with its yeast counterpart.

Our data also suggest that the threshold interaction strength required for robust transcriptional activation is similar in both organisms. In both the yeast and bacterial systems, full activation appears to require an interaction affinity between bait and prey fusion proteins defined by a dissociation constant in the \(10^{-7}\) to \(10^{-8}\) M range. Although our results demonstrate a sharp transition between no activation and full activation of the reporter genes, previous studies in both systems have demonstrated that the magnitude of transcriptional activation observed can be correlated with the
affinity of the bait-prey interaction (Dove et al. 1997; Estojak et al. 1995). While we
do not know the precise reason for this difference in our results compared with
previous studies, we note that Estojak and coworkers assessed interactions using a
series of reporters of varying stringency (i.e. containing differing numbers of binding
sites for the baits) to expand the detection range: there is no technical limitation to
using a similar strategy with this new system. Overall, our results strongly suggest
that use of the current system as a selection tool will work best for detecting
interactions with dissociation constants in the mid-to-high nanomolar range.

We note that, to our knowledge, this is the first description of a promoter
combination that is potent in both yeast and bacterial milieus. In fact, we have found
that our pGLS plasmids express sufficient levels of bait fusion proteins for activity in
the bacterial two-hybrid system even without inducing the strong bacterial promoter.
Levels of bait expression can be further regulated by the choice of prey plasmids:
pBR-AMP-αLPL shares the same pBR origin of replication with the pGLS bait
plasmids. Hence, by co-transformation in E.coli, the total number of copies of bait is
lower, which might be advantageous if the overexpression of the bait protein is
deleterious for the E.coli cell. Using pAC-AMP-αLPL, which has a different
(pACYC) origin of replication, allows the full copy number of bait plasmid and
therefore a potential for higher expression levels. Lastly, while this work focuses on
the use of the pGLS plasmids in a two-hybrid context, we anticipate that our general
promoter design might also be useful in other functional characterization studies.

References.

a method to identify and clone genes for proteins that interact with a protein of


While this invention has been disclosed with reference to specific embodiments, other embodiments and variations of this invention may be devised by those of skill in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.
What is claimed is:

1. An isolated nucleic acid comprising a promoter sequence of SEQ ID NO: 1 which drives expression of an operably linked coding sequence in both yeast and bacteria.

2. A plasmid comprising the promoter of claim 1 selected from the group consisting of SEQ ID NO: 1 (pGLS20), SEQ ID NO: 2 (pGLS22) and SEQ ID NO: 3 (pGLS23) and pBR-AMP-αLPL (SEQ ID NO: 5).

3. A host cell expressing a plasmid of claim 3.

4. The host cell of claim 3, selected from the group consisting of eukaryotic cells and bacterial cells.

5. The host cell of claim 4 selected from the group consisting of S. cerevisiae PRT50, and diploid strains resulting from the mating of said SKY191 and PRT50 strains with the appropriate partner strains.

6. The host cell of claim 4 which is E.coli KJ1567 or E. coli AG58A(RP28).

7. A method for comparing binding interactions between a first protein and a second protein in both bacterial and yeast organismal milieux using a constructs which function in both organisms, comprising

   a) providing a yeast and bacterial host cells, each comprising,
      i) a reporter gene operably linked to a DNA sequence comprising a protein binding site;
      ii) a first fusion gene which expresses a first fusion protein, said first fusion protein comprising said first protein covalently bonding to a binding moiety which is capable of specifically binding to said protein binding site which is driven by a the promoter element as claimed in claim 1 and
      iii) a second fusion gene which expresses a second fusion protein, said second fusion protein comprising said second protein covalently bonded to gene activating moiety;
b) allowing said first and second proteins to interact; and

c) measuring expression of said reporter gene as a measure of said
interaction between said first and second proteins in both E.coli and yeast.

8. The method of claim 7, wherein said promoter sequence comprises the
TEF1 and E.coli lacUV5 promoters.

9. A kit for practicing the method of claim 7, comprising:
   a) a plasmid selected from the group consisting of (pGLS20), SEQ ID
      NO: 2 (pGLS22) and SEQ ID NO: 3 (pGLS23);
   b) at least one of pAC-AMP-αLPL, pBR-AMP- αLPL;
   c) an E. coli strain selected from the group consisting of E.coli
      KJ1567 or E. coli AG58A(RP28); and
   d) a yeast strain selected from the group consisting of PRT50, and

PRT475.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEATURES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>terminator</td>
<td>1336.1528</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rep_origin</td>
<td>1663.2820</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frag</td>
<td>3078..4280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frag</td>
<td>3078..4280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promoter</td>
<td>3122..3465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS</td>
<td>3466..4275</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>terminator</td>
<td>4298..4570</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rep_origin</td>
<td>4605..4874</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frag</td>
<td>1..542</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promoter</td>
<td>10..422</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDS</td>
<td>4677..542</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>terminator</td>
<td>468..526</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASE COUNT</td>
<td>1414 A</td>
<td>1177 C</td>
<td>1160 G</td>
<td>1545 T</td>
<td>0 OTHER</td>
<td></td>
</tr>
</tbody>
</table>

**ORIGIN**

1 TCAATGAGGGG ATCCCCCACA CACCATAATG TCAAAATGCT TCACTACTGCT TTTTACTCTT
61 CGCGATTCGCC TCGAGCCTCG CCGATGCGCC TACCCACCTA AAAACCCCAA GCCAGCATCA
121 CTAAATTTCC CCAACAACCT CACTCTAGGT GTCGCTAATT ACCTGCTTC ACCAGGTGGA
181 AAAAAAAA GAGGCCGCTCT CTTTCCTTCT TCTCTCGCTA AAAAGCGCGT AAAAATTTTT
241 ATCAGCTTCT TTTTCTCTGA AAAATTTTCT TTCTGTTCCT TTCTCTTGCT ATGACCTCCC
301 ATTTGATTTT AAGGTAAATG AGCGTCTCTC ATTTCTCAAG TTTTCTTCTTT ATTTTTCTCG
361 TTATATTACA ACTTTTTTTA CTTCGTCCTC ATTAAGAGAA AAGCACTAGC ATCTAATCTA
421 AGGGCCGCTTG TGACAAATTTA TCAACTGCAT ATGTATATGC CCTAGCTTTT ATCAATTTGG
481 CTCTCGCGCTC GTATAAATGT TGGAGTTGAG GCGGATACAC AATTTCTAC GAGAAACGCG
541 GTATGAGACG CACAAAAAAA CCAATTAACAC AGAGACGAGC TGGAGACGGC GCTGCCTTA
601 AAGCAAATTTA TGAAGAAAG AAAAAATGAC TTTGCTTTACT CCGAGAATCT GCTGCGACA
661 AGATGGGGGAT GGGCAGCTCA GGCGCTTTTG CCTTTTATCTA GGGCATCTATGC ATATTTAGT
721 CTTTATAAGCC CGGATGCTTCT GCAAAAAATT TCAAAATTGT GGTATGAAA TTTAGTCTTC
781 CAACTCGACG AGGAATCTAC GAGATGTAGT AAGCGGTTAG TATGCAGCGC TCACTTTAGA
841 GGGAGATAGA GTACCCGGCGT TTTCTCTAGT TCCAGCGAGG ATGTTCTCTA CTTGACCTTA
901 GAGACCTTTAC CAAAAGGTGTAT GGGGAGAGAT GGTAGTACAC AAAACAAAAA GCCAGTGATT
961 CCTTACATAG CTTTTGAGTT GAGAGTAAAT CACAGCGGCG ACCAAAGGCG TCCAAGCCCA
1021 GCCCTTCTGTA CGGAATGTTTA ATTTCTGGTG ACCCTGAGCA GCCGCTTTAG GCCAGTTGAA
1081 CTTGCACTAGC CAGACTTGTTG GAGATGAGAT TTACCTTCTA AAAAATGACG AGGGTATAGC
1141 GTCAGGTTT TTTTACAACCA CTAAACCCAC AGTACCCCAAT GATCCCGAGC AATGAGAGT
1201 GTGCCCTTGTG GGGGAAGATT ATGCGTATGT AGTGGCGCTGA GAGACGTCTA GCGCAATTTA
Figure 1E

4561 TGCAAGCTGG AGACCAACAT GTGAGCAAAC AGCCAGCAA AGGCCAGGAA CCCTAAAAAG
4621 GCCGCCTTGC TGCCGTCTTC CCATAGGCTC CGCCCCCCTG AGAAGCATCA CAAAATCGA
4681 CGCTCAAGTC AGAGGTTGGC AAAAAAGCAGA AAAAAAGCAA GATACCAAGGC GTTTCCCTCT
4741 GGAAGCTCCC TCGTCGCTTC TCTGTTCTCC ACCGCCTGCC TTACCGGATA CCTGTGCCGC
4801 TTTCCTCCTTT CGGGAGAGGT GCCGCTTTCCT CAAATCTCAG GCCTGAGGTA TCTCAGTTGC
4861 TTGTGGCTCG GCCGCTCCAA GCTGGGCTGT GTGCAACGAC CCCCCCTCA GCGCGCGC
4921 TCACGCTTTAT CGGTGAACCTC TGCTCTTCCAG TCCAAACCAGG TAAAGACAGA CTGATCGCCA
4981 CTGGGACGCA CGACTGGTAA AAAGGATTAC AGAGCCAGGT ATGTAAGCGG TGCTACAGAG
5041 TTCTTTGAGT GTTGCGCTCAAT CAGAGGCTAC ACTAGAAAGGA CAGGTTCAGG TATCAGCGCT
5101 CTGCTGAGGC CAGTTACCTT AAAAAAGAGA GGGGTTAGGCT CTTGATCGCGG CAAACAAACC
5161 ACCGCTGGTTA CGGGTGGTTTT TTGGTGTGTC AAGCAGCAAG TTAGCGCGCA AAAAAAGGA
5221 TCTCAAGAAG ATCCTTTTGTG AAAAAAGAGA GGGTGCTGACG GTACGTGGAA CGAAAACTCA
5281 CGTTAAGGGA TTTTGG

//
**Figure 1F**

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>PGLS20HC</th>
<th>6187 BP</th>
<th>DS-DNA</th>
<th>CIRCULAR SYN</th>
<th>15-JAN-1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEATURES</td>
<td>Location/Qualifiers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>join(543..&gt;1252,&lt;1254..1256)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;cI gene (Split)&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>terminator 1336..1528</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;ADH Terminator&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rep_origin 1663..2820</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;2p origin&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rep_origin 5496..7565</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;ColEl ori&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>frag 1..542</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;4 to 545 of TEF + UVlac promoter&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>promoter 10..422</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;TEF promoter&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>frag 467..542</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;1555 to 1630 of pBT.mv&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>promoter 468..526</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;lacUV5-lacO&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>frag 2834..5472</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;1 to 2639 of Untitled1&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>frag complement(&gt;2834..5472)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;67 to 2706 of CamR-HIS5 cassette&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CDS complement(3718..4884)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;HIS5 yeast&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>promoter complement(4885..5470)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;HIS5 prom&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>frag complement(&gt;2834..3710)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;95 to 1017 of CamR cassette [Split]&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CDS 2941..3708</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;CamR&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>promoter 2835..2940</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/note=&quot;Cam R promoter&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASE COUNT</td>
<td>1773 A</td>
<td>1297 C</td>
<td>1281 G</td>
<td>1836 T</td>
<td>0 OTHER</td>
</tr>
<tr>
<td>ORIGIN</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>TCATGAGGGG</th>
<th>ATCCCCCAAC</th>
<th>CACCATAGCT</th>
<th>TCAAAATGTT</th>
<th>TCTACTCCCT</th>
<th>TTTTACTCTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>CCAGATTTTC</td>
<td>TGCGACCTCG</td>
<td>CGCATCGCGC</td>
<td>TACACCTCTA</td>
<td>AAAACCCCA</td>
<td>GCACAGGATA</td>
</tr>
<tr>
<td>3</td>
<td>CTAATTTTCC</td>
<td>CTCTTCTTCT</td>
<td>CCTCTAGGT</td>
<td>GCAGTTAAT</td>
<td>ACCCGTACTA</td>
<td>AAGGTTGGA</td>
</tr>
<tr>
<td>4</td>
<td>AAAAAGAAAA</td>
<td>GAGACCGCCT</td>
<td>TGTTTTTTTTT</td>
<td>TCTCTCGTA</td>
<td>AAAAAAGCAA</td>
<td>AAAAAATTTTT</td>
</tr>
<tr>
<td>5</td>
<td>ATTCAGGGGG</td>
<td>ATTTTTTTTTT</td>
<td>TTTTATTTTTT</td>
<td>TTTATTTTTT</td>
<td>TTTTTTTTTT</td>
<td>TTTTTTTTTT</td>
</tr>
<tr>
<td>6</td>
<td>ATTCATATTT</td>
<td>ATGTTAATTA</td>
<td>ACCGCGCCAT</td>
<td>ATTTTTTTTTT</td>
<td>TTTTATTTTTT</td>
<td>TTTTATTTTTT</td>
</tr>
<tr>
<td>7</td>
<td>TTCATTTTTC</td>
<td>CTTTCCGCTT</td>
<td>ATTTTTTTTTT</td>
<td>TTTTATTTTTT</td>
<td>TTTTATTTTTT</td>
<td>TTTTATTTTTT</td>
</tr>
<tr>
<td>8</td>
<td>ATGGCGGTTG</td>
<td>TGACAAATTA</td>
<td>ATGCGCCGAT</td>
<td>ATGTTATATTG</td>
<td>ACCGATAAC</td>
<td>AATTTCCAC</td>
</tr>
<tr>
<td>9</td>
<td>TGCAGTGCT</td>
<td>GTAAAATGTT</td>
<td>TGGAAATGTT</td>
<td>ACCGATAAC</td>
<td>AATTTCCAC</td>
<td>AGAAGACGC</td>
</tr>
<tr>
<td>10</td>
<td>GCTATGACGAC</td>
<td>AGAAAAAGAA</td>
<td>CCATTAACAC</td>
<td>AGAAGACGC</td>
<td>ATGCGACCTG</td>
<td>TGGAAATGTT</td>
</tr>
<tr>
<td>11</td>
<td>AAGAAATTTA</td>
<td>TGAAAAAAAG</td>
<td>AAAATGAA</td>
<td>TGTGGCTATC</td>
<td>CCCAGAATCT</td>
<td>GTGCCAGACA</td>
</tr>
<tr>
<td>12</td>
<td>AGATGCGGTAT</td>
<td>GGCGGGATC</td>
<td>GCAGTTGGTG</td>
<td>TTATAATTAA</td>
<td>TGCCATGCA</td>
<td>CTGTTAAATG</td>
</tr>
<tr>
<td>13</td>
<td>CTTATCAAGCT</td>
<td>GCATTTAGCT</td>
<td>GCAGTTGGTG</td>
<td>TTATAATTAA</td>
<td>TGCCATGCA</td>
<td>CTGTTAAATG</td>
</tr>
<tr>
<td>14</td>
<td>GTGAGATGAT</td>
<td>GTACACCTGT</td>
<td>TTATATATAT</td>
<td>TTATATATAT</td>
<td>TTATATATAT</td>
<td>TTATATATAT</td>
</tr>
<tr>
<td>15</td>
<td>TCTCCTTGCG</td>
<td>CCAGCTATCT</td>
<td>TGACCTTATG</td>
<td>AGAGATGAT</td>
<td>AAGCGGTTG</td>
<td>ATGCGACCTG</td>
</tr>
<tr>
<td>16</td>
<td>CATTATTTAC</td>
<td>AGCGAAGTAC</td>
<td>CGATATGAT</td>
<td>AGAGATGAT</td>
<td>AAGCGGTTG</td>
<td>ATGCGACCTG</td>
</tr>
<tr>
<td>17</td>
<td>GAGGAGAAAT</td>
<td>TGGCGATGAA</td>
<td>CCCAGAATCT</td>
<td>GTGCCAGACA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>CTGCATTCTG</td>
<td>GCTTTAGGTT</td>
<td>GAGGATTTAC</td>
<td>CCATGACGCG</td>
<td>ACCCAACAGC</td>
<td>TCCAGGGGAA</td>
</tr>
<tr>
<td>19</td>
<td>1021</td>
<td>GCTTTCTG</td>
<td>CGGAATGTGTA</td>
<td>ATCTCCCCCC</td>
<td>ACCCTGAGCA</td>
<td>GGCTGTGAGC</td>
</tr>
</tbody>
</table>
Figure 1H

```
4441 GTCCGAAACA GTTAAGGAGC ATTTGGACGAC TTCTATATCA TTATAATTTTG CACAAACAGA  
4501 ATACATAGAAA TATGGTTGTTG GAAGAACACG AACTTTTCTCT TTCCCCGGGAA CACAGCATGC  
4561 TCTAATAATAA GCTAACATAC TCTCATACAG CCCACACATT ACGCAGCAAG ATGCAGCAGT  
4621 TAAAGGTTTTT ACCTCTGGGT CATGGCCATA ACTGCTTTGT TGGTTCTCTG ATTTCCGTCAT  
4681 TGGGCTCTTG AACTCCAAAT GTGAGGAGGC TGGGGAAGCC TGGGTTATGG TCTGCGTCAA  
4741 TCTCACTGGA GTAGGCTCAT CCGCTTTTGA ATTTGGGCTT ACGAAATATAC CCGGTCGGGA  
4801 ATCATCTCTTT GCCAGGCAAT AAGGTTCCAA GTTATAAAAT TGTTAGCTCA CAATCCTTTT  
4861 CAAATCAAAA ACCATAGGTG AATTTTTATA ATTACGCCAC ACAGGCCTTG CGGTGGTTATC  
4921 AGGAAACAAC ATTAAGAGTAG TCAATTGGAAG AACTGTTTCT ATCATATTAGA GGTTATATTA  
4981 AACAGAACCTG TGCTGACCTCT TTACTAGTTA TATAAGCTGA GAGATAAAT ATCAAGTATG  
5041 TCATGTCAGG GAAAGAAACA TCAATTGGAAG TGAGTCAAAC GATCCAAAGA AAAAAAGCAGC  
5101 TAACACTGCTA ACTACACCTAT GAATATTTTGA AAATGTTTAG TTAAAGCTATG TCAAGATCCA  
5161 TTTACTGGGA AAAACAGTGT ATACATAACA ATTAACCTCG GATTAGCTCA TAATATTTTC  
5221 ATATGGTCTT TCCCTTTGGA ATAGTACATC CCGGACTGTT GTACAGTATA CAAATAAAAT  
5281 ATATGCTTTAC CAATGGCTTG TGTATTAGTT ACTCTCCAGG AATGTCCTCA ACATAACCGG  
5341 TCACTCATATT TATGTAACCA ATTTTTAACC ATTTACCCCT TATTTTTTGA AAGTTATAGC  
5401 CTTTGGAATAAC GACGAGAGAG AAAAAATTTG AAGAGTAAAC ATCAGAAGAG GTTCCGCGAAA  
5461 AGACAGATCA TGTAAGCAAAA AGGCGCAAGA AAGCCAGAGA ACCGTAAGAA GGCAGCGTTG  
5521 CTGGCGGTTTT TCCCAATCTTG CCGGCCTCTT GACGCACACG AAAAAAATCG ACACCTCAAGT  
5581 CAGGAGTGCGC AGACGCCACG AGGCGCTATAA AGATACCCAG GCTTTCCTCC TGGAGCTGCC  
5641 CTGCTGCTCTT CCCCTGCTAC GACCTGCGGC ACCGTCACAG AACTCTGCTG CTTTCCTCCCT  
5701 TGCGGCGGCC TGGCGCTTTG TCATCTTCTCA CGCTGTAGTT CAATCTGCTC GGTGGAGGCT  
5761 GTGGCTCTCCA AGCTGGCGCG TGTCGACGAA CCCCCGCTC AGCCGACGCC CTGGGCGCTTA  
5821 TCCGTTAACT ATGTTCTTGA GCTGAAACCG GTAAGACAG ACTTTATCCG ACTCGGACCA  
5881 GCCACTTGTAA AAGAGGAGTG CAGAGGCGAG TATGAGGCG GTGCCACAG GCTCTGAAAG  
5941 TGCTGGCGCTA ATCAGCCGTT CACATGAAGG AACGTATTTG GATCTGCGCC TCGCTGAAAG  
6001 CCAGTGCTCTG CGGAAGAAAG AGTGGTGGATG TCTTGTACCG CCAAACAAAC CACCCGCTGT  
6061 AGCCGGTGGTT TTTTTGTGTG CAGGGCAAGG ATTACGGGGA GAAAGGTGGA ATCTCAGGAA  
6121 GATCCCTTTGA TCTTTTCTAC GGGGCTGTGA CACTAGGTGA ACGAAAACCT ACAGTTAAGGG  
6181 ATTTTGG //
```
**FIGURE II**

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>PBR_UVS_AL</th>
<th>4320 BP</th>
<th>DS-DNA</th>
<th>CIRCULAR</th>
<th>SYN</th>
<th>18-DEC-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASE COUNT</td>
<td>1088 A</td>
<td>1061 C</td>
<td>1085 G</td>
<td>1086 T</td>
<td>0 OTHER</td>
<td></td>
</tr>
<tr>
<td>ORIGIN</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 GAATTCGGAC CCGTCCTAGC ACGTCTAAGA AACCATTATC ATCATGACAT TAACCTATAA
61 AAATAGGGGTT ATCGAGGAGG CCGTTGGAAT ACCAGAAGCA ATAAAATAATC AATTCGGATT
121 TCAGTATATA ATCTCAGCTT ATCTAAGATG ATCCAGCGAG ACGCATCGTG TTTTCTTTC
181 GCTTCCCCATC AAAAAATAAT TTAGGACATAA ACAAAATTTG TTGATATCATG TAAACCTAC
301 ATGGGAGATTT ATCTGCTACTC TTGGATAGAG CTTTACATTT TATGCTCTCG GCTGCTATAA
361 TCTGGTGGTAA ACGAGACGGA TAAACAGATG ACAGCAGGA AACCTATGAG CATGATTTCT
421 GATCAGCTGG AACTATGACTG ACAACAGACG ACGCAGCTGC CAGCTCTGCTG AGCTGATTCT
481 TAAAACCGGG CCGTGTTGATG ATCGGAGGAG TGGATCTCA ACACCCGAGA TGCTCGCTGG
541 AGCCTTCTAGA GCGTGCTGTAG TGCACTTACTC TGGGAACCTA AGCTCGCGCTAG ATCCCTGCTC
601 CATCTGACGAT GCGCTGCAAG GCGTACGGAC TTGATGCTAC CTTGTTGCTAC CTAAGATGTA
661 GCACCAACAG AGCGGCTTAC GAGAAATCAT TTGAATCTCT GTGAACACTG AAAGGCAGAC
721 CCGTGGGACT ATCCGGACAC ACGGATCTTC TGTATCTCT CTGCGCTGTAT GCATGATGCT
781 CTCGCGCTTC GGCGTCACTG TTGCTCATTG CGCTGGAGCA TATGCTGAGG CCGAGCAAGG
841 TGATCTGCTCA CGCGTCACTAA GTGACAGGAT CTTACCCAA AAGGAGATGG CCGACCTGAA
901 GTCTGGTGTGA TGGGAGTGTA CTGACGCTAAT TTACCAAGC GAAGACTAGT CGAGGCTACTG
961 GCGCTCTGCTG ATGACGGAGT TTTGGCCTTA AGCTGATGCTG TACCTAGTTA AAAGCTGACT
1021 AGACGCGATG AGAAGACTAG GAGCGCGATT AACTAAGATG TATCGCTGATG TACGCTACTG
1081 GCACAATCAGA TCCTGAAAGG CGGATCTGAT GCTCGGCAAC CATTCTGTCTG GAAACACTTG
1141 AAGCTTCTTGT TATGTTAGCT ATGCTGTGCTG GAGAAGAGATG AAAGCAAGAG GAAACAGAGG
1201 CGCGCGCGCA ATGGGAGCTC ATCGAGGACT CTTAAGGATCT CTTATGCTCTG GAGGAGACAG
1261 TAGCTCTGAG ACGCTAAGGG TATGCTGATG ACGCTGATCTG TGGGCTCTCA GCTGAAATCA
1321 CGGAGAGACG GAGCGGCGCC CGTCACTTTC AATCCAGTTA CTCGGATGTC CGGACTTCTG
1381 AGCGGCCCTTG AACTCGGCGA ATAGGTTGCTG CGGGAGTGTTG TGCTGCTGATG GGGGAGGAC
1441 CTATGGTGTG ACGCTGCTAT GTGCTGGCTGC TCTGCTGCTG CTTGCTGCTG CTGCTGCTG
1501 CTGCCTGGGT CCTTCTCTGT AGCTATCTTG TGGGAGGCTG CTCGCTGCTG CTTGCTGCTG
1561 GTTGTTCTGG ACGAGCTGCTT CCGAAAAATGC ATGTATGGAG TGTGGTGTGC CTTGCTGCTG
1621 CATGCTGCTCG GCTACGGTGT TTTGGCCTTA TGCGCTGATG TACCTAGTTA AAAGCTGACT
1681 GACTCTTGTG TCCACGTGGA ATCACTGAATC AACTACTCTC GTGCTTACTC TTCTATGTAT
1741 AAGGATTTTT GCGCTTTGAT TAAAGAATGA TCTGATTTAG TAAAGAATGA TAAAGAATGA
1801 AGGCGAATTG TAAACAAATAA TTGAGGCCTTT TAAATAAGGT GGGCGCTTTG GGGGAGGT
1861 GCGGGAAACA GCCCTTTGCTT ATGACTACGA AATACGCTGG CTTGACGAGC
1921 ACAATTCCCA CTATGACGGT TTTCTATGCT TTGAGGAAGG CTTAAGATGC TATCTGACTA
1981 TTGCGCTGCTG GCTTCTGTGCT TCTGCTTGTG CTTGCTGCTG CTTGCTGCTG CTTGCTGCTG
2041 AGAAGAGCTG CTGAGGAGTAA AAGAACGTGA AGAAGACGTA GCATGACAGA TGAGTGAAGA
2101 CGCAGCTGCTG CACTCACGTT GAATAGCACT CGTGGATTTC GTCGCCCAGG AGACCTTCTA
2161 AAGGGAGGTC ATTTAAAAAG GCGAGCTCGT TTGCGCTGCTG CTTGCGCTGCT TCTGCTGCTG
2221 GCGAGGAGCG AATGGATGTTT CTGAACTGATG CTAATGGAGA ATCTTACAGA GTTCTGCTG
2281 AGTACGAGA AAGCTCAGTTA CGTGGATGCT CAAGATAGGA GAATAGCAGA TTGCTGCTGCT
2341 AACAGTGATG TATTTTTTAC AGTCCAGACT CTCTGCTGCTG CTTGACTGAG CAGCAGAGAG
2401 GCTAAACCCT GTGGCTGCTT ACATGGGAGA TGCTGCTGCTG CGTGGAGTCTC TGGTGCTGCT
2461 GACCTAGCTT ATGAGCGGCTA CAAACGCAAG GCTGGAGAGA ATCTTACAGA GCATGACAGA
2521 AACACTGCTT CGAACAAGTA AAGAACGTGA ATGAGCAGA AGAAGACGTA GCATGACAGA
2581 AATAGGACTG CTTGAGGCGCG AATAGGGTTAC GGCCAGTCTG CTTGCTGCTG CTTGCTGCTG
Figure 1J

2641 TGGCTGGTTT ATTGCTGATA AATCTGGAGC CGGTGAGCG GTATCATTCG
2701 AGCAGCTGGG CGCAGTGGTA AGCCCTCGCC TATCGTGTAT ATCTACAGGA CGGGAGTCA
2761 GGAACACTAG TGAGACAGAA ATAGACAGAT CGCTGAGATA GGGCTCTCGA TGATTAAGCA
2821 TCCTTTACCTG TCAGCAACAG TTTATCTCAT TATACCTTAG ATTTAGTTAA AACTCTATTT
2881 TTAATTTAAA AGGACTTAGG TGAAAGATCT TTGTGAATA CTCATGACCA AAATCCCCTA
2941 AGCTGAGTTTT TCCTTCCACT GACGTCGAG CCCCCGAGAA AAGATCAAGA GATCTCTTCG
3001 AGATCCCTTTTT TTCTCCTGCC TAAATCTGCTG CTCTGAAACA AAATAAACACC CGCTACACGA
3061 GTGGTTTTGT TTGGCGGATC AAGAGCTACC AACTCTTTTT CCGAAGGTAA CTGGCTTCAG
3121 CAGAGGCGCAG ATACCAAAATA CTGTCCTTCT AGTGTAAGCG TAGTTAGGCC ACCACTTCAA
3181 GAACCTCTGTA GCACCCCGCCT AATACACTCC CTCGTAATCT CTGTGACCCAG TGGCTGTGCG
3241 CAGTCGTGGAT AAGTGCGTCT TCTACGGGTT GACCTCAGAA GATATGTTAC CGGATAAGGC
3301 GCACGCCGTCG GGCTGAAAGG GGGGTTCTGTG CACACAGCCC AGCTGGAGCC GAAGGACTTA
3361 CACCGCAACTG AGGATACCTAC AGGATGAGCT ATGAGAAAGA GCCACGCTTC CCGAGGGAG
3421 AAAGGAGGGCG AGGTATCCGG GTCAGGGCGA GGCTGGAAAGA GGAAAGAGCA CGGGAGACGT
3481 GCCAGGGGGA ATCGCTGTTT ATATTTATAG TCTGTGAGGG TTTGCCCTCC TCTGACTTGA
3541 GCCGTAATTT TTGCTGACGG GTACAGGGGG CCGAGGCTCA TTGAAAACAG CAACGCAAGC
3601 GCCCTTTTAA CGGTCGGTCC CTTCTGCTTG AAGCTCTGAT CACATGTTCT TCTGGGCGTT
3661 ATCCCCCCTGAT TCTCTGGAATA AGCCGATTTC CCCTTGGAG TGGACTCTGA CCCTCCGGCG
3721 CACCCGACCG ACCAGGCGCA GGCGATCAGT GAGCGAGGA ACGAGAAGAC GCCTCTGATC
3781 GTATTTTCTCT CTATGCAGGT CTGGCGGTAT TACACACCGC ATGAGGTGCA CTCTCACGAT
3841 AAATCTGCTCT GATCAGCGCA ATGGTACGCA CCTAAAACCT CGCTACTGCT ACCTGAGCG
3901 GCAGCTGCTG CGCCGCAACA CCGCCGAACA CCCGGCTACG CGCCGCGACG GGCTGTCTCG
3961 CTCCCCGCGAT CGCAGCTGCA ACAAGCTGCTG ACCGTCTCCG GAGCTGGCAT GCTGGAGGG
4021 CTGGCCTGCT GATCCACGGAA AGCCGCGCGG CAGCTGCGGTT AAAGCTCAGT AGCGGCGTCC
4081 TGAGCCGGAT CAGATCGCTG GCCTCCTCTG TCCGCCGCCC GCCTTCCTG TCTCCACGAG
4141 AGGCAATAAG TCGCGGCTCT GTTAAAGCGG GCCATGGTAA GGGGTTTTT TTCCCTGTGTT
4201 GTCACTGATG CCTCGGATGA AGGGGTTTTG CTGTTCTAGG GGTTATGATG ACCAGTGAAG
4261 CGAGAGAGGGA TGCTCAGCAT ACCTGTACCT CATGATGAAC ATGCCCGGGTT ACTGGAAACG

//
Figure 6

A hybrid promoter element consisting of the yeast TEF promoter (yellow) and bacterial (E.coli) lacUV5 promoter (bold) (SEQ ID NO: 1)

TCATGAGGGGATCCCCCACACACCATAGCTTCAAAATGTCTTTTACTCCATT
TTTACTTTCCAGATTCTTCGGACTCCGCACTCCGACCTTCAA
ACACCCCAAGCAGCATACTAAATTTCCCTTCCTTTTCTTCTTAGGGTGT
GTAAATTACCGTACTAAAGGTTTTTGAAAAGAAGAGACCGCCCTGCT
TTCTTTTCTTCTCGAAAGAACAGAAAATTTTTATCAAGGTTCTTTTCT
TTTTTTCCGATACTCCCAATTGATATT
TTAAGTTAATAAACGGTCTTCAATTCTCAAGTTTCAGTTTCATTATTTTCTTG
TTACTTACAHTTTTTTACTCTTGCTCTAGAAAGAAAGCATAGCAATCT
TAGGCTTTACACTTTATGCTTCCGGCTCGATAATGTGTGGAAATTGTG
AGCGGATAACAATTCACACAGAAACACCGT