



(86) Date de dépôt PCT/PCT Filing Date: 2001/11/01
(87) Date publication PCT/PCT Publication Date: 2002/05/10
(45) Date de délivrance/Issue Date: 2011/10/04
(85) Entrée phase nationale/National Entry: 2003/04/25
(86) N° demande PCT/PCT Application No.: AU 2001/001407
(87) N° publication PCT/PCT Publication No.: 2002/036784
(30) Priorités/Priorities: 2000/11/01 (AU PR 1161);
2001/05/10 (AU PR 4901)

(51) Cl.Int./Int.Cl. *C12N 15/63* (2006.01),
A01K 67/027 (2006.01), *A61K 49/00* (2006.01),
C12N 15/85 (2006.01), *C12N 5/10* (2006.01),
C12N 9/02 (2006.01), *C12Q 1/68* (2006.01)
(72) Inventeurs/Inventors:
LIDDLE, CHRISTOPHER, AU;
GOODWIN, BRYAN JAMES, US;
ROBERTSON, GRAHAM, AU
(73) Propriétaire/Owner:
THE UNIVERSITY OF SYDNEY, AU
(74) Agent: GOWAN INTELLECTUAL PROPERTY

(54) Titre : ANIMAUX TRANSGENIQUES UTILISES POUR ANALYSER LA REGULATION DU GENE P450 DU
CYTOCHROME CYP3A4
(54) Title: TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450 GENE REGULATION

(57) **Abrégé/Abstract:**

The invention relates to the generation of non-human transgenic animals comprising a reporter construct for producing a detectable amount of a reporter molecule operably linked to a transcriptional regulatory nucleic acid molecule from the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located 13,000 nucleotides upstream from the site. The invention also relates to the use of these animals for determining the effect of a compound, particularly, but not exclusively, a xenobiotic or steroid, on the regulation of expression of the CYP3A4 gene in a human.



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36784 A1

(51) International Patent Classification⁷: **C12N 15/63**,
5/10, C12Q 1/68, A01K 67/027, A61K 49/00

(74) Agent: **GRIFFITH HACK**; GPO Box 4164, SYDNEY,
New South Wales 2001 (AU).

(21) International Application Number: PCT/AU01/01407

(22) International Filing Date:
1 November 2001 (01.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 1161 1 November 2000 (01.11.2000) AU
PR 4901 10 May 2001 (10.05.2001) AU

(71) Applicant (*for all designated States except US*): **THE UNIVERSITY OF SYDNEY** [AU/AU]; Parramatta Road, Sydney, New South Wales 2006 (AU).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **LIDDLE, Christopher** [AU/AU]; 7 Hillside Street, Chatswood, New South Wales 2067 (AU). **GOODWIN, Bryan, James** [AU/US]; Apartment B10, Sterling Bluff, 140 BPW Club Road, Carrboro, NC 27510 (US). **ROBERTSON, Graham** [AU/AU]; c/o The University of Sydney, Parramatta Road, Sydney, New South Wales 2006 (AU).

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

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

Published:

— *with international search report*

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

(54) Title: TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450 GENE REGULATION

(57) Abstract: The invention relates to the generation of non-human transgenic animals comprising a reporter construct for producing a detectable amount of a reporter molecule operably linked to a transcriptional regulatory nucleic acid molecule from the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located 13,000 nucleotides upstream from the site. The invention also relates to the use of these animals for determining the effect of a compound, particularly, but not exclusively, a xenobiotic or steroid, on the regulation of expression of the CYP3A4 gene in a human.



WO 02/36784 A1

TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450 GENE REGULATION**TECHNICAL FIELD OF THE INVENTION**

5 The invention relates to the generation of a transgenic animal and to the use of the animal for determining the effect of a compound, particularly, but not exclusively, a xenobiotic or steroid, on the regulation of expression of a P450 gene in a human.

10

BACKGROUND OF THE INVENTION

Many endogenous and exogenous compounds are observed to have a therapeutic effect in drug development trials *in vitro*. However, the intended therapeutic effect is often
15 not realised in clinical practice, for example, when compounds are co-administered, because certain compounds induce the expression of the CYP3A4 gene. This induction generates CYP3A4 cytochrome P450 molecules which metabolise compounds before the intended therapeutic
20 effect of each compound can be realised. Accordingly, induction of expression of the CYP3A4 gene interferes with intended dosage, leading to therapeutic failure or sub-optimal treatment.

25 Induction of CYP3A4 gene expression is a significant problem for drug development because time, resources and expense are wasted in the development of candidate drugs for therapy of particular disease conditions which will ultimately fail or perform sub-optimally in clinical
30 practice.

It would be advantageous to have an animal model for use in drug development trials from which, at an early stage of drug development, one could determine whether a

candidate drug would be likely to achieve an intended therapeutic effect in a human.

Such an animal model would not be useful unless at least
5 some of the aspects of the regulation of CYP3A4 gene
expression in the human, especially tissue specific
expression, are reproduced. This is because in the human,
the CYP3A4 gene is expressed in specific tissues,
including liver and small intestine, which many compounds
10 inevitably come into contact with when administered for
the purpose of therapy. Accordingly, one would be unable
to determine whether the bio-availability of a candidate
drug would be sufficient for achieving an intended
therapeutic effect in clinical practice in a model which
15 does not reproduce the constitutive and xenobiotic induced
tissue specific expression of the CYP3A4 gene that is
observed in the human.

WO99/61622 and Goodwin B et al, "The orphan human pregnane X
20 receptor mediates the-transcriptional activation of CYP3A4 by
rifampicin through a distal enhancer module." Molecular.
Pharmacology, 56(6): 1329-1339 (December 1999) disclose a nucleic
acid molecule located 8 kb upstream from the initiation of
transcription site of the of the CYP3A4 gene which regulates
25 transcription of the CYP3A4 gene in response to xenobiotic
compounds. These documents do not disclose elements for regulating
the constitutive and xenobiotic inducible tissue specific and
developmental expression of the CYP3A4 gene observed in a human.

30 There is a need for an animal model which reproduces at least some
aspects of the expression of the CYP3A4 gene in a human, for
determining whether a compound, for example, one identified in a
drug development trial, would be likely to induce CYP3A4, and hence
cause drug-drug interactions, or auto-induction of the metabolism
of the drug under study.

DESCRIPTION OF THE INVENTION

The invention seeks to address the above identified need and in a first aspect provides a non-human mammal comprising:

5

(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to a sequence of the human CYP3A4 gene located between the
10 initiation of transcription site of the gene and a position located at least 13,000 nucleotides upstream from the site; and

(b) a reporter nucleic acid molecule for producing a
15 detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules
20 are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

As described herein, the inventors have found that the
25 incorporation of a region of the human CYP3A4 gene that is located between the initiation of transcription site of the gene and a position 13,000 nucleotides upstream of the initiation of transcription site into an animal model provides the animal with sufficient genetic information
30 for reproducing the constitutive and xenobiotic induced tissue specific expression of the CYP3A4 gene that is observed in humans. More specifically, the inventors have generated animal models which contain a transgene comprising this region and have observed that these models
35 provide constitutive and xenobiotic inducible expression of a transgene in a tissue pattern which reproduces the

tissue specific expression of CYP3A4 which is observed in a human. Importantly, the level of constitutive expression is sufficient to allow one to observe the effect on the regulation of tissue specific transgene
5 expression, of administration of a compound, for example, a xenobiotic or steroid, to the animal.

Further, the inventors have observed that the animal models described herein also reproduce aspects of the
10 constitutive and xenobiotic inducible developmental expression of the CYP3A4 gene that is observed in humans.

These findings are unanticipated because prior to the invention, there was no suggestion that the genetic
15 information required for simulating the constitutive and xenobiotic induced tissue specific or developmental expression of the CYP3A4 gene that is observed in a human would be contained in the region of the human CYP3A4 gene between the initiation of transcription site of the gene
20 and a position 13,000 nucleotides upstream of the initiation of transcription site.

Further, prior to the invention, differences in the induction profile of the mouse CYP3A11 and the human
25 CYP3A4 gene had been observed, and differences had also been observed in the ligand binding profile of mouse transcription factors, especially PXR and CAR, and human PXR and CAR. Accordingly, there was no suggestion that a non-human animal would have factors sufficient for
30 interacting with a region of the CYP3A4 gene for reproducing the constitutive and xenobiotic induced tissue specific or developmental expression of CYP3A4 observed in a human.

35 Further, prior to the invention, mechanisms associated with transgene integration had been observed, such as gene

silencing and mosaic transgene expression which limited the extent to which an a transcriptional enhancer element incorporated into a transgenic model could reproduce regulation of gene expression observed in a human.

5 Accordingly, there was no suggestion that a region of the human CYP3A4 gene would be capable of reproducing the regulation of expression of the CYP3A4 gene that is observed in a human. However, as described herein, the inventors have shown in 2 separate founder lines that the
10 expression of the transgene reproduces aspects of CYP3A4 gene expression that are observed in humans.

Thus in a second aspect, the invention provides a non human mammal comprising:

15

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the nucleotide sequence of the human CYP3A4 gene that extends about 13,000 nucleotides upstream from the initiation of
20 transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid
25 molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic
30 acid molecule.

In one embodiment, the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:1.

35 Further, as described herein, the inventors have generated transgenic animals which contain a region of the human

CYP3A4 gene between the initiation of transcription site and a position about 3,200 nucleotides upstream of the initiation transcription site and observed that the transgene is not constitutively expressed or inducible by xenobiotics in these animals. Accordingly, the inventors have found that the genetic information required for reproducing the constitutive and xenobiotic induced tissue specific and developmental expression of CYP3A4 observed in a human is contained in the region of the human CYP3A4 gene between the position located about 3,200 nucleotides upstream of the initiation of transcription site of the gene and a position 13,000 nucleotides upstream of the initiation of transcription site.

Thus, in a third aspect, the invention provides a non-human mammal comprising:

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 8,000 nucleotides upstream from a position about 3,000 nucleotides upstream from the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

In one embodiment, the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:2.

In a fourth aspect, the invention provides a non-human mammal comprising:

(a) a regulatory nucleic acid molecule which is capable
5 of regulating transcription of the human CYP3A4 gene and
which comprises a nucleotide sequence that is identical to
the sequence of the human CYP3A4 gene that extends about
600 nucleotides upstream from a position about 7,200
nucleotides upstream of the initiation of transcription
10 site of the gene; and

(b) a reporter nucleic acid molecule for producing a
detectable amount of a reporter molecule for indicating
regulation of transcription of the reporter nucleic acid
15 molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules
are arranged to permit the regulatory nucleic acid
molecule to regulate transcription of the reporter nucleic
20 acid molecule.

In one embodiment, the regulatory nucleic acid molecule
comprises the sequence shown in SEQ ID NO:3.

25 In another embodiment, the regulatory nucleic acid
molecule has the sequence of any one of the following
fragments of the CYP3A4 gene:

- (i) a fragment consisting of from nucleotide positions
-13,000 to +53;
- 30 (ii) a fragment consisting of from nucleotide positions
-13,000 to -12,700 contiguous with -8000 to +53;
- (iii) a fragment consisting of from nucleotide positions
-13,000 to -5,100 contiguous with -1,200 to +53;
- (v) a fragment consisting of from nucleotide positions
35 -7,800 to -6,000 contiguous with -362 to +53;
- (vi) a fragment consisting of from nucleotide positions

-7,500 to -6,000 contiguous with -362 to +53;

A regulatory nucleic acid molecule which has the sequence of a fragment consisting of from nucleotide positions
5 -7836 to -7207 contiguous with -362 to +53 is particularly preferred, as this construct contains the minimal sequences necessary for regulating transcription of the human CYP3A4 gene, more specifically, an element
10 responsive to xenobiotics (the "Xenobiotic Response Element Module" or "XREM") and the proximal promoter of the CYP3A4 gene.

The regulatory nucleic acid molecule of the invention typically contains at least one enhancer capable of
15 regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor. Examples of such enhancers are those capable of regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor bound to a ligand, such as a xenobiotic or steroid. Other
20 examples are those capable of regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor consisting of a heterodimer of PXR (pregnane X receptor, otherwise known as SXR (steroid and xenobiotic receptor)) and RXR (9-cis retinoic acid receptor), or CAR
25 (constitutive androstane receptor- β) and RXR.

The inventors believe that certain nucleic acid molecules which have substantially the same nucleotide sequence as a regulatory nucleic acid molecule of the invention would
30 also have sufficient genetic information for reproducing the constitutive and xenobiotic induced tissue specific and developmental expression of the CYP3A4 gene that is observed in a human. Accordingly, it will be understood that nucleotides could be modified or deleted in regions
35 of the regulatory nucleic acid molecule, more specifically, those regions which do not contain an

enhancer such as those described above, without significantly limiting the capacity of the molecule to regulate transcription of the human CYP3A4 gene.

5 The inventors recognise that it would be advantageous to provide an animal model further capable of reproducing the expression of other human genes, specifically those genes encoding products which modify or modulate the therapeutic activity of exogenous and endogenous compounds used for
10 therapy and cause drug-drug interactions, for example, cytochrome P450 genes or ABC transporter superfamily genes, for example, P-glycoprotein (otherwise known as MDR-1). The regions controlling the constitutive and xenobiotic induced tissue specific expression of some of
15 these genes are known, and in some instances, non-human animal models have been generated. The inventors recognise that the genetic background of these animals could be incorporated into the non-human mammal of the present invention, for example, by conventional breeding
20 techniques.

Thus in a fifth aspect, the invention provides a non-human mammal of any one of the first to fourth aspects of the invention, further comprising:

25

(c) a further regulatory nucleic acid molecule which is capable of regulating transcription of a human gene; and

(d) a further reporter nucleic acid molecule for
30 producing a detectable amount of a further reporter molecule for indicating regulation of transcription of the further reporter nucleic acid molecule by the further regulatory nucleic acid molecule

35 wherein the further reporter and further regulatory nucleic acid molecules are arranged to permit the further

regulatory nucleic acid molecule to regulate transcription of the further reporter nucleic acid molecule.

In one embodiment, the at least one further regulatory
5 nucleic acid molecule has a sequence shown in SEQ ID NO:4.
In another embodiment, the at least one further regulatory
nucleic acid molecule has a sequence shown in SEQ ID NO:5.

Although the regulatory nucleic acid molecule of the
10 invention described herein is sufficient for reproducing
the constitutive tissue specific and developmental
expression of the CYP3A4 gene that is observed in a human,
the inventors recognise that aspects of the xenobiotic
inducibility of the gene could be better reproduced in an
15 animal by incorporating at least one human transcription
factor that is capable of interacting with the regulatory
nucleic acid molecule for regulating transcription of the
human CYP3A4 gene. Examples of such factors are nuclear
receptors. These receptors may be those capable of
20 regulating CYP3A4 gene transcription in a human when the
receptor is bound to a ligand, such as a xenobiotic or
steroid. One example of such a receptor is the human PXR
(pregnane X receptor, otherwise known as SXR (steroid and
xenobiotic receptor)). Another suitable receptor is the
25 human CAR (constitutive androstane receptor- β). Non-human
animals comprising a human PXR or CAR receptor are known.
The inventors recognise that the genetic background of
these animals could be incorporated into the non-human
mammal of the present invention, for example, by
30 conventional breeding techniques.

Thus in a sixth aspect, the non-human animal of the
invention further comprises at least one human
transcription factor for regulating transcription of a
35 human CYP3A4 gene. Preferably the transcription factor is
a nuclear receptor. Preferably, the nuclear receptor is a

heterodimer of the human PXR (pregnane X receptor, otherwise known as SXR (steroid and xenobiotic receptor)) and human RXR (9-cis retinoic acid receptor) or human CAR (constitutive androstane receptor- β) and human RXR.

5

It follows that the reporter nucleic acid molecule can be any molecule which is capable of detection when the reporter nucleic acid molecule is transcribed. For example, the reporter nucleic acid molecule could be the
10 CYP3A4 cytochrome, or the mRNA transcript which is translated to produce the cytochrome. Those reporter molecules which are commercially available, including firefly luciferase, β -galactosidase, alkaline phosphatase, green fluorescent protein or chloramphenicol acetyl
15 transferase can be used.

Thus in one embodiment, the reporter nucleic acid molecule is capable of producing a reporter molecule selected from the group of reporter molecules consisting of firefly
20 luciferase, β -galactosidase, alkaline phosphatase, green fluorescent protein or chloramphenicol acetyl transferase.

While the non-human mammal of the invention, as exemplified below, is a mouse, the inventors believe that
25 any other non-human mammal could be used in the invention, especially those for which standard transgenic techniques have been developed including for example, rat and rabbit. However, typically the non-human mammal is a mouse.

30 In another aspect, the invention provides a tissue of a non-human mammal of the invention.

In one embodiment, the tissue is an embryo capable of producing a non-human mammal of the invention.

35

In a further aspect, the invention provides a method of determining whether a compound is capable of effecting the transcription of a human CYP3A4 gene the method comprising the following steps:

5

(a) administering the compound to a non human mammal according to the invention and

(b) determining whether the reporter molecule is produced by the reporter nucleic acid molecule in the mammal.

10

In one embodiment, the production of the reporter molecule indicates that the binding compound is capable of effecting the transcription of the human CYP3A4 gene.

15

Any compound can be tested in the method however, preferred compounds are xenobiotic or steroid compounds.

20

The inventors recognise that a non human animal which comprises a 5' flanking region of CYP3A4 gene but which is deficient for the region from -7836 to -7207 would be useful as a negative control in a method for determining whether a compound is capable of regulating transcription of the human CYP3A4 gene.

25

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. CYP3A4/*lacZ* transgene constructs used to generate transgenic mice. The upstream regions of the human CYP3A4 gene are depicted as open boxes with the position of the XREM at approximately -7.5kb of the CYP3A4 gene indicated by cross-hatching. The 5'-flanking region extended from 56bp downstream of the transcription initiation site to a HindIII site at -3,213 in the construct designated -
35 3CYP3A4/*lacZ* and to a KpnI site at -12,926 kb in construct

-13CYP3A4/lacZ. The coding region of the E.coli *lacZ* gene together with eukaryotic translational initiation and termination signals, transcription termination and polyadenylation sites are indicated by a solid box.

5

Figure 2. Xenobiotic induction of hepatic transgene expression. Female mice from line 9/4 harbouring the -13CYP3A4/*lacZ* transgene were treated with various reagents. Histochemical staining of liver slices with X-gal revealed an increased zone of blue staining cells containing β -galactosidase after treatment with rifampicin, phenobarbital and pregnenolone 16 α -carbonitrile compared with corn oil treated mice.

15 **Figure 3.** Comparison of the xenobiotic induction profile of the -13CYP3A4/*lacZ* transgene with the mouse *Cyp3a11* gene. Transgenic mice from line 9/4 were treated with a range of xenobiotic reagents and naturally occurring steroids. **A.** Transgene expression was assessed by
20 determining β -galactosidase activity in total liver lysates using the ONPG assay. The units of β -galactosidase activity are given as A_{420} /mg liver/minute. Dexamethasone and pregnenolone 16 α -carbonitrile were the most potent xenobiotic activators of the -13CYP3A4/*lacZ* transgene,
25 while rifampicin treatment resulted in relatively low levels. The steroids pregnenolone and 17 α -progesterone were very weak inducers. **B.** Hepatic expression of the endogenous mouse *Cyp3a11* gene was examined in the same mice by Northern analysis. A similar pattern of induction to
30 the CYP3A4/*lacZ* transgene was observed with both xenobiotic and endogenous regulators. The data are presented as the mean +/- the standard deviation for 3 animals.

Figure 4. Dose response of -13CYP3A4/lacZ transgene expression after treatment with dexamethasone. **A.** Male mice from line 9/4 were treated with from 1 to 100mg/kg dexamethasone. Higher doses of dexamethasone resulted in increased β -galactosidase activity (determined in liver lysates as described in Fig. 3). **B.** Zonal expansion of transgene expression with increasing doses of dexamethasone. X-gal staining of frozen liver sections revealed greater numbers of hepatocytes containing transgene-derived β -galactosidase activity after treatment with 1, 10 and 100 mg/kg dexamethasone. At low doses there are limited numbers of transgene -expressing cells immediately adjacent to the central vein. With higher doses there are more cells committed to transgene expression extending across the liver lobule towards the portal tract.

Figure 5. (SEQ ID NO:1) Sequence of the CYP3A4 5'-flanking region included in the -13 CYP3A4/lacZ construct. This sequence corresponds to -12,926 to +56 base pairs relative to the transcription initiation site of the CYP3A4 gene.

Figure 6. (SEQ ID NO:2) Sequence of the 5'-flanking region of the CYP3A4 gene extending from -12,926 to -3,213 base pairs and representing the difference in sequence between the -13 CYP3A4/lacZ and the -3 CYP3A4/lacZ constructs.

Figure 7. (SEQ ID NO:3) The "Xenobiotic-Responsive Enhancer Module" (XREM) of the human CYP3A4 gene. This region encompasses -7836 to -7207 base pairs relative to the transcription initiation site of the CYP3A4 gene.

Figure 8. (SEQ ID NO:4) The 5'-flanking region of the human CYP3A7 gene (Genbank Accession No. AF329900). The extent of the sequences is -11,133 to +52 base relative to the transcription initiation site of the CYP3A7 gene.

Figure 9. (SEQ ID NO:5) Sequence of the 5'-flanking region of the human MDR1 gene (p-glycoprotein gene) encompassing -10,000 to +200 base pairs relative to the transcription initiation site of the MDR1 gene. Sequence derived from within Genbank sequence Accession Number AC002457.

An embodiment of the invention is now described in the following Example which will be understood to merely exemplify and not to limit the scope of the invention.

EXAMPLE

MATERIALS AND METHODS

Transgene constructs. Two transgene constructs were synthesized with the upstream 5' flank of the human cytochrome P450 CYP3A4 gene linked to the E. coli *lacZ* reporter gene (Figure 1). The first construct, designated -3CYP3A4/*lacZ*, contained the region of the CYP3A4 gene from the HindIII site at -3213bp relative to the transcription start site to nucleotide +56bp downstream of the transcription start site. The other construct, designated -13CYP3A4/*lacZ*, included the region of the CYP3A4 gene from the KpnI site at -12,926bp upstream to +56bp downstream of the transcription start site. It includes the DNA sequences of the XREM region located between -7836 and -7208 in addition to the proximal promoter of the CYP3A4 gene. The DNA sequence of the CYP3A4 gene between -10468bp and +906bp has been determined and deposited with the

GenBank/EMBL/DDJB database under accession number AF185589. Additional sequence information covering the region -10,469bp to -12,926bp was obtained from publically accessible Genbank files. The E.coli *lacZ* reporter gene comprises the coding region for the bacterial enzyme β -galactosidase flanked by DNA sequences for eukaryotic translational start and stop signals, SV40 transcriptional termination and polyadenylation signals and an intron. The CYP3A4/*lacZ* transgene constructs were released from vector sequences and purified on agarose gels prior to microinjection

Generation of transgenic mouse lines. Mice carrying the CYP3A4/*lacZ* transgenes were created by microinjection of the DNA constructs into the pro-nuclei of zygotes harvested from FVB/N strain mice. Microinjection and manipulation of embryos were carried by standard techniques. Stable transgenic mouse lines were established by breeding from transgenic founders identified by Southern analysis.

Administration of xenobiotics to mice. 8-10 week old male and female mice hemizygous for the -3CYP3A4/*lacZ* and 13CYP3A4/*lacZ* transgenes were used to test the ability of a range of xenobiotics and hormones to activate expression of transgene-derived β -galactosidase. Mice were administered the following reagents and vehicles by single daily intraperitoneal injection for 4 days: rifampicin/corn oil; dexamethasonephosphate/H₂O; pregnenolone 16 α -carbonitrile/2% Tween® 20 in H₂O; phenobarbital/H₂O; clotrimazole/2% Tween® 20; phenytoin/2% Tween® 20; 17 α -OH progesterone/2% Tween® 20; pregnenolone/2% Tween® 20. All reagents were supplied by Sigma Chemical Co. (St Louis, MO) except for dexamethasone phosphate which was obtained from Faulding (Mulgrave, Australia) and pregnenolone 16 α -

carbonitrile from Upjohn Co. (Kalamazoo, MI). The dose used for all reagents to test for induction of the transgene was 100mg/kg body weight. Dose response studies were carried out in the range of 1-100mg/kg with male hemizygous transgenic mice.

Analysis of transgene and mouse Cyp3a gene expression. β -galactosidase activity was visualised in slices and frozen sections of liver and other tissues by staining with X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside).

10. Tissues were fixed in 0.25% glutaraldehyde, 0.1M phosphate buffer pH7.3, 5mM EGTA, 2 mM MgCl_2 : washed in 0.1M phosphate buffer pH7.3, 0.01% sodium deoxycholate, 0.025% NP40, 2mM MgCl_2 and stained by incubation at 37°C in wash solution supplemented with 1mg/ml X-gal, 5mM potassium ferricyanide, and 5mM potassium ferrocyanide. The level of β -galactosidase activity was determined in whole liver homogenates [100mg fresh tissue/ml 0.25M Tris-HCl (pH 7.3)] using the O-nitrophenyl- β -D-galactopyranoside (ONPG) assay according to standard techniques. After appropriate dilution the homogenate was incubated with β -galactosidase assay reagent (0.1M sodium phosphate buffer (pH7.3)/1mM MgCl_2 /50 mmol β -mercaptoethanol/0.88mg/ml ONPG) at 37°C, quenched by the addition of 1M Na_2CO_3 and the absorbance at 420nm determined. The units of β -galactosidase activity are given as A_{420} /mg liver/minute.

The levels of endogenous mouse Cyp3a mRNA expression were determined by Northern analysis using a riboprobe complementary to nucleotides 852-1061 of the mouse Cyp3a11 cDNA. Filters were stripped and reprobed with an 18S rRNA oligonucleotide to normalise loading.

RESULTS

4 transgenic lines were generated with the construct containing the -3.2kb region of the human *CYP3A4* gene linked to *lacZ*. Transgene-derived β -galactosidase activity was not detected in kidney, large and small intestine, spleen, lung and liver tissue from mice for all 4 - 3CYP3A4/*lacZ* transgenic lines treated with vehicle or xenobiotics (Table 1). In contrast, transgene expression was readily detected in 3 of the 4 lines carrying the - 13CYP3A4/*lacZ* construct. Line 9/4 had a very low constitutive level in the liver, with β -galactosidase detected only in isolated hepatocytes adjacent to major blood vessels. Administration of xenobiotics resulted in robust expression in a zone of cells surrounding the central vein (Figure 2). As the basal level of transgene expression in untreated mice in line 9/4 is extremely low, induction is obvious and is essentially an off/on process. Expression in other tissues in mice from line 9/4 was restricted to the gut, predominantly in the villi of the small intestine.

The relative degree of induction for a range of xenobiotics was analysed by determining the transgenic β -galactosidase activity in liver lysates of mice from line 9/4 (Figure 3A). Dexamethasone and pregnenolone 16 α -carbonitrile were the most potent inducers, while rifampicin activated the transgene to relatively modest levels. Phenobarbital, clotrimazole and phenytoin were intermediate inducers. The induction profile of the transgene in line 9/4 was similar to that observed for the endogenous *Cyp3a11* gene in the same mice (Fig 3B), likely reflecting the activation profile of the mouse rather than the human PXR. Activation of the transgene was observed with naturally occurring

steroids such as pregnenolone and 17 α -progesterone, however the induction was weak compared with xenobiotics.

There was a marked gender difference in hepatic transgene expression, with lower levels observed in females than in males for most reagents. Such a male-predominant pattern was not evident in the induction profile of the mouse *Cyp3a11* gene. Indeed higher levels of *Cyp3a11* mRNA were observed in females than males after treatment with rifampicin and pregnenolone 16 α -carbonitrile. The reason for this apparent reversal in gender-related transgene expression pattern is not known. However, as *Cyp3a11* mRNA is only just detectable in males of the FVB/N strain of mice, it may be attributed to the relatively greater degree of induction of the mouse *Cyp3a11* gene in males compared to females (Figure 3B).

The other line which showed significant transgene expression - 15/10, had a higher constitutive level in both the liver and small intestine in untreated mice.

Expression was not detected in other organs, confirming the tissue specificity observed in line 9/4. The same set of reagents were capable of increasing hepatic and intestinal transgene expression to the same levels as in mice from line 9/4. However, the overall degree of induction was not as great as observed in line 9/4 due to the higher basal level in line 15/10. The induction profile was similar with dexamethasone being the most potent activator and rifampicin the least (data not shown).

Dose response of xenobiotic induction. The activation of transgene expression in line 9/4 by dexamethasone was dose-dependent over the range 1 to 100 mg/kg (Figure 4A). The higher transgene-derived β -galactosidase activity in liver homogenates from mice treated with increasing doses of

dexamethasone was associated with an expanded zone of cells which were stained by X-gal. At low doses of dexamethasone a ring of hepatocytes only 1-2 cells thick around the central vein expressed the transgene (Figure 4B). With
5 100mg/kg dexamethasone the zone of X-gal positive hepatocytes increased to up to 10 cells, approximately midway between the central vein and the portal triad. A similar dose-dependent expansion of hepatocytes expressing the transgene was observed with other reagents and also in
10 line 15/10 which also contained the -13CYP3A4/lacZ construct.

GENERAL INFORMATION

APPLICANT: THE UNIVERSITY OF SYDNEY

TITLE OF INVENTION: TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450
GENE REGULATION

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS: Aird & Berlis LLP
627 Lyons Lane, Suite 204
Oakville, Ontario
Canada
L6J 5Z7

COMPUTER-READABLE FORM

COMPUTER: Dell Dimension L400c

OPERATING SYSTEM: Windows NT

SOFTWARE: PatentIn version 3.1

CURRENT APPLICATION DATA

APPLICATION NUMBER: 2,426,959

INTERNATIONAL APPLICATION: PCT/AU01/01407

FILING DATE: 2001-11-1

CLASSIFICATION:

PRIOR APPLICATION DATA

APPLICATION NUMBER: AU PR 1161

FILING DATE: 2000-11-1

APPLICATION NUMBER: AU PR 4901

FILING DATE: 2001-05-10

CLASSIFICATION:

PATENT AGENT INFORMATION

NAME: Aird & Berlis LLP, Dr. C.B. Barlow

REFERENCE NUMBER: 1483

INFORMATION FOR SEQ ID NO.: 1

SEQUENCE CHARACTERISTICS

LENGTH: 12983

TYPE: nucleotide

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ORIGIN: homo sapiens

ANTI-SENSE:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: CYP3A4 - 5' flanking region included in -13CYP3A4/1aZ
construct; corresponds to -12,926 to + 56 base pairs relative to transcription
initiation site of CYP3A4 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 1

CTGGTTCATC TCATTGGGAC TGGTTGGACA AGAGGGTGCA GCCCACGGAG GGTGAGCCAA	60
AGCAGGGTGG GGCGTCGCCT CACCTGGGAA GCACAAGGGG TCGTGGAATT TTCTCCCCTA	120
CCCAAGGAAA GCCATAAGGG ACTGAGCCTG AGGAACTGTG CACTCTGGCC CAGATACTGC	180
ACTTTTCCCA TGGTCTTTGC AACCCGCAGA CCAGGAGATT CCCTCCGGTG CCTATGCCAC	240
CAGGGCCCTG GGTTTCAAGC ACAAACCTGG GCAGCCATTT GGGCAGACAC CGAACTAGCT	300
GCAGGAGTTT TTTTTTTTTT TTTCCATACC CCATTGGCAC CTGGAACGCC AGTGAGACAG	360
AACCGTTCAC TCCCCTGGAA AGGGGGCTGA AACCAGGGAT CCAAGTGGTC TGGCTCGGTG	420
GGCCCCACCC CCATGGAGCC CAGCAAACAA AGATTCACCTT GGCTTGAAAT TCTTGCTGCC	480
AGCACAGCAG CAGTCTGAGA TTGACCTGGG ACCCTCGAAC TTGGTTGGGT GCTGTGGGGG	540
GGCATCTTCC ATTGCTGAGG CTTGAGTAGG TGGTTTTACC TTCGCGGTGT AAACAAAGCT	600
GCTGGGAAGT TTGAACTGGG TGGAGCTCAC CACAGCTCAG TAAGGCCACT GTGGCCAGAC	660
TGCCTCTCTG GATTTCTCCT CTCTGGGAAG GATATCTCTG AAAAAAAGGC AGCAGCCCCA	720
GTCAGGGACT TATAGATGAA ACCCCCATCT CCCTGGGACA GAGCCCCTCG GGAAGAGGT	780
GGCTTCCACC ATTGTGGAAG ACTGTGTGGC AATTCCTCAC GGATTTAGAA CTAGAGATAC	840
CATTTGACCC AGCAATCCCA TTACTGGGTG TATACCCATA GGATTATAAA TCATTCTACT	900
ATAAAGACAC ATGCACACTT ATGTTTATTG TAACACTATT TACAATAGCA ATGACCTGGA	960
ACCAATCCAA AAGCCCATCA ATGATAGACT GAATAAGAA AATGTGGCAC ATATACACTG	1020
TGGAATACTA TGCAGCCATA AAAAAGGATG AGTTCATGTC CTTTGCAGAG ACATGGATGA	1080
AGCTGGAAAC CATCATCTC AGCAAACCTAG CACAATAACA GAAAACCAAA CACTGCATGT	1140
TGTCACCTCAT AAGTGGGAGT TAAACAATGA GAACACATGG ACACAGGGAG GGGAACGTCA	1200
CACACTGGGG CATGTCGGGG AGTGGGGGCC TACGGGAGGG ATAGCATTAG CAGAAATACC	1260
TAATGTAGGT GACGGGTTGA TGGGTGCAGC AAACCACCAT GGCACATATA CACCTATGTA	1320
ATAAACTGC ACGTTCTGCA CATGTACCCC AGAACTTAAA GTATAATTAA TAATAATAAT	1380
AATTTCTGGG CATGTAAGTA GCTGTCTTTC AGGTCTACT TTGATACATA TTCTGAGAGA	1440
ATTAAACCTG TCAAAGAAAC CTTGACTTTC AATGGCAGGC ACTGGAATTG ACCCTAATAA	1500
TGTGTTTTGG GGTAAGCCTA CTCATATTCT CAACCTGTCT GCAGTAGTCG TTAGAATCTG	1560
AACTTCCTGA AGTTCATGTG CAAAGTTGAG TTAATTGTTT AATATTCAAC AAGGATTATG	1620

CCAGTAAGAT	GGTAGGAAAA	TATTAGATAT	GTGTCATCAC	TGCTGGTATT	ATTTAAACTG	1680
CAACATATTT	TAGCTGGCTG	CTGATCTCAG	CCACCATGCC	TGCATTTTAT	CTCTGTCTCG	1740
TGGTCTGCAA	CCTTGGAAGC	TTTGAACCTA	GCTCATAGAA	TCCTGGGCAT	CAAGAACATG	1800
TGGTTCTAAT	GGCTAGATAG	GGAATGAGAG	TAAAAGGATT	TTGCCCACGG	TCACGTGAGT	1860
AAACAACAGA	TTTGGAGGGG	TCTGGACTAC	TGTGATGACT	TCATTCTGAC	AATATGTTCC	1920
AGTTGTCCTT	TCATTTCCCTC	CTAATCACAT	GTCTGGTCTG	ATCTGGCTGT	TTCCCACCTT	1980
CCAATTCCTG	CCTTCTCCAA	TGCTCCCTTC	CGTAGGTCAC	TCTGTGGCTC	AGAGACCCTG	2040
CTTAGCAAGC	GCCCAACCTT	TCAATTATTT	G TTCAGTAAA	ACTTGAACTC	ATGTCTCCCC	2100
TTCTTGATAA	AAAGAAAATA	CGTTATGTAA	TGTCGGGTTA	CTCTATAACT	CTTGTCCTGT	2160
CTCTCGGCAA	CTACTGAACT	AACTGTTTTC	ATATTGAGCA	AACGTTTATG	GAAGGACTGC	2220
CAAGAGTCAG	GTACTAGGCT	TGGTAATATT	CCCCGTTCTC	TCTAGTCAAA	GCCAACACCA	2280
GCCAGACTTG	CAGATCTAGG	TCCCAAGCCC	ACTGCAGATC	ACAGGCCAGG	GTCTGGTCTC	2340
CTCTGAGCTC	CTTTGGGAGG	GAAAGACAGA	ATTATTAACA	CCCATTTTGT	AGATTAGGCA	2400
ACTGAGGCTG	AGGAAGTTTA	AATAACTCAG	ACAGGGCCTG	CACGTCAGTC	ATATTCCAAG	2460
GATCCCTACT	CACTGTCTTC	TCTCTACAGA	ACGAGATGTC	TCTGGAGTCC	ATAGAAAGCC	2520
CAGGAGCCTG	GCTGGGCACG	GTGGCTCCTG	CCTGTAATCC	CAGCACTTTG	GGAGGCCGAG	2580
GCAGGCAGAT	CACCTGAGCT	CAGGAGTTCA	AGACCAGCCT	GGGCAACATG	GCAAAACCCC	2640
ATCTCTACTA	AAAATACAAA	AAATTAGCTG	GGCGTGGTGG	TGCATGCCTC	TAATCCCAGC	2700
TACTTGGGAG	GCTGAGGCAC	AAGAATTGCT	TGAGCCCAGG	AGGCAGCAGT	TGCAGTGAGC	2760
TGAGATTGTG	CCAGTGCACT	CCAGCCTGGG	CAACAGAGCA	AGATTCCATT	TCAAAAACAA	2820
AAACAAACAC	AAACAAACAA	ACAAAAATAG	AAAGCCCAGG	GACCACCTGC	GTCAGGTTCC	2880
CAGCCACACC	TTTTTCTTGT	CCTCCTCTGT	CTCTGGCATC	TTCTCACAGG	TTCCTAATTG	2940
TTTGTGGTTG	CACAAATTCA	AAATCCCAGA	AAAATTACCA	CTTCACACCC	ACTCAGATGG	3000
CTATTTTTTT	TTTGAAGGAA	GATAACAAGT	GTTGACAAGA	ACATGGAGAA	ATTGGAATTC	3060
TCACCCATTG	CTGGTGAGAA	TGTAATACGG	TGCTGCTGCT	ATGGAAAACA	GCTTGGAGTT	3120
TCCTCAAAAA	GTTCAACAGA	ATTTCAATGT	GACCCAGCAA	TTCCCCTCTA	AGTTATAGAT	3180
CTGAGAGGAT	TAAAAACAGT	TACTAAAATA	CACGGACTCA	CATATTTCTA	ACAGTCCAAT	3240
TCACAAGGGC	CAAAAGGTGC	TAATAGCCCA	CATGTCCATC	GATGGATGGA	TAAATAAATT	3300
GTGGTCTATC	CATACAATGG	AATATTATTC	GGCCATAAAT	GGAATGAAGT	ACTGACGCAT	3360

GCTACAGAAT	GGATGAACCG	CAAAAAAAT	GGATGAACAC	ATGCTACAGA	ATGGATAGCC	3420
TCACTTTACT	ATGAAGTGAA	GGCCAGAAAC	GAAGTCCATA	TATTGCATCA	TACAAAATAT	3480
CCAGAAGAGG	GAAGCCCACA	GAGACAGAAT	GTGCAATGGT	GGATGCCAGG	GTCTGGGGAG	3540
AGGGGAGAGT	GGGGAGAAAC	TGCTCAACTG	GTACAGGCTT	TATTTTGGAA	TGATGGGAAC	3600
ATTTTGCAAC	TAGATAGAGG	TAGTGATTGC	AGAACACAGA	ATGTACTGAA	TTCCACTGAT	3660
TTTTTTCACC	TTAAAATGGT	TAATTTTCAG	TCCTGAGATT	GGATAATCAT	AAAAAATGG	3720
TTAATTTTAT	GTTATGTGAA	TTTCATCCCT	ATACATATTT	TAAACCTCAG	AAATATACAC	3780
TAGCAGGCAT	GGAACAGGTC	ACTGTGGTGC	CTGCCAAGCC	CGGTGATGTT	ATCTGGGGTC	3840
CCCGGCCAGC	CTTAAGCCTC	TTGCTGACCG	GTGGAGGGCA	GAACCTTTGC	CCTAAAAGTA	3900
TAATATCCAC	ATGCTGGCAT	GATTCCTGGC	CAGATGGCTT	CTTTATTAGC	AGTAATTGAA	3960
ACTGCCTCGA	TACAGACACT	GTACCTTGCA	ACCAAAAAAT	GACTCAACAA	TGATAATAAG	4020
GGTTAAGCTG	GGCCTTTCTC	TCTTTGCCAG	TTAAATTATA	TTTATTATAG	CTTGACATGA	4080
AAAACAAAGC	AACTCCAACA	GGTATCACAA	GGGCAAAGGA	CATGAACATT	TTATCAAAGA	4140
AGAAATGCAG	CTGTCAAAAA	TACAGAAATA	TTCAACCTTG	TTCATAATAA	AGTGGCTGGG	4200
CTCAGTGGTT	CATGCCTGTA	ATCCCAGTGC	TTTGCAAGGC	TGAGACAGGA	GGATCATTTG	4260
AAGCCAGAAG	TTCAAGACCA	TCCTAGGCAA	GTCAGTTCAA	TACCAGACTT	CATGTCTACA	4320
AAACATCAAA	AAATTAGCCA	GGCATGGTGA	TGCATGCCTG	TTGTCCCAGC	TACTCAGGAG	4380
GCTGAGGCAG	GAGAATTGCT	TGAGCCTGGG	AGGCTGCGGT	GGCGGTGAGC	CATGATTGTG	4440
CCATTGTACT	CCAGCCTGGG	CAATGCAGCA	AGACTGTCTA	AATAACAAAA	ATAATAGTAA	4500
AGAAAAGGAT	TGGGATGCCA	TTTACTTGCG	TATTCAATAC	ACAGAGTTAA	AAGTAATTTT	4560
TACGTTTTCT	ATTTTTTTAT	TACTAAAAAA	AGCTGGACCA	TTCTCACAGC	CTGAAATGCT	4620
TCTCACTTTC	CCTTCTTCTG	TCCAAACACT	TCTCTATGAT	AATGCAAACA	GTCACTCCTT	4680
TAGGAAGACT	TCACCCCAGG	TAGTTCCAGA	TCCCCTTATC	TCTGCCTTCC	CAGAACTCCT	4740
GGTGTCTCTC	CAGTTCCCTC	CGTGTGGTGA	AGTACCCTAC	CTAGGGTTTC	AGTATGGCTC	4800
TGTCTGCAAA	GGTCTTGTTT	ACACCTTCCC	TTATGGTTCT	GTTGCCCTGT	GTTGTGTCAT	4860
AGCACAGGGC	ACAGTGGAGA	ACCCATTAC	ACTGATAGAG	AGGGCCCCAT	GGTCCTGGAG	4920
ATAACCATGT	AACCGATCAG	AATAAGGCAT	TGAGGGCTGG	GTGTCAGGCG	TGGGCTGCAC	4980
TTGGGTGGGC	AGGTCCCCTG	GAAAGTCACT	GGGTTTGGCA	AGCTTCCTAG	TAACATGTCT	5040

CTCTGGGGTC	CCCCTTGGAA	CTTCATGCAA	AAATGCTGGT	TGCTGGTTTA	TTCTAGAGAG	5100
ATGGTTCATT	CCTTTCATTT	GATTATCAAA	GAAACTCATG	TCCCAATTAA	AGGTCATAAA	5160
GCCCAGTTTG	TAAACTGAGA	TGATCTCAGC	TGAATGAACT	TGCTGACCCT	CTGCTTTCCT	5220
CCAGCCTCTC	GGTGCCCTTG	AAATCATGTC	GGTTCAAGCA	GCCTCATGAG	GCATTACAAA	5280
GTTTAATTAT	TTCAGTGATT	ATTAAACCTT	GTCCTGTGTT	GACCCCAGGT	GAATCACAAG	5340
CTGAACTTCT	GACAAGAACA	AGCTATCATA	TTCTTTTCAA	TTACAGAAAA	AAGTAAGTTA	5400
ATTGATAGGA	TTTTTTTTGT	TTAAAAAAA	TGTTACTAGT	TTTGAAAAGG	TAATATGTGC	5460
ACATGGTAAA	CACTAAGAAG	GTATAAGAGC	ATAATGCTTT	TATACTACTA	AGAATAATGT	5520
TTTCTCTAAG	TTTTTTTTTG	TAGATGCTTT	CATCAGATTA	AGAAAATTCC	CTGCTATTAG	5580
TTGTTGAAGG	TTTTTATATC	ATAAATGAAA	GTTGAATATT	ATTATCATAT	ATTATTAATA	5640
TATTGTTATT	GAACTATCAA	AGCCTTTTCC	TAAAACCATT	GAGATGATCT	TATAACCATT	5700
CTCCTTTAAC	CTGTTGACGA	GATCATTTGG	ATTTATACTA	TTTCTCTGTT	AACCATTCTT	5760
GAGTCTCAGG	TTTAAATTCA	ACTTGGTCAT	GGTGTGTCAT	CTTTGATCAT	TGCTGTCTGT	5820
GGCTTGCTAC	TGTTTTGTTT	AGGATTTTTG	CACTGATGCT	CATCAATGAG	ACTGGCATGC	5880
CATCTTCCTT	TGCAGTCCTG	ATTTTTTTCT	GATTTGGATC	ATGTGGTTAT	GGCCCTCATG	5940
GAATGAGTTG	GGCATGATGC	CTTTTTTTCA	TGTCTCTGGA	TTGATGGGAC	ACTTTGGATT	6000
CTCTCCAGAT	GGCCCTCAAT	GGTCCCTGCC	TCCTCATTGT	TAGGCCCTG	GGCAAGCCCT	6060
TCTCATTTCT	GGTAGGCCCA	GGAACCTGTG	GGGGTTTTGT	TTGTTTGTTT	GTTTCTTGAG	6120
TCGGAGTCTC	ACTCTGTCAC	CCAGGCTGGA	GTTGGAGTGC	AATGGCCCGA	TCTTGGCTCA	6180
CTGCAACCTC	CACCTCCCAG	ATTCAAGCAA	TTCTCCTGCC	TCAGCCTCCT	GAGTAGCTGG	6240
AATTACAGGC	ACCCACCGAC	ACACCCTGCT	AATTTTTGTA	TTTTTAGTAC	AGATGGGGTT	6300
TCACAATATT	GGCCAAGCTG	GTCTCGAACT	CCTGATCTCA	TGATCTGCCC	GGCTTGGCCT	6360
CCCAAAGTGT	TGAGATTACA	AGCATGAGCC	ACCACACCCA	GTGAACCTGT	GGTTTTTAGA	6420
AGCTCCCCAT	GCATGTGAAT	GCTGTGAGCA	TCCCAGGATG	ACAGCCACTG	TGTGTTTCAGC	6480
TGTTGGAACT	GTGAGAAAGC	ACCAGTGGGA	CCTTCTCCAG	CACCTGCCTG	CTGAGTTCAT	6540
GGAAGAGGCT	TGTTGGGGAG	ATGATGCCCT	GGCTGACTCC	TGAAGGATGG	TTAGGAATGC	6600
ACCAGATGGA	AGCTGGGTTG	GACCCACTCT	ATGCTGAAGA	ACAGCTTGTTG	TGGACACAAG	6660
GAGACACGGA	TATGTCATTT	TTGTAGAGCC	TGAGGAGTGT	CCAATCACAC	CATTTGCTTA	6720
AAACATCATG	CACACTTGGA	AAAGTGGACT	GAGACCGAAT	GAAGAAGCTA	ACAGTGGCCA	6780

GATCAGAAAG	GGTCTTGTGT	TACTTCCTAG	AGATACTTAG	ATTTTATCCT	GTGGGTGATA	6840
GGAGCAGTTG	GAGGGACTGA	AGACAAGGAA	AGAAACATGT	TTCAAGATCT	ATGTTTTTCA	6900
AGACGCTTTT	CTGGTGGCTG	AGTAGGGAAT	TCCCTGGATA	AGTCCTGCCC	AGGGTCAGGC	6960
AAAACAAGTT	AGGGGGTTAC	TGAAATAAGG	AGTATGAGAA	ATGGTGTAGG	TTGTGCTGAC	7020
GTTTTGTAAC	ACATCTCATG	ATGATCTTCA	TTTCCTTCAC	TAATTTCTCT	TTTCATTAAT	7080
TCCCTTCCAC	GTGCTCTTCT	GAAATTTGCC	TCACATTCTC	TGATTTCTCT	TTTACCTGTT	7140
GGTTTCATCA	CCTTTTACTT	TTTGCTTTCC	TGGAAACACA	AATGATTCTG	ATTGTGACAT	7200
GTCAGAATTA	TTTGCAACAT	TTGCCTTTCT	GCTGAAACCA	TGAGTTCACT	GAATACACAA	7260
TTTAGTAAAG	TGTAGGATGC	ACATGTCGTT	TTCGTGGTCA	CAACCAGCTC	TGTAGCATTT	7320
TATAACTACA	CTGGCAGTGT	GCTGGGAGGT	GTAGAGAGAA	ATATTTATCA	CATGTGTGGC	7380
TGACACAACC	TGCCAAGTTA	TTTTAGGAGC	CTCCTTGGA	TCCCAGCAAG	AATGCTACCG	7440
GCACAATTTG	TAATCACAGC	ATCCTGCTCC	ATGCCTTGGC	TTCATGGCAT	AGTCACTTCT	7500
GCAAGTCTCT	TTCCAGCTGT	CTGTTCCCAT	GTCTATAAAG	TATGAGTTAA	ATCATCCTAA	7560
CACTACTCAT	CTTACAAAGT	TTTCTTGCTG	ATGTTAAGAG	AGTTGGGAAA	GAAGTGTATA	7620
AACTGTGAAG	TGCCATGGAG	ATGTTAGTGG	TTACTTTATC	AAGAAATAGA	CACTCTAGAA	7680
TGGAGTAGAA	AGCCAACAGT	TATGATTGAG	TCCTCCTCCT	CTTCTTCTTT	TTATTAATTT	7740
ATAAAGAAAA	GAGGTTTAAT	TGACTCACAG	TTCCATATGG	CTGGGGAGGC	CTCGGGAAAC	7800
TCTCAGTCAT	AGCAGGAGGC	AAAGGGGAAG	AAGGCACCTT	CTTCACAAGG	CGGCAGGAGA	7860
GAGAGAGCTC	CTGTTCTTTT	TTGTCATAAA	GTCTACAGAA	GTGCTTATAC	TTCAGGACAA	7920
GGGCAGGCAG	AGAGAAGGAA	GGACATTGCT	TCACCCCAGC	CCTCACTGAC	GAGTTTGCTA	7980
GGGGACCTCA	CTTTGTCCCA	GAGTAGGGCA	GAAGTCTGGC	CACTACCCAT	TCAGAAGGCC	8040
TGGGCTGCAC	TGCTAGTTCC	TCACTAACTC	TGTGTGGCCT	TGGGCAAGGT	TGGGCCTGTG	8100
TTAACAGATT	ATGACCCTGG	GCTCTCAAGC	TAGAGGATCT	AAATTTGAAT	CCTGGCTCTG	8160
CTAAAGCAAT	TAGTGATGTA	AACTTTAATG	GGTCAGTTAA	CCTTCCTGTG	GCTTAGTTTG	8220
CTCATCTGTA	AAATAGGGAT	CATAACAGTA	TCAATACCAC	ATGATTGTTG	GACAGATTGA	8280
ATCAGTTAAT	GCAGGGGAAG	TACTTAGCAT	GACACGTATT	CACTATCATT	TCCTGGAGTA	8340
AGAGCTGTGT	GTGAGTGGGT	GTGAGCATGT	GTGAAACCTT	TTCTCTGCAA	TCTCAGTTAA	8400
GAAACCAATC	CAGAATTTAA	AGTTCAGGGC	CTAAATGGGT	GGTTATCTTC	TCCCAGTTCC	8460

ATCCTATCCC	ACCTTTGCTC	TTCCTCCCGC	CCACAGGAGC	TGTTGGTCCT	TGATTGGGCT	8520
GGAAGACCTG	GTGGACCCTA	AGTGATCTAT	AAGAGGAGAA	TAGAGAACAG	GGAATGTCTT	8580
CAAAAATCTA	GAGGGACACA	GAGGCTGAGA	GGCAGGCAGT	CCTGCAGGGT	CTTCTGATTG	8640
GGACAAGGAG	AACCTTGGTC	TTCACAGGCC	AATTCTGGTC	AGTTTCCCCC	ATGGACAGAT	8700
GAGGAAACAG	GCCCAGGAAT	ATCCAAGGTC	TCACACTTCC	CATCTGTCAA	GTCTTGTTGA	8760
TTCTGTTGTA	TTCATGTCTC	TCAAAGGGAG	ATAGAGTTTA	GGGAAGAAAG	AAGGATCAAC	8820
TGTGTCTGAT	ACCACTGGGA	GCTTAAGTAA	AGGGTTCTTT	TACTTCATAG	CATTTATCCC	8880
AATTTGTAAT	TCAGTATTAT	TTGTGTGGCT	GTTTGGTGTC	TCTTTCTCCT	ATATGAGTGC	8940
TAGCTTCATA	AGGGCAAGGA	TTTTGATTCT	TTAATATTTA	GTGCTTGCCA	CATGCCCTGA	9000
ACACAGCAGG	CATACAGGCT	AACCAACATA	CAGTGGCATG	AAAGTCATGA	AAGTGAGACA	9060
CCTACCTCCT	CCAGTGCCAA	GAGAGCATAA	CCATGCACCT	GTCACTCTCC	TCAACACCAC	9120
CCCCAAGCAT	GAGGCCCAA	AGCATTAGCT	AATCCCCTCC	TCCAGCCACT	AAAAC TTAAA	9180
GGCCAGGTGT	GGTGGCTCCC	ATCTGAAATC	CCAGAACTTC	AGGAGACAGC	AGCAGGAGGA	9240
TCACTTGAGG	CCAGGAGTTT	GAGATCAGCC	TGGGCAACAT	AGCTAGGTCC	CATCTGTACT	9300
AAAAATTAGC	TGGGCGTTGT	TGCATGCCTG	TAGTCCCAGC	TACTAAGGAG	GCTGAGGTGG	9360
GAGGATCACT	TGAGCCCAGG	AGGTGGAAAC	AACAGTAAGC	TATAATCACA	GCACTGAACT	9420
CTAGCCTGGG	CAACAGAGTG	ACACCCTGCC	TCAAAACAAT	TTTAAAAATA	AATAAGAGCA	9480
AAACTTAGAT	ACCACGTGGT	CACCCCAACA	TGCAAAATCA	AGTTTTCCCC	TACTGAGAAG	9540
AATGGGGACT	TGACAGCTGA	GTTACAGAGA	GATAATCTTC	TTCTTCTTTT	TTTTTTTTTG	9600
GTTTACATCC	TCAAGATCAT	GACTTGTGAA	ATTTGAATCG	AATACACATG	TAATTCCAGA	9660
GCAATGTTGC	CTCCGCATAC	CATCAGCAAT	TCACTTGGCT	ACTGGAAGTC	AGGATAAGCT	9720
TCCCAGAAGA	GAGGTACCAC	TTGGGCTACC	AATATAAAAG	GATGAAAATA	TCAGAGTGAT	9780
GGTGTTCTTT	ACAACGTTGA	GTCCCTGGAC	AGCCTGTCCA	CTGATGCTGA	TATCTGAGCC	9840
TAATGCTTCT	CTGAATGTTG	AGATTGAACT	TTGATCCAAT	GAAACTAGAA	CGAGAAAGAA	9900
GATAAGTCTT	TCATTGTTGA	TAAGGACATT	ATGTTTCTCA	TACTTGTATG	ATTATTTTTC	9960
CTTAGCTGTA	CTATAATTAT	CTGCTTATTT	GTCTCTGCTC	TATGTGCTTA	GGGTACAAAG	10020
TTGACCAAGA	CCAAC TTTGG	TTGGAAGCAT	AGTACTAAGA	GCACAGTACT	GAGAGCACAG	10080
TATTGAGAGC	ACAGCTTTAA	AAAACATGAT	GAAGGCTTTA	ATACAGGAAA	TGAGCAGGGG	10140
AGAGGCATGT	GGTGGTTGGA	TGTATCTTCC	TTGACACAGT	CAGTGCAGCT	CTCAGTAGTC	10200

AAGTCCCTAC	ATGTTAGAAG	ATGTTACCTT	CTGTGGAATT	AAGTGGCAGA	ACTTGCCTTC	10260
AATTATTTTC	CTTTGCAGAA	CAACACCAAC	TGCATTAGTT	AGGACACAGT	GCTGGCTGCA	10320
TTTAAGTCCC	AAGCGATGAT	TAGTCTCTCA	CTGTTGGTAT	AGATTCAAAC	CAATCAGACC	10380
ACCTCCTAAA	GTTTGTAGGG	CAGGTAAATC	CTCATCTTAG	AATAAAAATC	ATCTTACCAA	10440
GTATGTGTTT	TAGAGGCAAG	AAGAAAACAT	ATTTGTTTCT	GTAAGAGTTT	TGTTTAAAAA	10500
AAATATAAGA	AAGGCTCTCG	GTTTAGGTGA	GGTAATGAAG	TTGTTGATAG	TTATCAGATG	10560
ACACTGGAAT	CTTTACTTCT	CTGAACGTGT	TCTGTGCATC	TCTCAGTGTG	GGAACATAGA	10620
GAGGGAGATC	CTCCAGCAAT	GCCACTGATA	TGGTCAGAAA	CTGCATCTTT	CTTTCTCCCT	10680
GCTGAGATGA	GATGGAGTCC	TTTGTCTAG	AAGACCCATG	GTGGTGCCGC	TGGGAGTAAC	10740
CCTTGAGACA	GGAACACAAA	TCCCAACCAA	TTTGTGGTTG	CAGCCTTGAG	TCTCACTATT	10800
TCCCATAGTG	ATGCGTAGCA	GGGAATGGCA	GGTGCACCAG	AGCAGGAGAG	GACCTAATAT	10860
CTCCCTTCCT	GTTAGCTTTT	TATAAAGTTT	TATTGTGATC	AGTAGCAGTT	GGGAAGCTAC	10920
TTGCAGTCAC	TGAGCCTCAG	TTTCTACATC	TGTAAACTGG	GGATAGTAGC	ATGGCCCCTA	10980
CTTAATGTGC	TCAGCAAAGC	CACTGAAAGG	AGACAGAAAT	GTATCTAAAT	TACCCTGGAC	11040
TTTTATCCTA	CCTCTCTTGG	GGATTGTCAC	CACCTTCCCA	TGTTTGTCTT	TTTTGGTTTG	11100
ATGCTTGCTG	TCACTTCTTT	CCTTAGGTGC	CTCTCTGTAC	GGCTCTTTTA	TCCCAGGGAT	11160
TCCAGAGTTA	CAGCACATGC	ATACCACCAT	CCAAGCATGT	TTATTTGTCT	CCTGCTTCAC	11220
TAGGCTGTCC	CCAAGGAACA	TGTGGCTCCC	GGCACACACC	TGGCACAACA	CTGCACATGA	11280
CATTCACCCA	CTTGGCCTTG	AATCTGACAA	GGAATCTGGC	ATGATGTTCA	CCCACTCAGG	11340
CCAGGTGCCG	AGCAGCCCTG	GAGGCTTAGG	GGCCAGAGGG	ATGGGAAAAG	GTGTCTTTCT	11400
GGGGTGAGTA	TCAGTTTCTG	CAGGAGGGCT	GAATGTGAGA	AAGAATAAAG	AGAGAAGGAA	11460
GCGAACAAGC	ACAGCTTAAA	CATCGCCTAT	TTCTATTGAG	TTTTAAGAAC	GCTGTGATTT	11520
TGTTTGTGTCAT	GCAATCCATT	CATCAGGCCA	GGCAGACACA	GAACTTGGGT	GTGAGTGACG	11580
ATAATGAGCT	GATATAATTT	TCACACCCTC	ATCACTGAGA	TCTCTCCCAT	CAGGAATGGG	11640
TCAGGGAGCT	CACAGGTGGC	AGCAACTGCT	ATTACAGGCC	TCATCTCTAC	CAGCTCCTGG	11700
GGCCTGCCCT	CCTCCCATTA	GAAAATCCTC	CACTTGTCAA	AAAGGAAGCC	ATTTGCTTTG	11760
AACTCCAATT	CCACCCCCAA	GAGGCTGGGA	CCATCTTACT	GGAGTCCTTG	ATGCTGTGTG	11820
ACCTGCAGTG	ACCACTGCCC	CATCATTGCT	GGCTGAGGTG	GTTGGGGTCC	ATCTGGCTAT	11880

```
CTGGGCAGCT GTTCTCTTCT CTCCTTTCTC TCCTGTTTCC AGACATGCAG TATTTCCAGA 11940
GAGAAGGGGC CACTCTTTGG CAAAGAACCT GTCTAACTTG CTATCTATGG CAGGACCTTT 12000
GAAGGGTTCA CAGGAAGCAG CACAAATTGA TACTATTCCA CCAAGCCATC AGCTCCATCT 12060
CATCCATGCC CTGTCTCTCC TTTAGGGGTC CCCTTGCCAA CAGAATCACA GAGGACCAGC 12120
CTGAAAGTGC AGAGACAGCA GCTGAGGCAC AGCCAAGAGC TCTGGCTGTA TTAATGACCT 12180
AAGAAGTCAC CAGAAAGTCA GAAGGGATGA CATGCAGAGG CCCAGCAATC TCAGCTAAGT 12240
CAACTCCACC AGCCTTTCTA GTTGCCCACT GTGTGTACAG CACCCTGGTA GGGACCAGAG 12300
CCATGACAGG GAATAAGACT AGACTATGCC CTTGAGGAGC TCACCTCTGT TCAGGGAAAC 12360
AGGCGTGGAA ACACAATGGT GGTAAAGAGG AAAGAGGACA ATAGGATTGC ATGAAGGGGA 12420
TGGAAGGTGC CCAGGGGAGG AAATGGTTAC ATCTGTGTGA GGAGTTTGGT GAGGAAAGAC 12480
TCTAAGAGAA GGCTCTGTCT GTCTGGGTTT GGAAGGATGT GTAGGAGTCT TCTAGGGGGC 12540
ACAGGCACAC TCCAGGCATA GGTAAAGATC TGTAGGTGTG GCTTGTTGGG ATGAATTTCA 12600
AGTATTTTGG AATGAGGACA GCCATAGAGA CAAGGGCAAG AGAGAGGCGA TTTAATAGAT 12660
TTTATGCCAA TGGCTCCACT TGAGTTTCTG ATAAGAACCC AGAACCTTG GACTCCCCAG 12720
TAACATTGAT TGAGTTGTTT ATGATACCTC ATAGAATATG AACTCAAAGG AGGTCAGTGA 12780
GTGGTGTGTG TGTGATTCTT TGCCAACTTC CAAGGTGGAG AAGCCTCTTC CAACTGCAGG 12840
CAGAGCACAG GTGGCCCTGC TACTGGCTGC AGCTCCAGCC CTGCCTCCTT CTCTAGCATA 12900
TAAACAATCC AACAGCCTCA CTGAATCACT GCTGTGCAGG GCAGGAAAGC TCCATGCACA 12960
TAGCCCAGCA AAGAGCAACA CAG 12983
```

INFORMATION FOR SEQ ID NO.: 2

SEQUENCE CHARACTERISTICS

LENGTH: 9715

TYPE: nucleotide

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ORIGIN: homo sapiens

ANTI-SENSE:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: 5' flanking region of CYP3A4 gene extending from -12,926 to -3,213 base pairs

SEQUENCE DESCRIPTION: SEQ ID NO: 2

CTGGTTCATC TCATTGGGAC TGGTTGGACA AGAGGGTGCA GCCCACGGAG GGTGAGCCAA	60
AGCAGGGTGG GCGTCGCCT CACCTGGGAA GCACAAGGGG TCGTGGAATT TTCTCCCCTA	120
CCCAAGGAAA GCCATAAGGG ACTGAGCCTG AGGAACTGTG CACTCTGGCC CAGATACTGC	180
ACTTTTCCCA TGGTCTTTGC AACCCGCAGA CCAGGAGATT CCCTCCGGTG CCTATGCCAC	240
CAGGGCCCTG GGTTC AAGC ACAA AACTGG GCAGCCATTT GGGCAGACAC CGAACTAGCT	300
GCAGGAGTTT TTTTTTTTTT TTTCCATACC CCATTGGCAC CTGGAACGCC AGTGAGACAG	360
AACCGTTCAC TCCCCTGGAA AGGGGGCTGA AACCAGGGAT CCAAGTGGTC TGGCTCGGTG	420
GGCCCCACCC CCATGGAGCC CAGCAAACAA AGATTCACTT GGCTTGAAAT TCTTGCTGCC	480
AGCACAGCAG CAGTCTGAGA TTGACCTGGG ACCCTCGAAC TTGGTTGGGT GCTGTGGGGG	540
GGCATCTTCC ATTGCTGAGG CTTGAGTAGG TGGTTTTACC TTCGCGGTGT AAACAAAGCT	600
GCTGGGAAGT TTGAACTGGG TGGAGCTCAC CACAGCTCAG TAAGGCCACT GTGGCCAGAC	660
TGCCTCTCTG GATTTCTCCT CTCTGGGAAG GATATCTCTG AAAAAAAGGC AGCAGCCCCA	720
GTCAGGGACT TATAGATGAA ACCCCCATCT CCCTGGGACA GAGCCCCTCG GGGAAGAGGT	780
GGCTTCCACC ATTGTGGAAG ACTGTGTGGC AATTCCTCAC GGATTTAGAA CTAGAGATAC	840
CATTTGACCC AGCAATCCCA TTA CTGGGTG TATACCCATA GGATTATAAA TCATTCTACT	900
ATAAAGACAC ATGCACACTT ATGTTTATTG TAACACTATT TACAATAGCA ATGACCTGGA	960
ACCAATCCAA AAGCCCATCA ATGATAGACT GAATAAAGAA AATGTGGCAC ATATACACTG	1020
TGGAATACTA TGCAGCCATA AAAAAGGATG AGTTCATGTC CTTTGCAGAG ACATGGATGA	1080
AGCTGGAAAC CATCATTCTC AGCAA ACTAG CACAATAACA GAAAACCAAA CACTGCATGT	1140
TGTCACTCAT AAGTGGGAGT TAAACAATGA GAACACATGG ACACAGGGAG GGGAACGTCA	1200
CACACTGGGG CATGTCGGGG AGTGGGGGCC TACGGGAGGG ATAGCATTAG CAGAAATACC	1260
TAATGTAGGT GACGGGTTGA TGGGTGCAGC AAACCACCAT GGCACATATA CACCTATGTA	1320
ATAAACTGC ACGTTCTGCA CATGTACCCC AGAACTTAAA GTATAATTAA TAATAATAAT	1380
AATTTCTGGG CATGTAAGTA GCTGTCTTTC AGGTTCTACT TTGATACATA TTCTGAGAGA	1440
ATTAAACCTG TCAAAGAAAC CTTGACTTTC AATGGCAGGC ACTGGAATTG ACCCTAATAA	1500
TGTGTTTTGG GGTAAGCCTA CTCATATTCT CAACCTGTCT GCAGTAGTCG TTAGAATCTG	1560
AACTTCCTGA AGTTCATGTG CAAAGTTGAG TTAATTGTTT AATATTCAAC AAGGATTATG	1620
CCAGTAAGAT GGTAGGAAAA TATTAGATAT GTGTCATCAC TGCTGGTATT ATTTAAACTG	1680

CAACATATTT	TAGCTGGCTG	CTGATCTCAG	CCACCATGCC	TGCATTTTAT	CTCTGTCTCG	1740
TGGTCTGCAA	CCTTGGAAGC	TTTGAACCTA	GCTCATAGAA	TCCTGGGCAT	CAAGAACATG	1800
TGGTTCTAAT	GGCTAGATAG	GGAATGAGAG	TAAAAGGATT	TTGCCCACGG	TCACGTGAGT	1860
AAACAACAGA	TTTGGAGGGG	TCTGGACTAC	TGTGATGACT	TCATTCTGAC	AATATGTTCC	1920
AGTTGTCCTT	TCATTTTCCTC	CTAATCACAT	GTCTGGTCTG	ATCTGGCTGT	TTCCACCTT	1980
CCAATTCCTG	CCTTCTCCAA	TGCTCCCTTC	CGTAGGTCAC	TCTGTGGCTC	AGAGACCCTG	2040
CTTAGCAAGC	GCCCAACCTT	TCAATTATTT	GTTCAGTAAA	ACTTGAACTC	ATGTCTCCCC	2100
TTCTTGATAA	AAAGAAAATA	CGTTATGTAA	TGTCGGGTTA	CTCTATAACT	CTTGTCCTGT	2160
CTCTCGGCAA	CTACTGAACT	AACTGTTTTT	ATATTGAGCA	AACGTTTATG	GAAGGACTGC	2220
CAAGAGTCAG	GTACTIONGCT	TGGTAATATT	CCCCGTTCTC	TCTAGTCAAA	GCCAACACCA	2280
GCCAGACTTG	CAGATCTAGG	TCCCAAGCCC	ACTGCAGATC	ACAGGCCAGG	GTCTGGTCTC	2340
CTCTGAGCTC	CTTTGGGAGG	GAAAGACAGA	ATTATTAACA	CCCATTTTGT	AGATTAGGCA	2400
ACTGAGGCTG	AGGAAGTTTA	AATAACTCAG	ACAGGGCCTG	CACGTCAGTC	ATATTCCAAG	2460
GATCCCTACT	CACTGTCTTC	TCTCTACAGA	ACGAGATGTC	TCTGGAGTCC	ATAGAAAGCC	2520
CAGGAGCCTG	GCTGGGCACG	GTGGCTCCTG	CCTGTAATCC	CAGCACTTTG	GGAGGCCGAG	2580
GCAGGCAGAT	CACCTGAGCT	CAGGAGTTCA	AGACCAGCCT	GGGCAACATG	GCAAAACCCC	2640
ATCTCTACTA	AAAATACAAA	AAATTAGCTG	GGCGTGGTGG	TGCATGCCTC	TAATCCCAGC	2700
TACTTGGGAG	GCTGAGGCAC	AAGAATTGCT	TGAGCCCAGG	AGGCAGCAGT	TGCAGTGAGC	2760
TGAGATTGTG	CCAGTGCACT	CCAGCCTGGG	CAACAGAGCA	AGATTCCATT	TCAAAAACAA	2820
AAACAAACAC	AAACAAACAA	ACAAAAATAG	AAAGCCCAGG	GACCACCTGC	GTCAGGTTCC	2880
CAGCCACACC	TTTTTCTTGT	CCTCCTCTGT	CTCTGGCATC	TTCTCACAGG	TTCCTAATTG	2940
TTTGTGGTTG	CACAAATTCA	AAATCCCAGA	AAAATTACCA	CTTCACACCC	ACTCAGATGG	3000
CTATTTTTTT	TTTGAAGGAA	GATAACAAGT	GTTGACAAGA	ACATGGAGAA	ATTGGAATTC	3060
TCACCCATTG	CTGGTGAGAA	TGTAATACGG	TGCTGCTGCT	ATGGAAAACA	GCTTGGAGTT	3120
TCCTCAAAAA	GTTCAACAGA	ATTTCAATGT	GACCCAGCAA	TTCCCCTCTA	AGTTATAGAT	3180
CTGAGAGGAT	TAAAAACAGT	TACTAAAATA	CACGGACTCA	CATATTTCTA	ACAGTCCAAT	3240
TCACAAGGGC	CAAAGGTGC	TAATAGCCCA	CATGTCCATC	GATGGATGGA	TAAATAAATT	3300
GTGGTCTATC	CATACAATGG	AATATTATTC	GGCCATAAAT	GGAATGAAGT	ACTGACGCAT	3360
GCTACAGAAT	GGATGAACCG	CAAAAAAAT	GGATGAACAC	ATGCTACAGA	ATGGATAGCC	3420

TCAC TTTACT	ATGAAGTGAA	GGCCAGAAAC	GAAGTCCATA	TATTGCATCA	TACAAAATAT	3480
CCAGAAGAGG	GAAGCCCACA	GAGACAGAAT	GTGCAATGGT	GGATGCCAGG	GTCTGGGGAG	3540
AGGGGAGAGT	GGGGAGAAAC	TGCTCAACTG	GTACAGGCTT	TATTTTGGAA	TGATGGGAAC	3600
ATTTTGCAAC	TAGATAGAGG	TAGTGATTGC	AGAACACAGA	ATGTACTGAA	TTCCACTGAT	3660
TTTTTTCACC	TTAAAATGGT	TAATTTTCAG	TCCTGAGATT	GGATAATCAT	AAAAAAATGG	3720
TTAATTTTAT	GTTATGTGAA	TTTCATCCCT	ATACATATTT	TAAACCTCAG	AAATATACAC	3780
TAGCAGGCAT	GGAACAGGTC	ACTGTGGTGC	CTGCCAAGCC	CGGTGATGTT	ATCTGGGGTC	3840
CCCGGCCAGC	CTTAAGCCTC	TTGCTGACCG	GTGGAGGGCA	GAACCTTTGC	CCTAAAAGTA	3900
TAATATCCAC	ATGCTGGCAT	GATTCCTGGC	CAGATGGCTT	CTTTATTAGC	AGTAATTGAA	3960
ACTGCCTCGA	TACAGACACT	GTACCTTGCA	ACCAAAAAAT	GACTCAACAA	TGATAATAAG	4020
GGTTAAGCTG	GGCCTTTCTC	TCTTTGCCAG	TTAAATTATA	TTTATTATAG	CTTGACATGA	4080
AAAACAAAGC	AACTCCAACA	GGTATCACAA	GGGCAAAGGA	CATGAACATT	TTATCAAAGA	4140
AGAAATGCAG	CTGTCAAAAA	TACAGAAATA	TTCAACCTTG	TTCATAATAA	AGTGGCTGGG	4200
CTCAGTGGTT	CATGCCTGTA	ATCCCAGTGC	TTTGCAAGGC	TGAGACAGGA	GGATCATTTG	4260
AAGCCAGAAG	TTCAAGACCA	TCCTAGGCAA	GTCAGTTCAA	TACCAGACTT	CATGTCTACA	4320
AAACATCAAA	AAATTAGCCA	GGCATGGTGA	TGCATGCCTG	TTGTCCCAGC	TACTCAGGAG	4380
GCTGAGGCAG	GAGAATTGCT	TGAGCCTGGG	AGGCTGCGGT	GGCGGTGAGC	CATGATTGTG	4440
CCATTGTACT	CCAGCCTGGG	CAATGCAGCA	AGACTGTCTA	AATAACAAAA	ATAATAGTAA	4500
AGAAAAGGAT	TGGGATGCCA	TTTACTTGCG	TATTCAATAC	ACAGAGTTAA	AAGTAATTTT	4560
TACGTTTTCT	ATTTTTTTAT	TACTAAAAAA	AGCTGGACCA	TTCTCACAGC	CTGAAATGCT	4620
TCTCACTTTC	CCTTCTTCTG	TCCAAACACT	TCTCTATGAT	AATGCAAACA	GTCACCTCCT	4680
TAGGAAGACT	TCACCCCAGG	TAGTTCCAGA	TCCCCTTATC	TCTGCCTTCC	CAGAACTCCT	4740
GGTGTCTCTC	CAGTTCCCTC	CGTGTGGTGA	AGTACCCTAC	CTAGGGTTTC	AGTATGGCTC	4800
TGTCTGCAAA	GGTCTTGTTT	ACACCTTCCC	TTATGGTTCT	GTTGCCCTGT	GTTGTGTCAT	4860
AGCACAGGGC	ACAGTGGAGA	ACCCATTCACT	ACTGATAGAG	AGGGCCCCAT	GGTCCTGGAG	4920
ATAACCATGT	AACCGATCAG	AATAAGGCAT	TGAGGGCTGG	GTGTCAGGCG	TGGGCTGCAC	4980
TTGGGTGGGC	AGGTCCCCTG	GAAAGTCACT	GGGTTTGGCA	AGCTTCCTAG	TAACATGTCT	5040
CTCTGGGGTC	CCCCTTGGA	CTTCATGCAA	AAATGCTGGT	TGCTGGTTTA	TTCTAGAGAG	5100

ATGGTTCATT	CCTTTCATTT	GATTATCAAA	GAAACTCATG	TCCCAATTAA	AGGTCATAAA	5160
GCCCAGTTTG	TAAACTGAGA	TGATCTCAGC	TGAATGAACT	TGCTGACCCT	CTGCTTTCCT	5220
CCAGCCTCTC	GGTGCCCTTG	AAATCATGTC	GGTTCAAGCA	GCCTCATGAG	GCATTACAAA	5280
GTTTAATTAT	TTCAGTGATT	ATTAAACCTT	GTCCTGTGTT	GACCCCAGGT	GAATCACAAG	5340
CTGAACTTCT	GACAAGAACA	AGCTATCATA	TTCTTTTCAA	TTACAGAAAA	AAGTAAGTTA	5400
ATTGATAGGA	TTTTTTTTGT	TTAAAAAAA	TGTTACTAGT	TTTGAAAAGG	TAATATGTGC	5460
ACATGGTAAA	CACTAAGAAG	GTATAAGAGC	ATAATGCTTT	TATACTACTA	AGAATAATGT	5520
TTTCTCTAAG	TTTTTTTTTG	TAGATGCTTT	CATCAGATTA	AGAAAATTCC	CTGCTATTAG	5580
TTGTTGAAGG	TTTTTATATC	ATAAATGAAA	GTTGAATATT	ATTATCATAT	ATTATTAATA	5640
TATTGTTATT	GAACTATCAA	AGCCTTTTCC	TAAAACCATT	GAGATGATCT	TATAACCATT	5700
CTCCTTTAAC	CTGTTGACGA	GATCATTGGT	ATTTATACTA	TTTCTCTGTT	AACCATTCTT	5760
GAGTCTCAGG	TTTAAATTCA	ACTTGGTCAT	GGTGTGTCAT	CTTTGATCAT	TGCTGTCTGT	5820
GGCTTGCTAC	TGTTTTGTTT	AGGATTTTTG	CACTGATGCT	CATCAATGAG	ACTGGCATGC	5880
CATCTTCCTT	TGCAGTCCTG	ATTTTTTTCT	GATTTGGATC	ATGTGGTTAT	GGCCCTCATG	5940
GAATGAGTTG	GGCATGATGC	CTTTTTTTCA	TGTCTCTGGA	TTGATGGGAC	ACTTTGGATT	6000
CTCTCCAGAT	GGCCCTCAAT	GGTCCCTGCC	TCCTCATTGT	TAGGCCCTG	GGCAAGCCCT	6060
TCTCATTTCT	GGTAGGCCCA	GGAACCTGTG	GGGGTTTTGT	TTGTTTGTTT	GTTTCTTGAG	6120
TCGGAGTCTC	ACTCTGTCAC	CCAGGCTGGA	GTTGGAGTGC	AATGGCCCGA	TCTTGGCTCA	6180
CTGCAACCTC	CACCTCCCAG	ATTCAAGCAA	TTCTCCTGCC	TCAGCCTCCT	GAGTAGCTGG	6240
AATTACAGGC	ACCCACCGAC	ACACCCTGCT	AATTTTTGTA	TTTTTAGTAC	AGATGGGGTT	6300
TCACAATATT	GGCCAAGCTG	GTCTCGAACT	CCTGATCTCA	TGATCTGCCC	GGCTTGGCCT	6360
CCCAAAGTGT	TGAGATTACA	AGCATGAGCC	ACCACACCCA	GTGAACCTGT	GGTTTTTAGA	6420
AGCTCCCCAT	GCATGTGAAT	GCTGTGAGCA	TCCCAGGATG	ACAGCCACTG	TGTGTTCAGC	6480
TGTTGGAACT	GTGAGAAAGC	ACCAGTGGGA	CCTTCTCCAG	CACCTGCCTG	CTGAGTTCAT	6540
GGAAGAGGCT	TGTTGGGGAG	ATGATGCCCT	GGCTGACTCC	TGAAGGATGG	TTAGGAATGC	6600
ACCAGATGGA	AGCTGGGTTG	GACCCACTCT	ATGCTGAAGA	ACAGCTTGTG	TGGACACAAG	6660
GAGACACGGA	TATGTCATTT	TTGTAGAGCC	TGAGGAGTGT	CCAATCACAC	CATTTGCTTA	6720
AAACATCATG	CACACTTGGA	AAAGTGGACT	GAGACCGAAT	GAAGAAGCTA	ACAGTGGCCA	6780
GATCAGAAAG	GGTCTTGTGT	TACTTCCTAG	AGATACTTAG	ATTTTATCCT	GTGGGTGATA	6840

GGAGCAGTTG	GAGGGACTGA	AGACAAGGAA	AGAAACATGT	TTCAAGATCT	ATGTTTTTCA	6900
AGACGCTTTT	CTGGTGGCTG	AGTAGGGAAT	TCCCTGGATA	AGTCCTGCCC	AGGGTCAGGC	6960
AAAACAAGTT	AGGGGGTTAC	TGAAATAAGG	AGTATGAGAA	ATGGTGTAGG	TTGTGCTGAC	7020
GTTTTGTAAC	ACATCTCATG	ATGATCTTCA	TTTCCTTCAC	TAATTTCTG	TTTCATTAAT	7080
TCCCTTCCAC	GTGCTCTTCT	GAAATTTGCC	TCACATTCTC	TGATTTCTCT	TTTACCTGTT	7140
GGTTTCATCA	CCTTTTACTT	TTTGCTTTCC	TGGAAACACA	AATGATTCTG	ATTGTGACAT	7200
GTCAGAATTA	TTTGCAACAT	TTGCCTTTCT	GCTGAAACCA	TGAGTTCACT	GAATACACAA	7260
TTTAGTAAAG	TGTAGGATGC	ACATGTCGTT	TTCGTGGTCA	CAACCAGCTC	TGTAGCATTT	7320
TATAACTACA	CTGGCAGTGT	GCTGGGAGGT	GTAGAGAGAA	ATATTTATCA	CATGTGTGGC	7380
TGACACAACC	TGCCAAGTTA	TTTTAGGAGC	CTCCTTGGA	TCCCAGCAAG	AATGCTACCG	7440
GCACAATTTG	TAATCACAGC	ATCCTGCTCC	ATGCCTTGGC	TTCATGGCAT	AGTCACTTCT	7500
GCAAGTCTCT	TTCCAGCTGT	CTGTTCCCAT	GTCTATAAAG	TATGAGTTAA	ATCATCCTAA	7560
CACTACTCAT	CTTACAAAGT	TTTCTTGCTG	ATGTTAAGAG	AGTTGGGAAA	GAAGTGTATA	7620
AACTGTGAAG	TGCCATGGAG	ATGTTAGTGG	TTACTTTATC	AAGAAATAGA	CACTCTAGAA	7680
TGGAGTAGAA	AGCCAACAGT	TATGATTGAG	TCCTCCTCCT	CTTCTTCTTT	TTATTAATTT	7740
ATAAAGAAAA	GAGGTTTAAT	TGACTCACAG	TTCCATATGG	CTGGGGAGGC	CTCGGGAAAC	7800
TCTCAGTCAT	AGCAGGAGGC	AAAGGGGAAG	AAGGCACCTT	CTTCACAAGG	CGGCAGGAGA	7860
GAGAGAGCTC	CTGTTCTTTT	TTGTCATAAA	GTCTACAGAA	GTGCTTATAC	TTCAGGACAA	7920
GGGCAGGCAG	AGAGAAGGAA	GGACATTGCT	TCACCCCAGC	CCTCACTGAC	GAGTTTGCTA	7980
GGGGACCTCA	CTTTGTCCCA	GAGTAGGGCA	GAAGTCTGGC	CACTACCCAT	TCAGAAGGCC	8040
TGGGCTGCAC	TGCTAGTTCC	TCACTAACTC	TGTGTGGCCT	TGGGCAAGGT	TGGGCCTGTG	8100
TTAACAGATT	ATGACCCTGG	GCTCTCAAGC	TAGAGGATCT	AAATTTGAAT	CCTGGCTCTG	8160
CTAAAGCAAT	TAGTGATGTA	AACTTTAATG	GGTCAGTTAA	CCTTCCTGTG	GCTTAGTTTG	8220
CTCATCTGTA	AAATAGGGAT	CATAACAGTA	TCAATACCAC	ATGATTGTTG	GACAGATTGA	8280
ATCAGTTAAT	GCAGGGGAAG	TACTTAGCAT	GACACGTATT	CACTATCATT	TCCTGGAGTA	8340
AGAGCTGTGT	GTGAGTGGGT	GTGAGCATGT	GTGAAACCTT	TTCTCTGCAA	TCTCAGTTAA	8400
GAAACCAATC	CAGAATTTAA	AGTTCAGGGC	CTAAATGGGT	GGTTATCTTC	TCCCAGTTCC	8460
ATCCTATCCC	ACCTTTGCTC	TTCTCCCGC	CCACAGGAGC	TGTTGGTCCT	TGATTGGGCT	8520

GGAAGACCTG GTGGACCCTA AGTGATCTAT AAGAGGAGAA TAGAGAACAG GGAATGTCTT	8580
CAAAAATCTA GAGGGACACA GAGGCTGAGA GGCAGGCAGT CCTGCAGGGT CTTCTGATTG	8640
GGACAAGGAG AACCTTGGTC TTCACAGGCC AATTCTGGTC AGTTTCCCCC ATGGACAGAT	8700
GAGGAAACAG GCCCAGGAAT ATCCAAGGTC TCACACTTCC CATCTGTCAA GTCTTGTTGA	8760
TTCTGTTGTA TTCATGTCTC TCAAAGGGAG ATAGAGTTTA GGAAGAAAG AAGGATCAAC	8820
TGTGTCTGAT ACCACTGGGA GCTTAAGTAA AGGGTTCTTT TACTTCATAG CATTTATCCC	8880
AATTTGTAAT TCAGTATTAT TTGTGTGGCT GTTTGGTGTC TCTTTCTCCT ATATGAGTGC	8940
TAGCTTCATA AGGGCAAGGA TTTTGATTCT TTAATATTTA GTGCTTGCCA CATGCCCTGA	9000
ACACAGCAGG CATAAGGCT AACCAACATA CAGTGGCATG AAAGTCATGA AAGTGAGACA	9060
CCTACCTCCT CCAGTGCCAA GAGAGCATAA CCATGCACCT GTCACTCTCC TCAACACCAC	9120
CCCCAAGCAT GAGGCCCAA AGCATTAGCT AATCCCCTCC TCCAGCCACT AAAACTTAAA	9180
GGCCAGGTGT GGTGGCTCCC ATCTGAAATC CCAGAACTTC AGGAGACAGC AGCAGGAGGA	9240
TCACTTGAGG CCAGGAGTTT GAGATCAGCC TGGGCAACAT AGCTAGGTCC CATCTGTACT	9300
AAAAATTAGC TGGGCGTTGT TGCATGCCTG TAGTCCCAGC TACTAAGGAG GCTGAGGTGG	9360
GAGGATCACT TGAGCCCAGG AGGTGGAAAC AACAGTAAGC TATAATCACA GCACTGAACT	9420
CTAGCCTGGG CAACAGAGTG ACACCCTGCC TCAAAACAAT TTAAAAATA AATAAGAGCA	9480
AAACTTAGAT ACCACGTGGT CACCCCAACA TGCAAAATCA AGTTTCCCCC TACTGAGAAG	9540
AATGGGGACT TGACAGCTGA GTTACAGAGA GATAATCTTC TTCTTCTTTT TTTTTTTTGT	9600
GTTTACATCC TCAAGATCAT GACTTGTGAA ATTTGAATCG AATACACATG TAATTCCAGA	9660
GCAATGTTGC CTCCGCATAC CATCAGCAAT TCACTTGGCT ACTGGAAGTC AGGAT	9715

INFORMATION FOR SEQ ID NO.: 3

SEQUENCE CHARACTERISTICS

LENGTH: 632

TYPE: nucleotide

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ORIGIN: homo sapiens

ANTI-SENSE:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: Xenobiotic Responsive Enhancer Module of CYP3A4 gene; -7836 to -7207 base pairs relative to transcription initiation site of CYP3A4 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 3

```

TCTAGAGAGA TGGTTCATTC CTTTCATTTG ATTATCAAAG AAACATCATGT CCCAATTAAA      60
GGTCATAAAG CCCAGTTTGT AAACATGAGAT GATCTCAGCT GAATGAACTT GCTGACCCTC      120
TGCTTTCCTC CAGCCTCTCG GTGCCCTTGA AATCATGTCTG GTTCAAGCAG CCTCATGAGG      180
CATTACAAAG TTTAATTATT TCAGTGATTA TTAAACCTTG TCCTGTGTTG ACCCCAGGTG      240
AATCACAAGC TGAACCTTCTG ACAAGAACAA GCTATCATAT TCTTTTCAAT TACAGAAAAA      300
AGTAAGTTAA TTGATAGGAT TTTTTTTGTT TAAAAAAAAT GTTACTAGTT TTTGAAAAGG      360
TAATATGTTG CACATGGTAA AACTAAGAA GGTATAAGAG CATAATGCTT TTATACTACT      420
AAGAATAATG TTTTCTCTAA GTTTTTTTTG GTAGATGCTT TCATCAGATT AAGAAAATTC      480
CCTGCTATTA GTTGTTGAAG GTTTTTATAT CATAAATGAA AGTTGAATAT TATTATCATA      540
TATTATTAAT ATATTGTTAT TGAACATCA AAGCCTTTTC CTAAAACCAT TGAGATGATC      600
TTATAACCAT TCTCCTTTAA CCTGTTGACG AG                                     632

```

INFORMATION FOR SEQ ID NO.: 4

SEQUENCE CHARACTERISTICS

LENGTH: 11186

TYPE: nucleotide

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ORIGIN: homo sapiens

ANTI-SENSE:

LOCATION:

IDENTIFICATION METHOD: 5' flanking region of CYP3A7; -11,133 to +52 base pairs relative to transcription site of CYP3A7 gene

OTHER INFORMATION:

SEQUENCE DESCRIPTION: SEQ ID NO: 4

```

GGATCCAGTT TCAGCTTTCT ACATATGGCT AGCCAGTTTT CCCAGCACCA TTTATTAAAT      60
AGGGAATCCT TTCCCCATTG CTTGTTTTTG TCAGGTTTGT CAAAGATCAG ATGGTTGTAG      120
ATGTGTGGTG TTTGTTCTGA GGCCTCTGTT CTGTTCCATT GGTCCATATC CCTGTTTTTG      180
TACTAGTACC ATGCTCTTTT GGTACTGTA GCCTTGTAGT ATAGTTTGAA GTCAGGTAGC      240
GTGATTCCTC CAGCTTTGCT CTTTTTGCTT AGGATTGTCT TGGGAATGTG GGCTCTTTTT      300
TGGTTCCATA TGAAATTTAA AGTAGTTTTT TTTCCAATTC TATGAAGAAA GTCATTGGTA      360
ACTTGATGGG GATGGCATTG AATCTATAAA TTACCTTGGG AAGTATGGCC ATTTTCACGA      420
TATTGATTCT TCCTATCCAT GAGCATGGAA CATTCTTCCA TTTGTTTGTG TCCTCTTTGA      480

```

TTTTGTTGAG CAGTGGTTTG TAGTTCCTCCT TGAAGAAGTC CTTACCTCC CTTTAATTTG	540
GATTACTAGA TATTTTATTC TCTTAGTAAC AATTGCAAAT GGGAGTTCAC TCATGATTTG	600
GCTCTCTTTC TGTTATTGGT GTATAGGAAT GCTTGTGATT TTTGCGCATT AATTTTGTAT	660
CCTGAGACTT TGCTGAAGTT GCTTATCAGC TTAAAAGGAT TTTGGGCTGA GACGATGGGG	720
TTTTCTAAAT ATACAATCAT GGCATCTGCA AACAGGAACA ATTTGACTTC CTCTTTTCCT	780
AATTGAATAC CCTTTATTTT TTTTCTTGC CTGATTGCCC TGGCCAGAAC TTCCAATACT	840
ATGTTGAATA AGAGTCATGA GTGAGGGCAT CGTTGTCTTG TGCTGGTTTC AAAGTTTTTG	900
CCCATTCAGT ATGATTTTGG CTGTGGTTTT GCCATAAATA GCTCTTATTA TTTTGAGATA	960
CGTTCCACCA ATACCTACTT TATTGAGAGT TTTTAGCAGG AAGGGCTGTT GAATTTTGTC	1020
GAAGGCCTTT TCTACATCTA TTGAGACAAT TATGTGGTTT TTTAATCGTT GATTCTGTTT	1080
ATGTGATGGA TTACATTTAT TAATTTGCAT ATGTTGAACC AGCCTTGCAT CCCAGGGATG	1140
AAGCCCACTT GATTGTAGTG GATAAGCTTT TTGATGTGCT GCTGGATTCA GTTTGCCAGT	1200
ATTTTATTGA GGATTTTGGC ATCAATGTTC ATCAGGGATA TTGGTCTAAA ATTCTCTTTT	1260
TTTGTGTGT CTCTGCCAGG CTTTGGTATC AGGATGATGC AGGCCTCAGA AACTGAGTTA	1320
GGGAGGATTC CCTCATTTTC TATTGATTGG AATAGTTTCA GAAAGAATGG TACCAGCTAC	1380
TCTTTGTACC TCTGGTAGAA TTCAGCTGTG AATCCATCTG GTCCTGGACT TTTTGGTTGG	1440
TAGGCTATTA ATTATTGCCT CAATTTTAGG GCCTGTTATT GGTCTATTCA GACATTCAAC	1500
TTCTTCCCGG TTTGGTCTTG GGAGGGTTTA TGTGTCCAGG AATTTATCCA TTTCTTCTAG	1560
ATTTTCTAGT TTATTTGTGT AGAGGTGTTT ATAGTATTGT CTGATGGTAG TTTGTATTTT	1620
TGTGAGATCG GTGGTGATAT CCCCTTTATC ATTTTTTATT GCATCTATTT AATTCTTCTC	1680
TCTTTTCTTC TTTATTATTC TGGCTGGCGG TCTGTCAATT TTTTGTATCT TTTCAAAAAA	1740
CCAGCTCCTG GGTTTCACTG ATTATTTGAA GGGTTTTTTG TGTCTCTATT TCTTTCAGTT	1800
CTCCTGTGAT CTTAGTTATT TCTTGCCTTC TGCTAGCTTT TGAATGTGTT TGCTCTTCCT	1860
TCTCTAGTTC TTTGAATTGT GATGTTACAG TGTTGATTTT AGATCTTTCC TGCTTTCTCT	1920
TGTGGTCATT TAGTGCTATA AATTTCCCTC TACACATTGG TTTACATGTG TCTCAGAGAT	1980
TCTGGTATGT TGTGTCTTTG TTCTCATTCA TTTCAAGAAC ATCTTTACTT CTGCCTTCAT	2040
TTTGTTATTT GCCCAGTAGT CATTCAGGAG CAGGTTGTTC AGTCTTCATG TAGTTGTGTG	2100
GTTTTGAGTG AGTTTCTTAA TCCTGAGTTC TAATTTGATT GCACTGTTGT CTGAGAGACA	2160

GTTTGTTGTG	ATTTCCATTC	TTTTACATTT	ACTGAGCATG	CTTTATGTCC	CATTATGTGG	2220
TCAATTTTAG	AATAAGTGTG	ATGTGATGCT	GAGAAGAATG	TATATTCTGT	TGATTTGGGG	2280
TGTGGAGTTC	TGTAGATGTC	TATTCAGTCC	ACTGGGTGCA	GAGCTGAGTG	GACATGAACA	2340
TTTTATCAAA	GAAGAAACAC	AGCTATCAAA	AATCCAGAAA	TATTGAACCT	TGTTAATAAT	2400
AAAGTGGCTG	GCCTCTGGTT	CATTCCTGTA	ATCTCAGTCC	TTTGAAAGGC	TGAGAAAGGA	2460
GGATCACTTG	AGGCCACAAG	TTCAAGACCA	TCCTAGACAA	GTCAGTTCAA	GACCAGACTT	2520
CATGTCTACA	AAACATCAAA	AAATTAGCCA	GGCATGGTGA	TGCATGCCTG	TCATCCCAGC	2580
TACTCAGGAG	GCTGAGGCAG	GAGGATTGCT	TGAGCCTGGG	AGATTGAAGT	GGCAGTGAGC	2640
CATGATTGTG	CCATTGCACT	CCAGCCTGGG	CAATGCATCA	AGACTCTGTC	TAAACAATAA	2700
TAATAATAAT	AGTAATAGTA	ATAATAATAA	TAATAAAGAA	AACGGTTGGG	ACGCCATTCC	2760
TTACTTATTC	AATACACAAA	GTTAAAAGCA	ATTTCTACTT	TCTCTATTTT	TTTATTACTA	2820
AAAAAAGCTG	AACCATTCTC	ACAGCCTGAA	ATGCTTCTCA	CCTTCCCCTC	TTCTATACAA	2880
ACACTTCTCT	GTTGATGATA	ATGCAGACAG	TCTCTCCTTT	AGGAATACTT	CACACCAGGT	2940
AGTTCCAGAT	CCCCTTATCT	CTGCCTTCCC	AGAGCTCCTG	GTGTCTCCCC	AGTTCCCTCT	3000
GTGTGGTGAA	GTACCCCCAC	CTTGGGTCTC	AGCATGACTC	GTTCTTTGAA	GGTCTTG TTC	3060
ACATTTTCCC	TTATGGTTCT	GTTCCCCTGT	GTTGTGTCAC	AGCACTGGGC	AGAGTGGACA	3120
ACCCATTAC	ACCGATAGAG	AGGGCCCCAT	GGTTCCTGGAG	ATAACCATGT	AACTGATCAG	3180
AATAGGGCAT	TGAGGGCTGG	GTGTCAGGCA	TGGGCTGCAC	TTGGGTGGGC	AGGCCCCCTG	3240
GAAAGTCACA	GGATTTGGCA	AGCTTCCTAG	TAACATCTCT	CCCTGGGGTC	CTCTTGGAAC	3300
TTCATGCCCG	ATGCTGGATG	CTGGTTTATT	CTCGAGAGAT	GGTTCATTCC	AATAATCAAT	3360
GAAACTCATG	TCCCAACTAA	AGTTCATAAA	CTCCAGTTTG	TAAACTGAGA	TAATCTCAGC	3420
TGAATGAACT	TGCTGACCCT	CTGCTTTCCC	CCAGCCTCTC	AGTGCCCTTG	AAATCATGTC	3480
AGTTCAAGCA	GCCCCATGAG	GCATTACAAT	GTTTAGTTAT	TTCAGTGTTT	ATTAAACCTT	3540
GCCCTATGCT	GACCCCAGGT	GAATCACAAG	CTGGACTTCT	GACAAGGACA	AGCTATGATA	3600
TTCTTTTCAA	TTACAGAAAA	AGTAAGTTAA	CTGATAGGAT	TTTTTAAAGA	TGTTACTAGT	3660
TTTGGAAGG	TAATTTGTGC	ACATGGTAAA	CAAGAAGGTA	TAAGAGGATA	ATGCTTTTAT	3720
ACTGCTGAGA	ATAATGTTTT	CTCTAATTTT	TTTTGGTAGA	TGCTTTCATC	AGATTAATAA	3780
AATTCACTGC	TGTTAGTTGT	TGAAGGTTTT	TTATATCATG	AATGGGAGTT	GAATATTATC	3840
ATGTATTATT	AATATATTAT	TATTGAACTA	GCAAAGGCTC	TTCCTAAAAC	AATTGAGATG	3900

ATCTTATAAT	CGTTCTCCTT	TAATCTGTTG	ATGAGATCAT	TGGTATTTAT	ACTTTTCTC	3960
TGTTAACTAT	TCTTGAGTCT	CAGGTTTAAA	TTCAACTTGG	TCATGGTGTA	TCATCTTTGA	4020
ACACTCCTGT	CTCTGGCTTG	CTACTATTGT	G TTCAGCATT	TTTGCACTGA	TGCCGATGAA	4080
TGAGACTGGC	ATGTCATCTT	CCTTTGCGGT	CCTGATTTTT	TTCAGATTTG	GATCATGTGG	4140
CCCTCATTGA	ATGAGTTGGG	TGTGATGCCT	TCTTTTTTCAT	GTATCTGGAT	TGATGGGACA	4200
CTTTGGAGTC	TCTCCAGATG	GCCCTCAATG	GTCCCTGCCT	CCTCATTGTT	AGGCTCCTAG	4260
GCAACCCTTT	CTCATTTCTG	G TAGGCCAG	GAACCTGTGG	GTTTTATGTT	TGTTTGTTTG	4320
TTTGTTTGTT	TGTTTTTTGA	GTTGGAGTCC	TGCTTTGTCT	CCCAGGCTGG	GGTTGGAGTG	4380
CAATGGCCTG	ATCTCGGCCC	ACTGCAACCT	CCACCTCCTG	GGTTCAAGTG	ATTCTCCTGC	4440
CTCAGCCTTC	TGTGTAGCTG	GGATTACAGG	CATCCACCAC	CACTCCTGGC	TAATTTTTGT	4500
ATTTTTAGTA	GAGACGGGGT	TTTACAATAT	AGGCCATTGT	GATCTCTTGG	ACAGGCTAGT	4560
CTCAAATTCC	TGACCTCATG	ATCTGCCTGC	CTCAGCCTCC	CAAAGTGCTG	AGATTACAGT	4620
TTTGTGCCTC	CACACACAGT	GAATCTGTGG	TTTTTAAAAG	CTCCTCATGC	ATGTGAATTC	4680
TGTGAGCATC	CCGGGATGAC	AGCCACTGTG	TGTCCAGCTG	TTAAAACTGT	GAGAAAGCAC	4740
CAGCGGGACC	CTCTCCAGCA	TTTGCTTGCT	GTGGTCATGA	AAGAGGCTTG	TTGGGGAGAT	4800
GATGCCCTGG	TTGACTCCTG	AAGGATGGTT	AGGAATGCAC	CAGATGGAAG	CTGGGTGGA	4860
CCCAGTCTAT	GCTAAAGAAC	AGCTTGTGTG	GACACAAGGA	GACACGAACA	CATCATTTTT	4920
GCAGAGCCTG	GGGAGTAGCC	AATCGCACCA	TTTGCTTAAA	ACACCGTGTA	CAGTTGGAGA	4980
AGTGGACTGA	GACAGGCTGA	AGAAGCTAAC	AGTGGCCAGA	TGAGAAAGGG	TCTTGTGTTA	5040
CTTCCTAGAT	ATACTTAGAT	TTTATCCTGT	GAGTGATAGG	AACAGTTGCA	GGGACTGAAG	5100
CCAAGGAAGC	ATGCTTTAAG	ATTCCATGTT	TTTTGAGATG	CTGTCTGGTG	GCTGAGTAGG	5160
GAATTCCCTG	GATAAGTACT	GCCCAGGGTA	GGCAAAAGAA	GCTAGGAGGT	TACTGAAATA	5220
AGGAGTATGA	GAAATGGTGT	AGGTTTTGCT	GATGTTTTGT	AACACATCTC	ATGACAATCT	5280
TCATTTCCCTT	CACCAATTTT	CTGTTTCATT	AATTCCCTTC	CACGTGCTCT	TCTGAAATTT	5340
GCCTCATATT	CTTTGATTTT	TCTTTTACAT	GTTGGTTTCA	TCACCTTTTA	CTTTTGTGCTT	5400
TCCTGGAAAC	ACAAATGATT	CTGATTGTGA	CATGTCAGAA	TTATTTGCAA	CATTCCCCTT	5460
TCTGCTGAAA	CATGAGCTCA	CTGAATACAC	AATTTAGTAA	AGTGTAGGAT	GCACATGTTG	5520
TTTTCATGGT	CATAACCAGC	TCTGTAGCAT	TTTATAACTA	CACTGGCAGT	GTGCTGGGAG	5580

GTGTAGAGAG	AAATATTTAT	CTCATGTGTG	GCTGACACAA	CCTGCCAAGT	TGTTTTAGGA	5640
GCCTTCTTGG	AATCCCAGCA	AGAACACCAC	TGATGCAATT	TGAAATCACA	ATGTCCTGCT	5700
CCATGCCCTG	GCTTCATGGC	TTAGTCACGT	CTGAAGTCTA	TTTCTAACTA	TCTGTTTCCA	5760
CATCTATAAA	GTATGAGTTA	AATCATCCTA	ATACTACTCA	TCTTACAAAG	TTTTCTTGCT	5820
GATATTAGGA	GAGTTGGGAA	AGAACTGTAT	AAATTATGAA	GTGCCATGGA	GATGTTGGTG	5880
GTTACTTTAT	CAAGAAATAG	ACACTCCAGA	ATAGAGTAGA	AAGAAAACAG	TTATGATTAA	5940
GTCCTCCTCC	TCTTCTTTTT	TTTAAATTTA	CAAAGAAAGG	TTTAATTGAG	TCACAGTTCC	6000
ATATGGTTGG	GGAGGCTCAG	AAAACCTGCA	ATCATGGCAG	TTGGCAAAGT	GGAAGAAGGC	6060
ACCTTCTTCA	CAAGGTGGCA	GGAGAGAGAG	AGCTCCTCTT	CTTTTTTGTT	GTAAAGTCTA	6120
CAGAAGTGCA	TATACTTCAG	GGCAAGGGCA	GGCAGGGAGA	AGAAAGGACA	TTGCTTCACC	6180
CCAGTCCTCA	CTGACAAGTT	TGCTTTGGGA	CTTCATTTTG	TCCCAGCATA	TGGGACAGAG	6240
CTCTGGCCAC	TACCCATTCA	GAAGGCCTGA	GCTGCATTGC	TAGTTCCCCA	CTAACTCTGT	6300
GTGTCCTTGG	GCAAGGCTGG	GCTTATGTCA	AAAGATTATG	ACCCTGGGCT	CTCCAGCTAC	6360
AGAATCTACA	TATGAATCCT	GGCTCTGCTA	GAGCAATTAG	TGACGTAACC	TTGGATGGGT	6420
CAGTTAACCT	TCCTGTGGCT	TAGTTTGCTC	ATCTGTAAAA	TAGGGATCAT	AACAACATCA	6480
ATACCATGGG	TTGTTAGACA	GATTGAATCA	GTTAATGCAG	GGTAAATACT	TAGCATGACA	6540
CGTATTCACT	ATCATTTCTT	TGAGTAAAAG	CTGAGTGTGA	GTGGGTGTGA	GAATGTGTGA	6600
AACCCTTTCA	CTGCAATCTC	AGTTAAGAAA	CCCATCCATA	ATTTAAAGTT	CAGGGCCTAA	6660
ATGGGTGGTT	ATCTTCTCCC	AGTTGCATCC	TATCCCACCT	TTGCTCTTCT	CCTGCCCCGT	6720
GGAGCTGTTG	GTCTTTGATT	GGGCTGGAAG	ACCTGGTGGA	CCCTAAGTGA	TCTATAAGAG	6780
AATGAGAATA	GAGGACAGGG	AATGTCTTCA	AAACTCCTAG	AGGGACACAG	AGGCTGAGAG	6840
GCAGGCAGTC	CTGCAGGGGT	CTTCTGATTG	GGACAAGGAG	GACCTTGGTC	TTCATAGGCC	6900
AATTCTGGTC	AATTTCCCCC	ATGGACAGAT	GAGGAAACAG	ATCCAGGAAT	ATCCAAGGTC	6960
TCACACTTCC	CATCTGTCAA	GTCTTGTTGA	TTCTGTTGTA	TTCATGTCTT	TCAAAGAGAG	7020
AGAGAGTTTA	AGGAAAGAAA	GAAGGATCAA	CTGTGTCTGA	TATCACTGGG	AGCTTAAGTA	7080
AAGGGTTCTT	TTACTTCATA	GCATTTTTTC	CAATTTGTAA	TTCAGTATTA	TTTTTGTCAC	7140
TGTTTAGTAT	CTCTTTGTCC	TATTAGAGAG	ATAGCTTCAT	CAGGACAAGG	ATTTTGATTC	7200
TTTAATATTT	AGTGCTTGCC	ACATGCCCTG	AACACAGCAG	GCATACAGAC	TAACCAACAT	7260
ACAGTGGCAT	CGAAGTGAGA	CACCTACCTC	CTCCAGTGCC	TAGAGTACAT	GTCCATGGAC	7320

CTGTCACTCT	CCTCAACACC	ACCCCTAAGC	ATGAGGCCCG	AAAGCATTGC	TAATCCCCTC	7380
CTCCAGCCAC	CAAAACTTAA	AGGCCAGGTG	TGGTGGCTCC	TATCTGAAAT	CTCAGAACTT	7440
TAGGAGACAG	CAGCAGGAGG	ATCACTTGAG	GCCAGGAATT	TGAGACGAGC	CTGGGCAACA	7500
TAGCTAGACA	CCATCTGTAC	TAAAAATTAG	CTGGGCATGG	TGGTATACCT	GTAGTACCAG	7560
CTACTAAGGA	GGCTGAGGTA	GGAGGATCAC	TTGAACCCAG	GAGGTGGAAG	CTACAGTGAG	7620
CTATAACCAC	AGCACTGAAC	TCCAGCCTGA	GCAACAGAGT	GAGACCCTGC	CTCAAAACAA	7680
TTTCAAAAAT	AAATAAATAA	AAACAAAAC	TAGATACCAC	GTGGTCACCC	CAACATGCAA	7740
AATCAAGTTT	TCCCCTACTG	AGAAGAATGG	GGACTTGAGA	GCTGAGTTAC	AGAGAGATAA	7800
TCTGCCTTTT	TTTTTTTTTT	TTTGGTTTAC	ATCCTCAAGA	TCATGACCTG	TGAAATTTGA	7860
ATCTAATACA	CAAATCATTC	CAGAGCAATG	TTGCTTCTGC	CTACCACGAG	TAATTCACTT	7920
GGCCACTGGA	AGTCAGAACA	AGCTTCCCAG	AAGAGAGGTA	CCACTTGGAC	TACCAATATA	7980
AAAGGATGAA	AATATCGGAG	TGAAGGTGTT	CCTTGCATCA	CTGAGTCCCT	GGACAGCCTG	8040
TCCACTCATG	CTGATATCTG	AGCCTAATGC	TTCTCTGAAT	GTTGAGATTT	AACTTTGATC	8100
CAATGAAACC	AGACCAAGAA	AGAAGAAACG	TCTTTCATTG	TTGATAAGGA	CATGATTTTT	8160
CTCACAATTT	TATGATTATT	TTTCCTTAGC	TGTCCTATAA	TTATCTGCTT	ATTTGTCTCT	8220
TCTCCATGTG	CTTAGGGTAC	AAAGTTGACC	AAGACCAAGA	ATAATGTCTG	GGAGCACAAT	8280
ACTGACAGCA	CAGCTTTAAA	AACATGATGA	ATGCTTTAAT	ACAGGAAATG	AGTAGGGGAG	8340
AGGCAAGTGG	TGCTTGGGTG	TTCTTCCAAT	GCATAGTATC	TTCTTGACA	CAGTCAGTGC	8400
AGCTCTCAGT	AGGCAAGTCC	CTACATGTTA	GAAGATGTTA	CTTTCTGTGG	AATTAGGTGG	8460
CAGAACATGC	CTTCAATTAT	TTTCCTTTGC	AGAACAACAC	CAATTTCATT	AGTTAGGACA	8520
GAGTGCTGGC	TGCATTTGAA	TTCCAAGCAA	CGATTAGTCT	ATCACTGTTG	GTATAGATTC	8580
CAACCAGTCA	CACCACCTCC	TGAAGTTTGT	TGGGCAGGTA	AATCTTCATC	TTAGAATAAA	8640
AATCATCTTA	GCCAAGTAAG	TGTTTTAGAG	GAAAGAAGAA	AACATAATCG	TTCCATAAG	8700
AGTTTTGTTT	CTAAAAAAT	AAGAAAGGCT	CTTTGTTTAG	GTGAGCTAAT	GAAGTTGTTG	8760
ATAGTTATCA	GATGACACTG	GAATCTTTAC	TTGCCAGAAT	GTGTTCTGTG	CACCTCTCGG	8820
TGTGGCAACA	TAGAGAGGGA	GATCCTCCAG	CAATGCCATT	GATATGGTCA	GAAACTGCAT	8880
CTTTCTTTCT	CCCTGCTGAG	ATGGGGTCCT	TTGTTCTAGA	AAACCCAGGG	GGTGCCACTG	8940
GGAGTAACCC	TTGAGACAGG	AACACGAATC	TCAACCAATT	TCTGGTTGCA	GCCTTGAGTC	9000

TTACTATTTG	CCATAGTGAT	GCTTAGCAAG	GAATGGCAGG	TGCACCAGAG	CAGCAGAGGA	9060
CCTAATATCT	CCCTTCCTGT	TAACTTTTTA	TAATATTTTA	TTGTGATCAG	TATCAGTTGG	9120
GAAGCTACTT	GCAGTCACTG	AGCCTCAGTT	TCTACATCTG	TAAACTGGGG	ATAGTAGCAT	9180
GGCCCTATTT	AATGTGCTCA	GCGAAGCCAC	TGAAAGGAGA	CAGAAATGTA	CCAGAATTCC	9240
CTGGACTTTT	ATCCTACTTC	TCCTGGGGAT	TGTCACCCAC	CTACCCGTGT	CTGTCCTTTG	9300
TTGCTTTGAC	GCTGTCACTT	CTTTTCTTAG	GTACCTCTCT	GTAGGGCTCC	ATTATTCCAG	9360
GGATTCCAGA	GTTACAGCAC	ATGCATACCT	CCATCCAAGC	ATGTTTATTT	GTCTCCTGCT	9420
TCACTAGGCT	GTCCCCAAGG	AACATGTGGC	TCCCGGCACA	TACCTGGCAC	AACACTGCAC	9480
ATGACATTCA	CCCCTTGGC	CTTGAATCTG	ACAAGGAATC	TGGCATGATG	TTCACCTGCT	9540
GAGGCCAGGT	GCCGAGCAGC	CCTGGAGGCT	TAGGGGCCAG	AGGGATGGGA	AAAGGTGTCT	9600
TTCTGGGGTG	AGTATCAGTT	TCTGCAGGAG	TGCTGAACCT	GAGAAAGAAT	AAAGAGAGAA	9660
GGAAGTGAAC	AAGCACAGCT	TAAACATCAT	CTGTTTCTAC	TGAGTTTTAA	CAACTCTGAG	9720
ATTTTGTTTG	TCATGGAATC	CATTTCTCAG	GCCAAGCAGA	CACAGAACTT	GGGTGTGAGT	9780
GATGATAATG	AGCTGATATA	ATTTTCACAC	CCTCATCACT	GAGATCTCTC	CCATCAGGAA	9840
TGGGTCACAG	GGCTCACAGG	TGGCAGCAAC	TGTTATTACA	GGCCTCATCT	CTACCAGCTC	9900
CTGGCACCTG	CTCTCCTCTC	ATTAGAAAAT	CCTCCACTTG	TCAAAAAGGA	AGCCATTTGC	9960
TTTGAATTCC	AATTCCACCC	TCAAGAGGCT	GGGACCACCT	CATTGGAGTC	CTTGATGCTG	10020
TGTGACCTGC	AGTGACCACT	GCCCCATTGT	TGCTGGCTGA	GGTGGTTTGG	GTCAACCTGG	10080
CCATCTGGGC	AGCTGTTCTC	TTCTCTTCTT	TCTCCCCTAC	TGTTTCCAGA	CATGCAGTAT	10140
TTCCAGAGAG	AAGGGGCCAC	TCTTTGGCAA	AGAACCTGTC	TAACTTTCTA	TCTACGGCAG	10200
GACTTTTGAA	AGCTACAGAG	GAAGAAGCAC	AAATTGATGC	TATTCCACTA	AGCCATCAGC	10260
TCCATCTCAT	CCATGCCATG	TCTCTTTTTT	AGGGGTCCTC	TTGCCAACAG	AATCACAGAG	10320
GACAAATCTG	AAAGTGCAGA	GACAGCAGCT	GAGGCACAGC	CAAGAGCTCT	GGCTGTATTA	10380
ATGACCTAAG	AAGATGGAGT	GGTCACCAGA	AAGTCAGAGG	AAGTGACACA	CAGGGGCCCA	10440
GCAATCTCAG	CCAAGTCAAC	TCCACCAGCC	TTTCTGGTCC	CCACTGTGTG	TACAGCACCC	10500
TGATAGGGAC	CAGAGCCATG	AGAGTGAGTA	AGACCAGACT	ATGCCCTTGA	GGAGCTCACC	10560
TCTGCTAAGG	GAAACAGGCC	TGGAAACACA	CAATGGTGGT	AAAGAGGAAA	GAAGACAATA	10620
GAACTGCATG	AAGGGGATGG	AAAGTGCCCA	GGGGAGGAAA	TGGTTACTTC	TGTGTGAGGG	10680
GGTTGGTGAG	GAAAGACTCT	AAGAGAAGGC	TCTGTCTGGC	TGGGTATGAA	AGGATGTGTA	10740

```

GGAGTCTTCT AGGGGGCACA GGCACACTCC AGGCATAGGT AAAGATCTGT AGGCATGGCT 10800
TGTTGGGATG AGTTTCAAGT ATTCTGGAAT GAGGACAGCC ATAGAGACAA GAGGAGAGTT 10860
AATAGATTTT ATGCCAATGG CTCCACTTGA GTTTGTGATA AGAACCCAGA ACCCTTGGAC 10920
TCCCCAGTAA CATTGATTGA GTTGTGTATG ATTCTACATA GAATATTAAC TCAATGGAGG 10980
TCAGTGAGTG GTGTGTGTGT GATTATTTGC CAACTGCCGA GGTGGAGAAG CCTCTTCCGA 11040
CTGCAGGCAG AGCACGGGGG CCCTGCTACT GGCTGCAGCT CCAGCCCTGC CTCCTTCTCC 11100
AGCATATAAA CAATCCAACA GCCTCACTGA ATCACTGCTG TGCAGGGCAG GAAAGCTCCA 11160
CACACACAGC CCAGCAAACA GCAGCA 11186

```

INFORMATION FOR SEQ ID NO.: 5

SEQUENCE CHARACTERISTICS

LENGTH: 10200

TYPE: nucleotide

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ORIGIN: homo sapiens

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: 5' flanking region of human MDR1; -10,000 to +200 relative to transcription site of MDR 1 gene

SEQUENCE DESCRIPTION: SEQ ID NO: 5

```

GCCAATTAGA AGAAACACAA CTACAAGGTC AGGGCATATT ATTCAAACAG TAGAGACAAT 60
ACAGTCAAAT ATTTGGCAGA ATTACAAAAT ATCTCATTGG AAAAGACACG CAAGGGGAAAT 120
CAACAAAAAG ATATGAATCA GAATTCATCT GTGTCTCAAG AAAAGGTCAT GCGATAAATT 180
AAGTTCTGCT AGTGTTTCTA CACTACCGTT AGCCTCATTA CCTTATTTTT TAAGTGTTAA 240
TATAGTTTTA GGTATTTTAC ATACATTTTT ATTATTAATT ACAACCAAAG TGCAACTTGT 300
AATAGCAATT CCTTCACATT TTTTTTTTCA AATCTTGCAC CTTAAAATCC ACCTCGGGCC 360
TCAGTTGGCC AGCTTTGGTA TCTGATACTT GGACTACAGA TACCACTAAG GCAAGTAGAT 420
AAAATGTACT CTAGGACCTA CAGCCCTTCT GCTAGATCCT GAAGAATGAT CATTAAAACA 480
AGCTGGTCTA GCTGGTCAAG AGCAAAAATA AAATCAAGAT GACAGAAAAT TGATGCAAAA 540
GTGAAGTAAA ATAGCTAGAG AATATGATTG CGCCTGTCCC CTTAGCATGG ATTCCCATGC 600

```

TAGCCAATCT	AAAATCCTCA	CTGTTAGAAT	CCTCCTGTCA	ATATGATAGA	ATGAACAGCA	660
AGCTCAGTGT	CAGAAAACCT	GTGTTGTTAA	CTTGGCCCTC	TTTCTAGCTG	AATGTGTGTT	720
TTTGGTCAAG	TTCTTTGGCA	TTTCAGAGAC	TCAGAGTAGT	GAAGGAAGTG	GATAAGATGA	780
CCTCTACATT	CTCTTGCAAG	CTCAAACATC	TATGAATCCA	GAGAGAAAAA	CTAGAGCATG	840
AAATTAAGGT	TATTTTAAAG	AAATAACCTT	AAAATTATTA	GTATTCGAGG	ATCTCCAATA	900
TATTCATGGC	ACCACTCAAA	ACTTTCCTTC	TGCTCTATCC	CGTCTTGGCT	CAAAGTTATC	960
TCCTTAATGA	GGTCTGCCCT	GACTATCCTA	CTTAAAATTG	TAAACTTTGC	CCACCTGGTA	1020
CTTCCACTCT	CTTTCCCCTG	CTCTGTTTTT	CACCGTAATA	CTTTACTCTT	TTTAACATAC	1080
AAAATCACTT	ATTTACTGTG	TTGTTATCTA	TCTGCCTACT	CTTACCATCA	AATATAAGTT	1140
CTACCTAGGC	AGGGATTTTT	GTATGTTTTG	CTCATGGATA	TATACGAAGC	ACTTAGAGTA	1200
ATATGTGACA	TATACAGGGT	ACTTGATTAA	TACTGTTGAG	TGAATGAATG	AGTTTCCAAT	1260
ACAAATTTAA	AATAAAATAT	TTCCTAACTT	AAAATTGTAA	AGTCAGATCT	AACCAACTGT	1320
TCATTGGTCT	GCTAGCAGTG	TTTCTTGTAT	ATGGAAATAT	ATTTTAAATA	GATATGTCCT	1380
GTGAAATAAT	ACTAAGTGTT	CTAAAGAAAT	AAGTGAGTGA	ACGTTACCTC	ATTGAACTAA	1440
CTTGACCTTG	CTCCTGGGAG	AGAGTTCATT	TGAGATTAAA	CAAGTTCAAA	GTCTATGAAT	1500
CATAAAACGA	TAAAAAAAAC	TAAAAGGGAA	ATGGTGTTTT	TATAAGCTCT	GCAATTCAAA	1560
AGCCATTTTCG	GGTAATATTG	TTATTTTTTAT	GTCAGGAATT	CCTCAGTGCT	GATATCTTAG	1620
GGCAAAGGGT	TTGGTTATAA	ATTAAGAGAA	TGAGGAAATA	GGTACATAGT	AGGATTGTTC	1680
CAACCAAATA	TGTGTTGAAT	GTCAAAGGAA	TTTCCTTGAG	GAATAATCTT	CAGAATAATT	1740
TGCTAAGCAC	AGGAGAAAAT	TTGGCTTATT	ACTTTATAGC	CAGATTTTCAT	TTTTAATTGA	1800
AACTTCTTTC	AAGCAAATCA	CTTACTAGTC	TATTAACAAT	AACAACATAA	ACACAAGTAA	1860
ACATTCGGAA	TATAGACATC	CAGGTACTAA	GCTGATTGCT	TTACACTCAC	TGTCTTATTT	1920
TACAAGTAAG	GAGTTTTAGT	TGCAGCAAAA	GAAATAAATT	TTCCAATGTC	AAATGACCAG	1980
AACTTAAACC	CAATCTGTTT	GGTGCTAAAG	CCAATGTTCT	TTACTGCAAT	GTTGGGTTAT	2040
CTTGTTTCTA	AACTTAAAT	TTATCAGTAA	AAGGCAAAAT	TTGCTATTAT	TGAGGACATT	2100
AAAATCATAT	TTTTGTAGAC	TCTGAGGACA	AATCCAACAA	AAAAGTTCCA	ACTATTTCTT	2160
GGCAGGCATC	ATTGAAATTG	GTATATAGCT	TCCTTGGGTA	TTGACTTTGA	AAAGGAAGTT	2220
GGTCACTTTA	GATATATAAG	TTCAGTCTGT	TTGTAAAAAC	AAAATGAAAA	CAAAACAGTT	2280
GCCTTATATG	CTAAAATTAT	CCTAATCGTT	TTCACCTTTA	ACAACATATA	CACACAGAAC	2340

TTGAGGAACT	TTACACGGCT	CATCTTCATA	TTGTCAGCAT	CTAGCAAAGT	ACTTGCCACA	2400
TAGTGATCAA	TAAAAGTTTT	AGCCAGCCTG	GGCAACATAG	TGACAGCCTA	TCTCTACAAA	2460
AAAAAATTAG	CCAGGCAAGG	TGGCGCACAC	CTTTGGTCCC	AGCTACTTGA	GAGGAGGATG	2520
TGGGGAGATC	CTTTGAATGC	AGGAGGTTGA	GGCTGCAGTG	AGCTGTGATT	GCGCCACTGT	2580
GCTTCAACCT	TGGTGACAGA	GCAAGACCCT	GTCTCACACA	CACACAAATT	AAGTAAAGAA	2640
AAAAAAAAGA	ATCAAAGAAA	AAAATAATTC	CCCAGCTTAA	GTCCATCTTT	ATTTGTTTGG	2700
ATAAGCTATA	AAGTGTCAAA	TAATGCTGTT	AATGGACATT	TCTCTAGCTC	TCCCAAAGGA	2760
GGAATTGAGC	ACATAGTATG	TGCTGTATTT	TATATACAGA	ATAAAAATAG	AGACAAGATT	2820
TCTACCCTCA	CAGAACTTAA	ATTCTTCAGG	AGAATGACAC	TGAAGTCCTT	AATTGGACTT	2880
CTCTCTTCTG	TATTATCTTC	CTCAAGTGGA	GGTATATGGT	GCTTAGTTAT	GAAAAATACC	2940
TCCAGGGCTT	TGATCTTCTC	AATAACTCTT	TGAGGCTGAT	ATGAAAACAG	TAATTAGAAA	3000
AAAACCATGT	ATCCAACCTA	TATAGACAGT	TGATGACCAA	AGCTAGAATC	CAGTTATTTT	3060
AGCCTCCCAT	GTATTTTCTT	ATTACTTAAG	GAGAATCTCT	ATCTCTACCT	CTTTCTCTCT	3120
TCTTCCTCTC	TCACTTTTCT	TAGAAACATG	GGTAAGATTT	TCAGAAATAT	GAGAACTTA	3180
TTAATAAATG	AAAAATACTG	GGAATTCTCA	ATGTTTCTTG	TTTtagccag	TTAATTTTGG	3240
CCTTCATTCA	ATGTGAGTGT	CCCTTAATAA	GGAGCAAAC	CCACTGAGAG	ATAGATACTA	3300
ATAACCAGGA	TTCTGAAAAT	GCATTCTCAT	CCCCATCTCC	AACTTTTAT	AAAAAATATT	3360
ATAAAATAAT	ACACTTTTAA	TATAGGAAAT	TTCTCAAATA	CAGAAAAAAT	TAAAGTAAAC	3420
TCAGACCTAA	CTTTCATCAC	TAAAAGATAA	TCACTCTTGA	CACTTTGATA	TCTTTATCTC	3480
TAATATTACA	CCAATTAATT	TGCTTGATAT	AGTGAATATT	ATGCTATTAT	AATTTTCCCC	3540
TGCCTTTTGT	CCTTGCATAT	TAGCATAGGT	ATTTTCTGAG	GTTATCACAA	ACTCTGTAAG	3600
CACGTTTTAT	ATTACTACTT	TTTTAAAGAG	GATGTATAAT	AATTAATTCA	TCCATATATA	3660
TGTGGTTAAG	TATTCAGGTC	ACTGCTCATT	TTTCACTGTT	ATAAAATAAA	GCAGCAATGA	3720
ATACCTTTGG	CTGATATTTT	TTTCTGTACT	TGGAATTATT	TCTTTAAGAT	AGATTTCCTA	3780
AAATTGAATT	ACTGAGTCAA	ACAGACTTAA	GTTTTCCTTA	TGTATGTTTC	CTTATTCATT	3840
TGAATAATTT	TCAACTCCTA	CTTGTTTATT	TAACTCTTGT	GAGCATGTGA	TAGTCTCATT	3900
TTTCAAATA	TCTTTGCTGT	TGTAATTTGC	ATTTCTTTGT	AGTTAGCATG	AATATTTTCA	3960
TATGTTTTCT	TCTGTGTACC	AGTATACTAC	ATACTTTTTT	ATATGAATTG	CCTATTTGCA	4020

TCTTTTGCTC	GGTTCTATTA	GACCTTTGAA	TTTTTTCTTA	TCCATTTATA	TAAGCTCTTT	4080
ATATATTAAG	AATATTAACC	TATTGTGATA	TTTGCAATAA	ATAGCTATAT	GGTTTGTTGT	4140
TGGTTTTAAA	TGTGAATTTA	TTCAATTTTC	TTCATAATTT	TGTTGTTTTT	ATAGATTTCT	4200
TTAAAAGTAA	TAAAATTATT	GCCATATTAT	TATTTATTTT	CAATGTCCAT	TACTAGCCCT	4260
TTCCAGTCAT	CACTTTTACT	CTCATGTTCT	TAATTTTAT	TCATATCTTG	GCTCCATTGA	4320
CCATCTTTAT	TAGGATATTT	GGGAAATAAT	ACAATTTATA	CTAAACACAC	ATCAAATCTT	4380
ACCTTATTTT	CTTCTTACAA	AAACCCTAGG	AATATGCTGT	TTTTGTCTTT	ATTTGAATGA	4440
CACAGAAATC	AAGGTTTTTG	AGCAGTGGAG	TATTTCTTCA	AATGACACAG	AAATCAAGTC	4500
TTTTGAGCAG	TGGAGTATTT	CTTCAAAGCC	ACACAGCTAG	TAAGTCATGA	AGCTGGAATT	4560
CCAAGAGTTG	CCACTTCATT	TTCTTCTTTC	CCTTTATCTT	ACTCAGTTGT	CTTCTCTCCT	4620
CTTAATTTTG	TCATTCATTT	AAAAACATTT	CTTGTGCTAT	TATGGTAGAT	TTATTTTAAT	4680
AGGGGGCAGT	GACTTACTCA	GAGAGATGAT	TCTCTAATGG	AGTTTTTAAAG	ATCTTAGAAG	4740
TTGATAGAGG	AGGCTGGGCG	TGGTGGCTCA	TGCCTATAAT	CCCAGCACTT	TGGGAGGCTG	4800
AGGTGGGTGG	ATCACTTGAG	GCCAGAAGTT	CAAGATCAGC	CTAGTCAATG	TGGTGAAACC	4860
CCATTTCTAC	TAAAATTACA	AAAATTAGCC	TGGTGTGCTG	GTGCACTCCC	ATAATCCCAG	4920
CTACTCAGGA	GGCTGAGGCA	TGAGAATTGC	TGGAACCCAG	GAGGCGGAGG	TTGCAGTGCG	4980
CCTGGATTAT	TAACACTACA	CTCCAGCCTG	GGAGACAGAG	TAATACTCCA	TCTCAATAAA	5040
AAGAAGTTGA	TAAGGGAGAT	AGTTCATGGC	AACGGATCTT	TGAAGGCACG	CTAATGATAA	5100
CTTAGGCATT	TAGCCTACTA	GTGTAATTTT	CATAAATCTG	CCTCTGATGT	CATACTCTCA	5160
GCACCTAATA	TTTTCTACAA	ACATTTATTG	AAACTTTATT	TTGTATAAGT	CTCTGTCCAG	5220
TTTGAATATT	TAAAAAATTC	ATAATCATAT	GAAACATTAA	TAATAAATAC	AAAATGAGAG	5280
ATGCCGATAC	TGAAAAGTAG	GATTGCGGAG	TGGTAGAAAA	TATTTCTGGC	TGTAGTAGAT	5340
GGGGAAGTGT	TCAAAGAGGA	GTATAATTCA	GGTTTCCATT	TGCCATCGAC	TTATCACATG	5400
GCTAACTCAC	TAAGCGACTT	AATTAAAATT	AAATTAATTT	ATCATCATCT	GATCACCATT	5460
TCACACAACT	CATGTCTGTT	GCTGTATTGG	CTAAATGATG	GCAAGACAAA	CGACCTCTGA	5520
AAATGATCCT	ATTGACCTTC	GGAATCTGGA	TTTTTTTTTC	AATGCAGGTG	TCCATAGAAG	5580
CAATCTGATG	TAATCCAACA	TGAGTTCAAG	CACAGTCATT	TAATATCCCC	TATCAAGTAC	5640
AGTCATTTAA	TATCCCCTAT	CATCACATGT	CCTTCATACA	TAAAAATCAT	TACATGTGAA	5700
AGGGTGGAGA	GTGTGTGGAT	CCCTTATTAT	TGTGTTATTG	TAACACAATA	ACAATATTGT	5760

GTTATTATTG	TTAACACAAG	TGAGTCATAT	GTCTTGCTCT	TTGGACTGAG	TGGAAACTTG	5820
TATTCTTTCT	CTGCCTCAGG	TCAATTAAC	TCATTGAGGT	GAGTTGCATT	CCTTCTTTAA	5880
GCATGTTGAA	CCTTCAATCT	GGACTCAGAT	GGGCTAAATA	GAGGAGCTAG	GAAAAATACA	5940
GAAAATAAAT	TATTAGAGAG	ATCAGAGAAA	GATACATAAG	ATTTACAATA	AAAGAATTAT	6000
GAGAAAAACA	TCCAAAAGAA	TTAAAAACCA	TAGGAGAAGG	AAAAATAGGT	GAACAGCTTT	6060
TTAATTTCTA	TAAATGTGTT	GTTAATACTC	ATAATAAAGG	ACTCAGAGCT	GGGATATGAG	6120
AATAATAGGT	CAAACGTATA	TGGATACATA	GATGTGACTA	CATACATGAG	TTGCAAAGAA	6180
TGCTAAGGAG	GGCAAAAAGA	GATTGAGAAG	AGGGCATTAT	TACTAATATA	TAGCAATGTT	6240
GAATGTTTAG	GGTGTTTCAG	GCACTGTACA	AATCTTTTAA	ATACACAAAT	CACTTAATCT	6300
TGCCATAACA	TTAGAAGATA	CTATCTACTC	TTTACCAAAA	AGGTAAGTGT	GGGATAGCAG	6360
AGTTAAGTAA	CCCGTTCAAA	CCTATGCATA	ATAATCAGCA	GAGATGGTCC	TTATCTAAGT	6420
TTTTCTGCCT	TTGAAGTCCA	AATAGTTTAA	TGCAGCCAGG	TACTAAAGAA	GAAAACTTTT	6480
GTAAATTAGT	TTAGTTTAAT	GATTTACATG	TGGAAAAGCA	CAGAGTGAAA	AGCACATATC	6540
ATCTGAGAAG	CCCAGTGAGT	TTGGCTGAAA	TGGAGTGAAC	ATGTACATGT	TGAGGGTGAG	6600
GAAGATGATT	AGAGAGAAGT	GATGTGTTGT	GGTTCTTAAA	AGCTAAGAGG	AAACTGTTAG	6660
ATATGATATA	GTCTGTGGCA	GGGAGCCATT	GCAGGTTTTC	CAAGATGGAC	TCATAAGGAG	6720
AAAACACTTC	GTGGGACTGG	AAAAGGCAGT	GAAGTGGTGT	GTCCTATGAT	AATGTACTAC	6780
AGTGTAGGAT	AGTGGTTAAG	AGTACAGTTA	TGAGAGAGGG	ACTACTGGTT	AGCACCTTAC	6840
CTGCTGTGTG	ACTGGGCAAA	TAATGCAAAC	CTCAGTGTCT	TTTATTGTAA	TATGGGAGTA	6900
ACAAAAATAG	TAACTACTTC	ATAGGATTCT	TGTAAAGATT	AAATGACTTA	ATTTCTTTGA	6960
AGTGCTTGGC	AGTTCCTGAT	AAATGACCAG	TAGTTAATAA	ATGTTAGTTG	TTATTATTAT	7020
CATTATATAT	TATTACTCCC	ATAGATACAT	ATAGAACAGA	CTGCAGCAGA	GAGGCAAATC	7080
TTTAATGTTG	TCAGAGTATA	GACAAGTTGG	TGAAATGGCT	ACATGAGAGC	GGAGGACAAG	7140
AAGGTGCAGA	TTGTGGCAGT	CACTTCAAAT	GGAAATATCA	CCGCTTGAAT	GAAGGTATAT	7200
GAGTGTCAAC	TTGCAAGGGG	ACCAGGTAGG	TTTCATCAGA	AATTAAGGAA	GCTTAAGGAG	7260
AACAGCCAAG	TTCAGCTTGA	CAGAAGTGGT	GGTGGCACAA	ATGCAAGACT	GGTGTCTTTC	7320
AAGAAACCAA	GGACTGTTGA	AAGTAGCAAG	AGCTAGTTTG	TTTTAGGTCC	ATCATGTTTT	7380
ATATTCACAC	TTTCATGTCA	GTGGAGCAAA	GAAATGGAAT	ACAATATAAT	AGAATGGTAG	7440

AATCTTATTT	TTAAAATCTG	TGTTATTCTG	ATCTTTAACT	TACTTATATC	TTTGATAGAG	7500
ATCTTTACCT	GATGCTCAAG	ATTGTAGAAA	TAGTATAATC	AACATAACAG	TATAGCACTG	7560
TATTTATATC	CTGCACTGTT	TAGGGAGGGT	TTAAGGCCAT	TCAAAAGGAT	ACATAAAATA	7620
CAACAAGATT	ACATAAATGA	AAGGTGAGAT	AAAGCAACAA	AGCAAAACAA	AAGTGAAAAC	7680
AGAGATCATA	GGCACAAATA	AGATTAAAAA	CGCATGTAAT	GAAGATGAAA	GCTTTTACAT	7740
TTACCCCAGA	TGGACCACAG	GGTTGTTGTT	AAGCCTTTAA	ACAGTGAACA	ATGCTGTACA	7800
CTTGCAATATG	CAATTAGAAC	ATGTGGAAAA	AATAGTGGCC	TGTTAGAAGC	CTAATTAACA	7860
ATTTGTGAAA	AAAAAAAAAA	AAAAAAAAAA	AAGAGGCCGA	GCTGTAGCTC	ACGCCTGTAA	7920
TCCCTGCACT	TTGGGAGGCC	GAGGCGGGCG	GATCACGAGG	TCAGGAGATC	AAGACCATCC	7980
TGGCTAACAC	AGTGAAACCC	AGTCTCTACG	AAAAATACAA	AAAATTAGCC	GGGCGTGGTG	8040
GCGGGAGCCT	GTAGTCCCAG	CTACCTGGGA	GGCTGAGGCA	GGAGAATGGT	GTGAACCCGG	8100
GAGGCGGAGC	TTGCAGTGAG	CCGAGATCCT	GCCACTGCAC	TCCAGCCTGG	GCGACAAAGC	8160
AAGACTCCGT	CTCAAAAAG	AAAAAAGAAA	GAAAAACAAA	AGAAAACCTC	ATTGTATTGT	8220
AAGGCCAAGA	ACAAAATATA	TCAAGATAAG	GAAAATTTGT	AGTCAAGAAT	AGAAAAAAAT	8280
TATGGCTTTG	AAGTATGAGT	TATTTAAAGA	AAGTGGAAC	ATCCTCAGAC	TATGCAGTAA	8340
AAAACAAAGT	GATTTTCTTC	TTCTAAACTT	ATGCAATAAA	CTGATAGGTA	ATATGTGAAA	8400
GTCATAGAAT	GTAGACTAGA	GGATACAACA	AACCTATTTT	CTCTATGTTC	ATAAGAAGTA	8460
AGAAAAGCTC	TGATGTGAGT	TAGCATTGCT	TTACAATTTT	GAATTGTGCA	GATTGCACGT	8520
ACTTTTCCTC	AGTTTGAAGT	AAATAGTGGA	CAGGAAAAAA	TATTAAATGT	TGGCAGTAAA	8580
TATGGAAGGA	AATTACAAC	AATGTAATAT	GCTAAAACAT	GCTATGTTTA	TTTTACTAAT	8640
TTGAATTAAA	ATGTAAGAAT	TTAAAATGCC	CTGGAAAAAC	ACGGGCATTG	ATCTGACGTC	8700
TGAAGTTTTA	AAATATTACA	CACTTTGAAA	TAGCATTTGT	ACCTTGAAAT	ACCTGTCTCT	8760
ATATATTTTT	TAAAACCTCC	TTTTTCTTTC	ATTCCATTTA	TCATCAAATA	AAGGATGAAC	8820
AGATGTAACT	CAGAACTGT	CAAGCATGCT	GAAGAAAGAC	CACTGCAGAA	AAATTTCTCC	8880
TAGCCTTTTC	AAAGGTGTTA	GGAAGCAGAA	AGGTGATACA	GAATTGGAGA	GGTCGGAGTT	8940
TTTGTATTAA	CTGTATTAAA	TGCGAATCCC	GAGAAAATTT	CCCTTAACTA	CGTCCTGTAG	9000
TTATATGGAT	ATGAAGACTT	ATGTGAACTT	TGAAAGACGT	GTCTACATAA	GTTGAAATGT	9060
CCCCAATGAT	TCAGCTGATG	CGCGTTTCTC	TACTTGCCCT	TTCTAGAGAG	GTGCAACGGA	9120
AGCCAGAACA	TTCCTCCTGG	AAATTCAACC	TGTTTCGCAG	TTTCTCGAGG	AATCAGCATT	9180

CAGTCAATCC	GGGCCGGGAG	CAGTCATCTG	TGGTGAGGCT	GATTGGCTGG	GCAGGAACAG	9240
CGCCGGGGCG	TGGGCTGAGC	ACAGCCGCTT	CGCTCTCTTT	GCCACAGGAA	GCCTGAGCTC	9300
ATTCGAGTAG	CGGCTCTTCC	AAGCTCAAAG	AAGCAGAGGC	CGCTGTTCGT	TTCCTTTAGG	9360
TCTTTCCACT	AAAGTCGGAG	TATCTTCTTC	CAAAATTTCA	CGTCTTGGTG	GCCGTTCCAA	9420
GGAGCGCGAG	GTAGGGGCAC	GCAAAGCTGG	GAGCTACTAT	GGGACAGTTC	CCAAGTGTCA	9480
GGCTTTCAGA	TTTCCTGAAC	TTGGTCTTCA	CGGGAGAAGG	GCTTCTTGAG	GCGTGGATAG	9540
TGTGAAGTCC	TCTGGCAAGT	CCATGGGGAC	CAAGTGGGGT	TAGATCTAGA	CTCAGGAGCT	9600
CCTGGAGCAG	CGCCCAAACC	GTAGTGGCAC	TGGACCATGT	TGCCCCGAGC	GCGCACAGCC	9660
CGCGCGGTGC	GGGGACCTGC	TCTCTGAGCC	CGCGGGCGGT	GGGTGGGAGG	AAGCATCGTC	9720
CGCGGCGACT	GGAACCGGGA	GGGAGAATCG	CACTGGCGGC	GGGCAAAGTC	CAGAACGCGC	9780
TGCCAGACCC	CCAACTCTGC	CTTCGTGGAG	ATGCTGGAGA	CCCCGCGCAC	AGGAAAGCCC	9840
CTGCAGTGCC	CATCGCGGCC	AGAGCAGCTG	GGGCATCAAC	GGCGGGCGCT	CCCTCTTACT	9900
GCTCTCTGGC	TTCGACGGGG	GACTAGAGGT	TAGTCTCACC	TCCAGCGCGC	CTGAGGCTCA	9960
TGCATTTGGC	TAATGAGCTG	CGGTTTCTCT	TCAGGTCGGA	ATGGATCTTG	AAGGGGACCG	10020
CAATGGAGGA	GCAAAGAAGA	AGAACTTTTT	TAAACTGAAC	AATAAAAGGT	AACTAGCTTG	10080
TTTCATTTTC	ATAGTTTACA	TAGTTGCGAG	ATTTGAGTAA	TTTATTTCTA	GCCTCCAGCT	10140
CTGAAATAAA	TGACATGTTG	TTGTTTTTAA	TTATTTTTTAA	GAAACGCAAG	CTAGCCTTTG	10200

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

1. Use of a non-human mammal comprising:

(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to a sequence of the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located about 13,000 nucleotides upstream from the site; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule,

the mammal being characterised in that it has a pattern of constitutive and xenobiotic inducible expression of the reporter molecule that reproduces the tissue specific expression of CYP3A4 that is observed in a human,

to determine whether a compound is likely to induce the CYP3A4 gene.

2. Use of a non human mammal comprising:

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the nucleotide sequence of the human CYP3A4 gene that extends from the transcription initiation site of the gene to about

13,000 nucleotides upstream from the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule

the mammal being characterised in that it has a pattern of constitutive and xenobiotic inducible expression of the reporter molecule that reproduces the tissue specific expression of CYP3A4 that is observed in a human,

to determine whether a compound is likely to induce the CYP3A4 gene.

3. A use according to claim 2 wherein the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:1.

4. Use of a non-human mammal comprising:

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 8,000 nucleotides upstream from a position about 3,000 nucleotides upstream from the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule,

5 the mammal being characterised in that it has a pattern of constitutive and xenobiotic inducible expression of the reporter molecule that reproduces the tissue specific expression of CYP3A4 that is observed in a human,

10 to determine whether a compound is likely to induce the CYP3A4 gene.

5. The use according to claim 4 wherein the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:2.

6. Use of a non-human mammal comprising:

15 (a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 600 nucleotides upstream from a position about 7,200 nucleotides upstream of the initiation of transcription site of the gene; and

20

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;

25 wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule,

the mammal being characterised in that it has a pattern of constitutive and xenobiotic inducible expression of the reporter molecule that reproduces the tissue specific expression of CYP3A4 that is observed in a human,

5 to determine whether a compound is likely to induce the CYP3A4 gene.

7. A use according to claim 6 wherein the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:3.

8. Use of a non-human mammal comprising:

10 (a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which has the sequence of a fragment of the CYP3A4 gene consisting of from nucleotide positions -7836 to -7207 contiguous with nucleotide positions -362 to +53
15 wherein the nucleotide positions are relative to the transcription initiation site of the CYP3A4 gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic
20 acid molecule by the regulatory nucleic acid molecule;

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule,

25 the mammal being characterised in that it has a pattern of constitutive and xenobiotic inducible expression of the reporter molecule that reproduces the tissue specific expression of CYP3A4 that is observed in a human,

to determine whether a compound is likely to induce the CYP3A4 gene.

9. A use according to any one of claims 1 to 8 wherein the regulatory nucleic acid molecule is capable of regulating transcription of a reporter molecule selected from the group of reporter molecules consisting of firefly luciferase, β -galactosidase, alkaline phosphatase, green fluorescent protein and chloramphenicol acetyl transferase.

10. A use according to any one of claims 1 to 9, further comprising:

(c) at least one further regulatory nucleic acid molecule which is capable of regulating transcription of a human gene; and

(d) a further reporter nucleic acid molecule for producing a detectable amount of a further reporter molecule for indicating regulation of transcription of the further reporter nucleic acid molecule by the further regulatory nucleic acid molecule;

wherein the further reporter and further regulatory nucleic acid molecules are arranged to permit the further regulatory nucleic acid molecule to regulate transcription of the further reporter nucleic acid molecule.

11. A use according to claim 10 wherein the at least one further regulatory nucleic acid molecule has the sequence shown in SEQ ID NO:4.

12. A use according to claim 10 wherein the at least one further regulatory nucleic acid molecule has the sequence shown in SEQ ID NO:5.

13. A use according to any one of claims 1 to 12, further comprising at least one human transcription factor for regulating transcription of the human CYP3A4 gene.

14 A use according to claim 13 wherein the transcription factor is a nuclear receptor.

15. A use according to claim 14 wherein the nuclear receptor is a heterodimer of the human pregnane X receptor and human 9-cis retinoic acid receptor or a heterodimer of human constitutive androstane receptor- β and human 9-cis retinoic acid receptor.

16. A use according to any one of claims 10 to 15 wherein the further regulatory nucleic acid molecule is capable of producing a further reporter molecule selected from the group of reporter molecules consisting of firefly luciferase, β -galactosidase, alkaline phosphatase, green fluorescent protein and chloramphenicol acetyl transferase.

17. A use according to any one of claims 1 to 16 wherein the mammal is a mouse.

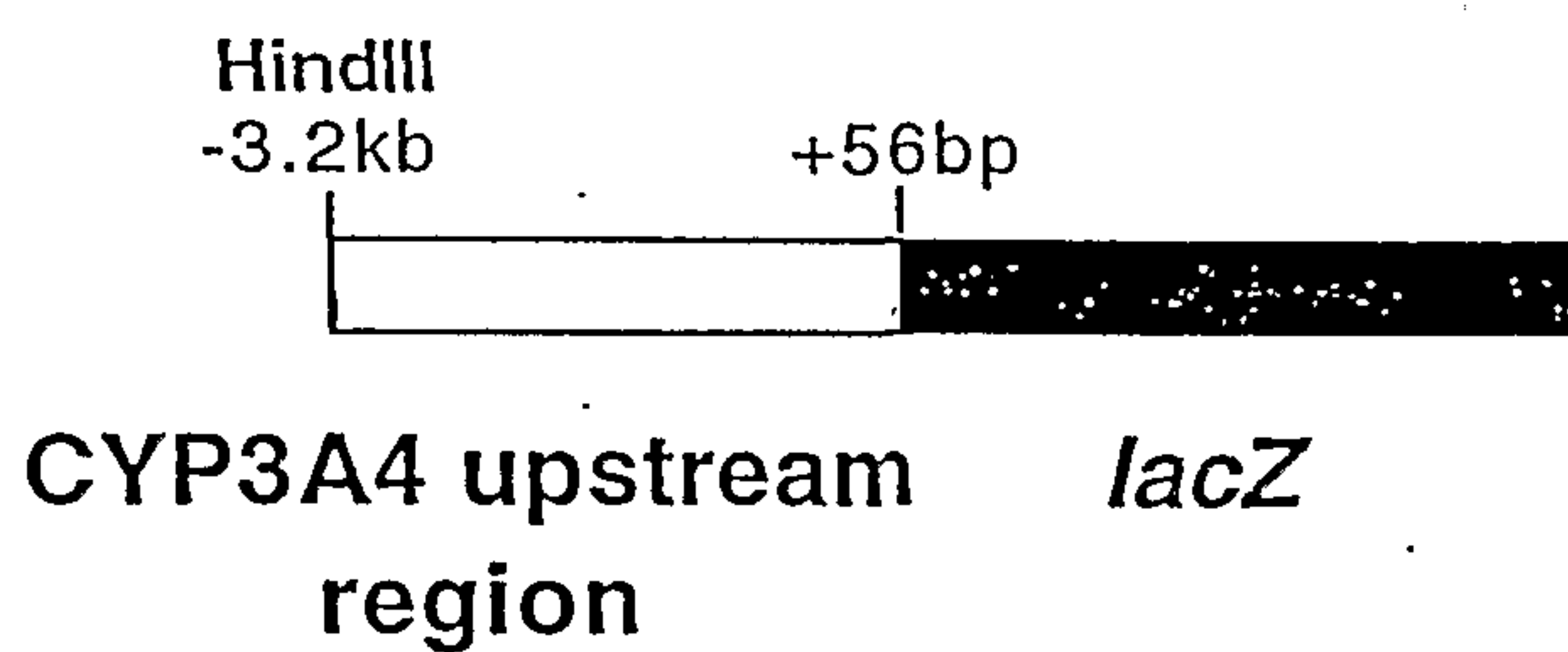
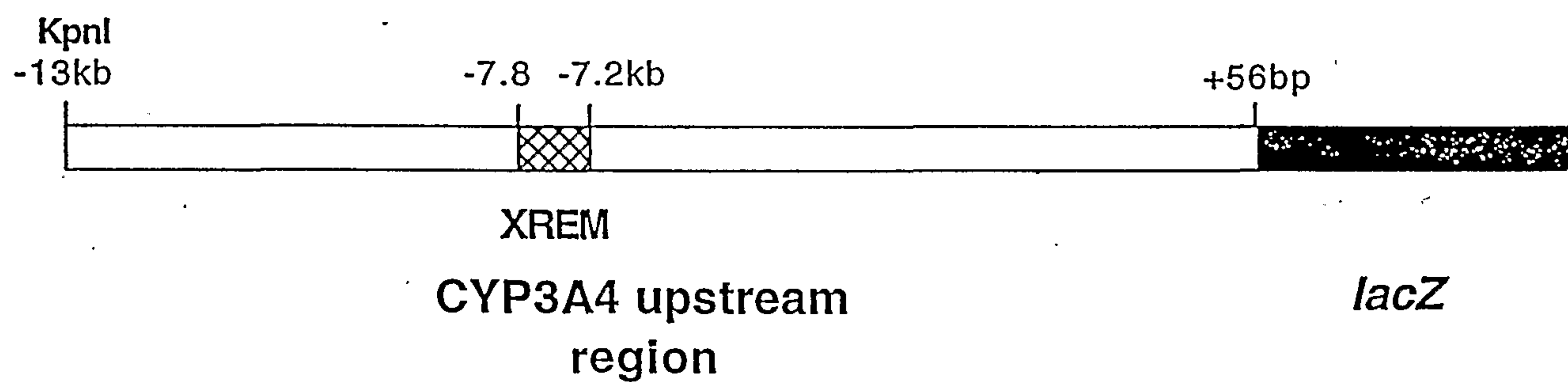
18. A somatic cell derived from a tissue obtained from an organ or tissue of a non-human mammal as defined in any one of claims 1 to 17.

19. A method of determining whether a compound is capable of effecting the transcription of a human CYP3A4 gene the method comprising the following steps:

- (a) administering the compound to a non human mammal used according to any one of claims 1 to 17; and
- (b) determining whether the reporter molecule is produced by the reporter nucleic acid molecule in the mammal.

20. A method according to claim 19 wherein the production of the reporter molecule indicates that the compound is capable of effecting the transcription of the human CYP3A4 gene.

1/22

Fig 1. Human CYP3A4/*lacZ* transgene constructs**-3 CYP3A4/*lacZ*****-13 CYP3A4/*lacZ***

2/22

Figure 2. Xenobiotic induction of CYP3A4/lacZ transgene



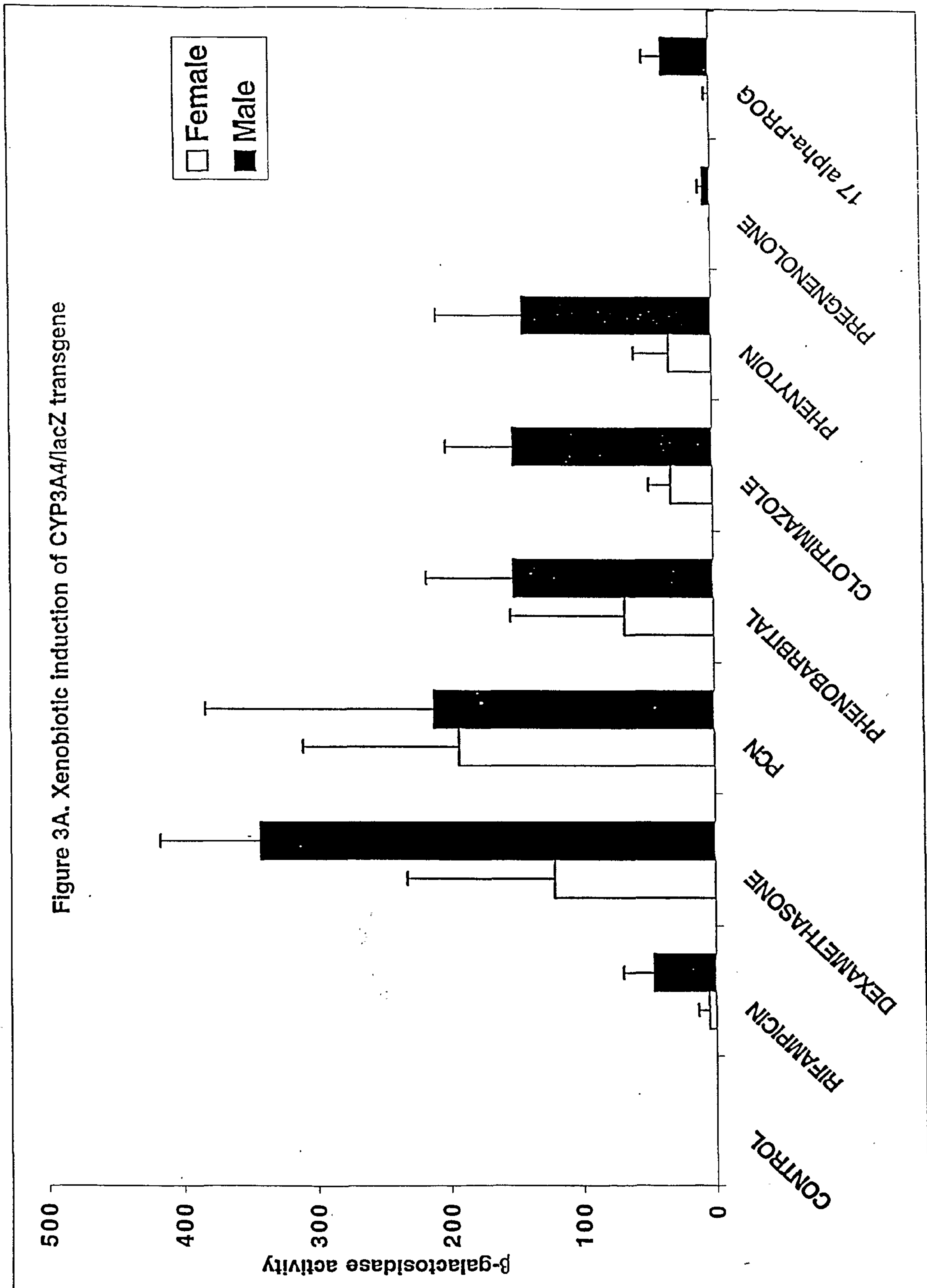
Pregnenolone-16 α
carbonitrile

Phenobarbital

Rifampicin

Control

3/22



4/22

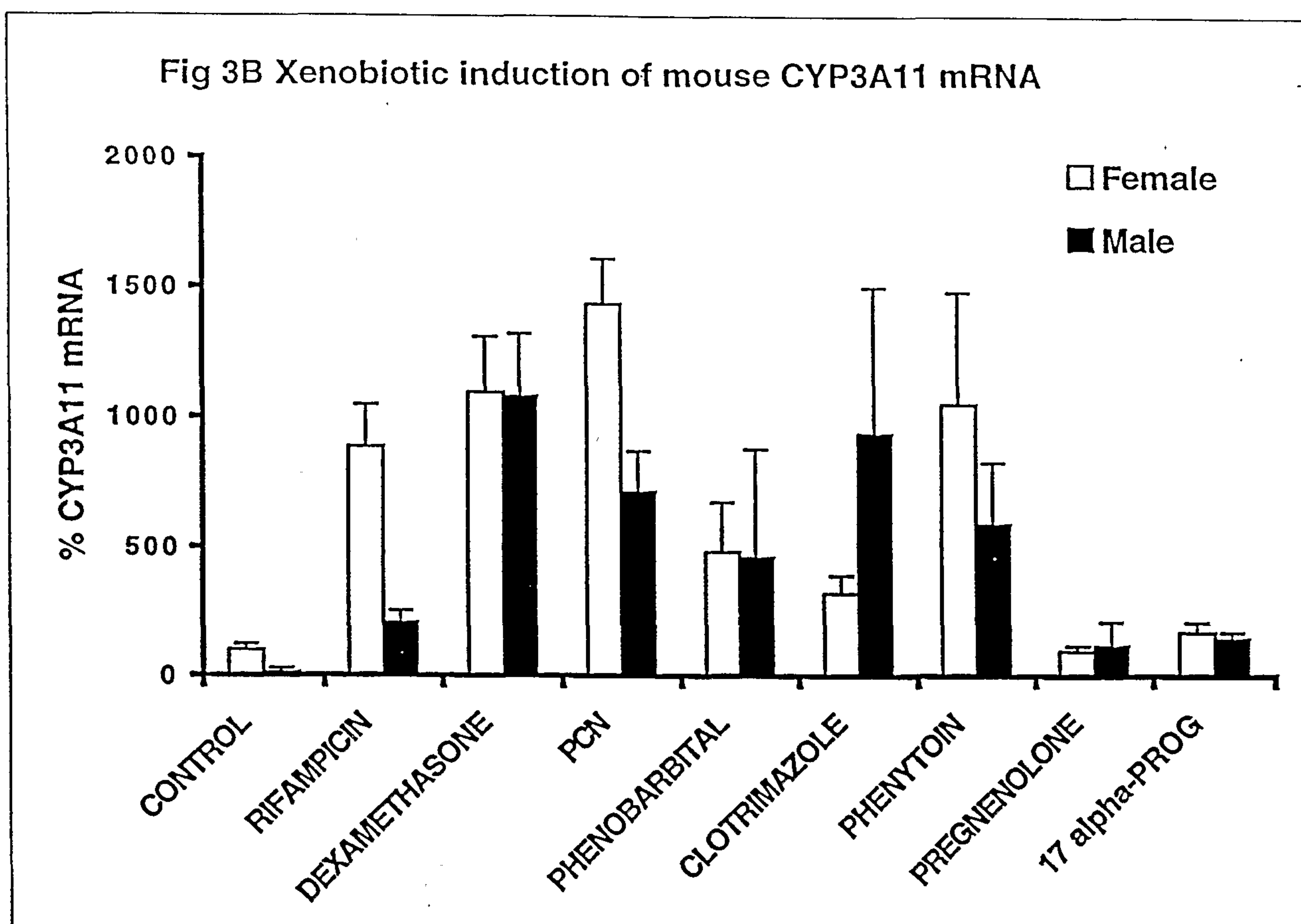
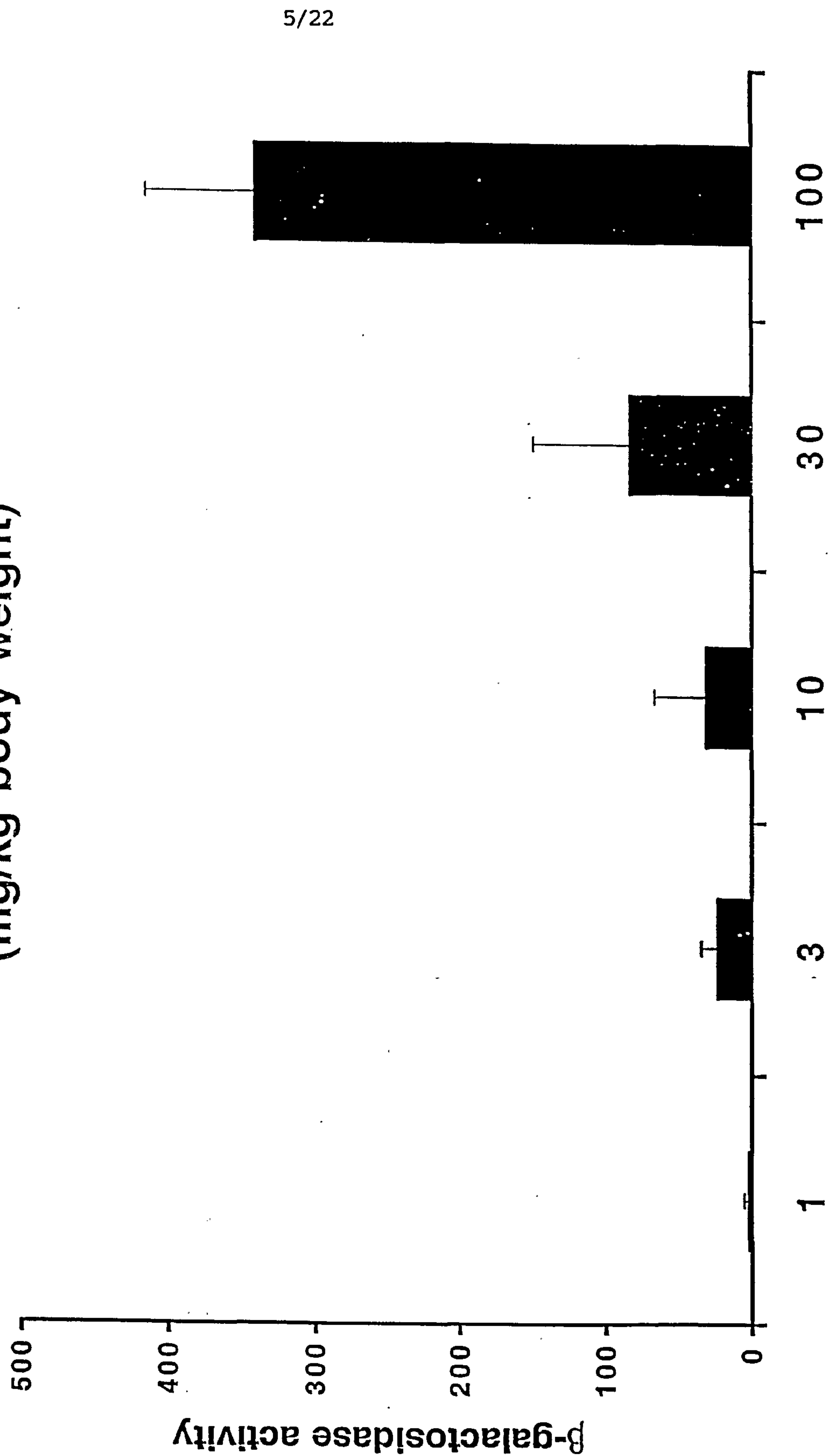
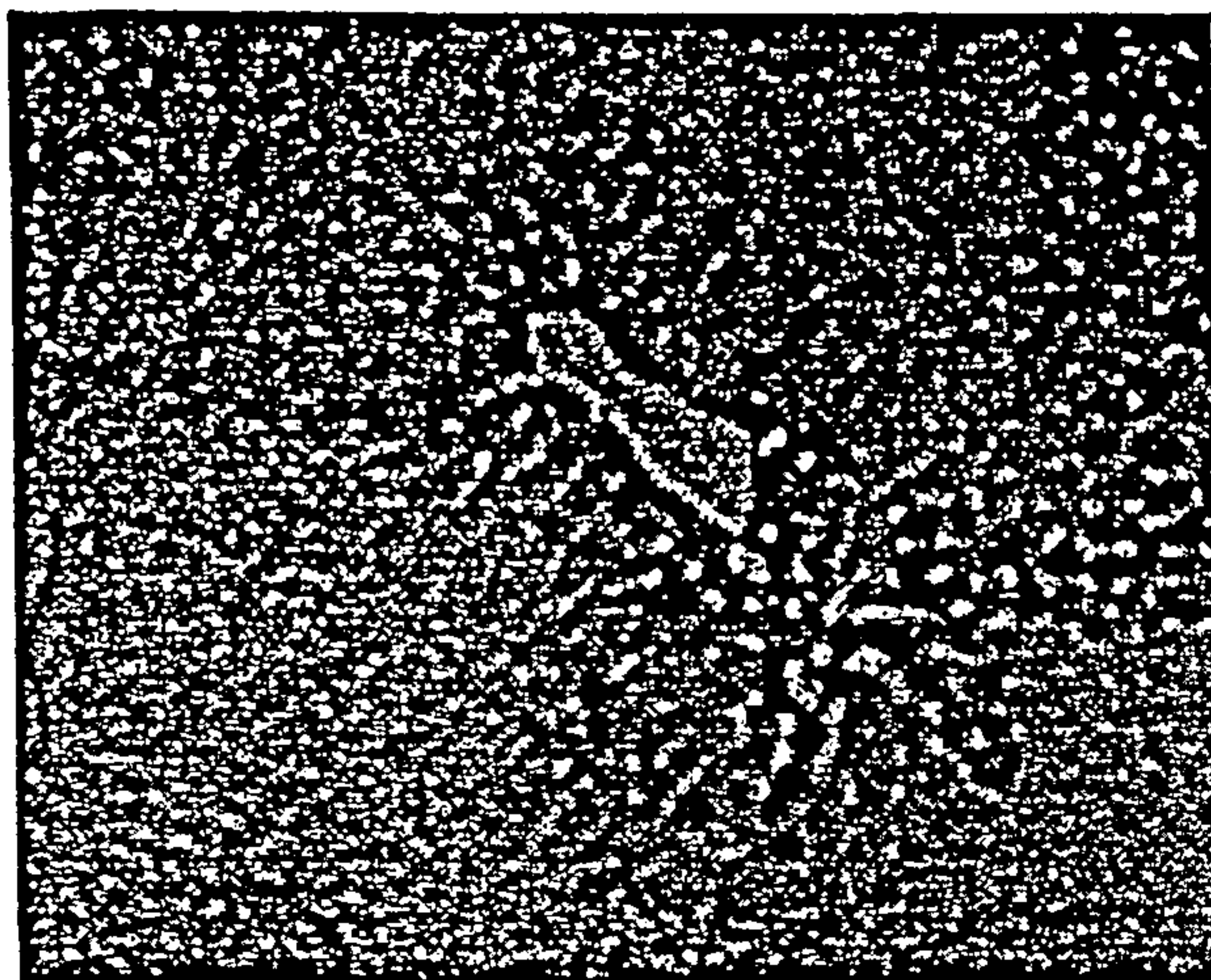


**Figure 4A Dexamethasone dose-response
(mg/kg body weight)**

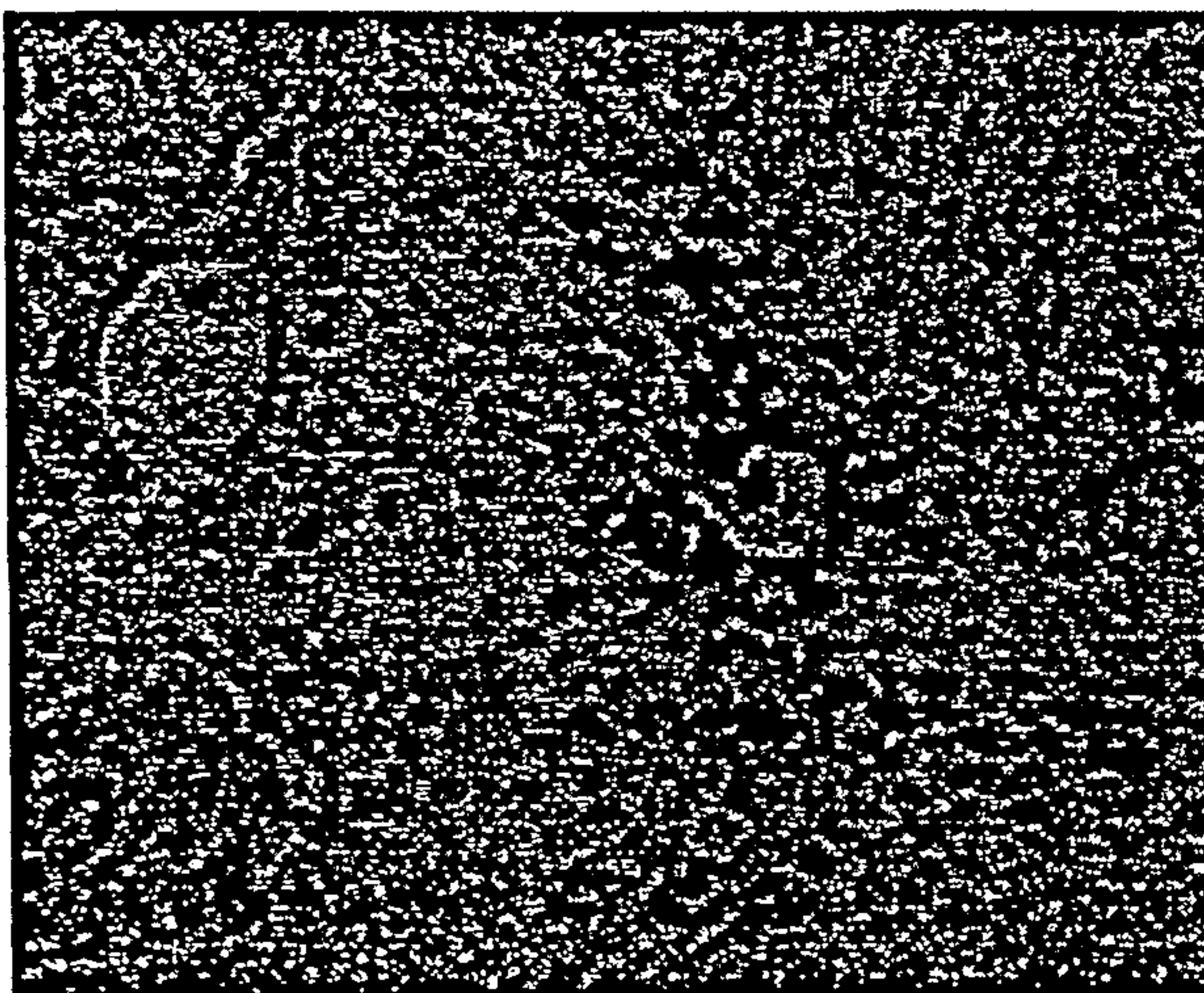


6/22

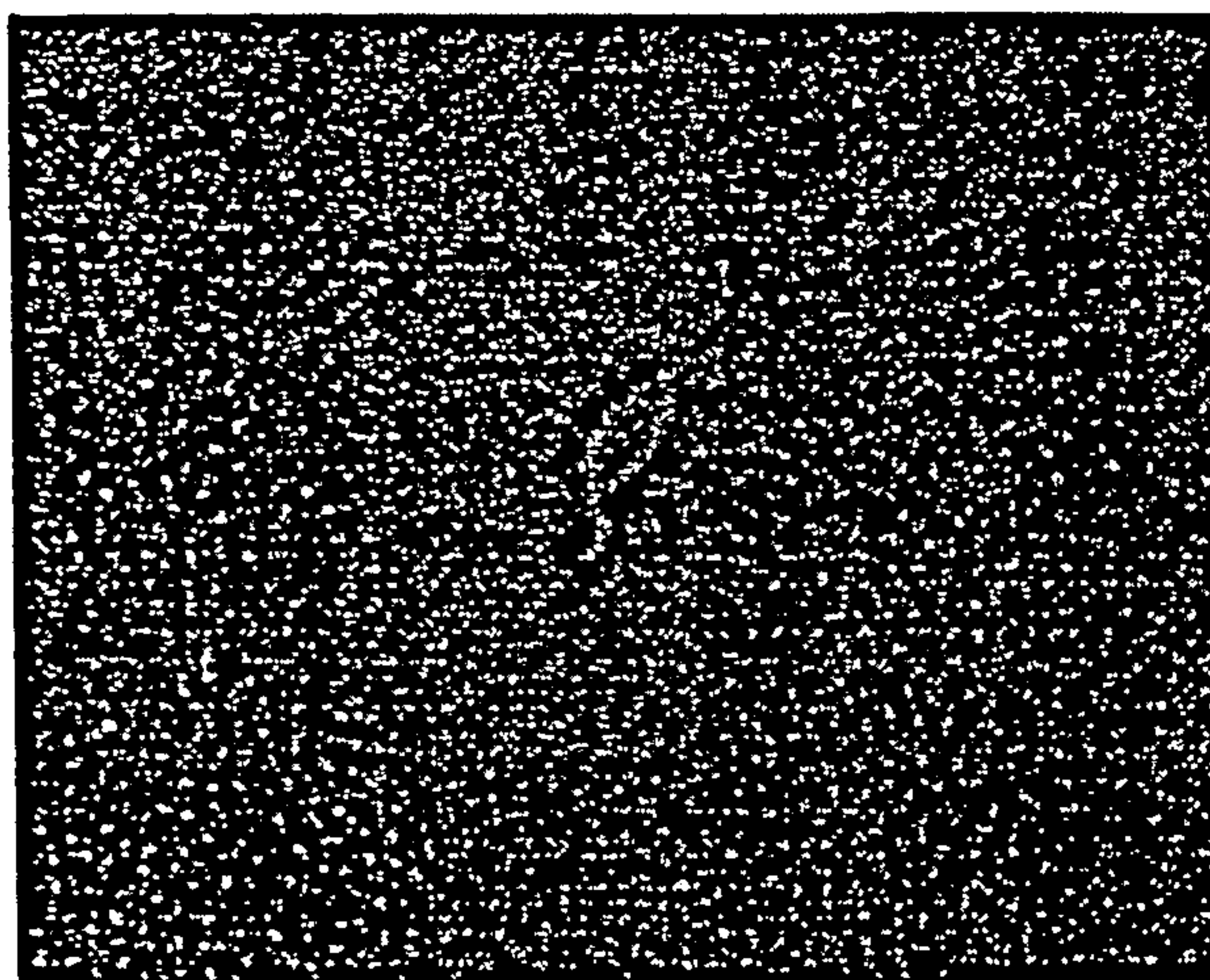
Figure 4B. Dose response of transgene expression



100 mg/kg



10 mg/kg



1 mg/kg

Figure 5.

CTGGTTCATCTCATTGGGACTGGTTGGACAAGAGGGTGCAGCCCACGGAGGGTGAGCCAAAGCAGGGTGGG
GCGTCGCCTCACCTGGGAAGCACAAGGGGTCGTGGAATTTTCTCCCCTACCCAAGGAAAGCCATAAGGGAC
5 TGAGCCTGAGGAACTGTGCACTCTGGCCCAGATACTGCACTTTTCCCATGGTCTTTGCAACCCGCAGACCA
GGAGATTCCCTCCGGTGCCTATGCCACCAGGGCCCCTGGGTTTCAAGCACAAAAGTGGGCAGCCATTTGGGC
AGACACCGAACTAGCTGCAGGAGTTTTTTTTTTTTTTTTTCCATACCCCATTTGGCACCTGGAACGCCAGTGA
GACAGAACCGTTCACTCCCCTGGAAAGGGGGCTGAAACCAGGGATCCAAGTGGTCTGGCTCGGTGGGCCCC
ACCCCATGGAGCCCAGCAAACAAAGATTCACCTGGCTTGAAATTCTTGCTGCCAGCACAGCAGCAGTCTG
10 AGATTGACCTGGGACCCTCGAACTTGGTTGGGTGCTGTGGGGGGGCATCTTCCATTGCTGAGGCTTGAGTA
GGTGGTTTTACCTTCGCGGTGTAAACAAAGCTGCTGGGAAGTTTGAACTGGGTGGAGCTCACCACAGCTCA
GTAAGGCCACTGTGGCCAGACTGCCTCTCTGGATTTCTCCTCTCTGGGAAGGATATCTCTGAAAAAAGGC
AGCAGCCCCAGTCAGGGACTTATAGATGAAACCCCATCTCCCTGGGACAGAGCCCCCTCGGGGAAGAGGTG
GCTTCCACCATTTGTGGAAGACTGTGTGGCAATTCCTCACGGATTTAGAAGTAGAGATACCATTGACCCAG
15 CAATCCCATTACTGGGTGTATACCCATAGGATTATAAATCATTTCTACTATAAAGACACATGCACACTTATG
TTTATTGTAACACTATTTACAATAGCAATGACCTGGAACCAATCCAAAAGCCCATCAATGATAGACTGAAT
AAAGAAAATGTGGCACATATACACTGTGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTCCTTTG
CAGAGACATGGATGAAGCTGGAAACCATCATTTCTCAGCAAAGTAGCACAATAACAGAAAACCAAACACTGC
ATGTTGTCACTCATAAGTGGGAGTTAAACAATGAGAACACATGGACACAGGGAGGGGAACGTCACACACTG
20 GGGCATGTCTGGGAGTGGGGGCCTACGGGAGGGATAGCATTAGCAGAAATACCTAATGTAGGTGACGGGTT
GATGGGTGCAGCAAACCACCATGGCACATATACACCTATGTAATAAACTGCACGTTCTGCACATGTACCC
CAGAACTTAAAGTATAATTAATAATAATAAATTTCTGGGCATGTAAGTAGCTGTCTTTCAGGTTCTACT
TTGATACATATTCTGAGAGAATTAAACCTGTCAAAGAAACCTTGACTTTCAATGGCAGGCACTGGAATTGA
CCCTAATAATGTGTTTTGGGGTAAGCCTACTCATATTCTCAACCTGTCTGCAGTAGTCGTTAGAATCTGAA
25 CTTCTGAAGTTCATGTGCAAAGTTGAGTTAATTGTTTAATATTCAACAAGGATTATGCCAGTAAGATGGT
AGGAAAATATTAGATATGTGTCATCACTGCTGGTATTATTTAAACTGCAACATATTTTAGCTGGCTGCTGA
TCTCAGCCACCATGCCTGCATTTTATCTCTGTCTCGTGGTCTGCAACCTTGGAAGCTTTGAACTTAGCTCA
TAGAATCCTGGGCATCAAGAACATGTGGTTCTAATGGCTAGATAGGGAATGAGAGTAAAGGATTTTGCCC
ACGGTCACGTGAGTAAACAACAGATTTGGAGGGGTCTGGACTACTGTGATGACTTCATTCTGACAATATGT
30 TCCAGTTGTCCTTTCATTTCCCTCCTAATCACATGTCTGGTCTGATCTGGCTGTTTCCCACCTTCCAATTCC
TGCCTTCTCCAATGCTCCCTTCCGTAGGTCACTCTGTGGCTCAGAGACCCTGCTTAGCAAGCGCCCAACCT
TTCAATTATTTGTTTCAAGTAAACTTGAAGTCAATGTCTCCCTTCTTGATAAAAAGAAAATACGTTATGTAA
TGTCGGGTACTCTATAACTCTTGTCTCTCTCGGCAACTACTGAACTAACTGTTTTTCATATTGAGCAA
ACGTTTATGGAAGGACTGCCAAGAGTCAGGTACTAGGCTTGGTAATATTCCCCGTTCTCTCTAGTCAAAGC
35 CAACACCAGCCAGACTTGCAATCTAGGTCCCAAGCCCACTGCAGATCACAGGCCAGGGTCTGGTCTCCTC
TGAGCTCCTTTGGGAGGGAAAGACAGAATTATTAACACCCATTTTGTAGATTAGGCAACTGAGGCTGAGGA
AGTTTAAATAACTCAGACAGGGCCTGCACGTCAGTCATATCCAAGGATCCCTACTCACTGTCTTCTCTCT
ACAGAACGAGATGTCTCTGGAGTCCATAGAAAGCCAGGAGCCTGGCTGGGCACGGTGGCTCCTGCCTGTA
ATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAAC
40 ATGGCAAAACCCCATCTCTACTAAAAATACAAAAAATTAGCTGGGCGTGGTGGTGCATGCCTCTAATCCCA

GCTACTTGGGAGGCTGAGGCACAAGAATTGCTTGAGCCCAGGAGGCAGCAGTTGCAGTGAGCTGAGATTGT
GCCAGTGCCTCCAGCCTGGGCAACAGAGCAAGATTCCATTTCAAAAACAAAACAAAACAAAACAA
ACAAAAATAGAAAGCCCAGGGACCACCTGCGTCAGGTTCCCAGCCACACCTTTTTCTTGTCTCCTCTGTC
TCTGGCATCTTCTCACAGGTTCTTAATTGTTTGTGGTTGCACAAATTCAAATCCCAGAAAAATTACCACT
5 TCACACCCACTCAGATGGCTATTTTTTTTTTTGAAGGAAGATAACAAGTGTTGACAAGAACATGGAGAAATT
GGAATTCTCACCCATTGCTGGTGAGAATGTAATACGGTGCTGCTGCTATGGAAAACAGCTTGGAGTTTCCT
CAAAAAGTTCAACAGAATTTCAATGTGACCCAGCAATTCCTCTAAGTTATAGATCTGAGAGGATTAAAA
ACAGTTACTAAAATACACGGACTCACATATTTCTAACAGTCCAATTCACAAGGGCCAAAAGGTGCTAATAG
CCCACATGTCCATCGATGGATGGATAAATAAATTGTGGTCTATCCATACAATGGAATATTATTTCGGCCATA
10 AATGGAATGAAGTACTGACGCATGCTACAGAATGGATGAACCGCAAAAAAATGGATGAACACATGCTACA
GAATGGATAGCCTCACTTTACTATGAAGTGAAGGCCAGAAACGAAGTCCATATATTGCATCATACAAAATA
TCCAGAAGAGGGAAGCCCACAGAGACAGAATGTGCAATGGTGGATGCCAGGGTCTGGGGAGAGGGGAGAGT
GGGGAGAACTGCTCAACTGGTACAGGCTTTATTTTGAATGATGGGAACATTTTGCAACTAGATAGAGGT
AGTGATTGCAGAACACAGAATGTACTGAATTCCACTGATTTTTTTTACCTTAAAATGGTTAATTTTCAGTC
15 CTGAGATTGGATAATCATAAAAAAATGGTTAATTTTATGTTATGTGAATTTTCATCCCTATACATATTTTAA
ACCTCAGAAATATACACTAGCAGGCATGGAACAGGTCACTGTGGTGCCTGCCAAGCCCGGTGATGTTATCT
GGGGTCCCCGGCCAGCCTTAAGCCTCTTGCTGACCGGTGGAGGGCAGAACCTTTGCCCTAAAAGTATAATA
TCCACATGCTGGCATGATTCCTGGCCAGATGGCTTCTTTATTAGCAGTAATTGAAACTGCCTCGATACAGA
CACTGTACCTTGCAACCAAAAAATGACTCAACAATGATAATAAGGGTTAAGCTGGGCCTTTCTCTCTTTGC
20 CAGTTAAATTATATTTATTATAGCTTGACATGAAAAACAAAGCAACTCCAACAGGTATCACAAGGGCAAAG
GACATGAACATTTTATCAAAGAAGAAATGCAGCTGTCAAAAATACAGAAATATTCAACCTTGTTTCATAATA
AAGTGGCTGGGCTCAGTGGTTCATGCCTGTAATCCCAGTGCTTTGCAAGGCTGAGACAGGAGGATCATTTG
AAGCCAGAAGTTCAAGACCATCCTAGGCAAGTCAGTTCAATACCAGACTTCATGTCTACAAAACATCAAAA
AATTAGCCAGGCATGGTGATGCATGCCTGTTGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTG
25 AGCCTGGGAGGCTGCGGTGGCGGTGAGCCATGATTGTGCCATTGTACTCCAGCCTGGGCAATGCAGCAAGA
CTGTCTAAATAACAAAAATAATAGTAAAGAAAAGGATTGGGATGCCATTTACTTGCGTATTCAATACACAG
AGTTAAAGTAATTTCTACGTTTTCTATTTTTTTTATTACTAAAAAAGCTGGACCATTCTCACAGCCTGAA
ATGCTTCTCACTTTCCCTTCTTCTGTCCAAACACTTCTCTATGATAATGCAAACAGTCACTCCTTTAGGAA
GACTTCACCCCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAACTCCTGGTGTCTCTCCAGTTCC
30 CTCCGTGTGGTGAAGTACCCTACCTAGGGTTTCAGTATGGCTCTGTCTGCAAAGGTCTTGTTACACCTTC
CCTTATGGTTCTGTTGCCCTGTGTTGTGTCATAGCACAGGGCACAGTGGAGAACCCATTACACTGATAGA
GAGGGCCCCATGGTCCTGGAGATAACCATGTAACCGATCAGAATAAGGCATTGAGGGCTGGGTGTCAGGCG
TGGGCTGCACTTGGGTGGGCAGGTCCCCTGGAAAGTCACTGGGTTTGGCAAGCTTCCTAGTAACATGTCTC
TCTGGGGTCCCCCTTGGAACCTTCATGCAAAAATGCTGGTTGCTGGTTTATTCTAGAGAGATGGTTCATTCC
35 TTTCAATTTGATTATCAAAGAACTCATGTCCCAATTAAAGGTCATAAAGCCCAGTTTGTAAGTGAAGATGA
TCTCAGCTGAATGAACTTGCTGACCCTCTGCTTTCCTCCAGCCTCTCGGTGCCCTTGAAATCATGTCGGTT
CAAGCAGCCTCATGAGGCATTACAAAGTTTAATTATTTTCAGTGATTATTAAACCTTGTCCTGTGTTGACCC
CAGGTGAATCACAAGCTGAACTTCTGACAAGAACAAGCTATCATATTCTTTTCAATTACAGAAAAAAGTAA
GTTAATTGATAGGATTTTTTTTTTGTAAATAAATAAGTACTAGTTTTTGAAAAGGTAATATGTGCACATGGT
40 AAACACTAAGAAGGTATAAGAGCATAATGCTTTTATACTACTAAGAATAATGTTTTCTCTAAGTTTTTTTTT

9/22

GGTAGATGCTTTCATCAGATTAAGAAAATTCCCTGCTATTAGTTGTTGAAGGTTTTTATATCATAAATGAA
AGTTGAATATTATTATCATATATTATTAATATATTGTTATTGAACTATCAAAGCCTTTTCCTAAAACCATTT
GAGATGATCTTATAACCATTTCTCCTTTAACCTGTTGACGAGATCATTGGTATTTATACTATTTCTCTGTTA
ACCATTTCTTGAGTCTCAGGTTTAAATTCAACTTGGTCATGGTGTGTCATCTTTGATCATTTGCTGTCTGTGG
5 CTTGCTACTGTTTTGTTTAGGATTTTTGCACTGATGCTCATCAATGAGACTGGCATGCCATCTTCCTTTGC
AGTCCTGATTTTTTTCTGATTTGGATCATGTGGTTATGGCCCTCATGGAATGAGTTGGGCATGATGCCTTT
TTTTCATGTCTCTGGATTGATGGGACACTTTGGATTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTC
ATTGTTAGGCCCTGGGCAAGCCCTTCTCATTTCTGGTAGGCCAGGAACCTGTGGGGGTTTTGTTTGT
GTTTGTCTTCTGAGTCGGAGTCTCACTCTGTCACCCAGGCTGGAGTTGGAGTGCAATGGCCCGATCTTGGC
10 TCACTGCAACCTCCACCTCCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGAATTACAG
GCACCCACCGACACACCTTGCTAATTTTTGTATTTTTAGTACAGATGGGGTTTCACAATATTGGCCAAGCT
GGTCTCGAACTCCTGATCTCATGATCTGCCCGGCTTGGCCTCCCAAAGTGTTGAGATTACAAGCATGAGCC
ACCACACCCAGTGAACCTGTGGTTTTTTAGAAGCTCCCCATGCATGTGAATGCTGTGAGCATCCCAGGATGA
CAGCCACTGTGTGTTTCACTGTTGGAAGTGTGAGAAAGCACCAGTGGGACCTTCTCCAGCACCTGCCTGCT
15 GAGTTCATGGAAGAGGCTTGTGGGGAGATGATGCCCTGGCTGACTCCTGAAGGATGGTTAGGAATGCACC
AGATGGAAGCTGGGTGGACCACTCTATGCTGAAGAACAGCTTGTGTGGACACAAGGAGACACGGATATG
TCATTTTTGTAGAGCCTGAGGAGTGTCCAATCACACCATTTGCTTAAAACATCATGCACACTTGGAAGT
GGACTGAGACCGAATGAAGAAGCTAACAGTGGCCAGATCAGAAAGGGTCTTGTGTTACTTCCTAGAGATAC
TTAGATTTTATCCTGTGGGTGATAGGAGCAGTTGGAGGGACTGAAGACAAGGAAAGAAACATGTTTCAAGA
20 TCTATGTTTTTCAAGACGCTTTTCTGGTGGCTGAGTAGGGAATTCCTTGATAAGTCTGCCAGGGTCAG
GCAAAACAAGTTAGGGGGTTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTGTGCTGACGTTTTGTAA
CACATCTCATGATGATCTTCATTTCCCTTCACTAATTTCCCTGTTTCATTAATTCCTTCCACGTGCTCTTCT
GAAATTTGCCTCACATTCTCTGATTTCTCTTTTACCTGTTGGTTTTCATCACCTTTTACTTTTTGCTTTCCT
GGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTTGCCTTTCTGCTGAAACCATG
25 AGTTCACCTGAATACACAATTTAGTAAAGTGTAGGATGCACATGTCGTTTTCGTGGTCACAACCAGCTCTGT
AGCATTTTATAACTACACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCACATGTGTGGCTGAC
ACAACCTGCCAAGTTATTTTAGGAGCCTCCTTGGAAATCCCAGCAAGAATGCTACCGGCACAATTTGTAATC
ACAGCATCCTGCTCCATGCCTTGGCTTCATGGCATAAGTCACTTCTGCAAGTCTCTTTCCAGCTGTCTGTTT
CCATGTCTATAAAGTATGAGTTAAATCATCCTAACACTACTCATCTTACAAAGTTTTCTTGCTGATGTTAA
30 GAGAGTTGGGAAAGAACTGTATAAACTGTGAAGTGCCATGGAGATGTTAGTGGTTACTTTATCAAGAAATA
GACACTCTAGAATGGAGTAGAAAGCCAACAGTTATGATTGAGTCCTCCTCCTCTTCTTTTTTATTAATT
TATAAAGAAAAGAGGTTTAATTGACTCACAGTTCCATATGGCTGGGGAGGCCTCGGGAAACTCTCAGTCAT
AGCAGGAGGCAAAGGGGAAGAAGGCACCTTCTTCAAGGCGGCAGGAGAGAGAGAGCTCCTGTTCTTTTT
TGTCATAAAGTCTACAGAAGTGCTTATACTTCAGGACAAGGGCAGGCAGAGAGAAGGAAGGACATTGCTTC
35 ACCCCAGCCCTCACTGACGAGTTTGCTAGGGGACCTCACTTTGTCCCAGAGTAGGGCAGAACTCTGGCCAC
TACCCATTGAGAAGGCCTGGGCTGCACTGCTAGTTCCCTCACTAACTCTGTGTGGCCTTGGGCAAGGTTGGG
CCTGTGTTAACAGATTATGACCCTGGGCTCTCAAGCTAGAGGATCTAAATTTGAATCCTGGCTCTGCTAAA
GCAATTAGTGATGTAACTTTAATGGGTCAGTTAACCTTCCTGTGGCTTAGTTTGCTCATCTGTAAAATAG
GGATCATAACAGTATCAATACCACATGATTGTTGGACAGATTGAATCAGTTAATGCAGGGGAAGTACTTAG
40 CATGACACGTATTCATATCATTTTCTGGAGTAAGAGCTGTGTGTGAGTGGGTGTGAGCATGTGTGAAACC

SUBSTITUTE SHEET (RULE 26)

10/22

TTTTCTCTGCAATCTCAGTTAAGAAACCAATCCAGAATTTAAAGTTCAGGGCCTAAATGGGTGGTTATCTT
CTCCCAGTTCCATCCTATCCCACCTTTGCTCTTCCTCCCGCCCACAGGAGCTGTTGGTCCTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGGAGAATAGAGAACAGGGAATGTCTTCAAAAATCTAG
AGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGTCTTCTGATTGGGACAAGGAGAACCTTGGTCTT
5 CACAGGCCAATTCTGGTCAGTTTCCCCCATGGACAGATGAGGAAACAGGCCCAGGAATATCCAAGGTCTCA
CACTTCCCATCTGTCAAGTCTTGTTGATTCTGTTGTATTTCATGTCTCTCAAAGGGAGATAGAGTTTAGGGA
AGAAAGAAGGATCAACTGTGTCTGATACCACTGGGAGCTTAAGTAAAGGGTTCTTTTACTTCATAGCATTT
ATCCCAATTTGTAATTCAGTATTATTTGTGTGGCTGTTTGGTGTCTCTTTCTCCTATATGAGTGCTAGCTT
CATAAGGGCAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGCATACAG
10 GCTAACCAACATACAGTGGCATGAAAGTCATGAAAGTGAGACACCTACCTCCTCCAGTGCCAAGAGAGCAT
AACCATGCACCTGTCACTCTCCTCAACACCACCCCCAAGCATGAGGCCCAAAGCATTAGCTAATCCCCTC
CTCCAGCCACTAAAACTTAAAGGCCAGGTGTGGTGGCTCCCATCTGAAATCCCAGAACTTCAGGAGACAGC
AGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGATCAGCCTGGGCAACATAGCTAGGTCCCATCTGTACTA
AAAATTAGCTGGGCGTTGTTGCATGCCTGTAGTCCCAGCTACTAAGGAGGCTGAGGTGGGAGGATCACTTG
15 AGCCCAGGAGGTGGAACAACAGTAAGCTATAATCACAGCACTGAACTCTAGCCTGGGCAACAGAGTGACA
CCCTGCCTCAAAACAATTTTAAAAATAAATAAGAGCAAACTTAGATACCACGTGGTCACCCCAACATGCA
AAATCAAGTTTTCCCTACTGAGAAGAATGGGGACTTGACAGCTGAGTTACAGAGAGATAATCTTCTTCTT
CTTTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACTTGTGAAATTTGAATCGAATACACATGTAATTC
CAGAGCAATGTTGCCTCCGCATACCATCAGCAATTCACCTGGCTACTGGAAGTCAGGATAAGCTTCCCAGA
20 AGAGAGGTACCACTTGGGCTACCAATATAAAAGGATGAAAATATCAGAGTGATGGTGTTCCTTACAACGTT
GAGTCCCTGGACAGCCTGTCCACTGATGCTGATATCTGAGCCTAATGCTTCTCTGAATGTTGAGATTGAAC
TTTGATCCAATGAACTAGAACGAGAAAGAAGATAAGTCTTTCATTGTTGATAAGGACATTATGTTTCTCA
TACTTGTATGATTATTTTTCTTAGCTGTACTATAATTATCTGCTTATTTGTCTCTGCTCTATGTGCTTAG
GGTACAAAGTTGACCAAGACCAACTTTGGTTGGAAGCATAGTACTAAGAGCACAGTACTGAGAGCACAGTA
25 TTGAGAGCACAGCTTTAAAAACATGATGAAGGCTTTAATACAGGAAATGAGCAGGGGAGAGGCATGTGGT
GGTTGGATGTATCTTCCTTGACACAGTCAGTGCAGCTCTCAGTAGTCAAGTCCCTACATGTTAGAAGATGT
TACCTTCTGTGGAATTAAGTGGCAGAACTTGCCTTCAATTATTTCCCTTGCAGAACAAACCAACTGCAT
TAGTTAGGACACAGTGCTGGCTGCATTTAAGTCCCAAGCGATGATTAGTCTCTCACTGTTGGTATAGATTC
AAACCAATCAGACCACCTCCTAAAGTTTGTAGGGCAGGTAAATCCTCATCTTAGAATAAAAATCATCTTAC
30 CAAGTATGTGTTTTAGAGGCAAGAAGAAAACATATTTGTTTCTGTAAGAGTTTTGTTTAAAAAAAATATAA
GAAAGGCTCTCGGTTTAGGTGAGGTAATGAAGTTGTTGATAGTTATCAGATGACACTGGAATCTTTACTTC
TCTGAACGTGTTCTGTGCATCTCTCAGTGTGGGAACATAGAGAGGGAGATCCTCCAGCAATGCCACTGATA
TGGTCAGAACTGCATCTTTCTTTCTCCCTGCTGAGATGAGATGGAGTCCTTTGTTCTAGAAGACCCATGG
TGGTGCCGCTGGGAGTAACCCTTGAGACAGGAACACAAATCCCAACCAATTTGTGGTTGCAGCCTTGAGTC
35 TCACTATTTCCCATAGTGATGCGTAGCAGGGAATGGCAGGTGCACCAGAGCAGGAGAGGACCTAATATCTC
CCTTCCTGTTAGCTTTTTATAAAGTTTTATTGTGATCAGTAGCAGTTGGGAAGCTACTTGCAGTCACTGAG
CCTCAGTTTCTACATCTGTAAACTGGGGATAGTAGCATGGCCCCCTACTTAATGTGCTCAGCAAAGCCACTG
AAAGGAGACAGAAATGTATCTAAATTACCCTGGACTTTTATCCTACCTCTCTTGGGGATTGTCACCACCTT
CCCATGTTTGTCTTTTTTGGTTTGATGCTTGCTGTCACTTCTTTCCTTAGGTGCCTCTCTGTACGGCTCTT
40 TTATCCCAGGGATTCCAGAGTTACAGCACATGCATACCACCATCCAAGCATGTTTATTTGTCTCCTGCTTC

11/22

ACTAGGCTGTCCCCAAGGAACATGTGGCTCCCGGCACACACCTGGCACAACACTGCACATGACATTCACCC
ACTTGGCCTTGAATCTGACAAGGAATCTGGCATGATGTTCACCCACTCAGGCCAGGTGCCGAGCAGCCCTG
GAGGCTTAGGGGCCAGAGGGATGGGAAAAGGTGTCTTCTGGGGTGAGTATCAGTTTCTGCAGGAGGGCTG
AATGTGAGAAAGAATAAAGAGAGAAGGAAGCGAACAAGCACAGCTTAAACATCGCCTATTTCTATTGAGTT
5 TTAAGAACGCTGTGATTTTGTGTTGTCATGCAATCCATTCATCAGGCCAGGCAGACACAGAACTTGGGTGTG
AGTGACGATAATGAGCTGATATAATTTTCACACCCTCATCACTGAGATCTCTCCCATCAGGAATGGGTCAG
GGAGCTCACAGGTGGCAGCAACTGCTATTACAGGCCTCATCTCTACCAGCTCCTGGGGCCTGCCCTCCTCC
CATTAGAAAATCCTCCACTTGTCAAAAAGGAAGCCATTTGCTTTGAACTCCAATTCCACCCCCAAGAGGCT
GGGACCATCTTACTGGAGTCCTTGATGCTGTGTGACCTGCAGTGACCACTGCCCCATCATTGCTGGCTGAG
10 GTGGTTGGGGTCCATCTGGCTATCTGGGCAGCTGTTCTCTTCTCTCCTTTCTCTCCTGTTTCCAGACATGC
AGTATTTCCAGAGAGAAGGGGCCACTCTTTGGCAAAGAACCTGTCTAACTTGCTATCTATGGCAGGACCTT
TGAAGGGTTCACAGGAAGCAGCACAAATTGATACTATTCCACCAAGCCATCAGCTCCATCTCATCCATGCC
CTGTCTCTCCTTTAGGGGTCCCCTTGCCAACAGAATCACAGAGGACCAGCCTGAAAGTGCAGAGACAGCAG
CTGAGGCACAGCCAAGAGCTCTGGCTGTATTAATGACCTAAGAAGTCACCAGAAAGTCAGAAGGGATGACA
15 TGCAGAGGCCCAGCAATCTCAGCTAAGTCAACTCCACCAGCCTTTCTAGTTGCCCACTGTGTGTACAGCAC
CCTGGTAGGGACCAGAGCCATGACAGGGAATAAGACTAGACTATGCCCTTGAGGAGCTCACCTCTGTTCAG
GGAAACAGGCGTGGAACACAATGGTGGTAAAGAGGAAAGAGGACAATAGGATTGCATGAAGGGGATGGAA
GGTGCCCAGGGGAGGAAATGGTTACATCTGTGTGAGGAGTTTGGTGAGGAAAGACTCTAAGAGAAGGCTCT
GTCTGTCTGGGTTTGGAAGGATGTGTAGGAGTCTTCTAGGGGGCACAGGCACACTCCAGGCATAGGTAAAG
20 ATCTGTAGGTGTGGCTTGTTGGGATGAATTTCAAGTATTTTGGAATGAGGACAGCCATAGAGACAAGGGCA
AGAGAGAGGCGATTTAATAGATTTTATGCCAATGGCTCCACTTGAGTTTCTGATAAGAACCAGAACCCCTT
GGACTCCCCAGTAACATTGATTGAGTTGTTTATGATACCTCATAGAATATGAACTCAAAGGAGGTCAGTGA
GTGGTGTGTGTGTGATTCTTTGCCAACTTCCAAGGTGGAGAAGCCTCTTCCAAGTGCAGGCAGAGCACAGG
TGCCCTGCTACTGGCTGCAGCTCCAGCCCTGCCTCCTTCTCTAGCATATAAACAATCCAACAGCCTCACT
25 GAATCACTGCTGTGCAGGGCAGGAAAGCTCCATGCACATAGCCCAGCAAAGAGCAACACAG

12/22

Figure 6.

CTGGTTCATCTCATTGGGACTGGTTGGACAAGAGGGTGCAGCCACGGAGGGTGAGCCAAAGCAGGGTGGG
GCGTCGCCTCACCTGGGAAGCACAAGGGGTCGTGGAATTTTCTCCCCTACCCAAGGAAAGCCATAAGGGAC
TGAGCCTGAGGAACTGTGCACTCTGGCCCAGATACTGCACTTTTCCCATGGTCTTTGCAACCCGCAGACCA
5 GGAGATTCCCTCCGGTGCCTATGCCACCAGGGCCCTGGGTTTCAAGCACAAAAGTGGGCAGCCATTTGGGC
AGACACCGAACTAGCTGCAGGAGTTTTTTTTTTTTTTTTTCCATACCCCATTTGGCACCTGGAACGCCAGTGA
GACAGAACCGTTCACCTCCCCTGGAAAGGGGGCTGAAACCAGGGATCCAAGTGGTCTGGCTCGGTGGGCCCC
ACCCCATGGAGCCCAGCAAACAAAGATTCACCTGGCTTGAAATTCTTGCTGCCAGCACAGCAGCAGTCTG
AGATTGACCTGGGACCCTCGAACTTGGTTGGGTGCTGTGGGGGGGCATCTTCCATTGCTGAGGCTTGAGTA
10 GGTGGTTTTACCTTCGCGGTGTAAACAAAGCTGCTGGGAAGTTTGAACTGGGTGGAGCTCACCACAGCTCA
GTAAGGCCACTGTGGCCAGACTGCCTCTCTGGATTTCTCCTCTCTGGGAAGGATATCTCTGAAAAAAGGC
AGCAGCCCCAGTCAGGGACTTATAGATGAAACCCCCATCTCCCTGGGACAGAGCCCCTCGGGGAAGAGGTG
GCTTCCACCATTGTGGAAGACTGTGTGGCAATTCCTCACGGATTTAGAACTAGAGATACCATTTGACCCAG
CAATCCCATTACTGGGTGTATACCCATAGGATTATAAATCATTCTACTATAAAGACACATGCACACTTATG
15 TTTATTGTAACACTATTTACAATAGCAATGACCTGGAACCAATCCAAAAGCCCATCAATGATAGACTGAAT
AAAGAAAATGTGGCACATATACACTGTGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTCCTTTG
CAGAGACATGGATGAAGCTGGAAACCATCATTCTCAGCAAAGTAGCACAATAACAGAAAACCAAACACTGC
ATGTTGTCACCTCATAAGTGGGAGTTAAACAATGAGAACACATGGACACAGGGAGGGGAACGTCACACACTG
GGGCATGTCGGGGAGTGGGGGCCTACGGGAGGGATAGCATTAGCAGAAATACCTAATGTAGGTGACGGGTT
20 GATGGGTGCAGCAAACCACCATGGCACATATACACCTATGTAATAAAAGTGCACGTTCTGCACATGTACCC
CAGAACTTAAAGTATAATTAATAATAATAATTTCTGGGCATGTAAGTAGCTGTCTTTCAGGTTCTACT
TTGATACATATTCTGAGAGAATTAAACCTGTCAAAGAAACCTTGACTTTCAATGGCAGGCACTGGAATTGA
CCCTAATAATGTGTTTTGGGGTAAGCCTACTCATATTCTCAACCTGTCTGCAGTAGTCGTTAGAATCTGAA
CTTCCTGAAGTTCATGTGCAAAGTTGAGTTAATTGTTTTAATATTCAACAAGGATTATGCCAGTAAGATGGT
25 AGGAAAATATTAGATATGTGTCATCACTGCTGGTATTATTTAAAGTGCACATATTTTAGCTGGCTGCTGA
TCTCAGCCACCATGCCTGCATTTTATCTCTGTCTCGTGGTCTGCAACCTTGGAAGCTTTGAACTTAGCTCA
TAGAATCCTGGGCATCAAGAACATGTGGTTCTAATGGCTAGATAGGGAATGAGAGTAAAAGGATTTTGCCC
ACGGTCACGTGAGTAAACAACAGATTTGGAGGGGTCTGGACTACTGTGATGACTTCATTCTGACAATATGT
TCCAGTTGTCCTTTTCATTTCCCTCCTAATCACATGTCTGGTCTGATCTGGCTGTTTCCCACCTTCCAATTCC
30 TGCCTTCTCCAATGCTCCCTTCCGTAGGTCACTCTGTGGCTCAGAGACCCTGCTTAGCAAGCGCCCAACCT
TTCAATTATTTGTTTCAGTAAAGTGAAGTCAATGTCTCCCTTCTTGATAAAAAGAAAATACGTTATGTAA
TGTCGGGTACTCTATAACTCTTGTCCTGTCTCTCGGCAACTACTGAACTAACTGTTTTTCATATTGAGCAA
ACGTTTATGGAAGGACTGCCAAGAGTCAGGTACTAGGCTTGGTAATATCCCCGTTCTCTCTAGTCAAAGC
CAACACCAGCCAGACTTGCAGATCTAGGTCCCAAGCCCACTGCAGATCACAGGCCAGGGTCTGGTCTCCTC
35 TGAGCTCCTTTGGGAGGGAAAGACAGAATTATTAACACCCATTTTGTAGATTAGGCAACTGAGGCTGAGGA
AGTTTAAATAACTCAGACAGGGCCTGCACGTGAGTCATATTCCAAGGATCCCTACTCACTGTCTTCTCTCT
ACAGAACGAGATGTCTCTGGAGTCCATAGAAAGCCAGGAGCCTGGCTGGGCACGGTGGCTCCTGCCTGTA
ATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAAC
ATGGCAAAACCCCATCTCTACTAAAAATACAAAAAATTAGCTGGGCGTGGTGGTGCATGCCTCTAATCCCA
40 GCTACTTGGGAGGCTGAGGCACAAGAATTGCTTGAGCCCAGGAGGCAGCAGTTGCAGTGAGCTGAGATTGT

13/22

GCCAGTGC ACTCCAGCCTGGGCAACAGAGCAAGATTCCATTTCAAAAACAAAAACAAACACAAACAAACAA
ACAAAAATAGAAAGCCCAGGGACCACCTGCGTCAGGTTCCCAGCCACACCTTTTTCTTGTCCTCCTCTGTC
TCTGGCATCTTCTCACAGGTTCTTAATTGTTTGTGGTTGCACAAATTCAAATCCCAGAAAAATTACCACT
TCACACCCACTCAGATGGCTATTTTTTTTTTTGAAGGAAGATAACAAGTGTTGACAAGAACATGGAGAAATT
5 GGAATTCTCACCCATTGCTGGTGAGAATGTAATACGGTGCTGCTGCTATGGAAAACAGCTTGGAGTTTCCT
CAAAAAGTTCAACAGAATTTCAATGTGACCCAGCAATTCCCCCTCTAAGTTATAGATCTGAGAGGATTAAAA
ACAGTTACTAAAATACACGGACTCACATATTTCTAACAGTCCAATTCACAAGGGCCAAAAGGTGCTAATAG
CCCACATGTCCATCGATGGATGGATAAATAAATTGTGGTCTATCCATACAATGGAATATTATTTCGGCCATA
AATGGAATGAAGTACTGACGCATGCTACAGAATGGATGAACCGCAAAAAAATGGATGAACACATGCTACA
10 GAATGGATAGCCTCACTTTACTATGAAGTGAAGGCCAGAAACGAAGTCCATATATTGCATCATACAAAATA
TCCAGAAGAGGGAAGCCCACAGAGACAGAATGTGCAATGGTGGATGCCAGGGTCTGGGGAGAGGGGAGAGT
GGGGAGAACTGCTCAACTGGTACAGGCTTTATTTTGAATGATGGGAACATTTTGCAACTAGATAGAGGT
AGTGATTGCAGAACACAGAATGTACTGAATTCCTGATTTTTTTTTCACCTTAAAATGGTTAATTTTCAGTC
CTGAGATTGGATAATCATAAAAAAATGGTTAATTTTATGTTATGTGAATTTTCATCCCTATACATATTTTAA
15 ACCTCAGAAATATACACTAGCAGGCATGGAACAGGTCACTGTGGTGCCTGCCAAGCCCGGTGATGTTATCT
GGGGTCCCCGGCCAGCCTTAAGCCTCTTGCTGACCGGTGGAGGGCAGAACCTTTGCCCTAAAAGTATAATA
TCCACATGCTGGCATGATTCCTGGCCAGATGGCTTCTTTATTAGCAGTAATTGAAACTGCCTCGATACAGA
CACTGTACCTTGCAACCAAAAAATGACTCAACAATGATAATAAGGGTTAAGCTGGGCCTTTCTCTCTTTGC
CAGTTAAATTATATTTTATTATAGCTTGACATGAAAAACAAAGCAACTCCAACAGGTATCACAAGGGCAAAG
20 GACATGAACATTTTATCAAAGAAGAAATGCAGCTGTCAAAAATACAGAAATATTCAACCTTGTTTATAATA
AAGTGGCTGGGCTCAGTGGTTCATGCCTGTAATCCCAGTGCTTTGCAAGGCTGAGACAGGAGGATCATTTG
AAGCCAGAAGTTCAAGACCATCCTAGGCAAGTCAGTTCAATACCAGACTTCATGTCTACAAAACATCAAAA
AATTAGCCAGGCATGGTGATGCATGCCTGTTGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTG
AGCCTGGGAGGCTGCGGTGGCGGTGAGCCATGATTGTGCCATTGTACTCCAGCCTGGGCAATGCAGCAAGA
25 CTGTCTAAATAACAAAAATAATAGTAAAGAAAAGGATTGGGATGCCATTTACTTGCGTATTCAATACACAG
AGTTAAAAGTAATTTCTACGTTTTCTATTTTTTTTATTACTAAAAAAGCTGGACCATTTCTCACAGCCTGAA
ATGCTTCTCACTTTCCCTTCTTCTGTCCAAACACTTCTCTATGATAATGCAAACAGTCACTCCTTTAGGAA
GACTTCACCCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAACTCCTGGTGTCTCTCCAGTTCC
CTCCGTGTGGTGAAGTACCCTACCTAGGGTTTCAGTATGGCTCTGTCTGCAAAGGTCTTGTTTACACCTTC
30 CCTTATGGTTCTGTTGCCCTGTGTTGTGTCATAGCACAGGGCACAGTGGAGAACCCATTCACACTGATAGA
GAGGGCCCCATGGTCCTGGAGATAACCATGTAACCGATCAGAATAAGGCATTGAGGGCTGGGTGTCAGGCG
TGGGCTGCACTTGGGTGGGCAGGTCCCCTGGAAAGTCACTGGGTTTGGCAAGCTTCCTAGTAACATGTCTC
TCTGGGGTCCCCCTTGGAACCTTCATGCAAAAATGCTGGTTGCTGGTTTATTCTAGAGAGATGGTTTCATTCC
TTTCATTTGATTATCAAAGAACTCATGTCCCAATTAAAGGTCATAAAGCCCAGTTTGTAAGTGAAGATGA
35 TCTCAGCTGAATGAACCTTGCTGACCCTCTGCTTTCCCTCCAGCCTCTCGGTGCCCTTGAAATCATGTCGGTT
CAAGCAGCCTCATGAGGCATTACAAAGTTTAATTATTTTCACTGATTATTAAACCTTGTCCTGTGTTGACCC
CAGGTGAATCACAAGCTGAACTTCTGACAAGAACAAAGCTATCATATTTCTTTTCAATTACAGAAAAAAGTAA
GTTAATTGATAGGATTTTTTTTTTGTTTAAAAAATGTTACTAGTTTTTGAAAAGGTAATATGTGCACATGGT
AAACACTAAGAAGGTATAAGAGCATAATGCTTTTATACTACTAAGAATAATGTTTTCTCTAAGTTTTTTTT
40 GGTAGATGCTTTCATCAGATTAAGAAAATTCCTGCTATTAGTTGTTGAAGGTTTTTATATCATAAATGAA

14/22

AGTTGAATATTATTATCATATATTATTAATATATTGTTATTGAACTATCAAAGCCTTTTCCTAAAACCATT
GAGATGATCTTATAACCATTCTCCTTTAACCTGTTGACGAGATCATTGGTATTTATACTATTTCTCTGTTA
ACCATTCCTGAGTCTCAGGTTTAAATTCAACTTGGTCATGGTGTGTCATCTTTGATCATTGCTGTCTGTGG
CTTGCTACTGTTTTGTTTAGGATTTTTGCACTGATGCTCATCAATGAGACTGGCATGCCATCTTCCTTTGC
5 AGTCCTGATTTTTTTCTGATTTGGATCATGTGGTTATGGCCCTCATGGAATGAGTTGGGCATGATGCCTTT
TTTTCATGTCTCTGGATTGATGGGACACTTTGGATTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTC
ATTGTTAGGCCCTGGGCAAGCCCTTCTCATTTCTGGTAGGCCCAGGAACCTGTGGGGGTTTTGTTTGT
GTTTGTCTCTTGAGTCGGAGTCTCACTCTGTCACCCAGGCTGGAGTTGGAGTGCAATGGCCCGATCTTGGC
TCACTGCAACCTCCACCTCCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGAATTACAG
10 GCACCCACCGACACACCCTGCTAATTTTTGTATTTTTTAGTACAGATGGGGTTTCACAATATTGGCCAAGCT
GGTCTCGAACTCCTGATCTCATGATCTGCCCCGGCTTGGCCTCCCAAAGTGTTGAGATTACAAGCATGAGCC
ACCACACCCAGTGAACCTGTGGTTTTTAGAAGCTCCCCATGCATGTGAATGCTGTGAGCATCCCAGGATGA
CAGCCACTGTGTGTTTCAGCTGTTGGAACCTGTGAGAAAGCACCAGTGGGACCTTCTCCAGCACCTGCCTGCT
GAGTTCATGGAAGAGGCTTGTGGGGAGATGATGCCCTGGCTGACTCCTGAAGGATGGTTAGGAATGCACC
15 AGATGGAAGCTGGGTTGGACCCACTCTATGCTGAAGAACAGCTTGTGTGGACACAAGGAGACACGGATATG
TCATTTTTGTAGAGCCTGAGGAGTGTCCAATCACACCATTGTCTTAAAACATCATGCACACTTGGAAAAGT
GGACTGAGACCGAATGAAGAAGCTAACAGTGGCCAGATCAGAAAGGGTCTTGTGTTACTTCCTAGAGATAC
TTAGATTTTATCCTGTGGGTGATAGGAGCAGTTGGAGGGACTGAAGACAAGGAAAGAAACATGTTTCAAGA
TCTATGTTTTTCAAGACGCTTTTCTGGTGGCTGAGTAGGGAATCCCTGGATAAGTCCTGCCCAGGGTCAG
20 GCAAAACAAGTTAGGGGGTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTGTGCTGACGTTTTGTAA
CACATCTCATGATGATCTTCATTTCCCTTCACTAATTTCCCTGTTTCATTAATTCCTTCCACGTGCTCTTCT
GAAATTTGCCTCACATTCTCTGATTTCTCTTTTACCTGTTGGTTTCATCACCTTTTACTTTTTGCTTTTCT
GGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTTGCCTTTCTGCTGAAACCATG
AGTTCACTGAATACACAATTTAGTAAAGTGTTAGGATGCACATGTCGTTTTTCGTGGTCACAACCAGCTCTGT
25 AGCATTTTATAACTA~~C~~ACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCACATGTGTGGCTGAC
ACAACCTGCCAAGTTATTTTAGGAGCCTCCTTGGAATCCCAGCAAGAATGCTACCGGCACAATTTGTAATC
ACAGCATCCTGCTCCATGCCTTGGCTTCATGGCATAGTCACTTCTGCAAGTCTCTTCCAGCTGTCTGTTC
CCATGTCTATAAAGTATGAGTTAAATCATCCTAACACTACTCATCTTACAAAGTTTTCTTGCTGATGTTAA
GAGAGTTGGGAAAGAACTGTATAAACTGTGAAGTGCCATGGAGATGTTAGTGGTTACTTTATCAAGAAATA
30 GACACTCTAGAATGGAGTAGAAAGCCAACAGTTATGATTGAGTCCTCCTCCTCTTCTTTTATTAATT
TATAAAGAAAAGAGGTTTAATTGACTCACAGTTCCATATGGCTGGGGAGGCCTCGGGAACTCTCAGTCAT
AGCAGGAGGCAAAGGGGAAGAAGGCACCTTCTTCACAAGGCGGCAGGAGAGAGAGAGCTCCTGTTCTTTTT
TGTCATAAAGTCTACAGAAGTGCTTATACTTCAGGACAAGGGCAGGCAGAGAGAAGGAAGGACATTGCTTC
ACCCAGCCCTCACTGACGAGTTTGCTAGGGGACCTCACTTTGTCCCAGAGTAGGGCAGAACTCTGGCCAC
35 TACCCATTGAGAAGGCCTGGGCTGCACTGCTAGTTCCTCACTAACTCTGTGTGGCCTTGGGCAAGGTTGGG
CCTGTGTTAACAGATTATGACCCTGGGCTCTCAAGCTAGAGGATCTAAATTTGAATCCTGGCTCTGCTAAA
GCAATTAGTGATGTAACTTTAATGGGTGAGTTAACCTTCTGTGGCTTAGTTTGCTCATCTGTAAAATAG
GGATCATAACAGTATCAATACCACATGATTGTTGGACAGATTGAATCAGTTAATGCAGGGGAAGTACTTAG
CATGACACGTATTCATATCATTTCTGGAGTAAGAGCTGTGTGTGAGTGGGTGTGAGCATGTGTGAAACC
40 TTTTCTCTGCAATCTCAGTTAAGAAACCAATCCAGAATTTAAAGTTCAGGGCCTAAATGGGTGGTTATCTT

15/22

CTCCCAGTTCCATCCTATCCCACCTTTGCTCTTCCTCCCGCCACAGGAGCTGTTGGTCCTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGGAGAATAGAGAACAGGGAATGTCTTCAAAAATCTAG
AGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGTCTTCTGATTGGGACAAGGAGAACCTTGGTCTT
CACAGGCCAATTCTGGTCAGTTTCCCCCATGGACAGATGAGGAAACAGGCCAGGAATATCCAAGGTCTCA
5 CACTTCCCATCTGTCAAGTCTTGTGATTCTGTTGTATTTCATGTCTCTCAAAGGGAGATAGAGTTTAGGGA
AGAAAGAAGGATCAACTGTGTCTGATACCACTGGGAGCTTAAGTAAAGGGTTCTTTTACTTCATAGCATTT
ATCCCAATTTGTAATTCAGTATTATTTGTGTGGCTGTTTGGTGTCTCTTTCTCCTATATGAGTGCTAGCTT
CATAAGGGCAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGCATACAG
GCTAACCAACATACAGTGGCATGAAAGTCATGAAAGTGAGACACCTACCTCCTCCAGTGCCAAGAGAGCAT
10 AACCATGCACCTGTCACTCTCCTCAACACCACCCCCAAGCATGAGGCCCAAAGCATTAGCTAATCCCCTC
CTCCAGCCACTAAAACCTTAAAGGCCAGGTGTGGTGGCTCCCATCTGAAATCCCAGAACTTCAGGAGACAGC
AGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGATCAGCCTGGGCAACATAGCTAGGTCCCATCTGTACTA
AAAATTAGCTGGGCGTTGTTGCATGCCTGTAGTCCCAGCTACTAAGGAGGCTGAGGTGGGAGGATCACTTG
AGCCCAGGAGGTGGAAACAACAGTAAGCTATAATCACAGCACTGAACTCTAGCCTGGGCAACAGAGTGACA
15 CCCTGCCTCAAAACAATTTTAAAAATAAATAAGAGCAAACTTAGATACCACGTGGTCACCCCAACATGCA
AAATCAAGTTTTTCCCCTACTGAGAAGAATGGGGACTTGACAGCTGAGTTACAGAGAGATAATCTTCTTCTT
CTTTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACTTGTGAAATTTGAATCGAATACACATGTAATTC
CAGAGCAATGTTGCCTCCGCATACCATCAGCAATTCACCTTGGCTACTGGAAGTCAGGAT

20

16/22

Figure 7.

TCTAGAGAGA TGGTTCATTC CTTTCATTTG ATTATCAAAG AAACATCATGT CCCAATTAAA
GGTCATAAAG CCCAGTTTGT AAACATGAGAT GATCTCAGCT GAATGAACTT GCTGACCCCTC
TGCTTTCCTC CAGCCTCTCG GTGCCCTTGA AATCATGTCTG GTTCAAGCAG CCTCATGAGG
5 CATTACAAAG TTTAATTATT TCAGTGATTA TTAAACCTTG TCCTGTGTTG ACCCCAGGTG
AATCACAAGC TGAACCTCTG ACAAGAACAA GCTATCATAT TCTTTTCAAT TACAGAAAAA
AGTAAGTTAA TTGATAGGAT TTTTTTTGTT TAAAAAAAAT GTTACTAGTT TTTGAAAAGG
TAATATGTTG CACATGGTAA AACTAAGAA GGTATAAGAG CATAATGCTT TTATACTACT
AAGAATAATG TTTTCTCTAA GTTTTTTTTG GTAGATGCTT TCATCAGATT AAGAAAATTC
10 CCTGCTATTA GTTGTTGAAG GTTTTTTATAT CATAAATGAA AGTTGAATAT TATTATCATA
TATTATTAAT ATATTGTTAT TGAACATCA AAGCCTTTTC CTAAAACCAT TGAGATGATC
TTATAACCAT TCTCCTTTAA CCTGTTGACG AG

17/22

Figure 8.

GGATCCAGTTTCAGCTTTCTACATATGGCTAGCCAGTTTCCCAGCACCATTATTATTAATAGGGAATCCTT
TCCCCATTGCTTGTGTTTTGTCAGGTTTGTCAAAGATCAGATGGTTGTAGATGTGTGGTGTGTTGTTCTGAGG
5 CCTCTGTTCTGTTCCATTGGTCCATATCCCTGTTTTGGTACTAGTACCATGCTCTTTTGGTTACTGTAGCC
TTGTAGTATAGTTTGAAGTCAGGTAGCGTGATTCCCTCCAGCTTTGCTCTTTTTTGCTTAGGATTGTCTTGGG
AATGTGGGCTCTTTTTTGGTTCCATATGAAATTTAAAGTAGTTTTTTTTTCCAATTCTATGAAGAAAGTCAT
TGGTAACTTGATGGGGATGGCATTGAATCTATAAATTACCTTGGGAAGTATGGCCATTTTCACGATATTGA
TTCTTCCTATCCATGAGCATGGAACATTCTTCCATTTGTTTGTGTCTCTTTGATTTTGTGAGCAGTGGT
10 TTGTAGTTCTCCTTGAAGAAGTCCTTCACCTCCCTTTAATTTGGATTACTAGATATTTTATTCTCTTAGTA
ACAATTGCAAATGGGAGTTCACTCATGATTTGGCTCTCTTTCTGTTATTGGTGTATAGGAATGCTTGTGAT
TTTTGCGCATTAATTTTGTATCCTGAGACTTTGCTGAAGTTGCTTATCAGCTTAAAGGATTTTGGGCTGA
GACGATGGGGTTTTCTAAATATACAATCATGGCATCTGCAAACAGGAACAATTTGACTTCCTCTTTTCCTA
ATTGAATACCCTTTATTTCTTTTTCTTGCTGATTTGCCCTGGCCAGAACTTCCAATACTATGTTGAATAAG
15 AGTCATGAGTGAGGGCATCGTTGTCTTGTGCTGGTTTCAAAGTTTTTGCCCATTCAGTATGATTTTGGCTG
TGGTTTTGCCATAAATAGCTCTTATTATTTTGAGATACGTTCCACCAATACCTACTTTATTGAGAGTTTTT
AGCAGGAAGGGCTGTTGAATTTTGTGCAAGGCCTTTTCTACATCTATTGAGACAATTATGTGGTTTTTTAA
TCGTTGATTCTGTTTATGTGATGGATTACATTTATTAATTTGCATATGTTGAACCAGCCTTGCATCCCAGG
GATGAAGCCCACTTGATTGTAGTGGATAAGCTTTTTGATGTGCTGCTGGATTTCAGTTTGCCAGTATTTTAT
20 TGAGGATTTTGGCATCAATGTTTCATCAGGGATATTGGTCTAAAATTCTCTTTTTTTGTTGTGTCTCTGCCA
GGCTTTGGTATCAGGATGATGCAGGCCTCAGAACTGAGTTAGGGAGGATTCCCTCATTTTCTATTGATTG
GAATAGTTTCAGAAAGAATGGTACCAGCTACTCTTTGTACCTCTGGTAGAATTCAGCTGTGAATCCATCTG
GTCCTGGACTTTTTGGTTGGTAGGCTATTAATTATTGCCTCAATTTTAGGGCCTGTTATTGGTCTATTTCAG
ACATTCAACTTCTTCCCGGTTTGGTCTTGGGAGGGTTTATGTGTCCAGGAATTTATCCATTTCTTCTAGAT
25 TTTCTAGTTTATTTGTGTAGAGGTGTTTATAGTATTGTCTGATGGTAGTTTGTATTTCTGTGAGATCGGTG
GTGATATCCCCCTTATCATTTTTTTTATTGCATCTATTTAATTCTTCTCTCTTTTCTTCTTTATTATTCTGGC
TGGCGGTCTGTCAATTTTTTTTGATCTTTTCAAAAACCAGCTCCTGGGTTCAGTGATTATTTGAAGGGTT
TTTTGTGTCTCTATTTCTTTCAGTTCTCCTGTGATCTTAGTTATTTCTTGCCCTTCTGCTAGCTTTTGAATG
TGTTTGCTCTTCTTCTCTAGTTCTTTGAATTGTGATGTTACAGTGTGATTTTAGATCTTTCCTGCTTTC
30 TCTTGTGGTCAATTTAGTGCTATAAATTTCCCTCTACACATTGGTTTACATGTGTCTCAGAGATTCTGGTAT
GTTGTGTCTTTGTTCTCATTTCAATTTCAAGAACATCTTACTTCTGCCTTCATTTTGTATTGTTGCCAGTAG
TCATTTCAGGAGCAGGTTGTTTCAGTCTTCATGTAGTTGTGTGGTTTTGAGTGAGTTTCTTAATCCTGAGTTC
TAATTTGATTGCACTGTGTCTGAGAGACAGTTTGTGTGATTTCCATTCCTTTTACATTTACTGAGCATGC
TTTATGTCCCATTATGTGGTCAATTTTAGAATAAGTGTGATGTGATGCTGAGAAGAATGTATATTCTGTTG
35 ATTTGGGGTGTGGAGTTCTGTAGATGTCTATTTCAGTCCACTGGGTGCAGAGCTGAGTGGACATGAACATTT
TATCAAAGAAGAAACACAGCTATCAAAAATCCAGAAATATTGAACCTTGTTAATAATAAAGTGGCTGGCCT
CTGGTTCATTCTCTGTAATCTCAGTCCCTTTGAAAGGCTGAGAAAGGAGGATCACTTGAGGCCACAAGTTCAA
GACCATCCTAGACAAGTCAGTTCAAGACCAGACTTCATGTCTACAAAACATCAAAAATTAGCCAGGCATG
GTGATGCATGCCTGTCTATCCCAGCTACTCAGGAGGCTGAGGCAGGAGGATTGCTTGAGCCTGGGAGATTGA
40 AGTGGCAGTGAGCCATGATTGTGCCATTGCACCTCCAGCCTGGGCAATGCATCAAGACTCTGTCTAAACAAT
AATAATAATAAGTAATAAGTAATAATAATAATAATAAAGAAACGGTTGGGACGCCATTCTTACTTATT
CAATACACAAAGTTAAAAGCAATTTCTACTTTCTCTATTTTTTTTATTACTAAAAAAGCTGAACCATTCTC
ACAGCCTGAAATGCTTCTCACCTTCCCCCTCTTCTATACAAACACTTCTCTGTTGATGATAATGCAGACAGT
CTCTCCTTTAGGAATACTTCAACACCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAGCTCCTGGT
45 GTCTCCCCAGTTCCCTCTGTGTGGTGAAGTACCCCCACCTTGGGTCTCAGCATGACTCGTTCTTTGAAGGT
CTTGTTTCACATTTTCCCTTATGGTTCTGTTCCTCTGTGTTGTGTCTACAGCACTGGGCAGAGTGGACAACCC
ATTACACCCGATAGAGAGGGCCCCATGGTTCTGGAGATAACCATGTAACCTGATCAGAATAGGGCATTGAGG
GCTGGGTGTGAGGCATGGGCTGCACTTGGGTGGGCAGGCCCCCTGGAAAGTCAAGGATTTGGCAAGCTTC
CTAGTAACATCTCTCCCTGGGGTCTCTTGGAACTTCATGCCCGATGCTGGATGCTGGTTTATTCTCGAGA
50 GATGGTTTCATTCCAATAATCAATGAACTCATGTCCCAACTAAAGTTCATAAACTCCAGTTTGTAAGTGA
GATAATCTCAGCTGAATGAACCTTGCTGACCCTCTGCTTTCCCCCAGCCTCTCAGTGCCCTTGAAATCATGT
CAGTTCAAGCAGCCCCATGAGGCATTACAATGTTTAGTTATTTTCAGTGTTTATTAAACCTTGCCCTATGCT
GACCCCAAGTGAATCACAAGCTGGACTTCTGACAAGGACAAGCTATGATATTTCTTTTCAATTACAGAAAA
GTAAGTTAACTGATAGGATTTTTTAAAGATGTTACTAGTTTTTGAAAGGTAAATTTGTGCACATGGTAAACA
55 AGAAGGTATAAGAGGATAATGCTTTTTTACTGCTGAGAATAATGTTTTCTCTAATTTTTTTTGGTAGATGC
TTTCATCAGATTAATAAAATTCAGTGTCTGTTAGTTGTTGAAGGTTTTTTTATATCATGAATGGGAGTTGAAT
ATTATCATGTATTATTAATATATTATTGAAGTACAAAGGCTCTTCTTAAACAATTGAGATGATCTT
ATAATCGTTCTCCTTTAATCTGTTGATGAGATCATTTGGTATTTATACTTTTTCTCTGTTAACTATTCTTGA
GTCTCAGGTTTTAAATTCAACTTGGTCATGGTGTATCATCTTTGAACACTCCTGTCTCTGGCTTGCTACTAT

18/22

TGTGTTTCAGCATTTTTGCACTGATGCCGATGAATGAGACTGGCATGTCATCTTCCTTTGCGGTCCTGATTT
TTTTTCAGATTTGGATCATGTGGCCCTCATTGAATGAGTTGGGTGTGATGCCCTTCTTTTTTCATGTATCTGGA
TTGATGGGACACTTTGGAGTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTCATTGTTAGGCTCCTAG
GCAACCCTTTCTCATTTCTGGTAGGCCCAGGAACCTGTGGGTTTTATGTTTGTGTTGTTTGTGTTTGTGTTT
5 GTTTTTTGGAGTTGGAGTCCTGCTTTGTCTCCAGGCTGGGGTTGGAGTGCAATGGCCTGATCTCGGCCCCAC
TGCAACCTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTTCTGTGTAGCTGGGATTACAGGCAT
CCACCACCACTCCTGGCTAATTTTTTGTATTTTTTAGTAGAGACGGGGTTTTACAATATAGGCCATTGTGATC
TCTTGACAGGCTAGTCTCAAATTCCTGACCTCATGATCTGCCTGCCTCAGCCTCCCAAAGTGCTGAGATT
ACAGTTTTGTGCCTCCACACACAGTGAATCTGTGGTTTTTAAAAGCTCCTCATGCATGTGAATTCTGTGAG
10 CATCCCGGGATGACAGCCACTGTGTGTCCAGCTGTTAAACTGTGAGAAAGCACCAGCGGGACCCTCTCCA
GCATTTGCTTGCTGTGGTCATGAAAGAGGCTTGTGGGGAGATGATGCCCTGGTTGACTCCTGAAGGATGG
TTAGGAATGCACCAGATGGAAGCTGGGTTGGACCCAGTCTATGCTAAAGAACAGCTTGTGTGGACACAAGG
AGACACGAACACATCATTTTTGCGAGAGCCTGGGGAGTAGCCAATCGCACCATTTGCTTAAAACACCGTGTA
CAGTTGGAGAAGTGGACTGAGACAGGCTGAAGAAGCTAACAGTGGCCAGATGAGAAAGGGTCTTGTGTTAC
15 TTCCTAGATATACTTAGATTTTTATCCTGTGAGTGATAGGAACAGTTGCAGGGACTGAAGCCAAGGAAGCAT
GCTTTAAGATTCCATGTTTTTTGAGATGCTGTCTGGTGGCTGAGTAGGGAATTCCTTGATAAGTACTGCC
CAGGGTAGGCAAAAGAAGCTAGGAGGTTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTTTGCTGATG
TTTTGTAACACATCTCATGACAATCTTCATTTCCCTTCACCAATTTCCCTGTTTCATTAATCCCTTCCACGT
GCTCTTCTGAAATTTGCCTCATATTCTTTGATTTCTCTTTTACATGTTGGTTTCATCACCTTTTACTTTTT
20 GCTTTCCTGGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTCCCCTTTCTGCTG
AAACATGAGCTCACTGAATACACAATTTAGTAAAGTGTAGGATGCACATGTTGTTTTCATGGTCATAACCA
GCTCTGTAGCATTTTTATACTACACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCTCATGTGT
GGCTGACACAACCTGCCAAGTTGTTTTAGGAGCCTTCTTGAATCCCAGCAAGAACACCACTGATGCAATT
TGAAATCACAATGTCCTGCTCCATGCCCTGGCTTCATGGCTTAGTCACGTCTGAAGTCTATTTCTAACTAT
25 CTGTTTCCACATCTATAAAGTATGAGTTAAATCATCCTAATACTACTCATCTTACAAAGTTTTCTTGCTGA
TATTAGGAGAGTTGGGAAAGAAGTGTATAAATTATGAAGTGCCATGGAGATGTTGGTGGTTACTTTATCAA
GAAATAGACACTCCAGAATAGAGTAGAAAGAAAACAGTTATGATTAAAGTCTCCTCCTCTTCTTTTTTTTT
AATTTACAAAGAAAGGTTTAATTGAGTCACAGTTCCATATGGTTGGGGAGGCTCAGAAAACCTGCAATCAT
GGCAGTTGGCAAAGTGGAAGAAGGCACCTTCTTCACAAGGTGGCAGGAGAGAGAGAGCTCCTCTTCTTTTT
30 TGTTGTAAAGTCTACAGAAGTGCATATACTTCAGGGCAAGGGCAGGCAGGGAGAAGAAAGGACATTGCTTC
ACCCAGTCTCACTGACAAGTTTGCTTTGGGACTTCATTTTGTCCCAGCATATGGGACAGAGCTCTGGCC
ACTACCCATTGAGAAGGCCTGAGCTGCATTGCTAGTTCCCCACTAACTCTGTGTGTCTTGGGCAAGGCTG
GGCTTATGTCAAAGATTATGACCCTGGGCTCTCCAGCTACAGAATCTACATATGAATCCTGGCTCTGCTA
GAGCAATTAGTGACGTAACCTTGGATGGGTGAGTTAACCTTCCTGTGGCTTAGTTTGTCTCATCTGTAAAT
35 AGGGATCATAACAACATCAATACCATGGGTTGTTAGACAGATTGAATCAGTTAATGCAGGGTAAATACTTA
GCATGACACGTATTCATCTATCATTTCTTGTAGTAAAGCTGAGTGTGAGTGGGTGTGAGAATGTGTGAAAC
CCTTTCATGCAATCTCAGTTAAGAAACCCATCCATAATTTAAAGTTTCAGGGCCTAAATGGGTGGTTATCT
TCTCCCAGTTGCATCCTATCCCACCTTTGCTCTTCTCCTGCCCCGTAGGAGCTGTTGGTCTTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGAATGAGAATAGAGGACAGGGAATGTCTTCAAACTC
40 CTAGAGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGGTCTTCTGATTGGGACAAGGAGGACCTTG
GTCTTCATAGGCCAATTCTGGTCAATTTCCCCCATGGACAGATGAGGAAACAGATCCAGGAATATCCAAGG
TCTCACACTTCCCATCTGTCAAGTCTTGTTGATTCTGTTGTATTTCATGTCTTTCAAAGAGAGAGAGAGTTT
AAGGAAAGAAAGAAGGATCAACTGTGTCTGATATCACTGGGAGCTTAAGTAAAGGGTCTTTTACTTCATA
GCATTTTTTCCAATTTGTAATTCAGTATTATTTTTGTCACTGTTTAGTATCTCTTTGTCTTATTAGAGAGA
45 TAGCTTCATCAGGACAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGC
ATACAGACTAACCAACATACAGTGGCATCGAAGTGAGACACCTACCTCCTCCAGTGCCTAGAGTACATGTC
CATGGACCTGTCACTCTCCTCAACACCACCCCTAAGCATGAGGCCCCGAAAGCATTGCTAATCCCCTCCTCC
AGCCACCAAAACTTAAAGGCCAGGTGTGGTGGCTCCTATCTGAAATCTCAGAACTTTAGGAGACAGCAGCA
GGAGGATCACTTGAGGCCAGGAATTTGAGACGAGCCTGGGCAACATAGCTAGACACCATCTGTACTAAAAA
50 TTAGCTGGGCATGGTGGTATACCTGTAGTACCAGCTACTAAGGAGGCTGAGGTAGGAGGATCACTTGAACC
CAGGAGGTGGAAGCTACAGTGAGCTATAACCACAGCACTGAACTCCAGCCTGAGCAACAGAGTGAGACCCT
GCCTCAAAACAATTTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
AAATCAAGTTTTCCCCTACTGAGAAGAATGGGGACTTGAGAGCTGAGTTACAGAGAGATAATCTGCCTTTT
TTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACCTGTGAAATTTGAATCTAATACACAAATCATTC
55 AGAGCAATGTTGCTTCTGCCTACCACGAGTAATTCATTGGCCACTGGAAGTCAGAACAAGCTTCCCAGAA
GAGAGGTACCACTTGGACTACCAATATAAAAGGATGAAAATATCGGAGTGAAGGTGTTCTTGCATCACTG
AGTCCCTGGACAGCCTGTCCACTCATGCTGATATCTGAGCCTAATGCTTCTCTGAATGTTGAGATTTAACT
TTGATCCAATGAAACCAGACCAAGAAAGAAAGAAAGCTTTTCATTGTTGATAAGGACATGATTTTTCTCAC
AATTTTATGATTATTTTTTCTTAGCTGTCTTATAATTATCTGCTTATTTGTCTCTTCTCCATGTGCTTAGG
60 GTACAAAGTTGACCAAGACCAAGAATAATGTCTGGGAGCACAACTACTGACAGCACAGCTTTAAAACATGA

SUBSTITUTE SHEET (RULE 26)

19/22

TGAATGCTTTAATACAGGAAATGAGTAGGGGAGAGGCAAGTGGTGCTTGGGTGTTCTTCCAATGCATAGTA
TCTTCCTTGACACAGTCAGTGCAGCTCTCAGTAGGCAAGTCCCTACATGTTAGAAGATGTTACTTTCTGTG
GAATTAGGTGGCAGAACATGCCTTCAATTATTTTCCTTTGCAGAACACACCAATTTTCATTAGTTAGGACA
GAGTGCTGGCTGCATTTGAATTCCAAGCAACGATTAGTCTATCACTGTTGGTATAGATTCCAACCAGTCAC
5 ACCACCTCCTGAAGTTTGTGGGCAGGTAAATCTTCATCTTAGAATAAAAATCATCTTAGCCAAGTAAGTG
TTTTAGAGGAAAGAAGAAAACATAATCGTTTCCATAAGAGTTTGTCTTCTAAAAAATAAGAAAGGCTCTT
TGTTTAGGTGAGCTAATGAAGTTGTTGATAGTTATCAGATGACACTGGAATCTTACTTGCCAGAATGTGT
TCTGTGCACCTCTCGGTGTGGCAACATAGAGAGGGAGATCCTCCAGCAATGCCATTGATATGGTCAGAAAC
TGATCTTTCTTTCTCCCTGCTGAGATGGGGTCTTTGTTCTAGAAAACCCAGGGGGTGCCACTGGGAGTA
10 ACCCTTGAGACAGGAACACGAATCTCAACCAATTTCTGGTTGCAGCCTTGAGTCTTACTATTTGCCATAGT
GATGCTTAGCAAGGAATGGCAGGTGCACCAGAGCAGCAGAGGACCTAATATCTCCCTTCCTGTTAACTTTT
TATAATATTTTATTGTGATCAGTATCAGTTGGGAAGCTACTTGCAGTCACTGAGCCTCAGTTTCTACATCT
GTAAACTGGGGATAGTAGCATGGCCCTATTTAATGTGCTCAGCGAAGCCACTGAAAGGAGACAGAAATGTA
CCAGAATTCCTGGACTTTTATCCTACTTCTCCTGGGGATTGTCACCCACCTACCCGTGTCTGTCTCTTGT
15 TGCTTTGACGCTGTCACCTTCTTTTCTTAGGTACCTCTCTGTAGGGCTCCATTATTCCAGGGATTCCAGAGT
TACAGCACATGCATACCTCCATCCAAGCATGTTTATTTGTCTCCTGCTTCACTAGGCTGTCCCAAGGAAC
ATGTGGCTCCCGGCACATACCTGGCACAACACTGCACATGACATTCACCCACTTGGCCTTGAATCTGACAA
GGAATCTGGCATGATGTTTACCTGCTGAGGCCAGGTGCCGAGCAGCCCTGGAGGCTTAGGGGCCAGAGGGA
TGGGAAAAGGTGTCTTTCTGGGGTGAGTATCAGTTTCTGCAGGAGTGCTGAACCTGAGAAAGAATAAGAG
20 AGAAGGAAGTGAACAAGCACAGCTTAAACATCATCTGTTTCTACTGAGTTTAAACAACTCTGAGATTTTGT
TTGTATGGAATCCATTTCTCAGGCCAAGCAGACACAGAACTTGGGTGTGAGTGATGATAATGAGCTGATA
TAATTTTCACACCCTCATCACTGAGATCTCTCCCATCAGGAATGGGTACAGGGCTCACAGGTGGCAGCAA
CTGTTATTACAGGCCTCATCTCTACCAGCTCCTGGCACCTGCTCTCCTCTCATTAGAAAATCCTCCACTTG
TCAAAAAGGAAGCCATTTGCTTTGAATTCCAATTCACCCCTCAAGAGGCTGGGACCACCTCATTGGAGTCC
25 TTGATGCTGTGTGACCTGCAGTGACCACTGCCCCATTGTTGCTGGCTGAGGTGGTTTGGGTCAACCTGGCC
ATCTGGGCAGCTGTTCTCTTCTCTTCTTTCTCCCCTACTGTTTCCAGACATGCAGTATTTCCAGAGAGAAG
GGGCCACTCTTTGGCAAAGAACCTGTCTAACTTTCTATCTACGGCAGGACTTTTGAAAGCTACAGAGGAAG
AAGCACAAATTGATGCTATTTCCACTAAGCCATCAGCTCCATCTCATCCATGCCATGTCTCTTTTTTAGGGG
TCCTCTTGCCAAACAGAATCACAGAGGACAAATCTGAAAGTGCAGAGACAGCAGCTGAGGCACAGCCAAGAG
30 CTCTGGCTGTATTAATGACCTAAGAAGATGGAGTGGTCACCAGAAAGTCAGAGGAAGTGACACACAGGGGC
CCAGCAATCTCAGCCAAGTCAACTCCACCAGCCTTTCTGGTCCCCACTGTGTGTACAGCACCTGATAGGG
ACCAGAGCCATGAGAGTGAGTAAGACCAGACTATGCCCTTGAGGAGCTCACCTCTGCTAAGGGAAACAGGC
CTGGAAACACACAATGGTGGTAAAGAGGAAAGAAGACAATAGAAGTGCATGAAGGGGATGGAAAGTGCCCA
GGGGAGGAAATGGTTACTTCTGTGTGAGGGGGTTGGTGAGGAAAGACTCTAAGAGAAGGCTCTGTCTGGCT
35 GGGTATGAAAGGATGTGTAGGAGTCTTCTAGGGGGCACAGGCACACTCCAGGCATAGGTAAAGATCTGTAG
GCATGGCTTGTGAGGATGAGTTTCAAGTATTCTGGAATGAGGACAGCCATAGAGACAAGAGGAGAGTTAAT
AGATTTTATGCCAATGGCTCCACTTGAGTTTGTGATAAGAACCAGAACCTTGGACTCCCCAGTAACATT
GATTGAGTTGTGTATGATTCTACATAGAATATTAATCAATGGAGGTCAGTGAGTGGTGTGTGTGTGATTA
TTTGCCAACTGCCGAGGTGGAGAAGCCTCTTCCGACTGCAGGCAGAGCACGGGGGCCCTGCTACTGGCTGC
40 AGCTCCAGCCCTGCCTCCTTCTCCAGCATATAAACAATCCAACAGCCTCACTGAATCACTGCTGTGCAGGG
CAGGAAAGCTCCACACACACAGCCCAGCAAACAGCAGCA

20/22

Figure 9.

GCCAATTAGAAGAAACACAACCTACAAGGTCAGGGCATATTATTCAAACAGTAGAGACAATACAGTCAAATA
 TTTGGCAGAATTACAAAATATCTCATTGGAAAAGACACGCAAGGGAAATCAACAAAAAGATATGAATCAGA
 5 ATTCATCTGTGTCTCAAGAAAAGGTCATGCGATAAATTAAGTTCTGCTAGTGTCTTACTACACTACCGTTAGC
 CTCATTACCTTATTTTTTTAAGTGTTAATATAGTTTTAGGTATTTTACATACATTTTTTATTATTAATTACAA
 CCAAAGTGCAACTTGTAATAGCAATTCCTTCACATTTTTTTTTTCAAATCTTGACCTTAAATCCACCTC
 GGGCCTCAGTTGGCCAGCTTTGGTATCTGATACTTGGACTACAGATACCACTAAGGCAAGTAGATAAAATG
 TACTCTAGGACCTACAGCCCTTCTGCTAGATCCTGAAGAATGATCATTAACAAGCTGGTCTAGCTGGTC
 0 AAGAGCAAAAATAAAATCAAGATGACAGAAAATTGATGCAAAAGTGAAGTAAATAGCTAGAGAATATGAT
 TGC GCCTGTCCCCCTTAGCATGGATTCCCATGCTAGCCAATCTAAAATCCTCACTGTTAGAATCCTCCTGTC
 AATATGATAGAATGAACAGCAAGCTCAGTGTGAGAAAACCTGTGTTGTTAACTTGGCCCTCTTTCTAGCTG
 AATGTGTGTTTTTGGTCAAGTTCTTTGGCATTTCAGAGACTCAGAGTAGTGAAGGAAGTGGATAAGATGAC
 CTCTACATTCTCTTGCAAGCTCAAACATCTATGAATCCAGAGAGAAAACTAGAGCATGAAATTAAGGTTA
 5 TTTTAAAGAAATAACCTTAAATTTATAGTATTCGAGGATCTCCAATATATTCATGGCACCCTCAAACT
 TTCCTTCTGCTCTATCCCGTCTTGGCTCAAAGTTATCTCCTTAATGAGGTCTGCCCTGACTATCCTACTTA
 AAATTGTAAACTTTGGCCACCTGGTACTTCCACTCTCTTTCCCTGCTCTGTTTTTCACCGTAATACTTTA
 CTCTTTTTTAACATACAAAATCACTTATTTACTGTGTTGTTATCTATCTGCCTACTCTTACCATCAAATATA
 AGTTCTACCTAGGCAGGGATTTTTGTATGTTTTGCTCATGGATATATACGAAGCACTTAGAGTAATATGTG
 0 ACATATACAGGGTACTTGATTAATACTGTTGAGTGAATGAATGAGTTTCCAATACAAATTTAAATAAAAT
 ATTTCCCTAACTTAAATTTGTAAAGTCAGATCTAACCAACTGTTTCATTGGTCTGCTAGCAGTGTTCCTTGTA
 TATGGAAATATATTTTAAATAGATATGTCCTGTGAAATAATACTAAGTGTTCATAAGAAATAAGTGAGTGA
 ACGTTACCTCATTGAACTAAGTTGACCTTGCTCCTGGGAGAGAGTTTCATTGAGATTAAACAAGTTCAAAG
 TCTATGAATCATAAAACGATAAAAAAACTAAAAGGGAAATGGTGTTTTTATAAGCTCTGCAATTCAAAAG
 5 CCATTTCCGGGTAATATTGTTATTTTTATGTCAGGAATTCCTCAGTGCTGATATCTTAGGGCAAAGGGTTTG
 GTTATAAATTAAGAGAATGAGGAAATAGGTACATAGTAGGATTGTTCCAACCAATATGTGTTGAATGTCA
 AAGGAATTTCCCTGAGGAATAATCTTCAGAATAATTTGCTAAGCACAGGAGAAAATTTGGCTTATTACTTT
 ATAGCCAGATTTTCAATTTTAAATTGAACTTCTTTCAAGCAAATCACTTACTAGTCTATTAACAATAACAAC
 ATAAACACAAGTAAACATTCGGAATATAGACATCCAGGTACTAAGCTGATTGCTTTTACACTCACTGTCTTA
 0 TTTTACAAGTAAGGAGTTTTAGTTGCAGCAAAAGAAATAAATTTTCCAATGTCAAATGACCAGAACTTAAA
 CCAATCTGTTTGGTGCTAAAGCCAATGTTCTTTACTGCAATGTTGGGTATCTTGTTTCTAAAACCTTAAA
 TTTATCAGTAAAAGGCAAAATTTGCTATTATTGAGGACATTAAATCATATTTTTGTAGACTCTGAGGACA
 AATCCAACAAAAAAGTTCCAACCTATTTCTTGGCAGGCATCATTGAAATTTGGTATATAGCTTCCTTGGGTAT
 TGACTTTGAAAAGGAAGTTGGTCACTTTAGATATATAAGTTCAAGTCTGTTTGTAAAAACAAATGAAAACA
 5 AAACAGTTGCCTTATATGCTAAAATTATCCTAATCGTTTTTCACTTTTAAACAACATATACACACAGAACTTG
 AGGAACTTTACACGGCTCATCTTCATATTGTCAGCATCTAGCAAAGTACTTGCCACATAGTGATCAATAAA
 AGTTTTAGCCAGCCTGGGCAACATAGTGACAGCCTATCTCTACAAAAAAAATTAGCCAGGCAAGGTGGCG
 CACACCTTTGGTCCCAGCTACTTGAGAGGAGGATGTGGGGAGATCCTTTGAATGCAGGAGGTTGAGGCTGC
 AGTGAGCTGTGATTGCGCCACTGTGCTTCAACCTTGGTGACAGAGCAAGACCCTGTCTCACACACACACAA
 0 ATTAAGTAAAGAAAAAAAAGAATCAAAGAAAAAATAATTTCCCAGCTTAAGTCCATCTTTATTTGTTT
 GGATAAGCTATAAAGTGTCAAATAATGCTGTTAATGGACATTTCTCTAGCTCTCCCAAAGGAGGAATTGAG
 CACATAGTATGTGCTGTATTTTATATACAGAATAAAAAATAGAGACAAGATTTCTACCCTCACAGAACTTAA
 ATTCTTCAGGAGAATGACACTGAAGTCCTTAATTGGACTTCTCTCTTCTGTATTATCTTCCTCAAGTGGAG
 GTATATGGTGCTTAGTTATGAAAAATACCTCCAGGGCTTTGATCTTCTCAATAACTCTTTGAGGCTGATAT
 5 GAAAACAGTAATTAGAAAAAAACCATGTATCCAACCTTATATAGACAGTTGATGACCAAAGCTAGAATCCAG
 TTATTTTCAAGCTCCCATGTATTTTCTTATTACTTAAGGAGAATCTCTATCTCTACCTCTTTCTCTCTTCTT
 CCTCTCTCACTTTTCTTAGAAACATGGGTAAGATTTTTCAGAAATATGAGAACTTATTAATAAATGAAAAA
 TACTGGGAATTCTCAATGTTTCTTGTTTTAGCCAGTTAATTTTGGCCTTCATTCAATGTGAGTGTCCCTTA
 ATAAGGAGCAAACCTCCACTGAGAGATAGATACTAATAACCAGGATTTCTGAAAATGCATTCTCATCCCCATC
 0 TCCAAACTTTTATAAAAAATATTATAAAATAATACACTTTTAAATATAGGAAATTTCTCAAATACAGAAAAA
 ATTAAGTAAACTCAGACCTAACTTTTCATCACTAAAAGATAATCACTCTTGACACTTTGATATCTTTATCT
 CTAATATTACACCAATTAATTTGCTTGATATAGTGAATATTATGCTATTATAATTTTCCCCTGCCTTTTGT
 CCTTGCAATTTAGCATAGGTATTTTCTGAGGTTATCACAACTCTGTAAGCACGTTTTTATATTACTACTTT
 TTTAAAGAGGATGTATAATAATTAATTCATCCATATATATGTGGTTAAGTATTCAGGTCACTGCTCATTTT
 5 TCACTGTTATAAAATAAAGCAGCAATGAATACCTTTGGCTGATATTTTTTTCTGTACTTGGAATTATTTCT
 TTAAGATAGATTTCTAAAATTGAATTACTGAGTCAAACAGACTTAAGTTTTCTTATGTATGTTTCCTTA
 TTCATTTGAATAATTTTCACTCCTACTTGTATTAACTCTTGTGAGCATGTGATAGTCTCATTTTTCA
 AAATATCTTTGCTGTTGTAATTTGCATTTCTTTGTAGTTAGCATGAATATTTTCAAGTATGTTTTCTTCTGTG
 TACCAGTATACTACATACTTTTTTATATGAATTGCCTATTTGCATCTTTTGCTCGGTTCTATTAGACCTTT

21/22

GAATTTTTTCTTATCCATTTATATAAGCTCTTTATATATTAAGAATATTAACCTATTGTGATATTTGCAAT
AAATAGCTATATGGTTTGTGTTGGTTTAAATGTGAATTTATTCAATTTTCTTCATAATTTGTTGTTTT
TATAGATTTCTTTAAAAGTAATAAAATTATTGCCATATTATTATTTTCAATGTCCATTACTAGCCCT
TTCCAGTCATCACTTTTACTCTCATGTTCTTAATTTTTATTTCATATCTTGGCTCCATTGACCATCTTTATT
5 AGGATATTTGGGAAATAATAACAATTTATACTAAACACACATCAAATCTTACCTTATTTTCTTCTTACAAAA
ACCTAGGAATATGCTGTTTTTGTCTTTATTTGAATGACACAGAAATCAAGGTTTTTGAGCAGTGGAGTAT
TTCTTCAAATGACACAGAAATCAAGTCTTTTGAGCAGTGGAGTATTTCTTCAAAGCCACACAGCTAGTAAG
TCATGAAGCTGGAATTCCAAGAGTTGCCACTTCATTTTCTTCTTTCCCTTTATCTTACTCAGTTGTCTTCT
CTCCTCTTAATTTTGTCAATTCATTTAAAAACATTTCTTGTGCTATTATGGTAGATTTATTTTAATAGGGGG
10 CAGTGACTTACTCAGAGAGATGATTCTCTAATGGAGTTTTAAAGATCTTAGAAGTTGATAGAGGAGGCTGG
GCGTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCACTTGAGGCCAGAAG
TTCAAGATCAGCCTAGTCAATGTGGTGAAACCCCATTTCTACTAAAATTACAAAAATTAGCCTGGTGTGCT
GGTGCACCTCCATAATCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTGCTGGAACCCAGGAGGCGGAGG
TTGCAGTGCGCCTGGATTATTAACACTACACTCCAGCCTGGGAGACAGAGTAATACTCCATCTCAATAAAA
15 AGAAGTTGATAAGGGAGATAGTTTCATGGCAACGGATCTTTGAAGGCACGCTAATGATAACTTAGGCATTTA
GCCTACTAGTGTAATTTCCATAAATCTGCCTCTGATGTCATACTCTCAGCACCTAATATTTTCTACAAACA
TTTATTGAACTTTATTTTGTATAAGTCTCTGTCCAGTTTGAATATTTAAAAAATTCATAATCATATGAAA
CATTAATAATAAATACAAAATGAGAGATGCCGATACTGAAAAGTAGGATTGCGGAGTGGTAGAAAATATTT
CTGGCTGTAGTAGATGGGGAAGTGTTCAAAGAGGAGTATAATTCAGGTTTCCATTTGCCATCGACTTATCA
20 CATGGCTAACTCACTAAGCGACTTAATTAATAATTAAATTAATTTATCATCATCTGATCACCATTTACACA
ACTCATGTCTGTGCTGTATTGGCTAAATGATGGCAAGACAAACGACCTCTGAAAATGATCCTATTGACCT
TCGGAATCTGGATTTTTTTTTTCAATGCAGGTGTCCATAGAAGCAATCTGATGTAATCCAACATGAGTTCAA
GCACAGTCATTTAATATCCCCTATCAAGTACAGTCATTTAATATCCCCTATCATCACATGTCCTTCATACA
TAAAAATCATTACATGTGAAAGGGTGGAGAGTGTGTGGATCCCTTATTATTGTGTTATTGTAACACAATAA
25 CAATATTGTGTTATTATTGTTAACACAAGTGAGTCATATGTCTTGCTCTTTGGACTGAGTGGAACTTGTA
TTCTTTCTCTGCCTCAGGTCAATTAACCTTCATTGAGGTGAGTTGCATTCCTTCTTTAAGCATGTTGAACCT
TCAATCTGGACTCAGATGGGCTAAATAGAGGAGCTAGGAAAAATACAGAAAATAAATTATTAGAGAGATCA
GAGAAAGATACATAAGATTTACAATAAAAGAATTATGAGAAAAACATCCAAAAGAATTAAAAACCATAGGA
GAAGGAAAAATAGGTGAACAGCTTTTTTAATTTCTATAAATGTGTTGTTAATACTCATAATAAAGGACTCAG
30 AGCTGGGATATGAGAATAATAGGTCAAACGTATATGGATACATAGATGTGACTACATACATGAGTTGCAAA
GAATGCTAAGGAGGGCAAAAAGAGATTGAGAAGAGGGCATTATTACTAATATATAGCAATGTTGAATGTTT
AGGGTGTTTCAGGCACCTGTACAAATCTTTTAAATACACAAATCACTTAATCTTGCCATAACATTAGAAGAT
ACTATCTACTCTTTTACCAAAAAGGTAACCTGTGGGATAGCAGAGTTAAGTAACCCGTTCAAACCTATGCATA
ATAATCAGCAGAGATGGTCCTTATCTAAGTTTTTCTGCCTTTGAAGTCCAAATAGTTTAAATGCAGCCAGGT
35 ACTAAAGAAGAAAACCTTTTGTAATTAGTTTAGTTTAAATGATTTACATGTGGAAAAGCACAGAGTGAAAAG
CACATATCATCTGAGAAGCCCAGTGAGTTTGGCTGAAATGGAGTGAACATGTACATGTTGAGGGTGAGGAA
GATGATTAGAGAGAAGTGATGTGTGTGGTTCTTAAAAGCTAAGAGGAACTGTTAGATATGATATAGTCT
GTGGCAGGGAGCCATTGCAGGTTTTCCAAGATGGACTCATAAGGAGAAAACACTTCGTGGGACTGGAAAAG
GCAGTGAAGTGGTGTGCTTATGATAATGTACTACAGTGTAGGATAGTGGTTAAGAGTACAGTTATGAGAG
40 AGGGACTACTGGTTAGCACCTTACCTGCTGTGTGACTGGGCAAATAATGCAAACCTCAGTGTCTTTTATTG
TAATATGGGAGTAACAAAAATAGTAACCTTTCATAGGATTCTTGTAAGATTAAATGACTTAATTTCTTT
GAAGTGCTTGGCAGTTCCTGATAAATGACCAGTAGTTAATAAATGTTAGTTGTTATTATTATCATTATATA
TTATTACTCCCATAGATACATATAGAACAGACTGCAGCAGAGAGGCAAATCTTTAATGTTGTGTCAGAGTATA
GACAAGTTGGTGAAATGGCTACATGAGAGCGGAGGACAAGAAGGTGCAGATTGTGGCAGTCACTTCAAATG
45 GAAATATCACCGCTTGAATGAAGGTATATGAGTGTCAACTTGCAAGGGGACCAGGTAGGTTTCATCAGAAA
TTAAGGAAGCTTAAGGAGAACAGCCAAGTTCAGCTTGACAGAAGTGGTGGTGGCACAATGCAAGACTGGT
GTCTTTCAAGAAACCAAGGACTGTTGAAAGTAGCAAGAGCTAGTTTGTTTTAGGTCCATCATGTTTTATAT
TCACACTTTCATGTCAAGTGGAGCAAAGAAATGGAATACAATATAATAGAATGGTAGAATCTTATTTTTTAAA
ATCTGTGTTATTCTGATCTTTAACTTACTTATATCTTTGATAGAGATCTTTACCTGATGCTCAAGATTGTA
50 GAAATAGTATAATCAACATAACAGTATAGCACTGTATTTATATCCTGCACTGTTTAGGGAGGGTTTTAAGGC
CATTCAAAGGATACATAAAATACAACAAGATTACATAAATGAAAGGTGAGATAAAGCAACAAAGCAAAAC
AAAAGTGAAAACAGAGATCATAGGCACAAATAAGATTAAAAACGCATGTAATGAAGATGAAAGCTTTTACA
TTTACCCCAGATGGACCACAGGGTTGTTGTTAAGCCTTTAAACAGTGAACAATGCTGTACACTTGCATATG
CAATTAGAACATGTGGAAAAAATAGTGGCCTGTTAGAAGCCTAATTAACAATTTGTGAAAAAATAAAAAA
55 AAAAAAAGAGAGCCGAGCTGTAGCTCACGCCTGTAATCCCTGCACTTTGGGAGGCCGAGGCGGGCGGA
TCACGAGGTGAGGAGATCAAGACCATCCTGGCTAACACAGTGAAACCCAGTCTCTACGAAAAATACAAAAA
ATTAGCCGGGCGTGGTGGCGGGAGCCTGTAGTCCCAGCTACCTGGGAGGCTGAGGCAGGAGAATGGTGTGA
ACCCGGGAGGCGGAGCTTGCAAGTGGCCGAGATCCTGCCACTGCACTCCAGCCTGGGCGACAAAGCAAGAC
TCCGTCTCAAAAAAGAAAAAAGAAAGAAAAACAAAGAAACTTCATTGTATTGTAAGGCCAAGAACAAAA
60 TATATCAAGATAAGGAAAATTTGTAGTCAAGAATAGAAAAAATTTATGGCTTTGAAGTATGAGTTATTTAA

22/22

AGAAAGTGGAAACATCCTCAGACTATGCAGTAAAAAACAAAGTGATTTTCTTCTTCTAACTTATGCAATA
AACTGATAGGTAATATGTGAAAGTCATAGAATGTAGACTAGAGGATACAACAAACCTATTTCCCTCTATGTT
CATAAGAAGTAAGAAAAGCTCTGATGTGAGTTAGCATTTGCTTTACAATTTTGAATTGTGCAGATTGCACGT
ACTTTTCCTCAGTTTGAAGTAAATAGTGGACAGGAAAAAATATTAATGTTGGCAGTAAATATGGAAGGAA
5 ATTACAACATAATGTAATATGCTAAAACATGCTATGTTTATTTTACTAATTTGAATTAAAATGTAAGAATTT
AAAATGCCCTGGAAAAACACGGGCATTGATCTGACGTCTGAAGTTTTTAAAATATTACACACTTTGAAATAG
CATTTGTACCTTGAAATACCTGTCTCTATATATTTTTTTTAAACTTTCCTTTTTCTTTCATTCCATTTATCAT
CAAATAAAGGATGAACAGATGTAACAGAACTGTCAAGCATGCTGAAGAAAGACCACTGCAGAAAAATT
TCTCCTAGCCTTTTCAAAGGTGTTAGGAAGCAGAAAGGTGATACAGAATTGGAGAGGTCGGAGTTTTTGTA
10 TTAAGTGTATTAAATGCGAATCCCGAGAAAATTTCCCTTAACTACGTCTGTAGTTATATGGATATGAAGA
CTTATGTGAACTTTGAAAGACGTGTCTACATAAGTTGAAATGTCCCAATGATTCAGCTGATGCGCGTTTC
TCTACTTGCCCTTTCTAGAGAGGTGCAACGGAAGCCAGAACATTCCTCCTGGAAATTCAACCTGTTTCGCA
GTTTCTCGAGGAATCAGCATTTCAGTCAATCCGGGCCGGGAGCAGTCATCTGTGGTGAGGCTGATTGGCTGG
GCAGGAACAGCGCCGGGGCGTGGGCTGAGCACAGCCGCTTCGCTCTCTTTGCCACAGGAAGCCTGAGCTCA
15 TTCGAGTAGCGGCTCTTCCAAGCTCAAAGAAGCAGAGGCCGCTGTTTCGTTTCCTTTAGGTCTTTCCACTAA
AGTCGGAGTATCTTCTTCCAAAATTTACGTCTTGGTGGCCGTTCGAAGGAGCGCGAGGTAGGGGCACGCA
AAGCTGGGAGCTACTATGGGACAGTTCCCAAGTGTGAGGCTTTTCAGATTTCTTGAACCTTGGTCTTTCAGGG
AGAAGGGCTTCTTGAGGCGTGGATAGTGTGAAGTCTCTGGCAAGTCCATGGGGACCAAGTGGGGTTAGAT
CTAGACTCAGGAGCTCCTGGAGCAGCGCCCAAACCGTAGTGGCACTGGACCATGTTGCCCGGAGCGCGCAC
20 AGCCCGCGCGGTGCGGGGACCTGCTCTCTGAGCCCGCGGGCGGTGGGTGGGAGGAAGCATCGTCCGCGGCG
ACTGGAACCGGGAGGGAGAATCGCACTGGCGGCGGGCAAAGTCCAGAACGCGCTGCCAGACCCCCAACTCT
GCCTTCGTGGAGATGCTGGAGACCCCGCGCACAGGAAGCCCTGCAGTGCCCATCGCGGCCAGAGCAGCT
GGGGCATCAACGGCGGGCGCTCCCTCTTACTGCTCTCTGGCTTCGACGGGGGACTAGAGGTAGTCTCACC
TCCAGCGCGCCTGAGGCTCATGCATTTGGCTAATGAGCTGCGGTTTTCTTTCAGGTCCGAATGGATCTTGA
25 AGGGGACCGCAATGGAGGAGCAAAGAAGAAGAACTTTTTTAACTGAACAATAAAGGTAAGTAGCTTGTT
TCATTTTCATAGTTTACATAGTTGCGAGATTTGAGTAATTTATTTCTAGCCTCCAGCTCTGAAATAAATGA
CATGTTGTTGTTTTTAATTATTTTTTAAGAAACGCAAGCTAGCCTTTG

30

35