



- (51) International Patent Classification:
C12N 15/53 (2006.01) *C12Q 1/00* (2006.01)
- (21) International Application Number:
PCT/AU2015/050847
- (22) International Filing Date:
24 December 2015 (24.12.2015)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
2014905277 24 December 2014 (24.12.2014) AU
- (71) Applicant: THE UNIVERSITY OF QUEENSLAND
[AU/AU]; -, St Lucia, Queensland 4072 (AU).
- (72) Inventors: GILLAM, Elizabeth; The University of Queensland, St Lucia, Queensland 4072 (AU). GUMULYA, Yosephine; The University of Queensland, St Lucia, Queensland 4072 (AU).
- (74) Agent: FISHER ADAMS KELLY CALLINANS; Level 6, 175 Eagle Street, Brisbane, Queensland 4000 (AU).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,

[Continued on next page]

(54) Title: IMPROVED OXIDOREDUCTASES FOR BIOCATALYSIS

X₁₈₂ is position 397
X₁₈₄ is position 399
X₁₈₅ is position 403
X₁₈₆ is position 405
X₁₈₇ is position 412
X₁₈₈ is position 415
X₁₈₉ is position 416
X₁₉₀ is position 418
X₁₉₁ is position 421
X₁₉₇ is position 423
X₁₉₈ is position 427
X₁₉₉ is position 439
X₁₉₉ is position 440
X₁₉₉ is position 441
X₁₉₇ is position 444
X₁₉₈ is position 446
X₁₉₉ is position 447
X₂₀₀ is position 448
X₂₀₁ is position 449
X₂₀₂ is position 452
X₂₀₃ is position 456
X₂₀₄ is position 457
X₂₀₅ is position 462
X₂₀₆ is position 466
X₂₀₇ is position 467
X₂₀₈ is position 468
X₂₀₉ is position 469
X₂₁₀ is position 470
X₂₁₁ is position 472
X₂₁₇ is position 473
X₂₁₇ is position 474
X₂₁₄ is position 476
X₂₁₅ is position 477
X₂₁₆ is position 480
X₂₁₇ is position 483
X₂₁₈ is position 484
X₂₁₉ is position 485

(57) Abstract: Engineered ancestral P450 enzymes and P450 reductase enzymes are provided. Also provided are methods for the production of these engineered ancestral enzymes. The engineered ancestral enzymes may have one or more improved or enhanced properties relative to one or more existing P450 enzymes or P450 reductase enzymes. One or more of these enzymes can be used to perform chemical reactions, such as for structure- activity relationship analysis; pharmacological testing; bioremediation; or biosensor technology.

Figure 1



WO 2016/101045 A1



PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,

Published:

— *with international search report (Art. 21(3))*

TITLE

IMPROVED OXIDOREDUCTASES FOR BIOCATALYSIS

TECHNICAL FIELD

THE present invention relates to the use of oxidoreductases for biocatalysis. More particularly, the invention relates to isolated, engineered oxidoreductase proteins with improved characteristics for use in biocatalysis, and a process for developing these proteins.

BACKGROUND

Biocatalysis, the use of enzymes to perform chemical transformations of industrial relevance, is emerging as an attractive means of addressing bottlenecks in the synthesis and modification of new chemicals. The basic principle underpinning this is that enzymes can catalyse reactions with greater chemo-, regio- and stereo-selectivity than can be accomplished by purely chemical means (Walsh 2001). Additionally, biocatalysts have substantial utility in biosensor technology (Ronkainen et al. 2010) and for bioremediation (Wood 2008). Oxidoreductases comprise a major proportion of biocatalysts exploited industrially (Straathof et al. 2002). The P450 oxidoreductases are amongst the most versatile enzymes known, a property which has made them a high-priority target for exploitation as biocatalysts (Guengerich 2002). Furthermore, P450 oxidoreductases can be used in the preparation of small quantities of metabolites for use as experimental standards, and P450 oxidoreductase-based biocatalytic assays have substantial potential for screening drug candidates.

In microorganisms, P450s tend to be specialized to interact in a highly efficient manner with a relatively small set of structurally related substrates. However, efficiency is generally markedly diminished when microbial P450 forms act on unnatural substrates. In multicellular organisms, certain P450s perform key roles in xenobiotic metabolism, including the metabolism of drugs. These P450s show unusual characteristics compared to other more typical enzyme catalysts in that they act on an extraordinarily wide spectrum of substrates (Rendic 2002); such xenobiotic-metabolizing P450s may therefore be useful in a broader range of biocatalytic applications. However, known xenobiotic-metabolizing P450s possess relatively poor catalytic efficiency and inefficient coupling of product formation to cofactor consumption when acting on any one substrate, as well as sub-optimal interactions with their principal redox partner, NADPH-cytochrome P450 reductase (referred to variously as NPR or CPR) (Shimada et al. 2005; Tan et al. 1997).

Process efficiency considerations mean that it is usually desirable to add high concentrations of substrates and generate high concentrations of product in bioreactors (Straathof et al. 2002). Achieving high substrate concentrations may necessitate the use of organic solvents at concentrations that are deleterious to the activity or stability of most P450s (e.g. Chauret et al. 1998; Wong et al. 2004). Another inherent difficulty with using P450s in an industrial context is their thermal and mechanical instability; this is particularly a problem for the lipophilic, membrane-bound, mammalian P450s, which typically need to be processed and stored in a high concentration of glycerol or some other cryoprotectant (Guengerich 1995; Guengerich et al. 1996). Conducting reactions under increased temperature may also improve efficiency, however at temperatures above which a P450 is stable the deactivation of enzyme will negate efficiency gains.

The identification and/or development of P450s with enhanced properties for industrial applications is clearly desirable. To date, screening libraries of existing P450s for desired properties has been seen as more expedient than engineering a new catalyst from a nonspecialized precursor (Straathof et al. 2002). However, improvements in protein engineering offer new scope to rapidly customize enzymes (Burton et al. 2002). Certain well-characterized and relatively efficient bacterial P450 forms have been engineered to redefine their substrate specificity, regio and enantio-selectivity, and catalytic efficiency. However, one potential limitation of substrate-specific microbial P450-derived mutant proteins, particularly for medicinal chemistry applications, is that they will be unlikely to produce many minor metabolites that may be of interest. Thus, the xenobiotic-metabolising P450s (e.g. mammalian enzymes), may be useful for engineering proteins that can produce a wider range of metabolites.

SUMMARY

The invention is broadly directed to the production of “engineered ancestral” P450 enzymes and/or redox partners suitable for use with P450 enzymes. It is a preferred object of the invention to provide “engineered ancestral” P450 enzymes and/or redox partners suitable for use with P450 enzymes that display or possess one or more desirable properties that are at least partly absent in extant P450 enzymes and/or redox partners or are relatively enhanced compared to extant P450 enzymes and/or redox partners.

In a first aspect, the invention provides an isolated protein comprising the amino acid sequence set forth in SEQ ID NO:1, or an amino acid sequence at least 80%

identical to SEQ ID NO:1, wherein residues X_1 - X_{219} in SEQ ID NO:1, or in the amino acid sequence at least 80% identical to SEQ ID NO:1, may be any amino acid.

In an embodiment of this aspect, one or more of the residues X_1 - X_{219} is selected from the respective group of variable amino acids as set forth in Table 9.

5 In particular embodiments, the isolated protein comprises an amino acid sequence set forth in any one of SEQ ID NO:2-41 or SEQ ID NOS:544-578.

Suitably, the isolated protein of this aspect has P450 enzyme activity.

This aspect also provides fragments, variants and/or derivatives of the isolated protein.

10 In one embodiment the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:2-41 or SEQ ID NOS:544-578.

In a second aspect, the invention provides an isolated protein comprising the amino acid sequence set forth in SEQ ID NO:180, or an amino acid sequence at least 80% identical to SEQ ID NO:180, wherein residues X_1 - X_{163} in SEQ ID NO:180, or in
15 the amino acid sequence at least 80% identical to SEQ ID NO:180, may be any amino acid.

In an embodiment of this aspect, one or more of the residues X_1 - X_{163} is selected from the respective group of variable amino acids set forth in Table 10.

In particular embodiments, the isolated protein comprises an amino acid
20 sequence set forth in any one of SEQ ID NOS:181-250.

Suitably, the isolated protein of this aspect has P450 enzyme activity.

This aspect also provides fragments, variants and/or derivatives of the isolated protein.

One embodiment provides an isolated protein comprising an amino acid
25 sequence at least 80% identical to the amino acid sequence of any one of SEQ ID NOS:181-250.

In a third aspect the invention provides an isolated protein comprising the amino acid sequence set forth in SEQ ID NO:321, or an amino acid sequence at least 80% identical to SEQ ID NOS:321, wherein residues X_1 - X_{11} in SEQ ID NO:321, or in the
30 amino acid sequence at least 80% identical to SEQ ID NO:321, may be any amino acid.

In an embodiment of this aspect, one or more of the residues X_1 - X_{11} is selected from the respective group of variable amino acids set forth in Table 11.

In particular embodiments, the isolated protein comprises an amino acid set forth in SEQ ID NOS:322-431.

Suitably, the isolated protein of this aspect has P450 reductase enzyme activity.

This aspect also provides fragments, variants and/or derivatives of the isolated protein.

One embodiment provides an isolated protein comprising an amino acid
5 sequence at least 80% identical to the amino acid sequence of any one of SEQ ID
NOS:322-431.

In a fourth aspect, the invention provides a method of producing or constructing
an isolated protein, said method including the step of producing or constructing an
engineered ancestral amino acid sequence of at least a fragment of a P450 protein or
10 P450 reductase protein from one or more P450 protein or P450 reductase amino acid
sequences that are different to the engineered ancestral amino acid sequence.

Suitably, the engineered ancestral P450 protein or P450 reductase protein
displays or possesses one or more increased or enhanced properties compared to at least
one of the one or more P450 proteins or P450 reductase proteins comprising amino acid
15 sequences that are different to the engineered ancestral amino acid sequence.

In a fifth aspect, the invention provides an isolated P450 protein or P450
reductase protein produced according to the method of the fourth aspect.

In certain embodiments of the fourth or fifth aspects, the isolated protein having
P450 enzyme activity is the isolated protein of the first aspect. In one particular
20 embodiment, the isolated protein comprises an amino acid sequence set forth in SEQ ID
NO:1 or SEQ ID NO:2.

Non-limiting examples of the one or more P450 protein amino acid sequences
that are different to said engineered ancestral amino acid sequence are set forth in SEQ
ID NOS:42-179.

25 In another embodiment, the isolated protein having P450 enzyme activity is the
isolated protein of the second aspect. In one particular embodiment, the isolated protein
comprises an amino acid sequence set forth in SEQ ID NO:180 or SEQ ID NO:181.

Non-limiting examples of the one or more P450 protein amino acid sequences
that are different to said engineered ancestral amino acid sequence are set forth in SEQ
30 ID NOS:251-320.

In another embodiment, the isolated protein having P450 reductase enzyme
activity is the isolated protein of the third aspect. In one particular embodiment, the
isolated protein comprises an amino acid sequence set forth in SEQ ID NO:321 or SEQ
ID NO:322.

Non-limiting examples of the one or more P450 reductase protein amino acid sequences that are different to said engineered ancestral amino acid sequence are set forth in SEQ ID NOS:432-540.

5 In a sixth aspect, the invention provides a method of producing or constructing a modified engineered ancestral P450 protein or P450 reductase protein, said method including the step of introducing one or more amino acid substitutions in an amino acid sequence of the engineered ancestral P450 protein or P450 reductase protein to thereby produce or construct the modified engineered ancestral P450 protein or P450 reductase protein.

10 In certain preferred embodiments, said amino acid substitutions are non-conservative amino acid substitutions.

Suitably, the modified engineered ancestral P450 protein or P450 reductase protein displays or possesses one or more increased or enhanced properties compared to the engineered ancestral P450 protein or P450 reductase protein.

15 In a seventh aspect, the invention provides a modified engineered ancestral P450 protein or P450 reductase protein produced according to the method of the sixth aspect.

In certain embodiments of the sixth or seventh aspect, the modified P450 enzyme comprises an amino acid sequence set forth in SEQ ID NO:1, any one of SEQ ID NOS:3-41 or SEQ ID NOS:544-578.

20 In other particular embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:3-41 or SEQ ID NOS:544-578.

In other embodiments, the modified P450 enzyme comprises an amino acid sequence set forth in SEQ ID NO:180 or any one of SEQ ID NOS:182-250.

In still other embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:182-250.

In yet other embodiments, the modified P450 reductase enzyme comprises an amino acid sequence set forth in SEQ ID NO:321 or any one of SEQ ID NOS:323-431.

30 In still yet other embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:323-431.

In an eighth aspect, the invention provides an isolated nucleic acid encoding an isolated protein, fragment or derivative of any one of the aforementioned aspects, or produced according to the fourth aspect.

In a ninth aspect, the invention provides a genetic construct comprising an isolated nucleic acid of the eighth aspect.

In a tenth aspect, the invention provides a host cell comprising the genetic construct of the ninth aspect.

5 In an eleventh aspect, the invention provides an antibody or antibody fragment which binds and/or is raised against an isolated protein of any one of the aforementioned aspects.

Suitably, in embodiments wherein the isolated protein is of the first aspect, the antibody or antibody fragment shows at least partial specificity for an isolated protein
10 comprising any one of the amino acid sequences set forth in SEQ ID NOS:2-41 or SEQ ID NOS:544-578.

Suitably, in embodiments wherein the isolated protein is of the second aspect, the antibody or antibody fragment shows at least partial specificity for an isolated protein comprising any one of the amino acid sequences set forth in SEQ ID NOS:181-
15 250.

Suitably, in embodiments wherein the isolated protein is of the third aspect, the antibody or antibody fragment shows at least partial specificity for an isolated protein comprising any one of the amino acid sequences set forth in 322-431.

A twelfth aspect of the invention relates to a composition for performing a
20 chemical reaction, said composition comprising one or more isolated proteins according to the aforementioned aspects and one or more buffers, solvents and/or other reagents suitable for performing the chemical reaction.

A thirteenth aspect of the invention relates to a method of performing a chemical reaction, said method including the step of exposing one or more substrate molecules to
25 one or more isolated proteins according to the aforementioned aspects to thereby perform the chemical reaction.

A fourteenth aspect of the invention relates to a reaction product produced according to the method of the thirteenth aspect.

A fifteenth aspect of the invention relates to use of the isolated protein or
30 composition of the aforementioned aspects for biocatalysis, preferably for structure-activity relationship analysis; pharmacological testing; bioremediation; or biosensor technology.

It will be appreciated that the indefinite articles “a” and “an” are not to be read as singular indefinite articles or as otherwise excluding more than one or more than a

single subject to which the indefinite article refers. For example, “a” protein includes one protein, one or more proteins or a plurality of proteins.

As used herein, unless the context requires otherwise, the words “comprise”, “comprises” and “comprising” will be understood to mean the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

BRIEF DESCRIPTION OF THE FIGURES

In order that the invention may be readily understood and put into practical effect, preferred embodiments will now be described by way of example with reference to the accompanying figures, wherein:

Figure 1 sets out the amino acid sequences SEQ ID NOS:1 and 2.

Figure 2 sets out the amino acid sequences SEQ ID NOS:180 and 181.

Figure 3 sets out the amino acid sequences SEQ ID NOS:321 and 322.

Figure 4 is a schematic representation of the construction of the bicistronic expression vector for CYP3_N1, pCW/3_N1His/hNPR.

Figure 5 is a schematic representation of the construction of the monocistronic expression vector for CYP3_N1, pCW/3_N1His.

Figure 6 is a schematic representation of the construction of the bicistronic vector for coexpression of the CYP3_N1-EGFP fusion with hCPR, pCW/3_N1-EYFP His/hNPR.

Figure 7 is a schematic representation of the construction of the bicistronic expression vector for CYP2D_N1.

Figure 8 sets out the N-terminal sequences of two inferred ancestors (CYP2D_N1-nat; and CYP2D_N1-FL which is herein referred to as CYP2D_N1) and the extant CYP2D proteins. The modifications made include changing the second residue to alanine (Gillam *et al.* 1995, *Arch. Biochem. Biophys.* 319, 540-550) and using the MAKKTSSKGK leader sequence (von Wachenfeldt *et al.* 1997, *Arch. Biochem. Biophys.* 339, 107-114) before the conserved PPGP motif. The CYP2D6 modifications used previously have also been included (Gillam *et al.* 1995, *Arch. Biochem. Biophys.* 319, 540-550). The “FL” sequences are full-length and the “trunc” sequences truncated.

Figure 9 is a schematic representation of the construction of the bicistronic expression vector for CPR_N1.

Figure 10 sets out thermostability profiles of extant (CYP3A4, CYP3A5, CYP3A27, and CYP3A37, labelled as 3A4, 3A5, 3A27, and 3A37, respectively) and

engineered ancestral (CYP3_N1, labelled as 3_N1) CYP3 proteins, as measured by the percentage of folded protein after treatment with various temperatures. Heat treatment comprised heating the protein at the indicated temperature for 60 min, followed by cooling at 4°C and equilibration to room temperature for 5 min.

5 Figure 11 sets out thermostability profiles of extant (2D22 from mouse) and ancestor (2D_N1) CYP2Ds based on percentage of folded protein after heat treatment. 2D_N1 refers to the engineered ancestral CYP2D_N1 protein, as described herein. Heat treatment comprised heating the protein at the indicated temperature for 60 min, followed by cooling at 4°C and equilibration to room temperature for 5 min.

10 Figure 12 sets out an Arrhenius plot of extant (CYP3A4, CYP3A5, CYP3A27, and CYP3A37, labelled as 3A4, 3A5, 3A27, and 3A37, respectively) and engineered ancestral (CYP3_N1, labelled as 3_N1) CYP3 proteins based on the percentage of folded protein after heat treatment. Samples were periodically taken and the remaining folded protein was subsequently measured. Heat treatment comprised heating the
15 protein at the indicated temperature for 60 min, followed by cooling at 4°C and equilibration to room temperature for 5 min.

 Figure 13 sets out thermostability profiles of CYP3_N1 expressed in mono- (labelled as 3_N1) and bicistronic (labelled as 3_N1_hNPR) format based on percentage of folded protein after heat treatment. Heat treatment comprised heating the protein at
20 the indicated temperature for 60 min, followed by cooling at 4°C and equilibration to room temperature for 5 min.

 Figure 14 sets out thermostability profiles of engineered ancestral CPR_N1 and extant human CPR (labelled as hNPR). Heat treatment comprised heating the protein at the indicated temperature for 60 min, followed by cooling at 4°C and equilibration to
25 room temperature for 5 min. Data are shown for three preparations of CPR_N1 derived from separate cultures (CPR.N1 #1, CPR.N1 #2, and CPR.N1 #3) compared to a pooled preparation of hNPR.

 Figure 15 sets out solvent stability profiles of engineered ancestral CYP3_N1 (labelled as TS) and extant CYP3A4 (labelled as 3A4) in methanol, DMSO, or
30 acetonitrile at various concentrations.

 Figure 16 sets out solvent stability profiles of engineered ancestral CYP3_N1 (labelled as TS) and extant CYP3A4 (labelled as 3A4) in 10% methanol, 10% DMSO, or 10% acetonitrile.

Figure 17 is a schematic representation of the construction of the CYP3_N1 library.

Figure 18 sets out ligand binding data for engineered ancestral CYP3_N1 (labelled as CYP_TS) and extant CYP3A4 (labelled as Parent). ‘-’ indicates that data is not presented.

Figure 19 sets out CYP3_N1 variant protein amino acid sequences SEQ ID NOS:3-41 respectively, in FASTA format in descending order.

Figure 20 sets out CYP2D_N1 variant protein amino acid sequences SEQ ID NOS:182-250 respectively, in FASTA format in descending order.

Figure 21 sets out CPR_N1 variant protein amino acid sequences SEQ ID NOS:323-431 respectively, in FASTA format in descending order.

Figure 22 sets out extant animal CYP3 protein amino acid sequences SEQ ID NOS:42-179 respectively, in FASTA format in descending order.

Figure 23 sets out extant animal CYP2D protein amino acid sequences SEQ ID NOS:251-320 respectively, in FASTA format in descending order.

Figure 24 sets out extant animal CPR protein amino acid sequences SEQ ID NOS:432-540 respectively, in FASTA format in descending order.

Figure 25 sets out exemplary nucleotide sequences SEQ ID NOS:541-543, encoding CYP3_N1, CYP2D_N1, and CPR_N1, respectively, in FASTA format in descending order.

Figure 26 sets out certain compounds, and the structure thereof, that may act as substrates for P450 enzymes.

Figure 27 sets out an overview of the activity of CYP3_N1 (labelled as CYP3_TS) on various substrates.

Figure 28 sets out the N-terminal “membrane anchor” region within certain extant CYP3 enzymes and the engineered ancestral CYP3_N1.

Figure 29 sets out substrate saturation curves for 6 β -hydroxylation of testosterone by CYP3_N1 and CYP3A4.

Figure 30 sets out CYP3_N1 variant protein amino acid sequences SEQ ID NOS:544-578 respectively, in FASTA format in descending order.

Figure 31 sets out metabolism of tamoxifen (100 μ M) to its major demethylated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes

indicated plus human NADPH-cytochrome P450 reductase were incubated with 100 μ M tamoxifen. At 20 and 120 minutes, respectively, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were lyophilised then resuspended for analysis by LC-MS. Results are the means $\pm$ standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 32 sets out metabolism of erythromycin (100 μ M) to its major demethylated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 100 μ M erythromycin. At 20 and 120 minutes, respectively, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were lyophilised then resuspended for analysis by LC-MS. Results are the means $\pm$ standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 33 sets out metabolism of erythromycin (10 μ M) to its major demethylated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 10 μ M erythromycin. At 20 and 120 minutes, respectively, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were lyophilised then resuspended for analysis by LC-MS. Results are the means $\pm$ standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 34 sets out metabolism of erythromycin (100 μ M) to its minor demethylated metabolite by CYP2D_N1 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 100 μ M erythromycin. At 20 and 120 minutes,

reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were lyophilised then resuspended for analysis by LC-MS. Results are the means +/- standard deviation of three independent replicates. No activity was observed with CYP3_N1 or CYP3A4. The percent
5 conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 35 sets out metabolism of ticlopidine (100 μ M) to its desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was
10 included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 100 μ M ticlopidine. At the times indicated, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were lyophilised then resuspended for analysis by LC-MS. Results are the means +/-
15 standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 36 sets out metabolism of ticlopidine (10 μ M) to its desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1
20 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 10 μ M ticlopidine. At the times indicated, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts were
25 lyophilised then resuspended for analysis by LC-MS. Results are the means +/- standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

Figure 37 sets out metabolism of ticlopidine (100 μ M) to its doubly desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1
30 (green triangle). hNPR expressed in the absence of any P450 (purple cross) was included as a control. Bacterial membranes containing 0.1 μ M of the P450 enzymes indicated plus human NADPH-cytochrome P450 reductase were incubated with 100 μ M ticlopidine. At the times indicated, reactions were quenched by addition of two volumes of acetonitrile and protein was removed by sedimentation. Reaction extracts

were lyophilised then resuspended for analysis by LC-MS. Results are the means +/- standard deviation of three independent replicates. The percent conversion is based on the mass spectrometer response and the ratio of metabolite to parent in that particular sample.

5

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

	SEQ ID NO:1	Amino acid sequence of generic engineered ancestral CYP3 sequence.
10	SEQ ID NO:2	Amino acid sequence of engineered ancestral CYP3 sequence CYP3_N1.
	SEQ ID NOS:3-41	Amino acid sequences of CYP3_N1 variants.
15	SEQ ID NOS:42-179	Amino acid sequences of extant animal CYP3 sequences.
	SEQ ID NO:180	Amino acid sequence of generic engineered ancestral CYP2D sequence.
20	SEQ ID NO:181	Amino acid sequence of engineered ancestral CYP2D sequence CYP2D_N1.
	SEQ ID NOS:182-250	Amino acid sequences of CYP2D_N1 variants.
25	SEQ ID NOS:251-320	Amino acid sequences of extant animal CYP2D sequences.
	SEQ ID NO:321	Amino acid sequence of generic engineered ancestral CPR sequence.
30	SEQ ID NO:322	Amino acid sequence of engineered ancestral CPR sequence CPR_N1.
	SEQ ID NOS:323-431	Amino acid sequences of CPR_N1 variants.
35	SEQ ID NOS:432-540	Amino acid sequence of extant animal CPR sequences.
	SEQ ID NO:541	Exemplary nucleotide sequence encoding CYP3_N1.
40	SEQ ID NO:542	Exemplary nucleotide sequence encoding CYP2D_N1.
	SEQ ID NO:543	Exemplary nucleotide sequence encoding CPR_N1.
45	SEQ ID NOS:544-578	Amino acid sequences of further CYP3_N1 variants.

DETAILED DESCRIPTION

This invention relates to the design and production of engineered P450 enzymes and/or P450 redox partners that represent enzymes ancestral to extant P450 enzymes and/or P450 redox partners, respectively. The invention is at least partly predicated on the surprising discovery that said engineered ancestral enzymes P450 enzymes and/or P450 redox partners may have one or more increased or enhanced properties relative to one or more extant P450 enzymes and/or P450 redox partners.

For example, in certain embodiments an engineered ancestral enzyme of the invention has substantially increased thermal stability. This is a highly surprising discovery because the temperature conditions to which the hypothetical engineered ancestral enzyme may have been exposed are likely to be substantially similar to the temperature conditions to which one or more corresponding extant enzymes are exposed.

In particular embodiments, “engineered ancestral” CYP3 and CYP2D P450 enzymes have been created which are putative or hypothetical ancestors of extant P450 enzymes. Furthermore, amino acid residues in the engineered ancestral CYP3 and CYP2D P450 enzymes have been identified that may be modified to confer, modify or remove one or more properties of the P450 enzymes. These include substrate specificity or promiscuity, thermal stability, pH stability and kinetic properties, although without limitation thereto.

In another aspect, the invention relates to a P450 redox partner suitable for use with a P450 enzymes such as CYP3 and CYP2D, although without limitation thereto. Suitably, the P450 redox partner is a P450 reductase. In a particular embodiment, the P450 redox partner is an “engineered ancestral” CPR P450 reductase. Furthermore, amino acid residues in the engineered ancestral CPR P450 reductase have been identified that may be modified to confer, modify or remove one or more properties of the P450 reductase enzymes. These include substrate specificity or promiscuity, thermal stability, pH stability and kinetic properties, although without limitation thereto.

For the purposes of this invention, by “*isolated*” is meant material that has been removed from its natural state or otherwise been subjected to human manipulation. Isolated material may be substantially or essentially free from components that normally accompany it in its natural state, or may be manipulated so as to be in an artificial state together with components that normally accompany it in its natural state. Isolated material may be in native, chemical synthetic or recombinant form.

By “*protein*” is meant an amino acid polymer. The amino acids may be natural or non-natural amino acids, D- or L- amino acids as are well understood in the art.

A “*peptide*” is a protein having no more than fifty (50) amino acids.

A “*polypeptide*” is a protein having more than fifty (50) amino acids.

5 As used herein terms such as “*oxidizing*” and “*oxidation*” refer to increasing the oxidation state of an atom, molecule or ion through a loss or transfer of one or more electrons from the atom, molecule or ion. Typically, oxidation is accompanied by an associated reduction in oxidation state of another atom, molecule or ion through a gain or transfer of one or more electrons by or to the atom, molecule or ion. These associated
10 or linked changes in oxidation state are referred to as “*redox*” reactions.

As used herein a “*P450 enzyme*” or a “*protein having P450 enzyme activity*” is a protein having one or more activities of a P450 enzyme. A preferred activity is monooxygenase activity according to the reaction: $\text{RH} + \text{O}_2 + \text{NAD(P)H} + \text{H}^+ \rightarrow \text{ROH} + \text{H}_2\text{O} + \text{NAD(P)}^+$ where R is a carbon-containing heteroatom. Particular, non-limiting
15 examples of such reactions include hydroxylation at aromatic and aliphatic centres, epoxidation, N-, O-, S-dealkylation and N-, and S-oxidation, acyl migration, oxidative dehalogenation, ring expansion, contraction and cleavage, C-C bond cleavage, denitrosation of N-nitrosamines, oxidative ester cleavage, aldehyde scissions (*e.g.* to alkenes and HCOOH), ipso attack on aromatic ring substituents and N- or O-
20 dearylation. One or more other activities may include NO synthase-like activity, reductase activity (*e.g.* reductions of alkyl halides, N-oxides, nitro compounds, inorganic molecules such as SO_2 , Cr(VI) or NO, desaturations (*e.g.* dehydrogenations), one electron oxidations, isomerizations and/or phospholipase D activity (*e.g.* phosphate ester hydrolysis). As used herein, “*CYP3 protein*” or “*CYP3 enzyme*” refers to a
25 particular class or family of proteins having P450 enzyme activity. Additionally, “*CYP2D protein*” or “*CYP2D enzyme*” refers to another particular class or family of proteins having P450 enzyme activity.

Typically, a redox partner is required for redox activity of a P450 enzyme. As used herein, a “*redox partner*” is a protein that facilitates the transfer of electrons from
30 an electron donor molecule to a P450 enzyme. Generally, the redox partner of a P450 enzyme is a P450 reductase, although without limitation thereto.

As used herein, a “*P450 reductase*” or a “*protein having P450 reductase activity*” is a protein having one or more activities of a P450 reductase enzyme. A preferred activity is the transfer of electrons from an electron donor molecule to a P450

enzyme. As used herein, “CPR protein” or “CPR enzyme” refers to the “cytochrome P450 reductase” class or family of proteins having P450 reductase activity. Generally, a CPR enzyme comprises a “flavin adenine dinucleotide” (“FAD”) -binding domain and a “flavin mononucleotide” (“FMN”) -binding domain. The general scheme of electron
5 flow during redox reactions involving CPR and P450 enzymes is: $\text{NADPH} \rightarrow \text{FAD} \rightarrow \text{FMN} \rightarrow \text{P450} \rightarrow \text{O}_2$, although without limitation thereto.

It will be appreciated that a P450 enzyme may have a diverse array of substrates including testosterone, progesterone, midozalam, nifedipine, tamoxifen, cyclosporin A, erythromycin, cyclophosphamide, paracetamol, lignocaine, ethosuximide, codeine,
10 lovastatin, 7-benzyloxy-4-(trifluoromethyl)-coumarin, 7-benzyloxyresorufin, terfenadine, S-omeprazole and benzyloxyluciferin; pesticides such as organochlorine pesticide, an organophosphate pesticide, or a pyrethroid pesticide; solvents such as perchloroethylene (PCE) or trichloroethylene (TCE); and food contaminants such as carbamate or organophosphate pesticide residues, although without limitation thereto.

15 Without limitation, certain compounds and the chemical structure thereof, that may be substrates of a P450 enzyme, are set out in FIG. 26 and/or FIG. 27.

Isolated proteins

The 138 amino acid sequences respectively set forth in SEQ ID NOS:42-179
20 (FIG. 22) are the amino acid sequences of certain isolated CYP3 proteins of animals (i.e. “extant” CYP3 proteins).

The amino acid sequence set forth in SEQ ID NO:1 is the amino acid sequence of an isolated engineered ancestral CYP3 enzyme, as set out in FIG. 1. SEQ ID NOS:2-41 (FIGS. 1 and 19) and SEQ ID NOS:544-578 (FIG. 30) are particular amino acid
25 sequences of “variants” of SEQ ID NO:1, comprising variations at one or more of the amino acid residues designated X_1 - X_{219} .

The location of the residues X_1 - X_{219} in SEQ ID NO:1 with reference to SEQ ID NO:2 is given in table form in Table 9. Certain preferred amino acids for each of the residues X_1 - X_{219} , respectively, are set forth in Table 9 under the heading ‘Variable’.

30 In one embodiment, one or more of the residues X_1 - X_{219} is an amino acid sequence selected from the respective groups consisting of the variable amino acids set forth in Table 9.

It will be appreciated that in some embodiments, the isolated protein comprising the amino acid sequence set forth in SEQ ID NO:1, and variants thereof, may be

referred to as a protein having P450 activity that is “ancestral” to each of the proteins set forth in SEQ ID NOS:42-179.

Suitably, the isolated engineered ancestral CYP3 enzyme comprising the sequence set forth in SEQ ID NO:1, and variants thereof, have one or more improved or enhanced properties compared to one or more of the isolated proteins comprising sequences set forth in SEQ ID NOS:42-179.

The particular engineered ancestral CYP3 enzyme variant comprising the amino acid sequence set forth in SEQ ID NO:2, and described in the EXAMPLES, is herein referred to as “CYP3_N1”. Additionally, the isolated proteins comprising the amino acid sequences set forth in SEQ ID NOS:3-41 and SEQ ID NOS:544-578 are herein referred to as “CYP3_N1 variants”.

The CYP3_N1 variants set forth in SEQ ID NOS:3-41, are variants of CYP3_N1 which represent the CYP3_N1 variant library constructed as set forth in EXAMPLES 9-10.

Furthermore, The CYP3_N1 variants set forth in SEQ ID NOS:544-578 are variants of CYP3_N1 which represent nodes of the evolutionary tree calculated for construction of the engineered ancestral CYP3 protein as described in EXAMPLE 1.

In certain preferred embodiments, a CYP3_N1 variant comprising an amino acid sequences set forth in SEQ ID NOS:3-41 or SEQ ID NOS:544-578 has one or more improved or enhanced properties compared to CYP3_N1.

Non-limiting examples of the one or more improved properties of an engineered ancestral CYP3 enzyme as herein described include: thermal stability, stability in solvents (e.g. organic solvents), metabolite production, catalytic versatility, catalytic efficiency (e.g. the efficiency of coupling of product formation to cofactor consumption), substrate specificity (e.g. increased specificity or increased genericity, as desired) and enzyme kinetic properties (e.g increased $V_{\max}$, lower K_m), although without limitation thereto.

The 70 amino acid sequences respectively set forth in SEQ ID NOS:251-320 (FIG. 23) are the amino acid sequences of certain isolated CYP2D proteins of animals (i.e. “extant” CYP2D proteins).

The amino acid sequence set forth in SEQ ID NO:180 is the amino acid sequence of an isolated engineered ancestral CYP2D protein, as set out in FIG. 2. SEQ ID NOS:181-250 (FIGS. 2 and 20) are particular amino acid sequences of variants of SEQ ID NO:180, comprising variations at one or more of the amino acid residues

designated X_1 - X_{163} The location of the residues X_1 - X_{163} in SEQ ID NO:180 with respect to SEQ ID NO:181 is given in table form in Table 10.

Certain preferred amino acids for each of the residues X_1 - X_{163} , respectively, are set forth in Table 10 under the heading 'Variable'.

5 In one embodiment, one or more of the residues X_1 - X_{163} is an amino acid sequence selected from the respective groups consisting of the variable amino acids set forth in Table 10.

10 It will be appreciated that in some embodiments, the isolated protein comprising the amino acid sequence set forth in SEQ ID NO:180, and variants thereof, may be referred to as a protein having P450 activity that is ancestral to each of the proteins set forth in SEQ ID NOS:251-320.

Suitably, the isolated engineered ancestral CYP2D protein comprising the sequence set forth in SEQ ID NO:180, and variants thereof, have one or more improved or enhanced properties compared to one or more of the isolated proteins comprising
15 sequences set forth in SEQ ID NOS:251-320.

The particular engineered ancestral CYP2D enzyme variant comprising the amino acid sequence set forth in SEQ ID NO:181 is referred to herein as "CYP2D_N1". Additionally, the isolated proteins comprising the amino acid sequences set forth in SEQ ID NOS:182-250 are herein referred to as "CYP2D_N1 variants".

20 The CYP2D_N1 variants set forth in SEQ ID NOS:182-250 are variants of CYP2D_N1 which represent nodes of the evolutionary tree calculated for construction of the engineered ancestral CYP2D protein as described in EXAMPLE 3.

In certain preferred embodiments, a CYP2D_N1 variant comprising an amino acid sequences set forth in SEQ ID NOS:182-250 has one or more improved or
25 enhanced properties compared to CYP2D_N1.

Non-limiting examples of the one or more improved properties of an engineered ancestral CYP2D enzyme as herein described include: thermal stability, stability in solvents (e.g. organic solvents), metabolite production, catalytic versatility, catalytic efficiency (e.g. the efficiency of coupling of product formation to cofactor
30 consumption), ligand binding capacity (e.g. increased or decreased strength of binding; and/or increased or decreased specificity of binding, as desired), substrate specificity (e.g. increased specificity or increased genericity, as desired) and enzyme kinetic properties (e.g. increased V_{max} , lower K_m), although without limitation thereto.

The 109 amino acid sequences set forth in SEQ ID NOS:432-540 (FIG. 24) are the amino acid sequences of certain isolated CPR reductase proteins of animals (i.e. “extant” CPR reductase proteins).

5 The amino acid sequence set forth in SEQ ID NO:321 is the amino acid sequence of an isolated engineered ancestral CPR protein, as set out in FIG. 3. SEQ ID NOS:322-431 (FIGS. 3 and 21) are particular amino acid sequences of variants of SEQ ID NO:321, comprising variations at one or more of the amino acid residues designated X_1 - X_{11} . The location of the residues X_1 - X_{11} in SEQ ID NO:321 with respect to SEQ ID NO:322 is given in table form in Table 11.

10 Certain preferred amino acids for each of the residues X_1 - X_{11} , respectively, are set forth in Table 11 under the heading ‘Variable’.

In one embodiment, one or more of the residues X_1 - X_{11} is an amino acid sequence selected from the respective groups consisting of the variable amino acids set forth in Table 11.

15 It will be appreciated that in some embodiments, the isolated protein comprising the amino acid sequence set forth in SEQ ID NO:321, and variants thereof, may be referred to as a protein having P450 reductase activity that is ancestral to each of the proteins set forth in SEQ ID NOS:432-540.

Suitably, the isolated engineered ancestral P450 reductase protein comprising
20 the sequence set forth in SEQ ID NO:321, and variants thereof, have one or more improved or enhanced properties compared to one or more of the isolated proteins comprising sequences set forth in SEQ ID NOS:432-540.

The particular engineered ancestral CPR enzyme variant comprising the amino acid sequence set forth in SEQ ID NO:322 is referred to herein as “CPR_N1”.
25 Additionally, the isolated proteins comprising the amino acid sequences set forth in SEQ ID NOS:323-431 are herein referred to as “CPR_N1 variants”.

The CPR_N1 variants set forth in SEQ ID NOS:323-431 are variants of CPR_N1 which represent nodes of the evolutionary tree calculated for construction of the engineered ancestral CPR protein as described in EXAMPLE 4.

30 In certain preferred embodiments, a CPR_N1 variant comprising an amino acid sequence set forth in SEQ ID NOS:323-431 has one or more improved or enhanced properties compared to CPR_N1.

Non-limiting examples of the one or more improved properties of an engineered ancestral CPR enzyme as herein described include: thermal stability, stability in

solvents (e.g. organic solvents), enzyme kinetic properties (e.g. increased $V_{\max}$, lower K_m), ligand binding capacity (e.g. increased or decreased strength of binding; and/or increased or decreased specificity of binding, as desired), substrate specificity (e.g. increased specificity or increased genericity, as desired), ability to couple to a diversity of proteins (e.g. P450s and/or other electron acceptor proteins), and ability to use a diversity of electron donors (e.g. both NADPH and NADH, or NADH preferentially as an electron donor), although without limitation thereto.

Certain embodiments also relate to fragments of the isolated proteins disclosed herein.

10 In one embodiment, a protein “*fragment*” includes an amino acid sequence that constitutes less than 100%, but at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90-99% of said isolated protein.

In another embodiment, a protein fragment comprises no more than 6, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 200, 250, 300, 350 or 400 contiguous amino acids of the isolated protein.

In a preferred embodiment, the protein fragment has one or more activities of a P450 enzyme or P450 reductase enzyme as hereinbefore described.

In certain embodiments, the protein fragment does not comprise an N-terminal “membrane anchor” region. By way of example, with specific regard to CYP3_N1, the N-terminal membrane anchor region comprises the amino acid positions 1-38 set forth in SEQ ID NO:2, as set forth in FIG. 28.

Certain embodiments also relate to variants of an isolated protein of the invention. In a preferred embodiment, the protein variant has one or more activities of a P450 or P450 reductase enzyme as hereinbefore described.

25 As used herein “*variant*” proteins of the invention have one or more amino acids deleted or substituted by different amino acids. It is well understood in the art that some amino acids may be substituted or deleted without an expectation of changing the activity of the protein substantially (“*conservative*” substitutions). More substantial changes to activity may be made by introducing substitutions or deletions that are less conservative (“*non-conservative*” substitutions).

Preferably, protein variants share at least 70% or 75%, preferably at least 80% or 85% or more preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with an amino acid sequence of the isolated protein.

In particular preferred embodiments relating to protein variants of engineered ancestral CYP3 enzymes as herein described, said protein variants share at least 70% or 75%, preferably at least 80% or 85% or more preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with at least one of the amino acid sequences set forth in SEQ ID NOS:2-41 or SEQ ID NOS:544-578.

In particular preferred embodiments relating to protein variants of engineered ancestral CYP2D enzymes as herein described, said variants share at least 70% or 75%, preferably at least 80% or 85% or more preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with at least one of the amino acid sequences set forth in SEQ ID NOS:181-250.

In particular preferred embodiments relating to protein variants of engineered ancestral CPR enzymes as herein described, said variants share at least 70% or 75%, preferably at least 80% or 85% or more preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with at least one of the amino acid sequences set forth in SEQ ID NOS:322-431.

In other preferred embodiments, protein variants share at least 80% or 85% or more preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with an amino acid sequence of the isolated protein, excluding the N-terminal membrane anchor region of said protein. By way of example, with specific regard to CYP3_N1, the N-terminal membrane anchor region comprises the amino acid positions 1-38 of the amino acid sequence set forth in SEQ ID NO:2, as set forth in Figure 28 and Table 9. Additionally, the N-terminal membrane anchor region of CYP2D_N1 comprises the amino acid positions 1-36 as set forth in Table 10.

As will be understood by one skilled in the art, the N-terminal membrane anchor region of a P450 enzyme or P450 reductase enzyme of the invention may be substantially modified without substantially affecting the function of said protein.

In certain preferred embodiments, the protein variant comprises a modified N-terminal “membrane anchor” region. By way of example, CYP3_N1 comprises the modified N-terminal sequence MALLLAVFL at amino acid positions 1-9. Such a modified N-terminal membrane anchor region may assist with protein expression, although without limitation thereto.

Terms used generally herein to describe sequence relationships between respective proteins and nucleic acids include “comparison window”, “sequence identity”, “percentage of sequence identity” and “substantial identity”. Because

respective nucleic acids/proteins may each comprise (1) only one or more portions of a complete nucleic acid/protein sequence that are shared by the nucleic acids/proteins, and (2) one or more portions which are divergent between the nucleic acids/proteins, sequence comparisons are typically performed by comparing sequences over a “comparison window” to identify and compare local regions of sequence similarity. A “*comparison window*” refers to a conceptual segment of typically 6, 9 or 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence for optimal alignment of the respective sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerised implementations of algorithms (Geneworks program by Intelligenetics; GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA, incorporated herein by reference) or by inspection and the best alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, Nucl. Acids Res. **25** 3389, which is incorporated herein by reference. A detailed discussion of sequence analysis can be found in Unit 19.3 of CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel *et al.* (John Wiley & Sons Inc NY, 1995-1999).

The term “*sequence identity*” is used herein in its broadest sense to include the number of exact nucleotide or amino acid matches having regard to an appropriate alignment using a standard algorithm, having regard to the extent that sequences are identical over a window of comparison. Thus, a “*percentage of sequence identity*” is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For example, “*sequence identity*” may be understood to mean the “match percentage” calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA).

Certain embodiments also relate to derivatives of an isolated protein of the present invention. In a preferred embodiment, the derivative protein has one or more activities of a P450 or P450 reductase enzyme as hereinbefore described.

As used herein, “*derivative*” proteins have been altered, for example by
5 conjugation or complexing with other chemical moieties, by post-translational modification (*e.g.* phosphorylation, acetylation *etc*), modification of glycosylation (*e.g.* adding, removing or altering glycosylation) and/or inclusion of additional amino acid sequences as would be understood in the art.

Additional amino acid sequences may include fusion partner amino acid
10 sequences which create a fusion protein. By way of example, fusion partner amino acid sequences may assist in detection and/or purification of the isolated fusion protein. Non-limiting examples include metal-binding (*e.g.* polyhistidine) fusion partners, maltose binding protein (MBP), Protein A, glutathione S-transferase (GST), fluorescent protein sequences (*e.g.* GFP), epitope tags such as myc, FLAG and haemagglutinin tags.

15 For the particular purpose of fusion polypeptide purification by affinity chromatography, relevant matrices for affinity chromatography include glutathione-, amylose-, and nickel- or cobalt-conjugated resins respectively. Many such matrices are available in kit form, such as the QIAexpressTM system (Qiagen) useful with (HIS₆) fusion partners and the Pharmacia GST purification system.

20 Preferably, the fusion partners also have protease cleavage sites, such as for Factor X_a or Thrombin, which allow the relevant protease to partially digest the fusion polypeptide of the invention and thereby liberate the recombinant polypeptide of the invention therefrom. The liberated polypeptide can then be isolated from the fusion partner by subsequent chromatographic separation.

25 Other derivatives contemplated by the invention include, but are not limited to, modification to amino acid side chains, incorporation of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the isolated protein, fragments and variants disclosed herein.

30

Methods for producing engineered ancestral proteins

In another aspect, the invention provides a method of producing or constructing an isolated protein, said method including the step of producing or constructing an engineered ancestral amino acid sequence of at least a fragment of a P450 protein or

P450 reductase protein from one or more P450 protein or P450 reductase amino acid sequences that are different to the engineered ancestral amino acid sequence.

Suitably, the engineered ancestral P450 protein or P450 reductase protein displays or possesses one or more increased or enhanced properties compared to at least one of the one or more P450 protein or P450 reductase amino acid sequences that are different to the engineered ancestral amino acid sequence.

Non-limiting examples of the one or more increased or enhanced properties of the protein include thermal stability, stability in solvents (e.g. organic solvents), metabolite production, catalytic versatility, catalytic efficiency (e.g. the efficiency of coupling of product formation to cofactor consumption), ligand binding capacity (e.g. increased or decreased strength of binding; and/or increased or decreased specificity of binding, as desired), substrate specificity (e.g. increased specificity or increased genericity, as desired), enzyme kinetic properties (e.g. increased $V_{\max}$, lower K_m), ability to couple to a diversity of proteins (e.g. P450s and/or other electron acceptor proteins), and ability to use a diversity of electron donors (e.g. both NADPH and NADH, or NADH preferentially as an electron donor), although without limitation thereto.

Suitably, the engineered ancestral protein having P450 or P450 reductase enzyme activity is distinct from any of a plurality of corresponding “*extant*” enzymes encoded by respective genomes of different organisms, such as those proteins used to construct or produce the engineered ancestral protein. In this regard, it will be appreciated by persons skilled in the art that a protein comprising an amino acid sequence comprised by the engineered ancestral protein having P450 or P450 reductase enzyme activity may or may not ever have actually existed until its construction according to the method of the invention.

In some embodiments of the method, the step of producing or constructing an engineered ancestral amino acid sequence of at least a fragment of a P450 protein or P450 reductase protein comprises the use of one or more computational methods for the reconstruction of engineered ancestral DNA and/or amino acid sequences. As will be understood by one skilled in the art, such methods can include “maximum parsimony” methods, “maximum likelihood” methods, and “Bayesian inference” methods. In some embodiments, said one or more computational methods include a “marginal likelihood” method and/or a “joint likelihood” method.

In some embodiments of the method, the construction of a sequence of a hypothetical engineered ancestral protein includes the use of one or more software tools for the reconstruction of engineered ancestral DNA and/or amino acid sequences. Said software tools may include, although without limitation thereto, FastML (Ashkenazy *et al.* 2012), Phylobayes 3 (Lartillot *et al.* 2009) and/or one or more of those listed at http://topicpages.ploscompbiol.org/wiki/Engineeredancestral_reconstruction#Software, incorporated herein by reference.

According to some embodiments of the method, “*evolutionary intermediate sequences*” may be calculated or constructed using one or more software tools in the process of constructing the engineered ancestral sequence, wherein said evolutionary intermediate sequences may or may not have ever existed prior to construction according to the method.

In some embodiments of the method, an evolutionary tree is constructed for the engineered ancestral P450 or P450 reductase protein using one or more of the aforementioned computational methods and/or software tools, and amino acid sequences which represent nodes within the evolutionary tree are identified. As used herein, a “*node*” within an evolutionary tree represents a common ancestor of the descendants that share or are linked by the node. In one embodiment, the isolated protein having P450 enzyme activity is an isolated engineered ancestral CYP3 protein as hereinbefore described. In one particular embodiment, the isolated protein comprises an amino acid sequence set forth in SEQ ID NO:1 or SEQ ID NO:2.

Non-limiting examples of the one or more P450 protein amino acid sequences that are different to said engineered ancestral amino acid sequence are set forth in SEQ ID NOS:42-179.

In another embodiment, the isolated protein having P450 enzyme activity is an isolated engineered ancestral CYP2D protein as herein described. In one particular embodiment, the isolated protein comprises an amino acid sequence set forth in SEQ ID NO:180 or SEQ ID NO:181.

Non-limiting examples of the one or more P450 protein amino acid sequences that are different to said engineered ancestral amino acid sequence are set forth in SEQ ID NOS:251-320.

In another embodiment, the isolated protein having P450 reductase enzyme activity is an isolated engineered ancestral CPR protein, as herein described. In one

particular embodiment, the isolated protein comprises an amino acid sequence set forth in SEQ ID NO:321 or SEQ ID NO:322.

Non-limiting examples of the one or more P450 protein amino acid sequences that are different to said engineered ancestral amino acid sequence are set forth in SEQ
5 ID NOS:432-540.

A related aspect provides a modified engineered ancestral P450 protein or P450 reductase protein produced according to the method of this aspect.

In yet another aspect, the invention provides a method of producing or constructing a modified engineered ancestral P450 protein or P450 reductase protein,
10 said method including the step of introducing one or more amino acid substitutions in an amino acid sequence of the engineered ancestral P450 protein or P450 reductase protein to thereby produce or construct the modified engineered ancestral P450 protein or P450 reductase protein. In certain preferred embodiments, said amino acid substitutions are non-conservative amino acid substitutions.

15 Suitably, the modified engineered ancestral P450 protein or P450 reductase protein displays or possesses one or more increased or enhanced properties compared to the engineered ancestral P450 protein or P450 reductase protein.

Non-limiting examples of the one or more improved or enhanced properties of the protein include thermal stability, stability in solvents (e.g. organic solvents),
20 metabolite production, catalytic versatility, catalytic efficiency (e.g. the efficiency of coupling of product formation to cofactor consumption), substrate specificity (e.g. increased specificity or increased genericity, as desired) and enzyme kinetic properties (e.g. increased V_{max} , lower K_m).

In certain preferred embodiments, said non-conservative amino acid
25 substitutions comprise functional or functionally important amino acids comprised by an engineered ancestral P450 protein or P450 reductase protein. In certain embodiments a functional or functionally important amino acid is one which is known or predicted to substantially affect a protein property or function, such as hereinbefore described.

In certain other preferred embodiments, said non-conservative amino acid
30 substitutions comprise amino acid positions wherein there is low prediction confidence within an engineered ancestral P450 protein or P450 reductase protein amino acid sequence constructed using one or more computational methods and/or software tools, or discrepancy between ancestral P450 protein or P450 reductase protein amino acid sequences constructed using one or more computational methods and/or software tools,

such as the computational methods and/or software tools as hereinbefore described. Said modifications may comprise amino acid positions wherein there is variation amongst a plurality of hypothetical engineered ancestral protein sequences constructed for the P450 or P450 reductase enzyme using different software tools and/or computational methods, although without limitation thereto.

By way of example, as set forth in Table 11, the positions X₁-X₁₁ of the generic engineered ancestral CPR protein sequence set forth in SEQ ID NO:322 represent amino acid positions at which there is variation amongst a sequence constructed using a 'Joint Maximum Likelihood' and 'Marginal Maximum Likelihood' method, as hereinabove described. In this respect, CPR_N1 (SEQ ID NO:321) represents the sequence constructed using the joint likelihood method.

In certain other preferred embodiments, said non-conservative amino acid substitutions comprise amino acid positions wherein there is variation amongst extant and/or evolutionary intermediate sequences used to construct a hypothetical engineered ancestral protein sequence, and/or sequences which represent nodes within an evolutionary tree constructed for the hypothetical engineered ancestral P450 or P450 reductase protein, as hereinbefore described.

By way of example, as set forth in Table 9, the positions X₁-X₂₁₉ of the generic engineered ancestral CYP3 protein sequence set forth in SEQ ID NO:1 represent amino acid positions at which there is variation amongst corresponding extant and/or evolutionary intermediate sequences used to construct the engineered ancestral CYP3 protein, and/or sequences which represent nodes within the evolutionary tree constructed for the engineered ancestral CYP3 protein.

Similarly, as set forth in Table 10, the positions X₁-X₁₆₃ of the generic engineered ancestral CYP2D protein sequence set forth in SEQ ID NO:180 represent amino acid positions at which there is variation amongst corresponding extant and/or evolutionary intermediate sequences used to construct the engineered ancestral CYP2D protein, and/or sequences which represent nodes within the evolutionary tree constructed for the engineered ancestral CYP2D protein.

In certain preferred embodiments, the modified engineered ancestral P450 protein is a modified engineered ancestral CYP3 protein. In one embodiment, the modified engineered ancestral CYP3 protein enzyme comprises an amino acid sequence set forth in SEQ ID NO:1 or any one of SEQ ID NOS:3-41 or SEQ ID NOS:544-578.

In other embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:3-41 or SEQ ID NOS:544-578, including at least 85% identical, at least 90% identical, and at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, and 99% identical.

- 5 In certain other preferred embodiments, the modified engineered ancestral P450 protein is a modified engineered ancestral CYP2D protein. In one embodiment, the modified engineered ancestral CYP2D protein comprises an amino acid sequence set forth in SEQ ID NO:180 or any one of SEQ ID NOS:182-250.

- 10 In other embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:182-250, including at least 85% identical, at least 90% identical, and at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, and 99% identical.

- 15 In certain other preferred embodiments, the modified engineered ancestral P450 reductase protein is a modified engineered ancestral CPR protein. In one embodiment, the modified engineered ancestral CPR protein comprises an amino acid sequence set forth in SEQ ID NO:321 or any one of SEQ ID NOS:323-431.

- 20 In other embodiments the isolated protein comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOS:323-431, including at least 85% identical, at least 90% identical, and at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, and 99% identical.

In certain embodiments, an isolated engineered ancestral protein of the invention has enhanced thermal stability compared to one or more corresponding extant animal proteins.

- 25 In certain embodiments, thermal stability is the ability of a protein to resist “thermal deactivation”. As used herein, “*thermal deactivation*” refers to the loss of a properly folded state of a protein as a result of exposure to heat, and/or the loss of enzyme activity of a protein as a result of exposure to heat.

- 30 A measure of thermal stability may be expressed as a $^xT_{50}$ temperature ($^{\circ}C$) or “T50” for folded protein, wherein 50% of protein is properly folded after exposure to said temperature for a duration of x minutes. For example, the $^{60}T_{50}$ for CYP3_N1 is about 66 ± 2 (Table 1). Another measure of thermal stability may be expressed as a $^xT_{50}$ temperature ($^{\circ}C$) or T50 for enzyme activity, wherein there is 50% of enzyme activity after exposure to said temperature for a duration of x minutes.

Thermal stability may also be measured as the percentage of properly folded protein (%) after exposure to a given temperature (°C) for a given duration (min), and/or the percentage of enzyme activity (%) after exposure to a given temperature (°C) for a given duration (min).

5 Yet another measure of thermal stability may be “*thermal deactivation energy*” (E_a) which can be defined as the minimum energy that is required to cause thermal deactivation of a protein. For example, E_a for CYP3_N1 is 251 kJ/mol (Table 3).

10 In one particular embodiment, the isolated engineered ancestral CYP3_N1 protein of the invention has enhanced thermal stability compared to one or more of the isolated proteins comprising SEQ ID NOS:42-179. As will be evident from FIG. 10 and EXAMPLE 5 herein, isolated CYP3_N1 displayed substantially greater thermal stability than extant animal CYP3 proteins.

15 Additionally, it will be evident from EXAMPLE 10 that the CYP3_N1 variants comprising amino acid sequences set forth in SEQ ID NOS:3-28 displayed thermostability similar or greater than CYP3_N1.

20 In another particular embodiment, the isolated CYP2D_N1 protein of the invention has enhanced thermal stability compared to one or more of the isolated proteins comprising SEQ ID NOS:251-320. As will be evident from FIG. 11 and EXAMPLE 5 herein, isolated CYP2D_N1 protein displayed substantially greater thermal stability than an extant animal CYP2D protein.

25 In another particular embodiment, the isolated CPR_N1 protein of the invention has enhanced thermal stability compared to one or more of the isolated proteins comprising SEQ ID NOS:432-540. As will be evident from FIG. 14 and EXAMPLE 5 herein, isolated CPR_N1 protein displayed substantially greater thermal stability than extant animal CPR proteins.

In certain particular embodiments, an isolated engineered ancestral protein of the invention has enhanced solvent stability compared to one or more of the corresponding extant animal proteins.

30 One measure of solvent stability may be the ability of a protein to adopt a properly folded state at a given concentration of solvent. Another measure of solvent stability may be the enzyme activity of a protein at a given concentration of solvent.

Preferably, said solvent is an organic solvent including, but not limited to: a polar protic solvent including an alcohol, e.g. ethanol, methanol, and isopropanol; a polar aprotic solvent e.g. dimethyl sulfoxide (DMSO), and acetonitrile; and a non-polar

solvent including saturated and unsaturated hydrocarbon molecules and aromatic molecules.

In one particular embodiment, the isolated engineered ancestral CYP3_N1 molecule of the invention has enhanced stability in an organic solvent comprising methanol, compared to one or more of the isolated proteins comprising SEQ ID NOS:42-179.

As will be evident from EXAMPLE 7 and FIG. 15 and FIG. 16, CYP3_N1 displayed substantially higher velocity in methanol compared to extant animal CYP3 proteins.

10 In another particular embodiment, the isolated engineered ancestral CYP3_N1 molecule of the invention has enhanced stability in an organic solvent comprising acetonitrile, compared to one or more of the isolated proteins comprising SEQ ID NOS:42-179.

As will be evident from EXAMPLE 7 and FIG. 15 and FIG. 16, CYP3_N1
15 displayed substantially increased velocity in solvents comprising various concentrations of acetonitrile, compared to extant animal CYP3 proteins.

In certain other embodiments, an isolated engineered ancestral protein of the invention has altered binding affinity for one or more ligands as compared to one or more corresponding extant animal proteins.

20 One measure of ligand binding may be a “dissociation constant” (K_d) in molar units. As will be understood by those skilled in the art, K_d indicates the concentration of ligand at which the binding site on a particular protein is half occupied. Therefore, it will be understood that a lower K_d value indicates a higher binding affinity of a protein for a given ligand.

25 In some preferred embodiments, said one or more ligands comprise a macrolide (e.g. azithromycin, cyclosporin A, troleandomycin, clarithromycin, erythromycin, and telithromycin) and/or a benzodiazepine (e.g. 2-keto benzodiazepines such as diazepam, and imidazo benzodiazepines such as midazolam, although without limitation thereto).

In yet other preferred embodiments, an isolated engineered ancestral protein of
30 the invention demonstrates altered metabolism of chemicals, such as drugs, as compared to one or more corresponding extant animal proteins.

As will be evident from EXAMPLE 6 and FIGS. 31-37, CYP3_N1 demonstrated altered metabolism of tamoxifen, erythromycin, and ticlopidine compared to extant CYP3A4.

In some embodiments, an isolated engineered ancestral protein of the invention may demonstrate more rapid conversion of a chemical, such as a drug, to a metabolite of that chemical; and/or result in a greater conversion of a chemical, such as a drug, to a metabolite of that chemical at completion (or near completion).

5 As set forth in EXAMPLE 6 and FIG. 31, as compared to CYP3A4, CYP3_N1 demonstrated more rapid metabolism of tamoxifen to its major demethylated metabolite, and resulted in a greater conversion of tamoxifen to its major demethylated metabolite after 120 minutes.

10 Furthermore, as set forth in EXAMPLE 6 and FIG. 35-36, as compared to CYP3A4, CYP3_N1 demonstrated more rapid metabolism of ticlopidine to its desaturated metabolite, and resulted in a greater conversion of ticlopidine to its desaturated metabolite after 120 minutes. In certain other embodiments, the activity of an isolated protein of the invention in a reaction comprising a given substrate produces a different metabolite profile, as compared to the activity of one or more extant animal
15 proteins in a corresponding reaction comprising said substrate.

In certain embodiments, the different metabolite profile may comprise one or more metabolites that are present or absent, and/or a relative increase or decrease in one or more metabolites.

20 **Isolated nucleic acids**

Another aspect of the invention provides an isolated nucleic acid that encodes an isolated protein of the invention, inclusive of fragments, variants and derivatives of the isolated protein.

25 For embodiments of the aspect relating to isolated CYP3 proteins, said nucleic acid is exemplified in SEQ ID NO:541 (FIG. 25).

For embodiments of the aspect relating to isolated CYP2D proteins, said nucleic acid is exemplified in SEQ ID NO:542 (FIG. 25).

For embodiments of the aspect relating isolated CPR proteins, said nucleic acid is exemplified in SEQ ID NO:543 (FIG. 25).

30 The term “*nucleic acid*” as used herein designates single-or double-stranded DNA and RNA. DNA includes genomic DNA and cDNA. RNA includes mRNA, RNA, RNAi, siRNA, cRNA and autocatalytic RNA. Nucleic acids may also be DNA-RNA hybrids. A nucleic acid comprises a nucleotide sequence which typically includes nucleotides that comprise an A, G, C, T or U base. However, nucleotide sequences may

include other bases such as inosine, methylcytosine, methylinosine, methyladenosine and/or thiouridine, although without limitation thereto.

A “*polynucleotide*” is a nucleic acid having eighty (80) or more contiguous nucleotides, while an “*oligonucleotide*” has less than eighty (80) contiguous
5 nucleotides.

A “*probe*” may be a single or double-stranded oligonucleotide or polynucleotide, suitably labelled for the purpose of detecting complementary sequences in Northern or Southern blotting, for example.

A “*primer*” is usually a single-stranded oligonucleotide, preferably having 15-50
10 contiguous nucleotides, which is capable of annealing to a complementary nucleic acid “template” and being extended in a template-dependent fashion by the action of a DNA polymerase such as *Taq* polymerase, RNA-dependent DNA polymerase or SequenaseTM.

Another particular aspect of the invention provides a variant of an isolated
15 nucleic acid that encodes an isolated protein, variant, fragment or derivative disclosed herein. In a preferred embodiment, nucleic acid variants share at least 60% or 65%, preferably at least 70% or 75%, more preferably at least 80% or 85%, and even more preferably at least 90% or 95% nucleotide sequence identity with an isolated nucleic acid that encodes one or more of the isolated proteins of the invention.

20 In yet another embodiment, nucleic acid variants hybridize to isolated nucleic acids of the invention, under at least low stringency conditions, preferably under at least medium stringency conditions and more preferably under high stringency conditions.

“*Hybridize and Hybridization*” is used herein to denote the pairing of at least partly complementary nucleotide sequences to produce a DNA-DNA, RNA-RNA or
25 DNA-RNA hybrid. Hybrid sequences comprising complementary nucleotide sequences occur through base-pairing between complementary purines and pyrimidines as are well known in the art.

In this regard, it will be appreciated that modified purines (for example, inosine, methylinosine and methyladenosine) and modified pyrimidines (thiouridine and
30 methylcytosine) may also engage in base pairing.

“*Stringency*” as used herein, refers to temperature and ionic strength conditions, and presence or absence of certain organic solvents and/or detergents during hybridisation. The higher the stringency, the higher will be the required level of complementarity between hybridizing nucleotide sequences.

“*High stringency conditions*” designates those conditions under which only nucleic acid having a high frequency of complementary bases will hybridize.

Reference herein to high stringency conditions include and encompass:-

- 5 (i) from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridisation at 42°C, and at least about 0.01 M to at least about 0.15 M salt for washing at 42°C;
- 10 (ii) 1% BSA, 1 mM EDTA, 0.5 M NaHPO₄ (pH 7.2), 7% SDS for hybridization at 65°C, and (a) 0.1 x SSC, 0.1% SDS; or (b) 0.5% BSA, 1mM EDTA, 40 mM NaHPO₄ (pH 7.2), 1% SDS for washing at a temperature in excess of 65°C for about one hour; and
- (iii) 0.2 x SSC, 0.1% SDS for washing at or above 68°C for about 20 minutes.

15 In general, washing is carried out at $T_m = 69.3 + 0.41 (G + C) \% - 12^\circ\text{C}$. In general, the T_m of a duplex DNA decreases by about 1°C with every increase of 1% in the number of mismatched bases.

Notwithstanding the above, stringent conditions are well known in the art, such as described in Chapters 2.9 and 2.10 of. Ausubel *et al.*, *supra*. A skilled addressee will
20 also recognize that various factors can be manipulated to optimize the specificity of the hybridization. Optimization of the stringency of the final washes can serve to ensure a high degree of hybridization.

In other embodiments, isolated nucleic acid variants may be produced using a nucleic acid amplification technique. Suitable nucleic acid amplification techniques are
25 well known to the skilled addressee, and include polymerase chain reaction (PCR); strand displacement amplification (SDA); rolling circle replication (RCR); nucleic acid sequence-based amplification (NASBA), Q-β replicase amplification and helicase-dependent amplification, although without limitation thereto.

As used herein, an “*amplification product*” refers to a nucleic acid product
30 generated by nucleic acid amplification.

Particularly for analytical purposes, nucleic acid amplification techniques may include quantitative and semi-quantitative techniques such as qPCR, real-time PCR and competitive PCR, as are well known in the art.

Suitably, isolated nucleic acid variants may be produced using nucleic acid amplification techniques using one or more degenerate primers based on, or derived from, a nucleotide sequence of an isolated nucleic acid disclosed herein. By way of example, the degenerate primer(s) may be designed to anneal to one or more nucleotide sequences of a variant nucleic acid to thereby facilitate amplification of the variant nucleic acid, or a fragment thereof.

Genetic constructs

Yet another aspect of the invention provides a genetic construct that comprises an isolated nucleic acid or variant as herein described and one or more additional nucleotide sequences.

Suitably, the genetic construct may be in the form of, or comprise genetic components of, a plasmid, bacteriophage, a cosmid, or a yeast or bacterial artificial chromosome as are well understood in the art.

Genetic constructs may be suitable for maintenance and propagation of the isolated nucleic acid in bacteria or other host cells, for manipulation by recombinant DNA technology and/or expression of the nucleic acid or an encoded protein of the invention.

For the purposes of host cell expression, the genetic construct is an expression construct. Suitably, the expression construct comprises one or more nucleic acid or variants disclosed herein operably linked to one or more additional sequences in an expression vector.

An “*expression vector*” may be either a self-replicating extra-chromosomal vector such as a plasmid, or a vector that integrates into a host genome.

By “*operably linked*” is meant that said additional nucleotide sequence(s) is/are positioned relative to the nucleic acid of the invention preferably to initiate, regulate or otherwise control transcription.

In one embodiment, the additional nucleotide sequences are regulatory sequences. Regulatory nucleotide sequences will generally be appropriate for the host cell used for expression. Numerous types of appropriate expression vectors and suitable regulatory sequences are known in the art for a variety of host cells.

Typically, said one or more regulatory nucleotide sequences may include, but are not limited to, promoter sequences, leader or signal sequences, ribosomal binding

sites, transcriptional start and termination sequences, translational start and termination sequences, and enhancer or activator sequences.

Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be either naturally occurring promoters, or hybrid
5 promoters that combine elements of more than one promoter.

In another embodiment, the additional nucleotide sequence is a selectable marker gene to allow the selection of transformed host cells. Selectable marker genes are well known in the art and will vary with the host cell used.

The expression construct may also include an additional nucleotide sequence
10 encoding a fusion partner (typically provided by the expression vector) so that the recombinant polypeptide of the invention is expressed as a fusion protein, as hereinbefore described.

Isolated proteins of the invention (inclusive of fragments, derivatives and homologs) may be prepared by any suitable procedure known to those of skill in the art.
15 Preferably, the isolated protein is a recombinant protein.

By way of example only, a recombinant isolated protein of the invention may be produced by a method including the steps of:

- (i) preparing an expression construct which comprises an isolated
20 nucleic acid of the invention, operably linked to one or more regulatory nucleotide sequences;
- (ii) transfecting or transforming a suitable host cell with the expression construct;
- (iii) expressing a recombinant protein in said host cell; and
- (iv) isolating the recombinant protein from said host cell.

25 Suitable host cells for expression may be prokaryotic or eukaryotic. For example, suitable host cells may be mammalian cells, plant cells, yeast cells, insect cells or bacterial cells. One preferred host cell for expression of an isolated protein according to the invention is a bacterium.

Introduction of genetic constructs into host cells (whether prokaryotic or
30 eukaryotic) is well known in the art, as for example described in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel *et al.*, (John Wiley & Sons, Inc. 1995-2009), in particular Chapters 9 and 16.

The recombinant protein may be conveniently prepared by a person skilled in the art using standard protocols as for example described in Sambrook, *et al.*,

MOLECULAR CLONING. A Laboratory Manual (Cold Spring Harbor Press, 1989), in particular Sections 16 and 17; CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel *et al.*, (John Wiley & Sons, Inc. 1995-2009), in particular Chapters 10 and 16; and CURRENT PROTOCOLS IN PROTEIN SCIENCE Eds. Coligan *et al.*, (John Wiley & Sons, Inc. 1995-2009), in particular Chapters 1, 5 and 6.

Antibodies

Another aspect of the invention provides an antibody or antibody fragment which binds, or has been raised against, an isolated protein disclosed herein.

10 In certain preferred embodiments wherein the isolated protein is an engineered ancestral CYP3 protein, the antibody or antibody fragment shows at least partial specificity for an isolated protein comprising any one of the amino acid sequences set forth in SEQ ID NOS:2:41 or SEQ ID NOS:544-578. Suitably, said isolated protein does not bind, or demonstrates substantially reduced binding against, one or more of the
15 isolated proteins set forth in SEQ ID NOS:42-179.

In certain preferred embodiments wherein the isolated protein is an engineered ancestral CYP2D protein, the antibody or antibody fragment shows at least partial specificity for an isolated protein comprising any one of the amino acid sequences set forth in SEQ ID NOS:182-250. Suitably, said antibody or antibody fragment does not
20 bind, or demonstrates substantially reduced binding against, one or more of the isolated proteins comprising the amino acid sequences SEQ ID NOS:251-320.

In certain preferred embodiments wherein the isolated protein is an engineered ancestral CPR protein, the antibody or antibody fragment shows at least partial specificity for an isolated protein comprising any one of the amino acid sequences set forth in 322-431. Suitably, the antibody or antibody fragment does not bind, or demonstrates substantially reduced binding against, one or more of the isolated proteins comprising the amino acid sequences SEQ ID NOS:432-540.

As used herein an “antibody” is or comprises an immunoglobulin. The term “immunoglobulin” includes any antigen-binding protein product of a mammalian
30 immunoglobulin gene complex, including immunoglobulin isotypes IgA, IgD, IgM, IgG and IgE and antigen-binding fragments thereof. Included in the term “immunoglobulin” are immunoglobulins that are chimeric or humanised or otherwise comprise altered or variant amino acid residues, sequences and/or glycosylation, whether naturally occurring or produced by human intervention (*e.g.* by recombinant DNA technology).

Antibody fragments include Fab and Fab'2 fragments, diabodies, triabodies and single chain antibody fragments (e.g. scVs), although without limitation thereto. Typically, an antibody comprises respective light chain and heavy chain variable regions that each comprise CDR 1, 2 and 3 amino acid sequences. A preferred antibody
5 fragment comprises at least one light chain variable region CDR and/or at least one heavy chain variable region CDR.

Antibodies and antibody fragments may be polyclonal or preferably monoclonal. Monoclonal antibodies may be produced using the standard method as for example, described in an article by Köhler & Milstein, 1975, Nature **256**, 495, or by
10 more recent modifications thereof as for example described in Chapter 2 of Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, by immortalizing spleen or other antibody producing cells derived from a production species which has been inoculated with an isolated protein or a fragment thereof. It will also be appreciated that antibodies may be produced as recombinant synthetic antibodies or antibody fragments for
15 example by expressing a nucleic acid encoding the antibody or antibody fragment in an appropriate host cell. Recombinant synthetic antibody or antibody fragment heavy and light chains may be co-expressed from different expression vectors in the same host cell or expressed as a single chain antibody in a host cell. Non-limiting examples of recombinant antibody expression and selection techniques are provided in Chapter 17 of
20 Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY and Zuberbuhler *et al.*, 2009, Protein Engineering, Design & Selection **22** 169.

In some embodiments, the antibody or antibody fragment is labelled.

The label may be selected from a group including a chromogen, a catalyst, biotin, digoxigenin, an enzyme, a fluorophore, a chemiluminescent molecule, a
25 radioisotope, a drug or other chemotherapeutic agent, a magnetic bead and/or a direct visual label.

It will be appreciated that the antibody or antibody fragment may be used for the detection and/or purification of an isolated protein disclosed herein.

30 **Methods of use of the isolated protein**

Another aspect of the invention relates to a method of performing a chemical reaction, said method including the step of exposing a molecule to one or more isolated proteins disclosed herein to thereby perform a chemical reaction in the molecule.

An aspect of the invention also provides a composition suitable for performing the chemical reaction. The composition suitably comprises one or more buffers, salts, solvents and/or other reagents that facilitate or allow the reaction to proceed. It will be understood that such pH buffers, salts, solvents and/or other reagents are well known in the art and may be selected according to the particular type of chemical reaction.

Suitably, the method includes exposing the molecule to a protein having P450 activity and/or a redox partner for the protein having P450 activity.

In some preferred embodiments, said method includes the step of exposing the molecule to at least one of:

- i) an isolated engineered ancestral CYP3 protein disclosed herein;
- ii) an isolated engineered ancestral CYP2D protein disclosed herein; and
- iii) an isolated CPR protein disclosed herein.

In preferred embodiments of the method wherein the protein having P450 activity comprises an isolated engineered ancestral CYP3 protein disclosed herein, the redox partner for said CYP3 protein comprises an isolated engineered ancestral CPR protein disclosed herein.

In other embodiments, the redox partner for said CYP3 protein comprises any other suitable redox partner(s), which may include, but is not limited to one or more of: cytochrome b5 and cytochrome b5 reductase; a ferredoxin (e.g. adrenodoxin); and a ferredoxin reductase (e.g. adrenodoxin reductase) or flavodoxin reductase.

In a preferred embodiment of the method wherein the protein having P450 activity comprises an isolated engineered ancestral CYP2D protein disclosed herein, the redox partner for said CYP2D protein comprises an isolated engineered ancestral CPR protein disclosed herein.

In other embodiments of the method wherein the protein having P450 activity comprises an isolated engineered ancestral CYP2D protein disclosed herein, the redox partner for said CYP2D protein comprises any other suitable redox partner(s), which may include, but is not limited to one or more of cytochrome b and cytochrome b5 reductase; and ferredoxin (e.g. adrenodoxin) and ferredoxin reductase (e.g. adrenodoxin reductase) or flavodoxin reductase.

In a preferred embodiment of the method wherein the redox partner for a protein having P450 reductase activity comprises an isolated engineered ancestral CPR protein disclosed herein, the protein having P450 activity comprises an isolated engineered ancestral CYP3 protein disclosed herein.

In another preferred embodiment of the method wherein the redox partner for a protein having P450 activity comprises the CPR protein disclosed herein, the protein having P450 activity comprises an isolated engineered ancestral CYP2D protein disclosed herein.

5 In other embodiments wherein the redox partner for a protein having P450 activity is an isolated engineered ancestral CPR protein disclosed herein, the protein having P450 activity comprises any other suitable P450 protein, which may include one or more animal P450 proteins, e.g. CYP1, CYP2, CYP3, CYP4, CYP5, CYP6, CYP7, CYP8, CYP9, CYP11, CYP12, CYP17, CYP19, CYP20, CYP21, CYP24, CYP26,
10 CYP27, CYP39, CYP46, and CYP51 animal P450 enzyme classes; microbial P450 proteins, e.g. CYP101, CYP107A1, and CYP119 microbial P450 enzyme classes; and/or plant P450 proteins, e.g. CYP51, CYP74, CYP97, CYP710, CYP711, CYP727, CYP746 plant P450 enzymes classes, although without limitation thereto.

Particular, non-limiting examples of such chemical reactions include redox
15 reactions, hydroxylation at aromatic and aliphatic centres, epoxidation, N-, O-, S-dealkylation and N-, and S-oxidation, acyl migration, oxidative dehalogenation, ring expansion, contraction and cleavage, C-C bond cleavage, denitrosation of N-nitrosamines, oxidative ester cleavage, aldehyde scissions (*e.g.* to alkenes and HCOOH), ipso attack on aromatic ring substituents and N- or O-dearylation, reductions of alkyl
20 halides, N-oxides, nitro compounds, inorganic molecules such as SO₂, Cr(VI) or NO, desaturations (*e.g.* dehydrogenations), one electron oxidations, isomerizations and/or phospholipase D activity (*e.g.* phosphate ester hydrolysis).

These reactions may have applications for: the production of fine chemicals (*e.g.* pharmaceuticals, agrichemicals, fragrances, and dyes); gene therapy; bioremediation;
25 biosensors; diagnostics; plant biotechnology; and/or medicinal chemistry (*e.g.* drug discovery and pharmacological testing). With particular regard to drug discovery, an isolated protein disclosed herein may be used for structural diversification of molecules present in molecular libraries, such as natural product libraries, synthetic combinatorial libraries, and/or rationally designed structure-based libraries, although without
30 limitation thereto.

As will be understood by those skilled in the art, structural diversification may facilitate structure-activity relationship ("SAR") analysis to thereby generate an array of improved lead compounds having one or more improved properties such as: enhanced pharmacodynamic properties; reduced toxicity; enhanced bioavailability; enhanced half-

life; enhanced formulation properties; and/or reduced production cost, although without limitation thereto. As will be understood by those skilled in the art, in the aforementioned context a “*lead*” compound refers to a chemical compound that has pharmacological or biological activity likely to be useful for a given purpose; for example, for therapeutic and/or industrial application, although without limitation thereto; but may still have suboptimal properties for said purpose.

With particular regard to pharmacological testing, an isolated protein described herein may be used to metabolize a xenobiotic, for example a drug, to produce metabolites that are produced during human metabolism of said xenobiotic. It is anticipated that said metabolites can be used for purposes including the assessment of toxicity to an animal, for example “general toxicity”, “genotoxicity”, “embryo-fetal toxicity”, as will be understood by those skilled in the art; and “carcinogenicity”, although without limitation thereto.

With particular regard to bioremediation, an isolated protein described herein may be used to metabolize an environmental pollutant, for example, although without limitation thereto, a hydrocarbon pollutant such as a diesel, a gasoline, or an oil; a pesticide pollutant such as an organochlorine pesticide, an organophosphate pesticide, or a pyrethroid pesticide; and a solvent pollutant such as perchloroethylene (PCE) or trichloroethylene (TCE). In some embodiments, the isolated protein may be expressed by a microorganism or a plant, thereby allowing said microorganism or plant to metabolize an environmental pollutant, or improving the efficiency with which said microorganism or plant metabolizes an environmental pollutant.

With particular regard to biosensor technology, an isolated protein described herein may be used to detect metabolites in a human blood sample, for example, drugs or drug metabolites; and/or food contaminants, such as carbamate or organophosphate pesticide residues, although without limitation thereto.

EXAMPLES

EXAMPLE 1. *Ancestral sequence reconstruction of the cytochrome P450 family 3 (CYP3).*

A total of 138 CYP3 sequences from 42 animal species were collected from Uniprot, NCBI and the cytochrome P450-nomenclature homepage (http://dnelson.utm.edu/Cytochrome_P450.html) database. The sequences were aligned using MAFFT (Kato *et al.* 2002). The alignment was fine-tuned manually to

improve its reliability at gap positions. Phylogenetic relationships of these CYP3 sequences were reconstructed using the Maximum Likelihood method under the JTT substitution model (Jones *et al.* 2002) with Phyml using T-REX (Alix *et al.* 2012). The initial tree was determined by neighbour-joining (BIONJ) (Gascuel 1997).
5 Bootstrapping analysis was performed to evaluate the tree. Prediction of ancestral nodes of the tree (SEQ ID NOS:544-578) was performed using FastML (Pupko *et al.* 2000). The last common ancestor of CYP3 was identified and designated 'N1'.

A codon-optimized nucleotide sequence for this N1 amino acid sequence was synthesized by GeneArt (Germany), in which the first 19 codons of N1 were replaced
10 with the sequence 5' ATG GCT CTG TTA TTA GCA GTT TTT CTG 3' encoding the peptide, MALLLAVFL, which was known to facilitate the expression of P450 CYP3A4 in *E. coli* (Gillam *et al.* 1993). This nucleotide sequence encodes CYP3_N1, set forth in SEQ ID NO:2. A C-terminal hexa-His tag was also added to the coding sequence of the engineered ancestral protein linked to the C-terminal end of the P450 coding sequence
15 by a Ser-Thr linker encoding a SalI site.

EXAMPLE 2. *Construction of CYP3_N1 expression plasmids.*

The CYP3_N1 open reading frame was subcloned from the cloning vector, pMA_RQ/3_N1 (GeneArt, Germany) into the pCW expression vector (Muchmore *et al.*
20 1989) containing the human NADPH-cytochrome P450 reductase (hNPR) sequence, using 5' NdeI and 3' SalI sites (FIG. 4). The sequence of the CYP3_N1 open reading frame in the expression plasmid was verified by automated dideoxy sequencing at the Brisbane Node of the Australian Genome Research Facility (St. Lucia, Australia). Subsequently, a monocistronic expression construct (lacking the hNPR open reading
25 frame) and a CYP3_N1/enhanced yellow fluorescent protein (EYFP) fusion construct (retaining the hCPR expression cassette) were constructed by the strategies shown in FIG. 5 and FIG. 6. For the monocistronic construct, a synthetic, double-stranded, oligonucleotide linker containing XbaI, SmaI, BlnI, SacI, NsiI, HindIII, and NheI restriction sites was first ligated into the bicistronic construct (pCW/3_N1His/hNPR) to
30 facilitate the subsequent removal of hNPR via BlnI digestion followed by relegation of the linearized monocistronic vector (FIG. 5). The CYP3_N1 EYFP fusion was obtained by ligating the EYFP fragment from pCW/2C19_{FL}-EYFPHis/hNPR digested with SalI into the bicistronic construct (FIG. 6).

EXAMPLE 3. *Ancestral sequence reconstruction of the CYP2D subfamily (CYP2D)*

A total of 70 CYP2D sequences from vertebrate species were collected from Uniprot, NCBI and the cytochrome P450-nomenclature homepage (http://dnelson.utm.edu/Cytochrome_P450.html) database. The sequences were aligned using MEGA 6 (Beta 2) (Tamura *et al.* 2013), using a Multiple Sequence Comparison by Log-Expectation (MUSCLE) alignment with the following parameters: gap open: -2.9, gap extend: -1.01, hydrophobicity multiplier: 1.2. The alignment was fine-tuned manually to improve its reliability at gap positions. Phylogenetic relationships of these CYP2D sequences were reconstructed using the Maximum Likelihood method under the JTT substitution model (Jones *et al.* 1992) with Phyml using T-REX (Alix *et al.* 2012). The initial tree was determined by neighbour-joining (BIONJ) (Gascuel 1997). Bootstrapping analysis was performed to evaluate the tree. Prediction of ancestral nodes of the tree (SEQ ID NOS:182-250) was performed using FastML (Pupko *et al.* 2000) and the last common ancestor of CYP2D and the ancestor of the mammalian CYP2D forms were identified, with the last common ancestor designated 'N1'. A nucleotide sequence encoding the full length (FL) CYP2D_N1 amino acid sequence (which amino acid sequence is set forth in SEQ ID NO:181) was designed in which the second codon was changed to GCT (Ala) and the first 12 codons were optimised to minimise mRNA secondary structure formation (Gillam *et al.* 1995) without further alteration of the amino acid sequence. Sequences were analysed from -20 bases preceding to +96 following the start codon (Goodman *et al.* 2013) using the online software Nupack (Zadeh *et al.* 2011) to calculate free energy values.

A C-terminal hexa-His tag was also added to the coding sequence of the engineered ancestral protein linked to the C-terminal end of the P450 coding sequence by a Ser-Thr linker encoding a SalI site. Two 80 bp flanking sequences complementary to the destination expression vector were added to the open reading frame for this FL sequence, and the resultant sequence was synthesised as a GeneBlock by Integrated DNA Technologies (Singapore). As set forth in FIG. 7, the final bicistronic expression vector was generated by Gibson Assembly (Zadeh *et al.* 2011) of this Geneblock with the backbone of the pCW'/2D22/hNPR vector from which the CYP2D22 insert had been removed by digestion with NdeI and XbaI.

A truncated version of the CYP2D_N1, as set forth in FIG. 8, was generated by PCR amplification using mutagenic primers; Forward (2DN1MAKFWD), 5'AGGTCATATGGCAAAAAAACATCATCAAAGGA

AAATTCCCACCAGGCCCTATGTCATT3' and Reverse (hNPR5'REV),
5'GACACGGTGGAGCTGGTGTCCACGTGG3'. As set forth in FIG. 8, the
MAKKTSSKGK leader sequence (von Wachenfeldt *et al.* 1997; Rowland *et al.* 2006)
was added to the coding sequence upstream of the proline-rich region. The PCR product
5 was digested with NdeI and XbaI and cloned into the cognate sites of
pCW'/2D22/hNPR to generate the pCW'/2D_N1trunc/hNPR bicistronic expression
vector.

EXAMPLE 4. *Ancestral sequence reconstruction of the NADPH-cytochrome P450
10 reductase (CPR)*

CPR sequences were obtained from Uniprot, and by BLASTing the Uniprot sequences
against the NCBI (Altschul *et al.* 1990) database to retrieve additional sequences of high
homology, then aligned using the CLUSTALW method (Kyoto University
Bioinformatics Center; <<http://www.genome.jp/>>) Anomalous sequences were removed
15 (e.g. obvious sequencing errors, clearly incomplete open reading frames, etc.), realigned
and minor allelic variants and sequences that deviate significantly from the family
characteristics (e.g. <55% sequence identity to all other forms, either globally or over
any section of ~20 residues) were pruned such that the ultimate alignment included only
one sequence per enzyme. The remaining 109 sequences were then aligned with the
20 house fly CPR sequence (designated as the outgroup sequence) and the evolutionary
tree derived in MEGA (Tamura *et al.* 2013) by the ML method, with the designated
outgroup as the root. (ML has been shown to be more accurate than MP; Gadagkar *et al.* 2005). This tree was then imported along with the alignment into the FastML web
server (Ashkenazy *et al.* 2012) and prediction of ancestral nodes of the tree (SEQ ID
25 NOS:182-250) and the last common ancestor using joint maximum likelihood and
marginal maximum likelihood methods was performed. The last common ancestor made
using the joint reconstruction method was designated 'N1' and used for further
experiments.

The engineered ancestral amino acid sequence was reverse translated using the
30 Geneart codon optimisation algorithm (Thermo Fisher Scientific: Life Technologies
2014) with codon usage optimised for *Escherichia coli*. The propensity for the
corresponding mRNA to fold into stable secondary structure was analysed over the N-
terminal nucleotide sequence (from -21 to +96 with respect to start codon) using
NUPACK (Zadeh *et al.* 2011) and minimised by iterative silent changes to the

nucleotide sequence. The resultant sequence with two flanking sequences complementary to the destination expression vector added to the final open reading frame, encoding CPR_N1 set forth in SEQ ID NO:322, was synthesised as a GeneBlock by Integrated DNA Technologies (Singapore). As set forth in FIG. 9, the final
5 bicistronic expression vector was generated by Gibson Assembly (Zadeh *et al.* 2011) of this synthetic gene fragment using the backbone of the pCW'/4A11/hNPR vector from which the hNPR insert had been removed by digestion with XbaI and HindIII.

EXAMPLE 5. *Expression and thermostability measurement of the CYP3_N1, CYP2D_N1, and CPR_N1 engineered ancestral proteins*
10

Recombinant proteins were expressed as described previously (Notley *et al.* 2002) with minor modifications. Starter cultures (5 mL LB media supplemented with 11.1 mM glucose, 100 µg/mL ampicillin and 20 µg/mL chloramphenicol, in 25 mL flasks) were inoculated from glycerol stocks or freshly transformed colonies and incubated overnight
15 at 37°C, with shaking at 400 rpm. Expression cultures were set up in 100 mL flasks by inoculating 100 µL of starter culture into 10 mL Terrific Broth (TB) (Sambrook *et al.* 1989) containing 100 µg/mL ampicillin, 20 µg/mL chloramphenicol, 1mM thiamine and trace elements (Bauer *et al.* 1974; TB expression medium). Cultures were incubated at 25°C with shaking at 180 rpm for an initial 5h. Recombinant protein expression was
20 then induced by adding arabinose (4 mg/mL), δ-aminolaevulinic acid (0.5 mM) and isopropyl-β-thiogalactopyranoside (IPTG) (1mM) and the flasks were incubated at 25°C with shaking at 180 rpm for a further 43 h. As controls, the following recombinant extant proteins were expressed in parallel: for CYP3 experiments – CYP3A4 (Gillam *et al.* 1993), CYP3A5 (Gillam *et al.* 1995b), CYP3A27 (Behrendorff 2011) and CYP3A37 (Rawal *et al.* 2010); for CYP2D experiments – CYP2D22, CYP2D6, CYP2D12, CYP2D45, and CYP2D49; for CPR experiments – hNPR (extant human CPR).
25

For CYP3 and CYP2D experiments, cultures were harvested by centrifugation at 3000 x g for 10 min then resuspended in whole cell assay buffer (WCAB: 100 mM potassium phosphate, 6 mM magnesium acetate, 10 mM (+) glucose, pH 7.4). The
30 resuspended enzyme solution was heated at various temperatures (30-75°C) for 1h in Bio-Rad MyCycler Thermal Cycler, cooled at 4°C for 5 min and equilibrated at room temperature for 5 min. The residual folded protein was then quantified by Fe(II):CO versus Fe(II) difference spectroscopy (Johnston *et al.* 2008). In this way measures of thermal stability of CYP3 and CYP2D proteins based on the percentage of folded

protein after treatment for 60 minutes at various temperatures were obtained. Results for CYP3 and CYP2D proteins are set forth in FIG.10 and FIG. 11, respectively.

As will be evident from FIG. 10, the percentage of folded protein was substantially higher for CYP3_N1 (3N_1) after temperature treatment as compared to the percentage of folded protein for extant CYP3 proteins after temperature treatment. For example, after temperature treatment of 55°C, the percentage folded protein was ~ 100% for CYP3N_1, as compared to ~ 0% for the extant CYP3 proteins. Additionally, T50 values for folded protein were calculated, as presented in Table 1.

As will be evident from FIG. 11, the percentage of folded protein was substantially higher for CYP2D_N1 (2DN_1) after temperature treatment as compared to the percentage of folded protein for extant CYP3 proteins after temperature treatment. For example, after 60 minutes temperature treatment of 55°C, the percentage folded protein was ~ 100% for CYP2DN_1, as compared to ~ 0% for the extant CYP2D protein.

For CYP3 proteins, the half-life of folded protein was estimated by plotting the exponential decay graph of extant and ancestor proteins heated at various temperatures (50, 55, 60, 65, 70, and 75°C); results are set forth in Table 2. As will be evident from Table 2, the CYP3_N1 (3_N1) ancestor protein displayed substantially elevated half life of folded protein as compared to the extant CYP3 proteins. For example, after heating at 55°C, the half life of CYP3_N1 was ~ 566.5 minutes, while the half life for the extant CYP3 proteins was between ~ 0.5 minutes and ~ 1.6 minutes. Furthermore, CYP3 protein samples were taken at specific time-points and the residual folded protein was quantified. The percentage of folded enzyme remaining after heat treatment was compared to the non-heat-treated samples and the data obtained was used to generate an Arrhenius plot as set forth in FIG. 12. As will be evident from FIG. 12, the slope of the Arrhenius plot substantially was steeper for CYP3_N1 than for the extant CYP3 animal proteins. It will therefore be understood from FIG. 12 that the thermal deactivation energy was higher for CYP3_N1 than for the extant CYP3 animal proteins, as set forth in Table 3.

In order to evaluate whether there are stability effects being contributed by the interaction of CYP3_N1 with its redox partner, the thermostability of proteins produced by monocistronic (pCW/3_N1 His) and bicistronic (pCW/3_N1 his/hNPR) constructs were compared. As will be evident from FIG. 13, the thermostability of proteins produced by monocistronic and bicistronic constructs was similar.

For CPR_N1 experiments, cells were subjected to sub-cellular fractionation and membranes were prepared for analysis using a cytochrome c reductase assay; CPR_N1 activity was then assessed via the reduction of the surrogate electron acceptor cytochrome *c*, as previously described by Guengerich (1994) before and after heat treatment as described above for P450s. Results are set forth in FIG. 14; data are shown for three preparations of CPR_N1 derived from separate cultures compared to a pooled preparation of extant human CPR (hNPR).

As will be evident from FIG. 14, the percentage of active enzyme was substantially increased for CPR_N1 after temperature treatment as compared to the percentage of active enzyme for extant CPR protein after temperature treatment. For example, the percent active enzyme after 60 minutes heating at 55°C was ~ 50% for CPR_N1, as compared to ~ 10% for extant CPR. Additionally, T50 enzyme activity values were obtained for CPR proteins. $^{60}\text{T}_{50}$ values were as follows: CPR_N1, $45.0 \pm 1.0^\circ\text{C}$; hNPR, 38.5°C . Similar measurements on housefly reductase ($^{60}\text{T}_{50} = 41.2^\circ\text{C}$) and two *Arabidopsis thaliana* reductases (ATR1, $^{60}\text{T}_{50} = 39.1^\circ\text{C}$; ATR2, $^{60}\text{T}_{50} = 27.4^\circ\text{C}$) expressed in *E. coli* also showed these evolutionarily divergent CPRs to be less thermostable than CPR_N1.

EXAMPLE 6. *Characterization of enzyme activity for CYP3_N1 and CYP2D_N1*

Initial screens for activity towards fluorogenic or luminogenic or steroid marker substrates were done with intact cells washed and resuspended as above in WCAB. For detailed kinetic comparisons of CYP3_N1 and CYP2D_N1 and the respective extant P450 control forms (i.e. CYP3A4, CYP2D22, etc. as indicated for individual experiments), membrane fractions were isolated from bacteria expressing either CYP3_N1, CYP2D_N1 or an extant protein, with or without hCPR. Recombinant enzymes were expressed in 50 ml cultures, harvested and fractionated according to established procedures. An overview of the activity of CYP3_N1 on certain substrates is provided in FIG. 27.

30 *P450-Glo assays*

Activity assays were performed as described previously (Johnston *et al.* 2007) with luminogenic P450-Glo™ substrates at the following final concentrations: luciferin-CEE (15 μM), luciferin-ME (50 μM), luciferin-H (50 μM), luciferin-ME-EGE (15 μM), luciferin-H-EGE (5 μM), luciferin PFBE (25 μM), luciferin-BE (25 μM) and luciferin-

IPA (1.5 μ M). Briefly, cells harvested from P450 expression cultures were resuspended in an equal volume of WCAB then 2 μ L (~20 nM P450) were incubated with P450-GloTM substrate in a total volume of 40 μ L WCAB in white flat-bottomed 96 well plates. The reactions were allowed to proceed at 25°C for 1 h before quenching with
5 luciferin detection reagent (20 μ L). Samples were then incubated at room temperature for 20 min before quantification of luminescence using a Microbeta Trilux luminescence counter (Perkin-Elmer, USA). Results are presented in Table 4 and Table 5.

10 *Resorufin O-dealkylation assays*

Resorufin O-dealkylation assays were carried out as described in Chang and Waxman (2006). Results are presented in Table 4.

HPLC analysis of steroid hydroxylase activity

15 Analysis of steroid metabolism was performed as described previously (Hunter *et al.* 2011). Briefly, four 1 mL microaerobic cultures (Johnston *et al.* 2008) of each variant were combined, centrifuged (3000 x g, 10 min), and resuspended in 1 mL 100 mM potassium phosphate pH 7.4. Aliquots of each sample were taken to measure the P450 content, then the remaining cells were centrifuged and resuspended again in 1 mL 100
20 mM potassium phosphate pH 7.4 containing 10 mg/ml (+) glucose and substrate (testosterone or progesterone to final concentration 100 μ M). The cultures were incubated for 48 h at 25°C, with shaking at 180 rpm, and supplemented at 4 h, 20 h, 28 h and 44 h with glucose (2.75 μ moles per addition, in 50 μ L aliquots of a 10 mg/mL stock) and with ammonium hydroxide (5 μ moles per addition as 5 μ L aliquots of a 1 M
25 solution) in order to provide additional carbon and nitrogen sources. After 48 h, the cells were removed by centrifugation (3000 x g, 10 min) and 200 μ L of supernatant from each incubation were added to an equal volume of HPLC-grade acetonitrile. Aliquots (25 μ L) were then subjected to HPLC analysis on a C18 column (4.6 x 150 mm, 5 μ M, Agilent Technologies), initially equilibrated in 24% (v/v) acetonitrile:water
30 at a flow rate of 1.5 mL/min. The following gradient was used for elution of testosterone metabolites: 0 – 5.5 min, 24% acetonitrile; 5.5-12 min, linear gradient to 62% acetonitrile; 12-16 min, 62% acetonitrile; 16-18 min linear gradient to 24% acetonitrile; 18-22 min, 24% acetonitrile. Metabolites were detected by absorbance at 247 nm. Under these conditions, the major metabolite 6 β -hydroxytestosterone eluted at

approximately 7 min. Products formed were identified by comparison to the previously calibrated retention times of 6 β -hydroxytestosterone and other identified testosterone metabolites.

Kinetic analyses were carried out with isolated membrane fractions as described previously using 1 μ M P450 and 10 - 500 μ M testosterone for 10 min. Reactions (250 μ L) were quenched by the addition of 25 μ L progesterone 1 mM (as internal standard) and metabolites were extracted with 1 ml of ethyl acetate prior to analysis by HPLC as described above. Results are presented in Table 4, Table 8, and Figure 29.

10 *HPLC analysis of nifedipine oxidation*

Analysis of nifedipine metabolism was performed as described previously using bacterial membranes containing 0.1 μ M P450 in 100 mM Tris buffer pH 7.4 and 5 - 200 μ M nifedipine added from a methanolic stock such that the final methanol concentration was 1 % v/v. Reactions were quenched after 5 min with 50 μ L tetrahydrofuran containing 100 μ g/mL nordazepam as internal standard then metabolites were extracted by addition of 800 μ L of ethyl acetate and 200 μ L sodium carbonate pH 10. Extracts were desiccated under a gentle stream of nitrogen then resuspended in 60 μ L of mobile phase of which 25 μ L were analysed on a C18 column (4.6 x 150 mm, 5 μ M, Agilent Technologies), eluted isocratically with mobile phase (55% methanol, 45% water, 0.02% triethylamine, pH 5.0) at flow rate of 0.75 ml/min. Results are presented in Table 4.

Analysis of drug metabolism by LC-MS^E

Bacterial membranes containing 0.1 μ M CYP3_N1 (SEQ ID NO:2) coexpressed with human hNPR (human CPR); extant animal CYP3 protein CYP3A4 (Gillam et al. 1993); CYP2D_N1 (SEQ ID NO:181); or hNPR (human CPR) in the absence of any P450 as a control, were incubated with the substrates listed in Table 7 at two concentrations, 10 and 100 μ M, except for cyclosporin A, which was used at 2 and 20 μ M. Incubations were prepared in 250 μ L total volume of 0.1 M potassium phosphate buffer pH 7.4 and initiated by the addition of an NADPH-generating system consisting of 10 mM glucose-6-phosphate, 250 μ M NADP⁺, and 0.5 U/ml glucose-6-phosphate dehydrogenase. Reactions were quenched at 0, 20 and 120 mins by the addition of two volumes of ice-cold acetonitrile. Precipitated protein was removed by centrifugation and samples were

lyophilized then resuspended in 50 % v/v acetonitrile in water. Samples (10 µl) were injected into a Waters ACQUITY UPLC liquid chromatography system coupled to a Waters Synapt HDMS instrument (Waters, Milford, MA, USA) with an electrospray ionisation (ESI) source. Separation was achieved with a 130 Å, 1.7 µm x 2.1 mm x 100 mm Waters ACQUITY UPLC BEH C18 column (Waters, Milford, MA) at a column temperature of 45 °C.

Mobile phases consisting of ultra-pure water supplemented with formic acid (0.1% v/v; mobile phase A) and pure acetonitrile (mobile phase B) were employed at a flow rate of 0.5 mL min⁻¹. The gradient used was as follows: 0.0–6.0 min (10–70% mobile phase B); 6.0– 6.7 min (70–90% mobile phase B), then a return to the initial mobile phase composition over 0.01 min. The MSE analysis was performed with a Waters Synapt HDMS operating in V-mode positive electrospray ionization (ESI) conditions. A routine method with two scan functions was used, comprising an m/z range of 80–1000, cone voltage of 20 V and scan time of 0.1 s. The trap collision energy (CE) in function 1 was 20 V; in function 2 an energy ramp of 15–45 V was applied, and the transfer cell CE was 12 V. Data were collected in a centroid mode. Leu-enkephalin (250 pg mL⁻¹) was used as a lock mass (m/z 556.2771) for internal calibration at a flow rate of 0.04 mL min⁻¹. The MSE data were processed in MetaboLynx 4.1 (Waters, Milford, MA) using the dealkylation tool and mass defect filter and the proposed metabolites were manually reviewed.

Results of metabolism of tamoxifen at a concentration of 100 µM to its major demethylated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in Figure 31.

Results of metabolism of erythromycin at a concentration of 100 µM to its major demethylated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in Figure 32. Similarly, results of metabolism of erythromycin at a concentration of 10 µM to its major demethylated metabolite by CYP3_N1 (blue diamond), CYP3A4 (red square), and CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in Figure 33.

Results of metabolism of erythromycin at a concentration of 100 µM to its minor demethylated metabolite by CYP2D_N1 (green triangle) and control results for hNPR

(purple cross) are given in Figure 34. Note that no metabolism of erythromycin to its minor demethylated metabolite by CYP3_N1 or CYP3A4 was detected.

Results of metabolism of ticlopidine at a concentration of 100 μ M to its desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and
5 CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in Figure 35.

Results of metabolism of ticlopidine at a concentration of 10 μ M to its desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and
10 CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in Figure 36.

Results of metabolism of ticlopidine at a concentration of 100 μ M to its doubly desaturated metabolite by CYP3_N1 (blue diamond); CYP3A4 (red square); and
CYP2D_N1 (green triangle), and control results for hNPR (purple cross), are given in
15 Figure 37.

EXAMPLE 7. *Solvent stability of CYP3_N1*

Testosterone assays were carried out as described earlier. However solvents (methanol, acetonitrile or DMSO) were included in incubation mixtures to the concentrations indicated in the figures. Results are presented in FIG. 15 and FIG. 16. As will be
20 evident from FIG. 15 and FIG. 16, CYP3_N1 (TS) displayed substantially increased velocity in solvents comprising various concentrations of acetonitrile or methanol than extant animal CYP3 protein CYP3A4 (3A4). For example, velocity (pmol/min/pmol P450) in 10% methanol was ~ 2 for CYP3_N1 and ~ 1.5 for CYP3A4; and velocity (pmol/min/pmol P450) in 10% acetonitrile was ~ 0.8 for CYP3_N1 and ~ 0.25 for
25 CYP3A4 (FIG. 16).

EXAMPLE 8. *Ligand binding assays for CYP3_N1*

Spectral binding assays were performed as described elsewhere (Isin *et al.* 2008) with bacterial membranes containing CYP3_N1 or CYP3A4 to a final concentration of 0.1
30 μ M P450 in 1x TES using an OLIS-modified Aminco DW2A spectrophotometer. Binding data were analysed in Prism using the quadratic form of the binding equation (Isin *et al.* 2006). Results are presented in FIG. 18. In FIG. 18 CYP3_N1 is labelled as CYP3_TS.

EXAMPLE 9. *Creation of a CYP3_N1 variant library*

Overlap extension PCR was used for constructing a mutant library containing CYP3_N1 variant-encoding nucleic acids with codon substitutions at multiple sites (Williams *et al.* 2014). Each fragment A1, A2, B1, B2, C1, C2, D1, D2, E1, E2, as set
5 forth in Table 6, was amplified individually with the following: 10 ng template DNA (pMA_RQ/3_N1), polymerase buffer, 200 µM of each dNTP, 0.5 µM of forward and reverse primers (silent or mutagenic primers, see Table 1), 0.2 U Phusion® high fidelity DNA polymerase (New England Biolabs, USA) and sterile water to a total volume of 20 µl). Fragment F was amplified using pCW/2C19FL His EYFP/hNPR as template). PCR
10 cycling conditions were as follow: an initial “hot start” at 98°C for 1 min; 29 cycles of 98°C for 10s, T_{ann} for 20 s, and 72°C for 10 s; a polishing stage at 72°C for 10 min; and finally, storage at 4°C until use). PCR products were analysed by using agarose gel electrophoresis and purified using the Wizard® SV Gel and PCR clean-up system (Promega, Australia). Product yields were quantified using a Nanodrop
15 spectrophotometer.

Equal quantities of matched fragments from each PCR (A1 and A2, B1 and B2, C1 and C2, D1 and D2, E1 and E2) were combined in a primerless reassembly PCR containing polymerase buffer, 200 µM of each dNTP and 0.6 U Phusion® high fidelity DNA polymerase (New England Biolabs, USA) in a total volume of 30 µl. Cycling
20 conditions consisted of an initial hot start at 98°C for 2 min followed by 14 cycles of 98°C for 10s and 72°C for 30 s before storage at 4°C until use. These 30 µL PCR mixtures were then supplemented with additional polymerase buffer, dNTP and Phusion® high fidelity DNA polymerase (New England Biolabs, USA) plus each of the two flanking primers to a final concentration of 0.5 µM in a final volume of 50 µL. The
25 fragments were then amplified using the following conditions: an initial hot start at 98°C for 2 min; followed by 14 cycles of 98°C for 10s, and 72°C for 30 s; a polishing stage at 72°C for 10min; then storage at 4°C until use.

Fragments A, B, C, D, E and F were combined in stepwise manner. Equal quantities of fragment A were first combined with fragment B in a primerless
30 reassembly PCR followed by PCR with flanking primers as above. After gel purification, fragment AB was then combined with fragment C. The process was continued with the addition of fragments D, E and F in sequence, until the full-length open reading frame was obtained. The full-length mutant sequences were then

subcloned into the NdeI and XbaI sites of the pCW/2C19_{FL}His/hNPR plasmid as set forth in FIG. 17.

EXAMPLE 10. *Expression and thermostability screening of the CYP3_N1 first generation library*

Library variants were expressed as fusion proteins with EYFP from a bicistronic expression vector that allowed co-expression of hCPR. This vector was chosen to enable facile screening to eliminate mutants containing frame-shift mutations or premature stop codons in the subsequent library. The general protocol for P450 expression described above was followed but modified for high throughput format. Each CYP3 library variant (bicistronic, monocistronic and EYFP fusion in bicistronic format with hCPR) was grown in duplicate from two independent colonies. All media were as described above but starter cultures were set up in 96 well plates, inoculated from single colonies and incubated overnight at 37°C, with shaking at 400 rpm in a 5 mm orbit microplate shaker. Expression cultures were set up in 24-well plates by inoculating 20 µL starter cultures into 1 mL of TB expression medium. Cultures were incubated at 25°C with shaking at 350 rpm for an initial 5h, then recombinant protein expression was induced by adding arabinose (4 mg/mL), δ-aminolevulinic acid (0.5 mM) and isopropyl-β-thiogalactopyranoside (IPTG) (1mM) as above. The plates were incubated at 25°C with shaking at 350 rpm for a further 43 h. In some cases, plates were sealed with BreathEasy membranes (Diversified Biotech, Boston, USA) to restrict oxygen availability (microaerobic conditions) during P450 expression.

Cultures were harvested by centrifugation at 2000 g for 10 min then resuspended in WCAB. The resuspended enzyme solution was heated to either 72°C or 73°C for 10 min, cooled at 4°C for 5 min and equilibrated at RT for 5 min. The residual folded protein was then quantified by Fe(II).CO versus Fe(II) difference spectroscopy (Johnston *et al.* 2008). Variants which showed 50-60% residual folded protein after heating at 73°C or 60-70% after heating at 72°C were also analysed at 74°C. Finally, 39 mutants were selected for determination of ¹⁰T₅₀ values (temperature at which 50% of the protein remain folded after being heated for 10 min) as described above. Of these, 26 variants (comprising amino acid sequences set forth in SEQ ID NOS:3-28) showed the best thermostability in initial tests at 72-74°C and the remaining 13 variants (comprising amino acid sequences set forth in SEQ ID NOS:29-41) were chosen to reflect mutants with lower thermostability.

TABLES

Variant	⁶⁰ T ₅₀ (°C)
3A4	35 ± 3
3A5	38 ± 2
3A27	37 ± 3
3A37	36 ± 4
3_N1	66 ± 2

Table 1. ⁶⁰T₅₀ folded protein values of extant and ancestor CYP3s.

T (°C)	Variants	t _{1/2} (min)
50	3A4	2.1 ± 1.2
	3A5	1.1 ± 0.4
	3A27	4.5 ± 2.4
	3A37	2.4
	3_N1	595.4
55	3A4	0.5 ± 0.2
	3A5	0.9 ± 0.1
	3A27	1.6 ± 0.6
	3A37	0.8
	3_N1	566.5
60	3A4	0.4 ± 0.2
	3A5	0.5 ± 0.2
	3A27	0.7
	3A37	0.6
	3_N1	54.7
65	3A4	0.2 ± 0.1
	3A5	0.3
	3A27	0.8
	3A37	0.3
	3_N1	26.6
70	3A4	n.a
	3A5	0.2
	3A27	0.1
	3A37	0.2

	3_N1	11.7 ± 1.4
75	3A4	0.1
	3A5	0.1
	3A27	n.a
	3A37	0.1
	3_N1	1.3

Table 2. The estimated half-life of extant and ancestor CYP3s.

Variants	E _a (KJ/mol)
3A4	91
3A5	116
3A27	151
3A37	116
3_N1	251

Table 3. Thermal deactivation energy (E_a) of extant and ancestor CYP3s.

Substrates	CYP3_N1	CYP3A4	Reaction conditions
Luciferin derivatives			
Luc-PFBE	0.23 ± 0.06 pmol / 60 min	0.12 ± 0.02 pmol / 60 min	P450 = 22 nM, Luc-PFBE = 25 μM, Luc-IPA = 1.5 μM, Luc-
Luc-IPA	0.16 ± 0.03 pmol / 60 min	0.60 ± 0.10 pmol / 60 min	CEE = 30 μM, Luc-ME = 100 μM, Luc-H = 100 μM, Luc-ME-
Luc-CEE	n.a	n.a	EGE = 30 μM, Luc-H-EGE = 10 μM, 37°C, 60 min, whole cell
Luc-ME	n.a	n.a	
Luc-H	n.a	n.a	
Luc-ME-EGE	n.a	n.a	
Luc-H-EGE	n.a	n.a	
Resorufins			
7-benzoyloxyresorufin (7-BR)	0.154 ± 0.008*	0.192 ± 0.008*	P450 = 100 nM, hNPR = 0.2 μM, total protein = 0.8 mg/ml, resorufin = 5 μM, 25°C, 60 min, membrane
7-ethoxyresorufin (7-ER)	n.a	n.a	
7-methoxyresorufin (7-MR)	n.a	n.a	
7-pentoxoresorufin (7-PR)	0.025 ± 0.006*	0.036 ± 0.003*	
Coumarin	n.a	n.a	P450 = 150 nM, coumarin = 100 μM, 37°C, 120 min, whole cell
p-nitrophenol	n.a	n.a	P450 = 40 nM, pNP = 8.3 μM 37°C, 120 min, whole cell
Steroids			

Testosterone	16.3 ± 2.9 pmol/min/pmol	12.3 ± 0.6 pmol/min/pmol	Testosterone = 100 µM, 37°C, 10 min, membrane
	P450 (6β)	P450 (6β)	
	6β : 2β: 1β : 15β =	6β : 2β: 1β : 15β =	
	100 : 0 : 0 : 0	80 : 11: 4.5 : 4.5	
Nifedipine	1.79 pmol/min/pmol P450	3.74 pmol/min/pmol P450	Nifedipine = 100 µM, 37°C, 5 min, membrane
n.a = no activity , * = initial rate			

Table 4. Activities of CYP3_N1 and CYP3A4 towards marker substrates.

Substrates	CYP2D_N1	Reaction conditions
Luciferin derivatives		
Luc-ME-EGE	5.22 ± 0.08 pmol / 30 min	P450 = 40 nM, Luc-ME-EGE = 30 µM, 37°C, 60 min, CYP2D_N1/hNPR membranes, 50 µl reaction.

Table 5. Activities of CYP2D_N1 towards a luciferin derivative.

Entry	Fragment	Primer	Position	N1- marginal reconstruction	Sequence
1	A1	1F1	-	-	GCTTAGGAGGTCATATGGCTCTGTATTAGCAGTTTTTCTGACCCCTGCT
2		1R1	57	IF	GCATTGGGTATCAAAATCCAGGAWGCCTTTACGATATTCCAGAAAGGTAC
3	A2	1F2	57	IF	GTACCTTTCTGGAATATCGTAAAGGCWTCCTGGATTTTGATACCGAATGC
4		1R2	74	IL	CCATAATTGCCAGAACCGGCTGACGACCATCATAGAKGCCCCCAGATTTTGCCG
5	B1	1F3	-	-	CGTCAGCCGGTTCTGGCAATTATGGATCCGGCAATTATC
6		1R3	101	NT	CCAAAATTGCGACGATTGGTAAAGKTGCTGTAGCATTTCTTC
7	B2	1F4	101	NT	GAAAGAATGCTACAGCAMCTTTACCAATCGTCGCAATTTTGG
8		1R4	164, 172	VA, VM	ATAGGCGCCAAAAATATCTTTCAATCCACCGGCTCATCTTTTCCRCCTTTTCTGAAT
9	C1	1F5	-	-	TGAAAGATATTTTGGCGCCTATAGCATGGATGTTATTACAGCACCCAGCTTTAGCG
10		1R5	232	IV	ACAGTGAAATATTCAATTTTCCAGGAYCGGTGTGAGAAACGGAAAGATCAC
11	C2	1F6	232	IV	GTGATCTTTCCGTTTCTGACACCGRTCTCTGGAAAAAATGAATATTTCACTGT
12		1R6	250	MK	GCGATCTTCTTTTGATTTTTTTCACGGCGTTTCWTGAAAAAGTCCATCAC
13	D1	1F7	-	-	GTGAAAAAATCAAGAAGATCGCAAAAAAGACAAACCACAAAGATCGTGTGATTTCCCT
14		1R7	284	SN	CATCGGTCAGTGCTTTATGGCTATGGYTTTCTTCGCTGGTCTGGC
15	D2	1F8	284	SN	GCCAGACCCAGCGAAGAAARCCATAGCCATAAAGCACTGACCCGATG
16		1R8	451	ST	CTGCAGCAGACGCACAAATTGCAACTTTTCATCGWAGCAGTGCAAAAGCGC
17	E1	1F9	-	-	ATGAAAGTTGCAATTGTGCGTCTGCTGCAGAAATTTTAGCTTTCGTACCTGT

56

18	1R9-a	484	QT	CAATCGGTTTTTTTCGGCGTCAGAAAAACCCTGGGTATTTCAGTTCCAGCGGGATCTG
19	1R9-b			CAATCGGTTTTTTTCGGCTGCAGAAAAACCCTGGGTATTTCAGTTCCAGCGGGATCTG
20	E2	484	QT	CAGATCCCGCTGGAACTGAATACCCAGGGTTTTTCTGACGCCGAAAAAACCCGATTG
21	1F10-b			CAGATCCCGCTGGAACTGAATACCCAGGGTTTTTCTGACGCCGAAAAAACCCGATTG
22	1R10	-	-	TTGCTCACCATGCTCGAGCGCGCCGGTCGACCGAGGAACCA
23	F	-	-	GCCTCGAGCATGGTGAGCAAGGCGGAGGAGCTGTTTAC
24	1R12	-	-	CTTAAAGTTAAACAAAAATTATTTCTAGAGCTCAGCTGCATCAATGATGATGATGATG

Table 6. Mutagenic and silent oligonucleotides used in generating the CYP3 ancestors population.

Substrates	Substrates
Caffeine	Aminopyrine
Chlorzoxazone	Amlodipine
Dextromethorphan	Amodiaquine
Diclofenac	Atorvastatin
Erythromycin	Benzbromarone
Metoprolol	Carbamazepine
Phenytoin	Chlorzoxazone
Pravastatin	Clozapine
Propranolol	Cyclosporin A
Tamoxifen	Diltiazem
Ticlopidine	Donepezil
R, S, and R/S Warfarin	Fluoxetine

Table 7. List of representative drug substrates used to compare catalytic activities of CYP3_N1 and CYP2D_N1 with CYP3A4 and CYP2D6 respectively

Substrate	Enzyme	k_{cat} (min^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{min}^{-1} \mu\text{M}^{-1}$)
Testosterone	CYP3A4	23.9 ± 0.6	97 ± 6	0.25
	CYP3 _{TS}	18.1 ± 0.7	30 ± 4	0.61
Nifedipine	CYP3A4	5.4 ± 0.1	42 ± 3	0.13
	CYP3 _{TS}	2.0 ± 0.1	9 ± 1	0.22

Table 8. Comparison of kinetic parameters for 6 β -hydroxylation of testosterone and nifedipine oxidation by CYP3_{N1} and CYP3A4

X residue	Alignment Position	Position Relative to CYP3_N1	Conserved	Variable	Frequencies
	1		M		M:100%
	2			DNGSEMFLA	D:48% N:18% G:13% S:5% E:5% M:4% I:3% T:2% L:2% A:1%
	3			LFYDISN	L:61% F:15% Y:11% D:6% I:4% S:1% N:1%
	4			ILFVDS	I:55% L:24% F:13% V:4% D:3% S:1%
	5			PSLTFV	P:73% S:12% L:10% T:4% F:1% V:1%
	6			SNFDALYHG	S:32% N:24% F:13% D:11% A:6% L:5% Y:4% H:3% G:1%
	7			FL	F:64% L:36%
	8			SAT	S:79% A:20% T:1%
	9			MAIVLTGP	M:31% A:21% I:21% V:10% L:9% T:6% G:1% P:1%
	10			EGTVQDAI	E:79% G:7% T:6% V:3% Q:1% D:1% A:1% I:1%
	11	1	T		T:100%
	12	2	W		W:100%
	13	3		TVLIAM	T:39% V:32% L:19% I:7% A:2% M:1%
	14	4	L		L:99% I:1% P:1%
	15	5	L		L:88% I:7% V:5%
	16	6		AVGIL	A:59% V:36% G:3% I:1% L:1%

17	7		LTVIA	L:28% T:28% V:24% I:11% A:9%
18	8		SFLVIT	S:45% F:27% L:17% V:5% I:4% T:1%
19	9		LIFV	L:72% I:14% F:8% V:6%
20	10		VTAICS	V:51% T:35% A:7% I:4% C:1% S:1%
21	11	L		L:98% I:2%
22	12	L		L:88% I:7% F:4% M:1%
23	13		YLIVFTAM	Y:57% L:12% I:10% V:8% F:6% T:3% A:2% M:1%
24	14		LVIRYQGM	L:64% V:13% I:13% R:5% Y:2% Q:1% G:1% M:1%
25	15	Y		Y:98% F:2%
26	16	G		G:90% A:7% V:4%
27	17		TIYSVCN	T:58% I:22% Y:12% S:3% V:2% C:1% N:1%
28	18		WHYRA	W:42% H:21% Y:20% R:14% A:3%
29	19		PSTN	P:40% S:36% T:23% N:1%
30	20		HYF	H:61% Y:35% F:4%
31	21		GRKNSHQ	G:79% R:10% K:7% N:1% S:1% H:1% Q:1%
32	22		LIFVTH	L:65% I:12% F:11% V:6% T:6% H:1%
33	23	F		F:100%
34	24	K		K:96% T:2% E:1% R:1%
35	25	K		K:85% N:8% R:7%

	36	26	L		L:84% Q:7% M:5% Y:2% V:1% N:1%
	37	27	G		G:100%
	38	28	I		I:96% V:4%
	39	29	P		P:96% S:2% A:1% H:1%
	40	30	G		G:100%
	41	31	P		P:100%
X ₁	42	32		TKRMQA	T:40% K:29% R:26% M:2% Q:1% A:1%
	43	33	P		P:100%
	44	34	L		L:82% I:10% W:4% V:4%
	45	35	P		P:99% M:1%
	46	36	F		F:84% Y:9% L:7%
X ₂	47	37		FLIV	F:46% L:25% I:21% V:8%
	48	38	G		G:100%
X ₃	49	39		TNS	T:73% N:26% S:1%
X ₄	50	40		VFIMAC	V:46% F:18% I:18% M:13% A:4% C:1%
	51	41	L		L:96% F:3% G:1%
X ₅	52	42		SEGNARYFHQ	S:33% E:18% G:14% N:13% A:9% R:4% Y:3% F:3% H:2% Q:1%
	53	43	Y		Y:88% F:7% N:2% Q:1% S:1% H:1%
	54	44	R		R:81% Y:7% L:4% T:2% H:2% S:1% N:1% K:1% G:1%

X ₆	55	45		KQREDHNM	K:79% Q:7% R:4% E:3% D:3% H:2% N:1% M:1%
	56	46	G		G:99% V:1%
X ₇	57	47		FLVMIYN	F:49% L:26% V:10% M:8% I:4% Y:2% N:1%
X ₈	58	48		WLFHMCVYTDAI	W:43% L:12% F:9% H:8% M:7% C:7% V:5% Y:2% T:2% D:1% A:1% I:1%
X ₉	59	49		KNEDTAHQVSLPMG	K:21% N:17% E:17% D:16% T:10% A:4% H:4% Q:3% V:2% S:1% L:1% P:1% M:1% G:1%
	60	50	F		F:99% D:1%
	61	51	D		D:99% E:1%
X ₁₀	62	52		MTEKRQYSILVN	M:30% T:17% E:15% K:12% R:9% Q:6% Y:4% S:2% I:2% L:1% V:1% N:1%
X ₁₁	63	53		EKGQDQWM	E:72% K:16% G:4% D:4% Q:1% W:1% M:1%
	64	54	C		C:100%
X ₁₂	65	55		FYQHANSRK	F:46% Y:31% Q:6% H:6% A:4% N:3% R:2% S:1% K:1%
X ₁₃	66	56		KQAERS	K:77% Q:8% A:6% E:5% R:3% S:1%
	67	57	K		K:97% Q:2% T:1%
	68	58	Y		Y:100%
	69	59	G		G:96% R:4%
X ₁₄	70	60		KRDNE	K:79% R:7% D:6% N:4% E:3%
X ₁₅	71	61		MVIT	M:37% V:35% I:19% T:9%

	72	62	W		W:100%
	73	63	G		G:99% S:1% R:1%
X ₁₆	74	64		FILTMVS	F:38% I:30% L:23% T:4% M:3% V:1% S:1%
X ₁₇	75	65		YFQ	Y:76% F:23% Q:1%
	76	66	D		D:85% E:15%
	77	67	G		G:93% C:4% A:2% L:1%
X ₁₈	78	68		RQPKLT	R:76% Q:15% P:5% K:2% L:1% T:1%
X ₁₉	79	69		QLTKRM	Q:78% L:11% T:5% K:3% R:1% M:1%
	80	70	P		P:97% A:3%
X ₂₀	81	71		VLIMA	V:74% L:13% I:7% M:5% A:1%
	82	72	L		L:86% F:6% M:6% I:1% V:1%
X ₂₁	83	73		ACMVT	A:66% C:17% M:9% V:7% T:1%
	84	74	I		I:87% V:13%
X ₂₂	85	75		TMLASV	T:55% M:31% L:6% A:3% S:3% V:2%
	86	76	D		D:99% E:1%
	87	77	P		P:85% T:7% L:4% Q:3% V:1% K:1%
X ₂₃	88	78		DEASVNQLG	D:42% E:24% A:21% S:4% V:3% N:3% Q:1% L:1% G:1%
X ₂₄	89	79		MITL	M:65% I:27% T:7% L:1%
	90	80	I		I:99% L:1%
	91	81	K		K:99% R:1%

	92	82	T		T:85% N:6% A:4% I:2% S:2%
X ₂₅	93	83		VI	V:74% I:26%
	94	84	L		L:94% M:6%
	95	85	V		V:90% I:10%
	96	86	K		K:100%
	97	87	E		E:100%
	98	88	C		C:99% V:1%
	99	89	Y		Y:91% F:8% L:1%
	100	90	S		S:88% T:11% N:1%
X ₂₆	101	91		VTLNYASHIF	V:60% T:14% L:11% N:8% Y:1% A:1% S:1% H:1% I:1% F:1%
	102	92	F		F:100%
	103	93	T		T:100%
	104	94	N		N:100%
	105	95	R		R:99% H:1%
	106	96	R		R:97% M:3%
X ₂₇	107	97		NSPDVMTTER	N:29% S:24% P:20% D:17% V:4% M:3% T:1% E:1% R:1%
X ₂₈	108	98		FLVSTI	F:76% L:16% V:3% S:2% T:1% I:1%
X ₂₉	109	99		GRVDEIHH	G:76% R:13% V:5% D:3% E:1% I:1% H:1%
X ₃₀	110	100		PLFVIE	P:58% L:35% F:3% V:1% I:1% E:1%

X ₃₁	111	101		VNSAMDRPEL	V:45% N:26% S:10% A:8% M:4% D:1% R:1% P:1% E:1% L:1%
	112	102	G		G:100%
X ₃₂	113	103		FPIVDAQRL	F:38% P:32% I:17% V:5% D:3% A:1% Q:1% R:1% L:1% E:1%
X ₃₃	114	104		MLF	M:50% L:48% F:2%
X ₃₄	115	105		KEYSRAGDF	K:47% E:22% Y:13% S:5% R:5% A:3% G:2% D:1% F:1%
X ₃₅	116	106		SDKNE	S:41% D:26% K:18% N:13% E:1%
	117	107	A		A:86% G:8% S:6%
X ₃₆	118	108		IVL	I:66% V:29% L:5%
	119	109	S		S:82% T:13% L:3% A:1% F:1%
X ₃₇	120	110		IVLFMTSAR	I:57% V:17% L:17% F:3% M:2% T:1% S:1% A:1% R:1%
X ₃₈	121	111		ASVT	A:60% S:28% V:11% T:1%
X ₃₉	122	112		EKQV	E:79% K:18% Q:2% V:1%
	123	113	D		D:100%
	124	114	E		E:84% D:16%
X ₄₀	125	115		QERDKHT	Q:47% E:39% R:7% D:3% K:1% H:1% T:1%
	126	116	W		W:100%
	127	117	K		K:88% R:12%

	128	118	R		R:99% E:1%
	129	119	I		I:89% L:5% Y:5% M:1%
	130	120	R		R:100%
X ₄₁	131	121		TSAGN	T:49% S:41% A:7% G:1% N:1%
X ₄₂	132	122		LVTISQM	L:51% V:35% T:7% I:3% S:1% Q:1% M:1%
	133	123	L		L:93% I:7%
	134	124	S		S:99% C:1%
	135	125	P		P:100%
X ₄₃	136	126		TSCANY	T:65% S:18% C:7% A:4% N:3% Y:2%
	137	127	F		F:100%
	138	128	T		T:100%
	139	129	S		S:99% T:1%
	140	130	G		G:97% V:3%
X ₄₄	141	131		KRRH	K:69% R:30% H:1%
	142	132	L		L:96% F:3% I:1%
	143	133	K		K:100%
	144	134	E		E:92% Q:8%
	145	135	M		M:91% V:4% I:3% A:2%
	146	136	F		F:83% V:13% Y:3% L:1%
X ₄₅	147	137		PGDSQTE	P:79% G:6% D:5% S:4% Q:3% T:2% E:1%
	148	138	I		I:95% L:2% V:2% T:1%

X ₄₆	149	139		IMVA	I:58% M:33% V:7% A:2%
X ₄₇	150	140		KAGSNERMLI	K:35% A:21% G:14% S:7% N:7% E:7% R:3% M:3% L:3% I:1%
X ₄₈	151	141		QHRTKDSN	Q:56% H:24% R:7% T:5% K:4% D:3% S:1% N:1%
X ₄₉	152	142		YHCF	Y:75% H:21% C:3% F:1%
X ₅₀	153	143		GSATP	G:68% S:21% A:10% T:1% P:1%
X ₅₁	154	144		DEASQKHG	D:74% E:9% A:7% S:4% Q:2% K:1% H:1% G:1%
X ₅₂	155	145		VINRMLK	V:44% I:22% N:14% R:9% M:5% L:4% K:2%
	156	146	L		L:96% F:3% I:1%
X ₅₃	157	147		VITL	V:77% I:18% T:4% L:1%
X ₅₄	158	148		KRNSAQD	K:44% R:38% N:8% S:5% A:2% Q:1% D:1%
X ₅₅	159	149		NSKYFLG	N:70% S:15% K:6% Y:6% F:1% L:1% G:1%
X ₅₆	160	150		LMIV	L:63% M:21% I:15% V:1%
X ₅₇	161	151		RQKEHGW	R:52% Q:20% K:13% E:6% H:4% G:4% W:1%
X ₅₈	162	152		KQRGSV	K:54% Q:24% R:18% G:2% S:1% V:1%
X ₅₉	163	153		EKTYSMRQNG	E:56% K:29% T:4% Y:2% S:2% M:2% R:1% Q:1% N:1% G:1%
	164	154	A	VA	A:82% V:13% E:3% I:2% S:1%
X ₆₀	165	155		EDNTKAQ	E:72% D:17% N:6% T:2% K:1% A:1% Q:1%
X ₆₁	166	156		KNTLARPHE	K:65% N:10% T:9% L:6% A:3% R:3% P:2% H:1% E:1%

X ₆₂	167	157		GDSEKN	G:54% D:32% S:7% E:4% K:2% N:1%
X ₆₃	168	158		KEQSD	K:54% E:40% Q:4% S:1% D:1%
	169	159	P		P:82% S:7% V:4% N:3% T:1% L:1% A:1% Y:1%
X ₆₄	170	160		VILCT	V:68% I:17% L:10% C:3% T:2%
X ₆₅	171	161		TDNEAKQSP	T:47% D:24% N:13% E:7% A:3% K:2% Q:1% S:1% P:1%
X ₆₆	172	162		MLVIT	M:39% L:34% V:13% I:12% T:3%
	173	163	K		K:99% R:1%
X ₆₇	174	164		DEQHN	D:65% E:26% Q:5% H:3% N:1%
X ₆₈	175	165		IVFYL	I:52% V:23% F:22% Y:2% L:1%
	176	166	F		F:87% V:7% L:4% I:3%
	177	167	G		G:92% A:8%
X ₆₉	178	168		APS	A:73% P:24% S:4%
	179	169	Y		Y:100%
	180	170	S		S:97% T:3%
	181	171	M		M:89% L:10% I:1%
	182	172	D		D:100%
	183	173	V		V:96% I:2% T:1% A:1%
X ₇₀	184	174		IVM	I:69% V:29% M:1%
	185	175	T		T:94% L:2% A:2% I:1%
X ₇₁	186	176		SGA	S:79% G:19% A:2%

	187	177	T		T:90% A:5% S:3% C:1%
X ₇₂	188	178		SAL	S:74% A:24% L:3%
	189	179	F	F	F:100%
X ₇₃	190	180		GS	G:62% S:38%
	191	181	V	VI	V:95% I:5%
X ₇₄	192	182		NDEH	N:61% D:31% E:4% H:4%
X ₇₅	193	183		IVALTM	I:65% V:24% A:4% L:4% T:1% M:1%
	194	184	D		D:100%
	195	185	S		S:99% C:1%
X ₇₆	196	186		LIM	L:74% I:20% M:6%
	197	187	N		N:96% G:2% S:1% Q:1%
	198	188	N		N:94% K:3% H:3%
	199	189	P		P:97% Q:1% M:1%
X ₇₇	200	190		QSNKDEH	Q:43% S:18% N:17% K:10% D:8% E:2% H:1%
	201	191	D		D:97% S:2% N:1%
	202	192	P		P:99% T:1%
	203	193	F		F:95% V:3% L:2%
	204	194	V		V:90% I:4% L:3% N:3%
X ₇₈	205	195		ETKVRAIQS	E:44% T:24% K:8% V:5% R:5% A:4% I:4% Q:3%
					S:2%
X ₇₉	206	196		NKEHSQ	N:63% K:14% E:10% H:7% S:4% Q:1%

X ₈₀	207	197		ITAVSM	I:41% T:25% A:19% V:6% S:4% M:4%
	208	198	K		K:87% R:10% Q:2% S:1%
	209	199	K		K:95% N:5%
X ₈₁	210	200		LMFVIA	L:69% M:14% F:7% V:4% I:4% A:1%
	211	201	L		L:82% F:5% I:4% V:4% T:3% S:1%
X ₈₂	212	202		KRNTS	K:67% R:26% N:3% T:2% S:1%
	213	203	F		F:92% L:3% T:3% I:1% V:1%
X ₈₃	214	204		DNSGK	D:60% N:26% S:9% G:4% K:2%
	215	205	F		F:82% L:9% I:5% S:4% P:1%
X ₈₄	216	206		LFWVM	L:51% F:41% W:5% V:1% M:1%
X ₈₅	217	207		DNS	D:57% N:32% S:10%
	218	208	P		P:99% S:1%
X ₈₆	219	209		LFVP	L:51% F:39% V:7% P:3%
X ₈₇	220	210		FLIV	F:65% L:23% I:10% V:3%
X ₈₈	221	211		LIFA	L:76% I:16% F:6% A:1%
X ₈₉	222	212		SLIAF	S:54% L:33% I:9% A:2% F:1%
X ₉₀	223	213		IVLSQTK	I:56% V:27% L:7% S:4% Q:3% T:2% K:1%
X ₉₁	224	214		IVATSKMFLG	I:27% V:24% A:17% T:15% S:5% K:4% M:2% F:2% L:1% G:1%
X ₉₂	225	215		LIFVM	L:55% I:31% F:9% V:4% M:1%
	226	216	F		F:96% C:3% I:1%

	227	217	P		P:100%
	228	218	F		F:98% S:2%
X ₉₃	229	219		LIMITGSC	L:79% M:8% I:4% T:3% G:3% S:2% C:1%
X ₉₄	230	220		TAIGVRNQLSE	T:57% A:12% I:8% G:8% V:4% R:3% N:3% Q:2% L:1% S:1% E:1%
	231	221	P		P:89% R:6% N:2% Q:1% T:1% Y:1%
X ₉₅	232	222		VILA	V:40% I:39% L:20% A:1%
X ₉₆	233	223		LFYM	L:43% F:43% Y:12% M:3%
X ₉₇	234	224		EKDAG	E:79% K:14% D:6% A:1% G:1%
X ₉₈	235	225		KAVMLGFI	K:36% A:21% V:19% M:12% L:7% G:2% F:2% I:1%
X ₉₉	236	226		LMF	L:71% M:26% F:4%
X ₁₀₀	237	227		NDGEKS	N:68% D:13% G:7% E:7% K:4% S:1%
X ₁₀₁	238	228		IFVL	I:71% F:20% V:8% L:1%
X ₁₀₂	239	229		STCWFDANGE	S:54% T:16% C:8% W:4% F:4% D:3% A:3% N:3% G:3% E:3%
X ₁₀₃	240	230		LFMVIT	L:37% F:20% M:18% V:13% I:12% T:1%
	241	231	F		F:87% L:10% A:1% V:1% I:1%
	242	232	P		P:95% S:5%
X ₁₀₄	243	233		KRTLASNG	K:60% R:24% T:8% L:2% A:2% S:2% N:1% G:1%
X ₁₀₅	244	234		DSKEHQTYA	D:40% S:38% K:7% E:7% H:3% Q:2% T:1% Y:1%

							A:1%
X ₁₀₆	245	235			VSATIFN	V:52% S:23% A:14% T:4% I:3% F:3% N:1%	
X ₁₀₇	246	236			TMVIL	T:46% M:21% V:15% I:13% L:6%	
X ₁₀₈	247	237			DNSEQARKH	D:46% N:30% S:9% E:6% Q:3% A:2% R:2% K:1% H:1%	
	248	238		F		F:96% Y:4%	
X ₁₀₉	249	239			FL	F:75% L:25%	
X ₁₁₀	250	240			KYTRMANDSQF	K:34% Y:25% T:18% R:8% M:5% A:4% N:2% D:1% S:1% Q:1% F:1%	
X ₁₁₁	251	241			KNADRTQE	K:53% N:17% A:13% D:7% R:6% T:3% Q:1% E:1%	
X ₁₁₂	252	242			SAFVI	S:50% A:26% F:16% V:7% I:1%	
X ₁₁₃	253	243			VLIF	V:63% L:21% I:15% F:1%	
X ₁₁₄	254	244			KQEDRNATYHML	K:54% Q:11% E:7% D:4% R:4% N:4% A:3% T:3% I:3% Y:2% H:1% M:1% L:1%	
X ₁₁₅	255	245			RKSQ	R:50% K:41% S:6% Q:3%	
X ₁₁₆	256	246			IMFLT	I:49% M:41% F:7% L:1% T:1%	
	257	247		K		K:96% R:4%	
X ₁₁₇	258	248			ESKDA	E:60% S:16% K:14% D:9% A:1%	
X ₁₁₈	259	249			SDNEQKGTR	S:35% D:18% N:16% E:12% Q:6% K:5% G:5% T:2% R:1%	

	260	250	R		R:89% H:11%
X ₁₁₉	261	251		LKNEQVHMTS	L:50% K:18% N:7% E:7% D:5% Q:5% V:4% H:2% M:1% T:1% S:1%
X ₁₂₀	262	252		KDATNQEHG	K:61% D:10% A:10% T:7% N:5% Q:2% E:2% H:1% G:1%
X ₁₂₁	263	253		DSGEN	D:49% S:25% G:11% E:9% N:6%
X ₁₂₂	264	254		KNTDSEYVQL	K:40% N:29% T:20% D:5% S:1% E:1% Y:1% V:1% Q:1% L:1%
X ₁₂₃	265	255		QHEKS	Q:56% H:25% E:11% K:5% S:3%
	266	256	K		K:81% S:7% Q:3% N:3% E:2% L:1% T:1% M:1%
X ₁₂₄	267	257		HDNSGTQKPRI	H:49% D:10% N:10% S:9% G:7% T:5% Q:3% K:3% P:2% R:1% I:1%
	268	258	R		R:99% S:1%
	269	259	V		V:89% A:4% G:4% L:1% P:1% I:1%
	270	260	D		D:97% N:3%
	271	261	F		F:95% L:5%
	272	262	L		L:94% F:3% M:3%
	273	263	Q		Q:97% K:1% H:1% R:1%
	274	264	L		L:88% V:6% Q:3% M:2% H:1%
	275	265	M		M:99% L:1%
X ₁₂₅	276	266		IMVL	I:70% M:15% V:13% L:3%

X ₁₂₆	277	267		DNQE	D:55% N:35% Q:7% E:2%
	278	268	S		S:88% A:7% T:3% N:3%
	279	269	Q		Q:85% H:7% E:6% R:2% G:1%
X ₁₂₇	280	270		NTIKSMQD	N:52% T:20% I:10% K:9% S:4% M:3% Q:1% D:1%
X ₁₂₈	281	271		SLNIAVKEFP	S:68% L:7% N:6% I:6% A:3% V:3% K:2% E:2% F:1% P:1%
X ₁₂₉	282	272		KESDGTN	K:66% E:16% S:7% D:7% G:2% T:1% N:1%
X ₁₃₀	283	273		EDGQSKVNL	E:55% D:15% G:15% Q:6% S:4% K:2% V:1% N:1% L:1%
X ₁₃₁	284	274		TKESDNMQGVLA	T:32% K:15% E:13% S:10% D:10% N:6% M:6% Q:2% G:2% V:1% L:1% A:1%
X ₁₃₂	285	275		EDHQKSNVLA	E:34% D:18% H:18% Q:16% K:6% S:3% N:3% A:1% V:1% L:1%
X ₁₃₃	286	276		STQPNGR	S:57% T:17% Q:10% P:9% N:4% G:2% R:1%
X ₁₃₄	287	277		HDQYPSEVNKTL	H:60% D:9% Q:9% Y:6% P:4% S:3% E:3% V:2% N:1% K:1% T:1% L:1%
	288	278	K		K:99% T:1% G:1%
X ₁₃₅	289	279		AGDEVS	A:61% G:29% D:4% E:2% V:1% S:1%
	290	280	L		L:100%
X ₁₃₆	291	281		STN	S:70% T:28% N:2%

	292	282	D		D:93% E:6% N:1%
X ₁₃₇	293	283		LHVEMQATNKI	L:32% H:24% V:15% E:7% M:6% Q:5% A:4% T:3% N:2% K:2% I:1%
	294	284	E		E:100%
X ₁₃₈	295	285		ILMV	I:60% L:36% M:3% V:1%
X ₁₃₉	296	286		LVMAIT	L:47% V:35% M:7% A:4% I:4% T:3%
X ₁₄₀	297	287		AST	A:73% S:26% T:1%
	298	288	Q		Q:100%
X ₁₄₁	299	289		SAG	S:64% A:35% G:1%
X ₁₄₂	300	290		IMFVLT	I:63% M:17% F:10% V:5% L:4% T:1%
	301	291	I		I:93% M:2% A:1% V:1% L:1% S:1% G:1%
	302	292	F		F:97% I:3%
	303	293	I		I:96% V:3% L:1%
	304	294	F		F:93% L:4% M:2% V:1%
	305	295	A		A:89% G:11%
	306	296	G		G:97% A:3%
	307	297	Y		Y:100%
	308	298	E		E:94% D:6%
	309	299	T		T:94% A:3% P:2% S:1%
	310	300	T		T:86% S:14%
	311	301	S		S:98% A:1% T:1%

X ₁₄₃	312	302		STNI	S:69% T:29% N:1% I:1%
X ₁₄₄	313	303		TSVA	T:49% S:25% V:21% A:4%
	314	304	L		L:96% I:4% M:1%
	315	305	S		S:81% T:7% G:5% P:3% M:2% A:1%
	316	306	F		F:89% Y:7% L:2% N:2%
X ₁₄₅	317	307		LITVFM	L:57% I:29% T:6% V:5% F:2% M:1%
X ₁₄₆	318	308		MALFVS	M:28% A:26% L:18% F:12% I:10% V:4% S:2%
	319	309	Y		Y:99% H:1%
X ₁₄₇	320	310		NESLHQICG	N:37% E:35% S:9% L:7% H:5% Q:4% C:2% I:1% G:1%
	321	311	L		L:98% I:2%
	322	312	A		A:100%
	323	313	T		T:95% I:4% L:1% R:1%
X ₁₄₈	324	314		HN	H:68% N:32%
	325	315	P		P:99% H:1%
	326	316	D		D:81% E:11% H:5% K:2% A:1%
X ₁₄₉	327	317		VIAC	V:75% I:13% A:9% C:2% T:1%
X ₁₅₀	328	318		QML	Q:77% M:17% L:6%
X ₁₅₁	329	319		QKEHRT	Q:57% K:32% E:6% H:3% R:2% T:1%
	330	320	K		K:86% R:8% T:6%
	331	321	L		L:99% V:1%

	332	322	Q		Q:90% H:5% V:3% E:2%
X ₁₅₂	333	323		EKDQN	E:74% K:17% D:7% Q:2% N:1%
	334	324	E		E:100%
	335	325	I		I:99% V:1%
	336	326	D		D:100%
X ₁₅₃	337	327		ATRSEKV	A:61% T:8% R:8% S:7% E:7% K:6% V:3%
X ₁₅₄	338	328		TVANFLI	T:40% V:26% A:16% N:10% F:6% L:2% I:1%
X ₁₅₅	339	329		LFI	L:63% F:37% I:1%
	340	330	P		P:99% Q:1%
	341	331	N		N:87% K:5% D:3% L:2% G:2% S:1%
	342	332	K		K:90% D:9% Q:1%
	343	333	A		A:87% E:4% T:3% V:2% G:2% S:1%
	344	334	P		P:84% A:5% S:4% R:3% L:1% T:1% F:1%
X ₁₅₆	345	335		PVAIL	P:50% V:24% A:11% I:10% L:5%
X ₁₅₇	346	336		TSQEDNVH	T:76% S:7% Q:7% E:4% D:3% N:2% V:1% H:1%
	347	337	Y		Y:100%
X ₁₅₈	348	338		DEQAN	D:79% E:14% Q:6% A:1% N:1%
X ₁₅₉	349	339		ATDVIP	A:63% T:20% D:8% V:4% I:3% P:2%
X ₁₆₀	350	340		LVM I	L:74% V:22% M:3% I:1%
X ₁₆₁	351	341		MVLAN	M:49% V:32% L:12% A:4% N:3%
X ₁₆₂	352	342		QEGKDNR	Q:72% E:11% G:6% K:3% D:3% N:3% R:2%

	353	343	M		M:90% L:10%
	354	344	E		E:89% D:7% Q:4%
	355	345	Y		Y:100%
	356	346	L		L:98% M:2%
	357	347	D		D:98% E:2%
X ₁₆₃	358	348		MCQSAI	M:79% C:7% Q:6% S:4% A:3% I:1%
	359	349	V		V:94% A:5% S:1%
X ₁₆₄	360	350		VIL	V:43% I:33% L:24%
	361	351	N		N:94% Q:3% L:1% C:1%
	362	352	E	E	E:100%
X ₁₆₅	363	353		TSN	T:63% S:34% N:4%
	364	354	L		L:87% M:11% Q:2%
	365	355	R		R:100%
	366	356	L		L:94% M:4% I:1%
X ₁₆₆	367	357		YFIL	Y:60% F:33% I:4% L:3%
	368	358	P		P:100%
X ₁₆₇	369	359		IVPTL	I:63% V:14% P:13% T:7% L:2%
X ₁₆₈	370	360		AGVT	A:78% G:17% V:4% T:1%
X ₁₆₉	371	361		GPNMAISVT	G:35% P:19% N:13% M:10% A:7% I:7% S:4% V:2% T:1%
	372	362	R		R:100%

	373	363	L		L:85% I:13% V:3%
	374	364	E		E:93% D:4% T:3%
	375	365	R		R:100%
	376	366	V		V:88% M:5% L:2% T:2% A:1% I:1% E:1%
X ₁₇₀	377	367		CAST	C:74% A:18% S:7% T:1%
	378	368	K		K:100%
	379	369	K		K:83% A:7% T:4% Q:3% E:1% N:1% R:1%
X ₁₇₁	380	370		DTSNA	D:54% T:40% S:3% N:3% A:1%
	381	371	V		V:91% I:4% M:4% F:1%
	382	372	E		E:93% Q:5% D:2%
	383	373	I		I:88% L:10% V:1% F:1%
	384	374	N		N:82% H:13% S:5% D:1%
	385	375	G		G:100%
X ₁₇₂	386	376		VLIM	V:61% L:15% I:15% M:9%
X ₁₇₃	387	377		FTVSLY	F:44% T:34% V:9% S:7% L:4% Y:2%
	388	378	I	IVFM	I:96% V:2% F:1% M:1%
	389	379	P	P	P:100%
	390	380	K	KEA	K:92% E:6% A:2%
	391	381	G	GD	G:85% D:15%
X ₁₇₄	392	382		TMVSALC	T:60% M:14% V:12% S:7% A:4% L:3% C:1%
X ₁₇₅	393	383		VTMLAIGC	V:79% T:6% M:4% L:3% A:3% I:2% G:1% C:1%

	394	384	V		V:92% T:2% I:2% M:2% A:1%
X ₁₇₆	395	385		MVGALI	M:79% V:8% G:7% A:4% L:1% I:1%
X ₁₇₇	396	386		IVA	I:68% V:28% A:4%
	397	387	P		P:100%
X ₁₇₈	398	388		TVSPIAL	T:37% V:24% S:17% P:8% I:7% A:4% L:3%
X ₁₇₉	399	389		YFWHS	Y:60% F:22% W:8% H:6% S:4%
X ₁₈₀	400	390		AVPLT	A:48% V:32% P:10% L:6% T:5%
	401	391	L		L:100%
	402	392	H		H:93% Q:5% N:1% S:1%
X ₁₈₁	403	393		RHKLN	R:66% H:23% K:8% L:2% N:1%
	404	394	D		D:93% N:5% A:2%
	405	395	P		P:93% S:4% Q:1% L:1%
X ₁₈₂	406	396		EKQDRASN	E:42% K:27% Q:9% D:9% R:6% A:3% S:3% N:1%
X ₁₈₃	407	397		YLFHI	Y:52% L:29% F:10% H:5% I:4%
	408	398	W		W:100%
X ₁₈₄	409	399		PTSLEAD	P:68% T:18% S:10% L:1% E:1% A:1% D:1%
	410	400	E		E:88% S:9% D:2% K:1% N:1%
	411	401	P		P:100%
	412	402	E		E:93% Q:3% D:3% K:1%
X ₁₈₅	413	403		EKALS	E:76% K:13% A:4% L:3% S:3%
	414	404	F		F:100%

X ₁₈₆	415	405		RKLCQH	R:63% K:18% L:10% C:5% Q:4% H:1%
	416	406	P		P:100%
	417	407	E		E:100%
	418	408	R		R:100%
	419	409	F		F:100%
	420	410	S		S:97% T:3%
	421	411	K		K:96% P:2% R:1%
X ₁₈₇	422	412		EKGQ	E:55% K:40% G:3% Q:1%
	423	413	N		N:99% M:1%
	424	414	K		K:83% R:5% Q:4% E:4% T:2% N:1% S:1%
X ₁₈₈	425	415		DEGKSAQ	D:47% E:34% G:12% K:2% S:2% A:1% Q:1%
X ₁₈₉	426	416		SNTEK	S:59% N:21% T:10% E:9% K:1%
	427	417	I		I:82% V:5% L:3% F:3% Q:3% H:2% R:1%
X ₁₉₀	428	418		DNTL	D:62% N:33% T:3% L:2%
	429	419	P		P:97% Q:2% L:1%
	430	420	Y		Y:97% F:3%
X ₁₉₁	431	421		TIVAMRN	T:38% I:32% V:12% A:10% M:4% R:3% N:2%
	432	422	Y		Y:80% F:20%
X ₁₉₂	433	423		LMTI	L:54% M:29% T:15% I:3%
	434	424	P		P:100%
	435	425	F		F:100%

	436	426	G			G:100%
X ₁₉₃	437	427		ATNLS DVI	A:32% T:25% N:15% L:12% S:9% D:3% V:2% I:2%	
	438	428	G		G:100%	
	439	429	P		P:100%	
	440	430	R		R:100%	
	441	431	N		N:100%	
	442	432	C		C:100%	
	443	433	I		I:93% V:6% L:1%	
	444	434	G		G:100%	
	445	435	M		M:99% L:1% N:1%	
	446	436	R		R:100%	
	447	437	F		F:93% Y:5% M:1%	
	448	438	A		A:100%	
X ₁₉₄	449	439		LIQV	L:79% I:17% Q:3% V:1%	
X ₁₉₅	450	440		MLVIT	M:52% L:23% V:18% I:4% T:3%	
X ₁₉₆	451	441		NMSTVIL	N:52% M:13% S:12% T:11% V:7% I:4% L:2%	
	452	442	M		M:89% I:4% L:4% V:3%	
	453	443	K		K:100%	
X ₁₉₇	454	444		LVMF	L:73% V:21% M:6% F:1%	
	455	445	A		A:91% V:7% L:2%	

X ₁₉₈	456	446		LIVM	L:46% I:34% V:15% M:6%
X ₁₉₉	457	447		VITSA	V:57% I:25% T:10% S:4% A:3%
X ₂₀₀	458	448		REKTGVIS	R:57% E:18% K:10% T:4% G:4% V:4% I:2% S:1%
X ₂₀₁	459	449		VLIA	V:50% L:31% I:16% A:3%
	460	450	L		L:100%
	461	451	Q		Q:100%
X ₂₀₂	462	452		NRSEHQK	N:73% R:10% S:7% E:4% H:4% Q:2% K:1%
	463	453	F		F:85% Y:15%
	464	454	S		S:85% T:10% D:3% R:2%
	465	455	F		F:81% V:17% L:2%
X ₂₀₃	466	456		KRSQEVY	K:37% R:18% S:15% Q:13% E:12% V:3% Y:1%
X ₂₀₄	467	457		PTV	P:63% T:22% V:15%
	468	458	C		C:100%
	469	459	K		K:87% E:7% A:3% Q:2% D:1%
	470	460	E		E:90% D:10%
	471	461	T		T:100%
X ₂₀₅	472	462		QEPLMD	Q:71% E:13% P:7% L:4% M:4% D:1%
	473	463	I		I:92% V:8%
	474	464	P		P:100%
	475	465	L		L:88% M:7% F:5%
X ₂₀₆	476	466		KEVQR	K:62% E:31% V:3% Q:3% R:1%

X ₂₀₇	477	467		LMFI	L:73% M:11% F:9% I:7%
X ₂₀₈	478	468		SDNGRQCV	S:35% D:21% N:18% G:12% R:9% Q:2% C:2% V:1%
X ₂₀₉	479	469		STNRLWIFAKVP	S:18% T:18% N:18% R:10% L:10% W:6% I:5% F:4% A:3% K:3% V:2% P:2%
X ₂₁₀	480	470		QKGELIVSFN	Q:63% K:10% G:10% E:4% L:4% I:3% V:2% S:2% F:1% N:1%
	481	471	G		G:88% P:10% A:1%
X ₂₁₁	482	472		LFIVY	L:63% F:24% I:9% V:3% Y:2%
X ₂₁₂	483	473		LIFMVT	L:68% I:11% F:11% M:5% V:4% T:1%
X ₂₁₃	484	474		QTMARKGLE	Q:63% T:17% M:6% A:5% R:3% K:3% G:1% L:1% E:1%
	485	475	P		P:89% T:9% A:1% S:1%
X ₂₁₄	486	476		EKTLQAVSI	E:51% K:31% T:7% L:6% Q:2% A:1% V:1% S:1% I:1%
X ₂₁₅	487	477		KRNVDIT	K:76% R:8% N:6% V:5% D:3% I:1% T:1%
	488	478	P		P:100%
	489	479	I		I:98% V:1% L:1%
X ₂₁₆	490	480		VKITQRFL	V:57% K:23% I:13% T:2% Q:2% R:1% F:1% L:1%
	491	481	L		L:100%
	492	482	K		K:88% Q:4% S:2% R:2% T:1% N:1% M:1%

X ₂₁₇	493	483		VLFAIM	V:46% L:32% F:10% A:9% I:2% M:1%
X ₂₁₈	494	484		VEKILHDT	V:54% E:29% K:5% I:4% L:3% H:3% D:1% T:1%
X ₂₁₉	495	485		PSLT	P:54% S:35% L:10% T:1%
	496	486	R		R:98% K:2%

Table 9. Table of the amino acid sequence positions X₁ – X₂₁₉ of SEQ ID NO:1 with reference to CYP3_N1 (SEQ ID NO:2). Preferred amino acid sequence variants at X₁ – X₂₁₉ are presented. ‘Alignment Position’ represents the position within an alignment of extant and/or evolutionary intermediate sequences used to construct CYP3_N1. ‘Position Relative to CYP3_N1’ represents the respective position of CYP3_N1 (SEQ ID NO:2). ‘Conserved’ indicates amino acids at the respective positions which were conserved amongst the extant and/or evolutionary intermediate sequences used to construct CYP3_N1. ‘Variable’ indicates amino acids identified at the respective positions which varied amongst the extant and/or evolutionary intermediate sequences used to construct CYP3_N1. ‘Frequency’ indicates the percentage frequency of the individual amino acids detected amongst the extant and/or evolutionary intermediate sequences at the variable positions. The residues which form the N-terminal anchor region (alignment positions 1-38) are shaded grey. Note that, as set forth in SEQ ID NO:2, CYP3_N1 features a modified N-terminal region such that positions 1-9 (corresponding to alignment positions 11-19) are M, A, L, L, A, V, F, and L, respectively.

X residue	Alignment position	Position Relative to CYP2D_N1	Conserved	Variable	Frequencies
	1	1	M		M:100%
	2	2		GERSTA	G:68% E:18% R:6% S:4% T:3% A:1%
	3	3	L		L:99% M:1%
	4	4	L		L:96% P:4%
	5	5		TSLIN	T:76% S:12% L:4% I:4% N:3%
	6	6	G		G:97% W:3%
	7	7		DTSEAC	D:59% T:15% S:9% E:9% A:4% C:4%
	8	8		AGNEDTL	A:57% G:16% N:12% E:6% D:4% T:3% L:1%
	9	9	L		L:93% V:4% I:3%
	10	10		GWVFSa	G:34% W:28% V:16% F:16% S:4% A:1%
	11	11		PASTVI	P:76% A:7% S:7% T:4% V:3% I:1%
	12	12		LVIMG	L:65% V:26% I:4% M:3% G:1%
	13	13	A		A:84% G:12% V:4%
	14	14		IVL	I:53% V:46% L:1%
	15	15		AFIT	A:50% F:31% I:12% T:7%
	16	16		VTLCFM	V:57% T:28% L:9% C:3% F:1% M:1%
	17	17		AVTLI	A:57% V:21% T:18% L:3% I:1%
	18	18		ILV	I:66% L:26% V:7%
	19	19	F		F:96% C:3% L:1%

20	20	L				L:81% I:15% T:3% A:1%
21	21	L				L:97% I:3%
22	22	L				L:100%
23	23	V				V:93% L:7%
24	24	D				D:96% N:4%
25	25	L				L:88% F:9% Y:3%
26	26			MLVT		M:76% L:21% V:1% T:1%
27	27	H				H:88% K:12%
28	28	R				R:96% W:1% K:1% Q:1%
29	29	R				R:96% H:4%
30	30			QRKSP		Q:56% R:24% K:13% S:6% P:1%
31	31			RFPKH		R:66% F:25% P:4% K:3% H:1%
32	32	W				W:100%
33	33			ATCR		A:60% T:26% C:12% R:1%
34	34			ASPT		A:66% S:24% P:6% T:4%
35	35	R				R:94% N:3% H:1% D:1%
36	36			YF		Y:74% F:26%
37	37	P				P:100%
38	38	P				P:100%
39	39	G				G:96% S:4%
40	40	P				P:100%

X ₁	41	41			MVLPT	M:59% V:21% L:13% P:4% T:3%
	42		P			P:87% S:12% G:1%
X ₂	43	43			LWVRI	L:72% W:21% V:3% R:3% I:1%
	44		P			P:100%
X ₃	45	45			GLVF	G:41% L:24% V:24% F:12%
	46	46	L			L:88% V:12%
	47		G			G:100%
	48	48	N			N:96% T:4%
	49	49	L			L:96% M:4%
	50	50	L			L:97% P:3%
	51	51	Q			Q:81% H:16% Y:3%
X ₄	52	52			VMI	V:71% M:28% I:1%
	53	53	D			D:97% N:3%
	54	54	F			F:81% L:13% Y:4% P:1%
X ₅	55	55			QEDRKHSNG	Q:57% E:13% D:7% R:6% K:4% H:4% S:3% N:3% G:1%
X ₆	56	56			NDG	N:63% D:35% G:1%
X ₇	57	57			TMPLI	T:44% M:37% P:10% L:7% I:1%
X ₈	58	58			PLHRA	P:57% L:24% H:12% R:6% A:1%
X ₉	59	59			YFLANSPC	Y:50% F:21% L:6% A:6% N:6% S:4% P:4% C:3%
X ₁₀	60	60			SYCG	S:56% Y:24% C:15% G:6%

X ₁₁	61	61			FLCM	F:78% L:19% C:1% M:1%
X ₁₂	62	62			DYNQKIRE	D:38% Y:28% N:19% Q:6% K:4% I:1% R:1% E:1%
X ₁₃	63	63			QKR	Q:41% K:38% R:21%
	64	64		L		L:96% F:4%
X ₁₄	65	65			RQS	R:71% Q:21% S:9%
X ₁₅	66	66			RNKCHGQ	R:53% N:16% K:12% C:10% H:6% G:1% Q:1%
	67	67		R		R:88% Q:7% K:4%
X ₁₆	68	68			FY	F:71% Y:29%
	69	69		G		G:100%
	70	70		D		D:93% N:7%
	71	71		V		V:91% L:6% I:3%
	72	72		F		F:96% M:4%
	73	73		S		S:100%
	74	74		L		L:100%
	75	75		Q		Q:100%
X ₁₇	76	76			LMQNVF	L:63% M:21% Q:6% N:4% V:3% F:3%
X ₁₈	77	77			AGCVFI	A:69% G:12% C:7% V:6% F:4% I:1%
	78	78		W		W:94% S:4% F:1%
X ₁₉	79	79			TKERN	T:56% K:35% E:6% R:1% N:1%
	80	80		P		P:91% S:4% N:4%
X ₂₀	81	81			VM	V:79% M:21%

	82	82	V			V:100%	
	83	83	V			V:97% I:3%	
X ₂₁	84	84			LIV	L:78% I:18% V:4%	
	85	85	N			N:100%	
	86	86	G			G:97% R:3%	
	87	87	L			L:87% F:9% Y:3% P:1%	
X ₂₂	88	88			AEKTVLQP	A:34% E:31% K:24% T:4% V:3% L:1% Q:1% P:1%	
	89	89	A			A:93% V:4% T:3%	
X ₂₃	90	90			VMLI	V:79% M:10% L:6% I:4%	
	91	91	R			R:81% K:16% Q:3%	
	92	92	E			E:100%	
X ₂₄	93	93			AVLM	A:79% V:18% L:1% M:1%	
	94	94	L			L:100%	
	95	95	V			V:88% L:7% I:4%	
X ₂₅	96	96			TYQNKSH	T:44% Y:29% Q:9% N:9% K:6% S:1% H:1%	
X ₂₆	97	97			HYCKRN	H:54% Y:16% C:13% K:12% R:3% N:1%	
X ₂₇	98	98			SG	S:62% G:38%	
	99	99	E			E:93% Q:4% T:1% H:1%	
	100	100	D			D:99% E:1%	
X ₂₈	101	101	T		TF	T:96% F:4%	
	102	102	A			A:96% S:4%	

	103	103	D			D:100%	
	104	104	R			R:100%	
	105	105	P			P:100%	
	106	106	P			P:91% Q:3% Y:3% L:1% T:1%	
X ₂₉	107	107			VMLFPA	V:51% M:26% L:13% F:4% P:3% A:1%	
	108	108	P			P:84% H:10% A:3% F:1% N:1%	
	109	109	I			I:87% T:4% F:4% V:3% L:1%	
X ₃₀	110	110			YFNTL	Y:62% F:16% N:12% T:6% L:4%	
X ₃₁	111	111			EQDNKA	E:72% Q:13% D:6% N:4% K:3% A:1%	
X ₃₂	112	112			HIVYLCQ	H:62% I:10% V:9% Y:9% L:6% C:3% Q:1%	
	113	113	L			L:99% S:1%	
	114	114	G			G:100%	
X ₃₃	115	115			FYVI	F:65% Y:18% V:15% I:3%	
	116	116	G			G:85% K:12% V:3%	
X ₃₄	117	117	P		RP	P:84% H:7% G:4% R:4%	
X ₃₅	118	118			RKGNH	R:66% K:19% G:7% N:4% H:3%	
	119	119	S			S:85% A:10% N:4%	
	120	120	Q			Q:81% E:13% K:6%	
	121	121	G			G:97% A:3%	
	122	122	V			V:96% I:4%	
X ₃₆	123	123			VFI	V:40% F:37% I:24%	

X ₃₇	124	124				LMGF	L:69% M:25% G:3% F:3%
	125	125		A			A:100%
X ₃₈	126	126				RPNSY	R:68% P:24% N:4% S:3% Y:1%
	127	127		Y			Y:100%
	128	128		G			G:100%
X ₃₉	129	129				PRHQKD	P:53% R:26% H:7% Q:7% K:4% D:1%
X ₄₀	130	130				AESVQ	A:49% E:43% S:4% V:3% Q:1%
	131	131		W			W:100%
	132	132		R			R:88% K:12%
	133	133		E			E:90% Q:6% D:4%
	134	134		Q			Q:88% L:12%
	135	135		R			R:100%
	136	136		R			R:97% K:3%
	137	137		F			F:100%
X ₄₁	138	138				SAT	S:68% A:21% T:12%
X ₄₂	139	139				VL	V:62% L:38%
	140	140		S			S:97% T:3%
	141	141		T			T:99% S:1%
	142	142		L			L:93% F:4% M:3%
	143	143		R			R:99% H:1%
	144	144		N			N:81% D:10% H:6% T:3%

	145	145	F			F:85% L:15%	
	146	146	G			G:100%	
	147	147	L			L:81% M:15% V:4%	
	148	148	G			G:100%	
	149	149	K			K:100%	
	150	150	K			K:100%	
	151	151	S			S:100%	
	152	152	L			L:100%	
	153	153	E			E:100%	
X ₄₃	154	154			QED	Q:75% E:18% D:7%	
	155	155	W			W:85% R:15%	
	156	156	V			V:100%	
X ₄₄	157	157	T		RT	T:94% R:6%	
	158	158	E			E:85% K:12% Q:1% D:1%	
	159	159	E			E:100%	
	160	160	A			A:100%	
X ₄₅	161	161			SGARTN	S:44% G:28% A:16% R:9% T:1% N:1%	
X ₄₆	162	162			CHFV	C:68% H:21% F:7% Y:4%	
	163	163	L			L:100%	
	164	164	C			C:97% S:3%	
X ₄₇	165	165			ADS	A:74% D:22% S:4%	

	166	166	A			A:100%	
	167	167	F			F:96% I:4%	
X ₄₈	168	168			ATSQ	A:63% T:25% S:7% Q:4%	
X ₄₉	169	169			DASNKQ	D:56% A:16% S:12% N:12% K:3% Q:1%	
X ₅₀	170	170			HQE	H:57% Q:26% E:16%	
X ₅₁	171	171			ASEQTGN	A:75% S:10% E:6% Q:4% T:1% G:1% N:1%	
	172	172	G			G:99% E:1%	
X ₅₂	173	173			RHQFC	R:68% H:18% Q:12% F:1% C:1%	
	174	174	P			P:96% S:4%	
	175	175	F			F:88% I:12%	
X ₅₃	176	176			SNRDC	S:51% N:21% R:15% D:12% C:1%	
	177	177	P			P:100%	
X ₅₄	178	178			NRHKSTM	N:51% R:24% H:12% K:6% S:4% T:1% M:1%	
X ₅₅	179	179			ATGFDVSPN	A:43% T:24% G:12% F:7% D:4% V:4% S:3% P:1% N:1%	
	180	180	L			L:88% M:12%	
	181	181	L			L:94% V:3% I:3%	
X ₅₆	182	182			NDS	N:79% D:19% S:1%	
X ₅₇	183	183	K		NK	K:85% N:10% T:4%	
	184	184	A			A:97% S:3%	
	185	185	V			V:94% L:3% T:3%	

X ₅₈	186	186			SC	S:62% C:38%
	187		N			N:100%
	188		V			V:100%
	189		I			I:100%
X ₅₉	190				ACTS	A:78% C:12% T:7% S:3%
	191		S			S:99% T:1%
	192		L			L:88% I:12%
X ₆₀	193				TIL	T:53% I:41% L:6%
X ₆₁	194				YFC	Y:56% F:32% C:12%
X ₆₂	195				GA	G:63% A:37%
X ₆₃	196				RECH	R:76% E:12% C:10% H:1%
	197		R			R:100%
	198		F			F:100%
	199		E			E:93% D:7%
	200		Y			Y:100%
X ₆₄	201				DENG	D:69% E:21% N:7% G:3%
	202		D			D:100%
	203		P			P:87% K:7% H:6%
X ₆₅	204				RYKFM T	R:63% Y:15% K:9% F:6% M:4% T:3%
X ₆₆	205				FLIVM	F:54% L:38% I:4% V:1% M:1%
X ₆₇	206				ILHQNK T	I:32% L:29% H:21% Q:9% N:4% K:3% T:1%

X ₆₈	207	207		RK	R:56% K:44%
	208	208	L		L:81% M:19%
X ₆₉	209	209		LIFM	L:75% I:21% F:3% M:1%
X ₇₀	210	210		DKTRCEG	D:63% K:19% T:4% R:4% C:4% E:3% G:1%
X ₇₁	211	211		LGVRTFM	L:65% G:15% V:6% R:6% T:4% F:3% M:1%
X ₇₂	212	212		LATVIF	L:37% A:28% T:22% V:4% I:4% F:4%
X ₇₃	213	213		ELQKH	E:46% L:22% Q:16% K:13% H:3%
X ₇₄	214	214		EDKQNT	E:47% D:24% K:18% Q:7% N:3% T:1%
X ₇₅	215	215		GSTAI	G:47% S:32% T:15% A:3% I:3%
	216	216	L		L:93% F:4% V:3%
X ₇₆	217	217		KQTGENRI	K:46% Q:21% T:12% G:7% E:6% N:4% R:3% I:1%
	218	218	E		E:96% A:4%
X ₇₇	219	219		EDVISN	E:51% D:32% V:10% I:3% S:1% N:1%
	220	220	S		S:90% A:4% T:3% V:1% F:1%
	221	221	G		G:97% S:1% N:1%
	222	222	F		F:81% L:13% A:3% I:1% P:1%
X ₇₈	223	223		LIVAKM	L:49% I:21% V:15% A:13% K:1% M:1%
X ₇₉	224	224		PRALI	P:47% R:46% A:4% L:1% I:1%
X ₈₀	225	225		EQMKG	E:57% Q:34% M:6% K:1% G:1%
X ₈₁	226	226		VAIFL	V:60% A:22% I:7% F:7% L:3%
	227	227	L		L:94% V:3% I:3%

	228	228	N			N:90% E:6% S:4%
X ₈₂	229	229			ASVT	A:43% S:29% V:15% T:13%
X ₈₃	230	230			IFVLM	I:54% F:21% V:19% L:4% M:1%
	231	231	P			P:100%
X ₈₄	232	232			VIWLMA	V:65% I:13% W:9% L:6% M:4% A:3%
	233	233	L			L:96% S:4%
	234	234	L			L:94% S:4% R:1%
X ₈₅	235	235			RHKSQN	R:76% H:15% K:4% S:1% Q:1% N:1%
	236	236	I			I:99% V:1%
	237	237	P			P:100%
	238	238	G			G:87% A:7% R:3% W:1% C:1%
	239	239	L			L:96% V:4%
	240	240	A			A:82% P:9% V:6% L:1% C:1%
X ₈₆	241	241			GDAQK	G:41% D:37% A:10% Q:7% K:4%
	242	242	K			K:93% L:3% R:3% M:1%
	243	243	V			V:90% F:4% A:3% L:1% I:1%
X ₈₇	244	244			FL	F:78% L:22%
X ₈₈	245	245			PSQR	P:31% S:29% Q:21% R:19%
X ₈₉	246	246			GSAFPKNCR	G:43% S:25% A:10% F:7% P:4% K:4% N:3% C:1% R:1%
	247	247	Q			Q:85% K:6% L:4% R:4%

	248	248	K			K:84% R:7% I:4% N:3% T:1%
X ₉₀	249	249		ASTH		A:76% S:13% T:7% H:3%
X ₉₁	250	250		FLMY		F:68% L:25% M:4% Y:3%
X ₉₂	251	251		MLIV		M:57% L:29% I:9% V:4%
X ₉₃	252	252		ATKDV		A:51% T:38% K:4% D:4% V:1%
X ₉₄	253	253		LQIFMYS		L:51% Q:15% I:12% F:9% M:7% Y:4% S:1%
	254	254	L			L:87% I:6% T:4% V:1% N:1%
X ₉₅	255	255		DNQSEY		D:71% N:19% Q:4% S:3% E:1% Y:1%
X ₉₆	256	256		ENKVM		E:75% N:15% K:7% V:1% M:1%
	257	257	L			L:91% I:4% M:4%
X ₉₇	258	258		LVI		L:53% V:29% I:18%
X ₉₈	259	259		TQAKINE		T:53% Q:24% A:7% K:4% I:4% N:4% E:3%
	260	260	E			E:97% K:3%
	261	261	H			H:84% N:13% Q:3%
X ₉₉	262	262		RKQN		R:65% K:31% Q:3% N:1%
X ₁₀₀	263	263		MTIEQV		M:40% T:26% I:21% E:7% Q:4% V:1%
	264	264	T			T:99% G:1%
X ₁₀₁	265	265		WR		W:65% R:35%
	266	266	D			D:94% N:4% A:1%
	267	267	P			P:94% S:4% L:1%
X ₁₀₂	268	268		ATDGV		A:57% T:31% D:6% G:4% V:1%

	269	269	Q			Q:88% H:10% Y:1%
	270	270	P			P:99% A:1%
	271	271	P			P:88% T:12%
	272	272	R			R:93% Q:7%
	273	273	D			D:84% N:12% H:3% G:1%
	274	274	L			L:88% F:12%
	275	275	T			T:94% I:6%
	276	276	D			D:84% E:16%
	277	277	A			A:100%
	278	278	F			F:100%
	279	279	L			L:96% M:4%
X ₁₀₃	280	280			ADLKTH	A:51% D:35% L:7% K:3% T:1% H:1%
	281	281	E			E:99% Q:1%
X ₁₀₄	282	282			MIV	M:38% I:35% V:26%
	283	283	E			E:88% K:6% A:3% H:3%
	284	284	K			K:94% E:6%
	285	285	A			A:97% G:3%
X ₁₀₅	286	286			KRQ	K:78% R:21% Q:1%
	287	287	G			G:96% E:4%
	288	288	N			N:90% V:4% A:3% S:1% R:1%
	289	289	P			P:88% K:4% S:3% E:3% V:1%

X ₁₀₆	290	290	E	ED	E:96% D:4%
	291	291	S		S:97% N:3%
	292	292	S		S:93% N:4% G:3%
	293	293	F		F:100%
	294	294	N		N:100%
	295	295	D		D:85% E:12% K:3%
	296	296	E		E:87% K:3% Q:3% D:3% N:3% S:1%
	297	297	N		N:100%
X ₁₀₇	298	298		LM	L:79% M:21%
X ₁₀₈	299	299		RLSQ	R:69% L:25% S:4% Q:1%
X ₁₀₉	300	300		MLIVT	M:49% L:35% I:13% V:1% T:1%
	301	301	V		V:93% T:6% L:1%
X ₁₁₀	302	302		VT	V:65% T:35%
X ₁₁₁	303	303		ASVIGR	A:78% S:15% V:3% I:1% G:1% R:1%
	304	304	D		D:100%
	305	305	L		L:100%
	306	306	F		F:99% L:1%
X ₁₁₂	307	307		STLMAIG	S:37% T:22% L:21% M:7% A:6% I:4% G:3%
	308	308	A		A:100%
	309	309	G		G:100%
	310	310	M		M:82% T:6% I:6% S:6%

X ₁₁₃	311	311			VLEI	V:72% L:15% E:12% I:1%
X ₁₁₄	312	312			TS	T:76% S:24%
	313	313		T		T:100%
	314	314		S		S:96% A:3% T:1%
X ₁₁₅	315	315			TAIV	T:75% A:21% I:3% V:1%
	316	316		T		T:100%
	317	317		L		L:100%
X ₁₁₆	318	318			ASRT	A:54% S:28% R:12% T:6%
	319	319		W		W:100%
X ₁₁₇	320	320			AG	A:79% G:21%
	321	321		L		L:97% F:3%
	322	322		L		L:97% M:3%
	323	323		L		L:87% F:13%
	324	324		M		M:100%
	325	325		I		I:88% L:12%
	326	326		L		L:100%
	327	327		H		H:85% Y:15%
	328	328		P		P:100%
	329	329		D		D:97% E:3%
	330	330		V		V:97% I:3%
	331	331		Q		Q:100%

	332	332	R			R:93% S:6% C:1%
	333	333	R			R:87% K:12% H:1%
	334	334	V			V:100%
	335	335	Q			Q:88% H:12%
X ₁₁₈	336	336			QREK	Q:66% R:21% E:10% K:3%
	337	337	E			E:100%
	338	338	I			I:100%
	339	339	D			D:100%
X ₁₁₉	340	340			EDKQNA	E:69% D:12% K:7% Q:4% N:4% A:3%
	341	341	V			V:100%
X ₁₂₀	342	342			IL	I:79% L:21%
	343	343	G			G:100%
X ₁₂₁	344	344			QPRH	Q:65% P:22% R:12% H:1%
X ₁₂₂	345	345			VEDTGAN	V:75% E:6% D:6% T:4% G:4% A:3% N:1%
X ₁₂₃	346	346			RQ	R:76% Q:24%
X ₁₂₄	347	347			RHCKSQP	R:69% H:12% C:9% K:4% S:3% Q:1% P:1%
	348	348	P			P:100%
X ₁₂₅	349	349			EAT	E:68% A:21% T:12%
	350	350	M			M:99% I:1%
X ₁₂₆	351	351			AGTK	A:50% G:46% T:3% K:1%
	352	352	D			D:100%

	353	353	Q			Q:96% V:4%
X ₁₂₇	354	354		ATVL		A:66% T:22% V:7% L:4%
X ₁₂₈	355	355		RHLQNK		R:66% H:16% L:6% Q:4% N:4% K:3%
	356	356	M			M:100%
	357	357	P			P:100%
X ₁₂₉	358	358		YF		Y:59% F:41%
	359	359	T			T:100%
X ₁₃₀	360	360		NMTVL		N:49% M:29% T:18% V:3% L:1%
	361	361	A			A:100%
	362	362	V			V:100%
X ₁₃₁	363	363		IMVL		I:65% M:21% V:13% L:1%
	364	364	H			H:100%
	365	365	E			E:100%
X ₁₃₂	366	366	V	VI		V:96% I:4%
	367	367	Q			Q:100%
	368	368	R			R:100%
	369	369	F			F:88% Y:9% C:3%
X ₁₃₃	370	370		GA		G:75% A:25%
	371	371	D			D:100%
X ₁₃₄	372	372		IL		I:79% L:21%
X ₁₃₅	373	373		VALI		V:62% A:28% L:6% I:4%

	374	374	P			P:100%	
	375	375	L			L:87% V:7% M:4% T:1%	
X ₁₃₆	376	376			GNS	G:75% N:21% S:4%	
X ₁₃₇	377	377			VLIM	V:69% L:21% I:6% M:4%	
	378	378	P			P:84% T:15% A:1%	
	379	379	H			H:87% R:13%	
X ₁₃₈	380	380			MIKQ	M:79% I:13% K:6% Q:1%	
	381	381	T			T:100%	
	382	382	S			S:88% Y:12%	
	383	383	R			R:96% C:3% H:1%	
	384	384	D			D:100%	
	385	385	I			I:84% T:15% V:1%	
	386	386	E			E:94% H:4% K:1%	
X ₁₃₉	387	387	V		LI	V:82% L:12% I:4% F:1%	
	388	388	Q			Q:96% K:4%	
	389	389	G			G:88% D:12%	
	390	390	F			F:100%	
X ₁₄₀	391	391			LFRVH	L:75% F:12% R:7% V:3% H:3%	
	392	392	I			I:100%	
	393	393	P			P:100%	
	394	394	K			K:100%	

	395	395	G			G:100%	
	396	396	T			T:100%	
	397	397	T			T:87% I:9% V:3% M:1%	
	398	398	L			L:88% V:6% I:6%	
X ₁₄₁	399	399			IFML	I:75% F:18% M:4% L:3%	
X ₁₄₂	400	400			TPI	T:79% P:18% I:3%	
	401	401	N			N:100%	
	402	402	L			L:93% M:7%	
	403	403	S			S:97% T:3%	
	404	404	S			S:100%	
	405	405	V			V:97% A:1% M:1%	
	406	406	L			L:100%	
	407	407	K			K:100%	
	408	408	D			D:100%	
X ₁₄₃	409	409			EK	E:75% K:25%	
X ₁₄₄	410	410			TAKS	T:78% A:16% K:4% S:1%	
	411	411	V			V:100%	
	412	412	W			W:100%	
X ₁₄₅	413	413			EK	E:72% K:28%	
	414	414	K			K:100%	
	415	415	P			P:100%	

X ₁₄₆	416	416			FLNH	F:66% L:29% N:3% H:1%
X ₁₄₇	417	417			RQHEC	R:79% Q:9% H:7% E:3% C:1%
	418	418	F			F:100%
	419	419	H			H:87% Y:13%
	420	420	P			P:100%
	421	421	E			E:97% G:3%
	422	422	H			H:100%
	423	423	F			F:100%
	424	424	L			L:100%
	425	425	D			D:99% N:1%
	426	426	A			A:94% R:4% E:1%
	427	427	Q			Q:84% D:9% E:4% N:3%
	428	428	G			G:100%
X ₁₄₈	429	429			HQRKN	H:37% Q:28% R:26% K:4% N:4%
	430	430	F			F:100%
	431	431	V			V:99% T:1%
	432	432	K			K:99% R:1%
X ₁₄₉	433	433			QPHR	Q:35% P:31% H:26% R:7%
	434	434	E			E:100%
	435	435	A			A:100%
	436	436	F			F:100%

X ₁₅₀	437	437			MLI	M:79% L:18% I:3%
	438	438	P			P:96% A:4%
	439	439	F			F:100%
	440	440	S			S:100%
	441	441	A			A:100%
	442	442	G			G:100%
	443	443	R			R:97% P:3%
	444	444	R			R:100%
X ₁₅₁	445	445			AVIS	A:53% V:22% I:15% S:10%
	446	446	C			C:100%
	447	447	L			L:100%
	448	448	G			G:100%
	449	449	E			E:100%
	450	450	P			P:87% Q:12% A:1%
	451	451	L			L:100%
	452	452	A			A:100%
	453	453	R			R:100%
	454	454	M			M:100%
	455	455	E			E:100%
X ₁₅₂	456	456			LF	L:79% F:21%
	457	457	F			F:100%

	458	458	L			L:97% I:3%	
	459	459	F			F:99% L:1%	
	460	460	F			F:100%	
	461	461	T			T:100%	
X ₁₅₃	462	462			CST	C:78% S:18% T:4%	
	463	463	L			L:100%	
	464	464	L			L:99% M:1%	
	465	465	Q			Q:100%	
	466	466	R			R:85% H:12% K:3%	
	467	467	F			F:100%	
	468	468	S			S:96% T:4%	
	469	469	F			F:99% I:1%	
	470	470	S			S:90% Q:6% V:3% C:1%	
	471	471	V			V:88% I:12%	
	472	472	P			P:100%	
X ₁₅₄	473	473			AVDET	A:41% V:21% D:16% E:12% T:10%	
	474	474	G			G:97% D:3%	
	475	475	Q			Q:96% E:4%	
	476	476	P			P:100%	
	477	477	R			R:82% Q:13% S:3% C:1%	
	478	478	P			P:100%	

	479	479	S			S:88% R:12%
X ₁₅₅	480	480		DHENT		D:60% H:15% E:13% N:9% T:3%
X ₁₅₆	481	481		HDRQLSY		H:78% D:12% R:3% Q:3% L:1% S:1% Y:1%
	482	482	G			G:94% P:4% R:1%
	483	483	V			V:90% H:4% A:3% F:1% N:1%
	484	484	F			F:93% P:3% Y:3% S:1%
X ₁₅₇	485	485		ASGVRLY		A:62% S:24% G:6% V:4% L:1% R:1% Y:1%
X ₁₅₈	486	486		FILAT		F:71% I:13% L:7% A:6% T:3%
	487	487	L			L:84% P:9% I:4% V:1% S:1%
X ₁₅₉	488	488		VMLTQN		V:60% M:21% L:6% T:6% Q:4% N:3%
X ₁₆₀	489	489		TIASVLF		T:49% I:22% A:16% S:6% V:4% L:1% F:1%
	490	490	P			P:100%
X ₁₆₁	491	491		SPAHYLCTRQ		S:43% P:21% A:13% H:12% Y:3% L:3% C:1% T:1% R:1% Q:1%
	492	492	P			P:93% D:6% R:1%
	493	493	Y			Y:100%
X ₁₆₂	494	494		QEK		Q:79% E:16% K:4%
	495	495	L			L:91% I:9%
	496	496	C			C:96% R:4%
	497	497	A			A:100%
	498	498	V			V:81% L:4% K:4% M:4% T:3% A:1% S:1%

X ₁₆₃	499	499		PVASML	P:72% V:15% A:4% S:4% M:3% L:1%
	500	500	R		R:100%

Table 10. Table of the amino acid sequence positions $X_1 - X_{163}$ of SEQ ID NO:180 with reference to CYP2D_N1 (SEQ ID NO:181). 'Alignment Position' represents the position within an alignment of extant and/or evolutionary intermediate sequences used to construct CYP2D_N1. 'Position Relative to CYP2D_N1' represents the respective position of CYP2D_N1 (SEQ ID NO:181). 'Conserved' indicates amino acids at the respective positions which were conserved amongst the extant and/or evolutionary intermediate sequences used to construct CYP2D_N1. 'Variable' indicates amino acids identified at the respective positions which varied amongst the extant and/or evolutionary intermediate sequences used to construct CYP2D_N1 at the respective residues $X_1 - X_{163}$. 'Frequency' indicates the percentage frequency of the individual amino acids detected amongst the extant and/or evolutionary intermediate sequences at the variable positions. The residues (alignment positions 1-36) which form the N-terminal anchor of the protein are shaded grey. Note that, as set forth in FIG. 8, in CYP2D_N1 The the second residue has been modified to alanine.

X index	CPR_N1 Position	Variable (Joint, Marginal)
X_1	2	G, D
X_2	9	G, S
X_3	10	P, T
X_4	28	V, I
X_5	56	P, K
X_6	62	P, S
X_7	64	L, V
X_8	285	P, S
X_9	287	T, S
X_{10}	414	A, E
X_{11}	674	C, S

Table 11. Table of the amino acid sequence positions $X_1 - X_{11}$ of SEQ ID NO:321 with reference to CPR_N1 (SEQ ID NO:322). 'Variable' represents the amino acids at the respective positions that were present in ancestral CPR sequences constructed using joint likelihood and marginal likelihood methods as hereinabove described. Note that CPR_N1 represents the sequence constructed using the joint maximum likelihood method.

REFERENCES

- Alix, B., Boubacar, D.A. & Vladimir, M. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. Nucl. Acids Res. 40, W573-W579 (2012).
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. J Mol Biol 215, 403-410 (1990).
- Ashkenazy, H. et al. FastML: a web server for probabilistic reconstruction of ancestral sequences. Nucleic Acids Res 40, W580-584 (2012).
- Bauer, S. & Shiloach, J. Maximal exponential growth rate and yield of E.coli obtainable in a bench-scale fermentor. Biotechnol. Bioeng. 16, 933-941 (1974).
- Behrendorff, J.B.Y.H. in School of Chemistry and Molecular Biosciences, Vol. Ph.D. (The University of Queensland, 2011).
- Chang, T. K. and Waxman, D. J. (2006) Enzymatic Analysis of cDNA-Expressed Human CYP1A1, CYP1A2, and CYP1B1 With 7-Ethoxyresorufin as Substrate. In Cytochrome P450 Protocols (Phillips, I. R. and Shephard, E. A., eds.). pp. 85-90, Springer, Humana Press, Totowa, N.J.
- Gadagkar, S.R. & Kumar, S. Maximum likelihood outperforms maximum parsimony even when evolutionary rates are heterotachous. Mol Biol Evol 22, 2139-2141 (2005).
- Gascuel, O. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14, 685-695 (1997).
- Gillam, E.M.J., Baba, T., Kim, B.-R., Ohmori, S. & Guengerich, F.P. Expression of modified human cytochrome P450 3A4 in Escherichia coli and purification and reconstitution of the enzyme. Arch. Biochem. Biophys. 305, 123-131 (1993).
- Gillam, E.M.J., Guo, Z.Y., Martin, M.V., Jenkins, C.M. & Guengerich, F.P. Expression of cytochrome P450 2D6 in Escherichia coli: Purification, and spectral and catalytic characterization. Arch. Biochem. Biophys. 319, 540-550 (1995).
- Gillam, E.M.J. et al. Expression of cytochrome P450 3A5 in Escherichia coli: Effects of 5' modification, purification, spectral characterization, reconstitution conditions, and catalytic activities. Arch. Biochem. Biophys. 317, 374-384 (1995).

- Goodman, D.B., Church, G.M. & Kosuri, S. Causes and Effects of N-Terminal Codon Bias in Bacterial Genes. *Science* 342, 475-479 (2013).
- Guengerich, F.P. in *Principles and Methods of Toxicology*, Edn. 3. (ed. A.W. Hayes) 1259-1313 (Raven Press, Ltd., New York; 1994).
- 5 Hunter, D.J.B. et al. Facile production of minor metabolites for drug development using a CYP3A shuffled library. *Metab. Eng.* 13, 682-693 (2011).
- Isin, E.M. & Guengerich, F.P. Kinetics and thermodynamics of ligand binding by cytochrome P450 3A4. *J. Biol. Chem.* 281, 9127-9136 (2006).
- Isin, E.M. & Guengerich, F.P. Substrate binding to cytochromes P450. *Analytical and Bioanalytical Chemistry* 392, 1019-1030 (2008).
- 10 Jones, D.T., Taylor, W.R. & Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8, 275-282 (1992).
- Johnston, W.A., Huang, W., De Voss, J.J., Hayes, M.A. & Gillam, E.M.J. A shuffled CYP1A library shows both structural integrity and functional diversity.
- 15 Drug Metab. Dispos. 35, 2177-2185 (2007).
- Johnston, W.A., Huang, W., Hayes, M.A., De Voss, J.J. & Gillam, E.M.J. Quantitative whole cell cytochrome P450 measurement suitable for high throughput application. *J. Biomol. Screen.* 13, 135-141 (2008).
- Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid
- 20 multiple sequence alignment based on fast Fourier transform. *Nucl. Acids Res.* 30, 3059-3066 (2002).
- Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. & Dahlquist, F.W. Expression and N-15 labeling of proteins for proton and N-15 nuclear magnetic resonance. *Methods Enzymol.* 177, 44-73 (1989).
- 25 Notley, L.M., de Wolf, C.J.F., Wunsch, R.M., Lancaster, R.G. & Gillam, E.M.J. Bioactivation of tamoxifen by recombinant human cytochrome P450 enzymes. *Chem. Res. Toxicol.* 15, 614-622 (2002).
- Pupko, T., Pe'er, I., Shamir, R. & Graur, D. A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol Biol Evol* 17, 890-896
- 30 (2000).
- Rawal, S., Yip, S.S.M. & Coulombe, R.A. Cloning, expression and functional characterization of cytochrome P450 3A37 from turkey liver with high aflatoxin B-1 epoxidation activity. *Chem. Res. Toxicol.* 23, 1322-1329 (2010).
- Rowland, P. et al. Crystal structure of human cytochrome P450 2D6. *J. Biol.*

Chem. 281, 7614-7622 (2006).

Sambrook, J., Fritsch, E.F. & Maniatis, T. Molecular cloning : a laboratory manual, Edn. 2nd. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; 1989).

- 5 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30, 2725-2729 (2013).

- von Wachenfeldt, C., Richardson, T.H., Cosme, J. & Johnson, E.F. Microsomal P450 2C3 is expressed as a soluble dimer in Escherichia coli following
10 modifications of its N-terminus. Arch. Biochem. Biophys. 339, 107-114 (1997).
Williams, E.M., Copp, J.N. & Ackerley, D.F. in Directed Evolution Library Creation, Vol. 1179. (eds. E.M.J. Gillam, J.N. Copp & D.F. Ackerley) 83-101 (Humana Press, Totowa, N.J.; 2014).

- Zadeh, J.N. et al. NUPACK: Analysis and design of nucleic acid systems. Journal
15 of Computational Chemistry 32, 170-173 (2011).

Throughout the specification, the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Various changes and modifications may be made to the
20 embodiments described and illustrated without departing from the present invention.

The disclosure of each patent and scientific document, computer program and algorithm referred to in this specification is incorporated by reference in its entirety.

CLAIMS

1. An isolated protein comprising the amino acid sequence set forth in SEQ ID NO:1, or an amino acid sequence at least 80% identical to SEQ ID NO:1, wherein X_1 - X_{219} in SEQ ID NO:1, or in the amino acid sequence at least 80% identical to SEQ ID NO:1, may be any amino acid.
5
2. The isolated protein of Claim 1, wherein one or more of the residues X_1 - X_{219} is selected from the respective group of variable amino acids as set forth in Table 9.
3. The isolated protein of Claim 1 or Claim 2, wherein said isolated protein
10 comprises an amino acid sequence set forth in any one of SEQ ID NOS:2-41 or SEQ ID NOS:544-578, or an amino acid sequence at least 80% identical to any one of SEQ ID NOS:2-41 or SEQ ID NOS:544-578.
4. An isolated protein comprising the amino acid sequence set forth in SEQ ID
15 NO:180, or an amino acid sequence at least 80% identical to SEQ ID NO:180, wherein residues X_1 - X_{163} in SEQ ID NO:180, or in the amino acid sequence at least 80% identical to SEQ ID NO:180, may be any amino acid.
5. The isolated protein of Claim 4, wherein one or more of the residues X_1 - X_{163} is
20 selected from the respective group of variable amino acids as set forth in Table 10.
6. The isolated protein of Claim 4 or Claim 5, wherein said isolated protein comprises an amino acid sequence set forth in any one of SEQ ID NOS:181-250, or an amino acid sequence at least 80% identical to any one of SEQ ID NOS:181-250.
25
7. The isolated protein of any preceding claim, wherein said isolated protein has P450 enzyme activity.
8. An isolated protein comprising the amino acid sequence set forth in SEQ ID
30 NO:321, or an amino acid sequence at least 80% identical to SEQ ID NO:321, wherein residues X_1 - X_{11} in SEQ ID NO:321 or in the amino acid sequence at least 80% identical to SEQ ID NO:321, may be any amino acid.

9. The isolated protein of Claim 8, wherein one or more of the residues X_1 - X_{11} is selected from the respective group of variable amino acids as set forth in Table 11.
- 5 10. The isolated protein of Claim 8 or Claim 9, wherein said isolated protein comprises an amino acid set forth in any one of SEQ ID NOS:322-431, or an amino acid sequence at least 80% identical to any one of SEQ ID NOS:322-431.
- 10 11. The isolated protein of any one of Claims 8-10, wherein said isolated protein has P450 reductase enzyme activity.
12. An isolated protein that comprises an amino acid sequence set forth in any one of SEQ ID NOS:2-41, 181-250, 322-431, or 544-578, or an amino acid sequence at least 80% identical to any of these.
- 15 13. A method of producing or constructing an isolated protein, said method including the step of producing or constructing an engineered ancestral amino acid sequence of at least a fragment of a P450 protein or P450 reductase protein from one or more P450 protein or P450 reductase amino acid sequences that are different to the
- 20 engineered ancestral amino acid sequence.
14. The method of claim 13, wherein the engineered ancestral P450 protein or P450 reductase protein displays or possesses one or more increased or enhanced properties compared to at least one of the one or more P450 proteins or P450 reductase proteins
- 25 comprising amino acid sequences that are different to the engineered ancestral amino acid sequence.
15. An isolated engineered ancestral P450 protein or P450 reductase protein produced according to the method of Claim 13 or Claim 14.
- 30 16. The method of Claim 13 or Claim 14, or the protein of Claim 15, wherein the engineered ancestral P450 protein or P450 reductase protein comprises an amino acid sequence set forth in SEQ ID NOS: 1, 2, 180, 181, 321 and/or 322.

17. A method of producing or constructing a modified engineered ancestral P450 protein or P450 reductase protein, said method including the step of introducing one or more amino acid substitutions in an amino acid sequence of the engineered ancestral P450 protein or P450 reductase protein to thereby produce or construct the modified engineered ancestral P450 protein or P450 reductase protein.

18. The method of Claim 17, wherein the amino acid substitutions are non-conservative amino acid substitutions.

19. The method of Claim 18, wherein the modified engineered ancestral P450 protein or P450 reductase protein displays or possesses one or more increased or enhanced properties compared to the engineered ancestral P450 protein or P450 reductase protein.

20. A modified engineered ancestral P450 protein or P450 reductase protein produced according to the method of any one of Claims 17-19.

21. The method of any one of Claim 17-19, or the protein of Claim 20, wherein the modified engineered ancestral P450 protein or P450 reductase protein comprises an amino acid sequence set forth in SEQ ID NOS: 1, 3-41, 180, 182-250, 321, 323-431 and/or 544-578, or an amino acid sequence at least 80% identical to SEQ ID NOS: 1, 3-41, 180, 182-250, 321, 323-431 and/or SEQ ID NOS: 544-578.

22. An isolated nucleic acid encoding the isolated protein of any one of Claims 1-12, 15, 16, or 20, or produced according to the method of any one of Claims 13, 14 or 16-21, or a fragment or derivative of said isolated protein.

23. A genetic construct comprising the isolated nucleic acid of Claim 22.

24. A host cell comprising the genetic construct of Claim 23.

25. An antibody or antibody fragment which binds and/or is raised against an isolated protein of any one of Claims 1-12, 15, 16, 20, or 21.

26. A composition for performing a chemical reaction, said composition comprising one or more of the isolated proteins of any one of Claims 1-12, 15, 16, 20, or 21, and one or more buffers, solvents and/or other reagents suitable for the chemical reaction.
- 5 27. A method of performing a chemical reaction, said method including the step of exposing a substrate molecule to one or more of the isolated proteins of any Claims 1-12, 15, 16, 20, or 21, to thereby perform the chemical reaction.
28. A reaction product produced according to the method of Claim 27.
- 10 29. The composition of Claim 26 or the method of Claim 27, wherein the one or more isolated proteins include a protein having P450 enzyme activity and a protein having P450 reductase activity.
- 15 30. Use of the isolated protein of any one of Claims 1-12, 15, 16, 20, or 21, or the composition of Claim 26 or Claim 29 for structure-activity relationship analysis; pharmacological testing; bioremediation; or biosensor technology.

SEQ ID NO:2 (CYP3_N1)

MALLLAVFLTLLILYGIWPYGLFKKLGI PGPRPLPFIGTFLEYRKGFLDFDTECFKKYGKIW
 GFYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFG LNGLPESAI SIAEDEQWKRI RTVLS
 PTFTSGKLKEMFPI MKHYGDILVKNIQKKA EKDEPVDMKDIFGAYSMDVITSTSF SVDIDSI
 NNPNDPFVTNIKKLLKFNFNPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFKNAVKKIKE
 DRKKDNHKKDRVDFLQLMIDSQTSEENHSHKALTDVEILAQAII FIFAGYETTSSTLSFLAYN
 LATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLD MVINETLRLYPIGGRLERVCKKTVE
 INGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTFLPFGAGPRNCIG
 MRFALLTMKVAIVRLLQNF SFR TCKETQIPLELNTQGFLTPKKPIVLKLVPR

SEQ ID NO:1

With reference to the above sequence (SEQ ID NO:2)

X ₁	is position	32
X ₂	is position	37
X ₃	is position	39
X ₄	is position	40
X ₅	is position	42
X ₆	is position	45
X ₇	is position	47
X ₈	is position	48
X ₉	is position	49
X ₁₀	is position	52
X ₁₁	is position	53
X ₁₂	is position	55
X ₁₃	is position	56
X ₁₄	is position	60
X ₁₅	is position	61
X ₁₆	is position	64
X ₁₇	is position	65
X ₁₈	is position	68
X ₁₉	is position	69
X ₂₀	is position	71
X ₂₁	is position	73
X ₂₂	is position	75
X ₂₃	is position	78
X ₂₄	is position	79
X ₂₅	is position	83
X ₂₆	is position	91
X ₂₇	is position	97
X ₂₈	is position	98
X ₂₉	is position	99
X ₃₀	is position	100
X ₃₁	is position	101
X ₃₂	is position	103
X ₃₃	is position	104
X ₃₄	is position	105
X ₃₅	is position	106

X ₃₆	is position	108
X ₃₇	is position	110
X ₃₈	is position	111
X ₃₉	is position	112
X ₄₀	is position	115
X ₄₁	is position	121
X ₄₂	is position	122
X ₄₃	is position	126
X ₄₄	is position	131
X ₄₅	is position	137
X ₄₆	is position	139
X ₄₇	is position	140
X ₄₈	is position	141
X ₄₉	is position	142
X ₅₀	is position	143
X ₅₁	is position	144
X ₅₂	is position	145
X ₅₃	is position	147
X ₅₄	is position	148
X ₅₅	is position	149
X ₅₆	is position	150
X ₅₇	is position	151
X ₅₈	is position	152
X ₅₉	is position	153
X ₆₀	is position	155
X ₆₁	is position	156
X ₆₂	is position	157
X ₆₃	is position	158
X ₆₄	is position	160
X ₆₅	is position	161
X ₆₆	is position	162
X ₆₇	is position	164
X ₆₈	is position	165
X ₆₉	is position	168
X ₇₀	is position	174
X ₇₁	is position	176
X ₇₂	is position	178
X ₇₃	is position	180
X ₇₄	is position	182
X ₇₅	is position	183
X ₇₆	is position	186
X ₇₇	is position	190
X ₇₈	is position	195
X ₇₉	is position	196
X ₈₀	is position	197
X ₈₁	is position	200
X ₈₂	is position	202
X ₈₃	is position	204
X ₈₄	is position	206

X ₈₅	is position	207
X ₈₆	is position	209
X ₈₇	is position	210
X ₈₈	is position	211
X ₈₉	is position	212
X ₉₀	is position	213
X ₉₁	is position	214
X ₉₂	is position	215
X ₉₃	is position	219
X ₉₄	is position	220
X ₉₅	is position	222
X ₉₆	is position	223
X ₉₇	is position	224
X ₉₈	is position	225
X ₉₉	is position	226
X ₁₀₀	is position	227
X ₁₀₁	is position	228
X ₁₀₂	is position	229
X ₁₀₃	is position	230
X ₁₀₄	is position	233
X ₁₀₅	is position	234
X ₁₀₆	is position	235
X ₁₀₇	is position	236
X ₁₀₈	is position	237
X ₁₀₉	is position	239
X ₁₁₀	is position	240
X ₁₁₁	is position	241
X ₁₁₂	is position	242
X ₁₁₃	is position	243
X ₁₁₄	is position	244
X ₁₁₅	is position	245
X ₁₁₆	is position	246
X ₁₁₇	is position	248
X ₁₁₈	is position	249
X ₁₁₉	is position	251
X ₁₂₀	is position	252
X ₁₂₁	is position	253
X ₁₂₂	is position	254
X ₁₂₃	is position	255
X ₁₂₄	is position	257
X ₁₂₅	is position	266
X ₁₂₆	is position	267
X ₁₂₇	is position	270
X ₁₂₈	is position	271
X ₁₂₉	is position	272
X ₁₃₀	is position	273
X ₁₃₁	is position	274
X ₁₃₂	is position	275
X ₁₃₃	is position	276

X ₁₃₄	is position	277
X ₁₃₅	is position	279
X ₁₃₆	is position	281
X ₁₃₇	is position	283
X ₁₃₈	is position	285
X ₁₃₉	is position	286
X ₁₄₀	is position	287
X ₁₄₁	is position	289
X ₁₄₂	is position	290
X ₁₄₃	is position	302
X ₁₄₄	is position	303
X ₁₄₅	is position	307
X ₁₄₆	is position	308
X ₁₄₇	is position	310
X ₁₄₈	is position	314
X ₁₄₉	is position	317
X ₁₅₀	is position	318
X ₁₅₁	is position	319
X ₁₅₂	is position	323
X ₁₅₃	is position	327
X ₁₅₄	is position	328
X ₁₅₅	is position	329
X ₁₅₆	is position	335
X ₁₅₇	is position	336
X ₁₅₈	is position	338
X ₁₅₉	is position	339
X ₁₆₀	is position	340
X ₁₆₁	is position	341
X ₁₆₂	is position	342
X ₁₆₃	is position	348
X ₁₆₄	is position	350
X ₁₆₅	is position	353
X ₁₆₆	is position	357
X ₁₆₇	is position	359
X ₁₆₈	is position	360
X ₁₆₉	is position	361
X ₁₇₀	is position	367
X ₁₇₁	is position	370
X ₁₇₂	is position	376
X ₁₇₃	is position	377
X ₁₇₄	is position	382
X ₁₇₅	is position	383
X ₁₇₆	is position	385
X ₁₇₇	is position	386
X ₁₇₈	is position	388
X ₁₇₉	is position	389
X ₁₈₀	is position	390
X ₁₈₁	is position	393
X ₁₈₂	is position	396

X ₁₈₃	is position	397
X ₁₈₄	is position	399
X ₁₈₅	is position	403
X ₁₈₆	is position	405
X ₁₈₇	is position	412
X ₁₈₈	is position	415
X ₁₈₉	is position	416
X ₁₉₀	is position	418
X ₁₉₁	is position	421
X ₁₉₂	is position	423
X ₁₉₃	is position	427
X ₁₉₄	is position	439
X ₁₉₅	is position	440
X ₁₉₆	is position	441
X ₁₉₇	is position	444
X ₁₉₈	is position	446
X ₁₉₉	is position	447
X ₂₀₀	is position	448
X ₂₀₁	is position	449
X ₂₀₂	is position	452
X ₂₀₃	is position	456
X ₂₀₄	is position	457
X ₂₀₅	is position	462
X ₂₀₆	is position	466
X ₂₀₇	is position	467
X ₂₀₈	is position	468
X ₂₀₉	is position	469
X ₂₁₀	is position	470
X ₂₁₁	is position	472
X ₂₁₂	is position	473
X ₂₁₃	is position	474
X ₂₁₄	is position	476
X ₂₁₅	is position	477
X ₂₁₆	is position	480
X ₂₁₇	is position	483
X ₂₁₈	is position	484
X ₂₁₉	is position	485

Figure 1

SEQ ID NO:181 (CYP2D_N1)

MALLSGSNLGFALGVALTVFLLLVD FMKRRKRWC SRFP GPMSLPFVGNLLQVDFHNLHYSFN
 QLSKRYGDVFSLQLCWKPVVVLNGFEAVREALVQKSEDTADRPPLPIYEH LGFGRKSEGVVL
 ARYGQAWKELRRFTLSTLRDFGMGKKSLEERVREEAGFLCAAFSSEEGRPFDPHFLNNAVS
 NVICSIIYGERFEYDDKKFQKLLRLLEESLKEESGFLPQILNVVPWLLRIPGLAQKVFQGQK
 ALLAFLELIKEHKETWDP AHPTRDFIDAF LLEMEKAKGNPESSFNEENLRLTTADLFSAGT
 ETTSTTLRWGLLFMLLYPDVQRKVHEEIDKVIGRDRRPTMGDQVKMPYTN AVIHEVQRYGDI
 VPLGVPHMTYRDTELQGFFIPKGTTVITNLSSVLKDETVWEKPFQFYPEHFLDADGQFVKPE
 AFMPFSAGRRVCLGEQLARMEFLFFTCLLQRF SFQIPEGQPRPREDGVFAFL LAPHDYQIC
 AVPRSTHHHHHH

SEQ ID NO:180

With reference to the above sequence (SEQ ID NO:181)

X ₁	is position	41
X ₂	is position	43
X ₃	is position	45
X ₄	is position	52
X ₅	is position	55
X ₆	is position	56
X ₇	is position	57
X ₈	is position	58
X ₉	is position	59
X ₁₀	is position	60
X ₁₁	is position	61
X ₁₂	is position	62
X ₁₃	is position	63
X ₁₄	is position	65
X ₁₅	is position	66
X ₁₆	is position	68
X ₁₇	is position	76
X ₁₈	is position	77
X ₁₉	is position	79
X ₂₀	is position	81
X ₂₁	is position	84
X ₂₂	is position	88
X ₂₃	is position	90
X ₂₄	is position	93
X ₂₅	is position	96
X ₂₆	is position	97
X ₂₇	is position	98
X ₂₈	is position	101
X ₂₉	is position	107
X ₃₀	is position	110
X ₃₁	is position	111
X ₃₂	is position	112
X ₃₃	is position	115

X ₃₄	is position	117
X ₃₅	is position	118
X ₃₆	is position	123
X ₃₇	is position	124
X ₃₈	is position	126
X ₃₉	is position	129
X ₄₀	is position	130
X ₄₁	is position	138
X ₄₂	is position	139
X ₄₃	is position	154
X ₄₄	is position	157
X ₄₅	is position	161
X ₄₆	is position	162
X ₄₇	is position	165
X ₄₈	is position	168
X ₄₉	is position	169
X ₅₀	is position	170
X ₅₁	is position	171
X ₅₂	is position	173
X ₅₃	is position	176
X ₅₄	is position	178
X ₅₅	is position	179
X ₅₆	is position	182
X ₅₇	is position	183
X ₅₈	is position	186
X ₅₉	is position	190
X ₆₀	is position	193
X ₆₁	is position	194
X ₆₂	is position	195
X ₆₃	is position	196
X ₆₄	is position	201
X ₆₅	is position	204
X ₆₆	is position	205
X ₆₇	is position	206
X ₆₈	is position	207
X ₆₉	is position	209
X ₇₀	is position	210
X ₇₁	is position	211
X ₇₂	is position	212
X ₇₃	is position	213
X ₇₄	is position	214
X ₇₅	is position	215
X ₇₆	is position	217
X ₇₇	is position	219
X ₇₈	is position	223
X ₇₉	is position	224
X ₈₀	is position	225
X ₈₁	is position	226
X ₈₂	is position	229

X ₈₃	is position	230
X ₈₄	is position	232
X ₈₅	is position	235
X ₈₆	is position	241
X ₈₇	is position	244
X ₈₈	is position	245
X ₈₉	is position	246
X ₉₀	is position	249
X ₉₁	is position	250
X ₉₂	is position	251
X ₉₃	is position	252
X ₉₄	is position	253
X ₉₅	is position	255
X ₉₆	is position	256
X ₉₇	is position	258
X ₉₈	is position	259
X ₉₉	is position	262
X ₁₀₀	is position	263
X ₁₀₁	is position	265
X ₁₀₂	is position	268
X ₁₀₃	is position	280
X ₁₀₄	is position	282
X ₁₀₅	is position	286
X ₁₀₆	is position	290
X ₁₀₇	is position	298
X ₁₀₈	is position	299
X ₁₀₉	is position	300
X ₁₁₀	is position	302
X ₁₁₁	is position	303
X ₁₁₂	is position	307
X ₁₁₃	is position	311
X ₁₁₄	is position	312
X ₁₁₅	is position	315
X ₁₁₆	is position	318
X ₁₁₇	is position	320
X ₁₁₈	is position	336
X ₁₁₉	is position	340
X ₁₂₀	is position	342
X ₁₂₁	is position	344
X ₁₂₂	is position	345
X ₁₂₃	is position	346
X ₁₂₄	is position	347
X ₁₂₅	is position	349
X ₁₂₆	is position	351
X ₁₂₇	is position	354
X ₁₂₈	is position	355
X ₁₂₉	is position	358
X ₁₃₀	is position	360
X ₁₃₁	is position	363

X ₁₃₂	is position	366
X ₁₃₃	is position	370
X ₁₃₄	is position	372
X ₁₃₅	is position	373
X ₁₃₆	is position	376
X ₁₃₇	is position	377
X ₁₃₈	is position	380
X ₁₃₉	is position	387
X ₁₄₀	is position	391
X ₁₄₁	is position	399
X ₁₄₂	is position	400
X ₁₄₃	is position	409
X ₁₄₄	is position	410
X ₁₄₅	is position	413
X ₁₄₆	is position	416
X ₁₄₇	is position	417
X ₁₄₈	is position	429
X ₁₄₉	is position	433
X ₁₅₀	is position	437
X ₁₅₁	is position	445
X ₁₅₂	is position	456
X ₁₅₃	is position	462
X ₁₅₄	is position	473
X ₁₅₅	is position	480
X ₁₅₆	is position	481
X ₁₅₇	is position	485
X ₁₅₈	is position	486
X ₁₅₉	is position	488
X ₁₆₀	is position	489
X ₁₆₁	is position	491
X ₁₆₂	is position	494
X ₁₆₃	is position	499

Figure 2

SEQ ID NO: 322 (CPR_N1)

MGASDTESGPSLPEPAAEEEPFLSMLDVFLLSLILGLLVYWFFFRKKKKEEIPFEKPIQPVAP
TLRETSFIEKMKKTGKNIVVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEECDMEELSR
LGEIDNSLAVFCMATYGECDPTDNAQDFYDWLQEDDADLSGLKYAVFGLGNKTYEHFNAMGK
YVDKRLEELGAQRIFDLGLGDDDDANIEEDFITWRERFWPAVCEHFGIEASGEDGSMRQYSLV
EHTDINKNKIYTGEMARLHSYENQKPPFDAKNPFLAPVTVNRELHKGGGRSCMHIELDITGS
KIRYEAGDHSVAVYPTNDAALVEELGKLLNTNLDTVFSLNNVDEESSKKHPPFCPTTYRTALT
HYLDITNPPRTNVLHELAEYASDPKDKEFLRKMASSSPEGKKQYQSWVVESRRNILAILQDL
PSCRPPIDHLCCELLPRLQARYYSIASSSKVHPNSIHICAVVVEYETKTGRVNKG VATNWLKN
KVPATDNGHKPRVPIYVRKSQFRLPFKPSTPVIMIGPGTGLAPFIGFIQERAYLKEQGKEVG
ETVLYFGCRKKNEDFIYQEELEQYQKEGVLTQLNMAFSRDQAQKVYVQHLLIKQNKESVWKL
HEDNAHIYVCGDARNMARDVQNALYEIVAEGKNMTQTQAVDYIKKLMTKGRYSCDVWS

SEQ ID NO: 321

With reference to the above sequence (SEQ ID NO:322)

X₁ is position 2
X₂ is position 9
X₃ is position 10
X₄ is position 28
X₅ is position 56
X₆ is position 62
X₇ is position 64
X₈ is position 285
X₉ is position 287
X₁₀ is position 414
X₁₁ is position 674

Figure 3

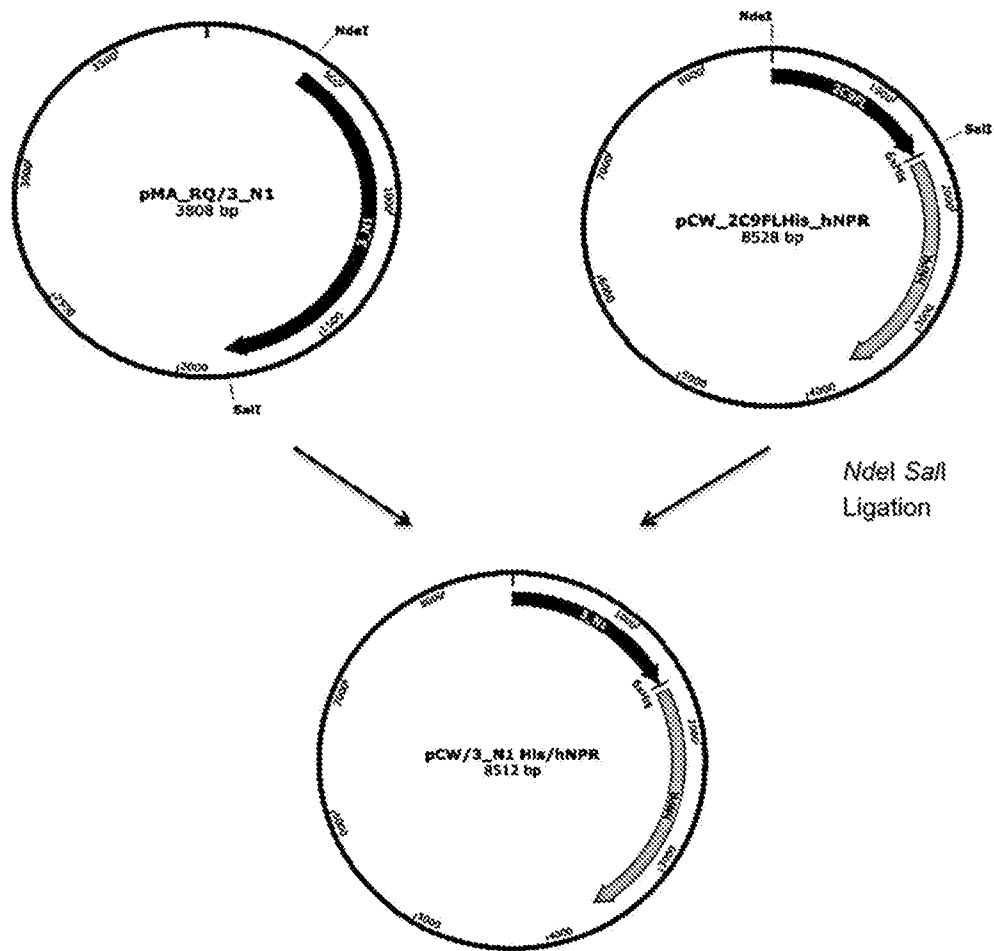


Figure 4

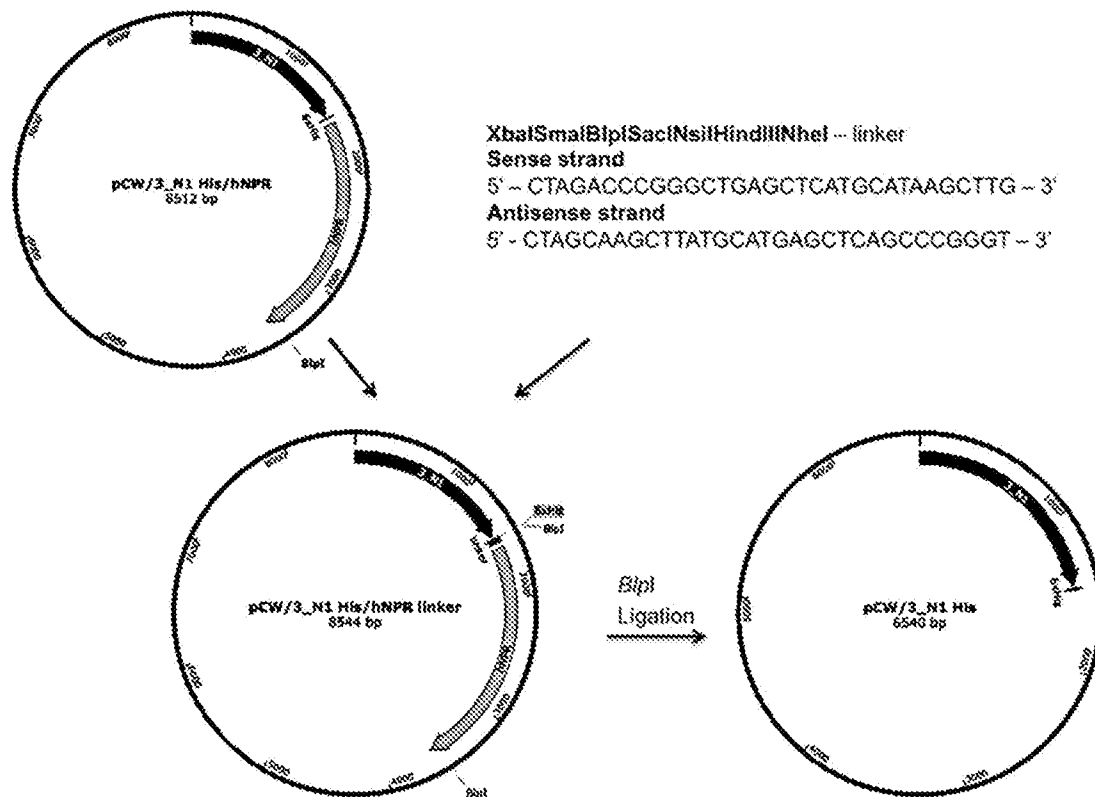


Figure 5

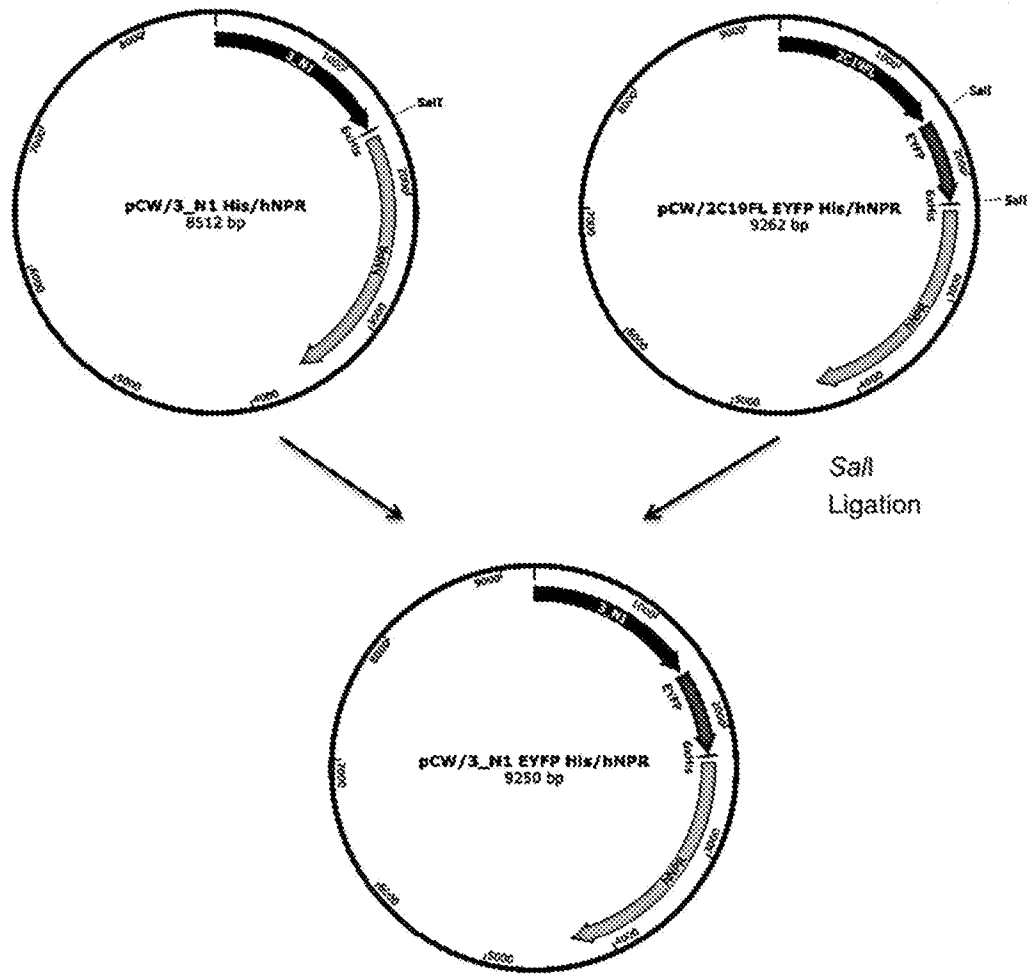


Figure 6

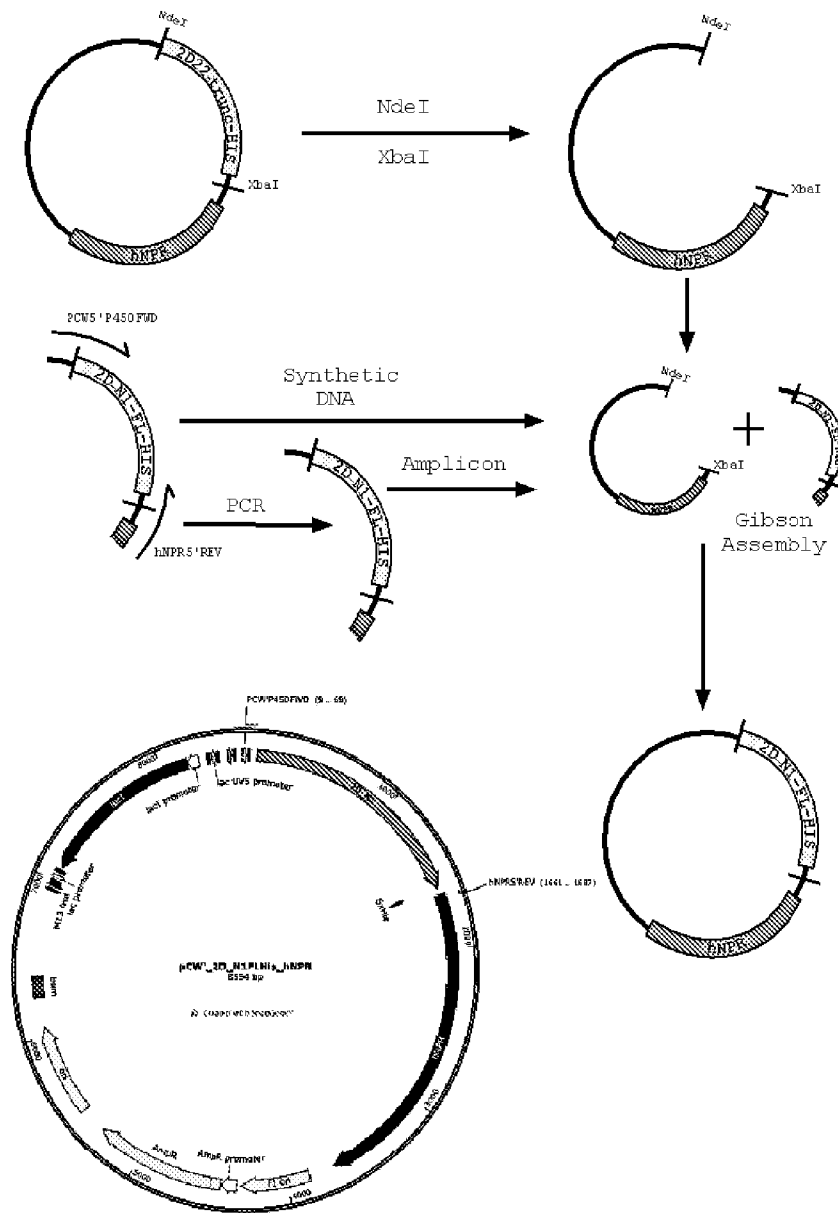


Figure 7

```

CYP2D_N1-nat      MGLLSGSNLFALGVALTVFLLLVDLMKRRKRWCSEPPPGPMSLPFVGNLLQVDFHNL
CYP2D_N1-FL       MALLSGSNLFALGVALTVFLLLVDLMKRRKRWCSEPPPGPMSLPFVGNLLQVDFHNL
CYP2D_N1-trunc    MAKKTSSKGK-----FPPGPMSLPFVGNLLQVDFHNL
CYP2D_N4-nat      MGLLTGDALFPLAVAVAIFFLLVLDLMHRRQRWAARYPPGPMPPLPGLGNLLQVDFQNM
CYP2D_N4-FL       MALLTGDALFPLAVAVAIFFLLVLDLMHRRQRWAARYPPGPMPPLPGLGNLLQVDFQNM
CYP2D_N4-trunc    MAKKTSSKGK-----YPPGPMPPLPGLGNLLQVDFQNM

CYP2D6 DB2:       M-----ALLAVFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB3:                               MHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB4:       MGL---EALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB5:                               MHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB6:       MAL---EALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB7:                               MARRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB8:       MALLLAVFL-----AIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB9:       MALLLAVFL---AVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB10:      MALLLAVFL---MHRRQRWAARYPPMAARRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6 DB11:      MALLLAVFL---MHRRQRWAARYPPMAARRQRWAARYPPGPLPLPGLGNLLHVDFQNT
CYP2D6-trunc      MAKKTSSKGK-----YPPGPLPLPGLGNLLHVDFQNT

CYP2D22-FL        MALPTGAELWPPIAIFTVIFILIVNLMHWRQRWTAHYPPGPMPWPVLGNLLHMDFQNM
CYP2D12-FL        MALLTGTDLWSVAIFTVIFILLVDLMHRRQSWTSCYPPGPVPWPVLGNLLQVDLNNM
CYP2D12-trunc     MAKKTSSKGK-----YPPGPVPWPVLGNLLQVDLNNM

```

Figure 8

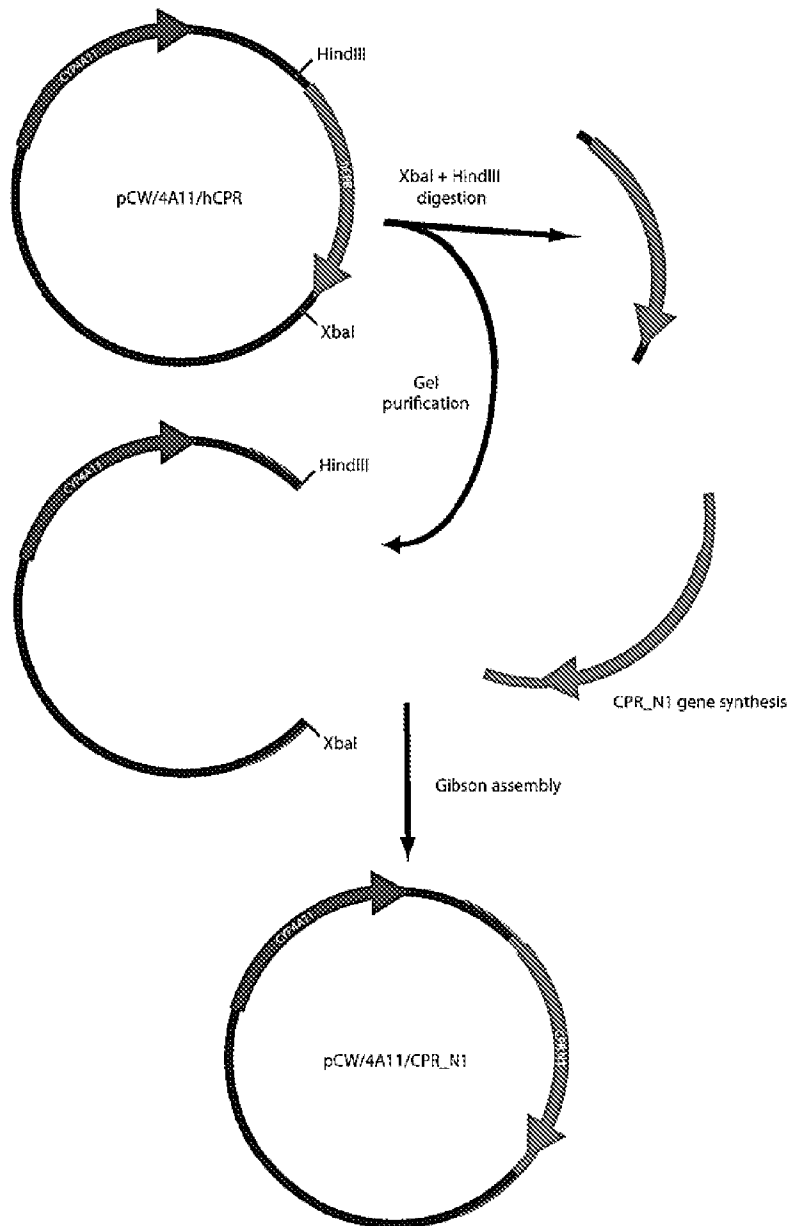


Figure 9

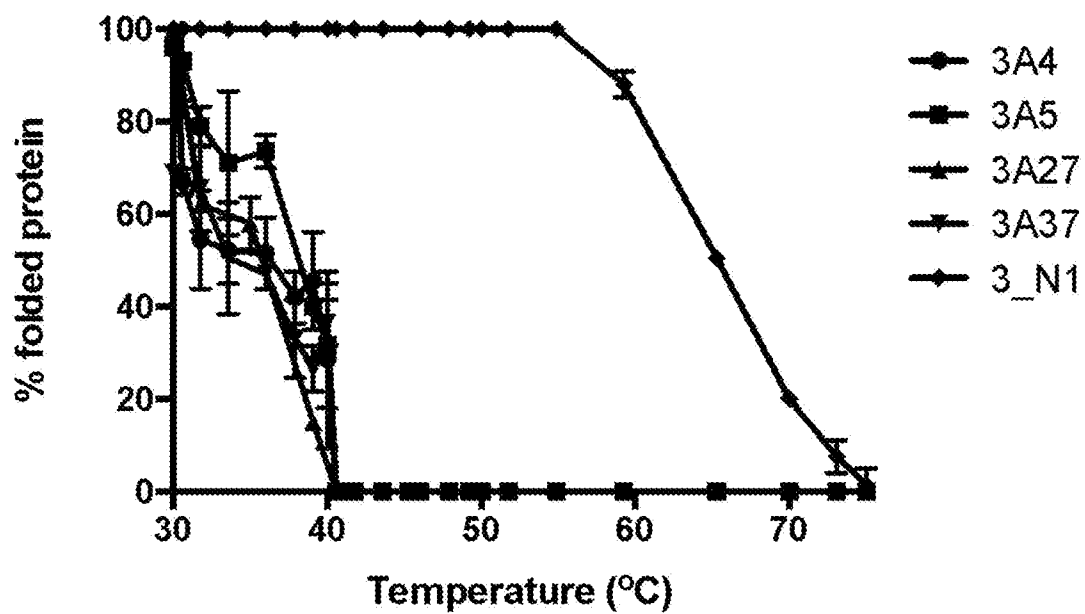


Figure 10

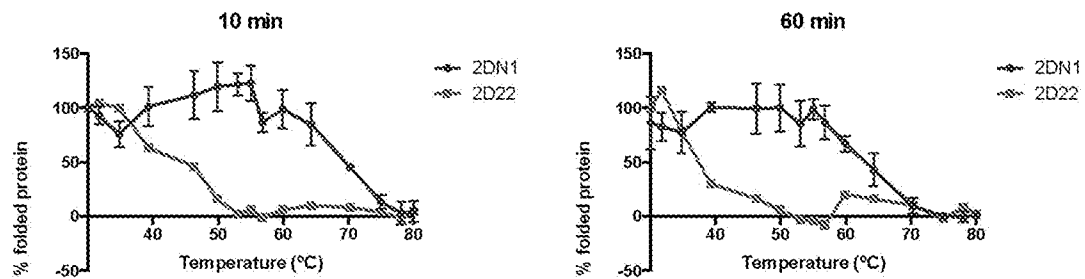


Figure 11

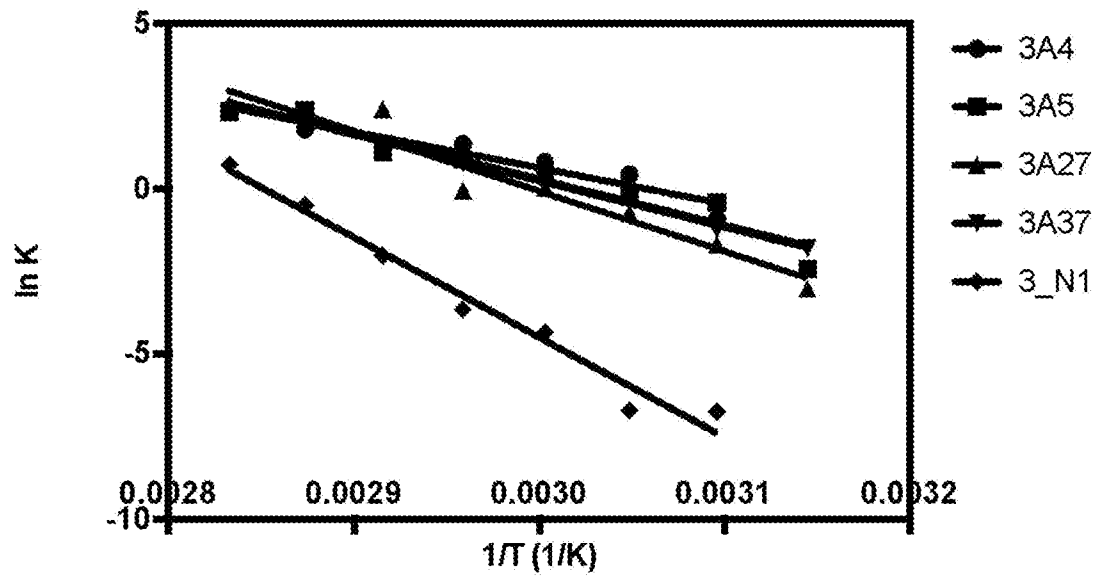


Figure 12

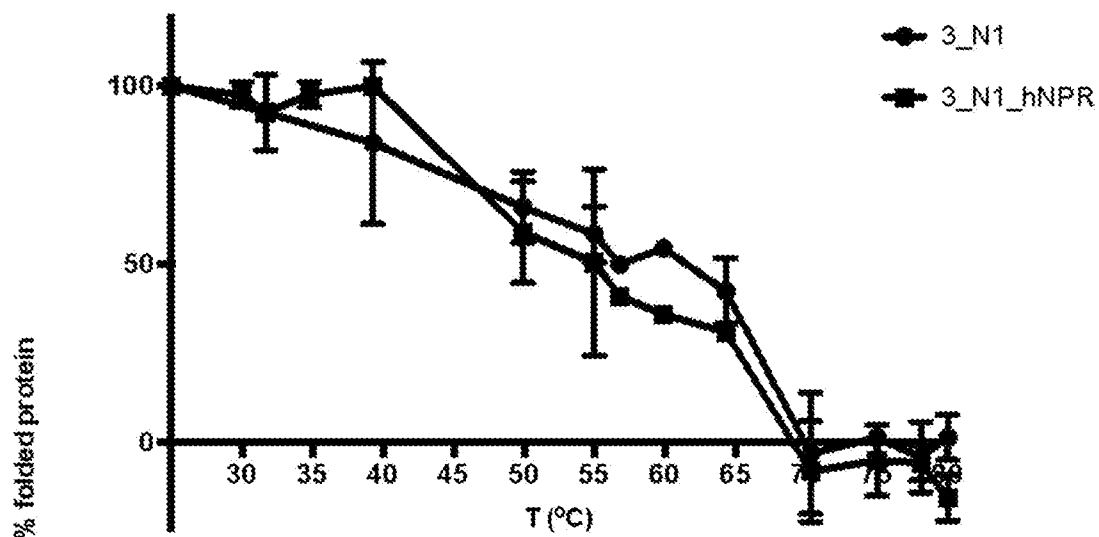


Figure 13

60 Minute Thermostability Assay

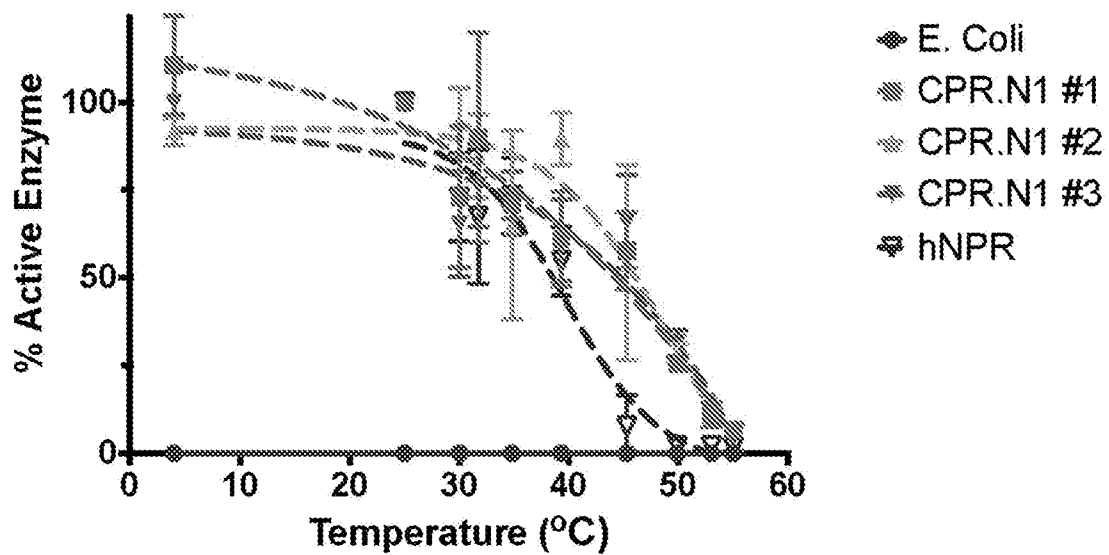


Figure 14

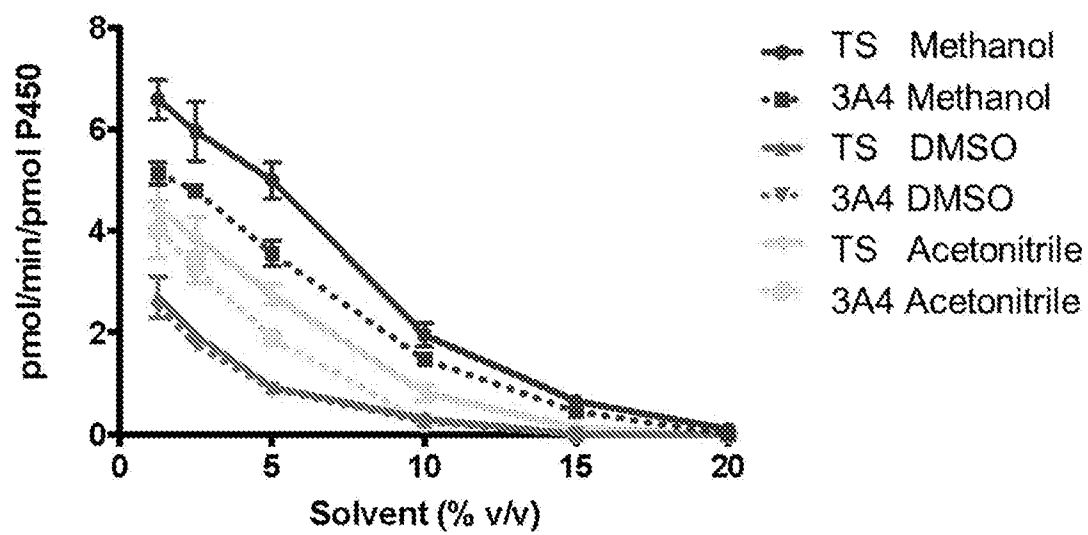


Figure 15

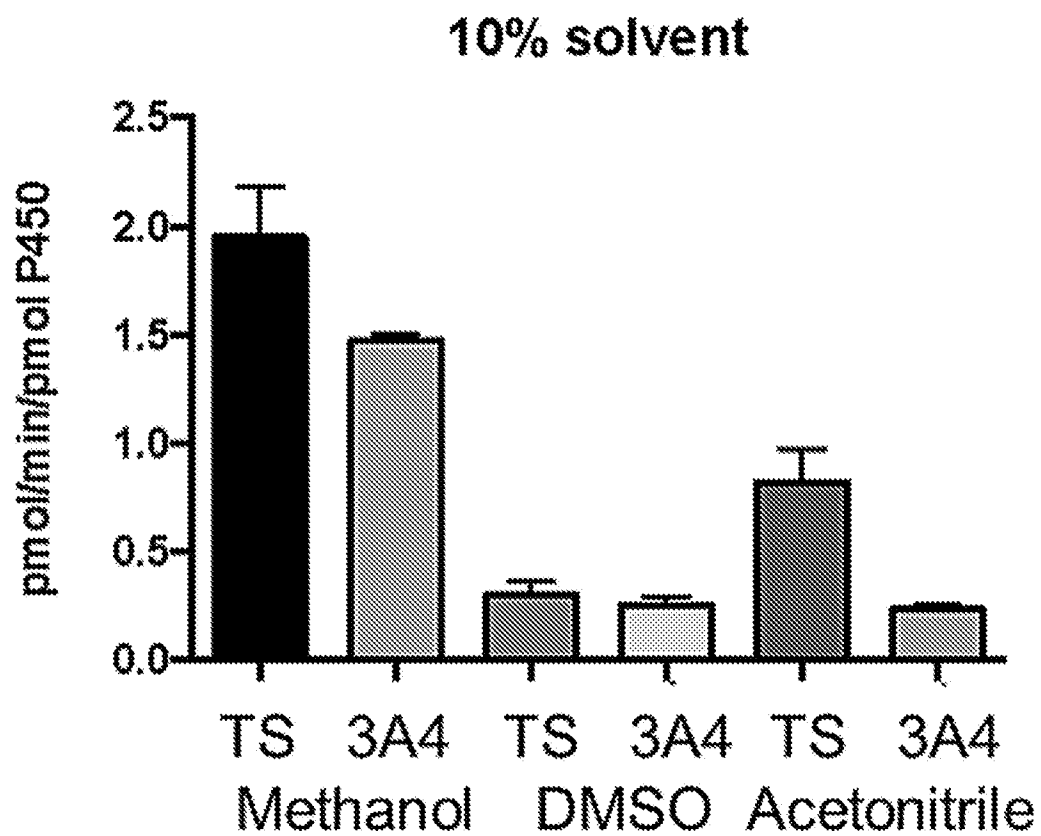


Figure 16

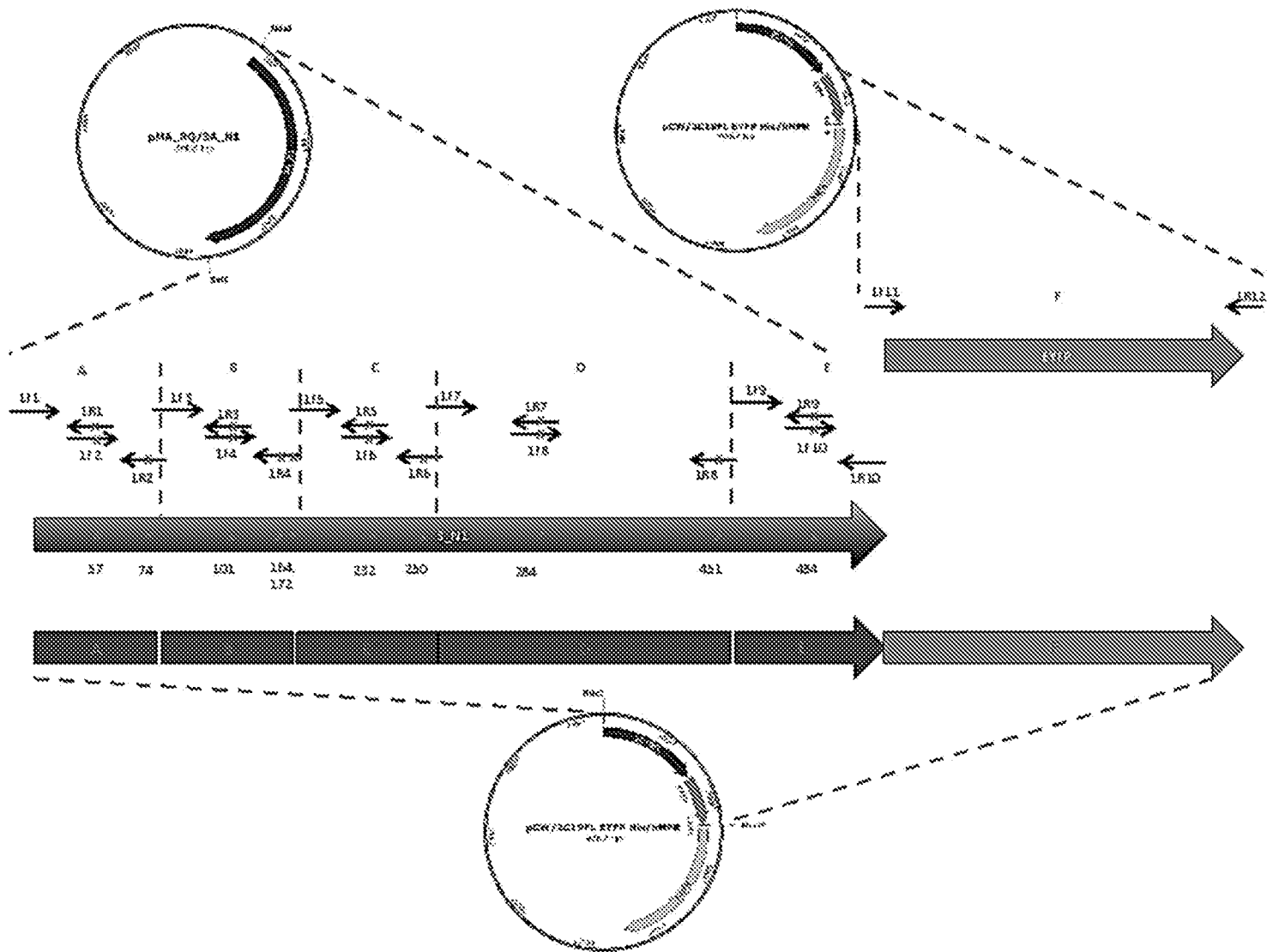


Figure 17

Ligand binding – summary

Class	Ligand	CYP _{TS}	Parent
Steroids	Cholesterol	Weak binding	Weak binding
	Progesterone	13 μ M	126 μ M
	Androstenedione	31 μ M	20 μ M
	DHEA	30 μ M	6.8 μ M
	Testosterone	24 μ M	6.2 μ M
	DHT	4.9 μ M	18 μ M
	Estradiol	4.6 μ M	6.6 μ M
	Cortisol	37 μ M	40 μ M
Macrolides	Erythromycin	8.3 μ M	3.9 μ M
	Cyclosporin A	4 μ M	3.6 μ M
	Troleandomycin	9 μ M	-
Benzodiazepines	Midazolam	4.2 μ M	-
	Diazepam	22 μ M	-
Inhibitor (Type II)	Ketoconazole	0.1 μ M	0.05 μ M

Ligand binding – continued

Class	Ligand	CYP _{TS}	Parent
Misc. drugs	Caffeine	no spectra	-
	Theophylline	no spectra	no spectra
	Quinidine	weak binding	weak binding
	Pravastatin	weak binding	weak binding
	Etoposide	no spectra	weak binding
	Tamoxifen	weak binding	weak binding
	R-warfarin	weak binding	weak binding
	S-warfarin	no spectra	-

Figure 18

>YG_3N1_1_001/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLQPKKPIVLKLVPR

>YG_3N1_1_002/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLQPKKPIVLKLVPR

>YG_3N1_1_003/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

> YG_3N1_1_004/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_005/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLQPKKPIVLKLVPR

>YG_3N1_1_006/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLQPKKPIVLKLVPR

>YG_3N1_1_007/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_008/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_009/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF

NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_010/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_011/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_012/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_013/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_014/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSNFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_015/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_016/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_017/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSNFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LVKNIQKKA EKDEPVDVKDIFGAYSMDVITSTSFSDVIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDPMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_018/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_019/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_020/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_021/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_022/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_023/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_024/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_025/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYP IGGRLERVCKKTVEINGVTIPKGTVMIPVYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_026/1-486

MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF

NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_027/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_028/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_029/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_030/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_031/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_032/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_033/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GIYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR
>YG_3N1_1_034/1-486
MALLLAVFLTLILYGIWPHYGLFKKLGIPIPRPLPFIGTFLEYRKGFLDFDTECFKKYKGIW
GLYDGRQPVLAIMDPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRIRTVLSPTFTSGKLKE
MFPIMKHYGDIIVKNIQKAEKDEPVDMDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNFF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_035/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYGKIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDILVKNIQKKVEKDEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLDMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_036/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGFDFDTECFKKYGKIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDILVKNIQKKADEPVDMDKIDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLDMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_037/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGFDFDTECFKKYGKIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDILVKNIQKKADEPVDMDKIDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLDMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_038/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGFDFDTECFKKYGKIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDILVKNIQKKADEPVDVKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLDMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLSMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

>YG_3N1_1_039/1-486
MALLLAVFLTLLILYGIWPHYGLFKKLGIPGPRPLPFIGTFLEYRKGILDFDTECFKKYGKIW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDILVKNIQKKADEPVDMDKIDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPILEKMNISLFPKDVMDFFMNAVKKIKEDRKKDNHKDRVDFLQLMIDSQTSEESHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQQKLQEEIDATLPNKAPPTYDALMQMEYLDMVI
NETLRLYPIGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRTCKETQIPLLELNTQGFLTPKKPIVLKLVPR

Figure 19

>N1

MGLASGMLLVALLVCLTVMVLM SVWQQRKSKWCGKLPPGPTPLPFI GNYLQLNTEQM YNSLMKIS
 ERYGPVFTIHLGPRRVVVL CGHDAVREALVDQAE EFSGRGEQATFDWVG YGRKSEG VVVSNGER
 AKQLRRFSIATLRDFGVGKRGIEERI QEEAGFLIDALRG TGGANIDPTFFLSRTVSNVISSIVFG
 DRFDYKDKEFLSLRLMMLGIFQFTSTSTGQLYEMFSVMKHLPGPQQQAFQLLQGLEDFIAKKVEH
 NQRTLDPNSPPRDFIDSFLIRMQEE EKNPNTEFY LKNLVM TLTNLFI GGTETVSTTLRYGFLLLM
 KHPEVEAKVHEEIDRVI GKNRQPKFEDRAKMPYMEAVI HEIQRF GDIVPMS LARRVKKDTKFRDF
 FLPKGTEVF PMLGSLVRDP SFFSNPQDFNPQHFLNEKGQFKKSDAFV PFSIGKRNCFGEGLARME
 LFLFFTVMQNFR LKSSQSQPKDIDVSHVGFATIPRNYTMSFLPR

>N2

MGLLSGSNLFALGVALTVFLLLVD FMKRRKRWC SRFP PGMSLPFVGNLLQVDFHNLHYSFNQLS
 KRYGDVFSLQLCWKPVVVLNGFEAVREALVQKSEDTADRPPLP IYEH LGFGRKSEG VVLARYGQA
 WKELRRFTLSTLRDFGMGKKSLEERVREEAGFLCAAFS SEEGRPDPHFLNNAVSNVICSI IYG
 ERFEYDDKKFQKLLRLLEESLKEESGFLPQILNVVPWLLRIPGLAQKVFGQKALLAFLELIKE
 HKETWDPAPHTRDFIDAF LLEMEKAKGNP ESSFNEENLRLTTADLF SAGTETTSTTLRWGLLFML
 LYPDVQQRKVHEEIDKVI GRDRRPTMGDQVKMPYTNAVI HEVQRYGDIVPLGVPHMTYRDTELQGF
 FIPKGTTVITNLSSVLKDET VWEKPFQFYPEHFLDADGQFVKPEAFMPFSAGRRVCLGEQLARME
 LFLFFTCLLQRF SFQIPEGQPRPREDGVFAFL LAPHDYQICAVPR

>N3

MGLLSGSNLFALGVALTVFLLLVD FMKRRKRWC SRYP PGMSLPFVGNLLQVDFHNLHYSFNQLS
 KRFGDVFSLQLCWKPVVVLNGFEAVREALVQKSEDTADRPPLP IYEH LGFGRKSEG VVLARYGQA
 WKELRRFTLSTLRNFGMGKKSLEERVTEEAGFLCAAFS SEEGRPDPHFLNNAVSNVICSI IYG
 ERFEYDDKKFQKLLRLLEESLKEESGFLPQILNVVPWLLRIPGLAQKVFGQKALLAFLELIKE
 HKETWDPAPHTRDFIDAF LLEMEKAKGNP ESSFNEENLRLVTADLF SAGTETTSTTLRWALLFML
 LYPDVQQRKVHEEIDKVI GRDRRPTMGDQVKMPYTNAVI HEVQRYGDIVPLGVPHMTYRDTELQGF
 FIPKGTTVITNLSSVLKDET VWEKPFQFYPEHFLDADGQFVKPEAFMPFSAGRRVCLGEQLARME
 LFLFFTCLLQRF SFQIPEGQPRPREDGVFAFL LAPHPYQICAVPR

>N4

MGLLTGDALFPLAVAVAI FLLLVDLMHRRQRWAARYPP GPMPLPGLGNLLQVDFQNM PYSFNQLR
 RRF GDVFSLQLAWKPVVVLNGLAAVREALVTHSEDTADRPPVP IYEH LGFGPRS QGVILARYGPA
 WREQRRFSVSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPNALLNKAVSNVIASLIYG
 RRF EYDDPRFLRLLDLLEDGLKEESGFLPQVLNAIPVLLRIPGLAGKVFP GQKAFMALLDELLE
 HRMTWDPAPPRDLTDAFLAEME KAKGNP ESSFN DENLRLVADLF SAGMVTTSTTLAWALLMI
 LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPYTNAVI HEVQRF GDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDET VWEKPF RFHPEHFLDAQGRFVKPEAFMPFSAGRRACLGEPLARME
 LFLFFTCLLQRF SFVSPAGQPRPSDHGVFAFLVTPSPYQLCAVPR

>N5

MGLLTGDALVPLAVAVAI FLLLVDLMHRRQRWAARYPP GPMPLPGLGNLLQVDFQNT PYSFNQLR
 RRF GDVFSLQLAWTPVVVLNGLAAVREALVTHSEDTADRPPVP IYEH LGFGPRS QGVILARYGPA
 WREQRRFSVSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPNALLNKAVSNVIASLT YG
 RRF EYDDPRFLRLLDLTEDGLKEESGFLPQVLNAIPVLLRIPGLAGKVFP GQKAFMALLDELLE
 HRMTWDPAPPRDLTDAFLAEME KAKGNP ESSFN DENLRLVADLF SAGMVTTSTTLAWALLMI
 LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPYT TAVI HEVQRF GDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDET VWEKPF RFHPEHFLDAQGRFVKPEAFMPFSAGRRACLGEPLARME
 LFLFFTCLLQRF SFVSPAGQPRPSDHGVFAFLVTPSPYQLCAVPR

>N6

MGLLTGDALGPLAVAVAI FLLLVDLMHRRQRWAARYPP GPMPLPGLGNLLQVDFQDTP SSFNQLR
 RRF GDVFSLQLAWTPVVVLNGLAAVREALVTHSEDTADRPPVP IYEH LGFGPRS QGVILARYGRA
 WREQRRFSVSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPNALLNKAVSNVIASLT FG
 RRF EYDDPRFLRLLDLTEDGLKEESGFLPQVLNAIPVLLRIPGLAGKVFP GQKAFMALLDELVTE
 HRMTRDDAPPRDLTDAFLDEMEKAKGNP ESSFN DENLRLVADLF SAGMVTTSTTLAWALLMI
 LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPF TMAVVHEVQRF GDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDET VWEKPF RFHPEHFLDAQGRFVKQEA FMPFSAGRRACLGEPLARME
 LFLFFTCLLQRF SFVSPAGQPRPSDHGVFAFLVTPAPYQLCAVPR

>N7

MGLLTGDALGPLAVAVAFIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQDTPSSFNQLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPVPIYHLGFGPRSQGVILARYGRA
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSNALLNKAVSNVIASLTFG
RRFEYDDPRFLKLLDLTEDGLKEESGFLRQVLNAIPVLLRIPGLAGKVFPGQKAFMALLDELVTE
HRMTRDPAQPPRDLTDAFLDEMEKAKGNPESSFNENLRLVVADLF SAGMVTSTSTLAWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPFTMAVVHEVQRFGDIVPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVWEKPFRRFHPEHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVSPAGQPRPSDHGVFAFLVTPAPYQLCAVPR

>N8

MGLLTGDALGPLAVAVAFIFLLLVDLMHRRRRWAARYPPGPMPLPLLGNLLQMDFQDTLCYFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVILARYGRE
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTFG
RRFEYDDPRLLKLLDLTQQGLKEDSGFLREVLNAIPVLLRIPGLAGKVFSAQKAFMTLLDELVQE
HRMTRDPTQPPRDLTDAFLDEIEKAKGNPESSFNENLSMVTADLF IAGMVSTSTSTLAWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPFTMAVVHEVQRFGDIVPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVWKKPFRFHPEHFLDAQGRFVKQEAFFMPFSAGRRVCLGEPLARME
LFLFFTCLLQRFSSVSPAGQPRPSDHGVFSFLVTPAPYQLCAVPR

>N9

MGLLTGDALGPLAVAVAFIFLLLVDLMHRRRRWAARYPPGPMPLPLLGNLLQMDFQDTLCYFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVFLARYGRE
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLLKLLDLTQQGLKEDSGFLREALNSIPVLLRIPGLAGKVFSAQKAFMTLLNELVQE
HRMTRDPTQPPRDLTDAFLDEIEKAKGNPESSFNENLSMVTADLF IAGMVSTSTSTLAWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQTRMPFTMAVVHEVQRFGDIVPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVWKKPFRFHPEHFLDAQGRFVKQEAFFMPFSAGRRVCLGEPLARME
LFLFFTCLLQRFSSVSPAGQPRPSDHGVFSFLVIPAPYQLCAVPR

>N10

MGLLTGDALGPLAIAVALFIFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDTLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRHLKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLF LAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGDLAPLGVPHTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRFSSVSPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N11

MGLLTGDALGPLAIAVALFIFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDTLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRHLKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLF LAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGDLAPLGVPHTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVSPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N12

MGLLTGDALGPLAIAVALFIFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDTLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRHLKLIDGALKGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLF LAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGDLAPLGVPHTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVSPVQPRPSDHGVFSFLMIPPPYQLCAMPR

>N13

MGLLTGDALGPLAIAVALFIFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDTLFYFDRLR
RRFGDVFSLQLAWTPMVVLNGLAAREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRHLKLIDGALKGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE

HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAMPR

>N14

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPMVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLSAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALQGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFASAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAMPR

>N15

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFVREALNSIPVLLHIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCATPR

>N16

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFVREALNSIPVLLNIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGSPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCATPR

>N17

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAVPR

>N18

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAVPR

>N19

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVNLGLEAVREALVYHSEDTADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVITSLTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPQDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAVPR

>N20

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
 RRFQDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
 WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVITSLTYG
 RRFQYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVLSSQKALMTLLNELVQE
 HRITRDPTQPPQDLTEAFLDEIEKARGNPESFNENMLMTADLFLAGIVSTSATLTWALLLMI
 LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFQDGLAPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRVCLGEPLARME
 FFLFFTCLLQRFSSVPPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N21

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
 RRFQDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
 WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVITSLTYG
 RRFQYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLNELVQE
 HRITRDPTQPPQDLTDAFLDEIEKARGNPESFNENMLMTADLFLAGMLSTSATLAWALLLMI
 LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFQDGLVPLGMPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRVCLGEPLARME
 FFLFFTCLLQRFSSVPPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N22

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
 RRFQDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
 WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVITSLTYG
 RRFQYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLSELVQE
 HRITRDPTQPPQDLTDAFLDEIEKARGNPESFNENMLMTADLFLAGMLSTSATLAWALLLMI
 LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFQDGLVPLGMPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRVCLGEPLARME
 FFLFFTCLLQRFSSVPPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N23

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
 RRFQDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
 WREQRRFALSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVITSLTYG
 RRFQYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLSELVQE
 HRITRDPTQPPQDLTDAFLDEIEKARGNPESFNENMLMTADLFLAGMLSTSATLAWALLLMI
 LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFQDGLVPLGMPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGFVKQEAFFMPFSAGRRVCLGEPLARME
 FFLFFTCLLQRFSSVPPVQPRPSDHGVFSFLMIPPPYQLCAVPR

>N24

MGLLTGDALGPLAVAVVIFLLLVDLMHRRSRWAPRYPPGPMPLPGLGNLLQVDFQDPRPSFNQLR
 RRFQDVFSLQQVWTPVVVLNGLAAVREALVYHSQDADRPPVPIYHLGYGPRSEGVIILARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSNDLLNKAVSNVIASLTFG
 RRFQYNDPRILKLLDLTEDGLKEESGLLRQVLEAIPVLLRIPGLAAKVFPQKAFMALIDELITE
 HKMTRDPAQPPRDLTDAFLDEMKEAKGNPESFNENMLMTADLFLAGMLSTSATLAWALLLMI
 LHPDVQRRVQREIDEVIGQVRRPEMGDQALMPFTMAVMHEVQRFQDGLVPLGMPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVEKPPFRFHPEHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
 LFLFFTSLQRFSSVPPAGQPRPSDHGVFAFLVTPAPYQLCAVPR

>N25

MGLLTGDLLGILALAMVIFLLLVDLMHRRSRWAPRYPPGPMPLPGLGNLLQVNFQDPRLSFIQLR
 RRFQDVFSLQQIWRPVVVVLNGLAAVREALVSHSHETSDRPPVPIYHLGYGPRSEGVIILARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEEWVTQEASCLCAAFADQAGRPFSNNLLNKAVSNVIASLTFA
 RRFQYNDPRMLKLLDLVLEGLKEEVGLMRQVLEAMPVLRHIPGLCAKLFPRQKAFVLMIDELITE
 HKMTRDLAQPPRDLTDAFLDEMKEAKGNPESFNENMLMTADLFLAGMLSTSATLAWALLLMI
 LHPDVQRRVQREIDEVIGHVRQPEIKDQALMPFTLAVLHEVQRFQDGLVPLGVAHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVEKPPFRFYPEHFLDAQGRFTKQEAFFMPFSAGRRSCLGEPLARME
 LFLFFTLLQRFSSVPTGQPRPSDHGVFAFLFPSPYQLCAVPR

>N26

MGLLSGDTLGPLAVAVLIFLLLVDLMHRRSRWAPRYPPGPTPLPVLGNLLQVDFEDPRPSFNQLR
 RRFQGNVFSLQQVWTPVVVLNGLAAVREALVYRSQDADRPPPAVYHLGYGPRAEGVIILARYGKA
 WREQRRFSLSTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSKDLLNKAVSNVIASLTFG

CRFEYNDPRI IKLLDLTEDGLKEESGLVRQVVEAVPVLLRIPGLAAKVFPQGKAFMALIDELITE
QKMTRDPTQPPRHLDLDAFLDEMKEAKGNPESSFNENLRLVVADLF SAGMVTSTTLAWALLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQALMPFTVAVVHEVQRFADIVPLGLPHMTSRDIEVQGF
HIPKGTTLITNLSSVLKDETVWEKPFRRFHPEHFLDAQGRFVKQEAFFIPFSAGRRACLGEPLARME
LFLFFTSLQLQHFSSVPAGQPRPSDHGVFAFLVTPAPYQLCAVPR

>N27

MGLLSGDTLGLAVALLIFLLLLDLMHRRSRWAPRYPPGPTPLPVLGNLLQVDFEDPRPSFNQLR
RRFGNVFSLQQVWTPVVVLNGLAAVREALVYRSQDTADRPPPAVYEHLGYGPRAEGVILARYGDA
WREQRRFSLTTLRNFLGLGKKSLEQWVTEEASCLCAAFADQAGRPFSPMDLLNKAVSNVIASLTFG
CRFEYNDPRI IKLLDLTEDGLKEEFNLVRKVVEAVPVLLSIPGLAARVFPAQKAFMALIDELIAE
QKMTRDPTQPPRHLDLDAFLDEVKEAKGNPESSFNENLRLVVADLF SAGMVTSTTLAWALLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQALMPFTVAVVHEVQRFADIVPLGLPHMTSRDIEVQGF
HIPKGTTLITNLSSVLKDETVWEKPFRRFHPEHFLDAQGRFVKQEAFFIPFSAGRRACLGEPLARME
LFLFFTSLQLQHFSSVPAGQPRPSEHGVSFAFLVTPAPYQLCAVPR

>N28

MGLLTGDALGPLAVAVAFIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQDTPSSFNQLR
RRFGDVFSLQLAWTPVVVLNGLAAIREALVTHGEDTADRPPVPIEHLGFGPHAQGVILARYGHA
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADHAGRPFSPNALLNKAVSNVIASLTFG
RRFEYDDPRFLRLDLTEDILKEESGFLPQVLNAIPVLLRIPGLVGKVFPGQKAFMALLDELVTE
HRMTRDPAQPPRDLTDAFLDEVEKAKGNPESSFNDDNLRLLVVADLFAAGMVTSTTLAWALLMI
LHPDVQRRVQQEIDEVIGQARRPEMGDQARMPFTMAVVHEVQRFQDIVPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVWEKPFRRFHPEHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDHGVFATLVTPAPYQLCAVPR

>N29

MGLLSGEALAPLAVAVAFIFLLLVDLMHRRQRWAARYPPGPVGIPGLGNLLQVDFRGIPNCFRQLR
RRYGDVFSLQLAWTPVVVLNGLPAAIREALVITYGEDTADRPPAHTLEHLGFGPHAQGVIMARYGHA
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEATCLCAAFADHAGCPFSPSALLNKAVCNVIASLTFG
CRFEYDDHRLTRLMDLTETILKESTGILPQVLNVIPILLRIPGLVDKVFVRGQKAFMALLDELVTE
HRMTRDPAQPPRDLTDAFLDQVEKAKGNPESSFNDDNLRLLVVADLFAAGMVTTSITLSWALLMI
LHPDVQRRVQQEIDEVIGPARRPEMGDQARMPYTTAVVHEVQRFADIVPLGVPHQTSRDIEVQGF
LIPKGTTLITNLSSVLKDEAVWEKPFRRFHGPHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDQGAPATLVTPAPYQLCAVAR

>N30

MGLLTGDALVPLAVTVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLHVDFQNTYPYFNQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYEVLFVGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLC AAFADHAGRPFPRNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLDLTQEGLEESGFLREVLNAIPVLLRIPGLAGKVLRSQKAF LAQLDELLE
HRMTWDPAQPPRDLTEAFLAEME KAKGNPESSFNENLRLVVADLF SAGMVTSTTLAWGLLMI
LHPDVQRRVQQEIDDVIGQVRRPEMGDQARMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDEAVWEKPFRRFHPEHFLDAQGRFVKPEAFLPF SAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>N31

MGLLTGDALVPLAVTVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTYPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYQVLGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLC AAFADHAGRPFPRNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLDLTQEGLEESGFLREVLNAIPVLLRIPGLAGKVLRSQKAF LTQLDELLE
HRMTWDPAQPPRDLTEAFLAEME KAKGNPESSFNENLRIVVADLF SAGMVTSTTLAWGLLMI
LHPDVQRRVQQEIDDVIGQVRRPEMGDQARMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDEAVWEKPFRRFHPEHFLDAQGRFVKPEAFLPF SAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>N32

MGLLTGDALVPLAVTVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFKNTYPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTCGEDTADRPPVPIYQVLGIGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLC AAFADQAGRPFPRNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLDLTQEGLEESGFLREVLNAVPLLLRIPGLAGKVLRSQKAF LTQLDELLE
HRMTWDPAQPPRDLTEAFLAEME KAKGNPESSFNENLRIVVADLF SAGMVTSTTLAWGLLMI
LHPDVQRRVQQEIDNVI GQVRRPEMGDQARMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVQGF

LIPKGTTLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGRFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>N33

MELLTGDAVLPLAVTVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFKNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTCGEDTADRPPVPIYQVLGIGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFTDQAGRPFPRPNSLLDKAVSNVIASLTYG
RRFEYDDPRFLRLFDLTALKEESGFLREVLNAVPLLLRIPGLAGKVLRSQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNEENLRIVVADLF SAGMVTSTSTLAWGLLLMI
LHPDVQRRVQQEIDNVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIELQGF
LIPKGTTLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGRFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>N34

MELLTGDAVLPLAVTVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFKNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTCGEDTADRPPVPIYQVLGVGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFTDQAGRPFPRPNSLLDKAVSNVIASLTYG
RRFEYDDPRFLRLFDLTALKEESGFLREVLNAVPLLLRIPGLAGKVLRSQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNEENLRIVVADLF SAGMVTSTSTLAWGLLLMI
LHPDVQRRVQQEIDNVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIELQGF
LIPKGTTLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>N35

MGLLTGEALVPLAVIVTIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYQVLGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFADHAGRPFPRPNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLLDLAQEGLEESGFLREVLNAIPVLLHIPALAGKVLRFQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNENLRIVVADLF SAGIVTTSTSTLAWGLLLMI
LHPDVQRRVQQEIDDDVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIEVQGF
RIPKGTTLFTNLSSVLKDKAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTSLQHFSSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

>N36

MGLLTGEALVPLAVIVTIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYQILGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLLDLAQEGLEESGFLREVLNAIPVLLHIPALAGKVLRFQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNENLRIVVADLF SAGIVTTSTSTLAWGLLLMI
LHPDVQRRVQQEIDDDVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIEVQGF
RIPKGTTLFTNLSSVLKDKAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTSLQHFSSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

>N37

MGLLTGEALVPLAVIVTIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYQILGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLLDLAQEGLEESGFLREVLNAIPVLLHIPALAGKVLRFQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNENLRIVVADLF SAGIVTTSTSTLAWGLLLMI
LHPDVQRRVQQEIDDDVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIEVQGF
RIPKGTTLFTNLSSVLKDKAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTSLQHFSSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

>N38

MGLLTGEALVPLAVIVTIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQLR
RRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIYQILGFGPRSQGVFLARYGPA
WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDKAVSNVIASLTG
RRFEYDDPRFLRLLDLAQEGLEESGFLREVLNAVPLVLLHIPALAGKVLRFQKAFLTLQDELLETE
HRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNENLRIVVADLF SAGMVTSTSTLAWGLLLMI
LHPDVQRRVQQEIDDDVI GQVRRPEMGDQARMPYTTAVIHEVQRFGDIVPLGVTHMTSRDIEVQGF
RIPKGTTLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
LFLFFTSLQHFSSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

>N39

MGLLTGEALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGFLPLPGLGNLLHVDFQNTPYCFDQLR
 RRFQDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPIQTILGFGPRSQGVFLARYGPA
 WREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFPRPNGLLDKAVSNVIASLTG
 RRFQYDDPRFLRLLDLAQEGLEESGFLREVLNAVVPVLLHIPALAGKVLRFQKAFLTQDELLE
 HRMTWDPAQPPRDLTEAFLAEMEKAAGNPESFNENLRIVVADLF SAGMVTSTSTLAWGLLLMI
 LHPDVQRRVQQEIDDDVIGQVRRPEMGDQAHMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVQGF
 RIPKGTTLITNLSSVLKDEAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRACLGEPLARME
 LFLFFTSLQLQHFSSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

>N40

MGLLTGDALFPVAVAVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 CRFGDVFSLQLAWKPVVVLNGLAAVREALVTHSEDTADRPPVPIYHLGFGPRSQGVILAPYGP
 WREQRRFSVSTLRNFGLGKKSLEQWVTEEAACLCAAFADHAGHPFSPNALLNKAVCNVIASLIYA
 RRFQYDDPRFIRLLDLLEDTLKEESGFLPQVLNAIPVLLRIPGLAGKVFPGQKAFMALLDELLE
 HRMTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRIVVADLF TAGMVTSTSTLAWALLMI
 LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPYTNVAVIHEVQRFQDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVEKPFRRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
 LFLFFTCLLQRFSSVPAQGPQPSDHGVFAFLVTPSPYQLCAVPR

>N41

MGLLTGDALFPVAVAVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 CRFGDVFSLQLAWKPVVVLNGLAAVREALVNHSEDTADRPPVPIYHLGFGPRSQGVIMAPYGP
 WREQRRFSVSTLRNFGLGKKSLEQWVTEEAACLCAAFANHAGHPFSPNALLNKAVCNVIASLIYA
 RRFQYDDPRFIRLLDLLEDTLKEDSGFLPQVLNAIPVLLRIPGLAGKVFPAQKAFMALLDELLE
 HRMTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRIVVADLF TAGMVTSTSTLAWALLMI
 LHPDVQRRVQQEIDEVIGQVRRPEMGDQARMPFTNAVIVIHEVQRFQDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVEKPFRRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
 LFLFFTCLLQRFSSVPAQGPQPSDHGVFAFLVTPSPYQLCAVSR

>N42

MGLLTGDALFPVAVAVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 CRFGDVFSLQLAWKPVVVLNGLVAVREALVNHSEDTADRPPVINELLGFGPRAQGVIMAPYGP
 WREQRRFSVSTLRNFGLGKKSLEQWVTEEAACLCAAFANHAGHPFSPNALLNKAVCNVIASLIYA
 RRFQYDDPMFIRLLDLLEDTLKEDSGLLPQVLNAIPVLLRIPGVAAKVFPQKAFMALLDELLE
 HRTTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRIVVSDLF TAGMVTSTSTLAWALLMI
 LHPDVQRRVQQEIDEVIGQGRPEMGDQARMPFTNAVIVIHEVQRFQDIVPLGVPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVEKPFRRFHPEHFLDAEGHFVKHEAFMPFSAGRRACLGEPLARME
 LFLFFTCLLQRFSSVPEGPQPSDHGVFAFLVTPSPYQLCAVSR

>N43

MGLLTGDALFSVAVAVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 HQFGDVFSLQFAWTPVVVLNGLVAVREALVNHSEDTSDRPPVTNELLGFGPKAQGVIGAPYGP
 WREQRRFSVSTLRNFGLGKKSLEQWVTEEAACLCAAFANHAGHPFSPNALLNKAVCNVIASLIYA
 RRFQYDDPMFIRLLEFLEDTLREDSGLLLQVLNIPVLLRIPWVAAKVLPQRSFMALLDELLE
 HKTTWDPDQPPRDLTDAFLAEVHKAKGNSESSFNENLRIVVSDLF TAGMVTSTSTLSWALLMI
 LHPDVQRRVQQEIDEVIGQGRPEMGDQAHMPFTNAVIVIHEVQRFQDIVPMGVPHMTSRDIEVQGF
 LIPKGTILFTNLSSVLKDETVEKPFRRFHPEHFLDAEGHFVKREAFMPFSAGPRICLGEPLARME
 LFLFFTCLLQRFSSVPEGPQPSDRGVSLLLVTPSPYQLCAVSR

>N44

MGLLTGDALFSVAVAVAIFLLLVDLMHRRQRWAARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 HQFGDVFSLQFVWTPVVVVLNGLLAVREALVNNSTDTSDRPTLPTNALLGFGPKAQGVIGAYYGP
 WREQRRFSVSSLRNFGLGKKSLEQWVTEEAACLCAAFTNHAGQPFPCPKALLNKAVCNVISSLIYA
 RRFQYDDPMVRLLEFLEETLRENSSLKIQVLNIPVLLRIPCVAKVLSAQRSFIALNDKLLAE
 HNTGWAPDQPPRDLTDAFLTEMHKAQGNSESSFNENLRLLVSDLF GAGMVTSTSTLSWALLMI
 LHPDVQRRVQQEIDEVIGQGRPEMGDQAHMPFTNAVIVIHEVQRFADIVPMGVPHMTSRDIEVQGF
 LIPKGTMLFTNLSSVLKDETVEKPLHFFHFGHFLDAEGRFVKREAFMPFSAGPRICLGEPLARME
 LFLFFTSLQLQRFSSVPEGPQPSDRGAPYLVLVTPSPYQLCAVLR

>N45

MGLLTGDALWPVAIFTAIFLLLVDLMHRRQRWTARYPPGPMPLPGLGNLLQVDFQNMPLYSFYKLR
 CRFGDVFSLQLAWKPVVVLNGLAAVREALVTYSEDTADRPPVPIYHLGFGPRSQGVVLAPYGP
 WREQRRFSVSTLRNFGLGKKSLEQWVTEEAACLCAAFADHAGHPFSPNTLLNKAVCNVIASLIYA
 RRFQYDDPRFIRLLDLLKDTLEESGFLPQVLNAIPVLLRIPGLAGKVFPGQKAFMALLDELLE

HKMTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVADLFTAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMADQARMPYTNAVIHEVQRFADIVPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPAGQPRPSDHGVFAFLVTPSPYQLCAVPR

>N46

MRLLTGAELWPVAIFTAIFLLLVNLMHRRQRWTARYPPGPMPLPVLGNLLQVDFQNMPPAGFQKLR
CRFGDLFSLQLASEPVVVLNGLTALREALVKYSEDTADRPPPLHINDHLGFGPRSQGVVLARYGPA
WRQQRRESVSTFRHFGLGKKSLEQWVTEEARCLCAAFADHSGHPFSPNTLLNKAVCNVIASLLYA
CRFEYDDPRFIRLLDLLKDTLEESGFLPMFLNVFPVLLHIPGLAGKVFSGKRAFMAMLDLLETE
HKMTWDPTQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVADLFMAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVRCPEMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPGQPRPSDHGVFGALTTPSPYQLCALPR

>N47

MRLPTGAELWPVIAIFTVIFLLLVNLMHRRQRWTARYPPGMPWPVVLGNLLQVDFQNMPPAGFQKLR
CRFGDLFSLQLASESVVVLNGLTALREALVKYSEDTADRPPPLHFNDHLGFGPRSQGVVLARYGPA
WRQQRRESVSTFRHFGLGKKSLEQWVTEEARCLCAAFADHSGHPFSPNTLLDKAVCNVIASLLYA
CRFEYDDPRFIRLLDLLKDTLEESGFLPMFLNVFPMLLHIPGLVGKVFSGKRAFMAMLDLLETE
HKMTWDPTQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVADLFMAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVRCPEMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPGQPRPSDHGVFGALTTPSPYQLCALPR

>N48

MRMPTGSELWPVIAIFTIIFLLLVNLMHRRQRWTSRYPPGPVWPVVLGNLLQIDFQNMPPAGFQKLR
CRFGDLFSLQLAFESVVVLNGLPALREALVKYSEDTADRPPPLHFNDQSGFGPRSQGVVLARYGPA
WRQQRRESVSTFRHFGLGKKSLEQWVTEEARCLCAAFADHSGHPFSPNTLLDKAVCNVIASLLFA
CRFEYNDPRFIRLLDLLKDTLEESGFLPMMLNVFPMLLHIPGLLGKVFSGKKAFAVAMLDLLETE
HKVTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRVVVADLFMAGMVTSTTTLTWALLFMI
LHPDVQCRVQQEIDEVIGQVRRPEMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIEVQGF
LIPKGTTLITNLSSVLKDETVEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPTGQPRPSDYGVFGALTTPRPYQLCASPR

>N49

MRLPTGAELWPVIAIFTVIFLILVNLMHWRQRWTAHYPPGMPWPVVLGNLLHMDFQNMPPAGFQKLR
GRYGDLSLQLASESVVVLNGLTALREALVKHSEDTADRPPPLHFNDLLGFGPRSQGIIVLARYGPA
WRQQRRESVSTMHHFGLGKKSLEQWVTEEARCLCAAFADHTGHHPFSPNTLLDKAVCNVIASLLYA
CRFEYDDPRFIRLLGLLKETLKEEAGFLPMFLNVFPMLLRIPGLVGKVFPGKRAFMALDDELLAE
HKTTWDPTQPPRDLTDAFLAEVEKAKGNPESSFNENLRVVGDLFSAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVQCPMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIELQGF
LIPKGTTLITNLSSALKDETVEKPLCFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVDPGQPPSDHGVFRALTTPCPYQLCALPR

>N50

MGLLTGDGLWPVIAIFTAIFLLLVNLMHRRQRWTARYPPGPMPLPGLGNLLQVDFENMPYSLYKLQ
NRYGDVFSLQMAWKPVVINGLKAVREVLVTYGEDTADRPPVPIYHLGYGPKSQGVVLAPYGP
WREQRRESVSTLRNFGGLGKKSLEQWVTEEAGHLCDFTAQAGHPFNPNNTLLNKAVCNVIASLIYA
RRFEYEDPYFIRMLKLLKESLGEESGFIEVLNAIPVLLRIPGLAGKVFPGQKAFMALLDKLLTE
HKTTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVADLFTAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMADQARMPYTNAVIHEVQRFADIVPLNIPHMTSRDIEVQGF
LIPKGTTLIPNLSSVLKDETVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPAGQPRPSDHGVFAFLVTPSPYQLCAVPR

>N51

MGLLIGDGLWAVVIFTAIFLLLVNLMHRRQFWTARYPPGPMPLPGLGNLLQVDFENMPYSLYKLQ
NRYGDVFSLQMAWKPVVINGLKAVREVLVTYGEDTADRPPVPIYNHLGYGHKSQGVVLAPYGP
WREQRRESVSTLRDFGVGKKSLEQWVTEEAGHLCDFTKEAGHPFNPNSTLLNKAVCNVIASLIYA
RRFEYEDPFFNRMLKTLKESFGEDSGFIAEVLNAIPVLLRIPGLPGKVFPKLNAFMALLDKMLIE
HKTTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVADLFMAGMVTSTTTLWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMADQARMPYTNAVIHEVQRFADIVPNIPHMTSRDIEVQGF
LIPKGTTLIPNLSSVLKDETVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPAGQPRPSDHGVFAFLVTPSPYQLCAVAR

>N52

MGLLI GDDLWAVVIFTAIFLLLVLDLVHRRQFWTARYPPGPVPLPGLGNLLQVDFENMPYSLYKLQ
NRYGDVFSLQMAWKPVVINGLKAVRELLVTYGEDTADRPLMPIYNHLGYGHKSKGVVLAPYGPE
WREQRRFSVSTLRDFGVGKKSLEQWVTEEAGHLCDRAFTKEAEHPFNPSTLLSKAVSNVIASLIYA
RRFEYEDPFFNRMLKTLKESFGEDTGFIAEVLNAIPILLQIPGLPGKVFPKLNSFIALVDKMLIE
HKTTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVIDLDFMAGMVTSTSTLSWALLLMI
LHPDVQRRVHQEIDEVIGQVRRPEMADQARMPYTNAVIHEVQRFADIVPTNIPHMTSRDIKFQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDHGVYALPVTQPYPQLCAVAR

>N53

MALLIGDGLWSGVIFTALFLLLVLDLMHRRKFWRARYPPGPMPLPGLGNLLQVDFENMPYSLYKFQ
QRYGDVFSLQMAWKPVVINGLKAVREVLVNCGEDTADRPPVP IFNHLGYGHKSQGVVFARYGPQ
WREQRRFSVSTMTRDFGVGKKSLEQWVTEEAGHLCDRAFTQEAGHPFNPTTLLNKSVCNVISSLIYA
HRFDYEDPFFNKLLKTLQESFGEDSGFIAEVLNAVVPVLLRIPGLPGKAFPKLTAFMDSLYKMLIE
HKTTWDPAQPPRGLTDAFLAEVEKAKGRPESSFNENLRMVVADLFIAGMVTSTSTLSWALLLMI
LHPDVQSRVQQEIDDVIGQVRRPEMADQARMPYTNAVIHEVQRFQDIAPVNI PHMTSRDIEVQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGRFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDQGVFALPVTPTPYELCAVVR

>N54

MELLTGTGLWPVAIFTVIFILLVLDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLENMPYSLYKLQ
NRYGDVFSLQMGWKPVVINGLKAVREVLVTYGEDTADRPPVP IFEHLGYGPKSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEAGHLCDRAFTAQAGRPFNPNTLLNKAVCNVIASLIYA
RRFEYEDPYFIRMLKLLKESLGEESGFIEVLNTFPVLLRIPGLADKVFPQGKTFMAILDNLLTE
HKTTWDPAQPPRDLTDAFLAEIEKAKGNPESSFNENLRMVVSDLFTAGMVTSTSTLSWALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFADIVPLNLP HMTSRDIEVQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDHGVFAFLVSPSPYQLCAVPR

>N55

MELLTGTGLWPVAIFTVIFILLVLDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLENMPYSLYKLQ
NRYGDVFSLQMGWKPMVVINGLKAVREVLVTYGEDTADRPPVP IFEHLGVGPKSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEAGHLCDRAFTAQAGRPFNPNTLLNKAVCNVIASLIFA
RRFEYEDPYLIRMLKLLLEESL TEVSGFIEVLNTFPVLLRIPGLADKVFPQGKTFMAILDNLLTE
NRTTWDPAQPPRDLTDAFLAEIEKAKGNPESSFNENLRMVVSDLFTAGMVTSTSTLSWALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFADIAPLNLPRI TS RDIEVQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDHGVFAILVSPSPYQLCAVPR

>N56

MELLTGTGLWPVAIFTVIFILLVLDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMGWKPMVVINGLKAVKEVLVTYGEDTADRPPVP IFEYLVGKPRSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEEWVTEAGHLCDRAFTAQAGQPINPNTMLNKAVCNVIASLIFA
RRFEYEDPYLIRMLKLLLEESL TEVSGFIEVLNTFPVLLRIPGLADKVFPQGKTFMAILDNLLTE
NRTTWDPAQPPRDLTDAFLAEIEKAKGNPESSFNENLRMVVSDLFTAGMVTSTSTLSWALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIAPLNLPRI TS RDIEVQDF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSNHGVFAILVAPSPYQLCAVVR

>N57

MELLNGTGLWPMAIFTVIFILLVLDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLSNMPYSLYKLQ
HRYGDVFSLQMGWKPMVIVNRLKAVQEVLVTHGEDTADRPPVP IFKCLGVKPRSQGVVLASYGPE
WREQRRFSVSTLRFTFGMGKKSLEEWVTEAGHLCDRAFTAQAGQ SINPKAMLNKALCNVIASLIFA
RRFEYEDPYLIRMLKLV EESL TEVSGFIEVLNTFPALLRIPGLADKVFPQGKTFMALLDNLLAE
NRTTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRMVVVDLFTAGMVTATTLTWALLLMI
LYPDVQRRVQQEIDEVIGQVRCP EMTDQAHMPYTNAVIHEVQRFQDIAPLNLPRI TS CDIEVQDF
VIPKGTTLI INLSSVLKDETVWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSTHGFFAIPVAPLPYQLCAVVR

>N58

MELLNGTGLWPMAIFTVIFILLVLDLMHRHQRWTSRYPPGPVPWPVLGNLLQVDP SNMPYSMYKLQ
HRYGDVFSLQMGWKPMVIVNRLKAVQEVLVTHGEDTADRPPVP IFKCLGVKPRSQGVVFASYGPE
WREQRRFSVSTLRFTFGMGKKSLEEWVTEAGHLCDRAFTAQNGRSINPKAMLNKALCNVIASLIFA

RRFEYEDPYLIRMLTLVEESLIEVSGF IPEVLNTFPALLRIPGLADKVFQGGQKTFMAFLDNLLAE
NRTTWDPAPPPRNLDAFLAEVEKAKGNPESSFNENLRMVVDLFTAGMVTATTTLTWALLLMI
LYPDVQRRVQQEIDEVIGQVRCPMTDQAHMPYTNAVIHEVQRFQDIAPLNLPRITSCDIEVQDF
VIPKGTTLIINLSSVLKDETVWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQHFSSVPAGQPRPSTLGNFAISVAPLPYQLCAAVR

>N59

MELLTGTGLWPVAIFTVIFILLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMGWKPMVVINGLKAMKEVLLTCGEDTADRPPVP IFEYLGVKPGSQGVVLAPYGP
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEAGHLCDAFQAQAGQPINPNTMLNKAVCNVIASLIFA
RRFEYEDPYLIRMLKVLEESLIEVSGF IPEVLNTFPILLRIPGLADKVFQGGQKSFMALDNLLTE
NRTTWDPAPPPRNLDAFLAEIEKAKGNPESSFNENLRMVSDLFTAGMVTSTTLTALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIAPLNLPRITSRDIEVQDF
LIPKGTILIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVPDQGPQPSNHGVFAILVAPSPYQLCAVVR

>N60

MELLTGTGLWSVAIFTVIFILLVDLMHRHQWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMGWKPMVVINGLKAMKEVLLTCGEDTADRPPVP IFEYLGVKPGSQGVVLAPYGP
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEARHLCDAFQAQAGQPINPNTMLNNAVCNVIASLIFA
RRFEYEDPYLIRMLKVLEESLIEISGF IPEVLNTFPILLRIPGLADKVFQGGQKSFMALDNLLTE
NRTTWDPDQPPRNLDAFLAEIEKAKGNPESSFNENLRMVSDLFTAGMVTSTTLTALLLMI
LHPDVQRRVQQEIDAVIGQVRHPEMADQARMPYTNAVIHEVQRFQDIAPLNLPRITSRDIEVQDF
LIPKGTILIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVPDQGPQPSNHGVFAIPVAPYPYQLCAVVR

>N61

MELLTGTGLWSVAIFTVIFILLVDLMHRHQWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMGWKPMVVINGLKAMKEVLLTCGEDTADRPPVP IFEYLGVKPGSQGVVLAPYGP
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEARHLCDAFQAQAGQPINPNTMLNNAVCNVIASLIFA
RRFEYEDPYLIRMLKMLEESLIEISGFIPGVNTFPILLRIPGLADMVFGGQKSFMALDNLLTE
NRTTWDPDQPPRNLDAFLAEIEKAKGNPESSFNENLRMVSDLFTAGMVTSTTLTALLLMI
LHPDVQRRVQQEIDAVIGQVRHPEMADQARMPYTNAVIHEVQRFQDIAPLNLPRITSRDIEVQDF
LIPKGTILIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVPDQGPQPSNHRVFAIPVAPYPYQLCAVVR

>N62

MELLTGTDLWPVAIFTVIFILLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMAWKPMVVINGLKAMKEVLLTCGEDTADRPPVP IFEYLGVKPGSQGVVLAPYGP
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEAGHLCDAFQAQAGQPINPNTMLNKATCNVIASLIFA
RRFEYEDPYLIRMLKVLEESLIEVSGLIPEVLNTFPILLRIPRLADKVLQGGQKSFAILDNLLTE
NRTTWDPAPPPRNLDAFLAEIEKAKGNPESSFNENLRMVSDLFTAGMLTTSTTLTALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIVPLNLPRITSRDIEVQDF
LIPKGTILIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVPDQGPQPSNHGVFAILVAPSPYQLCAVVR

>N63

MELLTGTDLWPVAIFTVIFILLVDLTHQRQRWTSRYPPGPVPWPVLGNLLQVDLDNMPYSLYKLQ
NRYGDVFSLQMAWKPMVVINGLKAMKEMLTCGEDTADRPPVP IFEYLGVKPGSQGVVLAPYGP
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEANHLCDAFQAQAGQPINPNTMLNKSTCNVIASLIFA
RRFEYEDPFLIRMLKVLEQSLTEVSGLIPEVLNAFPILLRIPRLADKALQGGQKSFAILDNLLTE
NRTTWDPVQAPRNLDAFLAEIEKAKGNPESSFNENLMLVVRDLFGAGMLTTSTTLTALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIVPNLPRITSHDIEVQDF
LIPKGTILIPNMSSMLKDESVEKPLRFHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEALARME
LFLFFTCLLQRFSSVPDQGPQPSNSGVYGILVAPSPYQLCAVVR

>N64

MGLLLGSNLSALGVALTVFALLVDFMKRRKRWC SRYPPGPMSLPFVGTMLQVDYHNPHLSFNQLS
KKFGNVFSLQNCWNNVVVLNGFQAVREALVQKSEDFADRPPLP IYHLGYGRKSEGVLARYGHA
WKELRRFTLSTLRNFGMGKKSLEERVTEEAGFLCSAISSEGRPFDPHFLINNAVSNVICSI IYG
ERFDYDDKKFQKLLRLFEEESLNEESGFLPQFLNVVPWLLRIPGLPQKVFQGGQKALLAFTDELIKE
HKETWNPAHPTRDFTDAFLHEMEKAKENVESSFNENNLRLVTADLF SAGTETTSTTLRWALLFML
LYPDVQSKVHEEIDKVI GRDRSP TMGDQVNMPYTNAVIHEVQRYGDIVPVGVPHMTYRDTLQGF

FIPKGTTVITNLSSVLKDETVWEKPHQFYPEHFLDADGQFVKPEAFMPFSAGRRVCLGEQLARME
LFLFFTCLLQRFTEFCIPEGQPRPREDGHHAFILAPHPYQIRAVPR

>N65

MTLLLWSNISVLGVFLTVFTLLVDFMKRRKKWCSTRYPPGPMSLPFVGTMPIYVDYNNPHLSFNKFR
KKFGNIFSLQNCWTNVVVLNGYKTVKEALVHKSEDFADRPYLP IYEHLGYGHKSEGIVLARYGHV
WKELRKFTLSTLRNFGMGKKSLEERVTEEAGFLCSAISSEEGRPFDPHFLVNNAVCNVCISIIYG
ERFDYGDKTFKLLTLFENSLNEEAGFLPQFLNVVPILLRIPGLPQKVFRGQKAYIDFTDVLIEK
HKETWNPAHPTRDFTDAFLKEMAKGKEAEENGFNKNNLRLVTADLFAAGSETTSTTLRWAFLFML
LYPEIQSKVHKEIDKVI GRDRSPTMADQVNMPYTNNAVIHEVQRCGDIVPVGLPHMTYRDTELQGF
FIPKGTTVITNLSSVLKDETVWKKPNEFYPEHFLDANGQFVKPEAFMPFSAGRRVCLGEQLARME
LFIFFTLLQKFTFVIPEDQPRPREDGHHAFINSPHPYQLRAVPR

>N66

MTLLLWSNISVLGVFLTVFTILVDFMKRRKKWCSTRYPPGPMSLPFVGTMPIYVNNPHLSFEKFR
KKFGNIFSLQNCWTNVVVLNGYKTVKEALVNKSEDFADRPYMP IYEHLGYGHKSEGIVLARYGHV
WKELRKFTLTLTRNFGMGKKSLEERVTEEAGFLCSAISSEGGRPFDPHFLVNNAVCNVCITITYG
ERFDYGDKTFKLLTLFENSLNEEAGFLPQLLVNVPILLRIPGLPQKIFRCQKAYIDFTEMLIEK
HKETWNPAHPTRDFTDAFLKEMAKGKEAEENGFNKSNLSLVTADLLAAGSETTSTTLRWAFLFML
LYPEIQSKVHKEIDKVI GRNRPTMADQVNMPYTNNAVIHEVQRCGDIVPMGLPHMTYRDTELQGF
FIPKGTIIITNLSSVLKDETVWKKPNEFYPEHFLNENGQFVRPEAFLPFSAGRRACLGEQLARME
LFIFFTLLMQKFTFVIPEDQPRPREDGHHAFINSPHPYQLRAVPR

>N67

MSLLSGCNVFTLGIICTLCLLLLDYMKRRKPWCTNFPPSPPSRPFVGNLLQVDFRNLHNSFKQLS
KQYGDVMSLQFLWFKPMVVLNGFEVMKEALIQKSEDTADRPPFHIYEILGFGGNNKGVVLANYGQS
WKDLRRFTLSTLRDFGMGKKSLEERVREEAGYLCAAFQSEQGRPFDPHVLLNTAVSNVICSIIFG
ERFEYDDHKFQKLLCLIEESLKAESGAVPQILSSLPWSSKIPGLAKRFFQPRISMLKYLQEIINE
HQQTWDSGHPTRDFIDAFMLEMEKAKGVKDSNFNEENLQLTTADLFSAGTETTSTTLRWGLLFML
LYPDVQQRKVHEEIDQVI GRTRKPTMGDVLQMPYTNNAVIHEIQRYADIIPLSVPHMTYRDTHIKGF
FIPKGTTIMTNLSSVLKDEKVWEKPFQFYPEHFLDRDGKFKVREAFMAFSAGRRVCLGEQLARME
LFLFFTSLQRFQIPDGEPSPREDPVFVFLQVPHDYKICAKVR

>N68

MSLLSGCNVFTLGIICTLCLLLLDYMKRRKPWCTNFPPSPPSRPFVGNLLQMDFRNLHNSFKQLS
KQYGDVMSLQVFWKPMVVLNGFEVMKEALIQKSEDTADRPPFHIYEILGFVGNNAVVLANYGQS
WKDLRRFTLSTLRDFGMGKKSLEERVREEAGYLCAAFQSEQGRPFDPHVLLNTAVSNVICSIIFG
ERFEYDDHKFQKLLCLIEESVKAESGAVPQIISSLPWSSKIPGLAKLFFQPRIHMLKYLQEIINE
HQQTWDSGHPTRDFIDAFMLEMEKAKGVKDSNFNEQNLLLTADLFSAGSETTSTTLRWGLLFML
LYPDVQQRKVHEEIDQVI GRTRKPTMGDVLQMPYTNNAVIHEIQRYADIIPLSVPHMTYRDTHIKGF
FIPKGTTIMTNLSSVLKDEKVWEKPFQFYPEHFLDRDGKFKVREAFMAFSAGRRVCLGEQLARME
LFLFFTSLQRFQIPDGEPSPREDPVFVFLQVPHDYKICAKVR

>N69

MSLLSGCNVFTLGIIFTLLLLLLDFMKRRKPWCTDFPPSPPSWPFVGNLLQMDFRDLHNSFKQLS
KQYGDVMSLQVFWKPMVVLNGFEVKEALIQKSEDTADRPPFNLYEILGFVGNNAVVLANYGQS
WKDLRRFTLSTLRDFGMGKKSLEERVREDEAGYLCAAFQSEQGRPFDPHVLLINTAVSNVICSIIFG
ERFEYDDHKFQKLLCLIEESVKAESGPVPQIISSLPWSSKVPGLAKLFFQPRIHMLKYLQEIINE
HKQTWDSGHPTRDFIDAFMLEMEKAKGVKDSNFNDQNLLLTADLFSAGSETTTTTTLRWGLLFML
LYPDVQQRKVQEEIDQVI GRTRKPTMGDVLQMPYTNNAVIHEIQRYADIIPLSVPHMTYRDTHIKGF
FIPKGTIVIMTNLSSVLKDEKVWEKPFQFYPEHFLDRDGKFKVREAFMAFSAGRRVCLGEQLARME
LFLFFTSLQRFQIPDGEPSPREDPVFVFLQVPHDYKICAKVR

Figure 20

>N1

MGASD TESG PSLPEPAAEEEEPF L SMLDVFLSLILGLLVYWFFFRKKKEE IPEFKPIQP VAPT LR
ETS FIEKMKKTGKNIVVFYGSQTGTAE EFANRLAKDAQRYGMKGMAADPEE CDMEELSRLGEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQEDDADLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFDLGLGDDDDANIEEDFI TWRE RFWPAVCEHFGIEASGEDGSMRQYSLVEHTDINKNKIYT
GEMARLHSYENQKPPFDAKNPFLAPVTVNRELHKGGGRSCMHIELDITGSKIRYEAGDHVAVYPT
NDAALVEELGKLLNTNLDTVFSLNNVDEESSKKHPFPCPTTYRTAL THYLDITNPPRTNVLHELA
EYASDPKDKEFLRKMASSSPEGKKQYQSWVVESRRNILAILQDLPSCRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNKG VATNWLKNKVPATDNGHKPRVP IYVRKSQFRL
PFKPSTPVMIGPGTGLAPFIGFIQERAYLKEQGKEVGETVLYFGCRKKNE DFIYQEELEQYQKE
GVLTQLNMAFSRDQAQKVYVQHLLKQNKESVWKLIHEDNAHIYVCGDARNMARDVQNALYEIVAE
KGNMTQTQAVDYIKKLMTKGRYSCDVWS

>N2

MGASD TESG PSLPEPAAEEEEPF L SMLDVFLSLILGLLVYWFFFRKKKEE IPEFKPIQP VAPT LR
ETS FIEKMKKTGKNIVVFYGSQTGTAE EFANRLAKDAQRYGMKGMAADPEE CDMEELSRLGEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQEDDADLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFDLGLGDDDDANIEEDFI TWRE RFWPAVCEHFGIEASGEDGSMRQYSLVEHTDINKNKIYT
GEMARLHSYENQKPPFDAKNPFLAPVTVNRELHKGGGRSCMHIELDITGSKIRYEAGDHVAVYPT
NDAALVEELGKLLNTNLDTVFSLNNVDEESSKKHPFPCPTTYRTAL THYLDITNPPRTNVLHELA
EYASDPKDKEFLRKMASSSPEGKKQYQSWVVESRRNILAILQDLPSCRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNKG VATNWLKNKVPATDNGHKPRVP IYVRKSQFRL
PFKPSTPVMIGPGTGLAPFIGFIQERAYLKEQGKEVGETVLYFGCRKKNE DFIYQEELEQYQKE
GVLTQLNMAFSRDQAQKVYVQHLLKQNKESVWKLIHEDNAHIYVCGDARNMARDVQNALYEIVAE
KGNMTQTQAVDYIKKLMTKGRYSCDVWS

>N3

MGASD TESG PSLPEPAAEEESFFSMDVDFLFSLIAGLLVYWFFFRKKKEE IPEFKPIQP VAPT LR
ETS FIEKMKKTGKNIVVFYGSQTGTAE EFANRLSKDAQRYGMRGMAADPEE YDMSELSRLSEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDADLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFDLGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEDSSIRQYELVVHTDINMNKVYT
GEMGR LKSYEIQKPPFDAKNPFLAPVTVNRLKNKGGERHLMHLELDITGSKIRYESGDHVAVYPT
NDAALVNELGKILNTDLDTVFSLNNLDEESNKKHPFPCPTTYRTAL THYLDITNPPRTNVLYELA
QYATDPKDQEFRLRKMASSSPEGKALYLSWVVEARRNILAILQDLP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNKG VATNWLKNKVPATDNGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRHKNE DFLYQEELEQFQKE
GVLTQLNVAFSRDQAQKVYVQHLLKKNKEHVWKLIHEDNAHIYVCGDARNMARDVQNAFYEIVAE
FGNMTQTQAVDYIKKLMTKGRYSCDVWS

>N4

MGDS D TESDTSPPEPAAEEESFFSMDVDFLFSLIAGLLIYWFFFRKKKEE IPEFKKIQP VAPSVR
ETS FIEKMKKTGKNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEE YDMSELSRLSEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDADLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFDLGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESIRQYELVVHTDINMNKVYT
GEMGR LKSYETQKPPFDAKNPFLAPVTVNRLKNKGGERHLMHLELDITGSKIRYESGDHVAVYPT
NDAALVNKLGEILGADLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDITNPPRTNVLYELA
QYATDPKDQENLRKMASSSPEGKALYLSWVVEARRNILAILQDLP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNKG VATNWLKNKVPATDNGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRHKNE DFLYQEELEQFQKE
GVLTQLNVAFSRDQAQKVYVQHLLKKNKEHVWKLIHEDNAHIYVCGDARNMARDVQNAFYEIVAE
FGNMTQTQAVDYIKKLMTKGRYSQDVWS

>N5

MGDS S TESDTSPPE SAAEEESFFSMTDVF LFSLI VGLLYWFFFRKKKEE IPEFTKIQT TSSVR
ESS FIEKMKKTGKNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEE YDMSDL SRLSEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETD DVL SGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESIRQYELVVHTDINMNKVYT
GEMGR LKSYETQKPPFDAKNPFLAPVTVNRLNQGGERHLMHLELDITGSKIRYESGDHVAVYPA

NDTALVNKLGEILGADLDTVISLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPKEQENLRKMASSSPEGKALYLSWVVEARRNILAILQDVPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNGVATNWLKNKQPANDNGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRHENEEDFLYKEELEQFHKE
GVLTLQNLVAFSRDQAQKVYVQHLLKKNKEHVWKLIIHEDNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMNQTAQVDYIKKLMTKGRYSQDVWS

>N6

MGDSSTESDTSPPEASAAEEASLFSMTDVFLFSLIVGLLTYWFFFRKKKEEIEPFTTKIQPTTSSVR
ESSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSRLSEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHVAVYPA
NDTSLVNQLGEILGADLDTVISLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPNEQEHLRKMMASSSAEGKALYLSWVVEARRNILAILQDVPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNGVATNWLKNKQPANENGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRHENEEDFLYKEELEAQFHKE
GALTQNLVAFSRDQAQKVYVQHLLKKNKEHVWKLIIHEDNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMNQTAQVDYIKKLMTKGRYSQDVWS

>N7

MGDSSTSTSPPEASAAEEASLFSMTDVFLFSLIVGLLTYWFFFRKKKEEIEPFTTKIQPTTSSVR
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSRLSEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
NDTSLVNQLGEILGADLDTVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPNEQEHLRKMMASSSAEGKALYLSWVVEARRNILAILQDVPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRVNGVATNWLKNKQPANENGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRHENEEDFLYREELAQFHKE
GALTQNLVAFSRDQAQKVYVQHLLKKNKEHVWKLIIHEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMQTAQVDYIKKLMTKGRYSQDVWS

>N8

MGDSNM DATATVSEAAAEVSLFSMTDMFLFSLIVGLLTYWFFFRKKKEEIEPFTTKIQPTTSSVK
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNRLKNQGGGERHLMHLELDISDSKIRYESGDHVAVYPA
NDSSLVNQLGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATEPSEQEHLRKMMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRVNGVATNWLRAKEPANENGHKATVPMFVRKSQFRL
PFKPSTPVMIGPGTGIAPFMGFIQERAWLKQQGKEVGETVLYYGCRANEDFLYREELAQFHKD
GALTQNLVAFSREQPHKVYVQHLLKKNKEHLWKLIIHEGSAHIYVCGDARNMARDVQNTFYEIVAE
FGGMSQAQVDYIKKLMTKGRYSQDVWS

>N9

MGDSNM DATATVSEAAAEVSLFSMTDMI LFLSLIVGLLTYWFLFRKKKEEVPEFTTKIQTTTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNRLKNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATEPSEQEHLRKMMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKAGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVMIGPGTGIAPFMGFIQERAWLRQQGKEVGETLLYYGCRSDEDLYREELAQFHKD
GALTQNLVAFSREQPHKVYVQHLLKDKKEHLWKLIIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQVDYIKKLMTKGRYSQDVWS

>N10

MGDSNM DTSATVSEAAAEVSLFSMTDMI LFLSLIVGLLTYWFLFRKKKEEVPEFTTKIQTTTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL

GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N11

MGDSNMDSATVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N12

MGDSHMDTSATVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
ESSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVF CMATYGECDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N13

MGDSHMDTSSTVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
ESSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVF CMATYGECDPTDNAQDFYDWLQETDVDLSGVKFVAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N14

MGDSHVDTSSSTVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKEEVPEFTKIQT STSSVK
ESSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVF CMATYGECDPTDNAQDFYDWLQETDVDLSGVKFVAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDAAKVYM
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N15

MGDSHVDTSSSTVSEAVAAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K E E V P E F T K I Q T L T S S V R
E S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M S A D P E E Y D L A D L S S L P E I D N
A L V V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K F A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y M
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G K I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E L L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D C P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K A G R I N K G V A T N W L R A K E P A G E N G G R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P G T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H R D
G A L T Q L N V A F S R E Q S H K V Y V Q H L L K R D R E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
L G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N16

MGDSHVDTSSSTVSEAVAAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K E E V P E F T K I Q T L T S S V R
E S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M S A D P E E Y D L A D L S S L P E I D N
A L V V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K F A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y M
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G K I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E L L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D C P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K A G R I N K G V A T N W L R A K E P A G E N G G R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P G T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H R D
G A L T Q L N V A F S R E Q S H K V Y V Q H L L K R D R E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
L G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N17

MGDSHVDTSSSTVSEAVAAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K E E V P E F T K I Q T L T S S V R
E S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M S A D P E E Y D L A D L S S L P E I D N
A L V V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K F A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G K I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E L L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D C P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K A G R I N K G V A T N W L R A K E P A G E N G G R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P G T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H R D
G A L T Q L N V A F S R E Q S H K V Y V Q H L L K R D R E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
L G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N18

MGDSHVDTSSSTVSEAVAAEEVSLFSMTDVI LFS V I V G L L T Y W F L F R K K K E E V P E F T K I Q T L T S S V R
E S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M S A D P E E Y D L A D L S S L P E I D N
A L V V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K F A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G K I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E L L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D C P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K A G R I N K G V A T N W L R A K E P A G E N G G R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P G T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H R D
G A L T Q L N V A F S R E Q S H K V Y V Q H L L K R D R E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
L G A M E H A Q A V D Y V K K L M T K G R Y S L D V W S

>N19

MGDSHVDTSSSTVSEAVAAEEVSLFSMTDVI LFS V I V G L L T Y W F L F R K K K E E V P E F T K I Q T L T S S V R
E S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M S A D P E E Y D L A D L S S L P E I D N
A L V V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K F A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G K I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E L L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D C P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K A G R I N K G V A T N W L R A K E P A G E N G G R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P G T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H R D

GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N20

MGDSHVDTSSTVSEAAVEEVSLSFMTDVI LFSVIVGLLTYWFLFRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPF IGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N21

MGDSHVDTSSTVSEAAVEEVSLSFMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPF IGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N22

MGDSNVDTSSSTVSEAAAEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQTSTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLVVFCMATYGE GPD TDNAQDFYDWLQETDMDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHSDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPFMGFIQERAWLQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N23

MGDSHMDTSATVTEAAVEEVSLSFMTDMI LFSLIVGVLT YWFLFRKKKEEVPEFTKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGVAPF IGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N24

MGDSHMDTSATVTEAAVEEVSLSFMTDMI LFSLIVGVLT YWFLFRKKKEEVPEFTKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHTDIDPAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY

YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLRQQGKEVGETLLYYGCRRSEDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N25

MGDSHMDTGATNTEAVAAEEVSLFSMTDMI LFSLIVGVLTWFLFRKKKEEVPEFTKISTTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHTDIDPAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKDPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLRQQGKEVGETLLYYGCRRSEDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIAAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N26

MGDSHMDTGATNTEAVAAEEVSLFSMTDMI LFSLIMGVLTWFLFRKKKEEVPEFTKISTTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGECDPTDNAQDFYDWLQETDLDLDSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVLHTDIDPAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKDPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLRQQGKEVGETLLYYGCRRSEDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDQEH LWKLIHEDGAHIYVCGDARNMARDVQNTFCDIAAE
LGGMEHAQAVDYIKKLMTKGRYSLDVWS

>N27

MGDSHMDTGATNTEAVAAEEVSLFSMTDMI LFSLIMGVLTWFLFRKKKEEVPEFTKISTTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGECDPTDNAQDFYDWLQETDLDLDSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVLHADIDPAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKDPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLRQQGKEVGETLLYYGCRRSEDEDYLYREELAQFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDQEH LWKLIHEDGAHIYVCGDARNMARDVQNTFCDIAAE
LGGMEHAQAVDYVKKLMTKGRYSLDVWS

>N28

MGDSHKDTSATVTEAVAAEEVSLFSMTDMVLFSLIVGVLTWFLFRKKKEEVPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGECDPTDNAQDFYDWLQETDVLDTGVKF AVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHTDIDPAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYEA KSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGIAPFMGFIQERAWLRQ GKEVGETLLYYGCRRSEDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS

>N29

MGDSHKDTSATMTEAVAAEEVSLFSTTDMVLFSLIVGVLTWFLFRKKKEEVPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVFCMATYGECDPTDNAQDFYDWLQETDVLDTGVKF AVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDIDTAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTNNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA

NDSTLVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGTIAPFMGFIQERAWLREQGKEVGETLLYYGCRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS

>N30

MGDSHKDTSATMTEAVAAEEVSLFSTTDMVLFSLIVGVLTYYWFI FRKKKEEVPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVFCMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDIDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGTIAPFMGFIQERAWLREQGKEVGETLLYYGCRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS

>N31

MGDSHKDTSATMTEAVAAEEVSLFSTTDMVLFSLIVGVLTYYWFI FRKKKEEVPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVFCMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGTIAPFMGFIQERAWLREQGKEVGETLLYYGCRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS

>N32

MGDSHEDTSATMPEAVAAEEVSLFSTTDMVLFSLIVGVLTYYWFI FRKKKEEIPEFSKIQT TAPPVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVFCMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVIMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGTIAPFMGFIQERAWLREQGKEVGETLLYYGCRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHTQAVDYVKKLMTKGRYSLDVWS

>N33

MGDSNM DTSATVSETVAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQAMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETD VDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGTGVAFFIGFIQERAWLRQOGKEVGETLLYYGCRSDEDYLYREELA QFHKD
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N34

MGDSNM DTSATVSETVAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETD VDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL

GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAAVTNTRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N35

MGDSNMDSATVSETVAEEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAAVTNTRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGSMEHAQAVDYIKKLMTKGRYSLDVWS

>N36

MGDSNIDTSATGSETVAEEVSLFSMTDMI LFS LIVGLLTYWFFFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTNTRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS

>N37

MGDSNTDTSATGSETVAEEVSLFSMTDMI LFS LIVGLLTYWFFFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTNTRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS

>N38

MGDSNTDTSATGSETVAEEVSLFSMTDMI LFS LIVGLLTYWFFFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPACVEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTNTRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS

>N39

MGDSNMDSATVSETVAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K D E I P E F T K I Q P M T S S V K
D S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L G S L P E I E N
S L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G M G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D M A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K L R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E H L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K S G R I N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H K D
G S L T Q L N V A F S R E Q P H K V Y V Q H L L K R D K E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
V G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N40

MGDSNMDSATVSETVAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K D E I P E F T K I Q P V T S S V K
D S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L G S L P E I E N
S L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G M G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D M A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K L R Y E S G D H V A V Y P A
N D S A L V S Q L G E I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E H L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K S G R I N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H K D
G S L T Q L N V A F S R E Q P H K V Y V Q H L L K R D K E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
V G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N41

MGDSNVDSATVSETVAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K D E I P E F T R I Q P V T S S V K
D S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L G S L P E I E N
S L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D M D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G M G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D V A K V Y M
G E M G R L K S Y E N Q K P P F D A K N P F L A E V T T N R K L N Q G T E R H L M H L E L D I S D S K L R Y E S G D H V A V Y P A
N D S A L V S Q L G K I L G A D L D V I M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E H L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K S G R I N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K D V G E T L L Y Y G C R R S D E D Y L Y R E E L T Q F H K D
G S L T Q L N V A F S R E Q P H K V Y V Q H L L K R D K E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
V G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N42

MGDSSMDASATVSETVAEEVSLFSMTDMI LFS L I V G L L T Y W F L F R K K K D E I P E F T K I Q P V T S S V K
D S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L G S L P E I E N
S L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G M G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D I D P A K V Y V
G E M G R L K S Y E N Q K P P F D A K N P F L A A V T T N R K L N Q G T E R H L M H L E L D I S D S K L R Y E S G D H V A V Y P A
N D S T L V S Q L G E I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E Q L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K S G R V N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A G F H K D
G S L T Q L N V A F S R E Q P H K V Y V Q H L L K R D K E H L W Q L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
G G A M E H A Q A V D Y I K K L M T K G R Y S L D V W S

>N43

MGDSHMDASATVSETVAEEVSLLSMTDMI LFS L I V G L L T Y W F F R K K K D E V P E F T K I Q P V T S S V K
D S S F V E K M K K T G R N I I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L G S L P D I E N
S L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D L D L T G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F E L G M G D D D G N L E E D F L T W R E Q F W P A V C E H F G V E A T G E E C S I R Q Y A L V V H P D I D P A K V Y V
G E V G R L K S Y E N Q K P P F D A K N P F L A A V T A N R K L N Q G T E R H F M H L E L D I S D S K L R Y E S G D H V A I Y P A
N D S T L V S Q L G E I L G A D L D V V M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q H A S E P S E Q E Q L R R M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K S G R V N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I L V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A G F H Q D

GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAD
GGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N44

MGDSSMDASATVSETVAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEIPEFTKIQPVTS SVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDPAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKLRYESGDHVAVYPA
NDSTLVSQ LGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAGFHKD
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
GGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N45

MGDSNMDSATVSEAVAEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTS SVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N46

MGDSNMDSATVSEAVAEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEVPEFTKIQT TTS SVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAGVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N47

MGDSNMDSATVSEAVTEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEVPEFTKIQT TTS SVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAGVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGHRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWQLIHEAGAHYVCGDARNMARDVQNTFYDIVAE
FGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N48

MGDSSMDTSTTVSEAVTEEMSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEVPEFTKIQT TTS SVK
DSSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLTEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAGVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCCELLPRLQARY

YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGHRALVPMFVRKSQFRL
PFKATTPVIMVGPGTGVPFFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEAGAHYVCGDARNMARDVQNTFYDIVAE
FGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N49

MGDSADASTTVSEAVTEEMSLFSMTDMI LFS LIVGLLTYWFLFRKKKDEIPEFTKIQTMTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLTEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQETDVDLSGVKFVAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAGVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGHRALVPMFVRKSQFRL
PFKATTPVIMVGPGTGVPFFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEAGAHYVCGDARNMARDVQNTFYDIVAE
FGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N50

MGDSNMDTGATMSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLFRKKKEEVPEFTKIQTMTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGPGTGVPFFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIIHEGGAHYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N51

MGDSNVDTGATMSEAVAEVSLFSMTDMVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQTMTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDTAKVYT
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGPGTGVPFFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIIHEGGAHYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N52

MGDSNVDTSATMSEAVAEVSLFSMTDMVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQTMTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDTAKVYT
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVGPGTGVPFFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEGGAHYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N53

MGDSNVDTSATMSEAVAEVSLFSMTDMVLFSLIVGLLTYWFLFRKKKDEVPEFTKIQTMTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDTAKVYT
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA

NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N54

MGDSNMDSATMSETVAEEVSLFSMTDMLLFSLIVGLLTYWFLFRKKKDEVPEFTKIQTSTTSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GSLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N55

MGDSNVDTSATMSEAVTEEMSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEVPAFTKIQPTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLAEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GARRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDIAKVYT
GEMGR LKSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPVEENGHRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQGKDVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N56

MGDSHGDTSATMPEAAVEEAAAFSMTDAVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQSTTASVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSEIDN
ALVVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDPAKVYH
GEMGR LKSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLSVAFSREQAHKVYVQHLLRKDKEHLWKLIIHEGGAHIYVCGDARNMAKDVQNTFYEIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N57

MGDSHGDTSATMPEAAAEAAAFSMTDAVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQSTTASVK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSEIDN
ALVVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDPAKVYQ
GEMGR LKSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGALLGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAI EYETKTGRVNKG VATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQQGKEVGETLLYYGCRSEEDYLYREELAQFHKD
GALTQLSVAFSREQAHKVYVQHLLRKDKEHLWKLIIHEGGAHIYVCGDARNMAKDVQNTFYEIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N58

MGDSNVDTGATMSEAAVEEVS LFSMTDMVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQTSTTSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL

GAQRIFELGMGDDDDGNLEEDFI TWREQFWP AVCEHFGVEATGEESSIRQYELVVHTDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N59

MGDSNVDTGTTTSEMVAEEVSLFSATDMVLFSLIVGLLTYWFI FRKKKDEVPEFTKIE TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSSLPEIEN
ALAVFCMATYGE GPD TDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRKLEQL
GAQRIFDLGLGDDDDGNLEEDFI TWREQFWP AVCEHFGVEATGEESSIRQYELVVHTDMDTAVVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGTDLDIVMSLNNLDEESNKRHFPFCPTTYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLQE QGKEVGETLLYYGCRSDEDYLYREELAQFHAK
GALTRLSVAFSREQPQKVYVQHLLKRDKEHLWKLIHDGGAHIYICGDARNMARDVQNTFCDIVAE
QGPMEHAQAVDYVKKLMTKGRYSLDVWS

>N60

MGDSSVDTGATMSEAAAEVSLFSMTDMVLFSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLAVFCMATYGE GPD TDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWP AVCEHFGVEATGEESSIRQYELVVHTDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGSERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYAAEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKAGRINKGVATSWLRTKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N61

MGDSNMDTGATMSEAVAAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRKLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWP AVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N62

MGDSNMDTGATMSEAVAAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRKLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWP AVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N63

MGDSNMDTGTTTSEAVAEVSLFSTTDMI L FSLIVGVMTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DRSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEK
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFALGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPAEQEQLRKMASSSGEGKELYLRWVLEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAGFHKD
GALTQLNVAFSREQPQKVYVQHLLKKDKHEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N64

MADSNMDAGTTTSEMVAEEVSLFSTTDVIL FSLIVGVMTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DRSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEK
ALAIFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFALGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELMVHTDMDMAKVYT
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDIIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPAEHEQLRKMASSSGEGKELYLRWVLEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATSWLRAKEPAGENGGRALVPMYVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAGFHKD
GALTQLNVAFSREQPQKVYVQHLLKKDKHEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N65

MGDSNMDTGATMSEAVAEVSLFSMTDMI L FSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYIKKLMTKGRYSLDVWS

>N66

MGDSNVDAAGATASETVAEVSLFSTTDMI L FSLIVGLLTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAKFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N67

MGDSNVDAAGATAAETVTEEVS L FSTMDMI L FSLIVGLLTYWFFRKKKEEIP EFTKIQPM TSSVK
DSSFVEKMKKTGRNIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQR IYELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHPDIEAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLVELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQD SP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGIAPFMGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAKFHKD

GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N68

MGDSSVEAGATASDTVAEEVSLFSTTDMIMFSLIVGLLTYWFLFRKKKDEVPEFTKIQTSSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEDSNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGRALVPMFVRKSQFRL
PFKATVPVIMVPGTGVAPFIGFVQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAKFHQD
GALTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N69

MGDSSVEAGATASDTVAEEVSLFSTTDMIMFSLIVGLLTYWFLFRKKKDEVPEFTKIQTSSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEDSNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGRALVPMFVRKSQFRL
PFKATVPVIMVPGTGVAPFIGFVQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAKFHQD
GALTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N70

MGDSSVEAGATASDTVAEEVSLFSTTDMIMFSLIVGLLTYWFLFRKKKDEVPEFTKIQTSSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHFMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEDSNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGRALVPMFVRKSQFRL
PFKATVPVIMVPGTGVAPFIGFVQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAKLHQD
GALTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N71

MGDSSVEAGATASDTVAEEVSLFSTTDMIMFSLIVGLLTYWFLFRKKKDEVPEFTKIQTSSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHFMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEDSNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGRALVPMFVRKSQFRL
PFKATVPVIMVPGTGVAPFIGFVQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAKLHQD
GALTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>N72

MADVDTSATATVSEAGAEDVSLFSMTDMFLFSLIVGLLTYWFFFRKKKEEIPFTTKIQPVSSNVK
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGEDPTDNAQDFYDWLQETDADLSGVKFVAVFGLGNKTYEHFNAMGKYVDNRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDENMNKVYT
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGERHLMHLELIDISNKKIRYESGDHVAVYPA
NDASLVNQLGEILGADLDVVMVSLNNLDEESNKKHFPFCPTTYRTALTYLDITNPPRTNVLYELA
QYAGEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYPSLRPPIDHLCCELLPRLQARY

YSIASSSSKVHPNSVHICAVVVEYETKAGRVNKG VATNWLRAKEPAQENGHKATVPMFVRKSQFRL
PFKPTTPVIMIGPGTGIAPFMGFIQERAWLKEQGKEVGETVLYYGCRANEDYLYREELAQFHKD
GALTQLNVAFSREQPHKVYVQHLMKQNKHEHLWKLIHEGSAHIYVCGDARNMARDVQNVFYEIVSE
FGGMSQAQAVDFIKKLMAKGRYSLDVWS

>N73

MGDSSMDSTISPPESAAQEASLFSMTDVFLFSLIVGLLTYWFFFRKKKEEIPFTTKIQPTTSSVR
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSRLSEIDN
SLAVFCMATYEGEDPTDNAQDFYDWLQETDVDLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
NDTSLVNQIGEILGADLDTIMSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPNEQEHLRKMMASSAAEGKALYLSWVVEARRNILAILQDVP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHIICAVVVEYETKTGRNLNKG VATNWLKNKVPANENGHKSTVPMYVRKSQFRL
PFKPTTPVIMIGPGTGIAPFIFGFIQERAWLKQQGKEVGETVLYYGCRHENEDYLYREELAQFHKE
GALTQLNVAFSRDQAQKVYVQHLLKKNKEGVWKLIIHEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N74

MGDSAMDSTLPPPEASVPEGALFSMTDVFLFSLIVGLLSYWFLFRKKKEEIPFTTKIQPTTSSVR
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSRLSEIDN
SLAVFCMATYEGEDPTDNAQDFYDWLQETDVDLTGLKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYESQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
NDTSLVNQIGEILGADLDTIMSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPNEQEHLRKMMASSAAEGKTLYLSWVVEARRNILAILQDVP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHIICAVVVEYETKTGRNLNKG VATTWLKNKVPANENGHKLTVPMYVRKSQFRL
PFKPTTPVIMIGPGTGIAPFVGFVQERAWLKQQGKEVGETVLYYGCRHENEDYLYREELAQFHKE
GALTQLNVAFSRDQAQKVYVQHLLKKNKEGVWKLIIHEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGSMTQPQAVDYVKKLMTKGRYSLDVWS

>N75

MGDSAMDSTLPPPEASVPEGALFSMTDVFLFSLIVGLLSYWFLFRKKKEEIPFTTKIQPTTSSVR
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLSDLSRLSEIDN
SLAVFCMATYEGEDPTDNAQDFYDWLQETD TDLTGLKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGRKLSYESQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
NDTSLVNQIGEILGADLDTIMSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELA
QYATDPNEQEHLRKMMASSAAEGKTLYLSWVVEARRNILAILQDVP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHIICAVVVEYETKTGRNLNKG VATTWLKNKVPANENGHKLTVPMYVRKSQFRL
PFKPTTPVIMIGPGTGIAPFVGFVQERAWLKQQGKEVGETVLYYGCRHENEDYLYREELAQFHKE
GALTQLNVAFSRDQAQKVYVQHLLKKNKEGVWKLIIHEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGSMTQPQAVDYVKKLMTKGRYSLDVWS

>N76

MGDSGMDPTISPPESAAQEASLFSMTDVFLISLIVGLFTYWFFFRKKKEEVPEFTTKIQPTTSSVR
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLSDLSRLSEIDN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAPVTENRLKNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
NDSSLVNQIGEILGADLDTIMSLNNLDEESNKKHPFPCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDANEQERLRKMMASSAAEGKALYLSWVVEARRNILAILQDMP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHIICAVTVEYETKTGRNLNKG VATNWLKNKVPANENGKSLVPMYVRKSQFRL
PFKPTTPVIMIGPGTGIAPFIFGFIQERAWLKQQGKEVGETVLYYGCRHEKEDYLYREELAQFHKE
GALTQLNVAFSRDQAQKVYVQHLLKKNKEGVWKLIIHEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N77

MGDSGMDPTVSPPESTAQEASFLSMTDMFLISLIIGLFSYWFFFRKKKEEVPEFTTKIQQTSSGG
DSNFVDKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGLGADPEEYDLSDLSRLSEIDN
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGDESSIRQYELVVHTDMNMNKVYT
GEMGRKLSYESQKPPFDAKNPFLAPVTKNRKNQGGGERHLMHLELDISGSKVRYESGDHVAVYPA

NDSTLVNQIGEILGADLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDANEQDRLRKMASSAAEGKEFYLSWVVDSSRNILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVTVEYETKTGRNLKGVATTWLKNKVPANENGKNSLVP MYVRKSQFRL
PFKPSTPVMIGPGTGIAPFVGFIQERAWLKQQGKEVGD TVLYYGCRHEKEDYLYQEELTQFHKE
GALTQLNVAFSRDQAQKVYVQHLLKQNKKEGVWKL IHEGNAHIYVCGDARNMARDVQNTFYEIVAE
YGNMSQPQAVDYVKKLMTKGRYSMDVWS

>N78

MGDAGMEPNISPPDSAAQQDSLFSMTDVF LISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYEGEDPTDNAQDFYDWLQEADVDLSGLRF AVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMNMNKVYT
GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARRN ILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVTVEYETKTGRNLKGVATNWLKNKVPANENGNNSLVP MYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFHQE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N79

MGDAGMEPNISPPDSPAQQDSLFSMTDVF LISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYEGEDPTDNAQDFYDWLQEADADLSGLRF AVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARRN ILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVTVEYETKTGRNLKGVATNWLKNKVPANENGNSLVP MYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFHQE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N80

MGDAGMEPNISPPDSPAQQDSLFSMTDVF LISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYEGEDPTDNAQDFYDWLQEADADLSGLRF AVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARRN ILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNAIHICAVTVEYETKTGRNLKGVATNWLKNKVPANENGNSLVP MYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFHQE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N81

MGDAGMEPNISPPDSPAQQDSLFSMTDVF LISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYEGEDPTDNAQDFYDWLQEADADLSGLRF AVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARRN ILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNAIHICAVTVEYETKTGRNLKGVATNWLKNKVPANENGNSLVP MYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFHQE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N82

MGGAGMEQNI SPANSPAQQDSLFSMMDVF LISLITGLFTYWFFFRKKKEEVPDLPKMQSVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYEGEDPTDNAQDFYDWLQEADADLSGLRF AVFGLGNKTYEHFNAMGKYVDKRLEEL

GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALMYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNAIHC AVTVEYETKTGRNLKGVATNWLKNKVPANENGSSSLVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCGREREDYLYREELARFHQE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGDARNMARDVQNTFYEIVAE
FGNMGQPQAVDYVKKLMTKGRYSLDVWS

>N83

MGEPSMDPTISAADSPAQQDSLFSMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYGECDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSIHC AVTVEYETKTGRNLKGVATNWLKSKVPANQNGSSSLVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFQKE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLVNEGMAHIYVCGDARNMARDVQNTFYEIVSE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N84

MGEPSMDPTISAADSPAQQDSLFSMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYGECDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSIHC AVTVEYETKTGRNLKGVATNWLKSKVPANQNGSSSLVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFQKE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLVNEGMAHIYVCGDARNMARDVQNTFYEIVSE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N85

MGEPSMDPTISAADSPAQQDSLFSMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYGECDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSIHC AVTVEYETKTGRNLKGVATNWLKSKVPANQNGSSSLVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFQKE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLVNEGMAHIYVCGDARNMARDVQNTFYEIVSE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N86

MGEPSMDPTISAADSPAQQDSLFSMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYGECDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDRLEEL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASDAGEQERLRKMASSAAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSIHC AVTVEYETKTGRNLKGVATSWLRSKVPAGQNGSSSLVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRREREDYLYREELARFQKE
GVLTQLNVAFSRDQAEKVYVQHLLKKNKENVWKLNEGMAHIYVCGDARNMARDVQNTFYEIVSE
FGNMSQPQAVDYVKKLMTKGRYSLDVWS

>N87

MGSASSKDDSSPVKSAPEASVFPFSMTDVF LFSIIIGLLTYWFFFRKKKEEIPFTTKIQTPTQSAN
ESSFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGMRGMSADPEEYDLSDSLSEIDK
SLAVFCMATYGECDPTDNAQDFYDWLQETDADLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHTDINANKVYT
GEMGRKLSYENQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHVAVYPA
NDTSLVNQLGEILGADLDTVISLNNLDEESNKKHPFPCTTYRTALTYYLDITNPPRTNVLYELA
QYATDASEQEHLRKMMASSAEGKALYLSWVVEARRNILAILQDTPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYTTKTGRVNKG VATNWLKNKQPANENGHKSTVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQERGWLKQQGKEVGETVLYYGCRHENEDYLYKEELAKFLKE
GALTQLNVAFSRDQAQKVYVQHLLKKNKEHVWKL I HEDNGHIYVCGDARNMARDVQNTFYEIVAE
FGKMNQPPQAVDYIKKLMTKGRYSLDVWS

>N88

MGTASSKEESSSVKSAPEASVAFSMDVFLFSIIVIGLLTYWFFFRKKKEEIPFSKIQTPTQSAN
ENFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGMRGMSADPEEYDVSDLSHLSEIDK
SLAIFCMATYGECDPTDNAQDFYDWLQD TDANLSGIKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHKDISPSKVYM
GEMGRKLSYENQKPPFDAKNPFLAPVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHIAVYPA
NDTSLVNQLGEILSADLDTVISLNNLDEESNKKHPFPCTSYRTALTYYLDITNPPRTNVLYELA
QYATDASEQEQLRKMMASSAEGKTLYLSWVVEARRNILAILQDTPSLRPPIDHLCCELLPRLQARY
YSIASTSKVHPNSIHICAVVVEYTTKTDNRVNKG VATNWLKNKQPANENGHKSTVPLYVRKSQFRL
PFKPNT PVMIGPGTGIAPFIGFIQERGWLKQQGKEVGETVLYYGCRHEKEDIYIYKEELAKFVKE
GALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKL I HEENGHIYVCGDARNMARDVQNTFYEIVAE
YGKMNQQQAVDYVKKLMTKGRYSLDVWS

>N89

MGTASSKEEPSVKSAPASVAFSMDVFLFSIIVIGLLTYWFFFRKKKEEIPFSKIQTPTQSAN
ENFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGMRGMSADPEEYDVSDLTHLSEIDK
SLAIFCMATYGECDPTDNAQDFYDWLQD TDAKLSGIKYAVFGLGNKTYEHFNAMGKFVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHKDIAPAKVYM
GEMGRKLSYENQKPPFDAKNPYLAPVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHIAVYPA
NDTSLVNQLGEILSADLDTVISLNNLDEESNKKHPFPCTSYRTALTYYLDITNPPRTNVLYELA
QYATNASEQEQLRKMMASSAEGKTLYLSWVVESRRNILAILQDTPSLRPPIDHLCCELLPRLQARY
YSIASTSKVHPNSIHICAVVVEYTTKTDNRVNKG VATNWLKNKQPANENGHKSTVPLYVRKSQFRL
PFKPNT PVMIGPGTGIAPFIGFIQERGWLKQQGKEVGETVLYFGCRHEKEDIYIYKEELAKFVKE
GALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKL I NEENGHIYVCGDARNMARDVQNTFYEIVAE
YGKMNQQQAVDYVKKLMTKGRYSLDVWS

>N90

MGTATSKEEPSVKSAPASVAFSMDVFLFSIIVIGLLTYWFFFRKKKEEIPFSKIQTPTQSAN
ENFIEKMKKTGKNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDVSDLTHLSEIEK
SLAIFCMATYGECDPTDNAQDFYDWLQD TDAKLSGIKYAVFGLGNKTYEHFNAMGKFVDKRLEEL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHKDIAPAKVYM
GEMGRKLSYENQKPPFDAKNPYLAPVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHIAVYPA
NDTSLVNQLGEILSADLDTVISLNNLDEESNKKHPFPCTSYRTALTYYLDITNPPRTNVLYELA
QYATNASEQEQLRKMMASSAEGKTLYLSWVVESRRNILAILQDTPSLRPPIDHLCCELLPRLQARY
YSIASTSKVHPNSIHICSVVVEYTTKTDNRVNKG VATNWLKNKQPANENGHKSTVPLYVRKSQFRL
PFKPNT PVL MIGPGTGIAPFIGFIQERGWLKQQGKEVGETVLYFGCRHEKEDIYIYKEELAKFVKE
GALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKL I NEENGHIYVCGDARNMARDVQNTFYEIVAE
YGKMNQQQAVDYVKKLMTKGRYSLDVWS

>N91

MGESCTEQDTSEQGNSSPEEAFFSMADMFLLSLIVGLLTYWFFFRKKKEEIPFTTKIQTPTSSSVR
ESSFIEKMKKTGKNIVVFYGSQTGTGEEFANRLAKDAHRYGMRGMAADPEEFEMADLSRLTEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDIDLTGLKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAERIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYELVVHTDENMNVYT
GEMGRKLSYETQKPPFDAKNPFLANVTNRLKNQGGGERHLMHLELDITGSKIRYESGDHVAVYPA
NDTALVNKLGEILGADLDTVISLNNLDEESNKKHPFPCTTYRTALTYYLDITNPPRTNVLYELA
QYATDSKEQENLRKMMASSAQDGKALYLSWVVESRRNILAILEDVPSLRPPLDHLCCELLPRLQARY
YSIASSSKVHPNSIHVCAVLVEYETKTGRENKG VATNWLKNKQPASDNHGKSSVPMYVRKSQFRL
PFKPSTPVMIGPGTGIAPFIGFIQEREWLKQQGKDVGGETVLYYGCRHEHEDFLYKDELKRYHKE

GVLTQLNVAFSRDQAHKVYVQHLLKNNKEMVWKL I HEDNAHIYVCGDARNMARDVQNTFYDIVAE
YGKIDHAQAVDYIKKLMTKGRYSQDVWS

>N92

MGDMDESSTQPPPEVAEEEEFFSNVDLFLFSLIAGLLIYWFFFRKKKEE I PEFKKIQTVPAPSTR
ETSFI EKMKKTGKNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDMS ELSRLSEIDN
SLAIFCMATY GEGDPTDNAQDFYDWLQETDADLSGVNYTVFGLGNKTYEHYNAMGKYVDKRLEEL
GAQRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
GEMGR LKSFETQKPPFDAKNPFLAPVTVNRKLNKGGDRHLMHLELDITGSKIRYESGDHVAVYPT
NDAAI VNKLG EILGVLDLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDITNPPRTNVL YELA
QYATDPKDQENLRKMASSPEGKALYQSWVLDARRNILAILEDLPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYETKTGRV NKG VATNWLKNKV VATDNGHKSTVPMYVRKSQFRL
PFKASNPVIMIGPGTGIA PFI GF IQERAWLKEQGKEVGETVLYY GCRHKNE DFLYQEELEEF EKE
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKL I HTD NAHIYICGDARNMARDVQNAFYEIVAE
LGSMTHTQAVDYIKKLMTKGRYSQDVWS

>N93

MADMDTESSTQPPPEVEEEEEFFSNLDLFLFSLIAGLLIYWFMRKKAEP I PEFKKLETPAPSTR
ETSFI EKMKKTGKNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMS ELSRLSEIDN
SLAIFCMATY GEGDPTDNAQDFYDWLQETDADLDGVNYTVFGLGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEATGEESSIRQYELKVHTDINMNKVYT
GEMGR LKSFETQKPPFDAKNPFLAPVTVNRKLNKGGDRHLMHLELDITGSKIRYESGDHVAVYPT
NDSAI VNKLG EILGVLDLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDITNPPRTNVL YELA
QYASDPKDQENMRKMASSPEGKALYQSWVLDASRNIL AILEDLPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRV NKG VATNWLKNKV VATDNGHKSTVPMYIRKSQFRL
PFKASNPVIMIGPGTGIA PFMGFI QERAWLKEQGKEVGETVLYY GCRHKNE DFLYQEELEEF EKE
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKL I HTD NAHIYICGDARNMARDVQTA FYEIAEE
LGGMTHTQAVDYIKKLMTKGRYSQDVWS

>N94

MADMDTESSTQPPPEVEEEEEFFSNLDLFLFSLIAGLLIYWFMRKKAEP I PEFKKLETPAPSTR
ETSFI EKMKKTGKNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMS ELSRLSEIDN
SLAIFCMATY GEGDPTDNAQDFYDWLQETDADLDGVNYTVFGLGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEATGEESSIRQYELKVHTDINMNKVYT
GEMGR LKSFETQKPPFDAKNPFLAPVTVNRKLNKGGDRHLMHLELDITGSKIRYESGDHVAVYPT
NDSAI VNKLG EILGVLDLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDITNPPRTNVL YELA
QYASDPKDQENMRKMASSPEGKALYQSWVLDASRNIL AILEDLPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRV NKG VATNWLKNKV VATDNGHKSTVPMYIRKSQFRL
PFKASNPVIMIGPGTGIA PFMGFI QERAWLKEQGKEVGETVLYY GCRHKNE DFLYQEELEEF EKE
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKL I HTD NAHIYICGDARNMARDVQTA FYEIAEE
LGGMTHTQAVDYIKKLMTKGRYSQDVWS

>N95

MADMDTESSTQPPPEVEEEEEPLFSNLDLFLFSLIAGLLIYWFMSRKKAEPI PEFKKLDTPAPSTR
ETSFI EKMKKTGKNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMS ELSRLSEIDN
SLAIFCMATY GEGDPTDNAQDFYDWLQETDAELDGVNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGEESSIRQYELKVHTDINMNKVYT
GEMGR LKSFETQKPPFDAKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVYPT
NDSAI VNKLG QILGVLDLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDITHPRTNVL YELA
QYASDPKDQENMRKMASSPEGKALYQSWVLDASRNIL AILEDLPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRV NKG VATNWLKNKL VATDNGHKSTVPMYIRKSQFRL
PFKASNPVIMIGPGTGIA PFMGFI QERGWLKEQGKEVGETVLYY GCRHKNE DYLYQEELEEEAEKN
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKL I HTD NAHIYICGDARNMARDVQTA FYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N96

MADVEAESSTQPPPEAMEEEEEPLFSNLDLFLFTLIAGLLIYWFMSRKKTEPI PEFKKLDQPAPSTR
ETSFI EKMKKTGRNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMS ELSRLAEIDN
SLAIFCMATY GEGDPTDNAQDFYDWLQETDGELDGVNYTVFALGNKTYEHYNAMGAYVDKRLEEL
GAKRVFDLGMGDDDDGNLEEDFVTWREQFWPAMCEHFGVEASGEDSSIRQYELKEHNDINMNKVYT
GEMGR LKSFETQKPPFDAKNPFLAPVTVNRKLNKAGDRHLMHLEVDITGSKIRYESGDHVAVYPT
NDTAI VNR LGQILGVLDLDTVISLNNLDEESNKKHPFPCPTTYRTAL THYLDI IHPRTNVL YELA
QYATDPKDQENMRKMASSAPEGKALYQSFVLEDNRNILAILEDLPSLRPPIDHLC ELM PRLQARY

YSIASSSSKVHPNSIHICAVVVEYTTKTGRLTGKVATTWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKASNPVIMVPGGTGIAPFMGFIQERGWLKEQGKEVGETVLYFGCRHKKNEDYLYQEELEEAECT
GVITKLNVAFSRDQEQKVYVQHLLRTNKEDLWRQIHTDNAHIYICGDARNMARDVQTAFYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N97

MADVEAESSTQPPEAMEEENPLFSNLDLFLFTLIAGLIYWFMSRKKKEPIPEFKKLDQPAPSTR
ETSFIEKMKKTGRNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMSELSRLAEIGN
SLAIFCMATYEGEDPTDNAQDFYDWLQETDGELDGVNYTVFALGNKTYEHYNAMGAYVDKRLEEL
GAKRVFDLGMGDDDDGNLEEDFVWREQFWPAMCEHFGVEASGEDSSIRQYELKEHNDINMNKVYT
GELGRLKSFETQKPPFDAKNPFLAPVTVNRKLNKAGDRHLMHLEVDITGSKIRYESGDHVAVYPT
NDTAIVNRLGQILGVDLDSVISLNNLDEESNKKHFPFCPTTYRTALTHYLDIHPPTNVLYELA
QYATDPKDQENMRKMASSAPEGKALYQSFVLEDNRNILAILEDLP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHICAVLVEYTTKTGRLTGKVATTWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKASNPVIMVPGGTGIAPFMGFIQERGWLKEQGKEVGETVLYCGCRHKKNEDYLYQEELEEAECM
GVITKLNVAFSRDQEQKVYVQHLLRTNKEDLWRQIHTDNAHIYICGDARNMARDVQTAFYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N98

MADMDTETTTTTPPHPVEEEEEPLFSNLDLFLSLIVGLVIYWFMSRKKKEPIPEFKKLDTPPTSTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLSEIEN
SLAIFCMATYEGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDINMNKVYT
GEIGRLKSFVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQILGVDLDVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENMRKMASSSPEGKALYQSWVLDASRNILAILEDMP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSIHICAVVVEYKTKTGRVNKG VATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKATNPVIMIGPGGTGIAPFMGFIQERGWLKEQGKEVGETVLYFGCRHKKNEDYLYQEELEEAECN
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKL IHTDNAHIYICGDARNMAKDVQTAFYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N99

MPDMDTETHTTEPAPVDEEASLFSMDLLLSVIVGFVVYWFMSRKKKEPIPEFKKLDTPPTTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMVADPEEYSMGELSRLSEIKN
SLAIFCMATYEGEDPTDNAQDFYDWLQENDDELSGLNYTVFGLGNKTYEHYNAMGKYVDKRLEEL
GATRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKEHTDINMNKVYT
GEIGRLKSFVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQILGVDLDVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENMRKMASSSPEGKALYQSWVLDASRNILAILEDMP SLRPPVDHLCCELLPRLQARY
YSIASSSSKVHPNSIHICAVVVEYKTKTGRTNKG VATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKSTNPVIMVPGGTGIAPFMGFIQERGWLKQKGKEVGETVLYFGCRHKKNEDYIYQEELEGAECN
GVLTQLNVAFSRDQDHKVYVQHLLTKNKEHVWKL IHTDNAHIYICGDARNMAKDVQTAFYEIAEE
VGGMTRTQAVDYVKKLMTKGRYSQDVWS

>N100

MPDMDPETHTTEPAPVDEEASLFSMDLLLSVIVGFVVYWFMSRKKKEPIPEFKKLDTPPTTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMVADPEEYSMGELSRLSEITN
SLAIFCMATYEGEDPTDNAQDFYDWLQENDDELSGLNYTVFGLGNKTYEHYNAMGKYVDKRLEQL
GATRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKEHTDINMNKVYT
GEIGRLKSFVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQILGVDLDVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENMRKMASSSPEGKALYQSWVLDASRNILAILEDMP SLRPPVDHLCCELLPRLQARY
YSIASSSSKVHPNSIHICAVVVEYETKTGRTNKG VATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKSTNPVIMVPGGTGIAPFMGFIQERGWLKQKGKEVGETVLYFGCRHKKNEDYIYQEELEGSEKN
GVLTQLNVAFSRDQDHKVYVQHLLTKNKEHVWKL IHTDNAHIYICGDARNMAKDVQTAFYEIAEQ
VGGMTRTQAVDYVKKLMTKGRYSQDVWS

>N101

MADMDTETTTTTPAHPVEEEEEPLFSNLDLFLSLIVGLVIYWFMSRKKKEPIPEFKKLDTPPTSTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLSEIEN
SLAIFCMATYEGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDINMNKVYT
GEIGRLKSFVQKPPFDSKNPFLAPVTVNRRLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT

NDSALVNKLGQILGVLDLVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENMRKMASSPEGKALYQSWVLDACRNILAILLEDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFGCRHKKNEDYLYQEELEEEAEKN
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKLIIHTDNAHIYICGDARNMAKDVQTAFYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N102

MADMDTETTTTPAHPVEEEEP LFSNLDIFLFLSLIVGLVIYWFMSRKKPEPIPEFKKLETPTPSTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLSEIEN
SLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGMDGDDGNLEEDFVSWREQFWPAVCEHFGVEALGDESSIRQYELKVHTDINMNKVYT
GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRRLNKGGRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQVLGVLDLVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENLRKMASSPEGKALYQSWVLDACRNILAILLEDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFGCRHKKNEDYLYQEELEEEAEKN
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKLIIHSDNAHIYICGDARNMAKDVQTAFYEIAEE
LGGMTRTQATDYIKKLMTKGRYSQDVWS

>N103

MADMDRPTLTTRAHPVEEVEP LFSNLDIFLFLSLIVGIVYWFMSRKKPEPVPEFKKLETQAPSTR
ETSFIEKMKMGKNIIVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMAELSRLEIEN
SLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGLGDDGNLEEDFVSWREQFWPAVCEHFGVEALGDESSIRQYELKVHTDINMNKVYT
GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRRLNKGGRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQVLGVLDLVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITQPPRTNVLYELA
QYASDGKDQENLRKMASSPEGKALYQSWVLDACRNILAILLEDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYQTKTGRVNGVATNWLKNKLIIASDNGHKSTVPMYIRKSQFRL
PFKATNPVLMVGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFGCRHKKNEDYLYQEELEEEAEKN
GVLTQLNVAFSRDQEQKVYVQHLLKANKEHVWKLIIHSENAHIYICGDARNMAKDVQQLAFYEIAEE
QGEMTRSQATDYIKKLMTKGRYSQDVWS

>N104

MADMDPETTTNPAHTVEEEEP LFSNLDLFLFLSLIVGLVIYWFMSRKKPEPIPEFKKLDTPTPSTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLSEIEN
SLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGLGDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDINMNKVYT
GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRRLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGQILGVLDLVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITHPPRTNVLYELA
QYASDPKDQENMRKMASSPEGKALYQTWVLDSCRNIIAILLEDMPSLKPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFGCRHKKNEDYLYQEELEEEAEKN
GVLTQLNVAFSRDQEQKVYVQHLLKKNKEHVWKLIIHTDNAHIYVCGDARNMAKDVQTAFYEIAEE
LGGMTRTQATDYVKKLMTKGRYSQDVWS

>N105

MADMDPETTTNPAHTVGEEEP LFSNLDLFLFLSLIVGLVIYWFMSRKKPEPIPEFKKLDTPTPSTR
ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYSMGELARLSEIEN
SLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
GAKRIFDLGLGDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDVNMNKVYT
GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRRLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
NDSALVNKLGEILGVLDLVVISLNNLDEESNKKHFPFCPTTYRTALTHYLDITNPPRTNVLYELA
QYASDPKDQENMRKMASSPEGKALYQTWVLDSCRNIIAILLEDMPSLKPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
PFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFGCRHKKNEDYLYQEELEEEAEKN
GVLTQLNVAFSRDQDHKVYVQHLLKKNKEHVWKLIIHTDNAHIYVCGDARNMAKDVQTAFYEIAEE
LGGMTRTQATDYVKKLMTKGRYSQDVWS

>N106

MADMDPETTTNPAHTAGEEEP LFSNLDLFLFLSLIVGLVIYWFMSRKQPEPIPEFKKLDTPPTTR
ETSFVEKMKKTGKNIIVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEEYSMGELARLAEIEN
SLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL

GAKRIFDLGLGDDDDGNLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDVNMNKVYT
 GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
 NDSELVNKLGEILGVLDLVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELA
 QYASDPKDQENMRKMASSPEGKALYQNWVLDSRNILAVLEDMPSLKPPIDHLCELLPRLQARY
 YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGKVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
 PFKASNPVIMIGPGTGIAPFMGFIQERGWLKQQGKEVGETVMYFGCRHKNEEDYIYQEELEEAEN
 SVLTQLNVAFSRDQDHKVYVQHHLKKNKEHIWKLHITDNAHIYVCGDARNMAKDVQTAFYEIAEE
 YGSMTRTQATDYVKKLMTKGRYSQDVWS

>N107

MADMDPETATNPAHTAGEEEPLFSNLDLFLSLIVGLVIYWFMSRKQPEPIPEFKKLDTPPTTR
 ETSFVEKMKKTGKNIIVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEEYSMGELARLAEIEN
 SLAIFCMATYGECDPTDNAQDFYDWLQENDDELSTGLNYTVFALGNKTYEHYNAMGKYVDKRLEEL
 GAKRIFDLGLGDDDDGNLEEDFISWREQFWPAVCEHFGVEASGDESSIRQYELKVHTDVNMNKVYT
 GEIGRLKSFEVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPT
 NDSELVNKLGEILGVLDLVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELA
 QYASDPKEQENMRKMASSPEGKALYQNWVLDSRNILAVLEDMPSLKPPIDHLCELLPRLQARY
 YSIASSSKVHPNSIHICAVVVEYKTKTGRVNGKVATNWLKNKLVATDNGHKSTVPMYIRKSQFRL
 PFKASNPVIMIGPGTGIAPFMGFVQERGWLKQQGKEVGETVMYFGCRHKNEEDYIYQEELEEAEN
 SVLTQLNVAFSRDQDHKVYVQHHLKKNKEHIWKLHITDNAHIYVCGDARNMAKDVQTAFYEIAEE
 YGSMTRTQATDYVKKLMTKGRYSQDVWS

>N108

MGASDTEGSLPEPAAEEESFFSMVDVFLSLIAGLLVYWFFFRKKKEEIPFCKPIQPVAPTLR
 ETSFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMRGMAADPEEYDMSLSRLSEIDN
 SLAVFCMATYGECDPTDNAQDFYDWLQETDADLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEEL
 GAQRIFDLGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYELVVHTDINMNKVYT
 GEMGRKLSYEIQKPPFDAKNPFLAPVTVNRKLNKGGGRHLMHLELDITGSKIRYESGDHVAVYPT
 NDAALVNELGKILNTDLDTVFLSNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELA
 QYATDPKDQEFRLKMASSPEGKAQYLSWVVEARNILAILQDLPSLRPPIDHLCELLPRLQARY
 YSIASSSKVHPNSIHICAVVVEYETKTGRVNGKVATNWLKNKVPATDNGHKSTVPMYVRKSQFRL
 PFKPSTPVIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYGCRRHKNEEDFYQEELEQFQKE
 GVLTQLNVAFSRDQAQKVYVQHLLKKNKEHVWKLHEDNAHIYVCGDARNMARDVQNAFYEIVAE
 FGNMTQTQAVDYIKKLMTKGRYSQDVWS

>N109

MGASDTEGSLPEPAAEEEPFLSMLDVFLSLIILGLLVYWFFFRKKKAEIPFCKPIQPVAPTLR
 ETSFIDRMKKTGKNIIVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEECDMEELSRLGEIDN
 SFAVFCMATYGECDPTDNAQDFYDWLQEDDALSSGLRYAVFGLGNKTYEHFNAMGKFVDKRLEEL
 GAQRLFDLGLGDDDDANIEEDFVTWRERLWPAVCEHFGIEASGEDGSMRQYSLVEHEDINKNKIYT
 GEMARLHSYENQKPPYDAKNPFLAPVKVNRELHKGGGRSCMHIEFDITGSRIRYEAGDHVAVYPT
 NDAAMVEELGKLLNTNLDTVFTLNNVDEESSKKHPFPCPTTYRTALLHYLDITNPPRTNVLHELI
 EYASDPKDKEFLRMASSPEGKKQYQEWVESRRNLAAILKDLPSCRPPIDHLCELLPRLQARY
 YSIASSPKVHPNSIHICAVVVEYETKTGRVNGKVATNWLKNKVPATDNGHKPRVPYIYVRKSQFRL
 PFKPSTPVIMIGPGTGLAPFIGFIQERAYLKEQGKEVGETVLYYGCRRKNEEDFYQEELEQYQKE
 GVLTQLNMAFSRDQAQKVYVQHLLKQNKESVWKLHEDNAHIYVCGDARHMARDVQNALYEIVVEE
 KGNMTQTQAVDFIKGLMNKGRYSQDVWS

Figure 21

>eca3a89/1-496

MDLIPSFSMETWVLLATSLVLLYLYGTYTHGLFKKLGIPGPTPLPFFGNVLSYHKGWDFDKKCFEYKGMW
GTYHGTPVLAITDPDMIKTVLVKECYSVFTNRRPFGPFGFMKSAISLSEDEQWKRIITLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKPIITLKDIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVENTKKLFSDFDL
DPLLLSITLFPFLNAVFEVLNVFVPKSVTDFFIKSVKRMKESRLKDKKQHRVDFLQLMINSQNSKELDTHK
ALSDLELVAQSIIFIFAGYETTSSLSFLMYFLATHPDVQQKLQEEIDVTFPNKAPPTYDALVQMEYLDMLV
NESLRLFPVAVRLERVCKKDEINGVFIPKGTVMVPTFALHKKHPEFWPEPEEFRPERFSKENKDSINPYIY
LPFGAGPRNCIGMRFALNMMLALVRMLQNFSEFKPCKETQIPLKLGNOGLLPQKPIVLKVESR

>eca3a93/1-496

MDLIPSFSTETWVLLATSLVLIYLYGTYTHGLFKKLGIPGPTPLPFFGNVLSYRKGWDFDKKCFKYGKMW
GVYDGRRPVLAITDPDMIKTVLVKECYSVFTNRRPFRPVGFMKSAISLSEDEEWKRIITLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKPVTLKDIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVENSNNKLLRDFDL
NPLILMLVLPFLQPIFEVLNISLFPKSAIDFFTKSVKRMKESRLKDKKQHRVDFLQLMINSQNSKELDTHK
GLSDLELVAQSIVFIFAGYETTSTLSFLYLLATHPDVQQKLQEEIDAIFPNKAPPTYDALVQMDYLDMLV
NESLRLFPVAVRLERVCKKDAEINGVFIPKGTVMVPTFSLHRASEFWPEPEEFRPERFSKKNKDNINPCIY
MPFGNGPRNCIGMRFAMVNMMLALVRVLQNFSEFKPCKETQIPLKLGNOGLLPQKPIVLKVESR

>eca3a97/1-496

MDLIPNFSTETWVLLATSLVFLYLYGTYTHGLFKKLGIPGPTPLPFFGNVLSYRKGWDFDKKCFKYGKMW
GFYEGRLPVLAITDPDMIKTVLVKECYSVFTNRRTFGPEGFMKNAIIRSEDEQWKRIITLLSPFTTSGKLKE
MFPIIGHYGDVLRNLRNEAEKGKPVTLKNIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVNAKRLRLDFDL
DPLILSITLFPFLRPVYEALNISVFPKSVTDFFIKSVKRMKESRLKKNKETNRVDFLQLMINSQNSKEMDTHK
ALSDLELVAQSIVFIFAGYETTSTLSFLMYLLATHPDVQQKLQKEIDVTFPNKVPPTYDALVQMDYLDMLV
NESLRLFPVAVRLERVCKKDELNGVFIPKRTLVTPTFVLHRASEFWPEPEEFRPERFSKKNKDNINPCIY
LPFGTGPRNCIGMRFALNMMLAVVRVLQNFSEFKPCKETEIPIKLGTEAIVKPQKPIVLKVEPR

>eca3a94/1-496

MDLIPSFSMETWVLLATSLVLLYLYGTYTHGLFKKLGIPGPTPLPFLGTVLGYLKGFWDFDKKCFKYGSMW
GIYDGPQPVLAITDPDMIKTVLVKECYSVFTNRRSLGPVGFMKNAIISLSEDEQWKRIITLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKETEKGPVTLKNIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVENTKKLLRFNFDL
DPFFLSITLFPFLKAVFEVIHIYMFPRKRVTDFFTKSVKRMKESRLKDKKQHRVDFLQLMINSQNSKEMDTHK
ALSDLELVAQSIIFIFAGYETTSTLSFLMYLLATHPDVQQKLQDEIDVTFPNKALPTYDILLQMEYLDMLV
NESLRLFPVAVRLERVCKKDEINGVLIPKGTVMVPTFILHRASEFWPEPEEFCPERFSKNNKDNINPYIY
LPFGTGPRNCIGMRFAMNMMLAIVRVLQNFSEFKPCKETQIPLKIGNQGLIQPKPIVLKVESR

>eca3a96/1-496

MDLIPSFSVETWVLLATSLALLYLYGIYTHGLFKKLGIPGPKPLPFLGTVLGYRKGFGGFDTECLKYRKMW
GFYDGRQPVLAITDPDLIKTVLVKECYSVFTNRRSFGPVGFMKNAIISLSEDEQWKRIITLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKPIALKDIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVENTKKLLRFDFDL
DPFFLSITLFPFLNPVFELLNIFLFPKSVTDFFTKSVKRIKESRLKDKKQHRVDFLQLMINSQNSKEMDTHK
ALSDLELVAQSIIFIFAGYEPVSSLSFLYLLATHPDVQQKLQEEIDATFPNMAPPTYDALVQMEYLDMLV
NESLRLFPVAGRIERTCKKDELGGVFIPKGTVMVPTFALHRDELWPQPEEFHPERFSKENKDSINPYIY
MPFGNGPRNCIGMRFALNMMLAVVRVLQNFSEFKPCKETQIPLKLVTYGFLOPEKPIVLKVESR

>eca3a95/1-496

MDLIPSLSMETWVLLATSLVLLYLYGTYSHGHFKKLGIPGPTPLPFFGTVLGYIKGVLAFAFDKKCFKYGKMW
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPMGLMKNAITIAEDEKWKRIITLLSPFTTSGKLKE
MFPIIGQYGDMLVRNLRKEAEKGKPIITMKDIFGAYSMDVITSTSFSGVNIDSLNPNQDPFVENTKKLLRFHFDL
DPFFLSILFPFLIPVYEALNIFVFPKCVTDFFIKSVKMKESRLKDKKQHRVDFLQLMINSQNSKEIDNLK
VLSDELELVAQSITFIFAGYEPVSTLSFLYLLATHPDVQQKLQEEIDATFPNMAPPTYDALVQMEYLDMLV
NETLRLFPVADRLERVCKKDEINGVFIPKGTVMMPFVFLHQPWFSEPEEFRPERFSKKNKDSINPYTY
LPFGTGPRNCIGMRFAMVNIKLAVVRVLQNFSEFKPCKETQIPLKLVSQGLIRPEKPIVLKVESR

>cfa3a26/1-496

MDLIPSFSTETWLLAISLVLLYLYGTYTHGIFRKLGPPTPLPFGVTALGYRNGFYVFDKCFKYGKMW
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRTLGPVGFMKSAISLSEDEEWKRIITLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKAINLKDVGAYSMDVITSTSFSGVNIDSLNHPQDPFVENTKNLLKDFDL
DPFFLSILFPFLTPVFEILNIWLFPPKVTDFFRKSVERMKESRLKDKKQHRVDFLQLMINSQNSKEMDTHK
ALSDLELVAQSIVFIFAGYETTSSCLSFLMYELATHRDVQQKLQEEIDATFPNKAAPTYEALVQMEYLDMLV
NETLRLYSVAGRLERVCKKDEISGVFIPKGTVMVPTFILHRDQNLWPEPEEFRPERFSRKNKDSINPYTY
LPFGTGPRNCIGMRFAMNMMLALVRVLQNFSEFKSCKETQISLRINTRGIIQPEKPVVLKVEPR

>cfa3a12/1-496

MDLIPSFSTETWLLAISLVLLYLYGTYTHGIFRKLGPPTPLPFGVTALGYRNGFYVFDKCFKYGKMW
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRTLGPVGFMKSAISLSEDEEWKRMRTLLSPFTTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKAINLKDVGAYSMDVITSTSFSGVNIDSLNHPQDPFVENTKNLLKDFDL
DPFFLSILFPFLTPVFEILNIWLFPPKVTDFFRKSVERMKESRLKDKKQHRVDFLQLMINSQNSKEMDTHK
ALSDLELVAQSIIFIFAGYETTSTLSFLMYELATHPDVQQKLQEEIDATFPNKAAPTYDALVQMEYLDMLV
NETLRLYPVAGRLERVCKKDEISGVFIPKGTVMVPTFTLHRDQSLWPEPEEFRPERFSRKNKDSINPYTY
LPFGTGPRNCIGMRFAMNMMLALVRVLQNFSEFKPCKETQIPLKLNAQGIQPEKPIVLKVEPR

>cfa3a98/1-496
MDLIPFSMETWLLLATSLVLLLYLYGTYTHGVFKKLGIPGPTPLPFVGTALGYRKGFVFDENCFRKYGRMW
GFYDGRQPVLAITDPDPMIKTVLVKECYSVFTNRRSFGPVGFMKSAISLSEDEEWKRIRTLTLLSPTFTSGKLKE
MFPIIGQYGDVLRNLRKEAEKGKSINLKDIFGAYSMDVITSTSFVNIDSLNNPQDPFVENIKKLLKDFDL
DPFFLSILLFPFLTPVFEVLNIWLFPPKSVITDFTKSVKRMKENRLKDKQKHRVDFLQLMINSQNSKETDTHK
ALSDLELVAQSIIFIFAGYETTSTLSFLMYELATHPDVQQKLQEEIDATFPNKALPTYDALVQMEYLDMLV
NETLRLYP IAGRLERVCKKDVEISGVFIPKGTVMVPTFTLHRDQSLWPEPEEFRPERFSRKNKDSINPYTY
LPFGTGRNCIGMRFAIMNMKLALVRVLQNFSEFKPCKETQIPLKLNAGQIIQPEKPIVLKVEPR
>cfa3a99/1-496
MDLIPFSSTETWLLLATSLVLFYLYGTYTHGLFKKLGIPGPTPLPFLGTVLGYRNGFCDFDEKCFRKYGRMW
GFYDGRQPVLAIMDPDPMIKTVLVKECYSVFTNRRSFGPVGFMKSAITVSEDEEWKRIRTLTLLSPTFTSGKLKE
MFPIIGQYGDMLVRNLRKEAEKGKSAISLKDIFGAYSMDVITSTSFVNIDSLNNPQDPFVENAKKLLKDFDP
DPFLLSIIILFPFLTPLYEMLNIWLFPPKITDFTKSVKRMKESRLKDKQKHRVDFLQLMINSQNSKEMNTHK
ALSDLELVAQSIIFVAGYETTSTSLCLMYELATHPDVQQKLQKEIDATFPNKAAPTYDTLVQMEYLDMLV
NESLRLYPITGRLVRVCKKDVEISGVFIPKGTVMVPTFTLHQDPDIWPEPEKFPQPERFSKKNKDSINPYTY
LPFGTGRNCLGMRFAIMNMKLALIKVLQNFSEFKPCKETQIPLKLSSQGLIRPEEPIILNVEPR
>bta3a74/1-496
MELILSFSTETWVLLATGLVLLLYLYGTYSYGLFKKLGVPGRPLPYFGNLSYRKGVCFEDECFFKKGKIW
GIFEGKQPLLVIITDPDPMIKTVLVKECYSVFTNRRVFGPSGMKNASVAEDEQWKIRTLTLLSPTFTSGKLKE
MFPIIGKYGDVLRNLRKEAEKGTSDIKDIFGAYSMDVITSTSFVNIDSLGNPQDPFVENVKLLRFSIL
DPFLLAVVLPFLVPIILDVNIITIFPKSAVNFTKSVKRIKESRLKDNQKPRVDFLQLMINSQNSKETDNHK
ALSDQELIAQSIIFIFAGYETTSTLSFLLYIILATHPDVQQKLQEEIDATFPNKAPPTYDVLAQMEYLDMLV
NETLRMFPIAIRLERLCKKDVEIHGVSIPKGTVMVPI SVLHKDPQLWPEPEEFRPERFSKKNKDSINPYVY
LPFGTGRNCIGMRFAIMNMKLAI VRVLQNFSEFKPCKETQIPLKISSQGVLRPEKPVVLKVVLR
>bta3a100/1-496
MELIPFSMETWVLLSISLVLLLYLYGTYSHGLFKKLGVPGRPLPYFGNVLSYRKGVCFEDECFFKKGKMW
GIFEGKHPLLVIITDPDVIKTVLVKECYSVFTNRKVFGPRGIMKNASVAEDEQWKIRTLTLLSPAFTSGKLKE
MFPIIGKYGDVLRNLRKEAEKGKSVNMKEVFGAYSMDVITSTSFVNIDSLGNPQDPFVENAKKLLRFDIL
NPFLLSVVLPFLVPIILDVNIITIFPKSAVNFTKSVKRIKESRLKDNQKPRVDFLQLMINSQNSKETDNHK
ALSDQELIAQSIIFIFAGYETTSTLSFLLYIILATHPDVQQKLQEEIDATFPNKAPPTYDVLAQMEYLDMLV
NETLRMFPIAIRLERLCKKDVEIHGVSIPKGTVMVPI SVLHKDPQLWPEPEEFRPERFSKKNKDSINPYVY
LPFGTGRNCIGMRFAIMNMKLAI VRVLQNFSEFKPCKETQIPLKISSQGVLRPEKPVVLKVVLR
>bta3a76/1-496
MELIPNFSVETWVLLAISLVLLLYLYGTYSHGLFKKLGVPGRPLPLFGNVLSYRKGVCFEDECFFKKGKMW
GIFEGKHPLLVIITDPDPMIKTVLVKECYSVFTNRRVFGPMGMKNASVAEDEQWKIRTLTLLSPTITSGKLKE
MFPIIGKYGDVLRNLRKEAEKGTSDMKEVFGAYSMDVITSTSFVNIDSLGNPQDPFVENAKKLLRFDIL
DPFLLSVVLPFLPIFEVLNISIIFPKSAVNFLTTSVKKIKESRLKDTQKPRVDFLQLMINSQNSKETDNHK
ALSDQELMAQSIIFIFAGYETTSTLSFLIYELATHPDVQQKLQEEIDATFPNKAPPTYDVLAQMEYLDMLV
NETLRMFPIAVRLERFCKKDVEIHGVSIPKGTVTVPISVLHRDPQLWPEPEEFRPERFSKKNKDSINPYVY
LPFGTGRNCIGMRFAIMNMKLAVVRVLQNFSEFKPCKETQIPLKIKSQGLLRPEKPIFLKVVLR
>oar3a24/1-496
MELIPFSLETWVLLAISLVLLLYLYGTYSHGLFKKLGVSGRPLPYFGNVLSYRKGVCFEDECFFKKGKMW
GVFEGKQPLLVIITDPDVIKTVLVKECYSVFTNRRVFGPMGMKNASVAEDEQWKIRTLTLLSPTFTSGKLKD
MFPIIGKYGDVLRNLRKEAEKGKSVNMKIDIFGAYSMDVITSTSFVNIDSLGNPQDPFVENAKKLLRFDIL
DPFLLSVVLPFLVPIFEVLNITMFPKSAVDFTLTKSVKRIKESRLKDNQKPRVDFLQLMINSQNSKETDNHK
ALSDQELMAQSVIFIFAGYETTSTLSFLLYIILATHPDVQQKLQEEIDATFPNKAPPTYDVLAQMEYLDMLV
NETLRMFPIAVRLDRCKKDVEIHGVSIPKGTAVTVPISVLHRDPQLWPEPEEFRPERFSKKNKDSINPYVY
LPFGTGRNCIGMRFAIMNMKLAI VRVLQNFSEFKPCKETQIPLKINSQGLIRPEKPIFLKVVLR
>bta3a28/1-496
MELIPFSMETWVLLATSLVLLLYLYGTYSYGLFKKLGIPGRPVYFGSTMAYHKGIPFDNQCFKKGKMW
GFYEGRQPM LAITDPDIKTVLVKECYSVFTNRRIFGPMGIMKYAISLAWDEQWKIRTLTLLSPAFTSGKLKE
MFPIIGQYGDMLVRNLRKEAEKGNPVMKDMFGAYSMDVITGTAFGVNIDSLNNPHDPFVEHSKNLLRFRPF
DPFILSIIILFPFLNPVFEILNITLFPKSTVDFTKSVKIKESRLTDKQMNVRDQLQLMINSQNSKEIDNHK
ALSDIELVAQSTIFIFGGYETTSTLSFLIYELTTHPHVQQKLQEEIDATFPNKAPPTYDALVQMEYLDMLV
NETLRMFPIIAGRLERVCKKDVEIHGVITPKGTTVLVPLFVLHNNPELWPEPEEFRPERFSKKNKDSINPYVY
LPFGTGRNCLGMRFAIMNIKLALVRILQNFSEFKPCKETQIPLKLYTQGLTQPEQPVILKVVPR
>ssc3a29/1-496
MDLIPGFSTETWVLLATSLVLLLYLYGTYSHGLFKKLGIPGRPLPYFGNILGYRKGVDFHDFKCKFQQYKGMW
GVYDGRQPLLAVTDPNMIKSVLVKECYSVFTNRRSFGPLGAMRNALSLAEDEEWKRIRTLTLLSPTFTSGKLKE
MFPIISHYGDLLVSNLRKEAEKGKPVMTMKDIFGAYSMDVITSTAFGVNIDSLNNPQDPFVENS KLLKFSFF
DPFLLSLIFFPFLTPIFEVLNITLFPKSVNFTKSVKRMKESRLTDQKRRVDLLQLMINSQNSKEMDPHK
SLSNEELVAQGIIFIFAGYETTSSALSLLAYELATHPDVQQKLQEEIEATFPNKAPPTYDALAQMEYLDMLV
NETLRLYP IAAERLACKKDVEIHGVFVPKGTVVVVPFVLHRDPDLWPEPEEFRPERFSKKNKDSINPYTY
LPFGTGRNCIGMRFAIMNMKLALVRVLQNFSEFKPCKETQIPLKLTTQGLTQPEKPVVLKILPR
>ssc3a39/1-496
MDLIPGFSTETWVLLATSLVLLLYLYGTYSHGLFKKLGIPGRPLPYFGNILGYRKGVDFHDFKCKFQQYKGMW
GFFDGRQPVLAITDPDPMIKTVLVKECYSVFTNRRSFGPRGAMRTALS LAEDEEWKRIRTLTLLSPTFTSGKLKE
MFPIISHYGDLLVSNLRKEAEKGKPVMTMKDIFGAYSMDVITSTAFGVNTDFLNNPQDPFVENS KLLKFSFF

SPLFLSIIFFPFLTPILEVLNVTLPKSVVNFMRSIKRMKESRLKDKQTHRVDLQQLMINSQNSKETDTHK
 GLSDEELVAQGVFFIFAGYETTSSSLSLVYELATHPDVQQLQEEIDATFPSPKALPSYDALAQMEYLDMMV
 NEILRLYP I AARLERVCKKDVEIHGVSVPKGTVMMPVFSIHRDPELWPEPEEFRPERFSKKNKDSINPYTY
 LPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQTPLKLSQGLIQPEKPILLKVVR
 >ssc3a22/1-496
 MDLIPSFVETWLLLATSLVLLYLYGTYSHGLFKKLGPGLPYLGSVLEYRKGMMWHDNCKFKYKGMW
 GFYDGRQPVLMARDPEMIKAVLVKECYSFTNRRNFGPAGVMKNASIAKDEQWKRVRTLLSPFTTSGKLKE
 MFPIIAQYGDMLVRNLRKEAEKGNISINTKEIFGAYSMDVITSTAFGVDVDSLNNPRDPFVQYARKLLRDFDL
 DPFIILSIFPFLSPFFEVLDITLFPRSSVKFFTQSVKRMKESRLKDDQMMRRVDLLQQLMINSQNSKETDAHK
 ALSDQELVAQSIIFICAGYETTSSSLSLFLAYILATHPDVQQLQEEIDATFPSPKALPSYDALAQMEYLDMMV
 NEILRLYP I AARLERVCKKDVEIHGVSVPKGTVMMPVFSIHRDPELWPEPEEFRPERFSKKNKDSINPYTY
 LPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQTPLKLSQGLIQPEKPILLKVVR
 >hsa3a7/1-496
 MDLIPNLAVETWLLLAVALSLVLLYLYGTRTHGLFKKLGPGLPPLFGLNALSFRKGYWTFDMECYKKYRKVW
 GIYDCQQPMLAITDPDMIKTVLVKECYSVFTNRRPFGPVGMKNASIAEDEEWKRIRSLSPFTTSGKLKE
 MVIPIAQYGDVLVRNLRREAEETGKPVTLKHVF GAYSMDVITSTSFVGSIDSLNNPQDPFVENTKKLLRFNPL
 DPFVLSIKVFPFLTPILEALNITVFPKVISFSLTKSVKQIKEGRKLTQKHRVDLQQLMIDSQNSKDSETHK
 ALSDLELMAQSIIFIFAGYETTSSVLSFIYELATHPDVQQLQKEIDTVLPNKAPPTYDITVLQLEYLDMV
 NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIDPIYI
 TPFSGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLRFGGLLLEKPIVLKAESR
 >ptr3a67/1-496
 MALIPNLAVETWLLLAVALSLVLLYLYGTHSHGLFKKLGPGLPPLFVGNVWSFRKGFWTFDMECYKKYRKVW
 GFYDGLQPVLAITDPDMIKTVLVKECYSVFTNRRPLGPVGMKSAISRAEDEEWKRIRSLSPFTTSGKLKE
 MVIPIAQYGDVLVRNLRREADTGMPVTLKDVFGAYSMDVITSTSFVGNIDSLNNPQDPFVENTKKLLRFNPL
 DPFILSIKIPFLTPILEALNITVFPKVISFSLTKSVKQIKEGRKLTQKHRVDLQQLMIDSQNSKDMESHK
 ALSDLELVAQSIIFIFAGYETTSSVLSFIYELATHPDVQQLQEEIDAVLPNKAPPTYDITVLQLEYLDMV
 NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIDPIYI
 TPFNGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLRFGGLLLEKPIVLMAESR
 >hsa3a4/1-496
 MALIPDLAMETWLLLAVALSLVLLYLYGTHSHGLFKKLGPGLPPLFGLNLSYHKGFCMFDMECHKKYKGVW
 GFYDQQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGMKSAISIAEDEEWKRIRSLSPFTTSGKLKE
 MVIPIAQYGDVLVRNLRREAEETGKPVTLKDVFGAYSMDVITSTSFVGNIDSLNNPQDPFVENTKKLLRDFDL
 DPFFLSITVFPFLIPILEVLNICVFPREVTFNLRKSVKRMKESRLKDTQKHRVDLQQLMIDSQNSKETESHK
 ALSDLELVAQSIIFIFAGYETTSSVLSFIYELATHPDVQQLQEEIDAVLPNKAPPTYDITVLQMEYLDMMV
 NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYALHHDPKYWTEPEKFLPERFSKKNKDNIDPIYI
 TPFSGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLSLGLLQPEKPVVLKVESR
 >cja3a21/1-496
 MDFIPNLAVETWLLLAVALSLVLLYLYGTHSHGLFKKLGPGLPPLFGLTVLYYRQGFWKFDMECYKKYKGMW
 GIYDGRQPVLAITDPNIKTIVLVKECYSVFTNRRPFGPVGMKSAISIAQDDEWKIRIRSLSPFTTSGKLKE
 MVIPIAQYGEVLVRNLRREAEKGPINMKDIFGAYSMDVITGTSTFGVNIDSLNNPQDPFVESTKKLLRDFDL
 DPFFLSITIFPFLTPILEALNISMFPDSTSLRKSIRKIKESRLKDTQKHRVDLQQLMIDSQNSKETESDK
 ALSDLELVAQSIIFIFAGYETTSSVLSFIYELATHPDVQQLQEEIDAVLPNKAPATYDITVLQMEYLDMMV
 NETLRLFPVAMRLERVCKKDVEINGVFIPKGVVVMIPSYALHHDPKYWTEPEKFLPERFSKKNKDNIDPIYI
 TPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLRGLLQTEKPIVLKVEPR
 >cja3a5/1-496
 MDLIPNLAVETWLLLAVALSLVLLYLYGTRSHGIFKKLGPGLPPLFVGNILSYRQGIWKFDSECHKKYKGMW
 GSYDGLQPVLAITDPDIKAVLVKECYSVFTNRRPFGPVGLMKSASIAQDDEWKIRIRSLSPFTTSGKLKE
 MFPIIAQYGDVLVRNLRREAGKGPVTMKDIFGAYSMDVITGTSTFGVNIDSLNNPKDPFVESVKKFLKDFDL
 DPLFLSTIFFPFLTPVFEALNFSLPKDAINF LKQSVNRMKKSRLNDKQKHRVDLQQLMIDSQNSNETASHK
 ALSDLELLAQSIIFIFAGYETTSSVLSFIYELATNPVQQLQEEIDVLPNKAPATYDAVQMEYLDMMV
 NETLRLYP I AVRLERVCKKDVEINGVFIPKGAIVPTIYALHHDPKYWTEPEKFRPERFSKKNKDSIDPIYI
 TPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLVQGLLQAEKPIILKVESR
 >hsa3a5/1-495
 MDLIPNLAVETWLLLAVALSLVLLYLYGTRTHGLFKRLGPGLPPLLGNVLSYRQGLWKFDTECYKKYKGMW
 GYEGQLPVLAITDPDVIRTVLVKECYSVFTNRRSLGPVGMKSAISLAEDEEWKRIRSLSPFTTSGKLKE
 MFPIIAQYGDVLVRNLRREAEKGPVTMKDIFGAYSMDVITGTSTFGVNIDSLNNPKDPFVESTKKFLKFGFL
 DPLFLSIIFPFLTPVFEALNVSFLPKDTINFLSKSVNRMKKSRLNDKQKHRDLQQLMIDSQNSKETESHK
 ALSDLELLAQSIIFIFAGYETTSSVLSFIYELATHPDVQQLQKEIDAVLPNKAPPTYDAVQMEYLDMMV
 NETLRLFPVAVRLERTCKKDVEINGVFIPKGSVMVPTIYALHHDPKYWTEPEEFRPERFSKKNKDSIDPIYI
 TPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLDLQGLLQPEKPIVLKVDSR
 >hsa3a43/1-496
 MDLIPNFAMETWVLVATSLVLLYLYGTHSHKLFKKLGPGLPPLFGLTILFYLRGLWNFDRECNEKYGEMW
 GLYEGQPMVLIMDPDMIKTVLVKECYSVFTNQMLPGMGLKLSALSFAEDEEWKRIRSLSPFTTSGKLKE
 MVIPIISQCGDMLVRSRQEAENSKSINLKDFFGAYTMDVITGTTLFGVNLDLNNPQDPFLKNMKLLKLDL
 DPFLLLISLFPFLTPVFEALNIGLFPKDVTHFLKNSIERMKESRLKDKQKHRVDFFQQMIDSQNSKETKSHK
 ALSDLELVAQSIIFIAAYDTTSTLPLFIMYELATHPDVQQLQEEIDAVLPNKAPVTYDALVQMEYLDMMV
 NETLRLFPVVSRTVCKKDIEINGVFIPKGLAVMPIYALHHDPKYWTEPEKFCPERFSKKNKDSIDLYRY
 IPFGAGPRNCIGMRFALTNIKLAVIRALQNFSEFKPCKETQIPLKLDNLPILQPEKPIVLKVHLR

>oga3a92/1-496
MDLIPEFATETWVLLATSLVLLYLFSGTHSHGLFKKLGIPGPTPLPILGNVLSYRKGFWKFDDEECYKKGKWT
GLYDGRQPLLAITDPEMIKTVLVKECYSFTNRRPFGPVGFMKSAISISEDEEWKRIRALLSPFTTSGKLKE
MFPIMSQYADVLVRNLQEAADKGGTITLKDVFSGAYSMDVITGSTSGVNVDSLNNPQDPFVEKTKLLRFDFL
DPLFLSILFPFLTPVFEALNISMFPKDVNFNLKNSVQKMKESRLHDKQKHRVDLQQLMIDSQNSKETESHK
VLSDELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQKQLQDEIDAALPNKTPATYDALAQMECLDMVV
NETLRLFP IAGRLERVCKKDVEIKGVLPKGTVMVPTIYALHRDPKYWTEPEEFRRPERFSKKNKDSIDPYTY
MPFGSGPRNCIGMRFALVNMKLALVRVLQNFSEFKPCKETQIPMKLSPKGMLQPEKPIVLKIESR
>oga3a91/1-496
MDVILNFAMETWVLLAASLVLLYIYGTHSHGLFKKLGIPGPTPLPFFGTFLNYHRGVWDSLECYKKGKMW
GLYDQQPVLAITDPEMIKTVLVKECHSVFTNRRSFGPLGLMKSAISLSVDEWKRI RTLLSPFTTSSGKLKE
MFPIMSQYGDVLRNLQEAANKGKTIVNLKDI FGAYSMDVITSTSGVNVDSLNNPQDPFVENIKKFLKFN
TPFFLSILFPFLIPVFEALNIFVFPKDVMDFFEKSIKRIKESRLQDKQKHRVDLLQQLMIDSQNSKETESHK
VLSDELVAQSIIFIFAGYETTSSALSFLMYELATHPDVQKQLQDEIDAALPSKTPATYDAMVQMEYLDMMV
NETLRLFP IAGRLERVCKKDVEVNGVLIPKGTVMIPITFALHRDPKHWTPEEFRRPERFSKKNKDSIDPYMY
MPFGSGPRNCIGMRFALNMKIALIRILKNFSFKPCKETQIPLKLSKKA ILQPEKPIVLKIESR
>mmus3a44/1-496
MNLFSALS LDTLVLLAIILVLLYRYGTRTHGLFKKQGIPGPKPLPFLGTVLNYYTGIWKFDMECYEKYGKWT
GLFDGQTPLLVITDPETIKNVLVKDCLSVFTNRRREFGPVGIMSKAISISKDEEWKRYRALLSPFTTSGRLKE
MFPVIEQYGDILVKYLRQEAEGKMPVAMKDVLGAYSMDVITSTSGVNVDSLNNPEDPFVEEAKKFLRVDF
DPLLSVVLFPPLTPVYEMLNICMFPND SIEFFKKFVDRMQESRLDSNQKHRVDLQQLMNSHNSKDKDSHK
AFSNMEITVQSIIFISAGYETTSSSTLSFTLYCLATHPD IQKKLQAEIDKALPNKATPTCDTVMEMEYLDMMV
NETLRLYPIVTRLERVCKKDVELNGVYIPKGSVMVIPSIALHHDHPQHWPDPPEEFQPERFSKENKGSIDPYVY
LPFGIGPRNCIGMRFALNMKLA VTKVLQNF SFQPCQETQIPLKLSRQGLQPEKPIVLKVVPR
>mmus3a41/1-496
MNLFSALS LDTLVLLAIILVLLYRYGTRTHGLFKKQGIPGPTPLPFLGTVLNYYKGLWKFDMECYEKYGKWT
GLFDGQMP LFVITDPEMIKNVLVKECF SVFTNRRREFGPVGIMSKAISISKDEEWKRYRALLSPFTTSGKLKE
MFPVIEQYGDILVKYLRQEAEGKGPVTMKDVLGAYSIDVITSTSGVNVDSLNNPEDPFVEKAKGILRVDF
DPLVFSVVLFPFLTPVYEMLNICMFPKDSIEFFKKFVDRMKE SRLDSKQKHRVDLQQLMNAHNSKDKDSHK
ALSDMEITAQSIIFIFAGYETTSSSTLSFTLYCLATHPD IQKKLQEEIDETLPNKAPPTYDTVMEMEYLDMMV
NETLRLYPIGNRLERVCKKDVELNGVYIPKGSVMVIPSIALHHDHPQHWPDPPEEFQPERFSKENKGSIDPYLY
MPFGIGPRNCIGMRFALNMKLA LTKVMQNF SFQPCQETQIPLKLSRQGLQPEKPIVLKVVPR
>mmus3a16/1-496
MNLFSALS LDTLVLLAIILVLLYRYGTYTHGLFKKQGIPGPKPLPFLGTVLNYYKGLWKFDMECYEKYGKWT
GLFDGQIP LFVITDPETIKNVLVKECF SVFTNRQDFPVGIMSKSISLAKDEEWKRYRALLSPFTTSGNLKE
MFPVIEQYGDILVKYLRQEAEGKGPVAVKDVLGAYSMDVISTTSGVNVDSLNNPEDPFVENAKKVLRFDYF
DPLSLSVALFPFLTPVYEMLNICMFPKDSIEFFKKFVDRMTENRLDSKQKHRVDIYLMMEAYNSKDKDSHK
ALSEIETIAQSIIFIFAGYETTSSILSFTVYSLATHPD IQKKLQEEIDEALPNKAPPTYDTVMAMEYLDMMV
NETLRLYPI TNRLQRVCKKDVEINGIYIPKGS TVIIPSYVLHHDHPQHWPDPPEEFQPERFSKENKGSIDPYVY
LPFGNGPRNCIGMRFALNMKLALIKVLQNF SFQPCQETQIPLKLSRELLQPVKPIVLKVVPR
>mmus3a11/1-496
MDLVSALS LDTLVLLAISLVLLYRYGTRKHGLFKKQGIPGPKPLPFLGTVLNYYKGLWKFDMECYKKGKWT
GLFDGQTPLLAVTDPETIKNVLVKECF SVFTNRRDFGPVGIMSKAISISKDEEWKRYRALLSPFTTSGKLKE
MFPVIEQYGDILVKYLRQAKKGPVTMKDVLGAYSMDVITSTSGVNVDSLNNPEDPFVEKAKKLLRFDF
DPLFLSVVLFPFLTPVYEMLNICMFPKDSIEFFKKFVDRMKE SRLDSKQKHRVDLQQLMNSHNSKDKVSHK
ALSDMEITAQSIIFIFAGYETTSSSTLSFTLHSLATHPD IQKKLQDEIDEALPNKAPPTYDTVMEMEYLDMMV
NETLRLYPIANRLERVCKKDVELNGVYIPKGS TVMIPSYALHHDHPQHWPDPPEEFQPERFSKENKGSIDPYVY
LPFGNGPRNCLGMRFALNMKLALTKIMQNF SFQPCQETQIPLKLSRQGLQPEKPIVLKVVPR
>rno3a73/1-496
MDLVSALS LDTWLLAIILVLFYRFGTRTHGIFKKQGIPGPKPLPFLGTVLNYYRGLWKFDMECYKKGKWT
GLFDGQTPVFAIMDTEMIKSVLVKECF SVFTNRRNIGPVGIMSKSISVAKDEEWKRYRAFLSPFTTSGRLKE
MFPVIEQYGDILVKYLRQKVEKGPVAMKEVFGAYSMDVITSTSEVNINSINNP KDPFVEKVKKQRFDF
DPLFLSVVLFPFLTPVYEMLNICLFPKDSVAFFQKFVYRMKQTRLD SKHKHRVDLQQLMNAHNSKDKVSHK
ALSDIEIVAQAIFIFASYETTSSSTLSFVLYSLATHPDSQKKLQEEIDRALPNKAPPTYDTVMEMEYLDMMV
NETPRLYPIGYRLERVCKDKLDGVFIPKGSVVMIPFYTLQHDHPQHWPDPPEEFQPERFSKENKGSIDPYVY
LPFGNGPRNCIGMRFALNMKLALTKVLQNF SFQPCQETQIPLKLSRQRLFGPEKPIVLKVVPR
>rno3a2/1-496
MDLLSALTLETWVLLAVILVLLYRLGTHRHGIFKKQGIPGPKPLPFLGTVLNYYKGLGRFDMECYKKGKWT
GLFDGQTPVFAIMDTEMIKSVLVKECF SVFTNRRDFGPVGIMKAVSVAKDEEWKRYRALLSPFTTSGRLKE
MFPVIEQYGDILVKYLRQEAETGKPVMTMKVFGAYSMDVITSTSGVNVDSLNNPKDPFVEKTKLLRFDF
DPLFLSVVLFPFLTPVYEMLNICMFPKDSIAFFQKFVHRIKETRLDSKHKHRVDLQQLMNAHNSKDEVSHK
ALSDVEIIAQSVIFIFAGYETTSSSTLSFVLYFLATHPD IQKKLQEEIDGALPSKAPPTYDI VMEMEYLDMMV
NETLRLYPIGNRLERVCKKDIELDGLFIPKGSVVTIPTYALHHDHPQHWPDPPEEFHPERFSKENKGSIDPYVY
LPFGNGPRNCIGMRFALNMKLALTKVLQNF SFQPCQETQIPLKLSRQAILEPEKPIVLKVLPR
>rno3a1/1-496
MDLLSALTLETWVLLAVVLVLLYGFGRTHGLFKKQGIPGPKPLPFFGTVLNYYMGLWKFDVECHKKYGKWT
GLFDGQMP LFAITDTEMIKNVLVKECF SVFTNRRDFGPVGIMKAVSVAKDEEWKRYRALLSPFTTSGRLKE
MFPVIEQYGDILVKYLRQEAETGKPVMTMKVFGAYSMDVITSTSGVNVDSLNNPKDPFVEKTKLLRFDF

DPLFSLVVLFPFLTPITYEMLNICMFPKDSIEFFKKFVYRMKETRLDSVQKHRVDFLQMLMNAHDSKD KESHT
 ALSDMEITAQSIIFIFAGYEPTSSSTLSFVLHSLATHPDTQKKLQEEIDRALPNKAPPTYDVTVMEMEYLD MVL
 NETLRLYPIGNRLERVCKKDVEINGVFMFKGVSVMIPSYALHRDPQHWPEPEEFRPERFSKENKGSIDPYVY
 LPFGNGPRNCIGMRFALNMKLALTKVLQNF'SFQPKETQIPLKLSRQGLLQPTKPIILKVVR
 >rno3a23/1-494
 MDLLSALTLETWVLLAVVLVLLYGFGTRTHGLFKKQGIPGPKPLPFFGTVLNYYMGLWKFDVECHKKYGKIW
 GLFDGQMPFLAITDTEMIKNVLVKECYSVFTNRRDFGPGVIMGKAISVSKDEEWKRYRALLSPFTTSGRLKE
 MFPVIEQYGDILVKYLRQ--EKGKPVVKEVF GAYSMDVITSTSGVNVDSLNNPKDPFVEKAKKLLRIDFF
 DPLFSLVVLFPFLTPVYEMLNICMFPKDSIEFFKKFVYRMKETRLDSVQKHRVDFLQMLMNAHDSKD KESHT
 ALSDMEITAQSIIFIFAGYEPTSSSTLSFVLHSLATHPDTQKKLQEEIDRALPNKAPPTYDVTVMEMEYLD MVL
 NETLRLYPIGNRLERVCKKDVEINGVFMFKGVSVMIPSYALHRDPQHWPEPEEFRPERFSKENKGSIDPYVY
 LPFGNGPRNCIGMRFALNMKLALTKVLQNF'SFQPKETQIPLKLSRQGLLQPTKPIILKVVR
 >mau3a31/1-493
 MDLTSVLSLETWVLLAISLVLLYRLGTHKYDIFKKQGIPGPKPLPFLGNVLNYYKGIWTFDIECHKKYGKIW
 GLFEGQRPFLFTVTDTEMIKNVLVKECYSIFTNRRDFGPGVIMSKAVSISKDEEWKRIRALLSPFTTSGKLKE
 MFPPIEQYGDILVKFLRREAEGKNPVTTKEVF GAYSMDVITSTAFGVSVDSLNNPKD-LLWKGRKLLRFDF
 DPLFSLVVLFPFLIPIYEKLNVS MFPKDSISFRKFVDKTKENRLDYNQKHRVDFLQMLMNSHDSKD--SHK
 ALSDMEITAQSIIFIFAGYDTSSTLSFALYLLATHPDVQKKLQEEIDIALPNKARPSYDKVMMEYLD MVL
 NETLRLYPIGSRLERVCKQDVEMDGVFVPKGSIVMVPVAFALHYDPQYWPEPEKFRPERFSKENKGSIDPYIF
 LPFGNGPRNCIGMRFALNMKLALTKVLQNF'SLQPKETQIPMKLSRKAMLQPEKPIILKVVR
 >mmus3a59/1-496
 MELIPNLSIETWVLLVTSLVFFYIYGTYSHGLFKKLGIPGPKPLPFGTIFNYFDGMWKFDDEDCYKKYGKIW
 GFYVGPKPILAIMDPEI IKIVLVKECYSVFTNRP TLGPVGFLKKSITISEDEEWKRLRTLLSPFTTSGKLKE
 MFPIMRQYGDILVRNLRREEEKGEPINMKDIFGAYSMDVITGTSGVNVDSLNNPQNPFVQTKKILKFKIF
 DPFLLSIILFPFLTPITYEMLNLSIFPRDSMNF FKKFVKRMKKERLASKQKNRVDFLQMLMNTQNSKGQESQK
 ALSDLEMAAQAVIFIFGGYDATST SISLIMYELATHPDVQKKLQDEIDRTLPNKAPVTYDALMDMEYLD MVL
 NESLRLYPTAIRLERVSKKDVEINGVFIPKGTVMVPIPIPLHRNPEYWPEPEFCPGRFSKENEGNIDPYIY
 MPFGNGPRNCIGMRFALISMKLAVIGVLQNF'TVQPCETQIPLKISREPVFQPEKPIILKVVR
 >mmus3a25/1-496
 MELIPNLSIETWVLLVTSLVLFYIYGTYSHGLFKKLGIPGPKPLPLLGITIFNYFDGMWKFDDEDCYKKYGKIW
 GFYEGQPILAIMDPEI IKIVLVKECYSVFTNRRFGPGVGMKKAITISEDEEWKRLRTLLSPFTTSGKLKE
 MFPIMRQYGDILVRNLRREEEKGEPISMKDIFGAYSMDVITGTSGVNVDSLNNPQDPFVQKAKKILKFKIF
 DPFLLSIILFPFLTPITYEMLNLSIFPRDSMNF FKKFVKRMKKERLASNQKNRVDFLQMLMNTQNSKGQESQK
 ALSDLEMAAQAVIFIFGGYDATST SISLIMYELATHPDVQKKLQDEIDRTLPNKAPVTYDALMDMEYLD MVL
 NESLRLYPTAIRLERVSKKDVEINGVFIPKGTVMVPIPIPLHRNPEYWPEPEFCPGRFSKENEGNIDPYIY
 MPFGNGPRNCIGMRFALISIKLAVIGVLQNF'TVQPCETQIPLKISREPVFQPEKPIILKVVR
 >mmus3a57/1-496
 MELIPNLSIQTWVLLVTSLVFFYIYGTYSHGLFKKLGIPGPKPLPFGNIINYLDGMWKFDDEDCYKKYGKIW
 GFYEGQPILAIMDPEI IKIVLVKECYSVFTNRRFGPGVGMKKAITISEDEEWKRLRTLLSPFTTSSKLKE
 MFPIMRQYGDILVRNLRREEEKEEPIINMKDIFGAYSMDVITGTSCGVNDSLNNPQDPFVQKIKKILKFKFF
 DPFLLSVVLFPFLTPITYEMLNLSIFPRQSMNF FKKFVKMKKDRLDYNQKNRVDFLQMLMNTQNSKGQESLK
 ALSDLEMAAQTLIFIFGGYDGTST SISFIMYELATHPDVQKKLQDEIDRALPNKAPVTYDALMDMEYLD MVL
 NESLRLYPVSRIDRF SKKDVEIHGVFIPKGTIVMIPPIPLHRDPEYWPEPEDFCPERFSKENEGNIDPYIY
 MPFGNGPRNCIGMRFALISVKLAVTGVLQNF'TVQPCETQIPVKISRQPIFRPEKPIILKFVSR
 >rno3a18/1-496
 MEIIPNLSIETWVLLATSLMLFYIYGTYSHGLFKKLGIPGPKPVPLFGTIFNYDGMWKFDDEDCYKKYGKIW
 GFYEGQPFLAIMDPEI IKMVLVKECYSVFTNRRCFGPMGMKKAITMSEDEEWKRLRTILSPFTTSGKLKE
 MFPLMRQYGDITLLKNLRREEAKGEPINMKDIFGAYSMDVITGTSGVNVDSLNNPQDPFVQKAKKILKFKIF
 DPFLLSVVLFPFLTPITYEMLNLSIFPRQSMNF FKKFVKTMKKNRDLSNQKNRVDFLQMLMNTQNSKGQESQK
 ALSDLEMAAQAIIFIFGGYDATST SISFIMYELATRPVQKKLQNEIDRALPNKAPVTYDALMEMEYLD MVL
 NESLRLYPIATRLDRVSKKDVEINGVFIPKGTIVTIPPIPLHRNPEYWLEPEEFNPERFSKENKGSIDPYVY
 LPFGNGPRNCIGMRFALISMKLAVIGVLQNFNIQPCETQIPLKISRQPIFQPEGPIILKLVR
 >mau3a10/1-496
 MELIPNLSIETWVLLAISLVLIYIYGTYSHGTFKKLGIPGPKPLPFFGTILGYHDGTWKFDETCYKKYGKIW
 GFYDGRVPVLAITADPEI IKTVLVKECYSNFTNRRSFGPGVGMKKSITISKDEEWKRLRTLLSPAFTSGKLKE
 MFPPIIGQYGDITLVKNLRREEEKGKPVNMKEILGAYSMDVITGTSGVNVDSLNNPEDPFVQKARKILKFNFF
 DPFILSIILFPFLTTIYDLRFSIFPRQSTNFFKKFIFTMKNRLHSNQKTRMDFFQLMMNTQNSKGKESQK
 ALSDLEMAAQAIIFIFAGYESTST SICLVLYELATHPDVQKKLHDEIDSALPNKAPVTYDVLMGMGEYLD MVI
 NEGLRLYPIANRLERISKKAIVEINGLFI PKGTIVMVPTYPLHRDPEYWPEPEEFRPERFSKENKGSIDPYVY
 MPFGNGPRNCIGMRFALLSMKLAVSVLQNF'TLQTCETENHLKFARQIILQPENPIILKIISR
 >cap3a14/1-496
 MDLVPSFSLETWVLLALS LVLLYRYATYSHGFFKKLGIPGPKPLPFGNVLSYRKGMSFDEICRKKYGNMW
 GLYDGPQPVLAITEPDIKAVLVKECYSVFTNRRSLVPVGMKKAIVSLSEDEEWKRIRQTQLSPNFTSGKLKE
 MFPPIIKQYGDVLVKNLRQEAEGKPVQLKEIFGAYSMDIIVATAFGVNVDSLNNPHDPFVSKARKLRFDFL
 SPFLLSIVMFPFLTQLYEMLSISIFPRDSLKFFTKFVKKTKENHLESNKKQRVDFLQMLMNSQNFKD TESHK
 ALSDVEILAQSIIFIFAGYETTSSTLSFIMYSLATHPDVQKKLQEEIDKTLPNKAPPTYDVMEMEYLD MVL
 NETLRLYPTVNIERMSKKDVEINGMSFPKGTGVMIPSFALHRDSKYWPEPEDEFRPERFSKKNKENIDPYIY
 MPFGNGPRNCIGMRMAMNLKLALIRLLQNF'SFYPCKETQIPLRLGSEALLQPAKPIILKVVR

>cap3a15/1-496
MDLIPSFSLTETWVLLALSIVLLLYQYATYSHGFFKKLGIPGPKPLPLFGNVLSYRKGWMSFDFIECRKKYGNMW
GLYDGRQPVLAITEPDMDIKAVLVKECYSVFTNRRSLVPVGMKKSVLSLSEDEEWKRIRITQLSPNFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVQLKEIFGAYGMDIIATAFGVNVDSLNNPHDPFVSKANKLFRDFDL
SPFLLSVMPFPLTQLYEMLSISIFPRDSLKFFTKFVKRTKENHLESNKKQVRNVLQMMLSQNFKDTESHK
ALSDVEILAQSIIFIFAGYETSSSTLSCIMYSLATHPDVQKKLHQEIDKTLPNKAFPTYDVMMEYLDMMV
NETLRLYPVANRIERMSKKDFEINGMSFPKGTLMVIPSFALHRDSKYWPEPDEFRRPERFSKKNKENIDPYIY
MPFGNGPRNCIGRRMALMNLKALIRLLQNF SFYPCKETQIPLRLSSEALLQPEKPLILKVVSR

>cap3a17/1-496
MDLIPSFSLTETWVLLALSIVLLLYQYATYSHGFFKKLGIPGPKPLPLFGNVLSYRKGWNSFDFIECHKKYGNMW
GLYDGRPVPVLSITEPMDIKAVLVKECYSVFTNRRSIFPAGFMKKALSISKDEEWKRIRITQLSQNFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVQLKEIFGAYSMDIIATAFGVNVDSLNNPHDPFVSKASKLFRDFDL
SPFLLSVVFPPFLTQLYEMLNISIFPRDSLNF TKFVKRTKENHLESNEKQVRNVLQMMLSQNFKDTESHK
ALSDVEILAQSIFFIVAGYETTSSTLCFIMYSLATHPDVQKKLQOEIDKTLPNKAFPTYDVMMEYLDMMV
NETLRLYPVTNRIERMSKKDFEINGMSFPKGTGVMIPSFALHRDSKYWPEPDEFRRPERFSKKNKENIDPYIY
MPFGNGPRNCIGRRMALMNLKALIRLLQNF SFYTCKETQIPLRLSGSQVILQPAKPIILKVVSR

>mmus3a13/1-496
MDLIPNFSMETWMLLATSLVLLLYLYGTHSHGIFKKLGIPGPKPLPFLGTILAYQKGFWECDIQCHKKYGNMW
GLYDGRQPVLAITDPDI IKTVLVKECYSFTNRRRFGPVGILKKAISISENEEWKRIRALLSPTFTSGRLKE
MFPIINQFTDVLVRNMRQGLGEGKPTSMKDI FGAYSMDVITATSFGVNI DSLNNQDPFVEKIKKLLKFDIF
DPLFLSVTLFPFLTPVFDALNVSLFPRDVISFFTTSVERMKNRMKEKEKQVRDVLQLMINSQNYKT KESHK
ALSDVEIVAQSVIFIFAGYETTSALSALYLLAIHPDVQKKLQOEIDAALPNKAPATYDTLLQMEYLDMMV
NETLRLYPVAGRLEVRCKTDVEINGLFI PKGTVMVMTFALHKDPKYWPEPEEFRPERFSKKNQDSINPYMY
LPFGSGPRNCIGRMALNMKVALVRVLQNF TVQPKETEIP LKLSKQGLLQPENPLLLKVVSR

>rno3a9/1-496
MDLIPNFSMETWMLLVISLVLLLYLYGTHSHGIFKKLGIPGPKPLPFLGTILAYRKGFWEFDKYCHKKYGNMW
GLYDGRQPVLAITDPDI IKTVLVKECYSFTNRRRFGPVGILKKAISISEDEEWKRIRALLSPTFTSGKLKE
MFPIINQYTDMLVRNMRQSEEGKPTSMKDI FGAYSMDVITATSFGVNVDSLNNQDPFVEKVKLLKFDIF
DPLFLSVTLFPFLTPLEFALNVSMFPRDVIDF FKT SVERMKNRMKEKEKQVRDVLQLMINSQNSKV KDSHK
ALSDVEIVAQSVIFIFAGYETTSALSALYLLAIHPDIQKKLQOEIDAALPNKAHATYDTLLQMEYLDMMV
NETLRLYPVAGRLEVRCKTDVEINGVFI PKGTVMVMTFALHKDPHYWPEPEEFRPERFSKKNQDNINPYMY
LPFGNGPRNCIGRMFALNMKVALVRVLQNF SFQPKETQIPLKLSKQGLLQPEKPLLLKVVSR

>rno3a62/1-496
MDLIPNISLETWMLLATILVLLLYLYGTSTHGNFKKLGISGPKPLPFVGNILAYRHGFWEFDRHCHKKYGNMW
GFYEGRQPLAITDPDI IKTVLVKECYSFTNRRSFGPAGILKKAITLSEDEEWKRIRALLSPTFTSGKLKE
MFPIINQYADLLVKNVKEAEKGNPITMKDI FGAYSMDVITGT SFGVNVDSLNNQDPFVQVKLLKFNFL
DPLFLSVTLFPFLTPVEAFDITVFPKDVMMKFRTSVERMKNRMQEKVKQRDLDFLQLMINSQSSGDKESHQ
GLTDVEIVAQSIFFIFAGYETTSALSALYLLAIHPDLQKKLQOEIDAALPNKAPVTDVLSVMEYLDMMV
NETLRLFPVGGRLERVRCKDVEINGVFI PKGTVMVMTFALHKDPKCWPEPEEFCPERFRKKNQDSINPYIY
LPFGNGPRNCIGRMFALNMKIALVRVLQNF SFGLCKETQIPLKLRKKGFFQPEKPIILRAVSR

>ocu3a6/1-494
MDLI--FSLETWVLLAASLVLLLYLYGTSTHGLFKKMGIPGPTPLPFFIGTILEYRKGWDFDIECRKKYGNMW
GLFDGRQPLMVIDTDPMDIKTVLVKECYSVFTNRRSFGPVGFMKKAVSISEDEDWKRVRITLLSPTFTSGKLKE
MLPIIAQYGDVVLVKNLRQEAEGKGPVLDKEIFGAYSMDVITGT SFGVNI DSLRNQDPFVKNVRLLKFSF
DPLLLSTLFPFLTPIFEALHISMFPKDVMDFLKTSVEIKIDDRLDKQKRRVDLQLMINSQNSKEIDSHK
ALDDIEVVAQSIIFIFAGYETTSSTLSFIMHLLATHPDVQKKLQOEIDTLLPNKELATYDTLVKMEYLDMMV
NETLRLYPVAGRLEVRCKDVDINGTFIPKGTIVMMPTYALHRDPQHWTEPDEFRRPERFSKKNKDNINPYIY
HPFGAGPRNCLGMRFALNMKIALVRVLQNF SFKLCKETQVPLKLGKQGLLQPEKPIVLKVVSR

>mdo3a103/1-494
MNIIPSLSAGTWTLIVFLTLTLLYLYGTRTHKLFKNLGIPGPTPLPFFGTIVLAYRKGFADFYGCFKKYKGTW
GFFDGRQAVLAIMDPETIKTVLVKECYSVFTNRRMFGPSGFLEKAISIAKDEQWKIRITVLSPTFTSGKLKE
MFPIINQYGDVVLVKNMKKEAEKSKPVTMKEIFGAYSMDVITSTSGFIHVDSLNNQNDPFFVREIRKLIRFSFL
DPLILSITIFPFLIPLFKKLDITVFPKEATDFLAKSIIKIKEERTKSTEK--ANFLQLMMSDQTSKESHSQK
DLSDEEILAQSIIFIFAGYETTSVLSFLFYHLATNP KIQEKLQOEIDAFLPNKEAVTYDALVQMEYLDMMV
NENLRLYPVAGRIERVAKKTVELNGLTIPKGTVMVAPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYVY
LPFGAGPRNCIGRMFALMSMKVAVSRLLQEF SFRPCKETQIPLKLSNQPLLTPVPIVLQAE LR

>mdo3a105/1-494
MNIIPSLSAGTWTLIVFLTLTLLYLYGTRTHKLFKNLGIPGPTPLPFFGTIVLAYRKGFADFYGCFKKYKGTW
GFFDGRQAVLAIMDPETIKTVLVKECYSVFTNRRMFGPSGFLEKAISIAKDEQWKIRITVLSPTFTSGKLKE
MFPIINQYGDVVLVKNMKKEAEKSKPVTMKEIFGAYSMDVITSTSGFIHVDSLNNQNDPFFVREIKKLIRFSFL
DPLILSTITFPFLIPLFNKLDITVFPKEATDFLAKSIIKIKEERTKSTEK--ANFLQLMMSDQTSKESHSQK
DLSDEEILAQSIIFIFAGYETTSVLSFLFYHLATNP KIQEKLQOEIDAFLPNKEAVTYDALVQMEYLDMMV
NENLRLYPVAGRIERVAKKTVELNGLTIPKGTVMVAPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYVY
LPFGAGPRNCIGRMFALMSMKVAVSRLLQEF SFRPCKETQIPLKLSNQPLLTPVPIVLQAE LR

>mdo3a104/1-494
MNIIPSLSAGTWTLIVFLTLTLLYLYGTRTHKLFKNLGIPGPTPLPFFGTIVLAYRKGFADFYGCFKKYKGTW
GFFDGRQAVLAIMDPETIKTVLVKECYSVFTNRRMFGPSGFLEKAISIAEDEQWKIRITVLSPTFTSGKLKE
MFPIINQYGDVVLVKNMKKEAEKSKPVTMKEIFGAYSMDVITSTSGFIHVDSLNNQNDPFFVREIRKLIRFSFL

DPLILSITIFPFLIPLFKKLDITVFPKEATDFLAKSIKIKEERKKSTGK--ANFLQLMMSQTSKESHQSQK
DLSDEEILAQSIIFIFAGYETTSSVLSFLFYHLATNPKIQEKLQKEIDAFLPNKEAVTYDALVQMEYLDPMVI
NENLRLYPIAGRIERVAKKTVELNGLTIPKGTVMVAPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYVY
LPFGAGPRNCIGMRFALMSMKVAVSRLQEFSEFRPCKETQIPLKLSNQPLLTPTVPIVLQAEALR
>mdo3a101/1-478
-----IFTLPYRYGTRTHKLFKNLGIPTPLPFFIGTVLSYHKGIEGFDYGCFFKYGKTW
GFFDGTGPVLAIMDPETIKTVMVKECYSVFTNRRMLGLSGILEKAISIAEDEQWKMRTVLSAFTGGKLKE
MFPIIKQYGDVVLKVMKKEAEKSKPVTLDILGAYSMDVITSTSFGEIHVDSLNNPNDPFVRKIRKLIRFSFR
DPLILSVIFPFLIPLFKKLDITIFSKDVTDFLAKSILKIKEERKK--TEKHRVDFLQLMMSQTSKESHQSQK
DLSDEEILAQSIIFIFAGYESTSSVLCFLFYQLATNPGIQEKLQKEIDAFLPNKEAVTYDALVQMEYLDPMVI
NENLRLYPITGRIERIAKKPVELNGLMIPKGTVMVAPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYVY
LPFGVGPRNCLGMRFALMSMKVAVSRLQEFSEFRPCKETQIPLKLSYRPLLAPSVPIVLQAVLR
>mdo3a102/1-496
MNIIPNLSAGTWTLLIIFLTILYLYGTRTHKLFKNLGIPTPFPFFIGTILYRKGIVGFDYGCYKKGKTW
GFFDGTGPVLAIMDPETIKTVMVKECYSVFTNRRMLGLSGILEKAISIAEDEWKRIRTVLSAFTSGKLKE
MFPIINQYGDVVLKVMKKEAEKSKPVTMKEIFGAYSMDIISTSFGEIHVDSLNNPNDPFVREIRKLIRFSFL
DPLILSITIFPFLIPLFKKLDITVFSKDATDFLGKSILRIKEERKKSTEKHRVDFLQLMMSQTSKESHQSQK
DLSDEEILAQSIIFIFAGYESTSSVLSFLFYHLATNLEIQEKLQKEIDAVLPNKEAVTYDALVQMEYLDPMVI
NENLRLYPISGRIERIAKKTVQLNGLTIPQGTVMTPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYVY
LPFGAGPRNCIAMRFALMGMKVAVSRLQEFSEFRPCKETPPLKLNQPLLTSPVPIVLQAVLR
>mdo3a106/1-496
MNIIPSLSAGTWTLLIVFLTLILYLYGTRTHKLFKNLGIPTPKPLPFFGTVFSYRKGLVNFYDCFKKYGKTW
GFYDGRQPVLATMDPETIKTVMVKECYSVFTNRRSFGPVGSLSAITVAKDDQWKIRTVLSPTFTSGKLKE
MFPIINQYGDVVLKVMKKEAEKKNPVTMKDILGAYSMDVITSTSFGEIHVDSLNNPNDPFVREIKKLIRFNFL
DPLILSVAIFPFLTPLFNKLDITVFPKDAINFALAKSLISMFSSTILHYFFQHRVDLLQLMMSQSSKEPHSHK
DLSDEEILAQSIIFLIFAGYETTSSVLSFLFYHLATNPEIQEKLHKEIDAVLPNKEAVTYDALVQMEYLDPMVI
NENLRLYPIAGRVERVAKKTVELNGLTIPKGTVMTPPYVLHREPEYWPEPEKHFHPERFSKENKESINPYVY
LPFGAGPRNCIGMRFALNMKVAVSRLQEFSEFRPCKETQVPLKLGNGPILPTIPIVLQAEALR
>mgi3a70/1-496
MNIIPSLSAGTWALLVFLTLILYLYGTWTHKLFKDLGIPGPAPLPFFGTVLSYRKGFDTDFDMTCFKKYGKTW
GFFDGRKPVLSIMDPETIKLVLVKECFVFTNRRSFGPAGIMESAISIAEDDQWKIRITILSPTFTSGKLKE
MFPIIKQYGDVVLKVMKKEAEKSKPVSLKDFGAYSMDVITSTSFGEIKTDSLNNQDDPFVREIRKLIRFNFL
DPLLLTVSIFPFPVPLKMDVTIFPKETLDFLTKSIIKKIKEDRKNSQKNRVDFLQLMMSQTSKEMHSQK
DLSDEEILAQSVIFIFAGYETTSSVLCFLFYHLATHPEIQKQLQKEIDAVLPNKEAVTYDALVQMEYLDPMVI
HEILRLYPIAGRIERVAKKTVEINGVKIPKGTVMVSPFVLHRDPDYWPEPEEFRPERFSKENRESMNPYVY
LPFGAGPRNCIGRRFALMSMKVATSLVLQEFSEFRPCKETQIPVKLNNQPLISPTIPIVLQAVPR
>oan3a109/1-496
MSSILFSIETWILLGTFLALMILYGIWPGVFTKLRIPGPRPIPLGNILDYSQGMDEFDCKCFQKYGKIW
GLYEGRKAKIAVLDLDIISVVLVKEVFTNFTNRQAGIEGIMSESIFSAKDERWKELRTLVSPTFTSGKLKE
MVPMMIYGEVLVKNLQNKVEKNQNVTIKDYFGAYNLDTITSTSFGEIDIDNLNKMNTFVKEIKKLTSSINFF
SPLFALISVLPFMIPLKLLNINTFPNDSEYFVKRFMDKIKERTQQGNPQQRMNQLQLMLDTQNKEGKLKNG
GLSDAEILAQAFSIFAGYETSSITLSFISYNLATHLDVQKKLQEEIDEVLNPKAPLTYDALVQMEYLDMMI
NETLRLYTVGSRLEKCKTKMKNGLTIPKGMTVIVPVSCLHSDPNYWNPDDEFHPERFSKEMRDKRDPYTF
MPFGAGPRNCIGMRFALLNKKVAIIGLLQNFSLRTECQETQIPLEVNSFGLLTTKKPIVLRIFPR
>oan3a108/1-496
MSSIFSFSIETWILLGTFLALMILYGIWPGVFTKRNIPGPWPVPFPLGNIQEYSQGLMVFDKCKCEKYGKIW
GLYEGRQAIIVLDLDIISVVLVKEVFSSTNRRNTGIEGILELAVSSAEDEHWKELRTLMSTRFTSGKLKE
MLPIMTHYGEVLVRNIWKAENNVTKDYFGAYSLDVITSTSFGEVDVDSLNMNDTFVKEIKKISSIDFF
SPLFALISTFPFMIPLKLLNINLFPNDSEYFVKRFLHKVKERRHQDNQQRVDFLQLMLDSQNKESKSGKG
GLSDAEILAQAFIFAGYETSSNTLSFISYNLATHPDVQEKLOAEIDTVLPNKSPLTYDALVRMEYLDPMVI
NESLRLYPISSRLERMCKKTTEINGLTIPKGTCVVIPLFLQKDPNYWLEPDEFHPERFSKEMRDKRDPYTF
MPFGAGPRNCIGMRFALLSIKVMIGLLKNFSLRTECQETQIPPEINSFGLLTAKKPIVLRIFPR
>oan3a107/1-496
MSSILNFSIETWILLGTFLALMILYGIWPFPRVFTKSGIPGPRPVPFFGNILEYSQGIIEFDKCKCFQKYGKIW
RLFEGRQAMVAVLDLEIITVVLVKEVFSSTNRRRTDLGVLNRNAVSLAEDDQWKRLRTVLSPTFTSGKLKE
MLRIMIHYGEILVKTIWKAENNVIVKDYYGAYSLDVITSTSFGEVDVDSLNMNDTFVKEVKKLSSVDFF
SPLALTSIFPFIPTLLNLLDLSLFPDSVEYIKRFVHKIRESRHQDNQQRADLLKLLDSQNKELGKGG
GLSDGTILAQAVAFVFLAGYETTSSALSFSYNLATHPDVQKKLQEEIDKVLNKSPLTYDALVQMEYLDPMVI
NETLRLYPIGRRLERMCKKTLEINGLTIIHKGTCVVIPIAVLQHDPNYWTEPDEFRPERFSKEMREKRDYPTF
IPFGVGPRNCIGMRFALLNMKVAIIGLLQNFSLRTECQETQIPLELNSMGFLTPKKPIVLQFIPR
>gga3a37/1-496
MNFLPFFSIETWILLIFVVLIIYGTWPFGLFKKLGIPTGPRPLPFFGTCLFYRKGFDFDTECFQKYGKIW
GIYDGRQPAVVVMDPQIIKTVVLVKECYSTFTNRRRTDLGVLNRNAVSLAEDDQWKRLRTVLSPTFTSGKLKE
MFPTMKYFGEMLVKNIQKRVKNSVVPVKDFFGSYSMDVITSTSFGEVNDISMNPKSPFVREMQLTKDFDF
DPVFILSFVCPFLTPLMAKMNISFFPSDAVDFFMRSIDKIKKDRERETHGRVDFLQMMIESQKSDAKHSYK
ALSDIEVLSQAFIFIFAGYEPTSNLTGLYLAQLALHPDVQKKVNEIDTILPNKAPLTYEAIMQLEYLDMAV
NETLRLYPLGGRLEKRDVEINGVTIPKGTIVIPPYTLHRSPEYWPNEEFRPERFSKENKDNIDPYTY
LPFGAGPRNCIGMRFALLTLKVAITAVLQHFTFQVCKETQIPLKLQSLGLTTPKEPIVLKLVPR

>mga3a37/1-496
MNLLPFFSIETWILLIFVALLIAYGTWPFGLFKKLGIPGPRPLPFFGTCLFYRKGFLEFDTECFKKYKGIW
GIYDGRQPAVVVMDPQIIKTIVLVKECYSTFTNRRRVDLAGELRNAVSLAEDDQWKRLRTVLSPTFTSGKLKE
MFPTMKHFGEMLVKNVQKRVEKSSVPVKDFFGSYSMDVVTSTSFVGNIDSMNNPKSPFVREMOKLTKDFDF
DPLFILAEVCPFLTPMLAKMNVSFPPSDAVDFFLRSIDKIKKDRKETHKGRVDFLQMMIEESQKSDGKNSHK
ALSDIEVLSQAFIFIFAGYEPTSNLTCLYLALHDPVQKQVNEIDTILPNKAPLTYEAIMQLDYLDMAV
NETLRLYPLGGRIERTCKDVEINGVTIPKETIVVIPYTLHRNSEYWPNEEFRRPERFSKENKDNIDPYTY
LPFGAGPRNCIGMRFALLTLKVAITSILQHFTFQVCKETQIPLKLLSMGLMTPEKPIILKLVRP
>aca3a37/1-496
MELLPSFSIETWSFLVLLGLLLLYGIWPFVFGKLGIPGPRPWPFLGTLLQYRKGFVKFDSWCFQKYGRIW
GIYDGRQPVMAILDPAIIKTIVLVKECYSTFTNRRNFAGTGLKSAISLAEDDEHKKRVRTVLSPTFTSGKLKE
MFPIIKHYGEALMRNVQKQVMDDEPLVVVDVFGAYAMDVVTSTSFVGNIDSMNNPKDPFVKETQKLTFRSF
DPRIILFTFPIILRTVFEALNIQIFSAEAVDFFTRSFAKMKAGRAGETKEGRVDFLQLMIDSQTNGIDHTYK
GLTDDEILAQAQAFIFVAGYEPTSNLSGYMAYMLATHPDVQKQLQDEIDSVLPNKAPLTYDAIMQLEYLDMVL
NETQRLYPLGGRIERVCKDVEINGVTIPKGVVVMIPPYNLHRDQEYWPNEEFRRPERFSKENKESVDPYTY
LPFGAGPRNCIGMRFALVTMKAITILLQHFSFRVCKETPIPLELSSQAFLVPEKPIVLKLVRP
>mga3a80/1-496
MGFLPDLSPITWLLLVAFSLVLLVLYGIWPHYQTFKKLGIPGPQVPVFLGTFLGYQQGILNFDQMCFEKYKGIW
GIFDGRQPIMAILDPALIKTILVKECYSTFTNRRVFGLSGRLESASVMAVDDQWKIRTVLSPTFTSGKLKE
MFPIIESYGEKLVKNIEKKVANEELDMKNIFGAYSMDVVASTSFSDVDSDMSKPSDPFVTNIRKFLKFSFL
NPLLMFIVLFPFMIPVLEKMNVTLLSNEVMDFFDVFMMKMKEREKGYHMNRVDFLQLMIDSQSAEEKDSHK
SLSDEEILAQAQAFIFVAGYETTSTLSYISYHLAIHPDVQKQLQDEIDANLPNKAAPTYNVMMQMEYLDMMV
NESLRLYPAGGRIDRVCKKTVEFNGVTIPKDMVVMIPYVLHHDPEYWPKEEFRRPERFSKENRENIDPYTY
LPFGAGPRNCIGMRFALLVVKVAVVALLQNFSTRPCKDTPIPVLVLDTKGFMQPKKPIVLKMVPR
>gga3a80/1-496
MGLLPDLPPITWLLLAFLGLLVLYGIWPHYQTFKKLGIPGPRPLPFLGTFLRYRQGLLNFDQMCFEKYKGIW
GIYDGRQPVMAILDPVLIKTILVKECYSTFTNRRVFGLSGRLESALSAMDDQWKIRTVLSPTFTSGKLKE
MFPIIESYGDKLKLVKNIEKKVANEELDKSIFGAYSMDVVASTSFSDVIDSDMSKPSDPFVTNIRKFLKFSFL
NPLLIIFIVLFPFMIPVLEKMNVTLLPTKVMDFDAVFMKMKEREKGYHVDVDFLQLMIDSQSAREKDSYK
SLSDEEILAQAQAFIFVAGYETTSTLSYISYHLAIHPDVQKQLQDEIDANLPNKAAPTYNVMMQMEYLDMMV
NESLRLHPPGGRIERICKTVEFNGVTIPKDMVVMIPYVLHHDPEYWPKEEFRRPERFSKENGENDIDPYTF
LPFGAGPRNCIGMRFALLIVKVAMVLLQNFSTRPCKDTPIPVLVLDTKGFMQPKKPIILKMVPR
>mga3a80/1-496
MTSYSGLSTETWLLVAFVTLILYGIWPHYGIFKKLGIPGPRPLPFFGTFLFYRKGFLEFDKQCFQKYGKMW
GFYDGRPLILAILDPDIIKTIVLVKEFYTLFTNRRNFGLNGILDSGITVAEGEKWKIRSIISPTFTTSGKLKE
MFPIIKHHIDVLVNNIEKKVAQDESVNMKDIFGAYSMDVITSTSFVGDIDSINKPNPDLVHIKKLLSFSFM
SPLLILILVFPASLVPVLEKMNLSLLSKENIDFLVNASRIIADRKKNSNRDRVDFLQLMIDSQVEKRNNSPK
ALTEMEIVAQAVTIFAGYETTSTLNFITYNLATHPEIKQLLEEIDSTLBNKAVPTYDTIFQMEYLDMMV
NETLRLFLPLGGRIERICQKTAENGITIPKGTVMIPVYVLHHDPEYWPKEEFRRPERFDQEGRSIDPYVF
LPFGAGPRNCVGMRFALLTLKTALVTLLQNFTEPCKETPIPLELETKGFMQPKKPIILKLVRP
>aca3a112/1-496
MDFFSFFSIQTWLLGIFLLLLMYGIWPHYNFFKKLGIPGPRPLPFIGTFLEYNKGVLDFDERCRQKYKGIW
GFFDGRQPVMAIMDPVIIKTILVKECYTHFTNRRDFGLNGDFDASVLVAQDEHKKIRTVLSPNFTSGRLKE
MLPIINHYGEMLVKNVQKKGENG EYIDTKDIFGAYSMDVITSTSFSDIDSINNPKDPFVSMKKIITKNFV
NPLLLIVIFPFLRPLFVLEALNLSLAPTSATGFLRGALRIKSDRQKNVHQDRVDFLQLMVDSQIDGNTSYK
ALNDREILAQAFTFVFGGYETTSTLGYMAYCLATHPDVQKQLQDEIDEVLDPQATPVYDALLQMEYLDMMVI
NETLRLYPPGGRIERVCKNTVEINGVTIPKGTVMIPAFVLHRDPEFWPEPEEFRRPERFSKENKEFLNPYVF
LPFGAGPRNCIGMRFALLSLKVATVVLQRFSTRCTKETPIPLELESKGFMRPKNPIRLKFVPR
>aca3a113/1-496
MDLLSFLSPQTWVFLGVFLVLLMYGIWPHYNFFKKLGIPGPRPLPFIGTFLEYRKGVLDDEMCHKKYKGIW
GFFDGRQPVMAILDPVIIKTILVKECYTHFTNRRDFGLNGDFDSSILTVQDEQWKIRTVLSPNFTSGRLKE
MFPIINHYGEMLVKNIQKQVENGEYIDTKDIFGAYSMDVITSTSFSDIDSINNPNPDPFVTNIIKKIVKFNFV
NPLLVILVIFPFLRPLFVLEALNLSLTPSTVNTF IKESIKKLLKDRQKNLHKDRVDFLQLMVDSQILDDSKSYK
ALNDKEILTQAIIFVAGYETTSTLSYMAYCLATHPDVQKQLQEEIDEALPNQATPVYEVLLQMEYLDMMVI
NETLRLYPPGGRIERVCKNTVEINGVTIPKGTVMIPAFVLHRDPEFWPEPEEFRRPERFSKENKETLDPYTF
LPFGAGPRNCIGMRFALLSLKVAMIVLLQRFSTRCTETPIPLELQSQGFMTPKKPIRLKFVPR
>aca3a114/1-472
-----YGIWPHYNFFKKLGIPGPRPLPFIGTLHKHRIGVNNFDQKSHEKYKGIW
GLFDGRQPVIAILDPAIIKIILVKHFYTHFTNRRDFGLNGELDTSAITRDEHKKIRTVLSPNFTSGRLKE
MLPIINHYGKILVESAQKKVENGEAIDLSFFGGYSMDVITSTSFSDIDSINNPNPDPFVMIKTFLKFDTF
DPVILLVVFPLRPLFRLNENFVAPSSFTTFIDVLKVKKDRQKNLHKNRVDFLQLMVDSQIPEEAKMYK
ALTDKEILTQAIIFVAGFETTSTLSFLSHCLATNPVQKQLQEEIDATLPDQATPTYNAIQMEYLDMMVI
NETLRLYPPGGRIERDCKNTVEINGVTIPKGTVMIPAFVLHRDPEFWPEPEEFRRPERFSKENKETLDPYTF
LPFGAGPRNCIGMRFALLVVKVAIVLLQKHTLRPCKETQIPLEVHGVLLPEPKKTIKLKLVRP
>xtr3a82/1-496
MNLPHLSTGTWILLAVLLVLLLYGIWPHYGLFKKMGIPGPTPLPFIGTFLEFRKGMIFQDIECFKKYGRMW
GMYDGRQPVLAIMDPAAIKTILVKECYTNFTNRRNFGLNGPLESAITVAEDEQWKIRSVLSPTFTSGKLKE
MFQIMKDYSDILVKNVQVSDKDEPCATKDVIGAYSMDVITSTSFSDIDSINNPNPDPFVTHIKNLLKTGFL

SPVIFAVIFPFLRPIFEVLNISFFPKDFTQFFMNAVTSFREKRKKGDHSGRVDLLQLMVDSGTNDSSNQHK
ALTDAEIMAQSLIFIFAGYETTSTALSIFYNLATHPDVQQLHHEIDSFLPDKASPTYDILMQMEYLDPMVI
QETLRLFPFAGRLERSKQNEINGVSIIPKGTVMIPAYVLQORDPEYWPEPEEFRPERFSKENRATQTPFTF
LPFGDGRNCIGLRFALLSMKVAIATLLQNF SVRPCAETLIPMEFSTIGVLQPKKPIVLKFLSR
>xtr3a83/1-496
MTFLPDFSMATWTLVLLLTLLAYYAIWPYKLFKRYGIPGPTPIPFIGTFLGNRHGLMEFDMCECFKKGKGVW
GIYEGQKPLLAIVDPVVIKSIMVKECYTNFTNRRDFGLSGPLKSSVLISKDEQWKRI RTVLSPTFTSGKLKQ
MFPLMNHYGELLVKNIHKINKEPLDMKHIFGSYSMDIVLSTSF SVNVDSMNPNNDPFTVNARNLFTFSFF
NPLFLISILCPFLVPLLDKMNFCFLSLKILNFFKDAVASIKKKRQKGTHERVDFLQLMVDAQSEEEKHGYK
ELSDTEILAQSLIFIMAGYETTSTTLMFLAYNIATHPDVQQRKLEEEIDALLPNKAPPTYDALMKMEYMDPMVI
NESLRMFPPAIRIDRVCKTMEINGVTIPAGVVIVVPLFALHLNPEIWPEPEEFQPERFSKENQKNQDPYNF
LPFGVGRNCIGMRFALVNMKVALTILLQNFRLKETCKDTPVPLKICTKGYLKPTKPIILNLIPK
>xtr3a120/1-496
MTFLPDFSMATWTLVLLLTLLAYYAIWPYRFLFKRYGIPGPTPIPFIGTFLGNRKGLMEFDMCECFKKGKGVW
GIYEGQKPLLAIVDPVVIKSIMIKECYSNFTNRRDLGLTGPLKSSVLVSKDEQWKRI RTVLSPTFTSGKLKQ
MFPLMKHYGELLVKNIHKIDNKEPLDMKHIFGSYSMDIVLSTSF SVNVDSMNPNNDPFTVNARNLFTFSFF
NPLFLIAILCPFLVPLLDKMNFCFLSLKILNFFKDAVASIKKKRQKGTHERVDDLQLMVDAQSEEEKHGYK
ELSDTEILAQSLIFIMAGYETTSTTLMFLAYNIATHPDVQQRKLEEEIDALLPNKAPPTYDALMKMEYMDPMVI
NETLRMFPPAIRIDRVCKTMEINGVTIPAGVVIVVPLFVHLNPEIWPEPEEFQPERFSKENQKNQDPYNF
LPFGVGRNCIGMRFALVNMKVLVTILLQNF RFETCKDTPVPLKICTDGYLKPTKPIILNLIPK
>xtr3a121/1-496
MTFLPDFSKBTWTLVLLLTLLAYYAIWPYRFLFKRYGIPGPTPIPFIGTFLGNRHGLMEFDMCKCFKKGKGVW
GFYDGPQPVLAIDLDPVVIKSIMVKECYTNFTNRRDFGLSGPLKSSVLMSKDEQWKRI RTVLSPTFTSGKLKQ
MFPLMKHYGELLVKNIHKIDNKEPLDMKSIFGSYSMDITLSTSF SVNVDSMNPNNDPFTVNARNLFTFSFF
KPLFLITILCPFLVPLLDKMNFCFLSSKFLKFFKDAVASIKKKRQKGAHMDRVDFLQLMVDAQSEEEKHRSK
ELSDTEILAQSLIFIMAGYETTSTTLMFLTYNIARYPDVQQRKLEEEINTLLPNKAPPTYDALMKMEYMDPMVI
NETLRLPPAIRIDRVCKTMEINGVTIPAGVVIVVPLFALHLNPEIWPEPEEFQPERFSKENQKNQDPYNF
LPFGVGRNCIGMRFALVNIKLALTILLQNF KLETCKDTPVPLKVCTRGNLKPTKPIILNLVPK
>xtr3a122/1-496
MNLIPHLSTGTWILLAVLLVLILLYGIWPYGLFKKMGIPTPLPFIGTFLEFRKGMQFDIECFKKGKGRMW
GMYDGRQPVLAIMDPAI IKTILVKECYTNFTNRRNFGLSGPLESAITSAEDEQWKRI RSVLSPTFTSGKLKE
MFQIMKDYSDILVKNVQVYVDKDEPCATKDVI GAYSMDVITSTSF SVNIDSLQNP SNPFVVIHKNLKLTGLL
SPVILVVFPPFLRPIFEVLNISFFPKNFTQFFMNAVTSFREKRKKGDHSGRVDLLQLMVDSGTTEGNDQHK
ALTDAEIMAQSVIFIFAGYETTSTALSIFYNLATHPDVQQLHHEIDSFLPDKASPTYDILMQMEYLDPMVI
QETLRLFPFAGRLERSKQNEINGVSIIPKGTVMIPAYVLQORDPEYWPEPEEFRPERFSKENRATQTPFTF
LPFGDGRNCIGLRFALLSMKVAIATLLQNF SVRPCAETLIPMEFSTIGVLQPKKPIVLKFLSR
>xla3a124/1-496
MTFLPDFSMETWTLVLLLTLLAYYAIWPYRFLFKRHGIPGPTPIPFIGTFLGNRNGIMEFDMCECFKKGKGVW
GFYDGPKPLLAIVDPVVIKNIIMIKECYTNFTNRRDFGLSGPLKSSVLISKDEQWKRI RTVLSPTFTSGKLKQ
MFPVMKHYGELLVKNIQKKIDNKEPLNMKYIFGSYSMDITLSTSF SVNVDSMNPNNDPFTVNARNLFTFSFF
NPLFLITILCPFLVPLLDKMNFCFLSLKILNFFKDAVASIKKKRQKDIHEDRVDFLQLMVDAQNEGEKQRYK
ELSDDEILAQSLIFIMAGYETTSTTLMFLAYNIAMHPDVQSRLEEEIDTLLPNKAPPTYEALMKMEYMDPMVI
NETMRLFPSAIRIDRVCKTMEINGVTIPAGVVIVVPLFVHLNPEVWPEPEKQPERFSKENQKNQDPYSF
LPFGTGRNCIGMRFALVNMKLALTILLQNF RFKCEDTPDPLNICTKGYLKPTKPIILKLVPK
>xla3a123/1-496
MNLIPHLSTGTWILLAAFLLLILLYGIWPYGLFKKMGIPTPLPFIGTFLEFRKGLVQFDTECFKKGKGTW
GLYDGRQPVLAIMDPAI IKTILVKECYTNFTNRRNFGLSGPLESAITAAEDEQWKRI RSVLSPTFTSGKLKE
MFQIMKDYSDILVKNIQDNVGDPCMTKDVI GAYSMDVITSTSF SVNIDSLKNP SHPFVVIHKNLKKGGLL
NPLIILVVFPPFLRPIFEGNLNINFFPKNFTQFFTNAVTSFREKRKKGDHGTGRVDFLQLMVDSSVNDSGNQHK
ALTDVEIMAQSIIFIFAGYETTSTALSIFYCLATHPDVQQRLEHEIDSFLPDKASPTYDILMQMEYLDPMVI
QETLRLYPPAGRLERSKQNEINGVSIIPKGTVTMIPAYVLQORDPDYWPEPEEFRPERFSKENRETHTPFTF
LPFGDGRNCIGMRFALLSMKVAIVSMLQNF SVRPCAETLIPMEYSTIGVLQPKKPIVLKFLSR
>xtr3a81/1-496
MNLIPHLSTGTWILLAAFLVLLILLYGIWPYGYFKKMGIPTPLPFIGTFLEFRKGMVQFDTECFKKGKGMW
GTYDGRQPVLAIMDPAI IKTILVKECYTNFTNRRNFGNGPFESAITIAEDEQWKRI RSVLSPTFTSGKLKE
MFQIMKDYSDILVKNIQGYVEKDEPCATKDVI GAYSMDVITSTSF SVNIDSLNKP SDPFVVIHMKKLLTGLL
SPLLIILVVFPPFLRPILEGLNLNFPKDFTEFFMNAVTSFREKRKKGDHSGRVDLLQLMMDSRNDLSNKHK
ALTDAEIMAQSVIFIVAGYETTSTALSIFYNLATHPDVQQRLEHEIDSFLPDKASPTYDILMQMEYLDPMVI
QETLRLFPFAGRLERSKQNEINGVSIIPKGTVTMIPAYVLQORDPEYWPEPEEFRPERFSKENRATHTPFTF
LPFGDGRNCIGLRFALLSMKVAIVTLLQNF SVRPCAETLIPMEFSTIGVLQPKKPIVLKFLSR
>xla3a125/1-496
MNLIPHLSTGTWILLAAFLVLLILLYGIWPYGYFKKMGIPTPLPFIGTFLEFRKGMVQFDTECFKKGKGMW
GTYDGRQPVLAIVMDPAI IKTILVKEFYTNFTNRRNFGNGPFESAITIAEDEQWKRI RSVLSPTFTSGKLKE
MFQIMKDYSDILVKNIQGYVVDKDEPCATKDVI GAYSMDVITSTSF SVNIDSLNKP SDPFVVIHMKKLLTGLM
SPVLILVVFPPFLRPILEGLNLNFPNDFTQFFMNAVTSFREKRKKGDHSGRVDLLQLMVDSTNDLSNQHK
ALTDAEIMAQAVIFIVAGYETTSTALSIFYCLATHPDVQQRLEHEIDSFLPNKASPTYDILMQMEYLDPMVI
QETLRLYPPAGRLERSKQNEINGVSIIPKGTVTMIPAYVLQORDPEYWPEPEEFRPERFSKENRETHTPFTF
LPFGDGRNCIGMRFALLSMKVAIVTMLQNF SVRPCAETLIPMEFSTIGVLQPKKPIVLKFLSR

>gac3a118/1-496
MGFSLFLSAETWTLTLLVALVTLLFVYVCWNFGIFKKLGIPGPQPIPFPGTMLAYRKGITTFDEDCSKKYGKVV
GIFDGRQPVLCTIDPAMIKTILIKECYSLFTNRRNFRNLNGQLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLTSSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVANIKKMLKFDIL
SPFLFLIVAFFPFLGPIMEKMEFSFFPKSVTDFFYASALQKIKSNRVHSEQKSRVDFLQLMIDSQKLSKEESNK
GLSDHEILSQAMIFLIFAGYETSSSTLSFLAYNLATNPHIMKRLQEEVDSTFPNKGPEYQALMQMEYLD SVI
NESLRLYP IILRLERVAKATVEINGIVIPKDMVAVPTWPLHRDPSLWPEEAFKPERFSKENKESIDPYTY
MPFGAGPRNCIGMRFALVMMKLMAMVENLQRFSSFSVCKETEIPLEMDVQDILLMPKRPIKLLVSR
>gac3a117/1-496
MGFSLFLSAETWTLTLLVALVTLLFMYVCWNFGIFKKLGVPGPQPIPFPGTMLAYRKGFTTFDEDCSKKYGNVW
GIFEGRQPVLCTIDPAMIKTILIKECYSLFTNRRNFRNLNGQLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLTSSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVANIKKLLNIDFF
SPVFLIVAFFPFLGPIMEKMEFSFFPKSVTDFFYALQKIKSNRVHSEQKSRVDFLQLMIDSQKLSKEESNK
GLSDHEILSQAMIFLIFAGYETSSSTLSFLAYNLATNPHIMKRLQEEVDSTFPNKGPEYQALMQMEYLD SVI
NESLRLYP IIVRLERVAKATVEINGIVIPKDMVVMVPTWPLHRDPSLWPEEAFKPERFSKENKDSIDHYTY
MPFGAGPRNCIGMRFALMMKLMAMVEILQRFSSLSVCKETEIPLEMDTQAVVMPKRPIKLLVSR
>gac3a119/1-494
MGF--FLSAETWTLTLLVALVTLLFVYVCWNFGIFKKLGVPGPQPIPFPGTMLAYRKGFTTFDEDCSKKYGNVW
GIFDGRQPVLCTIDPAMIKTILIKECYSLFTNRRNFRNLNGQLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLTSSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVANIKKMLKFDFF
SPVFLIVAFFPFLGPIMEKMEFSFFPKSVTDFFYALQKIKSNRVHSEQKSRVDFLQLMIDSQKLSKEESNK
GLSDHEILSQAMIFLIFAGYETSSSTLSFLAYNLARNPHIMKRLQEEVDSTFPNKGPEYQALMQMEYLD SVI
NESLRLYP IAPRLERVAKATVEINGIVIPKDMVVMVPTWPLHRDPSLWPEEAFKPERFSKENKDSIDPYTY
MPFGAGPRNCIGMRFALVMMKLMAMVEILQRFSSFSVCKETEIPLEMDVQGLLMPKRPIKLLVPR
>msa3a68/1-496
MGYFFFSAEWTLLVALITLIFVYVCWPFGTFFKKMGVPGPRPIPFPGTMLAYKKGMHFDSECFKKYKGTW
GIFDGRQPVLCTIDPAMIKTVLKECYSLFTNRRNFRNLNGELYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLTSSMKKTADKDEPLELKEFFGPYGMVVTSTAFSVDIDSLNNP TDPFVANIKKMLKFDLF
NPILFLIAFFPFLGPIMEKLEFSFFPKSVLDFFYALKKIKANRETTKLN RVDFLQLMINSQKLSKVENDK
GLNDHEILSQAMIFLIFAGYETSSSTLTFLAYNLARNPHVMKRLQEEVDATFPNKAPEYQALMQMEYLD SVV
NESLRLYP IAPRLERVAKATVEINGLVIPKGMVVMVPTWPLHRDPSLWPEEAFKPERFSKENKETIDPYTY
MPFGAGPRNCIGMRFALVLMKLAIVEILQTYSSFSVCKETEIPLEMDIQGLISPKRPIKLLVPR
>dla3a116/1-496
MGYFFFSAEWTLLVTLITLFFVYVWYPGTFFKKMGIPGPKPIPFPGTMLAYRKGMNFDSECFQKYKGTW
GIYDGRQPVLCTIDVSMIKTVLKECYSLFTNRRNFRNLNGPLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFGIMKQHSANLTSSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDLDSLNNP SDPFVTNKKMLKFDLF
SPFLFLIVAFFPFLGPVFEKMEFSFFPTYVTDFYALQKIKSNRET SKQKSRVDFLQLMIDSQKLSGVEQDK
GLNDHEILSQAMIFLIFAGYETSSSTLTFLAYNLATNPVAVMKRLQEEIDSTFPNKAPEYQALMQMEYLD SVI
NESLRLYP IAPRLERVAKATVELNGIVIPKDMVVMPTWPLHRDPSLWPEEAFKPERFSKENKETIDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEVLQRYSSFSVCKETEVPFEMDIQGLLMPKRPIKLLVPR
>fhe3a56/1-493
MGYF-YLTAETWTLTLLVAFVTLLLIYAYWYPGTFFKRLGISGPKVPVFFGTMLHYRRGFFTFDEECKKKYKGVW
GIYDGRQPVLCTDPEI IKA VLKVECLSFFTNRRNFRNLNGPLYDALSAIEDDQWRRIRSVLSPSFTSGRLKE
MFEIMKNHSANLTSSMKKKADKDEPLDLKEFFGSYSMDVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDLF
NPLFLAVAFFPFLGPILKFEFSFFPKSVTDFFYASLEKIKSNREASQQKSRVDFLQLMIDSQK--GAQQDK
SLTDHEILSQSMIFIFAGYETSSSTLTFLAYNLATNPEVMKKLQEEIDATFPNKAPEYQALMEMEYLD CVI
NESLRLFP IAA RLERVAKAAVEINGIVIPKDMVVMPTWPLHRDPEI WPEEAFKPERFSKKNKDNIDPYIY
MPFGSGPRNCIGMRFALVLIKLAVVEILQQYSFSVCKETEVPFEMDIQGLLAPKRPIQLKLVP
>fhe3a30/1-493
MGYF-YLTAETWTLTLLVAFVTLLLVYAYWYPGTFFKRLGISGPKVPVFFGTMLHYRRGFFTFDEECKKKYKGVW
GIYDGRQPVLCTDPEI IKA VLKVECLSFFTNRRNFRNLNGPLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFEIMKNHSANLTSSMKKKADKDEPLDLKEFFGSYSMDVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDLF
NPLFLAVAFFPFLGPILKFEFSFFPKSVTDFFYASLEKIKSNREASQQKSRVDFLQLMIDSQK--GAQQDK
SLTDHEILSQSMIFIFAGYETSSSTLTFLAYNLATNPEVMKKLQEEIDATFPNKAPEYQALMEMEYLD CVI
NESLRLFP IAA RLERVAKAAVEINGVIVIPKDMVVMPTWPLHRDPEI WPEEAFKPERFSKKNKDNIDPYIY
MPFGSGPRNCIGMRFALVLIKLAVVEILQQYSFSVCKETEVPFEMDIQGLLAPKRPIQLKLVP
>ola3a40/1-494
MGNFLYFSAETWTLTLLVAFITLLLVYAYWYPGTFFKRLGIPGPKVPVFFGTMLAYRKGFNFDDCECRKKYKGTW
GIYDGRQPVLCTIDPATIKAVLVKECYSLFTNRRNFRNLNGPLYDAVSAIEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLIQSMKKKADKDEPLEMKEYFGPYSDMVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDLF
NPLFLAVAFFPFLGPILKFEFSFFPKSVTDFFYASLEKIKSNREASQQKSRVDFLQLMIDSQK--NGLQDK
GLSDHEILSQAMIFIFAGYETSSSTLTFLAYNLATNPEVMKKLQKEIDATFPNKAPEYQALMEMEYLD CVV
NESLRLYP IAA RLERVAKATVEVNGLVIPKDMVVMPTWPLHRDPELWPEEAFKPERFSKENKDTFDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEILQTYDFSVCKETEVPFEMDVQGLFAPKRPIQLKLTP
>ola3a38/1-494
MGNFLYFSAETWTLTLLVAFITLLLVYAYWYPGTFFKRLGIPGPKVPVFFGTMLAYRKGFPIFDEECKKKYKGTW
GIYDGRQPVLCTIDPATIKAVLVKECYSLFTNRRNFRNLNGPMYDSVFNAEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSANLVNSMKKKADKDEPLNLKDHFGPYSDMVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDLF

NPLFLLVAFPPFLGPIMEKMELSFFPASAFDFFYSALQKIKSDHKNNKQKSRIDFLQLMIDSQK--NGLQDK
GLSDHEILSQAMIFIFAGYETTSSSLTFLAYNLATNPEVMKKLQKEIDATFPNKAPIQYAPLLEMEYLD SVV
NESRLRYP I AARLERVAKATVEVNGLVIPKDMVVMVPTWPIQRDPELWPEPEKFKPERFSKETKETFDPTYT
MPFGAGPRNCIGMRFALVSMKALVEILQQYSFSVCETEIPFELDIQGLIAPKRPIFLKLVRP
>omy3a45/1-496
MSFLPYFSAETWTLLALLITLIVVYGYWPHYGVFTKMGIPGPKPLPYFGTMLEYKKGITNFDTECFQKYGKIW
GIYDGRQPVLCIMDKSI IKT VLIKECYNIFTNRRDFHNGEMFDALS VVEDDAWRRIRSVLSPSFTSGRLKE
MFGIMKQHSANLLNGMKKQADKDQTI EVKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVSNVKKMVKFDLL
NPLFLLVALFPFTGP ILEKMKFSFFPTAVLDFFYASLIKIKSGRDTGNSTNRVDFLQLMIDSQKKTGEEQTK
GLTDHEILSQAMIFIFAGYETSSSTMSFLAYNLATNHHVMTKLQEEIDTVFPNKAPIQYEALMQMDYLD CVL
NGSLRLYPVAPRLERVAKTVEINGIVIPKDCVVLVPTWALHRDPEIWSDP EEFKPERFSKENKEPIDPTYT
MPFGAGPRNCIGMRFAMIMMKLAMVEILQSFTFSVCDETEIPLEIDNQGLLMPKRPIKLRLEPR
>omy3a27/1-496
MSFLPYFSAETWTLLALLITLIVVYGYWPHYGVFTKMGIPGPKPLPYFGTMLEYKKGFTNFDTECFQKYGRIW
GIYDGRQPVLCIMDKSMIKTVLIKECYNIFTNRRNFHNGELFDALSVAEDDTWRRIRSVLSPSFTSGRLKE
MFGIMKQHSSTLLSGMKKQADKDQTI EVKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVSNVKKMLKFDLF
NPLFLLVALFPFTGP ILEKMKFSFFPTAVTDFFYASLAKIKSGRDTGNSTNRVDFLQLMIDSQKKTGEEQTK
GLTDHEILSQAMIFIFAGYETSSSTMSFLAYNLATNHHVMTKLQEEIDTVFPNKAPIQYEALMQMDYLD CVL
NESRLRYP I APRLERVAKTVEINGIVIPKDCIVLVPTWTLHRDPEIWSDP EEFKPERFSKENKESIDPTYT
MPFGAGPRNCIGMRFALIMIKLAMVEILQSFTFSVCDETEIPLEMDNQGLLMPKRPIKLRLEAR
>dre3a65/1-492
----MFSAETWTLLAFLVALLVIYGSWPHRYFKKLGIPGPKTI PFFGTMLRYKDG FHVFDMDCFE KYGRIW
GIYDGRQPVLCIMDQTI IKT VLIKECYSLFTNRRNFRLNGALYDAVSIAEDDDWRRIRSVLSPSFTSGRLKE
MFGIMKHHSHILVDSMGKTAKRGESADIKEFFGAYSM DVVTSTAFSVDIDSLNNPKDPFVTNIKKMLKFDLL
NPLFLLIAFFPFMAPVLEKMDFAFPSTVDFFYAALQKIKSDRDTNTNEERVDFLQLMVDQSOTSDGEHTEK
GLSDHEILSQSMIFIFAGYETSSSTLSFLYNLATNPETMKKLQEEIDETFPNQAPVDYETLMSMDYLD AAL
SESLRLYPVAARLERVCKKTVEINGLLIPKDLVVMVPTVYLHRDPDYWSEPE SFKPERFTKGNKESIDPYMY
MPFGLGRNCIGMRF AQVIMKLAIVEILQRFDVSVCDDETQVPLELGFNGLLSPKDP I KKLQRP
>msa3d1/1-490
MGYLPFLSSETWILLFTFICIFVTYGCWTFGTFEKLGIPGPKPLMYFGTIGRHNPNVYYLADNECAQKYGRVW
GMYELRRPMLAVMDPDM LKT ILVKECFTYFTNRRNIRLNGDLYDAVSTVEDDHWRIRNILSPFTSGHIKE
MFSIMKHHSHKLTASLQSKVHDDEVITIKDIFGPYSIDVMASCAFSVDMDTINNPSSPLITHASKLKF----
-SLFLFQGCFFFLP LLELLGV SFFPKTSTGFFKT VVEKIRAERNKSSQQNSGDVLQHHMNSQQNKEKQN-K
G-TDHEITSQATMFVAGYETSATTLVFLAYNLARNPEVMKHLQEEIESTFPNKGVPQYEALMQMEYLD SVV
HESRLRYPSPRLERIAKETVKISGITIPKDMLVMTPVYALHRDPELWDP EEFKPERFSKQNRQSINPYTY
LPFGIGPRNCLGMRFAVMIKLALVEVLQNY SFSVCEETEFLKMDPLSLVGPLQPIK LKLVR
>gac3d1/1-491
MGCLPFFSVETWILLITFICIFVIYGNWTHGIFENLGIPGPKPYTYWGTIGRHNKVYYKDDQOCAEKYGVW
GMYEFRSPMLAVMDPDM LKT ILVKQCFTYFTNRRNIRLNGDLYDSVSTVEDDHWRIRNVLTPSFTSGHIKE
MFKIMKHLSSKLTASLQSKAQNEVVTIKD YFGAYSVDVMASCAFSVDMDSINNHSSPIIFHASRFFKF----
-FVLILQGIFPVLIP LPELLGISFFSPTSTAFFKTILAKIAREGSTHKNSGDFLQSMINRTSKGEKN-K
GLTDHEILTQATMFVAGYETSATTLAFLAYNLARNPEVMKRLQEEIDSTFPNKGAVRYEELLQMEYLD SVV
NESMRLYPPAGRLDRVAKTVEIRGITIPKDMIVMIPVYALHRDPELWDP EEFKPD RFSKQNKRSINPYTY
LPFGTGRNCLGMRFAVLVLKALVEVLQNY SFSVCETEIPLTMDPGGLVGPLNPIK LKVTR
>tru3d1/1-483
MNYLPFALETWILLITFTCLFVMYGKRTFGIFEKLGISGPKPTMYFGSICKYNNVYYLDDQECAQKYGKIW
GTYELRKPM LVMDPDL LKT ILVKECFTHTNRRNFCLNGDLYDAVNIAEDDDWRRIRNVLSP LFTTGRIKQ
IFSLMKQSSKLTSSLEPKAENEEIISIKDFFGAYNMNVATGVLF GME--P SLP-IK HASKLKF----
-PLFMIQGASPRCFPIQLMGFSLFPRDSFAFVKKIVEKIRAERDGGSHQE--DFLQHMI-STQGIYAVT-S
GLTDHEIVSQLT VLLTGGYETSTLALTL SIYSLATNPGSMNRLQEEIDATFPDAAPVTYEALMQMEYVDCVI
NECLRLYP AARLERTAKETVEISGITIPKNMTVTPIFALHRDPEHWPEPEEFKPD RFSKQNKGRINPYTY
LPFGIGPRKCLGMRLALVIVKALVETLQKYSFSVCETEIPFKMDPNTFVGP INPIK LRVVR
>tru3a49/1-496
MDLLPNFSIETWTLLIALVITLITVYGYAPYGF FKKVGISGPKPWPFIGTFLNYKKGVH HFD EECYK KYGVW
GLYDGRQPLMCIMDTGI IKT ILVKECYSNFTNRRDLGLNGPLRDAVSVAEDEQWKIRIGILSPSFTSGRLKE
MYTIMLQHSKNLLNFLHKKVEADEVIDVKDVG PYSMDVVTSTAFSADIDSINHP SDPFVANIKKMVKFNFL
NPLLIFFVVLFPFTQPIFDKVDFSFFPAHVIDFFYNFLRQIKSDRNKDKKSRVDFMQLMVNAQMNEEGSSQK
GLTDNEILAQAMIFIFAGYETTSSSLGFLAYNLATNP KIQKQLQEEIDKTFPGKVRPNYDDLMQLEYLDMV
NESMRLYPIANRLERMTKTSVEINGLTIPKGTVAIPVYALQRDPVLWPEPEAFKPERFSKENKDNVDPYAY
LPFGAGPRNCIGNRFALVLMKLAIAEILQHYSFVTCPETDIPMVLGNEG FVTPKNPIK LKLKL
>tru3a48/1-493
MDLLPNFSIETWTLLT LIFTLIIVYGYAPYGF FKKVGIPGPKPWPFIGTFLNYRRGIH HFD EECYK KYGVW
GLYDGRQPLMCIMDTGMIKT VLVKECYSNFTNRRDLGVNGPLSDAVSVAEDEQWKIRIGILSPSFTSGRLKE
MYTIMLQHSKNLLNFLNKKVEADEVIDVKDVG PYSMDVVTSTAFSVDIDSINNP SDPFVANIKEMTQFSFL
NPL----VVLFPFLVPIFKKMNVTFFPAHVIDFFFNFLRQIKSDRNKDKKSRVDFMQLMVNAQMNEEGSSQK
GLTDNEILAQAMIFIFGGYETTSSSLGFLAYNLATNP KIQKQLQEEIDKTFPGKVRPNYDDLMQLEYLDMV
NESMRVFPILSRLERMTKTSVEINGFTIPKGTVAIPVYVLQHDKAYWPEPEAFKPERFSKENKDNVDPYAY
LPFGAGPRNCIGNRFALVLMKLAIAEILQHYSFVPCKETDIPMVLNTEGLVAPKNPIK LKLKRP

>tni3a48/1-496

MDLLPDFSIETWTILIVLFTLITVYGYAPYGF FKNLGIHGPKPWPF IGTFNLNRYKGVHHFDQECYKKYGKVW
SLFDGRQPLMCIMDTLMIKTILIKECYSNFTNRRDLGLNGPLRDAVSITEDEQWKIRRSILCPSFTSGRLKE
MYTIMLQHPQNLINLLQKKVKADDVIDVKDVGYPYSMDVVTSTAFSVDIDSINHPSDPFVANIKKMKVFNFM
NPLLILVVLFPFTQTIFDKMDVSLFPGHVLDF FYNFLRQIKSDRSKEKKKSRVDFMQMLMVNAHMEKSSSLK
GLTDKEILAQAMIFIFAGYETTSSSLGFLAYSLATNPDIQKKLQEEIDKTFPGKVRPNYDDLMQLEYLDMVV
NESMRLYPIANRLERMAKTSVEINGVTIPKGTVAIPVYALHRDPALWPEPEAFKPERFSKENSANIDPYTY
LPFGAGPRNCIGMRFALVVMKLAIAEILQHYSFVPCKETIIPMVLGSEGFTPKTPIKCLKKPR

>tni3a49/1-496

MDLLPDFSIETWTILIVLFTLITVYGYAPYGF FKNLGIHGPKPWPF IGTFNLNRYKGVHHFDQECYKKYGKVW
SLFDGRQPLMCIMDTLMIKTILIKECYSNFTNRRDLGLNGPLRDAVSITEDEQWKIRRSILCPSFTSGRLKE
MYTIMLQHPQNLINLLQKKVKADDVIDVKDVGYPYSMDVVTSTAFSVDIDSINHPSDPFVANIKKMKVFNFM
NPLLILVVLFPFTQTIFDKMDVSLFPGHVLDF FYNFLRQIKSDRSKEKKKSRVDFMQMLMVNAHMEKSSSLK
GLTDKEILAQAMIFIFAGYETTSSSLGFLAYSLATNPDIQKKLQEEIDKTFPGKVRPNYDDLMQLEYLDMVV
NESMRLYPIANRLERMAKTSVEINGVTIPKGTVAIPVYALHRDPALWPEPEAFKPERFSKENSANIDPYTY
LPFGAGPRNCIGMRFALVVMKLAIAEILQHYSFVPCKETIIPMVLGSEGFTPKTPIKCLKKPR

>gac3a48/1-496

MGFLPELSIQWTILIVLFTLITVYGYAPYGF FKNLGIHGPKPWPF IGTFLEYRKGHNFDECEYKKYGKVW
GIYDGRVPIVLTMDTAMIKSILVKECYSNFTNRRDFGLNGALRDAITVVEDEEWKIRRSILSPSFTSGRLKE
MYSIMLQHSGLIKSLHKKAEAGEVVDVKDIFGPYSMDVVASTAFSVDIDSINNP SDPFVANIRKLVKFTFL
TPLFLLLVLPFPLEPMLDKMNVTLFPAELLDFFYSFLKTIKSDRSKNEHKNRVDIMQLMVD SQMKEDPSAHK
GLSDSEILSQAMIFIFAGYETTSSGTLGFLAYNLATNPDIQKTLQEEIDETFPNKARPTYDALMQMEYLDMLV
NESMRLYPSGNRLDRVSKSSVEIEGVTIPKGTVMIPVYTLHREPTLWPEPDDFKPERFSKENKDSIDPYSY
LPFGAGPRNCIGMRFALLMMKLSVVEILQNF SVTCKETQIPLLELSTDGFTPKDSINLKLEPR

>ola3b5/1-486

--MLFLFSATTWTLLALIFTILLFGIWPFRIFQNMGISGPMMPYVGNLLQQGKGFFEFDEHCQTKYGDVW
GVFEGRPILMVADPEI IKTVMVKECYSFTNRRNFEGDGLTEAVASAKDEKWKIRRSSLSPCFTSGRLKQ
VFSIVSRVADRLVEKLGET-YLNKAIDIRQFVAPYTLDDVVASASFVDPDCINNPDDPINVQAQKATRFNFL
--PTILQSIIPFGGRLLKLFKIETLPSANVEFFIDIIRNLKDQHKTE---NRPDFLQVLI ESEMIKQQQPSK
GLTETEILSQALGFI VGGYDTTSTALSIFYCLATNTDAMHTLQKEIDSNLQKNSSISYEDLNGLEYLDQVI
NESRLYPVAPRLDRECKMTVETQGFTVPEGMIVGIPVYLLHKDPRFWSSPELFRPERFSKENEGELNPYAY
MPFGLGRNCVGMRFVAVLMMKMIIVRLLRYSVETCKYTLIPMQFNW--FQPIKPVKLFVPR

>ola3b6/1-485

-----MFSATTWTLLALFFT VLLLYGVWPYRVFRKLGISGPMPLPFVGNVLQMGRDILSFRDQCLDKYGDVW
GVFEGRPMLMVANPELVKAVMKECYSFTNRRNFEPDGLADGVAVVDEKWKIRRSSLTPCFTSGRLKQ
VYPLVARYAENLVKKLGRM-NSDEPIDVRQFVAPYSMDVVASTSFSVDTD C INNPDDPINVQAQKATNFNFW
--PVILQTMIPFGDLLLKLKFETLPRSNVDFFYDLIKRLRDQHKAG-KLTRPDFLQVLIDSEIVKNEQPSK
GLTEHEVLSQAFGVFLAGYETMSNTLCF ILYNLATNPDMACTLQKEIDSNLQNDGFSFYEELSGLEYLDQVL
CESMRFFSPTPRLDRECKKTIEIHGVTIPKGTILVAIPVSLHMDPRFWSSPEIFRPERFSKENASELNPYAY
MPFGLGRNCVGMRFVAVLMMKMIIVRLLRYSLETCKDTLIPMQYDWQ--FQPIKPIKLFVPR

>ola3b4/1-489

-MFWSWFGATWTLLALLLTI LIFYGLWPYRVFKNLGISGPMPLPFVGTVPQQMKGLFAFNRECQAKYGDVW
GLYEGRSPLLMIADPKI IKTVLVKECYSTFTNRRNFDDDGPLNGVA AVKDERWKIRRSITISPCFTSGRLKQ
VFPFVARYADRLVNKLQKM-NLNEPIDIKQFVAPYSLDVVTSASFVETDCINNPDPVNVQAIAKATNFNLW
--PMIIRMIFFLGLKHFKEIIPRSVDF FCDVIKKIKDQHNAG-KLTRPDFLQVMIESEIVKNEQPSK
GLTEQEILFQAFSFI LGGYETTSTTLSHIFYNLAINPDAMHTLQEEIDANLQKDVPSYDKLNGLEYLDQVL
CESRLRIPPAPLLDRVCKTMQINGLTIP EGTVVGIPVHLLHRDPRFWDSPDLFRPDRFSKENEGALNPYAY
MPFGLGRNCVGMRYAILVMKMIIVRLLRYSLETCKDTLIPMQFNW--FQPLKPIKLFVPR

>ola3b3/1-488

-MFFTLLSPTTWTLLALLFALLWLYGTWPHSVFKNLNI PGPKLPFVGTIFGVRKGLFAFDSECCQAKYGDVW
GLYDGRSPVLMVADPEMIKTIMVKECYSVFTNRRRELI-PGPFEDGITVVKDDRWKIRRSITISPCFTSGRLKQ
VFPVIVASYADRLKKLEQT-NLDEPIHVQFVAPYSLD TVASASFVEMDSINNSDPVNVHIIQKLFNINFW
--PFILLMICPFAGRLNLSNIGMLPKAESDFFLNIKKFKDQHKAG-ESTRQDFLQVMVQSEI IKQE QPSK
GLTEPEIVAQAFTFILGGYETTSTTSLF ILYNLATNP HAMHTLQKEIDANLQKNASFYEELNDLHYLDQVI
LES MRVLPTAPRLRGCKKTVEIQGFTIPEGTVVGIPVNLHKDPRYWSSPELFRPERFSKENEGELNPYAY
MPFGLGRNCVGMRYAILVMKMIIVRLLRYSLETCKDTLIPMQFDWK--FQPLKPIKLFVPR

>gac3b7/1-489

MFLDFWFPATSWTLLALFFTLLILYGIWPYRFFKNLGIKGRPLPF IGTFHLFRKGWLSFDRECCQAKYGDVW
GNFDARLPVLMVADPEI IKTVMVKECYSVFTNRRNAVPGPMGDAITAVKDERWKIRRTSLSPCFTSGRMKQ
VFPVIVARYADRLTEKLGM-H-SVPIDVKQFMAPYSLDVVTSASFVANSINNPDDPLNVHLKNI LNFKLW
--PVFILMVFPFGARLLRLKLEMLDRVSVDYFYDVIKHFKHQHGED-ESTRADFLQVMIQNEI IKEDQPSK
GLTEHEILSQALIFIFGGYETTSTTLLF LLYNLATNPDMARTLQEEIDANIPRDAPIRYEDLLGLQYLDQVL
LESQRILPTAPRLRVCCKTTQVHGLTIPKDTLVGIPVQLLHMDPRFWSSPELFRPERFSKESGEEVNPYAY
MPFGLGRNCVGMRYAVLTMKTVLVRLLQSYNVETCKDTMIPLEFDWK--AQPVQPVKLFVPR

>tru3b2/1-488

MFVFMCFSAATTWTILVLFSTLLLYGHWPYRLFRKNGIPGPMPLPF IGTMWNLLKGNMVFDRCCSKYGDVW
GVFEGRTPLMVSDPGMIKTILVKECYSVFTNHRVRIFSGPLETSVFS AKDETWKRMRTISIPCFTSGRLRQ
AFPIIARYADRFIAKLEQT-KLEDSTDIKKLFGPYSLDVIASSFSVDADSINNPDDPFIINKKVLNLFNFW

--LLLIKNVLPFSSYLFEFLHIDIIPRSIVDYFFNLIKQLKAQH--D-QSIKGDFLHVMIQNEI IKQDQPPK
GLTEQEILSQSVLFI FG GYDTTITITITYLLYNLA INPDVLQILHSEIDS NF PKDTPFSYEDLVGFQYLDQVL
NESQRLIPTAPALERFCKKTQVIHGLTIPEGTVVAVPVHLLHKDPRFWSSPEEFRPERFSKDNIEEVNPYAF
MPFGLGPRNCIGMRYAILVMKMILVRVLQSYTVETCKDTMIPLEFDWK--LQPTKQIKLRLVPR
>tni3b2/1-485
MLDFTTFSATTWTILVLF TLLVYGTWPYRIFKKNIGPGMPLPFI GTMLYSINKNLI F DRECAKYGDVW
GTFEAMPVLMVTNPEMIKTIMVKECYS AFTNH-TDFFSGPLENSV FVAKDEKWKMRASISPCFSSGRIKQ
AFPIIARYADR FIAKLSQ--NLEEATDIKQLFGPYSLDVVTSSSFSDADSINNPDDPFILNVKKILQFKFW
--VLLIKSMFPFSSYL FKLGMELTPRP IVDYFMSVIKRFKDQHKAD-ESI QGDFLHVMIQNEI IKQDQPSK
GLTEHEILSQSILFI FG GYDTTTSV SITNALYNLA INPDALQTLHKEIDTCLPKDAPISYEDLMGLQYLDQVL
NESQRLMPTAIRLERVCKKT VQVEGV T I PKGATVVIPLLLLHKDPRFWSSPEEFRPERFSK---EEVNPYAF
MPFGLGPRNCVGMRYALLVMKMVIARVLQIYTVETCKDTMIPMELDWK--FQPVKPIKLSLVPR
>tni3b1/1-490
MLDLWGLSGTTWTLLALFFALFFLYGIWPNHFKKLGI PGPSLPF IGSLFYRKGLVPFERWCQEYQYDGIW
GIFDGRLPILMVSDPEMVKT VLVKECYS AFTNRRDSLFGSLDDGVT VVKDERWKRI RSTMSPCFTSGRLKQ
AFPIVARYADRVT K KLEQS-DLDKPIDVKEFLAPYSLDTVTSASF SVEADAINNPNDPIIVNKKVLKFN FV
--VLFLIAFFPFCRLLSFLGIDPLPRASVNYFYNLIKKLKDEH HAD-ESVRGDFLQVLIQSEI IKEDQPPK
GLTEHEILSQAFIFIFGGYETTTTTLTNLLYCLAVNPDALETLHREIDANTSGDAPISYEDLMGLQYLDQVL
NEAQRLYPTAPRLERMCKKT VQVQGV T IPEGT VVGIPSHLLHKDPRYWSSPEEFRPERFSKENAEVNPYAF
VPFGLGPRNCVGMRYALLVMKMVI VRLQSYTVKTC KDTMIPLEMDWK--YQPLKPIKLSFVPR
>tru3b1/1-488
MFEFVLFSGTTWALLALFFALLLLYGVWPHYHFKKLGI RPRPLPFMGSTFYRKGIIPFESWCQAEYGDVW
GMFEGRTPLVMVSDPEILKT VLVKECYS VFTNRRDSGFAGPLEDSVSAVKDERWKRI RSTISPCFTSGRLKN
AFPIVARYADRIT K KLEQS-NLDEP INVKEFLAPYSLDAVTSVSFSVEADSINNPNDPLIVNLKKVFKFN FV
--VFFLVAFFPFCARLFQFLGIDLIPRSVNYFYNVIKNFKDQHHAD---TRGDFLQVLIQSEI IKQDQPPK
GLTEHEILSQAFIFIFGGYETTTTTLTNVLYGLA INPDVLQVLHKEIDTNIPSDAPISYEDLMGLQYLDQVL
NESQRLYPTAPRLERACKKT VQIHGLTI LEGTIVGIPVHLLHKDPRFWSSPEEFRPERFSK DSTEEVNPYAF
MPFGLGPRNCVGMRYAILVMKMLI VRLQSYTVETCKDTMIPLEFDWK--SQPLKPIKLSFIPR
>dre3c3/1-490
MIDLSSLS-VTWTLVVLVITLLYMYGVWPHGFFKKLGI PGPRPWPFFGTFLSYTKGFFNFDMECAKKYGVW
GIYDGRLPILMVTDLEMIKTIMVKECYS TFTNRREVDLAGPLADGISAVKDERWKRI RGTLSPYFTSGRLKE
IFPIAVTHADR FIKNMQKR-DHEQPVKI KEVVAPYSLDVVTSSSFSDIDSINNPDDPFVNIKNFLKFSLF
SPLILFLTLP SIANLLGKMGISVFSRSIMNFY TALRKIKDEH-KD-SNGRVDFLRLMIQNQIAKSEQPVK
GLTDHEILSQSFIFILGGYETTTSTLSFLHNLATNPDCLEKLVEEIDKNFPLDTPISYDALMKMDYLEMSI
NESMRLLPTAPRLERVCKKTVELNGITIPKDTLVAIPTVYVLRNDPQLWDS PQEF RPERFSPENKSEFLQYAF
MPFGLGPRNCIGMR FALMIVKLLVVKLLQNF SLETCKETQIPELTPV--FQPKVPITLKLTPR
>dre3c2/1-490
MIDLSSLS-VTWTLVVLVITLLYMYGVWPHGFFKKLGI PGPRPWPFFGTFLSYTKGFFNFDMECAKKYGVW
GIYDGRLPILMVTDLEMIKTIMVKECYS TFTNRREVDLAGPLADGITAVKDERWKRI RGTLSPYFTSGRLKE
IFPIAMTHADR FIKNMQKG-DHEQPVKI KEVMASYSLDVVTSSSAFSDIDSINNPDDPFVNIKNFLKFNLF
GLFVLLLTVP SIANLLGKMGLSMFSKPTVDFFYKALRKIKEEH-KD-SNGRVDFRLMIQNEIVKSEQPEK
GLTDHEILSQSFIFIMGGYETTTSTLTFLLYNLATNPDCLEKLVEEIDKNFPLDTPISYDALMKIDYLEMSI
NESMRLLPTAPRLERVCKKT VEMNGITIPKDTMITIPTVYVLRNDPQLWDS PHEFRPERFSPENKSEFLQYAF
MPFGLGPRNCIGMRYALMIVKLLVVKLLQNF SVETCKETQIPELNPV--FQPKVPITLKLTPR
>dre3c4/1-491
MIDLSSLS-VTWTLVVLVITLLFIYGVWPHRFFIKLGI PGPRPWPFFVGTLLSYAKGLCNFDMECAKKYGVW
GIYDGRLPILMVTDLEVIKVVLVKECYSNF INRRKTGLAGPFDGIIMVKDERWKRM RSTLSPYFTSGRLKE
IFPIAMTHADR FIKNMQKK-DPEQPVKI KEVVAPYSLDVITSSSFSDIDSINNPEDPFVINVKKFIQFNMF
NPLLLLILFP SLAIP LKKMGVTLF S K SAMDFFYSVLKKIKDDHNKE-SEGRVDFLKLMLQNQIIQSEQPAK
GLTDHEILSQSLVFI LGGYETTTSTLTFLLYNLATNPDCLEKLVEEIDKNFPLDTPITYDALMRMDYLEMAI
NESMRLLPTAPRLERSAKKT VVINGLT IPEGT LVGIPTVYVLSHDPDIWESPGEFRPERFSPENKSEFLQYAF
IPFGLGPRNCIGMRYAVMIMKLFVVKLLQNF SVETCKETQIPELINV T--FQPKVPITLKLIPR
>dre3c1/1-490
MFDLSSLS-VTWTLVVLVITLLLIYGVWPHGFFKKLGI PGPRPLPFVGTALSYSKGICNF DIESCKKYGVW
GIYDGRLP LLLVTDLEMIKTILVKDCYSTFTNRRNPDLVGPFADGITLVKDERWRRIRSSLSPYFTSGRLKE
IFPIAMTHADR FIKNMEKK-DPNLPLKIKDVVAPYSLDV VASSSFSDIDSINNPDDPFVTSIKSFFNINPL
SPLFLLLAFCPSAANLLAKMGISVFSRSTDFYKALRKIKDEH-NE-SNGRVDFLKLMIQNQIVKSEQPVK
GLTDHEILSQSFIFILGGYETTTSTLSYLLYNLATNPDCLEKLVEEIDKNFPLDIPITYDALMRMDYLEMAI
HESMRVFPAGRLERVCKKTVEINGITIPKNTLVGIPLYVLSRDPDLWESPNEFKPERFSPESKTEINQCAF
MPFGLGPRNCIGMR FALMMKLLVVKLLQKYTVETCKETQIPVQLNFF--FQPKVPITLKLIPR
>ami3a77/1-452
-----PFFGTALSYRNGFMDFDNECFQKYGNLW
GIYDGRQPVLATVDPTVIKT VLVKECYTTFTNRRSFGPTGVLES AISIAEDEEWKRIRTVLSPFTTSGKLKE
MFPVMKHGDI LVRNVRKQVEKNDVIGVKEVF GAYSMDVVTSTSGFVN TDSMNPKDPFVKEAKKLKFDFF
SPVFVLLYVLPCLNPVMKKG VNI FPKDALEFFMQSVTKI KEARKEEARADRVDFLQLMIDSQRSNNGHSSK
VLTDIEILAQAII FIFAGYETISSTLRYLVYELATHPDVQKQLQEEVDRVLLNKDSLTYDALMQMDYLDMTL
NETLRFPPFGGRLERVCKKDVEVNGVSI LKGMVVMVPLYVLHNNPEYWPEPEFRPERFSKENKESMDPYLY
LPFGAGPRNCIGMR FALLAMKIAVARLVQNFTFKTCKETPIPLKLD SKGFLNPVKPIILKL VPR

>bac3a72/1-496
MDLIPSFVETWVLLATSLVLLYLYGTYSHGLFKKLGIPGPRPLPYFGNLSYRKGLWEFDNRCFKKYGKMW
GCYEGQQPPLLITDPIIKTVLVKECYSVFTNRRAFGPGVGMKNAISLSEDEQWKIRITLLSPAFTSGKLKE
MFPIISQIGDVLVRNLKKEAEKGKPISMKDFFGAYSMDVITGIAFGVNIDSLNNPDPPFVEYSKKILKFNLL
SPLLLSIVLPFLTPVFEALNITLFPKSLNFLT KSVKRIKESRLKDKQTHRVDLQLMINSQNSKETDTHK
VLSDELMAQSIIFIFAGYETTSNSLSFIIYELATHPDVQQLQEEIDATFPNKAPPTYEALVQMEYLDMLV
NETLRIFFIAGRLEVRCKKDVEVNGVFIPKGT TVVVPLFVLHNDPELWPEPQEFRRPERFSKKNKDSINPYTY
LPFGTGPRNCIGMRVIMNMKLALVRILQNFSEFKPCKETQIPLKLLTQGLTQVEKPVVLKVVSR
>cae3a4/1-496
MDLIPDLAVETWLLAVTLVLLYLYGTHSHGLFKKLGIPGPTPLPLLGNVLSYRKGFQWQFDMECYKKYGKVW
GFYDGRQPVLAITDPNMIKT VLVKECYSVFTNRRPFGPGVGMKSAISIAEDEEWKRIRSLSPFTTSGKLKE
MVPPIIAQYGDVLVRNLREAEETGKPVTLKDFGAYSMDVITSTSGVNVDSLNNPDPPFVENTKLLRFDFL
DPFFLSITIFPFLIPILEVLNISIFPRKVTSLRKS VKRIKESRLKDTQKHRVDLQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSVLSFIMYELATHPDVQQLQEEIDTVLPNKAPPTYDTVLQMEYLDMLV
NETLRIFFIAMRLERVCKKDVEINGIFIPKGVVVMIPSYALHHDPKYWAEPEKFLPERFSKKNNDNIDPYIY
TPFGSGPRNCIGMRFALNMMLKALIRVLQNFSEFKPCKETQIPLKLRGLGLQTEKPIVLKIESR
>cae3a7/1-496
MALIPDLAMETWLLAVSLILLYLYGTHSHGLFKKLGIPGPTPLPFFGNVLSFHKGLWTFDMECYKKYRKVW
GFYDCLQPVLAITDPNMIKT VLVKECYSVFTNRRPLGPGVGMKSALSRAQDEEWKRIRSLSPFTTSGKLKE
MVPPIIAQYGDVLVRNLREAEETGKPVTLKDFGAYSMDVITSTSGVNVDSLNNPDPPFVENTKLLRFNSL
DPFFLSIKIFPFLTPIFEALNITMFSRKVINFLRKS VNKIKEDRLKETQKRRVDLQLMIDSQNSKDTESHK
ALSDLELVAQSIIFIFAGYETTSVLSFIMYELATHPDVQQLQKEIDAVLPNKAPPTYDTVLQMEYLHMLV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYVLHHDPKHWTEPEKFLPERFSKKNKDNIDPYIY
TPFGSGPRNCIGMRFALNMMLKALIRVLQNFSEFKPCKETQIPLKLRFGGLLQPEKPIVLTAESR
>cja3a90/1-496
MDLIPNLAVETWLLALLSLVLLYLYGTRSHGIFKKLGIPGPAPLPFVGNVLSYRQGIWKFDSECHKKYGKMW
GSYDGLQPLVLSITDPMIKAVLVKECYSIFTNRRPLGPGVGMKSAITVAQDEWKIRIRSLSPFTTSGKLKE
MFPIIAQYGDVLVRNLREAEKGKLVTKDIFGAYSMDVITGTSGVNVDSLNNPKDPFVENVKFLKFDVL
APLFLSVVLPFLIPVFEALNFSLPKDAINF LKQSVNRMKKSRLNDKQKHRVDLQLMIDSQNSNETASHK
ALSDLELVAQSIIFIFAGYETTSVLSFTIYELATNPVQQLQEEIDVLPNKAPATYDAVVQMEYLDMLV
NETLRLYPITVRHERVCKKDVEINGVFIPKGA LVIPITYALHHDPKYWTEPEKFRPERFSKKNKDSIDPYIY
TPFGTGPRNCIGMRFALNMMLKALIRVLQNFSEFKPCKETQIPLKLG VQGLQAEKPIVLKVESR
>cid3a135/1-496
MSYDLFFSAETWALLVLVLF GALLVIYGSWTHSIFKKLGIAGPKPIPFPGTMLRYREGFHNFDLECFKKYGRVW
GIYDARQPVLCIMDQSIITKILIKECYSLFTNRRNFRNLNGPLYDAVSIVEDDDWRRIRSVLSPFTTSGRLKE
MFGIMKAH SKILVENLGKSATRGENVDIKEFFGAYSMDVVTSTAFSVDIDSLNNPKDPFVTNKKMLKSDFL
NPLFLISALFPFVPLEKMDFAFPSTVDFFYAALQIKIKSERVANDHKKRDFLQLMVD SQTAGKTQHDE
ELSDHEILSQSMIFIFAGYETSSSTLMFFFYNLATNPETMKKLQEEIDETFPDKAPVDYEA VMNMDYDLAAL
NESLRLYPVAARLERVCKKTVDINGLIIPKDMVVMVPTFVLHRDPDYWSDPESFKPERFTKGNKELIDPYMY
MPFGLGPRNCIGMRF AQVTMKLAIVEILQRFDVSVCEETQVPLELGLNGLLAPKDPIKLRKPR
>cgi3a136/1-496
MSYGLFFSAETWALLILFVTLLFIYGSWPHGVFKKLGP GPKPLPFFGTMLRYRKGFHNFDMECFKKNGRVW
GIYDARQPVLCIMDLSIITKILVKECYSLFTNRRNFRNLNGPLYDAVSIVEDDDWRRIRSVLSPFTTSGRLKE
MFGIMKTHSHTLVDNLGKTATRGEAVEIKEFFGAYSMDVVTSTAFSVDIDSLNNPKDPFVTNKKMLKFDLF
NPVFLISAVFPFITPVLKMGFAFPSTVDFFYAALQIKIKSERVSSEQKRVDFLQLMVD SQTASEEHTEK
GLSDHEILSQSMIFIFGGYETSSSTLSFFFYNLATHPEAMEKLQGEIDQSF SREDPVDYEGIMNMEYLDAAL
NESLRLYP IVARLERVCKKTVDINGLLVPKDVVMIPTFALHRDPDYWSEPDSFRPQRF SKDNRESIDPYMF
MPFGLGPRNCIGMRF AQVSIKLAIVEILQRFDVSVCEETQVPLELDTSGLLAPKSP IKLQFKPR
>dla3a79/1-493
MGYFFYFSAETWLLVTLITLFFVYVWPYGT FKKMGIPGPKPIPFPGTMLAYRKGFNMFDSECFQKYGKTW
GIYDGRQPVLCITDVSMIKTVL IKECYSLFTNRRNFRNLNGPLYDAVSIAEDDQWRRIRSVLSPFTTSGRLKE
MFGIMKQHSANLISSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDLDSLNNPDPPFVTNKKMLKFDLF
SPLFLIVAFFPFGMPVF EKMEFSFFPTYVTDFFYAALQIKIKSNRET SKQKSRVDLQLMIDSQKNNDL---S
GLNDHEILSQAMIFIFAGYETSSSLTFLAYNLATNP AVMKRLQGEIDSTFPNKAPIEYQALMQMEYLD SVI
NESLRLYP IAPRLERVA KATVELNGIVIPKDMVVMIP TWPLHRDPDLWPEPEEFKPERFSKENKETIDPYTY
MPFAGPRNCIGMRFALVMMLKLAIVEVLQRYSF SVCKETEVPFEMDIQGLLMPKRPIKLEPR
>fca3a131/1-496
MDLIPSFFTETWLLLATSLVLLYLYGTHSHGLFKKLGIPGPKPLPFLGTALGYRQGFCEFEKCFRTYGRMW
GFYDRRQPVLAITDPMIKTVLVKECYSVFTNRRSFGPGVGMKG AISLSEDEQWKIRITVLSPTFTSGKLKE
MFPIVQYGDVLVRNLKKEAEKGKPVNLKDILGAYSMDVITSTSGVNVDSLNNPRDPFVENTKLLKFSFL
DPFFFLSLLFPFLTPIFDLLNIWLFPKSVINFFT KSVKRMKESRLKDKQKHRVDLLQLMINSQNSKETDTHK
ALSDLELVAQSVIFIFAGYEPTSTSLSLVYELATHPDVQQLQEEIDATFPNKAPPTYDALVQMEYLDMLV
NETLRLYP IAGRLEVRCKRDVEISGVFIPKGT VVMVPTFTLHRDLDLWPEPEEFHPERFSKKNKDSINPYIY
LPFGTGPRNCIGMRFVAVNMMLKALVTLLQNF SFQPCETQIPLKLNAQGIIQPEKPIVLKVELR
>fca3a132/1-496
MDLIPSFSEAWLLLATSLVLLYLYGTYTHGLFKKLGIPGPKPLPFGVTALGYRQGFCAFDEKCFKTYGRMW
GFYDGRQPVLAITDPMIKTVLVKECYSVFTNRRSLGPMGMKSAISLSEDEQWKIRITMLSPFTTSGKLKE
MFPIVQYGDVLVRNLKKEAEKGKPVNLKDIFGAYSMDVITGTSGVNVDSLNNPDPPFVENTKNLKF SFL

DPLFFSIIILFPFLTPIFELLNIWIFPKKVTDFFTKSVKSMKESRLKDKQKHRVDLLQLMINSQNSKETDTHK
ALSDLELVAQSIMFIFGGYETTSTSLSFVLYDLATHPDVQQLQEEIDATFPDKAPPTYDALVQMEYLDMMV
NETLRLYP IAGRLERVCKRDVEISGVFIPKGTVMVMTFTLHRDLDLWPEPEEFCPERFSKKNKDSINPYIY
LPFGSGPRNCIGMRFAIMNMKLALVRVLQNFSEFKPCKETQIPLKLNLSIEPEKPIVLKVELR
>mfa3a4/1-496
MDLIPDLAVETWLLAVTLVLLYLYGTHSHGLFKKLGIPGPTPLPLGNILSYRKGFWTFDMECYKKGKVV
GFYDGRQPVLAITDPNMIKTVLVKECYSVFTNRRPFGPVGMKNAISIAEDEEWKRIRSLSPFTTSGKLKE
MVPPIIAKYGDVLRNLRREAETGKPVTLKDVFGAYSMDVITSTSFVGNIDSLNNPQDPFVENTKKLLRFDFL
DPFFLSITIFPFIIPILEVLNISIFPREVTSFLRKS VKRIKESRLKDTQKHRVDLQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIIYELATHPDVQQLQEEIDTVLPNKAPPTYDTVLQMEYLDMMV
NETLRIFFIAMRLERVCKKDVEINGVIFIPKGVVVMIPSYALHHDPKYWPEPEKFLPERFSKKNNDNIDPYIY
TPFGSGPRNCIGMRFAIMNMKLAIIRVLQNFSEFKPCKETQIPLKLRGLLQTEKPIVLKIESR
>mfa3a5/1-496
MDLIPNLAMETWLLAVSLVLLYLYGTRSYGLFKRQGIPGPTPLPFLGNILSYRQGLWKFDTECYKKGKMM
RTQDGGQPLVLTITDPDMIKTVLVKECYSVFTNRRPLGPVGLMKS AISIAEDEEWKRIRSLSPFTTSGKLKE
MFPIIAQYGDMLVRNLRREAEGKGPVTLKDIFGAYSMDVITSTSFVGNIDSLNNPKDPFVESVKKFLKDFL
DPLFLTLIFPFLIPAFEALNVSLFPKDAINF LNKS VNMSMKSR LNDKQKHRVDLQLMIDSQNSKETESHK
ALSDQELVAQSIIFIFAGYETTSSVLSFIIYELATHPDVQQLQKEIDAVLPNKAPATYDAMVQMEYLDMMV
NETLRLFP IAIRLERACKKDVEINGVIFIPKGVVVMIPSYALHHDPKYWTEPEEFRPERFSKKNKDSIDPYIY
TPFGSGPRNCIGMRFAIMNMKLAI IKVLQNFSEFKPCKETQIPLKLGNOGLLQSEKPIVLKVESR
>mfa3a7/1-496
MDLIPDLAVETWLP LAVSLI LLYLYGTHSHGLFKKLGIPGPTPLPFFGNVLSFRKGFWTFDMECYKKGKVV
GFYDCRQPVLAITDPNMIKTVLVKECYSVFTNRRPLGPVGMKS AITVAEDEEWKRIRSLSPFTTSGKLKE
MVPPIIAQYGDVLRNLRREAETGKPVTLKDVFGAYSMDVITSTSFVGNVDSLNNPQDPFVENTKKLLRFNSL
DPFFLSIKIFPFLTPIFEALNITVFPKVTISFLRKS VKRIKEGR LKETQKRRVDLQLMIDSQNSKDTASHK
VLSDELMQAQSIMFIFAGYETTSSVLSFIIYELATHPDVQQLQKEIDTVLPNKAPPTYDTMLQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMSIPKGVVVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIDPYIY
TPFGNGPRNCIGMRFAIMNMKLALIRVLQNFSEFKPCKETQIPLKLSFGGLLKTEKPIVLTAESR
>mfa3a43/1-496
MDLIPNFAMETWVLVATSLVLLYLYGTHSHKLFKKVGIPGPTPLPFLGTLIFLYRLGLWKFDRECNEKYGEMW
GLYEGQPPMLVMDPDMIKTVLVKECYSVFTNRMPLGPMGFMKSALSFAEDEEWKRIRTLSPAFTSVKFKE
MVPPIISQCGDMLVRSRLREAENSKPTNLKDDFGAYTMDVITGTFLFGVNLDLNNPQDPFLKNMKLLKLDL
DPFFLSISLFPFLTPVFEVLNIGLFPKDVTRFLKNSIERMKESRLKDKQKHRVDFFQQMIDSQNSKETKSHK
ALSDLELVAQSIIFIAAYDTTSTTLPIFYELATHPDVQQLQEEIDAVLPNKATVTYDALVQMEYLDIVV
NETLRLFPVVS RVTRVCKKDIEINGVIFIPKGLAVMVP IYALHHDPKYWTEPEKFCPERFSKKNKDSIDPYRY
IPFGAGPRNCIGMRFALTNIKLAVIKALQNFSEFPECEETQIPLKLNNLPILQPEKPIVLKVHLR
>macmu3a4/1-496
MDLIPDLAVETWLLAVTLVLLYLYGTHSHGLFKKLGIPGPTPLPLGNILSYRKGFWTFDMECYKKGKVV
GFYDGRQPVLAITDPNMIKTVLVKECYSVFTNRRPFGPVGMKNAISIAEDEEWKRIRSLSPFTTSGKLKE
MVPPIIAKYGDVLRNLRREAETGKPVTLKDVFGAYSMDVITSTSFVGNIDSLNNPQDPFVENTKKLLRFDFL
DPFFLSITIFPFIIPILEVLNISIFPREVTSFLRKS VKRIKESRLKDTQKHRVDLQLMIDSQNSKETESHK
ALSDQELVAQSIIFIFAGYETTSSVLSFIIYELATHPDVQQLQEEIDTVLPNKAPPTYDTVLQMEYLDMMV
NETLRIFFIAMRLERVCKKDVEINGVIFIPKGVVVMIPSYALHHDPKYWPEPEKFLPERFSKKNNDNIDPYIY
TPFGSGPRNCIGMRFAIMNMKLAIIRVLQNFSEFKPCKETQIPLKLRGLLQTEKPIVLKIESR
>macmu3a5/1-496
MDLIPNLAMETWLLAVSLVLLYLYGTRSYGLFKRQGIPGPTPLPFLGNILSYRQGLWKFDTECYKKGKMM
RTQDGGQPLVLTITDPDMIKTVLVKECYSVFTNRRPLGPVGLMKS AISIAEDEEWKRIRSLSPFTTSGKLKE
MFPIIAQYGDMLVRNLRREAEGKGPVTLKDIFGAYSMDVITSTSFVGNIDSLNNPKDPFVESVKKFLKDFL
DPLFLTLIFPFLTPAFEALNVSLFPKDAINF LNKS VNMSMKSR LNDKQKHRVDLQLMIDSQNSKETESHK
ALSDQELVAQSIIFIFAGYETTSSVLSFIIYELATHPDVQQLQKEIDAVLPNKAPATYDAMVQMEYLDMMV
NETLRLFP IAIRLERACKKDVEINGVIFIPKGVVVMIPSYALHHDPKYWTEPEEFRPERFSKKNKDSIDPYIY
TPFGTGRNCIGMRFAIMNMKLAI IKVLQNFSEFKPCKETQIPLKLGNOGLLQSEKPIVLKVESR
>macmu3a7/1-496
MDLIPDLAVETWLP LAVSLI LLYLYGTHSHGLFKKLGIPGPTPLPFFGNVLSFRKGFWTFDMECYKKGKVV
GFYDCRQPVLAITDPNMIKTVLVKECYSVFTNRRPLGPVGMKS AITVAEDEEWKRIRSLSPFTTSGKLKE
MVPPIIAQYGDVLRNLRREAETGKPVTLKDVFGAYSMDVITSTSFVGNVDSLNNPQDPFVENTKKLLRFNSL
DPFFLSIKIFPFLTPIFEALNITVFPKVTISFLRKS VKRMKEGR LKETQKRRVDLQLMIDSQNSKDTESHK
VLSDELMQAQSIMFIFAGYETTSSVLSFIIYELATHPDVQQLQKEIDTVLPNKAPPTYDTMLQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMSIPKGVVVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIDPYIY
TPFGNGPRNCIGMRFAIMNMKLALIRVLQNFSEFKPCKETQIPLKLSFGGLLKTEKPIVLTAESR
>macmu3a43/1-447
MDLIPNFAMETWVLVATSLVLLYM-----
-LYEGQPPMPVIMDPDMIKTVLVKECYSVFTNRMPLGPMGFMKSALSFAEDEEWKRIRTLSPAFTSVKFKE
MVPPIISQCGDMLVRSRLREAENSKPTNLKDDFGAYTMDVITGTFLFGVNLDLNNPQDPFLKNMKLLKLDL
DPFFLSISLFPFLTPVFEVLNIGLFPKDVTRFLKNSIERMKESRLKDKQKHRVDFFQQMIDSQNSKETKSHK
ALSDLELVAQSIIFIAAYDTTSTTLPIFYELATHPDVQQLQEEIDAVLPNKAPVTYDALVQMEYLDIVA
NETLRLFPVVS RVTRVCKKDIEINGVIFIPKGLAVMVP IYALHHDPKYWTEPEKFCPERFSKKNKDSIDPYRY
IPFGAGPRNCIGMRFALTNIKLAVIKALQNFSEFPECEETQIPLKLNNLPILQPENPIVLKVHLR

>ppy3a4/1-490
-----LAMETWILLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPLLGNVFSYRKGFWRFDMECHKKYGKVV
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISIAEDEEWKRIRSLLSPTFTSGKLKE
MVPPIIARYGDVLRNLRREAETGKPVTLKDVFGAYSMVDITSTSFVGNIDSLNNPQDPFVENTKKLLRFDFL
DPFFLSIIVFPFLTPILEILNISVFPREVTSFLRKSVKRMKESRLKDTQKHRVDLQQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHDPVQKRLQEEIDAVLPNKAPPTYDVLQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYALHHDPKYWTEPEKFLPERFSKKNKDNIPIY
TPFGTGRNCIGMRFALNMMLAVIRVLQNFSEFKPCKETQIPLKLRGLGGLLQTEKPIVLKVESR
>ppy3a5/1-484
-----WLLAVSLVLLYLYGTRTHGLFKKLGIPGPTPLPLLGNVLSYRQGLWKFDTECYKKYKGMW
GTYDQQLPVLAITDPDMIKTVLVKECYSVFTNRRSLGPGVGMKSAISIAEDEEWKRIRSLLSPTFTSGKLKE
MFPIIAQYGDVLRNLRREAEGKGPVTLKDIFGAYSMVDITGTSGVGNIDSLNNPQDPFVESTKKFLKFSFL
DPLFLSIILFPFLTPVFEALNVSLFPKDTINF LNKSVMRMKKSRLNDKQKHRDLFLQQLMIDSQNSKETESHK
ALSDLELAAQSIIFIFAGYETTSSVLSFTLYELATHDPVQKQLQEEIDAVLPNKAPPTYDAVVQMEYLDMMV
NETLRLFPVAIRLERCTCKKDVEINGVFIPKGMVVIPTIALHHDPKYWTEPEEFRRPERFSKK-KDSIDPIY
TPFGTGRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLDTOGLLQPEKPIVLKVDSR
>ppy3a43/1-496
MDLIPNFAMETWVLVATSLVLLYIYGTHSHKLFKKLGIPGPTPLPFLGTILFYLRGLWKFDRECNEKYGEMW
GLYEGQQPMLVIMDPDMIKTVLVKECYSVFTNRMPLGPMGLKSALSFAEDEEWKRIRTLSPAFTSVKFKKE
MVPPIISQCGDMLVRSRQEAENSKSINLKDFFGAYAMDVITGTSLFGVNLDSLNNPQDPFLKNMKLLKLDLFL
DPFLLLSLFPFLTPVFEALNIGLFPKDVTRFLKNSIERMKESRLKDKQKHRVDFFQQMIDSQNSKETESHK
ALSDLELVAQSIIFIAAYDTTSTLPLFIMYELATHDPVQKQLQEEIDAVLPNKAPVTYDALVQMEYLDMMV
NETLRLFPVVSRTVRVCKKDIEINGVLIPKGLAVMVPYIALHHDPKYWTEPEKFCPERFSKKNKDSIDPIY
IPFGAGPRNCIGMRFALTNIKLAVIRALQNFSEFKPCKETQIPLKLDNLPILQPEKPIVLKVHLSR
>ppy3a67/1-496
MDLIPNLAVETWLLAVSLIILYLYGTHSHGLFKKLGIPGPTPLPFVGNALSFKHGFWTFDMECYKKYRKVV
GFYDGLQPVLAITDPDMIKTVLVKECYSVFTNRRPLGPGVGMKSAISRAEDEEWKRIRSLLSPTFTSGKLKE
MVPPIIAQYADVLRNLRREAETGKPVTLKDFFGAYSMVDITSTSFVGNIDSLNNPQDPFVENTKKLLRFNSL
DPLTSLIKIFPFLIPILEALNVTVPREVTSFLTKSVKRIKEGRLKETQKHRVDLQQLMIDSQNLKDTESHK
ALSDLELVAQSIIVFIFAGYETTSSVLSFIIYELATHDPVQKQLQEEIDAVLPNKAPPTYDVLQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYVLHHDPKYWIEPEKFLPERFSKKNKDNIPIY
TPFGNGPRNCIGMRFALNMMLFALVRVLQNFSEFKPCKETQIPLKLRFGGLLRTEKPIVLMAESR
>ptr3a4/1-496
MALIPDLAMETWLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYRKGFVCFDMECHKKYGKVV
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISIAEDEEWKRIRSLLSPTFTSGKLKE
MVPPIIAQYGDVLRNLRREADTGKPVTLKDVFGAYSMVDITSTSFVGNIDSLNNPQDPFVENTKKLLRFDFL
DPFFLSIIVFPFLTPILELVNICVFPREVTFNLRKSVMRMKESRLKDTQKHRVDLQQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHDPVQKQLQEEIDAVLPNKAPPTYDVLQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYALHHDPKYWTEPEKFLPERFSKKNKDNIPIY
TPFGSGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLRGLGGLLQPEKPIVLKVESR
>ptr3a5/1-495
MDLIPNLAVETWLLAVSLVLLYLYGTRTHGLFKKLGIPGPTPLPLLGNVLSYRQGLWKFDTECYKKYKGMW
GMYDQQLPVLAITDPDMIRTVLVKECYSVFTNRRSLGPGVGMKSAISIAEDEEWKRIRSLLSPTFTSGKLKE
MFPIIAQYGDVLRNLRREAEGKGPVTLKDIFGAYSMVDITGTSGVGNIDSLNNPQDPFVESTKKFLKFGFL
DPLFLSIILFPFLTPVFEALNVSLFPKDTINF LNKSVMRMKKSRLNDKQKHRDLFLQQLMIDSQNSKETESHK
ALSDLELAAQSIIFIFAGYETTSSVLSFTLYELATHDPVQKQLQEEIDAVLPNKAPPTYDAVVQMEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGSVMVVIPTIALHHDPKYWTEPEKFLPERFSKKNKDNIPIY
TPFGTGRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLRGLGGLLQPEKPIVLKVESR
>ptr3a7/1-496
MDLIPNLAVETWLLAVSLIILYLYGTRTHGLFKKLGIPGPTPLPFLGNALSFRKGYWTFDMECYKKYRKVV
GIYDCQQPMLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISIAEDEEWKRIRSLLSPTFTSGKLKE
MVPPIIAQYGDVLRNLRREAETGKPVTLKHVFGAYSMVDITSTSFVGNIDSLNNPQDPFVENTKKLLRFNPL
DPFVLSIKVFPFLTPILEALNITVFPKVISFLTKSVKQIKEGRLKETQKHRVDLQQLMIDSQNSKEDSETHK
ALSDLELMAQSIIFIFAGYETTSSVLSFIIYELATHDPVQKQVQKEIDTVLPNKAPPTYDVLQLEYLDMMV
NETLRLFPVAMRLERVCKKDVEINGMFIPKGVVVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIPIY
TPFGSGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLRFGGLLLTEKPIVLKAESR
>ptr3a43/1-496
MDLIPNFAMETWVLVATSLVLLYIYGTHSHKLFKKLGIPGPTPLPFLGTILFYLRGLWNFDRECNEKYGEMW
ELYEGQQPMLVIMDPDMIKTVLVKECYSVFTNRLPLGPMGLKSALSFAEDEEWKRIRTLSPAFTSVKFKKE
MVPPIISQCGDMLVRSRQEAENSKSINLKDFFGAYTMDVITGTSLFGVNLDSLNNPQDPFLKNMKLLKLDLFL
DPFLLLSLFPFLTPVFEALNIGLFPKDVTHFLKNSIERMKESRLKDKQKHRVDFFQQMIDSQNSKETESHK
ALSDLELVAQSIIFIAAYDTTSTLPLFIMYELATHDPVQKQLQEEIDAVLPNKAPVTYDALVQMEYLDMMV
NETLRLFPVVSRTVRVCKKDIEINGVFIPKGLAVMVPYIALHHDPKYWTEPEKFCPERFSKKNKDSIDPIY
IPFGAGQNCIGMRFALTNIKLGVIRALQNFSEFKPCKETQIPLKLDNLPILQPEKPIVLKVHLSR
>ssc3a46/1-496
MDLIPGFSTETWVLLATSLVLLYLYGTYSGLFKKLGIPGRPLPYFGNIGYRKGVDFDKKCFQYKGMW
GFFDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPRGAMRSVSLAEDEEWKRIRTLSPAFTSVKFKKE
MFPIISHYGDLLVSNLRREAEGKGPVTMKDIFGAYSMVDITSTAFGVNVDLNNPQDPFVENSRLKLSFF

SPFFLSIIFFPFLTPILEVLNITLFPKSVNFFTRSICKRMKESRLKDTQKHRVDLLQLMINSQNSKEMDAHK
GLSNEELVAQGVIFIFAGYETTSSSLSLVYELATHPDVQQKLQEEIDATFPNKAPTYDGLAQMEYLDMMV
NESLRIFPVTPLRLERSCKKDVEIHGVFVPGKTVMMVPIFALHRAPELWPEPEEFRPERFSKKNKDTINPYTY
LPFGTGPRNCIGMRFALNMNKLALVRVLQNFSEFKPCRETQIPLKISSQGLIQPEKPIVLMVVR
>ppr3a126/1-496
MSYDLFFSAETWALLSLLVALLVLYGHWPHRQFKKLGIAGPKPIPFFGTMLRYREGFHNFDLECFKKYGRIW
GIYDARQPVLCTDQSIKTIILIKECYSLFTNRRNFRLNGPLYDAVSIVEDDDWRRIRSVLSPSFTSGRLKE
MFGIMKAHSQILVQNLGKSATRGENVDIKEVFGAYSMDVVTSTAFSVDIDSLNNPKDPFVTNIKKMLKFDL
SPVFLISVLFPTIPILEKMELGFFPISVTDFFYAALQIKISERVANDHKKRVDFLQLMVDSQTAGEEHTEK
GLSDHEILSQAMIFIFAGYETSSSTLSFFFYNLATNPEAMKKLQEEVDETFPNKAPVDYEAVNMNDYLDAL
NESLRFPVAARLERVCKKTVINGVVIIPKDMVVMIPTFALHRDPDYWSDPESFKPERFTKGNKEAIDPYMY
MPFGLGPRNCIGMRFAQVTMCLAIVEILQRFDVSVCAETQVPLKLGKSGLLAPKDPINLKFEHR
>gra3a127/1-486
-----TWALLVLLVTLVLYGWSWTHSIFKKLGIQGPPIPFFGTMLRYREGFHNFDLECFRKYGRVW
GIYDARQPVLCTDQSIKTIILIKECYSLFTNRRNFRLNGPLYDAVSIVEDDDWRRIRSVLSPSFTSGRLKE
MFGIMKTHSQILVQNLGKSATRGENVDIKEFFGAYSMDVVTSTAFSVDIDSLNNPKDPFVTNIKKMLKFDLL
NPVFLISAMFPFITPLMEKMDFAFFPTSVDFFYAALQIKISERVVSDHKKRVDFLQLMVDSQTAGEDHTEK
GLSDHEILSQSMIFIFAGYETSSSTLAFFFYNLATNPEAMEKLQEEIDETFPNKAAPVDYEAVNMNDYLDAL
NESLRFPVAARLERVCKKTVINGLLIPKDMVVMIPTYALHRDPDYWSDPESFKPERFTKGNKESIDPYMY
MPFGLGPRNCIGMRFAQVTMCLAIVEILQRFDVSVCAETQVPLELGLSGLLAPKDPKIKLKLKPR

Figure 22

>2A6_3_HUMAN_P11509

M-LASGMLLVALLVCLTVMVLSVWQQRKSK--
GKLPPGPTPLPFIGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVL CGHDAVREALVDQA
EEFSGRGEQATFDWVGYG-----
VVFSNGERAKQLRRFSIATLRDFGVGKRGIEERI QEEAGFLIDALRGTTGGANIDPTFFLSRTVSN
VISSIVFGDRFDYKDKEFLSLLRMMLGIFQFTSTSTGQLYEMFSVMKHLPGPQQQAFQLLQGGLED
FIAKKVEHNQRTLDPN--
PRDFIDSFLIRMQEEEEKNPNTFEFYLKNLVMTTNLNFIGGTETVSTTLRYGFLLLMKHPEVEAKVH
EEIDRVIGKNRQPKFEDRAKMPYMEAVIHEIQRFQDVIPMSLARRVKKDTKFRDFFLPKGTEVFP
MLGSLVRDPSFFSNPQDFNPQHFLNEKGQFKKSDAFVPPFSIGKRNCFGEGGLARMELFLFFTTVMQ
NFRKSSQS-PKDIDVSHVGFATIPRNYTMSFLPR

>2D1_RAT_P10633

MELLNGTGLWSMAIFTVIFILLVDLMHRRHRWTSRYPPGPVPWPVLGNLLQVDLSNMPYSLYKLQ
HRYGDVFSLQKGWKPMVIVNRLKAVQEVLVTHGEDTADRPPVP IFKCLGVKPRSQGVILASYGPE
WREQRRFSVSTLRFTFGMGKKSLEEWVTK EAGHLCDAF TAQAGQSINPKAMLNKALCNVIASLIFA
RRFEYEDPYLIRMVKLVEESL TEVSGF IPEVLNTFPALLRIPGLADKVFQGGQKTFMALLDNLLAE
NRTTWDPAQPPRNLDAFLAEVEKAKGNPESSFN DENLRMVVDLFTAGMVT TATTLTWALLLMI
LYPDVQRRVQQEIDEVIGQVRCPEMTDQAHPYTN AVIHEVQRFQDIAPLNLPRFTSCDIEVQDF
VIPKGTTLI INLSSVLKDET VWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRF SFSVPVQGPRPSTHGFFAFPVAPLPYQLCAVVR

>2D10_RAT_P12939

MELLNGTGLWPMAIFTVIFILLVDLMHRHRQRWTSRYPPGPVPWPVLGNLLQVDP SNMPYSMYKLQ
HRYGDVFSLQMGWKPMVIVNRLKAVQEVLVTHGEDTADRPPVP IFKCLGVKPRSQGVV FASYGPE
WREQRRFSVSTLRFTFGMGKKSLEEWVTK EAGHLCDAF TAQNGRSINPKAMLNKALCNVIASLIFA
RRFEYEDPYLIRMLTLVEESLIEVSGF IPEVLNTFPALLRIPGLADKVFQGGQKTFMAFLDNLLAE
NRTTWDPAQPPRNLDAFLAEVEKAKGNPESSFN DENLRMVVDLFTAGMVT TATTLTWALLLMI
LYPDVQRRVQQEIDEVIGQVRCPEMTDQAHPYTN AVIHEVQRFQDIAPLNLPRITSCDIEVQDF
VIPKGTTLI INLSSVLKDET VWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQHFSFSVPAGQPRPSTLGNFAISVAPLPYQLCAAVR

>2D3_RAT_P12938

MELLAGTGLWPMAIFTVIFILLVDLMHRHRQRWTSRYPPGPVPWPVLGNLLQVDLCNMPYSMYKLQ
NRYGDVFSLQMGWKPMVIVNRLKAVQELLVTCGEDTADRPEMP IFQHIGYGHKAKGVVLAPYGPE
WREQRRFSVSTLRNFGVGKKSLEQWVTDEASHLCDALTA EAGRPLDPYTLLNKAVCNVIASLIYA
RRFDYGDPDFIKVLKILKESMG EQTGLFPEVLNMFPVLLRIPGLADKVFPGQKFTLTMVDNLVTE
HKKTWDPDQPPRDLTDAFLAEIEKAKGNPESSFN DANLRVLVNDLFGAGMVTTSITLTWALLLMI
LHPDVQCRVQQEIDEVIGQVRHPEMADQAHPFTNAVIHEVQRFADIVPMNLPHKTSRDIEVQGF
LIPKGTTLIPNLSSVLKDET VWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRF SFSVPTGQPRPSDYGVFAFL LSPSPYQLCAFKR

>2D5_10_RAT_Q5I0P9

MELLNGTGLWPMAIFTVIFILLVDLMHRHRQRWTSRYPPGPVPWPVLGNLLQVDP SNMPYSMYKLQ
HRYGDVFSLQMGWKPMVIVNRLKAVQEVLVTHGEDTADRPPVP IFKCLGVKPRSQGVV FASYGPE
WREQRRFSVSTLRFTFGMGKKSLEEWVTK EAGHLCDAF TAQNGRSINPKAMLNKALCNVIASLIFA
RRFEYEDPYLIRMLTLVEESLIEVSGF IPEVLNTFPALLRIPGLADKVFQGGQKTFMAFLDNLLAE
NRTTWDPAQPPRNLDAFLAEVEKAKGNPEGSFNDEKL RMVVVDLFTGGMVTTSSTLTWALLLMI
LYPDVQRRVQQEIDEVIGQVRCPEMTDQAHPYTN AVIHEVQRFQDIAPLNLPRITSCDIEVQDF
VIPKGTTLI TNLSSVLKDET VWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQHFSFSVPAGQPRPSTLGNFAISVAPLPYQLCAAVR

>2D4_RAT_P13108

MRMPTGSELWPIAIFTIIFLLLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQIDFQNM PAGFQKLR
CRFGDLFSLQLAFESVVVLNGLPALREALVKYSEDTADRPP LHFNDQSGFGPRSQGVVLARYGPA
WRQRRFSVSTFRHFGLGKKSLEQWVTEEARCLCAAFADHSGFPFSPNTLLDKAVCNVIASLLFA
CRFEYNDPRFIRLLDLLKDTLEESGFLPMLLNVPFMLLHIPGLLGKVFSGKKA FVAMLDELLTE
HKVTWDPAQPPRDLTDAFLAEVEKAKGNPESSFN DENLRVVVADLFMAGMVTTSSTLTWALLFMI
LHPDVQCRVQQEIDEVIGQVRRPEMADQARMPFTNAVIHEVQRFADILPLGVPHKTSRDIEVQGF
LIPKGTTLI TNLSSVLKDET VWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRF SFSVPTGQPRPSDYGIFGALTTPRPYQLCASPR

>2D26_2D2_RAT_P10634

MGLLI GDDLWAVVIFTAIFLLLVDLVHRHKFWTAHYPPGPVPLPGLGNLLQVDFENMPYSLYKLR
SRYGDFVSLQIAWKPVVINGLKA VRELLVTYGEDTADRPLLP IYNHLGYGNKSKGVVLAPYGPE
WREQRRFSVSTLRDFGVGKKSLEQWVTEEAGHLCDTFAKEAEHPFNPSILLSKAVSNVIASLVYA
RRFEYEDPFFNRMLKTLKESFGEDTGMAEVLNAIPILLQIPGLPGKVFPKLNSFIALVDKMLIE
HKKSWDPAQPPRDMTDAFLAEMQKAKGNPESSFNENLRLVVIDLFMAGMVTTSTTL SWALLLMI
LHPDVQRRVHEEIDEVIGQVRRPEMADQARMPFTNAVIHEVQRFADIVPTNIPHMTSRDIKFQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVLAGRPRPSTHGVALPVTQPYPQLCAVAR

>2D18_RAT_Q64680

MRMPTGSELWPIAIFTIIFLLLVDLMHRRQRWTSRYPPGPVWPVPLGNLLQIDFQNM PAFQKLR
CRFGDLFSLQLAFESVVVLNGLPALREALVKYSED TADRPP LHFNDQSGFGPRSQGVVLARYGPA
WRQRRFSVSTFRHFGGLGKKSLEQWVTEEARCLCAAFADHSGFPFSPNTLLDKAVCNVIASLLFA
CRFEYNDPRFIRLLDLLKDTLEESGFLPMLLNVPFMLLHIPGLLGKVFSGKKA FVAMLDLLEL TE
HKVTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENL RVVADLFMAGMVTTSTTLTWALLFMI
LRPDVQCRVQQEIDEVIGQVRRPEMADQARMPFTNAVIHEVQRFADILPLGVPHKTSRDI EVQGF
LIPKGTTLIINLSSVLKDETVWEKPLRFHPEHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSNYGVFGALTTPRPYQLCASPR

>2D34_MOUSE_L7N463

MELLTGTGLWSVAIFTVIFLILVDLMHRRQHWT SRYPPGPVWPVPLGNLLQVDLDNIPYSLYKLO
NRYGDFVSLQMAWKPVVINGLKAMQEVL LTCGKDTADHPPVP IFEYLGFKSKSQGVVLASYGPE
WREQRRFSVSTLRNFGGLGKKSLEEWVTKEAKHLCDAF TARAQQSINPNTMLNNAVCNVIASLIFA
RRFEYEDPFLIRMLKMREESLKEVTGFIPGVNLTFPILLRIPGLADMVFQSQKTFMAILDNLVTE
NRRTWDPDQPPRNLADAF LAEIQKAKGNPESSFNENL CMVVSDF TAGMVTTSTTLSCALLLMI
LHPDVQRRVQQEIDAVIGQVRCP EMADQARMPYTNAVIHEVQRF GDIIPLNIPRITSRDI EVQDF
LIPKGTILIPNMSSMLKDETVWEKPLRFYPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVPAGQPQPSDHRIFAIPVAPYPYQVCAIMR

>2D10_MOUSE_P24456

MELLTGAGLWSVAIFTVIFILLVDLMHRHQHWT SRYPPGPVWPVPLGNLLQVDLDNMPYSLYKLO
NRYGDFVSLQMGWKPMVVINGLKAMKEVLLTCGEDTADR PQVP IFEYLGVKPGSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEARHLCDAF TAQAGQPINPNTMLNNAVCNVIASLIFA
RRFEYEDPYLIRMQKVLEDSLTEISGLIPEVLNMFPILLRIPGLPGKV FQGQKSLLAIVENLLTE
NRRTWDPDQPPRNLDAFLAEIEKVKGNAESSFNENL RMVVLDF TAGMVTTSTTL SWALLLMI
LHPDVQRRVQQEIDAVIGQVRHPEMADQARMPYTNAVIHEVQRF GDIAPLNLPRIITSRDI EVQDF
LIPKGSILIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQHFSFSVPNGQPRPRNLGVFPFPVAPYPYQLCAVMR

>2D11_MOUSE_P24457

MELLTGAGLWSVAIFTVIFILLVDLMHRHQHWT SRCPPGPVWPVPLGNLLQVDLDGNMPYSLYKLO
NRYGDFVSLQMGWKPMVVINGLKAMKEVLLTCGEDTADR PQVP IFEYLGVKPGSQGVVLAPYGPE
WQEQRRFSVSTLRNFGGLGKKSLEDWVTKEARHLCDAF TAQAGQPINPNTMLNNAVCNVIASLIFA
RRFEYEDPYLIRMLKMLKECFTEISGFIPGVNLEFP IFLRIPGLADMVFQ GQKSFMAILDNLLTE
NRRTWDPDQPPRNLDAFLAEIEKAKGNPESSFNENL RMVVGDLF TAGMVTTSTTL SWALLLMI
LHPDVQRRVQQEIDAVIGQVQHPEMADQARMPYTNAVIHEVQRF GDIAPLP LPRIITSRDI EVQDF
LVTKGSTLIPNMSSVLKDETVWEKPLRFHPEHFLDAQGHFVKPEAFMPFSAGHRSCLGEALARME
LFLFFTCLLQRFSSVDPGQPQPSNYRVHAIPVAPFPYQLCAVMR

>2D9_MOUSEii_Q3UNW2

MELLTGTDLWPVAIFTVIFILLVDLTHQRQRWTSRYPPGPVWPVPLGNLLQVDLDGNMPYSLYKLO
NRYGDFVSLQMAWKPMVVINGLKAMKEMLLTCGEDTADR PPVP IFEYLGVKPGSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEANHLCDAF TAQAGQPINPNMPLNKSTCNVIASLIFA
RRFEYEDPFLIRMLKVLEQSLTEVSGLIPEVLNAFPILLRIPRLADKALQGQKSFIAILDNLLTE
NRRTWDPVQAPRNLDAFLAEIEKAKGNPESSFNENL LMVVRDLF GAGMLTTSTTL SWALVLM I
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRF GDIVPNLPRIITSHDI EVQDF
LIPKGTILLPNMSSMLKDES VWEKPLRFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEALARME
LFLFFTCLLQRFSSVDPGQPQPSNSGVYGILVAPSPYQLCAVVR

>2D9_MOUSEi_P11714

MELLTGTDLWPVAIFTVIFILLVDLTHQRQRWTSRYPPGPVWPVPLGNLLQVDLDGNMPYSLYKLO
NRYGDFVSLQMAWKPMVVINGLKAMKEMLLTCGEDTADR PPVP IFEYLGVKPGSQGVVLAPYGPE
WREQRRFSVSTLRNFGGLGKKSLEDWVTKEANHLCDAF TAQAGQPINPNMPLNKSTCNVIASLIFA
RRFEYEDPFLIRMLKVLEQSLTEVSGLIPEVLNAFPILLRIPRLADKALQGQKSFIAILDNLLTE

NRTTWDVPVQAPRNLTDAFLAEIEKAKGNPESSFNENLLMVVRDLFGAGMLTTSTTLSWALMLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIPVNLPRITSHDIEVQDF
LIPKGTILLPNMSSMLKDESVWEKPLRFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEALARME
LFLFFTCLLQRFSSVDPDGPQPSNSGVYGILVAPSPYQLCAVVR
>2D26_MOUSE_Q8CIM7
MGLLVGDDLWAVVIFTAIFLLLVDLVHRRQRWTACYPPGPVPFPGGLGNLLQVDFENIPYSFYKLO
NRYGNVFSLQMAWKPVVVVNLKAVRELLVTYGEDTSDRPLMPIYNHIGYGHKSKGVILAPYGPE
WREQRRFSVSTLRDFGLGKKSLEQWVTEEAGHLCDRAFTKEAEHPFNPSPLLSKAVSNVIASLIYA
RRFEYEDPFFNRMLKTLKESLGEDTGFVGEVLNAIPMLLHIPGLPDKAFPKLNSFIALVNKMLIE
HDLTWDPAQPPRDLTDAFLAEVEKAKGNPESSFNENLRIVVIDLFMAGMVTSTTLSWALLLMI
LHPDVQRRVHQEIDEVIGHVRHPEMADQARMPYTNAVIHEVQRFADIVPTNLPHMTSRDIKFQDF
FIPKGTTLIPNLSSVLKDETVWEKPLRFYHPEHFLDAQGHFVKHEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVDPDGPQPRPSDYGIYTMPVTPEPYQLCAVAR
>2D22_MOUSE_EMG
MRLPTGAELWPIAIFTVIFLILVNLMHWRQRWTAHYPPGMPWPVPLGNLLHMDFQNMPPAGFQKLR
GRYGDLSLQQLASESVVVLNGLTALREALVKHSEDTADRPPPLHFNDLLGFGPRSOGIVLARYGPA
WRQQRFSVSTMHHFGLGKKSLEQWVTEEARCLCAAFADHTGHPFSPNTLLDKAVCNVIASLLYA
CRFEYDDPRFIRLLGLLKETLKEEAGFLPMFLNVFPMLLRIPGLVGKVPFGKRAFTVMTLDELLAE
HKTTWDPTQPPRDLTDAFLAEVEKAKGNPESSFNENLRTVVGDLSAGMVTSTTLSWALMLMI
LYPDVQRRVQQEIDEVIGQVQCPMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIELQGF
LIPKGTTLITNLSSALKDETVWEKPLCFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVDPDGPQPSDHGVFRALTTPCPYQLCALPR
>2D22_MOUSE_Q9JKY7
MRLPTGAELWPIAIFTVIFLILVNLMHWRQRWTAHYPPGMPWPVPLGNLLHMDFQNMPPAGFQKLR
GRYGDLSLQQLASESVVVLNGLTALREALVKHSEDTADRPPPLHFNDLLGFGPRSOGIVLARYGPA
WRQQRFSVSTMHHFGLGKKSLEQWVTEEARCLCAAFADHTGHPFSPNTLLDKAVCNVIASLLYA
CRFEYDDPRFIRLLGLLKETLKEEAGFLPMFLNVFPMLLRIPGLVGKVPFGKRAFTVMTLDELLAE
HKTTWDPTQPPRDLTDAFLAEVEKAKGNPESSFNENLRTVVGDLSAGMVTSTTLSWALMLMI
LHPDVQRRVQQEIDEVIGQVQCPMADQARMPYTNAVIHEVQRFADILPLGVPHKTSRDIELQGF
LIPKGTTLITNLSSALKDETVWEKPLCFHPEHFLDAQGHFVKPEAFMPFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVDPDGPQPSDHGVFRALTTPCPYQLCALPR
>2D12_MOUSE_Q8BVD2
MELLTGTDLWSVAIFTVIFILLVDLMHRRQSWTSCYPPGPVPWPVPLGNLLQVDLNNMPYSLYKLO
NRYGDVFSLQMAWKPMVVINRMKAMKEVLLTCGEDTADRPPVPIFEHLGFKPRSOGMIFAPYGPE
WREQRRFSLSLRLNGLGRKSLEEWVIKEAGHLCDRAFTTQAGQYINPNTMLKKATCNVIASLIFA
RRFEYEDPYLIRMLKVLEDSLTELGLIPEVINTFPILLHIPRLADKFLQSQKSFAIIVDNLLE
NRTTWDPAQSPRNLTDAFLAEIEKAKGNPESSFNENLRMVVIDLFTAGILTSTTLSWALLLMI
LHPDVQRRVQQEIDEVIGQVRHPEMADQAHMPYTNAVIHEVQRFQDIPVNLPRITSRDIEVQDF
LIPKGTILLPNMSSVHMDTVWEKPLRFHPEHFLDAQGHFVKHEAFITFSAGRRSCLGEPLARME
LFLFFTCLLQRFSSVDPDGPQPSDHRVFSIMVAPSPYQLCAVIR
>2D281_B_GOLDEN_HAMSTER_Q9QUJ1
MELLTGHLWPVAMFTVILILLVDLLHRRQRWASRYPPGPVPLPLLGNLLQVDLKNMQYSVHKLQ
QHYGDVFSLQMAWKHMVMINGLKAVREVLVNYGEYTADRPKIPYEHGSLGPKARGVILAPYGPE
WREQRRFSVSTLRNLGLGKKSLEQWVTDEAGHLCDAFKDAQGRPFNPSTLLNKAVCNVITSLIFA
RRFEYEDANLIRMLRLLEEALTNISGFIPFELNTPVLLHIPGLFDKVFSGQKTFAAIIVDNLLE
NRRTWDPEQPPRGLTDAFLAEVEKAKGNPESSFNENLRMVVNDLFIAGTVSTSTTLSWALLFMI
QYPDVQRRVQQEIDIDILGPGRSPMADQARMPYTNAVIHEVQRFADIAPLNLPICITSRDIKVQGF
LIPKGTTLITNLSSVLKDETVWEKPLRFHPEHFLDAQGCFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPAGQPRPSDHVVLGVLSKAPYQLCAVPR
>2D20_GOLDEN_HAMSTER_Q9QYG5
MVLLIGDGLWSGVIFTALFLLLVDLMHRRKFWRARYPPGMPPLPGLGNLLQVDFENMPYSLYKFQ
QRYGDVFSLQMAWKPVVVINGLKAVREVLVNCGEDTADRPPVPIFNHLGYRPKSQGVVFARYGPQ
WREQRRFSVSTMTRDFGVGKKSLEQWVTEEAGHLCDRAFTQEAGHPFNITLLNKSVCNVISSLIYA
HRFDYEDPFFNKLLKTLQESFGEDSGFIAEVLNAVVPVLLRIPGLPGKAFPKLTAFMDSLYKMLIE
HKTTWDPAQPPRGLTDAFLAEVEKAKGRPESSFNENLHVVVADLFIAGMVTSTTLSWALLLMI
LHPDVQSRVQQEIDDVIGQVRRPEMADQARMPYTNAVIHEVQRFQDIPVNVPHMTSRDIEVQGF
LIPKGTTLIPNLSSVLKDETVWEKPLRFHPEHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVDPAGQPRPSDQGVFALPVTPTPYELCAVVR

>2D4_CHINESE_HAMSTER_G3I0F2

MRLLTGADLWSVAIFMAIFLLLVLNLMHRRHRWTARYPPGPMPLPVLGNLLQVDFQNMPPAGFQKLR
SRFGDLFSLQLASEPVVVLNGLTALREALVRYSED TADRPLLTNDHDLGFGPRSQGVVLAQYGSA
WRQORRF SVSTFRHFLGLGKKSLEQWVTEEARCLCAAF TDYNGLPFSPNTLLNKAVCNVIASLLYA
CRFEYDDPRFIRLLDLLKDTLEESGFLPMFLNVFPVLLHIPGLPGKVFSGKRAFMMVMDLLELLE
HKMTRDPTQPPRDLTDAFLAEVEKAKGKPESSFN DENLRMVVADLFMAGMVTSTTLSWALLLMI
LYPDVQSRVQKEIDEVIGQERCEMADQAHMPYTNAVIHEVQRFSDILPLGVPHKTSRDIEVQGF
VIPKGTTLITNLSSVLKDET VWEKPLLFHPEHFLDAQGRFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFIFSVPDQGPRPSDQGVFGALTTPSPYQLCAFSR

>2D27_GOLDEN_HAMSTER_Q9QYG6

MALLIGDGLWSGVIFTALFLLLVDLMHRRKFWRARYPPGPMPLPGLGNLLQVDFEHMPYSLYKFR
QRYGDVFSLQMAWKPVVINGLKAVREVLVNCGED TADRPPVP IFNHVGF GHNSQGVAFARYGPQ
WREORRF CVSTMRDFGVGKKSLEQWVTEEAHGLCDAFTQEAGHPFNPTTLLNKSVCNVISSLIYA
HRFDYEDPFFNSLLKMLQESFGEDTGFIAEVLNAVVPVLLRIPGLPGKAFPKLTAFMDSLYKMLIE
HKTTWDPAQPPRGLTDAFLAEVEKAKGRPESSFN DENLRMVVADMF IAGMVTSTTLSWALLLMI
LHPDVQSRVQQEIDDVIGQVRRPEMADQARMPYTNAVIHEVQRF GD IAPVNI PHMTSHDVEVQGF
LIPKGTTLIPNLSSVLKDET VWEKPLLFHPEHFLDAQGRFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFIFSVPAGQPRPSDQGFALPVTPTPYELCAVVR

>2D16_GUINNEA_PIGi_Q64403

MGLLTGDALFSVAVAVAIIFLLLVDLMHRRQRWAARYPPGPVPVPGGLGNLLQVDFENMAYSCDKLR
HQFGDVFSLQFVWTPVVVVNGLLAVREALVNNSTDTSDRPTLPTNALLGFGPKAQGVIGAYYGPA
WREORRF SVSSLRNFLGLGKKSLEQWVTEEAACLC AAF TNHAGQPF C PKALLNKAVCNVISSLIYA
RRFDYDDPMVLRLLLEFLEETLRENSSLKIQVLNSIPLLLRIPCVA AKVLSAQRSFIALNDKLLAE
HNTGWAPDQPPRDLTDAFLTEMHKAQGNSESSFN DENLRLLVSDLF GAGMVTSTTLSWALLLMI
LHPDVQRHVQEEIDEVIGQGRCEMADQAHMPFTNAVIHEVQRFADIVPMGVPHMTSRDTEVQGF
LIPKGTMLFTNLSSVLKDET VWEKPLLFHHPGHFLDAEGRFVKREAFMPFSAGPRICLGEPLARME
LFLFFTSLQRF SFSVPEGQPRPSDRGAPYLVLVLPSPYQLCAVLR

>2D16_GUINNEA_PIGii_Q64403

MGLLTGDALFSVAVAVAIIFLLLVDLMHRRQRWAARYPPGPVPVPGGLGNLLQVDFENMAYSCDKLR
HQFGDVFSLQFVWTPVVVVNGLLAVREALVNNSTDTSDRPTLPTNALLGFGPKAQGVIGAYYGPA
WREORRF SVSSLRNFLGLGKKSLEQWVTEEAACLC AAF TNHAGQPF C PKALLNKAVCNVISSLIYA
RRFDYDDPMVLRLLLEFLEETLRENSSLKIQVLNSIPLLLRIPCVA AKVLSAQRSFIALNDKLLAE
HNTGWAPDQPPRDLTDAFLTEMHKAQGNSESSFN DENLRLLVSDLF GAGMVTSTTLSWALLLMI
LHPDVQRHVQEEIDEVIGQVRCPEMADQAHMPFTNAVIHEVQRFADIVPMGVPHMTSRDTEVQGF
LIPKGTMLFTNLSSVLKDET VWEKPLLFHHPGHFLDAEGRFVKREAFMPFSAGPRICLGEPLARME
LFLFFTSLQRF SFSVPEGQPRPSDRGAPYLVLVLPSPYQLCAVLR

>2D4_NAKED_MOLE_RAT_G5BPR6

MGLLTGDALFLVAVAMAVFLLLVDLMHRRQRWGARYPPGPVALPGLGNLLQLDLHNIPYSFYKLR
CHFGDVFSLQLAWNPAVVLNGLVAVREVLVNHSEDTADRPPVP INELLGFGPRAQGLVMAPYGP
WREORRF SVSTLRNLGLGKKSLEQWVTEEA SCLCVAFAKQDGCPSPGALVNNAV CNVIASLIYA
RRFEYNDP LLVRLDLLLEDTLTEGLELIPQVLNAFPVLLRIPGVAAKVFPQAQAF LALMDELVTE
HRATFDLSQPPQDLTGAFLAEVEKAKGNPESTFN NENLRLLVSDLF MAGMVTSTTTLAWALLLMI
LHPDVQRRVQKEIDEVIGQGRPEMADQARLPFTNAVIHEVQRF GD IVPVNLPHMASRDIEVQGF
RIPKGTGLVNLSSVLKDET VWEKPLLFHPEHFLDAEGHFVKHEAFMPFSAGRRACLGEPLARME
LFLFFTCLLQRF SFSVPEGQPPSDQGVFAFLNTPSPYQLCAVSR

>2D16_NAKED_MOLE_RAT_G5BPR7

MGLLNGDALLSVAIAVAVFLLLVDLMHRRQRWAARYPPGPMVPGLGNLLQLDFQNLPSFYKLR
HQFGDVFSLQFAWTPVVVFNGLVAVREALVNHGEDTSDRPPVPTNELMGFGPKAKGVIGAPYGP
WREORRF SVSTLRNLGLGKKSLEQWVTEEA TCLCDAFANHARHPFNPSALLSKAVCNVIASLIYA
HRFEYNDP IFIRMLKFLEDALREDSGLLLQVLNSMPVLLHIPWVAANVLP AQRSFMALLNELLTE
HKTTWDPEPPRGLADAF LAEVHKTKEKSESSFN DENLRLLVISELFTAGMVTSTITMSWALLLMI
LHPDVQCRVHQEIDEVIGQGRPEMADQAHMPFTNAVIHEVQRF GD IVPMSVPHMTSRDTEVQGF
LIPKGTILFINLSSVLKDET IWEKPLLFHPEHFLDAEGCFVKREAFMPFSAGPRICLGEPLARME
LFLFFTCLLQRF SFSVPKGNPQPSDRGMSLPLVAPSPYEICAVSR

>2D23_RABBIT_G1T2D2

MGLLSGEALAPLAVAVAIIFLLLVDLMHKKRPRWAARYPPGPVGIPGLGNLLQVDFRGIPNCFRQLR
CRYGDVFSLQLAWTPVVVLNGLPAAMREALVTYGEDTADRPSLSLEHLGFGPQAQGVIMACYGHA
WREORRF SVSTLRNFGMGKKSLEHWVTEEAACLC AAF SEHAGHPFSPKALLNKAI GNVIASLTFG

CRFEYDDHRLTRLMDLIEIMLEESTGILPLVLNMIPILLRIPGLVDKVFVRGQKAFMALLDELVTE
HRMTRDPAQPPRDLTDAFLDQVEKAKGNPESSFNDDNLRLLVADLFVAGMFTTSFTLSWALLLMI
LHPDVQRRVQQEIDEVIGPARRPEMGDQARMPYTTAVVHEVQRFADIVPLGVPHQTLRDIEVQGF
LIPKGTMLFTNLSSVLKDEAVWEKPFRRFHPGHFLDAQGRFVKQEAFFMPFSAGHRACLGEPLARME
LFLFFTCLLQRFSSVVPAGQPPSDQGAPATLVTPAPYQLCAVAR

>2D24_RABBIT_G1TU02

MGLLSGEALAPLAVAVAIFLLLVDLMHRRQRWAARYPPGPVGIPLGNLLQVDFRGIPNCFRQLR
RRYGDVFSLQLAWTPVVVLNNGPAVIREALVTYGEDTADRPPAHTLEPLGFGPHAQGVVMARYGPA
WREQRRFSVSTLRNFGGLGKKSLEQWVTEEATCLCAAFADHAGCPFSPSMLLNKAVCNVIASLTHG
CRFEYDDHRLTRLMDLTQTILKESTGNLPQVLNVIPILLRIPGLVDKVFVRGQKAFMALLDELVTE
HRMTRDPAQPPRDLTDAFLDQVEKAKGNPESSFNDDNLRLLVTDLFAAGMVTTSITLSWALLLMI
LHPDVQRRVQQEIDEVIGPARRPEMGDQARMPYTTAVVHEVQRFADIIPLGVPHQTSRDIEVQGF
LIPKGTVLFTNLSSVLKDEAVWEKPFRRFHPGHFLDAQGRFVKQEAFFMPFSAGRRACLGEPLARME
LFLFFTCLLQRFSSVVPVGPQPRPSDQGAPATLVTPAPYQLCAVAR

>2D6_SQUIRREL_I3N994

MGLLTG----

PLCVAVAIFLLLVDLMHRRQRWAARYPPGGLPLPLGLGNLLQIDFQNMPPYPRYKMRFRFGDVFSLQ
LAWKPVIVLNGLAAREALVNHNEYTSDRPWMP IYEHGFGPRSQGVIMATYGHAWREQRRFSVS
TMRNFGGLGKKSLEQWVTKEASCLCTAFANKAGSPFNPMSSLKRAVCNVISSLIYARRFEYDEQRL
AKILDLEDTVKEDSGFMPMVLNVVPLLRIPGLPGRVFPQAQAFMAMVDELLEEHRTWDPAQP
PRDLTDAFLAEVEKAKGNPESSFNDDNLRIMVLDLFTAGMVTSTTTLAWALLLMI LHPDVQRRVQ
QEIDEVIGQVRQPEMGDQDRMPFTMAVIHEVQRFGLVPLGLPHSTTRDIELQGFFIPKGTTLIT
NLSSVLKDETVWEKPLHFHPGHFLDAQGHFVKREAFMPFSAGRRVCLGEPLARMELFLFFTCLLQ
RFSFSVPPGQAQPSDYGVFGTIVTPSPYQLCAKSR

>2D17_CRAB_EATING_MACAQUEi_D9DB13

MEL---

DALVPLAVTVAIFLLLVDLMHRRQRWAARYPPGGLPLPLGLGNLLHVDKNTPYCFDQLRRRFGDV
FSLQLAWTPVVVLNGLAAREALVTCGEDTADRPPVPIYQVLGIGPRSQGVFLARYGPAWREQRR
FSVSTLRNLGLGKKSLEQWVTEEAACLAFTDQAGRPFRRPNSLLDKAVSNVIASLTYGRRFEYD
DPRFLRLFDLTALKEESGFLREVLNAVPLLLRIPGLAGKVLRSQKAFLTQLDELLETEHRMTWD
PAQPPRDLTEAFLAEMEKAAGNPESSEFNENLRIVVADLFSAGMVTSTTTLAWGLLLMI LHPDVQ
RRVQQEIDNVIGQVRRPEMGDQARMPYTTAVIHEVQRFGLVPLGVTHMTSRDIELQGFLIPKGT
TLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGRFVKPEAFLPFSAGRHAACLGEPLARMELFLFFT
CLLQRFSSVVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>2D17_CRAB_EATING_MACAQUEii_Q29488

MEL---

DALVPLAVTVAIFLLLVDLMHRRQRWAARYPPGGLPLPLGLGNLLHVDKNTPYCFDQLRRRFGNV
FSLQLAWTPVVVLNGLAAREALVTCGEDTADRPPVPIINQVLGFGPRSQGVFLARYGPAWREQRR
FSVSTLRNLGLGKKSLEQWVTEEAACLAFTDQAGRPFRRPNSLLDKAVSNVIASLTYGRRFEYD
DPRFLRLFDLTALKEESGFLREVLNAIPLLLRIPGLAGKVLRSQKAFLTQLDELLETEHRMTWD
PAQPPRDLTEAFLAEMEKAAGNPESSEFNENLRMVVADLFSAGMVTSTTTLAWGLLLMI LHPDVQ
RRVQQEIDDVIGQVRRPEMGDQARMPYTTAVIHEVQRFGLVPLGVTHMTSRDIELQGFLIPKGT
TLFTNLSSVLKDEAVWEKPFRRFHPEHFLDAQGHFVKPEAFLPFSAGRRAACLGEPLARMELFLFFT
CLLQRFSSVVPAGQPRPSHHGVFAFLVTPSPYELCAVPR

>2D19_MARMOSETii_H9KVS3

MGL---

DALVPLAVTVAIFVLLVDLMHRRQRWAARYPPGPMPLPLGLGNLLHVDKNTPNPNSFNQLRRRFGDV
FSLQLAWTPVVVLNGLAAREALVTRGEDTADRPPVPIITEMLGFGPHSQGLFLARYGPAWREQRR
FSVSTLRNLGLGKKSLEQWVTEEAATYLCAAFADHAGRPFRPNGLLDKAVSNVIASLTCRRRFEYN
DPCLLRLLDLTMEGLKEESGLLREVLNAIPVLLRIPGLAGKVLRSQKAFLAQLDELLETEHRMTWD
PAQPPRDLTEAFLAEMEKTGNPESSFNDDNLRLLVADLFSAGMVTSTITLAWGLLLMI LHPDVQ
RRVQQEIDDVIGRVRPEMGDQTYMPYTTAVIHEVQRFADIVPLGVTHMTSRDIEVQGFIPKGT
TLFTNLSSVLKDEANWEKPFRRFHPEHFLDAQGRFVKPEAFLPFSAGRRAACLGEPLARMELFLFFT
CLLQRFSSVVPAGQPRPSPHGVFAFLVTPSPYELCAVPR

>2D6_HUMAN_P10635

MGL---

EALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGGLPLPLGLGNLLHVDKNTPYCFDQLRRRFGDV
FSLQLAWTPVVVLNGLAAREALVTHGEDTADRPPVPIITQILGFGPRSQGVFLARYGPAWREQRR

F SVSTLRNLGLGKKSLEQWVTEEAACLC AAFANHSGRPFRPNGLLDKAVSNVIASLTCGRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAV PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRIVVADLF SAGMVT TSTTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLITNLSSVLKDEAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVSP SPYELCAVPR

>2D6_BONOB O_Q2XNC9

MGL---

EALVPLAVIVTIFLLLVDLMHRRQRWAARYPPG PLPLPGLGNLLHVDFQNTPYCFDQLRRRFGDV
 FSLQLAWTPVVVLNGLA AVREALVTHGEDTADRPPVPITQILGFGPRSQGVFLARYGPAWREQRR
 FSVSTLRNLGLGKKSLEQWVTEEAACLC AAFANHSGRPFRPNGLLDKAVSNVIASLTCGRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAV PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRIVVADLF SAGMVT TSTTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLFTNLSSVLKDEAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVTP SPYELCAVPR

>2D6_CHIMP_Q2XNC8

MGL---

EALVPLAVIVTIFLLLVDLMHRRQRWAARYPPG PLPLPGLGNLLHVDFQNTPYCFDQLRRRFGDV
 FSLQLAWTPVVVLNGLA AVREALVTHGEDTADRPPVPITQILGFGPRSQGVFLARYGPAWREQRR
 FSVSTLRNLGLGKKSLEQWVTEEAACLC AAFANHSGRPFRPNGLLDKAVSNVIASLTCGRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAI PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRIVVADLF SAGIVT TSTTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLFTNLSSVLKDKAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVTP SPYELCAVPR

>2D7_CHIMP i_H2R1S5

MGL---

EALVPLAVIVAIFLLLVDLMHRRQRWAARYPPG PLPLPGLGNLLHVDFQNTPYCFDQLRRRFGDV
 FSLQLAWTPVVVLNGLA AVREALVTHGEDTADRPPVPITQILGFGPRSQGVFLARYGPAWREQRR
 FSVSTLRNLGLGKKSLEQWVTEEAACLC AAFANHSGRPFRPNGLLDKAVSNVIASLTCGRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAI PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRIVVADLF SAGIVT TSTTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLFTNLSSVLKDKAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVTP SPYELCAVPR

>2D7_CHIMP ii_H2QLT8

MGL---

EALVPLAVIVTIFLLLVDLMHRRQRWAARYPPG PLPLPGLGNLLHVDFQNTPYCFDQLRRRFGDV
 FSLQLAWTPVVVLNGLA AVREAMVTRGEDTADRPPAPIYQVLGFGPPPTGVILARYGPAWREQRR
 FSVSTLRNLGLGKKSLEQWVTEEAACLC AAFAD EAGRPFRPNGLLDKAVSNVIASLTCERRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAI PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRIVVADLF SAGIVT TSTTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLFTNLSSVLKDKAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVTP SPYELCAVPR

>2D7_CHIMP iii_H2RB23

MGL---

EALVPLAVIVAIFLLLVDLMHRRQRWAARYPPG PLPLPGLGNLLHVDFQNTPYCFDQLRRRFGDV
 FSLQLAWTPVVVLNGLA AVREALVTHGEDTADRPPVPITQILGFGPRGQGVFLARYGPAWREQRR
 FSVSTLRNLGLGKKSLEQWVTEEAACLC AAFANHSGRPFRPNGLLDKAVSNVIASLTCGRFFEYD
 DPRFLRLLDLAQEG LKEESGFLREVLNAI PVLLHIPALAGKVLRFQKAF LTQLDELLTEHRMTWD
 PAQPPRDLTEAFLAEMEKAKGNP ESSFN DENLRMVVADLF LAGMVT TSVTLAWGLLLMI LHPDVQ
 RRVQQEID DVIGQVRRPEMGDQAHPYTTAVIHEVQRF GDIVPLGVTHMTSRDIEVQGFRI PKGT
 TLITNLSSVLKDEAVWEKPF RFHPEHFLDAQGHFVKPEAF LPFSAGRRACLGEPLARME LFLFFT
 SLLQHF SF SVPTGQPRP SHHGVFAFLVTP SPYELCAVPR

>2D42_RHESUS_MACAQUE_Q6GUQ5

MEL---

DALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFKNTPYCFDQLRRRFGDV
FSLQLAWTPVVVLNGLAAREALVTCGEDTADRPPVPIYQVLGVGPRSQGVFLARYGPAWREQRR
FSVSTLRNLGLGKKSLEQWVTEEAACLCFAFTDQAGRPFPRNSLLDKAVSNVIASLTYGRRFEYD
DPRFLRLFDLTHEALKEESGFLREVLNAVPLLLRIPGLAGKVLRSQKAFLTQLDELLEHRMTWD
PAQPPRDLTEAFLAEMEKAKGNPESSFNEENLRIVVADLFSAGMVTSTTLAWGLLLMI LHPDVQ
RRVQQEIDNVIGQVRRPEMGDQARMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIELQGFLIPKGT
TLFTNLSSVLKDEAVWEKPFQFYPEHFLDAQGHFVKPEAFPLPFSAGRRACLGEPLARMELFLFFT
CLLQRFQFSPVAGQPRPSHHGVFAFLVTPSPYELCAVPR

>2D44_CRAB_EATING_MACQUE_D8KWQ4

MGL---

DALVPLAMTVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFKNTPYCFDQLRRRFGDV
FSLQLAWTPVVVLNGLAAREALVTCGEDTADRPPVPIYQVLGIGPRSQGVFLARYGPAWREQRR
FSVSTLRNLGLGKKSLEQWVTEEAACLCFAFADQAGCPFRPSGLLDKAASNVIASLTGCRFEYD
DPRFLRLNLNAQEGKKEESGFLREVLNAVPLLLRIPALAGKVLRSQRAFLTQLDELLEHRMTWD
PAQPPRDLTDAFLAEMEKAKGNPESSFNEENLRMVVADLFSAGMVTSTTLAWGLLLMVLHPDVQ
RRVQQEIDNVIGQVRRPEMRDQARMPYTTAVIHEVQRFQDIVPLNMPHMTSRDIEVQGFIPKGT
TLFTNLSSVLKDEAVWEKSFRFQFYPEHFLDAQGRFVKPEAFPLPFSAGRRACLGEPLARMELFLFFT
CLLQRFQFSPVAGRPRPSHSHVVGFVTPSPYELCAVPR

>2D45_FROGi_F6VCB3

MSLLSGCNVFTLGIIFTLLLLLLDFMKRRKP-

CTDFPPSPPSWPFVGNLLQMDFRDLHNSFKQLSKQYGDVMSLQVFWKSMVVLNGFEVIKEALIQK
SEDTADRPPFNLYEILGFVGNNAKAVVLANYGQSWKDLRRFTLSTLRDFGMGKKSLEERVREAGY
LCDAFQSEQGRPFDPQVLINTAVSNVICSIIIFGERFEYDDHKFQKLLRLTEEIVTSESQKVTQVI
TLFAWISKFPGLAKPFFQTRMQLHKYLQEIINEHKQTWDSGH-
TRDFIDAFILEMEKAKGVKDSNFNDQNLILLIADLFAAGTETTTTTLRWGLLFMMLYPDVQKKVQ
EEIDQVIGRTRKPTMGDVLQMPYTNVIAHEIQRYADIIPLSVPHMTYRDTHIKGFFIPKGTVMIT
NLSSVLKDEKVVWEKPFQFYPEHFLDRDGKFKVKREAFMAFSAARRVCLGEQLARMELFLFFSSLLQ
RFSFQIPDGEPREDPEFVYMQFPHPHYKICAKVR

>2D45_FROGi_i_Q5FW30

MSLLSGCNVFTLGIIFTLLLLLLDFMKRRKP-

CTDFPPSPPSWPFVGNLLQMDFRDLHNSFKQLSKQYGDVMSLRVFWKPTVVVLNGFEVIKEALIQK
SEDTADRPPFNLYEILGFVGNNAKAVVLANYGQSWKDLRRFTLSTLRDFGMGKKSLEERVREAGY
LCDAFQSEQGGPFDPHVLINTAVSNVICSIIIFGERFEYDDHKFLKLLCLIEESIKAESGPVPQII
SSLPWSSKVPGLARLFFQPRIHMLQYLQEIINEHKQTWDSGH-
TRDFIDAFMLEMKKAKGVKDSNFNDQNLILLTTADLFSAGSETTTTTLRWGLLFMMLYPDVQRKVQ
EEIDQVIGRTRKPTMGDVLQMPYTNVIAHEIQRYADIIPLSVPHMAYRDTHIKGFFIPKGTVMIT
NLSSVLKDEKVVWEKPFQFYPEHFLDRDGKFKVKREAFMAFSAARRVCLGEQLARMELFLFFFTSLLO
RFSFQIPDGEPCLREDPVFVFLQVPHDYKICAKVR

>2D45a_CLAWED_FROG_Q6DCR5

MSLLSGCNVFTLGIICTLCLLLLDYMKRRKP-

CKNFPPSPPSKPFVGNLLQLNFRNLNNSFKQLSKQYGDVMSLQVFWKPMVVVLNGLEVMEALIQK
SEDTADRPEFHVLEILGFVGNNAKAVVLANYGQSWKDLRRFTLSTLRDFGMGKKSLEERVREEAGY
LCAAFQSEQGRPFDPHILLNTAVSNVICSIIIFGERFEYDDHTFQKLLCLIEESVKAESGAVPQII
ASLPWSSKIPGLAKMFFQPRIHMLKYLQEIINEKHQQTWDSGH-
TRDFIDAFMLEMEKAKGVKDSNFNEQNLLTTADLFSAGSETTNTTLRWGLLFMMLYPNVQRKVH
EEIDHVIIGRTRKPTMGDVLQMPYTNVIAHEIQRYVDIVPLSVPHMTYRDTHIQGFFIPKGTVMIT
NLSSVLKDEKAWKPFQFYPEHFLDRDGKFKVKREAFMAFSAARRVCLGEQLARMELFLFFFTLLQ
RFSFQIPNGEPSPREDPVFVFLQLPHDYKMCACVR

>2D45ii_FROG_Q7SYW2

MSLLSGCNVFTLGIICTLCLLLLDYMKRRKP-

CTNFPPSPPSRPFVGNLLQVDLKNLHNSIKQLSKQYGDVISLQVFWKPMVVVLNGFEVMEALIQK
SEDIADRPTIYIFDIFGFGANNRGVMFANYGQSWKDLRRFTLSTLRDFGMGKKSLEERVGEEAGY
LCAAFQSEQGRPFYPNVLLNTAVSNIICSIIIFGERFEYDDHKFQKLLSLTEEILISGETMPQVL
CLLPWSAKFPSLAKRFFKPRISMEKYLKEIINEHQQTWDSGH-
TRDFIDAFILEMEKEKAVKDSNFNEENLQTLTIADLFSAGTETTSSTLRWGLLFMMLYPDVQRKVN
AEIDQVIGRTRKPTMGDVQMPYTNVIAHEIQRYADIIPLSVPHVTYRDTHIKGFFIPKGTILIMIT

NLSSVLKDERVWEKPFQFYPEHF LDRDGKFKVREAFMAF SAGRRVCLGEQLARMTLFLFFTSLLQ
HFSFQIPDGEPSPREDPVIVYNQIPH DYKICAKVR
>2D6_AMUR_LEOPARD_CATi_E3VVX2
MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITS LT YG
RRFEYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLLSELVQE
HRITRDPTQPPQDLTDAFLDEIEKARGNP ESSFN DKNMLMTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLVPLGMPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAVPR
>2D6_AMUR_LEOPARD_CATii_E3VVX3
MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITS LT YG
RRFEYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLLSELVQE
HRITRDPTQPPQDLTDAFLDEIEKARGNP ESSFN DKNMLMTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLVPLGMPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAVPR
>2D6_BOB_CATi_E3VVX0
MGLLTGDVGLPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITS LT YG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVLSSQKALMTLLNELVQE
HRITRDPTQPPQDLTEAFLDEIEKARGNP ESSFN DENMLMTADLFLAGIVSTSATLTWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLAPLGVP HMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAVPR
>2D6_BOB_CATii_E3VVW8
MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITS LT YG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVLSSQKALMTLLNELVQE
HRITRDPTQPPQDLTEAFLDEIEKARGNP ESSFN DENMLMTADLFLAGIVSTSATLTWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLAPLGVP HMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAVPR
>2D6_CATii_E0D568
MGLLI G DVLGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLCAAFADHAGRPFSPKALLNKAVSNVITS LT YG
RRFGYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLLNELVQE
HRITRDPAQPPQDLTDAFLDEIEKARGNP ESSFN DENMLMTADLFLAGMLSTSTTLAWALLLMI
LHPDVQRRVQQEIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLVPLGMPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFMPFSAGRRVCLGEPLARME
FFLFFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAVPR
>2D6_CHEETAHi_E3VVW0
MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLFFYFDRLR
RRFGDVFSLQLAWTPMVVLNGLEAVREALVYHSED TADRPPMPIYEH LGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFG LGKKSLEQWVTEEASCLSAAFADHAGRPFSPRALLNKAVSNVIASLT YG
RRFEYDDPRLHKLIDGALQGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNP ESSFN DENMLMTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRF GDLAPLGVP HMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRICLGEPLARME
FFLLFTCLLQRF SFSVPVGQPRPSDHGVFSFLMIPPPYQLCAMPR
>2D6_CHEETAHi_E3VVW3

MGLLTGDVLGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
RRFGDVFSLQLAWTPMVVLNGLEAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLSAAFADHAGRPFSPRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALQGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRICLGEPLARME
FFLFTCLLQRFSSVPPVGQPRPSDHGVFSFLMIPPPYQLCAMPR

>2D6_CHEETAHiii_E3VWV1

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
RRFGDVFSLQLAWTPMVVLNGLEAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLSAAFADHAGRPFSPRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALQGLQEDSGFVREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRICLGEPLARME
FFLFTCLLQRFSSVPPVGQPRPSDHGVFSFLMIPPPYQLCAMPR

>2D6_FERRET_M3YY17

MGLLSGVALGPLAVAVAFLLLVDLMHRRRCWAARYPPGPVRMPLLGNLLQMDFQNALCYFDQLR
QRFGNVFTLYLAWTPFVVLNGLDAVREALVYHSEDADRPPMPIYHLGFGKQYSQGVILARYGRE
WREQRRFSLSALRNFGLGKKSLEQWVTEEASSLCAAFADYTGRPFSPRALLDKAVSNVISSLTFG
RRFEYDDPRLKLLDRLHAGLEEDQGLLEVLNAIPVLLHIPGLAGKLFSAQKAFMTLIDELVRE
HRTSRDPTQPPRDLTDAFLDEVEKAKGNPKSSFNENMLSMVTSDLFVAGMVSTSTTLTWALLLMI
LHPDVQRRVQREIDEVIGQARQPEMGDQAHMPFTMAVVHEVQRYADIVPLGVPHMTTRDIKVQGF
LIPKGTTLITNLSSVLKDKTAWKKPFRFDPEHFLDAEGRFVKQEAFLPFSAGRRVCLGEPLARME
LFLFTTRLLQRFDFSVPVGQPRPSDHGTFSLVTPAPYQLCAVPR

>2D6_LEOPARD_CATi_E3VVX1

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITSLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALLTLLSELVQE
HRITRDPTQPPQDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGLVPLGMPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRVCLGEPLARME
FFLFTCLLQRFSSVPPVGQPRPSDHGVFSFLMIPPPYQLCAVPR

>2D6_LION_E3VVX7

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFRDITLIFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFAREALNSIPVLLHIPGLADKVFSSQKALMTLLNELVQE
HRITWDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRICLGEPLARME
FFLFTCLLQRFSSVPPVGQPRPSDHGVFSLLMIPPPYQLCATPR

>2D6_MOUNTAIN_LIONi_E3VVX4

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFLPFSAGRRICLGEPLARME
FFLFTCLLQRFSSVPPVGQPRPSDHGVFSFLMIPPRYQLCAVPR

>2D6_MOUNTAIN_LIONii_E3VVX6

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDTLIFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSPRALLNKAVSNVITSFTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE

HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNDENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEIQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAVPR

>2D6_PUMA_E3VVX5

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQRVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDGALKGLQEDSGFAREALNSIPVLLRIPGLADKVFSSQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGNPESFNDENMLMVTADLFLAGMVSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPVQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCAVPR

>2D6_SIBERIAN_TIGERi_E3VY1

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGRE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFVREALNSIPVLLNIPGLTDKVF SNQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGSPESFNDENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCATPR

>2D6_SIBERIAN_TIGERii_E3VY0

MGLLTGDALGPLAIAVALFLLLVDLLHRRRFWAARFPPGPMPLPLLGNLLQMDFQDITLFYFDRLR
RRFGDVFSLQLAWTPVVVLNGLEAVREALVYHSEDADRPPMPIYHLGFGPRSQGVFMARYGHE
WREQRRFALSTLRNFGLGKKSLEQWVTEEASCLCAAFADHAGRPFSRALLNKAVSNVIASLTYG
RRFEYDDPRLHKLIDRALKGLQEDSGFVREALNSIPVLLNIPGLADKVF SNQKALMTLLNELVQE
HRITRDPTQPPRDLTDAFLDEIEKARGSPESFNDENMLMVTADLFLAGMLSTSATLAWALLLMI
LHPDVQRRVQREIDEVLGPEQRPAMADQTRMPFTMAVMHEVQRFGLAPLGVPHMTSRDIEVQGF
LIPKGTTLITNLSSVLKDKTVWKKPFRFHPEHFLDAQGQFVKQEAFFMPFSAGRRICLGEPLARME
FFLFFTCLLQRFSSVPGQPRPSDHGVFSFLMIPPPYQLCATPR

>2D49_CHICKEN_F1NJG4

MTLLLWSNISVLGVFLTFTVFTILVDFMKRRKKW-
SRYPPGPMPLPFVGTMPYVNYNPHLSFEKFRKKFGNIFSLQNCWTNVVVLNGYKTVKEALVNKS
EDFADRPYMPVYEHLYGHKSEGLVLARYGHLWKELRKFTLTTLRNFGMGKKSLEERVTEEAGFL
CSAISSEGGHPFDPRFLVNNAVCNVICTITYGERFDYGDKTFKLLTLFENSLNEEAGFLPQLLN
VAPVLLRIPGLPQKIFPCQKAYVDF TQMLIDKHKETWNPAY-
IRDFTDAFLKEMAKGKEAEENGFNKSNLT LV TADLLVAGSETTATTLRWAF LFMLLYPEIQSKVH
KEIDKVIGRNRPTMADQVNMPYTNVIHEVQRF GDVVPMLPHMTYRDTELQGF FIPKGTIIIT
NLTSVLKDETAWKKNPEFYPEHFLNENGQFVRPEAF LPFSAGRRACLGEQLTRMELF IFFTLLMQ
KFTFVFPEDQPRPREDSHFAFTNSPHPYQLRAVPR

>2D6_DUCK_U3IEM6

MTLLLWSNIAVLGVFLTFTVFTLVDFMKRRKKW-
SRYPPGPMPLPFVGTMPYIDFNNPHLSFNKFHKKFGNIFSLQNCWTNVVVLNGYKTVKEALVHKA
DDFADRPYFPTYEHVGYGDKSEGIVVARYGHVWKEIRKFTLTSTLRDFGMGKKSLEEQVTEEAGFL
CSAIAEEGRPFDLRFLVNNAVCNMICSIIYGRFDYGDDEVFRKMVALFENSLSEEAGFLTQFLN
VMPILLRIPGLPQKVFRGQKAYMDFIDVLIQKHMETWNPEH-
TRDFTDAFLKEMSKGKEAEENGFNYNLRLVTADLFAAGSDTTSATLRWAF LYMLLHPGTQSKI H
KEIDMVIGRERSPTMMDQVNMPYTNVIHEAQRCDIAPVGLPHMAYRDTELQGYFIPKGTTVIT
NLSSVLKDES VWKKPNEFYPEHFLDVNGQFVKPEAFMPFSAGRRVCLGAQLARME LF IFFTLLQ
QFTFMLPEDQPRPRTDGYFALIKCPHPFKLRAIPR

>2D6_TURKEY_G1NHK6

MTLLLWSNISVLGVFLTFTVFTLVDFMKRRKKW-
SCYPPGPVSLPFVGTMPYVNYNPNILSFEKFRKKFGNVFSLQNCWTNVVVLNGYNTVKEALVNKS
EDFADRPYMPIYHLGYGHKSEGLVLAKYGHVWKEIRKFTLTITLRNFGMGKKSLEERVTEEAGFL
CSAISSEGGRPFDLHLLVNNAVSNVICTITYGERFEYGD ETFKLLTLFENSLNEEIGLLPELLO
VVPILLRIPGLPQMIFRCQKEYIDFTEMLIEKHKETWNPAY-
TRDFTDAFLKEMAKGNEAEVNGFNRTNLSFVTSLLAAGSQTTSTTLRWAF LFMLLYPEIQSKVH

KEIDKVIGRNRPPTMADQANMPYTNNAVIHEVQRCGDIIPMGLPHMTYRDTKLQGGFFIPKGTTIIT
 NLTSVLKDETVWKKPNEFYPEHFLNENGEFVRPEAFLLPFSAGRRACLGEQLAKMELFIFFTTLMQ
 KFTFVIPEDQPRPREDDGHFAFINSHPYQLQAVPR

>2D6_TURTLE_K7FT09

MELLLSKILPALGIALIVFALLTDFVKKRKS-
 SRYPPGPTSLPFIGTMLQIDYHNPFLSFTQLSKKFGNVYSLQNCWNNLVVVNGFQAIREALVQKA
 EDFADRPPLPIYKHLGYGKNSEGLVVARYGHAWKEQRRFTLSTLRNFGMGKKSLEQRVIEETGFL
 CSAICSEKGRSFDPHFLINNAVSNNVICFLVYGERFDYNNKKFQRLLRVFEEALNEESGFLPQFLI
 VVPWLLHIPGVPQTTFVQSQAQLLAFTDELVEKHATWNPQT-
 KRDFTDVFLHEIEKAKEHVGSSFNENNQRLVTVDLFSAGTETTSTTLRWALLYILLYPDVQSKVQ
 EEIDKVIGRDRSPQMEDQVNMPYTNNAVIHETQRYGDIVPVAVPHMTYRDTLQGGFFIPKGTTVIT
 NLSSVLKDETVWEKPHQFYPEHFLDADGQFVKPEAFLLPFSAGRRVCLGEQLARMELFLFFTCLLQ
 HFTFCLPEGQPRPREDDGHFAATLAPYPYQIQAVPR

>2D25_PIGi_F1SRG4

MGLLTGGLLGILALAMVIFLLLVDLMHRRSRWAPRYPPGPMPLPGLGNLLQVNFQDPRLSFIQLR
 RRFQDVFSLQQIWRPVVVLNGLAAVREALVSHSHETS DRPPVFIEHLGYGPRSEGVI LARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEEWVTQEASCLCAAFADQAGRPFSPNNLLNKAVSNVIASLTFA
 RRFQYNDPRMLKLLDLVLEGLKEEVGLMRQVLEAMPVLRHIPGLCAKLFPRQKAFVLMIDELITE
 HKMTRDLAQPPRDLTDAFLDEMKEAKGNPESSFNENLRLVVADLF SAGMI TTSTTLAWALLLMI
 LHPDVQRRVQQEIDEVIGHARQPEIKDQALMPFTLAVLHEVQRFQDIVPLGVAHMTSRDIEVQGF
 LIPKGTTLITNLTSVLKDETVWKKPFRFYPEHFLDAQGHFTKQEAFFMPFSAGRRSCLGEPLARME
 LFLFFTLLQRFSSVPTGQPRPSDHGVFAFLFLPSPYQLCAVPR

>2D25_PIGii_O46658

MGLLTGDLGLGILALAMVIFLLLVDLMHRRSRWAPRYPPGPMPLPGLGNLLQVNFQDPRLSFIQLR
 RRFQDVFSLQQIWRPVVVLNGLAAVREALVSHSHETS DRPPVFIEHLGYGPRSEGVI LARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEEWVTQEASCLCAAFADQAGRPFSPNNLLNKAVSNVIASLTFA
 RRFQYNDPRMLKLLDLVLEGLKEEVGLMRQVLEAMPVLRHIPGLCAKLFPRQKAFVLMIDELITE
 HKMTRDLAQPPRDLTDAFLDEMKEAKGNPESSFNENLRLVVAHLF SAGMI TTSTTLAWALLLMI
 LHPDVQRRVQQEIDEVIGHVRQPEIKDQALMPFTLAVLHEVQRFQDIVPLGVAHMTSCDIEVQGF
 LIPKGTTLITNLTSVLKDETVWKKPFRFYPEHFLDAQGRFTKQEAFFMPFSAGRRSCLGEPLARME
 LFLFFTLLQAFSFSVPTGQPCPSDHGVFAFLFLPSPYQLCAVPR

>2D6_SHEEP_F1CGV1

MGLLSGDTLGLPLAVAVLIFLLLVDLMHRRSRWAPRYPPGPTPLPVLGNLLQVDFDDPCPSFNQLR
 RRFQGNVFSLQQVWTPVVVLNGLAAVREALVYCNQDTS DRPPPAVYEHVGYGPRAEGVI LARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSPKDLLHKAVSNVIASLTFG
 CRFEYNDPRIVKLLDKLEDLLKEESGLVRQVVEAVPMLLRIPGLAAKVFPGLKAFMALIDELITE
 QKMTRDPTQPPRHLTDAFLDEMKEAKGNPESSFNENLRLVVTDLFTAGMVTTSTTLAWALLLMI
 LHPDVQRRVQQEIDEVIGKVRPEMGOQAFMPFTVAVVHEVQRFADIIPLGLPHMTSRDIEVQGF
 YVPKGTTLITNLSSVLKDETVWEKPFRRFHPHFLDAQGRFVKQEAFFIPFSAGRRACLGEPLARME
 LFLFFTSLQHFSSVPAQPRPSDHGVFAFLVTPAPYQLCAVPR

>2D50_HORSE_A8W4X2

MGLLTWDKLGPAVAVAVLIFLLLVDLMHRRQWRAPRYPPGPMPLPGLGNLLQVDFQDQTVSSFTRLR
 RRFQGNVFSLQLAWTPVVVLNGLAAVREALVYRSQDTADRPPPAVYEHVGYGPRAEGVI LARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSPDALLNKAVSNVIASLTFG
 GRFDYNDPHFLEILDLTEDILKEQSGFLPQVLNAIPMLLHIPGLVAKVFPQRAFMAQLDELVAE
 RRMTRDPAQPPRDLTDAFLDEVQKAKGNPESSFNDDNRLRVSDLF AAGMVTTSTALAWALLLMI
 LHRDVQRRVQQEIDEVIGQARRPEMGOQARMPTMAVVHEVQRFQDIAPVGAPHMTSRDIEVQGF
 LIPKGTTLITNLSSVLKDETVWKKPFRFHPHFLDAQGRFVKQEAFFIPFSAGRRACLGEPLARME
 LFLFFTCLLQHFSSVPAQPRPSDHGVFGTLVSPAPYQLCAEPR

>2D14_COWi_Q01361

MGLLSGDTLGLPLAVALLIFLLLVDLMHRRSRWAPRYPPGPTPLPVLGNLLQVDFEDPRPSFNQLR
 RRFQGNVFSLQQVWTPVVVLNGLAAVREALVYRSQDTADRPPPAVYEHVGYGPRAEGVI LARYGKA
 WREQRRFSVSTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSPMDLLNKAVSNVIASLTFG
 CRFEYNDPRIIKLLDLTEDGLKEEFNLVRKVVEAVPVLLSIPGLAARVFPQAQAFMALIDELIAE
 QKMTRDPTQPPRHLTDAFLDEVKEAKGNPESSFNENLRLVVADLF SAGMVTTSTTLAWALLLMI
 LHPDVQRRVQQEIDEVIGQVRPEMGOQALMPFTVAVVHEVQRFADIVPLGLPHMTSRDIEVQGF
 HIPKGTTLITNLSSVLKDETVWEKPFRRFHPHFLDAQGRFVKQEAFFIPFSAGRRACLGEPLARME
 LFLFFTSLQHFSSVPAQPRPSEHGVFAFLVTPAPYQLCAVPR

>2D14_COWii_G1K127

MGLLSGDTLGPLAVALLIFLLLLDLMHRRSRWAPRYPPGPTPLPVLGNLLQVDFEDPRPSFNQLR
RRFGNVFSLQQVWTPVVVLNGLAAVREALVYRSQDTADRPPPAVYEHLYGYPRAEGVILARYGDA
WREQRRFSLTTLRNFGGLGKKSLEQWVTEEASCLCAAFADQAGRPFSPMDLLNKAVSNVIASLTFG
CRFEYNDPRIIKLLDLTEDGLKEEFNLVRKVVEAVPVLLSIPGLAARVFPAQKAFMALIDELIAE
QKMTRDPTQPPRHLDLDAFLDEVKEAKGNPESSFNENLRLVVADLF SAGMVTSTTLAWALLLMI
LHPDVQRRVQQEIDEVIGQVRRPEMGDQALMPFTVAVVHVEERFADIVPLGLPQKASRDIEVQGF
HIPKGTTLITNLSSVLKDETVWEKPFRRFHPHF LDAQGRFVKQEAFIPFSAGRRACLGEPLARME
LFLFFTSLQLQHFSFSVPAGQPRPSEHG VF AFLVTPAPYQLCAVPR

>2D15_DOG_Q29473

MGLLTGDTLGPLAVAVAFIFLLLVDLMHRRRRWATRYPPGPTPVPMVGNLLQMDFQEPICYFSQLQ
GRFGNVFSLLEAWTPVVVLNGLAVREALVHRSEDADRPPMPIYDHLGLGPESQGLFLARYGRA
WREQRRFSLSTLRNFGLGRKSLEQWVTEEASCLCAAFAEQAGRPFPGGALLNKAVSNVISSLTYG
RRFEYDDPRLLQLLELTQQALKQDSGFLREALNSIPVLLHIPGLASKVFS AQKAIITLTNEMIQE
HRKTRDPTQPPRHLDLDAFVDEIEKAKGNPKTSFNEENLCMVTSDLFIAGMVSTSTTLTWALLLMI
LHPDVQRRVQQEIDEVIGREQLPEMGDQTRMPFTVAVIHEVQRFGDIVPLGVPHMTSRDTEVQGF
LIPKGTTLITNLSSVLKDEKVWKKPFRFYPEHF LDAQGHFVKHEAFMPFSAGRRVCLGEPLARME
LFLFFTCLLQRFSSVPAGQPRPSDHGVFTFLKVPAPFQLCVEPR

Figure 23

>BRAFL_C3XVJ1_1675_1-675

MDPDETQSVPLPATEEEFPFLGMDIFLVCMLGLTIYWFFRDRKAAATPQQLVKPLTVVLSTAS
R-
DRMKKTKKNVAVFYGSQTGTAEFFAGRLARDCIRYGLKGFVDPPECEMEELTRMGEVESSFALF
CLATYGEDPTDNAQEFFDWLQEDHDLSE-
LRYAVFGLGNKTYEHFNAMGIFVDKRLEEMGAVRLCTAGMGDDDANIEEDFVTWRETLWPAVCEH
FGIESVGEEGVMRQYTLEKHEDINQNEIFTGEPTRLNSLTQQKPPYDAKNPYMSSVKVNRELHKG
GDRSCMHIEFDISGSRIRYEAGDHVAVYPTNDPSMVEEIAHLLNVNLDEVFTLNNVDDEAAKKHP
FPCPTTYRTALLHYVDITHPPRTNVLHDIIDYASDPKDREFLTRITA-S-
AGKKEFAEWVTESHRDILQILKDLPSLKPPLDHLCELLPRLHARYYSSISSPKEHPERIHITAVL
VDYHISKIDRRIKGVATNWLATKIPNGPDG--
PRVPIYVRKSPFRLPFKTTTPVIMIGPGTGLAPFRGFIQERHHAKTEGKPIGDTVLFSGCRKKSE
DYIYQEELEGYEGDGTLSQLNMAF'SRDQAQKVYVQHLLIKQKSESVWNLL-
NDGAHLYVCGDARHMAHDVDMVLHEIVEKNGNMTQSQAKDFIKGLRNKGRYACDVWS
>CIOIN_F6T1J8_1673_1-671

MDAPVEEAAPFMSTLDLFLLSALVGFVGYWFFFRKDDTPSDDFKKLSVVQSTIRSESF-
DRMKKSGRNIAILYGSQTGTAEFFANRLAKDAQRYGMKALVLDPEECNVEDLPQLTEIQNSFLIF
CMATYGEDPTDNAQDFHDWLQGD TMLT-
GIRYAVFGLGNKTYEHYNSMGKFVDKRVEELGGERLYELGLGDDDTNIEEDFVTWKEGLWPAVCQ
RFGIEASADDGSFRAYSFTQYEDLDKKTITYTGEVARLHSDYNQRPPYDAKNPFLAPVQTNKELHK
GGGRSCMHIEFDITGSRIKYEAGDHVAVYPTNDEAMVEELGERLKTNLDTLFTMNNVDEDSSKRH
PFPCPTTFRTALLHYLDIASPPRTNVLSEFIEYTSDEKDKQFLKLLSSATPEGKKLYQEWMDSR
RTLLAVLKELPSCCPAMDHICELLPRLQARYYSSASSPRAHPNSIHICAVVVYKYNAGYDNFGV
ATNWLSGKVPAT-
GTAIPRAPIYVRKSQFRLPFKVSHPVVMIGPGTGLAPFIFGFIQDRAYHREQGKDVGKTILYFGCR
KRSEDFIYEEMLENWKSEEVLTELNLAF'SRDAEKKTYVQHLLIAQNKDSLWEVI-
EKHGHIIYVCGDARHMARDVHDAIVTIVEEKGDKSHQQAIDFVKGLMNKGYSADVWS
>ASTME_XP_007258471.1_1644_1-644
MADSHTEDSTQPEL-
MEDGEPFLSNLDIFLFLSLIAGLIYWMFRKKAETPEF'KKLETPAPAARESSFIEKMKKTGKNI
VVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMSELSRLKEIDNSLAIFCMATYGEDP
TDNAQDFYDWLQESDADLDGVNFTV----SFS-----
SMGDDDGNNLEEDFVSWREQFWPAVCEHFGVEATGEESSIRQYELKVHEDLNMNKIYTGEVPLLF
FFF--
RPFDSKNPFLSPVTVNRKLNKGGDRHLMHLELDITGSKIRYESGDHVAVYPTNDSNIVNKLGEVL
GVDLDTVISLNNLDEESNKKHPFPCPTTYRTALTHYLDVTSCPTNVLYELAQYASDPKDQENMR
KMASSSPEGKALYQTWVLDSSRNILAILEDLPSLRPPIDHLCELLPRLQARYYSSASSSKVHPNS
IHICAVVVEYQTKTGRTNKG VATNWLKNKVP-
TDNGHKATVPMYIRKSQFRLPFKSSNPVIMIGPGTGIAPFMGFIQERAWLKQOGKEVGETVMYFG
CRHKNEDFLYQEELEAFEKEEVLTQLNVAF'SRDQEKKVYVQHLLKNNKEEVWRQIHKDNAHIYIC
GDARNMARDVQTVFYEIAEELGGMTHYTQAVDYIKKLMTKGRYSQDVWS
>DANRE_CONSENSUS_1674_1-674

MEDTEPQPVSDQDQTPDPLSSLDIFLFLSLIAGLLIYWFFFRKKAEPPEMKPFTLVTAPIRETS
FIEKMKKTNRISIVVFYGSQTGTGEEFANRLSKDAHRYGMKGMAADPEEYDMSELSRLKEIPKSMA
VFCMATYGEDPTDNAQEFYDWLQGTDDDLEGVNFAVFGNGKTYEHYNAMGKYTDKRLAELGGK
RVFDLGLGDDDSNLEEDFISWKEQFWPAVCEYFGVEATGEDSSIRQFELVVHNDINMNQVYTGM
GRLKSFTQKPPFDAKNPFLASVAVNRKLNKGGNRHLMHIELDITESKIRYDSGDHVAVYPTNDS
AMVNRIGERLGVLDLDAVISLKNLDEESNKKHPFPCPTTYRTALTHYLDINNMPRTNVLYELAQYA
SDPQEENMRKMASASPEGKALYQSWVLDSESNILAILEDLPSLNPPIDHLCELLPRLQARYYSS
ASSSKVHPHCIHICAVVIEYNTKTGRVFKGVATNWLKGKGV-
TDNGHKPTVPMYVRRSQFRLPFKPSNPVIMIGPGTGIAPFMGFIQERAWQKEQKDVGETILYFG
CRHSNEDFLYQEELEEFERAGVLTQLNVAF'SRDQEKKVYVQHLLKNNKQQLWKLIHTDNAHIYVC
GDARNMARDVHAASFSEIAEQEGGLTHTQALDYFKKLMTKGRYSQDVWS
>POEFO_XP_007564212.1_1673_1-673

MADMDPETATN---
 TAGEEEPLFSNLDLFLFLSLIVGLVIYWFMSRKQPEPIPEFKKLDTPPTTTRETSFVEKMKKTGKN
 IIVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEEYSMGELARLAEIENSLAIFCMATYGED
 PTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEELGAKRIFDLGQGGDD
 GNLEEDFISWREQFWPAVCEHFGEASGDESSIRQYELKVHTDVNMNKVYTGEIGRLKSFEVQKP
 PFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSELVNKLGEILG
 VDLDVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELAQYASDPKEQENMRK
 MASSSPEGKALYQNWVLDSRNILAVLEDMPSLKPPIDHLCCELLPRLQARYYSIASSSKVHPNSI
 HICAVVVEYKTKTGRVNKG VATNWLKNKLV---
 NGHKSTVPMYIRKSQFRLPFKASNPVIMIGPGTGIAPFMGFVQERGWLKQQGKEVGETVMYFGCR
 HKNEDYIYQEELEEA EKNSVLTQLNVAFSRDQDHKVYVQHHLKKNKEHIWKLIHTDNAHIYVCGD
 ARNMAKDVQMAFYEIAEEYGSMTRTQATDYVKKLMTKGRYSQDVWS
 >POERE_CONSENSUS_1673_1-673
 MADMDPETATN---
 TAGEEEPLFSNLDLFLFLSLIVGLVIYWFMSRKQPEPIPEFKKLDTPQTTRTSFVEKMKKTGKN
 IIVFYGSQTGTAEFANRLAKDAQRYGMKGMAADPEEYSMGELGRLAEIENSLAIFCMATYGED
 PTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGNYVDKRLEELGAKRIFDLGLGDDDD
 GNLEEDFISWREQFWPAVCEHFGEASGDESSIRQYELKVHSDVNMNKVYTGEIGRLKSFEVQKP
 PFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSELVNKLGEILG
 VDLDVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELAQYASDPKEQENMRK
 MASSSPEGKALYQNWVLDSRNILAVLEDMPSLKPPIDHLCCELLPRLQARYYSIASSSKVHPNSI
 HICAVVVEYNTKTGRVNKG VATNWLKNKLV---
 NGHKSTVPMYIRKSQFRLPFKASNPVIMIGPGTGIAPFMGFVQERGWLKQQGKEVGETVMYFGCR
 YKNEDYIYQEELEEA EKNSVLTQLNVAFSRDQDHKVYVQHHLKKNKEHIWKLIHTDNAHIYVCGD
 ARNMAKDVQTAFFEIAEEYGSMTRTQATDYVKKLMTKGRYSQDVWS
 >XIPMA_XP_005803425.1_1673_1-673
 MADMDPETTTN---
 TAGEEEPLFSNLDLFLFLSLIVGLVIYWFMSRKQPEPIPEFKKLDTPPTTTRETSFVEKMKKTGKN
 IIVFYGSQTGTAEFANRLAKDSQRYGMKGMAADPEEYSMGELARLAEIENSLAIFCMATYGED
 PTDNAQDFYDWLQENDDELSGLNYTVFSLGNKTYEHYNAMGKYVDKRLEELGAKRIFDLGLGDDDD
 GNLEEDFVSWREQFWPAVCEHFGEASGDESSIRQYELKVHTDVNMNKVYTGEIGRLKSFEVQKP
 PFDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSELVNKLGEILG
 VDLDVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITNPPRTNVLYELAQYASDPKDQENMRK
 MASSSPEGKALYQNWVLDSRNILAVLEDMPSLKPPIDHLCCELLPRLQARYYSIASSSKVHPNSI
 HICAVVVEYKTKTGRFNKG VATNWLKNKLV---
 NGHKSTVPMYIRKSQFRLPFKASNPVIMIGPGTGIAPFMGFIQERGWLKQQGKEVGETVMYFGCR
 HKNEDYIYQEELEEA EKNNFLTQLNVAFSRDQDHKIYVQHHLKKNKEHLWKLIHTDNAHIYVCGD
 ARNMAKDVQTAFFEIAEEYGSMTRTQATDYVKKLMTKGRYSQDVWS
 >STEPA_XP_008280826.1_1678_1-677
 MADMDPEGTTTTEHMMEEEEESLSSLDLILFTLLAGIVLYFFMNRKKPEPIPEFKKLDTPA-
 TARETSFIEKMKKTGKN IIVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLSE
 ITNSLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFSLGNKTYEHYNAMGKYVDKRL
 EELGAKRIFDLGLGDDDDGNLEEDFVSWREQFWPAVCEHFGEALGDESSIRQYELKEYTDINMNK
 VYTGEIGRLKSFEVQKPPFDSKNPFLAPVTVNRRLNKAGDRHLMHLELDITGSKIRYESGDHVAV
 FPTNDTALVNKLQILGVLDLDMVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDITHPPRTNVLY
 ELAQYASDPKDQENMRKMASSSPEGKSLYQTWVLDSRNILHILEDMPSLKPPIDHLCCELLPRLQ
 ARYYSIASSSKVHPNSIHICAVVVEYKTKTNRVNKG VATNWLKNKLV-
 TDNGHKSTVPMYIRKSQFRLPFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVMYFG
 CRHKNEDYLYQEELEEA EKNAVLTQLNVAFSRDQEQKVYVQHHLKQNKENIWKLVHSDNAHIYVC
 GDARNMAKDVQTAFFEIAEELGGMTRTQATDYVKKLMTKGRYSQDVWS
 >ORYLA_XP_004075861.1_1623_1-623
 ---MEPETHTEP---
 VGDEEPLFSNLDLFLFLSLIIGLVVYFMSRKKPEPIPEFKKIDTPQVSQRETSFIEKMKKTGKN I
 IVFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYTMGELARMSEIKNSLAIFCMATYGEDP
 TDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEELGAKRIFDLGLGDDDD
 NLEEDFVSWREQFWPAVCEHFGEASADELSIRQYELKLHNDVNMNKVFTGEIGRLKSFEVQKPP
 FDAKNPFLAPVTVNRRLNKAGDRHLMHLELDITGSKLRYESGDHVAVFPTNDSALVNRLGEILGV
 DLDVVISLNNLDEESNKKHPFPCPTTYRTALHTHYLDIMNPPRTNVLYELAQYASEPKDQENMRKM

ASSSPEGKALYQTWVLDSCRNILAILEDMPSLKPPVDHLCELLPRLQARYYSIASSSKVHPNSIH
ICAVVVEYSTKTGRVNGKVATNWLKNKLV---
NGHKSTVPMFIRKSQFRLPFKATNPVIMIGPGTGIAPFMGFIQERGWLKEQGKEVGETVLYFGCR
HKNEDYIYQEELEDAEKSGVLTQLNVAFSRDQDHKVYVQHLLKKNKEHVWKLIHTDNAHLYVCG-

>CYNSE_XP_008308055.1_1673_1-673

---MDTETTQTMV---

EEEEPLLNLDFLFLSLIIIGLIYWFMSRKKSEPIPEFKKLETVTPSSRETSTFIEKMKKTGKTI
VFYGSQTGTAEFANRLSKDAQRYGMKGMAADPEEYDMGELSRLGEIPNSLAIFCMATYGEDPT
DNAQDFYDWLQENDDELSGLNYSIFGLGNKTYEHYNAMGKYVDKRLEELGAKRIYDLGMGDDGDN
IEEDFISWREQFWPAVCEHFGVESLGDDSSIRQYELKVHTDINMNKVFTGEIGRLKSFEVQKPPY
DAKNPFLAPVTVNRRLNKGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSMLVNKLGGQVLGVD
LDVVISLNNLDEESNKKHPFPCPSTYRTALTHYLDITHPPRTNVLYELAQYASDPKDQEHRLKMA
SSSPEGKALYHSWVLDACRNILAVLEDMPSLQPPIDHLCELLPRLQARYYSIASSSKVHPNSIH
ICAVVVEYKNTTGRTNKGVATNWLKNKLV-

TDNGHKSTVPMYIRKSQFRLPFKATNPVIMIGPGTGIAPFMGFIQESGWLKEQGKEVGETVMFFG
CRHKNEDYLYQEELEEAKEKNEVLTQLNVAFSRDQEQKVYVQHLMKNNKEHLWRLIHSDNAHIFIC
GDARNMAKDVQAIFYEIAEEVEGMTRTQATDYIKKLMTKGRYSQDVWS

>ORENI_I3KIU5_1677_1-677

MPDMDTETHTTEPA-

VDEEASLFSTMDLLFSVIVGFVYWFMSRKKPEPIPEFKKLDTPPTTRETSTFIEKMKKTGKNI
IVFYGSQTGTAEFANRLSKDAQRYGMKGMVADPEEYSMGELSRLSEIKNSLAIFCMATYGEDP
TDNAQDFYDWLQENDDELSGLNYTVFGLGNKTYEHYNAMGKYVDKKLEELGATRIFDLGMGDDG
NLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKEHTDINMNKVYTGEIGRLKSFEVQKPP
FDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSALVNKLGGQILGV
DLDVVISLNNLDEESNKKHPFPCPTTYRTALTHYLDISHPPRTNVLYELAQYASDPKDQENMRKM
ASSSPEGKALYQSWVLDASRNILAILEDMPSLRPPVDHLCELLPRLQARYYSIASSSKVHPNSIH
ICAVVVEYKTKTGRTNKGVATNWLKNKLV-

TDNGHKSTVPMYIRKSQFRLPFKSTNPVIMVGP GTGIAPFMGFIQERGWLKQQGKEVGETVLF
CRHKNEDYIYQEELEGAEKNGVLTQLNVAFSRDQDHKVYVQHLLTKNKEHVWKLIHTDNAHIYIC
GDARNMAKDVQTAFYEIAEEVGGMTRTQAVDYVKKLMTKGRYSQDVWS

>PUNNY_XP_005746976.1_1677_1-677

MPDMDPETHTTEPA-

VDEEASLFSTMDLLFSVIVGFVYWFMSRKKPEPIPEFKKLDTPPTTRETSTFIEKMKKTGKNI
IVFYGSQTGTAEFANRLSKDAQRYGMKGMVADPEEYSMGELSRLSEITNSLAIFCMATYGEDP
TDNAQDFYDWLQENDDELSGLNYTVFGLGNKTYEHYNAMGKYVDKRLEQLGATRIFDLGMGDDG
NLEEDFVSWREQFWPAVCEHFGVEASGDESSIRQYELKEHTDINMNKVYTGEIGRLKSFEVQKPP
FDSKNPFLAPVTVNRKLNKAGDRHLMHLELDITGSKIRYESGDHVAVFPTNDSALVNKLGGQILGV
DLDVVISLNNLDEESNKKHPFPCPTTYRTALTHYLDITHPPRTNVLYELAQYASDPKDQENMRKM
ASSSPEGKALYQSWVLDASRNILAILEDMPSLRPPVDHLCELLPRLQARYYSIASSSKVHPNSIH
ICAVVVEYETKTGRTNKGVATNWLKNKLV-

TDNGHKSTVPMYIRKSQFRLPFKSTNPVIMVGP GTGIAPFMGFIQERGWLKQQGKEVGETVLYFG
CRHKNEDYIYQEELEGSEKNGVLTCLNVAFSRDQDHKVYVQHLLTKNKEHVWKLIHTDNAHIYIC
GDARNMAKDVQTAFYEIAEQVGGMTRTQAVDYVKKLMTKGRYSQDVWS

>HAPBU_XP_005933201.1_1635_1-635

MGCVFSLPEDRRDAAERTPTTRETSTFIEKMKKTGKNIIVFYGSQTGTAEFANRLSKDAQRYGMK
GMVADPEEYSMGELSRLSEITNSLAIFCMATYGEDPTDNAQDFYDWLQENDDELSGLNYTVFGL
GNKTYEHYNAMGKYVDKRLEQLGATRIFDLGMGDDGDNLEEDFVSWREQFWPAVCEHFGVEASGD
ESSIRQYELKEHTDINMNKVYTGEIGRLKSFEVQKPPFDSKNPFLAPVTVNRKLNKAGDRHLMHL
ELDITGSKIRYESGDHVAVFPTNDSALVNKLGGQILGVDLDVVISLNNLDEESNKKHPFPCPTTYR
TALTHYLDITHPPRTNVLYELAQYASDPKDQENMRKMMASSSPEGKALYQSWVLDASRNILAILED
MPSLRPPVDHLCELLPRLQARYYSIASSSKVHPNSIHICAVVVEYETKTGRTNKGVATNWLKNKL
V-

TDNGHKSTVPMYIRKSQFRLPFKSTNPVIMVGP GTGIAPFMGFIQERGWLKQQGKEVGETVLYFG
CRHKNEDYIYQEELEGSEKNGVLTCLNVAFSRDQDHKVYVQHLLTKNKEHVWKLIHTDNAHIYIC
GDARNMAKDVQTAFYEIAEQVGGMTRTQAVDYVKKLMTKGRYSQDVWS

>TAKRU_H2T155_1673_1-673

---MLRPPL--

RAHPWEEVEPLFSNLDIFLFLSLIVGIVIYWFMRKKPEPVPEFKKIETQAPSTRETSFIEKMKKM
GKNIIVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMAELSRSEIENSLAIFCMATYGE
EGDPTDNAQDFYDWLQENDDELSGLNYTVFALGNKTYEHYNAMGKYVDKRLEELGAKRIFDLGLG
DDGDNLEEDFVSWREQFWPACVCEHFGVEPLGDESSIRQYELKVHTDINMNKVYSGEIGRLKSFEV
QKPPFDSKNPFLAPVTVNRRNLNKGGERHLMHLELDITGSKIRYESGDHVAVFPTNDSALVNKLGG
VLGVDLDVVISLNNLDEESNKKHFPFCPTTYRTALHYLDITQPPRTNVLYELAQYATDGKDQEN
LRKMASSSPEGKSLYQNWVLDA CRN ILAILEDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVHP
NSIHC AVVVEYQTKTGRLNKG VATNWLNKLI-

SDNGHKSMVPMYIRKSQFRLPFKATNPVLMVGP GTGIAPFVGF IQER GWLKEQGKEVGETVMYFG
CRHK NEDYLYQEELEEA EKNAALTQ LNVAFSRDQEQKVYVQHLLKANKEHVWKLIHSENAHIYIC
GDARNMAKDVQLAFYEIAEEQGEMTRS QATDYIKKLMTKGRYSQDVWS

>TETNI_CAF91751.1_1713_1-678

MA-VKKVKVAG-RWRITQPSTAALAGTITKVCACMCVG-

YGCERRPLLWSPAATEPSKNAPSTRETSFIEKMKKMGKNIIVFYGSQTGTGEEFANRLSKDAQR
YGMKGMAADPEEYDMAELSRSEIENSLAIFCMATYGE G DPTDNAQDFYDWLQENDDELSGLNYT
VFALGNKTYEHYNAMGKYVDKRLEELGAKRIFDLGLGDDGDNLEEDFVSWREQFWPACVCEHFGVE
ALGDESSIRQYELKVHTDINMNKVYTGEIGRLNSFEVQKPPFDSKNPFLAPVTVNRRNLNKGGRH
LMHLELDITGSKIRYESGDHVAVFPTNDAAALVNKLGGVVLGVDLDVVISLNNLDEESNKKHFPFCP
TTYRTALHYLDITQPPRTNVLYELAQYASDGKDQENLRKMASSSPEGKALYQSWVLDSCRNILA
ILEDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPNSIHC AVVVEYQTKTGRLNKG VATTWL
KNKLI-

SDNGHKSTVPMFIRKSQFRLPFKASNPVLMVGP GTGIAPFMGFIQER GWLKEQGKEVGETVMYYG
CRHRMEDYLYQEELEEAERSGVVTQ LNVAFSRDQDHKVYVQHLLKANKEHVWKLINSENAHIYIC
GDARNMAKDVQLAFSEIAEEQGDMRS QATDYIKKLMTKGRYSQDVWS

>ONCMY_CONSENSUS_1677_1-677

MADVEAESSTQ-

PEAMEEEEEPLFSNLDLFLFTLIAGLI IYWFMSRKKTEPIPEFKKLDQPAPSTRETSFIEKMKKTG
RNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAADPEEYDMSELSRLAEIDNSLAIFCMATYGE
GDPTDNAQDFYDWLQETDGELDGVNYTVFALGNKTYEHYNAMGAYVDKRLEELGAKRVFDLGMGD
DDGDNLEEDFVTWREQFWPAMCEHFGVEASGEDSSIRQYELKEHNDINMNKVYTGE MGR LKS FETQ
KPPFDAKNPFLAPVTVNRRNLNKGDRHLMHLEVDITGSKIRYESGDHVAVYPTNDTAIVNRLGQI
LGVDLDTVISLNNLDEESNKKHFPFCPTTYRTALHYLDI IHPRTNVLYELAQYATDPKDQENM
RKMASSAPEGKALYQSFVLEDNRNILAILEDLP SLRPPIDHLC ELM PRLQARYYSIASSSKVHPN
SIHC AVVVEYTTKTGRLTKGVATTWLKNKLVA-

DNGHKSTVPMYIRKSQFRLPFKASNPVIMVGP GTGIAPFMGFIQER GWLKEQGKEVGETVLYFGC
RHKNEDYLYQEELEEA EKTGVI TKLNVAFSRDQEQKVYVQHLLRTNKEDLWRQIHTDNAHIYICG
DARNMARDVQTAFYEIAEELGGMTRTQATDYIKKLMTKGRYSQDVWS

>SALSA_C0H8T5_1678_1-676

MADVEAESSTQ-PEAMEEENPLFSNLDLFLFTLIAGLI IYWFMSRK-

KEPIPEFKKLDQPAPSTRETSFIEKMKKTGRNIVVFYGSQTGTGEEFANRLSKDAQRYGMKGMAA
DPEEYDMSELSRLAEIGNSLAIFCMATYGE G DPTDNAQDFYDWLQETDGELDGVNYTVFALGNKT
YEHYNAMGAYVDKRLEELGAKRVFDLGMGDDGDNLEEDFVTWREQFWPAMCEHFGVEASGEDSSI
RQYELKEHNDINMNKVYTGE LGR LKS FETQKPPFDAKNPFLAPVTVNRRNLNKGDRHLMHLEVDI
TGSKIRYESGDHVAVYPTNDTAIVNRLGQILGVDLDSVISLNNLDEESNKKHFPFCPTTYRTALT
HYLDI IHPRTNVLYELAQYATDPKDQENMRKMASSAPEGKALYQSFVLEDNRNILAILEDLP SL
RPPIDHLC ELM PRLQARYYSIASSSKVHPNSIHC AVLVEYTTKTGRLTKGVATTWLKNKLVA-

DNGHKSTVPMYIRKSQFRLPFKASNPVIMVGP GTGIAPFMGFIQER GWLKEQGKEVGETVLYCGC
RHKKEDYLYQGELEEA EKMGVI TKLNVAFSRDQEQKVYVQHLLRTNKEDLWRQIHTDNAHIYICG
DARNMARDVQTAFYEIAEELGGMTRTQAIDYIKKLMTKGRYSQDVWS

>SALTR_P19618_1601_1-600

KLDQPAPSTQETSFIEKMKKTGRNIVVFYGSQTGTGEEFANRLSKDAHRYGMGSMAADPEEYDMS
ELSR LAEIGNSLAIFCMATYGE G DPTDNAQDFYDWLQETDGQLSGVNYPVFALGDKTYEHYNAMG
AYVDKRLEELGAKRVFDLGMGDDGDNLEEDFVTWREQFWPAMCEHFGVEASGEDSSVRQYELKEH
NDINMNKVYTGE LGR LKS FETQKPPFDAKNPFLAPVTVNRRNLNKA GELH KMHLEVDITGSKIRYE
SGDHVAVYPTNNTVIVNRLGQILGVDLDSVISLNNLDEESNKKHFPFCPTTYRTALHYLDI IHP
RTNVLYELAQYATDLKDQENTDSMASSAPEGKALYQSFVLEDNRNILAILEDLP SLRPPIDHLC

ELMPRLQARYYSIASSSKVHPNSIHICAVLVEYT-----KGVATTWLK-----
-
YIRKSQFRLPFKASNPVIMVGP GTGIAPFMGFIQERGWLKESGKEVGETVLYCGCRHKEEDYLYQ
EELEQAHKKGALTCLNVAFSREQDQKVYVQHLLRKNKVDLWRQIHEDYAHYICGDARNMARDVQ
TAFYEIAEELGGMTRTQATDYIKKLMTKGRYSQDVWS
>LEPOC_XP_006640872.1_1670_1-670

MENQSPEPVAEETSFLSSVDLFLFSIISGVLIYWFFFRKKKEEIPDFKKISTVAPSTRETSFIEK
MKKTGKNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDMAELSRSEIDNSFAIFCM
ATYGECDPTDNAQDFYDWLQETD TDLSGVSYTVFGLGNKTYEHYNAMGKYVDKRLEELGAQRIFD
LGMGDDDDGNLEEDFVSWREQFWPAVCEHFGEATGEESSIRQYELVVHTDVNMNKVYTGMGRLLK
SFETQKPPFDAKNPFLASVTVNRKLNRRGSRHLMHLELDITGSKIRYESGDHVAVYPVNAAIVN
KIGEILGVDLDTVISLNNLDEESNKKHPFPCPTSYRTALHYLDISNPPRTNVLYELAQYATDPK
DQETLRKMASSSPEGKALYHQWVPDARNIVAILEDMPSLRPPDLHLCELLPRLQARYYSIASSS
KVHSNSVHICAVVVEYETKTGRINKGVATNWLKNKV-
TDNGHKSTVPMYVRKSQFRLPFKASNPVIMIGPGTGIAPFIGFIQERGWLKEQGKEVGETVLYYG
CRHKNEFDLYQEELEAFEKAGVITQLNVAFSRDQDQKVYVQHLLKANKEHLWKLIHHENAHYIC
GDARNMARDVQNTFYEIVTELGSMTHTQAVDYIKKLMTKGRYSQDVWS
>CALMI_V9KCB0_1678_1-678
MEAPELPVPEGLDSSAVSSESFFSFFDIFLFLSLIAAFTVYWFFFRKKKEEIPDYRPLQTPAATLR
ETSFIEKMKKTGKNIVVFYGSQTGTAEFANRLSKDTQRYGMRGMATDPEEYDMSELRLTEIEN
SLAMFCVATYGECDPTDNAQDFYDWLQETDADLSGVKYTVFGLGNKTYEHFNAMGKYLDKRLEEL
GAERIFDLGVGDDDDGNLEEDFITWREQFWPALCEYFGEATDEDSIRQYELLVHNDVNMNKVYT
GMGRLLKSYEIQKPPFDAKNPFLASVSVNRKLNKGGERHLMHLELDISGSKIRYDSGDHVAVFPT
NDAKIVTDIGKILNTDLDIVFSLNNLDEESNKKHPFPCPTTYRTALSHYLDITNPPRTNVLYELA
QYASDSKAQEFLRKMASSSDGKSQYLKWVEDRRNILAILQDLPSLRPPIDHLCELLPRLQARY
YSIASSSKVHADSVHVCVVVEYITKAGRTNRGVATTWLRHQIP-
PENGRHPTVPVYVRKSQFRLPFKPSTPIIMVGP GTGIAPFIGFIQERAWLQQQGKQLGETVLYCG
CRFQAEDFLYQEQLEQWQQEGVLTQLNVAFSRDQPNKVYVQHLLKKNKENVWHLIHDANAHYIC
GDAKNMARDVQNAFCEIVAELGNMTQTQAVDYVKKLMTKGRYSSDVWS
>LATCH_H3BGP9_1676_1-676
MDASNVESGPPVADANTET-
TIFSMVDVFLVSLISGLLVYWL VFRKKKEEIPDFKPIKPVALTPRETSFVEKMKKTGKNIVVFY
SQTGTAEELANRLSKDAQRYGMRGMAADPEEYQMCELSRLSEIDNSLAVFCMATYGECDPTDNAQ
DYFDWLQETDEDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLELGAQRMFDLGVGDDDDGNLEED
FISWREQFWPAVCEHFGEATGEDSSIRQYELVVHMDINMNKVYSGEMGRLLKSYEIQKPPFDAKN
PYLGVVSVNRKLNKGGERHLMHIELDIAGSKIRYESGDHVAVYPSNDIALVNEMGKILNTDLDTV
FSLNNLDEESNKKHPFPCPTTYRIALHYLDITNPPRTNVLYELAQYATNAQDQEF LRKMASSSP
EGKEQYLSWVVEARRNILAILQDLPSLRPPIDHLCELLPRLQARYYSIASSSKVHPNSIHICAVV
VEYETKTGRINKGVATNWLKNKVP-
TDNGHKSSVPMYVRKSQFRLPFKPSTPVIMIGPGTGIAPFIGFIQERAWQQQGKEVGETVLYYG
CRHKKEDYLYQEELEQFQGDGVLTQLNVAYS RDQAEKVYVQHLLKKNKEHIWKLIHQDNAHIYVC
GDARNMARDVQNAFYDIVMEFGEMTRTQAVDYVKKLMTKGRYSCDVWS
>ALPAC_CONSENSUS_1678_1-678
MGDSSVDAGATAAETVTEEVS LFTMDMILFSLIVGLLTYWFFFRKKKEEIPFTTKIQPMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETD VDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIYELGLGDDDDGNLEEDFITWREQFWPAVCEHFGEATGEESSIRQYELVVHPDIEAAKVYV
GMGRLLKSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVSQ LGEILGADLDVVM SLNNLDEESNKKHPFPCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQELRKMASSSGEGKELYLSWVVEARRHILAILQDSPSLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGIAPFMGFIQERAWLQQQGKEVGETLLY YGCRSDEDYLYREELAKFHKD
GALTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHDGGAHIYVCGDARNMARDVQNTFYDIVAE
QGAMEHTQAVDYVKKLMTKGRYSLDVWS
>CAMEL_CONSENSUS_1633_1-633
MGDSSVDAGATAAETVTEEVS LFTMDMILFSLIVGLLTYWFFFRKKKEEIPFTTKIQPMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN

ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAPVEHLGVEATGEESSIRQYELVVHPTDIEAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDSSLRPPIDRLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGIAPFMGFIQERAWLQQQGE PQ-----

VYVQHLLKRDKEHLWKLIDHGGAHIVYCGDARNMARDVQNTFYDIVAEQGAMEHAQAVDYVKKLM
TKGRYSLDVWS

>ORCA_XP_004268859.1_1678_1-678

MGDSSVEAGATASDTVAEEVSLFSTMDMIMFSLIVGLLTYWFLFRKKKDEVPEFTRIQTTTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAPVEHLGVEATGEESSIRQYELVVHTDIDVAKVYV
GEMGRKLSYENQKPPFDAKNAFLAVVTTNRKLNQGTERHFMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGHRALVPMFVRKSQFRL
PFKATVPVIMVGP GTGVAPF IG FVQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAKLHQD
GALTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIDHGGAHIVYCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>TURTR_XP_004326617.1_1679_1-679

MGDSSVEAGATASDAVAEEVSLFSTTMDMIMFSLIVGLLTYWFLFRKKKDEVPEFTRIQTTTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVSPCGLVQPPEPVSKRNSSGSHLSEL
SRQPHLEPGSSPLVPSLEEDFITWREQFWPAPVEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHFMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGHRALVPMFVRKSQFRL
PFKATVPVIMVGP GTGVAPF IG FVQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAKLHQD
GVLTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIDHGRGAHIVYCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>LIPVE_XP_007468859.1_1678_1-678

MADARVEAGATASDSAAEEASLFSATDVIAFSLIVGLLTYWFLFRKKDEVPEFTKIQTTPSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAPVEHLGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHFMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPSGENGRRALVPMFVRKSQFRL
PFKATVPVIMVGP GTGVAPF IG FVQERAWLRQQGKEVGQTLLYYGCRSDEDYLYREELAKLHQD
GVLTQLNVAFSREQPQKVYVQHLLKRDREHLWKLIDHGGAHIVYCGDARNMARDVQNTFCDVVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>MINKE_XP_007186623.1_1678_1-678

MGDSSVEAGATVSDTAAEEVSLFSTTDVIMFSLIVGLLTYWFLFRNKKDEVPEFTKIQTTTSSVK
DSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLSEIEN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAPVEHLGVEATGDESSIRQYELVVHTDIDAAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELIDISDKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETQSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATVPVIMVGP GTGVAPF IG FVQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAKFHQD
GALTQLNVAFSREQPQKVYVQHLLKRDGEHLWKLIDHGGAHIVYCGDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS

>PHYCA_XP_007104040.1_1678_1-678

MGDS SAEAGATASDTVAEEVSLFSTTDMI LFS L I VGL LTYWFLFRKKKDEVPEFTKIETTTSSVK
DSSFVEKMKKTGRN I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M V A D P E E Y D L G D L S S L S E I E N
A L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A R R I F D L G L G D D D G N L E E D F I T W R E Q F W P A V C E H L G V E A T G E E S S I R Q Y E L V V H T D I D A A K V Y V
G E V G R L K S Y E S Q K P P F D A K N P F L A V V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G A D L D V I M S L N N L D E D S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q G Q L R K M V S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y S S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T Q S G R V N K G V A T S W L R A K E P S G E N G R R A L V P M F V R K S Q F R L
P F K A T V P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A K F H Q D
G T L T Q L N V A F S R E Q P Q K V Y V Q H L L K R D R E H L W K L I H D G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
Q G A M E H A Q A T D Y V K K L M T K G R Y S L D V W S

>BOAR_NP_001123431.1_1678_1-678

MGDS NVDTGTTTSEMVAEEVSLFSATDMVLFSLIVGLLTYWFIFRKKKDEVPEFTKIETTTSSVK
DSSFVEKMKKTGRN I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L S D L S S L P E I E N
A L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E A D V D L S G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F D L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D M D T A V V Y T
G E M G R L K S Y E N Q K P P F D A K N P F L A V V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G T D L D I V M S L N N L D E E S N K R H P F P C P T T Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E Q L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K S G R V N K G V A T S W L R A K E P A G E N G R R A L L P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L Q E Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H A K
G A L T R L S V A F S R E Q P Q K V Y V Q H L L K R D K E H L W K L I H D G G A H I Y I C G D A R N M A R D V Q N T F C D I V A E
Q G P M E H A Q A V D Y V K K L M T K G R Y S L D V W S

>PIG_P04175_1678_1-678

MGDS NVDTGTTTSEMVAEEVSLFSATDMVLFSLIVGLLTYWFIFRKKKDEVPEFSKIETTTSSVK
DSSFVEKMKKTGRN I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L S D L S S L P E I E N
A L A V F C M A T Y G E G D P T D N A Q D F Y D W L Q E A D V D L T G V K Y A V F G L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F D L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L V V H T D M D T A V V Y T
G E M G R L K S Y E N Q K P P F D A K N P F L A V V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G T D L D I V M S L N N L D E E S N K R H P F P C P T T Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P S E Q E Q L R K M A S S S G E G K E L Y L S W V V E A R R H I L A I L Q D Y P S L R P P I D H L C E R L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V V V E Y E T K S G R V N K G V A T S W L R A K E P A G E N G R R A L V P M F V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L Q E Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A Q F H A K
G A L T R L S V A F S R E Q P Q K V Y V Q H L L K R D K E H L W K L I H D G G A H I Y I C G D A R N M A R D V Q N T F C D I V A E
Q G P M E H A Q A V D Y V K K L M T K G R Y S L D V W S

>BOVIN_CONSENSUS_1678_1-678

MADSNMDAGTTTSEMVAEEVSLFSTTDVILFSLIVGVMTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DRSFVEKMKKTGRN I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L S S L P E I E K
A L A I F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F A L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F D L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L M V H T D M D M A K V Y T
G E M G R L K S Y E N Q K P P F D A K N P F L A V V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G A D L D I I M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P T E H E Q L R K M A S S S G E G K E L Y L R W V L E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K T G R I N K G V A T S W L R A K E P A G E N G R R A L V P M Y V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A G F H K D
G A L T Q L N V A F S R E Q P Q K V Y V Q H L L K K D K E H L W K L I H E G G A H I Y V C G D A R N M A R D V Q N T F Y D I V A E
Q G A M E H A Q A V D Y V K K L M T K G R Y S L D V W S

>YAK_XP_005892891.1_1681_1-679

MADSNMDAGTTTSEMVAEEVSLFSTTDVILFSLIVGVMTYWFLFRKKKEEVPEFTKIQT TTTSSVK
DRSFVEKMKKTGRN I V F Y G S Q T G T A E E F A N R L S K D A H R Y G M R G M A A D P E E Y D L A D L S S L P E I E K
A L A I F C M A T Y G E G D P T D N A Q D F Y D W L Q E T D V D L S G V K Y A V F A L G N K T Y E H F N A M G K Y V D K R L E Q L
G A Q R I F D L G L G D D D G N L E E D F I T W R E Q F W P A V C E H F G V E A T G E E S S I R Q Y E L M V H T D M D M A K V Y T
G E M G R L K S Y E N Q K P P F D A K N P F L A V V T T N R K L N Q G T E R H L M H L E L D I S D S K I R Y E S G D H V A V Y P A
N D S A L V N Q L G E I L G A D L D I I M S L N N L D E E S N K K H P F P C P T S Y R T A L T Y Y L D I T N P P R T N V L Y E L A
Q Y A S E P A E H E Q L R K M A S S S G E G K E L Y L R W V L E A R R H I L A I L Q D Y P S L R P P I D H L C E L L P R L Q A R Y
Y S I A S S S K V H P N S V H I C A V A V E Y E T K T G R I N K G V A T S W L R A K E P A G E N G R R A L V P M Y V R K S Q F R L
P F K A T T P V I M V G P T G V A P F I G F I Q E R A W L R Q Q G K E V G E T L L Y Y G C R R S D E D Y L Y R E E L A G F H K D

GALTQLNVAFSREQPQKVYVQHLLKKDKEHLWKLIIHEGGCRLPSARDARNMARDVQNTFYDIVAE
QGAMEHAQAVDYVKKLMTKGRYSLDVWS
>PANHO_XP_005972302.1_1634_1-631
MGDSNVDGTGTTTPEAVAEVSLFSATDTILFSLIVGVMTYWFLFRKKKEEVPEFTKIQTATSSVK
DRSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEK
ALAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFALGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDMAKVYT
GEMGRLLKSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLRWVLEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATSWLRREEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQRP-----
-----QVYVQHLLKKDKEHLWKLIIHEGGAHIYVC-
DARNMARDVQNTFYDIAAEQGAMEQAQAVDYIKKLMTKGRYSLDVWS
>FELCA_M3VV27_1678_1-678
MGDSNVDASATVSETVAEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEIPEFTRIQPVTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATYGEDPTDNAQDFYDWLQETDMDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDVAKVYM
GEMGRLLKSYENQKPPFDAKNPFLAEVTNRKLNQGTERHLMHLELDISDSKLRYESGDHVAVYPA
NDSALVSQ LKILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKDVGETLLYYGCRRSDEDYLYREELTQFHKD
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
VGAMEHAQAVDYIKKLMTKGRYSLDVWS
>TIGER_XP_007074057.1_1628_1-628
MGDSNVDASATVSETVAEEVSLFSMTDMI LFSLIMGLLTYWFLFRKKKDEIPEFTRIQPVTSPVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATYGEDPTDNAQDFYDWLQETDMDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDVAKVYT
GEMGRLLKSYENQKPPFDAKNPFLAEVTNRKLNQGTERHLMHLELDISDSKLRYESGDHVAVYPA
NDSALVSQ LKILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEAGR HILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKCVGTGWGSLGLRVPG-----

YLLMPPPPSCPDARNMARDVQNTFYDIVAEVGAMEHAQAVDYIKKLMTKGRYSLDVWS
>CANFA_CONSENSUS_1678_1-678
MEDSSMDASATVSETVVEEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKDEIPEFTTIQPMTSSAK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATYGEDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRLLKSYENQKPPFDAKNPFLAAVTNRKLNQGTERHLMHLELDISDSKLRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPTEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYQTRSGRINKGVATSWLRRAKEPAGENGRRALVPMFVRKSQFRL
PFKAATPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELA QFHQD
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEAGAHYVCGDARNMARDVQNTFYDIVAE
VGAMEHAQAVDYIKKLMTKGRYSLDVWS
>LEPWE_XP_006733968.1_1678_1-678
MGDSHMDASATVSETVAEEVSLLSMTDMI LFSLIVGLLTYWFFRKKKDEVPEFTKIQTTS SVK
DSSFVEKMKKTGRNII VVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGS LPEIEN
SLAVFCMATYGEDPTDNAQDFYDWLQETNLDLTGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEECSIRQYALVVHPDIDPAKVYM
GEVGRLLKSYENQKPPFDAKNPFLAAVTANRKL NQGTERHFMHLELDISDSKLRYESGDHVAIYPA
NDSTLVSQ LKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QHASEPSEQEQLRRMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY

YSIASSSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVILVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAGFCQD
GSLTQLNVAVSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAD
GGAMEHAQAVDYIKKLMTKGRYSLDVWS
>WALRU_XP_004399140.1_1678_1-678
MGDSHMDASATVSESVAAEVSLLSMTDMI LFSLI VGLLTYWFFLRKKKDEVPEFTKI QPVTSSVK
DSSFVEKMKKTGRNI IVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLGSLPDIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDLDLTGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFLT WREQFWPAVCEHFGVEATGEECSIRQYGLVVHPDIDPAKVYV
GEVGR LKSYENQKPPFDAKNPFLAAVTANRKL NQGTERHFMHLELDISDSKLRYESGDHVAIYPA
NDSTLVSQ LGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QHASDPSEQEQLRRMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYKTKSGRINKGVATSWLRAKEPTGENGRRALVPMFVRKSQFRL
PFKATTPVILVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAGLHQD
GFLTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNAFYDIVAD
GGPMEHAQAVDYIKKLMTKGRYSLDVWS
>MUSPF_M3YX63_1678_1-678
MGDSSMDASATVSETVAAEVSLFSMTDMI LFSLI VGLMTYWFLFRKKKDEIPEFTKI QPMTSSVK
DSSFVEKMKKTGRNI IVFYGSQTGTAE EFANRLSKDAHRYGLRGMAADPEEYDLADLGSLPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHPDIDPAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKLRYESGDHVAIYPA
NDSTLVSQ LGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAVLQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKAGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRADEDYLYREELAGFQKD
GSLTQLSVAFSREQPHKVYVQHLLRRDKEHLWQLIHEAGAH IYVCGDARNMARDVQNTFYDIVAE
GGAMEHAQAVDYVKKLMTKGRYSLDVWS
>PANDA_CONSENSUS_1679_1-678
MGDSSMDASTTVSETVAAEVSLFSMTDMI LFSLI VGVLTYWLLFRKKKDEIPEFTKI QPVTSSVK
DSSFVEKMKKTGRNI IVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLGDLGSLPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDLDPAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKLRYESGDHVAIYPA
NDSTLVSQ LGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDAQVELLPRLQARY
YSIASSSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELA AFHRD
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWQLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
GGTMAHAQAVDYIKKLMTKGRYSLDVWS
>MYOBR_S7Q4E9_1731_1-678
MGDSNTDTSATGSETVAAEVSLFSMTDMI LFSLI VGLLTYWFFFRKKKDEIPEFTKI QAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGR LKSYENQKPPFDAKNPFLAVVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQ IGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELA QFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS
>MYOLU_XP_006088353.1_1675_1-666
MGDSNTDTSATGSETVAAEVSLFSMTDMI LFSLI VGLLTYWFFFRKKKDEIPEFTKI QAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEES-----
RDMKVYVVGEMGR LKSYENQKPPFDAKNPFLAVVT TNRKL NQGTERHLMHLELDISDSKIRYESG

DHVAIYPANDSALVNQIGIILGADLDVMSLNNLDEESNKKHFPFCPTSYRTALTYYYLDITNPPR
TNVLYELAQYASDPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCEL
LPRLQARYYSIASSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCRALVPMF
VRKSQFRLPFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYRE
ELAQFHKEGSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQN
TFYDIVSELGSMEHTQAVDYIKKLMTKGRYSLDVWS

>MYODA_CONSENSUS_1678_1-678

MGDSNTDTSATGSETVAEEVSLFSMTDMI LFS LIVGLLTYWFFFRKKKDEIPEFTKIQAMTPSAK
DSSFVEKMKKTGRNVIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQIGIILGADLDVMSLNNLDEESNKKHFPFCPTSYRTALTYYYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHANSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS

>EPTFU_XP_008139684.1_1678_1-678

MGDSNIDTSATGSETVAEEVSLFSMTDMI LFS LIVGLLTYWFFFRKKKDEIPEFTKIQAMTSSVK
DSSFVEKMKKTGRNVIVFYGSQTGTAEFANRLSKDAQRYGMRGMSADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVLDSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMAKVYV
GEMGRKLSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQIGIILGADLDVMSLNNLDEESNKKHFPFCPTSYRTALTYYYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHNSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGCKALVPMFVRKSQFRL
PFKATTPVIMIGPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GSLTQLNVAFSREQPQKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVSE
LGSMEHTQAVDYIKKLMTKGRYSLDVWS

>HORSE_B1PL89_1678_1-678

MGDSNMDASAPTSETVAEEVSLFSMMDMFLFS LIVGLLTYWFLFRKKKDEIPEFTKIQTTTTSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGSLSEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELLVHTDIDAAKVYV
GEMGRKLSYETQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYYYLDITNPPRTNVLYELA
QYASEPFEEQQLRKMASSSGEGKELYLTWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATTWLRAKEPAKENGRRALVPMFVRKSQFRL
PFKATTPVIMVPGTGIAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYRDELAQFHRD
GSLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGTMEHAQAVDYIKKLMTKGRYSLDVWS

>RHINO_XP_004442186.1_1678_1-678

MRDSNMDASTTMSETVAEEVSLFSMTDVLFS LIVGLLTYWFLFRKKKDEVPEFTKIQTTTTSVK
DSTFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLGSLPEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYELVVHTDIDTAKVYT
GEMGRKLSYENQKPPFDAKNPFLAVVTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYYYLDITNPPRTNVLYELA
QYASEPTEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLQAKEPAGENGRRALVPMFVRKSQFRL
PFKATTPVVMVPGTGIAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GSLTQLNVAFSREQAHKVYVQHLLQRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFSDIVAE
LGAMEHAQAVDYIKKLMTKGRYSLDVWS

>LOXAF_G3SMI9_1678_1-678

MGDSNVDTSATMSEAVTEEMSLFSMTDMI LFS LIVGLLTYWFLFRKKKDEI PAFTKIQPTTSSAK
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLAEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL

GARRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDIAKVYM
GEIGRLKSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPVEENGHRPLVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFIGFIQERAWLKQQGKDVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHKGGAHIYVCGDARNMARDVQNTFYDIVAE
LGDMHAQAQVDYIKKLMTKGRYSLDVWS

>TRIMA_XP_004385994.1_1681_1-678

MGDSNVDTSATVSEAVTEEMSLFSMMDVILFSLIVGLLSYWFMRKRKDEVVPFTKIQPATSSVT
ESSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLAEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQEANVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
GASRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDIAKVYT
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASDPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYKTKTDRVNKG VATSWLRAKEPVEENGHQALVP IFVRKSQFRL
PFKASTPVIMVPGGTGIAPFIGFIQERAWLKQQGKDVGETVLYYYGCRRSDEDYLYREELAQFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEASAHYVCGDARNMARDVQNTFCDIVAE
LGAMEHAQAQVDYIKKLMTKGRYSLDVWS

>CHRAS_XP_006859747.1_1682_1-678

MGDSNVDTSATMSEAVTEEVSLFSMTDMVLFSLIVGLLTYWFLFRKKKDEVPEFSKIQTITSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLAVFCMATYGECDPTDNAQDFYDWLQETDVELSGVKYAVFGLGNKTYEHFNAMGKYVDTRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDIAKVYT
GEMGRKLSYENQKPPFDAKNPFLAGVTVNRLKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NNSTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTTYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPVAENGHRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLQQGKDVGETVLYYYGCRRADEDYLYREELTQFHKD
GALTQLNVAFSREQAQKVYVQHLLKRDKEHLWKLIHEAGAHYVCGDARNMAKDVQNAFYDIVAE
FGAMEHTQAQVDYIKKLMTKGRYSLDVWS

>ECHTE_XP_004705681.1_1678_1-678

MGDT SADASTTVSEAVTEDVSLFSMTDMILFSLIVGLLTYWFLFRKKKDEIPEFTKIQTMTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGLAADPEEYDLADLSSLADIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETD TDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHSDMDIAKVYT
GEMGRKLSYENQKPPFDAKNPFLAGVT TNRKLNQGMERHLMHLELDISDSKIRYESGDHVAVYPA
NDSSIVKQLGETLGADLDVIMSLKNLDEDNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKG VATSWLRAKEPAGENGHRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAHFHKE
GTLTQLNVAFSREQAHKVYVQHLLKRDQEHLWKLIHEAGAHYVCGDARNMARDVQNTFYDLVAE
LGGMEHAQAQVDYIKKLMTKGRYSLDVWS

>ELEED_XP_006889753.1_1678_1-678

MGDS SADASTTVSEAVTDEM SLFSGTDMILFSLIVGLLTYWFLFRKKKDEIPEFTKIQTLTSSVK
DSSFVEKMKKTGRNII VFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLTEIEN
SLAVFCMATYGECDPTDNAQDFYDWLQETD VDL SGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDMARVYT
GEMGRKLSYENQKPPYDAKNPFLAGVT TNRKLNQGNERYLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKSGRINKG VATSWLRAKEPVGENGHRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGIAPFMGFIQERAWLKQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHDAGAHYVCGDARNMAKDVQNTFYDIVAE
FGAMEHAQAQVDYVKKLMTKGRYSLDVWS

>ORYAF_XP_007938439.1_1681_1-678

MEDSSMDTSTTVSEAVTEEMSLFSMTDMI LFS LIVGLLTYWFLRKKKDEVPEFSKIQTTS SVK
ESSFVEKMKKTGRNIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLGDLSSLTEIEN
SLVVF CMATY GEGDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFDLGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDTAKVYS
GEMGR LKSYENQKPPFDAKNPFLAGVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDPTLVNQLGEILGADLVNIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASDPSEQEHLRKMASSSGEGKELYLSWVVEARRNILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKEPAGENGHRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFIGFIQERAWLQQQKDVGETLLYYGCRRSNEDYLYREELAQFHKD
GTLTQLNVAFSREQPHKVYVQHLLRRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
FGAMEHTQAVDYVKKLMTKGRYSLDVWS

>HUMAN_WT_P16435_1677_1-677

MGDSHVDTSSSTVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVF CMATY GEGDPTDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRD
GALTQLNVAFSREQSHKVYVQHLLKQDREHLWKLI-

EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS

>PANTR_H2QUS7_1680_1-677

MGDSHVDTSSSTVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVF CMATY GEGDPTDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDIVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-

EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS

>GORIL_XP_004045657.1_1680_1-677

MGDSHMDTSSSTVSEAVAEVSLFSMTDMI LFS LIVGLLTYWFLRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIEN
ALVVF CMATY GEGDPTDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASETSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELVQFHRD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-

EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHTQAVDYIKKLMTKGRYSLDVWS

>MACFA_G7P166_1680_1-677

MGDSHVDTSSSTVSEAVAEVSLFSMTDVI LFSVIVGLLTYWFLRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNIVFYGSQTGTAEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVF CMATY GEGDPTDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDRLEQL
GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSRVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGTGVAPFRGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRD

GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-
EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYVKKLMTKGRYSLDVWS
>MACMU_CONSENSUS_1680_1-677
MGDSHVDTSSSTVSEAAAEVSLFSMTDVI LFSVIVGLLTYWFLFRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGPTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFH RD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-
EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYVKKLMTKGRYSLDVWS
>PAPAN_XP_003895904.1_1680_1-677
MGDSHVDTSSSTVSEAAAEVSLFSMTDVI LFSVIVGLLTYWFLFRKKKEEVPEFTKIQTLTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDAAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGPTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFH RD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-
EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYVKKLMTKGRYSLDVWS
>CHLSA_XP_008016563.1_1680_1-677
MGDSHVDTSSSTVSEAAAEVSLFSMTDVI LFSVIVGLLTYWFLFRKKKEEVPEFTKIQTSTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDAAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGPTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFH RD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-
EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYVKKLMTKGRYSLDVWS
>NOMLE_M3Z9U2_1686_1-677
MGDSHVDTSSSTVSEAAAEVSLFSMTDMI LFSLIVGLLTYWFLFRKKKEEVPEFTKIQTSTSSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDN
ALVVFCMATYGE GPD TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDTAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISNSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGPTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFH SD
GALTQLNVAFSREQSHKVYVQHLLKRDREHLWKLI-
EGGAHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
>CALJA_U3C910_1680_1-677
MGDSNVDTSSSTVSEAAAEVSLFSMTDMI LISLIVGLVTYLF LFRKKKEEVPEFTKIQTSTSSVK
ESSFVEKMKKTGRNIVIF YGSQTGTAE EFANRLSKDAHRYGMRGMTADPEEYDLADLSSLPEIEN
SLVVFCMATYGE GPD TDNAQDFYDWLQETDMDLSGVKF SVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFDLGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHSDIDAAKVYM
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGKILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTY YLDITNPPRTNVL YELA
QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY

YSIASSSKVHPNSVHICAVVVEYESKAGRINKGVATNWLRAKEPARENGGRALVPMFVRKSQFRL
 PFKATTPVIMVPGTGVAPFMGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
 GTLTQLNVAFSREQAHKVYVQHLLKRDREHLWKLI-
 EGGAHYVCGDARNMARDVQNTFCDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
 >SAIBO_XP_003934174.1_1677_1-674
 MGDSNVDTSSSTVSEAAATEEVSLFSMTDMI LFSLI VGLLTYWFLFRKKKEDVPEFTKIQTSTTSVK
 ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
 SLVVFCMATYGE GDP TDNAQDFYDWLQETDMDLSGVKF AVFGLGNKTYEHFNAMGKYVDRRLEQL
 GAQRIFDLGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHSDIDTAKVYV
 GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
 NDSALVNQLGKILGTDLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
 QYASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCCELLPRLQARY
 YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRL
 PFKATTPVIMVPGTGVAPFMGFIQERAWLQQQGKKGW---
 ASGVGCRSDEDYLYREELARFHKDGALTQLNVAFSREQPHKVYVQHLLKRDREHLWKLI-
 EGGAHYVCGDARNMARDVQNTFCDIAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
 >OTOGA_XP_003794235.1_1681_1-678
 KGDSHMDTSTTMSEAMVEEVSLFSMTDVI LFSLI VGLLTYWFLFRKKKEDVPEFTKIQTATSSVK
 ESSFVEKMKKTGKNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
 ALVVFCMATYGE GDP TDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTD VDMAKVYV
 GEMGR LKSYENQKPPFDAKNPFLAAVATNRLNQGTERHFMHMELDISDSKIRYESGDHVAVYPA
 NDSALVNQLGEILGADLDVIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
 QYASAPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
 YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGHRALVPMFVRKSQFRL
 PFKATTPVIMVPGTGVAPFIFGFIQERAWLRQQGKEVGETLLYYGCRSEEDYLYREELAQFHRD
 NALTHLNVAFSREQSHKVYVQHLLKRDREHLWKLIHEEGAHYVCGDARNMARDVQNTFYDIVAE
 FGGMEHAQAVDYIKKLMTKGRYSLDVWS
 >TUPCH_L9K4S3_EDIT_1691_1-677
 MGDSHADSSATVSETVAEEVSLFSMTDMI LFSLI LGLLTYWFLFRKKKEEVPEFTKIQAMTSSVK
 ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSDIDN
 SLAVFCMATYGE GDP TDNAQDFYDWLQETDVDLSGLKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDPAKVYK
 GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
 NDSALVTQLGEILGADLDVVM SLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
 QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
 YSIASSSKVHPNSVHICAVAVEYETKSGRINKGVATSWLRAKEPAGENGRRALVPMFVRKSQFRL
 PFKATTPVIMVPGTGVAPFIFGFIQERAWLQQQGKEVGETLLYY-
 CRRSDEDYLYREELAQFHKGSLTQLNVAFSREQPHKVYVQHLLKRDKEHLWKLIHDGGAHYVCG
 GDARNMARDVQNTFCDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
 >TARSY_XP_008056208.1_1677_1-677
 MGDSQMDSSSTVSEAAAEVSLFSMTDMI LFSLI VGLLTYWFLFRKKKEEVPEFTKIQATTSSVK
 ESSFVEKMKKTGRNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
 SLVVFCMATYGE GDP TDNAQDFYDWLQETDVDLSGVKF AVFGLGNKTYEHFNAMGKYVDRRLEQL
 GAQRIFELGMGDDDDGNLEEDFI TWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDVAKVYT
 GEMGR LKSYENQKPPFDAKNPFLAAVT TNRLNQAERHFMHLELDISDSKIRYESGDHVAVYPA
 NDSALVSQ LGEILGADLDAVISLNNIDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
 QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDCRSLRPPVDHLCCELLPRLQARY
 YSIASSSKVHPNSVHICAVAVEYETKTGR TGKVATTWLRAKEPAPENGRRALVPMFVRKSQFRL
 PFKATTPVVMVPGTGVAPFLGFIQERAWLRQQGKEVGETLLYYGCRSDEDYLYREELAQFQKE
 GALTHLNVAFSREQPQKVRRAGRSGR-
 REHLWKLVHEGGAHYVCGDARNMARDVQNTFCDIVAELGAMEHAQAVDYVKKLMTKGRYSLDVW
 S
 >CAVPO_CONSENSUS_1678_1-678
 MGDSHVDTGATSTEAAAEVSLFSMTDMI LLSVLVGLTYFFLFRKKKEEIPFTKISTTTSSVK
 ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
 SLAVFCMATYGE GDP TDNAQDFYDWLQETDLDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
 GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVLHADTDPAKVYT

GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLHKMASSSGEGKELYLTWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLQAKDPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFIGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GTLTQLNVAFSREQAQKVYVQHLLKRDQEH LWKLIHEDGAHIYVCGDARNMARDVQNTFCNIAAE
LGGMEHTQAVDYVKKLMTKGRYS LDVWS

>OCTDE_XP_004630501.1_1678_1-678

MGDSHVDTAATSTEAVAEEMSLFSMTDMI LLSLIMGILTYWFFFRKKKEEIPFTKISAEPSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLTEIDN
SLVVFCMATYGE GDP TDNAQDFYDWLQETDLDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGQGGDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQYELVLHTDMDPAKVYT
GEVGR LKSYK NQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAIYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKDPAGENSHRALVPMFVRKSQFRL
PFKTTTPVIMVPGTGVAPFMGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GALAQLNVAFSREQAHKVYVQHLLKRDQEH LWKLIHEDNAHIYVCGDARNMARDVQNTFCDIAAE
LGGMEHPQAVDYIKKLMTKGRYS LDVWS

>CHILA_CONSENSUS_1678_1-678

MGDSHMDTGASNAEAVAEVSLFSMTDMI LFSLIMGVLTYWFLFRKKKEEVPEFTKISTTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIDN
SLVVFCMATYGE GDP TDNAQDFYDWLQETDLDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQYELVLHADIDLAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSNLVNQLGKILGADLDTVMSLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYATEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLRAKDPAGENGHRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDREHLWKLIHEDGAHIYVCGDARNMARDVQNAFCDIVAE
LGGMEHAQAVDYVKKLMTKGRYS LDVWS

>HETGA_G5BG65_1882_1-678

MGDSHMDTGATNAEAVAEVSLFSTTDMI VFSV IIGVLTYWFLFRKKKEEMPEFTKISTTTTSVR
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLAEIDN
SLVVFCMATYGE GDP TDNAQDFYDWLQETDLDLSGIKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQYELVHMDIDEARVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGT ERYLMHLELDISDSKIRYESGDHVAIYPA
NDSNLVNQLGKILGADLDVVM SLKNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVVVEYETKSGRINKGVATSWLQAKNPAGDNRRALVPMFVRKSQFRL
PFKSTTPVIMVPGTGVAPFMGFIQERAWLQQQGKEVGETLLYYGCRSDEDYLYREELAQFHKE
GALSGLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMARDVQNTFCDIAAE
LGAMEHAQAVEYIKKLMTKGRYS LDVWS

>CRIGR_CONSENSUS_1678_1-678

MGGSHKDSSATMTEAVAEVSLFSTTDVVLFSLIVGVLTYWFI FRKKKEEVPEFSKIQTATPSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVFCMATYGE GDP TDNAQDFYDWLQETDLDLTGVKF AVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQYELLVHEDIDAAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVVM SLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVLYELA
QYASEPSEQEQLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSVTPVIMVPGTGIAPFMGFIQERAWLREQGKEVGETLLYYGCRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYS LDVWS

>MICOC_XP_005344666.1_1678_1-678

MGDSHKDTSATMTEAVAEVSLFSTTDMVLFSLIVGVLTYWFI FRKKKEEVPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNII VFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK

SLVVF CMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAV FGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHENVDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVSQIGEILGADLDIVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVL YELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMIGPGTGIAPFMGFIQERAWLREQGKEVGETVLYYGCRRSDEDYLYREELARFHKD
GTLTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHTQAVDYVKKLMTKGRYSLDVWS
>MOUSE_CONSENSUS_1678_1-678
MGDSHEDTSATVPEAAVEEVSLFSTTDIVLFSLIVGVLT YWFIFRKKKKEE IPEFSKIQT TAPPVK
ESSFVEKMKKTGRNIIVFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVF CMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAV FGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVL YELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRTKEPAGENGRRALVPMFVRKSQFRL
PFKPTTPVIMVGP GTGVAPFMGFIQERAWLREQGKEVGETLLYYGCRRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHTQAVDYVKKLMTKGRYSLDVWS
>RAT_P00388_1678_1-678
MGDSHEDTSATMPEAAVEEVSLFSTTDMVLFSLIVGVLT YWFIFRKKKKEE IPEFSKIQT TAPPVK
ESSFVEKMKKTGRNIIVFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVF CMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAV FGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDMDVAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVTANRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVL YELA
QYASEPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGVAPFMGFIQERAWLREQGKEVGETLLYYGCRRSDEDYLYREELARFHKD
GALTQLNVAFSREQAHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHTQAVDYVKKLMTKGRYSLDVWS
>PERMA_XP_006971344.1_1678_1-678
MGDSHKDASATMTEAAVEEVSLFSTTDMVLFSLIVGVLT YWFIFRKKKKEE VPEFSKIQT TTTSSVK
ESSFVEKMKKTGRNIIVFYGSQTGTAE EFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDK
SLVVF CMATY GEGDPTDNAQDFYDWLQETD VDLTG VKFAV FGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHEDMDTAKVYT
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSTLVNQIGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVL YELA
QYASKPSEQEHLHKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGVAPF IGFIQERAWLREQGKEVGETLLYYGCRRSNEDYLYREELA AFHKA
GALTQLNVAFSREQEHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS
>JACJA_XP_004666327.1_1678_1-678
MEASHEDASATAPETAANEVSLFSMTDMVLFSLIMGLLT YWFFRKKKKEE VPEFSKIQT TTTTSVK
ESSFVEKMKKTGRNIIVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEN
SLVAF CMATY GEGDPTDNAQDFYDWLQETD VDLTG VRF AVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGLGDDDDGNLEEDFI TWREQFWPAVCEFFGVEATGEESSIRQYELVVHPDIDPAKVYV
GEMGR LKSYENQKPPFDAKNPFLAAVT TNRKL NQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVSQIGEILGADLDVMSLNNLDEESNKKHFPFCPTTYRTALTYYLDITNPPRTNVL YELA
QYASEPSEQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYEAKSGRVNKG VATTWLRNKEPAGENGRRALVPMFVRKSQFRL
PFKSTTPVIMVGP GTGIAPFMGFIQERAWLREQGKEVGETVLYYGCRRSDEDYLYREELARFHKD
GALTQLRVAFSREQAQKVYVQHLLRRDKEHVWKLIHEGGAHIYVCGDARNMAKDVQNTFYDIVAE
FGPMEHAQAVDYVKKLMTKGRYSLDVWS
>ICTTR_XP_005328569.1_1606_1-606

MKKTGRNIIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIHNSLVVFCM
ATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQLGAQRIFE
LGLGDDDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQYELVVHTDMDMAKVYVGEMGRK
SYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPANDCTLVN
QLGEILGADLDVVM SLNNLDEESNKKHPFPCPTSYRTALTYYLDTNPPRTNVLYELAQAASEPA
EQEHLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARYYSIASSS
KVHPNSVHICAVAVEYETKSGRVNKG VATSWLRAKEPAGENGRRALVPMFVRKSQFRLPFKPTTP
VIMVGP GTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRADEDYLYREELAQFHKDGTTLTQLN
VAFSREQAHKVYVQHLLKRDAEHLWELIHNGGAHIYVCGDARNMARDVQNTFCDIVAELGAMEHA
QAVDYIKKLMTKGRYSLDVWS

>PTEAL_L5K7Y7_1657_1-657

MRPAAVLLSGGDGAQFRPLESARMVEKGTVKCGFTAENG NLSVKDS SFVEKMKKTGRNIIIVFYGS
QTGTAEFANRLSKDAHRYAMRSMADPEEYDLADLGS LPEIENSLAVFCMATYGECDPTDNAQD
FYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQLGAQRIFELGLGDDDDGNLEEDF
ITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDIDVAKVYVGEMGRKLSYENQKPPFDAKNP
FLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPANDSALVNQLGEILGVLDVVM
SLNNLDEESNKKHPFPCPTSYRTALTYYLDTNPPRTN ILYELAQAASEPSEQEHLRKMASSSGE
GKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARYYSIASSSKVHPNSVHICAVAV
EYETKSGRINKGVATSWLRAKEPTKENGRRALVPMFVRKSQFRLPFKPTTPVIMVGP GTGVAPFI
GFIQERAWLQHKGKEVGETLLYYGCRSEDEDYLYREELAQFHKEGSLTQLNVAFSREQPQKVYVQ
HLLKRDK EHLWKL IHEGGAHIYVCGDARNMARDVQNTFYDIVAELGSMEHAQAVDYIKKLMTKGR
YSLDVWS

>EOSBA_E6Y366_1678_1-678

MGDSHGDTSATMPEAAAEAAAFSMTDAVLFS L I VGLLTYWFLFRKKKEEVPEFTKIQSTTASVK
ESSFVEKMKKTGRNIIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSEIDN
ALVVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDPAKVYQ
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGALLGADLDVVM SLNNLDEESNKKHPFPCPTTYRTALTYYLDTNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAIEYETKTGRVNKG VATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGIAPFIGFIQERAWLQQGKEVGETLLYYGCRSEEDYLYREELAQFHKD
GALTQLS VAFSREQAHKVYVQHLLKRDK EHLWKL IHEGGAHIYVCGDARNMAKDVQNTFYEIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>OCHCU_E6Y368_1678_1-678

MGDSHGDTSATMPEAAAEAAAFSMTDAVLFS L I VGLLTYWFLFRKKKEEVPEFTKIQSTTASVK
ESSFVEKMKKTGRNIIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSEIDN
ALVVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDPAKVYQ
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGALLGADLDVVM SLNNLDEESNKKHPFPCPTTYRTALTYYLDTNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAIEYETKTGRVNKG VATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVGP GTGIAPFIGFIQERAWLQQGKEVGETLLYYGCRSEEDYLYREELAQFHKD
GALTQLS VAFSREQAHKVYVQHLLKRDK EHLWKL IHEGGAHIYVCGDARNMAKDVQNTFYEIVAE
LGAMEHAQAVDYVKKLMTKGRYSLDVWS

>OCHPR_XP_004587043.1_1681_1-678

MGDSHGDTSATMPEAAAEAAAFSMTDAVLFS L I VGLLTYWFLFRKKKEEVPEFTKIQSTTASVK
ESSFVEKMKKTGRNIIIVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLSEIDN
ALVVFCMATYGECDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMDPAKVYH
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQIGAILGADLDVVM SLNNLDEESNKKHPFPCPTTYRTALTYYLDTNPPRTNVLYELA
QYASEPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKTGRINKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRL

PFKATTPVIMVPGGTGIAPFIGFIQERAWLQQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKD
GALTQLSVAFSREQAHKVYVQHLLRKDKHEHLWKLIQEGGAHIYVCGDARNMAKDVQNTFYEIVAE
QGAMEHAQAVDYVKKLMTKGRYSQDVWS
>RABIT_P00389_1679_1-679
MADSHGDTGATMPEAAQEASVFSMTDVVLFSLIVGLITYWFLFRKKKEEVPEFTTKIQPTSSSVK
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEINN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVDLSGVKYAVFGLGNKTYEHFNAMGKYVDQRLEQL
GAQRIFELGMGDDDDANLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVLHTDIDVAKVYQ
GEMGRKLSYENQKPPFDAKNPFLATVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVMSSLNLDDESNKKHFPFCPTSYRTALTYLYLDITNPPRTNVLYELA
QYAADPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVAVEYETKAGRLNKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRL
PFKATTPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRAAEDYLYREELAGFQKD
GTLSQLNVAFSREQAQKVYVQHLLRRDKHEHLWRLIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHAQAVDYVKKLMTKGRYSQDVWS
>DASNO_XP_004448327.1_1678_1-678
MADSNADTSATASEAVAEVSLFSMTDVVLFSLIVGLITYWFLFRKKKDEVPEFTTKIQPTTSSIK
ESSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLAEIDN
ALAVFCMATYGECDPTDNAQDFYDWLQETDVDLSGVKFVAVFGLGNKTYEHFNAMGKYVDKRLEQL
GAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVLHTDIDVAKVYT
GEMGRKLSYENQKPPFDSKNPFLAGVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPA
NDSALVNQLGEILGADLDVVMSSLNLDDESNKKHFPFCPTSYRTALTYLYLDITNPPRTNVLYELA
QYASEPTEQEHLRKMASSSGEGKELYLSWVVDARRHILAILQDYP SLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSVHICAVTVEYETKSGRVNKGVATSWLRAKDPAGENGRRALVPMFVRKSQFRL
PFKAATPVIMVPGGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRADEDYLYREELARFHQD
GALTQLNVAFSREQAHKVYVQHLLRKDKQHLWELIHEGGAHIYVCGDARNMARDVQNTFYDIVAE
LGAMEHDQAVDYVKKLMTKGRYSQDVWS
>ERIEU_XP_007518012.1_1678_1-677
MGDSSPDAAASMSSETAAEEVSLFSMTDMVLFSLIIGLLITYWFLFRKKKEDIPEFTTKIQPA-
SSVKDSSSFVEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSLP
EIENSLAVFCMATYGECDPTDNAQDFYDWLQEADADLSGVKFVAVFGLGNKTYEHFNAMAKYVDRR
LEQLGAQRIFELGMGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELAVHTDIDTA
KVYTGEMGRKLSYENQKPPYDAKNPFLAAVTTNRKLNQGSEHLMHLELDISDSKIRYESGDHVA
VYPANDSTLVNQIGEILGADLDIVMSLNLDDESNKKHFPFCPTSYRTALTYLYLDITNPPRTNVLY
YELAQYAAQPAEQEQLRKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPPIDHLCCELLPRL
QARYYSIASSSKVHPSSVHICAVAVEYETKAGRLNKGVATSWLRTKEPAGENGCRALVPMFVRKS
QFRLPFKATTPVIMVPGGTGVAPFIGFIQERAWLQQQGKEVGETLLYYGCRRSDEDYLYREELER
FHQEGALTQLNVAFSREQPQKVYVQHLLRKDKHEHLWKLIHQGAHIYVCGDARNMAKDVQNAFYN
IVEELGTMEHAQAVDYVKKLMTKGRYSQDVWS
>CONCR_XP_004688336.1_1676_1-676
MGDSSMDATA--
SEAAAEVSLFSMTDMILFSLIVGLITYWFLFRKKKEEVPEFTTKIQTTTSSVKDSSSFVEKMKKTG
RNIVVFYGSQTGTAEFANRLSKDAHRYGMRGLAADPEEYDLADLSSLPQIDNSLAVFCMATYGE
GDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQLGAQRIFELGMG
DDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVIHKDIDMAKVYMGEMGRKLSYENQ
KPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVYPANDSALVNQLGEI
LGADLDVVMSSLNLDDESNKKHFPFCPTTYRTALTYLYLDITNPPRTNVLYELAQYATEPSEQEHL
RKMASSSGEGKELYLNWVVEARRHILAILQDYQSLRPPIDHLCCELLPRLQARYYSIASSSKVHPN
SVHICAVAVEYETKAGRINKGVATSWLRAKDPAGENGRRALVPMFVRKSQFRLPFKATTPVIMIG
PGTGIAPFMGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHKDGF LTQLNVAFSR
EQPQKVYVQHLLKKDKHEHLWKLIHEEGAHIYVCGDARNMAKDVQNTFYDIVTELGAMEHAQAVDY
IKKLMTKGRYSQDVWS
>SORAR_XP_004620907.1_1676_1-676
MGGPSVDSGA--
PDSGPKEVSLFSTTDVVLFSLIVGLITYWFLFRKKKEEVPEFTTKIQQTSTSAKSSSFVEKMKKTG
RNIVVFYGSQTGTAEFANRLSKDAHRYGMRGMAADPEEYDLADLSSSLAEIENSLAVFCMATYGE
GDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDRRLEQLGARRIYELGLGD
DDGNLEEDFITWREQFWPAVCEHFGVEATGDEASIRQYELVHTDMDTAKVYTGEMGRKLSYENQ

KPPFDAKNPFLAAVTINRKLNQGGERHLMHLELDISDSKIRYESGDHVAVYPANDSALVQQLGEI
 LGADLDVVMSSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELAQYAAEPAEQEQL
 RKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPP IDHLCCELLPRLQARYYSIASSSKVHPN
 SVHICAVAVEYETKAGRVNKG VATSWLRTKDPAGDNGRRALVPMFVRKSQFRLPFPKATTPVIMVG
 PGTGVAPFMGFIQERAWLKQOGKEVGETLLYYGCRSDEDYLYRDELAQFHKDGALSRLSVAFSR
 EQPQKVYVQHLLKKDKHLWKL IHEDGAHIYVCGDARNMAKDVQNTFYDMVAELGAMEHTQAVDY
 IKKLMTKGRYSLDVWS

>PHACI_CONSENSUS_1677_1-677

MADVDP SATA--

SQAGAEDVSLFSVTDVFLFSFIVGLMTYWFFFRKKKEE IPEFTKI QPVSTNVKD NSFVEKMKKTG
 RNIVVFYGSQTGTAE EFANRLAKDAHRYGMRGMAADPEEYDLADLSSLPEIDNALALFCMATYGE
 GDPTDNAQDFYDWLQETDADLSGVKFVAVFALGNKTYEHFNAMGKYVDNRLEQLGAQRIFELGLGD
 DDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDENMNKVYTGEMGR LKS YENQ
 KPPFDAKNPFLATVTTNRKLNQGGERHLMHLELDISNSKIRYESGDHVAVYPANDPTLVNQLGEI
 LGADLDTVISLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELAQYAGEPFGTGAA
 AQDGSSSGEGKELYLNWVVEARRN ILAILQDYP SLRPP IDHLCCELLPRLQARYYSIASSSKVHTN
 SIHICAVVVEYETKAGRVNKG VATNWLRAKEPALENGHKATVPMFVRKSQFRLPFPKPTTPVIMIG
 PGTGIAPFVGFIQERAWLKEQGKEVGETVLYYGCRANEDYLYREELAQFQKDGVLSQLNVAFSR
 EQPHKVYVQHLMKQNK EHLWRLMHEGSAHVYVCGDARNMAREVQNVFYEIVSEFGMSRAQAVDF
 IKKLMAKGRYSLDVWS

>SARHA_XP_003770264.1_1625_1-625

MADVDT SATA--

SEAGTEDVSLFSMTDMFLFSLIIVGLLLTYWFFFRKKKEE IPEFTKI QPVPPNVKES SFVEKMKKTG
 RNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIENSLAVFCMATYGE
 GDPTDNAQDFYDWLQETDADLSGVRFAVFG LGNKTYEHFNAMGKYVDNRLDQLGAQRIFELGLGD
 DDGNLEEDFITWREQFWPAVCEYF GVEATGEESSIRQYELVVHTDENMNKVYTGEMGR LKS YENQ
 KPPFDAKNPFLAIVTTNRKLNQGGERHLMHLELDISNSKIRYESGDHVAVYPANDALLVNQLGEI
 LGADLDVVMSSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELAQYAGEPAEQEQL
 RKMASSSGEGKELYLSWVVEARRHILAILQDYP SLRPP IDHLCCELLPRLQVRYYSIASSSKVHPN
 SVHICAVVVEYKTKAGRVNKG VATNWLRAKEPAQENGKATVPMFVRKSQFRLPFPKPTTPVIMIG
 PGTGIAPFMGFIQERAWLKEQ GK-----

HLMKQNK EHLWKL IHGSAHIYVCGDARNMARDVHNVFCEIASEFGMSQS QAVDF IKKLMAKGR
 YSLDVWS

>ALLMI_XP_006274812.1_1681_1-678

MGDSGMDPTVSPPESTAQEASFLSMTDMFLISLIIGLFSYWFFFRKKKEEVPEFTKIQQTTSSGG
 DSNFVDKMKKTGRNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGFGADPEEYDLSDLSRLSEIDN
 SLAVFCMATYGE GDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHTDMNMNKVYT
 GEMGR LKS YESQKPPFDAKNPFLAPVTKNRKLNQGGERHLMHLELDISGSKVRYESGDHVAVYPA
 NDSTLVNQIGEV LGADLDTIMSLNNLDEESNKKHPFPCPTSYRTALTYYLDITNPPRTNVLYELA
 QYATDANEQDRLRKMASSAAEGKEFYLSWVVD SRRNILAILQDMP SLRPP IDHLCCELLPRLQARY
 YSIASSSKVHPNSIHICAVTVEYETKTGR LNKGVATTWLKNKVP-

NENGKNSLVP MYVRKSQFRLPFPKSTPVIMIGPGTGIAPFVGFIQERAWLKQOGKEV GDTVLYYG
 CRHEKEDYLYQEELTQFHKEGALTQLNVAFSRDQAQKVYVQHLLKQNK EGVWKL IHAGNAHIYVC
 GDARNMARDVQNTFYEIVAEYGNMSQPQAVDYVKKLMTKGRYSMDVWS

>ALLSI_XP_006033298.1_1681_1-678

MGDSGMDPTVSPPESTAQEASFLSMTDMFLISLIIGLFSYWFFFRKKKEEVPEFTKIQQTTSSGG
 DSNFVDKMKKTGRNIVVFYGSQTGTAE EFANRLSKDAHRYGMRGLGADPEEYDLSDLSRLSEIDN
 SLAVFCMATYGE GDPTDNAQDFYDWLQEADVDLSGVKYAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRIFELGLGDDDDGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHTDMNMNKVYT
 GEMGR LKS YESQKPPFDAKNPFLAPVTKNRKLNQGGERHLMHLELDISGSKVRYESGDHVAVYPA
 NDSTLVNQIGEILGADLDTIMSLNNLDEESNKKHPFPCPTSYRTALTYYLDITNPPRTNVLYELA
 QYATDANEQDRLRKMASSAAEGKEFYLSWVVD SRRNILAILQDMP SLRPP IDHLCCELLPRLQARY
 YSIASSSKVHPNSIHICAVTVEYETKTGR LNKGVATTWLKNKVP-

NENGKNSLVP MYVRKSQFRLPFPKSTPVIMIGPGTGIAPFVGFIQERAWLKQOGKEV GDTVLYYG
 CRHEKEDYLYQEELTQFHKEGALTQLNVAFSRDQAQKVYVQHLLKQNK EGVWKL IHAGNAHIYVC
 GDARNMARDVQNTFYEIVAEYGNMSQPQAVDYVKKLMTKGRYSMDVWS

>ANAPL_CONSENSUS_1664_1-664

MGDAGVG---

APPAGSVQQDSFFSMTDVFLISLITGLFTYWFFFRKKKEEIPDLPKIQTAPASPRDSSFIEKMKK
TGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDKSLAVFCMATY
GEGDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEELGAQRIFELGL
GDDDGNNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMSKVYTGEMGRKLSYE
NQKPPFDAKNPFLAMVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPANDSSLVNQIG
EILGTDLDTIMSLNNLDEESNKKHPFPCPTSRYTALTYLDITNPPRTNVLYELAQYATDSSEQE
RLRKMASSSAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVH
PNAIHICAVTVEYETKTGRLNKG VATNWLKSKVP-

NENGSNLSLVP MYVRKSQFRLPFPSTP VIMIGPGTGIAPFIGFIQERAWLKQQGKEVGE-----

REDYLYRQELAQFQQEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLLHDGNAHIYVCGDAR
NMARDVQNTFYEIVA EYGNMSQS QATDYVKKLMTKGRYSLDVWS

>CHICK_CONSENSUS_1676_1-676

MGDAGMESTVSPPEGTA-

QDSFLSMTDVFLISLITGLFTYWFFFRKKKEEIPDLPKIQTVSSPARDSSFIEKMKKTGRNIVVF
YGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDKSLAVFCMATY GEGDPTDN
AQDFYDWLQEADTDL SGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEELGAQRIFELGLGDDDGNNLE
EDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYTGEMGRKLSYENQKPPFDA
KNPFLAVVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPANDASLVNQLGEILGTDLD
TVMSLNNLDEESNKKHPFPCPTSRYTALTYLDITNPPRTNVLYELAQYATDTGEQEQLRKMMASS
SAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPNSIHICA
VTVEYETKTGRLNKG VATNWLKDKVP-

NENGRNSLVP MYVRKSQFRLPFPSTP VIMIGPGTGIAPFIGFIQERAWLKEQGKEVGETVLYYG
CRREREDYLYRQELARFKQEGVLTQLNVAFSRDQAEKVYVQHLLKKNKEHIWKLVNDGNAHIYVC
GDARNMARDVQNTFYEIVSEYGNMNQS QAVDYVKKLMTKGRYSLDVWS

>FALCH_XP_005439856.1_1675_1-675

MGGAGMEQNISPANSPAQQDSLFSMDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQSVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATY GEGDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDGNNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIG EILGTDLDTIMSLNNLDEESNKKHPFPCPTSRYTALMYLDITNPPRTNVLYELA
QYATDAGEQERLRKMMASSAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNAIHICAVTVEYETKTGRLNKG VATNWLKNKVP-

NENGSNLSLVP MYVRKSQFRLPFPSTP VIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETGLYY-
-

GLRAGGYLYREELARFHQEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCG
DARNMARDVQNTFYEIVA EFGNMGQPQAVDYVKKLMTKGRYSLDVWS

>FALPE_XP_005240527.1_1674_1-674

MGGAAMEQNISPANSPAQQDSLFSMDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQSVAAPVR
DSSFIEKMKKTGRNIVVFYGSQTGTAEFANRLAKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATY GEGDPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDGNNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGRKLSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIG EILGTDLDTIMSLNNLDEESNKKHPFPCPTSRYTALMYLDITNPPRTNVLYELA
QYATDAGEQERLRKMMASSAEGKALYLSWVVEARNILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNAIHICAVTVEYETKTGRLNKG VATNWLKNKVP-

NENGSNLSLVP MYVRKSQFRLPFPSTP VIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVGCR-
--

REREDYLYREELARFHQEGILTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLINEGNAHIYVCGD
ARNMARDVQNTFYEIVA EFGNMGQPQAVDYVKKLMTKGRYSLDVWS

>COLLI_CONSENSUS_1677_1-677

MGDAGMELNVSPPDNAVQQDSLFSMTDVFLISLITGLFTYWFFFRQKKEEIPDLPKMQTVVAPVR
ESSFIEKMKKTGRNIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATY GEGDPTDNAQDFYDWLQEADVDLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDDDGNNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDMNMNKVYT

GEMGR LKSYENQKPPFDAKNPFLAHVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYSTDAGEQERLRKMASSAADGKALYLSWVVEARRNILAILQDMPSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPDSIHICAVTVEYETKTGRVNGKVATNWLKKNKP-
NEKGNNSLVPMYVRKSQFRLPFKSSTPVIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYG
CRREHEDYLYREELARFHQEGVLTQLNVAFSRDQAEKVYVQHLLKKNKESIWKLINEGNAHIYVC
GDARNMARDVQNTFYEIIVAEFGNMSQPQAVDYIKKLMTKGRYSLDVWS
>FICAL_XP_005056571.1_1676_1-676
MGEP SMDPTISA-
DSTAQQDSLF SMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVRDSSFIEKMKKTG
R NIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDKSLAVFCMATYGE
G DPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEELGAQRIFELGLGDD
DGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT GEMGR LKSYENQK
PPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA NDSSLVNQIGEIL
GTDLDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA QYATDAGEQERLR
KMASSAAEGKALYLSWVVEARRNILAILQDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPNS
IHICAVTVEYETKTGR LNKGVATNWLKSKVP-
QQGGSSSLVPMYVRKSQFRLPFKPSTPVIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYG
CRHEREDYLYREELARFQKEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENIWKLVNEGMAHIYVC
GDARNMARDVQNTFYEIIVSEFGSMSQPQAVDYVKKLMTKGRYSLDVWS
>PSEHU_XP_005525894.1_1677_1-677
MGEP SMDPTISAADSPAQQDSLF SMDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQSVAAPVR
DSSFIEKMKKTG R NIVVFYGSQTGTAEFANRLSKDAHRYGLRGMAADPEEYDLSDSLRLSEIDK
SLAVFCMATYGE G DPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEEL
GAQRIFELGLGDD DGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT
GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA
NDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA
QYATDAGEQERLRKMASSAAEGKALYLSWVVEARRNILAILQDMSSLRPPIDHLCCELLPRLQARY
YSIASSSKVHPNSIHICAVTVEYKTKTGR LNKGVATNWLKSKVP-
DQNGSSSLVPMYVRKSQFRLPFKPSTPVIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYG
CRRQHEDYLYREELAHFQKEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENVWKL VNEGMAHIYVC
GDARNMARDVQNTFYEIIVSEFGNMSQPQAVDYVKKLMTKGRYSLDVWS
>GEOFO_XP_005425451.1_1676_1-676
MGEP SMDPTIAA-
DSPAQQDSLF SMTDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQTVAAPVRDSSFIEKMKKTG
R NIVVFYGSQTGTAEFANRLAKDAHRYGLRGMAADPEEYDLSDSLRLSEIDKSLAVFCMATYGE
G DPTDNAQDFYDWLQEADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDKRLEELGAQRIFELGLGDD
DGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDVNMNKVYT GEMGR LKSYENQK
PPFDAKNPFLAQVTENRKLNEGGERHLMHLELDISNSKIRYESGDHVAVYPA NDSSLVNQIGEIL
GTDLDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELA QYASDSGEQERLR
KMASSAAEGKALYLSWVVEARRNILAILQDMPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPNS
IHICAVTVEYETKTGR LNKGVATSWLRSKVP-
GQQGSSSLVPMYVRKSQFRLPFKPSTPVIMIGPGTGIAPFMGFIQERAWLKQQGKEVGETVLYYG
CRHEREDYLYREELARFQKEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENVWKL VNEGMAHIYVC
GDARNMARDVQNTFYEIIVSEFGNMSQPQAVDYVKKLMTKGRYSLDVWS
>TAEGU_XP_002194936.1_1677_1-676
MGEP GMDPTMPAADSPAQQDSLF SMDVFLISLITGLFTYWFFFRKKKEEVPDLPKMQSVAAPVR
DSSFIEKMKKTG R NIVVFYGSQTGTAEFANRLSKDAHRYGLRGMTADPEEYDLSDSLRLSEIDK
SLAVFCMATYGE G DPTDNAQDFYDWLQEADADLSGLRFVAVSG-
GNKTYEHFNAMGKYVDKRLEELGAQRIFELGLGDD DGNLEEDFITWREQFWPAVCEHFGVEATGE
ESSIRQYELVVHTDVNMNKVYT GEMGR LKSYENQKPPFDAKNPFLAQVTENRKLNEGGERHLMH
ELDISNSKIRYESGDHVAVYPA NDSSLVNQIGELLGTDLDTVMSLNNLDEESNKKHFPFCPTSYR
TALTYLDITNPPRTNVLYELA QYASDTGEQERLRKMASSAAEGKALYLSWVVEARRNILAILQD
MPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPNSIHICAVTVEYETKTGR LNKGVATSWLRSK
VP-
EQNGSSSLVPMYVRKSQFRLPFKPSTPVIMIGPGTGIAPFIGFIQERAWLKQQGKEVGETVLYYG
CRREHEDYLYREELARFQKEGVLTQLNVAFSRDQAEKVYVQHLLKKNKENVWKL VNEGMAHIYVC
GDARNMARDVQNTFYEIIVSEFGNMSQPQAVDYVKKLMTKGRYSLDVWS

>CALAN_XP_008493329.1_1617_1-617

 AAPVRDSSFIEKMKKTGRNIVVFYGSQTGTAEFFANRLSKDAHRYGLRGMAADPEEYDLSDSLRL
 SEIDKSLAVFCMATYGECDPTDNAQDFYDWLQEQADADLSGLRFVAVFGLGNKTYEHFNAMGKYVDK
 RLEELGAQRIFELGLGDDGDNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDLNM
 NKVYTGEMGRLKSYENQKPPFDAKNPFLAKVTENRKLNEGGERHLMHLALDISNSKIRYESGDHV
 AVYPANDSSLVNQIGEILGTDLDTIMSLNNLDEESNKKHPFPCPTSYRTALTYYYLDITNPPRTNV
 LYELAQYASDTGEQERLRMASSAAEGKALYHSWVVEARRNILAILQDMPSLRPPIDHLCCELLPR
 LQARYYSIASSSKVHPNAIHCIVTVEYETKTGRLNKG VATNWLKNKVP-
 DENGRRSSMVP MYVRKSQFRLPFPKSTPVMIGPGTGIAPFVGFVQERAWLKQQGKEVGETVLYYG
 CRHQREDYLYREELAHFHQEGILTQLNVAFSRDQAQKVYVQHLLKKNKEHIWKLINEGNAHIYVC
 GDARNMARDVQNTFYEIVSEFGNMSQAQATDYVKKLMTKGRYSLDVWS

>CHEMY_M7AU06_1685_1-677

MGDSAMDSTLPIPEASVPEGALFSMTDVFLFSLIVGLLSYWFLFRKKKEEIPFTTKIQPVTSSVR
 ETSFVEKMKKTGRNIVVFYGSQTGTAEFFSNRLSKDAHRYGMRGMSADPEEYDMSDSLRLSEIDN
 SLAVFCMATYGECDPTDNAQDFYDWLQETDLDLTGLKFVAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRIFELGMGDDGDNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
 GEMGRLKSYESQKPPFDAKNPFLAPVTLNRKLNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
 NDTSLVNQIGEILGADLDMIMSLNNLDEESNKKHPFPCPTTYRTALTYYYLDITNPPRTNVLYELA
 QYATDPNEQEHLRKMMASSAAEGKTLYLSWVVEARRNILAILQDVSSLRPPIDHLCCELLPRLQARY
 YSIASSSKVHPHSIHCIVVVEYETKLGRVNLKG VATTWLKNKVP-
 NENGHKLTVPMYVRKSQFRLPFPKSSTPVMIGPGTGIAPFVGFVQERAWLKQQGKAVGETVLYYG
 CRHENEDYLYREELAQFHKEGALTQLNVAFSRDQAQKVYVQHLLKKNKAGVWKLIHEGNAHIYVC
 GDARNMARDVQNAFYEIVAEFGSMTQPAVDYVKKLMTKGRYSLDVWS

>CHRP1_XP_008164419.1_1677_1-677

MGDSAMDSTLPPPEASVPEGTLFSMTDVFLFSLIVGLLSYWFLFRKKKEEIPFTTKIQPTTSSVR
 ESSFVEKMKKTGRNIVVFYGSQTGTAEFFANRLSKDAHRYGMRGMSADPEEYDLSDSLRLSEIDN
 SLAVFCMATYGECDPTDNAQDFYDWLQETDLDLTGLKFVAVFGLGNKTYEHFNAMGKYVDKRLEQL
 GAQRMFELGMGDDGDNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYELVVHTDINMNKVYT
 GEMGRLKSYESQKPPFDAKNPFLAPVTLNRKLNQGGGERHLMHLELDISGSKIRYESGDHVAVYPA
 NDTSLVNQIGEILGADLDVIMSLNNLDEESNKKHPFPCPTTYRTALTYYYLDITNPPRTNVLYELA
 QYATDPNEQEHLRKMMASSAAEGKTLYLSWVLEARNILAILQDVPSLRPPIDHLCCELLPRLQARY
 YSIASSSKVHPHAIHCIVVVEYETKTGRLNKG VATTWLKNKVP-
 NENGHKLTVPMYVRKSQFRLPFPKSTPVMIGPGTGIAPFVGFVQERAWLKQQGKDVGETVLYYG
 CRHENEDYLYREELAQFHKEGALTQLNVAFSRDQAQKVYVQHLLKKNKEGVWKLIHEGNAHIYVC
 GDARSMARDVQNTFYEIVAEFGSMTQPAVDYVKKLMTKGRYSLDVWS

>PELSI_K7G8U9_1674_1-674

MEDSAMDSTP---

ESAGPEGALFSMTDVFLFSLIIGLLSYWFLFRKKKEEIPFTTKIQPITTSARESSFVEKMKKTSR
 NIVVFYGSQTGTAEFFANRLSKDAHRYGMRGMAADPEEYDLSDSLRLSEIDNSLAVFCMATYGECD
 DPTDNAQDFYDWLQETDLDLTGLKFVAVFGLGNKTYEHFNAMGKYVDKRLEQLGAQRIFEVGMGDD
 DGNLEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHTDINMNKVYTGEMGRLKSYESQK
 PPFDAKNPFLAPVTLNRKLNQAGERHLMHLELDISGSKIRYESGDHVAVYPANDTTLVNQIGEIL
 GADLDTIMSLNNLDEESNKKHPFPCPTTYRTALTYYYLDITNPPRTNVLYELAQYATDLGEQEHLR
 KMASSAAEGKTLYLSWVVEARRNILAILQDVPSLRPPIDHLCCELLPRLQARYYSIASSSKVHPHS
 IHCIVVVEYETKTGRLNRGVATTWLKNKVP-
 SENGHKLTVPMYVRKSQFRLPFPKSTPVMIGPGTGIAPFVGFVQERAWLKQQGKEVGETVLYYG
 CRHENEDYLYREELSQFHKEGALTQLNVAFSRDQAQKVYVQHLLKRNKEGVWRLIHEGNAHIYVC
 GDARNMARDVQNAFYEIVAEFGSMTQPAVDYVKKLMTKGRYSLDVWS

>CROAD_J3SE91_1681_1-681

MGTATSKEEPSSVKS-

PEASVAFSMDVFLFSIVIGLLTYWFFFRKKKEEIPFESKIQTPTQSANENSFIEKMKKTGKNIV
 VFYGSQTGTAEFFANRLSKDAHRYGMRGMPADPEEYDVSDLTHLSEIEKSLAIFCMATYGECDPT
 DNAQDFYDWLQDQDAKLSGIKYAVFGLGNKTYEHFNAMGKFVVDKRLEELGAQRIFELGLGDDGDN
 LEEDFITWREQFWPAVCEHFGVEATGDESSIRQYELVVHVDIAPAKVYMGEMGRLKSYENQKPPF
 DAKNPYLAPVTLNRKLNQGGGERHLMHLELDITGSKIRYESGDHIAVYPANDTSLVNQIGEILGAD
 LDTVISLNNLDEESNKKHPFPCPTSYRTALTYYYLDITNPPRTNVLYELAQYATNASEQEQLRKMA
 SSSAEGKALYLSWVVECRRNILAILQDTPSLRPPIDHLCCELLPRLQARYYSIASTSKVHPNSIHI

CSVVVEYTTKTNRVNKG VATNW LKNKQP-
NENGHKSTVPLYVRKSQFRLPFPKNT PVL MIGPGTGIAPFIGFIQERGW LKQQGKEVGETVLYFG
CRHEKEDYIYKEELAKFVKEGALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKLINEGNCHIYVC
GDARNMARDVQNTFYEI VAEY GKM NQQQAVDYVKKLMTKGRYSLDVWS
>OVOK_BAN82201.1_1658_1-658

MIDVFLFSIVIGLLTYWFFFRKKKEEIPFESKIQTPTQSANENSFIEKMKKTGKNIVVFYGSQTG
TAEEFANRLSKDAHRYGMRGMSADPEEYDVSDLTHLSEIEKSLAIFCMATYGEDPTDNAQDFYD
WLQD TDAKLSGIKYAVFGLGNKTYEHFNAMGKFVDKRLEELGAQRIFELGLGDDDNLEEDFITW
REQFWPAMCEHFGEATGDESSIRQYELVVHKDIAPAKVYIGEMGRLKSYENQKPPFDAKNPYLA
PVTLNRLKNQGGGERHLMHHELDITGSKIRYESGDHIAVYPANDTSLVTQLGEILSADLDAVISLN
NLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELAQYATNASEQEQLRKMASSAEGKT
LYLSWVVESRRNILAILQDTPSLRPPIDHLCELLPRLQARYYSIASTSKVHPNSIHICSVVVEYT
TKTNRVNKG VATNW LKNKQL-
NENGHKSTVPLYVRKSQFRLPFPKNT PVL MIGPGTGIAPFIGFIQERAWLKEQGKEVGETVLYFG
CRHEKEDYIYKEELAKFVKEGALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKLI SEGNGHIYVC
GDARNMARDVQNTFYEI VAEY GKM NQQQAVDYVKKLMTKGRYSLDVWS
>MICFL_U3F8N0_1681_1-681
MGTASSKDEPSSLKS-
PEASVSFSMMDIFLFSIVIGLLTYWFFFRKKKEEIPFESKIQTPTQSANENSFIEKMKKTGRNIV
VFYGSQTGTAEEFANRLAKDAQRYGMRGMSADPEEYDVSDLTHLSEIDKSLAVFCMATYGEDPT
DNAQDFYDWLQD TDAKLSGIKYAVFGLGNKTYEHFNAMGKFVDKRMEELGAQRIFELGLGDDDN
LEEDFITWREQFWPAVCEHFGEATGDETSIRQYELVVHKDIAPDKVYSGEMGRLKSYENQKPPF
DAKNPYLAPVSLNRKNQGGGERHLMHLELDITGSKIRYESGDHIAVYPANDTSLVNQLGEILSAD
LDTVMSLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELAQYATNASEQEQLRKMA
SSSAEGKTLYLSWVVESRRNILAILQDTPSLRPPIDHLCELLPRLQARYYSIASTSKVDPNSIHI
CAVVVEYTTKTNRVNKG VATNW LNMNKP-
NENGHKSTVPLYVRKSQFRLPFSNTPVIMIGPGTGIAPFIGFIEERGW LKQQGKEVGETVLYFG
CRHEKEDYIYKEELAKFVKEGALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKLINEGNNGHIYVC
GDARNMARDVQNAFYEI VAEY GKM NQQQAVDYVKKLMTKGRYSLDVWS
>PYTBI_XP_007433824.1_1680_1-680
MGTASSKEESSSVKS-
PEASVAFSMDVFLFSVIVIGLLTYWFFFRKKKEEIPFESKIQTTPAQSANENSFVEKMKKTGRNIV
VFYGSQTGTAEEFANRLAKDAHRYGMRGMSADPEEYDVSDLSHLSEIDKSLAIFCMATYGEDPT
DNAQDFYDWLQD TDANLSGIKYAVFGLGNKTYEHFNAMGKYVDKRLEELGAQRIFELGLGDDDN
LEEDFITWREQFWPAVCEHFGEATGDESCIRQYELVVHKDISQSKVYMGEMGRLKSYGNQKPPF
DAKNPFLAPVAVNRKNQGGGERHLMHLELDITGSKIRYEAGDHIAVYPANDTSLVNRLGEILSAD
LDTVISLNNLDEESNKKHFPFCPTSYRTALTYLDITNPPRTNVLYELAQYATDASEQEKLKMA
SSSPGKTLYLSWVVEARNILAILQDTPSLRPPIDHLCELLPRLQARYYSIASTSKVHTNSIHI
CAVVVEYTTKTDRVNKG VATNW LKNKQP-
NENGHKSTVPLYVRKSQFRLPFPKNT PVL MIGPGTGIAPFIGFIQERGW LKQQGKDVGETVLYYG
CRHEKEDYIYKEELAKFVKEGALTQLNVAFSRDQPEKIYVQHLLKKNKEHVWKLIHEENGHIYVC
GDARNMARDVQNTFYEI VAEH GKM NQQQAVDYVKKLMTKGRYSLDVWS
>ANOCA_XP_008120878.1_1678_1-678
MGSASSKDD SAPVKT-
PEASVPFSLTDVFLFSIIIGLLTYWFFFRKKKEEVPDFAKIQTSPQSTNDSSFIEKMKKDGKNIV
VFYGSQTGTAEEFANRLAKDAHRYGMRGMSADPEEYNLSDLNRLPEIDKSLAVFCMATYGEDPT
DNAQDFYDWLQETDADLSGLKFAVFGGLGNKTYEHFNAMGKYVDKRMEELGAQRIFELGLGDDDN
LEEDFITWREQFWPAVCEHFGEATGDESSIRQYELVVHTDINANKVYTGMGRLKSFENQKPPF
DAKNPFLAPVLVNRKNQGGGERHLMHLEVDITGSKIRYESGDHVAVYPCNDTSLVNQLGEILGAD
LDTVISLNNLDEESNKKHFPFCPTTYRTALTYLDITNPPRTNVLYELAQYASAASEQEHLKMA
SSSAEGKALYLSWVVEARNILAILQDTPSLRPPIDHLCELLPRLQARYYSIASSSKVHSNSIHI
CAVVVEYTTKTGRVNKG VATNW LKSKQP-
MDNGHKSVPMPYVRKSQFRLPFSSTPVL MIGPGTGVAPFIGFIQERGLHKKQGKEVGETVLYYG
CRHENEDYLYKEELANFLKEGALTQLHVAFSRDQAQKVYVQHLLKKNKESVWKLIHEDNGHIYVC
GDARNMARDVQNTFYEI VAEFGMDQPQAVDYIKKLMKGRYSLDIWS
>XENLA_Q6PCH9_1680_1-680

MGESCTEQDTSEQNGSPEEAFFSMADMFLLSLIVGLLTYWFFFRKKKEEI -
 EFTKIQPTNNSVRESSFIEKMKKTGKNIVVFYGSQTGTGEEFANRLAKDAHRYGVRGMAADPEEF
 EMADLSRLTEIENALAVFCMATYGECDPTDNAQDFYDWLQETDIDLTGLKYAVFGLGNKTYEHFN
 AMGKYVDRLEELGAERIFELGMGDDGDNLEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYEL
 VVHTDENMNKVYTGEMGRKLSYETQKPPFDAKNPFLANVTVNRKLNEGGDRHLMHLELDVTGSKI
 RYESGDHVAVYPANDTALVNKLGEILGADLDTVISLNNLDEESNKKHPFPCPTTYRTALTYYLDI
 TNPPRTNVLYELAQAATDSKEQENLRKMASSAQDGKGLYLSWVVESRRNILAILEDVPSLRPPLD
 HLCELLPRLQARYYSIASSSKVHPSSIHVCAVLVEYETKTGRENKGVATNWLKKNKQP -
 SDNGHKSSVPMYVRKSQFRLPFPSTPVMIGPGTGIAPFIGFIQEREWLKQQGKDVGETVLYYG
 CRHEHEDFLYKDELKRYHKDGVLTQLNVAFSRDQDRKVYVQHLLKDNKEMVWKLIHEDNAHIYVC
 GDARNMARDVQNTFYDIVAEGKIDHAQAVDYIKKLMTKGRYSQDVWS
 >XENST_XP_002933864.1_1680_1-680
 MGESCAEQDTSEQNGSSPEEAFSSMADMFLLSLIVGLLTYWFFFRKKKEEI -
 EFTKIQPTSSSVRESSFIEKMKKTGKNIVVFYGSQTGTGEEFANRLAKDAHRYGMRGMAADPEEF
 EMTDLSRLTEIENALAVFCMATYGECDPTDNAQDFYDWLQETDIDLAGLKYAVFGLGNKTYEHFN
 AMGKYVDRLEELGAERIFELGMGDDGDNLEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYEL
 VVHTDENMNKVYTGEMGRKLSYETQKPPFDAKNPFLANATVNRKLNEGGDRHFMHLELDITGSKI
 RYESGDHVAVYPANDAALVNKLGEILGADLETVISLNNLDEESNKKHPFPCPTTYRTALTYYLDI
 TNPPRTNVLYELAQAATDSKEQENLRKMASSAQDGKALYLSWVVESRRNILAILEDIPSLRPPLD
 HLCELLPRLQARYYSIASSSKVHPNSIHVCAVLVEYETKTGRENKGVATNWLKKNKQP -
 SDNGHKSSVPMFVRKSQFRLPFPSTPVMIGPGTGIAPFMGFIQEREWLKQQGKDVGETVLYYG
 CRHEHEDFLYINELKRYHKEGVLTQLNVAFSRDQAHKVYVQHLLKNNKEMVWKLIHEDNAHIYVC
 GDARNMARDVQNIFYDIVEEYGKLDHAQAVDYIKKLMTKGRYSQDVWS
 >FLY_Q07994_1-671

 MSAEHVEEVVSEEPFLGTLTDIALLVLLVGATWYFMRSRKKEEAPIRSSIQPTTSTVTENSFIKK
 LKASGRSLVVFYGSQTGTAEFFAGRLAKEGLRYRMKGMVADPEECDMEELLQMKDIPNSLAVFCL
 ATYGECDPTDNAMEFYEWITNGEVDLTGLNYAVFGLGNKTYEHYNKVAIYVDRLEELGATRVFE
 LGLGDDDDANIEDDFITWKDRFWPSVCDFFGIEGSGEEVLMRQFRLLEQPDVQPDRIYTGEIARLH
 SMQNQRPPFDAKNPFLASVIVNRELHKGGGRSCMHIELDIDGSKMRYDAGDHIAMYPINDKILVE
 KLGLCDANLDTVFLINTDTSKHKHPFPCPTTYRTALTHYLEITAIPRTHILKELAEYCSDEK
 DKEFLRNMA SITPEGKEKYQNWIQNSSRNIVHILEDIKSCRPPIDHICELLPRLQPRYYSISSSS
 KLYPTNVHITAVLVQYETPTGRVNGGVATSYMKEKNP ---
 SVGEVKVPVFIRKSQFRLPTKSEIPIIMVGPGTGLAPFRGFIQERQFLRDGGKVVGDTILYFGCR
 KKDEDFIYREELEQYVQNGTLT -
 LKTAFSRDQQEKIYVTHLIEQDADLIWKVIGEQQKHFIYICGDAKNMAVDVRNVLKILSTKGNMN
 ESDAVQYIKKMEAQKRYSDVWS

Figure 24

>CYP3_N1_nucleotide

ATGGCTCTGTTATTAGCAGTTTTTCTGACCCTGCTGATTCTGTATGGTATTTGGCCGTATGGTCTGTTCA
AAAAAAGTGGGAATTCGGGGTCCGCGTCCGCTGCCGTTTATTGGTACCTTTCTGGAATATCGTAAAGGC
tttCTGGATTTTGATACCGAATGCTTCAAAAAATACGGCAAAATCTGGGGCtttTATGATGGTCTGCAG
CCGGTTCTGGCAATTATGGATCCGGCAATTATCAAAAACCATCCTGGTGAAAGAATGCTACAGCaccTTT
ACCAATCGTCGCAATTTTGGTCTGAATGGTCCGCTGGAAAGCGCAATTAGCATTGCCGAAGATGAACAG
TGGAACCGTATTTCGTACGGTTCTGAGCCCCGACCTTTACAAGCGGTAAACTGAAAGAAATGTTTCCGATT
ATGAAACACTATGGCGACATCCTCGTGAAAAACATTAGAAAAAAAgccGAAAAAGATGAGCCGGTGGAT
atgAAAGATATTTTGGCGCCTATAGCATGGATGTTATTACCAGCACCAGCTTTAGCGTTGATATTGAT
AGCATTAAACAATCCGAACGATCCGTTTCGTGACCAATATCAAAAAGCTTCTGAAATTCAACTTTTTTAAC
CCGCTGTTTCATCTGATTGTGATCTTTCCGTTTCTGACACCGgtgCTGGAAAAAATGAATATTTCACTG
TTTTCCCAAGGATGTGATGGACTTTTTTCaaaAACGCCGTGAAAAAAATCAAAGAAGATCGCAAAAAAGAC
AACCACAAAGATCGTGTGATTTCCTGCAGCTCATGATTGATAGCCAGACCAGCGAAGAAaatCATAGC
CATAAAGCACTGACCGATGTTGAGATCTTGGCACAGGCCATTATTTTCATTTTTGCCGGTTACGAAACC
ACTAGTAGCACCCTGAGCTTTCTGGCATATAATCTGGCAACCCATCCTGATGTTGAGCAGAAACTCCAG
GAAGAAATTGACGCAACCCTGCCGAATAAAGCACCGCCTACCTATGATGCACTGATGCAGATGGAGTAC
TTGGATATGGTTATTAATGAAACCCTGCGCCTGTATCCGATTGGTGGTCTGGAACGCGTGTGTAA
AAAACCGTTGAAATTAATGGTGTGACCATTCCGAAAGGCACCGTTGTTATGATTCCGGTTTATGTGCTC
CATCGTGATCCGGAATATTGGCCTGAACCGGAAGAATTCGTCCCGAGCGTTTTAGCAAAGAAAAACAA
GAAAGCATCGACCCGTATACGTTTCTGCCGTTTGGTGCAGGTCCGCGTAATTGTATTGGCATGCGCTTT
GCACTGCTGaccATGAAAGTTGCAATTGTGCGTCTGCTGCAGAATTTAGCTTTCTGTACCTGTAAAGAA
ACCCAGATCCCCTGGAAGTGAATACCCAGGGTTTTCTGacgCCGAAAAAACCGATTGTTCTTAAGCTG
GTTCTCGG

>CYP2D_N1_nucleotide

ATGGCATTATTATCAGGATCAAAATTTATTCGCATTAGGCGTAGCCCTGACCGTCTTCCTGCTTCTGGTG
GACTTCATGAAGCGGCGCAAGCGCTGGTGTCTCGCTTCCCACCAGGCCCTATGTCATTGCCTTTCTGTG
GGCAACCTGCTGCAGGTGGACTTCCATAATCTGCACTACAGCTTCAACCAGCTGAGCAAGCGCTATGGG
GACGTGTTTACGCTACAGCTGTGCTGGAAGCCAGTGGTGTGTGCTCAATGGGTTTCGAGGCCGTGCGAGAG
GCGCTGGTGC AAAAGAGCGAGGACACTGCTGACCGCCCACTGCCCCATCTATGAGCATCTGGGTTTTT
GGGCGGAAGTCCGAAGGTGTGGTCCTAGCACGCTATGGGCAAGCGTGGAAAGGAGCTGAGCGCTTCACT
CTGTCTACCTTGCGCGACTTTGGCATGGGCAAGAAGTCACTGGAGGAGAGGGTGC GCGAGGAGGCCGGC
TTCCTCTGTGCCGCTTCTCTAGCGAGGAAGGACGCCCTTCGACCCCCACTTCTTATTGAACAACGCA
GTGAGCAACGTGATCTGCTCCATCATCTATGGCGAGCGCTTTGAGTACGATGACAAGAAGTTCCAGAAG
CTGCTGCGCCTATTGGAGGAGAGTCTGAAGGAGGAGTCTGGATTCTTACCCAGATTCTGAATGTAGTC
CCGTGGCTCCTGCGCATCCCAGGGCTGGCTCAGAAAGTCTTCCAGGGTCAAAAGGCCTTACTGGCCTTC
TTGGATGAGCTGATCAAGGAACACAAGGAGACCTGGGACCCTGCCACCCAACCCGAGATTTTATTGAT
GCCTTCTCTGCTAGAGATGGAGAAGGCCAAGGGGAATCCTGAGAGCAGCTTCAATGAGGAGAACCTGCGC
CTGACGACGGCTGACCTGTTCTCTGCAGGGACGGAGACCACCTCAACCACACTGCGCTGGGGCCTGCTG
TTCATGCTCCTGTATCCAGATGTGCAGCGCAAGTCCATGAGGAAATCGATAAGGTATAGGGCGGGAC
CGGCGTCCAACGATGGGAGACCAGGTAAAGATGCCCTACACCAATGCTGTCTTTCATGAGGTGCAGCGC
TATGGGGACATTGTTCCACTGGGTGTACCTCACATGACATATCGTGACACTGAACTACAGGGCTTCTTC
ATCCCTAAGGGGACGACCGTCAACCAACCTGTCTCTGTGCTGAAGGATGAGACTGTCTGGGAGAAG
CCCTTCCAGTTCTATCTGAACACTTCTTGATGCGGACGGCCAGTTTGTGAAGCCTGAGGCCTTCATG
CCATTCTCAGCAGGCCGAGAGTATGCCTGGGGGAGCAACTGGCCCGCATGGAGCTCTTCTCTTCTTC
ACCTGCCTCCTGCAGCGCTTTAGCTTCCAAATCCCCGAGGGACAGCCCCGGCCCCGCGAGGATGGAGTC
TTTGCTTTTCTGCTGGCCCCACGACTACAGATATGGTGTGCTGTGCCCCGtgcactCATCAGCTTAC
CATCACTGA

>CPR_N1_nucleotide

ATGGGAGCATCAGACACAGAATCAGGACCATCACTACCAGAACCAGCAGCAGAAGAAGAACCATTCTTA
TCAATGCTAGACGTTTTCTTGCTATCACTGATTCTGGGTCTGCTGGTTTATTGGTTTTTCTTTTCGAAA
AAAAAAGAAGAGATCCCGAATTCAAACCGATTACGCCGGTTGCACCCGACCCTGCGTGAAACCAGCTTT
ATTGAAAAAATGAAAAAACCGGCAAAAACATCGTGGTGTTTTATGGTAGCCAGACCGGTACAGCCGAA
GAATTTGCCAATCGTCTGGCAAAAGATGCACAGCGTTATGGTATGAAAGGTATGGCAGCAGATCCGGAA
GAATGTGATATGGAAGAACTGAGCCGTCTGGGTGAAATGATAATAGCCTGGCAGTTTTTTGTATGGCC
ACCTATGGTGAAGGTGATCCGACCGATAATGCACAGGATTTTTATGATTGGCTGCAAGAGGATGATGCA
GATCTGAGCGGTCTGAAATATGCAGTTTTTGGTCTGGGCAACAAAACCTATGAACATTTTAAATGCCATG
GGCAAAATATGTGGATAAACGTCTGGAAGAACTGGGTGCACAGCGTATTTTTGATCTGGGTCTGGGTGAT
GATGATGCCAATATTGAAGAAGATTTTATCACCTGGCGTGAAACGTTTTTGGCCTGCAGTTTGCGAACAT
TTTGGTATTGAAGCAAGCGGTGAAGATGGTAGCATGCGTCAGTATAGCCTGGTTGAACATACCGATATC
AACAAAAACAAAATCTATACCGGTGAAATGGCACGCTGCATAGTTATGAAAAATCAGAAACCGCCTTTC
GATGCCAAAAATCCGTTTCTGGCACCGGTTACCGTTAATCGTGAACCTGCATAAAGGTGGTGGTCTGAGC

TGTATGCATATTGAACTGGATATCACCGGTAGCAAAATTCGTTATGAAGCCGGTGATCATGTTGCAGTT
TATCCGACCAATGATGCAGCACTGGTTGAGGAACTGGGTAAACTGCTGAATACCAATCTGGATACCGTT
TTTAGCCTGAACAATGTGGATGAAGAGAGCAGCAAAAAACACCCGTTTCCGTGTCCGACCACCTATCGT
ACCGCACTGACCCATTATCTGGATATTACCAATCCGCCTCGTACCAATGTTCTGCATGAACTGGCAGAA
TATGCCAGTGATCCGAAAGATAAAGAATTCCTGCGTAAAAATGGCAAGCAGCTCACCGGAAGGTAAAAAA
CAGTATCAGAGCTGGGTTGTTGAAAGCCGTCGTAATATTCCTGGCGATTCTGCAGGATCTGCCGAGCTGT
CGTCCGCCTATTGATCATCTGTGTGAACTGCTGCCTCGTCTGCAGGCACGTTATTATAGCATTGCAAGC
AGCAGCAAAGTTCATCCGAATAGCATTATATTTGTGCCGTTGTGGTTGAATACGAAACCAAACCGGT
CGTGTTAATAAAGGTGTTGCAACCAATTGGCTGAAAAACAAAGTTCGGCAACGGATAATGGTCATAAA
CCGCGTGTTCCGATTTATGTTTCGTAAAAGCCAGTTTCGTCTGCCGTTTAAACCGAGCACACCGGTTATT
ATGATCGGTCCGGGTACAGGCCTGGCACCGTTTATTGGCTTTATTCAAGAACGTGCCTATCTGAAAGAA
CAGGGTAAAGAAGTTGGCGAAACCGTTCTGTATTTTGGTTGTCGGAAAAAAAACGAGGACTTCATCTAT
CAAGAAGAAGTGAACAGTACCAGAAAAGAAGGTGTTCTGACCCAGCTGAATATGGCATTATAGCCGTGAT
CAGGCACAGAAAGTTTATGTTTCAGCATCTGATCAAAACAGAACAAAAGAAAGCGTGTGGAAACTGATCCAT
GAAGATAACGCCCATATTTATGTTTGTGGTGATGCACGTAATATGGCACGTGATGTTTCAGAATGCACTG
TATGAAATTGTTGCCGAAAAAGGCAATATGACCCAGACACAGGCAGTGGATTATATCAAAAAACTGATG
ACCAAAGGTCGCTATAGCTGTGATGTTTGGAGCTGA

Figure 25

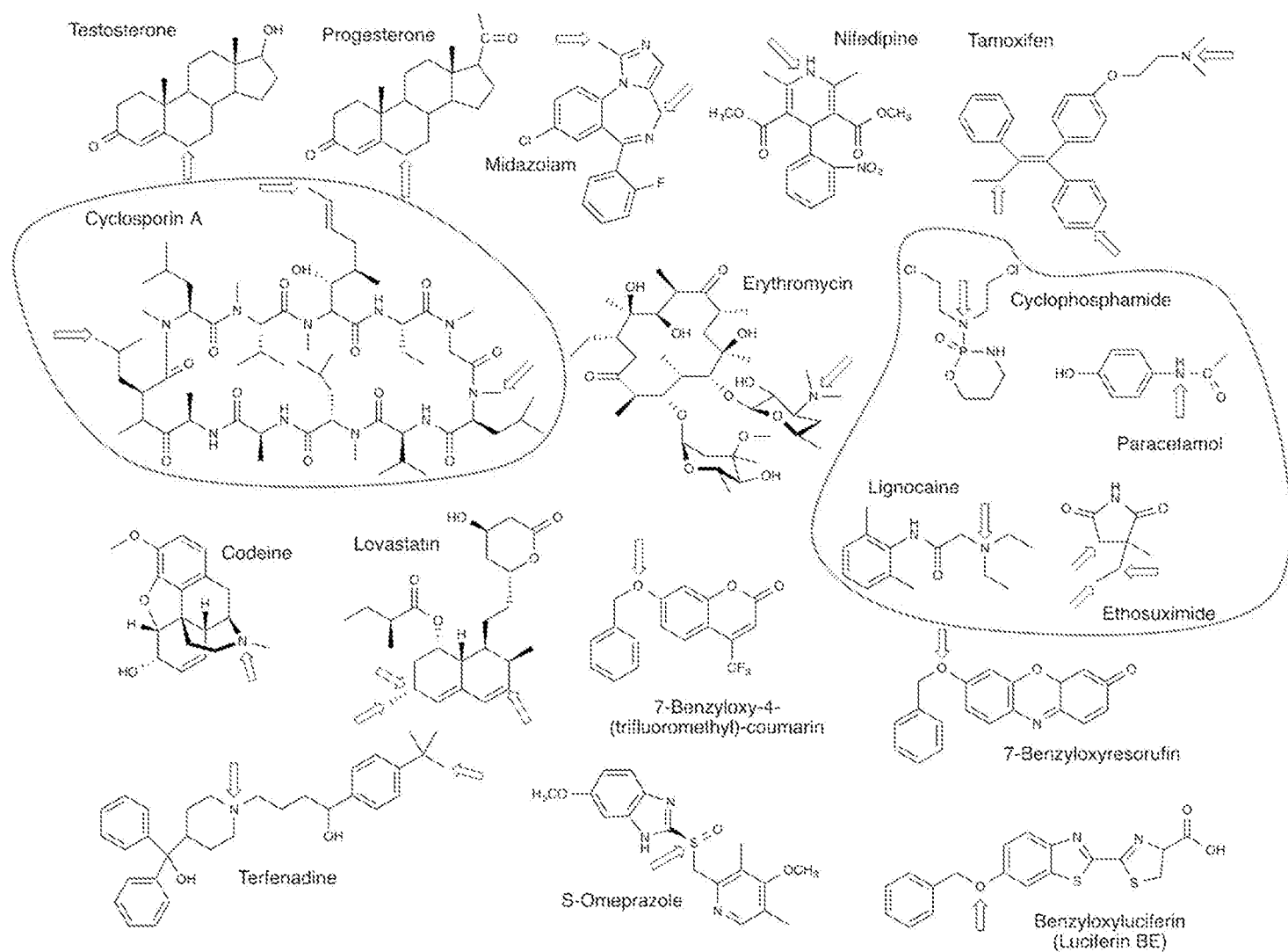
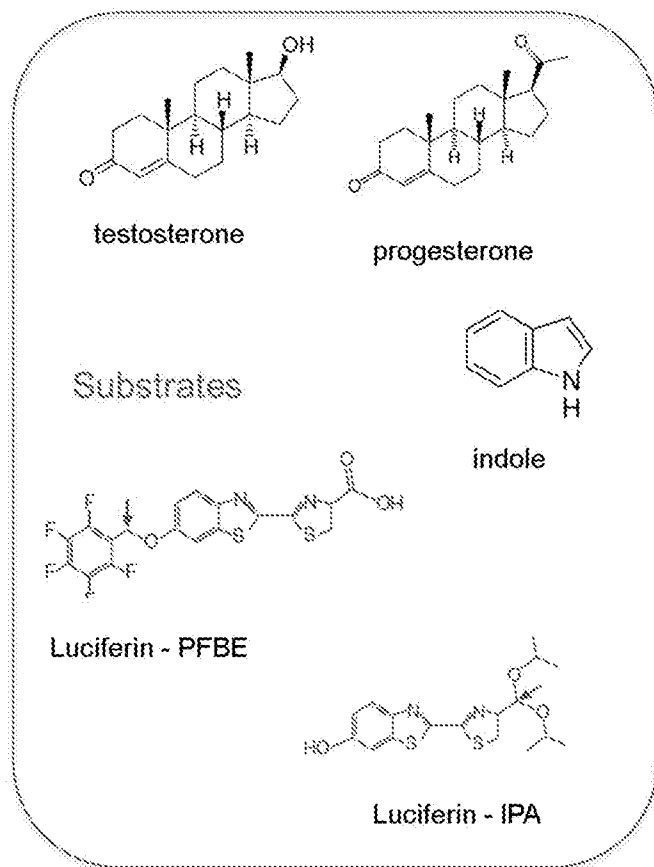
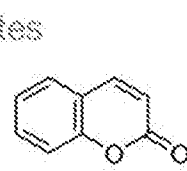


Figure 26

Activity of CYP3_{TS}



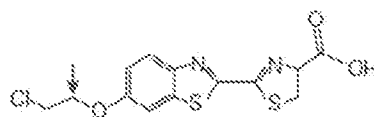
Non - substrates



coumarin

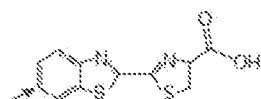


p-nitrophenol

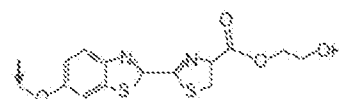


Luciferin - CEE

Luciferin - ME



Luciferin - H



Luciferin - ME- EGE

Figure 27

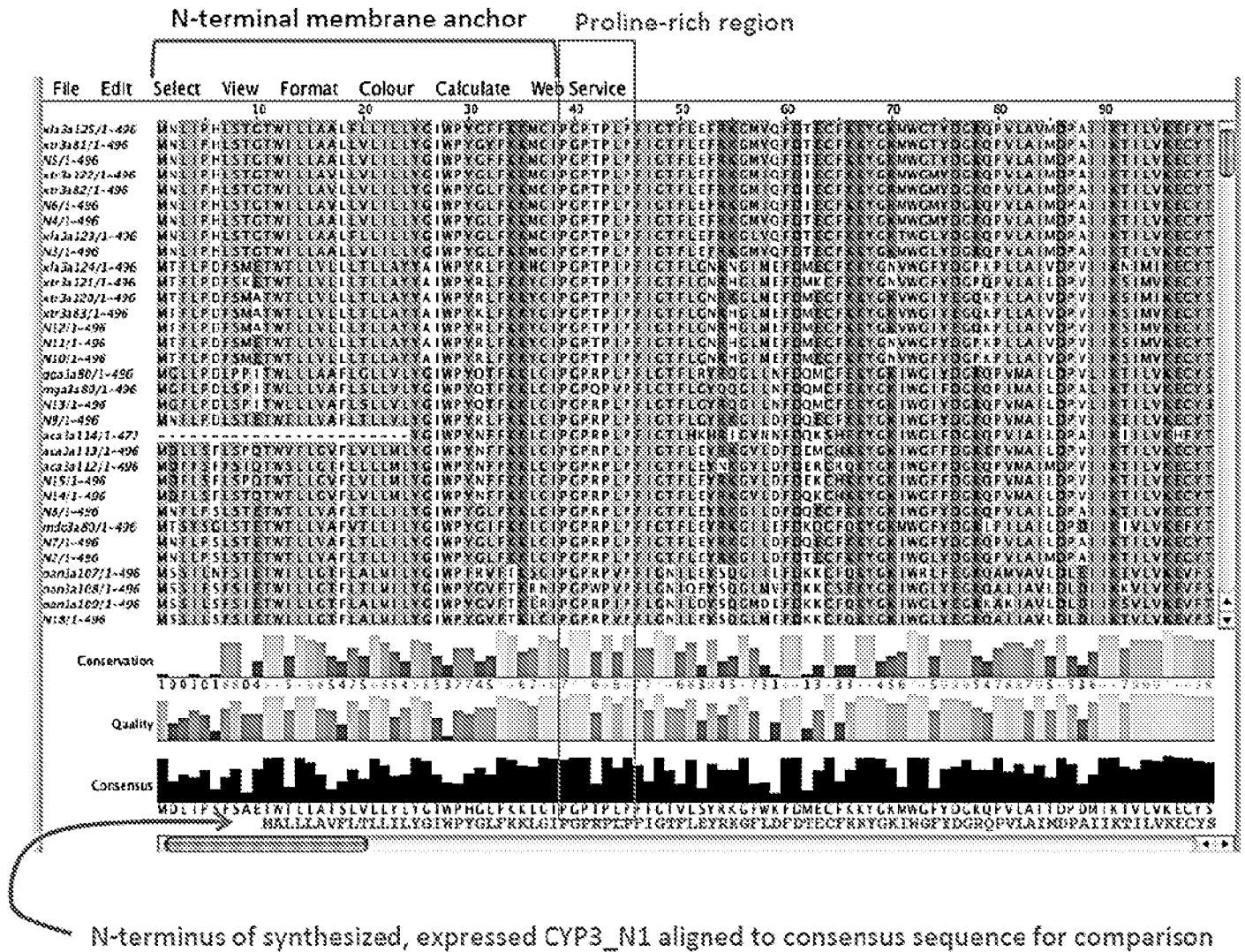


Figure 28

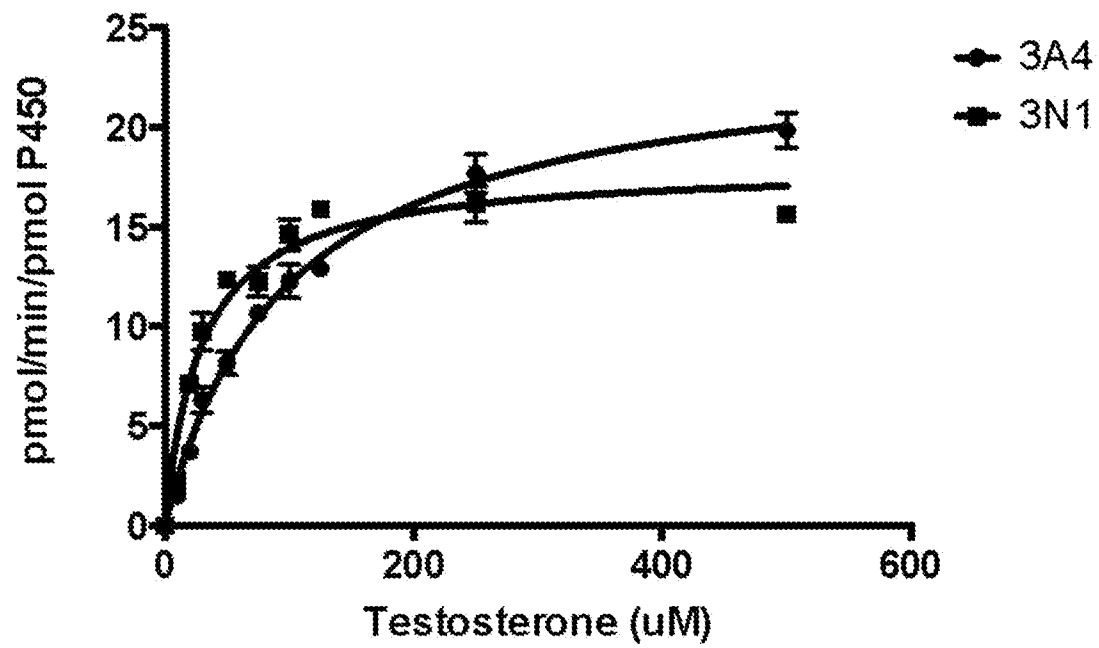


Figure 29

>N1/1-496

MNFLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFIGTFLEYRKGFDFDTECFKKYKGIW
GFYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKVNQKKA EKDEPDMKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLFILIVIFPFLTPVLEKMNISLFPKDVMDFFKNAVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRCTKETQIPLELNTQGFLTPKKPIVLKLVPR

>N2/1-496

MNFLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFIGTFLEYRKGLDFDTECFKKYKGIW
GFYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKVNQKKA EKDEPDMKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLILIVIFPFLTPVLEKMNISLFPKDVMDFFKNAVKKIKEDRKKGNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVTLLQNFSFRPCKETPIPLELNTQGFLQPKKPIVLKLVPR

>N3/1-496

MNLIPHLSTGTWILLALLVLILYGIWPHYGLFKKMGIPGPTPLPFFIGTFLEFRKGLVQDFDTECFKKYKGMW
GLYDGRQPVLAIMDPAAIKTILVKECYSTFTNRRNFGLNGPLESAITVAEDEQWKRI RSVLSPTFTSGKLKE
MFQIMKDYSDILVKNIQGNVDKDEPCAMKDVIGAYSMDVITSTSFVNIDSLNNP SDPFVHIHKNLLKTGLL
NPLIILVIFPFLRPIFEGNLNINFFPKDFTQFFMNAVTSFREKRKKGDHSGRVDFLQLMVDSRTNDSSNQHK
ALTDVEIMAQSIIFIFAGYETTSSTALSFLYNLATHPDVQQLRHEEIDSLPDKASPTYDILMQMEYLDMMVI
QETLRLYPAGRLERVSKQNEINGVSI PKGTVTMIPAYVLQRDPEYWPEPEEFRPERFSKENRETHTPFTF
LPFGDGRNCIGMRFALLSMKVAIVTLLQNFSVRPCAETLIPMEFSTIGVLQPKKPIVLKFLSR

>N7/1-496

MNFLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFIGTFLEYRKGLDFDQECFKKYKGIW
GFYDGRQPVLAIDLPAIIKTILVKECYSTFTNRRNFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIIKHYGDI LKVNQKKA EKDEPDMKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLILIVIFPFLTPVLEKMNISLLPKDVMDFFKNAVKKIKEDRKKGNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPGGRIERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVTLLQNFSFRPCKETPIPLELNTQGFMPKPKPIVLKLVPR

>N8/1-496

MNFLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFIGTFLEYRKGLDFDQECFKKYKGIW
GFYDGRQPVMAILDPAAIKTILVKECYSTFTNRRDFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIIKHYGEILVKNQKKA EKDEPDMKDIFGAYSMDVITSTSFSDIDSINNPNDPFVTNIKKLLKFNF
NPLILIVIFPFLTPVLEKMNISLLPTDVMDFFKDAVKKIKKDRQKGNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPGGRIERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAMVLLQNFSFRPCKETPIPLELNTQGFMPKPKPIVLKLVPR

>N9/1-496

MNFLPDFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFIGTFLEYRKGLDFDQECFKKYKGIW
GFYDGRQPVMAILDPAAIKTILVKECYSTFTNRRDFGLNGPLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIIKHYGEILVKNQKKA EKDEPDMKDIFGAYSMDVITSTSFSDIDSMNNPNDPFVTNIKKLLKFNF
NPLILIVIFPFLTPVLEKMNISLLPTDVMDFFKDAVKKIKKDRQKGNHKKDRVDFLQLMIDSQSSSEKHSYK
ALSDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDATLPNKAPPTYDALMQMEYLDMMVI
NETLRLYPGGRIERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKENIDPYTF
LPFGAGPRNCIGMRFALLTMKVAMVLLQNFSFRPCKDTPIPLELNTQGFMPKPKPIVLKLVPR

>N16/1-496

MNLLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFGTVLEYRKGFDFDTECFKKYKGIW
GFYDGRQPVLAIMDPAAIKTVLVKECYSTFTNRRNFGLAGVLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKVNQKKA EKDEPVTMKDIFGAYSMDVITSTSFVSDIDSINNPNDPFVKEIKKLLKFDF
NPLFILIVIFPFLTPVLEKLNISLFPKAMDFFKRSVKKIKEDRKKDNHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYNLATHPDVQKQLQEEIDAVLPNKAPLTYDALMQMEYLDMMVI
NETLRLYPGGRLERVCKKTVEINGVTIPKGTVMIPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTF
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRCTKETQIPLELNSQGFLTPKKPIVLKLVPR

>N19/1-496

MNLLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFGTVLEYRKGFDFDTECFKKYKGIW
GFYDGRQPVLAIMDPAAIKTVLVKECYSTFTNRRNFPGAGVLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKVNQKKA EKDEPVTMKDIFGAYSMDVITSTSFVNIDSLNNPNDPFVKEIKKLLKFDF
DPLFILIVIFPFLTPVLEKLNISLFPKAMDFFKRSVKKIKEDRKKDTHKKDRVDFLQLMIDSQTSEENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLSFLAYQLATHPDVQKQLQEEIDAVLPNKAPLTYDALMQMEYLDMMVI
NETLRLYPGGRLERVCKKTVEINGVTIPKGTVMIPPVYVLRDPEYWPEPEEFRPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALLTMKVAIVRLQNFSFRCTKETQIPKLNSQGFLTPEKPIVLKLVPR

>N20/1-496

MNLLPSFSIETWTLVFLTLILYGIWPHYGLFKKLGPGRPLPFFGTVLEYRKGFDFDTECFQKYKGIW
GIYDGRQPVLAIMDPAAIKTVLVKECYSTFTNRRNFPGAGVLESAISIAEDEQWKRI RTVLSPTFTSGKLKE
MFPIMKHYGDI LKVNQKKA EKDEPVTMKDIFGAYSMDVITSTSFVNIDSMNNPKDPFVKEIKKLLKFDF

DPLFILLFIFPFLTPVMEKLNISIFPKDAMDFFMRSVKKIKEDRKKETHKDRVDFLQLMIDSQTSDENHSHK
ALTDVEILAQAIIFFIFAGYETTSSTLGYLAYQLATHPDVQKQLQEEIDAVLPNKAPLTYDALMQMEYLDMMVI
NETLRLYPLGGRLERVCKKDVEINGVTIPKGTVMIPPYVLHRDPEYWPEPEEFRPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALLTMKVAITRLLQNF'SFRTCKETQIPLKLNQSGFLTPEKPIVLKLVPR
>N23/1-496
MNLIPSFSETWTLLVFLTLTLLYLYGTHSHGLFKKLGIPGPTPLPFFGTVLSYRKGFDFDMECFKKYKGIW
GFYDGRQPVLAIMDPEIITVLVKECYSVFTNRRSFGPAGILESASISAEDEQWKRIRTVLSPTFTSGKLKE
MFPIIKQYGDVVLVKNMRQEAEGKGPVTMKDIFGAYSMDVITSTSGVNIIDSLNNPNDPFFVEIKKLLKDFDF
DPLFLSIIIFPFLTPVFEKLNISIFPKDAMDFFKKS VKKIKEDRKKDTQKHRVDFLQLMIDSQTSKESHSHK
ALSDVEILAQSIIFIFAGYETTSSTLSFLMYQLATHPDVQKQLQEEIDAVLPNKAPVTYDALVQMEYLDMMVI
NETLRLYPIAGRLEVRCKKTVINGVTIPKGTVMIPPYVLHRDPEYWPEPEEFRPERFSKENKESINPYIY
LPFGAGPRNCIGMRFALMSMKVAIVRLLQNF'SFRTCKETQIPLKLNQGLLTPEKPIVLKLVPR
>N30/1-496
MDLIPNFSMETWVLLATSLVLLYLYGTHSHGLFKKLGIPGPTPLPFFGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVTMKDIFGAYSMDVITGTSGVNIIDSLNNPQDPFVENVKLLKDFDF
DPFFLSIILFPFLTPVFEALNISMFPKDVMMNFKKSVKRMKENRLKDKQKHRVDFLQLMINSQNSKETESHK
ALSDVEIVAQSIIFIFAGYETTSSTLSFIMYELATHPDVQKQLQEEIDAALPNKAPATYDALVQMEYLDMMV
NETLRLYPIAGRLEVRCKKDVEINGVFIPKGTVMIPYALHRDPKYWPEPEEFRPERFSKKNKDSINPYIY
LPFGTGRNCIGMRFALNMMLALVRVLQNF'SFKPCKETQIPLKLSNQGLLQPEKPIVLKVVSR
>N31/1-496
MDLIPNFSMETWVLLATSLVLLYLYGTHSHGLFKKLGIPGPTPLPFFGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVTMKDIFGAYSMDVITGTSGVNIIDSLNNPQDPFVENVKLLKDFDF
DPFFLSIILFPFLTPVFEALNISMFPKDVMMNFKKSVKRMKENRLKDKQKHRVDFLQLMINSQNSKETESHK
ALSDVEIVAQSIIFIFAGYETTSSTLSFIMYLLATHPDVQKQLQEEIDAALPNKAPATYDTLVQMEYLDMMV
NETLRLYPIAGRLEVRCKKDVEINGVFIPKGTVMIPYALHRDPKYWPEPEEFRPERFSKKNKDSINPYIY
LPFGTGRNCIGMRFALNMMLALVRVLQNF'SFKPCKETQIPLKLSKQGLLQPEKPIVLKVVSR
>N32/1-496
MDLIPNFSMETWVLLATSLVLLYLYGTHSHGLFKKLGIPGPKPLPFFGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVTMKDIFGAYSMDVITGTSGVNIIDSLNNPQDPFVENVKLLKDFDF
DPFFLSIILFPFLTPVFEALNISMFPKDVMMNFKKSVKRMKENRLKDKQKHRVDFLQLMINSQNSKDKESHK
ALSDVEIVAQSIIFIFAGYETTSSTLSFIMYLLATHPDVQKQLQEEIDAALPNKAPATYDTLVQMEYLDMMV
NETLRLYPIAGRLEVRCKKDVEINGVFIPKGTVMIPYALHRDPKYWPEPEEFRPERFSKKNKDSINPYIY
LPFGNGPRNCIGMRFALNMMLALVRVLQNF'SFQPKCKETQIPLKLSKQGLLQPEKPIILKVVSR
>N33/1-496
MDLIPNFSMETWMLLATSLVLLYLYGTHSHGIFKKLGIPGPKPLPFLGTILAYRKGFWEFDKQCHKKYGKMW
GLYDGRQPVLAITDPDIITVLVKECYSTFTNRRSFGPVGILKKAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIINQYADVLVKNMRQEAEGKGPVTMKDIFGAYSMDVITGTSGVNIIDSLNNPQDPFVEKVKLLKDFDF
DPFFLSVILFPFLTPVFEALNISMFPKDVMMNFKTSVERMKENRMKEKEKQKRVDFLQLMINSQNSKDKESHK
ALSDVEIVAQSIIFIFAGYETTSALSFLYLLATHPDVQKQLQDEIDAALPNKAPATYDTLVQMEYLDMMV
NETLRLYPIAGRLEVRCKKDVEINGVFIPKGTVMIPTFALHKDPKYWPEPEEFRPERFSKKNQDSINPYIY
LPFGNGPRNCIGMRFALNMMLALVRVLQNF'SFQPKCKETQIPLKLSKQGLLQPEKPIILKVVSR
>N34/1-496
MDLIPNFSMETWMLLATSLVLLYLYGTHSHGIFKKLGIPGPKPLPFLGTILAYRKGFWEFDKQCHKKYGKMW
GLYDGRQPVLAITDPDIITVLVKECYSTFTNRRSFGPVGILKKAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIINQYTDVLVRNMRQGESEEGKPTSMKDIFGAYSMDVITATSGVNIIDSLNNPQDPFVEKVKLLKDFDIF
DPLFLSVTLFPFLTPVFEALNVSMFPRDVINFKTSVERMKENRMKEKEKQKRVDFLQLMINSQNSKVKESHK
ALSDVEIVAQSVIFIFAGYETTSALSFLYLLATHPDVQKQLQDEIDAALPNKAPATYDTLLQMEYLDMMV
NETLRLYPIAGRLEVRCKTDVEINGVFIPKGTVMIPTFALHKDPKYWPEPEEFRPERFSKKNQDSINPYIY
LPFGNGPRNCIGMRFALNMMLALVRVLQNF'SFQPKCKETQIPLKLSKQGLLQPEKPIILKVVSR
>N35/1-496
MDLIPNFSLETWVLLAISLVLLYLYGTHSHGLFKKLGIPGPKPLPFFGTVLNRYKGMWKFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVTMKEIFGAYSMDVITGTSGVNIIDSLNNPQDPFVEKAKKLLKDFDF
DPFFLSIILFPFLTPVFEALNISMFPKDSMNFKKFVKRMKENRLDSNQKHRVDFLQLMINSQNSKDKESHK
ALSDVEIVAQSIIFIFAGYETTSSTLSFIMYSLATHPDVQKQLQEEIDRALPNKAPPTYDTLMEMEYLDMMV
NETLRLYPIANRLERVCKKDVEINGVFIPKGTVMIPYALHRDPKYWPEPEEFRPERFSKKNKGSIDPYIY
LPFGNGPRNCIGMRFALNMMLALIRVLQNF'SFQPKCKETQIPLKLSRQGLLQPEKPIILKVVSR
>N36/1-496
MDLIPNFSLETWVLLAISLVLLYLYGTHSHGLFKKLGIPGPKPLPFFGTVLNRYKGMWKFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLLSPFTSGKLKE
MFPIIKQYGDVVLVKNLRQEAEGKGPVTMKEIFGAYSMDVITGTSGVNIIDSLNNPQDPFVEKAKKLLKDFDF
DPFFLSIILFPFLTPVFEALNISMFPKDSMNFKKFVKRMKENRLDSNQKHRVDFLQLMINSQNSKDKESHK
ALSDVEIVAQSIIFIFAGYETTSSTLSFIMYSLATHPDVQKQLQEEIDRALPNKAPPTYDTLMEMEYLDMMV
NETLRLYPIANRLERVCKKDVEINGVFIPKGTVMIPYALHRDPKYWPEPEEFRPERFSKKNKGSIDPYIY
LPFGNGPRNCIGMRFALNMMLALIRVLQNF'SFQPKCKETQIPLKLSRQGLLQPEKPIILKVVSR

>N51/1-496

MDLIPNFSMETWVLLATSLVLLLYLYGTHSHGLFKKLGIPGPTPLPFFGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKSAISISEDEEWKRIRTLTLLSPTFTSGKLKE
MFPFIISQYGDVLRNLRQEAEGKGPVTMKDIFGAYSMDVITGTSFGVNIDSLNPNQDPFVENTKLLKDFDL
DPFFLSIILFPFLTPVFEALNISMFPKDVTNFFKKSVMKMSRLKDKQKHRVDFLQLMINSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQKQLEEIDAALPNKAPATYDALVQMEYLDMMV
NETLRLYP IAGRLERVCKKDVEINGVFIPKGTVMIMPTIYALHRDPKYWPEPEEFRPERFSKKNKDSINPYIY
LPFGTGPRNCIGMRFALNMMLALVRVLQNFSEFKPCKETQIPLKLSNQGLLQPEKPIVLKVESR

>N52/1-496

MDLIPNFAMETWVLLATSLVLLLYLYGTHSHGLFKKLGIPGPTPLPFLGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISISEDEEWKRIRTLTLLSPTFTSGKLKE
MFPFIISQYGDVLRNLRQEAEGKGPVTMKDIFGAYSMDVITGTSFGVNIDSLNPNQDPFVENTKLLKDFDL
DPFFLSIILFPFLTPVFEALNISMFPKDVTNFFKKSVMKMSRLKDKQKHRVDFLQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQKQLEEIDAALPNKAPATYDALVQMEYLDMMV
NETLRLFP IAGRLERVCKKDVEINGVFIPKGTVMIMPTIYALHRDPKYWTEPEEFRPERFSKKNKDSIDPYIY
MPFGTGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSNQGLLQPEKPIVLKVESR

>N53/1-496

MDLIPNFAMETWVLLATSLVLLLYLYGTHSHGLFKKLGIPGPTPLPFLGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISIAEDEEWKRIRTLTLLSPTFTSGKLKE
MFPFIISQYGDVLRNLRQEAEGKGPVTMKDIFGAYSMDVITGTSFGVNIDSLNPNQDPFVENTKLLKDFDL
DPFFLSIILFPFLTPVFEALNISMFPKDVTNFFKKSVMKMSRLKDKQKHRVDFLQLMIDSQNSKETESHK
ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQKQLEEIDAALPNKAPATYDALVQMEYLDMMV
NETLRLFP IAGRLERVCKKDVEINGVFIPKGTVMIMPTIYALHHDPKYWTEPEEFRPERFSKKNKDSIDPYIY
MPFGTGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSNQGLLQPEKPIVLKVESR

>N78/1-496

MDLIPNFAMETWVLLATSLVLLLYLYGTHSHGLFKKLGIPGPTPLPFLGTVLSYRKGFWEFDMECYKKYGKMW
GLYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISISEDEEWKRIRTLTLLSPTFTSGKLKE
MFPIMSQYGDVLRNLRQEAEGKGPVTMKDIFGAYSMDVITGTSFGVNIDSLNPNQDPFVENTKLLKDFDL
DPFFLSIILFPFLTPVFEALNISMFPKDVTNFFKKSVMKMSRLQDKQKHRVDFLQLMIDSQNSKETESHK
VLSDELELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQKQLEEIDAALPNKTPATYDALVQMEYLDMMV
NETLRLFP IAGRLERVCKKDVEINGVLPKGTVMIMPTIYALHRDPKYWTEPEEFRPERFSKKNKDSIDPYIY
MPFGSGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSNQGLLQPEKPIVLKIESR

>N79/1-496

MDLIPFSMETWVLLATSLVLLLYLYGTYSHGLFKKLGIPGPTPLPFFGTVLYRKGFWEFDDKCFKKYGKMW
GFYDGRQPVLAITDPDMIKTVLVKECYSVFTNRRSFGPVGFMKNAISLSEDEQWKIRTLTLLSPTFTSGKLKE
MFPFIIGQYGDVLRNLRKEAEGKGPVTMKDIFGAYSMDVITSTSGVNIDSLNPNQDPFVENTKLLKDFDL
DPFFLSIILFPFLTPVFEVLNISLFPKSVTNFFTKSVKRMKESRLKDKQKHRVDFLQLMINSQNSKETDTHK
ALSDLELVAQSIIFIFAGYETTSSLSFLMYELATHPDVQKQLEEIDAATFNPAPPTYDALVQMEYLDMMV
NETLRLYP IAGRLERVCKKDVEINGVFIPKGTVMVPTFALHRDPKELWPEPEEFRPERFSKKNKDSINPYTY
LPFGTGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSQGLIPEKPIVLKVEPR

>N100/1-496

MNFLPSFSIETWTLVLVLFITLLILYGIWYPYGFKKLGIPGPRPLPFIPTFLEYRKGFNFDETECFKKYGKVW
GIYDGRQPVLCIMDPAMIKTILVKECYSTFTNRRNFGNLNGPLEDAISIAEDEQWKIRSVLSPSFTSGRLKE
MFPIMKHHSDNLIKMQKKADEKDEPVDIKDIFGPYSMDVVTSTSFSDIDSINNPDPFVTNIKLLKFNFF
NPLFLVLVLPFLAPILEKMDISLFPKSVMDFFYNALKKIKSDRNKDNHKNRVDFLQLMIDSQTSSEEDQSHK
GLTDHEILSQAMIFIFAGYETTSSVLSFLAYNLATNPVQKQLEEIDAATFNPAPPTYDALVQMEYLDMMV
NESMRLYP IAPRLERVCKKTVEINGVTIPKGTVMVPTFALHRDPKELWPEPEEFRPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSQGLIPEKPIVLKVEPR

>N101/1-496

MGFLPSFSIETWTLVLVLFITLLIVYGYWYPYGFKKLGIPGPRPLPFIPTFLEYRKGFNFDETECFKKYGKVW
GIYDGRQPVLCIMDPAMIKTILVKECYSLFTNRRNFGNLNGPLEDAISIAEDEQWKIRSVLSPSFTSGRLKE
MFSIMKHHSGNLIKSMQKKADEKDEPVDIKDIFGPYSMDVVTSTAFSVDIDSINNPDPFVTNIKLLKFNFF
NPLFLVLVLPFLAPILEKMDISLFPKSVMDFFYNALKKIKSDRNKDNHKNRVDFLQLMIDSQTSSEEDQSHK
GLTDHEILSQAMIFIFAGYETTSSVLSFLAYNLATNPVQKQLEEIDAATFNPAPPTYDALVQMEYLDMMV
NESMRLYP IAPRLERVAKTVEINGVTIPKGTVMVPTFALHRDPKELWPEPEEFRPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSQGLIPEKPIVLKVEPR

>N102/1-496

MGYLPFFSVETWTLVLVLFITLLVYGYWYPYGFKKLGIPGPKPLPFFGTMLLEYRKGFNFDETECFKKYGKVW
GIYDGRQPVLCIMDPAMIKTILVKECYSLFTNRRNFGNLNGPLYDAVIAEDDQWRRIRSVLSPSFTSGRLKE
MFSIMKHHSSNLINSMQKKADEKDEPVDIKDFFGPYSMDVVTSTAFSVDIDSINNPDPFVTNIKLLKDFDF
NPLFLVLVLPFLAPILEKMDISLFPKSVMDFFYALQKIKSDRDTSNHKNRVDFLQLMIDSQTSSEGEQSDK
GLTDHEILSQAMIFIFAGYETTSSVLSFLAYNLATNPVQKQLEEIDAATFNPAPVQYEAALVQMEYLDMMV
NESMRLYP IAPRLERVAKTVEINGITIPKDMVVMIPVYALHRDPKELWPEPEEFRPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALNMMLALIRVLQNFSEFKPCKETQIPLKLSQGLIPEKPIVLKVEPR

>N103/1-496

MGYLPFFSAETWTLVLVLFITLLVYGYWYPYGFKKLGIPGPKPLPFFGTMLLEYRKGFNFDETECFKKYGKVW
GIYDGRQPVLCIMDPAMIKTILIKECYSLFTNRRNFGNLNGPLYDAVIAEDDQWRRIRSVLSPSFTSGRLKE
MFGIMKHHSSNLINSMKKADEKDEPVDIKEFFGPYSMDVVTSTAFSVDIDSINNPDPFVTNIKMLKDFDF

NPLFLLVALFPFMAPILEKMDFSFFPKSVTDFFYAALQKIKSDRDTSNHKNRVDFLQLMIDSQTSEGEQSDK
GLTDHEILSQAMIFIFAGYETSSSTLSFLAYNLATNPEVMKKLQEEIDATFPNKAQVQYEALMQMEYLDQCVI
NESRLRYP IAPRLERVAKTVEINGIVIPKDMVVMIPYALHRDPELWPEPEEFKPERFSKENKESIDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEILQRF'SFSVCKETEIPLEMDTQGLLTPKNPIKLLKLVPR
>N108/1-496
MGYLPFFSAETWTLLVLLITLLVVYGYWPGYGFKKLGIPGPKPLPFFGTMLAYRKGFNMFDTECFKKYGVW
GIYDGRQPVLCIMDPAMIKTVLIKECYSFTNRRNFRNLNGPLYDAVSAEDDQWRRIRSVLSPSFTSGRLKE
MFGIMKHHSSNLINSMKKKADKDEPVEIKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDFF
NPLFLLVALFPFMGPPILEKMEFSFFPKSVTDFFYAALQKIKSDRDTSNHKNRVDFLQLMIDSQKSSGEQQDK
GLTDHEILSQAMIFIFAGYETSSSTLSFLAYNLATNPHVMKKLQEEIDATFPNKAQVQYEALMQMEYLDQCVI
NESRLRYP IAPRLERVAKTVEINGIVIPKDMVVMVPTWALHRDPELWPEPEEFKPERFSKENKESIDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEILQRF'SFSVCKETEIPLEMDTQGLLMPKRPIKLLKLVPR
>N109/1-496
MGYFLFFSAETWTLLVALITLLFVYAYWPGYGFKKLGIPGPKPIPFFGTMLAYRKGFNMFDDEECFKKYGVW
GIYDGRQPVLCITDPAMIKTVLIKECYSFTNRRNFRNLNGPLYDAVSAEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSSANLINSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDFF
NPLFLIVAFFPFMGPPILEKMEFSFFPKSVTDFFYAALQKIKSNRET SKQKSRVDFLQLMIDSQKLSGEEQDK
GLSDHEILSQAMIFIFAGYETSSSTLTFLAYNLATNPHVMKKLQEEIDATFPNKAQVQYQALMQMEYLDQCVI
NESRLRYP IAPRLERVAKTVEINGIVIPKDMVVMVPTWPLHRDPELWPEPEEFKPERFSKENKETIDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEILQRF'SFSVCKETEIPFEMDIQGLLMPKRPIKLLKLVPR
>N110/1-496
MGYFLFFSAETWTLLVALITLLFVYVYWPGYGFKKLGIPGPKPIPFFGTMLAYRKGFNMFDDEECFKKYGVW
GIYDGRQPVLCITDPAMIKTVLIKECYSFTNRRNFRNLNGPLYDAVSAEDDQWRRIRSVLSPSFTSGRLKE
MFDIMKHHSSANLISSMKKKADKDEPLELKEFFGPYSMDVVTSTAFSVDIDSLNNP SDPFVTNKKMLKFDFF
NPLFLIVAFFPFMGPPILEKMEFSFFPKSVTDFFYAALQKIKSNRET SKQKSRVDFLQLMIDSQKLSGEEQDK
GLSDHEILSQAMIFIFAGYETSSSTLTFLAYNLATNPHVMKRLQEEIDATFPNKAQVEYQALMQMEYLDQCVI
NESRLRYP IAPRLERVAKTVEINGIVIPKDMVVMVPTWPLHRDPELWPEPEEFKPERFSKENKETIDPYTY
MPFGAGPRNCIGMRFALVMMKLAIVEILQRF'SFSVCKETEIPFEMDIQGLLMPKRPIKLLKLVPR
>N119/1-496
MGYLPFFSVETWILLITFICLFVIYGNWTFGIFEKLGIPGPKPLMYFGTIGRYNKVYYLDDQECAQKYGVW
GMYELRRPMLAVMDPDMKTILVKECTYFTNRRNFRNLNGPLYDAVSTAEDDHWRRIRNVLSPTSFTSGRIKE
MFSIMKHHSSKLTASLQSKAENDEVITIKDFFGPYSMDVMTSCAFSVDMDSLNNP SSPLITHASKLKFDF
NPLFLLQGVFPFLLPILMGLISFFPKTSTAFFKTIVEKIRAERDGS SHQNSGDFLQHMINSQQSKGEQNDK
GLTDHEILSQATMFVAGYETSATTLAFLAYNLATNPEVMKRLQEEIDSTFPNKAQVQYEALMQMEYLDQCVI
NESRLRYP IAPRLERVAKTVEISGITIPKDMVVMIPVYALHRDPELWPEPEEFKPERFSQKQKQINPYTY
LPFGIGPRNCIGMRFALVMMKLAIVEILQRF'SFSVCKETEIPKMDPQGLVGPLNPIKLLKLVTR
>N121/1-496
MGFLPNFSIETWTILVLLITLITVYGYAPYGFKKLGIPGPRPLPFIPTFLEYRKGIHNFDETCYKKYGVW
GIYDGRQPVLCIMDTAMIKTILVKECYSFTNRRDGLNGPLRDAISVAEDEQWKRIRSVLSPSFTSGRLKE
MYSIMLQHSGLNLIKSLHKKAEAEVVDKDFGPYSMDVVTSTAFSVDIDSLNNP SDPFVANIKKLVKFNFL
NPLFLLVVLFPFLEPILDKMDVSLFPAHVLDFFYNFLKIKSDRNKDKHKNRVDFMQLMVD SQMNEEDSSHK
GLTDNEILSQAMIFIFAGYETSSSTLGLFLAYNLATNPDIQKKLQEEIDETFPNKAQPTYDALMQMEYLDQCVI
NESMRLLYP IANRLERVAKTVEINGVTIPKGTVMIPVYALHRDPALWPEPEEFKPERFSKENKESIDPYTY
LPFGAGPRNCIGMRFALLMMKLAIVEILQNF'SFVTCKETQIPLELSTEGFVTPKNPIKLLKLEPR
>N122/1-496
MDLLPNFSIETWTILVLFVTLITVYGYAPYGFKKLGIPGPKPWPFIFTFLNYRKGVHFFDEECYKKYGVW
GLYDGRQPLMCMIDTAMIKTILVKECYSFTNRRDLGLNGPLRDAVSAEDEQWKRIRSVLSPSFTSGRLKE
MYTIMLQHSKNLINFLHKKVEAEVIDVDVFGPYSDVVTSTAFSVDIDSLNNP SDPFVANIKKLVKFNFL
NPLLLVVLFPFPTQPIFDKMDVSLFPAHVLDFFYNFLRQIKSDRNKDKKSRVDFMQLMVNAQMNEEDSSQK
GLTDNEILSQAMIFIFAGYETSSSTLGLFLAYNLATNPDIQKKLQEEIDKTFPGKVRPNYDDLMQLEYLDQCVI
NESMRLLYP IANRLERMAKTVEINGVTIPKGTVAIPVYALHRDPALWPEPEAFKPERFSKENKDNIDPYTY
LPFGAGPRNCIGMRFALVMMKLAIAEILQHSFVTCKETEIPMVLGTEGFVTPKNPIKLLKLEPR
>N125/1-496
MIDLSSFSIVTWTLLVLFITLLLLYGIWPGYGFKKLGIPGPRPLPFIPTFLSYRKGFNFDECAKKYGVW
GIYDGRQPLVLMVTDPEMIKTILVKECYSTFTNRRNFGSLGPLEDGITIVKDERWKIRISTLSPCFTSGRLKE
IFPIVMRHADRLIKNMQKADPDEPVDIKEIVAPYSLDVVTSTSFSDIDSLNNP SDPFVANIKKLVKFNFL
NPLFLLVALFPFLARLLKKMGISLFPSTMDFFYNALKKIKDEHNKDNHNNRVDFLQLMIQSQIVKEDQPPK
GLTDHEILSQAFIFIFAGYETTTSTLFLYLNATNPDLVQLQEEIDANFPNDAPITYDALMRMEYLDQCVI
NESMRLLYP IAPRLERVCKTVEINGITIPKGTVVGIPVYVLRDPELWPSPEEFRPERFSKENKEINPYAF
MPFGLGPRNCIGMRFALLIMKLVI VRLQNF'SFVTCKETQIPLELNLQGFQPKKPIKLLKLVPR
>N126/1-496
MIDLSSLSIVTWTLLVVLVITLLFIYGVWPHGFFKKLGIPGPRPWPFVGTFLSYTKGFFNFDMCAKKYGVW
GIYDGRQPLVLMVTDLEMIKTILVKECYSTFTNRRNVDLAGPLADGITMVKDERWKIRISTLSPYFTSGRLKE
IFPIAMTHADRFIKNMQKADPEQPVKIKEVVAPYSLDVVTSSSFSDIDSLNNP SDPFVANIKKLVKFNFL
SPLFLLVALFPFLANLLKKMGISVFSRSTMDFFYKALRKIKDEHNKDNHNNRVDFLQLMIQSQIVKSEQPVK
GLTDHEILSQSFIFILGGYETTTSTLFLYLNATNPDCLEKLVEEIDKNFPLDTPITYDALMRMDYLEMAI
NESMRLLYP IAPRLERVCKTVEINGITIPKGTLVGIPTVYVLRDPELWESPQEFRPERFSPENKSEFLQYAF
MPFGLGPRNCIGMRFALLIMKLVLVVKLLQNF'SFVTCKETQIPLELNLVGFQPKVPI TLKLI PR

>N129/1-496

MMDFTLFSATTWTLLALFFTL LLLYGIWPHYRLFKKLGI PGPRPLPFIGTILYYRKGLFAFDRECQAKYGDVW
GIFDGRLPVLMVSDPEMIKTVLVKECYSVFTNRRDFVFSGPLEDGVT VVKDERWKRIIRSTISPCFTSGRLKQ
VFPIVARYADR LIKKLEQTANLDEPIDVKQFVAPYSLDVVTSASF SVEADSINNPD DPFIVNIKKVLKF NFW
NPVFLLLAIFPFSARLLKLLGIDL LPRS SVDFFY NVIKKFKDQHKADNE STRADFLQVMIQSEI I KQDQPPK
GLTEHEILSQAFIFIFGGYETTSTT LSFLLYNLATNP DALQTLQKEIDANLPKDAPISYEDLMGLQYLDQVL
NESMR LIPTAPRLERVCKKT VQIHGLTIPEGTVVGIPVHLLHKDPRFWSSPEEFRPERFSKENE EEVN PYAF
MPFGLGPRNCVGMRYAILVMKMVIVRLLQSYTVETCKDTMIPLEFDWKGFQPLKPIKLFVPR

Figure 30

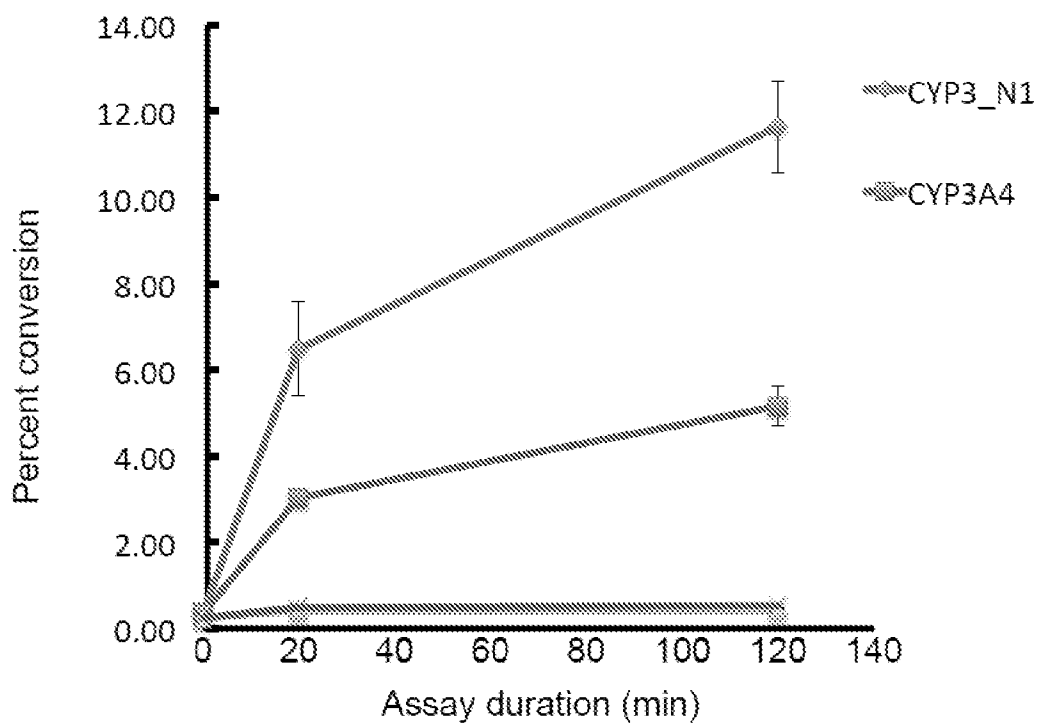


Figure 31

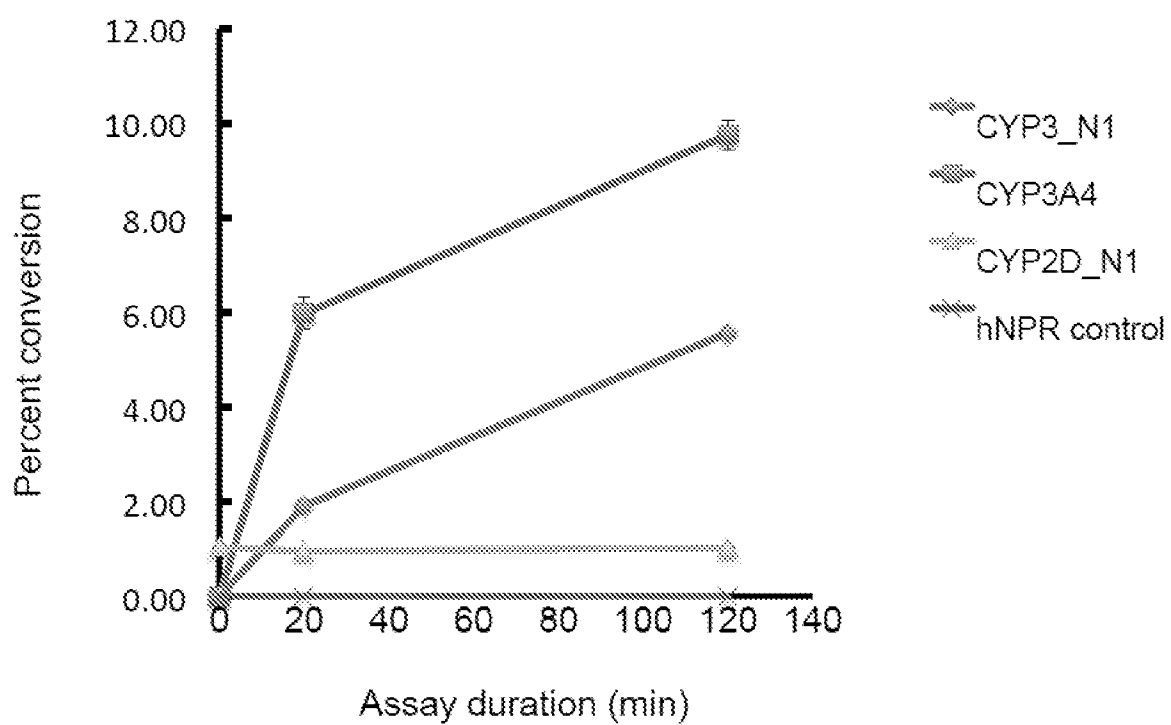


Figure 32

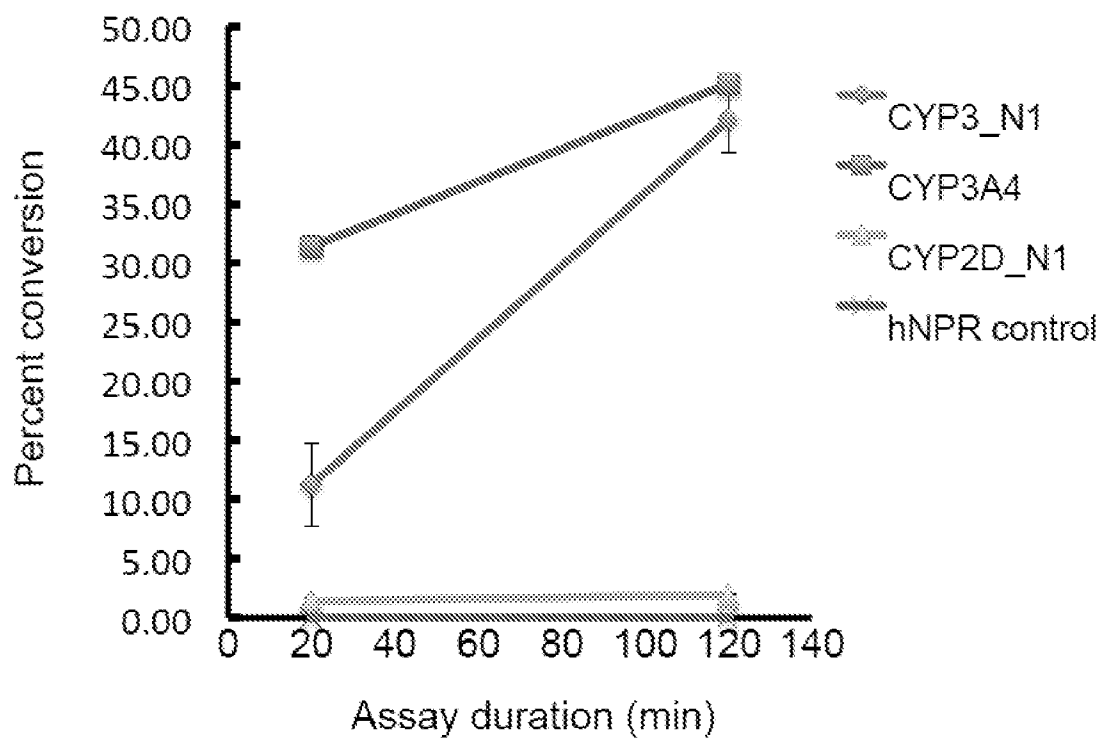


Figure 33

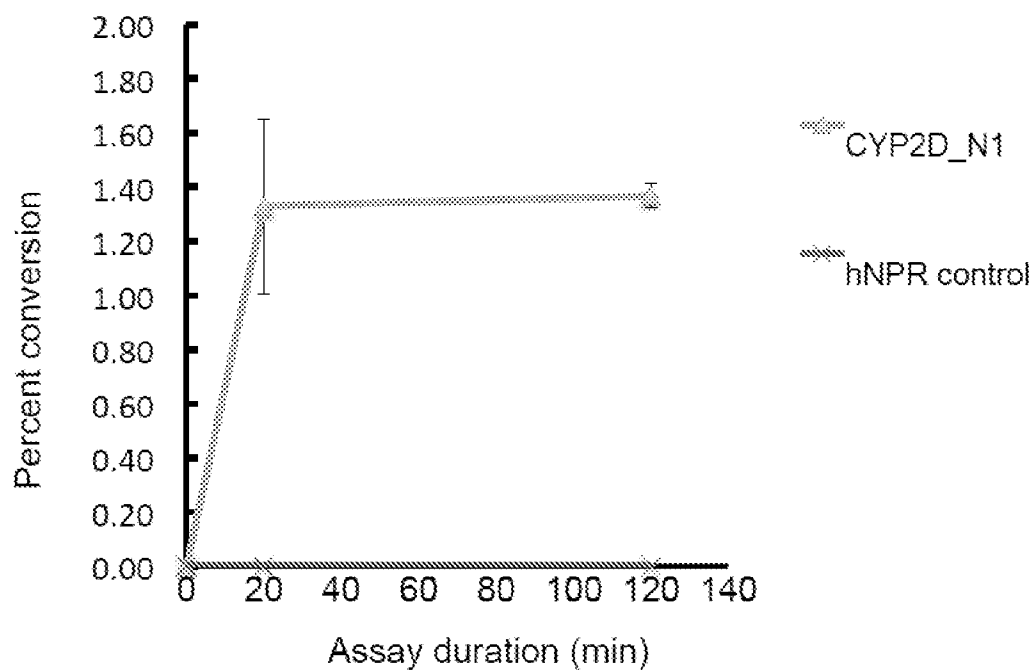


Figure 34

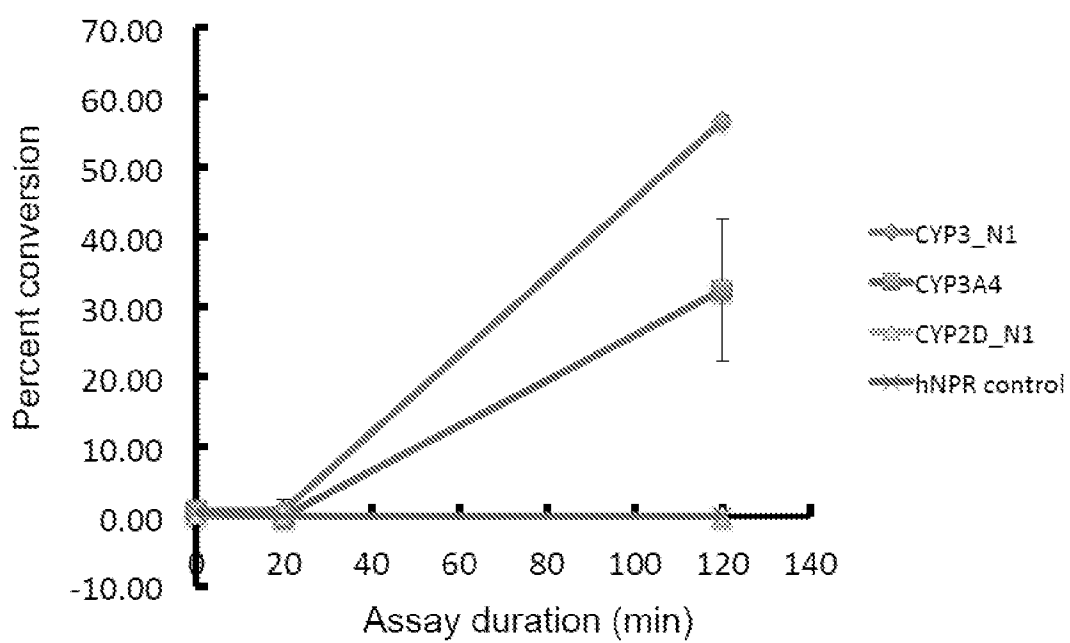


Figure 35

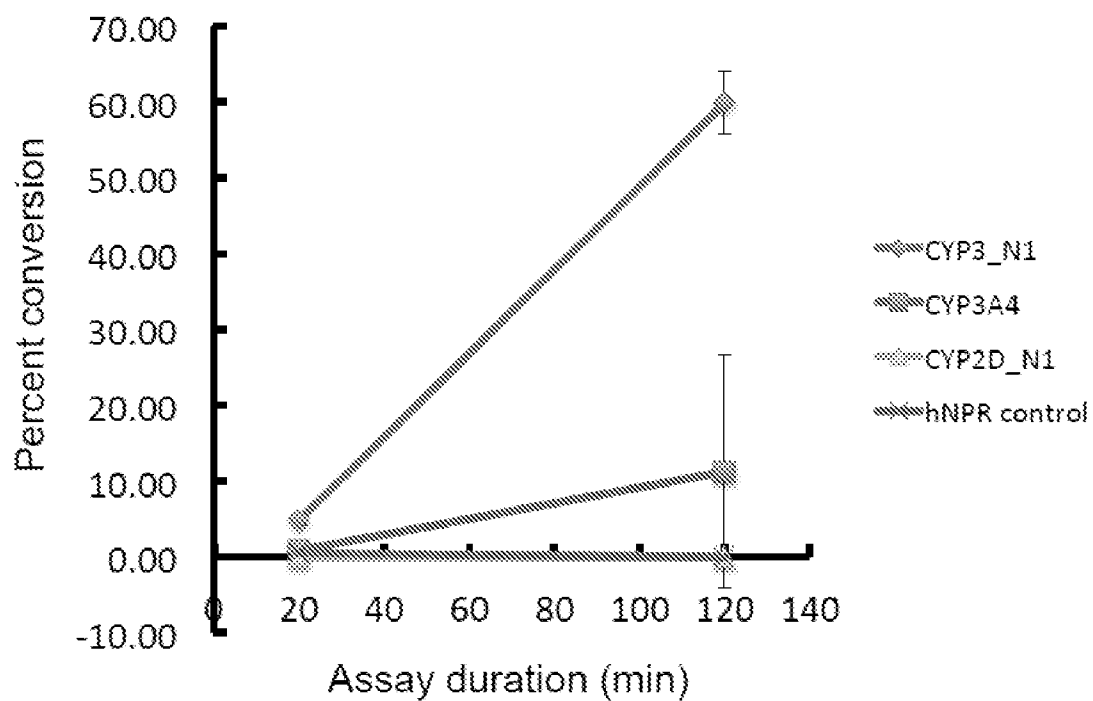


Figure 36

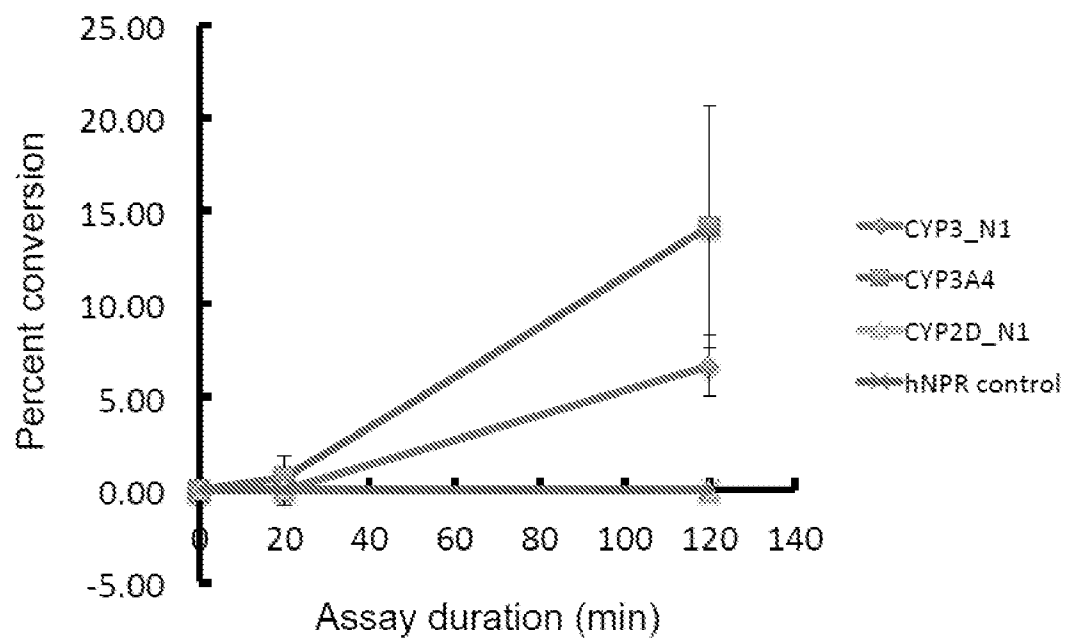


Figure 37

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2015/050847

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/53 (2006.01) C12Q 1/00 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Genome Quest Live: Subsequence search SEQ ID NO. 1, 2-41, 321, 321, 544-578 at 80% and 70%

STN Registry Keyword search, Databases MEDLINE, CAPLUS, WPIDS, BIOSIS, BIOTECHABS, Keywords: P450 OR CYP, ENGINEER? OR DESIGN?, EVOLUTION? OR EVOLVE? OR ANCESTRAL? OR ANCESTOR? OR CONSENS?

Applicant Inventor Search: INTESS, Auspat, Eponline, Patentscope: THE UNIVERSITY OF QUEENSLAND and GILLAM, Elizabeth; GUMULYA, Yosephine

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	



Further documents are listed in the continuation of Box C



See patent family annex

* "A"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	earlier application or patent but published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
11 March 2016Date of mailing of the international search report
11 March 2016

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
Email address: pct@ipaustalia.gov.au

Authorised officer

Ricky Fung
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. 0262223648

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2015/050847

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:
 - a. (means)
☐ on paper
☒ in electronic form
 - b. (time)
☐ in the international application as filed
☒ together with the international application in electronic form
☐ subsequently to this Authority for the purposes of search
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/AU2015/050847
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2012/038508 A1 (AUTODISPLAY BIOTECH GMBH) 29 March 2012 See abstract SEQ ID NO 6, figure 14, page 30 line 19-22, page 19 line 30- page 20 line 5	1, 2 , 7. 22
X	GenBank: ADO60898.1, 2 May 2011 See Whole Document	4, 5, 7
X	UniProtKB - H3BGP9 (H3BGP9_LATCH)'s NADPH--cytochrome P450 reductase See Whole Document	8, 9, 11
A	GILLIAM E. M. J. "Engineering Cytochrome P450 Enzymes" Chemical Research. Toxicology. 2008, Vol 21, pages 220-231 See Whole Document epsically abstract	1-30
A	KABUMOTO H., ET AL "Directed Evolution of the Actinomycete Cytochrome p450 MoxA (CYP105) for Enhanced Activity." Bioscience Biotechnology Biochemistry, 2009, Vol. 73 No. 9, pages 1922-1927. See Whole Document epsically abstract	1-30
A	KUMAR S., et al "Directed Evolution of Mammalian Cytochrome P450 2B1" THE JOURNAL OF BIOLOGICAL CHEMISTRY, 2005, Vol. 280, No. 20, Issue of May 20, pp. 19569-19575, See Whole Document epsically abstract	1-30

Form PCT/ISA/210 (fifth sheet) (July 2009)

INTERNATIONAL SEARCH REPORT Information on patent family members		International application No. PCT/AU2015/050847	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2012/038508 A1	29 March 2012	WO 2012038508 A1 EP 2619306 A1 EP 2619306 B1 US 2014011706 A1	29 Mar 2012 31 Jul 2013 20 May 2015 09 Jan 2014
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(July 2009)			