Whole Genome Analysis Reveals That Conservation and Duplication Gave Rise to the Current V2R Gene Repertoire in Mammals

By
Alexander S. Berry
April, 2012



Allegheny College
Biochemistry Program

Whole Genome Analysis Reveals That Conservation and Duplication Gave Rise to the Current V2R Gene Repertoire in Mammals

By
Alexander S. Berry
April, 2012

A Senior Comprehensive Project in Partial Fulfillment of the Requirements for a Bachelor of Science Degree from Allegheny College

I hereby recognize and pledge to fulfill my responsibilities as defined in the Honor Code and to maintain the integrity of both myself and the College community as a whole.

_____
Pledge

Approval by the Senior Thesis Committee:

_____
Dr. Kristen Webb

_____
Dr. Martin Serra

Table of Contents

Abstract……..1

Introduction……...……1

Methods……10
V2R Gene Identification……..10
Reconstruction of Neighbor-Joining Tree……...13

Results……..14

Discussion……17
Number of Intact V2R Genes Found……...17
Few Pseudogenes Found……18
Domestication May Reduce Number of Intact Genes……19
Duplications in Rodent Lineages……20
Conservation of Specific Genes……..22
Limitations……..23
Future Studies……24
Conclusions……26

Literature Cited……...26

Acknowledgements……..32

Appendix A-Perl script used to trim sequences……..33

Appendix B-Amino acid sequence of intact platypus V2R gene…..…...34

Appendix C-Class C GPCR Database……..34

Appendix D-500 genes used to reconstruct the Neighbor-Joining Tree.64

Abstract

The vomeronasal organ (VNO) is the part of the olfactory system that detects specific semiochemicals known as pheromones, which allow for communication between members of a species, and also affect social and reproductive behaviors. The chemodetection function of the VNO in mammals is facilitated by G protein-coupled receptors (GPCRs) encoded by members of two gene superfamilies: vomeronasal 1 receptor (V1R) and vomeronasal 2 receptor (V2R). Previously, members of the V2R gene family from eight mammalian genomes had been identified and their evolutionary relationships reconstructed. The objective of this study was to use bioinformatics techniques to survey four additional mammalian genomes from the African savannah elephant (*Loxodonta africana*), the European rabbit (*Oryctolagus cuniculus*), the guinea pig (*Cavia porcellus*), and the domestic horse (*Equus caballus*) for V2R genes and to include these genes in evolutionary analyses. These four mammals were chosen to diversify the group of mammals surveyed for V2R genes. Ninety and fifty-two intact genes were identified in *O. cuniculus* and *C. porcellus*, respectively, and these genes appear to have mostly evolved through duplications after the rodent lineages diverged. Two and five intact genes were found in *E. caballus* and *L. africana*, respectively, and appear to be conserved, as they are also present in several other species.

Introduction

Semiochemicals are a broad class of chemicals or chemical compounds that carry messages. The messages transmitted within and between species of mammals are detected by the olfactory system. The mammalian olfactory system governs the sense of smell, which is vital for survival. The reception of olfactory signals allows organisms to

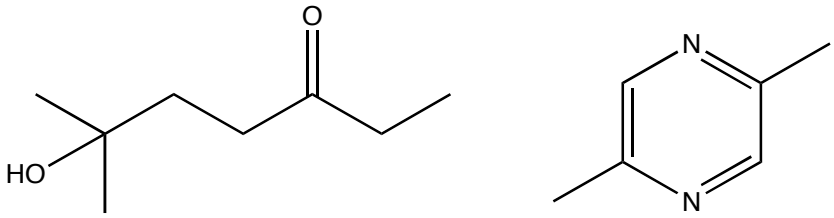
find food, identify mates and offspring, recognize territory, and avoid danger (Niimura and Nei, 2006). The responses evoked by these chemicals and their importance for survival suggest that specific chemosensory receptors assure the proper response to each chemosensory signal (Liberles et al., 2009).

Olfactory receptors (ORs) are the proteins that are responsible for detecting odor molecules. Mammals have a large repertoire of ORs allowing for the detection of a large variety of scents in their environments. In mammals, OR genes comprise a large gene family that includes about 1,000 genes that can be found on almost all chromosomes. Niimura and Nei (2003) used bioinformatics techniques to search the human (*Homo sapiens*) genome for intact OR genes and OR pseudogenes and found that humans have fewer than 400 total OR genes, 52 percent of which are pseudogenes. Niimura and Nei (2005a) used similar techniques to search the mouse (*Mus musculus*) genome and found that mice have greater than 1,000 intact OR genes and 354 pseudogenes. The number of intact OR genes is about 2.7 times greater in the mouse compared to the human, and 75 percent of total mouse OR genes are intact while only 48 percent of human OR genes are intact, the lowest percentage among all mammals surveyed (Niimura and Nei, 2003; Niimura and Nei, 2005a; Niimura and Nei, 2005b; Quignon et al., 2005). This is likely because humans do not rely on olfaction as heavily as other mammals and instead rely more on the visual system (Gilad et al. 2004). The OR gene repertoire has been surveyed using bioinformatics techniques in the rat (*Rattus norvegicus*), dog (*Canis lupus familiaris*), chicken (*Gallus gallus*), frog (*Xenopus tropicalis*), green spotted pufferfish (*Tetraodon nigroviridis*), and zebrafish (*Danio rerio*) in addition to mouse and human (Niimura and Nei, 2003; Niimura and Nei, 2005a; Niimura and Nei, 2005b; Quignon et

al., 2005). The data showed that the OR gene repertoire in tetrapods underwent many gene duplications as well as gene losses in order to adapt to terrestrial life, while changes in fish OR gene repertoire have been small due to less environmental changes (Niimura and Nei, 2006). The OR gene family is susceptible to extensive gene birth and death, culminating in a variable number of intact genes between species (Niimura and Nei, 2006). Niimura et al. (2006) recognized that the number of species currently surveyed is too small to suggest a conclusive trend in OR gene evolution.

The mammalian vomeronasal organ (VNO) is part of the olfactory system and is located on the bottom of the nasal cavity, separate from the main olfactory system that detects odor molecules (Keverne, 1999). The VNO is used to detect specific semiochemicals known as pheromones that evoke social and reproductive behaviors (Halpern, 1987; Keverne, 1999; Young et al., 2005). For example, pheromone detection by the VNO occurs in ungulates such as the horse when they curl their upper lip to touch pheromone molecules to the VNO receptors, an action known as the Flehmen response (Sankar and Archunan, 2004).

Pheromones are highly variable in their chemical structures and as such are normally grouped by function (Tirindelli et al., 2009). An example of this variability is illustrated in Figure 1, which shows the distinct chemical structures of two known pheromones detected by the VNO.



6-Hydroxyl-6-methyl-3- heptanone 2,5-Dimethyl pyrazine

Figure 1. Two semiochemicals detected by the VNO (Tirindelli et al., 2009). 6-Hydroxyl-6-methyl-3- heptanone (left), which evokes an electrical response and calcium influx in VNO neurons is found in male mouse urine and is involved in puberty acceleration. 2,5-Dimethyl pyrazine (right), which affects IP3 production in the VNO, is found in female mouse urine and is involved in the suppression of the estrus cycle (Tirindelli et al., 2009).

The chemodetection function of the VNO in mammals is facilitated by G protein-coupled receptors (GPCRs) from the vomeronasal 1 receptor (V1R) and vomeronasal 2 receptor (V2R) gene superfamilies (Boschat et al., 2002). V1R genes are coexpressed with the G protein $G\alpha_{i2}$ in neurons in the upper VNO neuroepithelium, and V2R genes are coexpressed with the G protein $G\alpha_{0}$ in neurons in a lower region (Dulac and Axel, 1995; Herrada and Dulac, 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997; Pantages and Dulac, 2000). The number of intact genes and pseudogenes in each superfamily varies from nine to over 800 across a diverse set of mammals including human, mouse, rat, dog, cow (*Bos taurus*), opossum (*Monodelphis domestica*), and platypus (*Ornithorhynchus anatinus*) (Shi and Zhang, 2007; Young and Trask, 2007; Grus et al., 2007).

Five classes of GPCRs have been identified in mammals based on amino acid sequence similarity (Kolakowski, 1994; Bockaert and Pin, 1999; Fredriksson et al., 2003; Pin et al., 2005). Rhodopsin-like receptors form class A, secretin-like receptors form class B, metabotropic glutamate-like receptors form class C, and frizzled receptors and adhesion receptors form the last two unnamed classes. V2R genes are members of class C GPCRs. Other class C receptors include metabotropic glutamate receptors, GABA-B

receptors, calcium-sensing receptors, putative pheromone receptors, odorant receptors, and taste receptors. Class C GPCRs are structurally more complex than other classes of GPCRs due to their three structural domains and the association of class C GPCRs into constitutive dimers (Livnah et al., 1999; He et al., 2001; Pin et al., 2005) (Fig. 2).

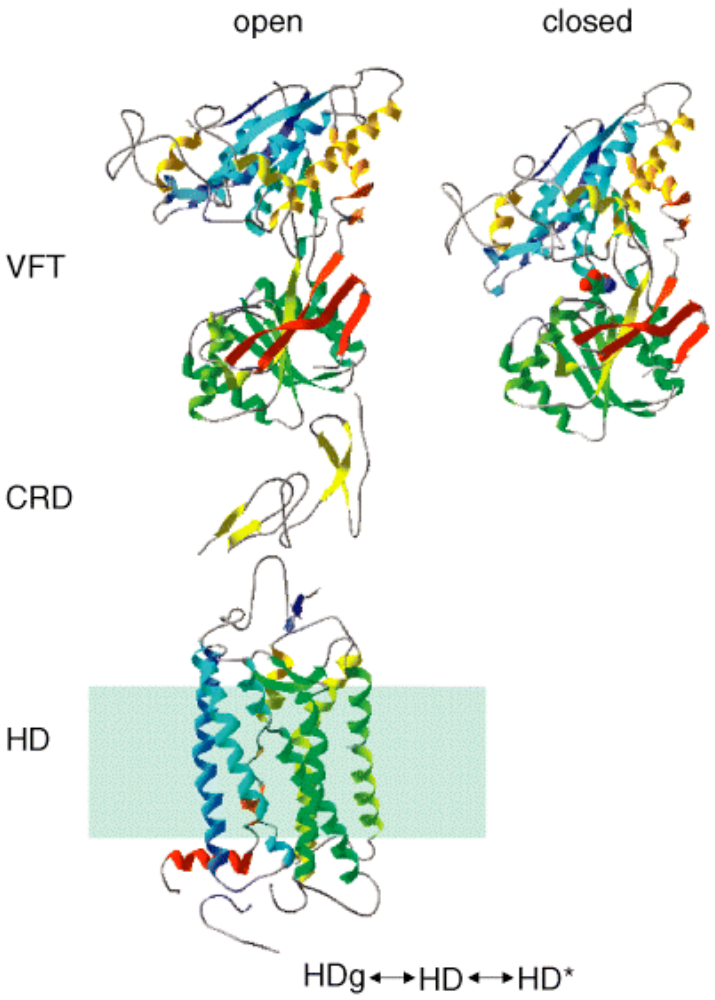


Figure 2. The three main domains of class C GPCRs and their conformation states (Pin et al., 2005). Class C GPCRs are made up from three domains, the Venus flytrap domain (VFT), the cystein-rich domain (CRD), and the heptahelical domain (HD.) Agonists and competitive antagonists bind to the VFT, while the CRD connects the VFT to the rhdopsin-like HD (Pin et al., 2005).

Young et al. (2005) surveyed complete mammalian genomes for intact and pseudo V1R genes using bioinformatics tools and analyzed the evolution of the V1R gene family. A large number of intact V1R genes were found in the mouse and rat

genomes compared to the genomes of dogs, humans, and chimpanzees. Pseudogenes were also identified in the mouse, rat, chimpanzee (*Pan troglodytes*), human, and dog genomes (Young et al., 2005) (Table 1). It has been proposed that the large number of V1R genes in rodents is due to the conservation of ancestral genes, as well as the development of new genes after the lineages of rodents and primates diverged about 100 million years ago (Li et al., 1990; Young et al., 2005).

Young and Trask (2007) used similar techniques to study the V2R gene family in rodents, primates, and canines, and found a similar trend, however, the V2R superfamily has undergone more extensive decimation than V1R. A fairly large number of functional genes in the V2R superfamily exist in the mouse and rat genomes (although fewer than the number of genes in the V1R family), but there are no intact V2R genes in the genomes of the cow, dog, human, chimpanzee, or macaque (*Macaca mulatta*) (Shi and Zhang, 2007; Young and Trask, 2007) (Table 1). The expanded pheromone receptor gene repertoire observed in both gene families in rodents allows for the detection of a wider variety of pheromones that other mammals are unable to detect (Young and Trask, 2007). The decimation of the V2R gene family in the cow, dog, and primates suggests that the pheromones detected by V2R are less important for the survival of higher mammals than those detected by V1R (Shi and Zhang, 2007; Young and Trask, 2007).

Table 1. Number of Intact and Pseudogenes of the V1R and V2R superfamilies in Eleven Species

Organism	Intact V1R Genes	V1R Pseudogenes	Intact V2R Genes	V2R Pseudogenes
Mouse	165[a]	165[a]	123[b]	160[b]
Rat	106[a]	110[a]	87[b]	157[b]
Dog	8[a]	54[a]	0[b]	10[b]
Human	2[a]	115[a]	0[b]	20[b]
Chimpanzee	0[a]	102[a]	0[b]	18[b]
Cow	40[c]	45[c]	0[b]	16[b]
Opossum	98[c]	30[c]	86[b]	79[b]
Platypus	270[d]	579[d]	15[d]	112[d]
Chicken	0[c]	0[c]	0[c]	0[c]
Frog	21[c]	2[c]	249[e]	448[d]
Zebrafish	2[c]	0[c]	44[e]	8[f]

References. ([a]Young et al., 2005; [b]Young and Trask, 2007; [c]Shi and Zhang, 2007; [d]Grus et al., 2007; [e]Niimura et al., 2006; [f]Gloriam et al., 2005)

The ratio of V1R to V2R genes has been shown to be about 50 times greater in terrestrial vertebrates than in aquatic vertebrates (Shi and Zhang, 2007) (Table 1). As a result, it was proposed that V1Rs bind small airborne chemicals while V2Rs bind water-soluble molecules (Shi and Zhang, 2007). Despite this claim, Grus et al. (2007) found that the semiaquatic platypus had the largest V1R repertoire among mammals, and a small V2R repertoire (Table 1).

It is possible that the deterioration of these pheromone receptor genes seen in certain mammals, such as dogs and primates, is due to another set of genes adopting the same function, eliminating the need for these two specific gene superfamilies. Liberles et al. (2009) has reported a possible third family of mouse VNO receptors that is part of the formyl peptide receptor (FPR) family, although the role of these receptors is unknown. Another hypothesis is that there is less reliance on pheromone reception in higher mammals such as primates, who rely more on the visual system than olfactory cues for survival and reproduction (Young et al., 2005). In order to accept or reject either

hypothesis, however, V1R and V2R gene evolution must be studied across a more diverse group of species than is currently documented.

Young et al. (2005), Young and Trask (2007), Shi and Zhang (2007), Grus et al. (2007), Niimura et al. (2006), and Gloriam et al. (2005) completed the first step in studying the evolution of V1R and V2R genes by identifying all V1R and V2R genes in 11 diverse vertebrates (Table 1). These data can be used to find homologous genes in other vertebrates based on amino acid sequence similarity. Since the three domains of class C GPCRs form a complex protein, V2R amino acid sequences must be conserved across vertebrates in order to maintain its function.

Here we used bioinformatics techniques similar to Young et al. (2005) and Young and Trask (2007) to identify intact V2R genes and pseudogenes from genome sequences based on sequence similarity to those previously identified for four previously unsurveyed mammals: the African savannah elephant (*Loxodonta africana*), the European rabbit (*Oryctolagus cuniculus*), the guinea pig (*Cavia porcellus*), and the domestic horse (*Equus caballus*). We then analyzed the evolution of the V2R gene repertoire of these four mammals as compared to the V2R gene repertoire of the other eight mammals previously studied by Young and Trask (2007). The four mammals used in this study were chosen based on genome sequence availability and because they represent a diverse set of mammals. The elephant was chosen to quantify the number of pheromone receptor genes of a terrestrial mammal that is closely related to aquatic mammals. The domestic horse was chosen because although the Flehmen response has been observed in horses, the genes of the vomeronasal organ remain unstudied. Studying the rabbit and guinea pig allowed for the comparison of the V2R gene repertoire between

mammals closely related to the previously studied rodents. The diversity of this group of mammals allowed us to further the understanding of where the loss of the V2R gene family originated, and to observe cases of gene births and deaths in this family across mammalian evolution (Fig. 3).

Rasmussen et al. (1982) extracted components of the urine of estrous cow elephants and showed that they evoke pheromone responses in adult elephants. Johnson and Rasmussen (2002) showed that 6-week-old Asian elephants have a developed vomeronasal organ, which supports the view that the elephant has a functional VNO. Due to the functionality of the VNO, we hypothesized that the elephant would have several intact V2R genes. Taniguchi and Mikami (1985) showed that the horse VNO is similar in structure to the VNO of cattle. Young and Trask (2007) found no intact V2R genes in the cow genome and Shi and Zhang (2007) found 40 V1R genes in the cow genome. Based on these data, we hypothesized that the horse would have very few, if any V2R genes, and that the pheromone receptor genes would mostly be part of the V1R gene repertoire. We expected the guinea pig and the European rabbit to have a similar number of intact V2R genes to the mouse and rat because they are closely related species that live in similar environments.

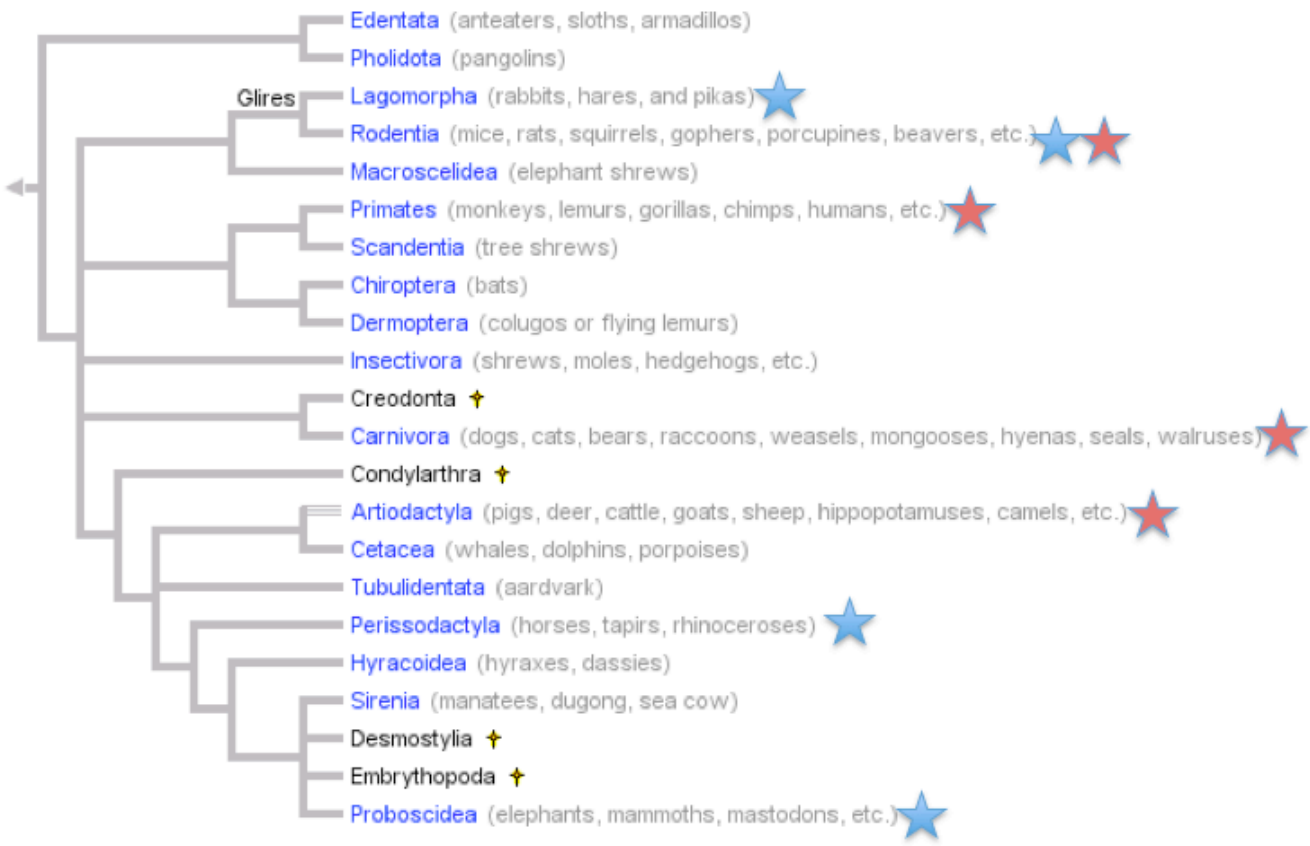


Figure 3. Phylogeny of Placental Mammals (http://tolweb.org/Eutheria/15997). Mammals whose genomes had been previously surveyed for V2R genes are labeled with red stars. Mammals whose genomes are surveyed in this project are labeled with blue stars.

Materials and Methods

V2R Gene Identification

V2R genes were identified using a modified version of the methods used by Young et al. (2005) and Young and Trask (2007) to collect V1R and V2R genes. A file containing the amino acid sequences of six diverse, intact V2R genes was created including three genes from mouse (*Mus musculus*), two from rat (*Rattus norvegicus*), and one from platypus (*Ornithorhynchus anatinus*) (Table 2; Appendix B). Using the tblastn function of the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/), each of the six amino acid sequences of the known V2R

genes was aligned to the translated nucleotide sequences of the complete genomes of the African savannah elephant (*Loxodonta africana*) (Loxafr3.0), the European rabbit (*Oryctolagus cunuculis*) (OryCun2.0), the guinea pig (*Cavia porcellus*) (Cavpor3.0), and the domestic horse (*Equus caballus*) (EquCab2.0) deposited in the National Center for Biotechnology Information (NCBI) genome database (http://www.ncbi.nlm.nih.gov/). The Genbank accession number, start, and stop position were recorded for all hits within a genome with an expect value (E-value) of 10^{-5} or lower for all six genes in the data set.

Table 2. Accession Numbers corresponding to Amino Acid Sequences of Five Intact V2R Genes from the Mouse and Rat.

Species	Accession Number
Mouse	AAC53404.1
Mouse	AAC53412.1
Mouse	AAC53408.1
Rat	AAC08417.1
Rat	AAC08415.1

Each accession number that was recorded represented a larger nucleotide sequence: an entire chromosome or a supercontig. The start and stop coordinates for each hit correspond to a potential V2R gene that must be trimmed from the larger sequence. Each file that included the nucleotide sequence of a potential V2R gene was downloaded from the NCBI database. Coordinates were modified by subtracting 1000 bases (1kb) from the start position and adding 1kb to the stop position of each potential V2R gene such that a 1kb flanking region was located on either side of the potential V2R gene sequence. Using a perl script (Appendix A), all potential V2R regions were trimmed from the large nucleotide sequences using the modified coordinates. When the modified start and stop positions of two neighboring sequences overlapped by less than 5kb, the sequences were combined.

All trimmed potential V2R sequences were classified as V2R genes or non-V2R genes based on amino acid sequence similarity to 44 V2Rs or 144 other class C GPCRs including 46 metabotropic glutamate receptors (MGR), 29 taste receptors, 23 GABA-B receptors, 20 calcium sensing receptors, 11 retinoic acid-induced receptors, 11 putative pheromone receptors, and 4 odorant receptors as assessed using the blastx function within BLAST (Appendix C). Those sequences that matched one of the class C GPCRs with an E-value of 10^{-80} or less were classified as the GPCR to which it matched most closely. The 10^{-80} E-value cutoff was based on the lowest sequence similarity between known V2R genes as determined by an alignment of the amino acid sequences of 44 V2R genes to each other using BLAST. The remaining sequences were then classified as intact or pseudogenes based on the absence or presence, respectively, of stop codons disrupting the reading frame. The coordinates of the start and stop codons of the newly classified genes were recorded, as well as if the sequence was in the same orientation as the previously identified V2R sequence (+/+) or the reverse orientation (+/-).

The above mentioned perl script was again used to trim the intact and pseudo V2R genes from the recorded start and stop positions to include only the gene without flanking regions. Those sequences in the reverse orientation were combined into a multifasta file, and the reverse complement of each sequence was generated using BCM Search Launcher (http://searchlauncher.bcm.tmc.edu/seq-util/Options/revcomp.html) in order to ensure all sequences were in the same orientation to allow for a successful multiple alignment. V2R sequences that were +/+, and the reverse complements of the sequences that were +/- were combined into a multifasta file, and translated to amino acid sequences using Virtual Ribosome (Wernersson, R.,2006).

Reconstruction of Neighbor-Joining Trees

The amino acid sequences of the intact and pseudo V2R genes from the mouse, rat, opossum, dog, cow, human, chimpanzee, and macaque genomes included in the phylogenetic analyses performed by Young and Trask (2007) were added to the file containing the newly translated V2R amino acid sequences from the horse, guinea pig, rabbit, and elephant. Partial sequences, defined as those sequences missing greater than 40% of the sequence data relative to the length of the longest sequence in the dataset, were excluded from the alignment in order to ensure all sequences in the alignment had sufficient overlap to reduce spurious groupings due to long-branch attraction as a result of missing data (Young and Trask, 2007, Wiens and Morrill, 2011). An alignment was created with MUSCLE (Edgar, R.C., 2004) using Mega_5.0 (Tamura et al., 2011) including amino acid sequences of all 491 trimmed V2R genes collected in this study and those collected by Young and Trask (2007), and including 4 taste receptors and 5 calcium-sensing receptors. Alignment parameters included: no presets, a gap opening of -12, gap extension of -1, hydrophobicity multiplier of 1, max memory of 1745 megabytes, 16 max iterations, the clustering method was UPGMB, the max diagonal length was 24, and standard genetic code when using cDNA. A neighbor-joining tree with bootstrap values based on 500 replicates was reconstructed from the alignment of the 500 sequences using Mega_5.0 (Tamura et al., 2011). Parameters for the neighbor-joining reconstruction included: phylogenetic reconstruction using all selected taxa, bootstrap values based on 500 replications, amino acids were the substitution type, Poisson model, uniform rates among sites, homogeneous pattern among lineages, and pairwise deletion of gaps/missing data. Taste receptor sequence M_AAK394381_TR was

used to root the tree. Fig Tree 1.3.1 (Rambaut and Drummond, 2010) and GIMP 2.6 (http://www.gimp.org/) were used to edit the tree.

Results

After initially searching each genome for potential V2R gene-containing sequences using BLAST, 26 sequences were downloaded for the elephant, 230 for the rabbit, 232 for the guinea pig, and 45 for the horse. A subsequence alignment revealed the identity of each sequence. The number of each non-V2R class C GPCRs identified in each genome is shown in Table 3. The number of intact V2Rs and V2R pseudogenes identified in each genome assembly, as well as eight from Young and Trask (2007) is shown in Table 4.

The number of intact V2R genes in the rabbit was similar to the number previously reported in the mouse, rat, and opossum, while slightly fewer intact genes were found in the guinea pig. The number of intact genes in the elephant and horse was very low, but greater than the number of intact V2Rs present in the cow, dog, macaque, human, and chimpanzee, in which none were found. The number of V2R pseudogenes found in the rabbit was much lower than in the mouse, rat, and opossum. The number of pseudogenes found in the guinea pig, elephant, and horse was lower than the number found in any previously studied species (Table 4). The sequences that were not classified as a class C GPCR did not align to known GPCR sequences with high correlation and were excluded from further analyses.

Table 3. Number of non-V2R class C GPCRs identified in the genomes of four newly surveyed organisms. The number of metabotropic glutamate receptors (MG), calcium-sensing receptors, and taste receptors found from the sets of potential V2R sequences from each genome were determined for the four organisms shown.

Species	MG	Calcium-Sensing	Taste
Rabbit	9	1	2
Guinea Pig	8	1	3
Elephant	8	2	2
Horse	7	2	2

Table 4. Number of V2R-like genes found in the genomes of twelve organisms. The number of intact V2R genes, pseudogenes, and their sums were calculated for the twelve organisms shown. Data collected by Young and Trask (2007) are denoted with an asterisk (*).

Species	Assembly	Intact Gene	Pseudogene	Total
Rabbit	January 2010	90	17	107
Guinea Pig	March 2008	52	8	60
Elephant	July 2009	5	1	6
Horse	July 2008	2	2	4
Opossum*	January 2006	86	79	165
Rat*	November 2004	79	142	221
Mouse*	February 2006	121	158	279
Cow*	March 2005	0	16	16
Dog*	March 2005	0	9	9
Macaque*	January 2006	0	11	11
Human*	March 2006	0	20	20
Chimpanzee*	March 2006	0	17	17

A neighbor-joining tree was reconstructed and bootstrap values based on 500 replicates are labeled on basal nodes (Fig. 4). Amino acid sequences of all genes included in Figure 4 are shown in Appendix D. All taste receptors sequences resolved basal to the V2R sequences, while the calcium-sensing receptors formed a monophyletic clade near the base of the intact V2R gene clade. High bootstrap support is seen at basal nodes. Generally, sequences did not group by species, however, clusters of highly similar sequences from the same species are seen. Most species have V2R genes distributed throughout the tree (Fig. 4).

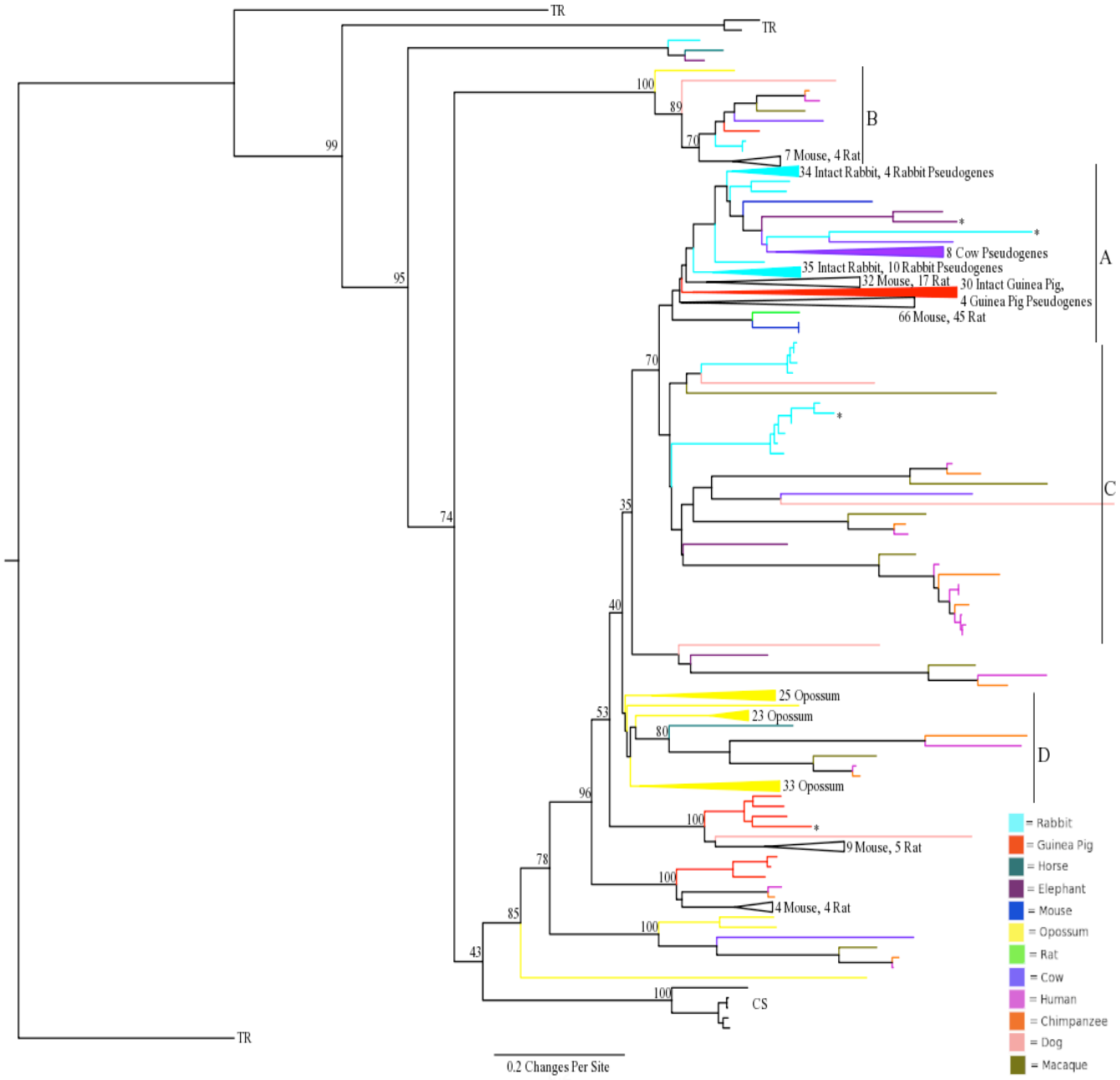


Figure 4. Rooted Neighbor-Joining tree including V2R genes from twelve organisms. 491 V2R-like sequences are included: 4 intact and 1 pseudogene from the elephant, 45 intact and 6 pseudogenes from the guinea pig, 2 intact from the horse, 84 intact and 16 pseudogenes from the rabbit, 86 intact from the opossum, 77 intact from the rat, 121 intact from the mouse, 12 pseudogenes from the cow, 5 pseudogenes from the dog, 8 pseudogenes from the macaque, 14 pseudogenes from the human, and 10 pseudogenes from the chimpanzee. Four taste receptors from the mouse and rat are labeled on the tree as TR. Five calcium sensing receptors from *Takifugu rubripes*, mouse, rat, cow, and human that share an evolutionary origin with V2R genes, but have different function, were also included, and are labeled on the tree as CS. Bootstrap values are labeled on basal nodes and are based on 500 replicates. Taste receptor gene sequence, M_AAK394381_TR, is used to root the tree. Pseudogenes from the rabbit, guinea pig and elephant are notated with asterisks (*) to distinguish them from intact genes from the same species. Scale bar indicates changes per site. A, B, C, and D refer to clades referenced in the discussion section.

Discussion

Number of Intact V2R Genes Found

An objective of this study was to identify the intact and pseudo V2R genes in the genomes of the rabbit, guinea pig, elephant, and horse. We expected that the rabbit and guinea pig would have the most intact V2R genes, that the elephant would have several, and that the horse would have very few. The number of intact V2R genes found in the rabbit and guinea pig supported the hypotheses. The rabbit had a number of V2R genes between those of the mouse and rat, while the guinea pig had less, but still a number comparable to those of the other rodents surveyed. The large number of intact V2R genes in the guinea pig and rabbit are consistent with observations of a highly functional VNO in rodents and rabbits (Hudson and Distel, 1986; Leypold et al., 2002; Stowers et al., 2002; Young et al., 2005).

The numbers of V2R genes in the horse and elephant were consistent with the hypotheses. Five intact V2R genes were found in the elephant, and two in the horse. These numbers are significantly less than the numbers of V2R genes found in the rodents and rabbit, but the existence of any intact V2R genes at all is more than the dog, cow, and primates. The horse VNO is similar in structure to the cow VNO, which may indicate that the majority of the pheromone receptor genes in the horse are located in the V1R gene family, similar to the cow, which had no intact V2R genes but 40 intact V1R genes (Taniguchi and Mikami, 1985; Shi and Zhang, 2007; Young and Trask, 2007). The number of intact V2R genes in the elephant genome is consistent with the observation of a functional VNO in the elephant (Rasmussen et al., 1982). Since the elephant does have a functional VNO, yet only five intact V2R genes were found in the genome, it is

possible that the majority of pheromone receptor genes in the elephant are located in the V1R gene family. This would be surprising because of the small number of V1R genes and the large number of V2R genes found in the closely related aquatic organisms that are thought to use water-soluble molecules (which are received by V2R proteins) to send signals (Gloriam et al., 2005; Niimura et al., 2006; Grus et al., 2007; Shi and Zhang, 2007).

Few Pseudogenes Found

Fewer V2R pseudogenes were identified in this study compared to previous studies performed by Young and Trask (2007) and Shi and Zhang (2007), which created large discrepancies in total V2R gene counts. This is likely due to differences in methods, or differences in completeness of genome assemblies, rather than differences in the actual genome composition.

Most pseudogenes contain genetic lesions that prevent translation of the sequence into a functional protein (Vanin, 1985). These lesions can be caused by single base substitutions that cause termination codons in the reading frame or insertions and deletions of bases that shift the reading frame, causing premature termination (Vanin, 1985). Examples of these types of pseudogenes have been recorded in many different gene families in many mammalian organisms, including the mouse α globin and the cytochrome *c* genes (Nishioka et al., 1980; Vanin et al., 1980; Limbach and Wu, 1985; Vanin, 1985). There is no indication that the four species surveyed in this study are less susceptible to mutations that cause pseudogenes, therefore it is unlikely that they do have fewer V2R pseudogenes than previously studied species (Shi and Zhang, 2007; Young and Trask, 2007).

Domestication May Reduce Number of Intact Genes

Normally, changes, or mutations in genomes arise based on the relative reproductive success of individuals with different genotypes. These normal evolutionary changes, however, may be interrupted in certain species, including the guinea pig and the domestic horse, through the process of domestication. Through this process, humans select for or against certain genes that may go against the natural tendencies of these genomes. For example, wild strains of mice display diurnal cycles of melatonin synthesis, allowing them to monitor daylight and adjust reproductive behaviors accordingly (Tamarkin et al., 1985; Olson, 1999). Due to human interaction, however, the reproductive cycles of certain laboratory strains are not linked to melatonin synthesis and have accumulated recessive mutations in two genes, eliminating melatonin synthesis (Ebihara et al., 1986; Goto et al., 1994; Olson, 1999). These mutated genes have DNA sequences similar to their functional counterparts, and are thus recognizable, but they would be pseudogenes due to their lack of function (Olson, 1999).

In the wild, V2R genes are important for survival of an individual. An individual lacking one or more functional V2R genes will likely have a reduced chance of reproductive success, reducing the likelihood of a debilitating mutation in a V2R gene being incorporated into the genome. With humans controlling reproduction in domesticated mammals, however, V2R genes would not be selected for in a population of these mammals, and many intact genes may accumulate mutations that render them nonfunctional and thus, pseudogenes. The low number of V2R genes found in the guinea pig as compared to the other rodents could indicate an effect of domestication, but this seems unlikely because of the low number of pseudogenes found in the guinea pig. It is

more likely that this organism never required as large a number of V2R genes for survival as compared to the mouse and rat. The number of intact V2R genes in the domestic horse may also be a result of domestication, however, since only two pseudogenes were found in its genome it is likely that this species never acquired a large V2R gene repertoire.

Duplications in Rodent Lineages

In species including the mouse, rat, and opossum, V2R genes have been found on almost all chromosomes, and are located in clusters, or in close proximity to one another, within these chromosomes (Shi and Zhang, 2007; Young and Trask, 2007). The rabbit and guinea pig genomes are not assembled into chromosomes at this time, which makes it impossible to determine the spread of V2R genes across the chromosomes, however, the number of V2R genes that appear in the same supercontig within the genomes of these species was determined. Although these supercontigs are smaller than whole chromosomes, many V2R genes within the rabbit genome are located in clusters including up to 12 genes spanning 500,000 base pairs, which is consistent with previous observations of clusters of rodent V2R genes (Yang et al., 2005; Young and Trask, 2007). Similarly, V2R genes in the guinea pig appear to be arranged in clusters, some including up to 19 V2R genes spanning two million base pairs. These numbers are underestimates because if all chromosomes were assembled, contigs that contain a large number of genes may, for example, be adjacent to one another, and therefore the number of V2R genes in that cluster would be doubled.

Due to variable length of contigs, and the variable distance between sequences on these contigs, it is impossible to assess the extent of these clusters. A high percentage of

eukaryotic genomes contains introns, therefore, although the length of genome spanned by these clusters of V2R genes is apparently very large, much of this region may contain non-coding information, making the effective span of these genes smaller (Venter et al., 2001).

The large number of clusters in these genomes is likely due to duplications of genes after the Glires (referring to the clade that includes both lagomorphs and rodents) lineages diverged (Fig. 1). A single ancestral gene could have duplicated multiple times, and the paralogs acquired mutations that allowed for the detection of a wider variety of pheromones. Orkin (1978) used restriction endonuclease mapping to show that duplication of the human α globin gene lead to multiple α globin genes in close proximity to one another. Genes in close proximity on the chromosome also allow for easier gene regulation because genes that are adjacent to one another can be transcribed together. For example, Boutanaev et al. (2002) identified 1,661 testes-specific genes in *Drosophila*, one third of which were located in clusters, suggesting that the clustering of genes allows for transcriptional co-regulation in higher eukaryotes.

Clade A (Fig. 4) shows a duplication event on a large scale that includes genes from the mouse, rat, guinea pig, and rabbit. Most of the V2R genes located in each of these genomes evolved from one ancestral gene possessed by the common ancestor of all members of clade A. This observation is consistent with that of Young and Trask (2007) that included only the mouse and rat in this event. The addition of another rodent and a close relative of rodents provided by this study further supports the observation that after the rodent lineages diverged, duplication events within this clade gave rise to the majority of the V2R genes located in each of these four genomes.

The relatively low number of V2R genes in the elephant and horse genomes compared to those of the rodents is partially due to the lack of duplication events within the V2R gene repertoire of these organisms after the divergence of the different species.

Conservation of Specific Genes

While most mammals surveyed, both here and previously (excluding the rodents, rabbit, or opossum) have very few V2R genes in their genomes, it appears that these few genes have been conserved across evolutionary time (Shi and Zhang, 2007; Young and Trask, 2007). Clade B (Fig. 4) shows an ancestral V2R gene that was conserved in ten of the twelve mammals surveyed, excluding only the elephant and the horse. Each species has a gene that is a recent descendant of this ancestor. The amino acid compositions of each of the 20 sequences located within this clade are very similar, and likely make up proteins with the same or similar functions. Proper function of this protein must be vital for the survival of most mammals. Neuhaus et al. (2005) identified an odorant receptor in the fruit fly (*Drosophila melanogaster*), DOR83b, that is unusually highly conserved among insects, similar to the way in which the ancestral V2R gene of clade B is conserved among mammals. DOR83b forms complexes with other odorant-receptor proteins in a way that enhances functionality of that odorant receptor (Neuhaus et al., 2005). Young and Trask (2007) suggested that the V2R gene in clade B may have a similar function, binding to and increasing the functionality of other V2R proteins.

Clade C (Fig. 4) shows another example of conservation of an ancestral gene by the primates, dog, cow, elephant, and rabbit. This conservation event spans a wide variety of mammals, and the protein that these sequences code for must detect a pheromone that is present in the diverse environments of all of these mammals. Most of the large

opossum V2R gene repertoire descended from the ancestor located at clade D (Fig. 4), when this gene duplicated multiple times. In addition to the opossum, four other mammals conserved this ancestral gene, including the primates and the horse. It has been determined that several V2R genes are involved in the detection of chemicals present in the urine of various species, therefore it is possible that these conserved genes facilitate detection of these types of chemicals, since it is likely that many mammals would conserve this function (Tirindelli et al., 2009).

Limitations

The methods presented in this study have several inherent limitations. The genomes surveyed were not yet assembled into complete chromosomes at the time of this study, and were instead broken into many contigs of variable length. It is possible that a number of V2R genes spanned two contigs and thus went unrecognized as a V2R gene because the region was not of sufficient length to provide itself with a probable match to other known V2R genes. Although it is unlikely that this would account for a large number of unrecognized V2R genes, the numbers presented in this study should be considered an underestimate of the total number of V2R genes within each genome.

In this study, pseudogenes were classified by the presence of a stop codon that interrupted the reading frame of the DNA sequence. This would come about by a nonsense mutation: a mutation that changes a codon that normally codes for an amino acid to a stop codon, or a frameshift mutation: an insertion or deletion that shifts the reading frame, and causes a premature stop codon to be translated. These mutations almost always render the protein non-functional, however, these are not the only mutations that can cause pseudogenes. A mutation that changes the identity of an amino

acid by altering a base, called a missense mutation, or a frameshift mutation that changes the identities of several integral amino acids, but does not cause a premature stop codon, could also cause a pseudogene if the affected region was vital for proper functioning of the protein, such as in a substrate binding site. These mutations would have gone undetected by the approach taken in this study, and would have caused the gene to be misclassified as intact. Although it is less likely to occur than a nonsense mutation, the number of pseudogenes presented should be considered an underestimate, while the number of intact genes should be considered an overestimate. The number of pseudogenes presented in this study should be further considered an underestimate because V2R genes that accumulated many mutations throughout evolutionary time may no longer be recognizable as V2R genes due to low sequence similarity compared to intact V2R genes.

Future Studies

It is difficult to make conclusions about the fate of the V2R gene repertoire with only twelve mammalian species studied. Future studies expanding on the number of species surveyed for V2R genes would be useful in order to observe a more conclusive trend of V2R gene birth and death. Although a wide variety of mammals have now been surveyed, it is important to continue to fill in the gaps between orders of mammals. Young and Trask (2007) found a large number of intact V2R genes within the opossum genome. The only way to identify exceptions to the apparent decimation of the V2R gene repertoire is to survey all orders of mammals, regardless of how likely or unlikely they are to contain intact V2R genes, because even the identification of V2R pseudogenes tells a lot about the fate of this gene family.

The V2R gene repertoire of an organism only represents half of the potential chemodetection of pheromones. V1R genes are responsible for the detection of pheromones separate from those detected by V2R genes. The number of V2R genes present in a genome does not necessarily correlate to the number of V1R genes present in the genome, as seen in the cow, which has 40 intact V1R genes and no intact V2R genes, therefore V1R genes should be surveyed to the same extent as V2R genes (Shi and Zhang, 2007; Young and Trask, 2007). This will allow for conclusions to be made about the relative importance of the two gene families in different species, and the necessity of pheromone reception overall in these species.

The V2R genes that seem to be conserved across species of mammals should be studied to determine which semiochemical the corresponding receptors detect. This can be done by performing a targeted deletion of the gene of interest and determining which pheromone the organism is no longer able to detect. It is likely that these genes were conserved because the chemicals they detect are still found in the environment and used by many species for survival.

Finally, it should be determined which chromosomes contain V2R genes in each species, which could not be determined in this study due to the incomplete assembly of the genomes. This will allow for a more effective survey of clusters of V2R genes, and a determination of the extent that V2R genes have proliferated throughout the genomes of these organisms. Previous research has shown that V2R genes are located on most chromosomes and found in clusters within chromosomes, suggesting multiple duplication events (Shi and Zhang, 2007; Young and Trask, 2007). This finding should be confirmed or rejected in the four species surveyed in this study.

Conclusions

All of the intact V2R genes found in the elephant and horse are part of larger conservation events among diverse sets of mammals, which suggests that the semiochemicals detected by these proteins are encountered by various mammals. Although a large number of V2R genes were found in the genomes of the Glires, most of the genes in the V2R gene repertoire of these species were contributed to by duplication events after the divergence from one common ancestral gene. The large number of V2R genes in these organisms in consistent with the high functionality of their vomeronasal organ and indicates the importance of pheromone detection in this group.

Literature Cited

Bockaert, J., Pin, J.P. (1999). Molecular tinkering of G-protein coupled receptors: an evolutionary success. *The EMBO Journal*, 18: 1723–1729.

Boschat, C., Pélofi, C., Randin, O., Roppolo, D., Lüscher, C., Broillet, M. C., et al. (2002). Pheromone detection mediated by a V1r vomeronasal receptor. *Nature Neuroscience,* 5(12): 1261-1262.

Boutanaev, A.M., Kalmykova, A.I., Shevelyov, Y.Y., Nurminsky, D.I. (2002). Large clusters of co-expressed genes in the *Drosophila* genome. *Nature*, 420: 666-669.

Del Punta, K., Leinders-Zufall, T., Rodriguez, I., Jukam, D., Wysocki, C. J., Ogawa, S., et al. (2002). Deficient pheromone responses in mice lacking a cluster of vomeronasal receptor genes. *Nature,* 419(6902): 70-74.

Dulac, C., Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. *Cell,* 83: 195–206.

Ebihara, S., Marks, T., Hudson, D.J., Menaker, M. (1986). Genetic control of melatonin synthesis in the pineal gland of the mouse. *Science*, 231: 491-493.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5): 1792-1797.

Fredriksson, R., Lagerstrom, M.C., Lundin, L.G., Schioth, H.B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology*, 63: 1256–1272.

Gilad, Y., Wiebe, V., Przeworski, M., Lancet, D., Pääbo, S. (2004). Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biol*ogy, 2: E5.

Gloriam, D.E., Bjarnadottir, T.K., Yan, Y.L., Postlethwait, J.H., Schioth, H.B., Fredriksson, R. (2005). The repertoire of trace amine G-protein-coupled receptors: large expansion in zebrafish. *Molecular Phylogenetics and Evolution*, 35:470–482.

Goto, M., Oshima, I., Hasegawa, M., Ebihara, S. (1994). The locus controlling pineal serotonin N-acetyltransferase activity (Nat-2) is located on mouse chromosome 11. *Brain Research Molecular Brain Research*, 21: 249-354.

Grus, W. E., Shi, P., & Zhang, J. (2007). Largest vertebrate vomeronasal type 1 receptor gene repertoire in the semiaquatic platypus. *Molecular Biology and Evolution,* 24(10): 2153.

Halpern, M. (1987). The organisation and function of the vomeronasal system. *Annual Review of Neuroscience*, 10: 325-362.

He, X.L., Chow, D.C., Martick, M.M., Garcia, K.C. (2001). Allosteric activation of a

spring-loaded natriuretic peptide receptor dimer by hormone. *Science*, 293: 1657–1662.

Herrada, G., Dulac, C. (1997). A novel family of putative pheromone receptors in mammals with a topologically organized and sexually dimorphic distribution. *Cell,* 90: 763–773.

Hudson, R., Distel, H. (1986). Pheromonal release of suckling in rabbits does not depend on the vomeronasal organ. *Physiology and Behavior*, 37: 123-128.

Johnson, E., Rasmussen, L. (2002). Morphological Characteristics of the Vomeronasal Organ or the Newborn Asian Elephant (*Elephas maximus*). *The Anatomical Record*, 267: 252-259.

Keverne, E. B. (1999). The Vomeronasal Organ. *Science*, 286: 716–720.

Kolakowski, L.F. (1994). GCRDb: a G-protein-coupled receptor database. *Receptors Channels*, 2: 1-7.

Leypold, B.G., Yu, C.R., Leinders-Zufall, T., Kim, M.M., Zufall, F., Axel, R. (2002). Altered sexual and social behaviors in trp2 mutant mice. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 6376-6381.

Li, W.H., Gouy, M., Sharp, P.M., O'hUigin, C., Yang, Y.W. (1990). Molecular Phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. *Proceedings of the National Academy of Sciences of the United States of America*, 87: 6703-6707.

Liberles, S.D., Horowitz, L.F., Kuang, D., Contos, J.J., Wilson, K.L., Siltberg-Liberles, J., Liberles, D.A., Buck, L.B. (2009). Formyl peptide receptors are candidate chemosensory receptors in the vomeronasal organ. *Proceedings of the National*

Academy of Sciences of the United States of America, 106: 9842–9847.

Limbach, K.J., Wu, R. (1985). Characterization of a mouse somatic cytochrome C gene and three cytohrome C pseudogenes. *Nucleic Acids Research*, 13: 617-630.

Livnah, O., Stura, E.A., Middleton, S.A., Johnson, D.L., Jolliffe, L.K., Wilson, I.A. (1999). Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science*, 283: 987–990.

Matsunami, H., Buck, L.B. (1997). A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell,* 90: 775–784.

Neuhaus, E.M., Gisselmann, G., Zhang, W., Dooley, R., Störtkuhl, K., Hatt, H. (2005). Odorant receptor heterodimerization in the olfactory system of *Drosophila melanogaster*. *Nature Neuroscience*, 8: 15-17.

Niimura, Y., & Nei, M. (2003). Evolution of olfactory receptor genes in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, 100: 12235–12240.

Niimura, Y., Nei, M. (2005a). Comparative evolutionary analysis of olfactory receptor gene clusters between humans and mice. *Gene,* 346: 13–21.

Niimura, Y., Nei, M. (2005b). Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. *Proceedings of the National Academy of Sciences of the United States of America,* 102: 6039–6044.

Niimura, Y., & Nei, M. (2006). Evolutionary dynamics of olfactory and other chemosensory receptor genes in vertebrates. *Journal of Human Genetics*, 51(6): 505-517.

Nishioka, Y., Leder, A., Leder, P. (1980). An unusual alpha globin-like gene that has

cleanly lost both globin intervening sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 77: 2806-2809.

Olson, M.V. (1999). When Less Is More: Gene Loss as an Engine of Evolutionary Change. *American Journal of Human Genetics*, 64: 18-23.

Orkin, S.H. (1978). The duplicated human α globin genes lie close together in cellular DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 75, 12: 5950-5954.

Pantages, E., Dulac, C. (2000). A novel family of candidate pheromone receptors in mammals. *Neuron,* 28: 835–845.

Pin, J.P., Kniazeff, J., Liu, J., Binet, V., Goudet, C., Rondard, P., Prezeau, L. (2005). Allosteric functioning of dimeric class C G-protein-coupled receptors. *The FEBS Journal,* 272(12): 2947-2955.

Quignon, P., Giraud, M., Rimbault, M., Lavigne, P., Tacher, S., Morin, E., Retout, E., Valin, A.S., Lindblad-Toh, K., Nicolas, J., Galibert, F. (2005). The dog and rat olfactory receptor repertoires. *Genome Biology*, 6: R83.

Rambaut, A., Drummond, A. (2010) FigTree v1.3.1. *Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom*. [http://tree.bio.ed.ac.uk/software/figtree/].

Rasmussen, L.E., Schmidt, M.J., Henneous, R., Groves, D., Daves, Jr, G.D. (1982). Asian bull elephants: Flehmen-like responses to extractable components in female elephant estrous urine. *Science*, 217: 159-162.

Ryba, N.J., Tirindelli, R. (1997). A new multigene family of putative pheromone receptors. *Neuron,* 19: 371–379.

Sankar, R., Archunan, G. (2004). Flehmen response in bull: role of vaginal mucus and other body fluids of bovine with special reference to estrus. *Behavioural Processes*, 67: 81–86

Shi, P., & Zhang, J. (2007). Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. *Genome Research,* 17(2): 166.

Stowers, L., Holy, T.E., Meister, M., Dulac, C., Koentges, G. (2002). Loss of sex discrimination and male-male aggression in mice deficient for trp2. *Science*, 295: 1493-1500.

Tamarkin, L., Baird, C.J., Almeida, O.F. (1985). Melatonin: a coordinating signal for mammalian reproduction? *Science*, 227, 4688: 714-720.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Likelihood, Distance, and Parsimony methods. *Molecular Biology and Evolution*.

Taniguchi, K., Mikami, S. (1985). Fine structure of the epithelia of the vomeronasal organ of horse and cattle. *Cell Tissue Research*, 240: 41-48.

Tirindelli, R., Dibattista, M., Pifferi, S., Menini, A. (2009). From pheromones to behavior. *Physiological Reviews,* 89: 921-956.

Vanin, E.F. (1985). Processed Pseudogenes: Characteristics and Evolution. *Annual Review of Genetics*, 19: 253-272.

Vanin, E.F., Goldberg, G.I., Tucker, P.W., Smithies, O. (1980). A mouse α-globin-related pseudogene lacking intervening sequences. *Nature*, 286: 222-226.

Venter, J.C., Adams, M.D., Myers, E.W., et al. (271 co-authors). (2001). The sequence of

the human genome. *Science*, 291: 1304-1351.

Wernersson, R. (2006) Virtual Ribosome - a comprehensive DNA translation tool with support for integration of sequence feature annotation. *Nucleic Acids Research.* 34: W385-W388.

Wiens, J.C., Morrill, M.C. (2011) Missing Data in Phylogenetic Analysis: Reconciling Results from Simulations and Empirical Data. *Systematic Biology,* 60(5): 719-31

Yang, H. et al. (2005) Composition and evolution of the V2r vomeronasal receptor gene repertoire in mice and rats. *Genomics,* 86: 306–315.

Young, J. M., Kambere, M., Trask, B. J., & Lane, R. P. (2005). Divergent V1R repertoires in five species: Amplification in rodents, decimation in primates, and a surprisingly small repertoire in dogs. *Genome Research, 15*(2): 231.

Young, J. M., & Trask, B. J. (2007). V2R gene families degenerated in primates, dog and cow, but expanded in opossum. *Trends in Genetics, 23*(5): 212-215.

Acknowledgements

I would like to thank my advisor, Dr. Kristen Webb, for her direction, patience, and expertise in bioinformatics, throughout the course of this project. I would also like to thank Dr. Martin Serra for his dedication to biochemistry and counsel during this project. I would like to thank Dr. Janet Young for her assistance with the unfamiliar methodology. I would also like to acknowledge Allegheny College and the Class of 1939 for funding this project. Lastly, I wish to thank Megan Silbaugh and Dan Hawkins for their advice during the writing process.

Appendix A - Perl script used to trim sequences

```
#Read in file and store in array
my @FILE = <FASTQFILE>;
close FASTQFILE;

#Remove Fasta header
my @FILE2;

foreach my $line (@FILE) {
   if ($line =~ /^>/) {
      next;
   }
   else {
      push (@FILE2, $line);
   }
}

#From the lines of the AA file, put the AA sequence into a single string
my $FILE2 = join('',@FILE2);

#Remove whitespace
$FILE2 =~ s/\s//g;

#Explode sequence into array where each AA is an element in the array
@FILE2 = split('',$FILE2);

#Initialize counts
my $count_AA = 0;

my $AA;
my $i;

#For loop
for ($i=$start-1; $i<$end; $i++) {
   $AA = $AA . $FILE2[$i];
}

open OUTFILE, ">>$out_name";
print OUTFILE "\n>$out_name\n";
print OUTFILE "$AA";
```

Appendix B - Amino acid sequence of intact platypus V2R gene

>platypus_3_V2R3
GLSSYSYTQIIQGRVPSFHYRYFTKHYQHLMALVFAVEEINKDANLLPNISLGFHL
FNVHFLGPIAAESSMALLSGRAGLVPNYNCEPERWDKLLAVIGGMSTTVSMTSS
WLLGLHKLPQVGYNPCPTWSVLSSSPFHSYTQIIQGRVPSFHYRYFTKHYQHLM
ALVFAVEEINKDANLLPNISLGFHLFNVHFLGPIAAESSMALLSGRAGLVPNYNC
EPERWDKLLGCHMKGYNTFFQGSLLFSLHRKEIPGFRDFERILNPRKYPLDAFLP
KFWQWAFECQYLESFLSMTCDFNDTLENRPLDLSAQSYSVYSTVYAIAHALHQE
IWAEKDMETSGDGPSRIPPPWQLHQFLKNTLSHTGTGEQIHLVDPGWVTTKFDIL
NYQILENGSKGYVKVGEFMPLAPAGQDFAISDEKIMWNKLYSETPRSVCSESCPS
GFRKTPREGKAACCYDCVPCAEGQISNQTDADQCITCLQDEFPNPQKDQCILKA
VVFLSYQEPLGIVLVSLALFFSLVTTLDLGIFIRHRDTPIGKANNRSLSYTLLTALL
LCFLSSLIFVGHPSLATCTLRQTAFGVIFSVAISTLLSKTVTVVLAFRATGPGSRMR
MWLGPRVPSSIVWVCSLVQVGICTVWLGTFPPFPDVDVASETGVIIVQCNEGSAI
AFYCVLGYMGLLALITLTVAFLARKLLDSFNEAKFITFSMLVFCSVWFTFLPTYL
STKGKAMVGVEVFSILASSAGILGCIFGPKVYMILLR

Appendix C - Class C GPCR Database (from Young et al., 2007)

>Q75JT4_DICDI_GABA_B
D--
NKKLLKGDSISKVIDFGQYSIPIEFIDYPNSLKYGVTIVSGVCIFICLVCMTLVVVF
KKARVIKSSSPAFLLLILLGCCIIFAACILFAQSPTNQTCSARIWLLSLGYTLFLGNL
LVKNWRIWLLFDNPKLKK---RA-ITNWKLYPWVFAILAIDVMILAIWQGLGNINA-
ESRIGYDSLTQY-QYKNVCDQGS-IALYLLLVFHGLVLLVACFISFKIK-
DIEEFNESKPITTSVYIITFCLFIVIPLMVSPQSL---
TSQTTIICVCAIVTTLISMLLLFGSKFYKMATQGLANET
>Q86HH3_DICDI_GABA_B
DV-
NGCVLKGDKYSKTISFGTYTIPIEFVDYSSSMKLGITITSSICIFLCIISIIIVLVFRTAR
IIKSASPAFLFLILMGCILIFIGCIIFSQSPNEGTCRARVWLLSIGYTIFLGSLLVKNW
RIWLLFDNPKLKK---RS-ITNWKLYPWVAGILAADVLILAFWQGLGNIRS-
ESRIGIDSLTKY-QYTNVCDQGS-IALYILLVFHGIKLLVACFISFKIK-
DIDEFNESKPIASSVYIITFCLFIVIPLMVSPQSV---
TSQVTTICVCAIVTTLISIILLFGSKFYKMITQGAANQT
>Q55AP1_DICDI_GABA_B
N--
TKLLLKGTDISKTINLGLYTIPIEFVDYSNSMKLGLTIVSGFCILFCIISMVLVIMFR
HAKIIKSASPIFCLLILFGCIIIFSGCIIFSLSPTDGICGARVWLLSIGYTIFLGSLLVKN
WRIWLLFDNPKLKK---RS-ITNWKLYPFVAGILAADVLILALWQGLGDIRS-
ESRIGIDSLTKY-QYANVCS-GS-VALYILLVFHGIKLLAACFISFKIK-
DIEEFNESKPIASSIYIITFCLFIVIPLMVSPQSV---
ASQVITIVVCAIVTTLISISLLFGSKFYMMATQGLANTK
>Q54SW3_DICDI_GABA_B

DP-ILESIGSGNWINDTQFRANTKI---
LPYPESLKIGVTVVSGFCIFLCLISMIIVIKFKEAKVIKSSSPIFCLLILFGCIVIFVGCI
MFARSPTDGSCRSRVWLLSLGYTIFLGNLMVKNWRIWLLFDNPKLKK---RA-
ITNWKLYPWVSGIVIIDIVILSIWQALGDIVA-ESRTGIDSLTKY-EYRNVCDQGS-
IALYLLLVFHGLILLVACFISFKIK-DIEEFNESKPITTSVYIITFCLFIVIPIMVSSPTV-
--TTQTTIICICALITTMLSIILLFGTKFFKMIT-PKINVL
>Q8WTM9_CAEEL_GABA_B
DT-TSGDLEWGKGPPPDSTVIKKHA---MTVSNEFYYPTILFAVLGIAACV-
FIYLFTQKHHRLIIFQSQPECNNILLIGCSLCLFSLFLIGL-P-
PLLCHARVTILLFGFTFAYGSMFAKVWIVHRMG--RKSSGQAYGA-
LSSSKFYVIVAALTAVDVFVCFVWVLI-DPPLFTPADSEEDEMIM--PVLQCQ-SN-
QWIGIIMGFKCLLLVFGTFLSYETRNLKLRFNDSRFVGLAIYNVAVMTLVTAPVV
TLLIHG---K------VDANFAFISLTVGLIYGPKIRHIIKVPP-NVG
>Q9BML5_DROME_GABA_B
DF-RCPRC----
RPVKWHSGQVPIRVFKLRVATIAFYTIATLSSVGIALAITFLAFNLHFRKLKAIKLS
SPKLSNITAVGCIFVYATVILLGL-
PSATVCTARVYLLSAGFSLAFGSMFAKTYRVHRIFT-RTGSVFKDKM-
LQDIQLILLVGGLLLVDALLVTLWVVT-
DPMERHLHNLTDRSVVYQPQVEVCSQHTQTWLSVLYAYKGLLLVVGVYMAWE
TR-IP-ALNDSQYIGVSVYSVVITSAIVVVLANLISERV--
TLAFITITALILTSTTATLCLLFIPKLHDIWARNDINTR
>Q9VPS7_DROME_GABA_B
DF-RCPRC----
RPVKWHSGQVPIRVFKLRVATIAFYTIATLSSVGIALAIAFLAFNLHFRKLKAIKLS
SPKLSNITAVGCIFVYATVILLGL-
PSATVCTARVYLLSAGFSLAFGSMFAKTYRVHRIFT-RTGSVFKDKM-
LQDIQLILLVGGLLLVDALLVTLWVVT-
DPMERHLHNLTDRSVVYQPQVEVCSQHTQTWLSVLYAYKGLLLVVGVYMAWE
TR-IP-ALNDSQYIGVSVYSVVITSAIVVVLANLISERV--
TLAFITITALILTSTTATLCLLFIPKLHDIWARNDINTR
>Q8IPW4_DROME_GABA_B
DF-RCPRC----
RPVKWHSGQVPIRVFKLRVATIAFYTIATLSSVGIALAIAFLAFNLHFRKLKAIKLS
SPKLSNITAVGCIFVYATVILLGL-
PSATVCTARVYLLSAGFSLAFGSMFAKTYRVHRIFT-RTGSVFKDKM-
LQDIQLILLVGGLLLVDALLVTLWVVT-
DPMERHLHNLTDRSVVYQPQVEVCSQHTQTWLSVLYAYKGLLLVVGVYMAWE
TR-IP-ALNDSQYIGVSVYSVVITSAIVVVLANLISERV--
TLAFITITALILTSTTATLCLLFIPKLHDIWARNDINTR
>Q8MSP9_DROME_GABA_B
DF-RCPRC----
RPVKWHSGQVPIRVFKLRVATIAFYTIATLSSVGIALAIAFLAFNLHFRKLKAIKLS
SPKLSNITAVGCIFVYATVILLGL-
PSATVCTARVYLLSAGFSLAFGSMFAKTYRVHRIFT-RTGSVFKDKM-
LQDIQLILLVGGLLLVDALLVTLWVVT-

DPMERHLHNLTDRSVVYQPQVEVCSQHTQTWLSVLYAYKGLLLVVGVYMAWE
TR-IP-ALNDSQYIGVSVYSVVITSAIVVVLANLISERV--
TLAFITITALILTSTTATLLSAFHPKTP----------
>Q5TRW6_ANOGA_GABA_B
------LKWVGRSPPKDRT--L-
RIIEHSQVNITIFVVLASTSCVGIIMATVFLAVNITFRNQRYIKMSSPHLNNLIIIGCI
LTYLSVIFLGL-PY--ICTARAWLLMAGFSLAFGAMFSKTWRVHSIFTDLKLNK---
KV-IKDYQLFIVVGVLLAIDLAIMTTWQIA-DPF-YRETKYIDVIIVQ--
ENEYCSSKMTIFIGVIYAYKGLLLIFGAFLAWETR-IP-
ALNDSKHVGLSVYNCVIMCVMGAAIALVLSDRK--
DAMFILISVFIIFCTTATLCLVFVPKLVELKRNPSGVVD
>Q9BML6_DROME_GABA_B
------VKWVGKTPPKDRT--L-
IYIEHSQVNPTIYIVSASASVIGVIIATVFLAFNIKYRNQRYIKMSSPHLNNLIIVGC
MITYLSIIFLGL-PY--ICTARAWILMAGFSLSFGAMFSKTWRVHSIFTDLKLNK---
KV-IKDYQLFMVVGVLLAIDIAIITTWQIA-DPFYRETKQLEDVLVIP--
ENEYCSEHMTIFVSIIYAYKGLLLVFGAFLAWETR-IP-
ALNDSKHIGFSVYNVFITCLAGAAISLVLSDRK--
DLVFVLLSFFIIFCTTATLCLVFVPKLVELKRNPQGVVD
>Q9Y133_DROME_GABA_B
------VKWVGKTPPKDRT--L-
IYIEHSQVNPTIYIVSASASVIGVIIATVFLAFNIKYRNQRYIKMSSPHLNNLIIVGC
MMTYLSIIFLGL-PY--ICTARAWILMAGFSLSFGAMFSKTWRVHSIFTDLKLNK---
KV-IKDYQLFMVVGVLLAIDIAIITTWQIA-DPFYRETKQLEDVLVIP--
ENEYCSEHMTIFVSIIYAYKGLLLVFGAFLAWETR-IP-
ALNDSKHIGFSVYNVFITCLAGAAISLVLSDRK--
DLVFVLLSFFIIFCTTATLCLVFVPKLVELKRNPQGVVD
>GABR2_RAT_GABA_B
----DTIRFQGSEPPKDKT----
IILEQLRISLPLYSILSALTILGMIMASAFLFFNIKNRNQKLIKMSSPYMNNLIILGG
MLSYASIFLFGL-
TFETLCTVRTWILTVGYTTAFGAMFAKTWRVHAIFKNVKMKK---KI-
IKDQKLLVIVGGMLLIDLCILICWQAV-DPLRRTVERYS---
DISIRPLLECNTHMTIWLGIVYAYKGLLMLFGCFLAWETRNVSIALNDSKYIGMS
VYNVGIMCIIGAAVSFLTRDQP--
NVQFCIVALVIIFCSTITLCLVFVPKLITLRTNPDAATQ
>GABR2_HUMAN_GABA_B
----DTIRFQGSEPPKDKT----
IILEQLRISLPLYSILSALTILGMIMASAFLFFNIKNRNQKLIKMSSPYMNNLIILGG
MLSYASIFLFGL-
TFETLCTVRTWILTVGYTTAFGAMFAKTWRVHAIFKNVKMKK---KI-
IKDQKLLVIVGGMLLIDLCILICWQAV-DPLRRTVEKYSMERDISIRP-
LHCNTHMTIWLGIVYAYKGLLMLFGCFLAWETRNVSIALNDSKYIGMSVYNVGI
MCIIGAAVSFLTRDQP--NVQFCIVALVIIFCSTITLCLVFVPKLITLRTNPDAATQ
>Q9BML7_DROME_GABA_B

DNLSW-QWIGGKVPQDRVTHVL-RTV-----
SLPLFVCMCTISSCGIFVAFALIIFNIWNKHRRVIQSSHPVCNTIMLFGVIICLISVIL
LGI-PEPKICQARAWLLSTGFTLAYGAMFSKVWRVHR-FT-TKAKTDPKKK-
VEPWKLYTMVSGLLSIDLVILLSWQIF-DPLQRYLETFPLE-
DIKIRPELECSQRNSMWLGLVYGFKGLILVFGLFLAYETR-
SIKQINDSRYVGMSIYNVVVLCLITAPVGMVIASQQ--
DASFAFVALAVIFCCFLSMLLIFVPKVIEVIRHPKDKAE
>Q9V3Q9_DROME_GABA_B
DNLSW-QWIGGKVPQDRVTHVL-RTV-----
SLPLFVCMCTISSCGIFVAFALIIFNIWNKHRRVIQSSHPVCNTIMLFGVIICLISVIL
LGI-PEPKICQARAWLLSTGFTLAYGAMFSKVWRVHR-FT-TKAKTDPKKK-
VEPWKLYTMVSGLLSIDLVILLSWQIF-DPLQRYLETFPLE-
DIKIRPELECSQRNSMWLGLVYGFKGLILVFGLFLAYETR-
SIKQINDSRYVGMSIYNVVVLCLITAPVGMVIASQQ--
DASFAFVALAVIFCCFLSMLLIFVPKVIEVIRHPKDKAE
>GABR1_RAT_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKT-
LEPWKLYATVGLLVGMDVLTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSKKMNTWLGIFYGYKGLLLLLGIFLAYETK-
STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>Q8IW08_HUMAN_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKT-
LEPWKLYATVGLLVGMDVLTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSRKMNTWLGIFYGYKGLLLLLGIFLAYETK-
STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>Q5SUJ8_HUMAN_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKT-
LEPWKLYATVGLLVGMDVLTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSRKMNTWLGIFYGYKGLLLLLGIFLAYETK-
STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>GABR1_HUMAN_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKT-
LEPWKLYATVGLLVGMDVLTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSRKMNTWLGIFYGYKGLLLLLGIFLAYETK-

STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>Q6PGJ2_MOUSE_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKERRKT-
LEPWKLYATVGLLVGMDILTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSKKMNTWLGIFYGYKGLLLLLGIFLAYETK-
STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>GABR1_MOUSE_GABA_B
DDWSKTDKWIGGSPPADQTLVI-KTFRFL--
SQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSLALAA
VFPLGL-PF--VCQARLWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKT-
LEPWKLYATVGLLVGMDILTLAIWQIV-DPLHRTIETFIDVSILP--
QLEHCSKKMNTWLGIFYGYKGLLLLLGIFLAYETK-
STEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQ--
DAAFAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSE
>O96954_GEOCY_GABA_B
NP--------GMWPDGIPNDGIPIEVVTV--SVPLTVVYVALAVGGLVFAI-
VCVFFTVIRKRKLIRLSSPNLNYLIGLGAIILYFNVITLVI-
PTTDTCNINPWLTSLGYSLCYGTILAKTIRIWFIFNKPRVPS-VTKS-
IVIKALALFVVSLVVIDVIILGIFAIVEGLRGELAVHSNKE-
NIEDTIGPTCSKGQVALFTVLFGYKGLLQVTALILAFNTRKVKVKLDDSKYIAAAI
YVTSIVLAVAAISTYTLRDYV--
NIYPAVVGIGFLLGTTMILGLVFVPRMVGLYQDPQ-NIK
>RAI3_HUMAN_RAI
----------DKAEAWGIVLETVATAGVVTSVAFM-LTLPI-----
VQDSNRRTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAV
SLTKLVRGRKPLS---LL-V----ILGLAVGFSLVQDVIAIEYIVLTMN--RTNVNVFSE--
------LSAP-RRNEDFVLLLTYVLFLMALTFLMSSFTFCGS--
FTGWKHIYLTMLLSIAIWVAWITLLMLPDFDR--
RWDDTILSSALAANGWVFLLAYVSPEFWLLTK--QRNPM
>RAI3_MOUSE_RAI
----------DSRYH-RLC--------DLAEGWGIALETLAAVGAVATVAFA-LVFLICK---
VQDSNLPAQFLFLLGVLGVFGLTFAFIIKLDGATGPTRFFLFGVLFAICFSCLLAHA
FNLIKLVRGRKPLS---WL-V----ILSLAVGFSLVQDVIAIEYLVLTMN--
RTNVNVFSE--------LPAP-RRNEDFVMLLIYVLVLMVLTFFTSFLVFCGS--
FSGWKHICFTSFLSIAIWVAWIVLLLIPDIDR--
KWDDTILSTALVANGWVFLAFYILPEFRQLPR-PTDNRA
>Q6PA25_XENLA_RAI
----------DTTAAWGIVLETLAAAGIVFSIILI-
LALLIRISDPAKRAVSPVQLIFLIGTFGIFGLTFAFIVELTEQTCPTRFFLFGVLFAIC
FSCLLAHASKLVRLVRGGHGIC---WW-M----
MLLMALFLPLVQVVIAILYIVLGLV--RSG----------
TPCTVFNGADFVLILIYVFLLMAITFLVSLISLCGPCKYRHGAHIYVTMFFSIGIWV

AWICMLLRGNANQGNTWDDPVLSIALVANGWVFLMMYMVPELCLMTR-
PHLNTN
>Q5M7Q6_XENTR_RAI
----------DLNAAWGIVLQAITSAGIVST-FIL-
TIVLVASTPFIQNSLLGTQVFFLLGTLGIFGLVFDMVVKKDFATCASRRFLFGLLF
AICFACLWVHGVSLNLIRRNTQPS----GW-W----
IFGLAVCLSLVEAIINTEWLIIPPA----------DPCLIGN-
ADFVMALIYVMFLILAGFITAWPALCGRYWRKHAIFILLTLFLSISIWIVWIVMYV
YGNVQVGNYWDDPTLAIALVSNGWAFIFFYIIPEISQLTK-PTLEQT
>GPC5C_HUMAN_RAI
----------DRSGAWGIVLEAVAGAGIVTT-FVL-
TIILVASLPFVKRSLLGTQVFFLLGTLGLFCLVFACVVKPDFSTCASRRFLFGVLFA
ICFSCLAAHVFALNFLARKNHGPR---GW-V----IFTVALLLTLVEVIINTEWLIIGPQ-
-GNSSAGWAVA-------SPCA-
IANMDFVMALIYVMLLLLGAFLGAWPALCGRYWRKHGVFVLLTTATSVAIWVV
WIVMYTYGNKQHNSTWDDPTLAIALAANAWAFVLFYVIPEVSQVTK-PEQSYQ
>GPC5C_MOUSE_RAI
DL----------DPLYY-NLC--------DRSGAWGIVSEAVAGAGIITT-FVL-
TIILVASLPFVKRSLLGTQVFFLLGTLGLFCLVFACVVKPDFSTCASRRFLFGVLFA
ICFSCLVAHVLSLNFLTRKNHGPR---GW-V----IFTVALLLTLVEVIINTEWLIISPL--
GNVSADSTMT-------SPCA-IANMDFVMALIYVMLLLLTAFLGAWPTLCGR--
FKRWRFVLLTTVISIAIWVVWIVMYTYGNEQHHSTWDDPTLAIALAANAWTFVL
FYVIPEVSQVTK-PEQSYQ
>GPC5B_MOUSE_RAI
DL----------LPQYV-SLC--------DLDAIWGIVVEAVAGAGALITLLLM-LILLVR-
LPFIKDKERCLHFLFLLGTLGLFGLTFAFIIQMDETICSIRRFLWGVLFALCFSCLLS
QAWRVRRLVRQGTSPA---SW-Q----LVSLALCLMLVQVIIATEWLVLTVL--RDTK-
----------PACA-YEPMDFVMALIYDMVLLAITLAQSLFTLCGK--
FKRWKFILVTTFLSALIWVVWMTMYLFGNSLI--
KQSDPTLAITLAASGWVFVIFHAIPEIHYTLLPPLQENP
>GPC5B_HUMAN_RAI
DL----------LPQYV-SLC--------DLDAIWGIVVEAVAGAGALITLLLM-LILLVR-
LPFIKESPVGLHFLFLLGTLGLFGLTFAFIIQEDETICSVRRFLWGVLFALCFSCLLS
QAWRVRRLVRHGTGPA---GW-Q----LVGLALCLMLVQVIIAVEWLVLTVL--
RDTR----------PACA-YEPMDFVMALIYDMVLLVVTLGLALFTLCGK--
FKRWKLNGAFLLITAFLSVLIWVAWMTMYLFG--
NVNDPTLAITLAASGWVFVIFHAIPEIHCTLL-PAENTP
>BOSS_DROVI_Boss
DNHTS--LELGLVFNCQFTAGENRRYPFL-KLDTWVATGLTAAILGLIATLAIL-
VFIVVRSLGDVFEGNPVTSILLLLSLILVFCSFVPFSM-
TLNTLCGVRVFIMTLVYCFVFSLLLCRAVMLASIGSEGGFLSHVNGY-I----
QAIICVLSVFVQVGMSVQLLVVMHL--ASES----------
VSCNIYYGRWLWGLLAYDFLLLCSLVSLVPFIYRSQRNYREGILIVIGAVLILIIWS
VWIALSMFGDE-----WRDAAIPLGMQASGWAVLVGILIPRTFLIVRGIERNNQ
>Q8MSJ2_DROME_Boss

DNHTS--LELGQVFNCQYSAGDNRRYPFL-KMDTWVAAGLTAAILGLIATLAIL-
VFIVVRSLGDVFEGNPTTSILLLLSLILVFCSFVPYSI-
TLNTLCAVRVFIMTLVYCFVFSLLLCRAVMLASIGSEGGFLSHVNGY-I----
QAVICAFSVVAQVGMSVQLLVVMHV--ASET----------
VSCNIYYGRWLWGLLAYDFALLCCVGALIPSIYRSQRNYREGILIVIGSVLIMVIW
VAWIALSLFGDE-----WRDAAIPLGLQASGWAVLVGILIPRTFLIVRGIERNNQ
>BOSS_DROME_Boss
DNHTS--LELGQVFNCQYSAGDNRRYPFL-KMDTWVATGLTAAILGLIATLAIL-
VFIVVRSLGDVFEGNPTTSILLLLSLILVFCSFVPYSI-
TLNTLCAVRVFIMTLVYCFVFSLLLCRAVMLASIGSEGGFLSHVNGY-I----
QAVICAFSVVAQVGMSVQLLVVMHV--ASET----------
VSCNIYYGRWLWGLLAYDFALLCCVGALIPSIYRSQRNYREGILIVIGSVLIMVIW
VAWIALSLFGDE-----WRDAAIPLGLQASGWAVLVGILIPRTFLIVRGIERNNQ
>Q4SE66_TETNG_MGR

AELCPYEQRPDRNRTGCRPIPVVRLEWRSAWALLPVFLSVLGLVATSFVAVTFLR
YRHTPLVRASGRELSFLLLCGIFLCYLATFPMLAAPGVLVCSLRRVLLGLGLSCS
YAALLTKTNRIHRIFQQGRKAVLAPRC-
ISPASQLLVAFSLISVQLLGILVWFAADPPHSFVDYGERRT--P-ARGVLRCD-
ISDLALICSLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIIWLAFIPIFF
GTAQ-SLQHIQTAMLTVSLSLSASVSLGMLFVPKVYIIIFHPEQNVG
>Q68EF4_MOUSE_MGR
DRYTCKTCPYDMRPTENRTSCQPIPIVKLEWDSPWAVLPLFLAVVGIAATLFVVV
TFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGM
SISYAALLTKTNRIYRIFEQGKRSVSAPRF-
ISPASQLAITFVLISLQLLGICVWFVVDP--SVVDFQDQRTLDPFARGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSADK----------VTSEALPVEFSPPLLAHN--------
>Q62916_RAT_MGR
DRYTCKTCPYDMRPTENRTSCQPIPIVKLEWDSPWAVLPLFLAVVGIAATLFVVV
TFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGM
SISYAALLTKTNRIYRIFEQGKRSVSAPRF-
ISPASQLAITFILISLQLLGICVWFVVDP--SVVDFQDQRTLDPFARGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSADKYIQTTTLTVSVSLSASVSLGMLYMPKVYIILF-PSSLSC
>MGR4_RAT_MGR
DRYTCKTCPYDMRPTENRTSCQPIPIVKLEWDSPWAVLPLFLAVVGIAATLFVVV
TFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGM
SISYAALLTKTNRIYRIFEQGKRSVSAPRF-
ISPASQLAITFILISLQLLGICVWFVVDP--SVVDFQDQRTLDPFARGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSADKYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPEQNVP
>Q6ZMQ2_HUMAN_MGR
DRYTCKTCPYDMRPTENRTGCRPIPIIKLEWGSPWAVLPLFLAVVGIAATLFVVIT
FVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGMSI
SYAALLTKTNRIYRIFEQGKRSVSAPRF-ISPASQLAITFSLISLQLLGICVWFVVDP-

-SVVDFQDQRTLDPFARGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSADKYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPEQNVP
>MGR4_HUMAN_MGR
DRYTCKTCPYDMRPTENRTGCRPIPIIKLEWGSPWAVLPLFLAVVGIAATLFVVIT
FVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGMSI
SYAALLTKTNRIYRIFEQGKRSVSAPRF-ISPASQLAITFSLISLQLLGICVWFVVDP-
-SVVDFQDQRTLDPFARGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSADKYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPEQNVP
>Q4S5Q2_TETNG_MGR
DTYTCKMCRFDLRPNGNHTGCDPIPIVKLEWSSPWAVIPVLIAVLGIMATLFVVA
TFVRYNDTPIVKASGRELSYVLLTGIFLCYATTFLMISAPDVFICSLRRIFLGLGMS
ISYAALLTKTNRIYRIFEQGKMSVSAPRF-
ISPASQLVITFSLISVQLLGVCIWFGVDP--AIIDYEDQRTANPKMAGVLKCD-
ISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPI
FFGTSQSAEK----------SNLSKFKQL----------
>Q4RJZ9_TETNG_MGR
DETSCRLCAYNMRPNPNRTACQPIPIVKLEWHSPWAVIPVFLAMLGIIATIFVMAT
FVRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCAFRRIFLGLGMCI
SYAALLTKTNRIYRIFEQGKQTVTAPRF-
ISPTSQIAITSSLICVQLLGVLVWFAVDPPNTIVDYDEQKTIPMLARGVLKCD-
ITDLQIICSLGYSILLMVTCTIYAIKTRDVPEDFNEAKPIGFTMYTTCIVWLAFIPIFF
GTAQSAEKYIQTTTLTISMNLSASVALGMLYMPKVYIIIFHPELNVQ
>MGR7_PONPY_MGR
DEMTCQHCPYDQRPNENRTGCQDIPIIKLEWHSPWAVIPVFLAMLGIIATIFVMAT
FIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRVFLGLGMCI
SYAALLTKTNRIYRIFEQGKKSVTAPRL-
ISPTSQLAITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNP-ARGVLKCD-
ITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTMYTTCIVWLAFIPIF
FGTAQSAEKYIQTTTLTISMNLSASVALGMLYMPKVYIIIFHPELNVQ
>MGR7_HUMAN_MGR
DEMTCQHCPYDQRPNENRTGCQDIPIIKLEWHSPWAVIPVFLAMLGIIATIFVMAT
FIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRVFLGLGMCI
SYAALLTKTNRIYRIFEQGKKSVTAPRL-
ISPTSQLAITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNP-ARGVLKCD-
ITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTMYTTCIVWLAFIPIF
FGTAQSAEKYIQTTTLTISMNLSASVALGMLYMPKVYIIIFHPELNVQ
>MGR7_RAT_MGR
DEMTCQHCPYDQRPNENRTGCQNIPIIKLEWHSPWAVIPVFLAMLGIIATIFVMAT
FIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRVFLGLGMCI
SYAALLTKTNRIYRIFEQGKKSVTAPRL-
ISPTSQLAITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNP-ARGVLKCD-
ITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTMYTTCIVWLAFIPIF
FGTAQSAEKYIQTTTLTISMNLSASVALGMLYMPKVYIIIFHPELNVQ
>MGR7_MOUSE_MGR

DEMTCQHCPYDQRPNENRTGCQNIPIIKLEWHSPWAVIPVFLAMLGIIATIFVMAT
FIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRVFLGLGMCI
SYAALLTKTNRIYRIFEQGKKSVTAPRL-
ISPTSQLAITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNP-ARGVLKCD-
ITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTMYTTCIVWLAFIPIF
FGTAQSAEKYIQTTTLTISMNLSASVALGMLYMPKVYIIIFHPELNVQ
>MGR8_HUMAN_MGR
DELSCELCPLDQRPNMNRTGCQLIPIIKLEWHSPWAVVPVFVAILGIIATTFVIVTF
VRYNDTPIVRASGRELSYVLLTGIFLCYSITFLMIAAPDTIICSFRRVFLGLGMCFS
YAALLTKTNRIHRIFEQGKKSVTAPKF-
ISPASQLVITFSLISVQLLGVFVWFVVDPPHIIIDYGEQRTLDPKARGVLKCD-
ISDLSLICSLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIIWLAFIPIFF
GTAQSAEKYIQTTTLTVSMSLSASVSLGMLYMPKVYIIIFHPEQNVQ
>MGR8_MOUSE_MGR
DELSCELCPLDQRPNINRTGCQRIPIIKLEWHSPWAVVPVFIAILGIIATTFVIVTFV
RYNDTPIVRASGRELSYVLLTGIFLCYSITFLMIAAPDTIICSFRRIFLGLGMCFSYA
ALLTKTNRIHRIFEQGKKSVTAPKF-
ISPASQLVITFSLISVQLLGVFVWFVVDPPHTIIDYGEQRTLN--ARGVLKCD-
ISDLSLICSLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIIWLAFIPIFF
GTAQSAEKYIQTTTLTVSMSLSASVSLGMLYMPKVYIIIFHPEQNVQ
>MGR8_RAT_MGR
DELSCELCPLDQRPNINRTGCQRIPIIKLEWHSPWAVVPVFIAILGIIATTFVIVTFV
RYNDTPIVRASGRELSYVLLTGIFLCYSITFLMIAAPDTIICSFRRIFLGLGMCFSYA
ALLTKTNRIHRIFEQGKKSVTAPKF-
ISPASQLVITFSLISVQLLGVFVWFVVDPPHTIIDYGEQRTLN--ARGVLKCD-
ISDLSLICSLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIIWLAFIPIFF
GTAQSAEKYIQTTTLTVSMSLSASVSLGMLYMPKVYIIIFHPEQNVQ
>MGR6_HUMAN_MGR
DEFTCEACPGDMRPTPNHTGCRPTPVVRLSWSSPWAAPPLLLAVLGIVATTTVV
ATFVRYNNTPIVRASGRELSYVLLTGIFLIYAITFLMVAEPGAAVCAARRLFLGLG
TTLSYSALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFSLTSLQVVGMIAWLGARPPHSVIDYEEQRT----ARGVLKCD-
MSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIIWLAFVP
IFFGTAQSAEKYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQNVQ
>MGR6_RABIT_MGR
DEFTCEACPRDMRPTPNHTGCRPTPVVRLTWSSPWAAPPLLLAVLGIMATTTVV
GTFVRHNNTPIVRASGRELSYVLLTGIFLIYAVTFLMVAEPGAAVCATRRLFLGL
GTTLSYSALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFSLTSLQVVGVIAWLGAQPPHSVIDYEEQRTVDV-----LKCD-
MSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIVWLAFV
PIFFGTAQSAEKYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQNVQ
>Q8CFQ7_MOUSE_MGR
DEFTCEACPGHMRPTPNHTGCRPTPVVRLTWSSPWAALPLLLAVLGIMATTTIIA
TFMRHNDTPIVRASGRELSYVLLTGIFLIYAITFLMVAEPCAAVCASRRLLLGLGT
TLSYSALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFGLTSLQVVGVIAWLGAQPPHSVIDYEEQRTVDV-----LKCD-

MSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIIWLAFVP
IFFGTAQSAEKYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQNVQ
>MGR6_MOUSE_MGR
DEFTCEACPGHMRPTPNHTGCRPTPVVRLTWSSPWAALPLLLAVLGIMATTTIIA
TFMRHNDTPIVRASGRELSYVLLTGIFLIYAITFLMVAEPCAAVCASRRLLLGLGT
TLSYSALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFGLTSLQVVGVIAWLGAQPPHSVIDYEEQRTVDV-----LKCD-
MSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIIWLAFVP
IFFGTAQSAEKYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQNVQ
>MGR6_RAT_MGR
DEFTCEACPGDMRPTPNHTGCRPTPVVRLTWSSPWAALPLLLAVLGIMATTTIM
ATFMRHNDTPIVRASGRELSYVLLTGIFLIYAITFLMVAEPCAAICAARRLLLGLG
TTLSYSALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFGLTSLQVVGVIAWLGAQPPHSVIDYEEQRTVDV-----LKCD-
MSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIIWLAFVP
IFFGTAQSAEKYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQNVQ
>Q8NHA9_HUMAN_MGR
DEFTCEACPGDMRPTPNHTGCRPTPVVRLSWSSPWAAPPLLLAVLGIVATTTVV
ATFVRYNNTPIVRASGRELSYVLLTGIFLIYAITFLMVAEPGAAVCAARRLFLGLG
TTLSYFALLTKTNRIYRIFEQGKRSVTPPPF-
ISPTSQLVITFSLTSLQESVTFRDVAVDP-
WMVESGVPQPGSTVGTYQSLQIDKTNQLSLTCFSGIQGFLASVNSLVMDKGKNS
PTIFGEALLQKFPAKDLGCILLQYVDDLLLGDSMADGDLRDTSAQKAKLIALTRA
LELREESKGRGCLGQPGKCEG
>Q75QW6_APIME_MGR
DVTQCISCRQGTVPDETHSACRDIPEEFLRPESGWAIGAMSFSATGILITLFVCGVF
LKHNDTPVVRASGRELSYVLLSGILLCYLVTFALVLRPTDIVCGIQRFAAGFCFTV
VYAALLTKTNRISRIFNASKHSAKRPSF-ISPRSQLIICSGLVFVQILINGVWMIIDP-
AKAMHHYPTREDNL-----
LVCNSYVDASYMIAFAYPIMLIVVCTVYAVLTRKIPEAFNESKHIGFTMYTTCVI
WLAFVPLYFGT-GNN--ALRITSMSVTISLSASVTIACLFSPKLYIILIRPERNVR
>Q9V4U3_DROME_MGR
DETHCKLCKLGTLPDAHKQYCRPIPEIYLRPESAWAIGAMAFSATGILVTLFVMG
VFVRHNDTPIVRASGRELSYILLAGIFMCYGVTFALVLKPTNIVCAIQRFGVGFCF
TVVYAALLTKTNRIARIFKAGKQSAKRPSF-
ISPKSQLVICACLVSVQILINGVWMVIAP-SHAMHHYPTREDNL-----
LVCDSYIDASYMIAFSYPIFLIVICTVYAVLTRKIPEAFNESKHIGFTMYTTCVIWL
AFVPLYFGTANHV--PLRITSMSVTISLSASVTIACLFSPKLYIILIRPERNVR
>Q4REZ5_TETNG_MGR
DEFTCSPCAPGQWPTDDLTSCYDLPEEYIMWEDAWAIGPISIACVGFMCTGLVIW
VFIRHNHTPLVKASGRELCYILLSGVFMSYAMTFLFLAKPSPAICALRRLGLGTSF
AVCYSALFTKTNRIARIFNGVKDGAARPRF-
ISPSSQVFICLSLISVQLVMVSVWLLLEVPGTRRFTLPERRQTVI----LKCN-
VRDSSMLLSLGYDVFLVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLP
IFYVTSSDY--RVQTTTMCISVSLSGFVVLGCMFAPKVHIIMFQPQKNVT
>MGR3_MOUSE_MGR

DEFTCMDCGPGQWPTADLSGCYNLPEDYIRWEDAWAIGPVTIACLGFMCTCIVIT
VFIKHNNTPLVKASGRELCYILLFGVSLSYCMTFFFIAKPSPVICALRRLGLGTSFA
ICYSALLTKTNCIARIFDGVKNGAQRPKF-
ISPSSQVFICLGLILVQIVMVSVWLILETPGTRRYTLPERETVI-----LKCN-
VKDSSMLISLTYDVVLVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPI
FYVTSSDY--RVQTTTMCISVSLSGFVVLGCLFAPKVHIVLFQPQKNVV
>MGR3_RAT_MGR
DEFTCMDCGPGQWPTADLSGCYNLPEDYIKWEDAWAIGPVTIACLGFLCTCIVIT
VFIKHNNTPLVKASGRELCYILLFGVSLSYCMTFFFIAKPSPVICALRRLGLGTSFA
ICYSALLTKTNCIARIFDGVKNGAQRPKF-
ISPSSQVFICLGLILVQIVMVSVWLILETPGTRRYTLPERETVI-----LKCN-
VKDSSMLISLTYDVVLVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPI
FYVTSSDY--RVQTTTMCISVSLSGFVVLGCLFAPKVHIVLFQPQKNVV
>MGR3_PONPY_MGR
DEFTCMDCGPGQWPTADLTGCYDLPEDYIRWEDAWVIGPVTIACLGFMCTCMV
VTVFIKHNNTPLVKASGRELCYILLFGVGLSYCMTFFFIAKPSPVICALRRLGLGSS
FAICYSALLTKTNCIARIFDGVKNGAQRPKF-
ISPSSQVFICLGLILVQIVMVSVWLILEAPGTRRYTLAERETVI-----LKCN-
VKDSSMLISLTYDVILVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPI
FYVTSSDY--RVQTTTMCISVSLSGFVVLGCLFAPKVHIILFQPQKNVV
>MGR3_HUMAN_MGR
DEFTCMDCGSGQWPTADLTGCYDLPEDYIRWEDAWAIGPVTIACLGFMCTCMV
VTVFIKHNNTPLVKASGRELCYILLFGVGLSYCMTFFFIAKPSPVICALRRLGLGSS
FAICYSALLTKTNCIARIFDGVKNGAQRPKF-
ISPSSQVFICLGLILVQIVMVSVWLILEAPGTRRYTLAERETVI-----LKCN-
VKDSSMLISLTYDVILVILCTVYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPI
FYVTSSDY--RVQTTTMCISVSLSGFVVLGCLFAPKVHIILFQPQKNVV
>MGR2_HUMAN_MGR
DEFTCADCGLGYWPNASLTGCFELPQEYIRWGDAWAVGPVTIACLGALATLFVL
GVFVRHNATPVVKASGRELCYILLGGVFLCYCMTFIFIAKPSTAVCTLRRLGLGT
AFSVCYSALLTKTNRIARIFGGAREGAQRPRF-
ISPASQVAICLALISGQLLIVVAWLVVEAPGTGKETAPERREVVT----LRCN-
HRDASMLGSLAYNVLLIALCTLYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFL
PIFYVTSSDY--RVQTTTMCVSVSLSGSVVLGCLFAPKLHIILFQPQKNVV
>MGR2_RAT_MGR
DEFTCADCGLGYWPNASLTGCFELPQEYIRWGDAWAVGPVTIACLGALATLFVL
GVFVRHNATPVVKASGRELCYILLGGVFLCYCMTFVFIAKPSTAVCTLRRLGLGT
AFSVCYSALLTKTNRIARIFGGAREGAQRPRF-
ISPASQVAICLALISGQLLIVAAWLVVEAPGTGKETAPERREVVT----LRCN-
HRDASMLGSLAYNVLLIALCTLYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFL
PIFYVTSSDY--RVQTTTMCVSVSLSGSVVLGCLFAPKLHIILFQPQKNVV
>MGR1_CAEEL_MGR
NETHCVGCEQGWWPTKDRKGCFDLSLSQLKY-
SMYSLVPTILAVFGIIATLFVIVVYVIYNETPVVKASGRELSYILLISMIMCYCMTF
VLLSKPSAIVCAIKRTGIGFAFSCLYSAMFVKTNRIFRIFS--TRSAQRPRF-
ISPISQVVMTAMLAGVQLIGSLIWLSVVPPGWRHHYPT-RDQVV-----LTCN-

VPDHHFLYSLAYDGFLIVLCTTYAVKTRKVPENFNETKFIGFSMYTTCVVWLSWI
FFFFGTGSDF--QIQTSSLCISISMSANVALACIFSPKLWIILFEKHKNVR
>Q8MXU2_CAEEL_MGR
NETSCEECAVGMVP--
DRTLCVPIPPVSMQWDTTWSLIPAAFSTLGIASTIFVVSVFLKFSNTPVIMASGREL
CYCMMSGIGMCYTLTFFLVSQPTVITCSMTRILMGLSMSAIYAAIITKTNRLARVF
K--PDSAQRPRF-ITPKAQVGICMGIVSVQLIGTFVWILFDPPGTMIVFPTRTEAV----
--LTCK-
ATTSHLLISLLYNILLIVACTVYAFKTRKIPENFNETRHIGFTMYSTCILWLAFGPIY
FATQSDF--RIQITSLCMCISLSGTVALICFFAPKVYIVLFQPYKNVR
>Q75QW7_APIME_MGR
DEYTCMDCGPGKWPHEDKRGCYQLAINHIRWNSAFAIAPAVISCLGIVATMAVA
CLLFHHRDTPVVRASGRELTIILLAGVLVCYLNTFLLLATPTTVTCILQRFGVGVS
FSAVYGALLTKTNRIARIFDSASRTAVRPRY-
ISPASQVCIAAALIALQIVLTLVWMIIEPPGTRFFYPDRKQVI------LKCN-
IQDMSFLFSQLYNALLILISTVYAVKTRKIPENFNESKFIGFTMYTTCIIWLAFVPIY
FGTGNAH--ETQITTLCVAISLSATVTLVCLYSPKIYIILFQPDKNIR
>MGR_DROME_MGR
DEFTCKDCGPGLWPYADKLSCYALDIQYMKWNSLFALIPMAIAIFGIALTSIVIVL
FAKNHDTPLVRASGRELSYTLLFGILVCYCNTFALIAKPTIGSCVLQRFGIGVGFSI
IYSALLTKTNRISRIFHSASKSAQRLKY-
ISPQSQVVITTSLIAIQVLITMIWMVVEPPGTRFYYPDRREVI------LKCK-
IQDMSFLFSQLYNMILITICTIYAIKTRKIPENFNESKFIGFTMYTTCIIWLAFVPIYF
GTGNSY--EVQTTTLCISISLSASVALVCLYSPKVYILVFHPDKNVR
>MGR5_RAT_MGR
DEYTCKACQLGSWPTDDLTGCDLIPVQYLRWGDPEPIAAVVFACLGLLATLFVT
VIFIIYRDTPVVKSSSRELCYIILAGICLGYLCTFCLIAKPKQIYCYLQRIGIGLSPAM
SYSALVTKTNRIARILAGSKKKICTKKF-
MSACAQLVIAFILICIQLGIIVALFIMEPPDIMHDYPSIRE-VY-----LICN-
TTNLGVVTPLGYNGLLILSCTFYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFVP
IYFGSNY----K--IITMCFSVSLSATVALGCMFVPKVYIILAKPERNVR
>MGR5_HUMAN_MGR
DEYTCKACQLGSWPTDDLTGCDLIPVQYLRWGDPEPIAAVVFACLGLLATLFVT
VVFIIYRDTPVVKSSSRELCYIILAGICLGYLCTFCLIAKPKQIYCYLQRIGIGLSPA
MSYSALVTKTNRIARILAGSKKKICTKKF-
MSACAQLVIAFILICIQLGIIVALFIMEPPDIMHDYPSIRE-VY-----LICN-
TTNLGVVTPLGYNGLLILSCTFYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFVP
IYFGSNY----K--IITMCFSVSLSATVALGCMFVPKVYIILAKPERNVR
>Q4R3P0_MACFA_MGR
DEYTCKACQLGSWPTDDLTGCDLIPVQYLRWGDPEPIAAVVFACLGLLATLFVT
VVFIIYRDTPVVKSSSRELCYIILAGICLGYLCTFCLIAKPKQIYCYLQRIGIGLSPA
MSYSALVTKTNRIARILAGSKKKICTKKF-
MSACAQLVIAFILICIQLGIIVALFIMEPPDIMHDYPSIRE-VY-----LICN-
TTNLGVVTPLGYNGLLILSCTFYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFVP
IYFGSNY----K--IITMCFSVSLSATVALGCMFVPKVYIILAKPERNVR
>Q98UC4_CHICK_MGR

DEYTCKACQLGSWPNDELTGCDLIPVQYLRWGDPEPIAAVVFACLGLLATLFVT
AIFIMYRDTPVVKSSSRELCYIILAGICLGYLCTFCLIAKPQQIYCYLQRIGIGLSPA
MSYSALVTKTNRIARILAGSKKKICTKKF-
MSACAQLVIAFILICIQLGIIVALFIMEPPDIMHDYPSIRE-VY-----LICN-
TTNLGVVTPLGYNGLLILSCTFYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFVP
IYFGSNY----K--IITMCFSVSLSATVALGCMFVPKVYIILAKPERNVR
>MGR1_HUMAN_MGR
DEFTCKACDLGWWPNADLTGCEPIPVRYLEWSNIESIIAIAFSCLGILVTLFVTLIF
VLYRDTPVVKSSSRELCYIILAGIFLGYVCPFTLIAKPTTTSCYLQRLLVGLSSAMC
YSALVTKTNRIARILAGSKKKICTRKF-
MSAWAQVIIASILISVQLTLVVTLIIMEPPMPILSYPSIKE-VY-----LICN-
TSNLGVVAPLGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFV
PIYFGSNY----K--IITTCFAVSLSVTVALGCMFTPKMYIIIAKPERNVR
>MGR1_RAT_MGR
DEFTCRACDLGWWPNAELTGCEPIPVRYLEWSDIESIIAIAFSCLGILVTLFVTLIF
VLYRDTPVVKSSSRELCYIILAGIFLGYVCPFTLIAKPTTTSCYLQRLLVGLSSAMC
YSALVTKTNRIARILAGSKKKICTRKF-
MSAWAQVIIASILISVQLTLVVTLIIMEPPMPILSYPSIKE-VY-----LICN-
TSNLGVVAPVGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFV
PIYFGSNY----K--IITTCFAVSLSVTVALGCMFTPKMYIIIAKPERNVR
>MGR1_MOUSE_MGR
DEFTCRACDLGWWPNAELTGCEPITIRYLEWSDIESIIAIAFSCLGILVTLFVTLIFV
LYRDTPVVKSSSRELCYIILAGIFLGYVCPFTLIAKPTTTSCYLQRLLVGLSSAMCY
SALVTKTNRIARILAGSKKKICTRKF-
MSAWAQVIIASILISVQLTLVVTLIIMEPPMPILSYPSIKE-VY-----LICN-
TSNLGVVAPVGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFV
PIYFGSNY----K--IITTCFAVSLSVTVALGCMFTPKMYIIIAKPERNVR
>Q5TZ45_BRARE_MGR
DEFTCRACDLGWWPDDELAECKPLPLKYLEWGSVESIIAVVFSCVGILVTLFVTFI
FFQYSDTPVVKSSSRELCYIILAGIFLGYICPFTLIARPTVISCYLQRILVGLSSAMC
YSALVTKTNRIARILAGSKKKICTRKF-
MSAWAQVVIAFMLISLQLAVVVTLMVLEPPEPVKSYPSIRE-VY-----LICN-
TSNVGVVAPLGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFV
PIYFGSNY----K--IITTSFSVSLSVTVALGCMFTPKMYIIIAKPERNVR
>Q90ZF3_ONCMA_MGR
DEFTCTACDLGWWPDPELEGCEPITLRYLEWGNPESIVQVVFACLGILVTSFVTFI
FVLYRDTPVVKSSSRELCYIILAGIFLGYICPFTLIAQPTVASCYLQRLLVGLSATM
CYSALVTKTNRIARILAGSKKKICTRKF-
MSAWAQLVIAGLLVSVQLTLEVTLIILEPPMPVKSYPSIRE-VF-----LICN-
TSTVGMVAPLGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAF
VPIYFGSNY----K--IITTSFSVSLSVTVALGCMFSPKIYIILAKPERNVR
>Q4REI8_TETNG_MGR
DEFTCKACELGWWPDKELEGCELIPVHYLEWSNPESIIQVVFSCMGILVTSFVAFI
FVLYRDTPVVKSSSRELCYIILAGIFLGYLCPFTLIARPTVASCYLQRLLVGLSTAM
CYSALVTKTNRIARILAGSKKKICTRKF-
MSAWAQVVIASILISLQLSLEVTLIILEPPEPIKSYPSIRE-VF-----LICN-

TSNMGVVAPLGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAF
VPIYFGSNY----K--IITTSFSVSLSVTVALGCMFTPKMYIILAKPERNVR
>Q90WL6_SPAAU_CS
DASVCTKCPNNSWSNGNHTFCFLKEIEFLSWTEPFGIALAICAVLGVVLTAFVMG
VFVRFRNTPIVKATNRELSYVLLFSLICCFSSSLIFIGQPQDWMCRLRQPAFGISFV
LCISCILVKTNRVLLVFEAKIPTSLHRKW-
WGLNLQFLLVFLCTFVQVMICVVWLYNAPPSSYMNH-DIDEIIF-----
ITCNEGSVMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWIS
FIPAYFSTYG----KFVSAVEAIAILASSFGMLACIFFNKVYIILFKPSRNTI
>Q5YEV6_OREMO_CS
DASVCAKCPNNSWSNGNHTFCFLKEVEFLSWTEPFGIALAICAVLGVVLTAFVIG
VFVRFRNTPIVKATNRELSYVLLFSLICCFSSSLIFIGEPQDWTCRLRQPAFGVSFV
LCISCILVKTNRVLLVFEAKIPTSLHRKW-
WGLNLQFLLVFLCTFVQVMICVVWLYNAPPSSYQNH-DIDEIIF-----
ITCNEGSVMALGFLIGYTCILAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWISF
IPAYFSTYG----KFVSAVEVIAILASSFGMLACIFFNKVYIILFKPSRNTI
>O73635_FUGRU_CS
DASICTKCPNNSWSSGNHTFCFLKEIEFLAWSEPFGIALAICAVLGVLLTAFVMGV
FVRFRNTPIVKASNRELSYVLLLSLICCFSSSLIFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSIHRKW-
WGLNLQFLLVFLCTFVQVMICVVWLYNAPPSSYRNH-DIDEIIF-----
ITCNEGSVMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAKFITFCMLIFFIVWIS
FIPAYFSTYG----KFVSAVEAIAILASSYGMLACIFFNKVYIILFKPCRNTI
>Q6XAF3_SALSA_CS
DASVCTKCPNDSWSNENHTSCFLKEIEFLSWTEPFGIALALCSVLGVFLTAFVMG
VFIKFRNTPIVKATNRELSYLLLFSLICCFSSSLIFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSLHRKW-
WGLNLQFLLVFLFTFVQVMICVVWLYNAPPASYRNH-DIDEIIF-----
ITCNEGSMMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWIS
FIPAYFSTYG----KFVSAVEVIAILASSFGLLACIFFNKVYIILFKPSRNTI
>Q6XAF1_SALSA_CS
DASVCTKCPNDSWSNENHTSCFLKEIEFLSWTEPFGIALALCSVLGVFLTAFVMG
VFIKFRNTPIVKATNRELSYLLLFSLICCFSSSLIFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSLHRKW-
WGLNLQFLLVFLFTFVQVMICVVWLYNAPPASYRNH-DIDEIIF-----
ITCNEGSMMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWIS
FIPAYFSTYG----KFVSAVEVIAILASSFGLLACIFFNKVYIILFKPSRNTI
>Q6XAF2_SALSA_CS
DASVCTKCPNDSWSNENHTSCFLKEIEFLSWTEPFGIALALCSVLGVFLTAFVMG
VFIKFRNTPIVKATNRELSYLLLFSLICCFSSSLIFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSLHRKW-
WGLNLQFLLVFLFTFVQVMICVVWLYNAPPASYRNH-DIDEIIF-----
ITCNEGSMMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWIS
FIPAYFSTYG----KFVSAVEVIAILASSFGLLACIFFNKVYIIHQ-PQDQ--
>Q8JI04_SQUAC_CS

DASACTKCPNDFWSNENHTSCIAKEIEYLSWTEPFGIALTIFAVLGILITSFVLGVFI
KFRNTPIVKATNRELSYLLLFSLICCFSSSLIFIGEPRDWTCRLRQPAFGISFVLCISC
ILVKTNRVLLVFEAKIPTSLHRKW-
VGLNLQFLLVFLCILVQIVTCIIWLYTAPPSSYRNHELEDEVIF-----
ITCDEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWIS
FIPAYVSTYG----KFVSAVEVIAILASSFGLLGCIYFNKCYIILFKPCRNTI
>Q4R5V3_MACFA_CS
DASACNKCPDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKW-
WGLNLQFLLVFLCTFMQIVICVIWLYTAPPSSYRNHELEDEIIF-----
ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWIS
FIPAYASTYG----KFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTI
>CASR_HUMAN_CS
DASACNKCPDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKW-
WGLNLQFLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF-----
ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWIS
FIPAYASTYG----KFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTI
>CASR_MOUSE_CS
DASACDKCPDDFWSNENYTSCIAKEIEFLAWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKW-
WGLNLQFLLVFLCTFMQIVICIIWLYTAPPSSYRNHELEDEIIF-----
ITCHEGSLMALGSLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWIS
FIPAYASTYG----KFVSAVEVIAILAASFGLLACIFFNKVYIILFKPSRNTI
>CASR_RAT_CS
DASACDKCPDDFWSNENHTSCIAKEIEFLAWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKW-
WGLNLQFLLVFLCTFMQILICIIWLYTAPPSSYRNHELEDEIIF-----
ITCHEGSLMALGSLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWIS
FIPAYASTYG----KFVSAVEVIAILAASFGLLACIFFNKVYIILFKPSRNTI
>Q93564_CAEEL_MGR
--
NQCMNCTLGSWPNANRTGCEYIIPEVVSWTSFGHILALVLAVTGIITSMATLAVF
LRHNSTPVVKSTTRELSYIILSGLVACYAVSFALLATPSTTSCFITRVIPPIAFAVVY
SALLTKTNRIARILAGSKKLTKKPRF-
LTTFSQVVITWILVAVQCVIVGVGLMRDWPDATYAKYALPRKLI-----LECD-
TETKSFLIPFFWDFFLITLCTLYAFKTRNLPENFNEAKFIGFTISQPAKSLPFFQKEH
VIIFQYCTVTHKALVMSFSYSLSASVALALLFFPKLYIILMHPEKNIR
>Q622H2_CAEBR_MGR
--
DQCMNCSLGSWPNANRTGCEDIIPEVVSWTSFGHILALVLAVTGIITSLATLAVFL
RHNSTPVVKSTTRELSYIILSGLVACYAVSFALLATPSSSSCFITRVIPPIAFAVVYS

ALLTKTNRIARILAGSKKLTKKPRF-
LTTFSQVVITWILVAVQCVIVGVGLMRDWPDATYAKYALPRKLI-----LECD-
TETKSFLIPFFWDFFLISLCTLYAFKTRNLPENFNEAKFIGFTMKTPRIMKKRKTLV
FRYCTVVVWIAFKALVMSFSYSLSASVALALLFFPKLYIILMHPEKNIR
>GPC6A_HUMAN_V2R
DMPHCLLCN-
THWAPVRSTMCFEKEVEYLNWNDSLAILLLILSLLGIIFVLVVGIIFTRNLNTPVV
KSSGGRVCYVILLCHFLNFASTSFFIGEPQDFTCKTRQTMFGVSFTLCISCILTKSL
KILLAFSFDPKLQKFLKC-LYR--PILIIFTCTGIQVVICTLWLIFAAPTVEVNV-
SLPRVII-----LECEEGSILAFGTMLGYIAILAFICFIFAFKGK--
YENYNEAKFITFGMLIYFIAWITFIPIYATTFG----
KYVPAVEIIVILISNYGILYCTFIPKCYVIICKQEINTK
>GPC6A_RAT_V2R
DMDHCLVCN-
THWAPVRSTMCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAVGIIFTRNLKTPV
VKSSGGVVCYVMLACHALNFASTGFFIGEPQDFTCKTRQTLFGVSFTLCVSCILT
KSLKILLAFSFDPTLKTFLKC-LYR--PVPIVLTCTGIQVVICTLWLVLAAPTVEENT-
SLPRVII-----
LECEEGSALAFGTMLGYIAVLAFICFVFAFKGRKLPENYNEAKFLTFGMLIYFIA
WITFIPVYATTFG----KYLPAVEIIVILISNYGILCCTFFPKCYIILCKQKTNTK
>GPC6A_MOUSE_V2R
DMDHCLLCN-
THWAPVRSTTCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAIGIIFTRNLKTPVV
KSSGGVVCYVMLICHALNFASTGFFIGEPQDFACKTRQTLFGVSFTLCVSCILTKS
LKILLAFSFDPKLTMFLKC-LYR--PVPIVLTCTGIQVVICTLWLVLAAPSVEENI-
SLPRVII-----
LECEEGSALAFGTMLGYITVLAFICFVFAFKGRKLPENYNEAKFLTFGMLIYFIAW
ITFIPVYTTTFG----KYLPAVEIIVILISNYGILCCIFFPKCYIILCKQKTNTK
>GPC6A_BRARE_V2R
DADECSPCD-
SMWSLANSTECHPKVFEYFDWNSGFAIVLLILAALGVLLLFFMSALFFWQRHSP
VVKAAGGPLCHLILVSLLGSFISVVFFVGEPSDLTCRARQVIFGFSFTLCVSCILVK
SLKILLAFEMNFELKELLCM-LYK--
PYMIVSVGMGVQIIICTVWLTLYKPFKDKEV--QTESIL-----
LECNEGFYVMFWLMLGYIALLALFCFTFAYIGRKLPQKYNEAKFITFSMVICLMA
WIIFIPIHVTTSG----KYVPAVEMVVILISNYGILSCHFLPKSYIILFKKEHNTK
>GPC6A_CARAU_V2R
DADTCSKCD-
SLWSNANSSKCYPKFYEYFEWNSGFAIALLTLAALGILLLISMSALFFWQRNSLV
VKAAGGPLCHLILFSLLGSFISVIFFVGEPSNESCRVRQVIFGLSFTLCVSCILVKSL
KILLAFQMNLELKELLRK-LYK--PYVIVCMCMGLQVTICTLWLTLHRPFIEKVV--
QPKSIL-----
LECNEGSDLMFGLMLGYIVLLALICFTFAYKGRKLPQKYNEAKFITFGMLIYLMA
WVIFIPVHVTTSG----KYVPAVEVVVILISNYGILSCHFLPKCYIIIFKKEYNTK
>Q767K3_XENLA_V2R

DSERCARCPDKEWPNENKDQCITKTEEFLAN-
DMMSVFFSTISVLIFLVTLLILGVFITYRDSPIVRANNRTQ-
LLLLVSIKLSFLSVFLFLGRPVDITCMLRNITFGITFSIAVSSLLAKTIMVCVAFKAT
KPGSTWRKW-VGAKLSYSVVLFCSSIQIIICMTWLAISPPFQELDLHTYPGTII-----
IQCNEGS-
YWLLSVIGYMGLLAAVSFVLAFLARTLPDSFNEAKYITFSMLLFCSVWITMIPAY
LSTKG----KNTVCVEIFAILTSSAGLLACIFLPKCYIILFTPEMNTK
>O35267_RAT_V2R
NMNQCIKCLNDQYANPGGTRCLKKVIVFLGYEDPLGMSLAILALCFSALTAFVLS
IFLKHQETPTVKANNRTLSYVLLISLISCFLCSLLFIGHPSFTTCIMQQTTFAVVFTV
AASTVLAKTIIVILAFKVTNTSRKMRWL-
LVSGAPKFIIPICTMIQLILCGIWLGTSPPFVDADGHVEKGHIL-----
IFCNKGSILAFYCVLGYLVSIAIASFTLAFFARNLPDTFNEAKFLTFSMLVFCSVW
VTFLPVYHSTKG----KSMVAVEVFCILASSAGLLFCIFAPKCFIILLRPEKNIH
>AAC53403.1_V2R
DMEQCVKCPYDKYANIEKTHCLSRAVSFLAYEDPLGIALGCIALSFSAITILVLITF
LKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQVSCVLQQTTFGVFFTVS
VSTVLAKTITVVMAFKLTTPGRRMREM-
LVTGAPKLVIPICTLIQFVLCGIWLITSPPFIDRDIQSEHGKIV-----
ILCNKGSVIAFHVVLGYLGSLALGSFTLAFLARNLPDTFNEAKFLTFSMLVFCSV
WITFLPVYHSTRG----KVMVVVEVFSILASSAGLLMCIFVPKCYVILVRPDSNFI
>O35189_MOUSE_V2R
DMEQCVRCPDDKYANIEQTHCLSRAVSFLAYEDSLGMALGCMALSFSAITILILV
TFVKYKDTPTVKANNRILSYILLISLVFCFLCSLLFIGPPDQVTCIFQQTTFGVLFTV
SVSTVLAKTITVVMAFKLTTPGRRMRGM-
MMTGAPKLVIPICTLIQLVLCGIWLVTSPPFIDRDIQSEHGKIV-----
ILCNKGSVIAFHVVLGYLGSLALGSFTLAFLARNLPDTFNEAKFLTFSMLVFCSV
WITFLPVYHSTRG----RVMVVVEVFSILASSAGLLMCIFVPKCYVILIRPDSNFI
>O35190_MOUSE_V2R
DMEQCVRCPDNKYANLEQTHCLQRTVSFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPDQVTCILQQTTFGVLFT
VSVSTVLAKTITVVMAFKLTTPGRRMRGM-
MMTGAPKLVIPICTLIQLVLCGIWLVTSPPFIDRDIQSEHGKIV-----
ILCNKGSVVAFHVVLGYLGSLALGSFTLAFLARNLPDTFNEAKFLTFSMLVFCSV
WITFLPVYHSTRG----KVMVVVEVFSILASSAGLLMCIFVPKCYVILIRPDSNFI
>AAC08419.1_V2R

----------KTITVVMAFRITAPGIKMRGM-
LASGTPNLVIPICTLIQIVLCGIWLVASPPFIERDTQSEHGKTI-----
IVCNKGSAIAFHFVLGYLGSLALGSFTVAFMARNLPDRLNEAKFLTFSMLVFCSV
WITFLPVYHSTRG----KFMVVVEVFSILASTAGLLGCIFVPKCYIILVRPDLNAL
>AAC53407.1_V2R
DVDQCVKCPESHYANTEKSNCFPKSVSFLAYEDPLGMALASIALCLSALTVFVIG
IFVKNRDTPIVKANNRTLSYILLITLTFCFLCSLNFIGQPNTAACILQQTTFAVAFT
MALATVLAKAITVVLAFKISFPGRMLRWL-
MISRGPRYIIPICTLIQLLLCGIWMATSPPFIDQDVNTEDGYII-----

LLCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTFNESKFLSFSMLVFFCV
WVTFLPVYHSTKG----KV----------
>AAC53406.1_V2R
DVDQCVKCPESHYANTEKSNCYQKSVSFLGYEDPLGMALASIALCLSALTAFVI
GIFVKHKDTPIVKANNQALSYTLLITLKFCFLCSLNFIGQPNTVACILQQTTFAVAF
TMALATVLAKAITVVLAFKVSFPGRMVRWL-
MISRGPNYIIPICTLIQLLLCGIWMAISPPYIDQDAHIEHGHII-----
ILCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTFNESKFISLSMLVFFCV
WITFLPVYHSTKG----KV----------
>O35269_RAT_V2R
DVDQCVKCPETHYANIEKIHCLQKTVTFLYYDDPLGKTLCFMSLGFSSLTAAVL
VVFLKNRDTPIVKANNLALSYTLLITLMLCFLCPLLFIGRPSTASCILQQNIFGLLF
TVALSTVLAKTITVVIAFKITSPGRIRRWL-
LISRAPNFIIPLCTLLQVFLSGIWLTTSPPFIDKDAHSEHGHII-----
IICNKGSAVAFHCNLGYLGALALVSYFMAFLSRNLPDTFNEAKFLAFSMLVFCSV
WVTFLPVYHSTKG----KNMVAMEVFSILASSTSLLGIIFAPKCYLILLRPERNSL
>Q80Z09_MOUSE_V2R
--
DHCVHCPETHYANAEKRHCLKKTVTFLYYNDPLGKGITLLSIGFSTLTALFIWVF
VDHRGTPIIKANNRSLSYILLITLILCFLCPLLFIGLPNTATCIMQQYMFGLLFTVAL
STVLAKTITVVMAFKITAPGRKIRWL-
LTSQVPKFAIPICTLIQVLLSGIWLGTSPPFIDKDYHSEHGHII-----
ILCNKGSYIYFYCTLAYLGVMAFGSYLLAFLSRRLPDRFNESKSLAFSMLVFCSV
WVMFLPVYHSTSG----KVMVAMEMFSILASSASILIIVFAPKCYIVLFRPELNIL
>O35265_RAT_V2R
DVGQCVKCPESHYANTEKSHCLKKTMTFLDYNDSLGTGLTLMSLGFFVVTGLVI
GVFIIHRNTPIVKANNRSLSYILLITLTLCFLCPLLFIGLPNTATCILQQNLFGLLFTV
ALSTVLAKTITVVMAFKITAPGRKTRWL-
LILRAPQFIIPLCALMQILFSGIWLGTSPPFVDMDAHSEHGHII-----
ILCNKGSAIGFYCTLAYLGVMAFGSYLLAFMSRNLPDTFNESKALAFSMLMFCS
VWVTFLPVYHSTTG----KVRVAMEMFSILASSASILTLIFVPKCYIVLFRPERNIL
>AAC08418.1_V2R

ITFGVVFTVSVSTVLAKTITVVLAFEITASQRMMKYI-
LVSRASNYIILVCTLIQVILCAVWLGASPPSVDIDAQSEHGHII-----
IVCHKGSVTAFYCVLGYLGTLAFGSFTLAFFSRIFLVPFNEAKSITFSMLVFCSVW
VTFIPVYHSSKG----KMMVAVEIFSTLASSAGMLGCIFVPKCYTVLFRPEQNSL
>O35268_RAT_V2R
NVELCVQCPEDQYANQEQNHCIHKARIFLSYDEPLGMALSLMALCLAALTVVVL
GVFVKHHRTPIVKANNCTLTYILLIALIFCFLCPLFFIGHPNSATCILQQITFGVVFT
VAISTVLAKTTTVILAFRVTAPHRMMKYF-
LVSRASNYIIPICTLIQIIVCAIWLGASPPSVDIDAQSEHGHII-----
IACNKGSVTAFYCVLGYLACLAFVSFTLAFLSRNLPVTFNEAKSMTFSMLVFCSV
WVTFLPVYHGTKG----KVMVAVEIFSTLASSAGMLGCIFAPKCYTILFRPDRNSL
>Q9QWK0_RAT_V2R

CCFVCSPCPDNEISNE----------
TTVVLWVFVKHHDTPIVKANNRILSYILIMSLMFCFLCSFFFIGHPNRGTCILQQIT
FGIVFTVAVSTVLAKTITVLLAFQVTDTGRKLRNFLVSGT-
PNYIIPICSLLQCTLCAIWLAVSPPFVDIDEHSEHGHII-----
IVCNKGSVMAFYCVLGYLAFLALGSFTMAFLAKNLPDTFNEAKFLTFSMLVFCS
VWITFLPVYHSTKG----RVMVAVEIFSILTSSAGMLGCVFAPKIYIILMKPERLSK
>O35271_RAT_V2R
CCFVCNPCPENEISNE----------
TMVVFWVFVKHHDTPIVKANNRILSYLLIVSLMFCFLCSFFFIGYPNRATCILQQI
TFGIFFTVAISTVLAKTITVVLAFKVTDPGRQLRIF-
LVSGTPNYIIPICSLLQCILCAIWLAVSPPFVDIDEHSEHGHII-----
IVCNKGSITAFYCVLGYLACLAFGSFTIAFLAKNLPDTFNEAKFLTFSMLVFCAV
WVTFLPVYHSTKG----KVMVAVEIFSILASSAGMLGCIFAPKVYIILMRPDRNSI
>AAC53333.1_V2R
NMDQCANCPEYQYANTEKNKCIQKGVIVLSYEDPLGMALALIAFCFSAFTVVVF
WVFVKHHDTPIVKANNRILSYLLIVSLMFCFLCSFFFIGYPNRATCILQQITFGIFFT
VAISTVLAKTITVVLAFKVTDPGRQLRIF-
LVSGTPNYIIPICSLLQCILCAIWLAVSPPFVDIDEHSEHGHII-----
IVCNKGSITAFYCVLGYLACLAFGSFTIAFLAKNLPDTFNEAKFLTFSMLVFCAV
WVTFLPVYHSTKG----KVMVAVEIFSILASSAGMLGCIFAPKVYIILMRPDRNSI
>O35266_RAT_V2R
--
DQCVNCPEYQYANTEKNKCIQKDVIFLSYEDPLGMALALIAFCLSAFTAVVLWV
FVKHHDTPIVKANNRILSYILIMSLMFCFLCSFFFIGHPNRGTCILQQITFGIVFTVA
VSTVLAKTITVILAFKLRDPGRSLRNF-
LVSGAPNYIIPICSLLQCILCAIWLAVSPPFVDIDEHSEHGHIM-----
IVCNKGSIMAFYCVLGYLACLALGSFTTAFLAKNLPDTFNEAKFLTFSMLVFCSV
WVTFLPVYHSTRG----RVMVAVEIFSILASSAGMFGCIFAPKIYIILMKPERNSI
>O35202_MOUSE_V2R
CCFVCSPCPENEISNE----------
TTVVLCVFVKHHDTPIVKANNRSLSYLLLMSLMSCFLCSFFFIGLPNRAICVLQQI
TFGIVFTMAVSTVLAKTVTVVLAFKVTDPGRRLRNF-
LVSGTPNYIIPICSLLQCVLCAIWLAVSPPFVDIDEHTLHGHII-----
IVCNKGSVTAFYCILGYLACLALGNFSVAFLAKNLPDTFNEAKFLTFSMLVFCSV
WVTFLPVYHSTKG----KHMVAVEIFSILASSAGILGCIFVPKIYIILMRPERNST
>O35192_MOUSE_V2R
--
DQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVVLC
VFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNKVICVLQQITFGIVFT
VAVSTVLAKTVTVVLAFKVTVPGRRLRYF-
LVSGTLNYIIPICSLLQCVLCAIWLAVSPPFVDIDEHSQHGHII-----
IVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPDAFNEAKFLTFSMLVFCSV
WVTFLPVYHSTKG----KHMVAVEIFSILASSAGMLGCIFVPKIYIILMRPERNST
>Q80Z08_MOUSE_V2R
--
DQCVNCPEYQYANTEQNKCIKKGVTFLSYEDPLGMVLALMAFCFSAFTAVVLC

VFVKHHDTPIVKANNRILSYILLLSLMSCFLCSFFFIGHPNRATCVLQQITFGIVFT
VAVSTVLAKTVTVVLAFKVTDPGRRLRHF-
LVSGTPNYIIPICSLFQCVLCAIWLAVSPPFVDIDEHSQHGHII-----
IVCNKGSVTAFYCVLGYLACLALGSFIVAFLAKNLPDTFNEAKFLTFSMLVFFSV
WVTFLPVYHSTKG----KVMVAVEIFSILASSAGMLGCIFAPKIYIILMRPERNST
>Mm_V2R1d.frame1pep.fasta_V2R
GASLPFGIP--
QYPNKQRNQCLPKITAFLSHEDTLGAVLGSMAISFAAFSAMILGLFIHYRDTPIVR
ANNRNLSYLLLVSLKLCFFCSLMFIGEPRRVTCLLRQMIFGVVFSVALSAVLAKT
FIVVVAFKVIKPGSTLQMR-
MATQISNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWYNEGSTLAFYCVLGYLGFLASLSLFIAFLVRSLPDSFNEAKTITFSMLVFCSVW
ISFVFTYLRSKA----KTMVAVEIVSILASSAGLLSFIFLPKCYVILLRSGGYSR
>Rn_V2R1d.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNKQRNWCHPKITAFLSHEDTLGAVLVSVAISLSVFSAMILG
LFICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQTIFGVVFSI
AISAILAKTFIVVVAFKAIKPGSTLQTW-
MVTRVSNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KTMVAVEILTILASSAGLLGCIFLPKCFVILLKSGGLPR
>Rn_V2R1c.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNKQRNQCLPKITAFLSHEDTLGAVFISLAISFSAFSAMILGL
FICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQTIFGVVFSIA
ISAILAKTFIVVVAFKAIKPGSTLQTW-
MVTRVSNAIVCCGSSIQVCICAVWLGTYPPFPDVDMQSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KNVVAVEILAILASSAGLLGCIFLPKCFVILLRSDDLPR
>AAC08415.1_V2R

----------KTFIVVVAFKAIKPGSTLQTW-
MVTRVSNAIVCCGSSIQVCICAVWLGTYPPFPDVDMQSEFGQII-----
LWCNEGSTLGFYCVLGYLGFLGSLSLLIAFLARRLPDTFNEAKTITFSMLVFCRV
WISFVPAYLSSKA----KNVVAVEILAILASSAGLLGCIFLPKCFVILLRSDDLPR
>Rn_V2R1a.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNKQRNQCLPKITAFLSCEDPLGAVLVSVAISLSAFSAMILG
LFIHYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFICQPTTVTCILRQMIFGVVFSI
TISAILAKTFIVVVAFKVIKPGSTLQTW-
MVTRVSNAIVCCGSSIQVCICAVWLGTYPPFPDVDMQSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFNEAKTITFSMLVFCSV
WISFVPTYLSSKG----KTMVAVEILSILASSSTLLGCIFLPKCYVILMRSGGHSR
>AAC08412.1_V2R
DMHQCIKCPEDQYPNKQRNQCLPKIILFLAHEDPLGTVLVSLAISLSAFSAMILGL
FICYRETPIVRANNRNLSYLLLISLKLCFFCSLIFIGQPRTVTCVLRQIIFGIVFSIVIS
AILAKTFIVVVAFKAIKPGSILKMG-
MVTRLSNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----

LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILLKSGGHSR
>O70409_MOUSE_V2R
DMHQCIKCPEDQYPNKQRNQCLPKIILFLAHEDPLGTVLVSLAISLSAFSAMILGL
FICYRETPIVRANNRNLSYLLLISLKLCFFCSLIFIGQPRTVTCVLRQIIFGIVFSIVIS
AILAKTFIVVVAFKAIKPGSILKMG-
MVTRLSNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILLKSGGHSR
>Mm_V2R1c.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNKQRNQCLPKIMVFLAHEDPLGTVLVSLAISLSAFSAMILG
LFICYRETPIVRANNRNLSYLLLISLKLCFSCSLMFIGQPRTVTCVLRQIIFGIVFSIV
ISAILAKTFIVVMAFKAIKPGSILKMG-
MVTRLSNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILLKSGGHSR
>Q6TAC4_MOUSE_V2R
DMDQCIKCPEDQYPNKQRNQCLPKIMVFLAHEDPLGTVLVSLAISLSAFSAMILG
LFICYRETPIVRANNRNLSYLLLISLKLCFSCSLMFIGQPRTVTCVLRQIIFGIVFSIV
ISAILAKTFIVVMAFKAIKPGSILKMG-
MVTRLSNAIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFCSV
WISFVPAYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILLRSGGHSR
>Rn_V2R1b.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNRQRNQCLPKITEFLSHEDTLGAVFVSIAISLSALSAMILGL
FICYRDTSIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQMIFGIVFSV
AVSTILAKTFIVVVAFKAIKPGSALKMS-
MVIRLSNTIVFCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNETKTITFSMLVFCSVW
ISFVPTYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYMILMSSGGHSR
>Rn_V2R1e.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNTQRNQCLPKITDFLSHEDTLGAVLVFVALSFSAFSAMILR
LFICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPKTVTCVLRQVIFGVVFS
VAVSTILAKTFIVVVAFKAIKPGSALKMS-
MVIRLSNTIVFCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWCNEGSTFAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFFSVW
ISFVPTYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILMKSGDHSR
>AAC08414.1_V2R
DMDQCIKCPEDQYPNTQRNQCLPKITDFLFHEDTLGAVLVFVALSFSAFSAMILR
LFICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPKTVTCVLRQMIFGVVFS
VAVSTILAKTFIVVVAFKAIKPGSALKMS-
MVIRLSNTIVFCGSIIQVCICAVWLGTYPPFPDVDMHSEFGQII-----
LWSNEGSTFAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEAKTITFSMLVFFSVW
ISFVPAYLSSKG----KTMVAVEILSILASSAGLLGCIFLPKCYVILMKSGDHSR
>Mm_V2R1a.frame1pep.fasta_V2R

DMDQCIKCPEDQYPNKQRNQCLPKIMAFLSHEDTLGALLVCTAISLSAFAAMILG
LFIHYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPGTVTCVLRQMIFGVVFS
IAVSAILAKTFIVVVAFKSIKPRSTLQMW-
IVTRLSITIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGHII-----
LCCNEGSTLAFYCVLGYLGFLASLSLLIAF-
VRRLPDSFNEAKTITFSMLVFCSVWISFIPTYLSSTG----
KTMVAVEIFCILASSAGLLGCMFLPKCYVILLRSGDHSR
>Mm_V2R1b.frame1pep.fasta_V2R
DMDQCIKCPEDQYPNKQRNQCLPKIMVFLSHEDTLGAVLVSLAISLSAFSAMILG
LFVCYRDTSIVRANNRNLSYLLLISLMFCFFCSLIFIGQPSTVTCVLRQMIFGVVFS
VAVSAILAKTFIVVVAFKAIKPGSRLQMW-
MVMRLSNVIVCCGSIIQVCICAVWLGTYPPFPDVDMHSEFGHII-----
LCCNEGSTLAFYCVLGYLGFLASLSLLIAFLVRRLPDSFNEAKTITFSMLVFCSVW
ISFIPTYLSSTG----KTMVAVEIFSILASSAGLLGCMFLPKCYVILLRSADLSR
>O73637_FUGRU_CS
DSMECTSCPEDFWSSPQRDHCVPKKTEFLSYHEPLGICLTAASLLGTVISVVVLGI
FIHHRSTPVVRANNSELSFLLLVSLKLCFLCSLLFIGRPRLWTCQLRHAAFGISFVL
CVSCILVKTMVVLAVFRASKPGGGATKW-
FGAVQQRGTVLGLTSIQAAICFAWLLSSSPKPHKNIQYHKDKIV-----
FECVVGSTVGFAVLLSYIGLLAILSFLLAFLARNLPDNFNEAKLITFSMLIFCAVW
VAFVPAYINSPG----KYADAVEVFAILTSSFGLLVALFGPKCYIILFRPERNTK
>Q5I5D1_9TELE_CS
DATKCTVCPDDFWSNLYKDQCIPKEVEFLSYNDPLGISLTTASLLGTCFCAFVMV
IFVHHHNTPIVRANNSELSFLLLFSLKLCFLCVLLFIGRPQLWTCQLRHAVFGISFV
LCISSILVKTMVVIAVFKSSRPGKMAMKW-
FGAAQQRCTILALTAIQVVICVVWLSTASPTPHKNNLYIRSIIV-----
YECTIGSVTGFSMLLGYIGLLAAVSFLIAFLARNLPDNFNEAKFITFSMLIFCAVWI
TFVPAYVSSPG----KYSVAVEIFAILASSFGLLVSIFAPKCYIILLHPERNNK
>Q5I5C9_9TELE_CS
DAIKCTVCPDEFWSNFNKDQCVPKEIEFLSYEDPLGISLTTASLLGTCFCALVMIIF
AFHHNTPIVRANNSELSFLLLLSLKLCFLCVLLFIGQPQLWTCQLRHAVFGISFVL
CISSILVKTMVVIAVFNSSRPGKGAMQW-
FGAAQQRCTVLVLTVLQVVICAVWLSTSSPTPHKNNQYVRSKIV-----
YECAIGSVAGFSLLLGYIGLLAAISFLLAFLARNLPDNFNEAKFITFSMLIFCAVWI
AFVPAYVSSPG----KYAVAVEIFAILASSFGVLVAIFAPKCYIIILHPERNTK
>Q5I5D0_9TELE_CS
DAIKCTVCPDEFWSNLNKDQCVPKEIDFLSYEDPLGISLTTTSLLGTCFCALVMIIF
TFHRNTPIVRANNSELSFLLLLSLKLCFLCVLLFIGQPQLWTCQLRHAVFGISFVLC
ISSILVKTMVVIAVFKSSRPGSGAMKW-
FGTAQQRCTVLVLTALQIVICAVWLSTSSPTPYKNNQSIRSKIV-----
YECAIGSLAGFSLLLGYIGLLAAISFLLAFLARNLPDNFNEAKFITFSMLIFCAVWV
AFVPAYMSSSG----KYAVAMEIFAILASSFGVLVAIFAPKCYIIILHPERNTK
>Q5I5D2_9TELE_CS
DAIKCTVCPDEFWSNLNKDRCVPKEIDFLSYEDPLGISLTTTSLLGTCFCALVMIIF
TFHRNTPIVRANNSELSFLLLLSLKLCFLCVLLFIGQPQLWTCQLRHAVFGISFVLC
ISSILVKTMVVIAVFKSSRPGSGAMKW-

FGTAQQRCTVLVLTALQIVICAVWLSTSSPTPYKNNQSIRSKIV-----
YECAIGSVAGFSLLLGYIGLLAAISFLLAFLARNLPDNFNEAKFITFSMLIFCAVW
VAFVPAYMSSSG----KYAVAMEIFAILASSFGVLVAIFAPKCYIIILHPERNTK
>Q5I5C7_9TELE_CS
DAIECMLCPDEFWSSPDKDQCVPKEVEFLSYEDPLGISLTTASLLGSCFCVLVMII
FGLHRNTPIVRANNSELSFMLLLSLKMCFLCVLLFIGRPQLWTCQLRHAVFGISFV
LCISSILVKTMVVIAVFKSSRPGKGAMKW-
FGAAQQRCTVLVLTVLQVVICSVWLSTASPTPYKNNQYIRSKIV-----
YECAIGSVAGFSLLLGYIGLLAAVSFLLAFLARNLPDNFNEAKFITFSMLIFCAVW
IAFVPAYVSSPG----KYAVAVEIFAILASSFGLLVAIFAPKCYIILLHPERNTK
>Q5I5C5_9TELE_V2R

DEFWSSPKKNQCVPKDVEFLTYGDPLGISLTTASLLGSCICSAVVVIFAHHRHTPV
VRANNSELSFLLLVSLKLCFLCVLLFIGQPQLWTCRLRHAVFGISFVLCISSILVKT
MVVIAVFKSSRPGKSAMKW-
FGSHQQRCTVLVLTALQVVICAVWLTNASPKPYKNNQYTSSKIV-----
YECTIGSVVGFAMLLGYIGILAAVSFLLAFLARNLPDHFNEAKFITFSMLIFCAVW
IAFVPAYVSSPG----KYAVAVEIFAILASSFGLLAAIFAPKCYIIILHPERNTK
>Q5I5C4_9TELE_V2R

ECLKCPEYYWPNKERVSCVAGIEEFLSYHDVMGIILISLSLFGVAVATVVMVIFFI
FRSTPIVKANNSEMSFLLLLSLKLCFLCSLVFVGRPSPLTCRARQAAFGISFVLCIS
CILVKTIVVLLAFRSTVPGSISLKV-
FGPPQQRVFIFCCTTGQVILCACWLALAPPYPYKNTSYQDGRII-----
LECKDILPLGFYLVLGYIGLLSCMCFVLAFLGRKLPDTFNEAKLITFSMLIFCAVW
ISFIPAYNSSPG----KYTVAVEIFAILASTFGLLFSIFIPKCYVILLRPDLNTK
>Q5I5C8_9TELE_PPR
GSIKCYRCPKEFWSNSYKNECVARETEFLSVKETMGITLMSVAVSGAVMTTTVA
VIFLYHRNTPIVKANNSELSFLLLLSLKLCFLCALVFVGQPSLWSCRIQQAAFGISF
VLCISCILVKTFVVLIAFHSTRPSSALIKW-
FGLGKQRGIVLVFTCVQVVICAIWLCVSPPLPYQNFGMHRSKVI-----
LECTIGSVVGFTCVLGYIGFLATVCFLLAFFARKLPDNFNEAKFITFSMLIFCAVW
VAFVPAYVSSPG----KYSVVVEVFAILASSFGLLFCIFVPKCYIILLKPENNTK
>O73640_FUGRU_CS
GSLQCDRCPPEFWSNDGRTACVPRQLDFLSFNETLGVALTAVAVSGAVVTTAVF
VVFLHYRHTPMVRANNSELSFLLLLSLKLCFLCSLVFIGRPSVWSCRFQQAAFGIS
FVLCVSCLQVKTIVVLAAFRSARPGAGALKWFGPQQRGSVCIFTCVQARVIICIV
WLSLSPPVPQADLDVPGLQVT-----
LECAMASVVGFSLVLGYIGLLACTCLLLAFLARKLPDNFNEAKLITFSMLIFCAV
WVAFVPAYISSPG----KYSVAVEIFAILASSYGLLFCIFAPKCFIILLRPEKNTK
>Q501X9_BRARE_OR
DQAECLLCPEDFWSNKHRNYCVPKEIEFLSYLEAFGMVLAAIAILGAVAAISVGV
VFFRHRDTPLVRANNSELSFLLLISPTLCFICALTFLGQPSHWACPLRRISFGLTFA
LCLSCLLSKTLVVLIAFKSTLPGNNTARW-
FRPPQQRLGVFICSLLQGGVCIAWLTTASPYPVKNTWLYRDRII-----

LECHLGSVVYFCCVLGYIGCLAAFCFILAFLARKLPDNFNEAKFITFSMLIFCAVW
ITFIPAYVSSPG----KFTVAVEIFAILASSYGVLLCIFAPKCYIIIFMPEKNSK
>O73638_FUGRU_PPR
NSAECLQCPLEFWSNEDHSQCVPKVIEFLSFEETMGALLAAVSLFGAALTSLVFC
VFFRFRHTPLVKASNSELSFLLLFSLTLCFLCSLTFIGRPSRWSCVLRHTAFGITFA
LCMSCVLAKTVAVLFAFTAKRPGNTVFYC-SVP-
LQRTSVFACITLQVIICVLWLTLAPPHPHKNTAHAKERII-----
LECNLGSPVWFWVVLGYIGLLAVICFILAFLARKLPDNFNEAKFITFSMLIFCAVW
VTFIPAYVSSPG----KFTVAVEIFAILASSFGLLFCIFAPKCYILILKPEKNTK
>Q6UNX3_ICTPU_OR
NAVTCLPCPLEYKSNGNKTKCDLKSIEFLNFNELMGIPADGILFVWWMSYNHSW
VHFLLLQDLPIVRANNSELSFLLLSSLSLCFLCSLTFIGPPSEWSCMLRHTAFGIAF
VLCISCVLGKTVVVLMAFRATLPGSNIMEVVWTCTAETQCNFLHLCIQVFICSLW
LTISPPFPYKNMHHYKDKII-----
LECHVGSALGFWAVPGYIGILAFLCFILAFLARKLPDNFNEAKYITFSMLIFCVVW
ITFIPAYVSSPG----KFIVAVEIFAILTSSSGLLFCIFFPKCYIILLQPDKNTK
>Q5I5D6_9TELE_PPR
DSINCKQCLEEYWSNAEKNKCVLKAVEFLSFTEIMGIILVFFSLFGVVLTTLVTVL
FYRKKDTPIVKANNSELSFLLLFSLSLCFLCSLTFIGRPTEWSCMLRHTVFGITFVL
CISCVLGKTIVVLMAFKATLPGSNVMKW-
FGPVQQRLSVLAFTLIQIIICVVWITLSPPFPYKNMEYYKEKII-----
IECNLGSTLGFWAVLGYISLLSVLCFVLAFLARKLPDNFNEAKFITFSMLIFCAVW
ITFIPAYVSSPG----KLTVAVEIFAILSSSFGLLFCIFAPKCYIILLKPEQNTK
>Q5I5D4_9TELE_PPR
DSNNCKQCPEEYWPNTEKSKCVLKIVEFLSFTELMGIVLVIFSLFGVTLTVLVTFL
FYSKKDTPIVKANNSELSFLLLFSLTLCFLCSLTFIGRPTEWSCMLRHTTFGITFVL
CISCILGKTIVVLMAFRATIPGNNMIKW-
FGPVQQRLSVLAFTLIQILICVVWLTISPPFPSKNIKYYKEKII-----
LECSLGSTIGFWIVLGYIGLLAVLCFILAFLARTLPDNFNEAKFITFSILIFCAVWITF
IPAYVSSPG----KLTVAVEIFAILASSFGLLFCIFAPKCYIILFKPEQNTK
>Q5I5C3_9TELE_PPR
DSVNCKQCQGEYWSNVEKNKCVLKAVEFLSFTDFMGIVLVFFSLFGVGLTVLVA
ILFYSKKDTPIVKANNSELSFLLLLSLTLCFLCSLTFIGRPTEWSCMLRHTAFGITF
VLCISCVLGKTIVVLMAFKATLPGSNVMKW-
FGPVQQRLSVLAFTLIQVLICVLWLTMSSPFPYKNMKYYKEKII-----
LECSLGSVIGFWAVLGYIGLLAALCFILAFLARTLPDNFNEAKFITFSMLIFCAVW
VTFIPAYVSSPG----KYTVAVEIFAILASTFGLLFCIFVPKCHIIIFKPEQNTK
>Q5I5D3_9TELE_PPR
DSVNCKQCPLEYWSNTEKNKCVFKSVEFLSFTEVMGVVLVFFSLFGVGLTMLVA
ILFYNKKDTPIVKANNSELSFLLLFSLTLCFLCSLTFIGRPTEWSCMLRHTTFGITF
VLCISCVLGKTIVVLMAFKATLPGSNIMKW-
FGPVQQRLSVLAFTLIQVIICVLWLIISPPFPYKNMKYYKEKII-----
LECSLGATICFSAVLFYIGLLAILCFILAFLARKLPDNFNEAKFITFSMLIFCAVWIT
FIPAYVSSPG----KFTVAVEIFAILASSFGLLFCIFVPKCYIILFKNEQNTK
>O93552_CARAU_OR

NE-
TCKPCPWEYWSNAEKNKCVLKAVEFLSFTEIMGVVLVFFSLFGVGLTLLVAILFY
NKKDTPMVKANNSELSFLLLFSLTLCFLCSLTFIGRPTEWSCMLCHTAFGITFVLC
ISCVLGKTIVVLMAFKATLPGNNIMKW-
FGPAQQRLSVLAFTLIQVIICVLWLTISPPFPYKNMKYYKEKII-----
LECSLGSTIGFWAVLTYISLLAFLCFILAFLARTLPDKFNEAKFITFSMLIFCAVWIT
FIPAYVSSPG----KFTVAVEIFAILSSSFGLLFGIFAPKCYIILLKPEQNTK
>Q5I5D5_9TELE_PPR
DSNNCKQCPGEYWSNDKKNKCVLKAVEFLSFTEVMGVLLLFVSLFGVGLTTLV
AILFYNKKDTPIVRANNSELSFLLLFSLILCFLCSITFIGRPTEWSCMLRHTAFGITF
VLCISCVLGKTIVVLMAFKATHPGSNVMKW-
FGPVQQRLSVLVFTFIQVIICVLWLTISPPFPYRNMKYYKEKII-----
LECSLGSIVGFWAVLGYIGLLAALCFILAFLARKLPDNFNEAKFITFSMLIFCAVWI
TFIPAYFSSPG----KLTVAVEIFAILASSIGLLFCIFAPKCYIILCKPEQNTK
>O93553_CARAU_OR
DSVTCLRCHQDFWSNMQNDGCVKKETEFLSYEEIMGILLTTISLVGAFITIIIAVIF
FRYKNTPIVKANNSELSFLLLFSLMLCFLCSLTFIGRPTEWSCMLRHTAFGITFVLC
ISCVLGKTIVVLMAFRATLPGSNVMKW-
FGPPQQRLSVFSFTLIQVIICVLWLTIYPPFPFKNLNYFKEKII-----
LECNVGSVVGFWAVLGYIGLLAILCFFLAFLARKLPDNFNEAKFITFSMLIFCAV
WIAFIPAYVSSPG----KFTVAVEVFAILASTYGMLFCIFIPKCYIILLKPDKNSK
>Q4T805_TETNG_PPR
D---
CVPCINNFWPNPVSNACFPKPVEFLSFSEVLGMILAVFSVGGACLAVIIAAVFFHH
RACPIVRANNSELSFLLLFSLTLCFLCSLTFIGAPSQWSCMLRHTAFGVTFVLCISC
VLGKTVVVLMAFRATLPGSNVMKW-
FGPPQQRLTVFTFTAIQVLICIVWLVVSPPYPVKNLSTYKDRII-----
LECALGSALGFWAVLGYIGLLAAVCLVLAVLARKLPDNFNEAKLITFSMLIFCAV
WITFIPCLCQLTW----KVKKIDEQEIIIIS--GMSSLH-IPKVPQFWVIACVC--
>O73636_FUGRU_PPR
DSPECIPCLDDFWPNPERNACFPKPVEFLSFNEVLGIILAVFSVGGACLAVITAAVF
FHHRTSPIVRANNSELSFLLLFSLTLCFLCSLTFIGAPSHLSCMLRHTAFGITFVLCI
SCVLGKTVVVLMAFRATLPGSNVMKW-
FGPPQQRMTVVTFTSIQVLICIVWLVVNPPFPVRNLTTYKERII-----
LECALGSSVGFWAVLGYIGLLAAVCLVLAVLARKLPDNFNEAKMITFSMLIFCA
VWITFIPAYVSSPG----KFTVAVEIFAILASSFGLILCIFAPKCFIILFKPEKNSK
>Q5I5C6_9TELE_PPR
DALECVKCSPEFWSNTRRDTCIPKVVEFLSYNDTMGITLLSVALLGSCFTLAVTLI
FALKRHTPLVRANNSEISFLILSSLLLCFLCALAFIGRPTTWSCGLRHTAFGIAFSL
CLSCVLGKTLVVLMAFKASLPGSNVMRW-
FGPLQQRGIIFMCTAVQVVICCTWLCVAPPIPHMLMTHETARII-----
LLCDVGSSFAFSLVLGYIGLLAAVCFLLAFFARKLPDNFNEAKFITFSMLIFCAVW
IAFVPAYISSPG----KYTVAVEIFAILASSYGLLLCIFTPKCYIILLKPEKNTK
>O73639_FUGRU_PPR
DAIECARCLPEFWSNADRTACVPKQVEFLSFGDTIGIALLVVSLIGSFLTCAVALV
FFYHRTSPIVRANNSDLSFLLLFSLTLCFLCSLTFISPPSQWSCMLRHTAFGITFVLC

ISCILGKTIVVLMAFRATLPGSDVMKW-
FGPGKQKAIITFSTLVQVVICTVWLVVAPPTPRQYMPRESIII------
LLCDEGSTIAFSLVLGYIGVLACMCFLLAFLARKLPDNFNEARLIAFSMLIFCAVW
VAFVPAYISSPG----KYSTLTEIFAILASSYGLLGCIFAPKCYIILMKSEKNTR
>Q8NGV9_HUMAN_CS
GERECEQCGEDYWSNAQKSECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVT
AVYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLFIDKPHNWSCMAGQVTLAL
GFSLCLSCLLGKTSSLFLAYRISKSKTQLTSM--
HPLYRKIIVLISVLAEIGICTAYLILEPPMVYKNMESQNTKII-----
LGCNEISIEFLYSMFGIDAFLALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIIWM
SFVPVYLSTKG----KFKMAVEIFAILASSHGLLGCIFAPKCLIILLRPERNTS
>Q8TDU1_HUMAN_CS

DYWSNAQKSECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAVTAVYVIHRHTP
LVNASDWQLGFLIQVSLIIMLLSSMLFIDKPHNWSCMAGQVTLALGFSLCLSCLL
GKTSSLFLAYRISKSKTQLTSM--
HPLYRKIIVLISVLAEIGICTAYLILEPPMVYKNMESQNTKII-----
LGCNEISIEFLYSMFGIDAFLALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIIWM
SFVPVYLSTKG----KFKMAVEIFAILASSHGLLGCIFAPKCLIILLRPERNTS
>AAC08417.1_V2R

NWTCMARQVTLALGFCLCLSSILGKTISLFFAYRISKSKTRLISM--
HPIFRKLIVLICVVGEIGVCTAYLMLKPPRMVKNIEPQNVKII-----
FECNEGSIEFLCSIFAFDVLLALLCFLTTFVARKLPDNYYEGKCITFGMLVFFIVWI
SFVPAYLSTKG----KFKVAVEIFAILASSYGLLGCIFLPKCFIILLRPKRNTD
>AAC08416.1_V2R
GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYEEALGFTLVILSIFGALVVLAVTV
VYVIYRHTPLVKANDRELSFLIQMSLVITVLSSMLFIGKSCNWTCMARQVTLALG
FCLCLSSTLGKTISLFFAYRISKSKTRLISM--
SPIYRKLIVLICVVGEIGVCAAYLVLEPPRMFKNIEPQNVKII-----
FECNEGSIEFLCAIFGFDVFLALLCFLTTFVARKLPDNYYEGKCITFGMLVFFIVWI
SFVPAYLSTKG----KFKVAVEIFAILASSYGLLGCIFLPKCFIILLRPKRNTD
>O70410_MOUSE_V2R
GQRECDPCGEDDWSNAGKSKCVPKLVEFLAYGEALGFTLVILSIFGALVVLAVT
VVYVIHRHTPLVKANDRELSFLIQMSLVITVLSSLLFIGKPCNWSCMARQITLALG
FCLCLSSILGKTISLFFAYRISVSKTRLIS----
PIFRKLIVLVCVVGEIGVCAAYLVLEPPRMFKNIEIQNVKII-----
FECNEGSVEFLCSIFGFDVLRALLCFLTTFVARQLPDNYYEGKCITFGMLVFFIVW
ISFVPAYLSTKG----KFKVAVEIFAILASSYGLLGCLFLPKCFIILLRPKRNTD
>TS1R3_GORGO_TR
DDVTCTSCGQDEWSPERSTRCFHRRSRFLAWGEPAVLLLLLLLSLALGLVLAAL
GLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPAQCLAQQPLSHL
PLTGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALC
TWYLVAFPPEVVTDWHMLPTEAL-----
VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFI

TWVSFVPLLANVQV----
VLRPAVQMGALLLCVLGILAAFHLPRCYLLIRQPGLNTP
>TS1R3_HUMAN_TR
DDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALG
LFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLP
LTGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCT
WYLVAFPPEVVTDWHMLPTEAL-----
VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFI
TWVSFVPLLANVQV----
VLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTP
>TS1R3_PANTR_TR
DDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALG
LFIHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPL
TGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTW
YLVAFPPEVVTDWHMLPTEAL-----
VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFI
TWVSFVPLLANVQV----
VLRPAVQMGALLLCVLGILAAFHLPRCYLLMWQPGLNTP
>Q49KI5_FELCA_TR
DDLLCTQCDQDQWSPDRSTRCFARKPMFLAWGEPAVLLLLALLALALGLALAA
LGLFLWHSDSPLVQASGGPRACFGLACLGLVCLSVLLFPGQPGPASCLAQQPLFH
LPLTGCLSTFFLQAAEIFVGSELPPSWMRGR---
LRGPWAWLVVLLAMLAEAALCAWYLVAFPPEVVTDWRVLPTEAL-----
VHCHVHSWISFGLVHATNAMLAFLCFLGTFLVQSRPGRYNGARGLTFAMLAYFI
TWISFVPLFANVHV----AYQPAVQMGTILLCALGILATFHLPKCYLLLQRPELNTP
>Q49HH9_CANFA_TR
DDLLCTQCDQNQWSPDRSTRCFPRRLTFLAWGQPAVLVLLILLALALGLVLVAL
GLFIRHRDSPLVQASGGPRACFGLACLGLVCLSVLLFPGQPGPASCLAQQPLLHL
PLTGCLSTLFLQAAQIFVGSELPSSLRRCL-----
GPWAWLLVLLALLAEAALCAWYLVAFPPEVVTDWWVLPTQVL-----
VHCRMRSWISFGLLHAINAMLAFLCFLGTFLVQSRPGRYNGARGLTFAMLAYFI
TWISFVPLFANVHV----AYQPTVQMAAILLCALGILATFHLPKCYLLLQQLELNNP
>TS1R3_MOUSE_TR
DDFTCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAAL
GLSVHHWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHL
PLTGCLSTLFLQAAETFVESELPLSWANWL---
CSYLRAWLVVLLATFVEAALCAWYLIAFPPEVVTDWSVLPTEVL-----
EHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFI
TWVSFVPLLANVQV----
AYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQ
>TS1R3_RAT_TR
DDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLCLVLGLTLAALG
LFVHYWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLP
LTGCLSTLFLQAAEIFVESELPLSWANWLCSYLRGPWAWLVVLLATLVEAALCA
WYLMAFPPEVVTDWQVLPTEVL-----

EHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFII
WVSFVPLLANVQV----AYQPAVQMGAILFCALGILATFHLPKCYVLLWLPELNTQ
>Q2MHK9_BRARE_TR
DDIQCKSCPNGQWSTLRSTSCVYPIYTYLDWTNYESIGVILGGIVVLASHVWVGA
LFFKHRGTPLVKTAGGSLCGLTLLSLAGGCMSLVLFLGQPGDTVCRLQEPLNAFF
PTVALSVILSSSLQIVCVTEFPEQSSEHLEN-
LRGRGSWFVILGCCGLQAGLCGWYGLEGPSLTQYVASLDVTYVK---
TFLRCPVEPMLNFGLMLGFNVILALMSFMSTFMALKPPGQYNLARDITISTLSYC
VMWVMFIPIYTSLDD----
KNKSLAQVGVSLLSNMGLVAAYFFPKCHLLVKQPELNTD
>Q2MHJ8_FUGRU_TR
-
DIQCTPCPPRQWSLARSSRCTDPIYDYLSWDTPEALLLTLAIVLVVLFMGSVVVV
FVNHRETVLVTASGGTLSIVVLLGLMGACLSLLLFLGQPGDTVCRLQLPLISAFQT
VPLSIIMSISLQIFFVSEFPNLAASYLHV-
LRGPGTWLLLLTCCAVQAGICGWFVQDGPSLSEYLADRRVDFVR---
SFLACPVSP-
LSFALMQGFSAAMALMSFMCTFMATKPLHQYNLARDITFSSLIYCVIWVTFIPIYI
GLEE----KRRAIIHVSFILASDLGLVAVYYIPKCYFLLKTPELNTA
>Q2MHK2_ORYLA_TR
DDLQCTTCPDRQWSSIRSNSCIAPSFDVLEWGAPVSLYLIVAGVILLICQGSVFVL
FLMHRGTPVVLASGGALTFATLLSLMGACLSLLLFFGQPGDVVCHIQLPLISIFQT
VALSVILAVSLQLLLVTEFPKSVAPHLRT-
LRGPGSWVFVLICCLVQAGLCGWFVQDVPSLSEYKNNMTINFLH---
SFLACPVPS-
EKLAIMQGLNGSMALVSFMCTFMAVKPLHQYNLARDITFSTLIYCVTWVIFIPLY
TGFSTSQSTKKRSIAHVSFTLVGISALVLAYYLPKCFLLLRKPNLNTP
>TS1R1_HUMAN_TR
DLYRCQPCGKEEWAPEGST-
CFPRTVVFLALREHTSWVLLAANTLLLLLLLGTAGLFAWHLDTPVVRSAGGRLC
FLMLGSLAAGSGSLYGFFGEPTRPACLLRQALFALGFTIFLSCLTVRSFQLIIIFKFS
TKVPTFYHAWVQNHGAGLFVMISSAAQLLICLTWLVVWTPLPAREYQRFPHLV
M-----
LECTETNSLGFILAFLYNGLLSISAFACSYLGKDLPENYNEAKCVTFSLLFNFVSW
IAFFTTASVYDG----KYLPAANMMAGLSSLSSGFGGYFLPKCYVILCRPDLNST
>TS1R1_MOUSE_TR
NTSECQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANTLLLLLLIGTAGL
FAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQPLFSLGF
AIFLSCLTIRSFQLVIIFKFSTKVPTFYHTWAQNHGAGIFVIVSSTVHLFLCLTWLA
MWTPRPTREYQRFPHLVI-----
LECTEVNSVGFLVAFAHNILLSISTFVCSYLGKELPENYNEAKCVTFSLLLHFVSW
IAFFTMSSIYQG----SYLPAVNVLAGLATLSGGFSGYFLPKCYVILCRPELNTE
>Q3U5H1_MOUSE_TR
NTSECQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANTLLLLLLIGTAGL
FAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQPLFSLGF
AIFLSCLTIRSFQLVIIFKFSTKVPTFYHTWAQNHGAGIFVIVSSTVHLFLCLTWLA

MWTPRPTREYQRFPHLVI-----
LECTEVNSVGFLVAFAHNILLSISTFVCSYLGKELPENYNEAKCVTFSLLLHFVSW
IAFFTMSSIYQG----SYLPAVNVLAGLATLSGGFSGYFLPKCYVILCRPELNTE
>TS1R1_RAT_TR
NMSECQPCGTEEWAPKESTTCFPRTVEFLAWHEPISLVLIAANTLLLLLLVGTAG
LFAWHFHTPVVRSAGGRLCFLMLGSLVAGSCSFYSFFGEPTVPACLLRQPLFSLG
FAIFLSCLTIRSFQLVIIFKFSTKVPTFYRTWAQNHGAGLFVIVSSTVHLLICLTWL
VMWTPRPTREYQRFPHLVI-----
LECTEVNSVGFLLAFTHNILLSISTFVCSYLGKELPENYNEAKCVTFSLLLNFVSWI
AFFTMASIYQG----SYLPAVNVLAGLTTLSGGFSGYFLPKCYVILCRPELNTE
>Q2MHL2_BRARE_TR
GYTSCQPCPIDDWSEAESEVCLPRAELYLSWGAPLTTALIIYLAVTLFVTLGTTLV
FLLNLSTPVVKSAGGKTCLLMLVSMIVACCSTLCHFSRPSRVGCLLKQPLFVISCT
VCLACVTVRSFQVVCIFKWSSKLPRSYETWAKNRGPEMFILITTVVETFISVLRM
LLDPPFPSRDYDFYHDSIV-----
LECSKTLSLSAFAELFFVCALSLVCFCLSYMGKDLPANYNEAKCITFSLMIYMIS
WITFFTAYCISRG----SFVMALNVGAILLSVLGILGGYFLPKVYIILIKPQLNTT
>Q2MHK6_ORYLA_TR
DPTECQPCLPEQWAPPSSEQCLNRTVLLLAWDNPLSIALLFFLATCLLLTSGTAIIL
LVNLNTPVAKSAGGRTCLLMLAALTAAALSTLCHFGEPSPLACILKQPLFTVSFT
VCLACITMRSLQIICIFKFSSKLPPAYDKWVKKRGPEITILLVSVTILFISVLRVSLD
TPRPSQDLNFYPDKIV-----
LECSKTLSVGSGVELAYVSLLSVLCFSFSYMGKDLPANYNEAKCVTFSLLVYMIS
WISFFTLYLINRE----TFTMAAQVFATLFSVLAFIGGYFVPKIYIIVLKPKMNTT
>Q2MHK1_FUGRU_TR
DPTSCQDCLPEEWAPKSSERCLKRTPLLLEWDHHLSIALLFFLACCLLMTSSSAVI
LLLNINTPVAKSAGGRTCLLMLAALTAAAMSSLCHFGQPSPLACILKQPLFTFSFT
VCLACITVRSLQVVCIFKFASRLPPAYDKWSKKQGPEVTIFIVSVTILCISVLRVAV
GPPEPSQDLDFYMDSIV-----
LECSNTLSPGSFIELCYVCVLSVLCFFFSYMGKDLPANYNEAKCVTFSLMVYMIS
WISFFTVYLISRG----PFTVAAYVCATLVSVLAFFGGYFLPKIYIIVLKPQMNTT
>Q4S833_TETNG_TR
DPYSCLPCEKHQWSTAGSTSCLPRTVEWVPFTDPAAVVILAGACLLVALTAAMC
VLFAVNYNTPVVRSAGGPMCFLILGCLSLCSISVFFYFERPTEAFCILRFMPFLLFY
AVCLACFAVRSFQIVIIFKIAAKFPRVHSWWMKYHGQWLVISVAFVLQAVLIVIA
FSSDPPLPYSDFLSYPDKII-----LGCDVNLNMAS-
TSFFFASIVVHPLLYFLLHGERPPEEHNLLLASADPHLDHICHRDHALPRQIHPHG
RALQRLLSAPEHPEVLPGSDSGLHQNHQPRGAGQKLHPPASYK
>Q2MHK3_ORYLA_TR
DPYNCINCKETEWSTEGSTSCTIRLLEYIPFTDTAAVVIMVGALVLVALTVVMSV
LFALNYNTPVVRSAGGPMCFLILGCLSLCSLSVFFYFGEPTETYCTLRFLPFLLFYT
VCLACFVVRSFQIVCIFKISAKFPKLLSYWTKYHGHWLIITVAFFIQALLLITGYSI
APPKPYSDTFWYSDIIV-----LGCDLDLATSS--
SFILLSVLCCLSFIFSYMGKNLPKNYNEAKTITFCLLLLILTWITFATVSMLYHG----
KYIHAMNALAILSSLYSFLLWYFFPKCYIIIFQPKKNTQ
>Q2MHK0_FUGRU_TR

DQVSCHHCEEHQWSAAGSTSCLPRTVEYVAFTDTAAVVILVGAGLLLALTLAM
CVLFAINYNTPVVRSAGGPMCFLILGCLSLCSISIFFFFDKPTVAFCVLRFLPFVLF
YTVCLACFVVRSFQIVSIFKIAAKFPQAHSWWMKYHGQWLVISVTFVIQAFLIIVS
LSSDPPSPFRDIVSYPDKII-----LGCFMNLKTSS-
GSTVLLLSLGSLCFIFSYMGKDLPKNYNEAKAITFCLLLLILTWIIFTTASLLYQG---
-KYIHSLNALAVLSSIYSFLLWYFLPKCYIIIFQPQKNTQ
>Q2MHJ9_FUGRU_TR
DPFNCIDCEEHQWSAAGSTSCLPRTVEYVAFTDTAAVVILVGAGLLLALTLAMC
VLFAINYNTPVVRSAGGPMCFLILGCLSLCSISIFFFFDKPTVAFCVLRFLPFVLFYT
VCLACFVVRSFQIVSIFKIAAKFPQAHSWWMKYHGQWLVISVTFVIQAFLIIVSVS
SDPPSPFRDIVSYPDKII-----LGCFMNLKTSS-
VSFVLLLLLCLLCFIFSYMGKDLPKNYNEAKAITFCLLLLILTWIIFTTASLLYQG---
-KYIHSLNALAVLSSIYSFLLWYFLPKCYIIIFQPQKNTQ
>Q4S836_TETNG_TR
DPYSCLPCEKHQWSTAGSTSCLPRTVEWVPFTDPAAVVILAGACLLVALTAAMC
VLFAVNYNTPVVRSAGGPMCFLILGCLSLCSISVFFYFERPTEAFCILRFMPFLLFY
AVCLACFAVRSFQIVIIFKIAAKFPRVHSWWMKYHGQWLVISMTFVLQAVVIVIG
FSSDPPLPYSNFLSYPDKII-----LGCDVNLNMAS-
TSFFLLLLLCILCFTFSYMGKDLPKNYNEAKAITFCLLLLILTWIIFATAFMLYHG--
--KYIHTLNALAVLSSAYCFLLWYFLPKCYIIIFQPHKNTQ
>Q2MHK4_ORYLA_TR
NPYKCVLCKSTEWSEEGSTSCNLRQLEYIQLTDTVAIAIMVTTWVIVGLTLSVSV
LFSINYNTPVVRSAGGPMCFLILGCLGLCSISVFFHFGLPTPASCILRYFPFIWFYTV
CLACFVVRSFQIVCIFRISAKFPSLLIWWKKYHGQWLLISAAFVVQAVLLLTFSYD
-PPRPYNETSWRPEQIV-----LFCYMNL-
KATSCSLVLPISFACLCFVFSYMAKDLPKNYNEAKAITFSLLLLTFTWTIFITVYM
LYRG----KYIQTLNAVTVLLSLYYFLVGYFLPKCYIILFQPNKNTQ
>Q2MHK5_ORYLA_TR
DPFKCISCKDTEWAAEGSTSCLPRLLEYVSLTDIGAVLTMVLEGLLLGLTVSVSV
VFAINYNTPVVRSAGGPICFLNLGCLSLCSISVFFYFGVPTPSSCVLRYLPFILFYTV
CLACFVVRSFQIVCIFKISAQIPNIHIWWVKYHGQWLLISAAFMFQALSLTISFCCD
PPRPYNEMLWYPEQII-----LSCNLNI-
QASCFSLFLLVSLCCLCFIFSYMAKDLPKNYNEAKAITFCLLLLICTWVIFITEYML
YRG----KYIQIFNALAVLSSLYSFLLWYFLPKCYIIIFKPDKNTQ
>Q2MHL1_BRARE_TR
DPYTCFPCAVSDWSDEGSTACKTRAVVYLEFTEITSIAVMISVLFLIILLIGILGLFT
YNFNTPVVRSAGGDMCLLMLLCLTISSISVFFFFGKPSSVHCLVRNAIFAFFFTVC
LSCLTVRSFQIICVFKMAAQFPRLHSLWVKHNGQWLFIAFSSFIHLISCVIWTTVSP
PIA--DSWTFKDQTL-----
LMCEMVNTITFTVVLFISWFLGFLCLVFSYMGRDLPKNYNEAKSITFSLILYYLS
WIVYFTAYLSIKS----KYIMLLNAMAQISSIYGILFSYFIPKSYIMIFQPQKNTA
>TS1R2_HUMAN_TR
DEYECQACPNNEWSYQSETSCFKRQLVFLEWHEAPTIAVALLAALGFLSTLAILV
IFWRHFQTPIVRSAGGPMCFLMLTLLLVAYMVVPVYVGPPKVSTCLCRQALFPL
CFTICISCIAVRSFQIVCAFKMASRFPRAYSYWVRYQGPYVSMAFITVLKMVIVVI
GMLA-SPTTRTDPDDPKITI------

VSCNPNYRNSLLFNTSLDLLLSVVGFSFAYMGKELPTNYNEAKFITLSMTFYFTSS
VSLCTFMSAYSG----VLVTIVDLLVTVLNLLAISLGYFGPKCYMILFYPERNTP
>TS1R2_RAT_TR
DEFNCLSCPGSMWSYKNDITCFQRRPTFLEWHEVPTIVVAILAALGFFSTLAILFIF
WRHFQTPMVRSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQAFFTVCF
SICLSCITVRSFQIVCVFKMARRLPSAYSFWMRYHGPYVFVAFITAIKVALVVGN
MLA-NPIGRTDPDDPNIMI------
LSCHPNYRNGLLFNTSMDLLLSVLGFSFAYMGKELPTNYNEAKFITLSMTFSFTS
SISLCTFMSVHDG----VLVTIMDLLVTVLNFLAIGLGYFGPKCYMILFYPERNTS
>TS1R2_MOUSE_TR
DEFNCLSCPGSMWSYKNNIACFKRRLAFLEWHEVPTIVVTILAALGFISTLAILLIF
WRHFQTPMVRSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQAFFTVCF
SVCLSCITVRSFQIVCVFKMARRLPSAYGFWMRYHGPYVFVAFITAVKVALVAG
NMLA-NPIGRTDPDDPNIII------
LSCHPNYRNGLLFNTSMDLLLSVLGFSFAYVGKELPTNYNEAKFITLSMTFSFTSS
ISLCTFMSVHDG----VLVTIMDLLVTVLNFLAIGLGYFGPKCYMILFYPERNTS
>Q49HI0_CANFA_TR
DEFDCQPCPSYEWSHRNDTSCFKRRLAFLEWHEPSTIFVVMLTILGFLSTLAIMVI
FWRHLHTPVVRSAGGPMCFLMLVPLLLAYAMVPMYIGQPTFFSCLWRQTFFTLC
FTICISCITVRSFQIVCIFKMARRLPRAYGYWVRCHGPYVFVASFMVLKVVIVAG
NVLA-NPTARPDPDDPNIMV------LSCN--
YRRALLFNTSLDLLLSVAGFSFAYMGKELPTNYNEAKFITLCMTFYFTSSVSLCTF
MSVYDG----VLVTILDLLITVLNLLGISFGYFGPKCYMVLFYPERNTQ

Appendix D - 500 genes used to reconstruct the Neighbor-Joining Tree

>E_2R
DIPIIKANNRSLSYILCFLCSLLFTDHPTTVTCLLRQTTFGVVFTVAASTILAKTITV
VLAFKVRRPGGRVRYYIGPRASNSILTCSLVQVILCGIRLVSSLLFPDTDSYSEPGF
IIVQCNEGSASALYCVLGYMDFLALASFSMAFVARNLPDTFNEAKFCSVWVSFLP
TYHSTKGKAMVAVEIFSILASSLGLL
>G_10R
DMKSCLQCPEEQYSSQSKDYCLPRTETFLAFDEPLGLVLTVVSLSLASLACLVLG
VFLWFQQTPVVRANNRALSYVLLVSLTLCALSTLLFLGRPSPTTCLLRQITFAVVF
TVAVSCVLAKTLTVVLAFRATKPGQRVHMCLGPSASSFVVLVASLVQIILCGIWL
GTSPPFPHRDTASEPGHLVLQCHEGSGIAFSCVLGYLGLLAIGTFSVAFLARGLPD
TFNETKFLTFSMLLFCSVWTAFLPLYFSARGKATVAVEVFSILASTAGLLGGIFFP
KCYLILLKPERN
>G_11R
EQCIKCPEHHYANAQQNHCLQKSVTFLAYEDPVGKALAGTALILTVLTAAVLGIF
VKHQDTPIVKANNQALSYTLLISLIFCFLCSLLFIGHPNTATCILQQTTFGVVFTVA
VSTVLAKTITVMLAFKVTTPGRRMRQFLVSGAPNYIIPICTLIQFTLCAVWMGTN
PPYVDTDAHSEHGHVIIMCNKGSLTAFYCVLGYLSSLALVSFTVAFLVRNLPDTF
NEAKFLTFSMLVFCSVWVTFFPVYHSRRGKVMVAMEVFSILASSAGLLGCIFVPK
CYIILLRPEKNS
>G_12R

DMEQCVKCSDDHYATIQQDLCLQKSVTFLAHEDLMGKMLAGTALSLTVLTAAV
LGLFVKHQDTPIVKANNRALSYTLLISLIFCFLCPLLFIGQPNTATCILQQTTFAVV
FTVAVSTVLAKTITVVLAFKVTAPGRRMRQLLVSGAPNYIIPICTMIQLTLCGVW
MGTNPPYIDTDAHSEHGHIVIMCNKGSLAAFYCVLGYLGSLAMLSFTVAFLVRN
LPDSFNEAKFLTFSMLVFCSVWVTFLPMYHSSKGKVIVAMEVFSILVSGAGLLGC
IFAPKCYIILLRPERNSL
>G_13R
CMKCPDHQYASIQRNQCLPKAVTFLAYEDSMGMTLAGTALSLTSLTAAVLVLF
VKHQDTPIVKANNRALSYTLLISLMFCFLCSLLFIGCPNTISCILQQTTFGVVFTVS
VSTVLAKTVTVMLAFKFTVPGSRMRQLLVSGAPNYIIPTCTLIQLTICGVWMGTS
PPFVDTNAHSEHGHIIIVCNKGSLTAFYFILGYLGSLSLISFIVAFMARNLPDAFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVTMEVFCILASSSGLLGCIFFPKCY
IILLKPERNS
>G_14R
CMKCPDHQYASIQRNQCLPKAVTFLAYEDSMGMTLTGTALSLTSLTAAVLVLFV
KHQDTPIVKANNPALSYTLLISLMFCFLCSLLFIGCPNTISCILQQTTFGVVFTVSV
STVLAKTVTVMLAFKFTVPGSRMRQLLVSGAPNYIIPTCTLIQLTICGVWMGTSP
PFVDTNAHSEHGHIIIVCNKGSLTAFYFILGYLGSLSLISFIVAFMARNLPDAFNEA
KFLTFSMLVFCSVWVTFLPVYHSSKGKVMVTMEVFCILASSSGLLGCIFFPKCYII
LLKPERNS
>G_16R
EQCVMCPDQEYANIQRTRCLQKSVIFLSYEDPLGKTLAGIALSFTIITSAVLGLFV
KHRDTPIAKANNRALSYTLLISLTFCFLCTLLFIGHPNTATCILQHTTFAVVFTVAV
STVLAKTITVVLAFKVIAPQRSMRQLLVSGAPNSIIPICSLIQLSICGVWMGTNPPY
LDTDAHSEHGHIIIMCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSMLASSAGLLVCIFAPKCF
IILFRPERNSL
>G_18R
DMKSCLQCPEKQYPSQTRDHCIPRTEIFLTFDEPLGLALTLVSLALASLACLVFGV
FLWFRHTPVVRANNRALSYVLLVSLTLCALSPLLFLGRPSPTTCLLRQTTFAVVF
TVAVSCVLAKTLTVDLAFRATRPGQRVRMCLGNSASSSVVLVSSLMQIILCGVW
LGTSPPFPHRDTASEPGHLVLQCQEGSGMAFSCMLGYLGLLAIGTFSVAFLARDL
PDAFNEAKFLTFSMLLFCSVWTAFLPLYHSARGKATVAVEVFSILASTAGLLGGI
FLPKCYLILLKPEQN
>G_19R
EQCVKCLGHQYANTQRNQCLHKIVTFLAYEDPLGKALAGTALSLTALTAAVLGL
FVKHQDTAIVKANNRTLSYTLLISLLSCFLCSLLFIGHPNTATCILQQTIFAVAFTV
AVSTVLAKTITVVLAFKITAPDKRMRWLLVSGTPNFIIPICSLIQLTLYAVWMGTN
PPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALVSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVPK
CYIILLRPERNS
>G_20R
EQCVKCLGHQYANTQRNQCLHKIVTFLAYEDPLGKALAGTALSLTALTAAVLGL
FVKHQDTAIVKANNRTLSYILLISLLSCFLCSLLFIGQPNTATCILQQTIFAVAFTVA
VSTVLAKTITVVLAFKITAPAKRIRRFLVSGTPNFIIPICSLIQLTLYAVWMSTNPPY
IDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNEAK

FLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVPKCYII
LLRPERNS
>G_21R
EQCVKCLGHQYANTQRNQCLHKTVTFLAYEDPLGKALAGTALSLTALTAAVLG
LFVKHQNTAIVKANNRTLSYTLLISLLSCFLCSLLFIGQPNTATCTLQQTTFALAFT
VAVSTVLAKTITVVLAFKITAPGKRMRRLLVSGTPNFIIPICSLIQLTLCAVWMGT
NTPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLLFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVP
KCYIILLRPERNS
>G_22R
EQCVKCLGHQYANTQRNQCLHKTVTFLAYEDPLGKALAGTALSLTALTTAVLG
LFVKHQHTAIVKANNRTLSYTLLISLLSCFLCSLLFIGQPNTATCILQQTIFAVAFT
VAVSTVLAKTITVVLAFKITAPGKRMRRLLVSGTPNFIIPICSLIQLTLYAVWMGT
NPPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVP
KCYIILLRPERNS
>G_25R
EQCSQCPKHEYPNRNRDRCLPKTVDCLSFEDPLGMALNLIALCFSVLTALILGIFV
KHRDTAIVKANNRTLSYTLLISLLLCFLCSLLFIGRPHRITCILQQMTFGLVFTVAV
SSVLAKTITVVLAFKVTAPGRRMRNLLISGAPNYIIPICTLIQLTLCGIWMGTNPPY
IDTDVHSEHGHIIIVCNKGSLTAFYCVLGYLGFLALLSFTVAFLARNLPDTFNEAK
FLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFAPKCYIIL
TRSANNSL
>G_29R
EQCVKCPDDQYATIQQDHCLQKSVIFLAHEDLMGKILTGTALSLTLLTAAVLGLF
VKHQDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPNTATCILQQTTFAIVFTVA
VSTVLTKTITVVLAFKITAPGRRMRHLLVSGAPNYIIPICTLVQLTLCGIWMGTNP
PYLDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLAKNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEIFSILVSSAGLLGFIFAPKCYI
ILLRPERNS
>G_30R
EQCTQCPEHEYTNRNRDRCLPKILDFLSFEDSLGMALTHIALCFSVLTAVILAIFV
KHRDTAIVKANNRTLSYTLLISLLLCFLCCLLFIGRPNTATCILQQMTFGLVFTVA
VSCVLAKTITVVLAFKVTGPEKRMRQLLLLGAPNYIIPICTLIQLTLCGFWMGTNP
PYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASGVGLLGCIFLPKC
YVILIRSENNT
>G_32R
DMEQCMKCLDHQYANIQRNHCLQKAVIFLAYEDPLGKVLAGTALSLPVLTAAV
LGLFVKHRDTPIVKANNRALSYTLLISLTFCFLCTLLFIGRPSTASCILQWTTFAVV
FTVTVSTVLAKTITVVLAFKVTGRRMRHLLVSGAPNYVIPICTLIQLTLCGIWMGT
NPPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASGAGLLGCIFAP
KCYIILLRPQGNS
>G_34R

IATLGLLVKHRDIPNVKANNRALSYTLLFSLTFCFLCSLLSIGRPSTDTCILQQTTF
AVVFTVVVSTVLAKTITVVLAFKVTAPGRRMRQLLVSGAPNYIIPICTLIQLTLCA
AWMGTNSPYSDTDPHSEHGHIIIMCNKGSLTAFYCVLGYLGSLALLSFTVTFLAR
NLPDTFNEANFLTFSMLVFCSVWVTFLPMYHSSKGKVMVAIEVFSILASGAGLLG
CIFAPKCYIILLRPQGN
>G_36R
EQCVKCLGHQYANTQRNQCLHKIVTFLAYEDPLGKALAGTALSLTALTAAVLGL
FVKHQDTAIVKANNRTLSYILLISLLSCFLCSLLFIGQPNTATCILQQTTFGLAFTV
AVSTVLAKTITVVLAFKITAPGKRMRRLLVSGKPNFIIPICSLIQLTLCGFWMGTN
PPYIDTDAHAEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVPK
CYIILLRPERNS
>G_37R
DMEQCAKCLGHQYANTQRNHCLHKTVTFLAYEDPLGKVLAGTALSLTVLTAAV
LWLFVKHQDTPIVKANNQALSYTLLISLLSCFLCSLLFIGHPNTTTCILQQTIFAVA
FTVAVSTVLAKTITVVLAFKITAPGKRMRQLLVSRTPNSIIPICSLIQLTLCGIWMG
TNPPYIDRDTHSVHGHIIIVCNKGSLTAFYCVLGYLGSLALVSFTVAFLARNLPDT
FNEAKFLTFSMLVFCGVWVTFLPVYHSSKGKEMVAMEVFSILASGAGLLVCIFV
PKCYIILLRPERNSI
>G_40R
RANNRNLSYVLLISLILCFFCSLMFIGLPSPVTCILRQTFFGVVFSIAVSAILAKTFIV
LVAFKSIKPGNSVPMWMVTRFSNVIVCSCSLTQVCICTVWLTTSPPFPDTDTQSEF
GQIILQCNEGSTVAFYCVLGYLGFLAFLSLVLAFLARTLPDRFNESKMITFSMLVF
CSVWISFIPSYISTKGKTMVALEIFAILSSSAGLLGCIFLPKCYVILLRSDNASR
>G_42R
QCMKCPDHQYANTQRNHCLQKSVTFLAYEDSLGKALTCTALSLAILTAFVLGLF
VKYQNTPIVIANNHTLSYILLISYTFSFLCSFLFIGYPNTATCIFQQTTFRIVFTVAVS
TALVKTITVVLAFHVTAPGKGMRQQLISRAPNYIILICSLIQLTLCGIWVGTNPPFI
DTDTHSEQGHIIIICNKGSIIAFYCVLGYLGTLAIVS
>G_45R
DMEQCMKCPDHQYANTQRNHCLQKSVTFLAYEDSLGKALTCTALSLAILTAFVL
GLFVKYRNTPIVKANNRTLSYILLISLTFCFLCSLLFIGRPNTATCILQQTTFGIVFT
VAVSTVLAKTITVVLAFKVTAPGKGMRQLWISGAPNYIILICSFIQLTLCGIWVGT
NPPFIDTDTHSEQGHIIIICNKGSITAFYYVLGYLGSLALVSFSVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAMEVFSILASSVGLLSCIFAPKCY
IILLRPDRNYL
>G_5R
DIEQCVMCPEQEYANIQRTRCLQKSVMFLAYEDPLGKTLAGIALSFTIITSAVLGL
FVKHRDTPIVKANNRALSYTLLISLTFCFLCTLLFIGHPNTATCILQHTTFAVVFTV
AVSTVLAKTITVVLAFKVTAPGRRMRQMLILGAPNSIIPICSLVQLTLCGVWMGT
NPPYIDKDAHSEHGHIIIMCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWITFLPVYHSSKGKVMVAMEVFSMLASGAGLLVCIFAP
KCFIILFRPERN
>G_8R
DMEECVKCPDHQYADIHHKHCVQKSVTFLAYEDSLGRALAGTALSLTILTVAIL
GLFVKHCSSPIVKANNRSLSYILLLSLSFCFLCALLFIGRPNTATCILQQVTFGIVFT

VAVSTILAKTITVVLAFKVTTPDRRIRQFLVSGTPNCIIPICSLIQFTLCGIWMGTSP
PFIDRDAHFEHGFIIIVCNKGSLTAFYCMLGYLGLLALVSFTVAFLARNLPDTFNE
AKFITFNMLVFCSVWVTFLPVYHSTKGRVMVAMEVFSILASSAGLLGCIFAPKCY
IILFKKDRCSL
>G_9R
DMKSCLECPEEQYPSQTRDHCIPRTENFLAFDEPLGLALTSVSLALASLACLVFG
VFLWFRHTPVVRANNRALSYVLLVSLTLCALSPLLFLGRPSPTTCLLRQTTFAVV
FTVAVSCVLAKTLTVVLAFRATRPGQRVRMCLGNSASSSVVLVSSLMQIILCGV
WLGTSPPFPHRDTASEPGHLVLQCQEGSGVAFSCMLGYLGLLAIGTFSVAFLARD
LPDAFNEAKFLTFSMLLFCSVWTAFLPLYHSARGKATVAVEVFSILASTAGLLGG
IFLPKCYLILLKPEQN
>G_P1R
DMEH*VMCAD*KYTNIQKNHCLQKSVIFLAY*RPLGEETGGHHSESHCPHICCPC
ALCEALRHSHCQGSQLYTANLPHPLFPLHLALHRPNTASCIRQQNIFPDVFTVAV
STVLAQTISVVLAFKVTAQGRRMRQLLVPGA*NYIIPICSLVQLTLCGVWMGTNP
PYIDNDLHSEYGHIIIMCNKDSLTALHCVLAYLGSLAQLCFNVAFRARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPMYHSNKRKVMVAMEVFSILAPSAGLLGCIFVHQ
CLIILLQTGRN
>G_P2R
DMEQCVKCPDDQYATIQQDHCLQKLVTFLAYENLMGKMLAGTALSITVLTVAV
LVLFVKN*DTPIVKTNNQALSYTLLISLTFCFLCSLLFIGHPNTATCILQQTTFAVV
FTVAVSNVLTKTITVVLAVKVTAPGRRMRLLLVSRAPNYIIPIFTLVQLTLCGVW
MGTNPSYIDVYAHSEPGHIIIMCNKGSLIDFYCVLGYLGSSALVSYMVAFIARNLP
DTFSEAKFLTFSMLVFCSVWITFLPVYHSSKGNVMVAMEVFSILASGAGMLDCIF
VPKCYIILLRPERNS
>G_P5R
LRLFVKHRDTPIVKANNQALSYTLLISLTFCFPCSLFFIGCPSTASFILQWTTIAVVF
TVAISTVLAKTITVVLACKVTAPGRRMRQLLVSGPPIPSSPSAP*LNSLCAIWMGT
NAPYIDTDAHSEHSHIIVVCNKGSLTAFYYVLGYLGSLALFSFILVFLVRNLPDTF
NEAKYLTFSMLVFCGVWVPFLPVYHSSKGKVMVALEVFSILASSAGLLGCIFVPK
CFIVLFRSERN
>H_1R
THWAPVRSTVCFEKEVEYLSWNDSLAILLMALSLLGIMFVLAIGIIFTRNLNTPVV
KSSGGLTVCYVILLCHFLNFANTGFFVGEPQDFTCKTRQTLFGVSFTLCISCILMK
SLKILLAFSFDPKLQNFLKCLYKPIPIIFTCTGIQVVICTLWVILAAPTVEENVSLPR
VIILECEEGSVLAFGTMLGYIAILAFICFIFAFKGRKLPENYNEAKFITFGMLIYFIA
WIIFIPVYATTFGKYLPAVEIIVILISNYGILCCMFFPKCYVILCKQETNTK
>R_1R
THWAPVRSTMCFEKEVEYLSWSDSLAILLLAVSLLGILFVLAISIIFTRNLNTPVVK
SSGGLMVCYVILLCHLLIFASTGFFIGEPQDFTCKTRQTLFGVSFTLCVSCILTKSL
KILLAFSFDPKLQNFLKGLYKPIPIVLICTGIQVIICMLWLVFAAPAVEENVSLPRVI
ILECEEGSILAFGTMLGYIAILAFICFIFAFKGRKLPENYNEAKFITFGMLIYFIAWV
TFIPVYATTFGKYLPAVEIIVILISNYGILCCTFFPKCYVILCKQETNTK
>R_10R
DQCVNCPDLQYANTERDQCFHKKVTFLSFEDPLGMTLVCTALCFSILTVVFLGVF
VKHRDTPIVKANNRCLSYILLIALTFCFLCSLLFIGHPNTTTCVLQQTTFGVAFTA

AVSTILAKTITVLLAFKVTAPGRRMRHWLLSGVPNSIIPICSLIQLALCAIWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPKCY
IILLRPEKNA
>R_12R
EQCIQCAEHEYPNMERSRCLPKLMTFLAFEDTLGMALACMALCFSVLTAVVLW
VFVKHRDTPIVKANNRSLSYILLIALLLCFLCSLLFIGRPNTATCLLRQIAFGLVFT
VAISAVLAKTITVILAFKVMKPGRTIRRLLVSGASNFIIPICSLIQLTFCGIWLGISPP
FLELDAYSEHGHIILVCNKGSVTAFYCVLGYLGSLALFSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCY
IILIRPEKNSL
>R_13R
DQCMKCLDHQYASTERNQCFHKKVTFLSFEDPLGMTLVCTALCFSVLTAVVLG
VFVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFT
VAVSTILAKTITVVLAFKVTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLGTS
PPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALLSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKC
YIILFRPEKNA
>R_14R
EQCIPCADHEYPNLERNHCLPKLVTFLGFEEFLGMALTCMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPKTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRRLLVSGASNSVIPICFLIQLALCGIWLGSSPP
FIDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFFPKCY
IILLRPERNS
>R_16R
DVEQCVRCLDDQYPNPARDRCLPKVVVFLAYEDPLGAALAGVALCLSVLTALV
LGVFVKHRNTPLVKANNRNLSYTLLASLCLCFLCSLLFLGRPNTATCILQQTMFA
IVFTVAVSTILAKTITVVLAFKATAPGKRMRQWLVSGAPNSVIPICSLLQVMLCGI
WLGTCPPFIDTDTHAEPQRLLVLCNKGSATLFYLVLGYLGSLALGSFSVAFLARR
LPDTFNEAKYLTFSMLVFCSVWVTFIPVYHSTKGKVMVAVEVFSILASSAGLLGC
IFAPKCYILLVRPNRNV
>R_18R
YPNPARDRCLPKVVVFLAYEDPLGAALAGLALGFSVLTALVLVVFVKHRNTPLV
KANNHNLSYTLLTSLCLCFLCSLLFLGRPNTATCILQQTTFAIVFTVAVSTVLAKTI
TVVLAFKATAPGKRMRQWLVSGASNSVIPICSLLQVMLCGIWLGTCPPFIDTDTH
AEPQQLLVLCNKGSATLFYLVLGYL
>R_22R
DQCVNCPDLQYANTERDQCFHKKVTFLSFEDPLGMTLVCTALCFSVLTVVFLGV
FVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTL
AVSTVLAKTITVLLAFKVTVPGRRMRQWLLSGVPNSIIPICSLIQLALCGIWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDKFNE
AKFLTFSMLVFCSVWVTFLPVYQSTKGKVMVAVEVFSILCSSAGLLGCIFVPKCY
IILFRPEKNA
>R_23R

DQCMNCPDLQYANRERDQCFYKKLTFLSFEEPLGMTLVCTAVCFSVLTAVILGV
FVKHRDTPIVKANNRCLSYILLIALIFCFLCSLVFIGRPNTTTCVLQQTTFGMVFTL
AVSTVLAKTITVLLAFKVTAPGRRMRHWLLSGVPNSIIPICSLIQLALCGLWLGTS
PPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMMAVEVFSILASGAGLLGCIFVPK
CYIILFRPEKNA
>R_24R
QCMQCSDDEYPNQERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAVLWV
FVKH*DTPIVKANNRTLSYILLITLLLCFLCSLLFVGRPSTATCLLQQIAFGLVFTV
ALSTVLTKTITVILAFKVTKPGRTMRRLLVSGVYNSVIPICSLIQLALCGIWLGTYP
PFIELDTHSELGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFVPKCYI
ILFRPEKNA
>R_28R
DQCMKCPDHQMPPRSETSAYTNKVNFLFFEDPLGMTLVCTALCFSVLTTVVLGV
FVKHRDTPIVKANSRCLSYILLIALTFCFLCSLLFIGCPNTTTCVLQQTMLGLVFTV
AVSTVLAKTITVLLAFKVTAPGRRMRQWLISGAPNSIIPICSLIQLALCAIWLGTSP
PFIDTDAHSEHGHTILVCKKGSVTAFYCVLGFLGSLALVSFTVAFLARNLPDKFSE
AKFLTFSMLVFCSVWVTFLPVYQSTKGKVMMAVEVFSILASSVGLLGCIFVPKC
YIIIFRPEKNA
>R_29R
DQCVKCPDLQYANTERNKCFHKRVTFLSFEDALGMILVCTALCFSVLTVVVLGV
FVKYRDTPIVKANNRCLSYILLITLIFCFLCSLLFIGRPNKTTCVLQQTTFGVVFTV
AISTVLAKTITVLLAFRVTGPGRRMRHWLISGLPNSIIPICCLIQLALCGFWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTALYCVLGFLGSLALASFTVAFLARNLPDKFNE
AKFLTFSMLVFISVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPKCYI
ILLRPEKNA
>R_30R
DQCVKCPDLQYANTERDQCYHKRVTFLSYEEPLGMTLVCTALCFSVLTAVVLG
VFVKHRDTPIVKANNRCLTYILLIALIFCFLCSLLFIGHPNTTTCILQQTTFGVVFT
MAVSTVLAKTITVLLAFKVTAPGRRMRHWLLSGVPNSIIPICSLIQLALCAIWLGT
SPPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPK
CYIILLRPEKNA
>R_31R
DKLVKCPNHQYANKE*EQCFHKKVTFLSFEDPLGMTLVCTALCFSVLTAVVLGV
FVKHRDTPIVMANNRCLTCILLIALIFCLLCLLFTGCPNTTTSVLQQTTLGLVFTM
AVSTVLAKIITALLAFKVTAPGRRMKHCLISGVPNSIIPICFLIQLSLCAIWLGTSPP
FINTDAHSEHGHIILVCKKGSVTAFYCVLGFLFSLALASTTVAFLARNLPDTFNEA
KLLTFSMLVFCSVLVTFLPVYQSTKGKVMVAVEVFSILASSAGLLGCIFVPKCYII
LFRPEKNA
>R_32R
DQCVKCPHLQYANTERDQCFYKKLTFLSFEEPLGMTLVCTALCFSVLTAVVLGV
FLKHRDTPIVKANNRCLSYILLITLIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVLLAFNFTAPGRRMRHWLISGAPNSIIPICSLIQLALCAIWLGTSP
PFVDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDKFN

EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPKC
YIILFRPEKNA
>R_33R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEEFLGMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTI
AVSAVLAKTITVILAFKALKPGRTIRRLLVSGTSNFIIPICSMIQLVLCGIWLGTSPP
FIEVDKHSEHDYIILVCNKGSVTAFYCVLGYLGSLALASFIVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYV
ILIKPDKNS
>R_34R
EQCIPCADHEYPNMERNHCLPKLVTFLGFEEFLGMALTCMALCFSVLTAAVLWV
FVKHQDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKALRPGRTMRRLLVIGASNFIIPICSLIQLALCGIWLGISPPF
IELDAHSEPDHIILVCNKGSVTAFYCVLGYVVSLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYV
ILIKPDKNSL
>R_36R
QCVKCLDHQYANTEHDRCLHKAVDFLAYDDPLGMTLVCTALCFSAITAVVLGV
FVKHQDTPIVKANNRCLSYILLINLIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVVLAFKVTAPGRRMRQWLISGVPNSIIPICSLIQLTLCAIWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALLSFTAAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCY
IILFRPEKNA
>R_38R
DQCMKCLDHQYASTERNQCFHKKVTFLSFEDPLGMTLVCTALCFSVLTAVVLG
VFVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFT
VAVSTILAKTITVVLAFKVTAPGRRMRHWLLSGVPNSIIPICSLIQLALCAIWLGTS
PPFIDTDAHSEHGHIILVCNKGSITAFYCVLGYLGSLALLSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCY
IILFRPEKNA
>R_39R
CVLQQTTFGVVFTMAVSTVLAKTITVLLAFKVTAPGRRMRHWLLSRAPNSIIPIC
SLILLAPCGFCRGTSPPFVDTDAHSEHGHIILVCNKDSETAFYCVLGYLGSIATASF
TVAFLARNLPDTFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVQVFSIL
ASSSGLLGCIFVPKCYIILFRPEKNA
>R_40R
DQCVNCPDLQYANTERDQCFHKKVTFLSFEDPLGMTLVCTALCFSILTVVFLGM
FVKHRDTPIVKANNRCLSYILLIALTFCFLCSLLFIGHPNTTTCVLQQTMFGVAFT
AAVSTILAKTITVLLAFKFTAPGRRMRHWLLSGVPNSIIPICSLIQLALCAIWLGTS
PPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPKC
YIILFRPEKNA
>R_41R
FTMAVPTILAKTITVLLAFKVTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLG
TSPPFIDTDAHSEHGHTILVCNKGSVTAFYCVLGYLGSLALLSFTVAFLARNLPDT

FNEAKFLTFSMLVFCSVWVTFLPVYYSTKGKVMVPMEVFSILASGAGLLGCIFVP
KCYIILFRPEKNA
>R_42R
DQCVKCPDLQYANTERDQCFYKKLTFLSFEDPLGMTLVCTALCFSVLTAIVLGV
FLKHRDTPIVKANNRCLSYILLITLIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVLLAFNFTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLGTSP
PFVDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDKFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKC
YIILFRPEKNA
>R_43R
FTMAVSTVLAKTITVVLAFKVTAPGRRMRQWLLSGVLNSIIPICSLIQLALCAIWL
GTSPPFIDSDAHSEHSHIILVCNKGSVTAFYCVLGYLVSLALLSFTVAFLARNLPY
TFSEAKFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIF
VPKCYIILFRPEKNA
>R_45R
EQCIQCAEHEYPNMERSHCLPKMMTFLAFEDALGMALVCMALCFSVTTAAVLG
VFVKHRDTPIVKANNRSLSYILLIALLLCFLCSLLFIGRPNTATCLLRQITFGLVFT
VAISAVLAKTITVILAFKVMKPGRTIRRLLVSGASNFIIPICSLIQLALCGIWLGISPP
FIELDAYAQHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCY
IILIRPEKNS
>R_46R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAILWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQITFGLVFTV
AVSTVLAKT
ITVILAFKVTKPGRTIRRLLVSGASNSVIPICFLIQLALCGIWLGTSPPFIDTDAHSE
HGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNEAKFLTFSML
VFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFLPKCYIILLRPERNS
>R_47R
EQCIQCAEHEYPNMERSHCLPKLMTFLAFEDALGMALVCMALCFSVTTAAVLG
VFVKHRDTPIVKANNRSLSYILLIALLLCFLCSLLFIGRPNTATCLLRQITFGLVFT
VAISAVLAKTITVILAFKVMKPGRTIRRLLVSGASNFIIPICSLIQLALCGIWLGISPP
FIELDAYAEPDHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYI
ILIRPEKNS
>R_48R
DQCVKCPDLQYANTERNQCFHKRVTFLSFEDALGMILVCTALCFSVLTVVVLGV
FVKHRDTPIVKANNRYLSYILLIALIFCFLCSLLFIGRPNTATCVLQQTTFGVVFTV
AISTVLAKTITVLLAFRVTGPGRRMRQCLISGLPNSIIPICCLIQLALCGFWLGTSPP
FIDTDAHSEHGHIILVCNKGSVTALYCVLGFLGSLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWLTFLPVYHSTKGKVMVAVEVFSILCSSVALLGCIFVPKCYII
LLRPEKNA
>R_49R
DDPGLHSSVLLYAHSCFSWYFVKHRDTPIVKANNLCLSYILLIALTFCFLCSLLFIG
HPNTTTCVLQQTKFGVAFTAAVSTILAKTITVLLAFKVTAPGRRMRHWLISGVPK
SIIPICSLIQLALCAIWLGTSPPFIDTDAHSEQCHIILVCNKGSVTAFYCVLGYLGSL

ALVSFTVAFLARNLPDTFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVE
VFSILASSMGLLGCIFVPKCYIILFRPEKNAL
>R_50R
FTMAVSTVLAKTITVLLAFKVTDPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLG
TSPPFIDTDAHSEHGHTILVCNKGSVTAFYCVLGYLGSLALLSFTVAFLARNLPDT
FNEAKFLMFSMLVFCSVWVTFLPVYYSTKGKVMVAMEVFSILASGAGLLGCIFV
PKCYIILFRPEKNA
>R_52R
EQCIPCADHEYPNLERNHCLPKLVTFLGFEEFLGMALTCMALCFSVLTAAVLWV
FVKHQDTPIIKANNRTLSYILLIALLLCCLCSLLFLGHPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKALRPGRTMRRLLVIGASNFIIPICSLIQLALCGIWLGISPPF
IELDAHSEPDHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNMPDTFNET
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYV
ILIKPDKNSL
>R_53R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEDDLGKALASVALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPRTATCLLQQIAFCLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRQLLVSGASNFIIPICSLIQVALCVIWLGISPPF
IELDAHSEPDHIILVCNKGSATAFHCVLGYLGSLALASFTVAFLARNLPDTFNEAK
FLTFSVLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFLPKCYIILI
KPDKNSL
>R_54R
EQCIPCADHEYPNMERNRYLPKLVTFLAFEEFLGMALACMALCFSVLTAAVLW
VFVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFT
IAVSSVLAKT
ITVILAFKALKPGRTIRRLLVSGTSNFIIPICSMIQLVLCGIWLGTSPPFIEVDKHSEH
DYIILVCNKGSVTAFYCVLGYLGSLALASFIVAFLARNLPDTFNEAKFLTFSMLVF
CSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYVILIKPDKNS
>R_55R
QCMQCSDDEYPNQERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
ALSTVLAKTITVILAFKVTKPGRTMRRLLVSGVYNSVIPICSLIQLALCGIWLGTYP
PFIELDTHSELDHIILVCNKGSATAFYCVLGYLGSLALVSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSIKGKVMVAVEVFSILASSAGLLGCIFFPKCY
VILIRPEKNT
>R_56R
EQCIPCADHEYPNMERNHCLPKLVTFLAFEESLGMALACLALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQIAFGLLFTV
AVSTVLAKTITVILAFKALKPGRTTRRLLVSGASNSLIPICFLIQLALCGIWLGTYP
PFIELDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALVSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSVGLLGCIFFPKCYI
ILIKSDKNSL
>R_57R
EQCIPCADDEYPNMERNHCLPKLVTFLAFEESLGMALACVALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFIGRPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVMLAFKALKPGRTMRRLLVSGASNSLIPICFLIQLALCGIWLGTY

PPFIELDAHSEHGPIILVCNKGSVTAFYCVLGYLGSLALVSFTVAFLARNLPETFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILAFSAGLLGCIFAPKC
YIILIKSDKNSL
>R_58R
DQCVKCLDHQYANRERNQCFYKKLTFLSFEEPLGMILVCTALCFSVLTALVLGV
FVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGRPNTTTCVFQHTTFGVVFTV
AVSTVLAKTITVLLAFKVTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLGTSP
PFIDTDAHSEHGHIILVCNKGSVIAFYCVLGYLGSLALASFSVAFLARSLPGTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPKCY
IILLRPEKNA
>R_59R
DQCVKCPDLKYASTERDQCFYKKLTFLSFEEPLGMTLVCTAVCVSILTTVVLGVF
LKHRDTPIVKANNRCLSYILLIALIFFFLCSLLFIGHPNTTTCVLQQTTFGVVFNVA
VSTILAKTITVLLAFNFTAPGRRMRHWLISGVPNSIIPICSLIQLALCGIWLGTSPPFI
DTDAHSEHGHIILVCNKGSVTAFYCVLGSLGSLALASFTVASLARNLPDKFNETK
FLTFSMLVFCSVWVTFLPVYHSTKVKVMVAVEVFSILASSAGLLGCIFVPKCYIIL
FRPEKNA
>R_6R
DMKQCIKCLDDQYPNSKRDHCLPKVVTFLAYEDPVGMVLACAALSLSILTAVVL
GVFVKHRDTPIVKANNRTLSYILLLSLILCFLCPLFFIGQPNTATCILQQMTFGVLF
TVAVSTVLAKTITVVLAFKSTMPGMKMRQWLLSGAPNSIIPLCSLVQVTFGVVW
MVIAPPFIDTDAHSEPTLLIMVCNKGSVIMFYIFLGYLGSLAMGSFTLAFLARSLP
DTFNEAKFLTFSMLVFCSVWIAFLPVYNSTKGKFMVAVEVFSILASSAGLLVCIF
APKCYFILLKSERNIL
>R_60R
DQCMKCLDHQYASTERNQCFHKKVTFLSFEDPLGMTLVCTALCFSVLTAVVLG
VFVKHRDTPLVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVMF
TVAVSTVLAKTITVVLAFKVTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLG
NSPPFIDTDAHSEHGHIILMCNKGSVIAFYCVLGYLGSLAMASFTVAFLARNLPD
KFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIF
VPKCYIILLRPEKNA
>R_61R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEDDLGTALASVALSFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRRLLVSGASNFIIPICSLIQVALCVIWLGISPPF
IELDAHSEPDHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEAK
FLTFSVLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFLPKCYIILI
KPDKNSL
>R_62R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEEFLRMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFCLVFTV
AVSTVLAKTITVILAFKALKPGRTIRRLLVSGTSNFIIPICSMIQLVLCGIWLGTSPP
FIEVDKHSEHDYIILVCNKGSVTAFYCVLGYLGPLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVCHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCY
VILIKPDKNS
>R_63R

QCIPCEDHEYPNQDRSRCLPKLMTFLAFEESLGMALACMALCFSVLTAAVLWVF
VKHQDTPIVKANNRTLSYILLITLFLCFLCSLLFLGCPNTATCLLQQIAFGLVFTVA
VSTVLAKTITVILAFKVTKPGRTVRRLLISGVSNSVIPICFLLQLILCSIWLGISPPFIE
LDTHSEPGHIILVCNKGSVTAFYCVLGYLGSMALASFTVAFLARNLPDTFNEAKF
LTFSMLVFCSVWVTFLPGYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCYVILI
RPEMNS
>R_65R
DQCMKCPEQQYANTERDHCLHKAVAFLAYDDPLGMTLVCTALCFSVLTAVVL
GVFVKH*DTPIVKANNRCLSYILLIALIFCFLCSLLFIGRPSTATCVLQQITFGVVFT
MAVSTILAKTITVVLAFKVTAPGRKMRQWLLSGVLNSIIPICSLIQLVLCGIWLGT
SPPFVDTEAHSEHGHIILVCNKGSVISFYCVLGYLGSLALASFTVAFQARNLPDMF
NEAKFLTFSILVFWSVWVTFLPVYHSTKGKIMVAVEVFSILASSAGLLGCIFVPKC
YIILFRPEKNS
>R_66R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAILWV
FVKHRDTPIIKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQITFGLVFTVA
VSTVLAKTITVILAFKVTKPGRTIRRFLVSGASNSVIPICFLIQLVLCGIWLGTSPAF
TDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLAVASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCYII
LLRPERNS
>R_68R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEEFLGMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTI
AVSTVLAKTITVILAFKALKPGRTIRRLLVSGTSNFIIPICSMIQLVLCGIWLGTSPP
FIEVDKHSEHDYIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCY
VILIKPDKNS
>R_69R
DQCVKCPDLQYANTERDQCFYKKLTFLSFEEPLGMTLVCTAVCFSVLTAVVLGV
FVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTMFGVVFT
MAVCTVLAKTITVLLAFKVTAPGRRMRHWLISGLPNSIIPICCLIQLALCAIWLGT
SPPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDKF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFVPK
CYIILFRPEKNA
>R_71R
EQCIPCADDEYPNMERNHCLPKLVTFLAFEESLGMALACVALCFSVLTAAVLWI
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFIGRPNTATCLLQQIAFGLLFTVA
VSTVLAKTITVILAFKALKPGRTMRRLLVSGASNSLIPICFLIQLALCGIWLGTYPP
FIELDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALVSFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFLPKCYIIL
IKSDKNSL
>R_73R
SITVILAKTITVILAFKVTKPGRTMRRLLVSGVYNSVIPICSLIQLALCGIWLGTYPP
FIELDTHSELDHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSIWVTFLPVYHSTKGKVMVAMEVFSILASSAGLLGCIFFPKCYVI
LIRSEKNT

>R_76R
DHEYPNLERNHCLPKLVTFLGFEEFLGMALTCVALCFSVLTAAVLWVFVKHRDT
PIVKANNRTLSYILLIALLLCFLCSLLFLGHPNTATCLLQQIAFGLVFTVAVSTVLA
KTITVILAFKALRPGRTMRRLLVIGASNFIIPICSLIQLALCGIWLGSFPPFIDTDAHS
EPGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLVRNLPDTFNEAKFLIFSML
VFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYVILIKPDKN
S
>R_78R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEEFLGMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTIA
VSTVLAKTITVILAFKALKPGRTIRWLLVSGTSNFIIPICSLTQLVLCGIWLGTSPPFI
EVDKHSEHDYIILVCNKGSVTAFYCVLGYLGSLALASFIVAFLARNLPDTFNEAK
FLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFAPKCYVI
LIKPDKNS
>R_82R
EQCIPCADHEYPNMERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAILWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGHPNTATCLLQQITFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRRLLVSGASNSVIPICFLIQLALCGIWLGTSPP
FIDTDAHSEHGRIILVCNKGSATAFYCVLGYLGPLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCYII
LLRPERNS
>R_83R
QCIPCEDHEYPNQDRSRCLPKLMTFLAFEESLGMALACMALCFSVLTAAVLWVF
VKHQDTPIVKANNRTLSYILLITLLLCFLCSLLFIGCPNTATCLLQQIAFGLVFTVA
VSTVLAKTITVILAFKVTKPGRTVRRLLISGVSNSVIPICFLLQLILCSIWLGISPPFIE
LDTHSEPGHIILVCNKGSVTAFYCVLGYLGSMALASFTVAFLARNLPDTFNEAKF
LTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCYVILI
RPEMNS
>R_9R
DMKQCIKCLDDQYPNSKRDHCLPKVVTFLAYEDPVGMVLACAALSLSILTAVVL
GVFVKHRDTPIVKANNRTLSYILLLSLILCFLCPLLFIGQPNTATCILQQMTFGVLF
TVAVSTVLAKTVTVVLAFKSTMPGMKMRQWLLSGAPNSIIPLCSLVQVTFGVV
WMVIAPPFIDTDTNSEPTLLIMVCNKGSVIMFYIFLGYLGSLAMGSFTLAFLARSL
PDTFNEAKFLTFSMLVFCSVWIAFLPVYNSTKGKFMVAVEVFSILASSAGLLVCIF
APKCYFILLKSERNIL
>R_P1R
VAFLAYEDPLRMSQVCTALCFSVLTTGVLCVFVKH*DTPIVKANNHCLSYMLLI
ALIFCFLCSLLFIGHRNTAPCVLQQTTFGMVFTMALSTVLGKTITVVLAFKVTAPG
RRMRHWLISRAPNSIIPICSLILLAPCGFCLGTSPPFVDTDAHSEHGHIILVCNKDA
ETAFYCVLGYLGSIAMASFTVAFLARNLPDTFNEAKFLTFSMLVFCSLWVTFLPV
YHSTKGKVMVAVQVFSILASSSGLLGCIFVPKCYIILFRPEKNA
>R_P10R
DQCVKCLDHQYASTEHDHCLHKAVAF*AYEDPLGMTLVCTALCFSVLTAGVLGI
FVKHRHTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTATCVLQQTTFGVVFTM
AVSTVLAKTTTMVLAFKVTGPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLGTS
PPFIDTYAHSEHGHIILVCNKGSVIAFYCVLGYLGSLAMASFTVAFLGRNLPDKFN

EAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIFVPKC
YIILLRPEKNA
>R_P11R
DKLVKCPNHPYANTE*DQCFHKKMTFLSFEDPLGMILVCTALCCSILTPVVLGVF
VKHRETPIVKANNRCLSYILLIALIFCLLCLLFIGCPNTTTSVLQQTMLGLVFTMA
VSTVLAKIITVLLAFKVMAPGRRMRHWLLSGVTNSIIPICSLIQLALCAIWLGTSPP
FIDTDAHSEHGHIILVCNKSSVTAFYCVLGFLFSLALASFTVAFLARNLPDMFNEA
EFLTFSMLVFCSVWVTILPVYQSTKGR
>R_P12R
DQCVKCPEQQYANTEQDRCLHKAVAFLA*DDPLGMTLVCIAVCFSVLTAVVLG
VFVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGRPSTATCVLQQTTFGVVFT
MAVSTILAKTITVVLAFKVTAPGRRMRQWLLSGIPNNIIPICSLIQLVLCGIWLGTS
PPFIDTEAHSEHGHIILVCNKG*VIAFYCVLGYLGSLALASFTVAFLARNLPEMFN
EAKFLTFSILVFCSVWVTFLPVYHSTKGNIMVAVEVFSILCSSAGLLGCIFVPK*YI
ILFRPEKNS
>R_P13R
DQCVKCLDHQYASTEHDRCLHKSVAFLAYEDPL*MTLVCTALCFSVLTTGVLGV
FVKHRDTPIVKANNCCLSYTLLIALIFCYLCSLLFIGHPNTATCVLQQTTFGVVFT
MAVSTVLAKTTTVVLAFKVTAPGRRMRHWLISGLPNSIIPICSLIQLALCAIWLGT
SPPFIDTDAHSDHDHIILVCNKGSVTAFYCVLGYLGSLAMASFTVAFQARNLPDK
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIFVP
KCYIILLRPEKDA
>R_P14R
TE*DQCFHKKVTFLSFEDPLGMTLVCTALCFSALTAVVLGVFVKHRDTPIVKAN
NRCLTYILLIALIFCLLCLLFTGCPNTTTSVLQQTTLGLVFTMAVSTVLAKIITVFL
AFKVTAPGRRMRHWLISGVPNSIISICFLIQLALCGIWLGTSPFIDTDAHSEHGQIIL
VCNKGSVTAFYCVLGFLFSLALASITVAFLARNLPDKFNEAKLQTFSMLVFCSVL
VTFLPVYQSTKGKVMVAVEVFSILASSVGLLGCIFVPKCYIILFRPEKNA
>R_P2R
EQCIPCADHEYPNMERNHCLPKLVTFLAFEESLGMALACMALCFSVLTAAILWV
FVMHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGHPNTATCLLQQITFGLVFTV
AVSTVLAKTITVILAFKVTKPGITIR*LLVSGASNSVIPICFLIQLALCGIWLGTSPPF
IDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCYI
ILLRPERNS
>R_P3R
LCEALRHSHHEGQHRCLSYILLIALIFCFLCSLLFIGRPNTATPIFWKITFGLVFTVA
VSTVLANTITVILAFKVTKPGRTIR*LLVSGAFVIPIISLIQLVLGGFWLGIFPPFIDL
DSHSEHSHIILVCNKGSVTAFYCVLGYLGSLALFSFPSQEPA*HTQ*SQVPDIQHA
GVLQCLGDLPACLPQH*REGHGGCGGLLHLGLQCGAPGLHLCPQVLHYSYET*E
EL
>R_P4R
YPNPA*DRCLPKVVVFLAYEDPLGATLAGVALGFSVLTALVLGVFVKHRNTPLV
KANNRNLSYTLLTSLCLCFLCSLLFLGSPNTATCILQQTTFAIVFTVAVSTILAKIIT
VVLAFKATAPGKRMRQWLVSGASNSVIPICSLLQVMLCGIWLGTCPPFIDTDTHA
EPQQLLVLCNKGSATLFY

>R_P5R
MGMTLVCTAL*FSVLTAVVLGVFVKHQDTPIVKANNHSLSYILLIALIFCFLCSLL
FIGCPNTTTCVLQQTMFGVVFTVAVSTVLAKTITVVLAFKVTAPGRRMRHWLIS
GVPNSIIPSCSLIQLALCAIWLGTSPPFVDTDAHSDHGHIILVCNKGSVIAFYCVLG
YLGSLALVSFTVAFLARNLPDTFNEAKFLMFSMLVFCSVWVTFLLVYHSTKRKV
MVAVEVFSILASRAGLLGCIFVPKCYIILFRPEQKA
>R_P7R
DQCVKCPEQQYANTEQDHCLHKAVSFLAYDDPLGMTLVCTAVCFSVLTAVVLG
VFVKH*DTPIVKANNRCLSYILLIALIFCFLCSLLFIGRPSTATCVLQQTTFGVVFT
MAVSMVLAKTITVVLAFKVTAPGRRMRHWLISGVPNSIIPICSLIQLVLGGIWLGT
SPPFIDTDAHSEHGHIILVCKKVSVISFYCVLGYLGSLALASFTVAFLARNLPDMF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVQVFSILASSAGLLGCIFVPK
YYIILLRPEKNS
>R_P8R
DKLVKCPNHQYANTE*DQCLHKKVTFLSFEDPLGMTLVCTALCFSVLTAVVLGV
FVKHRDTPIVKANNRCLSYILLIALIFCLLCLLFIVCPNTTTCVLQQTTFGVAFTMA
VSTVLAKTITVLLAFKVSAPGRRMRHWLLSGVLNSIIPICSLIQLALCAIWLGTSPP
FIDTDAHSEHGHIILACNKGPVTAFYCVLGFLDILTLVSFTVAFLARNLTDTFNEA
KFMMFSMLVFCSVWVTFLPVYHSTKGKVMVAMEVFSILASGSGLLVCILVP**CI
VLFRPEKNA
>E_1
DMDHCLLCNNETHWAPVRSTTCYEKEVEYLNWNDSLAILLLGLSLLGIMFVLAI
GIIFTRNLNTPIVKSSGGLTVCYVILLSHFLNFASTGFFIGEPQSFTCKTRQTLFGVS
FTLCISCILMKSLKILLAFSFDPKLQNFLKRLYKPIPIIFTCTGIQVVICTLWLIFAAP
VVEENVSLPRVIILKCEEGSMVAFGSMLGYIAILAFICFIFAFKGRKLPENYNEAKF
ITFGMLIYFIAWITFIPVYVTTFGKYLPAIEIIVILISNYGILCCTFFPKCYVILWKQET
NTK
>E_3
QCV*CPAHEYPNSERNHCLPKVVTFLSFEDSLGVTLACIALCFSVVTAAVLWVFV
KHRDTPVVKANNHALSYMLLLSLLLASSAPYSSLAVPTQLCLLQRITFGVVFTVA
VSTVAKTITVILALKAMKPGRTMRWLLVSGASNSVVAICSLIQVIICGIWLGTSPP
FIDIDTHSEPRHLIIECNKGSVTAFYCILGYLGSLAMGTFSLASLAKNLPDTSVKPS
F
>E_4
FKEPLGTALASAALFLTGITAVVLGVFVKHRDTPIVKANNHTLSYLLTSRTLCLLC
SFLFIHRPCTATCNL*QTTFGVLFTVAVSSVLAKTITGVLAFKATAPDRRVRCWL
VSGAPNCVIPICSLIQVTLCGVWLGASPPFTDTDTHSKPGFVLLICNEGLVTTFYC
VLRYLGPLALGSFTVAFLARNLPDTFNEAKFLMFSVLVFCSIWVTFLPVYHSTKG
KVMEAVEIFAILASSTALLGCIFAPKCYIILLKPDRNTL
>E_P1
QCV*CPAHEYPNSERNHCLPKVVTFLSFEDSLGVTLACIALCFSVVTAAVLWVFV
KHRDTPIVKANNHALSYMLLLSLLLASSAPYSSLAVATQPPAFSDTSHLELCSLW
LFLLFLAKTLTVILALKAMKPGRTMRWLLVSGASNSVVAICSLIRVIICGIWLGTS
PPFIDIDTHSEPRHLIVECTKGSVTAFYCVLGYLGSLATGTFSLAYGARNLPDTFSE
AKCLTFSMLMFCSVWVTFRPVHHSTKGKAMVAMEIFSILTSSLGLLGCIFAPKCY
IILIKPDKNAL

>G_1
DIEQCVMCPEQEYANIQRTHCLQKSVMFLAYEDPLGKTLAGIALSFTIITSAVLGL
FVKHRDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPNTATCILQHTTFAVVFTV
AVSTVLAKTITVVLAFKVTAPGRRMRQLLVSGGPNYIIPICSLIQLTLCGVWMGTS
PPYIDTDAHSENGHIIITCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSMLASGAGLLVCIFAPK
CFIILFRPERNSL
>G_15
EQCVKCPDHQYANTARIQCLQKSVTFLAYEHPLGKMLAGTALSFTVLTAVVLGL
FVKHRDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPNTASCILRQTTFAVVFTV
AVSTVLAKTVTVVLAFKVTAPGRRMRQLLVSGAPNYIIPICSLIQLTLCGVWMGT
NPPYIDTDAHSEHGHIIITCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILISSAGLLGCIFGPK
CYIIFLRLDRNS
>G_17
DMEQCVKCPDQQYANDQKNQCLQKSVTFLAYKDTLGKALVSTALSLTILTAVV
LGLFVKHRDTPIAKANNRTLSYTLLISLTFCFLCALLFIGRPNTATCILQQTTFAVV
FTVAVSTVLAKTITVVLAFKVTAPGKRMRQFLVSGAPNYVIPICSLIQLTLCGTW
MGTNPPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALVSFTVAFLARNLP
DTFNEAKFLTFSMLVFCSVWITFLPVYHSTKGKVMVAMEVFSILASSAGLLGCIF
VPKCYIILLRPEKNAL
>G_2
PETNCLNPIMVLINTPIVKANNRTLSYTLLISLTFCFLCSLVFIGRPNTATCILQQTT
FAVVFTVAVSTVLAKTITVVLAFKVTAPGRRMRQFLVSGAPNYIIPICSLFQLTLC
GIWMGV
NPPYIDTDPHSEHGHIIIVCHKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWITFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFAPK
CYIILLRPERNSL
>G_23
EQCTQCPKHEYTNRNRDRCLPKILDFLSFEDPLGMALTQIALSFSVITAVILAIFVK
HRDTAIVKANNRTLSYTLLISLLLCFLCSLLFIGRPNTATCILQQMTFGLVFTVAVS
SVLAKTITVVLAFKVTGRRMRHLLVSGAPNYIIPICTLIQLTLCGFWMGTNPPYID
TDVHSEHGHIIIVCNKGSLTAFYCVLGYLGSLDLLSFTVAFLARNLPETFNEAKFL
TFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFLPKCYIILIR
SENNT
>G_26
EQCTQCPKHEYSNRNRDHCLPKILDFLSFEDPLGMALTQIALSFSLITAVILAIFVK
HRDTPIVKANNRTLSYTLLISLLLCFLCSLLFIGRPNTGTCILQQMTFGLVFTVAVS
SVLAKTITVVLAFKVTGPGRRMRHLLVSGAPNYIIPICTLIQLTICGFWMGTNPPYI
DTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLAKNLPDTFNEAKF
LTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFLPKCYVILI
RSENNT
>G_27
EQCIHCPKHEYPNRNRDRCLPKILDFLSFEDSLGMALTHVALSFSVLTVVIMAIFV
KHRDTAIVKANNRTLSYTLLISLLLCFLCSLLFIGRPNTATCILQQMTFGLVFTVA
VSSVLAKTITVVLAFKVTGPGRRMRHLLVSGAPNYIIPICTLIQLTICGFWMGTNP

PYIDTDAYSEHGHIILLCNKGSLTAFYCVLGFLGSLALLSFTVAFLAKNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFLPKCY
VIFIRSENNT
>G_28
DCCVKCPHHQYASTQRNHCLQKAVTFLDYEDPLGKTLAGTALSLTVLTVVVLG
LFVKHQDTPLVKANNRALSYTLLISLICCFICALLFIGRPSTGTCILQQTTFAVVFT
VAISTVLAQTITVVLAFKFTAPCRKLRQMLVSGAPNYIIPICCLIQLILCGIWMGTN
PPYLDTDAYSEHGHIIIMCNKGSLTAFYCVLGYLGFLALLTFTVAFLVRNLPDTFN
EAKFLTFSMLVFFSVWATFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFAPKC
YIILLKPERN
>G_3
EQCIKCPDHQYANSQRNNCVQKSITFMDYKDLSGKILAGTALSFTILTAAVLSLF
VKHRDTPIVKANNRALSYTLLVSLTFCFLCPLLFIGRPNTATCILQQTTFAVVFTV
AVSTVLAKTITVVLAFKVTAPGRRMRQLLVSGAPNYIIPICSLIQLTLCGVWMGT
NPPFVDTDAHSEYGHIIIMCNKGSLTAFYCVLGYLGFLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFVPK
CYIILLQPGRNS
>G_31
DMEQCTPCPDHQYANTERKHCLQKSETFLDYADSLGKTLAGSALSLTVFTAAVL
GLFVKHRDTPIVKANNRALSYTLLFSLTFCFLCTLLFIGRPSTASCILQWTTFAVVF
TVVVSTVLAKTITVVLAFKVTAPGRRMRQLLVSGAPNCIIPICTLVQLTLCGIWM
GTNPPYVDRDPHSEHGHIIIMCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLP
DTFNEAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIF
VPKCYIILLRPERNF
>G_33
EQCTQCPKHEYTNRNRDRCLPKILDFLSFEDALGMALTHIALSFSVFTTVILAIFV
KHRDTAIVKANNRTLSYTLLISLLLCFLCSLLFIGRPTTGTCILQQMTFGLVFTVAV
SSVLAKTITVVLAFKVTGPGRRMRHLLVSGAPNYIIPMCTLIQLTLCGFWMGTNP
PYIDTDAYSEHGHIIILLCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKLMVAMEVFSILASSAGLLGCIFLPKCY
VILIRSENNT
>G_35
EQCVKCLGHQYANTQRNQCLHKIVTFLAYEDPLGKALAGTALSLTALTAAVLGL
FVKHQDTAIVKANNRTLSYILLISLLSCFLCSLLFIGQPNTATCILQQTTFGLAFTV
AVSTVLAKTITVVLAFKITAPGKRMRQFVVSGTPNSIIPICSLIQLTLCAVWMGTN
PPYIDTDAHSEHGHIIIVCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASAAGLLGCIFVPK
CYIILLRPERNS
>G_38
EQYTQCPKDEYPNGNKDGCLPKVVDFLSFEDSLWVALAFIALRFSALTAVILTVF
VKHQDTPIVKANNRTLSYPLFISVLLCFLCTLLFIGHLDTGTCNLQQMTFGLVFTV
AVSSVLAKTIAVVLAFKVNAPGRRMRQLLVSGAANYITHLCSLIQLTLCAV*MG
TNPPYIDTDAHSEHGYIIIMCKKGSLTAFYCVLGYLSSLALLSFTVAFLARNLPDT
FNELKFLMFSMLVFCSVCITFLPMYYSSKGKVIVAMEIFSILASSAELLSCIMLPKC
YIVLIRSERNTL
>G_39

DMDQCVKCPEDQYPNMQRNQCLSKVMTFLSYEDTLGVALASTAISLSVLTALIL
GLFVQYRDTPIVRANNRNLSYILLISLILCFFCSLTFIGLPSPVTCILRQTFFGVVFSI
AVSAILAKTFIVVVAFKSVKPGNSMQMWMVTRFSNVIVCSCSLIQVCICTAWIIIS
PPFPDTDTQSEFEQIILQCNEGSNVAFYCVLGYLGFLAFLSLVLAFLARRLPDRFN
EAKMITFSMLVFCSVWISFIPSYLSSKGKTMVALEIFAILSSSAGLLGCIFLPKCYVI
LLRSDNISR
>G_4
DIEQCVMCPDQEYANIQRTHCLQKSVIFLSYEDPLGKTLAGIALSFTIITSAVLGLF
VKHQDTPIVKANNRALSYTLLISLTFCFLCTLLFIGHPNTATCILQHTTFAVVFTVA
VSTVLAKTITVVLAFKVTAPHRSMRQLLVSGAPNSIIPICSLIQLSLCGVWMGTNP
PYIDTDAHSEHGHIIIMCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSMLASSAGLLVCIFAPKC
FIILFRPERN
>G_41
DMHQCVKCPENQYPNMERYQCLPKVMTFLSYDDTLGTALASTAICLSVLTALIL
GLFIQYRNAPIVRANNRNLSYVLLISLILCFFCSLMFIGRPSPVTCILRQTFFGVVFS
IAVSAILTKTFIVLVAFKSIKPGSSIQMWMVTRISNAIVYICSLIQGCICTVWLATSP
PFPDTDTQSEFGQIILHCNEGSTVAFYCVLGYLGFLAFLSLVLAFLGRRLPDKFNE
AKMITFSMLVFCSVWISFIPSYLSTKGKTIVALEIFAILSSSAGLLGCIFLPKCYLILV
RPENNSR
>G_43
GQRECDQCGEDYWSNEQKSMCVLKEVEFLAYDEALGFTLVILSIFGALVILAVMI
VYVIYRQTPLVNANDRELSFLIQVSLIITVLSSMLFIGKPYNWSCMARQVTLALGF
SLCLSSTLGKTISLFFAYRISISKTRFISMRPIFRKIIVLISILVEIGICIAYLILQPPRVY
KNMESQNIKIILECNEGSIEFLCSMFGIDAFLALLCFLTTFVARQLPDNYYEGKCIT
FGMLVLFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCIFAPKCFIILLRPKR
NTD
>G_46
DMEHCVKCPDDKYATIQQDHCLQKSVTFLAHEDLIGKMLAGTALSLTVLTAAV
LGLFVKHQDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPNSATCILQQITFAVV
FTVVVSTVLAKTITVVLAFKITAPGKRMRQLLISGAPNYIIPICTLVQLTLCGVWM
GTNPPYIDTDAHSEHGHIIIILCNKGSLTAFYCVLGYLGSLALLSFTVAFLVRNLPD
TFNEAKFLTFSTLVFCSMWLTFLPIYHSSKGKVMVAMEVFSILVSGAGFLGCIFAP
KCYIILLRPERN
>G_47
DMEHCVKCPDDKYATIQKDHCLQKSVTFLAHEDLIGKMLAGTALSLTVLTGAV
LGLFVKHQDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPNSATCILQQITFAVV
FTVVVSTVLAKTITVVLAFKITATGRSMRQLLISRAPNYIIPICTLVQLTLCGVWM
GTNPPYIDKDAHSEHGHIIIILCNKGSLTAFYCVLGYLGSLALLSFTVAFLARNLPD
TFNEAKFLTFSMLMFCSMWLTFLPVYHSSKGKVMVAMEVFSILVSGAGLLGCIF
VPKCYIILLRPERNSL
>G_6
DMEHCIKCSDHEYANTQQNRCLKKAVTFLAYEDPLGKALAGTALSLTVLTAAA
LGIFVKHRDTPIVKANNPAMSYTLLISLIFCFLCSLLFIGRPSTVTCILQHTIFAIVFT
VAVSTVLAKTITVVLAFKATAPGRRMRQLLMSGALNYIIPICTLVQLTLCGIWMG
THPPYIDTDAHSEHGHIIIMCNKGSLRAFYSVLGYLGSLALLSFTVAFLARNLPDT

YNEAKFLTFSMLVFCSVWITLLPVYHSRKGKVIVAMEIFSILASSAGLLGCIFVPK
CYIMLFRKEKYS
>G_7
CVQCPPKEYANSERNHCLPKVITFLSLQDPLGMSLACTALCFSVITTLVLGLFVK
HRDTPMVKANNRGLSYTLLISLTFCFLCSLLFIGRPCTASCILQQTTFAVVFTVAV
STVLAKTITVVLAFKVTAPGRNMRQLLVSGAPNYIIPICTLIQLTLCAVWMGTNPP
FVDTDAHSEHGHIIIMCNKSSLTAFYCVLGYLGSLALLSFTVAFLARKLPGTFNEA
KFLTFSMLVFCSVWVTFLPVYHSSKGKVMVAMEVFSILASSAGLLGCIFLPKCYII
LIKPDKNSM
>G_P4
EQCTQCSKNVYPNRNRDSCLPKTVDLLSFEDPLGMALTHISLSISVFTAMILAIFA
KHQDTAIVDANNRTLSYTLLISLLLCFLCSLLFIGHPIIATCILQ*ITFGLVFTIAVSS
VLAKTITVVLAFKVTGPSRRMRQLLVSGAPNYIIPMCTLIQLTLCAIWMGRNPPY
VDTDAHSEHGHIILLCNKGSLTAFYCVLGYLGSLSLVSFTVAFLARNLPDNSSVA
KFFTFSMLVFCSMWVTFLPVYHSSKGKVMVAVEVFSILASSAGLLGCVFLPKCYI
ILIKSDNNS
>G_P6
DMDQCVKCPEDQYPNMERNQCLPKIMTFLSYEDTLGAALASTAISLSVLTALILG
LFIQYRDTPIVRANNRNLSYVLLISLILCFFCSLMFIGLPSPVTCILRQTFFGVVFSIA
VSAILAKTFIVVVAFKSIKPGNSIQMWMVTRVSNAIVCICSLIQGIICTIWLITTPPF
PDTDTQSEFGQIILQCNEGSTVSFYCVLGYLGFLAFLSLVLAFLARRLPDRFQ*GQ
K*SHFSMLVFCSVWISFILPT*ALKGKPMVA
>G_P7
DMEQCVKCLDHQYANTQRNHCLHK**LFWLMKIFWEKLWLAPPCLTALTAAIL
GLFVKH*DTPIVKANNQALSYTLLISLTFCFLCSLLFIGHPKTATCILQQTTFAVVF
TVAVSSVLAKTITVMLAFKITAPGSRMRQLLVSGAPNYIIPTCSMIQLTLCGIWMG
TIPPYIDTDAHSELGHIIIVCNKGSLTAFYRVLGYLDSLALLSFTMAFLARNLPDTF
DEAKFLTFSILMFCSVWITFLPVYHSSKGEVMLGMEVFSILASSTGLLGCILCP
>H_2
FYCSLLFISRPTKATCLLQQTGFGIVFTVAISSVLAKTIMVVLAFRAKKPGSRLRM
RMACTTLKTIVPVYSQIEATICGIWLGTSPPFLNKESHSERGVIIIECNKGSITAYYC
VLGYMGLLVLVSFIVAFLARSLPDTFTEAKFVTFSMLVFCSDWISFLPTHHSTKG
KAKVALEIFSILASSATLCTFIFICKCYVILLRPEKN
>R_11
QCVKCLDHQYASTEQDHCLHKAMTFLAYDEPLGMTLDCTALCFSVLTTVVLGV
FVKHQDTPIVKANNRCLSYILLIALIFCFLCSLVFIGRPSTATCVLQQTTFGMAFTV
AVSTVLAKTITVLLACKVTAPGRRMRHWLISGVPNSIIPICSLIQLALCAIWLGTSS
PFIDTDAHSEHGHIILVCNKGSVTAFYSVLGYLGSLALVSFTVAFLAGNLPDTFNE
AKFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVALMGCIFVPKC
YIILLRPKKNA
>R_15
DVEQCVRCLDDQYPNPARDRCLPKVVVFLAYEDPLGAALAGVALGFSVLTALV
LVVFVKHRNTPLVKANNRNLSYTLLASLCLCFLCSLLFLGRPNTATCILQQTTFAI
MFTVAVSTILAKTITVVLAFKATAPGKRMRQWLVSGAPNSVIPICSLLQVMLCGI
WLGTCPPFIDTDTHAEPQQLLVLCNKGSATLFYLVLGYLGSLALGSFSVAFLARS

LPDTFNEAKYLTFSMLVFCSVWVTFIPVYHSTKGKVMVAVEVFSILASSAGLLGC
IFAPKCYILLVRPNRNVL
>R_17
EQCVHCLDDQYPNPARDRCLPKVVVFLAYEDPLGAALAGLALGFSVLTALVLG
VFVKHRNTPLVKANNRNLSYTLLASLCLCFLCSLLFLGHPNTATCILQQTMFAIV
FTVAVSTILAKTITVVLAFKATAPGKRMRQWLVSGAPNSVIPICSLLQVMLCGIW
LGMCPAFIDTDTHAEPQRLLVLCNKGSATLFYLVLGYLGSLALGSFSVAFLARSL
PDTFNEAKYLTFSMLVFCSVWVTFIPVYNGTKGKVMVAVEVFSILASSAGLLGCI
FAPKCYILLVWPNRN
>R_19
DVEQCVRCLDDQYPNPA*DRCLPKVVVFLAYEDPLGAALAGVALCLTVLTALV
LGVFVKHRNTPLVKANNCNLSYTLIASLCLCFLCSLLFLGRPNTATCILQQTTFAL
VFTVAVSTILTKTITVVLAFKATAPGKRMRQWLVSGAPNSVIPICSLLQVMLCGI
WLGTCPPFIDTDTHAEPQQLLVLCNKGSATLFYLVLGYLGSLALGSFSVAFLARS
LPDTFNEAKYLTFSMLVFCSVWVTFIPVYHSTKGKVMVAVEVFSILASSAGLLGC
IFAPKCYILLMRPNRNV
>R_2
GQRECEQCGEDYWSNAEKSECVLKEVEFLAYDEALGFTLVILSIFGALVVLAVT
VVYVIYRHTPLVNANGRELSFLIQVSLTITVLSSLLFIGKPYNWSCMARQVTLALG
FSLCLACILGKTISLFLAYRISKSKTRLISIRPLYRKIIILISVLVEVGICTAYLVLEPP
RVYKNMESQNIKIIFECNEGSVEFLCSMFGVDVFLALLCFLTTFVARQLPDNYYE
GKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCLFAPKCFIIL
LRPKRNTD
>R_20
DQCVKCPDLQYANTERDQCFYKKLTFLSFEEPLGMTLVCTAVCFSVLTAVVLGV
FVKHRDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVLLAFKVTAPGRRMRHWLISGLPNSIIPICSLIQLALCAIWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDKFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIFVPKC
YIILFRPEKNA
>R_21
DQCVKCPEQQYANTEQDHCLHKAVSFLAYDDPLGMSLVCTALCFSVLTAVVLG
VFVKHRDTPIVKANNRCLSYILLIGLIFRFLCSLLFIGHPSTATCVLQQTTFGVVFT
MAVSTVLAKTITVLLAFKVTAPGRRMRQWLLSGVPNSIIPICSLIQLVLCGIWLGT
SPPFIDTDAHTEHGHIILVCNKGSVTAFYCVLGYLGTLALMSFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKIMVAMEVFSILASSAGLLGCIFVPK
CYIILFRPEKNS
>R_25
DQCMKCPYHQMQPRSETSACTNKVNFLFFEDPLGMTLVCTALCFSVLTTVVLGV
FVKHRDTPIVKANNRCLSYILLIALTFCFLCSLLFIGHPNTTTCVLLQTMLGLVFTV
AVSTVLAKTITVLLAFKVTAPGRRMRHWLISGAPNSIIPICSLIQLALCAIWLGTSP
PFIDTDAHSEHGHTILVCKKGSVTAFYCVLGFLGSLALVSFTVAFLARNLPDKFSE
AKFLTFSMLVFCSVWVTFLPVYQSTKGKVMMAVEVFSILCSSAGLLGCIFVPKC
YIIIFRPEKNA
>R_26

DQCVKCPALQYANTERDQCFYKKLTFLSFEEPLGMTLVCTALCFSVLTAVVLGV
FVKHRDTPIVKANNRGLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVLLAFKVTAPGRRMRQWLISGLPNSIIPICCLIQLALCGFWLGTSP
PFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASGAGLLGCIFVPKC
YIILFRPEKNA
>R_27
DQCVKCPDHQYANAERDQCFHKKVTFLSFEEPLGMTLVCTALCCSVLTAAVLG
VFLKHQDTPIVKGNNHSLSYILLFTLIFCFLCSSLFIGHPNTTTCVLQQTTFGMVFT
VAVSTVLAKTITVVLAFKVTAPGRRMRHWLISGLPNSIIPICSLIQLALCAIWLGTS
LPFIDTDDHSEHGHIILVCNKGSVTAHYCVLGYLGSLALGSCTVAFLGRNLPDKF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPK
CYIILFRPEENA
>R_3
GQRECEQCGEDYWSNAEKSECVLKEVEFLAYDEALGFTLVILSIFGALVVLAVT
VVYVIYRHTPLVNANGRELSFLIQVSLTITVLSSLLFIGKPYNWSCMACQVTLALG
FSLCLACILGKTISLFLAYKISKSKTRLISIRPLYWKIIILISVLVEVGICTAYLVLEPP
RVYKNMESQNIKIIFECNEGSVEFLCSMFGVDVFLALLCFLTTFVARQLPDNYYE
GKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCLFAPKCFIIL
LRPKRNTD
>R_35
QCIPCEDHEYPNQDRSRCLPKLMTFLAFEESLGMALACMALCFSVLTAAVLWVF
VKHQDTPIIKANNRPLSYILLITLLLCFLCSLLFLGRPNTATCLLRQIAFGLVSTVA
VSTGLAKTITVILAFKVTKPGRPVRRLLISGVSNSVIPICFLLQLILCSIWLGTSPPFI
ELDTHSEPGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEAK
FLMFSILVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCYVIL
IRPEMNS
>R_37
DQCVKCPEQQYANTEGDRCLHKAVSFLAYDDPLGMTLVCTAVCFSVLTAVVLG
VFVKHRDTPIVKANNRCLSYMLLIALIFCFLCSLLFIGRPSTATCVLQQTMFGVVS
TMAVSMVLAKTITVLLAFKVTAPGRRMRHWLIAGVPNSIIPICSLIQLVLCGIWLG
TSPPFIDTDAHSEHGHIILVCKKVSVNSFYCVLGYLGSLALASFTVAFLARNLPDM
FNEAKFLTFNMLVFISVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGGIFVP
KCYIILFRPEKNS
>R_4
QCNQCSEDEYPNSERNRCLSKLVTFLAFEESLGMALTCMALCFSVITAAVLGVFV
KHRETPIVKANNQSLSYILLISLLLCFLCSLLFIGRPNTATCILQQITFGLVFTVAVS
TVLAKTIIVILAFKVMKPGRTMRQLLISRASNSVIPICSMIQLIFCGIWLGTAPPFID
RDAHSEHGHIIIMCNKGSVTAFYCVLGYLGILAMGSFTVAFLSRNLPDTFNEAKF
LTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFVPKCYIIF
MRPEMNT
>R_44
EQCIPCADHEYPNLERNHCLPKLVTFLGFEEFLGMALTCMALCFSVLTAAVLWV
FVKHQDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPKTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRQLLVSGASNSVIPICFLIQLALCGIWLGSSPP
FIDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNE

AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFLPKCY
IILLRPERNS
>R_5
DQCIQCAEHEYPNSERNRCLSKLVTFLAFEDSLGMALTCMALCCSVITAAVLGVF
VKHRHTPIVKANNRSLSYILLISLLLCFLCSLLFIGRPQTATCILQQIMFGLVFTVA
VSTVLAKTITVILAFKVTKPGRTTRQLLISSVSNSVIPICSLIQLIFCGIWLGTSPPFID
RDAHAEHGHIIIVCNKGSVTAFYCVLGYLGFLALASFTVAFLARNLPDTFNEAKF
LTFSMVVFCSVWVTFLPVYHSTKGKIMVAVEVFSILASSAGLLGCIFAPKCYIILIR
PERNS
>R_51
QCVKCLDHQYASTEQDRCLNKAVAFLAYDEPLGMTLVFTALCFSVLTTVVLGV
FMKHQDTPIVKANNRCLSYILLIALVFCFLCSLLFIGRPNTATCVLQQTTFGVVFT
VAVSTVLAKTITVLLAFKVTAPGRRMRHWLLSGVPNSIIPICSLIQLALCAIWLGT
SPPFIDTDAHSEHGHIILVCNKGLVTAFYCVLGYLSSLALLSFTVAFLARNLPDTF
NEAKFLTFSMLVFCNVWVTFLPVYYSTKGKVMVAVEVFSILASSVALLGCIFVP
KCYIILFRPEKNA
>R_64
EQCIPSADHEYPNLERNHCLPKLVTFLGFEEFLGMALACMALCFSVLTAAVLWV
FVKHQDTPIIKANNRTLSYILLITLLLCFLCSLLFISHPNTATCLLQQIAFGLVFTVT
VSTVLAKTITVILAFKALKPGRTMRRLLVIGASNFIIPICSLIQLALCGIWLGSSPFID
TDAYSEPGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNEAEF
LTFSMLVFCSVWVTFLPVYHSTKGKVMVAVDVFSILVSSAGLLGCIFTPKCYVILI
KPDKNSL
>R_67
EQCIQCAEHEYPNMERSHCLPKMMTFLAFEDALGMALVCMALCFSVTTAAVLG
VFVKHRDTPIVKANNRSLSYILLIALLLCFLCSLLFIGRPNTATCLLRQITFGLVFT
VAISAVLAKT
ITVILAFKVMKPGRTMRRLLVSGASNFIIPICSLIQLALCGIWLGISPPFIELDAYSE
HGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEAKFLTFSML
VFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVPKCYIILIRPEKNS
>R_7
DMKQCIKCLDDQYPNSKRDHCLPKVVTFLAYEDPVGMVLACAALSLSILTAVVL
GVFVKHRDTPIVKANNRTLSYILLLSLILCFLCPLLFIGQPNTATCILQQMTFGVLF
TVAVSTVLAKTVTVVLAFKSTMPGMKMRQWLLSGAPNSIIPLCSLVQVTFGVV
WMVIAPPFIDTDTHSEPTLLIMVCNKGSVIMFYIFLGYLGSLAMGSFTLAFLARSL
PDTFNEAKFLTFSMLVFCSVWIAFLPVYNSTKGKFMVAVEVFSILASSAGLLVCIF
APKCYFILLKSERNIL
>R_70
EQCIPCADDEYPNMERNHCLPKLVTFLAFEESLGMALACVALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFIGRPNTATCLLQQIAFGLLFTVA
VSTVLAKTITVILAFKALKPGRTMRRLLVSGASNSLIPICFLIQLALCGIWLGTYPP
FIELDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALVSFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFFPKCYII
LIKSDKNSL
>R_72

DQCVKCPDLQYANTERDQCFYKKLTFLSFEEPLGMILVCTAVCFSVLTAVVLGV
FVKHQDTPIVKANNRCLSYILLIALIFCFLCSLLFIGHPNTTTCVLQQTTFGVVFTM
AVSTVLAKTITVLLAFKVTAPGRRMRHWLISGVPNSIIPICCLIQLALCAIWLGTSP
PFVDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKWKVMVAVEVFSILASGAGLLGCIFVPK
CYIILFRPEKNA
>R_74
EQCIPCADHEYPNLERNHCLPKLVTFLGFEEFLGMALTCMGLCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRRLLVSGASNSVIPIYFLIQLVLCGIWLGTSPP
FIDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNES
KFLTFSMLVFCSVWVTFLLVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCYII
LLRPERSS
>R_75
EQCIPCADHEYTNMERNRCLPKLVTFLAFEDALGMALACVALCFSVLTAAVLW
VFVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQITFGLVFT
VAVSTVLAKTITVILAFKVTKPGRTIRQLLVSGASNFIIPICSLIQVALCVIWLGISPP
FIELDAHSEPDHIILVCNKGSVTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFLPKYYII
LIKPDKNSL
>R_77
EQCIPCADHEYPNLERNHCLPKLVTFLGFEEFLGMALTCMGLCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
AVSTVLAKTITVILAFKVTKPGRTIRRLLVSGASNSVIPIYFLIQLVLCGIWLGTSPP
FIDTDAHSEHGHIILVCNKGSTTAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFFPKCYII
LLRPERNS
>R_79
QCMQCSDDEYPNQERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFVGRPNTATCLLQQIAFGLVFTV
ALSTVLAKTITVILAFKVTKPGRTMRRLLVSGIYNSVIPICSLIQLALCGIWLGTYP
PFIELDTHSELDHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFFPKCY
VILIRPEKNT
>R_8
DMKQCIKCLDDQYPNSKRDHCLPKVVTFLAYEDPVGMVLACAALSLSILTAVVL
GVFVKHRDTPIVKVNNRTLSYILLLSLILCFLCPLLFIGQPNTATCILQQMTFGVLF
TVAVSTVLAKTVTVVLAFKTTTPGMKMRQWLLSGAPNSIIPLCSLVQVTFGVVW
MVIAPPFIDTDTHSEPTLLIMVCNKGSVIMFYIFLGYLGSLAMGSFTLAFLARSLP
DTFNEAKFLTFSMLVFCSVWIPFLPVYNSTKGKFMVAVEVFSILASSAGLLVCIFA
PKCYFILLKSERNIL
>R_80
EQCIPCADHEYPNMERNRCLPKLVTFLAFEESLGMALACMALCFSVLTAAILWV
FVKHRDTPIIKANNRTLSYILLITLLLCFLCSLLFLGRPNTATCLLQQITFGLVFTVA
VSTMLAKTITVILAFKVTKPGRTIRRFLVSGASNSVIPICFLIQLALCGIWLGTSPPF
IDTDAHSEHGHIILVCNKGSATAFYCVLGYLGSLAVASFTVAFLARNLPDTFNES

KFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSVGLLGCIFLPKCYII
LLRPERNS
>R_81
DQCVKCPDLQYANTERNQCFHKRVTFLSFEDALGMTLVCTALCFSVLTVVVLG
VFVKHRDTPIVKANNRCLSYVLLITLIFCFLCSLVFIGRPNKTTCVLQQTTFGVVF
TVAISTVLAKTITVLLAFRVTGPGRRMRHWLISGLHNSIIPICCLIQLALCGFWLGT
SPPFIDTDAHSEHGHIILVCNKGSDTTFYCVLGYLGSLALASFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVPK
CYIILFRPEKNA
>R_84
IITMAVSTVLAKTITVVLAFKVTAPGRSMRQWLLSGVLNSIIPICSLIQLALCAIWL
GTSPPFIDTDAHSEHSHIILVCNKGSVTAFYCVLGYLVSLALLSFTVAFLARNLPD
TFSEAKFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSMGLLGCIF
VPKCYIILFRPEKNA
>R_P15
EQCIPCADHEYPNMERNRCLPKLVTFLAFEDDLGTALASVALSFSVLTAAVLWV
FVKHRDTPIVKANNRTLSYILLITLLLCFLCSLLFLGRPNTAT*LLQQIAFGLVFTV
AVSTVLAKTITVLLAFKVTKPGRTIRRLLVSGASNFIIPICSLIQVALCVIWLGISPP
FIELDAHSEPDHIILVCNKGSATAFYCVLGYLGSLALASFTVAFLARNLPDTFNEA
KFLTFSVLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGLLGCIFLPKCYIIL
IKPDKNSL
>R_P16
EQCIPCADHEYPNMERNHCLPKLVTFLAFEESLGIALACMALCFSVLTAAVLWV
FVKH*DTPIVKANNRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFTV
AVSAVLAKTITVILAFKVTKPGRTIRRLLVSGASNSFIPICFLIQLALCGIWLGISPPF
TELDTHSEHGHIILVCNKGSVTAFYCVLGYLGSLALVSFTLAFLARNLPDTFNEA
KVLTFSMLVFCSVWVSFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCYII
LIRPEKNS
>R_P6
DQCVKCPDHQYANKE*DQCFHKMVTFLSFEEPLGMTLVCTALCFSVLTAAVLG
VFLKHQDTPIVKANNHSLSYILLFTLIFCFLCSLLFIGQPNTTTCVLQQTTFGMLFT
VSVSTALTKTITVLLAFKVTAPGRRMRHWLISGVPKSIIPICSLIQLALCGFWLGTS
PPFIDTDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALGSCTVAFLARNLPDKF
NEAKFLMFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILCSSAGLLGCIFVP
KCYIILLRPEKNA
>R_P9
EQCTPCADHE*PKMERNHCLPKLVTFLAFEESLGMALACMALCFSVLTAVVLW
VFVKHRDTPIIKANKRTLSYILLIALLLCFLCSLLFLGRPNTATCLLQQIAFGLVFT
VAVSTVLAKTITVILAFKFTKPGRTIRQLLVSGASNSLIPICFLIQLAVCGIWLGISP
PFTELDAHSEHGHIILVCNKGSVTAFYCVLGYLGSLALASFTLAFLARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFLPKCY
IILIRPEKNSL
>M_AAK394381_TR
DEFNCLSCPGSMWSYKNNIACFKRRLAFLEWHEVPTIVVTILAALGFISTLAILLIF
WRHFQTPMVRSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQAFFTVCF
SVCLSCITVRSFQIVCVFKMARRLSAYGFWMRYHGPYVFVAFITAVKVALVAGN

MLATTINPIGRTDPDDPNIIILCHPNYRNGLLFNTSMDLLLSVLGFSFAYVGKELPT
NYNEAKFITLSMTFSFTSSISLCTFMSVHDGVLVTIMDLLVTVLNFLAIGLGYFGP
KCYMILFYPERNTSAYFN
>M_NP_114078_TR
DDFTCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAAL
GLSVHHWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHL
PLTGCLSTLFLQAAETFVESELLSWANWLCSYLRGLWAWLVVLLATFVEAALCA
WYLIAFPPEVVTDWSVLPTEVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLV
QSQPGRYNRARGLTFAMLAYFITWVSFVPLLANVQVAYQPAVQMGAILVCALGI
LVTFHLPKCYVLLWLPKLNTQEFFLGRNAKKA
>R_AAK516011_TR
DDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLCLVLGLTLAALG
LFVHYWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLP
LTGCLSTLFLQAAEIFVESELLSWANWLCSYLRGPWAWLVVLLATLVEAALCA
WYLMAFPPEVVTDWQVLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFL
VQSQPGRYNRARGLTFAMLAYFIIWVSFVPLLANVQVAYQPAVQMGAILFCALG
ILATFHLPKCYVLLWLPELNTQEFFLGRSPKEA
>M_NP_114073_TR
ELHTCQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANTLLLLLLIGTAGL
FAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQPLFSLGF
AIFLSCLTIRSFQLVIIFKSTKVPTFYHTWAQNHGAGIFVIVSSTVHLFLCLTWLAM
WTPRPTREYQRFPHLVILECTEVNSVGFLVAFAHNILLSISTFVCSYLGKELPENY
NEAKCVTFSLLLHFVSWIAFFTMSSIYQGSYLPAVNVLAGLATLSGGFSGYFLPK
CYVILCRPELNNTEHFQASIQDYTRRCGTT
>T_BAA261221_CS
DASICTKCPNNSWSSGNHTFCFLKEIEFLAWSEPFGIALAICAVLGVLLTAFVMGV
FVRFRNTPIVKASNRELSYVLLLSLICCFSSSLIFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSIHRKWWGLNLQFLLVFLCTFVQVMICVVWLYNA
PPSSYRNHDIDEIIFITCNEGSVMALGFLIGYTCLLAAICFFFAFKSRKLPENFTEAK
FITFCMLIFFIVWISFIPAYFSTYGKFVSAVEAIAILASSYGMLACIFFNKVYIILFKP
CRNTIEEVRCSTAAHAFRVAAKATLKHRTTVRKKSNSIGS
>RN_NP_058692_CS
DASACDKCPDDFWSNENHTSCIAKEIEFLAWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKWWGLNLQFLLVFLCTFMQILICIIWLYTAPP
SSYRNHELEDEIIFITCHEGSLMALGSLIGYTCLLAAICFFFAFKSRKLPENFNEAK
FITFSMLIFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFKPS
RNTIEEVRSSTAAHAFKVAARATLRRPNISRKRSSSLGG
>M_AAD283711_CS
DASACDKCPDDFWSNENYTSCIAKEIEFLAWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKWWGLNLQFLLVFLCTFMQIVICIIWLYTAPP
SSYRNHELEDEIIFITCHEGSLMALGSLIGYTCLLAAICFFFAFKSRKLPENFNEAK
FITFSMLIFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFKPS
RNTIEEVRSSTAAHAFKVAARATLRRPNISRKRSSSLGG
>C_NP_776427_CS

DASACDKCPDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKWWGLNLQFLLVFLCTFMQIVICAIWLNTAP
PSSYRNHELEDEIIFITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEA
KFITFSMLIFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFK
PSRNTIEEVRCSTAAHAFKVAARATLRRSNVSRQRSSSLGG
>HS_NP_000379_CS
DASACNKCPDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGV
FIKFRNTPIVKATNRELSYLLLFSLLCCFSSSLFFIGEPQDWTCRLRQPAFGISFVLC
ISCILVKTNRVLLVFEAKIPTSFHRKWWGLNLQFLLVFLCTFMQIVICVIWLYTAP
PSSYRNQELEDEIIFITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEA
KFITFSMLIFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKP
SRNTIEEVRCSTAAHAFKVAARATLRRSNVSRKRSSSLGG
>RN_Vom2r59
DVDQCMQCLDHQYANTAQNHCLQKSMTFLHYEDPMGISLSCIALCFSALTSLVL
GIFLKQKNTPIVKANNRTLTYILLIALNFCFLCSFLFIGHPNTVTCILQQTTFAVVFT
VAVSTVLAKTITVVLAFKVTAPGRRIRWLLVTRAPNFIVPICTLIQIILCGVWLGNS
PPFIDTDTHSENEHIIIMCNKGSVTAFYCVLGYLGALALGSFTLAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVVVEVFSILVSSAGLLGCIFAPKC
YIILLRPQRNSLQSHRHKIHYK
>RN_Vom2r58
DVDQCMQCLDNQYANTAQNHCLQKSITFLDYEDPMGIGLSCIALCFSALTSLVL
GIFLKQKNTPIVKANNRTLTYVLLIALNFCFLCSFLFIGHPNTATCILQQTTFAVLF
TVAVSTVLAKTITVVLAFKVTAPGRRIRWLLVTRAPNFTVPICTLIQIVLCGVWLG
NSPPFIDTDTRSEHEHIIIMCNKGSVTAFYCVLGYLGALALGSFTLAFLARNLPDT
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVVVEVFSILVSSAGLLGCIFAP
KCYIILLRPQRKSLQGHRDKIHYNYKT
>RN_Vom2r60
DVDQCMQCLDHQYANTAQNHCLQKSMTFLDYEDPMGISLSCIALCFSALTSLVL
SIFLKQKNTPIVKANNRTLTYILLIALNFCFLCSFLFIGHPNTATCILQQTTFAVVFT
VAVSTVLAKTITVVLAFKVTAPGRRIRWLLVTRAPNFIVPICTLIQIILCGVWLGNS
PPFIDTDTRSEHEHIIIMCNKGSVTAFYCVQGYLGALALGSFTLAFLARNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVVVEVFSILVSSAGLLGCIFAPKC
YIILLRPQRNSIQGHRDKIHYNYKTS
>C_VN2R401
PGIRECYQCPDQGPKQGEKLLPPRTCHLLASEDSLGMSLTCMALRFVIAALVLW
VFLKHQGTPIVKANNWVLSYVLLLFLLLCLLCSLLFIGHPNTATCALQQITFGVVF
TMAVATVLVNTPTVILAFKVRKPRRTMRWLVAGASNYVILVYFLICGVCLGIFPP
FLEMDIHSEPKDLIIVYKGLVTALYYILGYLGSLACSFSLAFLAKNLPETFSTKGK
VMVAVEIFSMLASIAGLLGCIFVKCYIILIRPDKN
>C_VN2R412
PDIKQCVQCPAQEHPNREENHCLPKVVTFLVFDDFLGISLTNMTLCFCVITGVVL
RVFVKHRDTPIVKANIQAFSYVLFISLLLCFLCSLLFIGHPQTAICILQQITFGIALTV
TVATVLAKTLTVILAFKTRKPRITMRRLLVTGAANVIPMCSLIQVIIYGALESSLPF
LEMDTHSEHKELLIVCNKGSITAFYCVLGFLGSITLGTFSLAFLAMNLPDTFNEAK
FLTFSTLVFCSVWVTFLPIYHSTKGKVMVAVEIFSILASSAGLLGCTFAPKLF

>C_VN2R402P
ELCVQCPDQEYTNWERNHCHPKIVTFLAFEDSWRMSLTCMALWFFVMTAMVL
WVFVKH*DIPVVKANSWTLSYVLLISLLLCFLCSLLFIGHPNRATCILQQITLGTVF
TVAAATGLARSLTVILAFMARKLGRTMRRLLETGAPNDVIPLCSKIQVILCGVWR
GTSPPFLEIDTHSEPKELIIVCKKDLVTDFYSVLGYLGSLALV
>C_VN2R406
PLFISLLPCFLCSLLFISRPHTATSILQQTTFGAVFTVAVVTVLAKIVPVILAFKARK
PGRTMRRLLLTGPYVFPMYSMIQVIICGVWLGASPPFLEMDTHSEPKELLIVCNK
GSVTAFYCVLGYLGSLALGTFSLAFLARNLPDTFSEAKFLTFSRLVLLSVWVTFL
PICHSTKGKVMVAVEIISILASSAGLLGCIFAPKYYIVLIRTDKNSLKSLKNKLSSK
GNKHS
>C_VN2R405
PGAQHCVQRPDQEYPNRKRNRCLPKVLTFVALEGSLGMSLACTAPCLVLMAVV
LWVFVKHRHTPIVKANNRALSHVLLISFLLCFLRSLLFINHPPTATCILQQITLEAV
FTVAVATVLAKTLAVILAFKARKPGRTMRRLLLTGPYVFPMYSMIQVIICGVWL
GASPLFLEMDTHSEPKELLIVCNKGSVTAFYCVLDYLGSLALGTFSLAFLARKLP
DTFSEANFLTFSMLVLLSVWVTFLPICHSTKGKVMVAVEIISILASSAGLLRCIFAP
KCYTLL
>C_VN2R404
PDAEQCAPCPDREYPNRERNHCLPRVVTFLAFEGSLGMSLACTALCFCVLTAVV
LWVFLKHRDTPIVKANNRALSHVLLISLLLYFLRSLLFINHPHTATCILQQITFGAV
FTVAVATVLAKTLAVILAFKARKPGRTMR*LLVTGATNYVFPMCSLIQVIICGVW
LGTSPPFLEIDAHSEPKELLILCNKGSVTAFYHVLGYLGSLALGTFSLAFLGRNLP
DAFSEAKFPTFRMVVFLSIWVTLLLVYHSTKGKVMAAVETFSILA
>C_VN2R403
PDAEHCILCPAQEHPNSERNRCLPKDLTSLGFDDPLGVSLACTALCLSAITAEVL
WVFVK*QDTPIVKANNWTLS*VLLISLLLCFLCSLLFIGHPNRATCILQQITFGAGL
SIAVASVLAKTLTVNLAFKTRKPGRTMRW*LVTGASNDVIPLCSLIQVILCGVWL
GTSPPFLEMDTHSEPKELIIMCNKGSVTAFYCVLGYLGSLALGTFSLAFLARNLPD
TFSETKFLAFRMLVFCSVWVTFLPYRTKGKVMVAVEILSMLASSAGLQGRIFAPE
CYIILLRPEKITLKVLKN
>C_VN2R411
PDTEQCVQCLAHEYPNIEKSLSPQIVIFLAFEDILGMSLACMALCFCVITSVVFWV
FVKHRDTPIIKANNWPLSYVLLISLLLCFLCSLLFIGCPNIVICIIQQITFGSVFTVAL
AIVLAKTLIVILAFKAMKPGRTMR*LLVIGASNYVIPICSLIQLIICGIWLGTTPPFLE
KSHSETMELIIICNKCLVTAFYHVLGYLGFLALGTFFLAILARNLPDIFNEAKFLTF
SMLVFCNVQVTFLPVYHSTKGKVMVAVKIFSILASSAGLLGCIFAPKCY
>C_VN2R410
PVKANNWPLSYVLLISLLLCFLCSLLFIGCPNIVICIIQQITFGSVFTVALAIVLAKTL
IVILAFKAMKPGRTMR*LLVIGASNYVIPICSLIQLIICGIWLGTTPPFLEKSHSETM
ELIIICNKCLVTAFYHVLGYLGFLALGTFFLAILARNLPDIFNEAKFLTFSMLVFCN
VQVTFLPVYHSTKGKVMVAVKIFSILASSAGLLGCIFAPKCYTILIRHEKNSLKSF
RNKIDSKRNR
>MQ_VN2R208
PMTIWACASVTFLWGVFGVFVTHQDTPIVKANNQTLRCVLFIALNLYFHCPLFIG
YPNIATCILWNITFELMFTCLHSLDKTITVVLTFQVTTQETDMHWLVSGVSNSVIA

VSSLIEAEVGGISWGMTTAFMHVDTSEPEYVIFTKTTIWLLTVL*ILGSLAHGSFS
VAFLARSLPDIFNNPILKFSMLVFFSVWVTVLLVYQNPK
>O_VN2R644
DTEQCLKCPEDQYPSKEKDRCLPKTVTYLAYEETLGMALAAVALSLSLLTTLVL
GIFVMHRDTPIVKANNRSLSYTLLTSLTLCFLSSLLFIGRPTPATCLLRQSTFAIVFT
VTIASILAKTITMVLAFRATKQVSKLRRLMGTRVSYFLIFICTLIQMCLCGIWLGT
YPPFLEKDTHSEFSFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTF
NEAKYLTFSMLVFCSVWIAFLPAYQSTKGKAMVAMEIFSILVSGAGILTCIFVPKC
YVIFLRSDRNTLSLLKSKAN
>O_VN2R645
DTEQCLKCPEDQYPNKEKDRCLPKTVTYLAYEEPLGMTLTVVALSLSLLTTLVL
GIFVMHRNTPIVKVNNRSLSYTLLTSLTLCFLSSLLFIGRPTPATCLLRQTTFAIVFT
VTIASILAKTITVVLAFRATTPGSKLRRWLGTRVSYLLIFICTLIQMGLCGIWLGTY
PPFLEMDTQSEASFIVFQCNEGSLVAFYCVLGYMGFLALVSFTLAFLARNLPDTF
NEAKYLTFSMLVFCNVWISFLPTYQSTKGKVMVAVEIFSILVSSAGILICIFTPKCY
VILLRADRNTLSILKSKAN
>O_VN2R640
NAEECMQCPEDQYPSKEKDRCLPKTVTYLAYEETLGMILAFAALCFSILTTFVLG
VFVMYRDTPIVKANNRSLSYILLTSLTLCFLSSLLFIGRPTPAICLLRQTTFAIVFTV
AIASILAKTITVVLAFRATQPGSKLRRWLGTSVSYFLIFICTLIQMCLCGIWLGTYP
PFLEMDRHSEASFIIFQCNEGSLVTFYCVLGYMGFLAFVSFTLAFLARNLPDTFNE
AKYLTFSMLVFCSVWIAFLPAYQSTKGKAMVAMEIFSILVSSAGILICIFAPKCYII
LLRADRNTLEILRNKAN
>O_VN2R639
NAKECMQCPEDQYTTKEKDRCLPKTVTYLAYEETLGMILAFAALCFSILTTFVLG
VFVMYRDTPIVKANNRSLSYVLLTSLTLCFLSSLLFIGRPTPAICLLRQTTFAIVFT
VAIASILAKTITVVLAFRATQPGSKLRRWLGTSVSYFLIFICTFIQMCLCGIWLGTY
PPFLEMDRHSEASFIIFQCNEGSLVTFYCVLGYMGFLAFVSFTLAFLARNLPDTFN
EAKYLTFSMLVFCSVWIAFLPAYQSTKGKAMVAMEIFSIVVSSAGILICIFAPKCYI
ILLRADRNTLEILKNKANS
>O_VN2R641
NAEECMQCPEDQYTTKEKDRCLPKTVTYLAYEETLGMILASAALCFSILTTFVLG
VFVIHRDTPIVKANNRSLSYILLTSLTLCFLSSLLFLGRPTTAICLLRQTTFAIVFTV
AIASILAKTITVVLAFRATQPESKLRRWLGSSVSYFLIFICTLIQMGLCGIWLGTYP
PFLEKDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSYALAFLARNLPDTFN
EAKYLTFSMLVFCSVWIAFLPAYQSTKGKAMVAMEIFSILVSSAGILICIFAPKCYI
ILLRADRNTLEILKNKSN
>O_VN2R624
DAEQCTQCPEDHYTSKEKDRCLPKTITYLAYEETLGMILASAALCFSVLTTFVLG
VFIMHRDTPIVKANNRSLSYILLISLTLCFLSSLLFIGRPTPVTCILRQTTFAIVFTVA
IASILAKTITVVLAFRATKPGSQLRRWLRPAVSHSLIFICTLIQMCLCSIWLGTYPPF
LEMDTQSEVSFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNEA
KYLTFSMLVFCSVWISFLPTYQSTKGKTMVAMEIFSIVVSSAGILICIFAPKCYVIL
LRPDRNTLEILRNKANH
>O_VN2R623

DAEQCMQCPEDHYPSKEKDRCLPKTVTYLAYEETLGMILASLAFCLSILTTFVLG
VFVMHRDTPIVKANNRSLSYILLISLTLCFLSSLLFLGRPTPVTCILRQTTFAIVFTV
AIASILAKTITVVLAFRATKPGSQLRRWLRPAVSHALIFICTLIQMCLCSIWLGTNP
PFFEMDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNE
AKYLTFSMMVFCSVWIAFLPAYQSTKGKAMVALEIFSIVVSSAGILICIFAPKCYV
ILLRADRNTLQILKNKAS
>O_VN2R621
DAEQCMQCPEDQYPSKEKDRCLPRTVTYLAYEETLGMILASSAFCFSILTTFVLG
VFVMHRDTPIVKANNRSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVA
IASILAKTITVVLAFRATKPGSQLRRWLRPAVSHALIFICTLIQMCLCGIWLGTYPP
FLEMDKHSEVSFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNE
AKYLTFSMLVFCSVWIAFLPAYQSTKGKAMVALEIFSIVVSSAGILICIFAPKCYVI
LLRADRNTLQILRNKAN
>O_VN2R629
DAEECMKCPEDHYPSKEKDRCLPKTVTYLAYEETLGKFLASSAFCFSILTTFVLG
VFVMHRDTPIVKANNRSLSYILLISLTLCFLSSLLFLGRPTPVTCILRQTTFAIVFTV
AIASILAKTITVVLAFRATKPGSQLRRWLRPAVSHSLIFLCTLIQMCICGIWLGTYP
PFLEKDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNE
AKYLTFSMLVFCSVWIAFLPTYQSTKGKAMVAMEIFSILVSSAGILICIFAPKCYVI
LLRPDRNTLEILRNKAS
>O_VN2R627
DAEQCRQCPEDHYPSKEKDRCLPRTVTYLAYEETLGMILASSALCFSVLTTFVLG
VFVMHRDTPIVKANNCSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVA
IASILAKTITVVVAFRATKPGSQLRRWLRPAVSHALIFLCTLIQMCLCSIWLGTYPP
FLETDTHSEASFIVIECNEGSLVAFYCVLGYMGFLALVSFTLAFLARNLPDTFNEA
KYLTFSMLVFCSVWISFLPTYQSTKGKAMVAMEIFSILVSSAGILICIFAPKCYVIL
LRPDRNTLQILKNKAN
>O_VN2R619
DAEQCMQCPEDQYPSKQKDRCLPKTVTYLAYEETLGMILASAALCFSILTTFVLG
VFVMHRDTPIVKANNCSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVA
IASILAKTITVVLAFRATKPGSQLRRWLRPAVSHSLIFICTLIQMGLCGIWLGTDPP
FLEMDTHSESSFIVFQCNEGSLVTFFCVLGYMGFLALISFTLAFLARNLPDTFNEA
KYLTFSMLVFCSVWISFLLAYQSTKGKAMVAMEIFCILVSSAGILICIFAPKCYVIL
LKPGRNTLEILKNKAN
>O_VN2R630
DAEECIQCPEDQYPSQEKDRCLPRTVTYLAYEETLGMILTSAAFCFSILTIFVLRVF
VIHRDTPIVKANNRSLSYILLTSLTLCFLSSLLFLGRPTPAICILRQTTFAIVFTVAIA
SILAKTITVVLAFRATKPGSKLRRWLRPAVSHSLIFLCTLIQMFLCSIWLGTYPPFF
EMDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNEAK
YLTFSMLVFCSVWISFLPAYQSTKGKAMVAMEIFSILVSSAGILICIFAPKCYVILL
RPDRNTLEILKNKAS
>O_VN2R632
DAEECMQCPEDQYPSKEKDRCLPRTVTYLAYEETLGMILTSAAFCFSILTTFVLG
VFVMYRDTPIVKANNRSLSYILLTSLTLCFLSSLLFIGRPTPAICVLRQTTFAIVFTV
AIASILAKTITVVVAFRATKPGSQLRRWLRPAVSHSLIFLCTLIQMCLCGIWLGTY
PPFLEMDIYSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARKLPDTFN

EAKYLTFSMLVFCSVWISFLPAYQSTKGKAMVAMEIFSIMVSSAGILICIFAPKCY
VILLRSDRNTLGKLKNKAN
>O_VN2R622
DAEECMQCPEYQYPSKEKDRCLPRTVTYLAYEETLGMILASAALCFSILTMFVLG
VFVIHRDTPIVKANNRSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVAI
ASILAKTITVVLAFRATKPGSKLRRWLRPAVSHSLIFICTLIQMCLCGILLGTYPPF
LEMDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNEA
KYLTFSMLVFCSVWISFLPAYQSTKGKAMVAMEIFSILVSSAGILICIFASKCYVIL
LRPDRNTLEILKNKAS
>O_VN2R625
DAEECMQCPEYQYPSKEKDRCLPRTVNYLAYEETLGMILAFAALCFSILTTFVLR
VFVMYRDTPIVKANNCSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVA
IASILAKTITVVLAFRATKPGSKLRRWLRPAVSHSLIFICTLIQMCLCGFWLGTYPP
FLEMDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLVRNLPDTFNE
AKYLTFSMLVFCSVWISFLPAYQSTKGKAMVAMEIFSILVSSAGILICIFVPKCYVI
LLRPDRNTLEILKNKAN
>O_VN2R618
DTVQCMQCPEDQYPSKEKDRCLPKTMTYLTYEETLGMFLAFAALCFSILTTFVL
GVFVMYRDTPIVKANNRSLSYILLTSLILCFLSSLLFLGRPTPAICILRQTTFAIVFT
VSIGSILAKTITVVLAFRATKPGSQLRRWLRPAVSHSLIFICTLIQMCLCGIWLGTD
PPFLEMDTHSEASFIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFN
EAKYLTFSMLVFCSVWISFLPTYQSTKGKAMVAMEIFSIVVSSAGILICIFAPKCY
VILLRADRNTLEILKNKSNH
>O_VN2R635
DAEECIQCPEDQYPSKEKDRCLPKIVTYLAYEETLGMILASAALCFSILTTFVLGV
FVMHRDTPIVKANNRSLSYILLTSLSLCFLSSLLFIGRPTPANCVLRQTTFAIVFTV
AIASILAKTITVVVAFRATKPGSKLRRWLRPAVSYFLIFTCTLIQMCLCGIWLGTY
PPFLEMDIYSEVSSIVIQCNEGSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFN
EAKYLTFSMLVFCSVWISFLPAYQSTKGKVMVAMEIFSILVSSAGILTCIFAPKCY
VILLCSDRNTLKIMKNKAN
>O_VN2R634
DAIECMQCPEDQYPSKEKDRCLPKIVTYLAYEETLGMILASAALCFSILTTFVLGV
FVMHRDTPIVKANNCSLSYILLISLTLCFLSSLLFIGRPTPAICILRQTTFAIVFTVAI
ASILAKTITVVLAFRATKPGSKLRRWLGTSVSYFLIFICTLIQMCLCGIWLGTYPPF
LEMDIYSEFSSIVIQCNESSLVTFYCVLGYMGFLALVSFTLAFLARNLPDTFNEAK
YLTFSMLVFCSVWISFLPAYQSTKGKTMVAMEIFSILVSSAGILTCIFAPKCYVILL
CSDRNTLKITKNKADH
>O_VN2R652
DAQQCLQCPEDQYPSKERDRCLPKTITYLAYEEPLGMALASAALCFSLLTALVLG
VFVKHRDTPIVKANNRSLSYTLLISLTLCFLCSLLFIGLPTQATCLLRQISFAVVFT
VAIASILAKTLTVVLAFAATKPGSGLRKWLGSTTSYFLICICTLIQMCLCGIWLGT
HAPFLERDIHSEASLIVIQCNEGSLFFFYSVLGYMGFLALVSFMMAFLARKLPDTF
NEAKYLTFAMLVFCSVWISFLPTYQSTKGKAMVAVEIFSIVVSSAGILTCIFAPKC
YVILLRSDRNTLEILKNKVSF
>O_VN2R643

DAQQCLQCPEDQHPSKERDRCLPKTITYLAYEEPLGMVLASAALGLSLLTVLVL
GVFVKHRDTPIVKANNRSLSYTLLISLTLCFLCPLLFIGLPTQATCLLRQISFAVVF
TLAIASILAKTLTVVLAFAATKPGSGLRKWLGSTTSYFLICICTLIQMCLCGIWLGT
HAPFLERDIHSEASLIVIQCNEGSFFFFYSVLGYMGFLALVSFMMAFLARKLPDTF
NEAKYLTFAMLVFCSVWISFLPTYQSTKGKAMVAMEVLSILLSSAGILTCFFAPK
CYVILLRSDRNTLEILKNKVS
>O_VN2R647
DAQKCLQCPEDQHPSKERDRCLPKTITYLAYEEPLGMALASAALCFSLLTALVLG
VFVKHRDTPIVKANNRSLSYTLLISLTLCFLCSLLFIGLPTQATCLLRQISFAVVFT
LAIASILAKTLTVVLAFAATKPGSRLRKWLGSTTSYFLICICTLIQMCLCGIWLGT
HAPFLERDIHSEASLIVIQCNEGSLFFFYSVLGYMGFLALVSFMMAFLARKLPDTF
NEAKYLTFAMLVFCSVWISFLPTYQSTKGKAMVAMEVLSILLSSAGILTCFFAPK
CYVILLRSDRNTLEILKNKVS
>O_VN2R648
DAQKCLQCPEDQHPSKERDRCLPKTITYLAYEEPLGMALASAALCFSLLTALVLG
VFVKHRDTPIVKANNRSLSYTLLISLTLCFLCSLLFIGLPTQATCLLRQISFAVVFT
LAIASILAKTLTVVLAFAATKPGSGLRKWLGSTTSYFLICICTLIQMSLCGIWLGTH
APFLERDIHSEASLIVIQCNEGSLFFFYSVLGYMGFLALVSFMMAFLARKLPDTFN
EAKYLTFAMLVFCSVWISFLPTYQSTKGKAMVAMEVLSILLSSAGILTCFFAPKC
YVILLRSDRNTLEILKNKVS
>O_VN2R650
DAQKCLQCPEDQHPSKERDRCLPKTITYLAYEEPLGMVLASAALCFSLLTALVLG
VFVKHRDTPIVKANNRSLSYTLLISLTLCFLCSLLFIGLPTQATCLLRQISFAVVFT
VAIASILAKTLTVVLAFAATKPGSGLRKWLGSTTSYLLICICTLIQMCLCGIWLGT
HAPFLERDIHSEASLIVIQCNEGSLFFFYSVLGYMGFLALVSFMMAFLARKLPDTF
NEAKYLTFAMLVFCSVWISFLPTYQSTKGKAMVAVEIFSILLSSAGILTCFFAPKC
YVILLRSDRNTLEILKNKVS
>O_VN2R593
DMDQCIMCPEDEYPNKERNCCLQKMVTFLAYENPMGMTLACTALIFCILTFLILG
LFVKHQDTPIVKANNRTLSYILLLSLSQCFLSSLLFMGKPTSATCLLQQTIFGITFT
MAVSSILAKTVTVVLAFKATRPGSRSRKWVGSRAPISIILVCSLIQGILCGIWLLVS
PPFPDIDPHSEPGHMILKCNEGSLIAFYLVLGYMGFLALASFIVAFLARNLPDTFN
ETKFITFSMLVFCSVWVSFLPTYQSTKGKSMVAVEVFSILSSSAGLLGCIFIPKCYV
ILLRPQWNTQEVVNLKSYSRRRNHSKT
>O_VN2R594
DMDQCIMCAEDEYPNGEKNHCLPKMVTFLAYEDPMGMTLACTALVFCLLTLLI
LGLFVKHQDTPIVKANNRNLSYILLFSLSQCFLSSLLFMGKPTSVTCLLQQTIFGIS
FTMAVSSILAKTVTVVLAFKATRPGSRSRKWVGSRAPISLILVCSLIQGILCGIWLL
VSPPFPDIDPHSEPGHMILKCNEGSLIAFYLVLGYMAFLALASFIVAFLARNLPDTF
NETKFITFSMLVFCSVWASFLPTYQSTKGKMMVAVEVFSILSSSASLLCCIFIPKC
YVILLRPQWNTQNWINHKSYARRRNDSKAFSHAS
>O_VN2R570
SLDQCMKCPEDQYPNTEKIQCLPKRVIFLTYEETLGMALASIAASFSLLTALVLW
VFVKHKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPSAINCLLRQTTFGVMFT
VAISSILAKTITVVLAFRATRPSSRSRRWVGSRTPISLVVFCTLIQVMLCAIWLFLS
PPFPDADTYSDPEHIILECNEGSLIAFYLVLGYIALLALGSFTVAFLARNLPDTFNE

AKFITFSMLVFCSVWLSFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYV
ILLQPQKNTNKFLRN
>O_VN2R584
GKSCPEDQYPNRERIQCLPKRVTFLAYEENLGMALASIAASFSLLTVLVLWVFVK
HRDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTAISCLLRQTTFGVMFTVAISS
ILAKTITVVLAFRATRPGSRSRRWVGSRGPIYLVVFCTFIQVSLCTIWLFLSPPFPD
ADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTFNEAKFIT
FSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYVILLQP
QKNTKKFLKN
>O_VN2R585
GKDQCMKCPEDQYPNRERIHCLPKRVAFLAYEETLGMALASIAASFSLLTALVL
WVFVKHRDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTTTNCLLRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPGSRSRRWVGSRAPIYLVVFCTIIQVSLCIMWLF
LSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIAFLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKVMVAVEVFSILSSSAGLLGCIFIPKCY
VILVQPQKNTKKFLKN
>O_VN2R575
GKDQCMKCPVDHYPNRERIHCLPKSVTFLAYEETLGMALASIAASFSLLTALVL
WVFVKHKDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTTTNCLLRQTTFGVT
FTVAVSSILAKTITVVWAFRATRPGRRSRRWVGSRGPISLVVFCTFIQVTLCTIWL
FLSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILVQPQMNTKKFLKN
>O_VN2R607
GKDQCMKCPEDQYPNRERIHCLPKRVAFLAYEENLGTALASIAASFSLLTALVL
WVFVKHKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPSATNCLLRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPSSKRRRWVGSRAPISLVVFCTFTQVMICSMWL
FLSPPFPDVDTHSDTEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEARFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCY
VILLQPQKNTKKFLKN
>O_VN2R565
GKDQCMKCPEDQYPNRERNHCLPKGVNFLAYKETLGIALASIAASFSLLTALVL
WVFVKHKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPNAINCLLRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPSSKSRRWVGSRAPISLVFFCTFIQVVLCAIWLL
LSPPFPDVDTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NETKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCY
IILLQPQKNTKKFLKNNT
>O_VN2R572
GKVQCMKCPEDQYPNRERIQCLPKSVTFLAYEENLGIALASIAASFSLLTALVLW
VFVKHRDTPIVKANNRDLSYLLLFSLSLCFLCSLFFIGRPSAANCLLRQTTFGVMF
TVAISSILAKTITVVLAFRATRPGSRSIKWVGSRAPISLVVFCTFIQVTLCAMWLFL
SPPFPDVDTHSDPEHIILECNEGSLIAFYSVLGYIAFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWVSFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCY
VILVQPQKNTKKFLKNNT
>O_VN2R567

GKDQCMKCPEDQYPNRERIHCLPKRVTFLAYEETLGMALASIAASLSLLTVLVL
WVFVKHRDTPIVKANNRDLSYLLLFSLSLCFLCSLFFIGRPSAANCLLRQTTFGV
MFTVAISSILAKTITVVLAFRATRPGSRSRRWVGSRGPISLVVFCTFIQVMLCAIW
LLLSPPFPDVDTHSDPEHIILECNEGSLIAFYSVLGYIAFLALGSFTVAFLARNLPDT
FNEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILLQPQKNTTKFLKK
>O_VN2R571
GKDQCMKCPEDHYPNRERIHCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVFVNQKDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTAISCLLRQTAFGVM
FTVAISSILAKTITVVLAFRATRPGSSSRRWVGSRGPISLVVFCTFIQVMLCAIWLL
LSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWVSFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YIILLQPQKNTKKFLKK
>O_VN2R566
GKDQCMKCPEDQYPNRERIQCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVIVNQKDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTAINCLLRQTTFGVLF
TVAISSILAKTITVVLAFRATRPGSRSRRWVGSRGPISLVVFCTFIQVMLCAIWLLL
SPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYII
LLQPQKNTKK
>O_VN2R568
GKDQCMKCPEDQYPNRERIHCLPKRVTFLAYEETLGMALASIAVSFYLLTALVL
WVFVNQKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGHPSAINCLLRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPGSRSRRWVGSRGPISLVVFCTFIQVMLCAIWL
LLSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILLQPHKNTNKFLKK
>O_VN2R577
GKDQCMKCPEDHYPNTERIHCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVFVKHRDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPTATNCLLRQTTFGVL
FTVAVSSILAKTITVVLAFRATRPGSRSRRWVGSRGPIYLVVFCTFIQVSLCTMWL
FLSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILLQPQKNTKKCLKN
>O_VN2R599
GKDQCMKCPEDHYPNTERICCLPKRVTFLAYEENLGMALASIAASFSLLTALVL
WVFVKHKDTPIVKANNQDLSYILLFSLSLCFLCSLFFIGRPTAISCLLRQTTFGVMF
TVAISSILAKTITVVLAFRATRPGSRSRRWVGSRAPISLVVFCTFIQVMLCAIWLLL
SPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYII
LLQPQKNTKKFLKN
>O_VN2R603
GKDQCMKCPEDHYPNTERIHCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVFVKHRDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPTAINCLLRQTTFGVM
FTVAISSILAKTITVVLAFRATRPGSRSRRWVGSRAPISLVVFCIFIQVMLCAIWLL
LSPPFPDADTHSDPEHIILECNEGSLIAFYSVLSYIAVLALGSFTVAFLARNLPDTF

NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILLQPQKNTKKFLKN
>O_VN2R600
GKDQCMKCPEDQYPNTERIHCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVFVKHKDTPIVKANNRDLSYILLLSLSLCFLCSLFFIGRPTAINCLLRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPGSRSRRWVGSRAPISLVVFCTFIQVMLCAIWL
LLSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYIALLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YVILVQPQKNTKKFLKN
>O_VN2R583
GKDQCMKCPEDQYPNRERIHCLPKRVTFLAYEETLGMALASIAASFSLLTALVL
WVFVKNKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPTTINCLLRQTTFGVMF
TVAVSSILAKTITVVLAFRATRPGSRSRKWVGSRGPISLVVFCTFIQVTLCTMWLF
LSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYITFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCY
VILLQPQKNTKKFFKK
>O_VN2R604
GKDQCMKCPEDQYPNTERIHCLPKRVTFLAYEENLGMALASIAASFSLLTALVL
WVFVKHRDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGPPTAINCLLRQTTFGVMF
TVAVSSILAKTITVVLAFRATRPGSRSRRWVSSRTPISLVVFCTFIQVMLCAIWLLL
SPPFPDADIHSDPEHIILECNEGSLIAFYSVLGYIAFLAVGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYVI
LLEPQKNTKKFLKN
>O_VN2R606
GKGQCMKCPEDQYPNRERMHCLPKRVTFLAYEETLGMALASIAAFFSLLTALIL
WVFVKHKDTPIVKANNRDLSYILLFSLSLCFLCSLFFIGRPTAISCLFRQTAFGVMF
TVAISSILAKTITVVLAFRATRPGSKSRRWVGSRGPISLVVFCTTIQVILCTMWLLL
SPPFPDADTHSDPEHIILECNEGSLIAFYSVLSYIAFLALGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYIIL
LQPQKNIRKFLKN
>O_VN2R563
GKDQCMKCSEDQYPNREKIHCLPKGVTFLAYEDPLGMTLTSTAVCFSLLTALVL
WVFVKHKDTPIVKANNRDLTYILLFSLSLCFLCFLLFIGPPNTINCLLRQITFGVMF
TVAVSSILAKTITVVLAFRATRPGSRSRRWMGSRAPICLILFCTFIQAMLCAIWLL
LSPPFPDVDTHSDPEHIILDCNEGSLIAFYSVLGYIAFLALGSFAVAFLARNLPDTF
NEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKC
YIILLQPQRNTKENLKN
>O_VN2R590
GNEQCMKCPEDQYPNRERIYCLPKLAIFLDYEEPMGKTLASIAVCFSLLTVLVLW
VFVKHKDTPIVKANNRDLSYTLLISLSLCFLCSLLFIGPPTPTNCLLRQTTFGVMFT
VAVSSILAKTVTVVLAFRATRPGSRSRRWVGSKAPISLAVFCTFIQVMLCAIWLL
LSPPFPDADIHSDPEHIILKCNEGSLIAFYSVLGYMAFLALGSFTVAFLARNLPDTF
NEARFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILASSAGLLGCIFIPKC
YVILVQPQKNTKKPLKK
>O_VN2R591

GNEQCMTCPEDQYPNRDRIYCLPKLAIFLDYEEPMGKTLASIAVCFSLLTVLVLW
VFVKHKDTPIVKANNRDLSYTLLISLSLCFLCSLLFIGPPTPTNCLLRQTTFGVMFT
VAISSILAKTITVVLAFRATRPGSRSRRWVGSRAPISLVVFCTFIQVMLCAIWLLLS
PPFPDADTHSDPEHIILECNEGSLIAFYFVLGYIALLALGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCYVI
LVQPEKNTKKTLKN
>O_VN2R589
GKEQCMKCSEDQYPDRDRIYCLPKLVIFLDYEEPMGKTLASIAVCFSLLTVLVLW
VFVKHKDTPIVKANNRDLSYTLLISLSLCFLCSLLFIGPPTPTNCLLRQTTFGVMFT
VAISSILAKTITVVLAFRATRPGSRSRRWVGSRGPISLVVFCTFIQVMLCAIWLLLS
PPFPDADTHSDPEHIILECNEGSLIAFYSVLGYMAFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFIPKCY
VILVQPQKNTKKTLKNNTNF
>O_VN2R588
GNEQCMKCSEDQYPDRDRIYCLRKQVIFLDYEEPMGKTLASIAVCFSLLTVLVL
WVFVKHKDTPIVKANNRDLSYTLLISLSLCFLCSLLFIGPPTPTNCLFRQTTFGVM
FTVAVSSILAKTITVVLAFRATRPGSRSRRWVGSKAPISLVVFCTFIQVMLCAIWL
FLSPPFPDADTHSDPEHIILECNEGSLIAFYSVLGYMASLAFGSFTVAFLARNLPDT
FNEAKFITFSMLVFCSVWISFLPTYQSTKGKMMVAVEVFSILSSSAGLLGCIFVPK
CYVILVQPEKNTKKSLKNNTTT
>O_VN2R598
DMEHCEKCPEDQYSNSHRNHCVSKTEVFLGYEEPIGMALTCAALAFSLFTVFILG
VFVRHQGSPIVRSNNAVLSYILLISLALCFLSSLLFLGHPTTVTCLLRQTTFGTLFT
VAISSILSKTLTVVFAFQATRPGSTMRKWLRPGVFYSVVFILSSIQVILCAIWLGSF
PPFLDADTHSQPEYIILECNEGSFVAFYCVLGYIAILALGSFTVAFLARNLPDSFNE
AKFITFSMLVFCSVWISFLSTYQSTKGKATVAVEVFSILTSSAGLLGCIFIPKCYTIL
FRANGHTHEWLKRPPSGSQKVHLHILPSATSSSSCSAGRRPIK
>RN_Vom2r35
DMKRCLLCPKEQYSSNTRDHCLPRTEIFLAFEEPLGFILALMAILLAGLALLVLGV
FLKHRDIPVVRANNRTLSYFLLISLSLCSLCSLLFLGRPTVTTCLLRQTTFAVVFTV
AVSSVLAKTVTVVLAFRVTKPGSRIQVCLSPGASTSVILIASLIQTVLCGVWLATS
PPFPERDMISEPRHIVIQCQEGSGATFFCVLGFLGFLAVGTFSVAFLARDLPDVFN
ETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTDGLLCGIFIPKCYII
LLKPEKNT
>RN_Vom2r36
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFILTLMAILLAGLALLVLGV
FLKHRDIPVVRANNRTLSYFLLISLSLCSLCSLLFLGRPTVTTCLLRQTTFAVVFTV
AVSSVLAKTLTVVLAFRVTKPGSRIQVCLSPGASTSVILIASLIQAVLCGVWLATS
PPFPERDMISEPRHIVIQCQEGSGATFFCVLGFLGFLAGGTFSVAFLARALPDVFN
ETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTAGLLGGIFIPKCYII
LLKPEKNT
>RN_Vom2r34
DMKRCLLCPKEQYLRHTRDHCLPRTEIFLAFEEPLGFILALMAILLAGLALLVLGV
FLKHRDTPVVRANNRTLSCFLLISLSLCSLCSLLFLGRPTVTTCLLRQTTFAVVFT
VAVSSVLAKTVTVVLAFRVTKPGSRIQVCLSPGASTSVVLIASLIQAVLCGVWLA
TSPPFPERDMISEPQHIVIQCQEGSGATFFCVLGFLGFLAGGTFSVAFLARDLPDVF

NETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTAGLLGGIFIPKCY
IILLKPEKNT
>RN_Vom2r33
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFILALMAILLAGLALLVLGV
FLKHRDTPVVRVNNRTLSYFLLISLSLCSLCSLLFLGRPTVTTCLLRQTTFAVVFT
VAVSSVLAKTVTVVLAFRVTKSGSRIQVCLSPGASTSVVLIASLIQVVLCGVWLA
TSPPFPERDMISEPQHIVIQCQEGSGATFFCVLGFLGFLAGGTFSVAFLARDLPDVF
NETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTAGLLGGIFIPKCY
IILLKPEKNT
>HS_VN2R19
PGTKRCLLCPEEQYSSRTRDGCLPRTETFLAFDDPLGLMLALVALMLASLAVLVL
RLFKHRDTPVVGANNRALSCTPLTSLTLCALCPLPCLGCPTAATCRLYQTTVAVV
FTMAVSSVLAFRVPGDRVWVCLGPGASTSVVLAASLVQVVLCGIWLGTSPPFPD
GDMASEPSHIVIQC*EGSAVAFSCVLGYLGLLAGGTFFVAFLARGLPDAFNKSKF
LTFSVLLFCSFWTTFLPLYHSAQGKSTVAVEIFSILASMAGLLGGIFPKCYIILLKPE
RNA
>CH_VN2R116
PGTKRCLLCPEEQYSSRTRDGCLPRTETFLAFDDPLGLTLALVALMLASLAVLVL
RLFKHRDTPVVRANNRALSCTLLTSLTLCPLPFLGCPTAATCLLYQTTVAVVFTM
AVSSVLAFRVPGDRVWVCLGPGASTSVVLAASLVQVVLCGIWLGTSPPFPDGDM
ASEPSHIVIQC*EGSAVAFSCVLGYLGLLAGGTFFVAFLARGLPDAFNKSKFLTFS
MLLFCSFWTTFLPLYHSAQGKSTVAVEIFSILASMAGLLGGIFPKCYIILLKARAE
>CH_VN2R115
PNKKNKMEILEMKKLNKQMNNVFNRLINILDIVKEKVNELDVRTIKVKQRLRFIR
HLYVFHNLAILIFAHLYTLVVLYLLTYMCSYMCPLYYVFLPS*ATCILSSATFSVQ
FTVAVSSASAKTFTVVMAFQATRPASRMKVLLGTLKSHSVVLCPLIQVTISGI*LE
TSPSSLDKDHSPGIIIIQCNGGSITVFYCVLSYLGFLVLGTFAVVFLESNPPDIFNET
KCLTFSMLVFCSIWVAFLLSYSSNQGKAMVAVEIVSILTSSTGLHGCILLPCYVIL
LHPERNSLQCLKKKLRNHIQKDYTIFRLRGGTIMLISGREFVS
>HS_VN2R18
PSLQQRSHHLENPFKCRKCGKLFNRISPLMLHQRIHTSEKPYKSLVEDFLTIITFQK
FLFYIDFPMFHFNRFSVTFLLWFMLMCFFFGTLLYIKDCILVECLPRFATFSVQFT
VAVSSASAKTFTVVMAFQATRPASRMKVLLGTLKSHSVVLCPLIQVTISGI*LETS
PSSLDKDHSPGIIIIQCNGGSITVFYCVLSYLGFLVLGTFAVVFLESNPPDIFNETKC
LTFSMLVFCSIWVAFLLSYSSTQGKAMVAVEIVSILTSSTGLHGCILLPCYVILLHP
ERNSLQCLKKKLRNHKQKDYTIFRLRGGTIMLISGREFVS
>HS_VN2R17
PDIKQCRLCPEYHYPNKKREHTFSNRTFLAYDETLGLTLACTALFYSVLSVLILVE
FVWHQNTPIVKASNQTLSYTLLVSLTLCFLSSSLFIGRPSPATCLLSQTTFAAVFTV
AVFSAAFQAIRPESRIRKWMGPQKTNSVVFLCSFTQVTLCGIWLGTEPPFVNKDP
QFMPGYIIIQCNEGSVTAFYSVLGYLGFLVLGSLAVAFLARNLPDAFNEAKFLTFS
MLVSCSVWVAFLPSYKHPGQGQSGREIFSIVGSSTGLLGCIFAPKCYVILLHLERN
TIQCLKK
>CH_VN2R114
PDIKQCRLCPEDHYPNKKREHAFSNRTFLAYDETLGLTLACTALFYSALSVLILM
EFVWHQNTPIVKASNQTLSYTLLVSLTLCFLSSSLFIGRPSPATCLLSRTTFAAVFT

VAVFSAAFQAIRPESRIRKWMGPQKTHSVVFLCSFTQVTLCGIWLGTEPPFVNKD
PQFMPGYIIIQCNEGSVTAFYSVLGYLGFLVLGSLAVAFLARNLPDAFNEAKFLTF
SMLVSCSVWVAFLPSYKHPGQGQSGREIFSIVGSSTGLLGCIFAPKCYVILLHLER
NTIQCLKK
>MQ_VN2R210
PDIKQCRLCPEDHHPNKKREHAFQNRAFLAYDETLGLALGCTALFYSALSVLILV
ELVWHRNTPMVRTSNLTLSYPRVVSLTLCFLSSLLFISRPSPATCLL*RTTFAAMF
TVAVSSAAFQATRPASRIQKWMGPQKTHSVVFLCSCIQGTLCGIWLGTEPPFVDN
DPQSVPGYIIIQCNEGSVTAFHSVLGYLGFLVLGTLAVVFLGRNLPDAFNEAKLLT
FSMLVSCGVWVAFLPSYYSTQGKARVAEIFSIVGSSARLLGCIFAPKCYVILLHLE
RNILQCLKKK
>O_VN2R613
DMDNCMKCPEDQYPNREKDRCLPKVVTFLNTEEPLGWILTCMALSFSLLTLLILI
LFIKHKDTPIVKANNRVLSYTLLISLIFCFLCSLLFIGRPTTVTCLLRQTTFGIIFTVA
VSSILAKTIMVVLAFKATRPGSRSRIWLKPRVYNFLVLICSMIQVILCGIWLGTSPP
FLGTDTHSEPGHIIIECKEGSVIAFYCVLSYMGILALGSFMVAFLARNLPDTFNEA
KFITFSMLVFCSVWICFLPTYQSSKGKAMVAMEVFAILASSIGLLVCIFIPKCYVIL
LRPDRNILEGLRNKADFRANN
>O_VN2R614
NMDNCMKCPEDQYPNREKDRCLPKIVTFLNTEEPLGLILTCMALSFSLLTLLILIL
FVKHQDTPIVKANNRILSYTLLISLIFCFLCSLLFIGRPTTVTCLLRQTTFGVVFTM
AVSSVLAKTIMVVLAFKATRPGSKIRMWVQPQVYNFLVLICSMIQVILCGIWLGT
SPPFLGTDTHSEPGHIIIECKEGSVIAFYCVLSYMGILALGSFMVAFLARNLPDTFN
EAKFITFSMLVFCSVWVCFLPTYQSSKGKAMVAMEVFAILASSIGLLGCIFIPKCY
VILLRPDRNTSEWLRNKADSRVRN
>O_VN2R610
NMDSCMKCPEDQYPNRGKDRCLPKAVTFLSVEEPLGWTLTCMAISFSLLTLLILI
LFIKHQDTPIVKANNKVLSYTLLISLIFCFLCSLLFIGHPTTVTCLLRQTTFGIVFTV
AVSSILAKTIMVVLAFKATRPGSRIRMWVQPSVANSLVLICSGIQVILCGIWLGILP
PFLGTDPHSEPGHIIIECKEGSSIAFYCVLGYMGVLALASFIVAFLVRNLPDTFNEA
KFITFSMLVFCSVWICFLPTYQGSKGKAMVAMEIFAILASSVGLLGCIFIPKCYVIL
IRPDKNALEGLRNKAETRAKNLATEA
>O_VN2R597
DMDHCTKCPEDEYPNKNRDHCFPKTVTFLSFKEPLGLTLTCLALSFSFLTVIILVIF
VKCQDTPIVKANNRILSYILLISLTLCFLCSILFIGPPTSAKCLLQQTVFGIVFTVAIS
SILAKTITVVLAFKATKPGSKMKRWMLPRVSSFVVFICSMIQLALCCIWLGTSPPF
PGLDTHSEPDQIIIICEMGSAITFYSVLGFMGFLALGSFTVAFLARNLPDTFNEARFI
TFSMLIFCSVWMSFIPAYQNPKGQAMVTVEIFCILASSAGLLGCIFIPKCYVILLRP
ERNTLEWLRQKA
>O_VN2R560
DAEQCTKCPEDEYPNKHKDHCLPKVVTFLNMKEPLGMTLAFIAVSFSLLTALVL
GVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLIFIGHPTTATCLLQQTAFAVVFT
VAVSSVLAKTFIVVLAFKGIRPGSRIRIFMHPRVSVCIVLICSGLQVIFCGIWLGTSP
PFPEADTHSQYGHIILKCNEGSVLAFHCVLGYIALLALGSFSVAFLARNLPDTFNE
AKFITFSMLVFCSVWASFLPTYQSTKGKAMVAVEIFSILASSAGLLGCIFIPKCYVI
LLKPDRNTQRQIKNK

>O_VN2R552
NAEHCMKCPDVEYPNRERDRCIPKTITFLNAKDPLGMTLALLAVFFSLLTALVL
WVFVKFQDTPIVKANNRTLSYTLLISLTFCFLCSLLFIGHPTPATCLLQQTVFAVV
FTVAVSSILAKTVIVVLAFRSIKPGSRIRMFIQPRVSIYIVLICSGIQVIFCGIWLVTS
PPFPEADTHSEYGHIIIRCSEGSTFAFYSVLGYVALLALGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWASFLPTYQSTKGKAMVAVEIFSILASSAGLLGCIFIPKCYIIL
LKPDRNTQKKLKNKQVFKGKVFPRTQSYTSSSWNIRLEE
>O_VN2R544
DAEFCTKCPEDEYPNKHKDRCLPKVVTFLNGKETLGMTLGLVAVSFSLLTALVL
GVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPTTATCLLQQTVFAIVFT
VAVSSILAKTLIVVLAFRVIRPGSRIQMFIQPRVSNYIVFICSGIQVTFCGIWLGTSP
PFPEADTHSEYGHIIIRCNEGSAFAFYCVLGYMALLALGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWVSFLPTYQSSKGKAMVVVEIFSILASSAGLLGCIFIPKCYIIL
LKPDRNTQEQIKNK
>O_VN2R551
DAENCRKCPEIEYPNKQRDGCLPKTVIFLSMKEPLGMTLGLMAVSFSLLTALVLG
IFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGYPTTITCFLQHTVFGVVFTV
VVSSILAKTIIVVVAFRGIRPGSRIRMFIHPKVSIFVVLICSGIQVIFCGIWLGTSPPFP
EADAHSEYGHIIIRCNKGSTFAFYCVLGYMGFLAGGSFTVAFLARNLPDTFNEAK
FITFSMLVFCSVWATFLPTYQSTKGKSVVIVEIFSILASSAGLLGCIFIPKCYVILLK
PDRNTMGQIKK
>O_VN2R555
NAEHCTKCPEDEYPNKQRDRCLPKVVTFLNVQEPLGMTLALMAVSFSFLTALIL
WVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPTVATCLLRQTVFGVVF
TVAVSSILAKTMIVVLAFKGIRPGSRIRMFMHPRVSIYVVLFCSGIQVIFCALWLVI
SPPFPEADTHAEYGHIILRCNEGSDFVFYCVLGYMGFLALGSFSVAFLARNLPDTF
NEAKFITFSMLVFCSVWASFLPTYHSTKGKAMVVVEIFSILASSAGLLGCIFIPKC
YVILLCPDRNTLEQIQNKKDSRGKI
>O_VN2R554
NAEHCMKCPEDEYPNKQRDRCLPKVVTFLNVQEPLGMTLAFIAISLSFLTALILW
VFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPTPATCLLRQTVFGVVFT
VAVSSILAKTIIVVLAFKGIRPGSRIRMFMHPRVSIYVVLICSGIQVIFCALWLVISP
PFPEADTHSQYGHIILRCNEGSDFVFYCVLGYMGFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWASFLPTYQSAKGKAMVVVEIFSILASSAGLLGCIFIPKCY
VILLCPDRNTLEQIQNKQDSRGKI
>O_VN2R501
DAEHCKKCPENEYPTKQRDRCLSKVVTFLDVKEPLGMTLIFMAVSFSFLTALVL
WVFVKFRDTPIVKANNRSLSYTLLISLIFCFLCSLLFIGHPTTATCLLQQTVFGVVF
TVAVSCILAKTLMVVVAFRGIRPGSRIQMFIHPRMARYVVLMCSGIQVILCGIWL
GTYPPFPEADTHSEYGHIIIRCNEGSAFVFYCVLGYMGFLALGSFTVAFLARNLPD
TFNEAKFITFSMLVFCSVWATFLPTYQSTKGKAMVIVEIFSIMASSTGLLGCIFLPK
CYVILLKPNRNTQEQIKNKQDSRAK
>O_VN2R556
DAEHCMKCPEDTYPNKQRDHCVPKPVTFLNVKEPLGMTLMCMAVSFSLLTALV
LGVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGRPTTATCLLQQTVFGVV
FTVAISSIMAKTIIVVLAFRVIRPGSRIRMFVHPTASIYVVLICSGIQVTFCGIWLGT

SPPFPEADTHSEYGHIILRCNEGSVFAFYSVLGYMALLALGSFTVAFLARNLPDAF
NETKFITFSMLVFCSVWITFLPTYQSSKGKAMVTVEIFSILASSAGLLCCIFIPKCY
VILLKPDRNTQEQIKNK
>O_VN2R553
NAEHCRKCPEIEYPNKQRDRCLPKIVTFLNVKEPLGMTLTCMAISFSLLTALVLGI
FVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGRPTTATCLLQQTVFVVVFTV
AVSSILAKTIIVVVAFRGIRPGSRIRMFIHPRVSYYIVLICSGLQVIFCAFWLVTSPPF
PEADTHSQYGHIIIRCNEGSAFAFYSVLGYLGFLALGSFSVAFLARNLPDTFNEAK
FITFSMLVFCSVWVTFLPTYQSTKGKAMVAVEIFSILASSAGLLGCIFIPKCYVILL
KPDRNTQEQIKNKQASKGKI
>O_VN2R550
NADQCKKCPEDKYPNEQRDRCLSKAVTFLTVKEPLGMTLTCMAVSFSLLTALVL
GVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPTTATCLLRQTVFAVVF
TVAVSSILAKTIIVVVAFRGIRPGSRIRMFMHPRMSVYVVLICSGLQVTLCGIWLG
TSPPFPEADTHSEYAHIIIRCNEGSVFAFYSVLGYIAFLALGSFTVAFLARNLPDAF
NEAKFITFSMLVFCSVWVSFLPTYQSTKGKAMVAVEIFSILASSAGLLACIFIPKC
YVILLKPDRNTQRQIKNK
>O_VN2R558
DAEDCMKCPEIEYPNQQKDHCLSKIVTFLNMKESLGMMLTLVSISFSFLTALVLG
VFVKFQDSPIVKANNRTLSYTLLISLIFCFLCSLLFIGQPTTATCCLQQTVFAVVFT
VAVSSILSKTIIVLLAFRSISPGSRIRMFMHPRVSVYVVLIGSGLQVILCGIWLGTSP
PFPEADTQSEYGHIIIRCNEGSAFAFYCVLGYMGFLALGSFTVAFLARNLPDAFNE
AKFITFSMLVFCSVWATFLPTYQSSKGKAMVTVEIFSILASSAGLLGCIFIPKCYMI
LLKSDRNTQGQIKNKRDSRGKT
>O_VN2R557
DAEHCMKCPEIEYPNKQRDRCLSKVVTFLNMKEPLGMILTLMALSLSLLTVLVL
VVFVMFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPTTATCLLQQTLFAVVF
TVAISSILAKTIIVVVAFRGISPESRIRMFIHPRMSIYVVLICSGIQVIFCGIWLGTYPP
FPEADAHSEYGHIIIRCNEGSAFAFYCVLGYMGFLALATFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWATFLPTYQSTKGKAMVTVEIFSILTSSAGLLGCIFIPKCYVI
LLKLEKNAQEQMKNKRDSRGK
>O_VN2R510
NAERCTKCSEVEYPNKQRDHCLSKVVTFLNMKEPLGKMLALMAVSFSLLTALV
LGVFVKFRDTPIVKANNRTLSYTLLVSLIFCFLCSLLFIGQPTTATCLLQQTVFAV
VFTVAVSSILAKTMIVVLAFRGVRPGSRIRTFLRPRVSNYVVLLCSGIQVIFCGTW
LGTSPPFPEADTHSEYGYIIIRCNEGSVLAFYCVLGYMALLALGSFTVAFLARNLP
DTFNETKLITLSMLVFCSVWTTFLLTYQSTKGKAMVTVEIFSILVSSVGLLGCIFIP
KCYVILLKAERNTLEQVKNK
>O_VN2R505
QSQMCKKCPEDKYPNKEKDLCLPKMVTFLDIKEPLGIMLAFVAVSFSFLTALVL
WIFVKFQDTPIVKANNQMLSYTLLISLIFCFLCSLLFIGHPTTTTCLLQQTVLGVVF
TVAVSSILAKTLIVILAFRSIRPGSRIQMFLSPRVSNYVVLICSGIQMIFCGIWLGTY
PPFPEADTHSEYGHIIIRCNEGSAFAFYCVLGYMGFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWASFLPTYQSAKGKAMVTVEIFAILASSAGLLGCIFIPKCY
VILLKPDRNTPNRFR
>O_VN2R504

DAEHCKKCPEDKYPNKQRDLCLPKAVTFLSVKEPLGMTLAFLAVSLSFLTALVL
WVFVKFKDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGQPTTTSCLLQQTVFGVVF
TVAVSSILAKTIIVILAFRGIRPGSRIQMFLHPRVSTYVVLICSGIQVIFCGIWLGTYP
PFPEADTHSEYGHIIIRCNEGSAFAFYCALGYMGFLALGSFTVAFLARNLPDTFNE
AKFITFSMLVFCSVWASFLPTYQSTKGKAMVTVEIFAILASSAGLLGCLFIPKCYV
ILLKSDRNTQNRSKIKRITGAR
>O_VN2R506
DAEHCKKCPEDKYPNKQRDLCLPKMVTFLNIKEPLGMTLAFVAVSLSFLTALVL
WVFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGQPTTTSCLLQQTVFGVVF
TVAVSSILAKTIVVILAFRGIRPGSRIQMFLHPRVSTYVVLICSGIQMIFCGIWLGTY
PPFPEADTHSEYGHIIIRCNEGSAFAFYCVLGYMGFLALGSFTVAFLARNLPDTFN
EAKFITFSMLVFCSVWASFLPTYQSTKGKAMVTVEIFAILASSAGLLGCIFIPKCY
VILLKPDRNNQNRSKIKR
>O_VN2R543
DAEHCTKCPEDEYPNMQKDFCLPKVLTFLSMKEPLGMTLAFIAVSFSLLTALVLG
VFVKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGHPNSITCLLQQTVFAFVFTV
AVSSILAKTIIVILAFKGIRPGSRIRMFIHPRVSNYVVLICSGIQVIFCGIWLGTYPPF
PEADTHSEYGQIIIRCNEGSTSAFYCVLGYMGFLALGSFTVAFLVRNLPDTFNEAK
FITFSMLVFCSVWASFLPTYQSSKGKAMVIVEIFSILASSAGLLGCIFIPKCYVILLN
REKNTQGRIKNKR
>O_VN2R514
DMDQCAKCPEDEHPNKERDHCFPKVVTFLDFTGSLGTALTFVAASFSLLTALVL
WVFVKFRDTPIVKANNRALSYTLLVSLSFCFLCSLLFIGRPTAATCLLRQTTFGVV
FTVAVACILAKTIMVVLAFRVTRPGSSIRVWVQSRVSNCIVLVCSGIQVMLCGFW
LGTSPPFPDMDLHSEPGHIIIQCNEGSDIAFYCVLGYMGFLALGSFTVAFLARSLP
DTFNEAKFITFSMLVFCSVWVSFLPTYQSTKGQAMVVVEIFSILASGAGLLGCIFL
PKCYVILLRPERNVLAGMKRKVNS
>O_VN2R511
NMKHCVTCPEDEYPNQERNRCLPKVVIFLNVSEPLGVTLICVAFSFSLLTALVLW
VFVKFRDTPIVKANNRALSYTLLVSLSFCFLCSLLFIGRPTAATCLLRQTTFGIVFT
VAVSSILAKTIMVVLAFRLTIPGSRGRMWVQSRVSTSVVIICSGIQIIICGIWLGIAP
PFPDTDTHSEPGHIIIGCNEGSVIAFYSVLGYMGFLALGSFTVAFLARNLPDTFNE
AKLITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCWV
ILLRPDRNTQERIKNKVSSRGRNHS
>O_VN2R516
NMKHCVTCPEDEYPNKERSQCLAKVVIFLNVSEPLGVTLICVAFSFSLLTALVLW
VFVKFRDTPIVKANNRALSYTLLISLIFCFLCSLLFVGRPSSVSCLLRQTTFGIVFTV
AVASILAKTIMVVLAFRLTIPGSRGRMWVQSRVSNSVVIICSGIQVTLCGVWLGIA
PPFPDTDTHSEPGHIIIGCNEGSVIAFYSVLGYMGFLALCSFTVAFLARNLPDTFNE
AKLITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCWV
ILLRPDSNTQERIRNKVDSRGK
>O_VN2R519
NVKHCVKCQEDEYPNKERNQCLPKVVTFLDVSDSLGVILISVALSFSLMTVLVL
WIFLKFQDTPIVKANNRTLSYTLLISLIFCFLCSLLFIGRPSSVSCLLRQTTFGIVFTV
AVSSILAKTIMVVLAFRLTTPGSRGRMWVQFRVSNSVVIICSGIQVILCGIWLGIAP
PFPDTDTHSEPGHIIIGCNEGSVIAFYSVLGYMGFLALCSFTVAFLARNLPDTFNEA

KLITFSMLVFCSVWVSFLPTYQSTKGKAMVTVEIFSILASSAGLLSCIFIPKCYVIL
LKPDRNTQERIKNK
>O_VN2R541
NMKLCTKCPEDEYPNKERNYCLPKVISFLDIKEPLGMALACIALSFSVLTALILW
VFVKFCDTPIVKANNRALSYTLLISLTFCFLCSLLFIGRPTVATCLLRQTTFGLVFT
VAVSCILAKTITVILAFKVTIPGRRKMIWLQPGVSNCVVVLCSGIQVILCGIWLGIS
PPFPDINTHSQPDQIIIECNEGSVLGFYCVLGYMGFLALGSFTVAFLARSLPDTFNE
ARFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCYVI
LLRSDENTKECIRNKS
>O_VN2R539
NMKLCTKCPEDEYPNKERDHCLPKVISFLDIKEPLGMALACIALSFSVLTALILW
VFVKFRDTPIVKANNRALSYTLLISLIFCFLCSLLFIGRPTAATCLLRQTTFGLVFT
VAVSCILAKTITVILAFKVTIPGRRKMIWLQPGVSNCVVLFCSGIQVILCGIWLGIS
PPFPDTNTHSQPDQIIIECNEGSVLGFYSVLGYIGFLALGSFTVAFLARSLPDTFNE
AKFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCYVI
LLRSDENTKECIRNKS
>O_VN2R542
SMKHCMKCPEEEYPSPERDRCLPKAITFLDIKEPLGMALACVALFFSVLTALILGV
FVKFRDTPIVKANNRTLSYILLISLIFCFLCSLLFIGHPTTATCLLRQTTFGLVFTVA
VSCILAKTIMVILAFKVTIPGRRKMIWLRPGVSNCVVLLCSGIQVILCGIWLGISPP
FPDTDLYSQPNQIIIECNEGSVLGFYSVLGYMGFLALGSFMVAFLARSLPDTFNEA
KFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCYVIL
LRPHENTQEYIRNKSGQHS
>O_VN2R533
NMKHCVKCPEDEYPNKERNHCLPKAITFLDIKEPLGMSLASVALFFSVLTVLILG
VFVKFRDTSVVKANNRALSYTLLISLTLCFLCSLLFIGRPTAATCLLRQTIFGVVFT
LAVSSILAKTILVVLAFKVTKPGSRIKVWLQPRVSNCIVVLCSGIQVILCGIWLGIS
PPFPDTDIHSEPNQIIVECNEGSVFGFYSVLGYIGLLALASFIVAYLARSLPDTFNE
ARFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCYVI
LLRPHNARVRSI
>O_VN2R530
NMKHCMMCPEDEYPNKERDRCLPKVITFLDIKEPLGMILACVALFFSVLTALILG
VFVKFRDIPIVKANNRTLSYTLLISLTSCFLCSLLFIGRPTAATCLLRQTTFGLVFTV
AVSSILAKTILVVLAFRVTKPGSRIKVWLQPRVSNCVVVLCSGIQVILCSSWLGISP
PFPETDIHSEPNQILIECNEGSVFGFYSVLGYMGFLALGSFIVAFLARSLPDTFNET
KFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKCYVIF
LRPHNARVRSI
>O_VN2R537
NMKHCTKCPEDAYPNKKRDRCLPKAITFLDITEPLGMALACVALFFSVLTALILW
VFVKFRVTPIVKANNQTLSYTLLISLTFCFLCSLLFIGRPTAATCLLRQTIFGVVFT
VAVSSILAKTILVVLAFKVTIPASRGRVWLQPRVSNCVVLLCSGVQVILCGSWLG
MSPPFPDTDTHSQPNQIIIECNEGSVLGFYSVLGYIGFLALGSFTVAFLARNLPDTF
NEAKFITFSMLVFCSVWVSFLPAYQSTKGKAMVVVEIFSILASSAGLLGCIFIPKC
YVILLRPNENNQEFIRKKSDCKSGKHSSRLSQGQRLHKAKEIPE
>O_VN2R527

NMKHCMKCPEDEYPNKERDRCLPKAITFLDIKEPLGMTLACVALFFSVLTALILG
VFVKFRNSPIVKANNRTLSYTLLISLTLCFLCSLLFIGYPTAATCLLRQTTFGLVFT
VAVSSILAKTILVVLAFRVTKPGRRSKRWLQPGVSNCIVVLCSGIQVILCGIWLGI
SPPFPDTDRYSQSNQIIIECNEGSVLGFYCVLGYMGFLALGSFMVAFLARGLPDTF
NEAKFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASSAGLLCFIFIPKCY
VILLRPNGNTQECIRNVRVQEGKA
>O_VN2R528
NMKHCMKCPEDEYPNKERNRCLPKTVTFLDIKEPLGMALTCVALFFSVLTALIL
GVFVKFRNSPIVKANNRTLSYTLLISLTSCFLCSLLFIGRPTAATCLLRQTTFGLVF
TVAVSSILAKTILVVLAFKVTKPGSRSKRWLQPGVSNCVVVLCSGIQVILCSSWL
GISPPFPDTDIHSHPNKIIIECNEGSVLGFYSVLAYMGFLALGSFMVAFLARSLPDT
FNEAKFITFSMLVFCSVWVSFLPTYQSTKGKAMVVVEIFSILASGAGLLGCIFIPK
CYVILLRPNENTQKCMRNKSEDKRGKHSV
>D_VN2R302
PSFLLLTALVLGVFVRHWDTPLVKASKQSLSFILLIFLTLCFLCSLLFTNCPTTVTG
LFQTTFGTVFTLALPTILVKTISVVLAFTTIGPESRARHCMGLGASNTFIPIYSLVPL
MLCGVWLVNSPCFPDTDPHPEPGLLIIKCKEGSATTFYWVPSYMSILASASFSIAF
AARNLSDTFREAKSMCRMLVFCSVWVSFLLTCQSTGKATMAVEVFILASIPGLL
GCIFAFQCYLILPRPDRNTQSWVKSKYDKKRINSE
>HS_VN2R2
PPTTATCLSRPSTLGIIFPGAVPTMQKMIVVVLVCKTVRPGSSIWQLIAPRASNSCI
FTCSLVHVMLYGVASEPSSLPRRRPLSEPGHISLEGSEGPVTASYCVGCFLAVVSF
SVTFAAGNLCDTFCEAKFLMFCMPVVPSDRVSFLLANLRTRGKAMGVREVSML
ASSPGLLGCVFAPKCYVTLLRPDQNLVIKKDGLRNSIFSGALTMFAQIMSISEKKG
ALLYLY
>CH_VN2R102
PPTTATCLSRPSTLGIVFPGAVSTMPEDDVVVLLCKTVRPGSSIWQLIAPRASNSCI
FTCSLVHVMLYGWLVNPPLSPDAGPSEPGHISLEGSEGPVTASYCVGCFLAVVSF
SVTFAAGNLCDTFCEAKFLTFCMPVVPSVQASFLLTNPHTRGKAVAVREVSILAS
SLGLLGCVFAPKCYITLLRPDQNLVIKKEGLRNSIFSGASTMFAQIVSISEKKGALL
YLY
>MQ_VN2R203
PPITATCLSRPSTLGIVFPGAVSTMQKMITVVLACKTVRPGRSIRQLIAPRASNSCI
FTCSLVHVMLYGWLVKSPLFPDASLSEPGHVSLECSEWPVTASHCVPGCFPVVV
SFSVTFAAGNLSDTFYEARFLMFCMPVVPSVQASFLLTNLCTRGKAMGAREVST
LASSLGLLGCVFDLKCYITLLRPDQN
>D_VN2R309
PKQRNQCLPEIMTFLSHENLLGAVLALAIGFSLPTVLILGLFIQYNDICIIRANN*NL
SYILLVSLMLCFFHSLIFIGWPSLVIFVLR*TLFGVIASIVISAILAKTSIVAVAFKTIK
PGNTMNVVGDLILKCYSLVGFLIQVCICMIWLGTSSSFPNIDHNQSGKVFLQCNE
VSSTVFYCVLGNLGFLAFLTLTTAFLGLRLPSSFSKPQTITFSMLIICSVWIAFIPTY
LSTEGKATAALETF
>RN_Vom2r51
DMDQCIKCPEDQYPNKQRNWCHPKITAFLSHEDTLGAVLVSVAISLSVFSAMILG
LFICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQTIFGVVFSI
AISAILAKTFIVVVAFKAIKPGSTLQTWMVTRVSNAIVCCGSIIQVCICAVWLGTY

PPFPDVDMHSEFGQIILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFN
EAKTITFSMLVFCSVWISFVPAYLSSKGKTMVAVEILTILASSAGLLGCIFLPKCFV
ILLKSGGLPRKKF
>RN_Vom2r50
DMDQCIKCPEDQYPNKQRNQCLPKITAFLSHEDTLGAVFISLAISFSAFSAMILGL
FICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQTIFGVVFSIA
ISAILAKTFIVVVAFKAIKPGSTLQTWMVTRVSNAIVCCGSSIQVCICAVWLGTYP
PFPDVDMQSEFGQIILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFNE
AKTITFSMLVFCSVWISFVPAYLSSKGKNVVAVEILAILASSAGLLGCIFLPKCFVI
LLRSDDLPRKKFFK
>RN_Vom2r48
DMDQCIKCPEDQYPNKQRNQCLPKITAFLSCEDPLGAVLVSVAISLSAFSAMILG
LFIHYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFICQPTTVTCILRQMIFGVVFSI
TISAILAKTFIVVVAFKVIKPGSTLQTWMVTRVSNAIVCCGSSIQVCICAVWLGTY
PPFPDVDMQSEFGQIILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDTFN
EAKTITFSMLVFCSVWISFVPTYLSSKGKTMVAVEILSILASSSTLLGCIFLPKCYVI
LMRSGGHSRKKIFK
>RN_Vom2r49
DMDQCIKCPEDQYPNRQRNQCLPKITEFLSHEDTLGAVFVSIAISLSALSAMILGL
FICYRDTSIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPRTVTCVLRQMIFGIVFSV
AVSTILAKTFIVVVAFKAIKPGSALKMSMVIRLSNTIVFCGSIIQVCICAVWLGTYP
PFPDVDMHSEFGQIILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNE
TKTITFSMLVFCSVWISFVPTYLSSKGKTMVAVEILSILASSAGLLGCIFLPKCYMI
LMSSGGHSRIKFFK
>RN_Vom2r52
DMDQCIKCPEDQYPNTQRNQCLPKITDFLSHEDTLGAVLVFVALSFSAFSAMILR
LFICYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPKTVTCVLRQVIFGVVFS
VAVSTILAKTFIVVVAFKAIKPGSALKMSMVIRLSNTIVFCGSIIQVCICAVWLGT
YPPFPDVDMHSEFGQIILWCNEGSTFAFYCVLGYLGFLASLSLLIAFLARRLPDSF
NEAKTITFSMLVFFSVWISFVPTYLSSKGKTMVAVEILSILASSAGLLGCIFLPKCY
VILMKSGDHSRKKFFK
>O_VN2R657
DATTCERCPDDSWPSPSRDRCELQPDVFLQFGDPLGAVLASLGVWLSLVTGVVL
LVFVRHRHTPLVRASSPGLSALLLTSLFLCSLSSLLFLGRRSRFDCQLRQVAFSFTF
TVAVSCVLARTAVVLMAFRAVQPASFLQMCLGPGLPWAMVAGPFVVQACICIV
WLWAAHDDRNTSLAPNTAPPECREALSPGFWVIMGLLALAGFVLALLTRQLPG
GFCEAHLLTLSMAICVSVWLASLPAQIHSRGRVAAAVQAFSILTSNAALLGFIFLP
RCFIILLRPHRNNR
>MQ_VN2R205
PCQKCSDNQWPNVQKRKCIPKTLDFLFYHEPLDTALAVCTALLFLLALAILGIFI*
HHHTPII*ANNCQLSYLLLSSLALSFLCPFKFIGHPDPITCAVHQAAFGVTFTVCMS
TVLAKTIVVVAAFHATQPDTQLRR*AGPVLPSTILTACSLTQAALCALWVTRWPP
QT*NSKEPGPTVTVKCDKGSLELFYAMLGYLSLLDLVSLLVTFPTCQLPDTFSEA
KHITLSVVCSCVGVSLIPAHMHAHSKDTTAVEVFFILASGGGLMSSLFLPKCYIIL
LHPEKNTKD
>CH_VN2R105

PDSVACQKCSDNQWPNVQKGQRIPKTLDFLFYHKPLDTALAVCTALLFLLALAI
LGIFV*HHHTPIIQANNCQLSYLLLSSLALSFLCPFMFIGHPDPITCAVHQADFGVT
FMVCTSTVLAKTIVVVAAFHATQADTQLRGWAGTVLLSTILCSLTQAALCALWV
TRWPPQPVKLYRTLATVTVKCDKGSLELLLELGYLSLLDLVSLLVTFPTCRLPDT
FNEAKHITLSMVCSCVWVSFIPAHMHAHSKDTMAMEVFVILASAGGLMSSLFFS
KCYIILLHPEKNTKDQMFGRHLRKWEKLK
>HS_VN2R9
PDSVACQKCSDNQWPNVQKGECIPKTLDFLFYHKPLDTALAVCTALLFLLALAIL
GIFV*HHHTPIIRANNCQLSYLLLSSLALSFLCPFMFIGHPDPITCAVHQADFGVTF
MVCTSTVLAKTIVVVAAFHATQADTQLRGWAGTVLLSTILCSLTQAALCALWV
TRWPPQPVKLYRTLATVTVKCDKGSLELLLELGYLSLLDLVSLLVTFPTCRLPDT
FNEAKHITLSMVCSCVWVSFIPAHMHAHSKDTMAMEVFVILASAGGLMSSLFFS
KCYIILLHPEKNTKDQMFGRHHRKWEKLK
>C_VN2R414
PTVCPALHSLLTLAVWGIFIGLHHTPII*ARKPQISYLLLSTLALCFPFPFMFIGPPGP
LTCAVHQAAFGVTSTTCMSTVLAKTIVVVAAFHGTRLHAQMQVLGPVLPKTIPM
VCSLIQATSCSLWVTMTPTTCDLYRARAHSDLKCDEGSLELL*ATLGYLGITCLV
SLLVVFLTHQLPDTFGETQRITLSTLVCSCIWVTFILTHRSARGRDTVPLEVSAILA
SEAGLLFCLFLPKCYIILLQPEQNTKEQMLGRQHLK
>O_VN2R615
DSASCLKCPEDQWPNNEKNQCIPKVLDFLSYNEPLGMALAICTTLFFCLALTILGI
FIWHHHTPIVRANNVRLSYLLLCSLALCFLCPFMFIGFPVPLTCALRQATFGIIFTV
CVSSTLAKTIIVVSAFRATRPDARLRRWGGLSLPNAIPIGFSLVQVTLCGFWVTG
WPPVPMNSTEAGPKVTMLCSAGSSELFYAMLGYLGFLALVSLVVAFLARQLPDT
FNEAKHITLSMLVFSCVWISFIPAHLSAQGKDTVAVEIFAILASGAGLMCCLFFPK
CYIILLQPDKNTKDQMMGKSHFK
>O_VN2R616
DSTSCLKCPEDQWPNSKQDQCIPKNLDFLSYKEPLGMALATCTTLFFLLSLTILVI
FIWHHHTPIVRANNQRLSYILLCSLALCFLCPFMFIGHPVPLTCALRQAAFGIIFTV
CVSAILAKTIIVVSAFRTTRPNARFRKWAGSNLPHAILIACSLIQIVLCMFWVIVWP
PMPMNSMEAGPKVTVMCDIGSSEFFYAMLSFLGFLALVSLVVAFLARQLPDTFN
EAKHITLSMLVFSCVWISFIPAHLSAQGKDTVAVEIFAILASGAGLLCCLFFPKCYI
ILLQPAKNTRDQMMGKHYFKK
>D_VN2R308
PGQRECEQCNEDYWSNAQKRMCVLKEVEFFAYDEALGDTLVIRSIFGVLVLAVT
IVYVIYRHTPLVNASDREPGFLIQLLVIPLLSSMLFIGKPYNWSCVTCQVTAIGFSL
CLSYILGKTILLF*AYRISNSKIQLISIHSHYREVFADLVLLEIGICLAYLVLELPILY
KNIESQNRKIILECDKGSIEFLCSMFGIDIFLALLCFLTTFVARLLPDNYYEGKCISF
GMLVFFIIWISFVPACLSTEGKFKGAVEIFAILTSSYGLLRCVFAPECFIVLLRTTRT
TDEIVGGRVSTIDKSIQLTSASVSSKLNSTESTVVDD
>MQ_VN2R201
PGQRECEQCDEDYWSNAQKSKCVPKQVEYLAYDEALGFTPVILSVFGASVVLAI
TTVYVIHRHTPLVNANDQQLSFLIQFSLIITLLSSMLFIGKPHNWSCMARQVTLAL
GFSLCLSCILGKTISLFLAYRISKSKTRLTSMHPLYRIIIVLISVLFEIDICTAYLVLEP
PRVYKNMEPQNAKIIPGCNEGSIEFLCSMFGIDVFLASLCFLTTFVTRQLPDNYYE

GKCVTFGMLVFFIVWISFVPAYLSTKGKFKVAV*IFAILAFSYGLLGCIFASKCFIIL
LRPQRNTDEIVGRRVSATDKSIQLTSASVSSELHTTVSIVLDD
>CH_VN2R101
PGERECEQCGEDYWSNVQKSECVLKDVEYLAYDEALGFTLVILSVFGAFVVLAV
TAVYVIHRHTPLVNANDWQLGFLIQVSLIIMLLSSMLFIDKPHNWSCMAGQVTL
ALGFSLCLSCLLGKTSSLFLAYRISKSKTQLTSMHPLYRKIIVLISVLAEIGICTAYL
ILEPPMVYKNMEPQNTKIILRCNEISIEFLYSMFGIDAFLALLCFLTTFVARQLPDN
YYEGKCITFGMLVFFIIWISFVPVYLSTKGKFKVAVEIFAILASSYGLLGCIFAPKC
LIILLRPERNTNEIVCGRVSTTDNCIQLTSAFVSSELNTTVSTVLDD
>HS_VN2R1
PGERECEQCGEDYWSNAQKSECVLKEVEYLAYDEALGFTLVILSVFGAFVVLAV
TAVYVIHRHTPLVNASDWQLGFLIQVSLIIMLLSSMLFIDKPHNWSCMAGQVTLA
LGFSLCLSCLLGKTSSLFLAYRISKSKTQLTSMHPLYRKIIVLISVLAEIGICTAYLIL
EPPMVYKNMESQNTKIILGCNEISIEFLYSMFGIDAFLALLCFLTTFVARQLPDNY
YEGKCITFGMLVFFIIWMSFVPVYLSTKGKFKMAVEIFAILASSHGLLGCIFAPKC
LIILLRPERNTSEIVCGRVSTTDNCIQLTSAFVSSELNTTVSTVLDD
>C_VN2R407
PGQRECDQCGEDYWSNPQKSECVLKEVEFLAYDEALGFTLVILSIFGALLVLAVT
TVYVIHRHTPLVMANGRELSFLIQVSLVITLLSSMLFISKPYNWSCMACQITLVLG
FSLCLSCILGKTISLFSAYRISKSKTQLISIHPLYRKIIVLTSVLVEIGVVASYLVLNP
PRVYKNMESQNIKIILECNEGSIEPLCSMFGMDVFLALLGFLTTFVAHQLPDNYY
EGKSITFGMLVVLIIWIVPAYLSTKGKFKVAVETFAILASSYGLFGCTFAPKCFIIL
LRPKKNMNEIVGGRVPALHKSIQLTSASVKSKVNTSVSTVLDD
>O_VN2R653
GIVGMRKCHEDYWSNAAKDDCVLKEVEFLGYDEALGFTLVTLSIFGALVVVAV
TVVYVIYRKTPLAKANDRELSFLLQLSLIITLLTSLLFIGKPYNWSCMARQVTLAM
GFSLCLSCILGKTIALFLSYRVSKSKVRFTSIRPLYQKLIVLFSVLIESALCTAYLVM
MPPKVYKNMESQNIKIILECNEGSIEFLCSVFGFDAFLALLSFLTTFVARKLPDNY
NEAKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSFGLLGCVFAPKC
YIILLRPKRNTDEIVGGKVPTTDRSIQLTSASVSSEMNVTVSTVVLD
>RN_Vom2r46
GQRLCDPCGEDNWSNPQKNKCVPKLVEFLAYEEALGYALVILSIFGALVALAVT
VVYVIHRHTPLAKANDQELSFLIQMSLVITVLSSILFIGKPYNWTCMARQVTLAL
GFCLCLSSILGKTISLFFAYRISKSKTRLISMHPIFRKLIVLICVVGEIGVCTAYLML
KPPRMVKNIEPQNVKIIFECNEGSIEFLCSIFAFDVLLALLCFLTTFVARKLPDNYY
EGKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCIFLPKCFII
LLRPKRNTDETVGGRVPTVDRSIQLASTSVSSELNTTVSTVLDE
>RN_Vom2r47
GQRECDPCSEDDWSNAEKSKCVPKLVEFLSYEEALGFTLVILSIFGALVVMAVTV
VYVIHRHTPLVKANDRELSFFIQMSLVITVLSSMLFIGKPCNWSCMARQVTLALG
FCLCLSSILGKTISLFFAYRISKSKTRLISMHPIFRKLIVLICVVGEIGVCTAYLVLEP
PRMFKNIEPQNVKIIFECNEGSIEFLYSIFGFNVLLALLCFLTTFVARQLPDNYYEG
KCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCVFLPKCFIILL
RPKRNTDETVGGRVPTLDKSIQLTSASVSSELNTAESIVLDE
>RN_Vom2r44

GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYEEALGLTLVILSIFGALVVLAVTV
VYVIYRHTPLVKANDRELSFLIQMSLVITVLSSMLFIGKSCNWTCMARQVTLALG
FCLCLSSTLGKTISLFFAYRISKSKTRLISMSPIYRKLIVLICVVGEIGVCTAYLVLE
PPRMFKNIEPQNVKIIFECNEGSIEFLCAIFGFDVFLALLCFLTTFVARKLPDNYYE
GKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCIFLPKCFIIL
LRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTAVSTVLDE
>RN_Vom2r45
GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYEEALGFTLVILSIFGAIVVFAVTIV
YVIHRHTPLVKANDRELSFLIQMSLGITVLSSMLFIGKPSNWTCMARQVTLALGF
CLCLSSTLGKTISLFFAYRISKSKTRLISMHPVFRKLIVLFCVVGEIGVCAAYLVLE
PPRMFKNIESQNVKIIFECNEGSIEFLCAIFGFDVFLALLCFLTTFVARKLPDNYYE
GKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCIFLPKCFIIL
LRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>C_VN2R409
PDTHRCVRYTDHQFSNSQDHCLPKKVTFLDSQDPLRMALASTALCCSVLTAVLL
GVSVKHETLDHQGQQLTLSYIPLTFLLLAFFCLLSISPPNTGTCIAQQVTFPFVFTV
VISTV*AKIMTVVLAKNKGRRIRQWFISGALTLSLPSCLIQVRIRVRLGLSRFFTEM
NTHSESTHNIAGCTKGSVTAFYCVLECMGSLALGSTAVDSPVRNLPDTLGEAKF
GMRVSCSVWVTSFPVCHSTMVTMQTFFFSASSAGLPGCIFATRV
>D_VN2R305
PPYHQRSHSNQDHCLPKKATLLAYRPTGDSSACPDLCVSVPVAVVLGVAVKHQ
DIPLSRQQPTLSYLLLTSFPGFPSSFSTGPPNKGTYVFPQMMLASGFSGATSAASA
KIITVVLALMTTVPRKRRRQGCYQEHLAVLPICALIHVVTCGVLETASPFVEMST
HSTSAHITTGCSEASATAFCYVQGHLGSFALGSIAVARTPSDNLRGSQVSMRACC
SAWVTFLPVHHSTSGKADVAMEILSLLASSADRPGCLLAP
>HS_VN2R11
PNQSWALLALEDPLGMAVVSTALSFSALTAVVPVVFVKHGDSPSQGQRLLSCVI
SLTPSFLCSLLFTGRPNTATCPPANHMVVCTVAASTDLPSSLTVVLAFKAIAPGRR
LRCWLVSGPLHLSHLLLIQVTVCGVWLGTSPPSTARGSHSEPRYILVVCVLGYLS
SLALGSFSVTFLARNLSDTCVEQVPDVEREFCSVWVTFLPIRQGQVSSILAPRVGI
PGCIFAPR
>CH_VN2R107
PDADQCMRGPAGQHRNSERPLAPKQSWALLALEDPLGMAVVSTALSFSALTAV
VPVVFVKHGDSPSQGQRLLSCIISLTPSFLCSSLFTGRPNTATCPLQTTRVVGTVA
ASTDLPSSLTAVLAFKAVAPGRRLRCWLVSGPLHLSHLLLIQVTVCGVWLGTSPP
STARGSHSEPRYILVVCVLGYLSSLALGSFSLTFLARNLSDTCVEQVPDVEREFCS
VWVTFLPIRQGQVSSILAPRVGIPGCIFAPR
>MQ_VN2R207
PLVSTALSFSALTGVVPAVFLKHRLPRAKSRPTALSYVVSLTPSFLCSFLFTGCPG
TATCSPANHTSRVHCSCVQRLASVLTVVLVFKAIAPGRRLRCWLVSGHLGLSRS
LVIQVTVCGVRLATSPPSTARGSHSEPRYTLVVCVLGYLSSLALGSFSVTFLARNL
SDIRMEQLPDVQLEFCSVWVTFLPSFLPSR
>MQ_VN2R211
PRHERCEKCPDDQYQNTQRIYCLPKDVNFLTFEDPLGMALACTACFSVLTTMVL
GVFVKHQEIPSVKASTWTFRYILLKSLTLFFLCSALFIGHPPTQLPGLQQIIFRVGFI
ILTKTITVVLALKARMPGRIRW*LVSGSPNFVLFMFSLIQVLICRLGLRNFAPFIDT

DHSKLGSVVLTRNEGFSTAFYHVLGYLGSLTLGSFTLAFLAMGLSDTFNKSKFL
MFSMLVFCSAWVTSLPVAHSTKRKVLESMEVFFILASSAEVLGCIFAPKCYIILIRS
DMHTLQ
>CH_VN2R106
PDTIQYKKCPDDQYQSTQRIYCLPKDVNFLAFEGPLGMSLACTACLSVLTAMVL
GVFVKQQEIPSVKASTWAFSCTLLSLTLCFLCSALCIGHPRTATW*LQQTAFGVV
FTMAISIILTKTITVVLALKARMPGRIRW*LVSGSPNFVLFIFSLIQVLICQLGLRNF
PPFIDTDTLSKLGSVVLTRSEVSALFYHVLGYLGSLTLGSFTLAFLAMGLPDTFNK
KFLMFSMLVFCVTSLLADHSIKGKVLEATGFFFILAFSAELLGCIFAPIYYITLIR
>HS_VN2R10
PDTIQYKKCPDDQYQNSQRIYCLPKDVNFLAFEGPLGMSLACTACLSVLTAVVL
GVFVKQQEIPSVKASTWAFSCTLLSLTLCFLCSALCIGHPHTATW*LQQIAFGVVF
TMAISIILTKTITVVLALKARMPGRIRW*LVSGSPNFALFIFSLIQVLICQLGLRNFP
PFIDTDTLSKLGSVVLTRSEVSALFYHVLGYLGSLTLGSFTLAFLAMGLPDTFNK
KFLMFSMLVFCVTSLLADHSIKGKVLEAMGFFFILAFSAELLGCIFAPIYYVTLMR
YE
>CH_VN2R103
PNLIECAKCPDNQYSVREITALQVVTFLDFEDALGTALASIALCFSVLTAIDMVSL
CPHPNLILNYSSHNSHVLLSATSLSVFFCSLLFTGCTNTVTCIL*QTTFGIIFTVVVS
TVLAKTITVLIACKATASERKPKWFWGPGVFSTVIFICFLIQVTLWNLMGSKSAIR
TLNDEVFFAFHSVLGFLTFLALGSFILAFLARSLPDTFNEAKFLMFSTLMLCTV*IT
FQTV*HSTKGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPDKNTLGELREKSKF
>HS_VN2R8
PHRVVTFLDFEDALGTALASIALCFSVLTAIVLMIFVKHIHTVIVKTLSDISLSVFF
CSLLFTGCTNTVTCIL*QTTFGIIFTVVVSTALAKTITVLIACKATASERKPKWFWG
LGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEVFFAFHSVLGFLTFLALGSFILA
FLASSLPDTFNEAKLLMFSMLMLCTV*ITFQTV*HSTKGKVMVALEIFSTLTSSAV
LGGIFAPRGRIILIKPGKNTLGELGEKSKF
>HS_VN2R5
PDSIECAKCPDNQYSVERSLPHRVVTFLDFEDALGTALASIALCFSVLTAIVLMIF
VKHIHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCIL*QTTFGIIFTVVVSTALAKTIT
VLIACKATASERKPKWFWGLGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEVFF
AFHSVLGFLTFLALGSFILAFLASSLPDTFNEAKLLMFSMLMLCTV*ITFQTV*HST
KGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPGKNTLGELREKSKF
>HS_VN2R4
PDSIECAKCPDNQYSVERSLPHRVVTFLDFEDALGTALASIALCFSVLTAIVLMIF
VKHIHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCIL*QTTFGIIFTVVVSTALAKTIT
VLIACKATASERKPKWFWGLGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEVFF
AFHSVLGFLTFLALGSFILAFLASSLPDTFNEAKFLMFSMLMLCTV*ITFQTV*HST
KGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPGKNTLGELREKSKF
>HS_VN2R3
PDSNECAKCPDNQYSVERSLPHRVVTFLDFEDALGTALASIALCFSVLTAIVLMIF
VKHIHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCIL*QTTFGIIFTVVVSTALAKTIT
VLIACKATASERKPKWFWGLGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEVFF
AFHSVLGFLTFLALGSFILAFLASSLPDTFNEAKFLMFSMLMLCTV*ITFQTV*HST
KGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPDKNTLGELREKSKF

>CH_VN2R104
PDSNECAKCPDNQYSVERSLPHRVVTFLDFEDALGTAMASIALCFSVLTAMVLM
IFVKHKHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCIL*QTTFGIIFTVVVSTALAK
TITVLIACKATASERKPKWFWGPGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDE
VFFAFHSVLGFLTFLALGSFILAFLARSLPDTFNEAKFLMFSRLMLCTV*ITFQTV*
HSTKGKVMVALEIFSTLTSSAVLGGIFAPRGHIILIKPDKNTLGELREQSKF
>HS_VN2R6
PDSNECAKCPDNQYSVERSLPHRVVTFLDFEDALGTALASIALCFSVLTAMVLMI
FVKHIHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCNL*QTTFGIIFTVVVSTALAKT
ITVLIACKATASERKPKWFWGPGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEV
FFAFHSVLGFLTFLALGSFILAFLARSLPDTFNEAKFLMFSMLMLCTV*ITFHTV*H
STKGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPDKNTLGELRKKASSYMYKY
LISGLRKIKKRNVKNNCYLHA
>HS_VN2R7
PDSNECAKCPDNQYSVERSLPHRVVTFLDFEDALGTALASIALCFSVLTAMVLMI
FVKHIHTLIVKTLSDISLSVFFCSLLFTGCTNTVTCNL*QTTFGIIFTVVVSTALAKT
ITVLIACKATASERKPKWFWGPGVFSTVIFICFLIQVTLWNLMGSKSAIRTLNDEV
FFAFHSVLGFLTFLALGSFILAFLARSLPDTFNEAKFLMFSMLMLCTV*ITFHTV*H
STKGKVMVALEIFSTLTSSAVLGGIFAPRGRIILIKPDKNTLGELRKKASSYMYKY
LISGLRKIKKRNVKNNCYLHA
>MQ_VN2R206
PDSNECTKCPDNQYSVERSLPHRVVTFLDFEDALGTALASTALCFSVLTAMVLV
VFVKHIHTPIVKTLSYLFVFLCSLLFIGCTNTVTCIL*QRTFGIIFTVVVSTALAKTIT
VLACKATASERKPRWFWRPGVFSTVIFIYFLIQVTFWNGMGPKSAIRALNDEVFF
AFHSVLGFLVFSARGSFVLAFLARSLPDTFNEAKFLMFSMLMLCSWITFLPV*HST
KGKVMVALEIFSILTSSAGLGSIFAPRGHIILIKPDKNTLGELREKSKF
>D_VN2R301
PNMKQCMKCPDNQYPNSDGHCLLKVVIFLGPEDLGLVLACAALTTLVLGIFVKY
REIPVKANNQSLSYILFIYLILCSLCSLLFLGHPITATCILQQMTLGVLFTVAISTILA
KKISVVLAFKITTPGRKVRQWLLLGATKSAIPIYSLILVCSKSSVIMFYHSLGYSGP
LALASFTLAFLARSLPDTSTETKFLTFSILMFSSA*ITFLPVYNSTKGKVTAAVDIFS
IWASSAGLLGTIFASKCWNILPKPTKKNLPRLSGKTNSRRGAPGWLSQLNSLLIW
LRIMRSSPT
>RN_Vom2r77
DLAQCVKCPDGQYANIERNHCFPKAVTFLDYADPVGIVLVCMALCLSTLSALIL
GIFVKKQNTPIVKANNRTLSYILLVSLTLCFLCSLLFIGHPSTARCVLQQITFGTLFT
VAVSTILAKTITVLLAFSITVPGRRVTWPLFVASKVIIPICSLFQLMLCGLWLGLSP
PFVERDAHSEYGHIIIMCNKGSIIAFYCVLGYLGCLALASFIVAFLGRNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVFHSTKGKVMVAVEVFSILTSGAGLLGCIFIPKCYII
VLRQDRNFIKKVKGKNHS
>RN_Vom2r71
DMEQCVRCPDDKYANLEQTYCLQRVVSFLAYEDPLGMALGCTALCFSTLTVLIL
FTFLKYKDTPIVKANNCLLSYILLISLIFCFLCSLLFIGHPNHITCILQQTTFGVFFTV
AISTVLAKTITVVMAFKISTPGNRMKGMLTSGAPNLVIPICTLIQLVLCGIWLGTSP
PFIDRDIQSEPGKTIVICNKGSIFAFHFTLGYLDSLTLMSFAVAFLARNLPDRFNEA

KFLTFSMLVFCSVWITFLPVYHSTKGKVMVVVEVFSILASSAGLLGCIFVPKCYVI
LVRTDSNFLQKYKDK
>RN_Vom2r68
DMEQCVRCPDDKYANLEQTYCIQRHVSFLAYEEPLGLTLGCMALCFSALTILVL
VTFVKHKNIPVVKANNCILSYILLISLIFCFLCSLLFIGHPKKTTCILQQITFGIFFTM
AISTVLAKTITVVMAFKLTTPGRRMRRILASGAPNLIIPICTLIQLVICGTWLATSPP
FIDRDIQSEHGKVIIICNKGSVTAFHFVLGYLRSLALGSFAVAFLARNLPDRFNET
KFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCVFIPKCYVI
LVRPNSNFLQKYKNK
>RN_Vom2r70
DMEQCARCPDDKYANSEQTYCLQRAVTYLAYEDSLGMALGCMALFFSVTTILV
LVTFVKYKDTPIVKANNRILSYILLISLIFCFLCSLLFIGHPNKATCILQQITFGVFFS
VAVSAVLAKTVTVVMAFKLTTPGRRMRGMLALGAPNLVIPICTLIQLVLCGIWL
VISPPFIDRDIQSEHGKTIIICNKGSVIAFHFTLGYLGTLALMSFAVAFLARKLPDRF
NEAKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVPK
CYVILVRPDLNFLQK
>RN_Vom2r73
DMEQCARCPDDKYANSEQTYCLQRTVSFLAYEDPLGMALGCMALFFSAITILVL
VTFVKYKDTPIVKANNCILSYILLISLILCFLCSLLFIGHPNKATCILQQTTFGVFFT
VAVSAVLTKTVTVVMAFKLTTPGRRMRGMLASGTPNLVIPICTLIQLVLCGIWLV
ISPPFIDRDIQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSITVAFLARNLPDRFN
EAKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVPKC
YVILVRPDLNFLQK
>RN_Vom2r74
DMEQCARCPDDKYANSEQTYCLQRTVSFLAYEDPLGMALGCMALFFSAITILVL
VTFVKYKDTPIVKANNCILSYILLISLIFCFLWSLLFM?HPNKATCILQQTTFGVFFT
VAVSAVLAKTVTVVMAFKLTTPGRRMRGMLASGATNLVIPICTLIQLVLYGIWL
AISPPFIDRDIQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSITVAFLAR
>RN_Vom2r72
DMEHCVRCPDDKYANSEQTYCLQRAVSFLAYEDPLGMALGCVALSFSAITILVL
VTFVKYKDTPIVKANNCILSYILLISLIFCFLCSLLFIGQPNKATCILQQTTFGVFFT
VAVSTVLAKTVTVVMAFKLTTPGRRMRGMLSSGAPNLVIPICTLIQLVLYGIWLV
ISPPFIDRDIQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSIIVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVPKCY
VILVRPDLNFLQK
>RN_Vom2r67
DMEQCVRCPDDSYTNLEQTQCLQRTVSFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQDTCILQQTTFGVFFT
VAISTVLAKTITVVTAFKLTTPGRMIRRRMIKVATNFLIPICTLIQLALCGIWLGTS
PPFIDRDIQSEHGKVIIICNKGSIIAFYFVLGYLGSLALGSFTVAFLARNLPDRFNEA
KFLTFSMLVFYSVWITFLPVYHSTRGKVMIVVEVFSILASSAGLLGCIFVPKCYVL
LVRPDLNVLQK
>RN_Vom2r66
DMEQCVRCPDDSYANLEQTKCLQRAVTFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYRDTPIVKANNRILSYILLISLAFCFLCSLLFIGHPNQATCILQQTTFAIFFT
VAISTVLAKTITVVTAFKLTTPGRRIRKMMIKGATNLLIPICTLIQLVLCAIWLVTS

PPFIDRDIQSEHGKVIIICNKGSVTAFHFVLGYLGSLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTKGKVMVVVEVFSILASSAGLLGCIFVQKCY
VLLVRPDLNVLRK
>RN_Vom2r5
DMEQCVKCPDDSYANLEQSQCLQRVVTFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLYIGHPNQATCILQQTTFAIFFT
VAISTVLAKTITVVTAFKLTTPERRIKKIMIKGAINLLIPICTLIQLVLCAVWLVISPP
FIDRDIQSEHGKVIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRFNEAK
FLTFSMLVFYSVWITFLPVYHSTRGKDMVIVEVFSILASSAGLLGCIFFQKCYVLL
VRPDLNVLRK
>RN_Vom2r65
NMEQCVRCPDYSYANLEQTKCLQRVVTFLAYEDPLGMALGCLALSFSAITILVL
VTFVKYRDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQATFAIFFT
VAISTVLAKTVTVVTAFKLTTPGRRIRKMMIKGATNLLIPICTLIQLVLCGIWLVTS
PPFIDRDMQSEHGKIIIITCNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVLEVFSILASSAGLLGCIFIQKCYV
LLLRPDLNVLQK
>RN_Vom2r2
DMEQCVRCPDYSYANLEQTECLQRVVTFLAYEDPLGMALCCMALSFSAITILVL
VTFVKYRDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQATFAIFFT
VAISTVLAKTITVVTAFKLTTPGRRIRKMMIKGATNLLIPICTLIQLVLCGIWLVTS
PPFIDRDIQSEHGKIIITCNKGSVIAFHVVLGYLGSLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFIQKCY
VLLLRPDLNVLQK
>RN_Vom2r3
DMEQCVRCPDDSYANLEQTECLQRVVTFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYRDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQATFAIFFT
VAISTVLAKTITVVTAFKLTTPGRRIRKMMIKGATNLLIPICTLIQLVLCAIWLVTS
PPFIDRDIQSEHGKVIIICNKGSVTAFHFVLGYLSSLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVIVEVFSILASSAGLLGCIFFQKCYV
LLVRPDLNVLQK
>RN_Vom2r1
DMEQCVRCPDDSYANLEQTQCLQRAVTFLAYEDPLGMALGCMALSFSAITVLV
LVTFVKYRDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQATFAIFF
TVAISTVLAKTITVVTAFKLTTPGRRIRKMMIKGATHLLIPICTLIQLVLCAIWLVT
SPPFIDRDIQSEHGKVIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRFN
EAKFLTFSMLVFYSVWITFLPVYHSTRGKVMVIVEVFSILASSAGLLGCIFVQKCY
VLLVRPDLNVLQK
>RN_Vom2r4
DMEHCVRCPEDSYANLEQTECLQRVVTFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYRDTPIVKANNRILSYILLISLAFCFLCSLLFIGHPNQATCILQQTTFAIFFT
VAISTVLAKTITVVTAFKLTTPGRRIRKMMIKGAINLLIPICTLIQLVLCAIWLVTSP
PFIDRDIQSEHGKIIIIICNKGSVIAFHIVLGYLGSLALGSFTVAFLARNLPDRFNEAK
FLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVQKCYVL
LVQPDLNVLQK
>RN_Vom2r6

DMEQCVRCPDDSYANLEQTQCLQRAVTFLAYEDPLGMALGCMALSFSAITTLVL
VTFVKYRDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFAIFFT
VAISTVLAKTITVVTAFKLTTPGRRIRRMMIKGATNLLIPICTLIQLVLCGIWLVTS
PPFIDRDTQSEHGKIVIICNKGSDIAFHFVLGYLGSLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASCAGLLGCIFVQKCY
ILLFRTDLNFLQK
>RN_Vom2r69
DMEQCVRCPDDMYANLEQTQCLQRTVSFLAYDDPLGIALGCMALSFSAITILVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVFFT
VAISTVLAKTITVVMAFKLTTPGRRLRGMMITWASNFVIPFCTLIQLVLCGIWLVT
SPPFIDRDIQSEHGKTIIICNKGSVIAFHVVLGYLGSLALGSFTIAFLARNLPDRFNE
AKFLTFSMLVFYSVWITFLPVYHSTRGKVMVVVEVFSILASSADLLGCIFVQKCY
IILVSSNSNFLHK
>RN_Vom2r56
DSEECVKCPEDWYADTEQTHCLPKAVTFLAYEDPLGMALVCLALCFSVLTALVL
GVFLKHRETPIVKANNRTLSYILLTALIFCFLCSLLFVGNPNTTTCILQQTTFGILFT
VSVSTVLAKTLTVVLAFRITLPGRRMRGLMVSRIPKYIIPICTVVQIILYGIWLWTS
PPFVDIDTHSEHGCITITCSKGSVMAFYCVLGYLGSLALVSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWITFLPVYHSSKGKVMVAVEVFSILASSMGLLGCIFIPKCY
IILLRPDRNSIQKFRDKTQMNQ
>RN_Vom2r57
DSEECVKCPEDWYADTEQTHCLPKAVTFLAYEDPLGMALVCLALCFSVLTALVL
GVFLKHRETPIVKANNRTLSYILLTALIFCFLCSLLFVGNPNTTTCILQQTTFGILFT
VSVSTVLAKTLTVVLAFRITLPGRRMRGLMVSRIPKYIIPICTVVQIILYGIWLWTS
PPFVDIDTHSEHGCITITCSKGSVMAFYCVLGYLGSLALVSFTVAFLARNLPDTFN
EAKFLTFSMLVFCSVWITFLPVYHSSKGKVMVAVEVFSILASSMGLLGCIFIPKCY
IILLRPDRNSIQKFRDKTHS
>RN_Vom2r61
GKDQCVKCTDDQYANTDQTLCIQKAVTFLTFENPLGMTLASMALCFSILTSFVL
GIFVKHQNTPVVKANNSTLSYILLISLTFCFLCSFLFIGQPNAVTCILQQTMFGVVF
TVAVSTVLAKTITVILAFKSTAPGRRLRRLLISGAPKFIIPICTLIQTVLCGIWLGISP
PFVDTDAHSEHGHIIIMCNKGSVTAFYCVLGYLGALSLVSFTVAFMARNLPDTFN
EAKFLTFSMLVFLSVWITFLPVYHSTKGKIMVAVEVFSILASSTGLLGCIFIPKCYII
LARPDMNSLKSRRNRVQ
>RN_Vom2r62
GKDQCVKCTEDQYANTDRTVCIQKAVTFLALESPLGMTLASMALCFSILSSFVLG
IFVKHQNTPVVKANNSTLSYILLISLTFCFLCSFLFIGQPNAVTCILQQTTFGVVFT
VAVSTVLAKTITVILAFKSTAPGRRLRRLLISGAPKFIIPICTLIQTVLCGIWLGISPP
FVDTNAHSEHGHIIIMCNKGSVIAFYCVLGYLGVLSLVSFTVAFMARNLPDTFNE
AKFLTFSMLVFLSVWITFLPVYHSTKGKIMVAVEVFSILASSTGLLGCIFIPKCYIIL
ARPDMNSSKCQRNRVQ
>RN_Vom2r64
GKRQCVKCTEDQYANTDRTVCIQKAVTFLALENPLGMTLASMALCFSILTSFVL
GIFVKHQNTPVVKANNRTLSYILLISLTFCFLCSFLFIGQPNAVTCILQQTTFGVVF
TVAVSTVLAKTITVILAFKSTAPGRRLRWLLISGAPKFIIPICTLIQTVFCGVWLRIS
PPFVDTDAQSEHGHIIIMCNKGSVTAFYCVLGYLGALSLVSFTVAFMARNLPDAF

NEAKFLTFSMLVFLSVWVTFLPVYHSTKGKIMVAVEVFSILASSTGLLGCIFIPKC
YIILARPEMNSLKCRRNRVQ
>RN_Vom2r63
GKCVKCTEDQYANTDRTVCIQKAVTFLAFENPLGMTLASMALCFSILTSFVLGIF
VKHQSTLVVKANNSTLSYILLISLTFCFLCSFLFIGQPNAVTCILQQTTFGVVFTVA
VSTVLAKTTTVILAFKSTAPGRRLRGLLISGASKFIIPICTLIQIVLCGIWLGISPPFV
DTDAHSEHGHIIIMCNKGSVTAFYCVLGYLGALSLVSFTVAFMARNLPDIFNEAK
FLTFSMLVFLSVWITFFPVYHSTKGKIMVAVEVFSILASSTGLLGCIFIPKCYIILAR
PDMNSLKCRRNRVQ
>RN_Vom2r79
KMDDCVKCPEDHYANTVRTHCLRKAVIFLTYKNSLGMTFSGIALCLSLLTTLVL
GIFVKHQKTPVVKANNITFSYILLISLIFCFLCSFLFIGEPNTPICILQQTVFGIVFTV
AVSTVLAKTITVVLAFKISSPGRRMKGLLILLAPKFIILICTLIQIFFCGIWLGTSPPF
VDKDSKSEHGKIIILCNKGSVTAFYCVLGYLGSLSIVTFTVAFLARNLPDMFNEA
KFLTFSMLIFCSVWITFLPVYHSSKGRIMVVLEVLSILTSSTGLLGCIFFPKCYIILL
RPDRNSLKVMKD
>RN_Vom2r21
NIDLCARCPEDQYPNTEQSHCIPKAVVFLSYEEPLGMAISLISLCFSAFTTVVLGIF
MKHHSTPIIKANNCTLTYILLISLIFCFLCPLLFIGHPNSATCILQQITFGVVFSVSVS
TVLAKTITVVLAFEITASQRMMKYILVSRASNYIILVCTLIQVILCAVWLGASPPSV
DIDAQSEHGHIIIVCHKGSVTAFYCVLGYLGTLAFGSFTLAFFSRNLPGAFNEAKSI
TFSMLVFCSVWVTFIPVYHSSKGKMMVAVEIFSTLASSAGMLGCIFVPKCYTVLF
RPEQNSLQMVRVKSSSHVHIS
>RN_Vom2r19
NIDLCAQCPEDQYANTEQSHCIPKAVVFLSYEEPLGMAISLISLCFSAFTTVVLGV
FMKHHSTPIIKANNCTLTYILLISLIFCFLCPLLFIGHPNSAICILQQITFGVVFTVSV
STVLSKTITVVLAFEVTASQRMMKYILVSGASNYIILVCTLIQVTLCAVWLGASPP
SVDIDAQSEHGHIIIVCHKGSVTAFYCVLGYLGTLAFGSFTLAFFSRNLPGGFNEA
KSITFSMLVFCSVWVTFIPVYHSTKGKMMVAVEIFSTLASSAGMLGCIFVPKCYT
VLFRPEQNSLQMVRVKSSSHVHIS
>RN_Vom2r17
NVELCVQCPEEQYANQEQNHCIHKARIFLSYDEPLGMVLSLMALCLAALTVVVL
GVFVKHHRTPIVKANNCTLTYILLIALIFCFLCPLFFIGHPNSATCILQQITFGVVFT
VAISTVLAKTTTVILAFRITAPHRMMKYFLISRATNYIIPICTLIQIIVCAIWLGASPP
SVDIDAQSEHGHIIIVCNKGSVTAFYCVLGYLACLAFVSFTLAFLSRNLPVTFNEA
KSMTFSMLVFCSVWVTFLPVYHSTKGKVMVVVEIFSTLASSAGMLGCIFVPKCY
TILCRPDRNSLQMIKEKSSSHTHI
>RN_Vom2r18
NVELCVQCPEDQYANQEQNHCIHKARIFLTYDEPLGMALSLMALCLAALTVVVL
GIFVKHHRTPIVKANNCTLTYILLIALIFCFLCPLFFIGHPNSATCILQQITFGVVFTL
AISTVLAKTTTVILAFRVTAPHRMMKYFLVSRASNYIIPICTLIQIIVCAIWLGASPP
SVDIDAQSEHGHIIIVCNKGSITAFYCVLGYLACLAFVSFTLAFLSRNLPVTFNEAK
SMTFSMLVFCSVWVTFLPVYHSTKGKVMVVVEIFSTLASSAGMLGCIFVPKCYTI
LCRPDRNSLQMIKEKSSSHTHI
>RN_Vom2r15

NVEMCILCPEDQYANTEQNHCIQKPLIFLTYEEPLGITISLMALCFSALTAVILGVF
VKHHSTPIVKASNHILTYILLISLIFCFLCPLLFIGHPNSATCILQQITFGVVFTVAIST
VLAKTITVILAFKVTAPQRMMKYFLVQRASNYIIPMCTLIQVIVCAVWLGASPPF
VDIDTQSEHGHIIIVCNKGSVTAFYCVLGYLACLSFASFFLAFFSRNLPGNFNEAK
SMTFSMLLFCSVWITFLPVYHGTKGKVMVAVEIFSTLASSAGMLGCIFIPKCYTIL
FRPDKNSLQMIREKLSSHTHIS
>RN_Vom2r14
NVEMCILCPEDQYANTGQNHCIQKALIFLTYEDPLGRVLSLMALGFSAFTALVLG
IFLKHHHTPIVKANNSTLTYILLISLIFCFLCPLLFIGLPDLATCILQQITFGVVFTVAI
STVLAKTITVILAFKVTAPQRMMKYFLVSRVPNYIIPICTLLELLVCAVWLGAFPP
SVDIDDQSEHGHIIIICNKGSVTAFYCVLGYLACLAFVSFTLAFLSRNLPSTFNESK
SITFSMLVVCSVWVTFLPVYHSTKGKVMVAVEIFSILASSSGMLLCIFVPKCYIILF
RPDRNSFQMIREKSSSHTHIS
>RN_Vom2r12
NMEMCVLCPEDKYANTGQNDCIPKALIFLSYEDHLGMALSLMALGFSAFTAVVL
GIFVKHHSTPIVKANNRTLTYILLLSLIFCFLCPLLFIGHPDSVTCILQQITFGVLFTV
AISTVLAKTITVILAFKVTAPQRMMKYFLLSGVSMYIIPICTLIQVIVCAVCLGVSP
PSVDIDAQSEHGHIIVVCNNGSVTCFYSVLGYLACLTFVSFTLAFLSRNLPGAFNE
AKSITFSMLVVCSVWVTFLPVYHSTKGKIMVAVEIISILASSSGMLVCIFIPKCYIIL
FRPEKNSLKMIREKSPSHTHIS
>RN_Vom2r13
NLEMCVLCPEDKYANTGQNDCIPKALIFLSYEDHLGMALSLMALGFSAFTAVIL
GIFVKHHSTPIVKANNSTLTYILLLSLIFCFLCPLLFIGHPDSVTCILQQITFGVLFTM
AISTVLAKTITVILAFKVTAPQRMMKYFLLSGVSTYIIPICTLIQVIVCAVCLGVSPP
SVEIDAQSEHGHIIIVCNKGSVTCFYSILGYLACLTFVSFTLAFLSRNLPGVFNEAK
LITFSMLVVCSVWVTFLPVYHSTKGKIMVAGEIISILASSSGMLVCIFIPKCYIILFR
PEKNSLKMIREKSPSHTHIS
>RN_Vom2r16
NVDQCVKCPEDQYANPEQNHCIHKALIFLTYQETLGMTLSLVALCLSTLTAAILG
VFVKYHTTPIVKANNCRLTYILLITLIFCFLCPLLYIGHPNSVTCVLQQVTFGIVFT
VAISTVLAKTITVILAFKVTAPQRMIKYFLISGVFNYIIPICTLIQVFVCTLWLGTSP
PSVDFDPQSEHGHIIIVCNKGSVTAFYCMLGYLSCLALGSFTLAFLSRNLPGTFNE
AKFLTFSMLVFCSVWITFLPVYHSTKGKVMVAVEIFSTLASSAGMLGCIFIPKCYII
LLRPERNSLQRIREK
>RN_Vom2r75
NMDQCANCPEYQYANTEKNKCIQKGVIVLSYEDPLGMALALIAFCFSAFTVVVF
WVFVKHHDTPIVKANNRILSYLLIVSLMFCFLCSFFFIGYPNRATCILQQITFGIFFT
VAISTVLAKTITVVLAFKVTDPGRQLRIFLVSGTPNYIIPICSLLQCILCAIWLAVSP
PFVDIDEHSEHGHIIIVCNKGSITAFYCVLGYLACLAFGSFTIAFLAKNLPDTFNEA
KFLTFSMLVFCAVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAPKVYII
LMRPDRNSIHKIREKSYF
>RN_Vom2r30
NMDQCVNCPEYQYANKEKNKCIQKGVIFLSYEDPLGMVLALIAFCLSVFTAVVF
CVFVKNHDTPIVKANNRILSYLLVMSLMFCFLCSFFFIGHPNNATCILQQITFGITF
TVAVSTVLAKTITVVLAFKITDPGRMLRNFLVLGAPNYIIPICSLLQCILCAIWLAV
SPPFVDIDKHSEHGHIIIVCNKGSVMAFYCVLGYLACLALGSFTMAFLAKNLPDT

FNEAKYLTFSILVFFSVWVTFLPVYHSTKGRVMVAVEIFSILSSSAGMLGCIFTPKI
YIILVKPERNSIQRIREKSYF
>RN_Vom2r29
NMEECVNCPEYQYANTEKNKCIQKRVIFLSYEDPLGMALALIAFCLSAFTAVVL
WVFVKHHVTPIVKANNRILSYILIISLMFCFLCSFFFIGHPNRSTCVLQQITFGIVFT
VAVSTVLVKTITVVLAFKVTDPGRRLRNFLISGAPNYIIPICTLFQCILCAIWLAVS
PPFVDIDEHSEHGHIIIICNKGSIVAFYCVLGYLACLALGSFTLAFLAKNLPDTFNE
SKFLTFSMLVFCSVWVTFLPVYHSTKGRVMVAVEIFSILASSAGILGCIFAPKIYIIL
MKPDRNSIQKIMEKSNF
>RN_Vom2r32
NMDQCVNCPEYQYANTEKNKCIQKDVIFLSYEDPLGMALALIAFCLSAFTAVVL
WVFVKHHDTPIVKANNRILSYILIMSLMFCFLCSFFFIGHPNRGTCILQQITFGIVFT
VAVSTVLAKTITVILAFKLRDPGRSLRNFLVSGAPNYIIPICSLLQCILCAIWLAVSP
PFVDIDEHSEHGHIMIVCNKGSIMAFYCVLGYLACLALGSFTTAFLAKNLPDTFN
EAKFLTFSMLVFCSVWVTFLPVYHSTRGRVMVAVEIFSILASSAGMFGCIFAPKIY
IILMKPERNSIQKFREKSYF
>RN_Vom2r78
NMDQCVNCPEYQYANTEKNKCIQKDVIFLSYEDLLGMALALIAFCLSAFTAVVL
WVFVKHHDTPIVKANNRILSYILIMSLMFCFLCSFFFIGHPNRGTCILQQITFGIVFT
VAVSTVLAKTITVILAFKLRDPGRSLRNFLVSGAPNYIIPICSLLQCILCAIWLAVSP
PFVDIDEHSEHGHIMIVCNKGSIMAFYCVLGYLACLTLGSFTTAFLAKNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTRGRVMVAVEIFSILASSAGMLGSIFAPKIYII
LMKPERNCIQKFREKSYF
>RN_Vom2r31
NMDLCMNCPEYQYANTEKNRCIHKDMIFLSYEDPLGMALALIAFCLSSFTAVVL
WVFVKHHDTPIVKANNRILSYILIMSLMFCFLCSFFFIGHPNRGTCILQQITFGIVFT
VAVSTVLAKTITVLLAFQVTDTGRKLRNFLVSGTPNYIIPICSLLQCTLCAIWLAV
SPPFVDIDEHSEHGHIIIVCNKGSVMAFYCVLGYLAFLALGSFTMAFLAKNLPDTF
NEAKFLTFSMLVFCSVWITFLPVYHSTKGRVMVAVEIFSILTSSAGMLGCVFAPKI
YIILMKPERNSIQKIREKSRF
>RN_Vom2r28
NMDQCVNCPEYQYANTKRNKCIQKNVMFLSYEDPLGLVLAITVFVFSALTAVVL
WVFVKHHDTPIVKANNRILSYILIISLMFCFLCSFLFIGHPNITTCILQQITFGIVFTV
AVSTILAKTITVVLAFKVTNPGRRLRNFLVLGIPNYIIPICSLLQCILCAIWLAVSPP
FVDIDEHTERGHIIIMCNKGSVTAFYCVLGYLACLALASFTVAFLAKNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKIMVAVEIFSILASSAGMLGCIFAPKIYII
LMRPERNSIQKIREKSYF
>RN_Vom2r26
NMDQCMNCPEYQFANTKHNKCIQKGVLFLSYEDPLGMALALIAFCFSAFTAVVL
WVFVKHHDTPIVKANNRNLSYLLIISLMFCFLCTFFFIGHPNRATCILQQITFGIVF
TVAVSTVLAKTATVVLAFKVTDPGRRLRNVLVSGTPNYIIPICSLFQCILCAIWLA
VSPPFVDIDDYSEHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTVAFLAKNLPDT
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPDRNILK
>RN_Vom2r39

DMDQCVKCPNSQYANIKQSRCLKKVLTFLAYEDPLGMVLASLALCFSALTAVIL
GIFLKHQDNPIVKANNRLLSYILLISLIFCFLCSLLFIGHPYTATCILQQATFAVVFT
LAVSTILAKTITVVMAFKITDPRKKIRQMLVTRAPNYVIPICTMIQLILWATWLGT
SPPFVDTDPHFEHGHIIIVCNKGSVIAFF
>RN_Vom2r43
DMDYCYKCLDNQYANSEGTHCLSKTVIFLSYEDPIGVTLACLALCFSALTAVVF
GIFLRNQDTLIVKANNRALSYCLLISLIYCFLCSLLFIGKPHMVTCIMQQTIFAVAF
TVAVSTVLSKTITVVLAFKFTLPSKKMRWLMTSGTSVLIIPICTMIQLIICGIWLGT
YPPYVDSDGHIEHGHILIVCNKGSVTAFYCVLGYLGSIALASFTVAFLARNLPDTF
NEAKFLTFSMLVFCSVWLTFLPVYHSTSGKALVIVEVFCILTSTAGLLICIFAPKCY
IILLTPKNNSFHTFRHK
>RN_Vom2r42
NMNQCIKCLDDEYANPGGTNCFKKLIAFLGYEDPLGMSLTILALCFSALTAFVFG
IFLKHQESPTVKANNRTLSYVLLTSLISCFLCSLLFIGHPNFATCIMQQTTFAVVFT
VAASTVLAKTIIVILAFKVTNTSRKMRWLLVSGAPKFIIPICTMIQLILCGIWLGTSP
PFVDADGHVEKGHILIFCNKGSVLAFYCILGYLGSIAIGSFTLAFLARNLPNTFNE
AKYLTFSMLVFCSVWITFLPVYHSTKGKSMVAVEVFCILASSAGLLFCIFAPKCFII
LLRPEKKSSQKFQNIHSKM
>RN_Vom2r41
NMNQYIKCLDDQYANPGGTHCLKKVIVFLGYEDPLGMSLAILALCFSALTAFVL
SIFLKHQETPTVKANNRTLSYVLLISLISCFLCSLLFIGHPSFTTCIMQQTTFAVVFT
VAASTVLAKTIIVILAFKVTNTSRKMRWLLVSGAPKFIIPICTMIQLILCVIWLGTSP
PFVDADGHVEKGHILIFCNKGSILAFYCVLGYLVSIAIASFTLAYFARNLPDTFNE
AKFLTFSMLVFCSVWVTFLPVYHSTKGKSMVAVEVFCILASSAGLLFCIFAPKCFI
ILLRPEKKSFQKFQNIHSKI
>RN_Vom2r55
DMQQCMKCPETHYTNAERTQCLRKTITFLSYKDPLGMGLTIISLGFSALTGLIIGV
FVNYRDTPVVKANNRSLSYILLITLTHCFLCPLLFIGLPNTTTCILQENLLRLVFTV
ALSCVLAKSIIVVMAFQITAPGRKTRWLLISRAPNLIIPICTLIQILFTSIWLCTSPPF
VEVDAHSERGYIIILCNTGSTLAFHGTLAYLGVMAITSYFMAFLSRNLPDIFNEAK
FLAFSMLVFCCVWITFLPVYHSSTGKAVVVIEIFSILASSASILSLIFVPKCYVILFRS
EKNTIHAIFNKRLHRSKT
>RN_Vom2r54
DMNQCMKCPETHYANAEKRQCLKKTVTFLSYDDPLGMGLTIISMGFSALTALII
GVFVNYRDTAIVKANNRSLSYILLITLTHCFLCPLFFIGLPNTVTCILQENLLRLLFT
IALSCVLAKSITVVMAFQITAPRRKTRWLLFLRAPNLVIPVCLLVQVIFTSIWLGTS
PPFVEMDVQSEHGYIIILCNKGSGLAFHGTLAYLGVMAIASYIMAFLSRNLPDIFN
EAKFLAFSMLVFCSVWVTFLPVYHSTTGKTTVVMEIFSILASSASILSLIFAPKCYV
ILFRSEKNTLHQIRNKGHHRSKI
>RN_Vom2r53
DVGQCVKCPESHYANTEKSHCLKKTMTFLDYNDSLGTGLTLMSLGFFVVTGLVI
GVFIIHRNTPIVKANNRSLSYILLITLTLCFLCPLLFIGLPNTATCILQQNLFGLLFTV
ALSTVLAKTITVVMAFKITAPGRKTRWLLILRAPQFIIPLCALMQILFSGIWLGTSP
PFVDMDAHSEHGHIIILCNKGSAIGFYCTLAYLGVMAFGSYLLAFMSRNLPDTFN
ESKALAFSMLMFCSVWVTFLPVYHSTTGKVRVAMEMFSILASSASILTLIFVPKC
YIVLFRPERNILPLNREKRQHRSKNPET

>RN_Vom2r76
GMDQCFKCPETHYTSAEKKHCLKKTVTFLAYDDPLGKGLTFVSLGFSALTIFVIG
TFVNHRDTPIVKANNRTLSYILLVTLTLCFLCPLLFIGLPNTATCILQQNVFGILFT
VALSTVLAKTTTVVIAFKITAPGRKERWLLISQAPNFIIPVCTLIQVLLSGIWLGTSP
PFVDMDAHSEHGYIIILCNKGSTFAFYCTLAYLGVMAFGSYLMAFLSRNLPDTFN
EAKFLSFSMLVFCCVWVTFLPVYHSTTGKFMVAMEVFSILASSASILSLIFAPKCY
IILFRPERNTFHFIRDKRQNRCKN
>RN_Vom2r22
NMDQCVKCPETHYANTEKIHCLEKSVTFLAYDEPLGKAFSFMSLSFSSLTAVVLV
VFLKNRDTPIVKANNLALSYTLLIILFLCFLCPLLFIGRPSTASFILQQNIFGLLFTG
VLSTVLAKTITVVIAFKITASGKMRRWLLISRAPNYIIPICTLIQVALSGIWLTTSPP
FIDKDAHSEHGHIILICNKGSSIAFHCILGYLGTLALGSYFMAFLSRNLPDTFNEAK
FLAFSMLVFCSVWVTFLPVYHSTKGKIMVAMEVFSILASSISLLGIIFAPKCYIILL
RPEKNSLHYGRDKTYSRRQK
>RN_Vom2r25
DMDQCVKCPESHYGDTDKIHCLQKTVTFLAYGDPWGKTLCFLSLGFSSLTAAVL
GVFLKNRDTPIVKANNLPLSYTLLITLMLCFLCPLLFIGHPSTASCILQQNIFGLLFT
MALSTVLAKSITVVLAFKITAAGRIRRWLLISRAPNFIIPLCTLLQVFLSAIWLTTSP
PFVDKDAHSEHGHIIIICNKGSAIAFHCTLGYLGALALASYFMAFLSRNLPDTFNE
AKFLAFSMLVFCSVWVTFLPVYHSTKGKVTVAMEVFSILASSTSLLGIIFVPKCYL
ILLRPERNSLNYIREKSYARRKK
>RN_Vom2r23
DMDQCVKCPETHYANTNKIYCLPKTVTFLAYDDPLGKTLCFMSLGFSSLTAAVL
WVFLKNRDTPIVKANNLALSYTLLITLMLCLLCPLLFIGHPSTASCILQQNIFGLLF
TVSLSTVLAKTITVVIAFKITSPGRIRRWLLISRAPNFIIPLCTLLQVFLSAIWLTTSP
PFIDKDAQSEHGHIIIICNKGSAVAFHCNLGYLGALALVSYFMAFLSRNLPDTFNE
AKFLAFSMLVFCSVWITFLPIYHSTKGKTTVAMEVFSILASGTSLLGIIFAPKCYIIL
LRPERNSLHYIRDKKYAKRRK
>RN_Vom2r27
DVDQCVKCPETHYANIEKIHCLQKTVTFLYYDDPLGKTLCFMSLGFSSLTAAVL
VVFLKNRDTPIVKANNLALSYTLLITLMLCFLCPLLFIGRPSTASCILQQNIFGLLF
TVALSTVLAKTITVVIAFKITSPGRIRRWLLISRAPNFIIPLCTLLQVFLSGIWLTTSP
PFIDKDAHSEHGHIIIICNKGSAVAFHCNLGYLGALALVSYFMAFLSRNLPDTFNE
AKFLAFSMLVFCSVWVTFLPVYHSTKGKNMVAMEVFSILASSTSLLGIIFAPKCY
LILLRPERNSLSYIRDKTYAKSIK
>RN_Vom2r24
DMDQCVKCPQTHYANIEKTHCLQKTVTFLTYDDPLGKTLCVISLSFSSLTAAVLG
VFLKNRDTPIVKANNLALSYTLLITLMLCFLCPLLFIGRPSTASCILQQNIFGLLFT
MALSTVLAKTITVVIAFKITAPETVRRWLLISRAPNLIIPLCTLLQIFLSGIWLTTSPP
FIDKDVHSEHGHIIIICNKGSAVAFHCSLGYLGALALGSYFMAFLSRNLPDTFNEA
KFLAFSMLVFCSVWITFLPIYHSTKGKTTVAMEVFSILASGTSLLGIIFAPKCHIILL
RSEKNSLQYIKDKTHARRKT
>RN_Vom2r9
DVDQCVKCPETHFPNTEKNNCLQKSMSFLEYNDNLGMVLITIALCFSALTLAVL
GVFVKYRDTPIVKANNRTLSYILLMTLVVCFLCSLLFIGHPNTATCLMQQTTFGIV
FTVALSTVLAKVITVVMAFKFTIPGRITRWLMVSRAPNYVIPICTLIQIVLCGIYLG

TSPPFLDQDPYSEHDHILIVCNKGSDVVFHCILGYLCFLALGSCVMAFLSRNLPDT
FNESKLLSFSMQVFLCVWVTFLPVYHSTKGKVMVAMQVSSILASSAGILGFIFAP
KCYIILLRPDQNSLHHIRNTAYTRRNK
>RN_Vom2r11
DVDHCVKCPESYYANTEKNHCLQKAQSFLAYEDPLGMALTSISLLFSALTVIVV
GIFVRHRHTPMVKANNRALSYTLLISLNFCFLCSLNFIGQPNTATCILQQTIFGVAF
TVALATVLAKAITVVIAFKATFPGRIIRWLMISRAPNYTIPICTLIQLVLCGIWMAT
FPPFVDQDDHTEHGHVIIFCNKGSSVAFHCALGYLCFLALGSYAMAFFSRNLPDT
FNESKFLSFSMLVFFCVWITFLPVYHSTKGKIMLAMEVFCILASSAALLGFIFTPK
CYIILLRPEKNSFHRIRNKIHSKGNVSQTI
>RN_Vom2r10
DSDQCVRCPESHYPNTEKSNCYQKTSSFLAFEEPLGMALASIALCLSALTAFVIGI
FVKYRDTPIVKANNQALSYILLITLKFCFLCSLNFIGQPNTAACILQQTTFGVAFTV
ALATVLAKAITVVLAFKVSFPGRMVRWLMISRGPYYIVPICTLIQLLLCGIWMTM
SPPFIDQDPHTEHGHIIILCNKGSSLAFHAVLGYLCLLALGSYTMAFLSRNLPDTF
NESKFLSLSMLVFFCVWVTFLPVYYSTKGKVMVAMEVFSILASSTALLAFIFGPK
CYIILLRPDKNSFNHVRKKAHFRRKNS
>RN_Vom2r8
DVDQCLKCPESHYANTEKNHCLQKAVSFLAYQDPLGMSLASIALCLSTLTAFVIG
IFVKYRDTPIVKANNQALSYILLITLTFCFLCSLTFIGQPNKDTCIMQQITFGVVFT
VALATVLAKAITVVIAFKATFPGRMIRWLMKSRAPNYIIPICTLIQVFICGIWMATS
PPFIDQDFHAEHGHIIIFCNKGSSVAFHCTLGYLCFLALGGYTMAFLSRTLPDSFN
ESKFLSLSLLVFFCVWITFLPVYHSTMGKFMVATEIFSILASSIALLSFIFAPKCYIIL
FRPNENTFHYIRDKMHSRRNKS
>RN_Vom2r7
DVDQCMKCPESHYANPEKNHCFQKTVSFLAYEDPLGMILASISLCLSALTAFVIG
VFVKYRETPIVKANNRGLSYTLLITITFCLLCSLNFIGQPNAATCIFQQTIFGIAFTV
ALATVLAKAITVVIAFKATIPGRMVRWLMISRAPNYIIPICTLIQLLLCAIWMATSP
PFIDQDFHAEHGHILLLCNKGSSLAFHCVLGYLCSLAIGGYTMAFLSRNLPDTFNE
SKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSAALLGFIFAPKCY
IILLRPEKNSFHHMRDKTHSRKNKS
>RN_Vom2r38
DMDQCLKCPESHYANSEKNHCLRKSVSFLAYEEPLGMALTITALGFSVLTAVVL
VVFVKHRDTPIVKANNRSLSYILLLTLFICFLSSLLFIGQPSTATCILQQTAFGILFT
VALSTVLAKAITVVIAFKATVPARMVRWLMVSRAPNFIIPFCTLIQIIICGIWLVTS
PPFIDQDTHAERGHIIIVCNMGSIVAFHTVLGYLCSLALGSYAMAFLSRNLPDTFN
EAKFLSFSMQVFFCVWITFLPVYHSTKGKVMVSMEIFSTLASSIALLGLIFAPKCYI
ILLRPEKNSCIDIRHKRHSR
>RN_Vom2r37
DMDQCVKCPESHYANAEKNHCLRKSVSFLAYEEPLGMALTITALGFSVLTAVVL
VVFVKHRDTPIVKANNRSLSYILLLTLFICFLSSLLFIGQPRTATCILQQTAFGILFT
VALSTVLAKAITVVIAFKATVPARMVRWLMVSRAPNFIIPFCTLIQIIICGIWLVTS
PPFIDQDTHAELGHIIIVCNMGSTVAFHTVLGYLCSLALGSYTMAFLSRNLPDTFN
EAKFLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSILASSIALLGLIFAPKCY
IILLRTEKNSCPGIRHKKHSKSNI
>M_Vmn2r84

DMEQCVRCPDDQYANLEQTHCLQRAVSFLAYEDPLGMALGCMALSFTAITILVL
VTFVKYRDTPIVKANNCILSYILLISLVFCFLCSLLFIGHPNQVTCILQQTTFGLFFT
VAVSTVLAKTITVVMAFKLTTPGGRMRMMMMTGAPKLVIPICTLIQLVLCGIWL
LTSPPFIDRDIQSEHGKTVIFCNTGSVLAFHVVLGYLGSLALGSFTVAFLARNLPD
RFNEAKFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVP
KCYAILVRPDSNCIQKYKGK
>M_Vmn2r85
DMEQCVRCPDDKYANIEQTHCLQRAVSFLAYEDPLGMALGCMALSFTAITILVL
VTFVKYKDTPIVKANNCILSYILLISLVFCFLCSLLFIGHPNQVTCILQQTTFGLFFT
VAVSTVLAKTITVVMAFKLTTPGRRMRMMMMTGAPKLVIPICTLIQLVLCGIWL
LTSPPFIERDIQSEHGKTVIFCNTGSVLSFHVVLGYLGSLALGSFTVAFLARNLPDR
FNEAKFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVPK
CYVILVRPDSHFIQKYKGK
>M_Vmn2r86
DMEQCVRCPDDKYANLEQMHCLQRAVSFLAYEDPLGMALGCMALSFTAITILV
LVTFVKYKDTPIVKANNCILSYILLISLVFCFLCSLLFIGHPNQVTCILQQTTFGLFF
TVAVSTVLAKTITVVMAFKLTTPGKRMRGMMTGAPKLVIPICTLIQLVLCGIWLL
TSPPFIERDIQSEHGKTVIFCNKGSVLSFHVVLGYLGSLALGSFIVAFLARNLPDRF
NEAKFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLGCIFVPK
CYVILVRPDSNYIQKFKGK
>M_Vmn2r87
DMEQCVRCPDDKYANLEQSHCLQRTVSFLAYEDPLGMVLGSMALSSSAITILVL
VTFMKYKDTPIVKANNCILSYILLISLVFCFLCSLLFIEHPNQVTCILQQTTFGLFFT
VAVSTVLAKTITVVMAFKLTTPGKRMRGMMTGAPKLVIPICTLIQLVLCGIWLLT
SPPFIDRDIQSEHGKTVIFCNKGSVIAFHVVLGYLGSLALGSFTVAFLARNLPDRF
NEAKFLTFSMLVFCSVWITFLPVYHSTRGKIMVVVEVFSILASSAGLLGCIFVPKC
YVILVRPDSNFIQKYKGK
>M_Vmn2r80
DLDQCVHCPETHYANGEKRYCLKKRVTFLSYDDPLGKGITLLSIGFSTLTALFIW
VFVDHRGTPIIKANNRSLSYILLITLILCFLCPLLFIGLPNTAMCIMQQYIFGLLFTV
ALSTVLAKTITVVMAFKITAPGRKIRWLLTSQVPKFVIPVCTLIQILLSGIWLGTSP
PFIDKDCHSEHGHIIILCSKGSYIYFYCTLAYLGVMAFGSYLLAFLSRRLPDRFNES
KSLAFSMLVFCSVWVMFLPVYHSTSGKVMVAMEIFSTLASSASILIIVFAPKCYIV
LFRPEVNTLTHNRDKRQHRSKI
>M_Vmn2r81
GLDHCVHCPETHYANAEKRHCLKKTVTFLYYNDPLGKGITLLSIGFSTLTALFIW
VFVDHRGTPIIKANNRSLSYILLITLILCFLCPLLFIGLPNTATCIMQQYMFGLLFTV
ALSTVLAKTITVVMAFKITAPGRKIRWLLTSQVPKFAIPICTLIQVLLSGIWLGTSP
PFIDKDYHSEHGHIIILCNKGSYIYFYCTLAYLGVMAFGSYLLAFLSRRLPDRFNE
SKSLAFSMLVFCSVWVMFLPVYHSTSGKVMVAMEMFSILASSASILIIVFAPKCYI
VLFRPELNILTHNRDKRHHRSKT
>M_Vmn2r82
GMDQCVHCPETHYASAEKRHCLKKTITFLSYNDPLGKGITLLSIGFSTLTALVIW
VFVEHRGTPIIKANNRSLSYILLITLILCFLCPLLFIGLPNTATCIMQQYMFGLLFTV
ALSTVLAKTITVVMAFKITAPGRKIRWLLTSQVPKFVIPVCILIQVLLSGIWLGTSP
PFIDKDYQSEHGHIIILCSKGSYIYFYCTLAYLGVMAFGSYLLAFLSRRLPDRFNES

KSLAFSMLVFCSVWVMFLPVCHSTSGKVMVAMEMFSILASSASILIIVFAPKCYIV
LFRPELNTLTHNRDKRHHRSKT
>M_Vmn2r83
GAVQCVKCPETHYANAEKRHCLKKTVTFLDYNDPWGKGLTLMSLGFSALTTLII
GVFVNHRDTPIVKANNRSLTYILLITLALCFLCPLLFIGHPNTVTCIMQQNLFGLLF
TVVLSTVLAKTITVVMAFKIAVPGRKLRLLLISQVPNFIIPVCTLIQVCLSGIWLGT
FPPFIDMDAHSEYGHIIILCNKGSAIAFYCTLAYLGVMAIGSYLMAFLSRSLPDTF
NEAKFLAFSMLVFCAVWVTFLPVYHSTTGKVMVAMEMFSILASSSSLLILIFVPK
CYIILFRPERNTVHHIRDERHDRSKN
>M_Vmn2r88
DMEQCVKCPYDKYANIEKTHCLSRAVSFLAYEDPLGIALGCIALSFSAITILVLITF
LKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQVSCVLQQTTFGVFFTVS
VSTVLAKTITVVMAFKLTTPGRRMREMLVTGAPKLVIPICTLIQFVLCGIWLITSPP
FIDRDIQSEHGKIVILCNKGSVIAFHVVLGYLGSLALGSFTLAFLARNLPDTFNEA
KFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLMCIFVPKCYVI
LVRPDSNFIRKYKDK
>M_Vmn2r89
DIEQCVKCPDDKYANIEQTYCLSRAVSFLAFEEPLGMALGCMALSFSAITILVLVT
FVKYKNTPIVKANNRILSYILLISLVFCFLCSLLFIGHPDQVTCILQQTTFGVFFTVS
VSTVLAKTITVVMAFKFTTPGRRMRGMLVTGAPKLVIPICTLIQLVLCGIWLVTSP
PFIDRDIQSEHGKVVILCNKGSVIAFHIVLGYLGSLALGSFTLAFLARNLPDTFNEA
KFLTFSMLVFCSVWITFLPVYHSTRGKVMVVVEVFSILASSAGLLMCIFVPKCYVI
LVRPDSNFTKNRKGK
>M_Vmn2r90
DMDECVKCPESHYANSEKNHCLQKSVSFLAYEDSLGIALTTTALCLSVLTAVVL
VVFVKHRHTPIVKANNRTLSYTLLLTLIMCFLSSLLFIGQPNTTTCILQQTAFGIFF
TVALSTVLAKAITVVIAFKVTVPNRKVRRLMISRAPNFIIPICTLVQLLFCGIWLITS
PPFIDQNAHAEHGNIIILCNKGSAVAFHCVLGYLCFLALVSYTLAFFSRNLPDTFN
EAKFLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSMLASSTALLSLIFGPKC
YIILLRPDKNVCLDIRHKSHSKRNS
>M_Vmn2r91
DVDQCVKCPEGHYANTEKNNCFQKSVSFLAYEDPLGMALASIALCLSALTVFVI
GIFVKHRDTPIVKANNRVLSYILLITLTFCFLCSLNFIGQPNTAACILQQTTFAVAF
TVALATVLAKTITVVLAFKVSFPGRIVRWLMISRGPNYIIPICTLIQLLLCGIWMAT
SPPFNDQDAHSEYGHIIILCNKGSAVAFHSVLGYLCFLALVSYTMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMIAMEVFSILASSTALLAFIFGPKC
YIILLRPEKNLFRHTRKKNH
>M_Vmn2r92
DVDQCVKCPEGHYANSEKNNCFQKSVSFLAYEDPLGMALASIALCLSALTVFVI
GIFVKHRDTPIVKANNRVLSYILLITLTFCFLCSLNFIGQPNTAACILQQTTFAVAF
TVALATVLAKTITVVLAFKVSFPGRIVRWLMISRGPNYIIPICTLIQLLLCGIWMAT
SPPFIDQDAHSEYGHIIILCNKGSAVAFHSVLGYLCFLALVSYTMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMMAMEVFSILASSTALLAFIFGPK
CYIILLRPEKNSFSHTRKKKTHSKRKI
>M_Vmn2r93

DVDQCVKCPESHYANTERNNCLQKSVSFLAYDEPLGMALASIALCLSTITVFVIGI
FLRHRDTPIVKANNQGLSFILLITLTFCFLCSLNFIGQPNTSACILQQTTFAVAFTM
ALATVLAKAITVVLAFKVSFPGRMVRWLMISRGPNYIIPICTLIQLLLCGIWMATY
PPFIDQDAHTEHGHIILLCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTFN
ESKFLSFSMLVFFCVWVTFLPVYHSTNGKVLVAMEVFSILASSSALLAFIFGPKCY
IILLRPDKNSFNHIRRKQHTRRKNS
>M_Vmn2r94
DVDQCVKCPESHYANTEKDHCFPKSVSFLAYEDPLGMALGSIALCLSAFTAFVIG
IFVKHRDTPIVKANNRALSYILLITLTFCFLCSLNFIGKPNTVACILQQTTFAIAFTV
ALATVLAKAITVVLAFKVSFPGRMVRWLVISWGPRYIIPICTLIQLLICGIWMTTSP
PFIDQDAHTEHGHIILLCNKGSAVAFHSVLGYLCILALGSYTMAFFSRNLPDTFNE
SKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLAFIFGPKCYI
ILLRPEKNSFNHIKKKTRSKQKI
>M_Vmn2r95
DVDQCVKCPESHYANTEKSNCFQKSVSFLAYEDPLGMVLVSIALCLSTLTVFVIG
IFVKHRETPIVKANNRALSYILLVTLTFCFLCSLNFIGKPTTAACILQQTTFAVAFT
VALATVLAKAITVVLAFKVSFPGRMVRWLMISRGPNYIIPICTLIQLLLCGIWMAT
SPPFIDQDAHTEYGHIIILCNKGSAVAFHSVLGYLCFMALGSYAMAFLSRNLPDTF
NESKFLSFSMLVFFCVWITFLPVYHSTKGKVMVAMEIFSILASSTALLAFIFGPKC
YIILLRPDKNSFNHIRK
>M_Vmn2r96
DVDTCVKCPERHYANAEKSNCFQKSVSFLAYEDTLGMALASIALCLSVLTAFVI
GIFVKHRDTPIVKANNRALSYILLITLTFCFLCSLNFIGKPTTVSCILQQTTFAVAFT
MALATVLAKAITVVLAFKVSFPGRMIRWLMISRGPNYIIPICTLIQLLICGIWMATS
PPFIDQDAHTEYGHIIILCNKGSSVAFHSVLGYLCFMALGSYAMAFLSRNLPDTFN
ESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLAFIFGPKC
YIILLRPDKNSFNYIRKKTPSRKN
>M_Vmn2r97
DVDKCVKCQERHYANREKSNCFQKSVSFLAYEDPLGMVLASIAICLSALTVFVIG
IFMKHRDTPIVKANNRALSYILLITLTFCFLCSLNFIGKPTTASCILQQTTFAVAFT
MALATVLAKAITVVLAFKVSFPGRMVRWLMISRGPNYIIPICTLIQLLICGIWMAT
SPPFIDQDAHTEYGHIIILCNKGSAVAFHSVLGYLCLMALGSYAMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLIFIFGPKC
YIILLRPDKNSFNHIRKKTPTRKKN
>M_Vmn2r98
DVDKCVKCPERHYANKEKSNCFQKSVSFLAYEDPLGMAVASIALCLSALTAFVI
GIFVKHRDTPIVKANNRALSYILLITLTFCFLCSLNFIGKPTTASCILQQTTFAVAFT
MALGTVLAKAITVVLAFKISFPGRMVRWLMISRGPNYIIPISTLIQLIICGIWMATS
PPFIDQDAHTEYGHILILCNKGSAVAFHSVLGYLCFMALGSYAMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLVFIFGPK
CYIILLRPDKNSFNHIRKKTPSRKKK
>M_Vmn2r99
DVDKCVKCPERHYANAEKNNCFQKSVSFLAYEDPLGMALASIALCLSALTAFVI
GIFVKHRDTPIVKANNRALSYILLITLTFCFLCSLNFIGQPNTAACILQQTTFAVAF
TIALATVLAKAITVVLAFKVSFPGRMLRWLMISRGPNYIIPVCTLIQLLICGIWMA
TSPPFIDQDPQTEYGHIIILCNKGSAVAFHSVLGYLCFMALGSYAMAFLSRNLPDT

FNESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLVFIFGP
KCYIILLRPDKNSFNHIRKKTLSRRKN
>M_Vmn2r100
DVDQCVKCPESHYANTEKSMCYQKSVSFLAYEDPLGMALASIALCLSALTAFVI
GIFVKYRDTPIVKANNRALSYILLITLTFCFLCPLNFIGQPNIAACILQQTTFAVAFT
MALATVLAKAITVVLAFKVSFPGRMVRWLMISKGPNYIIPICTLIQLLICGIWMAT
SPPFIDQDAHTEHGHIIILCNKGSAVAFHSVLGYLCLLALGSYTMAFLSRNLPDTF
NESKFVSFSMLVFFCVWVTFLPVYHSTKGKIMVAMEVFSILVSSTALLAFIFGPKC
YIILLRPDKNSFNHIKKKTSH
>M_Vmn2r101
DVYQCVKCPEGHYANKERKSCFQKSMTFLAYEDPLGMALASIALCLSALTFFVI
GIFVKHRDTPIVKANNRTLSYILLITLTFCFLCSLNFIGQPNRATCILQQTTFAVTFT
VALATVLAKAITVVLAFKVSFPGKIVRWLMLSRGPNYIIPICTLIQLLLCGIWMTT
SPPFIDLDAHTEYGHIIILCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTF
NESKFLSFSMLVFFCVWLTFLPVFHSTKGKIMVAMEVFSILASTTALLAFIFGPKC
YIILLRPEENSFTHIQKKTHSRRKN
>M_Vmn2r102
DVYQCVKCPEGHFANTERNNCFQKSVIFLAYEDPLGMALASIALCLSSLTVFVIGI
FVKHRDTPIVKANNRALSYILLLTLTFCFLCSLNFIGQPNTATCILQQTTFAVTFTM
ALATVLAKAITVVLAFKVSFPGKIVRWLMLSRGPNYIIPICTLIQLLICGIWMATSP
PFIDEDAHTEYGQIIILCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTFNE
SKYLSFSMLVFFCVWVTFLPVYHSTKGKIMVAMEVFSILASSTALLAFIFGPKCYI
ILLRPENNSFTHIRKKTYSRRKN
>M_Vmn2r103
DVYQCVKCPEGHYANTERNNCFQKSVIFLAYEDPLGMTLASVALCLSALTVFVI
GIFVKHRDTPIVKANNRALSYILLLTLTFCFLCSLNFIGQPNTATCILQQTTFAVTF
TMALATVLAKAITVVLAFKVSFPGKIVRWLMLSRGPNYIIPICTLIQLLICGIWMT
TSPPFIDQDAHTEHGHIIILCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDT
FNESKYLSFSMLVFFCVWVTFLPVFHSTKGKIMVSMEVFSILASSIALLAFIFGPKC
YIILLKPENNSFTHVRKKTYSRKNYPKI
>M_Vmn2r104
DVDQCVKCPESHYANTEKSNCFPKSVSFLAYEDPLGMALASIALCLSALTVFVIG
IFVKNRDTPIVKANNRTLSYILLITLTFCFLCSLNFIGQPNTAACILQQTTFAVAFT
MALATVLAKAITVVLAFKISFPGRMLRWLMISRGPRYIIPICTLIQLLLCGIWMAT
SPPFIDQDVNTEDGYIILLCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSTALLAFIFGPK
CYIILLRPEKNSFNHIRKKTQSRRKKS
>M_Vmn2r105
DEDQCMKCPDSHYANTEKNYCLKKAVSFLAYEDLLGMILASIALCLSALTVFVI
GIFVKNRHTPIVKANNQALSYILLITLTFCFLCSLTFIGQPNKNTCILQQTTFGVAFT
AALATVLAKAITVVIAFKATFPGRIVRWLMISRAPNYIIPICTLIQVGFCGTWMAT
SPPFIDQDAHTEHGHIIIFCNKGSSVAFHCVLGYLCFLALCGYTVAFLSRTLPDTFN
ESKFLSFSLLVFFCVWVTFLPVYHSTKGKFMVATEIFSILASSTALLGFIFGPKCYI
VLLRPDENTFNYIRNKIHSRKNKSQKT
>M_Vmn2r106

DVDQCVKCPESHYSNTEKNHCFQKTVIFLAYEDPLGMTLACIALCLSALTAFVIG
IFVKYRDTPIVKANNRALSYTLLITLTFCFLCSLNFIGQPNLTTCILQQTTFGVAFT
MALATVLAKAITVVIAFRATFPGRMVRWLMMSRVPNYIIPICTLIQLLLCGIWMA
ISPPFIDNDVHTEHGYIIIFCNKGSSFAFHCVLGYLCSLAFGGYTMAFLSRNLPGTF
NESKFLSFSMLVFFCVWVTFLSVYHSTKGKFMVAMEVFAILASSSALLGFIFGPK
CYIILIRPDRNSFHHIRDKTHSRRNKT
>M_Vmn2r107
DVDQCVKCPESHYANTEKSNCYQKSVSFLGYEDPLGMALASIALCLSALTAFVI
GIFVKHKDTPIVKANNQALSYTLLITLKFCFLCSLNFIGQPNTVACILQQTTFAVAF
TMALATVLAKAITVVLAFKVSFPGRMVRWLMISRGPNYIIPICTLIQLLLCGIWM
AISPPYIDQDAHIEHGHIIILCNKGSAVAFHSVLGYLCFLALGSYTMAFLSRNLPDT
FNESKFISLSMLVFFCVWITFLPVYHSTKGKVMVAMEVFSILASSTALLAFIFGPK
CYIILLRPEKNSFHHIREKRTTHSRRK
>M_Vmn2r108
DVDQCVKCPESHYSNTEKNHCLQKTVIFLAYEDPLGMTLACIALCLSALTAFVIG
VFVKYRDTPIVKANNRGLSYTLLITITFCLLCSLNFIGQPNTASCILQQNIFGVAFT
VALATVLAKAITVVLAFKATFPGRMVRWLMTSRAPNYIIPLCTLIQLFICGIWMT
MSPPFIHQDNHAEHGHILLLCNKGSSVAFHCVLGYLCSLAFGGYTMAFLSRNLPD
TFNESKFLSFSMVVFFCVWVTFLPVYHSTKGKVMVAVEVFSILASSTALLGFIFAP
KCYIILFRPEKNSFHRIKDKTYSR
>M_Vmn2r109
DADQCIKCPESHYASIEKNRCFQKALSFLSYKDPLGVTLANISLCLSALTAFVIGV
FVKYRDTPIVKANNRGLSYTLLITITFCLLCSLNFIGQPNTASCILQQTIFGVSFTVA
LATVLAKAITVVIAFKATFPGRMVRWLMTSRAPNYIIPICTLIQLLLCGIWMATSP
PFIDQDVHAEHGQILLFCNKGSSVAFHCVLGYLCSLALGGYTMAFLSRNLPDTFN
ESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSAALLGFIFAPKC
YIILLRPEKNSFHHIRDKTHSRRNKS
>M_Vmn2r110
DVDQCVKCPESHYANTEKNHCFQKTVSFLAYEEPLGITLASITLCLSALTAFVIGV
FVKYRDTPIVKANNRGLSYTLLITITFCLLCSLNFIGQPNTASCILQQNIFGVAFTV
ALATVLAKAITVVIAFKATFPGRMVRWLMTSRAPNYIIPICTLIQLLLCGIWMATS
PPFIDQDVHAEHGQILLFCNKGSSVAFHCVLGYLCSLALGGYTMAFLSRNLPDTF
NESKFLSFSMLVFFCVWVTFLPVYHSTKGKVMVAMEVFSILASSAALLGFIFAPK
CYIILFRPDRNSFHHIRDKTHSQRNTS
>M_Vmn2r112
NVDNCVQCPEDQYANTEQNHCIHKAVVFLSYEEPLGMALSLLSLFFSAFTTVILG
VIVKHHNTPIVKANNRTLTYILLISLIFCFLCPLLFIGHPNSVTCFLQQITFGVVFTV
ALSTVLAKTITVVLAFKITASQRMMKYFLVSGASNYIILICTLIQVIVCAVWLGAS
PPFVDIDAQSEHGHIIIVCHKGSVNAFYSILGYLALLSFGSFTLAFLSRNLPGAFNE
AKSITFSMLVFCSVWVTFIPVYHSTKGKVMVAVEIFSTLASSAGILGCIFVSKCYTI
LFRPDRNSVEMTRVKSSFHAHIS
>M_Vmn2r111
NVDNCAQCPEDQYANTEQNHCIHKAVVFLNYEEPLGMALSLFSLCFSAFTTVILG
VIVKHHNTPIVKANNRTLTYILLISLIFCFLCPLLFIGHPNSVTCILQQITFGVVFTV
ALSTVLAKTIVVVLAFKITASQRMMKYFLVSGASNYIILICTLIQVIVCAVWLGAS
PPSVDIDAQSEHGHIIIVCHKGSVNAFYSILGYLALLSFGSFTLAFLSRNLPGAFNE

AKSITFSMLVFCSVWVTFIPVYHSTKGKVMVAVEIFSTLASSAGMLGCIFVSKCY
TILFRPDRNSVEMTRVKSSFHARVS
>M_Vmn2r113
NVELCVPCPEDQYANTEQNHCIHKALIFLTYKDPLGMALSLMALCFSAFTAVVL
LVFVKHHSTPIVKANNRTLTYILLISLIFCFLCPLLFIGHPNSVTCILQQVTFGVAFT
VAVSTVLAKTITILLAFRVTAPQRMMKYFLVSGTYIYIIPICTLIQVIVCAVWLGA
APPSVDIDARSEHGHIIIVCNKGSVSAFYCVLGYLACLAFGSFTLAFLSRNLPSTFN
EAKSITFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSTLASGAGMLICIFVPKC
YIILFMPDRNSLQMIRKKSSSHAH
>M_Vmn2r115
NVDYCVQCLEDQYANTEQNHCIRKIVVFLSYEEPLGVALSLISLFFSAFTIVVLGV
IVKHHSTPVVKANNRTLTYILLISLIFCFLCPLLFIGHPNAATCILQQITFGIVFTVSV
STVLAKTITVVLAFKIIASQRMMKYFLVSGAINYIIPVCILIQVIVCAVWLGASPPS
VDIDGQSEHGQIIIVCHKGSVNAFYCVLGYLATLAIGSFTLAFFSRNLPGAFNEAK
SITFSMLVFCSVWVTFIPVYHSTKGKVMVAVEIFSTLASSAGMLGCIFVPKCYTIL
FRPDQNSLEMIRVKSSSNARSS
>M_Vmn2r114
NVDYCVQCLEDQYANTEQNHCIRKTVVFLSYEEPLGVALSLISLFFSAVTIVVLG
VIVKHHSTPVVKANNRTLTYILLISLIFCFLCPLLFIGHPNAATCILQQITFGVVFTV
SVSTVLAKTITVVLAFKITASQRMMKYFLVSGAINYIIPFCILIQVIVCAVWLGASP
PSVDIDAHSEHGQIIIVCHKGSVNAFYCVLGYLATLAIGSFTLAFFSRNLPGAFNE
AKSITFSMLVFCSVWVTFIPVYHSTKGRVMVAVEIFSILASSAGMLGCIFVPKCYT
ILFRPDQNSLEMIRAKSSSRAHIS
>M_Vmn2r116
NVDNCVQCPEDQYANTEQNHCIRKAVVFLSYEEPLGVALSLLSLCFSAFTTVVL
GIFVKHHNTPIVKANNRTLTYLLLISLIFCFLCPLLFIGHPNSATCILQQLTFGVVFT
VSLSTVLAKTITVVLAFKIIASQRMMKYFLISGAINYIIPICILIQVIVCAVWLRASPP
SVDIDAHSEHGQIIIVCHKGSVNAFYCVLGYLAILAFGSFTLAFLSRNLPGAFNEA
KSITFSMLVFCSVWVTFIPVYHSTKGKVMVAVEIFSTLASSAGMLGCIFVPKCYTI
LFRQDQNSLEMIRVKSSSNVHVS
>M_Vmn2r117
NVDYCVQCLEDQYANTEQNQCIRKTVVFLSYEEPLGMALSLLSLCFSAFTTVVL
GVMVKHHSTPIVKANNRTLTYILLISLIFCFLCPLLFIGHPNAATCILQQITFGVVFT
VSVSTVLAKTITVVLAFKITASQRMMKYFLVSGAINYIIPICILIQVIVCAVWLRAS
PPSVDIDAHSEHGQIIIVCHKGSVNAFYCVLGYLATLAIGSFTLAFFSRNLPGAFNE
AKSITFSMLVFCSVWVTFIPVYYSTKGKVMVAVEIFSILASSAGMLGCIFVPKCYT
ILFRPDQNSLEMIRVKSSSNAHVS
>M_Vmn2r118
DLTQCVKCPDGQYANIERNHCFPKVVTYLAYADPLGIVLACMALFLSMLSVLVL
GIFVKNQYTPIVKANNRTLSYILLISLTLCFLCSLFFIGHPSTARCLLQQITFGILFTV
AVSAVLAKTITVLLAFSITIPGRRVKWLLLFGTHKVIIPICSLFQLIICGFWLGLSPP
FVERDTHTEHGHIIIMCNKGSLIAFYSVLGYLGCLALASFIMAFLVRNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVHHSTKGKAMVAVEVFSILASSAGLLGCIFIPKCYII
VLRPDRNFIQRSKGKNHS
>M_Vmn2r119

DLTQCVKCPDGQYANIERNHCFPKVVTYLAYADPLGIVLACMALFLSMLSVLVL
GIFVKNQYTPIVKANNRTLSYILLISLTLCFLCSLFFIGHPSTARCLLQQITFGILFTV
AVSAVLAKTITVLLAFSITIPGRRVKWLLLFGTHKVIIPICSLFQLIICGFWLGLSPP
FVERDTHTEHGHIIIMCNKGSLIAFYSVLGYLGCLALASFIMAFLVRNLPDTFNEA
KFLTFSMLVFCSVWVTFLPVHHSTKGKAMVAVEVFSILASSAGLLGCIFIPKCYII
VLRPDRNFIQRSKGKNHS
>M_Vmn2r120
DSIQCIKCPDDQYANTEQTHCIPKAVTFLAYEDPLGMALACMALCFSALTALVL
GVFVKHHETPIVKANNRTLSYILLISLILCFLCSLPFIGLPNMATCILQQMSFGVVF
TIAVSTVLAKTVTVVLAFRSTVPGRRMRWILVSPNLIIPICTLIQMVLCAIWLGAS
APFVDINAHNERGHITIMCNKGSATAFYCVLGYLGFLAMGTLTVAFWARNLPDI
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEVFSILASSAGLLGCIFVP
KCYIILIRPKNGSFK
>M_Vmn2r1
GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYEEALGFTLVILSIFGALVLLAVTV
VYVIHRHTPLVKANDRELSFLIQMSLVITVLSSLLFIGKPCNWSCMARQITLALGF
CLCLSSILGKTISLFFAYRISVSKTRLISMHPIFRKLIVLICVVGEIGICAAYLVLEPP
RMFKNIEIQNVKIIFECNEGSIEFLCSIFGFDVLLALLCFLTTFVARQLPDNYYEGK
CITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCLFLPKCFIILLR
PKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r2
GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYGEALGFTLVILSIFGALVVLAVT
VVYVIHRHTPLVKANDRELSFLIQMSLVITVLSSLLFIGKPCNWSCMARQITLALG
FCLCLSSILGKTISLFFAYRISVSKTRLISMHPIFRKLIVLVCVVGEIGVCAAYLVLE
PPRMFKNIEIQNVKIIFECNEGSVEFLCSIFGFDVLLALLCFLTTFVARQLPDNYYE
GKCITFGMLVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCLFLPKCFIIL
LRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r3
GQRLCNPCSEDDWSNAEKNKCVPKHVEFLAYEEALGFTLVILSIFGALVVLAVT
VVYVIHRHTPLVKANDRELSFLIQISLVITVLSTMLFIGKPCNWSCMARQVTLALG
FCLCLSSILGKTISLFFAYRISISKTQLISMHPSFRKLIVLLCVFGEIGVCASYLVLKP
PSMFKNIERQNVKIIFECNEGSVEFLCSIFGFDVLLALLCFVTTFMARQLPDNYYE
GKCITFGMLFFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCLFLPKCFIIL
LRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTMVSTVLDE
>M_Vmn2r4
GQRLCDPCGENDWSNAEKNKCVPKHVEFLAYEEALGFTLVIISIFGALVVLAVTV
VYVIHRHTPLVKANDRELSFLIQMSLGITVLSSMLFIGKPCNWSCKTRQVTLALG
FCLCLSSILGKTVSLFFAYRISISKTRLISMHPIFRKLIVLICVVGEIGVCTAYLVLEP
PSLFKNIEPQNVKIIFECNEGSIEFLCSIFGFDVLLALLCFATTFVARQLPDNYYEG
KCITFGMLVFFIVWISFVPAYLSTKGKFNVAVEIFAILASSYGLLGCLFLPKCFIILL
RPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r5
GQRECDPCGEDDWSNAEKSKCVPKLVEFLAYEEALGFTLVILSIFGALVVLAVTA
VYVVHRHTPLVKANDRELSFLIQMSLGITVLSSMLFIGKPCNWSCMARQITLALG
FCLCLSSILGKTISLFFAYRISISKTRLISMHPIFRKLIVLICVLAEICVCAVYLVLEPP
KMFKNIEIQNVKIIFECSEGSIEFLCSIFGFDVLLALLCFLTTFVARQLPDNYYEGK

CITFGILVFFIVWISFVPAYLSTKGKFKVAVEIFAILASSYGLLGCVFLPKCFIILLRP
KRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r6
GQRLCDPCGENDWSNAEKNKCVPKHVEFLAYEEALGFTLVILSIFGALVVLAVT
VVYVIHRHTPLVKANDRELSFLIQVSLGITVLSSMLFIGKPCNWSCKTRQVTLAL
GFCLCLSSILGKTVSLFFAYRISISKTRLISMHPIFRKLIVLICVVGEIGVCTTYLVLE
PPSLFKNIEPQNVKIIFECNEGSIEFLCSIFGFDVLLALLCFVTTFVARQLPDNYYEG
KCITFGMLVFFIVWISFVPAYLGTKGKFNVAVEIFAILASSYGLLGCVFLPKCFIIL
LRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r7
GQRLCDPCGENDWSNAEKNKCVPKHVEFLAYEEALGFTLVILSIFGALVVLAVT
VVYVIHRHTPLVKANDRELSFLIQVSLGITVLSSMLFIGKPCNWSCKTRQVTLAL
GFCLCLSSILGKTVSLFFAYRMSISKTRLISMHPIFRKLIVLICVVGEIGVCTAYLVL
EPPSLFKNIEPQNVKIIFECNEGSIEFLCSIFGFDVLLALLCFVTTFVARQLPDNYYE
GKCITFGMLVFFIVWISFVPAYLGTKGKFNVAVEIFAILASSYGLLGCVFLPKCFII
LLRPKRNTDETVGGRVPTVDRSIQLTSASVSSELNTTVSTVLDE
>M_Vmn2r8
DMEQCVRCPDDKYANLEKTHCLQRAVSFLAYEDPWGMALGCMALFLSALTVL
VLVTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVF
FTVAISTVLAKTITVLMAFKLTTPGRRMRGMLASGAPNLVIPICTLIQLFLCGIWL
VTSPPFIDRETQSEYGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDR
FNEAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFLPK
CCVILVRLDSNFLQKYKDK
>M_Vmn2r9
DMEQCVSCTGDKYANLEKTHCLQRAVSFLAYEDPLGMSLGCIALFLSALTSLVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVFFT
VAISTVLAKTITVLMAFKLTTPGRRMRGMLASGAPNLVIPICTLIQLVFCGIWLVT
SPPFIDRDTQSEYGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRFN
EAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFLPKCC
AILVRLDSNFLQKYKDK
>M_Vmn2r10
DMEQCMRCPDDKYANLEKTHCLQRAVSFLAYEDPLGMALGCTALSFSALTVLV
LVTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVFF
TVAISTVLAKTITVLMAFKLTTPGRRMRGMLASGAPNIVIPICTLIQLVLCGIWLV
TSPPFIDRDMQSEHGKTIIICNKGSVIAFHFVLGYLGALALGSFTVAFLARNLPDRF
NEAKFLTFSMLVFCSVWITFLPVYHSTQGTVMVVVEVFSILASSAGLLGCIFLPKC
CVLLRIQNSNFLHKYKHE
>M_Vmn2r11
DMEQCVRCPDDKYANLEKTHCLQRAVSFLAYEDPLGIALGCMALFLSALTVLVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVFFT
VAISTVLAKTITVLMAFKLTTPGRRMRGMLASGAPNLVIPICTLIQLVFCGIWLVT
SPPFIDRDTQSEYGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRFN
EAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFLPKCC
VILVRLDSNFLQKYEHK
>M_Vmn2r12

DMEQCVRCPDDKYANLEKTHCLQRAVSFLAYEDPLGMALGCMALFLSALMVL
VLVTFVKYKDTPIVKANNRILSYILLISLVFCFLCPLLFIGHPNLATCILQQTTFGVF
FTVAISTVLAKTITVLMAFKLTTPGRRMRGMLASGAPNLVIPICTLIQLVLCGIWL
VTSPPFIDRDTQSEYGKTIIICNKGSVIAFHFVLGYLGALALGSFTVAFLARNLPDR
FNEAKFLTFSMLVYCSVWITFLPVYHSTRGAVMVVVEVFSILASSAGLLGCIFLP
KCCVILL
>M_Vmn2r13
DMEQCVRCPDDKYANVEKTHCLQRVVSFLAYEDPLGMALGCMALSFSALTVLI
LVTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQITFGVFF
TVAISTVLAKTITVVLAFKLTTPGRRMRGMLASGAPNLAIPICTLIQLVLSGIWLV
KSPPFIDRDTQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRF
NEAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFLPKC
CVIL
>M_Vmn2r14
DMEQCVRCPDDKYANLEKTHCLQRAVSFLAYEDPLGMALGCMALTFSALTVLV
LVTFVKYKDTPIVKANNRILSNILLISLVFCFLCPLLFIGHPNKATCILQQTTFGVFF
TVAISTVLAKTITVLMAFKLTSPGGKMRGMLASGAPNLVIPICTLVQLVLCGIWL
VTSPPFIDRDTQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDR
FNEAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFVPK
CFVILVRPDSNFLQKYKDK
>M_Vmn2r16
DMEQCVRCPDDKYANLEKTHCLQRAVSFLAYEDPLGMALGSMTMCFSALTVL
VLVTFVKYKDTPIVKANNRILSYILLISLILCFLCSLFFIGQPNHATCILQQTTFGVF
FTVAISTVLAKTITVVMAFKLTTPGRRMRGMLASGTPNLVIPICSLVQLVLCGIWL
VTSPPFIDRDTQSEHGKTIIICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDR
FNEAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFVPK
CFVILVRPDSNFLQKYKDK
>M_Vmn2r15
DMEQCVRCPDDKYANLEKTQCLQRVVSFLAYEDPLGMTLGCMALTFSALTVLV
LVIFVKYKDTPIVKANNRIVSYILLISLVFCFLCSLLFIGHPTQATCILQQTTFGVFF
TVAISTVLAKTITVLMAFKLTTPGGRMRGMLASGAPNLVIPICTLIQLVLCGIWLV
TSPPFIDRDIQSEHGKTIIICNKGSVIAFHFVLGYLGALALGSFTVAFLARNLPDRF
NEAKFLTFSMLVFCSVWITFLPVYHSTRGTVMVVVEVFSILASSAGLLGCIFLPKC
CVILLRPDLTFLQKYKDK
>M_Vmn2r17
DMEQCVRCPDDKYANLEQTHCLQRAVSFLAYEDPLGMALGCMALCFSALTLLV
FFTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQATCILQQTTFGVFF
TVAISTVLAKTITVVMAFKLTAPGRRVRGMLVAGAPNLAIPICTLIQIVLCGIWLV
TSPPFIGRDTKSEHGKTILICNKGSVIAFHFVLGYLGSLALGSFTVAFLARNLPDRF
NEAKFLTFSMLMFCSVWITFLPVYHSTRGKAMVVVEVFSILASSAALLGFIFVPK
CYVILVKSDSNFREKYKNK
>M_Vmn2r18
DQQKCIACLIQEYPNPERNHCLPKSATFLSFEDPLGMSLACMALGFSVSTAAVLG
IFLKYQDSPIIKANNLTLSYILLISLLLCFLCSFLFIGHPNTVSCILQQITFALVFTLAL
STVLAKTITVILAFRALKPGRTMRRLFVLGIYNAVIPFCVLIQLITSGIWLATSPPYI
DTDSYSEHAHIIILCNNGSVTAFYCMLVYLGTLALGSFTVAFLARNLPDTFNEAK

FLTFSMVVFCSVWVTFVPVYQSTKAKAMVAVEVFSILASSAGLLACIFFPKCYIIL
REPDKRNLKCFKNKTRIKK
>M_Vmn2r19
DMDQCIKCPEDQYPNKQKNYCLPKITTYLSHGDTLGAVLGSVAISLSIFSAMILGL
FIHYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQTRTVTCVLRQMIFGVVFSV
TISAILAKTFIVVVAFNAIKPGSRLQMWMVIRLSYAIVCCGSIIQVCICAVWLGTYP
PFPDVDMHSEFGQIILWCNEGSTFAFYFVLGYLSFLSSLSLLIAFLARRLPDSFNEA
KTITFSMLVFCSVWISFVPTYLSSKGKTMVAVEIVSILVSSSILLLCIFLPKCYVILL
RSGGHSRKKFFK
>M_Vmn2r20
DMDQCIKCPEDQYPNKQKNNCLPKIKIFLSHEDTLGAVLVSLAISLSAFSAMILGL
FIHYRDTPIVRANNRNLSYVLLVSLMLCFFCSLIFIGQPSTVTCVLRQMIFGVVFSV
AVSAILAKTFIVVVAFTAIKPGSTLQMWMVTRLSNAIVCCGSIIQVCICAVWLGTS
PPFPDVDMHSEFGQIILWCNEGSTLAFYCVLGYLGFLSSLSLLIAFLARRLPESFNE
AKTITFSMLVFCTVWITFVPTYLSSKGKTMVAVEILSILASSSILLLCIFLPKCYVIL
LRSGGHSRKKFFK
>M_Vmn2r21
DMDQCIKCPEDQYPNKQKNNCLPKIKIFLSHEDTLGAVLVSLAISLSAFSAMILGL
FIHYRDTPIVRANNRNLSYVLLVSLMLCFFCSLIFIGQPSTVTCVLRQMIFGVVFSV
AVSAILAKTFIVVVAFTAIKPGSTLQMWMVTRLSNAIVCCGSIIQVCICAVWLGT
YPPFPDADMHSEFGQIILWCNEGSTLAFYCVLGYLGFLSSLSLLIAFLARRLPESFN
EAKTITFSMLVFCTVWITFVPTYLSSKGKTMVAVEILSILASSSILLLCIFLPKCYVI
LLRSGGHSRKKFFK
>M_Vmn2r23
DMDQCIKCPEDQYPNKQRNQCLPKIMAFLSHEDTLGALLVCTAISLSAFAAMILG
LFIHYRDTPIVRANNRNLSYLLLVSLMLCFFCSLIFIGQPGTVTCVLRQMIFGVVFS
IAVSAILAKTFIVVVAFKSIKPRSTLQMWIVTRLSITIVCCGSIIQVCICAVWLGTYP
PFPDVDMHSEFGHIILCCNEGSTLAFYCVLGYLGFLASLSLLIAFLVRRLPDSFNE
AKTITFSMLVFCSVWISFIPTYLSSTGKTMVAVEIFCILASSAGLLGCMFLPKCYVI
LLRSGDHSRKKVFK
>M_Vmn2r22
DMDQCIKCPEDQYPNKQKNNCLPKIKIFLSHEDTLGAVLVSLAISLSAFSAMILGL
FIHYRDTPIVRANNRNLSYVLLVSLMLCFFCSLIFIGQPSTVTCVLRQMIFGVVFSV
AVSAILAKTFIVVVAFTAIKPGSTLQMWMVTRLSNAIVCCGSIIQVCICAVWLGT
YPPFPDADMHSEFGQIILWCNEGSTLAFYCVLGYLGFLSSLSLLIAFLARRLPESFN
EAKTITFSMLVFCTVWITFVPTYLSSKGKTMVAVEILSILASSSILLLCIFLPKCYVI
LLRSGGHSRKKFFK
>M_Vmn2r24
DMDQCIKCPEDQYPNKQRNQCLPKIMVFLSHEDTLGAVLVSLAISLSAFSAMILG
LFVCYRDTSIVRANNRNLSYLLLISLMFCFFCSLIFIGQPSTVTCVLRQMIFGVVFS
VAVSAILAKTFIVVVAFKAIKPGSRLQMWMVMRLSNVIVCCGSIIQVCICAVWLG
TYPPFPDVDMHSEFGHIILCCNEGSTLAFYCVLGYLGFLASLSLLIAFLVRRLPDSF
NEAKTITFSMLVFCSVWISFIPTYLSSTGKTMVAVEIFSILASSAGLLGCMFLPKCY
VILLRSADLSRKKFFK
>M_Vmn2r26

DMDQCIKCPEDQYPNKQRNQCLPKIMVFLAHEDPLGTVLVSLAISLSAFSAMILG
LFICYRETPIVRANNRNLSYLLLISLKLCFSCSLMFIGQPRTVTCVLRQIIFGIVFSIV
ISAILAKTFIVVMAFKAIKPGSILKMGMVTRLSNAIVCCGSIIQVCICAVWLGTYPP
FPDVDMHSEFGQIILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEA
KTITFSMLVFCSVWISFVPAYLSSKGKTMVAVEILSILASSAGLLGCIFLPKCYVIL
LKSGGHSRKKFFK
>M_Vmn2r25
DMHQCIKCPEDQYPNKQRNQCLPKIMVFLAHEDPLGTVLVSLAISLSAFSAMILG
LFICYRETPIVRANNRNLSYLLLISLKLCFFCSLIFIGQPRTVTCVLRQIIFGIVFSIVI
SAILAKTFIVVVAFKAIKPGSILKMGMVTRLSNAIVCCGSIIQVCICAVWLGTYPPF
PDVDMQSEFGQVILWCNEGSTLAFYCVLGYLGFLASLSLLIAFLARRLPDSFNEA
KTITFSMLGFCSVWISFIPTYLSSKGKTMVAVEILSILASSAGLLGCIFLPKCYVILL
KSGGHSRKKFFK
>M_Vmn2r27
DMDQCIKCPEDQYPNEQRNQCLPKITAFLSHEDTLGAVLGSMAISFAAFSAMILG
LFIHYRDTPIVRANNRNLSYLLLVSLKLCFFCSLMFIGEPRRVTCLLRQMIFGVVF
SVALSAVLAKTFIVVVAFKVIKPGSTLQMRMATQISNAIVCCGSIIQVCICAVWLG
TYPPFPDVDMHSEFGQIILWYNEGSTLAFYCVLGYLGFLASLSLFIAFLVRSLPDS
FNEAKTITFSMLVFCSVWISFVFTYLRSKAKTMVAVEIVSILASSAGLLSFIFLPKC
YVILLRSGGYSRKKFFKSVS
>M_Vmn2r53
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFILALVALLLAGLAVLVLG
VFLKHRDTPVVRANNRTLSYFLLISLSLCALCALLFLGRPTVTTCLLRQTIFAVVF
TVAVSSVLAKTLTVVLAFRVTRPRSRIQVCLSSGASTSVVLIASLIQVVLCGIWLIT
SPPFPDRDMISEPRHTVIQCHEGSGANFFCVLGFLGFLAGGTFSVAFLARGLPDVF
NETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTAGLLGGIFIPKCY
IILLKPEKNTPAWLRQGHQVQHNRQ
>M_Vmn2r54
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFVLALVALLLAGLAVLVLG
VFLKHRDTPVVRANNRTLSYFLLISLSLCALCALLFLGRPTVTTCLLRQTTFAVVF
TVAVSSVLAKTLTVVLAFRVTRPRSRIQICLSPGTSTFVVLIASLIQVVLCGVWLA
TSPPFPDKDMISEPQHIVIQCQEGSGATFFCVLGYLGFLAGGTFSVAFLARGLPDIF
NETKFLTFSMLLFCSVWTAFLPLYHSARGKSTVAVEIFSILASTAGLLGGIFIPKCY
IILLKPEKNT
>M_Vmn2r55
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFILALVALLLAGLAVLVLG
VFLKHRDTPVVRANNRTLSYFLLISLSLCALCALLFLGRPTVTTCLLRQTTFAVVF
TVAVSSVLAKTLTVVLAFKVTRPRSRIQICLSPGTSTLVVLIASLIQVVLCGVWLA
TFPPFPDKDMLSEPQHIVIQCHDGSGATFFCVLGYLGFLAGGTFSVAFLARDLPD
VFNETKFLTFSMLLFCSVWTAFLPLYYSARGKSTVAVEIFSILASTAGLLGGIFIPK
CYIILLKPEKNT
>M_Vmn2r56
DMKRCLLCPKEQYSSHTRDHCLPRTEIFLAFEEPLGFILALVALLLAGLAVLVLG
VFLKHRDTPVVRANNRTLSYFLLISLSLCALCALLFLGRPTVTTCLLRQTTFAVVF
TVAVSSVLAKTLTVVLAFKVTRPRSRIQICLSPGTSTLVVLIASLIQVVLCGVWLA
TFPPFPDKDMLSEPQHIVIQCHDGSGATFFCVLGYLGFLAGGTFSVAFLARDLPD

VFNETKFLTFSMLLFCSVWTAFLPLYYSARGKSTVAVEIFSILASTAGLLGGIFIPK
CCIILLKPEKNTSSWLRQ
>M_Vmn2r57
DMDQCVQCPESHYANSEKNHCLQKSVSFLAYEEPLGIALTITALDFSVLTALVLV
VFVKHRDTPIVKANNRVLSYILLLTLIICFLSCLLYIGQPNTATCILQQTAFGTLFT
VALSTVLAKAIVVVTAFKVTSPSRMVRWLIVSRAPNLIIPICTLIQLIICGIWLATFP
PFIDQNAHVEHGHIIIMCNKGSAVAFHCVLGYLCSLALGSYTMAFLSRNLPDTFN
EAKCLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSILASSTSFLALIFVPKC
YIILLRPDKNSCLDIRNKIHSRRSS
>M_Vmn2r58
DMYQCMKCPESHYANSEKSHCLQKSVSFLAYDEPLGMALIITALCFSLLIAAVLV
VFVKHRDTPIVKANNRTLSYILLLTLFICLLSSFLFIGQPNTTTCILQQTAFGILFTV
ALSTVLAKAITVVIAFKATVPARMVRWLMVSRAPNFIIPICTLIQLIICGIWLVTSPP
FIDQDTQTEHGHIIIVCNMGSTVAFHTVLGYLCSLALGSYTMAFLSRNLPDTFNE
AKFLSFSMQVFFCVWITFLPVYHSSKGKVMVVMEVFSILASSVALFILIFAPKCYII
LLRPEKNSCHNVRHKKHSR
>M_Vmn2r60
DMEQCVQCPESHYANSEKNHCLQKSVSFLAYEEPLGMALTIIALCFTLLTAVVLV
VFVKHRDTPIVKANNRTLSYILLLTLFISFLSSLLFIGQPNTTTCILQQTAVGILFTV
ALSTVLAKAITVVIAFKATVPARMVRWLMVSRAPNFIIPICTLIQLIICGIWLVTSPP
FIDLDTHAEHGHIIIVCNMGSTVAFHTVLGYLCSLALGSYTMAFLSRNLPDTFNE
AKFLSFSMQIFFCVWITFLPVYHSTKGKVMVAMEVFCILASNIALLGLIFTPKCYII
LLRPEKNSYLGVRHK
>M_Vmn2r59
DMDQCVKCPESHYANSERNHCLQKSVSFLAYEEPLGMALTITAICFSLLTALVLA
VFVKNRDTPIVKANNRALSYILLLTLFICLLSSLLFIGQPNTITCILQQTAFGILFTV
ALSTVLAKAITVVIAFKATVPARMMRWLMISRAPNFIILICTLIQVIICGIWVVTSP
PFIDQDTQAEHGHIIIVCNMGSTVAFHTVLGYLCSLALGSYTMAFLSRNLPDTFNE
AKFLSFSIQVFFCVWITFLPVYHSTKGKVMVAMEVFSILASSIALLGLIFGPKCYII
LLRPEKISCLGIRHKKHSR
>M_Vmn2r61
DVDQCVKCSESHYVNSEKNYCLQKSVSFLAYEEPLGMALTITSLCFSLLTAVVLV
VFVKHRDTPIVKANNRTLSYILLLTLFISFLSSLLFVGQPNTTTCILQQTAFGILFTV
SLSTVLAKAITVVIAFKAIVPARMVRWLMVSRAPNFIIPICTLIQLIICGIWILTSPPF
IDQDTQAEHGHIIIVCNMGSTVAFHTVLGYLCSLALGSYTMAFLSRNLPDTFNEA
KFLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSILASSIALLGLNFAPKCYII
LLRPEKISCIDIRHTKYTRRNIHS
>M_Vmn2r62
DMDQCVQCPESHYANSEKNHCLQKSVSFLAYEEPLGMGLTITALDFSVLTAVVL
VVFVKHRDTPIVKANNRVLSYTLLLTLIICFLSCLLYIGQPNTATCILQQTAFGTLF
TVALSTVLAKAIVVVTAFKVTSPARMVRWLIVSGAPNSIIPICTLIQLIICGIWLATS
PPFIDQNAHVEHGHIIIMCNKGSAVAFHCVLGYLCFLALGSYSMAFLSRNLPDTF
NEAKYLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSVLASSIALLALIFAPK
CYIILLRPDKNSRLDIRHKIHSR
>M_Vmn2r64

DMDQCVQCPESHYANSEKNHCLQKSVSFLAYEEPLGMGLTITALDFSVLTAVVL
VVFVKHRDTPIVKANNRVLSYTLLLTLIICFLSCLLYIRQPNTATCILQQTAFGTLF
TVALSTVLAKAIVVVTAFKITSPARMVRWLIVSGAPNSIIPICTLIQLIIYGIWLATS
PPFIDQNAHVEHGHIIIMCNKGSAVAFHCVLGYLCFLALGSYSMAFLSRNLPDTF
NEAKYLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVFSVLASSIALLALIFAPK
CYIILLRPDKNSRLDIRHKIHSR
>M_Vmn2r63
DMDQCVQCPESHYANSEKNHCLQKSVSFLAYEEPLGMALTITALDFSVLTAVVL
VVFVKHRDTPIVKANNRVLSYTLLLTLIICFLSCLLYIGQPNTATCILQQTAFGTLF
TVALSTVLAKAIVVVTAFKVTSPARMVKWLIVIGAPNLIIPICTLIQLIICGIWLATS
PPFIDQNAHVEHGHIIIMCNKGSAVAFHCVLGYLCSLALGSYTMAFLSRNLPDTF
NEAKCLSFSMQVFFCVWITFLPVYHSTKGKVMVAMEVLSVLASSITLLVLIFAPK
CYIILLRPDKNSCLDIRHKIHSRKVVI
>M_Vmn2r28
NMDHCVNCPEYQYANTEQNKCIQKAVIFLSYEDPLGMALALMAFGFSAFTALV
LWAFVKHHDTPIVKANNRILSYLLLMSLMFCFLCTFFFIGHPNRVTCVLQQITFGI
VFTVAVSTVLAKTITVILAFKFTGPGRTLKNILVSGTPNYIIPICSLFQCILCAIWLA
VSPPFVDIDEHSEHGNIILVCNKGSVTAFYCVLGYLACLALGSFTVAFLARNLPD
AFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSSGILGCIFAP
KIYIIFLRPERNSIKKIREKSHF
>M_Vmn2r29
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYILLISLMSCFLCSFFFIGHPNRATCVLQQITFGIVF
TVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWLA
VSPPFVDIDEHTLHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPDT
FNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPERNSTQKIKEKSYF
>M_Vmn2r30
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACLALGSFSLAFLAKNLPDT
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVPK
IYIILMRPERNSTQKIREKSYF
>M_Vmn2r31
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYLLLISLMSCFLCSFFYIGHPNRATCVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWI
EVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD
TFNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFA
PKIYIILMRPERNSTQKIKVRNSTQKIKEK
>M_Vmn2r32
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNKVICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRYFLVSGTLNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD

AFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGMLECIFV
PKIYIILMRPERNSTQKIREKSYF
>M_Vmn2r33
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYILLISLMSCFLCSFFFIGHPNRATCVLQQITFGIVF
TVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWLE
VSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFSLAFLAKNLPDT
FNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPDRNSTQKIKVRNSTQKIKEK
>M_Vmn2r34
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMSCFLCSFLFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSVTAFYCILGYLACLALGNFSVAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIILMRPERNSTQKIREKSYF
>M_Vmn2r35
NMDHCVNCPEYQYASTEQNKCIQKVVTFLSYEDPLGMALALIAFCFSAFTAVVL
CVFVKHHDTPIVKANNRILSYILLLSLMSCFLCSFFFIGHPNRATCVLQQITFGIVF
TVALSTVLAKTVTVVLAFKITDPGRRLRHFLVSGTPNYIIPICSLLQCVLCAIWLA
VSPPFVDIDEHSQNGHIIIVCNKGSVTAFYCLLGYLACLALGSFTVAFLAKNLPDT
FNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASIAGMLGCIFAPK
IYIILMSPERNSTQKIKENNSTQKIKEK
>M_Vmn2r37
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMSCFLCSFFFIGLPNRAICVLQQITFGIV
FTMAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSVTAFYCILGYLACLALGNFSVAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIILMRPERNSTQKIREKSYF
>M_Vmn2r36
NMDQCMNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACLALGSFSLAFLAKNLPDT
FNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVPK
IYIILMLPERNSTQKIREKSYF
>M_Vmn2r38
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKYHDTPIVKANNRILSYILLISLMSCFLCSFFFIGHPNRATCVLQQITFGIVF
TVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWLE
VSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGFLACLALGSFSLAFLAKNLPDT
FNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPDRNSTQKIKVRNSTQKIKEK
>M_Vmn2r39

NMDQCVNCPEYQYANTEQNKCIKKGVTFLSYEDPLGMVLALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYILLLSLMSCFLCSFFFIGHPNRATCVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRHFLVSGTPNYIIPICSLFQCVLCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFIVAFLAKNLPD
TFNEAKFLTFSMLVFFSVCVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPERNSTQNIKESNSTQKIKEKSNS
>M_Vmn2r40
NMDQCMNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTALV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVIVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACMALGSFSLAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIILMRPERNSTQKIREKSYF
>M_Vmn2r41
NMDQCMNCPEYQYANTEQNKCIQKGVTFLSYENPLGMALALMAFCFSAFTALV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVIVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACMALGSFSLAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIILMRPERNSTQKIREKSYF
>M_Vmn2r42
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNKVICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTVPGRRLRYFLVSGTLNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD
AFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGMLGCIFV
PKIYIILMRPERNSTQKIREKSYF
>M_Vmn2r43
NMDQCVNCPEYQYANTEQNKCIKKGVTFLSYEDPLGMVLALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYILLLSLMSCFLCSFFFIGHPNRATCVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRHFLVSGTPNYIIPICSLFQCVLCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFIVAFLAKNLPD
TFNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFA
PKIYIILMRPERNSTQNIKESNSTQKIKEKSNS
>M_Vmn2r44
NMDKCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYLLLISLMSCFLCSFFYIGHPNRATCVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLFQCVLCAIWL
EVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD
AFNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFA
PKIYIILMRPERNSTQKIKVRNSTQKIKEK
>M_Vmn2r45
NMDQCMNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVALSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSVTAFYCILGYLACLALGSFSVAFLAKNLPD

AFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIILTRPERNSTQKIREKSYF
>M_Vmn2r65
DMDQCVKCPNDQYANIKQTHCLKKVVTFLAYEDPLGMTLTCLALLFSALTTVIL
SILLKHQDNPIVKANNRLLSYILLISLIFCFLCSLLYIGHPHMATCILQQTTFAVVFT
VAVSTILAKTITVVMAFKITDPRKKIRQMLVTRAPNYIIPICTMIQLILCGTWIGTSP
PFVDSDPHFEHGHIIIVCNKGSVIAFYCVLGYLGLLALGSFTVAFLARNLPDRFNE
AKFLTFSMLVFCSVWITFLPVYHSTKGKAMVVVEIFSILTSSAGLLVCIFFPKCFTI
LLKQKVNFTQKFSDTHSKIE
>M_Vmn2r66
NMDQCVKCPEDQYANEDQTLCLQKVVAVLDYRDPLGKALAGFALCFSVLTSVV
ICVFLKHRETPIVKANNQTLSYVLLISLIFCFMSSLLYIGHPTMVICILQQTTFAIVF
TVAASTILAKTITVVLAFKATVPGRRMRWLLVSGAPKYIILICTMVQLILCGIWLR
ISPPFVDADIHMLHGQIIIVCNKGSVVAFYCVLGYMGSIALASFTVAFLARNLPHT
FNEAKLLTFSMLVFCSVWITFIPVYHSTKGKTTVAVEVFCILASSAGLLLCIFAPK
CYLILLRPQTNSFYKFRRPHAKAENIS
>M_Vmn2r67
NMDQCVKCQEDQYANEDHTLCLQKVVSILDYRDPLGKALAGIALCFSVLTSGIL
CIFLKHRETPIVKANNQTLSYVLLISLVFCFICSLLYIGHPTTVICILQPITFAIAFTV
ATSTILAKTITVLLAFKITIPGRRMRWLLLSGVPKYIVLICTMIQLILCGIWLGTSPP
FVDADVYMVHGHIIIVCNKGSVIAFYCVLGYMGSVALASFTVAFLARNLPDTFN
EAKLLTFSMLVFCSVWITFIPVYHSTKGKTMVAVEVFSILASSAGLLLCIFAPKCY
IILLRPQTNSFHKFRKTHSKAENIS
>M_Vmn2r68
NMDQCVKCREDQYANEHHTLCLQKVVSILDYRDPLGKALAGIALCFSFLTSVVL
CIFLKHRETPIVKANNQTLSYVLLISLVFCFICSLLYIGHPTAVICILQQTTFVIAFTV
ATSTILAKTITVLLAFKITIPGRRMRWLLVSGAPKYIILICTMIQLILCGIWLGTSPPF
VDADAHMVHGHIIIVCNKGSVIAFYCVLGYMGSVALASFTVAFLARNLPDTFNE
AKLLTFSMLVFCSVWITFIPVYHSTKGKTMVAVEVFCILASSAGLLLCIFAPKCYII
LLRPDINSFQKFRK
>M_Vmn2r69
NMDQCVKCLEDQYANEDHTLCLQKVVAILDYRDPLGKALAGFALCFSLLTSVV
LFVFLKHRDTPIVKANNQTLSIVLLISLIFCFICSLLYIGHPTMVICILQQTTFAIAFT
VATSTILAKTITVVLAFKITVPGRRMRWLLVSGAPKYIILICTMIQLILCGIWLGTS
PPFVDADVQMVHGHIIIVCNKGSVIAFYCVLGYMGSVALAGFTVAFLARNLPDT
FNEAKLLTFSMLVFCSVWMTFIPVYHSTKGKTMVAVEVFCILASSAGLLLCIFAP
KCYFILRPQTNSFHKFRKPHAKTENKS
>M_Vmn2r70
NMDQCVKCPENQYANEDHTLCLEKVVAILDYKDPLGKALAGFALCFSVLTSVV
LGIFLKNRDTPIVKANNQSLSFVLLISLIFCFICSLLYIGHPTMVICILQQTTFAIVFT
VATSTILAKTVIVVLAFKITVPGRRMRWLLEIGAPKYIILICTIIQLILCGIWLGTSPP
FVDADVHMVHGHIIIVCNKGSVIAFYCVLGYMGSVALASFTVAFLARNLPDTFN
EAKLLTFSMLVFCSVWITFIPVYHSTKGKTMVAVEVFSILASSAGLLLCIFAPKCY
IILLKPQKNSFQKFRKPHAIADNIS
>M_Vmn2r71

NMDQCVKCLEDQYANEDHTLCLQKVVVIFDYRDPLGKALAGFALCFSVLTSVV
LFVFFKHRDTPIVKANNQTLSFVLLISLIFCFICSLLYIGHPTMVTCIVQQITFAIGFT
VATSTILAKTVIVVLAFKITVPGRRMRWMPESGAPKYIILICTMIQLILCGIWLGTS
PPFVDADVQMVHGHIIIICNKGSVIAFYCVLGYMGSLALASFTVAFLARNLPDTF
NEAKLLTFSMLVFCSVWITFIPVYHSTKGKIMAVVEVFCILASSAGLLLCIFAPKC
YIILLRPQTNYFHKFWKSHANSENKS
>M_Vmn2r72
NMDQCVKCREDQYANEDHTLFLQKVVSILDYREPLGKALAGFALCFSVLTSGIF
CVFLKHRETPIVKANNQTLSYVLLISLIFCFICSLLYIGHPTTVICILQQTTFAIAFTV
ATSTILAKTITVLLAFKITIPGRRMRWLLVSGAPKYIILICTMIQLILCGIWLGTSPP
YVDADVHMVHGHIIIVCNKGSVIAFYCVLGYMGSFALTSFTVAFLARNLPDTFNE
AKLLTFSMLVFCSVWITFIPVYQSTKGKTMVAVEVFCILASSAGLLLCIFAPKCYII
LLRPHINSFRNFRKPHAKTENIS
>M_Vmn2r73
NMDQCVKCPEDQYANEDHTLCLQKIVAILDYRDPLGKALAGFALCFCVITSGVL
CIFLKHRETPIVKANNQTLSYVLLISLIFCFICSLLYIGHPTIVICVLQQSTFAISFTV
AASTILAKTITVVLAFKVTVPGRMRWLLVAGAPKYIILICTMIQLILCGVWLGTSP
PFVDADLHMVHGHIIIVCNKGSVIAFYFVLGYMGSIALASFTVAFLARNLPDTFN
EAKLLTFSMLVFCSVWVTFIPVYYSTKGKTMVAVEVFCILASSAGLLLCIFAPKC
YNILLRPGKNSSHMFKKLHAKTENII
>M_Vmn2r74
NMDQCVKCPEDQYANEDHTVCLQKVVTILDYRDPLGKALAGFALCFSVLTSLV
LVVFLKHRDTPIVKANNQTLSFVLLISLIFCFICSLLYIGHPTMAICILQQTTFAIVFT
VATSTILSKTVIVILAFKITVPGRRMRWLLESGTPKYIIPICTIIQLILCGIWLGTSPPF
VDADLHMVHGHILIVCNKGSVIAFYCVLGYMGSVALASFTVAFLARNLPDTFNE
AKLLTFSMLVFCSVWITFIPVYHSTKGKTMVAVEVFCILASSAGLLLCIFAPKCYII
LLRPQKNYFYKFRKRRE
>M_Vmn2r75
NMDQCVKCPEDQYANEDQTLCLQKVVDVLDYRDPLGKSLAGFALCFSVLTSIV
LCVFLKHRESPIVKANNQTLSYVLLISLIFCFICSLLYIGHPTMFICILQQTAFAIAFT
VAASTVLAKTITVILAFKITVPGRMRWLLVSGAPKYIIFVCTMIQLIFCGIWLGTSP
PFVETDVHMTHGHIIIVCNKGSVIAFYCVLGYMGSVALASFTVAFLSRKLPDTFN
EAKLLTFSMLVFCSVWITFIPVYHSTKGKTMVAVEVFCILASSAGLLLCIYAPKCY
IILLRPQKNSFYKFRKPHSKSENIS
>M_Vmn2r76
DMDYCFKCPDNQYANPEGTHCLSKIVTFLSYEDPMGVVLACLALGFSALTAAVF
GIFLRNQDTPIVKANNRALSYCLLISLIYCFLCSLLFIGQPHILTCIMQQTIFAVVFT
VAISTVLTKTITVVLAFRFTLPSKRMRWLMTSGASILIIPICTMIQLIICGIWLGTSPP
FVDSDGHIEHGHILVVCNKGSVIAFYCVLGYLGSIALTSFTVAFLARNLPDTFNEA
KFLTFSMLVFCSVWLTFLPVYHSTNGKALVAVEVFCILTSTAGIFLCIFAPKCYIIL
LKPENKSFHKFRHKNANV
>M_Vmn2r77
DMDQCVKCPEDEYANTNQTKCFKKVVTFLAYEDPSGMALAVLAVCFSVLTSVV
LCVFLKHRNTPIVKANNETLSYVLLTSLIFCFICSLLYIGYPTMVSCILQQTTFAIVF
TVAASCVLAKTITVVLAFKITVPGRRLRWLLLSGAPNYIIPICTMIQMILCGIWLGT
SPPFVDADLHMVHGHIIIVCNKGSVIAFYCVLGYMGSLALSSFTVAFLARNLPDT

FNEAKLLTFSMLVFCSVWITFLPVYHSTKGKAMVAVEVFCILASSAGMLVCIFLP
KCYIILLRPQINSFHDFRKIHVKV
>M_Vmn2r78
DMDQCVKCPEDEYANTDQTNCLKKVVTFLDYKEPLGMALAVLAVFFSSLTFIIL
CVFLKHRDTPIVKANNETLSYVLLTSLIFCFICSLLYIGYPTMVSCILQQTTFALVF
TVAASCVLAKTITVVLAFKITVPGRRLRWLLLSGAPNYIIPICTTIQMILCGIWLGT
SPPFVDADLHMIHEHIIIVCNKGSVTAFYCVLGYMGSLALASFTVAFLARNLPDTF
NEAKLLTFSMLVFCSVWVTFLPVYHSTKGKAMVAVEVFCILASSAGMLLCIFVP
KCYLILLRPQINSFHEFRKIHAK
>M_Vmn2r79
DMDQCLKCPEDEYANTDQTNCLKKVVTFLDYKEPLGMALTGLAVLFSSLTVVV
LCVFLKHRDTPIVKANNETLSYVLLISLIFCFICSLLYIGYPSMVSCILQQTTFAMV
FTVSASCVLAKTITVVLAFKITVPGRRLRWLLLSGAPNYIIPICTTIQMMLCGIWLG
ISPPFVDADLHMVHGHIIIVCNKGSVIAFYCVLGYMGSLALASFTVAFLARNLPDT
FNEAKLLTFSMLVFCSVWITFLPVYHSTKGKAMVAVEVFCILASSAGMLLCIFVP
KCYLILWRPQINSFHEFRKMHVKNKNI
>M_Vmn2r46
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRILSYILLISLMSCFLCSFFFIGHPNRATCVLQQITFGIVF
TVAVSTVLAKTVTVVLAFKVTDPGRRLRNFVVSGTPNYIIPICSLLQCVLCAIWLA
VSPPFVDIDEHTLHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPDT
FNEAKFLTFSMLVFFSVWVTFLPVYHSTKGKVM
>M_Vmn2r47
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRAICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRNFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACMALGSFSLAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGILGCIFVP
KIYIIFMRPERNSTQKIREKSYF
>M_Vmn2r48
NMDQCMNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTALV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRVICVLQQITFGIV
FTVAVSTVLAKTVIVVLAFKVTDPGRRLRYFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACLALGSFSVAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGMLGCIFV
PKIYIILMRPERNSTQKIREKSYF
>M_Vmn2r49
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEDPLGMALALMAFCFSAFTAVV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNKVICVLQQITFGIV
FTVAVSTVLAKTVTVVLAFKVTDPGRRLRYFLVSGTLNYIIPICSLLQCELCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD
AFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGMLGCIFV
PKIYIILMRPERNSTQKIREKSYF
>M_Vmn2r50
NMDQCVNCPEYQYANTEQNKCIKKGVTFLSYEDPLGMVLALMAFCFSAFTALV
LCVFVKHHDTPIVKANNRSLSYILLLSLMSCFLCSFFFIGHPNRATCVLQQITFGIV

FTVAVSTVLAKTVTVVLAFKVTDPGRRLRQFLVSGTPNYIIPICSLFQCVLCAIWL
AVSPPFVDIDEHSQHGHIIIVCNKGSVTAFYCVLGYLACLALGSFTLAFLAKNLPD
TFNEAKFLTFSILVFFSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGMLGCIFAP
KIYIILMRPERNSTQNIKESNSTQKIKEK
>M_Vmn2r51
NMDQCVNCPEYQYANTEQNKCIQKGVTFLSYEEPLGMALALMAFCFSAFTALV
LCVFVKHHDTPIVKANNRSLSYLLLMSLMFCFLCSFFFIGLPNRVICVLQQITFGIV
FTVAVSTVLAKTVIVVLAFKVTDPGRRLRYFLVSGTPNYIIPICSLLQCVLCAIWL
AVSPPFVDIDEHTLHGHIIIVCNKGSDTAFYCILGYLACLALGSFSVAFLAKNLPD
TFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKHMVAVEIFSILASSAGMLGCIFV
PKIYIILMRPERNSTQKIREKSYF
>M_Vmn2r52
NMDHCVNCPEYQYANTEQNKCMQKAVIFLRYEDPLGMALALMAFCFSAFTAL
VLWAFVKHHDTPIVKANNRILSYLLLMSLMFCFLCTFFFIGHPNRVTCVLQQITF
GIVFTVAVSTVLAKTVTVILAFKFTGPGRRLRNILVSGTPNYIIPICCLFQCVLCAI
WLAVSPPFVDIDEHSEHGNIILVCNKGSVTAFYCVLGYLACLALGSFTVAFLARN
LPDTFNEAKFLTFSMLVFCSVWVTFLPVYHSTKGKVMVAVEIFSILASSAGILGCI
FAPKIYIIFLRPERNSIKKIREKSHF
>M_Vmn2r121
DMEQCVRCPDDKYANIEQTHCLSRAVSFLAYEDPLGMALGCMALSFSAITILVL
VTFVKYKDTPIVKANNRILSYILLISLVFCFLCSLLFIGHPNQVSCILQQTIFGVFFT
VSVSTVLAKTITVVMAFKLTTPGRRIRGMMMTGAPKLVIPICTLIQLVLCGIWLV
TSPPFIDRDIQSEHGKTVILCNKGSVIAFHVVLGYLGSLALGSFTLAFLARNLPDTF
NEAKFLTFSMLVFCSVWITFLPVYHSTRGRVMVVVEVFSILASSAGLLMCIFVPK
CYVILIRPDSNFIKKHKGK