3000 Ci/mM) (UTP) was then added, and the nuclear suspension was incubated at 30°C for 30 minutes, after which time  $15 \,\mu$ l of DNase I (5  $\mu$ g/ minutes, after which time 15 µJ of DNase 1(5 µg/ mi) in 10 mM CaCl\_(5 µg/ml) was added. After 5 minutes at 30°C, the reaction was made 1× SET (1 percent sodium dodecyl sulfate (SDS), 5 mM EDTA, 10 mM tris-HCl, pH 7.4), and proteinase K was added to a concentration of 200 µg/ml. After incubation at 37°C for 45 minutes, the solution was extracted with an equal volume of a mixture of phenol and chloroform, and the interphase was again extracted with 100  $\mu$ l of 1× SET. Ammonium acetate (10*M*) was added to the combined aqueous phases (original plus reextraction) to a final concentration of 2.3M, an recvataction of a linar concentration of 2.3%, an equal volume of isopropyl alcohol was added, and nucleic acid was precipitated ( $-70^{\circ}$ C for 15 minutes). The precipitate was centrifuged in a microcentrifuge for 10 minutes, and the pellet was resuspended in 100 µl of TE (10 mM tris-HCl, 1 mM EDTA) and centrifuged through a G-50 (medium) spin column. The aluate was made 50 (medium) spin column. The eluate was made 0.2*M* in NaOH and after 10 minutes on ice, HEPES was added to a concentration of 0.24*M*. Two and one-half volumes of ethanol were then added, and the solution containing the precipi-

tate held overnight at  $-20^{\circ}$ C. After centrifuga-tion in a microcentrifuge for 5 minutes, the pellet was resuspended in hybridization buffer, which consisted of [10 mM TES, pH 7.4, 0.2 percent SDS, 10 mM EDTA, 0.3M NaCl, 1× Den-hardt's, and *Escherichia coli* RNA (250 µg/ml)]. Nitrocellulose filters containing plasmid DNA's Nitrocellulose filters containing plasmid DNA's were prepared with a Schleicher & Schuell Slot Blot Apparatus under conditions suggested by S and S, except that wells were washed with 10× SSC (saline sodium citrate). These filters were first hybridized in the hybridization solution described above for a minimum of 2 hours at 65°C. After this preliminary hybridization, the filters were hybridized to the runoff products in hybridization solution for 36 hours. A typical reaction contained 2 ml of hybridization solution with  $1 \times 10^7$  cpm/ml. After hybridization, filters were washed for 1 hour in 2× SSC at 65°C. The were washed for 1 hour in  $2 \times SSC$  at  $65^{\circ}C$ . The filters were then incubated at  $37^{\circ}C$  in  $2 \times SSC$ with RNase A (10 mg/ml) for 30 minutes and were subsequently washed in  $2 \times SSC$  at  $37^{\circ}C$ for 1 hour. Alternatively, after hybridization the filters were washed twice for 15 minutes in 0.1 percent SDS,  $2 \times SSC$  at room temperature, and then washed at  $60^{\circ}C$  (0.1 percent SDS,  $0.1 \times$ 

SSC) for 30 minutes. Either protocol for proc-essing of the filters after hybridization yielded the same specificity in signal. Filters were then exposed to Kodak XAR film in cassettes con-taining Lightening-Plus screens at  $-70^{\circ}$ C for various times.

- C. Yanisch-Perron, J. Vierra, J. Messing, Gene 45.
- C. Yanisch-Perron, J. Vierra, J. Messing, Gene 33, 103 (1985).
   S. L. McKnight, E. R. Gavis, R. Kingsbury, R. Axel, Cell 25, 385 (1981).
   M. Groudine and C. Casimir, Nucleic Acids Res. 12, 1427 (1984).
   We thank many of our colleagues for discussion
- We thank many of our colleagues for discussion and suggestions during the course of this work; Hal Weintraub, Paul Neiman, and Craig Thomp-the state of the stat Son for comments on the manuscript; Craig Thompson for assistance in obtaining lympho-cyte preparations; Bill Schubach for plasmid pBK25; and Kay Shiozaki for assistance with the manuscript. Supported by NIH grants CA 18282 (M.L.) and CA 28151 (M.L. and M.G.), and NSF grant PCM 82-04696 (M.G.), and a scholarship from the Leukemia Society of America (M.G.) on for comments on the manuscript; Craig

30 July 1985; accepted 15 October 1985

## RESEARCH ARTICLE

## **Tyrosine Kinase Receptor with Extensive** Homology to EGF Receptor Shares Chromosomal Location with neu Oncogene

Lisa Coussens, Teresa L. Yang-Feng, Yu-Cheng Liao Ellson Chen, Alane Gray, John McGrath, Peter H. Seeburg Towia A. Libermann, Joseph Schlessinger, Uta Francke Arthur Levinson, Axel Ullrich

Growth factors and their receptors are involved in the regulation of cell proliferation, and several recent findings suggest that they also play a key role in oncogenesis (1-4). Of approximately 20 identified oncogenes, the three that have been correlated with known cellular proteins are each related to either a growth factor or a growth factor receptor. The B chain of platelet-derived growth factor (PDGF) is encoded by the proto-oncogene c-sis (2), the erb-B oncogene product gp68 is a truncated form of the epidermal growth factor (EGF) receptor (3), and the proto-oncogene c-fms may be related or identical to the receptor for macrophage colony-stimulating factor (CSF-1<sup>R</sup>) (4).

The receptor-related oncogenes are members of a gene family in that each has tyrosine-specific protein kinase activity, and is associated with the plasma membrane (5). Such features are also shared by several other polypeptide hormone receptors, including those for insulin (6), PDGF (7), and insulin-like growth factor 1 (IGF-1) (8); hence more connections may be found between tyrosine kinase growth factor receptors and tyrosine kinase oncogene products.

Comparison of the complete primary structure of the human EGF receptor (9) with the sequence of the avian erythroblastosis virus (AEV) transforming gene, v-erbB (10), revealed close sequence similarity; in addition, there were amino and carboxyl terminal deletions that may reflect key structural changes in the generation of an oncogene from the gene for a normal growth factor receptor (3, 9). Another oncogene, termed neu, is also related to v-erbB and was originally identified by its activation in ethylnitrosourea-induced rat neuroblastomas (11).

In contrast to v-erbB, which encodes a 68,000-dalton truncated EGF receptor, the neu oncogene product is a 185,000dalton cell surface antigen that can be detected by cross-reaction with polyclonal antibodies against EGF receptor (11); neu may itself be a structurally altered cell surface receptor with homology to the EGF receptor and binding specificity for an unidentified ligand.

Using v-erbB as a screening probe, we isolated genomic and cDNA clones coding for an EGF receptor-related, but distinct, 138,000-dalton polypeptide having all the structural features of a cell surface receptor molecule. On the basis of its structural homology, this putative receptor is a new member of the tyrosine-specific protein kinase family. It is encoded by a 4.8-kb messenger RNA (mRNA) that is widely expressed in normal and malignant tissues. We have localized the gene for this protein to q21 of chromosome 17, which is distinct from the EGF receptor locus, but coincident with the *neu* oncogene mapping position (12). We therefore consider the possibility that we have isolated and characterized the normal human counterpart of the rat neu oncogene.

Tyrosine kinase-type receptor gene and complementary DNA. As part of our attempts to isolate and characterize the chromosomal gene coding for the human cellular homologue of the viral erbB gp68 polypeptide, AEV-ES4 erbB sequences (2.5-kb Pvu II fragment of pAEV) (13) were used as a <sup>32</sup>P-labeled hybridization probe for the screening of a human genomic DNA library at reduced stringency

Lisa Coussens, Yu-Cheng Liao, Ellson Chen, Alarie Gray, Peter H. Seeburg, Arthur Levinson, and Axel Ullrich are in the Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080; John McGrath is currently with the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; Towia Libermann and Joseph Schlessinger are in the Department of Chemical Immunology at the Weizmann Institute of Science, Rehovot 76100, Israel; and Teresa L. Yang-Feng and Uta Francke are in the Department of Human Genetics at Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510.

(14). Clone  $\lambda c$ -erbB/1 was isolated; it contained a hybridizing 1.8-kb Bam HI fragment, which was subjected to DNA sequence analysis. The 1838-bp sequence contains three complete and one partial erbB-homologous exons separated by short intervening sequences (Fig. 1). Comparison of this human gene sequence with our complete cDNA-derived human EGF receptor protein sequence (9) revealed 32 differences (18.7 percent) within the 171 amino acid stretch of combined exons, suggesting that this gene fragment was not derived from the human EGF receptor gene. Since this gene may code for an unknown tyrosine kinase-type receptor that is closely related to the human EGF receptor, we named it HER2.

Northern blot analysis (15) with the <sup>32</sup>P-labeled 1.8-kb HER2 fragment as a hybridization probe revealed a 4.8-kb mRNA in human term placenta  $poly(A)^+$ RNA, distinct from the 5.8- and 10.5-kb EGF receptor mRNA's also present at high levels in this tissue (Fig. 2a, lane 1). Thus, we had isolated a portion of an EGF receptor-erbB-related but distinct gene. To obtain its complete primary structure, two single-stranded synthetic oligonucleotide probes (16) were prepared from HER2 exon sequence regions that differed sufficiently (less than 60 percent nucleotide sequence homology) from EGF receptor DNA sequences (Fig. 1, 1 and 2) and used to screen a term placenta complementary DNA (cDNA) library of  $2 \times 10^6$  independent recombinant clones in Agt10 (17). Fiftytwo clones were isolated; they hybridized strongly with both synthetic probes and weakly with an EGF receptor cDNA fragment (HER64-3) (9) containing the homologous region within the tyrosine kinase domain. One of these,  $\lambda$ HER2-436, had the longest cDNA insert (4.5 kb), consisting of three Eco RI fragments (1.4, 1.5, and 1.6 kb).

The complete cDNA sequence of this clone is shown in Fig. 3. The longest open reading frame starting with a methionine codon codes for a 1255 amino acid polypeptide (137,828 daltons) and contains the 171 residues encoded by the four exons in the 1.8-kp HER2 gene Bam HI fragment (Fig. 1). This 3765-bp coding sequence is flanked by 150 bp of 5' untranslated sequence and a TGA stop codon, followed by a 627-nucleotide 3' untranslated sequence. No stop codon is found in the 5' untranslated region. In support of our assignment, however, the initiation codon at position 151 is flanked by sequences that follow perfectly Kozak's rule (18) for translation initiation. The 3' untranslated sequence contains a 6 DECEMBER 1985

potential poly(A) addition signal sequence (AATATA) 12 nucleotides upstream from a stretch of 15 adenylate residues. We are not certain if this  $(A)_{15}$ stretch is part of a poly(A) tail or represents an internal poly(A) stretch of a longer 3' untranslated sequence.

those for EGF and insulin (9, 19). Such features are apparent in the hydropathy profile (20) comparison (Fig. 4a). On the basis of this comparison, and on amino acid sequence alignment with the EGF receptor (Fig. 4b, region 1), we predict a 21 amino acid signal sequence (Fig. 4b,

Abstract. A novel potential cell surface receptor of the tyrosine kinase gene family has been identified and characterized by molecular cloning. Its primary sequence is very similar to that of the human epidermal growth factor receptor and the v-erbB oncogene product; the chromosomal location of the gene for this protein is coincident with the neu oncogene, which suggests that the two genes may be identical.

Comparison of EGF receptor and HER2 sequence. As already indicated by the v-erbB sequence homology used to isolate HER2, the putative HER2 protein is very similar in its overall domain organization and sequence to the EGF receptor. Nevertheless, there are differences that are likely to define a specific biological role for the HER2 polypeptide.

The predicted HER2 polypeptide contains each of the domain features found in hormone receptor precursors, such as

1), an amino terminal serine residue, and a 632 amino acid putative extracellular ligand-binding domain; a highly hydrophobic, 22-amino acid transmembrane anchor domain separates the extracellular domain from a 580-residue-long carboxyl-terminal cytoplasmic domain, which possesses the highest homology to v-erbB and other members of the tyrosine kinase family.

The 632-amino acid, putative HER2 ligand binding domain is about 40 percent homologous with the 621-residue

1	740 Glu Lys Glu Ala 769 I leProAspGlyGluAsnValLysI leProValAlaIleLysValLeuArgGluAsnThrSerProLysAlaAsnLysGluIleLeuAsp GGATCCCTGATGGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAGCCAACAAAGAATCTTAGACGTAAGCCCCTCCACCCTCCTGCTAGG
121	${\tt AGGACAGGAAGGACCCCATGGCTGCAGGTCTGGGCTCTGGTCTCTCTTCATTGGGGTTTGGGGAGATATGACTCCCGCAAACCTAGACTATTTTTTGGAGACGGAGCTTGCTCTGTCAC}$
241	ccaggctggagtgcagtggcgttatctcggctcactgcaacctccaggctctaggcgattttcatgcctcaggctcctgagtagctgggattacaagcgcccgctaatttttttt
361	TTTTTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTAGAGTGAAATGGTGCGGTCTCAGCTCCAGCCTCCCAGGTTAAAGCGATTCTTCTCCCCTCAGTCTCCTGAGTAGCTGGGATTA
481	${\tt caggcgcgagccaccaccgcccggctaatttttgtatttttagtagagatgggatttcaccatgttggccaggttggtgtcaaactcctgacctcatgatccgccccgcctccgacctcccaggttggtgtgtgt$
601	$\label{eq:constructed} agtgctgcgattacaggtgtgagccacgtgccccggcctaatctttgtatttttagtagagacagggtttcaccatgttgtccaggctggtactttgagccttcacaggctgtgggccatggggccatggtggggccatggtgtggggccatggtggggccatggtggggccatggtggggccatggtggggccatggtggggccatggtgggggggg$
721	770 Ser AspAsn Hi GCTGTGGTTTGTGATGGTTGGGAGGCTGTGTGGTGTTTGGGGGTGTGGGTCTCCCATACCCTCTCAGCGTACCCTTGTGCAGAGAGCTTAGGTGGTGGTGGTGTTGGGGGTGTGGGTGTGGGTCCCCATA
841	s Cys IIe Phe Tyr HisLysAspAsnIle Ty rValSerArgLeuLeuG1yI1eCysLeuThrSerThrValG1nLeuVa1ThrG1nLeuMetProTyrG1yCysLeuLeuAspH1sVa1ArgG1uAsnArgG1yArgLeuG1ySerG1nAs TGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTTATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGG
961	r Val 831 pLeuLeuAsnTrpCysMetGInIleAlaLys CCTGCTGAACTGGTGTATGCAGATTGCCAAGGTATGCACCTGGGGCTCTTTGCAGGGTCTCTCCGGAGCAAACCCCTATGTCCACAAGGGGCTAGGATGGGGACTCTTGCTGGGCATGTGGC
1081	832 Asn Arg Thr G1yMetSerTyrLeuG1uAspValArgLeuValHisArgAspLeuAlaAlaArgAsnValLeuValLysSer CAGGCCCAGGCCCTCCCAGAAGGTCTACATGGGTGCTTCCCATTCCAGGGGATGAGCTACCTGGACAGTGCCGGTCGTACACAGGGACTTGGCCGCTCGGCAAGAGT 
1201	GIn Lys GIyAlaGlu Lys GIu 883 ProAsnHisValLysI]eThrAspPheGIyLeuAlaArgLeuLeuAspI]eAspGIuThrGIuTyrHisAlaAspGIyGIyLys CCCAACCATGTCAAAATTACAGACTTCGGCGFGCGGCGACAGATTGACGAGACAGAGTACCAAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1321	GGAGTGGTGTCTAGCCCATGGGAGAACTCTGAGTGGCCACTCCCACAACACACAGTTGGAGGACTTCCTCTTCTGCCCCCCCGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCC
1441	His Iletyr ArgArgArgPheThrHisGInSerAspValTrpSerTyrGIyVal CSCCSGCGGTCACCCACCAGAGTGATGTGTGGGAGTTATGGTGTGTGAGGGGGGGTGGGAGGGGGG
1561	CGGGGCCACCTCAGCATGTGAAGGGAGGGAAGGGGCTGCCTGTGCCCCACCTTGCAGGGTCTGTGCACTTCCCAGGATTAGGGAAAGACCGGGTAGGGTCTGTCT
1681	$\tt cccctgctacctgccatgatgctagactcctgagaacctctggccagaacctctggtacactaaagctccctctggccctcccactcctgaccctgtctctggcgtgggggggg$

1801 GCTGATGACTTTTGGGGGCCAAACCTTACGATGGGATCC

Fig. 1. Partial sequence of the HER2 gene. A partial Hae III-Alu I genomic library (14) of human fetal DNA in  $\lambda$  Charon 4A was screened using a radiolabeled 2.5-kb Pvu II fragment of pAEV (13) containing coding sequences for the tyrosine kinase domain. Hybridization was as described elsewhere (31), except that 30 percent formamide was used at 42°C. Three independent clones were isolated which shared a 1.8-kb hybridizing Bam HI fragment. This fragment and subsets thereof were isolated, subcloned into M13mp10 and M13mp11, and sequenced (32). The intron-exon organization was determined by comparison with v-erbB sequences (10). Amino acid numbering is based on the complete cDNA sequence shown in Fig. 3. Nucleotide sequence differences with the human EGF receptor sequence are shown in the regions that were used for the design of synthetic oligonucleotide probes 1 (30 nucleotides) and 2 (21 nucleotides). Amino acid sequence differences with the EGF receptor are shown above the HER2 sequence.

extracellular EGF binding domain of the EGF receptor. This homology includes two cysteine-rich subdomains of 26 and 21 regularly organized cysteine residues (Figs. 4a and 2c, subdomains 2 and 3), all of which are conserved in the EGF receptor. The cysteine residue spacing in this region is also homologous with the single cysteine-rich domain in the insulin receptor  $\alpha$  subunit (19). In contrast, HER2 contains only eight potential Nlinked glycosylation target sites (Asn-X-Thr or Ser) as compared to 12 in the corresponding region of the EGF receptor. Only five of these are conserved with respect to their relative position in each polypeptide.

The hydrophobic, putative membrane anchor sequence located between residues 653 and 676 (Fig. 4b, region 4) is flanked at its carboxyl terminus by a stretch of amino acids of predominantly basic character (KRRQQKIRKYTMRR) (21), as is found in the EGF receptor sequence (9) (Fig. 4b, region 5). This region of the EGF receptor contains Thr<sup>654</sup>, which plays a key role in protein kinase C-mediated receptor modulation (22). A homologous threonine residue is embedded in a basic environment in the HER2 sequence at position 685 (Fig. 4, a and b).

The region of most extensive homology (78.4 percent) between EGF receptor and HER2 (beginning at residue 687) extends over 343 amino acids and includes sequences specifying the adenosine triphosphate (ATP) binding domain (23) and tyrosine kinase activity (Fig. 4b. region 6) (5). This region is also the most conserved between v-erbB and EGF receptor (95 percent) (9). The collinear homology between the EGF receptorerbB and HER2 ceases at position 1032, but introduction of gaps into the EGF receptor or HER2 sequences reveals continued, although decreased, relatedness (Fig. 4b, region 7). This sequence alignment suggests that the two genes evolved by duplication of an ancestral receptor gene, and that subsequent nucleotide sequence divergence in this carboxyl terminal domain led to diverged biological roles for the encoded polypeptides.

The carboxyl terminal domain of HER2 is characterized by an unusually high proline content (18 percent) and predominant hydrophilicity (Fig. 4a). These general features are also found in the EGF receptor carboxyl terminal domain with a 10 percent proline content. The sequences in this region that are found to be conserved are almost exclusively centered around five tyrosine residues, which include the major (Tyr<sup>1173</sup>)

and two minor  $(Tyr^{1148}, Tyr^{1068})$  in vitro autophosphorylation sites in the human EGF receptor (24) (Fig. 4, a and b). Three of these tyrosine residues of HER2 (positions 1139, 1196, 1248) are flanked by homologous sequences PQPEYV, ENPEYL, and ENPEYL (21), respectively (Fig. 4b, region 7).

HER2 chromosomal location. In situ hybridization of two <sup>3</sup>H-labeled HER2 probes (legend, Fig. 5a) to human chromosomes resulted in specific labeling at bands  $ql2 \rightarrow q22$  of chromosome 17 (Fig. 5a). Metaphase cells (100) were analyzed for each probe; 40 percent of cells scored for HER2 probe 1 (HER2-1) had silver grains over  $17q12 \rightarrow q22$  (Fig. 5b). Of the 209 grains observed, 42 (20 percent) were found at this specific region, with no other site labeled above background. For HER2 probe 2, 36 percent of cells had silver grains over the  $q12 \rightarrow q22$ bands of chromosome 17. Of all silver grains, 17 percent (42/246) were localized to this chromosomal region. A secondary site of hybridization with 3.3 percent (8/246) of silver grains was detected at bands p13 $\rightarrow$ q11.2 of chromosome 7.

To test whether this secondary site represented cross-hybridization with the EGF receptor gene, in situ hybridization was carried out with <sup>3</sup>H-labeled EGF



Fig. 2. Northern blot hybridization analysis of normal and malignant human tissues. (a) Fetal tissues; (lane 1) term placenta, (lane 2) 20week placenta, (lane 3) 20-week liver, (lane 4) 20-week kidney, (lane 5) 20-week lung, (lane 6) 20-week brain. (b) Embryonic tumors; (lane 1) hepatoblastoma. (lanes 2 and 3) Ewing sarcoma, (lane 4) rhabdomyosarcoma, (lanes 5 and 6) neuroblastoma, (lane 7) Wilms' tumor. Total poly(A)<sup>+</sup> RNA was isolated as described (33); 4 µg per lane was analyzed on a 1 percent formaldehyde-agarose gel. <sup>32</sup>P-Labeled HER2-1 and HER-2 (legend to Fig. 5) were used as hybridization probes under high stringency conditions [50 percent formamide, 5× Denhardt's solution, 5× standard saline citrate (SSC), sonicated salmon sperm DNA (50  $\mu$ g/ml), 50  $\mu$ M sodium phosphate buffer (pH 6.8), 1 mM sodium pyrophosphate, and 10 µM ATP at 42°C for 16 hours; filters were washed three times for 15 minutes at 45°C with  $0.2 \times$  SSC]. The filters were exposed at -60°C with a Cronex Lightning Plus intensifying screen (Dupont) for 7 days. Rat ribosomal RNA's were used as size standards (28S, 4.8 kb; 18S, 1.8 kb). RNA sizes are given in kilobases.

receptor subclone 64-3. Of 100 cells examined, 30 had silver grains at bands  $p13\rightarrow q11.2$  of chromosome 7 and 3 percent (5/166) of total grains were found over  $q12\rightarrow q22$  of chromosome 17. With the other variant probe (HER2-1) no grain accumulation was observed at the EGF receptor site on chromosome 7.

Southern blot analysis (25) of DNA extracted from nine somatic cell hybrids from human and rodent cells confirmed the localization of HER2 sequences to chromosome 17. <sup>32</sup>P-labeled HER2-1 and HER2-2 probes were hybridized to the same set of Eco RI-digested DNA samples. With HER2-1, a 13-kb hybridizing band was detected in human DNA (Fig. 5c, lane 1) and in DNA samples from hybrids containing human chromosome 17 (Fig. 5c, lanes 6, 8, 10, and 12). Likewise, hybridization of HER2-2 to a 6.6-kb DNA fragment was observed in human control DNA (Fig. 5c, lane 1) and in hybrids containing human chromosome 17 (Fig. 5c, lanes 6, 8, 10, and 12). Chromosome 17 was the only chromosome with perfect concordant segregation; all other chromosomes were excluded by two or more discordant hybrids.

Regional localization to chromosome 17 was also confirmed by Southern blot analysis. In a mouse-human hybrid containing a rearranged human chromosome 17 with region  $17q21 \rightarrow qter$ , the human HER2 restriction fragments were detected (Fig. 5c, lane 4). The HER2 gene was therefore localized to region  $17q21 \rightarrow qter$ , in agreement with the localization made by in situ hybridization.

Even though a low level of hybridization with probe HER2-2 was seen at the site of the EGF receptor gene on chromosome 7, we were able to show that this finding represented cross-hybridization. In a control experiment an EGF receptor probe cross-hybridized to the same extent with the HER2 site on 17q.

Taken together, the results of the in situ and Southern blot hybridizations permit the site of the HER2 sequences to be further narrowed down to bands 17q21-q22, with the major peak of silver grains at band 17q21.

HER2 expression in normal and malignant tissues. To obtain further clues regarding the function of this receptor both in normal cells and in neoplasms, Northern hybridization analyses (15) were carried out with several normal human tissues and randomly collected tumors. A hybridizing 4.8-kb mRNA was detected in all human fetal tissues analyzed, including term placenta, 20-week placenta, liver, kidney, lung, and brain obtained

1	1 AATTCTCGAGCTCGTCGACCGGTCGACGAGCTCGAG	GTCGACGAGCTCGAGGGCGCGCGCCCGGCC	CCCACCCCTCGCAGCACCCCGCGCCCCGCGCC	CTCCCAGCCGGGTCCAGCCGGAGCC	ATGGGGCCGGAGCCGCAGTGAGCACC
		20		40	50
151	MetGluLeuAlaAlaLeuCysArglrpGlyLeuLeul	LeuAlaLeuLeuProProGiyAlaAlaSeri	nrginvallysinrgiyinrAspmetLysLeu/	Argleuproa laserprogiuinre	ACCTEGACATECTCCCCCCCCCCCTCTAC
151		TTOCCUTT TOCCCCCOGAGCCGCGAGCA			100
			60 bolouGlaAcaIleGlaGlaValGlaGlvTvc	/allouIloAlaHicAcnGlnVal/	raGlaValProLeuGlaAraleuAra
301		ACCTACCTGCCCACCAATGCCAGCCTGTCCT	TCCTGCAGGATATCCAGGAGGTGCAGGGCTAC	TGCTCATCGCTCACAACCAAGTGA	GGCAGGTCCCACTGCAGAGGCTGCGG
501		120	130	140	150
	IleValArgGlyThrGlnLeuPheGluAspAspTyr/	AlaleuAlaValleuAspAsnGlvAspProl	euAsnAsnThrThrProValThrGlvAlaSer	ProGlyGlyLeuAraGluLeuGlnL	euArgSerLeuThrGluIleLeuLvs
451	1 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTAT	SCCCTGGCCGTGCTAGACAATGGAGACCCGC	TGAACAATACCACCCCTGTCACAGGGGCCTCC	CAGGAGGCCTGCGGGAGCTGCAGC	TTCGAAGCCTCACAGAGATCTTGAAA
	160	170	180		200
	GlvGlvValLeuIleGlnArgAsnProGlnLeuCvs	TyrG1nAspThrIleLeuTrpLysAspIleP	heHisLysAsnAsnGlnLeuAlaLeuThrLeu	lleAspThrAsnArgSerArgAla	HisPro SerProMet Lys
601	1 GGÁGGGGTCTTGATCCAGCGGAACCCCCAGCTCTGC	TĂCCAGGACACGATTTTGTGGAĂGGACATCT	TCCACAĂGAACAACCAGCTGGCTCTCACACTG/	TAGACACCAACCGČTCTCGĞGCC	CACCCC
	210	220	230	240	250
	GlySerArgCysTrpGlyGluSerSerGluAspCys	G1nSerLeuThrArgThrVa1 <b>Cys</b> A1aG1yG	lyCysAlaArgCysLysGlyProLeuProThr/	Asp <b>CysCys</b> HisGluGln <b>Cys</b> AlaA	laGlyCon ThrGlyProLysHisSer
751	1 GGCTCCCGCTGCTGGGGGGGGGGGGGGGGGTTCTGAGGATTGT	CAGAGCCTGACGCGCACTGTCTGTGCCGGTG	GCTOTGCCCGCTGCAAGGGGGCCACTGCCCACT	GACTECTECCATGAGCAGTETGCTG	CCGGC
	260	270	280	290	300
	AspCysLeuAlaCysLeuHisPheAsnHisSerGly	Ile <b>Cys</b> GluLeuHis <b>Cys</b> ProAlaLeuVall	hrTyrAsnThrAspThrPheGluSerMetPro/	AsnProGluGlyArglyrlhrPhe	ilyAlaSer VallhrAla
901	I GALINGLIGGULINGLIULALIILALLALAGIGGU	ATC TO GAGE TOLAL TOLLAGUE TOGTCA	LUTALAALALAGALALGIIIIGAGILLAIGULLA	ATCCCGAGGGCCGGTATACATTCG	
	310 Tum Agen Tum Lou Son The Agen Val Clusses	320 Sulla Sur Drot sullis Asr Cla Cluvs IT	33U		
1051		TTCGTCTSCCCCCTGCACAACCAAGAGGTGA	CAGCAGAGGATGGAACACAGCGGTCTGAGAAG	AGCAAGCCCTOTECCCGAGTG	TATGGTCTGGGCATGGAGCACTTG
	360	370	380	300	400
	AraGluValAraAlaValThrSerAlaAsnIleGlu		eri euAlaPhei euProGluSerPheAsoGlw	AspProAlaSerAsnThrAlaProL	euGlnProGluGlnLeuGlnValPhe
1201	1 CGAGAGGTGAGGGCAGTTACCAGTGCCAATATCCAG	GAGTTTGCTGGCTGCAAGAAGATCTTTGGGA	GCCTGGCATTTCTGCCGGAGAGCTTTGATGGG	ACCCAGCCTCCAACACTGCCCCGC	TCCAGCCAGAGCAGCTCCAAGTGTTT
	410	420	430	440	450
	GluThrLeuGluGluIleThrGlyTyrLeuTyrIle	SerAlaTrpProAspSerLeuProAspLeuS	erValPheGlnAsnLeuGlnValIleArgGly/	\rgIleLeuHisAsnGlyAlaT <b>yr</b> S	erLeuThrLeuGlnGlyLeuGlyIle
1351	1 GAGACTCTGGAAGAGATCACAGGTTACCTATACATC	TCAGCATGGCCGGACAGCCTGCCTGACCTCA	GCGTCTTCCAGAACCTGCAAGTAATCCGĞGGĂ	CGĂATTCTGCACAATGGCGCCTĂCT	CGCTGACCCTGCAAGGGCTGGGCATC
	460	470	480	490	500
	SerTrpLeuGlyLeuArgSerLeuArgGluLeuGlyS	SerG1yLeuA1aLeuI1eHisHisAsnThrH	isLeu <b>Cys</b> PheValHisThrValProTrpAsp	GlnLeuPheArgAsnProHisGlnA	laLeuLeuHisThrAlaAsnArgPro
1501	1 AGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGGGC/	AGTGGACTGGCCCTCATCCACCATAACACCC	ACCTCTCTTCGTGCACACGGTGCCCTGGGAC	CAGCTCTTTCGGAACCCGCACCAAG	CTCTGCTCCACACTGCCAACCGGCCA
	510	520	530	540	550
1.001	GluAspGluCysValGlyGluGlyLeuAlaCysHis	GInLeuryAlaArgArgAlaLeuLeuGlyS	erGlyProThrGlnCysValAsnCysSerGlnl	heLeuArgGIyGInGIuCysValo	iluGluce ArgValLeuGlnGlyLeu
1651	I GAGGACGAGINIGIGGGCGAGGGCCIGGCLIGCLAC		LAGGGLLLALLLAG TO ILAAL TO AGLLAG		AGGAA
		5/U	580 Ser Val The PhoClu ProCluve 1 a Ace Clu	590 Valla Carola Hic Turola ve A	600
1801		STATSPFOGIOL SSGIPFOGINASHGISS	CAGTGACCTATTTGGACCGGAGGCTGACCAG		
1001		620	630	640	650
	ProSerGlvVall vsProAspl euSerTvrMetPro	lleTrnivsPheProAsnGluGluGlvAla	Glapro Pro LleAsa ThruisSer		ProAlaGluGlnArgAlaSerPro
1951	1 CCCAGCGGTGTGAAACCTGACCTCTCCTACATGCCC/	ATCTGGAAGTTTCCAGATGAGGAGGGCGCA	ECCAGCCT ECCCCATCAACTECACCCACTCC	<b>G</b> TGGACCTGGATGACAAGGGC	CCCGCCGAGCAGAGAGCCAGCCCT
				anaro 8	907008
	660	670	680	🕈 690	700
	660 LeuThrSer[I]eVa]SerA]aVa]G]yI]eLeul	670 Leu Val Val Val LeuGly Val Val PheGly I	680 leLeuIleLysArgArgGlnGlnLysIleArgl	.ysTyrThrMetArgArgLeuLeuG	700 InGluThrGluLeuValGluProLeu
2101	660 LeuThrSer[]eVa]SerA]aVa]Va]G]yI]eLeul 1 CTGACGTCCATCGTCTGCGGTGGTTGGCATTCTG	670 Leu Val Val Val LeuGly Val Val PheGly I CTGGTCGTGGTCTTGGGGGGTGGTCTTTGGGA	680 leLeuIleLysArgArgGlnGlnLysIleArgI TCCTCATCAAGCGACGGCAGCAGAAGATCCGG/	.ysTyrThrMetArgArgLeuLeuG AGTACACGATGCGGAGACTGCTGC	700 InGluThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG
2101	660 LeuThrSen[I]eVa]SerA]aVa]Va]G]yI]eLeul 1 CTGACGTCCATCCTCTCCGCGGTGGTTGGCATTCTG 710	670 Leu Val Val Val Leu Gly Val Val PheGly I CTGGTCGTGGTCTTGGGGGGTGGTCTTTGGGA 720	680 leLeuIleLysArgArgGInGInLysIleArg TCCTCATCAAGCGACGGCAGCAGAAGATCCGG 730	690 ysTyrThrMetArgArgLeuLeuG AGTACACGATGCGGAGACTGCTGC 740	700 InGluThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750
2101	LeuThrSer[I]eValSerAlaValValGlyI]eLeuf 1 CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG 710 ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/	LeuValValValVeuGlyValValPheGlyI CTGGTCGTGGTCTTGGGGGGTGGTCTTTGGGA 720 ArgIleLeuLysGluThrGluLeuArgLys	680 1eLeuIleLysArgArgGlnGlnLysIleArgl <u>TCCTCATG</u> AAGCGACGGCAGCAGAAGATCCGG 730 alLysValLeuGlySerClyAlaPheGlyThr TCATGATAGA	690 .ysTyrThrMetArgArgLeuLeuG AAGTACACGATGCGGAGACTGCTGC 740 /alTyrLysGlyIleTrpIleProA	700 SINGluThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 ISPGluAsnValLysIleProVal
2101 2251	660 LeuThrSer[IeValSerAlaValValGlyIleLeuf 1 CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG 710 ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/ 1 ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGTGG	670 Leu Val Val Val Leu Gly Val Val Phe Gly CTGGTCGTGGTCTTGGGGGGTGGTCTTGGGA 720 Arg I leLeu Lys Glu Thr Glu Leu Arg Lys V CGGATCCTGAAAGAGACGGAGCTGAGGAAGG	680 TeLeuIleLysArgArgG1nG1nLysIleArgi TCCTCATCAAGCGACGGCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGATCTGGCGCTTTGGCACA	♦ 690 LysTyrThrMetArgArgLeuLeuG AAGTACACGATGCGGGAGACTGCTGC 740 /alTyrLysG1yI1eTrpI1eProA STCTACAAGGGCATCTGGATCCCTG	700 InGluThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 IspGlyGluAsnValLysIleProVal ATGGGGAGAATGTGAAAATTCCAGTG
2101 2251	LeuThrSer[I TeValSerA1aVa1VaIG1y1 TeLeu CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGC 710 ThrProSerG1yA1aMetProAsnG1nA1aG1nMet/ ACACCTAGCGGAGCGCATGCCCCAACCAGCGCGCAGTGC ACACCTAGCGGAGCGCATGCCCCAACCAGCGCGCAGTGC	670 Eeu Val Val Val LeuG 1 yVal Val PheG 1 yI CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I leLeuLys G1uThrG1uLeuArgLys V CGGATCCTGAAAGAGACGGAGCTGAGGAAGG 770	680 1eLeuileLysArgArg6in6inLysIleArgi TCCTCATCAAGCGACGGCAGCAGAAGATCCGG 730 alLysValLeuGlySerGlyAlaPheGlyThri TGAAGGTGCTTGGATCTGGCCTTTTGGCACAI 780 000000000000000000000000000000000		700 11n61uThr61uLeuVa161uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 150 15061y61uAsnVa1Lys11eProVa1 ATGGGGAGAATGGAAAATTCCAGTG 800 2007
2101 2251 2401	LeuThrSer[I] EValSerAlaValValValGly1 I ELEU CTGACGTCCATCGTCTCTCCGGTGGTTGGCATTCTG ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/ ACACCTAGCGGAGGGGATGGCCCAACCAGGCGCAGGGGGAGG AlalleLysValLeuArgGluAsnThrSerProLysVa CCCATCGAAGTCTTGGCGAAAGCATCCCCCAAC	670 Leu Val Val Val Peo Filonov CTGGTCGTGGTCTTGGGGGGGGGTCTTTGGGA 720 Arg I leLeuLysGluThrGluLeuArgLysV CGGATCCTGAAAGAGACGGAGCTGAGGAAGG 770 70 AlaAsnLysGluI leLeuAspGluAl aTyrV	680 Icleuile ysArgArgG1nG1nLysIleArg ICCTCATCMAGCGACGSCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yAlaPheG1yThr IGAAGGGCTIGGACTCGGCGCTTTTGGCACA 780 1MetA1aG1yVa1G1ySerProTyrVa1Ser Teatacercecterceccroccatactrocc		700 SINGIUThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG SpGlyGluAsnValLysIleProVal ATGGGGAGAATGTGAAAATTCCAGTG 800 SerThrValGInLeuValThrGlnLeu CraccGCFGAGFGGTGAGFAGFGT
2101 2251 2401	660           LeuThrSer[I]eValSerAlaValValGlyI]eLeuI           CTGACGTCCATCGTCTCTCCGGGTGGTTGGCATTCTGG           ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/           ACACCTAGCGGGGGGGGAGGGCGATGCCCAACCAGGGCGCAGATGC           AlaIleLysValLeuArgGluAsnThrSerProLys/           GCCATCAAAGTGTTGGGGGAAAACACATCCCCCAAAC	670 LeuValValValLeuGlyValValPheGlyI CTGGTCGTGGTCTTTGGGGGTGGTCTTTGGGA ArgIleLeuLysGluThrGluLeuArgLysV CGGATCCTGAAAGAGGGGGGGGGGGGGGGGG 770 AlaAsnLysGluIleLeuAspGluAlaTyrV SCCAACAAAGAAATCTTAGACGAAGCATACG	680 IeLeuIIeLysArgArgG1nG1nLysIleArg TCTCATC AGCGACGGCACCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGATCTGGCGCT 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser/ TGATGGCTGGTGTGGGGCTCCCCATATGTCTCC	690 _ysTyrThrMetArgArgLeuLeuG AGTACACGATGCGGAGACTGCTGC 740 740 740 740 740 740 740 740	700 SINGIUThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG IspGlyGluAsnValLysIleProVal ATGGGGGAGAATGTGAAATTCCAGTG 800 SerThrValGInLeuValThrGInLeu CCACGGTGCAGCTGGTGACACAGCTT 850
2101 2251 2401	660           LeuThrSer[I]eValSerAlaValValGlyI]eLeuI           CTGACGTCCATCGTCTCCGCGTGGTTGGCATTCTGG           710           ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/           1 ACACCTAGCGGAGGGATGCCCAACCAGCGGCGAATGC           AlaIleLysValLeuArgGluAsnThrSerProLys/           GCCATCAAAGTGTTGAGGGAAAACCATCCCCCAAAC           Blo           MetProTyrGly	670 Eeu Val Val Val Leu Gi JV al Val Phe Gi JI CTGGTCGTGGTGTTTGGGGGGGGGTGTTTTGGA Arg I le Leu Lys Glu Thr Glu Leu Arg Lys V CGGATCCTGAAAGAGAGGGGGGGGGGGGGGGGGG 770 AlaAsn Lys Glu I le Leu Asp Glu Al a Tyr V GCCAACAAAGAAATCTTAGACGAAGCATACG Asparg Gl Vargl eu Gly Ser Gla Aspleul	680 1eLeuIIeLysArgArgG1nG1nLysIleArg TCCTCATC AAGCGACGGCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGGGCTTGGGCTCTTGGCACAL 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser/ TGATGGCTGGTGTGGGCTCCCCCATATGTCTCC 830 euAsnTrpCusMetG1u1leA1alysG1yMet		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCCCTG AGGGAGAACGGAGCTGGTGGAGCCCCTG 350G1yG1uAsnVa1Lys11eProVa1 34TGGGGGGAGAATGTGAAATTCCAGGTG 800 3erThrVa1G1nLeuVa1ThrG1nLeu (CCACGGTGCAGCTGGTGACACAGCTT 350 31HisArgAspLeuA1aA1aArgAsn
2101 2251 2401 2551	660           LeuThrSer[ITeValSerAlaValVaGIy11eLeu]           CTGACGTCCATCGTCTCGCGGTGGTTGGCATTCTGG           710           ThrProSerGIyAlaMetProAsnGInAlaGInMet/           ACACCTAGCGGATGCCCAACCAGGCGCCGATGG           ALaILeLysValLeuArgGluÄsnThrSerProLys/           GCCATCCAACGATGCCCAACACACCCCCAAA           810           MetProTyrGIySeLeuLeuAspHisValArgGlu/	200 2014 Jal Val LeuGiyVal Val PheGiyi CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I leLeuLysGluThrGluLeuArgLysV CGGATCCTGAAAGAGACGGAGCTGAGGAAGG 770 AlaAsnLysGluI leLeuAspGluAlaTyrV SCCAACAAAGAAATCTTAGACGAAGCATACG AsnArgGlyArgLeuGlySerGlnAspLeuL AccCGCGACGACCTGGGCTCCCAGGACCTGC	680 Telewile_ysArgArg6In6inLysIleArg TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 1LysValLeuGlySerGlyAlaPheGlyThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 780 alMetAlaGlyValGlySerProTyrValSer/ TGATGGCTGGGTGTGGGCTCCCCATATGTCTCC 830 euAsnTrpCysMetGlnIleAlaLysGlyMet2i TGAACTGGTGTATGCAAGGGGATG	♦ 690 _ysTyrThrMetArgArgLeuLeuG AGTACACGATGCGGAGACTGCTGC 740 /aTTyrLysGlyIleTrpIleProA STCTACAAGGGCATCTGGATCCCTG 790 ArgLeuLeuGlyIleTrtLeuThrS SGCCTTCTGGGCATCTGCCTGACAT 840 SerTyrLeuGluAspValArgLeuV AGCTACCTGGAGGATGTGCGGCTCC	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGAGAATGTGAAAATTCCAGTG 800 300 301 StargAspLeuA1aATAGAA 760 160 160 160 160 160 160 160 1
2101 2251 2401 2551	660         LeuThrSer[ITeValSerAlaValVaGIyI]teLeu[         CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCCGCAGTGG         AlaIleLysValLeuArgGluÄsnThrSerProLys/         GCACTCAAAGTGTTGAGGGAAAACACATCCCCCAAC         810         MetProTyrGIyTsLeuLeuAspHisValArgGlu/         ATGCCCTATGGCTCCCTTTAGACCAGTCCGGGAU         860	670 Eeu Val Val Val LeuGi yVal Val PheGi yi CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I leLeuLysGlu Thr GluLeuArgL ys V CGGATCCTGAAAGAGACGGAGGGACGGACGG 770 AlaAsnLysGlu I leLeuAspGluAl aTyr V GCCAACAAAGAAATCTTAGACGAAGGATACG AsnArgGlyArgLeuGlySerGlnAspLeuL AACCGCGGGACGCCTGGGCTCCCAGGACCTGG 870	680 1eLeuile_ysArgArgG1nG1nLysIleArgi TCCTCATCAAGCGACGACGACAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGATCTGGCCGTTTTGGCACA 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser TGATGGCTGGTGGGGCTCCCCATATGTCTCC 830 euAsnTrpCysMetG1nI1eA1aLysG1yMet: TGAACTGGGTATGCGATTGCCAAGGGGATG 880 △		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGGACACAGCTT 850 31H isArgAspLeuA1aA1aArgAsn TACACAGGGACTTGGCCGCTCGGACA 9000
2101 2251 2401 2551	660         LeuThrSer[ITeValSerAlaValValGly1leLeu]         CTGACGTCCTCTCCGGCTGGCTTGGCATTCTGG         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         ACACCTAGCGGAGGGGAGGGCGATGGCCGAACGAGGGGAGGGA	670 670 670 672 672 672 672 672 672 672 672	680 Iceuile ysArgArgG1nG1nLysIleArg ICCTCATCMAGCGACGGCAGCAGAAGATCCGG 730 al Lys ValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGACCTGGCGCTTTTGGCACA 780 al MetA1aG1yVa1G1ySerProTyrVa1Ser TGATGGCTGGTGGTGGGGCCCCCCATATGTCTCCC 830 euAsnTrpCysMetG1nIleA1aLysG1yMet TGAACGGGTATGCAAGGGATG 880 spG1uThrG1uTyrHisA1aAspG1yG1yLys'		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 150 150G1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGGACACAGCTT 850 141HisArgAspLeuA1aA1aArgAsn 17ACACAGGGACTTGGCCGCTCGGAAC 900 31uSerI1eLeuArgArgArgPheThr
2101 2251 2401 2551 2701	660         LeuThrSer[I]eValSerAlaValValValGlyI]eLeu[         1 CTGACGTCCTCTCCGGTGGTGGCATTCTGG         710         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         ACACCTAGCGGAGGGGAGGGCGATGGCCCAACCAGGGGCGAGGGG         ACACCTAGCGGGGAGAGGCCCAACCAGCGGCGACAGGGC         ACACCTAGCGGGAAAAACACATCCCCCAAA         AlaIleLysValLeuArgGluÁsnThrSerProLys/         GCATCAAAGTGTTGAGGGAAAAACACATCCCCCCAAAA         Blo         MetProTyrGlySsLeuLeuAspHisValArgGlu/         ATGCCCTCTTAGACCATGTCCGGGAAA         860         ValLeuValLysSerProAsnHisValLysIIeThr/         GGCTGCCAAGAGTCCCAACCATGTCAAAATTACAA	670 670 670 672 672 672 672 672 672 672 770 770 770 770 770 770 770 7	680 leLeuIIeLysArgArgG1nG1nLysIleArg TCCTCATC TCCTCATC AAGCGACGGCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGGATCTGGCGCTTTTGGCACAU 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser TGAAGGCTGGTGGTGGGGCTCCCCATATGCTCCC 830 euAsnTrpCysMetG1nIleA1aLysG1yMet: TGAACTGGTGTATGCAGATTGCCAAGGGGATG 880 △ spG1uThrG1uTyrHisA1aspG1yG1jLys' ACGAGACAGAGTACCATGCAGATGGGGGCAAGG	690 ysTyrThrMetArgArgLeuLeuC AGTACACGATGCGGAGACTGCTGC 740 740 740 740 740 740 740 740	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 550 15pG1yG1uAsnVa1LyS11eProVa1 1ATGGGGAGAATGTGAAATTCCAGTG 800 30 FThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCACCTGGTGACACAGCTT 850 7a1HisArgAspLeuA1aA1aArgAsn TACACAGGGACTTGGCCGCTCGGAAC 900 31USerI1eLeuArgArgArgPheThr AGTCCATTCTCCGCCGGGGGTTCACC
2101 2251 2401 2551 2701	660         LeuThrSer[ITeValSerAlaValVaGIy11eLeu]         CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyA1aMetProAsnG1nA1aG1nMet/         ACACCTAGCGGATGCCCAACCAGGCGCAGATGG         ALaIIeLysValLeuArgG1uÅsnThrSerProLys/         GCCATCCAACGGTGCCCCAACAGTCCCCCAAC         810         MetProTyrG1yCsLeuLeuAspHisValArgG1u/         ATGCCCTATGGCCCCCCCCAAGTGCCCCCAAGTGCCCCCAAGTGCCCCCAAGTGTTGCGGGAGACACTCCCCCCAAC         MetProTyrG1yCsLeuLeuAspHisValArgG1u/         ATGCCCTATGGCCCCCCCCAACTGCCCGGGAA         860         ValLeuValLysSerProAsnHisValLys11eThr/         GGCGTCCCAAGGGTCCCCAACCAGTGCCCAAATTACAC         910	20 20 20 20 20 20 20 20 20 20	1eLeuile_ysArgArg6in6inLys1leArgi         1eLeuile_ysArgArg6in6inLys1leArgi         730         1LysValLeuG1ySerG1yA1aPheG1yThri         1EAAGGTGCTTEGATCTEGCGCTTTTGGCACAI         780         alMetA1aG1yVa1G1ySerProTyrVa1Ser/         TGATGGCTGGGTGTGGGCTCCCCATATGTCTCCG         830         euAsnTrpCysMetG1n1leA1aLysG1yMet2         TGAACGGCGTATGCACATGCCGGGGTGGGCACGCAGATGCGGGGCAGGGGTGC         spG1uThrG1uTyrHisA1aAspG1yG1yG1yLys'         ACGGACACAGGTGCCACATGCAGATGGGGGCAGG         930       20	690 ys Jyr Thr Met ArgArgLeuLeuG AGTACACGATGCGGAGACTGCTGC 740 /a 1 Tyr Lys Gly 1 le Trp 1 le ProA STCTACAAGGGCATCTGGATCCCTG 790 ArgLeuLeuG 1 y 1 le Tre Leu Thr 5 Sec Tyr LeuG 1 uAspVa 1 ArgLeuV Scatt CTGGGCATCT 840 Ser Tyr LeuG 1 uAspVa 1 ArgLeuV Scatt CTGGGCAGGATGGCGCTCG 890 /a 1 Pro 1 le Lys Trp Met A 1 a LeuG 1 GCCCATCAAGTGGATGGCGCTGG 940	700 influThr6luLeuVal6luProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 isp6ly6luAsnValLys1leProVal ATGGGAGAATGTGAAAATTCCAGTG 800 ierThrVal6lnLeuValThr6lnLeu CCACGGTGCAGCTGGTGACACAGCTT 1AHisArgAspLeuAlaAlaArgAsn TACACAGGGACTTGGCCGCTCGGAC 900 iluSerIleLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCACC 950
2101 2251 2401 2551 2701	660         LeuThrSer[ITeValSerAlaValVaGIyI1eLeu]         CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCCGCAGTGG         AlaIleLysValLeuArgGluÄsnThrSerProLys/         GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAC         MetProTyrGlyCsLeuLeuAspHisValArgGlu/         AGCCCTATGGCTCCTCTTAGACCATGCCCGGAA/         ValLeuValLysSerProAsnHisValLysIleThr/         GGCTGGTCAAGGGTCCCCACACCATGCCAAACAAATTACAC         910         HisGInSerAspValTrpSerTyrGlyValThrVal	670 670 670 670 670 670 670 670	680 1eLeuI1eLysArgArgG1nGInLysI1eArg TCCTCATCHAGCGACGCACCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGATCTGGCGCTTTTGGCACA 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser TGATGGCTGGTGGGGCCCCCCATATGTCCCC 830 euAsnTrpCysMetG1nI1eA1aLysG1yMet: TGAACTGGGTATGCAGATTGCCAAGGGGATG 880 △ spG1uThrG1uTyrHisA1aAspG1yG1yLys ACGAGACAGAGTACCATGCAGATGGGGGCAAG 930 △ yrAspG1yI1eProA1aArgG1u11eProAspl		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGAGCCGCTG 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGGACACAGCTT 850 301 SerTheLeuA1aA1aArgAsn 17ACACAGGGACTTGGCCGCTGGACA 900 10 Ser11eLeuArgArgArgPheThr 10 Ser11eLeuArgArgArgPheThr 10 SerCATTCCCGCCGCGGGTTCACC 950 10 Ser11eLeuArgArgArgPheThr 10 Ser11eL
2101 2251 2401 2551 2701 2851	660         LeuThrSer[ITeValSerAlaValValGiy1ieLeu]         CTGACGTCCATCGTCTCTGCGGTGGCTTGGCATTCTGC         ThrProSerGiyAlaMetProAsnGinAlaGinMet/         ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGCGAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGAGGC         ALILLySValLeuArgGlu/AsnThrSerProLys/         1 GCCATCAAAGTGTTGAGGGAAAAACACATCCCCCAAAC         810         MetProTyrGlySteLeuLeuAspHisValArgGlu/         ATGCCTATGGCTCCCCTATGCCCCACCAGCGGCAAGAGACACACATCCCCGGAA/         ATGCCCTATGGCTCCCTCTAGACCATGTCCGGAA/         MetProTyrGlySteLeuLeuAspHisValArgGlu/         ATGCCCTATGGCTCCCCAACCATGTCCAGAAATTACAG         860         ValLeuValLySSerProAsnHisValLysIleThr/         GTGCTGGTCAAGAGTCCCAACCATGTCAAAATTACAG         910         HisGlnSerAspValTrpSerTyrGlyValThrValT         CACCAGAGTGATGTGTGGGAGGTTATGGTGGACTGTG	670 670 670 672 672 672 672 672 672 672 672	680 ILeLeuIIL ysArgArgGInGInLysIleArgi ICCTCATCHAGCGACGKCACAGAAGATCCGG 730 alLysValLeuGlySerGIyAlaPheGlyThr TGAAGGTGCTTGGACCTGGCGCTTTTGGCACA 780 alMetAlaGlyValGlySerProTyrValSer TGATGGCTGGTGGGGGCCCCCCATATGTCTCCC 830 euAsnTrpCosMetGInIleAlaLysGlyMet TGAACGGGTATGCAGATTGCCAGGGGCAAG 880 spGluThrGluTyrHisAlaAspGlyGlyLys ACGAGACAGAGTACCATGCAGATGGGGGCAAG 930 yrAspGlyIleProAlaArgGluIleProAspl	690 ysTyrThrMetArgArgLeuLeuG vaGTACACGATGCGGGAGACTGCTGC 740 740 740 740 740 740 740 740	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 350 350G1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300FTThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGGCACACAGCTT 850 31HisArgAspLeuA1aA1aArgAsn 31CACAGGGACTTGGCCGCCGCGGAAC 900 31USer11eLeuArgArgArgPheThr AGTCCATTCTCCGCCGCGGGTCACC 950 70501ProPro11e Thr11eAsp CCCAGCCCCCCATC ACCATTGAT
2101 2251 2401 2551 2701 2851	660         LeuThrSer[ITeValSerAlava]ValGly1leLeu[         1       CTGACGTCCATCGTCTGCGGTGGCATTCTGG         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         1       ACACCTAGCGGAGGGAGGATGCCCAACCAGGCGCAGATGG         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGG         7       ACACCTAGCGGAGGACGATGCCCCCAAC         810       B10         MetProTyrGlyCsLeuLeuAspHisValArgGlu/         1       ATGCCCTATGGCTCCTTTAGACCATGTCCCGGAA/         860       ValLeuValLySEPPOASnHisValLySIleThr/         1       GTGCTGGTCAAGAGTCCCAACCATGTCAAAATTACAG         910       HisGlnSerAspValTrpSerTyrGlyValThrValT         1       CACCAGAGTGATGTGGGAGTTGTGGGAGTTTGGGGTGACTGTG	20 20 20 20 20 20 20 20 20 20	1eLeuile_ysArgArg6in6inLys lleArgi         1ccrCATCAAGCGACGCAGCAGAAGATCCGG         730         1Lys ValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGATCTGGCGCTTTTGGACAU         780         alMetAlaGlyValGlySerProTyrValSer/         TGAAGGTGCTCGGACTCCCCATATGTCTCCC         830         euAsnTrpCysMetGInIleAlaLySGlyMet;         TGAACTGGTGTGGACTCCCCATATGTCTCC         830         euAsnTrpCysMetGInIleAlaLySGlyMet;         TGAACTGGTGTGTATGCAAGATTGCCAAGGGATG         880         930         yrAspGlyIPeroAlaArgGluIleProAlaCGGCAAG         930         yrAspGlyIPeroAlaArgGluIleProAlaCGGAAGACCCCGGGAAGACCCGGGAAGACCCGGGACGCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGACGCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGACGCAGACGCCCGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGAAGACCCCGGGACGCAGACGCCCGGAGACCCCGGAGACCCCGGGACGCAGACCCCCGGGACGCAGACCCCGGGAAGACCCCCGGGACGCAGACGAC	690 ysTyrThrMetArgArgLeuLeuG AGTACACGATGCGGGAGACTGCTGG 740 740 740 740 740 740 740 740	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 350 350G1yG1uAsnVa1Lys11eProVa1 34TGGGGAGAATGTGAAAATTCCAGTG 800 30 erThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGACACAGCTT 850 31 uSer11eLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC 900 31 uSer11eLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC 950 750 750 750 750 750 750 750 7
<ul> <li>2101</li> <li>2251</li> <li>2401</li> <li>2551</li> <li>2701</li> <li>2851</li> <li>2001</li> </ul>	660         LeuThrSer[ITeValSerAlaVal/ValGIy11eLeu]         1 cTGACGTCCGATCGTCTCGCGGTGGCATTCTGG         710         ThrProSerGIyA1aMetProAsnGInAlaGInMet/         ACACCTAGCGAGCGCAGATGG         ACACCAGCGAGCGCAGATGG         ACACCAGCGAGCGCAGATGG         ACACCAGCGAGCGCAGATGG         ACACCAGCGAGCGCAGATGG         ACACCAGCGAGCCCCAGATGGCCCAGAGCG         ACACCAGCGGAGAGCCCCAACCAGCCCCAGAGCGCCCCAACA         B10         MaisGInSerAspValTrpSerTyrGlyValTrpVal'         GGCGCGCCCCACCAGGGAGTATGGGTGGGAGCTATGGCGGAGTATGGGTGGTGGCGCGCTGCCAACCAGGTGGTGGGAGTATGGTGGGGGGCCGCCACCATGCCGACCAGGCGCGCCGCCACCAGGTGATGGGTGGG	20 20 20 20 20 20 20 20 20 20	680 1eLeuile_ysArgArg6in6inLys IIeArgi TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 80 alMetA1aG1yVa1G1ySerProTyrVa1Ser/ TGATGGCTGGTGGGGCCCCCATATGTCTCCC 830 euAsnTrpCysMetG1nI1eA1aLysG1yMet: TGAACTGGTTATGCAGATTGCCAGGGATG 880 5pG1uThrG1uTyrHisA1aAspG1yG1yLys' ACGAGACAGGTACCCAGCAGTGGGGGCAAG 930 yrAspG1yI1eProA1aArgG1u1PeroAsp1 ACGAGGACCCCGGGCAGGCCCCCGGAGATCCCTGAG 980 alSerG1uPheSerArgMetA1aArgAspPro1 980		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 15pG1yG1uAsnVa1Lys1leProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 60 60 60 60 60 60 60 60 60 60
2101 2251 2401 2551 2701 2851 3001	660         LeuThrSer[ITeValSerAlaValVaGIy1]teLeu[         1       CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCCGCAGTGG         AlaIleLysValLeuArgGluÄsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAC         MetProTyrGlyCsLeuLeuAspHisValArgGluÄ         1       ATGGCTGTCAGGGGTCCCCACACCATGCCCGGGAU         810         MetProTyrGlyCsLeuLeuAspHisValArgGluÄ         1       ATGGCTGTCAGGGTCCCACATGCACATGCCGGGAU         860         ValLeuValLySSerProAsnHisValLySIEThrA         1       GTGGTGGTCAAGAGTCCCAACATGCAAATTACAG         910       HisGInSerAspValTrpSerTyrGlyValThrValT         1       CACCAGAGTGATGTGTGGAGATTATGGTEGAACTGGC         1       GTGTACATGATCATGGTCAAAGTTATGGTTGAATTACGT         1       GTCTACATGATCATGGTCAAATGTTGGATGATTGGTTGAGATTGGT         2       ValTyrMetIleMetValLySCYSTrpMetIleAsp1         1       GTCTACATGATCATGGTCAAATGTTGGATGATTGGTTGAAATGTGTGAATTGACT	670 670 670 670 670 670 670 670	680         Ileuile, ysArgArgGinGinLys IleArgi         TCCTCATCAAGCGACGCACCAGAAGATCCGG         730         allysValLeuGlySerGiyAlaPheGlyThr         TGAAGGTGCTTGGACTCGCCCTTTGGCACAG         almetAlaGlyValGlySerCyTranscreating         almetAlaGlyValGlySerProTyrValSer         BAGGTGGGTGGGCCCCCCATATGCTCCG         830         euAsnTrpCySMetGInIleAlaySGlyMet:         TGAAGGGCAGGATATGCCAGATTGCCCAGGGGGATG         880         SpGluThrGluTyrHisAlaAspGlyGlyLys'         ACGAGACAGAGTACCCAGCCGGAGATGCGGGGGCAGG         930         yrAspGlyIlPeroAlaArgGluIlPeroAlaArgGluIPeroAspl         AGATGGGATCCCCGCCAGCCCGGGGGACCCCCCGCAGGGCCCCC         alserGluPheSerArgMetAlaArgAspPro         TGTCGGATTCCCCGCATGCCCCAGGGGCCCCC		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGAGCCGCTG 15pG1yG1uAsnVa1Lys11eProVa1 3ATGGGGAGAATGTGAAAATTCCAGTG 800 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGGACACAGCTT 850 301 SerTleLeuArgArgArgPheThr 16ATCCATTCCCGCCGCGGGTCACCA 900 10 SerTleLeuArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgArgArgArgPheThr 10 SerTleLeuArgArgArgArgArgArgArgArgArgArgArgArgArgA
2101 2251 2401 2551 2701 2851 3001	660         LeuThrSer[ITeValSerAlaValValGiy1ieLeu[         1 CTGACGTCCATCGTCTCGCGGTGGGTTGGCATTCTGG         710         ThrProSerGiyAlaMetProAsnGinAlaGinMet/         ACACCTAGCGGAGCGATGCCCAACCAGCGGCGAGGG         ALaIleLysValLeuArgGiuAsnThrSerProLys/         1 GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAAC         810         MetProTyrGiyCsLeuLeuAspHisValArgGIu/         1 ATGCCTATGGCTGCCCCCTCTTAGACCATGCCCGGAA/         810         MetProTyrGiyCsLeuLeuAspHisValArgGIu/         1 ATGCCTATGGCTGCCCCCTATGCCCCAACATGCCGGAA/         860         ValLeuValLysSerProAsnHisValLysIleThr/         1 GTGCTGGTCAAGAGTCCCAACCATGTCAAAATTACA(         910         HisGInSerAspValTrpSerTyrGlyValThrVal'         1 CACCAGAGTGATGTGTGGGAGTATGGTGGACTGTG         2         ValTyrMet11eMetValLysCysTrpMetT1eAspC         1 GTCTACATGATCATGGTCAAAGTTGGACTGGATTGACT         1010         4snSerthrpBeTyrArgCorter	670 670 670 670 670 670 670 670	680 1eLeuIIL ysArgArgG1nG1nLysI1eArgI TCCTCATCHAGCGACGSCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGACTTGGCGCTTTTGGCACA 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser TGATGGCTGGTGGGGCTCCCCATATGTCTCC 830 euAsnTrpCySMetG1nI1eA1aLysG1yMet: TGAACGGGGTATGCAGATTGCCAAGGGGATG 880 spG1uThrG1uTyrHisA1aAspG1yG1yLys ACGAGACAGGGATCCCAGACTGGCGGAAGTCCCTGACT 930 yrAspG1yI1eProA1aArgG1uI1eProAspI ACGATGGGGATCCCCGGCGAGACTCCCTGACT 930 alSerG1uPheSerArgMetA1aArgAspPro1 TGTCGAATTCCCCGCATGGCCAGGGGACCCCC 1030 yrLauVa1ProG1nc1nc20	690 ysTyrThrMetArgArgLeuLeuG VaTyrLysGlyIleTrpIleProA TCTACAAGGGCATCTGGACTGCTGC ArgLeuLeuGlyIleTrpIleProA SGCCTTCTGGGCATCTGGACATCCCTG CGCCTTCTGGGCATCTGCTGACAT 840 SerTyrLeuGluAspValArgLeuH AGCTACCTGGAGGATGGCCGGCCTGC 890 ValProIleLysTrpMetAlaLeuG STGCCCATCAAGTGGATGGCCGGCTGC 900 SIAArgCheValValIleGInAshG CAGCGCTTTGTGGTCATCCAGAATG 1040 SepTroAlaProCleVal261VG14V	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 150 150G1yG1uAsnVa1Lys1leProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGGCACAGCCTT 850 301 SargAspLeuA1aA1aArgAsn 17ACACAGGGACTTGGCGCTCGGAAC 900 301 Ser1leLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTTCACC 900 302 SargArgArgArgArgheThr AGTCCATTCTCCGCCGGCGGTTCACC 900 303 SargArgArgArgArgArgArgArgArgArgArgArgArgAr
<ul> <li>2101</li> <li>2251</li> <li>2401</li> <li>2551</li> <li>2701</li> <li>2851</li> <li>3001</li> <li>3151</li> </ul>	660         LeuThrSer[ITeValSerAlava]ValGly1leLeu[         1       CTGACGTCCATCGTCTGCGGTGGCATTCGC         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         1       ACACCTAGCGGATGCCCAACCAGGCGCAGATG         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         1       ACACCTAGCGGATGCCCAACCAGCGCGAGTG         6       760         AlaIleLysValLeuArgGluÅssnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA         810       810         MetProTyrGlyCysLeuLeuAspHisValArgGlu/         ATGCCTATGGCCTCCTTTAGACCATGTCCAGGAAA         860         ValLeuValLySSerProAsnHisValLysIleThr/         GTGCTGGTCAGAGGTCCCAACCATGTCAAAATTACAC         910         HisGlnSerAspValTrpSerTyrGlyValThrValT         CACCAGAGTGATGGAGTTATGGTGAGTATGGTGAACTATGC         ValTyrMet1leMetValLysCysTrpMet1leAsp2         1       GTCTACATGGTCAAATGGTCGAAAATGTTGGATGATGGACTATGGTGAGAGTGGAGTATGGTGGAGAGTATGGATGATGGATGACTGATGGACGATGATGGACGATCACGACCTCTACACCGCTCACACGCGACGATCAGACCGATGGAGACGATAGACCGACGATCACGACCGTCACACGCGACGATCGACGACGATCAGACCGACGATCACGACCGTCACACGCGACGATCGACGACGACGACGACGACGACGACGACGACGACGACGAC	20 20 20 20 20 20 20 20 20 20	680 1eLeuile_ysArgArg6in6inLys lieArgi TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser, TGATGGCTGGGTGGGGCTCCCCATATGTCTCCC 830 euAsnTrpCysMetG1n1leA1aLysG1yMet; TGAACTGGTATGCAGATTGCCAGAGGGATG SpG1uThrG1uTyrHisA1aAspG1yG1yLys' ACGAGACAGGATCCCAGCAGTGGGAGATCCCTGAC 930 yrAspG1y1 leProA1aArgG1u1 leProAspI ACGATGGGATCCCAGCAGGGACATCCCTGAC 980 alSerG1uPheSerArgMetA1aArgAspPro1 TGTCTGGATCCTCCCGGCAGGGACCCCC 1030 yrLeuVa1ProG1nG1nG1yPhePheCysPro4 ATCTGGTACCCGGAGCAGCTTGTCCAGGAGCACT	690 ys Jyr Thr Met ArgArgleul Leu ArgTACACGATGCGGAGACTGCTGC 740 Jal Tyr Lys Gly I le Trp I le ProA SICTACAAGGGCATCTGGATCCCTG 790 ArgLeuLeuGly I le Tre Leu Thr S SCCCTTCTGGGCATCTGCTGACAT 840 Ser Tyr LeuGluAspVal ArgLeuV MGCTACCTGGAGAGTGGTGCGGCTCC 890 Val ProI le Lys Trp Met Al aLeu STGCCCATCAAGTGGATGGCCCTG 940 LeuLeuGluLys GlyGluArgLeuF CTGCTGGAAAAGGGGGAGCGCCTG 990 GINArg Phe Val Val I le GINASNG CAGGCCTTTGTGGTCATCCAGATG 1040 SpProAl aProG JALaG JVG JW	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 60 160 160 160 160 160 160 160
2101 2251 2401 2551 2701 2851 3001 3151	660         LeuThrSer[ITeValSerAlaValVGIy11eLeu]         1       CTGACGTCCATCGTCTCGCGGTGGTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ALAILELYSValLeuArgGIuÄsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAC         810         MetProTyrGIyCsLeuLeuAspHisValArgGIu/         1       ATGCCCTATGGCCCCCCCAAC         860         ValLeuValLySSerProAsnHisValLyS1leThr/         1       GTGCTGGTCAAGGGTCCCAACCATGCCCGGAAU         860         ValLeuValLySSerProAsnHisValLyS1leThr/         1       GTGCTGGTCAAGGGTCCCAACCATGGTCGAAATTACAG         910         HisGInSerAspValTrpSerTyrGlyValThrVal'         1       CACCAGAGGAGTGTGTGGAGTTATGGTGGAGTTAGGTGGACGTTGG         ValTyrMet1leMetValLySCsTrpMet1leAsp1         1       GTCTACATGATCATGGTCAAGATGTGGAGGAAGATCTGTGGAGGACGAT         100       AspSerThrPheTyrArgSerLeuLeuGIuAspAsp/         1060       AspSerThrPheTyrArgSerLeuLeuGIuAspAsp/	20 20 20 20 20 20 20 20 20 20	680 16Leuile_ysArgArg6in6inLys IleArgi TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 alLysValLeuGlySerGlyAlaPheGlyThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 80 alMetAlaGlyValGlySerProTyrValSerr TGATGGCTGGTGGGGCCCCCATATGTCTCCC 830 euAsnTrpCysMetGlnIleAlaLysGlyMet: TGAACTGGTATGCGACTGCCAGGGGATG 880 spGluThrGluTyrHisAlaAspGlyGlyLys' ACGAGGACAGAGTACCCAGCAGTGGGGGGAAG 930 yrAspGlyIleProAlaArgGluIleProAspI ACGAGGACCCAGGGATCCCCGACGGGATCCCTGACU 980 alSerGluPheSerArgMetAlaArgAspProi TGTCTGGATTCTCCCCAGGGGCACGCCCGGAGGCCCCGCAGGGCCCCC 1030 yrLeuValProGlnGlnGlyPhePheCysProJ ATCTGGTACCCCAGCCCGGGGCAGCCTCCTCACU		700 infoluThr6luLeuVal6luProLeu AGGAAACGGAGCTGGTGGAGCCGCTG IspGlyGluAsnValLysIleProVal ATGGGGAGAATGTGAAAATTCCAGTG B00 SerThrValGInLeuValThrGInLeu CCACGGTGCAGCTGGTGGACACAGCTT IAHisArgAspLeuAlaAlaArgAsn TACACAGGGACTTGGCCGCTCGGAAC 900 iluSerIleLeuArgArgArgPheThr IAGTCCATTCTCCGCCGGCGGTCCACC 900 iluSerIleLeuArgArgArgPheThr IAGTCCATTCTCCGCCGGCGGTCACC 900 iluSerIleLeuArgArgArgPheThr IAGTCCATTCTCCGCCGGCGGTCACC 900 iluSerIleLeuArgArgArgPheThr IAGTCCATTCTCCGCCGGCGGTCCCC 900 iluSerIleLeuArgArgArgPheThr 1000 100 IuAspLeuGlyProAlaSerProLeu IAGGACTTGGGCCCACCGCCGCCGCCT 1000 100 100 100 100 100 100 1
2101 2251 2401 2551 2701 2851 3001 3151	660         LeuThrSer[ITeValSerA1aVa)ValG1y11eLeu[         1       CTGACGTCCATCGTCTCTGCGGTGGTGGCATTCTGG         710         ThrProSerG1yA1aMetProAsnG1nA1aG1nMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCCGCAGTGG         ACACCTAGCGGAGCGATGCCCCAACCAGCCCCCAAC         1       ACACCTAGGGGAAAACACATTCCCCCAAC         810         MetProTyrG1yCsLeuLeuAspHisVa1ArgG1u/         1       ATGCCTATGGCGCCCCCTCTTAGACCATGCCCGGAGU         810         MetProTyrG1yCsLeuLeuAspHisVa1ArgG1u/         1       ATGCCTATGGCCCCCCCACCATGCCAAATTACAG         910         1       ATGCCTAGGGTCCCAACCAGCCAAACAAATTACAG         910       HisG1nSerAspVa1TrpSerTyrG1yVa1ThrVa1         1       CACCAGAGTGATGTGTGGAGGTTATGGTGGACATTGCACTGGC         1       GTCTACATGATCATGGTCGAAAGTTATGGTGTGACTGTG'         1       GTCTACATGATCATGGTCAAATGTTGGAGTTATGGTGTGACTGTG'         1       GTCTACATGATCATGGTCAAATGTTGGAGGTGAGATTAGCT         1       GTCTACATGATCATGGTCAAATGTTGGAGGTGAGAGTGACTGGTGAGAGTGAGT	670 670 670 670 670 670 670 670	680 Teleuile ysArgArgGinGinLys IleArgi TCCTCATCAAGCGACGCACCAGAAGATCCGG 730 alLysValLeuGlySerGiyAlaPheGlyThri TGAGGGTGCTTGGATCTGGCCTTTTGGCACA 80 alMetAlaGlyValGlySerProTyrValSer TGATGGCTGGTGGGGCCCCCCATATGTCTCC 830 euAsnTrpCySMetGinIleAlaLysGlyMet: TGACTGGGTATGCAGATTGCCAAGGGGATG 880 △ spGluThrGluTyrHisAlaAspG1yG1yLys' ACGAGACAGAGTACCATGCAGATTGGCGGGCAAG 930 ✓ yrAspG1yIleProAlaArgG1uIleProAspl ACGATGGGATCCCCAGCCGGGGAATCCCTGCC 930 29 alSerGluPheSerArgMetAlaArgAspProl TGTCTGAATTCTCCCGCATGGCGGGGCACCCC 1030 yrLeuValProG1nG1nG1VphePheCyProA ATCTGGTACCCCAGCAGGGCTCCCCA 1080 erProLeuAlaProSerG1uG1vAlaG1ySer	690 ys Jyr Thr Met ArgArgLeuLeuG VaGTACACGATGCGGGAGACTGCTGG 740 741 Tyr Lys Gly I le Trp I le ProA SGC TTCTGGGGCATCTGGACCCCTG 790 ArgLeuLeuG I yI le Tr Leu Thr S CGC TTCTGGGCATCTGCTGACAT 840 Ser Tyr LeuG lu AspVa l ArgLeuV AGCTACCTGGAGGAGTGGCGGCTGC 890 741 Pro I le Lys TryMet Al aLeuG 5TGCCCATCAAGTGGATGGCGCGCCCG 890 940 LeuLeuG lu Lys GlyG lu ArgLeuF 790 Gln Arg Phe Va I Va I I le Gin Asn CAGCGCTTTGTGGTCATCCAGATG 290 Gln Arg Phe Va I Va I I le Gin Asn CAGCGCTTTGTGGTCATCCAGATG 290 Sap Pro Al a Pro GlyA la GlyG ly AspVa I Phe AspG I vAspLeuG I vA	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGAGCCGCTG 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 800 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGGACACAGCTT 850 301 SerThrVa1G1nLeuVa1ThrG1nLeu CCACGGTGCAGCTGGTGGCACACGGAC 900 310 Ser11eLeuArgArgArgArgPheThr 16ATCCATTCTCCGCCGCGGGTCACC 900 10 Ser11eLeuArgArgArgArgPheThr 16ATCCATTCTCCGCCGCGGGTCACC 900 10 Ser11eLeuArgArgArgArgPheThr 100 10 Ser11eLeuArgArgArgArgPheThr 100 10 Ser11eLeuArgArgArgArgPheThr 100 10 SuAspLeuG1yProA1aSerProLeu 1000 10 Ser11gGCCCAGCCAGCCAGTCCTTG 1000 10 Ser11gGCCCACCAGGCCCGCCGCCAGTCCTTG 1000 10 Ser11gGCCCACCAGGCCCGCCGCCGCCAGTCCTTG 1000 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 1000 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 1000 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 1000 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 100 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 100 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 100 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 100 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 100 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser11gGTCCACCACAGGCACCGCAGCTCA 10 Ser111gGTCCACCACAGGCACCGCAGCTCA 10 Ser111gGTCCACCACAGGCACCGCAGCTCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCACCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTCACCACGCAGCTC
2101 2251 2401 2551 2701 2851 3001 3151 3301	660           LeuThrSer[ITeValSerAlavalValGiy1leLeu]           CTGACGTCCATCGTCTGCGGTGGCATTCTGG           710           ThrProSerGiyAlaMetProAsnGInAlaGInMet/           ACACCTAGCGGAGGGAGCGATGCCCAACCAGGCGCAGATGG           760           AtaileLysValLeuArgGluAsnThrSerProLys/           GCCATCAAAGTGTTGAGGGAAACACATCCCCCCAAA           810           MetProTyrGlySteuLeuLeuAspHisValArgGlu/           ATGCCCTATGGCCCCTCTTAGACCATGTCCAGGAAA           810           MetProTyrGlySteuLeuLeuAspHisValArgGlu/           ATGCCCTATGGCCCCTCTTAGACCATGTCCAGGAAA           800           ValLeuValLySSerProAsnHisValLySIleThr/           GTGCTGGTCAAGAGTCCCAACCATGTCAAAATTACAA           910           HisGInSerAspValTrpSerTyrGlyGilaThrValT           CACCAGAGTGATGTGTGGAGATTTGGCTGAAATTGACT           910           HisGInSerAspValTrpSerTyrGlyGIJATTGAT           910           HisGInSerAspValTrgSerTyrGlyGIJATTGAT           910           HisGGTGATGTGTGGAGTTGTGGAGATTGGCGAAAGTTGGCGAAAGTGGTGAATTGACT           910           GCCACACGACTGTGTGGAGACTTGGCGAAAGTTGGCGAAGTGGUASATGACTGGTGAAGCACTGT           GCAAGACACCTTCTACCGCTCACTGCTGGAAGCAGCAT           1010           AspSerThrPheTyrArgSerLeuLeuGUAUASpASp/           GACAGCACCTTC	20 20 20 20 20 20 20 20 20 20	680 1eLeuIIL ysArgArgG1nG1nLysI1eArgI TCTCATCHAGCGACGSCAGCAGAAGATCCGG 730 alLysValLeuG1ySerG1yA1aPheG1yThr TGAAGGTGCTTGGACTTGGCGCTTTTGGCACA 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser TGATGGCGGGGTGGGGCCCCCCATATGTCTCC 830 euAsnTrpCySMetG1nI1eA1aLysG1yMet: TGAACGGGGTATGCAGATTGCCAAGGGGATG 880 spG1uThrG1uTyrHisA1aAspG1yG1yLys ACGAGACAGGGATCCCAGACTGGCGGAAGCCC 980 yrAspG1yI1eProA1aArgG1uI1eProAspI ACGATGGGATCCCAGCCGGGAACTCCCTGAC 980 alSerG1uPheSerArgMetA1aArgAspPro1 TGTCGAATTCCCCGACGGCCAGGGCACCCC 1030 yrLeuVa1ProG1nG1nG1yPhePheCyProv ATCTGGTACCCCAGCAGGGCTGCCTCTCCTTCCCA 1080 erProLeuA1aProSerG1uG1yA1aG1ySer/ CTCCACTGCGACCCCCCAAGGGCCTGCCCAAGGGCTCCCCCAAGGCCCCCCCC		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 350 350 350 367 3750 367 3750 375
2101 2251 2401 2551 2701 2851 3001 3151 3301	LeuThrSer[ITeValSerAlavalVaG04] LeuThrSer[ITeValSerAlavalVaG04] TGACGTCCATCGTCTGCGGTGGTTGGCATTCTGG 710 ThrProSerG1yAlaMetProAsnG1nAlaG1nMet/ ACACCTAGCGGAGCGCAGAGGCGCAGAGG AlaIleLysValLeuArgG1uÅsnThrSerProLys/ GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA 810 MetProTyrG1yCsLeuLeuAspHisValArgG1u/ ATGCCTATGGCTGCCTCTTAGACCATGCCCAAGTGGC ValLeuValLySSerProAsnHisValLysIleThr/ GTGCTGGTGGTCAGAGGTCCCACACGTGCACATGCCCAGGAGA 910 HisG1nSerAspValTrpSerTyrG1yValThrValT CACCAGAGGAGTGGGAGTTATGGTGACATGTC ValTyrMet1leMetValLysCysTrpMet1leAsp2 1 GTCACATGGTCCAGACATGTCGGAGATATGGCAGATGGC ValLeuValLySSerProAsnHisValLySTrpMet1leAsp2 1 GTCACAGAGGTGCCAAGTATGGTGGAGTATGGTGACATGTC ValTyrMet1leMetValLysCysTrpMet1leAsp2 1 GACAGCACCTTCTACCGCTCACTGCGGAGGACGATG 000 SerThrPheTyrArgSerLeuLeuG1uAspAsp4 1 GCCAGGAGGCGGGGGGGGGGGGCGCGGACCTAGGCACATGGCCACAGGACCATGGCGACCTATGGCGACCTATGGCGACCTACGGCCACAGGACGACCTCGGAGGACCTAGGGCGGCGGGGGGGCGGCGGCGGCGGCACCTGGACCATGGCCACAGGACCATGGCGACCTACGGCGACCTAGGGCGCGGGGGGCGCGGGGGCGCCACAGGACCTACGGCCACATGGGCGCGGGGGCGCGGGGGCGCTGGGCACCTAGGGC 1110	20 20 20 20 20 20 20 20 20 20	680 1eLeuile_ysArgArg6in6inLys1leArgi TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 1LysValLeuG1ySerG1yA1aPheG1yThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 780 alMetA1aG1yVa1G1ySerProTyrVa1Ser, TGATGGCTGGGTGGGGCTCCCCATATGTCTCCC 830 euAsnTrpCysMetG1n1leA1aLysG1yMet: TGAACTGGTTATGCAGATTGCCAGGGGATG 880 SpG1uThrG1uTyrHisA1aAspG1yG1ySyS ACGAGACAGAGTACCATGCAGATGGGGGCAGG 930 yrAspG1y1leProA1aArgG1u1leProAspI ACGAGGATCCCAGCCGGGAGATCCCCGATGGGGCACC 980 alSerG1uPheSerArgMetA1aArgAspProJ ATCTGGTACCCCGATGGCCAGGGACCCC 1030 yrLeuVa1ProG1nG1nG1yPhePheCyProJ ATCTGGTACCCCGGCAGGGGCCCGG 1080 erProLeuA1aProSerG1uG1yA1aG1ySer CTCCACTGGCACCCCGACGGGGGCTGCCTGCCAG	690 ys Jyr ThrMetArgArgLeuLeuG AGTACACAGTGCGGAGACTGCTGC 740 741 TyrLys Gly1leTrp1leProA SITCTACAAGGGCATCTGGACTCCTG 790 ArgLeuLeuGly1leTrtLeuThrS SGCCTTCTGGGCATCTGCTGCACAT 840 Ser TyrLeuGluAspValArgLeuV MGCTACCTGGAGGAGTGGCGGCTCC 890 741Pro1leLys TrpMetAlaLeuG 1040 Ser LeuGluLys GlyGluArgLeuF 1040 SaccctTGGGCAGGGGGCGCGGGGCAC 1040 SpProAlaProGlyAlaGlyGlyAspLeuGlyM SACCTGCCGGGGACSGGCTGGGGAC 1040 104	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 60 160 160 160 160 160 160 160
2101 2251 2401 2551 2701 2851 3001 3151 3301	660         LeuThrSer[ITeValSerAlaVa)ValGIy1IeLeu[         1       CTGACGTCCATCGTCTCGCGGTGGGTTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGTGG         ACACCTAGCGGAGCGATGCCCCAACCAGCCCCCAAC         810         MetProTyrGIyCsLeuLeuAspHisValArgGIu/         AGCCTATGGCCCCCCCTCTTAGACCATGCCCGGAA         860         ValLeuValLySSerProAsnHisValLys1leThr/         1       GCCGCGCCCCCCAACCATGCCCGGGAGCGCCGACCATGCCGGGAGTATGGGTGGACGTATGGTGGAGGTATGGTGGACGTATGGTGGACGTATGGTGGACGTATGGTGGACGTATGGTGGACGTATGGTGGACGTATGGTGGAGTGATGGATG	200 2014 21 Va 1 Va 1 Va 1 Va 1 Va 1 Phe Gi yī CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I 1 eL eu Lys Gl u Thr G 1 uL eu Arg Lys V CGGATCCTGAAAGAAGACGGAGCTGAGGAAGG 770 Al aAsnLys Gl u I 1 eL eu Asp Gi uAl a Tyr V SCCAACAAAGAAATCTTAGACGAAGCATACG AsnArg Gl yArg Leu Gl ySer G 1 nAsp Leu L AACCGCGGACGCCTGGGCTCCCAGGACTGC 870 AspPhe Gl yL eu Al aArg Leu Leu Asp I 1 eA SACTTCGGGCTGGCTGCGCTGCCGCAGACCTGC 920 TrpG 1 uL eu Met Thr Phe Gl yA 1 aL ys Pro T TGGGAGCTGATGCGCAGAGTTCGGGCTGCCAAACCT 970 Ser Gl u SArg Pro Arg Phe Arg Gl uL eu W 1020 AspMet Gl yAspL eu Va 1 AspAl aG 1 uG 1 uT 3ACATGGGGAGCCTGGTGGAGTCGG 1020 AspMet Gl yAspL eu Va 1 AspAl aG 1 uG 1 uT 3ACATGGGGAGCCTGGTGGAGTGGAGTCGG 1070 Leu Gl u Pro Ser Gl u Gl uG 1 uAl Pro Arg S CTGGAGCCCTCTGAAGAGGAGGGGCCCCCAGGT 1120 Ser Gl u AspPro Thr Va 1 Pro Leu Pro Ser G	680 16Lewild_ysArgArg6in6inLys IIeArgi TCCCATCAAGCGACGCAGCAGAAGATCCGG 730 alLysValLewG1ySerG1yA1aPheG1yThri TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI 80 alMetA1aG1yVa1G1ySerProTyrVa1Ser/ TGATGGCTGGTGGGGCCCCCATATGTCTCCC 830 ewAsnTrpCysMetG1nI1eA1aLysG1yMet: TGAACGGGTATGCGGATGCCAGGGGATG 880 5pG1uThrG1uTyrHisA1aAspG1yG1yLys' ACGAGGCACGAGTACCCAGCAGTGGGGGCAGG 930 yrAspG1yI1eProA1aArgG1u11eProAsp1 ACGAGGGATCCCAGCAGTGGGGGCAGG 930 yrAspG1yI1eProA1aArgG1u11eProAsp1 ACGAGGACCCGGCAGGGCCCGGCAGGGCCCCG 980 alSerG1wPheSerArgMetA1aArgAspProi TGTCTGGATTCTCCCGATGGCCAGGGACCC 1030 yrLewVa1ProG1nG1nG1yPhePheCysProA ATCTGGTACCCCCGCAGGCCAGGGCCCCCC 1030 erProLewA1aProSerG1wG1yA1aG1ySer/ CTCCACTGGCACCCCCGAAGGGCCGCCCCC 1130 luThrAspG1yTyrVa1A1aProLewThrCys	690 ys Jyr Thr Met ArgArgi eul eug AGTACACGATGCGGGAGACTGCTGG 740 /a Tyr Lys Gly I le Trp I le ProA STCTACAAGGGCATCTGGACCCCTG 790 ArgLeuLeuGly I le Tri Leu Thr S GCCTTCTGGGCATCTGCTGACAT 840 Ser Tyr LeuGluAspVal ArgLeuV AGCTACCTGGAGGAGGGGGGGCGCG 890 /a I Pro I le Lys Trp Met Al a LeuG 1040 AspCTACCTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGG	700 11nG1uThrG1uLeuVaTG1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 15pG1yG1uAsnVa1Lys11eProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 14TGGGGGGAGATGTGGAAAATTCCAGTG 14TGGGGGGCAGCTGGGGCACCAGCTT 14TGGGGGCAGCGGGGGGCACCAGCGT 14TGCACAGGGGCTGGGACCAGCGT 14TGCCACGGCGGCGGCGGCGGCGCGCGGCGCGCGCGGCGGCG
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451	660         LeuThrSer[ITeValSerAlaValVaGIy1] eLeu[         1       CTGACGTCCATCGTCTCTGCGGTGGTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         AlaIleLysValLeuArgGIuÄsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAG         810         MetProTyrGIyCsLeuLeuAspHisValArgGIu/         1       ATGCCTATGGCCCTCTTAGACCATGCCCGAAC         810         MetProTyrGIyCsLeuLeuAspHisValArgGIu/         1       ATGCCTATGGCCCCCTCTTAGACCATGCCGGAAATTACAC         860         ValLeuValLysSerProAsnHisValLysIleThr/         1       GTGCTGGTCAAGAGTCCCAACCATGCCAAAATTACAG         910         HisGInSerAspValTrpSerTyrGlyValThrVal'         1       CACCAGAGTGATGGTGAGGTATGGTGAAATTACGT         910       HisGInSerAspValTrpSerTyrGlyValThrVal'         1       CACCAGAGCCTTCTACGGTCAAATGTGGAGGAGCGATG         910       AspSerThrPheTyrArgSerLeuLeuGluAspAsp/         1       GACAGCACCTTCTCACGGCTCACTGCTGGAGGACCGATC         10       AspSerThrPheTyrArgSerLeuLeuGluAspAsp/         1       GACAGCACCTTCTACGGGGGGGGGGGACCTGACACTAGGGG         1010       SerThrArgSerGIyGIyGIyGIySIy	C70     CCAACAAAGAAACGGACGCACGCAGCGACG     C70     CCAACAAAGAAATCTTAGACGAAGG     C70     CCAACAAAGAATCTTAGACGAAGGACGCACGCC     CCACAAAGAAATCTTAGACGAAGGACGCCCCCCAGGACTGC     CCACAAAGAATCTAGACGAAGCCTGCC     S70     SchargGlyArgLeuGlySerGlnAspLeuL     AACCGGGGACGCCGGCTGGCTCCCAGGACTGC     S70     SpPheGlyLeuAlaArgLeuLeuAspIleA     GACTTCGGGCTGGCTCGGCTGCCGGCGCCAAACCT     S20     CTGGAGCCTGATGACTTTTGGGCCAAACCT     S70     SchargProArgPheArgGluLeuV     CTGAATGCGGGGACCTGGTGGGAGTTCG     S70     SchargProArgPheArgGluLeuV     CTGAATGCGGGGACCTGGTGGAGTTCCGGGGAGTTGG     S70     SpPhetGlyAspLeuValAspAlaGluGluT     SACATGGGGGACCTGGTGGAGATCCGGGGAGTTGG     1020     LeuGluProSerGluGluGluAlaProArgS     CTGGAAGCCCCCCAGGAGGGGGCCCCCAGG     ScrGluSpProTn'valProLeuProSerG     GGGAGGACCCCCACGGTACCCCGCCCCCCGC     ScrGluSpProTn'valProLeuProSerG     SGGAGGACCCCCCAGGTACCCCCGCCCCCCGC	680         Teleuile ysArgArgGinGinLys IleArgi         TCCTCATCAAGCGACGCAGCAGAAGATCCGG         730         allysValLeuGlySerGiyAlaPheGlyThr         TGAAGGGCTTGGACTCGCCGCTTTTGGCACAG         alMetAlaGlyValGlySerProTyrValSer         TGAGGGTGGGTGGGCCCCCCATATGCTCCG         830         euAsnTrpCySMetGInIleAlaySGlyMet:         TGAACGGGATATGCAGATTGCCAAGGGATG         spGluThrGluTyrHisAlaAspGlyGlyLys'         ACGAGGACAGAGTCCCCGAATGCGGGGCACC         930         yrAspGlyIlPProAlaArgGluIIPProAspl         ACGAGACCCCGGAGGATCCCCGACCCCG         930         yrAspGlyIlPProAlaArgGluIIPProAspl         ACGAGACCCCGGAGGCTCCCTACC         980         alserGluPheSerArgMetAlaArgAspSproi         TGTGGATCCCCCCCCGAAGGGCTCCTCCCC         1030         yrLeuValProGInGInGINGIYPhePheCyPro/         ATCTGGFACCCCCGGACCCCCCGAAGGGCTGCCTCCC         1030         yrLeuValProSINGINGIYGIYAIAlaFroLeuThrCs         1030         yrLuAlaProLeuThrSs         1030         yrLuAlaProSerGiUGIYAlaGIYSer         CTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCCCCCAGGGCCCCCCCC		700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGGGCCCCTC 350 350G1yG1uAsnVa1Lys11PProVa1 34TGGGGAGAATGTGAAAATTCCAGTG 300 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu 300 301 SerTheLeuArgArgArgPheThr 302 SerTheLeuArgArgArgPheThr 303 SerTheLeuArgArgArgPheThr 303 SerTheLeuArgArgArgPheThr 303 SerTheLeuArgArgArgPheThr 304 SerTheLeuArgArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 305 SerTheLeuArgArgArgArgPheThr 305 SerTheLeuArgArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 306 SerTheLeuArgArgArgArgPheThr 305 SerTheLeuArgArgArgPheThr 305 SerTheLeuArgArgArgArgPheThr 305 SerTheLeuArgArgArgArgArgArgArgArgArgArgArgArgArgA
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451	660         LeuThrSer[ITeValSerAlava]ValGly1leLeu[         1       CTGACGTCCATCGTCTGCGGTGGCATTCGGCATTCTGCGATGCCATCGTCGCGATGCCATCGCGTGGCATTCTGG         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGG         760         AlaIleLysValLeuArgGluAsnThrSeProLys/         GCCATCAAAGTGTTGAGGGAAACACATCCCCCCAAA         810         MetProTyrGlytsLeuLeuAspHisValArgGlu/         ATGCCCTATGGCCTCCTTAAACCATGTCCAGGAA         810         MetProTyrGlytsLeuLeuAspHisValArgGlu/         ATGCCCTATGGCCACCATGTCAAACAATTCCCCCAAA         810         MetProTyrGlytsLeuLeuAspHisValArgGlu/         ATGCCCTATGCCAACACATGTCCAAACCATGTCCGAGAA         810         MetProTyrGlytsLeuLeuAspHisValLySIleThr/         GGCGTGGTGAAGGTCCCAACCATGTCAAAATTACAA         910         HisGlnSerAspValTrpSerTyrGlytalThrValT         CACCAGAGTGATGTGTGAGACTTAGCTGGCGACTGTG         YalTyrMetIleMetValLyStysTrpMetIleAspi         GCTACATGGTCGACACTGCTGCAGCACTGCGGAGCGACT         YalTyrMetIleMetValLyStysTrpMetIleAspi         GCCAGGAGTGGCGGTGGGGGGGGGGGGGCGCCTGGAGACTAGGCG         YalTyrMetIleWetValLyStysTrpMetIleUsUAspAsp/         GACAGGACCTTACCGCTCACACTGCTGCGAGCCACTGGCGGCGACTGGGGGACCTAGGGG         SerThrPheTyrArgSerLeuLeuGUAspAspl <td>20 20 20 20 20 20 20 20 20 20</td> <td>680         1eLeuile_ysArgArgEinGinLys IleArgi         730         1LysValLeuGlySerGlyAlaPheGlyThri         TGAAGGGACGCACGCAGAGAGATCCGG         730         1LysValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGACTGGCGTTTTGGCACAI         780         alMetAlaGlyValGlySerProTyrValSer/         TGATGGCTGGGTGGGCTCCCCATATGTCTCC         830         euAsnTrpCysMetGInIleAlaLySGIyMet:         TGAACGGGTGGTATGCAAGAGATGCCAAGGGATG         880         SpGluThrGluTyrHisAlaAspGIyGlyGlyLys         ACGACGAGGATCCAACCGGGAAGACCCCTGCA         930         YrAspGlyI leProAlaArgGluI leProAspl         alSerGluPheSerArgMetAlaArgGsgACCCCTGCA         980         alSerGluPheSerArgMetAlaArgAspProi         ATCTGGTACCCAGCAGGGCTTGTTCTCTCACA         1030         yrLeuValProGInGInGIYPheProy         ATCTGGTACCCCAGCAGGGCTCTTCTCTCTCCA         1080</td> <td>690 yaTyrThrMetArgArgLeuLeuG VaTYrThrMetArgArgLeuLeuG VaTYrLysGIyIIeTrpIIeProA TCTACAAGGGCATCTGGATCCCTG ArgLeuLeuGIyIIeTrLeuThrS CGCCTTCTGGGCATCTGCTGACAT 840 SerTyrLeuGIuAspVaIArgLeuV AGCTACCTGAGGATGGCGGCGCCGC 890 VaIProIIeLysTrpMetAIaLeuG 21GCCCATCAAGTGGATGGCGGCGCCG 890 VaIProIIeLysTrpMetAIaLeuG 21GCCGCATCAAGTGGATGGCCGCTG 900 201ArgPheVaIVaIIIeGInAsnG CAGCGCTTTGTGGTCATCCAGATG 1040 35pProAIaProGIyAIaGIyGIyM SACCCTGCCCCGGGCGCTGGGGGC 1090 ASTGATTTGATGGTGACCTGGGAA SePProGINProGIUTyrVaIAsnG SGCCCCCGCCTGAATATGCACCC</td> <td>700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGAGCCGCTG 150G1uG1uAsnVa1Lys1leProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 160G102000 3000 30100000000000000000000000000</td>	20 20 20 20 20 20 20 20 20 20	680         1eLeuile_ysArgArgEinGinLys IleArgi         730         1LysValLeuGlySerGlyAlaPheGlyThri         TGAAGGGACGCACGCAGAGAGATCCGG         730         1LysValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGACTGGCGTTTTGGCACAI         780         alMetAlaGlyValGlySerProTyrValSer/         TGATGGCTGGGTGGGCTCCCCATATGTCTCC         830         euAsnTrpCysMetGInIleAlaLySGIyMet:         TGAACGGGTGGTATGCAAGAGATGCCAAGGGATG         880         SpGluThrGluTyrHisAlaAspGIyGlyGlyLys         ACGACGAGGATCCAACCGGGAAGACCCCTGCA         930         YrAspGlyI leProAlaArgGluI leProAspl         alSerGluPheSerArgMetAlaArgGsgACCCCTGCA         980         alSerGluPheSerArgMetAlaArgAspProi         ATCTGGTACCCAGCAGGGCTTGTTCTCTCACA         1030         yrLeuValProGInGInGIYPheProy         ATCTGGTACCCCAGCAGGGCTCTTCTCTCTCCA         1080	690 yaTyrThrMetArgArgLeuLeuG VaTYrThrMetArgArgLeuLeuG VaTYrLysGIyIIeTrpIIeProA TCTACAAGGGCATCTGGATCCCTG ArgLeuLeuGIyIIeTrLeuThrS CGCCTTCTGGGCATCTGCTGACAT 840 SerTyrLeuGIuAspVaIArgLeuV AGCTACCTGAGGATGGCGGCGCCGC 890 VaIProIIeLysTrpMetAIaLeuG 21GCCCATCAAGTGGATGGCGGCGCCG 890 VaIProIIeLysTrpMetAIaLeuG 21GCCGCATCAAGTGGATGGCCGCTG 900 201ArgPheVaIVaIIIeGInAsnG CAGCGCTTTGTGGTCATCCAGATG 1040 35pProAIaProGIyAIaGIyGIyM SACCCTGCCCCGGGCGCTGGGGGC 1090 ASTGATTTGATGGTGACCTGGGAA SePProGINProGIUTyrVaIAsnG SGCCCCCGCCTGAATATGCACCC	700 31nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGAGCCGCTG 150G1uG1uAsnVa1Lys1leProVa1 ATGGGGAGAATGTGAAAATTCCAGTG 160G102000 3000 30100000000000000000000000000
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451	660         LeuThrSer[ITeValSerAlava]ValGly1leLeu[         1       CTGACGTCCATCGTCTGCGGTGGCATTCGGATGCCATTCGT         710         ThrProSerGlyAlaMetProAsnGlnAlaGlnMet/         1       ACACCTAGCGGATGCCCAACCAGGCGCAGATG         ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATG         6CCATCAAAGGTTGGAGGAAAACCACTCCCCCAAA         810         MetProTyrGlyCsLeuLeuApgHisValArgGlu/AsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACCATCCCCCCAAA         810         MetProTyrGlyCsLeuLeuApgHisValArgGlu/         ATGCCCTATGGCCGCCTCTTAGACCATGCCCAGGAAA         860         ValLeuValLySSerProAsnHisValLyS1leThr/         GTGCTGGTCAGAGGTCCCACACGTGCAAAATTACA         910         HisGlnSerAspValTrpSerTyrGlyValThrVaT         1       CACCAGAGGGAGTGTGGGGAGTTATGGTGGAGTGAGGGGGACCTAGGCACT         910         HisGlnSerAspValTrpSerTyrGlyValThrVaT         1       CACCAGAGGAGTGTGGGGAGTATGGTGGAGGGAGCACTAGGGC         910       AspSerThrPheTyrArgSerLeuLeuGluAspAspJ         1       GACAGCACCTTCTCCCGCTCACCGCTGACACTAGGGC         1000       SepThrArgSerGlyGlyG1yG1yAspLeuThrLeuGly1         1       TCTACCAGGGAGGGGGGGGGGGGGGGGGGCCCTGACACTAGGGC         110       LeuProThrHisAspProSerProLeuGnArgTyr5         1100 <tdleupro< td=""><td>200 2014 21 Va 1 Va 1 Va 1 Va 1 Va 1 Phe Gi yi CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I 1 eL eu Lys Gl u Thr G lu Leu Arg Lys V CGGATCCTGAAAGAAGACGGGGCTGAGGAAGG 770 Al aAsnLys Gl u I 1 eL eu Asp G lu Al a Tyr V SCCAACAAGAAACATCTTAGACGAAGCATACG AsnArgGl yArg Leu Gl ySerG 1 nAspL eu L ACCGCGGACGCCTGGGCTCCCAGGACCTGC 870 AspPheGl yL eu Al aArg Leu Leu Asp I 1 eA SaCTTCGGGCTGGCTCGGCTCCCAGGACCTGC 920 TrpG 1 u Leu Met Thr PheG 1 yA 1 a Lys Pro T TGGGAGCTGATGGCCTGGGGCAGCCTGG 920 TrpG 1 u Leu Met Thr PheG 1 yA 1 a Lys Pro T TGGGAGCTGATGGCCTGGGGAGCTGCG 1020 AspMetG1 yAspL eu Va 1 AspA1 aG1 uG 1 U SACATGGGGGACCTGGGGAGGGGGCCCCCAGGACT 1020 AspMetG1 yAspL eu Va 1 AspA1 aG1 uG 1 U SACATGGGGGACCTGGGGAGGGGGCCCCCAGGT 1070 Leu G1 UProSerG1 uG 1 uG 1 uA 1 a Pro Arg S CTGGAGCCCTCTGAAGGAGGAGGGGCCCCCCAGGT 1120 SerG1 uAspPro Thr Va 1 Pro Leu Pro SerG GGGGGGCCCCCCAGTACCCCTCGCCCTCGCCCCCCCCCC</td><td>1eLeuile_ysArgArg6in6inLys1leArgi         1eLeuile_ysArgArg6in6inLys1leArgi         730         allysValLeuG1ySerG1yA1aPheG1yThri         1cAAgGTGCTTGGATCTGGCGCTTTTGGCACAI         780         alMetA1aG1yVa1G1ySerFoTyrVa1Ser,         TGATGGCTGGGTGTGGGCTCCCCATATGTCTCC         830         euAsnTrpCysMetG1n1leA1aLysG1yMet1         TGAACGGCACGACACGCAGAGACTGCCAGGGCAGG         spG1uThrG1uTyrHisA1aAspG1yG1yLys'         ACGAGCACGGACTCCCGCAGAGGGGCGCAGG         930         yrAspG1y1leProA1aArgG1u1leProAspI         ACGGACCCGGCAGGGCCCGGGAGACCCCGGGAGACCCCGGCAGGACCCCGGCAGGACCCCTGCCCTGCCCCGCAGGGCCAGGGCCCGGGACGCCGGGACCCCCGCAGGGCCCGGGACTCCTGCCCCGACGGGCCCGGGACCCCCCCGAGGGGCCGGGCCCGGGACTCCCCGACGGGCCCGGGACTCCCCGACGGGCCCGGGACCCCCCGCAGGGGCCGCGCCCGCGAGACCCCCC</td><td>690 ys Jyr Thr Met ArgArgu Eule uu ArgTACACGATGCGGAGACTGCTGG 740 741 Tyr Lys Glylle Trp IleProA SITCTACAAGGGCATCTGGACTCCTG 790 ArgLeuLeuGlylle Trp IleProA Sec Tyr LeuGluAspValArgLeuThrS Sec Tyr LeuGluAspValArgLeuGly Sec Tyr LeuGluAspValAspLeuGly Sat Gac TtGTGGACTGGGAC 1090 Sat Sec Thr Sec Tor Sec</td><td>700 influThr6luLeuVal6luProLeu AGGAAACGGAGCTGGTGGAGCCGCTG influThr6luLeuVal6luProLeu AGGAAACGGAACTGGTGAGACCGCTG isp6ly6luAsnValLys1leProVal ATGGGGAGAATGTGGAAAATTCCAGTG isp6ly6luAsnValLys1leProVal ATGGGGGGCCCGCGCGGGGGACACAGCTT CCACGGTGCAGCTGGTGCACCACGGCAG (a HisArgAspLeuAlaAlaArgAsn iTACACAGGGACTTGGCCGCCGCGGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATCGTCCGCCGCGGCGGTCCACC islvSpLeuGlyProAlaSerProLeu AGGACTTGGGCCCCACGCGGCGGCCGTCC islvSpLeuGlyProAlaSerProLeu AGGGCCTGGGCCCAGGCCGCCGCCGCCGCCA isloSer1 isloSerSerSer islGGGGGCACCCACGGGCGCGCCAGCCCA isloSerSerSer islGGGGGCACCCACGGGGCCGCCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCGCAGCCA isloSerSerSer islGGGGGCCCAGGCCGCGCGCCAGCCA isloSerSerSer islGGGGGCCCAGGCCGCCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGCGCCCAGGGCCCGCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCAGCCAGCCA isloSerSerSer islGGGCACCCAGGGCCCAGCCAGCCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCAGCCAGCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCCA isloSerSerSer islGGGCCCAGCCAGGGCCCAGCCAGCCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCCAGCCCA isloSerSerSer islSerSerSerSerSerSerSerSerSerSerSerSerSerS</td></tdleupro<>	200 2014 21 Va 1 Va 1 Va 1 Va 1 Va 1 Phe Gi yi CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I 1 eL eu Lys Gl u Thr G lu Leu Arg Lys V CGGATCCTGAAAGAAGACGGGGCTGAGGAAGG 770 Al aAsnLys Gl u I 1 eL eu Asp G lu Al a Tyr V SCCAACAAGAAACATCTTAGACGAAGCATACG AsnArgGl yArg Leu Gl ySerG 1 nAspL eu L ACCGCGGACGCCTGGGCTCCCAGGACCTGC 870 AspPheGl yL eu Al aArg Leu Leu Asp I 1 eA SaCTTCGGGCTGGCTCGGCTCCCAGGACCTGC 920 TrpG 1 u Leu Met Thr PheG 1 yA 1 a Lys Pro T TGGGAGCTGATGGCCTGGGGCAGCCTGG 920 TrpG 1 u Leu Met Thr PheG 1 yA 1 a Lys Pro T TGGGAGCTGATGGCCTGGGGAGCTGCG 1020 AspMetG1 yAspL eu Va 1 AspA1 aG1 uG 1 U SACATGGGGGACCTGGGGAGGGGGCCCCCAGGACT 1020 AspMetG1 yAspL eu Va 1 AspA1 aG1 uG 1 U SACATGGGGGACCTGGGGAGGGGGCCCCCAGGT 1070 Leu G1 UProSerG1 uG 1 uG 1 uA 1 a Pro Arg S CTGGAGCCCTCTGAAGGAGGAGGGGCCCCCCAGGT 1120 SerG1 uAspPro Thr Va 1 Pro Leu Pro SerG GGGGGGCCCCCCAGTACCCCTCGCCCTCGCCCCCCCCCC	1eLeuile_ysArgArg6in6inLys1leArgi         1eLeuile_ysArgArg6in6inLys1leArgi         730         allysValLeuG1ySerG1yA1aPheG1yThri         1cAAgGTGCTTGGATCTGGCGCTTTTGGCACAI         780         alMetA1aG1yVa1G1ySerFoTyrVa1Ser,         TGATGGCTGGGTGTGGGCTCCCCATATGTCTCC         830         euAsnTrpCysMetG1n1leA1aLysG1yMet1         TGAACGGCACGACACGCAGAGACTGCCAGGGCAGG         spG1uThrG1uTyrHisA1aAspG1yG1yLys'         ACGAGCACGGACTCCCGCAGAGGGGCGCAGG         930         yrAspG1y1leProA1aArgG1u1leProAspI         ACGGACCCGGCAGGGCCCGGGAGACCCCGGGAGACCCCGGCAGGACCCCGGCAGGACCCCTGCCCTGCCCCGCAGGGCCAGGGCCCGGGACGCCGGGACCCCCGCAGGGCCCGGGACTCCTGCCCCGACGGGCCCGGGACCCCCCCGAGGGGCCGGGCCCGGGACTCCCCGACGGGCCCGGGACTCCCCGACGGGCCCGGGACCCCCCGCAGGGGCCGCGCCCGCGAGACCCCCC	690 ys Jyr Thr Met ArgArgu Eule uu ArgTACACGATGCGGAGACTGCTGG 740 741 Tyr Lys Glylle Trp IleProA SITCTACAAGGGCATCTGGACTCCTG 790 ArgLeuLeuGlylle Trp IleProA Sec Tyr LeuGluAspValArgLeuThrS Sec Tyr LeuGluAspValArgLeuGly Sec Tyr LeuGluAspValAspLeuGly Sat Gac TtGTGGACTGGGAC 1090 Sat Sec Thr Sec Tor Sec	700 influThr6luLeuVal6luProLeu AGGAAACGGAGCTGGTGGAGCCGCTG influThr6luLeuVal6luProLeu AGGAAACGGAACTGGTGAGACCGCTG isp6ly6luAsnValLys1leProVal ATGGGGAGAATGTGGAAAATTCCAGTG isp6ly6luAsnValLys1leProVal ATGGGGGGCCCGCGCGGGGGACACAGCTT CCACGGTGCAGCTGGTGCACCACGGCAG (a HisArgAspLeuAlaAlaArgAsn iTACACAGGGACTTGGCCGCCGCGGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC islvSer1leLeuArgArgArgPheThr AGTCCATCGTCCGCCGCGGCGGTCCACC islvSpLeuGlyProAlaSerProLeu AGGACTTGGGCCCCACGCGGCGGCCGTCC islvSpLeuGlyProAlaSerProLeu AGGGCCTGGGCCCAGGCCGCCGCCGCCGCCA isloSer1 isloSerSerSer islGGGGGCACCCACGGGCGCGCCAGCCCA isloSerSerSer islGGGGGCACCCACGGGGCCGCCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCGCAGCCA isloSerSerSer islGGGGGCCCAGGCCGCGCGCCAGCCA isloSerSerSer islGGGGGCCCAGGCCGCCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGCGCCCAGGGCCCGCAGCCA isloSerSerSer islGGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCGCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCAGCCAGCCA isloSerSerSer islGGGCACCCAGGGCCCAGCCAGCCAGCCA isloSerSerSer islGGGGCACCCAGGGCCCAGCCAGCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCCA isloSerSerSer islGGGCCCAGCCAGGGCCCAGCCAGCCCA isloSerSerSer islGGGGCACCCACGGGCCCAGCCAGCCCAGCCCA isloSerSerSer islSerSerSerSerSerSerSerSerSerSerSerSerSerS
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451 3601	660         LeuThrSer[ITeValSerAlaValVGIy11eLeu]         1       CTGACGTCCATCGTCTCGCGGTGGTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         ALACTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         ALAILELYSValLeuArgGIuÄsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCCAAA         810         MetProTyrGIyGSLeuLeuAspHisValArgGIu/         1       ATGCCTATGGCTCCTCTAGACCATGCCCGGAAW         860         ValLeuValLySSerProAsnHisValLySIleThr/         1       GTGCTGGTCAAGAGTCCCAACCATGCTCAAAATTACAG         910         HisGInSerAspValTrpSerTyrGlyValThrVal'         1       CACCAGAGGTGATGTGTGGAGTTATGGTGGAGTTAGGTGGACGTTG         1       GTCTACATGATCATGGTCAAGATGTTGGAGGACCATC         1       GTCTACATGATCATGGTCAAGATGTCGGAGGACCATC         1       GACAGCACCTCTCTCACCGCTGCTGCTGGCGGACCTGGAGCACT         1       GACAGCACCTCTCTCACCGCGTGGGGACCTGACACTAGGGC         1       GTCACCAGAGTGGCGGGGGGGGGGGCCCTGACACCTAGGGCACCT         1       GCCCCCACACAGAGCCCCCGCCCCTCACCAGCGGTC/         1       GCCCCACAGGGCCCCCCACCCCCCCCCACCCACCACAGGGCCCCCC	200 2014 21 Va 1 Va 1 Va 1 Va 1 Va 1 Phe Gi yī CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I 1 eL eu Lys Gl u Thr G 1 uL eu Arg Lys V CGAATCCTGAAAGAAGAGGGGACGTGAGGAAGG 770 Al aAsnLys Gl u I 1 eL eu Asp Gi uAl a Tyr V SCCAACAAAGAAATCTTAGACGAAGCATACG AsnArg Gl yArg Leu Gl ySer G 1 nAsp Leu L AACCGCGGACGCCTGGCGCTCCCAGGACTGC 870 Asp Phe Gl yL eu Al aArg Leu Leu Asp I 1 eA SACTTCGGGCTGGCTCGCGCTGCCGCAGACTGC 920 TrpG 1 uL eu Met Thr Phe Gl yA 1 aL ys Pro T TGGGAGCTGATGGCCAGAGTTCGGCTGCCAAACCT 970 Ser Gl u SArg Pro Arg Phe Arg Gl uL eu W TCTGAATGTCGGCCAGGCTGGTGGATCGGGCTGCCCCAGGATGGU 1020 Asp Met Gl yAsp Leu Va 1 Asp Al aG 1 uG 1 uT SACATGGGGAGCCTGGTGGAGGCCCCCAGGT 1070 Leu Gl u Pro Ser Gl u Gl u Gl u Al a Pro Arg S CTGGAGCCCCCCAGATACCCCCCGCTGCCCCCGG 1120 Ser Gl u Asp Pro Thr Va 1 Pro Leu Pro Ser G 4GTGAGGACCCCCAGATACCCCTGGC 1170 1160 Na Gl yAl a Thr Leu Gl u Arg Al aL ys Thr L SCTGGTGCCACTCTGGAAAGGGCCAAGCTC	680         Telewille ysArgArg6inGinLys IIeArgi         TCCCATCAAGCGACGCAGCAGAAGATCCGG         730         allysValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGATCTGGCGCTTTTGGCACA         almetAlaGlyValGlySerProTyrValSer         TGATGGCTGGGTGGGGCCCCCATATGGCGCAT         regate         almetAlaGlyValGlySerProTyrValSer         GatGGCTGGTGTGGGCCCCCATATGTCCCG         euAsnTrpCySMetG1n1leAlaLySGlyMet:         TGAACGGGATATGCCAGATTGCCGAGGGATG         spGluThrGluTyrHisAlaAspG1yG1yLys'         ACGAGGACAGAGTACCCAGCAGAGTGCGAGGGCAAGG         930       ✓         yrAspG1y1leProAlaArgG1u1leProAspI         ACGAGGCCCCGGAGAGCCCCGGAGGACCCCTGCC         1030         yrLeuValProG1nG1nG1g1PhePheCysProj         ATCTGGTACCCCCGCCCCCCCCCCCCCCCCCCCCCCCCC		700 11nG1UThrG1uLeuValG1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 15nG1uThrG1uLeuValG1UProVal ATGGGGGAGAATGTGAAAATTCAGTG 15nG1GGGGGAGATGTGGAAAATTCAGTG 15nG1UProValG1nLeu 15CCACGGTGGCAGCGGTGGCACCAGCGT 15CCACGGTGGCAGCGGTGGCACCAGCGT 15CCACGGTGGAGACTTGGCCGCTCGGAAC 1000 15USer11eLeuArgArgArgPheThr 16AGTCCATTCTCCGCCGGCGGTCCACC 1000 15USer11eLeuArgArgArgPheThr 16AGTCCATTCTCCGCCGGCGGTCCACC 1000 15USer11eLeuArgArgArgPheThr 16AGTCCATTCTCCGCCGGCGGTCCACC 1000 15USer11eLeuArgArgArgPheThr 16AGTCCATCTCCGCCGCGCGGTCCACC 1000 15UASpLeuG1yProA1aSerProLeu 16GGGCCACCACGGCGCTCCTTG 1000 15UASpLeuG1yProA1aSerProLeu 1000 15UASpLeuG1yProA1aSerProLeu 1000 15UASpLeuG1yProA1aSerProLeu 1000 15UASpLeuG1yProA1aSerProLeu 1000 1050 10
<ul> <li>2101</li> <li>2251</li> <li>2401</li> <li>2551</li> <li>2701</li> <li>2851</li> <li>3001</li> <li>3151</li> <li>3301</li> <li>3451</li> <li>3601</li> </ul>	660         LeuThrSer[ITeValSerAlaValVaGIy11eLeu]         CTGACGTCCATCGTCTCTGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyAlaMetProAsnGInAlaGInMet/         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         AlaIleLysValLeuArgGIuÄsnThrSerProLys/         GCCTCAAAGTGTTGAGGGAAAACACATCCCCCAAA         MetProTyrGIyCsLeuLeuAspHisValArgGIu/         ATGCCTATGGCTCCTATGACCATGCCCCGAAC         WalLeuValLysSerProAsnHisValLysIleThr/         IGTGCTGGTCAAGAGTCCCCAACCATGCCCAAATTACAC         910         HisGInSerAspValTrpSerTyrGlyValThrVal'         CACCCAGAGTGATGTGTGGAGTTATGGTGGAGATTACGTGTGACATGGT         YalTyrMetIleMetValLysVsTrpMetIleAsp1         GTCTACATGATCATGGTCCAAATGTGGAGGAGCGTT         YalTyrMetIleMetValLysVsTrpMetIleAsp1         GACAGCACCTTCTACCGCTCACTGCTGGAGGAGCAGTT         SerThrPheTyrAgSerLeuLeuGIuAspAsp/         GACAGCACCTTCTACCGCTCACTGCTGAGAGCACGT         1010         AspSerThrPheTyrAgSerLeuGInArgTyr5         CTCCCCACACATGACCCAGCCCTCTCACAGCGGCACCTGACACTAGGCG         1110         LeuProThrHisAspProSerProLeuGInArgTyr5         CTCCCCCAACATGACCCCAGCCCTCTCTCACAGCGGACCCT         SerProArgGIuGIyProLeuProAlaAIaArgPro/         TCGCCCCGAGAGGCCCTCTCTCTCCCCGCCCCCCCACACT         SerProA	C70     C	680         Teleuile, ysArgArgGinGinLys IleArgi         TCCTCATCAAGCGACGCAGCAGAAGATCCGG         730         allysValLeuGlySerGiyAlaPheGlyThr         TGAAGGGCTTGGACTCGCGCTTTTGGCACAG         alMetAlaGlyValGlySerProTyrValSer         TGATGGCTGGTGGGCCCCCCATATGCCCC         830         alMetAlaGlyValGlySerProTyrValSer         GaGGCTGGGTGGGCCCCCCATATGCTCCC         830         euAsnTrpCySMetGInIleAlaLySGlyMet:         GAACGGGGATATGCCAGATTGCCAGGGATG         spGluThrGluTyrHisAlaAspGlyGlyLyS'         ACGAGACAGAGTACCCAGCAGGAGTCCCTACG         930       ✓         yrAspGlyI leProAlaArgGluIleProAspl         ACGAGACCCCGGGAGATCCCCGGGGGCACCCC         980         alSerGluPheSerArgMetAlaArgGspProl         TGTGGATCCCCGCAGGGGCTGCTCCC         1030         yrLeuValProGInGInGINGIVPhePhetyPro/         ATCTGGACCCCCGGCAGCGCCCCCCCCCCCCCCCCCCCC		700 11nG1uThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGGGGCCGCTG 15nG1uThrG1uLeuCa1GGAGCCGCTG 15nG1gG1uAsnVa1Lys1leProVa1 15nGGGGGGAGATGTGGAAAATTCCAGTG 15nG1uThrG1nLeuVa1ThrG1nLeu 15nCACGGGGCAGCTGGGGCACACAGCTT 100 101Ser1leLeuArgArgArgPheThr 15nG1nProPro1le 150Ser1leLeuArgArgArgPheThr 16GTCCATCTCCGCGCGGGGGTCCAC 100S 101Ser1leLeuArgArgArgPheThr 16GTCCATCTCCGCCGCGGGGTCCAC 100S 101Ser1leLeuArgArgArgPheThr 16GTCCATCTCCGCCGCGGGGTCCAC 100S 101Ser1leLeuArgArgArgPheThr 100OS 101Ser1leLeuArgArgArgPheThr 100S 101Ser1leLeuArgArgArgPheThr 1000S 101Ser1leLeuArgArgArgPheThr 1000S 101Ser1leCCCGCCGCGGGGTCCACG 100S 101Ser1GGGCCCAGCCGGCCGCCCCTTG 100S 101Ser1GGGCCCCAGCCAGCCCCCCTTG 1100 110S 110C 110
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451 3601	660         LeuThrSer[ITeValSerAlava]ValGly1leLeu[         1       CTGACGTCCATCGTCTGCGGTGGCATTCGGATGCCATTCGT         710         ThrProSerGlyAlaMetProAsnGlnAlaGIMet/         1       ACACCTAGCGGACGCATGCCCAACCAGGCGCAGATG         710         ThrProSerGlyAlaMetProAsnGlnAlaGIMet/         1       ACACCTAGCGGACGCATGCCCAACAGGCGCAGATG         6       260         AlaIleLySValLeuArgGluÄssnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA         810       810         MetProTyrGlySteLeuLeuAspHisValArgGlu/         1       ATGCCCTATGGCTCCTTTAGACCATGTCCAGGAAA         810       860         ValLeuValLySSerProAsnHisValLySIleThr/         1       GTGCTGGTCAAAGGTCCCAACCATGTCAAAATTACAA         910       HisGlnSerAspValTrpSerTyrGlyValThrValT         1       CACCAGAGTGATGTGGAGTTATGGTGAGAGTATGGCGAGTGAGACTAGGCGACTGGGCGAGTGGAGGCGACTGGAGACTAGGC         1       GTCTACATGGTCAACGTCGCAACTGTGCGAGCCTGGAGACTAGGCG         1       GTCTACAGGAGTGGCGGTGGGGGGGGGGGGGCGCCTGGACCTGGACCATAGGC         1       1010         AspSerThrPheTyrArgSerLeuLeuGluAspAsp/         1       GCCCCCAGGAGGGCGGTGGGGGGGGGGGCCCTGACATAGGC         1       1060         SerThrArgSerGlyGlyGlyGlyAspLeuThrLeuGly1 <t< td=""><td>20 20 20 20 20 20 20 20 20 20</td><td>680         Teleuile, ysArgArgEinelinLys lleArgi         TCCTCATCAAGCGACGCAGCAGAGAGATCCGG         730         TGAAGCGACGCAGCAGAGAGATCCGG         730         allys ValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI         780         alMetAlaGlyValGlySerProTyrValSer/         TGATGGCTGGGTGTGGGCTCCCCATATGTCTCCC         830         euAsnTrpCysMetGInIleAlaLysGlyMet:         TGAACGGGTGTATGCAAGATTGCCAAGGGGATG         880         SpGluThrGluTyrHisAlaAspGlyGlyGlyLys'         ACGAGACAGAGTACCATGCCAGGGACAGCCCGGGACATCCCCGGCAGGGCCCCGGGAGACCCCCTGCCAGCAGGGACTCCCTGCCAGGGGCCCGGGCCCCTGCCCGGGCCCCCTCCCCGAGGGGCCGCCCCCGGGCCCCCCCC</td><td></td><td>700 influThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 150 150 150 150 150 150 150 1</td></t<>	20 20 20 20 20 20 20 20 20 20	680         Teleuile, ysArgArgEinelinLys lleArgi         TCCTCATCAAGCGACGCAGCAGAGAGATCCGG         730         TGAAGCGACGCAGCAGAGAGATCCGG         730         allys ValLeuGlySerGlyAlaPheGlyThri         TGAAGGTGCTTGGATCTGGCGCTTTTGGCACAI         780         alMetAlaGlyValGlySerProTyrValSer/         TGATGGCTGGGTGTGGGCTCCCCATATGTCTCCC         830         euAsnTrpCysMetGInIleAlaLysGlyMet:         TGAACGGGTGTATGCAAGATTGCCAAGGGGATG         880         SpGluThrGluTyrHisAlaAspGlyGlyGlyLys'         ACGAGACAGAGTACCATGCCAGGGACAGCCCGGGACATCCCCGGCAGGGCCCCGGGAGACCCCCTGCCAGCAGGGACTCCCTGCCAGGGGCCCGGGCCCCTGCCCGGGCCCCCTCCCCGAGGGGCCGCCCCCGGGCCCCCCCC		700 influThrGluLeuValGluProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 750 150 150 150 150 150 150 150 1
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451 3601 3751	660         LeuThrSer[ITeValSerAlaValGIy11eLeu]         CTGACGTCCATCGTCTCGCGGTGGTTGGCATTCTGG         710         ThrProSerGIyA1aMetProAsnGInAlaGInMet/         ACACCTAGCGGAGCGCAGATGGCCAACCAGGCGCAGATGG         GCATCAAAGGTTGGGGGAGAGCCCAACCAGCGCCCAAAG         AlaIleLysValLeuArgGluÄsnThrSerProlys/         GCCATCAAAGTGTTGAGGGAAAACACATCCCCCCAAA         810         MetProTyrGlySeLeuLeuAspHisValArgGlu/         ATGCCTATGGCTGCCTCTTAGACCATGCCCGGAA         860         ValLeuValLysSerProAsnHisValLys1leThr/         GGCGTCAAGAGTCCCCACCATGCCGGAAGT         910         HisGInSerAspValTrpSerTyrGlyValThrVaT         1         CACCAGAGGGGCGGTGGAGTATGGTGGAGGTGAGGAGGAGCGGACCATG         910         HisGInSerAspValTrpSerTyrGlyValThrVaT         1         CACCAGAGGGCCCACCATGCGGGGGGGGGGGGGGGGGGG	200 2014 21 Va 1 Va 1 Va 1 Va 1 Va 1 Phe Gi y 1 CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGA 720 Arg I 1 eL eu Lys Gl u Thr G 1 uL eu Arg Lys V CGGATCCTGAAAGAGACGGAGCTGAGGAAGG 770 Al aAsnLys G 1 u I 1 eL eu Asp G 1 uA 1 a Tyr V SCCAACAAAGAAATCTTAGACGAAGCATACG AsnArg G 1 yArg Leu G 1 ySer G 1 n Asp L eu L ACCGCGGACGCCTGGGCTCCCAGGACCTGC Asp Arg I ya g Leu G 1 ySer G 1 n Asp L eu L AACCGCGGACGCCTGGGCTCCCAGGACCTGC 870 Asp Phe G 1 yL eu Al a Arg L eu L eu Asp 1 1 eA SaCTTCGGGCTGGCTCGGCTGCTGGACATTG 920 TrpG 1 u Leu Met Thr Phe G 1 yA 1 Lys Pro 1 TGGGAGCTGATGGCCAGATTCCGGCGCGCTGGGCACCTG 920 TrpG 1 u Leu Met Thr Phe G 1 yA 1 Lys Pro 1 TGGGAGCTGATGGCCAGATTCCGGGCGGCTGCGAAACCTT 970 Ser G 1 u TsArg Pro Arg Phe Arg G 1 LL eu V 1020 Asp Met G 1 yAsp L eu Va 1 Asp A1 aG 1 uG 1 U SaCATGGGGGACCTGGTGGGATGCTGGAGGAT 1020 Ser G 1 u Asp Pro Thr Va 1 Pro L eu Pro Ser G 405 GGCCCCCAGGATCCCCCCAGGATCCCCCCCAGGATCCCCCCCAGGT 1120 Ser G 1 u Asp Pro Thr Va 1 Pro L eu Pro Ser G 405 GGCCCCCTGGAAGGAGCAGGCCCCCCAGGT 1120 Ser G 1 u Asp Pro Thr Va 1 Pro L eu Pro Ser G 405 GGCCCCCCGGCAAGGCCCCCCAGGT 1120 Ser G 1 u Asp Pro Thr Va 1 Pro L eu Pro Ser G 405 GGCCCCCCGGCAAGGCCCCCAGGCCCCCAGGCCCCCCAGGCCCCCC	16Leuile_ysArgArg6inGinLys IIeArgi TCCTCATCAAGCGACGCAGCAGAAGATCCGG         730         alLysValLeuGlySerGlyAlaPheGlyThri TCAAGGTGCTTGGATCTGGCGCTTTTGGCACAI         780         alMetAlaGlyValGlySerProTyrValSer/ TGATGGCTGGTGGGGCTCCCATATGTCTCCG         euAsnTrpCysMetGlnIleAlaLysGlyMet2         TGAACGGCATGCGGTGTGGGCTCCCCATATGTCTCCG         sg0         almetAlaGlyValGlySerProTyrValSer/ TGAACGGCTGGTGTGGGCTCCCATATGTCTCCG         sg0         almetAlaGlyValGlySerProTyrValSer/ TGAACGGGTATGCGGCTCCCATATGTCCCA 830         euAsnTrpCysMetGlnIleAlaLysGlyMet2         SpGluThrGluTyrHisAlaAspGlyGlyLys' ACGAGACAGAGTACCATGCAGATGGGGGCGAGG 930         yrAspGlyIleProAlaArgGluIPeroAspA ACGAGGACCCAGCCGGAGACTCCCTGAC 1030         yrLeuValProGlnGlnGlyPhePheCyProCA 1030         yrLeuValProGlnGlnGlyPhePheCyProCA 1030         yrLeuValProGSrGLIGIYJNAJAIAGIYSerC 1030         uThrAspGlyTyrValAlaProLeuThrCyS AGACTGGCACCCTCCCGAGGGGCGCCCCCA 1130         1uThrAspGlyTyrValAlaProLeuThrCyS AGACTGAGCAGCAGAATGGGGCTGCTCCA 1180         euSerProGlyLySASnGIVValValLySASp7 TCTCCCCAGGGAACATGGGGCTCCAAGACC 1230         rpAspGlnAspProProGluArgGlyAlaProI GGGACCAGGACCCCACCACCAGAGCGGGGGGCTCCA	690 ys Jyr ThrMetArgArgLeuLeuG AAGTACACGATGCGGAGACTGCTGG 740 /a 11 yr Lys G1 y1 le Trp 11 eProA SICTACAAGGGCATCTGGACCCCTG 790 ArgLeuLeuG1 y1 le Tre 11 eProA SicCTTCTGGGCATC SICCTGACAT 840 Ser Tyr LeuG1uAspVa1ArgLeuV VaCCTACCTGGAGGATGGGCGGCGCG 940 241 Pro11 eLys TrpMetA1 aLeuG 240 241 Pro11 eLys TrpMetA1 aLeuG 240 241 Pro11 eLys TrpMetA1 aLeuG 240 241 Pro11 eLys TrpMetA1 aLeuG 240 241 Pro11 eLys TrpMetA1 aLeuG 240 250 250 250 250 250 250 250 25	700 31nG1UThrG1uLeuValG1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 35pG1yG1uAsnValLys1leProVal ATGGGGAGAATGTGAAAATTCCAGTG 300 301 SerThrVa1G1nLeuVa1ThrG1nLeu 302 303 SerThrVa1G1nLeuVa1ThrG1nLeu 303 SerThrVa1G1nLeuVa1ThrG1nLeu 304 314 SargAspLeuA1aA1aArgAsn 314 CACAGGGACTGGCCGCCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 305 SerThrVa1G1nLeuCCGCGCGGCGTCCACC 310 SerIleLeuArgArgArgPheThr 305 SerThrVa1G1nLeuCCGCGCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 305 SerThrVa1G1nLeuCCGCGCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 306 SerTCGGTCCACCCGGCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 306 SerTGGGCCCAGCCGCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 306 SerTGGGCCCAGCCGGCGGTCCACC 310 SerIleLeuArgArgArgPheThr 310 SerTGGGCCCAGCCGCGGCGGTCCCCTTG 310 SerTGGGCCCCAGCCCAGCCCCTTG 310 SerTGGGCCCCAGCCCAGCCCCCTTG 310 SerTGGGCCCCAGCCCAGCCCCCTTG 310 SerTGGGCCCCAGCCCAGCCCCCCT 310 SerTGGGCCCAGCCCAGCCCCCCCT 310 SerTGGIAJAIAA1aLySG1yLeuG1nSer 310 SerTGGIAPPOS1AFGProS1nSer 310 SerTGGIAPPOS1AFGProS1nSer 310 SerTGGGCCCAGCCCAGCCCCCCCCCCCCCCCCCCCCCCCCC
<ul> <li>2101</li> <li>2251</li> <li>2401</li> <li>2551</li> <li>2701</li> <li>2851</li> <li>3001</li> <li>3151</li> <li>3301</li> <li>3451</li> <li>3601</li> <li>3751</li> </ul>	660         LeuThrSer[ITeValSerA1aVa)ValG1y11eLeu[         1         CTGACGTCCATCGTCTCGCGGTGGGTTGGCATTCTGG         710         ThrProSerG1yA1aMetProAsnG1nA1aG1nMet/         ACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGG         ALaIleLysValLeuArgG1uÄsnThrSerProLys/         GCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA         810         MetProTyrG1yG1stLeuLeuAspHisVa1ArgG1u/         AGCCCTATGGCTCCTCTAGACCATGCCCGGAAV         860         ValLeuVa1LysSerProAsnHisVa1Lys11eThr/         1       GTGCTGGTCAAGGGTCCCAACCATGCTCGGGAGA         860         ValLeuVa1LysSerProAsnHisVa1Lys11eThr/         1       GTGCTGGTCAAGAGTCCCAACCATGGTCGAGAGTATGGTGGACGTTAGGTGGAGTTAGGTGGAGTTAGGTGGAGTTAGGTGGAGTTATGGTGG	20 20 20 20 20 20 20 20 20 20	680 1eLeuile_ysArgArgEineGinLysIleArg TCCTCATCAAGCGACGCAGCAGAAGATCCGG 730 alLysValLeuGlySerGlyAlaPheGlyThri TGAAGGTGCTTGGATCTGCGCGTTTTGGCACAI 80 alMetAlaGlyValGlySerProTyrValSerr TGATGGCTGGTGTGGGCCCCCATATGTCCC 830 euAsnTrpCySMetGInIleAlaLysGlyMet: TGAACTGGTATGCAGATTGCCAAGGGATG 880 spGluThrGluTyrHisAlaAspGlyGlyLys' ACGAGACAGAGTACCATGCAGGATGGGGGCAGG 930 yrAspGlyIleProAlaArgGluIleProAspI ACGAGGACCAGGTCCCGACTGCGGGGATG 930 yrAspGlyIleProAlaArgGluIleProAspI ACGAGGGCACGAGTCCCGGCAGGGCCCGG 980 alSerGluPheSerArgMetAlaArgAspProi TGTCTGGATTCCCCGCAGGGCCAGGGCCCCG 1030 yrLeuValProGInGInGInGlyPhePheCysProJ ATCTGGTACCCCCGCCCCCCCCCCCCCCCCCCCCCCCCC		700 31nG1UThrG1uLeuVa1G1uProLeu AGGAAACGGAGCTGGTGGAGCCGCTG 37GGGGGAATGTGAAAATTCCAGTG 3ATGGGGGAGAATGTGAAAATTCCAGTG 300 3erThrVa1G1nLeuVa1ThrG1nLeu 7CCACGGTGCAGCTGGTGGACACAGCTT 100 31USerIleLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTCCACC 900 31USerIleLeuArgArgArgPheThr AGTCCATTCTCCGCCGGCGGTTCACC 900 31USerIleLeuArgArgArgPheThr 1000 31UASPLEUG1yProAlaSerProLeu AGGACCTGGGCCGCCGCGGCGTCCACC 1000 100StuAspleuG1yProAlaSerProLeu 1050 104S0 104S0 1050
2101 2251 2401 2551 2701 2851 3001 3151 3301 3451 3601 3751 3901	660         LeuThrSer[ITeValSerA1aValFiy11eLeu]         1         CTGACGTCCATCGTCTCGCGGTGGTGGCATTCTGG         710         ThrProSerG1yA1aMetProAsnG1nA1aG1nMet/         1       ACACCTAGCGGAGCGATGCCCAACCAGCGCGCAGATGG         ACACCTAGCGGAGCGATGCCCAACCAGCGCCGAGATGG         AL11eLysVa1LeuArgG1uÅsnThrSerProLys/         1       GCCATCAAAGTGTTGAGGGAAAACACATCCCCCCAAA         810         MetProTyrG1yCsLeuLeuAspHisVa1ArgG1u/         1       ATGCCTATGGCTCCTCTAGACCATGCCCGGAAV         860       Va1LeuVa1LySSerProAsnHisVa1LyS11eThr/         1       GTGCTGGTCAAGAGTCCCAACCATGCCCAACATTACAG         910       HisG1nSerAspVa1TrpSerTyrG1yVa1ThrVa1         1       CACCAGAGGTCACATGGTGGAGGTTATGGTGGACTGTG         910       HisG1nSerAspVa1TrpSerTyrG1yVa1ThrVa1         1       CACCAGAGGTCACATGGTCGAACATGGCGAGGACGTT         1       GCTCTCACAGGAGTGCCCACCCTCACCTGCCGAGCACTGGCGACGTG         1010       AspSerThrPheTyrArgSerLeuLeuG1uAspAsp/         1       GACAGCACCTTCTACCGCTCACCTGCCGACACTAGGGG         1       TCTACCAGGAGTGGCGGGGGGGGGCACCTGACACTAGGGCACCT         110       LeuProThrHisAspProSerProLeuG1nArgTyr5         1       CTCCCCACACATGACCCCAGCCCTCTGCCTGCCTGCCCGCCGCCGCGCCGCGCCCCGCCCCCGCCCCGCCCCGCCCC	C70     C	680 16Leuile ysArgArgEineGinLys IleArg TCCTCATCAAGCGACGACAGCAGAAGATCCGG 730 allysValLeuGlySerGlyAlaPheGlyThr TGAAGGTGCTTGGATCTGCGCGTTTTGGCACA almetAlaGlyValGlySerProTyrValSer TGATGGCTGGTGTGGGCTCCCCATATGTCTCC 830 euAsnTrpCySMetGlnIleAlaLysGlyMet: TGAACTGGTATGCAGATTGCCAAGGGATG euAsnTrpCySMetGlnIleAlaLysGlyMet: TGAACTGGTATGCAGATTGCCAAGGGATG 880 290 yrAspGlyIleProAlaArgGluIleProAspl ACGAGGACAGAGTACCCAGCAGGGATCCCCGGAGCCCCCAC 980 alSerGluPheSerArgMetAlaArgGspProI TGTCTGGATTCTCCCGCATGGCGGGCACCCC 1030 yrLeuValProGlnGlnGlgYPhePheCySProJ ATCTGGTACCCCAGCCGGGGCTCCCC 1130 uThrAspGlyTyrValAlaProLeuThrCys AGACTGATGGCACCCTCCGAAGGGCTGCCCC 1180 euSerProGlyLysAsnGJVAlValValPaProJ TGCCCCAGGGAACCCACCAGAGGGCTCCCC 1180 euSerProGlyLysAsnGJVAlValValPaProJ GGGACCAGGACCCACCAGAGGGCTCCAAGAGCC 230 230 240 250 250 250 250 250 250 250 25		700 in Giu Thr Giu Leu Val Giu ProLeu AGGAAACGGAACTGGTGGGGGCCGCTGC isp Giy Giu Asn Val Lys I le ProVal ATGGGGAGAATGTGAAAATTCCAGTG 800 Ser Thr Val Gin Leu Val Thr Gin Leu (CACGGTGCAGCTGGTGGACACAGCTT 850 'al His Arg Asp Leu Al al a Arg Asn TACACAGGGACTTGGCCGCTCGGAAC 900 il USer I le Leu Arg Arg Arg Phe Thr 1AGTCCATTCTCCGCCGCGGGGTCCACC 900 10 LSer I le Leu Arg Arg Arg Phe Thr 10 AGTCCATTCTCCGCCGCGGGGTCCACC 900 10 LSer I le Leu Arg Arg Arg Phe Thr 10 AGTCCATTCTCCGCCGCGGGGTCCACC 900 10 LSer I le Leu Arg Arg Arg Phe Thr 10 AGTCCATTCTCCGCCGCCGGGGTCCACC 900 10 LSer I le Leu Arg Arg Arg Phe Thr 10 AGTCCATTCTCCGCCCGCGGCGTCCCCGGAC 900 10 LSER I Leu Arg Arg Arg Phe Thr 10 AGTCCATTCTCCGCCCGCGGCGTCCCCC 10 Stu Asp Leu Gi y ProAl a Ser Proleu 10 AGGCCAGACCCAGGGGCTGCAAGCC 11 Asp Leu Gi y ProAl a Ser Proleu 10 AGGCCAGAGCCAAGGGGCTGCAAGCC 11 Asp ProGlu Tyr Leu Thr ProGin AGAACCCCGGAGTACTTGACACCCCAG 11 Asp ProGlu Tyr Leu Thr ProGin AGAACCCCGAGTACTTGACACCCCAG 12 S0 hr Al a Giu Asp ProGlu Tyr Leu Giy 10 CGGCAGAGAACCCAGGCAGGCAGCAGCCAGCCAGCCAGC

Fig. 3. Complete nucleotide and amino acid sequences of HER2 (clone  $\lambda$ HER2-436). Synthetic probes 1 and 2 (Fig. 1) were used to screen 2  $\times$  10<sup>6</sup> Fig. 3. Complete hubble hubbl numbered starting with the initiation methionine (1). Underlining indicates the putative signal sequence (heavy) and the potential poly(A) addition signal (fine). Lines on top of the sequence indicate potential glycosylation sites, the black arrow demarcates the EGF receptor threonine 654 nalogue (Thr 686), shading indicates cysteine residues, boxing shows the putative transmembrane region, large open triangles indicate some ocations of introns in the HER2 gene, and the small triangle emphasizes a possible tyrosine autophosphorylation site by homology to Tyr 1173 in he EGF receptor sequence.

**DECEMBER 1985** 

from a single fetus (Fig. 2a). Two mRNA's, of 5.4 and 6.4 kb, were also detected in term placenta. No cross-hybridization with the 5.8-kb and 10.5-kb

EGF receptor mRNA's in term placenta mRNA was observed under these stringent hyridization conditions (legend, Fig. 2). Normal adult human tissues,



including kidney, liver, skin, lung, jejunum, uterus, stomach, and colon, contained lower but significant amounts of the same 4.8-kb mRNA. Because of the magnitude of fetal expression, we also examined several embryonic tumors (Fig. 2b); each expressed large amounts of the 4.8-kb transcript, although not more than that detected in normal fetal tissue.

Thus, it appears that the HER2 gene is widely expressed, in both normal adult tissues and in several normal fetal tissues. While detected in most embryonic tumors, the HER2 gene was not present at higher levels than in fetal tissues; thus, the particular level may reflect the state of differentiation of a given tumor.

HER2 structurally characterized as cell surface receptor. Using the transforming gene of the avian erythroblastosis virus, v-erbB, as a hybridization probe, we isolated genomic and cDNA sequences of an uncharacterized human gene. The 1255 amino acid polypeptide sequence

HER2	-24	MELAAL&RWGLLALLPPGAASTQV&TGTDMKLRLPASPETHLDMLRHLYQG&QVVQGNLELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVP	1
HER1 HER2	77 96	EN Q I NMYY NS SYA-K	
HER1	167	D SËPNGSË AGE NË K KIIË QQËSG ER KS S N RE VËRK RDEATËKDTË P ML PT YQMDV K S	2
HER2	196	SPMEKGSREWGESSEDEQSLTRTVEAGGEA-REKGPLPTDEEHEQEAAGETGPKHSDELAELHFNHSGIEELHEPALVTYNTDTFESMPNPEGRYTFGAS	
HER1	267	KK R VV H VRA GADSY ME- VRK EG RK N I IGEFKDSLSINAT KH KN TS S D HI VA R SFTH P D	
HER2	295	EVTAEPYNYLSTDVGSETLVEPLHNQEVTAEDGTQREEKESKPEARVEYGLGMEHLREVRAVTSANIQEFAGEKKIFGSLAFLPESFDGDPASNTAPLQP	
HER1	366	QE DILK VK F L Q ENRT HA E EI TKQH QF AVVS N TS K ISD DVI SG KN YAN IN KK GTSG KTK	_
HER2	395	EQLQVFETLEEITGYLYISAWPDSLPDLSVFQNLQVIRGRILHNGAYSLTLQGLGISWLGLRSLRELGSGLALIHHNTHLEFVHTVPWDQLFRNPHQALL	
HER1	466	IIS G NSKAT QV A SPEGEW PE RD SRNVS R DKENLE E FENSEIQ L AMNI TRGP NIQ I G	3
HER2	495	HTANRPEDEEVGEGLAEHQLEARRALLGSGPTQEVNESQFLRGQEEVEEERVLQGLPREYVNARHELPEHPEEQPQNGSVTEFGPEADQEVAEAHYKDPP	
erbB	1	MK FI G	
HER1	566	H KT A MGENNTLV- YA AGHV HL HP YG TGPGLE TNGPKI ATGM A LLLVVALGIG FM RHIV RL R	4/5
HER2	595	FEVAREPSGVKPDLSYMPIWKFPDEEGACOPEPINETHSEVDLDDKGEPAEORASPLTSIVSAVVGILLVVVLGVVFGILIKRROOKIRKYMRRLLQET	
erbB	11	H KA A LGENDTLVR- YA ANAV L HP RG KGPGLE NGSKT AAG G CL VVGLGIG YL RHIV R R	
HER1	663	EA LL FK I <b>* * *</b> L E K <b>*</b> E A S DN H C	
HER2	695	ELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSPYVSRLLGICLTSTV	
erbB	107	EA HL FK I L E K E A S DN H C	
HER1	763	I F Y HKDNI Y V N R T Q K GAE K E	6
HER2	795	QLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVRLVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESI	
erbB	207	I YI HKDNI Y V N ER T Q K GA K E	
HER1	863	H IY S S SSI ADS K II K YL GDER	
HER2	895	LRRRFTHQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWMIDSECRPRFRELVSEFSRMARDPQRFVVIQNEDL	
erbB	307	H IY S S SV ADS K IA K P YL GDER	
HER1 HER2 erbB	963 995 407	MHLP       T       N-       A       MDEE       D       I<	
HER1	1028	RN C IKED F S GA -T DSIDDTF V I -S PKR ÁGS-VQNPVYHNQPLNPAPSRD HYQDPHST G	•
HER2	1093	GAAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGWVAPLTCSPQPEYVNQPDVRPQPPSPREGPLPAARPAGATLERAKTLSPGKNGVVKDVFAFGGAVE	
erbB	472	RN - GH VRED FV S GNF-L ESIDDGF A V -LMPKK STAMVQNQIYNFISLTAISKLPM-DSRYQNSHST D	
HER1	1111	N V PTCVNSTFD HWAQKGSHQISL PDY-Q FF KEAK NGI S- A RVAPQSSEFIGA*	
HER2	1193	NPEYL-TPQGGAAPQPHPPPAFSPAFDNLYYWQQD-PPERGAPPSTFKGTPTAENPEYLGLDVPV*	
erbB	557	N N SPLAKTVFES -YWIQSGNHQINL PDY-Q FL TSCS*	
Fig. 4.	(a) Hydro	opathy analysis (20) of HER2 (1255 amino acids) coding sequences and comparison with EGF receptor (HER1; 1210 amino acids)	ids).

Fig. 4. (a) Hydropathy analysis (20) of HER2 (1255 amino acids) coding sequences and comparison with EGF receptor (HER1; 1210 amino acids). Different receptor domains and the extent of amino acids sequence homology are indicated. The autophosphorylation tyrosines in the EGF receptor sequence are shown, as is a potential analogue in HER2. The signal sequence is shown by fine shading, the cysteine-rich subdomains by hatching, the transmembrane region by a black bar, and the tyrosine kinase domain by coarse shading. (b) Comparison of HER2, EGF receptor (HER1), and v-erbB amino acid sequences. Identical residues are deleted, gaps are introduced to optimize alignment, cysteine residues are shaded, and carboxyl terminal tyrosines are in shaded boxes. The black and open triangles indicate the positions of the major and minor autophosphorylation sites in the EGF receptor, respectively. Asterisks indicate residues involved in ATP binding. Boxed regions include (i), signal sequence; (ii, iii), cysteine-rich domains; (iv), transmembrane domain; (v), protein kinase C modulation domain; (vi), tyrosine kinase domain; (vii), signal transfer domain.

derived from this cloned cDNA shows extensive homology to v-erbB and its cellular homologue, the human EGF receptor, and was therefore termed HER2. Its primary amino acid sequence displays all the structural features known to define a cell surface receptor for a polypeptide ligand. This finding represents the isolation and detailed structural characterization of a putative cell surface receptor before determination of its specific ligand, based solely on its structural homology to another receptor gene. In a similar case, the v-fms gene (26) was characterized because of its oncogenic activity, and was later determined to be identical or related to the receptor for CSF-1, a hematopoietic growth factor (4).

Examination of the HER2 primary sequence reveals extensive collinearity with the EGF receptor sequence. Detailed sequence comparison of these two closely related receptors, each having presumably different biological roles, provides important clues to the functional significance of the different regions of the proteins. A hydrophobic amino terminal sequence of HER2 presumably functions as a signal sequence (Fig. 4b, region 1) that, like the pre-EGF receptor, would be cleaved after alanine (22) to generate an amino terminal serine; however, cleavage after glycine (19) may occur, since signal peptidase frequently cleaves after small amino acid residues. Again, as in the case of the EGF receptor, the next 632 amino acids would represent the extracellular domain of the HER2 receptor that forms the binding pocket for a specific polypeptide ligand. Some structural features in this domain are highly conserved, while others have diverged extensively. Of the 51 cysteine residues of the EGF receptor ligand binding domain, 50 are conserved in HER2, of which 21 and 26, respectively, are clustered in subdomains 2 and 3 (Fig. 4b). Sequences between these conserved, regularly spaced cysteine residues, as well as the ones flanking subdomains 2 and 3 (Fig. 4b), have diverged to result in an overall homology of 40.6 percent (Fig. 4a).

Cysteine rich clusters (27) are found in receptors for EGF (9), low-density lipoprotein (28), and insulin (19), as well as in the EGF precursor (29), which is a potential receptor (Fig. 6). Since cysteine-rich domains are found in a family of cell surface receptors and potential receptors with very diverse functions, they are likely to form an essential structural backbone of the ligand binding domain, but not define ligand specificity. The ligand affinity of each of these receptors

may be specified by sequences flanking these cysteine residue clusters.

The presence of regularly repeated, extracellular cysteine clusters (27) defines a gene family that overlaps with another gene family, the membranespanning proteins that use an intrinsic tyrosine-specific kinase activity in the signal transduction process (5) (Fig. 6). v-fms, the oncogenic homologue of the M-CSF-1 (colony-stimulating factor) receptor, contains a tyrosine kinase domain (26), but lacks cysteine-rich clusters in its extracellular domain, and thus is only a member of this second gene family.

Only 5 of the 12 potential N-linked glycosylation sites of the EGF receptor are conserved with respect to their approximate position in HER2, which also contains only eight such sites in this extracellular domain. The role of carbohydrate side chain differences in defining the biological function of these proteins requires further investigation.

The cytoplasmic domains of EGF receptor and HER2 also exhibit striking regional differences in sequence homology. These differences (Fig. 4, a and b) most likely distinguish regions that define receptor-specific, ligand-induced signal generation from support functions such as ATP binding and enzymatic tyrosine kinase activity.

Directly adjacent to the putative transmembrane domain, a cluster of 14 predominantly basic amino acids is found in both HER2 and EGF receptor sequences, which may interact with the membrane phospholipid head groups. The Thr<sup>654</sup> residue, which is located within this region of the EGF receptor, is a phosphorylation substrate site for receptor activity down-modulation by protein kinase C (22). The HER2 sequence contains a threonine residue at an analogous position (686) (Figs. 3 and 4a and region 5, Fig. 4b) and may therefore also be subject to protein kinase C modulation.



bution along chromosome 17 [ideogram from International System for Human Cytogenetic Nomenclature 1981 (35)]. Solid line, HER2-1 (1620-bp Eco RI fragment of AHER2-436 containing coding sequences from extracellular, transmembrane and tyrosine kinase domains); dashed line, HER2-2 (1472-bp Eco RI fragment with carboxyl terminal coding and 3'untranslated sequences). Human metaphase



and prometaphase chromosomes were prepared from methotrexate-synchronized peripheral lymphocyte cultures (36) of two normal individuals. The probes for in situ hybridization (37) were <sup>3</sup>H-labeled by nick translation with three labeled nucleotides to specific activities of  $2.4 \times 10^7$  cpm/µg (HER2-1),  $3.4 \times 10^7$  cpm/µg (HER2-2), and  $2.5 \times 10^7$  cpm/µg (HER1 64-3); the probes (25 or 50 ng/ml) were hybridized to chromosome preparations (16 hours at 37°C). The emulsion-coated slides were exposed for 10 to 14 days at 4°C. Chromosomes were stained with Wright's stain for G-banding, and a second photograph was taken of the previously selected cells. G-banded chromosomes were analyzed for silver grain localizations. (b) Pairs of chromosome 17 from three cells, illustrating typical labeling at bands  $q12 \rightarrow q22$  for HER2-1 (top row) and HER2-2 (bottom row). (c) Southern blot hybridization of HER2-1 (a) and HER2-2 (b) to Eco RI-digested DNA from somatic cell hybrids of humans and rodents. Somatic cell hybrids of four different series were used. Their origin and characterization has been reported (38). Samples (10 µg) of hybrid, mouse, Chinese hamster, and human DNA were cleaved with Eco RI, subjected to electrophoresis on 0.8 percent agarose gels, transferred to nitrocellulose filters, and hybridized to <sup>32</sup>P-labeled HER2 probes according to standard procedures (26, 39). (Lane 1) Human lymphoblastoid control cell line; (lane 2) mouse 3T3 cells; (lane 3) Chinese hamster cell line; (lane 4) subclone of hybrid of mouse and human cells with a rearranged chromosome 17 containing only region q21->qter, retained by growth in hypoxanthine-amnopterin-thymidine (HAT) medium; (lane 5) same subclone as lane 4, except missing this rearranged chromosome after counterselection in bromodeoxyuridine medium; (lanes 6, 8, 10, and 12) clones of Chinese hamster and human hybrid cells containing human chromosome 17; (lanes 7, 9, and 11) Chinese hamster X human hybrid cell clones lacking human chromosome 17.



The subsequent 341 residues are not only highly homologous to the EGF receptor and the v-erbB oncogene, but also to all the other members of the src family of tyrosine kinase oncogenes (22), as well as the  $\beta$  subunit of the insulin receptor (19). This sequence similarity suggests that this HER2 region encompasses at least a portion of the kinase domain that may be activated by extracellular ligand binding. A Gly X Gly XX Gly sequence at HER2 position 727 and a lysine at position 753 are known from studies on other members of this gene family to be involved in ATP binding (23). Overall, this HER2 region (6 in Fig. 4b) is the most homologous to v-erbB and EGF receptor (78.4 percent; Fig. 4a), which explains why the initially isolated genomic clone (Fig. 1) contained exons coding for part of this tyrosine kinase domain (intron positions are shown in Fig. 3).

The region within the sequence of a tyrosine kinase receptor most likely to be responsible for translation of the activation signals intitiated by specific, extracellular ligand binding into physiological action is localized at the carboxyl terminus of the receptor polypeptide chain. This region contains the 32-amino acid sequence deletion that plays a crucial role in the generation of the AEV transforming gene v-erbB from the EGF receptor proto-oncogene (Fig. 6) (30), as well as the tyrosine residues that are major (1173) and minor (1068 and 1148) sites of ligand-induced autophosphorylation in the EGF receptor (24) (Fig. 4, a and b). This region contains the most extensive differences in length and primary structure between HER2 and EGF receptor/erbB. However, five of seven tyrosines are conserved in HER2, including short surrounding sequences. This suggests that, like the conserved cysteine residues in the ligand binding

domain, these tyrosine residues may serve an auxiliary activity-control function in the generation of a cytoplasmic signal. Other sequences in this signalcontrol and transfer domain are likely to define specificity, which may include recognition of possible substrates and/or cofactors. Future work will be needed to determine if tyrosine 1248 of HER2 plays a role analogous to tyrosine 1173, the major autophosphorylation site of the EGF receptor, and if carboxyl terminal truncation may lead to the generation of a HER2-derived oncogenic polypeptide.

pocket

EGFP.

Fig. 6. Schematic comparison

of HER2-related molecules.

Hatched boxes indicate cyste-

ine-rich regions within extra-

cellular or putatively extracel-

lular domains. The shaded re-

gion below the membrane bi-

layer indicates the location of

the tyrosine kinase domain.

Black circles within the human

insulin receptor (HIR) and v-

fms structures indicate cysteine residues that are not in-

cluded in cysteine clusters and

may be involved in heterote-

tramer (HIR) or ligand binding

formation

LDLR, LDL receptor; HER1,

EGF

(v-fms).

precursor;

The oncogenic potential of the HER2 gene is further suggested by its chromosomal location. Results of in situ and Southern blot hybridizations enabled us to assign the HER2 gene to chromosome 17 (17q21-q22), precisely coinciding with our previous assignment of the neu oncogene (12), a transforming sequence isolated from rat neuroblastoma (11). As documented elsewhere (12), our results with the rat neu probe were virtually identical to those obtained with the two HER2 probes described here, except for the lack of cross-hybridization of rat neu genomic sequence with the EGF receptor locus on chromosome 7. Our mapping results support the notion that the proto-oncogene neu and the HER2 gene, although both are very likely not distinct from each other. Further detailed analysis will be necessary to establish whether they actually represent the same gene.

Can we obtain clues to the cellular role of this putative receptor by examination of its gene expression? Northern blot hybridization analysis of poly(A)<sup>+</sup> RNA's from a variety of normal and malignant tissues implicates a role for the HER2 receptor-ligand system in many cell types in both fetal and adult life. Highest levels of the 4.8-kb mRNA were found in several fetal tissues, which rules out a cell lineage-specific receptor

for a growth or differentiation factor and is consistent with a more general role in fetal development. No clues were obtained from tumor RNA screening to suggest that increased expression was involved in tumor generation or growth. In some tissues, such as term placenta, fetal kidney, and embryonic tumors, we found additional larger (5.5, 6.4, and 7.4 kb) mRNA's hybridizing with our probe, suggesting transcriptional diversity or the presence of homologous mRNA's more closely related to HER2 than to the EGF receptor. Further experiments should establish the biological role of the HER2 gene, and its role in oncogenesis.

## **References and Notes**

- G. Guroff, Ed., Growth and Maturation Factors (Wiley, New York, 1983), vol. 1.
   M. D. Waterfield et al., Nature (London) 304, 35 (1983); R. F. Doolittle et al., Science 221, 275 (1983)
- 1983) 3. J. Downward et al., Nature (London) 307, 521
- (1984).
   C. J. Sherr et al., Cell 41, 665 (1985).
   T. Hunter and J. A. Cooper, Annu. Rev. Biochem. 54, 897 (1985).
- chem. 54, 897 (1985).
  6. M. Kasuga, F. A. Karlsson, C. R. Kahn, Science 215, 185 (1982); M. Kasuga et al., J. Biol. Chem. 257, 9891 (1982); M. Kasuga et al., J. Biol. Chem. 257, 9891 (1982); M. Kasuga, Y. Zick, D. L. Blithe, M. Crettaz, C. R. Kahn, Nature (London) 298, 667 (1982); Y. Zick, M. Kasuga, C. R. Kahn, J. Roth, J. Biol. Chem. 258, 75 (1983); L. M. Petruzzeli et al., Proc. Natl. Acad. Sci. U.S.A. 79, 6792 (1982); M. A. Shia and P. F. Pilch, Biochemistry 22, 717 (1983); R. A. Roth and D. J. Cassell, Science 219, 299 (1983); E. Van Obberghen, B. Rossi, A. Kowalski, H. Gazzano, G. Ponzio, Proc. Natl. Acad. Sci. U.S.A. 80, 945 (1983); J. Avruch, R. A. Nemenoff, P. J. Blackshear, M. W. Pierce, R. Osathanondh, J. Biol. Chem. 257, 15162 R. Osathanondh, J. Biol. Chem. 257, 15162 1982
- (1962).
  B. Ek, B. Westermark, A. Wasteson, C.-H. Heldin, Nature (London) 295, 419 (1982).
  S. Jacobs et al., J. Biol. Chem. 258, 9581 (1983);
  J. B. Rubin, M. A. Shia, P. F. Pilch, Nature (London) 305, 438 (1983).
- D. Rubini, M. A. Sina, F. F. Filch, *Value* (London) 305, 438 (1983).
   W. Weber et al., Science 224, 294 (1984); G. T. Merlino et al., *ibid.*, p. 417; A. Ullrich et al., *Nature (London)* 309, 418 (1984).
   T. Yamamoto et al., Cell 35, 71 (1983).
   A. L. Schechter et al., *Nature (London)* 312, 513 (1984); C. Shih, L. C. Padhy, M. Murray, R. A. Weinberg, *ibid.* 290, 261 (1981).
   A. L. Schechter et al., *Science* 229, 976 (1985).
   B. Vennstrom, L. Fanshier, C. Moscovici, J. M. Bishop, J. Virol. 36, 575 (1980).
   R. M. Lawn, E. F. Fritsch, R. C. Parker, G. B. Lake, T. Maniatis, Cell 15, 1157 (1978).
   H. Lehrach, D. Diamond, J. M. Wozney, H. Boedtker, Biochemistry 16, 4743 (1977).
   R. Crea and T. Horn, Nucleic Acids Res. 8, 2331 (1980).

- 17. T. Huynh, R. Young, R. Davis, in Practical Approaches in Biochemistry, D. Grover, Ed. (IRL, Oxford, 1984).
- M. Kozak, Nucleic Acids Res. 9, 5233 (1981) Ullrich et al., Nature (London) 313, 756 19. Α.
- (1985). 20. J. Kyte and R. F. Doolittle, J. Mol. Biol. 157, 105 (1982).
- 21. M. O. Dayhoff, Atlas of Protein Sequence and
- *Structure 5*, supplement 3 (National Biomedical Research Foundation, Washington, D.C., 1978). The one-letter abbreviations recommended by the IUPAC-IUB Commission on Biochemical Nomenclature are used: E, glutamic acid; I, isoleucine; K, lysine; L, leucine; M, methio-nine; N, asparagine; P, proline; Q, glutamine; R, arginine; T, threonine; V, valine; and Y, tyro-
- 22.
- sine.
  T. Hunter, N. Ling, J. A. Cooper, Nature (London) 311, 480 (1984).
  M. L. Privalsky, R. Ralston, J. M. Bishop, Proc. Natl. Acad. Sci. U.S.A. 81, 704 (1984); W. C. Barker and M. O. Dayhoff, *ibid.* 79, 2836 (1982); W. Moller and R. Amons, FEBS Letts.
  186, 1 (1985); M. J. E. Sternberg and W. R. Taylor, *ibid.* 175, 387 (1984). 23.

SCIENCE, VOL. 230

- 24. J. Downward, P. Parker, M. D. Waterfield, Nature (London) 311, 483 (1984).
- Nature (London) **311**, 483 (1984).
  25. E. Southern, Methods Enzymol. **68**, 152 (1979).
  26. S. J. Anderson et al., J. Virol. **44**, 696 (1982); R. Manger et al., Cell **39**, 327 (1984); C. W. Rettenmier et al., ibid. **40**, 971 (1985); M. F. Roussel et al., Mol. Cell. Biol. **4**, 1999 (1984).
  27. S. Pfeffer and A. Ullrich, Nature (London) **313**, 184 (1985)
- 184 (1985).
   T. Yamamoto *et al.*, *Cell* **39**, 27 (1984).
   A. Gray, T. J. Dull, A. Ullrich, *Nature (London)* **303**, 722 (1983); J. Scott *et al.*, *Science* **221**, 236 29
- 1983) 30. H. Riedel and A. Ullrich, unpublished data.
- A. Ullrich, C. H. Berman, T. J. Dull, A. Gray, J. M. Lee, *EMBO J.* 3, 361 (1984).
   F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463 (1977); J.

Messing and J. Vieira, Gene 19, 269 (1982); J. Messing, R. Crea, P. H. Seeburg, Nucleic Acids Res. 9, 309 (1981).
33. G. Cathala et al., DNA 2, 329 (1983).
34. J. M. Taylor, R. Illmensee, J. Summers, Biochim. Biophys. Acta 4, 324 (1976).
35. ISCN, "An international system for human cytogenetic nomenclature—high resolution banding," Cytogenet. Cell Genet. 31, 1 (1981).
36. J. J. Yunis, Science 191, 1268 (1976).
37. M. Harper and G. F. Saunders, Chromosoma

- 37.
- J. J. Yunis, Science 191, 1268 (1976).
  M. E. Harper and G. F. Saunders, Chromosoma 83, 431 (1981).
  U. Francke and B. de Martinville, Banbury Report 14, 175 (1983); U. Francke and M. A. Pellegrino, Proc. Natl. Acad. Sci. U.S.A. 7, 1147 (1977); U. Francke, Cytogenet. Cell Genet. 38, 298 (1984); \_\_\_\_\_\_ and B. Francke, Somat. Cell Genet. 7, 171 (1981). 38.
- G. M. Wahl et al., Proc. Natl. Acad. Sci. U.S.A. 76, 3683 (1979).
   C. R. King, M. H. Kraus, S. A. Aaronson, Science 229, 974 (1985).
- 41. We thank D. Smith, who was involved in the early stages of cDNA sequencing, W. Cavenee and X. Konfos for providing embryonic tumors and X. Breakefield for providing fetal tissues. We also thank M. Wechser and A. Douglas for technical assistance, J. Arch for preparation of the manuscript, and S. Pfeffer for discussions and help during the preparation of the manuand help during the preparation of the manuscript. After completion of this work, King *et al.* (40) published a partial sequence of an EGF receptor-related gene that matches our HER2 sequence perfectly.

20 September 1985; accepted 30 October 1985

## **AAAS-Newcomb** Cleveland Prize To Be Awarded for an Article or a Report Published in Science

The AAAS-Newcomb Cleveland Prize is awarded annually to the author of an outstanding paper published in Science. The 1985 competition starts with the 4 January 1985 issue of Science and ends with the issue of 27 December 1985. The value of the prize is \$5000; the winner also receives a bronze medal.

Reports and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the year, readers are invited to nominate papers appearing in the Reports or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS-Newcomb Cleveland Prize, AAAS, 1333 H Street, NW, Washington, D.C. 20005. Final selection will rest with a panel of distinguished scientists appointed by the editor of Science.

The award will be presented at a session of the AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.