

Sequences - Corrections of TERT protein sequence models

The model sequences of chimpanzee (*Pan troglodytes*), elephant (*Loxodonta africana*), opossum (*Monodelphis domestica*), zebra finch (*Taeniopygia guttata*), and anole (*Anolis carolinensis*) TERT proteins were retrieved from GenBank or Ensembl databases. Because no full-length cDNA clones are available for these genes we have manually checked these GenBank and Ensembl models against the GenBank whole genome sequence databases of corresponding species and we are suggesting following small scale corrections (shown in purple ink below). These corrections use data from following GenBank entries: [GenBank:NW_001235370.1, AAGU03019639.1, NW_001581900.1, NW_002198278.1, AAWZ02016664.1]. The corrections are based on the optimizations of the alignments with other vertebrate TERT proteins and the modification of features that would violate the conserved 16 exon-structure of the vertebrate *TERT* genes. The corrections include extensions or modifications of the N-terminal sequence, translation of the sequence that was considered intronic by the models and replacement of the amino acid sequences with X symbols where the underlying genomic sequence contains sequencing gaps. N-terminal extensions of zebra finch and anole TERT proteins are suggested based on the existence of N-terminal extensions in chicken and quail TERT proteins. All other corrections are strongly supported by the significant alignment (or misalignment for the X-substituted sequences) with corresponding regions in other TERT proteins.

```
>ptrTERT    Modified version of gi|114598857|ref|XP_001141663.1| PREDICTED:
telomerase reverse transcriptase isoform 2 [Pan troglodytes]
MPRAPRCRAVRSLRLSHYREVLPLATFVRRLLGPQGWRLLVQRGDPAAFRALVAQCLVCVPWDARPPPAAPSFRQVSCCLK
ELVARVLQRLCERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARC
ALFVLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARRRRGGSASRSLPL
PKRPRRGAAPEPERTPVGGQSWAHPGRTRGPSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSR
PPRPWDTPCPVYAETKHFYLYSSGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMR
PLFLELLGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHSSPWQVYGFVRAC
LRRLVPPGLWGSRHNNERRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGSPAAEHRLREEILAKFLH
WLMSVYVVELLRSSFYVTETTFQKNRLFYRKSVWSKLQSIGIRQHLLKRVQLRELSEAEVRQHQEARPALLTSRLRFI
PKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKALFSVLNRYERARRPGLLGASVLGLDDIHRARWTFVLRVRAQDP
PPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYXXXXXXXXXXXXXXXXXXXXXXXXXVSTLTDLQPYMRQFVAHL
QETSPLRDAVIEEQSSSLNEASSGLFDVFLRFVCRHAVRIRGKSYVQCQGIPOGSILSTLLCSLCYGD MENKLFAGIR
RDGLLLRLVDDFLLVTPHLTHAKAFLRTLVRGVPEYGCVVNLKRTVVNFPVEDEALGGTAFVQLPAHGLFPWCGLLLD
TRTLEVQSDYSSYARTSIRASLTFNRGFKAGRNMRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHA
CVLQLPFHQVQVWKNPTFFLRISDTSALCYSILKAKNAGMSLGAKGAAGPLPSEAMQWLCHQAFLLKLTRHRVTYVPL
LGLSLRTAQTLQSLRKLPGTTLSALEAAANPALPSDFKTILD
```

```
>lafTERT    Modified version of Ensembl gene: ENSLAFP00000010943 Description:
telomerase reverse transcriptase Species: Loxodonta africana
MPRAPRCRAVRALLRGRYREVLPLATFAQRLGAEGRRLLVRRGDPAAFRSLVAQCLVCVPWGAPPPPTAPSFRQVSCCLK
ELVARVVQRLCERGAKNVLAFGFALLDEARGGPPMAFTTSVRSYLPNTVTDTLRGSGAWGLLLQRVGGDDVLAHLLARC
SLYLLLVAPSCAYQVCGPLPLYELCARDTRGPEPGCGAPRRTVGGTRRRRWD SAGERHPLAKRSRHSLEALEPGQVPQSWAH
LDRACQSDSDPHAVTPGRRAAAKAVSWEGEIAGRNRSSLTGVKERGAGPSSAQPPLSLLPQAYVETKRFLYCLGAKE
RLPSTFLLSSLQGNLTGARRLVETIFLRTKPAWRKTQLRRLTPRYWRMRPLFQELLGNHARC PYSVLLRTHCPRLGSA
MATPAGTGGAGTGTPEQPPGPAVKPAPEEDAGLQCVVQLLRQHSSPWQVYTFVRACLHRLVPPGLWGSSHNKCRFLRN
VKMFIISLGKHAKLSLQELMWKMKVQDCEWLRRSPGDYAVPASEHRLREEILAKFLFWLMDTFVVGLLRAFFYVTETMF
QKNRLF FFRKSVWNKLQSIGIRQHFNKVRRLRELSEAEVQRRQQT KPSLLTSKLR FVPKAHGLRPIVNMDYVVGARTFR
REKKRLQHFN SQMKNLFSVLNRYERAVRPGLLGASVLGIDDVYKAWRAFVQVVRAGDPRPQLYFVKVDVTGAYDAIPHN
KLVEVIANVIKPKENIYCIIRYAVVQKTAQGYIRKSFKRHQVSTLADLLPYMKQFVEHLQETSSSLKNAV VIEQSSSLN
EPASSLFDFFLHLVHN NVIKIRGKYYVQCQGIPOGSILSTLLCSFCYGD MENKLFSGIQEDGLLLRLVDDFLLVTPHL
TQAKAFLSTLVRGVPEYGC MINLRKTVVNFVDDGPGQATILQLPAHCLFPWCGLLLDTRTLEVFC DYSSYARTSIKA
SLVFNHGFKAGRNMRRKLFVAVLRLKCHGLFLDLQVNSLQTVYINVYKIFLLQAYRFHACVLQLPFNQVVGKNPSFFLR
VISDTALRCYSILKAKNAGMSLGAPGAVGPF PSEARWLCHHAFLLLKLSRHRVTYKCLLGLTKTTKMQLG RKI PKATM
ALLEAAAAPSLCGEFKTILD
```

>mdoTERT Modified version of gi|126320879|ref|XP_001369432.1| PREDICTED:
similar to telomerase reverse transcriptase [Monodelphis domestica]
MASNPVSGSLLASVAFRAVRSVLQARYRDVLGLAEFVQRLGDEAQAEGGSGADVQLLRGGEPEVFQVQVFSQCVVCVP
WDARPPRPLTFQQLSSQKEVVARIVQRICEKKKKNILAFGYTLLEEKRMSPVMFTTNVYNYHPNTITETISVSALW
EMLLSRIGDDVMYMLEHCSLFMMVPPSCCYISGLPIYDLYLKDSTPPSGFVQRTYKQGANASLDNVRKISLLSKS
LAKSNLRKEMLSKEKAIKVVQQNQNSTEDPEDKSLRGRPEESCDQGEQVRTHSVHSTMALLSKRQREDEEKSEISAK
RSKTEELLQEKRELILGQELHENNESNLNDNESLTQRSMEICSSRPLFNEKSIHKEDGGGEGCFIKTKGSLSLLEHK
DGDLSRHNSSVIRSTLKGKAKARSSEGCLNRGAKEADATFSHPLCPGNEPPKGLKSTSSSIYINRKRFLYSAWNFRE
CLPTSFLNRLHDSLSGGQKLIVETIFLTSHLFGQKGDPPQQRTPWKKRRLPKRYWQMRPLFKELIQNHRKCPYFILLK
KNCPLRFSAAKADPSSQLTKPQAMASSHLTQAEGIQGETQDQGHQGEHTSRASNSMEEEPSSSLKDQISSKESSLSDA
VSKKALPASHEVGKTDKGGGDILQLLRWHSSPWQVYVFLRECLHRLVPAELWGSTHNKCRFFKKNVRKFLTLGKHKDFS
LRELMWKMVRKDCAWLALVRGNHFIPIASEHRLREEILAKFVYWLMMNTYVVELLRSFFYITETMFQKNRLLFFYRKCVWS
KLQNIIGIRKHFKEKVHLKLLSEEEIKQNPPEARFICLASRLRFVPKPTGLRPIVNVDSIIGARVSNKGTDKKIRCFNGQ
LKNLFSVLNYERTLHPEVLGSSSLFGIDGIYKEWRQFVLRVKPSKDKISNFYFVKADVTGAYDTIPHEKLEIVISSIIK
PKEKKVYICIRRYAVVQKNAHGNIQKSFKRHVSTYKDALPYMSQFVSHLQETTSLOQNAIIVEQXXXXXXTSINLFAFFL
HLIRNNILKIKNKYYVQCQGIPOGSILSTLLCSLCYGNMENKLLSGIQDGLLIRLIDDFLLVTPHLTQAKVFLRTL
KGIPEYGLINPKKTVVNFVVEEDILSDSFTQLPAHCLFPWCGLLLDTRTLNVFCDYSNYSRTSIRTSLSFDHSIKA
GRNMNRNKLIIVLRMKCHGLFLDLQVNSLQTIYINVYKIFLLQAYRFHACALKLPFNQQVRKNPGFFLSIISDIASCCY
SILKAKNPGVILGARGSGPFPFEARWLCYHAFLIKLVNHRIVYKCLLGTCLKMSKLQLFRKIPKATMQMLRVASDPS
LSQDFKMILD

>tguTERT Modified version of Ensembl gene: ENSTGUP00000008676 Description:
telomerase reverse transcriptase Species: Taeniopygia guttata
MPGGRAAWSQGRRGCARWTRAQFGRNMAGKEPFAAVLSALRRCYAEAVPLETFVRRLLGDGGAGDAEVLRADDPGYRN
FVGQCLVCVPRGARAI PRPFTFQQLSSQSEVTARVVQRLCEKKKKNILAYGYSLPDENSSQFPVMPLSKIHSYLPNTA
TETLCISGFWETLLSRIGDDVMYMLEHCAIFMLVPPNNCYQVCGQPVYELISHSVDSLSLVFVRQRFQSKYKRTSLKLY
MQKRLMFHRNYLLKSSRQRYEENVSRMRNERKNRSLVPTDQSSAKTVSKGSNQIRMVTEDEQEQSSSSSCVSATAL
SLKRKIHREQFEIPAKKAKMGEKVREEKACSLVPNVNQSSSERSETGYVAEHSESVIKTPCISERSNSAVSGPSLVHT
SCGRRKSVAVSLLQRFQSNKPLESNTMQAESHRKRVEIRMYESQLASGQTKPVKGSKCRQOESPQPHLSKKLPNRL
SSATYIERKSLLYSCRSFQECFPKSFVLNRLQGSQAGGRQLVEAIFLSQNVQQRHNQSLPKHKRRKTLPKRYWQMR
HTFQQLLMNHGKCAYLALLKKNCPVWISSEISMTKTKLSCQATLSEEAQVQKQAEQFGKEPAKCLTSSRCGSDHTDLPD
NLGAPLAESVRGELPPSEEQNPREADSALTELLKQHSSHWQVYMFVRDCEKVI PAELWGSNHNKCRFLKNVVFIS
RGRFAKVSLLQELMWRMRVNDCMWLRRLGKGDHFPADHECFREELLAKFLYWLMTYVVELLRSFFYITETMFQKNMFL
YYRKFIWGLQNIIGIRNHVFKVQLRPLSSEIEIETIRQKKIVPVALKLRFI PKPNGLRPIVKVSGVVEPQALSKE
KMNHYNTQLKNLFSVLNYERTINTSFIGSSVFGKDDIYKTKWQFVIKILESGGEIIPHYCVKADVSRAVDYTI PHNKL
EVISRVLKPEKRTVYICIRRYAVIMITPSGRAKRLYRRHVSTFKDFMPDMKQFVSQLENASLQNAIIVEQSLTFYETS
SSLFNFFLQMIHNIILEIRNRYYLQCCGIPQGSILSTLLCSLFYGD MENKWLPGIKQDGLVIRLIDDFLLFTPHLMKA
RTFLRTLTTGIPEYGLLINPNKTVVNFVDDIPGCSKFKQLPDCRLIPWCGLLLDIKTLEVYCDYSSYTCTSISSLS
FNSSVTAGKNMKYKLSAVLKLKCHSLFLDLQINSLRTVLINEYKIFLLQAYRFHACVLQLPFPNQVRKNPHFFLRIIS
QTASCCYAILKTINAGIAEGNGLSGIFPIQVAEWLCYHAFVTKLLNHKAVYKCLLTSCLKVCKRKLIRKIPEDTVALL
QAVTEPSLCQDFKAILD

>acaTERT Modified version of Ensembl gene: ENSACAP00000001407 Description:
telomerase reverse transcriptase Species: Anolis carolinensis
MQKIEGAAPALLPPRRRGGPGQAMRRSQVCRLLRGCFFEEVLPLEAFVKRLQEKEAKAGGLPAEPLIQDGDPKCFRV
LVERCLVGRPRGGKAPPPRLVFQQIFSQHDI IARVIRRICKKKNVLAFGYDLDLDENHFLPHMPNLYSYFPNNTTE
TICQSILWEKILNRVGDDFLMYILEHCSLFMLVPPSCCYQICGQPVYIEIAFKDSTSF PKFLRQRYPGPKHSTLSGYLR
RRRFSSYKQHTARGNRKKWHPRRKLGSKANNILEGSYQQLLIQTQVKNFTVSASECPESKQRTSECKSLTTRSLK
WKGHYEMSAKRMKIMKIEDGLQKETGNLVHTQSKHQLSLDGDNAASKSSTSFCLADQLTPVTSVLHSNECGEQISGVH
VAHLDRKSFLSSKTMTLVSGSKAHCEPSTKIDSVDSTKQEGGIRSTQMVSAGATARRDFRVHNSTSSSEKSSDNTS
FKRYSLLYCHRQLHECLPNSFVLNKLKSGSPSGGQSLVEIVFFTSQIPKQLDSSNQSHSKRKKRLPKRYWQMRGLFQEL
LQKHAKCPYLGILKRNCPWIWSDSIRYGTEEQACEEKVNQERRHSSQESQEANTIEYCPSSVTVGLGHFPRTS
SCGETELGQKVPEEQPILDSSTSNFRGLLKQHSSHWQVYTFVRECLQRVVPALWGSYNKCRFYKNVKKFISLGKLA
TFSMQELMWKMRVNDCTWLRSLKSGPGHFVPASEHHFRQDLMSKFFYWLMDSYVTELLRSFFYITETMFQKNLLFFFRK
TVWSKLETIGLRNHLAKVHLHALSEEKIKNLQKEKYVPLASKLRFIPKTNGLRPVVRLLDSVVGAKTFCEKIRERKVQL
FKTQLKNLFSVLNYERIKNPALLGSSVFGKDDIFAANKQFVLKILELNEEMPKFYFVKADVMGAYDSIPHDKLEEVL
QALSPNKKTITYSIRRYAVIIRTRNGLLRKYRRHASTYKEFKPEMNHVSHLQESTSLRNAVVEQSI SLKETSSHL
SEFFSRILIRNSILKIKDSYVQNCGIPQGSILSTLLCNMICYGDMENKLLRGIQKDGILMRLTDDFLVTPHLTQAKTF
LRTLAMGIPEYGFVINASKTVVNFVDEDI PGLSGFKQLPSHCMFPWCGLLIDTQTLEVYCDYSSYSCTSISSLSFN
SSVKAGVSMRNKLLDVLKLKCHSLFVLDLQINSLRTVCINVYKILLQAYRFHACVLQLPFDQKIKSNPSFFLGIISQT
ASCCFCILKTKHSDIPFSAAGISSPLTYKAVQWLCYHAFSVKLAIRHVIYKCLLVPLAQSKRRLQQMPEATVQLLKE
VTEPSIYNDFKTILD

Sequences - *De novo* generated models of elephant shark (*Callorhinchus milii*) and tammar wallaby (*Macropus eugenii*) TERT proteins

The models were derived by computational analysis based on the alignments of vertebrate TERT protein sequences with the whole genome sequences (WGS) of the respected species and on the conservation of 16 exon-structure of vertebrate *TERT* genes. Some exon sequences were not found in the current WGS databases. The sizes of these missing regions were estimated based on the expected similarity to respective regions of other vertebrate TERT proteins.

```
>cmiTERT    Predicted Callorhinchus milii telomerase reverse transcriptase
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXLS
AQSDVVLRLVIQRICEKGKKNVLAFGYANVTETSSMSVRFAPNLC SYQPNPTTVTIKNSILWKTL LSRIGDDVMMYLLE
HCSLFMLVKPSCCYQLCGVPIYSLITAGTRLPALWLRRRPIRARFNILLKIVQKRIRFHKGFLLKRKRRTLVD RD SCL
DARGTTECNKQKWTQIVSVHTRKPAKRAKLDTDIAKDTPRDVSSTGTKVTEVRPLKRL LHDEDKAESPAKKVKRDRSP
EGAIPHSTPPEILSAGNKMNESPENVEVSDSTSIHSTRGGLVEQKVIDAKTGVGCQRKYPQGSTRERKRTWQEGTIAG
YKRQAPGTDAKATHPTGAGGVEVKRSWHFMGVQEAVSFGKSVNSESNEEKVLLLEGGRSESKGKQAKVAKRKDESGKRC
AKFAAGKDKRGTKRAASNEGDDNDHFNPGRS AKHGQSTKCI AKNDYSWDGLNPDGQCRAAVGVSGQAKPSPIPADASD
AATVKNVRTWGSVYVERGHI IYCNDNRECLPKSFLLNCLQGCSSGGQRLVEAIFLSSDAFGNNGKKQPNNYWRKRRLP
KRYWKMKNVFFRLLRNYKRCPYRLLLNRC A V IIRKDNVSCSEGNPTPSVNESNSFSTVKMPCKPDS DHLEVNRVTG
GATNLIATQG PVLRRKNLEDHTSVNKSVPGLSCPNDLHLEAETGSLQPSEIPNLRCQH KRPTGI PAGKQKV N KELKVF
ADFDQTGMASGSDGDL LQLLKHYSSPLQVYRFVRECLLRV I SDELWGSNHNKYRFLKNVKKL I SLGKYDRFSLSELMW
KMRVNDCTWLQ L NNGQCSVSPSEHRLREEILSKFLLWLMGT YV VHL LKSFFYITETMFMKNM L FYYRKC VWNKIEKIG
IRNHLAKAQLQPLSKEEWARKQRQKT A I P FYALRF I PKRNGLRPIVKMRNITGSKKCVKGSNARKXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXADVTGAYDTI PHAKLVDVI SRVLDPKVQEN
YCIRRYASIWSNSAGQIRKSFKRQXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXVQRRGIPQGSILSTLLCSLCYGD MENQXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXMLTAGIPEY
GCFIHPNKT M VNFPLDNDLLGCINVKTLQEHCLFPWCGLLLNTQTLEVYCDYSSYANTSITSSLT FNCSSKAGQNL RQ
KLLAVLKLKCHQIFLDLEVSTLRTVSINVKIFLLQAYRFHACVMRF PFGQVRDNP SF FLSVISDMASCCYSILKTK
NEXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXX
```

This model (cmiTERT) uses data from the following GenBank WGS entries: [GenBank:AAVX01020497.1, AAVX01037780.1, AAVX01011403.1, AAVX01240754.1, AAVX01301389.1, AAVX01019362.1, AAVX01016929.1, AAVX01315420.1]. The model adds one nucleotide (N) to the sequence derived from AAVX01019362.1 to correct for the break in reading frame. Sizes of the missing regions are based on the corresponding regions of *Xenopus* TERT (the TERT protein with determined sequence phylogenetically most closely related - see Figure 2 of the article).

```
>meuTERT    Predicted Macropus eugenii telomerase reverse transcriptase
MASRPPAPRSL LASA AFRPVREVL RDSYRDVLGLAEFVQRLEDEARTARAPGDRPDQQVAVRLLRSGDPEVFQV FVSG
CVVCVPW GARPLPRPLTFRQLSSQKEVVARIVQRICEKKKKNILAFGYTLLEENRMSLPVKFTTNVYNYHPNTVTETI
SVSALWEILLSRIGDDVMMYMLEHCSIFMMVPPTCSYQISGLPIYELNLKDSVSPSGFLRKKYPKEEPNVSLDNVRKK
ISFQKKSSVKLDVRKAILRSKDKTLKVAGQNQNSAEDMVEEPL EDRPENSCGIEQGSNADQLEEPAHSAHSMMALLSK
RQRKDEE ESEI SAKRPKKEKCLLEKRKELILEQGFELHENESNDS AENITQTSMEICSSKHFCNEKSALEEDGGRED
CFIETKATKGGLSLLEHNEGDLSDHDTSI I KSTSEDI I EAKSSEEGLDGGTEQKDAIFTQPLCPGNEQPKRLRSASGS
IVYINRKRFLYSARNFRECLPTS FLLNHLHDSLSGGQRLVETIFLTSHLFEQKGD PQQRT P WKKRRLPKRYWQMRHL
FKELIKNHRKCPYFVLLKKNCPLRFSAAKADPSSQLKKPQSVASSQLTKAEVQHQSQDQGGQREQSSETPLSCGSNN
TEKPSCSLADQMPDKESVLASRISQNTQPVSCASGKSVFRGKGGSDVHQLLSWHSSPWQVYMF LRECLHRLVPAELWG
STYNKCRFFKNVRKFLTLGKHDKFSLREL MWKMRVKDCTWLALVKGNHFVPASEHRLREEILAKFVYWLMDTYVVELI
RSFFYVTETMFQKNRLFFYRKCVWSKLQNI G I R K H F D K V Q L R A L S E E E I M Q N L E T R F I C L A S R L R F I P K P N G L R P I V N
VDSIVGAKVSNKGTKNKKIRCFNGQLKNLFSVLNYERTLNPDILGSSLF G I D G I Y K E W R Q F V L R V R Q S K D Q V S N F Y F V
KVADLRGAYDTI PHDKLVEVISSIIKPEENKVYCI RRYAVVQKNAHGNIQKAFKRHXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXSSSLNETSMNLF T F F L H L I H N S I L K I K N K L Y V Q C R G I P Q G S V L S P L L C S L C Y G N M E N Q L L P G I Q Q D G L
L I R L I D D F L L V T P H L T Q A K V F L R T L A K G I P E Y G C L I N P K K T V V N F P V E E D I L S D S D F T Q L P A H C L F P W C G L L L D T R T L
NVFCDYSNYSRTSIRASLSFDHSVKAGRNM RNKLI A V L R L K C H G L F L D L Q X X X X X X X X X X X Y K I F L L Q A Y R F H A C V L K
```

LPFNQQVRKNPGFFLSVISDIASCCYSILKAKNPGFTLGARGASGPFPPFEAVLWLCYHAFLIKLVNHRLVYKCLLGTL
KMNKLRRLFRKI PRSTMQMLKTAMDPLLSQDFKMILD

This model (meuTERT) uses data from the following GenBank WGS entries: [GenBank:ABQO010931898.1, ABQO010102539.1, ABQO010292497.1, ABQO010291975.1, ABQO010152530.1, ABQO010711080.1, ABQO010883553.1, ABQO011205628.1, ABQO010179239.1, ABQO010292459.1]. The sequence of the ABQO010711080.1 entry was extended by 12 nucleotides derived from the trace file [GenBank:gnl|ti|1658238143]. The model adds one nucleotide (N) to the sequence derived from ABQO010102539.1 entry and removes one nucleotide (T) from the sequence derived from ABQO010292497.1 to correct for the breaks in reading frame. The nucleotide sequence encoding the ten amino acid region YKIFLLQAYR is derived from EST entry [GenBank:EX196206.1]. Sizes of the missing regions are based on length of the corresponding regions of opossum TERT.

Sequences – Human AS TERT variant pΔ2(136-end)

This sequence is too short to be accepted by GenBank database.

```

LOCUS                               55 bp    mRNA    linear    PRI 10-MAR-2011
DEFINITION Homo sapiens telomerase reverse transcriptase isoform
            Delta2(136-end)(TERT) mRNA, partial sequence, alternatively
            spliced.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 55)
  AUTHORS   Hrdlickova,R., Nehyba,J., Lim,S., Grutzner,F. and Bose,H.R. Jr.
  TITLE     Insights into the evolution of mammalian telomerase: Platypus TERT shares
            similarities with genes of birds and other reptiles and localizes on sex
            chromosomes
FEATURES             Location/Qualifiers
     source          1..55
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /cell_line="SAOS-2"
     gene            <1..>55
                     /gene="TERT"
     misc_feature    <1..>54
                     /gene="TERT"
                     /note="This variant is spliced using an alternative splice
                     donor site in exon 2. The variant lacks an internal
                     segment encoded by the 1219 nucleotides of the 3' portion
                     of exon 2. The open reading frame (<1..>55) that starts
                     with the ATG codon of the full-length TERT isoform
                     (cross-reference: NM_198253.2) ends in a premature
                     termination codon that is predicted to lead upon
                     translation to nonsense-mediated decay (NMD). This variant
                     is, therefore, represented as non-coding."
BASE COUNT          7 a      23 c      17 g      8 t
ORIGIN
    1 ccccccgagg ccttcaccac cagcggggttg gctgtgttcc ggccgcagag caccg
//

```