

**Supplement Fig. 3. Sequences of insulin genes for *Homo sapiens*, *Pan troglodytes*, *Pongo pygmaeus*, *Macaca mulatta*, *Callithrix jacchus*, *Microcebus murinus*, *Tupaia belangeri* and *Canis familiaris***

Coding sequences (CDS) are shown in blue (mature protein) or red (signal peptide and C-peptide) uppercase, non-coding sequences in orange (transcribed) or black lowercase. Below each sequence is shown the Chromosomal or scaffold location, or traces from which the sequence was assembled, the position of the ATG start codon (or exon 2 start), and any adjustments made. The IGF-II gene follows from the 3' end of the insulin gene. 'Complement' indicates that the sequence shown is the complement of that given in the genome assembly (and chromosomal location numbers therefore read backwards).

**Homo sapiens (man)**

gggcctcagctggggctgctgtcctaaggcaggggtgggaactaggcagccagcaggggagg 60  
ggaccctccctcactcccactctcccacccccaccaccttggcccatccatggcggcat 120  
cttgggccatccgggactggggacaggggtcctggggacaggggtgtggggacaggggtc 180  
ctggggacaggggtctggggacaggggtcctggggacaggggtgtggggacaggggtgtg 240  
gggacaggggtgtggggacaggggtcctggggacaggggtctggggacaggggtctgagg 300  
acaggggtgtggggacaggggtgtggggacaggggtgtggggacaggggtgtggggacag 360  
gggtctggggacaggggtccggggacaggggtgtggggacaggggtgtggggacagggg 420  
tgtggggacaggggtctggggacaggggtgtggggacaggggtcctggggacaggggtgt 480  
ggggataggggtgtggggacaggggtgtggggacaggggtgtggggacaggggtctgggg 540  
acagcagcgaagagccccgccctgcagcctccagctctcctggtctaagtggaaagt 600  
ggcccaggtgagggctttgctctcctggagacatttgccccagctgtgagcagggacag 660  
gtctggccaccgggccccctggttaagactctaataaccgctggtcctgaggaagaggtg 720  
ctgacgaccaaggagatcttcccacagaccagcaccagggaaatggtccggaattgca 780  
gcctcagccccagccatctgccgacccccccaccccaggccctaattgggccagggcgca 840  
ggggttgagaggtaggggagatgggctctgagactataaagccagcgggggcccagcagc 900  
cctcagccctccaggacagggctgcatcagaagaggccatcaagcagggtctgtccaaggg 960  
cctttgcgtcaggtgggctcaggattccaggggtggctggacccccaggccccagctctgca 1020  
gcagggaggacgtggctgggctcgtgaagcatgtgggggtgagcccaggggcccccaaggc 1080  
agggcacctggcctcagcctgcctcagccctgcctgtctcccagatcactgtccttctg 1140  
ccATGGCCCTGTGGATGCGCCTCCCTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTG 1200  
ACCCAGCCGACGGCTTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAGCTCTCT 1260  
ACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCAAGACCAGCCGGGAGGCAGAGG 1320  
ACCTGCAGGgtgagccaactgccattgctgcccctggccgccccagccaccctgct 1380  
cctggcgtcccaccagcatgggcagaagggggcagggaggtgccaccagcaggggggt 1440  
caggtgcacttttttaaaaagaagtctcttgggtcagtcctaaaagtgaccagctccct 1500  
gtggcccagtcagaatctcagcctgaggacggtgttggcttcggcagccccgagatacat 1560  
cagaggggtgggcagctcctccctccactcggccctcaacaaatgccccgagccatt 1620  
tctccaccctcatttgatgaccgcagattcaagtgtttgttaagtaaagtcctgggtga 1680  
cctggggtcacaggggtgccccacgctgcctgcctctgggcgaacccccatcacgcccgg 1740  
aggagggcgtggctgcctgcctgagtgggccagaccctgtcgccaggcctcacggcagc 1800  
tccatagtcaggagatggggaagatgctggggacagggcctggggagaagtactgggatc 1860  
acctgttcaggctcccactgtgacgctgccccggggcgggggaaggaggtgggacatgtg 1920  
ggcgttggggcctgtaggtccacaccagtggtgggtgacctccctctaacctgggtcca 1980  
gcccggctggagatgggtgggagtgcgacctagggctggcgggcagggcgggactgtgtc 2040  
tccctgactgtgtcctcctgtgtccctctgcctcgccgctgttccggaacctgctctcg 2100  
cggcacgtcctggcagTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTG 2160  
CAGCCCTTGGCCCTGGAGGGGTCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACC 2220  
AGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTagacgcagcccgcagggcagc 2280  
cccacaccggccgctcctgaccgagagagatggaataaagcccttgaaccagccctgc 2340  
tgtgccgtctgtgtgtcttgggggcccctgggccaagccccacttcccggcactgttgtga 2400  
gcccctcccagctctctccacgctctctgggtgccacaggtgccaacgcccggccaggcc 2460  
cagcatgcagtggtctccccaaagcggccatgcctgtcggtgcctgctgccccacc 2520  
tgtggctcagggctccagtatgggagctgcgggggtctctgaggggcccaggggtgggtggg 2580  
ccactgagaagtgacttcttgttcagtagctctggactcttgagtcccccagagaccttg 2640  
ttcaggaaagggaaatgagaacattccagcaattttccccccacctagccctcccaggttc 2700  
tatttttagagttatttctgatggagtcctgtggagggaggaggtgggctgagggagg 2760  
gggt 2820

From human genome sequence, chromosome 11: 2139919–2137156. Complement.

ATG start codon at 1143–1145

**Pan troglodytes (Chimpanzee)**

aggtgcctgttctggggagctgggagggccggaggggtgtaccccaggggctcagcccag 60  
atgacactatgggggtgatgggtgctgctgggacctggccaggagaggggagatgggctccc 120  
agaagaggagtaggggctgagaggggtgcctggggggcccgacggagctgggccaagtgca 180  
cagcttcccacacctgccacccccagagtcctgccgccacccccagatcacacggaaga 240  
tgaggctccgagtgacctgctgaggacttgcctgcttgcctcccggtccccgggtcatgccc 300  
tccttctgccaccctcgggagctgagggccacagctggggctgctgctctacggcggggt 360  
gggaactgggcagccagcagggaggggaccctccctcactcccactgtaccacccccac 420  
caccttggccatctatggcggcatcttggccatcagggactggggacaggggtcctgg 480  
ggacaggggtcctggggacaggggtctggggacaggggtcctggggacaggggtcctgggg 540  
acaggggtctggggacaggggtcctggggacaggggtcctgggaacaggggtcctgggga 600  
caggggtctggggacaggggtcctggggacaggggtctggggacaggggtcctggggaca 660  
ggggctctggggacaggggtcctggggacaggggtctggggacaggggtcctggggacagg 720  
ggtctggggacaggggtcctggggacaggggtcctggggacaggggtctggggacagcggt 780  
gcaaagagccccgccctgcagcctccagctctcctgggtctaagtggaaagtggcccagg 840  
tgacggctttgctctcctggagacatttgccccagctgtgagcagggacaggtctggcc 900  
accgggccccctggttaagactctaataaccgctggccctaaggaagaggtgctgacgac 960  
caaggagatcttcccacagaccagcaccagggaaatggctccggaatgacagcctcagc 1020  
caggaccatctgccgacccccacccccagccctaataggggcagggcaggggttga 1080  
caggtaggggagatgggctctgagactataaagccagcggggggcccagcagcctcagcc 1140  
ctccaggacagggctgcatcagaagaggccatcaagcagggtctgttccaagggcctttgcg 1200  
tcaggtgggctcaggggtccaggggtggctggacccccagggccccagctctgcagcagggag 1260  
gacgtggctgggctcctgaagcatgtgggggtgagcccaggggcccccaaggcagggcacc 1320  
tggccttcagccggcctcagccctgcctgtctcccagatcactgtccttctgccATGGCC 1380  
CTGTGGATGCGCCTCCGCCCCGCTGGTGGTGGTGGCCCTCTGGGGACCTGACCCAGCC 1440  
TCGGCCTTTGTGAACCAACACCTGTGCGGCTCCCACCTGGTGGAAAGCTCTCTACCTAGTG 1500  
TGCGGGGAACGAGGCTTCTTCTACACACCCAAGACC CGCCGGGAGGCAGAGGACCTGCAG 1560  
Ggtgagccaaccgcccgttgctgccccctggccacccccagccacccccctgctcctggcgc 1620  
tcccaccagcatgggcagaagggggcagggaggtgccaccagcaggggggtcaggtgca 1680  
ctttttaaaaaagaaatgaagttctcttggtcacatcctaaaagtgaccagctcccctgtg 1740  
gccagtcagaatctcagcctgaggacgggtgttggcttcggcagccccgagatacatcag 1800  
agggctgggcacgctcctcccctcactgccccccaacaaatgccccacagcccatttct 1860  
ccaccctcatttgatgaccgcagattcaagtgttttggtaagtaaaagtcctgggtgacct 1920  
ggggtcacaggggtgccacgctgcctgcctctgggcgaacacccccatcacgccctgagg 1980  
agggcgtggctgcctcccctgagtggggcagaccctgtcggcagccctcacggcagctcc 2040  
atagtcaggagatggggaagatgctggggacagggcctggggagaagtactggggccacc 2100  
tgttcaggctcccgtgtgacaccgccccggggcgggggaaggaggtaggacatgtgggc 2160  
ggtggggcctgtaggtccacaccagtggtgggtgacctccccttaacctgggtccagcc 2220  
cggctggagatgggtgggagtgcgacctagggctgggtgggcagggcgggactgtctctcc 2280  
ctgactgtgctcctcctgtgtccctctgcctcggcgtgttccggaacctgctctgcgcgg 2340  
cacgccctggcagTGGGGCAGGTGGAGCTGGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAG 2400  
CCCTTGGCCCTGGAGGGGTCCCTGCAGAAGCGTGGTATCGTGAACAATGCTGTACCAGC 2460  
ATCTGCTCCCTTACCAGCTGGAGAACTACTGCAACTagatggaataaagcccttgaacc 2520  
agccctgctgtgcccgtctgtgtgtcttgggggcccctgggccaagccccacttcccggcac 2580  
tgttgtagccccctcccagctctctccatgctctctgggtgccacaggtgccaatgccg 2640  
gccagcccagcatgcagtggtctctcccaaaagcggccatgcctgtcggctgcctgctac 2700  
ccccaccctgtggctcaggggtccagtatgggagctgcgggggtctctcagggggccagggg 2760  
tgggtcagccactgagaaatgacttctggttcagttagctctgggactcttggagtcccca 2820  
gagacctgttccaggaagggaaatgagaacattccagcaattttccccccacctagccct 2880  
cccaggttctatttttagagttatttctgatggagtcctgtggagggaggaggtgggc 2940  
tgagggaggggggtcctgcagggcaggggggtgggaaggtggggagaggtgccgagagcc 3000

From chimpanzee genome sequence, Chromosome 11:2212574–2209575. Complement

ATG start codon at 1377–1379

**Pongo pygmaeus (orangutan)**

agagcccatgccccctcactatgggtcagactgaacctccaggtgcctgtcccggggggc 60  
tgggagggccggaggggtgtaccccaggagctcagcccagatgacactatgggggtgatg 120  
gtgtcgtgggacctggccaggagaggggagatgggctcccagaagaggagtgggggctga 180  
gaggatgcctggggggcctggacggagctgggcccagtgacagcttcccacacctgccc 240  
ccccggagtcctgcccaccacccagatcacacgaaagatgaggctcctagtggcctgct 300  
gaggacttgcctgttgcctcccgggtcccgggtcatgcccgccttctgccaccgtgggga 360  
gctgagggcctcagctggggctgctgtcctaactggtgggggggaaactgggcagccagcag 420  
ggaggggaccctcccctcactcccactccccaccacattggccatccatggcggcttc 480  
ttgggcaatcagggactgaggacaggggtcctggggacaggggtcctggggacaggggtc 540  
tggggacagcggcgcacagagccccgcctgcagcctccagctctcctggctaatgtgg 600  
aaagtggcccaggtgagggctttgctctcctggagacatttgcctcccagctgcgagcagg 660  
gacaggtctggccaccagggccctgggtaagactcctaatgaccgcctggccctgaggaag 720  
aggtgctgacgaccaaggagatcttcccacagaccagcaccagggaaatgatccggaaa 780  
ttgcagcctcagccccagccatctgccgacccccccaccagccctaatgggccagg 840  
cggcaggggttggcaggtaggggagatgggctctgagactataaagccagtgggggccca 900  
gcagccctcagccctccgggacagggctgcatcagaagaggccatcaagcaggctctgttcc 960  
aagggcctttgctcaggtgggctcagggctgccccacttgggggttccaggggtggctgg 1020  
accacagctccagctgagctgggagggagctggctgggctcttgaagcatttggggg 1080  
tgagccagggggcccagggcagggcacctgcttccagccgacctcagctctgctgtc 1140  
tcccagatcactgtccttctgcccATGGCCCTGTGGATGCGCCTCCTGCCCTTGCTGGCGC 1200  
TGCTGGCCCTTGCGGACCTGACCCGGCCAGGCCTTTGTGAACCAGCACCTGTGCGGCT 1260  
CCCACCTGGTGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCA 1320  
AGACC CGCCGGGAGGCAGAGGACCTGCAGGgtgagccaaccgccccttgcctgcccctgg 1380  
cgccccagccatcccctgcttctgccactcccaccagcatgggcagaagggggcagga 1440  
ggctgccaccagcagggggcaggtgtacttttttaaaaagaaatgaagtctcttgggt 1500  
cacatcctaaaagtgaccagttccctgtggccattcagaatgggggcttcagcagcccc 1560  
gagatacatcagaggggtgggcacgctcctcccctcactcggccctcaaacaaatgcccc 1620  
cagcccatttctccaccctcatttgatgaccgcagattcaagtgttttgttaagtacaag 1680  
tcctgggtgacctggggctcacaggggtgccccatgctgcctgcctctgggcgaacgcccc 1740  
tcagccctgaggaggacatggccgctgctgagtgggccagaccctgtcaccagggc 1800  
tcacggcagctccgtagtcaggagatggggaagatgctggggacagggcctggggaggag 1860  
taccggggcaccctgttcaggtcctgctgtgacactgccccggggcgggggaaggaggt 1920  
aggacatgtggaccttggggcctgtaggtccacaccagtggtgggtgaccttccctataa 1980  
tctgggtccagcccggctggagatgggtgggagtgcgacctagggctgggtgggcaggtgg 2040  
gcagtgtctctcccctgactgtgtcctcctgtgtcctctgcctcaccgctgttccggaac 2100  
ctgctctgcgcggcagcccctggcagTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGC 2160  
AGGCAGCCTGCAGCCCTTGGCCCTGGAGGGGTCCCTGCAGAAGCGTGGTATCGTGAACA 2220  
ATGCTGTACCAGCATCTGCTCCCTTACCAGCTGGAGAACTACTGCAACtagatgtggcc 2280  
ggcagggcggccccgcaccctccgcctcctgcaccgagagagatggaataaagcccttga 2340  
ccagccctgctgtgccgtctgtgtgtccttggggggcctgggccaagccccacttcccggc 2400  
actgttgtgagccccctcccagctctctccacgctctctgggtgccacaggtgccaacgc 2460  
cggccagggcccagcatgcagtggtcttccccaaagcggcgatgcctgtcggctgcctgct 2520  
accctgccttggggctcagagtccagtgctcggagctgcgggggtctctgagggggccgg 2580  
gtggtggggccactgagaaatgacttcttgttcagtagctctggaccctcggagtcacca 2640  
gagaccttgttcaggaagggaatgagaacattccagcaattttccccccacctagccct 2700  
cccaggttctattttttagagttatttctgatggagtcctatggaggaaggaggatgggc 2760  
tgagggagggggctcctgcagggcagggggctgggaagggtggggagaggctgccgagagcc 2820  
accgctatccccagctccgggcagccccggggcagtcacacaccctggcctcgcggccc 2880  
aagctggcagccatctgcagccacagcttatgccagcccaggtccagccagacacctgag 2940  
ggaccactgggtgccttggaggaagcaggagagctcggacggcaccatgagccggggcag 3000  
gtgcagggaccgtggcagcactgggcagggcctcagaaccacgccttgggcaccctggc 3053

Assembled from traces (*Pongo pygmaeus abelii*):

PPAE-abu34f06.g1    PPAE-abu33a08.g1    PPAE-abu33a08.b1    PPAE-abu28f03.b1  
PPAE-abu25g08.g1    PPAE-abu25g08.b1    PPAE-abn64e08.b1    PPAE-abn64b05.b1

PPAE-abn60b06.g1    PPAE-abn55f02.b1    PPAE-abn55e11.b1    PPAE-abn55e04.b1  
PPAE-abn55b11.g1    PPAE-abn55b06.b1    PPAE-abn50h04.g1    PPAE-abn45h01.b1  
PPAE-abn45g11.b1    PPAE-abn45e02.b1    PPAE-abn22b12.b1    PPAE-abn43g08.g1  
PPAE-abn43f12.g1    PPAE-abn43f04.g1    PPAE-abn42h06.b1    PPAE-abn42g01.g1  
PPAE-abn42c09.g1    PPAE-abn42b10.g1    PPAE-abn41h09.g1    PPAE-abn41g04.b1  
PPAE-abn41e02.b1    PPAE-abn41c10.g1    PPAE-abn41b06.b1    PPAE-abn41b01.g1  
PPAE-abn40f10.g1    PPAE-abn40f08.b1    PPAE-abn40b10.b1    PPAE-abn40b01.b1  
PPAE-abn38g05.g1    PPAE-abn38f10.g1    PPAE-abn32e12.g1    PPAE-abn32c02.g1  
PPAE-abn27f02.b1    PPAE-abn27c04.b1    PPAE-abn22f11.g1    PPAE-abn22f03.g1  
PPAE-abn22e09.g1    PPAE-abn22b12.b1    PPAE-abn22b11.g1    PPAE-abn45b06.b1  
PPAE-abn16g03.b1    PPAE-abn16d12.b1    PPAE-abn16d09.b1    PPAE-abn11c06.b1  
PPAE-abn09c07.g1    PPAE-abn09b05.g1    PPAE-abn08g06.b1    PPAE-abn08g03.b1  
PPAE-abn08f01.g1    PPAE-abn07g10.g1    PPAE-abn07g03.b1    PPAE-abn07d12.b1  
PPAE-abn07b08.b1    PPAE-abn06f11.b1    PPAE-abn06d06.b1    PPAE-abn06c04.b1  
PPAE-abn06b04.b1    PPAE-abn06a09.g1    PPAE-abn05h08.g1    PPAE-abn05g03.b1  
PPAE-abn05e07.b1    PPAE-abn05c03.b1    PPAC-bhz34e11.g1    PPAC-awm58g11.g1  
PPAC-awm29a11.b1    PPAC-ary68h04.b1    PPAC-ajv23f06.b1    PPAC-aiu42g03.g1  
PPAC-adw11a03.g1

Possible polymorphism at 1713 (t/c)

ATG start codon at 1164-1166

**Macaca mulatta (rhesus macaque)**

cctcacctggggctgctgtcctaaggcgggtgggaactgggcagccagcagggaggggac 60  
ccttcctcactccccaccctgccaccttggcccatccacgggggcatggtgggcaacc 120  
agggacttagaaaaaggtcccaggacaggggatctggggatagtggcgcaaagagcccc 180  
gccctgcagccccagctctcctggtgtaatgtggaaagtgggcccgggagggccttgc 240  
tgctggagacatttgcctcagctgcgagcagggacaggcctggccaccagggcccctggt 300  
taagactctaatagcgcgctggccctgaggaagaggtgctgacgaccaaggagatcttcc 360  
cacagaccagcaccagggaaatgatccggaattgcagcctcagccccagccatctgc 420  
taacccccaccagggcctaataaggccagcagcaggggggtggcagggagggagat 480  
gggctctgagactataaagccagtgaggagcccagcagccctcaaccctccgggacaggct 540  
gcatcagaagaggccagcaagcaggctctgttccaagggccttcgcggttaggtgggctcag 600  
ggctgccccacttgggggttccaggggtggctggaccccaggccccagctctgcaacaggg 660  
aggacgtggctgggctcttgaagcgttgggggtaagcccaggggccccagggcagggca 720  
cctggccttcagccggcctcaggcctgcctgtctcccaggctactgtccttccgccATGG 780  
CCCTGTGGATGCGCCTCTTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCGG 840  
CCCCGGCTTTGTGAACCAGCACCTGTGCGGCTCCCACCTGGTGAAGCTCTCTACCTGG 900  
TGTGCGGGGAGCGAGGCTTCTTCTACACACCCAAGACC CGCCGGGAGGCAGAGGACCCTC 960  
AGGgtgagccccaccgcccagcgtgccccctggctgccccagccacccccctgctcctgg 1020  
cgctcccaccagcgtggacagaagggggcaggaggctgccaccagcaggggggtcagga 1080  
gtacttttttaaaaagaaatgaagttctcttggtcacctcctaaaagtgaccagctccc 1140  
tgtggcccagtcagaagctcagcttgagcatggcgttagcttcagcagctgagccctgac 1200  
atacgtcagaggggtgggcaggttccctccctccactgcccgtcacacaaatgccccgag 1260  
cccatttctccaccctcatttgatgaccacagattcaagtgttttgttgagtacaagtcc 1320  
tgggtgacttggggtcacaggggtgccccacgctgcctgcctgagtgggccagaccctgt 1380  
cgccaggcctcatggcagctccgtagtcaggagatggggaagatgctggggacaggccct 1440  
ggggagaagaaccgggggacactgttcaggctcccactatggcaccgccctggggcgggg 1500  
gaaggaggtaggacatgtgggcattggggcctgtgggtccacaccagtggtgggtgacc 1560  
tccctctaacctgggcccggcttggctggagacgggtgggagtgcgacctggggctggtg 1620  
ggcaggtgggcaactgtctctccctgactgtgtcctcctgtgtcctctgctcctgctg 1680  
ctccgacactgctctgcagggcctgcctggcagTGGGGCAGGTGGAGCTGGGCGGGG 1740  
CCCTGGCGCAGGCAGCCTGCAGCCCTTGGCGCTGGAGGGTCCCTGCAGAAGCGCGGCAT 1800  
CGTGGAGCAGTGTGCACCAGCATCTGTCCTCTACCAGCTGGGAGAAGTACTGCAACTa 1860  
gatgcgggcccgcaggcggccacaccctccgcctcctgccaagagagatcgaataaag 1920  
cccttgaaccagccctgctgtgccgtctgtgtcttggggccctgggacagggcccgc 1980  
ttccgggcagtggttgtagccccctccagctgtctccatgctctctgggtgccacaggt 2040  
gccaatgctggccaggccccgcatgagtggtctccccaaagcggcgatgcctgtcggc 2100  
tgctgctacccccaccctgtggctcaggggtccagtatcagagctgggggggtcgtgag 2160  
aggccaggggtggtgggaccactgagaaatgacttcttgttcagtagctctggacccttg 2220  
gagtcccagagacctgttcaggaaaggaatgagaacattccagcaactttccccac 2280  
ctagccctcccaggtcctattttagagctatcttctgttgagtcctgtggagggagga 2340  
ggctgggctgagggaggggttccctgcagagcagggggctgggaaggtggggagaggctgc 2400  
cgagtgccaccgctatccccagctctgggcagccccgggacagtcacataccctggtct 2460  
cgaggcccaagctggcagctgtctgcagccacagc 2595

From Macaca genome sequence, Chromosome 14:2170699–2168240 (Complement) (numbers do not match exactly because of indels introduced on correcting).

Completed and corrected using Traces:

zdm74c08.b1	147152767	147199241	147199199	147198316	147194076
147194005	147187144	wsy40g02.b1	RHNFQ23TF	RHNFN72TF	98858588
96856822	96849207	95932944	147134669	147183203	147164073
147137736	147152815	147187259	147211443	147211395	147206363
147204513	147200013	147199998	147199990	147199971	147199253
147123072	147187886	147187233	147116108	147175222	147143747
147164082	147187887	147110484	147234851	147183302	147152783
147123150	147193945	147193514	147187892	147123089	147175209

147123130	147193482	147187958	147187066	147128117	147183223
147183214	147128226	70218867	70212171	147234967	147226248
147226227	147226208	147226201	147183200	147187948	100131119
147137708	147223251	147223072	147223009	147221905	147221893
147221887	147221886	147217574	147216117	147211483	

MQAD0012h23.g1 MQAA-alh39a21.g1

Possible polymorphisms at 110 (g/c), 1171 (t/c), 2085 (g/a), 2438 (g/a)

ATG start codon at 777-779

**Callithrix jacchus (marmoset)**

ccctggcgctcaaaggtgggttaggggctgctgggcaatgaggctcggagcccttgactcc 60  
cgatgggtccagctgacctccagttgctccccagggctcagcccagatgacactgtggg 120  
ggtgatgggatcgtggggcctggccaggagaggcaatgggcaggggctcctggaagag 180  
gagtgggggcccaggaaggtgcctatggggccaggacggtgcttcccacacctgcccattgg 240  
agccctgcagccaccacagatcacaggaagatgaggctcaggcagcctgctggggacc 300  
gctgctcgtccctgggtccctgggtcctgcctcctgccaccaggggaactgagggcct 360  
cagctggggctgctgtcctgggggtggggggaactgggcagccagcaggaagggggcct 420  
tctcgtccccccccgcaacctctgccaccatgggggcatctcaggcaaccagaga 480  
tggagggcaggagtctggggacagcaacatgaagaacccccgccctgcagcctcctgctct 540  
cctggtctaagtggaaagtggcccagatgagggctgtgctctccgggagacatttacc 600  
ccagctgcgagcaggcacaggtctggccacgaggacctgggttaagaatctaataacc 660  
ctccctggccctgaggaccaggagatattcccgccgaccagcagccgggaaatgatct 720  
ggaaggtacagcctcagccccagccatctgccagccctgcaccccaggccctaatgggc 780  
caggcggcaggggtggcacgtagggaaagatgggctgtgggcctataaagccagcggggac 840  
ccagcagccttagccctccgggaccagctgcatcgaggagaccagccagcaggtgtgt 900  
tccaagggcctttgagccagcctggggccagggctgccccacttgggggttcagagcag 960  
ttggaccccaggcccagctctgcagtgggagggcatggctgggctctgaagcatttggg 1020  
tgagcccagggcccagggcagctgacctcagccggcctcagccctgctgtc 1080  
tcccagccgctgtgcttccaccATGGCCCCGTTGGATGCCCCCTGCCCCTGCTGGCAC 1140  
TGCTGGCCCCCTGGGGACCCGAGCCAGCCCCGGCGTTTGTGAACCAGCACCTGTGCGGCC 1200  
CCCACCTGGTGGAAAGCCCTCTACCTGGTGTGCGGCGAGCGAGGTTTCTTCTACGCCCA 1260  
AGACC CGCCGGGAGGCGGAGGACCTGCAGGgtgagccccaccgcccctactccccccac 1320  
cacccccagccacccttgcctcctgctgctcccaccagcctgggcaacaggtggactttt 1380  
taaaaagagatgaagttctcttggtcacatcctaaaacgtgaccagctccctgcccggccc 1440  
agcagaaactcagcctgaggacggtattggcttcggcagctgagctccgagatacctaag 1500  
agggcgggagggcagattcctccctcatgtgccctcaagcaagtgccccgagccca 1560  
cttctccaccctcaccgaaaactgcagcttccagtgttttgttgagtacatcaagtcct 1620  
gggtgacctggggtcacagggtcacaaacgccccatggcacctgaggaggacatggctgc 1680  
tggccacagtgtccctgggcttcacgacagctccaaagtcaggaaatgaggcagggcgtg 1740  
gggacaggccctgggaagacgtaccgggggtacctgttcagcctcccgccatggcaccac 1800  
ccagcgcattgaagccctctatgtccacatctggtgtgagccacccttctccacctggg 1860  
cccagcttggctggagaggggtgggagcgtgacctggggctgggagggcagggcagctg 1920  
tctcctcctgactgtgccatcctgtgctcctcttccctcactgctgttctggacctgctg 1980  
tgtggctcgcctggcagTGGGGCAGGTAGAGCTGGGTGGGGGCTCTATCACGGGCAGCC 2040  
TGCCGGCCCTGGAGGGTCCCCTGCAGAAGCGTGGCGTCGTGGATCAGTGTGCACCAGCA 2100  
TCTGCTCCCTCTACCAGCTGCAGAACTACTGCAACTagactggggccacaggccaccccg 2160  
tgcccactgccacctgcaccagcactgctccctctgccagctggagaaccgcagctagt 2220  
cgcggccccgaggcaggtcaaatgcccctgcacctctgcacctgcaccacagctgat 2280  
ggaataaa gcctcgaaccagctctgctgtgctgtctgtgtgttttcggggcctgggcca 2340  
ggccccgcttccctggcactgttatgagccccctcctagctctctagacactctctagatgc 2400  
ccataggtgccctcactggccagggccccccaaagtgggtgctgttggctgcctctt 2460  
gctactcccgcgctggggctcagggctccagtttcagagttggggggggtctcttgagagg 2520  
ccgggtggcggggccactgggaaatgacttcttgcctcagtagctctgggctcttggagtc 2580  
cccagagaccttgttgaggaaaggggatgagaagattccagcatttccccccacctagccc 2640  
tcccaggttccattttacagctatttctgatggagtcctgtggaagaaggaggctggg 2700  
ctgctggagcgggtcctgcagggcgggggctgggaagtggggggaggatgccgagtggtac 2760  
ctactatccccagctccaggcagctcccggagcagtcacacaccctggccttgaggcccaa 2820  
actggcagcggctctgcagccacagcccataacatccaggcccagacacctgggggccccca 2880  
ctcgtgccttggagggagcagggaaaggtcagatgggagcatgagctggggcaggtgtgag 2940

Assembled from sequence traces:

167500329	ghi19c09.y1	145105103	CXAF-aol07d01.g1	ghi17g02.y1
ghi13h03.y1	ghi14g02.y1	156711775	ghi13b05.y1	ghi01g12.y1
166757790	ghi20b01.x1	ghi14h03.x1	164282610	144403878
ghi10a09.y1	150978917	CXAG-1353D08.b1	ghi09b05.y1	ghi01h09.y1

ghi09d11.y1	146022714	ghi19d12.x1	145116400	ghi18g02.y1
ghi04d07.y1	ghi09d05.y1	ghi03g07.y1	153363437	ghi14g02.x1
167471919	ghi18d01.y1	144409249	ghi20d07.x1	161077821
146028179	142666551	167269906	ghi03g07.x1	ghi11c05.y1
155772517	ghi19d12.y1	156503136	142015260	ghi15g10.y1
ghi19d05.x1				

ATG start codon at 1104-1106

Microcebus murinus (gray mouse lemur)

**Microcebus insulin 1**

gaggggtgccgggaggggcacaaaggacacccccgggtccctgcagccccaggctctgccg 60  
actgctgcagaaaacgtctctgggagtcgggtggggccgtgctctccccggccaccctcgc 120  
ccccagctgtgacagggacagctctgcagtcagggcgtcagggcctcgtaagacgctaa 180  
tgaccgcgtggccccagcagagaggtgctgaccacggaggagatgctccccggccccgaag 240  
cagggaaaatggtccggaactgcagcctcagcgcacccccccccggccatctgccgacccc 300  
cccaggccctaattgggcccagcggccggcggggcagggaggtgggctcggggctata 360  
aagccggcagcgcggcagccccagccctgcggaccagctgtttccccggccgtcagc 420  
gagcaggtctgtgccaggggctccgggtccggggcgggtgggacccgggaccccagctctgc 480  
atggtggtggtggggagggacgtgggctcctctcgtggggcatttgggggagcaagcggg 540  
ggtcccggggcagggcgcccggccacgctggcctcagccccgctcctctcccaggtcttt 600  
gccccggccccggccATGGCCCTGTGGACGCGCCTCCTGCCCTGCTGGCCCTGCTGGCCC 660  
TGTGGGGGGCCGAGCCCGCCCCGCCCTTCGTCAACCAGCACCTGTGCGGCTCCCACCTGG 720  
TGGAGGCCCTCTACCTGGTGTGCGGCGAGCGGGGCTTCTTCTACACGCCCAAGAGCCGCC 780  
CGGAGGTGGAGGACGCCAGGgtgagcgcgcagggccgcggggcagagggggcgggagg 840  
tgccaccgaaggaaagggacctcttttggcctgccacgtcctgaaagcgcctgtggccc 900  
ggccagagactctgggcttcaggacagtggccgctgcacgagcagggccccgatggcacc 960  
ctgagaccactgtggcctcctcctgcacgtgcctccccaaacaacacccccagccccgt 1020  
ccctacctcacaggacagcagccttccaggggggactttagtaaagcccaaggccag 1080 Alu  
gtgcggtgactcacgcctgtaatcccagcactctgggaggccaaggtgggcagactgctc 1140  
gaggtcaggagtttgaaccagcctgagcaagagcagacccccatctctactataaatag 1200  
aaaattaattggccaactaatatataatagaaaaaattagccgggcatggtggcacatgcc 1260  
tgtagtcccagctactcgggaggctgaggcaggaggatcgcttgagcccaggagatggag 1320  
gttgctgtgagctaggctgacgccacggcactcgtcttagcctgggcaacaagcgagac 1380  
tctgtctcaaaaaataaataaaaaaacaagcgcgccaggtcctgggtgtgggtggtc 1440  
tcaggggtgaccttgggaagtggccccaccgcgccagtgggggcgcggtagcaataggagg 1500  
ctggcagtggggagtcaggggatgggcaacccccggggacgggtgcccggtggggggcacc 1560  
tgccacgttcccgcgcggcactgcccggggcggggacaggtggcaggggtgtgtgggct 1620  
cggggcacgectgtccctgccagcggccaccggggtagggggcggggcggctctgtgtc 1680  
cctgacgccttggcctgtctctctctctcctcctcctgtctgctgacctggcaccgcccggcg 1740  
gtagCGGGGCAGGTGGGGCCGGACGGCGGGCTGGGGCGGGCGGCCCTGCAGGCCCTGGCG 1800  
CTGGAGGGGGCCCGCAGAAGCGCGGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCG 1860  
CTGTACCAGCTGGAGAACTACTGCAACTagccaccggccccgccccgccccgggacggaa 1920  
taaacctcttgaatggcccccggtgtctgtcttccggtgtgtctccgagccccgccccgccc 1980  
cgccgcacccccctccccagcagcccccaacccccgagcgtcctccatgctccccggggcg 2040  
ccggctaggccccgggtaccccaaggagcgggggtgctgccactgccccccccggggctc 2100  
tggggggcgtgggcgagctgagatgccggggggctcttggggagctgacttccctcgttca 2160  
gaagccctggaccctcagggtccccagagaattttcaggaagagaatgagaacattcc 2220  
agtggcttcttcccaccta 2280

Assembled from traces:

mmi20g10.y1	mmi13e04.y1	mmi11b02.y1	mmi08h07.y1
mmi06h06.y1	mmi06b09.x1	mmi03a07.x1	mmi02e04.x1
mmi02d08.y1	mmi01c08.y1	mmi01a01.x1	G730P68654FL23.T0
G730P68569FI1.T0	G730P67274RM16.T0	G730P6704FA4.T0	G730P66105FO20.T0
G730P6585RO23.T0	G730P65361RI19.T0	G730P63604FG3.T0	G730P63375FG14.T0
G730P6325FI20.T0	G730P62785FD3.T0	G730P62436RA13.T0	

ATG start codon at 615–617; note Alu-type sequence in intron 2 (mauve)

There is evidence for a second (incomplete) gene sequence, differing from this at ~ 5% of nt.

**Tupaia belangeri (tree shrew)**

agaagtggcagaagtccctgggtgggactggcctggccctcttggggggccgggaggggtgga 60  
ggaggttgaccacaaggcctcaaccaggagcactggagggggcagtgaggactgtggg 120  
gcttaactcgaaggtggacgggtggctgggggtgccagaaagagggggacctgcgcggga 180  
ggctgctcttactgaaaatgatagatagggccggccttttggggccatagcttctcca 240  
tcagattccgccttcttcttggccccagggtagctggggcttctgtcctggaatggagtg 300  
ggaacaagaagaggggggaacacttttctactcttcccacccaagaccccgggaggcct 360  
gttgggcagaccagggatcctgggggagggcttggctggagcggagtgccccacctccccgc 420  
cacagccccagctctcctggaccgatgtgggaaatggcttggcaaggcccaggtaaggg 480  
tttgctgtcctggagacatttgccccagctgcgagcaggaccactctagccatccagtc 540  
gttaagactaatgaccccaggggccaagtgagagatgttgacgaccaagaagatgtcccc 600  
acaggccccacaccagggaatgatccagaaattgcagcctcagccccctggccatctgcc 660  
gacaccctaccactggaggccctaattggggcagggtggcagggcttggggggcaggggag 720  
atgggcttgggtgctataaagccagtggggaccagccatcctcaggcctgcgggaccag 780  
ctgcattgtgaggtcatcagcaagcaggtctgtactgggggtctccacgccagctctgcct 840  
gggggctccagggcactggactgcaggcttgagctctgcaaagagggacgacatggcttt 900  
tctctgggagtgtaaggggtggtcactgggaccaggagcggggcacttgccccctcacct 960  
ggcctcaaggccaccctgcacactcctaggttattgtctctcaccATGGCCCTGTGGACA 1020  
TGCTTCCTGCCCTGTGTGCCCTGTGCCCTGTGGGGCCCTGAGCCAGCCCCGGCCTTC 1080  
GTGAATCAGCACCTGTGTGGCTCCACCTGGGTGGAGGCGCTGTACCTGGTGTGTGGGGAG 1140  
CGAGGCTTCTTCTATACACCTAAGACC CGCCGGGAGGTGGAGGATTCCTCAGGgtgagcac 1200  
ccccagccctgctacccccagtagctccaccctggcaggagtggaggcgggagggatgct 1260  
gccaccgggtgggaagtcagggtgacttctataaaaaaggtaatggagttcccttggcc 1320  
ccctcctaaaagtgaccagctctctgaggccgagtcggagagtctgggaataaggacttc 1380  
tccctcctcgctccccctggaccatttcgccacctgcctgtggggaaggcaacttcagga 1440  
gaggaagtttactgagtgaatcaacccttgggcttggatgatctcggggcaacttggggg 1500  
ggtcacgggggtgccccatgctgcctgcatccctgtgggcaaagtccctgggggtgctctga 1560  
ggaggacgggtggagcacctgacagggagttgggaaacgggtgcaggccccggggatggg 1620  
actgacagggccctggcactgcccttgggccagagggtttgtgggccagtc caatcagc 1680  
aacaaggtgaaggggtaccctgggggggctaccctggggctctctgggtccccggtagT 1740  
GGGCAGGTGGAGCTGGGCGGGGGCCCTGGAGCAGCCTGCAGCCCTTGGCGCTGGA 1800  
GGTGCCCCCGCAGAAGCGCGGCATCGTGGAGCAGTGCACCAGCATCTGCTCCCTCTA 1860  
CCAGCTGGAGAACTACTGCAACTagacctgggcccggcctgccaccggggggcggagtgcc 1920  
gcttctctgcttccccgcccaccaagagttcaataaaacccccaaatgagccctactgtgtc 1980  
atctgtttgtcttggggggcagtgagagtgtgtgccatccacctgcctggaacatgct 2040  
ggcctcaggggcccgtctccctggctctcctcctgccgccccctgcatggacctgccagc 2100  
gttctgcatacctacaggggtctacaggggccagggtccaggaggccctccccaggatgcc 2160  
aattccccctctcagctgccactgccctacctggaggctctgggaagacagaggaggac 2220  
tgtgggggtctccggggctaggggtgacttcccccttgctgatagccctggattcgcaag 2280  
gtcctccgagcattttcccagaaaagagaatgagaacattccagtgacttcttctcacc 2340  
cagccctctcaggctcgattgttatagctatttctgatggatccccatgtggaaggaaga 2400  
ggctgggtgaggggagggagtctgccaggggagggagagaaagaaagggaccagagtctg 2460  
cccctttcagtgcttgggtgggaggggacagccccagccctcagaccagaccagcaa 2520  
acaatctgagggaaactgcttgtgtcatggacctcatgagctggggcagacggaaggccc 2580  
aaggcgggcccctgggaggggtttcttgagccccagttctcagaggccctgagaactgcag 2640  
cctacaagggccagaagaccttccgggcccacaagaccagaccagaggcaaggtctctc 2700  
cagtccttggccctcttgccttccccaagccgggcccactggaccagattcctctagtac 2760  
agacctggcagttgtccccagaagtgcccttggaccctgacaaagccacactgggtccag 2820  
gtccagcccagggtcacctccaggcaggtgacccaagccaccacaaggcctggccagcctg 2880  
cctgccaccaggccaaaccaatctgtgcctttctgaaaactccgccccgggctgggctg 2940

From scaffold\_5168: 93156-90221 (Complement) (numbers do not match exactly because of indels introduced on correcting).

Corrected using Traces:

G835P66833FF9.T0 G835P67157RD23.T0 G835P62308FK15.T0 G835P61931FB16.T0  
G835P66632FL22.T0 G835P61796RP12.T0 G835P69670FN4.T0 G835P64058RP24.T0  
G835P62926RK1.T0 G835P61594FC9.T0 G835P6603FI12.T0 G835P62308RK15.T0

Region between 1718 and 1733 very doubtful - may be additional sequence missing here.

ATG start codon at 1006-1008.

**Canis familiaris (dog)**

gtgCGGaaagagctccagaaacaggggCagacacggagagggagcgtgggCaggacatat 60  
ggggcccccacacaggggtgtctgcagccccacagtagagttgagaagtgcggcccccac 120  
ctgaatgtctgcacagcggccgggtccccgagctcggtcctctgcacggtactctgggggtg 180  
ctaggggtctcagggctgtgtgctggggctgggggggtgggctgtgagcagggaggagat 240  
gtccccccacaagcgcceccctcatggtcttgcaccatcctagggacctggagggcagac 300  
cagggactggctgggggtccagggccccgggcagaagaagccccagctctggccccagct 360  
cccctgcactaatgtggaaaatgtggctcaggcaaggccccaggtggggggctctgcccc 420  
gtggggacatttggccccagctgttagtgccatcaggccccctcgtaaagactctaataga 480  
ccccgaggccccgggtgtgtgacgaccaaggagctcttcccgcagaccagcactggggaa 540  
atgatccagaaattgcagcctcagcctccggccatctgccacccccctcatggccaggccg 600  
tgggctcgggagctataaagccaggaggggtccagcagccccagccccgggaccatctg 660  
caccggacacggccggcaaacaggtctgtccccacgggtccccgcgcgcctccccg 720  
ccagcctgtgtctcaaggcagcaggaggagaagagctgcctcggggcctttgcgggggtg 780  
ggctcaggggtggggggggccgtgcccttgccagcctcaaccctgcctgtccccaggtcg 840  
ccATGGCCCTCTGGATGCGCCTCCTGCCCTTGCTGGCCCTTGCTGGCCCTCTGGGCGCCCG 900  
CGCCACCCGAGCCTTCGTTAACAGCACCTGTGTGGCTCCCACCTGGTAGAGGCTCTGT 960  
ACCTGGTGTGCGGGGAGCGCGGCTTCTTCTACACGCCTAAGGCCCGCAGGGAGGTGGAG 1020  
ACCTGCAGGgtgagccccgcgccecccgccctggctccctacctggccccagggcag 1080  
ccaggtggaatattaaaaagaaaaatgactttcccttggctctacatcctgcaagggacc 1140  
agtccttggtcaggggtcgggcaccaaagcctgagggcagcctcccaccttggcaccac 1200  
cctggggcctgggagccactggcaggggtgggggtgggcggggcgcgctctctccctgac 1260  
cctgaccgctctccgctggctcctccgcagTGAGGGACGTGGAGCTGGCCGGGGCGCC 1320  
TGGCGAGGGCGGCCCTGCAGCCCCCTGGCCCTGGAGGGGGCCCTGCAGAAGCGAGGCATCGT 1380  
GGAGCAGTGTGCACCAGCATCTGCTCCCTTACCAGCTGGAGAATTACTGCAACTagggg 1440  
gcgcgggggggcaggacgtggcagcacctgtgtcaggtcacgggtggccgcaagccttcggc 1500  
tctctgcaccccaagtgattcaataaaaccctctgaatgagccctagtgggtgttgtctgtg 1560  
cggcgcaggggttgaggtgtgggccaggggcccttcagggacccccctgtttcctgttct 1620  
ctctacaccagccccccaaacttggccccctgccactgggtgcccaaaggcaagctcct 1680  
gcccattgtcggctgccccctgccccaggtgctaggaggtcccatggccccctggctgctt 1740  
cagcacaggacagagaaggacctcgggggctctggggggcctgggggctggttgggagca 1800  
gatttccttagctcaaaagccttgagctctcggggggccccagagagctttccaaggaaa 1860  
gggaataagaatattccagcagtttctccccacctaggcctgcagtgctctacctgtagt 1920

From dog genome sequence. Chromosome 18:49348635–49346757 (Complement)

Polymorphisms at 630 (t/c), 657 (t/g), 1008 (A/C), 1033 (insertion after -/GCCCTGCC), 1225 (c/t), 1233 (g/a), 1904 (a/g)

ATG start codon at 843–845.