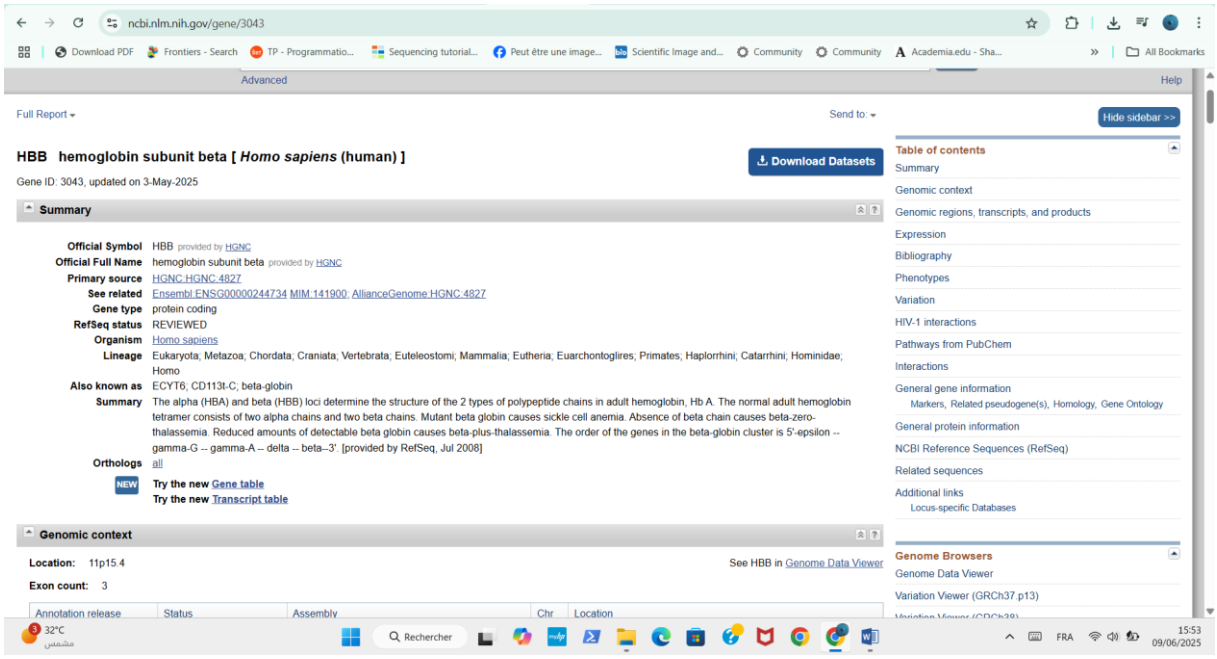
BIOINFROMHER MODULE 2 PROJECT

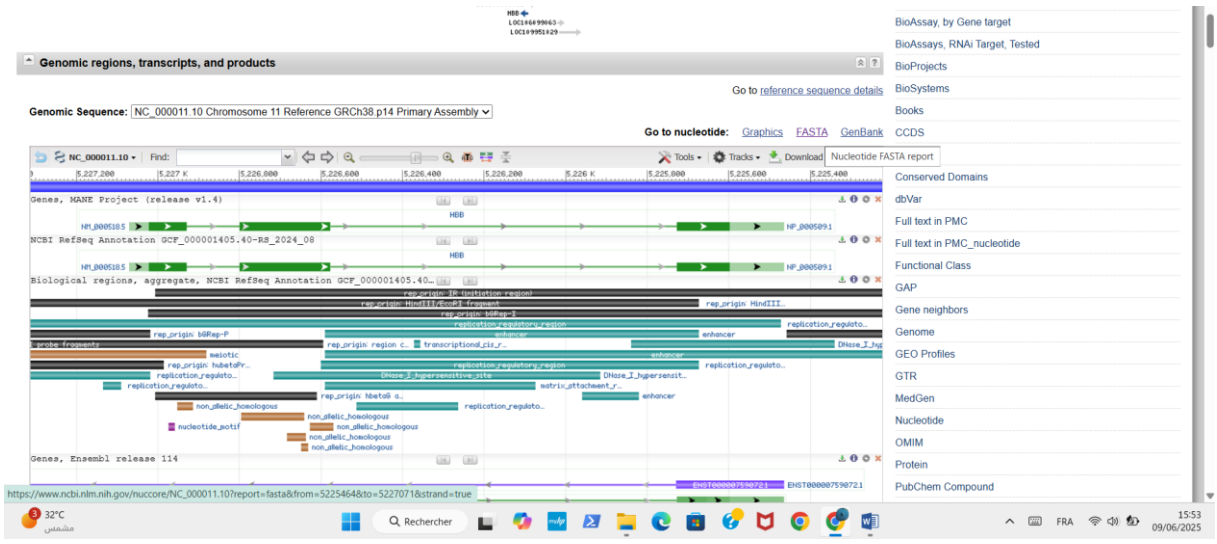
Tracking the Evolution of the Hemoglobin Beta (HBB) Gene across Species

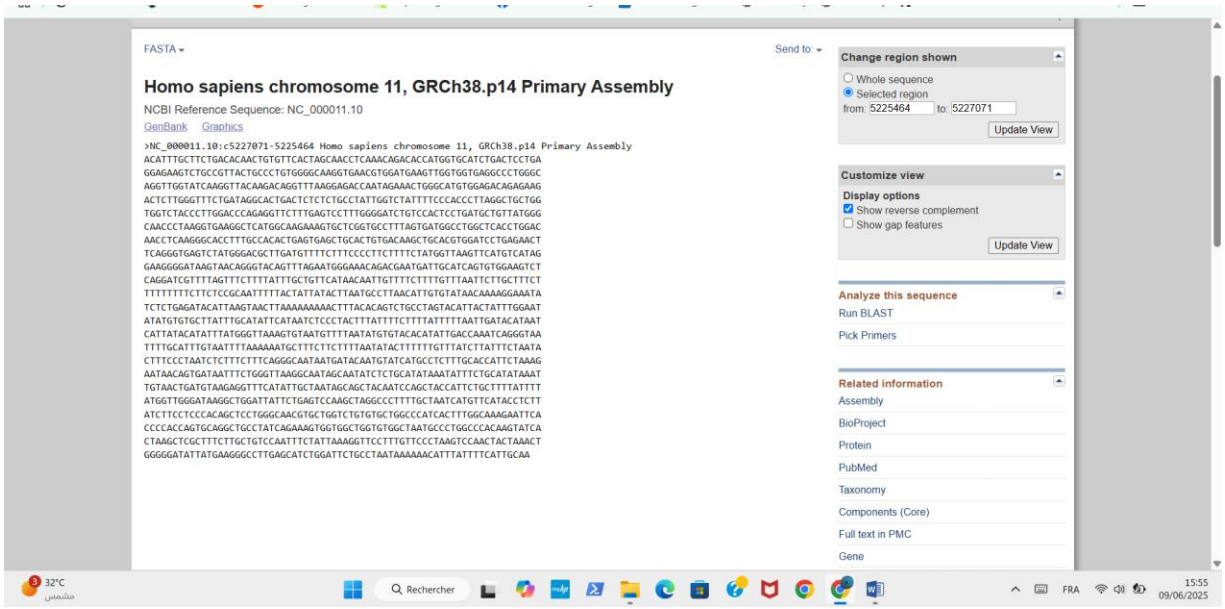
CONDUCTED BY :
Maram Nhaili

1 Sequence Retrieval and blast search:

1.1 Sequence Retrieval:







HBB for humain:

```
>NC_000011.10:c5227071-5225464 Homo sapiens chromosome 11, GRCh38.p14
Primary Assembly
ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTGA
GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGC
AGGTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAG
ACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTCCCACCCTTAGGCTGCTGG
TGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGG
CAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGAC
AACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACT
TCAGGGTGAGTCTATGGGACGCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAG
GAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCT
CAGGATCGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCT
TTTTTTTTCTTCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATA
TCTCTGAGATACATTAAGTAACTTAAAAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAAT
ATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAAT
CATTATACATATTTATGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTGACCAAATCAGGGTAA
TTTTGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATA
CTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCCTCTTTGCACCATTCTAAAG
AATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATAAATATTTCTGCATATAAAT
TGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTT
ATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTT
ATCTTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA
CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCA
CTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACT
GGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA
```

1.2 Fasta formet for other species:

1.2.1 COW:

Accession number: M63453.1

>M63453.1 Bovine Beta globin gene and globin (PSI-3) pseudogene, complete cds

GGAGAATAAAGTTTCTGAGTCTAGACACACTGGATCAGCCAATCACAGATGAAGGGCACTGAGGAACAGG
AGTGCATCTTACATTCCCCCAAACCAATGAACTTGTATTATGCCCTGGGCTAATCTGCTCAGAGCAGAGA
GGGCAGGGGGCTGGGTGGGGCTCACAAGCAAGACCAGGGCCCCTACTGCTTACACTTGCTTCTAACACAA
CTTGCAACTGCACAAACACACATCATGGTGCATCTGACTCTTGAGGGGAAGGCTACTTGTCACTGCCCTG
CAGACGAAAATGAGGGTGGCTGAAGTTGGTGTTGAAACCTTAGGCAGGCAGGTATTCAGCTTACAAGGCA
AGGAGAGTGAATGTCAGCTGGGTGTGTGGGGACAGAGCCATTGCCTGGGATTCTGGCAGGCATTGACTCC
CTCCTTCCTTATGCTGTTTTCACCCTGTAGGCTTCTGGTTGTCTACCCCTGGACTCAGAGGTTCTTTGAG
TCCTTTGGGAACTTGCCCTCTGCTGATCTATTATGGGCAACCTAAGGTGAAGGCCCATGACAAGAAGGTG
CTAGACTCCTTTACCAAAGGCCTGAAGCATGTTGACCACCTCAAGGGTGTCTTTGCTTTGCTAAGTGAGT
TGCACTGTAAGAATCTGCATGTCAGTCCTGAGAACATCAGTGTGAGTCTACGGGATGCTTAATATTCTCC
ATCTATTTTTTTTCTTCTGGTGGTTAAGTTCCTATCATGAGGAGAGAGTTAAGCAGCAGGATACAGTTCA
GAATGGAAAAGAGATATTCTGGTTACATCACTATGGATTCCTCAGGAACATTTAGTTTCCTTTACTTTCT
TTGTTCCCAGCCATCATTTCCTCTTACCCAATCTTGTTTTTTTCTGTTTGTTCTTTACAGTATCTTCTTT
TTATTCAAACATTTTGAGTATTTAAAAACACTTTTATATTTTAAGTCACTTAAAATTTTATCTCATATTT
TCCCCTTACCTCTTCCTTTCAAAGCAAGGGAGACAAAATGATGCATTGTGTCTTGAAATGGTTCAAAAGA
ATAAAAAATGATAACAGGCTATGGACTAAGACAGAAAGGCAGAAACATTTCTAAGAACAAGTTCAGGCTG
CTATCAATTCAGTTCAGTCACTCAGTCGTGTCCAACTCCTTTGCAACCCCATGAATCATAGTGCGCCAAG
CCTCCCTGTCCATCACCAACTCCCGACTCTAGAGGATCCCCGGAGTTCACTCAGACTCACGTCCATTGAG
TCAGTGATGCCATCCAGCCATCTCATCCTCTGTCATCCCCTTCTCCTCCTGCCCCCAATCCCTCCCAGCA
TCAGAGTCTTTTCCAATGAGTCAACTCTTCGCATGAGGTGGCCAAAGTATTGGAGTTTCAGCTTTAGCAT
CATTCCTTCCAAAGAAATCCCAGGGCTGATCTCCTTCAGAATGGACTGGCTGATATGGGTGGGTTCATAT
TAGAAGTAACATCTATACTTCAGCCATCTTTCTACTTATATTCTAGGGGCACAGCTTGGGATGAGACTGA
AATACTCTTTAGTCTGAATTGGGTGCCTCTGCTAACCATGTCCTTGTTTTTTTATTCCTTCCACACAGCT
CCTTGGCAACATACTGGTAATTACACTGGCTCAAAACTTTGGCAAGGAATTCACCCCGGAGTTCCTGGCT
GCCTATCAGAAGGTGGTGGCTGGTGTGCTAATGCCCTCACCTAAAAATACCACTGGGATCCTGGCCATTT
TCTTTAAAAAAGGAGAAAATTTATTTTTAATTGATTGATGATTGGTTTACAATACTGGTTTGATCACTAT
CACATATTAACAGGAATTAATCATAGGTGTACATATATCCCCTCCCATTTGAATCTCCCTCCCATCTCCC
ACCAATTCACACTCCTCTAGGTTATTACAGAGCCCTATTTGAATTCCTTGAGTGGATGAACATTGTGATA
GCTTACAGTGAACAGTGTCAGATTTGTGGCCTCCAGAGGGACAATTTCGATCTGGGGCCAGAGAAGAGGC
TTGATCACTCAGAGCTTTTGCATGGCAGAAATTTTATTACAGTGAAAAACGGACAGAGGAAGCTTCTGAG
ATAGACATCAGAAGGGGAACTGTAGAGTGCCCTACTTGCTAGTCTTATCAAGACCTTATATACTTTTTCC
TTTGGTTACTAACAATAGAAAGGTCTTATTGAACCCTCTCTCACAACATACATTTTAAGAGAACAGGATT
AGTCAGAAGAATCTTAAGAAGGAGAACCAGGTCCTCAAGTAAGATGCATTGTTGTTATATAATCATTAGT
ACAAAGTTAAGGAAAAACATACCTTTGAGCAAGATGAATTGTTTTGTTGTTTAATCATTGAAGTGAAGTG
AAGTTGCTCAGCCATGTCCAACTCTTTTCGACCCCATGGACTGTAGCTTACAAGGCTCCCCTGTCCATGG
AATTTTCCAGGCAAGAGTACTGGAGTGGGTTGCCATTTCCTTCTCCAGGGTATCTTCCCCACCCAGGGAT
CTAACCTGGGTATCCCACATTGCATGCAGACAATTTACCATCTGAGCTACCAGGGAAGCCATGGTAGAAG
GTCAGCCTTATTCCTTGGACTGACTTTTTCAGCAATTAGAGTATATTTTGAGCCTAACTTTAAAAATAGG
AGGTTTGTTGCCAAACAGAAAAGGAATTTCTTAACACTTTGATGAGGTAAAAATAAGGATTGACATGTTT
ATGAACAGCATACTATGGCAAGACAGCATAGATAACCCTGTACTGGTCTGCAGGTCCTGAACACTGCACT
GTCTGGTTCCCAAATTCCATCTGTAAAAAAGCAGCTTCACTCATTTCTCCAAGCAAATCTTAGGAGCAAA
CGAGTCCACGTAAGGCTTCTGTCTCTGGTGTAACTAAAATGGAGGAAAATGGACTATGAAAGGGATCATC
TACTTTTTCAGAGAAGACAAACGATGGGAGGTAGGAAAGCATTATGTTTTAAAACGACTTGAGTCTTGTA
GGTTTCTGTTGGCCTCTCATGCTCTTCCTATTTGGAGAGTAAGATCTAATCACTTGATGTACATAAGAGG
TTTGGTAAAGATCAATCAGTCTTTCTCTAAGAGACTTGGCCACTGATAATACATCCACCACAAAGGAGAA
TACCTAAACCATGAATATATCCTGTACTGGCACCATCATTAAGAATAGGCTTGGAAACAGGACTTGATGC
CACTTCCTTCTTTGTCCTTCATCTAGTAAATAGGCCTCTTTGTGTCTGTGTTAGTTGCTCAGTCAGTGTC
CCATTCTTGGCAACCCCATGGAGTCTGTCTGGTTTCTTCTGCCATGGGATTCTCCAGGCAAGAATACTGG
AGTGGGTTGCCATTCCCTCCTTCAGGGATCTTCCCAACCCAGGCATCGAAGCCAGGTTTCCTACATTGCT
GATAGATTCTTTACTGTCTGAGCCAGCAAGGAAGCCCTTTACCTTTACTTTAACCACCTAGGTTATGCAT
GGGTTTAAAGAGCTCTGGGCTTAAAGAAAGTTAGTCTTGAGATGTTGTCACAACTCAGAGTTGAAGAAAA
ATGTCTTATGTGACTAAGACTTTGCAATGTAGAACAAAACATTTTTCCTTTTCTCCTCCTTGAGAGCTCC
AGATCCCTTGAGCACCCTGGATTCCTTATCAGTCTATCTAAGAATTGACTCTCCAATTGGAAGACTTGAT
TTCCCTAGATTCTTCTATTTTTTACAGTTTAATTTTTAAGTCATTTTTCAGTTGTGATATTGCTTCTGTT
GTTTTTTATGTTCCTGTTTTCAGGCCACAAGGCATGTCAGATCTTAGTCCCCTGAACAGAAATAGAACCC
TCACGCCCTGCATTACACTAAGCCTTAAACACTAGACTGTCAGGAAGGCCCCAGTATTTCCCCAGATGTT
ATCTTCTGAACTTGGGGAAATAAAATCCACCCTCAAAGGTATGGCTTCTGCCTGATAAATAAATTTCAGC
TCAATTTCCTGATTCAGTTTACTCATTTTCCCCCCTAAGGATATGGGAAAGTCCCTAAGGGTCTATTGAT
AGGGAGGCTTTGTGTCTAATATGAGAGGACAAGGGGAATGAGAAAAGGAAGGTACCTCACATAGTCAATA

ATGGATGAAGCTTCTGTCTCAAAGCACAAGAACTATTAAGAAGGTTTAAAATATAGACAGGATACTGGCA
TTGCAGGGAATCATGGAAAACTTCAAACCTAGGCAGAAACTTCAGGCTAATTCTGATCCATATCCACTAA
TTACTCAAGAAATAGTGTGCTCAATCAGAGACATAATGATTACTCAAGAAACCATTATTAACGAAACCCT
TTTTCATACTTCCCTTGCCTGCACATTGTAGGCAACGATGCTAGATTACATCCGATATTCCCAATCTAAT
GCATAATCTCCCTCTCCCACTTGCCCCAAATAGCCTCCAAAATTAACCCCAACCTCCCATTTCAGGTAAA
GCTCAGTGCTTGCATAACACTTGCTTGAATCTTTTGACCTTGTCATCAGTGAAATTTCCAGTATTGCCTA
AGGAAATTCTTCCAAACCCTCCCACTCCAAACTTTGCACATTTCTTCCTGTCTAGCATACTCACCCTGTG
TGCTGTATAAGGGTTTCTGCCTTTTATACTATCCTACATGATTACCATGTCTTATACAGCAAATCCACCA
TAGACAGTAAGTGTACGTTTATTAAATGAAATAATCAATGAATGAACAATTATTTCTTGTGTCAAAACTC
CTGGGAGTAGAGGACAACTTTTAAAAATAGGCTATTGTACAGTGTTTCTTTGGTTTAGTGCTACAATATA
CAATGAGGAAGCACACAGGAAAGTAAATCACTGATTAGATATTTGCTTTCTTGCTTATAGGTAGTACAAT
ATCTTTAGAAAATTTTTAACTTTTTATTTTATGCTGGCTGCTGCTGCTAAGTCGCTTCATCGGTGTCCGA
CTCTGTGCGACCCCAGAGATGGCAGCCCACCAGGCTCCTCTTTCCCTGGGATTCTCCAGGCAAGAACACT
GGAGTGGGTTGCCATTTCCTTCTCCAATTTCATGAAAGTGAAAAGTGAAAGTCAAGTCGCTCAGTCATGT
CCGACTCTTAGCCACCCCATGGACTGCAGCCTACCAGGCTCCTCAATCCATGAGATTTTCCAGGCAAGAG
TACTGGAGTGGGGTACCATTGCCTTCTCCATTCCAATTAAAACTGTTGTGATAGTTTCAAGTGCACAGCA
AAGTGACTAAACCATTTATATACATGTTTCCATTCTCCCCCAAACTCCCTCCCATCCAAGCTGGCACATA
TCATTGATCGACTTCCCTGTGCTATACAGTAAGACCTTGTTAGTTATCCATTTTAAAGATAGCAATGTGT
ATATGTCCATCCCAAACTCCCTAACTATCCCTTGTGAAATTATTTTTGTCATATGTAAAATTGGAAGGAC
TTCTAGAACCTTCTAACTTAAATAAACTGAAGTGAAGTTGCTCAGTCGTGTCCAACTTTTTGTGAGCCCC
ATGGACGGAGAGCCTGCACCAGGCTACTCATCCAGTGATTTTTCAGGCAAGAATACTGGAGTGGGTTGCC
ATTTGCATGCATTTCTTTAAATCTATATCATGTTCCTGTTGATTAAAAAATTGTTGACTACCTGAAACAG
AACAAGTTGCTTCCCAGGTTATAGACAATGCATTACAGTTCACTTTTTCACAACTATATTTTGAATTATC
TGAAGGCAGATGAGCTTTTAAATATTTATTTTTTTCTCATGTCATTATACATAATAGGCACTGAGAATCC
TTTACACCCCAACACTATGTTTAAAAACCACAGGGTCACATCACAGAATCAGAGACATTTCAGGAAAATT
TTGGGAAGTTTCCTTTCACATTCTTGAGCACAGAAACGTCATGATAAAATTAGACATTAGGAGACAAAGA
GATTTTCAAAATGTTTGAAAGAGCGTAAGGGGAGACTGTGAGTTTCACAATCTCAAATGAGAAAACAAAG
TACACTTCTTCCTTCTAAATGTTCATAAAAGTTTAAAAAATATCTGAAGAACTTACTGTATTCTCTTTTG
TTCTTTCACCTTTGGAACTGCCAATCAGTCTAGGCAAGAAGTTGGGACTGAAAATCCCAATACTGGAGAA
TTCAAGAGAACATCACTGGGGTGGGACTTGGATGAAGTTCAAACAGTTGCTGTATTGCTGGTGTGGTGTC
TCCACGTTGGGAGTGAGGAAGTCAAGGGACGGAAAGAGACTTGGAAGCTTGTTATATGTTTTCATTCCTT
GTTAAACTGGAGGATCTGTAAAAGTCTAAGAGTAAAATTAATTCCTTAATATTTGATTTTATTTGATAAT
TTATCCAATAGTTACTTCTAGTTCCTCCCTATCATGTTGCTTTCTAGATCATGTTCAATATGTGTTAATG
CAGATGGGGTATGTGTGGAAGAAAGAAAGATGGAAAGAGACCAGAGGCAAGAGAGAAGAACAGAGGCAGG
GAGGGAGACACAGAGAGTCAGAGACAGAGAAAGGGAGGGAAGAAGAGGAGAGGGGAGAAAGGAGAGGAGA
GAGAGAGAAGTAAGTAATGTTGTGCATTGATTTCCATAATCATTTTAAAAATCAGCCAGTTGAATGTCAT
GGGAATTTAGGATGCTATTCTAGATTTTCTATCTCTTAAGCCTCAGAGACGGCAATGTTGGCTTTGGGAA
TTAATGCCTTATGTATTTGTCAAGGGTGACCAGTCTGAAGCCACGGTTCTGCCTTTCCAGACATACAGCT
TCAAATTTTACTTAAGAAAGACTGCCCCTTTGTTCTTGATAAATCAAGACCATTGCCCATTTTTGAGGTA
AATAAAAGCTAATTCTGAATGAATTAAAATCTGAGCCCAGTGGGGGAAATATTAATATGGAATTGATAAA
CAGACACTAACTGGTCTGTAATAATTAGGATATATTTTTTCAAAATTATGGCTATGTTGATTTCACTGTT
TTTAGAAAGAAATTTGACCTTTTTAAACATGAATTAGACTATAACATTCCTAAACTATTCTGTTTATGAC
ATCTGTTTCTATGAGTTCTGCTCTTCATGAAGCATTTACTTATTTTTTTGTTGTTTTATTTTATTTTAGT
TTATTTTTGAAACAGTGCTAAAACTGAGGTAACTAAGATATTGTACATCTGGGGAAGAAGAGTAAATAAT
AACTGCTTTATCTTGTAATAAATGCGAATGGAAAATAACCAAAGAAAATAGCTACTTTACCCAGAAAAAT
ATAGATTAGATTGCTGAAATATTCTACTTGGATAATTAAAACTTGAACAGCAGAGCATGTACAAAAGAAT
TCTAGCTGAATGAAATTTTTGAACCAAATGATGAAACCACACATCTCTGGATGAAATGTCATGATTCTCC
TGAAGAAAGTCAATATTCATCCTCTTAATTATACTCAGAATAGGCCCAGCCTGGTATAGAACATTAGGAA
TTAGGTCACCTTGCCAGACTGCAATAAAGGAAAAGAAAAAAAAAAAAAGTCTATCTCTCTCTCTGCGGACT
TGTGTGAGTTTGTGCATGTGTGTGTATGCTCAACAGGGCTTTAATCAAAGTAGAGCAGACATGCGGTAGA
TCCCTTCACAGAGTCCAAGAAAGCTCACGGTGGTGAAGATGGTAGTGGGCCTAAGAAGCATTCCAGAACT
TTTTAGTGGCGACCTGGATTTGCACTTTAGGGTCTGAACATAACCCCATTTATATAACTATTCACTCTTC
CTTATTTTTGCCTGAACTTTTGATGTCTGATAACAACTTTCAGACTTGTCCCTGCAGAGTTAGAAGGAAA
AACCCATTTCACTCTTTAATCATACCAAGTAGCAAGCGTGAAAGATAGCAACAACTCTTGCTCTGAGACT
TGCATGATGGTGGCTGCAGCAGCAATCAGTACAAAAGAGGGACTCATAGACCATTCATAGCAAAGGAAGA
TATAAAAAAGAAGACCCAGTAGAAGAATTTGTACATGTATACACACACACACACACACACACACACACAC
ACACACGTATACTGGAATATTACATATGACTCAGGTATAAAAAAGAATGAAATTTTGCCAGTTGCAACAA
CTTGGATGAACCTGAAGGATATTATGCTTGGTGAAATGTCACACAAAGAAAGACAAAATCTTGTATGTTG
CAACTTTTACGTGGAACACAAAAAGCAAAATAAACAAATTAATATAACAAAGCAGAAACAAACTCACAGG
TATAGAGAACAAACTAGTATTTACCCATGGGGAGAGGGGAGGGTCAGGATAAGGGTAGGGGATTCAGAAA
TACAAACTATTGTGTATAAAGTAGATTAGCAACAAGGATATATTGTACAGAACAGAAAACATAGTCAGTA

ATTTATTGCAACTTTAAATGGTGCATAGTCCTTAAAAACACTGAATCACTATGCTGTGCACCTGAAACTA
ATATAATATTCTAGATTAAATATTTAAAAATAATTAATTTAAAGCAAATGAATTAGGGAAATCATCTCTG
AAAGCAGAAAAGGATAGGGGAAGTCTTCAAGTACATGCAAATTTCTGTCAGTCAAGACAGCAAAATTGTA
GCCCTGACAATTTGACTCCTTTTACAAGAAAATATAGGAATAGACTATACTACTATTTAGTAAATTCTGT
GAGTTTTCCATAACACAAACATCCTGAAGAGGCATGTCCAAAGTTTCCCAGAAGAAGTCCTTTCACTCAC
ATATGCCTGACCTCTGATATCAGGGATGTGAATAGGGTGTTTAAAAGAGTACAAATATAAAACAATGATA
AAGCCAATGGAGATTTCAATTTAATCCATTGAAATGCAAAACCTAAAGGTCCAACCAAAAGTAGCTTTTT
GTTTTTAACATTTAAAACACTTGAAAGATTTCTTTTCCCAGAAGTCAACATATGGAAATTTTTTTCTAAG
TATAGACCAATAGTTAGAGCCTCTTTGTTTGAGGGTTGCAGAAAAGAGTGTATATAAATTGAAAATGTTT
AAATTCTATATTTTTTACCCCTATAATCTTTTGTGGGCATATAATTCTGCCCTAACTTCCAGAAAGCTAG
CAAGAGCCAAGACTTTTCTACATGTCTGAAAATATCAGAATGAATTTGTTCACAAGTACTTGAATGGGAA
TGGGCCAGGTTAAAGGGGTGAAGAATCTTCATTTACCTAGCAATGATAAAAAAAAAATATTTTGGAGGATG
GAGTAACGAAAGCTGGACATTTATGAATTTCTTTGTAACAACTAGAAATGCAGAAGTTCTTTCCTTCTAG
TTCACCTAAAAATATATTAGTTGAGTTACAGAGACATAGAGCTGAAAGAATGAGCATTTTTTTAAACATG
TTTTGCCAGTGATGGCAAAGGTGAGAGAGAGGTGATTCTAGACATCAAGTTGTCTGAATCAGGAGAAGGT
TCTAAGGAGAGAGATCATGGAAAGTAAATATATATGTTCTTTACTGAAAGGAAGTAGAAATTGAAGAAGG
GAAGATCAGAAAGTACAGATGGTAGAAAGTAGGGGAACCTCAGGGAACTCTGAAGTTCTGCTGAGCTCAG
AAATGCTTGCTGCTTTTGCTTTTTTCATTGAATCTTCTAGTAGTAAGAAACAGACCTCTATACCTCAGAA
GTTTTTTTTTTGCCTAAGAAGGTAAGAGGATTTATTATGTGGGTGGGAGGAGGGGCAAAGAGAGAAGAAA
TATTTGAAAAGCTGTAGATAATGAAGAATAAATTTTAGAGAAAGAAGGTATAAATAAAGTATATAGAGAA
AACAAAATTATGGATGAATAAAATAGAAGTAATTCTTATGAACACAGTATTCTTAGAGGCTTAGAGAAAG
AAAACAAATAATCCGATTCAAGCTGGACTGTCTTTCCACAATTTTATCCCTCCTTTTAAATCATGGAGAA
GCTTCTCTGTTCTTCCAGACACTTTCCCAGGTCAGTCCATGAAGAAACAAGTTATATGCCTCTACTAGTC
TTTGAACACATTACCTCTCTCGGGGTTATTGAAACCTTGGGGTTATGATTGATTTTTCATTTATTTGTAT
TTCTTACTACATGGAGAGGTCTCTATGTTTTCACATATGGATTAATGGAGTAAGTACTTCACAGGACCTC
TCGATCCATACTTATTCTATTTTTCTATCATTATTTATTTGGCCATAAATTAAATTAAGAAAAAATTAAA
TGAGTAGATAAGCAAATGAATATTTGTTTTCATACCAGAACGATTTAATCCCAAACAAGCAAACAAAAGA
GATGCGTATTTAGAACAAGGGCAGAGGTTGTATCCAGGTTGTACTTGTAATTCTTTTGCATATTCTGAAG
GCACAGGAGATGATCCATCCACATAGTCTTAAGTCGAATCATAGTGGACAAATCCTTTCCATTTTCTGGA
GCCCGGATTTTTCATCTGTATAATAAGAAAATTGAGGAGGTGGTATCCAAGAGATTACTTCGTTGCTGAT
TCTTAAATCTCTGTGAGCAAACTTGCCAAGGAAGATGATTTTAATAGCAATTTGTATTGCTGGAATGACT
GTGATCTAGAGATGCCCAGAAAGAGGGCTGATGGTCTAAAGTCAGTGCCAGGAAGACCAAGGAGAGCTAT
GACTATCATCGTTCAAGCCTCACCCTGTGGAACCACAACTTGGCATGAGCCAATCTGCTCACAGAAGCAG
GGAGGGCAGGAGGCAGGGCTGGGCATAAAAGGAAGAGCTGGGCCAGCTGCTGCTTACACTTGCTTCTGAC
ACAACCGTGTTCACTAGCAACTACACAAACAGACACCATGCTGACTGCTGAGGAGAAGGCTGCCGTCACC
GCCTTTTGGGGCAAGGTGAAAGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTAGGTATCCCACTTAC
AAGGCAGGTTTAAGGAGAGTGAAATGCACCTGGGCGTGTGAGGACAGAGCCGTCCCTGAGATTCAGAGAG
CTGCTGGCTTCCTCTGACCTTGTGCTGTTTTCTCCCCCTAGGCTGCTGGTTGTCTACCCCTGGACTCAGA
GGTTCTTTGAGTCCTTTGGGGACTTGTCCACTGCTGATGCTGTTATGAACAACCCTAAGGTGAAGGCCCA
TGGCAAGAAGGTGCTAGATTCCTTTAGTAATGGCATGAAGCATCTCGATGACCTCAAGGGCACCTTTGCT
GCGCTGAGTGAGCTGCACTGTGATAAGCTGCATGTGGATCCTGAGAACTTCAAGGTGAGTTTGTGGAATC
CTCAGTGTTCTCCTTCTTCTTTTTATGGTCAAGCTCATGTCATGAGGAGAAAGCTGAATGGCAGGACACA
GTTTAGAATGGAGAAGAGGTATTCTGGTTAGATTACTAAGGACTCCTCAGAACCGTTTAGACTCTTTTAA
CCTCTTTGTTCACAACCAGTATTTCCTCTGATTCATTCTTGTTCTCTGTTGTCTGCAATGTCCTCTTTTT
AATTATATTTTTTATTTTGAGGGTTTAATTTGAAAAAAAAATTATATATCAACTTTAAAAATTGTATCTA
ATATTTCCCCCTTATCTGTTCCTTTCAAGGAATAAGATGTTCTATTGCTTTTTGAAATGATTCAAAATAA
TAAAAATAATAACAAGTTCTGGATTAAGTTAGAAAGAGAGAAACATTTCTAAATATATATTCGGGAAGAT
ATAGGTAGATTCACATCAGTAGTAACAACTTCACTTCAGTCATCTTTGTGCTTATATCTACGGTCACAGC
TTGGGATAAGACTGAAATACCCTGAATCTAACCTTGGATTTCCCTCATAGCTCAGTTGGTTAAGCATCTG
CCTGCAATGCAAGAGATCCCAGTTCGATTCCTGGGTCGGGAAGGATGGCTGGAGAAGGGATAGGCACCCA
CTCCAGTATTCTTGGGTTTCCCTTGTGGCTCAGCTGGTAAAGAATCTGCCTGCAATGTGGGAGACCCAGC
TTCTATCCCTGAGTTTGGAAGATCCCCTGGAGAAGGGAAAGGCTACCCACTCCAGTATTCTGGCCTGGAG
AAATCTATGGACTGTAGAGTCATGGGGTTGCAAAGAATCAGACACGATTGAGAGACTCTCACTTCACTCA
CCTGCACTAACCCTGCCCTTGCTTAATGTCTTTTCCACACAGCTCCTGGGCAACGTGCTAGTGGTTGTGC
TGGCTCGCAATTTTGGCAAGGAATTCACCCCGGTGCTGCAGGCTGACTTTCAGAAGGTGGTGGCTGGTGT
GGCCAATGCCCTGGCCCACAGATATCATTAAGCTCCCTTTCCTGCTTTCCAGGAAAGGTTTTTTCATCCT
CAGAGCCCAAAGATTGAATATGGAAAAATTATGAAGTGTTTTGAGCATCTGGCCTCTGCCTAATAAAGAC
ATTTATTTTCATTGCACTGGTGTATTTAAATTATTTCACTGTCTCTTACTCAGATGGGCACATGGGAGGG
CAAAACACTGAAGACATAAAGAAATGAAGGGCTAGTCGAGACCTTGAGAAAATATATCAGTATCTTGGAC
CCCATGACAGCAGTGGTTGTAAATAGCTGATGTTATGGAAAACAGGCTTTGCTCCTTAGCCTTACTCTCC
CTTAAAGAATTC

1.2.2 Chimpanzee :

Accession number : X02345.1

>X02345.1 P.troglodytes beta-globin gene, exons 1-3
AATTCATGAGAACATCACCTGGATGGGACATGTGCCGAGCAACACAATTACTATATGCTAGGCATTGCTA
TCTTCATATTGAGGAGGAGGAGGTCAAGAGATGAAAAAAGACTTGGCACCTTGTTGTTATATTAAAATTA
TTTGTTAGAGTAGAGCTTTTGTAAGAGTCTAGGAGTGTGGGGGCTAAATGATGATACACATGGACACAAA
GAATAGATCAACAGACACCCAGGCCTACTTGAGGGTTGAGGGTGGGAAGAGGGAGACCATGAAAAAGTAC
CTATTGGGTATTAAGTTCATCACTGAGTGATGAAATAATCTGTACATCAAAACCCAGTGATATGCAATTT
ACCTATATAACTTGTACATGTACCCCCAAATTTAAAATGAAGTTAAAACAAAGTATAGGAGTGGAATTAA
TTCCTCAAGATTTGGCTTTAATTTTATTTGATAATTTATCAAATGGTTGTTTTTCTTTTCTCACTATGGC
GTTGCTTTATAAACTATGTTCAATATGTCTGAATGAAAGGGTGTGTGTGTGTGTGTGGAAGAGAGGGAGA
GAGGAAGGGAAGAGAGGACGTAATAACGTGAATTTGAGTTCATGAAAATTTTTAAATAAAATAATTTAAT
GTCAGGAGAATTAAGCATAATAGTCTCCTAAATCATCCATCTCTTGAGCTTCAGAGCAGTCCTCTGAATT
AATGCCTACATGTTTGTAAAGGGTGTTCAGACTGAAGCCAAGATTCTACCTCTAAAGAGATGCAATCTCA
AATTTCTCTGAAGACTGTACCTCTGCTCTCCATAAATTGACACCGTGGCCCACTTAATGAGGTTAAAAAA
AAGCTAATTCTGAATGAAAATCTGAGCCCAGTGGAGGAAATATTAATGAACAAGGTGTAGACTGAAATAT
AAACTTTTCTGTAATAATTATGCATATACTTTAGCAAAGTTCTGTCTATGTTGACTTTATTGCTTTTTGG
TAAGAAATACAACTTTTTAAAGTGAACTAAACTATCCTATTTCCAAACTATTTTGTGTGTATGCGGTTTG
TTTCTATGGGTTCTGGTTTTCTTGGAGCATTTTTATTTCATTTTAATTAATTAATTCTGAGAGATGCTGA
GTTGTGTTTACTGAGAGATTGTGTATCTGCGAGAGAAGTCTGTAGCAAGTAGCTAGACTGTGCTTGACCT
AGGAACATATACAGTAGATTGCTAAAATGTCTCACTTGGGGAATTTTAGACTAAACAGTAGAGCATGTAT
AAAAATACTCTAGTCAAGTGCTGCTTTTGAAACAGATGATAAAACCACACTCCCATAGATGAGTGTCATG
ATTTTCATGGAGGAAGTTAATATTCATCCTCTAAGTATACTCAGACTAGGGCCATTCTGATATAAAACAT
TAGGACTTAAGAAAGATTAATAGACTGGAGTAAAGGAAATGGACCTCTGTCTCTCTCGCTGTCTCTTTTT
TGAGGACTTGTGTGTGTGTGTGTGTGTTGTGGTCAGTGGGGCTGGAATAAAAGTAGAATAGACCTGCA
CCTGCTGTGGCATCCATTCACAGAGTAGAAGCAAGCTCACAATAGTGAAGATGTCAGTAAGCTTGAGTAG
TTTTTCAGGAACTTTGAATGCTGATTTAGATTTGAAACTGAGGCTCTGGCCATAACCAAATTTGCACTAT
TTATTGCTTCTTGAAACTTATTTTTGCCTGGTATGCCTGGGCTTTTGATGGTCTTAGTATAGCTTGCAGC
CTTGTCCCTGCAGGATATTATGGGTCATAGAAAGAAAAGTCTGCATTACACTCTAGTCACACTAAGTAAC
TACCATTGGAAAAGCAACCCCTGCCTTGAAGCCAGGATGATGGTATCTGCAGCAGTTGCCAACACAAGAG
AAGGATCCATAGTTCATCATTTAAAAAAGAAAACAAAAAAGAAAAAGGAAAACTATTTCTGAGCATAAGA
AGTTGTAGGGTAAGTCTTTAAGAAGGTGACAATTTCTGCCAATCAGGATTTCAAAGCTGTTGCTTTGACA
ATTTTGGTCTTTCAGAATACTATAAATATAACCTATATTATAATTTCATAAAGTCTGTGGATTTTCTTTG
ACCCGGGATATTTGCAAAAGACATATTCAAACTTCCGCAGAACACTTTATTTCACATATACATGCCTCTT
ATATCAGGGATGTGAAACAGGGTCTCGAAAACTGTCTAAATCTAAAACAATGCTAATGCAGGTTTAAATT
TGATAAAATAAAATCCAAAATCTAAAAGCCAAGTCAAATCTGTATGTTTTAACATTTAAAATATTTTAAA
GACGTCTTTTCCCAGGATTCAACACGTGAAATCTTTTCTCAGGGATACACGTGTGCCTAGATCCTCATTG
CTTTAGTTTTTTACAGAGGAATGAGCATACAAAGAAAATACTTAAATTTTATCTCTCTTACCTCTATCAT
CATACACAGGCATAATTTTTTAACCTAGGCTCCAGACAGCCATAGAAGAACCAAACACTTTCTGCGTGTG
TGAGAATAATCAGAGTGAGATTTTATTCACAAGTACCTGATGAGGGTTGAGACAGGTAGAAAAAGTGAGA
GATCTCTATTTATTTAGCAATAATAGAGAAAGCATTTAAGAGAATAAAGCAATGGAAATAAGAAATTTGT
AAATTTCCTTCTGATAACTAGAAATAGAGGATCCAGTTTCTTTTGGTTAACCTAAATTTTATTTCATTTT
ATTATTTTATTTTATTTTGTGTAATCGTAGTTTCAGAGTGTTACAGCTGAAAGGAAGAAGTAGGAACATG
GAAAGTAAAAGTATAACACTTTCCTTACTAAAACGACAGGGGTTTCCAGGTAGGGGCAGGATTCAGAATG
ACTGACAGGGCCCTTAGGAAACACTGAGACCCTACGCTGACCTCATAAAGGCTTGCTACCTTTGCTGTTT
TAATTACATCTTTTAATAGCAGGAAGCAGAACTCTGCACTTCAAAAGTTTTTCCTCACCTGAGGAGTTAA
TTTAGTACAAGGGGAAAAAGTACAGGGGGATGGGAGAGAGAGATCACGTTGGAAAGCTATAGAGAAAGAA
GAGTAAATTTTAGTAAAGGAGGTTTAAACAAACAAAATATAAAGAGAAATAGGAACTTGAATCAAGGAAA
TGATTTTAAAACGCAGTATTCTTAGTAGACTAGAGGAAAAAAATAATCTGAGCCAAGTAGAAGACCTTTT
CCCCTCCTACCCCTACTTTTTAAGTCACAGAGGCTTTTTGTTCCCCCAGACACTTCTGCAGATTAGTCCA
GGCAGAAACAGTTAGCTGTCCCCAGTTAACCTCCTATTTGACACCACTGATTACCCCATTGTTAGTCACA
CTTTGGGTTGTAAGTGACTTTTTATTTATTTGTATTTTTGACTGCATTAAGAGGTCTCTAGTTTTTTATC
TCTTGTTTCCCAAAACCTAATAAGTAACTAATGCACAGAGCACATTGATTTGTATTTATTCTATTTTTAG
ACATAATTTATTAGCATGCATGAGCAAATTAAGAAAAACAACAACAAATGAATGCATATATATGTATATG
TATGTGTGTATATATACATACATATATATATATATATATATATATTTCTTTTCTTACCAGAAGGTTTTAA
TCCAAATGAGGAGAAGATATGCTTAGAACTGAGGTAGAGTTTTCATCCATTCTGTCCTGTAAGTATTTTG

CATATTCTGGAGACACAGGAAGAGATCCATCTACACATCCCAAAGCTGAATTATGGTAGACAAAACTCTT
CCACTTTTAGTGCATCAATTTCTTATTTGTGTAATAAGAAAATTGGGAAAACGATCTTCAATATGCTTAC
CAAGCTGTGATTCCAAATATTACGTAAACACACTTGCAAAGGAGGATGTTTTTAGTAGCAATTTGTACTG
ATGGTGTGGGGCCAAGAGATATATCTTAGAGGGAGGGCTGAGGGTTTGAAGTCCAACTCCTAAGCCAGTG
CCAGAAGAGCCAAGGACAGGTATGGCTGTCATCACTTAGACCTCACCCTGTGGAGCCACACCCTAGGGTT
GGCCAATCTACTCCCAGGAGCAGGGAGGGCAGAAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATC
TATTGCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTG
ACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGG
CCCTGGGCAGGTTGGTATCAAGGTTACAAGACAGGCTTAAGGAGACCAGTAGAAACTGGGCATGTGGAGA
CAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGGCTATTTTCCCACCCTTAG
GCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCT
GTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTC
ACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCC
TGAGAACTTCAGGGTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCA
TGTCATAGGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGT
GGAACTCTCAGGATCATTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCT
TGCTTTCTTTTTTTTTCTTCTCTGCAATTTTTACTATTATACTTAATGCCTTAACGTTGTGTATAACAAA
AGGAAATATCTCTGAGATACATTAAGTAACTTAAAAAAAATGTTTACGCAGTCTGCCTAGTACATTACTA
TTTGGAATATATGTGTGCTTATTTGCATATTCATAATCTACCTACTTTATTTTCTTTTATTTTTAATTGA
TACATAATCATTATACATATTTATGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTGACCAAAT
CAGGGTAATTTTGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTAT
TTCTACTACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCCTCTTTGCACCA
TTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATTTCTGCATATAAATATTTCTGC
ATATAAATTGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGTTACCATTCTTCT
TTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCA
TACCTCTTGTCTTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAA
AG

1.2.3 Zebrafish :

No significant similarity found

1.2.4 Chicken :

No significant similarity found

1.2.5 Mouse :

>EF605360.1 Mus musculus isolate Lima138A beta-globin (Hbbt1) gene, complete cds

GCTCACACAGGATAGAGAGGGCAGGAGCCAGGGCAGAGCATATAAGGTGAGGTAGGATCAGTTGCTCCTC
ACATTTGCTTCTGACATAGTTGTGTTGACTCACAACCCCAGAAACAGACATCATGGTGCACCTGACTGAT
GCTGAGAAGGCTGCTGTCTCTGGCCTGTGGGGAAAGGTGAACGCCGATGAAGTTGGTGGTGAGGCCCTGG
GCAGGTTGGTATCCAGGTTACAAGGCAGCTCACAAGTAGAAGTTGGGTGCTTGGAGACAGAGGTCTGCTT
TCCAGCAGGCACTAACTTTGAGTGTCCCCTGTCTATGTTTCCCTTTTTAGGCTGCTGGTTGTCTACCCTT
GGACCCAGCGGTACTTTGATAGCTTTGGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGT
GAAGGCCCATGGCAAGAAAGTGATAACTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGC
ACCTTTGCCAGCCTCAGTGAGCTCCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAGGGTGAGTC
TGATGGGCACCTCCTGGGTTTCCTTCCCCTGGCTATTCTGCTCAACCTTCCTATCAGAAGGAAAGGGGAA
GCGATTCTAGGGAGCAGTCTCCATGACTGTGTGTGGAGTGTTGACAAGAGTTCGGATATTTTATTCTCTA
CTCAGAATTGCTGCTCCCCCTCACTCTGTTCTGTGTTGTCATTTCCTCTTTCTTTGGTAAGCTTTTAATT
TCCAGTCGCATTTTACTAAATTAATTAAGCTGGTTATTTACTTCCCATCCTGATATCAGCTTCCCCTCCT
CCTTTCCTCCCAGTCCTTCTCTCTCTCCTCTCTCTTTCTCTAATCCTTTCCTTTCCCTCAGTTCATTTCT
TCTTCTTTGATCTACGTTTGTTTGTCTTTTTAAATATTGCCTTGTAACTTGCTCAGAGGACAAGGAAGAT
ATGTCCCTGTTTCTTCTCATAGCTCTCAAGAATAGTAGCATAATTGGCTTTTATGCCAGGGTGACAGGGG
AAGAATATATTTTACATATAAATTCTGTTTGACATAGGATTCTTATAATAATTTGTCAGTAGTTTAAGGT
TGCAAACAAATGTCTTTGTAAATAAGCCTGCAGTATCTGGTATTTTTGCTCTACAGTTATGTTGATGGTT
CTTCCATCTTCCCACAGCTCCTGGGCAATATGATCGTGATTGTGCTGGGCCACCACCTGGGCAAGGATTT
CACCCCCGCTGCACAGGCTGCCTTCCAGAAGGTGGTGGCTGGAGTGGCCGCTGCCCTGGCTCACAAGTAC
CACTAAACCCCCTTTCCTGCTCTTGCCTGTGAACAATGGTTAATTGTTCCCAAGAGAGCATCTGTC

1.3 Blast search :

Species	Acession Number	Identity
Cow	M63453.1	81.09%
Mouse	EF605360	79.61%
Zebrafish		No significant similarity found
Chicken		No significant similarity found
Champanzee	X02345.1	98.71%

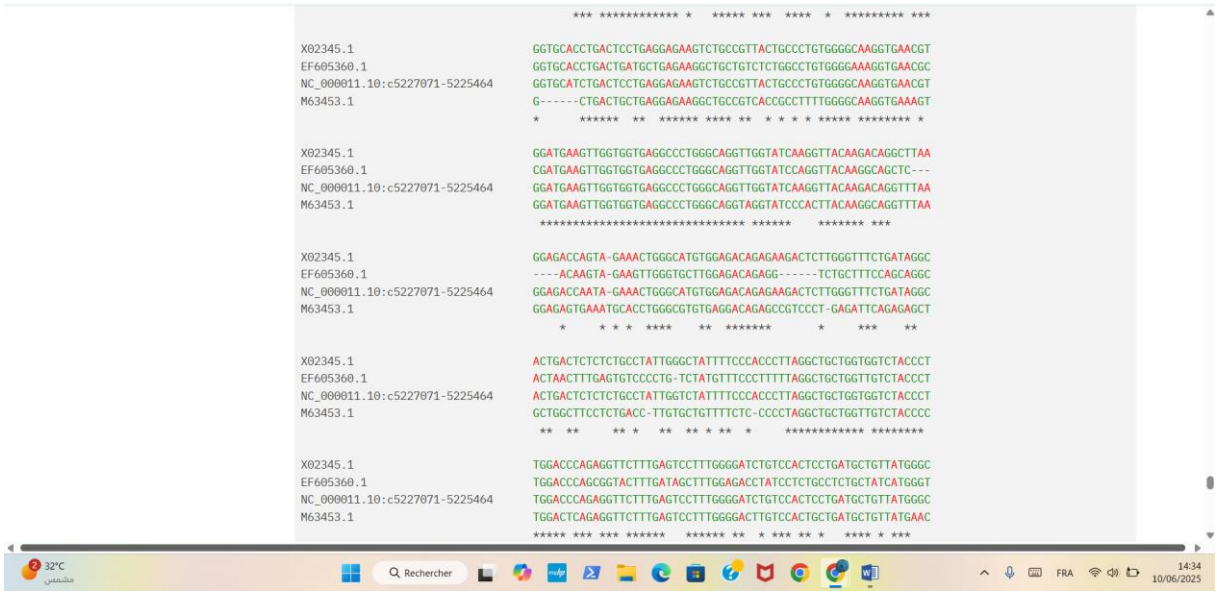
2 Pairwaise sequence Alignment :

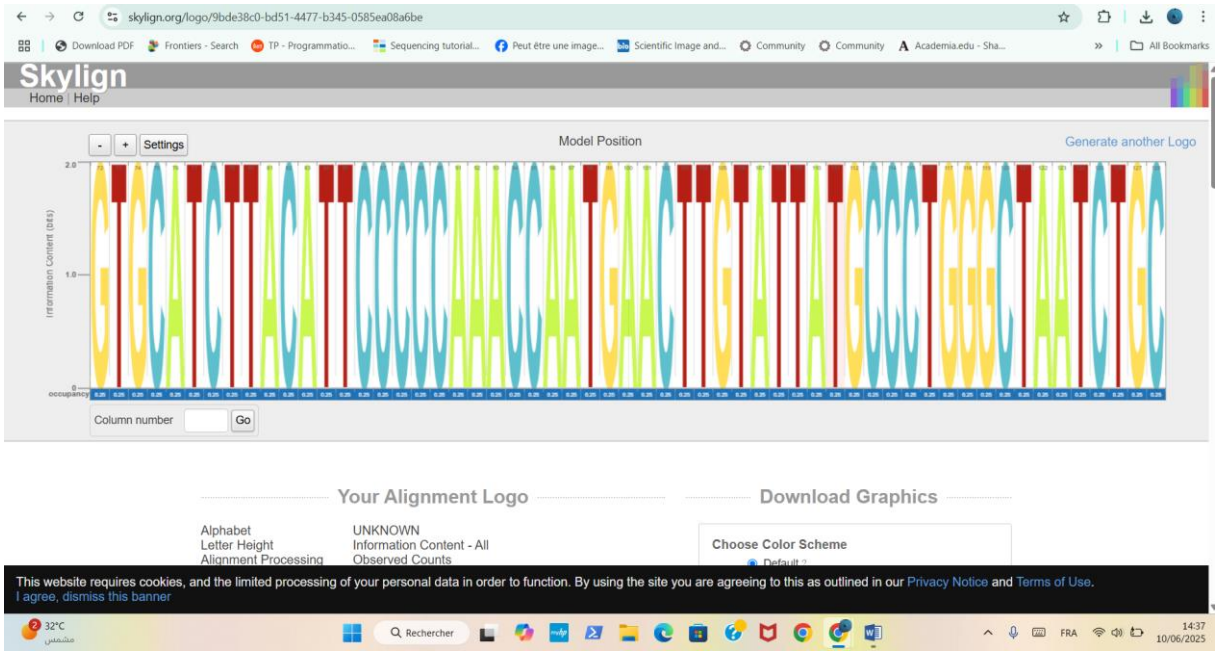
Species	Similarity	Identity	Number of gaps
Cow	8.9%	8.9%	89.1%
Champanzee	99%	99%	0%

Interpretation:

Based on the data provided from the table above we can confirm that the chimpanzee has more conserved regions that the cow because its similarity and identity (99%) is higher than the one found in the cow which represents 8.9%.Moreover, we can see that the number of gaps is higher in the cow that the chimpanzee.

3 Multiple sequence alignment:





The sequence logo shows that each columns correspond to a position in MSA (multiple sequence alignment).we observe high conserved regions due to the height of the nucleotide at each position where their total height correspond to 2 bits.

Additionally, there are highly conserved residues meaning there is little variation across sequences at those sites.

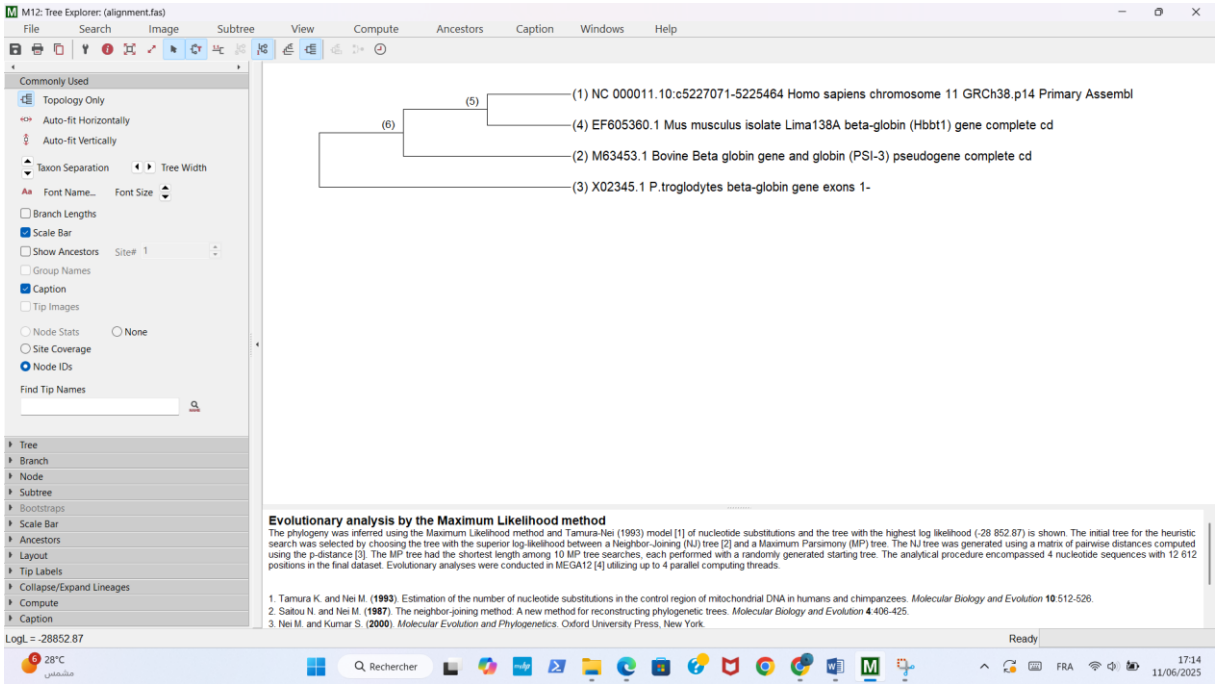
These conserved regions play critical role. They are functional importance where these regions could be :

- A promoter
- An enhancer
- A biding site for transcription factors or regulatory proteins

Structural stability

Evolutionary conservation

4 Phylogenetic tree construction:



- Human and Mouse beta-globin sequences cluster together more closely than with the other species, suggesting a more recent common ancestor between them for this gene region (which might reflect conserved functional regions rather than strict evolutionary distance).
- Chimpanzee (P. troglodytes) appears more distant, which is biologically unexpected given its close relation to humans. This suggests:

The chimpanzee sequence may be shorter (only "exons 1–"), or

There could be alignment gaps or divergence in the sequenced region used.

- Bovine (cow) gene shows greater divergence, as expected due to larger evolutionary distance from primates and rodents.

Does this tree match what you expect evolutionarily?

No, this tree does not match the expected evolutionary relationships of the species involved — especially not in the broad evolutionary sense.