Sequence of the salmon (*Oncorhynchus keta*) preproinsulin gene

A.P.Koval, A.I.Petrenko and V.M.Kavsan

Institute of Molecular Biology and Genetics, USSR Academy of Sciences, Kiev 252627, USSR

Submitted January 26, 1989

EMBL accession no. X13559

Preproinsulin gene were isolated from chum salmon λ EMBL 3 genomic DNA library. It possesses similar structural organization with the other sequenced insulin genes (1). The gene is split and contains two introns, the first intron within the 5'-untranslated region, the second intron within the C-peptide region. The 3'-untranslated region of the salmon insulin gene shares about 70% homology with the previously published cDNA sequences (2,3), thus suggesting the existence of at least two nonallelic insulin genes in salmon.

```

1 gtcgaccatcgcagggtcatcaatccctggaatcgctgtatatctcaga cacagggtcacatcagttatgtccgattcactggccctctgcttggtccac
101 taatgtcacttagtaagacaacatgcctaacgaggcgaaactgaatt gccacttaccatttgattcattgagagaattctcttttttcgctgaca
201 ctggcaagaccacaggccctgttccgagaccgtctgtctccaggggtat aaatggggggaggacctggtgttctgcctcaACTAATTATTCCTCAGCA
301 TTTTGTCTATTCTTCTACAGBCTTCTTACCACCACCACTCCATAGgc aagagcttgggggtctatgcgtttcacgtggaggattgggaacgacagta
401 ggggttagcgtggttagctcttggggatgtaactgcaactagaatatgaaa atgcctcagaactcttctggttttggtaacaatgtattctctatatac
501 tatgtactttacttcatctagaaccaataactgtagttttgttgatgta gggatattggtttattttagaacactgaatgacaatatatcaacaggat
601 atgggtcgtcagcattacacaaatgagttatatcatccaagaactaga atagagattgttctcatgcttgatgaaagctcccatagtgactgtaag
701 cagtgtgaacttatctgtttcattgtgttccctcatccagCCTACCATC ACCATG66CCTTCTG6CTCCAAGCTGCATCTCTGCTG6TGTGCTG6C6CT
801 CTCCCCG666TAGATGCTG6CAGCTG6CCAGCACCCTG6T66CTCTCACC T66T66AC6CCCTCTATCTG6T6T6T66AGAGAA66ATTCTTTTACCC
901 CCAAAG6AG6ATG66ATCCCCCTTATAGcgaagacaccttaacactgtc tgaattcaaaagacttcatagtctgcgacaactgaacttgatctaagac
1001 cacaagctactttactgtttttgtgtgtgtgaaatggcatgtaggaa actcttggctcactatgaataacaatgttgggaacaatgtataacctaca
1101 aatattgaatgaatgtgaaggcaatgatgagatacagcgaaggcatat acagtatatatctggtttatgtgtgtgctatagcactcactcaatatatt
1201 tttctctcttgcag66TTCTCTCTCCAAATCAGCAAG6AGAACBAA G6TACCCCTTCAAGACAGACG6AGATGATG6TAAG6AG66TATTGT
1301 AG6CAG6T6CTGTCAAGCCCTGCAACATCTT6ACCTGCAAACTACT GCAACT6AGACCTGTC66ACCACTTACCTG66TTCACCCACTG6CAG6C
1401 TCAATTTCTCAG6ACAGCATCTTACCTTCAAG6ATTGCAACATG6T6 T6CAACAG66AAG66TCTACAG666TTTATTTTATTTATTTATTT
1501 CCCAGAAAATAAGTTATATGAATTagcaaggactttgagtttctt tttgtaatggg

```

References

1. Steiner, D.F. et al (1985) Ann. Rev. Genet., 19, 463-484.
2. Kashuba, V.I. et al (1986). Mol. Biol. (Russian), 20, 845-852.
3. Sorokin, A. et al (1982) Gene, 20, 367-376.