

Sequence of a replication competent Hepatitis B virus genome with a preX open reading frame

Ivan Francisco Loncarevic, Hanswalter Zentgraf* and Claus Hobe Schröder

Institut für Virusforschung, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-6900 Heidelberg, FRG

Submitted May 9, 1990

EMBL accession no. X52939

The complete sequence of a replication competent Hepatitis B Virus (HPV) full length genome (p4a1;1) cloned via the unique XhoI restriction site from a hepatocellular carcinoma was 3215 base pairs long. As other sequenced HBV genomes it contains open reading frames (ORF) for the genes X (1247–1708), C (1774–2322), PreS/S (2586–705) and Pol (2180–1493). The genome can be attributed to the adr serotype and possesses an interrupted preC and an intact preX ORF, properties shared by 4 other HBV DNA clones from the same carcinoma (1). The presence of the preX ORF appears to be adr specific but not generally present on DNA of this serotype based on a comparison with published sequences (2, 3, 4, 5, 6).

REFERENCES

1. Loncarevic, I.F., Schranz, P., Zentgraf, H., Liang, X.-H., Herrmann, G., Tang, Z.-Y. and Schröder, C.H. (1990) *Virology* **174**, 158–168.
2. Fujiyama, A., Miyano, A., Nozaki, Z., Yoneyama, T., Ohtoma, N. and Matsubara, K. (1983) *Nucl. Acids Res.* **11**, 4601–4610.
3. Gan, R., Meijin, C., Lüping, S., Suwen, Q. and Zaiping, L. (1987) *Scientia Sinica B* **30**, 507–521.
4. Kobayashi, M. and Koike, K. (1984) *Gene* **30**, 227–232.
5. Ono, Y., Onda, H., Sasada, R., Igarashi, K., Sugino, Y. and Nishioka, K. (1983) *Nucl. Acids Res.* **11**, 1747–1757.
6. Rho, H.M., Kim, K., Hyun, S.W. and Yong, S.K. (1989) *Nucl.*

```

TCGAGGACTGGGGACCTGC ACCGAACATGGAGAGCACAA CATCAGGATTCTAGGACCC CTGCTCGTGTACAGGCGGG GTTTTCTTGTGACAAGAA TCCTCACAATACCACAGAT 120
CTAGACTCGTGGTTGACTTC TCTCAATTTCTAGGGGGAA CACCCAAGTGTCTGGCCAA AATTCGCAGTCCCAACCTC CAATCACTCACCAACCTTT GTCCTCCAACCTGTCTGGC 240
TATCGCTGGATGTGTCTGGC GCGTTTTATCATATTCCTCT TCATCCTGCTGCTATGCCCC ATCTTCTTGTGTGTTCTTCT GGACTACAAAGGTATGTTGC CCGTTTGTCTCTACTTCCA 360
GGAACATCACTACCAGCAC GGGACCATGCAAGACCTGCA CGATTCTGCTCAAAACACC TCTATGTTTCCCTCTTGTG CTGTACAAAACCTTCGGACC GAACTGCACTTGTATTCCC 480
ATCCCATCATCTGGGCTTT CGCAAGATTCCTATGGGAGT GGGCCTCAGTCCGTTTCTCC TGGCTCAGTTACTAGTGCC ATTTGTTCACTGCTGTCGAG GGCTTCCCCCACTGTTTGG 600
CTTTCAGTTATATGGACGAT GTGGTATTGGGGGCCAAGTC TGTACAACATCTTGAGTCCC TTTTACCTCTATTACCAAT TTTATGTTGCTTGGGCAT ACATTGAAACCTAATAAAA 720
CCAAGCGTTGGGGCTACTCC CTTAACTTCATGGGATATGT AATTGGAAGTTGGGGTACTT TACCACAAGAACATATTGTA CAAAAACTCAAGCAATGTTT TCGAAAACCTGCCTGCAATA 840
GACCTATTGATTGGAAGTA TGTAGAGAAATGTGGGTCT TTTAGGCTTGTCTGCCCTT TTACACAATGTGGCTATCCT GCCTTGATGCTTTATATGC CTGTATACAATCTAAGCAGG 960
CTTTCACCTTCTCGCCAACT TACAAGGCCCTTCTGTGTAA ACAATATCTGAACCTTTACC CCGTTGCCCGGCAACGGTCA GGCCTCTGCCAAGTGTGTC TGATGCAACCCCACTGGAT 1080
GGGCTTGGCCATGGCCAT CAGCGCATGCGTGGAACTTT TGTGGCTCCTCTGCCGATCC ATACTGCGGAACTCCTAGCA GCTTGTGTTGCTCGCAGCCG GTCTGGAGCGACTTATCG 1200
GAAACGACAACCTCTGTTGTC CTCTCTCGGAAATACACCTC CTTTCCATGGCTGCTAGGGT GTGCTGCCAAGTGGATCCTG CGCGGGACGTCCTTGTGCTA CGTCCCGTCGGCGCTGAATC 1320
CCGCGGACGACCCGTCGCG GGCCTTTGGGCTCTATCG TCCCTTCTTCTGCTGTGCTT TCCGGCGCACACGGGGCCG ACCTCTCTTTACCGCGTCTA CCGCTGTGCTTCTCATC 1440
TGCCGGACCCGTGCACTTC GCTTCACCTCTGCACGTCCG ATGGAGACCACCGTGAACGC CCACCAGTCTTGGCCAAAG TCTTATATAAGAGGACTCTT GGACTCTCAGCAATGTCAAC 1560
GACCGACCTTGAGGCATACT TCAAAGACTGTTGTTTAAA GACTGGGAGGAGTTGGGGGA GGAGATTAGGTTAATGATCT TTGTAAGTAGGAGCTGATAG CATAAATGGTCTGTCCACC 1680
AGCACCATGCAACTTTTTCA CCTCTGCCATAATCATCTCAT GTTCATGTCTACTGTTCAA GCCTCCAAGCTGTGCTTGG GTGGCTTAGGGCATGGACA TTGACCCGTATAAAGAATTT 1800
GGAGCTTCTGTGAGTTACT CTCTTTTTGGCCTCTGACT TCTTCCGCTGTGTCGAGT CTCCTGCACACCGCTCTGCT TCTCTATCGGGAGGCTTAG AGTCTCCGGAACATGTTCTA 1920
CCTCACCATACGCACTCAG GCAAGCTATTCTGTGTTGGG TTGAGTTGATGAATCTGGCC TCCGGCGCACACGGGGCCG AATTAGTAGTCAGCTATGTC GGAAGACCCAGCATCCAGGG AATTAGTAGTCAGCTATGTC 2040
AATGTTAATATGGGCTAAA AATCAGACAACATTTGCGT TTCACATTTCTGCTTACT TTTGGAAGAGAACTGTTTT AGAGTATTTGGTGTCTTTG GAGTGTGGATTCCGACTCCT 2160
CTCGTTACAGACCACAAA TGCCCTATCTTATCAACAC TTCGGAAACTACTGTTGTT AGACAACGAGGCGAGTCCCC TAGAAGAAGAACTCCCTCGC CTCGCAGACGAAGATCTCAA 2280
TCGCCCGCTCGCAGAAGATC TAAATCTCGGGAATCTCAAT GTTAGTATCCCTGGACTCA TAAGTGGGAACTTTACTG GACTTTATTTCTTACTGTA CCTGTCTTAACTCCTGAGTG 2400
GAAAACCTCTTTTCTTA ACATTATTACAGGAGGAC ATTATTGATAGATGTCAACA ATATGTGGCCCTCTTACAG TGAATGAAAAAGGAGATTA AAATTAATTATGCTGTAG 2520
GTTCTATCTAACCTTACCA AGTATTTACCCTTGGATAAA GGCATTAACCTTATTATCC TGAACATGCAGTTAATCATT ACTTCCAAACTAGGCATTAT TTACATACTCTGTGGAAGGC 2640
TGGCATTCTATATAAGAGAG AAATACAGGCACTGCTCA TTTTGTGGGTCAACATATTC TTGGGAACAAGAGCTACAGC ATGGGAGGTTGGTCTTCCAA ACCTCGACAAGGCATGGGGA 2760
CGAATCTTCTGTTCCTAAT CCTCTGGGATCTTTCCCGA TCACCAGTTGGACCTGCGT TCGGAGCAACTCAAACAAT CCAGATTGGGACTTCAACCC CAACAAGGATCACTGGCCAG 2980
AGCCAAGAGGTTAGGAGCG GCAAGCTTTCGGGCCAGGGTT CACCCACACACACGGCGGCC TTTTGGGGTGGAGCCCTCAG CACTAGGGCACATTGACAAC AGTCCAGCAGCGCCTCTCTC 3000
CTGCTCCACCAATCGGCAG TCAGGAAGACAGCCTACTCC CATCTCTCCACCTCTAAGAG ACAGTATCTCTCAGGCCATG CAGTGGAACTCCACAACATT CCACCAAGCTCTGCTAGATC 3120
CAAGAGTGAGGGGCTCTAT TTCCCTGCTGGTGGCTCCAG TCCGGAAACAGTAAACCTG TTCCGACTACTGCCTCACCC ATATCTGCTAATCTTC 3215

```

* To whom correspondence should be addressed