Complete nucleotide sequence of IC10, a retrovirus containing the R*mil* oncogene transduced in chicken neuroretina cells infected with avian retrovirus RAV-1

Alain Eychène, Maria Marx, Philippe Dezélée and Georges Calothy

Institut Curie-Biologie, Bâtiment 110, Centre Universitaire, 91405 Orsay, France
Submitted January 5, 1989

We report the complete nucleotide sequence (4213 bp) of IC10, an acutely mitogenic retrovirus which contains a novel oncogene named v-R*mil* (1). This transduced gene is a member of the *mil/raf* family. IC10 virus contains a single long open reading frame of 1079 amino acids   coding for a Δ *gag*-R*mil*-Δ *env* polyprotein with a calculated molecular weight of 116,946 daltons. This protein is composed of 645 amino acids of the N-terminal part of *gag* fused with 367 amino acids of v-R*mil*. The last 67 amino acids and the stop codon are provided by RAV-1 *env* sequences fused in frame with v-R*mil* sequences. IC10 contains RAV-1-derived sequences that have not been yet described. Therefore we compared these sequences with the published sequence of RAV-2 (1302bp), including the end of *env*, the beginning of *gag* and the complete LTR (2). This comparison extends from nucleotide 1 to nucleotide 593 and from nucleotide 3489 to 4213. The nucleotides differences found in RAV-2 are indicated below the IC10 sequence. The major divergence is located in the U3 region where a deletion of 11 nucleotides (indicated by asterisks) is found in RAV-2 sequence when compared to IC10. In addition, a deletion of one nucleotide at the end of IC10 *env* sequence results in a 22 bases shift of the stop codon position relative to RAV-2.

```
                                                  U3◄─┐                          ┌►U5
             GA ATTCCGCATTGCAGAGATATTGTATTTAAG TGCCTAGCTCGATACAATAAATGCCATTTG ACCATTCACCACATTGGTGTGCACCTGGGT      92
                                                U5◄─┐                     C
TGATGGCCGGACCGTTGATTCCCTGACGAC TACGAGCACATGCATGAAGCAGAAGGCTTC ATTTGGTGACCCCGACGTGATCGTTAGGGA ATAGTGGTCGGCCACAGACGGCCGTGGCGAT     212
                                   C
CCTGTCCTCATCCGTCTCGCTTATTCGGGG AGCGGACGATGACCCTAGTAGAGGGGGCTG CGGCTTAGGAGGGCAGAAGCTGAGTGGCGT CGGAGGGAGCTCTACTGCAGGGAGCCAACA     332
                                                                        C         G
TACCCTACCGAGAACTCAGAGAGTCGTTGG AAGACGGGAAGGAAGCCCGACGACTGAGCG GTCCACCCCAGGCGTGATTCCGGTTGCTCT GCGTGATTCCGGTCGCCCGGTGGATCAAGC     452
ATGGAAGCCGTCATAAAGGTGATTTCGTCC GCGTGTAAGACCTATTGCGGGAAAACCTCT CCTTCTAAGAAGGAAATAGGGGCTATGTTG TCCCTGTTACAAAAGGAAGGGTTGCTTACG     572
└►gag
TCCCCCTCAGACTTATATTCCCCGGGGTCC TGGGATCCCATTACCGCGGCGCTCTCTCAG CGGGCTATGGTACTTGGAAAATCGGGAGAG TTAAAAACCTGGGGATTGGTTTTGGGGGCA     692
TTGAAGGCGGCTCGAGAGGAACAGGTTACA TCTGAGCAAGCAAAGTTTTGGTTGGGATTA GGGGGAGGGAGGGTCTCTCCCCCAGGTCCG GAGAGCATCGAGAAACCAGCAACGGAGCGG     812
CGAATTGACAAGGGGGAGGAAGTGGGAGAA ACAACTGTGCAGCGAGATGCGAAGATGGCG CCGGAGGAAACGGCCACACCTAAAACCGTT GGCACATTCTGCTATCATTGCGGAACAGCT     932
ATTGGCTGTAATTGCGCCACAGCCTCGGCT CCTCCTCCTCCTTATGTGGGGAGTGGTTTG TATCCTTCCCTGGCGGGGGTGGGAGAGCAG CAGGGCCAGGGGGGTGACACACCTCCGGGG    1052
GCGGAACAGCCAAGGGCGGAGCAACAGCAT GCGGGTCAGGCTCCTGGGCCAGCCCTGACT GACTGGGCAAGGGTCAGGGAGGAGCTTGCG AGTACAGGTCCGCCCGTGGTGGCCATGCCT    1172
GTAGTGATTAAGACAGAGGGACCCGCCTGG ACCCCTCTGGAGCCAAAATTGATCACAAGA CTGGCTGATACGGTTAGGACCCAAGGGCTTA CGATCCCCGATCACTATGGCAGAAGTGGAA    1292
GCGCTTATGTCCTCCCCGCTGCTGCCGCAT GATGTTACGAATCTAATGAGACTTATTTTG GGACCTGCCCCATATGCCTTGTGGATGGAC GCTTGGGGCGTCCAACTACAGACGGTTATA    1412
GCGGCAGCCACTCGCGACCCCCGACACCCA GCGAATGGTCAAGGGCGGGGGGAACCGACT AACTTGGATCGCTTAAAGGGTTTAGCTGAT GGGATGGTGGGCAACCCACAGGGTCAGGCC    1532
GTATTATTAAGACCGGGGGAATTGGTTGCT ATTACGGCGTCGGCTCTCCAGGCGTTTAGA GAAGTTGCCCGGCTGGCGGAACCCGCAGGT CCATGGGCGGATATCACGCAGGGACCATCC    1652
GAGTCCTTTGTTGATTTTGCCAATCGGCTT ATAAAGGCGGTTGAGGGGTCAGATCTCCCG CCTTCCGCACGGGCTCCGGTGATCATTGAC TGCTTTAGGCAGAAGTCACAGCCAGATATT    1772
CAGCAGCTTATACGGGCAGCACCCTCCACG CTGACCACCCCAGGAGAGATAATCAAATAT GTGCTAGACAGGCAGAAGACTGCCCCTCTT ACGGATCAAGGCATAGCCGCGGCCATGTCG    1892
TCTGCTATCCAGCCCTTAGTTATGGCAGTA GTCAATAGAGAGAGGGACGGACAAACTGGG TCGGGTGGTCGTGCCCGAGGGCTCTGCTAC ACTTGCGGATCCCCGGGACATTATCAGGCG    2012
CAGTGCCCGAAAAAACGAAAGTCAGGAAAC AGCCGTGAGCGATGTCAGTTGTGTGACGGG ATGGGACACAACGCTAAACAGTGTAGGAAG CGGGATGGCAACCAGGGCCAACGCCCAGGA    2132
AGAGGTCTCTCTTCGGGGCCGTGGCCCGGC CCTGACCAGCCTGCCGTCTCGTTAGCGATG ACAATGGAACATAAAGACCGGCCCCTTGGTT AGGGTCATTCTGACTAACACTGGGAGTCAT    2252
CCGGTCAAACAGCGTTCGGTGTATATCCACC GCGCTGTTGGACTCTGGAGCGGACATCACT ATTATTTCAGAGGAGGACTGGCCCCACCGAT TGGCCAGTGATGGAGGCCGCGAACCCGCAG    2372
ATCCATGGGATAGGAGAGGGAGGTTCAACT GCAGGTTTGTCTGCAACACCTCCTGCATCT TTGCCTGGGTCACTTACCAATGTGAAAGCA TTACAGAAATCACCAGGCCCCCAACGGGAA    2492
└►v-R*mil*
AGGAAATCATCCTCATCCTCAGAAGACAGA AATAGGATGAAAACCCTTGGTCGACGAGAT TCAAGTGATGATTGGGAAATACCAGATGGG CAGATCACAGTTGGACAAAGGATAGGATCT    2612
GGATCATTTGGAACAGTCTACAAAGGAAAG TGGCATGGTGACGTGGCAGTGAAAATGTTG AATGTTACAGCACCCACACCTCAACAGTTA CAGGCTTTCAAAAATGAAGTAGGAGTGCTC    2732
AGGAAAACACGGCATGTGAATATCCTACTT TTTATGGGTTATTCAACAAAACCTCAGTTG GCTATTGTTACACAGTGGTGTGAGGGGTCC AGCTTATATCACCATCTGCACATAATTGAG    2852
ACCAAGTTTGAAATGATCAAACTAATTGAT ATTGCACGACAGACTGCACAAGGCATGGAT TATTTGCATGCCAAGTCAATCATCCACAGA GACCTCAAGAGTAATAATATTTTTTCTTCAT    2972
GAAGACCTCACAGTAAAAATAGGTGACTTC GGTCTGGCTACAGTGAAATCACGATGGAGT GGATCTCATCAATTTGAACAGTTATCTGGA TCAATTCTATGGATGGCACCGGAAGTGATC    3092
AGGATGCAAGACAAAAACCCATATAGCTTT CAGTCAGATGTGTATGCATTCGGGATTGTG CTTTATGAACTGATGACTGGACAGTTACCA TACTCAAACATCAACAACAGGGACCAGATA    3212
ATTTTTATGGTGGGACGAGGATATCTATCT CCAGACCTCAGTAAAGTAAGAAGTAACTGT CCCAAAGCTATGAAGAGACTAATGGCAGAA TGCTTGAAAAAGAAAAGAGATGAGAGACCT    3332
CTTTTTCCACAGATTCTTGCCTCCATTGAG CTTCTGGCCCGGTCGTTGCCAAAAATTCAC CGCAGTGCATCTGAGCCGTCACTAAACCGG GCTGGCTTCCAGACCGAGGATTTCAGTCTG    3452
TATGCTTGTGCTTCTCCAAAAACGCCCATC CAAGCAGGAGGAATAGGAGAATGGGCCGTA CACTTGCTGAAAGGACTGCTTTTGGGGCTT GTAGTTATCTTGTTGCTAGTGGTATGCTTG    3572
└►env
CCTTGCCTTTTGCAATGTGTATCTAGTAGT ATTCGAAAGATGATTGATAATTCACTCGGC TATCGCGAGGAATATAAAAAATTACAGGAG GCTTATAAGCAGCCCGAAAGAAGAGCCT*AG*    3692
                                         A
GCGAGTTCTTATATTCCGTGTGATAGCTGG TTGGATTGGTAATTGATCGGCTGGCACGCG GAATATAGGAGGTCGCTGAATAGTAAACTT GTAGACTTGGCTACAGCATAGAGTATCTTC    3812
                   G
TGTAGCTCTGATGACTGCTAGGAAATAATG CTACGGATAATGTGGGGAGGGCAAGGCTTG CGAATCGGGTTGTAACGGGCAAGGCTTGAC TGAGGGGACAATAGCATGTTTAGGCGAAAA    3932
GCGGGGCTTCGGTTGTACGCGGTTAGGAGT CCCCTCAGGATATAGTAGTTTCGCTTTTGC ATAGGGAGGGGGAAATGTAGTCTTATGCAA TACTCTTGTAGTCTTGCAACATGCTTATGT    4052
                                                             └►U3
AACGATGAGTTAGCAACACGCCTTATAAGG AGAGAAAAAGCACCGTGCATGCCGATTGGT GGAAGTAAGGTGGTATGATCGTGGTATGAT CGTGGTATGATCGTGCCTTATTAGGAAGGC    4172
              T                                              G                   ******* *** *  G
AACAGACGGGTCTAACACGGATTGGACGAA CCACTGAATTC                                                             4213
```

References (1) Marx, M., Eychène, A., Laugier, D., Béchade, C., Crisanti, P., Dezélée, P., Pessac, B. and Calothy, G. (1988) *EMBO J.* ,*7*, 3369-3373. (2) Bizub, D., Katz, R.A. and Skalka, A.M. (1984) *J. Virol.*, *49*, 557-565.