Transgenic Res (2007) 16:253–259
DOI 10.1007/s11248-006-9059-6

ERRATUM

Characterization and multi-generational stability of the growth hormone transgene (EO-1α) responsible for enhanced growth rates in Atlantic Salmon

Edward S. Yaskowiak · Margaret A. Shears · Alka Agarwal-Mawal · Garth L. Fletcher

Published online: 10 January 2007

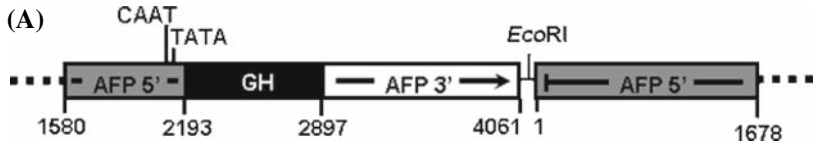
Erratum to: Transgenic Res
DOI: 10.1007/s11248-006-0020-5

Due to an unfortunate misunderstanding a distorted and incomplete version of Fig. 5 has been used in the above mentioned publication. On the following pages, the correct reproduction has been published that should be treated as definitive by the reader.

The online version of the original article can be found at http://dx.doi.org/10.1007/s11248-006-0020-5

E. S. Yaskowiak · M. A. Shears · A. Agarwal-Mawal · G. L. Fletcher
Aqua Bounty Canada, Inc., P.O. Box 13422, St. John's, Newfoundland, Canada, A1B 4B7

M. A. Shears · G. L. Fletcher (✉)
Ocean Sciences Centre, Memorial University of Newfoundland, St. John's, NL, Canada, A1C 5S7
e-mail: gfletcher@aquabounty.com



(B) ⇨ Genomic flanking sequence

```
F2     1 GTTGTACTGC TGATGCCTCT GATACCACAC AGACTGGGCC TCACGTTGTA
F4     1 GTTGTACTGC TGATGCCTCT GATACCACAC AGACTGGGCC TCACGTTGTA

F2    51 CTGCTGATAC CACACAGACT GGGCCTCACG TTGTACTACT GATACCACAC
F4    51 CTGCTGATAC CACACAGACT GGGCCTCACG TTGTACTACT GATACCACAC
                    ⇨ 5' ocean pout promoter region
C2  1580               TTGGCAGA TGTATGAGAA CTAAACCACT GACTGAACTT
F2   101 AGTAGTACAA CGTTGGCAGA TGTATGAGAA CTAAACCACT GACTGAACTT
F4   101 AGTAGTACAA CGTTGGCAGA TGTATGAGAA CTAAACCACT GACTGAACTT

C2  1618 GCACTAGAGG CATCTATTTT GTCTTTTCTC ATATGATGTT GGGATGGCAC
F2   151 GCACTAGAGG CATCTATTTT GTCTTTTCTC ATATGATGTT GGGATGGCAC
F4   151 GCACTAGAGG CATCTATTTT GTCTTTTCTC ATATGATGTT GGGATGGCAC

C2  1668 ATGGGAGTTT TTCCCCTGTC TCAGCTTGCT TTTTACCCCA AATATTGTAT
F2   201 ATGGGAGTTT TTCCCCTGTC TCAGCTTGCT TTTTACCCCA AATATTGTAT
F4   201 ATGGGAGTTT TTCCCCTGTC TCAGCTTGCT TTTTACCCCA AATATTGTAT

C2  1718 ATCTATTAGA ACCGTTGTCA CAGGGTTCAA ATTAACGTTT TAGTTTAGTT
F2   251 ATCTATTAGA ACCGTTGTCA CAGGGTTCAA ATTAACGTTT TAGTTTAGTT
F4   251 ATCTATTAGA ACCGTTGTCA CAGGGTTCAA ATTAACGTTT TAGTTTAGTT

C2  1768 TTGATCATGA TATACACATT TTATCCGTAA AGCATGTGCA TATACAGTAA
F2   301 TTGATCATGA TATACACATT TTATCCGTAA AGCATGTGCA TATACAGTAA
F4   301 TTGATCATGA TATACACATT TTATCCGTAA AGCATGTGCA TATACAGTAA

C2  1818 GGGCTTGTTA TTCGACAGCA AGAAGAAGAG GATATGTGTG CAGGCAGTCA
F2   351 GGGCTTGTTA TTCGACAGCA AGAAGAAGAG GATATGTGTG CAGGCAGTCA
F4   351 GGGCTTGTTA TTCGACAGCA AGAAGAAGAG GATATGTGTG CAGGCAGTCA

C2  1868 GCTAATGCAT GGATCACAAG TTATAGAATG CAAGCTTGTG ATAGTTTGGA
F2   401 GCTAATGCAT GGATCACAAG TTATAGAATG CAAGCTTGTG ATAGTTTGGA
F4   401 GCTAATGCAT GGATCACAAG TTATAGAATG CAAGCTTGTG ATAGTTTGGA

C2  1918 CAAAAACAAG TTATACTTTA CTTATAAGAA TATAAAATTT CCATTGCAAT
F2   451 CAAAAACAAG TTATACTTTA CTTATAAGAA TATAAAATTT CCATTGCAAT
F4   451 CAAAAACAAG TTATACTTTA CTTATAAGAA TATAAAATTT CCATTGCAAT

C2  1968 TGGCATAAGG AGGTGTGACA CAGTGACCTA CTTTCAGGCC AATAGGAAAC
F2   501 TGGCATAAGG AGGTGTGACA CAGTGACCTA CTTTCAGGCC AATAGGAAAC
F4   501 TGGCATAAGG AGGTGTGACA CAGTGACCTA CTTTCAGGCC AATAGGAAAC

C2  2018 GGGATATGCC GGTTAAGTCC TCCCACATAC TGTATATTAG ATGCAGCACA
F2   551 GGGATATGCC GGTTAAGTCC TCCCACATAC TGTATATTAG ATGCAGCACA
F4   551 GGGATATGCC GGTTAAGTCC TCCCACATAC TGTATATTAG ATGCAGCACA

C2  2068 TGGACCTGTC CTGTCAGAAG TCTCAGCTAC AGCTTTCACT TCGATCCAGA
F2   601 TGGACCTGTC CTGTCAGAAG TCTCAGCTAC AGCTTTCACT TCGATCCAGA
F4   601 TGGACCTGTC CTGTCAGAAG TCTCAGCTAC AGCTTTCACT TCGATCCAGA
```

Fig. 5 Schematic representation of the EO-1α integrate (**A**) and a comparison of opAFP-GHc2 (C2), EO-1α F2 generation (F2), and F4 generation (F4) consensus sequences (**B**). Abbreviations for specific sequences: AFP5′, Ocean pout antifreeze protein 5′ promoter region; GH, Chinook salmon growth hormone cDNA; AFP3′, Ocean pout antifreeze protein 3′ termination region; ——, pUC multiple cloning site; – – –, genomic flanking sequence. The CAAT and TATA boxes and bp numbers of the sequence in the injected construct are also indicated. The Atlantic salmon genomic flanking, chinook salmon GH cDNA, pUC 9 (P9) and pUC18 (P18) sequences are highlighted in grey. There were no ambiguities between the consensus sequences

```
C2 2118 TCTTTTCACT TCGATCTCCG ATAATTAATT AATTAATTAA TTATTAATTA
F2  651 TCTTTTCACT TCGATCTCCG ATAATTAATT AATTAATTAA TTATTAATTA
F4  651 TCTTTTCACT TCGATCTCCG ATAATTAATT AATTAATTAA TTATTAATTA
                                        ⇨ GH cDNA
C2 2168 ATTAAGTCTC AGCCACTGCA GGTCGTAAAA ATGGGACAAG TGTTTCTGCT
F2  701 ATTAAGTCTC AGCCACTGCA GGTCGTAAAA ATGGGACAAG TGTTTCTGCT
F4  701 ATTAAGTCTC AGCCACTGCA GGTCGTAAAA ATGGGACAAG TGTTTCTGCT

C2 2218 GATGCCAGTC TTACTGGTCA GTTGTTTCCT GAGTCAAGGG GCAGCGATAG
F2  751 GATGCCAGTC TTACTGGTCA GTTGTTTCCT GAGTCAAGGG GCAGCGATAG
F4  751 GATGCCAGTC TTACTGGTCA GTTGTTTCCT GAGTCAAGGG GCAGCGATAG

C2 2268 AAAACCAACG GCTCTTCAAC ATCGCGGTCA GCCGGGTGCA ACATCTCCAC
F2  801 AAAACCAACG GCTCTTCAAC ATCGCGGTCA GCCGGGTGCA ACATCTCCAC
F4  801 AAAACCAACG GCTCTTCAAC ATCGCGGTCA GCCGGGTGCA ACATCTCCAC

C2 2318 CTATTGGCTC AGAAAATGTT CAATGACTTT GACGGTACCC TGTTGCCTGA
F2  851 CTATTGGCTC AGAAAATGTT CAATGACTTT GACGGTACCC TGTTGCCTGA
F4  851 CTATTGGCTC AGAAAATGTT CAATGACTTT GACGGTACCC TGTTGCCTGA

C2 2368 TGAACGCAGA CAGCTGAACA AGATATTCCT GCTGGACTTC TGTAACTCTG
F2  901 TGAACGCAGA CAGCTGAACA AGATATTCCT GCTGGACTTC TGTAACTCTG
F4  901 TGAACGCAGA CAGCTGAACA AGATATTCCT GCTGGACTTC TGTAACTCTG

C2 2418 ACTCCATCGT GAGCCCAGTC GACAAGCACG AGACTCAGAA GAGTTCAGTC
F2  951 ACTCCATCGT GAGCCCAGTC GACAAGCACG AGACTCAGAA GAGTTCAGTC
F4  951 ACTCCATCGT GAGCCCAGTC GACAAGCACG AGACTCAGAA GAGTTCAGTC

C2 2468 CTGAAGCTGC TCCATATTTC TTTCCGTCTG ATTGAATCCT GGGAGTACCC
F2 1001 CTGAAGCTGC TCCATATTTC TTTCCGTCTG ATTGAATCCT GGGAGTACCC
F4 1001 CTGAAGCTGC TCCATATTTC TTTCCGTCTG ATTGAATCCT GGGAGTACCC

C2 2518 TAGCCAGACC CTGATCATCT CCAACAGCCT AATGGTCAGA AACGCCAACC
F2 1051 TAGCCAGACC CTGATCATCT CCAACAGCCT AATGGTCAGA AACGCCAACC
F4 1051 TAGCCAGACC CTGATCATCT CCAACAGCCT AATGGTCAGA AACGCCAACC

C2 2568 AGATCTCTGA GAAGCTCAGC GACCTCAAAG TGGGCATCAA CCTGCTCATC
F2 1101 AGATCTCTGA GAAGCTCAGC GACCTCAAAG TGGGCATCAA CCTGCTCATC
F4 1101 AGATCTCTGA GAAGCTCAGC GACCTCAAAG TGGGCATCAA CCTGCTCATC

C2 2618 ACGGGGAGCC AGGATGGCCT ACTGAGCCTG GATGACAATG ACTCTCAGCA
F2 1151 ACGGGGAGCC AGGATGGCCT ACTGAGCCTG GATGACAATG ACTCTCAGCA
F4 1151 ACGGGGAGCC AGGATGGCCT ACTGAGCCTG GATGACAATG ACTCTCAGCA

C2 2668 ACTGCCCCCC TACGGGAACT ACTACCAGAA CCTGGGGGGT GACGGAAACG
F2 1201 ACTGCCCCCC TACGGGAACT ACTACCAGAA CCTGGGGGGT GACGGAAACG
F4 1201 ACTGCCCCCC TACGGGAACT ACTACCAGAA CCTGGGGGGT GACGGAAACG

C2 2718 TCAGGAGGAA CTACGAGTTG TTGGCTTGCT TCAAGAAGGA CATGCACAAG
F2 1251 TCAGGAGGAA CTACGAGTTG TTGGCTTGCT TCAAGAAGGA CATGCACAAG
F4 1251 TCAGGAGGAA CTACGAGTTG TTGGCTTGCT TCAAGAAGGA CATGCACAAG

C2 2768 GTCGAGACCT ACCTGACCGT CGCCAAGTGC AGGAAGTCAC TGGAGGCCAA
F2 1301 GTCGAGACCT ACCTGACCGT CGCCAAGTGC AGGAAGTCAC TGGAGGCCAA
F4 1301 GTCGAGACCT ACCTGACCGT CGCCAAGTGC AGGAAGTCAC TGGAGGCCAA

C2 2818 CTGCACTCTG TAGACGTGGG CTGGAGAGGC AGCCAGCAAG AGCCCATCTC
F2 1351 CTGCACTCTG TAGACGTGGG CTGGAGAGGC AGCCAGCAAG AGCCCATCTC
F4 1351 CTGCACTCTG TAGACGTGGG CTGGAGAGGC AGCCAGCAAG AGCCCATCTC
                                        ⇨ 3' ocean pout termination region
C2 2868 CAGGGTTCGG TTTCCCAGAT ACAGATTAGG AACTGAACAT GTCAAAACCT
F2 1401 CAGGGTTCGG TTTCCCAGAT ACAGATTAGG AACTGAACAT GTCAAAACCT
F4 1401 CAGGGTTCGG TTTCCCAGAT ACAGATTAGG AACTGAACAT GTCAAAACCT
```

Fig. 5 continued

```
C2 2918 GTGGAGACTG TTGAGATTTG ATGTTCTGAA AAGATAAAGC CTATAAATAA
F2 1451 GTGGAGACTG TTGAGATTTG ATGTTCTGAA AAGATAAAGC CTATAAATAA
F4 1451 GTGGAGACTG TTGAGATTTG ATGTTCTGAA AAGATAAAGC CTATAAATAA

C2 2968 AATGTTGCCC AAATTTCCTG CCTGATGTTT TTCTTTGTCT TTGCTACATG
F2 1501 AATGTTGCCC AAATTTCCTG CCTGATGTTT TTCTTTGTCT TTGCTACATG
F4 1501 AATGTTGCCC AAATTTCCTG CCTGATGTTT TTCTTTGTCT TTGCTACATG

C2 3018 GCTTTGCTGC TCGGATCGGC TCACTCTGTG TATGCCACGT TCACTTTGTA
F2 1551 GCTTTGCTGC TCGGATCGGC TCACTCTGTG TATGCCACGT TCACTTTGTA
F4 1551 GCTTTGCTGC TCGGATCGGC TCACTCTGTG TATGCCACGT TCACTTTGTA

C2 3068 CTCTCCTTCT CACGGTAGGT TTATTATTTT TAGATGTGCA GTTAGTTTCT
F2 1601 CTCTCCTTCT CACGGTAGGT TTATTATTTT TAGATGTGCA GTTAGTTTCT
F4 1601 CTCTCCTTCT CACGGTAGGT TTATTATTTT TAGATGTGCA GTTAGTTTCT

C2 3118 GTGAAATAAC ACACCACACA CTGATATTGT CTGTGCATTG ACTTGGTGAG
F2 1651 GTGAAATAAC ACACCACACA CTGATATTGT CTGTGCATTG ACTTGGTGAG
F4 1651 GTGAAATAAC ACACCACACA CTGATATTGT CTGTGCATTG ACTTGGTGAG

C2 3168 TGCACATTGT TTTTGATCTT GACATATTTA TATTTGATTG ATCAGGTGAA
F2 1701 TGCACATTGT TTTTGATCTT GACATATTTA TATTTGATTG ATCAGGTGAA
F4 1701 TGCACATTGT TTTTGATCTT GACATATTTA TATTTGATTG ATCAGGTGAA

C2 3218 CTGTGTGAAT CTAAAGTGCT CCATACAGAT GTTCTGCATT GAAAATATTC
F2 1751 CTGTGTGAAT CTAAAGTGCT CCATACAGAT GTTCTGCATT GAAAATATTC
F4 1751 CTGTGTGAAT CTAAAGTGCT CCATACAGAT GTTCTGCATT GAAAATATTC

C2 3268 TCATTTTATT AGTGGAAGTG AGTGTATGCC ACATCCAATC AATTTCAGCA
F2 1801 TCATTTTATT AGTGGAAGTG AGTGTATGCC ACATCCAATC AATTTCAGCA
F4 1801 TCATTTTATT AGTGGAAGTG AGTGTATGCC ACATCCAATC AATTTCAGCA

C2 3318 AACACCCCAG TATGATTTAA TGCAAAAAAA TGAAGGTATC AAACACGCAT
F2 1851 AACACCCCAG TATGATTTAA TGCAAAAAAA TGAAGGTATC AAACACGCAT
F4 1851 AACACCCCAG TATGATTTAA TGCAAAAAAA TGAAGGTATC AAACACGCAT

C2 3368 TACTACTTTG CAGTTAAATA TTTAACATTT ATTCCAACAC GAAAAAAAGC
F2 1901 TACTACTTTG CAGTTAAATA TTTAACATTT ATTCCAACAC GAAAAAAAGC
F4 1901 TACTACTTTG CAGTTAAATA TTTAACATTT ATTCCAACAC GAAAAAAAGC

C2 3418 AGTAAATAAC ACTTTGACAA ACACGTCAGG ACATCTTATT TTTGTCACCC
F2 1951 AGTAAATAAC ACTTTGACAA ACACGTCAGG ACATCTTATT TTTGTCACCC
F4 1951 AGTAAATAAC ACTTTGACAA ACACGTCAGG ACATCTTATT TTTGTCACCC

C2 3468 TCACAGGCAA TTTAGTATAA TATATTATAT ATATATATAT ATCATATAAT
F2 2001 TCACAGGCAA TTTAGTATAA TATATTATAT ATATATATAT ATCATATAAT
F4 2001 TCACAGGCAA TTTAGTATAA TATATTATAT ATATATATAT ATCATATAAT

C2 3518 AATATTCAGT ATAATATATA TATATATATC ATATTATAAT ATTCAGTATA
F2 2051 AATATTCAGT ATAATATATA TATATATATC ATATTATAAT ATTCAGTATA
F4 2051 AATATTCAGT ATAATATATA TATATATATC ATATTATAAT ATTCAGTATA

C2 3568 ATATAAAACA CAAACACATA TATGTATAAT ATAATATAAC ATTTTTATTT
F2 2101 ATATAAAACA CAAACACATA TATGTATAAT ATAATATAAC ATTTTTATTT
F4 2101 ATATAAAACA CAAACACATA TATGTATAAT ATAATATAAC ATTTTTATTT

C2 3618 ATTGAGATGC CTCTATGGAC CGTGTTATAA GAAGTAAAGA TCAGGAGAAG
F2 2151 ATTGAGATGC CTCTATGGAC CGTGTTATAA GAAGTAAAGA TCAGGAGAAG
F4 2151 ATTGAGATGC CTCTATGGAC CGTGTTATAA GAAGTAAAGA TCAGGAGAAG
```

Fig. 5 continued

```
C2 3668 TAAACATGAA GTGTAATTAT GAATACTGAT GTTAAATTAA GCTATGATGA
F2 2201 TAAACATGAA GTGTAATTAT GAATACTGAT GTTAAATTAA GCTATGATGA
F4 2201 TAAACATGAA GTGTAATTAT GAATACTGAT GTTAAATTAA GCTATGATGA

C2 3718 GTTTTCACTG TTAATTTACC ATCTCAATTA AATGTTGATG CCTCCATGAC
F2 2251 GTTTTCACTG TTAATTTACC ATCTCAATTA AATGTTGATG CCTCCATGAC
F4 2251 GTTTTCACTG TTAATTTACC ATCTCAATTA AATGTTGATG CCTCCATGAC

C2 3768 CAAGTTAAGC AGATGAGACT GAGGACAACT GTAGAAGACA AGATGTTCAC
F2 2301 CAAGTTAAGC AGATGAGACT GAGGACAACT GTAGAAGACA AGATGTTCAC
F4 2301 CAAGTTAAGC AGATGAGACT GAGGACAACT GTAGAAGACA AGATGTTCAC

C2 3818 TTTGCTGAAT ATAGCTGGCT TGACAGTTAT CTATGACTCT ATAAATATAT
F2 2351 TTTGCTGAAT ATAGCTGGCT TGACAGTTAT CTATGACTCT ATAAATATAT
F4 2351 TTTGCTGAAT ATAGCTGGCT TGACAGTTAT CTATGACTCT ATAAATATAT

C2 3868 ATATATTTTT TTTTTTATAA AATGATTTAT TTATAACTAT ATATCCATTT
F2 2401 ATATATTTTT TTTTTTATAA AATGATTTAT TTATAACTAT ATATCCATTT
F4 2401 ATATATTTTT TTTTTTATAA AATGATTTAT TTATAACTAT ATATCCATTT

C2 3918 CTCAGACAGG TGCTTCATAT CCCTCACTCC CGTAGCTGTC CATGCTGGAT
F2 2451 CTCAGACAGG TGCTTCATAT CCCTCACTCC CGTAGCTGTC CATGCTGGAT
F4 2451 CTCAGACAGG TGCTTCATAT CCCTCACTCC CGTAGCTGTC CATGCTGGAT

C2 3968 CTGTCCCCGT TGTTTTTAAA AAGCTAAATA AGTTATTAAC ATGACTGCAT
F2 2501 CTGTCCCCGT TGTTTTTAAA AAGCTAAATA AGTTATTAAC ATGACTGCAT
F4 2501 CTGTCCCCGT TGTTTTTAAA AAGCTAAATA AGTTATTAAC ATGACTGCAT
                                                  ⇨ pUC 9
P9  239                                           GGCTGC
C2 4018 CCAGCGAGCC AAACCTGTCT GGTGTACAGC TACCAGAGAA GCTT
F2 2551 CCAGCGAGCC AAACCTGTCT GGTGTACAGC TACCAGAGAA GCTTGGCTGC
F4 2551 CCAGCGAGCC AAACCTGTCT GGTGTACAGC TACCAGAGAA GCTTGGCTGC

P9  245 AGGTCGACGG ATCCCCGGG⇨ pUC 18
P18 231                    A ATTCGAGCTC GGTACCCGG⇨ 5' ocean pout promoter
C2    1                                          G GATCCCCCAG
F2 2601 AGGTCGACGG ATCCCCGGGA ATTCGAGCTC GGTACCCGGG GATCCCCCAG
F4 2601 AGGTCGACGG ATCCCCGGGA ATTCGAGCTC GGTACCCGGG GATCCCCCAG

C2   12 AATGAGCTGG AACATGTTGC GGGGAGAGGG AAGTCTGGGT CAGCCTGCTT
F2 2651 AATGAGCTGG AACATGTTGC GGGGAGAGGG AAGTCTGGGT CAGCCTGCTT
F4 2651 AATGAGCTGG AACATGTTGC GGGGAGAGGG AAGTCTGGGT CAGCCTGCTT

C2   62 GGCCTGCTGC CACCGTGACC CGACCTCAGA TAAGCGGAGG AAAATGGATG
F2 2701 GGCCTGCTGC CACCGTGACC CGACCTCAGA TAAGCGGAGG AAAATGGATG
F4 2701 GGCCTGCTGC CACCGTGACC CGACCTCAGA TAAGCGGAGG AAAATGGATG

C2  112 GATGGATTGA ATCACAGAAT GTTTCTGAAG ACAGATATCA CCTTCGCTTC
F2 2751 GATGGATTGA ATCACAGAAT GTTTCTGAAG ACAGATATCA CCTTCGCTTC
F4 2751 GATGGATTGA ATCACAGAAT GTTTCTGAAG ACAGATATCA CCTTCGCTTC

C2  162 AAAGAGGTGC GCACCTGGGC AGGCACCCAC ACAGCCACAC AAATGGCATA
F2 2801 AAAGAGGTGC GCACCTGGGC AGGCACCCAC ACAGCCACAC AAATGGCATA
F4 2801 AAAGAGGTGC GCACCTGGGC AGGCACCCAC ACAGCCACAC AAATGGCATA

C2  212 TGAATCAACC AAGAAGACGG TTGGAACTGG TCAAAACCTT CACTATACCA
F2 2851 TGAATCAACC AAGAAGACGG TTGGAACTGG TCAAAACCTT CACTATACCA
F4 2851 TGAATCAACC AAGAAGACGG TTGGAACTGG TCAAAACCTT CACTATACCA
```

Fig. 5 continued

```
C2   262 TGTGTGACAG TTGTTTGTCA CAGTGTATAA AAGACAGGGA CTTAGAGACA
F2  2901 TGTGTGACAG TTGTTTGTCA CAGTGTATAA AAGACAGGGA CTTAGAGACA
F4  2901 TGTGTGACAG TTGTTTGTCA CAGTGTATAA AAGACAGGGA CTTAGAGACA

C2   312 GAGCTCTGAG CAGCTATGAG ATTGTAGTTT GGCCAGGATG CGCTTAAGAC
F2  2951 GAGCTCTGAG CAGCTATGAG ATTGTAGTTT GGCCAGGATG CGCTTAAGAC
F4  2951 GAGCTCTGAG CAGCTATGAG ATTGTAGTTT GGCCAGGATG CGCTTAAGAC

C2   362 CTTTGTGATG AAAAGTTATC AAATTCGTGA GTTTTCATGG AAGAACCTTG
F2  3001 CTTTGTGATG AAAAGTTATC AAATTCGTGA GTTTTCATGG AAGAACCTTG
F4  3001 CTTTGTGATG AAAAGTTATC AAATTCGTGA GTTTTCATGG AAGAACCTTG

C2   412 ACGTGGCGTG GTGGCCATTT TGCGTCATTC GGCATGGAAA AGGAAGTCGT
F2  3051 ACGTGGCGTG GTGGCCATTT TGCGTCATTC GGCATGGAAA AGGAAGTCGT
F4  3051 ACGTGGCGTG GTGGCCATTT TGCGTCATTC GGCATGGAAA AGGAAGTCGT

C2   462 TATAACTCCC AGGTACATTA TCTTATCTAC ACAAAATGTC TAATGCATGA
F2  3101 TATAACTCCC AGGTACATTA TCTTATCTAC ACAAAATGTC TAATGCATGA
F4  3101 TATAACTCCC AGGTACATTA TCTTATCTAC ACAAAATGTC TAATGCATGA

C2   512 TACTACTTAA AGCCTGAGCA TATTTCAAGG CCAGCACTTT TCAATAACTC
F2  3151 TACTACTTAA AGCCTGAGCA TATTTCAAGG CCAGCACTTT TCAATAACTC
F4  3151 TACTACTTAA AGCCTGAGCA TATTTCAAGG CCAGCACTTT TCAATAACTC

C2   562 ATAGGCCACC TGCTGGCAAA AGGAAATGCC ACATTTTATA CTTTTATTTA
F2  3201 ATAGGCCACC TGCTGGCAAA AGGAAATGCC ACATTTTATA CTTTTATTTA
F4  3201 ATAGGCCACC TGCTGGCAAA AGGAAATGCC ACATTTTATA CTTTTATTTA

C2   612 CTCCTAGACA GTTGACCTGA TCAGTCTCAA ATTTGGTAAG GATAGCCTTA
F2  3251 CTCCTAGACA GTTGACCTGA TCAGTCTCAA ATTTGGTAAG GATAGCCTTA
F4  3251 CTCCTAGACA GTTGACCTGA TCAGTCTCAA ATTTGGTAAG GATAGCCTTA

C2   662 AGACAATGAA GATGCTTCAT CAGGAATATT GTGAGTTGTC GTTGAACGTT
F2  3301 AGACAATGAA GATGCTTCAT CAGGAATATT GTGAGTTGTC GTTGAACGTT
F4  3301 AGACAATGAA GATGCTTCAT CAGGAATATT GTGAGTTGTC GTTGAACGTT

C2   712 GTTGCCGTGG CAACGCATCA TTCGCCATGA AAAAGAAGCT GATGGTTCAG
F2  3351 GTTGCCGTGG CAACGCATCA TTCGCCATGA AAAAGAAGCT GATGGTTCAG
F4  3351 GTTGCCGTGG CAACGCATCA TTCGCCATGA AAAAGAAGCT GATGGTTCAG

C2   762 TGGCTTGGGA TGCTCAAAAA GTCATGGAAC TTTGTACATG TGTCATAATT
F2  3401 TGGCTTGGGA TGCTCAAAAA GTCATGGAAC TTTGTACATG TGTCATAATT
F4  3401 TGGCTTGGGA TGCTCAAAAA GTCATGGAAC TTTGTACATG TGTCATAATT

C2   812 GATGGGAAGT TGTATGGGTT TTTGGCTTGC TTGTTATAAA TTGTCTCCAT
F2  3451 GATGGGAAGT TGTATGGGTT TTTGGCTTGC TTGTTATAAA TTGTCTCCAT
F4  3451 GATGGGAAGT TGTATGGGTT TTTGGCTTGC TTGTTATAAA TTGTCTCCAT

C2   862 AGCGCCCCCT ACAATATTTC AAAAGAGCAG CCCCAGTGCT ACGTACATGT
F2  3501 AGCGCCCCCT ACAATATTTC AAAAGAGCAG CCCCAGTGCT ACGTACATGT
F4  3501 AGCGCCCCCT ACAATATTTC AAAAGAGCAG CCCCAGTGCT ACGTACATGT

C2   912 ATGAAACTTA GTAGCCAGAT GTACCATATA GAGACTTACA AAAAGGTATC
F2  3551 ATGAAACTTA GTAGCCAGAT GTACCATATA GAGACTTACA AAAAGGTATC
F4  3551 ATGAAACTTA GTAGCCAGAT GTACCATATA GAGACTTACA AAAAGGTATC

C2   962 TTGGCCATGC TCTCAACCGT ACTGGAAGTC GGCCATTTTG ATTTTTGCAT
F2  3601 TTGGCCATGC TCTCAACCGT ACTGGAAGTC GGCCATTTTG ATTTTTGCAT
F4  3601 TTGGCCATGC TCTCAACCGT ACTGGAAGTC GGCCATTTTG ATTTTTGCAT

C2 1012 AATTTTTCAA TAGATTTTTG CACATTTGTA ATCGCTATAC TTTAACGAAC
F2  3651 AATTTTTCAA TAGATTTTTG CACATTTGTA ATCGCTATAC TTTAACGAAC
F4  3651 AATTTTTCAA TAGATTTTTG CACATTTGTA ATCGCTATAC TTTAACGAAC
```

Fig. 5 continued

```
C2 1062 TCCTCCAAGG AACTTTGTCT AATCAATTTC AAATTTTGTC AGTACAATCT
F2 3701 TCCTCCAAGG AACTTTGTCT AATCAATTTC AAATTTTGTC AGTACAATCT
F4 3701 TCCTCCAAGG AACTTTGTCT AATCAATTTC AAATTTTGTC AGTACAATCT

C2 1112 CAGTACTACA GTACCAAATC TACAGTTCTG CATCTCGTAG CTGCTCAGAG
F2 3751 CAGTACTACA GTACCAAATC TACAGTTCTG CATCTCGTAG CTGCTCAGAG
F4 3751 CAGTACTACA GTACCAAATC TACAGTTCTG CATCTCGTAG CTGCTCAGAG

C2 1162 GTCTGTCTCT AAGTCCCTGT CTTTTATACA CTGTGACAAA CAACTGTCAC
F2 3801 GTCTGTCTCT AAGTCCCTGT CTTTTATACA CTGTGACAAA CAACTGTCAC
F4 3801 GTCTGTCTCT AAGTCCCTGT CTTTTATACA CTGTGACAAA CAACTGTCAC

C2 1212 ACATGGTATA GTGAAGGTTT TGACCAGTTC CAACCGTCTT GTTGGTTGAT
F2 3851 ACATGGTATA GTGAAGGTTT TGACCAGTTC CAACCGTCTT GTTGGTTGAT
F4 3851 ACATGGTATA GTGAAGGTTT TGACCAGTTC CAACCGTCTT GTTGGTTGAT

C2 1262 TCATATGCCA TTCGTGTGGC TGTGTGGGTG CCTACCCAGA TGCGCACCTC
F2 3901 TCATATGCCA TTCGTGTGGC TGTGTGGGTG CCTACCCAGA TGCGCACCTC
F4 3901 TCATATGCCA TTCGTGTGGC TGTGTGGGTG CCTACCCAGA TGCGCACCTC

C2 1312 TTTGAAGCGA ATGTGATATC TGTCTTCATA AACATTCTGT TATTAGCAAG
F2 3951 TTTGAAGCGA ATGTGATATC TGTCTTCATA AACATTCTGT TATTAGCAAG
F4 3951 TTTGAAGCGA ATGTGATATC TGTCTTCATA AACATTCTGT TATTAGCAAG

C2 1362 TTCATATGAG ATTGAAGGCT GTATGCAAAC AGGTGCACAG TCTGTTTCTA
F2 4001 TTCATATGAG ATTGAAGGCT GTATGCAAAC AGGTGCACAG TCTGTTTCTA
F4 4001 TTCATATGAG ATTGAAGGCT GTATGCAAAC AGGTGCACAG TCTGTTTCTA

C2 1412 AGCATCATGG AAAAGTACAA GCAATTTGCA CAAATCATTC TGTATTTTTC
F2 4051 AGCATCATGG AAAAGTACAA GCAATTTGCA CAAATCATTC TGTATTTTTC
F4 4051 AGCATCATGG AAAAGTACAA GCAATTTGCA CAAATCATTC TGTATTTTTC

C2 1462 CAATAGCTAA CAATGTCACC GGGACATTGT GCTATTGGAT AGAAGAGACC
F2 4101 CAATAGCTAA CAATGTCACC GGGACATTGT GCTATTGGAT AGAAGAGACC
F4 4101 CAATAGCTAA CAATGTCACC GGGACATTGT GCTATTGGAT AGAAGAGACC

C2 1512 AGCTGATCTA GACAGTTGAT ATCATGATCA ACAGCCCCAA ACAACAAGTG
F2 4151 AGCTGATCTA GACAGTTGAT ATCATGATCA ACAGCCCCAA ACAACAAGTG
F4 4151 AGCTGATCTA GACAGTTGAT ATCATGATCA ACAGCCCCAA ACAACAAGTG

C2 1562 TGCATGCGCG AGGAGTGATT GGCAGATGTA TGAGAACTAA ACCACTGACT
F2 4201 TGCATGCGCG AGGAGTGATT GGCAGATGTA TGAGAACTAA ACCACTGACT
F4 4201 TGCATGCGCG AGGAGTGATT GGCAGATGTA TGAGAACTAA ACCACTGACT

C2 1612 GAACTTGCAC TAGAGGCATC TATTTTGTCT TTTCTCATAT GATGTTGGGA
F2 4251 GAACTTGCAC TAGAGGCATC TATTTTGTCT TTTCTCATAT GATGTTGGGA
F4 4251 GAACTTGCAC TAGAGGCATC TATTTTGTCT TTTCTCATAT GATGTTGGGA

C2 1662 TGGCACATGG GAGTTTT⇨ Genomic flanking sequence
F2 4301 TGGCACATGG GAGTTTTACT GGGCCTCACG TTGTACTGCT GATACCTCAC
F4 4301 TGGCACATGG GAGTTTTACT GGGCCTCACG TTGTACTGCT GATACCTCAC

F2 4351 AGACTGGGCC TCACGTTGTA CTACTGATAC CACACAGACT GGGCCTCACG
F4 4351 AGACTGGGCC TCACGTTGTA CTACTGATAC CACACAGACT GGGCCTCACG

F2 4401 TTGTACTACT GATACCACAC AGATTGGGCC TCACGTTGTA CTGCTGATAC
F4 4401 TTGTACTACT GATACCACAC AGATTGGGCC TCACGTTGTA CTGCTGATAC

F2 4451 CACACAGACT GGGTCTCACG TTGTCCTGCT GATACCTCAC AGACT
F4 4451 CACACAGACT GGGTCTCACG TTGTCCTGCT GATACCTCAC AGACT
```

Fig. 5 continued