

5'-AGAAGGAAGA GTATGTGGAA CATGTCTCTG ATGCCACGAT GTAATAACCT TGAAAACAAA GTAAAATTAC ACTAATGAGT CTTGCCTAAT TAAACTCATG CTCCTAGTGA TCACCACTTC TAGTTCAATT GTTCACATTC TTGCTCTGCT TTGAAAAATT AAAATTAAAT TTGCCTATCC TCTACTGACC ATAATTTCTA GAAGACGGCA TTCATCTCAT GGCAAGTTCT TCAGTACCCA AAGATGGAAT ACATAGATTA AAAAAGAACA TATATGTAGA TGCTTGTGAT GTTTTCCTAT CATAAATTGA ATTTCAAGTT CTTATAAA CG TATTAATATG TCCTACTCTT CTAGAGACAA GGATCAGGAA GTGTATTTAT CAATAGATAT TTACCAAGCA CCTGTCAAGC CAAAGTGGGG TTACAGAAAA GTAGGTATGG GCCCTGCACA CAAACAACCT GTATTAGCCA AAGGGACCCT TCCATAAAAT TTCCAATATG TAAACCCAAA TTTGGAACTT GCTGAAACAA GTACAGATGA GTACGTGAGG AAGCTGGGAA GTAAACACAG GTTGCTG-3'

Supplementary Fig. 2. Plots of G+C nucleotide percentage (upper panel) and observed versus expected ratio of CpG dinucleotide frequency (lower panel) for the FLG 5'-end. Nucleotide sequence (GenBank accession No. NC 000001.11) of the promoter region analyzed by bisulfite genomic sequencing and two key CpG sites (red box) analyzed by bisulfite pyrosequencing are indicated under the CpG island plot.