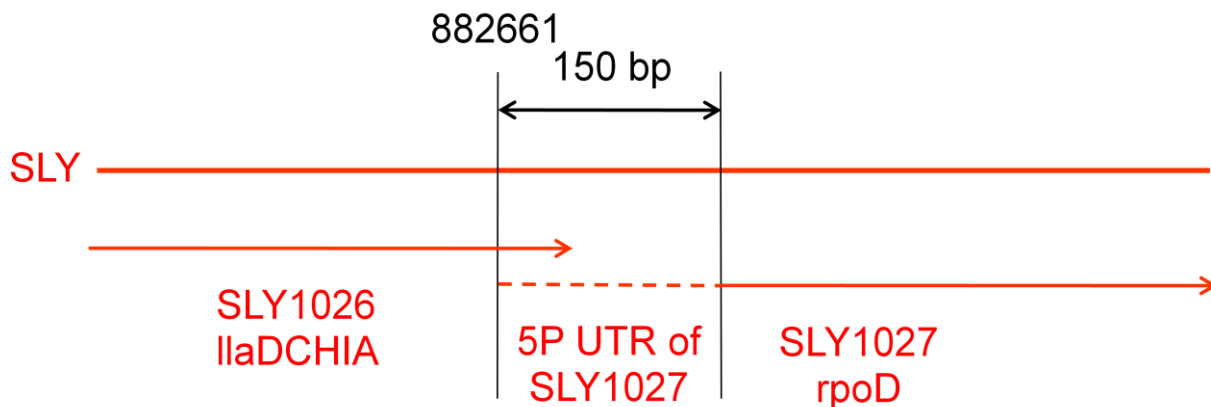


Figure S2. Example of an SLY Open Reading Frame (ORF) being altered by the intrusion of the 5' Untranslated Region (UTR) of the *rpoD* gene. (A) Diagram illustrating that the 150 bp 5' UTR of SLY1027 (*rpoD* Group B) overlaps with the 3' region of the ORF SLY1026. (B) A line-up of predicted amino acid sequences of SLY1026 and paralogues shows a high degree of sequence identity except for 11 residues of the carboxyl (3') terminus. (C) Nucleic acid alignment of SLY1026 and paralogues shows that the point at which sequences differ corresponds to the 5'UTR of SLY1027 (bottom row). Yellow colouring indicates exact nucleotide match and blue colouring is greatest consensus match.

A



B

```

SLY_1026 (190) WR-EGIKDKDFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGWIDEITRTLEV
SLY_1079 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_058 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_729 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_1088 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_171 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_582 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_693 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_959 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_985 (190) WR-EGIKDKYFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS SKITELFKKFHFIQPIQTTHKHLTSKGISQEVIIKNY-
SLY_218 (190) WR-EGTKDKDFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS PKIIELFKDFHFIQPLQTTTNVLTSGKITQEVIIKNY-
PA0250 (190) WR-EGTKDKDFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS PKIIELFKDFHFIQPLQTTTNVLTSGKITQEVIIKNY-
SLY_829 (129) WR-EGTKDKDFYQKPFTFEDHKRLYSLEMANYSKRWLYTNYS PKIIELFKDFHFIQPIQTTHKHLTSKGISQEVIIKNY-

```

C

```

(744) 744 750 760 770 780 790 800 819
SLY_058 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_729 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_1026 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_1088 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_171 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_582 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_693 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_985 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_959 (744) TCAAACCACTACAAACACAATTCTAACTTCAAAGGGATATCGCAAGAAGTTATTATTAAAAATTATTAA-----
SLY_218 (744) CCAAATACAAACCAATTACGCTTAACTTCAAAGGGAATAACACAAAGAAGTTATTATTAAAAATTATTAA-----
SLY_829 (561) TCAAATACAAACCAATTACGCTTAACTTCAAAGGGAATAACACAAAGAAGTTATTATTAAAAATTATTAA-----
SLY1027 (1) -----TGGACTGATGAAATAATTAGGACACTTTTGTGTTTTTAA

```