

5'...LVHEKDGSWCMCLDYRELNKLTIKDKFPIPVIDE~~LL~~DELHGSIYFTKLDLR **RVT**
 SGYHQIRMKMKDILKTTFRTHEGHYEFLVM**PFGL**TNAPSTFQDLMNSIFKPF
 LRKFVLVFF**DDILIY**IKSWKDHVEHVERVLQLLEEKQLYAKIFKCFGGVQEVEY
 LGHIVSHEGVKIDPSKIKSIKEWKISTSIKHLRGFLGLTGYYRKFKVKNYGRIAAP
 LTTLLKKDSFSWTPEATKAFKHLKEAMCQALVVATPDFTKTFIVE**CDAS**GNGI **RNaseH**
 GVVLMQDERPITFESRLIKGKFLRKDIYEKEMLAILHALKKWRPYLMGRHFKV
 KTDHDSLKYFLEQRLSSEEQQKWVTKMLGYDFEIIYKKGKQNVVADALSKDE
 DVEAFLCAISIIQPDWINEAREEWKNDDEEWALIRKLQQYSSTSETFSWKND
 SLWYKDHLYLCKNYQLKQKILMEF**HTSPLGGHSR**FLKTYHRVKKEFFWDGL
 KSDIQKFVAE**CLVCCQ**QNKVATIKTPGLLQPLSIPSQCWEDVSMDFITGLPKSE
 GKSVMVVV**DRL**TKYAHFCTLSHPFKASKVSTAFMETVQKLHGPNPKIIVS**DRD**
 PIFTGNFWSELFSC~~LG~~TQLAHSSSYHPQSDGKT**EIV**NKFLEGYLRCFVSDKQ **INT**
 TQWVKLPLAEWWYNTSFHTATKMTPFMALYGYQTPSTTSYLRENSKVQAV
 EHHIEHQQQVLQLLKDNLVLAQNRMKQQADQHRSE~~RS~~FDV**GD**WVFLRLQL
 YKKMSLKQAKKDN**KLSPKYYGSYKV**LQKIGTMAYKLELPAASRL**HPVFHVS**
 CLKKVIGDKLPIQTIFQNLTRKEKLYWNLESPKSLVTIFPELDEEGKIIFEPEAV
 TETSTRQVRNRSISEYLIKWKNLSTEDSTWEDENFMRKYPELLKREERK...3' **Chromo**

Additional file 1. Structural features of *Corky* retrotransposon.

Conserved amino acid (single letter code) domains of Reverse Transcriptase (underline), RNaseH, Integrase (underline) and Chromo (underline). In the Reverse Transcriptase two important motives PFGL and DDILIYS are in red. In RNaseH the CDAS motif is pointed in bold. In the integrase the three subdomains are in red: HH-CC; D,DE and G-(D/E)-X₁₀₋₂₀-KL-X₂/R(K)-F/Y(W)-X-G-P-(F/Y)-X-(I/V). The HPVFH(V)S motif is showed in bold.