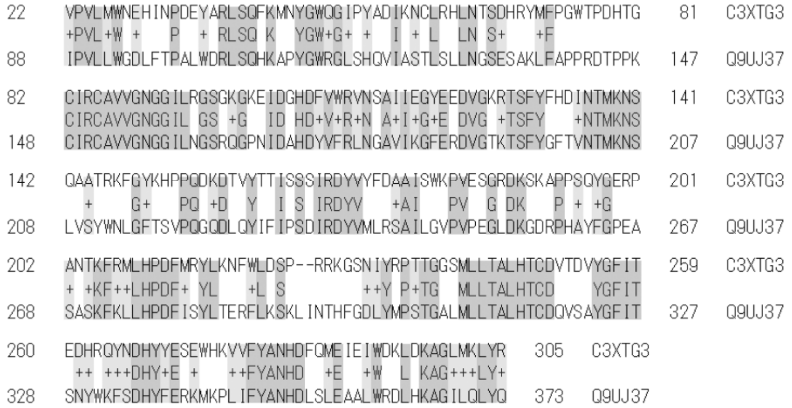
Supplementary Figure

(A) Query sequence:
>tr|C3XTG3|C3XTG3_BRAFL Putative uncharacterized protein (Fragment) OS=Branchiostoma floridae GN=BRAFLDRAFT_219600

Alignment against
>sp|Q9UJ37|SIA7B_HUMAN Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2 OS=Homo sapiens GN=ST6GALNAC2

Score: 666, E-value: 4.0×10^{-83}
Identity: 45.0%, Positives: 62.0%
Query length: 317, Match length: 374
Position: Q9UJ37 matches from 88 to 373 (286AA), in the query sequence from 22 to 305 (284AA)

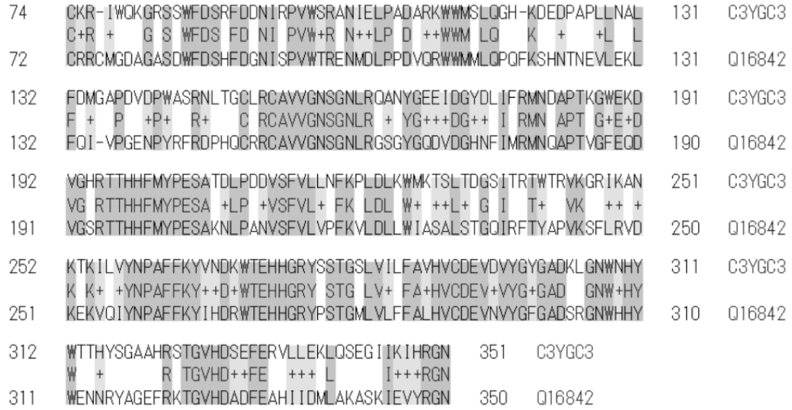


```
22    VPVLMWNEHINPDEYARLSQFKMNYGWQGIPYADIKNCLRHLNTSDHRYMFPGWTPDHTG   81   C3XTG3
      +PVL+W +   P  + RLSQ K  YGW+G+ +  I + L  LN S+   +F
88    IPVLLWGDLFTPALWDRLSQHKAPYGWRGLSHQVIASTLSLLNGSESAKLFAPPRDTPPK  147   Q9UJ37

82    CIRCAVVGNGGILRGSGKGKEIDGHDFVWRVNSAIIEGYEEDVGKRTSFYFHDINTMKNS  141   C3XTG3
      CIRCAVVGNGGIL GS +G  ID HD+V+R+N A+I+G+E DVG +TSFY   +NTMKNS
148   CIRCAVVGNGGILNGSRQGPNIDAHDYVFRLNGAVIKGFERDVGTKTSFYGFTVNTMKNS  207   Q9UJ37

142   QAATRKFGYKHPPQDKDTVYTTISSSIRDYVYFDAAISWKPVESGRDKSKAPPSQYGERP  201   C3XTG3
        +    G+   PQ +D  Y  I S IRDYV   +AI   PV  G DK   P + +G
208   LVSYWNLGFTSVPQGQDLQYIFIPSDIRDYVMLRSAILGVPVPEGLDKGDRPHAYFGPEA  267   Q9UJ37

202   ANTKFRMLHPDFMRYLKNFWLDSP--RRKGSNIYRPTTGGSMLLTALHTCDVTDVYGFIT  259   C3XTG3
      + +KF++LHPDF+ YL   +L S        ++Y P+TG  MLLTALHTCD    YGFIT
268   SASKFKLLHPDFISYLTERFLKSKLINTHFGDLYMPSTGALMLLTALHTCDQVSAYGFIT  327   Q9UJ37

260   EDHRQYNDHYYESEWHKVVFYANHDFQMEIEIWDKLDKAGLMKLYR   305   C3XTG3
       ++ +++DHY+E +   ++FYANHD  +E  +W  L KAG+++LY+
328   SNYWKFSDHYFERKMKPLIFYANHDLSLEAALWRDLHKAGILQLYQ   373   Q9UJ37
```

(B) Query sequence
>tr|C3YGC3|C3YGC3_BRAFL Putative uncharacterized protein OS=Branchiostoma floridae GN=BRAFLDRAFT_92867

Alignment against
>sp|Q16842|SIA4B_HUMAN CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase 2 OS=Homo sapiens GN=ST3GAL2

Score: 827, E-value: 5.0×10^{-107}
Identity: 53.0%, Positives: 71.0%
Query length: 355, Match length: 350
Position: Q16842 matches from 72 to 350 (279AA), in the query sequence from 74 to 351 (278AA)



```
74    CKR-IWQKGRSSWFDSRFDDNIRPVWSRANIELPADARKWWMSLQGH-KDEDPAPLLNAL  131   C3YGC3
      C+R +   G S WFDS FD NI PVW+R N++LP D ++WWM LQ   K  +   +L  L
72    CRRCMGDAGASDWFDSHFDGNISPVWTRENMDLPPDVQRWWMMLQPQFKSHNTNEVLEKL  131   Q16842

132   FDMGAPDVDPWASRNLTGCLRCAVVGNSGNLRQANYGEEIDGYDLIFRMNDAPTKGWEKD  191   C3YGC3
      F  +  P  +P+  R+   C RCAVVGNSGNLR + YG+++DG++ I RMN APT G+E+D
132   FQI-VPGENPYRFRDPHQCRRCAVVGNSGNLRGSGYGQDVDGHNFIMRMNQAPTVGFEQD  190   Q16842

192   VGHRTTHHFMYPESATDLPDDVSFVLLNFKPLDLKWMKTSLTDGSITRTWTRVKGRIKAN  251   C3YGC3
      VG RTTHHFMYPESA +LP +VSFVL+ FK LDL W+ ++L+ G I  T+  VK  ++ +
191   VGSRTTHHFMYPESAKNLPANVSFVLVPFKVLDLLWIASALSTGQIRFTYAPVKSFLRVD  250   Q16842

252   KTKILVYNPAFFKYVNDKWTEHHGRYSSTGSLVILFAVHVCDEVDVYGYGADKLGNWNHY  311   C3YGC3
      K K+ +YNPAFFKY++D+WTEHHGRY STG LV+ FA+HVCDEV+VYG+GAD  GNW+HY
251   KEKVQIYNPAFFKYIHDRWTEHHGRYPSTGMLVLFFALHVCDEVNVYGFGADSRGNWHHY  310   Q16842

312   WTTHYSGAAHRSTGVHDSEFERVLLEKLQSEGIIKIHRGN   351   C3YGC3
      W  +      R TGVHD++FE  +++ L     I+++RGN
311   WENNRYAGEFRKTGVHDADFEAHIIDMLAKASKIEVYRGN   350   Q16842
```

Supplementary Figure S1 Putative ST3Gal and ST6GalNAc family proteins in *Branchiostoma floridae*.
Unannotated sialyltransferases were searched for using BLAST search where the *B. floridae* proteome dataset were queried with human sialyltransferases. In four sialyltransferase families, *B. floridae* proteins that shared the highest homology with each human sialyltransferase were extracted as putative homologous sialyltransferases. All the members of the ST6GalNAc family (A) and ST3Gal family (B) lacked an intrinsic protein name and gene name. *B. floridae* C3XTG3 and C3YGC3 proteins had the lowest E-values in the ST6GalNAc family (A) and ST3Gal family (B), respectively.