

**S1 File. Multiple sequence alignment of the *F. hepatica* NEJ cathepsins.** Multiple sequence alignment of the FhCB (BN1106\_s5100B000033), FhCB1 (BN1106\_s6570B000050), FhCB2 (BN1106\_s4482B000044) and FhCB3 (BN1106\_s6570B000051) and ProCB2 (gi|27526823) were performed using Clustal Omega. Sequences are ordered according to aligned. Differences in the region containing the *N*-glycosylation site are highlighted in yellow. Consensus *N*-glycosylation is underlined.

```

BN1106_s5100B000033 -----
BN1106_s4482B000044 MNWLVFAIIAVVQAKPNHKPQFEAFSDELIRFVNEESGASWKAARSTRFSNVDFHKLHL
BN1106_s6570B000051 MSWLLIFAAIVVAQAKPNYKRFEPFSDLIHYINEESGASWKAAPSTRFNIDQVKQNL
gi|27526823 -----VVVQAAPNEKPFQFEPFSDLIHYINEKSGASWKAAPSSRFINIEHFKQHL
BN1106_s6570B000050 MSWLLIFVAIVVVQAAPNEKPFQFEPFSDLIHYINEKSGASWKAAPSSRFINIEHFKQHL

BN1106_s5100B000033 -----
BN1106_s4482B000044 GALSETPEERNALRPTIKHDISKNDLPESFDARSQWPQCWTISEIRDQASCSCWVIANE
BN1106_s6570B000051 GVLEETPEDRNTQRQTVRYVSENDLPESFDARQKWANCPSISEIRDQSSCSCWAVSSA
gi|27526823 GLLEETPEERQTRRPTVRYNVSDNDLPESFDAREKWPLCRSIRQIPDQSSCGSCWAVAGV
BN1106_s6570B000050 GLLEETPEERQTRRPTVKYNVSENDLPESFDAREKWP-CRSIRQIPDQSSCGSCWAVAGV

BN1106_s5100B000033 -----
BN1106_s4482B000044 HSTAEKEEIGQIYICVRHAQILWHRERESFVLLS--SYHRS-----GASWKAARSTR
BN1106_s6570B000051 SAITDR-----ICIHNSGQKPK---RLSAIDIVSCCAYCGYGCNGGIPAMSWDYWTREG
gi|27526823 GAMSDR-----VCIHNSGMQP---ELSAIDLVSCCSYCGNGCQGGSPPAAWDYWWRNG
BN1106_s6570B000050 GAMSDR-----VCIHNSGMQP---ELSAIDLVSCCSYCGNGCQGGSPPAAWDYWWRNG

BN1106_s5100B000033 -----
BN1106_s4482B000044 MNRLLA-----FAVIADVQA---TPSSKTRFDT-----YSD--QL
BN1106_s6570B000051 FSNVDHFKLHLGALSETPEE-RNALRPTIKHDISKNDLPESFDARSQWPQCWTISEIRDQ
gi|27526823 -----VVTGGTLENPTGCLPYFPFKCSHGVTPLPCCPRDIYPTPKCEKKCH--AG
BN1106_s6570B000050 -----IVTGGTLENPTGCLPYFPQCRHP-----
-----IVTGGTLENPTGCLPYFPQCRHP-----

BN1106_s5100B000033 -----
BN1106_s4482B000044 IHVNEEAFGAVEAMSDRVCIHNSGQMKPHLSARDLTSCCRTCGFGCKGGYPALAWDYWE
BN1106_s6570B000051 ASCGSCWATAAASAMSDRVCIHNSGQMRPRLAAADPLSCCTYCGQCRGGYPPKAWDYWM
gi|27526823 YNKTTVPAISSASAITDRICIHNSGQKPKRLSAIDIVSCCAYCGYGCNGGIPAMSWDYWT
BN1106_s6570B000050 -----

BN1106_s5100B000033 -----
BN1106_s4482B000044 SDGIVTGGSLLEEQTCAPYFPFKCSHYESSK-----
BN1106_s6570B000051 REGIVTGGTLENPTGCLPYFPFKCSHGVTPLPCCPRDIYPTPKCEKKCHAGYNKTYEQ
gi|27526823 -----GSRSQ-----
BN1106_s6570B000050 -----GSRSQ-----

BN1106_s5100B000033 -----
BN1106_s4482B000044 -----
BN1106_s6570B000051 DKVKGKSSYNVGEQETDIMMEIMKNGPVDGIFYMFEDFLVYXGCNGGIPAMSWDYWTREG
gi|27526823 -----
BN1106_s6570B000050 -----

BN1106_s5100B000033 -----
BN1106_s4482B000044 -----GLKPCPDELYPTPTCEMKCQAGYNKTYDEKDF
BN1106_s6570B000051 -----KYSRCPHYTYPTPPCARACQTGYNKTYEQDKF
gi|27526823 VVTGGTLENPTGCLPYFPFKCSHGVTPLPCCPRDIYPTPKCEKKCHAGYNKTYEQDKV
BN1106_s6570B000050 -----LNPCPRYTYPTPSCYPYCYQAGYDKTYEKDKV
-----LNPCP-----

BN1106_s5100B000033 -----
BN1106_s4482B000044 YGNSSYNVEKSEKNIMYEIMKNGPVEATLTVYEDFVYKSGIYHHVRSRLMGGHAVRILG
BN1106_s6570B000051 YGNSSYNVGEHESYIMQEIMKNGPVEVTFQDFGVYRSGIYHHVAGKFIGRHAVRMIG
gi|27526823 KGKSSYNVGEQETDIMMEIMKNGPVDGIFYMFEDFLVYKSGIYHYTTGRLVGGHAIRVIG
BN1106_s6570B000050 YGKTSYNVDRHEYTIMQEIMKNGPVEAGFIVYTDFAVYKSGIYHHVSGRYAGKHAIRIIG
-GKTSYNVDRNEYTIMQEIMKNGPVEAGFIVYTDFAVYKSGIYHHVSGRYAGKHAIRIIG

BN1106_s5100B000033 -----
BN1106_s4482B000044 WGVENGVKYWLIVANSWNEEWGENGYFRIRRGINECGIESEVTAGMPLL-----
BN1106_s6570B000051 WGVENGVNWYLMANSWNEEWGENGYFRMVRGRNECGIESEVVAGMPLR-----
gi|27526823 WGVENGVKYWLIVANSWNEEWGEGYFRMRGRNNECGIEARINAGLP-----
BN1106_s6570B000050 WGVENGVKYWLIVANSWNVGGENGYFRILRGTDECRIESIVVAGMPLRQKNIITNH
WGVENGVKYWLIVANSWNVGGENGYFRMLRGTDECRIESIVVAGMPLRQKNIITNRH

```