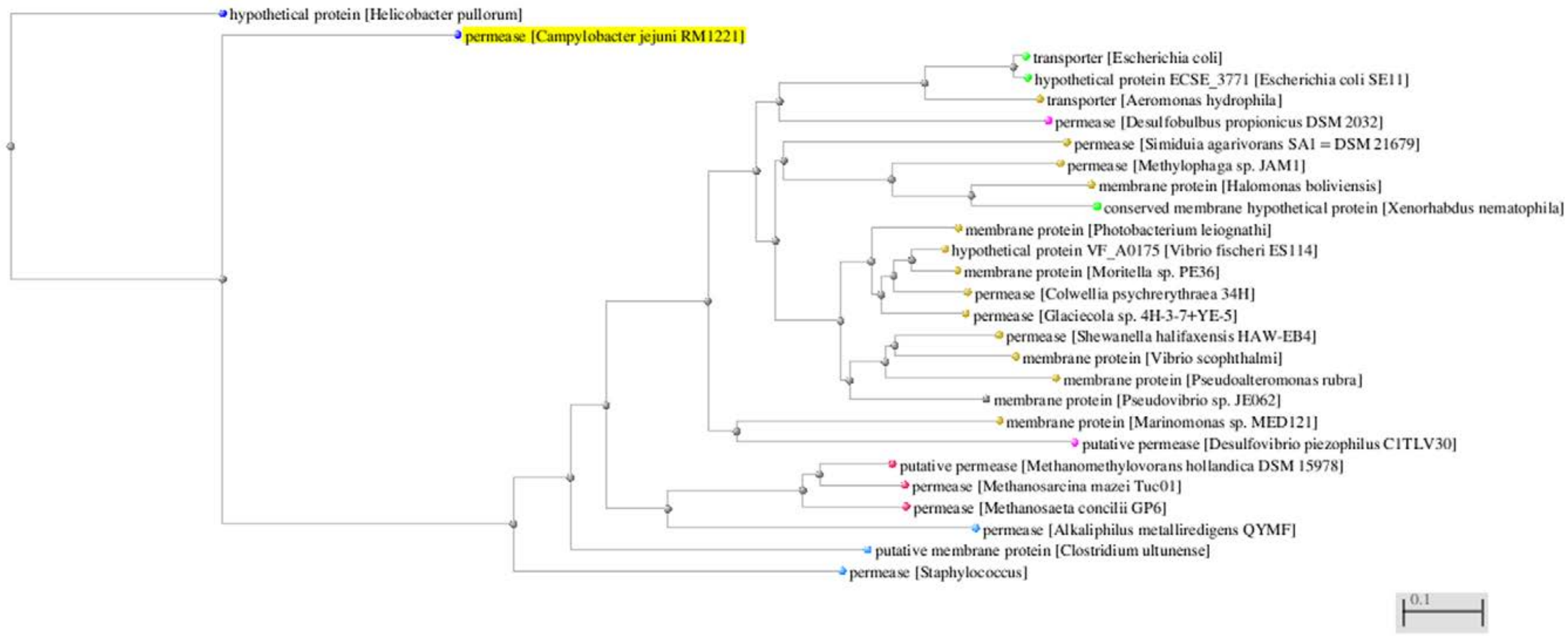
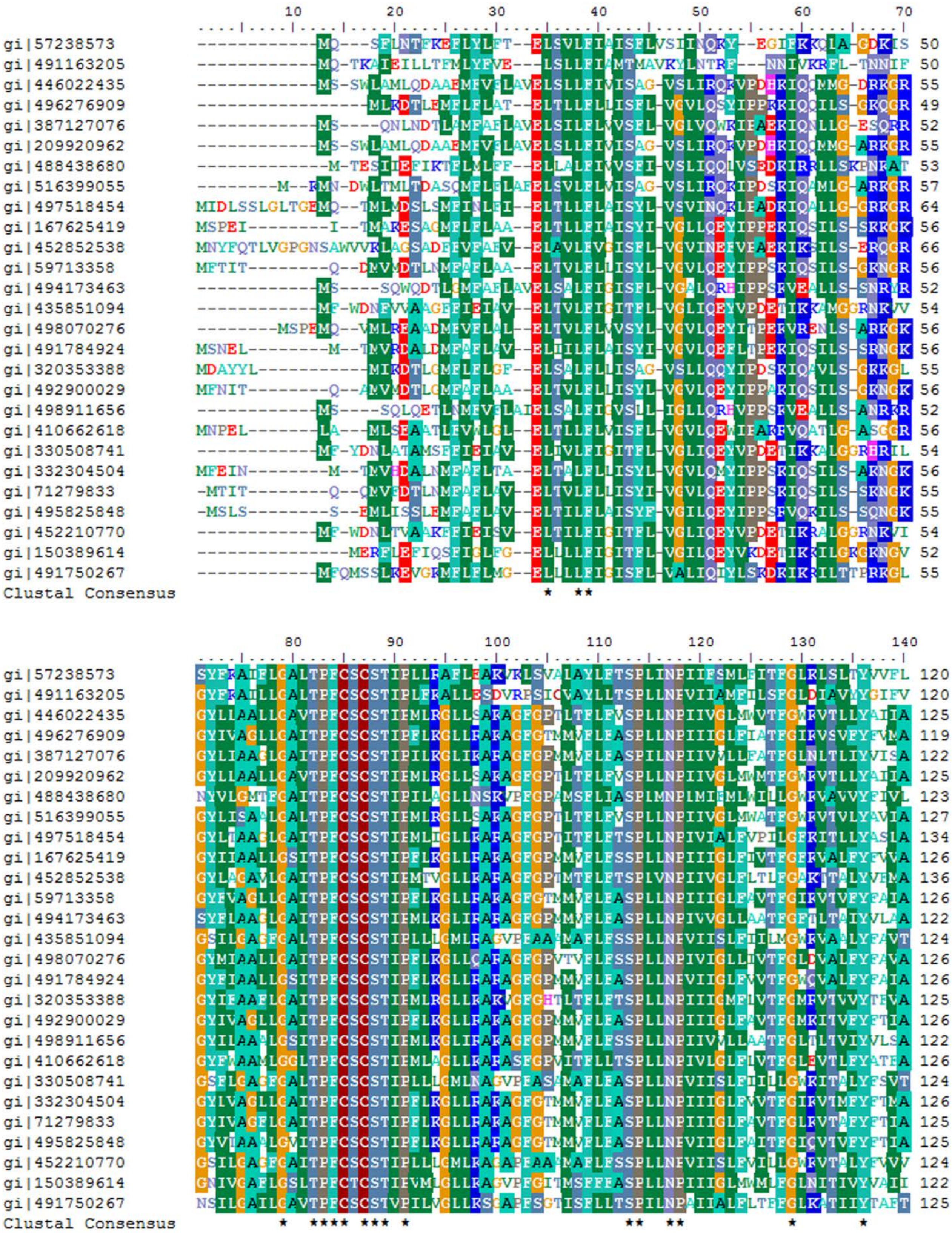
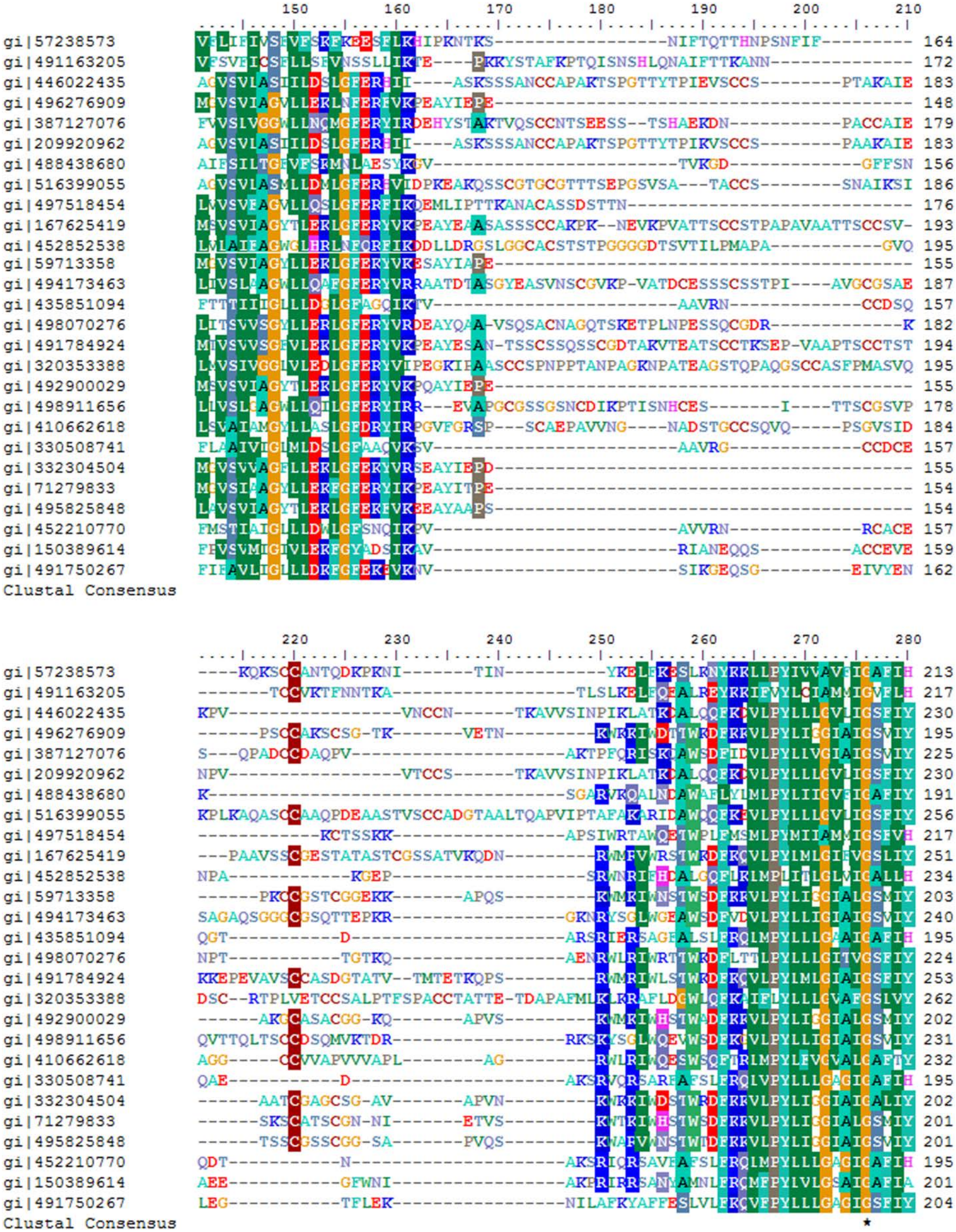


Supplementary FIG. 1. Phylogenetic analysis of full-length ArsP homologs in bacteria and archara. The tree was generated using ArsP homologs from representive species in 27 different genera and the Fast Minimum Evolution method with Constraint-based Multiple Protein Alignment Tool (COBALT) from National Center for Biotechnology information (NCBI). The colors of the dots indicate the phyla or classes of these species. Navy blue color represents Epsilonproteobacteria. Green color represents Enterobacteria. Yellow color represents Gammaproteobacteria. Pink represents Deltaproteobacteria. Grey color represents Alphaproteobacteria. Red color represents Euryarchaeotes. Blue color represents Firmicutes.



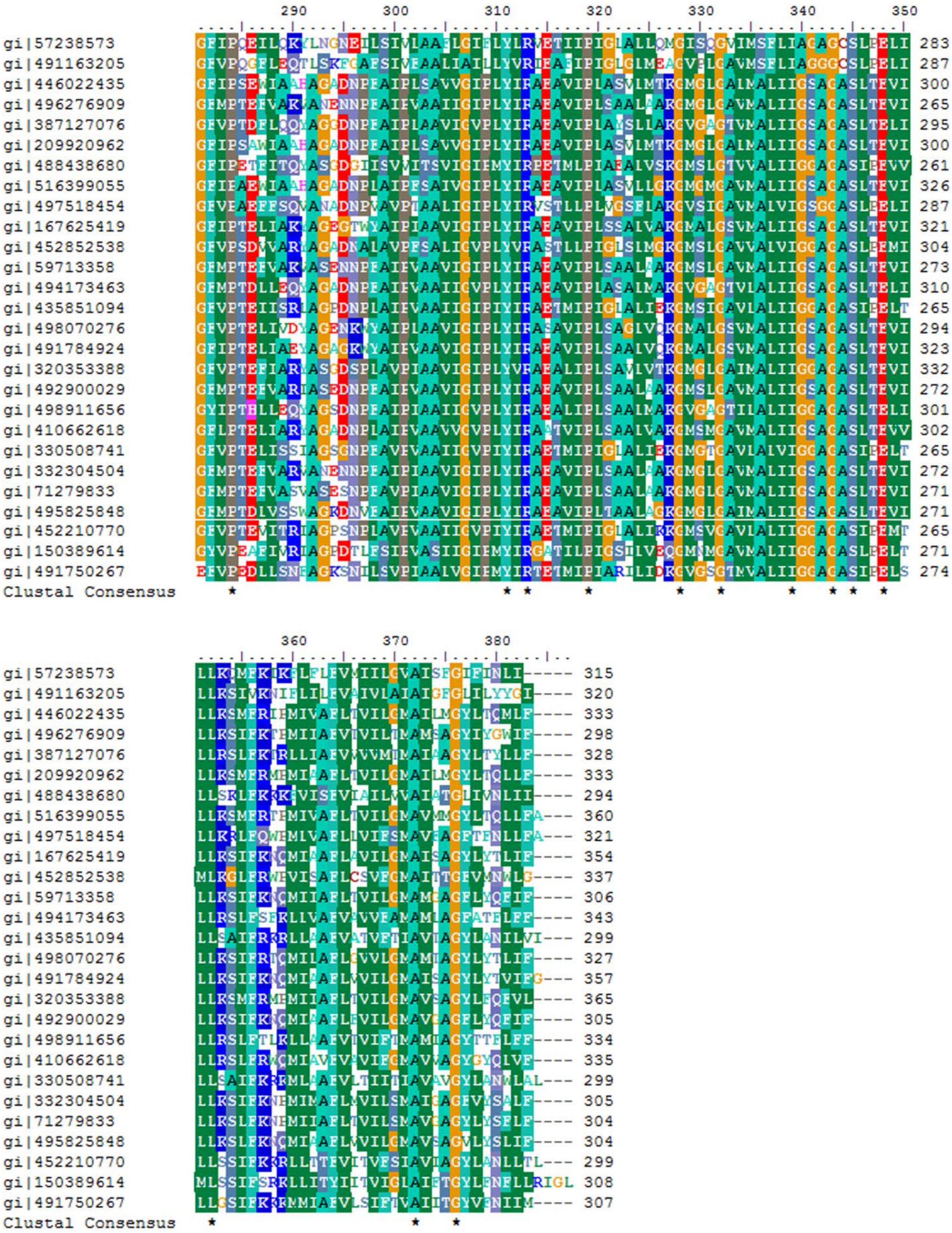
```
                           10        20        30        40        50        60        70
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
gi|57238573        ------------MQ---SFLNTFKEFLYLFT--ELSVLFIAISFLVSIINQKY--EGIFKKQLA-GDKIS 50
gi|491163205       ------------MQ-TKAIEILLTFMLYFVE---LSLLFIAMTMAVKYLNTRF---NNIVKRFL-TNNIF 50
gi|446022435       ------------MS-SWLAMLQDAAEMFVFLAVELSLLFIVISAG-VSLIRQKVPDHKIQQMMG-DRKGR 55
gi|496276909       -----------------MLKDTLEMFLFLAT--ELTLLFLLISFL-VGVLQSYIPPKKIQQILS-GKQGR 49
gi|387127076       ------------MS----QNLNDTLAMFAFLAVELSILFLVVSFL-VGLVQWKIPAEKIQNLLG-ESQRR 52
gi|209920962       ------------MS-SWLAMLQDAAEMFVFLAVELSLLFIVISAG-VSLIRQKVPDHKIQQMMG-ARKGR 55
gi|488438680       -------------M-TESIIEFIKTFLMLFF--ELLALFIVVSFI-VSLIQQIVSEDKIRRLLSKPNKAT 53
gi|516399055       --------M--KMN-DWLTMLTDASQMFLFLAFELSVLFLVISAG-VSLIRQKIPDSKIQAMLG-ARKGR 57
gi|497518454       MIDLSSLGLTGEMQ--TMLMDSLSMFINLFI--ELTLLFLAISYL-VSVINQKLFADKIQALLG-GRKGR 64
gi|167625419       MSPEI--------I--TMAKESAGMFLFLAA--ELTLLFIAISYI-VGLLQEYIPPEKIQSILS-SKKGK 56
gi|452852538       MNYFQTLVGPGNSAWVVKLAGSADFFVFAFV--ELAVLFVGISFL-VGVINEFVFAEKIKSILS-ERQGR 66
gi|59713358        MFTIT--------Q--DMVMDTLNMFAFLAA--ELTVLFLLISYL-VGVLQEYIPPSKIQSILS-GKNGR 56
gi|494173463       ------------MS----SQWQDTLGMFAFLAVELSALFIGISFL-VGALQRHIPPSKVEALLS-SNRYR 52
gi|435851094       ------------MF-WDNFVVAAGFFIEIAV--ELTVLFIGITFL-VGLIQEYVPDETIKKAMGGRNKVV 54
gi|498070276       --------MSPEMQ--VMLREAADMFVFLAL--ELTVLFLVVSYL-VGVLQEYITPEKVRENLS-ARKGK 56
gi|491784924       MSNEL--------M--TMVRDALDMFAFLAV--ELIILFLAISYI-VGVLQEFLTPEKIQSILS-SRNGK 56
gi|320353388       MDAYYL-----------MIKDTLGMFLFLGF--ELSALFLLISAG-VSLLQQYIPDSKIQAVLS-GKKGL 55
gi|492900029       MFNIT--------Q--AMVMDTLGMFAFLAA--ELTVLFLLISYL-VGVLQEYIPPAKIQSILS-GKNGK 56
gi|498911656       ------------MS----SQLQETLNMFVFLAIELSALFIGVSLL-IGLLQRHVPPSKVEALLS-ANRKR 52
gi|410662618       MNPEL-------LA---MLSEAATLFVWLGL--ELTLLFLVISFL-VGVLQEWIFAKFVQATLG-ASGGR 56
gi|330508741       ------------MF-YDNLATAMSFFIEIAV--ELIVLFIGITFL-VGLIQEYVPDETIKKALGGRHRIL 54
gi|332304504       MFEIN--------M--TMVHDALNMFAFLTA--ELTALFLLISYL-VGVLQMYIPPSKIQSILS-AKNGK 56
gi|71279833        -MTIT--------Q--QMVFDTLNMFAFLAV--ELTVLFLLISYI-VGVLQEYIPPSKIQSILS-SKNGK 55
gi|495825848       -MSLS--------S--EMLISSLEMFAFLAV--ELTILFLAISYF-VGILQEYIPPSFVQKILS-SQNGK 55
gi|452210770       ------------MF-WDNLTVAAKFFIEISV--ELTILFIGITFL-VGLIQEYVPDETIKRALGGRNKVI 54
gi|150389614       ---------------MERFLEFIQSFIGLFG--ELLLLFIGITFL-VGILQEYVKDETIKKILGKGKNGV 52
gi|491750267       ------------MFQMSSLKEVGKMFLFLMG--ELLLLFIGISFL-VALIQTYLSKDKIKRILTTPRRGL 55
Clustal Consensus                                    *  **

                           80        90       100       110       120       130       140
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
gi|57238573        SYFKAIFLGALTPFCSCSTIPLLRAFLEAKVRLSVALAYLFTSPLINPIIFSMLFITFGLKLSLTYVVFL 120
gi|491163205       GYFKAILLGALTPFCSCSTIPLFKALLESDVRPSICVAYLLTSPLLNPVIIAMFILSFGLDIAVYYGIFV 120
gi|446022435       GYLLAALLGAVTPFCSCSTIPMLRGLLSAKAGFGPTLTFLFVSPLLNPIIVGLMWVTFGWKVTLLYAIIA 125
gi|496276909       GYIVAGLLGAITPFCSCSTIPFLKGLLFAKAGFGTMMVFLFASPLLNPIIIGLFIATFGIKVSVFYFVMA 119
gi|387127076       GYLIAAGLGAITPFCSCSTIPILKGLIKAFAGFGPMMVFLFASPILNPIIVVLLFATFGLNLTLIYVISA 122
gi|209920962       GYLLAALLGAVTPFCSCSTIPMLRGLLSAKAGFGPTLTFLFVSPLLNPIIVGLMWMTFGWKVTLLYAIIA 125
gi|488438680       NYVLGMTFGAITPFCSCSTIPILAGLLNSKVPFGPAMSFLIASPLMNPLMIFMLWILLGWKVAVVYFIVL 123
gi|516399055       GYLISAALGALTPFCSCSTIPMLRGLLSAKAGFGPTLTFLFVSPLLNPIIVGLMWATFGWKVTVLYAVIA 127
gi|497518454       GYLIAAGLGAITPFCSCSTIPMLIGLLKAFAGFGPTITFLFTSPLLNPIVIALFVPILGFKITLLYASIA 134
gi|167625419       GYIIAALLGSITPFCSCSTIPFLKGLLFAKAGFGPMMVFLFSSPLLNPIIIGLFIVTFGFKVALFYFVVA 126
gi|452852538       GYLAGAVLGAITPFCSCSTIPMTVGLLKAFAGFGPTMTFLFTSPLVNPIIVGLFLTLFGAKTTALYVFMA 136
gi|59713358        GYFVAGLLGAITPFCSCSTIPFLKGLLFAKAGFGTMMVFLFASPLLNPIIIGLFAVTFGIKVTVFYFAIA 126
gi|494173463       SYFLAAGLGAITPFCSCSTIPMLKGLIFAFAGFGPMMVFLFASPLLNPIVVGLLAATFGFTLTAIYVLAA 122
gi|435851094       GSILGAGFGALTPFCSCSTIPLLLGMLFAGVPFAAAMAFLFSSPLLNPVIISLFIILMGWKVAALYFAVT 124
gi|498070276       GYMIAALLGAITPFCSCSTIPFLKGLLQAFAGFGPVTVFLFSSPLLNPIVIGLLIVTFGLDVALFYFAVA 126
gi|491784924       GYFIAALLGSITPFCSCSTIPFLKGLLFAFAGFGPMMVFLFASPLLNPVIIGLFVVTFGWQVALFYFAIA 126
gi|320353388       GYIFAAFLGAITPFCSCSTIPMLRGLLKAKVGFGHTLTFLFTSPLLNPIIIGMFLVTFGMFVTVVYTFVA 125
gi|492900029       GYIVAGLLGAITPFCSCSTIPFLKGLLFAKAGFGPMMVFLFASPLLNPIIIGLFAVTFGMKITVFYFTIA 126
gi|498911656       GYILAAALGSITPFCSCSTIPMLKGLIFAKAGFGPMMVFLFSSPLLNPIIVVLLAATFGLTLTVIYVLSA 122
gi|410662618       GYFVAAMLGGLTPFCSCSTIPMLAGLLKAFASFGPVITFLLTSPLLNPIVLGLFLVTFGLEVTLFYATFA 126
gi|330508741       GSFLGAGFGALTPFCSCSTIPLLLGMLNAGVPFASAMAFLFASPLLNPVIISLFIILLGWKITALYFSVT 124
gi|332304504       GYIVAGLLGAITPFCSCSTIPFLKGLLFAKAGFGTMMVFLFASPLLNPIIIGLFVVTFGIKVTMFYFTMA 126
gi|71279833        GYIVAGFLGAITPFCSCSTIPFLKGLLFAKAGFGTMMVFLFASPLLNPIIIGLFAVTFGLKVTAFYFTIA 125
gi|495825848       GYVIAAALGVITPFCSCSTIPFLKGLLFAFAGFGTMMVFLFASPLLNPVIIGLFAITFGIQVTVFYFTIA 125
gi|452210770       GSILGAGFGAITPFCSCSTIPLLLGMLKAGAPFAAAMAFLFSSPLLNPVIISLFVILLGWKVTALYFVVV 124
gi|150389614       GNIVGAFLGSLTPFCTCSTIPVMLGLLKAGVPFGITMSFFFASPLLNPIIIGLMWMLFGLNITIVYVAII 122
gi|491750267       NSILGAILGAVTPFCSCSTVPILVGLLKSGAPFSGTISFLLTSPILNPAIIALFLTFFGLKATIIYTAFT 125
Clustal Consensus          *  **** *** *                     **  **          *      *
```



```
                              150       160       170       180       190       200       210
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
gi|57238573           VFLIFIVSFVFSKFKEESFLKHIPKNTKS------------------NIFTQTTHNPSNFIF--------- 164
gi|491163205          VFSVFICSFLLSFVNSSLLIKTE----PKKYSTAFKPTQISNSHLQNAIFTTKANN-------------- 172
gi|446022435          AGVSVLASIILDSLGFERHII----ASKSSSANCCAPAKTSPGTTYTPIEVSCCS--------PTAKAIE 183
gi|496276909          MGVSVIAGVLLERLNFERFVKPEAYIEPE----------------------------------------- 148
gi|387127076          FVVSLVGGWLLNQMGFERYIRDEHYSTAKTVQSCCNTSEESS--TSHAEKDN-----------PACCAIE 179
gi|209920962          AGVSVLASIILDSLGFERHII----ASKSSSANCCAPAKTSPGTTYTPIKVSCCS--------PAAKAIE 183
gi|488438680          AIFSILTGFVFSKMNLAESYKGV------------------------TVKGD-------------GFFSN 156
gi|516399055          AGVSVLASMLLDMLGFERFVIDPKEAKQSSCGTGCGTTTSEPGSVSA---TACCS--------SNAIKSI 186
gi|497518454          LVVSVFAGVLLQSLGFERFIKQEMLIPTTKANACASSDSTTN---------------------------- 176
gi|167625419          MSVSVIAGYTLERLGFERYVKPEAYEAASASSSCCAKPK--NEVKPVATTSCCSTPAPAVAATTSCCSV- 193
gi|452852538          LVLAIFAGWGLHRLNFQRFIKDDLLDRGSLGGCACSTSTPGGGGDTSVTILPMAPA-----------GVQ 195
gi|59713358           MGVSVIAGYLLERLGFERYVKESAYIAPE----------------------------------------- 155
gi|494173463          LIVSLAAGWLLQAFGFERYVRRAATDTASGYEASVNSCGVKP-VATDCESSSCSSTPI----AVGCGSAE 187
gi|435851094          FTTTIIIGLLLDGLGFAGQIKTV------------------------AAVRN-------------CCDSQ 157
gi|498070276          LITSVVSGYLLERLGFERYVRDEAYQAA-VSQSACNAGQTSKETPLNPESSQCGDR-------------K 182
gi|491784924          MIVSVVSGFVLERLGFERYVKPEAYESAN-TSSCSSQSSCGDTAKVTEATSCCTKSEP-VAAPTSCCTST 194
gi|320353388          LMVSIVGGLVLEDLGFERYVIPEGKIPAASCCSPNPPTANPAGKNPATEAGSTQPAQGSCCASFPMASVQ 195
gi|492900029          MSVSVIAGYTLERLGFERYVKPQAYIEPE----------------------------------------- 155
gi|498911656          LLVSLGAGWLLQILGFERYIRR---EVAPGCGSSGSNCDIKPTISNHCES-------I----TTSCGSVP 178
gi|410662618          LSVAIAMGYLLASLGFDRYIRPGVFGRSP---SCAEPAVVNG----NADSTGCCSQVQ-----PSGVSID 184
gi|330508741          FLAAIVIGLMLDSLGFAAQVKSV------------------------AAVRG-------------CCDCE 157
gi|332304504          MGVSVVAGFLLERLGFERYVRSEAYIEPD----------------------------------------- 155
gi|71279833           MGVSIAAGYLLEKFGFERYIKPEAYITPE----------------------------------------- 154
gi|495825848          LAVSVIAGYTLERLGFEKFVKEEAYAAPS----------------------------------------- 154
gi|452210770          FMSTIAIGLLLDWLGFSNQIKPV------------------------AVVRN-------------RCACE 157
gi|150389614          FPVSVMIGIVLEKFGYADSIKAV------------------------RIANEQQS---------ACCEVE 159
gi|491750267          FIFAVLIGLLLDKFGFEKEVKNV------------------------SIKGEQSG---------EIVYEN 162
Clustal Consensus

                              220       230       240       250       260       270       280
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
gi|57238573           ----KQKSCCANTQDKPKNI-------TIN----------YKELFKESLRNYRKLLPYIVVAVFIGAFIH 213
gi|491163205          -------TCCVKTFNNTKA------------------TLSLKELFQEALREYRKIFVYLCIAMMIGVFLH 217
gi|446022435          KPV-----------------VNCCN------TKAVVSINPIKLATKCALQQFKCVLPYLLLGVLIGSFIY 230
gi|496276909          ------PSCCAKSCSG-TK-------VETN---------KWKKIWDTTWKDFRKVLPYLIGGIAIGSVIY 195
gi|387127076          S---QPADCCDAQPV---------------------AKTPFQRISKCAWSDFIDVLPYLLVGIAIGSVIY 225
gi|209920962          NPV-----------------VTCCS------TKAVVSINPIKLATKCALQQFKCVLPYLLLGVLIGSFIY 230
gi|488438680          K-----------------------------------SGARVKQALNDAWAFLYIMLPYLIIGVFIGAFIY 191
gi|516399055          KPLKAQASCCAAQPDEAASTVSCCADGTAALTQAPVIPTAFAKARIDAWQQFKFVLPYLLLGVLIGSFIY 256
gi|497518454          ------------KCTSSKK-----------------APSIWRTAWQETWPLFMSMLPYMIIAMMIGSFVH 217
gi|167625419          ---PAAVSSCGESTATASTCGSSATVKQDN---------RWMFVWRSTWKDFKQVLPYIMLGIFVGSLIY 251
gi|452852538          NPA------------KGEP-------------------SRWNRIFHDALGQFLKIMPLITLGLVIGALLH 234
gi|59713358           ------PKCCGSTCGGEKK-------APQS---------KWMKIWNSTWSDFRKVLPYLIGGIALGSMIY 203
gi|494173463          SAGAQSGGGCGSQTTEPKR-----------------GKNRYSGLWGEAWSDFVDVLPYLLIGIAIGSVIY 240
gi|435851094          QGT-----------D---------------------ARSRIERSAGFALSLFRQLMPYLLLGAAIGAFIH 195
gi|498070276          NPT-----------TGTKQ-----------------AENRWLRIWRTTWKDFLTTLPYLLLGITVGSFIY 224
gi|491784924          KKEPEVAVSCCASDGTATV--TMTETKQPS---------RWMRIWLSTWKDFKQVLPYIMLGIAIGSFIY 253
gi|320353388          DSC--RTPLVETCCSALPTFSPACCTATTE-TDAPAFMLKLKRAFLDGWLQFKAIFLYLLLGVAFGSLVY 262
gi|492900029          ------AKGCASACGG-KQ-------APVS---------KWMKIWHSTWADFRKVLPYLIGGIALGSMIY 202
gi|498911656          QVTTQLTSCCDSQMVKTDR-----------------RKSKYSGLWQEVWSDFKCVLPYLLIGIAIGSVIY 231
gi|410662618          AGG-----CCVVAPVVVAPL--------AG---------RWLRIWQESWSQFTRLMPYLFVGVALGAFTY 232
gi|330508741          QAE-----------D---------------------AKSRVQRSARFAFSLFRQLVPYLLLGAGIGAFIH 195
gi|332304504          ------AATCGAGCSG-AV-------APVN---------KWKKIWDSTWRDFRKVLPYLIGGIAIGALIY 202
gi|71279833           ------SKSCATSCGN-NI-------ETVS---------KWTKIWHSTWSDFRKVLPYLIGGIALGSMIY 201
gi|495825848          ------TSSCGSSCGG-SA-------PVQS---------KWAFVWNSTWTDFRKVLPYLIGGIAIGSVIY 201
gi|452210770          QDT-----------N---------------------AKSRIQRSAVFAFSLFRQLMPYLLLGAGIGAFIH 195
gi|150389614          AEE-----------GFWNI-----------------AKPRIRRSANNAMNLFRQMFPYLVLGSAIGAFIA 201
gi|491750267          LEG-----------TFLEK-----------------NILAFKYAFFESLVLFRQVFPYLLLGAGIGSFIY 204
Clustal Consensus                                                                      *
```



```
                            290       300       310       320       330       340       350
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
gi|57238573          GFIPQEILQKYLNGNEILSIVLAAFLGIFLYLRVETIIPIGLALLQMGISQGVIMSFLIAGAGCSLPELI 283
gi|491163205         GFVPQGFLEQTLSKFGAFSIVFAALIAILLYVRIEAFIPIGLGLMEAGVFLGAVMSFLIAGGGCSLPELI 287
gi|446022435         GFIPSEWIAAHAGADNPFAIPLSAVVGIPLYIRAEAVIPLASVLMTKGMGLGAIMALIIGSAGASLTEVI 300
gi|496276909         GFMPTEFVAKVANENNPFAIPVAAIIGIPLYIRAEAVIPLSAALAAKGMGLGAVMALIIGSAGASLTEVI 265
gi|387127076         GFVPTDFLQQYAGGDNPFAIPLAAVIGVPLYIRAEAVIPLAYSLLAKGVGAGTVMALIIGSAGASLTELI 295
gi|209920962         GFIPSAWIAAHAGADNPFAIPLSAVVGIPLYIRAEAVIPLASVLMTKGMGLGAIMALIIGSAGASLTEVI 300
gi|488438680         GFIPETFITQYASGDGIISVVITSVIGIPMYIRPETMLPIAEALVSKGMSLGTVVALIIGGAGASIPEVV 261
gi|516399055         GFIPAEWIAAHAGADNPLAIPFSAIVGIPLYIRAEAVIPLASVLLGKGMGMGAVMALIIGSAGASLTEVI 326
gi|497518454         GFVPAEFFSQVANADNPVAVPTAALIGIPLYIRVSTLLPLVGSFLAKGVSIGAVMALVIGSGGASLPELI 287
gi|167625419         GFIPTELIAKYAGEGTWYAIPIAAVIGIPLYIRAEAVIPLSSALVAKGMALGSVMALIIGSAGASLTEVI 321
gi|452852538         GFVPSDVVARYAGADNALAVPFSALIGVPLYVRASTLLPIGLSLMGKGMSLGAVVALVIGGAGASLPEMI 304
gi|59713358          GFMPTEFVAKVASENNPFAIPVAAVIGIPLYIRAEAVIPLSAALAAKGMSLGAVMALIIGSAGASLTEVI 273
gi|494173463         GFMPTDLLEQYAGADNPFAIPVAAIIGVPLYIRAEAVIPLASALMAKGVGAGTVLALIIGSAGASLTELI 310
gi|435851094         GFVPTEIISRIAGPDNPLAIPVAAIIGIPIYIRAETMIPIGLALIEKGMSIGAVLALIIGGAGASIPELT 265
gi|498070276         GFVPTELIVDYAGENKWYAIPLAAVIGIPLYIRASAVIPLSAGLVQKGMALGSVMALIIGSAGASLTEVI 294
gi|491784924         GFIPTELIAEYAGAGKWYAIPVAAVIGIPLYIRAEAVIPLSAALVQKGMALGSVMALIIGSAGASLTEVI 323
gi|320353388         GFVPTEFIARYASGDSPLAVPIAAVIGIPLYVRAEALIPLSAVLVTKGMGLGAIMALIIGGAGASLTEVI 332
gi|492900029         GFMPTEFVARIASEDNPFAIPVAAVIGIPLYIRAEAVIPLSAALAAKGMSLGAVMALIIGSAGASLTEVI 272
gi|498911656         GYIPTHLLEQYAGSDNPFAIPIAAIIGVPLYIRAEALIPLSAALMAKGVGAGTILALIIGGAGASLTELI 301
gi|410662618         GFLPTELIARYAGADNPLAIPVAAVVGVPLYIRAATVIPLSAALVAKGMSMGAVMALIIGSAGASLTEVV 302
gi|330508741         GFVPTELISSIAGSGNPFAVPVAAIIGVPIYIRAETMIPIGLALIEKGMGTGAVLALVIGGAGASIPELT 265
gi|332304504         GFMPTEFVARVANENNPFAIPIAAVIGIPLYIRAEAVIPLSAALAAKGMGLGAVMALIIGSAGASLTEVI 272
gi|71279833          GFMPTEFVASVASESNPFAVPIAAVIGIPLYIRAEAVIPLSAALAAKGMGLGAVMALIIGSAGASLTEVI 271
gi|495825848         GFMPTDIVSSWAGKDNVFAIPVAAVIGIPLYIRAEAVIPLTAALAGKGMGLGAIMALIIGSAGASLTEVI 271
gi|452210770         GFVPTEVITRIAGPSNPLAVPVAAIIGVPIYIRAETMIPIGLALIKKGMSVGAVLALIIGGAGASIPEMT 265
gi|150389614         GYVPEAFIVRIAGPDTLFSIPVASIIGIPMYIRGATILPIGSILVEQGMNMGAVMALIIGGAGASLPELT 271
gi|491750267         EFVPEDLLSNFAGKSNILSVPIAALVGIPMYIRTETMIPIARILIDKGVGSGTMVALIIGGAGASIPELS 274
Clustal Consensus       *                          * *     *        *   *      *   * *  *

                            360       370       380
                     ....|....|....|....|....|....|....|..
gi|57238573          LLKQMFKIKFLFLFVMIILGVAISFGIFINLI----- 315
gi|491163205         LLKSIVKNIFLILFVAIVLAIAIGFGLILYYGI---- 320
gi|446022435         LLKSMFRIFMIVAFLTVILGMAILMGYLTQMLF---- 333
gi|496276909         LLKSIFKTFMIIAFVTVILTMAMSAGYIYGWIF---- 298
gi|387127076         LLRSLFKTRLLIAFVVVVMTMAIAAGYLTYLLF---- 328
gi|209920962         LLKSMFRMFMIAAFLTVILGMAILMGYLTQLLF---- 333
gi|488438680         LLSKLFKKKFVISFVIAILVVAIATGLIVNLII---- 294
gi|516399055         LLKSMFRTFMIVAFLTVILGMAVMMGYLTQLLFA--- 360
gi|497518454         LLKRLFQWFMIVAFLIVIFSMAVFAGFTFNLLFA--- 321
gi|167625419         LLKSIFKNQMIAAFLAVILGMAISAGYLYTLIF---- 354
gi|452852538         MLKGLFRWFVISAFLCSVFGMAITTGFVMNWLG---- 337
gi|59713358          LLKSIFKNQMIIAFLTVILGMAMGAGFLYQFIF---- 306
gi|494173463         LLRSLFSFKLLVAFVAVVFAMAMLAGFATFLFF---- 343
gi|435851094         LLSAIFRKRLLAAFVATVFTIAVIAGYLANILVI--- 299
gi|498070276         LLKSIFRTQMILAFLGVVLGMAMIAGYLYTLIF---- 327
gi|491784924         LLKSIFKNQMIAAFLVVILGMAISAGYLYTVIFG--- 357
gi|320353388         LLKSMFRMFMIIAFLTVILGMAVSAGYLFQFVL---- 365
gi|492900029         LLKSIFKNQMIAAFLFVILGMAVGAGFLYQFIF---- 305
gi|498911656         LLRSLFTLKLLAAFVTVIFTMAMIAGYTTFLFF---- 334
gi|410662618         LLRSLFRWQMIAVFVAVIFGMAVVAGYGYQLVF---- 335
gi|330508741         LLSAIFKRRMLAAFVLTIITIAVAVGYLANWLAL--- 299
gi|332304504         LLKSIFKNFMIMAFLAVILSMAIGAGFVYSALF---- 305
gi|71279833          LLKSLFKNFMIIAFLTVILSMAVGAGYLYSFLF---- 304
gi|495825848         LLKSLFKNQMIAAFLVVILGMAVSAGVLYSLIF---- 304
gi|452210770         LLSSIFKKRLLTTFVITVFSIAVIAGYLANLLTL--- 299
gi|150389614         MLSSIFSRKLLITYIITVIGLAIFTGYLFNFLLRIGL 308
gi|491750267         LLGSIFKKRMMIAFVLSIFTVAIITGYVFNIIM---- 307
Clustal Consensus     *                   *   *
```

Supplementary FIG. 2. Sequence alignment of ArsP homologs in various bacteria and archara using COBALT. The lengths of these proteins range from 294 to 365 amino acids. Compared with *C. jejuni*, the identities are greater than 30% (>5E-43). The sequences are more conserved at the N- and C-terminal regions, while deletions and insertions occur in the middle region of the proteins, which corresponds to the large hydrophilic loop shown in Fig. 2. In order to see the conserved and diverse regions clearly, different amino acids are indicated by different colors. "*" indicates identical residues among the 27 homologous.