Supplemental data

```
                  20                 30        40                               50                    60
                  |                  |         |                                |                     |
NP_288269.1      MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002403065.1   MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002397981.1   MNK---------AEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002412850.1   MNK---------IEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_001880631.1   MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002391608.1   MNK---------AEFYADLNRDFNALMVGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_001730808.1   MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------INWA
ZP_03003290.1    MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_001743413.1   MNK---------AEFYADFNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002407226.1   MNK---------AEFYSDFNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002329476.1   MNK---------AGFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_002382399.1   MNK---------TEFYADLNRDFRALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
ZP_03445135.1    MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_688895.1      MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
YP_669680.1      MNK---------AEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------VNWA
NP_416346.4      MNK---------TEFYADLNRDFNALMAGETS-----------------------FLATLANTSALLYERLTD------------INWA
ZP_02903917.1    MNK---------TEFYADLNRDFNALMTGETS-----------------------FLATLANTSALLYERLAD------------VNWA
ZP_03282954.1    MNK---------SEFYADLNRDFQALMAGETS-----------------------FLATLANTSALLFERLSG------------VNWA
YP_001452718.1   MSK---------TELYADLNRDFQAVMAGETS-----------------------FLATLANTSALLFERLSG------------VNWA
YP_001570142.1   MSE---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
NP_460803.1      MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_150310.1      MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002141797.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002243298.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
NP_456352.1      MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
NP_804849.1      MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_001587603.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02666379.2    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02832604.2    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_001177122.1   MNK---------TEFYADLNRDFQALMTGETS-----------------------FLATLANTSALLFERLSD------------VNWA
YP_002226291.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
NP_754135.1      --------------------------MAGETS-----------------------FLXTLANTSALLYERLTD------------VNWA
ZP_02347214.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02574893.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02659310.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02663512.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02681881.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_02696584.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002041101.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002045892.1   MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
```

Supplemental data

```
ZP_03075263.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002114879.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002146178.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03165709.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002215294.1    MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03221812.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03342093.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03347154.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03352578.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03357060.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03371453.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03379701.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03383790.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03412299.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
ZP_03217260.1     MSK---------TELYAALNRDFQSLMAGETS-----------------------FLATLANTSALLFERLTE------------VNWA
YP_002237781.1    MSK---------TEFYADLNRDFQALMAGETS-----------------------FLAMIANTSALLFERLSE------------VNWA
YP_001336004.1    MSK---------TEFYADLNRDFQALMAGETS-----------------------FLAMIANTSALLFERLSE------------VNWA
YP_001437512.1    MNK---------TQFYAELNRDFHSLMAGETS-----------------------FLATLANTSALLFERLEG------------VNWA
ZP_03403067.1     --------------------------------------------------------MATLANTSALLFERLTE------------VNWA
YP_001907453.1    MTK---------AEFYTDLNRDMNALLSGETS-----------------------FLAVMSNCSALLYKRLDG------------VNWA
YP_001006127.1    MKK---------AEFYAELKRDLSALIAGETN-----------------------FIATLANASALIYERLDG------------LNWA
NP_669181.1       MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
NP_993079.1       MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_070893.1       MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_651729.1       MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_647856.1       MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001163103.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_01888376.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001400647.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001607069.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02221459.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02227087.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02229339.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02237158.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02307206.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02314388.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
ZP_02318544.1     MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001720525.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001872871.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_002346706.1    MKK---------TEFYAELKRDLCALVAGETN-----------------------FIATLANASALLYERLDG------------LNWA
YP_001478346.1    MTK---------EQFYAELKRDLCALLDGETN-----------------------FIAALSNASALINERLDD------------VNWV
YP_454994.1       MKK---------QQFYAELLQDMSALLDEETH-----------------------FIAALANASALLYERLEG------------VNWA
```

Supplemental data

```
YP_001217052.1    EVFIFQMEVIMNLEQYQRLTKQAVALLEGETN-----------------------LIANLANLSALLNMELTE------------LNWV
YP_002150764.1    MPQ---------LNFYQELTDNLSALLSGEND-----------------------FIASLANTSALIYEKLDK------------INWV
YP_856816.1       IEK---------QAFYRTLNQQAQALLEGEPD-----------------------LIANLANLSALLNQELAD------------INWV
YP_001093925.1    MST----------PCYTTLTRQVEALFSGEDN-----------------------LVTAMANFSALLNEHLDD------------INWV
YP_269525.1       MSN--------QAEFYHELLGQTKALINDESD-----------------------IIANMANISALLFERLAD------------VNWV
NP_718189.1       MKP----------QFYESLNRQVLALLEGEDD-----------------------LVAAMANFSALLNDNLTE------------LNWV
YP_001366642.1    MKS----------KFYESLNRQALALLEGEDD-----------------------LVAAMANFSALLNDNLTE------------LNWV
YP_001554990.1    MKS----------KFYESLNRQALALLEGEDD-----------------------LVAAMANFSALLNDNLTE------------LNWV
YP_737796.1       MKP----------QFYESLNRQALALLEGEDD-----------------------LVAAMANFSALLNDNLTE------------LNWV
YP_869415.1       MKP----------QFYESLNRQALALLEGEDD-----------------------LVAAMANFSALLNDNLTE------------LNWV
NP_797987.1       ----------MNIEQYQRLTKQAVALIESEPD-----------------------FIANLANLSSLLFMELED------------LNWA
YP_001501790.1    MVN----------AFYDTLNRQVLALLDGEDD-----------------------TIAALANFSALLNDNLTD------------LNWV
YP_927460.1       MKH----------AFYDTLNRQAEALLSGEDD-----------------------LVAAMANFSALLNEHMDD------------LNWV
YP_001474222.1    MIT----------EFYSSLDRQAKALLDGEDD-----------------------LIAAMANFSALINDALDD------------LNWV
YP_963195.1       MKS----------QFYESLNRQALALLEGEDD-----------------------VVAAMANFSALLNDNLTD------------LNWV
YP_001183724.1    MKS----------QFYESLNRQALALLEGEDD-----------------------VVAAMANFSALLNDNLTD------------LNWV
YP_733804.1       MKP----------QFYESLNRQALALLEGEDD-----------------------LVAAMANFSAFLNDNLTE------------LNWV
YP_001050811.1    MKS----------KFYESLNRQALALLEGEDD-----------------------LVAAMANFSTLLNDNLTE------------LNWV
YP_750914.1       MKT----------EFYQTLNRQAQALLEGEDD-----------------------LIAAMANFSALLNDSLED------------LNWA
YP_002126297.1    MK----------EPQYSQLINQAKSLVSGEHD-----------------------LIANMANISALLFNHLEE------------VNWA
YP_130129.1       ----------------------------DEPD-----------------------LIANISNISALLNMELED------------INWV
YP_001445537.1    ----------MNIEQYQRLTKQAVALIESEKD-----------------------YIANLSNISALLFMELEN------------LNWA
YP_661562.1       MTLS-------KTDLYQEVAQQAQALMAGEKD-----------------------LIANLANLSALMYMRLPE------------LNWA
YP_001674584.1    MVD----------AFYDTLNRQVQALFADEDD-----------------------TIAALANFSALINDSLDD------------LNWV
YP_002417027.1    ----------MKIEHYQRLTKQAVALIESEPD-----------------------LTANLANISSLLFMELDE------------LNWA
NP_934438.1       ----------MTIEHYQTLTKQAVALIESERD-----------------------LIANLSNISALLNMGMSD------------LNWV
NP_761466.1       ----------MTIEHYQTLTKQAVALIESERD-----------------------LIANLSNISALLNMGMSD------------LNWV
YP_001760510.1    MDL----------ECYDSLNRQAQALLSGEDD-----------------------LVAAMANFSALINDAVEQ------------LNWV
YP_002312133.1    MTA----------SFYDTLNRQVIALLSGEDD-----------------------QIAALANFSALLNDNLTD------------LNWV
YP_562727.1       MKA----------DFYDTLDRQASALLAGEDD-----------------------VIAAMANFSALLNESLED------------LNWV
NP_693123.1       MFEQKNYSG-EKAKDYELLIKQLDALSEGENN-----------------------KIALLSNASALLNQFLTE------------VNWV
YP_001699747.1    MFTQINYEG-SIADQYLTLSKQLDALLTGEKD-----------------------QIANLSNASALLHQFLTN------------INWV
YP_001343242.1    TIETHS--TGSSEEFYQELNIQLAALLSGERD-----------------------FICNAAQFSAFVMQTVTD------------LNWS
YP_001141806.1    IEK---------QAFYRGLNQQAQALLEGEPD-----------------------LIANLANLSALLALELTD------------INWV
YP_204661.1       ----------MELKHYARLTQQALGLIEGETD-----------------------YIANLANISALLNMELDD------------INWV
YP_301135.1       MTEVK-------ETNYNLLEKQVSSLIEDESN-----------------------LIAILSNVSALLNDSIDQ------------INWV
YP_002156075.1    ----------MELKHYARLTQQALGLIEGETD-----------------------YIANLANISALLNMELDD------------INWV
NP_244074.1       MFQPESYQG-TLEEKYSLVTKQLAALLEGESD-----------------------AIANLANASALLYHFLEE------------VNWV
YP_001559126.1    MDQIDSMYPENKVEMYELLLVQLKAMLEGEQD-----------------------VIANLSNVSALLNQALKD------------INWV
YP_080250.1       VEKTEG----SKEKKYELLLKQIDALTEGEPD-----------------------MIANTANASALLYQSLED------------VNWA
YP_253122.1       MPLIK-------ETNYNLLTKQLQSLIEDEQN-----------------------LIAILSNTSAILNDNLDQ------------INWV
YP_261333.1       MIDLQS--AGQGLDGYGLLCAQLESLLADERD-----------------------FIANAAQFSAFLYSQLED------------LNWA
```

Supplemental data

```
YP_001187407.1    MIDLSQ--TGAGMDGYPLLAAQLEALLAGERD-----------------------FIANAAQFSAFLFHELAD------------LNWA
YP_001422246.1    VEKQSG----DKEKSYQLLLKQLEAMTEDETD-----------------------AIANYANASALLYHSLHE------------VNWA
YP_001748360.1    MIDLNA--SGAGLDGYDLLAAQVQALFADERD-----------------------FIANAAQFSAFLYNQVDD------------LNWA
YP_155660.1       ----------MSSQLYSDLLAQSRALTEGEPD-----------------------LIANLANISALLYEHIED------------INWL
NP_744032.1       MIDLNA--SGAGLDGYNLLAAQVQALFADERD-----------------------FIANAAQFSAFLYNQVDD------------LNWA
YP_001269146.1    MIDLNA--SGAGLDGYNLLAAQVQALFADERD-----------------------FIANAAQFSAFLYNQVDD------------LNWA
YP_236680.1       MIDLQN--TGAGLDGYGLLAAQLEALLADERD-----------------------FIANAAQFSAFLYTQLDD------------LNWA
YP_001667695.1    MIDINA--SGAGLDGYNLLAAQVQALFADERD-----------------------FIANAAQFSAFLFNQVDD------------LNWA
YP_188854.1       MTKNI-------PTNYKLISKQLASLIEDEKN-----------------------LIAILSNTSALLNDTIDQ------------INWV
YP_002275225.1    ----MTTLPADHRVTLEDLLSTVESVLTG-------ERD----------------LIANMANISALLFEALPD------------INWA
YP_804779.1       MSE----------PINSIIVEQIDALLFNETN-----------------------LIANLSNASALLYSETKD------------VNWC
NP_470973.1       MIEIQKMTG-TKEENYALALKQVQAMIAGEPN-----------------------LIANLSNVSSILNQALSD------------INWV
YP_607257.1       MIDLNA--SGQGLDGYDLLAAQVQALFADERD-----------------------FIANAAQFSAFLYNQVDD------------LNWA
YP_001860169.1    MFEATINTALPKAEFYRELTSQARSLLEGESN-----------------------QIANAANLSALIFHSLPQ------------VNWA
NP_465120.1       MIEIKKMTG-TKEENYALALKQVQAMIHGEPN-----------------------LIANLSNVSSIINQALSD------------INWV
NP_372242.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLNQ------------INWV
NP_374828.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLNQ------------INWV
YP_001247140.1    MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLNQ------------INWV
YP_001316940.1    MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLNQ------------INWV
YP_001442294.1    MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLNQ------------INWV
YP_002314836.1    MFHTQTYTG-TREENYTLVIEQLKALISDEKS-----------------------VVANLANASALLNYFLHD------------INWV
YP_001814701.1    MFETTTYDG-DRTKQYEMLNKQLDALLTGETN-----------------------QIANLSNASALLGQFLDR------------TNWV
NP_646478.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_043764.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_186602.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_494359.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_500333.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_001332646.1    MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_001575592.1    MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
ZP_02759966.1     MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_001487827.1    VEKQAG----DPSKDYQLLVKQVEAITDGEPD-----------------------LIANLANAAALLYHSLPE------------VNWA
YP_417047.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSAILNDNLDQ------------VNWV
YP_041183.1       MTTIN-------PTNYTLLKKQAASLIEDEHH-----------------------MIAILSNMSALLNDNLDQ------------INWV
YP_176251.1       MFSKLDYPT-EQEAAYRLLEQQAVALLEDEPD-----------------------ALANYANLCALLGQFLDD------------INWV
YP_002349929.1    MIEIQKMTG-SKEENYGLALKQVQAMIHCEPN-----------------------LIANLSNVSSILNQALSD------------INWV
YP_001376537.1    MFTKESYVG-SREKQYETVIKQLDALLTGESN-----------------------TVANLANASALLNQFLER------------INWV
YP_784647.1       MFQASTLSTDFKPQFYQELAAQARALLDGEHD-----------------------RIANAANFAALIWQSVPD------------INWS
YP_557760.1       MFEVSTASHLPKAAHYEELVAQARSLLAGETD-----------------------WIANAANFSAFVFHSLSD------------LNWA
YP_275769.1       MIDLQN--TGAGLNGYGLLAAQLEALLADERD-----------------------SIANAAQFSAFLYTQIDD------------LNWA
YP_001118914.1    MFTLSTDPHASKAELYATLAEQARSLVESEPD-----------------------LIANAANFSALVYHSLDR------------LNWA
YP_001782545.1    MFALEDFKGLTEEQKYENMLLFLKGQLLGEKD-----------------------VIANLSNASAIIMALIDN------------LNWA
NP_882192.1       MFDATPISTVSKPVFYAELAAQARALLAGEHD-----------------------RIANAANFSALVYQALPD------------INWA
```

Supplemental data

```
YP_001894915.1    MFEVSSASHLPKAAHYEELVAQARSLLAGETD-----------------------WIANAANFSAFVFHSLSD------------LNWA
YP_001647268.1    MFTKESYAG-SREEQYKTVIKQLDALLTGEPN-----------------------VVANLSNASALLNQFLDR------------VNWV
NP_882465.1       MFDATPISTVSKPVFYAELAAQARALLAGEHD-----------------------RIANAANFSALVYQALPD------------INWA
YP_002448219.1    MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
NP_834359.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_002369448.1    MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_191278.1       ----MS--APIHRLSVDELVPMVRSVLQT-------ETD----------------MIANMANIAALLFETLPQ------------INWA
NP_886654.1       MFDATPISTVSKPVFYAELAAQARALLAGEHD-----------------------RIANAANFSALVYQALPD------------INWA
NP_847103.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------IVANLSNASALLNQFLDR------------VNWV
YP_030799.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------IVANLSNASALLNQFLDR------------VNWV
NP_981083.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_038702.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_002340711.1    MFTKESYTG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_085975.1       MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_002453717.1    MFTKESYAG-SRVQQYETVIKQLDALLTGESN-----------------------VVANLSNASALLNQFLDR------------VNWV
YP_772921.1       MFTLSNDPHASKAELYATLVEQARALVESERD-----------------------LIANAANFSALVYHSLDR------------LNWA
YP_001843726.1    KMMSK-------QADYQLLYQQALALFEGETD-----------------------WIANASNLAALLYNGLAD------------VNFA
YP_942902.1       ----------MTFKDYKLLSQQAVALIADETD-----------------------LTANLANLSALLNMNLDN------------INWV
YP_121651.1       SFTVAE-PTGDRVEDYRRLVAQAAALIEGEPD-----------------------RVANAANLSALIYHSLPE------------VNWA
YP_001633478.1    MFTATDLPTDDKPRFYAELAAQARALLDGETD-----------------------RIANAANFAALAYQALPR------------INWA
YP_001392189.1    MFALEDFKGLTEEQKYENMLLFLKGQLLGEKD-----------------------VIANLSNASAIIMALIDN------------LNWA
YP_795394.1       MSE----------SLNPIITKQLDALLYGETN-----------------------LVANLSNAAALLNQTLSD------------INWA
YP_002284153.1    MYATKNIEQGDKRSFYRELADQLQGLLQDERD-----------------------AVANAANLSALVFELVAD------------LNWA
YP_001255404.1    MFALEDFKELTEEQKYENMLLFLKGQLLGEKD-----------------------VIANLSNASAIIMALIDN------------LNWA
YP_001385170.1    MFALEDFKELTEEQKYENMLLFLKGQLLGEKD-----------------------VIANLSNASAIIMALIDN------------LNWA
YP_001388639.1    MFALEDFKELTEEQKYENMLLFLKGQLLGEKD-----------------------VIANLSNASAIIMALIDN------------LNWA
YP_001347703.1    MIDLQQ--TGAGLDGYALLAAQAEALFADERD-----------------------FIANAAQFCAFLFHELDD------------LNWA
NP_103622.1       MFAAKAIDTSDKAAFYSDLAAQLKALLDGESD-----------------------SIANAANTAALIFQMVPD------------LNWA
YP_001807739.1    MFTLSNDPHASKAELYATLVEQARALVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
YP_001580337.1    MFTLSHDPDASKADQYATLVEQARALVESERD-----------------------LIANAANFSALVYHSLDR------------LNWA
YP_002230176.1    MFTLSNDPTASKADQYATLVEQARALVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
NP_251512.1       MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----------------------FIANAAQFSAFLFHELGD------------LNWA
YP_459505.1       MFDFAPDETLPAEERYRQLVDAADALTAGEGD-----------------------GVANMANLAALMWEFIPD------------LNWA
YP_790351.1       MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----------------------FIANAAQFSAFLFHELGD------------LNWA
YP_002439846.1    MIDLQQ--SGAGLAGYALLAAQAEALFADERD-----------------------FIANAAQFSAFLFHELGD------------LNWA
YP_001970489.1    MFANAS-LTGSKPEQYAQLLEQARGLVYAESD-----------------------RIANAANLSALVYHALPD------------LNWV
YP_617023.1       MSYTLSFTATEAAERWAEAGEAAAALVAGEHD-----------------------GIANMANVAALIWQAIPD------------LNWA
YP_620554.1       MFTLSNDPTASKADQYATLVEQARALVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
YP_834795.1       MFTLSNDPTASKADQYATLVEQARALVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
YP_148653.1       MFRPMAYNG-TREENYALVIEQLKSLISGEPN-----------------------FIANLANAAALLNQFFTD------------INWV
YP_442277.1       MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_334263.1       MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
```

Supplemental data

```
YP_001067073.1   MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_368503.1      MFTLSNDPNASKADQYATLVEQARALVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
YP_109010.1      MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_993758.1      MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_001028775.1   MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_002026852.1   MFANAS-LTGSKPEQYAQLLEQARGLVYAESD-----------------------RIANAANLSALVYHALPD------------LNWV
YP_001764423.1   MFTLSNDPTASKADQYATLVEQARSLVESERD-----------------------LTANAANFSALVYHSLDR------------LNWA
NP_970249.1      SINYS-----DKTKFYKELHSEAEGLAEK--EWFVN-----------------------LANFSALLKQHLPEIN------------WV
YP_001059788.1   MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_001081316.1   MFALSEALPSSKPALYETLAAQARALVETETD-----------------------IVANAANFASLVYHSLDG------------LNWA
YP_001875000.1   ---------------MEIILKQIEMLLS-------PEID----------------AVANMANTAAVIFNSLPR------------LNWA
YP_278248.1      ------------------MLVQQYKSLIADEVK----------------------IYSILANTSAFIYQNFKN------------LNWA
YP_001309672.1   MFDIKIFEGMSTENKLDNMLIMLEGLIQGDEDS----------------------SITKLCNASALINSLIDR------------INWC
YP_001576818.1   MTAAK-------KEEYELLVEQAKALVAGESD-----------------------WIANTANLSALLFNSLNN------------VNFA
YP_193206.1      MAATT-------KENYELLVQQAQALIANESD-----------------------WIANTANLSALLFNSLSD------------VNFA
YP_199405.1      MFATSS-LTGSKPEQYAQLTAQAQALVHGEPD-----------------------RIANAANLAALIFHSLPS------------LNWA
YP_001915608.1   MFATSS-LTGSKPEQYAQLTAQAQALVHGEPD-----------------------RIANAANLAALIFHSLPS------------LNWA
YP_002509654.1   MEQ--------KEEKLREINEKLEYLIGNERD-----------------------WLANLANAAALLYNELDE------------VNWA
YP_757341.1      --------MSSKADAYQDVRKQIAAVIAGETS-----------------------RTARLATASCLLREMFGER-----------FFWT
YP_810624.1      YMETK-------QTRYSLLNDQAQSLLDGEHD-----------------------LIANMSNLVSLIFNSDQN------------INGT
YP_244708.1      MFATSS-LTGSKPEQYAQLTAQAQALVHGEPD-----------------------RIANAANLAALIFNSLPS------------LNWA
NP_635978.1      MFATSS-LTGSKPEQYAQLTAQAQALVHGEPD-----------------------RIANAANLAALIFNSLPS------------LNWT
XP_505859.1      YSRFSS---ESKSELYDTMASNVEALVDGQTNWVAN-----------------------LSNASSIIWHGLKSLPEPSSD-----INWA
NP_964546.1      MAATT-------EDNYQLLLKQAQALVEGEND-----------------------WIANTANISALLFNSLNN------------VNFA
YP_001261873.1   MYMFDIDRHGDKASLYADMAAALAALTADEPD-----------------------AIANMANAAALIWEYLPD------------LNWA
YP_814311.1      MSATT-------ENNYQLLVKQAEALVDGESD-----------------------WIANTANVSALLFNALND------------VNFA
NP_298520.1      MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----------------------RIANAANLAALIFHSLPG------------LNWV
YP_001775202.1   MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----------------------RIANAANLAALIFHSLPG------------LNWV
YP_001829224.1   MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----------------------RIANAANLAALIFHSLPG------------LNWV
NP_778732.1      MLAPFI-LTGNKPEQYAQLEAQAKALIHGETD-----------------------RIANAANLAALIFHSLPG------------LNWV
NP_246070.1      ------------MMNYALLLKQLEQILEPETY-----------------------IISRMANTSAFLYQHMPD------------LNWV
XP_645637.1      AEDLEISKG-SKEEQYENLLPQIEGLLTGENN-----------------------QIANLANVTAALKEQFN-------------FFWV
XP_456263.1      YSSFQT---TDRKKALEQLLISYEALAEGQDNWVCN-----------------------LANAASLIWHCYISLN---VD-----VNWA
YP_001300883.1   ETSFKQTTS-NKVERYRLFLPQFELLISDEKE-----------------------EISVLANTAAALREAFG-------------FFWV
YP_812434.1      MAAAS-------QAEYELLLEQAKALTAGESD-----------------------WIANAGNVSALLKAGMRN------------VNFA
NP_814796.1      QTREN------KKAAYELCLQQLSGLLEIETD-----------------------AIANLANSSALLAQTLPD------------TVFA
NP_358362.1      MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
YP_816246.1      MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
YP_001694316.1   MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
YP_001835526.1   MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
YP_002037500.1   MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
NP_345351.1      MLKSE------KQSRYQMLNEELSFLLEGETN-----------------------VLANLSNASALIKSRFPN------------TVFA
```

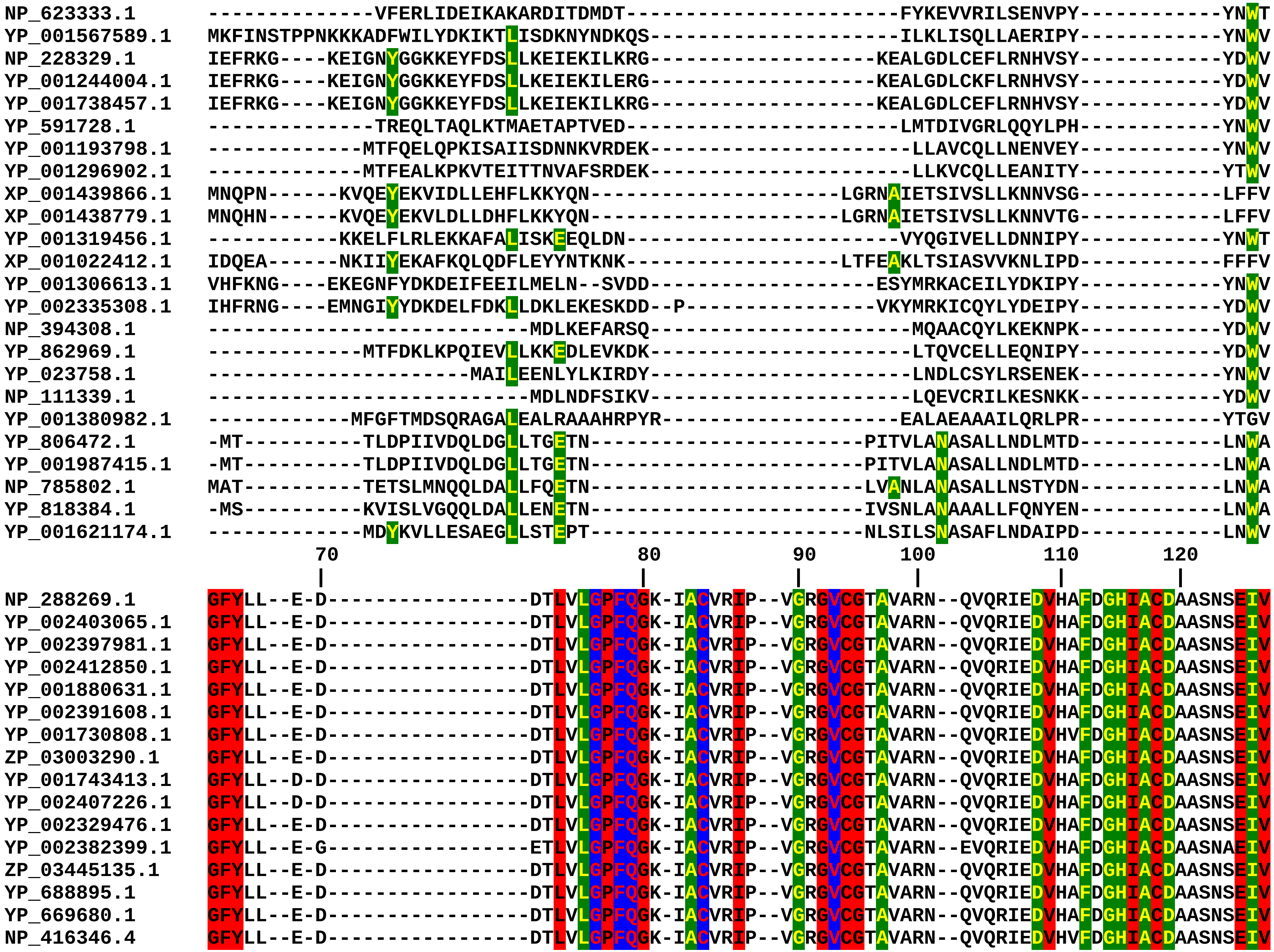
Supplemental data

```
XP_001385863.2   YSNVTG---LSKEETLQHIIDSYEALASDTTNWVAN-----------------------LSNCSSLVWHAYHSLN---VP-----VNWA
NP_985662.1      YSNFSA---SDRSEALIYMIESYQALAEHQDNMVCN-----------------------LANASSLLWHVYRNLN---VK-----VNWA
YP_759928.1      ---------MDKAEAYRQLKEEIDSVVAGETS-----------------------DTARYASAACLLAQAFKPR-----------FFWT
YP_002123053.1   MNKTT------KIEHYQLLLAQAKALFSNESN-----------------------ALANLSNASALLKMTLPH------------SVFA
YP_002131343.1   EAFNEIVLPASKAERYAAVAEEIASVLDGEPN-----------------------LVARMATVASMLANSFEH------------YFWT
XP_001683506.1   ---TPAPTFSSKKELYDWLVKQVEALVHDMSKRFTPQAN----------------LVIGLSNTAALCFYELNRFCNPEAALEKLRVNWF
NP_012854.1      HVNYSSN--LNKEEILEQLLLSYEGLSDGQVNWVCN-----------------------LSNASSLIWHAYKSLA---VD-----INWA
XP_460154.1      YTSYSN--DGGKEEILQSVVDSYEALSYDTRNWVAN-----------------------LANCSSLLWHAYHSMK---IN-----VNWS
XP_001801028.1   ASAFA--EGLSKKEVYAQVLEQARILFDGQRNWVCN-----------------------FSNTSSLLWHALHSLPSPSHS-----VNWA
YP_001201197.1   MTKQE------KISNYQLLLAQLEALLDGETN-----------------------ALANLSNASALLNQALPN------------SVFT
XP_001484046.1   YSNYKG--DETKEETLQSVVDSYEALSSDTQNWVAN-----------------------LANCSSLIWHAYHSLN---IP-----VNWA
NP_274882.1      MHALHFSAS-DKAALYREVLPQIESVVADEAD-----------------------WVANLANTAAVLKEAFG-------------WFWV
YP_001302571.1   AEELRIKNG-DKQEMYETLLPQIASLVGNETD-----------------------LIANMANVAAALKQTFG-------------FFWV
YP_002000650.1   MHALHFSAS-DKAALYREVLPQIESVVADEAD-----------------------WVANLANTAAVLKEAFG-------------WLWV
YP_001682978.1   EAFNDKTLTGDKAARYAELAQEIASVLDGETN-----------------------LTARMATVASMLASSFDH------------YFWT
YP_002342048.1   MHALHFSAS-DKAALYREVLPQIESVVADETD-----------------------WVANLANTAAVLKEAFG-------------WFWV
YP_207199.1      MHALHFSAS-DKAALYREVLPQIESVVADEAD-----------------------WVANLANTAAVLKEAFG-------------WLWV
XP_844799.1      GASTPTGGRSDKVTFYSRLVKEVQGIVEFPSKRLAPAAN----------------AIINLSNVAASLFYGLNRHANPEAVLQSPPVNWL
YP_001598496.1   MHALHFSAS-DKAALYREVLPQIESVVADETD-----------------------WVANLANTAAVLKEAFG-------------WFWV
YP_974447.1      MHALHFSAS-DKAALYREVLPQIESVVADETD-----------------------WVANLANTAAVLKEAFG-------------WFWV
XP_001465858.1   ---TPAPTFSSKKELYDWLVKQVEALVHDMSKRFTPQAN----------------LVIGLSNVAALCFYELNRFCNPEAALEKLKVNWF
YP_001280815.1   QAINPIDPNANKAEKYQLLLKQAEAIISSESD-----------------------LIANMANLTALLKDTFN-------------WFWI
NP_664843.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
NP_802083.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
YP_280574.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
YP_001128311.1   MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
YP_002286076.1   MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
XP_390238.1      ASNFA--EGVTKEDAYEQVLWQAEGLVTDQRNWVCN-----------------------LSNAASLLWHAYKSLGSPSKD-----VNWA
XP_001565241.1   ---LPAPIFTSKKELYDWLVKQVEALVDGMSKRFTPQAN----------------LIIGLSNVAALCFYELNRFCNPGAALEKLKVNWF
YP_596891.1      MNKSK------KIEQYQLMIEQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
YP_600772.1      MNKSK------KIEQYQLMIEQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFA
YP_001351938.1   AEVLSFNKLGGKADIYAELLPQIEALLAAEPD-----------------------LIANQANIAAVLHQAFG-------------FHWV
XP_001910908.1   ----K--IELTK---------------------GSN-----------------------LSNTSSLLWHLYRSLPSPSSS-----VNWA
XP_446236.1      YTNFNA---TDRSEALQLLIDSYEALAMDQRNWVCN-----------------------LANAASLLWHAYKSLN---VN-----VNWT
NP_607473.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
YP_060405.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
YP_602707.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
NP_269471.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
YP_282476.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
YP_598792.1      MNKSK------KIEQYQLMIAQAKELFANESN-----------------------ALANLSNASALLNMTLPN------------SVFT
YP_099090.1      AEELTFISG-SKEEQYLSLLPQVRSLIEGEVD-----------------------LVANLANVAAALKEAFD-------------FFWV
NP_421082.1      EAFNDITLSADKATRYAEVADEIASVLDGEPN-----------------------LTARMATVASMLANSFDS------------YFWA
```

Supplemental data

```
YP_001034962.1   MNKTE------KVSNYQILLAQLEALLEGETN-----------------------ALANLSNASALLNQALPH------------SVFT
YP_211509.1      AEELTFISG-SKEEQYLSLLPQVRSLIEGEVD-----------------------LVANLANVAAALKEAFD-------------FFWV
XP_001940256.1   ---------MSKKEVYAQVLEQARSL---------N-----------------------FSNTASLLWHAYHSLPSPSSS-----VNWA
YP_001450196.1   MNTKD------KISNYQILLAQLEALLEGETN-----------------------ALANLSNASALLNQALPH------------SVFT
XP_680937.1      SSYFG---SGSKAEIYAQVIQQAKGLVTGQRNWVRI----------------LISIVSNFSNVASLLWHAYAALPAPSSS-----VNWA
XP_001265696.1   SSYFSTSQEPTKAEVYAQVLEQAKGLVTGQRNWATQ--------------------PSNFSNIASLLWHAYAALPPPSSS-----VNWA
XP_367405.1      ASNFG--ANVTKEEAYKQVLEQAEGLFDGQRNWVCN-----------------------LANTASLLWHAYHSLPSPSSA-----VNWA
YP_001033719.1   MNKQE------KIEGYELLNLQLKALLSEQNY-----------------------TLSNLANASSLLWSFLPE------------QVYT
YP_812051.1      MNKQE------KIEGYELLNLQLKALLSEQNY-----------------------TLSNLANASSLLWSFLPE------------QVYT
NP_810723.1      AENLLIHTG-NKEEKYRELLPQLQALVSSETN-----------------------RIANLANIAAALKQTFH-------------FFWV
YP_141537.1      MTNEE------KKIAYDLMLAQANVLFANEDN-----------------------ALANFANASALLNTTLPN------------SVFT
NP_721920.1      MKNQE------KISAYQILLAQAQSMFANETN-----------------------ALANLANATAMLNKTLPN------------AIFT
YP_820528.1      MTNEE------KKIAYELMLAQAKVLFANEDN-----------------------ALANFANASALLNTTLPN------------SVFT
XP_001246405.1   -------------------------MTCG--IWKAY-------------------FPSNLASVSSLLWHAYASLPAPSNA-----VNWA
YP_612063.1      -----------MRVNYEDLSKVVAALTEGETD-----------------------EVALMATVVCELHHADDR------------FDWT
YP_444835.1      -PASSTATASDKAERYADTRRRIDALLDGPTD-----------------------WVSAMATVACELHHSFDH------------YDWT
NP_943356.1      -----------------------------------------------------------------MCWRRIS-----------------
YP_631535.1      AEVTLDLRGMPKAEAYAELKQHVRAVLEGIDD-----------------------DITGMATMSCLLHHAFG-------------HLWT
NP_687842.1      MNKSK------KIENYQLLLLQAQALFSDETN-----------------------ALANLSNASAMLNAMLPN------------SVFT
YP_329573.1      MNKSK------KIENYQLLLLQAQALFSDETN-----------------------ALANLSNASAMLNAMLPN------------SVFT
XP_001692446.1   PPLPVTASNEEREARYELVLESVKSVLEGEDD-----------------------WVAVMATVACLLHEAFEY------------YHWT
NP_390842.2      VEKQSG----DKEKDYQLLLKQLEAMTEDETD-----------------------QIANYANASALLYHSLPE------------VNWA
YP_511991.1      APHSCPQRMDLPMADYATLTKTIASLTEGETD-----------------------TITLMATVACELHHADPR------------FDWT
XP_719233.1      YSNLKSG-DFSKEETLQHVLDSYKALETD--NWVAN-----------------------LSNCSSLLWHAYKSLN---IN-----VNWT
NP_735295.1      MNKSK------KIENYQLLLLQAQALFSDETN-----------------------ALANLSNASAMLNAMLPN------------SVFT
YP_168601.1      -----------MTVDYPSLAKTIAALTEGETD-----------------------PVALMATVTCEVHHSDDR------------FNWT
YP_001169098.1   -------------MDYDRLDQTIDALCHGERD-----------------------TVALMATVACELHHAHPL------------SDWT
XP_001216799.1   SSYFG--GSGSKAEIYTQVLEQAKGLVTGQRNWVTHTNA--------PRHPSELTINSNFSNVASLLWHAYAALPSPSSS-----VNWA
YP_437091.1      DSLPLLSSEEEKSAHYQHTLRAIRHAIEGETD-----------------------WTAILSTVVCELHHRFEY------------FHWT
XP_001390389.1   SSYFGA-GNGSKSEVYTQVLEQARGLVYDQRNWLTN-------------HFSHRQQVSNFSNVASLLWHAYAALPAPSSS-----VNWA
XP_001550500.1   SSTFK--ADISKAEAN--LSNAASLLWHAYKSLPS--------------------------PSSEVNWAGFYTLDPTSTP-----SKPS
XP_001273080.1   SSYFASSGQASKADIYEQLLEQAKGLVTGQRNWVHLPSHPTLQPPPHKPSHLTHIQPSNLSNIASLLWHAYAALPPPSSS-----VNWA
YP_001042044.1   -------------MDYDQLDQTIEALCHGEQD-----------------------TVALMATVACEIHHAHPL------------SDWT
YP_916587.1      ------------MPDYDDLAARIQALTHDETD-----------------------EVALMATLACEIHHSDER------------FDWT
XP_961529.1      ASNFA--ANVTKEEAYQQVLDQAEGLFYEQRNWVCN-----------------------LSNCAALLWHAFKSLPAPSNQ-----VNWA
YP_351544.1      -------------MDYDQLDQTIEALCHGEQD-----------------------TVALMATVACEIHHAHPL------------SDWT
XP_001588803.1   ASTFK--ADISKEEAYAQVLEEARGLFDGQRNWIA--------------------------IEIEIS---FYTLDPSST-------KPS
XP_762195.1      AAALPS-YVVTKADFYDHLESSLISLIDPGTDWISA-----------------------LSNASSLVFNSMNRFPAWTSKR----INWA
XP_001537789.1   -------------------------------------------------------MVSNLANVSSLLWHAYAALPAPSSA-----VNWA
YP_001198994.1   MTKQE------KISNYQLLLAQLEALLDGETN-----------------------ALANLSNASALLNPALPN------------SVFT
NP_390842.1      VEKQSG----DKEKDYQLLLKQLEAMTEDETD-----------------------QIANYANASALLYHSLPE------------VNWA
YP_001469780.1   KSFRRG----ETMGSYTSKKQLFEELTNQIEEIIFQNNI-----------------HESMEEICNLLYENISY------------YNWV
```

Supplemental data

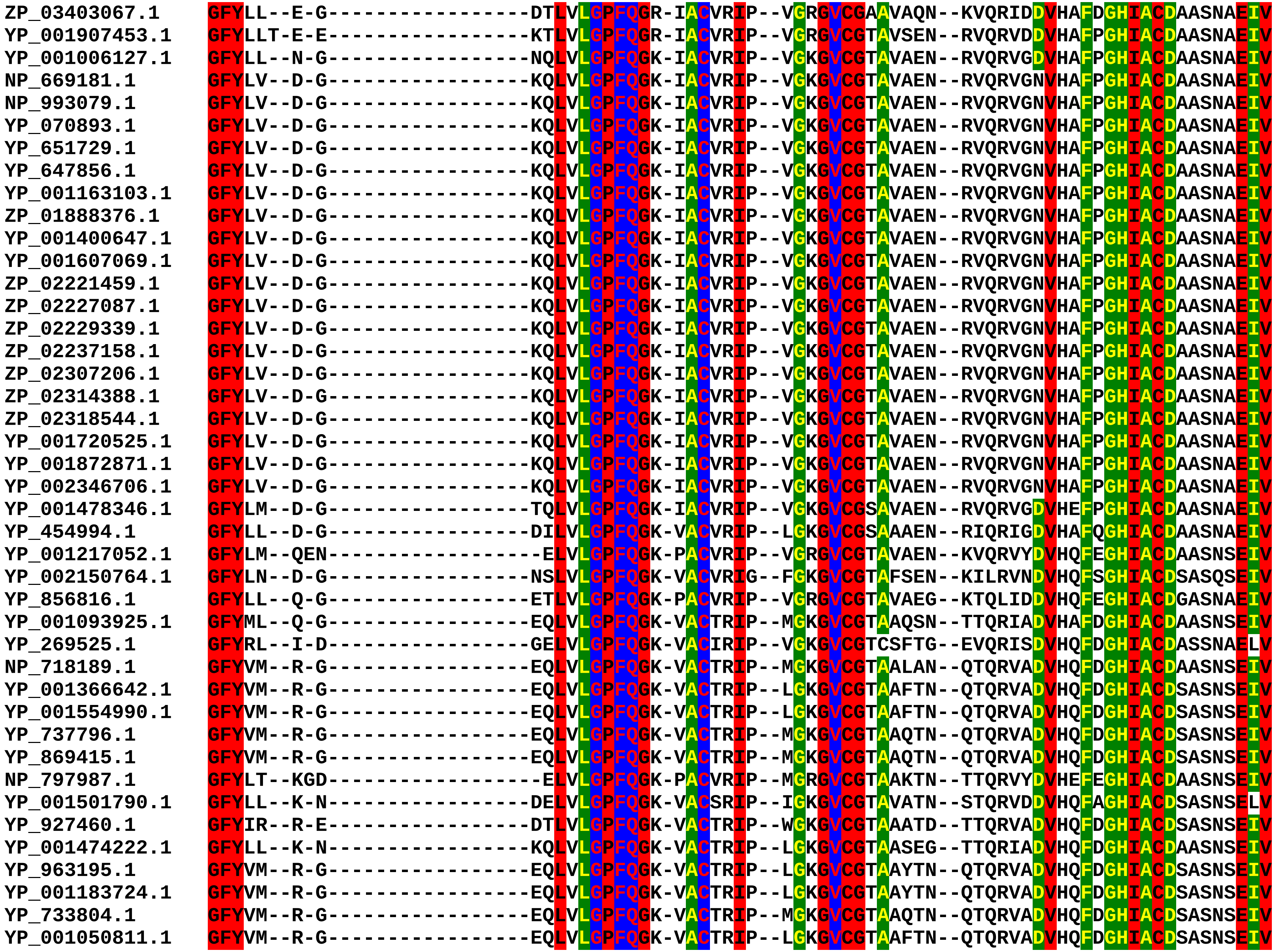


```
NP_623333.1      --------------VFERLIDEIKAKARDITDMDT-----------------------FYKEVVRILSENVPY------------YNWT
YP_001567589.1   MKFINSTPPNKKKADFWILYDKIKTLISDKNYNDKQS---------------------ILKLISQLLAERIPY------------YNWV
NP_228329.1      IEFRKG----KEIGNYGGKKEYFDSLLKEIEKILKRG-------------------KEALGDLCEFLRNHVSY------------YDWV
YP_001244004.1   IEFRKG----KEIGNYGGKKEYFDSLLKEIEKILERG-------------------KEALGDLCKFLRNHVSY------------YDWV
YP_001738457.1   IEFRKG----KEIGNYGGKKEYFDSLLKEIEKILKRG-------------------KEALGDLCEFLRNHVSY------------YDWV
YP_591728.1      --------------TREQLTAQLKTMAETAPTVED-----------------------LMTDIVGRLQQYLPH------------YNWV
YP_001193798.1   -------------MTFQELQPKISAIISDNNKVRDEK----------------------LLAVCQLLNENVEY------------YNWV
YP_001296902.1   -------------MTFEALKPKVTEITTNVAFSRDEK----------------------LLKVCQLLEANITY------------YTWV
XP_001439866.1   MNQPN------KVQEYEKVIDLLEHFLKKYQN---------------------LGRNAIETSIVSLLKNNVSG------------LFFV
XP_001438779.1   MNQHN------KVQEYEKVLDLLDHFLKKYQN---------------------LGRNAIETSIVSLLKNNVTG------------LFFV
YP_001319456.1   -----------KKELFLRLEKKAFALISKEEQLDN-----------------------VYQGIVELLDNNIPY------------YNWT
XP_001022412.1   IDQEA------NKIIYEKAFKQLQDFLEYYNTKNK------------------LTFEAKLTSIASVVKNLIPD------------FFFV
YP_001306613.1   VHFKNG----EKEGNFYDKDEIFEEILMELN--SVDD-------------------ESYMRKACEILYDKIPY------------YNWV
YP_002335308.1   IHFRNG----EMNGIYYDKDELFDKLLDKLEKESKDD--P----------------VKYMRKICQYLYDEIPY------------YDWV
NP_394308.1      ---------------------------MDLKEFARSQ----------------------MQAACQYLKEKNPK------------YDWV
YP_862969.1      -------------MTFDKLKPQIEVLLKKEDLEVKDK----------------------LTQVCELLEQNIPY------------YDWV
YP_023758.1      ----------------------MAILEENLYLKIRDY----------------------LNDLCSYLRSENEK------------YNWV
NP_111339.1      ---------------------------MDLNDFSIKV----------------------LQEVCRILKESNKK------------YDWV
YP_001380982.1   ------------MFGFTMDSQRAGALEALRAAAHRPYR--------------------EALAEAAAILQRLPR------------YTGV
YP_806472.1      -MT----------TLDPIIVDQLDGLLTGETN-----------------------PITVLANASALLNDLMTD------------LNWA
YP_001987415.1   -MT----------TLDPIIVDQLDGLLTGETN-----------------------PITVLANASALLNDLMTD------------LNWA
NP_785802.1      MAT----------TETSLMNQQLDALLFQETN-----------------------LVANLANASALLNSTYDN------------LNWA
YP_818384.1      -MS----------KVISLVGQQLDALLENETN-----------------------IVSNLANAAALLFQNYEN------------LNWA
YP_001621174.1   -------------MDYKVLLESAEGLLSTEPT-----------------------NLSILSNASAFLNDAIPD------------LNWV
                          70                         80           90       100         110       120
                          |                          |            |         |           |         |
NP_288269.1      GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002403065.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002397981.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002412850.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_001880631.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002391608.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_001730808.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHVFDGHIACDAASNSEIV
ZP_03003290.1    GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_001743413.1   GFYLL--D-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002407226.1   GFYLL--D-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002329476.1   GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_002382399.1   GFYLL--E-G-----------------ETLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--EVQRIEDVHAFDGHIACDAASNAEIV
ZP_03445135.1    GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_688895.1      GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
YP_669680.1      GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
NP_416346.4      GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHVFDGHIACDAASNSEIV
```

Supplemental data

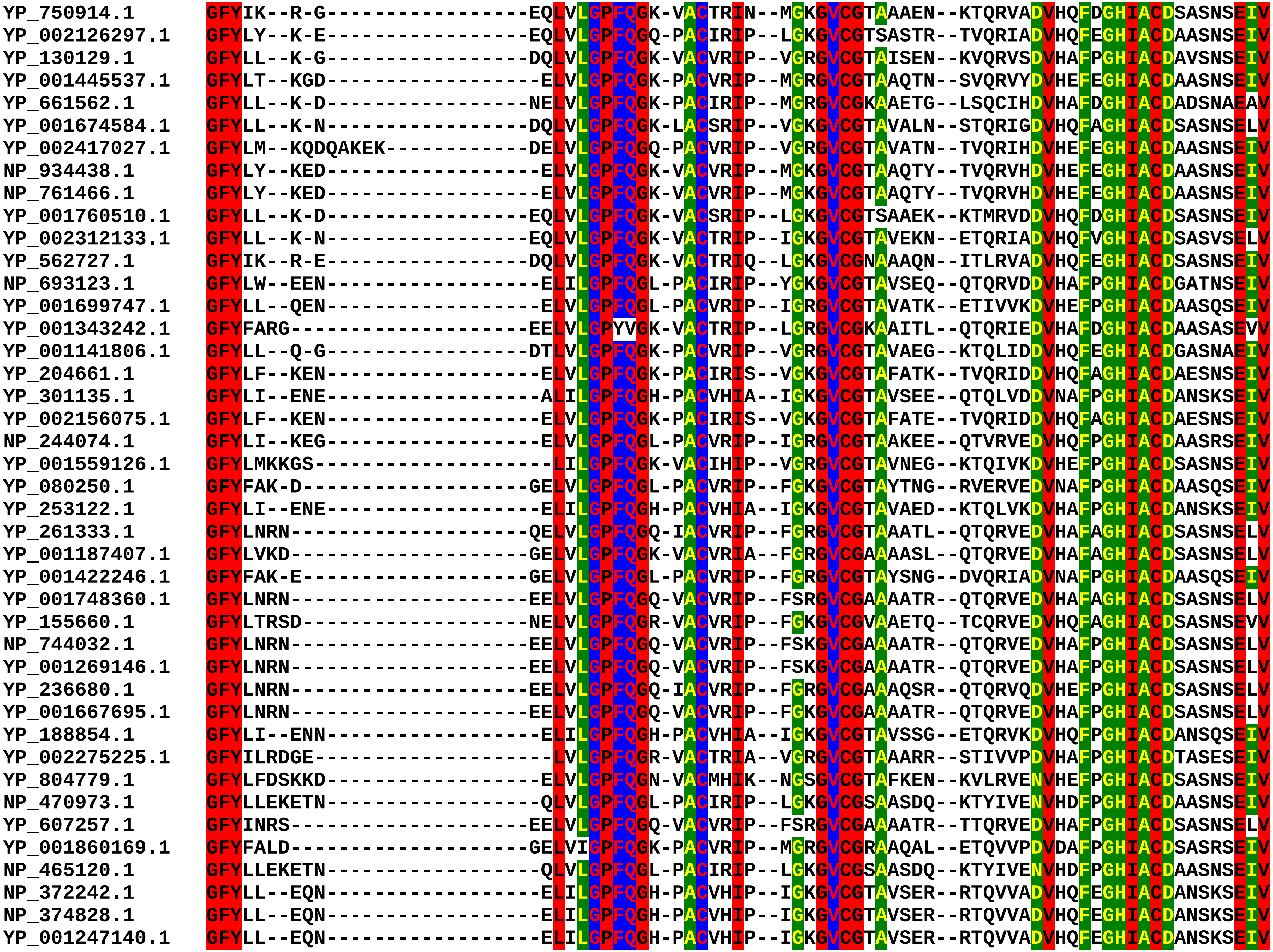
```
ZP_02903917.1    GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--EVQRIEDVHAFDGHIACDAASNSEIV
ZP_03282954.1    GFYLL--E-D-----------------NTLVLGPFQGK-IACVRIP--VGRGVCGTAVATN--QVQRVEDVHAFDGHIACDAASNSEIV
YP_001452718.1   GFYLL--E-N-----------------DTLVLGPFQGK-LACVRIP--VGRGVCGTAVAQN--QVQRIEDVHAFDGHIACDSASNAEIV
YP_001570142.1   GFYLL--E-G-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
NP_460803.1      GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_150310.1      GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002141797.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002243298.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
NP_456352.1      GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
NP_804849.1      GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_001587603.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02666379.2    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02832604.2    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_001177122.1   GFYLL--E-G-----------------NTLVLGPFQGK-LACVRIP--VGRGVCGTAVAQN--QVQRVEDVHAFDGHIACDSASNAEIV
YP_002226291.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASSAEIV
NP_754135.1      GFYLL--E-D-----------------DTLVLGPFQGK-IACVRIP--VGRGVCGTAVARN--QVQRIEDVHAFDGHIACDAASNSEIV
ZP_02347214.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02574893.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02659310.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02663512.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02681881.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_02696584.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002041101.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002045892.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03075263.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002114879.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002146178.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03165709.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002215294.1   GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03221812.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03342093.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03347154.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03352578.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03357060.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03371453.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03379701.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03383790.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03412299.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
ZP_03217260.1    GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_002237781.1   GFYLL--E-G-----------------DTLVLGPFQGK-LACVRIP--VGRGVCGAAVAQA--QVQRVEDVHAFDGHIACDAASNSEIV
YP_001336004.1   GFYLL--E-G-----------------DTLVLGPFQGK-LACVRIP--VGRGVCGAAVAQA--QVQRVEDVHAFDGHIACDAASNSEIV
YP_001437512.1   GFYLL--E-D-----------------KTLVLGPFQGK-IACVRIP--YGRGVCGTAVATG--EVQRVDDVHAFDGHIACDASSNAEIV
```

Supplemental data



```
ZP_03403067.1     GFYLL--E-G-----------------DTLVLGPFQGR-IACVRIP--VGRGVCGAAVAQN--KVQRIDDVHAFDGHIACDAASNAEIV
YP_001907453.1    GFYLLT-E-E-----------------KTLVLGPFQGR-IACVRIP--VGRGVCGTAVSEN--RVQRVDDVHAFPGHIACDAASNAEIV
YP_001006127.1    GFYLL--N-G-----------------NQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGDVHAFPGHIACDAASNAEIV
NP_669181.1       GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
NP_993079.1       GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_070893.1       GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_651729.1       GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_647856.1       GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001163103.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_01888376.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001400647.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001607069.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02221459.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02227087.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02229339.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02237158.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02307206.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02314388.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
ZP_02318544.1     GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001720525.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001872871.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_002346706.1    GFYLV--D-G-----------------KQLVLGPFQGK-IACVRIP--VGKGVCGTAVAEN--RVQRVGNVHAFPGHIACDAASNAEIV
YP_001478346.1    GFYLM--D-G-----------------TQLVLGPFQGK-IACVRIP--VGKGVCGSAVAEN--RVQRVGDVHEFPGHIACDAASNAEIV
YP_454994.1       GFYLL--D-G-----------------DILVLGPFQGK-VACVRIP--LGKGVCGSAAAEN--RIQRIGDVHAFQGHIACDAASNAEIV
YP_001217052.1    GFYLM--QEN------------------ELVLGPFQGK-PACVRIP--VGRGVCGTAVAEN--KVQRVYDVHQFEGHIACDAASNSEIV
YP_002150764.1    GFYLN--D-G-----------------NSLVLGPFQGK-VACVRIG--FGKGVCGTAFSEN--KILRVNDVHQFSGHIACDSASQSEIV
YP_856816.1       GFYLL--Q-G-----------------ETLVLGPFQGK-PACVRIP--VGRGVCGTAVAEG--KTQLIDDVHQFEGHIACDGASNAEIV
YP_001093925.1    GFYML--Q-G-----------------EQLVLGPFQGK-VACTRIP--MGKGVCGTAAQSN--TTQRIADVHAFDGHIACDAASNSEIV
YP_269525.1       GFYRL--I-D-----------------GELVLGPFQGK-VACIRIP--VGKGVCGTCSFTG--EVQRISDVHQFDGHIACDASSNAELV
NP_718189.1       GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--MGKGVCGTAALAN--QTQRVADVHQFDGHIACDAASNSEIV
YP_001366642.1    GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--LGKGVCGTAAFTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_001554990.1    GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--LGKGVCGTAAFTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_737796.1       GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--MGKGVCGTAAQTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_869415.1       GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--MGKGVCGTAAQTN--QTQRVADVHQFDGHIACDSASNSEIV
NP_797987.1       GFYLT--KGD------------------ELVLGPFQGK-PACVRIP--MGRGVCGTAAKTN--TTQRVYDVHEFEGHIACDAASNSEIV
YP_001501790.1    GFYLL--K-N-----------------DELVLGPFQGK-VACSRIP--IGKGVCGTAVATN--STQRVDDVHQFAGHIACDSASNSELV
YP_927460.1       GFYIR--R-E-----------------DTLVLGPFQGK-VACTRIP--WGKGVCGTAAATD--TTQRVADVHQFDGHIACDSASNSEIV
YP_001474222.1    GFYLL--K-N-----------------KQLVLGPFQGK-VACTRIP--LGKGVCGTAASEG--TTQRIADVHQFDGHIACDAASNSEIV
YP_963195.1       GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--LGKGVCGTAAYTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_001183724.1    GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--LGKGVCGTAAYTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_733804.1       GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--MGKGVCGTAAQTN--QTQRVADVHQFDGHIACDSASNSEIV
YP_001050811.1    GFYVM--R-G-----------------EQLVLGPFQGK-VACTRIP--LGKGVCGTAAFTN--QTQRVADVHQFDGHIACDSASNSEIV
```

Supplemental data

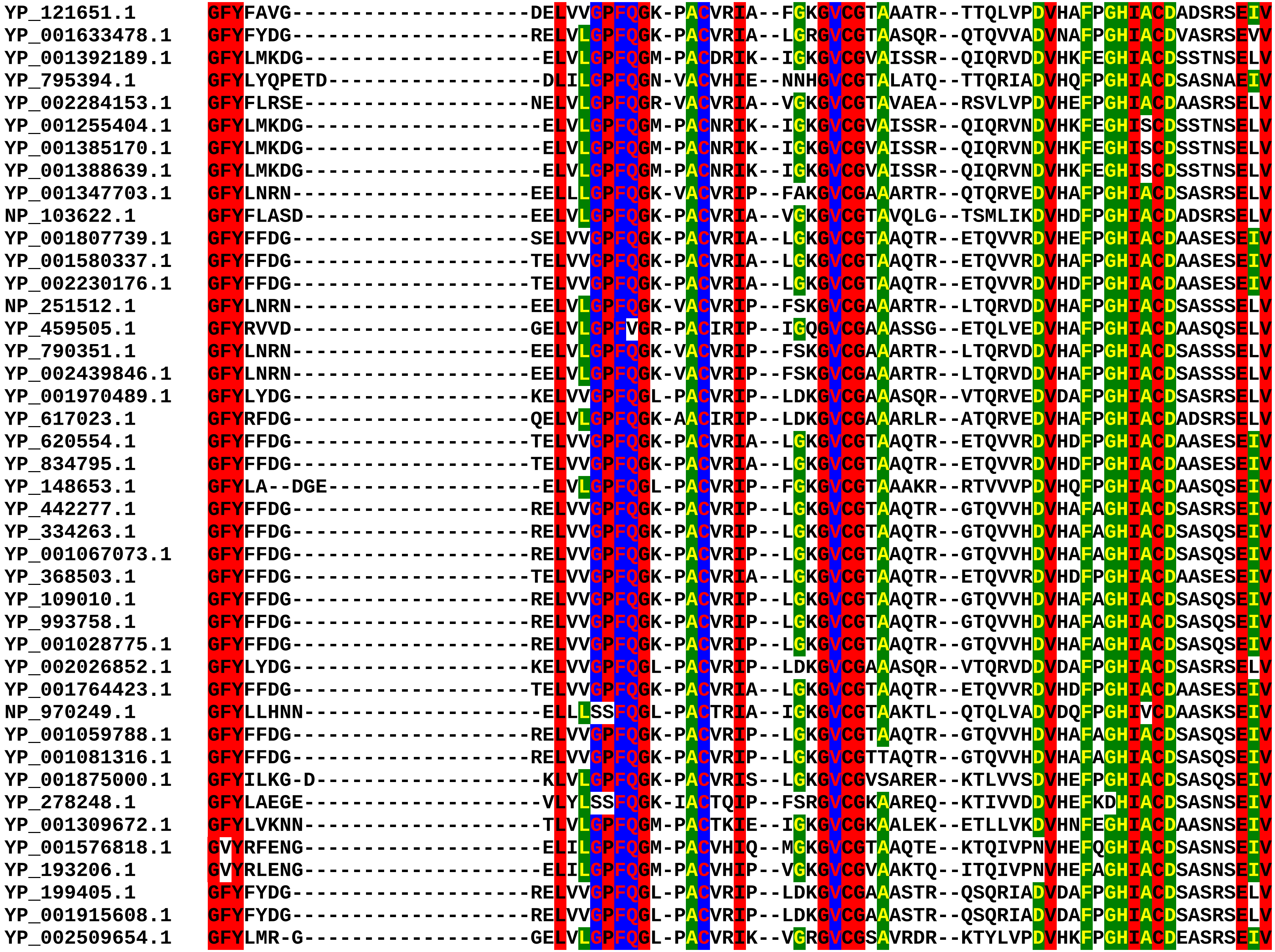


```
YP_750914.1      GFYIK--R-G-----------------EQLVLGPFQGK-VACTRIN--MGKGVCGTAAAEN--KTQRVADVHQFDGHIACDSASNSEIV
YP_002126297.1   GFYLY--K-E-----------------EQLVLGPFQGQ-PACIRIP--LGKGVCGTSASTR--TVQRIADVHQFEGHIACDAASNSEIV
YP_130129.1      GFYLL--K-G-----------------DQLVLGPFQGK-VACVRIP--VGRGVCGTAISEN--KVQRVSDVHAFPGHIACDAVSNSEIV
YP_001445537.1   GFYLT--KGD------------------ELVLGPFQGK-PACVRIP--MGRGVCGTAAQTN--SVQRVYDVHEFEGHIACDAASNSEIV
YP_661562.1      GFYLL--K-D-----------------NELVLGPFQGK-PACIRIP--MGRGVCGKAAETG--LSQCIHDVHAFDGHIACDADSNAEAV
YP_001674584.1   GFYLL--K-N-----------------DQLVLGPFQGK-LACSRIP--VGKGVCGTAVALN--STQRIGDVHQFAGHIACDSASNSELV
YP_002417027.1   GFYLM--KQDQAKEK------------DELVLGPFQGQ-PACVRIP--VGRGVCGTAVATN--TVQRIHDVHEFEGHIACDAASNSEIV
NP_934438.1      GFYLY--KED------------------ELVLGPFQGK-VACVRIP--MGKGVCGTAAQTY--TVQRVHDVHEFEGHIACDAASNSEIV
NP_761466.1      GFYLY--KED------------------ELVLGPFQGK-VACVRIP--MGKGVCGTAAQTY--TVQRVHDVHEFEGHIACDAASNSEIV
YP_001760510.1   GFYLL--K-D-----------------EQLVLGPFQGK-VACSRIP--LGKGVCGTSAAEK--KTMRVDDVHQFDGHIACDSASNSEIV
YP_002312133.1   GFYLL--K-N-----------------EQLVLGPFQGK-VACTRIP--IGKGVCGTAVEKN--ETQRIADVHQFVGHIACDSASVSELV
YP_562727.1      GFYIK--R-E-----------------DQLVLGPFQGK-VACTRIQ--LGKGVCGNAAAQN--ITLRVADVHQFEGHIACDSASNSEIV
NP_693123.1      GFYLW--EEN------------------ELILGPFQGL-PACIRIP--YGKGVCGTAVSEQ--QTQRVDDVHAFPGHIACDGATNSEIV
YP_001699747.1   GFYLL--QEN------------------ELVLGPFQGL-PACVRIP--IGRGVCGTAVATK--ETIVVKDVHEFPGHIACDAASQSEIV
YP_001343242.1   GFYFARG--------------------EELVLGPYVGK-VACTRIP--LGRGVCGKAAITL--QTQRIEDVHAFDGHIACDAASASEVV
YP_001141806.1   GFYLL--Q-G-----------------DTLVLGPFQGK-PACVRIP--VGRGVCGTAVAEG--KTQLIDDVHQFEGHIACDGASNAEIV
YP_204661.1      GFYLF--KEN------------------ELVLGPFQGK-PACIRIS--VGKGVCGTAFATK--TVQRIDDVHQFAGHIACDAESNSEIV
YP_301135.1      GFYLI--ENE------------------ALILGPFQGH-PACVHIA--IGKGVCGTAVSEE--QTQLVDDVNAFPGHIACDANSKSEIV
YP_002156075.1   GFYLF--KEN------------------ELVLGPFQGK-PACIRIS--VGKGVCGTAFATE--TVQRIDDVHQFAGHIACDAESNSEIV
NP_244074.1      GFYLI--KEG------------------ELVLGPFQGL-PACVRIP--IGRGVCGTAAKEE--QTVRVEDVHQFPGHIACDAASRSEIV
YP_001559126.1   GFYLMKKGS--------------------LILGPFQGK-VACIHIP--VGRGVCGTAVNEG--KTQIVKDVHEFPGHIACDSASNSEIV
YP_080250.1      GFYFAK-D-------------------GELVLGPFQGL-PACVRIP--FGKGVCGTAYTNG--RVERVEDVNAFPGHIACDAASQSEIV
YP_253122.1      GFYLI--ENE------------------ELILGPFQGH-PACVHIA--IGKGVCGTAVAED--KTQLVKDVHAFPGHIACDANSKSEIV
YP_261333.1      GFYLNRN--------------------QELVLGPFQGQ-IACVRIP--FGRGVCGTAAATL--QTQRVEDVHAFAGHIACDSASNSELV
YP_001187407.1   GFYLVKD--------------------GELVLGPFQGK-VACVRIA--FGRGVCGAAAASL--QTQRVEDVHAFAGHIACDSASNSELV
YP_001422246.1   GFYFAK-E-------------------GELVLGPFQGL-PACVRIP--FGRGVCGTAYSNG--DVQRIADVNAFPGHIACDAASQSEIV
YP_001748360.1   GFYLNRN--------------------EELVLGPFQGQ-VACVRIP--FSRGVCGAAAATR--QTQRVEDVHAFAGHIACDSASNSELV
YP_155660.1      GFYLTRSD-------------------NELVLGPFQGR-VACVRIP--FGKGVCGVAAETQ--TCQRVEDVHQFAGHIACDSASNSEVV
NP_744032.1      GFYLNRN--------------------EELVLGPFQGQ-VACVRIP--FSKGVCGAAAATR--QTQRVEDVHAFPGHIACDSASNSELV
YP_001269146.1   GFYLNRN--------------------EELVLGPFQGQ-VACVRIP--FSKGVCGAAAATR--QTQRVEDVHAFPGHIACDSASNSELV
YP_236680.1      GFYLNRN--------------------EELVLGPFQGQ-IACVRIP--FGRGVCGAAAQSR--QTQRVQDVHEFPGHIACDSASNSELV
YP_001667695.1   GFYLNRN--------------------EELVLGPFQGQ-VACVRIP--FGKGVCGAAAATR--QTQRVEDVHAFPGHIACDSASNSELV
YP_188854.1      GFYLI--ENN------------------ELILGPFQGH-PACVHIA--IGKGVCGTAVSSG--ETQRVKDVHQFPGHIACDANSQSEIV
YP_002275225.1   GFYILRDGE--------------------LVLGPFQGR-VACTRIA--VGRGVCGTAAARR--STIVVPDVHAFPGHIACDTASESEIV
YP_804779.1      GFYLFDSKKD------------------ELVLGPFQGN-VACMHIK--NGSGVCGTAFKEN--KVLRVENVHEFPGHIACDSASNSEIV
NP_470973.1      GFYLLEKETN------------------QLVLGPFQGL-PACIRIP--LGKGVCGSAASDQ--KTYIVENVHDFPGHIACDAASNSEIV
YP_607257.1      GFYINRS--------------------EELVLGPFQGQ-VACVRIP--FSRGVCGAAAATR--TTQRVEDVHAFPGHIACDSASNSELV
YP_001860169.1   GFYFALD--------------------GELVIGPFQGK-PACVRIP--MGRGVCGRAAQAL--ETQVVPDVDAFPGHIACDSASRSEIV
NP_465120.1      GFYLLEKETN------------------QLVLGPFQGL-PACIRIP--LGKGVCGSAASDQ--KTYIVENVHDFPGHIACDAASNSEIV
NP_372242.1      GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
NP_374828.1      GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_001247140.1   GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
```

Supplemental data

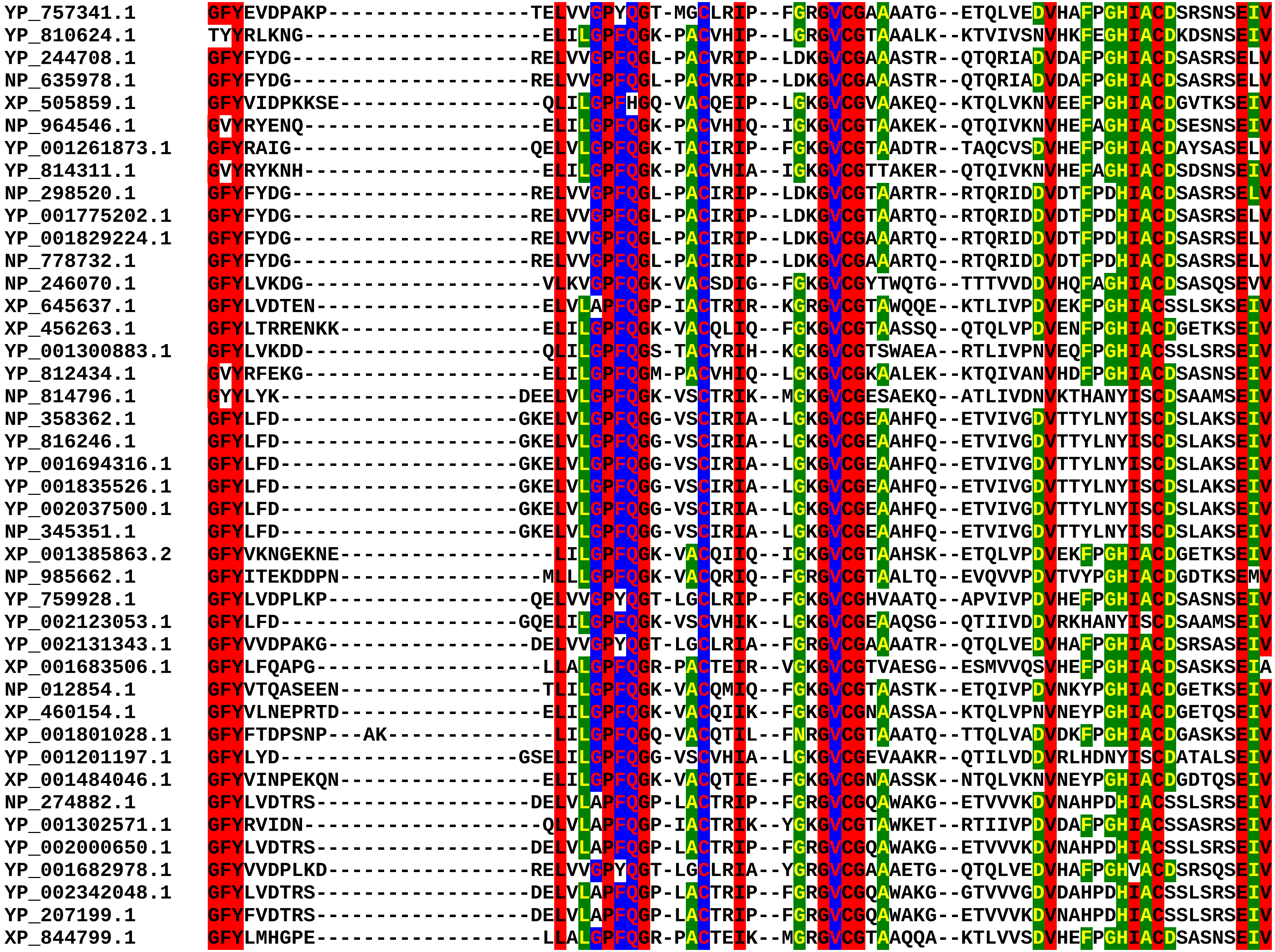
```
YP_001316940.1    GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_001442294.1    GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_002314836.1    GFYIA--EGN------------------ELVLGPFQGL-PACVRIP--FGKGVCGTAAAEK--RTVRVADVHAFPGHIACDAASQSEIV
YP_001814701.1    GFYLMDQEQN------------------ELVLGPFQGL-PACVRIP--LGRGVCGTAAQDQ--TTQLVMDVHAFPGHIACDAATNSEIV
NP_646478.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_043764.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_186602.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_494359.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_500333.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001332646.1    GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001575592.1    GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
ZP_02759966.1     GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQIVADVHQFEGHIACDANSKSEIV
YP_001487827.1    GFYLAK-G-------------------DELVLGPFNGL-PACVRIP--SGKGVCGTAFATG--EVQRIADVHAFPGHIACDAASQSEIV
YP_417047.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFEGHIACDANSKSEIV
YP_041183.1       GFYLL--EQN------------------ELILGPFQGH-PACVHIP--IGKGVCGTAVSER--RTQVVADVHQFKGHIACDANSKSEIV
YP_176251.1       GFYFL--KEN------------------ELVLGPFQGL-PACTRIA--IGKGVCGTAVAEN--KTMRIDDVHAFPGHIACDAASNAELV
YP_002349929.1    GFYLLEKETN------------------QLVLGPFQGL-PACIRIP--LGKGVCGSAADQ--KTYIVENVHDFPGHIACDAASNSEIV
YP_001376537.1    GFYVT--EGN------------------ELVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_784647.1       GFYFFDG--------------------QELVLGPFQGK-PACVRIA--LARGVCGAAASLR--QTQVVPDVHAFPGHIACDAASRSEIV
YP_557760.1       GFYFHDG--------------------QELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--QTQVVRDVHEFPGHIACDSASQSEIV
YP_275769.1       GFYLNRN--------------------EELVLGPFQGQ-IACVRIP--FGRGVCGAAAQSR--QTQRVQDVHEFPGHIACDSASNSELV
YP_001118914.1    GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--QTQVVRDVHAFPGHIACDAASESEIV
YP_001782545.1    GFYLMKDG--------------------ELVLGPFQGM-PACDRIK--IGKGVCGVAISSR--QIQRVDDVHKFEGHIACDSSTNSELV
NP_882192.1       GFYFHDG--------------------QELVLGPFQGK-PACVRIA--LSRGVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
YP_001894915.1    GFYFHDG--------------------KELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--QTQVVRDVREFPGHIACDSASQSEIV
YP_001647268.1    GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_882465.1       GFYFHDG--------------------QELVLGPFQGK-PACVRIA--LSRGVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
YP_002448219.1    GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_834359.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_002369448.1    GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_191278.1       GFYLWKEADR------------------QLVLGPFQGR-LACTRIP--LGRGVCGTVAQNR--ETLVVPDVHAFPGHIACDAASESEIV
NP_886654.1       GFYFHDG--------------------QELVLGPFQGK-PACVRIA--LSRGVCGAAASQR--RTQVVPDVHAFPGHIACDAASRAEIV
NP_847103.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_030799.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
NP_981083.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_038702.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_002340711.1    GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_085975.1       GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--ATQLVADVHQFPGHIACDSASNSEIV
YP_002453717.1    GFYVT--EGN------------------QLVLGPFQGM-PACVRIP--FGRGVCGVAAETK--TTQLVADVHQFPGHIACDSASNSEIV
YP_772921.1       GFYFFDG--------------------SELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHEFPGHIACDAASESEIV
YP_001843726.1    GFYRVQDG--------------------ELILGPFQGQ-VACVHIA--FGSGVCGTAAQTG--KTQLVKDVHQFAGHIACDAASNSEIV
YP_942902.1       GFYLL--KDN------------------QLVLGPFQGQ-PACIRIP--LNTGVCGKAFSTN--TIQRVADVQQFVGHIACDANSQSEIV
```

Supplemental data



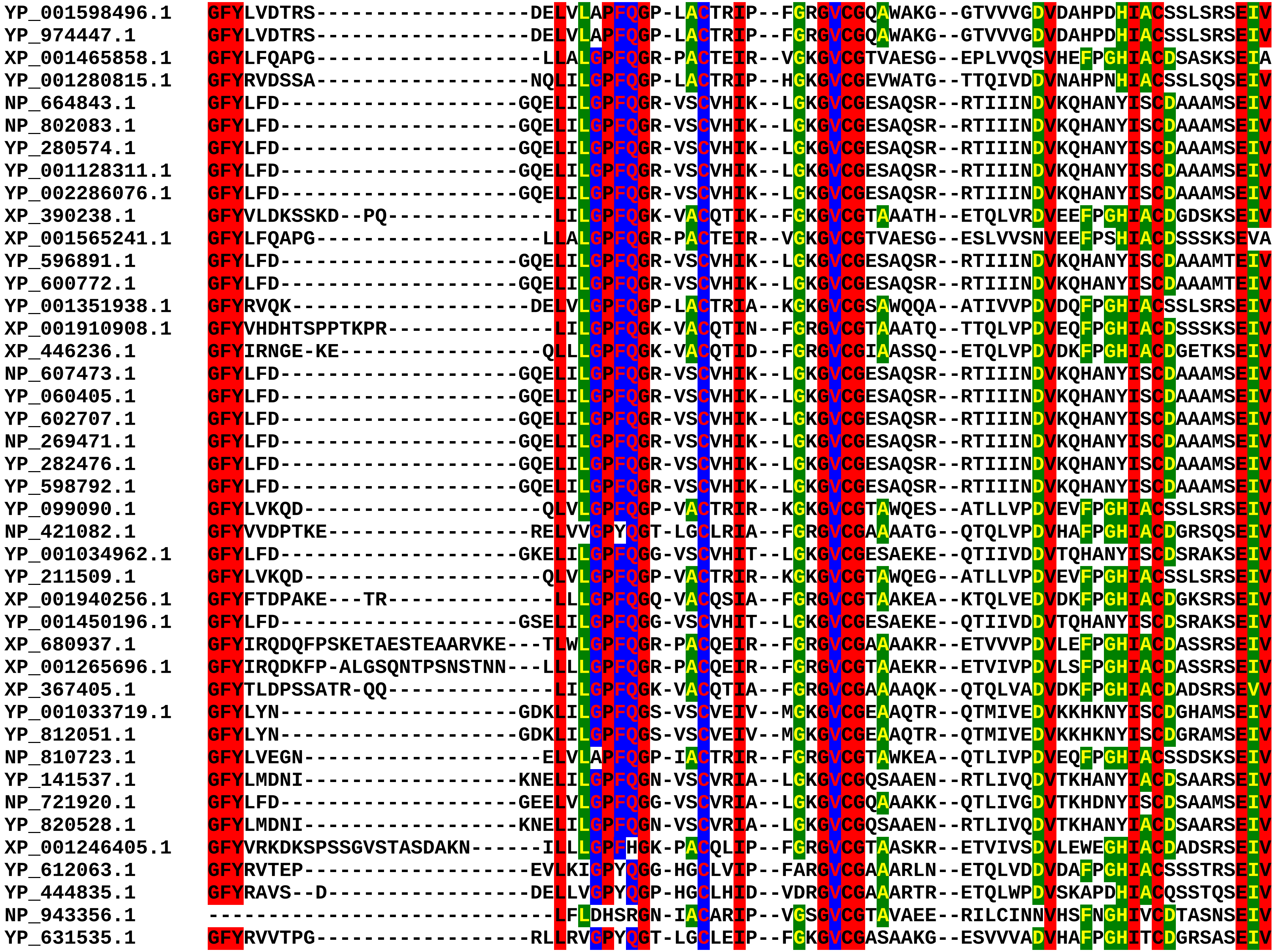
```
YP_121651.1       GFYFAVG--------------------DELVVGPFQGK-PACVRIA--FGKGVCGTAAATR--TTQLVPDVHAFPGHIACDADSRSEIV
YP_001633478.1    GFYFYDG--------------------RELVLGPFQGK-PACVRIA--LGRGVCGTAASQR--QTQVVADVNAFPGHIACDVASRSEVV
YP_001392189.1    GFYLMKDG--------------------ELVLGPFQGM-PACDRIK--IGKGVCGVAISSR--QIQRVDDVHKFEGHIACDSSTNSELV
YP_795394.1       GFYLYQPETD------------------DLILGPFQGN-VACVHIE--NNHGVCGTALATQ--TTQRIADVHQFPGHIACDSASNAEIV
YP_002284153.1    GFYFLRSE-------------------NELVLGPFQGR-VACVRIA--VGKGVCGTAVAEA--RSVLVPDVHEFPGHIACDAASRSELV
YP_001255404.1    GFYLMKDG--------------------ELVLGPFQGM-PACNRIK--IGKGVCGVAISSR--QIQRVNDVHKFEGHISCDSSTNSELV
YP_001385170.1    GFYLMKDG--------------------ELVLGPFQGM-PACNRIK--IGKGVCGVAISSR--QIQRVNDVHKFEGHISCDSSTNSELV
YP_001388639.1    GFYLMKDG--------------------ELVLGPFQGM-PACNRIK--IGKGVCGVAISSR--QIQRVNDVHKFEGHISCDSSTNSELV
YP_001347703.1    GFYLNRN--------------------EELLLGPFQGK-VACVRIP--FAKGVCGAAARTR--QTQRVEDVHAFPGHIACDSASRSELV
NP_103622.1       GFYFLASD-------------------EELVLGPFQGK-PACVRIA--VGKGVCGTAVQLG--TSMLIKDVHDFPGHIACDADSRSELV
YP_001807739.1    GFYFFDG--------------------SELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHEFPGHIACDAASESEIV
YP_001580337.1    GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHAFPGHIACDAASESEIV
YP_002230176.1    GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHDFPGHIACDAASESEIV
NP_251512.1       GFYLNRN--------------------EELVLGPFQGK-VACVRIP--FSKGVCGAAARTR--LTQRVDDVHAFPGHIACDSASSSELV
YP_459505.1       GFYRVVD--------------------GELVLGPFVGR-PACIRIP--IGQGVCGAAASSG--ETQLVEDVHAFPGHIACDAASQSELV
YP_790351.1       GFYLNRN--------------------EELVLGPFQGK-VACVRIP--FSKGVCGAAARTR--LTQRVDDVHAFPGHIACDSASSSELV
YP_002439846.1    GFYLNRN--------------------EELVLGPFQGK-VACVRIP--FSKGVCGAAARTR--LTQRVDDVHAFPGHIACDSASSSELV
YP_001970489.1    GFYLYDG--------------------KELVVGPFQGL-PACVRIP--LDKGVCGAAASQR--VTQRVEDVDAFPGHIACDSASRSELV
YP_617023.1       GFYRFDG--------------------QELVLGPFQGK-AACIRIP--LDKGVCGAAARLR--ATQRVEDVHAFPGHIACDADSRSELV
YP_620554.1       GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHDFPGHIACDAASESEIV
YP_834795.1       GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHDFPGHIACDAASESEIV
YP_148653.1       GFYLA--DGE------------------ELVLGPFQGL-PACVRIP--FGKGVCGTAAAKR--RTVVVPDVHQFPGHIACDAASQSEIV
YP_442277.1       GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASRSEIV
YP_334263.1       GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_001067073.1    GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_368503.1       GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHDFPGHIACDAASESEIV
YP_109010.1       GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_993758.1       GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_001028775.1    GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_002026852.1    GFYLYDG--------------------KELVVGPFQGL-PACVRIP--LDKGVCGAAASQR--VTQRVDDVDAFPGHIACDSASRSELV
YP_001764423.1    GFYFFDG--------------------TELVVGPFQGK-PACVRIA--LGKGVCGTAAQTR--ETQVVRDVHDFPGHIACDAASESEIV
NP_970249.1       GFYLLHNN--------------------ELLLSSFQGL-PACTRIA--IGKGVCGTAAKTL--QTQLVADVDQFPGHIVCDAASKSEIV
YP_001059788.1    GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTAAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_001081316.1    GFYFFDG--------------------RELVVGPFQGK-PACVRIP--LGKGVCGTTAQTR--GTQVVHDVHAFAGHIACDSASQSEIV
YP_001875000.1    GFYILKG-D-------------------KLVLGPFQGK-PACVRIS--LGKGVCGVSARER--KTLVVSDVHEFPGHIACDSASQSEIV
YP_278248.1       GFYLAEGE--------------------VLYLSSFQGK-IACTQIP--FSRGVCGKAAREQ--KTIVVDDVHEFKDHIACDSASNSEIV
YP_001309672.1    GFYLVKNN--------------------TLVLGPFQGM-PACTKIE--IGKGVCGKAALEK--ETLLVKDVHNFEGHIACDAASNSEIV
YP_001576818.1    GVYRFENG--------------------ELILGPFQGM-PACVHIQ--MGKGVCGTAAQTE--KTQIVPNVHEFQGHIACDSASNSEIV
YP_193206.1       GVYRLENG--------------------ELILGPFQGM-PACVHIP--VGKGVCGVAAKTQ--ITQIVPNVHEFAGHIACDSASNSEIV
YP_199405.1       GFYFYDG--------------------RELVVGPFQGL-PACVRIP--LDKGVCGAAASTR--QSQRIADVDAFPGHIACDSASRSELV
YP_001915608.1    GFYFYDG--------------------RELVVGPFQGL-PACVRIP--LDKGVCGAAASTR--QSQRIADVDAFPGHIACDSASRSELV
YP_002509654.1    GFYLMR-G-------------------GELVLGPFQGL-PACVRIK--VGRGVCGSAVRDR--KTYLVPDVHKFPGHIACDEASRSEIV
```

Supplemental data



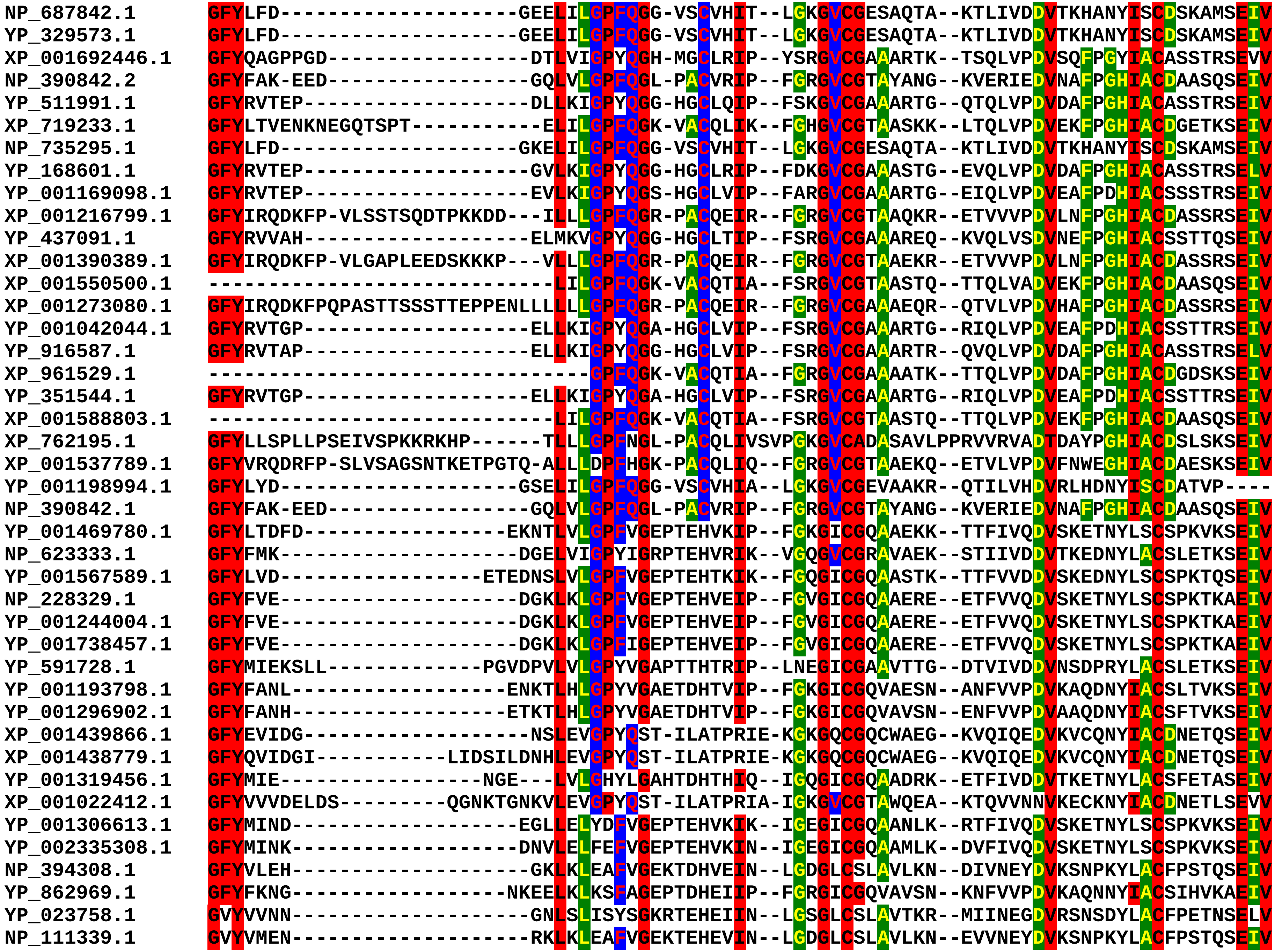
```
YP_757341.1      GFYEVDPAKP-----------------TELVVGPYQGT-MGCLRIP--FGRGVCGAAAATG--ETQLVEDVHAFPGHIACDSRSNSEIV
YP_810624.1      TYYRLKNG--------------------ELILGPFQGK-PACVHIP--LGRGVCGTAAALK--KTVIVSNVHKFEGHIACDKDSNSEIV
YP_244708.1      GFYFYDG--------------------RELVVGPFQGL-PACVRIP--LDKGVCGAAASTR--QTQRIADVDAFPGHIACDSASRSELV
NP_635978.1      GFYFYDG--------------------RELVVGPFQGL-PACVRIP--LDKGVCGAAASTR--QTQRIADVDAFPGHIACDSASRSELV
XP_505859.1      GFYVIDPKKSE-----------------QLILGPFHGQ-VACQEIP--LGKGVCGVAAKEQ--KTQLVKNVEEFPGHIACDGVTKSEIV
NP_964546.1      GVYRYENQ--------------------ELILGPFQGK-PACVHIQ--IGKGVCGTAAKEK--QTQIVKNVHEFAGHIACDSESNSEIV
YP_001261873.1   GFYRAIG--------------------QELVLGPFQGK-TACIRIP--FGKGVCGTAADTR--TAQCVSDVHEFPGHIACDAYSASELV
YP_814311.1      GVYRYKNH--------------------ELILGPFQGK-PACVHIA--IGKGVCGTTAKER--QTQIVKNVHEFAGHIACDSDSNSEIV
NP_298520.1      GFYFYDG--------------------RELVVGPFQGL-PACIRIP--LDKGVCGTAARTR--RTQRIDDVDTFPDHIACDSASRSELV
YP_001775202.1   GFYFYDG--------------------RELVVGPFQGL-PACIRIP--LDKGVCGTAARTQ--RTQRIDDVDTFPDHIACDSASRSELV
YP_001829224.1   GFYFYDG--------------------RELVVGPFQGL-PACIRIP--LDKGVCGAAARTQ--RTQRIDDVDTFPDHIACDSASRSELV
NP_778732.1      GFYFYDG--------------------RELVVGPFQGL-PACIRIP--LDKGVCGAAARTQ--RTQRIDDVDTFPDHIACDSASRSELV
NP_246070.1      GFYLVKDG--------------------VLKVGPFQGK-VACSDIG--FGKGVCGYTWQTG--TTTVVDDVHQFAGHIACDSASQSEVV
XP_645637.1      GFYLVDTEN-------------------ELVLAPFQGP-IACTRIR--KGRGVCGTAWQQE--KTLIVPDVEKFPGHIACSSLSKSEIV
XP_456263.1      GFYLTRRENKK-----------------ELILGPFQGK-VACQLIQ--FGKGVCGTAASSQ--QTQLVPDVENFPGHIACDGETKSEIV
YP_001300883.1   GFYLVKDD--------------------QLILGPFQGS-TACYRIH--KGKGVCGTSWAEA--RTLIVPNVEQFPGHIACSSLSRSEIV
YP_812434.1      GVYRFEKG--------------------ELILGPFQGM-PACVHIQ--LGKGVCGKAALEK--KTQIVANVHDFPGHIACDSASNSEIV
NP_814796.1      GYYLYK--------------------DEELVLGPFQGK-VSCTRIK--MGKGVCGESAEKQ--ATLIVDNVKTHANYISCDSAAMSEIV
NP_358362.1      GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
YP_816246.1      GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
YP_001694316.1   GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
YP_001835526.1   GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
YP_002037500.1   GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
NP_345351.1      GFYLFD--------------------GKELVLGPFQGG-VSCIRIA--LGKGVCGEAAHFQ--ETVIVGDVTTYLNYISCDSLAKSEIV
XP_001385863.2   GFYVKNGEKNE------------------LILGPFQGK-VACQIIQ--IGKGVCGTAAHSK--ETQLVPDVEKFPGHIACDGETKSEIV
NP_985662.1      GFYITEKDDPN-----------------MLLLGPFQGK-VACQRIQ--FGRGVCGTAALTQ--EVQVVPDVTVYPGHIACDGDTKSEMV
YP_759928.1      GFYLVDPLKP-----------------QELVVGPYQGT-LGCLRIP--FGKGVCGHVAATQ--APVIVPDVHEFPGHIACDSASNSEIV
YP_002123053.1   GFYLFD--------------------GQELILGPFQGK-VSCVHIK--LGKGVCGEAAQSG--QTIIVDDVRKHANYISCDSAAMSEIV
YP_002131343.1   GFYVVDPAKG-----------------DELVVGPYQGT-LGCLRIA--FGRGVCGAAAATR--QTQLVEDVHAFPGHIACDSRSASEIV
XP_001683506.1   GFYLFQAPG-------------------LLALGPFQGR-PACTEIR--VGKGVCGTVAESG--ESMVVQSVHEFPGHIACDSASKSEIA
NP_012854.1      GFYVTQASEEN-----------------TLILGPFQGK-VACQMIQ--FGKGVCGTAASTK--ETQIVPDVNKYPGHIACDGETKSEIV
XP_460154.1      GFYVLNEPRTD-----------------ELILGPFQGK-VACQIIK--FGKGVCGNAASSA--KTQLVPNVNEYPGHIACDGETQSEIV
XP_001801028.1   GFYFTDPSNP---AK--------------LILGPFQGQ-VACQTIL--FNRGVCGTAAATQ--TTQLVADVDKFPGHIACDGASKSEIV
YP_001201197.1   GFYLYD--------------------GSELILGPFQGG-VSCVHIA--LGKGVCGEVAAKR--QTILVDDVRLHDNYISCDATALSEIV
XP_001484046.1   GFYVINPEKQN-----------------ELILGPFQGK-VACQTIE--FGKGVCGNAASSK--NTQLVKNVNEYPGHIACDGDTQSEIV
NP_274882.1      GFYLVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--ETVVVKDVNAHPDHIACSSLSRSEIV
YP_001302571.1   GFYRVIDN--------------------QLVLAPFQGP-IACTRIK--YGKGVCGTAWKET--RTIIVPDVDAFPGHIACSSASRSEIV
YP_002000650.1   GFYLVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--ETVVVKDVNAHPDHIACSSLSRSEIV
YP_001682978.1   GFYVVDPLKD-----------------RELVVGPYQGT-LGCLRIA--YGRGVCGAAAETG--QTQLVEDVHAFPGHVACDSRSQSEIV
YP_002342048.1   GFYLVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--GTVVVGDVDAHPDHIACSSLSRSEIV
YP_207199.1      GFYFVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--ETVVVKDVNAHPDHIACSSLSRSEIV
XP_844799.1      GFYLMHGPE-------------------LLALGPFQGR-PACTEIK--MGRGVCGTAAQQA--KTLVVSDVHEFPGHIACDSASNSEIV
```

Supplemental data



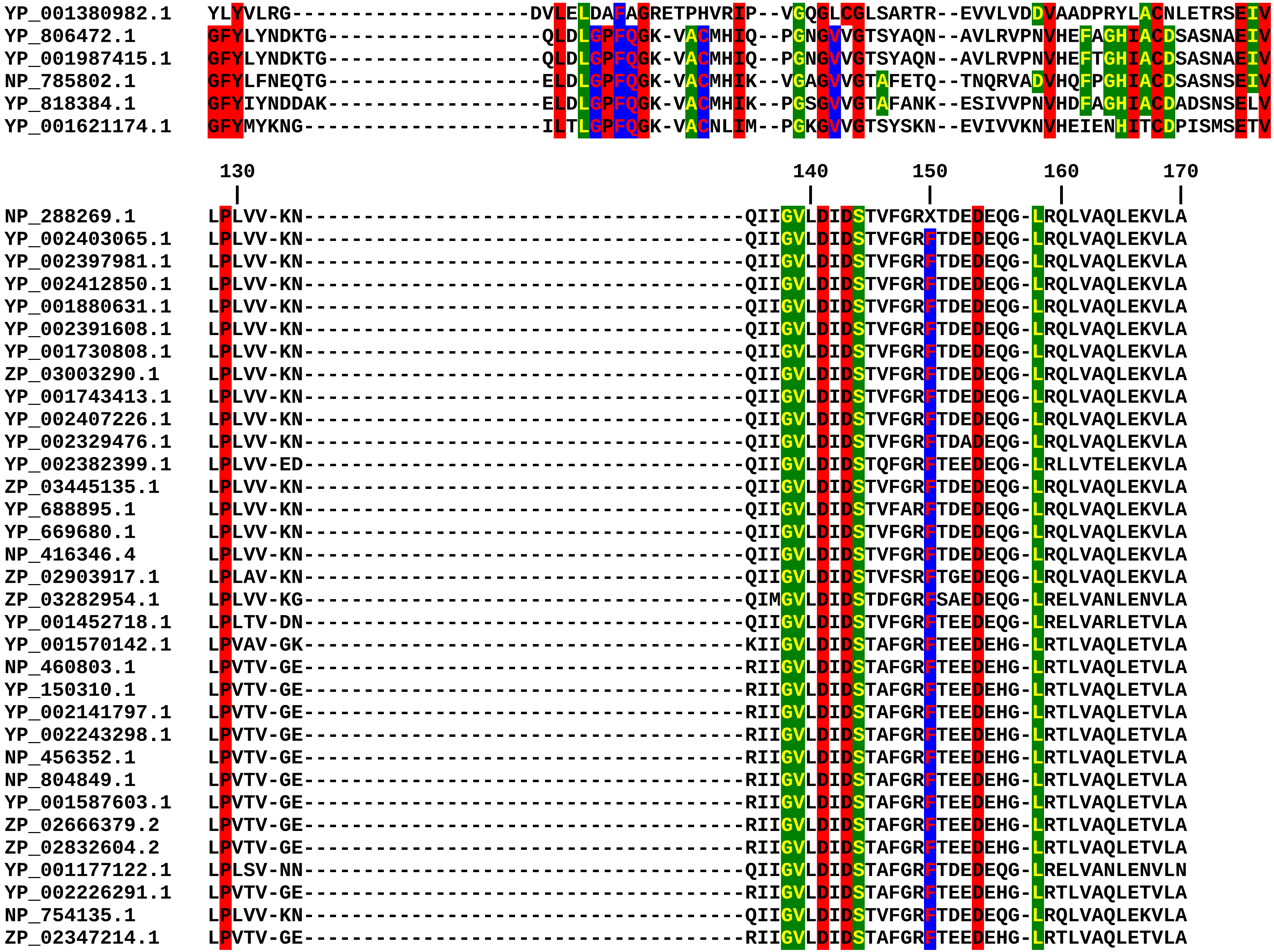
```
YP_001598496.1    GFYLVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--GTVVVGDVDAHPDHIACSSLSRSEIV
YP_974447.1       GFYLVDTRS------------------DELVLAPFQGP-LACTRIP--FGRGVCGQAWAKG--GTVVVGDVDAHPDHIACSSLSRSEIV
XP_001465858.1    GFYLFQAPG-------------------LLALGPFQGR-PACTEIR--VGKGVCGTVAESG--EPLVVQSVHEFPGHIACDSASKSEIA
YP_001280815.1    GFYRVDSSA------------------NQLILGPFQGP-LACTRIP--HGKGVCGEVWATG--TTQIVDDVNAHPNHIACSSLSQSEIV
NP_664843.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
NP_802083.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_280574.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_001128311.1    GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_002286076.1    GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
XP_390238.1       GFYVLDKSSKD--PQ--------------LILGPFQGK-VACQTIK--FGKGVCGTAAATH--ETQLVRDVEEFPGHIACDGDSKSEIV
XP_001565241.1    GFYLFQAPG-------------------LLALGPFQGR-PACTEIR--VGKGVCGTVAESG--ESLVVSNVEEFPSHIACDSSSKSEVA
YP_596891.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMTEIV
YP_600772.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMTEIV
YP_001351938.1    GFYRVQK--------------------DELVLGPFQGP-LACTRIA--KGKGVCGSAWQQA--ATIVVPDVDQFPGHIACSSLSRSEIV
XP_001910908.1    GFYVHDHTSPPTKPR--------------LILGPFQGK-VACQTIN--FGRGVCGTAAATQ--TTQLVPDVEQFPGHIACDSSSKSEIV
XP_446236.1       GFYIRNGE-KE-----------------QLLLGPFQGK-VACQTID--FGRGVCGIAASSQ--ETQLVPDVDKFPGHIACDGETKSEIV
NP_607473.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_060405.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_602707.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
NP_269471.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_282476.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_598792.1       GFYLFD--------------------GQELILGPFQGR-VSCVHIK--LGKGVCGESAQSR--RTIIINDVKQHANYISCDAAAMSEIV
YP_099090.1       GFYLVKQD--------------------QLVLGPFQGP-VACTRIR--KGKGVCGTAWQES--ATLLVPDVEVFPGHIACSSLSRSEIV
NP_421082.1       GFYVVDPTKE-----------------RELVVGPYQGT-LGCLRIA--FGRGVCGAAAATG--QTQLVPDVHAFPGHIACDGRSQSEIV
YP_001034962.1    GFYLFD--------------------GKELILGPFQGG-VSCVHIT--LGKGVCGESAEKE--QTIIVDDVTQHANYISCDSRAKSEIV
YP_211509.1       GFYLVKQD--------------------QLVLGPFQGP-VACTRIR--KGKGVCGTAWQEG--ATLLVPDVEVFPGHIACSSLSRSEIV
XP_001940256.1    GFYFTDPAKE---TR--------------LLLGPFQGQ-VACQSIA--FGRGVCGTAAKEA--KTQLVEDVDKFPGHIACDGKSRSEIV
YP_001450196.1    GFYLFD--------------------GSELILGPFQGG-VSCVHIT--LGKGVCGESAEKE--QTIIVDDVTQHANYISCDSRAKSEIV
XP_680937.1       GFYIRQDQFPSKETAESTEAARVKE---TLWLGPFQGR-PACQEIR--FGRGVCGAAAAKR--ETVVVPDVLEFPGHIACDASSRSEIV
XP_001265696.1    GFYIRQDKFP-ALGSQNTPSNSTNN---LLLLGPFQGR-PACQEIR--FGRGVCGTAAEKR--ETVIVPDVLSFPGHIACDASSRSEIV
XP_367405.1       GFYTLDPSSATR-QQ--------------LILGPFQGK-VACQTIA--FGRGVCGAAAAQK--QTQLVADVDKFPGHIACDADSRSEVV
YP_001033719.1    GFYLYN--------------------GDKLILGPFQGS-VSCVEIV--MGKGVCGEAAQTR--QTMIVEDVKKHKNYISCDGHAMSEIV
YP_812051.1       GFYLYN--------------------GDKLILGPFQGS-VSCVEIV--MGKGVCGEAAQTR--QTMIVEDVKKHKNYISCDGRAMSEIV
NP_810723.1       GFYLVEGN--------------------ELVLAPFQGP-IACTRIR--FGRGVCGTAWKEA--QTLIVPDVEQFPGHIACSSDSKSEIV
YP_141537.1       GFYLMDNI------------------KNELILGPFQGN-VSCVRIA--LGKGVCGQSAAEN--RTLIVQDVTKHANYIACDSAARSEIV
NP_721920.1       GFYLFD--------------------GEELVLGPFQGG-VSCVRIA--LGKGVCGQAAAKK--QTLIVGDVTKHDNYISCDSAAMSEIV
YP_820528.1       GFYLMDNI------------------KNELILGPFQGN-VSCVRIA--LGKGVCGQSAAEN--RTLIVQDVTKHANYIACDSAARSEIV
XP_001246405.1    GFYVRKDKSPSSGVSTASDAKN------ILLLGPFHGK-PACQLIP--FGRGVCGTAASKR--ETVIVSDVLEWEGHIACDADSRSEIV
YP_612063.1       GFYRVTEP-------------------EVLKIGPYQGG-HGCLVIP--FARGVCGAAARLN--ETQLVDDVDAFPGHIACSSSTRSEIV
YP_444835.1       GFYRAVS--D-----------------DELLVGPYQGP-HGCLHID--VDRGVCGAAARTR--ETQLWPDVSKAPDHIACQSSTQSEIV
NP_943356.1       -----------------------------LFLDHSRGN-IACARIP--VGSGVCGTAVAEE--RILCINNVHSFNGHIVCDTASNSEIV
YP_631535.1       GFYRVVTPG------------------RLLRVGPYQGT-LGCLEIP--FGKGVCGASAAKG--ESVVVADVHAFPGHITCDGRSASEIV
```

Supplemental data



```
NP_687842.1       GFYLFD--------------------GEELILGPFQGG-VSCVHIT--LGKGVCGESAQTA--KTLIVDDVTKHANYISCDSKAMSEIV
YP_329573.1       GFYLFD--------------------GEELILGPFQGG-VSCVHIT--LGKGVCGESAQTA--KTLIVDDVTKHANYISCDSKAMSEIV
XP_001692446.1    GFYQAGPPGD-----------------DTLVIGPYQGH-MGCLRIP--YSRGVCGAAARTK--TSQLVPDVSQFPGYIACASSTRSEVV
NP_390842.2       GFYFAK-EED-----------------GQLVLGPFQGL-PACVRIP--FGRGVCGTAYANG--KVERIEDVNAFPGHIACDAASQSEIV
YP_511991.1       GFYRVTEP-------------------DLLKIGPYQGG-HGCLQIP--FSKGVCGAAARTG--QTQLVPDVDAFPGHIACASSTRSEIV
XP_719233.1       GFYLTVENKNEGQTSPT-----------ELILGPFQGK-VACQLIK--FGHGVCGTAASKK--LTQLVPDVEKFPGHIACDGETKSEIV
NP_735295.1       GFYLFD--------------------GKELILGPFQGG-VSCVHIT--LGKGVCGESAQTA--KTLIVDDVTKHANYISCDSKAMSEIV
YP_168601.1       GFYRVTEP-------------------GVLKIGPYQGG-HGCLRIP--FDKGVCGAAASTG--EVQLVPDVDAFPGHIACASSTRSELV
YP_001169098.1    GFYRVTEP-------------------EVLKIGPYQGS-HGCLVIP--FARGVCGAAARTG--EIQLVPDVEAFPDHIACSSSTRSEIV
XP_001216799.1    GFYIRQDKFP-VLSSTSQDTPKKDD---ILLLGPFQGR-PACQEIR--FGRGVCGTAAQKR--ETVVVPDVLNFPGHIACDASSRSEIV
YP_437091.1       GFYRVVAH-------------------ELMKVGPYQGG-HGCLTIP--FSRGVCGAAAREQ--KVQLVSDVNEFPGHIACSSTTQSEIV
XP_001390389.1    GFYIRQDKFP-VLGAPLEEDSKKKP---VLLLGPFQGR-PACQEIR--FGRGVCGTAAEKR--ETVVVPDVLNFPGHIACDASSRSEIV
XP_001550500.1    -----------------------------LILGPFQGK-VACQTIA--FSRGVCGTAASTQ--TTQLVADVEKFPGHIACDAASQSEIV
XP_001273080.1    GFYIRQDKFPQPASTTSSSTTEPPENLLLLLLGPFQGR-PACQEIR--FGRGVCGAAAEQR--QTVLVPDVHAFPGHIACDASSRSEIV
YP_001042044.1    GFYRVTGP-------------------ELLKIGPYQGA-HGCLVIP--FSRGVCGAAARTG--RIQLVPDVEAFPDHIACSSTTRSEIV
YP_916587.1       GFYRVTAP-------------------ELLKIGPYQGG-HGCLVIP--FSRGVCGAAARTR--QVQLVPDVDAFPGHIACASSTRSELV
XP_961529.1       --------------------------------GPFQGK-VACQTIA--FGRGVCGAAAATK--TTQLVPDVDAFPGHIACDGDSKSEIV
YP_351544.1       GFYRVTGP-------------------ELLKIGPYQGA-HGCLVIP--FSRGVCGAAARTG--RIQLVPDVEAFPDHIACSSTTRSEIV
XP_001588803.1    -----------------------------LILGPFQGK-VACQTIA--FSRGVCGTAASTQ--TTQLVPDVEKFPGHIACDAASQSEIV
XP_762195.1       GFYLLSPLLPSEIVSPKKRKHP------TLLLGPFNGL-PACQLIVSVPGKGVCADASAVLPPRVVRVADTDAYPGHIACDSLSKSEIV
XP_001537789.1    GFYVRQDRFP-SLVSAGSNTKETPGTQ-ALLLDPFHGK-PACQLIQ--FGRGVCGTAAEKQ--ETVLVPDVFNWEGHIACDAESKSEIV
YP_001198994.1    GFYLYD--------------------GSELILGPFQGG-VSCVHIA--LGKGVCGEVAAKR--QTILVHDVRLHDNYISCDATVP----
NP_390842.1       GFYFAK-EED-----------------GQLVLGPFQGL-PACVRIP--FGRGVCGTAYANG--KVERIEDVNAFPGHIACDAASQSEIV
YP_001469780.1    GFYLTDFD-----------------EKNTLVLGPFVGEPTEHVKIP--FGKGICGQAAEKK--TTFIVQDVSKETNYLSCSPKVKSEIV
NP_623333.1       GFYFMK--------------------DGELVIGPYIGRPTEHVRIK--VGQGVCGRAVAEK--STIIVDDVTKEDNYLACSLETKSEIV
YP_001567589.1    GFYLVD-----------------ETEDNSLVLGPFVGEPTEHTKIK--FGQGICGQAASTK--TTFVVDDVSKEDNYLSCSPKTQSEIV
NP_228329.1       GFYFVE--------------------DGKLKLGPFVGEPTEHVEIP--FGVGICGQAAERE--ETFVVQDVSKETNYLSCSPKTKAEIV
YP_001244004.1    GFYFVE--------------------DGKLKLGPFVGEPTEHVEIP--FGVGICGQAAERE--ETFVVQDVSKETNYLSCSPKTKAEIV
YP_001738457.1    GFYFVE--------------------DGKLKLGPFIGEPTEHVEIP--FGVGICGQAAERE--ETFVVQDVSKETNYLSCSPKTKAEIV
YP_591728.1       GFYMIEKSLL-------------PGVDPVLVLGPYVGAPTTHTRIP--LNEGICGAAVTTG--DTVIVDDVNSDPRYLACSLETKSEIV
YP_001193798.1    GFYFANL------------------ENKTLHLGPYVGAETDHTVIP--FGKGICGQVAESN--ANFVVPDVKAQDNYIACSLTVKSEIV
YP_001296902.1    GFYFANH------------------ETKTLHLGPYVGAETDHTVIP--FGKGICGQVAVSN--ENFVVPDVAAQDNYIACSFTVKSEIV
XP_001439866.1    GFYEVIDG-------------------NSLEVGPYQST-ILATPRIE-KGKGQCGQCWAEG--KVQIQEDVKVCQNYIACDNETQSEIV
XP_001438779.1    GFYQVIDGI-----------LIDSILDNHLEVGPYQST-ILATPRIE-KGKGQCGQCWAEG--KVQIQEDVKVCQNYIACDNETQSEIV
YP_001319456.1    GFYMIE-----------------NGE---LVLGHYLGAHTDHTHIQ--IGQGICGQAADRK--ETFIVDDVTKETNYLACSFETASEIV
XP_001022412.1    GFYVVVDELDS---------QGNKTGNKVLEVGPYQST-ILATPRIA-IGKGVCGTAWQEA--KTQVVNNVKECKNYIACDNETLSEVV
YP_001306613.1    GFYMIND-------------------EGLLELYDFVGEPTEHVKIK--IGEGICGQAANLK--RTFIVQDVSKETNYLSCSPKVKSEIV
YP_002335308.1    GFYMINK-------------------DNVLELFEFVGEPTEHVKIN--IGEGICGQAAMLK--DVFIVQDVSKETNYLSCSPKVKSEIV
NP_394308.1       GFYVLEH--------------------GKLKLEAFVGEKTDHVEIN--LGDGLCSLAVLKN--DIVNEYDVKSNPKYLACFPSTQSEIV
YP_862969.1       GFYFKNG------------------NKEELKLKSFAGEPTDHEIIP--FGRGICGQVAVSN--KNFVVPDVKAQNNYIACSIHVKAEIV
YP_023758.1       GVYVVNN--------------------GNLSLISYSGKRTEHEIIN--LGSGLCSLAVTKR--MIINEGDVRSNSDYLACFPETNSELV
NP_111339.1       GVYVMEN--------------------RKLKLEAFVGEKTEHEVIN--LGDGLCSLAVLKN--EVVNEYDVKSNPKYLACFPSTQSEIV
```

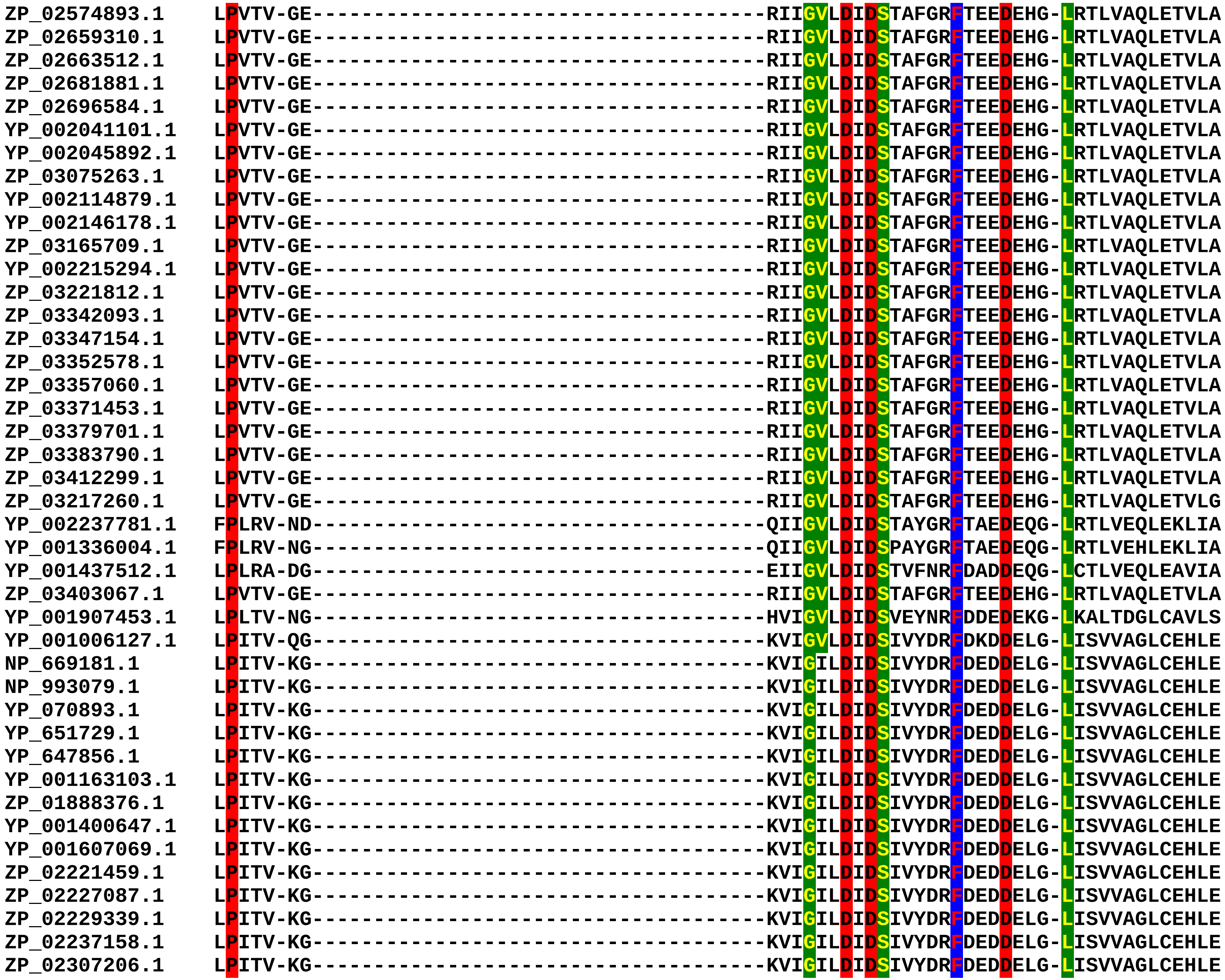
Supplemental data



```
YP_001380982.1    YLYVLRG--------------------DVLELDAFAGRETPHVRIP--VGQGLCGLSARTR--EVVLVDDVAADPRYLACNLETRSEIV
YP_806472.1       GFYLYNDKTG------------------QLDLGPFQGK-VACMHIQ--PGNGVVGTSYAQN--AVLRVPNVHEFAGHIACDSASNAEIV
YP_001987415.1    GFYLYNDKTG------------------QLDLGPFQGK-VACMHIQ--PGNGVVGTSYAQN--AVLRVPNVHEFTGHIACDSASNAEIV
NP_785802.1       GFYLFNEQTG------------------ELDLGPFQGK-VACMHIK--VGAGVVGTAFETQ--TNQRVADVHQFPGHIACDSASNSEIV
YP_818384.1       GFYIYNDDAK------------------ELDLGPFQGK-VACMHIK--PGSGVVGTAFANK--ESIVVPNVHDFAGHIACDADSNSELV
YP_001621174.1    GFYMYKNG--------------------ILTLGPFQGK-VACNLIM--PGKGVVGTSYSKN--EVIVVKNVHEIENHITCDPISMSETV

                   130                                             140       150        160       170
                    |                                               |         |          |         |
NP_288269.1       LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRXTDEDEQG-LRQLVAQLEKVLA
YP_002403065.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_002397981.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_002412850.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_001880631.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_002391608.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_001730808.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
ZP_03003290.1     LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_001743413.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_002407226.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_002329476.1    LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDADEQG-LRQLVAQLEKVLA
YP_002382399.1    LPLVV-ED-------------------------------------QIIGVLDIDSTQFGRFTEEDEQG-LRLLVTELEKVLA
ZP_03445135.1     LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
YP_688895.1       LPLVV-KN-------------------------------------QIIGVLDIDSTVFARFTDEDEQG-LRQLVAQLEKVLA
YP_669680.1       LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
NP_416346.4       LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
ZP_02903917.1     LPLAV-KN-------------------------------------QIIGVLDIDSTVFSRFTGEDEQG-LRQLVAQLEKVLA
ZP_03282954.1     LPLVV-KG-------------------------------------QIMGVLDIDSTDFGRFSAEDEQG-LRELVANLENVLA
YP_001452718.1    LPLTV-DN-------------------------------------QIIGVLDIDSTVFGRFTEEDEQG-LRELVARLETVLA
YP_001570142.1    LPVAV-GK-------------------------------------KIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
NP_460803.1       LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_150310.1       LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002141797.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002243298.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
NP_456352.1       LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
NP_804849.1       LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_001587603.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02666379.2     LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02832604.2     LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_001177122.1    LPLSV-NN-------------------------------------QIIGVLDIDSTAFGRFTDEDEQG-LRELVANLENVLN
YP_002226291.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
NP_754135.1       LPLVV-KN-------------------------------------QIIGVLDIDSTVFGRFTDEDEQG-LRQLVAQLEKVLA
ZP_02347214.1     LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
```

Supplemental data

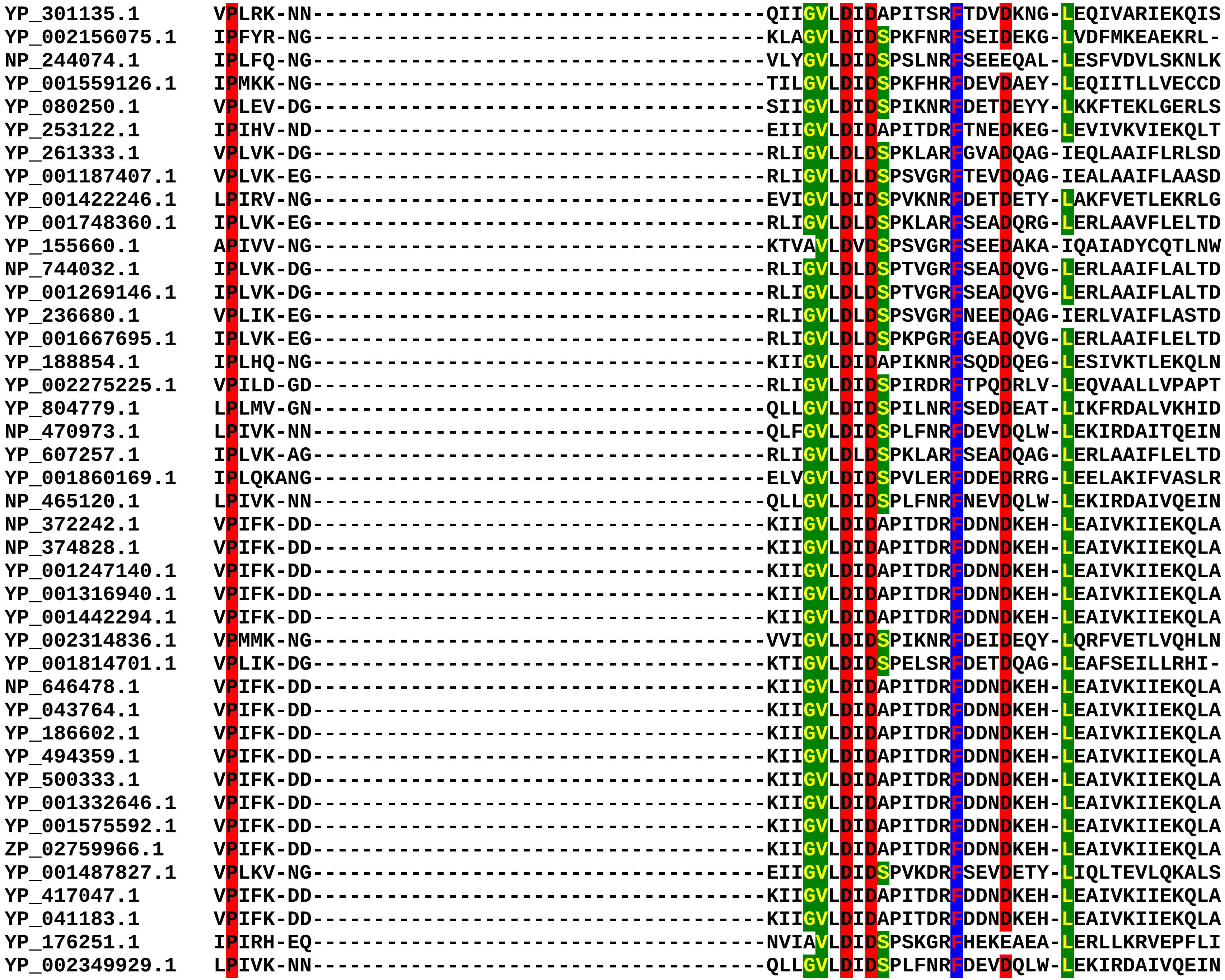


```
ZP_02574893.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02659310.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02663512.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02681881.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_02696584.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002041101.1   LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002045892.1   LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03075263.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002114879.1   LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002146178.1   LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03165709.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_002215294.1   LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03221812.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03342093.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03347154.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03352578.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03357060.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03371453.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03379701.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03383790.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03412299.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
ZP_03217260.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLG
YP_002237781.1   FPLRV-ND-------------------------------------QIIGVLDIDSTAYGRFTAEDEQG-LRTLVEQLEKLIA
YP_001336004.1   FPLRV-NG-------------------------------------QIIGVLDIDSPAYGRFTAEDEQG-LRTLVEHLEKLIA
YP_001437512.1   LPLRA-DG-------------------------------------EIIGVLDIDSTVFNRFDADDEQG-LCTLVEQLEAVIA
ZP_03403067.1    LPVTV-GE-------------------------------------RIIGVLDIDSTAFGRFTEEDEHG-LRTLVAQLETVLA
YP_001907453.1   LPLTV-NG-------------------------------------HVIGVLDIDSVEYNRFDDEDEKG-LKALTDGLCAVLS
YP_001006127.1   LPITV-QG-------------------------------------KVIGVLDIDSIVYDRFDKDDELG-LISVVAGLCEHLE
NP_669181.1      LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
NP_993079.1      LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_070893.1      LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_651729.1      LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_647856.1      LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001163103.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_01888376.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001400647.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001607069.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02221459.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02227087.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02229339.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02237158.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02307206.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
```

<u>Supplemental data</u>

```
ZP_02314388.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
ZP_02318544.1    LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001720525.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001872871.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_002346706.1   LPITV-KG-------------------------------------KVIGILDIDSIVYDRFDEDDELG-LISVVAGLCEHLE
YP_001478346.1   LPLNV-GG-------------------------------------QIIGVLDIDSTVYQRFDEQDERG-LEAVVAGLCAQLE
YP_454994.1      LPITL-DG-------------------------------------RVAGVLDIDSTIYQRFDQDDEEG-LKAVADYLCRQLG
YP_001217052.1   IPFSI-NG-------------------------------------KVAGVLDIDSPNIGRFSEIDEQG-LTYLMSEVEKLLN
YP_002150764.1   LPLEV-NG-------------------------------------QIIGVLDIDSPIFGRFDKKDENG-LKALCDVLCEHLK
YP_856816.1      IPLRR-GG-------------------------------------EIIGVLDIDSPIFNRFDAQDRIG-LEETVQILESML-
YP_001093925.1   IPLRR-DG-------------------------------------QVIGVLDIDSPSFSRFDEADQLG-LESCVKSLEKALF
YP_269525.1      VPVKI-AG-------------------------------------KVIAILDIDSTVFDRFNLEDEQG-IVSIVQAFEQNIT
NP_718189.1      IPVRA-DG-------------------------------------KVIAVLDIDSPIFDHFDEEDQKG-LELLVKSFEKCLF
YP_001366642.1   VPVCA-QG-------------------------------------NVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLF
YP_001554990.1   VPVCA-QG-------------------------------------NVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLF
YP_737796.1      IPVRA-DG-------------------------------------KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLF
YP_869415.1      IPVRA-DG-------------------------------------KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLF
NP_797987.1      IPFSI-KG-------------------------------------KVAGVLDIDSPSIGRFNETDEEG-LTHFMSEVEKLLN
YP_001501790.1   VPLRQ-NG-------------------------------------TVIGVLDIDSPSLSRFSEIDQQG-VETLVKSLEICLI
YP_927460.1      VPVRH-QG-------------------------------------RVIAVLDIDSPSFSRFDETDQAG-LETLVRTLEKALF
YP_001474222.1   IPIMR-DH-------------------------------------EVVAVLDIDSPSLARFAETDQVG-LEKLVNSFEKALF
YP_963195.1      VPVRA-NG-------------------------------------KVIAVLDIDSPSFDRFDEDDQKG-LELLVKSFENCLF
YP_001183724.1   VPVRA-NG-------------------------------------KVIAVLDIDSPSFDRFDEDDQKG-LELLVKSFENCLF
YP_733804.1      IPVRA-DG-------------------------------------KVVAVLDIDSPIFDRFDEEDQKG-LELLVKSFENCLF
YP_001050811.1   VPVCA-QG-------------------------------------NVVAVLDIDSPIFDRFDEDDQKG-LELLVKSFENCLF
YP_750914.1      IPVRQ-GN-------------------------------------EVVAVLDIDSPLLNRFDQDDQIG-LENMVKSFETCLF
YP_002126297.1   VPLVL-ND-------------------------------------SLVGVLDIDSPVFERFTEEDEKG-LVEIAQILMDSQK
YP_130129.1      IPLTV-KG-------------------------------------ELIGVLDIDSPSLSRFDQNDEDG-LVSFVKELQKSL-
YP_001445537.1   IPFTI-NG-------------------------------------EVAGVLDIDSPNIGRFNEIDEEG-LTYFMAEVEKLLN
YP_661562.1      LPVYQ-KG-------------------------------------QLLGVMDLDSPHTGRFDQDDMTG-LQNLVNILQDAIA
YP_001674584.1   VPLRK-NG-------------------------------------SVIGVLDIDSPSLSRFSEIDQQG-VETLVKSLEFCLF
YP_002417027.1   IPFSI-DG-------------------------------------KIAGVLDIDSPNVGRFSQIDEDG-LTFFMAEVEKLLN
NP_934438.1      IPFSI-NG-------------------------------------QLAGVLDIDSPSIGRFSEIDEQG-LTFFMAEVEKLLN
NP_761466.1      IPFSI-NG-------------------------------------QLAGVLDIDSPSVGRFSEIDEQG-LTFFMAEVEKLLN
YP_001760510.1   IPVFD-NN-------------------------------------RLIGVLDIDSPILNRFTKTDQIG-LENLVKSFESALF
YP_002312133.1   VPFRR-NG-------------------------------------DVIGVLDIDSPSLSRFDEIDQQG-VETLVKSLENCLF
YP_562727.1      IPVRR-GN-------------------------------------EVIAVLDIDSPSFDRFDSEDQLG-LEKLVKSLESCLF
NP_693123.1      VPITI-NN-------------------------------------QIYGVLDIDSPINNRFDEEDQYF-LEKFVLHLTKHLQ
YP_001699747.1   IPLMK-QG-------------------------------------EVMGVLDIDSPTVNRFSKEDQAG-LEHFVQTLLLHI-
YP_001343242.1   IPVVV-NG-------------------------------------QLIAVFDMDSPKVGRFSDVDQAG-LESFVATFIEATD
YP_001141806.1   IPVRR-GG-------------------------------------EIIGVLDIDSPIFNRFDEKDRIG-LEETVQILESML-
YP_204661.1      IPFYR-NG-------------------------------------ELAGVLDIDSPKFNRFSEIDEKG-LVDFMKEAEKRL-
```

Supplemental data



```
YP_301135.1       VPLRK-NN-------------------------------------QIIGVLDIDAPITSRFTDVDKNG-LEQIVARIEKQIS
YP_002156075.1    IPFYR-NG-------------------------------------KLAGVLDIDSPKFNRFSEIDEKG-LVDFMKEAEKRL-
NP_244074.1       IPLFQ-NG-------------------------------------VLYGVLDIDSPSLNRFSEEEQAL-LESFVDVLSKNLK
YP_001559126.1    IPMKK-NG-------------------------------------TILGVLDIDSPKFHRFDEVDAEY-LEQIITLLVECCD
YP_080250.1       VPLEV-DG-------------------------------------SIIGVLDIDSPIKNRFDETDEYY-LKKFTEKLGERLS
YP_253122.1       IPIHV-ND-------------------------------------EIIGVLDIDAPITDRFTNEDKEG-LEVIVKVIEKQLT
YP_261333.1       VPLVK-DG-------------------------------------RLIGVLDLDSPKLARFGVADQAG-IEQLAAIFLRLSD
YP_001187407.1    VPLVK-EG-------------------------------------RLIGVLDLDSPSVGRFTEVDQAG-IEALAAIFLAASD
YP_001422246.1    LPIRV-NG-------------------------------------EVIGVLDIDSPVKNRFDETDETY-LAKFVETLEKRLG
YP_001748360.1    IPLVK-EG-------------------------------------RLIGVLDLDSPKLARFSEADQRG-LERLAAVFLELTD
YP_155660.1       APIVV-NG-------------------------------------KTVAVLDVDSPSVGRFSEEDAKA-IQAIADYCQTLNW
NP_744032.1       IPLVK-DG-------------------------------------RLIGVLDLDSPTVGRFSEADQVG-LERLAAIFLALTD
YP_001269146.1    IPLVK-DG-------------------------------------RLIGVLDLDSPTVGRFSEADQVG-LERLAAIFLALTD
YP_236680.1       VPLIK-EG-------------------------------------RLIGVLDLDSPSVGRFNEEDQAG-IERLVAIFLASTD
YP_001667695.1    IPLVK-EG-------------------------------------RLIGVLDLDSPKPGRFGEADQVG-LERLAAIFLELTD
YP_188854.1       IPLHQ-NG-------------------------------------KIIGVLDIDAPIKNRFSQDDQEG-LESIVKTLEKQLN
YP_002275225.1    VPILD-GD-------------------------------------RLIGVLDIDSPIRDRFTPQDRLV-LEQVAALLVPAPT
YP_804779.1       LPLMV-GN-------------------------------------QLLGVLDIDSPILNRFSEDDEAT-LIKFRDALVKHID
NP_470973.1       LPIVK-NN-------------------------------------QLFGVLDIDSPLFNRFDEVDQLW-LEKIRDAITQEIN
YP_607257.1       IPLVK-AG-------------------------------------RLIGVLDLDSPKLARFSEADQAG-LERLAAIFLELTD
YP_001860169.1    IPLQKANG-------------------------------------ELVGVLDIDSPVLERFDDEDRRG-LEELAKIFVASLR
NP_465120.1       LPIVK-NN-------------------------------------QLLGVLDIDSPLFNRFNEVDQLW-LEKIRDAIVQEIN
NP_372242.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
NP_374828.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001247140.1    VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001316940.1    VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001442294.1    VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_002314836.1    VPMMK-NG-------------------------------------VVIGVLDIDSPIKNRFDEIDEQY-LQRFVETLVQHLN
YP_001814701.1    VPLIK-DG-------------------------------------KTIGVLDIDSPELSRFDETDQAG-LEAFSEILLRHI-
NP_646478.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_043764.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_186602.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_494359.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_500333.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001332646.1    VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001575592.1    VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
ZP_02759966.1     VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_001487827.1    VPLKV-NG-------------------------------------EIIGVLDIDSPVKDRFSEVDETY-LIQLTEVLQKALS
YP_417047.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_041183.1       VPIFK-DD-------------------------------------KIIGVLDIDAPITDRFDDNDKEH-LEAIVKIIEKQLA
YP_176251.1       IPIRH-EQ-------------------------------------NVIAVLDIDSPSKGRFHEKEAEA-LERLLKRVEPFLI
YP_002349929.1    LPIVK-NN-------------------------------------QLLGVLDIDSPLFNRFDEVDQLW-LEKIRDAIVQEIN
```

Supplemental data

```
YP_001376537.1   VPMIK-EG-------------------------------------NVIGVLDIDSPEKNRFDKVDQQY-LEKFVETLLKHM-
YP_784647.1      VPLLR-DG-------------------------------------QLLGVWDVDSPTPGRFDDADREG-MEALCQVFLHSLA
YP_557760.1      VPLVAPDG-------------------------------------TLIGVWDVDSPVVARFDEEDAKG-MEALCAVFIQAAL
YP_275769.1      VPLIK-DG-------------------------------------RLIGVLDLDSPSVGRFNEVDQAG-IERLAAIFLASTD
YP_001118914.1   VPLVAADG-------------------------------------TLIGVWDVDSPVAARFDDEDRSG-MEALCRVFVEHAW
YP_001782545.1   VPIIK-EN-------------------------------------KILGVLDLDSIEFGRFTELEEKY-FEKFVQILVENID
NP_882192.1      VPLVH-KG-------------------------------------SLLGVWDVDSPLPGRFDDEDRAG-MEALCEVFLHSLG
YP_001894915.1   VPLVAPDG-------------------------------------TLIGVWDVDSPSVARFDEEDAKG-MEALCAVFIEAAL
YP_001647268.1   VPIMK-EG-------------------------------------NIIGVLDIDSPEKNRFDEVDQHY-LEKFVETLLKHI-
NP_882465.1      VPLLH-KG-------------------------------------SLLGVWDVDSPLPGRFDDEDRAG-MEALCEVFLHSLG
YP_002448219.1   VPIVK-EG-------------------------------------AVIGVLDIDSPEKNRFDEVDQHY-LEKFVETLLKHM-
NP_834359.1      VPIVK-EG-------------------------------------AVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_002369448.1   VPIVK-EG-------------------------------------AVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_191278.1      LPVVA-RG-------------------------------------ALVGVLDIDSPVKDRFSDADRTR-LKEVVSLLVATLE
NP_886654.1      VPLVH-QG-------------------------------------SLLGVWDVDSPLPGRFDDEDRAG-MEALCEVFLHSLG
NP_847103.1      VPIIK-DG-------------------------------------NVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_030799.1      VPIIK-DG-------------------------------------NVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
NP_981083.1      VPIIK-DG-------------------------------------NVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_038702.1      VPIIK-DG-------------------------------------NVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_002340711.1   VPIIK-DG-------------------------------------NVIGVLDIDSPEKKRFDEVDQRY-LEKFVETLLKHM-
YP_085975.1      VPIIK-DG-------------------------------------NVIGVLDIDSPEKNRFDEVDQRY-LEKFVETLLKHM-
YP_002453717.1   VPIIK-DG-------------------------------------NVIGVLDIDSPEKKRFDEVDQRY-LEKFVETLLKHM-
YP_772921.1      VPLVAADG-------------------------------------TLIGVWDVDSPVAARFDDEDRNG-MEALCRVFVEHAW
YP_001843726.1   VPIHQ-GD-------------------------------------KLWGVLDIDSPKLARFDQEDQAG-LERLAPLFLPGA-
YP_942902.1      LPISL-QG-------------------------------------APALLLDIDSPIKNRFTQNDQDG-LAYLVEQIQQHL-
YP_121651.1      VPLVR-DD-------------------------------------TVIGVLDLDSPRPGRFDTVDRDG-LEAVARVFLDAL-
YP_001633478.1   VPLVH-NG-------------------------------------ELIGVWDVDSPEPGRFDEDDRQG-MQALCAVFLASLA
YP_001392189.1   VPIIK-ED-------------------------------------KILGGLDLDSIEFGRFTELEEKY-FKKFVQILVENID
YP_795394.1      VPLTI-GE-------------------------------------TKFGVLDIDSPTKDRFTAADQQV-LEEFARVLLSHID
YP_002284153.1   IPLVK-DG-------------------------------------GIVGVLDLDSPTPHRFDAEDQAG-FEALAAIYVAASA
YP_001255404.1   VPIIK-ED-------------------------------------KVLGVLDLDSIEFGRFTGLEEKY-FKKFVQILVENID
YP_001385170.1   VPIIK-ED-------------------------------------KVLGVLDLDSIEFGRFTGLEEKY-FKKFVQILVENID
YP_001388639.1   VPIIK-ED-------------------------------------KVLGVLDLDSIEFGRFTGLEEKY-FKKFVQILVENID
YP_001347703.1   VPLLK-DG-------------------------------------RLIGVLDLDSPSEGRFSARDQAG-IERLVEIFLRLSD
NP_103622.1      VLLED-DE-------------------------------------GVFGVLDLDSPLPGRFDKADQAG-IETLAAIYAAASS
YP_001807739.1   VPLVAADG-------------------------------------TLIGVWDVDSPVAARFDDEDRNG-MEALCRVFVEHAW
YP_001580337.1   VPLVADDG-------------------------------------TLIGVWDVDSPVAARFDDEDRRG-MEALCRVFVEHAW
YP_002230176.1   VPLVAGDG-------------------------------------TLIGVWDVDSPVAARFDDEDRQG-MEALCRVFVEHAW
NP_251512.1      VPLIK-DG-------------------------------------RLVAVLDLDSPSVGRFSAEDQAG-IEGLVEIFLRLTD
YP_459505.1      VPVVR-DD-------------------------------------TVVAVIDLDAPTKARFARADAEG-IEELARRIAGRI-
YP_790351.1      VPLLK-DG-------------------------------------RLVAVLDLDSPSVGRFSAEDQAG-IEGLVEIFLRLTD
YP_002439846.1   VPLLK-DG-------------------------------------RLVAVLDLDSPSVGRFSAEDQAG-IEGLVEIFLRLTD
```

Supplemental data

```
YP_001970489.1   VPLLR-GD-------------------------------------ELIGVFDIDSPKVGRFDADDQAG-LEAIARVFVEALG
YP_617023.1      VPVIA-ND-------------------------------------RLVGVLDLDSPLPARFTAADQAG-AEALVARIAGALA
YP_620554.1      VPLVADDG-------------------------------------TLIGVWDVDSPVAARFDDEDRKG-MEALCRVFVEHAW
YP_834795.1      VPLVADDG-------------------------------------TLIGVWDVDSPVAARFDDEDRKG-MEALCRVFVEHAW
YP_148653.1      VPLIK-DD-------------------------------------RVIGVLDIDSPVKNRFDDIDRRY-LEQFASVLISA--
YP_442277.1      VPLIARDG-------------------------------------ALIGVWDVDSPLIGRFDAEDATG-MEALCRVFVEVAW
YP_334263.1      VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDATG-MEALCRVFVEVAW
YP_001067073.1   VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDATG-MEALCRVFVEVAW
YP_368503.1      VPLVAADG-------------------------------------TLIGVWDVDSPVAARFDDEDRKG-MEALCRVFVESAW
YP_109010.1      VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDAIG-MEALCRVFVEVAW
YP_993758.1      VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDAIG-MEALCRVFVEVAW
YP_001028775.1   VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDAIG-MEALCRVFVEVAW
YP_002026852.1   VPLLR-GD-------------------------------------ELIGVFDIDSPKVGRFDADDQAG-LEAIARVFVEALG
YP_001764423.1   VPLVADDG-------------------------------------TLIGVWDVDSPVAARFDDEDRTG-MEALCRVFVEHAW
NP_970249.1      VPLIHNG--------------------------------------KLLGVLDVDAPVLNRFDSEDQKG-LQGLVQILISKTV
YP_001059788.1   VPLVACDG-------------------------------------ALIGVWDVDSPLVGRFDAEDATG-MEALCRVFVEVAW
YP_001081316.1   VPLVARDG-------------------------------------ALIGVWDVDSPLVGRFDAEDAIG-MEALCRVFVEVAW
YP_001875000.1   VPIIK-NG-------------------------------------TLFGVLDIDSPIKNRFSDEDRIF-FEKVVGLFSKYSD
YP_278248.1      IPIIV-NS-------------------------------------KLYGVLDLDSPLKSNFKDKEIVLTLEKIATELANKIK
YP_001309672.1   IPIIK-EG-------------------------------------NLIGVLDLDSEEFERFTEVEKTY-LEKAVVILSKYIE
YP_001576818.1   VPIFK-DD-------------------------------------QLWGVFDFDSTKLDNFDETDKKY-LTEIANVIFA---
YP_193206.1      IPIFK-NR-------------------------------------KLWGVFDFDSTKLDDFDDMDKKY-LTRITKMVFA---
YP_199405.1      IPLVK-GD-------------------------------------TLIGVLDLDSPELDRFDADDQRG-LEAIAQLFVDALA
YP_001915608.1   IPLVK-GD-------------------------------------TLIGVLDLDSPELDRFDADDQRG-LEAIAQLFVDALA
YP_002509654.1   IPVEV-DG-------------------------------------DIIGVLDIDSPVKNRFDEIDKRY-LEEFVKKIINLSD
YP_757341.1      VPVRDA-------------------------------------AGALVAVLDIDSTEFGAFDEVDQAG-LEAICGEL-----
YP_810624.1      VPIFS-KN-------------------------------------EFWGVLDLDSTKFNYFDKNDQEY-LEKIGNYIFGI--
YP_244708.1      IPLVR-GD-------------------------------------TLIGVLDLDSPELDRFDADDQRG-LEAIAQVFVGALT
NP_635978.1      IPLVQ-GD-------------------------------------TLIGVLDLDSPELDRFDADDQRG-LEAIAQVFVGALT
XP_505859.1      VPIV-LNG-------------------------------------KLFGVIDIDCASLEGFDEVDQQG-LEKIALILAKSCE
NP_964546.1      VPIFKANG-------------------------------------DLWGVFDFDSTKLANFDELDQKY-LEKISKVFKF---
YP_001261873.1   VPIVH-EG-------------------------------------RLIGVLDLDSPIKGRFDEQDVVG-CAELMAVLGPRIA
YP_814311.1      IPIFKKNG-------------------------------------ELWGVFDFDSTKIANFDELDQKY-LEAISNIFKF---
NP_298520.1      VPLLK-DD-------------------------------------TLIGVLDLDSPTCARFDSDDQTG-IESLARVFVSSLR
YP_001775202.1   VPLLK-DD-------------------------------------TLIGVLDLDSPTCARFDSDDQTG-IESLARVFVSSLR
YP_001829224.1   VPLLK-DN-------------------------------------TLIGVLDLDSPTCARFDSDDQTG-IESLARVFVSNLR
NP_778732.1      VPLLK-DN-------------------------------------TLIGVLDLDSPTCARFDSDDQTG-IESLARVFVSNLR
NP_246070.1      VPIFQ-GG-------------------------------------EMIAELDVDSPRLARFSAEDVAF-LENCAALICIKE-
XP_645637.1      LPLYK--------------------------------------QGNIIGVLDVDSDKLNSFDEIDEKY-LTQILKLLDN---
XP_456263.1      VPIV-QNG-------------------------------------ETVGVIDIDCLDYNGFTKLDQEF-LEKLAASVSKTCV
YP_001300883.1   VPILA--------------------------------------NGQVKGVLDIDSNLLNFFDETDRRY-LEEVTAIVAKT--
YP_812434.1      VPVFNRDG-------------------------------------SLWGVYDFDSTNLDNFDDLDQKY-LEQIAASVLGN--
```

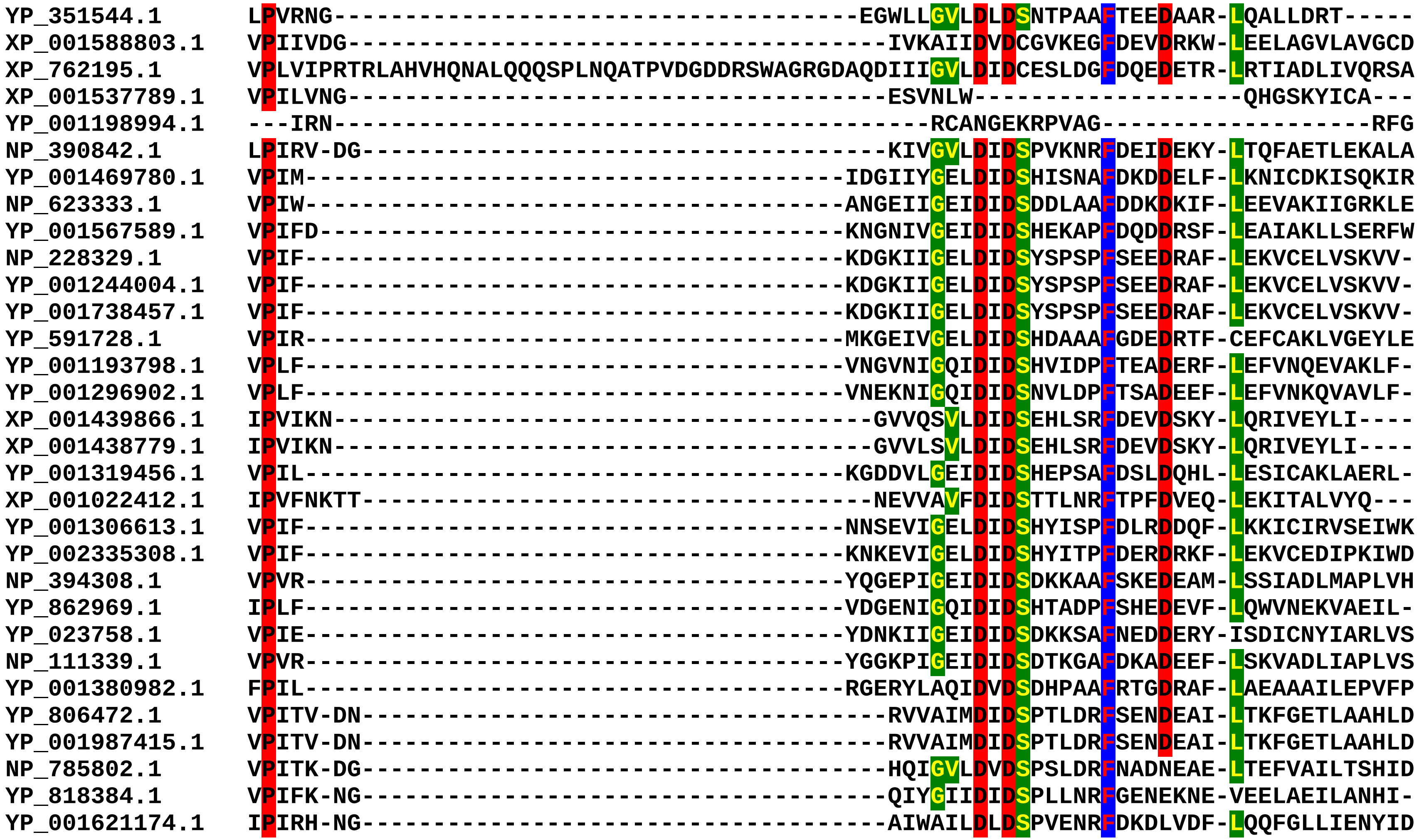
Supplemental data

```
NP_814796.1       VPMVKNHQ--------------------------------------LVGVLDIDSGVTNSYDAVDQLY-LEKFVTLLLEKSD
NP_358362.1       VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTT
YP_816246.1       VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTT
YP_001694316.1    VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTT
YP_001835526.1    VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTT
YP_002037500.1    VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTT
NP_345351.1       VPMMKNGQ--------------------------------------LLGVLDLDSSEIEDYDAMDRDY-LEQFVAILLEKTA
XP_001385863.2    VPLV-KEG-------------------------------------KILAVLDIDCLALDGFDESDKKF-LEELAVLIVRTCE
NP_985662.1       VPIL-QDN-------------------------------------RTVAIMDLDCSEHNGFTDLDKKY-LEKLAQEISRTCK
YP_759928.1       LPVFRA-------------------------------------DGSLAAVLDVDSTQFDAFDAVDQAG-LTAICESL-----
YP_002123053.1    VPMLKAGR--------------------------------------LVGVLDLDSSLVADYDQIDQDY-LEQFVALLLENTS
YP_002131343.1    VPVLDA-------------------------------------GGELIAVFDVDSETKAAFDSTDREW-LERILSET-----
XP_001683506.1    VPVLIENG-------------------------------------SVVAVIDVDSTELGHFSEEDREG-LERMAAVLAQHLD
NP_012854.1       VPIISNDG-------------------------------------KTLGVIDIDCLDYEGFDHVDKEF-LEKLAKLINKSCV
XP_460154.1       VPIV-NKG-------------------------------------QVVGVLDIDCLTLEGFDDIDVKY-LEQLATLIGNTCD
XP_001801028.1    VPIVQGG--------------------------------------RTVAIIDVDCAELSGFDGEDQRA-LEELACMLGEACD
YP_001201197.1    VPMVKNDQ--------------------------------------LLGVLDLDSRLVADYDTIDQDY-LEKFVALLVEKTD
XP_001484046.1    VPIV-YEN-------------------------------------KVVGVLDIDCLDLNGFDKVDQKY-LERLVEMISSSCK
NP_274882.1       VPLFS--------------------------------------DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
YP_001302571.1    VPVIW--------------------------------------ADKVIAVLDIDSDQPDSFDETDQIF-LEKIVELLPHQ--
YP_002000650.1    VPLFS--------------------------------------DGRCIGVLDADSEHLAQFDEADALY-LGELAKILERR--
YP_001682978.1    VPVFDA-------------------------------------AGTLLAVFDVDSDQPAAFDAVDQHG-IEIILKAT-----
YP_002342048.1    VPLFS--------------------------------------DGRCIGVLDADSEYLAQFDETDALY-LGELAKILEKR--
YP_207199.1       VPLFS--------------------------------------DGRCIGVLDADSEHLAQFDEADALY-LGELAKILERR--
XP_844799.1       VPIKSAAG-------------------------------------EVVGIIDVDSTQLNFFDDVDAQG-LQEVAEIVARHVD
YP_001598496.1    VPLFS--------------------------------------DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
YP_974447.1       VPLFS--------------------------------------DGRCIGVLDADSEHLAQFDETDALY-LGELAKILEKR--
XP_001465858.1    VPVLIENG-------------------------------------SVVAVIDVDSTELDHFSEEDKEG-LERMAAVLAQHLD
YP_001280815.1    VPITDE-------------------------------------AGEIIAVLDIDSDTLATFDEVDGEY-LQQLVDLLK----
NP_664843.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
NP_802083.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_280574.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_001128311.1    VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_002286076.1    VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
XP_390238.1       VPILVAQDDGT---------------------------------KKLVAIIDIDCAELNGFDEVDKAH-LEQLAALLATSCD
XP_001565241.1    VPVLIENG-------------------------------------SVVAVIDVDSTELGYFSEEDKEG-LERVAAVLEQHLN
YP_596891.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_600772.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_001351938.1    VPYFDR-------------------------------------NGEVAGVLDVDSEYPDAFSAVDQLW-LEKIAALLS----
XP_001910908.1    VPIMVEGK-------------------------------------GVVGIIDIDCTVENGFDQIDKEY-LEKLADFLAKSSD
XP_446236.1       VPIVSQSG-------------------------------------KTLGVIDLDCLDFEGFTDVDKEY-LEKLALAITKSCD
NP_607473.1       VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LETFVDLLLEKTT
```

Supplemental data

```
YP_060405.1      VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LETFVDLLLEKTT
YP_602707.1      VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LETFVDLLLEKTT
NP_269471.1      VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_282476.1      VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_598792.1      VPMVKEGH--------------------------------------LIGVLDLDSSLVADYDEVDQEY-LEAFVDLFLEKTT
YP_099090.1      VPLIK--------------------------------------DGEVWGVLDIDSDLLNFFDETDRKY-LEEMCGYLSK---
NP_421082.1      VPVFDP-------------------------------------DGRLIAVFDVDSDKPAAFDEIDQAW-LEKILKAT-----
YP_001034962.1   VPMVKVGR--------------------------------------LLGVLDLDSSLTGDYDKVDQEY-LEKFVQILLEKTT
YP_211509.1      VPLIK--------------------------------------DGEVWGVLDIDSDLLNFFDETDRKY-LEEMCGYLSK---
XP_001940256.1   VPVVQDG--------------------------------------KVVAIIDIDCAELNGFTEEDQEA-LEELAELLSQSCD
YP_001450196.1   VPMVKDGR--------------------------------------LLGVLDLDSALTGDYDKIDQEY-LEKFVQILLEKTI
XP_680937.1      VPILAGG--------------------------------------ETVAIIDIDCAEPSGFDDEDKEY-LEELAALLSESCD
XP_001265696.1   VPILVGG--------------------------------------ETVAIIDVDCTEPAGFDDEDKKW-LEELASLLSECCD
XP_367405.1      VPVLDSEG-------------------------------------QLRAIIDVDCALPDGFDEVDRVW-LESLAALVGRSCD
YP_001033719.1   VPMVKDGN--------------------------------------LVGVLDLDSSEVGFYDEIDQKY-LEEFASILCDMTD
YP_812051.1      VPMVKDGN--------------------------------------LVGVLDLDSSEVGFYDEIDQKY-LEEFASILCDMTD
NP_810723.1      VPIIK--------------------------------------EGNVIGVLDIDSDTPDSFDETDARY-LEEICTYIG----
YP_141537.1      VPMEKGGK--------------------------------------LVGVLDLDSHQVEDYNQVDQDY-LEAFVKILLEKTD
NP_721920.1      IPMVKDGQ--------------------------------------LIGVLDLDSHLVGDYDVIDQGY-LEQFVTILLEKTN
YP_820528.1      VPMEKGGK--------------------------------------LVGVLDLDSHQVEDYNQVDQDY-LEAFVKILLEKTD
XP_001246405.1   VPIVVDG--------------------------------------ETLALIDIDCTQPSGFDEIDKKG-LEHLAKLLAENCD
YP_612063.1      LPVRNR-------------------------------------AGEVIGVLDIDSNQPAAFDRDDQDA-LETLLHAV-----
YP_444835.1      VPVL----------TP---------------------------DDALLAVLDVDSDTLGAFDPTDREH-LEALCRDLG----
NP_943356.1      LPVRI-NN-------------------------------------TVVGALDIDSPAFSRFSSEDEL---------------
YP_631535.1      VPVFGR-------------------------------------NRELLAVLDIDSEHKNAFDEVDRRE-LEELVRWF-----
NP_687842.1      VPMFKNGK--------------------------------------LLGVLDLDSSLVADYDEIDQEY-LEKFVGILVEHTI
YP_329573.1      VPMFKNGK--------------------------------------LLGVLDLDSSLVADYDEIDQEY-LEKFVGILVEHTI
XP_001692446.1   VPVVRSHISFNDGRTPSI-------------------------PEKLVAVLDVDSDHPAAFTEVDQEC-LEELCEWLGDHLQ
NP_390842.2      LPIRV-DG-------------------------------------KIVGVLDIDSPVKNRFDEIDEKY-LTQFAETLEKALA
YP_511991.1      LPVRDA-------------------------------------SGAIIAVLDIDSNQPDAFTQDDATA-LESILRAT-----
XP_719233.1      VPIV-KNG-------------------------------------ETKGVIDLDCLDLEGFDQVDQKY-LEKLAELIAQSL-
NP_735295.1      VPMFKNGK--------------------------------------LLGVLDLDSSLVADYDEIDQEY-LEKFVGILVEHTI
YP_168601.1      LPVFGR-------------------------------------TGALIAVFDIDSDLPDAFTVADAEA-LGAILEQV-----
YP_001169098.1   LPVRNG-------------------------------------EGRLLGVLDLDSNTPAAFTEEDARR-LQALLDRT-----
XP_001216799.1   VPILAGG--------------------------------------ETVAIIDIDCTEPDGFDETDKKY-LEELAALLADSCD
YP_437091.1      LPLFNA-------------------------------------NGDVAAVLDIDSDHIDAFDAVDERF-LAELLTDL-----
XP_001390389.1   VPILSKG--------------------------------------ETVAIIDIDCAEPDGFDDEDKKY-LEELAALLAECCD
XP_001550500.1   VPIVVDG--------------------------------------VVKAIIDVDCGVKEGFDEVDRKW-LEELAGVLAEGCD
XP_001273080.1   VPVLVGG--------------------------------------ETVAIIDVDCTEPSGFDEVDRKY-LEELAALLAECCD
YP_001042044.1   LPVRNG-------------------------------------EGRLLGVLDLDSNTPAAFTEEDAAR-LQALLDRT-----
YP_916587.1      IPVIGG-------------------------------------GERLIGVLDIDSDQPDAFTRDDADR-LQEILDAT-----
XP_961529.1      VPIVVDD--------------------------------------RVVAIIDVDCAELNGFDEVDKVW-LEKLADLVGRSCD
```

Supplemental data



```
YP_351544.1        LPVRNG-------------------------------------EGWLLGVLDLDSNTPAAFTEEDAAR-LQALLDRT-----
XP_001588803.1     VPIIVDG--------------------------------------IVKAIIDVDCGVKEGFDEVDRKW-LEELAGVLAVGCD
XP_762195.1        VPLVIPRTRLAHVHQNALQQQSPLNQATPVDGDDRSWAGRGDAQDIIIGVLDIDCESLDGFDQEDETR-LRTIADLIVQRSA
XP_001537789.1     VPILVNG--------------------------------------ESVNLW-------------------QHGSKYICA---
YP_001198994.1     ---IRN------------------------------------------RCANGEKRPVAG-------------------RFG
NP_390842.1        LPIRV-DG-------------------------------------KIVGVLDIDSPVKNRFDEIDEKY-LTQFAETLEKALA
YP_001469780.1     VPIM--------------------------------------IDGIIYGELDIDSHISNAFDKDDELF-LKNICDKISQKIR
NP_623333.1        VPIW--------------------------------------ANGEIIGEIDIDSDDLAAFDDKDKIF-LEEVAKIIGRKLE
YP_001567589.1     VPIFD-------------------------------------KNGNIVGEIDIDSHEKAPFDQDDRSF-LEAIAKLLSERFW
NP_228329.1        VPIF--------------------------------------KDGKIIGELDIDSYSPSPFSEEDRAF-LEKVCELVSKVV-
YP_001244004.1     VPIF--------------------------------------KDGKIIGELDIDSYSPSPFSEEDRAF-LEKVCELVSKVV-
YP_001738457.1     VPIF--------------------------------------KDGKIIGELDIDSYSPSPFSEEDRAF-LEKVCELVSKVV-
YP_591728.1        VPIR--------------------------------------MKGEIVGELDIDSHDAAAFGDEDRTF-CEFCAKLVGEYLE
YP_001193798.1     VPLF--------------------------------------VNGVNIGQIDIDSHVIDPFTEADERF-LEFVNQEVAKLF-
YP_001296902.1     VPLF--------------------------------------VNEKNIGQIDIDSNVLDPFTSADEEF-LEFVNKQVAVLF-
XP_001439866.1     IPVIKN--------------------------------------GVVQSVLDIDSEHLSRFDEVDSKY-LQRIVEYLI----
XP_001438779.1     IPVIKN--------------------------------------GVVLSVLDIDSEHLSRFDEVDSKY-LQRIVEYLI----
YP_001319456.1     VPIL--------------------------------------KGDDVLGEIDIDSHEPSAFDSLDQHL-LESICAKLAERL-
XP_001022412.1     IPVFNKTT------------------------------------NEVVAVFDIDSTTLNRFTPFDVEQ-LEKITALVYQ---
YP_001306613.1     VPIF--------------------------------------NNSEVIGELDIDSHYISPFDLRDDQF-LKKICIRVSEIWK
YP_002335308.1     VPIF--------------------------------------KNKEVIGELDIDSHYITPFDERDRKF-LEKVCEDIPKIWD
NP_394308.1        VPVR--------------------------------------YQGEPIGEIDIDSDKKAAFSKEDEAM-LSSIADLMAPLVH
YP_862969.1        IPLF--------------------------------------VDGENIGQIDIDSHTADPFSHEDEVF-LQWVNEKVAEIL-
YP_023758.1        VPIE--------------------------------------YDNKIIGEIDIDSDKKSAFNEDDERY-ISDICNYIARLVS
NP_111339.1        VPVR--------------------------------------YGGKPIGEIDIDSDTKGAFDKADEEF-LSKVADLIAPLVS
YP_001380982.1     FPIL--------------------------------------RGERYLAQIDVDSDHPAAFRTGDRAF-LAEAAAILEPVFP
YP_806472.1        VPITV-DN-------------------------------------RVVAIMDIDSPTLDRFSENDEAI-LTKFGETLAAHLD
YP_001987415.1     VPITV-DN-------------------------------------RVVAIMDIDSPTLDRFSENDEAI-LTKFGETLAAHLD
NP_785802.1        VPITK-DG-------------------------------------HQIGVLDVDSPSLDRFNADNEAE-LTEFVAILTSHID
YP_818384.1        VPIFK-NG-------------------------------------QIYGIIDIDSPLLNRFGENEKNE-VEELAEILANHI-
YP_001621174.1     IPIRH-NG-------------------------------------AIWAILDLDSPVENRFDKDLVDF-LQQFGLLIENYID
```

FIGURE SI: Sequence alignment of all putative fRMsr proteins.
The sequences of 358 putative fRMsrs were generated by PEDANT database (http://pedant.gsf.de/) and aligned with Bioedit software. The numbering of amino acid residues is based on the *E. coli* fRMsr sequence. For reason of clarity, only the alignment between residues 19 and 170 is represented. The residues conserved at least at 95 % are indicated in black on red boxes (such as Y66, C94, C118, E125, D141 and D143), those conserved at least at 90% in red on blue boxes (such as F78, and C84), and finally, in yellow on green boxes (for example W62) those conserved at 80%.

Supplemental data

The name of each sequence correspond to NCBI reference:

C84/C94/C118 signature (333 sequences, 93%) :
NP_288269.1: *Escherichia coli O157:H7 EDL933*; YP_002403065.1: *Escherichia coli 55989*; YP_002397981.1 : *Escherichia coli ED1a* ; YP_002412850.1: *Escherichia coli UMN026*; YP_001880631.1: *Shigella boydii CDC 3083-94*; YP_002391608.1: *Escherichia coli S88*; YP_001730808.1: *Escherichia coli str. K12 substr. DH10B*; ZP_03003290.1 : *Escherichia coli 53638*; YP_001743413.1: *Escherichia coli SMS-3-5*; YP_002407226.1 : *Escherichia coli IAI39*; YP_002329476.1: *Escherichia coli O127:H6 str. E2348/69*; YP_002382399.1: *Escherichia fergusonii ATCC 35469*; ZP_03445135.1*: Escherichia coli O157:H7 str. TW14588*; YP_688895.1: *Shigella flexneri 5 str. 8401*; YP_669680.1: *Escherichia coli 536*; NP_416346.4: *Escherichia coli str.K-12 substr. MG1655*; ZP_02903917.1: *Escherichia albertii TW07627*; ZP_03282954.1: *Enterobacter cancerogenus ATCC 35316*; YP_001452718.1: *Citrobacter koseri ATCC BAA-895*; YP_001570142.1: *Salmonella enterica subsp. arizonae serovar 62:z4,z23*; NP_460803.1: *Salmonella typhimurium LT2*; YP_150310.1: *Salmonella enterica subsp. enterica serovar Paratyphi A str. ATCC 9150*; YP_002141797.1: *Salmonella enterica subsp. enterica serovar Paratyphi A str. AKU_12601*; YP_002243298.1: *Salmonella enterica subsp. enterica serovar Enteritidis str. P125109*; NP_456352.1: *Salmonella enterica subsp. enterica serovar Typhi str. CT18*; NP_804849.1: *Salmonella enterica subsp. enterica serovar Typhi str. Ty2*; YP_001587603.1*: Salmonella enterica subsp. enterica serovar Paratyphi B str. SPB7*; ZP_02666379.2: *Salmonella enterica subsp. enterica serovar Heidelberg str. SL486*; ZP_02832604.2: *Salmonella enterica subsp. enterica serovar Weltevreden str. HI_N05-537*; YP_001177122.1: *Enterobacter sp.* 638; YP_002226291.1: *Salmonella enterica subsp. enterica serovar Gallinarum str. 287/91*; YP_216830.1: *Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67*; NP_754135.1: *Escherichia coli CFT073*; ZP_02347214.1: *Salmonella enterica subsp. enterica serovar Saintpaul str. SARA29*; ZP_02574893.1: *Salmonella enterica subsp. enterica serovar 4,- str. CVM23701*; ZP_02659310.1: *Salmonella enterica subsp. enterica serovar Kentucky str. CDC 191*; ZP_02663512.1: *Salmonella enterica subsp. enterica serovar Schwarzengrund str. SL480*; ZP_02681881.1: *Salmonella enterica subsp. enterica serovar Hadar str. RI_05P066*; ZP_02696584.1: *Salmonella enterica subsp. enterica serovar Newport str. SL317*; YP_002041101.1: *Salmonella enterica subsp. enterica serovar Newport str. SL254*; YP_002045892.1: *Salmonella enterica subsp. enterica serovar Heidelberg str. SL476*; ZP_03075263.1: *Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188*; YP_002114879.1: *Salmonella enterica subsp. enterica serovar Schwarzengrund str. CVM19633*; YP_002146178.1: *Salmonella enterica subsp. enterica serovar Agona str. SL483*; ZP_03165709.1: *Salmonella enterica subsp. enterica serovar Saintpaul str. SARA23*; YP_002215294.1: *Salmonella enterica subsp. enterica serovar Dublin str. CT_02021853*; ZP_03221812.1: *Salmonella enterica subsp. enterica serovar Javiana str. GA_MM04042433*; ZP_03342093.1: *Salmonella enterica subsp. enterica serovar Typhi str. 404ty*; ZP_03347154.1*: Salmonella enterica subsp. enterica serovar Typhi str. E00-7866*; ZP_03352578.1: *Salmonella enterica subsp. enterica serovar Typhi str. E01-6750*; ZP_03357060.1: *Salmonella enterica subsp. enterica serovar Typhi str. E02-1180*; ZP_03371453.1: *Salmonella enterica subsp. enterica serovar Typhi str. E98-2068*; ZP_03379701.1: *Salmonella enterica subsp. enterica serovar Typhi str. J185*; ZP_03383790.1: *Salmonella enterica subsp. enterica serovar Typhi str. M223*; ZP_03412299.1: *Salmonella enterica subsp. enterica serovar Typhi str. E98-3139*; ZP_03217260.1: *Salmonella enterica subsp. enterica serovar Virchow str. SL491*; YP_002237781.1: *Klebsiella pneumoniae 342;* YP_001336004.1: *Klebsiella pneumoniae subsp. pneumoniae MGH 78578*; YP_001437512.1: *Enterobacter sakazakii ATCC BAA-894*; ZP_03403067.1: *Salmonella enterica subsp. enterica serovar Typhi str. AG3*; YP_001907453.1: Erwinia tasmaniensis Et1/99; YP_001006127.1: Yersinia enterocolitica subsp. enterocolitica 8081; NP_669181.1: Yersinia pestis KIM; NP_993079.1: Yersinia pestis biovar Microtus str. 91001; YP_070893.1: Yersinia pseudotuberculosis IP 32953; YP_651729.1: Yersinia pestis Antiqua; YP_647856.1: Yersinia pestis Nepal516; YP_001163103.1: Yersinia pestis Pestoides F; ZP_01888376.1: Yersinia pestis CA88-4125; YP_001400647.1: Yersinia pseudotuberculosis IP 31758; YP_001607069.1: Yersinia pestis Angola; ZP_02221459.1: Yersinia pestis

biovar Orientalis str. F1991016; ZP_02227087.1: Yersinia pestis biovar Orientalis str. IP275; ZP_02229339.1: Yersinia pestis biovar Antiqua str. E1979001; ZP_02237158.1: Yersinia pestis biovar Antiqua str. B42003004; ZP_02307206.1: Yersinia pestis biovar Antiqua str. UG05-0454; ZP_02314388.1: Yersinia pestis biovar Orientalis str. MG05-1020; ZP_02318544.1: Yersinia pestis biovar Mediaevalis str. K1973002; YP_001720525.1: Yersinia pseudotuberculosis YPIII; YP_001872871.1: Yersinia pseudotuberculosis PB1/+; YP_002346706.1: Yersinia pestis CO92; YP_001478346.1: Serratia proteamaculans 568; YP_454994.1: Sodalis glossinidius str. 'morsitans'; YP_001217052.1: Vibrio cholerae O395; YP_002150764.1: Proteus mirabilis HI4320; YP_856816.1: Aeromonas hydrophila subsp. hydrophila ATCC 7966; YP_001093925.1: Shewanella loihica PV-4; YP_269525.1: Colwellia psychrerythraea 34H; NP_718189.1: Shewanella oneidensis MR-1; YP_001366642.1: Shewanella baltica OS185; YP_001554990.1: Shewanella baltica OS195; YP_737796.1: Shewanella sp. MR-7; YP_869415.1: Shewanella sp. ANA-3; NP_797987.1: Vibrio parahaemolyticus RIMD 2210633; YP_001501790.1: Shewanella pealeana ATCC 700345; YP_927460.1: Shewanella amazonensis SB2B; YP_001474222.1: Shewanella sediminis HAW-EB3; YP_963195.1: Shewanella sp. W3-18-1; YP_001183724.1: Shewanella putrefaciens CN-32; YP_733804.1: Shewanella sp. MR-4; YP_001050811.1: Shewanella baltica OS155; YP_750914.1: Shewanella frigidimarina NCIMB 400; YP_002126297.1: Alteromonas macleodii 'Deep ecotype'; YP_130129.1: Photobacterium profundum SS9; YP_001445537.1: Vibrio harveyi ATCC BAA-1116; YP_661562.1: Pseudoalteromonas atlantica T6c; YP_001674584.1: Shewanella halifaxensis HAW-EB4; YP_002417027.1: Vibrio splendidus LGP32; NP_934438.1: Vibrio vulnificus YJ016; NP_761466.1: Vibrio vulnificus CMCP6; YP_001760510.1: Shewanella woodyi ATCC 51908; YP_002312133.1: Shewanella piezotolerans WP3; YP_562727.1: Shewanella denitrificans OS217; NP_693123.1: Oceanobacillus iheyensis HTE831; YP_001699747.1: Lysinibacillus sphaericus C3-41; YP_001343242.1: Marinomonas sp. MWYL1; YP_001141806.1: Aeromonas salmonicida subsp. salmonicida A449; YP_204661.1: Vibrio fischeri ES114; YP_301135.1: Staphylococcus saprophyticus subsp. saprophyticus ATCC 15305; YP_002156075.1: Vibrio fischeri MJ11; NP_244074.1: Bacillus halodurans C-125; YP_001559126.1: Clostridium phytofermentans ISDg; YP_080250.1: Bacillus licheniformis ATCC 14580; YP_253122.1: Staphylococcus haemolyticus JCSC1435; YP_261333.1: Pseudomonas fluorescens Pf-5; YP_001187407.1: Pseudomonas mendocina ymp; YP_001422246.1: Bacillus amyloliquefaciens FZB42; YP_001748360.1: Pseudomonas putida W619; YP_155660.1: Idiomarina loihiensis L2TR; NP_744032.1: Pseudomonas putida KT2440; YP_001269146.1: Pseudomonas putida F1; YP_236680.1: Pseudomonas syringae pv. syringae B728a; YP_001667695.1: Pseudomonas putida GB-1; YP_188854.1: Staphylococcus epidermidis RP62A; YP_002275225.1: Gluconacetobacter diazotrophicus PAl 5; YP_804779.1: Pediococcus pentosaceus ATCC 25745; NP_470973.1: Listeria innocua Clip11262; YP_607257.1: Pseudomonas entomophila L48; YP_001860169.1: Burkholderia phymatum STM815; NP_465120.1: Listeria monocytogenes EGD-e; NP_372242.1: Staphylococcus aureus subsp. aureus Mu50; NP_374828.1: Staphylococcus aureus subsp. aureus N315; YP_001247140.1: Staphylococcus aureus subsp. aureus JH9; YP_001316940.1: Staphylococcus aureus subsp. aureus JH1; YP_001442294.1: Staphylococcus aureus subsp. aureus Mu3; YP_002314836.1: Anoxybacillus flavithermus WK1; YP_001814701.1: Exiguobacterium sibiricum 255-15; NP_646478.1: Staphylococcus aureus subsp. aureus MW2; YP_043764.1: Staphylococcus aureus subsp. aureus MSSA476; YP_186602.1: Staphylococcus aureus subsp. aureus COL; YP_494359.1: Staphylococcus aureus subsp. aureus USA300; YP_500333.1: Staphylococcus aureus subsp. aureus NCTC 8325; YP_001332646.1: Staphylococcus aureus subsp. aureus str. Newman; YP_001575592.1 : Staphylococcus aureus subsp. aureus USA300_TCH1516; ZP_02759966.1: Staphylococcus aureus subsp. aureus USA300_TCH1516; YP_001487827.1: Bacillus pumilus SAFR-032; YP_417047.1: Staphylococcus aureus RF122; YP_041183.1: Staphylococcus aureus subsp. aureus MRSA252; YP_176251.1: Bacillus clausii KSM-K16; YP_002349929.1: Listeria monocytogenes HCC23; YP_001376537.1: Bacillus cereus subsp. cytotoxis NVH 391-98; YP_784647.1: Bordetella avium 197N; YP_557760.1: Burkholderia xenovorans LB400; YP_275769.1: Pseudomonas syringae pv. phaseolicola 1448A; YP_001118914.1: Burkholderia vietnamiensis G4; YP_001782545.1: Clostridium botulinum B1 str. Okra; NP_882192.1: Bordetella pertussis Tohama I; YP_001894915.1: Burkholderia phytofirmans PsJN; YP_001647268.1: Bacillus weihenstephanensis KBAB4; NP_882465.1: Bordetella parapertussis 12822; YP_002448219.1: Bacillus cereus G9842; NP_834359.1: Bacillus cereus ATCC 14579; YP_002369448.1:

Bacillus cereus B4264; YP_191278.1: Gluconobacter oxydans 621H; NP_886654.1: Bordetella bronchiseptica RB50; NP_847103.1: Bacillus anthracis str. Ames; YP_030799.1: Bacillus anthracis str. Sterne; NP_981083.1: Bacillus cereus ATCC 10987; YP_038702.1: Bacillus thuringiensis serovar konkukian str. 97-27; YP_002340711.1: Bacillus cereus AH187; YP_085975.1: Bacillus cereus E33L; YP_002453717.1: Bacillus cereus AH820; YP_772921.1: Burkholderia ambifaria AMMD; YP_001843726.1: Lactobacillus fermentum IFO 3956; YP_942902.1: Psychromonas ingrahamii 37; YP_121651.1: Nocardia farcinica IFM 10152; YP_001633478.1: Bordetella petrii DSM 12804; YP_001392189.1: Clostridium botulinum F str. Langeland; YP_795394.1: Lactobacillus brevis ATCC 367; YP_002284153.1: Rhizobium leguminosarum bv. trifolii WSM2304; YP_001255404.1: Clostridium botulinum A str. ATCC 3502; YP_001385170.1: Clostridium botulinum A str. ATCC 19397; YP_001388639.1: Clostridium botulinum A str. Hall; YP_001347703.1: Pseudomonas aeruginosa PA7; NP_103622.1: Mesorhizobium loti MAFF303099; YP_001807739.1: Burkholderia ambifaria MC40-6; YP_001580337.1: Burkholderia multivorans ATCC 17616; YP_002230176.1: Burkholderia cenocepacia J2315; NP_251512.1: Pseudomonas aeruginosa PAO1; YP_459505.1: Erythrobacter litoralis HTCC2594; YP_790351.1: Pseudomonas aeruginosa UCBPP-PA14; YP_002439846.1: Pseudomonas aeruginosa LESB58; YP_001970489.1: Stenotrophomonas maltophilia K279a; YP_617023.1: Sphingopyxis alaskensis RB2256; YP_620554.1: Burkholderia cenocepacia AU 1054; YP_834795.1: Burkholderia cenocepacia HI2424; YP_148653.1: Geobacillus kaustophilus HTA426; YP_442277.1: Burkholderia thailandensis E264; YP_334263.1: Burkholderia pseudomallei 1710b; YP_001067073.1: Burkholderia pseudomallei 1106a; YP_368503.1: Burkholderia sp. 383; YP_109010.1: Burkholderia pseudomallei K96243; YP_993758.1: Burkholderia mallei SAVP1; YP_001028775.1: Burkholderia mallei NCTC 10229; YP_002026852.1: Stenotrophomonas maltophilia R551-3; YP_001764423.1: Burkholderia cenocepacia MC0-3; NP_970249.1: Bdellovibrio bacteriovorus HD100; YP_001059788.1: Burkholderia pseudomallei 668; YP_001081316.1: Burkholderia mallei NCTC 10247; YP_001875000.1: Elusimicrobium minutum Pei191; YP_278248.1: Mycoplasma synoviae 53; YP_001309672.1: Clostridium beijerinckii NCIMB 8052; YP_001576818.1: Lactobacillus helveticus DPC 4571; YP_193206.1; Lactobacillus acidophilus NCFM; YP_199405.1: Xanthomonas oryzae pv. oryzae KACC10331; YP_001915608.1: Xanthomonas oryzae pv. oryzae PXO99A; YP_002509654.1: Halothermothrix orenii H 168; YP_757341.1: Maricaulis maris MCS10; YP_810624.1: Oenococcus oeni PSU-1; YP_244708.1: Xanthomonas campestris pv. campestris str. 8004; NP_635978.1: Xanthomonas campestris pv. campestris str. ATCC 33913; XP_505859.1: Yarrowia lipolytica; NP_964546.1: Lactobacillus johnsonii NCC 533; YP_001261873.1: Sphingomonas wittichii RW1; YP_814311.1: Lactobacillus gasseri ATCC 33323; NP_298520.1: Xylella fastidiosa 9a5c; YP_001775202.1: Xylella fastidiosa M12; YP_001829224.1: Xylella fastidiosa M23; NP_778732.1: Xylella fastidiosa Temecula1; NP_246070.1: Pasteurella multocida subsp. multocida str. Pm70; YP_001997158.1: Chloroherpeton thalassium ATCC 35110; XP_645637.1: Dictyostelium discoideum AX4; XP_456263.1: Kluyveromyces lactis; YP_001300883.1: Bacteroides vulgatus ATCC 8482; YP_812434.1: Lactobacillus delbrueckii subsp. bulgaricus ATCC BAA-365; NP_814796.1: Enterococcus faecalis V583; NP_358362.1: Streptococcus pneumoniae R6; YP_816246.1: Streptococcus pneumoniae D39; YP_001694316.1: Streptococcus pneumoniae Hungary19A-6; YP_001835526.1: Streptococcus pneumoniae CGSP14; YP_002037500.1: Streptococcus pneumoniae G54; NP_345351.1: Streptococcus pneumoniae TIGR4; XP_001385863.2: Pichia stipitis CBS 6054; NP_985662.1: Ashbya gossypii ATCC 10895; YP_759928.1: Hyphomonas neptunium ATCC 15444; YP_002123053.1: Streptococcus equi subsp. zooepidemicus MGCS10565; YP_002131343.1: Phenylobacterium zucineum HLK1; XP_001683506.1: Leishmania major strain Friedlin; NP_012854.1: Saccharomyces cerevisiae; XP_460154.1: Debaryomyces hansenii CBS767; XP_001801028.1: Phaeosphaeria nodorum SN15; YP_001201197.1: Streptococcus suis 98HAH33; XP_001484046.1: Pichia guilliermondii ATCC 6260; NP_274882.1: Neisseria meningitidis MC58; YP_001302571.1: Parabacteroides distasonis ATCC 8503; YP_002000650.1: Neisseria gonorrhoeae NCCP11945; YP_001682978.1: Caulobacter sp. K31; YP_002342048.1: Neisseria meningitidis Z2491; YP_207199.1: Neisseria gonorrhoeae FA 1090; XP_844799.1: Trypanosoma brucei TREU927; YP_001598496.1: Neisseria meningitidis 053442; YP_974447.1: Neisseria meningitidis FAM18; XP_001465858.1: Leishmania infantum JPCM5; YP_001280815.1: Psychrobacter sp. PRwf-1; NP_664843.1: Streptococcus pyogenes MGAS315; NP_802083.1: Streptococcus pyogenes SSI-1;

Supplemental data

YP_280574.1: Streptococcus pyogenes MGAS6180; YP_001128311.1: Streptococcus pyogenes str. Manfredo; YP_002286076.1: Streptococcus pyogenes NZ131; XP_390238.1: Gibberella zeae PH-1; XP_001565241.1: Leishmania braziliensis MHOM/BR/75/M2904; YP_596891.1: Streptococcus pyogenes MGAS9429; YP_600772.1: Streptococcus pyogenes MGAS2096; YP_001351938.1: Janthinobacterium sp. Marseille; XP_001910908.1: Podospora anserina DSM 980; XP_446236.1: Candida glabrata CBS 138; NP_607473.1: Streptococcus pyogenes MGAS8232; YP_060405.1: Streptococcus pyogenes MGAS10394; YP_602707.1: Streptococcus pyogenes MGAS10750; NP_269471.1: Streptococcus pyogenes M1 GAS; YP_282476.1: Streptococcus pyogenes MGAS5005; YP_598792.1: Streptococcus pyogenes MGAS10270; YP_099090.1: Bacteroides fragilis YCH46; NP_421082.1: Caulobacter crescentus CB15; YP_001034962.1: Streptococcus sanguinis SK36; YP_211509.1 : Bacteroides fragilis NCTC 9343; XP_001940256.1: Pyrenophora tritici-repentis Pt-1C-BFP; YP_001450196.1: Streptococcus gordonii str. Challis substr. CH1; XP_680937.1: Aspergillus nidulans FGSC A4; XP_001265696.1: Neosartorya fischeri NRRL 181; XP_367405.1: Magnaporthe grisea 70-15; YP_001033719.1: Lactococcus lactis subsp. cremoris MG1363; YP_812051.1: Lactococcus lactis subsp. cremoris SK11; NP_810723.1: Bacteroides thetaiotaomicron VPI-5482; YP_141537.1: Streptococcus thermophilus CNRZ1066; NP_721920.1: Streptococcus mutans UA159; YP_820528.1: Streptococcus thermophilus LMD-9; XP_001246405.1: Coccidioides immitis RS; YP_612063.1: Silicibacter sp. TM1040; YP_001727726.1: Leuconostoc citreum KM20;YP_444835.1: Salinibacter ruber DSM 13855; NP_943356.1: Klebsiella pneumoniae; YP_631535.1: Myxococcus xanthus DK 1622; YP_809822.1: Oenococcus oeni PSU-1; NP_687842.1: Streptococcus agalactiae 2603V/R; YP_329573.1: Streptococcus agalactiae A909; XP_001692446.1: Chlamydomonas reinhardtii; YP_511991.1: Jannaschia sp. CCS1; XP_719233.1: Candida albicans SC5314; NP_735295.1: Streptococcus agalactiae NEM316; YP_168601.1: Silicibacter pomeroyi DSS-3; YP_001169098.1: Rhodobacter sphaeroides ATCC 17025; XP_001216799.1: Aspergillus terreus NIH2624; YP_437091.1: Hahella chejuensis KCTC 2396; XP_001390389.1: Aspergillus niger CBS 513.88;XP_001550500.1: Botryotinia fuckeliana B05.10; XP_001273080.1: Aspergillus clavatus NRRL 1; YP_001042044.1: Rhodobacter sphaeroides ATCC 17029; YP_916587.1: Paracoccus denitrificans PD1222; XP_961529.1: Neurospora crassa OR74A; YP_351544.1: Rhodobacter sphaeroides 2.4.1; XP_001588803.1: Sclerotinia sclerotiorum 1980; XP_762195.1: Ustilago maydis 521; XP_001537789.1: Ajellomyces capsulatus NAm1; YP_001198994.1: Streptococcus suis 05ZYH33; NP_390842.1: Bacillus subtilis subsp. subtilis str. 168.

C94/C118 signature (20 sequences, 5.6%):
NP_288269.1: *Escherichia coli O157:H7 EDL933；* YP_001469780.1: *Thermotoga lettingae TMO*; NP_623333.1: *Thermoanaerobacter tengcongensis MB4*; YP_001567589.1: Petrotoga mobilis SJ95； NP_228329.1: Thermotoga maritima MSB8； YP_001244004.1: Thermotoga petrophila RKU-1; YP_001738457.1: Thermotoga sp. RQ2; YP_591728.1: Acidobacteria bacterium Ellin345； YP_001193798.1: Flavobacterium johnsoniae UW101； YP_001296902.1: Flavobacterium psychrophilum JIP02/86; XP_001439866.1: Paramecium tetraurelia strain d4-2; XP_001438779.1: Paramecium tetraurelia strain d4-2; YP_001319456.1: Alkaliphilus metalliredigens QYMF； XP_001022412.1: Tetrahymena thermophila SB210； YP_001306613.1: Thermosipho melanesiensis BI429； YP_002335308.1: Thermosipho africanus TCF52B; NP_394308.1: Thermoplasma acidophilum DSM 1728; YP_862969.1: Gramella forsetii KT0803; YP_023758.1: Picrophilus torridus DSM 9790; NP_111339.1: Thermoplasma volcanium GSS1; YP_001380982.1: Anaeromyxobacter sp. Fw109-5.

C84/C118 signature (5 sequences, 1.4%):
YP_806472.1: Lactobacillus casei ATCC 334; YP_001987415.1: Lactobacillus casei BL23; NP_785802.1: Lactobacillus plantarum WCFS1;YP_818384.1: Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293; YP_001621174.1: Acholeplasma laidlawii PG-8A