

Clover phyllosphere microbial rhodopsin sequences

D or **E** mark the proton donor and acceptor sites

L mark position 105

```
>contig08984_16 [247 - 2] (REVERSE SENSE)
KWIWYGWSCVAFLGVLYLLWGPLRLQAVSRGAALGTAYAKNLIFLTVIWFLYPIVFLVGP
EGLKLITDPASVWAFILMDIIA
>FQNZ4AG01BSQEZ_12 [244 - 11] (REVERSE SENSE)
ARNLVRKARLLTFASWGFYPIVYMAPYAGIGGGAATTVIQIGYAIADVAKAGLGVLIYM
ISVRKSQVEAHGVALDAH
>FQNZ4AG01CA2GU_11 [245 - 3] (REVERSE SENSE)
EQQPVAARDLVRRARLLTFASWGFYPIVYMAPYANLTGATATTVIQLGYTIADIVAKAGV
GVLIFMISVRKSQVEAHSPLV
>FQNZ4AG01CF7IA_11 [274 - 2] (REVERSE SENSE)
LRIALVGLWGVYPIAYLFPILGGDFFGGEGGFVLKQAGYSIADILAKAAYGLAIYKVARV
KSRLEDPEYDDHPDTVHDGSHEAISGETVGA
>FQNZ4AG01D0VFX_4 [3 - 248]
VPRIVMSGGHSETAPQPTPRTSTVFLWLNRSQVSPHYRTALTVAGLATAIALYHQVRIFD
SWTSAYDVKDGAITATGYAFND
>FQNZ4AG02JCLA8_6 [2 - 241]
LASSALPARDRGARGAMRSIEAPMDVVTPLQFSLAYNAFSFTLAAMGAATAFLWLSRSQV
APAYRTALTISGIVTAIAAY
>FQNZ4AG02GVS6U_7 [2 - 247]
DGSITYLARYLDWSVTTPILLAGLVLAAFHEHGKPGTVGGYLTAIIVLDVLMILTGLISS
LAPASPAKWWVYVWSCVAFVGV
>FQNZ4AG02FSHCG_13 [246 - 1] (REVERSE SENSE)
RGYLTSVIVLDVLMIVTGLISSLAAPSLAKWIWYGWSCVAFLGVLYLLWGPLRSQAVNRG
AALGTAYSKNLIFLTVIWFLYP
>FQNZ4AG02JHV6V_2 [2 - 250]
ACLLFTSLYLVVSQRQVAKQYRNAVIVSATVCAIAAYHYFRIFNNFEASYPSGETIDSAH
QLSNIEFNEGYRYVDWFLTVPLL
>FQNZ4AG02G8F9K_11 [234 - 1] (REVERSE SENSE)
LLHASVPFIAATAYLAMAFGIGTLFKLDGSVTYLARYLDWSVTTPILLSSLVLLAFHKHG
RTGEVGGHLTAIIVLDVL
>FSA20RK01DSYOG_16 [243 - 1] (REVERSE SENSE)
NGTTIYFARYIDWVTTPILLAGLVMIASFHEYGPSSQMNGYLTAIIGLDIIMVVTGLLSS
LADTFMVKAIWYLVSCMAFLG
>FSA20RK01DZ6C2_12 [241 - 2] (REVERSE SENSE)
RSRPRDLVRRARLLTFASWGFYPIVYMAPYANLTGATATTVIQLGYTIADIVAKAGVGVL
IFMISVRKSQVEAHSPLVHA
>FSA20RK01DKSZZ_10 [238 - 2] (REVERSE SENSE)
NQQPVAARGLVRKARLLTFASWGFYPIVYMAPYAGITGASATTVIQLGYTIADLVAKAGV
GILIFLISVRKSQAEAHDG
>FSA20RK01A0V76_2 [1 - 234]
ARSGAQGAPVSPSPPGVSTRSSTWRPMSGSPAGTATTVIQLGYTIADLLAKAGLGILIFM
ISSRKSRAEAHGDTGHYA
>FSA20RK02FJI45_6 [107 - 223]
RHRQRSSCRRCGLGDTTRSLSPSRRRSKRCDNAAKPATG
```

>FSA20RK02FJI45_12 [233 - 3] (REVERSE SENSE)
LWLNRSQVSPHYRTALTVAGLATAIALYHQVRIFDSWTSAYDVKDGAITATGYAFNDAYR
YVWLLTVPLLLLELVL

>FSA20RK02FIV8E_10 [192 - 1] (REVERSE SENSE)
VAMIVLGYPGEGAASNMARAFWSTASTIPFLYIMWQLFRGLGEAIERQPPEEVRGRIRQAR
LLTL

>FSA20RK02FSNPI_4 [1 - 201]
TFASWGFYPIVYMAPYANLTGATATTVIQLGYTIADIVAKAGVGVLFMISVRKSQVEAH
SPLVHAE

>FSA20RK02HORY1_4 [1 - 249]
KSIADQPVAQGLVRKARLLTFASWGFYPIVYMAPYAGLTGGTVTTTTIQIGYTIADIVAK
AGLGILIFLIAVRKSEAEADAHK

>FSA20RK02GEWCN_3 [1 - 243]
LGFTGLYLVISQRRVLKRYRNALIVSATVCFIATYHYFRIFNNFKESYQTQSQQGGTGTFT
LTNVEFNEAYRYVDWFLTVRC

>FSA20RK02HNTEKW_13 [228 - 1] (REVERSE SENSE)
IYATGKKTQPSGHHTLLHASVFPFIAATAYLAMAFGIGTLINSDGSLTYLARYVDWSVTTP
ILLGLVLLAFHEHGK

>FSA20RK02IW86B_5 [78 - 233]
PRSRRIITCASSRAGMRAYTLKDGVIAATGFAFNDAYRYVDWLLTVPLLLLE

>FSA20RK02GLOZA_10 [174 - 1] (REVERSE SENSE)
LVMRLSAEETSSKAVRLGSAALMIILGYPGEIATDIPTRALWGTLSSIPFVYIVWEL

>FSA20RK02HT1KJ_4 [1 - 228]
GIAAYHYRIFESWIAAYSFGNGVVTATPFAFNDAYRYVDWLLTVPLLLLELVLVMSLR
EETSSKAVRLGSAAL