

1 10 20 30 40 50 60

GCA_000337795.1_Haloferax_denitrificans
GCA_013407385.1_Nitrosarchaeum_sp._AC2
GCA_000380525.1_Thioalkalivibrio_sp._AKL8
GCA_000478095.1_Corynebacterium_sp._KPL1824
GCA_000775045.1_Burkholderia_pseudomallei
GCA_001458695.1_Nitrospira_inopinata
GCA_001723765.1_Scalindua_rubra
GCA_003097635.1_Flavobacterium_psychrotolerans
GCA_003135435.1_Nitrospira_sp.
GCA_003335105.1_Rhodanobacter_denitrificans
GCA_003495885.1_Neisseria_meningitidis
GCA_013267085.1_Pseudomonas_stutzeri
GCA_900492205.1_Rhizobium_naphthalenivorans
GCA_905339145.1_Anaerolineae_bacterium
Globobulimina_pacifica
GCA_00004985.1_Naegleria_gruberi
GCA_014843455.1_2_Fusarium_oxysporum
GCA_000204585.1_4_outgroup1
GCA_003113895.1_outgroup2
consensus>80

ANRIIAADPRDIPAPITRTSPETVSVE **L** ETLEQVAE **V** EPGVT **F** TYMT **F** NQGQ **V** PGP **F** IRT RVE
. ADSGRKVVEFNL **T** GESVTLPI **I** MGKTFNAMT **F** S **G** Q **V** PGP **T** LRV TQE
REGFAHDASLAP . . AANTTTHEETW **S** MTEEEIV **V** APGVKOMQWL **F** NGQ **A** PGP **T** LRGKV
. ALQP . . LDAARDQAIR **L** DTTHTVIR **I** APGIAFAA **W** T **F** GNQ **V** PGP **P** VHV KV
. SAGAKVHDVT **F** TATESEIVI **D** DGTKY **K** AWT **F** NGQ **M** PGP **V** VRV TQE
. GNIPKTVRVT **L** TAQEITATI **I** DSNVSFRY **F** T **F** GR **V** PGP **F** IRVME
MEAAELTAPPLVPKPIGSRDATKLIVN **M** EVKEMESE **L** ADGVKY **T** TYWT **F** FGGS **V** PGS **F** IRT RVE
IKAVVTTAPNVPPPITRTEPANVLVE **L** EAHEWVGN **L** SDDNKY **K** FWGFGT **V** PGP **M** IRV MV
IHAVLTSPPHPVPPPPIHRNYPAKVIVE **L** EVVEKEMP **I** SEGVST **T** FWT **F** FGFT **V** PGS **F** IRV RQ
IDAVTTHAPEPVPPAIDRDYPAKVRVK **M** ETVEKTMT **M** MEDGVE **V** RYWT **F** FDGD **V** PGP **M** IRV RE
GKDLTMMDPSKGVEPVGKREPKNLTID **L** RTTEEGR **L** SDGSSY **R** FWT **F** NGT **V** PGP **M** MRI RE
VKVDLVKPPFVHTQKAEGGPVKEFTM **T** IEKKLV **V** DDDGEV **H** AMT **F** DGS **V** PGP **L** MVV HE
. QNPALVGLPIQLRPGATPRTVTLN **L** KAEEVIA **E** VAMPNFSF **W** T **F** NGAV **V** PGP **M** LRV ME
. LPRRHSPVTHLIT **L** TSVEVVAE **L** EEEGH **F** EFT **Y** NAT **V** PGP **P** LRV RE
VEAILTKAPGIPPPPIDRDYPVQLVVNL **L** DTTIEVKP **I** SKQYK **Y** PYWT **F** NGT **V** PGP **F** IRV RV
EAALVTSAPNVPPPITRKHPVLLQVALATETKLAQ **L** TSQYK **Y** EQWT **F** NGT **V** PGP **F** IRV RV
. NTIRDFT **I** IAEDNKIE **I** ISPGVF **Y** NVN **W** T **F** NGT **V** PGP **P** IRATE
. KQENTIREFQL **I** TAGTSIIQL **N** SAVS **Y** NI **W** DLNGR **I** PGP **T** LRA KQ
. Y . . . f . g . v PGP . . . r . . .

70 80 90 100 110 120

GCA_000337795.1_Haloferax_denitrificans
GCA_013407385.1_Nitrosarchaeum_sp._AC2
GCA_000380525.1_Thioalkalivibrio_sp._AKL8
GCA_000478095.1_Corynebacterium_sp._KPL1824
GCA_000775045.1_Burkholderia_pseudomallei
GCA_001458695.1_Nitrospira_inopinata
GCA_001723765.1_Scalindua_rubra
GCA_003097635.1_Flavobacterium_psychrotolerans
GCA_003135435.1_Nitrospira_sp.
GCA_003335105.1_Rhodanobacter_denitrificans
GCA_003495885.1_Neisseria_meningitidis
GCA_013267085.1_Pseudomonas_stutzeri
GCA_900492205.1_Rhizobium_naphthalenivorans
GCA_905339145.1_Anaerolineae_bacterium
Globobulimina_pacifica
GCA_00004985.1_Naegleria_gruberi
GCA_014843455.1_2_Fusarium_oxysporum
GCA_000204585.1_4_outgroup1
GCA_003113895.1_outgroup2
consensus>80

130 140 150 160 170 180

GCA_000337795.1_Haloferax_denitrificans
GCA_013407385.1_Nitrosarchaeum_sp._AC2
GCA_000380525.1_Thioalkalivibrio_sp._AKL8
GCA_000478095.1_Corynebacterium_sp._KPL1824
GCA_000775045.1_Burkholderia_pseudomallei
GCA_001458695.1_Nitrospira_inopinata
GCA_001723765.1_Scalindua_rubra
GCA_003097635.1_Flavobacterium_psychrotolerans
GCA_003135435.1_Nitrospira_sp.
GCA_003335105.1_Rhodanobacter_denitrificans
GCA_003495885.1_Neisseria_meningitidis
GCA_013267085.1_Pseudomonas_stutzeri
GCA_900492205.1_Rhizobium_naphthalenivorans
GCA_905339145.1_Anaerolineae_bacterium
Globobulimina_pacifica
GCA_00004985.1_Naegleria_gruberi
GCA_014843455.1_2_Fusarium_oxysporum
GCA_000204585.1_4_outgroup1
GCA_003113895.1_outgroup2
consensus>80

190 200 210 220 230 240

GCA_000337795.1_Haloferax_denitrificans
GCA_013407385.1_Nitrosarchaeum_sp._AC2
GCA_000380525.1_Thioalkalivibrio_sp._AKL8
GCA_000478095.1_Corynebacterium_sp._KPL1824
GCA_000775045.1_Burkholderia_pseudomallei
GCA_001458695.1_Nitrospira_inopinata
GCA_001723765.1_Scalindua_rubra
GCA_003097635.1_Flavobacterium_psychrotolerans
GCA_003135435.1_Nitrospira_sp.
GCA_003335105.1_Rhodanobacter_denitrificans
GCA_003495885.1_Neisseria_meningitidis
GCA_013267085.1_Pseudomonas_stutzeri
GCA_900492205.1_Rhizobium_naphthalenivorans
GCA_905339145.1_Anaerolineae_bacterium
Globobulimina_pacifica
GCA_00004985.1_Naegleria_gruberi
GCA_014843455.1_2_Fusarium_oxysporum
GCA_000204585.1_4_outgroup1
GCA_003113895.1_outgroup2
consensus>80

250 260 270 280 290

GCA_000337795.1_Haloferax_denitrificans
GCA_013407385.1_Nitrosarchaeum_sp._AC2
GCA_000380525.1_Thioalkalivibrio_sp._AKL8
GCA_000478095.1_Corynebacterium_sp._KPL1824
GCA_000775045.1_Burkholderia_pseudomallei
GCA_001458695.1_Nitrospira_inopinata
GCA_001723765.1_Scalindua_rubra
GCA_003097635.1_Flavobacterium_psychrotolerans
GCA_003135435.1_Nitrospira_sp.
GCA_003335105.1_Rhodanobacter_denitrificans
GCA_003495885.1_Neisseria_meningitidis
GCA_013267085.1_Pseudomonas_stutzeri
GCA_900492205.1_Rhizobium_naphthalenivorans
GCA_905339145.1_Anaerolineae_bacterium
Globobulimina_pacifica
GCA_00004985.1_Naegleria_gruberi
GCA_014843455.1_2_Fusarium_oxysporum
GCA_000204585.1_4_outgroup1
GCA_003113895.1_outgroup2
consensus>80

1 10 20 30 40 50 60

ANRIIAADPRDIPAPITRTSPETVSVE **L** ETLEQVAE **V** EPGVT **F** TYMT **F** NQGQ **V** PGP **F** IRT RVE
. ADSGRKVVEFNL **T** GESVTLPI **I** MGKTFNAMT **F** S **G** Q **V** PGP **T** LRV TQE
REGFAHDASLAP . . AANTTTHEETW **S** MTEEEIV **V** APGVKOMQWL **F** NGQ **A** PGP **T** LRGKV
. ALQP . . LDAARDQAIR **L** DTTHTVIR **I** APGIAFAA **W** T **F** GNQ **V** PGP **P** VHV KV
. SAGAKVHDVT **F** TATESEIVI **D** DGTKY **K** AWT **F** NGQ **M** PGP **V** VRV TQE
. GNIPKTVRVT **L** TAQEITATI **I** DSNVSFRY **F** T **F** GR **V** PGP **F** IRVME
MEAAELTAPPLVPKPIGSRDATKLIVN **M** EVKEMESE **L** ADGVKY **T** TYWT **F** FGGS **V** PGS **F** IRT RVE
IKAVVTTAPNVPPPITRTEPANVLVE **L** EAHEWVGN **L** SDDNKY **K** FWGFGT **V** PGP **M** IRV MV
IHAVLTSPPHPVPPPPIHRNYPAKVIVE **L** EVVEKEMP **I** SEGVST **T** FWT **F** FGFT **V** PGS **F** IRV RQ
IDAVTTHAPEPVPPAIDRDYPAKVRVK **M** ETVEKTMT **M** MEDGVE **V** RYWT **F** FDGD **V** PGP **M** IRV RE
GKDLTMMDPSKGVEPVGKREPKNLTID **L** RTTEEGR **L** SDGSSY **R** FWT **F** NGT **V** PGP **M** MRI RE
VKVDLVKPPFVHTQKAEGGPVKEFTM **T** IEKKLV **V** DDDGEV **H** AMT **F** DGS **V** PGP **L** MVV HE
. QNPALVGLPIQLRPGATPRTVTLN **L** KAEEVIA **E** VAMPNFSF **W** T **F** NGAV **V** PGP **M** LRV ME
. LPRRHSPVTHLIT **L** TSVEVVAE **L** EEEGH **F** EFT **Y** NAT **V** PGP **P** LRV RE
VEAILTKAPGIPPPPIDRDYPVQLVVNL **L** DTTIEVKP **I** SKQYK **Y** PYWT **F** NGT **V** PGP **F** IRV RV
EAALVTSAPNVPPPITRKHPVLLQVALATETKLAQ **L** TSQYK **Y** EQWT **F** NGT **V** PGP **F** IRV RV
. NTIRDFT **I** IAEDNKIE **I** ISPGVF **Y** NVN **W** T **F** NGT **V** PGP **P** IRATE
. KQENTIREFQL **I** TAGTSIIQL **N** SAVS **Y** NI **W** DLNGR **I** PGP **T** LRA KQ
. Y . . . f . g . v PGP . . . r . . .

70 80 90 100 110 120

GDTV **D** IIRN **H** EDNSMV **H** NVDFH **H** ACRRGGGAETATN **V** AP **G** ERQLRF **K** VTYP **G** AF **F** YHC **A**
GDV **V** KM **T** TL **T** I PAGEVTCH **H** GNDM **H** ASQMSA . . GNFES **V** NP **G** ETSQYC **Y** IA **E** AA **G** VF **F** YHC **S**
GNI **V** RM **E** IS **N** SGD . . VM **H** GAS **I** HAA **A** YTQTSKHVGH **I** LP **G** QTKSIT **F** RATT **P** **G** VF **F** YHC **C**
GDKFKI **T** I **K** NEGS . . MA **H** SID **F** H AGEVSPDENMSK **I** QP **G** GEELTYEF **F** TANRA **G** IW **N** YHC **S**
GDRVR **L** SMT **N** RSDEPM **M** HSMDFH **H** AMVSPTDKRS **I** AP **G** QTMHLEF **F** THNYP **G** VF **F** YRC **G**
GDTVN **F** TL **I** GHKDNAF **H** SMDFH **H** AMVSPTDKRS **I** AP **G** QTMHLEF **F** THNYP **G** VF **F** YRC **G**
GDTLE **V** T **L** VN **P** KTNTE **T** **H** TVD **F** HAIKFRGGATRMM **A** PPG **Q** SRRSS **F** QI **T** R **P** **G** LY **F** YHC **V**
GDEVE **F** H **L** RN **H** PDNKLS **H** NIDL **H** AVTQGGGAASSFT **T** AP **G** HESQFT **F** KALNQ **G** IY **F** YHC **A**
GDTVE **I** N **L** K **N** DKNSKES **H** NIDL **H** AVNPGGGAAMLNT **E** PG **Q** ESKLR **F** KALNA **G** LY **F** YHC **A**
GDTVE **F** H **L** K **N** APDSKMF **H** NIDL **H** GTPGGGAASSFT **T** AP **G** HESQFT **F** KALNQ **G** IY **F** YHC **A**
GDTVE **V** EF **S** N **N** PSSTV **H** NVD **H** ATQGGGAASSFT **T** AP **G** RTSTFS **F** KALQ **P** **G** LY **F** YHC **A**
GDTVT **L** N **L** T **N** ELD **S** NH **H** SID **H** AVTPGGGAAVTQ **A** AP **G** QTRSFT **F** KALQ **P** **G** LY **F** YHC **A**
GDYVE **E** L **T** L **I** N **P** ETNTLQ **H** NIDL **H** STALGGGALT **I** VP **N** G EKTI **L** R **F** KATKA **G** VF **F** YHC **A**
GDTVI **V** N **I** S **N** DSKNTRAYA **I** DMP **A** ELPNRDTVT **N** LMP **G** ETATL **T** F **S** AAKS **G** AY **F** YGA **G**
GDWID **L** T **F** I **N** PNTSLH **E** SMDL **H** GMP **T** L **S** LDGGAASL **R** INPG **E** ARTV **W** Q **A** I **P** **G** MF **F** YHC **A**
GDMQV **N** Y **N** L **L** DETGM **A** H **N** ID **H** C **V** TPGGGAEM **L** A **E** K **D** E **E** KTGF **F** KLLT **P** **G** LF **F** YHC **A**
GDVVEL **T** L **T** N **K** DPAGNH **H** NIDL **H** AFTPGGAAVTT **V** EEN **N** ESKTAR **F** KLLY **P** **G** LY **F** YHC **A**
GDLVR **I** H **F** I **N** G **S** **N** S **G** M **F** G **L** L **V** N **G** A **I** H **E** M **D** G **V** F **E** V **A** GP **V** GV **H** I **Y** HC **H**
GDRIR **V** L **F** L **N** NAG . . HS **H** S **L** H **F** H **G** VPAEMDGIRP **I** S **N** G **S** ATIYE **F** D **A** EPY **G** V **H** I **Y** HC **H**

130 140 150 160 170 180

ANV **D** Y **H** I **S** A **G** M **F** G **I** I **L** V **P** EEGLPEVDH **E** FY **L** GQH **E** L **V** TNGKGQKGHHEFDTRMAMEDP
VKMDQ **H** V **L** S **G** M **Y** G **L** T **I** V **D** P **I** IDGYNADALE **E** FT **L** QYQ **N** QL **Y** L TPEGNYDAGKMFHQN
GGIPN **H** V **M** F **G** Q **Y** G **M** I **V** V **E** P **R** DPYKEPD **L** KLY **M** LQH **E** LY A SGKEAVEGDA
APMSI **H** I **A** G **M** G **A** N **G** V **I** I **D** P **P** D . . LKDVAE **E** YNFIA **D** V **F** L GEEKTGADAQRVADGEF
PMVLP **H** I **A** S **G** M

Fig. S1 Representative NirS alignment. Yellow boxes represent >70% conserved residues and red boxes represent identical residues. Numbered red boxes mark characteristic conserved motifs (I, III, IV, V, VII and IV distinguish NirS from NirN and NirF) and blue boxes mark functionally important residues, and asterisks beneath alignment mark specific residues where nirS can be differentiated from NirF and NirN. Notably, motif I on the *cyt_c* domain was generally characterised by -CaGCHg- in NirS, motif VI by -LHD- (vs. -PyD- in NirF and -LDD- in NirN), and motif VII at the *cyt_{d1}* domain is identified by the universal absence of a proline in the third position in NirN and NirF (i.e. -PHpGpG- vs. PH!PgeG). Clade 1c: *Hydrogenophilales* sp.; clade 3: *Sulfurimonas paralvinellae*; clade 4: *Calidifontibacillus erzurumensis*; clade 1d: *Brocadia* sp. WS118; clade 1e: *c. Magnetaquicoccus inordinatus*; clade 1a: *Stutzerimonas stutzeri*; clade 1f: *Delta proteobacteria bact. GWA2_43_19*; clade 5: *Levilinea saccharolytica*; clade 1h: *Scalindua brodae*; clade 1j: *Thermus oshimai*; clade 1b: *Thauera linaloolentis*; clade 2: *Methylomicrobium album*; clade 6: candidatus *Methylomirabilis oxyfera*; clade 1g: *Roseiflexus castenholzii*; clade 1k: *Colwellia psychrerythraea*; halophilic archaea NirS-like: *Natrinema pellirubrum*. The Fig. was prepared in ESPscript/ENDscript.