

The protein IDs of each subset (for prokaryotic extracellular sequences.)

| Subset 1 | Subset 2 | Subset 3 | Subset 4 | Subset 5 |
|------------|------------|------------|------------|------------|
| SPI_BACBR | XYNB_STRLI | XYNC_STRLI | LIPE_AERHY | PEL1_ERWCA |
| PRTB_ERWCH | PELE_ERWCH | STRK_STRGR | EBA1_FLAME | A85B_MYCKA |
| THER_BACST | BPRX_BACNO | PELB_ERWCA | SUBF_BACSU | AMYR_BACS8 |
| PRSG_ECOLI | PRTC_ERWCH | SNPA_STRLI | AMT4_PSEST | NPRS_BACST |
| GTFB_STRMU | HLT_VIBPA | PROA_XANCP | SACB_STRMU | NPRE_BACCE |
| GTF1_STRDO | PELD_ERWCH | PRTA_ERWCH | DRNE_VIBCH | SODF_MYCTU |
| CHOD_BREST | API_ACHLY | NUC_STAHY | NUCB_BACSU | SUBE_BACSU |
| HRPZ_PSESY | SEPA_STAEP | PIL1_SALTY | PAPG_ECOLI | AGAR_STRCO |
| NPRV_VIBPR | GTFD_STRMU | PRTS_SERMA | PELB_ERWCH | COMX_BACSU |
| EMPA_VIBAN | PROB_STRAG | PELA_ERWCH | PELC_ERWCH | ELAS_PSEAE |
| XYNA_STRLI | PELF_ERWCH | PELA_ERWCA | NPRE_BACAM | PAPF_ECOLI |
| NPRE_BACCL | AMYB_BACPO | PEL3_ERWCA | NPRM_BACME | LSTP_STASI |
| PRT1_ERWCA | SNPA_STRCO | SMP_SERMA | PIL1_ECOLI | TCPA_VIBCH |
| DEXT_ARTSP | PIL5_ECOLI | PHL3_BACCE | PBPA_STRPN | BPRV_BACNO |
| GTFC_STRMU | PHL1_BACCE | NPRE_BACSU | LSTP_STAST | A85B_MYCLE |
| NPRE_BACBR | APRA_PSEAE | AMT4_PSESA | AMT6_BACS7 | EBA2_FLAME |
| PAPA_ECOLI | LKTA_PASHA | CHOD_STRSQ | RNBR_BACAM | A85B_MYCBO |
| SNPA_STRSQ | THET_THEVU | PRTG_ERWCH | SUBV_BACSU | PAPE_ECOLI |
| RN_BACCO | A85B_MYCAV | NUC_SERMA | LIP_PSESP | MPR_BACSU |
| PHL_LEPIN | XYNC_PSEFL | PEL_BACSU | PRSE_ECOLI | DRNE_AERHY |
| EBA3_FLAME | HRPN_ERWAM | PAPH_ECOLI | PHB_ALCFA | PRTT_SERMA |
| | | | | CYAA_BORPE |
| | | | | PROA_LEGPN |