



Data Article

16S rRNA gene sequencing data of the human skin microbiome before and after swimming in the ocean



Marisa C. Nielsen^{a,*}, Sunny C. Jiang^{a,b}

^a Environmental Health Sciences, University of California, Irvine, United States

^b Civil and Environmental Engineering, University of California, Irvine, CA, United States

ARTICLE INFO

Article history:

Received 21 March 2021

Revised 7 June 2021

Accepted 8 June 2021

Available online 9 June 2021

Keywords:

Antibiotic resistance

Skin microbiome

Recreational water

Sequencing

Skin infections

ABSTRACT

These data represent the abundance, diversity and predicted function gene profiles of the microbial communities present on human skin before and after swimming in the ocean. The skin microbiome has been shown to provide protection against infection from pathogenic bacteria. It is well-known that exposure to ocean water can cause skin infection, but little is known about how exposure can alter the bacterial communities on the skin. Skin microbiome samples were collected from human participants before and after swimming in the ocean. These data were used to analyze the changes in abundance and diversity of microbial communities on the skin and the changes in the functional profiles of the bacteria, specifically focusing on genes involved in antibiotic resistance and bacterial virulence.

© 2021 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

DOI of original article: [10.1016/j.envres.2021.110978](https://doi.org/10.1016/j.envres.2021.110978)

* Corresponding author at: 709 Engineering Tower, University of California, Irvine, CA 92697, United States

E-mail address: mchattma@uci.edu (M.C. Nielsen).

<https://doi.org/10.1016/j.dib.2021.107207>

2352-3409/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject	Microbiology: microbiome
Specific subject area	Investigation of the changes in human skin microbial communities after ocean water exposure
Type of data	Table
How data were acquired	Sequencing Datasets (16S rRNA gene) 16S rRNA next-generation sequencing of extracted DNA from human skin microbiome swab samples. Instrument: Ion Torrent PGM at MRDNA/Molecular Research LP [1] Software: QIIME [2], PICRUSt [3]
Data format	Raw Analyzed
Parameters for data collection	Skin swab samples were collected from human volunteers at the beach. Participants had to consent and meet inclusion criteria (detailed in the experimental design section) to enroll.
Description of data collection	Skin microbiome samples were collected from human participants before swimming in the ocean, after swimming, and at 6h and 24 h post-swim.
Data source location	Institution: University of California, Irvine City/Town/Region: Irvine, CA Country: United States
Data accessibility	Samples were collected at Huntington Beach, CA Repository name: NCBI SRA BioProject ID: PRJNA735376 Direct URL to data: http://www.ncbi.nlm.nih.gov/bioproject/735376 Repository name: Mendeley Data Direct URL to data: https://data.mendeley.com/datasets/th7bfgfc6m/1 [4] and https://data.mendeley.com/datasets/hjnfzm949y/1 [5]
Related research article	M.C. Nielsen, N. Wang, S.C. Jiang, Acquisition of antibiotic resistance genes on human skin after swimming in the ocean, <i>Environ. Res.</i> 197 (2021) 110978 [6].

Value of the Data

- These data are valuable to marine, environmental and clinical microbiome researchers interested in understanding the role of the environment in skin microbial community function and alteration.
- These data can be used by public health officials interested in investigating mitigation and surveillance strategies for antibiotic resistant infections.
- The microbiome datasets provided can be used to further examine how changes in the microbiome, including additional function gene profiles, are affected by exposure to ocean water.

1. Data Description

The raw datasets contain 16S rRNA gene sequence data of swab samples taken from human skin before and after swimming in the ocean. Sequences are also provided for samples taken 6 h and 24 h post-swim from most of the participants. These data were used to investigate the changes in the skin microbiome after swimming in the ocean. QIIME was used to analyze changes in microbial communities and PICRUSt was used to analyze the predicted function gene profiles.

Table 1 lists the antibiotic resistance genes (ARGs) and their corresponding Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthologs (KOs) detected in the samples. Table 2 lists the KOs associated with virulence factor genes (VFGs) detected in the samples and their corresponding KEGG descriptions.

Table 1

Antibiotics and their corresponding ARGs and KOs detected in the samples.

Antibiotic	Gene (KO)
Vancomycin	vanX (K08641); vanY (K07260); vraR (K07694); vraS (K07681)
Tetracycline	tetA/tetG/H/J (K08151); tetK (K08168)
fluoroquinolone, quinolone, florfenicol, chloramphenicol, and amphenicol (FCA)	adeA/cmeA (K03585); catB3 (K00638); basR (K07771); qepA (K08167)
beta-lactams	acrA (K03585); ampC (K01467); ampG (K08218); blaI (K02171); blaR1 (K02172); cfxA (K01624); ftsI (K03587); mecA (K02545); mecR1 (K02547); metallo-beta-lactamase family protein (K07576); mrcA (K05366); mrdA (K05515); nagZ (K01207); ompU (K08720); ompC (K09475); ompF (K09476); pbpA (K12552); pbp1b (K03693); pbp2A (K12555); pbp2B (K00687); pbp2X (K12556); pbp3 (K12553); penA (K03587); tolC (K12340)
multidrug	emrE/qac/mmr/smr(K03297); MATE family (K03327); emrB(K03446); emrA(K03543); marC(K05595); mdtB(K07788); mdtC(K07789); mdtA(K07799); lmrP(K08152); blt(K08153); mdfA/cmr(K08160); mdtG(K08161); mdtH(K08162); mdtL(K08163); yebQ(K08169); norB/C(K08170); yitG/yfmfD/yfmO(K08221); oprJ(K08721); ebrA(K11814); ebrB(K11815)
aminoglycosides	aacC1(K03395); aacC2(K00662); aacC4(K00663); aadA1(K00984); aadE(K05593); ybcL(K08164)
macrolide, lincosamide, and streptogramin B (MLSb)	ermC/A(K00561); ereA_B(K06880); mph(K06979); mef(K08217); macA(K13888)

2. Experimental Design, Materials and Methods

2.1. Sample collection

Data collection methods were approved by the University of California, Irvine Institutional Review Board (IRB #2017-3751). Verbal consent was collected from participants before registration. Sample collection events occurred in April and September 2018 in Huntington Beach, CA. In brief, participants were given a detailed sample collection description if they met the participant criteria. Skin microbiome samples were obtained from 9 participants in April and 12 participants in September and included males and females with age ranges from 24–39, with no sunscreen application, infrequent exposure to the ocean and beach (once per month or less), no shower/bath in the past 12 hours, no antibiotic usage in the past 6 months, and no active infections. Rayon-tipped swabs moistened in sterile saline were used to swab an 8cm × 8cm section of skin on the back of the participants' calves before the individuals swam in the ocean. Participants were then instructed to swim or wade in the ocean for 10 minutes and air dry. At that time, a second set of samples were taken. The before and after samples were collected from the same calf but on different sections of the skin to ensure the sample collection taken before swimming did not remove bacteria from the section of skin swabbed after swimming. The participants were then instructed to leave the beach and not to shower or wash the leg area until after the 6 h and 24 h post-swim samples were collected. The 6 h samples were collected from the right side of the left calf and the 24 h samples were collected from the left side of the left calf. Samples were kept on ice after collection and all samples were processed within 24 h of collection. Ocean water samples (75mL each) were collected from each swim site at the time of experimental sample collection and were analyzed in the same manner as the experimental samples. [7].

Table 2

Descriptions of the KOs associated with VFGs detected in the samples.

#OTU ID	KEGG Description
K12984	["(heptosyl)LPS beta-1,4-glucosyltransferase [Enzyme commission number (EC):2.4.1.-]"]
K00655	["1-acyl-sn-glycerol-3-phosphate acyltransferase [EC:2.3.1.51]"]
K01771	["1-phosphatidylinositol phosphodiesterase [EC:4.6.1.13]"]
K00216	["2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase [EC:1.3.1.28]"]
K02510	["2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase [EC:4.1.2.-]"]
K01627	["2-dehydro-3-deoxyphosphooctonate aldolase (KDO 8-P synthase) [EC:2.5.1.55]"]
K01626	["3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]"]
K11211	["3-deoxy-D-manno-octulosonic acid kinase [EC:2.7.1.-]"]
K02527	["3-deoxy-D-manno-octulosonic-acid transferase [EC:2.-.-.-]"]
K00979	["3-deoxy-manno-octulosonate cytidyltransferase (CMP-KDO synthetase) [EC:2.7.7.38]"]
K00074	["3-hydroxybutyryl-CoA dehydrogenase [EC:1.1.1.157]"]
K01704	["3-isopropylmalate/(R)-2-methylmalate dehydratase small subunit [EC:4.2.1.33 4.2.1.35]"]
K00059	["3-oxoacyl-[acyl-carrier protein] reductase [EC:1.1.1.100]"]
K02372	["3R-hydroxymyristoyl ACP dehydrase [EC:4.2.1.-]"]
K00097	["4-hydroxythreonine-4-phosphate dehydrogenase [EC:1.1.1.262]"]
K06133	["4'-phosphopantetheinyl transferase [EC:2.7.8.-]"]
K14956	["6 kDa early secretory antigenic target"]
K00652	["8-amino-7-oxononanoate synthase [EC:2.3.1.47]"]
K01990	["ABC-2 type transport system ATP-binding protein"]
K01992	["ABC-2 type transport system permease protein"]
K10952	["accessory cholera enterotoxin"]
K10936	["accessory colonization factor AcfA"]
K10937	["accessory colonization factor AcfB"]
K10938	["accessory colonization factor AcfC"]
K10939	["accessory colonization factor AcfD"]
K07813	["accessory gene regulator B"]
K04072	["acetaldehyde dehydrogenase / alcohol dehydrogenase [EC:1.2.1.10 1.1.1.1]"]
K03896	["acetyl CoA:N6-hydroxylysine acetyl transferase [EC:2.3.1.102]"]
K00821	["acetylorntine/N-succinylidiaminopimelate aminotransferase [EC:2.6.1.11 2.6.1.17]"]
K01078	["acid phosphatase [EC:3.1.3.2]"]
K03191	["acid-activated urea channel"]
K02078	["acyl carrier protein"]
K13061	["acyl homoserine lactone synthase [EC:2.3.1.184]"]
K13060	["acyl homoserine lactone synthase [EC:2.3.1.184]"]
K00249	["acyl-CoA dehydrogenase [EC:1.3.99.3]"]
K07116	["acyl-homoserine-lactone acylase [EC:3.5.1.97]", """]
K00860	["adenylylsulfate kinase [EC:2.7.1.25]"]
K13735	["adhesin/invasin"]
K03274	["ADP-L-glycerol-D-manno-heptose 6-epimerase [EC:5.1.3.20]"]
K03894	["aerobactin synthetase subunit alpha [EC:6.3.2.27]"]
K03895	["aerobactin synthetase subunit beta [EC:6.3.2.27]"]
K12678	["AIDA-1 adhesin-like protein"]
K00697	["alpha,alpha-trehalose-phosphate synthase (UDP-forming) [EC:2.4.1.15]"]
K12994	["alpha-1,3-rhamnosyltransferase [EC:2.4.1.-]"]
K00766	["anthranilate phosphoribosyltransferase [EC:2.4.2.18]"]
K01657	["anthranilate synthase component I [EC:4.1.3.27]"]
K12687	["antigen 43"]
K06041	["arabinose-5-phosphate isomerase [EC:5.3.1.13]"]
K05372	["AraC family transcriptional regulator"]
K06599	["AraC family transcriptional regulator, chemosensory pili system protein ChpD"]
K10923	["AraC family transcriptional regulator, TCP pilus virulence regulatory protein"]
K12243	["AraC family transcriptional regulator, transcriptional activator of the genes for pyochelin and ferripyochelin receptors"]
K01953	["asparagine synthase (glutamine-hydrolysing) [EC:6.3.5.4]"]
K01579	["aspartate 1-decarboxylase [EC:4.1.1.11]"]
K01779	["aspartate racemase [EC:5.1.1.13]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K03224	["ATP synthase in type III secretion protein SctN [EC:3.6.3.14]"]
K11004	["ATP-binding cassette, subfamily B, bacterial HlyB/CyaB"]
K14698	["ATP-binding cassette, subfamily B, bacterial IrtA [EC:3.6.3.-]"]
K14699	["ATP-binding cassette, subfamily B, bacterial IrtB [EC:3.6.3.-]"]
K11085	["ATP-binding cassette, subfamily B, bacterial MsbA [EC:3.6.3.-]"]
K12530	["ATP-binding cassette, subfamily B, bacterial RtxB"]
K12531	["ATP-binding cassette, subfamily B, bacterial RtxE"]
K06147	["ATP-binding cassette, subfamily B, bacterial"]
K12536	["ATP-binding cassette, subfamily C, bacterial HasD"]
K06148	["ATP-binding cassette, subfamily C, bacterial"]
K03696	["ATP-dependent Clp protease ATP-binding subunit ClpC"]
K03697	["ATP-dependent Clp protease ATP-binding subunit ClpE"]
K01358	["ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]"]
K01401	["aureolysin [EC:3.4.24.29]"]
K12688	["autotransporter serine protease [EC:3.4.21.-]"]
K00785	["beta-galactosamide-alpha-2,3-sialyltransferase [EC:2.4.99.-]"]
K11609	["beta-ketoacyl ACP synthase [EC:2.3.1.-]"]
K13714	["bifunctional autolysin [EC:3.5.1.28 3.2.1.96]"]
K11936	["biofilm PGA synthesis N-glycosyltransferase PgaC [EC:2.4.-.-]"]
K11937	["biofilm PGA synthesis protein PgaD"]
K03561	["biopolymer transport protein ExbB"]
K03559	["biopolymer transport protein ExbD"]
K06011	["bontoxilysin [EC:3.4.24.69]"]
K08652	["C5a peptidase [EC:3.4.21.110]"]
K12107	["cag pathogenicity island protein 22"]
K10915	["CAI-1 autoinducer synthase [EC:2.3.-.-]"]
K11045	["cAMP factor"]
K07265	["capsular polysaccharide export protein"]
K07266	["capsular polysaccharide export protein"]
K09689	["capsular polysaccharide transport system ATP-binding protein"]
K10107	["capsular polysaccharide transport system permease protein"]
K09688	["capsular polysaccharide transport system permease protein"]
K03563	["carbon storage regulator"]
K01673	["carbonic anhydrase [EC:4.2.1.1]"]
K06131	["cardiolipin synthase [EC:2.7.8.-]"]
K03781	["catalase [EC:1.11.1.6]"]
K03782	["catalase/peroxidase [EC:1.11.1.6 1.11.1.7]"]
K12956	["cation-transporting ATPase V [EC:3.6.3.-]"]
K12950	["cation-transporting P-type ATPase C [EC:3.6.3.-]"]
K00523	["CDP-4-dehydro-6-deoxyglucose reductase [EC:1.17.1.1]"]
K12452	["CDP-6-deoxy-D-xylo-4-hexulose-3-dehydrase"]
K01709	["CDP-glucose 4,6-dehydratase [EC:4.2.1.45]"]
K09809	["CDP-glycerol glycerophosphotransferase [EC:2.7.8.12]"]
K04095	["cell filamentation protein"]
K05789	["chain length determinant protein (polysaccharide antigen chain regulator)"]
K12519	["chaperone protein PapD"]
K04077	["chaperonin GroEL"]
K06596	["chemosensory pili system protein ChpA (sensor histidine kinase/response regulator)"]
K06597	["chemosensory pili system protein ChpB (putative protein-glutamate methyltransferase)"]
K06598	["chemosensory pili system protein ChpC"]
K06600	["chemosensory pili system protein ChpE"]
K03411	["chemotaxis protein CheD [EC:3.5.1.44]"]
K03414	["chemotaxis protein CheZ"]
K00575	["chemotaxis protein methyltransferase CheR [EC:2.1.1.80]"]
K02556	["chemotaxis protein MotA"]
K02557	["chemotaxis protein MotB"]
K03933	["chitin-binding protein"]
K10928	["cholera enterotoxin subunit A [EC:2.4.2.36]"]
K10929	["cholera enterotoxin subunit B"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K01442	["chologylglycine hydrolase [EC:3.5.1.24]"]
K03496	["chromosome partitioning protein"]
K03497	["chromosome partitioning protein, ParB family"]
K01644	["citrate lyase subunit beta [EC:4.1.3.6]", "citrate lyase subunit beta / citryl-CoA lyase [EC:4.1.3.6 4.1.3.34]"]
K14201	["clumping factor A"]
K14192	["clumping factor B"]
K08605	["coccolysin [EC:3.4.24.30]"]
K02237	["competence protein ComEA"]
K14446	["crotonyl-CoA carboxylase/reductase"]
K10914	["CRP/FNR family transcriptional regulator, cyclic AMP receptor protein"]
K07796	["Cu(I)/Ag(I) efflux system outer membrane protein CusC"]
K04565	["Cu/Zn superoxide dismutase [EC:1.15.1.1]"]
K04337	["curli production assembly/transport component CsgE"]
K04338	["curli production assembly/transport component CsgF"]
K06214	["curli production assembly/transport component CsgG"]
K04336	["curli production protein"]
K02500	["cyclase HisF [EC:4.1.3.-]", "cyclase [EC:4.1.3.-]"]
K13688	["cyclic beta-1,2-glucan synthetase [EC:2.4.1.-]"]
K00574	["cyclopropane-fatty-acyl-phospholipid synthase [EC:2.1.1.79]"]
K11049	["CylG protein"]
K01697	["cystathionine beta-synthase [EC:4.2.1.22]"]
K01738	["cysteine synthase A [EC:2.5.1.47]"]
K02424	["cystine transport system substrate-binding protein"]
K02197	["cytochrome c-type biogenesis protein CcmE"]
K02198	["cytochrome c-type biogenesis protein CcmF"]
K11013	["cytolethal distending toxin subunit A"]
K11014	["cytolethal distending toxin subunit B"]
K11015	["cytolethal distending toxin subunit C"]
K07389	["cytolysin-activating lysine-acyltransferase [EC:2.3.1.-]"]
K11008	["cytotoxic necrotizing factor 1"]
K03367	["D-alanine-poly(phosphoribitol) ligase [EC:6.1.1.13]", "D-alanine-poly(phosphoribitol) ligase subunit 1 [EC:6.1.1.13]"]
K07262	["D-alanyl-D-alanine endopeptidase (penicillin-binding protein 7) [EC:3.4.99.-]", "D-alanyl-D-alanine endopeptidase (penicillin-binding protein 7) [EC:3.4.21.-]"]
K03272	["D-beta-D-heptose 7-phosphate kinase / D-beta-D-heptose 1-phosphate adenylyltransferase [EC:2.7.1.- 2.7.7.-]"]
K01494	["dCTP deaminase [EC:3.5.4.13]"]
K12202	["defect in organelle trafficking protein DotA"]
K12203	["defect in organelle trafficking protein DotB"]
K12204	["defect in organelle trafficking protein DotC"]
K12205	["defect in organelle trafficking protein DotD"]
K11039	["delta-hemolysin"]
K00689	["dextranucrase [EC:2.4.1.5]"]
K03273	["D-glycero-D-manno-heptose 1,7-bisphosphate phosphatase [EC:3.1.3.-]"]
K00836	["diaminobutyrate-2-oxoglutarate transaminase [EC:2.6.1.76]"]
K01586	["diaminopimelate decarboxylase [EC:4.1.1.20]"]
K12239	["dihydroaeruginic acid synthetase"]
K02073	["D-methionine transport system substrate-binding protein"]
K01972	["DNA ligase (NAD+) [EC:6.5.1.2]"]
K04483	["DNA repair protein RadA"]
K03631	["DNA repair protein RecN (Recombination protein N)"]
K03466	["DNA segregation ATPase FtsK/SpoIIIE, S-DNA-T family"]
K03746	["DNA-binding protein H-NS"]
K03530	["DNA-binding protein HU-beta"]
K01790	["dTDP-4-dehydrorhamnose 3,5-epimerase [EC:5.1.3.13]"]
K00067	["dTDP-4-dehydrorhamnose reductase [EC:1.1.1.133]"]
K01710	["dTDP-glucose 4,6-dehydratase [EC:4.2.1.46]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K03709	["DtxR family transcriptional regulator, Mn-dependent transcriptional regulator"]
K13461	["effector protein HopM1"]
K02358	["elongation factor EF-Tu [EC:3.6.5.3]", "elongation factor Tu"]
K01689	["enolase [EC:4.2.1.11]"]
K00209	["enoyl-[acyl-carrier-protein] reductase (NADPH2, B-specific) [EC:1.3.1.10]"]
K02363	["enterobactin 2,3-dihydroxybenzoate-AMP ligase / S-dihydroxybenzoyltransferase [EC:2.7.7.58 2.3.1.-]"]
K01252	["enterobactin isochorismatase [EC:3.3.2.1]"]
K02362	["enterobactin synthetase component D [EC:2.7.8.-]"]
K02364	["enterobactin synthetase component F [EC:2.7.7.-]"]
K01066	["esterase / lipase [EC:3.1.1.-]"]
K11041	["exfoliative toxin A/B"]
K01142	["exodeoxyribonuclease III [EC:3.1.11.2]"]
K12423	["fatty acid CoA ligase FadD21"]
K12427	["fatty acid CoA ligase FadD28"]
K12421	["fatty acid CoA ligase FadD9"]
K00666	["fatty-acyl-CoA synthase [EC:6.2.1.-]"]
K13255	["ferric iron reductase protein FhuF"]
K01772	["ferrochelataase [EC:4.99.1.1]"]
K04758	["ferrous iron transport protein A"]
K04759	["ferrous iron transport protein B"]
K14200	["fibrinogen-binding protein"]
K13734	["fibronectin-binding protein 1"]
K13732	["fibronectin-binding protein A"]
K13733	["fibronectin-binding protein B"]
K07346	["fimbrial chaperone protein"]
K08087	["fimbrial protein FimW"]
K08088	["fimbrial protein FimY"]
K07351	["fimbrial protein"]
K03773	["FKBP-type peptidyl-prolyl cis-trans isomerase FkB [EC:5.2.1.8]"]
K02386	["flagella basal body P-ring formation protein FlgA"]
K02399	["flagella synthesis protein FlgN"]
K13626	["flagellar assembly factor FliW"]
K02411	["flagellar assembly protein FliH"]
K02389	["flagellar basal-body rod modification protein FlgD"]
K02387	["flagellar basal-body rod protein FlgB"]
K02388	["flagellar basal-body rod protein FlgC"]
K02391	["flagellar basal-body rod protein FlgF"]
K02392	["flagellar basal-body rod protein FlgG"]
K02400	["flagellar biosynthesis protein FlhA"]
K02404	["flagellar biosynthesis protein FlhF"]
K04562	["flagellar biosynthesis protein FlhG"]
K04061	["flagellar biosynthesis protein"]
K02401	["flagellar biosynthetic protein FlhB"]
K02419	["flagellar biosynthetic protein FliP"]
K02420	["flagellar biosynthetic protein FliQ"]
K02421	["flagellar biosynthetic protein FliR"]
K13820	["flagellar biosynthetic protein FliR/FlhB"]
K02413	["flagellar FliJ protein"]
K02415	["flagellar FliL protein"]
K02390	["flagellar hook protein FlgE"]
K02396	["flagellar hook-associated protein 1 FlgK"]
K02407	["flagellar hook-associated protein 2"]
K02397	["flagellar hook-associated protein 3 FlgL"]
K02408	["flagellar hook-basal body complex protein FliE"]
K02414	["flagellar hook-length control protein FliK"]
K02393	["flagellar L-ring protein precursor FlgH"]
K02410	["flagellar motor switch protein FlgI"]
K02416	["flagellar motor switch protein FliM"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K02417	["flagellar motor switch protein FlIN/FliY"]
K02409	["flagellar M-ring protein FlIF"]
K02394	["flagellar P-ring protein precursor Flgl"]
K06602	["flagellar protein FlaF"]
K06603	["flagellar protein FlaG"]
K02385	["flagellar protein FlbD"]
K06601	["flagellar protein FlbT"]
K02395	["flagellar protein FlgJ"]
K03516	["flagellar protein FlhE"]
K02418	["flagellar protein FlhO/FlhZ"]
K02422	["flagellar protein FlhS"]
K02423	["flagellar protein FlhT"]
K06604	["flagellar rod protein Flal"]
K02402	["flagellar transcriptional activator FlhC"]
K02403	["flagellar transcriptional activator FlhD"]
K02406	["flagellin"]
K02412	["flagellum-specific ATP synthase [EC:3.6.3.14]"]
K02425	["FlhZ protein"]
K07533	["foldase protein PrsA [EC:5.2.1.8]"]
K01622	["fructose 1,6-bisphosphate aldolase/phosphatase [EC:4.1.2.13 3.1.3.11]"]
K13007	["Fuc2NAc and GlcNAc transferase [EC:2.4.1.-]"]
K03711	["Fur family transcriptional regulator, ferric uptake regulator"]
K00681	["gamma-glutamyltranspeptidase [EC:2.3.2.2]"]
K02377	["GDP-L-fucose synthase [EC:1.1.1.271]"]
K01711	["GDPmannose 4,6-dehydratase [EC:4.2.1.47]"]
K00066	["GDP-mannose 6-dehydrogenase [EC:1.1.1.132]"]
K03285	["general bacterial porin, GBP family"]
K02450	["general secretion pathway protein A"]
K02451	["general secretion pathway protein B"]
K02452	["general secretion pathway protein C"]
K02453	["general secretion pathway protein D"]
K02454	["general secretion pathway protein E"]
K02455	["general secretion pathway protein F"]
K02456	["general secretion pathway protein G"]
K02457	["general secretion pathway protein H"]
K02458	["general secretion pathway protein I"]
K02459	["general secretion pathway protein J"]
K02460	["general secretion pathway protein K"]
K02461	["general secretion pathway protein L"]
K02462	["general secretion pathway protein M"]
K02463	["general secretion pathway protein N"]
K02464	["general secretion pathway protein O [EC:3.4.23.43 2.1.1.-]"]
K02465	["general secretion pathway protein S"]
K01178	["glucoamylase [EC:3.2.1.3]"]
K00973	["glucose-1-phosphate thymidyltransferase [EC:2.7.7.24]"]
K01810	["glucose-6-phosphate isomerase [EC:5.3.1.9]"]
K12998	["glucosyltransferase [EC:2.4.1.-]"]
K01845	["glutamate-1-semialdehyde 2,1-aminomutase [EC:5.4.3.8]"]
K02501	["glutamine amidotransferase [EC:2.4.2.-]"]
K01915	["glutamine synthetase [EC:6.3.1.2]"]
K01318	["glutamyl endopeptidase [EC:3.4.21.19]"]
K02492	["glutamyl-tRNA reductase [EC:1.2.1.70]"]
K00134	["glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]"]
K00613	["glycine amidinotransferase [EC:2.1.4.1]"]
K07270	["glycosyl transferase, family 25"]
K13002	["glycosyltransferase [EC:2.4.1.-]"]
K01495	["GTP cyclohydrolase I [EC:3.5.4.16]"]
K00951	["GTP pyrophosphokinase [EC:2.7.6.5]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K12545	["heme acquisition protein HasA"]
K02193	["heme exporter protein A [EC:3.6.3.41]"]
K02194	["heme exporter protein B"]
K02196	["heme exporter protein D"]
K07215	["heme oxygenase"]
K11005	["hemolysin A"]
K07326	["hemolysin activation/secretion protein"]
K11017	["hemolysin activation/secretion protein??"]
K11035	["hemolysin BL binding component"]
K11037	["hemolysin BL lytic component L1"]
K11036	["hemolysin BL lytic component L2"]
K11003	["hemolysin D"]
K11139	["hemolysin E"]
K11032	["hemolysin II"]
K11068	["hemolysin III"]
K11016	["hemolysin"]
K10948	["hemolysin"]
K02498	["HemY protein"]
K02848	["heptose (I) phosphotransferase [EC:2.7.1.-]"]
K02841	["heptosyltransferase I [EC:2.4.-.-]"]
K12982	["heptosyltransferase I [EC:2.4.-.-]"]
K02843	["heptosyltransferase II [EC:2.4.-.-]"]
K02849	["heptosyltransferase III [EC:2.4.-.-]"]
K01590	["histidine decarboxylase [EC:4.1.1.22]"]
K10016	["histidine transport system permease protein"]
K13993	["HSP20 family protein"]
K00752	["hyaluronan synthase [EC:2.4.1.212]"]
K01727	["hyaluronate lyase [EC:4.2.2.1]"]
K01197	["hyaluronoglucosaminidase [EC:3.2.1.35]"]
K10814	["hydrogen cyanide synthase HcnA [EC:1.4.99.5]"]
K10815	["hydrogen cyanide synthase HcnB [EC:1.4.99.5]"]
K10816	["hydrogen cyanide synthase HcnC [EC:1.4.99.5]"]
K01749	["hydroxymethylbilane synthase [EC:2.5.1.61]"]
K09117	["hypothetical protein"]
K09790	["hypothetical protein"]
K09153	["hypothetical protein"]
K07226	["hypothetical protein"]
K09824	["hypothetical protein"]
K09860	["hypothetical protein"]
K07227	["hypothetical protein"]
K09152	["hypothetical protein"]
K01390	["IgA-specific metalloendopeptidase [EC:3.4.24.13]"]
K01347	["IgA-specific serine endopeptidase [EC:3.4.21.72]"]
K09607	["immune inhibitor A [EC:3.4.24.-]"]
K14196	["immunoglobulin G-binding protein A"]
K14197	["immunoglobulin G-binding protein Sbi"]
K11021	["insecticidal toxin complex protein TccC"]
K13730	["internalin A"]
K12206	["intracellular multiplication protein lcmB"]
K12207	["intracellular multiplication protein lcmC"]
K12208	["intracellular multiplication protein lcmD"]
K12209	["intracellular multiplication protein lcmE"]
K12211	["intracellular multiplication protein lcmG"]
K12212	["intracellular multiplication protein lcmJ"]
K12213	["intracellular multiplication protein lcmK"]
K12214	["intracellular multiplication protein lcmL"]
K12216	["intracellular multiplication protein lcmN"]
K12217	["intracellular multiplication protein lcmO"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K12218	["intracellular multiplication protein lcmP"]
K12219	["intracellular multiplication protein lcmQ"]
K12221	["intracellular multiplication protein lcmS"]
K12222	["intracellular multiplication protein lcmT"]
K12223	["intracellular multiplication protein lcmV"]
K12224	["intracellular multiplication protein lcmW"]
K12225	["intracellular multiplication protein lcmX"]
K13285	["invasin B"]
K13287	["invasin D"]
K14953	["ipoprotein LpqH"]
K02014	["iron complex outermembrane receptor protein"]
K02013	["iron complex transport system ATP-binding protein [EC:3.6.3.34]"]
K02015	["iron complex transport system permease protein"]
K02016	["iron complex transport system substrate-binding protein"]
K02010	["iron(III) transport system ATP-binding protein [EC:3.6.3.30]"]
K02011	["iron(III) transport system permease protein"]
K02012	["iron(III) transport system substrate-binding protein"]
K11704	["iron/zinc/copper transport system substrate-binding protein"]
K04782	["isochorismate pyruvate-lyase [EC:4.1.3.-]", "isochorismate pyruvate lyase [EC:4.2.99.21]"]
K02361	["isochorismate synthase [EC:5.4.4.2]"]
K01637	["isocitrate lyase [EC:4.1.3.1]"]
K13745	["L-2,4-diaminobutyrate decarboxylase [EC:4.1.1.86]"]
K02529	["LacI family transcriptional regulator"]
K08642	["LasA protease [EC:3.4.24.-]"]
K02654	["leader peptidase (prepilin peptidase) / N-methyltransferase [EC:3.4.23.43 2.1.1.-]"]
K02506	["leader peptidase HopD [EC:3.4.23.43]"]
K11038	["leukocidin/hemolysin toxin family protein"]
K02560	["lipid A biosynthesis (KDO)2-(lauroyl)-lipid iva acyltransferase [EC:2.3.1.-]"]
K02517	["lipid A biosynthesis lauroyl acyltransferase [EC:2.3.1.-]"]
K00748	["lipid-A-disaccharide synthase [EC:2.4.1.182]"]
K03800	["lipoate-protein ligase A [EC:2.7.7.63]"]
K05790	["lipopolysaccharide biosynthesis protein WzzE"]
K07271	["lipopolysaccharide cholinephosphotransferase [EC:2.7.8.-]"]
K09691	["lipopolysaccharide transport system ATP-binding protein"]
K09690	["lipopolysaccharide transport system permease protein"]
K14954	["lipoprotein LprG"]
K05803	["lipoprotein Nlpl"]
K12426	["long chain fatty acid CoA FadD26"]
K01909	["long-chain-fatty-acid-[acyl-carrier-protein] ligase [EC:6.2.1.20]"]
K10531	["L-ornithine N5-oxygenase [EC:1.13.12.-]"]
K07782	["LuxR family transcriptional regulator"]
K04333	["LuxR family transcriptional regulator, csgAB operon transcriptional regulatory protein"]
K03897	["lysine N6-hydroxylase [EC:1.14.13.59]"]
K01185	["lysozyme [EC:3.2.1.17]"]
K05303	["macrocin O-methyltransferase [EC:2.1.1.101]"]
K04334	["major curlin subunit"]
K12517	["major pilin subunit PapA"]
K07345	["major type 1 subunit fimbrin (pilin)"]
K01638	["malate synthase [EC:2.3.3.9]"]
K10112	["maltose/maltodextrin transport system ATP-binding protein"]
K11607	["manganese/iron transport system ATP-binding protein"]
K11605	["manganese/iron transport system permease protein"]
K11606	["manganese/iron transport system permease protein"]
K11604	["manganese/iron transport system substrate-binding protein"]
K00971	["mannose-1-phosphate guanylyltransferase [EC:2.7.7.22]"]
K01809	["mannose-6-phosphate isomerase [EC:5.3.1.8]"]
K01227	["mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase [EC:3.2.1.96]"]
K12989	["mannosyltransferase [EC:2.4.1.-]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K13001	["mannosyltransferase [EC:2.4.1.-]"]
K05375	["MbtH protein"]
K01273	["membrane dipeptidase [EC:3.4.13.19]"]
K07058	["membrane protein"]
K14743	["membrane-anchored mycosin MYCP [EC:3.4.21.-]"]
K08305	["membrane-bound lytic murein transglycosylase B [EC:3.2.1.-]"]
K08307	["membrane-bound lytic murein transglycosylase D [EC:3.2.1.-]"]
K13640	["MerR family transcriptional regulator, heat shock protein HspR"]
K00604	["methionyl-tRNA formyltransferase [EC:2.1.2.9]"]
K05874	["methyl-accepting chemotaxis protein I, serine sensor receptor"]
K05875	["methyl-accepting chemotaxis protein II, aspartate sensor receptor"]
K03406	["methyl-accepting chemotaxis protein"]
K07552	["MFS transporter, DHA1 family, bicyclomycin/chloramphenicol resistance protein"]
K02429	["MFS transporter, FHS family, L-fucose permease"]
K02575	["MFS transporter, NNP family, nitrate/nitrite transporter"]
K07784	["MFS transporter, OPA family, hexose phosphate transport protein UhpT"]
K05373	["MFS transporter, putative signal transducer"]
K01531	["Mg ²⁺ -importing ATPase [EC:3.6.3.2]"]
K01387	["microbial collagenase [EC:3.4.24.3]"]
K01174	["micrococcal nuclease [EC:3.1.31.1]"]
K04335	["minor curlin subunit"]
K07349	["minor fimbrial subunit"]
K07348	["minor fimbrial subunit"]
K07350	["minor fimbrial subunit"]
K04043	["molecular chaperone DnaK"]
K04080	["molecular chaperone IbpA"]
K03455	["monovalent cation:H ⁺ antiporter-2, CPA2 family"]
K12276	["MSHA biogenesis protein MshE"]
K12277	["MSHA biogenesis protein MshF"]
K12278	["MSHA biogenesis protein MshG"]
K12279	["MSHA biogenesis protein MshI"]
K12280	["MSHA biogenesis protein MshJ"]
K12281	["MSHA biogenesis protein MshK"]
K12282	["MSHA biogenesis protein MshL"]
K12283	["MSHA biogenesis protein MshM"]
K12284	["MSHA biogenesis protein MshN"]
K12285	["MSHA biogenesis protein MshO"]
K12286	["MSHA biogenesis protein MshP"]
K12287	["MSHA biogenesis protein MshQ"]
K10924	["MSHA pilin protein MshA"]
K10925	["MSHA pilin protein MshB"]
K10926	["MSHA pilin protein MshC"]
K10927	["MSHA pilin protein MshD"]
K02026	["multiple sugar transport system permease protein"]
K02025	["multiple sugar transport system permease protein"]
K02027	["multiple sugar transport system substrate-binding protein"]
K04793	["mycobactin lysine-N-oxygenase"]
K04789	["mycobactin peptide synthetase MbtE"]
K04792	["mycobactin peptide synthetase MbtF"]
K04788	["mycobactin phenylloxazoline synthetase"]
K04790	["mycobactin polyketide synthetase MbtC"]
K04791	["mycobactin polyketide synthetase MbtD"]
K04787	["mycobactin salicyl-AMP ligase [EC:6.3.2.-]"]
K08068	["N-acetylglucosamine-6-phosphate 2-epimerase and phosphatase [EC:5.1.3.-]"]
K05946	["N-acetylglucosaminyldiphosphoundecaprenol [EC:2.4.1.187]"]
K01447	["N-acetylmuramoyl-L-alanine amidase [EC:3.5.1.28]"]
K01654	["N-acetylneuraminate synthase [EC:2.5.1.56]"]
K00983	["N-acetylneuraminate cytidyltransferase [EC:2.7.7.43]"]
K00343	["NADH dehydrogenase I subunit N [EC:1.6.5.3]", "NADH-quinone oxidoreductase subunit N [EC:1.6.5.3]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K00336	["NADH-quinone oxidoreductase subunit G [EC:1.6.5.3]", "NADH dehydrogenase I subunit G [EC:1.6.5.3]"]
K02398	["negative regulator of flagellin synthesis FlgM"]
K01389	["nepriylisin [EC:3.4.24.11]"]
K00370	["nitrate reductase 1, alpha subunit [EC:1.7.99.4]"]
K00371	["nitrate reductase 1, beta subunit [EC:1.7.99.4]"]
K00373	["nitrate reductase 1, delta subunit [EC:1.7.99.4]"]
K00374	["nitrate reductase 1, gamma subunit [EC:1.7.99.4]"]
K02003	["None"]
K02004	["None"]
K00257	["None"]
K07001	["None"]
K07145	["None"]
K06998	["None"]
K07043	["None"]
K07154	["None"]
K07126	["None"]
K07011	["None"]
K06867	["None"]
K07018	["None"]
K07164	["None"]
K01795	["None"]
K07017	["None"]
K00786	["None"]
K06921	["None"]
K07031	["None"]
K06900	["None"]
K06887	["None"]
K06882	["None"]
K01732	["None"]
K11033	["non-hemolytic enterotoxin A"]
K11034	["non-hemolytic enterotoxin B/C"]
K12237	["nonribosomal peptide synthetase VibF"]
K00940	["nucleoside-diphosphate kinase [EC:2.7.4.6]"]
K02428	["nucleoside-triphosphate pyrophosphatase [EC:3.6.1.19]"]
K13012	["O-antigen biosynthesis protein WbqP"]
K03286	["OmpA-OmpF porin, OOP family"]
K07268	["opacity associated protein"]
K00611	["ornithine carbamoyltransferase [EC:2.1.3.3]"]
K01750	["ornithine cyclodeaminase [EC:4.3.1.12]"]
K12340	["outer membrane channel protein TolC"]
K12686	["outer membrane lipase/esterase"]
K02494	["outer membrane lipoprotein LolB"]
K12538	["outer membrane protein HasF"]
K10940	["outer membrane protein OmpT"]
K08720	["outer membrane protein OmpU"]
K07277	["outer membrane protein"]
K12518	["outer membrane usher protein PapC"]
K07347	["outer membrane usher protein"]
K02495	["oxygen-independent coproporphyrinogen III oxidase [EC:1.3.99.22]"]
K12973	["palmitoyl transferase [EC:2.3.1.-]"]
K00954	["pantetheine-phosphate adenyltransferase [EC:2.7.7.3]"]
K01918	["pantoate-beta-alanine ligase [EC:6.3.2.1]"]
K12267	["peptide methionine sulfoxide reductase msrA/msrB [EC:1.8.4.11 1.8.4.12]"]
K02032	["peptide/nickel transport system ATP-binding protein"]
K03640	["peptidoglycan-associated lipoprotein"]
K03767	["peptidyl-prolyl cis-trans isomerase A (cyclophilin A) [EC:5.2.1.8]"]
K03832	["periplasmic protein TonB"]
K13010	["perosamine synthetase"]
K03386	["peroxiredoxin (alkyl hydroperoxide reductase subunit C) [EC:1.11.1.15]"]
K12681	["pertactin"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K13063	["phenazine biosynthesis protein phzE [EC:2.6.1.86]"]
K12440	["phenolphthiocerol synthesis type-I polyketide synthase A"]
K12441	["phenolphthiocerol synthesis type-I polyketide synthase B"]
K12443	["phenolphthiocerol synthesis type-I polyketide synthase D"]
K12444	["phenolphthiocerol synthesis type-I polyketide synthase E"]
K04750	["PhnB protein"]
K00981	["phosphatidate cytidyltransferase [EC:2.7.7.41]"]
K13292	["phosphatidylglycerol:prolipoprotein diacylglycerol transferase [EC:2.-.-.-]"]
K03760	["phosphoethanolamine transferase"]
K01835	["phosphoglucomutase [EC:5.4.2.2]"]
K03431	["phosphoglucosamine mutase [EC:5.4.2.10]"]
K03271	["phosphoheptose isomerase [EC:5.-.-.-]"]
K01114	["phospholipase C [EC:3.1.4.3]"]
K01840	["phosphomannomutase [EC:5.4.2.8]"]
K01923	["phosphoribosylaminoimidazole-succinocarboxamide synthase [EC:6.3.2.6]"]
K14728	["phthioidiolone/phenolphthioidiolone dimycocerosates ketoreductase [EC:1.2.-.-]"]
K02279	["pilus assembly protein CpaB"]
K02280	["pilus assembly protein CpaC"]
K02281	["pilus assembly protein CpaD"]
K02282	["pilus assembly protein CpaE"]
K02283	["pilus assembly protein CpaF"]
K02651	["pilus assembly protein Flp/PilA"]
K13925	["plasmin and fibronectin-binding protein A"]
K08566	["plasminogen activator [EC:3.4.23.48]"]
K01729	["poly(beta-D-mannuronate) lyase [EC:4.2.2.3]"]
K07282	["poly-gamma-glutamate synthesis protein (capsule biosynthesis protein)"]
K12430	["polyketide synthase 1/15"]
K12436	["polyketide synthase 12"]
K12433	["polyketide synthase 5"]
K12434	["polyketide synthase 7"]
K00937	["polyphosphate kinase [EC:2.7.4.1]"]
K01991	["polysaccharide export outer membrane protein"]
K03328	["polysaccharide transporter, PST family"]
K01698	["porphobilinogen synthase [EC:4.2.1.24]"]
K02278	["prepilin peptidase CpaA [EC:3.4.23.43]"]
K02682	["prepilin peptidase dependent protein D"]
K03070	["preprotein translocase subunit SecA"]
K12537	["protease secretion protein HasE"]
K13571	["proteasome accessory factor A [EC:6.3.2.-]"]
K13527	["proteasome-associated ATPase"]
K11030	["protective antigen"]
K13743	["protein IpgB2"]
K14204	["protein Map"]
K02504	["protein transport protein HofB"]
K02505	["protein transport protein HofC"]
K02507	["protein transport protein HofQ"]
K01104	["protein-tyrosine phosphatase [EC:3.1.3.48]"]
K03305	["proton-dependent oligopeptide transporter, POT family"]
K00230	["protoporphyrinogen oxidase [EC:1.3.3.4]"]
K01399	["pseudolysin [EC:3.4.24.26]"]
K03408	["purine-binding chemotaxis protein CheW"]
K07804	["putative virulence related protein PagC"]
K02021	["putative ABC transport system ATP-binding protein"]
K02067	["putative ABC transport system substrate-binding protein"]
K06160	["putative ATP-binding cassette transporter"]
K06132	["putative cardiolipin synthase [EC:2.7.8.-]"]
K03606	["putative colanic acid biosynthesis UDP-glucose lipid carrier transferase"]
K06994	["putative drug exporter of the RND superfamily"]
K07386	["putative endopeptidase [EC:3.4.24.-]"]
K07010	["putative glutamine amidotransferase"]
K10039	["putative glutamine transport system substrate-binding protein"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K07225	["putative hemin transport protein"]
K06442	["putative hemolysin"]
K03699	["putative hemolysin"]
K07507	["putative Mg ²⁺ transporter-C (MgtC) family protein"]
K07497	["putative transposase"]
K07498	["putative transposase"]
K12238	["pyochelin biosynthesis protein PchD"]
K12242	["pyochelin biosynthetic protein PchC"]
K12240	["pyochelin synthetase"]
K03474	["pyridoxine 5-phosphate synthase [EC:2.6.99.2]"]
K00286	["pyrroline-5-carboxylate reductase [EC:1.5.1.2]"]
K06137	["pyrroloquinoline-quinone synthase [EC:1.3.3.11]"]
K00162	["pyruvate dehydrogenase E1 component subunit beta [EC:1.2.4.1]"]
K00171	["pyruvate ferredoxin oxidoreductase, delta subunit [EC:1.2.7.1]"]
K07740	["regulator of sigma D"]
K11749	["regulator of sigma E protease [EC:3.4.24.-]"]
K07272	["rhamnosyltransferase [EC:2.4.1.-]"]
K12997	["rhamnosyltransferase [EC:2.4.1.-]"]
K12990	["rhamnosyltransferase [EC:2.4.1.-]"]
K12996	["rhamnosyltransferase [EC:2.4.1.-]"]
K12995	["rhamnosyltransferase [EC:2.4.1.-]"]
K09761	["ribosomal RNA small subunit methyltransferase E [EC:2.1.1.-]"]
K03789	["ribosomal-protein-alanine N-acetyltransferase [EC:2.3.1.128]"]
K09748	["ribosome maturation factor RimpP", "hypothetical protein"]
K09889	["ribosome-associated protein", "hypothetical protein"]
K03087	["RNA polymerase nonessential primary-like sigma factor"]
K03086	["RNA polymerase primary sigma factor"]
K02405	["RNA polymerase sigma factor for flagellar operon FlIA"]
K03092	["RNA polymerase sigma-54 factor"]
K03088	["RNA polymerase sigma-70 factor, ECF subfamily"]
K03090	["RNA polymerase sigma-B factor"]
K10953	["RTX toxin RtxA"]
K12532	["RTX toxin transporter"]
K00121	["S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase [EC:1.1.1.284 1.1.1.1]"]
K01851	["salicylate biosynthesis isochorismate synthase [EC:5.4.4.2]"]
K04781	["salicylate synthetase [EC:5.4.4.2 4.2.99.21]", "salicylate synthetase [EC:5.4.4.2 4.1.3.-]"]
K06024	["segregation and condensation protein B"]
K00640	["serine O-acetyltransferase [EC:2.3.1.30]"]
K12684	["serine protease autotransporter [EC:3.4.21.-]"]
K04691	["serine protease DegS [EC:3.4.21.-]"]
K04771	["serine protease Do [EC:3.4.21.107]"]
K08884	["serine/threonine protein kinase, bacterial [EC:2.7.11.1]"]
K11915	["serine/threonine protein phosphatase Stp1 [EC:3.1.3.16]"]
K14949	["serine/threonine-protein kinase PknG [EC:2.7.11.1]"]
K11912	["serine/threonine-protein kinase PpkA [EC:2.7.11.1]"]
K11916	["serine/threonine-protein kinase Stk1 [EC:2.7.11.-]"]
K14194	["serine-aspartate repeat-containing protein C/D/E"]
K01406	["serralyisin [EC:3.4.24.40]"]
K12683	["serum resistance protein"]
K01186	["sialidase-1 [EC:3.2.1.18]"]
K11914	["sigma-54 dependent transcriptional regulator"]
K11908	["sigma-54 specific transcriptional regulator"]
K11917	["sigma-54 specific transcriptional regulator"]
K10941	["sigma-54 specific transcriptional regulator, flagellar regulatory protein A"]
K07315	["sigma-B regulation protein RsbU (phosphoserine phosphatase)"]
K03597	["sigma-E factor negative regulatory protein RseA"]
K03598	["sigma-E factor negative regulatory protein RseB"]
K03803	["sigma-E factor negative regulatory protein RseC"]
K03101	["signal peptidase II [EC:3.4.23.36]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K05813	["sn-glycerol 3-phosphate transport system substrate-binding protein"]
K07284	["sortase A", "sortase A [EC:3.4.22.70]"]
K08600	["sortase B", "sortase B [EC:3.4.22.70]"]
K01117	["sphingomyelin phosphodiesterase [EC:3.1.4.12]"]
K07173	["S-ribosylhomocysteine lyase [EC:4.4.1.21]"]
K08258	["staphopain A [EC:3.4.22.48]"]
K13715	["staphopain B [EC:3.4.22.-]"]
K11040	["staphylococcal enterotoxin"]
K14198	["staphylokinase"]
K04047	["starvation-inducible DNA-binding protein"]
K11046	["streptolysin S associated protein"]
K12685	["subtilase-type serine protease [EC:3.4.21.-]"]
K11042	["superantigen-like protein"]
K04564	["superoxide dismutase, Fe-Mn family [EC:1.15.1.1]"]
K09693	["teichoic acid transport system ATP-binding protein [EC:3.6.3.40]"]
K09692	["teichoic acid transport system permease protein"]
K00912	["tetraacyldisaccharide 4'-kinase [EC:2.7.1.130]"]
K11018	["thermolabile hemolysin"]
K11031	["thiol-activated cytolysin"]
K12510	["tight adherence protein B"]
K12511	["tight adherence protein C"]
K12512	["tight adherence protein D"]
K12513	["tight adherence protein E"]
K12514	["tight adherence protein F"]
K12515	["tight adherence protein G"]
K11043	["toxic shock syndrome toxin-1"]
K11063	["toxin A/B"]
K10930	["toxin co-regulated pilin"]
K10932	["toxin co-regulated pilus biosynthesis outer membrane protein C"]
K10931	["toxin co-regulated pilus biosynthesis protein B"]
K10933	["toxin co-regulated pilus biosynthesis protein D"]
K10934	["toxin co-regulated pilus biosynthesis protein E"]
K10935	["toxin co-regulated pilus biosynthesis protein F"]
K10919	["toxin co-regulated pilus biosynthesis protein H"]
K10961	["toxin co-regulated pilus biosynthesis protein I"]
K10966	["toxin co-regulated pilus biosynthesis protein J [EC:3.4.23.43 2.1.1.-]"]
K10920	["toxin co-regulated pilus biosynthesis protein P"]
K10962	["toxin co-regulated pilus biosynthesis protein Q"]
K10963	["toxin co-regulated pilus biosynthesis protein R"]
K10964	["toxin co-regulated pilus biosynthesis protein S"]
K10965	["toxin co-regulated pilus biosynthesis protein T"]
K07165	["transmembrane sensor"]
K01046	["triacylglycerol lipase [EC:3.1.1.3]"]
K03545	["trigger factor"]
K02659	["twitching motility protein PilI"]
K02660	["twitching motility protein PilJ"]
K02669	["twitching motility protein PilT"]
K02670	["twitching motility protein PilU"]
K02657	["twitching motility two-component system response regulator PilG"]
K02658	["twitching motility two-component system response regulator PilH"]
K10943	["two component system, response regulator FlrC"]
K07707	["two-component system, AgrA family, response regulator AgrA"]
K07706	["two-component system, AgrA family, sensor histidine kinase AgrC [EC:2.7.13.-]"]
K03412	["two-component system, chemotaxis family, response regulator CheB [EC:3.1.1.61]"]
K03415	["two-component system, chemotaxis family, response regulator CheV"]
K03413	["two-component system, chemotaxis family, response regulator CheY"]
K03407	["two-component system, chemotaxis family, sensor kinase CheA [EC:2.7.13.3]"]
K08083	["two-component system, LytT family, response regulator AlgR"]
K08082	["two-component system, LytT family, sensor histidine kinase AlgZ [EC:2.7.13.3]"]
K07689	["two-component system, NarL family, invasion response regulator UvrY"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K07695	["two-component system, NarL family, response regulator DevR"]
K07690	["two-component system, NarL family, response regulator EvgA"]
K02479	["two-component system, NarL family, response regulator"]
K07688	["two-component system, NarL family, response regulator, fimbrial Z protein, FimZ"]
K07678	["two-component system, NarL family, sensor histidine kinase BarA [EC:2.7.13.3]"]
K07682	["two-component system, NarL family, sensor histidine kinase DevS [EC:2.7.13.3]"]
K07679	["two-component system, NarL family, sensor histidine kinase EvgS [EC:2.7.13.3]"]
K11384	["two-component system, NtrC family, response regulator AlgB"]
K02667	["two-component system, NtrC family, response regulator PilR"]
K02481	["two-component system, NtrC family, response regulator"]
K07710	["two-component system, NtrC family, sensor histidine kinase AtoS [EC:2.7.13.3]"]
K02668	["two-component system, NtrC family, sensor histidine kinase Pils [EC:2.7.13.3]"]
K02482	["two-component system, NtrC family, sensor kinase [EC:2.7.13.3]"]
K07669	["two-component system, OmpR family, response regulator MprA"]
K07660	["two-component system, OmpR family, response regulator PhoP"]
K07671	["two-component system, OmpR family, response regulator PrrA"]
K07776	["two-component system, OmpR family, response regulator RegX3"]
K07661	["two-component system, OmpR family, response regulator RstA"]
K02483	["two-component system, OmpR family, response regulator"]
K07653	["two-component system, OmpR family, sensor histidine kinase MprB [EC:2.7.13.3]"]
K07637	["two-component system, OmpR family, sensor histidine kinase PhoQ [EC:2.7.13.3]"]
K07655	["two-component system, OmpR family, sensor histidine kinase PrrB [EC:2.7.13.3]"]
K07639	["two-component system, OmpR family, sensor histidine kinase RstB [EC:2.7.13.3]"]
K07768	["two-component system, OmpR family, sensor histidine kinase SenX3 [EC:2.7.13.3]"]
K02484	["two-component system, OmpR family, sensor kinase [EC:2.7.13.3]"]
K10942	["two-component system, sensor histidine kinase FlrB [EC:2.7.13.3]"]
K07357	["type 1 fimbriae regulatory protein FimB"]
K07358	["type 1 fimbriae regulatory protein FimE"]
K04049	["type III secretion protein SctB"]
K03219	["type III secretion protein SctC"]
K03220	["type III secretion protein SctD"]
K04050	["type III secretion protein SctE"]
K03221	["type III secretion protein SctF"]
K04051	["type III secretion protein SctG"]
K04052	["type III secretion protein SctH"]
K04053	["type III secretion protein SctI"]
K03222	["type III secretion protein SctJ"]
K04054	["type III secretion protein SctK"]
K03223	["type III secretion protein SctL"]
K04056	["type III secretion protein SctO"]
K04057	["type III secretion protein SctP"]
K03225	["type III secretion protein SctQ"]
K03226	["type III secretion protein SctR"]
K03227	["type III secretion protein SctS"]
K03228	["type III secretion protein SctT"]
K03229	["type III secretion protein SctU"]
K03230	["type III secretion protein SctV"]
K04058	["type III secretion protein SctW"]
K04059	["type III secretion protein SctX"]
K04060	["type III secretion protein SctY"]
K08084	["type IV fimbrial biogenesis protein FimT"]
K08085	["type IV fimbrial biogenesis protein FimU"]
K02487	["type IV pili sensor histidine kinase and response regulator"]
K02650	["type IV pilus assembly protein PilA"]
K02652	["type IV pilus assembly protein PilB"]
K02653	["type IV pilus assembly protein PilC"]
K02655	["type IV pilus assembly protein PilE"]
K02656	["type IV pilus assembly protein PilF"]
K02661	["type IV pilus assembly protein PilK"]
K02662	["type IV pilus assembly protein PilM"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K02663	["type IV pilus assembly protein PilN"]
K02664	["type IV pilus assembly protein PilO"]
K02665	["type IV pilus assembly protein PilP"]
K02666	["type IV pilus assembly protein PilQ"]
K02671	["type IV pilus assembly protein PilV"]
K02672	["type IV pilus assembly protein PilW"]
K02673	["type IV pilus assembly protein PilX"]
K02674	["type IV pilus assembly protein PilY1"]
K02676	["type IV pilus assembly protein PilZ"]
K03194	["type IV secretion system protein VirB1"]
K03195	["type IV secretion system protein VirB10"]
K03196	["type IV secretion system protein VirB11"]
K03197	["type IV secretion system protein VirB2"]
K03198	["type IV secretion system protein VirB3"]
K03199	["type IV secretion system protein VirB4"]
K03200	["type IV secretion system protein VirB5"]
K03201	["type IV secretion system protein VirB6"]
K03202	["type IV secretion system protein VirB7"]
K03203	["type IV secretion system protein VirB8"]
K03204	["type IV secretion system protein VirB9"]
K03205	["type IV secretion system protein VirD4"]
K11919	["type VI secretion system lysozyme-related protein"]
K11902	["type VI secretion system protein ImpA"]
K11901	["type VI secretion system protein ImpB"]
K11900	["type VI secretion system protein ImpC"]
K11899	["type VI secretion system protein ImpD"]
K11898	["type VI secretion system protein ImpE"]
K11897	["type VI secretion system protein ImpF"]
K11896	["type VI secretion system protein ImpG"]
K11895	["type VI secretion system protein ImpH"]
K11894	["type VI secretion system protein ImpI"]
K11893	["type VI secretion system protein ImpJ"]
K11892	["type VI secretion system protein ImpK"]
K11891	["type VI secretion system protein ImpL"]
K11890	["type VI secretion system protein ImpM"]
K11889	["type VI secretion system protein ImpN [EC:2.7.11.1]"]
K11906	["type VI secretion system protein VasD"]
K11907	["type VI secretion system protein VasG"]
K11909	["type VI secretion system protein VasI"]
K11910	["type VI secretion system protein VasJ"]
K11911	["type VI secretion system protein VasL"]
K11905	["type VI secretion system protein"]
K11913	["type VI secretion system protein"]
K11918	["type VI secretion system protein"]
K11903	["type VI secretion system secreted protein Hcp"]
K11904	["type VI secretion system secreted protein VgrG"]
K03269	["UDP-2,3-diacetylglucosamine hydrolase [EC:3.6.1.-]"]
K13017	["UDP-3-keto-D-GlcNAcA aminotransferase [EC:2.6.1.-]"]
K02536	["UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase [EC:2.3.1.-]"]
K02535	["UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase [EC:3.5.1.-]"]
K13018	["UDP-D-GlcNAc3NA acetyltransferase [EC:2.3.1.-]"]
K13020	["UDP-D-GlcNAcA oxidase [EC:1.1.1.-]"]
K13016	["UDP-D-GlcNAcA oxidase [EC:1.1.1.-]"]
K01854	["UDP-galactopyranose mutase [EC:5.4.99.9]"]
K13019	["UDP-GlcNAc3NAcA epimerase [EC:5.1.3.23]"]
K10011	["UDP-GlcUA decarboxylase/UDP-L-Ara4N formyltransferase [EC:1.1.1.- 2.1.2.-], "UDP-4-amino-4-deoxy-L-arabinose formyltransferase / UDP-glucuronic acid dehydrogenase (UDP-4-keto-hexauronic acid decarboxylating) [EC:2.1.2.13 1.1.1.305]"]
K01784	["UDP-glucose 4-epimerase [EC:5.1.3.2]"]
K00012	["UDPglucose 6-dehydrogenase [EC:1.1.1.22]"]

(continued on next page)

Table 2 (continued)

#OTU ID	KEGG Description
K03279	["UDP-glucose:(galactosyl)LPS alpha-1,2-glucosyltransferase [EC:2.4.1.58]"]
K02844	["UDP-glucose:(heptosyl)LPS alpha-1,3-glucosyltransferase [EC:2.4.1.-]"]
K02474	["UDP-N-acetyl-D-galactosamine dehydrogenase [EC:1.1.1.-]"]
K13015	["UDP-N-acetyl-D-glucosamine dehydrogenase [EC:1.1.1.-]"]
K02472	["UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase [EC:1.1.1.-]"]
K01791	["UDP-N-acetylglucosamine 2-epimerase [EC:5.1.3.14]"]
K02473	["UDP-N-acetylglucosamine 4-epimerase [EC:5.1.3.7]"]
K00677	["UDP-N-acetylglucosamine acyltransferase [EC:2.3.1.129]"]
K00075	["UDP-N-acetylmuramate dehydrogenase [EC:1.1.1.158]"]
K01138	["uncharacterized sulfatase [EC:3.1.6.-]"]
K00806	["undecaprenyl pyrophosphate synthetase [EC:2.5.1.31]", "undecaprenyl diphosphate synthase [EC:2.5.1.31]"]
K02851	["undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase [EC:2.7.8.-]"]
K00996	["undecaprenyl-phosphate galactose phosphotransferase [EC:2.7.8.6]"]
K03188	["urease accessory protein"]
K03190	["urease accessory protein"]
K03189	["urease accessory protein"]
K03187	["urease accessory protein"]
K01428	["urease alpha subunit [EC:3.5.1.5]", "urease subunit alpha [EC:3.5.1.5]"]
K14048	["urease subunit gamma/beta [EC:3.5.1.5]"]
K02496	["uroporphyrin-III C-methyltransferase [EC:2.1.1.107]"]
K01599	["uroporphyrinogen decarboxylase [EC:4.1.1.37]"]
K01719	["uroporphyrinogen-III synthase [EC:4.2.1.75]"]
K00963	["UTP-glucose-1-phosphate uridylyltransferase [EC:2.7.7.9]"]
K04778	["vibriobactin synthetase"]
K08604	["vibriolysin [EC:3.4.24.25]"]
K03980	["virulence factor"]
K04784	["yersiniabactin nonribosomal peptide synthetase"]
K04786	["yersiniabactin nonribosomal peptide/polyketide synthase"]
K04783	["yersiniabactin salicyl-AMP ligase [EC:6.3.2.-]"]
K04785	["yersiniabactin synthetase, thiazolinyl reductase component"]
K05374	["yersiniabactin synthetase, thioesterase component"]
K08598	["YopJ protease family"]
K08643	["zinc metalloprotease ZmpB [EC:3.4.24.-]"]
K09815	["zinc transport system substrate-binding protein"]
K02074	["zinc/manganese transport system ATP-binding protein"]
K02075	["zinc/manganese transport system permease protein"]
K02077	["zinc/manganese transport system substrate-binding protein"]
K10954	["zona occludens toxin"]

2.2. 16S rRNA gene sequencing

Samples were vortexed to loosen the bacteria from the swab, and then centrifuged to form a bacteria pellet. Cell pellets were frozen at $-80\text{ }^{\circ}\text{C}$ within 24 h of collection and submitted for analysis at MRDNA/Molecular Research LP (Shallowater, TX.). There, genomic DNA was extracted from the cell pellets using DNeasy PowerSoil Extraction Kit (Qiagen, USA) according to manufacturer instructions. A single-step 30 cycle PCR was performed for the 16S rRNA gene V4 variable region using PCR primers 515F/806R (515F: 5'-GTGCCAGCMGCCGCGTAA-3'; and 806R: 5'-GGACTACVSGGGTATCTAAT-3'). PCR conditions were: $94\text{ }^{\circ}\text{C}$ for 3 minutes, followed by 28 cycles of $94\text{ }^{\circ}\text{C}$ for 30 seconds, $53\text{ }^{\circ}\text{C}$ for 40 seconds and $72\text{ }^{\circ}\text{C}$ for 1 minute, and a final elongation step at $72\text{ }^{\circ}\text{C}$ for 5 minutes. NextGen Sequencing (NGS) for the 16S rRNA gene V4 variable region was performed on an Ion Torrent PGM at MRDNA/Molecular Research LP [1].

2.3. Data analysis

The raw sequencing data were analyzed using QIIME version 1.9.1 [2]. Raw sequencing reads were demultiplexed using split_libraries.py. The sequences were filtered using a cut-off

quality score of 25 and were then clustered into an open reference operational taxonomic unit (OTU) table using `pick_open_reference_otus.py` and taxonomically classified. The sequences were filtered using `filter_taxa_from_otu_table.py` to remove chloroplasts, and diversity analyses were performed using `core_diversity_analyses.py`. These data were further analyzed using PICRUSt version 1.1.4 to predict functional profiles of the bacteria and compared to shotgun metagenomic sequencing to reveal changes in ARGs, VFGs and antibiotic biosynthesis genes (ABSGs) after swimming. The following PICRUSt scripts were used to acquire gene counts: `normalize_by_copy_number.py`, `predict_metagenomes.py`, `categorize_by_function.py` and `metagenome_contributions.py`. Weighted nearest sequenced taxon index (NSTI) scores for each sample were calculated to assess prediction accuracy using the `predict_metagenomes.py` with the `-a` option. In order to obtain OTU-specific gene counts for ARGs and VFGs, we used the `metagenome_contributions.py` script with `-l` option for each gene of interest detected in the predicted profiles [8,9]. These gene counts, referred to as KOs [10], were compared to previously published literature on ARGs [8,11] and virulence factors [12], and with the KO database to determine which ARGs (Table 1) and VFGs (Table 2) were detected in the samples.

Ethics Statement

Data collection methods were approved by the University of California, Irvine Institutional Review Board (IRB #2017-3751). Participants gave verbal consent to enroll.

CRedit Author Statement

Marisa Nielsen: Conception, Data Curation, Formal Analysis, Visualization, Investigation, Writing-Original Draft Preparation; **Sunny Jiang:** Conception, Writing-Reviewing and Editing, Investigation, Supervision. All authors read and approve the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

Acknowledgements

Funding for this research was partially provided by NSF CBET 1806066, 1804166 and UCI Oceans Fellowship.

References

- [1] MR DNA Laboratory. www.mrdnalab.com Shallowater, TX., (n.d.). <http://www.mrdnalab.com/>. Accessed December 6, 2018.
- [2] J.G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Peña, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Turnbaugh, W.A. Walters, J. Widmann, T. Yatsunencko, J. Zaneveld, R. Knight, QIIME allows analysis of high-throughput community sequencing data, *Nat. Methods*. 7 (2010) 335–336, doi:10.1038/nmeth.f.303.
- [3] M.G.I. Langille, J. Zaneveld, J.G. Caporaso, D. McDonald, D. Knights, J.A. Reyes, J.C. Clemente, D.E. Burkepile, R.L. Vega Thurber, R. Knight, R.G. Beiko, C. Huttenhower, Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences, *Nat. Biotechnol.* 31 (2013) 814–821, doi:10.1038/nbt.2676.
- [4] M. Nielsen, Human Skin Microbiome Data (16S rRNA sequencing) dataset #1, V1, Mendeley Data (2020), doi:10.17632/th7bfg6m.1.

- [5] M. Nielsen, Human Skin Microbiome Data (16S rRNA sequencing) dataset #2, V1, Mendeley Data (2020), doi:[10.17632/hjnfzm949y.1](https://doi.org/10.17632/hjnfzm949y.1).
- [6] M.C. Nielsen, N. Wang, S.C. Jiang, Acquisition of antibiotic resistance genes on human skin after swimming in the ocean, *Environ. Res.* 197 (2021) 110978, doi:[10.1016/j.envres.2021.110978](https://doi.org/10.1016/j.envres.2021.110978).
- [7] M.C. Nielsen, S.C. Jiang, Alterations of the human skin microbiome after ocean water exposure, *Mar. Pollut. Bull.* 145 (2019), doi:[10.1016/j.marpolbul.2019.06.047](https://doi.org/10.1016/j.marpolbul.2019.06.047).
- [8] X.-Y. Fan, J.-F. Gao, K.-L. Pan, D.-C. Li, H.-H. Dai, X. Li, Functional genera, potential pathogens and predicted antibiotic resistance genes in 16 full-scale wastewater treatment plants treating different types of wastewater, *Bioresour. Technol.* 268 (2018) 97–106, doi:[10.1016/j.biortech.2018.07.118](https://doi.org/10.1016/j.biortech.2018.07.118).
- [9] B. Li, Y. Yang, L. Ma, F. Ju, F. Guo, J.M. Tiedje, T. Zhang, Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes, *ISME J* 9 (2015) 2490–2502, doi:[10.1038/ismej.2015.59](https://doi.org/10.1038/ismej.2015.59).
- [10] M. Kanehisa, KEGG: Kyoto Encyclopedia of Genes and Genomes, *Nucleic Acids Res* 28 (2000) 27–30, doi:[10.1093/nar/28.1.27](https://doi.org/10.1093/nar/28.1.27).
- [11] Z.Z. Yan, Q.L. Chen, Y.J. Zhang, J.Z. He, H.W. Hu, Antibiotic resistance in urban green spaces mirrors the pattern of industrial distribution, *Environ. Int.* 132 (2019) 105106, doi:[10.1016/j.envint.2019.105106](https://doi.org/10.1016/j.envint.2019.105106).
- [12] C.K.S.M.B. Pattaroni, Virulence Factor Database for 16S metagenomics, 1 (2018), doi:[10.17632/RVXNTBTSCW.1](https://doi.org/10.17632/RVXNTBTSCW.1).