Genomics Data 10 (2016) 91–96

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Draft genome sequence of *Staphylococcus hominis* strain Hudgins isolated from human skin implicates metabolic versatility and several virulence determinants

Shelby Calkins [a], M.B. Couger [a], Colin Jackson [a], Jordan Zandler [a], Garett C. Hudgins [a], Radwa A. Hanafy [a], Connie Budd [a], Donald P. French [b], Wouter D. Hoff [a], Noha Youssef [c,*]

[a] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, United States
[b] Department of Integrative Biology, Oklahoma State University, Stillwater, OK, United States
[c] 1110 S Innovation way, Stillwater, OK 74074, United States

ARTICLE INFO

Article history:
Received 19 July 2016
Received in revised form 29 September 2016
Accepted 5 October 2016
Available online 7 October 2016

Keywords:
Staphylococcus hominis
Draft genome sequence
Detailed annotation
Student Initiated Microbial Discovery (SIMD) project
Metabolic versatility
Capsule

ABSTRACT

Staphylococcus hominis is a predominant member of the human skin microbiome. We here report on the genomic analysis of *Staphylococcus hominis* strain Hudgins that was isolated from the wrist area of human skin. The partial genome assembly of *S. hominis* Hudgins consists of 2,211,863 bp of DNA with 2174 protein-coding genes and 90 RNA genes. Based on the genomic analysis of KEGG pathways, the organism is expected to be a versatile heterotroph potentially capable of hydrolyzing the sugars glucose, fructose, mannose, and the amino acids alanine, aspartate, glutamate, glycine, threonine, cysteine, methionine, valine, isoleucine, leucine, lysine, arginine, phenylalanine, tyrosine, and tryptophan for energy production through aerobic respiration, with occasional lactate and acetate fermentation. Evidence for poly-gamma glutamate capsule and type IV Com system pili were identified in the genome. Based on COG analysis, the genome of *S. hominis* Hudgins clusters away from the previously published *S. hominis* genome ZBW5.

1. Introduction

Staphylococcus hominis strain Hudgins was isolated from the surface of human skin in Stillwater, OK as part of the Student Initiated Microbial Discovery (SIMD) project (introduced in [1]). The genus *Staphylococcus* is a phylogenetically and physiologically diverse genus with members ubiquitously found as part of the skin microbiome [2–4]. Infections have been reported in patients with lowered immunity [3,5,6]. Members of *Staphylococcus hominis* were previously isolated from human-associated [7] and animal-associated microbiomes [8], as well as from the environment [9]. Genomic analysis of strains belonging to the *Staphylococcus hominis* can contribute to our understanding of the molecular mechanisms of opportunistic pathogenesis. Such knowledge could potentially help reduce the occurrence and the severity of such infections in the future. Here we report on the draft genomic sequence, and the detailed annotation and analysis of *Staphylococcus hominis* strain Hudgins with an emphasis on its virulence factors.

2. Materials and methods

2.1. Genome sequencing information

2.1.1. Genome project history

The draft assembly and annotation were completed in 2015–2016. Table 1 shows the genome project information.

2.1.2. Growth conditions and genomic DNA preparation

S. hominis Hudgins was isolated from wrist skin on Tryptic soy agar (TSA) and repeatedly streaked (three times) to obtain a pure culture. To have enough biomass for DNA extraction, the strain was grown overnight at 30 °C on TSA plates. Genomic DNA of high sequencing quality was isolated using the MPBio PowerSoil® DNA extraction kit according to manufacturer's instructions. Negative stain TEM micrographs were obtained using the services of the Oklahoma State University Microscopy Lab. Briefly, the sample was placed on a carbon film TEM grid and allowed to incubate for 2 min, after which the excess liquid was wicked off. Phosphotungestic acid (PTA; 2% w/v) was then added to the grid followed by a 45-second incubation. Excess PTA was wicked off and the grid was allowed to dry before it was visualized using JOEL JEM-2100 transmission electron microscope.

* Corresponding author.
E-mail address: noha@okstate.edu (N. Youssef).

http://dx.doi.org/10.1016/j.gdata.2016.10.003

Table 1
Project information.

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft
MIGS-28	Libraries used	2 × 300 paired end chemistry
MIGS 29	Sequencing platforms	Illumina
MIGS 31.2	Fold coverage	300×
MIGS 30	Assemblers	Velvet
MIGS 32	Gene calling method	Prodigal
	Genbank ID	MAYR00000000
	GenBank Date of Release	July 2016
	GOLD ID	Gp0126757
	BIOPROJECT	PRJNA327386
MIGS 13	Project relevance	Environmental

2.1.3. Genome sequencing and assembly

The genome of *S. hominis* Hudgins was sequenced using the Illumina MiSeq platform at the University of Georgia Genomics Facility using 2 × 300 paired end chemistry and an average library insert size of 700 bp. Quality filtered sequence data were assembled with the short read de brujin graph assembly program Velvet [10]. The assembly settings were a kmer value of 101 bp and a minimum contig coverage value of 7×. The genome project is deposited in GOLD (Genomes On-Line Database) and this Whole Genome Shotgun (WGS) project has been deposited in GenBank under the accession MAYR00000000. The version described in this paper is version MAYR01000000.

2.1.4. Genome annotation

Gene models were created using the prokaryotic gene calling software package prodigal [11], and as a result a total of 2270 gene models were predicted with average gene size of 877 bp. Translated protein sequences were functionally annotated using a combination of NCBI Blast C++ homology search and HMMER 3.0 [12] hmmscan against the PFAM 26.0 database [13]. Additional gene analysis and functional annotation were carried out through the Integrated Microbial Genomes Expert Review (IMG-ER) platform.

2.2. Phylogenetic analysis

A maximum likelihood phylogenetic tree was constructed using multiple sequence alignments of 16S rRNA genes sequences. Multiple sequence alignment was conducted in Mega using ClustalW, as were the selection of the best substitution model, and the maximum likelihood analysis [14]. The tree was obtained under Kimura 2- parameter model with evolutionary rate difference among sites (+G, α shape = 0.1836). The substitution rate for transitions were 0.172, and for transversions were 0.039. *Escherichia coli* isolate ECSD9 was used as the outgroup. Bootstrap values, in percent, are based on 100 replicates.

2.3. Comparative genomics

Previous reports of genomic sequences from human skin-associated *Staphylococcus hominis* include strain ZBW5 [7]. We sought to compare the genome of *Staphylococcus hominis* strain Hudgins to several *Staphylococcus hominis* (including strain ZBW5) as well as *Staphylococcus haemolyticus* genomes ($n = 16$, IMG IDs: 2648501332, 2657245414, 2617271042, 2617271041, 2609460194, 2602041636, 2657245129, 642555160, 2548876781, 2582580999, 2648501918, 651285007, 2548877063, 2623620554, 2654587786, 643886039, 2540341024). We used the "Genome clustering" function on the IMG-ER analysis platform to conduct genomic comparisons based on the COG profile. We also used principal component analysis to compare the genomes based on several genomic features including the genome size, the number of genes, the number of transporters identified, the GC content, the number of non-coding bases, the number of genes belonging to COG categories, as well as the number of genes belonging to each COG category. The PCA analysis was conducted using the "princomp" function in the labdsv library of R [15]. The results were visualized using a biplot, where genomes were represented by stars, and genomic features or COG categories used for comparison were represented by arrows. Further comparative genomics were conducted on strain Hudgins genome and five other *S. hominis* genomes (strains C80, ZBW5, RIT-PI-K,

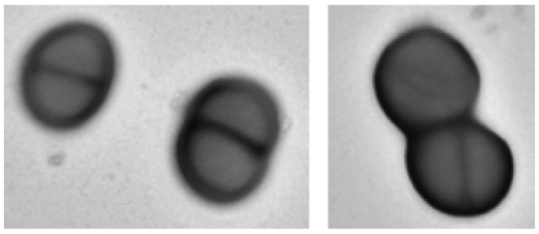

Fig. 1. Negative stain TEM micrographs of *S. hominis* Hudgins.

Table 2
S. hominis strain Hudgins 16S rRNA gene percentage similarity to other *Staphylococcus* species.

Staphylococcus species	Type strain	Hudgins strain % similarity
S. agnetis	DSM 23656^T	97.40%
S. argensis	DSM 29875^T	94.60%
S. argenteus	DSM 28299^T	97.90%
S. arlettae	ATCC 43957^T	97.80%
S. aureus	ATCC 12600^T	97.90%
S. auricularis	ATCC 33753^T	97.90%
S. capitis	ATCC 27840^T	98.20%
S. caprae	ATCC 35538^T	98.30%
S. carnosus	ATCC 51356^T	97.30%
S. caseolyticus	ATCC 13548^T	93.70%
S. chromogenes	ATCC 43764^T	97.60%
S. cohnii	ATCC 29974^T	98.00%
S. condimenti	DSM 11674^T	97.30%
S. delphini	ATCC 49171^T	97.10%
S. devriesei	LMG 25332^T	98.90%
S. epidermidis	ATCC 14990^T	98.50%
S. equorum	ATCC 43958^T	97.50%
S. felis	ATCC 49168^T	96.90%
S. fleurettii	DSM 13212^T	96.60%
S. gallinarum	ATCC 35539^T	97.80%
S. haemolyticus	ATCC 29970^T	98.70%
S. hominis	ATCC 27844^T	99.90%
S. hyicus	ATCC 11249^T	97.30%
S. intermedius	ATCC 29663^T	97.00%
S. jettensis	DSM 26618^T	99.20%
S. kloosii	ATCC 43959^T	97.90%
S. lentus	ATCC 29070^T	96.80%
S. lugdunensis	ATCC 43809^T	98.70%
S. lutrae	ATCC 700373^T	96.80%
S. massiliensis	5402776^T	96.70%
S. microti	DSM 22147^T	97.40%
S. muscae	ATCC 49910^T	97.80%
S. nepalensis	DSM 15150^T	97.40%
S. pasteuri	ATCC 51129^T	98.70%
S. petrasii	CCM 8418^T	99.10%
S. pettenkoferi	B3117^T	97.60%
S. piscifermentans	ATCC 51136^T	97.40%
S. pseudintermedius	ON 86^T	97.10%
S. pulvereri	ATCC 51698^T	94.20%
S. rostri	DSM 21968^T	97.30%
S. saccharolyticus	ATCC 14953^T	98.10%
S. saprophyticus	ATCC 15305^T	98.10%
S. schleiferi	ATCC 43808^T	97.10%
S. schweitzeri	DSM 28300^T	97.00%
S. sciuri	ATCC 29062^T	97.00%
S. simiae	CCM 51256^T	98.10%
S. simulans	ATCC 27848^T	97.20%
S. stepanovicii	CCM 7717^T	96.60%
S. succinus	ATCC 700337^T	97.70%
S. vitulinus	ATCC 3^T	96.80%
S. warneri	ATCC 27836^T	98.40%
S. xylosus	ATCC 29971^T.	98.00%

Table 3
Classification and general features of *Staphylococcus hominis* strain Hudgins [26].

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain Bacteria	TAS [27]
		Phylum Firmicutes	TAS [27]
		Class Bacilli	TAS [27]
		Order Bacillales	TAS [27]
		Family Staphylococcaceae	TAS [27]
		Genus *Staphylococcus*	TAS [27]
		Species *hominis*	TAS [27]
		(Type) strain: Hudgins	TAS [27]
	Gram stain	Positive	TAS [27]
	Cell shape	coccus	TAS [27]
	Motility	non-motile	TAS [27]
	Sporulation	Non-spore forming	TAS [27]
	Temperature range	Mesophile, 28–40 °C	TAS [27]
	Optimum temperature	35 °C	TAS [27]
	pH range; Optimum	Unknown	
	Carbon source	ND	
MIGS-6	Habitat	skin	
MIGS-6.3	Salinity	Up to 7.5%	TAS [27]
MIGS-22	Oxygen requirement	Facultative anaerobe	TAS [27,28]
MIGS-15	Biotic relationship	free-living	IDA
MIGS-14	Pathogenicity	ND	
MIGS-4	Geographic location	Stillwater, OK	IDA
MIGS-5	Sample collection	March 2015	IDA
MIGS-4.1	Latitude	36.1157	IDA
MIGS-4.2	Longitude	-97.0586	IDA
MIGS-4.4	Altitude	1 M	IDA

[a] Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [29].

SK119, and VCU122) and included computing the genomic average nucleotide identity (gANI), alignment fraction (AF) [16], and bidirectional best hits using the nucleotide similarity scanner NSimScan [17], as well as gene homology comparisons using Blastn.

3. Results and discussion

3.1. Classification and features

Cells of strain Hudgins appear to be Gram positive, non-motile aerobic cocci that were arranged in tetrads, pairs, as well as singles and clusters (Fig. 1). Colonies on TSA agar were beige in color.

Within the genus *Staphylococcus*, 52 species are described with validly published names. Strain Hudgins shares 93.7–99.9% 16S rRNA gene identities with other species in the *Staphylococcus* genus (Table 2). When compared to other *Staphylococcus hominis* strains with sequenced genomes, strain Hudgins shares 100% 16S rRNA gene similarity with *S. hominis* strain C80 (Genbank accession number ACRM01000000), and 99% similarity with strains H69 (Genbank accession number LVVO00000000), RIT-PI-K [18], ZBW5 [7], and ShAs1, ShAs2, and ShAs3 [8].

Phylogenetic analysis based on the 16S rRNA gene placed *Staphylococcus hominis* strain CV21 as the closest taxonomic relative of *Staphylococcus hominis* strain Hudgins (Table 3, and Fig. 2).

3.2. Genome properties

The genome assembly process produced a contig N50 of 1,549,674 bp with a total genome size of 2,211,863 bp. The GC content was 31.31%. There are 96 RNA genes including 15 ribosomal RNA and 60 tRNA genes. Since all rRNA operons identified in the genome were either located at the start or the end of a contig, the organization of the full ribosomal RNA operon could not be deciphered from the existing assembly. However in some of these operons the 5S and 23S rRNA genes were followed by several tRNA genes. Of the 2270 genes detected, 2174 were protein-coding, of which 79.07% had a function prediction, 73.44% represented a COG functional category, and 2.69% were predicted to have a signal peptide. PSORT [19] classified proteins as 51.3% cytoplasmic, 0.015% extracellular, and 25.99% associated with the membrane. Based on the presence of 139 single copy genes [20], the genome is predicted to be 84.9% complete. Genome statistics are shown in Table 4. The distribution of genes into COG functional categories is shown in Table 5.

3.3. Insights from the genome sequence

Genome analysis of *S. hominis* Hudgins identified a microorganism/bacterium with a typical Gram positive cell wall structure. We identified genes encoding for the biosynthesis of the polar lipid cardiolipin in the genome. Further genomic analysis suggested that strain Hudgins is a metabolically versatile microorganism with the capability to utilize glucose, fructose, and mannose, alanine, asparagine, glutamine, glycine, serine, cysteine, and proline as carbon and energy sources based on the presence of complete catabolic KEGG pathways in the genome. Aerobic respiration capability was evidenced by the presence of a complete

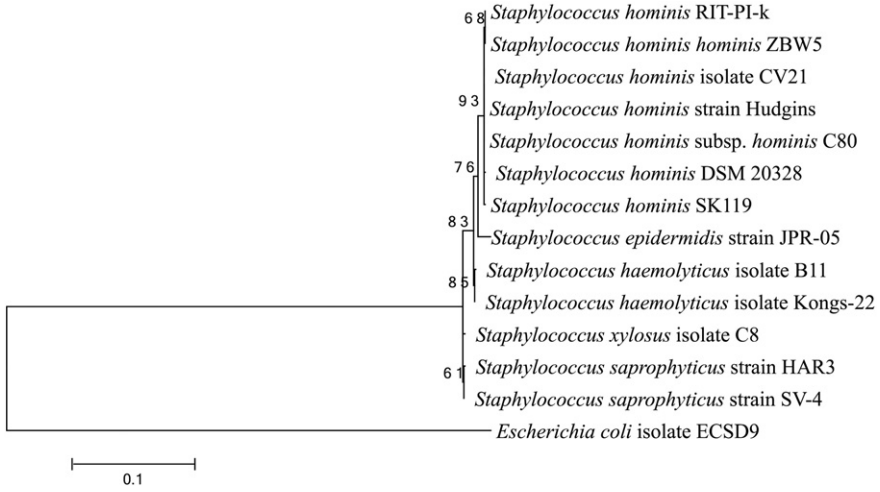

Fig. 2. Maximum likelihood phylogenetic tree constructed using multiple sequence alignments of 16S rRNA genes. *S. hominis* Hudgins is shown in bold. Reference sequences are also shown and Genbank accession numbers are given in parentheses. The tree was obtained under Kimura 2- parameter model with evolutionary rate difference among sites (+G, α shape = 0.1836). The substitution rate for transitions were 0.172, and for transversions were 0.039. *Escherichia coli* was used as the outgroup. Bootstrap values, in percent, are based on 100 replicates and are shown for branches with >50% bootstrap support. Multiple sequence alignment, model selection, and maximum likelihood analysis were carried out in Mega [14].

Table 4
Genome statistics.

Attribute	Value	% of total
Genome size (bp)	2,211,863	100.00%
DNA coding (bp)	1,916,307	86.64%
DNA G + C (bp)	692,578	31.31%
DNA scaffolds	14	100.00%
Total genes	2270	100.00%
Protein coding genes	2174	95.77%
RNA genes	96	4.23%
Pseudo genes	0	0
Genes in internal clusters	417	18.37%
Genes with function prediction	1795	79.07%
Genes assigned to COGs	1669	73.52%
Genes with Pfam domains	1878	82.73%
Genes with signal peptides	61	2.69%
Genes with transmembrane helices	556	24.49%
CRISPR repeats	1	

TCA cycle and electron transport chain with P/V/-type ATPase subunits confirming the aerobic nature of the microorganism. Facultative fermentation capability was also identified in the genome where genes for lactate and fermentation were present. Genomic analysis suggested auxotrophy for asparagine, serine, proline, histidine, vitamin B6, biotin, and lipoic acid. However, 113 ABC and 115 secondary transporters were identified in the genome via comparison of the protein-coding genes against the transporter database [21]. Some of the identified transporters could potentially address these auxotrophies. The above results were not based on biochemical tests and are only suggested by the genomic data available.

Genomic analysis also revealed the absence of genes encoding a complete flagellar assembly in line with the electron micrographs of the organism (Fig. 1). Other structures suggested by genomic analysis that could potentially contribute to the virulence of strain Hudgins include the type IV pili Com operon (*comGA-GB-GC-GD-GE-GF*) known to exist in Gram positive bacteria [22] in addition to homologues of the Gram positive sortase that would potentially be employed for activation and pilin polymerization [22,23]. The genome also harbors evidence for the biosynthesis of a poly-gamma glutamate capsule where a *cap* locus was identified in the genome with homologues of the four essential genes *capB-capC-capA-capE* [24]. The poly-gamma glutamate capsule is likely tethered to the cell wall based on the presence of the anchorage gene *capD* encoding for a gamma-glutamyl transpeptidase [24]. The halo around strain Hudgins cells in Fig. 1 could potentially be due to the presence of such capsule. Several other virulence factors are possible as well since comparison of the Hudgins genome against the virulence factor database [25] identified 475 virulence factor hits (21.8% of the protein-coding genes).

Table 5
Number of genes associated with general COG functional categories.

Code	Value	%age	Description
J	185	9.94%	Translation, ribosomal structure and biogenesis
A	0	0	RNA processing and modification
K	125	6.72%	Transcription
L	93	5%	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	26	1.4%	Cell cycle control, Cell division, chromosome partitioning
V	36	1.93%	Defense mechanisms
T	64	3.44%	Signal transduction mechanisms
M	97	5.21%	Cell wall/membrane biogenesis
N	4	0.21%	Cell motility
U	18	0.97%	Intracellular trafficking and secretion
O	71	3.82%	Posttranslational modification, protein turnover, chaperones
C	100	5.37%	Energy production and conversion
G	117	6.29%	Carbohydrate transport and metabolism
E	169	9.08%	Amino acid transport and metabolism
F	78	4.19%	Nucleotide transport and metabolism
H	124	6.66%	Coenzyme transport and metabolism
I	76	4.08%	Lipid transport and metabolism
P	134	7.2%	Inorganic ion transport and metabolism
Q	35	1.88%	Secondary metabolites biosynthesis, transport and catabolism
R	165	8.87%	General function prediction only
S	132	7.09%	Function unknown
–	601	26.48%	Not in COGs

The total is based on the total number of protein coding genes in the genome.

3.4. Insights from comparative genomics

We compared the genome of *Staphylococcus hominis* strain Hudgins to several *Staphylococcus hominis* (including the previously sequenced *Staphylococcus hominis* strain ZBW5) as well as *Staphylococcus haemolyticus* genomes. Based on their COG profile, the genomes clustered in agreement with their phylogeny. However, strain Hudgins genome clustered away from the ZBW5 genome (Fig. 3A). A closer look at the COG function profile of *S. hominis* strain Hudgins in comparison to only *S. hominis* strains is shown in Table S1. Further analyses using genomic features (including the genome size, the number of genes, the number of transporters identified, the GC content, the number of non-coding bases, the number of genes belonging to COG categories, as well as the number of genes belonging to each COG category) cluster *Staphylococcus hominis* strain Hudgins genome close to *S. hominis* genomes SK119, RIT-PI-K, C80 and VCU122 and away from the ZBW5 genome, likely due to the enrichment in the number of transporters identified in the former genomes in comparison to the ZBW5 genome (Fig. 3B). Compared to other *S. hominis* genomes, strain Hudgins genome shared 97.8–99.4% genome average nucleotide identity and 0.89–0.93 alignment fraction. Using NSimScan, 1920–1986 bidirectional hits were identified representing 84.6–87.5% of the total genes in the Hudgins genome (Table S2). These data are typical of strains belonging to the same species [16,17]. Blastn gene homology comparisons also revealed that the majority of the genes in the Hudgins genome (84–87% of the total number of genes) were ≥95% similar to genes in other *S. hominis* genomes (Table S3).

4. Conclusions

This study presents the draft genome sequence and annotation of *Staphylococcus hominis* strain Hudgins. The genome revealed extensive sugar and amino acid degradation machinery. Comparison to the virulence factor database identified 475 genes in the genome with potential virulence-associated function including type IV Com system for pili production, and the potential for the secretion of a membrane associated poly-gamma-glutamate capsule. Comparative genomics using general genomic features as well as the COG function profile coincided with the phylogenetic topology but placed strain Hudgins away from the previously published *S. hominis* genome ZBW5.

Competing interests

All authors declare no competing interests.

Funding

Staphylococcus hominis strain Hudgins was selected for sequencing as part of a project at Oklahoms State University funded by a grant from the Howard Hughes Medical Institute through the Science Education Program. The Student Initiated Microbial Discovery (SIMD) project is an undergraduate retention program that was introduced in 2015 with the aim of engaging undergraduates in authentic lab research. As part of the project, we encourage students to participate in a two--semester long research endeavor where undergraduate students isolate an environmental strain during the first semester in an introductory

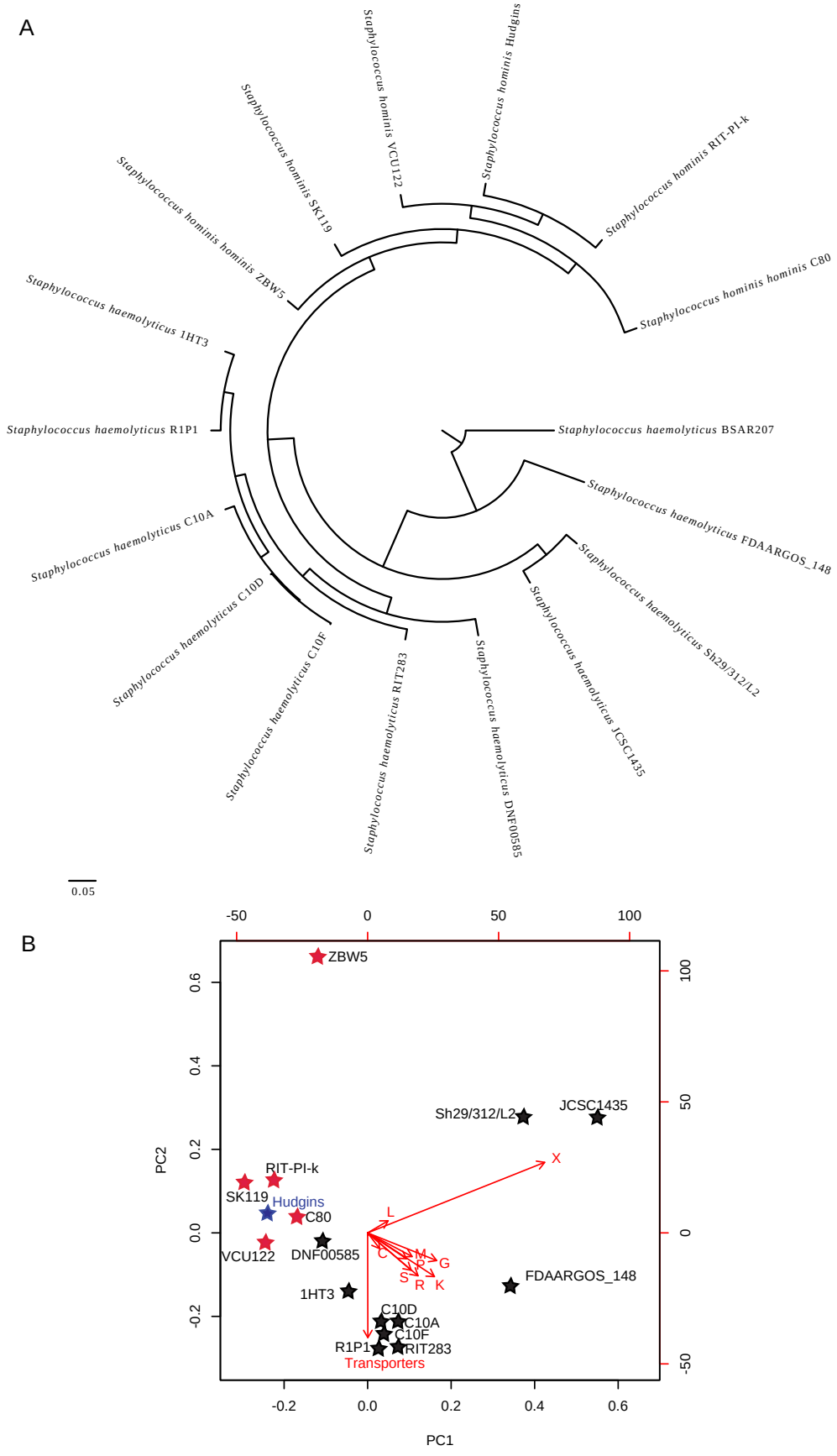

Fig. 3. (A) COG profile clustering of the genomes compared in this study. (B) Principal component analysis biplot of the genomic features and COG category distribution in the genomes compared. Genomes are represented by stars and are depicted by strain name. Arrows represent genomic features or COG categories used for comparison. The arrow directions follow the maximal abundance, and their lengths are proportional to the maximal rate of change between genomes. The first two components explained 75% of variation. Color coding: *S. haemolyticus*, black; *S. hominis*, red. Strain Hudgins is shown in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

microbiology course, followed by extraction of its genomic DNA and analysis of its genome sequence during the second semester in an upper division microbial genomics course. Here, strain Hudgins was isolated by an undergraduate student (GCH) and its genome was analyzed by a team of undergraduate (CJ and JZ) and graduate (SC) students. This is Draft Genome #8 in the SIMD project. WDH acknowledges support by NSF grants MCB-1051590, MRI-1338097, and CHE-1412500.

Authors' contributions

SC, CJ, JZ, MBC, and NY contributed to the analysis. SC, WDH, DPF, and NY wrote the manuscript. GCH, CB, and RAH performed the lab experiments.

Transparency document

The Transparency document associated to this article can be found in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2016.10.003.

References

[1] M.B. Couger, A. Hurlbut, C.L. Murphy, C. Budd, D.P. French, W.D. Hoff, M.S. Elshahed, Y. NH, Draft Genome Sequence of the Environmental Isolate Chryseobacterium sp. Hurlbut01. Genome Announc. 3 (5) (2015).
[2] C. Callewaert, F.-M. Kerckhof, M.S. Granitsiotis, M. Van Gele, T. Van de Wiele, N. Boon, Characterization of *Staphylococcus* and *Corynebacterium* clusters in the human axillary region. PLoS One 8 (8) (2013), e70538.
[3] R. Coates, J. Moran, M.J. Horsburgh, Staphylococci: colonizers and pathogens of human skin. Future Microbiol 9 (1) (2014) 75–91.
[4] P.M. Swe, M. Zakrzewski, A. Kelly, L. Krause, K. Fischer, Scabies mites Alter the skin microbiome and promote growth of opportunistic pathogens in a porcine model. PLoS Negl. Trop. Dis. 8 (5) (2014), e2897.
[5] K. Dansey, L. Stratton, B.D. Park, Staphylococcus Hominis carotid artery infection with septic embolization. J.Vasc. Surg. Cases 1 (2) (2015) 81–83.
[6] F. Chaves, M. García-Álvarez, F. Sanz, C. Alba, J.R. Otero, Nosocomial spread of a Staphylococcus Hominis subsp. Novobiosepticus strain causing sepsis in a neonatal intensive care unit. J. Clin. Microbiol. 43 (9) (2005) 4877–4879.
[7] S. Jiang, B. Zheng, W. Ding, L. Lv, J. Ji, H. Zhang, Y. Xiao, L. Li, Whole-genome sequence of *Staphylococcus hominis*, an opportunistic pathogen. J. Bacteriol. 194 (17) (2012) 4761–4762.
[8] G.L. Hughes, J.A. Raygoza Garay, V. Koundal, J.L. Rasgon, M. MM, Genome Sequences of Staphylococcus hominis Strains ShAs1, ShAs2, and ShAs3, Isolated from the Asian Malaria Mosquito *Anopheles stephensi*. Genome Announc. 4 (2) (2016).
[9] K. Marimuthu, Isolation and characterization of *Staphylococcus hominis* JX961712 from oil contaminated soil. J. Pharm. Res. 7 (3) (2013) 252–256.
[10] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18 (5) (2008) 821–829.
[11] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 11 (2010) 119.
[12] J. Mistry, R.D. Finn, S.R. Eddy, A. Bateman, M. Punta, Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41 (12) (2013), e121.
[13] R.D. Finn, P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, et al., The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res. 44 (D1) (2016) D279–D285.
[14] S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7) (2016) 1870–1874.
[15] D. Roberts, labdsv: Ordination and Multivariate Analysis for Ecology. 2007.
[16] N.J. Varghese, S. Mukherjee, N. Ivanova, K.T. Konstantinidis, K. Mavrommatis, N.C. Kyrpides, A. Pati, Microbial species delineation using whole genome sequences. Nucleic Acids Res. (2015).
[17] V. Novichkov, A. Kaznadzey, N. Alexandrova, D. Kaznadzey, NSimScan: DNA comparison tool with increased speed, sensitivity and accuracy. Bioinformatics (2016).
[18] P.N. Tran, N.E. Tan, Y.P. Lee, H.M. Gan, S.J. Polter, L.K. Dailey, A.O. Hudson, M.A. Savka, Whole-Genome Sequence and Classification of 11 Endophytic Bacteria from Poison Ivy (*Toxicodendron radicans*). Genome Announc. 3 (6) (2015).
[19] P. Horton, K.-J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier, K. Nakai, WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35 (Web Server issue) (2007) W585–W587.
[20] C. Rinke, P. Schwientek, A. Sczyrba, N.N. Ivanova, I.J. Anderson, J.F. Cheng, A. Darling, S. Malfatti, B.K. Swan, E.A. Gies, et al., Insights into the phylogeny and coding potential of microbial dark matter. Nature 499 (7459) (2013) 431–437.
[21] M.H. Saier Jr., V.S. Reddy, D.G. Tamang, A. Vastermark, The transporter classification database. Nucleic Acids Res. 42 (Database issue) (2014) D251–D258.
[22] S. Imam, Z. Chen, D.S. Roos, M. Pohlschröder, Identification of surprisingly diverse type IV pili, across a broad range of Gram-positive bacteria. PLoS One 6 (12) (2011), e28919.
[23] A.P. Hendrickx, J.M. Budzik, S.Y. Oh, O. Schneewind, Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. Nat. Rev. Microbiol. 9 (3) (2011) 166–176.
[24] T. Candela, A. Fouet, Poly-gamma-glutamate in bacteria. Mol. Microbiol. 60 (5) (2006) 1091–1098.
[25] L. Chen, J. Yang, J. Yu, Z. Yao, L. Sun, Y. Shen, Q. Jin, VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33 (Database issue) (2005) D325–D328.
[26] D. Field, G. Garrity, T. Gray, N. Morrison, J. Selengut, P. Sterk, T. Tatusova, N. Thomson, M.J. Allen, S.V. Angiuoli, et al., The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. 26 (5) (2008) 541–547.
[27] W.E. Kloos, C.G. George, J.S. Olgiate, L. Van Pelt, M.L. McKinnon, B.L. Zimmer, E. Muller, M.P. Weinstein, S. Mirrett, *Staphylococcus hominis* subsp. novobiosepticus subsp. nov., a novel trehalose- and N-acetyl-D-glucosamine-negative, novobiocin- and multiple-antibiotic-resistant subspecies isolated from human blood cultures. Int. J. Syst. Bacteriol. 48 (Pt 3) (1998) 799–812.
[28] W.E. Kloos, K.H. Schleifer, Isolation and characterization of *Staphylococci* from human skin II. Descriptions of four new species: *Staphylococcus warneri*, *Staphylococcus capitis*, *Staphylococcus hominis*, and *Staphylococcus simulans*. Int. J. Syst. Evol. Microbiol. 25 (1) (1975) 62–79.
[29] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25 (1) (2000) 25–29.