Genomic and phenotypic characterization of *Epilithonimonas diehli* sp. nov., from a freshwater creek

Kyle T. Jacobs, Jessica Hoffman, Jeffrey D. Newman, Lycoming College, Williamsport, PA, USA

Abstract:

A Gram-Staining-Negative, yellow-orange pigmented bacterial strain, designated FH1 was isolated from Fox Hollow in Williamsport, PA during an undergraduate microbiology course. The 16S rRNA sequences of strain FH1 was most similar to that of *Epilithonimonas lactis* (98.1%) and *Epilithonimonas ginsengisoli* (97.1%) in the family *Flavobacteriaceae*. The genomes of FH1 and *E. lactis* were sequenced, assembled, annotated and compared to each other and the type species of the genus, *E. tenax*. The estimated DNA-DNA hybridization value between FH1 and E. lactis was 33.1 as calculated by the DSMZ Genome-Genome Distance Calculator. The Average Nucleotide Identity (ANI) was 85.6. Both values are well below the threshold for separate species. Comparison of annotated gene sets identified. The 4.0 Mbp genome of FH1 contained 632 genes not found in either of the other *Epilithonimonas* sequenced genomes. Phenotypic comparisons identified a several distinguishing characteristics that support the hypothesis that isolate FH1 represents a novel species in the genus *Epilithonimonas* for which the name *Epilithonimonas diehli* sp. nov. is proposed.

f Epilithonimonas sp. FH1 Edit
nce :1490bp View query sequence
ess :99.9% (2 ~ 1493)

Rank	Name	Strain	Authors	Accession	Pairwise Similarity (%)	Diff/Total nt
1	Epilithonimonas lactis	LMG 24401(T)	Shaked et al. 2010	JPLY01000001	98.14	27/1454
2	Epilithonimonas ginsengisoli	DCY78(T)	Hoang et al. 2015	JX827729	97.13	41/1429
3	Epilithonimonas tenax	DSM 16811(T)	O'Sullivan et al. 2006	AUAA01000078	96.97	44/1454
4	Epilithonimonas xixisoli	S31(T)	Feng et al. 2014	KJ150596	96.56	50/1454
5	Chryseobacterium humi	ECP37(T)	Pires et al. 2010	EU360966	95.87	60/1454
6	Chryseobacterium xinjiangense	TSBY 67(T)	Zhao et al. 2011	DQ166169	95.87	60/1454
7	Chryseobacterium formosense	LMG 24722(T)	Young et al. 2005	JPRP01000004	95.74	62/1454
8	Chryseobacterium hominis	NF802(T)	Vaneechoutte et al. 2007	AM261868	95.60	64/1453

Figure 1. EZ-Taxon 16S rRNA pairwise similarity for FH1 (Kim et al 2012)

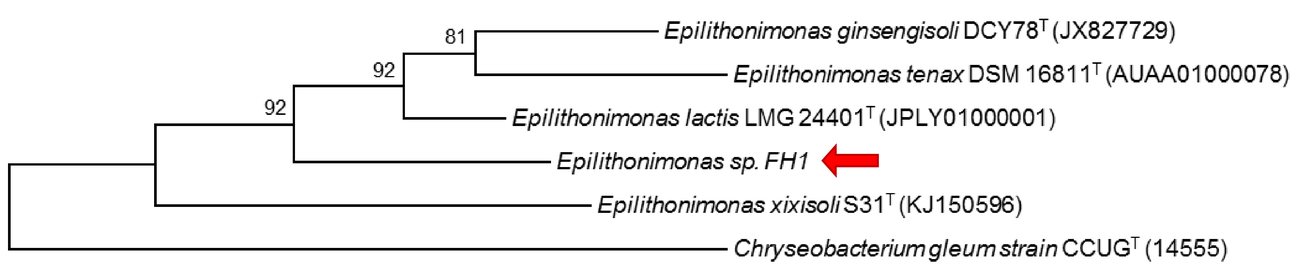

Figure 2. Neighbor Joining Tree of nearly full length 16s rRNA sequences

Genomic Comparisons:

	Table 4. Genome Sequencing Statistics (MiSeq v3 2x300 PE)	
	Epilithonimonas sp. FH1	*Epilithonimonas lactis*
Total Reads	1,366,100	1,394,644
Contigs	22	12
Average Coverage	93X	92X
Assembly Length	3,975,884 bp	4,236,390 bp
Accession	JPLZ01	JPLY01

Table 5. Estimated DNA-DNA Hybridization (species threshold = 70%) calculated by DSMZ Genome-Genome Distance Calculator (GGDC), Average Amino Acid Identify (AAI) calculated by Newman Lab Calculator, Average Nucleotide Identity (ANI) (species threshold = 95%) calculated by Kostas Lab Calculator, and Reciprocal Orthology Score Average (ROSA) calculated by Newman Lab ROSA calculator

					Estimated DNA-DNA Hybridization (eDDH)				
	16s	ANI	ROSA		1	2	3	4	5
Epilithonimonas diehli sp. FH1				1		33.1	26.2	22.4	20.1
Epilithonimonas lactis DSM 19921T	98.1	88.3	59.2	2	88.4		27.8	24.9	23.0
Epilithonimonas tenax DSM 16811T	97.0	83.3	51.9	3	85.4	86.7		23.7	22.1
Chryseobacterium aquaticum KCTC 12483T	95.2	79.6	33.3	4	69.9	70.5	70.6		21.7
Chryseobacterium gleum ATCC 35910T	95.0	78.0	28.3	5	68.1	68.9	70.3	77.6	
					Average Amino Acid Identity (AAI)				

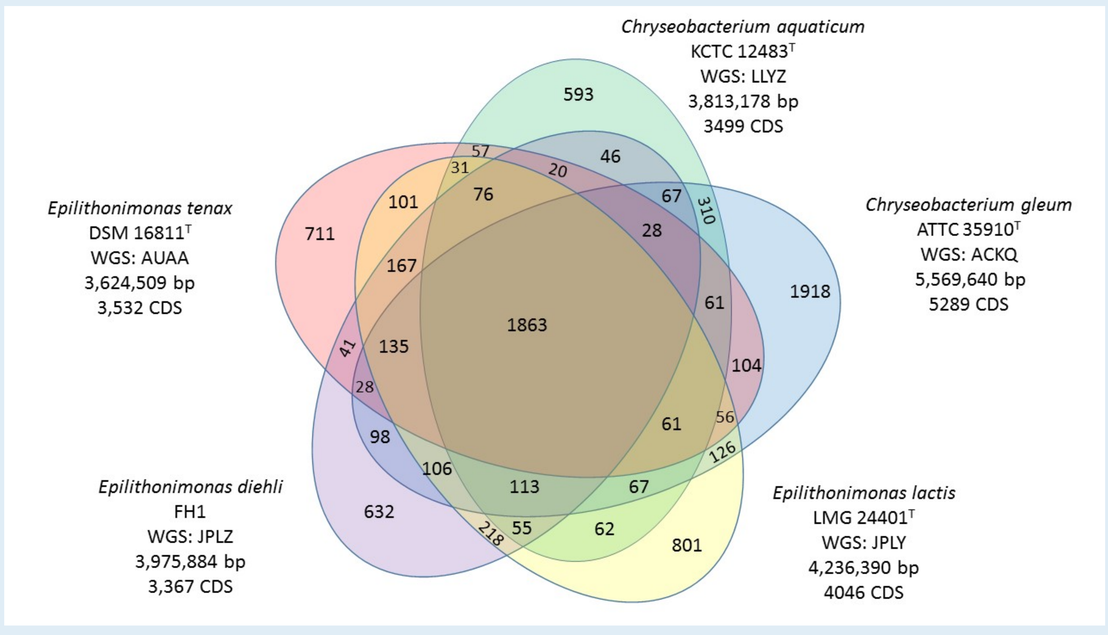

Figure 4. Venn Diagram comparing unique and shared genes between *Epilithonimonas diehli* and reference species

Phenotypic Analysis:

Table 1. MIDI-FAME

Fatty Acid Methyl Ester	FH1	*E. lactis*	*E. tenax*
iso-$C_{13:0}$	0.5	1.5	0.7
$C_{14:0}$	0.6	1.2	1.3
iso-$C_{15:0}$	19.1	18.1	15.1
anteiso-$C_{15:0}$	10.7	13.1	10.4
Summed Feature 2	0.4	0.7	1.1
Summed Feature 3	28.1	22.6	30.1
$C_{16:1}\omega5c$	1.2	0.9	1.2
$C_{16:0}$	8.4	9.7	10.8
iso-$C_{15:0}$ 3-OH	2.4	2.0	1.9
$C_{15:0}$ 2-OH	1.1	1.1	1.2
Summed Feature 9	1.1	1.7	2.1
anteiso-$C_{17:1}\omega9c$	2.8	3.4	1.4
iso$C_{16:0}$ 3-OH	1.5	2.0	1.8
$C_{16:0}$ 3-OH	6.1	7.5	7.9
iso$C_{17:0}$ 3-OH	12.7	9.7	7.5
$C_{17:0}$ 2-OH	ND	1.7	1.5

Table 2: HPLC Pigment Comparison

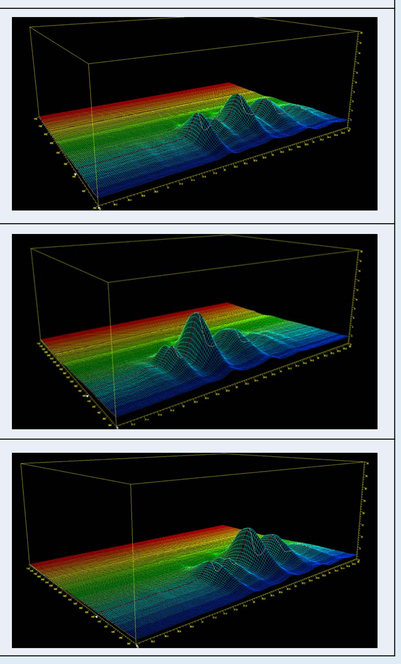

Table 3. BIOLOG GEN III

Well	Carbon Source/Condition	Epilithonimonas FH1	Epilithonimonas lactis	Epilithonimonas tenax	Well	Carbon Source/Condition	Epilithonimonas FH1	Epilithonimonas lactis	Epilithonimonas tenax
A01	neg control	41	30	26	G01	p-hydroxy-phenylacetic acid	12	16	14
A02	dextrin	100	99	100	G02	methyl pyruvate	87	95	12
A03	D-maltose	99	99	99	G03	D-lactic acid methyl ester	14	15	11
A04	D-trehalose	22	99	99	G04	L-lactic acid	15	13	10
A05	D-cellobiose	21	25	17	G05	citric acid	8	20	10
A06	gentiobiose	99	99	100	G06	a-keto-glutaric acid	8	19	10
A07	sucrose	28	22	100	G07	D-malic acid	8	11	10
A08	D-turanose	9	8	16	G08	L-malic acid	15	15	10
A09	stachyose	21	15	20	G09	bromo-succinic acid	8	6	10
A10	pos control	98	97	99	G10	nalidixic acid	44	11	13
A11	pH 6	98	96	98	G11	LiCl	10	11	14
A12	pH 5	90	94	29	G12	K-tellurite	97	93	20
B01	D-raffinose	14	12	20	H01	tween-40	96	98	89
B02	a-D-lactose	21	99	16	H02	g-amino-butyric acid	8	10	12
B03	D-melibiose	18	17	17	H03	a-hydroxy-butyric acid	11	12	12
B04	b-methyl-D-glucoside	14	9	14	H04	b-hydroxy-D,L-butyric acid	14	17	11
B05	D-salicin	99	98	12	H05	a-keto-butyric acid	8	6	10
B06	N-acetyl-D-glucosamine	9	8	10	H06	acetoacetic acid	19	86	10
B07	N-acetyl-b-D-mannosamine	9	9	12	H07	propionic acid	10	22	10
B08	N-acetyl-D-galactosamine	23	11	15	H08	acetic acid	93	98	99
B09	N-acetyl neuraminic acid	28	19	18	H09	formic acid	10	14	10
B10	1% NaCl	96	96	96	H10	aztreonam	10	16	98
B11	4% NaCl	16	11	27	H11	Na-butyrate	16	97	14
B12	8% NaCl	19	10	30	H12	Na bromate	12	14	15
C01	a-D-glucose	99	99	97	E01	gelatin	100	100	24
C02	D-mannose	8	6	97	E02	glycyl-L-proline	92	96	75
C03	D-fructose	98	16	10	E03	L-alanine	8	28	10
C04	D-galactose	17	98	13	E04	L-arginine	23	43	67
C05	3-methyl glucose	8	7	12	E05	L-aspartic acid	33	62	13
C06	D-fucose	15	19	11	E06	L-glutamic acid	99	98	99
C07	L-fucose	25	15	14	E07	L-histidine	11	34	10
C08	L-rhamnose	98	98	12	E08	L-pyroglutamic acid	11	10	10
C09	inosine	8	20	18	E09	L-serine	8	6	10
C10	1% Na-lactate	94	95	96	E10	lincomycin	10	62	12
C11	fusidic acid	10	9	13	E11	guanidine HCl	91	90	14
C12	D-serine	10	11	15	E12	niaproof 4	10	10	16
D01	D-sorbitol	16	16	18	F01	pectin	99	78	95
D02	D-mannitol	13	14	13	F02	D-galacturonic acid	99	98	100
D03	D-arabitol	13	12	11	F03	L-galacturonic acid lactone	8	6	12
D04	myo-inositol	26	23	13	F04	D-gluconic acid	19	23	12
D05	glycerol	10	13	10	F05	D-glucuronic acidi	34	28	11
D06	D-glucose-6-PO4	37	38	18	F06	glucuronamide	23	17	15
D07	D-fructose-6-PO4	34	13	14	F07	mucic acid	8	17	10
D08	D-aspartic acid	8	6	10	F08	quinic acid	12	17	10
D09	D-serine	8	6	10	F09	D-saccharic acid	12	17	14
D10	troleandomycin	10	8	11	F10	vancomycin	10	8	11
D11	rifamycin SV	14	10	14	F11	tetrazolium violet	44	38	99
D12	minocycline	11	13	17	F12	tetrazolium blue	99	99	97
E01	gelatin	100	100	24	G01	p-hydroxy-phenylacetic acid	12	16	14
E02	glycyl-L-proline	92	96	75	G02	methyl pyruvate	87	95	12
E03	L-alanine	8	28	10	G03	D-lactic acid methyl ester	14	15	11
E04	L-arginine	23	43	67	G04	L-lactic acid	15	13	10
E05	L-aspartic acid	33	62	13	G05	citric acid	8	20	10
E06	L-glutamic acid	99	98	99	G06	a-keto-glutaric acid	8	19	10
E07	L-histidine	11	34	10	G07	D-malic acid	8	11	10
E08	L-pyroglutamic acid	11	10	10	G08	L-malic acid	15	15	10
E09	L-serine	8	6	10	G09	bromo-succinic acid	8	6	10
E10	lincomycin	10	62	12	G10	nalidixic acid	44	11	13
E11	guanidine HCl	91	90	14	G11	LiCl	10	11	14
E12	niaproof 4	10	10	16	G12	K-tellurite	97	93	20
F01	pectin	99	78	95	H01	tween-40	96	98	89
F02	D-galacturonic acid	99	98	100	H02	g-amino-butyric acid	8	10	12
F03	L-galacturonic acid lactone	8	6	12	H03	a-hydroxy-butyric acid	11	12	12
F04	D-gluconic acid	19	23	12	H04	b-hydroxy-D,L-butyric acid	14	17	11
F05	D-glucuronic acidi	34	28	11	H05	a-keto-butyric acid	8	6	10
F06	glucuronamide	23	17	15	H06	acetoacetic acid	19	86	10
F07	mucic acid	8	17	10	H07	propionic acid	10	22	10
F08	quinic acid	12	17	10	H08	acetic acid	93	98	99
F09	D-saccharic acid	12	17	14	H09	formic acid	10	14	10
F10	vancomycin	10	8	11	H10	aztreonam	10	16	98
F11	tetrazolium violet	44	38	99	H11	Na-butyrate	16	97	14
F12	tetrazolium blue	99	99	97	H12	Na bromate	12	14	15

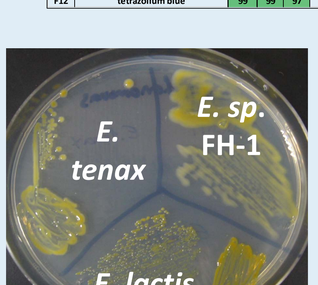

Figure 3. Mucoid Colony Morphology of *Epilithonimonas* sp. FH1 on R2A

Epilithonimonas specific genes:

gene #	protein	gene #	protein	gene #	protein
11	Lipolytic enzyme, G-D-S-L precursor	1014	lacX protein	2282	Protein vanZ
16	TonB-dependent receptor	1043	Subtilisin-like serine proteases	2283	Outer membrane lipoprotein omp16 precursor
83	Zinc metalloproteinase precursor (EC 3.4.24.29) / aureolysin	1053	FIG074102: hypothetical protein	2300	Bacterial leucyl aminopeptidase (EC 3.4.11.10)
85	Alpha-glucosidase (EC 3.2.1.20)	1071	Non-heme chloroperoxidase (EC 1.11.1.10)	2320	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.180)
91	LSU ribosomal protein L36p	1074	Oxidoreductase, aldo/keto reductase family	2425	Biopolymer transport protein ExbD/TolR
121	Leucine-rich-repeat protein	1094	Metallo-beta-lactamase precursor	2675	Putative inner membrane protein
151	putative TonB-dependent receptor	1165	Probable RND efflux membrane fusion protein	2708	Outer membrane assembly protein
179	N-formylglutamate deformylase (EC 3.5.1.68)	1179	NAD(FAD)-utilizing dehydrogenases	2872	OmpA
180	Formiminoglutamic iminohydrolase (EC 3.5.3.13)	1234	helix-turn-helix motif	2886	Formyltetrahydrofolate deformylase (EC 3.5.1.10)
260	D-lactate dehydrogenase (EC 1.1.1.28)	1246	Zinc metalloproteinase precursor (EC 3.4.24.29) / aureolysin	2888	putative two-component system sensor protein, no kinase domain
274	Rare lipoprotein A	1266	Zinc metalloproteinase precursor (EC 3.4.24.29) / aureolysin	2891	Histidinol-phosphate aminotransferase (EC 2.6.1.9)
338	YafQ toxin protein	1289	Altronate dehydratase (EC 4.2.1.7)	2903	GCN5-related N-acetyltransferase
366	Glycosyl transferase, family 2	1295	Pectin degradation protein KdgF	2916	putative esterase
445	Esterase/lipase/thioesterase family protein	1297	2-deoxy-D-gluconate 3-dehydrogenase (EC 1.1.1.125)	2972	COG1272: Predicted membrane protein hemolysin III homolog
493	TonB family protein / TonB-dependent receptor	1298	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (EC 5.3.1.17)	2982	response regulator
504	Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	1325	Protein YidD	2983	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)
540	Probable Co/Zn/Cd efflux system membrane fusion protein	1384	leucyl-tRNA synthetase(EC:6.1.1.4)	3005	L-sorbosone dehydrogenase
567	Possible heptosyltransferase	1438	Short-chain alcohol dehydrogenase family	3012	TPR repeat-containing protein
641	FIG074102: hypothetical protein	1443	ABC transporter, permease protein	3022	membrane protein
675	Sensory box histidine kinase	1451	A. fulgidus predicted coding region AF1548	3045	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
727	Glycosyl hydrolase, BNR repeat precursor	1494	tRNA-specific 2-thiouridylase MnmA	3114	Bleomycin resistance protein
738	putative stress-induced protein OsmC	1495	Chitinase (EC 3.2.1.14)	3142	PIN domain protein
743	Teicoplanin resistance protein vanZ	1509	Putative permease	3219	ATP-dependent DNA helicase RecQ
746	Ferric siderophore transport system, periplasmic binding protein TonB	1557	Pectate lyase precursor (EC 4.2.2.2)	3235	O-antigen acetylase
755	Transcriptional regulator, AraC family	1573	Transcriptional regulator, Cro/CI family	3263	putative membrane protein of unknown function
769	Protein-export membrane protein SecD (TC 3.A.5.1.1)	1577	FIG01019711: hypothetical protein	3304	Lipopolysaccharide core biosynthesis protein RfaS
802	Membrane associated protein	1580	Metal-dependent phosphohydrolase, HD subdomain	3374	Transcriptional regulator, HxlR family
823	Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids	1582	FIG00652395: hypothetical protein	3375	Transcriptional regulator, AraC family
845	Prolyl endopeptidase (EC 3.4.21.26)	1609	Putative acetyltransferase	3377	Internalin-like protein (LPXTG motif) Lmo0331 homolog
904	Integron integrase	1847	Endonuclease I	3455	Transcriptional regulator, AraC family
927	6-phosphogluconolactonase (EC 3.1.1.31)	1851	N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)	3473	Sensor histidine kinase of a two component response regulator
939	Outer membrane protein A precursor	1939	CTP:molybdopterin cytidylyltransferase	3518	FIG00574910: hypothetical protein
947	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	2071	Transcriptional regulator	3606	O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)
958	Internalin-like protein (LPXTG motif) Lmo0331 homolog	2090	FKBP-type peptidyl-prolyl cis-trans isomerase FkpA precursor (EC 5.2.1.8)		

Conclusions:

- When compared to the reference organisms, *Epilithonimonas diehli* contains enough genomic and phenotypic differences for it to be named a novel species

Acknowledgements

- Biolog and MIDI-FAME instruments were funded NSF Award DBI-0960114: MRI-R2: Acquisition of Instrumentation for Novel Microbe Characterization by Undergraduate Researchers to JDN.
- GCAT-SEEK is funded by NSF Award DBI-1248096:RCN-UBE GCAT-SEEK: The Genome Consortium for Active Undergraduate Research and Teaching Using Next-Generation Sequencing.

References:

- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H. & other authors. (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA Gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol 62, 716–721.
- Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-P., Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60, 2013.