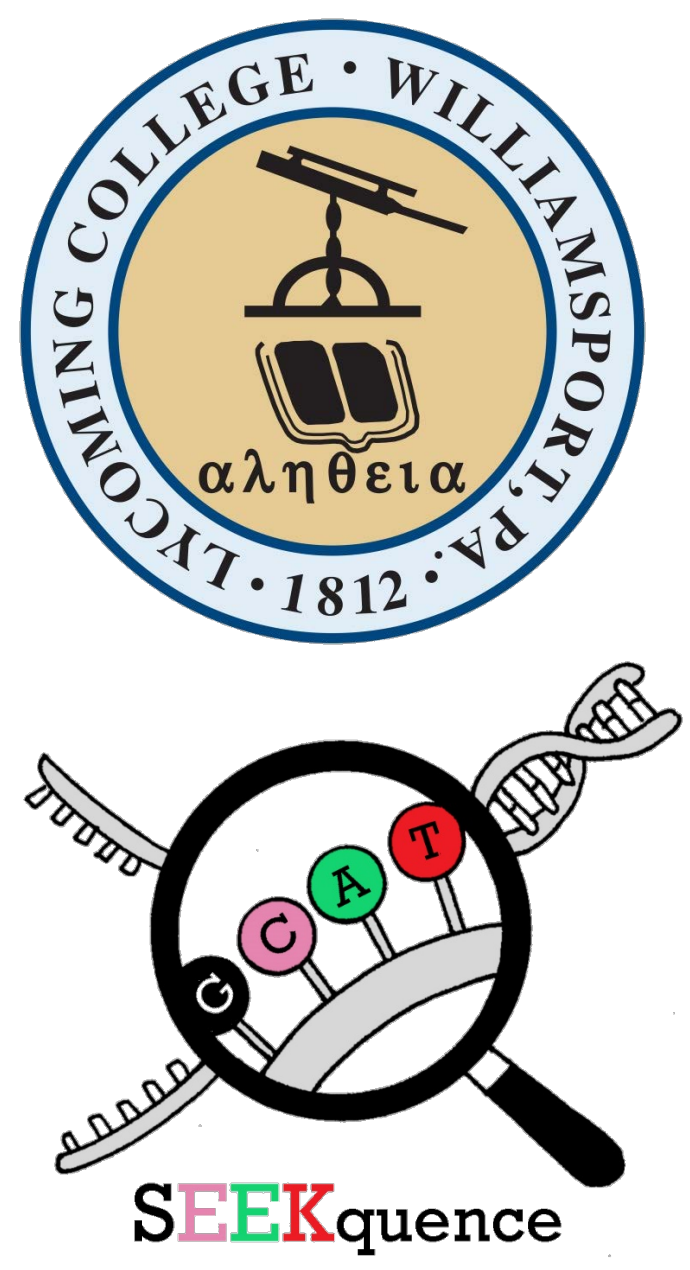
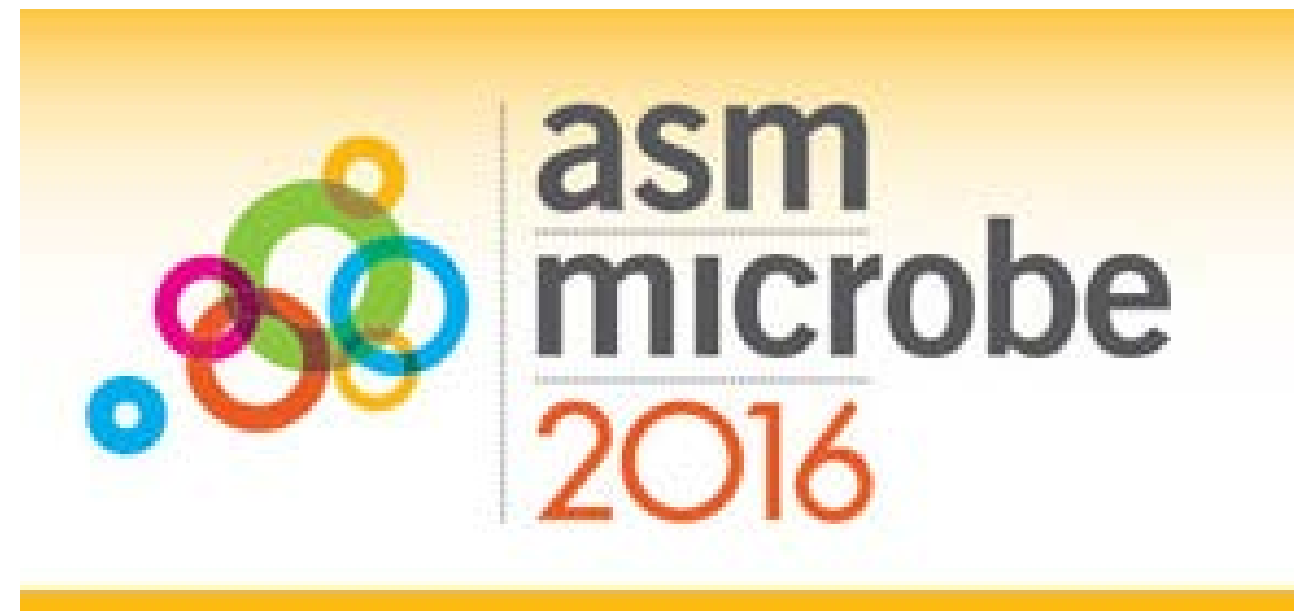


Genomic Analysis and Classification of Novel Species *Flavobacterium gabrieli* KJJ

Kirsten Fischer, Kathy Jacobs, Caitlin Dougherty & Jeffrey D. Newman

Biology Department, Lycoming College, Williamsport, PA 17701



Introduction

The classification and characterization of prokaryotes is formally regulated and has experienced rapid change over the last fifty years (Tindall *et al.*, 2010). The preliminary characterization of new species uses comparison of the 16S rRNA gene sequence. The widespread use of 16S rRNA sequences in bacterial identification is due to the fact that it is an essential gene, its conserved, and large enough for bioinformatics purposes at 1,500 base pairs (Janda & Abbott, 2007). 16S rRNA has continuously proven to be reliable for genus-level identification, but species level identification can be problematic at times due to undetectable species or strain differences as well as failure to recognize novel taxa. Other genes with greater resolution should be analyzed in addition to the 16S. Pairwise similarity alone is insufficient when describing a new species and more comprehensive methods, particularly genomic metrics, in conjunction with phenotypic comparison should be considered when characterizing novel taxa. (Tindall *et al.*, 2010).

Identity verification with EzTaxon

Rank Name Strain Authors Accession Pairwise Similarity (%) Diff/Total nt

- 1 *Flavobacterium chungangense* LMG 26729(T) Kim et al. 2009 JASY01000008 98.03 29/1472
- 2 *Flavobacterium chilense* LM-09-Fp(T) Kämpfer et al. 2012 FR774915 97.84 30/1388
- 3 *Flavobacterium saccharophilum* DSM 1811(T) (Reichenbach 1989) Bernardet et al. 1996 AM230491 97.73 33/1453
- 4 *Flavobacterium glaciei* 0499(T) Zhang et al. 2006 DQ515962 97.68 33/1422
- 5 *Flavobacterium aquidurensis* WB-1.1.56(T) Cousin et al. 2007 AM177392 97.67 34/1460
- 6 *Flavobacterium tractae* 435-08(T) Zamora et al. 2014 HE612100 97.67 34/1458
- 7 *Flavobacterium spartanense* T16(T) Loch and Faisal 2014 JX287799 97.54 34/1383
- 8 *Flavobacterium oncorhynchi* 631-08(T) Zamora et al. 2012 FN669776 97.53 35/1417
- 9 *Flavobacterium succinicans* LMG 10402(T) (Reichenbach 1989) Bernardet et al. 1996 JATV01000001 97.42 38/1472

Fig.1 - 16S rRNA Pairwise Identity of KJJ vs. type strains.

PCR product (27f-1492r) was sequenced (Sanger) and analyzed with EzTaxon (<http://www.ezbiocloud.net/>). If the closest match is below 98.5% similarity, it is potentially a novel species.

16S rRNA phylogeny with MEGA 6

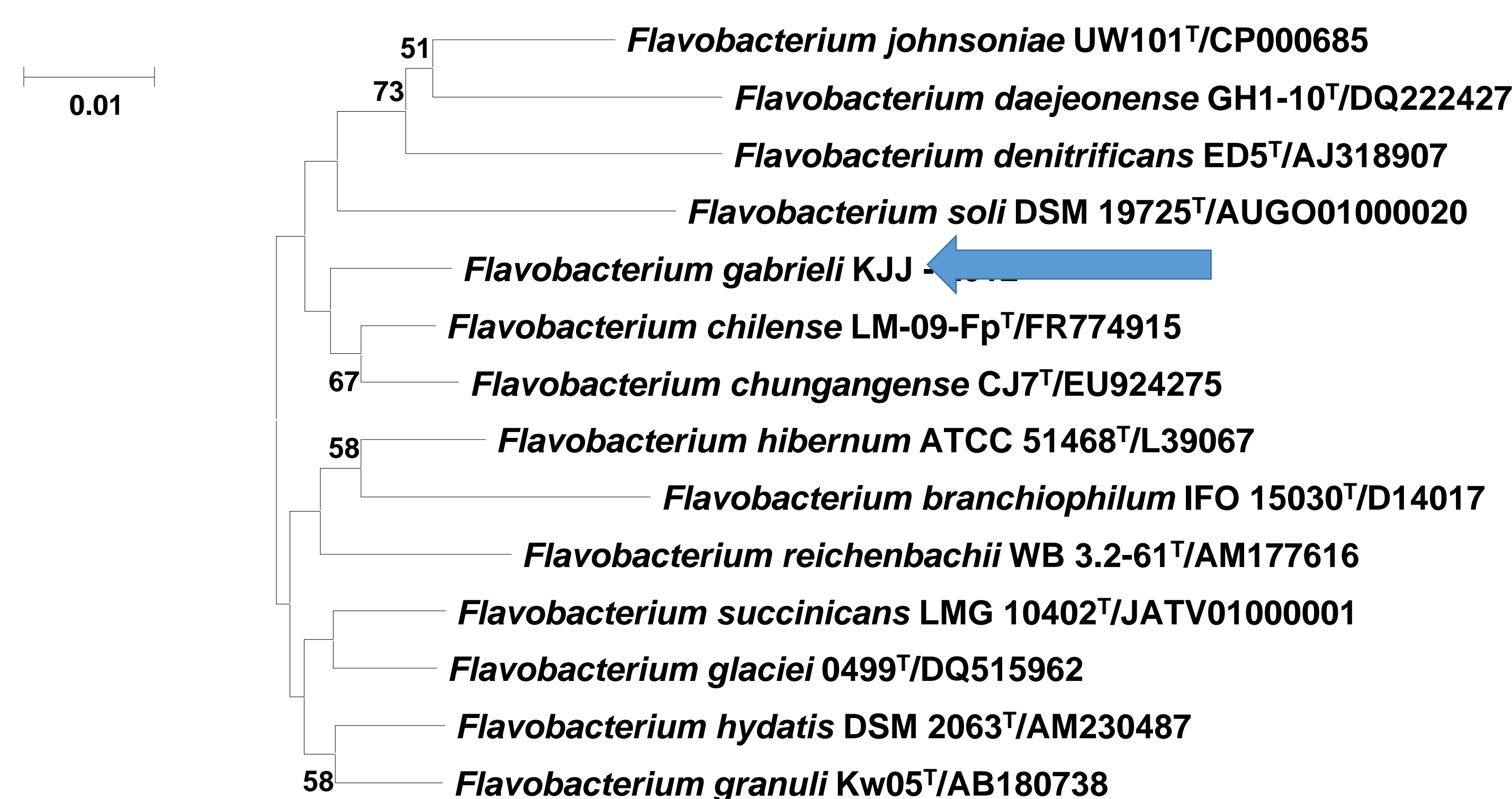


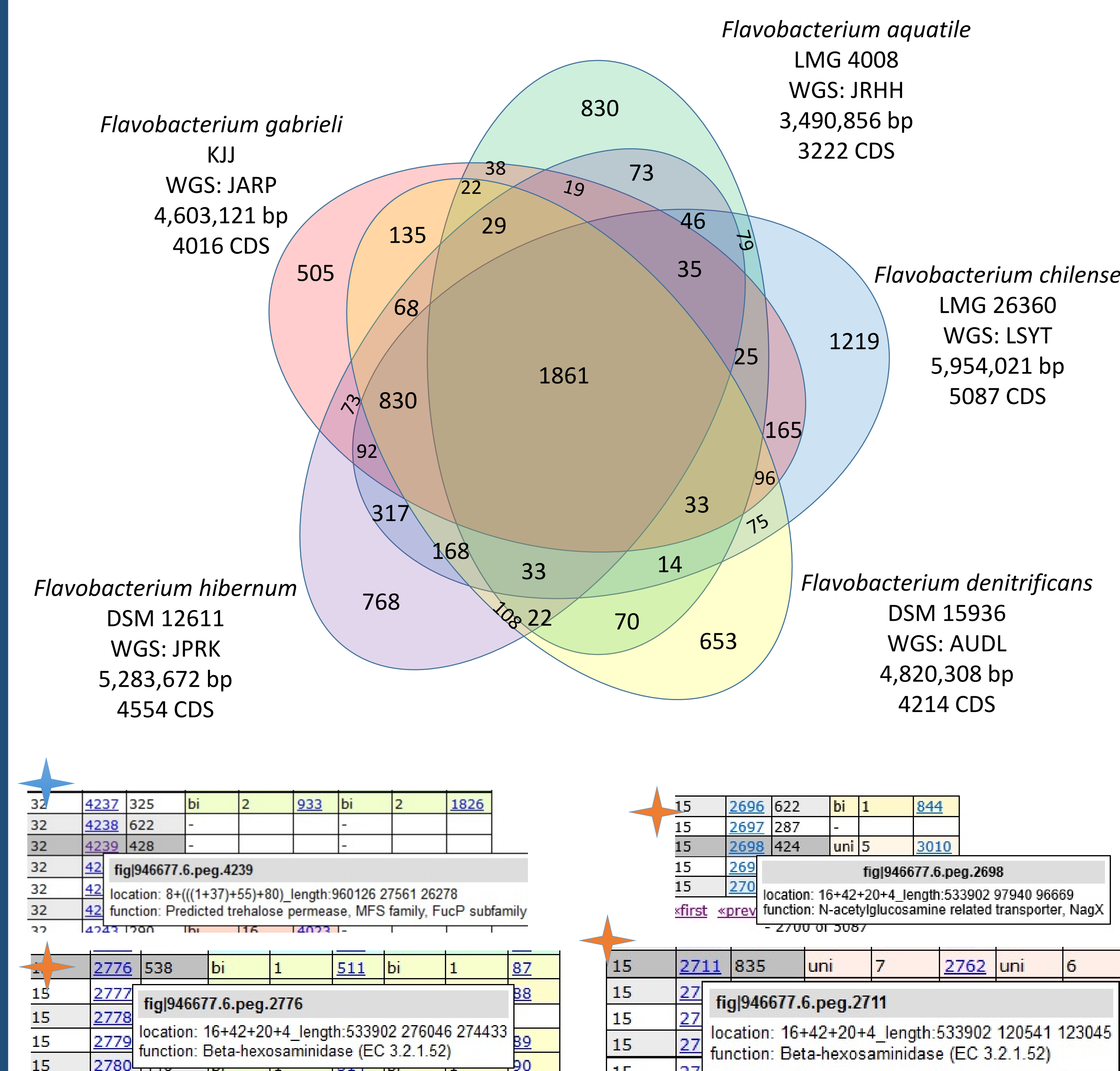
Fig.2 - 16S rRNA sequence neighbor-joining tree created with MEGA6. Strain KJJ clusters with *F. chilense* and *F. chungangense*.

Table 1. Phylogenomic matrices

Accession	Strain	16S ANI	Estimated DNA-DNA Hybridization (eDDH)																		
			1	2	3	4	5	6	7	8	9	10									
JARP	<i>F. gabrieli</i> KJJ	1	26.1	26.7	25.2	25.3	25.0	24.8	22.5	20.9	19.6										
LSYT	<i>F. chilense</i> LMG 26360 ^T	2	98.23	82.2	85.8	27.3	24.8	24.6	25.7	24.7	22.3	20.7	19.8								
JPRK	<i>F. hibernum</i> DSM 12611 ^T	3	97.14	82.6	85.2	85.9	24.6	25.3	24.9	24.7	23.3	21.1	20.1								
AUDL	<i>F. denitrificans</i> DSM 15936 ^T	4	96.55	82.4	85.8	84.1	83.6	24.3	25.9	24.1	22.0	20.5	20.0								
JPRL	<i>F. reichenbachii</i> LMG 25512 ^T	5	97.05	82.0	84.0	82.4	82.7	24.9	24.0	22.2	20.9	19.4									
CP000685	<i>F. johnsoniae</i> UW101 ^T	6	96.30	81.7	83.5	83.4	82.6	84.8	82.7	24.2	22.6	20.8	20.3								
JASY	<i>F. chungangense</i> LMG 26729 ^T	7	98.32	80.5	82.5	80.9	81.7	81.9	81.2	81.1	22.1	21.0	19.4								
JPRM	<i>F. hydatis</i> DSM 2063 ^T	8	97.56	77.6	78.3	78.0	79.7	78.0	76.3	77.1	77.0	21.4	20.1								
FQ859183	<i>F. glaciei</i> CGMCC 1.5380 ^T	9	97.73	75.8	75.3	74.8	74.8	74.7	74.4	74.1	74.3	76.0	19.9								
JRHH	<i>F. aquatile</i> LMG 4008 ^T	10	95.54	72.3	68.0	67.6	67.5	67.8	67.4	67.2	67.5	67.5	70.1								

All values are below the threshold for novel species, <70% for DDH and < 95% for ANI and AAI. *F. chilense*, *hibernum* and *denitrificans* were most similar in overall genome comparison. DDH calculated using the DSMZ Genome-Genome Distance Calculator (GGDC), ANI using EZ Taxon calculator, AAI with NewmanLab calculator.

Fig. 3 - Unique and shared genes via Rapid Annotation using Subsystems Technology (RAST)



Genomes annotated by RAST were explored for genetic differences to explain phenotypes (Table 2). Trehalose and hexosamine utilization were explored in particular. Blue star shows absence of trehalose permease in KJJ (right) found in *F. chilense* (left). Orange stars show absence of transporter, a bidirectional hexosaminidase, and a unidirectional hexosaminidase. Unidirectional hexosaminidase may explain the difference in specificity for various glucose derivatives between KJJ and *F. chilense*.

Connecting differences in phenotype to genotype using Biolog

well	carbon source or condition	Flavobacterium gabrieli KJJ	Flavobacterium chilense	Flavobacterium chungangense	Flavobacterium hibernum	well	carbon source or condition	Flavobacterium gabrieli KJJ	Flavobacterium chilense	Flavobacterium chungangense	Flavobacterium hibernum
A01	neg control	21	24	24	23	E01	gelatin	92	99	93	98
A02	dextrin	99	98	99	98	E02	glycyl-L-proline	53	95	95	91
A03	D-maltose	90	98	99	98	E03	L-alanine	44	83	7	33
A04	D-trehalose	12	98	10	97	E04	L-arginine	44	87	22	64
A05	D-cellobiose	57	98	98	99	E05	L-aspartic acid	44	97	91	93
A06	gentiobiose	99	98	99	97	E06	L-glutamic acid	70	97	97	95
A07	sucrose	14	11	97	98	E07	L-histidine	30	71	23	23
A08	D-turanose	14	12	20	9	E08	L-prolylglutamic acid	10	9	13	14
A09	stachyose	14	13	11	32	E09	L-serine	44	89	74	73
A10	pos control	98	97	98	98	E10	lincomycin	44	11	10	10
A11	pH 6	97	96	96	98	E11	guanidine HCl	40	18	9	39
A12	pH 5	27	67	12	16	E12	niaproof 4	16	14	14	12
B01	D-raffinose	17	14	14	96	F01	pectin	28	45	97	71
B02	α-D-lactose	16	21	19	22	F02	D-galacturonic acid	56	98	98	96
B03	D-melibiose	17	13	14	10	F03	L-galacturonic acid lactone	12	8	9	53
B04	β-methyl-D-glucoside	14	17	98	97	F04	D-gluconic acid	14	27	8	8
B05	D-salicin	10	98	98	98	F05	D-gluconic acid	18	47	97	51
B06	N-acetyl-D-glucosamine	99	97	6	96	F06	glucuronamide	17	21	33	18
B07	N-acetyl-β-D-mannosamine	77	16	10	12	F07	muic acid	40	10	15	7
B08	N-acetyl-D-galactosamine	48	97	36	95	F08	quinic acid	10	12	21	16
B09	N-acetylneuraminic acid	11	97	7	95	F09	D-saccharic acid	10	10	14	8
B10	1% NaCl	55	92	94	63	F10	vancomycin	63	94	95	94
B11	4% NaCl	14	13	13	12	F11	tetrazolium violet	90	99	45	70
B12	8% NaCl	19	16	16	11	F12	tetrazolium blue	99	99	95	96
C01	α-D-glucose	99	98	98	97	G01	p-hydroxy-phenylacetic acid	15	12	9	11
C02	D-mannose	99	97	98	97	G02	methyl pyruvate	14	71	12	85
C03	D-fructose	44	89	34	94	G03	D-lactic acid methyl ester	13	14	27	17
C04	D-galactose	99	97	51	97	G04	L-lactic acid	10	13	19	11
C05	3-methyl glucose	10	13	13	8	G05	citric acid	16	14	21	15
C06	D-fucose	12	16	6	8	G06	α-keto-glutaric acid	11	10	18	10
C07	L-fucose	16	20	68	87	G07	D-malic acid	13	11	16	12
C08	L-rhamnose	46	96	98	93	G08	L-malic acid	46	13	15	15
C09	inosine	11	8	7	8	G09	bromo-succinic acid	10	8	6	7
C10	1% Na-lactate	91	93	96	94	G10	nalidixic acid	16	13	12	13
C11	fusidic acid	11	10	9	9	G11	LiCl	11	11	9	11
C12	D-serine	16	15	10	12	G12	K-tellurite	21	18	16	18
D01	D-sorbitol	17	15	13	11	H01	tween-40	50	97	96	85
D02	D-mannitol	12	11	18	9	H02	γ-amino-butyric acid	16	12	12	16
D03	D-arabitol	10	12	11	7	H03	α-hydroxy-butyric acid	16	11	13	12
D04	myo-inositol	11	10	15	7	H04	β-hydroxy-D,L-butyric acid	13	11	19	13
D05	glycerol	44	8	8	8	H05	α-keto-butyric acid	12	8	7	7
D06	D-glucose-6-PO4	36	16	27	28	H06	acetoacetic acid	30	31	41	47
D07	D-fructose-6-PO4	27	21	90	34	H07	propionic acid	12	10	6	7
D08	D-aspartic acid	7	7	6	8	H08	acetic acid	51	91	89	60
D09	D-serine	7	7	6	6	H09	formic acid	16	11	8	9
D10	troleandomycin	12	10	9	9	H10	aztreonam	96	96	98	97
D11	rifamycin SV	96	94	96	96	H11	Na-butyrate	16	14	15	13
D12	minocycline	16	13	12	13	H12	Na bromate	15	14	16	12

Conclusions

- Flavobacterium gabrieli* KJJ shows sufficient genetic differences to be classified as a novel species.
- 16S rRNA analysis is not sufficient to identify organism's phylogeny. Phylogenomic matrices show *F. chilense* and *F. hibernum* to be more closely related than *F. chungangense*.
- F. chilense* and *F. hibernum* have more unique genes, possibly due to their larger genome sizes.

References

- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M. & other others (2008). The RAST server: rapid annotations using subsystems.
- 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.
- Newman Lab ROSA calculator: <http://lycofs01.lycoming.edu/~newman/ROSA/index.htm>
- Tindall, B.J., Rosselló-Móra, R., Busse, H.J., Ludwig, W., & Kämpfer. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Sys Evol Microbiol* 60, 249-266.