Use of Whole Genome Sequencing to Characterize a Cluster of Gram-Negative Bacilli of the Family *Xanthomonadaceae* Infecting Humans by Zoonotic Transmission

P2100

A-M. Bernier[1], K. Bernard [2, 3]

1. U. de Saint-Boniface, Winnipeg, Canada 2. National Microbiology Laboratory, PHAC, Winnipeg, Canada 3. U of Manitoba Dep't Medical Microbiology Winnipeg Canada

Introduction

Eight isolates collected over a 41 year period (1971- 2012) recovered from human infections associated with animal bites were studied. These were closest to each other (>99% identity) by 16S rRNA gene sequencing but distinct from other genera in the family *Xanthomonadaceae* by that approach. Strains were extensively characterized using conventional methods as well as by whole genome sequencing (WGS). They appeared to have caused at least localized infections in humans after being bitten by an animal and so may represent novel zoonotic agents. Isolates related to our strains by 16S sequencing have been reported in the feline and canine oral microbiome (1, 2).

Objectives

Perform a polyphasic taxonomic analysis of eight clinical sample isolates for the description of a novel genus with several new species of bacteria transmitted by zoonotic means which can cause human infections

Methods

- Eight clinical isolates closest to *Stenotrophomonas maltophilia* were studied using standard biochemical analyses (API-ZYM and API 20E) ; 16s rRNA sequencing and phylogenetic analyses were performed on all strains.
- Whole genome sequencing was performed in paired-end runs using the MiSeq Sequencer (Illumina1.9, 2x 300 cycles). Genomes were assembled using Spades (version 3.5.1). Annotation of the genomes was done with Prokka.
- Whole genome comparison was performed using JSpecies to calculate Average Nucleotide Identity values (ANIb). (3)
  - DNA-DNA hybridization % was calculated *in silico* (*is*DDH) using formula 2 of the Genome-to-genome distance calculator GGDC 2.1 (http://ggdc.dsmz.de/)
  - A Single Nucleotide Variant PHYLogenomics (SNVPhyl) pipeline was used to identify Single Nucleotide Variants (SNV) within the 8 strains and to construct a phylogenetic tree.
  - FFP – feature frequency phylogeny analysis for phylogenic tree based on core features (4)
- Proteome analysis was performed by MALDI-TOF MS on a Bruker Microflex, with analysis using Bruker's Biotyper 3.1 database (5999 spectra). Reference mass spectral profiles (MSPs) of each isolate were also constructed using Bruker software.

Acknowledgments

We thank the National Microbiology's DNA Core for 16S and whole genome sequencing, and Brittany Balcewich, Tina Zimmerman and Dr. Garrett Westmacott, Proteomic Core, for MALDI-TOF MS analyses

Results

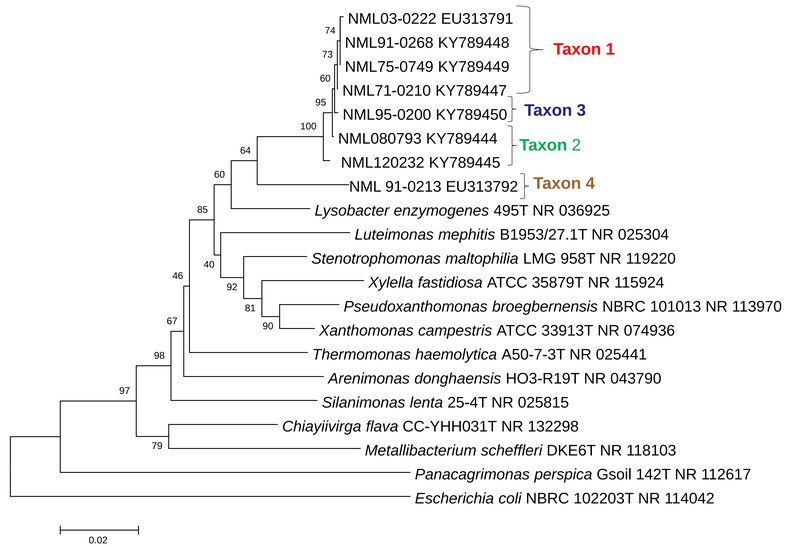

Figure 1. 16S rRNA gene sequence alignment using Clustal W (MEGA 6.06); relationship inferred using Neighbour-Joining algorithm (MEGA 6.06) for family *Lysobacteraceae* derived from *Xanthomonadaceae* (5). Scale indicates substitutions per nucleotide position; node values generated after 1000 replicates. Genbank accession numbers for 16S sequences

Strain	Number of contigs	# bp	G+C Content	Assembly coverage	Genes	rRNA	tRNA
NML08-0793	17	2,329,032	60,08%	72,9	2248	3	46
NML12-0232	36	2,358,721	60,00%	49,2	2288	3	38
NML03-0220	17	2,248,244	60,06%	85	2168	3	49
NML71-0210	1	2,219,593	60,17%	65,1	2092	3	47
NML75-0749	5	2,338,494	60,34%	55,6	2252	3	43
NML91-0268	19	2,164,729	60,31%	80,4	2087	3	46
NML95-0200	46	2,377,393	60,46%	61	2284	5	47
NML91-0213	15	3,236,205	69,30%	50	3020	3	50

Table 1 – WGS assembly and annotation results.

References

1. Dewhirst, F.E. et al. (2015). The feline oral microbiome: A provisional 16S rRNA gene based taxonomy with full length reference sequences. Vet. Micro. 175 : 294-303
2. Dewhisrt, F.E. et al. (2012). The canine oral microbiome. PLOS One 7(4):e36067
3. Richter, M. & Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. PNAS 106:19126-19131.
4. Sims, G.E. et al. (2009) Alignment-free genome comparison with feature frequency profiles (FFP) and optimal resolutions. PNAS 106(8):2677-2682
5. Naushad et al. (2015) A phylogenomic and molecular marker based taxonomic framework for the order *Xanthomonadales:* Ant Van Leeu et al.107:467-485

Results

Strain	NML08-0793	NML12-0232	NML03-0222	NML71-0210	NML75-0749	NML91-0268	NML95-0200	NML91-0213
NML08-0793	*	**83.00**	26.80	26.80	26.80	26.90	27.30	19.10
NML12-0232	**98.03**	*	27.10	27.10	27.00	27.20	27.40	18.90
NML03-0222	83.10	83.20	*	**78.10**	**78.70**	**77.80**	45.10	19.80
NML71-0210	82.94	83.04	**97.30**	*	**79.50**	**79.50**	45.30	19.50
NML75-0749	83.18	83.28	**97.43**	**97.61**	*	**98.30**	45.20	19.60
NML91-0268	83.03	83.16	**97.34**	**97.59**	**99.67**	*	45.10	19.60
NML95-0200	83.30	83.41	91.45	91.49	91.37	91.34	*	19.40
NML91-0213	71.16	71.16	71.50	71.69	71.65	71.63	71.69	*

Table 2. ANIb scores on bottom half of table and GGDC scores on top half of table. ANIb values > 94-96% represent the equivalent of 70% DDH values and are considered to be the same species (3). GGDC values represent DDH values, scores >70 are considered to be the same species.
GGCD ANIb values

- Strains were indistinguishable by biochemical analysis.
- **WGS** assembly and annotation results (Table 1).
- **16S sequence tree** (Figure 1) had the similar topology as the FFP and SNVphyl distance matrix trees (not shown) though whole genome trees show a clearer link between 120232 and 080793.
- **ANIb and *is*DDH values** >95% and 70% respectively identified 4 distinct taxonomic groups (Table 2, Figure 1) : 4 isolates (Taxon 1) closest by 16S to each other; a group of 2 isolates (Taxon 2) were found by WGS to fall into a distinct group/possible novel species (Table 2). Taxon 3 consisted of a single strain, Taxon 4 consisted of a single strain and possibly represents a new genus.
- **MALDI-TOF** analysis confirmed these WGS taxonomic results, with each taxon group being discernable from the other. Reference mass spectral profiles (MSPs) of each isolate were constructed using Bruker software. There were no matches in the Bruker Biotyper database. Locally made MSPs for each strain were compared to each and results confirm the 4 taxon groupings.

Conclusions

1) ANIb and isDDH values identified four new taxonomic groups distinct from know genus and species in the family *Xanthomonadaceae* (*Lysobacteraceae*) (5). These may represent genus and species nova.
2) These taxonomic groups were undistinguishable by biochemical analysis, therefore the use of molecular techniques is recommended.
3) The 4 new taxon groups had a range of G+C mol% of 60.00 to 69.3 and genome lengths from 2,164 to 3,24 Mb.
4) Isolate 91-0123 was very distinct from the other taxonomic groups studied here.