Journal of Microbiology (2015) Vol. 53, No. 8, pp. 503–510
DOI 10.1007/s12275-015-5142-0

Description of a novel pectin-degrading bacterial species *Prevotella pectinovora* sp. nov., based on its phenotypic and genomic traits[§]

Brigita Nograšek, Tomaž Accetto, Lijana Fanedl, and Gorazd Avguštin*

University of Ljubljana, Biotechnical Faculty, Animal Science Department, Division of Microbiology and Microbial Biotechnology, Groblje 3, 1230 Domžale, Slovenia

(Received Mar 16, 2015 / Revised Jun 17, 2015 / Accepted Jun 19, 2015)

Five strictly anaerobic Gram-negative bacterial strains, P4-65, P4-76^T, P5-60, P5-119, and P5-125, presumably belonging to the genus *Prevotella* were isolated from pig fecal samples. Strains were tested for various phenotypic traits and near-complete genome sequences were obtained and analyzed. Phylogenetic analysis based on 16S rRNA gene sequences and multilocus sequence analysis based on five conserved genes confirmed that the strains belong to the genus *Prevotella*, revealing that they represent a novel and discrete lineage distinct from other known species of this genus. The size of the genome of the isolated strains is 3–3.3 Mbp, and the DNA G+C content is 47.5–48.1 mol%. The isolates are strictly anaerobic, rod-shaped with rounded ends, non-motile and non-spore-forming. The main fermentation products are succinate and acetate, with minor concentrations of isovalerate, propionate and isobutyrate. Hydrogen is also produced. Major cellular fatty acids consist of anteiso-$C_{15:0}$ and iso-$C_{15:0}$, and a number of additional acids are present in lower concentrations. A substantial portion of genes involved in carbohydrate utilization is devoted to pectin degradation and utilization, while those supporting growth on xylan in ruminal *Prevotella* could not have been revealed. On the basis of the presented results, a novel species, *Prevotella pectinovora* sp. nov. is proposed. The type strain is P4-76^T (=DSM 29996^T =ZIM B1020^T).

Keywords: anaerobic bacterium, *Prevotella pectinovora* sp. nov., pectin degradation and utilization, genome analysis, pig feces, taxonomy

Introduction

Strictly anaerobic bacteria from the genus *Prevotella* represent dominant bacterial populations in terms of function and abundance in various niches of the gastrointestinal tracts of animals and humans. These bacteria are predominately inhabitants of the rumen of cattle and sheep (Van Gylswyk, 1990; Wood *et al*., 1998; Stevenson and Weimer, 2007) and human oral cavity (Alauzet *et al*., 2010). Since molecular biology-driven projects have increased, however, prevotellas were reported as major bacterial populations in the hindgut of certain human populations (De Filippo *et al*., 2010) and animal species too, and these studies also showed that a large proportion of ribosomal sequences from, for example, pig hindgut belonged to various uncultured members of the genus *Prevotella* (Leser *et al*., 2002; Kim *et al*., 2011; Lamendella *et al*., 2011; Pajarillo *et al*., 2014). Until now, a number of strains belonging to the genus *Prevotella* were isolated from pig feces; however, they remained unclassified (Robinson *et al*., 1981, 1984). Therefore, we set up isolation experiments that should yield novel species belonging to this genus. A combination of traditional and molecular methods has led to the successful isolation of a number of such organisms, and here we describe an evolutionarily related group of strains that represents a novel taxonomic lineage within the genus *Prevotella* and propose and describe a new species.

Materials and Methods

Isolation, bacterial strains, and culture conditions

Fecal samples were retrieved from Swedish Landrace and Large White crossbred pigs reared indoors at the Farme Ihan d.d. pig farm in Slovenia. Samples were transferred to the laboratory as soon as possible and immediately placed into the anaerobic glove chamber (Coy Laboratories) filled with CO_2/H_2. One gram of each fecal sample was suspended in dilution medium M2GSC without sugars and subsequently plated on to M2GSC agar media (Avguštin *et al*., 1997) with or without vancomycin (3.5 μg/ml) and incubated at 37°C. Grown colonies were collected with a 1-μl inoculation loop and the cells were transferred in 15 μl of sterile TE buffer (pH 7.5) and incubated at 97°C for 5 min. These cell extracts were then used as DNA templates in screening PCR specific for *Bacteroidales* using primers AllBac412r (Layton *et al*., 2006) and GenBac3 (Siefring *et al*., 2008), which was shortened by one base (G) at the 5′ end. PCR mixtures contained 1 μl of the cell extract containing DNA, 0.9 U of recombinant *Taq* DNA polymerase, 1× *Taq* buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix (all from Fermentas), 5 pmol of each primer and sterile water to a final volume of 25 μl (Sigma). The PCR mixture was first denatured at 94°C for 5 min and then subjected to 33 PCR cycles as follows: denaturation at 94°C for 30 sec, primer annealing at 61°C for 30 sec, and elongation at 72°C for 30 sec. The final elongation step

*For correspondence. E-mail: gorazd.avgustin@bf.uni-lj.si; Tel.: +386-1-320-3827; Fax: +386-1-724-10-05
[§]Supplemental material for this article may be found at http://www.springerlink.com/content/120956.

was 5 min at 72°C. The PCR products were analyzed by electrophoresis on a 1.2% (w/v) agarose gel. Positive colonies were purified by repeated streaking on M2GSC agar plates. Genomic DNA was retrieved from pure cultures as described above and amplification of near-full-length 16S rRNA genes was performed, using primers fD1 (Weisburg *et al.*, 1991) and 1492r (Lane, 1991) and the same reaction mixture composition as described above. The PCR mixture was first denatured at 94°C for 5 min and then subjected to 30 PCR cycles as follows: denaturation at 94°C for 30 sec, primer annealing at 52°C for 30 sec, and elongation at 72°C for 80 sec. The final elongation step was 7 min at 72°C. The PCR products were analyzed by electrophoresis on a 0.9% (w/v) agarose gel and subsequently purified with the High Pure PCR Product Purification Kit (Roche Diagnostics GmbH). The purified PCR products were sequenced at our request at the Microsynth GmbH using the 1492r sequencing primer. The isolated strains were designated PX-Y (X-number of the pig sample, Y-number of the strain isolated from the fecal sample). The strains were maintained on agar plates in an anaerobic chamber and anaerobic slants using the Hungate technique (Bryant, 1972). For long-term storage, the strains were grown in 0.75% agar medium M2GSC and frozen at -70 and -20°C.

Physiological and biochemical characterization

The growth of the strains was tested on M2GSC, BHI (brain heart infusion agar), RCA (reinforced clostridial agar) (both from Biolife Italiana) and PYG medium (B) (DSMZ medium 1139) solidified with 1.5% agar. The BHI, RCA, and PYG medium (B) were prepared with or without haemin and vitamin K_1 as used in modified PYG medium (DSMZ medium 104). The temperature and pH growth range of the isolated strains was tested using M2GSC and PYG (peptone yeast extract broth basal medium with glucose; Anaerobe Systems) liquid media at 25, 30, 37, 42, and 45°C, and within the pH range of 4.5–9.5 for M2GSC medium and 4.5–11 for PYG medium (increments of 0.5 pH units were used). Bile resistance was tested using solid and liquid media M2GSC and PYG, supplemented with up to 3% (w/v) Oxgall (Sigma).

Cell morphology and endospores were examined using an Olympus BX50 epifluorescence microscope equipped with phase-contrast optics (magnification ×1000). The microscopy images were recorded with a CCD camera Luca^{EM}R (Andor Technology) and analyzed with the Nis-Elements 4.0 program (Nikon). The presence of flagella was checked by using a Philips CM 100 transmission electron microscope.

Strains were tested for carbohydrate utilization capability using the API 50 CH and API RAPID ID 32 A kits (bioMérieux). All tests were repeated twice. Strains were grown anaerobically in M2GSC broth at 37°C until late logarithmic growth phase was reached. For API 50 CH, the cultures were recovered by centrifugation at 3,000 rpm for 10 min (T-23 centrifuge; Janetzki) and resuspended anaerobically in the same medium without sugars or resazurin but with 0.017% (w/v) bromocresol purple as a pH indicator as described previously (Avguštin *et al.*, 1997). The strips were incubated anaerobically at 37°C. For API RAPID ID 32 A, the cells were recovered as described above and then resuspended in sterile distilled H_2O until turbidity equivalent to 4 McFarland was reached. The strips were incubated aerobically for 4 h at 37°C as recommended by the manufacturer. Separate analyses were carried out on extracellular activities of xylanases, carboxymethyl cellulases and pectin hydrolases as described previously with minor modifications (Avguštin *et al.*, 1997). Growth in the presence of oxygen was tested by incubation of anaerobically inoculated M2GSC and PYG solid media in aerobic conditions at 37°C. Spore-forming capacity was tested as described by Logan and De Vos (2009). Analysis of fermentation products was carried out by gas chromatography. Short-chain fatty acids (SCFAs), succinate and lactate were analyzed following double ether extraction or methylation of the culture supernatants (Holdeman *et al.*, 1977) on a Shimadzu GC-14A chromatograph equipped with an FID detector and DB-WAXETR capillary column (Agilent Technologies). Fermentation gases were detected using a Shimadzu GC-14A chromatograph equipped with a TCD detector and Porapack Q column (Supelco). Analysis of the cellular fatty acid composition was carried out by the Identification Service of the DSMZ. Briefly, the fatty acid methyl esters were obtained using minor modifications of the method described by Miller (1982) and by Kuykendall *et al.* (1988) and analyzed by gas chromatography according to the standard protocol of the Sherlock Microbial Identification System (MIDI; Microbial ID). The profiles of cellular fatty acids were compared using the calculation method MOORE6 with library database version 6.00 (Microbial ID).

Genomic DNA isolation, whole genome sequencing and bioinformatics analysis

Genomic DNA was isolated as described by Ausubel *et al.* (1992) with minor modifications, from strains grown in M2GSC broth at 37°C. Whole genomes of five strains were sequenced at our request at Microsynth GmbH using Illumina MiSeq paired-end technology. Genomic libraries were prepared with the Nextera XT kit (Illumina) according to the manufacturer's instructions. The quality of raw reads was checked by the FASTQC tool, version 0.11.2 (Andrews *et al.*, 2011–2014) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). 5′ and 3′ sequence ends were trimmed using the FASTX toolkit (FASTX trimmer) (Hannon Lab, http://hannonlab.cshl.edu/fastx_toolkit/index.html), and overlapping paired-end reads were merged by using FLASH software, version 1.2.11 (Magoč and Salzberg, 2011). Assembly was done by SPAdes Assembler, version 3.1.1 (Bankevich *et al.*, 2012) and genome annotation of the P4-65, P4-76^T, P5-60, P5-119, and P5-125 genomes was performed using Prokka 1.7 (Seemann, 2014). The proteins active on carbohydrates as judged by the CAZy (Carbohydrate-Active enZYmes) database (Lombard *et al.*, 2014) were identified using dbCAN (DataBase for automated Carbohydrate-active enzyme Annotation) HMM release 3.0 models (Yin *et al.*, 2012) and the suggested cut off e-value was applied.

16S rRNA gene sequences were extracted from genomic sequences using RNAmmer version 1.2 (Lagesen *et al.*, 2007). Phylogenetic analysis of retrieved 16S rRNA gene sequences was performed with the ARB phylogenetic software package 5.5 (Westram *et al.*, 2011) using the All-Species Living Tree project database (release LTPs115, March 2014) (Yarza *et al.*, 2008). The final tree was reconstructed after multiple

analyses employing maximum parsimony, neighbor joining, and maximum likelihood algorithms. Confidence in the tree topology constructed using the maximum likelihood method was determined by the bootstrap analysis employing 1,000 resamplings of the analyzed sequences.

Multilocus sequence analysis (MLSA) was performed essentially as described by Sakamoto and Ohkuma (2011) using 16S rRNA, *rpoB*, *gyrB*, *recA*, and *dnaJ* gene sequences retrieved from sequenced genomes and from different *Prevotella* species' genomes deposited in the NCBI Genome database (NCBI Resource Coordinators, 28 January 2015) (list of the retrieved genomes is presented in Supplementary data Table S1). Sequences were analyzed with ARB phylogenetic software.

Comparative analysis of whole genomic sequences of isolated strains and different *Prevotella* species (Supplementary data Table S1) was performed using JSpecies software version 1.2.1 (Richter and Rosselló-Móra, 2009).

Nucleotide sequence accession numbers

These whole-genome shotgun projects have been deposited in DDBJ/EMBL/GenBank under the accession numbers JXQH00000000, JXQI00000000, JXQJ00000000, JXQK00000000, and JXQL00000000 for *P. pectinovora* P4-65, *P. pectinovora* P4-76^T, *P. pectinovora* P5-60, *P. pectinovora* P5-119, and *P. pectinovora* P5-125, respectively. The versions described in this paper are versions JXQH01000000, JXQI01000000, JXQJ01000000, JXQK01000000, and JXQL01000000.

Results and Discussion

Isolation and phenotypic characterization

Five strains (P4-65, P4-76^T, P5-60, P5-119, and P5-125) were isolated from pig fecal samples as part of a more comprehensive project concerned with the isolation of strictly anaerobic bacteria from the phylum *Bacteroidetes* from various gastrointestinal environments (in preparation). Grown strains were recovered after at least 6 days of incubation in an anaerobic atmosphere at 37°C on M2GSC agar plates with vancomycin. The strains were first screened by *Bacteroidales*-specific PCR and positive colonies were streaked repeatedly to obtain pure cultures. Subsequently, 16S rRNA genes were amplified from the pure cultures in almost full length and sequenced. Sequence analysis showed that the isolated strains belong to the genus *Prevotella*, and represent a novel and distinct lineage within this genus (data not shown).

Isolated strains were also subsequently grown on BHI, RCA, and PYG medium (B). Visible growth on M2GSC agar plates occurred after two days, and the colonies reached a maximum diameter of approximately 1 mm within 6 days of incubation. Growth on other solid media was substantially slower. The added haemin and vitamin K_1 did not contribute to faster growth. On the contrary, the growth was inhibited. When pectin was added to the agarized medium, growth substantially improved. In liquid medium M2GSC, growth was visible after 24 h of incubation, and it reached OD_{654} values up to 1.0. Other liquid media did not support growth without supplementation of haemin. The strains grew at 37 and 42°C, and at 45°C they grew in M2GSC medium only, whereas no growth was observed at 25 and 30°C. The strains were able to grow successfully in the presence of bile acids when grown in M2GSC broth, though the growth was barely noticeable when PYG medium was used. Strains were shown to be strictly anaerobic and non-spore-forming. The best growth occurred at pH 6.5–7.5. The colonies of isolated strains were uniform, round, raised, with entire margin, smooth and shiny, and slightly yellowish in color. Cells of the isolated strains were Gram-negative, rod-shaped with rounded ends (0.8–3.2-μm long and 0.5 ± 0.2-μm wide). Young cultures typically contained longer cells (up to 12 μm), whereas old cultures contained predominately shorter rods. Flagella were

Table 1. Metabolic products and cellular fatty acids of *Prevotella pectinovora* sp. nov. and related *Prevotella* species

Characteristics\species[a]	1	2[b]	3[b]
Metabolic end products[c]	S, A,	S, A,	S, A
	iv, p, ib	iv	
Cellular fatty acids[d]	anteiso-$C_{15:0}$	anteiso-$C_{15:0}$	anteiso-$C_{15:0}$
	iso-$C_{15:0}$	iso-$C_{15:0}$	$C_{18:1}$ *ω9c*
		$C_{18:1}$ *ω9c*	$C_{16:0}$
	anteiso-$C_{17:0}$	iso-$C_{17:0}$ 3-OH	$C_{16:0}$ 3-OH
	iso-$C_{17:0}$	$C_{16:0}$	iso-$C_{17:0}$ 3-OH
	iso-$C_{13:0}$	iso-$C_{14:0}$	$C_{15:0}$
	iso-$C_{14:0}$	iso-$C_{16:0}$	iso-$C_{15:0}$
	$C_{16:0}$	$C_{15:0}$ 3-OH	$C_{18:2}$ *ω6,9c*
	$C_{15:0}$	$C_{18:2}$ *ω6,9c*	iso-$C_{14:0}$
	iso-$C_{16:0}$	anteiso-$C_{17:0}$ 3-OH	anteiso-$C_{17:0}$
	Summed feature 11[e]	$C_{16:1}$ *ω7c*	$C_{16:0}$ dimethyl aldehyde
		Summed feature 10[f]	Summed feature 3[g]

[a] Strains: 1, *P. pectinovora* sp. nov. (n = 5); 2, *P. stercorea* DSM 18206^T; 3, *P. copri* DSM 18205^T;
[b] Data from Hayashi *et al.* (2007); [c] A, acetate; S, succinate; p, propionate; iv, isovalerate; ib, isobutyrate. Lower case letters indicate minor products; [d] Cellular fatty acids (CFAs) presented above the dotted line comprised more than 10% of the total fatty acids; CFAs that comprised less than 1.5% were omitted. For *P. copri* the average % of CFAs of five analyzed strains were taken into consideration; [e] Summed feature 11 contained $C_{18:2}$ dimethyl aldehyde and/or iso-$C_{17:0}$ 3-OH; [f] Summed feature 10 contained an unknown fatty acid of ECL 17.834 and/or $C_{18:1}$ *c*11/*t*9/*t*6 fatty acid methyl ester; [g] Summed feature 3 contained unknown fatty acid of ECL 13.570 and/or iso-$C_{15:0}$ ALDE.

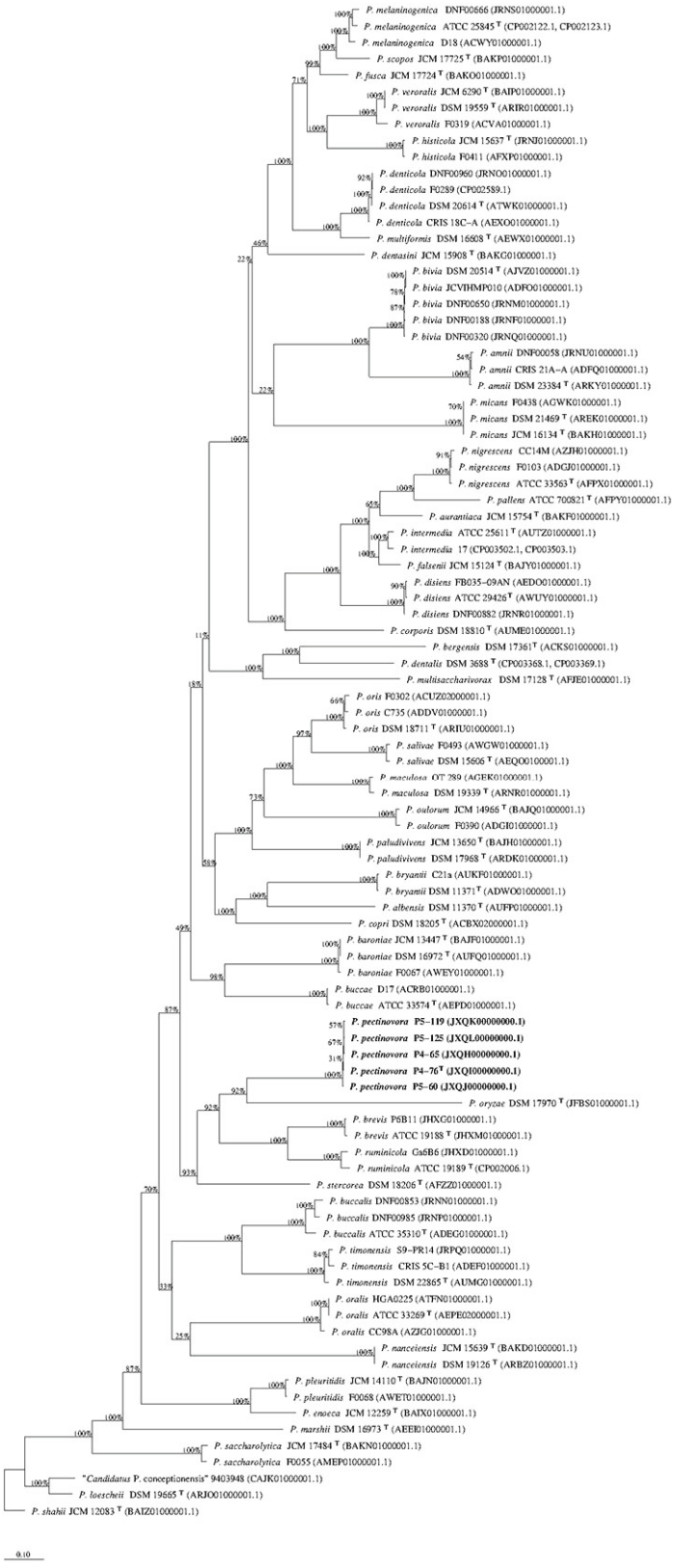

Fig. 1. Phylogenetic reconstruction based on the analysis of the concatenated sequences of 16S rRNA, *rpoB*, *recA*, *gyrB*, and *dnaJ* genes using maximum likelihood algorithm. The tree was constructed with the ARB phylogenetic software package. Bar indicates 10% of sequence divergence.

not observed.

Strains utilized L-arabinose, D-galactose, D-glucose, D-fructose, L-rhamnose, esculin, D-maltose, D-lactose, D-saccharose, inulin, D-raffinose, starch, and glycogen. Strain P5-119 also produced acid from D-xylose, which was only weakly utilized by P4-76^T, and showed a weak positive reaction for D-ribose. Other sugars included in API 50 CH test strips were not utilized by tested strains.

When the ability of the isolated strains to degrade polysaccharides was tested, a capacity for degradation of pectin was observed (data not shown). When grown on solid medium containing carboxymethyl cellulose and xylan, clearing zones after staining with Congo red indicating the degradation of the polysaccharide molecules were not observed.

The isolated strains produced as the main fermentation product succinate (2–5 g/L, on average 84%) and acetate (0.4 g/L, on average 11.5%), and as minor fermentation products (below 3%) isovalerate, propionate, and isobutyrate. Hydrogen was also produced (from 1 to 25% of the gas phase, on average 7%). Strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 were also tested with RAPID ID 32 A tests. Positive reactions were obtained for alkaline phosphatase, β-galactosidase, α-glucosidase, *N*-acetyl-β-glucosaminidase, α-fucosidase, leucyl glycine arylamidase, and alanine arylaminidase. Negative reactions were obtained for arginine dihydrolase, α-galactosidase, β-galactosidase-6-phosphate, β-galactosidase, α-arabinosidase, β-glucuronidase, for mannose and raffinose fermentation, glutamic acid decarboxylase, arginine, proline, phenylalanine, leucine, pyroglutamic acid, tyrosine, glycine, histidine, glutamyl glutamic acid and serine arylamidases. Strains did not reduce nitrate and no activity was detected for urease. The strains were positive for the production of indole.

The cellular fatty acids profile of *P. pectinovora* sp. nov. is given in Table 1. The major CFAs of strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 were anteiso-$C_{15:0}$ (42.7% on average) and iso-$C_{15:0}$ (20%), with lower amounts of anteiso-$C_{17:0}$ (6.4%), iso-$C_{17:0}$ (5.5%), iso-$C_{13:0}$ (2.7%), iso-$C_{14:0}$ (2.4%), and iso-$C_{16:0}$ (2.3%). The fatty acids profile of five strains differ from the fatty acids profiles of the most related species *Prevotella copri* and *Prevotella stercorea*, which contain fair amounts of $C_{18:1}$ *ω9c* fatty acid.

Phylogenetic analysis and genomic relatedness

The initial comparative analysis of the retrieved 16S rRNA gene sequences revealed that the isolated strains undoubtedly belonged to the genus *Prevotella* (data not shown). The isolated strains exerted a high degree of 16S rRNA gene sequence similarity (above 99%), however, the phylogenetic distance to the closest relative *Prevotella oryzae* is fairly large (10–12%, data not shown). The bootstrap values confirmed the stability of the node within the genus *Prevotella* (Supplementary data Fig. S1). Subsequently, genomic DNA of five selected strains was sequenced using Illumina

Table 2. Average Nucleotide Identity values (ANIb) of strains of *Prevotella pectinovora* sp. nov. and related *Prevotella* species

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	---	**98.85**	**98.81**	**98.83**	**98.83**	69.45	69.89	70.78	69.63	69.34	72.23	69.72	69.17	69.34	69.59	69.33
2	**98.93**	---	**98.81**	**98.94**	**98.72**	69.77	69.74	70.63	69.56	69.14	72.00	69.60	68.88	69.16	69.39	69.36
3	**98.81**	**98.73**	---	**98.82**	**98.85**	69.12	69.77	70.42	69.40	69.25	72.46	69.63	68.90	69.13	69.60	69.18
4	**98.90**	**98.90**	**98.78**	---	**98.75**	69.87	69.77	70.65	69.74	69.30	71.81	69.69	69.17	69.31	69.58	69.43
5	**98.87**	**98.67**	**98.83**	**98.73**	---	69.63	70.00	70.82	69.59	69.32	72.33	69.71	69.10	69.19	69.82	69.72
6	69.33	69.63	69.21	69.56	69.32	---	68.82	69.67	68.64	68.81	69.85	69.12	68.77	72.57	68.75	68.78
7	69.63	69.63	69.66	69.62	69.67	69.02	---	70.47	71.58	71.52	71.04	71.37	79.61	68.92	88.76	74.19
8	71.01	70.85	70.94	70.89	70.86	69.40	70.57	---	70.21	69.78	71.55	70.63	71.00	70.37	70.25	70.12
9	69.52	69.54	69.55	69.52	69.52	68.70	71.65	70.22	---	72.11	71.30	73.89	70.33	68.42	71.37	70.75
10	69.56	69.60	69.68	69.58	69.63	68.56	71.87	70.00	72.66	---	70.24	72.70	70.29	69.10	71.07	71.67
11	72.02	71.71	72.09	71.63	71.96	69.74	71.11	71.20	71.27	70.22	---	69.86	69.21	70.13	69.97	69.95
12	69.60	69.55	69.62	69.54	69.63	68.99	71.34	70.74	73.76	72.38	70.03	---	71.43	68.55	71.99	71.02
13	68.90	68.88	68.84	68.85	68.86	68.75	79.54	71.04	70.10	70.24	69.25	71.40	---	68.81	80.03	73.19
14	69.27	69.27	69.34	69.27	69.28	72.38	68.97	70.42	68.53	69.11	70.16	68.75	69.33	---	68.68	69.03
15	69.49	69.47	69.54	69.46	69.45	68.98	88.59	70.23	71.00	71.29	70.06	72.09	80.02	68.65	---	73.90
16	69.22	69.33	69.21	69.26	69.40	68.75	74.14	70.03	70.74	71.55	69.94	71.03	73.06	68.70	74.12	---

ANIb values were calculated using JSpecies software. The strains are as follows: 1, *P. pectinovora* P4-65; 2, *P. pectinovora* P4-76^T; 3, *P. pectinovora* P5-60; 4, *P. pectinovora* P5-119; 5, *P. pectinovora* P5-125, 6, *P. brevis* ATCC 19188^T; 7, *P. multiformis* DSM 16608^T; 8, *P. copri* DSM 18205^T; 9, *P. baroniae* JCM 13447^T; 10, *P. dentalis* DSM 3688^T; 11, *P. stercorea* DSM 18206^T; 12, P. *buccae* ATCC 33574^T; 13, *P. fusca* JCM 17724^T; 14, *P. ruminicola* ATCC 19189^T; 15, *P. denticola* DSM 20614^T; 16, *P. dentasini* JCM 15908^T.

MiSeq paired-end technology. Since phylogenetic analysis based on 16S rRNA gene sequences has limited resolution, the multilocus sequence analysis (MLSA) was performed with concatenated sequences of 16S rRNA, *rpoB*, *recA*, *gyrB*, and *dnaJ* genes, which were retrieved from assembled genomes and joined into a common database with sequences from 88 accessible genome projects containing genomic sequences from almost all of the known *Prevotella* species. MLSA confirmed the phylogenetic affiliation and the tight relationship of the isolated strains (Fig. 1). The sequence similarity of the isolated strains remained above 99%, whereas the distance to the *P. oryzae* and other related species increased to more than 15% (data not shown).

DNA:DNA relatedness of strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 was estimated with the JSpecies program using various methods. JSpecies ANIb (Average Nucleotide Identity based on BLAST) calculation confirmed 98.67–98.94% identity between five isolated strains. These numbers were essentially identical when ANIm (Average Nucleotide Identity based on MUMmer) was calculated (98.76–98.85% identity). When genomic sequences of strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 were compared to genomic sequences of other *Prevotella* species, a dramatic drop in identity values was observed (Table 2, Supplementary data Fig. S2). The highest ANIb values were observed when genomes were compared with *P. stercorea* (71.88%), though when ANIm was calculated, the highest identity was observed with *P. copri* (86.82%). Our results clearly show, that genomic se-

Table 3. Genomic characteristics of strains of *Prevotella pectinovora* sp. nov. and related strains of the genus *Prevotella*[a]

Strain\Characteristics	Level	Size (Mbp)	GC (%)	Gene	CDS	tRNA	CAZy[b]	Accession No.
P. pectinovora P4-65	contig	3.14	47.8	2456	2367	53	85	NZ_JXQH00000000.1
P. pectinovora P4-76^T	contig	3.26	47.5	2601	2508	53	86	NZ_JXQI00000000.1
P. pectinovora P5-60	contig	3.02	48.1	2345	2265	47	87	NZ_JXQJ00000000.1
P. pectinovora P5-119	contig	3.32	47.5	2657	2567	53	91	NZ_JXQK00000000.1
P. pectinovora P5-125	contig	3.12	47.8	2441	2372	51	93	NZ_JXQL00000000.1
P. stercorea DSM 18206^T	scaffold	3.09	49.0	2650	2571	48	84	NZ_AFZZ00000000.1
P. copri DSM 18205^T	scaffold	3.51	44.9	2958	2850	62	112	NZ_ACBX00000000.2
P. oryzae DSM 17970^T	scaffold	3.29	37.0	2632	2544	53	182	NZ_JFBS00000000.1
P. paludivivens DSM 17968^T	scaffold	3.45	37.3	2711	2647	47	181	NZ_ARDK00000000.1
P. bryantii C21a	scaffold	3.25	38.6	2608	2510	58	178	NZ_AUKF00000000.1
P. bryantii DSM 11371^T	contig	3.59	39.1	2960	2787	82	178	NZ_ADWO00000000.1
P. brevis ATCC 19188^T	scaffold	3.49	48.7	2736	2676	47	160	NZ_JHXM00000000.1
P. brevis P6B11	scaffold	3.21	49.2	2511	2453	46	147	NZ_JHXG00000000.1
P. ruminicola ATCC 19189^T	complete	3.62	47.7	3023	2939	53	191	NC_014033.1
P. ruminicola Ga6B6	contig	3.47	48.1	2862	2801	45	213	NZ_JHXD00000000.1
P. albensis DSM 11370^T	scaffold	2.67	41.2	2250	2163	49	107	NZ_AUFP00000000.1

[a] All data with the exception of CAZy were retrieved from the NCBI Reference Sequence database (Tatusova *et al.*, 2014);
[b] Only the number of genes coding enzymes from CAZy families of glycoside hydrolases, carbohydrate esterases and polysaccharide lyases are shown.

quences of five isolated strains are very similar (although not identical) and clearly different to genomic sequences of other prevotellas, and confirms the presumption, that a novel distinct species of the genus *Prevotella* was isolated.

Genome analysis

The estimated genome size of strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 is 3.022–3.319 Mbp, with G+C mol% ratios between 47.5 and 48.1 (Table 3). In this report, only the analysis of the glycoside hydrolases, carbohydrate esterases and polysaccharide lyases is shown, whereas other genome features will be presented elsewhere (in preparation). The reason for the focus on carbohydrate-active enzymes was the observation of this activity during growth on agarized medium containing pectin as the only energy source and discovery of clearing zones after staining with Congo red, indicating strong hydrolytic activity. The bioinformatics analysis showed that a plenitude of enzymes involved in pectin degradation is present, which is a clear genomic feature of this species. In Fig. 2, the number of genes coding for enzymes from different CAZy families in genomes of strains P4-65, P4-76^T, P5-60, P5-119, and P5-125 is presented (only families containing more than one CAZy-encoding gene are shown). It is apparent (i) that all five strains share very similar CAZy profiles, (ii) that a large proportion of genes belong to CAZy families known to be involved in pectin degradation, and (iii) that genes from CAZy families involved in xylan degradation, which are present in large numbers in other *Prevotella* species (Dodd *et al.*, 2011), are missing. We suggest that this profile of CAZy-encoding genes is unique among hitherto analyzed *Prevotella* species and could serve as a further diagnostic feature.

Description of *Prevotella pectinovora* sp. nov.

Prevotella pectinovora (pec.ti.no.vo'ra. N.L. n. *pectinum* pectin, methylated polygalacturonic acids in plant cell walls; L. v. *vorare* to devour; M.L. fem. adj. *pectinovora* pectin-destroying).

Cells are Gram-negative, rod-shaped with rounded ends (0.8–3.2-μm long and 0.5 ± 0.2-μm wide), non-motile, non-spore-forming. Strictly anaerobic. Colonies are uniform, round, raised, with entire margin, smooth and shiny, and slightly yellowish in color. Growth occurs in M2GSC and PYG medium at temperatures between 37 and 45°C, and is best in media at a pH between 6.5 and 7.5. Haemin and vitamin K_1 do not contribute to faster growth on solid media. When pectin is included, however, growth is improved. Visible colonies appear after 48 h of incubation under anaerobic conditions at 37°C, and reach a maximum diameter of approximately 1 mm after 6 days. In API 50 CH test strips, reactions are positive for D-glucose, D-lactose, D-galactose, D-fructose, L-arabinose, L-rhamnose, esculin, D-maltose, D-saccharose, inulin, D-raffinose, starch, and glycogen. Some strains also degrade D-xylose. In RAPID ID 32 A tests, reactions are positive for alkaline phosphatase, β-galactosidase, α-glucosidase, *N*-acetyl-β-glucosaminidase, α-fucosidase, leucyl glycine arylamidase, and alanine arylaminidase. The strains are positive for the production of indole. The main fermentation products are succinate and acetate, and minor fermentation products isovalerate, propionate, and isobutyrate are produced. Hydrogen is also produced. The major cellular fatty acids consist of anteiso-$C_{15:0}$ and iso-$C_{15:0}$ acids, and anteiso-$C_{17:0}$, iso-$C_{17:0}$,

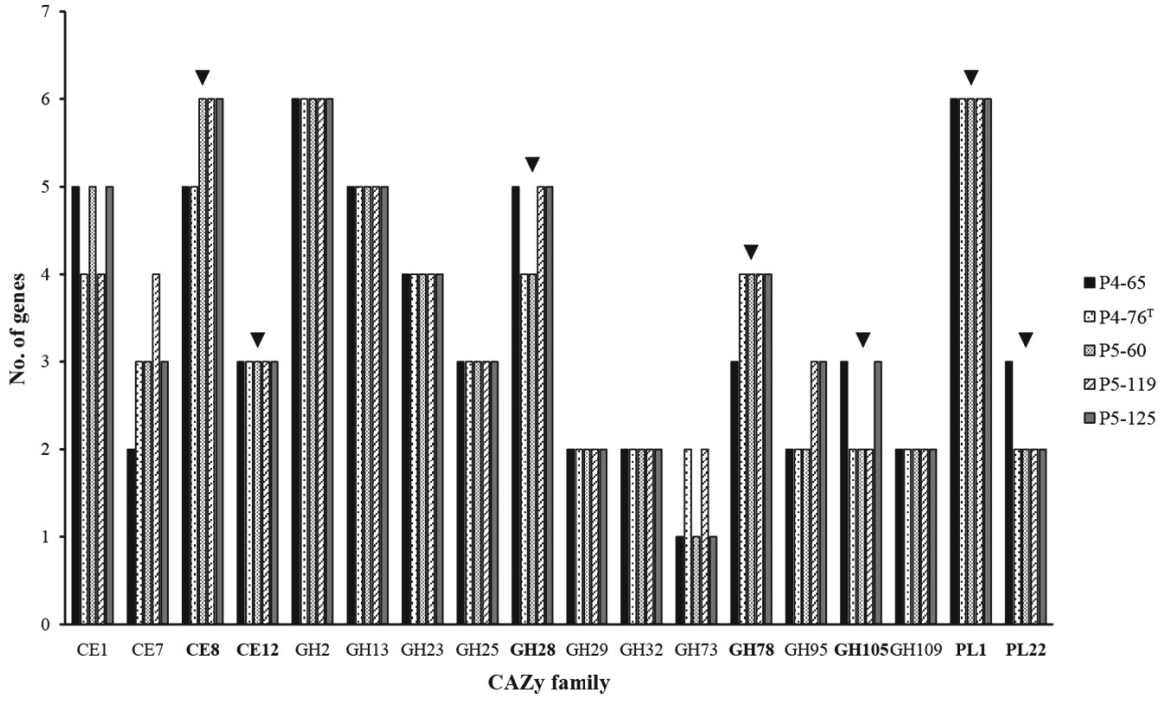

Fig. 2. Gene counts for CAZy families carbohydrate esterase (CE), glycoside hydrolase (GH), and polysaccharide lyase (PL) in genomes of strains of *Prevotella pectinovora* sp. nov. (P4-65, P4-76^T, P5-60, P5-119, and P5-125). Triangles above columns indicate CAZy families involved in pectin degradation.

iso-$C_{13:0}$, iso-$C_{14:0}$, $C_{16:0}$, $C_{15:0}$, and iso-$C_{16:0}$ acids are present in lower concentrations. The genome size is between 3 and 3.3 Mbp, and the mol% G+C is 47.5–48.1.

The type strain is *P. pectinovora* P4-76^T (=DSM 29996^T =ZIM B1020^T), isolated from the fecal sample of the Swedish Landrace and Large White crossbred pig.

Acknowledgements

We would like to express our gratitude to Marta Majdič for technical support in gas chromatography analysis and to Dr. Rok Kostanjšek for help in electron microscopy.

References

Alauzet, C., Marchandin, H., and Lozniewski, A. 2010. New insights into *Prevotella* diversity and medical microbiology. *Future Microbiol.* **5**, 1695–1718.

Andrews, S., Lindenbaum, P., Howard, B., and Ewels, P. 2011–2014. FastQC high throughput sequence QC report v. 0.11.2. Available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidmen, J.G., Smith, J.A., and Struhl, K. 1992. Current protocols in molecular biology, U2.4.1–U2.4.2, Vol. 1. John Wiley & Sons, New York, USA.

Avguštin, G., Wallace, R.J., and Flint, H.J. 1997. Phenotypic diversity among ruminal isolates of *Prevotella ruminicola*: proposal of *Prevotella brevis* sp. nov., *Prevotella bryantii* sp. nov., and *Prevotella albensis* sp. nov. and redefinition of *Prevotella ruminicola*. *Int. J. Syst. Bacteriol.* **47**, 284–288.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., *et al.* 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477.

Bryant, M.P. 1972. Commentary on the Hungate technique for culture of anaerobic bacteria. *Am. J. Clin. Nutr.* **25**, 1324–1328.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. USA* **107**, 14691–14696.

Dodd, D., Mackie, R.I., and Cann, I.K.O. 2011. Xylan degradation, a metabolic property shared by rumen and human colonic *Bacteroidetes*. *Mol. Microbiol.* **79**, 292–304.

Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S., and Benno, Y. 2007. *Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **57**, 941–946.

Holdeman, L.V., Cato, E.P., and Moore, W.E.C. 1977. Anaerobe Laboratory Manual, 4th ed. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA.

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., and Isaacson, R.E. 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. *Vet. Microbiol.* **153**, 124–133.

Kuykendall, L.D., Roy, M.A., O'Neill, J.J., and Devine, T.E. 1988. Fatty acids, antibiotic resistance, and deoxyribonucleic acid homology groups of *Bradorhizobium japonicum*. *Int. J. Syst. Bacteriol.* **38**, 358–361.

Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.H., Rognes, T., and Ussery, D.W. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108.

Lamendella, R., Santo Domingo, J.W., Ghosh, S., Martinson, J., and Oerther, D.B. 2011. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiol.* **11**, 103–119.

Lane, D.J. 1991. 16S/23S rRNA sequencing, pp. 115–148. *In* Stackebrandt, E. and Goodfellow, M. (eds.), Nucleic Acids Techniques in Bacterial Systematics. John Wiley & Sons, Chichester, UK.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., and Sayler, G. 2006. Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Appl. Environ. Microbiol.* **72**, 4214–4224.

Leser, T.D., Amenuvor, J.Z., Jensen, T.K., Lindecrona, R.H., Boye, M., and Møller, K. 2002. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* **68**, 673–690.

Logan, N.A. and De Vos, P. 2009. Genus *Bacillus* Cohn 1872, pp. 21–128. *In* De Vos, P., Garrity, G.M., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.R., Schleifer, K.H., and Whitman, W.B. (eds.), Bergey's Manual of Systematic Bacteriology, 2nd ed., Vol. 3. Springer, New York, USA.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495.

Magoč, T. and Salzberg, S.L. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963.

Miller, L.T. 1982. A single derivatization method for bacterial fatty acid methyl esters including hydroxy acids. *J. Clin. Microbiol.* **16**, 584–586.

NCBI Resource Coordinators. 2015. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **43**, D6–D17.

Pajarillo, E.A.B., Chae, J.P., Balolong, M.P., Kim, H.B., Seo, K.S., and Kang, D.K. 2014. Pyrosequencing-based analysis of fecal microbial communities in three purebred pig lines. *J. Microbiol.* **52**, 646–651.

Richter, M. and Rosselló-Móra, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* **106**, 19126–19131.

Robinson, I.M., Allison, M.J., and Bucklin, J.A. 1981. Characterization of the cecal bacteria of normal pigs. *Appl. Environ. Microbiol.* **41**, 950–955.

Robinson, I.M., Whipp, S.C., Bucklin, J.A., and Allison, M.J. 1984. Characterization of predominant bacteria from the colons of normal and dysenteric pigs. *Appl. Environ. Microbiol.* **48**, 964–969.

Sakamoto, M. and Ohkuma, M. 2011. Identification and classification of the genus *Bacteroides* by multilocus sequence analysis. *Microbiology* **157**, 3388–3397.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069.

Siefring, S., Varma, M., Atikovic, E., Wymer, L., and Haugland, R.A. 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. *J. Water Health* **6**, 225–237.

Stevenson, D.M. and Weimer, P.J. 2007. Dominance of *Prevotella* and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. *Appl. Microbiol. Biotechnol.* **75**, 165–174.

Tatusova, T., Ciufo, S., Fedorov, B., O'Neill, K., and Tolstoy, I. 2014. RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res.* **42**, D553–D559.

Van Gylswyk, N.O. 1990. Enumeration and presumptive identification of some functional groups of bacteria in the rumen of dairy cows fed grass silage-based diets. *FEMS Microbiol. Ecol.*

73, 243–254.

Weisburg, W.G., Barms, S.M., Pelletier, D.A., and Lane, D.J. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **173**, 697–703.

Westram, R., Bader, K., Pruesse, E., Kumar, Y., Meier, H., Glöckner, F.O., and Ludwig, W. 2011. ARB: a software environment for sequence data, pp. 399–406. *In* de Bruijn, F.J. (ed.), Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.

Wood, J., Scott, K.P., Avguštin, G., Newbold, C.J., and Flint, H.J. 1998. Estimation of the relative abundance of different *Bacteroides* and *Prevotella* ribotypes in gut samples by restriction enzyme profiling of PCR-amplified 16S rRNA gene sequences. *Appl. Environ. Microbiol.* **64**, 3683–3689.

Yarza, P., Richter, M., Peplies, J., Euzeby, J., Amann, R., Schleifer, K.H., Ludwig, W., Glöckner, F.O., and Rosselló-Móra, R. 2008. The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* **31**, 241–250.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451.