JOURNAL OF BACTERIOLOGY, June 2005, p. 3739–3751 Vol. 187, No. 11
0021-9193/05/$08.00+0 doi:10.1128/JB.187.11.3739–3751.2005

Novel Molecular Features of the Fibrolytic Intestinal Bacterium *Fibrobacter intestinalis* Not Shared with *Fibrobacter succinogenes* as Determined by Suppressive Subtractive Hybridization†

Meng Qi,[1] Karen E. Nelson,[2] Sean C. Daugherty,[2] William C. Nelson,[2] Ioana R. Hance,[2] Mark Morrison,[3] and Cecil W. Forsberg[1]*

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada[1]; The Institute for Genomic Research, Rockville, Maryland[2]; and Department of Animal Science, Ohio State University, Columbus, Ohio[3]

Received 18 February 2005/Accepted 2 March 2005

Suppressive subtractive hybridization was conducted to identify unique genes coding for plant cell wall hydrolytic enzymes and other properties of the gastrointestinal bacterium *Fibrobacter intestinalis* DR7 not shared by *Fibrobacter succinogenes* S85. Subtractive clones from *F. intestinalis* were sequenced and assembled to form 712 nonredundant contigs with an average length of 525 bp. Of these, 55 sequences were unique to *F. intestinalis*. The remaining contigs contained 764 genes with BLASTX similarities to other proteins; of these, 80% had the highest similarities to proteins in *F. succinogenes*, including 30 that coded for carbohydrate active enzymes. The expression of 17 of these genes was verified by Northern dot blot analysis. Of genes not exhibiting BLASTX similarity to *F. succinogenes*, 30 encoded putative transposases, 6 encoded restriction modification genes, and 45% had highest similarities to proteins in other species of gastrointestinal bacteria, a finding suggestive of either horizontal gene transfer to *F. intestinalis* or gene loss from *F. succinogenes*. Analysis of contigs containing segments of two or more adjacent genes revealed that only 35% exhibited BLASTX similarity and were in the same orientation as those of *F. succinogenes*, indicating extensive chromosomal rearrangement. The expression of eight transposases, and three restriction-modification genes was confirmed by Northern dot blot analysis. These data clearly document the maintenance of carbohydrate active enzymes in *F. intestinalis* necessitated by the preponderance of polysaccharide substrates available in the ruminal environment. It also documents substantive changes in the genome from that of *F. succinogenes*, which may be related to the introduction of the array of transposase and restriction-modification genes.

The genus *Fibrobacter* belongs to the poorly defined phylum *Fibrobacteres* (34) and is currently positioned between the *Spirochetes* and *Bacteroidetes*. The high cellulolytic activity and ability to degrade refractory plant structural polysaccharides of these organisms has given them a prominent role in gastrointestinal energy metabolism. Two species, *Fibrobacter intestinalis* and *Fibrobacter succinogenes* have been characterized based on 16S rRNA sequences and phenotypic properties. Both *F. intestinalis* and *F. succinogenes* are present in the rumen and ceca of cattle, pigs, and rats (3, 19, 30, 42). Of these two species *F. succinogenes* has received the most attention because of its substantially higher fibrolytic activity (28, 42). Genetic studies have identified seven cellulases and four xylanases, and have also shown that cellulose-binding proteins are important for the cellulolytic activity of this strain (9). The complexity of the cellulase system was further corroborated by sequence analysis of the *F. succinogenes* genome (31). Genes coding for over one hundred carbohydrate active enzymes, including cellulases, xylanases, polysaccharide lyases, and esterases, have been identified in the genome of *F. succinogenes*.

The most-studied strain of the *F. intestinalis* species, DR7, was shown to have an endoglucanase and a xylanase with high similarity to the respective enzymes of *F. succinogenes* S85 (18). In addition, multiple cellulose-binding proteins were identified (29). Furthermore, in vivo ^{13}C nuclear magnetic resonance study of glucose and cellobiose metabolism by *F. intestinalis* and *F. succinogenes* has revealed marked homogeneity in their carbon metabolism, with glucose being used for glycogen storage and energy production, while part of the cellobiose was diverted from glycolysis for synthesis of cellodextrins and the accumulation of glucose-6-phosphate (23). Despite these genotypic and phenotypic similarities between the two species the 16S rRNA similarity is only 92% (30), suggesting that there are substantial, yet undefined genomic differences between the two strains.

In the present study we have applied the technique of suppressive subtractive hybridization (SSH) (2) to explore the genomic differences between *F. intestinalis* and *F. succinogenes* by identifying genes present in *F. intestinalis* that either have low or no similarity to those of *F. succinogenes*.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *F. succinogenes* S85 (ATCC 19169) and *F. intestinalis* DR7 (ATCC 43855) strains were grown in a chemically defined medium prepared as described by Scott and Dehority (39), with either 0.3% (wt/vol) glucose or 0.3% (wt/vol) ball-milled cellulose as a carbon and energy source and with carbon dioxide as the gas phase. Cultures (10 ml) were grown at 37°C with reciprocating shaking at 150 rpm. To prepare genomic DNA, the

* Corresponding author. Mailing address: Department of Molecular and Cellular Biology, University of Guelph, Guelph, N1G 2W1 Ontario, Canada. Phone: (519) 824-4120, ext. 53433. Fax: (519) 837-1802. E-mail: cforsber@uoguelph.ca.

† Supplemental material for this article may be found at http://jb.asm.org/.

bacteria were grown in medium containing glucose for 12 h. For total RNA extraction, *F. intestinalis* DR7 was grown from a 5% inoculum for either 12 h in medium containing glucose or 24 h in medium containing ball-milled cellulose. Cells isolated after these incubation periods were in the mid to late exponential phase of growth. Ball-milled cellulose was prepared by mixing 200 ml of a 6% (wt/vol) aqueous suspension of Sigmacell-20 cellulose with flint balls in an 800-ml Mill Jar (Norton, Chemical Process Products Division, Ohio) at 70 rpm for 48 h.

Genomic DNA extraction. Genomic DNA was isolated from *F. intestinalis* DR7 and *F. succinogenes* S85 by using the cetyltrimethylammonium bromide procedure described by Ausubel et al. (5).

Suppression subtractive hybridization. SSH was performed by using the PCR-Select bacterial genome subtraction kit (Clontech), with *F. intestinalis* DR7 as the tester and *F. succinogenes* S85 as the driver. Briefly, genomic DNA (2 μg) from the tester (*F. intestinalis* DR7) was digested with RsaI and separated into two portions, each of which was subjected to a ligation reaction to attach a different set of PCR adaptors. The two portions were then separately hybridized to excess RsaI-digested *F. succinogenes* S85 driver DNA at 63°C for 1.5 h to allow all sequences common between the two strains to hybridize, leaving enriched tester-specific single-stranded DNA. The two preparations were then mixed together and hybridized at 63°C for 20 h. Only sequences specific for the tester strain that had different adaptors on each strand were amplified in the subsequent PCRs as describe in the Clontech subtraction protocol.

The mixture of PCR products was purified by phenol-chloroform extraction and ethyl alcohol precipitation, ligated to the T-A cloning vector pGEM-T (Promega, Madison, WI), and transformed into *Escherichia coli* JM109. The *E. coli* strains were grown in LB medium supplemented with ampicillin at 100 μg/ml and screened for clones by using X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)–IPTG (isopropyl-β-D-thiogalactopyranoside) plates containing 1.6% agar, 0.2 mM IPTG, and 40 μg of X-Gal/ml. A total of 1056 white colonies with inserts were recovered; they were restreaked for purity, and the plasmid inserts were sequenced at The Institute for Genomic Research (TIGR; Rockville, MD) as described by Nesbø et al. (32).

Sequence analysis. Sequences were examined by using the TIGR BLAST program (http://tigrblast.tigr.org/ufmg/) and the National Center for Biotechnology Information (NCBI) BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/). The expectation value of 0.001 was chosen as the cutoff, i.e., those with expectation values of >0.001 were considered to be lacking similarity. Similar criteria were used in several recent studies (32, 41). Genes with protein sequences lacking similarity according to this criterion to proteins encoded by either the *F. succinogenes* genome or gene in GenBank were further analyzed by position-specific iterative BLAST (PSI-BLAST) and by searches against the conserved domain databases. The genome of *F. succinogenes* S85 which is available at http://www.tigr.org/ was used for BLASTN and TBLASTX searches. The database used for BLASTX and PSI-BLAST searches was a combined database that included both the GenBank nonredundant amino acid database (nraa) and the translated coding sequences of *F. succinogenes* (http://www.tigr.org/tdb/ufmg/) to enable similarity comparisons by expectation value without the influence of database size. The PSI-BLAST was iterated up to three times for each sequence. Reverse-position-specific BLAST (RPS-BLAST) was used to search the conserved domain databases, including the NCBI's Conserved Domain Database, the Protein families database of alignments and HMMs (Pfam), Clusters of Orthologous Groups (COG and KOG), and the Simple Modular Architecture Research Tool (SMART), which are available at ftp://ftp.ncbi.nih.gov/pub/mmdb/cdd. The BLAST results were parsed and analyzed by using Microsoft Excel 2000 and Visual Basic for Application, version 6.3. Transposase genes from other organisms were downloaded from the UniProt database (http://www.uniprot.org/). A dendrogram showing the relationships of these genes were constructed by using CLUSTAL W (version 1.83) and visualized by using TreeView (version 1.6.6; http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

RNA extraction, cDNA synthesis, and digoxigenin labeling. Total RNA samples were prepared from a 10-ml culture of *F. intestinalis* DR7 grown in medium containing either glucose or cellulose by using the RNeasy minikit (Qiagen, Valencia, CA). Total RNA (10 μg) prepared from each culture was treated with 2 U of RNase-free DNase I (Amp Grade; Invitrogen) for 15 min at 22°C. The DNase I was then inactivated by the addition of 1 μl of 25 mM EDTA and heating for 10 min at 65°C. rRNA was removed from RNA preparations by using the MICROBExpress bacterial mRNA enrichment kit (Ambion). The enriched mRNA was treated a second time with 0.5 U of RNase-free DNase I for 15 min to ensure absence of DNA contamination. Reverse transcription reactions were performed by using Superscript II reverse transcriptase (Invitrogen) according to the recommended protocol. After 50 min, 1 U of RNase H was added to both samples, followed by incubation at 37°C for a further 20 min. cDNA was then purified by using a QIAquick PCR purification kit (Qiagen) and denatured at 96°C for 10 min. Denatured cDNA samples were chilled on ice and labeled with a digoxigenin (DIG) High-Prime labeling mixture (Roche), followed by incubation at 37°C for 20 h, and then stored at −20°C until used. The amounts of the DIG-labeled probe were determined as described in the DIG High-Prime labeling mixture manual.

Dot blot analysis to assess expression of selected genes. To assess differential mRNA synthesis of selected genes by *F. intestinalis*, inserts in selected SSH clones (see Table 4) were amplified by PCR with nested primers 1 and 2R (Clontech protocol) and Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen) by using 1-μl quantities of 16-h *E. coli* clone cultures as the templates. The reactions were performed in 30-μl volumes for 20 cycles with an annealing temperature of 52°C. To further dilute the cellular rRNA in the PCR, 0.3 μl of PCR product from each clone was amplified a further 20 cycles under the same conditions. The concentration of PCR product was determined by using PicoGreen reagent (Molecular Probes) as described by Ahn et al. (1). A 120-ng quantity of amplified insert from each of the clones was mixed with 0.4 M NaOH and heated at 100°C and then spotted on duplicate Hybond N+ nylon membranes (Amersham) by using a Bio-Dot SF Microfiltration Apparatus (Bio-Rad, Hercules, CA). The membranes were then baked for 2 h at 80°C. The hybridization and detection procedure that followed was carried out by using 400 ng of the DIG-labeled cDNA prepared with mRNA extracted from cells grown in medium containing either glucose or cellulose as an energy source as described by Ross et al. (37). The membranes were scanned, and the band profiles were plotted by using Scion Image software (version 4.0.2; Scion, Frederick, MD). The intensity of each band was represented by the area under the corresponding peak. A standard curve was plotted by spotting a range of 0.01 to 0.3 pg of standard DIG-labeled DNA onto a membrane and, after color development, the intensities of the bands were detected by the same procedure. The intensity increased in proportion to the amount of DIG-labeled probe up to a value of 900 arbitrary units with 0.3 pg of probe, which encompassed the intensity of the test sample blots. To determine the extent to which the genes were either up- or downregulated in cells during growth on glucose or on cellulose two constitutive house-keeping genes, RSGA_692, a phosphoglycerate kinase, and RSGA_128, a glucose-6-phosphate isomerase, were included on each blot to correct for differences in expression. The band intensities of the two housekeeping genes for cells grown on cellulose were divided by the values for cells grown on glucose. The average of the two ratios was 1.075, showing that there was slightly higher expression of the housekeeping genes on cellulose. To correct for this higher level of expression, the band intensity for each of the genes of cellulose grown cells was divided by the 1.075. Each gene selected was tested at least twice.

Codon usage analysis. The codon usage was analyzed by using the General Codon Usage Analysis program as described by McInerney (26). Pearson correlation coefficients between all pairs of cellulase genes and transposase genes were calculated and transformed into distances and clustered by using the unweighted pair group method with the arithmetic mean algorithm as previously described (10). The dendrogram was generated by the TreeView (version 1.6.6) program. Since short DNA fragments were cloned, complete open reading frames (ORFs) were not often obtained. The partial ORF sequences were identified by BLASTX matches to known proteins in the combined nraa databases of *F. succinogenes* and GenBank. All ORFs of the *F. succinogenes* genome were analyzed for codon usage.

Nucleotide sequence accession numbers. The sequences determined in the present study have been submitted to the GenBank nucleotide sequence database under accession numbers CZ 443203 to CZ 443914. The nucleotide sequences with annotations can also be accessed on the FibRumBa website (http://www.tigr.org/tdb/rumenomics/).

RESULTS

A total of 1,056 clones were sequenced in both directions, the sequences were assembled, and redundant sequences were removed, resulting in 712 contigs with an average length of 525 bp (Fig. 1). As determined by a BLASTN search of the genome of *F. succinogenes*, of the 712 nonredundant DNA contigs, 437 contained genes different from those in *F. succinogenes* (with an expectation value of >0.001). Another 206 had expectation values of ≤0.001, but with the DNA similarity to *F. succinogenes* of <80%, and 69 had sequence similarities of >80% over a ≥100-nucleotide region to the genes in *F. succinogenes.* None of the SSH contigs had a DNA similarity to *F. succinogenes*

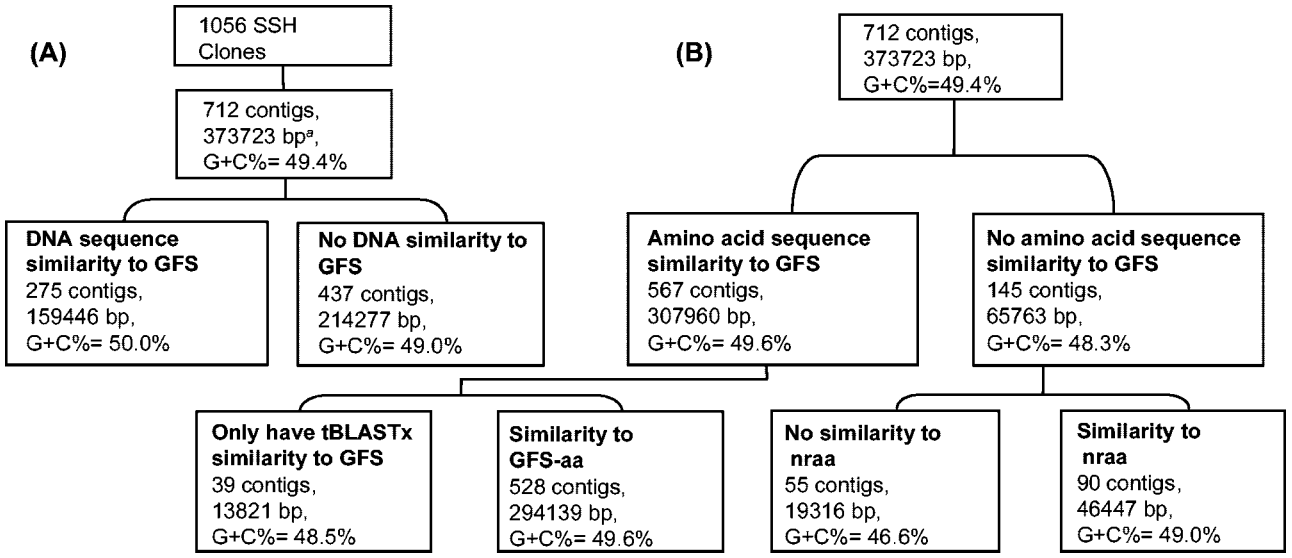

FIG. 1. Summary of the similarities of *F. intestinalis* contigs to genes present in the GenBank nonredundant amino acid database and the *F. succinogenes* S85 genome translated amino acid sequence database. The distribution of contigs is based on the gene in each contig with the lowest E-value. A possible second gene in a contig was not taken into account. The DNA inserts from clones were assembled into contigs, and redundant sequences were eliminated prior to BLAST analysis. (A) Similarity based on BLASTN search; (B) similarity based on BLASTX, tBLASTX, conserved domain database search, and PSI-BLAST search. Abbreviations: nraa, nonredundant amino acid database (which contains GenBank CDS translations, RefSeq Proteins, PDB, SwissProt, PIR and PRF); GFS, Genome of *F. succinogenes* database; GFS-aa, GFS amino acid sequence database; GFS-nraa, combined database contains both nraa and GFS-aa. Superscript *a*, total nucleic acid base pairs.

genome of >90% along the whole length of the sequence. The average G+C% of the unique clones was 49.4%, which is similar to the 48.1% reported for *F. succinogenes* (31). Using the estimate of redundancy as described by Nesbø et al. (32), the theoretical library size would be ca. 3,400 clones, with ca. 2,400 unique clones each containing 525 bp, which is equivalent to 1,260 kb of DNA. Therefore, assuming the genome size of *F. intestinalis* is similar to that of *F. succinogenes*, i.e., 3843 kb, the difference in genome content between the two strains is ca. 33%. A total of 145 contigs had no similarity to the genome of *F. succinogenes*, while among these contigs, BLASTX searches against the GenBank nraa showed similarity to 79 of them, and PSI-BLAST and conserved domain database searches found 11 more contigs that may share conserved domains with other proteins; the remaining 55 contigs were unique with no similarity to proteins in the nraa databases as of September 2004 (Fig. 1B). These 145 contigs demonstrate that at least 66 kb of DNA in the *F. intestinalis* chromosome is totally different from that in *F. succinogenes*.

The results of BLASTX analysis of the contigs, and the assignment of coding sequences to general functional roles, are shown in Table 1. The majority of contigs coded for proteins with partial rather than full-length similarities to proteins in the respective databases, a reflection of the digestion pattern resulting from the enzymes used in the SSH experiments. However, 137 (19%) of the contigs contained at least two genes and the intergenic region, i.e., each end of the contig encoding a different protein (see below). In all, 764 genes encoding different proteins were identified by BLAST searches and, among these, 109 genes coded for proteins that were unrelated to those of *F. succinogenes*, of which 45% had the highest similarities to the proteins of other gastrointestinal bacteria (Table 2). The remaining 655 genes coded for proteins that showed

TABLE 1. Roles of the unique *F. intestinalis* genes identified by SSH that exhibit similarity to protein sequences in GenBank and the genome of *F. succinogenes*[a]

Role(s)	No. of genes	
	Similarity to *F. succinogenes*[b] (n = 655)[c]	Absent from *F. succinogenes* (n = 109)
Amino acid biosynthesis	25	4
Purines, pyrimidines, nucleosides, and nucleotides	16	0
Fatty acid and phospholipid metabolism	8	0
Biosynthesis of cofactors and prosthetic groups	29	2
Central intermediary metabolism	3	0
Energy metabolism	63	6
Transport and binding proteins	33	13
DNA metabolism	28	9
Transcription	15	0
Protein synthesis	37	0
Protein fate	28	3
Regulatory functions	29	2
Cell envelope	77	11
Cellular processes	22	0
Mobile and extrachromosomal element functions	1	31
Other	0	7
Unknown function	270	21

[a] This table lists the functional role of each gene identified, showing the comprehensive range of genes that were identified by SSH. See Table 2 and Table S1 in the supplemental material for more detail.

[b] Genes with an E-value (which obtained either by BLASTX or PSI-BLAST searches after the third iteration) of ≤0.001 were regarded as similar to *F. succinogenes*, while genes with an E-value of >0.001 were regarded as absent from *F. succinogenes*.

[c] n, Total number of genes identified. The total number may exceed the sum of the genes included in each function role because some proteins may have more than one role.

TABLE 2. *F. intestinalis* SSH contigs that contain genes encoding putative proteins with no similarity to proteins in *F. succinogenes*[a]

Function and sequence[b]	Size (bp)[c]	Protein	Genome	Expect value	%Id[d]	Accession no.
Amino acid biosynthesis						
RSGA_238	155	CysQ protein	*Caulobacter crescentus*	2.0E-06	49	NP_422263.1
RSGA_399	717	Biotin carboxylase	*Dechloromonas aromatica*	3.0E-63	48	ZP_00152837.1
RSGA_682.a	1,254	Threonine dehydratase	*Fusobacterium nucleatum*	5.00E-47	58	NP_604305.1
RSGA_493.a	552	Aspartate aminotransferase	*Pirellula* sp.	1.00E-15	57	NP_867479.1
Biosynthesis of cofactors, prosthetic groups, and carriers						
RSGA_192	355	Ferrochelatase	*Xanthomonas axonopodis*	2.0E-07	44	Q8PEX0
RSGA_605	394	Ferrochelatase	*Xylella fastidiosa*	4.0E-08	28	NP_297856.1
Energy metabolism						
RSGA_528.a	405	Arsenate reductase	*Chlorobium tepidum*	8.0E-17	58	NP_662606.1
RSGA_058.a	689	Arsenate reductase	*Clostridium perfringens*	6.0E-24	76	NP_561754.1
RSGA_506.a	582	2,3-Bisphosphoglycerate-dependent phosphoglycerate mutase	*Ralstonia solanacearum*	2.0E-27	67	Q8Y2I3
RSGA_320	211	ATPase (AAA + superfamily) COG1373	*Desulfitobacterium hafniense*	2.0E-06	44	ZP_00098861.1
RSGA_419	767	Xylanase	*Streptomyces avermitilis*	8.0E-04	40	NP_827557.1
RSGA_456	270	COG3258 cytochrome *c*	Not applicable	6.0E-05[e]	54	CDD|12594
Transport and binding proteins						
RSGA_331	837	COG1629: outer membrane receptor proteins, mostly Fe transport	*Cytophaga hutchinsonii*	9.0E-09	25	ZP_00311071.1
RSGA_051	572	COG0370: Fe^{2+} transport system protein B	*Enterococcus faecium*	9.0E-15	38	ZP_00286591.1
RSGA_147	331	Phosphate-binding protein	*Bacillus halodurans*	4.0E-04	25	NP_241334.1
RSGA_531	328	COG0600: ABC-type nitrate/sulfonate/bicarbonate transport system, permease component	*Burkholderia cepacia*	2.0E-05	51	ZP_00224793.1
RSGA_207	393	COG1682: ABC-type polysaccharide/polyol phosphate export systems, permease component	*Magnetococcus* sp.	7.0E-32	50	ZP_00290976.1
RSGA_045.a	505	ABC transporter ATP-binding protein	*Nostoc* sp.	2.00E-04	32	NP_488818.1
RSGA_138	633	Na^+/H^+ antiporter	*Bacillus halodurans*	6.0E-14	30	NP_244906.1
RSGA_386	970	Na^+/H^+ antiporter NHAC	*Fusobacterium nucleatum*	2.0E-55	38	NP_603875.1
RSGA_338.a	1,073	Chromate transport protein	*Bacteroides thetaiotaomicron*	2.00E-53	60	NP_810648.1
RSGA_033.a	969	COG0701: permcases	*Clostridium thermocellum*	1.00E-24	50	ZP_00313782.1
RSGA_376	597	Transporter	*Bacteroides fragilis*	1.0E-21	32	AAK68924.1
RSGA_233	345	COG0609 ABC-type Fe^{3+} siderophore transporter system, permease component	Not applicable	0.001[e]	31	CDD|10479
RSGA_249	317	SMART00062 Bacterial periplasmic substrate-binding protein	Not applicable	0.001[e]	30	CDD|24199
DNA metabolism						
RSGA_017	358	Anaerobic ribonucleoside-triphosphate reductase	*Listeria monocytogenes*	3.0E-11	38	ZP_00233963.1
RSGA_686.a	1160	COG3077: DNA-damage-inducible protein J	*Desulfitobacterium hafniense*	2.00E-10	43	ZP_00100180.1
RSGA_009	219	Type I modification enzyme	*Haemophilus influenzae*	1.0E-10	43	NP_439439.2
RSGA_225	473	Type I restriction enzyme, modification subunit	*Helicobacter pylori*	1.0E-18	34	NP_223134.1
RSGA_194	302	HsdR (endonuclease subunit of type I restriction enzyme)	*Streptococcus thermophilus*	4.0E-09	38	AAG22017.1
RSGA_136	933	Type II restriction-modification system, restriction subunit	*Helicobacter pylori*	2.0E-32	33	YP_060476.1
RSGA_024	487	Type IIS restriction/modification enzyme	*Bacteroides thetaiotaomicron*	5.0E-20	46	NP_810844.1
RSGA_157	386	Methyltransferase, cytosine specific	*Methanothermobacter thermautotrophicus*	8.0E-33	57	NP_275638.1
RSGA_358	849	COG0322: nuclease subunit of the exonuclease complex	Not applicable	6.0E-04[e]	25	CDD|10196
Protein fate						
RSGA_573	514	Chaperonin, 10 kDa	*Chlorobium tepidum*	4.0E-04	50	NP_661469.1
RSGA_454	700	PrtD (metalloprotease)	*Photorhabdus* sp.	2.0E-09	24	AAS19413.1
RSGA_498.a	467	Serine protease htra	*Pyrococcus abyssi*	1.00E-05	41	NP_125928.1

Continued on following page

TABLE 2—*Continued*

Function and sequence[b]	Size (bp)[c]	Protein	Genome	Expect value	%Id[d]	Accession no.
Regulatory functions						
RSGA_547	368	Transcriptional regulator	*Bacteroides thetaiotaomicron*	3.0E-23	48	NP_813510.1
RSGA_230.a	711	Y4mF (low similarity to phage repressor proteins)	*Rhizobium* sp.	3.0E-07	52	NP_443972.1
Cell envelope						
RSGA_066	277	Glycosyltransferase	*Bacteroides thetaiotaomicron*	8.0E-10	43	NP_810265.1
RSGA_695	737	Tgh009 (glycosyltransferase)	*Campylobacter jejuni*	3.0E-13	38	AAS98971.1
RSGA_311	373	Capsular polysaccharide biosynthesis glycosyltransferase	*Erwinia carotovora*	3.0E-14	38	YP_048624.1
RSGA_689.a	1,092	Glycero-phosphotransferase	*Streptococcus pneumoniae*	3.00E-07	55	AAC38753.1
RSGA_469	221	Glycero-phosphotransferase	*Streptococcus pneumoniae*	2.0E-11	47	CAA59773.1
RSGA_524	702	Cell-wall-associated RhsD protein	*Bacteroides thetaiotaomicron*	1.0E-07	26	NP_812114.1
RSGA_343.a	1,201	GDP-mannose-dehydratase	*Geobacter metallireducens*	2.0E-54	58	ZP_00299411.1
RSGA_681	722	Capsule biosynthesis protein CapA	*Bacillus cereus*	1.0E-36	34	ZP_00235460.1
RSGA_221	691	Outer membrane protein precursor	*Bacteroides thetaiotaomicron*	1.0E-21	33	NP_810487.1
RSGA_504.a	1,026	Flippase	*Escherichia coli*	5.0E-18	21	AAK60454.1
RSGA_428	522	COG1538 Outer membrane protein	Not applicable	3.0E-07[e]	28	CDD|11252
Mobile and extrachromosomal element functions						
RSGA_001	213	COG3464: transposase	*Bifidobacterium longum*	1.0E-16	52	ZP_00206779.1
RSGA_421	850	COG3464: transposase	*Bifidobacterium longum*	4.0E-10	46	ZP_00121964.1
RSGA_113	368	Transposase	*Erwinia chrysanthemi*	9.0E-10	72	CAD22869.1
RSGA_520	217	COG3328: transposase	*Geobacter metallireducens*	3.0E-27	77	ZP_00298867.1
RSGA_532.a	741	ISPg4, transposase	*Porphyromonas gingivalis*	2.0E-12	34	NP_904514.19
RSGA_098	522	ISPg7, transposase	*Porphyromonas gingivalis*	2.0E-49	59	Q7MWX1
RSGA_201	879	ISPg7, transposase	*Porphyromonas gingivalis*	9.0E-50	59	Q7MWX1
RSGA_369.a	532	ISPg7, transposase	*Porphyromonas gingivalis*	7.0E-16	56	Q7MWX1
RSGA_517	717	ISPg7, transposase	*Porphyromonas gingivalis*	3.0E-74	60	Q7MWX1
RSGA_559	668	ISPg7, transposase	*Porphyromonas gingivalis*	8.0E-51	60	Q7MWX1
RSGA_618	311	ISPg7, transposase	*Porphyromonas gingivalis*	1.0E-33	65	Q7MWX1
RSGA_619	510	ISPg7, transposase	*Porphyromonas gingivalis*	9.0E-49	58	Q7MWX1
RSGA_621	652	ISPg7, transposase	*Porphyromonas gingivalis*	1.0E-49	59	Q7MWX1
RSGA_629	343	ISPg7, transposase	*Porphyromonas gingivalis*	2.0E-34	62	Q7MWX1
RSGA_640	508	ISPg7, transposase	*Porphyromonas gingivalis*	3.0E-51	60	Q7MWX1
RSGA_646	510	ISPg7, transposase	*Porphyromonas gingivalis*	8.0E-51	59	Q7MWX1
RSGA_649	372	ISPg7, transposase	*Porphyromonas gingivalis*	1.0E-29	70	Q7MWX1
RSGA_655	505	ISPg7, transposase	*Porphyromonas gingivalis*	7.0E-52	59	Q7MWX1
RSGA_660	510	ISPg7, transposase	*Porphyromonas gingivalis*	1.0E-49	59	Q7MWX1
RSGA_663	410	ISPg7, transposase	*Porphyromonas gingivalis*	2.0E-38	58	Q7MWX1
RSGA_666	665	ISPg7, transposase	*Porphyromonas gingivalis*	8.0E-36	59	Q7MWX1
RSGA_699	510	ISPg7, transposase	*Porphyromonas gingivalis*	3.0E-49	58	Q7MWX1
RSGA_700	205	ISPg7, transposase	*Porphyromonas gingivalis*	6.0E-16	65	Q7MWX1
RSGA_701	770	ISPg7, transposase	*Porphyromonas gingivalis*	5.0E-50	60	Q7MWX1
RSGA_708	205	ISPg7, transposase	*Porphyromonas gingivalis*	7.0E-16	65	Q7MWX1
RSGA_150	220	Transposase	*Salmonella enterica*	5.0E-29	78	NP_807108.1
RSGA_506.b	582	Transposase for IS*200*	*Salmonella typhimurium*	3.0E-12	58	NP_460910.2
RSGA_354	248	Transposase	*Sinorhizobium meliloti*	5.0E-05	41	NP_386487.1
RSGA_206	213	Transposase	*Terrabacter* sp.	2.0E-08	44	BAC54155.1
RSGA_115	957	Transposase	Tn*10* delivery vector pHV1248	1.0E-180	100	AAG37885.1
RSGA_633	331	COG4373 Mu-like prophage FluMu protein	Not applicable	5.0E-07[e]	58	CDD|13607
Other						
RSGA_187	864	CG3047-PA	*Drosophila melanogaster*	1.0E-05	24	NP_523475.2
RSGA_672.a	1,171	Thiosulfate-oxidizing enzyme	*Halothiobacillus neapolitanus*	5.00E-06	28	CAC82476.1
RSGA_373	622	Delayed anaerobic gene	*Saccharomyces cerevisiae*	4.0E-05	25	NP_012685.1
RSGA_036	252	COG5557: Polysulphide reductase	*Ralstonia eutropha*	2.0E-10	38	ZP_00202522.1
RSGA_448	370	KOG2548 SWAP mRNA splicing regulator	Not applicable	1.0E-06[e]	27	CDD|20334
RSGA_610	297	KOG4157: β-1,6-*N*-acetylglucosaminyltransferase	Not applicable	6.0E-04[e]	29	CDD|21934
RSGA_702	1,355	COG1487 nuclei acid-binding protein	Not applicable	3.0E-06[e]	54	CDD|11201

Continued on following page

TABLE 2—*Continued*

Function and sequence[b]	Size (bp)[c]	Protein	Genome	Expect value	%Id[d]	Accession no.
Unknown function						
RSGA_211	681	Hypothetical protein	*Amycolatopsis orientalis*	3.0E-07	33	T17479
RSGA_528.b	405	Hypothetical protein	*Aquifex aeolicus*	7.0E-04	30	AAC07358.1
RSGA_002	650	Conserved hypothetical protein	*Bacteroides thetaiotaomicron*	5.0E-58	50	NP_812957.1
RSGA_153	221	Conserved hypothetical protein	*Bacteroides thetaiotaomicron*	4.0E-13	56	NP_810397.1
RSGA_166	269	Conserved hypothetical protein	*Bacteroides thetaiotaomicron*	2.0E-22	61	NP_811318.1
RSGA_228	361	Conserved hypothetical protein	*Bacteroides thetaiotaomicron*	8.0E-36	62	NP_811318.1
RSGA_226	177	Hypothetical protein	*Bartonella henselae*	4.0E-14	68	YP_034241.1
RSGA_643.a	894	Unnamed protein product	*Candida glabrata*	5.00E-06	35	XP_447613.1
RSGA_462	342	Conserved hypothetical protein	*Cytophaga hutchinsonii*	3.0E-31	59	ZP_00307856.1
RSGA_134	416	Hypothetical protein	*Fusobacterium nucleatum*	1.0E-21	39	ZP_00145043.1
RSGA_058.b	689	Conserved hypothetical protein	*Methanothermobacter thermautotrophicus*	3.0E-11	52	AAB85392.1
RSGA_230.b	711	Conserved hypothetical protein	*Caulobacter crescentus*	5.0E-07	27	NP_421532.1
RSGA_527	684	Hypothetical protein	*Photobacterium profundum*	1.0E-15	34	CAG20251.1
RSGA_235	1,046	Conserved hypothetical protein	*Porphyromonas gingivalis*	4.0E-04	35	NP_905667.1
RSGA_388	658	Hypothetical protein	*Vibrio cholerae*	1.0E-04	25	NP_231406.1
RSGA_173	224	Hypothetical protein	*Shewanella oneidensis*	1.0E-24[f]	32	NP_719538.1
RSGA_670.a	927	Conserved hypothetical protein	*Sinorhizobium meliloti*	2.00E-08	32	NP_386983.1
RSGA_070	1,273	Hypothetical protein	*Vibrio parahaemolyticus*	8.0E-33	25	NP_797698.1
RSGA_503	1,310	Hypothetical protein	*Wolbachia* sp.	3.0E-12	36	NP_966293.1
RSGA_574	294	Hypothetical protein	*Zymomonas mobilis*	3.0E-04	43	NP_758980.1
RSGA_387	400	COG4645: uncharacterized conserved protein	Not applicable	1.0E-07[e]	24	CDD|13789

[a] The similarities, unless indicated by superscripts *e* and *f*, were based on the BLASTX search expectation values against the merged GenBank nonredundant protein database and the *F. succinogenes* translated coding sequence database. E-values denoted by the superscripts *e* and *f* were derived from the RPS-BLAST and the PSI-BLAST searches, respectively (see below). Only sequences with E-values higher than 0.001 in all of the above-mentioned three BLAST searches compared to the *F. succinogenes* proteins were regarded as absent from the genome of *F. succinogenes.*

[b] If more than one gene segment was present in a contig, they are listed as contigs a, b, and c, respectively. *F. succinogenes* genes in these contigs are italicized.

[c] That is, the length of the contig.

[d] %Id, percent amino acid sequence identity.

[e] Results obtained from RPS-BLAST searches against the conserved domain databases.

[f] E-value obtained from PSI-BLAST searches after the 3rd iteration.

similarity to those in *F. succinogenes* (Table S1), and ca. 93% of these showed greatest similarity to proteins in *F. succinogenes*.

Among the 137 contigs that contained at least two genes, 118 coded for proteins that showed similarities to two or three different proteins encoded by the *F. succinogenes* genome (Tables S1 and S2 in the supplemental material), 15 had one or two matches to the proteins from *F. succinogenes* and one match to the GenBank nraa excluding *F. succinogenes* proteins, and 4 contained two genes with no matches to proteins from *F. succinogenes* (Table 2). Of the 118 contigs encoding two different proteins with similarity to proteins from *F. succinogenes*, only 48 were encoded by genes that had the same gene order on the chromosome as those in *F. succinogenes* (Table S2 in the supplemental material), indicating substantial differences between the two species in terms of gene organization.

Carbohydrate-active enzymes. All of the cellulases and xylanases identified in *F. intestinalis* DR7 exhibited highest amino acid similarity to *F. succinogenes* S85 glycosyl hydrolases, although the percent identities did not exceed 89%. The glucanases corresponded to four family 5, two family 8, four family 9, one family 16, one family 18, one family 51 and one family 74 glycoside hydrolases. Two family 43 and one family 10 xylanases were also present in the library, and there was one family 26 mannanase gene and three polysaccharide lyases (Table 3).

Two clones contained gene fragments that encoded proteins exhibiting high similarity to the endoglucanase 3 encoded by the *cel-3* that was previously cloned from *F. intestinalis* DR7 (18) (accession L39840), and they accounted for ca. 61% of the coverage of the gene. The *cel-3* gene, which encodes a family 5 catalytic domain, had 68% DNA sequence similarity to its counterpart in *F. succinogenes*, and the similarity ranged from 30 to 80% in different regions.

Four genes showed highest similarities to the *F. succinogenes* endoglucanase gene *celF* (22). CelF, formerly known as endoglucanase 2 (24), is a major endoglucanase in the periplasm and extracellular culture fluid fraction of *F. succinogenes*. A putative *celF* gene fragment was cloned from *F. intestinalis* strain NR9 by Bera-Maillet et al. (6), which had 90% DNA similarity to the *celF* genes in *F. intestinalis* DR7, thereby documenting the universality of this gene in *Fibrobacter* sp.

Although all of the cellulase and xylanases genes identified encoded proteins showing the highest similarity to their counterparts in *F. succinogenes*, it is remarkable that 47% of these gene products showed the highest similarities (identities ranged from 28 to 44%) to translation products in the *Cytophaga* species when a BLASTX search was done against the GenBank nraa database, indicating the close relationship between the cellulase systems of the *Fibrobacter* and *Cytophaga* genera.

Genes involved in energy metabolism. Most of the genes identified with a role of energy metabolism coded for proteins with the highest similarity to those from *F. succinogenes*. However, two arsenate reductase genes were identified, which coded for proteins with 76% similarity over a 52-amino-acid

TABLE 3. Putative carbohydrate active enzymes of *F. intestinalis* identified by SSH

Sequence	Size (bp)	*F. succinogenes* ORF with closest match	CAZy family[a]	Expect value	%Id[b]
RSGA_555	809	ORF03598 endoglucanase 3 (A33598)[c]	GH05	4E-78	51
RSGA_146	555			5E-46	68
RSGA_490	604	ORF03025: endoglucanase	GH05	1E-21	37
RSGA_261	331	ORF01077: cellulase	GH05	8e-26	49
RSGA_243	534	ORFA00267: endoglucanase	GH05	2E-45	44
RSGA_190	503	ORF00143: endoglucanase	GH08	8E-58	65
RSGA_110	230			9E-24	65
RSGA_607	305	ORF01918: endoglucanase a	GH08	3E-09	36
RSGA_675	749	ORF02552: glycosyl hydrolase	GH09	1E-79	59
RSGA_678	564	ORF02861: glycosyl hydrolase	GH09	5E-71	60
RSGA_586	718			4E-62	49
RSGA_484	686			4E-44	41
RSGA_472	597			9E-44	49
RSGA_259	732			1E-39	35
RSGA_097	368			1E-17	33
RSGA_548	286	ORF02862: glycosyl hydrolase	GH09	2E-25	58
RSGA_422	156			4E-06	46
RSGA_280	403	ORFA00235: endoglucanase E (AAC44385)[c]	GH09	3E-43	60
RSGA_090	353	ORF02325: endo-1,4-β-xylanase	GH10	5E-35	57
RSGA_383	454	ORFA00451: β-glucanase precursor	GH16	7E-31	59
RSGA_360	822	ORF01086: glycosyl hydrolase	GH18	6E-73	57
RSGA_298	485	ORF00106: mannanase	GH26	7E-43	52
RSGA_029	390	ORF02754: xylanase	GH43	7E-20	76
RSGA_040	421			1E-72	92
RSGA_184	554	ORF03727: xylanase	GH43	4E-65	60
RSGA_705	1,144			2E-31	40
RSGA_501	1,137	ORF00990: endoglucanase	GH45	2E-16	56
RSGA_131	287	ORF00279: endoglucanase (*celF*) (AAC45377)[c]	GH51	4E-51	94
RSGA_183	757			1E-111	74
RSGA_282	440			6E-62	75
RSGA_705	1,144			1E-23	79
RSGA_608	489	ORF02265: glycosyl hydrolase	GH57	1E-47	51
RSGA_172	283			2E-12	67
RSGA_463	376	ORF02285: glycosyl hydrolase	GH57	2E-49	67
RSGA_536	648	ORF04367: glycosyl hydrolase	GH57	1E-34	35
RSGA_073	428	ORF04614: endoglucanase	GH74	3E-13	32
RSGA_448	370	ORF02936: pectate lyase	PL1	7E-03	77
RSGA_479	713	ORF00501: rhamnogalacturonan lyase	PL11	9E-70	54
RSGA_273	193	ORFA00711: pectate lyase	PL1	7E-32	87
RSGA_313	325	ORF02021: 4-α-glucanotransferase	GH77	3E-39	66
RSGA_395	361	ORF00267: 1,4-α-glucan branching enzyme	GH13	2E-66	90
RSGA_691	285	ORF02803: carbohydrate phosphorylase		3E-46	90
RSGA_657	497	ORF03715: α-galactosidase	GH27	1E-53	84
RSGA_507	548	ORF00504: acetylesterase	CE12	6E-44	52

[a] Carbohydrate-active enzyme families. GH, glycosyl hydrolase family; PL, pectate lyase family; CE, carbohydrate esterase family.
[b] Amino acid sequence similarity.
[c] GenBank accession numbers are indicated in parentheses.

sequence to the same enzyme in *Clostridium perfringens*, and 58% similarity over a 70-amino-acid sequence to the same enzyme in *Chlorobium tepidum* but had no counterpart in the *F. succinogenes* genome (Table 2). Arsenate is reduced to arsenite by some bacteria that grow anaerobically using the nonrespiratory substrate lactate as the electron donor (16). The existence of arsenate reductases may indicate the existence of arsenate respiration in *F. intestinalis*, although the possibility that it has a role limited to detoxification, or indeed, an unrecognized role, cannot be excluded.

Another gene that coded for a protein in this category that has no homolog with the *F. succinogenes* genome is a phosphoglycerate mutase (2,3-bisphosphoglycerate dependent) with 67% similarity over 86 amino acid residues to the C terminus of the 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase gene in *Ralstonia solanacearum* (38). In contrast to *F. intestinalis*, the *F. succinogenes* genome contains a putative phosphoglycerate mutase, which is a 2,3-bisphosphoglycerate-independent enzyme. In nature two types of phosphoglycerate mutases catalyze the conversion of 3-phosphoglycerate to 2-phosphoglycerate. One uses 2,3-bisphosphoglycerate as a cofactor and catalyzes an intermolecular reaction with a phosphorylated histidine residue intermediate (13). The other type is the cofactor-independent phosphoglycerate mutase. Animals mainly contain the first type, while the second type is often found in plants. However, both types are found in bacteria. Whether *F. intestinalis* DR7 has a functional 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase remains to be determined.

Genes involved in transport and binding. Thirteen of the genes identified coded for transport and binding proteins that

showed no similarity to *F. succinogenes*. These included three proteins involved in ferrous or ferric ion transport, a phosphate-binding protein, three ABC transporter proteins that may be responsible for nitrate/sulfonate/bicarbonate and polysaccharide/polyol phosphate transport, two Na^+/H^+ antiporters, one chromate transport protein, one periplasmic substrate-binding protein, one putative permease, and a transporter of unknown function (Table 2). The differences in transport capabilities are most likely a reflection of the different environmental niches that the two species occupy.

Genes involved in DNA metabolism. Nine of the genes in this functional group coded for proteins that showed greatest similarity to those in bacteria other than *F. succinogenes*. These included an anaerobic ribonucleoside-triphosphate reductase involved in DNA replication, a DNA damage-inducible protein, a nuclease subunit of the exonuclease complex, three type I, one type II, and one type IIS restriction modification enzymes and a cytosine-specific modification methyltransferase. The type IIS restriction-modification enzyme identified was most similar to a putative type IIS restriction-modification enzyme in *Bacteroides thetaiotaomicron* (41), whereas this kind of restriction-modification system was absent from the genome of *F. succinogenes*. Previous studies have shown the existence of a type II restriction-modification system in *F. succinogenes* (17), and type I and III restriction-modification systems have also been identified from the genome of *F. succinogenes* (31).

Genes involved in cell envelope metabolism. The putative proteins encoded by 11 genes with roles in cell envelope metabolism showed no similarity to proteins in *F. succinogenes*. Most of these proteins were involved in the biosynthesis of cell surface polysaccharides and included three glycosyl transferases, two glycerophosphotransferases, one capsule biosynthesis protein, one GDP-D-mannose dehydratase, one cell-wall-associated protein, two outer membrane proteins, and one putative flippase. Since no research has been conducted on the cell surface structure of *F. intestinalis*, the possible function of each protein is speculative. In addition, because of the observed differences in the genes between *F. intestinalis* and *F. succinogenes*, differences in the organization and structure of the cell surfaces would be expected.

Transposases of *F. intestinalis*. In the present study, 30 different transposase gene segments were identified in *F. intestinalis* (Table 2). Ten of the putative proteins had high similarity to transposases from various bacteria, while twenty of them were homologous to a transposase, encoded by ISPg7, from *Porphyromonas gingivalis*. Since the genes identified from the SSH contigs were short, none of the encoded proteins matched the whole ISPg7 transposase. However, 11 of the 20 protein products, which had >95% DNA identity within the group, were homologous to some 110 amino acids at the N terminus of the *P. gingivalis* genes with 58 to 65% similarities, indicating the existence of multiple copies of the same insertion sequence (IS) elements. Thus, altogether it would appear that at least 11 different types of IS elements exist in *F. intestinalis*. TBLASTX searches of the *F. succinogenes* genome with each of the *F. intestinalis* transposase gene fragments did not reveal homologous putative transposases (31).

BLAST searches indicated that the 30 transposase gene fragments belong to five IS families (http://www-is.biotoul.fr/). Three belong to the ISL3 family, three belong to the IS*4* family, one belongs to the IS*256* family, and three belong to the IS*605* family. The remaining 20 transposases, which were homologous to ISPg7 from *P. gingivalis*, belong to the IS*5* family (Fig. 2).

Transcription of glycosyl hydrolases, transposases, and restriction-modification enzymes in *F. intestinalis* DR7 that have low DNA similarity to *F. succinogenes* S85. To examine whether the genes identified in *F. intestinalis* were functional or simply vestigial DNA, mRNA synthesis for 23 glycosyl hydrolases, 9 transposases, and 7 restriction-modification enzymes were determined by dot hybridization of mRNA from cells grown on glucose and cellulose as an energy source, respectively (Table 4). Of the 23 glycosyl hydrolase genes, 16 were expressed and, of these, 8 were expressed at higher levels and 2 were expressed at a lower level during growth on cellulose compared to growth on glucose. Six genes exhibited no detectable differences in expression during growth with either substrate. The eight genes with higher levels of expression during growth on cellulose included six endoglucanases, one rhamnogalacturonan lyase, and one lichenase. An endoglucanase E homolog (21, 25), a endoglucanase 3 homolog (21, 25), and an uncharacterized gene (RSGA_190) were expressed at a much higher level (Table 4).

The *celF* gene homolog (RSGA_183), which exhibited the highest expression level, was not differentially expressed on either glucose or cellulose. A family 9 glycosyl hydrolase (RSGA_675) and a family 18 chitinase (RSGA_360) were also expressed at a relative high level. A 1,4-α-glucan branching enzyme (RSGA_395) was expressed at a low level and perhaps was slightly downregulated during growth on cellulose.

Among the transposases RSGA_532 and RSGA_150 were substantially upregulated by growth on cellulose, whereas the remaining transposases were expressed constitutively, except for RSGA_113 that was not expressed. The type I restriction enzyme gene RSGA_009 was substantially upregulated by growth on cellulose, whereas the type III restriction enzyme gene RSGA_245 was downregulated. The housekeeping genes phosphoglycerate kinase and glucose-6-phosphate isomerase used as the reference for assessing expression were expressed constitutively.

DISCUSSION

Suppressive subtractive hybridization was used to study the genomic differences between two species in the genus of *Fibrobacter* by determining the genes present in *F. intestinalis* that exhibit low similarity to those present in *F. succinogenes*, as well as those that are completely absent from the genome of *F. succinogenes*. As detemined by a BLASTN search of the *F. intestinalis* contigs, 61% of the genes that were identified by SSH had no DNA similarity to the genome of *F. succinogenes*, and only 9.7% had more than 80% DNA identity, indicating the enrichment of *F. intestinalis* specific genes by SSH and, as well, the level of genomic diversity between these two species. However, at the protein level, genes having low DNA identity still showed high amino acid sequence similarity with 80% of the genes having a higher BLASTX similarity to the *F. succinogenes* proteins than to other organisms in the GenBank nraa database (Table 1 and Table S1 in the supplemental material).

In the present study as in previous SSH studies with other

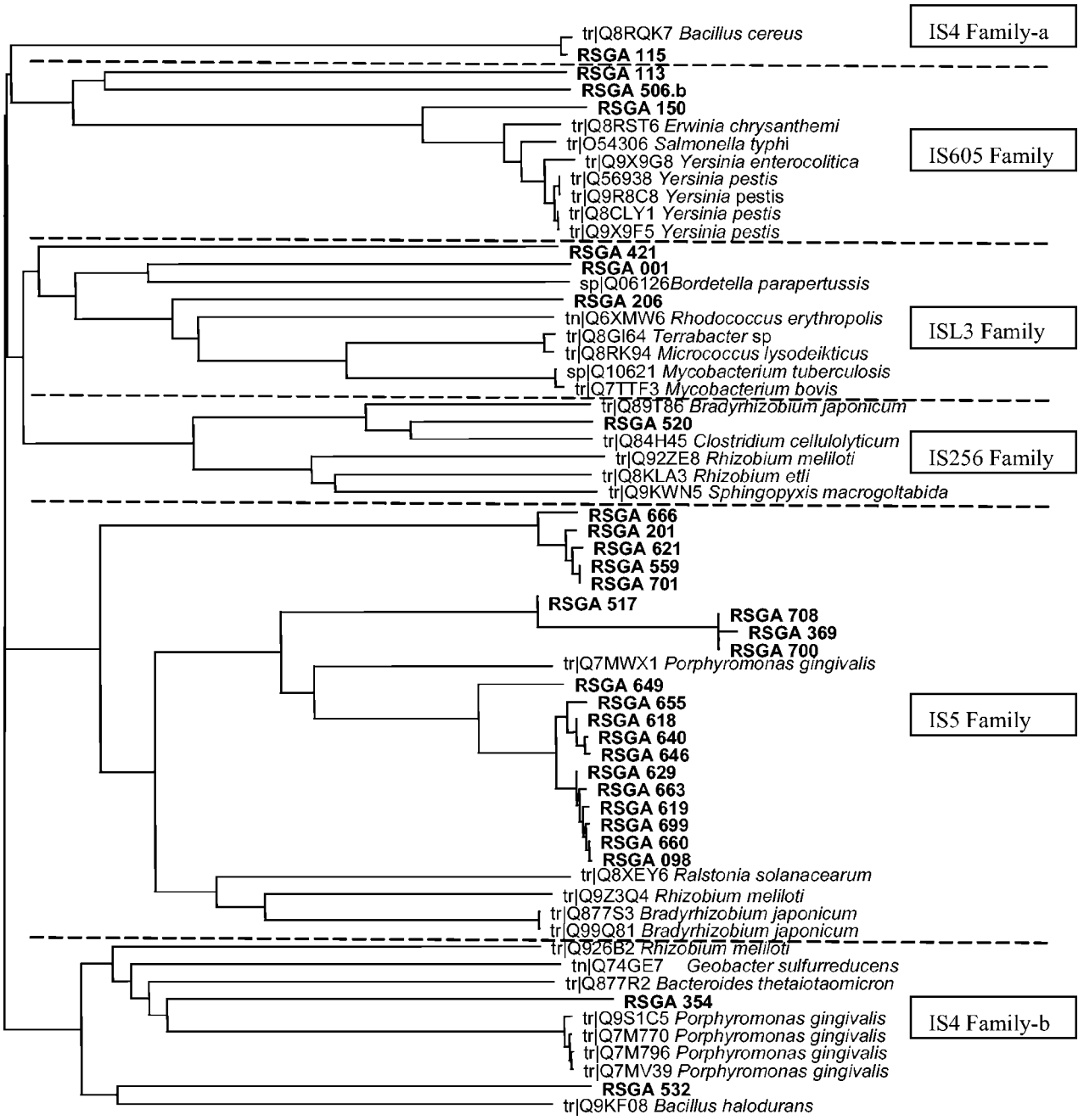

FIG. 2. Dendrogram of putative transposases and their closest relatives identified from the Swiss-Prot database (http://www.expasy.org/), showing that they belong to different insertion sequence families. The genes from *F. intestinalis* were highlighted as RSGA plus the number. The genes closest to the *F. intestinalis* clones were selected by using a BLASTP search against the Swiss-Prot database. The accession number and name of organisms are shown in the figure. The tree was constructed by using CLUSTAL W (version 1.83) and TreeView (version 1.6.6). The dashed line separates the different transposase families. The genes are clustered in different branches based on their sequence similarities. However, conclusive phylogenetic relationships between the genes cannot be accurately determined because the SSH transposase genes are incomplete sequences.

microbial species (4, 32), many of the *F. intestinalis* genes identified by SSH were more closely related to the genome that was used in the subtraction process, in this case *F. succinogenes*. Consequently, any gene detected by SSH as being present only in *F. intestinalis* should be treated as putatively unique, and other supporting evidence is essential for conclusive identification. In our case, the availability of the complete genome sequence of *F. succinogenes* (the driver) has greatly facilitated the interpretation of the SSH data, since the amino acid identities could be found by similarity searches.

SSH showed that, although divergent, *F. intestinalis* and *F. succinogenes* are common in several aspects. In *F. succinogenes*, 113 putative enzymes have been identified to have a role in plant cell wall degradation, including 40 cellulases and 29 xy-

TABLE 4. Differential expression of putative glycosyl hydrolases, transposases, and restriction modification genes in *F. intestinalis*

Enzyme type and clone	Name	Family	Relative expression		Expression ratio[a] (cellulose/glucose)
			Cellulose	glucose	
Carbohydrate active enzymes					
RSGA_183	ORF00279: endoglucanase (*celF*)	GH51	590	831	0.71
RSGA_190	ORF00143: endoglucanase	GH08	518	259	2.00
RSGA_280	ORFA00235: endoglucanase E	GH09	472	126	3.74
RSGA_073	ORF04614: endoglucanase	GH74	292	517	0.56
RSGA_146	ORF03598: endoglucanase 3	GH05	282	98	2.88
RSGA_490	ORF03025: endoglucanase	GH05	84	ND[b]	>10
RSGA_607	ORF01918: endoglucanase	GH08	71	ND	>10
RSGA_243	ORFA00267: endoglucanase	GH05	ND	ND	—[c]
RSGA_675	ORF02552: glycosyl hydrolase	GH09	488	457	1.07
RSGA_360	ORF01086: glycosyl hydrolase	GH18	258	304	0.85
RSGA_678	ORF02861: glycosyl hydrolase	GH09	107	ND	>10
RSGA_040	ORF02754: xylanase	GH43	ND	ND	—
RSGA_090	ORF02325: xylanase	GH10	ND	ND	—
RSGA_184	ORF03727: xylanase	GH43	ND	ND	—
RSGA_273	ORFA00711: pectate lyase	PL1	61	86	0.71
RSGA_479	ORF00501: rhamnogalacturonan lyase	PL11	148	ND	>10
RSGA_448	ORF02936: pectate lyase	PL1	ND	ND	—
RSGA_383	ORFA00451: β-glucanase	GH16	58	ND	>10
RSGA_298	ORF00106: mannanase	GH26	48	71	0.68
RSGA_395	ORF00267: 1,4-α-glucan branching enzyme	GH13	29	97	0.30
RSGA_657	ORF03715: α-galactosidase	GH27	ND	ND	—
RSGA_313	ORF02021: 4-α-glucanotransferase	GH77	66	71	0.93
RSGA_507	ORF00504: acetylesterase	CE12	ND	ND	—
Transposases					
RSGA_201	Transposase	IS*5*	478	569	0.84
RSGA_421	Transposase	ISL3	118	161	0.73
RSGA_001	IS*1557*, transposase	ISL3	113	148	0.76
RSGA_517	Transposase	IS*5*	56	113	0.49
RSGA_532	Transposase	IS*4*	42	ND	>10
RSGA_520	Transposase	IS*256*	39	85	0.46
RSGA_150	Transposase	IS*605*	34	ND	>10
RSGA_206	Transposase	ISL3	34	20	1.70
RSGA_113	Transposase	IS*605*	ND	ND	—
Restriction-modification enzymes					
RSGA_009	Type I restriction enzyme	—	59	ND	>10
RSGA_225	Type I restriction enzyme (modification subunit)	—	55	99	0.55
RSGA_157	Modification methyltransferase, cytosine specific	—	ND	45	<0.1
RSGA_245	ORF03805: type III restriction enzyme	—	ND	ND	—
RSGA_390	ORF03270: methyltransferase gidB/glucose-inhibited division protein B	—	ND	ND	—
RSGA_194	HsdR (endonuclease subunit of type I restriction enzyme)	—	ND	ND	—
RSGA_024	Type IIS restriction/modification enzyme	—	ND	ND	—
Other		—			
RSGA_692	ORF02881: phosphoglycerate kinase	—	146	161	0.91
RSGA_128	ORF00167: glucose-6-phosphate isomerase	—	36	32	1.12
RSGA_285	ORF01675: ABC transporter, ATP-binding protein	—	36	ND	>10
RSGA_706	Aminotransferase	—	ND	ND	—

[a] Ratios were calculated as described in Materials and Methods.
[b] Expression was not detected (ND) under the experimental conditions adopted.
[c] –, not applicable.

lanases (31). All of the subtractive cellulases and xylanases identified in *F. intestinalis* exhibit highest similarity to the comparable proteins in *F. succinogenes*. The *celF* gene, which codes for a predominant cellulase in *F. succinogenes* and belongs to family 51 glycosyl hydrolases, was present in *F. intestinalis* and was highly expressed in cells grown with either glucose or cellulose as the sole source of carbohydrate, just as it is in *F. succinogenes*. Most members in family 51 are arabinofuranosidases, which do not exhibit endoglucanase activity (7); however, CelF of *F. succinogenes* S85 had no activity on *p*-nitrophenol arabinofuranoside but had endoglucanase activity (22). Since the amino acid sequence identity between CelF and the

homologous protein from *F. intestinalis* was described above 75%, whereas the identity between CelF and other family 51 arabinofuranosidases were <25%, it is very likely that the *celF* counterpart is also cellulolytic rather than exhibiting arabinase activity, although further experiments are needed to confirm this suggestion.

Conservation of the cellulases between the two species indicates the unique nature of the cellulase system in the *Fibrobacter* genus. Distinct strategies are exploited by different bacteria for the complete degradation of cellulosic materials. Generally, in anaerobic bacteria the cellulases are cell associated and are an intricate part of a cellulosome structure (20). However, none of the genes in the *F. succinogenes* genome or the *F. intestinalis* SSH clones exhibit similarity to the classical cohesin/dockerin domain signatures or to the scaffolding protein of a cellulosome complex. It has been reported that some of the glycosyl hydrolases in *F. succinogenes* have a basic terminal domain but lack high amino acid sequence similarity (21). One *F. intestinalis* contig (contig RSG_705) encodes a protein that has high similarity to the C terminus of ORF03727 in *F. succinogenes* that encodes a xylanase with a basic terminal domain. The basic terminal domain was proposed to interact with the negatively charged cell surface rather than interact with a scaffolding protein (21). These data are insufficient to develop a model for cellulose digestion, but it does indicate that *Fibrobacter* species have a cellulase organization different from the typical cellulosomal structure present in many other anaerobic bacteria.

Besides conservation of the cellulase system, ca. 80% of the remaining genes that were identified had the highest similarity to genes in *F. succinogenes* (Table 1 and Table S1 in the supplemental material). Functions of many of the genes were assigned by similarity search, but some of them remain unknown and were annotated as hypothetical proteins. Among them are several hypothetical proteins that exist only within the two *Fibrobacter* species. As the genetics and physiology of this genus become more fully characterized, the true distribution and roles of these hypothetical and conserved hypothetical proteins will be determined.

SSH also showed clear differences between *F. intestinalis* and *F. succinogenes*. About 33% of the genome of *F. intestinalis* DR7 was absent or diverged greatly from the genome of *F. succinogenes* S85. Initial analyses reveal that there is at least 66 kb of DNA in the *F. intestinalis* genome that lacks similarity to *F. succinogenes* at the protein level.

Organization of the *F. intestinalis* DR7 genome. There is evidence of substantive genome rearrangements in *F. intestinalis* compared to *F. succinogenes*. If the organization of genes in *F. intestinalis* and *F. succinogenes* were identical, adjacent gene pairs should be the same in the two species. However, 89 of the 137 contigs or 65% of the contigs with segments of two adjacent genes did not contain the same gene order in *F. succinogenes* (Table 2 and Tables S1 and S2 in the supplemental material). Homologs of genes adjacent in *F. intestinalis* are 1.8 to 1,865 kbp apart in the *F. succinogenes* chromosome (Fig. S1 in the supplemental material). Furthermore, the fact that the difference in distribution of genes approached randomness suggests that the reorganization has been random rather than changes in restricted areas of the chromosomes of the species.

A fascinating question is the mechanism underlying the difference in genome organization between *F. intestinalis* and *F. succinogenes*. RecA-dependent recombination can result in rearrangements in the order of genes on the chromosome which may include duplication, translocation or inversion (12). IS elements with long repeats are often involved in chromosomal rearrangements. There is a positive correlation between the total number of IS elements and degree of rearrangement that is consistent with a role of IS elements in transposition and inversion events of DNA in a genome (27). Transposases are integral components of the IS elements. Since at least five transposase families have been identified in the genome, the repeat sequences of IS elements are very likely present within the *F. intestinalis* genome.

In some bacterial genomes, the restriction-modification elements were shown to be polymorphic (36). Similar polymorphisms were observed in *F. succinogenes* (31) and *F. intestinalis*. In an SSH study of *Ruminococcus flavefaciens* strain FD-1, Antonopoulos et al. (4) reported on several transposases and a restriction enzyme that were absent from strain JM1. These authors presented evidence of apparent domain swapping among the glycosyl hydrolases, which is probably related to the presence of transposases and the restriction enzyme. Similarly, in *F. intestinalis* the restriction-modification systems may have contributed to the genome rearrangement, since restriction-modification systems participate in generating bacterial diversity by promoting homologous recombination (11). Furthermore, a variety of mobile genetic units were found to be linked to restriction-modification gene complexes (15).

The difference in the gene complement of the *F. intestinalis* and *F. succinogenes* could also be due to horizontal gene transfer (HGT). HGT is commonly assumed to be an important force in shaping the prokaryote genome (33). Since limited information is known on the strains in the genus, HGT can only be inferred from the deviant base composition (including G+C content and codon usage) and abnormal sequence similarity. Since 80% of the genes identified are most closely related to the *F. succinogenes* genome, it is likely that most of these genes originate from the same ancestor and have evolved by internal shuffling, duplication, and/or accumulation of point mutation. The codon usage distance among the cellulases from *F. succinogenes* and *F. intestinalis* and transposases from *F. intestinalis* was calculated, and a dendrogram generated (Fig. 3). The similar codon usage, especially for the cellulases, indicates that most of the genes from both species originated from the same ancestor. In contrast, codon usage of the transposase genes is very different from that of the cellulase genes (Fig. 3). The clustering of transposases in three branches distinct from most of the cellulases suggests that they were recently introduced by HGT in contrast to practically all of the cellulases that have evolved slowly and are presumably restricted to the *Fibrobacter* genus. Differences in the G+C composition (G+C%) and G+C composition at the third codon (G+C3%) of genes compared to the host genome are also used as evidence for HGT (10). The mean and standard deviation values for G+C% and G+C3% for the transposase gene segments were 49.6% ± 0.1% and 65.7% ± 0.1%, respectively, whereas the values for the carbohydrate active enzyme gene segments were 49.6% ± 0.3% and 61.6% ± 0.1%, respectively. As a basis for these comparisons the mean values of the total G+C% and G+C3% of the SSH genes were 48.6% ± 3.7%

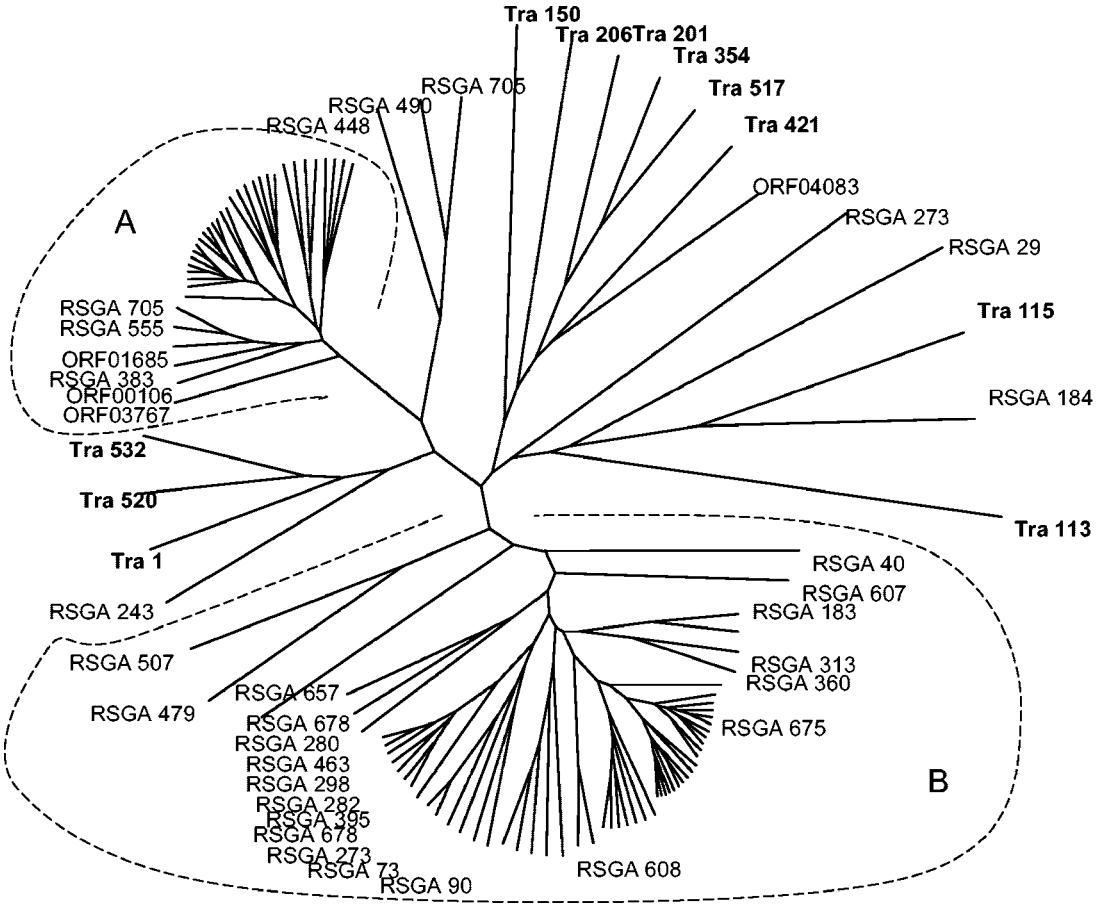

FIG. 3. Dendrogram of the relationships between the codon usage of *F. succinogenes* cellulolytic enzymes and *F. intestinalis* cellulolytic enzymes and transposases. Clones of SSH from *F. intestinalis* DR7 were named as RSGA plus the number for cellulolytic enzymes and Tra plus the number for transposases. The transposases showing the highest similarity to the ISPg7 transposase of *P. gingivalis* are represented by Tra201 and Tra532 because of their nearly identical DNA sequences. Genes from *F. succinogenes* were labeled with ORF numbers. Most of the genes from *F. succinogenes* were clustered in regions A and B as shown in the figures. Region A contains 33 *F. succinogenes* genes and includes ORFs 106, 267, 924, 1077, 1107, 1635, 1685, 1708, 1766, 2210, 2280, 2282, 2283, 2284, 2321, 2325, 2552, 2754, 2861, 2862, 3023, 3555, 3598, 3766, 3767, 3815, 3983, 4138, 4190, 4208, 4500, 4614, and A00451. Region B contains 41 *F. succinogenes* genes and includes ORFs 65, 143, 241, 279, 369, 601, 925, 990, 1079, 1086, 1088, 1258, 1538, 1785, 1918, 2003, 2115, 2519, 2520, 2686, 2715, 2748, 3025, 3444, 3674, 3685, 3687, 3690, 3711, 3713, 3720, 3726, 3727, 4081, 4082, 4341, 4367, 4534, 4628, A00235, and A00267.

and 56.9% ± 8.4%, respectively. From these calculations it can be seen that the G+C% value for the transposases was similar to that of the carbohydrate active enzymes and all of the SSH genes, although the variation of the latter was greater, as indicated by the greater standard deviation. Similarly, the G+C3% values for the transposases were higher and the variations were greater than for both the carbohydrate active enzymes and the SSH genes. However, at least half of the gene segments of the carbohydrate active enzymes showed as great a variation as that of the transposases, which indicates that these criteria are of limited value, at least for the *F. intestinalis* genome. However, the fact that BLAST searches showed that the transposases and restriction-modification enzymes have the highest similarities to proteins in other unrelated organisms and no similarity to proteins of *F. succinogenes* (Table 2) is supportive evidence of HGT. However, due to the lack of genetic information within the *Fibrobacter* genus, we cannot tell whether the differences among the genes are due to HGT into *F. intestinalis* or deletion from the genome of *F. succinogenes*. Further study is needed to assess the distribution of the specific genes among other phylogenetically related species. This conclusion is consistent with the complexity observed in previous studies to determine whether evolving differences in bacteria are due to HGT or gene loss (14, 35).

Vast and diverse communities of microbes colonize the gastrointestinal tracts of animals and are essential to the nutrition of the animal. Despite the difficulty in resolving the roles of HGT and gene loss in the evolution of genomes of bacteria in the intestinal environment, HGT among these organisms would seem very likely due to the high bacteriophage numbers (40) and close proximity of organisms in these densely populated ecosystems (8). As a result, it may not be a surprise that 45% of the genes identified by SSH had highest similarities to the genes of other gastrointestinal bacteria. Among them, 15 genes (Table 2 and unpublished data) that exist in *F. intestinalis* showed greatest similarity to genes from the ubiquitous intestinal bacterium *B. thetaiotaomicron* (41).

F. succinogenes strain S85 has been the model organism for

studying the mechanism of efficient cellulose degradation within the genus. SSH between the genomes of *F. intestinalis* and *F. succinogenes* has demonstrated that, although most of the divergent genes including those coding for cellulolytic enzymes still have a high degree of conservation, there is extensive genome rearrangement in *F. intestinalis* relative to *F. succinogenes*. Furthermore, we have identified novel genes in *F. intestinalis* whose presence may be attributed to the numerous IS elements, different restriction-modification systems in *F. intestinalis* and perhaps HGT. These data provide a fascinating snapshot of the evolving interrelationships of *F. intestinalis* with *F. succinogenes*, and with other microorganisms in the gastrointestinal tract, that provides a solid foundation for further studies of the genomics and physiology of fibrolytic organisms in this unique ecosystem.

ACKNOWLEDGMENTS

This research was supported by U.S. Department of Agriculture grants for the North American Consortium for Genomics of Fibrolytic Ruminal Bacteria (grant 00-52100-9618), CSREES, IFAFS, and the Natural Science and Engineering Research Council of Canada.

We thank Lewis Lukens for advice on Bioinformatics and Reggie Lo for proofreading the manuscript.

REFERENCES

1. **Ahn, S. J., J. Costa, and J. R. Emanuel.** 1996. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. Nucleic Acids Res. **24:**2623–2625.
2. **Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A. Lukyanov, E. D. Sverdlov, and D. E. Berg.** 1998. PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori.* Proc. Natl. Acad. Sci. USA **95:**13108–13113.
3. **Amann, R. I., C. H. Lin, R. Key, L. Montgomery, and D. A. Stahl.** 1992. Diversity among *Fibrobacter* isolates: toward a phylogenetic classification. Syst. Appl. Microbiol. **15:**23–31.
4. **Antonopoulos, D. A., K. E. Nelson, M. Morrison, and B. A. White.** 2004. Strain-specific genomic regions of *Ruminococcus flavefaciens* FD-1 as revealed by combinatorial random-phase genome sequencing and suppressive subtractive hybridization. Environ. Microbiol. **6:**335–346.
5. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Sidman, J. A. Smith, and K. Struhl.** 1994. Current protocols in molecular biology, p. 2.4.1–2.4.5. Greene Publishing Associates/John Wiley & Sons, Inc., New York, N.Y.
6. **Bera-Maillet, C., Y. Ribot, and E. Forano.** 2004. Fiber-degrading systems of different strains of the genus *Fibrobacter.* Appl. Environ. Microbiol. **70:**2172–2179.
7. **Beylot, M. H., V. A. McKie, A. G. Voragen, C. H. Doeswijk-Voragen, and H. J. Gilbert.** 2001. The *Pseudomonas cellulosa* glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity. Biochem. J. **358:**607–614.
8. **Flint, H. J.** 1994. Molecular genetics of obligate anaerobes from the rumen. FEMS Microbiol. Lett. **121:**259–267.
9. **Forsberg, C. W., E. Forano, and A. Chesson.** 2000. Microbial adherence to plant cell wall and enzymatic hydrolysis, p. 79–88. *In* P. B. Cronje (ed.), Ruminant physiology digestion, metabolism, growth, and reproduction. CABI Publishing, Wallingford, Oxon, United Kingdom.
10. **Garcia-Vallve, S., A. Romeu, and J. Palau.** 2000. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. Mol. Biol. Evol. **17:**352–361.
11. **Handa, N., Y. Nakayama, M. Sadykov, and I. Kobayashi.** 2001. Experimental genome evolution: large-scale genome rearrangements associated with resistance to replacement of a chromosomal restriction-modification gene complex. Mol. Microbiol. **40:**932–940.
12. **Hughes, D.** 2000. Evaluating genome dynamics: the constraints on rearrangements within bacterial genomes. Genome Biol. **1:**REVIEWS0006.
13. **Jedrzejas, M. J.** 2000. Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. Prog. Biophys. Mol. Biol. **73:**263–287.
14. **Jeltsch, A.** 2003. Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? Gene **317:**13–16.
15. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. **29:**3742–3756.
16. **Krafft, T., and J. M. Macy.** 1998. Purification and characterization of the respiratory arsenate reductase of *Chrysiogenes arsenatis.* Eur. J. Biochem. **255:**647–653.
17. **Lee, S. F., C. W. Forsberg, and A. M. Gibbins.** 1992. Type II DNA restriction-modification system and an endonuclease from the ruminal bacterium *Fibrobacter succinogenes* S85. J. Bacteriol. **174:**5275–5283.
18. **Lin, C., and D. A. Stahl.** 1995. Comparative analyses reveal a highly conserved endoglucanase in the cellulolytic genus *Fibrobacter.* J. Bacteriol. **177:**2543–2549.
19. **Lin, C. Z., B. Flesher, W. C. Capman, R. I. Amann, and D. A. Stahl.** 1994. Taxon specific hybridization probes for fiber-digesting bacteria suggest novel gut-associated *Fibrobacter.* Syst. Appl. Microbiol. **17:**418–424.
20. **Lynd, L. R., P. J. Weimer, W. H. van Zyl, and I. S. Pretorius.** 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. **66:**506–577.
21. **Malburg, L. M., Jr., A. H. Iyo, and C. W. Forsberg.** 1996. A novel family 9 endoglucanase gene (*celD*), whose product cleaves substrates mainly to glucose, and its adjacent upstream homolog (*celE*) from *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. **62:**898–906.
22. **Malburg, S. R., L. M. Malburg, Jr., T. Liu, A. H. Iyo, and C. W. Forsberg.** 1997. Catalytic properties of the cellulose-binding endoglucanase F from *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. **63:**2449–2453.
23. **Matheron, C., A. M. Delort, G. Gaudet, and E. Forano.** 1998. In vivo ^{13}C NMR study of glucose and cellobiose metabolism by four cellulolytic strains of the genus *Fibrobacter.* Biodegradation **9:**451–461.
24. **McGavin, M., and C. W. Forsberg.** 1988. Isolation and characterization of endoglucanases 1 and 2 from *Bacteroides succinogenes* S85. J. Bacteriol. **170:**2914–2922.
25. **McGavin, M. J., C. W. Forsberg, B. Crosby, A. W. Bell, D. Dignard, and D. Y. Thomas.** 1989. Structure of the cel-3 gene from *Fibrobacter succinogenes* S85 and characteristics of the encoded gene product, endoglucanase 3. J. Bacteriol. **171:**5587–5595.
26. **McInerney, J. O.** 1998. GCUA: general codon usage analysis. Bioinformatics **14:**372–373.
27. **Mira, A., L. Klasson, and S. G. Andersson.** 2002. Microbial genome evolution: sources of variability. Curr. Opin. Microbiol. **5:**506–512.
28. **Miron, J., and D. Ben Ghedalia.** 1993. Digestion of cell-wall monosaccharides of ryegrass and alfalfa hays by the ruminal bacteria *Fibrobacter succinogenes* and *Butyrivibrio fibrisolvens.* Can. J. Microbiol. **39:**780–786.
29. **Miron, J., and C. W. Forsberg.** 1999. Characterisation of cellulose-binding proteins that are involved in the adhesion mechanism of *Fibrobacter intestinalis* DR7. Appl. Microbiol. Biotechnol. **51:**491–497.
30. **Montgomery, L., B. Flesher, and D. Stahl.** 1988. Transfer of *Bacteroides succinogenes* (Hungate) to *Fibrobacter* gen. nov. as *Fibrobacter succinogenes* comb. nov. and description of *Fibrobacter intestinalis* sp. nov. Int. J. Syst. Bacteriol. **38:**430–435.
31. **Morrison, M., K. E. Nelson, I. K. O. Cann, C. W. Forsberg, R. I. Mackie, J. B. Russell, B. A. White, K. Amava, B. Cheng, M. Qi, H. Jun, S. Mulligan, K. Tran, H. A. Carty, H. Khouri, W. Nelson, S. Daugherty, and C. M. Fraser.** 2003. The *Fibrobacter succinogenes* strain S85 genome sequencing project. Abstr. 3rd ASM-TIGR Conference of Microbial Genomes, p. 33.
32. **Nesbø, C. L., K. E. Nelson, and W. F. Doolittle.** 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima.* J. Bacteriol. **184:**4475–4488.
33. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299–304.
34. **Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
35. **Philippe, H., and C. J. Douady.** 2003. Horizontal gene transfer and phylogenetics. Curr. Opin. Microbiol. **6:**498–505.
36. **Roberts, R. J., and D. Macelis.** 2001. REBASE–restriction enzymes and methylases. Nucleic Acids Res. **29:**268–269.
37. **Ross, R., X. L. Ross, B. Krueger, T. Lenin, and A. B. Reeks-Kunz.** 1999. No radioactive detection of differentially expressed genes using complex RNA or DNA hybridization probes. BioTechniques **26:**150–155.
38. **Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J. C. Camus, L. Cattolico, M. Chandler, N. Choisne, C. Claudel-Renard, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Siguier, P. Thebault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, and C. A. Boucher.** 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum.* Nature **415:**497–502.
39. **Scott, H. W., and B. A. Dehority.** 1965. Vitamin requirements of several cellulolytic rumen bacteria. J. Bacteriol. **89:**1169–1175.
40. **Swain, R. A., J. V. Nolan, and A. V. Klieve.** 1996. Natural variability and diurnal fluctuations within the bacteriophage population of the rumen. Appl. Environ. Microbiol. **62:**994–997.
41. **Xu, J., M. K. Bjursell, J. Himrod, S. Deng, L. K. Carmichael, H. C. Chiang, L. V. Hooper, and J. I. Gordon.** 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science **299:**2074–2076.
42. **Ziemer, C. J., R. Sharp, M. D. Stern, M. A. Cotta, T. R. Whitehead, and D. A. Stahl.** 2000. Comparison of microbial populations in model and natural rumens using 16S rRNA-targeted probes. Environ. Microbiol. **2:**632–643.