

Complete Genome Sequence of a thermotolerant sporogenic lactic acid bacterium, *Bacillus coagulans* strain 36D1

Mun Su Rhee^{1†}, Brélan E. Moritz^{1†}, Gary Xie^{2,3}, T. Glavina del Rio³, E. Dalin³, H. Tice³, D. Bruce^{2,3}, L. Goodwin^{2,3}, O. Chertkov^{2,3}, T. Brettin⁴, C. Han^{2,3}, C. Detter^{2,3}, S. Pitluck³, Miriam L. Land⁴, Milind Patel¹, Mark Ou¹, Roberta Harbrucker¹, Lonnie O. Ingram¹ and K. T. Shanmugam^{1*}

¹ Department of Microbiology and Cell Science, University of Florida, Gainesville, FL, USA

² Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

³ Joint Genome Institute, Walnut Creek, CA, USA

⁴ Oak Ridge National Laboratory, Oak Ridge, TN, USA

† These two authors contributed equally.

*Corresponding author: shan@ufl.edu

Keywords: *Bacillus coagulans*; genome sequence; lactic acid; fermentation; probiotics; thermotolerant bacterium

Bacillus coagulans is a ubiquitous soil bacterium that grows at 50-55 °C and pH 5.0 and ferments various sugars that constitute plant biomass to L (+)-lactic acid. The ability of this sporogenic lactic acid bacterium to grow at 50-55 °C and pH 5.0 makes this organism an attractive microbial biocatalyst for production of optically pure lactic acid at industrial scale not only from glucose derived from cellulose but also from xylose, a major constituent of hemicellulose. This bacterium is also considered as a potential probiotic. Complete genome sequence of a representative strain, *B. coagulans* strain 36D1, is presented and discussed.

Introduction

In addition to its use in food and cosmetics, lactic acid is increasingly used as a starting material for production of bio-based, renewable plastics [1-3]. Optically pure lactic acid required by the bioplastics industry is currently produced only by bacterial fermentation of sugars [3,4]. The main sugars currently used in such fermentations are glucose derived from corn starch or sucrose from sugar cane, sugar beets, etc. With increasing demand for renewable bio-based plastics, there is a shift away from food-based carbohydrates to non-food carbohydrates such as lignocellulosic biomass for lactic acid production [5,6]. Commercial fungal cellulases play a central role in the conversion of cellulose to glucose before fermentation to lactic acid and these enzymes function optimally at 50°C and pH 5.0 [7-10]. By matching the fungal enzyme activity optimum with that of the growth and fermentation optimum of the microbial biocatalyst, such as *Bacillus coagulans*, the amount of fungal cellulases required for simultaneous saccharification and fermentation (SSF) of cellulose to lactic acid can be reduced by a factor of three or higher compared to fermentation with lactic acid bacteria that grow optimally at

temperatures below 40°C [9]. Since fungal enzymes represent a significant cost component of the overall process of biomass conversion to fuels and chemicals [11], reducing the enzyme loading during SSF of cellulose to lactic acid by *B. coagulans* is expected to lower the overall process cost and help the bioplastics industry compete with petroleum-based non-renewable plastics.

Bacillus coagulans belongs to a group of bacteria classified as sporogenic lactic acid bacteria [12]. These facultative anaerobes ferment pentoses, a component of hemicellulose, to L(+)-lactic acid as the major fermentation product reaching yields of 90% and titers close to 100 g/L in about 48 hours [13,14]. In this regard, *B. coagulans* differs from other lactic acid bacteria, such as *Lactobacillus*, *Lactococcus*, etc., in its ability to ferment pentose sugars to lactic acid through the pentose-phosphate pathway in contrast to the phosphoketolase pathway used by the lactic acid bacteria that yield an equimolar mixture of lactate and acetate [14]. Because of the thermotolerant, acid-tolerant and pentose fermentation characteristics, there is significant commercial interest in developing *B. coagu-*

lans as a microbial biocatalyst for production of optically pure lactic acid as well as other fuels and chemicals. The higher operating temperature of *B. coagulans* is also expected to significantly reduce contamination of industrial fermentations that could lower product quality [15].

B. coagulans has been reported to function as a probiotic in animal trials and there is significant interest in the potential of this bacterium as a probiotic in humans [16]. These studies suggest that *B. coagulans* can readily achieve the GRAS (generally regarded as safe) status required for large scale industrial use. Genetic tools are being developed for manipulating *B. coagulans*, a genetically recalcitrant bacterium [17,18]. In order to fully explore the potential of *B. coagulans* as a microbial biocatalyst for production of fuels and chemicals, the entire genome of *B. coagulans* strain 36D1 was sequenced. Results from these experiments reveal that strain 36D1 has a single circular genome of 3,552,226 base pairs that encode 3,306 protein coding regions. Other characteristics of this bacterium, based on its genome composition, are presented and discussed.

Classification and features

B. coagulans was first isolated from coagulated milk by Hammer in 1915 [19]. Since then, several members of this group have been isolated from various sources [12,14]. *B. coagulans* strain 36D1 used in this study was isolated from a mud sample from an effluent stream of Old Faithful Geyser 1 near Calistoga, California, USA as an organism that can grow on xylose at 50°C and pH 5.0 both aerobically and anaerobically [14]. This bacterium is rod-shaped and produces endospore when cultured in nutrient broth (Fig. 1). Endospores are rarely observed when the bacterium was cultured in L-broth. Optimum temperature and pH for growth of strain 36D1 is 55°C and 5.5, respectively [10]. Corn steep liquor at 0.5% (w/v) provided the needed nutritional supplements for growth in mineral salts medium and the growth rate of the bacterium in that medium at 55°C was 1.67 h⁻¹. The main fermentation product of the bacterium is L-lactate. Pentose fermentation increases the level of acetate, ethanol and formate in the medium compared to hexose fermentation [14]. Anaerobic cultures started with sparging of the medium with N₂ require CO₂ for growth. Other characteristics of the bacterium are listed in Table 1. *B. coagulans* strain 36D1 is deposited in the American Type Culture Collection (PTA-5827).

The *B. coagulans* group is polydisperse [12] and among the *Bacillus* spp., strain 36D1 is phylogenetically close to *B. halodurans* based on 16S rRNA(DNA) sequences (Fig. 2). Although *B. coagulans* is similar to lactic acid bacteria in its ability to grow anaerobically and ferment sugars to lactic acid, it is distinct from the lactic acid bacteria based on 16S rRNA(DNA) sequence similarity.

Genome sequencing and annotation

Genome project history

This genome was selected for sequencing on the basis of the properties described above. The genome sequence is deposited in GenBank (Accession number, CP003056). Sequencing was initiated and completed to a level of four contigs and annotated by the DOE Joint Genome Institute (JGI). The original draft version was deposited in GenBank on February 7, 2007 and the final draft version with four contigs was deposited on Feb. 3, 2010, thereby updating previous releases to the database. Genome sequencing was completed at the University of Florida, annotated by the Oak Ridge National Laboratory, and processed by the Los Alamos National Laboratory and NCBI. A summary of the project information is shown in Table 2.

Growth conditions and DNA isolation

B. coagulans strain 36D1 was cultured in LB + glucose (10 g/L) medium (pH 5.0) at 50°C in a shaker at 200 RPM as described before [10]. Cells were harvested during mid-exponential phase of growth. Cell pellet from a 30 ml culture was resuspended in 2.1 ml of TE buffer (Tris, 10 mM; EDTA, 10 mM; pH 8.0) supplemented with lysozyme (1 mg/ml; Sigma Chemical Co., St. Louis, MO, USA) and RNase (0.1 mg/ml; Sigma Chemical Co.). The sample was incubated at 37°C for 20 minutes to remove the cell wall. Sodium dodecyl sulfate (SDS) was added to the lysed cells to achieve an SDS concentration of 1.4%. After 10 minutes on ice, the lysate was extracted with equal volume of TE-saturated phenol to remove cellular debris. After two more extractions of the aqueous phase with equal volumes of phenol-chloroform mixture (25:24:1 of phenol, chloroform and isoamyl alcohol), and one extraction with an equal volume of chloroform:isoamyl alcohol, the DNA was precipitated with ethanol and dried. The ratio of absorbance at 260 nm and 280 nm of the purified DNA was 1.99 and based on agarose gel electrophoresis and ethidium bromide staining, DNA contained only a trace amount of degraded RNA.

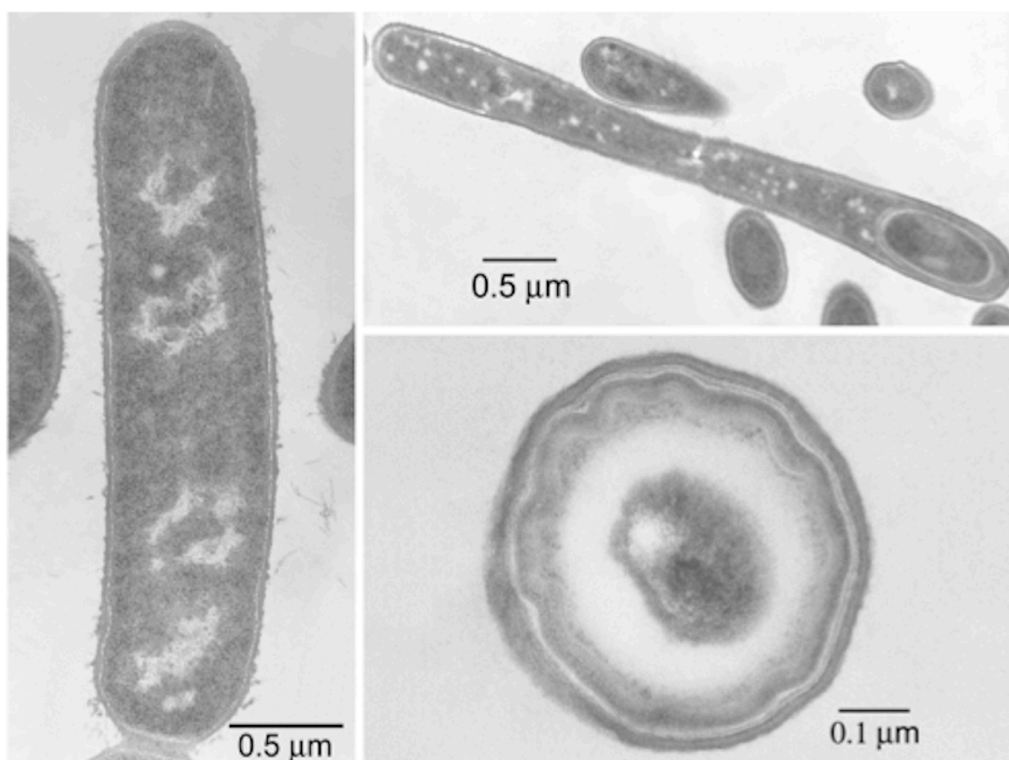


Figure 1. Thin section electron micrographs of *B. coagulans* strain 36D1. Left: vegetative cell. Upper right: sporulating cell. Lower right: mature spore.

Genome sequencing and assembly

The genome was sequenced using a combination of Sanger and 454 sequencing platforms. General aspects of library construction and sequencing can be found at the JGI website [32]. 454 pyrosequencing reads were assembled using the Newbler assembler version 1.1.02.15 (Roche). Large Newbler contigs were broken into 2 kb overlapping fragments (1 kb overlap) and entered into assembly as pseudo-reads. The sequences were assigned quality scores based on Newbler consensus q-scores with modifications to account for overlap redundancy and to adjust inflated q-scores. A hybrid 454/Sanger assembly was made using the Phrap assembler. Possible mis-assemblies were corrected with Dupfinisher or transposon bombing of bridging clones. Editing in Consed, custom primer walk or PCR amplification closed gaps between contigs. A total of 2,471 Sanger finishing reads were produced to close gaps, to resolve repetitive regions, and to raise the quality of the finished sequence. The error rate of the completed genome sequence was less than 1 in 100,000. Together all sequence types provided 9 x coverage of the genome. The final assembly contains a total of 35,357 Sanger and pyrosequence reads. This analysis yielded four

contigs with lengths of 2,712, 65,471, 565,365 and 2,917,758 base pairs for a total of 3,551,306 base pairs.

In order to close the gaps, a restriction map of *B. coagulans* strain 36D1 genome was constructed using BglII restriction enzyme. This optical mapping by OpGen (Gaithersburg, MD) yielded a circular map of approximately 3,521 kbp. Comparing the computed restriction map of the DNA sequence from the four contigs with the restriction map of the whole genome, the lengths of the gaps between the appropriate contigs were predicted. Using the sequence information from the contigs and appropriate restriction fragments, PCR primers were synthesized and the genomic DNA was sequenced using Sanger method by the Interdisciplinary Center for Biotechnology Research at the University of Florida. As needed, PCR primers were synthesized based on new sequence information for genome walking to fill-in the gaps and complete the genome sequence. Based on these analyses, the genome of *B. coagulans* strain 36D1 was determined to be circular with a length of 3,552,226 base pairs.

Table 1. Classification and general features of *Bacillus coagulans* strain 36D1

MIGS ID	Property	Term	Evidence code
MIGS-2	Classification	Domain <i>Bacteria</i>	TAS [20]
		Phylum <i>Firmicutes</i>	TAS [21-23]
		Class <i>Bacilli</i>	TAS [24,25]
		Order <i>Bacillales</i>	TAS [26,27]
		Family <i>Bacillaceae</i>	TAS [26,28]
		Genus <i>Bacillus</i>	TAS [26,29,30]
MIGS-3		Species <i>Bacillus coagulans</i>	TAS [12,19,26]
		Strain 36D1	
	Gram stain	Positive	IDA
	Cell shape	Rods	IDA
	Motility	Yes	IDA
	Sporulation	Yes	IDA
	Temperature range	Thermotolerant; 30 – 60 °C	TAS [10]
	Optimum temperature	55 °C	TAS [10]
MIGS-22	Oxygen requirement	Facultative anaerobe	TAS [14]
	Carbon source	Glucose, xylose, arabinose, galactose, maltose, fructose, cellobiose, yeast extract, peptone	IDA
	Energy source	Organic carbon compounds	IDA
	Terminal electron acceptor	O ₂ , acetate, pyruvate	TAS [14]
MIGS-6	Habitat	Soil	TAS [14]
MIGS-15	Biotic relationship	Free-living	TAS [14]
MIGS-14	Pathogenicity	None; potential probiotic	NAS
	Biosafety level	1	NAS
	Isolation	Soil	TAS [14]
MIGS-4	Geographic location	Near Calistoga, California, USA	TAS [14]
MIGS-5	Sample collection time	2000	IDA
MIGS-4.1	Latitude	38.58 N	NAS
MIGS-4.2	Longitude	-122.58	NAS
MIGS-4.3	Depth	Surface	TAS [14]
MIGS-4.4	Altitude	351'	NAS

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (i. e. a direct report exists in the literature); NAS: Non-traceable Author Statement (i. e. not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [31]. If the evidence code is IDA, then the property was directly observed for a living isolate by one of the authors or an expert mentioned in the acknowledgements.

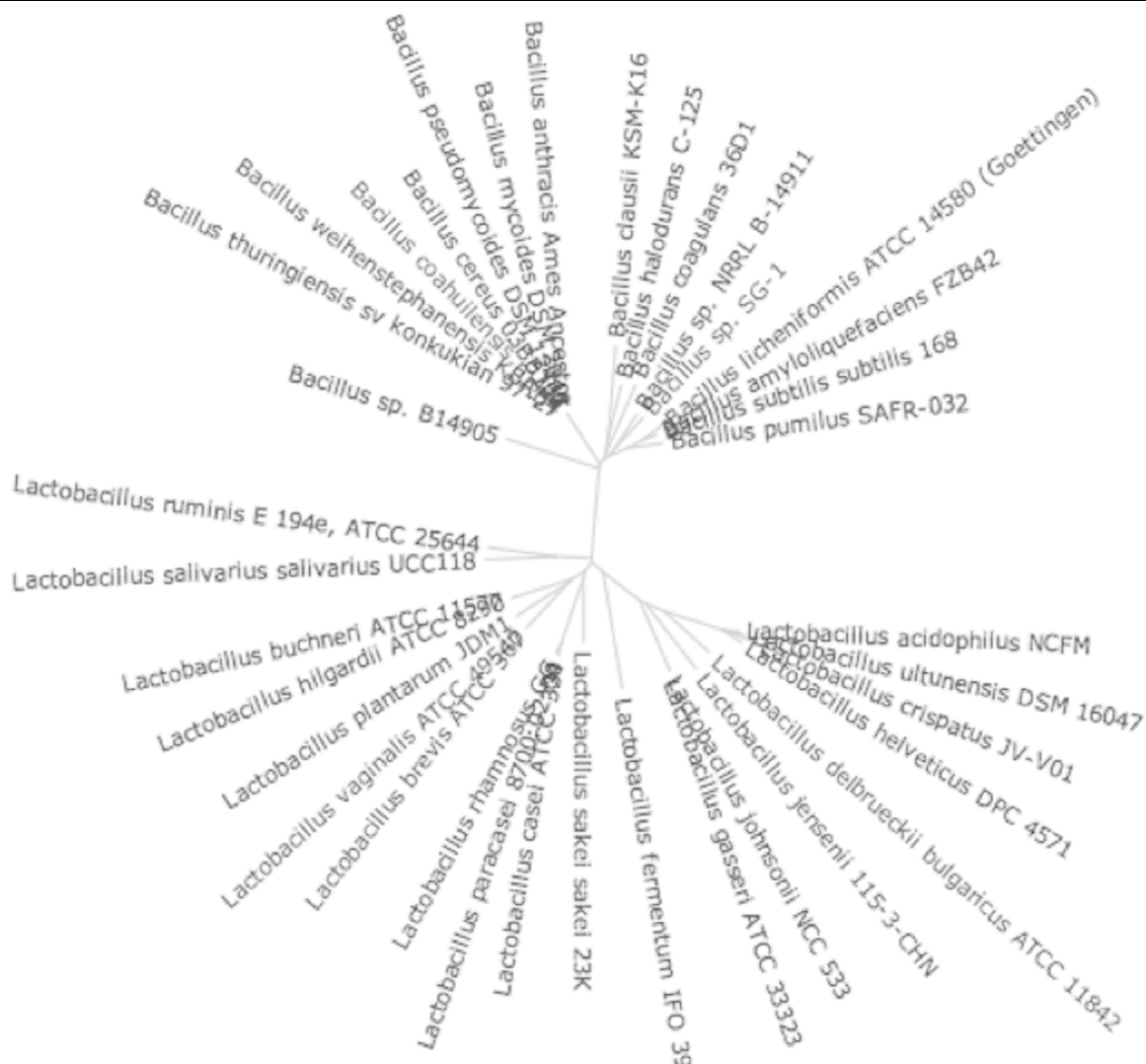


Figure 2. Unrooted phylogenetic tree based on 16S rRNA(DNA) of *B. coagulans* strain 36D1 and *Bacillus* spp. and *Lactobacillus* spp.

Genome annotation

Genes were identified using Prodigal [33] as part of the Oak Ridge National Laboratory genome annotation pipeline. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [34], RNAMmer [35], Rfam [36], TMHMM [37], and signalP [38].

Genome properties

The genome consists of a 3,552,226 bp long chromosome with a 46.5% GC content (Figure 3 and

Table 3). Of the 3,420 genes predicted, 3,306 were protein coding genes, and 114 encode RNAs. Among the 114 RNA genes, 10 each coded for 5S, 16S and 23S rRNAs and 84 can be accounted for tRNAs.

The majority of the protein-coding genes (74%) were assigned with a putative function while those remaining were annotated as hypothetical proteins. About 49 ORFs were identified as potential transposases. The distribution of genes into COGs functional categories is presented in Table 4. The first about 40% of the genome is predominantly transcribed from the lagging strand (as written) while the other 60% is transcribed from the leading strand (Fig. 4).

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Genomic libraries: Sanger (one each of 2.5, 8 and 37 kbp library); one 454CT library (2 kb); PCR; primer walk
MIGS-29	Sequencing platforms	Sanger; 454
MIGS-31.2	Sequencing coverage	9x
MIGS-30	Assemblers	Phrap, Newbler
MIGS-32	Gene calling method	Prodigal
	Genome Database	GenBank
	INSDC ID	CP003056
	Genbank Date of Release	September 26, 2011
	GOLD ID	GC01988
	NCBI project ID	15679
MIGS-13	Source material identifier	ATCC PTA-5827
	Project relevance	Biotechnological

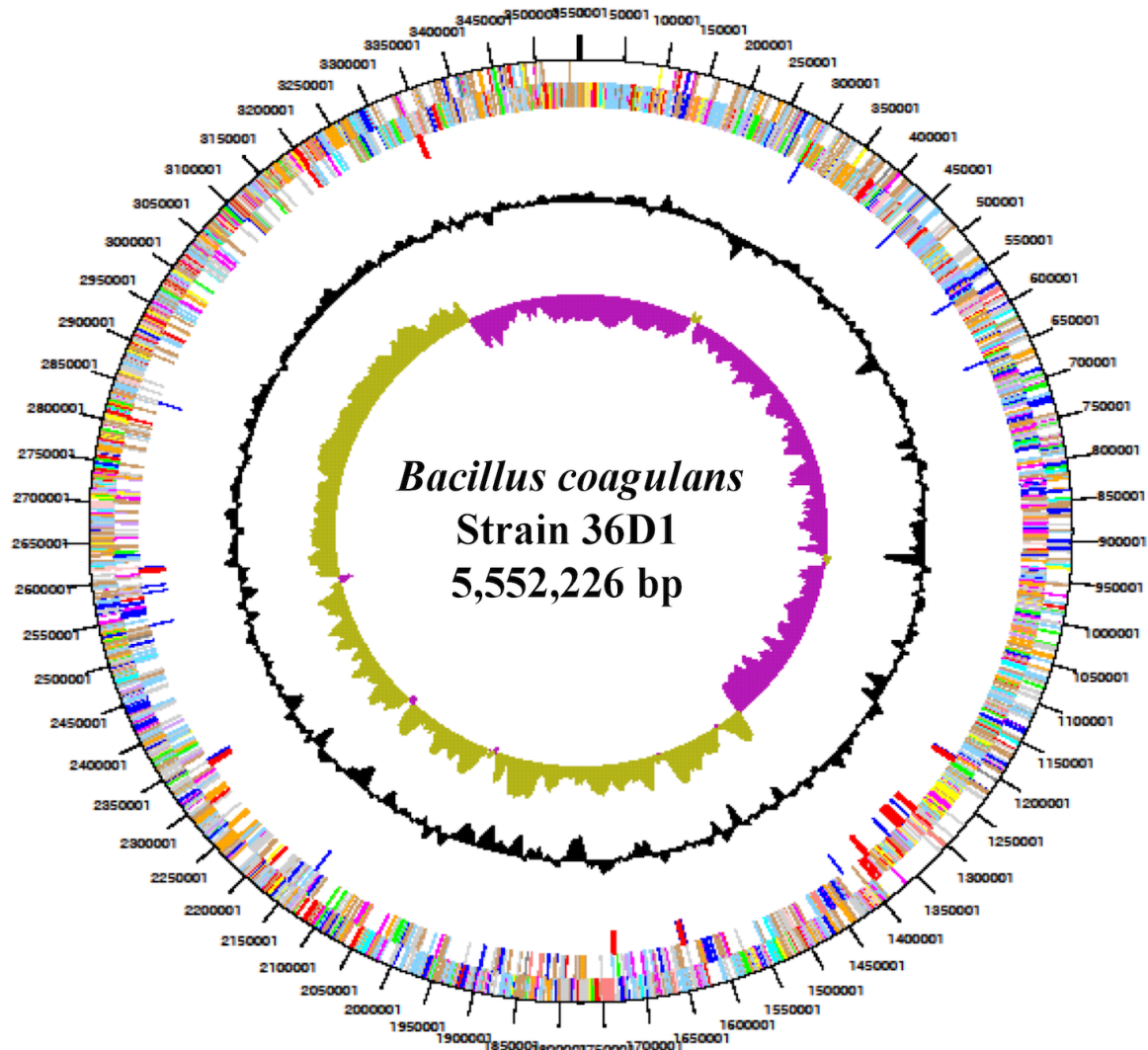


Fig. 3. Graphical circular map of the genome of *B. coagulans* strain 36D1. From outside to center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), pseudogenes, % G+C, GC skew.

Table 3. Genome statistics

Attribute	Value	% of total
Genome Size (bp)	3,552,226	100
DNA G+C content (bp)	1,651,327	46.5
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	3420	100
RNA genes	114	3.3
Protein-coding genes	3306	96.7
Genes in paralog clusters	541	15.6
Genes assigned to COGs	2456	71.0
Genes with signal peptides	858	24.8
Genes with transmembrane helices	863	24.9
Paralogous groups	199	5.8

Insights from genome sequence

Comparison of the predicted proteome of *B. coagulans* with that of a group of *Bacillus* spp genomes identified 491 unique proteins in *B. coagulans* that are not identified in other members of *Bacillus* spp. 404 of these genes are in the early part of the 36D1 genome as listed. This list includes 31 genes encoding putative transposases. Function of many of these gene products is not known. However, 413 of these unique proteins are found to be shared with *Lactobacillus* spp. In reverse, comparison of the *B. coagulans* genome with the genomes from a group of *Lactobacillus* spp. revealed that 423 ORFs are unique to *B. coagulans* and of these, 345 ORFs, mostly related to sporulation, are shared with *Bacillus* spp. Combining these two sets, a set of 78 ORFs coding for proteins with unknown function are unique to *B. coagulans* that are not present in either *Bacillus* or *Lactobacillus*. Based on principal components analysis, *B. coagulans* strain 36D1 groups with *Bacillus* but as an outlier and away from Lactic acid bacteria (Fig. 4). Although *B. coagulans* produced L-lactic acid as the fermentation product at an optical purity reaching close to 100%, the genome contains a gene encoding D-LDH.

Although some members of *B. coagulans* group are cellulolytic and xylanolytic, strain 36D1 is phenotypically unable to utilize cellulose and xy-

lan. However, genes encoding glycan hydrolases such as xylanase, xylosidase and α -amylase can be identified in the genome sequence. Presence of these genes suggest that the bacterium can be evolved to produce xylanase to reduce the severity of acid treatment during hydrolysis of hemicellulose from lignocellulosic biomass for production of optically pure lactic acid. *B. coagulans* strain 36D1 is an auxotroph for several amino acids and vitamins. Based on analysis of the genome sequence by Patric Comparative pathway tool [39], only histidine biosynthetic pathway appears to be incomplete among the amino acid biosynthesis pathways. Among the vitamins, the pathways for biosynthesis of biotin, pantothenic acid, nicotinamide and pyridoxine appear to be incomplete.

During the time of preparation of this manuscript, genome sequence for *B. coagulans* strain 2-6 was published [40]. The genome of this strain is 3,073,079 and is 479,147 bp smaller than the genome of strain 36D1. These two *B. coagulans* genomes share about 90% or higher nucleotide sequence identity in the regions that are present in both genomes. Additional comparative analysis of the two genomes is in progress.

Table 4. Number of genes associated with the general COG functional groups

Code	Value	%age	Description
J	172	5.14	Translation
A	0	0.00	RNA processing and modification
K	238	7.11	Transcription
L	144	4.30	Replication, recombination and repair
B	2	0.05	Chromatin structure and dynamics
D	106	3.16	Cell cycle control, mitosis and meiosis
Y	0	0.00	Nuclear structure
V	113	3.37	Defense mechanisms
T	162	4.84	Signal transduction mechanisms
M	240	7.17	Cell wall/membrane biogenesis
N	78	2.33	Cell motility
Z	0	0.00	Cytoskeleton
W	0	0.00	Extracellular structures
U	50	1.49	Intracellular trafficking and secretion
O	182	5.44	Posttranslational modification, protein turnover, chaperones
C	277	8.28	Energy production and conversion
G	310	9.26	Carbohydrate transport and metabolism
E	379	11.33	Amino acid transport and metabolism
F	100	2.98	Nucleotide transport and metabolism
H	217	6.48	Coenzyme transport and metabolism
I	104	3.10	Lipid transport and metabolism
P	214	6.39	Inorganic ion transport and metabolism
Q	168	5.02	Secondary metabolites biosynthesis, transport and catabolism
R	448	13.39	General function prediction only
S	242	7.23	Function unknown
-			Not in COGs (Total – 3345)

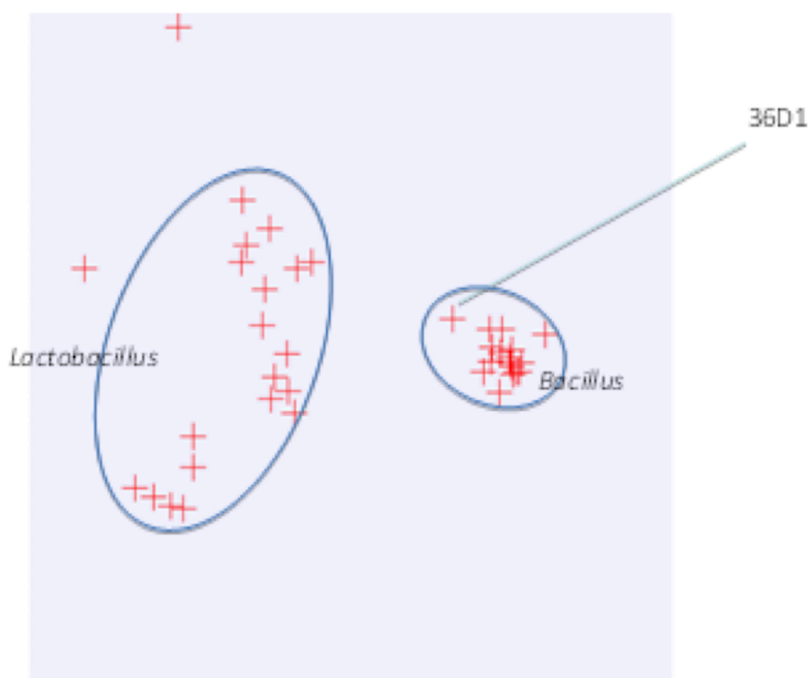


Figure 4. Principal component analysis based on COG profiles

Acknowledgements

This study was supported in part by a grant from the Department of Energy (DE-FG36-04G014019 and DE-FG36-08G088142), US Department of Agriculture, National Institute of Food and Agriculture (2011-10006-30358), the State of Florida, University of Florida Agri-

cultural Experiment Station and Florida Energy Systems Consortium. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

- Datta R, Henry M. Lactic acid: recent advances in products, processes and technologies - a review. *J Chem Technol Biotechnol* 2006; **81**:1119-1129. [doi:10.1002/jctb.1486](https://doi.org/10.1002/jctb.1486)
- Madhavan NK, Nair NR, John RP. An overview of the recent developments in polylactide (PLA) research. *Bioresour Technol* 2010; **101**:8493-8501. [PubMed doi:10.1016/j.biortech.2010.05.092](https://pubmed.ncbi.nlm.nih.gov/doi/10.1016/j.biortech.2010.05.092)
- Wee Y, Kim J, Ryu H. Biotechnological production of lactic acid and its recent applications. *Food Technol Biotechnol* 2006; **44**:163-172.
- Hofvendahl K, Hans-Hagerdal B. Factors affecting the fermentative lactic acid production from renewable resources. *Enzyme Microb Technol* 2000; **26**:87-107. [PubMed doi:10.1016/S0141-0229\(99\)00155-6](https://pubmed.ncbi.nlm.nih.gov/doi/10.1016/S0141-0229(99)00155-6)
- Carole TM, Pellegrino J, Paster MD. Opportunities in the industrial biobased products industry. *Appl Biochem Biotechnol* 2004; **115**:871-885. [PubMed doi:10.1385/ABAB:115:1-3:0871](https://pubmed.ncbi.nlm.nih.gov/doi/10.1385/ABAB:115:1-3:0871)
- Mooney BP. The second green revolution? Production of plant-based biodegradable plastics. *Biochem J* 2009; **418**:219-232. [PubMed doi:10.1042/BJ20081769](https://pubmed.ncbi.nlm.nih.gov/doi/10.1042/BJ20081769)
- Abe S, Takagi M. Simultaneous saccharification and fermentation of cellulose to lactic acid. *Biotechnol Bioeng* 1991; **37**:93-96. [PubMed doi:10.1002/bit.260370113](https://pubmed.ncbi.nlm.nih.gov/doi/10.1002/bit.260370113)
- Iyer PV, Lee YY. Product inhibition in simultaneous saccharification and fermentation of cellulose into lactic acid. *Biotechnol Lett* 1999; **21**:371-373. [doi:10.1023/A:1005435120978](https://doi.org/10.1023/A:1005435120978)
- Ou MS, Mohammed N, Ingram LO, Shanmugam KT. Thermophilic *Bacillus coagulans* requires less cellulases for simultaneous saccharification and fermentation of cellulose to products than mesophilic microbial biocatalysts. *Appl Biochem Biotechnol* 2009; **155**:379-385. [PubMed doi:10.1007/s12010-008-8509-4](https://pubmed.ncbi.nlm.nih.gov/doi/10.1007/s12010-008-8509-4)
- Patel MA, Ou M, Ingram LO, Shanmugam KT. Simultaneous saccharification and co-fermentation of crystalline cellulose and sugar cane bagasse hemicellulose hydrolysate to lactate by a thermotolerant acidophilic *Bacillus* sp. *Biotechnol Prog* 2005; **21**:1453-1460. [PubMed doi:10.1021/bp0400339](https://pubmed.ncbi.nlm.nih.gov/doi/10.1021/bp0400339)
- Leber J. Economics improve for first commercial cellulosic ethanol plants. *New York Times* 2010:Feb. 16, 2010.
- De Clerck E, Rodriguez-Diaz M, Forsyth G, Lebbe L, Logan NA, DeVos P. Polyphasic characterization of *Bacillus coagulans* strains, illustrating heterogeneity within this species, and emended description of the species. *Syst Appl Microbiol* 2004; **27**:50-60. [PubMed doi:10.1078/0723-2020-00250](https://pubmed.ncbi.nlm.nih.gov/doi/10.1078/0723-2020-00250)
- Ou MS, Ingram LO, Shanmugam KT. L: (+)-Lactic acid production from non-food carbohydrates by thermotolerant *Bacillus coagulans*. *J Ind Microbiol Biotechnol* 2011; **38**:599-605. [PubMed doi:10.1007/s10295-010-0796-4](https://pubmed.ncbi.nlm.nih.gov/doi/10.1007/s10295-010-0796-4)
- Patel MA, Ou MS, Harbrucker R, Aldrich HC, Buszko ML, Ingram LO, Shanmugam KT. Isolation and characterization of acid-tolerant, thermophilic bacteria for effective fermentation of biomass-derived sugars to lactic acid. *Appl Environ Microbiol* 2006; **72**:3228-3235. [PubMed doi:10.1128/AEM.72.5.3228-3235.2006](https://pubmed.ncbi.nlm.nih.gov/doi/10.1128/AEM.72.5.3228-3235.2006)
- Abdel-Banat BMA, Hoshida H, Ano A, Nonklang S, Akada R. High-temperature fermentation: how can process for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? *Appl Microbiol Biotechnol* 2010; **85**:861-867. [PubMed doi:10.1007/s00253-009-2248-5](https://pubmed.ncbi.nlm.nih.gov/doi/10.1007/s00253-009-2248-5)
- Drago L, De Vecchi E. Should *Lactobacillus sporogenes* and *Bacillus coagulans* have a future? *J Chemother* 2009; **21**:371-377. [PubMed doi:10.1007/s12010-008-8509-4](https://pubmed.ncbi.nlm.nih.gov/doi/10.1007/s12010-008-8509-4)
- Kovács AT, van Hartskamp M, Kuipers OP, van Kranenburg R. Genetic tool development for a new host for biotechnology, the thermotolerant bacterium *Bacillus coagulans*. *Appl Environ Microbiol* 2010; **76**:4085-4088. [PubMed doi:10.1128/AEM.03060-09](https://pubmed.ncbi.nlm.nih.gov/doi/10.1128/AEM.03060-09)
- Rhee MS, Kim JW, Qian Y, Ingram LO, Shanmugam KT. Development of plasmid vector and electroporation condition for gene transfer in spo-

- rogenic lactic acid bacterium, *Bacillus coagulans*. *Plasmid* 2007; **58**:13-22.
[PubMed doi:10.1016/j.plasmid.2006.11.006](#)
19. Hammer BW. Bacteriological studies on the coagulation of evaporated milk. *Iowa Agric. Exp. Station Res. Bull.* 1915; **19**:119-131.
 20. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579.
[PubMed doi:10.1073/pnas.87.12.4576](#)
 21. Gibbons NE, Murray RGE. Proposals Concerning the Higher Taxa of Bacteria. *Int J Syst Bacteriol* 1978; **28**:1-6. [doi:10.1099/00207713-28-1-1](#)
 22. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.
 23. Murray RGE. The Higher Taxa, or, a Place for Everything...? In: Holt JG (ed), *Bergey's Manual of Systematic Bacteriology*, First Edition, Volume 1, The Williams and Wilkins Co., Baltimore, 1984, p. 31-34.
 24. List Editor. List of new names and new combinations previously effectively, but not validly, published. List no. 132. *Int J Syst Evol Microbiol* 2010; **60**:469-472. [doi:10.1099/ijvs.0.022855-0](#)
 25. Ludwig W, Schleifer KH, Whitman WB. Class I. Bacilli class nov. In: De Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 3, Springer-Verlag, New York, 2009, p. 19-20.
 26. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. *Int J Syst Bacteriol* 1980; **30**:225-420. [doi:10.1099/00207713-30-1-225](#)
 27. Prévot AR. In: Hauderoy P, Ehringer G, Guillot G, Magrou J., Prévot AR, Rosset D, Urbain A (eds), *Dictionnaire des Bactéries Pathogènes*, Second Edition, Masson et Cie, Paris, 1953, p. 1-692.
 28. Fischer A. Untersuchungen über bakterien. *Jahrbücher für Wissenschaftliche Botanik* 1895; **27**:1-163.
 29. Cohn F. Untersuchungen über Bakterien. *Beitr Biol Pflanz* 1872; **1**:127-224.
 30. Gibson T, Gordon RE. Genus I. Bacillus Cohn 1872, 174; Nom. gen. cons. Nomencl. Comm. Intern. Soc. Microbiol. 1937, 28; Opin. A. Jud. Comm. 1955, 39. In: Buchanan RE, Gibbons NE (eds), *Bergey's Manual of Determinative Bacteriology*, Eighth Edition, The Williams and Wilkins Co., Baltimore, 1974, p. 529-550.
 31. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29.
[PubMed doi:10.1038/75556](#)
 32. The DOE Joint Genome tute. <http://www.jgi.doe.gov>.
 33. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119.
[PubMed doi:10.1186/1471-2105-11-119](#)
 34. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. [PubMed doi:10.1093/nar/25.5.955](#)
 35. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007; **35**:3100-3108.
[PubMed doi:10.1093/nar/gkm160](#)
 36. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. *Nucleic Acids Res* 2003; **31**:439-441.
[PubMed doi:10.1093/nar/gkg006](#)
 37. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**:567-580.
[PubMed doi:10.1006/jmbi.2000.4315](#)
 38. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795.
[PubMed doi:10.1016/j.jmb.2004.05.028](#)
 39. Patric Comparative pathway tool. <http://www.patricbrc.org/portal/portal/patric/PathwayFinder?cType=taxon&cId=&dm=>
 40. Su F, Yu B, Sun J, Ou HY, Zhao B, Wang L, Qin J, Tang H, Tao F, Jarek M, et al. Genome sequence of the thermophilic strain *Bacillus coagulans* 2-6, an efficient producer of high-optical-purity L-lactic acid. *J Bacteriol* 2011; **193**:4563-4564.
[PubMed doi:10.1128/JB.05378-11](#)