

## Complete genome sequence of *Cellulomonas flavigena* type strain (134<sup>T</sup>)

Birte Abt<sup>1</sup>, Brian Foster<sup>2</sup>, Alla Lapidus<sup>2</sup>, Alicia Clum<sup>2</sup>, Hui Sun<sup>2</sup>, Rüdiger Pukall<sup>1</sup>, Susan Lucas<sup>2</sup>, Tijana Glavina Del Rio<sup>2</sup>, Matt Nolan<sup>2</sup>, Hope Tice<sup>2</sup>, Jan-Fang Cheng<sup>2</sup>, Sam Pitluck<sup>2</sup>, Konstantinos Liolios<sup>2</sup>, Natalia Ivanova<sup>2</sup>, Konstantinos Mavromatis<sup>2</sup>, Galina Ovchinnikova<sup>2</sup>, Amrita Pati<sup>2</sup>, Lynne Goodwin<sup>2,3</sup>, Amy Chen<sup>4</sup>, Krishna Palaniappan<sup>4</sup>, Miriam Land<sup>2,5</sup>, Loren Hauser<sup>2,5</sup>, Yun-Juan Chang<sup>2,5</sup>, Cynthia D. Jeffries<sup>2,5</sup>, Manfred Rohde<sup>6</sup>, Markus Göker<sup>1</sup>, Tanja Woyke<sup>2</sup>, James Bristow<sup>2</sup>, Jonathan A. Eisen<sup>2,7</sup>, Victor Markowitz<sup>4</sup>, Philip Hugenholtz<sup>2</sup>, Nikos C. Kyrpides<sup>2</sup>, and Hans-Peter Klenk<sup>1\*</sup>

<sup>1</sup> DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany

<sup>2</sup> DOE Joint Genome Institute, Walnut Creek, California, USA

<sup>3</sup> Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

<sup>4</sup> Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA

<sup>5</sup> Lawrence Livermore National Laboratory, Livermore, California, USA

<sup>6</sup> HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany

<sup>7</sup> University of California Davis Genome Center, Davis, California, USA

\*Corresponding author: Hans-Peter Klenk

**Keywords:** non-motile, non-sporulating, aerobic, mesophile, Gram-positive, cellulolytic, xylan degrader, *Cellulomonadaceae*, GEBA

---

*Cellulomonas flavigena* (Kellerman and McBeth 1912) Bergey *et al.* 1923 is the type species of the genus *Cellulomonas* of the actinobacterial family *Cellulomonadaceae*. Members of the genus *Cellulomonas* are of special interest for their ability to degrade cellulose and hemicellulose, particularly with regard to the use of biomass as an alternative energy source. Here we describe the features of this organism, together with the complete genome sequence, and annotation. This is the first complete genome sequence of a member of the genus *Cellulomonas*, and next to the human pathogen *Tropheryma whippelii* the second complete genome sequence within the family *Cellulomonadaceae*. The 4,123,179 bp long single replicon genome with its 3,735 protein-coding and 53 RNA genes is part of the *Genomic Encyclopedia of Bacteria and Archaea* project.

---

### Introduction

Strain 134<sup>T</sup> (DSM 20109 = ATCC 482 = JCM 1489) is the type strain of the species *Cellulomonas flavigena* and was isolated from soil and first described in 1912 by Kellerman and McBeth [1], followed by a description in the first edition of Bergey's Manual in 1923 [2].

Because of the absence of a definite proof linking the deposited strains to the original description Stackebrandt and Kandler proposed in 1979 *C. flavigena* and six other *Cellulomonas* strains as neotype strains of their respective species [3]. Here *C. flavigena* cells are reported as Gram-

positive, non-motile and coryneform with snapping divisions [3].

In addition to the type species *C. flavigena*, the five *Cellulomonas* species, *C. biazotea*, *C. cellasea*, *C. gelida*, *C. fimi* and *C. uda* have been members of the genus since their original description in the first edition of Bergey's Manual in 1923 [2]. Because of the phenetic resemblance of the different species to each other *C. flavigena* was recognized as the only species in the genus *Cellulomonas* in the eighth edition of Bergey's Manual. This reduction to a single species was questioned by Braden and Thayer based on serological studies in 1976

[4] and by Stackebrandt and Kandler based on DNA reassociation studies in 1979 [3]. In 1980 the Approved Lists of Bacterial Names already listed six species: *C. flavigena*, *C. biazotea*, *C. gelida*, *C. uda*, *C. fimi* and *C. cellasea* [5]. Currently, 17 species belonging to the genus *Cellulomonas* are noted in the actual version of the List of Prokaryotic names with Standing in Nomenclature [6]. Due to the cellulolytic activity of these organisms, their preferred habitats are cellulose enriched environments such as soil, bark, wood, and sugar fields, but they were also successfully isolated from rumen and from activated sludge. Here we present a summary classification and a set of features for *C. flavigena* 134<sup>T</sup>, together with the description of the complete genomic sequencing and annotation.

### Classification and features

The 16S rRNA genes of the 16 other type strains in the genus *Cellulomonas* share between 92.2% (*C. bogoriensis* [7]) and 98.1% (*C. persica* [8]) sequence identity with strain 134<sup>T</sup>, whereas the other type strains from the family *Cellulomonadaceae*, which belong to the genera *Actinotalea*, *Oerskovia*, *Paraoerskovia* and *Tropheryma*, share less than

95.6% sequence identity [9]. Cultivated strains with highest sequence similarity include a so far unpublished strain 794 (Y09565) from human clinical specimen (99.7% sequence identity) and Everest-gws-44 (EU584517) from glacial meltwater at 6,350 m height on Mount Everest (98.1% sequence identity). The only reported uncultured clone with high sequence similarity (98.5%) originated from a diet-related composition of the gut microbiota of the earthworm *Lumbricus rubellus* [10]. Metagenomic surveys and environmental samples based on 16S rRNA gene sequences delivered no indication for organisms with sequence similarity values above 93-94% to *C. flavigena*, indicating that members of this species are not abundant in the so far screened habitats. The majority of these 16S rRNA gene sequences with similarity between 88% and 93% originate from marine metagenomes (status June 2010).

Figure 1 shows the phylogenetic neighborhood of *C. flavigena* 134<sup>T</sup> in a 16S rRNA based tree. The sequences of the two 16S rRNA gene copies in the genome differ by two nucleotides from each other and by up to four nucleotides from the previously published sequence generated from NCIMB 8073 (Z79463).



**Figure 1.** Phylogenetic tree highlighting the position of *C. flavigena* 134<sup>T</sup> relative to the other type strains within the family *Cellulomonadaceae*. The tree was inferred from 1,393 aligned characters [11,12] of the 16S rRNA gene sequence under the maximum likelihood criterion [13] and rooted with the type strain of the suborder *Micrococccineae*. The branches are scaled in terms of the expected number of substitutions per site. Numbers above branches are support values from 1,000 bootstrap replicates [14] if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [15] are shown in blue, published genomes in bold.

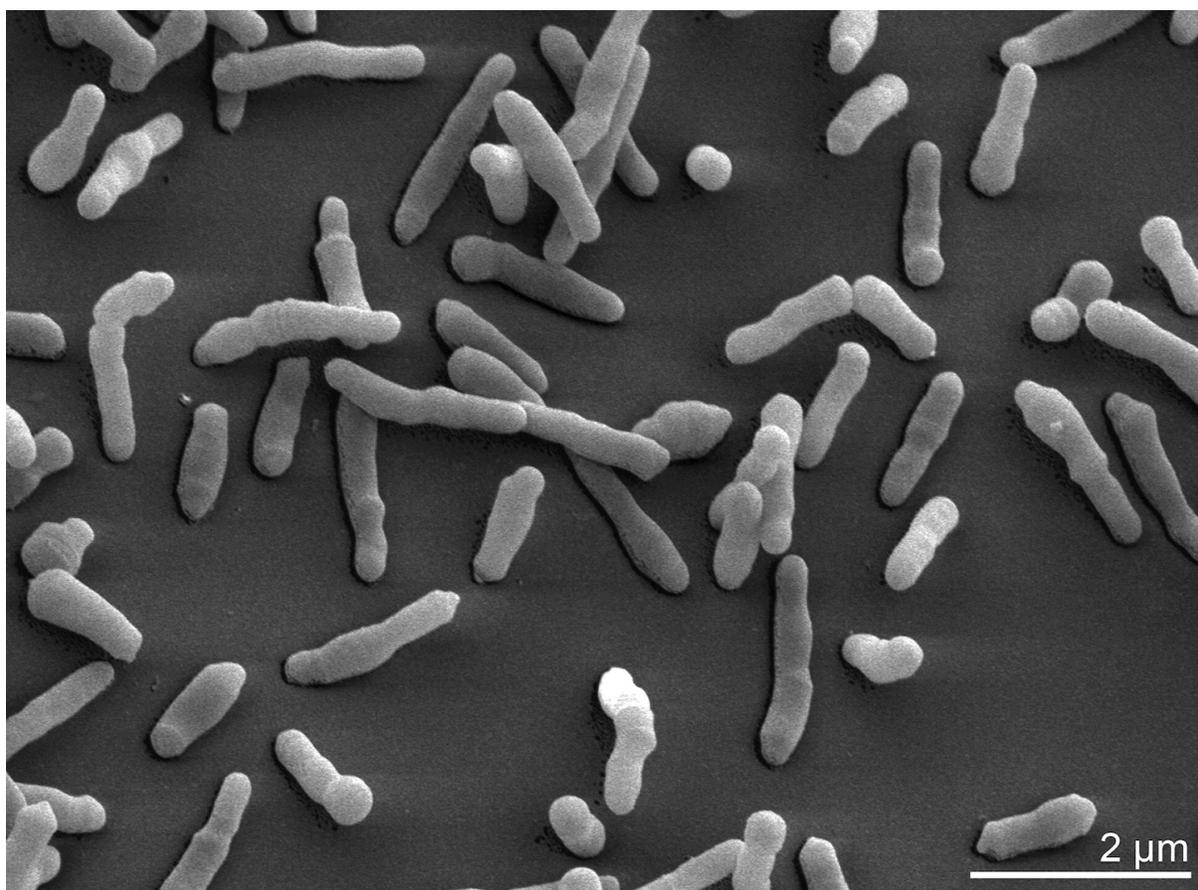
Cells of *C. flavigena* stain Gram-positive with a very fast rate of decolorization [3]. Cells in young broth cultures are typically coryneform with a snapping division (Table 1). In week old cultures a transformation to short rods can occur (Figure 2) [3]. On yeast extract-glucose agar *C. flavigena* forms smooth, glistening, yellow colonies about 5 mm in diameter. *C. flavigena* is described as non-motile [3,28], but according to Thayer *et al.* (1984) *C. flavigena* cells possess polar multitrichous flagella [31] (not visible in Figure 2). *C. flavigena* grows under aerobic conditions with an optimal growth temperature of 30°C [2] and an optimal pH of 7 [32].

Strain 134<sup>T</sup> is able to ferment glucose, maltose, sucrose, xylose and dextrin, but no fermentation of mannitol was observed [3]. While ribose, acetate and gluconate are utilized, there is no utilization of raffinose and L(+)-lactate [3]. It was shown by Kim *et al.* (1987) that gluconate is catabolized via the Entner-Doudoroff pathway and hexose

monophosphate shunt [33]. *C. flavigena* produces catalase but no urease [3]. Esculin and gelatin are hydrolyzed and nitrate is not reduced to nitrite [3].

### Chemotaxonomy

The peptidoglycan of *C. flavigena* contains as the diagnostic amino acid in position 3 of the peptide subunit ornithine with the interpeptide bridge containing D-aspartic acid. The major cell wall sugar is rhamnose, whereas mannose and ribose occur in minor amounts [34]. The major components of the fatty acid profile of *C. flavigena* are 12-methyltetradecanoic (*ai*-C<sub>15:0</sub>) and hexadecanoic (C<sub>16:0</sub>) acids; *i*-C<sub>15:0</sub>, *ai*-C<sub>17:0</sub>, C<sub>14:0</sub> and C<sub>15:0</sub> occur in lower amounts [35]. Menaquinone MK-9(H<sub>4</sub>) is the predominant isoprenoid quinone; minor amounts of MK-9(H<sub>2</sub>), MK-8(H<sub>4</sub>) and MK-7(H<sub>4</sub>) were detected [36]. The polar lipids consist of diphosphatidylglycerol, phosphatidylinositol and two so far unidentified phosphoglycolipids [37].



**Figure 2.** Scanning electron micrograph of *C. flavigena* 134<sup>T</sup>.

**Table 1.** Classification and general features of *C. flavigena* 134<sup>T</sup> according to the MIGS recommendations [16].

MIGS ID	Property	Term	Evidence code
		Domain <i>Bacteria</i>	TAS [17]
		Phylum <i>Actinobacteria</i>	TAS [18]
		Class <i>Actinobacteria</i>	TAS [19]
	Current classification	Order <i>Actinomycetales</i>	TAS [5,19-21]
		Family <i>Cellulomonadaceae</i>	TAS [19,21-25]
		Genus <i>Cellulomonas</i>	TAS [5,26,27]
		Species <i>Cellulomonas flavigena</i>	TAS [1,5,27]
		Type strain 134	
	Gram stain	positive	TAS [3]
	Cell shape	coryneform with snapping division	TAS [3]
	Motility	non-motile	TAS [3,28]
	Sporulation	non-sporulating	TAS [3]
	Temperature range	mesophile	TAS [2]
	Optimum temperature	30°C	TAS [2]
	Salinity	not reported	
MIGS-22	Oxygen requirement	aerobic	TAS [2]
	Carbon source	fermentation of glucose, maltose, sucrose, xylose and dextrin	TAS [3]
	Energy source	chemoorganotrophic	TAS [3]
MIGS-6	Habitat	soil	TAS [2]
MIGS-15	Biotic relationship	free living	NAS
MIGS-14	Pathogenicity	non pathogenic	NAS
	Biosafety level	1	TAS [29]
	Isolation	from soil	TAS [2]
MIGS-4	Geographic location	not reported	
MIGS-5	Sample collection time	in 1912 or before	NAS
MIGS-4.1	Latitude	not reported	
MIGS-4.2	Longitude	not reported	
MIGS-4.3	Depth	not reported	
MIGS-4.4	Altitude	not reported	

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 of the Gene Ontology project [30]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

## Genome sequencing and annotation

### Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [38], and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project [39]. The genome project is deposited in the Genome OnLine Database [15] and the

complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A summary of the project information is shown in Table 2.

**Table 2.** Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Two Sanger libraries - 8 kb pMCL200 and fosmids, one 454 pyrosequence standard library and one Solexa library
MIGS-29	Sequencing platforms	ABI3730, 454 Titanium, Illumina GAii
MIGS-31.2	Sequencing coverage	9.1× Sanger; 56.28× pyrosequence
MIGS-30	Assemblers	Newbler version 1.1.02.15, PGA
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	CP001964
	Genbank Date of Release	May 13, 2010
	GOLD ID	Gc01326
	NCBI project ID	19707
	Database: IMG-GEBA	2502422318
MIGS-13	Source material identifier	DSM 20109
	Project relevance	Tree of Life, GEBA

## Growth conditions and DNA isolation

*C. flavigena* 134<sup>T</sup>, DSM 20109, was grown in DSMZ medium 92 (Trypticase-Soy-Yeast Extract Medium) [40] at 30°C. DNA was isolated from 0.5-1 g of cell paste using Qiagen Genomic 500 DNA Kit (Qiagen, Hilden, Germany) following the standard protocol as recommended by the manufacturer.

## Genome sequencing and assembly

The genome was sequenced using a combination of Sanger and 454 sequencing platforms. All general aspects of library construction and sequencing can be found at the [JGI website](#). Pyrosequencing reads were assembled using the Newbler assembler version 1.1.02.15 (Roche). Large Newbler contigs were broken into 4,499 overlapping fragments of 1,000 bp 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 adjust inflated q-scores. A hybrid 454/Sanger assembly was made using PGA assembler. Possible mis-assemblies were corrected and gaps between contigs were closed by primer walks off Sanger clones and bridging PCR fragments and by editing in Consed. A total of 704 Sanger finishing reads were produced to close gaps, to resolve repetitive regions, and to raise the quality of the finished sequence. 12,171,379 Illumina reads were used to improve the final consensus quality using an in-house developed tool (the Polisher [41]). The error rate of the completed genome sequence is less than 1 in 100,000.

<http://standardsingenomics.org>

Together, the combination of the Sanger and 454 sequencing platforms provided 65.38× coverage of the genome. The final assembly contains 46,659 Sanger reads and 601,307 pyrosequencing reads.

## Genome annotation

Genes were identified using [Prodigal](#) [42] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI [GenePRIMP pipeline](#) [43]. 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. Additional gene prediction analysis and functional annotation was performed within the [Integrated Microbial Genomes - Expert Review](#) (IMG-ER) platform [44].

## Genome properties

The genome is 4,123,179 bp long and comprises one main circular chromosome with a 74.3% G+C content (Table 3 and Figure 3). Of the 3,788 genes predicted, 3,735 were protein-coding genes, and 53 RNAs; 57 pseudogenes were also identified. The majority of the protein-coding genes (71.1%) were assigned a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

**Table 3.** Genome Statistics

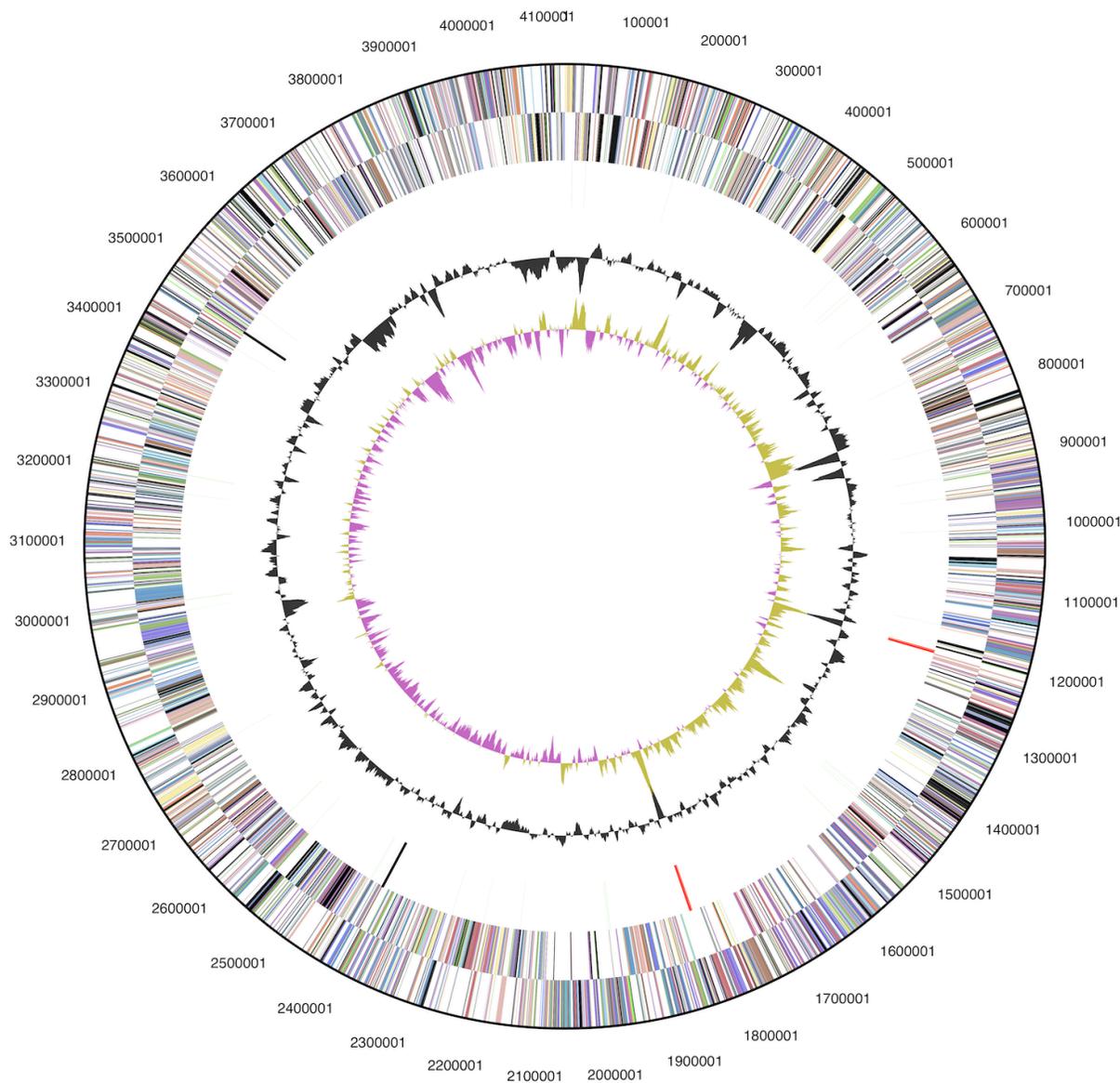
Attribute	Value	% of Total
Genome size (bp)	4,123,179	100.00%
DNA coding region (bp)	3,725,265	90.35%
DNA G+C content (bp)	3,063,259	74.29%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	3,788	100.00%
RNA genes	53	1.40%
rRNA operons	6	
Protein-coding genes	3,735	98.60%
Pseudo genes	57	1.50%
Genes with function prediction	2,692	71.07%
Genes in paralog clusters	435	11.48%
Genes assigned to COGs	2,572	67.90%
Genes assigned Pfam domains	2,758	72.81%
Genes with signal peptides	944	24.92%
Genes with transmembrane helices	1,004	26.50%
CRISPR repeats	0	

## Insights from genome sequence

A closer look on the genome sequence of *C. flavigena* revealed a set of genes which are probably responsible for the yellowish color of *C. flavigena* cells by encoding enzymes that are involved in the synthesis of carotenoids. Carotenoids are produced by the action of geranylgeranyl pyrophosphate synthase (Cfla\_2893), squalene/phytoene synthase (Cfla\_2892), phytoene desaturase (Cfla\_2891), lycopene cyclase (Cfla\_2890, Cfla\_2889) and lycopene elongase (Cfla\_2888). Cfla\_2893 is declared as a pseudo gene, but when ignoring the frame shift the deduced amino acid sequence shows significant similarity to geranylgeranyl pyrophosphate synthases. Geranylgeranyl pyrophosphate synthases start the biosynthesis of carotenoids by combining farnesyl pyrophosphate with C<sub>5</sub> isoprenoid units to C<sub>20</sub>-molecules, geranylgeranyl pyrophosphate. The phytoene synthase catalyzes the condensation of two geranylgeranyl pyrophosphate molecules followed by the removal of diphosphate and a proton shift leading to the formation of phytoene. Sequential desaturation steps are conducted by the phytoene desaturase followed by cyclisation of the ends of the molecules catalyzed by the lycopene cyclase [45]. It is remarkable that the genes belonging to the putative carotenoid biosynthesis clusters of *Beutenbergia cavernae* (Bcav\_3492-Bcav\_3488) [46], *Leifsonia xyli* subsp. *xyli* (*crtE*, *crtB*, *crtI*, *crtYe*,

*lctB*, *crtEb*) and *Sanguibacter keddieii* (Sked\_12750-Sked\_12800) [47] have a similar size and show the same organization as in the genome of *C. flavigena*.

In the eighth edition of Bergey's manual the members of the genus *Cellulomonas* are described as motile by one or a few flagella or non-motile, even within the genus both characteristics occur [32]. Regarding the motility of *C. flavigena* there are different observations described. Thayer *et al.* (1984) report the existence of polar multitrichous flagella [31], whereas Stackebrandt *et al.* (1979) and Schaal (1986) reported *C. flavigena* as non-motile [3,48]. In contrast to Thayer's observation we found no genes coding proteins belonging to the category 'flagellum structure and biogenesis' in the genome sequence. Kenyon *et al.* (2005) report for the genus *Cellulomonas* a coherency between the production of curdlan, a  $\beta$ -1,3-glucan, and non-motility. They observed that the production of curdlan EPS by the non-motile *C. flavigena* leads to a closer adherence to cellulose and hemicellulose. In contrast, cells of the motile *Cellulomonas* strain *C. gelida* produce no curdlan EPS and are not directly attached to the cellulose fibers [28]. The production of curdlan by *C. flavigena* is consistent with the observation of 17 glycosyl transferases (GT) belonging to family 2, as  $\beta$ -1,3-glucan synthases are often found in this GT family.



**Figure 3.** Graphical circular map of the genome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

The characteristic attribute of *C. flavigena* and the other members of the genus *Cellulomonas* is the ability to degrade cellulose, xylan and starch. The most molecular work has been done on cellulase and xylanase genes from *C. fimi*, but also cellulases, xylanases and chitinases of *C. flavigena* were identified and characterized [49-52]. The genome sequence and the subsequent annotation revealed that 9.6% of encoded proteins are classified into the COG category 'carbohydrate transport and metabolism'. Among them several genes coding for xylan degrading enzymes; 14 genes coding for putative endo-1,4- $\beta$ -xylanases belonging to glycoside hydrolase family 10 and five genes encoding

$\beta$ -xylosidases. For the hydrolysis of cellulose the concerted action of endo-1,4- $\beta$ -glucanases, 1,4- $\beta$ -cellobiohydrolases and  $\beta$ -glucosidases is necessary. Endo-1,4- $\beta$ -glucanases randomly cleave within the cellulose molecule and increase the number of non-reducing ends which are attacked by 1,4- $\beta$ -cellobiohydrolases. The released cellobiose is cleaved by  $\beta$ -glucosidases. In the genome of *C. flavigena* two genes coding endo-1,4- $\beta$ -glucanases (Cfla\_0016, Cfla\_1897), three genes encoding 1,4- $\beta$ -cellobiohydrolases (Cfla\_1896, Cfla\_2912, Cfla\_2913) and three genes coding  $\beta$ -glucosidases (Cfla\_1129, Cfla\_3027, Cfla\_2913) were identified.

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

Code	value	%age	Description
J	165	5.8	Translation, ribosomal structure and biogenesis
A	0	0.0	RNA processing and modification
K	270	9.5	Transcription
L	146	5.2	Replication, recombination and repair
B	1	0.0	Chromatin structure and dynamics
D	24	0.9	Cell cycle control, mitosis and meiosis
Y	0	0.0	Nuclear structure
V	64	2.3	Defense mechanisms
T	155	5.5	Signal transduction mechanisms
M	146	5.2	Cell wall/membrane biogenesis
N	8	0.3	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	38	1.3	Intracellular trafficking and secretion
O	99	3.5	Posttranslational modification, protein turnover, chaperones
C	163	5.8	Energy production and conversion
G	272	9.6	Carbohydrate transport and metabolism
E	209	7.4	Amino acid transport and metabolism
F	85	3.0	Nucleotide transport and metabolism
H	129	4.6	Coenzyme transport and metabolism
I	93	3.3	Lipid transport and metabolism
P	131	4.6	Inorganic ion transport and metabolism
Q	50	1.8	Secondary metabolites biosynthesis, transport and catabolism
R	358	12.7	General function prediction only
S	222	7.9	Function unknown
-	1,216	32.1	Not in COGs

## Acknowledgements

We would like to gratefully acknowledge the help of Gabriele Gehrlich-Schröter for growing *C. flavigena* cultures and Susanne Schneider for DNA extraction and quality analysis (both at DSMZ). This work was performed under the auspices of the US Department of Energy Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract

No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, UT-Battelle and Oak Ridge National Laboratory under contract DE-AC05-00OR22725, as well as German Research Foundation (DFG) INST 599/1-2.

## References

1. Kellermann FK, McBeth FM. The fermentation of cellulose. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt 2*. 1912; **39**:502-552.
2. Bergey DH, Breed RS, Hammer RW, Harrison FC, Huntoon FM. *Bergey's manual of determinative bacteriology*, 1<sup>st</sup> ed. The Willimas and Wilkins Co., Baltimore 1923.
3. Stackebrandt E, Kandler O. Taxonomy of the genus *Cellulomonas*, based on phenotypic characters and deoxyribonucleic acid - deoxyribonucleic acid homology, and proposal of seven neotype

- strains. *Int J Syst Bacteriol* 1979; **29**:273-282. [doi:10.1099/00207713-29-4-273](https://doi.org/10.1099/00207713-29-4-273)
4. Braden AR, Thayer DW. Serological study of *Cel-lulomonas*. *Int J Syst Bacteriol* 1976; **26**:123-126. [doi:10.1099/00207713-26-2-123](https://doi.org/10.1099/00207713-26-2-123)
  5. Skerman VBD, Mc Gowan V, Sneath PHA, eds. Approved Lists of Bacterial Names. *Int J Syst Bacteriol* 1980; **30**:225-420. [doi:10.1099/00207713-30-1-225](https://doi.org/10.1099/00207713-30-1-225)
  6. Euzéby JP. List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. [List of Prokaryotic names with Standing in Nomenclature. <http://www.bacterio.net>]. *Int J Syst Bacteriol* 1997; **47**:590-592. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/00207713-47-2-590/) [doi:10.1099/00207713-47-2-590](https://doi.org/10.1099/00207713-47-2-590)
  7. Jones BE, Grant WD, Duckworth AW, Schumann P, Weiss N, Stackebrandt E. *Cellulomonas bogoriensis* sp. nov., an alkaliphilic cellulomonad. *Int J Syst Evol Microbiol* 2005; **55**:1711-1714. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.63646-0/) [doi:10.1099/ijs.0.63646-0](https://doi.org/10.1099/ijs.0.63646-0)
  8. Elberson MA, Malekzadeh F, Yazdi MT, Kame-ranpour N, Noori-Daloui MR, Matte MH, Shaha-mat M, Colwell RR, Sowers KR. *Cellulomonas persica* sp. nov. and *Cellulomonas iranensis* sp. nov., mesophilic cellulose-degrading bacteria isolated from forest soils. *Int J Syst Evol Microbiol* 2000; **50**:993-996. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/00207713-50-5-993-996/)
  9. Chun J, Lee JH, Jung Y, Kim M, Kim S, Kim BK, Lim YW. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *Int J Syst Evol Microbiol* 2007; **57**:2259-2261. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.64915-0/) [doi:10.1099/ijs.0.64915-0](https://doi.org/10.1099/ijs.0.64915-0)
  10. Knapp BA, Podmirseg SM, Seeber J, Meyer E, In-sam H. Diet-related composition of the gut mi-crobiota of *Lumbricus rubellus* as revealed by a molecular fingerprinting technique and cloning. *Soil Biol Biochem* 2009; **41**:2299-2307. [doi:10.1016/j.soilbio.2009.08.011](https://doi.org/10.1016/j.soilbio.2009.08.011)
  11. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformat-ics* 2002; **18**:452-464. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101093/bioinformatics/18.3.452/) [doi:10.1093/bioinformatics/18.3.452](https://doi.org/10.1093/bioinformatics/18.3.452)
  12. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/00207713-17-3-540-552/)
  13. Stamatakis A, Hoover P, Rougemont J. A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Syst Biol* 2008; **57**:758-771. [PMID:18853362](https://pubmed.ncbi.nlm.nih.gov/101080/10635150802429642/) [doi:10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642)
  14. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How Many Bootstrap Replicates Are Necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. [doi:10.1007/978-3-642-02008-7\\_13](https://doi.org/10.1007/978-3-642-02008-7_13)
  15. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metage-nomic projects and their associated metadata. *Nucleic Acids Res* 2010; **38**:D346-D354. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101093/nar/gkp848/) [doi:10.1093/nar/gkp848](https://doi.org/10.1093/nar/gkp848)
  16. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, An-giuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101038/nbt1360/) [doi:10.1038/nbt1360](https://doi.org/10.1038/nbt1360)
  17. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms. Proposal for the do-mains *Archaea* and *Bacteria*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101073/pnas.87.12.4576/) [doi:10.1073/pnas.87.12.4576](https://doi.org/10.1073/pnas.87.12.4576)
  18. Garrity GM, Holt J. Taxonomic outline of the *Archaea* and *Bacteria*. In: *Bergey's Manual of Sys-tematic Bacteriology*, 2<sup>nd</sup> ed. vol. 1. *The Archaea, Deeply Branching and Phototrophic Bacteria*. Garrity GM, Boone DR and Castenholz RW (eds). 2001; 155-166.
  19. Stackebrandt E, Rainey FA, Ward-Rainey NL. Proposal for a new hierarchic classification sys-tem, *Actinobacteria* classis nov. *Int J Syst Bacteriol* 1997; **47**:479-491. [doi:10.1099/00207713-47-2-479](https://doi.org/10.1099/00207713-47-2-479)
  20. Buchanan RE. Studies in the nomenclature and classification of bacteria. II. The primary subdivi-sions of the *Schizomycetes*. *J Bacteriol* 1917; **2**:155-164. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/00207713-2-2-155-164/)
  21. Zhi XY, Li WJ, Stackebrandt E. An update of the structure and 16S rRNA gene sequence-based de-finition of higher ranks of the class *Actinobacte-ria*, with the proposal of two new suborders and four new families and emended descriptions of the higher taxa. *Int J Syst Evol Microbiol* 2009; **59**:589-608. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.65780-0/) [doi:10.1099/ijs.0.65780-0](https://doi.org/10.1099/ijs.0.65780-0)
  22. Stackebrandt E, Prauser H. The family *Cellulomo-nadaceae*. In: *The prokaryotes*. Balows A, Trüper HG, Dworkin M, Harder W, Schleifer KH (eds). 1991; 1323-1345.
  23. Stackebrandt E, Schumann P. Description of *Bo-goriellaceae* fam. nov., *Dermaococcaceae* fam. nov., *Rarobacteraceae* fam. nov. and *Sanguibac-teraceae* fam. nov. and emendation of some fami-

- lies of the suborder *Micrococccineae*. *Int J Syst Evol Microbiol* 2000; **50**:1279-1285. [PubMed](#)
24. Stackebrandt E, Prauser H. Assignment of the genera *Cellulomonas*, *Oerskovia*, *Promicromonospora*, and *Jonesia* to *Cellulomonadaceae* fam. nov. *Syst Appl Microbiol* 1991; **14**:261-265.
25. List Editor. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List no. 39. *Int J Syst Bacteriol* 1991; **41**:580-581. [PubMed](#)  
[doi:10.1099/00207713-41-4-580](https://doi.org/10.1099/00207713-41-4-580)
26. Stackebrandt E, Seiler H, Schleifer KH. Union of the genera *Cellulomonas* Bergey *et al.* and *Oerskovia* Prauser *et al.* in a redefined genus *Cellulomonas*. [Orig A]. *Zentralbl Bakteriol* 1982; **3**:401-409.
27. Bergey DH, Harrison FC, Breed RS, Hammer BW, Huntoon FM. In: Bergey DH, Harrison FC, Breed RS, Hammer BW, Huntoon FM (eds), *Bergey's manual of determinative bacteriology*, First Edition, The Williams and Wilkins Co., Baltimore, 1923, p. 1-442.
28. Kenyon WJ, Esch SW, Buller CS. The curdlan-type exopolysaccharide produced by *Cellulomonas flavigena* KU forms part of an extracellular glyco-calyx involved in cellulose degradation. *Antonie Van Leeuwenhoek* 2005; **87**:143-148. [PubMed](#)  
[doi:10.1007/s10482-004-2346-4](https://doi.org/10.1007/s10482-004-2346-4)
29. Classification of bacteria and archaea in risk groups. [www.baua.de](http://www.baua.de) TRBA 466
30. 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](https://doi.org/10.1038/75556)
31. Thayer DW. Motility and flagellation of cellulomonads. *Int J Syst Bacteriol* 1984; **34**:218-219.  
[doi:10.1099/00207713-34-2-218](https://doi.org/10.1099/00207713-34-2-218)
32. Keddie RM. Genus III. *Cellulomonas* Bergey *et al.* 1923, 154, emend. mut. char. Clark 1952, 50AL. In: *Bergey's Manual of Systematic Bacteriology*. Williams and Wilkins. Baltimore, MD. Sneath PHA, Mair NS, Sharpe ME, Holt JG (eds). 1974; 629-631.
33. Kim BH. Carbohydrate catabolism in cellulolytic strains of *Cellulomonas*, *Pseudomonas*, and *No-cardia*. *Kor J Microbiol* 1987; **25**:28-33.
34. Fieldler F, Kandler O. Die Mureintypen in der Gattung *Cellulomonas* Bergey *et al.* *Arch Microbiol* 1973; **89**:41-50. [doi:10.1007/BF00409398](https://doi.org/10.1007/BF00409398)
35. Funke G, Ramos CP, Collins MD. Identification of some clinical strains of CDC coryneform group A-3 and A-4 bacteria as *Cellulomonas* species and proposal of *Cellulomonas hominis* sp. nov. for some group A-3 strains. *J Clin Microbiol* 1995; **33**:2091-2097. [PubMed](#)
36. Collins MD, Jones D. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication. *Microbiol Rev* 1981; **45**:316-354. [PubMed](#)
37. Mimmikin DE, Collins MD, Goodfellow M. Fatty acid and polar lipid composition in the classification of *Cellulomonas*, *Oerskovia* and related taxa. *J Appl Bacteriol* 1979; **47**:87-95.
38. Klenk HP, Göker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* 2010; **33**:175-182. [PubMed](#)  
[doi:10.1016/j.syapm.2010.03.003](https://doi.org/10.1016/j.syapm.2010.03.003)
39. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, *et al.* A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature* 2009; **462**:1056-1060. [PubMed](#)  
[doi:10.1038/nature08656](https://doi.org/10.1038/nature08656)
40. List of growth media used at DSMZ:  
[http://www.dsmz.de/microorganisms/media\\_list.php](http://www.dsmz.de/microorganisms/media_list.php)
41. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008.
42. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal Prokaryotic Dynamic Programming Gene-finding Algorithm. *BMC Bioinformatics* 2010; **11**:119. [PubMed](#)  
[doi:10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119)
43. Pati A, Ivanova N, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: A gene prediction improvement pipeline for microbial genomes. *Nat Methods* 2010; **7**:455-457. [PubMed](#)  
[doi:10.1038/nmeth.1457](https://doi.org/10.1038/nmeth.1457)
44. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. [PubMed](#)  
[doi:10.1093/bioinformatics/btp393](https://doi.org/10.1093/bioinformatics/btp393)
45. Sandmann G. Carotenoid biosynthesis and biotechnological application. *Arch Biochem Biophys* 2001; **385**:4-12. [PubMed](#)  
[doi:10.1006/abbi.2000.2170](https://doi.org/10.1006/abbi.2000.2170)

46. Land M, Pukall R, Abt B, Göker M, Rohde M, Glavina Del Rio T, Tice H, Copeland A, Cheng JF, Lucas S, et al. Complete genome sequence of *Beutenbergia cavernae* type strain (HKI 0122<sup>T</sup>). *Stand Genomic Sci* 2009; **1**:21-28. [doi:10.4056/sigs.1162](https://doi.org/10.4056/sigs.1162)
47. Ivanova N, Sikorski J, Sims D, Brettin T, Detter JC, Han C, Lapidus A, Copeland A, Glavina Del Rio T, Nolan M, et al. Complete genome sequence of *Sanguibacter keddieii* type strain (ST-74<sup>T</sup>). *Stand Genomic Sci* 2009; **1**:110-118. [doi:10.4056/sigs.16197](https://doi.org/10.4056/sigs.16197)
48. Schaal KP. Genus *Actinomyces* Harz 1877, 133AL. In: *Bergey's Manual of Systematic Bacteriology*. Williams and Wilkins. Baltimore, MD. Sneath PHA, Mair NS, Sharpe ME, Holt JG (eds). 1986; 1383-1418.
49. Gutiérrez-Nava A, Herrera-Herrera A, Mayorga-Reyes L, Salgado LM, Ponce-Noyola T. Expression and characterization of the *celc1B* gene from *Cellulomonas flavigena* encoding an endo- $\beta$ -1,4-glucanase. *Curr Microbiol* 2003; **47**:359-363. [PubMed doi:10.1007/s00284-002-4016-y](https://pubmed.ncbi.nlm.nih.gov/1007/s00284-002-4016-y/)
50. Mayorga-Reyes L, Morales Y, Salgado LM, Ortega A, Ponce-Noyola T. *Cellulomonas flavigena*: characterization of an endo-1,4-xylanase tightly induced by sugarcane bagasse. *FEMS Microbiol Lett* 2002; **214**:205-209. [PubMed doi:10.1111/j.1574-6968.2002.tb11348.x](https://pubmed.ncbi.nlm.nih.gov/10.1111/j.1574-6968.2002.tb11348.x/)
51. Chen HC, Hsu MF, Jiang ST. Purification and characterization of an exo-N,N'-diacetylchitobiohydrolase-like enzyme from *Cellulomonas flavigena* NTOU 1. *Enzyme Microb Technol* 1997; **20**:191-197. [doi:10.1016/S0141-0229\(96\)00111-1](https://doi.org/10.1016/S0141-0229(96)00111-1)
52. Horcasitas CM, López JO, Plaza IM. Xylanases from *Cellulomonas flavigena*: purification and characterization. *Biotechnol Tech* 1998; **12**:663-666. [doi:10.1023/A:1008852618130](https://doi.org/10.1023/A:1008852618130)