Description of *Trichococcus ilyis* sp. nov. by combined physiological and *in silico* genome hybridization analyses

Nikolaos Strepis,[1,2] Irene Sánchez-Andrea,[1] Antonie H. van Gelder,[1] Henri van Kruistum,[1] Nicole Shapiro,[3] Nikos Kyrpides,[3] Markus Göker,[4] Hans-Peter Klenk,[4,5] Peter Schaap,[2] Alfons J. M. Stams[1,6] and Diana Z. Sousa[1,6]

Correspondence
Diana Z. Sousa
diana.sousa@wur.nl

[1]Laboratory of Microbiology, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands

[2]Laboratory of Systems and Synthetic Biology, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands

[3]DOE Joint Genome Institute, 2800 Mitchell Drive 100, Walnut Creek, CA 94598, USA

[4]Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Inhoffenstraße 7B, 38124 Braunschweig, Germany

[5]School of Biology, Newcastle University, Ridley Building 2, Newcastle, NE1 7RU, UK

[6]Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

Species of the genus *Trichococcus* share high similarity of their 16S rRNA gene sequences (>99 %). Digital DNA–DNA hybridization values (dDDH) among type strains of all described species of the genus *Trichococcus* (*T. flocculiformis* DSM 2094^T, *T. pasteurii* DSM 2381^T, *T. collinsii* DSM 14526^T, *T. palustris* DSM 9172^T, and *T. patagoniensis*DSM 18806^T) indicated that *Trichococcus* sp. strain R210^T represents a novel species of the genus *Trichococcus*. The dDDH values showed a low DNA relatedness between strain R210^T and all other species of the genus *Trichococcus* (23–32%). Cells of strain R210^T were motile, slightly curved rods, 0.63–1.40×0.48–0.90 μm and stained Gram-positive. Growth was optimal at pH 7.8 and at temperature of 30 °C. Strain R210^T could utilize several carbohydrates, and the main products from glucose fermentation were lactate, acetate, formate and ethanol. The genomic DNA G+C content of strain R210^T was 47.9 mol%. Based on morphological, physiological and biochemical characteristics along with measured dDDH values for all species of the genus *Trichococcus*, it is suggested that strain R210^T represents a novel species within the genus *Trichococcus*, for which the name *Trichococcus ilyis* sp. nov. is proposed. The type strain is R210^T (=DSM 22150^T=JCM 31247^T).

Abbreviations: DDH, DNA-DNA hydridization; dDDH, digital DNA-DNA hydridization.

The GenBank/EMBL/DDBJ accession number for the 16S rRNA gene sequence of *Trichococcus ilyis* strain R210^T is FJ374769. Whole genome sequences of strain R210^T, *T. pasteurii* DSM 2381^T, *T. flocculiformis* DSM 2094^T, *T. collinsii* DSM 14526^T, *T. palustris* DSM 9172^T and *T. patagoniensis* DSM 18806^T were determined in this study and are available at the European Nucleotide Archive (www.ebi.ac.uk/ena) under accession numbers PRJEB12709, PRJEB12708, PRJEB12705, PRJEB12704, PRJEB12707 and PRJEB13128, respectively.

Trichococcus sp. strain R210^T was previously isolated from a sulfate-reducing bioreactor treating municipal wastewater mixed with an industrial stream containing citrate (Stams *et al.*, 2009). Isolation was performed in basal medium supplemented with 10 mM citrate as sole energy and carbon source (without sulfate). The main fermentation products of citrate were acetate and formate, whereas the main products of glucose fermentation were lactate, ethanol, acetate and formate. This product profile is typically observed in heterolactic fermenters (Eiteman & Ramalingam, 2015).

At the time of writing, there are five species of the genus *Trichococcus* with validly published names, *Trichococcus*

flocculiformis (Scheff *et al.*, 1984), *T. pasteurii* (Schink *et al.*, 1984), *T. collinsii* (Liu *et al.*, 2002), *T. palustris* (Zhilina *et al.*, 1995), and *T. patagoniensis* (Pikuta *et al.*, 2006). All these species can use a broad range of sugars and polyols for growth. A peculiar feature of members of the genus *Trichococcus* is the ability to grow over a wide range of temperatures, from close to 0 °C to over 45 °C. This makes such strains particularly interesting for biotechnological applications at mildly low temperature. They are also tolerant to oxygen, which is advantageous because it eliminates the need of maintaining bioreactors completely oxygen-free.

Morphologically, members of the genus *Trichococcus* are pleomorphic (regular cocci, olive-like rods) and may have several cell arrangements (cells can grow individually, in pairs, in long chains, or in grape-like conglomerates) (Liu *et al.*, 2002). The similarity of the 16S rRNA gene sequence of described species of the genus *Trichococcus* (99–100 %) is above the threshold for the definition of a new species (Kim *et al.*, 2014). However, DNA–DNA hybridization (DDH) values among them are lower than 70 %, the threshold for the delineation of new species (Pikuta *et al.*, 2006). The presence of highly conserved 16S rRNA genes in different species is not so uncommon. Another example can be found in the genus *Bacillus*, where some species can have a 16S rRNA gene sequence similarity of 97–100 % (Liu *et al.*, 2015). For these cases, DDH is accepted as the gold standard for complete taxonomic classification and proposal of new species. Experimental DDH assays are based on optical measurements and are prone to experimental errors (Schleifer, 2009). The rapid progress in genome sequencing allows utilization of computational methods for genome comparison and taxonomic differentiation (Klenk & Göker, 2010). Genome BLAST Distance Phylogeny (GBDP) approaches have been optimized for calculation of digital equivalents for DDH values (dDDH), which are analogous to DNA– DNA hybridization (Meier-Kolthoff *et al.*, 2014).

To clarify the phylogenetic assignment of strain R210^T, the genome was sequenced along with the genomes of all five previously described species of the genus *Trichococcus* and dDDH values were calculated between these species using GBDP. Consequently, genome-guided physiological characterization was performed.

T. flocculiformis DSM 2094^T, *T. pasteurii* DSM 2381^T, *T. collinsii* DSM 14526^T, *T. palustris* DSM 9172^T, and *T. patagoniensis* DSM 18806^T were obtained from the German Collection for Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and cultivated on DSMZ recommended media. Strain R210^T was previously isolated by Stams *et al.* (2009), and since then was preserved in our laboratory. Strain R210^T was grown at 30 °C on an anaerobic bicarbonate-buffered medium (Stams *et al.*, 1993) supplemented with 0.1 g l^{-1} yeast extract and 20 mM D-glucose.

Genomic DNA of all the strains was extracted using MasterPure Gram-positive DNA purification Kit (Epicenter) according to manufacturer's instructions. The genomes of strain R210^T, *T. pasteurii* DSM 2381^T, *T. flocculiformis* DSM 2094^T, *T. collinsii* DSM 14526^T and *T. palustris* DSM 9172^T were sequenced at the JGI (Walnut Creek, CA) using an Illumina HiSeq2000 platform within the Genomic Encyclopaedia of archaeal and bacterial type strains, phase II (Kyrpides *et al.*, 2014). The genome of *T. patagoniensis* DSM 18806^T was sequenced at Baseclear (Leiden, The Netherlands) using an Illumina Hiseq2500 platform.

The genomes of strain R210^T (9 304 840 reads and 151 bp read length), *T. pasteurii* DSM 2381^T (10 101 940 reads and 151 bp read length), *T. flocculiformis* DSM 2094^T (10 454 276 reads and 151 bp read length), *T. collinsii* DSM 14526^T (9 948 768 reads and 151 bp read length), *T. palustris* DSM 9172^T (11 006 414 reads and 151 bp read length), and *T. patagoniensis* DSM 18806^T (5 740 601 reads and 126 bp read length) were assembled using a pipeline comprising: Ray (Boisvert *et al.*, 2012) to generate an initial assembly, followed by Opera (Gao *et al.*, 2011) for genome scaffolding, and CAP3 (Huang & Madan, 1999) for assembling optimization. For Ray assembler, the optimal kmer size was calculated with KmerGenie (Chikhi & Medvedev, 2014). In addition, the genomes were assembled by JGI (Walnut Creek, CA) using AllPATHS-LG (Gnerre *et al.*, 2011) and Velvet (Zerbino & Birney, 2008). Both assemblies were merged and improved with SSPACE (Boetzer *et al.*, 2011).

The dDDH values [including confidence intervals (CIs)] between strain R210^T and the type strains of established species of the genus *Trichococcus* were measured using the Genome-to-Genome Distance Calculator available at http://ggdc.dsmz.de (Meier-Kolthoff *et al.*, 2013).

Genome annotation of strain R210^T was performed using an in-house pipeline based on published tools and databases: Prodigal (Hyatt *et al.*, 2010), InterProsScan (Jones *et al.*, 2014), tRNAscan-SE (Lowe, 1997), RNAmmer (Lagesen *et al.*, 2007), UniRef50 (Suzek *et al.*, 2007), Swiss-Prot database (The UniProt Consortium, 2014), and combined with RAST (Aziz *et al.*, 2008). The merged results from the in-house annotation pipeline and RAST were used to curate the annotation of the genome by completing the full description of the encoding regions. The genomic sequences, as well the coding sequences, were used as input for BRIG (BLAST Ring Image Generator) (Alikhan *et al.*, 2011). Sequences were assigned with KEGG orthology (KO) identifiers using KEGG Automatic Annotation Server (KAAS) and further linked to metabolic pathways (Moriya *et al.*, 2007).

16S rRNA gene sequence analysis was performed with ARB v6.0 (Ludwig *et al.*, 2004) using the non-redundant SILVA database (Quast *et al.*, 2013). 16S rRNA gene sequences of strain R210^T and all the currently isolated species of the genus *Trichococcus* were implemented in the all-species living tree project (Yarza *et al.*, 2008). Three phylogenetic trees were generated based on maximum-likelihood (RAxML algorithm), neighbour-joining (Jukes–Cantor correction) and maximum-parsimony methods, and a consensus tree was generated with the tool integrated into ARB v6.0.

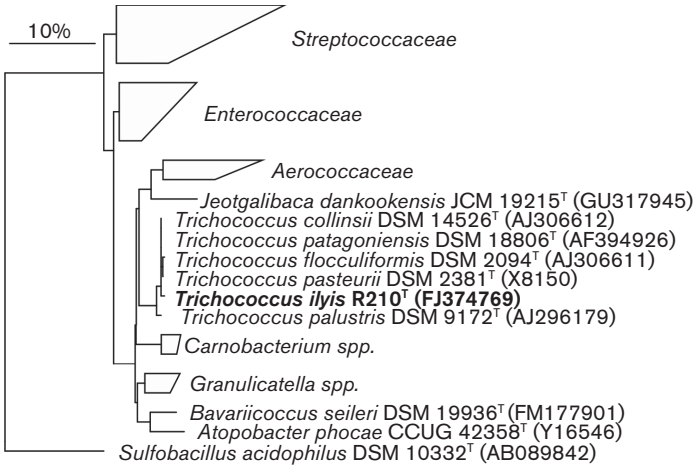

Fig. 1. Consensus tree of 16S rRNA gene sequences showing the phylogenetic affiliation of strain R210^T. Bar, 10 % sequence divergence.

Cell morphology of strain R210^T was observed using a phase-contrast microscope (Leica DM2000; Leica Microsystems) on glucose-grown cultures. Cell sizes were measured using the software Fiji-ImageJ (Schindelin *et al.*, 2012). Growth was estimated by measurements of optical density at 600 nm with a spectrophotometer (Hitachi U-1500; Labstuff).

The cellular fatty acids of strain R210^T were analysed at the DSMZ (Braunschweig, Germany). The analysis was done with the system of Sherlock MIS (MIDI) using previously described methods (Kuykendall *et al.*, 1988; Miller, 1982). Cells for lipids analysis were cultivated in pre-reduced tryptic soy broth (BD) as previously used for other *Trichococcus* strains (Liu *et al.*, 2002).

The growth of strain R210^T was tested for a temperature range from 0 to 45 °C and a pH range from 6.0 to 9.6. The ability of strain R210^T to utilize a set of substrates had been previously tested (van Gelder *et al.*, 2012). In this study, the following additional substrates were tested (at a concentration of 20 mM, unless otherwise mentioned): L-lactose, sucrose, D-galactose, D-fructose, D-glucose, D-arabinose, formate, D-galacturonate, D-glucuronate, D-mannose, L-rhamnose, raffinose, L-sorbose, D-gluconate, D-tagatose, D-xylose, starch, glycerol, methanol, ethanol, D-sorbitol, inositol, D-dulcitol and H_2/CO_2 [80/20 (v/v), 1.5×10^5 Pa]. Substrates for testing were selected based on genomic information, i.e. the presence of genes potentially encoding for the degradation of those substrates. Soluble substrates and intermediates (sugars and volatile fatty acids) were measured with a Thermo Electron HPLC system equipped with an Agilent Metacarb 67H column (300×6.5 mm) (Thermo) and a refractive index detector. The mobile phase used was sulfuric acid (5 mM) at a flow rate of 0.8 ml min^{-1}. The column temperature was set at 45 °C. Hydrogen and CO_2 were measured with a Shimadzu GC-2014 gas chromatograph equipped with a Molsieve 13X column (2 m×3 mm) (Shimadzu) and a thermal conductivity detector (TCD). Argon was used as carrier gas at a flow rate of 50 ml min^{-1}, and temperatures in the injector, column and detector were 80, 100 and 130 °C, respectively.

Table 1. dDDH values (%) with confidence intervals between type strains of species of the genus *Trichococcus*

dDDH values were calculated using the GGDC tool, http://ggdc.dsmz.de/distcalc2.php. 16S rRNA gene sequence similarities are shown in parentheses.

Strain	R210^T	*T. pasteurii* DSM 2381^T	*T. flocculiformis* DSM 2094^T	*T. collinsii* DSM 14526^T	*T. palustris* DSM 9172^T	*T. patagoniensis* DSM 18806^T
R210^T	100±0.0 (100)	32±3.0 (99.5)	30±3.0 (99.5)	30±3.0 (99.7)	23±2.8 (99.1)	31±2.5 (99.7)
T. pasteurii DSM 2381^T	32±3.0 (99.5)	100±0.0 (100)	48±3.3 (99.8)	40±3.1 (99.9)	23±2.9 (99.1)	39±2.5 (99.9)
T. flocculiformis DSM 2094^T	30±3.0 (99.5)	42±3.1 (99.8)	100±0.0 (100)	36±3.0 (99.7)	23±2.8 (99.0)	36±2.4 (99.7)
T. collinsii DSM 14526^T	30±3.0 (99.7)	40±3.2 (99.9)	36±3.0 (99.7)	100±0.0 (100)	23±2.9 (99.3)	41±2.5 (100)
T. palustris DSM 9172^T	23±2.8 (99.1)	23±2.9 (99.1)	23±2.8 (99.0)	23±2.0 (99.3)	100±00 (100)	23±2.5 (99.3)
T. patagoniensis DSM 18806^T	31±2.5 (99.7)	39±2.5 (99.9)	36±2.4 (99.7)	41±2.5 (100)	23±2.5 (99.3)	100±0.0 (100)

Phylogenetic analysis of species of the genus *Trichococcus* based on 16S rRNA gene sequences indicates a high similarity (>99 %) among all the species of the genus (Fig. 1 and Table 1). The closest relative of strain R210^T is *T. pasteurii* DSM 2381^T with a 16S rRNA gene sequence similarity of 99.5 % (Table 1). The 16S rRNA gene sequence of *T. pasteurii* DSM 2381^T is 99.9 % similar to the ones of *T. collinsii* DSM 14526^T and *T. patagoniensis* DSM 18806^T, while *T. patagoniensis* DSM 18806^T and *T. collinsii* DSM 14526^T have an identical 16S rRNA gene sequence (100 % similarity). This indicates that the 16S rRNA gene of species of the genus *Trichococcus* has not evolved and taxonomic differentiation of species within this genus needs to be based on DDH. In this work we calculated dDDH values of strain R210^T with type strains of other species of the genus *Trichococcus* and values were all lower than 70 % (Table 1), which is accepted as the threshold value for the definition of new species (Meier-Kolthoff *et al.*, 2013). Based on this, strain R210^T can certainly be classified as a novel species of the genus *Trichococcus*.

Substrate utilization of strain R210^T, and comparison with other species of the genus *Trichococcus*, is presented in Table 2. Strain R210^T can use several carbohydrates, which is corroborated by the presence of numerous putative genes from the Embden-Meyerhof-Parnas (EMP) pathway, pentose phosphate pathway, and phosphoketolase pathway. The main products of glucose fermentation by strain R210^T were acetate, formate, lactate and ethanol. An operon with genes coding for rhamnulokinase (EC 2.7.1.6), L-rhamnose

Table 2. Morphological and physiological characteristics of described species of the genus *Trichococcus*

Strains: 1, R210^T; 2, *T. pasteurii* DSM 2381^T; 3, *T. flocculiformis* DSM 2094^T; 4, *T. collinsii* DSM 14526^T; 5, *T. palustris* DSM 9172^T; 6, *T. patagoniensis* DSM 18806^T. Data are from this study unless otherwise indicated. All strains can grow on glucose and pyruvate, citrate (ND for *T. palustris*) and cellobiose (ND for *T. patagoniensis*). All strains did not grow with H_2/CO_2. Strain R210^T and *T. pasteurii* can grow on D-glucuronate, D-mannose and raffinose and cannot grow on D-galacturonate, starch, D-tagatose, L-sorbose, inositol and dulcitol. +, Positive; −, negative; ND, Not determined.

Characteristic	1	2	3	4	5	6
Source	Sulfate-reducing bioreactor*[a]	Septic pit[b]	Activated sludge[b]	Hydrocarbon-spill site[b]	Swamp[b]	Guano penguin[c]
Gram stain	+	+[b]	+[b]	+[b]	+[b]	Variable[c]
Cell length (μm)	0.63–1.40	1.0–1.5[b]	1.0–2.5*	1.0–2.5[b]	1.0–2.5[b]	1.3–2.0[c]
Cell shape	Cocci	Short chain of cocci[b]	Long chain of cocci[b]	Rods[b]	Cocci[b]	Cocci[c]
pH range (optimum)	6.0–9.6 (7.8)	5.5–9.0[b]	5.8–9.0[b]	6.0–9.0[b] (7.5)	6.2–8.4[b] (7.5)	6.0–10[c] (8.5)
Temperature range (optimum) (°C)	4–40 (30)	0–42[b] (25–30)	4–40[b] (25–30)	7–36[b] (23–30)	0–33[b]	−5–35[c] (28–30)
DNA G+C content (mol%)						
In vitro	ND	45.0[b]	48.0[b]	47.0[b]	48.0[b]	45.8[c]
In silico	47.9	45.7	44.3	44.3	45.9	46.9
Substrate utilization						
Formate	−	−	ND	ND	ND	−[c]
L-Lactate	−[a]	−[b]	−[a]	−[c]	+[c]	−[c]
L-Malate	+[a]	+[b]	−[a]	+[b]	−[b]	+[c]
D-Gluconate	−	+	ND	ND	ND	ND
Glycerol	−	+	−	−	−	−[c]
D-Arabinose	+[a]	+	+[a]	ND	ND	+[c]
D-Xylose	−	+	+[a]	ND	ND	ND
D-Galactose	+	+	ND	ND	ND	ND
L-Rhamnose	+	+	ND	ND	ND	ND
D-Fructose	+[a]	+	+[a]	ND	ND	+[c]
D-Lactose	+	+[b]	+[b]	−[b]	+[b]	−[c]
Maltose	+[a]	+[b]	+[a]	−[b]	+[b]	+[c]
D-Mannitol	+[a]	+[b]	−[a]	+[b]	+[b]	+[c]
Sucrose	+[a]	+	+[a]	ND	ND	+[c]
Methanol	−	−	ND	ND	ND	−[c]
Ethanol	−	−	ND	ND	ND	−[c]
D-Sorbitol	−	+	−[a]	+[b]	+[b]	ND

*Data from: *a*, van Gelder *et al.* (2012); *b*, Liu *et al.* (2002); *c*, Pikuta *et al.* (2006).

isomerase (EC 5.3.1.14), rhamnulose-1-phospate aldolase (EC 4.1.2.19) and lactaldehyde dehydrogenase (EC 1.2.1.22) was identified in strain R210^T (TR210_1063–1066). Lactaldehyde dehydrogenase is an essential enzyme for the conversion of rhamnose to 1,2-propanediol (Petit *et al.*, 2013), which was observed as a main product in the experimental growth tests of strain R210^T on rhamnose (together with acetate and formate). An additional operon potentially linked to rhamnose metabolism (TR210_973–982) contained genes encoding for L-rhamnose mutarotase, which may play a role in the first steps of rhamnose degradation (Petit *et al.*, 2013), together with tripartite ATP-independent periplasmic (TRAP) transporters, which may be involved in substrate import to the cell (Mulligan *et al.*, 2011). A mannose-specific IIC component gene (EC 2.7.1.69, TR210_1375) from the phosphotransferase system (PTS) is also present in the genome of strain R210^T. The PTS system is responsible for the uptake of carbohydrates and includes the mannose-specific IIC component, which can catalyse mannose to mannose 6-phosphate, during uptake (Okochi *et al.*, 2007). Sucrose was also metabolized by strain R201^T; genes encoding for beta-glucuronidase (EC 3.2.1.31, TR210_689 and TR210_756,), maltose-6'–phosphate glucosidase (EC 3.2.1.122, TR210_416) and glycogen operon protein (EC 3.2.1.-, TR210_2418) were identified.

Strain R210^T is a citrate-fermenting bacterium and its genome contains some, but not all, genes of the tricarboxylic acid cycle (TCA). The gene encoding for isocitrate dehydrogenase (EC 1.1.1.42), involved in the conversion of isocitrate (via oxalosuccinate) to 2-oxoglutarate, is present in the genome (TR210_2151), but also a NAD^+-dependent isocitrate dehydrogenase (EC 1.1.1.41, TR210_901). The latter enzyme is capable of direct converting of isocitrate to 2-oxoglutarate.

Table 3. Cellular fatty acid compositions of members of the genus *Trichococcus*

Strains: 1, R210^T (data from this study); 2, *T. pasteurii* DSM 2381^T (Liu *et al.*, 2002); 3, *T. flocculiformis* DSM 2094^T (Liu *et al.*, 2002); 4, *T. collinsii* DSM 14526^T (Liu *et al.*, 2002); 5, *T. palustris* DSM 9172^T (Liu *et al.*, 2002); 6, *T. patagoniensis* DSM 18806^T (Pikuta *et al.*, 2006). Values are presented as percentages of total fatty acids.

Fatty acid	1	2	3	4	5	6
$C_{12:0}$	0.7	0	1	12	0	0.7
$C_{14:0}$	16.1	14	28	57	21	11.2
$C_{15:0}$	0.6	0	0	0	0	0
$C_{16:1}\omega 9c$	35.5	46	0	18	20	0
$C_{16:0}$	19.7	15	16	14	15	16.5
$C_{18:1}\omega 9c$	19.7	18	6	0	22	21.8
$C_{18:1}\omega 7c$	0.6	0	0	0	0	0
$C_{18:0}$	3.0	2	2	0	4	3.3
$C_{20:1}\omega 9c$	0.8	0	0	0	0	0
$C_{20:1}\omega 7c$	0.8	0	0	0	0	0

T. pasteurii is the closest relative of strain R210^T, and it can use some polyalcohols, such as glycerol and sorbitol. Strain R210^T does not grow with these substrates, it lacks the genes encoding 1,3-propanediol dehydrogenase (EC 1.1.1.202) and glycerol dehydratase (EC 4.2.1.30) required for the conversion of polyalcohols (Wang *et al.*, 2003).

Cells of strain R210^T were slightly curved rods, 0.5–0.9 μm in diameter and 0.6–1.4 μm in length, were motile and occurred singly or in pairs. Growth was observed between 4 and 40 °C with optimum growth at 30 °C. The optimum pH for growth was 7.8 with range between pH 6.0 and pH 9.6. Major fatty acid methyl esters of strain R210^T were $C_{14:0}$ (16.1 %), $C_{16:1}\omega 9c$ (35.5 %), $C_{16:0}$ (19.7 %) and $C_{18:1}\omega 9c$ (19.7 %). These were also the major fatty acids in the type strains of *T. pasteurii* and *T. palustris* (Table 3). Membranes of *T. flocculiformis* DSM 2094^T and *T. patagoniensis* DSM 18806^T lacked the unsaturated fatty acid $C_{16:1}\omega 9c$ (the most abundant membrane fatty acid in strain R210^T and *T. pasteurii* DSM 2381^T). Unsaturated fatty acids are normally associated to increased membrane fluidity, and their relative abundance can vary with environmental factors, e.g. temperature (Zhang & Rock, 2008).

The data of dDDH analysis conducted for all species of the genus *Trichococcus* indicated that strain R210^T considerably differs from other species of the genus *Trichococcus*. Therefore, based on morphological, physiological and biochemical characteristics along with measured dDDH values for all species of the genus *Trichococcus*, it is suggested that strain R210^T represents a novel species within the genus *Trichococcus*, for which the name *Trichococcus ilyis* sp. nov. is proposed.

Description of *Trichococcus ilyis* sp. nov.

Trichococcus ilyis (i'ly.is. Gr. fem. n. *ilys, ilyos* mud, sludge; latinized Gr. gen. n. *ilyis* of sludge referring to the source of isolation).

Cells are motile, slightly curved rods, 0.63–1.40×0.48–0.90 μm. Cells grow in a temperature range of 4 to 40 °C (optimum 30 °C). The pH for growth is 6.0–9.6 (optimum pH 7.8). Utilizes a broad range of carbon sources for growth, such as pyruvate, L-malate, D-arabinose, citrate, D-fructose, D-galactose, D-glucuronate, D-glucose, D-mannitol, D-mannose, L-rhamnose, cellobiose, D-lactose, maltose, sucrose and raffinose. No growth was observed with formate, L-lactate D-galacturonate, D-gluconate, L-sorbose, D-tagatose, D-xylose, starch, glycerol, ethanol, methanol, inositol, D-dulcitol and D-sorbitol. Does not grow autotrophically on H_2/CO_2.

The type strain is R210^T (=DSM 22150^T=JCM 31247^T), isolated from sludge of a sulfate-reducing bioreactor with a citrate-containing waste stream. The genomic G+C content of the type strain is 47.9 mol%.

Acknowledgements

This research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/

2007–2013)/ERC Grant Agreement (323009) and by the Gravitation grant (024.002.002) of the Netherlands Ministry of Education, Culture and Science. The work conducted by the U.S. Department of Energy Joint Genome Institute (DOE-JGI), a DOE Office of Science User Facility, was supported by the Office of Science of the DOE under Contract No. DE-AC02-05CH11231. The authors gratefully acknowledge M. Huntemann, A. Clum, M. Pillay, K. Palaniappan, N. Varghese, N. Mikhailova, D. Stamatis, T.B.K. Reddy, C.Y. Ngan, C. Daum, V. Markowitz, N. Ivanova, and T. Woyke from JGI for their technical assistance.

References

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M. & other authors (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75.

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. & Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579.

Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F. & Corbeil, J. (2012). Ray Meta: scalable *de novo* metagenome assembly and profiling. *Genome Biol* **13**, 122.

Chikhi, R. & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics* **30**, 31–37.

Eiteman, M. A. & Ramalingam, S. (2015). Microbial production of lactic acid. *Biotechnol Lett* **37**, 955–972.

Gao, S., Sung, W. K. & Nagarajan, N. (2011). Opera: reconstructing optimal genomic scaffolds with high-throughput paired-end sequences. *J Comput Biol* **18**, 1681–1691.

Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F. J., Burton, J. N., Walker, B. J., Sharpe, T., Hall, G., Shea, T. P. & other authors (2011). High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* **108**, 1513–1518.

Huang, X. & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Res* **9**, 868–877.

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W. & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 199.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A. & other authors (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Kim, M., Oh, H. S., Park, S. C. & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**, 346–351.

Klenk, H. P. & Göker, M. (2010). En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* **33**, 175–182.

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

Kyrpides, N. C., Hugenholtz, P., Eisen, J. A., Woyke, T., Göker, M., Parker, C. T., Amann, R., Beck, B. J., Chain, P. S. & other authors (2014). Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol* **12**, e1001920.

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

Liu, J. R., Tanner, R. S., Schumann, P., Weiss, N., McKenzie, C. A., Janssen, P. H., Seviour, E. M., Lawson, P. A., Allen, T. D. & Seviour, R. J. (2002). Emended description of the genus *Trichococcus*, description of *Trichococcus collinsii* sp. nov., and reclassification of *Lactosphaera pasteurii* as *Trichococcus pasteurii* comb. nov. and of *Ruminococcus palustris* as *Trichococcus palustris* comb. nov. in the low-G+C gram-positive bacteria. *Int J Syst Evol Microbiol* **52**, 1113–1126.

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J. P., Wang, M., Sun, Y., Wang, L. & Shao, Z. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep* **5**, 14082.

Lowe, T. M. & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955–964.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar., Buchner, A., Lai, T., Steppi, S. & other authors (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**, 1363–1371.

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60.

Meier-Kolthoff, J. P., Klenk, H. P. & Göker, M. (2014). Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* **64**, 352–356.

Miller, L. T. (1982). Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. *J Clin Microbiol* **16**, 584–586.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C. & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**, W182–W185.

Mulligan, C., Fischer, M. & Thomas, G. H. (2011). Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* **35**, 68–86.

Okochi, M., Kurimoto, M., Shimizu, K. & Honda, H. (2007). Increase of organic solvent tolerance by overexpression of manXYZ in *Escherichia coli*. *Appl Microbiol Biotechnol* **73**, 1394–1399.

Petit, E., LaTouf, W. G., Coppi, M. V., Warnick, T. A., Currie, D., Romashko, I., Deshpande, S., Haas, K., Alvelo-Maurosa, J. G. & other authors (2013). Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. *PLoS One* **8**, 1–12.

Pikuta, E. V., Hoover, R. B., Bej, A. K., Marsic, D., Whitman, W. B., Krader, P. E. & Tang, J. (2006). *Trichococcus patagoniensis* sp. nov., a facultative anaerobe that grows at -5 degrees C, isolated from penguin guano in Chilean Patagonia. *Int J Syst Evol Microbiol* **56**, 2055–2062.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590–D596.

Scheff, G., Salcher, O. & Lingens, F. (1984). *Trichococcus flocculiformis* gen. nov. sp. nov. a new gram-positive filamentous bacterium isolated from bulking sludge. *Appl Environ Microbiol* **19**, 114–119.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S. & other authors (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682.

Schink, B. (1984). Fermentation of tartrate enantiomers by anaerobic bacteria, and description of two new species of strict anaerobes, *Ruminococcus pasteurii* and *Ilyobacter tartaricus*. *Arch Microbiol* **139**, 409–414.

Schleifer, K. H. (2009). Classification of *Bacteria* and *Archaea*: past, present and future. *Syst Appl Microbiol* **32**, 533–542.

Stams, A. J., Van Dijk, J. B., Dijkema, C. & Plugge, C. M. (1993). Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl Environ Microbiol* **59**, 1114–1119.

Stams, A. J. M., Huisman, J., Garcia Encina, P. A. & Muyzer, G. (2009). Citric acid wastewater as electron donor for biological sulfate reduction. *Appl Environ Microbiol* **83**, 957–963.

Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C. H. (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288.

The UniProt Consortium (2014). Activities at the universal protein resource (UniProt). *Nucleic Acids Res* **42**, 191–198.

van Gelder, A. H., Aydin, R., Alves, M. M. & Stams, A. J. (2012). 1,3-Propanediol production from glycerol by a newly isolated *Trichococcus* strain. *Microb Biotechnol* **5**, 573–578.

Wang, W., Sun, J., Hartlep, M., Deckwer, W. D. & Zeng, A. P. (2003). Combined use of proteomic analysis and enzyme activity assays for metabolic pathway analysis of glycerol fermentation by *Klebsiella pneumoniae*. *Biotechnol Bioeng* **83**, 525–536.

Yarza, P., Richter, M., Peplies, J., Euzeby, J., Amann, R., Schleifer, K. H., Ludwig, W., Glöckner, F. O. & Rosselló-Móra, R. (2008). The all-species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* **31**, 241–250.

Zerbino, D. R. & Birney, E. (2008). Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* **18**, 821–829.

Zhang, Y. M. & Rock, C. O. (2008). Membrane lipid homeostasis in bacteria. *Nat Microbiol Rev* **6**, 222–233.

Zhilina, T. N., Kotsyurbenko, O. R., Osipov, G. A., Kostrikina, N. A. & Zavarzin, G. A. (1995). *Ruminococcus palustris* sp. nov.: a psychoractive anaerobic organism from a swamp. *Microbiology* **64**, 574–579.