Genomic Insights Into A Novel, Alkalitolerant Nitrogen Fixing Bacteria, *Azonexus sp.* Strain ZS02.

Zohier B. Salah[1], Christopher J. Charles[1], Paul N. Humphreys[1], Andrew. P. Laws[2] and Simon P. Rout[1]*

1. Department of Biological and Geographical Sciences, University of Huddersfield, Queensgate Campus, Huddersfield, United Kingdom, HD1 3DH.
2. Department of Chemical Sciences, University of Huddersfield, Queensgate Campus, Huddersfield, United Kingdom, HD1 3DH.

*__Corresponding author__: Simon Rout, Department of Biological Sciences, University of Huddersfield, Huddersfield, UK, E-mail: S.Rout@hud.ac.uk

Article = Short Research Communication

Abstract

Alkaline environments represent a significant challenge to the growth of micro-organisms. Despite this, there are a number of alkaline environments which contain active microbial communities. Here we describe the genome of a diazotrophic, alkalitolerant strain of *Azonexus,* which was isolated from a microcosm seeded with hyperalkaline soils resulting from lime depositions. The isolate has a genome size 3.60 Mb with 3431 protein coding genes. The proteome indicated the presence of genes associated with the cycling of nitrogen, in particular the fixation of atmospheric nitrogen. Although closely related to *Azonexus hydrophilus* strain d8-1 by both 16S (97.9%) and *in silico* gDNA (84.1%) relatedness, the isolate demonstrates a pH tolerance above that reported for this strain. The proteome contained genes for the complete Na^+/H^+ antiporter (subunits A to G) for cytoplasmic pH regulation; this may account for the phenotypic characteristics of this strain which exhibited optimal growth conditions of pH 9 and 30°C.

Introduction

From a microbiological perspective alkaline environments fall into three broad categories based on their chemistry and origins. By far the most studied alkaline environments are the saline, sodium dominated Soda Lakes [1-3]. There are also calcium and magnesium dominated, alkaline environments generated by serpentinization processes [4] which are associated with groundwater [5, 6] and its subsequent environmental interactions [7-9]. Finally, a range of anthropogenic alkaline systems exist [10]; generated by the waste disposal practices of lime [11], steel [12], soda lime [13] and bauxite [14] manufacture.

Despite the harsh geochemical environment these sites often support extensive and diverse bacterial and archaeal populations [2, 15] capable of a wide range of metabolic and energy generating processes [9, 16]. Such processes have gained interest since they may be exploited

for industrial applications [17]. In the case of nitrogen, the fate of ammonia, nitrate and nitrite has received considerable attention [2, 18]. Nitrogen fixation in alkaline saline systems appears to be associated with free living diazotrophs from the Firmicute division [19-21], although anoxygenic photoautotrophs have also been implicated [22]. In serpentinization associated systems, some authors have reported a lack of nitrogen fixation [8] whilst it has been inferred from community data in others [23].

Our understanding of the nitrogen cycle within anthropogenic lime based environments is currently limited to that of denitrification processes through the measurement of nitrate/nitrite within the soil core [24]. It has also been observed through the amendment of microcosms, reaction vessels with constrained ecological parameters [25], with nitrate as a terminal electron acceptor [26, 27]. Nitrate enrichment of these soils resulted in the detection of *Azoarcus sp.* [26], whilst *Azonexus sp* and *Azovibrio sp* were detected in nitrate free systems [28]. This suggests that these free living diazotrophs play an important role in the nitrogen flow through these systems, potentially replacing the nitrogen fixing *Clostridium sp* [29]. When the studies of both Rizoulis, Milodowski [7] and Kyeremeh, Charles [28] are considered, *Clostridium sp* were below detection limits in the community analyses of these alkaline soil microcosms.

In this paper we provide the whole genome sequence, discussion of the associated proteome and a description of the metabolic capabilities of isolate ZS02, an alkaliphilic nitrogen fixing bacteria. ZS02 was isolated from an anaerobic microcosm.

Materials and Methods

Microcosm operation and characterisation

Hyperalkaline soil was obtained from Harpur Hill, Derbyshire and diluted in a 10 % mix of alkaline hydrolysed cellulose (prepared as per Rout, Radford [30]) and anoxic mineral media

[31]. The microcosm was kept under a nitrogen gas headspace and operated under a 10 % waste / feed ratio every 30 days at pH 9.5 for 12 months. Genomic DNA was then extracted and 16S rDNA library prepared as per Kyeremeh et al [28].

Isolation of pure cultures and phenotypic characterisation.

10 µL of the microcosm fluid was used to inoculate pH 9.5 fastidious anaerobe agar (LabM, UK) with growth observed following 5 days incubation at 25°C in anaerobic conditions ((10 % H_2:10 % CO_2: 80 % N_2; DW Scientific). Single colonies were further purified by sub-culture and cell morphology was determined via Gram staining and SEM (Quanta FEG 250, FEI). A pH and temperature profile was obtained using Bioscreen C technology (Oy Growth Curves). Briefly, fastidious anaerobe broth was prepared and pH values adjusted to 4-12 using 4M NaOH, each pH broth (180µl) was then transferred to duplicate wells and inoculated with 20µl of the isolate prepared to an OD of ~0.2. Manipulations were carried out in an anaerobic chamber and the plate sealed with clear, gas proof tape. Replica plates were prepared and incubated at a range of temperatures (10, 20, 25, 30, 35 and 40 °C), with OD_{600} measured every hour for 24 hours. The pH range and optimal pH for growth were then determined by calculating the fractional area of the growth curves compared with uninoculated control data.

Genome Sequencing, Assembly and Annotation

gDNA from the isolate was extracted using an UltraClean Microbial Isolation Kit. The concentration and purity of the extracted DNA was assessed using a NanoDrop ND1000 spectrophotometer. Sanger sequencing (Eurofins, Germany) of a partial fragment of the 16S rRNA gene coupled to BLASTn search allowed for initial identification of the isolate. For whole genome sequence generation, paired end sequence reads were prepared using an Illumina HiSeq 2500 system (Illumina, US) by Microbes NG (Birmingham, UK). A standard analysis pipeline assigned reads to the closest available reference genome using Kraken [32]

and BWA mem [33]. Further d*e novo* assembly of the reads using SPAdes [34] was also performed and the reads mapped back to the resultant contigs and quality checked using BWA mem. Gaps within the contigs were identified and filled using GapFiller v1.10 [35].

The average nucleotide and average amino acid identities of the genome to near neighbours were calculated using the methods of Rodriguez and Konstantinidis [36]. Annotation of the genome was carried out through RAST [37] and Pfam domain predictions carried out using the NCBI conserved domain database [38]. For the prediction of signal peptides and transmembrane domains, SignalP 4.1 server [39] and the TMHMM server v. 2.0 [40] were used respectively. CRISPRs were found online using the CRISPRfinder tool [41] and metabolic pathways were reconstructed from the annotated proteins using KAAS [42]. Proteins encoding carbohydrate active enzymes identified using the CAZy analysis toolkit [43].

Results and discussions

Community analysis on the microcosm suggested that it was dominated (94.5%) by sequences showing homology to *Azonexus hydrophilus* strain d8-1 [44], indicating that the microcosm was a source of nitrogen fixing bacteria by sequencing of the V4 region of the 16S rRNA gene via MiSeq platform. A pure culture of a facultative anaerobe was obtained, designated ZS02, with morphology of Gram negative rods ~1µm in length (Figure S1). Sanger sequencing generated a 934bp fragment of the 16S rRNA gene which showed 97.9% sequence homology to that of *Azonexus hydrophilus* strain d8-1 (DSM23864) and clustered with the other 16S sequences of this genus (Figure S2). The final assembly of the genome of ZS02 consisted of 3,596,517 bp within 17 scaffolds, representing a coverage of 116x and a GC content of 61.80 % containing no gaps or ambiguous nucleotide positions (Figure 1). *In silico* comparisons of available whole genome sequences of nitrogen fixing bacteria indicated that ZS-02 still showed greatest sequence identity to type strain d8-1 (84.1%, Table 1).

The genome contained 3,494 genes, of which 3,431 were protein coding, 63 were RNA coding, 43 were pseudogenes and 1 CRISPR repeat. 75.3% of the protein-coding genes were assigned a putative function, 2,443 were found to have Pfam domains, 819 transmembrane helicases and 309 signal peptides. When amino acid identities were compared with type strain nitrogen fixing bacteria, again ZS02 showed greatest homology to *Azonexus hydrophilus* d8-1 (84.2%, Table S1).

KAAS assignment indicated that the genome contained a number of nitrogen metabolism genes. Of these, ZS02 proteome includes the most well studied molybdenum dependent nitrogenases encoded by the catalytic n*ifHDK* cluster (Dos Santos, Fang [45]). Within this cluster, *nifH* is highly conserved and this was reflected in the nucleotide sequence homology of the *nifH* of ZS02 with that of other nitrogen fixing bacteria, with greatest homology (96.5%) seen *Azonexus hydrophilus* (DSM23864, Figure 2). In contrast, the remaining genes in the cluster showed greater homology to other organisms, with *nifD* nucleotide sequence showing greatest homology to *Zoogloea* sp LCSB751 (94.2%) and *Azovibrio restrictus* DSM23866 (93.1%). Comparisons of the *Nif*K resulted in homologies >90%, with greatest homology to *Canididatus accumulibacter* (87.9%) and *Azovibrio restrictus* (82.6%).

The proteome of ZS02 contained a complete *napAB* gene cluster encoding periplasmic nitrate reductase [46], with the presence of *nirBD* cluster suggesting that nitrite can then be further reduced to ammonia within the cytoplasm [47]. Although genes encoding NasAB were present, the absence of *nirA* suggests ZS02 is incapable of assimilatory nitrate reduction. The absence of *nirK* within the proteome also suggests that conversion of nitrite to nitric oxide is not a metabolic capability, however the presence of *norBC* and *nosZ* suggests that ZS02 is capable of the conversion of nitric oxide to molecular nitrogen [47]. The proteome also suggest that ZS02 is not capable of ammonium oxidation or nitrification of ammonia.

The presence of a complete Na^+/H^+ antiporter (subunits A to G) are a common feature of alkaliphilic microorganisms relating to cytoplasmic pH regulation [48], and were present in the proteome of ZS02. pH/temperature profiles (Figure 3) indicated that there was no growth above a pH of 10 or below 5, with the limit of temperature appearing to be 40°C. Optimum growth was observed at a temperature of 30°C and a pH of 9. Although moderate growth was observed at this temperature between a pH of 6 and 10. CAZy analysis indicated a number of carbohydrate active enzymes within the proteome of ZS02, and a greater number than the type strain (Figure S3). These increased numbers of enzymes may provide an advantage in harsh alkaline environments. Nutrient cycling in natural and anthropogenic alkaline environments requires micro-organisms that are capable of tolerating the harsh pH conditions. Since nitrate is limited in these environments, micro-organisms capable of utilising atmospheric nitrogen may play a pivotal role in delivering nitrogen. Here we present the whole genome sequence of an alkalitolerant *Azonexus hydrophilus* strain, containing nitrogen cycling abilities within its proteome.

References

1.   Grant WD. Alkaline Environments and Biodiversity. Extremophiles (Life Under Extreme Environmental Conditions): Encyclopedia of Life Support Systems (EOLSS), Unesco; 2007.
2.   Sorokin DY, Banciu HL, Muyzer G. Functional microbiology of soda lakes. Curr Opin Microbiol. 2015; 25. 88-96.
3.   Grant WD, Jones BE. Bacteria, archaea and viruses of Soda Lakes. Soda Lakes of East Africa: Springer; 2016. p. 97-147.
4.   Müntener O. Serpentine and serpentinization: A link between planet formation and life. Geology. 2010; 38. 959-60.
5.   Tiago I, Chung AP, Veríssimo A. Bacterial Diversity in a Nonsaline Alkaline Environment: Heterotrophic Aerobic Populations. Appl Environ Microbiol. 2004; 70. 7378-87.
6.   Veríssimo A, Tiago I. Bacterial Diversity in a Nonsaline Alkaline Environment. In: Nelson EK, editor. Encyclopedia of Metagenomics. New York, NY: Springer New York; 2013. p. 1-5.
7.   Rizoulis A, Milodowski AE, Morris K, Lloyd JR. Bacterial diversity in the hyperalkaline Allas Springs (Cyprus), a natural analogue for cementitious radioactive waste repository. Geomicrobiol J. 2016; 33. 73-84.

8.   Meyer-Dombard DAR, Woycheese KM, Yargıçoğlu EN, Cardace D, Shock EL, Güleçal-Pektas Y, et al. High pH microbial ecosystems in a newly discovered, ephemeral, serpentinizing fluid seep at Yanartaş (Chimera), Turkey. Front Microbiol. 2014; 5. 723.

9.   Cardace D, Meyer-Dombard DAR, Woycheese KM, Arcilla CA. Feasible metabolisms in high pH springs of the Philippines. Front Microbiol. 2015; 6. 10.

10.   Gomes HI, Mayes WM, Rogerson M, Stewart DI, Burke IT. Alkaline residues and the environment: a review of impacts, management practices and opportunities. J Clean Prod. 2016; 112. 3571-82.

11.   Milodowski AE, Shaw RP, Stewart DI. The Harpur Hill Site: its geology, evolutionary history and a catalogue of materials present British Geological Survey Commissioned Report CR/13/104 43pp, British Geological Survey, Keyworth, Nottinghamshire, UK; 2013.

12.   Riley AL, Mayes WM. Long-term evolution of highly alkaline steel slag drainage waters. Environ Monit Assess. 2015; 187. 463.

13.   Kalwasińska A, Felföldi T, Szabó A, Deja-Sikora E, Kosobucki P, Walczak M. Microbial communities associated with the anthropogenic, highly alkaline environment of a saline soda lime, Poland. Antonie Van Leeuwenhoek. 2017; 110. 945-62.

14.   Bray AW, Stewart DI, Courtney R, Rout SP, Humphreys PN, Mayes WM, et al. Sustained Bauxite Residue Rehabilitation with Gypsum and Organic Matter 16 years after Initial Treatment. Environ Sci Technol. 2018; 52. 152-61.

15.   Sorokin DY, Berben T, Melton ED, Overmars L, Vavourakis CD, Muyzer G. Microbial diversity and biogeochemical cycling in soda lakes. Extremophiles. 2014; 18. 791-809.

16.   Sorokin DY, Kuenen JG. Chemolithotrophic haloalkaliphiles from soda lakes. FEMS Microbiol Ecol. 2005; 52. 287-95.

17.   Borkar S. Alkaliphilic bacteria: diversity, physiology and industrial applications. Bioprospects of Coastal Eubacteria: Springer; 2015. p. 59-83.

18.   Dendooven L, Alcántara-Hernández RJ, Valenzuela-Encinas C, Luna-Guido M, Perez-Guevara F, Marsch R. Dynamics of carbon and nitrogen in an extreme alkaline saline soil: A review. Soil Biol Biochem. 2010; 42. 865-77.

19.   Zhilina TN, Kevbrin VV, Tourova TP, Lysenko AM, Kostrikina NA, Zavarzin GA. Clostridium alkalicellum sp. nov., an Obligately Alkaliphilic Cellulolytic Bacterium from a Soda Lake in the Baikal Region. Microbiology. 2005; 74. 557-66.

20.   Sorokin ID, Kravchenko IK, Doroshenko EV, Boulygina ES, Zadorina EV, Tourova TP, et al. Haloalkaliphilic diazotrophs in soda solonchak soils. FEMS Microbiol Ecol. 2008; 65. 425-33.

21.   Navarro-Noya YE, Luna-Guido M, Dendooven L. Cultivable Nitrogen Fixing Bacteria from Extremely Alkaline-Saline Soils. Adv Microbiol. 2016; 6. 412-3.

22.   Gorlenko VM, Bryantseva IA, Rabold S, Tourova TP, Rubtsova D, Smirnova E, et al. Ectothiorhodospira variabilis sp. nov., an alkaliphilic and halophilic purple sulfur bacterium from soda lakes. Int J Syst Evol Microbiol. 2009; 59. 658-64.

23.   Woycheese KM, Meyer-Dombard DAR, Cardace D, Argayosa AM, Arcilla CA. Out of the dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing seep (Manleluag, Pangasinan, the Philippines). Front Microbiol. 2015; 6.

24.   Burke I, Mortimer R, Palani S, Whittleston R, Lockwood C, Ashley D, et al. Biogeochemical reduction processes in a hyper-alkaline affected leachate soil profile. Geomicrobiol J. 2012; 29. 769-79.

25.   Barra Caracciolo A, Bottoni P, Grenni P. Microcosm studies to evaluate microbial potential to degrade pollutants in soil and water ecosystems. Microchem J. 2013; 107. 126-30.

26.   Bassil NM, Bryan N, Lloyd JR. Microbial degradation of isosaccharinic acid at high pH. The ISME Journal. 2014; 9. 310-20.

27.   Grant WD, Holtom GJ, Kelly NO, Malpass J, Rosevear A, Watkiss P, et al. Microbial Degradation of Cellulose derived Complexants Under Repository Conditions Harwell, Didcot, Oxfordshire, UK; 2002.

28.   Kyeremeh IA, Charles CJ, Rout SP, Laws AP, Humphreys PN. Microbial Community Evolution Is Significantly Impacted by the Use of Calcium Isosaccharinic Acid as an Analogue for the Products of Alkaline Cellulose Degradation. PloS one. 2016; 11. e0165832.

29.   Chen JS. Genomic Aspects of Nitrogen Fixation in the Clostridia. In: Palacios R, Newton WE, editors. Genomes and Genomics of Nitrogen-fixing Organisms. Dordrecht: Springer Netherlands; 2005. p. 13-26.

30.   Rout SP, Radford J, Laws AP, Sweeney F, Elmekawy A, Gillie LJ, et al. Biodegradation of the alkaline cellulose degradation products generated during radioactive waste disposal. PloS one. 2014; 9. e107433.

31.   B.S.I. BS ISO 14853:2005 Plastics. Determination of the ultimate anaerobic biodegradation of plastic materials in an aqueous system. Method by measurement of biogas production. London; 2005.

32.   Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014; 15. R46-R.

33.   Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997 [q-bioGN]. 2013.

34.   Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012; 19. 455-77.

35.   Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. Genome Biol. 2012; 13. R56.

36.   Rodriguez LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. Microbe. 2014; 9. 111-8.

37.   Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014; 42. D206-D14.

38.   Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic Acids Res. 2014. gku1221.

39.   Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods. 2011; 8. 785-6.

40.   Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001; 305. 567-80.

41.   Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007; 35. W52-W7.

42.   Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 2007; 35. W182-W5.

43.   Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology. 2010; 20. 1574-84.

44.   Chou J-H, Jiang S-R, Cho J-C, Song J, Lin M-C, Chen W-M. Azonexus hydrophilus sp. nov., a nifH gene-harbouring bacterium isolated from freshwater. Int J Syst Evol Microbiol. 2008; 58. 946-51.

45.   Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. BMC Genomics. 2012; 13. 162.

46.   Coelho C, Romão MJ. Structural and mechanistic insights on nitrate reductases. Protein Science. 2015; 24. 1901-11.

47.   Bedmar E, Bueno E, Correa D, Torres M, Delgado M, Mesa S. Ecology of denitrification in soils and plant-associated bacteria: CRC Press; 2016.

48.   Hamamoto T, Hashimoto M, Hino M, Kitada M, Seto Y, Kudo T, et al. Characterization of a gene responsible for the Na+/H+ antiporter system of alkalophilic Bacillus species strain C-125. Mol Microbiol. 1994; 14. 939-46.

Tables and Figures

Organism	Average Nucleotide Identity (%)							
	1	2	3	4	5	6	7	8
Azonexus hydrophilus ZS-02 (1)	100	84.1	81.1	79.0	77.6	77.2	76.0	75.6
Azonexus hydrophilus DSM23864 (2)		100	81.0	79.0	77.4	78.0	75.4	75.6
Dechloromonas denitrificans ATCC BAA-841 (3)			100	78.1	77.4	77.4	76.5	75.6
Azovibrio restrictus DSM23866 (4)				100	77.6	77.7	76.6	74.7
Azoarcus sp BH72 (5)					100	78.8	75.6	74.9
Zoogloea sp LCSB751 (6)						100	75.4	75.4
Aquaspirillium sp LM1 (7)							100	73.4
Candidatus accumulibacter sp SK-02 (8)								100

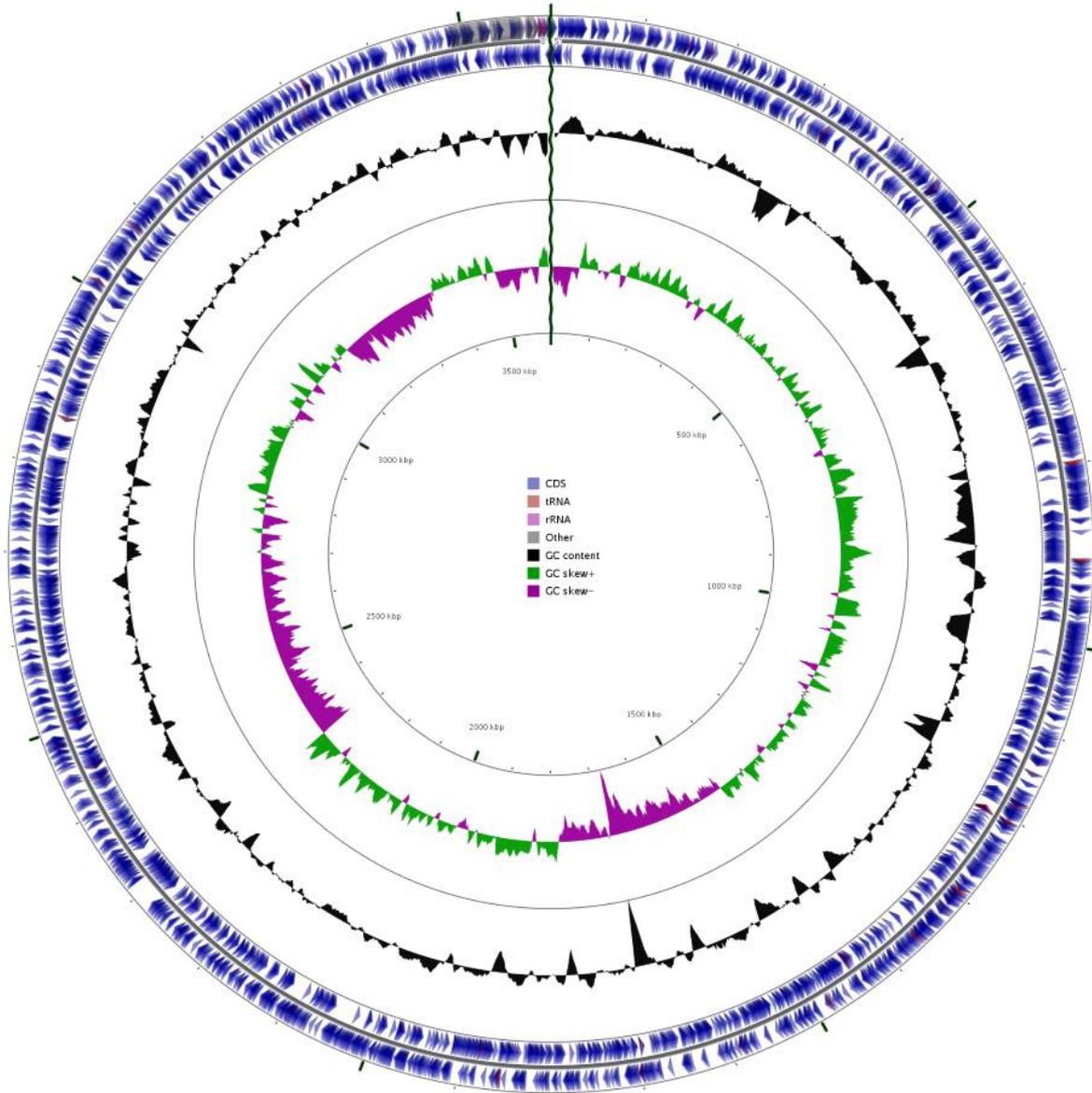Table 1: Average nucleotide identity comparisons between nitrogen fixing type strains.

Figure 1: Graphical representation of the features of the genome of isolate ZS-02, Circles (from outside to inside) 1 and 2 (protein) coding sequences on the forward and reverse strand, blue indicates CDS, peach shows tRNAs, pink rRNAs, and grey, other). Circle 3 shows GC content% and Circles 4 and 5 show positive (green) and negative (magenta) GC skew.

Organism	Amino acid sequence similarity to ZS-02 (%)		
	NifH	NifD	NifK
Azovibrio restrictus DSM23866	93.4	93.1	82.6
Aquaspirillium sp LM1	95.1	81.5	77.6
Candidatus accumulibacter sp SK-02	92.7	82.9	87.9
Azoarcus sp BH72	89.5	82.4	74.3
Zoogloea sp LCSB751	91.4	94.2	76.8
Dechloromonas denitrificans ATCC BAA-841	95.6	83.3	74.9
Azonexus hydrophilus ZS-02	100.0	100.0	100.0
Azonexus hydrophilus DSM23864	96.5	81.6	76.6

Figure 2: Amino acid similarities of *Azonexus hydrophilus* ZS02 NifHDK cluster. Estimate of phylogeny based on whole genome sequences using neighbour joining method is also shown (branch lengths not to scale).
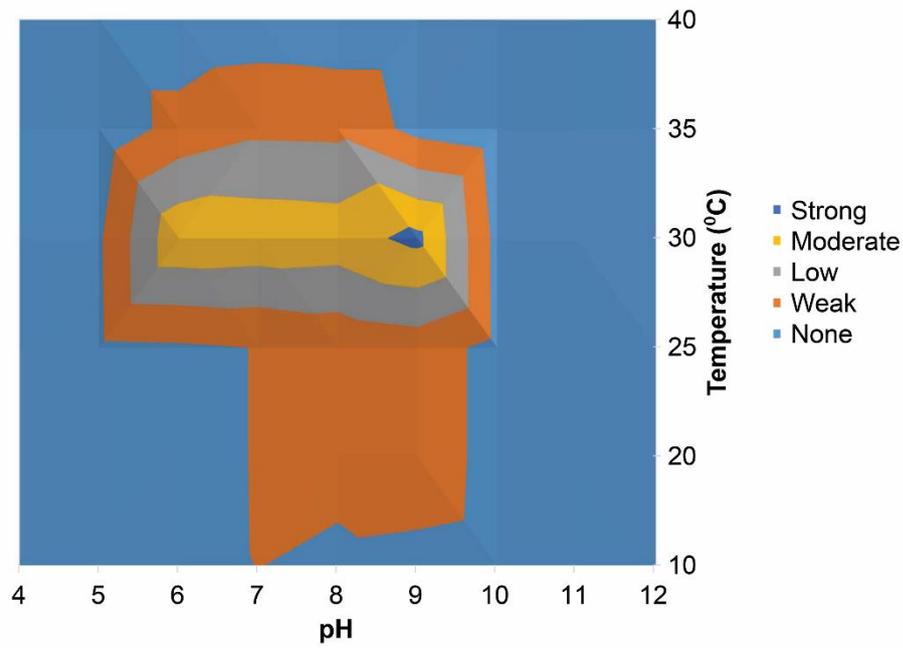
Figure 3: pH/temperature profile of isolate ZS02 indicated the isolate was capable of growth between pH 5 and 10, with an optimal pH for growth of 9 at 30°C