



Comparative Genome Analysis of the *nirK*-carrying *Azospirillum* sp. Strains

Jeonghwan Jang¹, Yoriko Sakai², Masahito Hayatsu², Keishi Senoo³, and Satoshi Ishii^{1,3,4}



¹BioTechnology Institute, University of Minnesota, St. Paul, MN, ²National Institute for Agro-Environmental Sciences, Tsukuba, Japan,

³Department of Applied Biological Chemistry, Tokyo, Japan, and ⁴Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN

Introduction

Azospirillum sp. strains are commonly found denitrifiers in soils.

Under O₂-limiting conditions, denitrifiers use nitrate as a respiratory electron acceptor. Nitrate is sequentially reduced to N₂ by denitrification (NO₃⁻ → NO₂⁻ → NO → N₂O → N₂).

The nitrite reductase, which catalyze the reduction of NO₂⁻ to NO, is considered the key enzyme for denitrification. There are two types of nitrite reductases: NirK and NirS, which contain Cu and cytochrome *cd*₁ as the reaction centers, respectively.

Previously, we isolated 41 *Azospirillum* sp. denitrifiers from a rice paddy soil in Tokyo, Japan. However, the *nir* gene could not be amplified by PCR from most of the *Azospirillum* strains probably due to the base mismatches of the PCR primers to the target gene sequences (Ishii *et al.* 2011).

Complete genome sequence of *Azospirillum* sp. B510 (Kaneko *et al.* 2010) suggested that this strain possesses *nirK* on a 681,723-bp plasmid named pAB510c (GenBank: AP010949), which is one of the seven replicons of the *Azospirillum* sp. B510 genome.

The objectives of this study are to:

- 1) Identify the *nirK* sequences of the 41 *Azospirillum* strains by using newly designed PCR primers;
- 2) Compare the *nirK* phylogeny with the 16S rRNA gene phylogeny to obtain the idea of the potential genetic exchange among the strains;
- 3) Identify the localization of *nirK* in the genomes of *Azospirillum* sp. strains by using pulsed-field gel electrophoresis (PFGE) followed by southern-blot hybridization;
- 4) Identify the denitrification functional gene clusters among *Azospirillum* strains by comparative genomics using the Illumina HiSeq system.

1. PCR primers for *nirK* in *Azospirillum* spp.

Table 1. Newly designed PCR primers for the amplification of *nirK* from the genus *Azospirillum*

name	sequence (5'→3')	amplicon size by pairing with B510_nirK_247F primer
B510_nirK_247F	ACACCTAYTGACSTTCAAC	-
B510_nirK_479R	CAGTGATAGACTAGACGCC	213
B510_nirK_598R	TAGATTCTGCTGCATCAC	352
B510_nirK_818R	GACSTTGTGGAAGATCTCGC	572
B510_nirK_918R	GATARTCSACCTGGAAGTCC	672

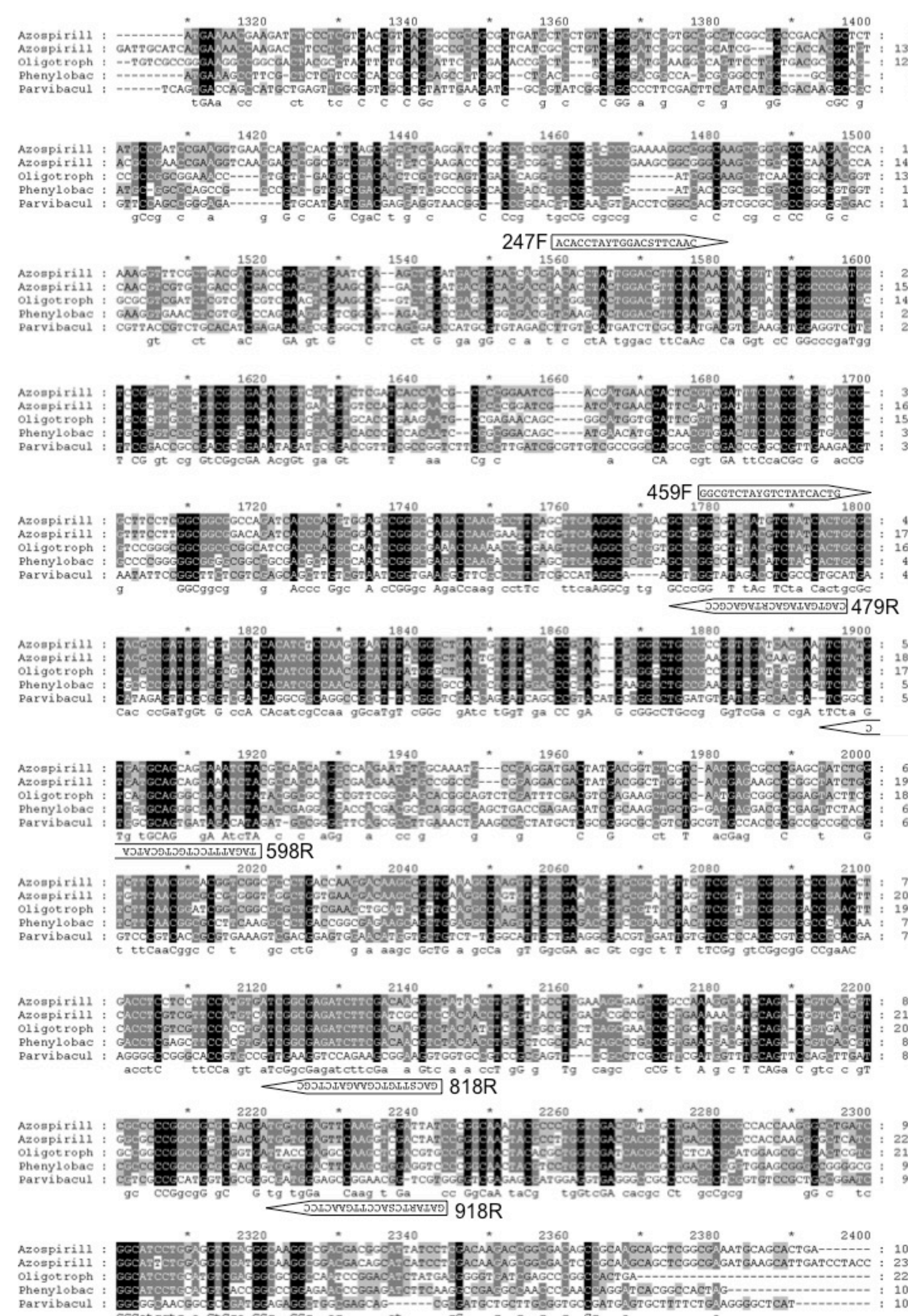


Figure 1. The annealing sites of the newly designed PCR primers on the *nirK* gene

➔ New *nirK* gene PCR primer sets were designed on conserved regions of the *nirK* from *Azospirillum*.

2. Phylogeny of *Azospirillum* sp. strains based on the 16S rRNA gene and *nirK* sequences

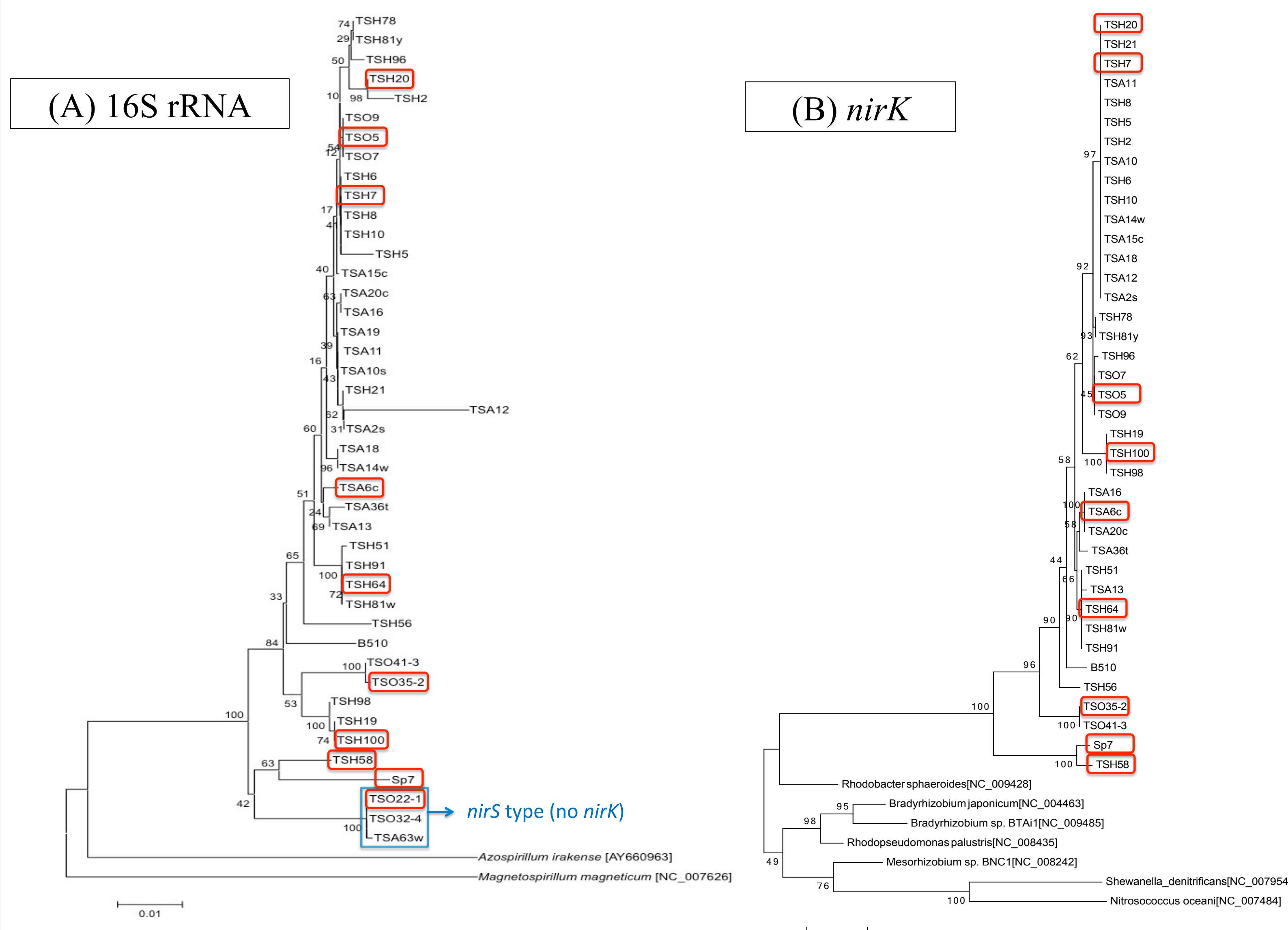


Figure 2. Phylogenetic trees constructed based on (A) 16S rRNA gene and (B) *nirK* sequences using the maximum likelihood method with bootstrap analysis. Strains marked as red were selected for genome sequencing.

- ➔ Sequence analysis of the *nirK* revealed the diversity within the gene of the genus *Azospirillum*
- ➔ The *nirK* phylogeny of many *Azospirillum* sp. strains did not match well with the 16S rRNA gene phylogeny, suggesting the possible occurrences of horizontal transfer of *nirK*.

3. Replicons of *Azospirillum* sp. strains containing *nirK* gene

Table 2. Number and size of replicons for each *Azospirillum* sp. strain based on pulsed-field gel electrophoresis. Replicons containing the *nirK* gene was determined using the southern-blot hybridization method.

strain ID	replicon size (Kbp)										
TSH2	*877	765	716	575	467	403	12				
TSH5	*857	770	703	581	442	403	11				
TSH6	*880	776	586	476	412	12					
TSH7	*862	764	713	570	476	437	397	12			
TSH8	*876	774	713	577	468	406	12				
TSH10	*868	768	716	583	450	405	12				
TSH19	*866	579	525	444	305	12					
TSH20	*866	766	716	579	463	399	11				
TSH21	*846	782	674	555	440	288	11				
TSH51	962	920	*847	761	484	379	273				
TSH56	922	832	*640	565	435	283	12				
TSH58	805	*642	549	383	168	68					
TSH64	951	905	*837	749	478	367	267	11			
TSH78	*935	761	693	573	536	427	322	283	241	193	7
TSH81w	959	914	*846	769	490	381	130	11			
TSH81y	*941	772	719	584	544	434	329	285	235	138	118
TSH91	1031	908	*840	757	486	380	283	11			
TSH96	*847	752	690	618	484	442	314				
TSH98	*863	576	521	491	307	11					
TSH100	*874	571	524	480	441	296	11				
TSA2s	949	*887	664	607	411						
TSA6c	*846	606	512	295	146	83					
TSA10	931	*876	775	547	452	332					
TSA13	981	*921	660	494	392	288	13				
TSA15c	926	*880	785	556	452	338					
TSA16	*819	585	497	278	66						
TSA19	899	858	756	552	458	341	302	23			
TSA20c	*800	583	496	290							
TSA36t	949	*876	616	441	286	10					
TSO5	*893	698	585	533	432	291	66				
TSO7	*915	720	606	553	449	303	8				
TSO9	*907	717	615	555	451	302					
TSO35-2	*850	544	477	361	9						
TSO41-3	874	561	488	373	10						
Sp7	794	*697	621	177	132						
B510	703	*662	612	521	243						

*: Replicon which contains the *nirK* gene

- ➔ Multiple replicons were identified in all *Azospirillum* sp. strains tested
- ➔ The *nirK* gene was identified in large replicons (640-950 kbp)

4. Comparative genome analysis

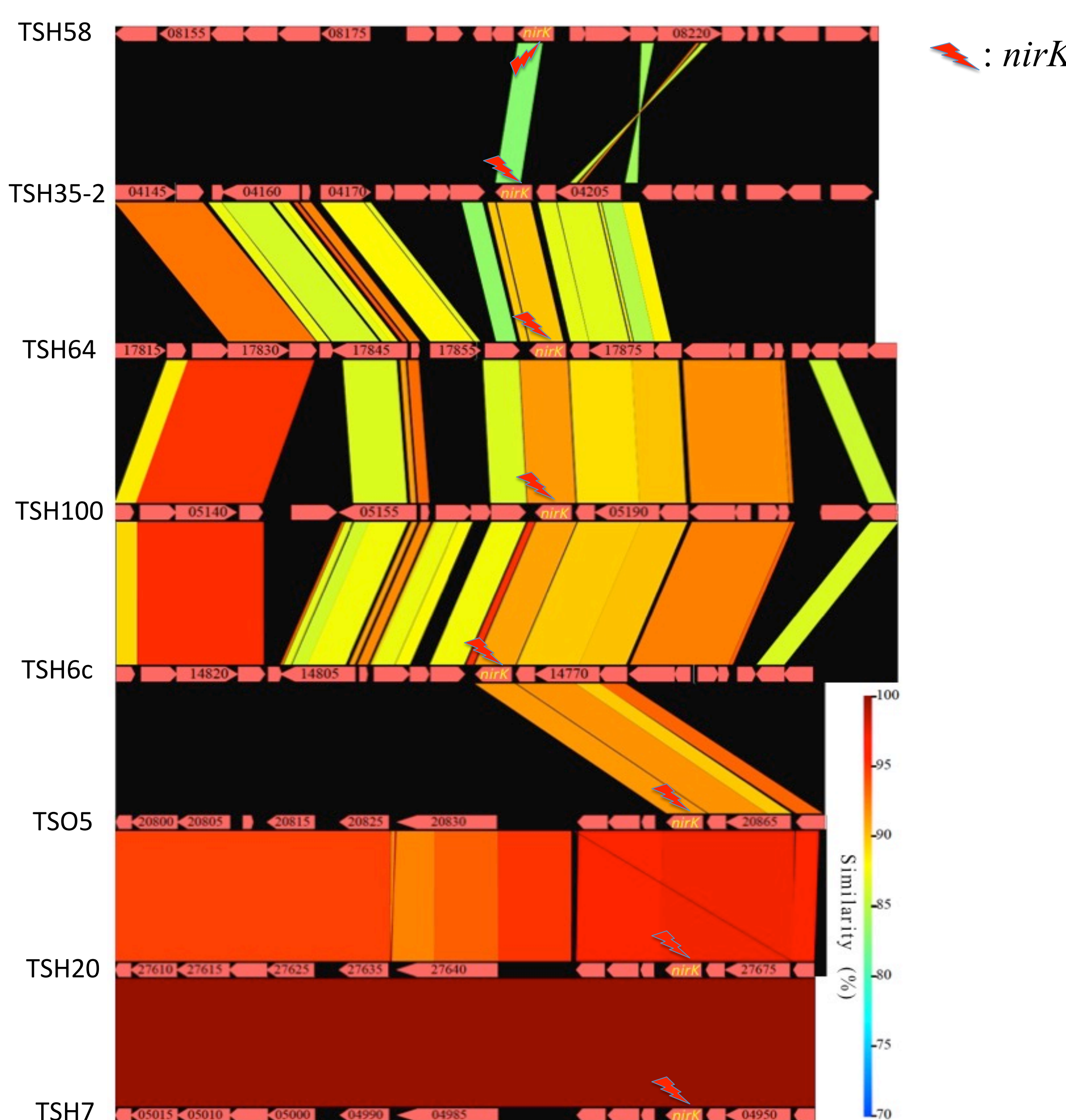


Figure 3. Comparison of the gene clusters around *nirK* among the *Azospirillum* sp. strains. Heatmap was generated using GenomeMatcher v2.03.

Table 3. Summary of the genome analysis of the ten representative *Azospirillum* sp. strains.

	Sp7	TSA6c	TSH7	TSH20	TSH58	TSH64	TSH100	TSO5	TSO22-1	TSO35-2
Total sequences (bp)	19,227,590	16,551,944	18,252,398	29,688,068	29,027,348	23,908,290	20,323,848	22,682,774	19,647,918	18,737,394
Number of contigs	704	395	810	762	654	245	406	616	752	201
N ₅₀ (bp)	84,704	7,635,475	57,847	61,723	67,868	7,543,664	119,003	62,252	60,411	1,296,767
Max length of contig (bp)	368,253	7,635,475	342,998	366,555	303,391	7,543,664	392,776	238,495	222,401	3,068,875
Total length of contigs (bp)	7,155,803	8,376,724	7,506,927	7,509,079	7,039,278	7,930,345	7,183,287	7,041,570	7,415,165	7,369,390
Genome coverage (fold)	271	219	245	399	416	320	285	325	267	256
CDS	6,005	6,929	6,344	6,358	5,883	6,536	6,021	5,954	6,299	6,067
tRNAs	69	72	64	62	67	66	61	62	53	72
rRNA operons	2	19	4	3	2	13	1	2	2	13
GenBank accession	SAMN3877357	SAMN3863722	SAMN3863725	SAMN3863731	SAMN3863732	SAMN3863780	SAMN3863786	SAMN3863787	SAMN3863794	SAMN3863795

➔ Strains with high *nirK* similarity tend to have similar gene organization patterns.

Conclusions

- The *nirK* sequence was diverse among *Azospirillum* sp. strains.
- The *nirK* might have been introduced via horizontal gene transfer. We need to look for more direct evidence for this through the analysis of the *Azospirillum* genomes.
- All *Azospirillum* sp. strains tested in this study had multiple replicons, and the *nirK* was observed in one of the large size replicons (640-950 kbp)

References

Ishii *et al.* 2011. *Appl Environ Microbiol.* 77:338-342
Kaneko *et al.* 2010. *DNA Res.* 17:37-50.