

University of Minnesota

• Professional Training • Research • Resources for Industry

<sup>1</sup>BioTechnology Institute, University of Minnesota, St. Paul, MN, <sup>2</sup>National Institute for Agro-Environmental Sciences, Tsukuba, Japan, <sup>3</sup>Department of Applied Biological Chemistry, Tokyo, Japan, and <sup>4</sup>Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN

Introduction

Azospirillum sp. strains are commonly found denitrifiers in soils.

Under  $O_2$ -limiting conditions, denitrifiers use nitrate as a respiratory electron acceptor. Nitrate is sequentially reduced to N<sub>2</sub> by denitrification (NO<sub>3</sub><sup>-</sup>  $\rightarrow$  $NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2).$ 

The nitrite reductase, which catalyze the reduction of  $NO_2^-$  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 southernblot 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* 

<b>I</b>			
name	sequence (5'>3')	amplicon size by pairing with B510_nirK_247F primer	
B510 nirK 247F	ACACCTAYTGGACSTTCAAC	-	
B510 nirK 479R	CAGTGATAGACRTAGACGCC	213	
B510 nirK 598R		352	
B510 nirK 818R		572	
	GATARTCSACCTTGAACTCC	672	
Azospirill : GATTGCATCA. Oligotroph :TGTCGCCGGG	ARAGESTICS-ETCICTCGSCACCGSCEAGGCSTEGCSCTFAC ACCAGECATECIGAGITCGGSGIGCSCETATIGAAGAISGCEGI Aa cc ct tc C C C Gc c G C g	CTGTCCGCGCGTCCGCGCGCGCGCGACAGGTCT : 91 CTGTCCGCGCGCGCGCGCGCGCGCGCGGCGGCGGCGGCGG	
Azospirill : ATOCATOGA Azospirill : ACOCAAOGA Oligotroph : CCOCOGCGGAA Phenylobac : ATOCOGCGGA Parvibacul : GTTCAGCOGG gCcg c	a g G c G CgaC t g c C C g t g	0 * 1480 * 1500 Constrained for the second	
Azospirill : AAAG TTCG Azospirill : CAAC CGTG Oligotroph : GCGC CGAT Phenylobac : GAAG GAAC Parvibacul : CGTTACCGTCTG gt ct	CACATEGAGAGAGECGEGGGCTEGTCAGEGAGCCATGCGTGTAGAGCT aC GA gt G C ct G ga gG c a t	CACOTATINGGA CITICAACACACCENTOCOGOCCCATER : 289 CACOTACINGGA GITCAACAACAACACENTOCOGOCCCATER : 1590 GGOTACINGGA GITCAACAGOAACAGOTACOGOCCCATEC : 1465 CAAGIACINGGA CITICAACAGOAACCIGOCOGOCCCATEC : 265 IGT CATGATCTCGCCGATGACGTGAAGTGAAGTGAAGTGA	
Azospirill : CCCCCCCCCCC Azospirill : CCCCCCCCCC Oligotroph : GCCCCCCCCCC Phenylobac : GCCCCCCCCC Parvibacul : TCCGACCCCCC 7 CG gt cg G	COP-CACACHERGALCE GTCCARGACOLLCG - COCCCGGATCG-	0 * 1680 * 1700 ACCATGAACCACTCCACGATTTCCACGCCCCACCG-: 382 ATCATGAACCATTCCATGATTTCCACGCCCCCACCG-: 1683 GCCATGCTCCATGGTCGACTTCCACGCCCCCCCCCCC	
Azospirill : CTTCCTCGGC Azospirill : TTTCCTTGGC Oligotroph : TCCGGGCGGC Phenylobac : CCCCGGGGGGGC Parvibacul : AATATTCCGGC g GGCg	1720 * 1740 * 176 GCCGCCA CATCACCCACCTCCACCCCGCCCACACCAACGCCCCACC GCCGCCACATCACCCACCCCGCCCGCCCACACCAACGAACCCCACC GCCCGCCACCCAC	0 * 1780 * 1800 TOAGGO OF ACCORDING CONTACTOR TATACTOR : 482 AGGO OF ACCORDING CONTACTOR TATACTOR : 1783 TOAGGO OF GTOCOGO OF CONTACTOR OF TATACTOR : 1658 TOAGGO OF CASC OF CONTACTOR OF CONTACTOR : 479 TCAAGGO T GTOCOGO OF CONTACTOR : 479 TCAAGGO T GCCCGG T TAC TCTA CACTOR : 479 TCAAGGO T GCCCGG T TAC TCTA CACTOR : 479 TCCAAGGO T GCCCGG T TAC TCTA CACTOR : 470 TCCAAGGO T GCCCGG T TAC TCTA CACTOR : 470 TCCAAGGO T GCCCGG T TAC TCTAAGGO TCTAAGG	
Azospirill : CACGOCATOG Azospirill : CACGOCATOG Oligotroph : CACGOCATOG Phenylobac : CGCCCCATOG Parvibacul : CATAGATTCCC Cac ccGaTgGt	1820 • 1840 • 186 CET CALACTERIA CEANER ACCOUNT ACCOUNT AND CALACTER ACCOUNT ACTION ACCOUNT AC	CONA	
	AAAL TACG CALCAL COGAL GAAD TOCCGCCOCGFAGGAC ACAL TATA GGIGC FCCGTTCGGUA CACGGCAGTCTCFATTC ACAL TACA CGAGG FACCOCGA GCCCAGGGCAGCTGACCGAG ACALAGAT GLOGGC TTC GCG TTGAAACT AAGCCCTATG A AtCTA C C AGg C CC g g g LLLY9VI 598R	GACTATCACCGTTCCACCACCACCGCCCGAGCTATCTCC : 676 GACTATCACCGCTTGCTC-AACCACGACCGCCGAGCTATCTCC : 1977 GACGTCCAGAAGCTGCTC-AATGACCGCCGGGGGAGTACTTCC : 1855 AGCATCCGCAAGCTGCTG-GACGACGACGCCGAGTTCTACC : 655 CTGCCCGGCGCCGTCTGCGTCCCCACCGGGCCGCCGCCG : 674 C G ct T acGag C t G	
Azospirill : CTTCAACGGCA Azospirill : GTTCAACGGCA Oligotroph : CTTCAACGGAT Phenylobac : CTTCAACGGAT Parvibacul : GTCCGTCACGGC t ttCaaCggc	2020 * 2040 * 206 GG CGGC CC CACCAACC CAACCACT CAACAC CT CA CC CGGC CGC CGTCAACCACT CAACCACT CAACAC CT CC GG CGGC CGC CT CAACGGC CAACCT CC T CC CT CAAGG CC CACCGGC CAACACT CC CC CAGG CC CA CT CAAGG CC CACCGACT CAACACT CC CC CC CACCCC CA CT CAAGG CC CACCGACT CAACACT CC CC CC CACCCC CA CT CAAGG CC CACCGACT CAACACT CC CC CC CACCCC CACCCCC CT CAAGG CC CC CACCGACT CC C	CALLER CONTENTED CONTENTED CONTENTED CONTENTS IN THE TRANSPORTED FOR THE CONTENT OF CONTENTS IN THE CONTENT OF CONTENT OF CONTENTS IN THE CONTENT OF CONTENT. TOTAL OF CONTENT OF CONTENT. OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT. TOTAL OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT. TOTAL OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT OF CONTENT. OF CONTENT OF CONTENT OF CONTENT OF CONTENT. OF CONTENT OF CONTENT OF CONTENT OF CONTENT. OF CONTENT OF CONTENT OF CO	
Azospirill : GACCTOCTO Azospirill : GACCTOCTO Oligotroph : GACCTOGTOG Phenylobac : GACCTOGAGCA Parvibacul : AGGGGCCGGGCA acctC tt	2120 2140 216 CCATOT CATOROGOGAGATOTTIC ACAAGETOTATACON GOTTICO CCATOT CATOROGOGAGATOTTIC ACAAGETOTATACON GOTTICO CCACC, GATOROGOGAGATOTTIC ACAAGETOTACAAT, COECEGOG CCACT, GATOROGOGAGATOTTIC ACAAGETOTACAACT GOTTICO CCACT, GATOROGOGAGATOTTIC ACAACETOTACAACT GOTTICO CCACT, GATOROGOGAGATOTTICO ACAACETOTACAACT GOTTICO CCACT, GATOROGOGOGAGATOTTICO ACAACETOTACAACT GOTTICO CCACT, GATOROGOGOGAGATOTTICO ACAACETOTACAACT GOTTICO CCACT, GATOROGOGOGAGATOTTICO ACAACETOTACAACT GOTTICO CCACT, GATOROGOGOGAGATOTTICO ACAACETOTACAACT GOTTICO CCACT, GATOROGOGOGOGOGOGOGOGOGOGOGOGOGOGOGOGOGOGO	0 * 2180 * 2200 TEGAALCE GAGEGGGCEARA GEARCEAGA-COSTCACCT : 875 TEGALE GGCGEGGGCEARA GEARCEAGA-COSTCACCT : 2176 TEGALE COCCEGGGEGGAAGEA GAGEAGGGCAGA-COSTCACCT : 2054 TEAC CASE COCCEGGGGAAGEAGGGCAGA-COSTCACCT : 854 TEAC CASE COCCEGGGGAAGEAGGACGAGACCAGA-COSTCACCT : 854 TEAC CASE COCCEGGGGAAGEAGGACGAGACCAGA-COSTCACCT : 854 TEAC CASE COCCEGGGGAAGEAGGACGAGAGTTCAAGATTCAAGTTGAAT : 870 Tg cage cCG t A g c T CAGa C gtc c gT	
* Azospirill : Crececco Azospirill : Grececco Oligotroph : Geeccecco Phenylobac : Crececce Parvibacul : Crececce ge CCGgegG	2220 * 2240 * 226 CACAC A CONSTRUCT OF CONTRAT CONSTRUCT CONSTRUCT OF CONTRAT CONSTRUCT CONSTRUCT OF CONSTRUCT OF CONSTRUCT CACCONSTRUCT OF CONSTRUCT OF CONSTRUCT OF CONSTRUCT CACCONSTRUCT OF CONSTRUCT OF CONSTRUCT OF CONSTRUCT CONSTRUCT OF CONSTRUCT	0 * 2280 * 2300 T GA GA CACCAAG C TGAT : 975 T GA CACCAAG GC TCAT : 2276 A CTCAC ATGAGCGC GA TCGT : 2154 GA GT GGGGGC CCGC T GGGGGGG C GGGGGGC : 954 GA GT GGGGC CCGC T GGGGGC C GGGGGCC : 969 tgGtcGA cacgc Ct gcCgcg gG c tc	
Azospirill : GGCATCCTCGAG Azospirill : GGCATCCTCGAG Oligotroph : GGCATCCTCCAT Phenylobac : GGCATCCTCCAC Parvibacul : GGCGGCAACGGC GGCatcctg a	GEORATER AARE GGALGACAGCA, CATCCT CACAA AGEGGEGA SECHAGER GCERICAATCCGGACA, CTATGACEGGGT FAT GAGCC	AGC GCALGCAGCTCGGCGLAATGCAGCACTGA : 1068 TCC GCALGCAGCTCGGCGLGATGAAGCATTGATCCTACC : 2376	
Figure 1. The a the <i>nirK</i> gene	annealing sites of the newly	designed PCR primers on	
New <i>i</i>	nirK gene PCR pri	mer sets were design	ed
	nserved regions of	e	
	e		
Azosp	oirilum.		
1			

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

## Jeonghwan Jang<sup>1</sup>, Yoriko Sakai<sup>2</sup>, Masahito Hayatsu<sup>2</sup>, Keishi Senoo<sup>3</sup>, and Satoshi Ishii<sup>1,3,4</sup>

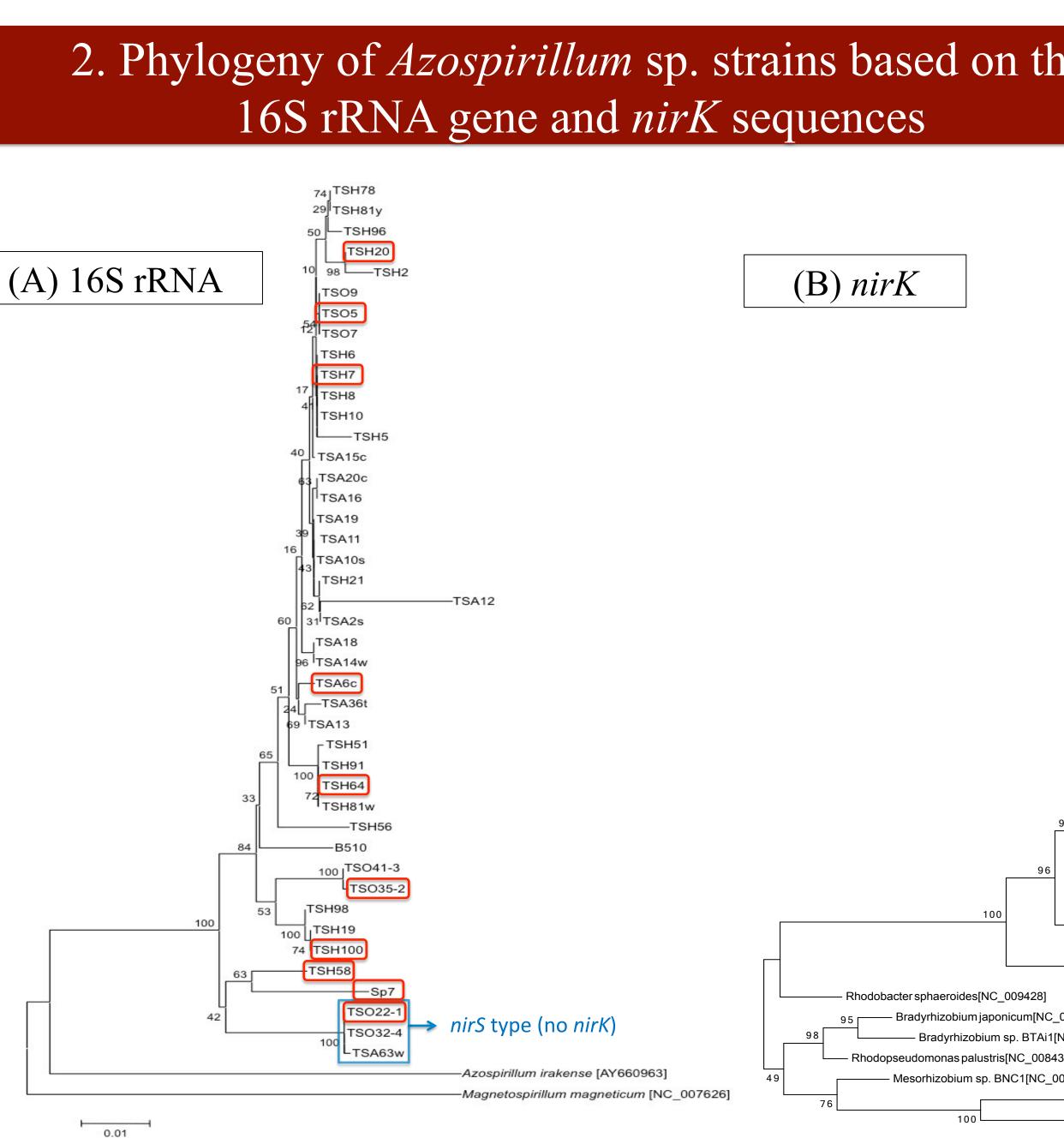
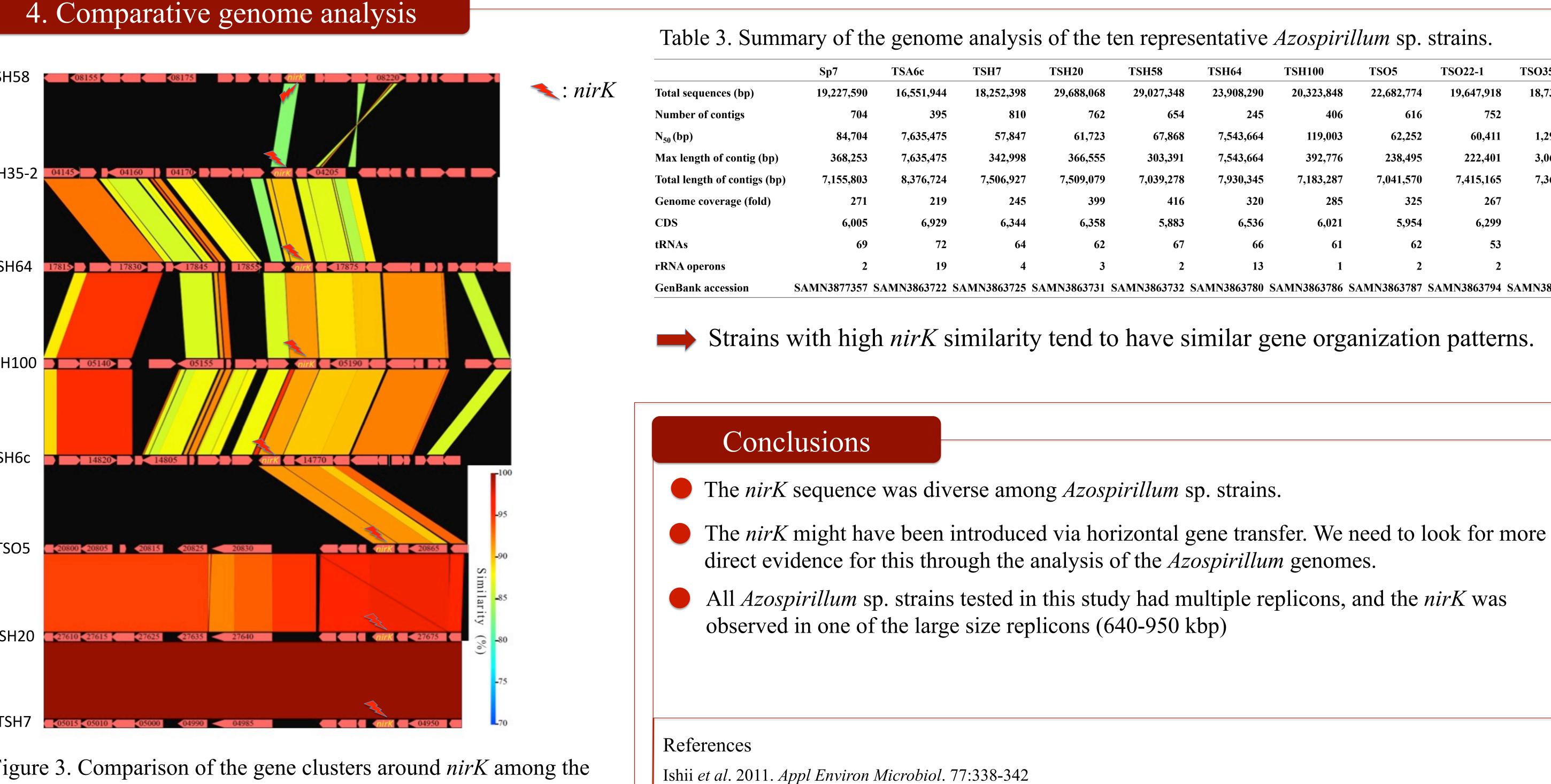


Figure 2. Phylogenetic trees constructed based on (A) 16S rRNA gene and sequences using the maximum likelihood method with bootstrap analysis. red were selected for genome sequencing.

- Sequence analysis of the *nirK* revealed the diversity within the gene or a second Azospirillum
- The nirK phylogeny of many Azospirillum sp. strains did not match w 16S rRNA gene phylogeny, suggesting the possible occurrences of hor transfer of *nirK*.



*zospirillum* sp. strains. Heatmap was generated using GenomeMatcher v2.03.

e		3. Re	1	ons cont					
	T 11								
TSH20 TSH21		2. Numb				-			
TSH7 TSA11	<b></b>	sed-field	U		-			<b>_</b>	
TSH8 TSH5	detern	nined usi	ng th	e sou	itherr	n-blo	ot hył	oridiz	za
<sup>97</sup> TSA10									
TSH6 TSH10		strain ID					replice	on size	(K
TSA14w TSA15c		TSH2	*877	765	716	575	467	403	1
92 TSA18		TSH5	*857	770	703	581	442	403	1
TSA12 TSA2s		TSH6	*880	776	586	476	412	12	~
TSH78		TSH7	*862 *076	764	713	570	476	437	3
93 TSH81y		TSH8	*876 *969	774 769	713 716	577	468 450	406 405	1
2 TSH96 TSO7		TSH10	*868 *866	768 570	716 525	583	450 205	405 12	1
4 5 TSO5		TSH19 TSH20	*866 *866	579 766	525 716	444 579	305 463	12 399	-
TSH10		TSH20 TSH21	*846	782	716 674	579	463 440	288	1
TSH19 TSH100		TSH21 TSH51	<b>846</b> 962	782 920	674 * <b>847</b>	555 761	440 484	288 379	1 2
100 TSH98		TSH51 TSH56	922	832	*640	565	435	283	1
TSA16 TSA6c		TSH58	805	*642	549	383	168	68 <u>68</u>	T
TSA20c		TSH64	951	905	*837	565 749	478	367	2
TSA36t		TSH78	* <b>935</b>	761	693	573	536	427	3
TSH51 - TSA13		TSH81w	959	914	* <b>846</b>	769	490	381	1
rSH64		TSH81y	* <b>941</b>	772	719	584	544	434	3
TSH81w		TSH91	1031	908	*840	757	486	380	2
<sup>I</sup> TSH91 — B510		TSH91 TSH96	* <b>847</b>	752	690	618	484	380 442	3
TSH56		TSH98	*863	576	521	491	307	11	5
TSO35-2		TSH100	*874	570	524	480	441	296	1
TSO41-3		TSA2s	949	*887	664	607	411		Ŧ
- TSH58		TSA6c	*846	606	512	295	146	83	
3]		TSA10	931	*876	775	547	452	332	
J 485]		TSA13	981	*921	660	494	392	288	1
		TSA15c	926	*880	785	556	452	338	_
Shewanella_denitrificans[NC_007954]		TSA16	*819	585	497	278	66		
– Nitrosococcus oceani[NC_007484]		TSA19	899	858	756	552	458	341	3
		TSA20c	*800	583	496	290			
		TSA36t	949	*876	616	441	286	10	
B) nirK		TSO5	*893	698	585	533	432	291	6
		TSO7	*915	720	606	553	449	303	8
rains marked as		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		
the genus		B510	703	*662	612	521	243		
		*: Replico	n whic	h conta	ains the	e nirK	gene		
ll with the	<b>→</b> \\/ı	iltinla ra	nlina	na m	ara id	ontil	fiad ;	n 11	Л
		ultiple re	pheo	112 M (				II all	A
zontal	📥 Th	e <i>nirK</i> g	ene w	vas id	entif	ied i	n lar	ge re	nl
					VIIII		11 101	5010	Ч

	Sp7	TSA6c	TSH7	TSH20	TSH58	TSH64	<b>TSH100</b>	TSO5	<b>TSO22-1</b>	TSO35-2
Fotal 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 <sub>50</sub> (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
Fotal 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
RNAs	69	72	64	62	67	66	61	62	53	72
<b>·RNA operons</b>	2	19	4	3	2	13	1	2	2	13
GenBank accession	SAMN3877357 S	SAMN3863722 S	AMN3863725 S	AMN3863731 S	SAMN3863732 S	SAMN3863780 S	SAMN3863786 S	AMN3863787 S	SAMN3863794	SAMN3863795

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

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



r sj	9. st	rair	IS		
ne					
ch A ns co	-	ning	the <i>r</i>	<b>_</b>	in based ene was
Kbp)					
12 11					
397 12 12	12				
11 11 273 12					
267 322 130	11	241	193	7	
329 283 314	285 11	235	138	118	
11					
13					
302	23				
66 8					
	<i>spiril</i> ons (6		•		tested