

Genome sequences of candidate wheat blast biocontrol bacteria

Emilie Chanclud¹, Joe Win¹, Jacob Malone², Musrat Zahan Surovy³, Dipali Rani Gupta³, Tofazzal Islam³, Sophien Kamoun¹

¹The Sainsbury Laboratory, Norwich Research Park, Norwich, UK

²Molecular Microbiology, John Innes Centre, Norwich, UK

³Department of Biotechnology, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur 1707, Bangladesh.

In an effort to combat wheat blast disease in Bangladesh, Prof. Tofazzal Islam and team have identified several biocontrol bacteria that have the ability to inhibit fungal growth in wheat ([Surovy et al., 2017](#)). They have isolated a number of these agents and we have sequenced the genomes of four bacterial strains to 30x coverage. The genome sequence data is now available to download from links in the tables (Table 1 and 2) below.

Table 1. Genome sequence data of biocontrol bacteria active against wheat blast disease

Bacterial isolate	Source	Appearance ¹	Identification from 16S rRNA sequence ²	Filtered Reads ³	Assemblies ⁴
BTS 3 (R2)	Seeds of Ranga binni (local rice)	White	<i>Bacillus subtilis</i>	10294_R2_1_trimmed.fastq.gz , 10294_R2_2_trimmed.fastq.gz 10294_R2_U1_trimmed.fastq.gz , 10294_R2_U2_trimmed.fastq.gz	fasta , gff , gbk
BTS 4 (R3)	Seeds of Ranga binni (local rice)	White, sticky	<i>Bacillus amyloliquefaciens</i>	10295_R3_1_trimmed.fastq.gz , 10295_R3_2_trimmed.fastq.gz 10295_R3_U1_trimmed.fastq.gz , 10295_R3_U2_trimmed.fastq.gz	fasta , gff , gbk
BTS 5 (SL3)	Seeds of Shakhorkora (local rice)	Light yellowish	<i>Staphylococcus saprophyticus</i>	10293_SL3_1_trimmed.fastq.gz , 10293_SL3_2_trimmed.fastq.gz 10293_SL3_U1_trimmed.fastq.gz , 10293_SL3_U2_trimmed.fastq.gz	fasta , gff , gbk
BTLK6A (K6A)	Seeds of Kanchan (wheat)	White	<i>Bacillus amyloliquefasciens</i>	10296_K6A_1_trimmed.fastq.gz , 10296_K6A_2_trimmed.fastq.gz 10296_K6A_U1_trimmed.fastq.gz , 10296_K6A_U2_trimmed.fastq.gz	fasta , gff , gbk

¹Appearance of bacteria in liquid cultures.

²Sequencing of 16S rRNA was done independently by Prof. Islam of whole genome sequencing.

³Four files for each sample: samplename_1_trimmed.fastq.gz (forward reads for sample), samplename_2_trimmed.fastq.gz (reverse reads for sample), samplename_U1_trimmed.fastq.gz (forward reads that are unpaired, the reverse read was lost during trimming), samplename_U2_trimmed.fastq.gz (reverse reads that are unpaired, the forward read was lost during trimming).

⁴Genome assemblies were made using [SPAdes](#), and annotations were performed using [Prokka](#).

Following data (tables 2-4) for sequencing and assembly statistics were provided by MicrobesNG

Table 2. Trimmed Reads

Sample id	Median insert size	Mean coverage	Mean coverage excluding 0s	Number of reads ¹	Number of reads w/ insert size > 300	Links to sequence reads ²
BTS 3 (R2)	510	91.20	91.20	888940	601581	10294_R2_1_trimmed.fastq.gz 10294_R2_2_trimmed.fastq.gz 10294_R2_U1_trimmed.fastq.gz 10294_R2_U2_trimmed.fastq.gz
BTS 4 (R3)	591	58.80	58.79	541941	393264	10295_R3_1_trimmed.fastq.gz 10295_R3_2_trimmed.fastq.gz 10295_R3_U1_trimmed.fastq.gz 10295_R3_U2_trimmed.fastq.gz
BTS 5 (SL3)	444	217.14	217.17	1251774	262243	10293_SL3_1_trimmed.fastq.gz 10293_SL3_2_trimmed.fastq.gz 10293_SL3_U1_trimmed.fastq.gz 10293_SL3_U2_trimmed.fastq.gz
BTLK6A (K6A)	573	79.03	79.03	731187	504962	10296_K6A_1_trimmed.fastq.gz 10296_K6A_2_trimmed.fastq.gz 10296_K6A_U1_trimmed.fastq.gz 10296_K6A_U2_trimmed.fastq.gz

¹The reads were trimmed using [Trimmomatic](#) and the quality was assessed using in-house scripts combined with the following software: [Samtools](#), [BedTools](#) and [bwa-mem](#).

²Four files for each sample: samplename_1_trimmed.fastq.gz (forward reads for sample), samplename_2_trimmed.fastq.gz (reverse reads for sample), samplename_U1_trimmed.fastq.gz (forward reads that are unpaired, the reverse read was lost during trimming), samplename_U2_trimmed.fastq.gz (reverse reads that are unpaired, the forward read was lost during trimming).

Table 3. Assembly Data

Sample id	#contigs (>= 0 bp) ¹	#contigs (>= 1000 bp) ¹	Total length (>= 0 bp) ¹	Total length (>= 1000 bp) ¹	#contigs ¹	Largest contig ¹	Total length ¹	GC (%) ¹	N50 ¹	N75 ¹	L50 ¹	L75 ¹	#N's per 100 kbp ¹
BTS 3 (R2)	49	16	4122253	4108665	23	1140720	4113250	43.48	1063829	1023790	2	3	0.00
BTS 4 (R3)	28	13	3907814	3901959	13	2032688	3901959	46.51	2032688	1024524	1	2	0.00
BTS 5 (SL3)	80	37	2644501	2626880	39	572144	2628542	33.50	128421	84554	6	12	0.00
BTLK6A (K6A)	33	15	3908827	3901415	16	1083238	3902222	46.51	1024542	947863	2	3	0.00

¹Genome assemblies were made using [SPAdes](#), and the assembly metrics are calculated using [QUAST](#), for further details about what they mean visit the [QUAST manual](#).

Table 4. Taxonomic Distribution

Sample	Unclassified (%) ¹	Most frequent Family ¹	Most frequent Family(%) ¹	2nd most frequent Family (%) ¹	Most frequent genus (%) ¹	2nd Most frequent genus ¹	Most frequent species (%) ¹	<i>Escherichia coli</i> (%) ¹
BTS 3 (R2)	Unclassified (1.73)	Bacillaceae (96.65)	96.65	Enterobacteriaceae (0.02)	<i>Bacillus</i> (96.63)	<i>Salmonella</i> (0.01)	<i>Bacillus subtilis</i> (91.86)	<i>Escherichia coli</i> (0.00)
BTS 4 (R3)	Unclassified (1.9)	Bacillaceae (97.98)	97.98	Rhizobiaceae (0.02)	<i>Bacillus</i> (97.96)	<i>Rhizobium</i> (0.02)	<i>Bacillus amyloliquefaciens</i> ² (32.10)	<i>Escherichia coli</i> (0.00)
BTS 5 (SL3)	Unclassified (72.25)	Staphylococcaceae (22.44)	22.44	Enterococcaceae (2.49)	<i>Staphylococcus</i> (22.22)	<i>Enterococcus</i> (2.47)	<i>Staphylococcus saprophyticus</i> ² (7.64)	<i>Escherichia coli</i> (0.00)
BTLK6A (K6A)	Unclassified (2.27)	Bacillaceae (97.62) ³	97.61	Rhizobiaceae (0.01)	<i>Bacillus</i> (97.59)	<i>Rhizobium</i> (0.01)	<i>Bacillus amyloliquefaciens</i> ² (32.10)	<i>Escherichia coli</i> (0.00)

¹Mapping of reads to families and genera calculated using the software [Kraken](#).

²Confirmation of the identity of these species will need community efforts, perhaps in analysing the genome data presented here.

Reference

Surovy, Musrat Zahan; Gupta, Dipali Rani; Chanclud, Emilie; Win, Joe; Kamoun, Sophien; Islam, Tofazzal (2017): Plant probiotic bacteria suppress wheat blast fungus Magnaporthe oryzae Triticum pathotype. figshare. <https://doi.org/10.6084/m9.figshare.5549278.v1>