

SUPPLEMENTAL MATERIAL

Whole-genome Sequence Analysis and Genome-wide Virulence Gene

Identification of *Riemerella anatipestifer* Strain Yb2

Xiaolan Wang^a, Chan Ding^a, Shaohui Wang^a, Xianghan Han^a, Shengqing Yu^{a,#}

Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences

(CAAS), Shanghai, P. R. China^a.

TABLE S1 Strains, plasmids and primers used in this study

Strains, plasmids or primers	Characteristics	Source or reference
Yb2	<i>Riemerella anatipestifer</i> serotype 2 strain, virulent isolate, Kanamycin resistant	1
BW19851	<i>Escherichia coli</i> , plasmid pEP4351 in BW19851, chloramphenicol resistant	2
pEP4351	pir-requiring R6K oriV, RP4 oriT, Cm ^r Tc ^r (Em ^r),vector for Tn4351 mutagenesis	2
Primers		
RA rRNA-F	16S 5' GAGCGGTAGAGTATCTTCGGATACT3'	This study
RA rRNA-R	16S 5' AATTCCTTTGAGTTTCAACCTTGCG3'	This study
Erm-F	5' GCCCGAAATGTTCAAGTTGT 3'	3
Erm-R	5' CTTGACAACCCGACTTT 3'	3
Tn4351-F	5' TGGCACCTTTGTGGTTCTTAC 3'	3
Tn4351-F	5' GAGAGACAATGTCCCCCTTTC 3'	3
SP1	5' CTCCCAGAAAATTTCCAAGACTCTCA 3'	2
SP2	5' TAAAGTGCTGACCCGTAACGAAC 3'	2
SP3	5' GTGGTAGCTATAGCATGGAGCTTGC 3'	2
340	5' GACTTGGATACCTCACGCC 3'	4
341	5' TTGGAAATTTTCTGGGAGG 3'	4
TN-1	5' GGACCTACCTCATAGACAA 3'	4
IS4351-F	5' TCAGAGTGAGAGAAAGG 3'	4

TABLE S2 Summary of *Riemerella anatipestifer* Yb2 gene prediction

Features	Genome
Total size, bp	2,184,066
Coding region size (bp)	1,984,800
Code percent (%)	90.88
GC content of coding region, %	35.73
Gene numbers	2,073
Average gene length (bp)	976
ORFs	2,021
Pseudo genes	7
CRISPR Arrays	1
rRNA operons	6
tRNAs	38
Other ncRNA	1
Fameshifted genes	3

TABLE S3 Numbers of gene associated with the general COG functional categories

Code	Genes	Percent (%)	Description ^a
A	0	0.0	RNA processing and modification
B	0	0.0	Chromatin structure and dynamics
C	80	6.59	Energy production and conversion
D	17	1.40	Cell cycle control, cell division, chromosome partitioning
E	102	8.40	Amino acid transport and metabolism
F	48	3.95	Nucleotide transport and metabolism
G	38	3.13	Carbohydrate transport and metabolism
H	75	6.18	Coenzyme transport and metabolism
I	47	3.87	Lipid transport and metabolism
J	127	10.46	Translation, ribosomal structure and biogenesis
K	52	4.28	Transcription
L	82	6.75	Replication, recombination and repair
M	119	9.80	Cell wall/membrane/envelope biogenesis
N	3	0.25	Cell motility
O	64	5.27	Posttranslational modification, protein turnover, chaperones
P	67	5.51	Inorganic ion transport and metabolism
Q	21	1.73	Secondary metabolites biosynthesis, transport and catabolism
R	140	11.53	General function prediction only
S	70	5.77	Function unknown
T	24	1.98	Signal transduction mechanisms
U	18	1.48	Intracellular trafficking, secretion, and vesicular transport
V	20	1.65	Defense mechanisms
W	0	0.0	Extracellular structures
Y	0	0.0	Nuclear structure
Z	0	0.0	Cytoskeleton
_	807	38.92	No related COGs

^aFunctional characterization of the proteins was predicted by searching against eggNOG (Version 3) database using BLASTP.

TABLE S4 Overview of the genomic islands in *Riemerella anatipestifer* Yb2

No.	CDS	Characteristics or putative functions
GI-1 ^a	<i>AS87_03450-</i> <i>AS87_03580</i>	hypothetical protein, N-acetylmuramoyl-L-alanine amidase, phage head protein, aminopeptidase, head morphogenesis protein, DNA-binding protein
GI-2	<i>AS87_05965-</i> <i>AS87_06030</i>	integrase, restriction endonuclease, ATPase AAA, transcriptional regulator, hypothetical protein
GI-3	<i>AS87_08275-</i> <i>AS87_08285</i>	hypothetical protein
GI-4	<i>AS87_08505-</i> <i>AS87_08565</i>	peptidase M23, serine/threonine protein kinase, transposase, ATPase AAA, membrane protein, hypothetical protein
GI-5	<i>AS87_09220-</i> <i>AS87_09230</i>	DNA-binding protein, restriction endonuclease, hypothetical protein
GI-6	<i>AS87_09605-</i> <i>AS87_09920</i>	antibiotic activity (LinF, chloramphenicol acetyltransferase, tetracycline resistance protein); transporters (mercury transporter, cation transporter, heavy metal transporter); transcriptional regulator (AraC, XRE); toxin (RelE, YoeB); multicopper oxidase, polymerase, restriction endonuclease MboI, DNA methylase, membrane protein, twitching motility protein PilT, prevent-host-death protein, esterase

^aGenomic islands (GIs) were detected by IslandViewer, and a total of six GIs were predicted.

TABLE S5 Single-nucleotide polymorphisms (SNPs) between *Riemerella anatipestifer* Yb2 and other reference strains DSM 15868, RA-GD, RA-CH-1 and RA-CH-2

		DSM 15868			
Yb2	A	C	G	T	
A		317 ^a	2276	545	
C	294		125	2330	
G	2366	122		328	
T	521	2298	382		
		RA-GD			
Yb2	A	C	G	T	
A		153	850	200	
C	137		45	877	
G	899	48		123	
T	210	880	126		
		RA-CH-1			
Yb2	A	C	G	T	
A		2057	14834	3537	
C	2273		742	15541	
G	15419	780		2307	
T	3511	15000	2131		
		RA-CH-2			
Yb2	A	C	G	T	
A		187	1059	290	
C	219		100	997	
G	1028	85		189	
T	256	972	166		

^aGenomic sequences with adjusted start position were submitted to NUCmer system (NUCleotide MUMmer Version 3.07) for alignment, and non-specific records were filtered by Delta-filter. After alignment of sequenced reads to a reference genome, SNP calling was performed using the Show-snps software.

REFERENCES

1. **Hu Q, Zhang Z, Miao J, Liu Y, Liu X, S D.** 2001. The Epidemiology study of *Riemerella anatipestifer* infection in Jiangsu and Anhui provinces. Chinese Journal of Veterinary Science and Technology **31**:12-13.
2. **Hu Q, Zhu Y, Tu J, Yin Y, Wang X, Han X, Ding C, Zhang B, Yu S.** 2012. Identification of the genes involved in *Riemerella anatipestifer* biofilm formation by random transposon mutagenesis. PloS one **7**:e39805.
3. **Wang X, Ding C, Wang S, Han X, Hou W, Yue J, Zou J, Yu S.** 2014. The AS87_04050 gene is involved in bacterial lipopolysaccharide biosynthesis and pathogenicity of *Riemerella anatipestifer*. PloS one **9**:e109962.
4. **Álvarez B, Secades P, Prieto M, McBride M, Guijarro J.** 2006. A mutation in *Flavobacterium psychrophilum tlpB* inhibits gliding motility and induces biofilm formation. Appl Environ Microbiol **72**:4044-4053.