Genome Sequence of Atyrau-5BJN(IL18), a Recombinant Lumpy Skin Disease Virus with Knockout of Virulence Genes

Aisha U. Issabek,a Mukhit B. Orynbayev,a Kulyaisan T. Sultankulova,a Asylulan Amirgazin,b Kunsulu D. Zakarya,a Zamira Omarova,a Olga V. Chervyakova,a

aResearch Institute for Biological Safety Problems, Gvardeysky, Republic of Kazakhstan
bNational Center for Biotechnology, Nur-Sultan, Republic of Kazakhstan

ABSTRACT Here, we present the coding sequence of the genome of the recombinant lumpy skin disease virus (LSDV) Atyrau-5BJN(IL18), obtained by knocking out four genes in the genome of a virulent field LSDV isolate. Genome sequencing confirmed the deletion of genes and the insertion of a foreign sequence in the viral genome.

Lumpy skin disease virus (LSDV), genus Capripoxvirus of the family Poxviridae, is the causative agent of an infectious disease in cattle. Rapid diagnostics and vaccination are the basis for controlling the disease spread. The development of molecular biology methods has made it possible to obtain attenuated viruses by targeted mutagenesis of the viral genome (1).

Here, we present the coding sequence of the genome of the recombinant lumpy skin disease virus (LSDV) Atyrau-5BJN(IL18), obtained by knocking out four genes in the genome of a virulent field LSDV isolate. The primary virulent virus was isolated from an infected cow during an outbreak of lumpy skin disease in Kazakhstan in July 2016 (ProMED 20160722.4363497). Initially, four passages were performed in the MDBK cell line, and then four genes were knocked out by homologous recombination with transient dominant selection of recombinant viruses (2). We anticipate that the recombinant Atyrau-5BJN(IL18) will be harmless, immunogenic, and protect cattle during challenge with virulent virus (3).

The recombinant virus was produced in lamb testicle cell culture and purified by centrifugation in a sucrose density gradient (4). Genomic DNA was isolated using TRIzol reagent (Invitrogen, USA) according to the manufacturer’s protocol.

A DNA library was prepared using the Nextera XT DNA library preparation kit (FC-131-1024). Sequencing was performed on the Illumina MiSeq platform using the MiSeq reagent kit v3 (600 cycles; MS-102-3003). Quality control of the reads was carried out using FastQC v0.11.15 (5). Overall, 2,268,136 reads were obtained and trimmed to at least Q30 from the 5’ end on 19 nucleotides and from the 3’ end on 3 nucleotides using the programs Sickle v1.33 (6) and SeqTK v1.3-r106 (7).

De novo genome assembly was performed using the program SPAdes v3.13 (8) with a k-mer length of 127 and the “--careful” option. As a result, a contig with a length of 145,167 bp was obtained; it was identified using blastn and the nucleotide collection (nt) database (9) as an LSDV genome.

To determine the depth of sequencing, the reads were mapped to the corresponding assemblies using the BWA-mem v0.7.17 algorithm (standard settings) (10), and the depths for each nucleotide and the median were calculated using SAMtools v1.15.1 (11). The sequencing depth was 255×.

Inverted terminal repeat sequences were PCR amplified using AccuPrime Taq DNA polymerase, high fidelity, and determined by Sanger sequencing using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Austin, TX) on the 3130xl genetic analyzer (Applied Biosystems, Hitachi, Japan). The resulting contigs were manually combined into a sequence of 148,628 bp with an average G+C content of 25.87%, evenly distributed. About
600 bp were missing from both the 5’ and 3’ ends of the genome compared to the genome sequences present in GenBank.

The BLAST results showed that Atyrau-5BJN(IL18) has 100% identity with LSDV field isolates from Israel (2012; GenBank accession number KX894508), Turkey (2014; MN995838), Greece (2015; KY829023), Kazakhstan (2016; MN642592). The identity with the NCBI Reference Sequence NC_003027 was 99.88%. The query coverage in both cases was 99%.

Genome annotation was performed using GATU software relative to the LSDV isolate Kubash/KAZ/16 (GenBank accession number MN642592) (12). In addition to the targeted mutations introduced into the genome of the recombinant virus, three nucleotide substitutions were identified that would lead to amino acid substitutions (Table 1).

Genome sequencing confirmed the deletions of genes and the insertion of a foreign sequence in the viral genome (Table 1).

### Data availability.

The genome sequence of the lumpy skin disease virus Atyrau-5BJN (IL18) has been deposited at GenBank under accession number ON000567, and the raw data have been submitted to the SRA under BioProject accession number PRJNA825391.

### ACKNOWLEDGMENT

This research was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (grant AP08856376).

### REFERENCES

1. Moss B. 2013. Reflections on the early development of poxvirus vectors. Vaccine 31:4220–4222. https://doi.org/10.1016/j.vaccine.2013.03.042.

2. Falkner FG, Moss B. 1990. Transient dominant selection of recombinant vaccinia virus. J Virol 64:3108–3111. https://doi.org/10.1128/JVI.64.6.3108-3111.1990.

3. Chernyakova OV, Sultankulova KT, Nissanova RK, Oynbayev MB. 2019. Lumpy skin disease virus: approaches to attenuation for vaccine development. Eur J Appl Biotechnol 1:3–14. https://doi.org/10.11134/btp.12019.6.

4. Zaitsev VL, Sandybayev NT, Sultankulova KT, Belousov VY, Chernyakova OV, Strochkov VM, Nikolenko SI, Pham S, Pijtalski AD, Pyshkin AV, Sirotkin AV, Vyakhri N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

5. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

6. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. https://github.com/najoshi/sickle.

7. Li H. 2012. seqtk: toolkit for processing sequences in FASTA/Q formats. https://github.com/lh3/seqtk.

8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyshkin AV, Pyotrkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/0022-2836(80)90360-2.

10. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997v2. https://arxiv.org/abs/1303.3997v2.

11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

12. Tcherepanov V, Ehlers A, Upton C. 2006. Genome Annotation Transfer Utility (GATU): rapid annotation of viral genomes using a closely related reference genome. BMC Genomics 7:150. https://doi.org/10.1186/1471-2164-7-150.