

DNA Sequence Analysis (Gene-Level)

1. Identify probable gene(s) from the provided DNA sequence

The provided 570 bp DNA sequence was analyzed using NCBI BLASTn. BLAST results revealed 100% query coverage, 100% identity, and E-value of 0.0 against multiple plasmid sequences of *Klebsiella pneumoniae*. This confirms that the sequence originates from a conserved plasmid-borne gene in *K. pneumoniae*.

Top 3 BLASTn Hits (Summary):

- 1) *Klebsiella pneumoniae* strain D17KP0018 plasmid pD17KP0018-1 (CP052337.1)
 - 2) *Klebsiella pneumoniae* isolate 11 plasmid P1 (OW969612.1)
 - 3) *Klebsiella pneumoniae* strain 51015 plasmid p51015_CTX_M_15 (CP050379.1)
- All hits show 570/570 identities (100%) with no gaps, indicating a highly conserved gene.

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

DownloadSelect columnsShow100?

☒ select all

100 sequences selected

GenBank

Graphics

Distance tree of results

MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Klebsiella pneumoniae isolate 11 genome assembly, plasmid: P1	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	220010	OW969612.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain D17KP0018 plasmid pD17KP0018-1, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	233970	CP052337.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain 51015 plasmid p51015_CTX_M_15, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	225540	CP050379.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain NK_H4_026 plasmid pNK_H4_026.1, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	144316	CP152610.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain F17KP0001 plasmid pF17KP0001-1, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	192688	CP052575.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae isolate INF327-sc-2280132 genome assembly, plasmid: 2	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	227807	LR890572.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain ARLG-4871 plasmid pC603_1, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	237925	CP067589.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae strain 2017HL-00030 plasmid Kpn-WC17-008-01, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	215001	CP094904.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae isolate KSB1_1H genome assembly, plasmid: 2	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	225937	LR890723.1
<input checked="" type="checkbox"/> Klebsiella pneumoniae subsp. pneumoniae strain LC-424/19 plasmid oIT-424-Flik2, complete sequence	Klebsiella pneu...	1053	1053	100%	0.0	100.00%	161141	CP110968.1

Feedback

Download	GenBank	Graphics	Next	Previous	Descriptions
Klebsiella pneumoniae isolate 11 genome assembly, plasmid: P1					
Sequence ID: OW969612.1 Length: 220010 Number of Matches: 1					
Range 1: 64980 to 65549 GenBank Graphics Next Match Previous Match					
Score	Expect	Identities	Gaps	Strand	
1053 bits(570)	0.0	570/570(100%)	0/570(0%)	Plus/Minus	
Query 1	ATGGACATCGATTTC	AAGAAGTTGGCTCCCTGGA	ACTGGTTCAAGAACGAGCAGCAAGAG	60	
Sbjct 65549	ATGGACATCGATTTC	AAGAAGTTGGCTCCCTGGA	ACTGGTTCAAGAACGAGCAGCAAGAG	65490	
Query 61	CAGCAGACCGTCTCT	CCCTGCCGGTGACGCGCAATGACCTGCCAGCGGCGGCGCA	120		
Sbjct 65489	CAGCAGACCGTCTCT	CCCTGCCGGTGACGCGCAATGACCTGCCAGCGGCGGCGCA	65430		
Query 121	GTCAGCCGATCCTGCA	ACTGCATCGGGAATCGACCGGCTGTTGATGACGATTCCGG	180		
Sbjct 65429	GTCAGCCGATCCTGCA	ACTGCATCGGGAATCGACCGGCTGTTGATGACGATTCCGG	65370		
Query 181	GGCTTCGGTTTTCCGG	CGTTGAACATGCCGAGTGGCCATCCGATTGGTCGGGCATGCTG	240		
Sbjct 65369	GGCTTCGGTTTTCCGG	CGTTGAACATGCCGAGTGGCCATCCGATTGGTCGGGCATGCTG	65310		
Query 241	AAGCCGGCCCTGGACAT	CCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTGCC	300		
Sbjct 65309	AAGCCGGCCCTGGACAT	CCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTGCC	65250		
Query 301	GGTGTGAGGAGGAGGACAT	CCAGATCACTCTCGACAACGACGTGCTGATGGTGGTGGC	360		

Klebsiella pneumoniae strain D17KP0018 plasmid pD17KP0018-1, complete sequence

Sequence ID: [CP052337.1](#) Length: 233970 Number of Matches: 1

Range 1: 62939 to 63508 [GenBank](#) [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1053 bits(570)	0.0	570/570(100%)	0/570(0%)	Plus/Minus
Query 1	ATGGACATCGATTTC	AAGAAGTTGGCTCCCTGGA	ACTGGTTCAAGAACGAGCAGCAAGAG	60
Sbjct 63508	ATGGACATCGATTTC	AAGAAGTTGGCTCCCTGGA	ACTGGTTCAAGAACGAGCAGCAAGAG	63449
Query 61	CAGCAGACCGTCTCT	TCCCTGCCGGTGCAGCGCAATGACCTGCCAGCGGCAGCGGGCCA	120	
Sbjct 63448	CAGCAGACCGTCTCT	TCCCTGCCGGTGCAGCGCAATGACCTGCCAGCGGCAGCGGGCCA	63389	
Query 121	GTCAGCCCGATCCTG	CAACTGCATCGGGAATCGACCGGCTGTTGATGACGCATTCCGG	180	
Sbjct 63388	GTCAGCCCGATCCTG	CAACTGCATCGGGAATCGACCGGCTGTTGATGACGCATTCCGG	63329	
Query 181	GGCTTCGGTTTTCC	GGCGTTGAACATGCCGCACTGGCCATCCGATTGGTCGGGCATGCTG	240	
Sbjct 63328	GGCTTCGGTTTTCC	GGCGTTGAACATGCCGCACTGGCCATCCGATTGGTCGGGCATGCTG	63269	
Query 241	AAGCCGGCCCTGGAC	ATCCAGGAAACGACAAGCAGTACAAGATTGCCCTGGAAGTGCCC	300	
Sbjct 63268	AAGCCGGCCCTGGAC	ATCCAGGAAACGACAAGCAGTACAAGATTGCCCTGGAAGTGCCC	63209	
Query 301	GGTGTCTGAGGAGA	AGGACATCCAGATCACTCTCGACAACGACGTGCTGATGGTGCCTGGC	360	
Sbjct 63208	GGTGTCTGAGGAGA	AGGACATCCAGATCACTCTCGACAACGACGTGCTGATGGTGCCTGGC	63149	
Query 361	GAGAAGCGCCAGGA	ACAGGAGAAGAAGGAAGGTGGCTTCCACCGTGTGGAGCGCTCCTAC	420	

2. Detect ORFs and determine the most likely coding region

ORF detection was performed using NCBI ORF Finder with the standard genetic code. Two ORFs were detected. ORF1 was selected as the most likely coding region based on its full-length coverage, presence of a start codon, and consistency with BLASTn results.

ORF Selection Reasoning:

- ORF1 length: 570 nt (189 amino acids)
- Located on the positive strand
- Covers the complete query sequence
- Matches known protein-coding regions in Klebsiella plasmids

ORF2 was shorter (318 nt) and located on the reverse strand, making it less likely to be the functional gene.

NIH National Library of Medicine
National Center for Biotechnology Information

< ORFfinder submitting page > Help

Open Reading Frame Viewer

Sequence

ORFs found: 2 Genetic code: 1 Start codon: 'ATG' only

ORF1 (189 aa) Display ORF as... Mark

Mark subset... Marked: 0 Download marked set as Protein FASTA

Label	Strand	Frame	Start	Stop	Length (nt aa)
ORF1	+	1	<1	570	570 189
ORF2	-	2	320	>3	318 105

Six-frame translation...

3. Translate the ORF to protein

The selected ORF (ORF1) was translated using an online transcription and translation tool. The resulting protein consists of 189 amino acids, indicating a functional protein product. This translation confirms that the DNA sequence encodes a protein rather than a non-coding region.

Total count, all bases:	570
Adenine (A) count:	133
Thymine (T) count:	96
Guanine (G) count:	178
Cytosine (C) count:	163
%G~C content:	59.8

4. GC content and repeat analysis

GC content analysis was performed using a simple online nucleotide composition calculator. The sequence showed a GC content of 59.8%, which is characteristic of Klebsiella pneumoniae plasmid DNA. No significant tandem or low-complexity repeats were detected within the sequence.

Nucleotide Composition:

Total length: 570 bp
Adenine (A): 133
Thymine (T): 96
Guanine (G): 178
Cytosine (C): 163
GC Content: 59.8%

Download ▾GenBankGraphics

▼Next▲Previous◀Descriptions

Klebsiella pneumoniae strain 51015 plasmid p51015_CTX_M_15, complete sequence

Sequence ID: [CP050379.1](#)Length: 225540Number of Matches: 1

Range 1: 49947 to 50516GenBankGraphics

▼Next Match▲Previous Match

Score	Expect	Identities	Gaps	Strand
1053 bits(570)	0.0	570/570(100%)	0/570(0%)	Plus/Plus
Query 1	ATGGACATCGATTTCAAGAAAGTTGGCTCCCTGGAACTGGTTCAAGAACGAGCAAGAG	60		
Sbjct 49947	ATGGACATCGATTTCAAGAAAGTTGGCTCCCTGGAACTGGTTCAAGAACGAGCAAGAG	50006		
Query 61	CAGCAGACCGTCTCTTCCCTGCCGGTGCAAGCAATGACCTGCCAGCGGCAGCGGGCCA	120		
Sbjct 50007	CAGCAGACCGTCTCTTCCCTGCCGGTGCAAGCAATGACCTGCCAGCGGCAGCGGGCCA	50066		
Query 121	GTCAGCCCGATCCTGCAACTGCATCGGGAATCGACCGGCTGTTGATGACGATTCGGG	180		
Sbjct 50067	GTCAGCCCGATCCTGCAACTGCATCGGGAATCGACCGGCTGTTGATGACGATTCGGG	50126		
Query 181	GGCTTCGGTTTTTCCGGGTTGAACATGCCGAGTGGCCATCCGATTGGTCGGGCATGCTG	240		
Sbjct 50127	GGCTTCGGTTTTTCCGGGTTGAACATGCCGAGTGGCCATCCGATTGGTCGGGCATGCTG	50186		
Query 241	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGAGTACAAGATTGCCCTGGAAAGTGCC	300		
Sbjct 50187	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGAGTACAAGATTGCCCTGGAAAGTGCC	50246		
Query 301	GGTGTGAGGAGAGGACATCCAGATCACTCTCGACAACGACGTGCTGATGGTGGTGGC	360		
Sbjct 50247	GGTGTGAGGAGAGGACATCCAGATCACTCTCGACAACGACGTGCTGATGGTGGTGGC	50306		