

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 51215 plasmid p51215-CTX_M_15 (CP052370.1)

3) *Klebsiella pneumoniae* strain 51015 plasmid p51015_C1_X_M_15 (CP050379.1)
All bits show 570/570 identities (100%) with no gaps, indicating a highly conserved gene

Descriptions		Graphic Summary	Alignments	Taxonomy						
Sequences producing significant alignments					Download	Select columns	Show	100	?	
<input checked="" type="checkbox"/> 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...	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...	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...	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...	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...	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...	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...	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...	Klebsiella_pneu...	1053	1053	100%	0.0	100.00%	215001	CP049040.1
<input checked="" type="checkbox"/>	Klebsiella pneumoniae isolate KSB1_1H genome assembly, plasmid: 2	Klebsiella_pneu...	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 pLT-424-FIIK2, complete sequence	Klebsiella_pneu...	Klebsiella_pneu...	1053	1053	100%	0.0	100.00%	161141	CP110968.1

Klebsiella pneumoniae isolate 11 genome assembly, plasmid: P1					▼ Next	▲ Previous	◀ Descriptions
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		Plus/Minus
1053 bits(570)		0.0	570/570(100%)	0/570(0%)	Plus/Minus		
Query 1	ATGGACATCGATTTCAGAAAGTTGGCTCCCTGGAAACTGGTCAAGAACGAGCAGCAAGAG	60					
Sbjct 65549	ATGGACATCGATTTCAGAAAGTTGGCTCCCTGGAAACTGGTCAAGAACGAGCAGCAAGAG	65490					
Query 61	CAGCAGACCGTCTTCCCCTGCCGTGCAGCGCAATGACCTGCCAGCGCGAGCGGGCCA	120					
Sbjct 65489	CAGCAGACCGTCTTCCCCTGCCGTGCAGCGCAATGACCTGCCAGCGCGAGCGGGCCA	65430					
Query 121	GTCAGCCCAGTCTGCAACTGCATCGGAAATCGACCGCTGTTGATGACGCACTCCGG	180					
Sbjct 65429	GTCAGCCCAGTCTGCAACTGCATCGGAAATCGACCGCTGTTGATGACGCACTCCGG	65370					
Query 181	GGCTTCGGTTTCGGCGTTGAACATGCCGAGTGGCATCGATTGGTCGGCATGCTG	240					
Sbjct 65369	GGCTTCGGTTTCGGCGTTGAACATGCCGAGTGGCATCGATTGGTCGGCATGCTG	65310					
Query 241	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTGC	300					
Sbjct 65309	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTGC	65250					
Query 301	GGTGTGAGGAGAAGGACATCCAGATCACTCTGAAACGACGCTGCTGATGGTGCGTGGC	360					

Klebsiella pneumoniae strain D17KP0018 plasmid pD17KP0018-1, complete sequenceSequence ID: [CP052337.1](#) Length: 233970 Number of Matches: 1Range 1: 62939 to 63508 [GenBank](#) [Graphics](#)[▼ Next Match](#) [▲ Previous Match](#)

Score 1053 bits(570)	Expect 0.0	Identities 570/570(100%)	Gaps 0/570(0%)	Strand Plus/Minus
Query 1	ATGGACATCGATTCAAGAAAGTTGGCTCCCTGGAACTGGTCAAGAACGAGCAGCAAGAG	60		
Sbjct 63508	ATGGACATCGATTCAAGAAAGTTGGCTCCCTGGAACTGGTCAAGAACGAGCAGCAAGAG	63449		
Query 61	CAGCAGACCGCTCTTCCCTGCCGTGCAGCGCAATGACCTGCCAGCGCGAGCGGGCCA	120		
Sbjct 63448	CAGCAGACCGCTCTTCCCTGCCGTGCAGCGCAATGACCTGCCAGCGCGAGCGGGCCA	63389		
Query 121	GTCAGCCCCATCCTGCAACTGCATCGGAAATCGACCGCTGTTGATGACGCATTCCGG	180		
Sbjct 63388	GTCAGCCCCATCCTGCAACTGCATCGGAAATCGACCGCTGTTGATGACGCATTCCGG	63329		
Query 181	GGCTTCGGTTTCCGGCTTGAAACATGCCGCAGTGGCATCCGATTGGTGGGCATGCTG	240		
Sbjct 63328	GGCTTCGGTTTCCGGCTTGAAACATGCCGCAGTGGCATCCGATTGGTGGGCATGCTG	63269		
Query 241	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTCCC	300		
Sbjct 63268	AAGCCGGCCCTGGACATCCAGGAAACCGACAAGCAGTACAAGATTGCCCTGGAAGTCCC	63209		
Query 301	GGTGTGAGGGAGAAGGGACATCCAGATCACTCTGACAAACGACGCTGCTGATGGTGGTGG	360		
Sbjct 63208	GGTGTGAGGGAGAAGGGACATCCAGATCACTCTGACAAACGACGCTGCTGATGGTGGTGG	63149		
Query 361	GAGAAGCGCCAGGAACAGGGAGAAGGAAGGGTGGCTTCCACCGTGTGGAGCGCTCTAC	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.

The screenshot shows the NCBI ORF Finder interface. At the top, there's a logo for the National Library of Medicine and a 'Log in' button. Below the header, it says 'ORFfinder submitting page' and 'Open Reading Frame Viewer'. Under 'Sequence', it indicates 'ORFs found: 2', 'Genetic code: 1', and 'Start codon: 'ATG' only'. The main view shows a horizontal timeline from 1 to 570. Two arrows above the timeline point right, labeled 'ORF1' and 'ORF2'. Below the timeline, a red bar represents the sequence, with arrows indicating the direction of each ORF. The left side shows the DNA sequence with labels for ORF1 and ORF2. On the right, there are various tools and tracks. At the bottom, a table provides detailed information about the ORFs, including their labels, strands, frames, starts, stops, and lengths (in nt and aa).

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

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%

