

M G S S H H H H H H S S G L V P ↓ R G S H M

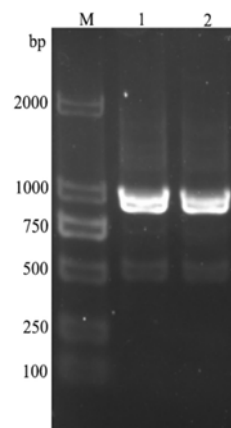
ATGGGCAGCAGCCATCATCATCATCACAGCAGCGCCTGGTGCCGCGCGGCAGCCATATG

Protein	K T Y E A C E L A R E L V D K G A A M D Q V G T W V C I A K
Original 1	AAGACTTACGAGGCTGCGAGCTCGCCGGAGCTGGTGGACAAGGGGCGCGATGGACCAGGTGGGCACCTGGGTCTGCATCGCCAAG
Optimized	AAGACTATGAAGCCTGTGAAGTGGCAGTGAAGTGGTAAAGGCGCAGCATGGATCAGGTGGTACCTGGGTGGTATGGCAAA
Protein	Y E S D F N T S A V G R L N G D G S A D H G I F Q I S D R Y
Original 91	TACGAGTCCGACTTCAACACCTCCGCCGTGGGAGGTGAACGGGACGGCAGCGCCAGCAGGCATCTTCCAGATCAGCGACCGCTAC
Optimized	TATGAAGCGATTTCACACCAAGCGCGTGGCCGCTGAATGGCGATGGTACCGCAGATCATGGTATTTTCAGATTAGCGATCGCTAT
Protein	W C S P Q G W A C N I P C D L L E D D D I S D D W K C A K R
Original 181	TGGTGCTCCCCAGGATGGGCTGCAACATACCTGCGACCTCTGGAGGACGACGACATCTCCGACGACTGGAAGTGCGCCAAGAGG
Optimized	TGGTGAGTCCGCGAGGTGGGCGATTAATATCCGTGCGATCTGCTGGAAGATGATGATATTAGCGATGATTGGAAATGTGCAAAACGC
Protein	I F R Q H K V L S G N G F R A W A V Y S S R C S G D N S G Y
Original 271	ATATTCAGGCAGCAAGGTCCTCAGCGGAATGGGTTACGGGCTGGGCGGTCTACTCCTCCAGGTGCTCGGGTGATAACAGTGGCTAT
Optimized	ATTTTCGTCAGCATAAAGTCTGAGCGCAATGGCTTCGCGCATGGGCGTTTATAGCAGTCGCTGAGTGGCGATAATAGCGGTTAT
Protein	V E G C F D S K A S V S Q S N T T V I T L P A F T T S T T T
Original 361	GTGAAGGTGTTTTGACAGTAAGGCTTCTGTTTCACAGTCCAACACCACAGTCATTACACTGCCTGCCTTCAAACTTCGACCACCACA
Optimized	GTGAAGGTGCTTTGATAGCAAAGCAAGTGTAGCCAGAGCAATACCACCGTGATTACCTGCCGCGCCTTTACCACCACTACCACCACC
Protein	S T S T T T E E P E R V E G S L S A R E R N N A E K V Q V I
Original 451	TCCACTTCTACCACAACGAAGAACCTGAGCGTGTGAAGGATCCCTTAGTGCCGAGAACGGAACAATGCAGAGAAAGTACAAGTTATA
Optimized	AGCACCAGCACCACCACGAAGAACCAGCAAGTGTGAAGGCAGCCTGAGTGCAAGCGAACGCAATATGCAGAAAAAGTGCAGGTTATT
Protein	D L E M M E K K G K S F V N V A S S S K K K W K P W T W E I
Original 541	GATCTAGAAATGATGGAAGAAAGGCAAGTCATTGTAAATGTGGCATCCTCATCAAAAAAGAAATGGAACCATGGACTTGGGAGATA
Optimized	GATCTGGAATGATGGAAGAAAGGTAAAGCTTTGTAACTGGCCAGCAGCAGCAAAAAAGAAATGGAACCGTGGACTGGGAATT

\*

TAAGAAATTC

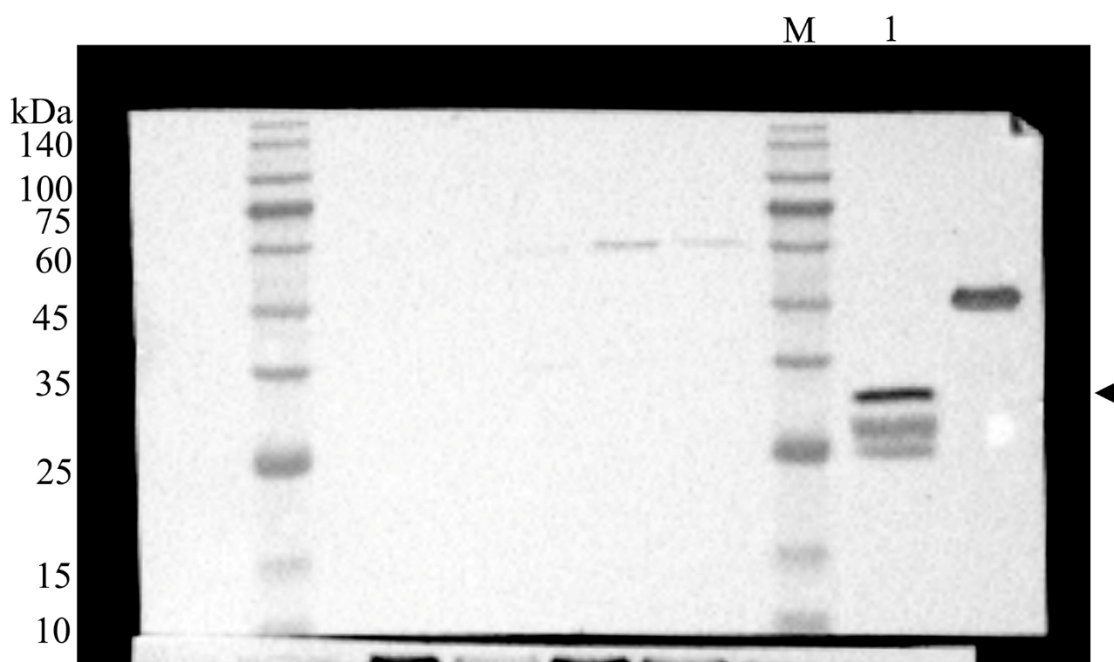
**Figure S1.** Optimized nucleotide and amino acid sequences of the mature CcLys2. The total length of the optimized gene is 633 bp, and red marks the mutated bases. The recognition sites for *Nde* I and *Eco*R I are underlined, and the red arrow marks the thrombin cleavage site. The N-terminal of the amino acid sequence contains a 6×His-tag. The asterisk marks the stop codon.



**Figure S2.** Identification of the recombinant expression vector by PCR. M: DL2000 DNA marker; Lanes 1 and 2: the *CcLys2* gene with a part of the pET-28b(+) vector sequence.

**Table S1.** Species name, protein name and GenBank accession number for multiple sequence alignment and phylogenetic analysis.

Protein name	Species name	GenBank accession number
CcLys2	<i>Coridius chinensis</i>	QPB70596
CcLLP3	<i>Coridius chinensis</i>	MW250790
HEWL	<i>Gallus gallus domesticus</i>	1LSZ_A
HLYZ1	<i>Homo sapiens</i>	NP_000230
PsLys	<i>Plautia stali</i>	BBE08150
HhLys	<i>Halyomorpha halys</i>	KAE8573139
TcLys1	<i>Tribolium castaneum</i>	ACV32411
TcLys2	<i>Tribolium castaneum</i>	ACV32412
HaLLP1	<i>Harmonia axyridis</i>	AIZ72680
HaLLP2	<i>Harmonia axyridis</i>	AIZ72679
HaLys3	<i>Harmonia axyridis</i>	AIZ72681
HaLys4	<i>Harmonia axyridis</i>	AIZ72682
AgLYSC1	<i>Anopheles gambiae</i>	Q17005
AgLys c-6	<i>Anopheles gambiae</i>	AAY21238
AgLys c-8	<i>Anopheles gambiae</i>	AAY21241
CqLys	<i>Culex quinquefasciatus</i>	XP_001843429
FoLys	<i>Frankliniella occidentalis</i>	KAE8737425
NilLys	<i>Nilaparvata lugens</i>	AGK40894
BmLZ	<i>Bombyx mori</i>	NP_001037448
BLLP1	<i>Bombyx mori</i>	ABP52099
ALLP1	<i>Antheraea mylitta</i>	ABP52098
CfLys3	<i>Coptotermes formosanus</i>	AFZ78838
SeLys3	<i>Spodoptera exigua</i>	AKJ54522
OfLys5	<i>Ostrinia furnacalis</i>	AGV28584



**Figure S3.** M: Protein molecular weight marker (10-140 kDa). Lane 1: blotting band of fusion CcLys2. The arrow indicates the band of fusion CcLys2.