#### **Supplementary Information**

## ALCAM is an Entry Factor for Severe Community Acquired Pneumoniaassociated Human Adenovirus Species B

Yusang Xie<sup>1\*</sup>, Hong Mei<sup>2\*</sup>, Wei Wang<sup>2</sup>, Xiao Li<sup>3</sup>, Pengfei Hu<sup>2</sup>, Xingui Tian<sup>3</sup>, Rong Zhou<sup>3, 4</sup>, Jia Liu<sup>2, 4, 5, #</sup>, Jieming Qu<sup>1, #</sup>

<sup>1</sup>Department of Pulmonary and Critical Care Medicine, Ruijin Hospital, Institutes of Respiratory Diseases, School of Medicine, Shanghai Jiao Tong University and Shanghai Key Laboratory of Emergency Prevention, Diagnosis and Treatment of Respiratory Infectious Diseases, Shanghai 200025, China <sup>2</sup>Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, P. R. China <sup>3</sup>State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou Medical University, Guangzhou 510182, China <sup>4</sup>Guangzhou National Laboratory, Guangzhou International Bio Island, No. 9 XingDaoHuanBei Road, Guangzhou, 510005, Guangdong, China. <sup>5</sup>Shanghai Clinical Research and Trial Center, Shanghai, 201210, China.

\*These authors contributed equally to this work.

Correspondence should be addressed to J.Q. (jmqu0906@163.com), J.L. (liujia@shanghaitech.edu.cn)

#### **Supplementary Figures**



**Supplementary Figure 1 Identification of SCAP-associated HAdV-B7. a**, Flowchart showing the procedure of multi-centre prospective study. The data are extracted from a previous publication<sup>8</sup> and re-analyzed and the figure re-drawn. **b**, Architecture of the genomic sequence of HAdV-B7. Fiber protein is encoded by L5 gene. **c**, Nucleotide and amino acid sequence alignment of fiber proteins in publicly available HAdV-B7 sequences. The alignment results are visualized using Highlighter online tools (https://www.hiv.lanl.gov/content/index). Source data are provided as a Source Data file.



**Supplementary Figure 2 Construction and validation of recombinant HAdV-C5/B7-GFP. a**, Flow chart showing cloning and production procedure of HAdV-C5/B7-GFP. **b-c**, Flow cytometry analysis of the EGFP positive ratio of WT, CAR knockout and DSG-2 knockout cells at 48 h post HAdV-C5 (**b**) or HAdV-C5/B7 (**c**) infection. For **b-c**, data are presented as mean  $\pm$  SD (n = 3) from three independent biological replicates. The significant difference is analyzed by two-tailed unpaired Student's *t* test. For **b**, the *p* values between GFP positive cells (%) of Nontarget and *CAR KO-1* and *CAR KO-2* are 6e-5 and 1e-5, respectively. Source data are provided as a Source Data file.



**Supplementary Figure 3 Analyses of the knockout efficiencies of CAR, DSG-2, ALCAM and CADM1 sgRNAs in HEK-293A. a**, Determination of the knockout efficiencies in CAR knockout and DSG-2 knockout cells by TIDE analysis of Sanger sequencing results. **b-c**, Flow cytometry analysis of cell surface expression of ALCAM and DSG-2 proteins in *ALCAM* knockout (**b**) and *DSG-2* knockout (**c**) cells respectively (as mixed population). Negative controls are mock staining with PBS. **d**, Evaluation of ALCAM and CADM1 knockout efficiencies, as determined by TIDE analysis of Sanger sequencing results. Source data are provided as a Source Data file.



Supplementary Figure 4 The design of sgRNAs in SfCRISPR-v2. a, Flow chart showing the sgRNA design.

**b-c**, The distribution of sgRNA on-target (**b**) and off-target (**c**) scores. Source data are provided as a Source Data file.



**Supplementary Figure 5 Quality validation of constructed SfCRISPR-v1 and -v2 as pooled plasmids or pooled cells. a**, Summary of the NGS analysis results of sgRNAs in constructed plasmid and HEK-293T cell libraries. **b-c**, Distribution of sgRNAs in SfCRISPR-v1 and -v2 plasmid (**b**) or HEK-293T cells (**c**) libraries. **d**, The fold change of non-targeting sgRNAs and sgRNAs targeting essential genes and nonessential genes in SfCRISPR-v1 and -v2 HEK-293T cell libraries. The sgRNA changes at 5 and 14 days after LV transduction are compared and shown as fold changes. The box plot illustrates the distribution of the data. The thick line within the box represents the median. The top and bottom edges of the box indicate the 25th percentile (lower quartile, Q1) and the 75th percentile (upper quartile, Q3), encompassing the interquartile range (IQR), which includes the middle 50% of the data. The vertical lines (whiskers) extend to the smallest and largest data points within Q1 - 1.5 \* IQR and Q3 + 1.5 \* IQR, respectively. Data points beyond this range are shown as individual dots, representing outliers. Statistical analysis is performed using Student's *t* test. Source data are provided as a Source Data file.



Supplementary Figure 6 Multi-centre prospective study of SCAP patients in China. a, Information of the five SCAP patients with HAdV infection. b, Representative axial and coronal planes of layered computer tomography images of Patient 1 on admission. The arrows dictate the consolidation in multiple lung lobes. The experiment is repeated three times independently and similar results are obtained. c, Analysis of CADM1expression in blood samples from SCAP patients with different pathogen infection, as determined by RNA-Seq. All available blood samples of HAdV-infected patients (5 in total) are analyzed. N.A., no available identifiable pathogen information. Others, all pathogens excluding HAdV. All, all analyzed SCAP patients. Data are presented as mean  $\pm$  SD ( $n \ge 3$ ) from different patients. The significant difference between HAdV and other pathogen groups is analyzed by two-tailed unpaired Student's t test. Source data are provided as a Source Data file.



Supplementary Figure 7 Validation of ALCAM function for HAdV-C5/B7-GFP infection in HEK-293A. a, RT-qPCR quantification of ALCAM expression in ALCAM knockdown cells treated with ALCAM siRNA-1, 2 and 3. GAPDH is used as an internal control. **b-c** (related to Fig. 2), Flow cytometry analyses (**b**) and representive fluorescence images (**c**) of WT and ALCAM knockdown cells at 48 h post HAdV-C5/B7-GFP infection. MFI, mean fluorescence intensity. Scale bar, 100  $\mu$ m. **d**, Proliferation of WT, non-targeting sgRNA and ALCAM knockout HEK-293A cells at 1, 3 and 5 day post seeding, quantified by CCK-8 assay. **e-g**, Characterization of *ALCAM*<sup>-/-</sup> single clone of HEK-293A by Sanger sequencing analysis of mutated alleles (**e**), western blot analysis of total ALCAM protein (**f**) and flow cytometry analysis of cell surface ALCAM protein (**g**). For **e**, the 20-bp CRISPR-Cas9 targeting sequence is highlighted in red and protospacer adjacent motif (PAM) underlined. **h**, Proliferation of WT, *ALCAM*<sup>-/-</sup> and overexpression-rescued *ALCAM*<sup>-/-</sup> HEK-293A cells at 1, 3 and 5 days post seeding, as quantified by CCK-8 assay. **i**, Representive bright-field and fluorescence

images of WT, knockout and overexpression-rescued cells at 48 h post HAdV-C5/B7-GFP infection. Scale bar, 100  $\mu$ m. For **a-b**, **d** and **h**, data are presented as mean +/- SD ( $n \ge 3$ ) from independent biological replicates. The significant difference is analyzed by two-tailed unpaired Student's *t* test. For **b**, the *p* value between siRNA-NC and *ALCAM*-siRNA-1 is 7e-5. For **d**, the *p* value between WT and *ALCAM* KO-mix at 1, 3 and 5 day post seeding is 4e-5. Source data are provided as a Source Data file.



Supplementary Figure 8. Comparison of fiber protein sequences and the infection efficiencies of chimeric HAdV-B and clinical HAdV-B3 and HAdV-B7 (related to Fig. 3). a, Amino acid sequence alignment of the fiber domains of C5-based chimeric HAdV-B3, B7, B14, B35 and B50. b, Flow cytometry analysis of the GFP positive ratio of non-targeting and CAR knockout HEK-293T cells at 48 h post infection with HAdV-C5 or C5-based chimeric HAdV-B3, -B11, -B14, -B35 and -B50. Data are presented as mean  $\pm$  SD (n = 3) from three independent biological replicates. The significant difference between non-targeting and knockout groups is analyzed by two-tailed unpaired Student's *t* test. **c-f**, Sanger sequencing results of the fiber sequences of clinical HAdV-B3, HAdV-B7, HAdV-B14 and HAdV-B35 used in this study for comparison with the deposited sequences in NCBI database. Source data are provided as a Source Data file.



**Supplementary Figure 9 Analysis of the dependency of HAdV-B infection on ALCAM expression in multiple cell lines. (related to Fig. 3). a**, RT-qPCR quantification of *ALCAM* mRNA expression in multiple human cell lines. **b**, RT-qPCR quantification of viral L5 DNA in different cells at 48 h post infection with clinical HAdV-B7. **a-b**, The significant difference between HeLa and other cells is analyzed. **c-d**, RT-qPCR (**c**) and flow cytometry (**d**) analyses of *ALCAM* mRNA and protein expression in WT and *ALCAM*-overexpressed HeLa cells. For flow cytometry analysis, cells are stained with anti-ALCAM antibody, and negative control is mock staining with PBS. **e**, The effects of *ALCAM* overexpression on HAdV-C5/B7-GFP and HAdV-C5-GFP infection in HeLa cells, as evaluated by RT-qPCR quantification of GFP mRNA. **f**, Protein sequence alignment of human ALCAM and CHO ALCAM, as analyzed by ClustalW. **g**, *Alcam* mRNA expression in CHO-K1, as determined by RT-qPCR. **h**, Clinical HAdV-B7 infection in CHO-K1 and HEK-293A cells., as determined by qPCR. **i**. Hamster *Alcam* mRNA expression in WT, *ALCAM*<sup>-/-</sup> and hamster's *Alcam*-overexpressed *ALCAM*<sup>-/-</sup> HEK-293A cells, as determined by RT-qPCR. **j**. Clinical HAdV-B7 infection in WT, *ALCAM*<sup>-/-</sup> HEK-293A cells, as determined by RT-qPCR. **j**. Clinical HAdV-B7 infection

293A cells, as determined by qPCR. For **a-c**, **e** and **g-j**, rplp0 is used as an internal control, and data are presented as mean  $\pm$  SD (n = 3) from three independent biological replicates, and the significant difference is analyzed by two-tailed unpaired Student's *t* test. For **a**, the *p* values between Hela and HEK293A, A549 and U-87MG are 3e-5, 3e-5 and 3e-7, respectively. For **b**, the *p* values between Hela and A549 and U-87MG are 5e-6 and 2e-5, respectively. For **c**, the *p* value is 5e-5. For **e**, the *p* value between WT and ALCAM-overexpressed HeLa cells after HAdV-C5/B7-GFP infection is 8e-5. For **g**, the *p* value is 6e-6. For **h**, the *p* value is 1e-5. For **i**, the *p* value between ALCAM<sup>-/-</sup> and Alcam-overexpressed ALCAM<sup>-/-</sup> cells is 6e-6. Source data are provided as a Source Data file.



Supplementary Figure 10 Dissection of the pattern of interaction between ALCAM and HAdV-B fiber proteins in HEK-293T cells (related to Fig. 4). a, Schematic diagram showing the procedure of attachment and internalization assays. b, Co-IP analysis of the interactions between ALCAM-myc and HAdV-fiber-HA or EGFP-HA. c, Co-IP analysis of the interaction of HAdV7-fiber-HA with ALCAM-myc or DSG-2-myc. d, Co-IP analysis of the interaction of ALCAM-myc with different domains of fiber protein, including the fulllength fiber-HA, shaft-HA, knob-HA, tail-shaft-HA, or shaft-knob-HA. e. Co-IP analysis of the interaction between ALCAM-C-domain-myc or ALCAM-V-domain-myc and HAdV7-fiber-HA, shaft-HA, knob-HA, tail-shaft-HA or shaft-knob-HA. For d-e, immunoprecipitation is conducted using anti-HA beads. For b-e, WCL, whole cell lysate. All Co-IP analysis is performed with denatured gels. The experiment is repeated three times independently and similar results are obtained. Source data are provided as a Source Data file.



Supplementary Figure 11 Characterization of gene knockout and knockdown cells. a, RT-qPCR quantification of DSG-2 mRNA expression in non-targeting sgRNA and ALCAM<sup>-/-</sup> HEK-293A cells. b, RT-qPCR quantification of ALCAM mRNA expression in non-targeting and *DSG-2* knockout cells. c-d, Flow cytometry analysis of ALCAM (c) and DSG-2 (d) protein expression on cell surface. e, Evaluation of the knockout efficiencies of EndoA3 sgRNAs in HEK-293A by TIDE analysis of Sanger sequencing results. f, RT-qPCR quantification of *ALCAM* mRNA expression in non-targeting, *ALCAM* knockdown, *DSG-2* knockout, *ALCAM* knockdown/*DSG-2* knockout cells. For a-b and f, Rplp0 is used as an internal control. Data are presented as mean  $\pm$  SD (n = 3) from three independent biological replicates. The significant difference between non-targeting and knockout groups is analyzed by two-tailed unpaired Student's *t* test. Source data are provided as a Source Data file.



Supplementary Figure 12. Characterization of the interaction between ALCAM and HAdV-B7 knob domain (related to Fig. 5). a, Expression and purification of His-tagged knob domain from SF9 cells. Arrow dictate target protein band. b, Identification of protein sequence of purified knob domain using mass spectrometry. An identified unique peptide of knob domain, GFMPSTTAYPFNVNSR, is shown. c-f, Fitted curves for SPR measurements of the interactions between fiber knob and ALCAM or DSG-2. c, ALCAM (ligand) and knob (analyte) (biological replicate to Fig. 5c-d). d, Knob (ligand) against ALCAM (analyte). e, Knob (ligand) against DSG-2 (ligand). f, Knob (ligand) against knob (analyte). g-i, ALCAM (ligand) (h), DSG-2 (ligand) (i) or knob (ligand) (j) against an irrelevant protein LbCas12 (analyte). Curve-fitting is performed using a steady-state affinity method. Source data are provided as a Source Data file.

## Supplementary Tables

	Sequences
NGS-F	5'-tctttccctacacgacgctcttccgatctccgtaacttgaaagtatttcga-3'
NGS-R	5'-gtgactggagttcagacgtgtgctcttccgatctctttttcaagttgataacggac-3'

Supplementary	v Table 1	. Primers fo	or søRNA	NGS an	alvses
Supprementary		• • • • • • • • • •	1 551111	1105 all	aryses

#### HAdV-B3

#### HAdV-B7

#### HAdV-B11

#### HAdV-B14

#### HAdV-B35

gacggagttettaetttaaaatgtttaaccecactaacaaccacaggeggatetetacagetaaaagtgggagggggaettacagtggatgacaetgatgg taeettacaagaaaacataegtgetacageacceattaetaaaaataateaetetgtagaactateeattggaaatggattagaaaeteaaaacaataaactat gtgecaaattgggaaatgggttaaaatttaacaaeggtgacatttgtataaaggatagtattaacaeettatggaetggaataaaeetecaaetgteaa attgtggaaaaeaetaataeaatgatggeaaaettaetttagtattagtaaaaaatggagggggttgttaaggetaegtgtetetagttggtgtateagaeaet 

#### HAdV-B50

Genes	Forward primers	<b>Reverse primers</b>
Nontarget sgRNA	caccgacggaggctaagcgtcgcaa	aaacttgcgacgcttagcctccgtc
CADM1-1	caccggatgctgaaggtgcacaagg	aaacccttgtgcaccttcagcatcc
CADM1-2	caccggatgctgaaggtgcacaagg	aaaccettgtgcacettcagcatec
ALCAM-1	caccgtgtgtgcatgctagtaactg	aaaccagttactagcatgcacacac
ALCAM-2	caccggattctgaggtacgtcaagt	aaacacttgacgtacctcagaatcc
DSG-2-1	caccgtagagetacccagtttacag	aaacctgtaaactgggtagctctac
DSG-2-2	caccgtagagetacccagtttacag	aaacctgtaaactgggtagctctac
CAR-1	caccgtacgcttagtcccgaagaccc	aaacggtcttcgggactaagcgtac
CAR-2	caccgacgcttagtcccgaagacca	aaactggtcttcgggactaagcgtc
EndoA3-1	caccggaaatacgggaaggagctcg	aaaccgagctccttcccgtatttcc
EndoA3-2	caccgacactgtgtcgaagatccga	aaactcggatcttcgacacagtgtc

Supplementary Table 3. Primers for construction of sgRNA plasmid

### Supplementary Table 4. Primers for PCR amplification of sgRNA targeted sites for Sanger sequencing

Genes	Forward primers	<b>Reverse primers</b>
<i>CADM1-1/2</i>	atggcgagtgtagtgctgc	gatcactgtcacgtctttcgt
ALCAM-1	cctgccgtgatggttatctga	tagttcctcactgcaaatctcct
ALCAM-2	agatgacaaccttcatcgaca	ccctccattttcagacccact
DSG-2-1	agctgagactgtttgggctt	catgtctctggaagcccctg
DSG-2-2	tccctccatcctcctgactc	gageteatecetaceaagee
<i>CAR</i> -1/2	ggactccattcctcgggacc	gtctcaacgtcatgcctgct
EndoA3-1	acttagcacccaccttggtaa	gctccaataaatgcaggagctt
EndoA3-2	acttagcacccaccttggtaa	gctccaataaatgcaggagctt

#### analyses

Genes	Forward primers	<b>Reverse primers</b>
$\beta$ -actin (human)	gtctgccttggtagtggataatg	tcgaggacgccctatcatgg
RPLP0 (human)	agcccagaacactggtctc	actcaggatttcaatggtgcc
RPLP0 (hamster)	agtgccacactctatcatca	gggcagcagccacaaa
EGFP	gaagaacggcatcaaggt	gctcaggtagtggttgtc
HAdV-C5-L3	acgatgacaacgaagacgaagtag	ggcgcctgcccaaatac
HAdV-C5-L5	taatgtagcaggaggactaag	atcaagtataaggcgtctgt
HAdV-B7-L3	gcaatggtcgttatgtgcct	gcaggaccatgttcacatcc
HAdV-B7-L5	caggcgggtctctacagtta	gggtcctagcgacaaaccta
HAdV-B3-L3	tacaacgtggcacaatgcaa	aggtgacggctttgtagtca
HAdV-B14-L5	agcccagacggagttcttac	ccggctcctagggataaacc
HAdV-B35-L5	atgaaagcacctcccaacac	agatccgcctgtggttgtta
DSG-2	ctaacaggttacgctttggatgc	gtgaacactggttcgttgtcat
ALCAM	tcctgccgtctgctcttct	ttctgaggtacgtcaagtcgg
ALCAM	agcaccctggagtacaagac	ctccttgatggcgttcttgg
(Codon optimized)		
ALCAM (hamster)	atcgagcccatccttttcta	gtgtcagtgactcccttttc
ALCAM (hamster Codon optimized)	ctgagcctgagcgagaacta	acgaatctcttctcgtcgct

## Supplementary Table 5. Primers for qPCR

Supplementary Table 6. siRNA sequences

Genes	Sense sequences (5'-3')	Anti-sense sequences (5'-3')
Nontarget	uucuccgaacgugucacgutt	acgugacacguucggagaatt
ALCAM-siRNA-1	gccuucagauccucuacaatt	uuguagaggaucugaaggctt
ALCAM-siRNA-2	ccaaggcugacauacaaautt	auuuguaugucagccuuggtt
ALCAM-siRNA-3	cucgguaauauccaagaaatt	uuucuuggauauuaccgagtt

#### ALCAM-myc-flag

gcctacggcgacaccatcattatcccctgcagactggacgtgcctcagaacctgatgttcggcaagtggaagtacgagaagcccgacggcagccccgtgttcatcgccttcagaagcagcacaaagaagagcgtgcagtacgacgacgtgcccgagtacaaggacagactgaacctgagcgagaactacaccctg agcagcctagcaagcccgagatcgtgagcaaggccctgttcctggagaccgagcagctgaagaagctgggcgactgcatcagcgaggacagctaccacagetgtacaccatgacaageaceetggagtacaagaceaceaaggeegacatteagatgeeetteacetgeagegtgacetaetaeggeeetageg ggcagaagaccatccacagcgagcaagccgtgttcgacatctactaccccaccgagcaagtgaccatccaagtgctgccccccaagaacgccatcaa ggagggcgacaacatcaccctgaagtgcctgggcaaccggcaacccccctcccgaggagttcctgttctacctgcccgggcagcccgagggcatccgg agcagcaacacctacaccctgaccgacgtgagaagaaacgctaccggcgattacaagtgcagcctgatcgacaagaagagcatgatcgctagcaccgccatcaccgtgcactacctggacctgaaccctagcggcgaggtgacaagacagatcggcgacgccctgcccgtgagctgcaccatcagcgc gagtagaaacgccaccgtagtgtggatgaaggacaacatcagactgagaagcagccctagcttcagcagcctgcactaccaagacgccggcaactac aagaccgaccctagcggcctgagcaagaccatcatctgccacgtggagggcttccccaagcccgccattcagtggaccatcaccggcagcggcagcg tgatcaatcagaccgaggagagcccctacatcaacggcagatactacagcaagatcattatcagccccgaggagaacgtgaccctgacctgcaccgcc gagaatcagctggagagaaccgtgaacagcctgaacgtgagcgccatcagcatccccgagcacgacgacgacgacgacgacgacaaaatag agaagagcaagaccgctagcaagcacgtgaacaaggacctgggcaacatggaggagaacaagaagctggaggagaataatcacaagaccgaggc cacgcggccggagcagaaactcatctcagaagaggatctggcagcaaatgatatcctggattacaaggatgacgacgataaggtttaa

#### ALCAM C domain-myc-flag

#### ALCAM V domain-myc-flag

# Supplementary Table 8. Human codon-optimized hamster *ALCAM* gene for overexpression experiments

#### Hamster ALCAM-myc-flag

gcctacggcgacaccatcgtgatgccctgcagactggacgtgcctcagaacctgatgttcggcaagtggaagtacgagaagcccgacggcagccccg tgttcatcgccttcagaagcagcacaaaaaagagcgtgcagtacgacgacgtgcccgagtacaaggacagactgagcctgagcgagaactacaccctg agcatcaacaacgccaagatcagcgacgagaagagattcgtgtgcatgctggtgaccgaggacaacgtgttcgaggcccccaccctggtgaaggtgtt acacagctgtacaccatgacaagcctggagtacaagaccaccaaggccgacattcagatgcccttcacctgcagcgtgacctactacggccctagc gggcagaagaccatccacagcgagcagaccgtgttcgacatctactaccccaccgagcaagtgaccatccaagtgctgccccccaagaacgccatca aggagggggacaacatcaccctgcagtgcctgggcaacggcaacccccctcccgaggagttcatgttctacctgcccgggcagcccgagggcatccg gagcagcaacacctacaccctgaccgacgtgagacggaacgctaccggcgactacaagtgcagcctgatcgacaagaagagcatgatcgctagcacc accatcaccgtgcactacctggacctgaaccctagcggcgaggtgaccaagcagatcggcgacgccctgcccgtgagctgcaccatcagcg ctagcagaaacgccaccgtggtctggatgaaggacaacatcagactgagaagcagccctagcttcagcagcctgcactaccaagacgccggcaacta gaagaccgaccctagcggcctgagcaagaccatcatctgccacgtggagggcttccccaagcccgccattcagtggaccatcaccggcagcggcagc gtgatcaatcagaccgaggagagcccctacatcaacggcagatactacagcaagatcattatcagccccgaggagaacgtgaccctgacctgcgccgc ctgggcaacatggaggagaacaaaaagctggaggagaataatcacaagaccgaggccacgcgggccggagcagaaactcatctcagaagaggatctg gcagcaaatgatatcctggattacaaggatgacgacgataaggtttaa

#### Fiber-HA

#### Knob-HA

#### Shaft-HA

atggatggcgtgctgaccctgaagtgcctgacacctctgaccacaaccggcggcagcctgcagctgaaagtgggaggcggactgaccatcgatgaca ccgacggcttcctcaaggaaaacattagcgccacgaccccactggtgaaaaccggccacagcatcggcctgtctctcggccctggcctggaaaccaat gagaacaagctgtgcgccaagctgggcgaaggtctgacattcaacagcaacaacatctgcatcaacgacaacattaacacatatccttacgacgtgcctg actacgcctga

#### Shaft-Knob-HA

atggatggcgtgctgaccctgaagtgcctgacacctctgaccacaaccggcggcagcctgcagctgaaagtgggaggcggactgaccatcgatgaca ccgacggcttcctcaaggaaaacattagcgccacgaccccactggtgaaaaccggccacagcatcggcctgtctctcggccctggcctggaaaccaat gagaacaagctgtgcgccaagctgggcgaaggtctgacattcaacagcaacaacatctgcatcaacgacaacattaacaacatgtggaccggagttaat cccaccagagccaactgccagatcatggcttcttccgagagcaatgactgcaagctgatcctgacactggtcaagaccggcgcactggtgaccgcttttg tgtacgtgatcggcgtgtccaacgacttcaacatgctgacgaccataagaacatcaacttcaccgccgagctgttcttcgactctacaggcaacctgctga ccagcctgagcagcctgaagacactctgaatcacaaaagcggacagaacatggccacaggcgccctgaccaacggcgtactagcctgacca cacagcctgagcagcctgaagacactctgaacaacaggaaaaaggaaaactacatctacggcaacctgttactacaccgccagcgaccaacagccttccccatcgac atcagcgtgatgctgaaccagaggacctgaacaacgagacaagctactgtatccgcgtgacatggtcctggaacaccggcgtggcccctgaggtgca gaccagtgctaccaccctggtcacatctccattcacctcacacagagaggatgattatccttacgacgtgcctgactacgcctga

#### HA-Tail-Shaft

#### His-Knob

atgcaccaccaccaccaccaccaccaccacctggaagtcctgttccagggtcccctgtggactggtgtgaaccctaccagggctaactgccagatc atggcttccagcgagtccaacgactgcaagctgatcctgatctggtgaagaccggcgctctggtcaccgctttcgtctacgtgatcggcgtcagcaacg acttcaacatgctgaccactcacaagaacatcaacttcactgctgaactgttcttcgactccactggcaacctgctgacttcctgtccagcctgaagactcc cctgaaccacaagagcggccagaacatggccaccggtgctttgactaacgctaagggtttcatgcctagcactactgcctaccctttcaacgtgaacagc gcggaaaaggagaactacatctacggcacctgctactacactgcctccgaccaccgcgttgctccagagtgccagaaccagcgcgcc ctgaacaacgaaaccagctactgcatccgtgtcacctggagctggaacaccggcgttgctcctgaggtccagacttccgtgaaccagcggcc ctgaacaacgaaaccagctactgcatccgtgtcacctggagctggaacaccggcgttgctcctgaggtccagacttccgctaccacctggtgactagcc ctttcaccttctactacatcgggaggacgactaa