

Association of a new 99-bp indel of the *CEL* gene promoter region with phenotypic traits in chickens

Xiangnan Wang[#] · Xinlei Wang[#] · Bingjie Chen · Yaping Guo · Hehe Tang · Donghua Li · Danli Liu · Yanbin Wang · Guoxi Li · Xiangtao Kang · Zhuanjian Li* & Ruili Han*

College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou, 450046, China

[#]: Xiangnan Wang and Xinlei Wang contributed equally to this work.

*: Correspondence Author: Zhuanjian Li and Ruili Han

Tel: +86-18037465363 and +86-15188312192

E-mail: lizhuanjian@163.com and rlhan@126.com

Address: College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou, 450046, China

Acknowledgments

This study was funded by NSFC-Henan joint grant (U1804107), Program for Innovation Research Team of Ministry of Education (IRT16R23), and the Tackle Key Problem of Science and Technology Project of Henan Province (182102110351).

```

1 M A H W A I L S F A L C C C L G V A Q A
1 atggetcaactggggcgattctgagctttgceettgtgetgetgeecteggggtagcaacaggee
21 A T L G V V L T E G C F V E G E S X R R
21 gcaactctgggtgtgggtgetcaecaggggaggtttttgtggaaggcgagagtagaaacgaegg
41 G L F G S Y V D I F R G I P F A A P P K
121 ggactctttgggagctatgtggatattctcaagagggatcccccttttgetgeccccgecaaaag
141 A L Q D P Q P H P G W D G T L K A X K F
181 geactgcaagaccccccaacctcaatcctgggtggggaacactgaaagcaaaaaatttt
81 K N R C M Q M T L T Q T D V R G K E D C
241 aagaatcgetgcatgcagatgacaacttaccocaaactgatgtccgtggggaaggaggactgc
101 L Y L N I W I P Q G K R E V S T N L P V
301 ctctatctgaaacatctggatcccccaaggggaagagagaagtctcccaaaacttgcacgtg
121 M V W I Y G C A F L L G C G C Q C A N F L
361 atgggtctggatctcaaggtgggtgcecttccctcttggagggggtcaaggagacaaacttccct
141 D N Y L Y D G E E I A V R G N V I V V T
421 gacaaactaacctctatgaggtgaggagatgcecggtgceggggcaatgtgattgtggtgacc
161 L N Y R V G P L G F L S T G D P N M P G
481 ctcaactatcgtgtggggccccctgggtctccctcaagcaactggagaccccccaatgcccaggg
181 N Y G L K D Q H M A I A W V K R N I K A
541 aactacggggtgaaggatcaagcaatggctattgcectgggtgaagaggaaatataaagggee
201 F C G C D P D N I T I F G E S A G A A S V
601 tttggagggegaccccagacaaacatcaacatctttggggagtcagctgggtgctgcccagtgctc
221 S L Q I L S P K N A C G L F K R A I S Q S
661 tccctgcaagatattgtcccccaagaaagcaggtctgttcaagagagccatcaagcaaaagc
241 C V S L C S W V I Q K D P L T W A K K V
721 ggtgtcagctctgtgcaagctgggtcaatcccccaaggaaccccaactcaettgggtcaaaaaggtt
261 G E Q V G C P T D N T T V L A N C L R A
781 ggagagcaggtgggtgcccccaagacaaaccccaaggtcttgggcaaaactgtctccgtgccc
281 T D P K A L T L A H H V E L I S L P G P
841 actgacccccaaagccctgacaactggcccccaacagctggaaactgatctccctgcectggctccc
301 L V H T L S I T P V V D G D F L P D M P
901 ctgggtcacaactctccatcaactcctctgtgttgatggagacttccctccctgcaatgccc
321 E N A F A N A A D I D Y I A G V N N M D
961 gagaacctctttgcccattgctgctgacatogaactacattgctgggggtcaaaacacatggat
341 G H F F A G F D L P A I N R P L Q K I T
1021 ggacattcttttgetggctttgatttaacctgctatcaaacctgcccacttccagaaaaatcaact
361 A S D V Y N L V K G L T A D R G E R G A
1081 gcgagcgatgtctataaacttgggtcaaaaggactaaactgcagacaggggtgagagaggagccc
381 N L T Y D L Y T E L W G D N P E Q Q V M
1141 aaactgacgtacgatctctacaacagagttgtgggggtgacaaacccagagcaaaagctcatg
401 K R T V V D L A T D Y I S L I P T Q W T
1201 aaagagaaacagtggtggacctggctaccgaactacatttccctgatcccccaacagtggaaca
421 L N L H H K N A R S C K T Y S Y L F S Q
1261 ctaaaaactgcaaccacaagaatgccccggagtggaagaacatacagetaacttgttctccccag
441 P S R M P I Y P S W V G A D H A D D L Q
1321 ccattctgaatgcccattctatccaaagctgggttaggggcagacccaagctgatgaacttgcaag
461 Y V F G K P F A T P L G Y L P K H R T V
1381 tacgtgtttgggaaaaacccctttgcccacccctctaggttaacctgcccacagcaacaggaacgctc
481 S S A M I A Y W T N F A R T G D P N S G
1441 tcaatctgcccattgattgcttattggaccccaattttgcccaggaactggggaaccccccaagtggg
501 N S E V P I T W P P Y T T E G C G Y Y L E
1501 aattcaagaggtgccccattacctggcccacccctacacccactgagggtgggttaactacctggaa
521 I N X I N S V X Q N L R T P Y V N
1561 atcaacaaacaaaataaaactataaattcaagtgaaaacagaatctgagaaacccccatattgtgaaac
541 Y W N S V Y L N L P L I A S T S *
1621 taactggaattcaagctctatctaaaatctgcccactgattgcccagcaatcctag

```

Figure S1. CDS and deduced amino acid sequence for chicken CEL. Nucleotides are shown in lowercase letters, and amino acids are shown in capital letters above the first nucleotide in each codon triplet. Shaded areas are mutation sites.

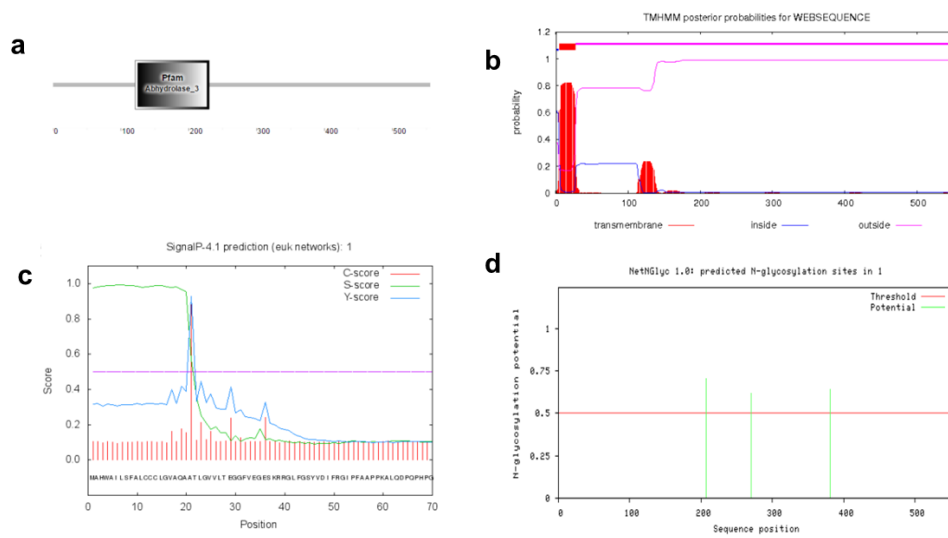


Figure S2. Bioinformatics analyses of CEL amino acid sequences. (a) The secondary structure of the CEL amino acid sequences. (b) TMHMM is the predicted transmembrane region of the CEL amino acid sequence. (c) Results of the CEL signal peptide P4.1 analysis. (d) NetNGlyc 1.0 predicts the glycosylation site of the CEL amino acid sequence.

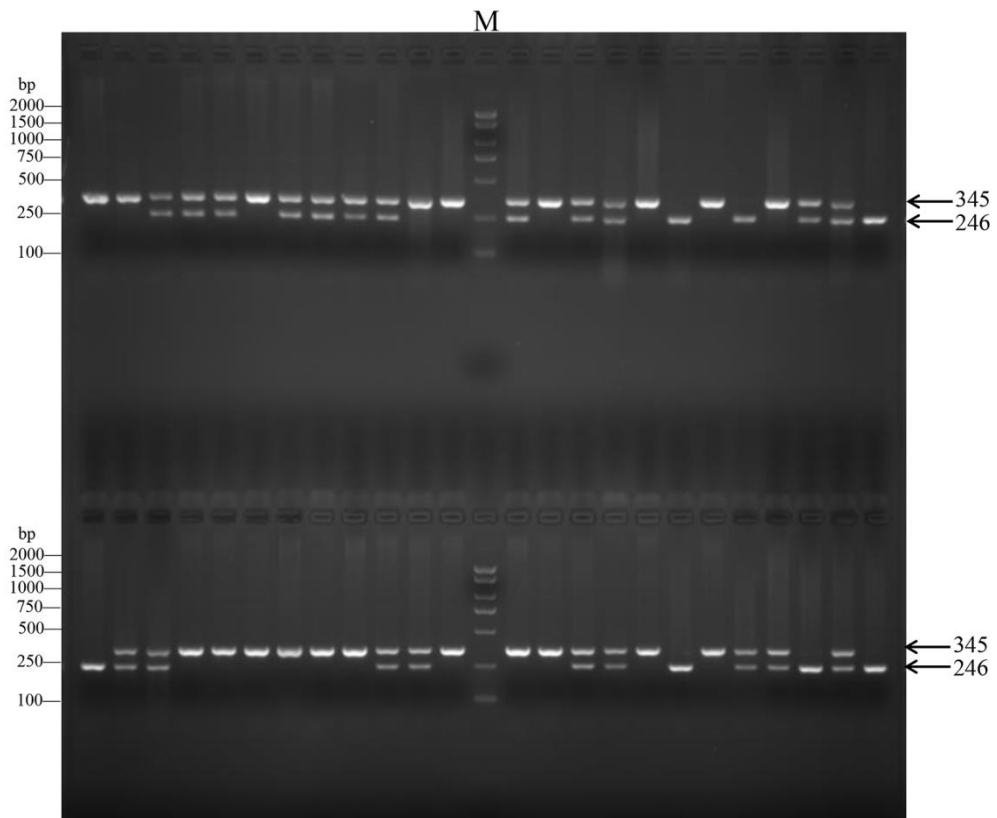


Figure S3. Electrophoresis pattern of the 99-bp indel locus within the chicken *CEL* gene (genotyped by 2.0% agarose gel electrophoresis).

Allele I

GTCACTCAGAGGCTCCCTTTATGGCACGTAACATATATATATGAATAAATATACACACACAGTGTGATCCAGCT
CCTCCTAAAAGTATGCATCTAAAAATAGGACTGAGAAGTCATGTCCTTAATGCACCTTGTCTCCCACCGATCAG
AAAGGGAGCCTAGCACATACAGATATCTAAAGTTA^{cgtgggatgagtcgactcttgggtctctgactgctctcaggctacgttctcacagct}
ggttctgttccagtcacagcctccagctctcctctgCTCTCCATGCAGAGAGGTTGCTCATGGCCATTTAAATTTCTGTGTAICG
TGGGTTCTAGCCA

Allele D

GTCACTCAGAGGCTCCCTTTATGGCACGTAACATATATATATGAATAAATATACACACACAGTGTGATCCAGCT
CCTCCTAAAAGTATGCATCTAAAAATAGGACTGAGAAGTCATGTCCTTAATGCACCTTGTCTCCCACCGATCAG
AAAGGGAGCCTAGCACATACAGATATCTAAAGTTA^{CTCTCCATGCAGAGAGGTTGCTCATGGCCATTTAAATTT}
CTCTGTGTAICGTTGGGTTCTAGCCA

Figure S4. The sequences of alleles I and D. Note: The underlined sequence represents the primer sequence, the lower case base represents the 99-bp insertion sequence, and the blue triangle represents the mutation loci.

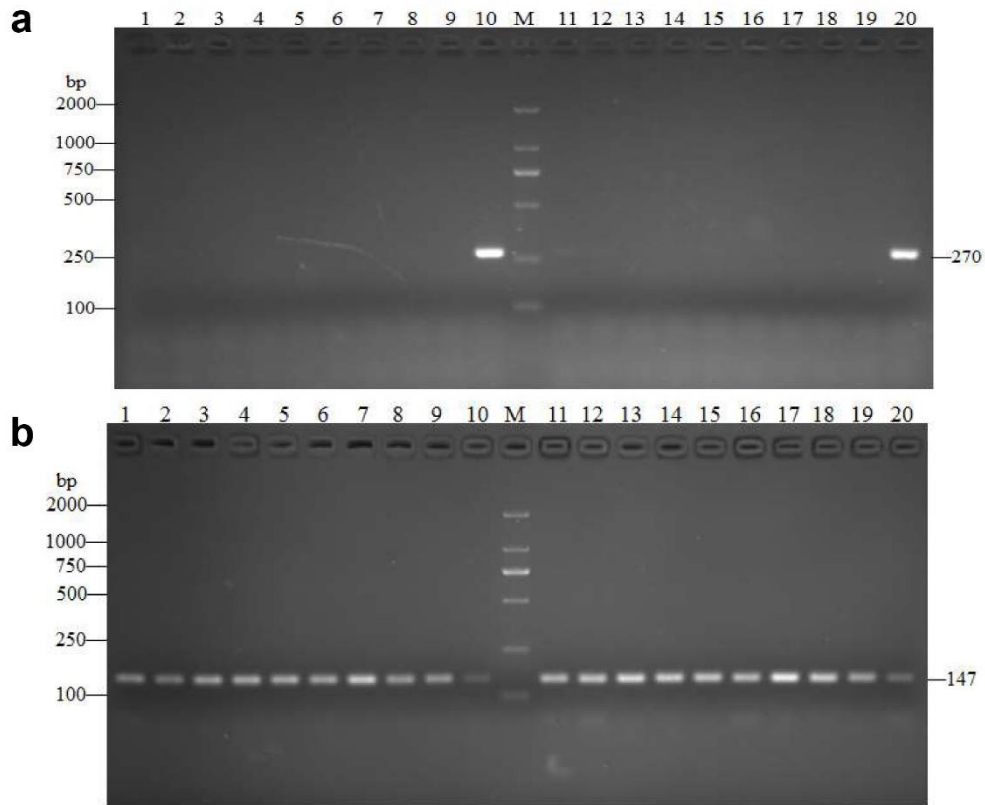


Figure S5. Electrophoresis pattern. (a) Electrophoresis pattern of chicken *CEL* gene expression in 10 tissues of different individuals. (b) Electrophoresis pattern of chicken *ACTB* gene expression in 10 tissues of different individuals.

Note: Figures 1 to 10 of the **a** and **b** diagrams represent the heart, liver, spleen, lung, kidney, duodenum, ovary, abdominal fat, breast muscle and pancreas of the same individual. Figures 11 to 20 represent ten samples from another individual. Figure. 4a is obtained from the late 11-20 lanes of the **a** and **b** diagrams, respectively.

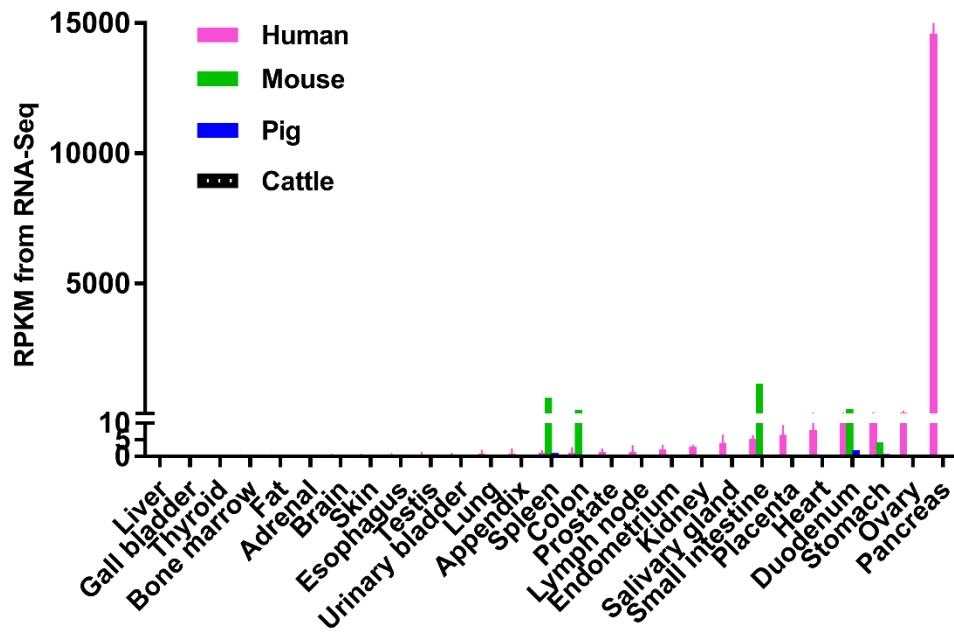


Figure S6. Tissue expression profiles of the *CEL* gene from different species.

Supplementary Table S1. The differences in genotypic frequencies of χ^2 and P -values among seven chicken breeds with *CEL* gene mutations

Breeds	XC	LS	CS	DX	Commercial laying hens	Commercial broiler
XC		0.450	4.359	4.031	100.410	172.757
LS	0.798		4.744	2.200	93.660	163.585
CS	0.113	0.093		11.223	118.467	195.664
DX	0.133	0.333	0.004		67.701	123.960
Commercial laying hens	0.000	0.000	0.000	0.000		13.869
Commercial broiler	0.000	0.000	0.000	0.000	0.000	

Note: χ^2 and P -value differences for genotypic frequencies between breeds are shown in the uptriangle and the downtriangle of this table, respectively. XC = Xichuan black-bone chicken, LS = Lushi blue-eggshell chicken, CS = Changshun blue-eggshell chicken, DX = Dongxiang blue-eggshell chicken. Commercial laying hens (Hy-Line Brown), commercial broiler (Ross 308 and Arbor Acres broiler).

Supplementary Table S2 Associations of the *CEL* indel with body weight at different stages in chicken

Body weight	(Mean \pm SE)			<i>P</i>
	II (n=243)	ID (n=415)	DD (n=136)	
BW0(g)	30.46 \pm 0.44	30.82 \pm 0.42	30.56 \pm 0.46	0.210
BW2(g)	122.44 \pm 2.57	122.43 \pm 2.44	121.59 \pm 2.76	0.890
BW4(g)	318.90 \pm 6.94	321.85 \pm 6.68	327.50 \pm 7.42	0.216
BW6(g)	553.95 \pm 13.12	562.99 \pm 12.60	573.91 \pm 13.91	0.092
BW8(g)	806.68 \pm 18.79	816.31 \pm 18.01	834.50 \pm 20.13	0.135
BW10(g)	1099.29 \pm 23.63	1110.73 \pm 22.65	1139.32 \pm 25.16	0.059
BW12(g)	1336.89 \pm 27.16	1354.02 \pm 25.87	1379.31 \pm 29.11	0.121

Note: BW0, 2, 4, 6, 8, 10, and 12 = body weight at the age of 0 days, 2, 4, 6, 8, 10 and 12 weeks, respectively.

Supplementary Table S3. Primers used for amplifying, cloning and expression analysis of chicken the *CEL* gene

Primers	Primer sequences (5'→3')	Tm	Product size (bp)
P1	F: GTCACTCAGAGGCTCCCTTT R: TGGCTAGGAACCCACGATAC	60°C	246/345
P2	F: CTGGAGACCCAAACATGCCA R: CTCTCTTGAACAGACCTGCGT	60°C	270
P3	F: TGCACACCATGGCTCACTG R: TTTTCAGCGGCTCCACAGTC	55°C	1698
<i>ACTB</i>	F: CAGCCAGCCATGGATGATGA R: ACCAACCATCACACCCTGAT	60°C	147
<i>APOB</i>	F: ATGTTCAAAAGATGCGGCC R: GCATGGCTCTTCTCTCACTG	60°C	224
<i>MTTP</i>	F: CAGGAGGGATGGAGTTCAGC R: TGGTCACGGAATGCCTGAAA	60°C	243
<i>APOV1</i>	F: CAATGAAACGGCTAGACTCA R: AACACCGACTTTTCTTCCAA	60°C	108
<i>SREBF1</i>	F: GAGACCATCTACAGCTCCGC R: CATCCGAAAAGCACCCCTCT	60°C	154