Bioinformatics analysis of KLK gene family in yak

Qingbo Zheng (zhengpingbo19@163.com)
Chinese Academy of Agricultural Sciences
Xiaolan Zhang
Chinese Academy of Agricultural Sciences
Pengjia Bao
Chinese Academy of Agricultural Sciences
Na Ye
Chinese Academy of Agricultural Sciences
Fubin Wang
Chinese Academy of Agricultural Sciences
Yi Wu
Chinese Academy of Agricultural Sciences
Min Chu
Chinese Academy of Agricultural Sciences
Xian Guo
Chinese Academy of Agricultural Sciences
Xiaoyun Wu
Chinese Academy of Agricultural Sciences
Heping Pan
Northwest Minzu University
Ping Yan
Chinese Academy of Agricultural Sciences

Research Article

Keywords: Yak, KLK, Bioinformatics analysis, Gene family

Posted Date: May 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1609143/v1

Abstract

Kallikrein-related peptidases (KLK) have 15 members, which are the largest family of human serine proteases known so far, but there is no report about KLK in yak. In this study, Six members of yak KLK gene family were selected as the research objects, including KLK1, KLK4, KLK5, KLK6, KLK7, and KLK15. The functions of yak KLK family members were predicted and analyzed by Expasy, STRING, R language and other software. The molecular weight of KLK proteins in yak ranged from 17276. 46-31884. 23 Da, and grand average of hydropathicity of the KLK proteins is less than 0, which indicated that they all belong to hydrophilic proteins; The phylogenetic tree showed that KLK5 and KLK1 are highly similar to the same family members, while different species of KLK7 differ in motif composition and gene structure. Protein–protein interaction analysis showed that KLK5, KLK6, and KLK7 interacted with serine protease inhibitor Kazal type 5(SPINK5), a serine protease inhibitor, while KLK7 and KLK6 interacted with corneodesmosin (CDSN); Go enrichment analysis showed that the KLK gene family was mainly enriched in the regulation of platelet alpha granules, endopeptidase activity, epidermal development and serine hydrolase activity. And the expression of KLK7 in yak skin was higher than that of KLK5, KLK6 and KLK10. Our results not only enriched the KLK gene family information and obtained the interaction gene enrichment pathway, but also provided a theoretical basis for further transition to in vitro experiments and formulation of experimental schemes.

Introduction

Kallikrein-related peptidases (KLK) are a family of extracellular serine proteases. Its gene is located on the long arm of chromosome 19 from 19q13.3□13.4, which is the largest known serine protease gene cluster, with a high degree of homology at the nucleotide and protein levels and a similar genome structure (Yousef et al, 2001). The KLK-encoded protein exists in the cell in the form of tissue prokallikrein, and is then transferred to the outside of the cell. It is hydrolyzed by other members of KLK or by matrix metalloproteinases to generate KLK with a mature catalytic conformation. As extracellular enzymes, KLK play roles in growth factors, proteases, membrane-bound receptors, initiation of intracellular signaling pathways, and the hydrolysis of cytokines (Dong et al, 2014). KLK family members are mainly secreted by epithelial cells and distributed in tissues such as pancreas, brain, skin, adrenal gland, colorectum, kidney, and prostate. However, how the different KLK proteases function in biological tissues, and how the associated factors and contributing factors occur, remains largely unknown. Therefore, we performed bioinformatics analysis of yak-associated KLK protein to explore the underlying molecular mechanism.

The expression levels of KLK1, KLK5-KLK8, and KLK10-KLK13 were slightly higher in the stratum corneum, upper granule, sebaceous glands, sweat glands and hair follicles, KLK2 and KLK3 are mainly highly expressed in the prostate and can be used as biological markers of prostate diseases (Prassas et al, 2015), so they are widely used in the screening, diagnosis and prognosis monitoring of prostate cancer. KLK4 is highly expressed in tooth development, when KLK4 gene mutation will lead to increased protein quality in the enamel space, but the enamel thickness does not change (Bartlett et al, 2014; Wrightet al, 2010). KLK6 is highly expressed in the brain and central nervous system, and is associated with some neurodegenerative diseases (Lawrence et al, 2010). Some studies have found that KLK7 can activate the skin defense mechanism by hydrolyzing human antimicrobial peptide hCAP18 and prevent pathogenic microorganism infection (Yamasaki et al, 2006). In addition, when the Th2 cells secrete interleukins IL-3 and IL-4 were highly expressed in the skin of patients with atopic dermatitis, the expression level of KLK7 in human keratinocytes was significantly increased (Morizane et al, 2012).

Yak is a unique genetic resource in the Qinghai-Tibet Plateau and adjacent areas, and has long lived in the alpine grassland with extremely harsh ecological environment. Yak is the most convenient and cheapest means of transport for herdsmen, so it has the reputation of "the boat on the plateau". At present, there are more than 14 million yaks in

China, accounting for more than 92% of the total number of yaks in the world (Guo et al, 2009). Under the special ecological environment and long-term natural selection, yak has gradually formed unique physiological characteristics and living habits in terms of morphology and physiology, such as tolerance to cold and heat, late maturity and low fecundity (Guo et al, 2006). The yak has strong cold resistance, which is benefited from its unique hair coat structure (Bao et al, 2020). Yak coat is a mixed coat composed of different lengths, fineness and fiber types. Although it is an open coat, it has a thick air layer and poor moisture absorption of the coat fiber. This feature is one of the protective mechanisms for yaks to adapt to alpine environmental conditions. Yak is one of the few cattle species that can adapt to harsh environments such as plateau cold and strong ultraviolet radiation. The skin covers the entire body surface and can protect deep tissues from external mechanical stimulation, photochemical radiation and resist cold. At present, KLK family has been extensively studied in humans and mice, and it is found that KLK gene family members are widely expressed in skin tissues, but there is no report on yak. Bioinformatics analysis of yak KLK family in this study can lay the foundation for further exploring the role of yak KLK gene family.

Materials And Methods

Origin of yak KLK protein sequence

Download known yak KLK protein sequences from the UniProt database (Table 1).

Table 1
Yak KLK protein information

Gene names	Protein names	UniProt accession number
KLK1	Kallikrein 1	A0A8B9XWN3
KLK4	Kallikrein related peptidase 4	A0A8B9XV38
KLK5	Kallikrein related peptidase 5	A0A8B9XSF1
KLK6	Kallikrein related peptidase 6	A0A8B9XXU1
KLK7	Kallikrein related peptidase 7	A0A8B9XXG3
KLK15	Kallikrein related peptidase 15	A0A8B9XXM5

Analysis Of Physicochemical Properties Of Yak Klk Protein

The online bioinformatics software Expasy Protoparam (Elisabeth et al, 2003) was used to analyze the physicochemical properties of the encoded protein, NetPhos3.1 Server (Gao et al, 2021) was used to predict the phosphorylation site, and SignalP-5.0 (Almagro et al, 2019) was used to analyze the signal peptide of the encoded protein. Prediction of protein secondary structure by SOPMA (Geourjon et al, 1995), using STRING (Damian et al, 2017) for protein interaction analysis; using Clustalw (Larkin et al, 2007) to align amino acid sequences; using ESPript 3 (Chakravarty et al, 2015) to check the alignment results; using MEME (Bailey et al, 2009) to predict yak klk protein conserved motif. Analysis of KLK gene expression in mouse dorsal skin tissue (GEO: GSE185268), Liaoning cashmere goat skin tissue (GEO: GSE182474) and yak skin tissue (preparing to upload scRNA-seq data of yak skin tissue to NCBI) using R language.

Results

Analysis of physicochemical properties of yak KLK protein

The analysis results of basic physical and chemical properties of yak KLK protein showed that the molecular weight of KLK protein was between 17276.46-31884.23 Da. The number of amino acids was between 160-293, the instability index was between 33.64-56.01. The instability indexes of KLK1, KLK4, KLK5, and KLK6 were less than 40, and the stability was high, while the instability indexes of KLK7 and KLK15 were greater than 40, and the stability was poor. the aliphatic index was between 68.94-86.00, the average hydrophilicity of six KLK proteins was less than 0, so they were hydrophilic proteins. KLK1, KLK4, KLK5, KLK6, and KLK15 have signal peptides, while KLK7 has no signal peptide (Table 2).

Table 2 Analysis of physicochemical properties of encoded proteins

Name	Formula	Isoelectric point	Molecular weight	Number of amino acids	Instability index	Aliphatic index	Grand average of hydropathicity
KLK1	$C_{1274}H_{1928}N_{332}O_{381}S_{16}$	4.81	28504.28	259	36.55	79.77	-0.175
KLK4	$C_{770}H_{1167}N_{207}O_{231}S_{8}$	4.99	17276.46	160	36.26	86.00	-0.03
KLK5	$C_{1392}H_{2180}N_{406}O_{417}S_{19}$	8.64	31884.23	293	38.42	73.62	-0.29
KLK6	$C_{1182}H_{1857}N_{347}O_{346}S_{17}$	6.93	27009.88	246	33.64	84.84	-0.244
KLK7	$C_{1325}H_{2089}N_{369}O_{405}S_{20}$	9.02	30309.59	284	56.01	68.94	-0.259
KLK15	$C_{1222}H_{1954}N_{362}O_{362}S_{18}$	8.23	28086.25	257	44.84	83.42	-0.187

Prediction of phosphorylation sites of yak KLK protein

The yak KLK protein has serine, threonine, and tyrosine phosphorylation sites (Table 3). The number of phosphorylation sites is serine > threonine > tyrosine. The phosphorylation sites of each family member are the number is between 15-53. Among them, KLK7 contains up to 53 phosphorylation sites, while KLK6 has only 15.

Table 3 Predicted phosphorylation sites of yak KLK protein

Name	Serine	Threonine	Tyrosine
KLK1	8	5	3
KLK4	13	2	2
KLK5	20	10	5
KLK6	8	4	3
KLK7	39	11	3
KLK15	15	8	2

Yak KLK protein sequence alignment and phylogenetic tree construction

The amino acid sequences of KLK family in yak were compared by Clustalw online software as shown in Fig. 1a. The amino acid sequence similarity of the KLK family is poor and the alignment score is low, among which the alignment score between KLK1, KLK5, and KLK15 is significantly higher than the overall comparison level. MEGA11 software was used to construct the phylogenetic tree of KLK protein sequences of yak, human, rattus norvegicus, sheep, pig and bos indicus ((Fig. 1b). It can be seen from the figure that KLK5 and KLK1 have high similarity in the same family members, and the relationship between KLK5 and KLK4 between family members is most recent, and they may originate from the same ancestor. However, the segregation of KLK7 among different species is remarkable. Interestingly, yak KLK7 and rattus norvegicus KLK7 are on the same branch, with high similarity.

Prediction of conserved regions of yak KLK protein

Six different motifs of yak KLK protein were predicted by MEME. The specific structures of the motifs are shown in Fig. 2a. The sequence lengths of the motifs are 50, 41, 21, 44, 15, and 16BP, respectively. Among them, KLK4 and KLK7 have less motifs than KLK1, KLK5, KLK6, and KLK15 (Fig. 2b). The frequencies of leucine, serine and alanine are the highest at 9.16%, 7.37%, and 7.31%, respectively. The lowest frequencies of cystine and tryptophan were 2.24%, 1.81%, and 1.33%, respectively.

Yak KLK protein interaction network analysis and enrichment analysis

The interaction network of six KLK proteins was obtained using the STRIN online tool (Fig. 3). KLK5, KLK6, and KLK7 all interact with serine peptidase inhibitor Kazal type 5 (SPINK5). At the same time, the regulatory networks of KLK5 and KLK7 are similar, and both interact with KLK8 and SPINK5. KLK7 and KLK6 interact with corneodesmosin (CDSN). While KLK15 has an independent regulatory network.

After the interaction analysis of KLK proteins, the interacting genes were used for GO enrichment analysis (Fig. 4). The results showed that the KLK1-interacting protein BP was mainly enriched in platelet degranulation, negative regulation of endopeptidase activity, negative regulation of peptidase activity, and negative regulation of proteolysis; CC was mainly enriched in platelet alpha granules, secretory granule lumen, cytoplasmic vesicle lumen, and vesicle lumen; MF is mainly enriched in molecular functions such as endopeptidase inhibitory activity, endopeptidase regulation activity, and peptidase inhibitory activity. GO enrichment analysis of KLK4 interacting proteins showed that its BP was mainly enriched in odontogenesis of dentin-containing tooth, biomineralized tissue development, and regulation of tooth mineralization and other pathways related to tooth development; CC was mainly enriched in the endoplasmic reticulum lumen, collagen-containing extracellular matrix, and platelet alpha granule lumen; MF is mainly enriched in molecular functions such as extracellular matrix structural constituent, endopeptidase inhibitor activity, peptidase inhibitor activity, and endopeptidase regulatory activity. GO enrichment analysis of KLK5 interacting proteins showed that its BP was mainly enriched in skin development-related pathways such as epidermal development, keratinization, keratinocyte differentiation, and epidermal cell differentiation; while CC was only enriched in one entry of lamellar bodies; MF is mainly enriched in molecular functions such as endopeptidase activity, serine-type endopeptidase inhibitor activity, and serine-type endopeptidase activity. GO enrichment analysis of KLK6 interacting proteins showed that its BP was mainly enriched in the negative regulation of endopeptidase activity and the negative regulation of endopeptidase activity; however, there were no enrichment items in cell composition, and MF was only enriched in six pathways related to enzyme function, including serine-type endopeptidase inhibitor activity, enzyme inhibitor activity, peptidase regulator activity, endopeptidase regulator activity, peptidase inhibitor activity, and endopeptidase inhibitor activity. GO enrichment analysis of KLK7 interacting proteins showed that its BP was mainly enriched in epidermal development, extracellular matrix organization, regulation of defense response to bacterium, and keratinocyte

differentiation; CC was mainly enriched in cell-cell junction, lamellar body, desmosomes, and cornified envelopes; MF is mainly enriched in molecular functions such as endopeptidase activity, serine hydrolase activity, serine-type peptidase activity, and serine-type endopeptidase activity. Finally, the GO enrichment analysis of KLK15 interacting proteins found that BP was mainly enriched in metabolic pathways such as organic acid catabolic process, carboxylic acid catabolic process, and small molecule catabolic process; MF was mainly enriched in G protein β-subunit binding, vitamin binding, and G-protein α-subunit binding and other signaling pathways. Among them, KLK5 and KLK7 are genes closely related to the development of yak epidermis.

KEGG enrichment analysis of KLK interaction genes using R language and visualization using Graphpad Prism software (Fig. 5), the results show that KLK4, KLK5, and KLK15 are only enriched in one pathway, complement and coagulation cascades, staphylococcus aureus infection and carbon metabolism pathway, KLK1, KLK6, and KLK7 is enriched in multiple pathways, and KEGG enrichment analysis of KLK1 interacting proteins found that they were mainly enriched in the complement and coagulation cascades, inflammatory mediator regulation of TRP channels, and the renin-angiotensin system; the first three pathways of KLK6 enrichment were histidine metabolic pathway, alzheimer ' s disease and β-alanine metabolism. The three pathways with the highest enrichment specificity of KLK7 were thyroid cancer, bacterial invasion of epithelial cells, and apelin signaling pathway.

Analysis of the expression of some KLK genes in yak, mouse and sheep skin tissue

After tSNE cluster analysis of scRNA-seq data downloaded from NCBI, the expression of KLK7 and other genes in skin tissues of yak, mice and sheep was observed (Fig. 6). The results showed that the expression of KLK gene in mice was significantly lower than that in yak and sheep, while the expression of KLK7 in yak was mainly high in cluster1 and cluster5 and the expression of KLK7 was significantly higher than that of KLK5, KLK6, and KLK10. KLK gene family members in mouse back skin tissue were mainly concentrated in cluster7, and the expression level of KLK7 was significantly higher than that of KLK11. KLK7 was mainly highly expressed in cluster1 and cluster3 in Liaoning cashmere goat skin tissue, while KLK11 was more widely distributed than KLK7, and it was distributed in cluster3, cluster4, cluster8, and cluster12.

Discussion

With the development of genomics, predicting the potential biological functions of target genes from the genomic level based on the existing genomic data has become a hot topic in the field of biology research. This study identified six members of the yak KLK gene family and found that KLK4 was mainly associated with tooth development, and KLK5 and KLK7 were mainly associated with skin function. Proteins interact in life processes such as gene expression regulation, growth and development, signal factor transmission, and cell cycle regulation, and together they form a protein interaction network diagram. From the results of protein interaction network, KLK5, KLK6, and KLK7 all interact with SPINK5. Loss of SPINK5 results in a loss of inhibition of KLK5 and KLK7, resulting in premature shedding of the stratum corneum by proteolysis of stratum corneum cadherin, while KLK5 activates PAR2, causing an inflammatory and allergic cascade (Hovnanian et al, 2014). Both KLK7 and KLK6 interact with CDSN protein. KLK7 in the skin can degrade CDSN in vitro, and KLK5 can activate KLK7, which can cause the degradation of desmosomes and then cause cell shedding (Caubet et al, 2004). Studies have found that KLK5 and KLK7 are highly expressed in acinar cells of pancreatic tissue, which may be involved in the digestive function of the pancreas or the activation of other secretases in the digestive pathway (Dong et al, 2008; Borgoño et al, 2004; Clements et al, 2004).

The skin is the largest tissue organ in mammals, which can account for more than 15% of the net weight of the human body. The skin is mainly composed of three parts, namely epidermis, dermis, and subcutaneous tissue. The epidermis is composed of keratinocytes with different differentiation stages, and its main function is to provide an effective

permeability barrier (Rawlings et al, 2004), abnormal skin barrier function may lead to skin barrier function-related diseases. The expression of KLK5, KLK6, KLK7, KLK8, KLK10, KLK13, and KLK14 in serum of patients with atopic dermatitis (AD) was significantly higher than that of normal skin, and the expression of KLK7 was higher than that of other KLK members (Komatsu et al, 2007). Overexpression of KLK7 in the skin can trigger abnormally elevated levels of proteolysis, leading to pathological desquamation and skin diseases such as Netherton syndrome (Descargues et al, 2006; Song, 2019). An important part of skin desquamation is the degradation of desmosomes, which causes keratinocytes to fall off and produce desquamation (Borgoño et al, 2007). So KLK7 plays a key role in skin desquamation (Egelrud, 1993). when KLK14 was overexpressed in the granular layer, it increased the proteolytic activity in the granular layer and hair follicles, resulting in no hair growth and showed hyperplastic hair follicle growth defects (Gouin et al, 2020). On the other hand, overexpression of KLK5 and KLK7 induces Netherton syndrome and hair shaft defects (Furio et al, 2014). Studies on mice with genetic background of Netherton syndrome showed that mice knocked out KLK5 and KLK7 genes had fully functional skin and normal hair growth in adulthood (Kasparek et al, 2017).

GO analysis of KLK interaction genes showed that yak KLK family members were enriched in serine protease inhibitor-related pathways. The serine protease inhibitor family is widely present in a variety of tissues and organs in organisms and is closely related to the occurrence of some diseases (Lute et al, 2009). For example, serine protease inhibitor Kazal-type 6 (Spink6) has inhibitory effect on KLK activity (Zheng et al, 2017). At the same time, Spink6, as a selective inhibitor of KLK, is differentially expressed in the skin of different parts of the human body, which can inhibit KLK5 and KLK7, but the expression level of Spink6 will decrease in the lesions of atopic dermatitis, in addition, the LEKTI-2 inhibitor encoded by the Spink9 gene can regulate the peeling process in the skin by specifically inhibiting KLK5, but has no inhibitory effect on other KLK proteases such as KLK7 (Meyer-Hoffert et al, 2010).

In summary, the amino acid sequence of yak KLK family was analyzed in terms of sequence characteristics, evolutionary relationship, physicochemical properties and GO enrichment. The results showed that yak KLK proteins were all hydrophilic proteins; except for KLK7, which had no signal peptide, the other five KLK members had signal peptides; the classification and evolutionary relationship of yak KLK gene family were further understood through phylogenetic tree analysis. At the same time, the GO and KEGG enrichment analysis of yak KLK interaction genes were carried out to predict the functional characteristics of the encoded protein, and provide reference for determining gene function. To lay the foundation for further study of biological function of KLK family in yak development.

Declarations

Acknowledgments

The authors thank Qi Bao for assistance with the data analyzing.

Funding

This research was funded by the Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (25-LZIHPS-01).

Conflict of interests

The authors declare no conflict of interest.

References

1. Almagro AJJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, Von HG, Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. Nature Biotechnology 37:420–423.
2. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research 37(Web Server issue):W202-208.
3. Bao P, Luo J, Liu Y, Chu M, Ren Q, Guo X, Tang B, Ding X, Qiu Q, Pan H, Wang K, Yan P (2020) The seasonal development dynamics of the yak hair cycle transcriptome. BMC genomics 21:355–367.
4. Bartlett JD, Simmer JP (2014) Kallikrein-related peptidase-4 (KLK4): role in enamel formation and revelations from ablated mice. Frontiers in physiology 5:1–7.
5. Borgoño CA, Michael IP, Diamandis EP (2004) Human tissue kallikreins: physiologic roles and applications in cancer. Molecular Cancer Research Mcr *2*:257–280.
6. Borgoño CA, Michael IP, Komatsu N, Jayakumar A, Kapadia R, Clayman GL, Sotiropoulou G, Diamandis EP (2007) A potential role for multiple tissue kallikrein serine proteases in epidermal desquamation. The Journal of biological chemistry 282:3640–3652.
7. Caubet C, Jonca N, Brattsand M, Guerrin M, Bernard D, Schmidt R, Egelrud T, Simon M, Serre G (2004) Degradation of corneodesmosome proteins by two serine proteases of the kallikrein family, SCTE/KLK5/hK5 and SCCE/KLK7/hK7. The Journal of investigative dermatology 122:1235–1244.
8. Chakravarty D, Chakraborti S, Chakrabarti P (2015) Flexibility in the N-terminal actin-binding domain: clues from in silico mutations and molecular dynamics. Proteins Structure Function and Bioinformatics 83:696–710.
9. Clements JA, Willemsen NM, Myers SA, Dong Y (2004) The tissue kallikrein family of serine proteases: functional roles in human disease and potential as clinical biomarkers. Critical Reviews in Clinical Laboratory Sciences 41:265–312.
10. Descargues P, Deraison C, Prost C, Fraitag S, Mazereeuw-Hautier J, D'Alessio M, Ishida-Yamamoto A, Bodemer C, Zambruno G, Hovnanian A (2006) Corneodesmosomal cadherins are preferential targets of stratum corneum trypsin- and chymotrypsin-like hyperactivity in netherton syndrome. The Journal of investigative dermatology 126:1622–1632.
11. Dong Y, Loessner D, Irving-Rodgers H, Obermair A, Nicklin JL, Clements JA (2014) Metastasis of ovarian cancer is mediated by kallikrein related peptidases. Clinical & experimental metastasis 31:135–147.
12. Dong Y, Matigian N, Harvey TJ, Samaratunga H, Hooper JD, Clements JA (2008) Tissue-specific promoter utilisation of the kallikrein-related peptidase genes, KLK5 and KLK7, and cellular localisation of the encoded proteins suggest roles in exocrine pancreatic function. Biological chemistry 389:99–109.
13. Egelrud T (1993) Purification and preliminary characterization of stratum corneum chymotryptic enzyme: a proteinase that may be involved in desquamation. The Journal of investigative dermatology 101:200–204.
14. Elisabeth G, Alexandre G, Christine H, Ivan I, Appel RD, Amos B (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research 31:3784–3788.
15. Furio L, Veer DS, Jaillet M, Briot A, Robin A, Deraison C, Hovnanian A (2014) Transgenic kallikrein 5 mice reproduce major cutaneous and systemic hallmarks of netherton syndrome. The Journal of experimental medicine 211:499–513.
16. Gao T, Gao Y, Liu X, Nie Z, Sun H, Lin K, Peng H, Wang S (2021) Identification and functional analysis of the SARS-COV-2 nucleocapsid protein. BMC Microbiology 21:58–67.
17. Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11:681–684.

18. Gouin O, Barbieux C, Leturcq F, Bonnet DCM, Petrova E, Hovnanian A (2020) Transgenic kallikrein 14 mice display major hair shaft defects associated with desmoglein 3 and 4 degradation, abnormal epidermal differentiation, and IL-36 signature. The Journal of investigative dermatology 140:1184–1194.
19. Guo X, Yan P, Liang CN, Pei J, Zeng YF, Bao PJ (2009) Developmental situations and countermeasures of yak industry in china. China Cattle Science 35:55–57.
20. Guo X, Yang BH, Li YS, Yan P. (2006). Ecological and physiological characteristics of yak. Chinese Journal of Animal Science 42:56–57.
21. Hovnanian Alain (2014) Netherton syndrome: new advances in the clinic, disease mechanism and treatment. Expert Review of Dermatology 7:81–92.
22. Kasparek P, Ileninova Z, Zbodakova O, Kanchev I, Benada O, Chalupsky K, Brattsand M, Beck IM, Sedlacek R (2017) KLK5 and KLK7 ablation fully rescues lethality of netherton syndrome-like phenotype. PLoS genetics 13:e1006566.
23. Komatsu N, Saijoh K, Kuk C, Liu AC, Khan S, Shirasaki F, Takehara K, Diamandis EP (2007) Human tissue kallikrein expression in the stratum corneum and serum of atopic dermatitis patients. Experimental dermatology 16:513–519.
24. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.
25. Lawrence MG, Lai J, Clements JA (2010) Kallikreins on steroids: structure, function, and hormonal regulation of prostate-specific antigen and the extended kallikrein locus. Endocrine reviews, 31:407–446.
26. Lute S, Wang H, Sanchez D, Barletta J, Chen Q, Brorson K (2009) Multiplex RT Q-PCR assay for simultaneous quantification of three viruses used for validation of virus clearance by biopharmaceutical production. Biologicals: journal of the International Association of Biological Standardization 37:331–337.
27. Meyer-Hoffert U, Wu Z, Kantyka T, Fischer J, Latendorf T, Hansmann B, Bartels J, He Y, Gläser R, Schröder JM (2010) Isolation of SPINK6 in human skin: selective inhibitor of kallikrein-related peptidases. The Journal of biological chemistry 285:32174–32181.
28. Morizane S, Yamasaki K, Kajita A, Ikeda K, Zhan M, Aoyama Y, Gallo RL, Iwatsuki K (2012) TH2 cytokines increase kallikrein 7 expression and function in patients with atopic dermatitis. The Journal of allergy and clinical immunology 130:259–261.
29. Prassas I, Eissa A, Poda G, Diamandis EP (2015) Unleashing the therapeutic potential of human kallikrein-related serine proteases. Nature reviews Drug discovery 14:183–202.
30. Rawlings AV, Harding CR (2004) Moisturization and skin barrier function. Dermatologic therapy 17, 43–48.
31. Song YQ (2019) Protease inhibitor A2ML1 regulates the differentiation and a dhesion of skin cells in healthy humans and patients with basal cell carcinoma via KLK7. Jinan University.
32. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C (2017) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research 45:362–368.
33. Wright JT, Daly B, Simmons D, Hong S, Hart SP, Hart TC, Atsawasuwan P, Yamauchi M (2010) Human enamel phenotype associated with amelogenesis imperfecta and a kallikrein-4 (g.2142G > A) proteinase mutation. European Journal of Oral ences 114:13–17.
34. Yamasaki K, Schauber J, Coda A, Lin H, Dorschner RA, Schechter NM, Bonnart C, Descargues P, Hovnanian A, Gallo RL (2006) Kallikrein-mediated proteolysis regulates the antimicrobial effects of cathelicidins in skin. FASEB

journal: official publication of the Federation of American Societies for Experimental Biology 20:2068–2080.

35. Yousef GM, Diamandis EP (2001) The new human tissue kallikrein gene family: structure, function, and association to disease. Endocrine reviews 22:184–204.

36. Zheng LS, Yang JP, Cao Y, Peng LX, Sun R, Xie P, Wang MY, Meng DF, Luo DH, Zou X, Chen MY, Mai HQ, Guo L, Guo X, Shao JY, Huang BJ, Zhang W, Qian CN (2017) SPINK6 promotes metastasis of nasopharyngeal carcinoma via binding and activation of epithelial growth factor receptor. Cancer research 77:579–589.

Figures

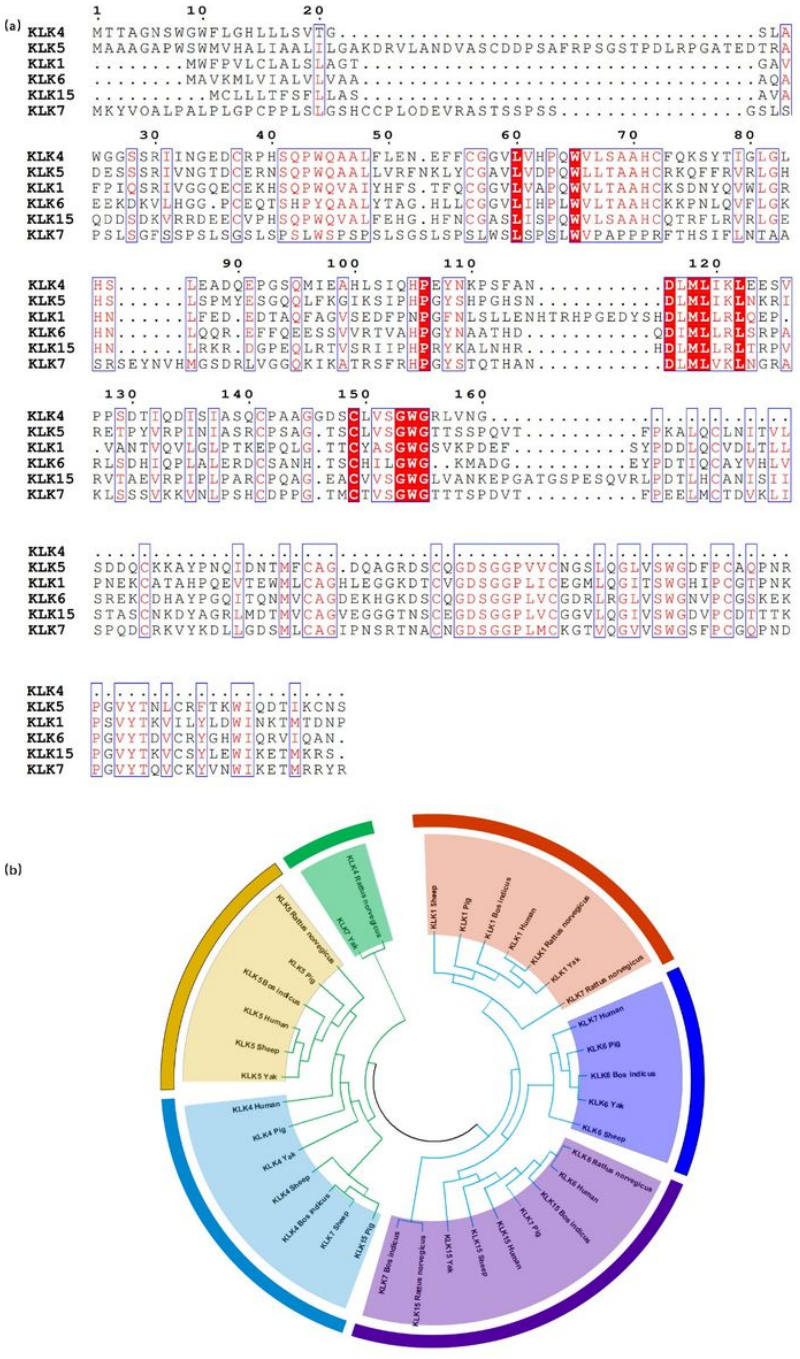

Figure 1

KLK sequence alignment and phylogenetic tree analyze. **a** Alignment of yak KLK protein sequences. **B** Phylogenetic analysis of yak KLK protein

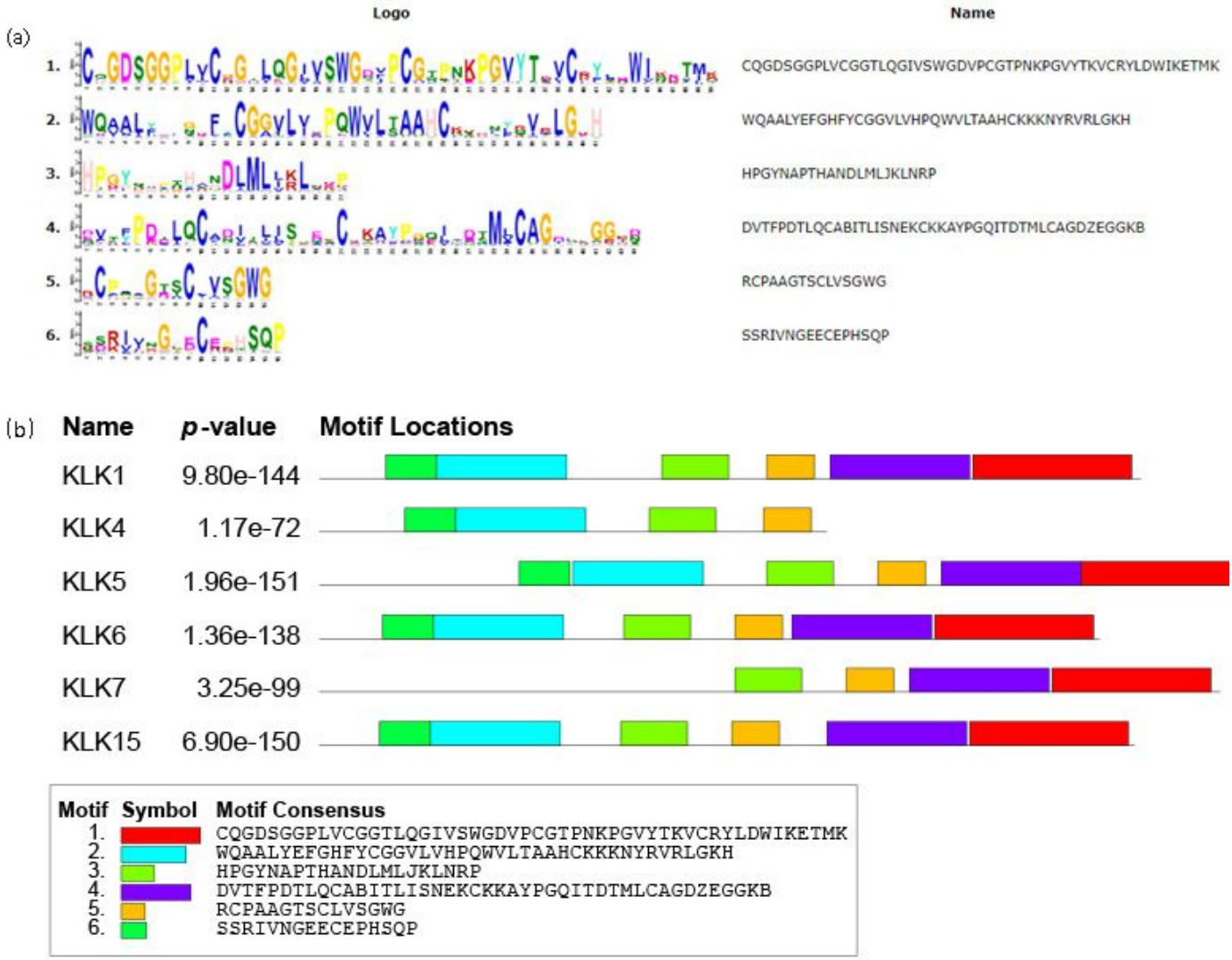

Figure 2

Sequence map and visualization of KLK motif in yak. **a** Sequence map of typical conserved motifs of yak KLK protein. **b** Visualization of yak KLK protein motif

Note: Each letter represents an amino acid, and the size of the letter represents the number of amino acids at the site, which is related to the consistency of amino acids. The larger the letter, the less the number, which means the better the consistency of the site and the more conservative it is. Good, otherwise it means that the consistency of the site is poor and the conservation is low

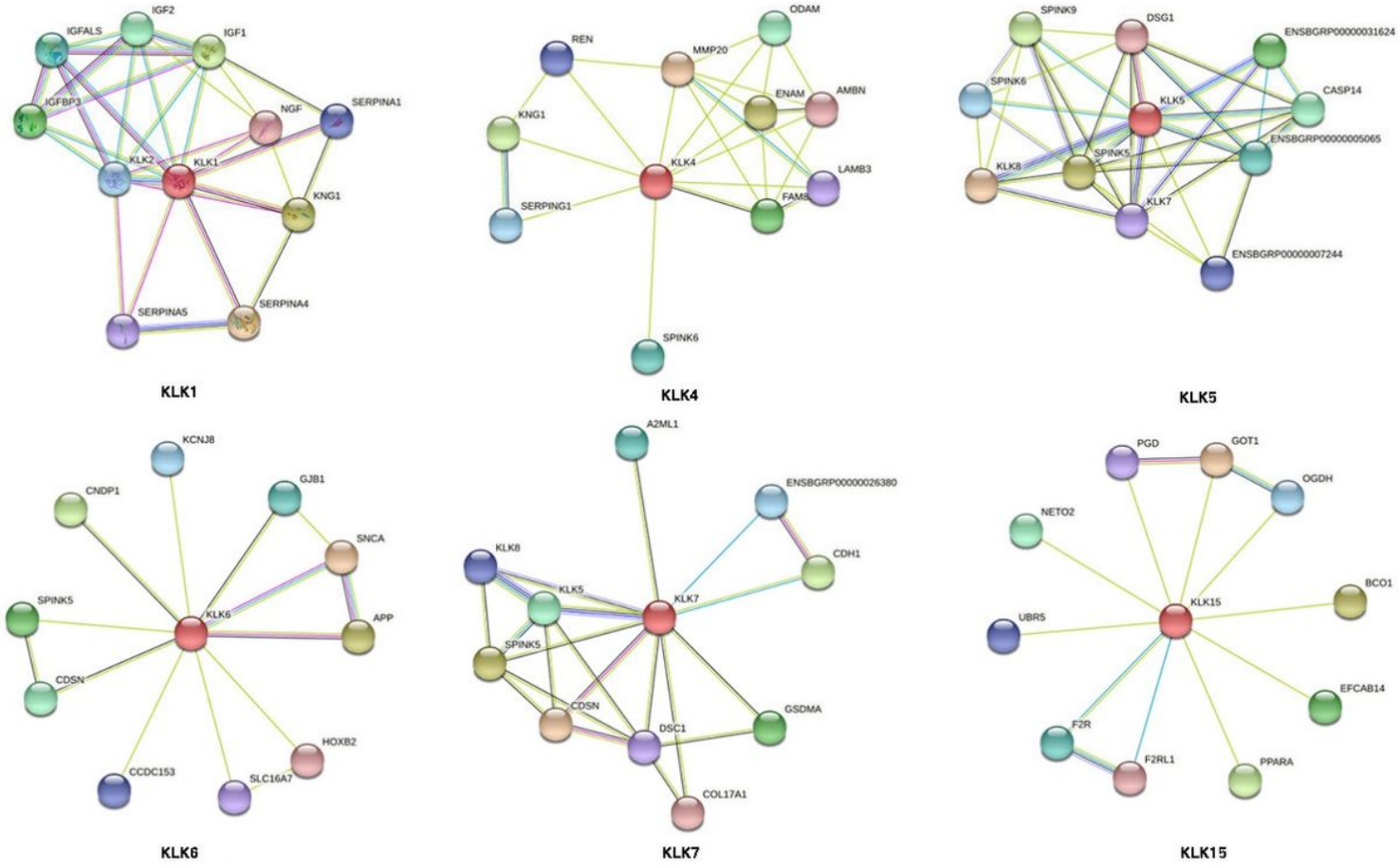

Figure 3

Analysis of yak KLK protein interaction network

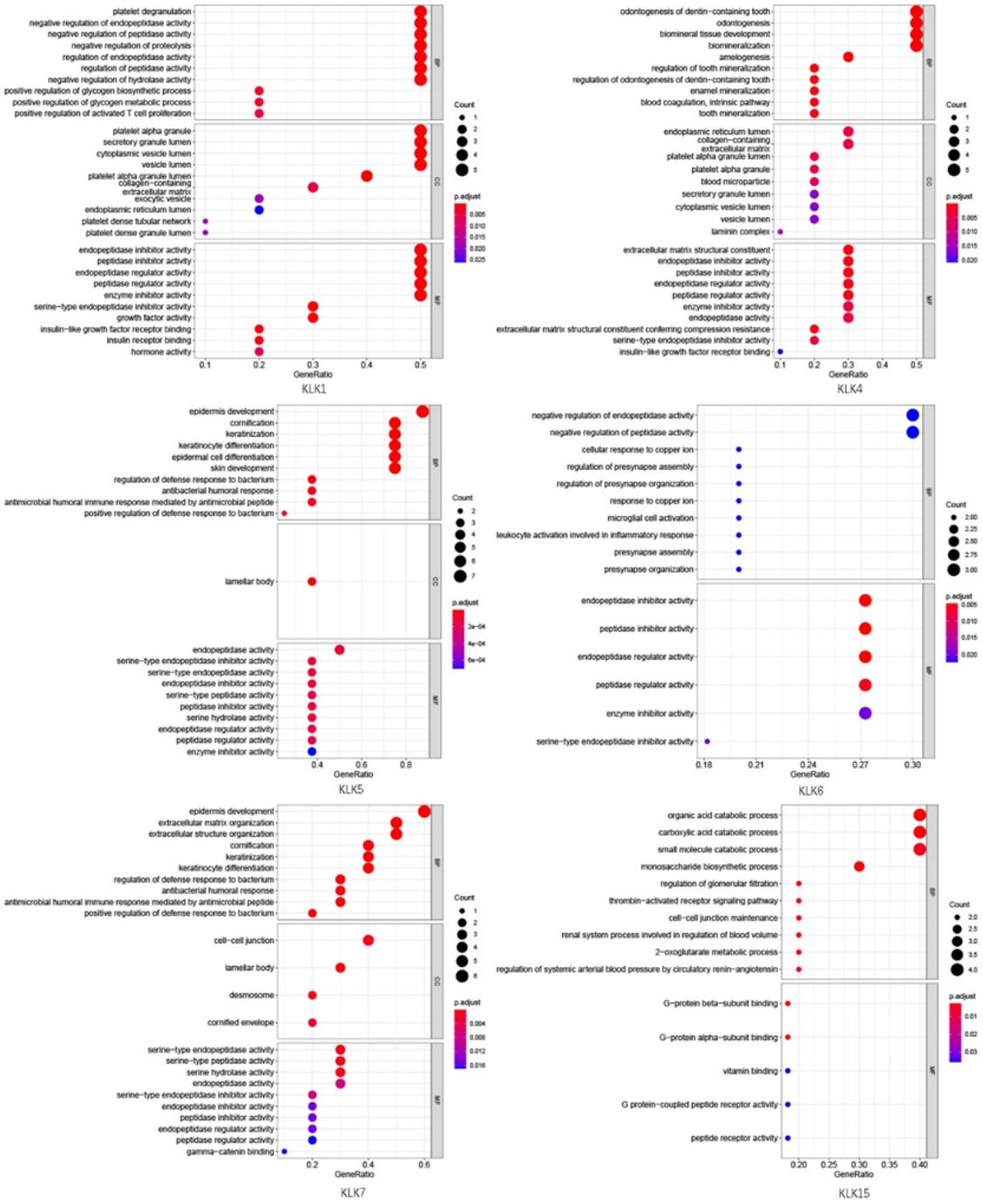

Figure 4

GO enrichment analysis of yak KLK genes

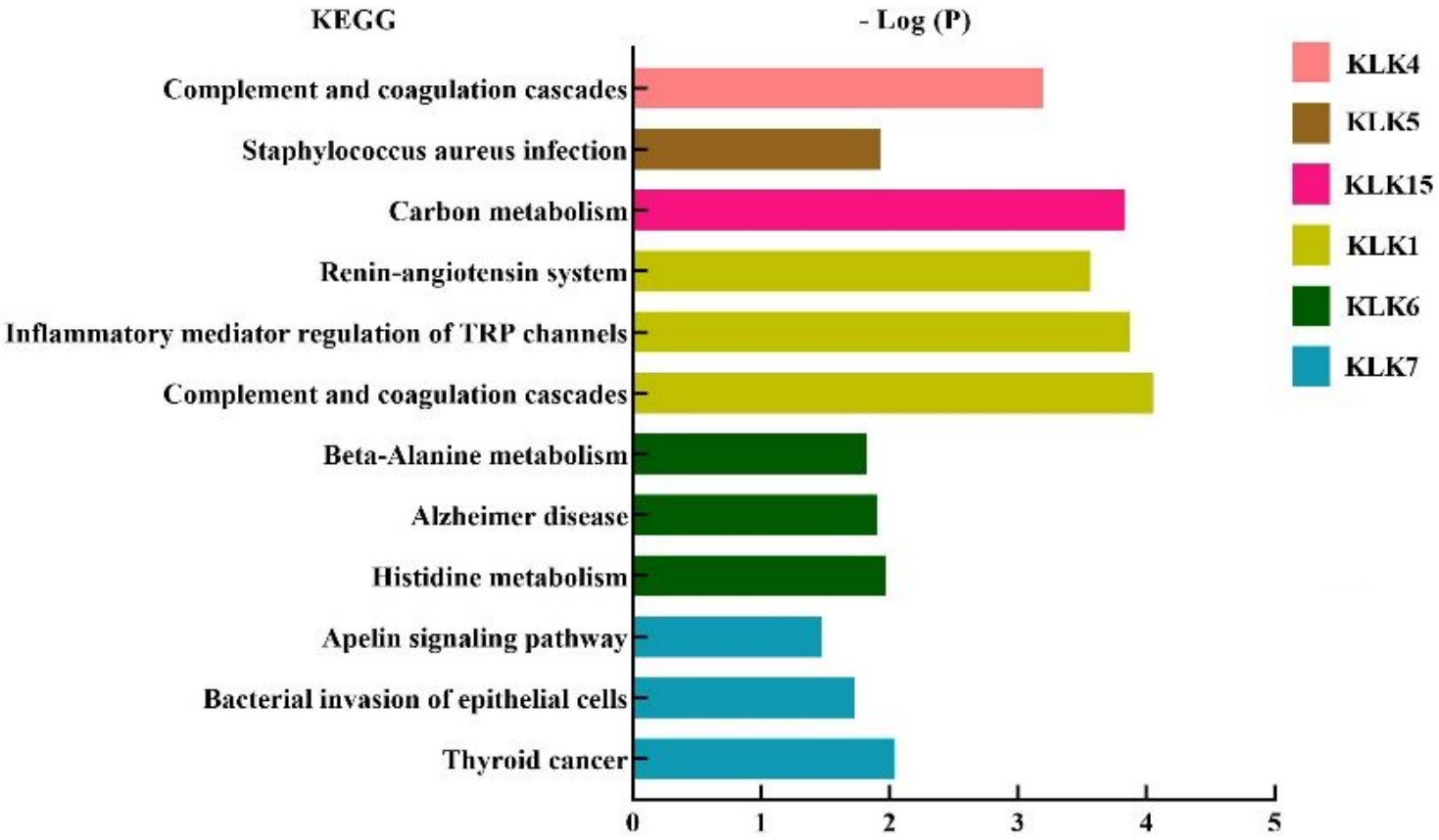

Figure 5

KEGG enrichment analysis of yak KLK genes

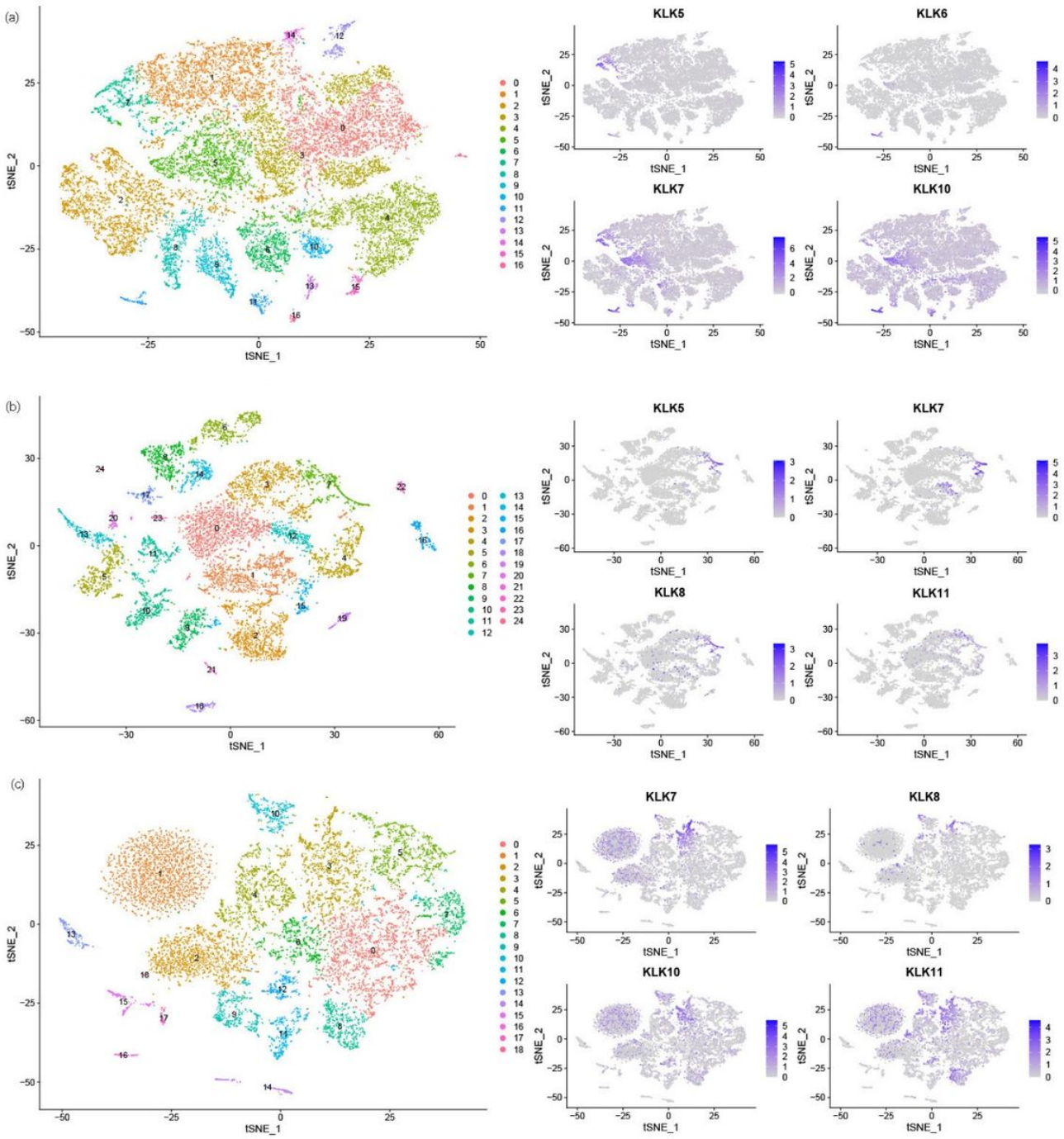

Figure 6

Expression of some KLK proteins in yak(a), mouse(b) and sheep(c) skin tissues