Proteomic profiling reveals KRT6C as the hereterodimer partner for KRT9 in epidermolytic palmoplantar keratoderma

Peiyao Li (lipeiyao11@hotmail.com)
Central South University https://orcid.org/0000-0002-1439-1036
Jialin Qi
Central South University
Yuhui Zhong
Hunan Normal University
Aoli Ding
Hunan Normal University
Ting Ding
Hunan Normal University
Pingping Xie
Hunan Normal University
Heng Xiao
Hunan Normal University https://orcid.org/0000-0002-4765-2597

Research Article

Keywords: epidermolytic palmoplantar keratoderma, hereterodimer, keratin 6C, keratin 9, mass spectroscopy, proteomics analysis

Posted Date: April 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1544457/v1

Abstract

Background: Epidermolytic palmoplantar keratoderma (EPPK), a highly penetrant autosomal dominant genodermatosis, is characterized by diffuse yellow keratoses on the palms and soles, with distinct erythematous border. The keratin 9 gene (*KRT9*) is restrictedly expressed in the suprabasal layer of palmoplantar epidermis, whose mutations are responsible for EPPK. To date, phenotypic therapy is the primary treatment for EPPK, so that the disease cannot be cured radically. Therefore, unveiling the pathogenic mechanism serves as an essential path to exploring the more fundamental treatment. Because KRT9 pairs with another type II keratin-binding partner to function as cytoskeleton protein in epidermis for EPPK, identifying the interaction partner is an essential first step in revealing their biological function.

Results: In this study, we obtained KRT9-interacting proteins in the human sole skin by using co-immunoprecipitation coupled with mass spectroscopy. Among the KRT9-interacting proteins, we proved that keratin 6C (KRT6C) is the hereterodimer partner for KRT9 by reverse co-immunoprecipitation, immunofluorescence assay and quantitative co-localization analysis. The *in silico* model for KRT6C/KRT9 built by SWISS-MODEL and PyMOL shows a typical coiled-coil structure in their 2B domains. The proteomics analysis by Metascape shows that KRT6C/KRT9 keratin pair is in a densely connected protein-protein network, where proteins participate jointly in regulating cytoskeleton organization and keratinization processes.

Conclusions: This study shows that co-immunoprecipitation coupled with mass spectroscopy provides a sensitive and reliable approach, which helps discover the previously unidentified hereterodimer partner KRT6C for KRT9 in a native endogenous environment. The acknowledgement of KRT6C/KRT9 pairwise relationship may help re-classify the genetic cutaneous disorders EPPK and PC-K6c at a molecular-based level, and lay foundation for deciphering the keratin network potentially contributing to function abnormalities in EPPK and even PC-K6c.

Background

Epidermolytic palmoplantar keratoderma (EPPK, OMIM: 144200), a highly penetrant autosomal dominant genodermatosis, is characterized by diffuse yellow keratoses on the palms and soles, with distinct erythematous border [1, 2]. The prevalence of EPPK is 1.0 to 4.4 per 100,000 without gender difference [3–5]. Patients usually manifest symptoms and signs shortly after birth [1]. Other manifestations include nail clubbing, knuckle pad-like keratoses on the finger joints [6], digital mutilation [7], hyperhidrosis [6, 8], camptodactyly [9], increased palm and sole sensitivity to mechanical trauma and decreased sensitivity to heat [6, 8, 10].

Keratins, the predominant subtype of intermediate filaments, are the major structural proteins in epithelial cells [11]. Keratins are encoded by a large multigene family consisting of 28 acidic type I and 26 basic to neutral type II evolutionarily conserved genes [12]. Each of the type I keratin is matched by a particular

type II keratin to assemble a coiled-coil heterodimer, then the dimers form tetramers by interacting along their lateral surfaces in an antiparallel orientation [13, 14]. Keratins share the tripartite structure: a conserved central α-helix rod domain flanked by variable nonhelical head and tail domains at both ends [15]. The α-helix rod domain consists of four coils (1A, 1B, 2A, and 2B) that are separated by three nonhelical linker segments (L1, L12, and L2) [16]. Most keratin gene mutations are located in the highly conserved boundary regions of 1A and 2B domains [17]. Keratins not only provide epithelial cells with remarkable structural stability and flexibility to withstand a variety of stresses [15, 17], but also take part in cell proliferation, migration, cell type-specific organelle transport, cell size determination and malignant transformation [18].

Among the keratin genes, the type I keratin 9 gene (*KRT9*, OMIM: 607606) encodes a protein that is exclusively expressed in the suprabasal cells and concentrated around intraepidermal sweat-gland ducts of palms and soles [19]. Mutations in the *KRT9* gene are responsible for EPPK. At present, limited literature speculate that the type II keratin-binding partner for KRT9 may be keratin 1 (KRT1, OMIM: 139350) or keratin 5 (KRT5, OMIM: 148040) [20, 21]. But there is a lack of detailed studies to prove the speculation, thus, the exact type II partner remains mysterious. In this study, we focus on identifying the type II keratin-binding partner for KRT9 and proving their interaction. Unveiling the unknown type II partner for KRT9 in palmoplantar skin may help investigate the pairwise regulation of keratin pair that is responsible for the pathogenic mechanism of EPPK.

Results

Identification of the potential type II keratin-binding partner for KRT9

After KRT9 was immunoprecipitated and isolated from human sole skin sample, the complex components were eluted and collected for protein identification using liquid chromatography-tandem mass spectrometry (LC-MS-MS) analysis (Fig. 1). LC-MS-MS identified 798 and 727 peptide sequences in the target sample and the IgG control, respectively, with peptide matches above homology or identity threshold. Peptide identification and data screening returned 175 proteins in the target sample and 128 proteins in the IgG control. By comparing the differential proteins in these two groups, we discovered that keratin, type II cytoskeletal 6C (KRT6C, UniProtKB/Swiss-Prot: P48668), encoded by the keratin 6c gene (*KRT6C*, OMIM: 612315), was the only type II keratin in the differential protein list of the target sample (Tables 1–2). By searching the Human Intermediate Filament Database (HIFD, http://www.interfil.org), no type I keratin was established as the hereterodimer partner for KRT6C. Furthermore, the *KRT6C* gene was reported to be responsible for a milder form of pachyonychia congenita (PC), named as PC-K6c, that was initially termed as palmoplantar keratoderma, non-epidermolytic, focal or diffuse (PPKNEFD, OMIM: 615735) [22, 23]. Therefore, KRT6C was selected as the hereterodimer partner candidate for KRT9.

Verification of KRT6C as the hereterodimer partner for KRT9

To demonstrate the putative relationship between KRT6C and KRT9, we performed the reverse co-immunoprecipitation using sole skin sample. By immunoprecipitating against the interactor KRT6C and probing the immunoblot against the original bait protein KRT9, we showed that endogenous KRT6C and KRT9 interact with each other in the native environment of sole skin (Fig. 2). To visualize the possible co-localization relationship between KRT6C and KRT9 and for further verification of their distribution in palm and sole skin, we immunostained palm and sole skin with antibodies and fluorophore mixture (Fig. 3). Confocal images of palm and sole tissue sections showed suprabasal cells and intraepidermal sweat-gland ducts in skin sample, in which yellow represented co-localization of both proteins. By analyzing five high-resolution confocal images, quantification of the fraction of KRT9 colocalizing with KRT6C was carried out using ImageJ software. The Pearson's correlation coefficient in the colocalized volume (R_{coloc}) was 0.8437 ± 0.024 (mean ± SD), showing a high degree of correlation [24]. Therefore, quantitative co-localization analysis (QCA) of KRT9 and KRT6C performed with ImageJ software suggests that KRT6C is colocalized with KRT9, and KRT6C is the hereterodimer partner of KRT9 in palm and sole skin.

In silico model for KRT6C/KRT9 interaction and multiple sequence alignment analysis

In the process of building *in silico* model for KRT6C/KRT9 interaction, the SWISS-MODEL returned six predicted models. The model with the best QMEANDisCo global and quaternary structure quality estimate (QSQE) values was built with a template that has a crystal structure of the heterocomplex between coil 2B domains of wild-type KRT1 and keratin 10 (KRT10, OMIM: 148080) containing mutation Cys401Ala. Thus, it was chosen to perform afterwards analysis. The QMEANDisCo global value for this KRT6C/KRT9 interaction model was 0.71 ± 0.06 (range 0–1), indicating a fine expected quality of the model. While the QSQE value was 0.46, implying that it is better to follow the predicted quaternary structure with skepticism. The predicted surface model of the KRT6C/KRT9 interaction was built by PyMOL based on the van der Waals radius of atoms in their 2B domains, which shows a typical coiled-coil structure (Fig. 4A).

The sequence conservation across different intermediate filament proteins was analyzed by performing multiple sequence alignment of vimentin, KRT1, KRT6C, KRT9 and KRT10. Their conserved regions in the 1A and 2B domains are shown in Figs. 4B–C. The heptad repeats ("*abcdefg*") in the intermediate filament rod domains form a repeated pattern of charged and hydrophobic amino acid residues, where residues at sites "*a*" and "*d*" are usually hydrophobic and buried [25, 26]. Residues at the "*d*" positions of KRT9 were partially predicted by PyMOL to interact with residues at the "*a*" positions of the corresponding heptad repeats in KRT6C. Residue p.Leu380 at the "*d*" position of KRT9 was thus predicted to interact with residue p.Ile395 at the "*a*" position of KRT6C (Fig. 4A).

Gene ontology enrichment and protein-protein interaction network analysis for KRT9-related differential protein data

To further analyze the biological processes and pathways that KRT9 was involved in, the relationships among 101 differential proteins of the target sample (vs IgG control) were examined using Metascape.

The analysis identified five most prominent enriched Gene ontology (GO) terms for KRT9-related differential protein data, including metabolism of RNA, ribonucleoprotein complex subunit organization, regulation of translation, regulation of proteolysis, and supramolecular fiber organization (Fig. 5A). The enriched GO terms with the best *P*-values were selected by Metascape and visualized in Cytoscape as a network plot (Fig. 5B), showing that nodes with the same cluster color are typically close to each other. Protein-protein interaction enrichment analysis was also performed in Metascape, which is shown as a general PPI network (Fig. 5C). The top-three functional enriched terms of the PPI network, shown underneath the network, are still ribonucleoprotein complex subunit organization, translation, and metabolism of RNA, as the GO terms in Fig. 5A, only with a slight shift in order. The Molecular Complex Detection (MCODE) algorithm identified five densely connected network components in the resultant PPI network, named as MCODE_1 to MCODE_5. To our excitement, in the MCODE_1 component, which contains a subset of 12 proteins, KRT9 and KRT6C are connected directly, implying that they form physical interaction with each other. The three functional descriptions of the corresponding MCODE have been reserved in the table besides the MCODE component plots within Fig. 5C, showing that the prominent functional descriptions for MCODE_1 are intermediate filament cytoskeleton organization, intermediate filament-based process, and formation of the cornified envelope. These three descriptions together with the sixth description Keratinization (R-HSA-6805567) for MCODE_1 all have the same gene hits as KRT9, KRT6C and KRT19 (Supplementary table 1), implying these three keratin proteins participate jointly in regulating intermediate filament-based cytoskeleton organization and keratinization processes.

Discussion

Keratins, together with microtubules and microfilaments, are cytoskeletal filament-forming proteins that are critical for cytoarchitecture and composition of skin and other epithelia [27]. The type I/type II keratin heterodimers are expressed in highly specific pairs (e.g., KRT1/KRT10, KRT3/KRT12, KRT4/KRT13, KRT5/KRT14, KRT6A/KRT16, KRT6B/KRT17) depending on tissue type and differentiation stage [27, 28]. A growing number of keratin gene mutations have been found to be causative factors for a wide range of skin, hair, and mucosal diseases [29].

EPPK is characterized by diffuse severe hyperkeratosis of palms and soles, whose pathogenic genes are identified as *KRT9* and *KRT1* [30]. At present, 29 *KRT9* and a few *KRT1* mutations identified in EPPK have been documented in the HIFD and Human Gene Mutation Database (HGMD, http://www.hgmd.cf.ac.uk) [29]. These mutations are disruptive to type I/type II filament assembly via dominant negative effect, causing abnormal perinuclear tonofilament clumping and large distorted keratohyalin granules [5].

PC is characterized by hypertrophic nail dystrophy, painful palmoplantar keratoderma, oral leukokeratosis, pilosebaceous cysts, and follicular keratoses [31]. The reported causative genes for PC are the keratin 6A gene (*KRT6A*, OMIM: 148041), the keratin 6B gene (*KRT6B*, OMIM: 148042), *KRT6C*, the keratin 16 gene (*KRT16*, OMIM: 148067), and the keratin 17 gene (*KRT17*, OMIM: 148069) [31]. Based on the pathogenic genetic factor, PC is classified as PC-K6a (OMIM: 615726), PC-K6b (OMIM: 615728), PC-K6c, PC-K16 (OMIM: 167200), and PC-K17 (OMIM: 167210) [31]. Data in the International PC Research

Registry (IPCRR, www.pachyonychia.org/) shows that only 3% (22/774) of PCs are attributed to pathogenic *KRT6C* mutations, while *KRT6A*, *KRT6B*, *KRT16* and *KRT17* mutations account for 39% (304/774), 9% (70/774), 32% (247/774), and 17% (130/774) of PCs, respectively. Although the clinical diagnostic criterion for PC is the triad of toenail thickening, plantar keratoderma, and plantar pain [31]. In the PC-K6c patients, only 59% (13/22) of them had mild toenail dystrophy, while 96%~99% of PC patients carrying mutations in the other four keratin genes had toenail thickened [31].

To date, most identified type I/type II keratin heterodimers were discovered by conforming empirical data with co-localization assay or using yeast-two hybrid analysis [32, 33]. In this study, we discovered that KRT6C is the type II keratin-binding partner for KRT9 in human sole skin by using co-immunoprecipitation coupled with mass spectroscopy. The *KRT6C* gene, belonging to the type II keratins clustered in chromosome 12q13, consists of nine exons, and encodes 564 amino acids (NCBI GenBank database). The three KRT6 isoforms (KRT6A, KRT6B and KRT6C), encoded by three functional isogenes (*KRT6A*, *KRT6B* and *KRT6C*), only differ by 7 separated amino acid substitutions. This fact has enhanced the difficulty in distinguishing them using effective antibodies [34]. Therefore, the keratin pair remains unidentified until now. By using co-immunoprecipitation coupled with mass spectroscopy, we identified KRT6C/KRT9 interaction in a native endogenous environment. In this way, this method functions as a sensitive and powerful way to identify novel protein interactions with physiological importance. By performing *in silico* model for KRT6C/KRT9 interaction and multiple sequence alignment of several related intermediate filament proteins, we propose that such interaction is significant for the formation of highly organized KRT6C/KRT9 heterodimers and tetramers.

Traditional classification of keratin-associated cutaneous disorders was generally based on morphologic appearance of skin lesions and presence or absence of extracutaneous symptoms and signs [35]. Growing evidence of pathogenic keratin mutations for hereditary cutaneous disorders and their overlapping phenotypes have facilitated a molecular-based classification of hereditary cutaneous disorders. For instance, mutations in either of the keratin pair KRT5/KRT14 are causative for epidermolysis bullosa simplex (EBS); KRT6A/KRT16 and KRT6B/KRT17 for PC; KRT1/KRT10 for epidermolytic ichthyosis (EI, OMIM: 113800) [27, 36].

In the case for keratin pair KRT6C/KRT9, underlying trace evidence for functional correlation between the two keratin proteins can be discovered by carefully reviewing previous studies. In the $Krt9^{+/mut}$, $Krt9^{mut/mut}$ and $Krt9^{-/-}$ mice, KRT6 increased significantly in response to KRT9 dysfunction in the footpad [17, 20]. In comparison, KRT1 showed no obvious expression level or location alteration in the $Krt9^{+/-}$, $Krt9^{+/mut}$ and $Krt9^{-/-}$ mice and increased in the $Krt9^{mut/mut}$ mice [17, 20]. KRT5 showed no obvious expression alteration in the $Krt9^{+/-}$ and $Krt9^{+/mut}$ mice and decreased in the $Krt9^{-/-}$ and $Krt9^{mut/mut}$ mice [17, 20]. In the PC-K6a patients, whose *KRT6A*, *KRT6B* and *KRT6C* transcripts were increased, *KRT9* showed dramatic decrease [37]. The relationship between KRT9 and its putative partner KRT1 or KRT5 seems ambiguous. The interesting reciprocal correlation between *KRT9* and *KRT6* expression suggest that the keratin pair KRT6C/KRT9 may be collectively involved in EPPK and PC-K6c pathogenesis.

The functional correlation between KRT9 and KRT6C may be used to explain the phenotypic overlap in EPPK and PC-K6c. EPPK is characterized by palmplantar hyperkeratosis, and PC-K6c by focal palmplantar hyperkeratosis with minor nail involvement [2, 38]. For EPPK kindreds, whose causative mutations were not identified in *KRT9* or *KRT1*, it is reasonable to carefully examine *KRT6C* to seek for potential mutations. Nonetheless, KRT9 seems to function as a more important cytoskeletal protein in keratinocytes to maintain cell structure and function than KRT6C. More *KRT9* mutations (about 29 mutations) have been identified than *KRT6C* mutations (about 4 mutations), indicating that occurrence of *KRT9* mutations in the conservative sites are more likely to cause disease phenotypes. *KRT6C* gene variants were also identified in participants with no clinical symptoms or signs of genodermatoses, implying that these variants were not severe enough to produce a clinical phenotype [22, 31].

Combined with the proteomics analysis result of the KRT9-related differential protein data produced by Metascape, which integrated a broad set of current biological datasets, we realized that the keratin pair KRT6C/KRT9 together with the keratin 19 (KRT19, OMIM: 148020) play an important role in intermediate filament-based cytoskeleton organization, keratinization, and formation of the cornified envelope. The dysfunction of KRT6C/KRT9 may trigger an imbalance in a subset of closely related keratin family and disrupt epidermal desquamation homeostasis. The produced structurally weakened keratinocytes and compromised epidermis may thus more likely prone to external stimuli, stress, and injury. Due to different KRT6C/KRT9 mutation type and different degree of conservativeness in various mutation sites, a certain range of subtly different phenotypes may be produced in the context of various genetic and environmental backgrounds [17, 36]. Therefore, it may be more reasonable and useful to consider disorders caused by the KRT6C/KRT9 keratin pair as a group of hereditary defects, as the role of KRT5/KRT14 for EBS and KRT1/KRT10 for EI mentioned above.

Conclusions

To sum up, combining co-immunoprecipitation, LC-MS-MS, and post-MS analyses, we have identified and verified KRT6C as the hereterodimer partner for KRT9. The proteomics analysis revealed the crucial role of KRT6C/KRT9 in the intermediate filament-based cytoskeleton organization and keratinization processes. The new understanding of KRT6C-KRT9 pairwise correlation may help re-classify the genetic cutaneous disorders EPPK and PC-K6c at a molecular-based level, and lay foundation for pathogenic mechanism research in EPPK and PC-K6c.

Methods

Sample preparation and co-immunoprecipitation

Human sole skin sample was collected from three foot trauma surgeries. Most of the sample was stored in liquid nitrogen before being processed for immunoprecipitation as previously described [39], the rest was embedded in paraffin for immunofluorescence staining. Informed consents were obtained from the involved subjects. This study was carried out in accordance with the Declaration of Helsinki and had

received approval from the Biomedical Research Ethics Committee of Hunan Normal University, Changsha, Hunan, P.R. China. Only epidermis part was used for this study. Skin epidermis was minced and lysed in ice-cold solubilization buffer (P0013, Beyotime). After debris was removed, part of cleared supernatant was transferred to immunoprecipitation using prepared beads bound with mouse monoclonal anti-keratin 9 antibody (ab19124, Abcam) or normal mouse IgG (074-1807, KPL). The rest cleared supernatant was saved for Western blotting. Proteins that co-precipitated with antibody-bead complex were washed with Tris-buffered saline (TBS). Then the beads were resuspended in sample buffer and heated for 5 min at 105℃ to elute proteins. The protein samples were centrifuged and collected for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and silver staining. Protein bands of interest were excised, reduced with dithiothreitol, alkylated with iodoacetamide, and destained. The gel pieces were digested, and the supernatant was withdrawed for further LC-MS-MS analysis [40, 41].

Liquid chromatography-tandem mass spectrometry (LC-MS-MS) and data analysis

Peptide sequences and intensity were analyzed using Thermo Scientific™ Q Exactive mass spectrometer as previously described [41]. Protein identification was performed using the Mascot search engine (http://www.matrixscience.com/) by searching LC-MS-MS fragmentation patterns from each peptide against the Swiss-Prot database in Uniprot (https://www.uniprot.org/). Qualified protein sequences that meet the peptide sequencing criterion were retained for afterwards analysis [41].

Type II keratin candidates that may be binding partner for KRT9 were searched against PubMed (https://pubmed.ncbi.nlm.nih.gov/). The predicted protein-protein interactions were analyzed using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, version 11.0) and HIFD. The protein-protein interaction between the retained candidate and KRT9 was further verified using reverse co-immunoprecipitation, Western blotting, immunofluorescence assay and quantitative co-localization analysis.

Western blotting

The protein samples that saved for verifying protein-protein interaction were immunoprecipitated using prepared beads bound with rabbit polyclonal anti-keratin 6C antibody (abx128996, Abbexa) or normal rabbit IgG (074-1506, KPL) in a similar way as described above. Then the collected protein samples were transferred to immunoblot analysis. Protein concentrations were measured using the bicinchoninic acid method (Pierce Biotechnology). Equal amounts of protein samples were subjected to SDS-PAGE and blotted onto polyvinylidene difluoride membranes (Millipore). The membranes were sequentially blocked and probed overnight at 4°C with primary antibodies, including mouse monoclonal anti-keratin 9 antibody (ab19124, Abcam) and rabbit polyclonal anti-keratin 6C antibody (abx128996, Abbexa). Protein bands were visualized using an enhanced chemiluminescence detection system (Advansta Corporation) after incubation with appropriate secondary antibodies as previously described [39].

Immunofluorescence staining

Paraffin sections of sole and palm skin sample were deparaffinized in xylene, hydrated with graded ethanol, rinsed in distilled water, and subjected to microwave-based antigen retrieval. Sections were incubated with 3% goat serum albumin for 30 min. Skin sections were incubated overnight at 4℃ with a mixture of primary antibodies, including mouse monoclonal anti-keratin 9 antibody (ab19124, Abcam, 1:500) and rabbit polyclonal anti-keratin 6C antibody (abx128996, Abbexa, 1:500). The sections were incubated with a mixture of secondary antibodies for simultaneous immunofluorescence double-labeling. Secondary antibodies were FITC-conjugated goat anti-mouse IgG (SA00003-1, Proteintech, 1:200) and Cy3-conjugated goat anti-rabbit IgG (SA00009-2, Proteintech, 1:100). Nuclei were stained with DAPI (Beyotime). A Leica TCS SP8 confocal laser scanning microscope was used to observe the cellular localization of KRT9 and KRT6C.

Statistical analysis

Colocalization analysis of different proteins in immunofluorescence images was performed using ImageJ (National Institutes of Health, Bethesda, MD), and colocalization coefficients were determined by Pearson's correlation analysis. Statistical analysis was performed using PASW Statistics 18 (SPSS Inc., Chicago, IL, USA).

In silico modeling of KRT6C/KRT9 interaction and multiple sequence alignment analysis

Schematic illustration of KRT6C/KRT9 interaction model was drawn with SWISS-MODEL and PyMOL (Schrödinger, LLC) [42]. The built model was evaluated using model quality measurement index, such as QMEANDisCo global. QMEANDisCo global score gives an overall model quality measurement from 0 to 1, with higher number indicating higher expected quality [43]. A QSQE score has a number between 0 and 1, which reflects the expected accuracy of the interchain contacts for a built model. A higher QSQE is generally "better", a value above 0.7 can be considered reliable to follow the predicted quaternary structure in the modelling process [44]. Multiple sequence alignment of vimentin (NP_003371.2), KRT1 (NP_006112.3), KRT6C (NP_775109.2), KRT9 (NP_000217.2) and KRT10 (NP_000412.4) was performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).

Proteomics analysis of the KRT9-related differential protein data in the target sample

The collection of KRT9-related differential protein data was analyzed with Metascape, a gene annotation and analysis resource (https://metascape.org/) [45]. Biological pathway and process enrichment analysis was performed in Metascape using an abundant of ontology sources, such as KEGG Pathway, GO Biological Processes, and Reactome Gene Sets, *etc*. GO terms with a *P*-value < 0.01 and an enrichment factor > 1.5 were collected and clustered based on their membership similarities. The top 20 clusters with their corresponding representative terms were documented. The enriched GO terms with the best *P*-values were selected by Metascape and rendered as a network plot, which was viewed in Cytoscape (https://cytoscape.org/) [46]. Protein-protein interaction (PPI) enrichment analysis was carried out in

Metascape using databases such as STRING and BioGrid [47, 48]. The MCODE algorithm was applied to identify densely connected network components in the PPI network [49]. Pathway and process enrichment analysis has been independently applied to each MCODE component, and the three best-scoring terms calculated by *P*-value have been reserved as the functional description of the corresponding components.

Declarations

Ethics approval and consent to participate: The study was conducted in accordance with the Declaration of Helsinki and had received approval from the Biomedical Research Ethics Committee of Hunan Normal University, Changsha, Hunan, P.R. China (protocol code 2022-008, 2022/1/21). Informed consent was obtained from all subjects involved in the study.

Consent for publication: Not applicable.

Availability of data and materials: The datasets analyzed during the current study are openly available in jPOST (https://jpostdb.org/, Announced ID: JPST001464) in February 2023.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by the National Natural Science Foundation of China (81903199, 81802871), the Natural Science Foundation of Hunan Province (2021JJ40373), and the Hunan Province College Students' Research Learning and Innovative Experiment Project (S202110542056), China. The funding bodies have no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions: HX and PL designed the study, supervised the whole process of experiments, and wrote the manuscript. JQ performed most of the experiments. YZ and AD performed part of the experiments. TD and PX analyzed the data. All authors read and approved the final manuscript.

Acknowledgements: We are grateful to the patients for their kind contribution in this work. Without them, this work cannot be accomplished.

References

1. Reis A, Hennies HC, Langbein L, Digweed M, Mischke D, Drechsler M, et al. Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). Nat Genet. 1994;6(2):174–9.
2. Grimberg G, Hausser I, Muller FB, Wodecki K, Schaffrath C, Krieg T, et al. Novel and recurrent mutations in the 1B domain of keratin 1 in palmoplantar keratoderma with tonotubules. Br J Dermatol. 2009;160(2):446–9.
3. Braun-Falco M. Hereditary palmoplantar keratodermas. J Dtsch Dermatol Ges. 2009;7(11):971–84. quiz 84 – 5.

4. Covello SP, Irvine AD, McKenna KE, Munro CS, Nevin NC, Smith FJ, et al. Mutations in keratin K9 in kindreds with epidermolytic palmoplantar keratoderma and epidemiology in Northern Ireland. J Invest Dermatol. 1998;111(6):1207–9.
5. Szalai S, Szalai C, Becker K, Torok E. Keratin 9 mutations in the coil 1A region in epidermolytic palmoplantar keratoderma. Pediatr Dermatol. 1999;16(6):430–5.
6. Kuster W, Reis A, Hennies HC. Epidermolytic palmoplantar keratoderma of Vorner: re-evaluation of Vorner's original family and identification of a novel keratin 9 mutation. Arch Dermatol Res. 2002;294(6):268–72.
7. Umegaki N, Nakano H, Tamai K, Mitsuhashi Y, Akasaka E, Sawamura D, et al. Vorner type palmoplantar keratoderma: novel KRT9 mutation associated with knuckle pad-like lesions and recurrent mutation causing digital mutilation. Br J Dermatol. 2011;165(1):199–201.
8. Li M, Yang LJ, Hua HK, Zhu XH, Dai XY. Keratin-9 gene mutation in epidermolytic palmoplantar keratoderma combined with knuckle pads in a large Chinese family. Clin Exp Dermatol. 2009;34(1):26–8.
9. Du ZF, Wei W, Wang YF, Chen XL, Chen CY, Liu WT, et al. A novel mutation within the 2B rod domain of keratin 9 in a Chinese pedigree with epidermolytic palmoplantar keratoderma combined with knuckle pads and camptodactyly. Eur J Dermatol. 2011;21(5):675–9.
10. Lopez-Valdez J, Rivera-Vega MR, Gonzalez-Huerta LM, Cazarin J, Cuevas-Covarrubias S. Analysis of the KRT9 gene in a Mexican family with epidermolytic palmoplantar keratoderma. Pediatr Dermatol. 2013;30(3):354–8.
11. Coulombe PA, Omary MB. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. Curr Opin Cell Biol. 2002;14(1):110–22.
12. Jacob JT, Coulombe PA, Kwan R, Omary MB. Types I and II Keratin Intermediate Filaments. Cold Spring Harbor perspectives in biology. 2018;10(4).
13. Herrmann H, Aebi U. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. Annu Rev Biochem. 2004;73:749–89.
14. Kim S, Coulombe PA. Intermediate filament scaffolds fulfill mechanical, organizational, and signaling functions in the cytoplasm. Genes Dev. 2007;21(13):1581–97.
15. Wang F, Zieman A, Coulombe PA. Skin Keratins. Methods Enzymol. 2016;568:303–50.
16. Irvine AD, McLean WH. Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype-genotype correlation. Br J Dermatol. 1999;140(5):815–28.
17. Lyu YS, Shi PL, Chen XL, Tang YX, Wang YF, Liu RR, et al. A Small Indel Mutant Mouse Model of Epidermolytic Palmoplantar Keratoderma and Its Application to Mutant-specific shRNA Therapy. Mol Ther Nucleic Acids. 2016;5:e299.
18. Magin TM, Vijayaraj P, Leube RE. Structural and regulatory functions of keratins. Exp Cell Res. 2007;313(10):2021–32.

19. Moll I, Heid H, Franke WW, Moll R. Distribution of a special subset of keratinocytes characterized by the expression of cytokeratin 9 in adult and fetal human epidermis of various body sites. Differ Res Biol Divers. 1987;33(3):254–65.
20. Fu DJ, Thomson C, Lunny DP, Dopping-Hepenstal PJ, McGrath JA, Smith FJD, et al. Keratin 9 is required for the structural integrity and terminal differentiation of the palmoplantar epidermis. J Invest Dermatol. 2014;134(3):754–63.
21. Kim D, Hossain MZ, Nieves A, Gu L, Ratliff TS, Mi Oh S, et al. To Control Site-Specific Skin Gene Expression, Autocrine Mimics Paracrine Canonical Wnt Signaling and Is Activated Ectopically in Skin Disease. Am J Pathol. 2016;186(5):1140–50.
22. Wilson NJ, Messenger AG, Leachman SA, O'Toole EA, Lane EB, McLean WH, et al. Keratin K6c mutations cause focal palmoplantar keratoderma. J Invest Dermatol. 2010;130(2):425–9.
23. Duverger O, Carlson JC, Karacz CM, Schwartz ME, Cross MA, Marazita ML, et al. Genetic variants in pachyonychia congenita-associated keratins increase susceptibility to tooth decay. PLoS Genet. 2018;14(1):e1007168.
24. Zinchuk V, Grossenbacher-Zinchuk O. Quantitative colocalization analysis of confocal fluorescence microscopy images. Curr Protoc Cell Biol. 2011;Chap. 4:Unit4 19.
25. Arin MJ. The molecular basis of human keratin disorders. Hum Genet. 2009;125(4):355–73.
26. Strelkov SV, Herrmann H, Geisler N, Wedig T, Zimbelmann R, Aebi U, et al. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. EMBO J. 2002;21(6):1255–66.
27. Knobel M, O'Toole EA, Smith FJ. Keratins and skin disease. Cell Tissue Res. 2015;360(3):583–9.
28. Li C, Chen P, Sun S, Zeng K, Liang J, Wang Q, et al. Exome sequencing identifies a KRT9 pathogenic variant in a Chinese pedigree with epidermolytic palmoplantar keratoderma. Mol Genet Genomic Med. 2019;7(7):e00703.
29. Szeverenyi I, Cassidy AJ, Chung CW, Lee BT, Common JE, Ogg SC, et al. The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. Hum Mutat. 2008;29(3):351–60.
30. Xiao H, Guo Y, Yi J, Xia H, Xu H, Yuan L, et al. Identification of a Novel Keratin 9 Missense Mutation in a Chinese Family with Epidermolytic Palmoplantar Keratoderma. Cell Physiol Biochem. 2018;46(5):1919–29.
31. Smith FJD, Hansen CD, Hull PR, Kaspar RL, McLean WHI, O'Toole E, et al Pachyonychia Congenita. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al., editors. GeneReviews((R)). Seattle (WA)1993.
32. Free RB, Hazelwood LA, Sibley DR. Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy. Curr Protoc Neurosci. 2009;Chap. 5:Unit 5 28.
33. Knapp AC, Franke WW, Heid H, Hatzfeld M, Jorcano JL, Moll R. Cytokeratin. No. 9, an epidermal type I keratin characteristic of a special program of keratinocyte differentiation displaying body site specificity. J Cell Biol. 1986;103(2):657–67.

34. Takahashi K, Paladini RD, Coulombe PA. Cloning and characterization of multiple human genes and cDNAs encoding highly related type II keratin 6 isoforms. J Biol Chem. 1995;270(31):18581–92.
35. Terron-Kwiatkowski A, Paller AS, Compton J, Atherton DJ, McLean WH, Irvine AD. Two cases of primarily palmoplantar keratoderma associated with novel mutations in keratin 1. J Invest Dermatol. 2002;119(4):966–71.
36. Lane EB, McLean WH. Keratins and skin disorders. J Pathol. 2004;204(4):355–66.
37. Cao YA, Hickerson RP, Seegmiller BL, Grapov D, Gross MM, Bessette MR, et al. Gene expression profiling in pachyonychia congenita skin. J Dermatol Sci. 2015;77(3):156–65.
38. Bowden PE. Mutations in a keratin 6 isomer (K6c) cause a type of focal palmoplantar keratoderma. J Invest Dermatol. 2010;130(2):336–8.
39. Li P, Feng J, Liu Y, Liu Q, Fan L, Liu Q, et al. Novel Therapy for Glioblastoma Multiforme by Restoring LRRC4 in Tumor Cells: LRRC4 Inhibits Tumor-Infitrating Regulatory T Cells by Cytokine and Programmed Cell Death 1-Containing Exosomes. Front Immunol. 2017;8:1748.
40. Ayachi O, Barlin M, Broxtermann PN, Kashkar H, Mauch C, Zigrino P. The X-linked inhibitor of apoptosis protein (XIAP) is involved in melanoma invasion by regulating cell migration and survival. Cell Oncol (Dordr). 2019;42(3):319–29.
41. Yi T. Identifying. RISC Components Using Ago2 Immunoprecipitation and Mass Spectrometry. Methods Mol Biol. 2018;1720:149–59.
42. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296–303.
43. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEANDisCo-distance constraints applied on model quality estimation. Bioinformatics. 2020;36(6):1765–71.
44. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep. 2017;7(1):10480.
45. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.
46. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
47. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13.
48. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006;34(Database issue):D535-9.

49. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2.

Tables

Table 1 is available in the Supplementary Files section.

Table 2
Identification of KRT6C based on Mascot search program.

Observed	Mr (expt)	Mr (calc)	ppm	Miss	Expect	Peptide sequence
473.2595	944.5045	944.5039	0.60	1	0.0092	R.GRLDSELR.N
473.2601	944.5056	944.5039	1.83	1	0.012	R.GRLDSELR.N
508.7725	1015.5304	1015.5298	0.60	0	0.041	R.QLDSIVGER.G
508.7729	1015.5313	1015.5298	1.45	0	0.99	R.QLDSIVGER.G
508.7733	1015.5320	1015.5298	2.16	0	0.00021	R.QLDSIVGER.G
508.7734	1015.5323	1015.5298	2.47	0	0.0013	R.QLDSIVGER.G
508.7737	1015.5329	1015.5298	3.06	0	0.0012	R.QLDSIVGER.G
513.7667	1025.5189	1025.5142	4.60	0	0.0023	R.SGFSSISVSR.S
554.2734	1106.5322	1106.5356	-3.06	0	0.00066	K.AQYEEIAQR.S
554.2748	1106.5350	1106.5356	-0.53	0	0.0012	K.AQYEEIAQR.S
554.2750	1106.5355	1106.5356	-0.10	0	0.004	K.AQYEEIAQR.S
583.2981	1164.5816	1164.5775	3.58	0	7.2	K.YEELQVTAGR.H
586.8228	1171.6309	1171.6309	0.02	1	0.0014	R.RQLDSIVGER.G
586.8234	1171.6322	1171.6309	1.07	1	0.0001	R.RQLDSIVGER.G
632.3515	1262.6884	1262.6870	1.13	0	1.1e-005	K.LALDVEIATYR.K
632.3519	1262.6892	1262.6870	1.71	0	9.6e-006	K.LALDVEIATYR.K
665.3678	1328.7210	1328.7187	1.75	0	0.0011	R.NLDLDSIIAEVK.A
704.3578	1406.7010	1406.7041	-2.22	0	0.00044	K.ADTLTDEINFLR.A

Figures

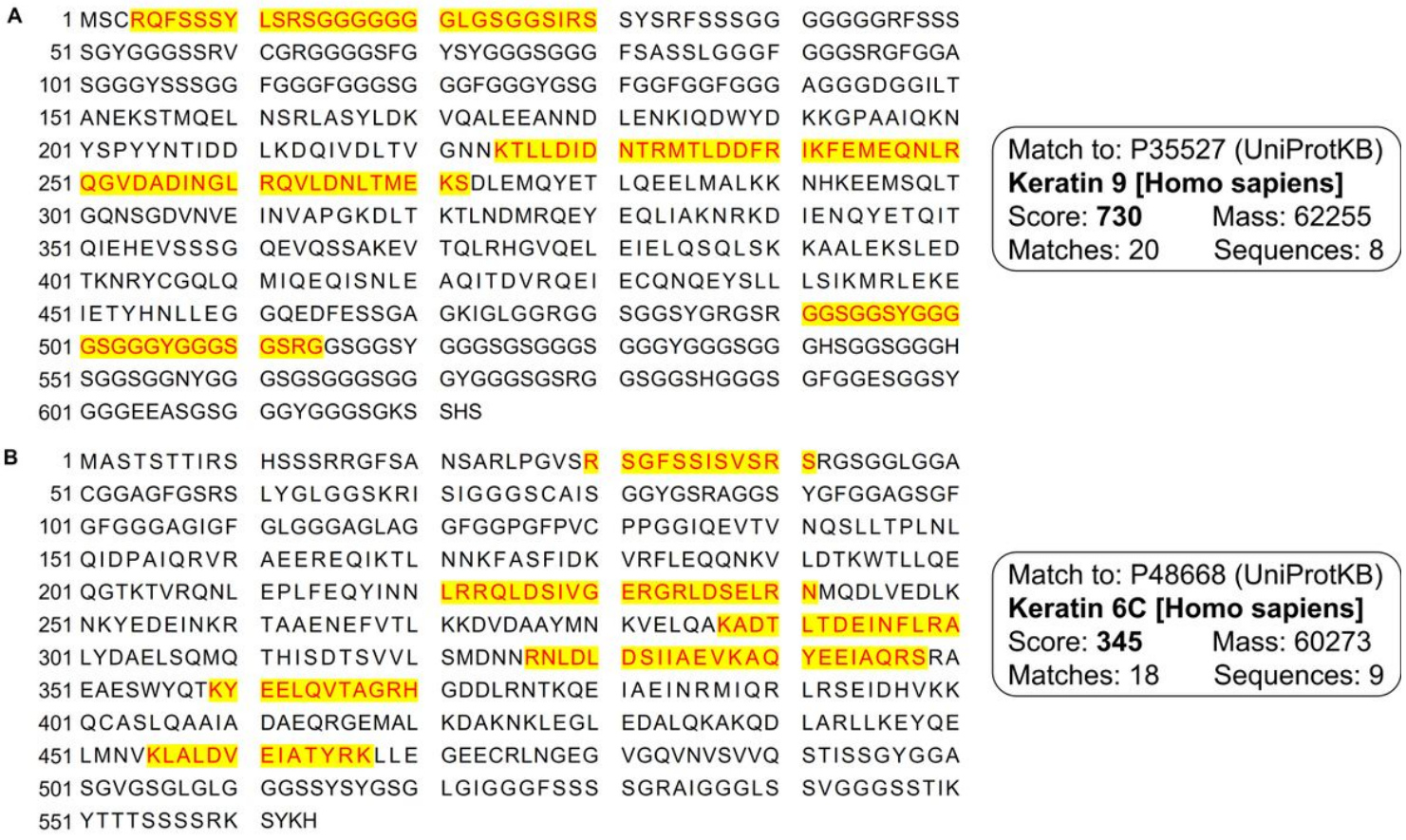


Figure 1

Representative peptide identification in the target sample based on liquid chromatography-tandem mass spectrometry (LC-MS-MS) analysis and Mascot search program. **A** The KRT9 peptide sequences identified by LC-MS-MS are highlighted in red with a yellow background, which are matched to the reference protein sequence obtained from the UniProtKB (P35527). **B** The KRT6C peptide sequences identified by LC-MS-MS are highlighted in red with a yellow background, which are matched to the reference protein sequence obtained from the UniProtKB (P48668)

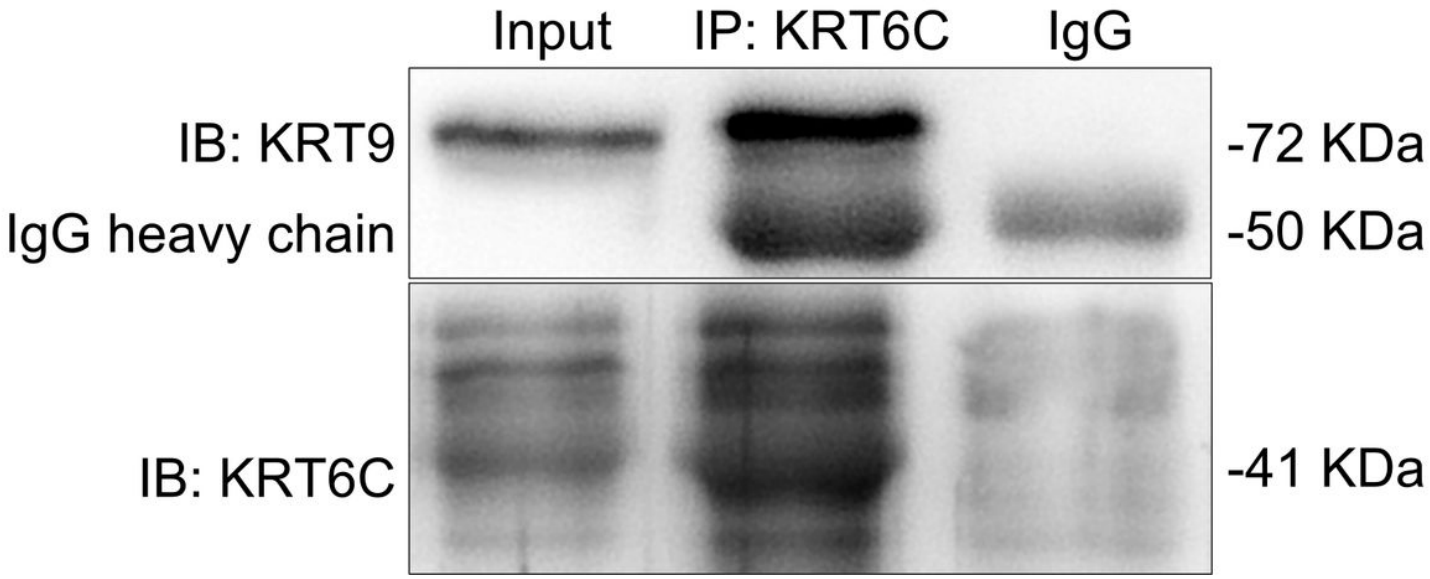


Figure 2

Interaction between endogenous KRT9 and KRT6C in human sole skin. Homogenates of human sole skin sample was incubated with antibody against KRT6C for immunoprecipitation. Skin sample lysate (20 μg) without co-immunoprecipitation (IP) served as a positive control (input), whereas IgG served as a negative control. The heavy chain of IgG served as a loading control. Resulting immunocomplexes were subjected to immunobloting (IB) with anti-keratin 9, anti-keratin 6C and anti-IgG antibodies, respectively

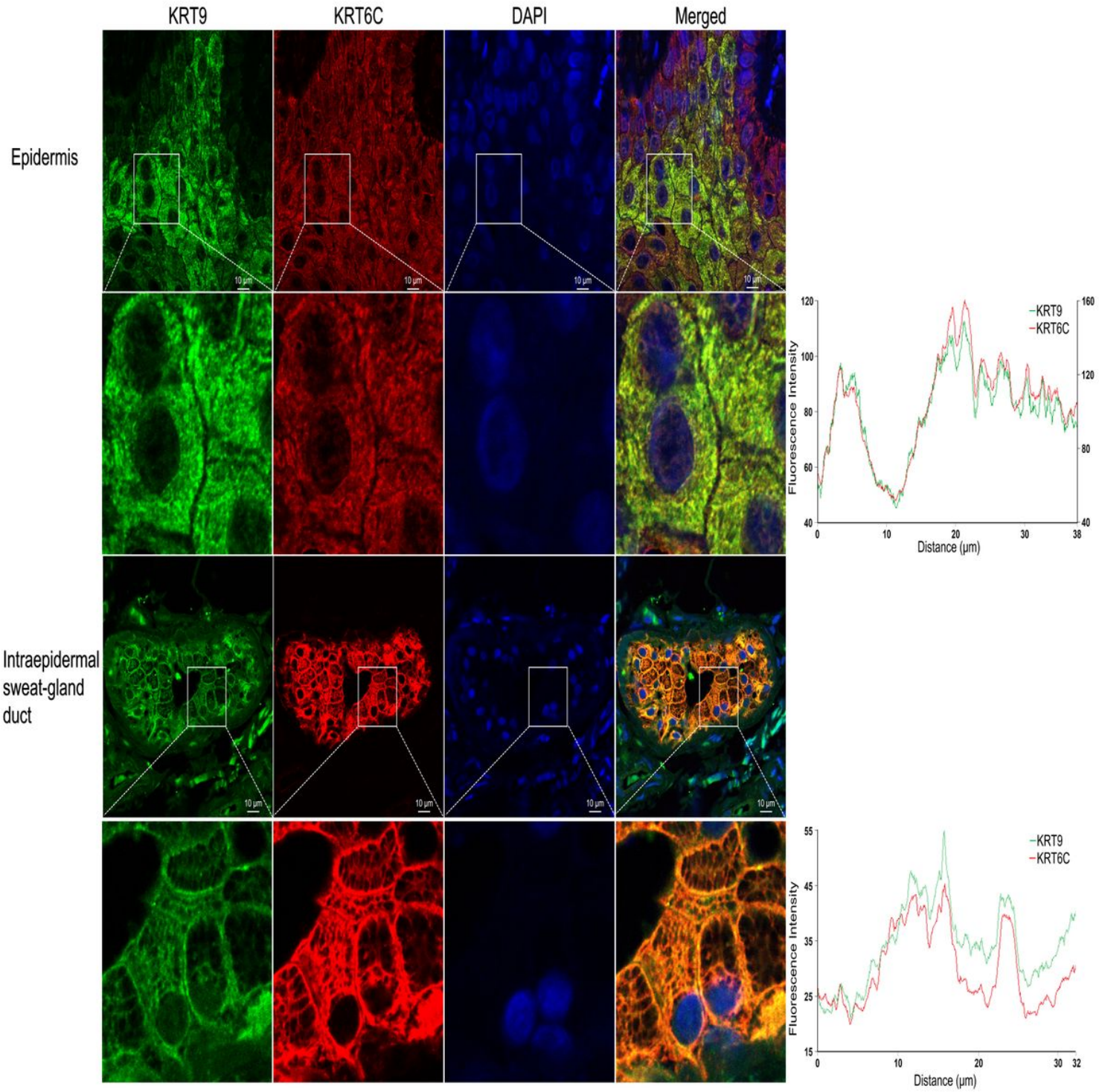
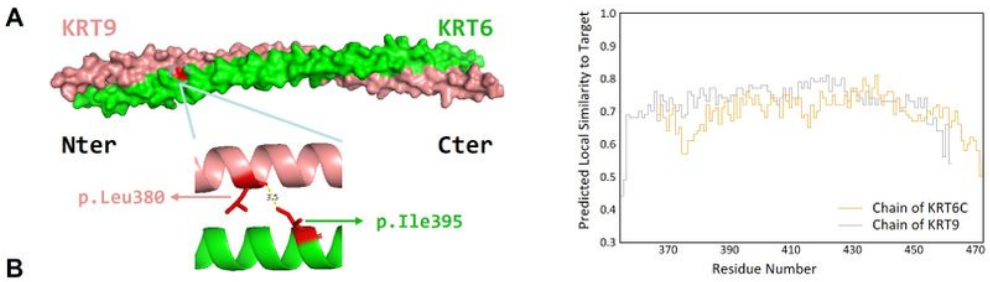


Figure 3

Distribution pattern and quantitative co-localization analysis of KRT9 and KRT6C in human palm and sole skin. KRT9 and KRT6C were simultaneously co-labeled with FITC (green) and Cy3 (red), respectively. Blue indicates DAPI staining in the nucleus. The top and second panels show suprabasal cells in palm skin sample; the third and bottom panels show intraepidermal sweat-gland ducts in sole skin sample. The boxed areas in the top and third panels are presented at higher magnification in the second and bottom panels, respectively. Fluorescence intensity profiles of the second and bottom panels for quantitative co-

localization analysis of KRT9 and KRT6C are shown on the right line charts. In the line chart for the second panel, left and right y axes of the fluorescence intensity profile correspond to KRT9 and KRT6C fluorescence intensity, respectively




```
B
vimentin   SLADAINTEFKNTRTNEKVELQELNDRFANYIDKVRFLEQQNKILLAE---  134
KRT1       PLNVEIDPEIQKVKSREREQIKSLNNQFASFIDKVRFLEQQNQVLQTK---  211
KRT6C      PLNLQIDPAIQRVRAEEREQIKTLNNKFASFIDKVRFLEQQNKVLDTK---  194
KRT9       ----AGGGDGGILTANEKSTMQELNSRLASYLDKVQALEEANNDLENKIQD  187
KRT10      ----GFGGDGGLLSGNEKVTMQNLNDRLASYLDKVRALEESNYELEGKIKE  180
                 .        .*:  :: **.::*.::***: **: *  *  :
           bcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefg
C
vimentin   LTCEVDALKGTNESLERQMREMEENFAVEAANYQDTIGRLQDEIQNMKEEMARHLREYQDLLNVKMALDIEIATYRKLLEG  406
KRT1       LRSEIDNVKKQISNLQQSISDAEQRGENALKDAKNKLNDLEDALQQAKEDLARLLRDYQELMNTKLALDLEIATYRTLLEG  488
KRT6C      LRSEIDHVKKQCASLQAAIADAEQRGEMALKDAKNKLEGLEDALQKAKQDLARLLKEYQELMNVKLALDVEIATYRKLLEG  471
KRT9       LEIELQSQLSKKAALEKSLEDTKNRYCGQLQMIQEQISNLEAQITDVRQEIECQNQEYSLLLSIKMRLEKEIETYHNLLEG  460
KRT10      LEIELQSQLALKQSLEASLAETEGRYCVQLSQIQAQISALEEQLQQIRAETECQNTEYQQLLDIKIRLENEIQTYRSLLEG  455
           *  *::        *:  : : : .        :  :  *:  : . : :      :*. *:. *: *: ** **:.****
           defgabcdefgabcdefgabcdefgxefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefab
```

Figure 4

KRT6C/KRT9 interaction and multiple sequence alignment analysis. **A** Surface model based on the van der Waals radius of atoms in the coiled-coil structure of KRT9 (pink) and KRT6C (green) 2B domains. The red spot circled with dashed line in the model is enlarged to show the predicted interaction between the residue p.Leu380 at the "*d*" position of KRT9 and the residue p.Ile395 at the "*a*" position of KRT6C, the distance between the two residues is about 3.5 Å. The line chart on the right shows the local quality estimate of KRT9 and KRT6C chains by SWISS-MODEL. The y axis is the predicted similarity of KRT9 and KRT6C chains to target templates. The x axis is the residue number in KRT9 and KRT6C 2B domains used for modeling. **B** Multiple sequence alignment of partial 1A segments in vimentin, KRT1, KRT6C, KRT9 and KRT10. **C** Multiple sequence alignment of partial 2B segments in vimentin, KRT1, KRT6C, KRT9 and KRT10. The heptad repeats in 1A and 2B segments are marked as "*abcdefg*", with core positions highlighted with yellow. The symbols below the 1A and 2B domain alignments show the conservation degree of residues: (*) fully conserved; (:) highly conserved; (.) conserved substitution

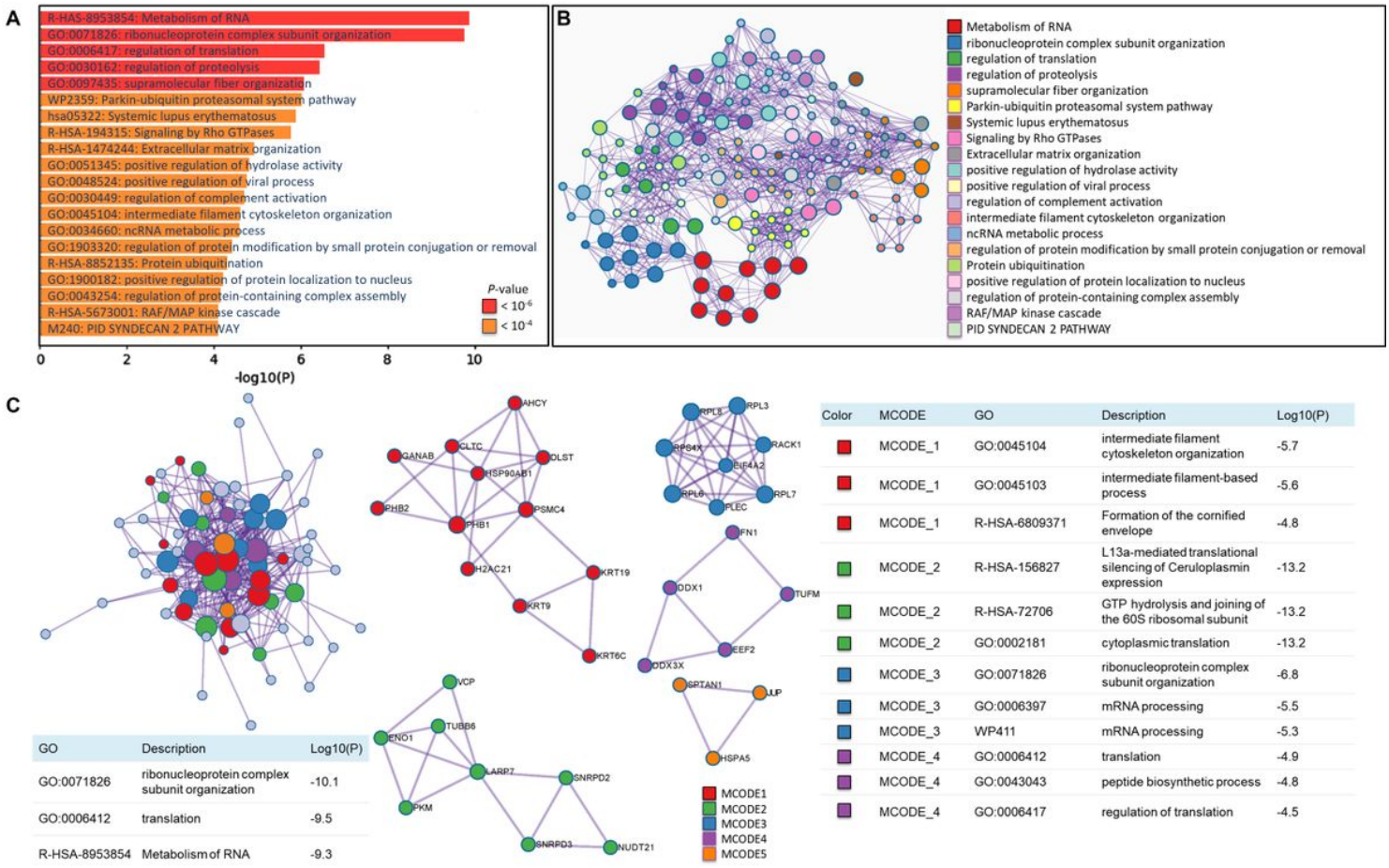


Figure 5

Characterization of KRT9-related differential protein data in the target sample. All 101 KRT9-related differential proteins together with KRT9 were analyzed using Metascape [45]. **A** Gene ontology (GO) analysis was performed for KRT9-related differential protein data in the target sample. The horizontal bar chart shows the top 20 enriched biological processes and pathways, using a discrete color scale to represent statistical significance. **B** The KRT9-related interactome network was formed by connecting subsets of enriched GO terms to identify densely related network components, where the Molecular Complex Detection (MCODE) complexes were colored according to their identities. The network was visualized using Cytoscape [46]. **C** The general KRT9-related protein-protein interaction (PPI) network and five MCODE components were identified in the differential protein list. The three best-scoring terms ranked by *P*-value have been reserved as the functional description of the PPI network and the first four MCODE components (the description for MCODE_5 was not displayed in Metascape), shown in the tables underneath or besides corresponding network plots

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx

- Additionalfile1.xls
- supplementaryIPWBRAWDATA.pdf